EVALUATING ATTRIBUTE SIGNIFICANCE IN AHP USING SHANNON ENTROPY

Ewa van Uden, Miroslaw Kwiesielewicz[1]
Faculty of Electrical and Control Engineering
Technical University of Gdansk
Narutowicza 11/12, 80-952 Gdansk – Poland
cotynato@ely.pg.gda.pl, mkwies@ely.pg.gda.pl

Keywords: decision-making, AHP, Shannon entropy

Summary: *Multi-attribute factor ranking in the sense of AHP is considered in this paper. It is assumed that in a multi-attribute decision-making process rankings of factors under particular attributes are given. Next, based on the Shannon entropy, an amount of information associated with each ranking is evaluated and an attribute ranking is fixed. The approach proposed is illustrated by a numerical example.*

1. Introduction

One of the widely used methods in multi-attribute (multi-criteria) decision making is AHP (Analytic Hierarchy Process) (Saaty, 1977, 1980). A hierarchical process of decision making has the following characteristics:

- its character is multi-leveled, tree structured,
- individual variants of the decision are characterized by many attributes,
- attributes may be given in numerical or linguistic form,
- the weights may be given for attributes,
- different experts or groups of experts may be involved in different decision levels; in such a case, we
  deal with group decision making,
- the base of the decision making process is determined on each level by a decision matrix, whose rows are equivalent to the alternatives (variants) and columns to attributes (criteria).

In AHP (Saaty, 1977, 1980), rankings for individual sub-attributes are created by the pairwise comparison method while aggregation is made by the simple additive weighting method (MacCrimmon, 1968). The process of ranking alternatives (factors) using pairwise comparison method was proposed by David (1963) and developed by Saaty (1977, 1980), who proposed the following procedure:

- evaluation by an expert of his preference of one factor compared to another for each pair of factors;
- providing, for each pair, a number from a previously defined scale;
- calculation of the ranking by the eigenvalue method (or by the logarithmic least-squared method or the ordinary least-squared method);
- arithmetic normalization of the result.

An example of a two-level decision process is presented in Fig. 1., which can however be simply developed into a tree structure. Thus a multi-criteria decision problem can be decomposed into separate sub-problems. These sub-problems can be solved independently, by applying the procedure described above.

[1] To Thomas Saaty thanks for a gentle support to participate at the Conference.
To Kris Wargan (NASA) thanks for inspiring discussions.

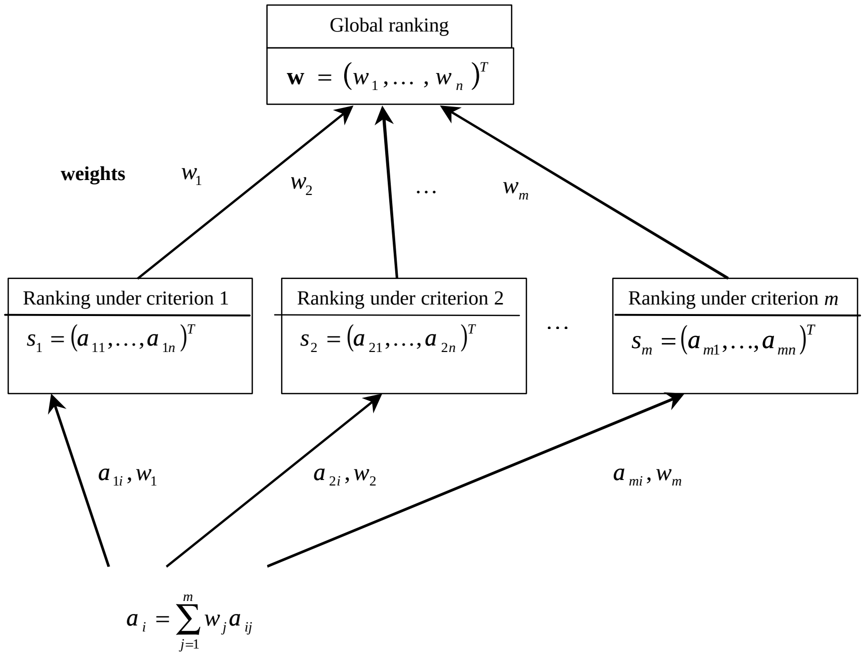

Fig. 1. Analytic Hierarchy Process

In recent years, there appeared some positions, which use Shannon's entropy in the AHP method in different aspects. Sanchez and Soyer (1998) applied Shannon's entropy as a criterion to stop pairwise comparisons for large size matrices with missing data, based on Harker's (1987) idea for estimating missing data. Mon and others (1994) used Shannon's entropy to obtain a ranking vector in a fuzzy version of AHP. Cheng (1996) proposed a new algorithm for evaluating weapon systems by AHP based on fuzzy scales. Harker (1987) noticed that in a decision problem consisting of many alternatives and criteria, the number of necessary opinions becomes very large, for example, with 9 alternatives and 5 criteria a group of experts must answer 190 questions. In such cases, an expert is not always able to evaluate each pair of factors, particularly for all criteria. Harker has proposed a method based on estimating missing data for such situations.

