Author(s): Brunelli, Matteo

Title: Introduction to the Analytic Hierarchy Process

Year: 2015

Version: Post print

Please cite the original version:

Brunelli, Matteo. 2015. Introduction to the Analytic Hierarchy Process. SpringerBriefs in Operations Research. P. 83. 978-3-319-12502-2 (electronic). 10.1007/978-3-319-12502-2.

Introduction to the Analytic Hierarchy Process *

Matteo Brunelli

Department of Mathematics and Systems Analysis, Aalto University

P.O. Box 11100, FIN-00076 Aalto, Finland

e-mail: matteo.brunelli@aalto.fi

January 22, 2015

*The final version of this draft has been published as: M. Brunelli, *Introduction to the Analytic Hierarchy Process*. SpringerBriefs in Operations Research (2015). Available at Springer via http://dx.doi.org/10.1007/978-3-319-12502-2

Contents

Preface

Why would anyone feel urged to write another book on the Analytic Hierarchy Process (AHP), given those already written? I felt urged because the existing books on the AHP are conservative, too anchored to the original framework, and do not cover recent results, whereas lots of questions have been addressed in the last years. Apparently, the interest in the AHP has not faded in the last years, and we shall see that this view is also supported by other studies, as well as by the years of publication of many of the references used in this booklet.

Now, the next question one should ask himself when writing a tutorial should regard to whom the tutorial is for. With the premise that a decision scientist might find these pages too simplistic, in my intentions, the readership should include the following categories.

- *Practitioners* and *consultants* willing to apply, and *software developers* willing to implement, the AHP. Some collateral issues, for instance the incompleteness of judgments, are usually neglected in didactic expositions, but remain fundamental in practical implementations. On the software development side, at present, there is still not a modern and free software which covers all the aspects of the AHP presented in this booklet

- Recent advances in the theory have been disseminated in different journals and, as research requires, are narrow, technical, and often use heterogeneous notation and jargon. Therefore, I also hope that *students* who have been introduced to the AHP and want to have an updated exposition on, and references to, the state of the art can find these pages useful

- Even the *applied mathematicians* might find it interesting. The mathematics behind the method is simple, but some of its extensions have been a fertile ground for the application of non-trivial concepts stemming from abstract algebra and functional analysis, just to mention two areas of interest.

The following pages assume neither previous knowledge of the AHP, nor higher mathematical preparation than some working knowledge of calculus and linear algebra with eigenvector theory. A brief tutorial on eigenvalues and eigenvectors is provided in the appendix. Moreover, some sections are marked with the symbol * to indicate that they contain further discussions and references to research literature. The reader interested in the fundamentals might want to skip them.
Ideally, this booklet is also articulated to suit different levels of readership. I believe that the following three can serve as approximate guidelines:

- A *basic* exposition is given in Chapter 1 with the exclusion of the section marked with *. The reader can then proceed examining Section 2.1 until the end of §2.1.1, Section 2.2 until the end of §2.2.1, and Section 3.2 with the exclusion of the subsection marked with *. A basic understanding allows the reader to use the AHP only at a superficial level.

- A *complete* exposition of the AHP can be gained by reading this booklet in its entirety, with the exclusion of the sections marked with *. A complete understanding allows the reader to choose between different tools to perform different tasks.

- An *advanced* understanding of the method is like the complete, but with the addition of the sections marked with *. Compared to the complete understanding, in the advanced, the reader will familiarize with the most recent results and the ongoing discussions, and will be able to orient through the literature.

I shall also spell out that I will not refrain from giving a personal perspectives on some problems connected with the AHP, as the method has been a matter of heated debate since its inception.

I hereby wish to thank those who helped me. Among them, I am particularly gratetul to Michele Fedrizzi, who also taught me much of the material contained in this booklet. I am also grateful to Springer, especially in the person of Matthew Amboy. Furthermore, this project has been financed by the Academy of Finland.

It goes without saying that I assume the paternity of all imprecisions and mistakes and that the reader is welcome to contact me.

Espoo, Finland, November 2014 Matteo Brunelli

1. Introduction and fundamentals

> Beauty started when people began to choose.
>
> Roberto Benigni

In a world whose complexity is rapidly growing, making the best decisions becomes an increasingly demanding task for managers of companies, governmental agencies and many other decision and policy makers. In recent years, this has gone arm-in-arm with the growth of what are now known as *decision analytics* methodologies. Namely, decision makers are more reluctant to make gut decisions based of feelings and hunches, and instead prefer to use analytic and quantitative tools, and base and analyze their decisions on a solid ground. Many methods stemming from applied mathematics and operations research have proved useful to help decision makers making informed decisions, and among these methods there are also those requiring, as inputs, subjective judgments from a decision maker or an expert. It is in this context that the Analytic Hierarchy Process (AHP) becomes a useful tool for analyzing decisions.

What is the AHP? Broadly speaking, the AHP is a theory and methodology for *relative measurement*. In relative measurement we are not interested in the exact measurement of some quantities, but rather on the *proportions* between them. Consider a pair of stones. In classical measurement we might be interested in knowing their exact weights and the pair of measurements $(2, 1)$ is not correct unless the weight of the first stone is 2kgs and the weight of the second is 1kg. Conversely, in relative measurement we confine our interest to the knowledge of how much heavier each object is compared to another. Hence, the pair of measurements $(2, 1)$ is correct as long as the weight of the first stone is double the weight of the second. It follows that, in this example, if we use relative measurement theory the pairs of measurements $(2/3, 1/3)$ $(4, 2)$, $(8, 4)$ are also correct for the two stones. Relative measurement theory suits particularly well problems where the best alternative has to be chosen. In fact, in many cases we are not really interested in the precise scores of the alternatives but it is sufficient to know their relative measurements to know which alternative is the best. Moreover, when attributes of alternatives are intangible, it is difficult to devise a measurement scale and using relative measurements simplifies the analysis. The ultimate scope of the AHP is that of using pairwise comparisons between alternatives as inputs, to produce a rating of alternatives, compatibly with the theory of relative measurement.

In what field of study do we stand when we talk of the AHP? It is the author's opinion that the AHP should be placed in the intersection between decision analysis and operations research. Keeney and Raiffa [76] gave the following definition of *decision*

analysis:

> The theory of decision analysis is designed to help the individual make a choice among a set of prespecified alternatives.

Hence, as long as the AHP is used as a technology for aiding decisions, it seems that its study belongs to decision analysis.
On the other hand, to justify its connection with operations research, without going too far, we can refer to some definitions reported by Saaty, the main developer of the AHP, in one of the first graduate textbooks in operations research [100]. In his book, curious and thought-provoking definitions can be found: operations research was defined as "quantitative common sense" and, perhaps in the intent of underlining its limitations, as "the art of giving bad answers to problems to which otherwise worse answers are given". Such definitions are surely thought-provoking but they capture the essence of quantitative methods, which is that of helping make better decisions. Consulting the Merriam-Webster dictionary one can find the following definition of *operations research*:

> The application of scientific and especially mathematical methods to the study and analysis of problems involving complex systems.

Hence, it is straightforward to conclude that the study of the AHP belongs to operations research too. Within operations research, two different types of studies appeared. The classical operations research, more mathematically oriented, which studies the modeling and solution of structured problems can be called 'hard' operations research. Conversely, especially recently, the effort of applying the reasoning of operations research to problems which, by nature, are unstructured, has gone under the name of 'soft' operations research. Perhaps the fact that the AHP mostly deals with subjective judgments and intangible attributes gave the false idea that it did not belong to the tools or 'hard' operations research, but rather to its 'soft' side. However, in recent discussions [90] the role of the AHP has been revisited and now it seems clearer that it has been a matter of study for 'hard' operations research. The positioning of the AHP is depicted in Figure 1.1.

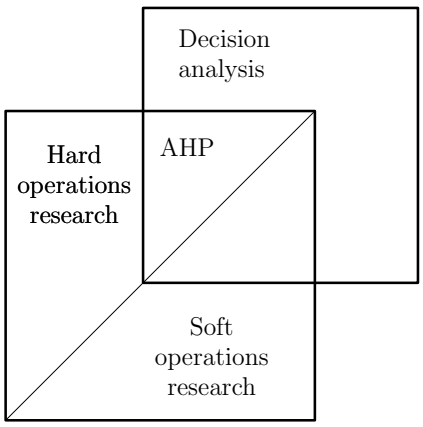


Figure 1.1.: The position of the AHP in the scientific debate.

Although the utility of the AHP is not limited to the following, it is safe to say that it has been especially advocated to be used with *intangible* criteria and alternatives, and thus used to solve *multi-criteria decision making* (MCDM) problems, which are choice problems where alternatives are evaluated with respect to multiple criteria. Tangible properties of alternatives, for example the weight of different stones or the salary of different employees, can be measured without ambiguity and subjectivity. Hence, the machinery of the AHP becomes unnecessary. Conversely, when the magnitude of some properties of alternatives, such as the dexterity of a sportsman or the aesthetic appeal of a bridge, cannot be easily grasped and measured we are in the domain of the intangibles, which is where the AHP gives its best.

The organization of this booklet is quite unorthodox and differs from the approach used in other expositions [19, 20, 74, 102, 103, 106]. Here, at the very beginning, the AHP is presented through a normative lens with lots of assumptions. That is, the AHP is introduced as a method which works in a rational world with full information. However, since this is clearly not the world we are living in, successively, by pointing out the limits of this normative approach, binding assumptions are relaxed and the AHP more fully explained. In this sense the reader should not be deceived: the exposition of the AHP contained in this first chapter is by no means complete, and it is even narrower than the one given originally by Saaty. But, as said, this little trick shall hopefully help to expose the AHP in a more natural and painless way.

In the following we shall use a standard notation where vectors are noted in boldface, e.g. $\mathbf{w} = (w_1, \ldots, w_n)^T$ and matrices (all square) in capital boldface, e.g. $\mathbf{A} = (a_{ij})_{n\times n}$. The set of real numbers is $\mathbb{R}$ and the set of positive real numbers is $\mathbb{R}_>$. We shall use open square brackets to indicate open intervals, e.g. $]0, 1[$.

1.1. Fundamentals

As already mentioned, in our framework, the AHP can be applied to a multitude of decision making problems involving a finite number of alternatives. Formally, in this setting, in a decision process there is one *goal* and a finite set of *alternatives*, $X = \{x_1, \ldots, x_n\}$, from which the decision maker, is usually asked to select the best one.
Explaining the AHP is like teaching a child how to tie the shoestrings: easier to show with an example than to explain with words. Hence, it is time to present a prototypical example which will accompany us for the rest of this section: a family has to decide which European city to visit during their holidays. Reasonably, the goal of the family is the highest satisfaction with their destination. Alternatives may be some cities, in our simple example

$$X = \{\underbrace{\text{Rome}}_{x_1}, \underbrace{\text{Barcelona}}_{x_2}, \underbrace{\text{Reykjavik}}_{x_3}\}, \tag{1.1}$$

and the structure of the problem represented in Figure 1.2.
Often, in decision processes, the decision maker is asked to assign a score to each alternative and then to choose the one with the maximum value. That is, given a set of

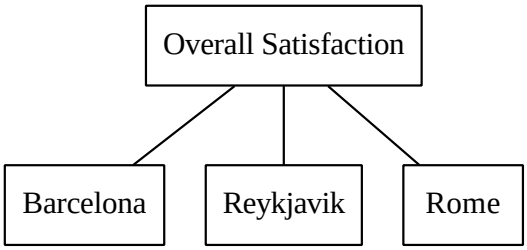


Figure 1.2.: Evaluating alternatives with respect to an overall goal.

alternatives, $X = \{x_1, \ldots, x_n\}$, the decision maker should provide a weight vector

$$\mathbf{w} = (w_1, \ldots, w_n)^T, \tag{1.2}$$

where w_i is a value which coherently estimates the score of alternative x_i. That is, the greater w_i, the better the ith alternative. Similarly to what happens for value theory [55], the rule is that alternative x_i is preferred to alternative x_j if and only if $w_i > w_j$. Weight vectors are nothing else but ratings, and their components w_i are called priorities, or weights, of the alternatives x_i. For example, $\mathbf{w} = (0.4, 0.2, 0.3, 0.1)^T$ implies $x_1 \succ x_3 \succ x_2 \succ x_4$ where $x_i \succ x_j$ means that alternative x_i is preferred to x_j. Possible ties are expressed as $x_i \sim x_j$.

Example 1. *Consider the example of the choice of the best site for holidays. If the vector* $\mathbf{w} = (0.3, 0.5, 0.2)^T$ *was associated with the set of alternatives*

$$X = \{Rome, Barcelona, Reykjavik\}$$

then we would have that

$$Barcelona \succ Rome \succ Reykjavik$$

because $w_2 > w_1 > w_3$.

Making decisions in this way seems easy, but it becomes a hard task when complexity increases. As we will see, complexity augments arm-in-arm with the number of alternatives and criteria.

From the priority vector to the pairwise comparison matrix

It is clear that a decision maker could run into troubles when asked to submit a rating in the form of a numerical vector for a large number of alternatives. Does not it often happen that we cannot decide among several alternatives? Even worse, do not we decide and eventually realize that it was not the best decision? This is a matter of fact and originates from our cognitive limits and the impossibility of effectively comparing several alternatives at the same time.
An effective way to overcome this problem is to use *pairwise comparisons*. The reason

for doing so, is that this allows the decision maker to consider two alternatives at a time. Thus, the strategy is that of decomposing the original problem into many smaller subproblems and deal with these latter ones. Formally, the pairwise comparisons are collected into a *pairwise comparison matrix*, $\mathbf{A} = (a_{ij})_{n\times n}$, structured as follows

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \tag{1.3}$$

with $a_{ij} > 0$ expressing the degree of preference of x_i to x_j. More precisely, according to Saaty's theory, each entry is supposed to approximate the ratio between two weights

$$a_{ij} \approx \frac{w_i}{w_j} \quad \forall i, j. \tag{1.4}$$

This means that, if the entries exactly represent ratios between weights, then the matrix $\mathbf{A}$ can be expressed in the following form,

$$\mathbf{A} = (w_i/w_j)_{n\times n} = \begin{pmatrix} w_1/w_1 & w_1/w_2 & \dots & w_1/w_n \\ w_2/w_1 & w_2/w_2 & \dots & w_2/w_n \\ \vdots & \vdots & \ddots & \vdots \\ w_n/w_1 & w_n/w_2 & \dots & w_n/w_n \end{pmatrix}. \tag{1.5}$$

Note that, as soon as we account for (1.4) and consider (1.5), a condition of multiplicative reciprocity $a_{ij} = 1/a_{ji}\ \forall i, j$ holds, and $\mathbf{A}$ can be simplified and rewritten,

$$\mathbf{A} = \begin{pmatrix} 1 & a_{12} & \cdots & a_{1n} \\ \frac{1}{a_{12}} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{a_{1n}} & \frac{1}{a_{2n}} & \cdots & 1 \end{pmatrix}. \tag{1.6}$$

In words, the simplified structure of pairwise comparison matrices in this form follows from the assumption that if, for example, x_1 is 2 times better than x_2, then we can deduce that x_2 is $1/2$ as good as x_1.
Let us now proceed with the example and imagine a pairwise comparison matrix for the set of cities X as defined previously, in §1.1. In this case, and only in this case, to facilitate the understanding, the labels x_1, x_2, x_3 are attached to the rows and columns of the matrix.

$$\mathbf{A} = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array} \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{pmatrix} 1 & 3 & 6 \\ 1/3 & 1 & 2 \\ 1/6 & 1/2 & 1 \end{pmatrix} \end{array} \tag{1.7}$$

From this matrix, in particular from entry a_{12}, one can figure out that x_1 (Rome) is considered three times better than x_2 (Barcelona). That is, $a_{12} = 3$ suggests us that

$w_1 = 3w_2$. Once a pairwise comparison matrix is completed, there are many methods to derive the priority vector $\mathbf{w}$. In the example it can be checked that the condition $a_{ij} = w_i/w_j \ \forall i, j$ is satisfied by, for instance, the following vector with its components summing up to one,

$$\mathbf{w} = \begin{pmatrix} 6/9 \\ 2/9 \\ 1/9 \end{pmatrix},$$

and thus Rome (x_1) is ranked the best. To summarize, whenever the number of alternatives is too large, pairwise comparing them is an effective way for obtaining a rating. Perhaps we have spent a bit more of our time but the rating of alternatives contained in $\mathbf{w}$ is now more robust than it would have been if it had been estimated directly, without using $\mathbf{A}$. We shall here ask the reader for a leap of faith and leave the issue of the weight determination open and discuss it later.

From the pairwise comparison matrix to the hierarchy

At this point, it is time to wonder why the pairwise comparison matrix $\mathbf{A}$ was filled in that particular way and what factors influenced the decision maker's judgments. Needless to say, such decision factors are few if the expert is choosing the type of bread to buy (mainly price and quality) whereas they are several when a member of a parliament has to vote a proposition (sake of the electors, own reputation, likelihood of reelection, and surely many others). First, we should start using the word criterion instead of factor and reckon that, if we can make decisions and account for multiple, and possibly conflicting criteria, we are in the realm of Multi Criteria Decision Making (MCDM) methods.
Formally, in the decision making process, the expert has to consider a set of *criteria* $C = \{c_1, \ldots, c_m\}$, which are characteristics making one alternative preferable to another with respect to a given goal. In the example, which regards the location for holidays, the set of criteria could be

$$C = \{\underbrace{\text{climate}}_{c_1}, \underbrace{\text{sightseeing}}_{c_2}, \underbrace{\text{environment}}_{c_3}\}, \tag{1.8}$$

At this point we need at least a graphical formalism to combine alternatives, criteria and goals and represent the structure of the problem in an intuitive way. In the AHP, a *hierarchy* serves this purpose and is compounded by:

- the goal
- the set of alternatives
- the set of criteria
- a relation on the goal, the criteria and the alternatives.

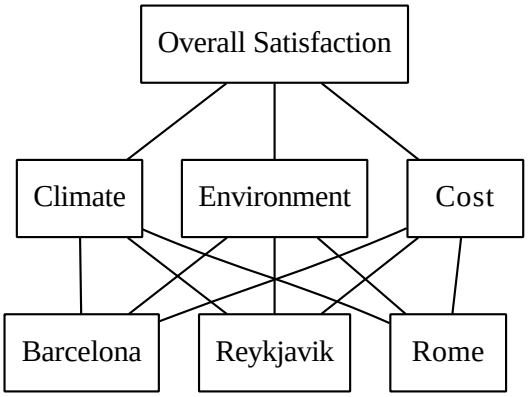


Figure 1.3.: Hierarchy for the European city selection problem. At the top level there is the goal, at the bottom there are the alternatives, and criteria are in intermediate levels. A line connecting two elements marks the existence of a relation of hierarchical dependence between them.

A graphical example of hierarchy for the decision on the European city is depicted in Figure 1.3. Note that in this booklet we shall not dwell on the hierarchy in more formal terms since it would be beyond its scope.

The main drawback of the pairwise comparison matrix $\mathbf{A}$ in (1.7) is that it compared alternatives without considering criteria. Simply, when filling it, the decision maker was only thinking about the overall satisfaction with the alternatives and did not make any separate reasoning about the criteria—cost, sightseeing and environment in the example—contributing to the global satisfaction.

Once again, complexity can be a problem and the solution is to decompose it. This is why, at this point, Saaty [101] suggested to build a different matrix for each criterion. Hence, in the following, a matrix $\mathbf{A}^{(k)}$ is the matrix of pairwise comparisons between alternatives according to criterion k. For example, using the convention c = climate, s = sightseeing, e = environment, entry a_{13} of matrix $\mathbf{A}^{(c)}$ below entails that the decision maker prefers Rome to Reykjavik if he compares these two cities *exclusively* under the climatic point of view. The following three matrices can be taken as examples of preferences expressed by a decision maker on the three cities according to the three different criteria.

$$\mathbf{A}^{(c)} = \begin{pmatrix} 1 & 1 & 4 \\ 1 & 1 & 4 \\ 1/4 & 1/4 & 1 \end{pmatrix} \qquad \mathbf{A}^{(s)} = \begin{pmatrix} 1 & 2 & 6 \\ 1/2 & 1 & 3 \\ 1/6 & 1/3 & 1 \end{pmatrix} \qquad \mathbf{A}^{(e)} = \begin{pmatrix} 1 & 1/2 & 1/8 \\ 2 & 1 & 1/4 \\ 8 & 4 & 1 \end{pmatrix}.$$

Then, we estimate (no worries for the moment about the method) their priority vectors

$$\mathbf{w}^{(c)} = \begin{pmatrix} 4/9 \\ 4/9 \\ 1/9 \end{pmatrix} \qquad \mathbf{w}^{(s)} = \begin{pmatrix} 6/10 \\ 3/10 \\ 1/10 \end{pmatrix} \qquad \mathbf{w}^{(e)} = \begin{pmatrix} 1/11 \\ 2/11 \\ 8/11 \end{pmatrix}.$$

Now we have three vectors instead of one! Their interpretation is at least twofold: (i) as they are 3 vectors of dimension 3, one can imagine them as 3 points in the 3-dimensional Euclidean space; (ii) vectors are ratings and they can be contradictory: climate-wise Barcelona is preferred to Reykjavik, but, on the other hand, the opposite is true if the criterion is the environment.

It is reasonable to assume that the solution should be a compromise between vectors $\mathbf{w}^{(c)}, \mathbf{w}^{(s)}, \mathbf{w}^{(e)}$. However, the simple arithmetic mean is not the best way to aggregate the vectors because, most likely, criteria have different degrees of importance. For instance, an old and rich man may not care much about the cost and just demand a quiet and peaceful place for his holidays—in this hypothetical case the criterion 'environment' would be judged more important than 'cost'. Hence, we need another type of averaging function and the compromise that we are looking for is the *weighted arithmetic mean*, in this case a convex (linear) combination of vectors. Now, the question is how to find the weights to associate to different vectors. The only thing we know is that the weight associated to a vector should be proportional to the importance of the criterion associated with it. The proposed solution is to use the same technique used so far. First, we build a pairwise comparison matrix $\hat{\mathbf{A}} = (\hat{a}_{ij})_{n\times n}$ which compares the importance of criteria with respect to the achievement of the goal. In the example, the matrix could be

$$\hat{\mathbf{A}} = \begin{pmatrix} 1 & 1/2 & 1/4 \\ 2 & 1 & 1/2 \\ 4 & 2 & 1 \end{pmatrix}.$$

Then, we derive a vector $\hat{\mathbf{w}} = (\hat{w}_1, \hat{w}_2, \hat{w}_3)^T$ (again, no worries on how it is derived)

$$\hat{\mathbf{w}} = \begin{pmatrix} 1/7 \\ 2/7 \\ 4/7 \end{pmatrix}$$

whose components are the weights of criteria. According to this vector the decision maker—in our case the family—is mainly interested in the third criterion, i.e. the environment. We proceed with the linear combination of $\mathbf{w}^{(c)}$, $\mathbf{w}^{(s)}$ and $\mathbf{w}^{(e)}$.

$$\begin{aligned} \mathbf{w} &= \hat{w}_1\mathbf{w}^{(c)} + \hat{w}_2\mathbf{w}^{(s)} + \hat{w}_3\mathbf{w}^{(e)} \\ &= \frac{1}{7}\begin{pmatrix} 4/9 \\ 4/9 \\ 1/9 \end{pmatrix} + \frac{2}{7}\begin{pmatrix} 6/10 \\ 3/10 \\ 1/10 \end{pmatrix} + \frac{4}{7}\begin{pmatrix} 1/11 \\ 2/11 \\ 8/11 \end{pmatrix} \\ &\approx \begin{pmatrix} 0.287 \\ 0.253 \\ 0.460 \end{pmatrix}. \end{aligned}$$

We have a final ranking and we can choose the best alternative, which is the one rated the highest, then x_3 which, in our example, is Reykjavik. Formally, the best alternative is any element of the set $\{x_i | w_i \geq w_j,\ \forall i, j\}$.

The role of the criteria weights can be stressed by a numerical example. Consider the priority vector for criteria $(1/7, 4/7, 2/7)^T$ instead of $(1/7, 2/7, 4/7)^T$. Then, the final priorities become $(0.43, 0.29, 0.28)^T$ and the best alternative is now x_1 (Rome).

Note that hierarchies can contain more levels of criteria. For example, for the selection of the best city for holidays, the criterion 'climate' could have been refined into subcriteria such as 'chance of rain', 'temperature', 'length of the daylight', each of which contributes to the concept of climate. For reasons of space we cannot provide a numerical example of a hierarchy with more criteria levels, but we invite the reader to consider the following exercise.

Problem 1. *Convince yourself that the AHP can work out the hierarchy in Figure 1.4. Note the complication that 'wheels' is a subcriterion of both 'mechanics' and 'aesthetics'.*

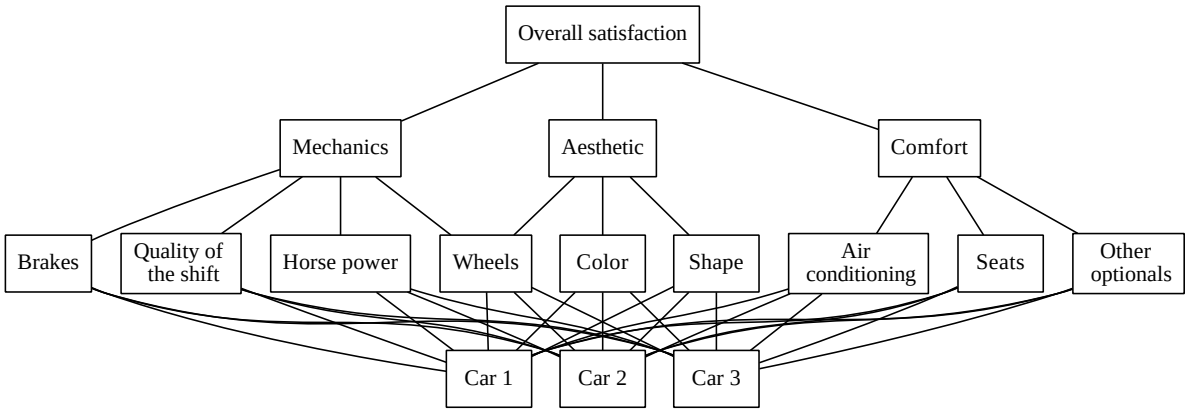


Figure 1.4.: Selection of an automobile.

