An analysis of the Analytic Hierarchy Process

by

Sándor Bozóki

Master's Thesis in Applied Mathematics

Supervisor: Róbert Fullér

Eötvös University, Budapest (ELTE)
Faculty of Sciences
Department of Operations Research

2001.

Contents

1 Introduction **3**

2 The Analytic Hierarchy Process **4**
2.1 Identity and Decomposition 4
2.2 Discrimination and Comparative Judgments 7

3 Synthesis of Priorities **12**
3.1 Eigenvector Method (EM) 12
3.2 Least Squares Method (LSM) 16
3.3 Logarithmic Least Squares Method (LLSM) 16
3.4 Chi Square Method (X^2M) 19
3.5 Singular Value Decomposition (SVD) 19

4 Critical factors **22**
4.1 Uniqueness of solution 22
4.2 Rank Reversals 22
4.3 Sensitivity Analysis 27
4.3.1 Changing the weights of criteria 28
4.3.2 Changing performances 29

5 Group Decision Making **32**
5.1 Aggregating individual judgments 32
5.2 Constructing a group choice 34
5.3 Group Decision based on Bridgman principle 34

6 Further observations **36**
6.1 A generalization of the discrete case 36
6.2 The Fundamental Scale 38
6.3 The Analytic Network Process (ANP) 40
6.4 Expert Choice 40

Appendix 1 **44**

Appendix 2 **45**

Appendix 3 **47**

Abstract

Multi-Criteria Decision Making (MCDM) is a prominent area of research in normative decision theory. The Analytic Hierarchy Process (AHP) is one of the popular methods employed in MCDM. It provides ratio-scale measurements of the priorities of elements in various levels of a hierarchy. These priorities are obtained through pairwise comparisons of elements in one level with reference to each element in the immediate higher level. This paper studies five methods for calculating the priorities. We also take some observations about non-uniqueness and rank reversals.

1 Introduction

Although most of the time decision making is almost a routine and does not require sophisticated analysis, we may occasionally be confronted with important decision-making problems. Correct decisions become important for our own welfare, future happiness and success. For instance, upon graduation from a university, we may be offered several kinds of jobs. Each job has a combination of elements including salary, advancement potential, working environment, living environment, and friendship possibilities with colleagues. This situation inevitably involves multiple criteria and the decision could be very important to our success and happiness. A careful analysis before choosing a job is therefore important.

Other examples that prevail in our daily lives might include the investment of our savings in security markets or real estate, allocation of resources to different projects, or the purchase of a house. Although this kind of decision problem does not occur every day, the impact of each important decision on the future is so strong that we must not ignore its careful analysis.

Acknowledgements

I would like to thank Róbert Fullér for his interesting lectures and advice. I would also like to thank Prof. Tamás Rapcsák[1] for his useful help in preparing the thesis.

[1]Head of Laboratory of Operations Research and Decision Systems, Computer and Automation Research Institute Hungarian Academy of Sciences.

2 The Analytic Hierarchy Process

The Analytic Hierarchy Process was developed by Thomas L. Saaty [9] in the 1970s. It is a systematic procedure for representing the elements of any problem, hierarchically.It organizes the basic rationality by breaking down a problem into its smaller and smaller constituent parts and then guides decision makers through a series of pairwise comparison judgments to express the relative strength or intensity of impact of the elements in the hierarchy. These judgments are translated to numbers. The AHP includes procedures and principles used to synthesize the many judgments to derive priorities among criteria and subsequently for alternative solutions.

Problem solving is a process of setting priorities in steps. One step decides on the most important elements of a problem, another how best to repair, replace, test and evaluate the elements, another on how to implement the solution and measure performance. The entire process is subject to revision and re-examination until one is satisfied that he/she has covered all the important features needed to represent and solve the problem. This process could be carried out in a sequence of hierarchies; i.e., by using the output of one hierarchy as the focus of concern of the next hierarchy. The AHP systematizes this process of problem solving.

Finally, if we were to assume that the unspoken feelings and experiences of people are the fundamental grounds on which an individual draws to articulate his/her creativity, then the judgments and their intensity can be used to express inner feelings and inclinations. They also enlarge the frameworks of the discourse itself by expanding the clusters and elements laid out in the hierarchy to deal with a particular problem.

This approach to system design and problem solving draws upon the innate capacity of humans to think logically and creatively, to identify events and establish relations among them. In this respect we note that people have two communicable attributes. One is their ability to impart and observe things, the other is to discriminate by establishing relations. These are the principle of *identity and decomposition*, the principle of *discrimination and comparative judgment* and the principle of *synthesis*.

2.1 Identity and Decomposition

We have used the first of the two principles identified above to structure problems in a hierarchic fashion. The principle of identity and decomposition calls for structuring problems hierarchically which is the first step one must complete when using the AHP. In its most elementary form, a hierarchy is structured from the top (objectives from a managerial standpoint,) through intermediate levels (criteria on which subsequent levels depend) to the lowest level (which is usually a list of alternatives.) The hierarchic portrayal of a problem is best illustrated by a simple example.

A family of average income decided to buy a house and identified eight criteria which they had to look for in a house. These criteria fall into three categories: economic, geographic and physical. Although one may have begun by examining the relative importance of these clusters, the family felt they wanted to prioritize the relative importance of all the factors without working with clusters. The problem was to decide of three candidate houses to choose. The first step is the *decomposition* or structuring of the problem as a hierarchy. In the first (or top) level is the overall goal of "Satisfaction with the house." In the second level are the eight factors or criteria which contribute to the goal, and in the third (or bottom) level are the three candidate houses which are to be evaluated in terms of the criteria in the second level. The definitions of the criteria and the pictorial representation of the hierarchy follow.

The criteria important to the individual family were:

(1) *Size of house*: Storage space, size of rooms, number of rooms, total area of house.
(2) *Location to bus lines*: Convenient, close bus service.
(3) *Neighborhood*: Little traffic, secure, nice view, low taxes, good condition of neighborhood.
(4) *Age of house*
(5) *Yard space*: Includes front, back and side, and space from neighbors.
(6) *Modern facilities*: Dishwashers, garbage disposals, air conditioning, alarm system, and other such items possessed by a house.
(7) *General condition*: Repairs needed; walls, carpet, drapes, cleanliness; wiring; roof; plumbing.
(8) *Financing available*: Assumable mortgage, seller financing available, or bank financing.

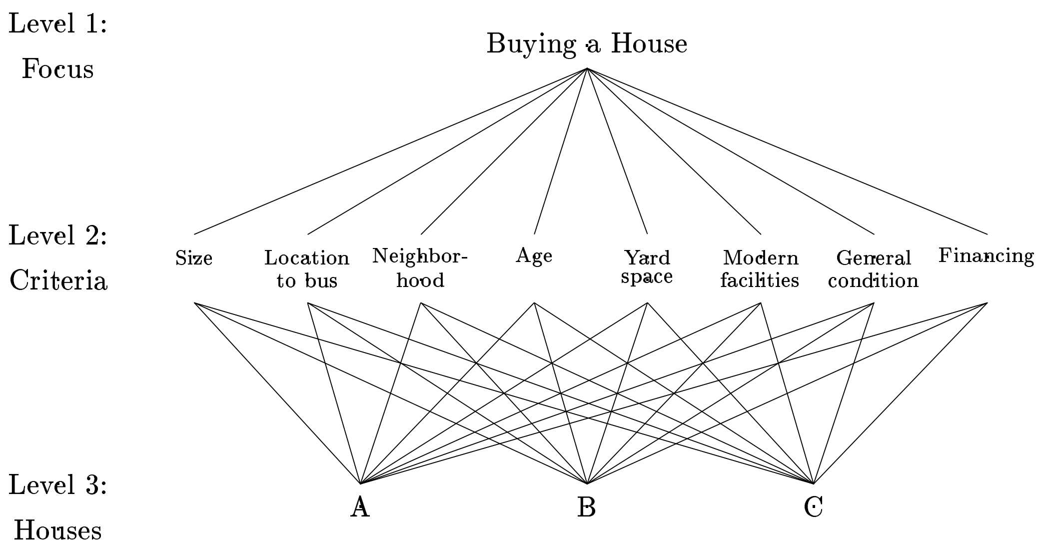


FIGURE 1

An other example with four levels of hierarchy can be found in Appendix 1.

This downward decomposition format can easily be used on a wide range of problems. The law of hierarchic continuity requires that the elements of the bottom level of the hierarchy be comparable in a pairwise fashion according to elements in the next level and so on up to the focus of the hierarchy.

For example, one must be able to provide meaningful answers to questions such as "With respect to neighborhood, what is the desirability of House A relative to House B or to House C?" or "With respect to satisfaction with the house, what is the importance of size relative to location to bus line?" and so on. The object is to derive priorities on the elements in the last level that reflect as best as possible their relative impact on the focus of the hierarchy.

It is important to note that the AHP demands that the problem be structured by the participants in the decision making process; in this simple example, the family members would devise the hierarchy in accordance with their perceived needs and their understanding of the constraints (i.e., limited funds) and opportunities (i.e., available houses) of the situation.

2.2 Discrimination and Comparative Judgments

Once a hierarchic or network representation of the problem has been achieved, we would like to go about establishing priorities among criteria and evaluating each of the alternatives on the criteria perceived to be most important.

A. Pairwise Comparisons

In the AHP, elements of a problem are compared in pairs with respect to their relative impact ("weight" or "intensity") on a property they share in common. We reduce the pairwise comparisons to a matrix form. When we compare a set of elements of a problem with each other a square matrix is produced that resembles the following:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix}$$

We observe that this matrix has reciprocal properties, that is:

$$a_{ji} = \frac{1}{a_{ij}}$$

where the subscripts i and j refer to the row and column, respectively, where any entry is located.

Now let $A_1, A_2, A_3, \dots, A_n$ be any set of elements and $w_1, w_2, w_3, \dots, w_n$ their corresponding weights or intensities. We want to compare the corresponding weights or intensities of each elements with the weights or intensities of every other element in the set with respect to a property or goal that they have in common. The comparison of weights can be represented as follows:

	A_1	A_2	A_3	$\dots$	A_n
A_1	1	$\frac{w_1}{w_2}$	$\frac{w_1}{w_3}$	$\dots$	$\frac{w_1}{w_n}$
A_2	$\frac{w_2}{w_1}$	1	$\frac{w_2}{w_3}$	$\dots$	$\frac{w_2}{w_n}$
A_3	$\frac{w_3}{w_1}$	$\frac{w_3}{w_2}$	1	$\dots$	$\frac{w_3}{w_n}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
A_n	$\frac{w_n}{w_1}$	$\frac{w_n}{w_2}$	$\frac{w_n}{w_3}$	$\dots$	1

B. The Need for a Scale of Comparison

We note that often a scale underlying a problem exists and the judgments in that case are expressed as ratios from the scale. For example, if one is comparing the relative weights of stones and one has a stone A of weight W_A and a stone B of weight W_B than the ratio $\frac{W_A}{W_B}$ is entered in the matrix for the relative weight of stone A over stone B. The reciprocal, $\frac{W_B}{W_A}$ is entered for the relative weight of stone B over stone A. In the matrix, one begins with an element on the left and asks how much more important it is than an element listed on the top. When compared with itself the ratio is 1. When compared with another element, either it is more important than that element, then an integer value, from a given scale, is used or its reciprocal in the opposite case. In either case the reciprocal ratio is entered in the position of the matrix. Thus we are always dealing with positive reciprocal matrices and need $\frac{n(n-1)}{2}$ judgments where n is the total number of elements being compared.

C. The recommended Scale of Relative Importance

Intensity of relative importance	Definition	Explanation
1	Equal importance	Two activities contribute equality to the objective
3	Moderate importance	Experience and judgment slightly favor one activity over another.
5	Essential or strong importance	Experience and judgment strongly favor one activity over another.
7	Demonstrated importance	An activity is strongly favored and its dominance is demonstrated in practice.
9	Extreme importance	The evidence favoring one activity over another is of the highest possible order of affirmation.
2,4,6,8	Intermediate values between the two adjacent judgments	When compromise is needed.
Reciprocal of above non-zero numbers	If an activity has one of the above numbers (e.g. 3) compared with a second activity, then the second activity has the reciprocal value (e.g. $\frac{1}{3}$) when compared to the first.	

