# Preference learning from interval pairwise data. A distance-based approach

Esther Dopazo[1], Mauricio Ruiz-Tagle[2], and Juan Robles[1]

[1] Facultad de Informática, Technical University of Madrid, Campus de Montegancedo, CP28660, Boadilla del Monte (Madrid), Spain
{edopazo,jrobles}@fi.upm.es
[2] Facultad de Ciencias de la Ingeniería, Universidad Austral de Chile, General Lagos 2086, Campus Miraflores, Valdivia, Chile
mruiztag@uach.cl

**Abstract.** Preference learning has recently received a lot of attention within the machine learning field, concretely learning by pairwise comparisons is a well-established technique in this field. We focus on the problem of learning the overall preference weights of a set of alternatives from the (possibly conflicting) uncertain and imprecise information given by a group of experts into the form of interval pairwise comparison matrices. Because of the complexity of real world problems, incomplete information or knowledge and different patterns of the experts, interval data provide a flexible framework to account uncertainty and imprecision. In this context, we propose a two-stage method in a distance-based framework, where the impact of the data certainty degree is captured. First, it is obtained the group preference matrix that best reflects imprecise information given by the experts. Then, the crisp preference weights and the associated ranking of the alternatives are derived from the obtained group matrix. The proposed methodology is made operational by using an Interval Goal Programming formulation.

**Key words:** Preference learning, pairwise comparison matrices, interval data, distance methods, interval goal programming.

## 1 Introduction

Preference learning has recently received a lot of attention within the machine learning literature [1] [2]. Learning by pairwise comparisons is a well-established technique in this field. In particular, it is a useful inference tool for assessing the relative importance of several alternatives. Formally, we consider the following preference learning scenario: let $X = \{x_1, ..., x_n\} (n \geq 2)$ be a finite set of alternatives (or labels) and $\{E_1, ..., E_m\}$ be a group of $m$ experts. We assume that preference information provided by the expert $E_k$ is structured by a pairwise comparison (pc) matrix $M^k = (m_{ij}^k)$, where $m_{ij}^k > 0$ represents the relative importance of the alternative $i$ over the alternative $j$, given by expert $k$. In many applications, due to incomplete information or knowledge, unquantifiable information, imprecise data, etc., the information cannot be assessed precisely in a quantitative form, so interval assessments are a natural way for expressing preferences. Therefore, we will consider interval pc matrices in our work.

In this scenario, we focus on the problem of scoring and ranking alternatives by computing their crisp preference weights that best reflect interval pairwise preferences, $M^1, ..., M^m$, given by multiple experts.

In this context, we face with different problems: the imprecision management problem (expert preferences are expressed by interval data), the group problem (i.e., how to integrate preferences from multiple experts) and the problem of consistency (i.e., how to derive preference weights from interval pc matrices without consistency properties).

In the Analytic Hierarchy Process (AHP) [14] context, the problem of consistency for interval assessments is analysed in [3], [4], [5], [6] and [7]. On the same context, the consensus problem has been studied under the fuzzy approach in [8], [9] and [10]. In [11], a consistency-driven logarithmic goal programming approach is applied for dealing with interval data for a particular distance. A distance model for interval rankings has been proposed in [12].

The related works deal with imprecise data, but they do not consider data certainty degree in the learning process. Because in multiple experts problems information is non-homogeneous, it is represented by interval-valued data with different precision degree, we propose to consider it. On the other hand, most of the methods dealing with interval data lead to interval weights. When the interval weights overlap, there is no a unique ranking of alternatives. In this case, additional ranking procedures are required in order to compare the final alternative scores.

We propose a two-stage method in a general distance-based framework, where the impact of the data certainty degree is captured. First, a method to retrieve the group preferences from the conflicting and imprecise individual preferences is proposed. To do this, we look for the crisp information that best reflects the multiple experts preferences by using a $l_p$-metric relative to the precision data. In the second step, the overall preference weights of the alternatives are computing from the group preference information obtained in the first phase. The proposed approach is made operational with the help of Interval Goal Programming (GP).

The paper is organized as follows. Section 2 focuses on the formulation of the problem and describes the first stage of the proposed model, group preference learning. In section 3, it is presented the second stage of the model and finally, the main conclusions derived from this work are included in section 4.

