USABILITY OF WEBSITES

by

ANGELA LILIANA MONTENEGRO VILLOTA

A thesis submitted to
The University of Birmingham
for the degree of
MASTER OF PHILOSOPHY

School of Mathematics
The University of Birmingham
October 2009

ABSTRACT

Due to its reach, acceptance and capability to share information, the World Wide Web has become in an important tool for business. Millions of websites have been developed and so inherently we can come across every kind of website from easy to hard-to-use. There are some so-called usability criteria, which should be respected by web designers in order to make websites useful. Using a multicriteria decision making approach, we evaluate the performance, based on 7 usability criteria, of 5 websites from which one can buy books online.

The complexity of multicriteria decision making is based on the fact that those multiple criteria are often contradicting with each other, and so a solution that optimises every criterion simultaneously, or an ideal solution, is generally unfeasible. In this situation making a decision implies giving an answer which without being optimal is still satisfactory.

Considering usability as a subjective matter, we use two well-known methodologies that deal with this issue: Analytic Hierarchy Process (AHP) and PROMETHEE. Based on pairwise comparison matrices AHP transforms subjective judgements into quantified ratios of importance. PROMETHEE relates the preference of a decision maker with specially defined criterion functions. We analyse deeply the mathematical background behind AHP.

ACKNOWLEDGMENTS

Firstly I would like to express my acknowledgements for the advice and patience of my supervisor Dr Sándor Németh.

A large debt of gratitude is also due to my friends and family especially to my mother for her invaluable encouragement and support.

Finally, I would like to acknowledge the foundation COLFUTURO of Colombia for their financial support.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOTATION AND TERMINOLOGY

We introduce notation and terminology here as a reference only. Some of this terminology will be introduced later as definitions, when they are needed.

Description	Notation
Complex numbers	$\mathbb{C}$
Real numbers	$\mathbb{R}$
Field or set of scalars ($\mathbb{R}$ or $\mathbb{C}$)	$\mathbb{F}$
Sets	A, B, C, etc.
Vector $\mathbf{v}$	$\mathbf{v} = (v_1, v_2, \ldots, v_n)^{\mathbf{T}}$
Vector with 1 in the ith component and zeros in all the other entries	$\mathbf{e}_i$
Length of a vector $\mathbf{v}$ (Euclidean norm)	$\|\mathbf{v}\| = (\|v_1\|^2 + \ldots + \|v_n\|^2)^{1/2}$
Inner product of vectors $\mathbf{v}$ and $\mathbf{w}$	$\langle \mathbf{v}, \mathbf{w} \rangle = v_1 w_1 + \ldots + v_n w_n$
Matrices	$\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, etc.
Elements of a matrix $\mathbf{A}$	a_{ij}
Determinant of a matrix $\mathbf{A}$	$\det(\mathbf{A})$
Rank of a matrix $\mathbf{A}$	$\mathrm{rank}(\mathbf{A})$
Nullity of a matrix $\mathbf{A}$	$\mathrm{nullity}(\mathbf{A})$
Inverse of a matrix $\mathbf{A}$	$\mathbf{A}^{-1}$
Transpose of a matrix $\mathbf{A}$	$\mathbf{A}^{\mathbf{T}}$
Conjugate transpose of a matrix $\mathbf{A}$	$\mathbf{A}^*$
Identity matrix	$\mathbf{I}$
Diagonal matrix	$\mathrm{diag}(x_1, \ldots, x_n)$
Principal submatrix of a matrix $\mathbf{A}$	$\mathbf{A}_i$
Sum of principal minors of a matrix $\mathbf{A}$	$E_k(\mathbf{A})$
Elementary symmetric polynomials in variables λ_i	$e_k(\lambda_1, \ldots, \lambda_n)$
Spectral radius of a matrix $\mathbf{A}$	$\rho(\mathbf{A}), \rho_{\mathbf{A}}$
Number of nonzero eigenvalues of a matrix $\mathbf{A}$	$e(\mathbf{A})$
Vector norm	$\| \cdot \|$
Matrix norm	$\boldsymbol{\| \cdot \|}$

PART I

INTRODUCTION

CHAPTER 1

INTRODUCTION

1.1 About the project

Usability of websites refers to the evaluation of the performance of a website based on usability criteria. Since there are several criteria involved, this problem has been formulated as a multicriteria decision making problem. Its solution has been achieved using two highly renowned techniques for multicriteria decision making problems, which are Analytic Hierarchy Process (AHP) and PROMETHEE methodology. We analyse deeply the AHP methodology.

1.2 Methodology of the project

The project comprises three main parts. In Part I we give a general introduction to multicriteria decision making problems. A rough description of two methodologies frequently used for solving multicriteria decision making problems, AHP and PROMETHEE, is presented in Chapter 2. In Chapter 3 we introduce usability criteria and we explain the most common criteria used.

Part II comprises the mathematical justification and explanation of the project. We analyse deeply the mathematical background behind AHP in Chapter 4 and we provide an explanation of PROMETHEE methodology in Chapter 5.

In Part III, we apply the methodologies described in Part II specifically to solve the problem of usability of websites stated in this study. In Chapter 6 a decision model of the problem is defined, specifying a goal, determining the evaluation criteria, and choosing some available websites as the alternatives. A solution to the problem by using AHP and PROMETHEE methodologies is presented in Chapter 7. To determine the weights of the criteria AHP is used. Then, having found the weights of the criteria, an evaluation of the alternatives with respect to the criteria, and the ranking of the alternatives to make the final decision, is made by both AHP and PROMETHEE methodologies. Chapter 8 provides a comparison of the results obtained by using those two methodologies.

CHAPTER 2

ABOUT MULTICRITERIA DECISION MAKING

2.1 General remarks

The conventional decision making problem takes place in a scenario limited by the availability of resources, which establishes the constraints of the problem. In this situation, the decision is restricted to consider only those feasible values that the decision variables can take without breaking any of the constraints. The decision making problem is associated with a defined criterion function f (Ballestero and Romero (1998)), and the decision variables are represented by x. A final decision is made considering the benefits that a particular choice of x gives, and so it is sensible to define the goal of the problem, without loss of generality, as a maximisation problem, i.e.

$$\text{Maximise} \quad f(x) : x \in \mathbb{R} \tag{2.1}$$

The first interest in solving decision making problems was based on the existence of a single criterion, and in this case the solution of the problem was the optimisation of that single criterion. However, decisions generated in this way, even if they are optimum, are still often far from a realistic scenario. In this manner, a multicriteria decision making area emerged to study those decision problems that are not only defined by one criterion, as an objective function, but by multiple criteria.

Considering those multiple criteria as n functions, the ideal goal of a multicriteria decision making problem can be represented as

$$\text{Maximise} \quad f(x) : x \in \mathbb{R} \tag{2.2}$$

with

$$f(x) = [f_1(x), f_2(x), \ldots, f_n(x)]^{\mathbf{T}} \tag{2.3}$$

The area of multicriteria decision making is in fact related with optimisation (Dyer et al. (1992)). Nevertheless, due to the nature of the systems analysed in this case, unsurprisingly, its complexity is based on the existence of multiple criteria that are often contradicting with each other. Then, the optimisation for all criteria simultaneously, or ideal solutions, are generally unfeasible (Zeleny (1998)). In this case, the concept of optimality has been reduced to the idea of finding a solution that, without being an optimal one, provides a satisfactory solution. Satisfactory here refers to the fact that it is not possible to find a strictly better solution for the problem.

Multicriteria decision making is one of the most studied cases in the decision making area (Triantaphyllou (2000)), and had an outstanding growth in the 1980s (Dyer et al. (1992)). Since then several investigations have been developed in terms of both theoretical and practical dimensions (Ballestero and Romero (1998)). Nevertheless there is still much to be done in the area towards the goal of finding optimal solutions, if they exist, or at least near-optimal solutions for this kind of problem.

There are many different approaches for solving multicriteria decision making problems. Larichev (2002) presents a classification of these methods into four categories:

1. *Methods based on quantitative measurements, generally based on utility theory.*
2. *Methods based on initial qualitative assessments that are transformed into quantitative variables.*
3. *Methods based on quantitative measurements but using a preference method to compare the alternatives.*
4. *Methods based on qualitative assessments but not using a transformation into quantitative variables.*

The problem stated in this study is going to be solved using two well-known techniques, which are Analytic Hierarchy Process, included in the second category, and the PROMETHEE method, included in the third category.

2.2 Analytic Hierarchy Process

The Analytic Hierarchy Process (AHP) is a methodology proposed by Thomas L. Saaty in the 1970s for modelling complex multicriteria decision making problems under a hierarchy of importance, analysing contradicting and interconnected components (Saaty (1990)), with the aim of making an informed decision.

Creating a hierarchy according with the relations of the elements of the problem, and synthesising the subjective judgements into quantitative judgements, AHP has captured the interest of researchers around the world, and has its applications in planning, selecting alternatives, resource allocations, resolving conflict, optimisation, in fields such as the manufacturing sector, politics, engineering, education, industry, government and others (Vaidya and Kumar (2006)). AHP is one of the most used techniques in multiple criteria decision making (Vaidya and Kumar (2006)) because of its simplicity in evaluating discrete alternative problems (Steuer and Na (2003)).

AHP methodology copies human behaviour when dealing with a complex decision, trying to decompose the complexity into simple entities that can be associated according to common characteristics. Then the method focuses on the relations between the identities by making a comparison between them as quantified judgements expressed in ratios of importance (Saaty (1990)).

Having obtained these ratios, the use of matrix theory helps to determine the specific value associated with each simple entry, and so the decision maker is able to make an informed decision. We will introduce a detailed description of AHP methodology in Chapter 4.

2.3 PROMETHEE

Preference Ranking Organisation Method for Enrichment Evaluation, better known as PROMETHEE, was introduced by Jean-Pierre Brans in 1982, to construct a relation in terms of importance or an outranking relation (Brans et al. (1986)) between some available alternatives. Since then, several versions of this methodology have been introduced

(I, partial ranking; II, complete ranking; III, ranking based on intervals; IV, continuous case; V, consideration of constraints and VI, sensitivity analysis procedure).

Due to its simplicity, adaptability (Goumas and Lygerou (2000)) and mathematical properties (Brans and Mareschal (2005)), PROMETHEE is widely recognised among the outranking methods (Wim De Keyser (1996)), and has been used in fields like banking, industrial location, manpower planning, water resources, investments, medicine, chemistry, health care, tourism, Operations Research and dynamic management, among others (Brans and Mareschal (2005)).

When dealing with decision models the degree of preference of the decision maker has to be determined. PROMETHEE models reflect the degree of preference of the decision maker by using utility functions which are supported by an economic foundation (Brans et al. (1986)). To estimate the parameters of these functions, the decision maker develops his own scale by fixing the selected parameters (Brans et al. (1986)). PROMETHEE methodology will be introduced in Chapter 5.

CHAPTER 3

ABOUT USABILITY CRITERIA

3.1 General remarks

About twenty years ago the World Wide Web was created, (Jacobs (2008)), and since then has been in continuous growth, with such success that nowadays it has become in an important tool to do business, due to its reach, and different interactive capabilities (Ranganathan and Ganapathy (2002)). In fact, as Porter (2001) expressed Internet technology offers the opportunity to make a company competitive in the market.

Nevertheless, in order to make that happen it is necessary to understand that the critical issue is not to make the decision of using the Internet but how to use it (Porter (2001)). This explains why several studies have focused on the interaction between users and computers which constitute the most basic principle that the use of the internet entails.

The interaction between users and computers arises with the purpose of the accomplishment of a task (Card et al. (1983)). Within this interaction, the concepts of efficiency and ease of use emerged as indicators of the need to design better human-computer interfaces. In this situation, in the approach to human computer interaction presented by Card et al. (1983) a new concept, *web usability analysis*, was developed.

Since then, and with the increasing global competition, usability has captured the interest not only of researchers but website managers and designers around the world. A number of studies have been undertaken to evaluate the performance of a website based on usability criteria. At the same time the concept of usability has been defined in different ways. However it has been seen always as a *quality* factor. Seffah et al. (2006) listed some definitions of usability from three different standards:

> *"A set of attributes that bear on the effort needed for use and on the individual assessment of such use, by a stated or implied set of users"*
>
> (ISO/IEC 9126, 1991)
>
> *"The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use"*
>
> (ISO 9241-11, 1998)
>
> *"The ease with which a user can learn to operate, prepare inputs for, and interpret outputs of a system or component"*
>
> (IEEE 610.12-1990)

In spite of the number of investigations related with usability, there are no standard usability criteria. This is due to the fact that researchers have proposed different usability criteria which vary in terms of number of dimensions, degree of generality or specificity, and level of precision (Park and Lim (1999)). Nonetheless, some common classifications used are those presented by Nielsen (2003), which establish learnability, efficiency, memorability, errors, and satisfaction as components; by Microsoft, who classified usability according to content, ease of use, promotion, made-for-the-medium, and emotional response as major categories (Keeker (1997)); or the classification presented by ISO who presented Standard usability criteria including effectiveness, efficiency, and satisfaction (International Organization for Standardization (1998)).

However what is similar in all the studies is that to estimate the value of usability it is necessary to consider a variety of different measures (Agarwal and Venkatesh (2002)), and since all criteria are not equally important, they have to be weighted according to their relative importance (Park and Lim (1999) and Nielsen (1993)). As a result, the context in which usability is being evaluated should be specified (Norros and Savioja (2004)). Moreover usability evaluation methods depend on subjective assessment in the form of user judgments (Agarwal and Venkatesh (2002)) that can be developed into quantitative measures (Norros and Savioja (2004)).

3.2 Usability criteria used in this project

For the purpose of this study we are going to consider as usability criteria accessibility, customisation and personalisation, download speed, ease of use, errors, navigation and site content, all of which are frequently mentioned in the literature (see Pearson and Pearson (2008), Turban and Gehrke (2000), and Keeker (1997)).

3.2.1 Accessibility

Accessibility refers to the availability of a website and is a necessary factor in order to let users access the content of a page (Pearson and Pearson (2008)). Moreover accessibility also refers to those different situations that designers should consider in order to make a page accessible no matter what agent is used by users, for example language, version of a browser, different browser, among others (World Wide Web Consortium (WC3) (1999)). Accessibility comprises the following subcriteria.

1. Availability to different agents (World Wide Web Consortium (WC3) (1999))
2. Alternatives for multimedia presentations (Texas A&M University (2004))
3. Readability (Texas A&M University (2004))
4. Frames identification (Texas A&M University (2004))
5. Skip-navigation links (to permit users to skip repetitive navigation links) (Texas A&M University (2004))

3.2.2 Customisation and personalisation

Customisation and personalisation is known as *made-for-the-medium* (e.g. Keeker (1997)) and refers to those characteristics of a website that fit a particular user's needs (Agarwal and Venkatesh (2002)). According to this, websites should provide dynamic content which has been adapted to a specific user (Pearson and Pearson (2008)). As subcriteria of customisation and personalisation we have the following.

1. Possibility of connection with other people (Keeker (1997))
2. Personalisation (Keeker (1997))
3. Refinement and addition of content over time(Keeker (1997))
4. Market research (Turban and Gehrke (2000))

3.2.3 Download speed

Suggested terms are *user response time* (e.g. Shneiderman (1998), Nielsen (2000)) or *download delay* (e.g. Rose et al. (1999), Palmer (2002), Erica S. Davis (2001)). This criterion is defined by Erica S. Davis (2001) as the "delay of instructional materials appearing on a web page after the page is accessed" and as a result can be affected by the content of a website (Pearson and Pearson (2008)). The importance of this criterion is due to the fact that users become frustrated if they have to wait more than a few seconds to access all the information on a website (Nielsen (1997)). Download speed is constituted by the following subcriteria.

1. Simple and meaningful use of graphics and tables (Turban and Gehrke (2000)).
2. Limited use of animation (Turban and Gehrke (2000))
3. Use of thumbnails (Turban and Gehrke (2000))

3.2.4 Ease of use

Ease of use is related with the effort that is required to use it (Venkatesh et al. (2003), Agarwal and Venkatesh (2002)). Ease of use has been seen as an important factor in determining user acceptance and behaviour in using a technology (Venkatesh (2000)). Among the subcriteria of ease of use we have the following.

1. Goals (prioritisation of the content) (Keeker (1997))
2. Structure of the website (Keeker (1997))
3. Feedback about the system status (Keeker (1997))

3.2.5 Errors

We consider the number of errors that users can make while using the webpage, how severe they are and how easy it is to recover from those errors (Nielsen (2003)). Among subcriteria of errors we have the following.

1. Number of errors (Nielsen (2003))
2. Severity of the errors (Nielsen (2003))
3. Recovering from errors (Nielsen (2003))

3.2.6 Navigation

Navigation is defined as the method used to find information within a web site (Koyani et al. (2004)) following a sequence of pages carefully organised (Palmer (2002)). In this situation users should be able to intuitively find what to do to follow the appropriate sequence into a website (Claiborne (2005)). Navigation comprises these subcriteria

1. Organisation (Palmer (2002))
2. Arrangement (Palmer (2002))
3. Layout (Palmer (2002))
4. Sequencing (Palmer (2002))

3.2.7 Site content

Site content concerns the accurate of the information provided. Also includes the quality of the content (Palmer (2002)).

1. Amount and variety of product information (Palmer (2002))

2. Relevance of the content (useful)(Keeker (1997))
3. Use of media (to make content attractive)(Keeker (1997))
4. Appropriate content (depth and breadth) (Keeker (1997))
5. Timely / current information (Keeker (1997))

PART II

MATHEMATICAL EXPLANATION AND JUSTIFICATION

CHAPTER 4

ANALYTIC HIERARCHY PROCESS METHODOLOGY

The AHP methodology to solve multicriteria decision problems requires the completion of the following steps (Johnson (1980), and Rapcsák (2007)).

1. Establishment of the decision hierarchy
2. Determination of the weights of the criteria
3. Evaluation of the alternatives
4. Ranking of the alternatives

We explain each step in Sections 4.1 to 4.4.

4.1 Decision hierarchy

In this stage the problem is structured according to a hierarchy, placing every entity or criterion in a level according to its influence in the stated problem or goal of the decision model. A hierarchy with three levels is shown as an example in Figure 4.1.

