Annals of Operations Research (2022) 311:967–995
https://doi.org/10.1007/s10479-020-03613-9

S.I.: MOPGP19

Solution approaches for equitable multiobjective integer programming problems

Bashir Bashir[1] · Özlem Karsu[1]

Published online: 29 April 2020

Abstract
We consider multi-objective optimization problems where the decision maker (DM) has equity concerns. We assume that the preference model of the DM satisfies properties related to inequity-aversion, hence we focus on finding nondominated solutions in line with the properties of inequity-averse preferences, namely the equitably nondominated solutions. We discuss two algorithms for finding good subsets of equitably nondominated solutions. The first approach is an extension of an interactive approach developed for finding the most preferred nondominated solution when the utility function is assumed to be quasiconcave. We find the most preferred equitably nondominated solution when the utility function is assumed to be symmetric quasiconcave. In the second approach we generate an evenly distributed subset of the set of equitably nondominated solutions to be considered further by the DM. We show the computational feasibility of the two algorithms on equitable multi-objective knapsack problem, in which projects in different categories are to be funded subject to a limited budget. We perform experiments to show and discuss the performances of the algorithms.

Keywords Equitable preferences · Equitable efficiency · Equitable dominance · Generalized Lorenz dominance · Multi-objective knapsack problem · Convex cones · Fairness · Multiobjective integer programming · Interactive algorithm

1 Introduction

Multi objective optimization problems (MOP) have been studied for many years. Different techniques have been used to successfully solve and analyse these problems in a wide range of application areas such as engineering design, medical treatments, logistics, resource allocation and facility location (Deb 2014; Ulungu and Teghem 1994; White 1990). In a typical MOP, multiple objective functions that correspond to decision criteria are simultaneously optimized over a feasible region. There are trade-offs between the multiple objectives

✉ Özlem Karsu
ozlemkarsu@bilkent.edu.tr

Bashir Bashir
bashir.ab@bilkent.edu.tr

1 Department of Industrial Engineering, Bilkent University, Ankara, Turkey

considered, hence, usually no single solution optimizes all of the objective functions simultaneously. Due to these trade-offs, the concept of optimality is replaced with the concept of Pareto optimality (nondominance).

General approaches to find non-dominated points for various multiobjective programming problems (multi-objective linear programming (MOLP), multi-objective integer programming (MOIP), multi-objective mixed integer programming (MOMIP) and multi-objective combinatorial optimization (MOCO) problems) have been proposed in the literature (see e.g. Antunes et al. 2016; Kirlik and Sayın 2014; Lokman and Köksalan 2013; Sylva and Crema 2004; Mavrotas and Diakoulaki 1998; Köksalan 2008).The solution methods can be classified as exact and approximate according to the type of solutions generated. Some of these works address the problem of finding all the non-dominated points while others suggest interactive approaches to find preferred points. Further works have been done in Mavrotas and Florios (2013), Zhang and Reimann (2014), Ozlen et al. (2014) and Özlen and Azizoğlu (2009) to reduce the computational times and number of models solved in previous algorithms. Surveys on the interactive and non-interactive solution approaches to some of these problems can be found in Alves and Clímaco (2007), Clímaco et al. (1997) and Ehrgott and Gandibleux (2000).

We consider problem settings where equity concerns over multiple categories/entities are involved. Hence the problems we consider are different in the sense that all objectives are of the same type (a single type of benefit), and it is the concern of maximizing the benefit received by each category (entity) that makes the problem a multi-objective one. We call this problem multi-objective optimization problem with equity concerns (E-MOP). Unlike a classical MOP, in E-MOP the values of the objective functions are comparable. Furthermore, the criteria are considered impartially, which makes the distribution of the criteria values more important than the assigned outcome to a specific criterion.

Equity concerns arise in various real life problems and it is vital to address them for the proposed solutions to be acceptable. Therefore, researchers have started to consider extensions of several classical problems like knapsack, assignment and location problems to incorporate equity concerns. The notion of equity is usually studied in allocation settings where one tries to attain a “fair” allocation of the resources or outcomes by treating the involved entities in an impartial manner.

In general, any system serving multiple users where the service quality for every individual user is taken into consideration can be assessed with equity concerns. The users or entities involved can be departments of an organization, people of different social classes, customers at different locations, etc. For example, public service location models strive to provide equitable access to different demand points (customers). The need for inequity averseness naturally occurs in various operational research (OR) applications, including but not limited to vehicle routing problems during disaster relief Beamon and Balcik (2008), workload allocation, queuing systems, bandwidth allocation and healthcare service provision (see Karsu and Morton 2015 and the references therein).

Incorporating equity concerns into the preference model makes some solutions which are non-dominated (in classical dominance sense) unattractive. Therefore, rather than focusing on the Pareto efficient (non-dominated) set of solutions, we focus on the more relevant set of *equitably efficient* solutions. Equitable efficiency was defined in Kostreva and Ogryczak (1999) and an approach to find non-dominated points for MOP with equity concerns by aggregating the objective functions has been studied in Kostreva et al. (2004). A two-step method to find equitably efficient solutions for MOP was developed in Baatar and Wiecek (2006).

Motivated by the observation that it may be computationally too expensive to find the whole set of equitably efficient solutions, we propose two algorithms that find subsets of it for equitable multiobjective integer programming problems. We exemplify their use for project portfolio selection problems where decision makers have fairness concerns. One example is the investment decision problem, in which projects that will provide different benefits to different beneficiary groups (different population groups or different geographical zones) are considered. Each project is associated with an output vector, showing the amount of benefit it provides to these different groups, which we call *entities* throughout the text. In such cases, a typical concern for decision makers is ensuring an equitable benefit allocation over the multiple entities and a total benefit maximizing approach is usually considered inapplicable as it may result in extreme inequity in the benefit distribution. Another example occurs when project proposals that belong to different categories are evaluated and it is important to ensure a balanced funding over the multiple categories involved. A total value maximizing approach may result in imbalanced funding decisions in the sense that the majority of the funded proposals might belong to a single category (Karsu and Morton 2014). Hence, we structure these problems as multiobjective programming problems, the objective functions of which correspond to the total benefit each entity receives.

This paper is structured as follows. In Sect. 2 we discuss the concepts of equitable dominance and equitable efficiency, alongside the underlying assumptions on the decision maker's preference model. We also provide mathematical models that can be used to find the set of equitably nondominated solutions. In Sect. 3 we propose an interactive approach that finds the most preferred equitably non-dominated point for a DM. This approach is based on a novel extension of the convex cones method to a symmetric environment. In Sect. 4 we discuss an approach to generate evenly distributed equitable non-dominated points. Finally, in Sect. 5 we provide the summary of our computational experiments, in which we demonstrate the performance of the algorithms using an equitable knapsack problem. In Sect. 6, we conclude our discussion and list some future research directions.

2 Equitable dominance (efficiency)

Consider a generic multiobjective integer programming model with p objectives:

$$\begin{aligned} &Model\ 1\\ &Max\text{``}z_1(x), z_2(x), \cdots, z_p(x)''\\ &s.t.\quad x \in \mathcal{X} \end{aligned} \tag{1}$$

x denotes the vector of decision variables, which are all integer, and $\mathcal{X}$ is the feasible decision space. In the problems we consider, each objective function, $z_j(x)$, denotes the total output received by entity j in a feasible solution x. $\mathcal{Z} \subset \mathbb{Z}^p := z(\mathcal{X})$ is the feasible set in the objective (criterion) space.

Note that the "max'' operator used in these settings is not a well-defined operator. Hence, solving these models refers to finding the most preferred solution or a set of "good" solutions that are candidates to be the most preferred solution. The solution concepts applied in multiple criteria decision making literature rely mainly on three ideas, namely: aggregating the multiple objectives into one and maximizing this aggregate function; using interactive methods that take preference information from the DM and reduce the solution space based on her responses; and finding the non-dominated frontier (or a subset of it) and presenting it to the DM for further consideration.

Unlike a classical multiobjective programming setting, we assume that the preference model of the DM reflects inequity-aversion, therefore we are interested in finding the set of equitably nondominated points. We now explain the equitable dominance relation that we use in this study.

The following dominance relation is used for a rational decision maker whose preferences can be modelled with a weak preference relation, which is reflexive, transitive and monotonic.

Definition 1 *(Weak Classical (Rational) Dominance)* Consider two solutions to $Model$ 1 with output vectors $z, z' \in \mathcal{Z}$. z rationally dominates z' ($z' \preccurlyeq_r z$) if and only if z is preferred to z' by all rational decision makers. $i.e$

$$z' \preccurlyeq_r z \iff z'_j \leq z_j \ , \ forall \ \ j \in P = \{1, 2, \ldots, p\}.$$

An output vector z is non-dominated ($x : z = z(x)$ is efficient) if there is no z' that dominates it. Note that any rational decision maker's preference relation is assumed to be in line with the classical dominance.

However, in our problem setting, we assume that the DM has equity concerns. To reflect these concerns, we assume two more properties for the preference model, namely: symmetry and Pigou-Dalton principle of transfers.

1. *Symmetry* This property states that the decision maker is indifferent between a feasible solution with an output vector z and any other feasible solution whose output vector is a permutation of the vector z. For example, the DM is indifferent among feasible solutions with output vectors (3, 5, 8), (5, 3, 8) and any other permutation of these.
2. *Pigou-Dalton principle of transfers* This property states that for any two solutions that have same total output, if one solution is obtained by transferring output from a better-off entity to a worse-off one in the other solution, then it is considered better. For example, the DM prefers (5, 5, 6) to (3, 5, 8).

A rational preference relation, which additionally satisfies symmetry and Pigou-Dalton principle of transfers properties, is called an equitable preference relation (Kostreva and Ogryczak 1999).

Definition 2 Consider two solutions to model 1 with output vectors $z, z' \in \mathcal{Z}$. z equitably dominates z' ($z' \preccurlyeq_e z$) if and only if z is preferred to z' by all decision makers with equitable preference relations.

A feasible output vector z is equitably nondominated ($x : z = z(x)$ is equitably efficient) if there is no z' that equitably dominates z. Note that equitable dominance is the generalized Lorenz (GL) dominance discussed in the economics literature (Shorrocks 1983). Hence we will refer to z as *equitably non-dominated* (meaning nondominated in the GL sense).

Theorem 1 [Kostreva and Ogryczak (1999)] $z' \preccurlyeq_e z \iff \sum_{j=1}^{k} \overrightarrow{z'}_j \leq \sum_{j=1}^{k} \overrightarrow{z}_j$ *for all* $k \in P$.
The vector $\overrightarrow{z}$ *is the ordered permutation of* z *with elements ordered in a non-decreasing fashion* $i.e$ $\overrightarrow{z}_j$ *is a vector whose elements express respectively: the minimum outcome, the second minimum outcome, the third minimum outcome, etc. of the outcome vector* z.