Certainly, this is an ad hoc toy-example, but the reader might be interested to know that the AHP has actually been proposed for automobile selection [32].

Let us conclude this section by remarking that, by using naïve examples we have already seen the three basic steps of the AHP:

1. Problem structuring and definition of the hierarchy
2. Elicitation of pairwise comparisons
3. Derivation of priority vectors and their linear combinations.

Nevertheless, so far, we have considered an idyllic situation which used various assumptions and only in the next chapters we shall see how the AHP can be used as a more flexible model. Even so, what we know is already sufficient to understand the next section on some applications.

1.2. Applications

Our previous example was simple and aimed at understanding the principles behind the AHP and, needless to say, real-world applications have presented a much higher level of complexity. In this section we shall skim through some applications to show their vast range and hopefully whet the readers' appetite for the AHP, whose full potential has not been revealed yet. Nowadays, applications are so many that no survey can be comprehensive enough. However, albeit not recent, the surveys by Golden *et al.* [64], Zahedi [140], and Vargas [130] remain the best reference points.

City evaluation and planning

Saaty [104] proposed to use the AHP to rank a set of cities from the most to the least livable. In his study, he considered a set of cities in the United States. Indeed, the satisfaction of the final goal 'livability' can be decomposed into the satisfaction of some criteria, such as 'environment', 'services', 'security', and each of these criteria can itself be decomposed into subcriteria. For instance, the 'services' criterion might depend on subcriteria such as 'transportation facilities', 'health care', and so forth. Some cities are undoubtedly more livable than others[1]. Interestingly, in this application, the AHP questionnaire was given to six decision makers representing different demographic groups and light was shed on differences of preferences between them. The research concluded presenting some conjectures on the reasons behind these discrepancies.

Another innovative application was proposed by Saaty and Sagir [113]. By looking at metropolitan areas, the authors were able to classify most of the world cities into one category, out of seven, each representing an alternative model of developing a city. Some alternatives were: compact, 3-dimensional (New York City), flat (Riyadh). The AHP was used to systematically take into account good and bad points of each type of city by means of an AHP-based cost-benefit analysis.

Country ranking

Until the late Eighties, ranking of countries was based on their gross domestic product per capita, or at least that was the most significant measure. More recently, starting in the early Nineties, a more inclusive and composite measure accounting for multiple criteria called Human Development Index has been popularized by economists such as the Nobel laureate Amartya Sen [4]. Few know that, in 1987, the AHP was already proposed to rank countries taking into account multiple criteria [97]. Clearly, in this study, the alternatives were the countries themselves, and the criteria simply all those characteristics which could make one country better than another. Indeed, with a suitable choice of criteria, this use of the AHP can be seen as a primer in the multivariate ranking of countries.

[1] It is the author's of this booklet half joke to say that the choice of San Francisco as the most livable served as a sure validation of the AHP.

Mobile value services

With the widespread use of (smart) mobile devices, mobile services and applications are becoming more and more successful and part of end-users' everyday life, but why are some devices and services successful while others are not? It is indeed of great importance to identify and understand critical success factors driving the acceptance and adoption of mobile devices and different mobile services. Traditional models mainly consider a limited set of adoption factors, focusing on the perceived values of mobile services (usefulness, ease of use, cost). Nikou and Mezei [92] proposed to use the AHP to determine the most important decision criteria driving the customers' adoption of mobile devices and mobile services. The main attributes considered include payment mode, functionality, added value, perceived quality, cost, and performance. The results of this type of studies can be essential for various service providers (operators, mobile handset manufacturers) to design profitable applications that generate value for the end-users.

Organ transplant

It is a fact that there are more people needing human organ transplants than available organs, and that different allocations of organs can make the difference between death and life. Some combinatorial optimization problems have been proposed to match donors with organs in the best possible way, and to be fair, such algorithms account for the fact that some patients require an organ in a shorter time than others. In a study, Lin and Harris [84] proposed to use the AHP to decompose the four criteria 'urgency', 'efficiency', 'benefit', and 'equity' into subcriteria and eventually estimate their importance in the donors-organs matching process. Patients were treated as alternatives, but it is clear that their huge amount would have made the use of subjective judgments impossible. Fortunately, in this case, the pairwise comparison matrices at the alternative level were filled automatically since different criteria were quite easily quantifiable. For instance, if the life expectancies of two patients are 1 and 2 years, it can be automatically derived that under that criterion, the first patient is two times more 'urgent' than the second.

Chess prediction

The AHP has been used for forecasting too. In sports, athletes can be seen as alternatives and their characteristics as criteria, and the player rated the highest shall be regarded as the most likely to win. Here we refer to an application of the AHP for the prediction of winners in chess matches—The AHP was used to evaluate the outcome of the Chess World Championships [118] as well as of the matches between Fischer and Spassky in 1972 and Karpov and Korchnoi in 1978 [116]. A possible hierarchy for this problem is represented in Figure 1.5. It is interesting to see that the values of the weights $w_1, \ldots, w_n$ in this sort of problems about forecasting can be interpreted as *subjective probabilities* [136]. For example, in this case, w_1 and w_2 could be interpreted as the subjective probabilities of the victories of the two chess players.

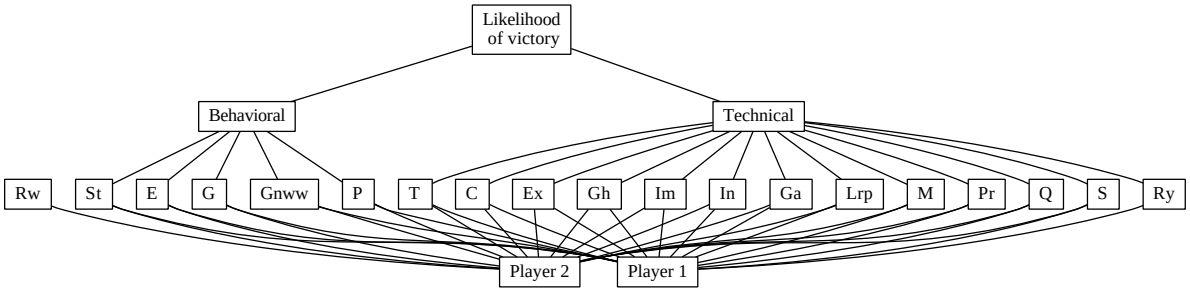


Figure 1.5.: A hierarchy for the chess competition problem. Abbreviations are as follows: Gamesmanship (G), Good Nerves and Will to Win (GN), Personality (P), Stamina (ST), Ego (E), Calculation (Q), Experience (EX), Good Health (GH), Imagination (IM), Intuition (IN), Game Aggressiveness (GA), Long Range Planning (LRP), Memory (M), Preparation (PR), Quickness (Q), Relative Youth (RY), Seconds (S), Technique (M). For a fuller description see the original paper [118].

Facility location

In Turku, a city in the South-West of Finland, the AHP has been used to find the best location the new ice hockey stadium (now called Turku Arena). Several criteria were used to evaluate different locations. Among the criteria one can find the accessibility of the arena, the possibility of having car parking, the quality of the soil on which the arena shall be built, and so forth. Carlsson and Walden [33] gave a frank political account of the decision process, which involved the local administration, and whose selected alternative was the third best ranked, and not the best.

1.3. Criticisms and open debates ⋆

Accounts of successful applications and empirical studies [71], have brought evidence on the AHP as an appealing method for decision making. Notwithstanding, as Shakespeare put it "All that glitters is not gold".
Thus, since any fair exposition must take into account its drawbacks and open issues, we should spell it out: the AHP is *not* a flawless method. Like the driver of a race car knows the limits of the machine, users of the AHP too shall be aware of its limitations and possible misuses. In this section we shall dwell on three of them. Further matters and open debates will be recalled later when they are related to specific topics of interest of some sections of this booklet. Even so, let us now focus on the three of them which can be already understood at this stage of the exposition.

Rank reversal

The most spirited criticisms against the AHP have been based on the *rank reversal* phenomenon. Since the treatise of Von Neumann and Morgenstern [91] some axioms

have been required to hold for decision analysis methodologies. One of these axioms requires that, if a new alternative is added to the original set of alternatives, then the order relation $\succ$ on the old set of alternatives should not change. Transposing this concept to our daily lives, if one has to select one meal and he prefers pasta to soup, when they offer him fish, this should *not* change his original preference of pasta to soup. Belton and Gear [14] proposed the following example to show that the AHP can suffer of rank reversal. Consider the matrices

$$\mathbf{A}^{(a)} = \begin{pmatrix} 1 & 1/9 & 1 \\ 9 & 1 & 9 \\ 1 & 1/9 & 1 \end{pmatrix} \qquad \mathbf{A}^{(b)} = \begin{pmatrix} 1 & 9 & 9 \\ 1/9 & 1 & 1 \\ 1/9 & 1 & 1 \end{pmatrix} \qquad \mathbf{A}^{(c)} = \begin{pmatrix} 1 & 8/9 & 8 \\ 9/8 & 1 & 9 \\ 1/8 & 1/9 & 1 \end{pmatrix}$$

which compare three alternatives with respect to three criteria, respectively—remember that a similar situation was proposed in the example of Figure 1.3. Assuming that the three criteria have equal weight, i.e. $1/3$, it follows that the final priority vector is $(0.45, 0.47, 0.08)^T$, and thus the alternatives are ranked $x_2 \succ x_1 \succ x_3$. So far so good, but suppose now that a new alternative, x_4, is added to the initial set, and the new judgments are

$$\mathbf{A}^{(a)} = \begin{pmatrix} 1 & 1/9 & 1 & 1/9 \\ 9 & 1 & 9 & 1 \\ 1 & 1/9 & 1 & 1/9 \\ 9 & 1 & 9 & 1 \end{pmatrix} \qquad \mathbf{A}^{(b)} = \begin{pmatrix} 1 & 9 & 9 & 9 \\ 1/9 & 1 & 1 & 1 \\ 1/9 & 1 & 1 & 1 \\ 1/9 & 1 & 1 & 1 \end{pmatrix} \qquad \mathbf{A}^{(c)} = \begin{pmatrix} 1 & 8/9 & 8 & 8/9 \\ 9/8 & 1 & 9 & 1 \\ 1/8 & 1/9 & 1 & 1/9 \\ 9/8 & 1 & 9 & 1 \end{pmatrix}.$$

Note that the preferences on the original three alternatives have been unchanged. However, still considering the criteria to be equally important, the new priority vector becomes $(0.37, 0.29, 0.06, 0.29)^T$ and thus the new ranking is $x_1 \succ x_2 \sim x_4 \succ x_3$. Now x_1 is ranked the best! The gravity of this drawback is made evident if we consider the initial example of the European city, where considering one more city, say Stockholm, might have changed the original ranking of the other three, let alone more important real-world problems. In a scientific context, and especially in decision analysis where everything should be justifiable, the rank reversal has been pivotal in the debate on the theoretical soundness of the AHP. On the other hand, many scholars ignore that the rank reversal is avoided if priority vectors are aggregated taking their component-wise geometric mean, instead of a convex linear combination. Although opposed by some [131], the use of this technique to avoid the rank reversal has been proven mathematically [11]. For a review of the rank reversal the interested reader can see the original discussion [14, 15, 118], a survey [89], and an account of the AHP versus Multi Attribute Utility Theory debate [61].

The nature of the AHP

The discovery of the rank reversal has been the spark for further discussions. One of the most important relates with the nature of the AHP. In other words, on what fundamental theory is the AHP based on? As whispered before, the AHP has something in common with *value theory*. In both these theories, there is a set of alternatives which

are eventually matched with real numbers such that $w_i \geq w_j \Leftrightarrow x_i \succeq x_j$. Ultimately, in value theory, there is a function $v : X \to \mathbb{R}$, where $v(x_i) = w_i$. It follows that, however complicated the function v is, rank reversal cannot happen, since alternatives are evaluated independently one from another and hence adding real or fictitious alternatives does not change the order of the existing ones. As seen before, this invariance is required by one axiom of value theory, which is violated by the AHP, because of the rank reversal. Thus, the existence of the rank reversal excludes that the AHP belongs to value theory. After Saaty's [105] attempt to axiomatize it and a debate initiated by Dyer [49], nowadays the AHP is considered to be grounded on *relative measurement theory* which can be seen as a theory where what matters are only the ratios between measurements of whatever entities under consideration. Hence, from a very high perspective, the AHP can be seen an a mathematical tool for relative measurement. The interested reader can refer to Saaty [104, 108] for an exposition of the AHP under this point of view and consider that, very recently, Bernasconi *et al.* [16] reinterpreted the AHP using the theory of psychological measurement.

Different scales

It does not take much to see that, in spite of the elegance of the relative measurement theory, a decision maker could have troubles to state that, under the climatic point of view, Barcelona is 4 times better than Reykjavik. In everyday life, people are more inclined to use linguistic expressions like "I *slightly* prefer pasta with salmon to pasta with cheese" or "I *strongly* prefer one banana to one apple". To help the decision maker, some linguistic expressions have been proposed and then linked to different values assignable to the entries a_{ij}. Hence, the decision maker can express opinions on pairs using linguistic terms, which are then associated to real numbers. In his original paper on the AHP, Saaty proposed an association between verbal judgments and values for pairwise comparisons. Other scales have been proposed and studied, among others, by Ji and Jiang [75] to which the reader can refer for a short overview. One of the foremost is the balanced scale proposed by Pöhjönen *et al.* [98]. The balanced scale, Saaty's scale, and their matching with verbal judgments are reported in Table 1.1.

Verbal description	Saaty's scale	Balanced scale
Indifference	1	1
—	2	1.22
Moderate preference	3	1.5
—	4	1.86
Strong preference	5	2.33
—	6	3
Very strong or demonstrated preference	7	4
—	8	5.67
Extreme preference	9	9

Table 1.1.: Two scales and their association with verbal judgments.

Which scale is better is still an open debate, but it is safe to say that, most likely, Saaty's scale is not optimal. It is a fact that it was introduced as a rule of thumbs, whereas other scales seem to have more supporting evidence. For instance, the balanced scale has been proposed on the basis of empirical experiments with people. Reasonably, this topic will require more research from the behavioral point of view than from the mathematical one.

One last remark is that, in spite of the open debate on the association between linguistic labels and numerical values, there is a meeting of minds on using bounded numerical scales, of which the most famous is the set of all integers up to 9 and their reciprocals,

$$\left\{\frac{1}{9}, \frac{1}{8}, \ldots, \frac{1}{2}, 1, 2, \ldots, 8, 9\right\}.$$

The main reason for this choice is our limited ability of processing information, also corroborated by psychological studies according to which our capacity of reckoning alternatives is limited to 7 ± 2 of them [112]. Nevertheless, although in practice this discrete scale is employed, in the following, unless otherwise specified, we shall not restrict the discussion and adopt the more general $\mathbb{R}_{>}$. In support of this approach comes also the fact that, mathematically speaking, the algebra of the AHP, and more generally of relative measurement theory, builds on positive real numbers.

2. Priority vector and consistency

> The average man's judgment is so poor, he runs a risk every time he uses it.
>
> Ed Howe

It is important to reflect on the fact that in the previous chapter, almost unconsciously, a number of very restrictive assumptions were imposed. Let us summarize them within one sentence, where the assumptions are highlighted in italic.

> A *single* decision maker is *perfectly rational* and can *precisely* express his preferences on *all pairs* of *independent* alternatives and criteria using *positive real numbers*.

Some of these assumptions had already been relaxed in Saaty's original works, and some others were relaxed later. In this and in the next chapter we shall present the ways in which these assumptions have been relaxed in the literature to provide the users of the AHP with a more flexible method. Everytime one assumption is relaxed, the previous box will be presented again and the assumption at stake emphasized in boldface. We are now ready to depart from a normative view on the AHP (how decisions should be made in a perfect world) to adopt a more descriptive view (how decisions are actually made).

2.1. Priority vector

We have seen that one pivotal step in the AHP is the derivation of a *priority vector* for each pairwise comparison matrix. Note that if each entry a_{ij} of the matrix is exactly the *ratio* between two weights w_i and w_j, then all the columns of $\mathbf{A}$ are proportional one another and consequently the weight vector is equal to any normalized column of $\mathbf{A}$ (see the matrices in Chapter 1). In this case the information contained in the matrix $\mathbf{A}$ can be perfectly synthesized in $\mathbf{w}$ and there is no loss of information.
However, we do not even bother dwelling on this case and technique to derive the weights, since it is hardly ever the case that a decision maker is so accurate and rational to give exactly the entries as ratios between weights. In this, and in the next section on consistency, we shall investigate how the AHP can cope with irrational pairwise

comparisons. Let us then represent again the box with the relaxed assumption now in boldface.

> A *single* decision maker is ***perfectly rational*** and can *precisely* express his preferences on *all pairs* of *independent* alternatives and criteria using *positive real numbers*.

When the entries of the matrix $\mathbf{A}$ are not obtained exactly as ratios between weights, there does *not* exist a weight vector which perfectly synthesize the information in $\mathbf{A}$. Nonetheless, since the AHP cannot make it without the weight vectors, it is necessary to devise some smart ways of estimating a 'good' priority vector.
Several methods for eliciting the priority vector $\mathbf{w} = (w_1, \ldots, w_n)^T$ have been proposed in the literature. Each method is just a rule for synthesizing pairwise comparisons into a rating, and mathematically is a function $\tau : \mathbb{R}_{>}^{n \times n} \to \mathbb{R}_{>}^{n}$. Clearly, different methods might lead to different priority vectors, except when the entries of the matrix are representable as ratios between weights, in which case all methods shall lead to the same vector $\mathbf{w}$. Needless to say, in the case of perfect rationality, the same vector $\mathbf{w}$ obtained with any method must be such that $(w_i/w_j)_{n \times n} = \mathbf{A}$.

2.1.1. Eigenvector method

The most popular method to estimate a priority vector is that proposed by Saaty himself, according to which the priority vector should be the principal eigenvector of $\mathbf{A}$. In linear algebra it is often called the Perron-Frobenius eigenvector, from the homonymic theorem [70]. The method stems from the following observation. Taking a matrix $\mathbf{A}$ whose entries are exactly obtained as ratios between weights and multiplying it by $\mathbf{w}$ one obtains

$$\mathbf{A}\mathbf{w} = \begin{pmatrix} w_1/w_1 & w_1/w_2 & \ldots & w_1/w_n \\ w_2/w_1 & w_2/w_2 & \ldots & w_2/w_n \\ \vdots & \vdots & \ddots & \vdots \\ w_n/w_1 & w_n/w_2 & \ldots & w_n/w_n \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} nw_1 \\ \vdots \\ nw_n \end{pmatrix} = n\mathbf{w}.$$

From linear algebra, we know that a formulation of the kind $\mathbf{A}\mathbf{w} = n\mathbf{w}$ implies that n and $\mathbf{w}$ are an *eigenvalue* and an *eigenvector* of $\mathbf{A}$, respectively [1]. Moreover, by knowing that the other eigenvalue of $\mathbf{A}$ is 0, and has multiplicity $(n-1)$, then we know that n is the largest eigenvalue of $\mathbf{A}$. Hence, if the entries of $\mathbf{A}$ are ratios between weights, then the weight vector is the eigenvector of $\mathbf{A}$ associated with the eigenvalue n. Saaty proposed to extend this result to all pairwise comparison matrices by replacing n with the more generic maximum eigenvalue of $\mathbf{A}$. That is, vector $\mathbf{w}$ can be obtained from any pairwise comparison matrix $\mathbf{A}$ as the solution of the following equation system,

$$\begin{cases} \mathbf{A}\mathbf{w} = \lambda_{\max}\mathbf{w} \\ \mathbf{w}^T\mathbf{1} = 1 \end{cases}$$

[1] A short overview of eigenvector theory in the AHP can be found in the Appendix.

where $\lambda_{\max}$ is the maximum eigenvalue of $\mathbf{A}$, and $\mathbf{1} = (1, \ldots, 1)^T$. Although this problem can easily be solved by mathematical software and also spreadsheets, its interpretation remains cumbersome for practitioners.

2.1.2. Geometric mean method

Another widely used method to estimate the priority vector is the *geometric mean method*, proposed by Crawford and Williams [43]. According to this method each component of $\mathbf{w}$ is obtained as the geometric mean of the elements on the respective row divided by a normalization term so that the components of $\mathbf{w}$ eventually add up to 1,

$$w_i = \left(\prod_{j=1}^{n} a_{ij}\right)^{\frac{1}{n}} \Bigg/ \underbrace{\sum_{i=1}^{n}\left(\prod_{j=1}^{n} a_{ij}\right)^{\frac{1}{n}}}_{\text{normalization term}}. \tag{2.1}$$

Example 2. *Let us take into account the following matrix*

$$\mathbf{A} = \begin{pmatrix} 1 & 1/2 & 1/4 & 3 \\ 2 & 1 & 1/2 & 2 \\ 4 & 2 & 1 & 2 \\ 1/3 & 1/2 & 1/2 & 1 \end{pmatrix} \tag{2.2}$$

for which, by using (2.1), one obtains

$$\mathbf{w} \approx (0.119, 0.208, 0.454, 0.219)^T$$

Problem 2. *Prove that, if $a_{ij} = w_i/w_j \ \forall i, j$, then the geometric mean method (2.1) returns the vector $\mathbf{w}$ whose ratios between components are the elements of $\mathbf{A}$.*

By looking at (2.1) it is apparent that the geometric mean method is very appealing for practical applications since, in contrast to the eigenvector method, the weights can be expressed as analytic functions of the entries of the matrix. Furthermore, even the final weights of the whole hierarchy can be expressed as analytic expressions of the entries of all the matrices in the hierarchy. This is particularly important since it opens avenues to perform efficiently some sensitivity analysis. Moreover, on a more mathematical note, it is interesting to note that the vector $\mathbf{w}$ obtained with this method, can equivalently be obtained as the argument minimizing the following optimization problem

$$\begin{aligned} \underset{(w_1,\ldots,w_n)}{\text{minimize}} \quad & \sum_{i=1}^{n}\sum_{j=1}^{n} \left(\ln a_{ij} + \ln w_j - \ln w_i\right)^2 \\ \text{subject to} \quad & \sum_{i=1}^{n} w_i = 1, \quad w_i > 0 \, \forall i \end{aligned} \tag{2.3}$$

Problem 3. *Prove that the argument optimizing (2.3) is the same vector (up to multiplication by a suitable scalar) which could be obtained with the geometric mean method.*

This optimization problem has some interpretations, the following being quite straightforward. We know that, in the case of perfect rationality, $a_{ij} = w_i/w_j \, \forall i, j$. Indeed, it is fair to consider $\sum_{i=1}^{n} \sum_{j=1}^{n} (a_{ij} - w_i/w_j)^2$ as a distance between $\mathbf{A}$ and the matrix $(w_i/w_j)_{n\times n}$ associated with the weight vector $\mathbf{w}$. Another metric can be found by using the natural logarithm ln, which is a monotone increasing function, thus obtaining $\sum_{i=1}^{n} \sum_{j=1}^{n} (\ln a_{ij} - \ln (w_i/w_j))^2$. The rest is done by observing that the logarithm of a quotient is the difference of the logarithms. Then the minimization problem (2.3) is introduced to find a suitable priority vector associated to a 'close' consistent approximation $(w_i/w_j)_{n\times n}$ of the matrix $\mathbf{A}$.

2.1.3. Other methods and discussion ⋆

A large number of alternative methods to compute the priority vector have been proposed in the literature. Choo and Wedley [39] listed 18 different methods and proposed a numerical and comparative study. Lin [83] reconsidered and simplified their framework. Another comparative study was offered by Ishizaka and Lusti [73]. Instead, Cook and Kress [40] presented a more axiomatic analysis where some desirable properties were stated. From all these studies it appears that, besides the eigenvector and the geometric mean method, other two methods have gained some popularity.

- The so-called *least squares method* where the priority vector is the argument solving the following optimization problem

$$\begin{aligned} \underset{(w_1,\ldots,w_n)}{\text{minimize}} \quad & \sum_{i=1}^{n} \sum_{j=1}^{n} \left(a_{ij} - \frac{w_i}{w_j} \right)^2 \\ \text{subject to} \quad & \sum_{i=1}^{n} w_i = 1, \quad w_i > 0 \, \forall i \, . \end{aligned} \tag{2.4}$$

  In spite of its elegance, this optimization problem can have local minimizers where the optimization algorithms get trapped. For a discussion on this method and its solutions the reader can refer to Bozóki [22].

- The other one is the *normalized columns method* which requires the normalization of all the columns of $\mathbf{A}$ so that the elements add up to 1 before the arithmetic means of the rows are taken and normalized to add up to 1 to yield the weights $w_1, \ldots, w_n$. This is the simplest method but lacks solid theoretical foundation.

  Example 3. *Consider the pairwise comparison matrix (2.2) already used to illustrate the geometric mean method. Then, the matrix with normalized columns and the priority vector are the following, respectively,*

$$\begin{pmatrix} 3/22 & 1/8 & 1/9 & 3/8 \\ 6/22 & 2/8 & 2/9 & 2/8 \\ 12/22 & 4/8 & 4/9 & 2/8 \\ 1/22 & 1/8 & 2/9 & 1/8 \end{pmatrix}, \qquad \mathbf{w} = \begin{pmatrix} 21/163 \\ 42/163 \\ 84/163 \\ 16/163 \end{pmatrix}.$$

Nevertheless, in spite of the great variety of methods, it is safe to say that the eigenvector and the geometric mean method have been the most used and therefore it is convenient to confine further discussions to these two. Saaty and Vargas [117] claimed the superiority of the eigenvector method and concluded that:

> In fact it is the only method that should be used when the data are not entirely consistent in order to make the best choice of alternative.

Saaty and Hu [111] proposed a theorem claiming the necessity of the eigenvector method, and Saaty [107] also proposed ten reasons for not using other methods. Fichtner [53] proposed some axioms and showed that the eigenvector method is the only one satisfying them. Curiously, supporters of the geometric mean method have used similar arguments. For instance, Barzilai at al. [10] proposed another axiomatic framework and proved that the geometric mean method is the only one which satisfies his axioms. Seemingly, the existence of two axiomatic frameworks leading to different conclusions suggest that the choice of the method depends on what set of properties we want the method to satisfy. Supporters of the geometric mean method also gave precise statements on the use of this method and, to summarize one of his papers, Barzilai [8] wrote:

> We establish that *the geometric mean is the only method* for deriving weights from multiplicative pairwise comparisons which satisfies fundamental consistency requirements.