D. An Illustration of Subjective Judgments Using the Scale

Table 1 represents the second level of the hierarchy. The cells of the matrix now have been filled in with the family's subjective judgments using the 1 to 9 scale and based on their preferences and perception of the constraints or opportunities of the situation. For example, when asked "With respect to overall satisfaction with the house, what is the importance of size relative to location to bus lines?" the family members agreed that size was strongly more important and therefore they entered the integer 5 in the corresponding cell; its reciprocal, or $\frac{1}{5}$ was automatically entered for the reverse comparison.

Overall satisfaction with house	Size	Location to bus	Neigh-borhood	Age	Yard space	Modern facili-ties	General condi-tion	Finan-cing
Size	1	5	3	7	6	6	$\frac{1}{3}$	$\frac{1}{4}$
Location to bus	$\frac{1}{5}$	1	$\frac{1}{3}$	5	3	3	$\frac{1}{5}$	$\frac{1}{7}$
Neighborhood	$\frac{1}{3}$	3	1	6	3	4	6	$\frac{1}{5}$
Age	$\frac{1}{7}$	$\frac{1}{5}$	$\frac{1}{6}$	1	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{7}$	$\frac{1}{8}$
Yard space	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	3	1	$\frac{1}{2}$	$\frac{1}{5}$	$\frac{1}{6}$
Modern facilities	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{4}$	4	2	1	$\frac{1}{5}$	$\frac{1}{6}$
General condition	3	5	$\frac{1}{6}$	7	5	5	1	$\frac{1}{2}$
Financing	4	7	5	8	6	6	2	1

TABLE 1

We move now on to the pairwise comparisons of the elements in the bottom level illustrated in Figure 1. The elements to be compared pairwise are the houses with respect to how much more desirable or better one is than the other in satisfying each criterion in the second level. There are eight 3×3 matrices of judgments since there are eight criteria in the second level and three houses to be compared for each criterion. Again the matrices contain the judgments of the family. In order to understand the judgments, a brief description of the houses is given below.

Size	A	B	C
A	1	6	8
B	$\frac{1}{6}$	1	4
C	$\frac{1}{8}$	$\frac{1}{4}$	1

Location to bus	A	B	C
A	1	7	$\frac{1}{5}$
B	$\frac{1}{7}$	1	$\frac{1}{8}$
C	5	8	1

Neighborhood	A	B	C
A	1	8	6
B	$\frac{1}{8}$	1	$\frac{1}{4}$
C	$\frac{1}{6}$	4	1

Age	A	B	C
A	1	1	1
B	1	1	1
C	1	1	1

Yard space	A	B	C
A	1	5	4
B	$\frac{1}{5}$	1	$\frac{1}{3}$
C	$\frac{1}{4}$	3	1

Modern facilities	A	B	C
A	1	8	6
B	$\frac{1}{8}$	1	$\frac{1}{5}$
C	$\frac{1}{6}$	5	1

General condition	A	B	C
A	1	$\frac{1}{2}$	$\frac{1}{2}$
B	2	1	1
C	2	1	1

Financing	A	B	C
A	1	$\frac{1}{7}$	$\frac{1}{5}$
B	7	1	3
C	5	$\frac{1}{3}$	1

House A. –This house is the largest of them all. It is located in a good neighborhood with little traffic and low taxes. Its yard space is comparably larger than houses B and C. However, the general condition is not very good and needs cleaning and painting. Also, the financing is unsatisfactory because it would have to be financed through a bank at a high rate of interest.

House B. –This house is a little smaller than House A and is not close to a bus route. The neighborhood gives one the feeling of insecurity because of traffic conditions. The yard space is fairly small and the house lacks the basic modern facilities. On the other hand the general condition is very good. Also an assumable mortgage is obtainable which means the financing is good with a rather low interest rate.

House C. –It is very small and has few modern facilities. The neighborhood has high taxes, but is in good condition and seems secure. The yard space is bigger than that of House B, but is not comparable to House A's spacious surroundings. The general condition of the house is good. Its financing is much better than A but not as good as B.

In the next section we show five methods for generating priorities.

3 Synthesis of Priorities

Now we have reciprocal matrices from each level of the hierarchy. We would like to generate priorities which express the relative impact of the set of elements on an element in the bottom level. To do this, at first we compute the local priorities in the level(s) above. Let us consider the case when we know all the importances (weights) of criteria, let denote them by $w_1, w_2, w_3, \ldots, w_n$. We can assume that $\sum_{i=1}^{n} w_i = 1$. Define the matrix of weight ratios as $W = [w_{ij}]_{n\times n}$:

$$\begin{pmatrix} 1 & \frac{w_1}{w_2} & \frac{w_1}{w_3} & \cdots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & 1 & \frac{w_2}{w_3} & \cdots & \frac{w_2}{w_n} \\ \frac{w_3}{w_1} & \frac{w_3}{w_2} & 1 & \cdots & \frac{w_3}{w_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \frac{w_n}{w_3} & \cdots & 1 \end{pmatrix}$$

Note that, for any i, j, k indexes

$$w_{ij} = \frac{1}{w_{ji}} \tag{1}$$

$$w_{ij} = w_{ik} w_{kj} \tag{2}$$

A matrix is called consistent if its components satisfy the $(1-2)$ equalities for any i, j, k.

3.1 Eigenvector Method (EM)

This is the original method suggested by Saaty [9]. If we take a consistent W matrix, we can observe that each row is a multiple of the first row, so the rank of W is one and thus there is only one non-zero eigenvalue. Since $w_{ii} = 1$ for $i = 1, 2, \ldots, n$ and the sum of all eigenvalues is equal to the trace of W ($tr(W) = \sum_{i=1}^{n} w_{ii}$), the eigenvalue is n. Let denote $w = (w_1, w_2, w_3, \ldots, w_n)^T$. We can easily check that $Ww = nw$:

$$\begin{pmatrix} 1 & \frac{w_1}{w_2} & \frac{w_1}{w_3} & \cdots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & 1 & \frac{w_2}{w_3} & \cdots & \frac{w_2}{w_n} \\ \frac{w_3}{w_1} & \frac{w_3}{w_2} & 1 & \cdots & \frac{w_3}{w_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \frac{w_n}{w_3} & \cdots & 1 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_n \end{pmatrix} = n \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_n \end{pmatrix},$$

therefore w is the eigenvector of W corresponding to the maximum eigenvalue n.

As a living system, human perception and judgment are subject to change when information inputs or psychological states of the decision maker change. A fixed weight vector is difficult to find.

Saaty suggested to estimate the weight ratio w_{ij} by a_{ij} and let $A = [a_{ij}]_{n\times n}$.

Note that as each $w_{ij} > 0$, we expect and shall assume that all $a_{ij} > 0$. Furthermore, as $w_{ij} = \frac{1}{w_{ji}}$, we assume that $a_{ij} = \frac{1}{a_{ji}}$, so only $a_{ij}, j > i$ need to be assessed.

Let denote λ_{max} the maximum eigenvalue of A, and $\hat{w}$ the normalized eigenvector corresponding to λ_{max} of A. Since A is found as an approximate for W, when the consistency conditions are almost satisfied for A, we would expect that $\hat{w}$ will also be close to w.

Theorem 1. *(Perron) Let M be a square matrix with positive components. Then M has a real positive simple (not multiple) eigenvalue λ_{max} which is not exceeded in modulus by any other eigenvalue of M. Moreover, the eigenvector of M corresponding to λ_{max} has positive components, and is essentially (within multiplication by a constant) unique.*

Theorem 2. *The maximum eigenvalue of A, λ_{max} is a positive real number and $\hat{w}_i > 0$ for all $i = 1, 2, \ldots, n$.*

Proof. It follows from the Perron theorem. □

Theorem 3. $\lambda_{max} \geq n$.

Proof. (Saaty, [9]) Since $\hat{w}$ is the eigenvector corresponding to λ_{max},

$$(A - \lambda_{max} I)\hat{w} = 0,$$

where the i-th equation can be written as

$$\sum_{j=1}^{n} a_{ij}\hat{w}_j - \lambda_{max}\hat{w}_i = 0.$$

Thus,

$$\lambda_{max} = \sum_{j=1}^{n} a_{ij}\frac{\hat{w}_j}{\hat{w}_i}.$$

Summing this over $i = 1, 2, \ldots, n$ we have

$$n\lambda_{max} = n + \sum_{i=1}^{n}\sum_{j=1, i\neq j}^{n} a_{ij}\frac{\hat{w}_j}{\hat{w}_i}.$$

Since

$$
\begin{aligned}
a_{ij} &= \frac{1}{a_{ji}} \\
a_{ji}\frac{\hat{w}_i}{\hat{w}_j} &= \frac{1}{a_{ij}\frac{\hat{w}_j}{\hat{w}_i}}
\end{aligned}
$$

and

$$\lambda_{max} = 1 + \frac{1}{n}\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left[a_{ij}\frac{\hat{w}_j}{\hat{w}_i} + \frac{1}{a_{ij}\frac{\hat{w}_j}{\hat{w}_i}}\right].$$

As

$$x + \frac{1}{x} = 2 + \frac{(x-1)^2}{x}, \quad \text{where } x \neq 0$$

with

$$x = a_{ij}\frac{\hat{w}_j}{\hat{w}_i}$$

we get

$$\lambda_{max} = n + \sum_{i=1}^{n}\sum_{j=i+1}^{n}\frac{(\hat{w}_j a_{ij} - \hat{w}_i)^2}{n\hat{w}_i\hat{w}_j a_{ij}}.$$

Now we see that $\lambda_{max} \geq n$ and the equality holds if and only if $a_{ij} = \frac{\hat{w}_i}{\hat{w}_j}$ for all i, j. We get that

$$\lambda_{max} = n \Leftrightarrow a_{ij} = \frac{\hat{w}_i}{\hat{w}_j} \quad \text{for all } i, j \Leftrightarrow A \text{ satisfies the consistency condition.}$$

□

In decision-making problems it may be important to know how good our consistency is, because we may not want the decision to be based on judgments that have such low consistency that they appear to be random.

Now we are in the position to define the consistency index (CI) for a given positive, reciprocal matrix of size $n \times n$,

$$CI = \frac{\lambda_{max} - n}{n - 1}$$

Since $\lambda_{max} \geq n$, CI is always non-negative.

Let us take the expected value of the consistency index, where matrices are of size $n\times\ n$, positive, reciprocal, and their elements are taken at random from the scale $\frac{1}{9},\frac{1}{8},\frac{1}{7},\ldots,\frac{1}{2},1,2,\ldots,9$. It is called mean random consistency index ($MRCI_n$).

$$MRCI_n=\frac{\bar{\lambda}_{max}-n}{n-1},$$

where $\bar{\lambda}_{max}$ denotes the expected value of λ_{max}.

Consistency ratio (CR) of an $n\times n$ matrix is defined as the ratio of its consistency index to the mean random consistency index:

$$CR=\frac{CI}{MRCI_n}.$$

If the matrix is consistent (as it was defined earlier,) then $\lambda_{max}=n$, so $CI=0$ and $CR=0$ as well. On the other hand, if the comparisons are carried out randomly, the expected value of CR is 1. Saaty suggested that a consistency ratio of about 10% or less should be usually considered acceptable.

Table 2 shows the results of computations for $MRCI_n$'s. The program written in MAPLE can be seen in Appendix 2.

n	$MRCI_n$
3	0.5245
4	0.8830
5	1.1085
6	1.2493
7	1.3405
8	1.4042
9	1.4511
10	1.4857

TABLE 2

Similar results were computed by Forman [3] and Tummala & Ling [16].