Utilizing Theorem 1, finding equitably nondominated solutions to $Model$ 1 is equivalent to finding nondominated solutions to $Model$ 2 below.

$$
\begin{aligned}
&Model\ 2\\
&Max\ \text{“}\overrightarrow{z}_1, \overrightarrow{z}_1 + \overrightarrow{z}_2, \cdots, \sum_{j=1}^{p} \overrightarrow{z}_j\text{"}\\
&s.t.\ x \in \mathcal{X}\\
&z_j = z_j(x)
\end{aligned}
$$

$Model$ 2 is not linear due to the use of the ordering operator $\overrightarrow{(.)}$. However, it has been shown in Ogryczak and Śliwiński (2003) that for any given output vector z, the cumulative ordered elements $\sum_{j=1}^{k} \overrightarrow{z}_j$ for any $k \in P$ can be found by solving the model below:

$$
\begin{aligned}
&Model\ 3\\
&\sum_{j=1}^{k} \overrightarrow{z}_j = Max\ kr_k - \sum_{j=1}^{p} d_{kj}\\
&s.t.\ r_k - d_{kj} \le z_j \quad \forall j \in P \qquad (2)\\
&d_{kj} \ge 0 \quad \forall j \in P \qquad (3)
\end{aligned}
$$

An optimal solution to Model 3 is as follows; Let $r_k^* = \overrightarrow{z}_k$ and

$$
d_{kj}^* = \begin{cases} \overrightarrow{z}_k - z_j, & if \quad \overrightarrow{z}_k \ge z_j \\ 0, & if \quad \overrightarrow{z}_k < z_j \end{cases}
$$

Hence the optimal value is $kr_k^* - \sum_{j=1}^{p} d_{kj}^* = k\overrightarrow{z}_k - \sum_{j:\overrightarrow{z}_k \ge z_j}(\overrightarrow{z}_k - z_j) = \sum_{j=1}^{k} \overrightarrow{z}_j$. Note that alternative optimal solutions can be found by making $r_k^* = \overrightarrow{z}_k + c$ where c is a positive constant. Consequently, we have $d_{kj}^* = \overrightarrow{z}_k + c - z_j$ for $j : \overrightarrow{z}_k \ge z_j$.

$Model$ 2 can be re-formulated as follows:

$$
\begin{aligned}
&Model\ 4\\
&Max\text{“}y_1, y_2, \cdots, y_p\text{"}\\
&s.t.\ y_k - \left(kr_k - \sum_{j=1}^{p} d_{kj} \right) = 0\ \forall k \in P\\
&x \in \mathcal{X}\\
&z_j = z_j(x)\ \forall j \in P\\
&ineq.\ (2), (3) \qquad (4)
\end{aligned}
$$

Here y_k is the sum of the “k" smallest components of any output vector z. For simplicity, let the feasible set of such y vectors in $Model$ 4 be represented by $\{y \in \mathbb{R}^p : y \in \mathcal{Y}\}$. The model transforms the criteria space into cumulative ordered criteria space. Finding the non-dominated solutions to this model is equivalent to finding equitably nondominated solutions to Model 1 (implied by Theorem 1). Such a transformation is illustrated in Fig. 1.

Figure 1a shows eight non-dominated points in the classical dominance sense plotted in the criteria space (in terms of the variables z_k). These points are Pareto optimal points in the rational dominance sense. However, we are interested in finding the equitably non-dominated points among these. To achieve this, we transform the space into the cumulative ordered criteria space (in terms of the variables y_k) and find the non-dominated cumulative ordered vectors (ys) as shown in Fig. 1b. It can be seen that the number of non-dominated

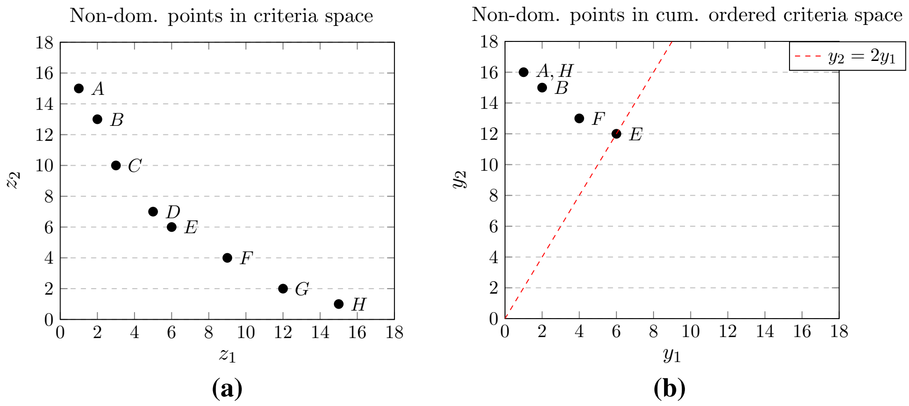


Fig. 1 Non-dom. points in criteria and cum. ordered criteria space

points in Fig. 1b is less than that of Fig. 1a (the set of equitably non-dominated points is a subset of the set of non-dominated points, see Baatar and Wiecek 2006). This is a direct result of the symmetry and Pigou-Dalton principle of transfers properties of the equitable preference relation. The points $(1, 15)$ and $(15, 1)$ in Fig. 1a correspond to the point $(1, 16)$ in Fig. 1b. The points $(5, 7)$ and $(3, 10)$ in Fig. 1a correspond to the points $(5, 12)$ and $(3, 13)$ respectively, in the cumulative ordered space. These points are dominated by $(6, 12)$ and $(4, 13)$ respectively. Moreover, it can be seen from Fig. 1b that for the case when $p = 2$ the transformed points lie in $y_2 \geq 2y_1$ region of the space.

Note that the algorithms developed to generate all non-dominated points for classical MOP can be used to find equitably non-dominated solutions. However, any such algorithm should be modified so that one works on the cumulative space, leading to the equitable MOP (Model 4). This modification is not always trivial due to the ordering operator.

There are many non-dominated points (both in the classical and equitable dominance sense) in large MOPs and it may not be practical or useful to generate them all. One way of handling this computational challenge would be finding the solutions that are of interest to the DM by incorporating her preferences into the solution procedure. We could employ interactive approaches that take the DM's preferences into account and use the information to converge to a single most preferred equitably non-dominated point. Another approach could be generating an evenly distributed subset of the equitably non-dominated points and present them to the DM. In this paper, we discuss two such algorithms.

We first develop an interactive algorithm that finds the most preferred solution for an inequity averse DM. In this approach, we assume that the social welfare function, which is a function of the allocation vectors (z) is symmetric quasiconcave (and hence in line with the properties of inequity-averse preference models). In the second approach, we work on the cumulative ordered space since finding nondominated points in this space is equivalent to finding equitably nondominated points in the original space. The second approach aims to generate evenly distributed equitably non-dominated points and can be used in cases where the DM is not available for interaction.

3 An interactive algorithm

In this section we discuss the algorithm we propose for finding the most preferred equitably nondominated point in equitable multiobjective optimization settings. The algorithm is based on theoretical results that extend previous works both in the classical multiobjective programming settings and equitable multicriteria decision making settings. It extends the results in the classical settings, which rely on the classical dominance concept, by incorporating symmetry and Pigou-Dalton principle of transfers assumptions. Unlike the assumption in a classical setting that the utility function is quasiconcave, we assume that it is symmetric quasiconcave, which increases complexity since the DM is indifferent between all permutations of a given allocation vector. To the best of our knowledge there is no interactive algorithm designed for equitable multiobjective programming problems. The algorithm extends previous work on interactive approaches in equitable multicriteria decision making settings since the problems considered so far in that area are not optimization problems. They assume that the alternatives (allocation vectors) are explicitly given rather than being implicitly defined by constraints. In this work we consider optimization problems.

The interactive algorithm gets preference information in terms of pairwise comparisons and at each iteration it generates an equitably nondominated point that is not inferior to the convex cones generated based on preference information. We first briefly introduce the convex cones concept in the multicriteria decision making settings: both the classical settings where there is no symmetry assumption (see Korhonen et al. 1984; Hazen 1983; Karsu 2013 for more information) and the symmetric setting where the alternatives are explicitly given (see Karsu et al. 2018). We then discuss use of convex cones in the classical optimization settings, where the underlying preference relation is rational. We finally provide our results that help us to extend these ideas to equitable optimization settings and present an interactive algorithm.

3.1 Convex cones in multicriteria evaluation settings

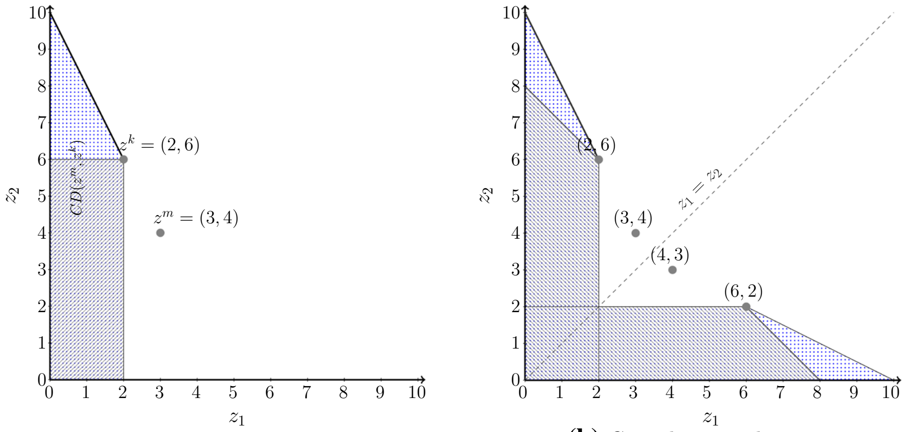
Assume that the DM has provided a pairwise comparison of the form $z^m \succ z^k$, i.e. "the DM prefers distribution z^m to z^k". Distribution z^m is referred to as the *upper generator* of the cone (and polyhedron) and z^k as the *lower* generator. The corresponding cone $C(z^m, z^k)$ and its dominated region $CD(z^m, z^k)$ are as follows.

$$C(z^m, z^k) = \{z | z = z^k + \mu(z^k - z^m), \mu \geq 0\}$$
$$CD(z^m, z^k) = \{z | z \leq z' \text{ for some } z' \in C(z^m, z^k)\}$$

Theorem 2 (Korhonen et al. 1984) *For any $z \in CD(z^m, z^k)$, $u(z) \leq u(z^k) \quad \forall u(.)$ such that $u(.)$ is increasing and strictly quasiconcave.*

An example case is illustrated in Fig. 2a. The figure shows the region of points that are dominated by (2, 6) in the absence of preference information (the region filled with diagonal lines) and the cone dominated region given information that (3, 4) is preferred to (2, 6) (The gray area). When the alternatives are given explicitly, for any alternative z, one can check whether $z \in CD((3, 4), (2, 6))$ by solving systems of linear inequalities and if so, eliminate it from further consideration.

A direct application of Theorem 2 to the symmetric case would lead to computational intractability since symmetry would necessitate checking a set of conditions with respect to every possible combination of all permutations of a set of distributions ((3, 4) and (2, 6)). Moreover, the underlying dominance relation is equitable dominance. In particular, given $z^m \succ z^k$, the cone dominated region in the symmetric case is defined as follows:



(a) Cone dominated region in an asymmetric setting

(b) Cone dominated region in a symmetric setting

Fig. 2 Cone dominated regions in asymmetric and symmetric settings

$$
\begin{aligned}
CD_{Symm}(z^m, z^k) = \{z| \quad & z \preceq_e z' \text{ for some } z' \in C(\Pi^r(z^m); \Pi^s(z^k)) \\
& \text{for some permutations } \Pi^r(z^m) \text{ and} \\
& \Pi^s(z^k) \text{ of } z^m \text{and} z^k \text{ or } z' \in C(\Pi^r(z^k); \Pi^s(z^k)) \text{ for some permutations } \Pi^r(z^k) \text{ and} \\
& \Pi^s(z^k) \text{ of } z^k\}.
\end{aligned}
$$

Figure 2b shows the cone dominated regions with and without preference information, in the symmetric case where the dominance is equitable dominance. Note that (2, 6) is considered equally good as (6, 2), by symmetry, and so there are now two dominated regions. Symmetry also dictates that any of (3, 4) or (4, 3) is preferred to any of (2, 6) and (6, 2) and so the dominated regions increase. For any distribution z such that z falls within any of the two (dotted) enlarged dominated regions we can again infer that $z \preceq z^k$ by the same DM.