## 2 Group preference learning

Let $X = \{x_1, ..., x_n\}(n \geq 2)$ be a finite set of alternatives and $\{E_1, ..., E_m\}$ a group of $m$ experts. We assume that expert $E_k$ is indecisive generating certain imprecision in his preferences. Consequently, he quantifies his preferences on the elements of $X$ giving an interval pc matrix on $X$, $M^k = ([\underline{m_{ij}^k}, \overline{m_{ij}^k}])$ as follows: he judges that alternative $i$ is between $\underline{m_{ij}^k}$ and $\overline{m_{ij}^k}$ times more important than alternative $j$ with $\underline{m_{ij}^k}, \overline{m_{ij}^k} > 0$ and $\underline{m_{ij}^k} < \overline{m_{ij}^k}$. Then, the interval comparison matrix $M^k$ is obtained

$$M^k = \begin{pmatrix} 1 & \left[\underline{m_{12}^k}, \overline{m_{12}^k}\right] & \cdots & \left[\underline{m_{1n}^k}, \overline{m_{1n}^k}\right] \\ \left[\underline{m_{21}^k}, \overline{m_{21}^k}\right] & 1 & \cdots & \left[\underline{m_{2n}^k}, \overline{m_{2n}^k}\right] \\ \vdots & \vdots & \vdots & \vdots \\ \left[\underline{m_{n1}^k}, \overline{m_{n1}^k}\right] & \left[\underline{m_{n2}^k}, \overline{m_{n2}^k}\right] & \cdots & 1 \end{pmatrix} \quad (1)$$

In practice, the complexity of the problem, imperfect and subjective judgements, different backgrounds of the experts, etc., lead to imprecise and incompatible pairwise information. Also, disjoint intervals could be assigned by different experts to the same objects, i.e. there could exist $i$ and $j$ and two different experts such that $\left[\underline{m_{ij}^{k_1}}, \overline{m_{ij}^{k_1}}\right] \cap \left[\underline{m_{ij}^{k_2}}, \overline{m_{ij}^{k_2}}\right] = \emptyset$.

In this phase, the challenge is to look for the group preferences that represent in some sense the multiple experts preference acting as a whole. The ideal solution meaning unanimous agreement, among all experts regarding all possible alternatives is difficult to achieve in real-life situations. Therefore, we provide a method for retrieving the group preference information that best reflects the multiple experts preferences $(M^1, ..., M^m)$ attending to the data precision. We provide a method to obtain a matrix $C$, named group preference matrix, such that all experts consider the information of $C$ to be close to their data. In order to measure the degree of closeness, we consider the $l_p$-distances family with weights relatives to data precision degree. Thus, we look for an $n \times n$ crisp positive matrix $C$, whose entries $c_{ii} = 1$ if $i = 1, .., n$ and $c_{ij}$ is obtained for each pair $(i, j)$ $(i \neq j)$ according to the following expression

$$\min_{c_{ij}>0} \left[ \sum_{k=1}^{m} \left( \frac{\left| m_{ij}^k - c_{ij} \right|}{\log \overline{m_{ij}^k} - \log \underline{m_{ij}^k}} \right)^p \right]^{1/p} \qquad if \;\; 1 \leq p < \infty \quad (2)$$

over the set of positive numbers.

For $p = \infty$ metric, for each $(i, j)$ $(i \neq j)$, the problem turns into the *minmax* problem.

$$\min_{c_{ij}>0} \left\{ \max_{k=1,..,m} \left\{ \frac{\left| m_{ij}^k - c_{ij} \right|}{\log \overline{m_{ij}^k} - \log \underline{m_{ij}^k}} \right\} \right\} \quad (3)$$

over the set of positive numbers.

We notice that in the above problems the input data are interval, $m_{ij}^k \in [\underline{m_{ij}^k}, \overline{m_{ij}^k}]$ that is the interval goal of the expert $k$ for each entry $(i, j)$.

The value $\log \overline{m_{ij}^k} - \log \underline{m_{ij}^k}$ is considered a measure of the imprecision degree of $m_{ij}^k$ data given by expert $k$. Because we work with pairwise estimations of the weight ratios, the logarithmic transformation is used to equalize the precision degree of the $m_{ij}^k$ data and its reciprocal $1/m_{ij}^k$ (assuming interval arithmetic).