Suppose that we have all the weights of criteria and all the performances respect to each criterion. Let $c_1,c_2,\ldots,c_M$ denote the weights of criteria and p_{ij}, $(i=1,\ldots,N,\ j=1,\ldots,M)$ the performance of ith alternative on jth criterion. Now the weight of the ith alternative can be obtained as a weighted sum of performances:

$$w_i=\sum_{j=1}^{M}p_{ij}\qquad i=1,2,\ldots,N.$$

Recalling the "Buying a House" example, in Appendix 3 we have computed the weights of criteria and the performances on each criterion, so we have the following table:

	Size	Location to bus	Neigh-borhood	Age	Yard space	Modern facili-ties	General condi-tion	Finan-cing	Global priority
	0.173	0.054	0.188	0.018	0.031	0.036	0.167	0.333	
A	0.754	0.233	0.754	0.333	0.674	0.747	0.200	0.072	0.396
B	0.181	0.054	0.065	0.333	0.101	0.060	0.400	0.649	0.341
C	0.065	0.712	0.181	0.333	0.226	0.193	0.400	0.279	0.263

Since House A's priority is the largest, we will buy House A.

3.2 Least Squares Method (LSM)

$$
\begin{aligned}
min \sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_{ij}-\frac{w_i}{w_j}\right)^2 & \\
\sum_{i=1}^{n} w_i = 1 & \\
w_i > 0 \qquad & \text{for } i = 1, 2, \ldots, n.
\end{aligned}
$$

3.3 Logarithmic Least Squares Method (LLSM)

$$
\begin{aligned}
\min \sum_{i=1}^{n}\sum_{i<j}\left[\log a_{ij}-\log\left(\frac{w_i}{w_j}\right)\right]^2 & \\
\prod_{i=1}^{n} w_i = 1 & \\
w_i > 0 \qquad & \text{for } i = 1, 2, \ldots, n.
\end{aligned}
$$

Theorem 4. *The optimal solution of LLSM can be obtained from the geometric mean of the elements of rows:*

$$
w_i^{LLSM} = \sqrt[n]{\prod_{j=1}^{n} a_{ij}} \qquad i = 1, 2, \ldots, n.
$$

Proof. Let $r_{ij} = \log a_{ij}$, $i,j = 1,2,\ldots,n$ and $x_i = \log w_i$, $i = 1,2,\ldots,n$. Since $w_i = e^{x_i}$, the condition $\prod\limits_{i=1}^{n} w_i = 1$ can be written as $\sum\limits_{i=1}^{n} x_i = 0$. Now we have a problem equivalent to the original one:

$$\min \sum_{1\le i<j\le n} (r_{ij} - x_i + x_j)^2 \tag{3}$$

$$\sum_{i=1}^{n} x_i = 0 \tag{4}$$

From the condition (4) instead of x_n we can write $-\sum\limits_{l=1}^{n-1} x_l$, so we get an equivalent minimization problem with $n-1$ variables and without a condition:

$$\min f(x_1,x_2,\ldots,x_{n-1}) = \sum_{1\le i<j\le n-1} (r_{ij} - x_i + x_j)^2 + \sum_{i=1}^{n-1}\left(r_{in} - x_i - \sum_{l=1}^{n-1} x_l\right)^2$$

Derive f with respect to x_k :

$$\begin{aligned}\frac{\partial f}{\partial x_k} &= \sum_{i=1}^{k-1} 2\left(r_{ik} - x_i + x_k\right) + \sum_{i=k+1}^{n-1} 2\left(-r_{ki} - x_i + x_k\right)\\ &\quad + \sum_{i=1, i\ne k}^{n-1} 2\left(r_{ik} - x_i + x_k\right) - 4\left(r_{kn} - x_k - \sum_{l=1}^{n-1} x_l\right).\end{aligned}$$

The original A matrix is reciprocal, i.e., $a_{ij} = \frac{1}{a_{ji}}$ for all i,j, with the new variables we have $r_{ij} = -r_{ji}$, and $r_{ii} = 0$ for all i,j :

$$\frac{\partial f}{\partial x_k} = 2\left[n\left(-\sum_{l=1, l\ne k}^{n-1} x_l\right) - \sum_{i=1}^{n} r_{ki} - \sum_{i=1}^{n} r_{in}\right]. \tag{5}$$

Observe that with $x_i = \frac{\sum\limits_{j=1}^{n} r_{ij}}{n}$, $i = 1,2,\ldots,n-1$, each partial derivative of f becomes to zero, which is due to the fact that $\sum\limits_{i=1}^{n-1}\sum\limits_{j=1}^{n-1} r_{ij} = 0$. Now we have to consider the second derivative matrix of f to be sure that we got the minimum place of f. From (5)

$$\frac{\partial^2 f}{\partial^2 x_k} = 4n \qquad \text{for } k = 1,2,\ldots,n-1$$

$$\frac{\partial^2 f}{\partial x_k \partial x_l} = 2n \qquad \text{for } k \ne l,\ k,l = 1,2,\ldots,n-1.$$

Let D_N denote the determinant of the following matrix:

$$\begin{pmatrix} 4n & 2n & \dots & 2n \\ 2n & 4n & \dots & 2n \\ \vdots & \vdots & \ddots & \vdots \\ 2n & 2n & \dots & 4n \end{pmatrix}_{N\times N}$$

Subtract the first row from each row, we get

$$\begin{pmatrix} 4n & 2n & 2n & \dots & 2n \\ -2n & 2n & 0 & \dots & 0 \\ -2n & 0 & 2n & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -2n & 0 & 0 & \dots & 2n \end{pmatrix}_{N\times N}$$

It can be seen, that $D_2 = 12n^2$, and for D_N we get a recursion formula:

$$D_N = 2nD_{N-1} + (2n)^N.$$

Leastwise, D_N is positive for all $N \geq 2$, which means that the second derivative matrix of f is positive definite. We get that the optimal solution of $(3-4)$ is:

$$x_i = \frac{\sum\limits_{j=1}^{n} r_{ij}}{n} \qquad i = 1, 2, \dots, n-1.$$

Recall the original $LLSM$ problem, its optimal solution has the form:

$$w_i^{LLSM} = \sqrt[n]{\prod_{j=1}^{n} a_{ij}} \qquad i = 1, 2, \dots, n,$$

which is the geometric mean of elements of rows of A. $\square$

Now we see an advantage of $LLSM$, which is the very easy way of computing the priorities.

3.4 Chi Square Method (X^2M)

$$min \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\left(a_{ij} - \frac{w_i}{w_j}\right)^2}{\frac{w_i}{w_j}}$$

$$\sum_{i=1}^{n} w_i = 1$$

$$w_i > 0 \qquad \text{for } i = 1, 2, \dots, n.$$

3.5 Singular Value Decomposition (SVD)

Singular value decomposition is an important tool of matrix algebra that has been applied to a number of areas, for example canonical correlation in statistics, the determination of Moore-Penrose generalized inverse, and low rank approximation of matrices.

The SVD of a general matrix A is a transformation into a product of three matrices, each of which has a simple special form and geometric interpretation. This SVD representation is given by the following theorem.

Theorem 5. *Any real $m \times n$ matrix A with rank k, can be expressed in the form of*

$$A = UDV^T, \tag{6}$$

where D is a $k \times k$ diagonal matrix with positive diagonal elements $\alpha_1, \alpha_2, \dots, \alpha_k$, U is an $m \times k$ matrix and V is an $n \times k$ matrix such that $U^T U = I$, $V^T V = I$, i.e., the columns of U and V are orthonormal in the Euclidean sense [5]

An equivalent formulation of (6) in terms of diads is

$$A = \sum_{i=1}^{k} \alpha_i u_i v_i^T, \tag{7}$$

where $u_1, u_2, \dots, u_k$ and $v_1, v_2, \dots, v_k$ are the columns of U and V, respectively. The diagonal numbers α_i of D are called singular values, while the vectors u_i, $i = 1, 2, \dots, k$, are termed the left and right singular vectors, respectively. The left and right singular vectors form an orthonormal basis for the columns and rows of A in m-dimensional and n-dimensional spaces, respectively.

We can assume that SVD (6) of A is such that the singular values α_i of D are arranged so that $\alpha_1 \le \alpha_2 \le \cdots \le \alpha_k$. If strict inequalities order the singular values, then there is no multiplicity of singular values and the SVD is uniquely determined up to reflections in corresponding singular vectors. If two singular values are identical, then the corresponding pairs of singular vectors are determined only up to rotations in their respective 2-dimensional subspace. This case leads to the instability of the associated singular vectors with respect to small changes in the elements of the matrix. However, it is rare in practice that singular values are equal.

If the singular values $\alpha_{k^\star+1}, \alpha_{k^\star+2}, \dots, \alpha_k$ are small compared to $\alpha_1, \alpha_2, \dots$ $\dots, \alpha_{k^\star}$ for some $k^\star < k$, then by dropping the last $k - k^\star$ terms of the right-hand-side of (7), a good approximation of A is obtained with a $k^\star$-dimensional matrix. The theorem of low rank approximation first stated and proved by Eckart and Young [1] is as follows:

Theorem 6. *Let*

$$A_{[k^\star]} = \sum_{i=1}^{k^\star} \alpha_i u_i v_i^T,$$

be the $m \times n$ matrix of rank $k^\star$ formed from the largest $k^\star$ singular values and the corresponding singular vectors of A. Then, $A_{[k^\star]}$ is the rank $k^\star$ least squares approximation of A in that it minimizes the function

$$\sum_{i=1}^{m} \sum_{j=1}^{n} (a_{ij} - x_{ij})^2 = trace\left((A - X)(A - X)^T \right) \tag{8}$$

for all matrices X of rank $k^\star$ or less.

Now consider the Singular Value Decomposition of pairwise comparison matrices. Gass and Rapcsák [4] proved the following theorem:

Theorem 7. *The SVD of a positive, consistent matrix consist of only one diad where the right singular vector is equal to the left eigenvector multiplied by a normalizing constant, and the left singular vector to the right eigenvector multiplied by a second normalizing constant.*

Proof. The rank of a positive, consistent A matrix is equal to one, and the general form of the matrix is as follows:

$$A = \begin{pmatrix} 1 & \frac{w_1}{w_2} & \dots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & 1 & \dots & \frac{w_2}{w_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \dots & 1 \end{pmatrix},$$

where $w_i > 0, i = 1, 2, \ldots, n$. The singular value decomposition for the matrix A can be obtained by the following diad:

$$A = \frac{c_1}{c_2} \begin{pmatrix} c_2 w_1 \\ c_2 w_2 \\ \vdots \\ c_2 w_n \end{pmatrix} \left(\frac{1}{c_1 w_1}, \frac{1}{c_1 w_2}, \ldots, \frac{1}{c_1 w_n} \right), \tag{9}$$

where c_1 and c_2 are positive constants and $c_1^2 = \sum\limits_{i=1}^{n} \frac{1}{w_i^2}$ and $c_2^2 = \frac{1}{\sum\limits_{i=1}^{n} w_i^2}$.

In formula (9), $\frac{c_1}{c_2}$ is the singular value. □

The singular value decomposition seems to be numerically stable and a generally satisfactory matrix algebra method. Moreover, the value of function (8) may measure the inconsistency of a matrix. By Theorem 7, the measure of inconsistency is equal to zero in the case of a consistent matrix. This method suggest the use of the left singular vector as weights in AHP for consistent pairwise comparison matrices.

4 Critical factors

4.1 Uniqueness of solution

There is a uniqueness problem with LSM. Consider the following 3×3 matrix:

	A	B	C	LSM Solutions w^{LSM_1}	w^{LSM_2}	w^{LSM_3}	EM Solution w^{EM}
A	1	9	$\frac{1}{9}$	0.670	0.242	0.088	0.333
B	$\frac{1}{9}$	1	9	0.242	0.670	0.242	0.333
C	9	$\frac{1}{9}$	1	0.088	0.088	0.670	0.333

This matrix is degenerated because ratios heavily contradict each other. It is impossible to justify the preference between A, B, C from these values. Note that $\lambda_{max} = 10.111$, and $CR = 6.13$, which is much higher than the 0.10 limit. This high inconsistency is not likely to be encountered in practice.