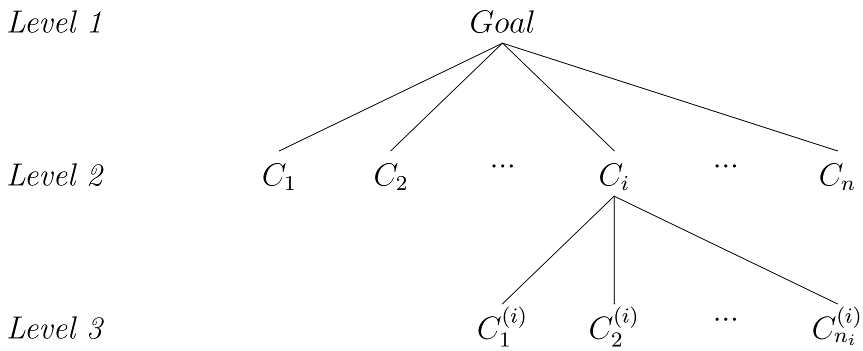


Figure 4.1: Hierarchy structure for AHP

4.2 Weights of the criteria

The second stage consists of determining the relative weights of the criteria that are at the same level in the hierarchy according to their influence on a higher level of the hierarchy. In this situation we could either assign the weights of the criteria directly, if we have reliable information, or we build pairwise comparison matrices for each level of the hierarchy and we find the eigenvector corresponding to the largest eigenvalue which will provide the weights of the criteria for each major criterion for every level in the hierarchy. In our example we have

Level 1			*Goal*				
Level 2	C_1	C_2	$\cdots$	C_i	$\cdots$	C_n	
	w_1	w_2		w_i		w_n	
Level 3			$C_1^{(i)}$	$\cdots$	$C_{n_i}^{(i)}$		
			$w_1^{(i)}$		$w_{n_i}^{(i)}$		

Specifically, the eigenvector is calculated by taking the limit of the normalised row sums of the pairwise comparison matrices generated. Let us analyse deeply the mathematical background behind this procedure.

4.2.1 Theoretical pairwise comparison matrices

In this section the characteristics of pairwise comparison matrices in the *theoretical* case are explained. In general, a pairwise comparison matrix reflects the preference of the decision maker when comparing two objects with respect to an evaluation criterion. In the theoretical case the weights of the criteria are known beforehand.

Definition (Set of evaluation criteria) Given C_i as the ith criterion for $i = 1 \ldots n$, the set of evaluation criteria is defined as

$$C = \Big\{C_1, \ldots, C_n\Big\} \tag{4.1}$$

Definition (Weights of the criteria) Let $C_i \in C$ be a given criterion. Then the value w_i represents the weight of the criterion according to the importance of the criterion with respect to the others criteria in C. We require $w_i > 0$ and $\sum_{i=1}^{n} w_i = 1$.

Now we introduce the formal definition of theoretical pairwise comparison matrices.

Definition (Theoretical pairwise comparison matrix) Let $C_i \in C$ for $i = 1, \ldots, n$, be given evaluation criteria and let w_i be their corresponding weights. Then the entries a_{ij}, of the pairwise comparison matrix $\mathbf{A}_{n\times n}$ are calculated by taking any two of the given criteria and assigning a value according to the importance of the ith criterion with respect to the jth criterion, for $i, j = 1, \ldots, n$. More specifically the entries of the matrix are

$$a_{ij} = \frac{w_i}{w_j} \qquad i, j = 1, \ldots, n \tag{4.2}$$

and the complete pairwise comparison matrix $\mathbf{A}_{n\times n}$ can be represented as

	C_1	C_2	$\cdots$	C_n
C_1	w_1/w_1	w_1/w_2	$\cdots$	w_1/w_n
C_2	w_2/w_1	w_2/w_2	$\cdots$	w_2/w_n
$\vdots$	$\vdots$	$\vdots$		$\vdots$
C_n	w_n/w_1	w_n/w_2	$\cdots$	w_n/w_n

Remark The entries a_{ij} form a set of positive numbers, i.e. $a_{ij} > 0$, for all $i, j = 1, \ldots, n$.

Let us recall the definition of non-negative and positive matrices.

Definition Let $\mathbf{A} = (a_{ij})$ be an $m \times n$ matrix. Then $\mathbf{A}$ is said to be *non-negative* (respectively *positive*) if and only if $a_{ij} \geqslant 0$ (respectively $a_{ij} > 0$), for all $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$.

Remark Theoretical pairwise comparison matrices are positive.

Observe that the entries a_{ij} and a_{ji} have a particular relation. That is, having calculated the value of one of those two, the other one as a consequence is equal to the inverse of the first, that is

$$\begin{aligned} a_{ij} &= \frac{w_i}{w_j} \\ &= \frac{1}{w_j/w_i} \\ &= \frac{1}{a_{ji}} \end{aligned}$$

Definition (Reciprocal matrix) A positive matrix $\mathbf{A}_{n\times n}$ is *reciprocal* if and only if

$$a_{ij} = \frac{1}{a_{ji}} \qquad i, j = 1, \dots, n$$

Remark Theoretical pairwise comparison matrices are reciprocal.

In the theoretical case, the comparison between the criteria preserves a consistency of relation, that is, for every i, j, k, it holds that

$$\begin{aligned} a_{ij}a_{jk} &= \frac{w_i}{w_j} \cdot \frac{w_j}{w_k} \\ &= \frac{w_i}{w_k} \\ &= a_{ik} \end{aligned}$$

This property is known as consistency.

Definition A positive matrix $\mathbf{A}_{n \times n}$ is *consistent* if and only if

$$a_{ik} = a_{ij} \cdot a_{jk} \qquad \text{for all } i, j, k = 1, \ldots, n \tag{4.3}$$

From equation (4.2) we have[1]

$$a_{ij} \cdot \frac{w_j}{w_i} = 1 \qquad i, j = 1, \ldots, n$$

and therefore

$$\sum_{j=1}^{n} a_{ij} w_j \frac{1}{w_i} = n \qquad i = 1, \ldots, n$$

or equivalently

$$\sum_{j=1}^{n} a_{ij} w_j = n w_i \qquad i = 1, \ldots, n \tag{4.4}$$

Remark Theoretical pairwise comparison matrices are consistent.

[1] Saaty (1990).

Remark A system of linear equations

$$
\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\
\vdots & & & & & & \vdots & & \vdots \\
a_{n1}x_1 & + & a_{n2}x_2 & + & \dots & + & a_{nn}x_n & = & b_n
\end{array}
$$

or

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \qquad i = 1, \dots, n$$

can be represented by the matrix equation

$$\mathbf{Ax} = \mathbf{b}$$

Thus, equation (4.4) may be represented by

$$\mathbf{Aw} = n\mathbf{w} \tag{4.5}$$

where $\mathbf{w} = (w_1, w_2, \dots, w_n)^{\mathbf{T}}$. An eigenvector, also called a characteristic vector, of $\mathbf{A}$ is a nonzero vector $\mathbf{x}$ which satisfies

$$\mathbf{Ax} = \lambda\,\mathbf{x}, \qquad \mathbf{x} \neq 0 \tag{4.6}$$

where λ is a scalar, known as an eigenvalue. In order to find λ and $\mathbf{x}$ which satisfy this relation, equation (4.6) may be rewritten as

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = 0, \qquad \mathbf{x} \neq 0$$

where $\mathbf{I}$ denotes the $n \times n$ identity matrix, and since we require $\mathbf{x} \neq 0$, then λ will be a solution of the relation (4.6) provided that the matrix $(\lambda\mathbf{I} - \mathbf{A})$ is singular, in other words if

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0 \tag{4.7}$$

Equation (4.5) gives the important result that in the consistent case n is an eigenvalue of $\mathbf{A}$ and $\mathbf{w}$ is its corresponding eigenvector. Now let us analyse deeply why this is the case.

Definition (Basis) Given a set of linearly independent vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ they will form a basis for a given vector space if by taking linear combinations of those vectors we can get every vector in the given vector space.

Definition (Range and null space of vector spaces) Let $\mathbf{A}$ be an $m \times n$ matrix representing a linear transformation from $\mathbb{F}^n$ to $\mathbb{F}^m$, then

(i) The *range* of $\mathbf{A}$ is defined as $\{\mathbf{y} \in \mathbb{F}^m \mid \mathbf{A}\mathbf{x} = \mathbf{y}, \text{ for some } \mathbf{x} \in \mathbb{F}^n\}$. The range is a subspace of $\mathbb{F}^m$ (for a proof see Andrilli and Hecker (2003) page 259). The dimension (number of linearly independent vectors needed to generate the subspace) of the range is known as the *rank* of the matrix.

(ii) The *null space* of $\mathbf{A}$ is defined as $\{\mathbf{x} \in \mathbb{F}^n \mid \mathbf{A}\mathbf{x} = \mathbf{0}\}$. The null space is a subspace of $\mathbb{F}^n$ (for a proof see Andrilli and Hecker (2003) page 259). The dimension of the null space is known as *nullity* of the matrix.

Let us consider the well known rank-nullity theorem.

Theorem 4.2.1. *(Rank-nullity) Let* $\mathbf{A}$ *be an* $m \times n$ *matrix. Then the rank and the nullity of the matrix add up to the number of columns of the matrix, i.e.*

$$\text{rank}(\mathbf{A}) + \text{nullity}(\mathbf{A}) = n \tag{4.8}$$

Proof. Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be the column vectors of the matrix $\mathbf{A}$. Such columns vectors may form a basis for the range if they are linearly independent, which may or may be not the case. Let $k \leqslant n$ be the maximum number of linearly independent columns of $\mathbf{A}$. We claim that those k vectors, without loss of generality denoted by $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$, form a basis for the range of $\mathbf{A}$ and $\text{rank}(\mathbf{A}) = k$. To show that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is a basis for the range we need to proof that every element in the range may be represented as a linear combination of those k linearly independent vectors, i.e. those k vectors span the range. The vector space $\mathbb{R}^n$ is equipped with the *standard basis* $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, in which, for each $i = 1, \ldots, n$, $\mathbf{e}_i$ is a vector with 1 in the ith component and zeros in all the other entries, then $\textit{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\} = \mathbb{R}^n$ and

$$\mathbf{A}\mathbf{e}_i = \mathbf{v}_i \qquad i = 1, \ldots, n$$

and since matrix multiplications preserve linear combinations then

$$\begin{aligned} \text{range}(\mathbf{A}) &= \textit{span}\{\mathbf{A}\mathbf{e}_1, \mathbf{A}\mathbf{e}_2, \ldots, \mathbf{A}\mathbf{e}_n\} \\ &= \textit{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} \end{aligned} \tag{4.9}$$

Therefore the range of $\mathbf{A}$ is the set of linear combinations of the columns in $\mathbf{A}$, Observe that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ forms a maximal set of linearly independent vectors of $\mathbf{A}$ then

$\{\mathbf{v}_{k+1}, \mathbf{v}_{k+2}, \ldots, \mathbf{v}_n\}$ depend on $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ and so for any column vector $k < j \leqslant n$ there exist coefficients α_{ji} such that

$$\mathbf{v}_j = \sum_{i=1}^{k} \alpha_{ji} \mathbf{v}_i \qquad k < j \leqslant n$$

Then $\mathbf{v}_j \in \mathit{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ for all $j > k$, and so

$$\mathit{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} = \mathit{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} \tag{4.10}$$

therefore from equations (4.9) and (4.10) we conclude

$$\text{range}(\mathbf{A}) = \mathit{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$$

Thus $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ forms a basis for the range and so $\text{rank}(\mathbf{A}) = k$. Now we need to show that there exists a set $\{\mathbf{w}_{k+1}, \mathbf{w}_{k+2}, \ldots, \mathbf{w}_n\}$ that forms a basis for the null space. Let

$$\mathbf{w}_j = \mathbf{e}_j - \sum_{i=1}^{k} \alpha_{ji} \mathbf{e}_i \qquad k < j \leqslant n$$

Observe that defined in this way each $\mathbf{w}_j$ contains one $\mathbf{e}_j$ and no other $\mathbf{e}_l$ for $j, l > k$. Consequently the set of column vectors $\{\mathbf{w}_{k+1}, \mathbf{w}_{k+2}, \ldots, \mathbf{w}_n\}$ is linearly independent and indeed $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k, \mathbf{w}_{k+1}, \mathbf{w}_{k+2}, \ldots, \mathbf{w}_n\}$. is linearly independent and so forms a basis for $\mathbb{R}^n$. Now we have

$$\begin{aligned}
\mathbf{A}(\mathbf{w}_j) &= \mathbf{A}\mathbf{e}_j - \sum_{i=1}^{k} \alpha_{ji}\mathbf{A}\mathbf{e}_i && k < j \leqslant n \\
&= \mathbf{v}_j - \sum_{i=1}^{k} \alpha_{ji}\mathbf{v}_i && k < j \leqslant n \\
&= \mathbf{v}_j - \mathbf{v}_j && k < j \leqslant n \\
&= \mathbf{0}
\end{aligned}$$

and so $\mathbf{A}\mathbf{w}_j = \mathbf{0}$ for $j = k+1, \ldots, n$. Thus

$$\text{nullity}(\mathbf{A}) \geqslant n - k \tag{4.11}$$

Now we need to show that $\{\mathbf{w}_{k+1}, \mathbf{w}_{k+2}, \ldots, \mathbf{w}_n\}$ spans null space of $\mathbf{A}$. We know that $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k, \mathbf{w}_{k+1}, \mathbf{w}_{k+2}, \ldots, \mathbf{w}_n\}$ is a basis for $\mathbb{F}^n$. Suppose $\mathbf{x}$ is in the null space of $\mathbf{A}$ and suppose

$$\mathbf{x} = \sum_{i=1}^{k} \beta_i \mathbf{e}_i + \sum_{j=k+1}^{n} \beta_j \mathbf{w}_j$$

Observe that since $\sum_{j=k+1}^{n} \beta_j \mathbf{w}_j$ is in the null space, also $\sum_{i=1}^{k} \beta_i \mathbf{e}_i$ is in the null space, i.e.

$$\begin{aligned}
\mathbf{A}\mathbf{x} = \mathbf{A}\left(\sum_{i=1}^{k} \beta_i \mathbf{e}_i + \sum_{j=k+1}^{n} \beta_j \mathbf{w}_j\right) &= \mathbf{0} \\
\mathbf{A}\left(\sum_{i=1}^{k} \beta_i \mathbf{e}_i\right) &= \mathbf{0} - \mathbf{A}\left(\sum_{j=k+1}^{n} \beta_j \mathbf{w}_j\right) \\
\sum_{i=1}^{k} \beta_i \mathbf{v}_i &= \mathbf{0} - \mathbf{0} \\
&= \mathbf{0}
\end{aligned} \tag{4.12}$$

But $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is linearly independent, so $\beta_i = 0$ for $i = 1, \ldots, k$ and

$$\mathbf{x} = \sum_{j=k+1}^{n} \beta_j \mathbf{w}_j$$

Then $\{\mathbf{w}_{k+1}, \mathbf{w}_{k+2}, \ldots, \mathbf{w}_n\}$ is a basis for the null space of $\mathbf{A}$ therefore

$$\text{null}(\mathbf{A}) \leqslant n - k \tag{4.13}$$

Equations (4.11) and (4.13) indicate that $\text{null}(\mathbf{A}) = n - \text{rank}(\mathbf{A})$ which completes the proof.

$\square$

Now let us introduce a theorem concerning the rank and the null space of a theoretical pairwise comparison matrix.

Theorem 4.2.2. *Let* $\mathbf{A}$ *be an* $n \times n$ *theoretical pairwise comparison matrix. Then the following holds.*

(i) The $\text{rank}(\mathbf{A})$ *is equal to 1.*

(ii) The $\text{nullity}(\mathbf{A})$ *is equal to* $n - 1$*.*

Proof. The real $n \times n$ matrix $\mathbf{A}$ represents a linear transformation $f : \mathbb{R}^n \to \mathbb{R}^n$ such that $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$. The linear transformation will depend on the basis chosen for $\mathbb{R}^n$. Recall that *span*$\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\} = \mathbb{R}^n$. Given $\mathbf{x}$, the values of $f(\mathbf{x})$ depend entirely on the values of the entries a_{ij}. Specifically, in our case we have

$$\begin{pmatrix} \frac{w_1}{w_1} & \frac{w_1}{w_2} & \cdots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & \frac{w_2}{w_2} & \cdots & \frac{w_2}{w_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \cdots & \frac{w_n}{w_n} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{w_1}{w_1} \\ \frac{w_2}{w_1} \\ \vdots \\ \frac{w_n}{w_1} \end{pmatrix} = \frac{1}{w_1} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \frac{1}{w_1} \mathbf{w}$$

$$\begin{pmatrix} \frac{w_1}{w_1} & \frac{w_1}{w_2} & \cdots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & \frac{w_2}{w_2} & \cdots & \frac{w_2}{w_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \cdots & \frac{w_n}{w_n} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{w_1}{w_2} \\ \frac{w_2}{w_2} \\ \vdots \\ \frac{w_n}{w_2} \end{pmatrix} = \frac{1}{w_2} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \frac{1}{w_2} \mathbf{w}$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$\begin{pmatrix} \frac{w_1}{w_1} & \frac{w_1}{w_2} & \cdots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & \frac{w_2}{w_2} & \cdots & \frac{w_2}{w_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \cdots & \frac{w_n}{w_n} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{w_1}{w_n} \\ \frac{w_2}{w_n} \\ \vdots \\ \frac{w_n}{w_n} \end{pmatrix} = \frac{1}{w_n} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \frac{1}{w_n} \mathbf{w}$$

Thus,

$$\mathbf{A}\,\mathbf{e}_i = \frac{1}{w_i}\,\mathbf{w} \qquad i = 1, \ldots, n \tag{4.14}$$

Now let us consider a vector $\mathbf{v} = (v_1, v_2, \ldots, v_n)^{\mathbf{T}}$. Such a vector $\mathbf{v}$ may be written as

$$\mathbf{v} = v_1\,\mathbf{e}_1 + v_2\,\mathbf{e}_2 + \ldots + v_n\,\mathbf{e}_n \tag{4.15}$$

where $v_i \in \mathbb{R}$ for $i = 1, \ldots n$. Using the fact that matrix multiplication preserves the operations of vector addition and scalar multiplication, with equations (4.14) and (4.15) we get:

$$\begin{aligned}
\mathbf{A}\,\mathbf{v} &= \mathbf{A}(v_1\,\mathbf{e}_1) + \mathbf{A}(v_2\,\mathbf{e}_2) + \ldots + \mathbf{A}(v_n\,\mathbf{e}_n) \\
&= v_1\,\mathbf{A}\,\mathbf{e}_1 + v_2\,\mathbf{A}\,\mathbf{e}_2 + \ldots + v_n\,\mathbf{A}\,\mathbf{e}_n \\
&= \frac{v_1}{w_1}\,\mathbf{w} + \frac{v_2}{w_2}\,\mathbf{w} + \ldots + \frac{v_n}{w_n}\,\mathbf{w} \\
&= c\,\mathbf{w} \qquad (4.16)
\end{aligned}$$

where c is a scalar and $\mathbf{w} = (w_1, w_2, \ldots, w_n)^{\mathbf{T}}$. Observe that in (4.16) the transformation of the vector $\mathbf{v}$ has been given in terms of the vector $\mathbf{w}$, regardless of the initial choice of $\mathbf{v}$, which indicates that the rank($\mathbf{A}$) is 1. Using Theorem 4.2.1, with the number of columns of $\mathbf{A}$ equal to n, the dimension of the null space of $\mathbf{A}$ can be obtained from the relation

$$\begin{aligned}
\text{nullity}(\mathbf{A}) &= n - \text{rank}(\mathbf{A}) \\
\text{nullity}(\mathbf{A}) &= n - 1
\end{aligned}$$

□

Recall equation (4.7)

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

The solution of (4.7) gives a polynomial in λ which is known as the characteristic polynomial. Consequently, the roots of the polynomial (4.7) are the eigenvalues of $\mathbf{A}$, denoted as $\lambda_1, \ldots, \lambda_n$. Now let us consider the definitions of elementary symmetric polynomials, principal submatrices and principal minors.