Bana e Costa and Vansnick [41] also moved a criticism against the eigenvector method based on what they called the *condition of order preservation* (COP). The COP states that, if x_i more strongly dominates x_j than x_k does with x_l, it means that $a_{ij} > a_{kl}$, and then it is natural to expect that the priority vector be such that $w_i/w_j > w_k/w_l$. Formally,

$$a_{ij} > a_{kl} \Rightarrow \frac{w_i}{w_j} > \frac{w_k}{w_l} \quad \forall i, j, k, l.$$

Bana e Costa and Vansnick showed some examples of cases where, given a pairwise comparison matrix $\mathbf{A}$, the eigenvector method does not return a priority vector satisfying the COP, although there exists a set of other vectors satisfying it.

On a similar note, a recent discovery related to what economists call Pareto efficiency. The reasonable idea behind this is suggested also by (2.3) and (2.4) and is that, having estimated the priority vector $\mathbf{w}$, the matrix $(w_i/w_j)_{n\times n}$ should be as near as possible to the original preferences in $\mathbf{A}$. Blanquero *et al.* [18] showed that, if $\mathbf{w}$ is estimated by the eigenvector method, in some cases there exists a vector $\mathbf{v} = (v_1, \ldots, v_n)^T \neq \mathbf{w}$ such that

$$\left| \frac{v_i}{v_j} - a_{ij} \right| \leq \left| \frac{w_i}{w_j} - a_{ij} \right| \quad \forall i, j.$$

The fact that $\mathbf{w} \neq \mathbf{v}$ implies that the inequality is strict for some i, j. To summarize, this means that there can be vectors which are closer than the eigenvector to the preferences expressed in $\mathbf{A}$. At the time of writing this manuscript, it seems that in some cases the differences between $\mathbf{v}$ and $\mathbf{w}$ can be relevant [23].

2.2. Consistency

A perfectly rational decision maker should be able to state his pairwise preferences exactly, i.e. $a_{ij} = w_i/w_j \ \forall i, j$. So, let us consider the ramifications of this condition on the entries of the pairwise comparison matrix $\mathbf{A}$. If we write $a_{ij}a_{jk}$ and apply the condition $a_{ij} = w_i/w_j \ \forall i, j$, then we can derive the following

$$a_{ij}a_{jk} = \frac{w_i}{w_j}\frac{w_j}{w_k} = \frac{w_i}{w_k} = a_{ik}.$$

Hence, we discovered that, if all the entries of the pairwise comparison matrix $\mathbf{A}$ satisfy the condition $a_{ij} = w_i/w_j \ \forall i, j$, then the following condition holds [2],

$$a_{ik} = a_{ij}a_{jk} \quad \forall i, j, k\,, \tag{2.5}$$

which means that each direct comparison a_{ik} is exactly confirmed by all indirect comparisons $a_{ij}a_{jk} \ \forall j$. Formally, a decision maker able to give perfectly consistent pairwise comparisons does not contradict himself. A matrix for which this transitivity condition holds is called *consistent*.

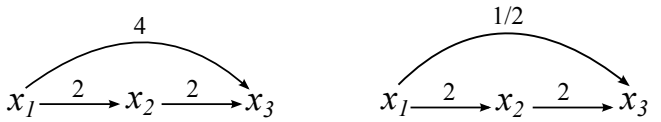
Example 4. *Consider the characteristic 'weight' of three stones x_1, x_2, x_3. If the decision maker says that x_1 is three times heavier than x_3 ($a_{13} = 3$), and then also says that x_1 is two times heavier than x_2 ($a_{12} = 2$), and x_2 is also two times heavier than x_3 ($a_{23} = 2$), then he contradicts himself, because he directly states that $a_{13} = 3$, but indirectly states that the value of a_{13} should be $a_{12}a_{23} = 2 \cdot 2 = 4$ and not 3.*

Evidently the whole reasoning can be translated into the language of pairwise comparison matrices.

Example 5. *Consider this other example with the two pairwise comparison matrices*

$$\begin{pmatrix} 1 & 2 & 4 \\ 1/2 & 1 & 2 \\ 1/4 & 1/2 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 1/2 \\ 1/2 & 1 & 2 \\ 2 & 1/2 & 1 \end{pmatrix}$$

for which we have the two diagrams in Figure 2.1 respectively.



(a) The matrix is consistent

(b) The matrix is inconsistent

Figure 2.1.: Examples of consistent and inconsistent transitivities.

[2] As we will see, the 'if' condition is in fact an 'if and only if'.

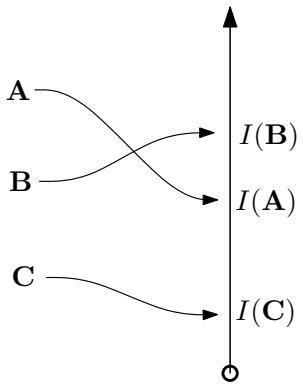


Figure 2.2.: An inconsistency index can be seen as a 'thermometer', which takes pairwise comparison matrices as inputs and evaluates how inconsistent the judgments are.

Being consistent is seldom possible because many factors can determine the emergency of inconsistencies. For instance, the decision maker might be asked to use integer numbers and their reciprocals; in this case if $a_{ij} = 3$ and $a_{jk} = 1/2$ it is impossible to find a consistent value for a_{ik}. Moreover, the number of independent transitivities (i, j, k) in a matrix of order n is equal to $\binom{n}{3}$, thus evidencing the difficulty of being fully consistent.

Example 6. *In a matrix of order 6, there are* $\binom{6}{3} = 20$ *independent transitivities; that is the number of possible assignments of values to* i, j, k *such that* $1 \leq i < j < k \leq 6$. *In a matrix of order 4, there are* $\binom{4}{3} = 4$ *transitivities. They are* $(1, 2, 3)$, $(1, 2, 4)$, $(1, 3, 4)$ *and* $(2, 3, 4)$.

In spite of the difficulty in being fully transitive, it is undeniable that consistency is a desirable property. In fact, an inconsistent matrix could be a symptom of the decision maker's incapacity or inexperience in the field. Additionally, it is possible to envision that violations of the condition of consistency (2.5) can be of different extent and gravity and imagine inconsistency as a gradual notion. Hence, on the ground that a matrix should deviate as less as possible from the condition of transitivity, a number of inconsistency indices have been proposed in the literature to quantify this deviation. Formally, an *inconsistency index* is a function mapping pairwise comparison matrices into the real line (see Figure 2.2 for an oversimplification).

There exist various inconsistency indices in the literature and this variety is in part justified by the fact that the condition of consistency can be formulated in many *equivalent* ways. Among them, it is the case to reckon the following four:

i) $a_{ik} = a_{ij}a_{jk}\ \forall i, j, k$,

ii) There exists a vector $(w_1, \ldots, w_n)^T$ such that $a_{ij} = w_i/w_j\ \forall i, j$,

iii) The columns of $\mathbf{A}$ are proportional, i.e. $\mathbf{A}$ has rank one,

iv) The pairwise comparison matrix $\mathbf{A}$ has its maximum eigenvalue, $\lambda_{\max}$, equal to n.

In this section we explore some inconsistency indices, each inspired by one of these equivalent consistency conditions.

2.2.1. Consistency index and consistency ratio

According to the result that given a pairwise comparison matrix $\mathbf{A}$, its maximum eigenvalue, $\lambda_{\max}$, is equal to n if and only if the matrix is consistent (and greater than n otherwise), Saaty [101] proposed the *Consistency Index*

$$CI(\mathbf{A}) = \frac{\lambda_{\max} - n}{n-1}. \tag{2.6}$$

However, numerical studies showed that the expected value of CI of a random matrix of size $n+1$ is, on average, greater than the expected value of CI of a random matrix of order n. Consequently, CI is not fair in comparing matrices of different orders and needs to be rescaled.
The *Consistency Ratio*, CR, is the rescaled version of CI. Given a matrix of order n, CR can be obtained dividing CI by a real number RI_n (*Random Index*) which is nothing else but an estimation of the average CI obtained from a large enough set of randomly generated matrices of size n. Hence,

$$CR(\mathbf{A}) = \frac{CI(\mathbf{A})}{RI_n} \tag{2.7}$$

Estimated values for RI_n are reported in Table 2.1. Note that the generation of random matrices requires the definition of a bounded scale where the entries take values, for instance the interval $[1/9, 9]$. According to Saaty [102], in practice one should accept matrices with values $CR \leq 0.1$ and reject values greater than 0.1. A value of $CR = 0.1$ means that the judgments are 10% as inconsistent as if they had been given randomly.

n	3	4	5	6	7	8	9	10
RI_n	0.5247	0.8816	1.1086	1.2479	1.3417	1.4057	1.4499	1.4854

Table 2.1.: Values of RI_n [3].

Example 7. *Consider the pairwise comparison matrix*

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 9 & 1 \\ 1/2 & 1 & 1/3 & 1/6 \\ 1/9 & 3 & 1 & 2 \\ 1 & 6 & 1/2 & 1 \end{pmatrix}. \tag{2.8}$$

It can be calculated that its maximum eigenvalue is $\lambda_{\max} = 5.28$. *Using the formula for* CI, *we obtain* $CI(\mathbf{A}) = 0.42667$. *Dividing it by* RI_4 *one obtains* $CR(\mathbf{A}) \approx 0.48$ *which is significantly greater than the threshold* 0.1. *In a decision problem it is common practice to ask the decision maker to revise his judgments until a value of* CR *smaller than* 0.1 *is reached.*

2.2.2. Index of determinants

The index of determinants was proposed by Peláez and Lamata [95] and comes from the following property of a matrix of order three. Expanding the determinant of a 3×3 real matrix one obtains

$$\det(\mathbf{A}) = \frac{a_{13}}{a_{12}a_{23}} + \frac{a_{12}a_{23}}{a_{13}} - 2.$$

If $\mathbf{A}$ is not consistent, then $a_{13} \neq a_{12}a_{23}$ and $\det(\mathbf{A}) > 0$, because, in general, $\frac{a}{b} + \frac{b}{a} - 2 > 0 \ \ \forall\, a \neq b,\ a, b > 0$.
It is possible to generalize this result to matrices of order greater than three and define this inconsistency index as the average of the determinants of all the possible submatrices $\mathbf{T}_{ijk}$ of a given pairwise comparison matrix, constructed as follow,

$$\mathbf{T}_{ijk} = \begin{pmatrix} 1 & a_{ij} & a_{ik} \\ a_{ji} & 1 & a_{jk} \\ a_{ki} & a_{kj} & 1 \end{pmatrix}, \quad \forall i < j < k.$$

The number of so constructed submatrices is $\binom{n}{3} = \frac{n!}{3!(n-3)!}$. The result is an index whose value is the average inconsistency computed for all the submatrices $\mathbf{T}_{ijk}$ $(i < j < k)$

$$CI^*(\mathbf{A}) = \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} \underbrace{\left(\frac{a_{ik}}{a_{ij}a_{jk}} + \frac{a_{ij}a_{jk}}{a_{ik}} - 2 \right)}_{\det(\mathbf{T}_{ijk})} \Bigg/ \binom{n}{3}. \tag{2.9}$$

Example 8. *Consider the matrix* $\mathbf{A}$ *in (2.8). It is then possible to calculate the average of the determinants of all the submatrices* $\mathbf{T}_{i\,j\,k}$ *with* $i < j < k$.

$$\begin{aligned} CI^*(\mathbf{A}) =& \frac{\overbrace{\det \begin{pmatrix} 1 & 2 & 9 \\ 1/2 & 1 & 1/3 \\ 1/9 & 3 & 1 \end{pmatrix}}^{\mathbf{T}_{1\,2\,3}} + \cdots + \overbrace{\det \begin{pmatrix} 1 & 1/3 & 1/6 \\ 3 & 1 & 2 \\ 6 & 1/2 & 1 \end{pmatrix}}^{\mathbf{T}_{2\,3\,4}}}{4} \\ =& (11.5741 + 1.3333 + 16.0556 + 34.0278)/4 = 15.7477. \end{aligned}$$

Interestingly, CI^* is proportional to another inconsistency index called c_3 [27]. The coefficient c_3 of the characteristic polynomial of a pairwise comparison matrix was proposed to act as an inconsistency index by Shiraishi and Obata [124] and Shiraishi *et al.* [125, 126]. By definition, the characteristic polynomial[3] of a matrix $\mathbf{A}$ has the following form

$$P_{\mathbf{A}}(\lambda) = \lambda^n + c_1\lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n,$$

with $c_1, \ldots, c_n$ that are real numbers and λ the unknown. Shiraishi *et al.* [125] proved that, if $c_3 < 0$, then the matrix cannot be fully consistent. In fact, this is evident if one

[3]See appendix on eigenvalues and eigenvectors

reckons that—in light of the Perron-Frobenius theorem—the only possible formulation of the characteristic polynomial which yields $\lambda_{\max} = n$, is

$$P_{\mathbf{A}}(\lambda) = \lambda^{n-1}(\lambda - n). \tag{2.10}$$

Thus, the presence of $c_3 < 0$ contradicts this last formulation and is certainly a symptom of inconsistency. Moreover, Shiraishi *et al.* [125] also proved that c_3 has the following analytic expression

$$c_3 = \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} \left(2 - \frac{a_{ik}}{a_{ij}a_{jk}} - \frac{a_{ij}a_{jk}}{a_{ik}} \right) \tag{2.11}$$

which highlights its proportionality with CI^*.

2.2.3. Geometric consistency index

This index was introduced by Crawford [42] and reexamined by Aguarón and Moreno-Jiménez [2]. It considers the priority vector to be estimated by the geometric mean method (2.1). With the so estimated weights it is possible to build a local quantification of inconsistency e_{ij} for each entry a_{ij} such that

$$e_{ij} = a_{ij}\frac{w_j}{w_i}, \quad i, j = 1, \ldots, n. \tag{2.12}$$

Obviously, for consistent matrices the value of e_{ij} is equal to 1 because it becomes the multiplication of an entry times its reciprocal. Note that,

$$a_{ij} = \frac{w_i}{w_j} \;\Rightarrow\; \ln e_{ij} = 0.$$

It is now possible to define an index of global inconsistency as the normalized sum of the local contributions to the inconsistency of $\mathbf{A}$. This index of global inconsistency, the Geometric Consistency Index (GCI), is the following:

$$GCI(\mathbf{A}) = \frac{2}{(n-1)(n-2)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left(\ln e_{ij} \right)^2 . \tag{2.13}$$

Example 9. *We refer to the matrix* $\mathbf{A}$ *presented in (2.8). By using the geometric mean method, the priority vector is* $\mathbf{w} \approx (2.06, 0.408, 0.904, 1.316)^T$. *Next, it could be convenient to collect values* e_{ij} *obtained with (2.12) into the following auxiliary matrix*

$$\mathbf{E} = (e_{ij})_{n\times n} = \left(a_{ij}\frac{w_j}{w_i} \right)_{n\times n} = \begin{pmatrix} 1 & 0.3964 & 3.9482 & 0.6389 \\ 2.5227 & 1 & 0.7379 & 1.8612 \\ 0.2538 & 1.3554 & 1 & 2.9129 \\ 1.5651 & 1.8612 & 0.3432 & 1 \end{pmatrix}.$$

Finally, the last computation is achieved applying formula (2.13), which yields $GCI(\mathbf{A}) \approx 1.52$.

2.2.4. Harmonic consistency index

If and only if $\mathbf{A}$ is a consistent pairwise comparison matrix, then its columns are proportional and $\text{rank}(\mathbf{A}) = 1$. Therefore, it is fair to suppose that the less proportional are the columns, the less consistent is the matrix. An index of inconsistency loosely based on proportionality between columns was then proposed by Stein and Mizzi [127]. Given a matrix $\mathbf{A}$, they proposed to construct an auxiliary vector $\mathbf{s} = (s_1, \ldots, s_n)^T$ with $s_j = \sum_{i=1}^{n} a_{ij}\ \forall j$. It was proven that $\sum_{j=1}^{n} s_j^{-1} = 1$ if and only if $\mathbf{A}$ is consistent, and smaller than 1 otherwise. The harmonic mean of the components of vector $\mathbf{s}$ is then the result of the following

$$HM(\mathbf{s}) = \frac{n}{\sum_{j=1}^{n} \frac{1}{s_j}}. \tag{2.14}$$

The function HM itself could be an index of inconsistency, but the authors, according to computational experiments, proposed a normalization in order to align its behavior with that of CI. The Harmonic Consistency Index is then

$$HCI(\mathbf{A}) = \frac{(HM(\mathbf{s}) - n)(n+1)}{n(n-1)}. \tag{2.15}$$

Example 10. *Considering the matrix* $\mathbf{A}$ *in (2.8), then the vector* $\mathbf{s}$ *is*

$$\mathbf{s} = \left(\frac{47}{18}, 12, \frac{65}{6}, \frac{25}{6}\right)^T$$

whose harmonic mean is

$$HM(\mathbf{s}) = \frac{4}{\frac{1}{\frac{47}{18}} + \frac{1}{12} + \frac{1}{\frac{65}{6}} + \frac{1}{\frac{25}{6}}} = \frac{733200}{146387} = 5.00864$$

Now it is possible to derive the value of the harmonic consistency index by plugging the value $HM(\mathbf{s})$ *into (2.15) and obtain* $HCI(\mathbf{A}) \approx 0.42$.

2.2.5. Ambiguity index

Salo and Hämäläinen [121, 122] proposed an *ambiguity index* which can be used as an inconsistency index too. It requires the construction of an auxiliary interval-valued matrix

$$\bar{\mathbf{A}} = (\bar{a}_{ij})_{n\times n} = \begin{pmatrix} [a_{11}^L, a_{11}^R] & \ldots & [a_{1n}^L, a_{1n}^R] \\ \vdots & \ddots & \vdots \\ [a_{n1}^L, a_{n1}^R] & \ldots & [a_{nn}^L, a_{nn}^R] \end{pmatrix}$$

where $a_{ij}^L = \min\{a_{ik}a_{kj} | k = 1, \ldots, n\}$ and $a_{ij}^R = \max\{a_{ik}a_{kj} | k = 1, \ldots, n\}$. Namely, each interval's lower (upper) bound a_{ij}^L (a_{ij}^R) is the smallest (greatest) possible value assignable to the entry if it was estimated indirectly using a transitivity of the pairwise comparison matrix. Surely, if $\mathbf{A}$ is consistent, then all the intervals collapse into real

numbers. From this insight, Salo and Hämäläinen deduced that the wider the intervals, the more inconsistent $\mathbf{A}$ should be. Hence, they presented their consistency measure,

$$CM(\mathbf{A}) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{a_{ij}^R - a_{ij}^L}{\left(1 + a_{ij}^R\right)\left(1 + a_{ij}^L\right)},$$

Example 11. *Salo and Hämäläinen proposed the following simple example [122]. Consider*

$$\mathbf{A} = \begin{pmatrix} 1 & 7 & 4 \\ 1/7 & 1 & 1/5 \\ 1/4 & 5 & 1 \end{pmatrix}$$

from which one can derive

$$\bar{\mathbf{A}} = \begin{pmatrix} 1 & [7, 20] & [7/5, 4] \\ [1/20, 1/7] & 1 & [1/5, 4/7] \\ [1/4, 5/7] & [7/4, 5] & 1 \end{pmatrix}. \tag{2.16}$$

It follows that

$$CM(\mathbf{A}) = \frac{2}{3 \cdot 2} \left(\frac{20 - 7}{(20 + 1) \cdot (7 + 1)} + \frac{4 - \frac{7}{5}}{(4 + 1) \cdot (7 + 1)} + \frac{\frac{4}{5} - \frac{1}{5}}{(\frac{4}{7} + 1) \cdot (\frac{1}{5} + 1)} \right) = 0.16\,.$$

2.2.6. Other indices and discussion ⋆

Many other inconsistency indices have been proposed. For instance, Koczkodaj [80] proposed an inconsistency index for matrices of order three which was later extended to matrices of greater order [48]. Golden and Wang formulated an index which considers a metric between the normalized columns of the matrix and the priority vector obtained either with the eigenvector or the geometric mean method [63]. Cavallo and D'Apuzzo proposed an interpretation of pairwise comparison matrices using group theory and introduced their own index [35, 36]. Barzilai, first transformed the entries of the matrix by means of a logarithmic function and then proposed another index [9]. Gass and Rapcsák [62] defined an index based on the singular value decomposition of matrices. Furthermore, consider that even the objective functions of the optimization problems of the logarithmic least squares (2.3) and the least squares (2.4) used in §2.1 to derive the priority vector can be considered inconsistency index. The interested reader can refer to a survey paper with numerical tests on various indices [26].

More broadly, and to include the most updated results, it is the case to remark that recently some questions have been answered.

- Some questions were open on the behavioral side of consistency. For instance, does the order in which the comparisons are asked affect the final inconsistency? Does inconsistency increase with the order of the matrix? These, and other questions, have been answered by means of empirical experiments with real decision makers [24].

- It used to be unclear whether different inconsistency indices were similar or choosing one or another really made a difference. Namely, the formulations of the indices are often so dissimilar that we cannot understand if they tend to give similar results. By means of numerical simulations it was discovered that some indices are very similar whereas some others can give very different results [26]. Curiously, some indices have even been proved to be proportional, and thus equivalent, in estimating inconsistency [27].

- Inconsistency indices have been introduced empirically and an inquisitive reader might not take their validity for granted. Clearly, functions like the product of all the entries of a matrix, $\Pi(\mathbf{A})$, or the trace of a matrix $\mathrm{tr}(\mathbf{A})$ cannot capture the inconsistency of a matrix. Five axioms were proposed and considered necessary to characterize inconsistency indices and it was proved that, in fact, some inconsistency indices fail to satisfy some of them and can be suspected of ill-behavior in some situations [28]. Figure 2.3 is a snapshot of the axiomatic system and its role.

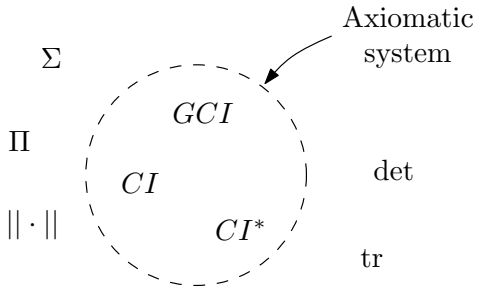


Figure 2.3.: The axiomatic system [28] defines a set of functions suitable to estimate inconsistency

In the literature it is often assumed that one inconsistency index together with a threshold value should be used to test whether the inconsistency of a matrix is tolerable or not. At least three points can be raised in connection with this standard procedure.
Firstly, unlike for the Consistency Index, thresholds have been rarely presented for inconsistency indices. That is, indices were introduced *tout court*, without discussions on acceptance rules for sufficiently consistency matrices. Hence, presently, without thresholds the use of many indices is limited to stating if a matrix is more (or less) inconsistent that another.
Secondly, a wise proposal would be that of using two indices, perhaps quite dissimilar, to test if a matrix is not too inconsistent. New acceptance rules might be devised too; considering two indices, one could for instance accept matrices respecting the threshold value for both the first *and* the second inconsistency index.
Thirdly, it is simple to envision that a decision maker can hardly ever be completely consistent, and therefore, instead of requiring him to achieve a sufficiently low value for an inconsistency index, one might want to lower the bar and introduce less demanding conditions which can realistically be fully satisfied. This kind of reasoning has pushed some authors to research on weaker conditions of consistency. One natural way to force

transitivity, but in a weaker sense, would be that of doing without the degrees of preference and simply require that if a decision maker prefers x_i to x_j and x_j to x_k, then it should also prefer x_i to x_k. This condition, which is nothing else but a restatement of ordinal transitivity for binary relations, can be formalized as follows:

$$a_{ij} > 1 \text{ and } a_{jk} > 1 \Rightarrow a_{ik} > 1 \quad \forall i, j, k.$$

This condition can be strengthened into the more restrictive *weak consistency* condition:

$$a_{ij} > 1 \text{ and } a_{jk} > 1 \Rightarrow a_{ik} > \max\{a_{ij}, a_{jk}\} \quad \forall i, j, k.$$

A deeper analysis of weaker consistency conditions, and their implications on the stability of the ranking of alternatives, can be found in [12, 13].

3. Missing comparisons and group decisions

3.1. Missing comparisons

Having, and manipulating, a complete and consistent pairwise comparison matrix means dealing with rich and reliable information and therefore it represents the most desirable situation in a decision making problem with the AHP. However, sometimes, it is not possible for the decision maker to express all the pairwise comparisons and therefore, it is nowadays common practice to accept that some entries of a pairwise comparison matrix be missing. Let us now reprise the famous box and highlight the assumption that we are going to relax

> A *single* decision maker is *perfectly rational* and can *precisely* express his preferences on ***all pairs*** of *independent* alternatives and criteria using *positive real numbers*.

In complex problems like those considered in §1.2, it may happen that the decision maker cannot complete a preference relation due to lack of time, the typology of problem, his incapacity in comparing two alternatives of different nature, and so forth [29]. Besides, sometimes, even if it was possible to obtain all the pairwise comparisons, doing so could be discouraged, since, due to information overload, the last ones could be given with less attention and care [34]. Certainly, it might be better to have few pairwise comparisons carefully given, than many, but given with scarce attention.
Hereafter, we shall call *incomplete pairwise comparison matrix* any pairwise comparison matrix with some missing entries. Considering all the diagonal elements of the matrix as given and the fact that, thanks to reciprocity, we only need to know a_{ij} to derive its reciprocal a_{ji}, then a pairwise comparison matrix of order n requires $n(n-1)/2$ *independent* comparisons.