The aim of this paper is to aid the decision-making process and reduce its complexity by qualifying the importance of each criterion. To avoid the pairwise comparisons of criteria, the quantity of information contained in each criterion is measured by Shannon entropy. In other words, an intention of proposed method is to simplify the process of giving weights to criteria, so that certain pairs of them need not be evaluated.

In chapter 2, the basic definitions and some properties of Shannon's entropy will be reviewed. In chapter 3, the formulation will be given of an optimisation problem concerning the selection of criteria with the most information content. In chapter 4, a numerical illustration of the proposed approach will be shown. At the end, remarks and conclusions will be stated and the proposal for further research in this area will be presented.

2. Shannon's entropy and its basic properties

Let us assume, following (Sanchez and Soyer, 1998), that $\mathbf{p} = (p_1, \ldots, p_n)$ denotes a priority vector according to a certain criterion, after arithmetic normalisation (so that the vector's co-ordinates sum up to 1). Entropy for this vector may be defined as:

$$H(\mathbf{p}) = -\sum_{i=1}^{n} p_i \ln(p_i) \quad . \tag{1}$$

In information theory entropy H is defined as a measure of uncertainty of a discrete random variable X, which can take finite values $(x_1, \ldots, x_n)$ such that $P(X = x_i) = p_i$. In the AHP context, the priority p_i can be interpreted as the probability that the i –th alternative will be preferred by the decision-maker.

Among the most important properties of entropy, we recall that:

$$H(X) \geq 0 \quad . \tag{2}$$

The entropy of a discrete distribution with finite support is nonnegative; it is equal to zero when all the components of the sum (1) are equal to zero simultaneously. It is possible only when one value of a discrete random variable appears with probability one and the other values have probability zero. In such a case, there is no uncertainty about which value random variable will take.

Entropy reaches its only maximum for the uniform distribution, which is given by

$$H(1/n, \ldots, 1/n) = \ln n \quad . \tag{3}$$

This property is consistent with the interpretation of entropy as an uncertainty measure – the maximal value is reached when all values of random variable X are equally probable. Moreover, entropy is a concave function.

In uncertainty theory three principles are applied: the principle of maximum uncertainty, the principle of minimum uncertainty and the principle of uncertainty invariance (Klir, Yan, 1995). Following the idea that the smaller the entropy the bigger the quantity of information, the principle of minimum uncertainty may be applied to determine which criterion will give the most information to the decision maker.

Criteria ranking may prompt the decision-maker as to what weights should be given to the criteria in the situation when his preferences are not specified precisely. For example, - the ranking vector of alternatives given in the form $\mathbf{v} = [1/n, \ldots, 1/n]$ does not provide any definite information – all alternatives are treated equally. When distinctiveness of alternatives increases, the entropy of such a vector decreases. Ranking of attributes from the point of view of the alternatives' distinctiveness could prompt the decision-maker to choose weights or particular attributes to ensure that the chosen decision differs fundamentally from the others. One can give low weights to those attributes that deliver the least information (by a uniform distribution of the alternatives' ranking vector) because their influence on the final vector is insignificant and there is no appreciable difference between alternatives. The ranking vector closer to uniform distribution means that the preferences between alternatives become indistinguishable.

3. The attribute importance problem

Let us assume that a given decision problem consists of n possible alternatives $\mathbf{A_i}, i = 1, \ldots, n$ considered according to m criteria $\mathbf{K_j}, j = 1, \ldots, m$.