Definition (Elementary symmetric polynomials) Given $\lambda_1, \ldots, \lambda_n$, the elementary symmetric polynomials are defined as

$$e_k(\lambda_1, \ldots, \lambda_n) = \sum_{1 \leqslant i_1 < \ldots < i_k \leqslant n} \prod \lambda_{i_j} \qquad k = 0, 1, \ldots, n. \tag{4.17}$$

or

$$\begin{aligned} e_0(\lambda_1, \ldots, \lambda_n) &= 1 \\ e_1(\lambda_1, \ldots, \lambda_n) &= \lambda_1 + \ldots + \lambda_n \\ e_2(\lambda_1, \ldots, \lambda_n) &= \lambda_1\lambda_2 + \lambda_1\lambda_3 + \ldots + \lambda_{n-2}\lambda_n + \lambda_{n-1}\lambda_n \\ e_3(\lambda_1, \ldots, \lambda_n) &= \lambda_1\lambda_2\lambda_3 + \lambda_1\lambda_2\lambda_4 + \ldots + \lambda_{n-3}\lambda_{n-1}\lambda_n + \lambda_{n-2}\lambda_{n-1}\lambda_n \\ \vdots \qquad & \vdots \qquad \vdots \\ e_n(\lambda_1, \ldots, \lambda_n) &= \lambda_1 \ldots \lambda_n \end{aligned} \tag{4.18}$$

Definition (Principal submatrix) Let $\mathbf{A} = (a_{ij})$ be an $n \times n$ matrix. Then a principal submatrix is any $k \times k$ matrix, $1 \leqslant k \leqslant n$, obtained by removing $n-k$ rows and columns of $\mathbf{A}$, such that if the ith row is removed then the ith column is removed too.

Definition (Principal minor) Given a principal submatrix, the corresponding principal minor is the determinant of this principal submatrix.

Observe that from the definition of a principal submatrix it follows that there are $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ principal submatrices.

Definition (Sum of principal minors) Let $E_k(\mathbf{A})$ denote the sum of the $k \times k$ principal minors of $\mathbf{A}$, i.e.

$$E_k(\mathbf{A}) = \sum_{i_1<\ldots<i_k} \Delta(i_1, \ldots, i_k) \qquad k = 1, \ldots, n \tag{4.19}$$

Where $\Delta(i_1, \ldots, i_k)$ is the determinant of the matrix composed by the intersection of rows $i_1, \ldots, i_k$ and columns $i_1, \ldots, i_k$.

Observe that in (4.19) if $k = 1$ then we have n, 1×1 principal submatrices corresponding to each element in the diagonal of $\mathbf{A}$, whose determinant is the element itself, and thus $E_1(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$.

Definition (Trace of a matrix) Given a matrix $\mathbf{A}$, the trace, denoted by $\mathrm{tr}(\mathbf{A})$, is given by

$$\mathrm{tr}(\mathbf{A}) = a_{11} + a_{22} + \ldots + a_{nn} = \sum_{i=1}^{n} a_{ii} \tag{4.20}$$

Remark $E_1(\mathbf{A}) = \mathrm{tr}(\mathbf{A})$.

In the same way, if $k = n$ then we have 1, $n \times n$ matrix which is $\mathbf{A}$ and thus $E_n(\mathbf{A}) = \det(\mathbf{A})$.

Now let us analyse the degree of the characteristic polynomial of a given matrix.

Lemma 4.2.3. *Let* $\mathbf{A} = (a_{ij})$ *be an* $n \times n$ *matrix. Let the entries* a_{ij} *be either constant or linear in* λ. *Then, for all* $m \leqslant n$, *if the matrix has* m *entries with nonzero* λ *coefficients and no two of these entries lie in the same row or same column. Then the characteristic polynomial* $\det(\lambda\mathbf{I} - \mathbf{A})$ *has degree* m, *i.e.*

$$\det(\lambda\mathbf{I} - \mathbf{A}) = \lambda^m + c_1\lambda^{m-1} + c_2\lambda^{m-2} + \ldots + c_n \tag{4.21}$$

Proof. We will show this by induction on m. As an induction hypothesis suppose we have a $n \times n$ matrix in which exactly m entries have nonzero λ coefficients and no two of these entries lie in the same row or same column. We say that such entries are *independent*. Then we claim that the determinant of such matrix is a polynomial in λ of degree m. If $m = 0$ then the result is trivial. Let us assume that the hypothesis holds for some $k = m-1$ and so we need to check that it works for $k = m$. Assuming that $1 \leqslant m \leqslant n-1$ and using Laplace expansion[2] with a row i containing a λ then every submatrix resulting by deleting the row i and a column j of the original matrix has at most $l \leqslant m-1$ entries with λ coefficients. Thus the determinant of every resulting submatrix in the expansion of row i is a polynomial of degree less than or equal to $m-1$.

However if we consider an entry with a λ, since such entries are independent, the submatrix has $l = m-1$ λ entries, and so the determinant of such submatrix is a polynomial of degree $m-1$. In this case when multiplied by the entry in the ith row that includes a λ the polynomial will have a degree equal to m, i.e. Using Laplace expansion

[2]Given an $n \times n$ matrix $\mathbf{A}$, the determinant can be calculated expanding by cofactors on row i or column j, i.e.

$$\begin{aligned}\det\mathbf{A} &= a_{i1}\cdot\mathbf{B}_{i1} + a_{i2}\cdot\mathbf{B}_{i2} + \cdots + a_{in}\cdot\mathbf{B}_{in}\\ &= a_{1j}\cdot\mathbf{B}_{1j} + a_{2j}\cdot\mathbf{B}_{2j} + \cdots + a_{nj}\cdot\mathbf{B}_{nj}\end{aligned}$$

With $\mathbf{B}_{ij} = (-1)^{i+j}|\mathbf{M}_{ij}|$ and $\mathbf{M}_{ij}$ is the matrix obtained by deleting the ith row and jth column of $\mathbf{A}$.

by the ith row containing one independent λ we get

$$\begin{aligned}\det(\lambda\mathbf{I}-\mathbf{A}) &= p_1^{(m)}(\lambda)+p_2^{(l_2)}(\lambda)+\ldots+p_n^{(l_n)}(\lambda)\\ &= p^{(m)}(\lambda)\end{aligned}$$

where $l \leqslant m-1$ and $p^{(s)}(\lambda)$ indicates a polynomial in λ of degree s. Therefore given a matrix with m independent λ coefficients then the determinant of such matrix is polynomial in λ of degree m.

□

Lemma 4.2.4. *Let* $\mathbf{A}$ *be an* $n\times n$ *matrix. Then the characteristic polynomial* $\det(\lambda\mathbf{I}-\mathbf{A})$ *has degree* n*, i.e.*

$$\det(\lambda\mathbf{I}-\mathbf{A}) = \lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + \ldots + c_n \tag{4.22}$$

Proof. The $\det(\lambda\mathbf{I}-\mathbf{A})$ has the form

$$\begin{vmatrix} -a_{11}+\lambda & -a_{12} & \ldots & -a_{1n}\\ -a_{21} & -a_{22}+\lambda & \ldots & -a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ -a_{n1} & -a_{n2} & \ldots & -a_{nn}+\lambda \end{vmatrix}$$

Observe that the matrix has n independent λ entries. By Lemma 4.2.3 with $m=n$ the characteristic polynomial in λ has degree n as claimed.

□

Now we are able to introduce a relation between the elementary symmetric polynomials $e_k(\lambda_1,\ldots,\lambda_n)$ and the sum of principal minors $E_k(\mathbf{A})$ $\qquad k=1,\ldots,n$.

Theorem 4.2.5. *Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of an $n \times n$ matrix $\mathbf{A}$, with multiplicity. Then*

$$e_k(\lambda_1, \ldots, \lambda_n) = E_k(\mathbf{A}) \qquad k = 1, \ldots, n \tag{4.23}$$

Proof. Considering the roots of the polynomial in (4.22) we have the expression

$$(\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_n) \tag{4.24}$$

Specifically, expanding (4.24) we get

$$\begin{aligned}(\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_n) = \lambda^n& \\ &- (\lambda_1 + \ldots + \lambda_n)\lambda^{n-1} \\ &+ (\lambda_1\lambda_2 + \lambda_1\lambda_3 + \ldots + \lambda_{n-2}\lambda_n + \lambda_{n-1}\lambda_n)\lambda^{n-2} \\ &- \ldots \\ &+ (-1)^n(\lambda_1 \ldots \lambda_n)\end{aligned} \tag{4.25}$$

Observe that in (4.25) every coefficient of λ^{n-k} corresponds to the equivalent $e_k(\lambda_1, \ldots, \lambda_n)$ for $k = 1, \ldots, n$ defined in (4.17). Therefore (4.25) can be written as

$$\begin{aligned}(\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_n) = \lambda^n& \\ &- e_1(\lambda_1, \ldots, \lambda_n)\lambda^{n-1} \\ &+ e_2(\lambda_1, \ldots, \lambda_n)\lambda^{n-2} \\ &- \cdots \\ &+ (-1)^n e_n(\lambda_1, \ldots, \lambda_n)\end{aligned} \tag{4.26}$$

On the other hand, by Lemma 4.2.4

$$\det(\lambda\mathbf{I} - \mathbf{A}) = \lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + \ldots + c_n$$

We have seen that

$$\det(\lambda\mathbf{I} - \mathbf{A}) = \begin{vmatrix} -a_{11} + \lambda & -a_{12} & \ldots & -a_{1n} \\ -a_{21} & -a_{22} + \lambda & \ldots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \ldots & -a_{nn} + \lambda \end{vmatrix}$$

Without loss of generality let us identify the λ in the ith row with λ_i, for $i = 1, \ldots, n$, such that

$$\begin{vmatrix} -a_{11} + \lambda_1 & -a_{12} & \ldots & -a_{1n} \\ -a_{21} & -a_{22} + \lambda_2 & \ldots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \ldots & -a_{nn} + \lambda_n \end{vmatrix}$$

We claim that the term in $\lambda_{i_1} \ldots \lambda_{i_k}$ with $k \leqslant n$ is equal to

$$\det\{\ldots\{\{\{-\mathbf{A}\}_{i_1}\}_{i_2}\}\ldots\}_{i_k} \tag{4.27}$$

where $\{\mathbf{B}\}_l$ indicates the resultant matrix after deleting the row l and the column l from $\mathbf{B}$. Let (4.27) be our induction hypothesis on k. For $k = 0$ the hypothesis holds. Assume that the hypothesis is true for $k - 1$ then we need to show that it also holds for k. Expanding around row i_k the term in $\lambda_{i_1} \ldots \lambda_{i_k}$ in $\lambda\mathbf{I} - \mathbf{A}$ is equal to the term in $\lambda_{i_1} \ldots \lambda_{i_{k-1}}$ in $\{\lambda\mathbf{I} - \mathbf{A}\}_{i_k}$. By the induction hypothesis this is equal to $\det\{\ldots\{\{\{-\mathbf{A}\}_{i_k}\}_{i_1}\}\ldots\}_{i_{k-1}}$

which is $\det\{\ldots\{\{\{-\mathbf{A}\}_{i_1}\}_{i_2}\}\ldots\}_{i_k}$. Then (4.27) holds. Let $d(i_1,\ldots,i_k)$ be the term with $\lambda_{i_1}\ldots\lambda_{i_k}$. Now to know the coefficient of a term with λ^k we need to sum up all the terms $\lambda_{i_1}\ldots\lambda_{i_k}$. Thus, using (4.27) we have

$$\sum_{i_1<\ldots<i_k} d(i_1,\ldots,i_k) = \sum_{i_1,\ldots,i_k} \det\{\ldots\{\{\{-\mathbf{A}\}_{i_1}\}_{i_2}\}\ldots\}_{i_k}$$

Let $d(i_1<\ldots<i_k) = d(S)$ where $S = \{i_1,\ldots,i_k\}$, and similarily let $\Delta(i_1,\ldots,i_k) = \Delta(S)$ Then $d(S) = \Delta\left([n]\backslash S\right)$. Let us define d_k as the the term of λ^k given by

$$d_k = \sum_{i_1<\ldots<i_k} d(i_1,\ldots,i_k)$$

then

$$d_k = \sum_{j_1<\ldots<j_{n-k}} \Delta(j_1,\ldots,j_{n-k}) \tag{4.28}$$

Thus since the term in λ^k is given by equation (4.28) therefore the term in λ^{n-k} is given by

$$\sum_{i_1<\ldots<i_k} \Delta(i_1,\ldots,i_k)$$

And so

$$\begin{aligned}\det(\lambda\mathbf{I}-\mathbf{A}) = \lambda^n &\\ &- \left(\sum_{j=1}^{n} \Delta(i_j)\right)\lambda^{n-1}\\ &+ \left(\sum_{\substack{i_j<i_k\\ j,k=1}}^{n} \Delta(i_j,i_k)\right)\lambda^{n-2}\\ &- \left(\sum_{\substack{i_j<i_k<i_l\\ j,k,l=1}}^{n} \Delta(i_j,i_k,i_l)\right)\lambda^{n-3}\\ &- \cdots\\ &+ (-1)^n\det(\mathbf{A})\end{aligned} \tag{4.29}$$

Which is equivalent to

$$\begin{aligned}\det(\lambda\mathbf{I}-\mathbf{A}) = \lambda^n &\\ &- E_1(\mathbf{A})\lambda^{n-1}\\ &+ E_2(\mathbf{A})\lambda^{n-2}\\ &- \cdots\\ &+ (-1)^n E_n(\mathbf{A})\end{aligned} \tag{4.30}$$

Consequently from (4.26) and (4.30) it follows that

$$e_k(\lambda_1,\ldots,\lambda_n) = E_k(\mathbf{A}) \qquad k=1,\ldots,n$$

□

In particular the next two corollaries follow from Theorem 4.2.5.

Corollary 4.2.6. *Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of an $n \times n$ matrix* $\mathbf{A}$ *with multiplicity, then*

$$\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_{ii} \tag{4.31}$$

Proof. Using Theorem 4.2.5 we know that $e_k(\lambda_1, \dots, \lambda_n) = E_k(\mathbf{A})$ for $k = 1, \dots, n$. This corollary corresponds to the particular case when $k = 1$.

□

Corollary 4.2.7. *Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of an $n \times n$ matrix* $\mathbf{A}$*, and suppose $a_{ii} = 1$, for all $i = 1, \dots, n$. Then*

$$\sum_{i=1}^{n} \lambda_i = n \tag{4.32}$$

Proof. From Corollary 4.2.6 we know that $\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_{ii}$ and with $a_{ii} = 1$ for all $i = 1, \dots, n$ we get $\sum_{i=1}^{n} a_{ii} = n$, therefore $\sum_{i=1}^{n} \lambda_i = n$.

□

Now we will see that the rank of $\mathbf{A}$ is at least the number of nonzero eigenvalues of A as explained in Lemma 4.2.8.

Definition (Number of nonzero eigenvalues) Let $e(\mathbf{A})$ be the number of nonzero eigenvalues of $\mathbf{A}$ with multiplicity.

Lemma 4.2.8. *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix. Then it holds that the number of nonzero eigenvalues, with multiplicity, is less than or equal to the rank of the matrix, i.e.*

$$e(\mathbf{A}) \leqslant \text{rank}(\mathbf{A}) \tag{4.33}$$

Proof. Let *nullity*$(\mathbf{A}) = k$. Then there exist k vectors in the basis for the null space of $\mathbf{A}$, and so there exist k zero eigenvalues. Also there are at most n eigenvalues in total, and so

$$e(\mathbf{A}) \leqslant n - k$$

However by Theorem 4.2.1 (rank-nullity) we know that

$$\text{rank}(\mathbf{A}) = n - k$$

Then

$$e(\mathbf{A}) \leqslant \text{rank}(\mathbf{A})$$

□

Lemma 4.2.9. *Let* $\mathbf{A}$ *be an* $n \times n$ *positive reciprocal consistent matrix. Then* $\mathbf{A}$ *is a theoretical pairwise comparison matrix.*

Proof. Consider the first column of a positive, consistent, reciprocal matrix, i.e.

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} \tag{4.34}$$

Note that since $\mathbf{A}$ is reciprocal then $a_{11} = 1$. Set $w_i = c \cdot a_{i1}$, where c is a chosen such that $\sum_{i=1}^{n} w_i = 1$, i.e.

$$c = \frac{1}{\sum_{i=1}^{n} a_{i1}}$$

Note that $c = w_1$ and therefore $a_{i1} = \frac{w_i}{c} = \frac{w_i}{w_1}$. Since $\mathbf{A}$ is reciprocal, also

$$a_{1i} = \frac{w_1}{w_i}$$

And since $\mathbf{A}$ is consistent,

$$\begin{aligned} a_{ij} &= a_{i1} \cdot a_{1j} && i, j = 1, \ldots, n \\ &= \frac{w_i}{w_1} \cdot \frac{w_1}{w_j} && i, j = 1, \ldots, n \\ &= \frac{w_i}{w_j} && i, j = 1, \ldots, n \end{aligned}$$

Therefore every positive reciprocal consistent matrix is a theoretical pairwise comparison matrix.

□

Now we are able to introduce the desired result of this section.

