Available online at www.sciencedirect.com
ScienceDirect
Computers & Operations Research 35 (2008) 1660–1670
computers & operations research
www.elsevier.com/locate/cor

Comparing methods for multiattribute decision making with ordinal weights

Byeong Seok Ahn[a], Kyung Sam Park[b,*]

[a]College of Business Administration, Chung-Ang University, 221 Heukseok, Dongjak, Seoul 156-756, South Korea
[b]Business School, Korea University, Anam-5Ga, Seongbuk, Seoul 136-701, South Korea

Available online 9 November 2006

Abstract

This paper is concerned with procedures for ranking discrete alternatives when their values are evaluated precisely on multiple attributes and the attribute weights are known only to obey ordinal relations. There are a variety of situations where it is reasonable to use ranked weights, and there have been various techniques developed to deal with ranked weights and arrive at a choice or rank alternatives under consideration. The most common approach is to determine a set of approximate weights (e.g., rank-order centroid weights) from the ranked weights. This paper presents a different approach that does not develop approximate weights, but rather uses information about the intensity of dominance that is demonstrated by each alternative. Under this approach, several different, intuitively plausible, procedures are presented, so it may be interesting to investigate their performance. These new procedures are then compared against existing procedures using a simulation study. The simulation result shows that the approximate weighting approach yields more accurate results in terms of identifying the best alternatives and the overall rank of alternatives. Although the quality of the new procedures appears to be less accurate when using ranked weights, they provide a complete capability of dealing with arbitrary linear inequalities that signify possible imprecise information on weights, including mixtures of ordinal and bounded weights.

Keywords: Multiattribute decisions; Imprecise information; Approximate weights; Dominance measures

1. Introduction

During the past several decades, there has been much research to deal with imprecise information in multicriteria decision analysis. Imprecise information means that decision parameters, such as marginal values (i.e., values of decision alternatives on each attribute) and attribute weights, are known only to the extent that the true values lie within prescribed bounds, while other parameters are known only to satisfy certain ordinal relations. After initial investigations by Fishburn [1], several approaches have been developed that obviate the need for precise preference information. For example, Kmietowicz and Pearman [2] show a linear programming approach to establishing dominance relations between alternatives, when given imprecise weights and exact marginal values. Sage and White [3] allow imprecise knowledge of both marginal values and weights in determining pairwise dominance. Recently, Park [4] provides a review and characterization of linear programming approaches to establishing several different dominance relations.

* Corresponding author. Tel.: +82 2 3290 2611; fax: +82 2 922 1380.
E-mail addresses: bsahn@cau.ac.kr (B.S. Ahn), sampark@korea.ac.kr (K.S. Park).

0305-0548/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cor.2006.09.026

Imprecise information is also referred to as incomplete information or partial information, and assumes that the decision maker may not be willing or able to provide exact estimations of decision parameters. This assumption is realistic in situations of time pressure, lack of knowledge, intrinsically intangible or non-monetary nature of criteria, and/or decision maker's limited attention and information processing capability [5,6]. In particular, there are more specific reasons why the assumption of exact weights is unrealistic. For instance, Barron and Barrett [7] state that various methods for eliciting exact weights from the decision maker may suffer on several counts, because the weights are highly dependent on the elicitation method [8–11] and there is no agreement as to which method produces more accurate results since the "true" weights remain unknown. Also, there may be no single decision maker and the decision making group may not be able to reach agreement on a set of exact weights.

In this paper, we are interested primarily in situations where the set of imprecise weights consists of a ranking of the importance of attributes. Ranking is a necessary first step in most procedures for eliciting more precise weights [12]. Rank-ordering the importance of attributes may be easier than describing other imprecise weights such as bounded weights. In the situation where there is a group of decision makers, it may be realistic to expect agreement only on a ranking of weights.

Using ranked attribute weights, two distinct approaches to determining a best alternative and/or ranking alternatives are often found in the literature. One approach is to develop a set of approximate weights from the given ranked weights for use with a multiattribute value function. Several methods for selecting approximate weights, including equal weights (EW) and rank-order centroid (ROC) weights, have been proposed and evaluated [7,13–16]. A common conclusion of these studies is that ROC weights have an appealing theoretical rationale and appear to perform better than the other rank-based schemes in terms of choice accuracy. Another approach attempts to eliminate inferior alternatives. For example, Kirkwood and Sarin [17] present a method to assess pairwise dominance in order to rank alternatives. Linear programming and closely related methods for establishing dominance [2,4,18,19] belong to this category. However, referring to Kirkwood and Corner [20], the resulting dominance relations often fail to determine the best alternative or to rank-order alternatives sufficiently or fully.

One purpose of the present study is to present a different approach that does not develop approximate weights and that extends the aforementioned linear programming approach. Using the intensity of dominance for each alternative, which can be obtained from employing linear programming to establish dominance, this approach makes an aggregated value of each multiattribute alternative. This enables it to recommend a best alternative and to rank alternatives fully under ranked weights, while the conventional linear programming approach does not guarantee this. Several different, intuitively plausible, procedures are demonstrated to rank alternatives. The primary purpose of this paper is to compare these new procedures, which we refer to as *dominance measuring* methods, with approximate weighting methods in terms of choice and ranking accuracy. Our simulation study provides a clear pattern of results in the relative performance of the evaluated procedures.