The definition of the cone dominated region implies that we need to perform checks by taking into account every permutation of the distributions over which preferences are provided, which may lead to prohibitively large computational effort. Karsu et al. (2018) provide a compact characterization of the cone dominated region which avoids the need for considering all permutational checks, and in some cases avoid them altogether, thus affording tractability by proving the following:

Theorem 3 *Consider a distribution* $z \in \mathbb{R}^p$. *Define* $\overline{CD}_{Symm}(z^m, z^k) = \{z| \ z \preceq_e z'$ *for some* $z' \in C(\overrightarrow{z^m}; \overrightarrow{z^k})\}$ *The following are equivalent:*

(i) $z \in CD_{Symm}(z^m, z^k)$.
(ii) $z \in \overline{CD}_{Symm}(z^m, z^k)$.

This reduces the computational burden since it shows that using the ordered vectors $\overrightarrow{z^m}$ and $\overrightarrow{z^k}$ is sufficient instead of permutational calculations. They use these results and design a ranking algorithm for a problem where the set of alternatives is explicitly given.

The theoretical results provided in the next section extend the previous work of Karsu et al. (2018), who considered the use of convex cones for multicriteria evaluation setting, to the multiobjective optimization setting. The results are also an extension of the work of

Lokman et al. (2016), where cone dominance concept is used for the multiobjective integer programming problems where rational dominance holds, hence there is not issue of symmetry. They prove the following theorem and suggest an effective interactive algorithm based on it.

Theorem 4 Lokman et al. (2016) *Given z^m, z^k such that $z^k \preceq z^m$, an alternative z is dominated by $C(z^m, z^k)$ if and only if the following conditions are satisfied:*

$$z_j^k - z_j \geq 0 \qquad \forall j \in P : z_j^m \geq z_j^k \tag{5}$$

$$(z_i - z_i^k)(z_j^m - z_j^k) \leq (z_j^k - z_j)(z_i^k - z_i^m) \quad \forall j \in P : z_j^m \geq z_j^k \qquad \forall i \in P : z_i^m < z_i^k \tag{6}$$

In the next section we provide our results that help us to extend the analysis to equitable optimization settings.

3.2 Extension for the equitable optimization settings

Based on Theorem 3, to check whether an alternative lies in the cone dominated region in the symmetric case, one should check whether there exists z' such that $z' \in C(\overrightarrow{z^m}, \overrightarrow{z^k})$ and $z \preceq_e z'$. Let $P = \{1, 2, \ldots, p\}$ be the set of entities.

Theorem 5 *Let z and z' be in $\mathbb{R}^p$. Let K^t be the set of t-subsets of P.*
Then $z \preceq_e z'$ if and only if the following holds:
$\sum_{i=1}^t \overrightarrow{z_i} \leq \sum_{j \in a} z'_j \quad \forall t, \forall a \in K^t$.

Proof From Theorem 1, we know the following:

$z \preceq_e z'$ if and only if $\sum_{i=1}^t \overrightarrow{z_i} \leq \sum_{i=1}^t \overrightarrow{z'_i} \quad \forall t \in P$. Note that $\sum_{i=1}^t \overrightarrow{z'_i}$ is the minimum value over all t-sums of z', i.e. $Min_{a \in K^t} \sum_{j \in a} z'_j$. Hence for any t the following holds: $\sum_{i=1}^t \overrightarrow{z_i} \leq \sum_{j \in a} z'_j \quad \forall t, \forall a \in K^t$. □

Example 1 Let $p = 3$. Then $P = \{1, 2, 3\}$, $K^1 = \{1, 2, 3\}$, $K^2 = \{(1, 2), (1, 3), (2, 3)\}$ and $K^3 = \{(1, 2, 3)\}$.

Consider $z = (2, 3, 6)$ and $z' = (2, 4, 5)$. For $t = 1$, we have: $2 \leq 2$, $2 \leq 4$, $2 \leq 5$. For $t = 2$, we have: $2 + 3 \leq 2 + 4$, $2 + 3 \leq 2 + 5$, $2 + 3 \leq 4 + 5$ and for $t = 3$ we have $2 + 3 + 6 \leq 2 + 4 + 5$, hence $z \preceq_e z'$. Now consider $z = (2, 3, 6)$ and $z' = (2.5, 3.5, 4)$ the conditions are violated for some t and some a. For example, $t = 3, 2+3+6 > 2.5+3.5+4$, hence z' does not equitably dominate z.

Definition 3 Consider $\overrightarrow{z^m}$ and $\overrightarrow{z^k}$. For any t-subset of the set $P = \{1, \ldots, p\}$, let K_{mk}^t be the set of t-subsets of P such that $\sum_{i \in b} \overrightarrow{z_i^k} \leq \sum_{i \in b} \overrightarrow{z_i^m} \quad \forall b \in K_{mk}^t$. That is, these are the criteria sets of size t such that the sum of the criteria values of $\overrightarrow{z^m}$ is at least as large as those of $\overrightarrow{z^k}$. Similarly, let K_{km}^t be the sets of t criteria such that $\sum_{i \in c} \overrightarrow{z_i^m} < \sum_{i \in c} \overrightarrow{z_i^k} \quad \forall c \in K_{km}^t$.

Example 2 Consider a case where $\overrightarrow{z^m} = (2, 3, 6)$ and $\overrightarrow{z^k} = (2.5, 3.5, 4)$. Then $K_{mk}^1 = \{3\}$ since $\overrightarrow{z_3^m} \geq \overrightarrow{z_3^k}$ $(6 \geq 4)$. $K_{mk}^2\{(1, 3), (2, 3)\}$ since $\overrightarrow{z_1^m} + \overrightarrow{z_3^m} \geq \overrightarrow{z_1^k} + \overrightarrow{z_3^k}$ $(2 + 6 \geq 2.5 + 4)$ and $\overrightarrow{z_2^m} + \overrightarrow{z_3^m} \geq \overrightarrow{z_2^k} + \overrightarrow{z_3^k}$ $(3 + 6 \geq 3.5 + 4)$. $K_{mk}^3 = \{(1, 2, 3)\}$ as $\overrightarrow{z_1^m} + \overrightarrow{z_2^m} + \overrightarrow{z_3^m} \geq \overrightarrow{z_1^k} + \overrightarrow{z_2^k} + \overrightarrow{z_3^k}$ $(2 + 3 + 6 \geq 2.5 + 3.5 + 4)$. Therefore, $K_{km}^1 = \{1, 2\}$, $K_{km}^2 = \{(1, 2)\}$ and $K_{km}^3 = \emptyset$.

Theorem 6 *Given z^m, z^k such that $z^k \preceq z^m$, $z \in C\bar{D}_{Symm}(z^m, z^k)$ (it is equitably dominated by $C(\overrightarrow{z^m}; \overrightarrow{z^k})$) if and only if the following conditions are satisfied:*

$$\sum_{j \in b} \overrightarrow{z_j^k} - \sum_{i=1}^{t} \overrightarrow{z_i} \geq 0 \quad \forall t \in P, \forall b \in K_{mk}^t \tag{7}$$

$$\left(\sum_{i=1}^{t_2} \overrightarrow{z_i} - \sum_{l \in c} \overrightarrow{z_l^k}\right)\left(\sum_{j \in b} (\overrightarrow{z_j^m} - \overrightarrow{z_j^k})\right) \leq \left(\sum_{j \in b} \overrightarrow{z_j^k} - \sum_{i=1}^{t_1} \overrightarrow{z_i}\right)\left(\sum_{l \in c} (\overrightarrow{z_l^k} - \overrightarrow{z_l^m})\right)$$
$$\forall t_1 \in P, \forall t_2 \in P,\ b \in K_{mk}^{t_1},\ c \in K_{km}^{t_2} \tag{8}$$

The proof of the Theorem is provided in "Appendix A". The theorem implies that for an outcome vector z to be not cone dominated at least one of these two conditions should be violated. In the next section we discuss the interactive algorithm we designed based on this observation.

3.3 The algorithm

The algorithm starts with finding an initial incumbent. Then a challenger to the incumbent is generated (another equitably nondominated point) and the decision maker is asked about her preferences on the pair. This will lead to eliminating the inferior one in the pair and declaring the superior one as the current incumbent (z^{inc}). Given preference information of the DM, another equitably nondominated point is generated solving a scalarization model with additional constraints restricting the alternative from being cone dominated, i.e. not being in the cone dominated areas. The generated alternative is compared to the incumbent and new preference information is obtained. This loop is repeated until there is no feasible equitably nondominated alternative that is not dominated by the cones so far (i.e., the scalarization model becomes infeasible.)

The scalarization model that aims to find a new equitably nondominated solution that is not cone dominated is as follows ($Pref$ is the set of pairs of alternatives over which the DM provided preference information):

$$(Scalarization \quad Model)$$

$$Max \quad \sum_{t \in P} w_t \sum_{i=1}^{t} \overrightarrow{z_i}$$

$$s.t.$$

$$x \in \mathcal{X}$$

$$z_i = z_i(x) \quad \forall i \in P \tag{9}$$

$$\left(\sum_{j \in b} \overrightarrow{z_j^k} + \epsilon\right) y_{tb}^{mk} \leq \sum_{i=1}^{t} \overrightarrow{z_i} \quad \forall (z^m, z^k) \in Pref, \forall t \in P, \forall b \in K_{mk}^t \tag{10}$$

$$\left[\sum_{j \in b} \overrightarrow{z_j^k} \left(\sum_{l \in c} (\overrightarrow{z_l^k} - \overrightarrow{z_l^m})\right) + \sum_{l \in c} \overrightarrow{z_l^k} \left(\sum_{j \in b} (\overrightarrow{z_j^m} - \overrightarrow{z_j^k})\right) + \epsilon\right] h_{t_1 t_2 bc}^{mk} \leq$$

$$\sum_{i=1}^{t_2} \overrightarrow{z_i} \left(\sum_{j \in b} (\overrightarrow{z_j^m} - \overrightarrow{z_j^k}) \right) + \sum_{i=1}^{t_1} \overrightarrow{z_i} \left(\sum_{l \in c} (\overrightarrow{z_l^k} - \overrightarrow{z_l^m}) \right)$$

$$\forall (z^m, z^k) \in Pref, \forall t_1 \in P, \forall t_2 \in P, b \in K_{mk}^{t_1}, c \in K_{km}^{t_2} \tag{11}$$

$$\sum_{t \in P} \sum_{b \in K_{mk}^t} y_{tb}^{mk} + \sum_{t_1 \in P} \sum_{t_2 \in P} \sum_{c \in K_{km}^{t_2}} \sum_{b \in K_{mk}^{t_1}} h_{t_1 t_2 bc}^{mk} = 1 \quad \forall (z^m, z^k) \in Pref \tag{12}$$

$$\left(\sum_{i=1}^{t} \overrightarrow{z_i^{inc}} + \epsilon \right) a^t \leq \sum_{i=1}^{t} \overrightarrow{z_i} \quad \forall t \in P \tag{13}$$

$$\sum_{t \in P} a^t = 1 \tag{14}$$

$$y_{tb}^{mk} \in \{0, 1\} \forall (z^m, z^k) \in Pref, \forall t \in P, \forall b \in K_{mk}^t \tag{15}$$

$$h_{t_1 t_2 bc}^{mk} \in \{0, 1\} \forall (z^m, z^k) \in Pref, \forall t_1 \in P, \forall t_2 \in P, \forall b \in K_{mk}^{t_1}, \forall c \in K_{km}^{t_2} \tag{16}$$

$$a^t \in \{0, 1\} \quad \forall t \in P \tag{17}$$

Constraint sets 10 and 11 control whether conditions 7 and 8 are satisfied through the use of binary variables y_{tb}^{mk} and $h_{t_1 t_2 bc}^{mk}$. Constraint 12 ensures that at least one of the conditions is violated for each pair of solutions in $Pref$, so that the solution is not in the corresponding cone dominated region. To guarantee that at least one of conditions 7 and 8 is violated, ensuring that one such binary variable y_{tb}^{mk} and $h_{t_1 t_2 bc}^{mk}$ for each pair takes a value of one is sufficient, hence constraint 12 is an equality constraint. This is because, even if there exist points that violate more of the conditions, one of the binary variables will take a value of one (ensuring nondominance) and the other binary variables (y_{tb}^{mk} or $h_{t_1 t_2 bc}^{mk}$) can take values of zero, making the corresponding constraints (in sets 10 and 11) redundant.