Example 12. *Consider a decision problem with 6 alternatives evaluated with respect to 5 criteria. Then, the number of independent comparisons is*

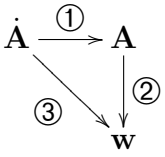
$$5\frac{6(6-1)}{2} + \frac{5(5-1)}{2} = 75 + 10 = 85.$$

Problem 4. *How many independent comparisons are required by the hierarchy in Figure 1.4 used in Problem 1?*

All in all, the range of reasons for leaving a matrix incomplete is wide and the real problem is how to derive a priority vector when there is not full information about the preferences on alternatives. In fact, we have seen that the eigenvector method and the geometric mean method were defined only for complete matrices. Several methods have been implemented to face this problem and, despite their diversities, considering $\mathbf{A}$ and $\dot{\mathbf{A}}$ to be a complete and an incomplete pairwise comparison matrix, respectively, the decision maker can proceed in one of the two following alternative ways:

- Complete the matrix by means of the information provided by the existing comparisons, ①. This operation is usually carried out following some principles of consistency, in the sense that the missing comparisons should be as coherent as possible with the existing ones. Having done this, it is possible to estimate the priority vector by means of one of the methods discussed earlier in §2.1 ②
- Estimate directly the priority vector by means of some modified algorithms which work even when some comparisons are missing, ③.

These two ways of proceeding are illustrated in the diagram below.



(3.1)

The following two subsections describe an algorithm of the first kind as well as one of the second kind.

3.1.1. Optimization of the coefficient c_3

The name c_3 refers to the coefficient of λ^{n-3} in the characteristic polynomial of the matrix $\mathbf{A}$. Shiraishi *et al.* [126] observed that c_3 can be considered an inconsistency index for a pairwise comparison matrix. This was already discussed in §2.2.2 and the analytic formula of c_3 was given in (2.11). Then, in order to complete $\dot{\mathbf{A}}$ following a principle of consistency, the authors considered the m missing comparisons as variables $\alpha_1, \ldots, \alpha_m$ and proposed to maximize c_3 (reckon that the greater c_3 the smaller the inconsistency) as a function of these variables, thus obtaining the values of the missing comparisons by solving

$$\begin{aligned} &\underset{(\alpha_1,\ldots,\alpha_m)}{\text{maximize}} \quad c_3 \\ &\text{subject to} \quad \alpha_1, \ldots, \alpha_m > 0\,. \end{aligned} \tag{3.2}$$

Note that there always exists an optimum for (3.2), but when there are too many missing comparisons uniqueness is not guaranteed [124].

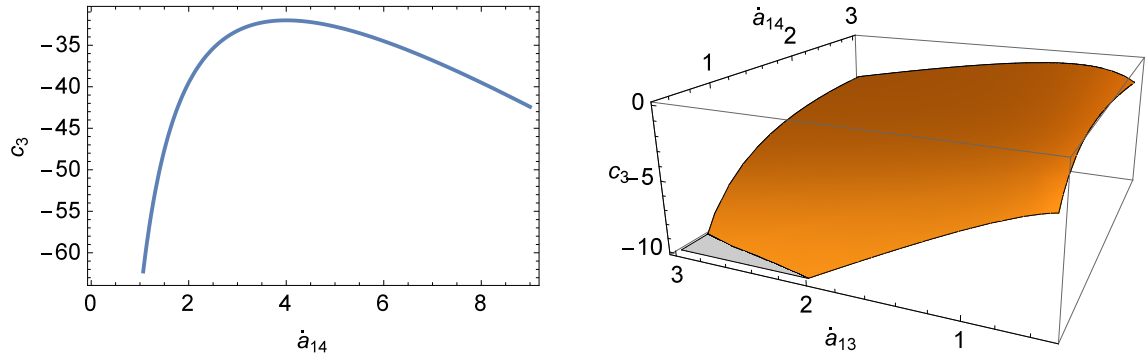
Example 13. *First, we present an incomplete pairwise comparison matrix* $\dot{\mathbf{A}}$

$$\dot{\mathbf{A}} = \begin{pmatrix} 1 & 2 & 4 & \dot{a}_{14} \\ 1/2 & 1 & 1/3 & 1 \\ 1/4 & 3 & 1 & 2 \\ 1/\dot{a}_{14} & 1 & 1/2 & 1 \end{pmatrix}. \tag{3.3}$$

Its missing comparison can be estimated by solving (3.2) with $\alpha_1 = \dot{a}_{14}$*. The plot of* c_3 *as a function of* $\dot{a}_{14}$ *is depicted in Figure 3.1a. The optimal solution is* $\dot{a}_{14} = 4$*. If we further assume that also* $\dot{a}_{13}$ *is missing, then the new incomplete pairwise comparison matrix becomes*

$$\dot{\mathbf{A}} = \begin{pmatrix} 1 & 2 & \dot{a}_{13} & \dot{a}_{14} \\ 1/2 & 1 & 1/3 & 1 \\ 1/\dot{a}_{13} & 3 & 1 & 2 \\ 1/\dot{a}_{14} & 1 & 1/2 & 1 \end{pmatrix}. \tag{3.4}$$

and the plot of c_3 *as a function of both* $\dot{a}_{13}$ *and* $\dot{a}_{14}$ *is in Figure 3.1b. In this case, the optimal solution is* $\dot{a}_{13} \approx 0.763143$ *and* $\dot{a}_{14} \approx 1.74716$*.*



(a) Plot of c_3 as a function of $\dot{a}_{14}$

(b) Plot of c_3 as a function of $\dot{a}_{14}$ and $\dot{a}_{13}$

Figure 3.1.: Plots of c_3

Problem 5. *Consider the matrix* $\dot{\mathbf{A}}$ *in (3.3). Find a way to recover* $\dot{a}_{14} = 4$ *as the* analytic *solution of the nonlinear optimization problem (3.2).*

The optimization of the coefficient c_3 provides values for the missing entries, and, upon completion of the matrix, it becomes straightforward to derive a priority vector.

3.1.2. Revised geometric mean method

This method, proposed by Harker [68], is not explicitly based on the optimization of an objective function, but refers to the eigenvector approach by Saaty. Practically, it extends Saaty's approach to non-negative quasi-reciprocal matrices, in order to apply it to the case of incomplete preferences. Unlike the optimization of c_3 proposed in (3.2), this method does not reconstruct the matrix but instead finds a priority vector using

less information. The algorithm requires to construct an auxiliary matrix $\mathbf{C} = (c_{ij})_{n\times n}$ as follows

$$c_{ij} = \begin{cases} 1 + m_i, & \forall i = j \\ \dot{a}_{ij}, & \forall i \neq j \text{ and } \dot{a}_{ij} \text{ not missing} \\ 0, & \dot{a}_{ij} \text{ is missing} \end{cases}$$

where m_i is the number of missing comparisons on the ith row. Having done this, the priority vector can be estimated by means of the eigenvector method. The following case, proposed by Harker [68], provides a numerical toy example and more insight on the method.

Example 14. *Consider the following pairwise comparison matrix,*

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & \dot{a}_{13} \\ 1/2 & 1 & 2 \\ \dot{a}_{31} & 1/2 & 1 \end{pmatrix},$$

and replace the missing comparison and its reciprocal entry with their theoretical values w_i/w_j so that the new matrix $\mathbf{B}$ is

$$\mathbf{B} = \begin{pmatrix} 1 & 2 & \frac{w_1}{w_3} \\ 1/2 & 1 & 2 \\ \frac{w_3}{w_1} & 1/2 & 1 \end{pmatrix}.$$

It is possible to observe what is obtainable through the operation $\mathbf{Bw}$

$$\mathbf{Bw} = \begin{pmatrix} 1 & 2 & \frac{w_1}{w_3} \\ 1/2 & 1 & 2 \\ \frac{w_3}{w_1} & 1/2 & 1 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} 2w_1 + 2w_2 \\ w_1/2 + w_2 + 2w_3 \\ w_2/2 + 2w_3 \end{pmatrix}.$$

We can reach the same result considering $\mathbf{Cw}$ with

$$\mathbf{C} = \begin{pmatrix} 2 & 2 & 0 \\ 1/2 & 1 & 2 \\ 0 & 1/2 & 2 \end{pmatrix}.$$

Finally, we can certainly state that

$$\mathbf{Bw} = \mathbf{Cw}. \tag{3.5}$$

We can proceed with the elicitation of weights, extending what was stated above. Therefore,

$$\mathbf{Bw} = \mathbf{Cw} = \lambda_{\max}\mathbf{w}. \tag{3.6}$$

Since $\mathbf{B}$ has some non-numerical entries, we can solve the eigenvector problem for $\mathbf{C}$. Needless to say, the result is $\mathbf{w} = (4, 2, 1)^T$.

Problem 6. *Formulate the auxiliary matrix $\mathbf{C}$ associated with $\mathbf{A}$ in (3.4).*

3.1.3. Other methods and discussion ⋆

There are other methods, of both types. For instance, Harker proposed the application of the concept of *connecting path*. A connecting path of length r between i and j is a product $a_{i\,i_1}a_{i_1 i_2}\cdots a_{i_{r-2}i_{r-1}}a_{i_{r-1}j}$, where the special case with $r = 2$ collapses to the consistency condition $a_{ik} = a_{ij}a_{jk}$. Harker proposed to compute the missing entries taking the geometric mean of all their connecting paths. Although sound, this proposal suffers of computational complexity when the size of the matrix is large enough, and is difficult to implement when several comparisons are missing.

It seems that one natural way to estimate missing comparisons is that of using some principles of consistency. For example, an inconsistency index can be optimized and the missing comparisons be used as variables. True, the foremost inconsistency index has been the CI which, fixed a value of n, is a positive affine transformation of the maximum eigenvalue $\lambda_{\max}$, which in turn is a root of a polynomial of degree n, hence impossible to be expressed analytically, except in few cases. In spite of this problem, using some convexity properties, Bozóki *et al.* [25] were able to formulate an optimization problem and a special algorithm to minimize $\lambda_{\max}$ keeping the missing comparisons as variables. It is indeed a very valuable proposal, but remains too cumbersome to be explained in this booklet.

On a more general level, a deeper discussion on missing comparisons goes back to the philosophy of the AHP and question how many comparisons the decision maker should provide. Is it carved in stone that the matrix has to be complete? Can, instead, some comparisons be missing? How many? This question has at least two possible answers, one algorithmic and one connected with common sense.

- From the *algorithmic* point of view, different methods for dealing with incomplete pairwise comparisons give different answers. Considering the revised geometric mean method presented in §3.1.2 one can observe that, in fact, this method works even when all the nondiagonal entries of $\dot{\mathbf{A}}$ are missing, in which case it returns a priority vector where all the weights are equal. More generally, it was also shown that one needs only $(n-1)$ independent comparisons to complete a matrix and make it consistent in a univocal way [69]. In fact, the knowledge of a set of comparisons, as for instance the set of entries right above the main diagonal, $\{a_{12}, a_{23}, \ldots, a_{n-1\,n}\}$, or the set of non-diagonal entries on, say, the first row, $\{a_{12}, a_{13}, \ldots, a_{1n}\}$ suffices to reconstruct the missing entries using the condition $a_{ik} = a_{ij}a_{jk}\ \forall i, j, k$.

  Problem 7. *Consider the following matrix* $\dot{\mathbf{A}}$ *and reconstruct its missing entries using the consistency condition* $a_{ik} = a_{ij}a_{jk}\ \ \forall i, j, k$.

$$\dot{\mathbf{A}} = \begin{pmatrix} 1 & 2 & 4 & 3 \\ 1/2 & 1 & \dot{a}_{23} & \dot{a}_{24} \\ 1/4 & 1/\dot{a}_{23} & 1 & \dot{a}_{34} \\ 1/3 & 1/\dot{a}_{24} & 1/\dot{a}_{34} & 1 \end{pmatrix}$$

- On the *common sense* side of the subject matter, one would surely refrain from estimating the priority vector from a totally incomplete matrix, and probably even question the convenience of reconstructing a (consistent) matrix from $(n-1)$ comparisons, since at that point the original $(n-1)$ comparisons would have been sufficient to estimate the priority vector directly. Moreover, by giving up a large number of comparisons one also gives up the possibility to estimate the inconsistency of preferences and thus to detect potential flaws in the decision maker's judgments. One last reason for not leaving too many comparisons missing is that evaluation errors can better compensate, and tend to cancel each other, when there are many comparisons then when there are few.

Another open question regards what comparisons should be elicited and what can be left missing. For example, knowing that the decision maker is willing to express his opinions on a subset of pairwise comparisons, but not all, then which ones should he express, and in what order? The quest for optimal completion rules and optimal completion paths has inspired some papers, as for instance those by Harker [67] and Fedrizzi and Giove [52].

There are various research papers on methods for dealing with incomplete preferences but very few investigated the relation between the number of missing comparisons and the stability of the obtained priority vector. One of these rare studies was by Carmone *et al.* [34] and it is safe to say that there is need and space for further investigation.

3.2. Group decisions

A further assumption was made regarding the number of decision makers: so far opinions have been given by a single decision maker. Even in the introductory exposition of the AHP given in Chapter 1, in the example of the European city, the family was considered as an unique entity and we did not account for possibly different opinions of family members. However, in many real-world contexts, decisions are made by groups of people, committees, boards, teams of experts, and so forth. Whenever there is a multitude of experts bringing diverse evidence on a problem, it is good practice to account for them.

> A ***single*** decision maker is *perfectly rational* and can *precisely* express his preferences on *all pairs* of *independent* alternatives and criteria using *positive real numbers*.

In his popular book *The Wisdom of the Crowd*, Surowiecki [128] argued that collective intelligence often outperforms individual one. An anecdote, originally by Galton [60], is reported in the introductory part of the book: at a county fair, individuals were asked to estimate the exact weight of an ox. Remarkably, by averaging the responses of the crowd members, they could obtain an estimate of the weight of the ox which was closer to the true weight then any of the individual judgments which were instead given by a

number of cattle experts. In other words, the collective wisdom of the crowd was more accurate than the estimates of true experts in the field. The problem of the ox was a problem of measurement, and we should not forget that the AHP is a theory of relative measurement.
The AHP for group decisions has been acclaimed by some researchers, for instance Dyer and Forman [50]. Peniwati [96] proposed some desirable properties, e.g. 'Scientific and mathematical generality', 'psychophysical applicability', and 'applicability to conflict resolution', for group decisions with MCDM methods, and according to her qualitative analysis, the AHP is a valuable decision methodology for group decisions.
The AHP can be used in many different ways as a group decision making method and it can be implemented in the so-called *Delphi method* [85]. In a nutshell, the Delphi method prescribes a number of meetings led by a moderator, where, after each meeting, the decision makers can revise their opinions. The role of the moderator is to make the opinions of different decision makers converge towards a consensual solution. Nonetheless, in spite of its seeming triviality, any short description of the Delphi method would be an oversimplification, and any lengthy one would go beyond the scope of this exposition. We shall therefore use some mathematical notation and focus on another way to make sense of the AHP in group decisions.

Suppose that m ($m \geq 2$) decision makers are involved in a decision and we want to take into account and synthesize their opinions, i.e. we want to *aggregate* them. This suggests that we ought to average different opinions. More specifically, starting from their pairwise comparison matrices

$$\underbrace{\left(a_{ij}^{(1)}\right)_{n\times n}}_{\mathbf{A}_1}, \ldots, \underbrace{\left(a_{ij}^{(m)}\right)_{n\times n}}_{\mathbf{A}_m}$$

we eventually want to obtain *one* representative group priority vector $\mathbf{w}^G = (w_1^G, \ldots, w_n^G)^T$. According to Forman and Peniwati [57] there are two methods to derive a vector $\mathbf{w}^G$ from a set of pairwise comparison matrices $\mathbf{A}_1, \ldots, \mathbf{A}_m$ and they differ in where the aggregation takes place.

- *Aggregation of individual judgments* (AIJ): Matrices $\mathbf{A}_1, \ldots, \mathbf{A}_m$ can be aggregated into a single pairwise comparison matrix, $\mathbf{A}^G = (a_{ij}^G)$, and then the priority vector be calculated from $\mathbf{A}^G$ with any of the methods described in §2.1. In this case the aggregation happens before the elicitation of the priorities

- *Aggregation of individual priorities* (AIP): Priority vectors $\mathbf{w}_1, \ldots, \mathbf{w}_m$ can be derived from the original set of matrices. These vectors are then aggregated into $\mathbf{w}^G$. In this case, the aggregation happens after the derivation of the priority vectors.

The following scheme shall clarify the difference between AIJ and AIP. Also, it should be evident that, either way, going from a set of pairwise comparison matrices to a single

priority vector entails a double process of aggregation.

$$\begin{array}{ccccc} \mathbf{A}_1, \ldots, \mathbf{A}_m & \xrightarrow{\S 2.1} & \mathbf{w}_1, \ldots, \mathbf{w}_m & \xrightarrow{\text{AIP}} & \mathbf{w}^G \\ \Big\downarrow \text{AIJ} & & & & \\ \mathbf{A}^G & & & & \\ \Big\downarrow \S 2.1 & & & & \\ \mathbf{w}^G & & & & \end{array} \tag{3.7}$$

Obviously, the crucial point is that of finding a suitable aggregation function. For the aggregation of individual *judgments* (AIJ), the reader can check that a basic function like the arithmetic mean fails since the resulting matrix $\mathbf{A}^G$ would not be reciprocal. Aczel and Saaty [1] and Saaty and Alsina [110] proposed a set of reasonable properties for the aggregation of preferences and, by using functional analysis, proved that in this context the only meaningful and non-trivial aggregation method is the weighted geometric mean. Namely, entries of the group matrix $\mathbf{A}^G = (a_{ij}^G)_{n\times n}$ are obtained using the following parametric formula,

$$a_{ij}^G = \prod_{h=1}^{m} a_{ij}^{(h)^{\lambda_h}}$$

with $\lambda_h > 0\ \forall h$ and $\lambda_1 + \cdots + \lambda_m = 1$. The most common interpretation of a given λ_h is that it should be proportional to the importance of the hth decision maker. When $\lambda_h = 1/m\ \forall h$ then all the decision makers have the same importance. Conversely, if $\lambda_h > \lambda_k$, then the relative importance of the hth decision maker is greater than that of the kth.

Example 15. *Consider the very simple case of two decision makers with preferences expressed as*

$$\mathbf{A}_1 = \begin{pmatrix} 1 & 2 & 1/2 \\ 1/2 & 1 & 3 \\ 2 & 1/3 & 1 \end{pmatrix} \qquad \mathbf{A}_2 = \begin{pmatrix} 1 & 7 & 2 \\ 1/7 & 1 & 1/4 \\ 1/2 & 4 & 1 \end{pmatrix} \tag{3.8}$$

and suppose that the first decision maker should be twice as influential as the second. This suggests the use of $\lambda_1 = 2/3$ and $\lambda_2 = 1/3$. Hence, the group matrix computed with AIJ is

$$\mathbf{A}^G = \begin{pmatrix} 1 & 2^{\frac{2}{3}}7^{\frac{1}{3}} & (1/2)^{\frac{2}{3}}2^{\frac{1}{3}} \\ (1/2)^{\frac{2}{3}}1/7^{\frac{1}{3}} & 1 & 3^{\frac{2}{3}}(1/4)^{\frac{1}{3}} \\ 2^{\frac{2}{3}}(1/2)^{\frac{1}{3}} & (1/3)^{\frac{2}{3}}4^{\frac{1}{3}} & 1 \end{pmatrix} \approx \begin{pmatrix} 1 & 3.04 & 0.79 \\ 0.22 & 1 & 1.31 \\ 1.26 & 0.76 & 1 \end{pmatrix},$$

from which a group priority vector be easily derived.

If we turn our attention to the aggregation of *priorities* (AIP), two formulas are accepted, the weighted geometric mean and the weighted arithmetic mean,

$$w_i^G = \left(\prod_{h=1}^{m} w_i^{(h)^{\lambda_i}} \right), \qquad w_i^G = \left(\sum_{h=1}^{m} \lambda_i w_i^{(h)} \right).$$

These two formulas clearly lead to different priority vectors, but they are both accepted in the literature, perhaps with a slight preference for the geometric mean [17].

Example 16. *Consider the two matrices in (3.8) and their priority vectors*

$$\mathbf{w}_1 \approx (0.331313, 0.379259, 0.289428)^T$$
$$\mathbf{w}_2 \approx (0.602629, 0.082342, 0.315029)^T$$

obtained with the eigenvector method (but we could have used any other method). At this point the geometric mean aggregation can be applied component-wise to $\mathbf{w}_1$ *and* $\mathbf{w}_2$ *to estimate the group priority vector*

$$\mathbf{w}^G = \begin{pmatrix} 0.331313^{\frac{2}{3}} 0.602629^{\frac{1}{3}} \\ 0.379259^{\frac{2}{3}} 0.082342^{\frac{1}{3}} \\ 0.289428^{\frac{2}{3}} 0.315029^{\frac{1}{3}} \end{pmatrix} \approx \begin{pmatrix} 0.404429 \\ 0.227945 \\ 0.297722 \end{pmatrix}.$$

It is noteworthy that, when the geometric mean method is used to derive the priorities and the geometric mean is used to aggregate judgments, the diagram (3.7) becomes commutative, as depicted in (3.9), and thus using AIP or AIJ makes no difference.

$$\begin{array}{ccc} \mathbf{A}_1, \ldots, \mathbf{A}_m & \xrightarrow{\S 2.1} & \mathbf{w}_1, \ldots, \mathbf{w}_m \\ {\scriptstyle \text{AIJ}}\downarrow & & \downarrow{\scriptstyle \text{AIP}} \\ \mathbf{A}^G & \xrightarrow{\S 2.1} & \mathbf{w}^G \end{array} \tag{3.9}$$

An interesting question could then refer to how much difference there is between weight vectors $\mathbf{w}^G$ when the geometric mean method is not used, and perhaps the eigenvector method is employed. According to a recent study [17], in these cases, the differences between vectors are often negligible.

Compatibility index

Very often, it is desirable that a sufficient level of consensus is reached, before opinions of different experts are aggregated. Namely, in many procedures different decision makers are encouraged to discuss, clarify issues and make their opinions converge towards a consensual solution. Therefore, it is important to define an index of similarity between opinions of decision makers. One of these indices was defined by Saaty [109] and goes under the name of *compatibility index*. Recall that, given two matrices of the same order $\mathbf{A} = (a_{ij})_{n\times n}$ and $\mathbf{B} = (b_{ij})_{n\times n}$, their Hadamard product $\mathbf{A} \circ \mathbf{B}$ is defined as the entry-wise multiplication, i.e. $\mathbf{A} \circ \mathbf{B} = (a_{ij} \cdot b_{ij})_{n\times n}$. At this point the compatibility index of two pairwise comparison matrices of order n was defined in a matrix form as

$$\text{comp}(\mathbf{A}, \mathbf{B}) = \frac{1}{n^2} \mathbf{1}^T \left(\mathbf{A} \circ \mathbf{B}^T \right) \mathbf{1},$$

where $\mathbf{1} = (1, \ldots, 1)^T$. Note that it can be rewritten as

$$\text{comp}(\mathbf{A}, \mathbf{B}) = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} (a_{ij} \cdot b_{ji}).$$

Example 17. *For illustrative simplicity, let us consider two matrices differing only for one entry and its reciprocal,*

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 1/2 & 1 & 5 \\ 1/3 & 1/5 & 1 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} 1 & 2 & 3 \\ 1/2 & 1 & 2 \\ 1/3 & 1/2 & 1 \end{pmatrix}.$$

By using the definition of compatibility index we obtain

$$comp(\mathbf{A}, \mathbf{B}) = \frac{1}{n^2}(1,1,1) \begin{pmatrix} 1 & 2 & 3 \\ 1/2 & 1 & 5 \\ 1/3 & 1/5 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 1/2 & 1/3 \\ 2 & 1 & 1/2 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Proceeding by solving the Hadamard product,

$$comp(\mathbf{A}, \mathbf{B}) = \frac{1}{n^2}(1,1,1) \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 5/2 \\ 1 & 2/5 & 1 \end{pmatrix}}_{\mathbf{A} \circ \mathbf{B}^T} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

At this point the remaining simplifies to taking the arithmetic mean of all the entries of $\mathbf{A} \circ \mathbf{B}^T$*, which returns* $comp(\mathbf{A}, \mathbf{B}) = 99/90 = 1.1$.

The example shows that the minimum value attained by the compatibility index is 1 and it represents perfect consensus. Additionally, it was proven by Saaty (see Theorem 1 in [109]) that there is a connection between this metric and the method of the eigenvector. Considering $\mathbf{W}$ the matrix constructed by using the priority vector of $\mathbf{A}$ obtained by using the eigenvector method, then $\text{comp}(\mathbf{A}, \mathbf{W}) = \lambda_{\max}/n$. In this sense, this quantification of consensus results appealing to those who prefer the eigenvector method. It goes without saying that many other metrics, e.g. matrix norms of $(\mathbf{A} - \mathbf{B})$, can be used to estimate the distances between preferences of experts.

3.2.1. Integrated methods ⋆

More models have been built to deal with many relaxations at once. These mathematical models can be called integrated, in the sense that they incorporate different purposes in the same model. Many times, integrated models can be formulated in very simple forms[1] and here we can even make up one of them for the purpose of the exposition. In this case we are interested to derive the priority vector from a set of *incomplete* pairwise comparison matrices $\dot{\mathbf{A}}_1, \ldots, \dot{\mathbf{A}}_m$. Hence, *two* problems, the incompleteness of preferences and the multiplicity of decision makers, can be accommodated in *one* single

[1] However, it is the opinion of the author of this manuscript that recently, in several papers, an apparent effort has been made to complicate things which could have been left simple.

optimization problem as, for example, the following

$$\begin{aligned} \underset{(w_1^G,\ldots,w_n^G)}{\text{minimize}} \quad & \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{h=1}^{m}\delta_{ij}^{(h)}\left(\log \dot{a}_{ij}^{(h)} + \log w_j^G - \log w_i^G\right)^2 \\ \text{subject to} \quad & \sum_{i=1}^{n} w_i^G = 1, \end{aligned} \tag{3.10}$$

with

$$\delta_{ij}^{(h)} = \begin{cases} 1, & \text{if } \dot{a}_{ij}^{(h)} \text{ is given} \\ 0, & \text{if } \dot{a}_{ij}^{(h)} \text{ is missing} \end{cases}$$

Note that the constrained optimization problem (3.10) aims at finding the closest (using a logarithmic metric) possible matrix $(w_i^G/w_j^G)_{n\times n}$ to the preferences expressed by the decision makers and the variables $\delta_{ij}^{(h)} \in \{0,1\}$ make all the terms containing a missing comparison vanish.