This paper is organized as follows. In Section 2, we briefly review linear programming methods to establish dominance relations. Section 3 presents decision making procedures using the approaches of dominance measuring and approximate weighting. In Section 4, we evaluate the procedures using a simulation study and present the evaluation results. Finally, Section 5 concludes the paper.

2. Preliminaries and definitions

Suppose that there are m available alternatives, $j=1,\ldots,m$, each of which is evaluated on n attributes, $i=1,\ldots,n$. Let $\mathbf{v}_j=(v_{1j},\ldots,v_{nj})^{\mathrm{T}}$ be the column vector of marginal values for alternative j over n attributes. Define $\mathbf{w}=(w_1,\ldots,w_n)\geqslant 0$ as the row vector of attribute weights. The overall value of alternative j can then be given by the following function of multiattribute additive value:

$$\mathrm{MAV}_j=\mathbf{w}\mathbf{v}_j=\sum_{i=1}^{n} w_i v_{ij}, \tag{1}$$

It is assumed that $0\leqslant v_{ij}\leqslant 1$, for all i and j, and sum of all weights equals one.[1]

[1] For the underlying assumptions of the additive value function in detail, see [21,22].

We assume throughout that the marginal values of all alternatives are known exactly and the weights are known only to satisfy ordinal relations, unless otherwise detailed. Specifically, we assume that the importance of attributes is arranged in a descending order from the most important attribute to the least important attribute, and let

$$\mathbf{w} \in W = \left\{ \mathbf{w} | w_1 \geqslant w_2 \geqslant \cdots \geqslant w_n \geqslant 0;\ \sum_{i=1}^{n} w_i = 1 \right\}. \tag{2}$$

Then the problem is to determine the best alternative, meaning the alternative for which MAV_j is a maximum, or to rank alternatives.

The concept of dominance is well known and widely utilized to address the problem. Linear programming is used for conceptual clarification and for a compact representation of the underlying problems. As shown in [2,4,18,19], the following linear program can be used to test dominance between different alternatives k and j:

$$PD_{kj} = \min\{\mathbf{w}(\mathbf{v}_k - \mathbf{v}_j) | \mathbf{w} \in W\}. \tag{3}$$

If the optimal objective value $PD_{kj} \geqslant 0$, then alternative k dominates j, and otherwise it does not. Note that PD_{kj} is obtained when alternative k achieves the worst scenario, but alternative j achieves the best scenario in the given $\mathbf{w} \in W$. It follows that the model determines if an alternative is better than another for *all* the given imprecise weights. This concept of dominance is called *pairwise dominance*.

Besides this, other kinds of dominance can be employed. For example, consider the following linear programs:

$$UB_j = \max\{\mathbf{w}\mathbf{v}_j | \mathbf{w} \in W\}, \tag{4.1}$$

$$LB_j = \min\{\mathbf{w}\mathbf{v}_j | \mathbf{w} \in W\}. \tag{4.2}$$

We can then regard the two measures UB_j and LB_j, respectively, as an upper and lower bound of true overall value that alternative j can have within the given imprecise weights. It can hence be said that if $LB_k \geqslant UB_j$, then alternative k dominates j. Following [23], this concept of dominance is referred to as *absolute dominance* in the sense that alternative k absolutely dominates j when the lower bound of alternative k exceeds the upper bound of j.

A relationship exists between pairwise dominance and absolute dominance. The set of dominated alternatives in paired comparisons is a superset of the set of absolutely dominated alternatives. This is because $PD_{kj} \geqslant LB_k - UB_j$. In other words, the use of common weights in checking pairwise dominance leads to possibly more discrimination of alternatives than the use of different sets of weights for individual alternative in identifying absolute dominance. Thus, if alternative k absolutely dominates alternative j, then k pairwisely dominates j, but the reverse does not hold.

Additionally, given ranked weights, we recall the computational aspect of the linear programming problems in (3) and (4) and suggest a simpler way instead of employing algorithms like the Simplex method. By the well-known properties of a linear program, only the extreme points of the ranked weights need to be considered to effect the desired optimum and they are readily available. For example, if there are three attributes and their weights are given in the form of (2), then the extreme points are $(1, 0, 0)$, $(\frac{1}{2}, \frac{1}{2}, 0)$, and $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. To test dominance of a pair of alternatives (k, j), we can simply calculate the optimum, $PD_{kj} = \min\{(v_{1k} - v_{1j}), \frac{1}{2}[(v_{1k} - v_{1j}) + (v_{2k} - v_{2j})], \frac{1}{3}[(v_{1k} - v_{1j}) + (v_{2k} - v_{2j}) + (v_{3k} - v_{3j})]\}$. An analogous way can apply to identifying absolute dominance in (4).[2]

The final important point to be noted is that this dominance approach, while formally elegant and generally acceptable, frequently results in almost no prioritization of alternatives or too many non-dominated alternatives [20]. It often fails to determine the best alternative or to rank alternatives sufficiently or fully. Nonetheless, this approach provides useful information, the paired dominance value, *PD* defined in (3), and the absolute dominance values, *UB* and *LB* in (4). In the next section, we show that these dominance values can be utilized to further prioritize competitive alternatives, and hence recommend a best alternative and to rank alternatives fully.

[2] See [19,24], for more detailed discussions of finding extreme points from various imprecise weights.