However, LSM non-uniqueness may arise even when there is no degeneracy, i.e., when the EM's eigenvector components are unequal suggesting a basis for differential row importance weighting. Let us see the following example:

	A	B	C	w^{LSM_1}	w^{LSM_2}	w^{EM}
A	1	9	3	0.740	0.355	0.692
B	$\frac{1}{9}$	1	3	0.100	0.568	0.231
C	$\frac{1}{3}$	$\frac{1}{3}$	1	0.169	0.077	0.077

The LSM in this case presents a dilemma. Now $\lambda_{max} = 4.333$ and $CR = 1.149$, which is still high but not extremely high. A is the best by w^{EM} and w^{LSM_1}, but B is the best by w^{LSM_2}. Moreover, the order by w^{EM} is A, B, C, while w^{LSM_1} gives A, C, B. We got three different orders.

4.2 Rank Reversals

In practice we are dealing with reasonable inconsistency (e.g., $CR \leq 0.10$,) when each method, including LSM, gives unique solution. However, the final order is not always the same using different methods. Consider the following example:

	A	B	C	w^{EM}	w^{X^2M}	w^{LSM}
A	1	2	7	0.559	0.498	0.427
B	$\frac{1}{2}$	1	9	0.383	0.444	0.514
C	$\frac{1}{7}$	$\frac{1}{9}$	1	0.058	0.058	0.059

Now $\lambda_{max} = 3.1$ and $CR = 0.086$, which must be considered very consistent. After all LSM yields different order.

As the following theorems state, if an alternative is at least as good as an other one compared to each alternative, then using any of EM,LSM, $LLSM$ or X^2M, its final preference can not be worse than the other one's.

Theorem 8. *If* $a_{ik} \geq a_{jk}$ *for* $k = 1, 2, \ldots, n$, *then* $w_i^{EM} \geq w_j^{EM}$.

Proof. (Saaty, [9]) Since w^{EM} is the eigenvector corresponding to λ_{max},

$$(A - \lambda_{max} I) w^{EM} = 0,$$

where the i-th equation can be written as

$$\sum_{k=1}^{n} a_{ik} w_k^{EM} - \lambda_{max} w_i^{EM} = 0.$$

Thus,

$$w_i^{EM} = \frac{1}{\lambda_{max}} \sum_{k=1}^{n} a_{ik} w_k^{EM} \geq \frac{1}{\lambda_{max}} \sum_{k=1}^{n} a_{jk} w_k^{EM} = w_j^{EM}.$$

□

Theorem 9. *If* $a_{ik} \geq a_{jk}$ *for* $k = 1, 2, \ldots, n$, *then* $w_i^{LSM} \geq w_j^{LSM}$.

Proof. (Jensen, [6]) Let $w = (w_1, w_2, \ldots, w_n)$ be a solution vector. Denote

$$ESS = \sum_{i=1}^{n} \sum_{j=1}^{n} \left(a_{ij} - \frac{w_i}{w_j} \right)^2.$$

Let i, j be fixed and

$$E_{ij} = \left(a_{ij} - \frac{w_i}{w_j} \right)^2 + \left(a_{ji} - \frac{w_j}{w_i} \right)^2 + \sum_{k \neq i,j} \left[\left(a_{ik} - \frac{w_i}{w_k} \right)^2 + \left(a_{jk} - \frac{w_j}{w_k} \right)^2 \left(a_{ki} - \frac{w_k}{w_i} \right)^2 + \left(a_{kj} - \frac{w_k}{w_j} \right)^2 \right].$$

Thus

$$ESS = E_{ij} + \sum_{\{k,l\}\cap\{i,j\}=\varnothing}^{n} \left(a_{kl} - \frac{w_k}{w_l}\right)^2.$$

Elements of E_{ij} have a subscript i or j, all other elements are in the second component. If we transpose w_i and w_j in w, we get w'. Denote

$$\begin{aligned}
E'_{ij} &= \left(a_{ij} - \frac{w_j}{w_i}\right)^2 + \left(a_{ji} - \frac{w_i}{w_j}\right)^2 \\
&+ \sum_{k\neq i,j} \left[\left(a_{ik} - \frac{w_j}{w_k}\right)^2 + \left(a_{jk} - \frac{w_i}{w_k}\right)^2 + \left(a_{ki} - \frac{w_k}{w_j}\right)^2 + \left(a_{kj} - \frac{w_k}{w_i}\right)^2\right].
\end{aligned}$$

and

$$ESS' = E'_{ij} + \sum_{\{k,l\}\cap\{i,j\}=\varnothing}^{n} \left(a_{kl} - \frac{w_k}{w_l}\right)^2.$$

Let ΔE_{ij} denote the change in ESS resulting from transposing w_i and w_j in w.

$$\Delta E_{ij} = ESS - ESS' = E_{ij} - E'_{ij},$$

because the second components are the same.

$$\begin{aligned}
&\Delta E_{ij} = E_{ij} - E'_{ij} \\
&= -2a_{ij}\left(\frac{w_i}{w_j} - \frac{w_j}{w_i}\right) - 2a_{ji}\left(\frac{w_j}{w_i} - \frac{w_i}{w_j}\right) \\
&\quad + 2\sum_{k\neq i,j}\left[-a_{ik}\left(\frac{w_i}{w_k} - \frac{w_j}{w_k}\right) - a_{jk}\left(\frac{w_j}{w_k} - \frac{w_i}{w_i}\right)\right] \\
&\quad + 2\sum_{k\neq i,j}\left[-a_{ki}\left(\frac{w_k}{w_i} - \frac{w_k}{w_j}\right) - a_{kj}\left(\frac{w_k}{w_j} - \frac{w_k}{w_i}\right)\right] \\
&= 2\left[(a_{ij} - a_{ji})\left(\frac{w_j}{w_i} - \frac{w_i}{w_j}\right)\right] \\
&\quad + 2\sum_{k\neq i,j}\left[(a_{ik} - a_{jk})\left(\frac{w_j}{w_k} - \frac{w_i}{w_k}\right) + (a_{kj} - a_{ki})\left(\frac{w_k}{w_i} - \frac{w_k}{w_j}\right)\right].
\end{aligned}$$

Suppose that $w_i < w_j$. Then

$$(a_{ij} - a_{ji})\left(\frac{w_j}{w_i} - \frac{w_i}{w_j}\right) \le 0,$$
$$(a_{ik} - a_{jk})\left(\frac{w_j}{w_k} - \frac{w_i}{w_k}\right) \le 0,$$
$$(a_{kj} - a_{ki})\left(\frac{w_k}{w_i} - \frac{w_k}{w_j}\right) \le 0.$$

We get that
$\Delta E_{ij} = 0$ if and only if $a_{ik} = a_{jk}$ for all $k = 1, 2, \ldots, n$, so Row i and Row j have the same elements. In this case no error change can be achieved by transposing w_i and w_j.
$\Delta E_{ij} > 0$ if and only if $a_{ik} > a_{jk}$ for at least one k among $k = 1, 2, \ldots, n$. In other words, if Row i strongly dominates Row j, then any solution in which $w_i < w_j$ cannot be LSM optimal, because a reduction in ESS can be achieved by transposing w_i and w_j in the solution vector. □

Theorem 10. *If* $a_{ik} \ge a_{jk}$ *for* $k = 1, 2, \ldots, n$, *then* $w_i^{LLSM} \ge w_j^{LLSM}$.

Proof. Proof of this entails only simple modification of the LSM proof. We should just replace a_{ij} and w_i symbols by $\log a_{ij}$ and $\log w_i$ counterparts. □

Theorem 11. *If* $a_{ik} \ge a_{jk}$ *for* $k = 1, 2, \ldots, n$, *then* $w_i^{X^2M} \ge w_j^{X^2M}$.

Proof. (Jensen, [6]) Let $w = (w_1, w_2, \ldots, w_n)$ be a solution vector. Denote

$$X^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\left(a_{ij} - \frac{w_i}{w_j}\right)^2}{\frac{w_i}{w_j}} = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_{ij} - \frac{w_i}{w_j}\right)^2 \frac{w_j}{w_i}.$$

Let i, j be fixed and

$$\begin{aligned} X_{ij}^2 &= \left(a_{ij} - \frac{w_i}{w_j}\right)^2 \frac{w_j}{w_i} + \left(a_{ji} - \frac{w_j}{w_i}\right)^2 \frac{w_i}{w_j} \\ &+ \sum_{k \ne i,j}\left[\left(a_{ik} - \frac{w_i}{w_k}\right)^2 \frac{w_k}{w_i} + \left(a_{jk} - \frac{w_j}{w_k}\right)^2 \frac{w_k}{w_j}\right] \\ &+ \sum_{k \ne i,j}\left[\left(a_{ki} - \frac{w_k}{w_i}\right)^2 \frac{w_i}{w_k} + \left(a_{kj} - \frac{w_k}{w_j}\right)^2 \frac{w_j}{w_k}\right]. \end{aligned}$$

Thus

$$X^2 = X_{ij}^2 + \sum_{\{k,l\}\cap\{i,j\}=\varnothing}^{n} \left(a_{kl} - \frac{w_l}{w_l}\right)^2 \frac{w_l}{w_k}.$$

Elements of E_{ij} have a subscript i or j, all other elements are in the second component. If we transpose w_i and w_j in w, we get w'. Denote

$$\begin{aligned}
{X_{ij}^2}' &= \left(a_{ij} - \frac{w_j}{w_i}\right)^2 \frac{w_i}{w_j} + \left(a_{ji} - \frac{w_i}{w_j}\right)^2 \frac{w_j}{w_i} \\
&+ \sum_{k\neq i,j} \left[\left(a_{ik} - \frac{w_j}{w_k}\right)^2 \frac{w_k}{w_j} + \left(a_{jk} - \frac{w_i}{w_k}\right)^2 \frac{w_k}{w_i}\right] \\
&+ \sum_{k\neq i,j} \left[\left(a_{ki} - \frac{w_k}{w_j}\right)^2 \frac{w_j}{w_k} + \left(a_{kj} - \frac{w_k}{w_i}\right)^2 \frac{w_i}{w_k}\right].
\end{aligned}$$

and

$$X^{2'} = {X_{ij}^2}' + \sum_{\{k,l\}\cap\{i,j\}=\varnothing}^{n} \left(a_{kl} - \frac{w_l}{w_l}\right)^2 \frac{w_l}{w_k}.$$

Let ΔX_{ij}^2 denote the change in X^2 resulting from transposing w_i and w_j in w.

$$\Delta X_{ij}^2 = X^2 - X^{2'} = X_{ij}^2 - {X_{ij}^2}',$$

because the second components are the same.

$$\begin{aligned}
\Delta X_{ij}^2 &= X_{ij}^2 - {X_{ij}^2}' \\
&= a_{ij}^2 \left(\frac{w_j}{w_i} - \frac{w_i}{w_j}\right) + a_{ji}^2 \left(\frac{w_i}{w_j} - \frac{w_j}{w_i}\right) \\
&+ \sum_{k\neq i,j} \left[a_{ik}^2 \left(\frac{w_k}{w_i} - \frac{w_k}{w_j}\right) + a_{jk}^2 \left(\frac{w_k}{w_j} - \frac{w_k}{w_i}\right)\right] \\
&+ \sum_{k\neq i,j} \left[a_{ki}^2 \left(\frac{w_i}{w_k} - \frac{w_j}{w_k}\right) + a_{kj}^2 \left(\frac{w_j}{w_k} - \frac{w_i}{w_k}\right)\right] \\
&= \left(a_{ij}^2 - a_{ji}^2\right) \left(\frac{w_j}{w_i} - \frac{w_i}{w_j}\right) \\
&+ \sum_{k\neq i,j} \left[\left(a_{ik}^2 - a_{jk}^2\right) \left(\frac{w_k}{w_i} - \frac{w_k}{w_j}\right) + \left(a_{kj}^2 - a_{ki}^2\right) \left(\frac{w_j}{w_k} - \frac{w_i}{w_k}\right)\right].
\end{aligned}$$