It should be noticed that as the range of the data (the vagueness) increases, less importance is given to this data in the objective function.

In the posed problems, the relative residual aggregation is affected by the parameter $p$ of the distance. Thus as $p$ increases, more importance is given to the largest relative residual value. The extremes of this set are the distance $l_1$, which minimizes the sum of relative residual and the Chebyshev or Minmax metric $l_\infty$, which minimizes the maximum relative deviation. Metric $p = 2$ corresponds to the Euclidean distance, generating a least square problem.

Once the analytical model has been established, we focus on solving the proposed minimization problems. In order to board them with interval data, we consider Interval Goal Programming ([12] and [13]). In this context, for each pair $(i,j)$ $(i \neq j)$, we consider the common deviational variables used in GP (see for example [12]):

$$\underline{n_{ij}^k} = \frac{1}{2}\left[\frac{\left|\underline{m_{ij}^k} - c_{ij}\right|}{\log \overline{m_{ij}^k} - \log \underline{m_{ij}^k}} + \frac{\underline{m_{ij}^k} - c_{ij}}{\log \overline{m_{ij}^k} - \log \underline{m_{ij}^k}}\right] \tag{4}$$

$$\underline{p_{ij}^k} = \frac{1}{2}\left[\frac{\left|\underline{m_{ij}^k} - c_{ij}\right|}{\log \overline{m_{ij}^k} - \log \underline{m_{ij}^k}} - \frac{\underline{m_{ij}^k} - c_{ij}}{\log \overline{m_{ij}^k} - \log \underline{m_{ij}^k}}\right] \tag{5}$$

$$\overline{n_{ij}^k} = \frac{1}{2}\left[\frac{\left|\overline{m_{ij}^k} - c_{ij}\right|}{\log \overline{m_{ij}^k} - \log \underline{m_{ij}^k}} + \frac{\overline{m_{ij}^k} - c_{ij}}{\log \overline{m_{ij}^k} - \log \underline{m_{ij}^k}}\right] \tag{6}$$

$$\overline{p_{ij}^k} = \frac{1}{2}\left[\frac{\left|\overline{m_{ij}^k} - c_{ij}\right|}{\log \overline{m_{ij}^k} - \log \underline{m_{ij}^k}} - \frac{\overline{m_{ij}^k} - c_{ij}}{\log \overline{m_{ij}^k} - \log \underline{m_{ij}^k}}\right] \tag{7}$$

$$k = 1, \cdots, m$$

where $\underline{n_{ij}^k}$ and $\underline{p_{ij}^k}$ measure the relative under-achievement and the relative over-achievement with respect to the low target $\underline{m_{ij}^k}$, respectively. Variables $\overline{n_{ij}^k}$ and $\overline{p_{ij}^k}$ play the same role for the high target $\overline{m_{ij}^k}$. They quantify in terms of the extremes of the intervals, "how relatively far" the solution $c_{ij}$ is from the interval target for the expert $k$.

Because we consider interval target for each $(i,j)$, the unwanted deviation variables are $\underline{n_{ij}^k}$ and $\overline{p_{ij}^k}$ and they have to be minimized. Therefore, for each pair $(i,j)$ $(i \neq j)$, we look for the $c_{ij}$ data that minimizes the objective function:

$$\left[\sum_{k=1}^{m}\left(\underline{n_{ij}^k} + \overline{p_{ij}^k}\right)^p\right] \quad if\ 1 \leq p < \infty \tag{8}$$

subject to

$$
\begin{aligned}
&\frac{\underline{m}_{ij}^k - c_{ij}}{\log \overline{m}_{ij}^k - \log \underline{m}_{ij}^k} - \underline{n}_{ij}^k + \underline{p}_{ij}^k = 0 \qquad k = 1, ..., m && (9)\\
&\frac{\overline{m}_{ij}^k - c_{ij}}{\log \overline{m}_{ij}^k - \log \underline{m}_{ij}^k} - \overline{n}_{ij}^k + \overline{p}_{ij}^k = 0 \qquad k = 1, ..., m && (10)\\
&\underline{n}_{ij}^k, \overline{n}_{ij}^k, \underline{p}_{ij}^k, \overline{p}_{ij}^k \geq 0 \qquad k = 1, ..., m\\
&c_{ij} > 0
\end{aligned}
$$

Expressions (9) and (10), connecting the variable $c_{ij}$ with the new variables $\underline{n}_{ij}^k$, $\underline{p}_{ij}^k$, $\overline{n}_{ij}^k$ and $\overline{p}_{ij}^k$, have been obtained by substracting (4) from (5), and substracting (6) from (7), respectively.

