Clustering and Group Selection of Multiple Criteria Alternatives with Application to Space-based Networks

Behnam Malakooti and Ziyong Yang
Electrical Engineering and Computer Science Department
Case Western Reserve University
Cleveland Ohio 44106-7070

Revised 4-29-2002
Abstract

In many real world problems the range of consequences of different alternatives are considerably different. Also, sometimes, selection of a group of alternatives (instead of only one best alternative) is necessary. Traditional decision making approaches treat the set of alternatives with the same method of analysis and selection. In this paper, we propose clustering alternatives into different groups so that different methods of analysis, selection, and implementation for each group can be applied. As an example, consider the selection of a group of functions (or tasks) to be processed by a group of processors. The set of tasks can be grouped according to their similar criteria, and hence each cluster of tasks to be processed by a processor.

The selection of the best alternative for each clustered group can be performed using existing methods; however, the process of selecting groups is different than the process of selecting alternatives within a group. We develop theories and procedures for clustering discrete multiple criteria alternatives. We also demonstrate how the set of alternatives is clustered into mutually exclusive groups based on: 1) similar features among alternatives; 2) ideal (or most representative) alternatives given by the Decision Maker; and 3) other preferential information of the Decision Maker.

The clustering of multiple criteria alternatives also has the following advantages: 1) it decreases the set of alternatives to be considered by the Decision Maker (for example different decision makers are assigned to different groups of alternatives); 2) it decreases the number of criteria; 3) it may provide a different approach for analyzing multiple decision makers problems. Each decision maker may cluster alternatives differently, and hence clustering of alternatives may provide a basis for negotiation.

The developed approach is applicable for solving a class of telecommunication networks problems where a set of objects (such as routers, processors, or intelligent autonomous vehicles) are to be clustered into similar groups. Objects are clustered based on several criteria and the decision maker’s preferences.

Key Words: Interactive Multiple Objective Optimization, Multiple Criteria Decision Making, Grouping, Screening, Ranking, Multiple Criteria Alternatives, Neural Networks, Group Technology, Aero-Space Applications
1. Introduction

In this paper, we present the problem of multiple criteria clustering of discrete alternatives where the decision is to choose one or more of the alternatives. Each set of clustered alternatives is unique and different than the other sets. For these problems, it is imperative that the decision maker be able to compare and contrast alternatives within each group and choose the best alternative within that group because each group may serve a different purpose for decision-making. While the general approach for generating alternatives may be the same; however, the selection and implementation process for each group is different and requires the use of identified clustered alternatives. As an example, consider the problem of selection of project contractors for repair of bridges in a given county. The decision maker (the county) can give all the bridges to one or more contractors. Since, the conditions and requirements of each project (each bridge) is different than the other ones, all the contractors are potential candidates to undertake the projects. Now, one may cluster these projects into different homogenous groups that have similarities in terms of complexity of the project, the cost of the project, and the expertise of the contractor to perform the project. The number of clusters and the number of projects to be assigned to a contractor may vary. Such clustering will simplify the process of selection of contractors to do different projects. In this paper, we show how multiple criteria alternatives can be clustered into homogenous groups such that appropriate solution for each group can be identified.

To complete clustering of alternatives, the program may generate a representative for each cluster. Or alternatively, we ask the Decision Maker to provide such representatives (ideal alternatives) for each cluster. Then, in the case that this method cannot determine the clusters for the alternatives, the program asks the Decision Maker to suggest their possible grouping. However, the method is designed to ask the minimum number of such questions. The preferential clustering of alternatives could be: completely known; completely unknown; or partially known. We also demonstrate the minimum partial preferential information needed to form exclusive groups of clustered alternatives. An example is given to explain how the developed theories and procedures are used. Results of computational experiments on several problems are reported.

In traditional multiple criteria decision making (MCDM) or multiple objective optimization (MOO) methods, the whole set of alternatives is considered for the selection of the best alternative by one decision maker. Whether such a selection is made based on certain principles, accepted practices, or heuristics; the method of selection and the decision maker remain the same in treating the whole set of alternatives. In many real world problems; however, all alternatives
cannot be treated the same way, using the same decision maker and the same method of selection. Furthermore, a number of different best alternatives (instead of one best alternative) should be selected. In these cases, the clustering of alternatives into different groups becomes necessary. As an example, consider the site selection of different nuclear power plants, and suppose that there are fifty possible sites that can be used for this purpose. Consider that depending on the type of nuclear power plant and the selected sites, the total number of power plants to be implemented may vary from one to five. The clustering of multiple criteria alternatives into different groups can simplify the process of analysis, selection, and implementation of the selected nuclear power plants. For example, one may cluster alternatives into five clustered groups and for each group identify the best alternative; and hence the five best alternatives are the ones to be implemented. Such a solution would be different if four, three, two, or one group of alternatives were to be selected. In this paper, we discuss how clusters can be selected and how different alternatives could be clustered into independent clustered groups.

Cluster analysis is concerned with the grouping of alternatives into homogeneous clusters (groups) based on certain features. The clustering of multiple criteria alternatives can also bring the following benefits: 1) It decreases the set of alternatives - since the Decision Maker may be interested in only those alternatives with similar kinds of features and discard other alternatives. 2) It decreases the number of criteria - when evaluating alternatives of one group, the Decision Maker does not need to consider all criteria since one or more criteria values of the alternatives in the same group are equal or very close. 3) It provides a basis for more in-depth evaluation of alternatives - once one set of clustered alternatives are selected then this set can be explored in more depth for analysis, selection, and implementation purposes. 4) It may provide a basis for analyzing multiple criteria problems. Each decision maker may cluster alternatives differently, and hence, clustering of alternatives may provide a basis for negotiation. 5) In case of selection of a group of alternatives, each decision maker can be in charge of one clustered group; hence the designated decision maker selects the best solution from each clustered group. For more background on clustering and Multiple Criteria Decision Making based on clustering see Miettinen and Salminen (1999) or Malakooti and Raman (2000).

While the concept of clustering alternatives for MCDM and MOO problems are new, this field has been explored by a variety of researchers. The main difference between multiple criteria decision making and other methods is that criteria or objectives are to be optimized (e.g. maximized or minimized). For other problems such assumptions are not necessary. Also, criteria present key decision making factors, for selection of the most preferred alternatives. In clustering methods, the key concept is clusters, and groups, and hence selection of the preferred alternatives is not central. Three well-known clustering strategies are hierarchical clustering (Pandit, Srivastava, and Sharma 2001; Dias, Costa, and Climaco 1995), graph-theoretic methods
(Matula 1977) and conceptual clustering (Michalski and Stepp 1982, 1983). Clustering methods based on fuzzy set theory and neural networks have also been proposed recently. Hierarchical clustering often presents its results in the form of a dendrogram, which is a sequence of increasingly refined partitions with corresponding values of a dissimilarity measurement. Conceptual clustering treats a cluster not as simply a subset of a given set of objects (alternatives), but as a specific concept. Each cluster represents a certain generalized description of conjunctive concepts, involving attributes of objects, and thus is supposed to have a clear interpretation (Genkin and Muchnik 1993).

Clustering has been applied in many areas, such as biology, data recognition, medicine, pattern recognition, production flow analysis, task selection, control engineering, automated systems, and expert systems; see Spath (1985) among others for reviews. Each of the above application fields can also use MCDM and MOO analysis; and hence use the developed concepts and procedures in this paper. Conventional clustering approaches involve two main factors: 1) distance (or dissimilarity) measurement and 2) cluster centers. In this paper, we discuss how clustering methods can be applied and used for MCDM and MOO problems and demonstrate the type of information needed to cluster alternatives. The final selection of the best alternative for each given cluster is not the subject of this paper; there are many MCDM and MOO methods that can be used to solve the clustered problem. Many approaches have been proposed for solving MCDM and MOO problems; see Keeney and Raiffa (1976), Zeleny (1982), Steuer (1986), and Von Winterfeldt and Edwards (1986). For a continuous set of alternatives, Steuer and Choo (1983) among others developed a gradient-based interactive approach requiring marginal rate of substitution questions to assess the gradient in each iteration; also see Sadagopan and Ravindran (1986) for the same problem. Other MCDM methods use numerical ratings to assess the weights of an additive multiple attribute value (utility) function and rank discrete sets of alternatives. Another approach for solving MCDM and MOO problems is the consecutive and systematic screening of efficient alternatives; see Steuer (1986) for references. For some other related MCDM and MOO screening and interactive methods see (Kirkwood and Sarin 1985; Malakooti and Subramanian 99; Lowe et al. 84), Malakooti (91,2000), Malakooti and Zhou 94).