Suppose that $w_i < w_j$. Then

$$\left(a_{ij}^2 - a_{jk}^2\right)\left(\frac{w_j}{w_i} - \frac{w_i}{w_j}\right) \leq 0,$$
$$\left(a_{ik}^2 - a_{jk}^2\right)\left(\frac{w_k}{w_j} - \frac{w_k}{w_i}\right) \leq 0,$$
$$\left(a_{kj}^2 - a_{ki}^2\right)\left(\frac{w_j}{w_k} - \frac{w_i}{w_k}\right) \leq 0.$$

We get that
$\Delta X_{ij}^2 = 0$ if and only if $a_{ik} = a_{jk}$ for all $k = 1, 2, \ldots, n$, so Row i and Row j have the same elements. In this case no error change can be achieved by transposing w_i and w_j.
$\Delta X_{ij}^2 > 0$ if and only if $a_{ik} > a_{jk}$ for at least one k among $k = 1, 2, \ldots, n$.
In other words, if Row i strongly dominates Row j, then any solution in which $w_i < w_j$ cannot be X^2M optimal, because X^2 can be reduced by transposing w_i and w_j in the solution vector. □

4.3 Sensitivity Analysis

Consider a decision problem with M alternatives and N criteria. Alternatives are denoted as A_i $(i = 1, 2, \ldots, M)$ and criteria as C_j $(j = 1, 2, \ldots, N)$. We assume that for each criterion C_j the decision maker has determined its importance, or weight, w_j. It is also assumed that the decision maker has determined a_{ij} $(i = 1, 2, \ldots, M, \quad j = 1, 2, \ldots, N)$; the measure of performance of alternative A_i in terms of criterion C_j. For example, using the Eigenvector Method, the vector $(w_1, w_2, \ldots, w_N)$ is the principal eigenvector of the $N \times N$ reciprocal matrix determined by pairwise comparisons of the criteria. In a similar way, the vector $(a_{1j}, a_{2j}, \ldots, a_{Mj})$ for each $j = 1, 2, \ldots, N$ is the principal eigenvector of an $M \times M$ reciprocal matrix which is determined by pairwise comparisons of the impact of the M alternatives on the ith criterion. So we can assume that these vectors are normalized:

$$\sum_{j=1}^{N} w_j = 1 \tag{10}$$
$$\sum_{i=1}^{M} a_{ij} = 1 \qquad \text{for } j = 1, 2, \ldots, N.$$

Let P_i $(i = 1, 2, \ldots, M)$ represent the final preference of alternative A_i, which is calculated according to the weighted sum principle:

$$P_i = \sum_{j=1}^{N} a_{ij} w_j \qquad \text{for } i = 1, 2, \ldots, M. \tag{11}$$

Without loss of generality, it can be assumed (by a simple rearrangement of the indexes) that the M alternatives are arranged in the following way:

$$P_1 \geq P_2 \geq \cdots \geq P_M.$$

4.3.1 Changing the weights of criteria

Definition 1. Let δ_{st}^k $(1 \leq s < t \leq M$ and $1 \leq k \leq N)$ denote the minimum change in the current weight w_k of criterion C_k such that the ranking of alternatives A_s and A_t will be reversed. Also, define as:

$$\delta_{st}^{k\,'} = \frac{\delta_{st}^k}{w_k} \cdot 100 \qquad (1 \leq s < t \leq M \text{ and } 1 \leq k \leq N),$$

which expresses changes in relative terms in percent.

Theorem 12. *(Triantaphyllou and Sánchez, [15]) The quantity* $\delta_{st}^{k\,'}$ $(1 \leq s < t \leq M$ *and* $1 \leq k \leq N)$, *by which the current weight* w_k *of criterion* C_k *needs to be modified (after normalization) so that the ranking of the alternatives* A_s *and* A_t *will be reversed, is given as follows:*

$$\begin{aligned} \delta_{st}^{k\,'} &< \frac{P_t - P_s}{a_{tk} - a_{sk}} \cdot \frac{100}{w_k} \qquad \textit{if } a_{tk} > a_{sk} \\ \delta_{st}^{k\,'} &> \frac{P_t - P_s}{a_{tk} - a_{sk}} \cdot \frac{100}{w_k} \qquad \textit{if } a_{tk} < a_{sk}. \end{aligned}$$

Proof. Let $w_k^\star$ denote the modified weight of criterion C_k :

$$w_k^\star = w_k - \delta_{st}^k.$$

To preserve property (10), it is necessary that all weights be normalized:

$$w_i{}' = \frac{w_i}{w_k^\star + \sum_{j \neq k} w_j} \qquad i = 1, 2, \ldots, k-1, k+1, \ldots, N \tag{12}$$

$$w_k{}' = \frac{w_k^\star}{w_k^\star + \sum_{j \neq k} w_j} \tag{13}$$

Given the new weights $w_i^{'}$ $(i = 1, 2, \ldots, N)$, let $P_s^{'}$ and $P_t^{'}$ denote the new final preference values for the two alternatives A_s and A_t, respectively. Since it is desired that the new ranking of these two alternatives be reversed, therefore:

$$P_s^{'} < P_t^{'}.$$

Using the formula (11) we get

$$\sum_{j=1}^{N} w_j^{'} a_{sj} = P_s^{'} < P_t^{'} = \sum_{j=1}^{N} w_j^{'} a_{tj}$$

Using $(12 - 13)$ we can write

$$\begin{aligned}
\frac{w_k^{\star} a_{sk}}{w_k^{\star} + \sum\limits_{j\neq k} w_j} + \frac{\sum\limits_{j\neq k} w_j a_{sj}}{w_k^{\star} + \sum\limits_{j\neq k} w_j} \quad &< \quad \frac{w_k^{\star} a_{tk}}{w_k^{\star} + \sum\limits_{j\neq k} w_j} + \frac{\sum\limits_{j\neq k} w_j a_{tj}}{w_k^{\star} + \sum\limits_{j\neq k} w_j} \\
w_k^{\star} a_{sk} + \sum_{j\neq k} w_j a_{sj} \quad &< \quad w_k^{\star} a_{tk} + \sum_{j\neq k} w_j a_{tj} \qquad (14)
\end{aligned}$$

Now from (10) and (14) we get

$$\begin{aligned}
-\delta_{st}^{k} a_{sk} + \underbrace{\sum_{j=1}^{N} w_j a_{sj}}_{P_s} \quad &< \quad -\delta_{st}^{k} a_{tk} + \underbrace{\sum_{j=1}^{N} w_j a_{tj}}_{P_t} \\
\delta_{st}^{k} (a_{tk} - a_{sk}) \quad &< \quad P_t - P_s
\end{aligned}$$

Dividing by $(a_{tk} - a_{sk})$,

$$\begin{aligned}
\delta_{st}^{k} \quad &< \quad \frac{P_t - P_s}{a_{tk} - a_{sk}} \qquad \text{if } a_{tk} > a_{sk} \text{ or} \\
\delta_{st}^{k} \quad &> \quad \frac{P_t - P_s}{a_{tk} - a_{sk}} \qquad \text{if } a_{tk} < a_{sk}
\end{aligned}$$

Multiplying by $\frac{100}{w_k}$, we get the theorem. $\square$

4.3.2 Changing performances

Definition 2. Let Δ_{st}^{k} $(1 \leq t < s \leq M$ and $1 \leq k \leq N)$ denote the threshold value of a_{sk}, which is the minimum change which has to occur on the current value of a_{sk} such that the ranking of alternatives A_s and A_t will be reversed. Also, define as:

$$\Delta_{st}^{k\,'} = \frac{\Delta_{st}^{k}}{a_{sk}} \cdot 100 \qquad (1 \leq t < s \leq M \text{ and } 1 \leq k \leq N),$$

which expresses changes in relative terms in percent.

Theorem 13. *(Triantaphyllou and Sánchez, [15]) The threshold value* $\Delta_{st}^{k'}$ *by which the measure of performance of alternative* A_s *in terms of criterion* C_k *needs to be modified so that the ranking of alternatives* A_s *and* A_t *will change, is given as follows:*

$$\Delta_{st}^{k'} = \frac{P_s - P_t}{P_s - P_t + +w_k\left(a_{tk} - a_{sk} + 1\right)} \cdot \frac{100}{a_{sk}}$$

Furthermore, the following condition should also be satisfied for the threshold value to be feasible:

$$\Delta_{st}^{k'} \leq 100.$$

Proof. Let $a_{sk}^{\star}$ denote the modified a_{sk} measure of performance:

$$a_{sk}^{\star} = a_{sk} - \Delta_{st}^{k}. \tag{15}$$

Suppose that before modification $P_t \geq P_s$, after changing a_{sk} to $a_{sk}^{\star}$ we have $P_t^{'} < P_s^{'}$. We also have to normalize the new $a_{sk}^{\star}$ values:

$$a_{sk}^{'} = \frac{a_{sk}^{\star}}{a_{sk}^{\star} + \sum\limits_{i \neq s} a_{ik}} \tag{16}$$

$$a_{lk}^{'} = \frac{a_{lk}}{a_{sk}^{\star} + \sum\limits_{i \neq s} a_{ik}} \qquad \text{for } l = 1, 2, \ldots, s-1, s+1, \ldots, M \tag{17}$$

Since $\sum\limits_{i=1}^{M} a_{ik} = 1$, and (15), we can write:

$$a_{sk}^{\star} + \sum_{i \neq s} a_{ik} = \sum_{i=1}^{M} a_{ik} - \Delta_{st}^{k'} = 1 - \Delta_{st}^{k'}.$$

We can simplify the equations $(16 - 17)$:

$$a_{sk}^{'} = \frac{a_{sk}^{\star}}{1 - \Delta_{st}^{k'}} = \frac{a_{sk} - \Delta_{st}^{k'}}{1 - \Delta_{st}^{k'}} \tag{18}$$

$$a_{lk}^{'} = \frac{a_{lk}}{1 - \Delta_{st}^{k'}} \qquad \text{for } l = 1, 2, \ldots, s-1, s+1, \ldots, M \tag{19}$$

Because of $P_t^{'} < P_s^{'}$:

$$\begin{aligned}
a_{tk}^{'}w_k + \sum_{j\neq k} a_{tj}w_j = P_t^{'} &< P_s^{'} = a_{sk}^{'}w_k + \sum_{j\neq k} a_{sj}w_j \\
a_{tk}^{'}w_k + (a_{tk} - a_{tk})w_k + \sum_{j\neq k} a_{tj}w_j &< a_{sk}^{'}w_k + (a_{sk} - a_{sk})w_k + \sum_{j\neq k} a_{sj}w_j \\
a_{tk}^{'}w_k - a_{tk}w_k + \underbrace{\sum_{j=1}^{N} a_{tj}w_j}_{P_t} &< a_{sk}^{'}w_k - a_{sk}w_k + \underbrace{\sum_{j=1}^{N} a_{sj}w_j}_{P_s} \\
a_{tk}^{'}w_k - a_{tk}w_k + P_t &< a_{sk}^{'}w_k - a_{sk}w_k + P_s. \qquad (20)
\end{aligned}$$

When we substitute $(18-19)$ on (20) we get:

$$\frac{a_{tk}w_k}{1-\Delta_{st}^k} - a_{tk}w_k + P_t < \frac{(a_{sk}-\Delta_{st}^k)w_k}{1-\Delta_{st}^k} - a_{sk}w_k + P_s,$$

which can be further reduced to:

$$\Delta_{st}^k < \frac{P_s - P_t}{P_s - P_t + w_k\left(a_{tk} - a_{sk} + 1\right)}. \qquad (21)$$

Note that the denominator on the right-hand-side is always positive, because of (20). Furthermore, the following condition should also be satisfied for the new $a_{sk}^{'}$ value to have a feasible meaning:

$$\begin{aligned}
0 \leq a_{sk}^{'} &\leq 1 \qquad \text{or} \\
\Delta_{st}^k &\leq a_{sk}. \qquad (22)
\end{aligned}$$

Multiplying (21) and (22) by $\frac{100}{a_{sk}}$, we get the statement of the theorem. □

5 Group Decision Making

5.1 Aggregating individual judgments

Let the function $f(x_1, x_2, \ldots, x_n)$ for synthesizing the judgments given by n judges.