3. Decision making procedures

3.1. Dominance measuring methods

First consider the paired dominance values, PD_{kj}, obtained when comparing ordered pairs of alternatives (k, j). We can then compute the *dominating* measure of alternative k by means of adding paired dominance values between this alternative and all other alternatives j. Similarly, we can obtain the *dominated* measure of alternative k which is the sum of the paired dominance values between each of the other alternatives and alternative k. Then the difference between the dominating and dominated measures can be regarded as a *net* dominance value that alternative k has over all other alternatives. We can thus use the net dominance value as a measure of strength of preference in the sense that a larger net value is better. Note that this idea is consistent with the sense of outranking dominance relationship [25,26]. This approach can be implemented by the following five steps:

Step 1: Obtain the paired dominance values, PD_{kj}, by solving problem (3) for $m(m-1)$ ordered pairs of alternatives.

Step 2: Compute the dominating measure ϕ_k^+ for each alternative:

$$\phi_k^+ = \sum_{\substack{j=1\\ j\neq k}}^{m} PD_{kj}, \quad k = 1, \ldots, m.$$

Step 3: Compute the dominated measure ϕ_k^- for each alternative:

$$\phi_k^- = \sum_{\substack{j=1\\ j\neq k}}^{m} PD_{jk}, \quad k = 1, \ldots, m.$$

Step 4: Calculate the net dominance value ϕ_k for each alternative:

$$\phi_k = \phi_k^+ - \phi_k^-, \quad k = 1, \ldots, m.$$

Step 5: Rank alternatives according to the ϕ_k values, where the best (rank 1) alternative is the one for which ϕ_k is a maximum and the worst (rank m) alternative is that for which ϕ_k is minimal.

Thus, without determining approximate weights a priori from the given imprecise weights, the proposed procedure can be used as a rule in selecting the best alternative or in ranking alternatives.[3]

It is important to note that the dominating value ϕ_k^+ in Step 2 is a larger-the-better measure, whereas the dominated value ϕ_k^- in Step 3 is a smaller-the-better measure. This implies that, besides the net dominance value $\phi_k = \phi_k^+ - \phi_k^-$ used in Step 4, we can use the ϕ_k^+ measure (or the ϕ_k^- measure) for ranking alternatives. We refer to the procedure using the ϕ_k^+ measure (i.e., Steps 1 and 2) as OUT I, and the other procedure using the net value ϕ_k (Steps 1–5) as OUT II. The ranks of the alternatives according to the two different procedures are not necessarily the same as they are based on different computations. Of interest may then be in comparing the relative efficacy of the two procedures OUT I and OUT II. The comparative study will be done in the next section, where some classical decision rules and approximate weighting methods described in the subsequent subsections are compared altogether.

3.2. Classical decision rules

In the course of decision analysis, different decision rules can be applied to help the decision maker make a decision. Three classical decision rules, which are considered to reflect the decision maker's propensity in a non-compensatory decision context, are modified to encompass an imprecise decision context [23,29]. These are (a) choosing an alternative with the largest possible overall value (i.e., maximax rule), (b) choosing an alternative for which the smallest possible

[3] For more detailed descriptions on this procedure, see [27,28]. In addition, the same ϕ_k value for different alternatives never happened in our simulation study (described in the next section), so we could rank-order alternatives fully in Step 5. However, if a tie occurs among the $\phi_k = \phi_k^+ - \phi_k^-$ values, then an appropriate tie-breaking rule is needed to make a complete rank-ordering of alternatives. For example, a higher rank might be given to the alternative whose ϕ_k^+ value is larger than the other alternative, when these two alternatives have the same ϕ_k value.

value is largest (maximin rule), and (c) selecting an alternative for which the maximum regret is smallest (minimax regret):

(a) maximax (OPTimistic): $\max_k\{UB_k, \forall k\}$,
(b) maximin (PESSimistic): $\max_k\{LB_k, \forall k\}$,
(c) minimax regret (REG): $\min_k\{MR_k, \forall k\}$.

Here, the UB_k and LB_k values are as defined in (4). The MR_k value to be minimized in (c) represents the maximum regret incurred when choosing alternative k, where

$$MR_k = \max\{\max\{\mathbf{w}(\mathbf{v}_j - \mathbf{v}_k) | \mathbf{w} \in W\}, \forall j \neq k\}.$$

In addition to the three classical decision rules (OPT, PESS, REG), we consider choosing an alternative for which the midpoint of the possible value interval is greatest:

(d) central values (CENT): $\max_k\{UB_k + LB_k, \forall k\}$.

3.3. Surrogate weighting methods

Given the ranked attribute weights as in (2), there have been several methods developed for determining approximate weights in decision making using the MAV function of (1). Stillwell et al. [16] present the following approximate weighting schemes, among others, that preserve the order of weights:

(a) rank sum (RS) weights: $w_i = \frac{n+1-i}{\sum_{j=1}^{n} j} = \frac{2(n+1-i)}{n(n+1)}, \quad i = 1, \ldots, n,$
(b) rank reciprocal (RR) weights: $w_i = \frac{1/i}{\sum_{j=1}^{n} 1/j}, \quad i = 1, \ldots, n.$

Barron and Barrett [7] suggest another weighting method,

(c) rank-order centroid (ROC) weights: $w_i = (1/n)\sum_{j=i}^{n} \frac{1}{j}, \quad i = 1, \ldots, n.$

While satisfying the weak ordinal relations in (2), another possible way of determining specific weights is

(d) equal weights (EW): $w_i = \frac{1}{n}, \quad i = 1, \ldots, n.$

Once a set of weights $\mathbf{w} = (w_i)$ is determined according to each of the weighting methods, the MAV function of (1) is then used to select the best alternative or rank the alternatives of interest.