For $p = \infty$, for each pair $(i, j)$ $(i \neq j)$, we get a mathematical programming problem $min\ \ D_{ij}$ over the nonnegative numbers subject to the above goals and constraints plus $\underline{n}_{ij}^k + \overline{p}_{ij}^k \leq D_{ij}, k = 1, ..., m$. $D_{ij}$ is an extra nonnegative variable that quantifies the maximum relative deviation for the $(i, j)$-entry.

For the most common values of $p$, $p = 1$ and $p = \infty$, the above formulations are reduced to linear programming problems that can be solved using the simplex method. The case $p = 2$ is a quadratic programming problem for which several numerical tools are available.

*Example 1.* Let us present a numerical example ([17]) to illustrate how the proposed methodology works. A group of four experts assess their preferences about four alternatives, through the pc interval matrices $M^1$, $M^2$, $M^3$ and $M^4$, on the Saaty's scale ([14]) as follows:

$$
\begin{aligned}
M^1 &= \begin{pmatrix} [1,1] & [5,9] & \left[\frac{1}{5},\frac{1}{3}\right] & [3,5] \\ \left[\frac{1}{9},\frac{1}{5}\right] & [1,1] & \left[\frac{1}{9},\frac{1}{5}\right] & \left[\frac{1}{7},\frac{1}{5}\right] \\ [3,5] & [5,9] & [1,1] & [3,7] \\ \left[\frac{1}{5},\frac{1}{3}\right] & [5,7] & \left[\frac{1}{7},\frac{1}{3}\right] & [1,1] \end{pmatrix}
\quad
M^2 = \begin{pmatrix} [1,1] & [1,3] & [3,5] & [3,5] \\ \left[\frac{1}{3},1\right] & [1,1] & \left[\frac{1}{7},\frac{1}{3}\right] & \left[\frac{1}{5},\frac{1}{3}\right] \\ \left[\frac{1}{5},\frac{1}{3}\right] & [3,7] & [1,1] & [1,5] \\ \left[\frac{1}{5},\frac{1}{3}\right] & [3,5] & \left[\frac{1}{5},1\right] & [1,1] \end{pmatrix}
\\
M^3 &= \begin{pmatrix} [1,1] & [3,5] & \left[\frac{1}{5},\frac{1}{3}\right] & [3,5] \\ \left[\frac{1}{5},\frac{1}{3}\right] & [1,1] & \left[\frac{1}{5},\frac{1}{3}\right] & \left[\frac{1}{7},\frac{1}{3}\right] \\ [3,5] & [3,5] & [1,1] & [5,7] \\ \left[\frac{1}{5},\frac{1}{3}\right] & [3,7] & \left[\frac{1}{7},\frac{1}{5}\right] & [1,1] \end{pmatrix}
\quad
M^4 = \begin{pmatrix} [1,1] & [3,5] & [1,3] & [5,7] \\ \left[\frac{1}{5},\frac{1}{3}\right] & [1,1] & \left[\frac{1}{5},\frac{1}{3}\right] & \left[\frac{1}{5},\frac{1}{3}\right] \\ \left[\frac{1}{3},1\right] & [3,5] & [1,1] & [3,5] \\ \left[\frac{1}{7},\frac{1}{5}\right] & [3,5] & \left[\frac{1}{5},\frac{1}{3}\right] & [1,1] \end{pmatrix}
\end{aligned}
\tag{11}
$$

We find that matrices given by the experts provide information discrepant and not compatible.

The proposed method is applied in order to find the consensus matrix $C$ (first phase) using $l_p$-metrics, for $p = 1$, $p = 2$ and $p = \infty$ and the results are listed in Table 1.