Theorem 4.2.10. *Let* $\mathbf{A} = (a_{ij})$ *be an* $n \times n$ *positive reciprocal consistent matrix, defined by weights* w_i*. Then* n *is the largest eigenvalue of* $\mathbf{A}$*, and* $\mathbf{w} = (w_1, w_2, \dots, w_n)^{\mathbf{T}}$ *its corresponding eigenvector, and all the other eigenvalues* λ_i *are equal to zero.*

Proof. Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of $\mathbf{A}$. Since $\mathbf{A}$ is positive reciprocal consistent matrix then by Lemma 4.2.9 $\mathbf{A}$ is a theoretical pairwise comparison matrix. Using Theorem 4.2.2 we know that the rank$(\mathbf{A}) = 1$. Thus from Lemma 4.2.8 we obtain the relation

$$e(\mathbf{A}) \leqslant 1$$

Consequently the number of nonzero eigenvalues of $\mathbf{A}$ is less than or equal to 1. However we have seen that n and $\mathbf{w}$ are a scalar and a nonzero vector respectively, satisfying the equation $\mathbf{Aw} = n\mathbf{w}$, and so n is an eigenvalue of $\mathbf{A}$ and $\mathbf{w}$ its corresponding eigenvector. Then combining both results we know that the only possible eigenvalues of $\mathbf{A}$ are now n and 0. Nevertheless from Corollary 4.2.7 we know that

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

Thus, because the sum of the eigenvalues is n and also n is an eigenvalue then we conclude that n is the largest eigenvalue of $\mathbf{A}$ with multiplicity 1, $\mathbf{w}$ its corresponding eigenvector and all the other eigenvalues λ_i are zero. (Alternatively note that Lemma 4.2.8 actually counts $e(\mathbf{A})$ with multiplicity, so we could not have another n.)

□

Corollary 4.2.11. *Let* $\mathbf{A} = (a_{ij})$ *be an* $n \times n$ *theoretical pairwise comparison matrix, defined by weights* w_i*. Then* n *is the largest eigenvalue of* $\mathbf{A}$*, and* $\mathbf{w} = (w_1, w_2, \ldots, w_n)^{\mathbf{T}}$ *its corresponding eigenvector, and all the other eigenvalues* λ_i *are equal to zero.*

Proof. Since $\mathbf{A}$ is positive, reciprocal and consistent, then by Theorem 4.2.10 n is the largest eigenvalue of $\mathbf{A}$, and $\mathbf{w} = (w_1, w_2, \ldots, w_n)^{\mathbf{T}}$ its corresponding eigenvector, and all the other eigenvalues λ_i are equal to zero.

□

Let us introduce a simple fact.

Lemma 4.2.12. *(The Arithmetic mean - geometric mean* $AM - GM$ *inequality) Let* x *and* y *be nonnegative numbers in* $\mathbb{R}$*. Then*

$$\frac{(x+y)}{2} \geqslant \sqrt{xy} \tag{4.35}$$

Equality holds in (4.35) *if and only if* $x = y$*.*

Proof. [3] If $x \neq y$ then

$$\begin{aligned} (x-y)^2 &> 0 \\ x^2 + 2xy + y^2 &> 4xy \\ \left(\frac{x+y}{2}\right)^2 &> xy \\ \frac{(x+y)}{2} &> \sqrt{xy} \end{aligned} \tag{4.36}$$

[3] This proof was introduced by Cauchy (1821).

If $x = y$ then the arithmetic mean is $(x+y)/2 = x$ and the geometric mean is equal to $\sqrt{xy} = x$.

□

Theorem 4.2.13. *Let* $\mathbf{A} = (a_{ij})$ *be an* $n \times n$ *positive reciprocal matrix, then* $\mathbf{A}$ *is consistent if and only if* $\lambda_{max} = n$.

Proof. If the matrix $\mathbf{A}$ is consistent, then $\lambda_{max} = n$ from Theorem 4.2.10. Suppose now that $\lambda_{max} = n$. Then n is an eigenvalue of $\mathbf{A}$ satisfying the characteristic equation and so

$$\mathbf{A}\mathbf{w} = n\mathbf{w}, \qquad \text{for some } \mathbf{w} \neq \mathbf{0}$$

Which can be rewritten as

$$\sum_{j=1}^{n} a_{ij} w_j = n w_i \qquad i = 1, \ldots, n$$

or

$$\sum_{j=1}^{n} a_{ij} w_j \frac{1}{w_i} = n \qquad i = 1, \ldots, n \tag{4.37}$$

However since $\mathbf{A}$ is a positive reciprocal matrix then

$$\sqrt{\left(a_{ij} \cdot \frac{w_j}{w_i}\right)\left(a_{ji} \cdot \frac{w_i}{w_j}\right)} = 1$$

and so using Lemma 4.2.12

$$a_{ij} \cdot \frac{w_j}{w_i} + a_{ji} \cdot \frac{w_i}{w_j} \geqslant 2 \qquad \forall\, i \neq j$$

and equality holds if and only if $a_{ij} \cdot \frac{w_j}{w_i} = a_{ji} \cdot \frac{w_i}{w_j} = 1$ for all $i \neq j$. Therefore considering that the number of entries a_{ij} with $i < j$ is $(n^2 - n)/2$ then

$$\sum_{i<j}^{n} \left(a_{ij} \cdot \frac{w_j}{w_i} + a_{ji} \cdot \frac{w_i}{w_j} \right) \geqslant 2 \left(\frac{n^2 - n}{2} \right)$$

$$\sum_{i<j}^{n} \left(a_{ij} \cdot \frac{w_j}{w_i} + a_{ji} \cdot \frac{w_i}{w_j} \right) \geqslant n^2 - n \tag{4.38}$$

and inequality in equation (4.38) holds if and only if $a_{ij} \cdot \frac{w_j}{w_i} = a_{ji} \cdot \frac{w_i}{w_j} = 1$ for all $i < j$. Observe that in (4.38) we are not only considering the entries with $i < j$ but also the entries with $i > j$. And so considering also the entries in the main diagonal a_{ii}, which are equal to 1 for all $i = 1, \ldots, n$, we get

$$\sum_{i<j}^{n} \left(a_{ij} \cdot \frac{w_j}{w_i} + a_{ji} \cdot \frac{w_i}{w_j} \right) + \sum_{i=1}^{n} a_{ii} \geqslant n^2 - n + n$$

which is equivalent to

$$\sum_{i,j}^{n} a_{ij} \cdot w_j \frac{1}{w_i} \geqslant n^2 \tag{4.39}$$

and equality in (4.39) holds if and only if $a_{ij} \cdot \frac{w_j}{w_i} = a_{ji} \cdot \frac{w_i}{w_j} = 1$ for all i, j. Also by (4.39)

$$\sum_{i=1}^{n} a_{ij} \frac{w_j}{w_i} = n^2$$

so equality holds therefore the entries

$$a_{ij} \cdot \frac{w_j}{w_i} = a_{ji} \frac{w_i}{w_j} = 1 \qquad i, j = 1, \ldots, n$$

Consequently we can see that all the entries a_{ij} satisfy the relation $a_{ik} = a_{ij} \cdot a_{jk}$ for all $i, j, k = 1, \ldots, n$, i.e. the matrix is *consistent*.

□

4.2.2 Practical pairwise comparison matrices

The properties described in Section 4.2.1 hold in the case of dealing with positive reciprocal consistent matrices. Let us analyse what happens in the practical case, in which the resultant matrices are positive and reciprocal but not always consistent. As in the theoretical case, practical pairwise comparison matrices are generated by the relation, in terms of importance, between some given evaluation criteria. However in the practical case, these evaluation criteria can be either *objective* or *subjective*. And so these preferences have to be given as comparisons because of the absence of the value, or weight, of each evaluation criterion. Therefore the aim of practical pairwise comparison matrices is to find a priority vector which guides the comparison in the elements of the matrix.

Observe that given a set of evaluation criteria, as defined in (4.1), a comparison scale needs to be defined to guide the comparison.

Remark A comparison scale is defined to associate a verbal gradation to a numerical value that belongs to a set of numbers or progression.

In the particular case of pairwise comparison matrices a comparison scale generally associates numerical values according to the degree of preference for one criterion over another.

Definition (Scale comparison in the case of pairwise comparison matrices) Let $l_1 \leqslant l_2 \leqslant \ldots \leqslant l_n$. Then a scale comparison is generally defined as

Intensity of importance	Description
l_1	Equally important
l_2	Weak importance of one over another
$\vdots$	$\vdots$
l_{n-1}	Very strong importance
l_n	Absolute importance

where $l_i > 0$ for $i = 1, \ldots, n$.

Now we introduce the definition of practical pairwise comparison matrices properly.

Definition (Practical pairwise comparison matrix) Let $\mathbf{P} = (p_{ij})$ be an $n \times n$ whose entries p_{ij} represent the subjective judgement given by experts, when considering the criterion $C_i \in C$ and the criterion $C_j \in C$, for $i, j = 1, \ldots, n$. More precisely, the entries of the matrix p_{ij} are equal to

$$p_{ij} = \begin{cases} \omega_{ij} & \text{if } C_i \text{ is more important than } C_j, \\ 1 & \text{if } C_i \text{ is equally important to } C_j, \\ 1/\omega_{ij} & \text{if } C_j \text{ is more important than } C_i. \end{cases} \tag{4.40}$$

where ω_{ij} is a value selected according to a given scale comparison and so $\omega_{ij} > 0$. The matrix $\mathbf{P}$ is called a practical pairwise comparison matrix.

From the definition of practical pairwise matrices, when comparing the C_i with C_j

there are only three possibilities

- C_i is more important than C_j
- C_i is equally important to C_j
- C_j is more important than C_i

Then for those entries p_{ij} in which the criterion C_i is considered to be more important than the criterion C_j, the values are taken from a given scale comparison which expresses the relation of importance, and belong to a set of positive numbers. To make the judgements of the experts reliable we expect that the criterion $C_i \in C$ for all $i = 1, \ldots n$, compared with itself, has the same importance. Then the entries p_{ij} with $i = j$ (i.e. the entries in the main diagonal) are equal to 1. The remaining entries p_{ij} are the reciprocal numbers of the elements already found.

Remark Let $\mathbf{P}_{n \times n}$ be a practical pairwise comparison matrix. Then the matrix is positive (i.e. $p_{ij} > 0$ for all $i, j = 1, \ldots n$).

In the practical case, only $n(n-1)/2$ (n^2 in total minus n elements in the diagonal divided by 2) elements are given by the expert (i.e. only those entries considering C_i be better than C_j or equally important).

Remark Let $\mathbf{P}_{n \times n}$ be a practical pairwise comparison matrix. Then the matrix is reciprocal (i.e. $p_{ij} = 1/p_{ji}$ for all $i, j = 1, \ldots n$).

Since the entries p_{ij} depend on subjective judgements, the comparisons between the criteria in C do not always preserve the consistency of relation presented in (4.3), and so

(4.5) cannot be used, and other concepts need to be introduced.

Starting from the theoretical case, in which matrices satisfy the properties of positiveness, reciprocity, and most importantly consistency, we will see, under different assumptions, how the theoretical background can be used in the practical case. Our first analysis concerns the sensitivity of the eigenvector arising from a perturbation of the data in $\mathbf{A}$. We have seen that the eigenvalues correspond to the roots of the characteristic polynomial. But then since the coefficients of this polynomial are calculated using addition and multiplication of the entries in $\mathbf{A}$ then the zeros of a polynomial are continuous functions of the coefficients of that polynomial. As a consequence the eigenvalues are continuous functions of the entries of a matrix, and so it is plausible that if a perturbation of the matrix $\mathbf{A}$ is small enough then the eigenvalues should stay close to their original values (Horn and Johnson (1999)). Theorem 4.2.14 illustrates this case.

Theorem 4.2.14. *Let $\mathbf{A} = (a_{ij})$ be an $n \times n$ matrix and let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $\mathbf{A}$ with multiplicity and let $\mathbf{B} = (b_{ij})$ be an $n \times n$ matrix with eigenvalues $\alpha_1, \ldots, \alpha_s$ with multiplicity m_i, for $i = 1, \ldots, s$ with $\sum_{i=1}^{s} m_i = n$. Then for an $\epsilon > 0$ sufficiently small, there is a $\delta = \delta(\epsilon) > 0$ such that if $|a_{ij} - b_{ij}| \leqslant \delta$ for $i, j = 1, \ldots, n$ then $\mathbf{A}$ has m_i eigenvalues satisfying $|\lambda_i - \alpha_i| < \epsilon$ for each $i = 1, \ldots, s$.*

Proof. For a proof see Franklin (1968) page 191. □

In a theoretical pairwise comparison matrix, by Theorem 4.2.10 it holds that

$$\mathbf{A}\mathbf{w} = n\mathbf{w}, \qquad \mathbf{w} \neq 0$$

where n is the largest eigenvalue, $\mathbf{w} = (w_1, \ldots, w_n)^{\mathbf{T}}$ its corresponding eigenvector and all the other eigenvalues λ_i are zero. Therefore if we can ensure that a practical pairwise comparison matrix $\mathbf{P}$ is close enough to some theoretical pairwise comparison $\mathbf{A}$ defined by weights $\mathbf{w}_i$ then by Theorem 4.2.14 we can ensure that the eigenvalues of such matrix $\mathbf{P}$ are close enough to the eigenvalues of $\mathbf{A}$ and therefore can give us a reliable value for the weights of the criteria in $\mathbf{P}$, and so

$$\mathbf{P}\mathbf{w} = \lambda_{max}\mathbf{w}, \qquad \mathbf{w} \neq 0$$

where λ_{max} is the largest eigenvalue, $\mathbf{w} = (w_1, \ldots, w_n)^{\mathbf{T}}$ its corresponding eigenvector, and all the other eigenvalues λ_i may or may be not zero.

Note that our major concern now is how to decide that the matrix $\mathbf{P}$ has an accepted level of consistency μ. Let us address this discussion.

We know that a practical pairwise comparison matrix is positive and reciprocal but not always consistent. However we know by Theorem 4.2.13 that a positive reciprocal matrix is consistent if and only if $\lambda_{max} = n$. Therefore a consistency index may be calculated as the deviation of λ_{max} from n.

Using Corollary 4.2.7 we know that

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

Without loss of generality let $\lambda_1 = \lambda_{max}$ then

$$\sum_{i=2}^{n} \lambda_i + \lambda_{max} = n$$

and

$$\lambda_{max} - n = -\sum_{i=2}^{n} \lambda_i \tag{4.41}$$

Therefore a consistency may be measured by considering the right hand side of Equation (4.41) which if **P** is *close* to being consistent, then $-\sum_{i=1}^{n} \lambda_i$ will be close to zero. Then let us determine a measure of consistency by determining the value of the remaining eigenvalues in (4.41), thus let[4]

$$\begin{aligned}
\mu &= \frac{-1}{n-1} \sum_{i=2}^{n} \lambda_i \\
&= \frac{-1}{n-1} \left(\sum_{i=1}^{n} \lambda_i - \lambda_1 \right) \\
&= \frac{-1}{n-1} \left(n - \lambda_{max} \right) \\
&= \frac{\lambda_{max} - n}{n-1}
\end{aligned}$$

Then the following definition gives us an equation to measure consistency.

Definition (Consistency index) Given an $n \times n$ practical pairwise comparison matrix **P**, a consistency index can be calculated by

$$CI = \frac{\lambda_{max} - n}{n-1} \tag{4.42}$$

where λ_{max} is the largest eigenvalue of **P** and n is the number of rows.

Now we are able to introduce a more specific result applicable to pairwise comparison matrices.

[4] Saaty (1990), page 180.

Theorem 4.2.15. *Let* $\mathbf{P} = (p_{ij})$ *be an* $n \times n$ *practical pairwise comparison matrix. If* $\mathbf{P}$ *has a consistency index* $CI \leqslant \mu$ *then* $\exists$ *a theoretical pairwise comparison matrix* $\mathbf{A}$ *defined by* $w_1, \ldots, w_n$ *such that* $\mathbf{P}$ *is close to* $\mathbf{A}$*. Then the largest eigenvalue of* $\mathbf{P}$*,* λ_{max}*, remains close to* n*,* $\mathbf{w} = (w_1, w_2, \ldots, w_n)^{\mathbf{T}}$ *is its corresponding eigenvector, and all the others eigenvalues* λ_i *remain close to zero.*

Proof. By Theorem 4.2.10 we know that given a theoretical pairwise comparison matrix $\mathbf{A}$ then n is the largest eigenvalue of $\mathbf{A}$, and $\mathbf{w} = (w_1, w_2, \ldots, w_n)^{\mathbf{T}}$ is its corresponding eigenvector, and all the other eigenvalues λ_i are equal to zero. Also by Theorem 4.2.14 we know that small variations in the entries a_{ij} will leave the largest eigenvalue λ_{max} close to n and all the others eigenvalues λ_i close to zero. If the level of consistency can be agreed to be less than some μ as to ensure that the entries a_{ij} will change by small amounts, then using Theorem 4.2.14.

□

4.2.3 Existence and uniqueness of the dominating eigenvalue

In this section we will show the existence and uniqueness of the largest eigenvalue λ_{max}. To do this we use a fundamental result presented by Frobenius (1912) for non-negative matrices (i.e. $\mathbf{A} \geqslant 0$), which is a generalisation of the theorem presented by Perron (1907) for positive matrices (i.e. $\mathbf{A} > 0$). Before presenting the Theorem we recall several definitions and prove several facts. Recall the definitions of an irreducible matrix and the spectral radius of a matrix.

Definition (Irreducible matrix) Let $\mathbf{A}$ be an $n \times n$ matrix. Then $\mathbf{A}$ is said to be irreducible if it *cannot* be changed by permutations of rows and columns to the form

$$\begin{pmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{A}_2 & \mathbf{A}_3 \end{pmatrix} \tag{4.43}$$

where $\mathbf{A}_1$ and $\mathbf{A}_3$ are square matrices, and $\mathbf{0}$ is the zero matrix.

Remark Since pairwise comparison matrices are strictly positive, they are irreducible.

Definition (Spectral radius of a matrix) Let $\mathbf{A}$ be an $n \times n$ matrix. Then the spectral radius of the matrix $\mathbf{A}$, denoted by $\rho(\mathbf{A})$, is given by

$$\rho(\mathbf{A}) = \max\Big\{ |\, \lambda \,| \, : \lambda \text{ is an eigenvalue of } \mathbf{A} \Big\} \tag{4.44}$$

By convention we use $\rho_{\mathbf{A}}$ instead of $\rho(\mathbf{A})$. Now we prove a simple fact.

Lemma 4.2.16. *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix and* $\mathbf{x}$ *be a vector in* $\mathbb{R}^n$*. If* $\mathbf{A} > 0$, $\mathbf{x} \geqslant 0$, *and* $\mathbf{x} \neq 0$ *then* $\mathbf{A}\mathbf{x} > 0$.