There have been several studies on comparing the decision quality of these approximate weighting methods [7,13,16]. Particularly, Barron and Barrett [7] compare the four methods (RS, RR, ROC, and EW) using a simulation study and report that the ROC weights appear to perform better than the other approximate weights. They have also shown that the ROC weights are given by the arithmetic mean of the extreme points of ranked weights, so the use of these weights in evaluating alternatives results in selecting the alternative whose expected MAV is largest. This is an appealing theoretical rationale of the method that the other weighting schemes published do not have.

However, the dominance measuring approach we have presented above has never been compared with the existing approximate weighting approach in order to see which approach is relatively more efficacious. In the next section, we compare the relative efficacy of the two different approaches together with the classical decision rules described in Section 3.2.

4. Comparative analysis

4.1. Simulation methods and design

Using a simulation study, we demonstrate the performance of all the decision making methods described in the previous section, which are approximate weighting methods (ROC, RR, RS, EW), dominance measuring methods

(OUT I, OUT II, CENT), and three classical decision rules (OPT, PESS, REG). The simulation study is conducted as follows:

Step 1: Create a simulated decision problem: Each random number sequentially generated from independent uniform distribution on the (0, 1) range constitutes a $m \times n$ matrix of marginal values. By convention, the marginal values in each column are normalized so that the smallest value is zero and the largest is one (see footnote 1).

Step 2: Remove the dominated alternatives from the generated value matrix: without considering attribute weights, if $\mathbf{v}_k \geqslant \mathbf{v}_j$ and $\mathbf{v}_k \neq \mathbf{v}_j$, then alternative k dominates alternative j, and otherwise it does not. If dominated alternatives exist, then drop them and go to Step 1. This renders the value matrix negatively correlated, to some extent, and yields a more realistic decision problem since attributes for alternatives in the non-dominated set are usually negatively correlated [16].

Step 3: Compute the attribute weights: The five different sets of weights need to be generated: ROC, RR, RS, and EW weights are calculated according to the formulae shown in the previous section. Finally, we randomly generate a set of weights from the weight space in (2). These are denoted as TRUE weights and the decision made by the TRUE weights is called the TRUE method. To generate the TRUE weights, we first select $n-1$ independent random numbers from a uniform distribution on (0, 1), and then rank these numbers. Suppose the ranked numbers are $1 > r_{n-1} \geqslant \cdots \geqslant r_2 \geqslant r_1 > 0$. The differences between adjacently ranked numbers are then used as the desired weights: $w_n = 1 - r_{n-1}$, $w_{n-1} = r_{n-1} - r_{n-2}, \ldots, w_1 = r_1$. The obtained weights will sum to 1 and be uniformly distributed on the weight space [13,15].

Step 4: Rank the alternatives: We first calculate the MAV of each alternative, using the value matrix obtained in Step 1 and the weights (TRUE, ROC, RR, RS, and EW) derived in Step 3. Then the alternative with the highest MAV is positioned in the first rank, one with the second highest in the second rank, and so on. For the dominance measuring methods (OUT I, OUT II, CENT) and the classical decision rules (OPT, PESS, REG), we determine the corresponding ranks of the alternatives according to their procedures in the previous section.

Step 5: Compare the decisions resulted by TRUE method and by each of the decision making methods in terms of efficacy measures.[4]

As shown in our simulation procedure, the TRUE weights are randomly generated from a uniform distribution, while satisfying the feasible weight space W defined in (2). Since the weights are known only to obey the given attribute ranks, no point in W may be considered more likely than another and the density of the weights is uniform over W. Any other distribution implies information beyond a rank-ordering [7]. The choice of the distribution for generating the TRUE weights is consistent with knowing the attribute ranks and no other information. The quality of the decisions resulted from each method is then assessed by comparisons with the decision made by the TRUE method using the TRUE weights. This in turn provides a relative ordering of efficacy among the decision making methods.

We employ two measures of efficacy, hit ratio and rank-order correlation. The primary or basic goal of multiattribute decision making is to determine the best alternative or rank-order alternatives. Thus, it is basically needed to evaluate the performance of a decision making method in terms of choice and ranking accuracy. Hit ratio evaluates, throughout the simulation runs, how frequently a decision making method selects the same best alternative as TRUE. More specifically, the hit ratio for each method is defined by the proportion of all cases in which that method selects the same best alternative as the TRUE method does.[5] Rank-order correlation represents the similarity of the overall rank structures of alternatives constructed by TRUE method and by a decision making method. This can be calculated by using Kendall's τ [30], defined as follows:

$$\tau = 1 - \frac{2(\text{Number of pairwise preference violations})}{\text{Total number of pairs of preferences}}.$$

The value +1 in Kendall's τ, which ranges from -1 to $+1$, means perfect correspondence between the two rank-orders.