Constraints 13 and 14 are used to ensure that the solution will not be equitably dominated by the current incumbent. That is, they ensure that the new solution will be strictly better than the incumbent in at least one component when their cumulative ordered maps are compared. Note that formulating 14 as an equality constraint does not leave out solutions that are better than the incumbent in multiple components. If the solution is better in more than one component, then all a^t variables except one will be zero, this will only make the corresponding constraints redundant. Hence, constraint 14 is sufficient for the model to return an equitably nondominated point (if it exists). The ϵ parameter used in constraint sets 10 and 11 is a sufficiently small number such that conditions 7 and 8 are violated and ϵ in constraint set 13 is a sufficiently small number such that the newly found point is strictly better than the current incumbent in one component.

Note that the model requires the cumulative ordered map of the decision variable vector z_i ($\sum_{i=1}^{t} \overrightarrow{z_i}$ values for all t), which requires it to be ordered endogenously. This can easily be ensured by adding the auxiliary variables r_t and d_{ts} as follows. Here, $\sum_{i=1}^{t} \overrightarrow{z_i}$ term is replaced with $tr_t - \sum_{s \in P} d_{ts}$ and additional constraints are added to make sure that the replacement works (Recall Model 3).

Linearized Scalarization Model (LSM)

$$Max \quad \sum_{t \in P} w_t \left(tr_t - \sum_{s \in P} d_{ts} \right)$$

$s.t.$

$x \in \mathcal{X}$

$$z_i = z_i(x) \quad \forall i \in P$$

$$\left(\sum_{j\in b} \overrightarrow{z_j^k} + \epsilon\right) y_{tb}^{mk} \leq \left(tr_t - \sum_{s\in P} d_{ts}\right) \quad \forall (z^m, z^k) \in Pref, \forall t \in P, \forall b \in K_{mk}^t$$

$$\left[\sum_{j\in b} \overrightarrow{z_j^k} \left(\sum_{l\in c} (\overrightarrow{z_l^k} - \overrightarrow{z_l^m})\right) + \sum_{l\in c} \overrightarrow{z_l^k} \left(\sum_{j\in b} (\overrightarrow{z_j^m} - \overrightarrow{z_j^k})\right) + \epsilon\right] h_{t_1 t_2 bc}^{mk} \leq$$

$$\left(t_2 r_{t_2} - \sum_{s\in P} d_{t_2 s}\right)\left(\sum_{j\in b} (\overrightarrow{z_j^m} - \overrightarrow{z_j^k})\right) + (t_1 r_{t_1} - \sum_{s\in P} d_{t_1 s})\left(\sum_{l\in c} (\overrightarrow{z_l^k} - \overrightarrow{z_l^m})\right)$$

$$\forall (z^m, z^k) \in Pref, \forall t_1 \in P, \forall t_2 \in P, b \in K_{mk}^{t_1}, c \in K_{km}^{t_2}$$

$$\left(\sum_{i=1}^{t} \overrightarrow{z_i^{inc}} + \epsilon\right) a^t \leq tr_t - \sum_{s\in P} d_{ts} \quad \forall t \in P$$

$$r_t - d_{ts} \leq z_s \quad \forall t \in P, \forall s \in P$$

$$r_t \geq 0 \quad t \in P$$

$$d_{ts} \geq 0 \quad t, s \in P$$

Constraint sets 12, 14, 15, 16, 17.

Note that any weight vector $w \in \mathbb{R}^p : w > 0$ can be used in the scalarization model. We, however find a (potentially new) w at each iteration using the preference information provided so far by solving the following model:

$$Weight \quad estimation \quad model \quad (WEM)$$

$$Max \quad \varepsilon$$

$$s.t.$$

$$\sum_{t\in P} w_t \sum_{i=1}^{t} (\overrightarrow{z_i^m} - \overrightarrow{z_i^k}) \geq \varepsilon \quad \forall (z^m, z^k) \in Pref$$

$$\sum_{t\in P} w_t = 1$$

$$w_t \geq \varepsilon$$

Note that this model finds weights of a linear value function over the cumulative ordered vectors that is in line with the given preference information. However, since the value function is not necessarily linear, the model may turn out to be infeasible. We use equal weights in that case.

The algorithm is as follows:

Step 1: Initialization: Set $z^{inc} = 0$, $Pref = \emptyset$ and $w_t = 1/p \quad \forall t$.
Solve LSM. If the model is infeasible, go to Step 4.
Otherwise let the solution be the incumbent, z^{inc}. Solve LSM. Let the solution be z^{ch} (This is a challenger solution).

Step 2: Present z^{inc} and z^{ch} to the DM. Let the preferred one be z^m and the less preferred one be z^k. $Pref = Pref \cup (z^m, z^k)$. Set $z^{inc} = z^m$. Solve WEM and let the solution be w.

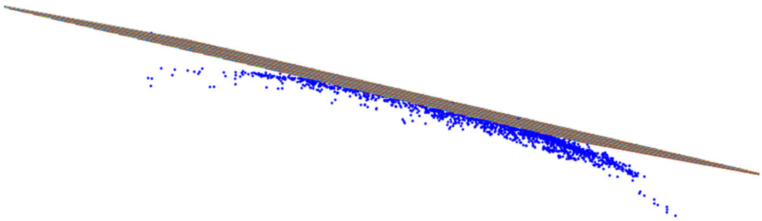


Fig. 3 Hyperplane fitted using point with maximum total outcome value

Step 3: Solve LSM.
If the model is infeasible, go to Step 4.
Otherwise let the solution be the z^{ch}. Go back to Step 2.
Step 4: Stop and return z^{inc}.

The interactive algorithm can be used whenever the DM is available for providing preference information. If not, we propose another algorithm that generates an evenly spread subset of the set of equitably nondominated points.

4 Algorithm for generating evenly distributed equitably nondominated points (The GEND algorithm)

This approach focuses on finding an evenly distributed subset of nondominated points to model 4, which are equitably nondominated solutions to the original problem (model 1). The approach is based on fitting a hyperplane function that is close to the nondominated frontier in the cumulative ordered space (hence nondominated frontier in the equitable dominance sense). We then select representative points on the hyperplane, generate regions around those points and search those regions for non-dominated points. This way, we generate a subset of the set of equitably non-dominated solutions that is well spread.

The hyperplane could be placed at different positions relative to the non-dominated frontier. Figure 3 below shows the hyperplane placed above the non-dominated frontier for a maximization setting.

In the next two subsections we explain the three main parts of the algorithm, fitting the hyperplane, defining the regions to be explored and finding the solutions within the specified regions.

4.1 Fitting the hyperplane and defining the regions

The hyperplane we fit is of the form $\sum_{k=1}^{p} y_k = T$. We explore the strategy of fitting a hyperplane that passes through the solution that has the maximum total outcome value. Hence, we set $T = \{max \sum_{k=1}^{p} y_k : y \in \mathcal{Y}\}$ to fit the hyperplane above the frontier.

We define well spread regions in the cumulative ordered criteria space around some selected representative points on the hyperplane fitted. As shown in Fig. 1b, in $\mathbb{R}^2$, the non-dominated points in the cumulative ordered space are restricted to the region defined by the polyhedron $Q = \{y \in \mathbb{R}^2 : y_2 \geq 2y_1\}$. A similar analogy can be made for higher dimensional real spaces. For example, in $\mathbb{R}^3$ the region is defined by $Q = \{y \in \mathbb{R}^3 : y_2 \geq 2y_1,\ y_3 \geq 3y_1,\ 2y_3 \geq 3y_2\}$. In general, we can state the proposition below:

Proposition 7 *For any real space $\mathbb{R}^p$, let $P = \{1, 2, \cdots, p\}$. The non-dominated points in the cumulative ordered space are restricted to the polyhedron $Q = \{y \in \mathbb{R}^p : jy_k \geq ky_j \ \ forall\ j, k \in P : k > j\}$.*

The proof of Proposition 7 is provided in "Appendix B". To define the regions, we select a number of representative points yr from the restricted polyhedron Q that lie on the hyperplane defined above since Proposition 7 implies that there is no need to focus on the region outside Q.
Any representative point $yr \in \mathbb{R}^p$ on the fitted surface will most likely be an infeasible or dominated point, so we use it as a reference point only to define a region around it and then generate feasible non-dominated points in the region. Note that the region defined around yr may or may not contain any non-dominated points in it, depending on the size of the region. In order to guarantee obtaining a set of non-dominated points in the region, we first find the non-dominated points $yrt \in \mathbb{R}^p$ and $yrl \in \mathbb{R}^p$, that are at minimum Tchebycheff and linear distance from the ideal point y^{IP} in the direction of the reference point yr by solving the problems M_{chev} and M_{linr}, respectively:

$$
\begin{aligned}
&(M_{chev})\\
&Min\ \ \rho_{max} - \varepsilon_1 * \sum_{k=1}^{p} y_k\\
&s.t.\ \ \rho_{max} \geq \lambda_k(y_k^{IP} - y_k) \qquad \forall k \in P\\
&y \in \mathcal{Y}
\end{aligned}
$$

$$
\begin{aligned}
&(M_{linr})\\
&Min\ \sum_{k=1}^{p} \lambda_k(y_k^{IP} - y_k) - \varepsilon_1 \sum_{k=1}^{p} y_k\\
&s.t.\ \ y \in \mathcal{Y},
\end{aligned}
$$

where ε_1 is a sufficiently small positive constant. The weight vector $\lambda \in \mathbb{R}^p$, corresponds to the diagonal direction for the reference point yr from the ideal point y^{IP} as follows (Steuer 1986, p. 425):

$$
\lambda_k = \begin{cases} \frac{1}{(y_k^{IP} - yr_k)} \left[\sum_{j=1}^{p} \frac{1}{(y_j^{IP} - yr_j)} \right]^{-1} & \text{if} \ \ yr_k \neq y_k^{IP} \ \ \forall k \in P \\ 1 & \text{if} \ \ yr_k = y_k^{IP} \\ 0 & \text{if} \ \ yr_k \neq y_k^{IP} \ \text{but} \ \exists j \in P : yr_j = y_k^{IP} \end{cases}
$$

For any reference point yr, let yrt and yrl be the optimal solutions obtained from solving M_{chev} and M_{linr} respectively. We determine the region by defining upper and lower bounds (UB and LB) as follows: $UB_k = max(yrt_k, yrl_k)\ LB_k = min(yrt_k, yrl_k) \forall k \in P$ (see Fig. 4). Note that we have multiple reference points, hence there is a possibility of generating intersecting regions. However, the size of the intersecting areas of the regions can be mitigated if the reference points are chosen in an appropriate manner. Furthermore, we could unify the

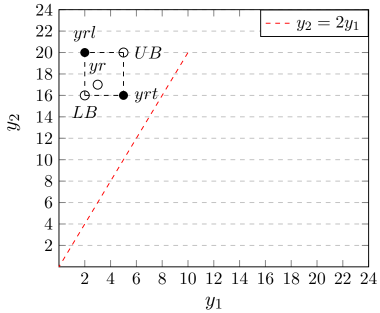


Fig. 4 Generating regions around reference points

intersecting regions in order to eliminate the possibility of generating the same solution from two or more generated regions.