This paper is organized as follows. Section 2 gives the definitions and notations for general discrete MCDM problems. Section 3 gives definitions and notations for the clustering problems. Section 4 develops definitions and theories for the clustering of the multiple criteria alternatives, and presents the procedure to cluster the alternatives. In this section, as an example of MCDM methods, additive multi-attribute value (utility) functions for finding the most satisfactory alternative for each clustered group are presented. Section 5 contains an example problem and also describes experimental results of several problems. Section 6 concludes the paper.
2. Definitions and Notation for Discrete MCDM Problems

In this section, we first give the definitions and notations for general Discrete Multiple Criteria Decision Making (MCDM) problems. Then we develop the concept of cluster-wise preference (utility) functions.

Let $X$ denote a discrete set of alternatives. Each alternative $x$ in $X$ is a $m$-tuple vector, $x = (x_1, x_2, ..., x_m)^T$, where $x_i$ is the $i$th decision variable. The discrete set, $X$, can be presented as $X = \{x^{(1)}, x^{(2)}, ..., x^{(p)}\}$ where $p$ is the number of alternatives. Over the set $X$, the objective function can be denoted as $f_1(x), f_2(x), ..., f_n(x)$, and are to be maximized (or minimized). The function $f_j(x)$ is called the $j$th objective (criterion) function. We assume that each objective function will be maximized from the view of a Decision Maker. Let the objective vector be $f(x) = [f_1(x), f_2(x), ..., f_n(x)]^T$.

Subject to: $x \in X$

In vector format, the MCDM problem can be re-written as:

$$\text{Maximize } \{f(x) \mid x \in X\}$$  \hspace{1cm} (2-1)

The objective functions usually conflict with each other: an increase in one objective may be possible only if another objective is reduced. Because of such conflicts, one commonly used concept in MCDM area is efficiency. An alternative $x^* \in X$ is efficient if and only if there does not exist an $x \in X$ such that $f(x) \geq f(x^*)$ and $f(x) \neq f(x^*)$. In the optimization process of problem (2-1), the Decision Maker seeks a good compromise alternative, which must be an efficient alternative. The Decision Maker's choice is subject to his/her preference over the set $X$. From the view of the Decision Maker's preference (utility) function, which may be known explicitly or implicitly, the MCDM problems can then be expressed as follows:

$$\text{Maximize } \{U_r(f) \mid f = f(x) \text{ and } x \in X_r\}$$  \hspace{1cm} (2-2)

where $U_r(f)$ is the Decision Maker's preference function for each subset $X_r$, $r = 1, 2, ..., R$, which is an increasing function of each variable $f_i(x)$, $i = 1, 2, ..., n$. Theoretically, the existence of the preference functions has been proved under certain requirements. From problem (2-2), it is clear that there are three factors, $U_r(f)$, $f(x)$, and $x$, involving a decision making process.

In this paper, we assume that the set of alternatives in the objective space is given. We define and use $f(x) = x$, hence $U_r(f) = U_r(x)$.

In solving discrete MCDM problems, an important step is to assess the Decision Maker's preference function $U_r(x)$. The utility function $U_r(x)$ can be of many forms; such as linear (additive), concave, convex, or a Chebychev function. The additive form is well-known and
widely used for utility functions. A nonlinear utility function is more flexible and versatile for representing the Decision Maker's preference. But in practical applications, identifying a nonlinear utility function structure and assessing its parameters is a difficult and time-consuming job for the Decision Maker. In our paper, we propose the Decision Maker to respond to a paired comparison of alternatives; hence evaluation by the Decision Maker would be the same for both additive and nonlinear functions. We chose additive functions because they are easier to be used and assessed.

In this paper, we assume that there are many alternatives that should be grouped into a small number of clusters, each having certain common features. We also assume that there exists a utility function $U_r$ for each subset $X_r$, $r = 1, 2, ..., R$; these utility functions are different from each other to signify that each subset poses a different problem and a different preferential function. In this paper, we assume each $U_r$ is an additive utility function because it may require much less number of questions to assess an additive (linear) utility function than a nonlinear one. However, methods using nonlinear utility functions (see References in the introduction) can be readily used for each cluster.

3. Definitions and Notation for Clustering Problems

Clustering seeks to group objectives into different subsets according to some kind of similarity among objectives. As we mentioned in Section 1, three well-known clustering strategies are hierarchical clustering, graph-theoretic methods, and conceptual clustering. In this paper, we adopt the conceptual clustering method, which treats a cluster not as simply a subset of a given set of objectives, but as a specific concept or feature. Definitions and notations for the clustering problems follow.

The similarity measurement and the number of clusters are two factors that should be decided in a clustering process. The decision maker specifies the number of clusters. However, the approach can be easily modified to help with the selection of the best number of clusters. Given a finite set $A$, the similarity measurement is defined as a real-valued function $d(y, z)$, $y, z \in A$, which satisfies the following conditions:

1) $d(y, z) \neq 0; d(y, y) = 0$;
2) $d(y, z) = d(z, y)$;
3) $d(w, z) \leq d(w, y) + d(y, z); w, y, z \in A$.

Any distance measurement that satisfies these conditions can be used as the similarity measurement in a clustering procedure.
After the similarity measurement \( d(y, z) \), \( y, z \in A \) is defined, and assuming that the number of clusters is known, a clustering process is to find the cluster memberships of all alternatives so as to minimize the following function:

\[
\min F(d) = \sum_{r=1}^{R} \sum_{s \in S_r} d(x_s, c_r)
\]

(3-1)

\( c_r \) is the cluster center of cluster \( r \), \( r = 1, 2, \ldots, R \). The term \( d(x_s, c_r) \) is the similarity measurement among object \( x_s \) and cluster center \( c_r \). \( S_r \) is the index set of all alternatives in cluster \( r \). Generally speaking, the procedure is to find cluster center \( c_r \), \( r = 1, 2, \ldots, R \), and then cluster all alternatives into \( R \) groups according to the following judgment:

An objective \( x \) belongs to cluster \( r \) if and only if

\[
d(x, c_r) < d(x, c_t) \text{ for } t = 1, 2, \ldots, R \text{ and } t \neq r.
\]

(3-2)

This strategy (3-2) will guarantee that the objective function \( F(d) \) is minimized.

In the next section, we will discuss the clustering of MCDM alternatives. We will utilize the Decision Maker’s ideal alternatives (goals) as the cluster centers and define the similarity measurement as the Euclidean distance.

4. Theories and Procedures for Clustering of Multiple Criteria Alternatives

In this section, we discuss the first stage of the two-stage approach: clustering of MCDM alternatives. We develop theories for clustering when the importance coefficient \( k \in \Lambda \) is unknown and for when it is partially known. We describe how to determine whether an alternative is classified or unclassified. We develop a method to assess the importance coefficient \( k \) according to partial information obtained from the Decision Maker. Finally, we give the computation algorithm to cluster alternatives based on partial information of \( k \).

We define the similarity measurement \( d(x, c) \) as the generalized Euclidean distance between m-tuple vectors \( x \) and \( c \). Thus,

\[
d(x, c) = \sqrt{k_1(x_1 - c_1)^2 + k_2(x_2 - c_2)^2 + \ldots + k_m(x_m - c_m)^2}
\]

where \( k_1, k_2, \ldots, k_m \) are the coefficients which are not known at the beginning of the clustering process. We denote \( k = \{k_i, i = 1, \ldots, m\} \). \( k \in \Lambda \), where \( \Lambda = \{ \sum_{i=1}^{m} k_i = 1, \ 0 \leq k_i \leq 1, \ i = 1, 2, \ldots, m \} \). In general, the distance between any two points \( x, y \) are as follows,

\[
d(x, y) = \sqrt{k_1(x_1 - y_1)^2 + k_2(x_2 - y_2)^2 + \ldots + k_m(x_m - y_m)^2}
\]

We can generalize the weighted Euclidian distance to \( t \) metric by

\[
d_t(x, y) = \sqrt{k_1(x_1 - y_1)^t + k_2(x_2 - y_2)^t + \ldots + k_m(x_m - y_m)^t}
\]
For this paper, we use equation $d(x, y)$ for simplicity of calculations, but one can also use for any given $t$ value $d_t(x, y)$.