Example 18. *Suppose that three decision makers, which could be the three members of the family of the initial example, express their preferences on three alternatives in the form of the following incomplete pairwise comparison matrices,*

$$\dot{\mathbf{A}}_1 = \begin{pmatrix} 1 & 2 & \dot{a}_{13}^{(1)} \\ 1/2 & 1 & 3 \\ \dot{a}_{31}^{(1)} & 1/3 & 1 \end{pmatrix} \quad \dot{\mathbf{A}}_2 = \begin{pmatrix} 1 & \dot{a}_{12}^{(2)} & 1/2 \\ \dot{a}_{21}^{(2)} & 1 & 3 \\ 2 & 1/3 & 1 \end{pmatrix} \quad \dot{\mathbf{A}}_3 = \begin{pmatrix} 1 & \dot{a}_{12}^{(3)} & \dot{a}_{13}^{(3)} \\ \dot{a}_{21}^{(3)} & 1 & 5 \\ \dot{a}_{31}^{(3)} & 1/5 & 1 \end{pmatrix}.$$

Then, solving the optimization problem (3.10), one obtains

$$\mathbf{w}^G = (0.312391, 0487379, 0.20023)^T.$$

Surely the reader can imagine more integrated models and the next section, on extensions of the AHP, will hopefully provide more food for thought also under this lens.

4. Extensions

In this chapter we shall proceed and analyze further extensions for pairwise comparison matrices. The common denominator of the following extensions is that they all involve the domain of representation of the pairwise comparisons a_{ij}, that is the set of possible values attained by a_{ij}.

4.1. Equivalent representations

So far we have expressed pairwise comparisons using the so-called *multiplicative* scale, i.e. the judgments have been expressed by means of positive real numbers, $a_{ij} > 0 \; \forall i, j$. The multiplicative scale is often taken from granted, but here we shall keep our minds open and observe that this should not be the case. Let us follow the tradition and highlight the assumption which will be relaxed in this section.

> A *single* decision maker is *perfectly rational* and can *precisely* express his preferences on *all pairs* of *independent* alternatives and criteria using ***positive real numbers***.

Alternative numerical representations have been proposed to model pairwise comparisons. The most popular and studied are the additive representation and the one based on reciprocal relations. In this section we shall discuss these two, see that concepts as reciprocity and consistency can be similarly replicated in these other two frameworks, and finally suggest that there is a deeper connection among these representations which can be formalized by using abstract algebra.

4.1.1. Additive pairwise comparison matrices

The so-called *additive* representation of preferences by means of additive pairwise comparison matrices was well-presented by Barzilai [9] and has been used in methods alternative, yet very similar, to the AHP such as the Simple Multi-Attribute Rating Technique (SMART) [86] and the Ratio Estimations in Magnitudes or deci-Bells to Rate Alternatives which are Non-Dominated Technique (REMBRANDT) [7, 93]. The domain of representation of preferences is the real line, indifference is represented by 0 and, if we call $\mathbf{P} = (p_{ij})_{n \times n}$ the *additive pairwise comparison matrix* containing the preferences in this form, then the condition of reciprocity becomes $p_{ij} + p_{ji} = 0 \; \forall i, j$, whence the name 'additive'. The condition of consistency becomes

$$p_{ik} = p_{ij} + p_{jk} \; \forall i, j, k. \tag{4.1}$$

If and only if a matrix is consistent, then there exists a priority vector $\mathbf{u} = (u_1, \ldots, u_n)^T$ such that $p_{ij} = u_i - u_j \ \forall i, j$. One natural question regards the relation between pairwise comparison matrices and their additive representations. Namely, is there a way to associate a pairwise comparison matrix to its additive version and vice versa? A minimum requirement is that this transformation could map consistent pairwise comparison matrices into their consistent counterparts. The answer is positive and any logarithmic function would make it. For instance, using the natural logarithm, given a pairwise comparison matrix $\mathbf{A}$, we can obtain its additive representation $\mathbf{P} = (p_{ij})_{n \times n}$ with $p_{ij} = \ln a_{ij}$. Conversely, to go back to the multiplicative representation one can use its inverse, the exponential transformation $a_{ij} = e^{p_{ij}}$.

Example 19. *Consider the consistent pairwise comparison matrix*

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 8 \\ 1/2 & 1 & 4 \\ 1/8 & 1/4 & 1 \end{pmatrix}. \tag{4.2}$$

Using the logarithm in base 2 *one obtains the following skew-symmetric matrix*

$$\mathbf{P} = \begin{pmatrix} \log_2 1 & \log_2 2 & \log_2 8 \\ \log_2 1/2 & \log_2 1 & \log_2 4 \\ \log_2 1/8 & \log_2 1/4 & \log_2 1 \end{pmatrix} \approx \begin{pmatrix} 0 & 1 & 3 \\ -1 & 0 & 2 \\ -3 & -2 & 0 \end{pmatrix},$$

for which the additive consistency condition (4.1) holds, e.g. $1 + 2 = 3$. *Moreover, one can check that the priority vector associated with* $\mathbf{P}$ *is*

$$\mathbf{u} = (2, 1, -1)^T.$$

Problem 8. *Can you find a way to derive the vector* $\mathbf{u}$ *from a consistent additive pairwise comparison matrix* $\mathbf{P}$*?*

One convenient fact about this representation is that, fixed a value for n, the set of all additive pairwise comparison matrices $\mathbf{P}$ of order n is a subspace of the linear space $\mathbb{R}^{n \times n}$. Note that also the set of *consistent* additive pairwise comparison matrices is a subspace of $\mathbb{R}^{n \times n}$ [81]. Thus, the possible loss of results that we get from giving up working with positive matrices is here compensated by the gain of the suite of tools from linear algebra. To explain it with an example, remember that the inconsistency index CI for pairwise comparison matrices was based on some results on positive square matrices (the fact that $\lambda_{\max}$ of $\mathbf{A}$ is always a real number) and therefore they are not directly replicable for additive pairwise comparison matrices. On the other hand, additive pairwise comparison matrices $\mathbf{P}$ form linear spaces, which allows us to draw from linear algebra to obtain original results. The interested reader might want to see, for instance, the inconsistency index defined by Barzilai [9] as it relies on considerations stemming from linear algebra. The priority vector $\mathbf{u}$ has a different interpretation than $\mathbf{w}$. In $\mathbf{u}$ the information is captured by the *differences* $(u_i - u_j)$ between priorities and not their ratios. Unlike for the components of $\mathbf{w}$, the ratio between u_i and u_j has no meaning. Consider that some

components u_i can be negative too.
One last remark regards the apparent similarity of this approach to the one with pairwise comparison matrices. In a consistent pairwise comparison matrix each column is equal to any other column multiplied times a suitable scalar. In the additive approach each column is equal to any other plus a suitable scalar. The same reasoning affects also the priority vectors. Priority vectors $\mathbf{w}$ of consistent pairwise comparison matrices are unique up to multiplication, whereas vectors $\mathbf{u}$ are unique, but up to addition.

4.1.2. Reciprocal relations

Another prominent representation of preferences is based on reciprocal relations [44], often called *fuzzy preference relations* [69] in the fuzzy sets literature. The notion of reciprocal relation became popular in the framework of fuzzy sets, but it can be verified that its inception dates back, at least, to the study by Luce and Suppes [88] on probabilistic preference relations.

A *reciprocal relation* can be represented by a matrix $\mathbf{R} = (r_{ij})_{n\times n}$ with $r_{ij} \in]0,1[$ satisfying the reciprocity condition $r_{ij} + r_{ji} = 1$ and with the indifference represented by the value 0.5. The *consistency* condition for reciprocal relations is

$$\frac{r_{ik}}{r_{ki}} = \frac{r_{ij}}{r_{ji}}\frac{r_{jk}}{r_{kj}} \quad \forall i,j,k. \tag{4.3}$$

Most of the references to this condition refer to Tanino [129] but the very same condition was already used by Luce and Suppes [88] and Shimura [123]. Furthermore, to make it more homogeneous with respect to the conditions of consistency for pairwise comparison matrices and additive pairwise comparison matrices, Chiclana *et al.* [38] showed that (4.3) can be equivalently written as

$$r_{ik} = \frac{r_{ij}r_{jk}}{r_{ij}r_{jk} + (1-r_{ij})(1-r_{jk})} \quad \forall i,j,k. \tag{4.4}$$

If and only if this consistency condition is satisfied, then there exists a weight vector $\mathbf{w}$ such that $r_{ij} = w_i/(w_i + w_j)$. The problem of finding the weight vector arises also for reciprocal relations and even in this environment many methods have been proposed. Among them, the most straightforward is probably the following,

$$w_i = \left(\prod_{j=1}^{n} \frac{r_{ij}}{1-r_{ij}}\right)^{\frac{1}{n}},$$

which was proven [51] to be the counterpart of the geometric mean method for pairwise comparison matrices. Similarly to what was established for additive pairwise comparison matrices, pairwise comparison matrices can be transformed into reciprocal relations by means of the following function

$$r_{ij} = \frac{a_{ij}}{1+a_{ij}}, \tag{4.5}$$

and its inverse $a_{ij} = r_{ij}/r_{ji}$ can be used to transform reciprocal relations to pairwise comparison matrices.

Example 20. *Consider the consistent pairwise comparison matrix* $\mathbf{A}$ *in (4.2). Using the transformation (4.5) one obtains*

$$\mathbf{R} \approx \begin{pmatrix} 1/2 & 2/3 & 8/9 \\ 1/3 & 1/2 & 4/5 \\ 1/9 & 1/5 & 1/2 \end{pmatrix},$$

for which the consistency condition (4.4) holds (check!). Moreover, one can check that the priority vector associated with $\mathbf{R}$ *is* $\mathbf{w} = (8/13, 4/13, 1/13)^T$ *and corresponds to the vector that would have been obtained from* $\mathbf{A}$*.*

The reader should be aware that another type of consistency condition, called *additive* consistency [129], for reciprocal relations was proposed and later developed, but we shall not dwell on it in this booklet. The reader can refer to [51] for an overview of transformations between pairwise comparison matrices and reciprocal relations and a method to derive the priority vector from these latter, and to Xu [135] for a survey which elaborates on different representations of pairwise preferences. Table 4.1 draws a parallel and summarizes the different representations of preferences and their main characteristics.

	Multiplicative	Additive	Reciprocal
Domain of representation	$\mathbb{R}_>$	$\mathbb{R}$	$]0,1[$
Reciprocity condition	$a_{ij} = 1/a_{ji}$	$p_{ij} = -p_{ji}$	$r_{ij} = 1 - r_{ji}$
Value for indifference between alternatives	1	0	0.5
Consistency condition	$a_{ik} = a_{ij}a_{jk}$	$p_{ik} = p_{ij} + p_{jk}$	$r_{ik} = \frac{r_{ij}r_{jk}}{r_{ij}r_{jk}+(1-r_{ij})(1-r_{jk})}$
Weight vector characterization	$a_{ij} = \frac{w_i}{w_j}$	$p_{ij} = u_i - u_j$	$r_{ij} = \frac{w_i}{w_i+w_j}$

Table 4.1.: Representations of pairwise preferences and their properties.

The transformations between different representations are instead depicted in Figure 4.1.

These three representations of preferences have different origins, but, if we look backwards, their similarities were already visible years ago. Consider, for example, that the problem of inconsistency and intransitivity, which can occur in all three representations, was in fact considered (and treated similarly) in each of them. As Gass [61] noted, Fishburn [56], whose skew symmetric representation of preferences is the progenitor of additive pairwise comparison matrices, wrote:

> Transitivity is obviously a great practical convenience and a nice thing to have for mathematical purposes, but long ago this author ceased to understand why it should be a cornerstone of normative decision theory.

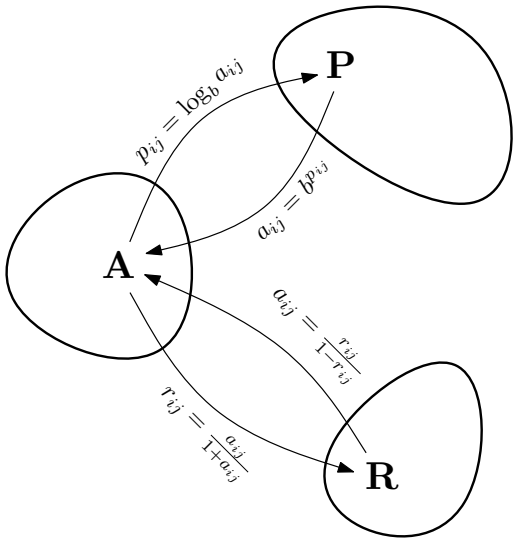


Figure 4.1.: Transformations between different representations of valued preferences.

Even Luce and Raiffa [87] whose work can be seen as an inception of reciprocal relations, wrote:

> No matter how intransitivities arise, we must recognize that they exist, and we can take a little comfort in the thought that they are an anathema to most of what constitutes theory in the behavioral sciences today.

The same view was also shared by Saaty who, already in his seminal paper on the AHP [101], wrote:

> As a realistic representation of the situation in preference comparisons, we wish to account for inconsistency in judgments because, despite their best efforts, people's feelings and preferences remain inconsistent and intransitive.

4.1.3. Group isomorphisms between equivalent representations ⋆

It is apparent that these three representations of preferences are very similar and we can shift from one approach to another, but to what extent are they interchangeable? The non-trivial, yet elegant answer, is to the extent to which the domains of representations of preferences, together with their conditions of consistency are isomorphic groups. Recall that, in group theory, a *group* is a set S equipped with a binary operator $* : S \times S \to S$ such that

- the set S is closed under the operator $*$, i.e. $a * b \in S \;\; \forall a, b \in S$
- the operator $*$ is associative, i.e. $a * (b * c) = (a * b) * c \;\; \forall a, b, c \in S$
- there exists an identity element e such that $s * e = s \;\; \forall s \in S$
- for each $s \in S$ there exists an inverse element $s^{-1} \in S$ such that $s * s^{-1} = e$.

A group is represented by a pair $(S, *)$ where the first component is the set and the second is the operator. Two groups $(S, *)$ and $(Q, \odot)$ are *group isomorphic* if and only if there exists a bijection (group isomorphism) $f : S \to Q$ such that, for all $x, y \in S$, it is

$$f(x) * f(y) = f(x \odot y)\,.$$

Now, if we look at Table 4.1 we shall check that each domain of representation together with its consistency operation is a group where the identity element e is the value expressing indifference between alternatives and where the inverse element is determined by means of the reciprocity condition. Moreover, it can be checked that they are isomorphic groups, the isomorphisms being the functions in Figure 4.1.

Example 21. *The logarithm relates* $(\mathbb{R}_>, \cdot)$ *with* $(\mathbb{R}, +)$ *and is perhaps the most famous group isomorphism. In fact, from basics of calculus we know the rule*

$$\log(x) + \log(y) = \log(x \cdot y) \;\; \forall x, y > 0\,,$$

which exposes the relation between pairwise comparison matrices and additive pairwise comparison matrices.

The reader familiar with group theory must have understood the strength and the implication of group isomorphism which, in words, was described by Fraleigh in his textbook [58] as "the concept of two systems being structurally identical, that is, one being just like the other except for names". The existence of group isomorphisms between different representations of preference is not a mere theoretical exercise but a precious result as it helps to naturally extend concepts from one representation to another one. For a deep and theoretical analysis of the group isomorphisms between these representations of preferences the reader might find the papers by Cavallo and D'Apuzzo [35, 36] enlightening.

4.2. Interval AHP

In §2.2.5 the reader was already presented with a pairwise comparison matrix whose entries were intervals instead of real numbers. In that case the interval-valued matrix was functional in the definition of an inconsistency index, but it is natural to imagine that a decision maker could express his judgments by means of intervals. This is natural to cope with uncertainty and imprecision. In this and in the next section we shall dwell on representations of preferences when the decision maker cannot state them precisely and with absolute certainty and see what the literature has to offer.

A *single* decision maker is *perfectly rational* and can ***precisely*** express his preferences on *all pairs* of *independent* alternatives and criteria using *positive real numbers*.

This section shall introduce the principles behind what probably is the most widely known extension of the AHP with intervals. Salo and Hämäläinen considered interval judgments $\bar{a}_{ij} = [a_{ij}^L, a_{ij}^R]$ as bounds for the values of the 'true' weights, i.e. the interval-valued comparison $\bar{a}_{ij} = [a_{ij}^L, a_{ij}^R]$, entails that $a_{ij}^L \leq w_i/w_j \leq a_{ij}^R$. At this point, it is important to know what values different weights can attain, given the constraints imposed by the interval pairwise comparisons. What is, for instance, the maximum possible value of w_i given an interval-valued matrix $\bar{\mathbf{A}}$? To solve this problem, we first need to define the set of all normalized priority vectors with n components as

$$W_n = \left\{ (w_1, \ldots, w_n)^T \middle| \sum_{i=1}^{n} w_i = 1,\ w_i > 0\ \forall i \right\},$$

Such a set is depicted in Figure 4.2 for the case with $n = 3$. Furthermore, the set of

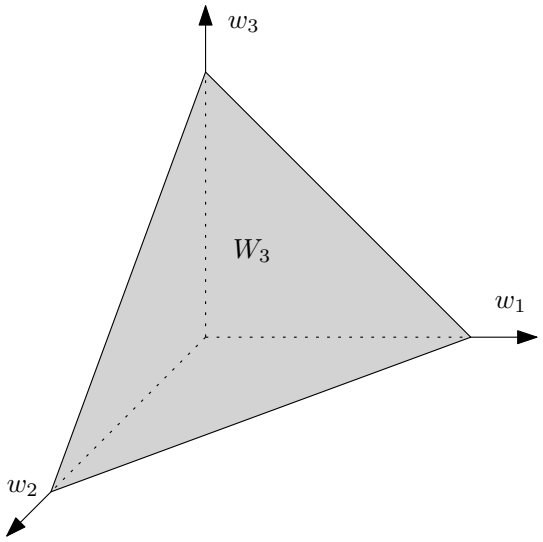


Figure 4.2.: Graphical representation of W_3

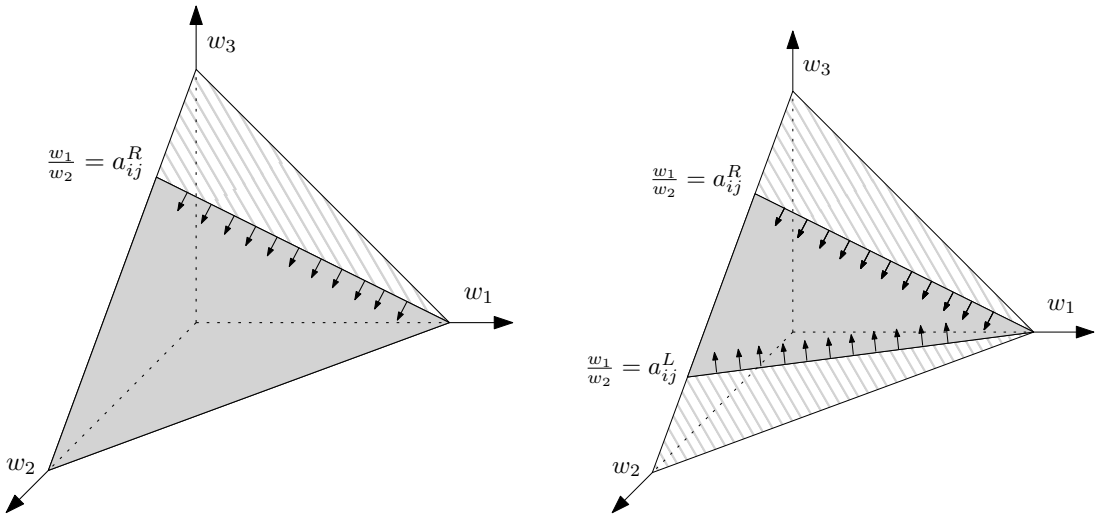
feasible weight vectors according to the interval-valued pairwise comparison matrix $\bar{\mathbf{A}}$ is

$$S_{\bar{\mathbf{A}}} = \left\{ (w_1, \ldots, w_n)^T \middle| a_{ij}^L \leq \frac{w_i}{w_j} \leq a_{ij}^R\ \forall i < j \right\}$$

As showed in Figure 4.3, adding the constraints characterizing $S_{\bar{\mathbf{A}}}$ to W obviously reduces the set of feasible solutions. It follows that the 'true' normalized weight vector must be an element of the set $W_n \cap S_{\bar{\mathbf{A}}}$, as pictured in Figure 4.4. Then it is possible to construct an interval-valued vector $\bar{\mathbf{w}} = (\bar{w}_1, \ldots, \bar{w}_n)^T$ with $\bar{w}_i = [w_i^L, w_i^R]$ where w_i^L and w_i^R are the smallest and the greatest possible values for w_i respectively. Hence, they can be computed as follows,

$$w_i^L = \underset{\mathbf{w} \in W_n \cap S_{\bar{\mathbf{A}}}}{\text{minimize}}\ w_i \qquad i = 1, \ldots, n, \tag{4.6}$$

$$w_i^R = \underset{\mathbf{w} \in W_n \cap S_{\bar{\mathbf{A}}}}{\text{maximize}}\ w_i \qquad i = 1, \ldots, n. \tag{4.7}$$



(a) Application of the constraint $w_1/w_2 \leq a_{ij}^R$ (b) Adding the constraint $w_1/w_2 \geq a_{ij}^L$

Figure 4.3.: The sequential application of the constraints reduces the region of feasible solutions

Example 22. *Considering the matrix*

$$\bar{\mathbf{A}} = \begin{pmatrix} 1 & [7,20] & [7/5,4] \\ [1/20,1/7] & 1 & [1/5,4/7] \\ [1/4,5/7] & [7/4,5] & 1 \end{pmatrix} \tag{4.8}$$

already used in (2.16). Then the weight w_1^R is the optimal value of the following optimization problem.

$$\begin{aligned} &\underset{(w_1,w_2,w_3)}{\textit{maximize}} && w_1 \\ &\textit{subject to} && \left.\begin{aligned} &7 \leq w_1/w_2 \leq 20, \\ &7/5 \leq w_1/w_3 \leq 4, \\ &1/5 \leq w_2/w_3 \leq 4/7, \end{aligned}\right\} \Rightarrow (w_1,w_2,w_3)^T \in S_{\bar{\mathbf{A}}} \\ &&& \left.\begin{aligned} &w_1 + w_2 + w_3 = 1, \\ &w_1, w_2, w_3 > 0 \end{aligned}\right\} \Rightarrow (w_1,w_2,w_3)^T \in W_3 \end{aligned} \tag{4.9}$$

By proceeding in this way, we ask what the greatest possible value achievable by w_1 is, when $(w_1,w_2,w_3) \in W_n \cap S_{\bar{\mathbf{A}}}$. The interval-valued priority vector derivable from $\bar{\mathbf{A}}$ in (2.16) is

$$\bar{\mathbf{w}} = \begin{pmatrix} [0.54,0.77] \\ [0.04,0.10] \\ [0.18,0.38] \end{pmatrix} \tag{4.10}$$

In the example, the vector (4.10) provided enough information and we knew that the best alternative was x_1 since its weight cannot be smaller than the weights of the

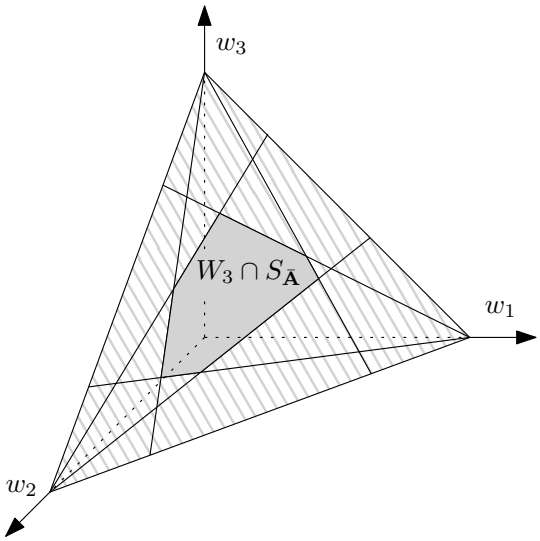


Figure 4.4.: The set $W_3 \cap S_{\bar{\mathbf{A}}}$.

other alternatives. However, in other cases, when intervals overlap, selection of the best alternative is non-trivial. To solve this problem one can use different strategies. Firstly, the decision maker can be asked to refine his judgments until the best alternative is clearly identified. Secondly, when this is not a viable solution, some methods for ranking intervals can be employed. Among such methods, there are the pairwise dominance [121] and the methods for ranking fuzzy quantities [133].
In this section we described a method for deriving weights which can be used on a single interval-valued pairwise comparison matrix and not on a whole hierarchy. The extension to the whole hierarchy is methodologically straightforward but quite lengthy to be explained, and therefore the reader can refer to the original contribution [121].

Euclidean center of $W_n \cap S_{\bar{\mathbf{A}}}$

The problem of ranking interval weights and their hierarchical composition can be solved by means of a shortcut, which is used to derive real valued weights from interval valued comparison matrices. The following is due to Arbel and Vargas [6]. Their solution is based on the fact that the set of constraints characterizing $W_n \cap S_{\bar{\mathbf{A}}}$ can be equivalently stated as a set of linear constraints since those containing ratios can be splitted into two linear constraints. Considering for instance the first constraint in (4.9), one can see that

$$\begin{aligned} 7 \leq w_1/w_2 \leq 20 &\Leftrightarrow 7w_2 \leq w_1 \text{ and } w_2 \leq 20w_1 \\ &\Leftrightarrow 7w_2 - w_1 \leq 0 \text{ and } w_2 - 20w_1 \leq 0. \end{aligned}$$

Hence, since $W_n \cap S_{\bar{\mathbf{A}}}$ is a bounded set defined by linear constraints it is a polytope. Arbel and Vargas proposed to take the real valued weights $w_1, \ldots, w_n$ as the coordinates of the Euclidean center of the polyhedron $W_n \cap S_{\bar{\mathbf{A}}}$. In words, the Euclidean center of a polytope is the center of the maximum radius ball which can be inscribed in the

polytope and it can be found by solving a linear optimization problem [1]. Figure 4.5 reports a graphical example of the Euclidean center of a 2-dimensional polytope, i.e. a polygon.