**Table 1.** Consensus matrix $C$ for $p = 1, 2, \infty$

| Matrix $C$ for $p = 1$ | Matrix $C$ for $p = 2$ | Matrix $C$ for $p = \infty$ |
|---|---|---|
| $\begin{pmatrix} 1.000 & 5.000 & 0.332 & 5.000 \\ 0.201 & 1.000 & 0.200 & 0.200 \\ 2.999 & 5.000 & 1.000 & 5.000 \\ 0.200 & 5.000 & 0.200 & 1.000 \end{pmatrix}$ | $\begin{pmatrix} 1.000 & 4.555 & 1.221 & 5.000 \\ 0.231 & 1.000 & 0.200 & 0.200 \\ 2.036 & 5.000 & 1.000 & 5.000 \\ 0.200 & 5.000 & 0.200 & 1.000 \end{pmatrix}$ | $\begin{pmatrix} 1.000 & 4.303 & 1.666 & 5.000 \\ 0.247 & 1.000 & 0.200 & 0.200 \\ 1.666 & 5.000 & 1.000 & 5.000 \\ 0.200 & 5.000 & 0.200 & 1.000 \end{pmatrix}$ |

## 3 Generating the group preference weights

Once the $l_p$-group preference matrix $C$ has been computed, the task is to obtain the crisp preference weights $w_1, ..., w_n$, of the alternatives from the matrix $C$. We assume that preference weights are positive and normalized, i.e. $\sum_{i=1}^{n} w_i = 1$.

Several procedures are available in the literature concerning this problem. The eigenvector method [14] is the standard method employed in the AHP context for reciprocal matrices. On the other hand, distance-based methods are provided by [15] and [16] among others. We adopt the distance-based approach followed in [16]. The idea is to look for $w = (w_1, ..., w_n)^t$ taking into account the consistency properties of the matrix $C$ in a $l_q$-distance framework. The priority vector is obtained by solving the following optimization problem:

$$\min_{w \in F} \left[ \sum_{i=1}^{n} \sum_{j=1}^{n} |c_{ij} w_j - w_i|^q \right]^{1/q} \qquad if\ \ 1 \le q < \infty \tag{12}$$

$$\min_{w \in F} \left[ \max_{i,j=1,..,n} \left( |c_{ij} w_j - w_i| \right) \right] \qquad if\ \ q = \infty$$

over the feasible set $F = \{w = (w_1, ..., w_n)^t / w_i > 0, \sum_{i=1}^{n} w_i = 1\}$.

We apply the second phase of the methodology to compute the preference weights associated to the matrices of the example given in section 2. For simplicity's sake, we assume the matrix $C$ has been obtained with $p = 2$. The obtained priority vectors and the associated rankings for the most usual values of $q$, $q = 1, 2, \infty$ are listed in Table 2.

We notice that there is a tie between options $x_2$ and $x_4$ for metric $q = 1$. This tie is solved in the results obtained with $q = 2$ and $q = \infty$ yielding dominance for option $x_4$ over $x_2$. We remark that as $q$ increases, the effect of greater deviations is emphasized.

We developed a GP matrix generator using MS Visual FoxPro® and problems were optimized using ILOG CPLEX® (Java classes).

**Table 2.** Preference weights and their associated rankings for $q = 1, 2, \infty$

| | Metric | | |
|---|---|---|---|
| | $q = 1$ | $q = 2$ | $q = \infty$ |
| $w$ | $(0.394, 0.087, 0.432, 0.087)^t$ | $(0.357, 0.067, 0.483, 0.093)^t$ | $(0.345, 0.069, 0.471, 0.115)^t$ |
| Ranking | $x_3 \succ x_1 \succ x_4 = x_2$ | $x_3 \succ x_1 \succ x_4 \succ x_2$ | $x_3 \succ x_1 \succ x_4 \succ x_2$ |

## 4 Conclusions

Methods for learning and predicting preferences in an automatic way is a topic in disciplines such as machine learning, recommendation systems and information retrieval systems. A problem in this field is the scoring and ranking of decision alternatives from imprecise preference information from different sources. We focus on the problem of learning the overall preference weights of a set of alternatives in a multiple interval pc matrices scenario.