When we use the generalized Euclidean distance, the coefficients $k = \{k_i, i = 1, \ldots, m\}$ can be used to represent indirectly the importance index of each attribute of the alternatives. A point has $m$ attributes (or criteria or objectives). As an example, examine two possible objectives for the power plant problem stated in the Introduction. Assume the objectives are to minimize $f_1$ and $f_2$ where $f_1 =$ Total Cost to build the power plant (millions) and $f_2 =$ Total time to complete the construction of the plant (months). For simplicity, we define variables and objectives to be the same (i.e. $f_1=x_1$, $f_2=x_2$). In this case $x_1$ corresponds with cost and $x_2$ corresponds with the time.

Assuming that cluster centers or cluster goals $c_1, c_2, \ldots, c_R$ are given by the Decision Maker as his/her ideal alternatives, each ideal alternative has some features the Decision Maker is looking for. For simplicity, in the following discussion, we assume that $R = 2$. But the developed approach is applicable for $R > 2$. Since the cluster goals are given, the distance measurement is a function of $k$, which means that the cluster memberships of some alternatives depend on the $k$. See Fig. 1 for an example: $d(x, c^1) = 0.5(k_2)^{1/2}$, and $d(x, c^2) = (k_1)^{1/2}$. If $k_2 < 4k_1$, then $d(x, c^1) < d(x, c^2)$, so $x$ belongs to cluster 1. If $k_2 > 4k_1$, then $d(x, c^1) > d(x, c^2)$, so $x$ belongs to cluster 2.

![Figure 1. An example to show that the distance is a function of $k$](image)

An alternative $x$ definitely belongs to cluster $c_1$ if and only if $d(x, c_1) < d(x, c_2)$ for all $k \in \Lambda$, and an alternative $x$ definitely belongs to cluster $c_2$ if and only if $d(x, c_2) < d(x, c_1)$ for all $k \in \Lambda$.

An alternative $x$ is a classified alternative if it definitely belongs to a cluster.

An alternative $x$ is an unclassified alternative if it does not definitely belong to a cluster.
See Figure 2 for an example showing the distance of a single point from two cluster centers that are being specified by the decision maker:

![Figure 2. Example of Cluster Memberships dependent on the importance coefficient k](image)

\[
d(x, c_1) = \sqrt{k_1 + 9k_2}
\]
\[
d(x, c_2) = \sqrt{9k_1 + k_2}
\]

From this example it is clear how the values of the importance coefficient \(k\) can decide the classification of an alternative to the appropriate cluster. If \(k_2 > k_1\), then \(x\) belongs to Cluster 2, and vice versa if \(k_1 > k_2\), then \(x\) belongs to Cluster 1.

Three possibilities about knowing \(k \in \Lambda\) are as follows:

1) \(k\) is completely known
2) \(k\) is completely unknown
3) \(k\) is partially known

4.1. \(k \in \Lambda\) is Completely Known
We assume that cluster goals or centers $c_1$ and $c_2$ are given by the Decision Maker. Clearly, if the scaling constant vector $k \in \Lambda$ is known, all alternatives can be clustered into two subsets with the computation of distance $d(x, c_2)$ and $d(x, c_1)$ for all $x \in X$.

### 4.2. \( k \in \Lambda \) is Completely Unknown

Please note that proofs of all remarks and propositions appear in Appendix 1.

**Remark 4-1** When $k \in \Lambda$ is unknown, an alternative $x$ definitely belongs to cluster $c_1$ if and only if
\[
\min [d(x, c_2) - d(x, c_1)] > 0, \text{ where } k \in \Lambda \text{ and } x \in X.
\]

**Remark 4-2** When $k \in \Lambda$ is unknown, an alternative $x$ definitely belongs to cluster $c_2$ if and only if
\[
\min [d(x, c_1) - d(x, c_2)] > 0, \text{ where } k \in \Lambda \text{ and } x \in X.
\]

**Proposition 4-1** When $k$ is unknown and $k \in \Lambda$, an alternative $x = [x_1, x_2, \ldots, x_m]$ definitely belongs to cluster $c_1 = [c_{11}, c_{12}, \ldots, c_{1m}]$ if and only if
\[
(x_1 - c_{11})^2 < (x_1 - c_{21})^2, (x_2 - c_{12})^2 < (x_2 - c_{22})^2, \ldots, (x_m - c_{1m})^2 < (x_m - c_{2m})^2
\]
and an alternative $x$ definitely belongs to cluster $c_2 = [c_{21}, c_{22}, \ldots, c_{2m}]$ if and only if
\[
(x_1 - c_{11})^2 > (x_1 - c_{21})^2, (x_2 - c_{12})^2 > (x_2 - c_{22})^2, \ldots, (x_m - c_{1m})^2 > (x_m - c_{2m})^2;
\]
otherwise, an alternative $x$ is an unclassified alternative.

**Proposition 4-2** When $k$ is unknown and $k \in \Lambda$, an alternative $x$ definitely belongs to cluster $c_1$ if and only if $Z < 0$ in Problem 1:

(Problem 1) \[
Z = \max \sum_{i=1}^{m} (x_i - c_{1i})^2 k_i + v
\]
subject to
\[
\sum_{i=1}^{m} (x_i - c_{2i})^2 k_i + v \leq 0
\]
\[
k \in \Lambda
\]
\[
k_i \geq \varepsilon, \quad i = 1, 2, \ldots, m
\]
In general, $v$ is a variable unrestricted in sign, but note that here $v$ is negative. The reason is that $k$ are positives, then $\Sigma(x_i-c_2)^2k_i$ is positive, hence $v$ is negative. $\epsilon$ is a small positive number ($\epsilon = 0.001$), $x = [x_1, x_2, ..., x_m]$, $c_1 = [c_{11}, c_{12}, ..., c_{1m}]$, and $c_2 = [c_{21}, c_{22}, ..., c_{2m}]$. See Appendix 2 for an explanation and example of the variable $v$ for the clustering linear program.

By using either Proposition 4-1 or Proposition 4-2, we may be able to cluster some alternatives even without any information about $k$. We also can find the unclassified alternatives whose cluster memberships depend on $k$.

We call an alternative $x$ a convex combination of other two alternatives $x_1$ and $x_2$ if $x = \mu x_1 + (1 - \mu) x_2$, where $\mu \in [0, 1]$. Convex combinations of more than two alternatives are defined in optimization literature: see Malakooti (1988) for details.

**Proposition 4-3** When two alternatives $x_1$ and $x_2$ definitely belong to cluster $c_1$, alternatives that are convex combinations of $x_1$ and $x_2$ also definitely belong to $c_1$.

### 4.3. $k \in \Lambda$ is Partially Known

When we do not know any information about $k$, some alternatives (classified alternatives) can still be clustered by using Proposition 4-1 and Proposition 4-2. Hence the alternative set $X$ can be divided into the classified alternative set $X_c$, and the unclassified alternative set $X_u$. We then have $X = X_c \cup X_u$ and $X_c \cap X_u = \emptyset$. In this section, we investigate how to group the unclassified alternatives $X_u$ according to the partial information obtained from the Decision Maker, especially the information from the paired comparison of alternatives.

There exists a closed and nonempty set of constraints on $k$ that is denoted by $\Lambda^*$, where $\Lambda^*$ is a subset of $\Lambda$, such that $k \in \Lambda^*$. Constraints in $\Lambda^*$ are constructed based on partial information obtained from the Decision Maker.