We design the simulation with four different levels of alternatives ($m = 3, 5, 7, 10$) and five different levels of attributes ($n = 3, 5, 7, 10, 15$). For each of the 20 design elements (alternatives $\times$ attributes), the process of generating and analyzing decision problems was repeated until 10 replications of 10,000 trials were obtained (see the Appendix for an illustration of the simulation for a specific value matrix).

[4] The simulation procedure was implemented using the Excel's programming language, Visual Basic for Applications (VBA) on an IBM compatible personal computer.

[5] Barron and Barrett [7] also employed the hit ratio to evaluate the choice accuracy of approximate weighting methods.

Table 1
Simulation results of the average hit ratios

Alternative	Attribute	Approximate weighting methods				Dominance measuring methods			Classical decision rules		
		ROC	RR	RS	EW	OUT I	OUT II	CENT	OPT	PESS	REG
3	3	0.891	0.881	0.873	0.721	0.884	0.883	0.841	0.737	0.812	0.832
	5	0.898	0.887	0.864	0.705	0.878	0.844	0.833	0.723	0.848	0.836
	7	0.889	0.866	0.857	0.686	0.848	0.830	0.815	0.692	0.836	0.826
	10	0.898	0.860	0.835	0.654	0.824	0.792	0.776	0.665	0.823	0.787
	15	0.875	0.835	0.829	0.648	0.762	0.720	0.719	0.620	0.759	0.723
5	3	0.846	0.830	0.816	0.654	0.816	0.833	0.782	0.590	0.680	0.759
	5	0.864	0.845	0.826	0.618	0.835	0.811	0.780	0.573	0.775	0.784
	7	0.855	0.830	0.805	0.593	0.809	0.786	0.753	0.558	0.773	0.756
	10	0.859	0.808	0.787	0.578	0.783	0.725	0.699	0.518	0.757	0.708
	15	0.876	0.799	0.773	0.565	0.730	0.680	0.652	0.486	0.719	0.658
7	3	0.815	0.793	0.792	0.622	0.785	0.815	0.695	0.576	0.605	0.701
	5	0.836	0.819	0.800	0.577	0.802	0.793	0.753	0.505	0.719	0.769
	7	0.825	0.799	0.773	0.551	0.782	0.767	0.717	0.483	0.730	0.714
	10	0.842	0.790	0.762	0.541	0.768	0.712	0.668	0.433	0.721	0.674
	15	0.831	0.781	0.757	0.525	0.713	0.652	0.620	0.411	0.704	0.618
10	3	0.760	0.735	0.734	0.578	0.729	0.737	0.719	0.477	0.718	0.711
	5	0.827	0.803	0.783	0.537	0.778	0.781	0.742	0.430	0.713	0.744
	7	0.806	0.779	0.748	0.507	0.727	0.723	0.678	0.402	0.696	0.692
	10	0.829	0.783	0.745	0.504	0.715	0.691	0.642	0.354	0.693	0.664
	15	0.849	0.774	0.740	0.489	0.702	0.645	0.611	0.350	0.681	0.608
Mean		0.849	0.815	0.795	0.593	0.784	0.761	0.725	0.529	0.738	0.728

4.2. Simulation results

For each decision making method, Table 1 exhibits the average hit ratio obtained for each of the 20 design elements (i.e., the hit ratio in each cell of Table 1 represents the average values of 10 replications of 10,000 trials).[6] The last row shows the mean of each of the columns. First looking at the approximate weighting methods, consistently, the ROC method is better than the RR method, which is better than the RS method, which is better than the EW method. Thus we have ROC > RR > RS > EW in terms of hit ratio. This result is identical to that of [7].

Next for the dominance measuring methods, OUT I method outperforms OUT II method, which outperforms CENT method: OUT I > OUT II > CENT in terms of hit ratio. The OUT I method maintains a correspondence of more than 80% with the TRUE method in cases with relatively fewer alternatives and attributes ($m = 3, 5$ and $n = 3, 5, 7$), but shows a decline as the numbers of both alternatives and attributes increase ($m = 10$ and $n = 10, 15$). Finally, for the classical decision rules (OPT, PESS, REG), there is no regular trend but all are worse than the other decision making methods except for EW and CENT.

On the whole, the ROC method is consistently superior to the OUT I method, and the RR and RS methods are also better than the OUT I method in almost all cases. The relative ordering of efficacy in terms of hit ratio can be summarized as ROC > RR > RS > OUT I > OUT II in this experiment. Throughout the different combinations of numbers of alternatives and attributes, ROC method appears to be the best performer in terms of hit ratio.

To make this more precise, we perform a statistical test. Table 2 summarizes the results of the paired t-tests to see if a significant difference occurs in the hit ratios produced between two different methods. For example, in comparison of the ROC and RR methods, we obtain a t-statistic of 7.75 and find that ROC significantly outperforms RR at a

[6] For example, 0.891 in the first cell is the mean of 10 values (obtained from 10 replications). Each of these 10 values is the mean of 10,000 values (from 10,000 trails). The range of these 10 values (the best value–the worst value) was 0.002. The ranges for the other cells were also very small and all less than 0.005. This implies that the analyses of the best and worst values do not provide a serious effect on the relative ordering of efficacy among the decision making methods under consideration. We stress, however, that after looking at the average values we also perform statistical tests to delve more precisely into the efficacy prioritizations (see Tables 2 and 4).