The proposed methodology is articulated into two phases. First, it is provided a $l_p$-distance model attending to data precision, that synthesizes expert's interval pc matrices into a crisp group matrix. This phase may prove to be useful in a group decision problem where difficulties in articulating consensus information from conflicting interests and different viewpoints are most common. In this context the parameter $p$ has a consensus meaning, it places more or less emphasis on the relative contribution of individual deviations. On the other hand, the effect of the precision degree, attached to deviations in the objective function considered in the paper, is to place more or less emphasis on the relative contribution of data deviations according to the precise knowledge of the data.

In the second phase, we deal with the problem of outputting crisp weights of the alternatives from the group information. Most of the methods dealing with interval data lead to interval weights. When the interval weights overlap, there is no unique ranking of alternatives. In this case, additional ranking procedures are required in order to compare the final alternative scores.

Another key characteristic of our approach is the ability of numerical methods for computing the proposed solution for the most usual metrics.

*Acknowledgements* The authors would like to thank to the anonymous reviewers for their valuable comments and suggestions. This research has been partly supported by Project MTM2007-67232, "Ministerio de Educación y Ciencia" of Spain.

## References

1. Doyle, J.: Prospects for preferences. Computational Intelligence, Vol. 20(2) (2004) 111-136

2. Fürnkranz, J., Hüllermeier, E.: Pairwise preference learning and ranking. Proceedings of the 14th European Conference on Machine Learning, Cavtat-Dubrovnik, Croatia, September 22-26, 2003, LNCS, Vol. 2837 (2003) 145-156
3. Arbel, A., Vargas L.: Preference simulation and preference programming: Robustness issues in priority derivation. European Journal of Operational Research, Vol. 69(2) (1993) 200-209
4. Salo A.A., Hämäläinen R.P.: Preference programming through approximate ratio comparisons. European Journal of Operational Research, Vol. 82 (1995) 458-475
5. Wang Y.-M., Yang J.-B., Xu D.-L.: Interval weight generation approaches based on consistency test and interval comparison matrices. Applied Mathematics and Computation, Vol. 167(1) (2005) 252-273
6. Islam R., Biswal M.P., Alam S.S.: Preference programming and inconsistent interval judgments. European Journal of Operational Research, Vol. 97(1) (1997) 53-62
7. Wang Y.-M., Yang J.-B., Xu D.-L.: A two-stage logarithmic goal programming method for generating weights from interval comparison matrices. Fuzzy Sets and Systems, Vol. 152 (2005) 475-498
8. Boender C.G.E., de Graan J.G., Lootsma F.A.: Multicriteria decision analysis with fuzzy pairwise comparisons. Fuzzy Sets and Systems, Vol. 29(2) (1989) 133-143
9. Yu, C.-S.: A GP-AHP method for solving group decision-making fuzzy AHP problems. Computers and Operations Research, Vol. 29 (2002) 1969-2001
10. Mikhailov L.: Group prioritization in the AHP by fuzzy preference programming method. Computers and Operations Research, Vol. 31(2) (2004) 293-301
11. Bryson N.K.M., Joseph A.: Generating consensus priority interval vectors for group decision-making in the AHP. Journal of Multi-Criteria Decision Analysis, Vol. 9(4) (2000) 127-137
12. González-Pachón J., Romero C.: Aggregation of partial ordinal rankings: an interval goal programming approach. Computers and Operations Research, Vol. 28(8) (2001) 827-834
13. Ignizio J.: Linear programming in single -& multiple- objective systems. Prentice-Hall, (1982)
14. Saaty T.: The analytic hierarchy process. McGraw-Hill, New York (1980)
15. Jones D.F., Mardle S.J.: A distance-metric methodology for the derivation of weights from a pairwise comparison matrix. Journal of the Operational Research Society, Vol. 55 (2004) 869-875
16. Dopazo E., González-Pachón J.: Consistency-driven approximation of a pairwise comparison matrix. Kybernetyca, Vol. 39(5) (2003) 561-568
17. Dopazo E, Ruiz-Tagle M.: A GP formulation for aggregating preferences with interval assessments. 7th International Conference on Multiobjective Programming and Goal Programming, (submitted) (2006)