It is possible to consider three types of partial information on $\Lambda^*$:

1) Lower and upper bounds. $LB_i$ and $UB_i$, for each scaling constant $k_i$ are provided by the Decision Maker:
   \[ LB_i \leq k_i \leq UB_i \text{ for } i = 1, 2, ..., m. \]  (type a)

2) Ranking of pairs of scaling constants is provided by the Decision Maker:
   \[ k_i > k_p \text{ for } i, p = 1, 2, ..., m, i \neq p. \]  (type b)

3) Paired comparisons of some alternatives are provided by the Decision Maker: that is, if an alternative $x_j$ is closer to cluster goal $c_1$ (or cluster goal $c_2$) than is another alternative $x_p$, then $d(x_j, c_1) < d(x_p, c_1)$, or
In this paper, we suggest using a type (c) paired comparison test for deciding between two multiple criteria alternatives. In the remainder of this section, we examine type c information and determine those alternatives that can be picked up for paired comparison so that we will get the most information on \( k \). Some other forms of questions can be used to generate partial information. In marginal rate of substitution questions (see Keeney and Raiffa 1976), the Decision Maker can provide bounds instead of exact values. Also, the Decision Maker can provide strength of preferences for paired comparisons of alternatives, e.g., he/she may state strong or weak preferences for paired comparisons: see Malakooti (1988). Hence all different types of partial information \( \Lambda^* \) can be represented by a set of linear constraints.

We arbitrarily choose an alternative from the set \( X_u \) as a reference alternative and denote it as \( x^0 \); then we call the distance between \( x^0 \) and cluster goal \( c_1 \), or \( d(x^0, c_1) \), the reference distance. If \( d(x^0, c_1) > d(x, c_1) \) for some \( k \in \Lambda \), where \( x \in X_u \), then alternative \( x \) is closer to cluster goal \( c_1 \) with respect to the reference alternative \( x^0 \). If \( d(x^0, c_1) > d(x, c_1) \) for all \( k \in \Lambda \), where \( x \in X_u \), then alternative \( x \) is definitely closer to cluster goal \( c_1 \) with respect to the reference alternative \( x^0 \).

**Remark 4-3** An alternative \( x, x \in X_u \), is definitely closer to cluster goal \( c_1 \) with respect to the reference alternative \( x^0 \) if and only if

\[
\min_k [d(x^0, c_1) - d(x, c_1)] > 0, \text{ where } k \in \Lambda \text{ and } x \in X_u.
\]

**Proposition 4-4** When \( k \) is unknown and \( k \in \Lambda \), an alternative \( x \in X_u \) is definitely closer to cluster goal \( c_1 \) with respect to reference alternative \( x^0 \) if and only if

\[
(x_1 - c_{11})^2 < (x_{1}^0 - c_{11})^2, (x_2 - c_{12})^2 < (x_{2}^0 - c_{12})^2, \ldots, (x_m - c_{1m})^2 < (x_m^0 - c_{1m})^2; \text{ where } x^0 = [x_1^0, x_2^0, \ldots, x_m^0], x = [x_1, x_2, \ldots, x_m], \text{ and } c_1 = [c_{11}, c_{12}, \ldots, c_{1m}].
\]

**Proposition 4-5** When \( k \) is unknown and \( k \in \Lambda \), an alternative \( x \in X_u \) is definitely closer to cluster goal \( c_1 \) with respect to reference alternative \( x^0 \) if and only if \( Z < 0 \) in Problem 2:

\[
(\text{Problem 2}) \quad Z = \max \sum_{i=1}^{m} (x_i - c_{1i})^2 k_i + v
\]
subject to \[ \sum_{i=1}^{m} (x_i^0 - c_{1i})^2 k_i + v \leq 0 \]

\[ k \in \Lambda \]

\[ k_i \geq \varepsilon, \quad i = 1, 2, ..., m \]

In general, \( v \) is a variable unrestricted in sign, but note that here \( v \) is negative as discussed below. \( \varepsilon \) is a small positive number (\( \varepsilon = 0.001 \)), \( x^0 = [x_1^0, x_2^0, ..., x_m^0] \), \( x = [x_1, x_2, ..., x_m] \), and \( c_1 = [c_{11}, c_{12}, ..., c_{1m}] \).

**Proposition 4-6** When two alternatives \( x_1 \in X_u \) and \( x_2 \in X_u \) are definitely closer to cluster goal \( c_1 \) with respect to reference alternative \( x^0 \), then alternatives which are convex combinations of \( x_1 \) and \( x_2 \) are also definitely closer to \( c_1 \) with respect to reference alternative \( x^0 \).

By using Proposition 4-4, 4-5, and 4-6, we can determine whether an unclassified alternative is definitely closer to cluster goal \( c_1 \) (or cluster goal \( c_2 \)) with respect to reference alternative \( x^0 \). Hence all unclassified alternatives can be divided into two groups according their closeness to cluster goal \( c_1 \) (or cluster goal \( c_2 \)) over the reference alternative. We call those alternatives, which are not definitely closer to cluster goal \( c_1 \) (or cluster goal \( c_2 \)) over the reference alternative, informative alternatives and denote them as \( X_T \). Clearly, the partial information about \( k \) can be obtained only by paired comparison of those alternatives \( x \in X_T \).

**Asking Direct Clustering Questions:**

We can generate constraints of type \( c \), by asking direct clustering questions. For example, we may ask the Decision Maker, to identify an alternative \( x_j \) as belonging to a particular cluster center (i.e. if it is closer to that cluster center, than to any other cluster center). Now, \( c_r, r = 1, 2, ..., R \) where \( R \) is the number of centers. We can generate \( R-1 \) number of constraints of type \( c \) and use them in the linear program as constraints. The above approach simplifies the paired comparison questions by asking simple clustering questions, while remaining very powerful in generating constraints in a way that could minimize the set size of the constraints on \( k \).

Let’s suppose \( x_j \) is closer to center \( c_p \), than any other centers, hence we generate \( R-1 \) constraints as follows:

\[
\sum_{i=1}^{m} (x_{ji} - c_{pi})^2 k_i < \sum_{i=1}^{m} (x_{ji} - c_{ri})^2 k_i
\]
Remark 4-4 We should choose informative alternatives $x \in X_T$ for presentation to the Decision Maker for soliciting clustering information (i.e. ask which cluster center alternative $x$ is closest to).

Remark 4-5 If there exist $q$ informative alternatives, then the Decision Maker needs to answer $q$ clustering questions.

Remark 4-6 The constraints $\Lambda^*$ on $k$ can be constructed by using the paired comparison information:

$$\sum_{i=1}^{m} (x_i^0 - c_{1i})^2 k_i < \sum_{i=1}^{m} (x_{ti} - c_{1i})^2 k_i$$

when the Decision Maker claims that the reference alternative $x^0$ is closer to the cluster goal $c_1$ than is $x_t$, or

$$\sum_{i=1}^{m} (x_i^0 - c_{1i})^2 k_i > \sum_{i=1}^{m} (x_{ti} - c_{1i})^2 k_i$$

when the Decision Maker claims that the reference alternative $x^0$ is farther from the cluster goal $c_1$ than is $x_t$.

where $x_t = (x_{t1}, x_{t2}, ..., x_{tm})$ is an informative alternative, $x_t \in X_T$.

$x^0 = (x_{10}, x_{20}, ..., x_{m0})$ is the reference alternative, and

$c_1 = (c_{11}, c_{12}, ..., c_{1m})$ is the cluster goal.

Figure 3. The relationship among alternative set $X$, unclassified alternative set $X_u$ and informative alternative set $X_T$.

Figure 3 depicts the relationship among alternative set $X$, unclassified alternative set $X_u$ and informative alternative set $X_T$.

After the informative alternative set $X_T$, consisting of $q$ informative alternatives, is identified, the Decision Maker will be asked to make $q-1$ paired comparisons with respect to a reference alternative, whereupon the constraints $\Lambda^*$ on scaling constant $k$ can be constructed. Then the
unclassified alternatives may be clustered according to the \( \Lambda^* \). After the last iteration of the method, if any alternative still cannot be clustered by the method, then the Decision Maker is asked to assign the cluster membership for such alternatives.
4.4. The Procedure of Clustering Alternatives

Using the theories developed in Section 4, we give the following procedure for the clustering of multiple criteria alternatives when partial information about the scaling constants is known from the Decision Maker.

**Procedure for Clustering Alternatives Using Paired Comparison of Two Alternatives to one Center**

1. **Step 1:** Identify the unclassified alternative set $X_u$ and cluster all classified alternatives.
2. **Step 2:** Arbitrarily choose a reference alternative $x_0$ from $X_u$. Identify the informative alternative set $X_T$ from the unclassified alternative set $X_u$.
3. **Step 3:** Have the Decision Maker make paired comparisons of the informative alternatives.
4. **Step 4:** Construct the constraints $L^*$ using the partial information obtained from the Decision Maker. Check the Decision Maker's consistency.
   - If there is no feasible solution of $L^*$, go to step 3 and have the Decision Maker revise his/her paired comparisons.
   - **Step 5:** Group unclassified alternatives according to the constraints $L^*$. If all unclassified alternatives are clustered, end. Otherwise, ask the Decision Maker to assign the cluster memberships for undecided alternatives, in $X_T$. Go to Step 1.