Table 2
Paired t-test results for the average hit ratios

Methods	RR	RS	EW	OUT I	OUT II	CENT	OPT	PESS	REG
ROC	7.75*	9.13*	21.09*	6.86*	6.16*	9.33*	14.60*	11.50*	9.41*
RR	–	8.52*	23.95*	5.48*	5.20*	9.62*	14.95*	8.79*	9.75*
RS		–	27.45*	2.32**	3.70*	8.56*	15.14*	6.13*	8.45*
EW			–	−25.08*	−18.41*	−17.03*	5.58*	−11.74*	−16.34*
OUT I				–	3.79*	10.76*	15.43*	4.44*	10.32*
OUT II					–	6.71*	13.93*	1.53***	5.57*
CENT						–	13.04*	−1.18	−1.55***
OPT							–	−10.69*	−12.90*
PESS								–	0.93

$^{*}p < 0.01$, $^{**}p < 0.05$, $^{***}p < 0.1$.

Table 3
Simulation results of the average rank-order correlations[a]

Alternative	Attribute	Approximate weighting methods				Dominance measuring methods		
		ROC	RR	RS	EW	OUT I	OUT II	CENT
3	3	0.864	0.853	0.843	0.611	0.843	0.852	0.809
	5	0.857	0.845	0.805	0.589	0.805	0.795	0.784
	7	0.848	0.821	0.790	0.558	0.789	0.772	0.748
	10	0.862	0.812	0.782	0.531	0.753	0.704	0.695
	15	0.870	0.775	0.773	0.525	0.701	0.650	0.619
5	3	0.861	0.847	0.836	0.637	0.829	0.846	0.795
	5	0.858	0.843	0.814	0.590	0.809	0.812	0.788
	7	0.855	0.832	0.801	0.570	0.784	0.771	0.747
	10	0.865	0.819	0.786	0.546	0.751	0.720	0.692
	15	0.875	0.793	0.775	0.529	0.724	0.671	0.630
7	3	0.857	0.842	0.835	0.647	0.825	0.846	0.770
	5	0.857	0.843	0.816	0.593	0.810	0.812	0.792
	7	0.856	0.832	0.797	0.562	0.775	0.774	0.743
	10	0.864	0.820	0.784	0.543	0.746	0.720	0.702
	15	0.877	0.797	0.778	0.534	0.695	0.668	0.640
10	3	0.836	0.822	0.813	0.613	0.823	0.826	0.817
	5	0.860	0.842	0.816	0.592	0.814	0.813	0.793
	7	0.858	0.833	0.798	0.566	0.787	0.771	0.749
	10	0.866	0.822	0.790	0.545	0.743	0.725	0.706
	15	0.880	0.802	0.782	0.542	0.716	0.694	0.645
Mean		0.861	0.825	0.801	0.571	0.776	0.762	0.733

[a]For the classical decision rules, the determination of a complete rank is excluded.

significance level of 0.01. The first row shows that the ROC method is significantly superior to all the other methods at the same significance level of 0.01. Similar interpretations apply to the other rows. In sum, we confirm that ROC > RR > RS > OUT I > OUT II, and OUT II outperforms the classical decision rules (OPT, PESS, REG) in this experiment.

Table 3 demonstrates the average rank-order correlations (or the average Kendall's τ values). The ROC method shows predominantly higher correlations than the other methods and, moreover, records the most robust performance. Throughout the simulated numbers of alternatives and attributes, the ROC method shows only a variation of about 2% in its rank-order correlations, which approximately range from 86% to 88%, whereas RR and RS methods record relatively rapid decreases of 7–8% as the numbers of alternatives and attributes increase. Table 3 also reports that the OUT I method, for which the rank-order correlations lie within 70% and 84%, is almost better than OUT II, which outperforms

Table 4
Paired t-test results for the average rank-order correlations

Methods	RR	RS	EW	OUT I	OUT II	CENT
ROC	6.00*	10.03*	30.45*	7.26*	6.24*	8.04*
RR	–	9.57*	48.08*	7.86*	6.12*	9.05*
RS		–	59.07*	4.07*	3.87*	6.51*
EW			–	−45.65*	−26.34*	−19.26*
OUT I				–	2.92*	8.58*
OUT II					–	7.91*

* $p < 0.01$.

the CENT method. The simulation results (Table 3) can thus be summarized as ROC > RR > RS > OUT I > OUT II in terms of rank-order correlation. Table 4, which is a summary of the paired t-test results for the rank-order correlations produced between two different methods, confirms this ordering of efficacy at a significance level of 0.01.

A couple of things with regard to the results can be mentioned. First, in 10 replications of 10,000 trials, no instance where the ROC method was surpassed by other methods in the two efficacy measures ever occurred. Second, as mentioned previously, the superiority of ROC weights under ranked attribute weights was previously determined by several researchers. Our simulation result is similar to those in favor of ROC weights.

A final point to be noted is that the dominance measuring approach (OUT I and OUT II) we have proposed is intuitively plausible and applicable to the problem of multiattribute decision making with imprecise weights, including ranked weights. Nevertheless, this approach has never been compared with the existing approximate weighting approach in order to see which approach is relatively more efficacious. Now from our comparative study, even though the result is negative regarding the performance of the proposed approach, we can rediscover the superiority of the approximate weighting approach (especially, the ROC method).