Proof. Let $\mathbf{y} = \mathbf{A}\mathbf{x}$ such that $y_i = \sum_{j=1}^{n} a_{ij} x_j$ for $i = 1, \ldots, n$. We need to prove that $y_i > 0$ for $i = 1, \ldots, n$. Suppose $y_i = 0$ for some i, then since $a_{ij} > 0$ for $i, j = 1, \ldots, n$ and since $x_j \geqslant 0$ then this implies that $\mathbf{x} = 0$ which contradicts the fact that $\mathbf{x} \neq 0$. Therefore $\mathbf{A}\mathbf{x} > 0$.

□

Now we introduce the definitions of maximum row sum and maximum column sum matrix norms, induced norms and we show that these two norms are induced.

Definition (Maximum row sum matrix norm) The maximum row sum matrix norm is defined by

$$\|\mathbf{A}\|_\infty = \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |a_{ij}| \tag{4.45}$$

Definition (Maximum column sum matrix norm) The maximum column sum matrix norm is defined by

$$\|\mathbf{A}\|_1 = \max_{1\leqslant j\leqslant n} \sum_{i=1}^{n} |a_{ij}| \tag{4.46}$$

Definition (Induced norms) Let $\| \centerdot \|$ be a vector norm in $\mathbb{C}^n$. Then the vector norm induces a norm on real or complex matrices $\boldsymbol{\| \centerdot \|}$ defined by

$$\|\mathbf{A}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\| \tag{4.47}$$

We say that a matrix norm is induced if there is a vector norm that induces it.

Remark If $\mathbf{x} \neq \mathbf{0}$ and letting $\mathbf{z} = \frac{\mathbf{x}}{\|\mathbf{x}\|}$ we have

$$\|\mathbf{Az}\| = \left\| \frac{\mathbf{Ax}}{\|\mathbf{x}\|} \right\| = \frac{1}{\|\mathbf{x}\|}\|\mathbf{Ax}\|$$

and the definition of induced norms is equivalent to

$$\|\mathbf{A}\| = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} \tag{4.48}$$

Definition (Max norm) The max norm is a vector norm denoted by $\|\mathbf{x}\|_\infty$ in $\mathbb{C}^n$ and is defined by

$$\|\mathbf{x}\|_\infty = \max\{|x_1|, \dots, |x_n|\} \tag{4.49}$$

Lemma 4.2.17. *The maximum row sum matrix norm is an induced norm.*

Proof. Let $\| \cdot \|_\infty$ be the max norm in $\mathbb{C}^n$. Using (4.48) we have for $\mathbf{x} = (x_1, \dots, x_n)^{\mathbf{T}} \neq 0$,

$$\begin{aligned}
\frac{\|\mathbf{Ax}\|_\infty}{\|\mathbf{x}\|_\infty} &= \max_{1\leqslant i\leqslant n} \frac{\left|\sum_{j=1}^n a_{ij}x_j\right|}{\|\mathbf{x}\|_\infty} \\
&\leqslant \max_{1\leqslant i\leqslant n} \frac{\sum_{j=1}^n |a_{ij}x_j|}{\|\mathbf{x}\|_\infty} \\
&\leqslant \max_{1\leqslant i\leqslant n} \frac{\sum_{j=1}^n |a_{ij}| \ \ \|\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} \\
&= \|\mathbf{A}\|_\infty
\end{aligned}$$

Observe that equality holds in both inequalities if $x_j = \epsilon_{ij} = sign(a_{ij})$ were i is the row with greatest sum in $\mathbf{A}$. Therefore

$$\max_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{Ax}\|_\infty}{\|\mathbf{x}\|_\infty} = \|\mathbf{A}\|_\infty$$

□

Definition (Sum norm) The sum norm is a vector norm denoted by $\|\mathbf{x}\|_1$ in $\mathbb{C}^n$ and is defined as

$$\|\mathbf{x}\|_1 = |x_1| + \ldots + |x_n| \tag{4.50}$$

Lemma 4.2.18. *The maximum column sum matrix norm is an induced norm.*

Proof. Let $\| \centerdot \|$ be the sum norm in $\mathbb{C}^n$ with $\mathbf{x} = (x_1, \ldots, x_n)^{\mathbf{T}}$ and $\mathbf{x} \neq 0$. Using (4.48) we have

$$\begin{aligned} \frac{\|\mathbf{Ax}\|_1}{\|\mathbf{x}\|_1} &= \frac{\sum_{i=1}^n |a_{ij}x_j|}{\|\mathbf{x}\|_1} \\ &\leqslant \max_{1\leqslant j\leqslant n} \frac{\sum_{i=1}^n |a_{ij}| \ \ \|\mathbf{x}\|_1}{\|\mathbf{x}\|_1} \\ &= \|\mathbf{A}\|_1 \end{aligned}$$

Observe that equality holds if $\mathbf{x} = \mathbf{e}_j$, where j is the index which maximises $\sum_{i=1}^n |a_{ij}|$. Therefore

$$\max_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{Ax}\|_1}{\|\mathbf{x}\|_1} = \|\mathbf{A}\|_1$$

□

Now we prove that $\rho_{\mathbf{A}} \leqslant \|\mathbf{A}\|$ holds for induced norms as stated in Lemma 4.2.19.

Lemma 4.2.19. *Let* $\| \centerdot \|$ *be any induced matrix norm. Then it holds that*

$$\rho_{\mathbf{A}} \leqslant \|\mathbf{A}\| \tag{4.51}$$

Proof. We start from the fact that

$$\| \mathbf{Ax} \| \leqslant \|\mathbf{A}\| \| \mathbf{x} \| \tag{4.52}$$

for if $\mathbf{x} = \mathbf{0}$ the assertion follows directly. If $\mathbf{x} \neq \mathbf{0}$, using (4.48) for any $\mathbf{y}$ it follows that

$$\|\mathbf{A}\| = \max_{\mathbf{y} \neq \mathbf{0}} \frac{\|\mathbf{Ay}\|}{\|\mathbf{y}\|} \geqslant \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|}$$

Next, recall that an eigenvector and an eigenvalue of $\mathbf{A}$ are a nonzero vector $\mathbf{x}$ and a scalar λ respectively, which satisfy $\mathbf{Ax} = \lambda \mathbf{x}$. Now, taking $\| \mathbf{Ax} \| = \| \lambda \mathbf{x} \|$ and recalling the properties of vector norms[5] and using (4.52) we get

$$| \lambda | \, \| \mathbf{x} \| = \| \mathbf{Ax} \| \leqslant \|\mathbf{A}\| \| \mathbf{x} \|$$

If we take the maximum $| \lambda |$ and using (4.44) we have $\rho_{\mathbf{A}} \leqslant \|\mathbf{A}\|$.

□

Combining Lemmas 4.2.17, 4.2.18 and 4.2.19 we get the following result.

[5]Let $\mathbf{x}$, $\mathbf{y}$ be real or complex vectors in a vector space V. A function $\| \centerdot \| : V \to \mathbb{R}^+$, is called a vector norm, if for all $\mathbf{x}$, $\mathbf{y}$ the following properties are satisfied.

(i) $\|\mathbf{x}\| \geqslant 0$, and $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = 0$ (Nonnegative)

(ii) $\|c\,\mathbf{x}\| = | c | \, \|\mathbf{x}\|$ for all real or complex scalars c (Homogeneous)

(iii) $\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$ (Triangle inequality)

Lemma 4.2.20. *Let $\mathbf{A} = (a_{ij})$ be an $n \times n$ nonnegative matrix, then*

(i) if the row sums of $\mathbf{A}$ are constant then $\rho_{\mathbf{A}} = \|\mathbf{A}\|_\infty$

(ii) if the column sums of $\mathbf{A}$ are constant then $\rho_{\mathbf{A}} = \|\mathbf{A}\|_1$

Proof. [6] Using Lemmas 4.2.17 and 4.2.19 we know that $\rho_{\mathbf{A}} \leqslant \|\mathbf{A}\|_\infty$, and we can see that if the row sums of $\mathbf{A}$ are constant then in the equation $\mathbf{A}\mathbf{x} = \lambda\,\mathbf{x}$, the vector $\mathbf{x} = (1, \ldots, 1)^{\mathbf{T}}$ and the scalar $\|\mathbf{A}\|_\infty$ are an eigenvector and an eigenvalue of $\mathbf{A}$ respectively, and therefore $\rho_{\mathbf{A}} = \|\mathbf{A}\|_\infty$.

Using the same argument as before, but with $\mathbf{A}^T$, we can see that if the column sums of $\mathbf{A}$ are constant then the vector $\mathbf{x} = (1, \ldots, 1)^{\mathbf{T}}$ and the scalar $\|\mathbf{A}\|_1 = \|\mathbf{A}^{\mathbf{T}}\|_\infty$ are an eigenvector and an eigenvalue of $\mathbf{A}^T$ respectively. Using Lemmas 4.2.18 and 4.2.19 we can conclude $\rho_{\mathbf{A}^T} = \|\mathbf{A}\|_1$. Note that $\rho_{\mathbf{A}} = \rho_{\mathbf{A}^T}$ and so $\rho_{\mathbf{A}} = \|\mathbf{A}\|_1$. □

Let us now see the relation of two nonnegative matrices and their spectral radius.

Lemma 4.2.21. *Let $\mathbf{A}, \mathbf{B}$ be two $n \times n$ matrices. If $0 \leqslant \mathbf{A} \leqslant \mathbf{B}$ then $\rho_{\mathbf{A}} \leqslant \rho_{\mathbf{B}}$.*

Proof. For a proof see Horn and Johnson (1999) page 491. □

Now we use the fact that the minimal row sum (or minimal column sum) of $\mathbf{A}$ is a lower bound for the spectral radius $\rho_{\mathbf{A}}$ as stated in Lemma 4.2.22. Although we will only need the lower bounds, similar upper bounds are given in the same Lemma because having proved the lower bounds the proof for the upper bounds is straightforward.

[6] This proof follows the ideas of Horn and Johnson (1999) page 492.

Lemma 4.2.22. *Let* $\mathbf{A} = (a_{ij})$ *be an* $n \times n$ *nonnegative matrix. Then*

$$\min_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} a_{ij} \leqslant \rho_{\mathbf{A}} \leqslant \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} a_{ij} \tag{4.53}$$

$$\min_{1 \leqslant j \leqslant n} \sum_{i=1}^{n} a_{ij} \leqslant \rho_{\mathbf{A}} \leqslant \max_{1 \leqslant j \leqslant n} \sum_{i=1}^{n} a_{ij} \tag{4.54}$$

Proof. [7] To prove the lower bound of equation (4.53), we construct an $n \times n$ matrix $\mathbf{B} = (b_{ij})$ such that $0 \leqslant \mathbf{B} \leqslant \mathbf{A}$. To do this we identify the minimal row sum of $\mathbf{A}$, let us say $\mu = \min_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} a_{ij}$ and then we assign values to b_{ij} such that $\sum_{j=1}^{n} b_{ij} = \mu$ for all i, i.e. the row sums of $\mathbf{B}$ are constant. Specifically we can take

$$b_{ij} = \frac{\mu}{\sum_{k=1}^{n} a_{ik}} a_{ij} \qquad i, j = 1, \ldots, n.$$

On the other hand to prove the upper bounds we will need a matrix $\mathbf{C}$ such that $0 \leqslant \mathbf{A} \leqslant \mathbf{C}$. Now we identify the maximal row sum of $\mathbf{A}$, let us say $\delta = \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} a_{ij}$ and then we assign values to c_{ij} such that $\sum_{j=1}^{n} c_{ij} = \delta$ for all i, i.e. the row sums of $\mathbf{C}$ are constant. Specifically we can take

$$c_{ij} = \frac{\delta}{\sum_{k=1}^{n} a_{ik}} a_{ij} \qquad i, j = 1, \ldots, n.$$

Having constructed matrices $\mathbf{B}$ and $\mathbf{C}$, applying Lemma 4.2.20 part *(i)* we get $\rho_{\mathbf{B}} = \mu$,

[7] This proof follows the ideas of Horn and Johnson (1999) page 492.

$\rho_{\mathbf{C}} = \delta$. Finally combining the last result and using Lemma 4.2.21 we get

$$\rho_{\mathbf{B}} \leqslant \rho_{\mathbf{A}} \leqslant \rho_{\mathbf{C}}$$
$$\mu = \min_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} a_{ij} \leqslant \rho_{\mathbf{A}} \leqslant \delta = \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} a_{ij}$$

In a similar way, to prove the lower and upper bound of equation (4.54) we construct an $n \times n$ matrix $\mathbf{B} = (b_{ij})$ but in this case such that $0 \leqslant \mathbf{B} \leqslant \mathbf{A}^T$. Now we identify the minimal column sum of $\mathbf{A}$, let us say $\mu^* = \min_{1 \leqslant j \leqslant n} \sum_{i=1}^{n} a_{ij}$ and the values of b_{ij} satisfy $\sum_{i=1}^{n} b_{ij} = \mu^*$, i.e. the column sums of $\mathbf{B}$ are constant. For the upper bound we need $\mathbf{C}$ such that $0 \leqslant \mathbf{A}^T \leqslant \mathbf{C}$. Then using the maximal column sum of $\mathbf{A}$, let us say $\delta^* = \max_{1 \leqslant j \leqslant n} \sum_{i=1}^{n} a_{ij}$ we assign the values c_{ij} such that $\sum_{i=1}^{n} c_{ij} = \delta^*$, i.e. the column sums of $\mathbf{C}$ are constant. Again using Lemmas 4.2.20 and 4.2.21 and since $\rho_{\mathbf{A}} = \rho_{\mathbf{A}^T}$ we conclude

$$\rho_{\mathbf{B}} \leqslant \rho_{\mathbf{A}} \leqslant \rho_{\mathbf{C}}$$
$$\mu^* = \min_{1 \leqslant j \leqslant n} \sum_{i=1}^{n} a_{ij} \leqslant \rho_{\mathbf{A}} \leqslant \delta^* \max_{1 \leqslant j \leqslant n} \sum_{i=1}^{n} a_{ij}$$

□

Let us now introduce a generalisation of Lemma 4.2.22.

Lemma 4.2.23. *Let* $\mathbf{A} = (a_{ij})$ *be an* $n \times n$ *nonnegative matrix and let* $\mathbf{x} = (x_1, \ldots, x_n)^{\mathbf{T}} > 0$ *be a vector in* $\mathbb{C}^n$. *Then*

$$\min_{1 \leqslant i \leqslant n} \frac{1}{x_i} \sum_{j=1}^{n} a_{ij} x_j \leqslant \rho_{\mathbf{A}} \leqslant \max_{1 \leqslant i \leqslant n} \frac{1}{x_i} \sum_{j=1}^{n} a_{ij} x_j \tag{4.55}$$

$$\min_{1 \leqslant j \leqslant n} x_j \sum_{i=1}^{n} \frac{a_{ij}}{x_i} \leqslant \rho_{\mathbf{A}} \leqslant \max_{1 \leqslant j \leqslant n} x_j \sum_{i=1}^{n} \frac{a_{ij}}{x_i} \tag{4.56}$$

Proof. [8] Let $\mathbf{z}$ be an eigenvector of $\mathbf{A}$ with eigenvalue λ. Let $\mathbf{Q}$ be an invertible matrix. Then it holds that

$$\mathbf{A}\mathbf{z} = \lambda \mathbf{z}$$
$$(\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q})\mathbf{Q}^{-1}\mathbf{z} = \lambda \mathbf{Q}^{-1}(\mathbf{Q}\mathbf{Q}^{-1})\mathbf{z}$$

Let $\mathbf{y} = \mathbf{Q}^{-1}\mathbf{z}$. Then

$$(\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q})\mathbf{y} = \lambda \mathbf{y}$$

Thus $\mathbf{y}$ is an eigenvalue of $(\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q})$ with eigenvalue λ. Then an eigenvalue of $\mathbf{A}$ is also an eigenvalue of $(\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q})$. On the other hand if λ is an eigenvalue of $(\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q})$ with eigenvector $\mathbf{y} = \mathbf{Q}^{-1}\mathbf{z}$ then

$$(\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q})\mathbf{Q}^{-1}\mathbf{z} = \lambda \mathbf{Q}^{-1}\mathbf{z}$$
$$(\mathbf{Q}\mathbf{Q}^{-1})\mathbf{A}(\mathbf{Q}\mathbf{Q}^{-1})\mathbf{z} = \lambda(\mathbf{Q}\mathbf{Q}^{-1})\mathbf{z}$$
$$\mathbf{A}\mathbf{z} = \lambda \mathbf{z}$$

Therefore any eigenvalue of $(\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q})$ is also an eigenvalue of $\mathbf{A}$, and so the two matrices

[8] This proof follows the ideas of Horn and Johnson (1999) page 493.

have exactly the same eigenvalues, which implies that $\rho_{(\mathbf{Q}^{-1}\mathbf{AQ})} = \rho_{\mathbf{A}}$. Define $\mathbf{Q} = diag(x_1, \ldots, x_n)$. Since $\mathbf{x} > 0$ and $\mathbf{A} \geqslant 0$ then $\mathbf{Q}^{-1}\mathbf{AQ} \geqslant 0$. Observe that the entries of $\mathbf{Q}^{-1}\mathbf{AQ}$ have the form $\left(\frac{a_{ij}x_j}{x_i}\right)$ for $i, j = 1, \ldots, n$, thus since $\rho_{(\mathbf{Q}^{-1}\mathbf{AQ})} = \rho_{\mathbf{A}}$ applying Lemma 4.2.22 to $(\mathbf{Q}^{-1}\mathbf{AQ}) \geqslant 0$ we get

$$\min_{1 \leqslant i \leqslant n} \frac{1}{x_i} \sum_{j=1}^{n} a_{ij}x_j \leqslant \rho_{\mathbf{A}} \leqslant \max_{1 \leqslant i \leqslant n} \frac{1}{x_i} \sum_{j=1}^{n} a_{ij}x_j$$
$$\min_{1 \leqslant j \leqslant n} x_j \sum_{i=1}^{n} \frac{a_{ij}}{x_j} \leqslant \rho_{\mathbf{A}} \leqslant \max_{1 \leqslant j \leqslant n} x_j \sum_{i=1}^{n} \frac{a_{ij}}{x_i}$$

□

From Lemma 4.2.23 we deduce the following.

Corollary 4.2.24. *Let $\mathbf{A} \geqslant 0$ be an $n \times n$ matrix, let $\mathbf{x} > 0$ be a vector in $\mathbb{R}^n$ and let $\alpha \geqslant 0$. If $\alpha\mathbf{x} < \mathbf{Ax}$ then $\alpha < \rho_{\mathbf{A}}$.*

Proof. [9] Since $\alpha\mathbf{x} < \mathbf{Ax}$ then $\alpha < \min_{1 \leqslant i \leqslant n} \frac{1}{x_i} \sum_{j=1}^{n} a_{ij}x_j$ and so by Lemma 4.2.23 $\alpha < \rho_{\mathbf{A}}$.