Now, if we want to use the direct clustering questions which are more efficient we use all steps of the above procedure except we change Step 3 as follows to Step 3’:

**Step 3’:** Have the Decision Maker identify to which clusters the informative alternatives belong to (hence generating R-1 number of constraints for each informative alternative based on Asking Direct Clustering Questions in Section 4.3).

In the following examples the steps of the above procedures are demonstrated.

4.5. Selection of the Most Satisfactory Alternative for Each Group

In Section 4, we discuss the theories and procedures for the first stage in our two-stage approach for solving MCDM problems, which is to cluster MCDM alternatives into groups having features similar to those of their corresponding ideal alternatives. The second stage of the proposed approach is to select the most satisfactory alternative for each group. As we mentioned before, we assume that the Decision Maker has an additive utility function for each group. To evaluate additive utility functions, many methods, mostly interactive ones, have been described in the literature. Our research emphasis in this paper is the clustering of MCDM alternatives, so we use existing methods to evaluate each subset and find the best alternative in each group, which is chosen as the one that maximizes the utility function corresponding to that group.
5. An Example and Experiments

In this section, we give an example to explain the developed theories and procedure for the clustering of multiple criteria alternatives. Then, we give the experimental results on several problems.

5.1. An Example to Explain the Developed Approach

Eight alternatives are given:
\[ x_1 = (2, 2.5) \quad x_2 = (4, 1.5) \quad x_3 = (3.5, 2.5) \quad x_4 = (3.1, 2.1) \]
\[ x_5 = (1.5, 0.5) \quad x_6 = (2.5, 4) \quad x_7 = (7/3, 3.5) \quad x_8 = (3.5, 4.1) \]

\( x_7 \) is a convex combination of \( x_1 \) and \( x_6 \) (as \( x_7 = (1/3) x_1 + (2/3) x_6 \)). The Decision Maker's two ideal alternatives (cluster goals) are \( c_1 = (1, 3) \) and \( c_2 = (5, 1) \).

![Figure 4. A Bi-criteria Example](image)

We summarize the solution procedure here. The complete details of the computations, interactions, and solution for this example appear in Appendix 3.

Summary of the Solution Procedure

Step 1: Identify the unclassified alternative set \( X_u \) and cluster all classified alternatives. \( X_u = \{ x_3, x_4, x_5, x_8 \} \). \( x_1, x_2, x_6, \) and \( x_7 \) are classified alternatives. \( x_1, x_6, \) and \( x_7 \) belong to cluster 1 and \( x_2 \) belongs to cluster 2.

Step 2: Choose a reference alternative \( x^0 = x_4 = (3.1, 2.1) \). Identify the informative alternative set \( X_T \) from the unclassified alternative set \( X_u \). \( X_T = \{ x_3, x_5 \} \).

Step 3: Have the Decision Maker make paired comparisons of the informative alternatives. The Decision Maker stipulates the following:
1) \( x_3 \) is farther from cluster goal \( c_1 \) with respect to reference alternative \( x^0 \);
2) \( x_5 \) is closer to cluster goal \( c_1 \) with respect to reference alternative \( x^0 \).
Step 4: Construct the constraints $\Lambda^*$ using the partial information from the Decision Maker. $\Lambda^* = \{ k_1 > 0.567, k \in \Lambda \}$.

Step 5: Group unclassified alternatives according to the constraint $\Lambda^*$. $x_3$ and $x_4$ are assigned to cluster 1, and $x_5$ to cluster 2. $x_8$ cannot be clustered according to $\Lambda^*$; the Decision Maker assigns it to cluster 2.

The final clustering of the eight alternatives is:
Cluster 1: $x_1, x_3, x_4, x_6, x_7$. Cluster goal $c_1 = (1, 3)$.
Cluster 2: $x_2, x_5, x_8$. Cluster goal $c_2 = (5, 1)$.

5.2. Computational Results on Several Problems

To further test our approach, we programmed it using C language and ran it on PC. We solved several problems. The alternatives were generated randomly with criterion values between 0 and 1. The Decision Maker’s ideal alternatives for clusters were inputted from the keyboard. Before using the Decision Maker’s preference information to cluster alternatives, the software identified the classified and unclassified alternatives. Then the Decision Maker selected an informative alternative for the cluster membership assignment. The resulting constraints were incorporated into $\Lambda^*$, and then the software checked whether more unclassified alternatives could be clustered according to $\Lambda^*$ automatically (without asking the Decision Maker). The procedure was repeated until all the alternatives were clustered.

Table 5.1 summarizes the experimental results. We note that in every problem, 75% or more of the alternatives were classified without soliciting any information from the Decision Maker. Among all the remaining unclassified alternatives, only 30% of them, on the average, were assigned membership by the Decision Maker. Overall, only less than 5% of the whole set of alternatives were directly assigned by the Decision Maker for the clustering to be completed.
6. Conclusions

In this paper, we presented the fundamental theories and algorithms for clustering multiple criteria alternatives by asking paired comparison question of alternatives. We developed an interactive method and demonstrated that the interactive method substantially reduces the number of paired comparison questions asked from the Decision Maker. We developed theories and procedures for the clustering of multiple criteria alternatives into classified, unclassified. We demonstrated how to identify the informative alternatives from the unclassified set in order substantially reduce the number of questions proposed to the decision maker. We demonstrated that complete clustering of alternatives can be accomplished by making a limited number of paired comparisons, and we showed how the minimum amount of information needed for complete clustering of alternatives could be identified. Examples were used to explain the developed theories and procedures. The future development of this work includes clustering physical objects presented graphically to the decision maker; applying the selection of the most preferred alternative for each cluster, and considering the problem of multiple decision makers.

Acknowledgement
The authors are indebted to the comments of the anonymous referees. We are also grateful to Mr. Michael Brunetti who helped with an appendix and commented on the paper. The authors would also like to thank Joel Mathewson for his input and time to revise the paper thoroughly.
References


Appendix 1
Proofs of Remarks and Propositions

Proof of Remark 4-1: 1. Because \(x\) definitely belongs to cluster \(c_1\), by definition, \(d(x, c_2) > d(x, c_1)\) for all \(k \in \Lambda\). Clearly, \(d(x, c_2) - d(x, c_1) > 0\) for all \(k \in \Lambda\). Therefore, \(\min_k \{d(x, c_2) - d(x, c_1)\} > 0\).

2. Because \(\min_k \{d(x, c_2) - d(x, c_1)\} > 0\), it follows that \(d(x, c_2) - d(x, c_1) > 0\) for all \(k \in \Lambda\), or \(d(x, c_2) > d(x, c_1)\) for all \(k \in \Lambda\). Therefore, \(x\) definitely belongs to cluster \(c_1\).

###

Proof of Proposition 4-1: 1. If an alternative \(x\) definitely belongs to cluster \(c_1\), then, from Remark 4-1, \(\min_k \{d(x, c_2) - d(x, c_1)\} > 0\).