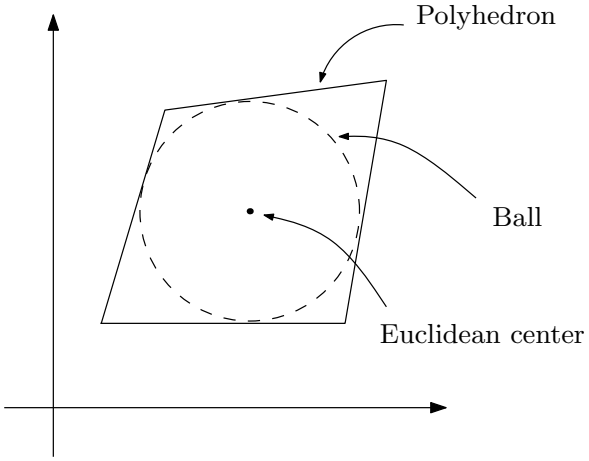


Figure 4.5.: Graphical example of the Euclidean center of a polygon.

Let us see how to write down a linear optimization problem to find the ball with the largest radius r in $W_n \cap S_{\bar{\mathbf{A}}}$. Since $W_n \cap S_{\bar{\mathbf{A}}}$ is a polytope, it can be defined by a set of inequalities $\mathbf{a}_i^T \mathbf{w} \leq b_i$. The problem is how to model the constraints with respect to the center of the ball. Consider a single constraint. A ball of radius r pointed in $\mathbf{w}$, $B(r, \mathbf{w})$, satisfies the ith inequality, if $\mathbf{a}_i^T \mathbf{y} \leq b_i \ \forall \mathbf{y} \in B(r, \mathbf{w})$. The trick is to write the inequality in such a way that we consider a point $\mathbf{y}^*$ which is the point in $B(r, \mathbf{w})$ with the greatest value when multiplied by $\mathbf{a}_i^T$, i.e. $\mathbf{a}_i^T \mathbf{y}^* \geq \mathbf{a}_i^T \mathbf{y} \ \forall \mathbf{y} \in B(r, \mathbf{w})$. This point is the point $\mathbf{y}^* = \mathbf{w} + r\frac{\mathbf{a}_i}{\|\mathbf{a}_i\|}$ (convince yourself graphically in 2-dimensions) and therefore the inequality can be written as

$$\mathbf{a}_i^T \underbrace{\left(\mathbf{w} + r\frac{\mathbf{a}_i}{\|\mathbf{a}_i\|}\right)}_{\mathbf{y}^*} \leq b_i \,.$$

Hence, the optimization problem becomes

$$\begin{aligned} &\underset{r,\mathbf{w}}{\text{maximize}} \quad r \\ &\text{subject to} \quad \mathbf{a}_i^T \left(\mathbf{w} + r\frac{\mathbf{a}_i}{\|\mathbf{a}_i\|}\right) \leq b_i \qquad i = 1 \ldots, N \end{aligned} \tag{4.11}$$

where N is the number of inequalities used to define P. The optimization problem (4.11) can be seen as max-min optimization problem too. In fact, the variable r, which is maximized, eventually is the distance between the center of the ball $\mathbf{w}$ and the closest (least distant) face of the polyhedron $W \cap S_{\bar{\mathbf{A}}}$.

[1] Note that in some other sources it is referred to as the Chebychev center. See, for instance, the book by Boyd and Vandenberghe [21].

Example 23. *For sake of homogeneity we still consider the matrix* $\bar{\mathbf{A}}$ *from (2.16). Then we have*

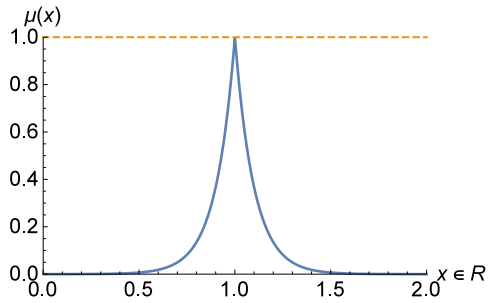
$$
\begin{array}{ll}
\underset{r,\mathbf{w}}{\textit{maximize}} & r \\
\textit{subject to} & \left.\begin{array}{l}
(-1,7,0)\left(\mathbf{w}+r\dfrac{(-1,7,0)^T}{5\sqrt{2}}\right)\leq 0, \\
(1,20,0)\left(\mathbf{w}+r\dfrac{(1,20,0)^T}{\sqrt{401}}\right)\leq 0, \\
(-1,0,7/5)\left(\mathbf{w}+r\dfrac{(-1,0,7/5)^T}{\sqrt{74}/5}\right)\leq 0, \\
(1,0,-4)\left(\mathbf{w}+r\dfrac{(1,0,-4)^T}{\sqrt{17}}\right)\leq 0, \\
(0,-1,1/5)\left(\mathbf{w}+r\dfrac{(0,-1,1/5)^T}{\sqrt{26}/5}\right)\leq 0, \\
(0,1,-4/7)\left(\mathbf{w}+r\dfrac{(0,1,-4/7)^T}{\sqrt{65}/7}\right)\leq 0,
\end{array}\right\} \Rightarrow (w_1,w_2,w_3)^T\in S_{\bar{\mathbf{A}}} \\
& \left.\begin{array}{l}
w_1+w_2+w_3=1, \\
w_1,w_2,w_3>0
\end{array}\right\} \Rightarrow (w_1,w_2,w_3)^T\in W_3
\end{array}
\tag{4.12}
$$

Note that the constraints defining W_3 *are left unchanged: the constraint* $w_1+w_2+w_3=1$ *is an equality and therefore it must hold exactly, and the positivity constraints* $w_1,w_2,w_3>0$ *could even be deleted since they are made redundant by those defining* $S_{\bar{\mathbf{A}}}$*. The vector maximizing* r *in the optimization problem is* $\mathbf{w}\approx(0.72,0.07,0.21)^T$*.*

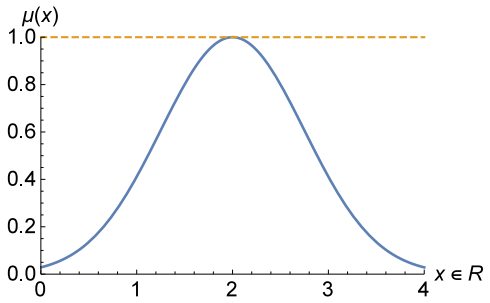
This approach to interval judgments, which considers intervals as implicitly defining bounds for weights, was initially proposed by Arbel [5]. Conversely, for a probabilistic approach to interval pairwise comparisons the reader can refer to Saaty and Vargas [119].

4.3. Fuzzy AHP

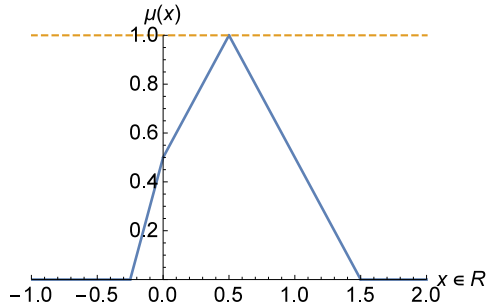
The *fuzzy AHP* is an even more popular methodology to account for uncertainty. In the fuzzy AHP entries of the pairwise comparison matrices are expressed in the form of fuzzy numbers. A function $\mu:\mathbb{R}\to[0,1]$ is a *fuzzy number* if and only if there exists an x_0 such that $\mu(x_0)=1$ and all the upper level sets of μ are convex, i.e. the set $\{x\in\mathbb{R}|\mu(x)\geq\alpha\}$ is convex for all $0<\alpha\leq 1$. Figure 4.6 reports some instances of fuzzy numbers. Also a real interval can be treated as a fuzzy number; considering the interval $[a,b]\subset\mathbb{R}$, then the value of its membership function is 1 for all $x\in[a,b]$ and 0 otherwise. The fuzzy AHP draws from the theory of fuzzy sets initiated by Zadeh [137] and described, for instance, in the excellent monographs by Klir and Yuan [79]



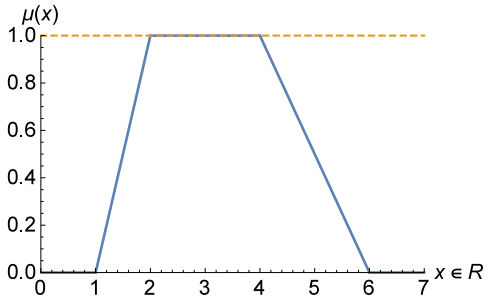
(a) Cuspidal fuzzy number



(b) Bell shaped fuzzy number



(c) Piecewise linear fuzzy number



(d) Trapezoidal fuzzy number

Figure 4.6.: Four examples of fuzzy numbers.

and Dubois and Prade [47]. Even so, to keep the description short and self-contained we shall here skip all the unnecessary details on fuzzy sets theory and go straight to the point.

4.3.1. Fuzzy AHP with triangular fuzzy numbers

One of the most used shapes of fuzzy numbers for modeling preferences, and more generally to represent uncertain quantities, is triangular. A *triangular fuzzy number* is defined by the following function

$$\mu(x) = \begin{cases} 0, & x \leq a^L \\ (x-a)/(b-a), & a^L \leq x \leq a^C \\ (c-x)/(c-b), & a^C \leq x \leq a^R \\ 0, & x \geq a^R \end{cases}$$

with $a^L \leq a^C \leq a^R$. Observe that there exists a one-to-one correspondence between triangular fuzzy numbers and triples $\tilde{a} = (a^L, a^C, a^R)$ with $a^L \leq a^C \leq a^R$. An example of triangular fuzzy number is reported in Figure 4.7. To many, the shape of a triangular fuzzy number might resemble a probability distribution, just with the normalization such that the area subtended by the curve is equal to one replaced by the condition $\sup_{x \in \mathbb{R}} \mu(x) = 1$. True, a fuzzy number can be seen as a distribution indicating the

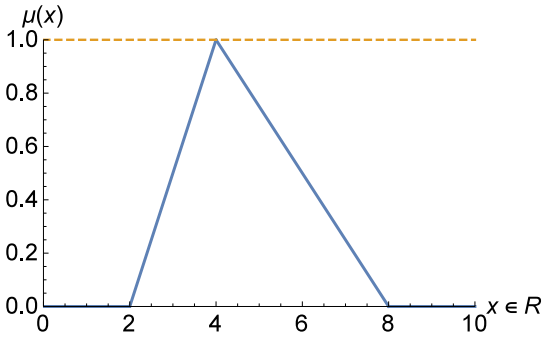


Figure 4.7.: Triangular fuzzy number $\tilde{a} = (2, 4, 8)$

likelihood of events, but within the framework of possibility theory [139], and not probability. According to Klir [77] the value $1 - \mu(x)$ can be interpreted as the degree of surprise to discover that x is the 'true' value of the variable under observation. Leaving aside the technicalities of this theory, for which the reader can be referred to the monograph by Klir [78], for practical purposes it is common to interpret the three values a^L, a^C, a^R characterizing a triangular fuzzy number as the smallest possible, the most likely, and the greatest possible values for the uncertain quantity under study. The use of triangular shapes for fuzzy numbers has been advocated by many, for instance Pedrycz [94], and a whole arithmetic has been developed to perform operations on fuzzy sets and fuzzy numbers in particular. In one of the first papers on fuzzy AHP, van Laarhoven and Pedrycz [82] defined the operations of addition ($\oplus$), multiplication ($\otimes$), logarithm ($\widetilde{\ln}$), inversion, and power as follows, respectively:

$$\begin{aligned} \tilde{a} \oplus \tilde{b} =& (a^L + b^L, a^C + b^C, a^R + b^R) \\ \tilde{a} \otimes \tilde{b} \approx& (a^L \cdot b^L, a^C \cdot b^C, a^R \cdot a^R) \\ \widetilde{\ln}(a_L, a_C, a_R) \approx& (\ln a_L, \ln a_C, \ln a_R) \\ \tilde{a}^{-1} \approx& \left(\frac{1}{a^R}, \frac{1}{a^C}, \frac{1}{a^L}\right) \\ e^{(a^L, a^C, a^R)} =& (e^{a^L}, e^{a^C}, e^{a^R}) \end{aligned} \tag{4.13}$$

The primal issue with a fuzzy pairwise comparison matrix is that of deriving the priority vector, and one straightforward approach could be that of using these operations on matrices with triangular fuzzy entries, i.e. $\tilde{\mathbf{A}} = (\tilde{a}_{ij})_{n \times n} = (a^L_{ij}, a^C_{ij}, a^R_{ij})$, in the same way their corresponding operations were used with pairwise comparison matrices. Hereafter, we shall focus on the problem of finding a suitable priority vector for a fuzzy pairwise comparison matrix. To this scope, we should distinguish *a priori* between two types of methods:

- Methods to derive a vector of fuzzy weights.
- Methods to derive a vector of weights expressed as real numbers.

We shall here dwell a bit more on these two methodologies by explaining how they have been treated in the literature.

Obtaining fuzzy weights

One straightforward solution to this problem was recently suggested by Ramík and Korviny [99]. According to this method, the components of the priority vector are fuzzy numbers and can be estimated by an extension of the geometric mean method. Namely, the priority vector appears as $\tilde{\mathbf{w}} = (\tilde{w}_1, \ldots, \tilde{w}_n)^T$, where the components $\tilde{w}_i = (w_i^L, w_i^C, w_i^R)$ themselves are triangular fuzzy numbers. Following this method, the priority vector with triangular fuzzy components is estimated as the minimizer of the following constrained optimization problem.

$$\begin{aligned} \underset{(\tilde{w}_1,\ldots,\tilde{w}_n)}{\text{minimize}} \quad & \sum_{i=1}^{n}\sum_{j=1}^{n}\Big(\big(\ln a_{ij}^L - \ln w_i^L + \ln w_j^L\big)^2 + \big(\ln a_{ij}^C - \ln w_i^C + \ln w_j^C\big)^2 + \\ & \big(\ln a_{ij}^R - \ln w_i^R + \ln w_j^R\big)^2\Big) \\ \text{subject to} \quad & \sum_{i=1}^{n} w_i^C = 1, \\ & w_i^U \geq w_i^C \geq w_i^L > 0 \ \forall i. \end{aligned} \tag{4.14}$$

Ramík and Korviny proved (see Theorem 1 in their paper [99]) that the analytic solution of this optimization problem is

$$w_k^L = c_{\min} \cdot \frac{\left(\prod_{j=1}^{n} a_{ij}^L\right)^{\frac{1}{n}}}{\sum_{i=1}^{n}\left(\prod_{j=1}^{n} a_{ij}^C\right)^{\frac{1}{n}}} \ \forall k, \tag{4.15}$$

$$w_k^C = \frac{\left(\prod_{j=1}^{n} a_{ij}^C\right)^{\frac{1}{n}}}{\sum_{i=1}^{n}\left(\prod_{j=1}^{n} a_{ij}^C\right)^{\frac{1}{n}}} \ \forall k, \tag{4.16}$$

$$w_k^R = c_{\max} \cdot \frac{\left(\prod_{j=1}^{n} a_{ij}^R\right)^{\frac{1}{n}}}{\sum_{i=1}^{n}\left(\prod_{j=1}^{n} a_{ij}^C\right)^{\frac{1}{n}}} \ \forall k, \tag{4.17}$$

where

$$c_{\min} = \min_{i=1,\ldots,n}\left\{\frac{\left(\prod_{j=1}^{n} a_{ij}^C\right)^{\frac{1}{n}}}{\left(\prod_{j=1}^{n} a_{ij}^L\right)^{\frac{1}{n}}}\right\} \qquad c_{\max} = \max_{i=1,\ldots,n}\left\{\frac{\left(\prod_{j=1}^{n} a_{ij}^C\right)^{\frac{1}{n}}}{\left(\prod_{j=1}^{n} a_{ij}^U\right)^{\frac{1}{n}}}\right\}$$

Example 24. *Consider the following matrix*

$$\tilde{\mathbf{A}} = (\tilde{a}_{ij})_{3\times 3} = \begin{pmatrix} (1,1,1) & (1/2,2,3) & (1,1,2) \\ (1/3,1/2,2) & (1,1,1) & (1/3,2,4) \\ (1/2,1,1) & (1/4,2,3) & (1,1,1) \end{pmatrix} \tag{4.18}$$

Then, the weight vector obtained by using (4.15)–(4.17) is

$$\tilde{\mathbf{w}} = (\tilde{w}_1, \tilde{w}_2, \tilde{w}_3)^T = \begin{pmatrix} (0.412599, 0.412599, 0.412599) \\ (0.249914, 0.32748, 0.454124) \\ (0.259921, 0.259921, 0.32748) \end{pmatrix} \tag{4.19}$$

For a critical analysis of this method and a broader overview on the use of fuzzy sets in decision making, the interested reader can refer to the recent paper by Dubois [45].

One method was proposed by van Laarhoven and Pedrycz [82] themselves, but a lot has happened since then and their proposal has been refined a number of times. Here we should present one of the most recent refinement, which can be seen as a fuzzy extension of the geometric mean method in the optimization form that we encountered in (2.3) and seemingly resembles the optimization problem (4.14). Note that, again, the solution is itself a priority vector whose components are triangular fuzzy numbers and is here denoted as $\tilde{\mathbf{w}} = (\tilde{w}_1, \ldots, \tilde{w}_n)^T$ with $\tilde{w}_i = (w_i^L, w_i^C, w_i^R)$.

$$\begin{aligned} \underset{(\tilde{w}_1,\ldots,\tilde{w}_n)}{\text{minimize}} \quad & \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} \Big(\big(\ln a_{ij}^L - \ln w_i^L + \ln w_j^L\big)^2 + \big(\ln a_{ij}^C - \ln w_i^C + \ln w_j^C\big)^2 + \\ & \big(\ln a_{ij}^R - \ln w_i^R + \ln w_j^R\big)^2 \Big) \\ \text{subject to} \quad & w_i^L + \sum_{j=1, j\neq i}^{n} w_j^U \geq 1, \ \forall i \\ & w_i^U + \sum_{j=1, j\neq i}^{n} w_j^L \leq 1, \ \forall i \\ & \sum_{i=1}^{n} w_i^C = 1, \\ & \sum_{i=1}^{n} \big(w_i^L + w_i^R\big) = 2, \\ & w_i^U \geq w_i^C \geq w_i^L > 0 \ \forall i. \end{aligned} \tag{4.20}$$

Since all these methods return a vector $\tilde{\mathbf{w}}$ whose components are fuzzy numbers, the question on how to select the best alternative remains. In fact, if there exists a universally accepted order on the set $\mathbb{R}$—that is, given two different real numbers we can always say which one is the greatest—the situation is more ambiguous in the case of fuzzy numbers. Consider, for instance, the weights in (4.19). If it is intuitive to say that $\tilde{w}_1$ is greater than $\tilde{w}_3$, then the situation between $\tilde{w}_1$ and $\tilde{w}_2$ is much more ambiguous. Which one

should be considered greater, and which one be the best between x_1 and x_2? Although much research has been done on the topic, there is still not a meeting of minds on how to order fuzzy numbers. The interested reader can refer to Wang and Kerre [133] and Brunelli and Mezei [30] for an axiomatic and a numerical study of methods for ranking fuzzy numbers, respectively.

Obtaining a real-valued priority vector

From the literature, it seems that the problem of ranking fuzzy numbers and its ambiguity can be bypassed by using methods which recover real valued priority vectors. There are few doubts that the most popular method for deriving a real valued priority vector $\mathbf{w}$ for a pairwise comparison matrix with fuzzy entries $\tilde{\mathbf{A}}$ is the so called *extent analysis*, proposed by Chang [37]. The extent analysis can be described in five algorithmic steps.

1. For each row, calculate its sum $\tilde{s}_i = \tilde{a}_{i1} \oplus \cdots \oplus \tilde{a}_{in}$.
2. Normalize all the $\tilde{s}_i$'s in the following way: $\tilde{r}_i = \tilde{s}_i \otimes (\tilde{s}_1 \oplus \cdots \oplus \tilde{s}_n)^{-1}$.
3. Calculate the degree of possibility that $\tilde{r}_i$ be greater than $\tilde{r}_j$ as follows

$$\text{Pos}\left(\tilde{r}_i \gtrsim \tilde{r}_j\right) = \begin{cases} 1, & \text{if } r_i^C \geq r_j^C \\ \frac{r_i^R - r_j^L}{(r_i^R - r_i^C)+(r_j^C - r_j^L)}, & \text{if } r_i^C < r_j^C \text{ and } r_j^L \leq r_i^U \\ 0, & \text{otherwise} \end{cases}$$

   The second case looks cumbersome but has a simple geometric interpretation: it is the value of the membership function for which the 'right leg' of $\tilde{r}_i$ and the 'left leg' of $\tilde{r}_j$ intersect. The concepts is illustrated in Figure 4.8.

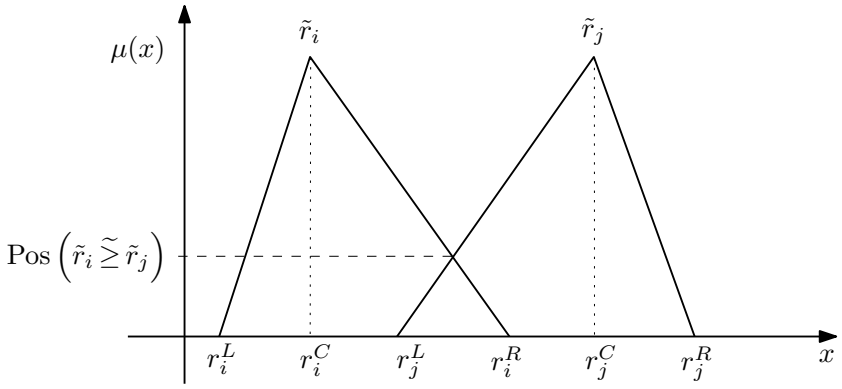


Figure 4.8.: Assessing the degree of possibility that the fuzzy number $\tilde{r}_i$ be greater than $\tilde{r}_j$.

4. Generalize the previous step by considering that

$$\text{Pos}\left(\tilde{r}_i \gtrsim \tilde{r}_j | j = 1, \ldots, n\,, j \neq i\right) = \min_{j \in \{1,\ldots,n\}, j \neq i} \text{Pos}\left(\tilde{r}_i \gtrsim \tilde{r}_j\right).$$

5. The real valued priority vector $\mathbf{w}$ is obtained by normalizing the values obtained in the previous steps:

$$w_i = \frac{\text{Pos}\left(\tilde{r}_i \widetilde{\geq} \tilde{r}_j | j = 1, \ldots, n\,, j \neq i\right)}{\sum_{k=1}^{n} \text{Pos}\left(\tilde{r}_i \widetilde{\geq} \tilde{r}_j | j = 1, \ldots, n\,, j \neq k\right)}.$$

Let us check the extent analysis method with a numerical example.

Example 25. *Consider the matrix* $\tilde{\mathbf{A}}$ *in (4.18) as the starting point. Then, the sums of the fuzzy numbers on the rows are calculated by means of the operation at step 1 and can be collected in the following vector,*

$$\begin{pmatrix} \tilde{s}_1 \\ \tilde{s}_2 \\ \tilde{s}_3 \end{pmatrix} = \begin{pmatrix} (2.5, 4, 6) \\ (5/3, 3.5, 7) \\ (7/4, 4, 5) \end{pmatrix}.$$

To normalize the components of this vector, one calculates $(\tilde{s}_1 \oplus \cdots \oplus \tilde{s}_n) = \left(\frac{71}{12}, 10, 18\right)$ *where* $71/12 = 2.5 + 5/3 + 7/4$ *and uses it to obtain, as described in the step 2,*

$$\begin{pmatrix} \tilde{r}_1 \\ \tilde{r}_2 \\ \tilde{r}_3 \end{pmatrix} = \begin{pmatrix} (0.138889, 0.4, 1.01408) \\ (0.0925926, 0.35, 1.1831) \\ (0.0972222, 0.25, 0.84507) \end{pmatrix}$$

Then we can construct the matrix of possibilities according to step 3

$$\mathbf{V} = \begin{pmatrix} - & 1 & 1 \\ 0.954305 & - & 1 \\ 0.824804 & 0.882695 & - \end{pmatrix} \tag{4.21}$$

where each nondiagonal entry is a value $\text{Pos}\left(\tilde{r}_i \widetilde{\geq} \tilde{r}_j\right)$*. Now, considering the algorithmic steps 4 and 5 together we can obtain the following priority vector,*

$$\mathbf{w} = \begin{pmatrix} \frac{1}{1+0.954305+0.824804} \\ \frac{0.954305}{1+0.954305+0.824804} \\ \frac{0.824804}{1+0.954305+0.824804} \end{pmatrix} = \begin{pmatrix} 0.359828 \\ 0.343385 \\ 0.296787 \end{pmatrix}.$$

More on the extent analysis will follow in the next section. For the moment it is sufficient to observe that, although an algorithm for ranking fuzzy numbers has not been explicitly mentioned, it has nevertheless been implicitly used. The matrix $\mathbf{V}$ in (4.21) is, *de facto*, a representation of a fuzzy ordering relation [138] which does induce a ranking on the fuzzy numbers $\tilde{r}_1, \tilde{r}_2, \tilde{r}_3$. Hence, we can conclude that, even by using the extent analysis, the ambiguity inherent to the ranking of fuzzy numbers is not avoided, but rather swept under the carpet.

4.3.2. Is the fuzzy AHP valid? ⋆

The question posed in the title of this subsection is as provocative as still standing, and it has definitely been answered in a negative sense by many. Since the seminal papers by van Laarhoven and Pedrycz [82] and Buckley [31], the fuzzy AHP has attracted most of the criticisms directed to the AHP plus a good deal of original others due to the (mis)use of fuzzy sets.

The first criticism is that the operations for triangular fuzzy numbers commonly used in the fuzzy AHP, and here reported in (4.13), are only approximations of the correct operations. The correct operations are defined by means of the extension principle and, according to these latter, for instance, the product of two triangular fuzzy numbers is *not* a triangular fuzzy number, but something nonlinear. Part of the scientific community accepts the approximations (4.13) as a necessary compromise to mitigate computational complexity while the other part does not. The reader can refer to Dubois and Prade [46] and Klir and Yuan [79] for a correct definition of arithmetic operations with fuzzy sets and fuzzy numbers.