(i) Separability condition (S): $f(x_1, x_2, \ldots, x_n) = g(x_1)g(x_2)\ldots g(x_n)$ for all $x_1, x_2, \ldots, x_n$ in an interval I of positive numbers, where g is a function mapping I onto a proper interval J and is a continuous, associative and cancellative operation. *(S)* means that the influences of the individual judgments can be separated as above.

(ii) Unanimity condition (U)*:* $f(x, x, \ldots, x) = x$ for all x in I. (U) means that if all individuals give the same judgment x, that judgment should also be the synthesized judgment.

(iii) Homogeneity condition (H)*:* $f(ux_1, ux_2, \ldots, ux_n) = uf(x_1, x_2, \ldots, x_n)$, where $u > 0$ and $x_k, ux_k, (k = 1, 2, \ldots, n)$ are all in I. (H) means that if all individuals judge a ratio u times as large as another ratio, then the synthesized judgment should also be u times as large.

(iv) Power condition (P_p)*:* $f(x_1^p, x_2^p, \ldots, x_n^p) = f^p(x_1, x_2, \ldots, x_n)$. (P_2) for example means that if the kth individual judges the length of a side of a square to be x_k, the synthesized judgment on the area of that square will be given by the square of the synthesized judgment on the length of its side.

Special case $(R = P_{-1})$*:*

$$f(\frac{1}{x_1}, \frac{1}{x_2}, \ldots, \frac{1}{x_n}) = \frac{1}{f(x_1, x_2, \ldots, x_n)}.$$

(R) is of particular importance in ratio judgments. It means that the synthesized value of the reciprocal of the individual judgments should be the reciprocal of the synthesized value of the original judgments.

Aczél and Saaty (see [13],[14]) proved the following theorem:

Theorem 14. *The general separable* (S) *synthesizing functions satisfying the unanimity* (U) *and homogeneity* (H) *conditions are*
the geometric mean: $f(x_1, x_2, \ldots, x_n) = \sqrt[n]{x_1 x_2 \cdots x_n}$, *and*
the root-mean-power: $f(x_1, x_2, \ldots, x_n) = \sqrt{\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}}$.
If moreover the reciprocal property (R) *is assumed even for a single* n*-tuple* $(x_1, x_2, \ldots, x_n)$ *of the judgments of* n *individuals, where not all* x_k *are equal, then only the geometric mean satisfies all the above conditions.*

In any rational consensus, those know more, should, accordingly, influence the consensus more strongly than those are less knowledgeable. Some people are clearly wiser and more sensible in such matters than others, others may be more powerful and their opinions should be given appropriately greater weight. For such unequal importance of voters not all g's in (S) are the same function. In place of (S), the *Weighted Separability condition* (WS) is now:

$$f(x_1, x_2, \ldots, x_n) = g_1(x_1)g_2(x_2)\ldots g_n(x_n).$$

(WS) implies that not all judging individuals have the same weight when the judgments are synthesized and the different influences are reflected in the different functions $(g_1, g_2, \ldots, g_n)$.

Aczél and Alsina (see [14]) proved the following theorem:

Theorem 15. *The general weighted-separable* (WS) *synthesizing functions with the unanimity* (U) *and homogeneity* (H) *properties are the weighted geometric mean:* $f(x_1, x_2, \ldots, x_n) = x_1^{q_1}x_2^{q_2}\cdots x_n^{q_n}$, *and the weighted root-mean-powers:*

$$f(x_1, x_2, \ldots, x_n) = \sqrt[\gamma]{q_1x_1^\gamma + q_2x_2^\gamma + \cdots + q_nx_n^\gamma},$$

where $q_1 + q_2 + \cdots + q_n = 1, \quad q_k > 0 \quad (k = 1, 2, \ldots, n)$ *and* $\gamma > 0$, *but otherwise* $q_1, q_2, \ldots, q_n, \gamma$ *are arbitrary constants.*
If f *also has the reciprocal property* (R) *and for a single* n*-tuple* $(x_1, x_2, \ldots, x_n)$ *of the judgments of* n *individuals, where not all* x_k *are equal, then only the weighted geometric mean applies.*

We give the following theorem which is an explicit statement of the synthesis problem that follows from the previous results (Saaty, [12]):

Theorem 16. *If* $x_1^{(i)}, x_2^{(i)}, \ldots, x_n^{(i)} \quad i = 1, \ldots, m$ *are rankings of* n *alternatives by* m *independent judges and if* a_i *is the importance of* i*-th judge developed from a hierarchy for evaluating the judges, and hence* $\sum_{i=1}^{m} a_i = 1$, *then*

$$\left(\prod_{i=1}^{m} x_1^{a_i}\right), \left(\prod_{i=1}^{m} x_2^{a_i}\right), \ldots, \left(\prod_{i=1}^{m} x_n^{a_i}\right)$$

are the combined ranks of the alternatives for the m *judges.*

5.2 Constructing a group choice

Given a group of individuals, a set of alternatives (with cardinality greater than 2,) and individual ordinal preferences for the alternatives. Arrow proved with his Impossibility Theorem that it is impossible to derive a rational group choice (construct a social choice function that aggregates individual preferences) from ordinal preferences of the individuals that satisfy the following four conditions, i.e., at least one of them is violated:

Decisiveness: the aggregation procedure must produce a group order.

Unanimity : if all individuals prefer alternative A to alternative B, then the aggregation procedure must produce a group order indicating that the group prefers A to B.

Independence of irrelevant alternatives : given two sets of alternatives which both include A and B, if all individuals prefer A to B in both sets, then the aggregation procedure must produce a group order indicating that the group, given any of the two sets of alternatives, prefers A to B.

No dictator : no single individual preferences determine the group order.

5.3 Group Decision based on Bridgman principle

This method is for the group decision phase that consist of the aggregation of the individual expert weight vectors. In our case, the decision problem is to aggregate the weight vectors of the m experts so that the voting powers of the decision makers can be taken into account. Let $w_{ij} > 0$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$ denote the weight of alternative A_j given by the ith expert; $v_i > 0$, $i = 1, 2, \ldots, n$ the jth expert's voting power; and x_j, $j = 1, 2, \ldots, n$ the unknown components of the synthesized weight vector. The data of this decision problem can be written in tabular form:

$$
\begin{array}{cc|cccc}
 & & x_1 & x_2 & \ldots & x_n \\
 & & A_1 & A_2 & \ldots & A_n \\
\hline
v_1 & E_1 & w_{11} & w_{12} & \ldots & w_{1n} \\
v_2 & E_2 & w_{21} & w_{22} & \ldots & w_{2n} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
v_m & E_m & w_{m1} & w_{m2} & \ldots & w_{mn}
\end{array},
$$

where $E_1, E_2, \ldots, E_m$ denote the experts.
Let $a_i = (w_{i1}, w_{i2}, \ldots, w_{in})^T$, $i = 1, 2, \ldots, m$, and $v = \sum_{l=1}^{m} v_l$.

Definition 3. The Hölder-Young distance of vector $u = (u_1, u_2, \ldots, u_k)$ and vector $z = (z_1, z_2, \ldots, z_k)$ for a given parameter α is given in the form of

$$H_\alpha (u \parallel z) = \frac{1}{\alpha (1-\alpha)} \sum_{j=1}^{k} \alpha u_j + (1-\alpha) z_j - u_j^\alpha z_j^{1-\alpha}, \quad \alpha \in \mathbb{R} \setminus \{0,1\},$$

and

$$H_0 (u \parallel z) = \lim_{\alpha \to 0} H_\alpha (u \parallel z), \quad H_1 (u \parallel z) = \lim_{\alpha \to 1} H_\alpha (u \parallel z)$$

The decision principle is to minimize the weighted sum of one of the Hölder-Young distances of the given weight vectors and the unknown vector $x = (x_1, x_2, \ldots, x_n)^T$:

$$\begin{aligned} & min \frac{\sum\limits_{i=1}^{m} v_i H_\alpha (a_i \parallel x)}{v} \\ & x_j \geq 0, \quad j = 1, 2, \ldots, n. \end{aligned} \tag{23}$$

The following concept is also used as the objective function in decision models.

Definition 4. The generalized Kullback's I-divergence of vector $u = (u_1, u_2, \ldots, u_k) \in \mathbb{R}_+^k$ and vector $z = (z_1, z_2, \ldots, z_k) \in \mathbb{R}_+^k$ is

$$D (u \parallel z) = \sum_{j=1}^{k} u_j \log \left(\frac{u_j}{z_j} \right) - \sum_{j=1}^{k} u_j + \sum_{j=1}^{k} z_j .$$

We note that the case of Kullback's I-divergence can be obtained from the Hölder-Young distance with $\alpha = 1$, and if we consider the (23) minimization problem with $\alpha = 1$, then the optimal solution has an explicit form that yields the weighted geometric mean for aggregating the weight vectors:

$$x_j^{opt} = \prod_{i=1}^{m} w_{ij}^{\frac{v_i}{v}}, \quad j = 1, 2, \ldots, n.$$

6 Further observations

6.1 A generalization of the discrete case

The relative ratio scale derived from a pairwise comparison reciprocal matrix of judgments is derived by solving:

$$\sum_{j=1}^{n} a_{ij} w_j = \lambda_{max} w_i \qquad \text{for } i = 1, 2, \ldots, n \tag{24}$$

$$\sum_{j=1}^{n} w_j = 1 \tag{25}$$

where $a_{ij} = \frac{1}{a_{ji}}$, $a_{ij} > 0$. The matrix $A = (a_{ij})$ is said to be consistent, if $a_{ij}a_{jk} = a_{ik}$ for all i, j, k. The discrete formulation given in (24) and (25) above may be generalized to the continuous case through Fredholm's integral equation of the second kind and is given by:

$$\lambda \int_a^b K(s,t) w(t) dt = w(s) \tag{26}$$

$$\int_a^b w(s) = 1$$

where instead of matrix A we have as a positive kernel, $K(s,t) > 0$ and $\lambda = \frac{1}{\lambda_{max}}$. Note that the entries in a matrix depend on the two variables i and j which assume discrete values. Thus the matrix itself depends on these discrete variables, and its generalization, the kernel function also depends on two (continuous) variables. Here also we have the reciprocal property (27), and as in the finite case, the kernel $K(s,t)$ is consistent if it satisfies the relation (28) :

$$K(s,t) = \frac{1}{K(t,s)} \tag{27}$$

$$K(s,t)K(t,u) = K(s,u) \qquad \text{for all } s, t, u. \tag{28}$$

It follows by putting $s = t = u$, that $K(s,s) = 1$ for all s, which is analogous to having ones down the diagonal of the matrix in the discrete case.
An example of this type of kernel is $K(s,t) = \frac{e^s}{e^t} = e^{s-t}$.
A value of λ for which Fredholm's equation has a nonzero solution $w(t)$, is called a characteristic value (or its reciprocal is called an eigenvalue) and

the corresponding solution is called an eigenfunction. An eigenfunction is determined to within a multiplicative constant. If $w(t)$ is an eigenfunction corresponding to the characteristic value λ and if C is an arbitrary constant, we can easily see by substituting in the equation that $Cw(t)$ is also an eigenfunction corresponding to the same λ. The value $\lambda = 0$ is not a characteristic value because we have the corresponding solution $w(t) = 0$ for every value of t, which is the trivial case, excluded in our discussion.