Let \(J = d(x, c_2) - d(x, c_1)\)

\[
J = \frac{k_1(x_1 - c_{21})^2 + k_2(x_2 - c_{22})^2 + \cdots + k_m(x_m - c_{2m})^2 - d(x, c_1)}{\sqrt{k_1(x_1 - c_{11})^2 + k_2(x_2 - c_{12})^2 + \cdots + k_m(x_m - c_{1m})^2}} \tag{A-1}
\]

The first-order conditions are

\[
\frac{\partial J}{\partial k_1} = \frac{(x_1 - c_{21})^2}{2k_1(x_1 - c_{21})^2 + k_2(x_2 - c_{22})^2 + \cdots + k_m(x_m - c_{2m})^2} - \frac{(x_1 - c_{11})^2}{2k_1(x_1 - c_{11})^2 + k_2(x_2 - c_{12})^2 + \cdots + k_m(x_m - c_{1m})^2}
\]

\[
\frac{\partial J}{\partial k_2} = \frac{(x_2 - c_{22})^2}{2k_1(x_1 - c_{21})^2 + k_2(x_2 - c_{22})^2 + \cdots + k_m(x_m - c_{2m})^2} - \frac{(x_2 - c_{12})^2}{2k_1(x_1 - c_{11})^2 + k_2(x_2 - c_{12})^2 + \cdots + k_m(x_m - c_{1m})^2}
\]

\[
\frac{\partial J}{\partial k_m} = \frac{(x_m - c_{2m})^2}{2k_1(x_1 - c_{21})^2 + k_2(x_2 - c_{22})^2 + \cdots + k_m(x_m - c_{2m})^2} - \frac{(x_m - c_{1m})^2}{2k_1(x_1 - c_{11})^2 + k_2(x_2 - c_{12})^2 + \cdots + k_m(x_m - c_{1m})^2} \tag{A-2}
\]

Let \(\frac{\partial J}{\partial k_1} = 0\); then,

\[
\sqrt{k_1(x_1 - c_{21})^2 + k_2(x_2 - c_{22})^2 + \cdots + k_m(x_m - c_{2m})^2} = \frac{[(x_1 - c_{21})^2 / (x_1 - c_{11})^2] k_1(x_1 - c_{11})^2 + k_2(x_2 - c_{12})^2 + \cdots + k_m(x_m - c_{1m})^2}{\sqrt{k_1(x_1 - c_{11})^2 + k_2(x_2 - c_{12})^2 + \cdots + k_m(x_m - c_{1m})^2}}
\]
Substituting (A-3) into (A-1), we get
\[
\min J = \left[ \frac{(x_1 - c_{21})^2}{(x_1 - c_{11})^2} - 1 \right] \sqrt{k_1(x_1 - c_{11})^2 + k_2(x_2 - c_{12})^2 + \ldots + k_m(x_m - c_{1m})^2} \\
k
\]
Because \( \min J > 0 \), we have \( \left[ \frac{(x_1 - c_{21})^2}{(x_1 - c_{11})^2} - 1 \right] > 0 \), so \( (x_1 - c_{11})^2 < (x_1 - c_{21})^2 \).

Similarly, let \( \partial J/\partial k_2 = 0 \), \( \partial J/\partial k_3 = 0 \), \ldots, \( \partial J/\partial k_m = 0 \). We get
\[
(x_2 - c_{12})^2 < (x_2 - c_{22})^2, (x_3 - c_{13})^2 < (x_3 - c_{23})^2, \ldots, (x_m - c_{1m})^2 < (x_m - c_{2m})^2.
\]
2. Because \( (x_1 - c_{11})^2 < (x_1 - c_{21})^2 \), \( (x_2 - c_{12})^2 < (x_2 - c_{22})^2 \), \ldots, \( (x_m - c_{1m})^2 < (x_m - c_{2m})^2 \), we have
\[
\sqrt{k_1(x_1 - c_{11})^2 + k_2(x_2 - c_{12})^2 + \ldots + k_m(x_m - c_{1m})^2} > \sqrt{k_1(x_1 - c_{21})^2 + k_2(x_2 - c_{22})^2 + \ldots + k_m(x_m - c_{2m})^2}
\]
or \( d(x, c_2) > d(x, c_1) \) for all \( k \in \Lambda \). Therefore an alternative \( x \) definitely belongs to cluster \( c_1 \).

We can prove the conditions for which an alternative \( x \) definitely belongs to cluster \( c_2 \) by using the same procedures.

**Proof of Proposition 4-2:** We first prove that if \( Z < 0 \), then alternative \( x \) definitely belongs to cluster \( c_1 \). If \( Z < 0 \), then from constraint (A-5) we have \( d(x, c_2) = -v \). In the optimal solution, constraint (A-5) is equality because \( v \) is maximized and all \( k_i \) are positive, i.e., \( d(x, c_2) = -v \). Furthermore, \( -v \) is minimized in (A-4). Since in (A-4), \( Z = d(x, c_1) - (-v) < 0 \), we have \( Z = \max \{ d(x, c_1) - d(x, c_2) \mid k \in \Lambda \} < 0 \), hence \( d(x, c_1) < d(x, c_2) \).

We now prove that if an alternative \( x \) definitely belongs to cluster \( c_1 \), then \( Z < 0 \). Let us consider the optimal solution of Problem 1. From (A-5) we have \( d(x, c_2) = -v \) and the equality holds. We know \( d(x, c_1) - d(x, c_2) = d(x, c_1) - (-v) \). Hence if \( d(x, c_1) - d(x, c_2) < 0 \), then \( d(x, c_1) - (-v) < 0 \), which implies \( Z < 0 \).

**Proof of Proposition 4-3:** Because \( x_1 \) and \( x_2 \) definitely belong to cluster \( c_1 \), from Proposition 4-1, we have
\[
(x_{11} - c_{11})^2 < (x_{11} - c_{21})^2, (x_{12} - c_{12})^2 < (x_{12} - c_{22})^2, \ldots, (x_{1m} - c_{1m})^2 < (x_{1m} - c_{2m})^2 \quad (A-8)
\]
\[
(x_{21} - c_{11})^2 < (x_{21} - c_{21})^2, (x_{22} - c_{12})^2 < (x_{22} - c_{22})^2, \ldots, (x_{2m} - c_{1m})^2 < (x_{2m} - c_{2m})^2 \quad (A-9)
\]
From (A-8) and (A-9), we have
\[
-2x_{11}c_{11} + c_{11}^2 < -2x_{11}c_{21} + c_{21}^2
\]
\[-2x_{12}c_{12} + c_{12}^2 < -2x_{12}c_{22} + c_{22}^2\]

\[\ldots\]

\[-2x_{1m}c_{1m} + c_{1m}^2 < -2x_{1m}c_{2m} + c_{2m}^2 \quad (A-10)\]

\[-2x_{21}c_{11} + c_{11}^2 < -2x_{21}c_{21} + c_{21}^2\]

\[-2x_{22}c_{12} + c_{12}^2 < -2x_{22}c_{22} + c_{22}^2\]

\[\ldots\]

\[-2x_{2m}c_{1m} + c_{1m}^2 < -2x_{2m}c_{2m} + c_{2m}^2 \quad (A-11)\]

then we compute \(\mu\) (A-10) + (1 - \(\mu\)) (A-11):

\[-2\mu x_{11}c_{11} + \mu c_{11}^2 - 2(1 - \mu)x_{21}c_{11} + (1 - \mu)c_{11}^2 < -2\mu x_{11}c_{21} + \mu c_{21}^2 - 2(1 - \mu)x_{21}c_{21} + (1 - \mu)c_{21}^2\]

\[-2\mu x_{12}c_{12} + \mu c_{12}^2 - 2(1 - \mu)x_{22}c_{12} + (1 - \mu)c_{12}^2 < -2\mu x_{22}c_{22} + \mu c_{22}^2 - 2(1 - \mu)x_{22}c_{22} + (1 - \mu)c_{22}^2\]

\[\ldots\]

\[-2\mu x_{1m}c_{1m} + \mu c_{1m}^2 - 2(1 - \mu)x_{2m}c_{1m} + (1 - \mu)c_{1m}^2 < -2\mu x_{2m}c_{2m} + \mu c_{2m}^2 - 2(1 - \mu)x_{2m}c_{2m} + (1 - \mu)c_{2m}^2 \quad (A-12)\]

from (A-12), we have

\[[\mu x_{11} + (1 - \mu)x_{21}]^2 - 2[\mu x_{11} + (1 - \mu)x_{21}]c_{11} + c_{11}^2 < [\mu x_{11} + (1 - \mu)x_{21}]^2 - 2[\mu x_{11} + (1 - \mu)x_{21}]c_{21} + c_{21}^2\]

\[[\mu x_{12} + (1 - \mu)x_{22}]^2 - 2[\mu x_{12} + (1 - \mu)x_{22}]c_{12} + c_{12}^2 < [\mu x_{12} + (1 - \mu)x_{22}]^2 - 2[\mu x_{12} + (1 - \mu)x_{22}]c_{22} + c_{22}^2\]

\[\ldots\]

\[[\mu x_{1m} + (1 - \mu)x_{2m}]^2 - 2[\mu x_{1m} + (1 - \mu)x_{2m}]c_{1m} + c_{1m}^2 < [\mu x_{1m} + (1 - \mu)x_{2m}]^2 - 2[\mu x_{1m} + (1 - \mu)x_{2m}]c_{2m} + c_{2m}^2 \quad (A-13)\]