In the next section, we explain the algorithm we used to generate the non-dominated points in the regions.

4.2 Generating non-dominated points in the defined regions

The method we use to generate equitably nondominated points in the defined regions is based on the epsilon-constraint scalarization (see Lokman and Köksalan 2013, 2014). The algorithm generates non-dominated points in evenly distributed subsets of the feasible set, i.e. the regions generated around reference points. To explore each region, we solve scalarization models with additional constraints $LB_k \leq y_k \leq UB_k$ $k = 1, 2, \ldots, p$.

We initialize the algorithm by arbitrarily choosing a region, r to begin with and a criterion, n. We then find the point that maximizes the nth criterion value in the region by solving

$$
\begin{aligned}
&(M_n^0)\\
&\quad Max \;\; y_n + \varepsilon_1 \sum_{k \neq n} y_k\\
&s.t.\\
&y_k \geq LB_k \;\forall k \in P && (18)\\
&y_k \leq UB_k \;\forall k \in P\\
&y \in \mathcal{Y} && (19)
\end{aligned}
$$

where ε_1 is as in Sect. 4.1 and the augmented part of the objective function (term with ε_1) is used to make sure the model returns a non-dominated point in the region. The optimal solution to the model (M_n^0) above, denoted by $\hat{y}^0 \in \mathbb{R}^p$, may or may not be dominated by a feasible point outside the region. We solve the (MD_n^ω) model below with $\omega = 0$ to check whether the point $\hat{y}^0$ is dominated or not.

$$
\begin{aligned}
&(MD_n^\omega)\\
&Max \; y_n + \varepsilon_1 \sum_{k \neq n} y_k\\
&s.t\\
&y_k \geq \hat{y}_k^w \;\; \forall k \in P\\
&y \in \mathcal{Y} && (20)
\end{aligned}
$$

Let the optimal solution to (MD_n^0) be $\overline{y}^0 \in \mathbb{R}^p$. If $\overline{y}_k^0 = \hat{y}_k^0\ \forall k \in P$, then there does not exist a feasible point dominating $\hat{y}^0$. Then $\hat{y}^0$ is placed in the set of non-dominated points Ω_r that are in this region. We repeatedly generate new points in this region and check whether every obtained point is dominated or not. At every iteration ω, we use the epsilon constraint scalarization to find a new point y^ω. To make sure that the scalarization model provides a new solution, we utilize additional constraints that ensure that the new solution is different from (and not dominated by) the previously found nondominated solutions, including the ones generated outside the region.

At every iteration ω of the algorithm (for $\omega > 0$) a new non-dominated point is generated solving model (M_n^ω), until it becomes infeasible.

$$(M_n^\omega)$$

$$Max\ \ y_n + \varepsilon_1 \sum_{j \neq n} y_j$$

$$s.t.$$

$$y_k \geq (\hat{y}_k^\tau + 1)h_{\tau k} - BM(1 - h_{\tau k})\ \forall \tau = 0, \cdots, \omega\ \forall k \neq n \tag{21}$$

$$\sum_{k \neq n} h_{\tau k} = 1\ \forall \tau = 0, \cdots, \omega \tag{22}$$

$$h_{\tau k} \in \{0, 1\}\ \forall \tau = 0, \cdots, \omega\ \forall k \neq n,$$

$$y \in \mathcal{Y}$$

$$ineq.\ (18), (19) \tag{23}$$

If (M_n^ω) is feasible, the solution found, $\hat{y}^\omega$ is a non-dominated point in this region and it is not identical to any of previously generated non-dominated points from this region (in set Ω_r). This is guaranteed by using auxiliary binary variables $h_{\tau k}$ and constraints 21 and 22, which ensure that the solution is better than the previously found solutions in at least one criterion. Then MD_n^ω is solved to see if $\hat{y}^\omega$ is a nondominated solution of the original model and if so, it is added to Ω_r. BM in condition 21 is a sufficiently large number.

We implement the algorithm in every region and generate the non-dominated points in the regions. The set $\Omega = \Omega_1 \cup \Omega_2 \cup \cdots \cup \Omega_m$ is the set of generated non-dominated points. Hence, we obtain subsets of the non-dominated frontier that lie on different parts of the frontier.

In a nutshell, the GEND algorithm is as follows:

Step 1: Fit a hyperplane to the frontier.
Step 2: Select m reference points on the fitted surface.
Step 3: Generate non-dominated solutions in the neighbourhood of the selected points. For each reference point, find the non-dominated points yrt and yrl and define a region with its upper and lower bound vectors.
Step 4: Generate non-dominated points in the regions defined in step 4 above.

a: (Initialization). Enumerate the m regions. Select the first region (set $r = 1$) to explore and a criterion, n, to maximize. Set $\omega = 0$, $\Omega = \emptyset$ and $\Omega_1 = \emptyset$.
b: (Generating a new point). Solve the (M_n^ω) model. If (M_n^ω) is feasible, denote the optimal point as $\hat{y}^\omega \in \mathbb{R}^p$ and go to step $4c$. Otherwise, go to step $4d$.
c: (Checking for dominance). Solve (MD_n^ω) to check whether $\hat{y}^\omega$ is non-dominated. Let the optimal solution be $\overline{y}^\omega$. If $\overline{y}_k^\omega = \hat{y}_k^\omega\ \forall k \in P$, then $\Omega_r \cup \hat{y}^\omega \rightarrow \Omega_r$. Go to step $4b$.
d: Stop. Ω_r is the entire set of non-dominated points in region r.

e: If $r = m$, stop, $\Omega = \Omega_1 \cup \Omega_2 \cup \cdots \cup \Omega_m$. Else, set $r + 1 \longrightarrow r$, set $\omega = 0$, $\Omega_r = \emptyset$ and go to Step $4b$.

5 Computational experiments

In this section, we illustrate the two approaches on equitable multi-objective knapsack problems. The classical multi-objective binary knapsack problem is an extension of the single-objective binary knapsack problem (Silvano and Paolo 1990; Kellerer et al. 2003) where each item is associated with a vector of outputs instead of a single output value. There have been recent attempts to develop fast and efficient exact and approximate solution algorithms to multi(bi)-objective knapsack problems (Visée et al. 1998; Klamroth and Wiecek 2000; Bazgan et al. 2009; Figueira et al. 2013; Mansour and Alaya 2015; Mansour et al. 2018, 2019).

We will consider multi-objective binary knapsack problems where the decision maker has equity concerns (E-MOBKP). We assume that the preference model of the decision maker satisfies properties related to inequity-aversion and try to find the set of equitably efficient portfolios that result in equitably non-dominated output vectors.

Consider a setting where there are n project proposals that provide outputs to p entities. Let $P = \{1, 2, \cdots, p\}$ be the set of entities and $N = \{1, 2, \cdots, n\}$ be the set of proposals. Every project i is expected to generate an output value of o_{ij} for entity j and consumes c_i units of resource. Assume that the decision maker would like to select and fund a portfolio of these projects, which results in an equitable distribution of outputs across the p entities. The total amount of available resource is denoted by B, which is generally not enough to initiate all the projects. The decision to be made here, is whether to initiate a project or not, i.e. partial funding is not possible. The decision variables are as follows:

$$x_i = \begin{cases} 1, & if\ project\ i\ is\ initiated \\ 0, & otherwise \end{cases}$$

The aim is finding equitably nondominated solutions to the following problem:

$$Max \text{ ``}z_1, z_2, \cdots, z_p\text{''}$$

$$s.t. \quad \sum_{i=1}^{n} c_i x_i \leq B \tag{24}$$

$$z_j - \sum_{i=1}^{n} o_{ij} x_i = 0 \quad \forall j \in P \tag{25}$$

$$x_i \in \{0, 1\} \quad \forall i \in N \tag{26}$$

Note that in GEND we work in the cumulative ordered space and aim to find the nondominated solutions to the following problem:

$$Max \text{ ``}y_1, y_2, \cdots, y_p\text{''}$$

$$s.t. \quad y_k - \left(k r_k - \sum_{j=1}^{p} d_{kj} \right) = 0 \, \forall k \in P \tag{27}$$

$$r_k - d_{kj} - z_j \leq 0 \quad \forall j, k \in P \tag{28}$$

$$d_{kj} \geq 0 \quad \forall j, k \in P \tag{29}$$

$$\text{constraints } 24, 25, 26 \tag{30}$$

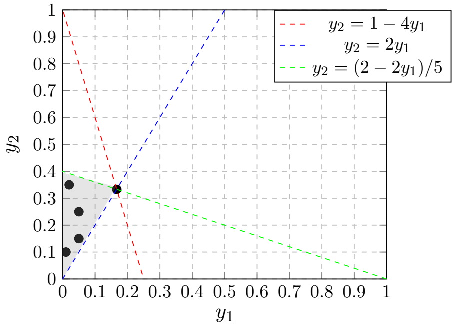


Fig. 5 Selected weight values in Q_μ

We performed experiments to see whether the proposed algorithms provide satisfactory results. The experiments were conducted on randomly generated multi-objective knapsack problems with three objectives (entities (p)). In these instances, the cost and the output values are generated randomly using uniform distributions.

The algorithms are coded in Visual C++ and solved by a computer with an Intel Xeon E5 3.60 GHz processor and 32 GB RAM. The solution times are expressed in central processing unit (CPU) seconds. All mathematical models are solved with CPLEX 12.7.

We created three sets of problem instances by generating integer c_i and o_{ij} values in the ranges [1, 10], [1, 50] and [1, 1000], respectively. Different values are used for the total number of items n. For each n, 10 problem instances are generated. For every instance, the total budget B is set as $\sum_{i=1}^{n} c_i/2$.

For the interactive algorithm, we simulated the responses of the inequity-averse decision maker by using a symmetric quasi-concave function of the following form: $u(z) = \sum_{j=1}^{p-1} \sum_{k \in P: k>j} min(z_j, z_k)$. That is, the utility score is assumed to be equal to the sum of pairwise minima.

We implemented the GEND algorithm with a set of five reference points in the polyhedron Q that have a total benefit of T. For the tri-objective case, the points are in the set $\{yr_1, yr_2, yr_3 \in \mathbb{R}_+ : yr_1 + yr_2 + yr_3 = T,\ yr_3 \geq 3yr_1,\ 2yr_3 \geq 3yr_2,\ yr_2 \geq 2yr_1\}$. Moreover, we can define each element of the reference point yr_k as a fraction of T, i.e, we define a weight vector $\mu \in \mathbb{R}^3$ where $yr_k = \mu_k T$, $\sum_{k=1}^{m} \mu_k = 1$ and $\mu_k \geq 0 \ \ \forall k \in P$. In $\mathbb{R}^3$, the problem reduces to that of finding μ values that lie in the polyhedron $Q_\mu = \{\mu \in \mathbb{R}^3 : \sum_{k=1}^{3} \mu_k = 1,\ \mu_3 \geq 3\mu_1,\ 2\mu_3 \geq 3\mu_2,\ \mu_2 \geq 2\mu_1,\ \mu_k \geq 0 \ \ \forall k \in P\}$. We chose five weight values in Q_μ that are spread. Figure 5 below shows Q_μ and the selected points.

We use these points as the reference points to generate regions as shown in Sect. 4.1. Moreover, we scale the regions by a factor α in the interval [0, 1] to enlarge the generated region. For any given region we scale it by making its lower bound and upper bound $(1-\alpha)LB$ and $(1+\alpha)UB$ respectively. We set $\alpha = 0.005$.