□

Now we are going to mention the fundamental discovery made by Perron (1907) for positive matrices.

[9]This proof follows the ideas of Horn and Johnson (1999) page 493.

Theorem 4.2.25. *(Perron) Let* $\mathbf{A}$ *be an* $n \times n$ *positive matrix, then:*

(i) $\mathbf{A}$ *has an eigenvalue equal to* $\rho_{\mathbf{A}}$.

(ii) The eigenvalue $\rho_{\mathbf{A}}$ *is positive.*

(iii) The eigenvalue $\rho_{\mathbf{A}}$ *is a simple (i.e. not multiple) eigenvalue of* $\mathbf{A}$.

(iv) There exists a positive eigenvector $\mathbf{x}$*, which corresponds to the eigenvalue* $\rho_{\mathbf{A}}$.

(v) If λ *is an eigenvalue of* $\mathbf{A}$ *and* $\lambda \neq \rho_{\mathbf{A}}$ *then* $|\lambda| < \rho_{\mathbf{A}}$ *(i.e. unique eigenvalue of maximum modulus).*

Proof. For a proof see Horn and Johnson (1999) page 500.

□

Let us consider the definition of a sequence and a monotone decreasing sequence.

Definition (Sequence) A sequence, denoted by $\{a_k\}_{k=1}^{n}$, is an ordered list of objects.

Definition (Monotone decreasing sequence) A sequence is monotone decreasing if each term is less than or equal to the preceding term in the sequence.

Lemma 4.2.26. *Let* $\mathbf{A} = (a_{ij})$ *be an* $n \times n$ *irreducible matrix. Then* $\mathbf{A}$ *cannot have a zero row or a zero column.*

Proof. Let a^i denote the ith row of $\mathbf{A}$. Suppose $\mathbf{A}$ has at least one row of zeros, let us say α, i.e. $a^i = (0, \ldots, 0)$ for some $i = 1, \ldots, n$. Now by using permutations matrices, it is possible to transform **A**, by moving the row α to the upper part of the matrix, into the form

$$\left(\begin{array}{c|c} 0 & \mathbf{0} \\ \hline \mathbf{A}_1 & \mathbf{A}_2 \end{array}\right) \tag{4.57}$$

where $\mathbf{A}_1$ is an $(n-1) \times 1$ matrix, and $\mathbf{A}_2$ is an $(n-1) \times (n-1)$ matrix. On the other hand, if the matrix a has a column of zeros then the matrix can be transformed, into the form

$$\left(\begin{array}{c|c} \mathbf{A}_1 & \mathbf{0} \\ \hline \mathbf{A}_2 & 0 \end{array}\right) \tag{4.58}$$

where $\mathbf{A}_1$ is, in this case, an $(n-1)\times(n-1)$ matrix, and $\mathbf{A}_2$ is an $1\times(n-1)$ matrix. Observe that in both cases (4.57) and (4.58) the new matrices generated are in the reduced form, contradicting the fact that **A** is irreducible. Therefore **A** has no zero rows or columns, as claimed.

□

Now let us prove a fact about the relation between the spectral radius of a principal submatrix and the spectral radius of **A**.

Lemma 4.2.27. *Let* $\mathbf{A} = (a_{ij})$ *be an* $n \times n$ *irreducible matrix, and let* $\mathbf{A}_i$ *be any principal submatrix of* **A**. *Then* $\rho_{\mathbf{A}_i} < \rho_{\mathbf{A}}$.

Proof. [10] Since the permutation of rows and columns does not change $\rho_{\mathbf{A}}$ then without

[10] This proof follows the ideas of Bapat and Raghavan (1997) page 37.

loss of generality we can take any principal square submatrix, say $\mathbf{A}_1$ such that

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_3 & \mathbf{A}_4 \end{pmatrix}$$

Where $\mathbf{A_1}$ and $\mathbf{A}_4$ are principal square submatrices of $\mathbf{A}$. And set

$$\mathbf{A}' = \begin{pmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{4.59}$$

Observe that $\rho_{\mathbf{A}_1} = \rho_{\mathbf{A}'}$. Now take $\mathbf{B} = (\mathbf{A} + \mathbf{A}')/2$, i.e.

$$\mathbf{B} = \begin{pmatrix} \mathbf{A}_1 & \frac{\mathbf{A}_2}{2} \\ \frac{\mathbf{A}_3}{2} & \frac{\mathbf{A}_4}{2} \end{pmatrix}$$

Observe that $0 \leqslant \mathbf{A}' \leqslant \mathbf{B} \leqslant \mathbf{A}$ and so by Lemma 4.2.21 $\rho_{\mathbf{A}_1} = \rho_{\mathbf{A}'} \leqslant \rho_{\mathbf{B}} \leqslant \rho_{\mathbf{A}}$. Since $\mathbf{A}$ is irreducible and nonnegative then $\mathbf{B}$ is also irreducible and nonnegative. Let $\mathbf{x} > 0$ be an eigenvector of $\mathbf{B^T}$ with eigenvalue $\rho_{\mathbf{B}}$ and let $\mathbf{y} > 0$ be an eigenvector of $\mathbf{A}$ with eigenvalue $\rho_{\mathbf{A}}$. Using Lemma 4.2.16 and because $\mathbf{x} > 0$ and $\mathbf{y} > 0$ and since $\mathbf{B} \neq \mathbf{A}$ then

$$\mathbf{B} \leqslant \mathbf{A}$$

$$\mathbf{x^T B y} < \mathbf{x^T A y}$$

$$\rho_{\mathbf{B}}\, \mathbf{x^T y} < \rho_{\mathbf{A}}\, \mathbf{x^T y}$$

Thus $\rho_{\mathbf{B}} < \rho_{\mathbf{A}}$. However, since $\rho_{\mathbf{A}_1} = \rho_{\mathbf{A}'} \leqslant \rho_{\mathbf{B}} \leqslant \rho_{\mathbf{A}}$, then $\rho_{\mathbf{A}_1} < \rho_{\mathbf{A}}$.

□

Lemma 4.2.28. *Let* $\mathbf{A} = a_{ij}$ *be an* $n \times n$ *matrix. If* $\mathbf{A} \geqslant 0$ *then* $\mathbf{A}$ *is irreducible if and only if*

$$(\mathbf{I} + \mathbf{A})^{n-1} > 0 \tag{4.60}$$

Proof. For a proof see Horn and Johnson (1999) page 362.

□

Now we are able to introduce the Perron-Frobenius Theorem.

Theorem 4.2.29. (Perron-Frobenius) *Let* $\mathbf{A}$ *be an* $n \times n$ *nonnegative irreducible matrix, then:*

(i) $\mathbf{A}$ *has an eigenvalue equal to* $\rho_{\mathbf{A}}$.

(ii) The eigenvalue $\rho_{\mathbf{A}}$ *is positive.*

(iii) The eigenvalue $\rho_{\mathbf{A}}$ *is a simple (i.e. not multiple) eigenvalue of* $\mathbf{A}$.

(iv) There exists a positive eigenvector $\mathbf{x}$*, which corresponds to the eigenvalue* $\rho_{\mathbf{A}}$.

(v) If λ *is an eigenvalue of* $\mathbf{A}$ *with a nonnegative eigenvector, then* $\lambda = \rho_{\mathbf{A}}$ *(i.e. uniqueness).*

Proof. *(i)* [11] Let $\mathbf{A}_\epsilon = (a_{ij} + \epsilon) > 0$ for some appropriate $\epsilon > 0$. By Theorem 4.2.25 there exists $\mathbf{x}_\epsilon > 0$ such that $\mathbf{A}_\epsilon \mathbf{x}_\epsilon = \rho_{\mathbf{A}_\epsilon} \mathbf{x}_\epsilon$ with $\|\mathbf{x}\|_1 = 1$. Since in $\mathbb{C}^n$ the set $\{\mathbf{x} \ : \ \|\mathbf{x}\|_1 = 1\}$ is compact there is a monotone decreasing sequence $\{\epsilon_k\}_{k=1,\ldots,\infty}$ such that

$$\lim_{k \to \infty} \epsilon_k = 0$$

[11] This part of the proof follows the ideas of Horn and Johnson (1999) page 503.

and such that

$$\mathbf{x} = \lim_{k\to\infty} \mathbf{x}_{\epsilon_k}$$

exists. But then since $\mathbf{x}_{\epsilon_k} > 0$ for $k = 1, \ldots, \infty$ then $\mathbf{x} \geqslant 0$. Also observe that $\mathbf{x} \neq 0$ because

$$\sum_{i=1}^{n} \mathbf{x}_i = \lim_{k\to\infty} \sum_{i=1}^{n} (\mathbf{x}_{\epsilon_k})_i = 1$$

Observe that $\mathbf{A}_{\epsilon_k} \geqslant \mathbf{A}_{\epsilon_{k+1}} \geqslant \ldots \geqslant \mathbf{A}$ for $k = 1, \ldots, \infty$ is also a monotone decreasing sequence, and so by Lemma 4.2.21 we get $\rho_{\mathbf{A}_{\epsilon_k}} \geqslant \rho_{\mathbf{A}_{\epsilon_{k+1}}} \geqslant \ldots \geqslant \rho_{\mathbf{A}}$, thus

$$r = \lim_{k\to\infty} \rho_{\mathbf{A}_{\epsilon_k}}$$

also exists. Note that

$$r \geqslant \rho_{\mathbf{A}} \tag{4.61}$$

Now, observe that

$$\begin{aligned}
\mathbf{A}\mathbf{x} &= \lim_{k\to\infty} \mathbf{A}_{\epsilon_k} \mathbf{x}_{\epsilon_k} \\
&= \lim_{k\to\infty} \rho_{\mathbf{A}_{\epsilon_k}} \mathbf{x}_{\epsilon_k} \\
&= \lim_{k\to\infty} \rho_{\mathbf{A}_{\epsilon_k}} \lim_{k\to\infty} \mathbf{x}_{\epsilon_k} \\
&= r\mathbf{x}
\end{aligned}$$

Thus since $\mathbf{x} \neq 0$ then r is an eigenvalue of $\mathbf{A}$ and as such satisfies

$$r \leqslant \rho_{\mathbf{A}} \tag{4.62}$$

Combining (4.61) and (4.62) we get $r = \rho_{\mathbf{A}}$, and so $\rho_{\mathbf{A}}$ is an eigenvalue of $\mathbf{A}$.

(ii) [12] Since $\mathbf{A}$ is a nonnegative matrix then for every entry it holds that $a_{ij} \geqslant 0$, therefore using Lemma 4.2.26 we know that

$$\sum_{i=1}^{n} a_{ij} > 0, \qquad j = 1, \ldots, n \qquad \left(\sum_{j=1}^{n} a_{ij} > 0, \qquad i = 1, \ldots, n \right)$$

Combining this result with Lemma 4.2.22 we get

$$0 < \min_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} a_{ij} \leqslant \rho_{\mathbf{A}} \quad \left(0 < \min_{1 \leqslant j \leqslant n} \sum_{i=1}^{n} a_{ij} \leqslant \rho_{\mathbf{A}} \right)$$

Thus $\rho_{\mathbf{A}} > 0$.

(iii) [13] We need to show that the eigenvalue $\rho_{\mathbf{A}}$ of $\mathbf{A}$ has multiplicity one in the solution of the characteristic polynomial $\det(\lambda \mathbf{I}_n - \mathbf{A}) = 0$. If $\mathbf{A}_i$ is any principal submatrix of $\mathbf{A}$ then by Lemma 4.2.27 $\rho_{\mathbf{A}_i} < \rho_{\mathbf{A}}$ and also by *(ii)* we know that $\rho_{\mathbf{A}} > 0$. Therefore if we calculate $\det(\lambda \mathbf{I}_{n-1} - \mathbf{A}_i)$, this value cannot be zero for any $\lambda \geqslant \rho_{\mathbf{A}}$, thus

$$\det(\rho_{\mathbf{A}} \mathbf{I} - \mathbf{A}_i) > 0 \tag{4.63}$$

Therefore by (4.63) the sum of principal subminors $E_{(n-1)}(\rho_{\mathbf{A}} \mathbf{I}_{n-1} - \mathbf{A}_i) > 0$. In other words the matrix $(\rho_{\mathbf{A}} \mathbf{I}_{n-1} - \mathbf{A}_i)$ is nonsingular for any principal submatrix $\mathbf{A}_i$ with $1 \leqslant i \leqslant n$ and the polynomial determined by $\det(\rho_{\mathbf{A}} \mathbf{I}_{n-1} - \mathbf{A}_i)$ does not have a solution for those submatrices. Thus $\rho_{\mathbf{A}}$ corresponds only to one zero in the solution of $\det(\lambda \mathbf{I}_n - \mathbf{A})$.

(iv) [14] We have seen in *(i)* that there exists an eigenvector $\mathbf{x} \geqslant 0$ with $\mathbf{x} \neq 0$ such that

[12] This part of the proof follows the ideas of Horn and Johnson (1999) page 492.
[13] This part of the proof follows the ideas of Varga (1962) page 31.
[14] This part of the proof follows the ideas of Horn and Johnson (1999) page 508.

$\mathbf{Ax} = \rho_{\mathbf{A}}\mathbf{x}$. And so we get the equivalence

$$(\mathbf{I}+\mathbf{A})^{n-1}\mathbf{x} = (1+\rho_{\mathbf{A}})^{n-1}\mathbf{x}$$

therefore

$$\mathbf{x} = \frac{(\mathbf{I}+\mathbf{A})^{n-1}\mathbf{x}}{(1+\rho_{\mathbf{A}})^{n-1}} \tag{4.64}$$

Using Lemma 4.2.28 we know that $(\mathbf{I}+\mathbf{A})^{n-1} > 0$. But then with $(\mathbf{I}+\mathbf{A})^{n-1} > 0$ and $\mathbf{x} \geqslant 0$ Lemma 4.2.16 indicates that $(\mathbf{I}+\mathbf{A})^{n-1}\mathbf{x} > 0$. Also, $\rho_{\mathbf{A}} > 0$, thus from (4.64) we get $\mathbf{x} > 0$.

(v) [15] We know by *(i)* and *(iv)* that $\exists\ \mathbf{x} > 0$ satisfying $\mathbf{A^T x} = \rho_{\mathbf{A}}\mathbf{x}$ (observe that this holds because $\mathbf{A^T} \geqslant 0$ is irreducible and $\rho_{\mathbf{A^T}} = \rho_{\mathbf{A}}$). Let λ and $\mathbf{y} \geqslant 0$, with $\mathbf{y} \neq 0$ be a scalar and a vector such that $\mathbf{Ay} = \lambda\mathbf{y}$. Note that $\mathbf{yTx} \neq 0$. Then

$$\begin{aligned} \mathbf{A^T x} &= \rho_{\mathbf{A}}\mathbf{x} \\ \mathbf{y^T A^T x} &= \rho_{\mathbf{A}}\mathbf{y^T x} \\ \lambda\,\mathbf{y^T x} &= \rho_{\mathbf{A}}\mathbf{y^T x} \end{aligned}$$

Thus $\lambda = \rho_{\mathbf{A}}$, which means that there are no nonnegative eigenvectors for $\mathbf{A}$ corresponding to the eigenvalues other than $\rho_{\mathbf{A}}$, even those corresponding to different eigenvalues. However by *(iii)* we know that $\rho_{\mathbf{A}}$ has multiplicity 1, and so $\rho_{\mathbf{A}}$ and its corresponding eigenvector $\mathbf{x}$ are uniquely determined (up to a scalar multiple).

□

[15] This part of the proof follows the ideas of Meyer (2000) page 673.

4.2.4 Determination of the eigenvalue λ_{max} and its corresponding eigenvector $\mathbf{w}$

Having proved the existence and uniqueness of the principal eigenvector for pairwise comparison matrices, we will see how the eigenvector $\mathbf{w}$ corresponding to largest eigenvalue λ_{max} can be calculated in a nonnegative irreducible matrix. However we will start by showing how the largest eigenvalue λ_{max} is calculated given a positive matrix. Let us introduce the definition of the Jordan block and the existence of a Jordan block canonical form of a matrix.

Theorem 4.2.30. *Let $\mathbf{A}_{n\times n} > 0$. If $\lambda_i \neq \lambda_j$ for all i and j, then*

$$\lim_{k\to\infty} \frac{\mathbf{A}^k\mathbf{e}}{\mathbf{e}^{\mathbf{T}}\mathbf{A}^k\mathbf{e}} = c\mathbf{w}_1 \tag{4.65}$$

Where $\mathbf{e} = (1,\dots,1)^{\mathbf{T}}$, $\mathbf{w}_1$ is the eigenvector corresponding to $\lambda_1 = \lambda_{max}$ and c is a constant.

Proof. [16] Since $\mathbf{w}_i$ is the eigenvector corresponding to λ_i, then

$$\mathbf{A}\mathbf{w}_i = \lambda_i\mathbf{w}_i \qquad i = 1,\dots,n$$

$$\mathbf{A}^k(a_1\mathbf{w}_1 + \dots, a_n\mathbf{w}_n) = a_1\lambda_1^k\mathbf{w}_1 + \dots + a_n\lambda_n^k\mathbf{w}_n$$

[16] See Saaty (1990) page 176.