5. Concluding remarks

We have proposed dominance measuring methods (including OUT I and OUT II) for selecting the best alternative and/or ranking alternatives under ranked attribute weights in multicriteria decision analysis. Using a simulation study, the performance of these new methods is compared with existing methods—approximate weighting schemes (including ROC, RR, RS) and some classical decision rules. The simulation result shows that the approximate weighting approach outperforms the dominance measuring approach in terms of selecting the best alternative and ranking alternatives. In particular, the ROC method appears to be the best performer throughout the simulation. The result is negative with regard to the performance of the new proposed approach, but we support the superior performance of the ROC method.

Finally, we would remiss if we did not remark that there may exist advantages to the dominance measuring approach proposed. This is because, besides ranked attribute weights, other various types of imprecise weights possibly occur with this approach. For example, if the decision maker can assess a holistic judgment for a pair of alternatives, then we may have $\mathbf{w}(\mathbf{v}_k - \mathbf{v}_j) \geqslant 0$, where $\mathbf{v}_k$ and $\mathbf{v}_j$ can be the vectors of marginal values of hypothetical and/or actual alternatives. In other circumstances, the decision maker may say that the importance of attribute 1 is more than two times and less than three times that of attribute 2, so we have a ratio bound $2 \leqslant w_1/w_2 \leqslant 3$. Similar judgments are found in the study of analytic hierarchy process [31]. Interval judgments are also possible [32], for example, $0.2 \leqslant w_1 \leqslant 0.3$. Given a mixture of the above types of imprecise weights, it is not easy to obtain approximate weights (ROC, RR, RS weights) using the corresponding formulae. In contrast, the dominance measuring approach employs a linear program from which the necessary information is readily achievable so that the best alternative selection and rank of alternatives can be accomplished. However, given such general forms of imprecise weights, a further comparative study is needed on the efficacy of the approach.

Acknowledgments

We thank the editor and two anonymous referees for their helpful comments and suggestions. The work of the first author was completed while he was a visiting professor in the Graduate School of Management, University of California at Irvine (UCI) in the year 2003. The second author is also grateful for financial support from a Korea University Grant.

Appendix

Table 5 demonstrates an example of applying approximate weighting methods to a specific value matrix, where there are five alternatives and five attributes. The value matrix is constructed by repeated random number generations and simple dominance tests. The random TRUE weights are generated from the weight space that satisfies $w_1 \geqslant w_2 \geqslant w_3 \geqslant w_4 \geqslant w_5$, and four approximate weights in the bottom rows are derived using the corresponding formulae. The final ranks of alternatives are then determined according to their MAV, as shown at the right. In terms of choice accuracy (or hit ratio), all methods select the same best alternative (alternative 1) as TRUE. In terms of rank-order correlation, only RR weights rank the alternatives (1, 4, 2, 5, 3) identical to that of the TRUE method in this example.

Table 6 exhibits an example of applying dominance measuring methods (OUT I and OUT II) to the same value matrix as in Table 5. These methods do not need to generate any approximate weights but, as recorded at the left of Table 6, first need to obtain the paired dominance values using linear program (3) for 20 ordered pairs of the five alternatives (see Step 1 of the procedures). Next, as shown in the right, the dominating and dominated measures are calculated, and then the net dominance values are computed (see Steps 2–4 of the procedures). Using the dominating measures alone, the OUT I method produces the rank-order of the alternatives (1, 4, 2, 5, 3) shown in the bottom. As can be seen, the OUT II method produces the same order, which is based on the net dominance values recorded in the final column of the table. This result is the same as TRUE exhibited in Table 5, meaning that in terms of both choice and ranking accuracy, the OUT I and OUT II methods are accurate in this example.

Table 5
Example of the decision process of approximate weighting methods

Alternatives	Attributes ($w_1 \geqslant w_2 \geqslant w_3 \geqslant w_4 \geqslant w_5$)					MAV				
	1	2	3	4	5	TRUE	ROC	RR	RS	EW
1	1	0.86	0.59	0.51	0	0.789	0.817	0.768	0.748	0.592
2	0.58	0.05	1	0.46	0.76	0.565	0.507	0.528	0.519	0.570
3	0	0.62	0.31	1	0.06	0.246	0.300	0.295	0.365	0.398
4	0.82	1	0	0	1	0.655	0.672	0.666	0.607	0.564
5	0.34	0	0.54	0.17	0.9	0.346	0.291	0.425	0.304	0.390

Methods	Approximate weights					Rank of alternatives
TRUE	0.539	0.132	0.129	0.119	0.081	1, 4, 2, 5, 3
ROC	0.457	0.257	0.157	0.090	0.040	1, 4, 2, 3, 4
RR	0.438	0.219	0.146	0.109	0.088	1, 4, 2, 5, 3
RS	0.333	0.267	0.200	0.133	0.067	1, 4, 2, 3, 4
EW	0.200	0.200	0.200	0.200	0.200	1, 2, 4, 3, 5

Table 6
Example of the decision process of dominance measuring methods

Alternatives	Alternatives					OUT I		OUT II
	1	2	3	4	5	ϕ^+	ϕ^-	ϕ
1	–	0.022[a]	0.194	0.020	0.202	0.438	−2.660	3.098
2	−0.615	–	0.005	−0.595	0.145	−1.060	−0.886	−0.175
3	−1	−0.580	–	−0.820	−0.340	−2.740	−0.049	−2.692
4	−0.285	−0.068	−0.028	–	0.174	−0.206	−2.135	1.929
5	−0.760	−0.260	−0.220	−0.740	–	−1.998	0.181	−2.161
Rank-order						1, 4, 2, 5, 3		1, 4, 2, 5, 3

[a]This entry represents the paired dominance value between alternatives 1 and 2, PD_{12}.