As we know, A matrix is consistent if and only if it has the form $A = (\frac{w_i}{w_j})$ which is equivalent to multiplying a column vector $(w_1, w_2, \ldots, w_n)^T$ by the row vector $(\frac{1}{w_1}, \frac{1}{w_2}, \ldots, \frac{1}{w_n})$.

Saaty and Vargas [12] proved the following theorems:

Theorem 17. *K(s,t) is consistent if and only if it is separable of the form:*

$$K(s,t) = \frac{k(s)}{k(t)}$$

Theorem 18. *If K(s,t) is consistent, the solution of (26) is given by:*

$$w(s) = \frac{k(s)}{\int\limits_S k(s)ds}$$

In the discrete case, the normalized eigenvector was independent of whether all the elements of the pairwise comparison matrix A are multiplied by the same constant a or not, and thus we can replace A by aA and obtain the same eigenvector. Generalizing this result we have:

$$K(as, at) = aK(s,t)$$

which means that K is a homogeneous function of order one. We can replace $k(s)$ above by $k(as)$ and obtain $w(as)$. We have now derived from considerations of ratio scales the following condition to be satisfied by a ratio scale:

Theorem 19. *A necessary and sufficient condition for $w(s)$ to be an eigenvector solution of Fredholm's equation of the second kind, with a consistent kernel that is homogeneous of order one, is that it satisfies the functional equation*

$$w(as) = bw(s)$$

where $b = \alpha a$.

6.2 The Fundamental Scale

During the paired comparison, instead of assigning two number w_i and w_j and forming the ratio $\frac{w_i}{w_j}$ we assign a single number drawn from the $1-9$ scale of absolute numbers to represent the ratio $\frac{w_i}{w_j}$. The derived scale will reveal what the w_i and w_j are. This is a central fact about the relative measurement approach of the AHP and the need for a fundamental scale.

In 1846 Weber found that people while holding in their hand different weights, could distinguish between a weight of 20g and weight of 21g, but could not distinguish between 40g and 41g. They could between 40g and 42g, and so on at higher levels. We need to increase a stimulus s by a minimum amount Δs to reach a point where our senses can first discriminate between s and $s + \Delta s$. This Δs is called the just noticeable difference. The ratio $r = \frac{\Delta s}{s}$ does not depend on s. Weber's law states that change in sensation is noticed when the stimulus is increased by a constant percentage of the stimulus itself. This law holds in ranges where Δs is small when compared with s, and hence in practice it fails to hold when s is either too small or too large.

In 1860 Fechner considered a sequence of just noticeable increasing stimuli. He denoted the first one by s_0. The next just noticeable stimulus is given by

$$s_1 = s_0 + \Delta s_0 = s_0 + \frac{\Delta s_0}{s_0} s_0 = s_0(1+r)$$

based on Weber's law.
Similarly $s_2 = s_1 + \Delta s_1 = s_1(1+r) = s_0(1+r)^2 = s_0\alpha^2$, by denoting $\alpha = 1+r$.
In general $s_n = s_{n-1}\alpha = s_0\alpha^n$ $(n = 0, 1, 2, \dots)$.
Thus stimuli of noticeable differences follow sequentially in a geometric progression. Fechner noted that the corresponding sensations should follow each other in an arithmetic sequence at the discrete points at which just noticeable differences occur. But the latter are obtained when we solve for n. We have $n = \frac{\log s_n - \log s_0}{\log \alpha}$ and sensation is a linear function of the stimulus. Thus if M denotes the sensation and s the stimulus, the psychophysical law of Weber-Fechner is given by

$$M = a \log s + b, \quad a \neq 0$$

Now we assume that the stimuli arise in making pairwise comparisons of relatively comparable activities. We are interested in responses whose numerical values are in the form of ratios. Thus $b = 0$, from which we must have $\log s_0 = 0$ or $s_0 = 1$, which is possible by calibrating a unit stimulus. Here the unit stimulus is s_0. The next noticeable stimulus is $s_1 = s_0\alpha = \alpha$ which yields a response $a \log \alpha$. The third noticeable stimulus is $s_2 = s_0\alpha^2$ which yields a response of $2a \log \alpha$. Thus we have for the different responses:

$$\begin{aligned} M_0 &= a \log s_0 = 0 \\ M_1 &= a \log \alpha \\ M_2 &= 2a \log \alpha \\ &\vdots \\ M_n &= na \log \alpha \end{aligned}$$

While the noticeable ratio increases geometrically, the response to that stimulus increases arithmetically. Note that $M_0 = 0$ and there is no response. By dividing each M_i by M_1 we obtain the sequence of absolute numbers $1, 2, 3, \ldots$ of the fundamental $1 - 9$ scale.

But why is 9 the upper value of the scale? One reason can be that people have a capacity to divide their response to stimuli into three categories: high, medium and low. They also have the capacity to refine this division by further subdividing each of these intensities of responses into high, medium and low, thus yielding in all nine subdivisions.

It may be useful to mention that instead of the $1 - 9$ linear scale, other scales were also observed. Consider the following function:

$$\Phi : t \mapsto e^{\tanh^{-1}\left(\frac{t-1}{9}\right)}$$

where $t = 1, 2, \ldots, 9$. Then we get

Original scale	Modified scale
1	1.000
2	1.118
3	1.254
4	1.414
5	1.612
6	1.871
7	2.236
8	2.828
9	4.123

As empirical results of Tung and Tang [17] show, we can not state that this modified AHP is better than the original AHP with respect to rank reversals and consistency.

6.3 The Analytic Network Process (ANP)

The Analytic Network Process [11] is built on the AHP. It is a general theory of relative measurement used to derive composite priority ratio scales from individual ratio scales that represent relative measurements of the influence of elements that interact with respect to control criteria. Through its supermatrix whose elements are themselves matrices of column priorities, the ANP captures the outcome of dependence and feedback within and between clusters of elements. The Analytic Hierarchy Process with its dependence assumptions on clusters and elements is a special case of the ANP. The ANP provides a general framework to deal with decisions without making assumptions about the independence of higher level elements from lower level elements and about the independence of the elements within a level. In fact the ANP uses a network without the need to specify levels as in a hierarchy.

The ANP is a coupling of two parts. The first consist of a control hierarchy or network of criteria and sub-criteria that control the interactions. The second is a network of influences among the elements and clusters. The network varies from criterion to criterion and a different supermatrix of limiting influence is computed for each control criterion. Finally, each of these supermatrices is weighted by the priority of its control criterion and the results are synthesized through addition for all the control criteria.

6.4 Expert Choice

Since 1983, Expert Choice has been improving the decision-making of businesses. In the early 1980s, Dr. Ernest Forman developed Expert Choice by adapting AHP for use with PCs and was granted one of the first computer software patents.

Expert Choice was used in a variety of critical decisions in the 1980s for different organizations. They used Expert Choice to do strategic planning, project prioritization, resource allocation and technology vendor selection.

Today, Expert Choice is used to allocate over $30 billion a year for government and commercial agencies worldwide. The application has now been adapted for Internet use, allowing for greater efficiency in group decision-making. By allowing users to give critical input from anywhere around the world, the software eliminates the need for unnecessary business travel and excess meetings.

References

[1] Eckart, C. and Young, G., The approximation of one matrix by another of lower rank, *Psychometrika* 1 (1936): 211-218.

[2] Expert Choice Home Page, *http://www.expertchoice.com*

[3] Forman, E.H. (1990), Random Indices for incomplete pairwise comparison matrices, *European Journal of Operational Research* 48: 153-155.

[4] Gass, S.I. and Rapcsák, T., A note on synthesizing group decisions, *Decision Support Systems* 22 (1998) 59-63.

[5] Golub, G.H. and Van Loan, C.F. Matrix computations *The Johns Hopkins University Press, Baltimore, Maryland,* 1983.

[6] Jensen, R.E. (1983), Comparison of Eigenvector, Least squares, Chi square and Logarithmic least square methods of scaling a reciprocal matrix, *Working Paper 153 http://www.trinity.edu/rjensen/127wp/127wp.htm*

[7] Kumar, N. Vinod and Ganesh, L.S., A simulation-based evaluation of the approximate and the exact eigenvector methods employed in AHP (1996) *European Journal of Operational Research* 95: 656-662.

[8] Kwiesielewicz, M., The logarithmic least squares and the generalized pseudoinverse in estimating ratios, *European Journal of Operational Research* 93 (1996): 611-619.

[9] Saaty, Thomas L., The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation, *McGraw-Hill, New York,* 1980.

[10] Saaty, Thomas L. (1988), Decision Making for Leaders -The Analytical Hierarchy Process for decisions in a complex world.

[11] Saaty, Thomas L. Fundamentals of The Analytic Network Process, *ISAHP 1999, Kobe, Japan, August 12-14.*

[12] Saaty, Thomas L. The seven pillars of the Analytic Hierarchy Process, *ISAHP 1999, Kobe, Japan, August 12-14.*

[13] Saaty, Thomas L., Multicriteria Decision Making: The Analytic Hierarchy Process, (1990) *RWS Publications, 4922 Ellsworth Avenue, Pittsburgh, PA 15213.*

[14] Saaty, Thomas L., Fundamentals of Decision Making and Priority Theory, (1994) *RWS Publications, 4922 Ellsworth Avenue, Pittsburgh, PA 15213.*

[15] Triantaphyllou, E. and Sánchez, A., A Sensitivity Analysis Approach for Some Deterministic Multi-Criteria Decision-Making Methods, *Decision Sciences,* Volume 28, Number 1, pp. 151-194, Winter 1997

[16] Tummala, V.M.R. and Ling, Hong (1998), A note on the computation of the mean random consistency index of the Analytic Hierarchy Process (AHP), *Theory and Decision* 44: 221-230.

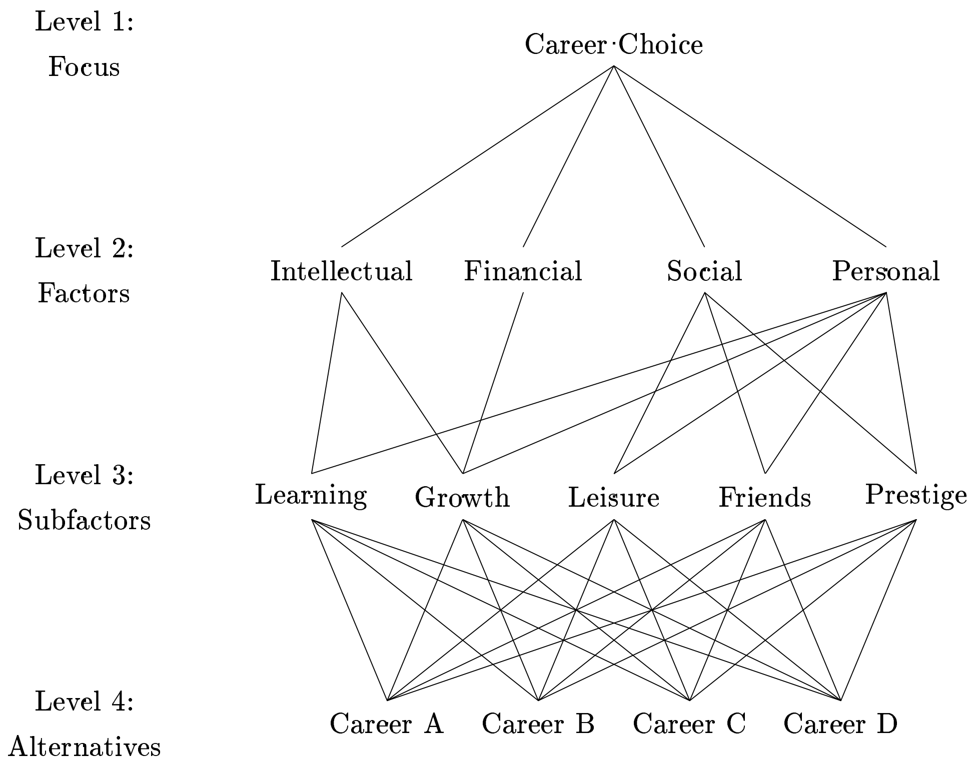
[17] Tung, S.L. and Tang, S.L. A comparison of the Saaty's AHP and modified AHP for right and left eigenvector inconsistency *European Journal of Operational Research* 106 (1998): 123-128.