finally, we have

\[[\mu x_{11} + (1 - \mu)x_{21}]^2 - c_{11}^2 < ([\mu x_{11} + (1 - \mu)x_{21}]^2 - c_{21})^2\]

\[[\mu x_{12} + (1 - \mu)x_{22}]^2 - c_{12}^2 < ([\mu x_{12} + (1 - \mu)x_{22}]^2 - c_{22})^2\]

\[\ldots\]

\[[\mu x_{1m} + (1 - \mu)x_{2m}]^2 - c_{1m}^2 < ([\mu x_{1m} + (1 - \mu)x_{2m}]^2 - c_{2m})^2\]

we conclude that alternatives which are convex combination of \(x_1\) and \(x_2\) also definitely belong to \(c_1\). Thus, the proof is finished.
Proof of Remark 4-3: 1. Because \( x \) is definitely closer to cluster goal \( c_1 \) with respect to reference alternative \( x^0 \), then, by definition, \( d(x^0, c_1) > d(x, c_1) \) for all \( k \in \Lambda \). Clearly, \( \sum_{k} [d(x^0, c_1) - d(x, c_1)] > 0 \) for all \( k \in \Lambda \). Therefore, \( \min_{k} [d(x^0, c_1) - d(x, c_1)] > 0 \).

2. Because \( \min_{k} [d(x^0, c_1) - d(x, c_1)] > 0 \), then \( \sum_{k} [d(x^0, c_1) - d(x, c_1)] > 0 \) for all \( k \in \Lambda \), that is \( d(x^0, c_1) > d(x, c_1) \) for all \( k \in \Lambda \). Therefore, \( x \) is definitely closer to cluster goal \( c_1 \) with respect to reference alternative \( x^0 \).

Appendix 2

An Example and Explanation of the variable \( v \) in the linear program

Consider a linear program with an alternative \( x = (3, 4) \) and two cluster centers \( c_1 = (1,2) \) and \( c_2 = (5,8) \), and assume that \( k_1 = k_2 = .5 \). The linear program objective function determines the sum of the distances between each attribute of alternative \( x \) and the corresponding alternatives of a selected cluster. Then the constraint of another sum of distances between the attributes of the same alternative with the remaining clusters is applied to the program. The linear program will be max \( Z = (3-1)^2(.5) + (4-2)^2(.5) + v = 4 + v \) subject to the constraints \( (3-5)^2(.5) + (4-8)^2(.5) + v = 10 + v \leq 0 \). The \( v \) variable provides a comparison between the distances calculated in the objective function and the constraints; here \( v = 10 \) which means the maximum \( Z = -6 \).

Thus, the linear program proves that the alternative \( x \) is closer to the cluster center \( c_1 \) because \( Z < 0 \). The \( v \) value represents the distance between the alternative \( x \) and the other cluster center \( c_2 \), but expresses the number as a negative value. This explains why the \( v \) value must be less than or equal to zero. Applying the \( v \) value to the objective function allows for the two distances to be compared to each other. If the objective function is greater than zero, this means that the alternative is closer to the cluster center \( c_2 \), and vice versa: if the value is less than zero then the alternative \( x \) is closer to \( c_1 \).
Appendix 3

Details for the Example of Section 5.1

Following are the details of the computations and interactions of solving the example in Section 5.1:

Step 1  Identify the unclassified alternative set \( X_u \) and cluster all classified alternatives.

\[ x_1 = (2, 2.5) : \]
\[ (x_{11} - c_{11})^2 = (2 - 1)^2 = 1; \quad (x_{11} - c_{21})^2 = (2 - 5)^2 = 9 \]
\[ (x_{12} - c_{12})^2 = (2.5 - 3)^2 = 0.25; \quad (x_{12} - c_{22})^2 = (2.5 - 1)^2 = 2.25 \]

Because \( (x_{11} - c_{11})^2 < (x_{11} - c_{21})^2 \) and \( (x_{12} - c_{12})^2 < (x_{12} - c_{22})^2 \), from Proposition 4-1, we know that \( x_1 \) definitely belongs to cluster 1. If we use Proposition 4-2, \( Z = \max \{ k_1(x_{11} - c_{11})^2 + k_2(x_{12} - c_{12})^2 + v \} = -2.006 < 0 \), and we reach the same conclusion.

\[ x_2 = (4, 1.5) : \]
\[ (x_{21} - c_{11})^2 = (4 - 1)^2 = 9; \quad (x_{21} - c_{21})^2 = (4 - 5)^2 = 1 \]
\[ (x_{22} - c_{12})^2 = (1.5 - 3)^2 = 2.25; \quad (x_{22} - c_{22})^2 = (1.5 - 1)^2 = 0.25 \]

Because \( (x_{21} - c_{11})^2 > (x_{21} - c_{21})^2 \) and \( (x_{22} - c_{12})^2 > (x_{22} - c_{22})^2 \), from Proposition 4-1, we know that \( x_2 \) definitely belongs to cluster 2. If we use Proposition 4-2, \( Z = \max \{ k_1(x_{21} - c_{21})^2 + k_2(x_{22} - c_{22})^2 + v \} = -0.756 < 0 \), and we reach the same conclusion.

\[ x_3 = (3.5, 2.5) : \]
\[ (x_{31} - c_{11})^2 = (3.5 - 1)^2 = 6.25; \quad (x_{31} - c_{21})^2 = (3.5 - 5)^2 = 2.25 \]
\[ (x_{32} - c_{12})^2 = (2.5 - 3)^2 = 0.25; \quad (x_{32} - c_{22})^2 = (2.5 - 1)^2 = 2.25 \]

Because \( (x_{31} - c_{11})^2 > (x_{31} - c_{21})^2 \) and \( (x_{32} - c_{12})^2 < (x_{32} - c_{22})^2 \), from Proposition 4-1, we know that \( x_1 \) is an unclassified alternative. If we use Proposition 4-2, \( Z = \max \{ k_1(x_{31} - c_{11})^2 + k_2(x_{32} - c_{12})^2 + v \} = 5.994 > 0 \), and we reach the same conclusion.

**Note:** \( d(x_3, c_1) = (6.25k_1 + 0.25k_2)^{1/2} \) and \( d(x_3, c_2) = (2.25k_1 + 2.25k_2)^{1/2} \). We can see that \( x_3 \) belongs to cluster 1 when \( k_1 < 1/3 \) and \( k_2 > 2/3 \), but \( x_3 \) belongs to cluster 2 when \( k_1 > 1/3 \) and \( k_2 < 2/3 \), so \( x_3 \) is an unclassified alternative.

\[ x_4 = (3.1, 2.1) : \]
\[ (x_{41} - c_{11})^2 = (3.1 - 1)^2 = 4.41; \quad (x_{41} - c_{21})^2 = (3.1 - 5)^2 = 3.61 \]
\[ (x_{42} - c_{12})^2 = (2.1 - 3)^2 = 0.81; \quad (x_{42} - c_{22})^2 = (2.1 - 1)^2 = 1.21 \]

Because \( (x_{41} - c_{11})^2 > (x_{41} - c_{21})^2 \) and \( (x_{42} - c_{12})^2 < (x_{42} - c_{22})^2 \), from Proposition 4-1, we know that \( x_4 \) is an unclassified alternative. If we use Proposition 4-2, \( Z = \max \{ k_1(x_{41} - c_{11})^2 + k_2(x_{42} - c_{12})^2 + v \} = ** > 0 \), and we reach the same conclusion.