Tables 1 and 2 summarize the results of the approaches for Set 1 and Set 2 instances whose parameters are generated in the ranges [1, 10] and [1, 50], respectively. We first generated the whole equitably non-dominated frontier for these problem instances using the epsilon

Table 1 Results of epsilon constraint approach, interactive approach and GEND: Set 1

n	Epsilon constraint approach				Interactive approach				GEND algorithm			
	# of solutions		Solution time (s)		# of questions		Solution time (s)		# of solutions		Solution time (s)	
	Avg	Max	Avg	Max	Avg	Max	Avg	Max	Avg	Max	Avg	Max
250	26.70	102	10.58	48.92	4.90	15	1.79	8.96	7.4	25	2.98	7.84
300	69.4	203	22.63	47.76	8.10	18	5.36	19.83	21.3	48	12.82	42.12
350	32.4	79	13.66	35.7	4.70	10	1.79	5.54	9.8	27	6.67	15.09
400	23.8	52	8.01	28.44	4.20	10	1.71	3.74	10.3	24	6.26	14.15
450	26.4	144	11.20	53.92	5.00	19	14.60	85.43	8.3	41	12.21	83.34
500*	74	268	28.48	89.54	6.40	10	5.12	9.55	30.56	154	27.52	162.68

* The epsilon constraint algorithm failed to return the solution set in one of the instances, hence the results are reported for 9 instances

Table 2 Results of epsilon constraint approach, interactive approach and GEND: Set 2

n	Epsilon constraint approach				Interactive approach				GEND algorithm			
	# of solutions		Solution time (s)		# of questions		Solution time (s)		# of solutions		Solution time (s)	
	Avg	Max	Avg	Max	Avg	Max	Avg	Max	Avg	Max	Avg	Max
50	16.9	49	3.49	12.57	8.7	22	2.17	8.67	5.4	10	0.52	1.10
100	43.1	106	23.32	82.06	10.9	27	7.26	37.72	8.7	20	2.60	5.18
150	55.7	304	43.22	305.55	9.5	23	28.72	186.29	11.1	39	11.09	39.73
200	285.2	1676	341.62	2201.05	79	685	687.02	6178.46	32.7	133	105.84	664.64
250	93.4	328	96.62	410.07	16.2	74	223.59	1762.05	20.9	91	52.01	182.97

constraint method in the cumulative ordered space (Laumanns et al. 2006). It shows the average and the maximum values for the number of equitably non-dominated solutions, and the time it takes to generate these solutions. As for the interactive approach, we report the average and maximum values for the number of questions asked and the solution time. The results of the GEND algorithm are also summarized.

5.1 Discussion

It is observed that the interactive algorithm provides promising results since it returns the most preferred solution of the DM in significantly less time compared to the time it takes to generate the whole set and after asking a relatively small number of pairwise comparison questions for Set 1 instances. The GEND algorithm performs comparably to the epsilon-constraint approach in terms of the solution time but the advantage is not that clear as in some instances the solution times of the algorithm are closer to those of the epsilon constraint approach. In Set 2 instances, the computational advantage of the GEND algorithm is more clear. It is seen that, solution time of the interactive algorithm exceeds that of the epsilon constraint approach, especially in instances with very large number of equitably nondominated solutions. However, one should note that the interactive approach guides the DM to her most preferred solution while epsilon constraint and GEND only present a (sometimes quite large) set of solutions to the DM.

Tables 3 and 4 summarize the results of the approaches for the randomly generated instances where c_i and o_{ij} are generated in the range [1, 1000] (Set 3 instances). For each combination of p and n, 10 problem instances are generated.

The results for the GEND algorithm are shown in Table 3. It is seen that the average number of solutions found and hence the computational effort increases significantly as the number of projects (n) increases. In the case where $n = 200$, there is a problem instance with 2143 solutions. Removing this instance will reduce the average solution times to 1764.82 and 915.45 seconds for the epsilon constraint method and the GEND algorithm respectively. We can observe that on average we generate a significant portion of the equitably non-dominated set using GEND approach in a fraction of the average time it takes to generate the whole set.

The interactive algorithm involves relatively large parameters derived from multiplication, therefore in these instances where the parameters of the original knapsack problem is large, numerical issues might arise. We resolve this issue by dividing the problem parameters c_i and o_{ij} by 100, hence the parameters are not integer any more.

Table 4 presents the results of the interactive algorithm. We report the average and maximum values for the number of questions asked and the solution time for ϵ values of 0.09 and 0.9. We report the number of questions asked and the solution times for the ϵ value of 0.09 where the most preferred solution of the DM is always returned. However, the solution times and number of questions asked are significantly increasing as the number of projects (n) increases and significantly higher than in the case where ϵ is 0.9. Increasing the value of ϵ to 0.9 leads to fewer number of questions albeit may not lead to the most preferred solution of the DM. Hence we may consider solving the interactive approach with higher ϵ values to obtain meaningful heuristics.

We evaluate the quality of the solutions generated by the GEND algorithm by three performance measures namely, the coverage error Sayın (2000), coverage gap Ceyhan et al. (2019) and another spread measure.

We use a Tchebycheff distance-based coverage error and coverage gap to measure how well the set of solutions generated by the GEND algorithm ($G \subseteq \mathcal{Y}$) covers the whole set of

Table 3 Results of the epsilon-constraint approach and GEND: Set 3

n	Epsilon constraint approach				GEND algorithm			
	Solution time (s)		# of solutions		Solution time (s)		# of solutions	
	Avg	Max	Avg	Max	Avg	Max	Avg	Max
50	13.66	62.4	19.9	58	0.91	1.98	9.2	18
100	89.2	257.02	44.5	105	3.6	7.79	17.8	43
150	180.71	1222.12	47.6	205	18.22	93.92	22.2	84
200*	44535.85	429475.14	363.9	2143	2441.88	16179.78	76.2	291

*The GEND algorithm for these instances was implemented with $\alpha = 0.003$

Table 4 Results of the interactive approach: Set 3

n	Interactive approach $\epsilon = 0.09$				Interactive approach $\epsilon = 0.9$			
	# of questions		Solution time (s)		# of questions		Solution time (s)	
	Avg	Max	Avg	Max	Avg	Max	Avg	Max
50	9.9	20	2.187	5.79	5.6	12	0.863	2.72
100	12.1	25	15.809	50.91	6.2	15	2.389	7.89
150	7.5	22	9.821	69.65	4	10	1.327	4.28
200	36.1	92	1998.093	7961.11	11.8	30	317.617	2880.13

solutions. Let $r \in \mathbb{R}^p$ be a solution such that $r \notin G$, the measure of how well G covers r is given by $\beta_{CE}(r) = min_{g \in G}\{max_{i \in P}|r_i - g_i|\}$ and $\beta_{CG}(r) = min_{g \in G}\{max_{i \in P}(r_i - g_i)\}$ for coverage error and coverage gap respectively. Then the coverage error (gap) of G, β_{CE} (β_{CG}), given by the worst covered point is $\beta_{CE} = max_{r \in \mathcal{Y}}\beta_{CE}(r)$ ($\beta_{CG} = max_{r \in \mathcal{Y}}\beta_{CG}(r)$). We report the average and standard deviation of the scaled coverage error (gap) values for every n. For every instance, we calculate the efficient ranges of the objective functions and take the ratio between the coverage value and the maximum of the efficient ranges as the *scaled coverage error (gap) value*.

To have further information on the spread of the solutions found by the GEND algorithm, we divided the cumulative criteria space into 125 boxes of equal dimensions. As expected, only some of the boxes contain equitably non-dominated points. We report the percentage of the non-empty boxes that at least one solution is found by the GEND algorithm. Tables 5 and 6 below report results of the GEND algorithm for Set 1 and Set 2 instances. As seen in the Tables the scaled coverage gap values are satisfactory. Moreover, the algorithm is able to find representative solutions in 40% to 50% of the non empty boxes on average.

As seen in Table 7 the algorithm is able to find representative solutions in 35% to 60% of the non empty boxes on average for the set 3 instances. The percentages tend to drop as the problem gets larger.

We also performed experiments for the five objectives ($p = 5$) setting. We generated three sets of problems, sets 4, 5 and 6 whose parameters were generated as in that of sets 1, 2 and 3 respectively. We use three reference points generated with weights $(0, 0, 0, 0.22, 0.78)$, $(0, 0, 0.33, 0.25, 0.42)$ and $(0.05, 0.1, 0.15, 0.2, 0.5)$ in Q_μ. Table 8 below summarizes the results of the GEND algorithm on these sets of problems. It is observed that for problems with the same number of items (n), the solution times increase as p increases from 3 to 5,

Table 5 Quality of solutions returned by GEND: Set 1

n	Max range	Scaled coverage error		Scaled coverage gap		Spread (% of) boxes with GEND soln.		
	Avg	Avg	Std Dev.	Avg	Std Dev.	Min	Avg	Max
250	24.10	0.22	0.15	0.14	0.10	14.29	49.39	100
300	37.20	0.29	0.12	0.14	0.07	20.83	48.70	100
350	34.40	0.32	0.10	0.14	0.07	16.67	35.64	52.38
400	28.30	0.26	0.13	0.12	0.06	22.73	52.06	100
450	19.60	0.19	0.13	0.10	0.09	18.52	62.78	100
500	37.89	0.20	0.11	0.10	0.06	27.59	52.54	100

Table 6 Quality of solutions returned by GEND: Set 2

n	Max range	Scaled coverage error		Scaled coverage gap		Spread (% of) boxes with GEND soln.		
	Avg	Avg	Std Dev.	Avg	Std Dev.	Min	Avg	Max
50	70.90	0.22	0.15	0.17	0.13	15.38	53.60	100
100	86.70	0.34	0.17	0.21	0.11	15.38	38.59	100
150	95.60	0.30	0.10	0.14	0.05	28.00	41.28	70
200	515.50	0.37	0.25	0.14	0.07	15.00	38.01	53.3
250	160.10	0.27	0.11	0.13	0.04	20.00	32.92	50

Table 7 Quality of solutions returned by GEND: Set 3

n	Max range	Scaled coverage error		Scaled coverage gap		Spread (% of) boxes with GEND soln.		
	Avg	Avg	Std Dev.	Avg	Std Dev.	Min	Avg	Max
50	1401.00	0.27	0.14	0.14	0.08	31.25	51.63	100
100	2060.20	0.30	0.16	0.17	0.11	24.00	48.44	100
150	1835.50	0.24	0.18	0.10	0.07	26.32	62.18	100
200	3482.20	0.28	0.08	0.15	0.06	20.83	35.39	71.43

as expected. We put a time limit of 18000 seconds (5 hours) and do not increase n if 3 out of 10 instances can not be solved within this time limit. Nevertheless, the results show that GEND algorithm still returns solutions in reasonable amount of time for most instances.