Let $\mathbf{e} = a_1\mathbf{w}_1 + \ldots + a_n\mathbf{w}_n$ with $a_1, \ldots, a_n$ constants. And so

$$\begin{aligned}\mathbf{A}^k\mathbf{e} &= a_1\lambda_1^k\mathbf{w}_1 + \ldots + a_n\lambda_n^k\mathbf{w}_n \\ &= \lambda_1^k\left[a_1\mathbf{w}_1 + a_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\mathbf{w}_2 + \ldots + a_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\mathbf{w}_n\right]\end{aligned}$$

and also

$$\begin{aligned}\mathbf{e}^{\mathbf{T}}\mathbf{A}^k\mathbf{e} &= a_1\lambda_1^k\mathbf{e}^{\mathbf{T}}\mathbf{w}_1 + \ldots + a_n\lambda_n^k\mathbf{e}^{\mathbf{T}}\mathbf{w}_n \\ &= b_1\lambda_1^k + \ldots + b_n\lambda_n^k \\ &= \lambda_1^k\left[b_1 + b_2\left(\frac{\lambda_2}{\lambda_1}\right)^k + \ldots + b_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\right]\end{aligned}$$

where $b_i = a_i\mathbf{e}^{\mathbf{T}}\mathbf{w}_i$ therefore

$$\frac{\mathbf{A}^k\mathbf{e}}{\mathbf{e}^{\mathbf{T}}\mathbf{A}^k\mathbf{e}} = \frac{\lambda_1^k\left[a_1\mathbf{w}_1 + a_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\mathbf{w}_2 + \ldots + a_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\mathbf{w}_n\right]}{\lambda_1^k\left[b_1 + b_2\left(\frac{\lambda_2}{\lambda_1}\right)^k + \ldots + b_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\right]}$$

and so as $k \to \infty$, the terms with $\left(\frac{\lambda_i}{\lambda_1}\right)^k$ for $i = 2, \ldots, n$ tend to zero and so

$$\begin{aligned}\lim_{k\to\infty}\frac{\mathbf{A}^k\mathbf{e}}{\mathbf{e}^{\mathbf{T}}\mathbf{A}^k\mathbf{e}} &= \frac{\lambda_1^k a_1\mathbf{w}_1}{\lambda_1^k b_1} \\ &= c\mathbf{w}_1 \qquad (4.66)\end{aligned}$$

□

Definition (Jordan block) A Jordan block is a $k \times k$ upper triangular matrix with the eigenvalue λ in every element of the main diagonal, with ones in the superdiagonal and zeros in all the other entries, i.e.

$$\mathbf{J}_k(\lambda) = \begin{pmatrix} \lambda & 1 & & \mathbf{0} \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ \mathbf{0} & & & \lambda \end{pmatrix} \tag{4.67}$$

Lemma 4.2.31. *Let* $\mathbf{A}$ *be an* $n \times n$ *complex matrix. Then there exists a nonsingular* $n \times n$ *matrix* $\mathbf{S}$ *such that*

$$\mathbf{A} = \mathbf{SJS}^{-1} \tag{4.68}$$

with

$$\mathbf{J} = \begin{pmatrix} \mathbf{J}_{n_1}(\lambda_1) & & & 0 \\ & \mathbf{J}_{n_2}(\lambda_2) & & \\ & & \ddots & \\ 0 & & & \mathbf{J}_{n_r}(\lambda_r) \end{pmatrix}$$

where $\mathbf{J}_{n_i}$, *for* $i = 1, \ldots, r$, *corresponds to a Jordan block. Also* $n_1 + n_2 + \ldots + n_r = n$, *and the values of* λ_i *are not necessarily distinct.*

Proof. For a proof see Horn and Johnson (1999) page 121. □

Now let us prove the binomial theorem.

Theorem 4.2.32. *(Binomial theorem) The expansion of a polynomial* $(x+y)^n$ *has the form*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \tag{4.69}$$

Proof. We will prove this by induction. Our induction hypothesis is that the expansion of a polynomial $(x+y)^n$ has the form of (4.69). For $n = 1$ using (4.69) we have

$$\binom{1}{0} x^1 y^0 + \binom{1}{1} x^0 y^1 = x + y$$

Expanding directly in the expression we get $(x+y)^1 = x+y$, and the coefficients of k are the same as found before. Assuming that the hypothesis holds for some n we are going to prove that it holds for $n+1$ as well. Using (4.69) we have

$$\begin{aligned}
(x+y)^{n+1} = (x+y)(x+y)^n &= (x+y) \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \\
&= \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k+1} + \sum_{i=1}^{n+1} \binom{n}{i-1} x^i y^{n-i+1} \\
&= y^{n+1} + x^{n+1} + \sum_{j=1}^{n} \left(\binom{n}{j} + \binom{n}{j-1} \right) x^j y^{n-j+1} \\
&= y^{n+1} + x^{n+1} + \sum_{j=1}^{n} \binom{n+1}{j} x^j y^{n-j+1}
\end{aligned}$$

where $i = k+1$. Therefore

$$(x+y)^{n+1} = \sum_{j=0}^{n+1} \binom{n+1}{j} x^j y^{n+1-j}$$

which completes the proof.

□

Theorem 4.2.33. *Let* $\mathbf{A}$ *be a* $n \times n$ *positive matrix. Then*

$$\lim_{k\to\infty} \frac{\mathbf{A}^k\mathbf{e}}{\mathbf{e}^{\mathbf{T}}\mathbf{A}^k\mathbf{e}} = c\mathbf{w} \tag{4.70}$$

where $\mathbf{e} = (1, \ldots, 1)^{\mathbf{T}}$, $\mathbf{w} = \mathbf{w}_1$ *is the principal eigenvector corresponding to the largest eigenvalue* $\lambda_1 = \lambda_{max}$ *and* c *is a constant.*

Proof. [17]

Remark Theorem 4.2.33 is stronger than Theorem 4.2.30 because in the latter we do not assume that all the eigenvalues are distinct.

Let $\mathbf{A} = \mathbf{SJS}^{-1}$ be the Jordan factorisation of $\mathbf{A}$ according with Lemma 4.2.31, such that $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_r$. Since $\mathbf{A}$ is positive then λ_1 is simple by Theorem 4.2.25, and so

$$\mathbf{J} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \mathbf{J}_{n_2}(\lambda_2) & & \\ & & \ddots & \\ 0 & & & \mathbf{J}_{n_r}(\lambda_r) \end{pmatrix}$$

[17] This proof follows the ideas of Saaty (1990) page 176.

or specifically,

$$\mathbf{J} = \begin{pmatrix} \lambda_1 & & & & & & & & & \mathbf{0} \\ & \lambda_2 & 1 & 0 & \ddots & & & & & \\ & 0 & \lambda_2 & \ddots & 0 & & & & & \\ & \vdots & \ddots & \ddots & 1 & & & & & \\ & 0 & \cdots & 0 & \lambda_2 & & & & & \\ & & & & & \ddots & & & & \\ & & & & & & \lambda_r & 1 & 0 & \ddots \\ & & & & & & 0 & \lambda_r & \ddots & 0 \\ & & & & & & \vdots & \ddots & \ddots & 1 \\ \mathbf{0} & & & & & & 0 & \cdots & 0 & \lambda_r \end{pmatrix}$$

Now for each Jordan block $J_{ni}(\lambda_i)$, for $i = 1, \ldots, r$, we define the vectors

$$\begin{aligned} &\mathbf{w}_1 = \mathbf{S}\mathbf{e}_1 \\ &\quad \mathbf{w}_{21} = \mathbf{S}\mathbf{e}_2 \\ &\quad \mathbf{w}_{22} = \mathbf{S}\mathbf{e}_3 \\ &\quad \vdots \\ &\quad \mathbf{w}_{2n_2} = \mathbf{S}\mathbf{e}_{n_2+1} \\ &\qquad \ddots \\ &\qquad\quad \mathbf{w}_{r1} = \mathbf{S}\mathbf{e}_{n_1+n_2+\ldots+n_{r-1}+1} \\ &\qquad\quad \mathbf{w}_{r2} = \mathbf{S}\mathbf{e}_{n_1+n_2+\ldots+n_{r-1}+2} \\ &\qquad\quad \vdots \\ &\qquad\quad \mathbf{w}_{rn_r} = \mathbf{S}\mathbf{e}_n \end{aligned} \tag{4.71}$$

Such that for each Jordan Block we have the basis vectors

$$\begin{aligned}
&\{\mathbf{w}_1\} \\
&\{\mathbf{w}_{21}, \ldots, \mathbf{w}_{2n_2}\} \\
&\vdots \\
&\{\mathbf{w}_{r1}, \ldots, \mathbf{w}_{rn_r}\}
\end{aligned} \tag{4.72}$$

Now let

$$\begin{aligned}
\mathbf{e} = {} & a_1\mathbf{w}_1 \\
& + a_{21}\mathbf{w}_{21} + \ldots + a_{2n_2}\mathbf{w}_{2n_2} \\
& + \ldots \\
& + a_{r1}\mathbf{w}_{r1} + \ldots + a_{rn_r}\mathbf{w}_{rn_r}
\end{aligned}$$

where $a_1, a_{21}, \ldots, a_{rn_r}$ are constants such that $\mathbf{e} = (1, \ldots, 1)^{\mathbf{T}}$. Then

$$\begin{aligned}
\mathbf{A}\mathbf{w}_1 &= (\mathbf{S}\mathbf{J}\mathbf{S}^{-1})(\mathbf{S}\mathbf{e}_1) \\
&= \mathbf{S}\lambda_1\mathbf{e}_1 \\
&= \lambda_1\mathbf{w}_1
\end{aligned}$$

and also

$$\mathbf{A}\mathbf{w}_{ij} = \begin{cases} \lambda_i\mathbf{w}_{ij} + \mathbf{w}_{i(j-1)}, & \text{if } j > 1; \\ \lambda_i\mathbf{w}_{ij}, & \text{if } j = 1. \end{cases}$$

And so using Lemma 4.2.32 we know that

$$\begin{aligned}
\mathbf{A}^k\mathbf{e} = a_1\lambda_1^k\mathbf{w}_1 & \\
& + a_{21}\lambda_2^k\mathbf{w}_{21} \\
& + a_{22}\binom{k}{0}\lambda_2^k\mathbf{w}_{22} + a_{22}\binom{k}{1}\lambda_2^{k-1}\mathbf{w}_{21} \\
& \vdots \\
& + a_{2n_2}\binom{k}{0}\lambda_2^k\mathbf{w}_{2n_2} + a_{2n_2}\binom{k}{1}\lambda_2^{k-1}\mathbf{w}_{2(n_2-1)} + \ldots + a_{2n_2}\binom{k}{n_2-1}\lambda_2^{k-n_2+1}\mathbf{w}_{21} \\
& \vdots \\
& + a_{r1}\binom{k}{0}\lambda_r^k\mathbf{w}_{r1} \\
& + a_{r2}\binom{k}{0}\lambda_r^k\mathbf{w}_{r2} + a_{r2}\binom{k}{1}\lambda_r^{k-1}\mathbf{w}_{r1} \\
& \vdots \\
& + a_{rn_r}\binom{k}{0}\lambda_r^k\mathbf{w}_{rn_r} + a_{rn_r}\binom{k}{1}\lambda_r^{k-1}\mathbf{w}_{r(n_r-1)} + \ldots + a_{rn_r}\binom{k}{n_r-1}\lambda_r^{k-n_r+1}\mathbf{w}_{r1}
\end{aligned}$$

and so

$$\begin{aligned}
\mathbf{e^T}\mathbf{A}^k\mathbf{e} = a_1\mathbf{e^T}\lambda_1^k\mathbf{w}_1 & \\
& + a_{21}\mathbf{e^T}\lambda_2^k\mathbf{w}_{21} + \ldots + a_{2n_2}\mathbf{e^T}\binom{k}{n_2-1}\lambda_2^{k-n_2+1}\mathbf{w}_{21} \\
& \vdots \\
& + a_{r1}\mathbf{e^T}\lambda_r^k\mathbf{w}_{r1} + \ldots + a_{rn_r}\mathbf{e^T}\binom{k}{n_r-1}\lambda_r^{k-n_r+1}\mathbf{w}_{rn_r}
\end{aligned}$$

Therefore

$$\frac{\mathbf{A}^k\mathbf{e}}{\mathbf{e^T}\mathbf{A}^k\mathbf{e}} = \frac{\lambda_1^k\left[a_1\mathbf{w}_1 + \sum_{i,j,l} a_{ij}\binom{k}{l}\left(\frac{\lambda_i^{k-l}}{\lambda_1^k}\right)\mathbf{w}_{ij}\right]}{\lambda_1^k\left[c_1 + a_{21}\mathbf{e^T}\left(\frac{\lambda_2^k}{\lambda_1^k}\right)\mathbf{w}_{21} + \ldots + a_{rn_r}\mathbf{e^T}\binom{k}{n_r-1}\left(\frac{\lambda_r^{k-n_r+1}}{\lambda_1^k}\right)\mathbf{w}_{rn_r}\right]} \tag{4.73}$$

where $c_1 = a_1\mathbf{e}^{\mathbf{T}}\mathbf{w}_1$, $i = 2, \ldots, r$, $j = 1, \ldots, n_i$ and $l = 0, \ldots, j$. Recall that since $\mathbf{A}$ is positive then by Theorem 4.2.25 the largest eigenvalue $\lambda_{max} = \lambda_1$ is positive. Therefore since λ_1 is the largest eigenvalue, as $k \to \infty$ tends to infinity we will get $(a_1/c_1)\mathbf{w}_1$ and all the other terms tend to zero and so

$$\lim_{k\to\infty} \frac{\mathbf{A}^k\mathbf{e}}{\mathbf{e}^{\mathbf{T}}\mathbf{A}^k\mathbf{e}} = c\mathbf{w}$$

where $c = a_1/c_1$ and $\mathbf{w} = \mathbf{w}_1$.

□

Now, to extend the applicability of the previous theorem to nonnegative irreducible matrices it is necessary to see the geometrical interpretation of an irreducible matrix and the definition of a primitive matrix.

Definition (Graph of a matrix) Let $\mathbf{A}$ be an $n \times n$ matrix. Then $G(\mathbf{A})$ represents the graph of $\mathbf{A}$ and is defined as a directed graph with n nodes and a directed edge from node v_i to v_j if and only if $a_{ij} \neq 0$.

Definition (Strongly connected graph) A graph $G(\mathbf{A})$ with nodes n_i for $i = 1, \ldots, n$ is strongly connected if there exists a directed path connecting n_i with n_j for any $i, j = 1, \ldots, n$.

Lemma 4.2.34. *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix. Then* $\mathbf{A}$ *is irreducible if and only if its directed graph* $\mathrm{G}(\mathbf{A})$ *is strongly connected.*

Proof. Suppose that $\mathbf{A}$ is reducible. Then by permutations it can be transformed into the form

$$\begin{pmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{A}_2 & \mathbf{A}_3 \end{pmatrix} \tag{4.74}$$

where $\mathbf{A}_1$ is a $k_1 \times k_1$ matrix, $\mathbf{A}_3$ is a $k_2 \times k_2$ matrix, and $k_1 + k_2 = n$. If $j \in \{1, \ldots, k\}$ and $i \in \{k_1 + 1, \ldots, n\}$, then $a_{ij} = 0$ and so $\nexists$ edge from v_i to v_j in $G(\mathbf{A})$, that is the graph is not strongly connected. Then there is no edge from S to S', where $S = \{\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ and $S' = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$. Then $\nexists$ path from v_i to v_j for any $v_i \in S$, $v_j \in S'$. On the other hand, if $G(\mathbf{A})$ is not strongly connected, then $\exists$ some j, i such that $\nexists$ path from v_j to v_i. Let S be the set of nodes that can be reached from v_j. Move S to the bottom of the matrix by permutations of rows. Let $S' = \{1, \ldots, n\} \setminus S$ then $\nexists$ edge from S' to S and the matrix looks like (4.74), and so $\mathbf{A}$ is reducible.

□

Definition (Primitive matrix) Let $\mathbf{A}$ be an $n \times n$ nonnegative irreducible matrix. Then $\mathbf{A}$ is *primitive* if and only if it has only one eigenvalue of maximum modulus.

In the case of pairwise comparison matrices, since $a_{ii} > 0$ then $\mathbf{A}^{n-1} > 0$. The following lemma generalises this property.

Lemma 4.2.35. *Let $\mathbf{A} = (a_{ij})$ be an $n \times n$ nonnegative irreducible matrix. If $a_{ii} > 0$ for $i = 1, \ldots, n$ then $\mathbf{A}^{n-1} > 0$.*

Proof. [18] Let $\gamma = \frac{1}{2}\min\{1, a_{11}, \ldots, a_{nn}\}$ so $\gamma > 0$. Also let $\mathbf{B} = \frac{\mathbf{A}}{2}$. Observe that by

[18]This proof follows the ideas of Varga (1962) page 41.

construction, $\mathbf{B}$ is an irreducible nonnegative matrix. Also it holds that

$$\mathbf{A} \geqslant \gamma(\mathbf{I} + \mathbf{B})$$
$$\mathbf{A}^{n-1} \geqslant \gamma^{n-1}(\mathbf{I} + \mathbf{B})^{n-1}$$

Since $\mathbf{B}$ is irreducible then by Lemma 4.2.28 it holds that

$$(\mathbf{I} + \mathbf{B})^{n-1} > 0$$

And since $\gamma > 0$ then

$$\mathbf{A}^{n-1} \geqslant \gamma^{n-1}(\mathbf{I} + \mathbf{B})^{n-1} > 0$$

□

Lemma 4.2.36. *Let $\mathbf{A} = (a_{ij})$ be an $n \times n$ positive matrix. Then $\mathbf{A}$ is primitive.*

Proof. Since $\mathbf{A} > 0$ then by Theorem 4.2.25 $\rho_{\mathbf{A}}$ is a simple eigenvalue of $\mathbf{A}$, and so $\mathbf{A}$ is primitive.

□

Considering now nonnegative and irreducible matrices we will see that this kind of matrices are primitive if and only if there exists an integer $m \geqslant 1$ with $\mathbf{A}^m > 0$ as stated in Lemma 4.2.37.

Lemma 4.2.37. *Let $\mathbf{A} = (a_{ij})$ be an $n \times n$ nonnegative irreducible matrix. Then $\mathbf{A}$ is primitive if and only if there exists an integer $m \geqslant 1$ such that $\mathbf{A}^m > 0$.*

Proof. [19] Suppose that $\mathbf{A}$ is not primitive. Then there exist $k > 1$ eigenvalues of maximum modulus $\rho_{\mathbf{A}}$. And so $\mathbf{A}^m$ has k eigenvalues with modulus $(\rho_{\mathbf{A}})^m$. By the hypothesis of the Lemma we have that $\mathbf{A}^m > 0$ which contradicts Lemma 4.2.36, and so if $\mathbf{A}^m > 0$ then $\mathbf{A}$ is primitive. On the other hand if $\mathbf{A}$ is primitive then it has only one eigenvalue of maximum modulus and $\mathbf{A}$ is irreducible and so by Lemma 4.2.34 there exists a path in $G(\mathbf{A})$ from every n_i to n_j for $i, j = 1, \ldots, n$ and so there is a closed path starting and ending in each v_i. Observe that if $a_{ij} \cdot a_{jk} > 0$ then $(a^{(2)})_{ik} > 0$. Therefore if the path has length k then in $\mathbf{A}^k$, $(a^{(k)})_{ii} > 0$ for $i = 1, \ldots, n$. Using Lemma 4.2.35 we can see that $\mathbf{A}^m > 0$ for some m which completes the proof.

□

With Lemmas 4.2.35 and 4.2.37 we get the following result.

Lemma 4.2.38. *Let $\mathbf{A}$ be an $n \times n$ practical pairwise comparison matrix. Then $\mathbf{A}$ is primitive.*

Proof. By Lemma 4.2.37 $\mathbf{A}$ is primitive if and only if there exists an integer $m \geqslant 1$ such that $\mathbf{A}^m > 0$ and by Lemma 4.2.35 we know that such integer can be $n - 1$ and so $\mathbf{A}$ is primitive.