Saaty and Tran [115] criticized the fuzzy AHP by saying that the traditional and real valued AHP suffices to account for all the imprecision in human judgments. Ramík and Korviny [99] rebutted that the traditional AHP can be seen as a special case of the fuzzy AHP—and not as a different method—and therefore it is difficult to see how the fuzzy AHP, which is more general, could perform worse than the AHP, which is the less general of the two.

A delicate point in the fuzzy AHP is that of ranking the components of the priority vector, when these are fuzzy numbers. Although it is not a real drawback, the fact that different ranking methods for fuzzy numbers could give very different results [30] can be perceived as a lack or robustness of the method, especially because there is not a prime ranking method. On the other hand, also methods which derive a real valued priority vector such as the extent analysis are not immune to criticisms. For instance, the extent analysis was criticized by Wang *et al.* [134] as, among many criticisms, they pointed out that the final weights are surely useful to rank alternatives but they cannot be interpreted as weights in a multiplicative sense. It is sufficient to see that in the extent analysis some weights can be equal to zero. However, in the case of null weights this does not mean that one alternative is infinitely better than another.

Recently, Zhü [141] moved some criticisms to the AHP, but it seems that many of them are pretentious and unsupported and others stem from a very narrow view of the method, which differs in large part from the more open minded view offered in this booklet. For instance, one of the criticisms moved by Zhü is that there is not an unique inconsistency index for fuzzy pairwise comparison matrices. This, clearly, stems from a vision of the AHP (very much *à la* Saaty) where, as Zhü [141] admitted, the Consistency Index CI is considered as the only reasonable consistency index and all others considered inferior.

5. Conclusions

As reported by Saaty and Sodenkamp [114], in 2008 Saaty was awarded by the INFORMS for the inception and development of the AHP. Part of the motivation for the award was the following:

> The AHP has revolutionized how we resolve complex decision problems... the AHP has been applied worldwide to help decision makers in every conceivable decision context across both the public and private sectors, with literally thousands of reported applications.

Moreover, from a survey by Wallenius *et al.* [132] it seems that the AHP has been by far the most studied and applied MCDM method, at least judging by the number of publications. The reader should have noticed that only basic mathematical and technical knowledge is required to use the AHP. For instance, if we consider that the priority vector can be derived using the method of the normalized columns, mentioned in §2.1.3, and consistency can be estimated by using the harmonic consistency index, in §2.2.4, then one can use the AHP at a basic level by using only elementary operations! Nevertheless, in spite of this possible simplicity, it is difficult to find an aspect of the AHP, or of pairwise comparison matrices, which has not been object of heated debates. Many of these debates are still open and probably will be so for much longer. However, even if inconclusive, it would be a mistake to regard them as pointless, since they contributed to create awareness around the AHP. Still, for the same sake of awareness, in this last part we shall overview some aspects of the AHP which have not been considered in the exposition.

Analytic Network Process

The observant reader might have also noticed that one of the assumptions has not been relaxed yet. Let us do it now.

> A *single* decision maker is *perfectly rational* and can *precisely* express his preferences on *all pairs* of ***independent*** alternatives and criteria using *positive real numbers*.

It is possible that, in some decisions, two criteria might affect each other. For instance, considering the selection of a resort for holidays, one can envision that the two criteria 'cost' and 'environment' are not independent since, probably, *ceteris paribus*, the resort in the best environment will also be more expensive. The best-known methodology for

dealing with interdependencies between parts of the hierarchy is the *Analytic Network Process* (ANP), which can be seen as a generalization of the AHP. That is, the AHP is a special case of the ANP without dependencies. Although more general than the AHP, the ANP still lacks a fundamental discussion and an axiomatization. Also, a self contained exposition of the ANP would require the introduction of new concepts, some of which of difficult digestion for those who are not in the field. For these reason, and the fear of making a sloppy job and possibly not do justice to the method, we shall here not dwell on the ANP. The interested reader can refer to a dedicated book by Saaty and Vargas [120] or, for an easier and more superficial treatment, to the book by Ishizaka and Nemery [74]. Let us incidentally note that the term Analytic Network Process was not coined by Saaty but, instead, by Hämäläinen and Seppäläinen [66].

Alternative methods

The AHP is a decision analysis methodology, but it is not the only one. Although nowadays geographical distinctions are arguably meaningless, for historical reasons, in decision analysis there has been two schools, the American and the French [54]. Here we shall touch upon one method of each type.

- *Multi-Attribute Value Theory* (MAVT) belongs to the so-called American school of decision analysis [76] and assumes that alternatives are fully described by their attributes. Then, each attribute state is mapped into a real number, and finally the numerical expressions of the different attributes are aggregated into a unique representative value. We shall now change notation and consider x the alternative and x_i as the state of the ith attribute in the alternative x. Consider a car, represented by the following list of characteristics:

  $$x = (x_1, x_2, x_3) = (\text{blue}, 180, 3)$$

  where the attributes are 'color', 'max speed in km/h' and 'safety level', respectively. Consider X_i as the set of possible states of the ith attribute. Then, according to value theory, for each attribute, there is a function $u_i : X_i \to [0, 1]$. The greater the value, the greater the satisfaction of the attribute. Given the existence of these functions, we can suppose that the car represented by x be mapped into the following vector,

  $$\mathbf{x} = (u_1(\text{blue}), u_2(180), u_3(3)) = (0.5, 0.7, 0.6) \in [0, 1]^3.$$

  At this point, the values attained for the attributes, in this case three, are aggregated into a single value by means of a function $v : [0, 1]^3 \to [0, 1]$ and a single real number is used to synthesize the value of an alternative. This process is represented in Figure 5.1. Functions $u_1, \ldots, u_n, v$ are defined once for all, and therefore their application is automatic when new alternatives are considered. The selection rule is simple: the greater the value, the better the alternative, i.e. with

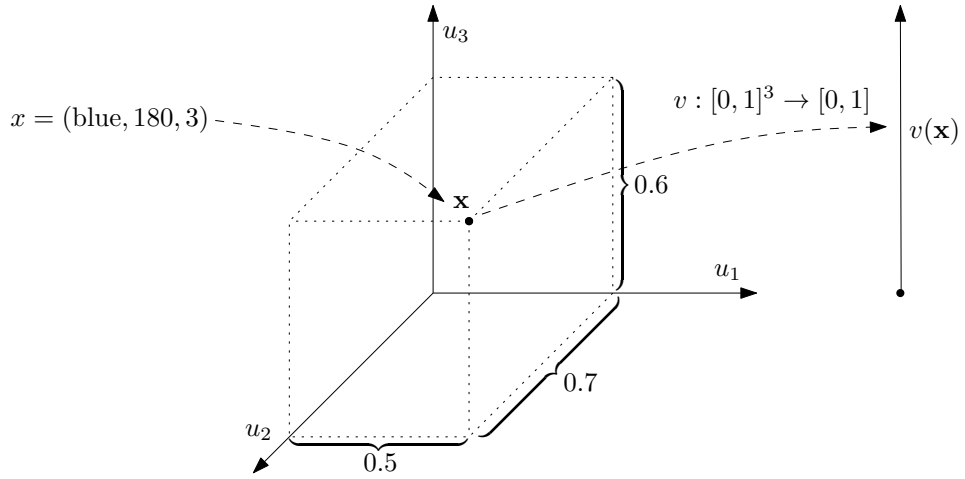


Figure 5.1.: A list of attribute states x is mapped into an $\mathbf{x} \in [0,1]^3$, which, in turn, is synthesized into $v(\mathbf{x})$.

$x = (x_1, \ldots, x_m)$ and $y = (y_1, \ldots, y_m)$, which are two alternatives described by m attributes,

$$x \succeq y \Leftrightarrow v(\underbrace{u_1(x_1), \ldots, u_m(x_m)}_{\mathbf{x}}) \geq v(\underbrace{u_1(y_1), \ldots, u_m(y_m)}_{\mathbf{y}})$$

A strength of value theory is its elegance and explanatory power on how decisions are made. Conversely, practical uses of this theory are limited by the difficulties in the estimation of the functions $u_1, \ldots, u_n, v$.

- The acronym *ELECTRE* stands for *ELimination Et Choix Traduisant la REalité*, and it is used to denote a family of methods from the French school. Nowadays many variants of the original ELECTRE methods exist and are applied to problems of ranking and also sorting. These methods are based on pairwise comparisons between alternatives, and to each comparison degrees of concordance and discordance are attached. A number of parameters and a non-trivial algorithm are necessary for the implementation of these methods, whose interpretation, possibly due to the aforementioned reasons, is not as straightforward as the one of the AHP.

Software

It is difficult to say whether much software appeared thanks to the popularity of the AHP or the popularity of the AHP is due to the wealth of software. Perhaps both propositions are to some extent true and the popularity of the method and of its software have gone arm-in-arm and boosted each other.

The foremost software is called *Expert Choice* and was first developed by Saaty and Forman in 1983. Expert Choice adopts Saaty's approach, according to which the priority vector is calculated with the eigenvector method and CI is used to estimate the inconsistency of preferences. Expert Choice was described and discussed by Ishizaka

and Labib [72]. Expert Choice's natural evolution and generalization to the ANP is called *SuperDecisions*. The name of the software comes from the 'supermatrix', which is a special matrix used in the ANP.
A direct and recent competitor of Expert Choice is *MakeItRational*, which was described by Ishizaka and Nemery [74]. One of the characteristics of MakeItRational is its ease of use, together with a captivating interface.
The software listed so far is not free and the user has to pay for its use. Among the free available software there is *HIerarchical PREference analysis on the World Wide Web* (Web-HIPRE) software which is part of Decisionarium [65], an online platform offering software for decision-making. Web-HIPRE allows the use of both the original scale of Saaty and the balanced scale (see §1.3). Two inconsistency indices can be used in Web-HIPRE: Saaty's CI and the index CM by Salo and Hämäläinen [122]. Web-HIPRE was the first online platform for decision making with the AHP and has a module which supports group decision making.
A comparative study between three software for the AHP was proposed by French and Xu [59]. Although other software exist, at present there is not an updated *and* free software for the AHP. An auspicable characteristic of such a free software is that it include different inconsistency indices, prioritization method, and methods to deal with with incomplete pairwise comparison matrices.

Sensitivity analysis

A module which is included in most AHP software allows for sensitivity analysis. In mathematical modeling, sensitivity analysis studies how the output of a mathematical model reacts to variations in the inputs. In the introductory chapter we encountered a numerical case where three weight vectors rating alternatives with respect to three criteria were aggregated using the weighs of criteria as factors in a linear combination.

$$\begin{aligned}\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} &= \hat{w}_1\mathbf{w}^{(c)} + \hat{w}_2\mathbf{w}^{(s)} + \hat{w}_3\mathbf{w}^{(e)} \\ &= \frac{1}{7}\begin{pmatrix} 4/9 \\ 4/9 \\ 1/9 \end{pmatrix} + \frac{2}{7}\begin{pmatrix} 6/10 \\ 3/10 \\ 1/10 \end{pmatrix} + \frac{4}{7}\begin{pmatrix} 1/11 \\ 2/11 \\ 8/11 \end{pmatrix} \approx \begin{pmatrix} 0.287 \\ 0.253 \\ 0.460 \end{pmatrix}.\end{aligned}$$

Now, we can assume that we want to see what happens to the final ranking of alternatives if we allow the weight of the third criterion to take values in $[0, 1]$ and rescale the weights of the other two criteria accordingly. In this case, the final rating can be expressed as follows and it becomes a function of $\hat{w}_3$,

$$\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \frac{1}{3}(1-\hat{w}_3)\begin{pmatrix} 4/9 \\ 4/9 \\ 1/9 \end{pmatrix} + \frac{2}{3}(1-\hat{w}_3)\begin{pmatrix} 6/10 \\ 3/10 \\ 1/10 \end{pmatrix} + \hat{w}_3\begin{pmatrix} 1/11 \\ 2/11 \\ 8/11 \end{pmatrix}.$$

For example the weight of the first alternative is

$$w_1 = \frac{74(1-\hat{w}_3)}{135} + \frac{\hat{w}_3}{11} = \frac{74}{135} - \frac{679}{1485}\hat{w}_3,$$

that is, an affine function of the weight of the third criterion $\hat{w}_3$. The same property of affinity holds also for w_2 and w_3 and, when the dimension of the problem allows it, sensitivity analysis lends itself nicely to graphical interpretations. In this case the graphical interpretation of w_1, w_2 and w_3 as functions of $\hat{w}_3$ is in Figure 5.2, which we should briefly comment.

The original weight assigned to $\hat{w}_3$ was $4/7 \approx 0.51$. From the picture we can see that, if

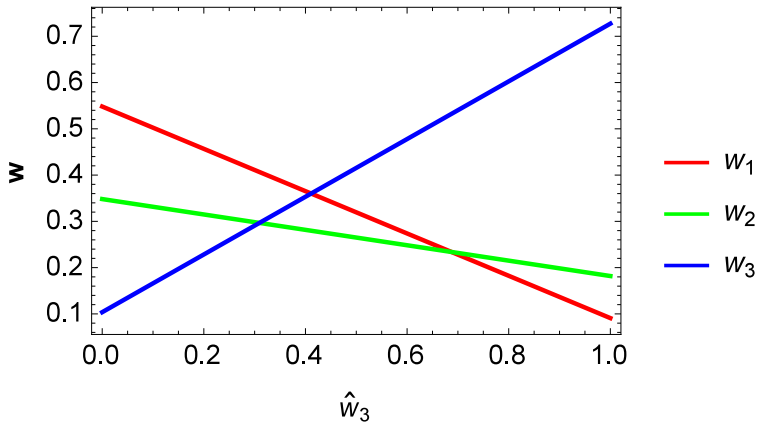


Figure 5.2.: Sensitivity analysis.

$\hat{w}_3 \geq 4/7$, then the solution is stable and alternative x_3 is always the best. Conversely, if the weight $\hat{w}_3$ is decreased, then, at some point, alternative x_1 will prevail. Sensitivity analysis is a precious tool for testing the robustness of solutions and their stability with respect to the inputs, in this case subjective judgments of experts. Moreover, here we have only presented the most popular way of performing sensitivity analysis, but it is easy to figure out that, by using the geometric mean method as the prioritization method, we can make the final ranking depend directly on entries of pairwise comparison matrices.

Future studies

The AHP is a fundamentally simple method which, in its simplest implementations, consists of three steps:

1. Problem structuring and definition of the hierarchy
2. Elicitation of pairwise comparisons
3. Derivation of priority vectors and their linear combinations.

In spite of its ease of interpretation, research has been going on for the last forty and more years and although many issues are still open, and perhaps are bound to be open for very long, nowadays it is safe to say that this technology has reached the maturity. We have seen in this booklet that a wide range of methods have been proposed to perform tasks within the AHP. Consider, for example, the wide range of methods for estimating the priority vector or the wealth of the inconsistency indices.

Unlike for some other areas of applied mathematics and mathematical modeling, in the case of the AHP, more often than not, new methods, indices, and extensions have been introduced heuristically and without results showing their originality and superiority. This practice generated an overabundance of material. In the future, it is auspicable that new numerical and axiomatic studies clarify and polish the state of the art, and when new methods are introduced, clear evidence on their originality and feasibility be provided.

Index

Bibliography

[1] Aczél, J., Saaty, T.L.: Procedures for synthesizing ratio judgments. Journal of Mathematical Psychology **27**(1), 93–102 (1983)

[2] Aguarón, J., Moreno-Jiménez, J.M.: The geometric consistency index: Approximated thresholds. European Journal of Operational Research **147**(1), 137–145 (2003)

[3] Alonso, J.A., Lamata, M.T.: Consistency in the analytic hierarchy process: a new approach. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems **14**(4), 445–459 (2006)

[4] Anand, S., Sen, A.: Human development index: methodology and measurement. Tech. rep., Human Development Report Office (HDRO), United Nations Development Programme (UNDP) (1994)

[5] Arbel, A.: Approximate articulation of preference and priority derivation. European Journal of Operational Research **43**(3), 317–326 (1989)

[6] Arbel, A., Vargas, L.: Interval judgments and Euclidean centers. Mathematical and Computer Modelling **46**(7), 976–984 (2007)

[7] Barfod, M.B., Leleur, S.: Scaling transformation in the REMBRANDT technique: examination of the progression factor. International Journal of Information Technology & Decision Making **12**(5), 887–903 (2013)

[8] Barzilai, J.: Deriving weights from pairwise comparison matrices. The Journal of the Operational Research Society **48**(12), 1226–1232 (1997)

[9] Barzilai, J.: Consistency measures for pairwise comparison matrices. Journal of Multi-Criteria Decision Analysis **7**(3), 123–132 (1998)

[10] Barzilai, J., Cook, W.D., Golany, B.: Consistent weights for judgements matrices of the relative importance of alternatives. Operations Research Letters **6**(3), 131–134 (1987)

[11] Barzilai, J., Golany, B.: AHP rank reversal, normalization and aggregation rules. INFOR-Information Systems and Operational Research **32**(2), 57–64 (1994)

[12] Basile, L., D'Apuzzo, L.: Weak consistency and quasi-linear means imply the actual ranking. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems **10**(3), 227–240 (2002)

[13] Basile, L., D'Apuzzo, L.: Transitive matrices, strict preference order and ordinal evaluation operators. Soft Computing **10**(10), 933–940 (2006)

[14] Belton, V., Gear, T.: On a short-coming of Saaty's method of analytic hierarchies. Omega **11**(3), 228–230 (1983)

[15] Belton, V., Gear, T.: The legitimacy of rank reversal—a comment. Omega **13**(3), 143–144 (1985)

[16] Bernasconi, M., Choirat, C., Seri, R.: The analytic hierachy process and the theory of measurement. Management Science **56**(4), 699–711 (2010)

[17] Bernasconi, M., Choirat, C., Seri, R.: Empirical properties of group preference aggregation methods employed in AHP: Theory and evidence. European Journal of Operational Research **232**(3), 584–592 (2014)

[18] Blanquero, R., Carrizosa, E., Conde, E.: Inferring weights from pairwise comparison matrices. Mathematical Methods of Operations Research **64**(2), 271–284 (2006)

[19] Bodin, L., Gass, S.I.: On teaching the analytic hierarchy process. Computers & Operations Research **30**(10), 1487–1497 (2003)

[20] Bodin, L., Gass, S.I.: Exercises for teaching the analytic hierarchy process. INFORMS Transactions on Education **4**(2), 1–13 (2004)

[21] Boyd, S.P., Vandenberghe, L.: Convex Optimization. Cambridge University Press (2004)

[22] Bozóki, S.: Solution of the least squares method problem of pairwise comparison matrices. Central European Journal of Operations Research **16**(4), 345–358 (2008)

[23] Bozóki, S.: Inefficient weights from pairwise comparison matrices with arbitrarily small inconsistency. Optimization: A Journal of Mathematical Programming and Operations Research **63**(12), 1893–1901 (2014)

[24] Bozóki, S., Dezső, L., Poesz, A., Temesi, J.: Analysis of pairwise comparison matrices: an empirical research. Annals of Operations Research **211**(1), 511–528 (2013)

[25] Bozóki, S., Fülöp, J., Rónyai, L.: On optimal completion of incomplete pairwise comparison matrices. Mathematical and Computer Modelling **52**(1), 318–333 (2010)

[26] Brunelli, M., Canal, L., Fedrizzi, M.: Inconsistency indices for pairwise comparison matrices: a numerical study. Annals of Operations Research **211**(1), 493–509 (2013)

[27] Brunelli, M., Critch, A., Fedrizzi, M.: A note on the proportionality between some consistency indices in the AHP. Applied Mathematics and Computation **219**(14), 7901–7906 (2013)

[28] Brunelli, M., Fedrizzi, M.: Axiomatic properties of inconsistency indices for pairwise comparisons. Journal of the Operational Research Society **66**(1), 1–15 (2014)

[29] Brunelli, M., Fedrizzi, M., Giove, S.: Reconstruction methods for incomplete fuzzy preference relations: A numerical comparison. In: WILF, pp. 86–93 (2007)

[30] Brunelli, M., Mezei, J.: How different are ranking methods for fuzzy numbers? A numerical study. International Journal of Approximate Reasoning **54**(5), 627–639 (2013)

[31] Buckley, J.J.: Fuzzy hierarchical analysis. Fuzzy Sets and Systems **17**(3), 233–247 (1985)

[32] Byun, D.H.: The AHP approach for selecting an automobile purchase model. Information & Management **38**(5), 289–297 (2001)

[33] Carlsson, C., Walden, P.: AHP in political group decisions: A study in the art of possibilities. Interfaces **25**(4), 14–29 (1995)

[34] Carmone Jr, F.J., Kara, A., Zanakis, S.H.: A Monte Carlo investigation of incomplete pairwise comparison matrices in AHP. European Journal of Operational Research **102**(3), 538–553 (1997)

[35] Cavallo, B., D'Apuzzo, L.: A general unified framework for pairwise comparison matrices in multicriterial methods. International Journal of Intelligent Systems **24**(4), 377–398 (2009)

[36] Cavallo, B., D'Apuzzo, L.: Characterizations of consistent pairwise comparison matrices over Abelian linearly ordered groups. International Journal of Intelligent Systems **25**(10), 1035–1059 (2010)

[37] Chang, D.Y.: Applications of the extent analysis on fuzzy AHP. European Journal of Operational Research **95**(3), 649–655 (1996)

[38] Chiclana, F., Herrera-Viedma, E., Alonso, S., Herrera, F.: Cardinal consistency of reciprocal preference relations: A characterization of multiplicative transitivity. IEEE Transactions on Fuzzy Systems **17**(1), 14–23 (2009)

[39] Choo, E.U., Wedley, W.C.: A common framework for deriving preference values from pairwise comparison matrices. Computers & Operations Research **31**(6), 893–908 (2004)

[40] Cook, W.D., Kress, M.: Deriving weights from pairwise comparison ratio matrices: An axiomatic approach. European Journal of Operational Research **37**(3), 355–362 (1988)

[41] Bana e Costa, C.A., Vansnick, J.C.: A critical analysis of the eigenvalue method used to derive priorities in AHP. European Journal of Operational Research **187**(3), 1422–1428 (2008)

[42] Crawford, G.: The geometric mean procedure for estimating the scale of a judgment matrix. Mathematical Modelling **9**(3-5), 327–334 (1989)

[43] Crawford, G., Williams, C.: A note on the analysis of subjective judgment matrices. Journal of Mathematical Psychology **29**(4), 387–405 (1985)

[44] De Baets, B., De Meyer, H., De Schuymer, B., Jenei, S.: Cyclic evaluation of transitivity of reciprocal relations. Social Choice and Welfare **26**(2), 217–238 (2006)

[45] Dubois, D.: The role of fuzzy sets in decision sciences: Old techniques and new directions. Fuzzy Sets and Systems **184**(1), 3–28 (2011)

[46] Dubois, D., Prade, H.: Operations on fuzzy numbers. International Journal of Systems Science **9**(6), 613–626 (1978)

[47] Dubois, D., Prade, H.: Fuzzy Sets and Systems: Theory and Applications, *Mathematics in Science and Engineering*, vol. 144. Academic Press (1980)

[48] Duszak, Z., Koczkodaj, W.W.: Generalization of a new definition of consistency for pairwise comparisons. Information Processing Letters **52**(5), 273–276 (1994)

[49] Dyer, J.S.: Remarks on the analytic hierarchy process. Management Science **36**(3), 249–258 (1990)

[50] Dyer, R.F., Forman, E.H.: Group decision support with the analytic hierarchy process. Decision Support Systems **8**(2), 99–124 (1992)

[51] Fedrizzi, M., Brunelli, M.: On the priority vector associated with a reciprocal relation and with a pairwise comparison matrix. Soft Computing **14**(6), 639–645 (2010)

[52] Fedrizzi, M., Giove, S.: Optimal sequencing in incomplete pairwise comparisons for large-dimensional problems. International Journal of General Systems **42**(4), 366–375 (2013)

[53] Fichtner, J.: On deriving priority vectors from matrices of pairwise comparisons. Socio-Economic Planning Sciences **20**(6), 341–345 (1986)

[54] Figueira, J., Greco, S., Ehrgott, M.: Multiple Criteria Decision Analysis: State of the Art Surveys, *International Series in Operations Research & Management Science*, vol. 78. Springer (2005)

[55] Fishburn, P.C.: Utility Theory for Decision Making. R. E. Krieger Pub. Co. (1979)

[56] Fishburn, P.C.: Preference relations and their numerical representations. Theoretical Computer Science **217**(2), 359–383 (1999)

[57] Forman, E., Peniwati, K.: Aggregating individual judgments and priorities with the analytic hierarchy process. European Journal of Operational Research **108**(1), 165–169 (1998)

[58] Fraleigh, J.B.: A First Course in Abstract Algebra, 7th edn. Pearson (2002)

[59] French, S., Xu, D.L.: Comparison study of multi-attribute decision analytic software. Journal of Multi-Criteria Decision Analysis **13**(2-3), 65–80 (2005)

[60] Galton, F.: Vox populi. Nature **75**, 450–451 (1907)

[61] Gass, S.I.: Model world: The great debate — MAUT versus AHP. Interfaces **35**(4), 308–312 (2005)

[62] Gass, S.I., Rapcsák, T.: Singular value decomposition in AHP. European Journal of Operational Research **154**(3), 573–584 (2004)

[63] Golden, B.L., Wang, Q.: An alternate measure of consistency. In: B.L. Golden, E.A. Wasil, P.T. Harker (eds.) The Analytic Hierarchy Process: Applications and Studies, pp. 68–81. Springer-Verlag (1989)

[64] Golden, B.L., Wasil, E.A., Levy, D.E.: Applications of the analytic hierarchy process: A categorized, annotated bibliography. In: The Analytic Hierarchy Process: Applications and Studies, pp. 37–58. Springer-Verlag (1989)

[65] Hämäläinen, R.: Decisionarium—aiding decisions, negotiating and collecting opinions on the web. Journal of Multi-Criteria Decision Analysis **12**(2-3), 101–110 (2003)

[66] Hämäläinen, R.P., Seppäläinen, T.O.: The analytic network process in energy policy planning. Socio-Economic Planning Sciences **20**(6), 399–405 (1986)