Note that even the single objective binary knapsack problem is NP-hard. Hence generating the equitably non-dominated frontier of the multi-objective version is computationally challenging (as seen in Table 3). The algorithms we propose aim to tackle this challenge. In settings where the DM is available to provide preference information, one can utilize the interactive approach and guide her to her most preferred solution. The sample results that we provide in Tables 1, 2 and 4 show that it is possible to detect the most preferred solution by asking a small number of comparison questions to the DM. If the DM is not available for providing preference information, one can generate a good subset of the equitably nondom-

Table 8 Results of the GEND Algorithm on Sets 4, 5, and 6

Set 4					Set 5					Set 6				
n	# of solutions		Solution time (s)			# of solutions		Solution time (s)			# of solutions		Solution time (s)	
	Avg	Max	Avg	Max	n	Avg	Max	Avg	Max	n	Avg	Max	Avg	Max
250	3.20	9	7.92	38.92	50	5.00	11	0.67	1.68	50	3.5	6	1.05	2.96
300	6.50	18	24.88	132.78	100	13.30	47	10.16	60.39	100	19.6	87	17.29	112.01
350	5.20	10	26.11	113.33	150	13.00	59	58.88	488.95	150	57.8	209	740.36	6223.67
400	13.50	65	169.60	926.39	200	11.60	36	48.47	261.32	200	53.5	223	2177.34	16122.66
450	14.78	52	4560.02	18000	250	9.00	39	51.28	420.57					

inated points. We observe that the GEND algorithm generates a good representative subset in only a fraction of the time it takes to generate the whole frontier.

6 Conclusion

We consider multi objective optimization problems where the decision maker is inequity averse, hence she is interested in finding equitably efficient (nondominated) points. We discuss two solution approaches that differ in terms of the timing of preference articulation.

The first approach is interactive and relies on input from the DM during the solution process. This algorithm is based on the assumption of a symmetric quasiconcavity utility function, which is a widely accepted form in the literature focusing on fair allocations (Sen and Foster 1997). The assumption of symmetry, however, makes the problem computationally very challenging as permutational calculations are involved. We extend the current results in the literature so as to provide an algorithm for such problems. To the best of our knowledge, this is the first study proposing such an interactive approach for equitable multi-objective programming problems. Equitable problems are highly relevant in operational research applications, especially in public sector, hence this work is expected to contribute both to the theory and practice of OR.

In the second approach, we aim to generate an evenly spread subset of the set of equitably non-dominated solutions to be presented to the DM for further consideration. We analyse the cumulative criteria space and fit a simple function close to the Pareto in the cumulative ordered criteria space. We then select reference points on the fitted function and generate regions around these points. Finally, we generate the equitably non-dominated points in these regions.

We illustrate the computational feasibility of the algorithms on equitable knapsack problems in which one funds projects that benefit multiple entities subject to a limited budget. Such problems are especially relevant in public service provision as entities may correspond to various population groups benefiting from the service. The experiments demonstrate that the proposed algorithms are computationally very efficient compared to the epsilon-constraint approach that finds the whole set of equitably non-dominated solutions.

The GEND algorithm can still be used to find a representative subset of the set of equitably nondominated points when the number of entities increases. However, the interactive approach involves permutational calculations, which could lead to exponentially growing number of constraints in the models when the number of objectives increases. One can handle this issue by grouping the entities so as to obtain a smaller number of objectives. It is well-known that even the classical MOP problems become computationally intractable when the number of objectives increases. To the best of our knowledge, there is still relatively limited work on finding exact solutions to multiobjective integer optimization problems with more than three objective functions (see Kirlik and Sayın 2014; Holzmann and Smith 2018 and the references therein for some recent work in this area). As in the classical multi-objective optimization problems, designing computationally efficient exact and interactive solution algorithms for equitable optimization problems with more than three objectives is a very promising future research area. Specifically, future research can address the technical challenges involved in dealing with the large number of t-subsets for larger number of objectives.

This study can also be extended by working on developing faster algorithms for generating non-dominated points in the defined regions for larger problem instances (in terms of p and

n) in reasonable time. We can also investigate the application of some evolutionary and meta-heuristic approaches in approximating the equitably non-dominated frontier and generating diverse solutions.

Appendices

A Proof of Theorem 6

Proof PART 1: We will show that if $z \in C\bar{D}_{Symm}(z^m, z^k)$ (it is equitably dominated by $C(\overrightarrow{z^m}; \overrightarrow{z^k}))$ then Eqs. 7 and 8 hold. Let z be cone dominated, i.e. $\exists\ \mu \geq 0 : z' = \overrightarrow{z^k} + \mu(\overrightarrow{z^k} - \overrightarrow{z^m})$ and $z \preceq_e z'$. Then by Theorem 5, $\sum_{i=1}^{t} \overrightarrow{z_i} \leq \sum_{j \in a} z'_j \quad \forall t, \forall a \in K^t$. That is:

$$\sum_{i=1}^{t} \overrightarrow{z_i} \leq \sum_{j \in a} (\overrightarrow{z_j^k} + \mu(\overrightarrow{z_j^k} - \overrightarrow{z_j^m})) \quad \forall t, \forall a \in K^t \tag{31}$$

Note that $K^t = K^t_{km} \cup K^t_{mk}$ and $K^t_{km} \cap K^t_{mk} = \emptyset \quad \forall t$. Then:

$$\sum_{i=1}^{t_1} \overrightarrow{z_i} \leq \sum_{j \in b} (\overrightarrow{z_j^k} + \mu(\overrightarrow{z_j^k} - \overrightarrow{z_j^m})) \quad \forall t_1 \in P, \forall b \in K^{t_1}_{mk} \tag{32}$$

$$\sum_{i=1}^{t_2} \overrightarrow{z_i} \leq \sum_{l \in c} (\overrightarrow{z_l^k} + \mu(\overrightarrow{z_l^k} - \overrightarrow{z_l^m})) \quad \forall t_2 \in P, \forall c \in K^{t_2}_{km} \tag{33}$$

Since $\sum_{j \in b}(\overrightarrow{z_j^k} - \overrightarrow{z_j^m}) \leq 0$ by definition, Eq. 32 can be rewritten as:

$$\mu \sum_{j \in b} (\overrightarrow{z_j^m} - \overrightarrow{z_j^k}) \leq \sum_{j \in b} \overrightarrow{z_j^k} - \sum_{i=1}^{t_1} \overrightarrow{z_i} \ \forall t_1 \in P, \forall b \in K^{t_1}_{mk} \tag{34}$$

Since $\mu \geq 0$ and $\sum_{j \in b}(\overrightarrow{z_j^m} - \overrightarrow{z_j^k}) \geq 0$, $\sum_{j \in b} \overrightarrow{z_j^k} - \sum_{i=1}^{t_1} \overrightarrow{z_i} \geq 0 \quad \forall t_1 \in P, \forall b \in K^{t_1}_{mk}$ should hold. That is, Eq. 7 holds.

Equation 33 can be rewritten as:

$$\sum_{i=1}^{t_2} \overrightarrow{z_i} - \sum_{l \in c} \overrightarrow{z_l^k} \leq \mu \sum_{l \in c} (\overrightarrow{z_l^k} - \overrightarrow{z_l^m}) \ \forall t_2 \in P, \forall c \in K^{t_2}_{km} \tag{35}$$

Since $\sum_{l \in c}(\overrightarrow{z_l^k} - \overrightarrow{z_l^m}) > 0$ and $\sum_{j \in b}(\overrightarrow{z_j^m} - \overrightarrow{z_j^k}) \geq 0$, from Eqs. 34 and 35 we have the following (by multiplying Eq. 34 by $\sum_{l \in c}(\overrightarrow{z_l^k} - \overrightarrow{z_l^m})$ and Eq. 35 by $\sum_{j \in b}(\overrightarrow{z_j^m} - \overrightarrow{z_j^k})$):

$(\sum_{i=1}^{t_2} \overrightarrow{z_i} - \sum_{l \in c} \overrightarrow{z_l^k})(\sum_{j \in b}(\overrightarrow{z_j^m} - \overrightarrow{z_j^k})) \leq \mu(\sum_{l \in c}(\overrightarrow{z_l^k} - \overrightarrow{z_l^m}))(\sum_{j \in b}(\overrightarrow{z_j^m} - \overrightarrow{z_j^k})) \leq (\sum_{j \in b} \overrightarrow{z_j^k} - \sum_{i=1}^{t_1} \overrightarrow{z_i})(\sum_{l \in c}(\overrightarrow{z_l^k} - \overrightarrow{z_l^m})) \forall t_1 \in P, \forall t_2 \in P, \forall b \in K^{t_1}_{mk}, \forall c \in K^{t_2}_{km}$

PART 2: Now suppose that Eqs. 7 and 8 hold. We will show that z is cone dominated, i.e. $\exists \mu \geq 0 : z \preceq_e (\overrightarrow{z^k} + \mu(\overrightarrow{z^k} - \overrightarrow{z^m}))$. Note that $K^{t_1}_{mk}$ consists of two subsets $K^{t_1}_{strict\_mk}$ and $K^{t_1}_{equal\_mk}$ as follows: $\sum_{j \in b} \overrightarrow{z_j^m} > \sum_{j \in b} \overrightarrow{z_j^k}$ for all $b \in K^{t_1}_{strict\_mk}$ and $\sum_{j \in b} \overrightarrow{z_j^m} = \sum_{j \in b} \overrightarrow{z_j^k}$ for all $b \in K^{t_1}_{equal\_mk}$.

Since for $b \in K^{t_1}_{strict\_mk}$ we have $\sum_{j\in b}(\overrightarrow{z^m_j} - \overrightarrow{z^k_j}) > 0$ and for $c \in K^{t_2}_{km}$ $\sum_{l\in c}(\overrightarrow{z^k_l} - \overrightarrow{z^m_l}) > 0$, Eq. 8 implies the following:

$$\frac{(\sum_{i=1}^{t_2}\overrightarrow{z_i} - \sum_{l\in c}\overrightarrow{z^k_l})}{\sum_{l\in c}(\overrightarrow{z^k_l} - \overrightarrow{z^m_l})} \leq \frac{(\sum_{j\in b}\overrightarrow{z^k_j} - \sum_{i=1}^{t_1}\overrightarrow{z_i})}{\sum_{j\in b}(\overrightarrow{z^m_j} - \overrightarrow{z^k_j})} \ \forall t_1 \in P, \forall t_2 \in P, b \in K^{t_1}_{strict\_mk}, c \in K^{t_2}_{km} \tag{36}$$

We will find a $\mu \geq 0$ that makes z equitably dominated by $\overrightarrow{z^k} + \mu(\overrightarrow{z^k} - \overrightarrow{z^m})$). One can define

$$\bar{\mu} = \min_{t_1\in P, b\in K^{t_1}_{strict\_mk}} \frac{(\sum_{j\in b}\overrightarrow{z^k_j} - \sum_{i=1}^{t_1}\overrightarrow{z_i})}{\sum_{j\in b}(\overrightarrow{z^m_j} - \overrightarrow{z^k_j})}. \tag{37}$$

Since $(\sum_{j\in b}\overrightarrow{z^k_j} - \sum_{i=1}^{t}\overrightarrow{z_i}) \geq 0$ (see Eq. 7) and $\sum_{j\in b}(\overrightarrow{z^m_j} - \overrightarrow{z^k_j}) > 0$ for all $b \in K_{strict\_mk}$, $\bar{\mu} \geq 0$ holds.

Note that the following holds for this $\bar{\mu}$:

$$\frac{(\sum_{i=1}^{t_2}\overrightarrow{z_i} - \sum_{l\in c}\overrightarrow{z^k_l})}{\sum_{l\in c}(\overrightarrow{z^k_l} - \overrightarrow{z^m_l})} \leq \bar{\mu} \leq \frac{(\sum_{j\in b}\overrightarrow{z^k_j} - \sum_{i=1}^{t_1}\overrightarrow{z_i})}{\sum_{j\in b}(\overrightarrow{z^m_j} - \overrightarrow{z^k_j})}$$
$$\forall t_1 \in P, \forall t_2 \in P, b \in K^{t_1}_{strict\_mk}, c \in K^{t_2}_{km} \tag{38}$$

$\sum_{i=1}^{t_2}\overrightarrow{z_i} \leq \sum_{l\in c}\overrightarrow{z^k_l} + \bar{\mu}\sum_{l\in c}(\overrightarrow{z^k_l} - \overrightarrow{z^m_l})\ \forall t_2 \in P, \forall c \in K^{t_2}_{km}$(From the left side of Eq. 38)

$\sum_{i=1}^{t_1}\overrightarrow{z_i} \leq \sum_{j\in b}\overrightarrow{z^k_j} + \bar{\mu}\sum_{j\in b}(\overrightarrow{z^k_j} - \overrightarrow{z^m_j})\ \forall t_1 \in P, \forall b \in K^{t_1}_{strict\_mk}$(From the right side of Eq. 38).