As a result of the alternatives' pairwise comparisons with respect to particular criteria one obtains the ranking vectors concerning given criteria in the form:

$$\mathbf{v}_j = (v_{j1}, \ldots, v_{jn}), \quad j = 1, \ldots, m \quad . \tag{4}$$

Next, the aggregation according to criteria is done. The simplest method is a weighted sum normalised to one. In the most usual case, the decision-maker himself chooses the weights for the criteria, - that

express his preferences. Sometimes criteria are also compared in pairs and a vector of weights is obtained as a result of these comparisons, for example in (Laarhoven, Pedrycz, 1983, Saaty , 1980).
Let us assume that weights $a_1,\ldots,a_m$ are unknown and satisfy the arithmetic normalisation condition:

$$\sum_{j=1}^{m} a_j = 1 \quad , \tag{5}$$

and

$$a_j \geq 0 \quad \forall j \in 1,\ldots,m \quad . \tag{6}$$

The result of aggregation is a final vector $\mathbf{v}$ in form:

$$\mathbf{v} = a_1\mathbf{v}_1 + \ldots + a_m\mathbf{v}_m \quad . \tag{7}$$

The question posed by the authors is: for what values of $a_1,\ldots,a_m$ will the entropy of a vector $\mathbf{v}$ reach its minimum value? This question allows the following interpretation: for what values $a_1,\ldots,a_m$ will the decision-maker get the maximum amount of information.

The above considerations led to the following optimisation problem:

$$\min\left\{H(\mathbf{v}) = -\sum_{i=1}^{n} v_i \ln v_i\right\} \tag{8}$$

where:

$$v_i = \sum_{j=1}^{m} a_j v_{ji} \tag{9}$$

taking into an account conditions (5) and (6).

4. A sample calculation

Left Let us assume a decision problem with 3 alternatives and 4 criteria. The ranking vectors corresponding to given criteria have the form:

$$\mathbf{v}_1 = [1/2,1/6,1/3]\,,\ \mathbf{v}_2 = [1/4,1/4,1/2]\,,\ \mathbf{v}_3 = [1/5,3/5,1/5]\,, \mathbf{v}_4 = [1/3,1/3,1/3]\,.$$

As a result of calculating the minimal value of function (1) with the constraints (5) and (6) the values of the weights equalled: $a_1 = 0$, $a_2 = 0$, $a_3 = 1$, $a_4 = 0$. The following conclusion can be made: the third criterion contains the greatest amount of information, the best distinguishing the alternatives. One notes the basic preference for the second alternative compered to the first and third ones.

The same procedure may be applied to the remaining vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4$. Now the values equal $a_1 = 1$, $a_2 = 0$, $a_4 = 0$. Vector $\mathbf{v}_1$ may be ranked in the second position from the point of view of the amount of information. The last application of the procedure concerns the vectors $\mathbf{v}_2, \mathbf{v}_4$. In this case the following values were obtained: $a_2 = 1$, $a_4 = 0$. The final ranking of the criteria is: $\mathbf{K_3}, \mathbf{K_1}, \mathbf{K_2}, \mathbf{K_4}$, starting from the criterion with the most information. It is worth noting that the fourth criterion appeared in the last position, which follows from (3).

5. Concluding remarks

This paper presents an algorithm, based on the entropy measure, to determine the criterion that provides the greatest amount of information. Consequently multiple application of the proposed algorithm allows for criteria ranking as well as elimination the criteria with the small amount of information from decision process. Presented method can be particularly useful when the quantity of criteria is very big and the decision-maker would like to reduce some of them. The plans of developing this approach to fuzzy version of AHP, where rankings may be given in a quality form, using linguistic variables, is left for future.

References

Cheng C.H. (1996) Evaluating naval tactical missile systems by fuzzy AHP based on the grade value of membership function. *European Journal of Operational Research* 96, 343–350.

Cheng C.H., Mon D.L., Lin J.–C. (1994) Evaluating weapon system using fuzzy analytic hierarchy process based on entropy weight. *Fuzzy Sets and Systems* 62, 127–134.

David H.A. (1963) *The Method of Paired Comparisons*, Griffin, London.

Harker P.T. (1987) Incomplete pairwise comparisons in the analytic hierarchy process. *Mathematical Modelling* 9 (11), 837–848.

Klir G.J., Yan B. (1995) *Fuzzy sets and fuzzy logic theory and applications*. Prentice–Hall International, Inc. London.

Laarhoven P.J.M van, Pedrycz W. (1983) A fuzzy extention of Saaty's priority theory. *Fuzzy Sets and Systems* 11, 229–241.

MacCrimmon K.R. (1968) *Decision Making Among Multiple–Attribute Alternatives: A Survey and Consolidated Approach*. RAND Memorandum, RM-4823-ARPA.

Saaty T.L. (1977) A scaling method for priorities in hierarchical structures, *Journal of Mathematical Psychology* 15, 234–281.

Saaty T.L. (1980) *The Analytic Hierarchy Process*, Mc-Graw Hill, New York.

Sanchez P.P., Soyer R. (1998) Information concepts and pairwise comparison matrices. *Information Processing Letters* 68, 185–188.

Sobczak W.: Metody statystyczne w elektronice. Wydawnictwa Naukowo–Techniczne, Warszawa 1971.