References

[1] Fishburn PC. Analysis of decisions with incomplete knowledge of probabilities. Operations Research 1965;13:217–37.
[2] Kmietowicz ZW, Pearman AD. Decision theory, linear partial information and statistical dominance. Omega 1984;12:391–9.
[3] Sage AP, White CC. ARIADNE: a knowledge-based interactive system for planning and decision support. IEEE Transactions on Systems, Man and Cybernetics 1984;14:35–47.
[4] Park KS. Mathematical programming models for characterizing dominance and potential optimality when multicriteria alternative values and weights are simultaneously incomplete. IEEE Transactions on Systems, Man and Cybernetics—Part A 2004;34:601–14.
[5] Kahneman D, Slovic P, Tversky A. Judgment under uncertainty: heuristics and biases. Cambridge: Cambridge University Press; 1982.
[6] Weber M. Decision making with incomplete information. European Journal of Operational Research 1987;28:44–57.
[7] Barron FH, Barrett BE. Decision quality using ranked attribute weights. Management Science 1996;42:1515–23.
[8] Borcherding KS, von Winterfeldt D. The effect of varying value trees on multiattribute evaluations. Acta Psychologica 1988;68:153–70.
[9] Fischer GW. Range sensitivity of attribute weights in multiattribute value models. Organization Behavior and Human Decision Processes 1995;62:252–66.
[10] Jaccard J, Brinberg D, Ackerman LJ. Assessing attribute importance: a comparison of six methods. Journal of Consumer Research 1986;12: 463–8.
[11] Shoemaker PJH, Waid CD. An experimental comparison of different approaches to determining weights in additive utility models. Management Science 1982;28:182–96.
[12] Watson SR, Buede DM. Decision synthesis. Cambridge: Cambridge University Press; 1986.
[13] Butler J, Olson DL. Comparison of centroid and simulation approaches for selection sensitivity analysis. Journal of Multi-Criteria Decision Analysis 1999;8:146–61.
[14] Edwards W, Barron FH. Smarts and smarter: improved simple methods for multiattribute utility measurement. Organization Behavior and Human Decision Processes 1994;60:306–25.
[15] Jia J, Fischer GW, Dyer JS. Attribute weighting method and decision quality in the presence of response error: a simulation study. Journal of Behavioral Decision Making 1998;11:85–105.
[16] Stillwell WG, Seaver DA, Edwards W. A comparison of weight approximation techniques in multiattribute utility decision making. Organization Behavior and Human Decision Processes 1981;28:62–77.
[17] Kirkwood CW, Sarin RK. Ranking with partial information: a method and an application. Operations Research 1985;33:38–48.
[18] Malakooti B. Ranking and screening multiple criteria alternatives with partial information and use of ordinal and cardinal strength of preference. IEEE Transactions on Systems, Man and Cybernetics—Part A 2000;30:355–68.
[19] Marmol AM, Puerto J, Fernandez FR. The use of partial information on weights in multicriteria decision problems. Journal of Multi-Criteria Decision Analysis 1998;7:322–9.
[20] Kirkwood CW, Corner JL. The effectiveness of partial information about attribute weights for ranking alternatives in multiattribute decision making. Organization Behavior and Human Decision Processes 1993;54:456–76.
[21] Keeney RL, Raiffa H. Decisions with multiple objectives: preferences and value tradeoffs. New York: Wiley; 1976.
[22] von Winterfeldt D, Edwards W. Decision analysis and behavioral research. Cambridge: Cambridge University Press; 1986.
[23] Salo AA, Hamalainen RP. Preference ratios in multiattribute evaluation (PRIME)—elicitation and decision procedures under incomplete information. IEEE Transactions on Systems, Man and Cybernetics—Part A 2001;31:533–45.
[24] Solymosi T, Dombi J. A method for determining the weights of criteria: the centralized weights. European Journal of Operational Research 1986;26:35–41.
[25] Brans JP, Vincke PH. A preference ranking organization method. Management Science 1985;31:647–56.
[26] Roy B. The outranking approach and the foundation of ELECTRE methods. Theory and Decision 1991;31:49–73.
[27] Ahn BS. Extending Malakooti's model for ranking multi-criteria alternatives with preference strength and partial information. IEEE Transactions on Systems, Man and Cybernetics—Part A 2003;33:281–7.
[28] Ahn BS, Park KS, Han CH, Kim JK. Multi-attribute decision aid under incomplete information and hierarchical structure. European Journal of Operational Research 2000;125:431–9.
[29] Puerto J, Marmol AM, Monroy L, Fernandez FR. Decision criteria with partial information. International Transactions in Operational Research 2000;7:51–65.
[30] Winkler RL, Hays WL. Statistics: probability, inference, and decision. New York: Holt, Rinehart & Winston; 1985.
[31] Saaty TL. The analytic hierarchy process. New York: McGraw-Hill; 1980.
[32] Arbel A. Approximate articulation of preference and priority derivation. European Journal of Operational Research 1989;43:317–26.