[67] Harker, P.T.: Derivatives of the Perron root of a positive reciprocal matrix: with application to the analytic hierarchy process. Applied Mathematics and Computation **22**(2), 217–232 (1987)

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

[69] Herrera-Viedma, E., Herrera, F., Chiclana, F., Luque, M.: Some issues on consistency of fuzzy preference relations. European Journal of Operational Research **154**(1), 98–109 (2004)

[70] Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press (1985)

[71] Ishizaka, A., Balkenborg, D., Kaplan, T.: Does AHP help us make a choice? An experimental evaluation. Journal of the Operational Research Society **62**(10), 1801–1812 (2011)

[72] Ishizaka, A., Labib, A.: Analytic hierarchy process and expert choice: Benefits and limitations. OR Insight **22**(4), 201–220 (2002)

[73] Ishizaka, A., Lusti, M.: How to derive priorities in AHP: a comparative study. Central European Journal of Operations Research **14**(4), 387–400 (2006)

[74] Ishizaka, A., Nemery, P.: Multi-criteria Decision Analysis: Methods and Software. Wiley (2013)

[75] Ji, P., Jiang, R.: Scale transitivity in the AHP. Journal of the Operational Research Society **54**(8), 896–905 (2003)

[76] Keeney, R.L., Raiffa, H.: Decisions with Multiple Objectives: Preferences and Value Tradeoffs. Wiley, New York (1976)

[77] Klir, G.J.: Uncertainty in economics: The heritage of G.L.S. Shackle. Fuzzy Economic Review **7**(2), 3–21 (2002)

[78] Klir, G.J.: Uncertainty and Information: Foundations of Generalized Information Theory. Wiley (2005)

[79] Klir, G.J., Yuan, B.: Fuzzy Sets and Fuzzy Logic: Theory and Applications. Pretience Hall (1995)

[80] Koczkodaj, W.W.: A new definition of consistency of pairwise comparisons. Mathematical and Computer Modelling **18**(7), 79–84 (1993)

[81] Koczkodaj, W.W., Orlowski, M.: An orthogonal basis for computing a consistent approximation to a pairwise comparisons matrix. Computers & Mathematics with Applications **34**(10), 41–47 (1997)

[82] van Laarhoven, P.J.M., Pedrycz, W.: A fuzzy extension of Saaty's priority theory. Fuzzy Sets and Systems **11**(1-3), 229–241 (1983)

[83] Lin, C.C.: A revised framework for deriving preference values from pairwise comparison matrices. European Journal of Operational Research **176**(2), 1145–1150 (2007)

[84] Lin, C.S., Harris, S.L.: A unified framework for the prioritization of organ transplant patients: Analytic hierarchy process, sensitivity and multifactor robustness study. Journal of Multi-Criteria Decision Analysis **20**(3-4), 157–172 (2013)

[85] Linstone, H.A., Turoff, M.: The Delphi Method: Techniques and Applications, vol. 29. Addison-Wesley Massachussets (1979)

[86] Lootsma, F.A.: Multi-Criteria Decision Analysis via Ratio and Difference Judgement, *Applied Optimization*, vol. 29. Springer (1999)

[87] Luce, R.D., Raiffa, H.: Games and Decisions. John Wiley and Sons (1957)

[88] Luce, R.D., Suppes, P.: Preference, utility, and subjective probability. Handbook of Mathematical Psychology **3**, 249–410 (1965)

[89] Maleki, H., Zahir, S.: A comprehensive literature review of the rank reversal phenomenon in the analytic hierarchy process. Journal of Multi-Criteria Decision Analysis **20**(3-4), 141–155 (2013)

[90] Mingers, J.: Soft OR comes of age—but not everywhere! Omega **39**(6), 729–741 (2011)

[91] von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press (1944)

[92] Nikou, S., Mezei, J.: Evaluation of mobile services and substantial adoption factors with analytic hierarchy process (AHP). Telecommunications Policy **37**(10), 915–929 (2013)

[93] Olson, D.L., Fliedner, G., Currie, K.: Comparison of the REMBRANDT system with analytic hierarchy process. European Journal of Operational Research **82**(3), 522–539 (1995)

[94] Pedrycz, W.: Why triangular membership functions? Fuzzy Sets and Systems **64**(1), 21–30 (1994)

[95] Peláez, J.I., Lamata, M.T.: A new measure of consistency for positive reciprocal matrices. Computers & Mathematics with Applications **46**(12), 1839–1845 (2003)

[96] Peniwati, K.: Criteria for evaluating group decision-making methods. Mathematical and Computer Modelling **46**(7-8), 935–947 (2007)

[97] Peniwati, K., Hsiao, T.: Ranking countries according to economics, social and political indicators. Mathematical Modelling **9**(3-5), 203–209 (1987)

[98] Pöhjönen, M.A., Hämäläinen, R.P., Salo, A.A.: An experiment on the numerical modelling of verbal ratio statements. Journal of Multi-Criteria Decision Analysis **6**(1), 1–10 (1997)

[99] Ramík, J., Korviny, P.: Inconsistency of pair-wise comparison matrix with fuzzy elements based on geometric mean. Fuzzy Sets and Systems **161**(11), 1604–1613 (2010)

[100] Saaty, T.L.: Mathematical Methods of Operational Research. McGraw Hill (1959)

[101] Saaty, T.L.: A scaling method for priorities in hierarchical structures. Journal of Mathematical Psychology **15**(3), 234–281 (1977)

[102] Saaty, T.L.: The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation. McGraw-Hill, New York (1980)

[103] Saaty, T.L.: The Logic of Priorities: Applications in Business, Energy, Health and Transportation. Kluwer-Nijhoff, Boston (1982)

[104] Saaty, T.L.: Absolute and relative measurement with the AHP. The most livable cities in the United States. Socio-Economic Planning Sciences **20**(6), 327–331 (1986)

[105] Saaty, T.L.: Axiomatic foundation of the analytic hierarchy process. Management Science **32**(7), 841–855 (1986)

[106] Saaty, T.L.: Decision Making for Leaders. The Analytic Hierarchy Process for Decisions in a Complex World. University of Pittsburgh Press (1988)

[107] Saaty, T.L.: Eigenvector and logarithmic least squares. European Journal of Operational Research **48**(1), 156–160 (1990)

[108] Saaty, T.L.: What is relative measurement? The ratio scale phantom. Mathematical and Computer Modelling **17**(4-5), 1–12 (1993)

[109] Saaty, T.L.: Relative measurement and its generalization in decision making why pairwise comparisons are central in mathematics for the measurement of intangible factors the analytic hierarchy/network process. RACSAM-Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas **102**(2), 251–318 (2008)

[110] Saaty, T.L., Alsina, C.: On synthesis of judgements. Socio-Economic Planning Sciences **20**(6), 333–339 (1986)

[111] Saaty, T.L., Hu, G.: Ranking by eigenvector versus other methods in the analytic hierarchy process. Applied Mathematics Letters **11**(4), 121–125 (1998)

[112] Saaty, T.L., Ozdemir, M.S.: Why the magic number seven plus or minus two. Mathematical and Computer Modelling **38**(3), 233–244 (2003)

[113] Saaty, T.L., Sagir, M.: Global awareness, future city design and decision making. Journal of Systems Science and Systems Engineering **21**(3), 337–355 (2012)

[114] Saaty, T.L., Sodenkamp, M.: The analytic hierarchy and analytic network measurement processes: the measurement of intangibles. In: Handbook of Multicriteria Analysis, pp. 91–166. Springer (2010)

[115] Saaty, T.L., Tran, L.T.: On the invalidity of fuzzifying numerical judgements in the analytic hierarchy process. Mathematical and Computer Modelling **46**(7-8), 962–975 (2007)

[116] Saaty, T.L., Vargas, L.G.: Hierarchical analysis of behavior in competition: Prediction in chess. Systems Research and Behavioral Science **25**(3), 180–191 (1980)

[117] Saaty, T.L., Vargas, L.G.: Comparison of the eigenvalue, logarithmic least squares and least squares methods in estimating ratios. Mathematical Modelling **5**(5), 309–324 (1984)

[118] Saaty, T.L., Vargas, L.G.: Modeling behavior in competition: The analytic hierarchy process. Applied Mathematics and Computation **16**(1), 49–92 (1985)

[119] Saaty, T.L., Vargas, L.G.: Uncertainty and rank order in the analytic hierarchy process. European Journal of Operational Research **32**(1), 107–117 (1987)

[120] Saaty, T.L., Vargas, L.G.: Decision Making with the Analytic Network Process: Economic, Political, Social and Technological Applications with Benefits, Opportunities, Costs and Risks, *International Series in Operations Research & Management Science*, vol. 195, 2nd edn. Springer (2013)

[121] Salo, A.A., Hämäläinen, R.: Preference programming through approximate ratio comparisons. European Journal of Operational Research **82**(3), 458–475 (1995)

[122] Salo, A.A., Hämäläinen, R.: On the measurement of preferences in the analytic hierarchy process. Journal of Multi-Criteria Decision Analysis **6**(6), 309–319 (1997)

[123] Shimura, M.: Fuzzy sets concept in rank-ordering objects. Journal of Mathematical Analysis and Applications **43**(3), 717–733 (1973)

[124] Shiraishi, S., Obata, T.: On a maximization problem arising from a positive reciprocal matrix in the AHP. Bulletin of Informatics and Cybernetics **34**(2), 91–96 (2002)

[125] Shiraishi, S., Obata, T., Daigo, M.: Properties of a positive reciprocal matrix and their application to AHP. Journal of the Operations Research Society of Japan **41**(3), 404–414 (1998)

[126] Shiraishi, S., Obata, T., Daigo, M., Nakajima, N.: Assessment for an incomplete comparison matrix and improvement of an inconsistent comparison: computational experiments. In: ISAHP, pp. 200–205 (1999)

[127] Stein, W.E., Mizzi, P.J.: The harmonic consistency index for the analytic hierarchy process. European Journal of Operational Research **177**(1), 488–497 (2007)

[128] Surowiecki, J.: The Wisdom of Crowds. Random House LLC (2005)

[129] Tanino, T.: Fuzzy preference orderings in group decision making. Fuzzy Sets and Systems **12**(2), 117–131 (1984)

[130] Vargas, L.G.: An overview of the analytic hierarchy process and its applications. European Journal of Operational Research **48**(1), 2–8 (1990)

[131] Vargas, L.G.: Comments on Barzilai and Lootsma: Why the multiplicative AHP is invalid: A practical counterexample. Journal of Multi-Criteria Decision Analysis **6**(3), 169–170 (1997)

[132] Wallenius, J., Dyer, J.S., Fishburn, P.C., Steuer, R.E., Zionts, S., Deb, K.: Multiple criteria decision making, multiattribute utility theory: Recent accomplishments and what lies ahead. Management Science **54**(7), 1336–1349 (2008)

[133] Wang, X., Kerre, E.E.: Reasonable properties for the ordering of fuzzy quantities (i). Fuzzy Sets and Systems **118**(3), 375–385 (2001)

[134] Wang, Y.M., Luo, Y., Hua, Z.: On the extent analysis for fuzzy AHP and its applications. European Journal of Operational Research **186**(2), 735–747 (2008)

[135] Xu, Z.: A survey of preference relations. International Journal of General Systems **36**(2), 179–203 (2007)

[136] Yager, R.R.: An eigenvalue method of obtaining subjective probabilities. Systems Research and Behavioral Science **24**(6), 382–387 (1979)

[137] Zadeh, L.A.: Fuzzy sets. Information and Control **8**, 338–353 (1965)

[138] Zadeh, L.A.: Similarity relations and fuzzy orderings. Information Sciences **3**(2), 177–200 (1971)

[139] Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems **1**(1), 3–28 (1978)

[140] Zahedi, F.: The analytic hierarchy process—A survey of the method and its applications. Interfaces **16**(4), 96–108 (1986)

[141] Zhü, K.: Fuzzy analytic hierarchy process: Fallacies of the popular methods. European Journal of Operational Research **236**(1), 209–217 (2014)

A. Eigenvalues and eigenvectors

The AHP is an important field of application of linear algebra, and especially of its theory regarding positive matrices. This appendix contains an introduction to eigenvalues and eigenvectors focused on their relevance for the AHP. At present, there are many ways to work out the AHP without getting dirty with eigenvalues and eigenvectors. Thus, in a certain sense knowing about them is superfluous. However, by knowing them the reader will figure out the connection between AHP and linear algebra and hopefully see the AHP from a higher observation point.

Definition 1 (Eigenvalues and eigenvectors)**.** *Consider an $n \times n$ square matrix $\mathbf{A}$ and an n-dimensional vector $\mathbf{w}$. Then, $\mathbf{w}$ and λ are an eigenvector and an eigenvalue of $\mathbf{A}$, respectively, if and only if*

$$\mathbf{A}\mathbf{w} = \lambda\mathbf{w}. \tag{A.1}$$

Example 26. *Consider the matrix and the vector as follows*

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 1/2 & 1 \end{pmatrix} \qquad \mathbf{w} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Then, one reckons that $\mathbf{w}$ is an eigenvector of $\mathbf{A}$ for $\lambda = 2$. In fact

$$\begin{pmatrix} 1 & 2 \\ 1/2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

.

Note that, if $\mathbf{w}$ is an eigenvector of $\mathbf{A}$, then all vectors $\alpha\mathbf{w}$ for $\alpha \in \mathbb{R}$ are also eigenvectors of $\mathbf{A}$, we call this set of vectors the *eigenspace* of $\mathbf{A}$ associated to that eigenvector (or its respective eigenvalue). Now one natural question arises; how to find the eigenvalues and the eigenvectors of a given matrix. By considering the identity matrix $\mathbf{I}$ and the null vector $\mathbf{0} = (0, \ldots, 0)^T$, we can rewrite (A.1),

$$\begin{aligned} \mathbf{A}\mathbf{w} &= \lambda\mathbf{w} \\ \mathbf{A}\mathbf{w} - \lambda\mathbf{w} &= \mathbf{0} \\ \mathbf{A}\mathbf{w} - \lambda\mathbf{I}\mathbf{w} &= \mathbf{0} \\ (\mathbf{A} - \lambda\mathbf{I})\,\mathbf{w} &= \mathbf{0} \end{aligned} \tag{A.2}$$

Now, from the basics of linear algebra we know that, if $\det(\mathbf{A} - \lambda\mathbf{I}) \neq 0$, then there is only one solution to (A.2), which is the trivial solution $\mathbf{w} = (0, \ldots, 0)^T$. We are instead interested in the case where other solutions exists, then to the case $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. Hence, by changing notation $\rho_{\mathbf{A}}(\lambda) := \det(\mathbf{A} - \lambda\mathbf{I})$, we need to find the roots of $\rho_{\mathbf{A}}(\lambda)$. Such a polynomial is called the *characteristic polynomial* of $\mathbf{A}$.

Example 27. *Reprising the matrix of the previous example*

$$\begin{aligned}\rho_{\mathbf{A}}(\lambda) &= \det\begin{pmatrix}1-\lambda & 2\\ 1/2 & 1-\lambda\end{pmatrix}\\ &= (1-\lambda)(1-\lambda) - \frac{1}{2}2\\ &= \lambda(\lambda-2)\end{aligned}$$

and by imposing $\lambda(\lambda-2)=0$ *it follows that* $\rho_{\mathbf{A}}(\lambda)=0$ *for* $\lambda=0,2$*. Now, considering for example the eigenvalue* $\lambda=2$ *the associated eigenvector can be found by solving*

$$\begin{pmatrix}1 & 2\\ 1/2 & 1\end{pmatrix}\begin{pmatrix}w_1\\ w_2\end{pmatrix} = 2\begin{pmatrix}w_1\\ w_2\end{pmatrix}$$

from which we derive that $w_1 = 2w_2$ *and that* $\mathbf{w} = (2,1)^T$ *is the eigenvector associated to* $\lambda=2$*. Clearly, also all the eigenvectors of the eigenspace spanned by* $\mathbf{w}$ *are eigenvectors of* $\lambda=2$*, e.g.* $(1,0.5)^T$*.*

Note that the eigenvalues can be ordered from the greatest to the smallest according to their absolute value. We call *maximum eigenvalue* the one with the greatest absolute value and we denote it as $\lambda_{\max}$. In Example 27, we have $\lambda_{\max}=2$. Going back to the computational part, with the increasing size of a matrix, things get more complicated, especially when it comes to find the roots of the characteristic polynomial. However, the idea remains the same.

Example 28. *Consider the following matrix*

$$\mathbf{A} = \begin{pmatrix}1 & 2 & 8\\ 1/2 & 1 & 4\\ 1/8 & 1/4 & 1\end{pmatrix}.$$

Then, by putting its characteristic polynomial equal to 0*, and by skipping the elementary steps, one recovers*

$$\rho_{\mathbf{A}}(\lambda) = \lambda^2(3-\lambda) = 0 \tag{A.3}$$

The eigenvalues are then $\lambda=0,3$*. In this case we say that the algebraic multiplicity of* $\lambda=0$ *is equal to* 2*. Roughly speaking, with* algebraic multiplicity *we indicate the number of times that a solution appears in the equation. In this case the multiplicity* 2 *of* $\lambda=0$ *is obvious if we rewrite (A.3) as follows,*

$$\rho_{\mathbf{A}}(\lambda) = \lambda\lambda(3-\lambda) = 0 \tag{A.4}$$

Note that in the previous example one eigenvalue was equal to n, and the other, with multiplicity $(n-1)$ was equal to 0. This is not a case, but a more general result.

Proposition 1. *Given a pairwise comparison matrix* $\mathbf{A}$*, if and only if* $\mathbf{A}$ *is consistent, then one eigenvalue,* $\lambda_{\max}$ *is equal to* n *and the other is equal to* 0*, with multiplicity* $(n-1)$*.*

Proceeding further, another question arises and regards the behavior of $\lambda_{\max}$ when $\mathbf{A}$ is not consistent. As $\lambda_{\max}$ cannot be equal to n, then what else can it be? Eigenvalues are roots of polynomials and it is natural to suspect that $\lambda_{\max}$ could be a complex number. Fortunately, this cannot happen for pairwise comparison matrices and we can restrict the search to real numbers. This is formalized in the following theorem.

Theorem 1 (Perron-Frobenius)**.** *Given a square matrix* $\mathbf{A}$*, if* $\mathbf{A}$ *is positive, i.e.* $a_{ij} > 0\ \forall i, j$*, then its maximum eigenvalue is real,* $\lambda_{\max} \in \mathbb{R}$*.*

Example 29. *Consider the following matrix*

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 8 \\ 1/2 & 1 & 1/4 \\ 1/8 & 4 & 1 \end{pmatrix}$$

Using the rule of Sarrus we compute

$$\begin{aligned} \rho_{\mathbf{A}}(\lambda) &= (1-\lambda)^3 + \left(2 \cdot \frac{1}{4} \cdot \frac{1}{8}\right) + \left(8 \cdot \frac{1}{2} \cdot 4\right) - (1-\lambda) - (1-\lambda) - (1-\lambda) \\ &= (1-\lambda)^3 + \frac{2}{32} + \frac{32}{2} - 3(1-\lambda) \\ &= \frac{225}{16} + 3\lambda^2 - \lambda^3. \end{aligned}$$

By solving $\frac{225}{16} + 3\lambda^2 - \lambda^3 = 0$ *we find that* $\lambda_{\max} \approx 3.9167$*. The other two roots are conjugate complex and we are not interested in them. Such solution can be easily found by any mathematical software. Now, with this solution, we need to solve the equation system*

$$\begin{pmatrix} 1 & 2 & 8 \\ 1/2 & 1 & 1/4 \\ 1/8 & 4 & 1 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = 3.9167 \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}.$$

To aid the process and avoid the problem of infinitely many solutions we add the condition $w_1 + w_2 + w_3 = 1$ *and solve*

$$\begin{cases} w_1 + 2w_2 + 8w_3 = 3.9167w_1 \\ \frac{1}{2}w_1 + w_2 + \frac{1}{4}w_3 = 3.9167w_2 \\ \frac{1}{8}w_1 + 4w_2 + w_3 = 3.9167w_3 \\ w_1 + w_2 + w_3 = 1 \end{cases}$$

from which we obtain

$$\mathbf{w} \approx (0.660761, 0.131112, 0.208127)^T.$$

Note that in the this last example $\mathbf{A}$ was inconsistent and $\lambda_{\max} > n$. The following proposition clarifies the range of possible values attained by $\lambda_{\max}$.

Proposition 2 (Saaty [101])**.** *Let* $\mathbf{A}$ *be a pairwise comparison matrix. Then* $\lambda_{\max} = n$ *if and only if* $\mathbf{A}$ *is consistent and strictly greater than* n *otherwise.*

Nowadays, all textbooks on linear algebra cover the theory of eigenvalues and eigenvectors. For a less didactic and more involving exposition of eigenvalues and eigenvectors with an eye on positive matrices the reader can refer to the book by Horn and Johnson [70].

B. Solutions

Solution to Problem 2: Consider that we assumed $a_{ij} = w_i/w_j \ \forall i, j$. Then we write w_i and w_j as the respective geometric means and see what happens if we account for the assumption.

$$a_{ij} = \frac{\left(\prod_{k=1}^{n} a_{ik}\right)^{1/n}}{\left(\prod_{k=1}^{n} a_{jk}\right)^{1/n}} = \left(\frac{a_{i1}a_{i2}\cdots a_{in}}{a_{j1}a_{j2}\cdots a_{jn}}\right)^{\frac{1}{n}}.$$

Since we assumed that $a_{ij} = w_i/w_j$ we can substitute these in the equation and rewrite it as

$$a_{ij} = \left(\frac{\frac{w_i}{w_1}\frac{w_i}{w_2}\cdots\frac{w_i}{w_n}}{\frac{w_j}{w_1}\frac{w_j}{w_2}\cdots\frac{w_j}{w_n}}\right)^{\frac{1}{n}} = \left(\frac{\frac{w_i^n}{w_1 w_2 \cdots w_n}}{\frac{w_j^n}{w_1 w_2 \cdots w_n}}\right)^{\frac{1}{n}} = \frac{w_i}{w_j}.$$

The original assumption is correctly recovered and therefore, when $a_{ij} = w_i/w_j \ \forall i, j$, the geometric mean method returns the correct vector.

Solution to Problem 3: The proof was provided by the Crawford and Williams [43]. See Theorem 3 in their paper.

Solution to Problem 4: Underbraced are the numbers of independent comparisons for each level of the hierarchy, starting from the top.

$$\underbrace{\frac{3(3-1)}{2}}_{3} + \underbrace{\frac{4(4-1)}{2} + \frac{3(3-1)}{2} + \frac{3(3-1)}{2}}_{12} + \underbrace{9\frac{3(3-1)}{2}}_{27} = 42$$

Solution to Problem 5: Consider the analytic formula of c_3, that is,

$$c_3 = \sum_{i=1}^{n-2}\sum_{j=i+1}^{n-1}\sum_{k=j+1}^{n}\left(2 - \frac{a_{ik}}{a_{ij}a_{jk}} - \frac{a_{ij}a_{jk}}{a_{ik}}\right).$$

At this point, consider the matrix $\dot{\mathbf{A}}$ with the entry $\dot{a}_{ij}$ missing. The sum contains four transitivities and we can expand it

$$\begin{aligned} c_3 &= 6\Bigg(8 \underbrace{- \frac{\dot{a}_{13}}{\dot{a}_{12}\dot{a}_{23}} - \frac{\dot{a}_{12}\dot{a}_{23}}{\dot{a}_{13}}}_{-\frac{37}{6}} \underbrace{- \frac{\dot{a}_{14}}{\dot{a}_{12}\dot{a}_{24}} - \frac{\dot{a}_{12}\dot{a}_{24}}{\dot{a}_{14}}}_{-\frac{\dot{a}_{14}}{2} - \frac{2}{\dot{a}_{14}}} \underbrace{- \frac{\dot{a}_{14}}{\dot{a}_{13}\dot{a}_{34}} - \frac{\dot{a}_{13}\dot{a}_{34}}{\dot{a}_{14}}}_{-\frac{\dot{a}_{14}}{8} - \frac{8}{\dot{a}_{14}}} \underbrace{- \frac{\dot{a}_{24}}{\dot{a}_{23}\dot{a}_{34}} - \frac{\dot{a}_{23}\dot{a}_{34}}{\dot{a}_{24}}}_{-\frac{13}{6}}\Bigg) \\ &= 6\left(-\frac{5x}{8} - \frac{10}{x} - \frac{1}{3}\right) = -\frac{15\dot{a}_{14}}{4} - \frac{60}{\dot{a}_{14}} - 2. \end{aligned}$$

Let us inspect the first and second derivatives of c_3 in $\dot{a}_{14}$:

$$\frac{\partial c_3}{\partial \dot{a}_{14}} = \frac{60}{(\dot{a}_{14})^2} - \frac{15}{4},$$
$$\frac{\partial^2 c_3}{\partial \dot{a}_{14}^2} = -\frac{120}{(\dot{a}_{14})^3}.$$

The second derivative is strictly negative for positive values of $\dot{a}_{14}$, which means that the function is strictly concave for $\dot{a}_{14} > 0$ and that, if there is a maximum, then it is unique. By equating the first derivative to zero, we recover that $(\dot{a}_{14})^2 = 16$. Of the two solutions we take the positive one, which is $\dot{a}_{14} = 4$.

Solution to Problem 6:

$$\mathbf{C} = \begin{pmatrix} 3 & 2 & 0 & 0 \\ 1/2 & 1 & 1/3 & 1 \\ 0 & 3 & 2 & 2 \\ 0 & 1 & 1/2 & 2 \end{pmatrix}$$

Solution to Problem 7:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 4 & 3 \\ 1/2 & 1 & 2 & 3/2 \\ 1/4 & 1/2 & 1 & 3/4 \\ 1/3 & 2/3 & 4/3 & 1 \end{pmatrix}$$

Solution to Problem 8: If the matrix is consistent, then any column can act as the priority vector. An alternative method, which is also used to derive the vector from inconsistent matrices, is the arithmetic mean of the rows,

$$u_i = \frac{1}{n} \sum_{j=1}^{n} p_{ij}.$$

In the case of consistent matrices, this method returns exactly the correct vector. The proof is similar to the one used to solve Problem 2.