[18] Wierzbicki, A.P. and Makowski, M. Multi-Objective Optimization in Negotiation Support, *System Analysis Techniques for International Negotiations, Scoping Conference, held at IIASA on October 9-10, 1991.*

[19] Yu, Po-Lung, Multi-criteria decision making, *Plenum Press, New York, 1985.*

Appendix 1

Hierarchy for Choosing a Career



Appendix 2

```
> #Computation of Random Consistency for n=3,4,...,10
> #Maple V Release 4
> with(linalg):
> Digits:=16:
> r:=proc()
> t:=rand(1..17):
> s:=t():
> if s>9 then 1/(s-9+1) else s fi end:
> c:=proc(m)
> if m>9 then 1/(m-9+1) else m fi end:
> maxroot:=proc(A)
> s:=charpoly(A,x):
> n:=degree(charpoly(A,x), x);
> h:=1;
> lb:=n-0.1; # lower limit for lambda_max
> # from Gersgorin's theorem
> for i while h>10^(-12) do
> t:=(lb+ub)/2;
> h:=(ub-lb)/2;
> if evalf(subs(x=t,s))>0 then ub:=t else lb:=t fi; od;
> t; end:


> sam[4]:=1000000:  #sample of 4x4 matrices
> sam[5]:=1000000:  # " 5x5 "
> sam[6]:= 700000:  # " 6x6 "
> sam[7]:= 700000:  # " 7x7 "
> sam[8]:= 400000:  # " 8x8 "
> sam[9]:= 300000:  # " 9x9 "
> sam[10]:=200000:  # " 10x10 "
```

```
> n:=3:  #for n=3 we can compute all the 17^3 matrices
> sumlambda:=0:
> for i to 17 do
> for j to 17 do
> for k to 17 do
> A:=matrix([[1,c(i),c(k)],
> [1/c(i),1,c(j)],
> [1/c(k),1/c(j),1]]);
> sumlambda:=sumlambda+maxroot(A);
> od; od; od; mrci:=(sumlambda/4913 - n)/(n-1):
> print(`n=3`, `MRCI` =mrci , `sample=4913`);
> for n from 4 to 10 do
> sumlambda:=0:
> for counter to sam[n] do
> for i to n-1 do for j from i+1 to n do
> a[i][j]:=r():a[j][i]:=1/a[i][j] od:  od:
> for i to n do a[i][i]:=1 od:
> A:=matrix([seq([seq(a[w][q],q=1..n)],w=1..n)] ):
> sumlambda:=sumlambda+maxroot(A);
> od:  mrci:=(sumlambda/sam[n] - n)/(n-1):  sm:=sam[n]:
> print(`n=`.n , `MRCI` =mrci , `sample=`.sm); od:
```

$$n = 3,\ MRCI = .5244574542911805,\ sample = 4913$$

$$n = 4,\ MRCI = .8830132353452345,\ sample = 1000000$$

$$n = 5,\ MRCI = 1.1085416778272295,\ sample = 1000000$$

$$n = 6,\ MRCI = 1.2492742025565432,\ sample = 700000$$

$$n = 7,\ MRCI = 1.3405613627172790,\ sample = 700000$$

$$n = 8,\ MRCI = 1.4042148583424556,\ sample = 400000$$

$$n = 9,\ MRCI = 1.4511271015999108,\ sample = 300000$$

$$n = 10,\ MRCI = 1.4856984317151252,\ sample = 200000$$

Remark. In fact this program was runned by dividing into parts, the total computing time was 150 hours using P II 350MHz processor.

Appendix 3

```
> #Eigenvector Method
> #Maple V Release 4
> with(linalg):
> mrci[3]:=0.5245:  #Mean Random Consistency Indexes
> mrci[4]:=0.8830:
> mrci[5]:=1.1085:
> mrci[6]:=1.2493:
> mrci[7]:=1.3405:
> mrci[8]:=1.4042:
> mrci[9]:=1.4511:
> mrci[10]:=1.4857:

> ordering:=proc(m)
> si:=nops(m); for ii to si do
> mm[ii][1]:=m[ii];mm[ii][2]:=ii;od;
> for ii to si do for jj to si-1 do
> if (mm[jj][1] < mm[jj+1][1]) then
> h1:=mm[jj][1];h2:=mm[jj][2];mm[jj][1]:=mm[jj+ 1][1];
> mm[jj][2]:=mm[jj+1][2];mm[jj+1][1]:=h1;mm[jj+ 1][2]:=h2;
> fi od:  od:
> o:=matrix([[seq(mm[q][2],q=1..si)],[seq(mm[q] [1],q=1..si)]]);
> end:

> priorEM:=proc(M)
> EV:='EV':v:=[seq(evalf(Eigenvals(M,EV))[q],q= 1..n)];
> s:=nops(v);
> for i to s do if Re(v[i])=v[i] then rv[i]:=v[i]
> else rv[i]:=0:  fi od:
> j:=1:
> for i to s do if rv[i]>rv[j] then j:=i:  fi:  od:
> lambda:=v[j]; #lambda_max of M
> for i to s do
> ev:=[seq(evalf(EV[q,j]),q=1..s)]; od;
> #The eigenvector corresponding to lambda-max of M
> h:=sum(ev[q],q=1..s);
> ev:=ordering([seq(ev[q]/h,q=1..s)]); #normalizing
> ci:=evalf((lambda-s)/(s-1)); #Consistency Index
> cr:=ci/mrci[n]:
```

```
> for i to s do print(Alt[ev[1,i]],ev[2,i]) od:
> print(`lambda_max`=lambda):
> print(`CR`=cr):
> end:

> n:=8:
> A:=matrix([
> [1,5,3,7,6,6,1/3,1/4],
> [1/5,1,1/3,5,3,3,1/5,1/7],
> [1/3,3,1,6,3,4,6,1/5],
> [1/7,1/5,1/6,1,1/3,1/4,1/7,1/8],
> [1/6,1/3,1/3,3,1,1/2,1/5,1/6],
> [1/6,1/3,1/4,4,2,1,1/5,1/6],
> [3,5,1/6,7,5,5,1,1/2],
> [4,7,5,8,6,6,2,1]]);
> priorEM(A);
```

$$A := \begin{bmatrix} 1 & 5 & 3 & 7 & 6 & 6 & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{5} & 1 & \frac{1}{3} & 5 & 3 & 3 & \frac{1}{5} & \frac{1}{7} \\ \frac{1}{3} & 3 & 1 & 6 & 3 & 4 & 6 & \frac{1}{5} \\ \frac{1}{7} & \frac{1}{5} & \frac{1}{6} & 1 & \frac{1}{3} & \frac{1}{4} & \frac{1}{7} & \frac{1}{8} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & 3 & 1 & \frac{1}{2} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{4} & 4 & 2 & 1 & \frac{1}{5} & \frac{1}{6} \\ 3 & 5 & \frac{1}{6} & 7 & 5 & 5 & 1 & \frac{1}{2} \\ 4 & 7 & 5 & 8 & 6 & 6 & 2 & 1 \end{bmatrix}$$

$$Alt_8,\ .3331889793$$

$$Alt_3,\ .1881102989$$

$$Alt_1,\ .1730174748$$

$$Alt_7,\ .1668336738$$

$$Alt_2,\ .05397572306$$

$$Alt_6,\ .03632220209$$

$$Alt_5,\ .03104373640$$

$$Alt_4,\ .01750791133$$

$$lambda\_max = 9.66888729$$

$$CR = .1697852657$$

```
> n:=3:  #Size of house
> A:=matrix([[1,6,8],[1/6,1,4],[1/8,1/4,1]]); priorEM(A);
```

$$A := \begin{bmatrix} 1 & 6 & 8 \\ \frac{1}{6} & 1 & 4 \\ \frac{1}{8} & \frac{1}{4} & 1 \end{bmatrix}$$

$$Alt_1,\ .7535538991$$

$$Alt_2,\ .1811354644$$

$$Alt_3,\ .06531063634$$

$$lambda\_max = 3.135610853$$

$$CR = .1292763136$$

```
> n:=3:  #Location to bus
> A:=matrix([[1,7,1/5],[1/7,1,1/8],[5,8,1]]); priorEM(A);
```

$$A := \begin{bmatrix} 1 & 7 & \frac{1}{5} \\ \frac{1}{7} & 1 & \frac{1}{8} \\ 5 & 8 & 1 \end{bmatrix}$$

$$Alt_3,\ .7124870368$$

$$Alt_1,\ .2330592338$$

$$Alt_2,\ .05445372922$$

$$lambda\_max = 3.246954582$$

$$CR = .2354190486$$

```
> n:=3:  #Neighborhood
> A:=matrix([[1,8,6],[1/8,1,1/4],[1/6,4,1]]); priorEM(A);
```

$$A := \begin{bmatrix} 1 & 8 & 6 \\ \frac{1}{8} & 1 & \frac{1}{4} \\ \frac{1}{6} & 4 & 1 \end{bmatrix}$$

$$Alt_1,\ .7535538991$$

$$Alt_3,\ .1811354644$$

$$Alt_2,\ .06531063634$$

$$lambda\_max = 3.135610853$$

$$CR = .1292763136$$

```
> n:=3:  #Age
> A:=matrix([[1,1,1],[1,1,1],[1,1,1]]); priorEM(A);
```

$$A := \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$Alt_1,\ .3333333335$$

$$Alt_3,\ .3333333335$$

$$Alt_2,\ .3333333331$$

$$lambda\_max = 2.999999999$$

$$CR = .9532888465\ 10^{-9}$$

```
> n:=3:  #Yard space
> A:=matrix([[1,5,4],[1/5,1,1/3],[1/4,3,1]]); priorEM(A);
```

$$A := \begin{bmatrix} 1 & 5 & 4 \\ \frac{1}{5} & 1 & \frac{1}{3} \\ \frac{1}{4} & 3 & 1 \end{bmatrix}$$

$$Alt_1,\ .6738105712$$

$$Alt_3,\ .2255354991$$

$$Alt_2,\ .1006539297$$

$$lambda\_max = 3.085766691$$

$$CR = .08176042993$$

```
> n:=3:  #Modern facilities
> A:=matrix([[1,8,6],[1/8,1,1/5],[1/6,5,1]]); priorEM(A);
```

$$A := \begin{bmatrix} 1 & 8 & 6 \\ \frac{1}{8} & 1 & \frac{1}{5} \\ \frac{1}{6} & 5 & 1 \end{bmatrix}$$

$$Alt_1,\ .7466065849$$

$$Alt_3,\ .1933233541$$

$$Alt_2,\ .06007006103$$

$$lambda\_max = 3.197275844$$

$$CR = .1880608618$$

```
> n:=3:  #General condition
> A:=matrix([[1,1/2,1/2],[2,1,1],[2,1,1]]); priorEM(A);
```

$$A := \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 2 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix}$$

$$Alt_2,\ .4000000001$$

$$Alt_3,\ .4000000001$$

$$Alt_1,\ .2000000002$$

$$lambda\_max = 3.000000000$$

$$CR = 0$$

```
> n:=3:  #Financing
> A:=matrix([[1,1/7,1/5],[7,1,3],[5,1/3,1]]); priorEM(A);
```

$$A := \begin{bmatrix} 1 & \frac{1}{7} & \frac{1}{5} \\ 7 & 1 & 3 \\ 5 & \frac{1}{3} & 1 \end{bmatrix}$$

$$Alt_2,\ .6491180045$$

$$Alt_3,\ .2789545655$$

$$Alt_1,\ .07192742997$$

$$lambda\_max = 3.064887579$$

$$CR = .06185660534$$