**Note:** \( d(x_4, c_1) = (4.41k_1 + 0.81k_2)^{1/2} \) and \( d(x_4, c_2) = (3.61k_1 + 1.21k_2)^{1/2} \). We can see that \( x_4 \) belongs to cluster 1 when \( k_1 < 1/3 \) and \( k_2 > 2/3 \), but \( x_4 \) belongs to cluster 2 when \( k_1 > 1/3 \) and \( k_2 < 2/3 \), so \( x_4 \) is an unclassified alternative.

\[ x_5 = (1.5, 0.5) : \]
(x_{51} - c_{11})^2 = (1.5 - 1)^2 = 0.25; \ (x_{51} - c_{21})^2 = (1.5 - 5)^2 = 12.25
(x_{52} - c_{12})^2 = (0.5 - 3)^2 = 6.25; \ (x_{52} - c_{22})^2 = (0.5 - 1)^2 = 0.25

Because (x_{51} - c_{11})^2 < (x_{51} - c_{21})^2 and (x_{52} - c_{12})^2 > (x_{12} - c_{22})^2, from Proposition 4-1, we know that \( x_1 \) is an unclassified alternative. If we use Proposition 4-2, \( Z = \max \{k_1(x_{51} - c_{11})^2 + k_2(x_{52} - c_{12})^2 + v \} = ** > 0, \) and we reach the same conclusion.

Note: \( d(x_5, c_1) = (0.25k_1 + 6.25k_2)^{1/2} \) and \( d(x_5, c_2) = (12.25k_1 + 0.25k_2)^{1/2} \) We can see that \( x_5 \) belongs to cluster 1 when \( k_1 > 1/3 \) and \( k_2 < 2/3, \) but \( x_4 \) belongs to cluster 2 when \( k_1 < 1/3 \) and \( k_2 > 2/3, \) so \( x_4 \) is an unclassified alternative.

\( x_6 = (2.5, 4): \)
\[ (x_{61} - c_{11})^2 = (2.5 - 1)^2 = 2.25; \ (x_{61} - c_{21})^2 = (2.5 - 5)^2 = 6.25 \]
\[ (x_{62} - c_{12})^2 = (4 - 3)^2 = 1; \ (x_{62} - c_{22})^2 = (4 - 1)^2 = 9 \]

Because \( (x_{61} - c_{11})^2 < (x_{61} - c_{21})^2 \) and \( (x_{62} - c_{12})^2 < (x_{62} - c_{22})^2, \) from Proposition 4-1, we know that \( x_6 \) definitely belongs to cluster 1. If we use Proposition 4-2, \( Z = \max \{k_1(x_{61} - c_{11})^2 + k_2(x_{62} - c_{12})^2 + v \} = -4.004 < 0, \) and we reach the same conclusion.

\( x_7 = (7/3, 3.5): \)
Note that \( x_7 = (7/3, 3.5) \) and is a convex combination of \( x_1 \) and \( x_6, \) which are members of the cluster \( c_1. \) The following calculation verifies the preceding statement.

Solving for \( \mu_1 \) and \( \mu_2 \) of the above set of equations, it can be determined that the \( \mu_1 = 1/3 \) and \( \mu_2 = 2/3. \)
\[ x_7 = \mu_1 x_1 + \mu_2 x_6 \]
\[ = \frac{7}{3} \begin{bmatrix} 2/5 \\ 2/5 \end{bmatrix} + \frac{2}{3} \begin{bmatrix} 25/4 \\ 25/4 \end{bmatrix} = \begin{bmatrix} 4k_1 + 2.5k_2 \\ 25k_1 + 4k_2 \end{bmatrix} \]
The values of the two weights satisfy the constraints of the convex combination theorem because \( x_7 \) is a convex combination of \( x_1 \) and \( x_6, \) that is \( x_7 = (1/3) x_1 + (2/3) x_6. \) Since \( x_1 \) and \( x_6 \) definitely belong to cluster 1, from Proposition 4-3, we conclude that \( x_7 \) also definitely belongs to cluster 1.

\( x_8 = (3.5, 4.1): \)
\[ (x_{81} - c_{11})^2 = (3.5 - 1)^2 = 6.25; \ (x_{81} - c_{21})^2 = (3.5 - 5)^2 = 2.25 \]
\[ (x_{82} - c_{12})^2 = (4.1 - 3)^2 = 1.21; \ (x_{82} - c_{22})^2 = (4.1 - 1)^2 = 9.61 \]

Because \( (x_{81} - c_{11})^2 > (x_{81} - c_{21})^2 \) and \( (x_{82} - c_{12})^2 < (x_{82} - c_{22})^2, \) from Proposition 4-1, we know that \( x_8 \) is an unclassified alternative. If we use Proposition 4-2, \( Z = \max \{k_1(x_{81} - c_{11})^2 + k_2(x_{82} - c_{12})^2 + v \} = ** > 0, \) and we reach the same conclusion.

Note: \( d(x_8, c_1) = (6.25k_1 + 1.21k_2)^{1/2} \) and \( d(x_8, c_2) = (2.25k_1 + 9.61k_2)^{1/2}. \) We can see that \( x_8 \) belongs to cluster 1 when \( k_1 < 2/3 \) and \( k_2 > 1/3, \) but \( x_8 \) belongs to cluster 2 when \( k_1 > 2/3 \) and \( k_2 < 1/3, \) so \( x_8 \) is an unclassified alternative.

Hence, when \( k \in \Lambda \) is unknown, \( x_1, x_2, x_6, \) and \( x_7 \) are clustered; \( x_1, x_6, \) and \( x_7 \) belong to cluster 1, and \( x_2 \) belongs to cluster 2. We also know that \( x_3, x_4, x_5, \) and \( x_8 \) are unclassified.
alternatives; their cluster memberships will depend on the importance coefficient $k$. We identify the unclassified alternative set $X_u = \{x_3, x_4, x_5, x_8\}$.

**Step 2** Identify the informative alternative set $X_T$ from the unclassified alternative set $X_u$.

We use the theories developed in Section 4.2 to investigate the four unclassified alternatives $x_3$, $x_4$, $x_5$, and $x_8$. From the results in step 1, we know that alternatives $x_1$, $x_6$, $x_7$ belong to cluster $c_1$ and alternative $x_2$ belongs to cluster $c_2$.

We choose $x_4 = (3.1, 2.1)$ as a reference alternative, hence $x^0 = x_4$.

**Step 3** Paired comparison of the informative alternatives by the Decision Maker.

The Decision Maker gives the following results of paired comparison:
1) $x_3$ is farther from cluster goal $c_1$ with respect to reference alternative $x^0$;
2) $x_5$ is closer to cluster goal $c_1$ with respect to reference alternative $x^0$.

**Step 4** Construct the constraints $\Lambda^*$ using the partial information from the Decision Maker. Check the Decision Maker’s consistency.

We can construct the following constraint on $k$ according Remark 4-6:

$$\Lambda^* = \{k_1, k_2 \mid d(x_3, c_1) > d(x^0, c_1) \text{ and } d(x_5, c_1) < d(x^0, c_1), \ k \in \Lambda \}$$
= \{ k_1, k_2 \mid k_1(x_{31} - c_{11})^2 + k_2(x_{32} - c_{12})^2 > k_1(x_{10}^0 - c_{11})^2 + k_2(x_{20}^0 - c_{12})^2;\ k_1(x_{51} - c_{11})^2 + k_2(x_{52} - c_{12})^2 < k_1(x_{10}^0 - c_{11})^2 + k_2(x_{20}^0 - c_{12})^2, \ k \in \Lambda \}

= \{ k_1, k_2 \mid 6.25k_1 + 0.25k_2 > 4.41k_1 + 0.81k_2;\ 0.25k_1 + 6.25k_2 < 4.41k_1 + 0.81k_2, \ k \in \Lambda \}

or \ \Lambda^* = \{ k_1 > 0.567, \ k \in \Lambda \}.

Because the feasible set of \( \Lambda^* \) exists, the Decision Maker’s paired comparisons are consistent.

\textbf{Step 5} Group unclassified alternatives according to the constraint \( \Lambda^* \).

According to the partial information from the Decision Maker, the constraint \( \Lambda^* = \{ k_1 > 0.567, \ k \in \Lambda \} \). We cluster the unclassified alternatives \( x_3, x_4, x_5, \) and \( x_8 \) according to \( \Lambda^* = \{ k_1 > 0.5676, \ k \in \Lambda \} \).

\( x_3, \) and \( x_4 \) are clustered into cluster 1, and \( x_5 \) is clustered into cluster 2. \( x_8 \) cannot be clustered according \( \Lambda^* \); the Decision Maker assigns it to cluster 2.

The final clustering of the eight alternatives is:

Cluster 1: \( x_1, x_3, x_4, x_6, x_7 \). Cluster goal \( c_1 = (1, 3) \).

Cluster 2: \( x_2, x_5, x_8 \). Cluster goal \( c_2 = (5, 1) \).