$\sum_{i=1}^{t}\overrightarrow{z_i} \leq \sum_{j\in b}\overrightarrow{z^k_j} + \bar{\mu}\sum_{j\in b}(\overrightarrow{z^k_j} - \overrightarrow{z^m_j}) = \sum_{j\in b}\overrightarrow{z^k_j}\ \forall t \in P, \forall b \in K^{t}_{equal\_mk}$ (From condition 7).

Note that $K^t = K^t_{strict\_mk} \cup K^t_{equal\_mk} \cup K^t_{km}$. Therefore the conditions of Theorem 5 are satisfied, making z equitably dominated by $\overrightarrow{z^k} + \bar{\mu}(\overrightarrow{z^k} - \overrightarrow{z^m})$. □

B Proof of Proposition 7

For $P = \{1, 2, \cdots, p\}$:

$$\begin{aligned} y_1 &= \overrightarrow{z}_1 \\ y_2 &= \overrightarrow{z}_1 + \overrightarrow{z}_2 \\ &\vdots \\ y_p &= \overrightarrow{z}_1 + \overrightarrow{z}_2 + \cdots + \overrightarrow{z}_p \end{aligned}$$

and $\overrightarrow{z}_1 \leq \overrightarrow{z}_2 \leq \cdots \leq \overrightarrow{z}_p$

We will prove Proposition 7 by induction. The base case is at $p = 2$ where $P = \{1, 2\}$. To show that Proposition 7 holds for the base case, we need to show that $y_2 \geq 2y_1$.

Since by definition $y_1 = \overrightarrow{z}_1$ and $y_2 = \overrightarrow{z}_1 + \overrightarrow{z}_2$ where $\overrightarrow{z}_2 \geq \overrightarrow{z}_1$, then $y_2 = \overrightarrow{z}_1 + \overrightarrow{z}_1 + \epsilon \geq 2\overrightarrow{z}_1 = 2y_1$ where $\epsilon \geq 0$. Hence $y_2 \geq 2y_1$.

Hypothesis 1 *Assume that Proposition 7 holds for $p = s$.*

To complete the proof, we need to show that Proposition 7 holds for $p = s + 1$. Due to Hypothesis 1, we just have to show that Proposition 7 holds for all $(j, s + 1) : 1 \leq j \leq s$. $jy_s \geq sy_j$ for all $1 \leq j \leq s$ (Due to Hypothesis 1)

$$\begin{aligned}
&\overrightarrow{z}_{s+1} \geq \overrightarrow{z}_j = y_j - y_{j-1} \\
&jy_s + j\overrightarrow{z}_{s+1} \geq sy_j + j\overrightarrow{z}_j \\
&jy_s + j\overrightarrow{z}_{s+1} \geq sy_j + j(y_j - y_{j-1}). \\
&j(y_s + \overrightarrow{z}_{s+1}) \geq (s+1)y_j + \underbrace{(j-1)y_j - jy_{j-1}}_{\geq 0,\ \text{Due to Hypothesis 1}}. \\
&\text{Hence } jy_{s+1} \geq (s+1)y_j.
\end{aligned}$$

Therefore, Proposition 7 holds for any p.

References

Alves, M. J., & Clímaco, J. (2007). A review of interactive methods for multiobjective integer and mixed-integer programming. *European Journal of Operational Research*, *180*(1), 99–115.

Antunes, C. H., Alves, M. J., & Clímaco, J. (2016). *Multiobjective linear and integer programming*. Berlin: Springer.

Baatar, D., & Wiecek, M. M. (2006). Advancing equitability in multiobjective programming. *Computers & Mathematics with Applications*, *52*(1–2), 225–234.

Bazgan, C., Hugot, H., & Vanderpooten, D. (2009). Solving efficiently the 0–1 multi-objective knapsack problem. *Computers & Operations Research*, *36*(1), 260–279.

Beamon, B. M., & Balcik, B. (2008). Performance measurement in humanitarian relief chains. *International Journal of Public Sector Management*, *21*(1), 4–25.

Ceyhan, G., Köksalan, M., & Lokman, B. (2019). Finding a representative nondominated set for multi-objective mixed integer programs. *European Journal of Operational Research*, *272*(1), 61–77.

Clímaco, J., Ferreira, C., & Captivo, M. E. (1997). Multicriteria integer programming: An overview of the different algorithmic approaches. In J. Clímaco (Ed.), *Multicriteria analysis* (pp. 248–258). Springer.

Deb, K. (2014). Multi-objective optimization. In E. Burke, & G. Kendall (Eds.), *Search methodologies* (pp. 403–449). Springer.

Ehrgott, M., & Gandibleux, X. (2000). A survey and annotated bibliography of multiobjective combinatorial optimization. *OR-Spektrum*, *22*(4), 425–460.

Figueira, J. R., Paquete, L., Simões, M., & Vanderpooten, D. (2013). Algorithmic improvements on dynamic programming for the bi-objective {0, 1} knapsack problem. *Computational Optimization and Applications*, *56*(1), 97–111.

Hazen, G. (1983). Preference convex unanimity in multiple criteria decision making. *Mathematics of operations research*, *8*(4), 505–516.

Holzmann, T., & Smith, J. (2018). Solving discrete multi-objective optimization problems using modified augmented weighted tchebychev scalarizations. *European Journal of Operational Research*, *271*(2), 436–449.

Karsu, Ö. (2013). Using holistic multicriteria assessments: The convex cones approach. In J. J. Cochran, L. A. Cox Jr., P. Keskinocak, J. P. Kharoufeh, & J. C. Smith (Eds.), *Wiley encyclopedia of operations research and management science* (pp. 1–14). Wiley.

Karsu, Ö., & Morton, A. (2014). Incorporating balance concerns in resource allocation decisions: A bi-criteria modelling approach. *Omega*, *44*, 70–82.

Karsu, Ö., & Morton, A. (2015). Inequity averse optimization in operational research. *European Journal of Operational Research*, *245*(2), 343–359.

Karsu, Ö., Morton, A., & Argyris, N. (2018). Capturing preferences for inequality aversion in decision support. *European Journal of Operational Research*, *264*(2), 686–706.

Kellerer, H., Pferschy, U., & Pisinger, D. (2003). *Knapsack problems. 2004*. Berlin: Springer.

Kirlik, G., & Sayın, S. (2014). A new algorithm for generating all nondominated solutions of multiobjective discrete optimization problems. *European Journal of Operational Research*, *232*(3), 479–488.

Klamroth, K., & Wiecek, M. M. (2000). Dynamic programming approaches to the multiple criteria knapsack problem. *Naval Research Logistics*, *47*(1), 57–76.

Köksalan, M. (2008). Multiobjective combinatorial optimization: Some approaches. *Journal of Multi-Criteria Decision Analysis*, *15*(3–4), 69–78.

Korhonen, P., Wallenius, J., & Zionts, S. (1984). Solving the discrete multiple criteria problem using convex cones. *Management Science*, *30*(11), 1336–1345.

Kostreva, M. M., & Ogryczak, W. (1999). Linear optimization with multiple equitable criteria. *RAIRO-Operations Research*, *33*(3), 275–297.

Kostreva, M. M., Ogryczak, W., & Wierzbicki, A. (2004). Equitable aggregations and multiple criteria analysis. *European Journal of Operational Research*, *158*(2), 362–377.

Laumanns, M., Thiele, L., & Zitzler, E. (2006). An efficient, adaptive parameter variation scheme for metaheuristics based on the epsilon-constraint method. *European Journal of Operational Research*, *169*(3), 932–942.

Lokman, B., & Köksalan, M. (2013). Finding all nondominated points of multi-objective integer programs. *Journal of Global Optimization*, *57*(2), 347–365.

Lokman, B., & Köksalan, M. (2014). Finding highly preferred points for multi-objective integer programs. *IIE Transactions*, *46*(11), 1181–1195.

Lokman, B., Köksalan, M., Korhonen, P. J., & Wallenius, J. (2016). An interactive algorithm to find the most preferred solution of multi-objective integer programs. *Annals of Operations Research*, *245*(1–2), 67–95.

Mansour, I. B., & Alaya, I. (2015). Indicator based ant colony optimization for multi-objective knapsack problem. *Procedia Computer Science*, *60*, 448–457.

Mansour, I. B., Alaya, I., & Tagina, M. (2019). A gradual weight-based ant colony approach for solving the multiobjective multidimensional knapsack problem. *Evolutionary Intelligence*, *12*(2), 253–272.

Mansour, I. B., Basseur, M., & Saubion, F. (2018). A multi-population algorithm for multi-objective knapsack problem. *Applied Soft Computing*, *70*, 814–825.

Mavrotas, G., & Diakoulaki, D. (1998). A branch and bound algorithm for mixed zero-one multiple objective linear programming. *European Journal of Operational Research*, *107*(3), 530–541.

Mavrotas, G., & Florios, K. (2013). An improved version of the augmented ε-constraint method (augmecon2) for finding the exact pareto set in multi-objective integer programming problems. *Applied Mathematics and Computation*, *219*(18), 9652–9669.

Ogryczak, W., & Śliwiński, T. (2003). On solving linear programs with the ordered weighted averaging objective. *European Journal of Operational Research*, *148*(1), 80–91.

Özlen, M., & Azizoğlu, M. (2009). Multi-objective integer programming: A general approach for generating all non-dominated solutions. *European Journal of Operational Research*, *199*(1), 25–35.

Ozlen, M., Burton, B. A., & MacRae, C. A. (2014). Multi-objective integer programming: An improved recursive algorithm. *Journal of Optimization Theory and Applications*, *160*(2), 470–482.

Sayın, S. (2000). Measuring the quality of discrete representations of efficient sets in multiple objective mathematical programming. *Mathematical Programming*, *87*(3), 543–560.

Sen, A., & Foster, J. (1997). *On economic inequality*. Oxford: Clarendon Press. (expanded Edition).

Shorrocks, A. (1983). Ranking income distributions. *Economica*, *50*, 3–17.

Silvano, M., & Paolo, T. (1990). *Knapsack problems: Algorithms and computer implementations*. Wiley.

Steuer, R. (1986). *Multiple criteria optimization: Theory, computation, and application*. New York: Willey.

Sylva, J., & Crema, A. (2004). A method for finding the set of non-dominated vectors for multiple objective integer linear programs. *European Journal of Operational Research*, *158*(1), 46–55.

Ulungu, E. L., & Teghem, J. (1994). Multi-objective combinatorial optimization problems: A survey. *Journal of Multi-Criteria Decision Analysis*, *3*(2), 83–104.

Visée, M., Teghem, J., Pirlot, M., & Ulungu, E. L. (1998). Two-phases method and branch and bound procedures to solve the bi-objective knapsack problem. *Journal of Global Optimization*, *12*(2), 139–155.

White, D. (1990). A bibliography on the applications of mathematical programming multiple-objective methods. *Journal of the Operational Research Society*, *41*, 669–691.

Zhang, W., & Reimann, M. (2014). A simple augmentedε-constraint method for multi-objective mathematical integer programming problems. *European Journal of Operational Research*, *234*(1), 15–24.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.