□

The limit of the normalised row sums of a primitive matrix gives us the eigenvector $\mathbf{w}$ corresponding to the largest eigenvalue λ_{max} as stated in Theorem 4.2.39.

[19] This proof follows the ideas of Varga (1962) page 41.

Theorem 4.2.39. *Let* $\mathbf{A}$ *be a* $n \times n$ *primitive matrix. Then*

$$\lim_{k \to \infty} \frac{\mathbf{A}^k \mathbf{e}}{\mathbf{e}^{\mathbf{T}} \mathbf{A}^k \mathbf{e}} = c\mathbf{w} \tag{4.75}$$

where $\mathbf{e} = (1, \ldots, 1)^{\mathbf{T}}$, $\mathbf{w}$ *is the principal eigenvector corresponding to the largest eigenvalue* λ_{max}, *and* c *is a constant.*

Proof. For a proof see Saaty (1990) page 178. □

Remark Given a practical pairwise comparison matrix the eigenvector $\mathbf{w}$, corresponding to the weights of the criteria, is calculated using equation (4.75).

Observe that so far we have yet not calculated the largest eigenvalue corresponding to the principal eigenvector.

Remark Given a matrix $\mathbf{A}$ and an eigenvector $\mathbf{w}$ of $\mathbf{A}$, then its corresponding eigenvalue is calculated by solving the system

$$\mathbf{A}\mathbf{w} = \lambda \mathbf{w} \tag{4.76}$$

4.3 Evaluation of the alternatives

We have seen that the aim of AHP is to make an informed decision according to a given choice of possibilities and so since we have a limited number of possibilities then the set of feasible solutions to the problem is *finite*. Let us define the set of possible solutions as the alternatives that we are interested in evaluating.

Definition (Set of alternatives) Given A_j as the jth alternative for $j = 1 \dots m$, the set of alternatives is defined as

$$A = \left\{ A_1, \dots, A_m \right\} \tag{4.77}$$

The selected alternatives are evaluated with respect to each criterion at every level in the hierarchy. We start the evaluation with the criteria in the lowest level and then we move up until we reach the second level of the hierarchy.

To start the evaluation in the lowest level of the hierarchy (level 3 in our example of Figure 4.1) we generate pairwise comparison matrices according to the performance of each alternative A_j for $j = 1, \dots, m$ in A with respect to a specific criterion C_i for $i = 1, \dots, n$ where n is the number of criteria to be evaluated at this particular level. Then we determine the eigenvector corresponding to the largest eigenvalue of each pairwise comparison matrix. This is done using the same methodology explained in the Section 4.2.

Observe that at this point we will have l pairwise comparison matrice one for each criterion C_i. And so the eigenvector of each pairwise comparison matrix will correspond to the evaluation of the alternative A_j on the criterion C_i. Then let x_{ij} for $i = 1, \dots, n$ and $j = 1, \dots, m$ be the evaluation of the alternative j on the criteria i. Then the value of x_{ij} will be given by the eigenvector of the corresponding pairwise comparison matrix. If we arrange all the values of x_{ij} in a matrix we get what is called a decision table. Such that each row of the decision table will corresponds to the eigenvector obtained by the corresponding pairwise comparison matrix.

Definition (Decision table) A decision table is an $n \times m$ matrix $\mathbf{X} = (x_{ij})$ whose entries x_{ij} correspond to the evaluation of the alternative A_j, for $j = 1, \ldots m$, with respect to the criterion C_i, for $i = 1, \ldots n$, i.e.

$$\begin{array}{c c} & \begin{array}{ccc} A_1 & \cdots & A_m \end{array} \\ \begin{array}{c} C_1 \\ \vdots \\ C_n \end{array} & \begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix} \end{array} \tag{4.78}$$

where $\sum_{j=1}^{n} x_{ij} = 1$ for $i = 1, \ldots, n$.

Observe that so far we have evaluated the alternatives only considering the lowest level of the hierarchy. And so in order to move up in the hierarchy we need a relation between one level in the hierarchy and an upper one. This relation is given by the weights of the criteria that were calculated in Section 4.2. Recall our example and observe that the weigths of the subcriterion will sum up to 1 for each major criteria, that is

Level 1			*Goal*				
Level 2	C_1	C_2	$\cdots$	C_i	$\cdots$	C_n	
	w_1	w_2		w_i		w_n	$\sum_{i=1}^{n} w_i = 1$
Level 3			$C_1^{(i)}$	$\cdots$	$C_n^{(i)}$		
			$w_1^{(i)}$		$w_n^{(i)}$		$\sum_{j=1}^{n} w_j^{(i)} = 1$

Therefore using the corresponding weight for each subcriterion $C_k^{(i)}$ within its corresponding criterion C_i the entries x_{ij} of a decision table in the next level of the hierarchy are then calculated by

$$x_{ij} = \sum_{k=1}^{n_i} x_{kj}^{(i)} w_k^{(i)} \qquad j = 1, \ldots, n \tag{4.79}$$

We continue this process until we reach the second level in the hierarchy and we get a final decision table like the one presented in the equation (4.78).

4.4 Ranking of the alternatives

The final decision of the problem is made by ranking the alternatives, which is ordering the alternatives according to a given score. Different approaches have been proposed to rank the alternatives (e.g. Rapcsák (2007)). We are going to use the *distributive method* which distributes the value of 1 between the criteria and alternatives with respect to their importance.

Definition (Distributive ranking) Given a decision table the ranking vector $\mathbf{x}$ of the alternatives using the distributive model is given by

$$x_j = \sum_{i=1}^{m} \frac{w_i}{w} \frac{x_{ij}}{\sum_{k=1}^{n} x_{ik}} \tag{4.80}$$

The vector $\mathbf{x}$ will provide the ranking of the alternatives, that is the score obtained by each alternative in the evaluation. Thus we made a final decision selecting the alternative which corresponds to the largest entry in the vector.

CHAPTER 5

PROMETHEE METHODOLOGY

We have seen that AHP methodology involves four main steps: (1) the modelling of the problem as a hierarchy, (2) the determination of the weights of the evaluation criteria, (3) the evaluation of the alternatives, and (4) the ranking of the alternatives to make the final decision. PROMETHEE methodology starts when the weights of the evaluation criteria have been calculated and focuses on steps (3) and (4). In this sense, PROMETHEE is a methodology to evaluate the alternatives with respect to the given criteria and to rank the alternatives to make a final decision. PROMETHEE methodology requires the completion of the following steps (Brans et al. (1986), Goumas and Lygerou (2000), and Rapcsák (2007)).

1. Establishment of the generalised criterion function
2. Determinations of the generalised criterion value
3. Establishment of the preference index
4. Ranking of the alternatives, better known in PROMETHEE as outranking of the alternatives

The concept of preference relations it is the basic idea of PROMETHEE, let us explain such concept. Recall the definition of the set of alternatives.

Definition (Set of alternatives) Given A_j as the jth alternative for $j = 1 \dots m$, the set of alternatives is defined as

$$A = \Big\{A_1, \dots, A_m\Big\} \tag{5.1}$$

Definition (Preference relation) The preference relation between A_k and A_l is a binary relation, represented by $A_k \succeq A_l$ where A_k is preferred over A_l and satisfy the following properties

(i) It is reflexive, i.e.

$$A_j \succeq A_j \qquad \text{for all } \ A_j \in A \tag{5.2}$$

(ii) It is transitive, i.e.

$$A_j \succeq A_k \ \text{ and } \ A_k \succeq A_l \Longrightarrow A_j \succeq A_l \qquad \text{for all } \ A_j, A_k, A_l \in A \tag{5.3}$$

(iii) It is complete, i.e.

$$A_j \succeq A_k \quad \text{or} \quad A_k \succeq A_j \qquad \text{for all } \ A_j, A_k \in A \tag{5.4}$$

We describe each step of the PROMETHEE methodology in Sections 5.1 to 5.4.

5.1 Criterion function

PROMETHEE methodology starts with the decision of how a specific criterion is going to be evaluated. This is done by defining a generalised criterion function that better represents the preference of the decision maker when comparing two alternatives on a specific criterion.

Definition (Generalised criterion function) The generalised criterion function, denoted by P, establishes the parameters of evaluation assigned to a specific criterion according with the preference of the decision maker. The range of this function is between zero and one, i.e. $P(d) \in [0, 1]$, where $d \in \mathbb{R}$.

In order to be completely defined, criterion functions generally require the definition of one or more parameters from among indifference threshold, strict preference threshold and standard deviation.

Definition (Indifference threshold) Denoted by q, the indifference threshold represents the value in the criterion function below which there is no preference between alternative $A_k \in A$ and alternative $A_l \in A$.

Definition (Strict preference threshold) Denoted by p, the strict preference threshold represents the value in the criterion function below which the alternative $A_k \in A$ it is preferred over the alternative $A_l \in A$.

Definition (Standard deviation) Given a set of data represented by a normal distribution, the standard deviation, denoted by σ, indicates how tightly the data in the normal distribution are clustered around the mean. To calculate the standard deviation, if we have x_i for $i = 1, \ldots, n$ data then we use

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n-1} \tag{5.5}$$

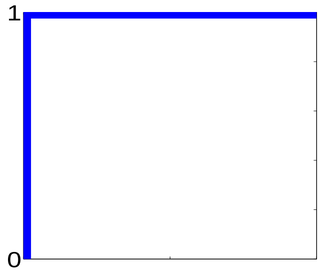
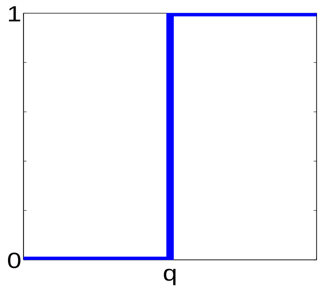
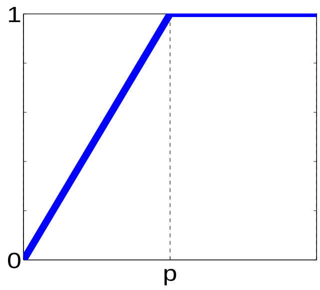
where μ corresponds to the mean of the data, i.e.

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{5.6}$$

Vincke and Brans (1985) proposed six basic types of criterion functions: (1) Simple criterion, (2) U-shaped criterion, (3) V-shaped criterion, (4) Step criterion, (5) Trapezoid criterion and (6) Gaussian criterion. In this project, considering the continuity of the data we will use only Gaussian criterion.[1] However we introduce the others only for completeness and as a reference of the existent types. A definition, the parameters required and a graphic representation of each of these functions is given in Tables 5.1 and 5.2. Observe in those tables that some of the functions are special cases of others (e.g. V-shaped is special case of trapezoid for $q = 0$).

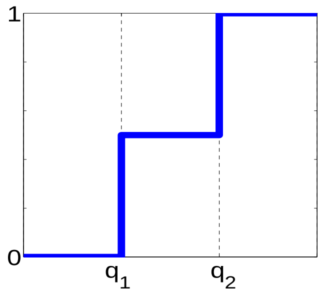
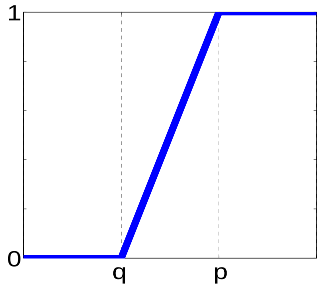
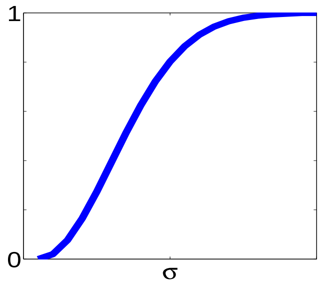
[1]According to Brans (1984) Gaussian criterion had been selected more by users for practical applications considering continue data.

Table 5.1: Generalised criterion functions

Function type	Definition	Graphic representation*
Simple	$P(d) = \begin{cases} 0 & d \leqslant 0, \\ 1 & d > 0. \end{cases}$ Parameters to define *none*	
U-shaped	$P(d) = \begin{cases} 0 & d \leqslant q, \\ 1 & d > q. \end{cases}$ Parameters to define q	
V-Shaped	$P(d) = \begin{cases} 0 & d \leqslant 0, \\ d/p & 0 \leqslant d \leqslant p, \\ 1 & d > p. \end{cases}$ Parameters to define p	

*The graphics are included as a reference of the shape only.

Table 5.2: Generalised criterion functions *(cont.)*

Function type	Definition	Graphic representation*
Step	$P(d) = \begin{cases} 0 & d \leqslant q_1, \\ 1/2 & q_1 < d \leqslant q_2, \\ 1 & d > q_2. \end{cases}$ Parameters to define $q_1,\ q_2$	
Trapezoid	$P(d) = \begin{cases} 0 & d \leqslant q, \\ \frac{d-q}{p-q} & q < d \leqslant p, \\ 1 & d > p. \end{cases}$ Parameters to define $p,\ q$	
Gaussian	$P(d) = \begin{cases} 0 & d \leqslant 0, \\ 1 - e^{-d^2/2\sigma^2} & d \geqslant 0. \end{cases}$ Parameters to define σ	

*The graphics are included as a reference of the shape only.

5.2 Criterion value

A comparison between the alternatives is made to establish the preference between pairs of alternatives with respect to each criterion. From the comparison, a number between 0 and 1, with 0 as no preference and 1 for strict preference, is assigned. In order to make this comparison we will again need the definition of a decision table.

Definition (Decision table) A decision table is an $n \times m$ matrix $\mathbf{X} = (x_{ij})$ whose entries x_{ij} correspond to the evaluation of the alternative A_j, for $j = 1, \ldots m$, with respect to the criterion C_i, for $i = 1, \ldots n$, i.e.

$$\begin{array}{c c} & \begin{array}{ccc} A_1 & \cdots & A_m \end{array} \\ \begin{array}{c} C_1 \\ \vdots \\ C_n \end{array} & \begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix} \end{array} \tag{5.7}$$

where $\sum_{j=1}^{n} x_{ij} = 1$ for $i = 1, \ldots, n$.

Assuming that a decision table has been given then we calculate a deviation of the alternatives based on pairwise comparison to express the deviation in the evaluation of the alternatives A_k and A_l according to the criterion C_i.

Definition (Deviation in the evaluation of the alternatives) Given a decision table represented by an $n \times m$ matrix $\mathbf{X} = (x_{ij})$, then a deviation of the alternatives is calculated using

$$d_i(A_k, A_l) = x_{ik} - x_{il} \tag{5.8}$$

where x_{ij} represents the entries as given in the decision table.

Having determined the deviations based on pairwise comparison, the criterion values for those deviations are calculated according with the criterion functions defined in Section 5.1.

Definition (Criterion values) The criterion value is denoted by $P_i(A_k, A_l)$ and it is the value given by the respective criterion function. It represents the preference given by the decision maker when comparing the alternative A_k with the alternative A_l on the ith criterion, as a function of the distance given by (5.8). Thus,

$$P_i(A_k, A_l) = P_i\big(d_i(A_k, A_l)\big) \qquad i = 1, \dots, n \tag{5.9}$$

where P_i is the chosen preference function of the criterion C_i for $i = 1, \dots, n$.

5.3 Preference index

A preference index is determined for each pair of alternatives, by taking the weighted average of the criterion values given for each alternative by using (5.9). The preference index takes into account the relative importance of each criterion.

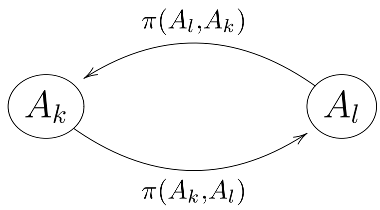
Definition (Preference index) The preference index of the alternative $A_k \in A$ over the alternative $A_l \in A$ is calculated by taking the weighted sum of $P_i(A_k, A_l)$, i.e.

$$\pi(A_k, A_l) = \sum_{i=1}^{n} w_i P_i(A_k, A_l) \qquad k, l = 1, \ldots, m \tag{5.10}$$

where w_i is the weight corresponding to the criterion $C_i \in C$ for $i = 1, \ldots, n$.

5.4 Ranking of the alternatives

The preference index $\pi(A_k, A_l)$, calculated in the last Section, provides information about the preference of the alternative A_k over the alternative A_l. However does not provide information about the disadvantage of the alternative A_k over the alternative A_l and that is why it is necessary to calculate $\pi(A_l, A_k)$ as well. But then the difference between $\pi(A_k, A_l)$ and $\pi(A_l, A_k)$ provides information about the mutual preference and the degree of difference (Rapcsák (2007)). This is known as the *outranking* relation, and graphically is represented by



Using the outranking relation every two alternatives can be compare and so an *out-*

ranking flow will provide information about how strong or weak is an alternative with respect to the others. We will explain the types of outranking flow in the next Section.

There are two steps to provide a ranking using PROMETHEE. The partial ranking and the complete ranking. The partial ranking provides a comparison only between the alternatives that are comparable with each other, therefore it may be incomplete, and the total ranking gives us a whole relation of the alternatives even if they are incomparable with each other.

Having done the partial ranking, if some of the alternatives are incomparable then it is necessary to do the complete ranking in order to make a final decision.

5.4.1 Partial ranking

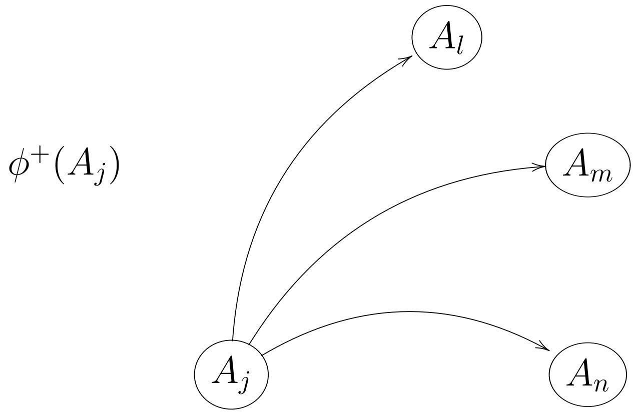
In this case the alternatives are ranked by considering a positive and a negative outranking flow.

Definition (Positive outranking flow) The positive flow indicates the preference of the alternative A_j with respect to all the other alternatives $A_k \neq A_j \in A$. It is given by

$$\phi^+(A_j) = \frac{1}{m-1} \sum_{k=1}^{m} \pi(A_j, A_k) \qquad j = 1, \dots, m \tag{5.11}$$

where m is the number of alternatives.

Graphically the positive outranking flow is represented as

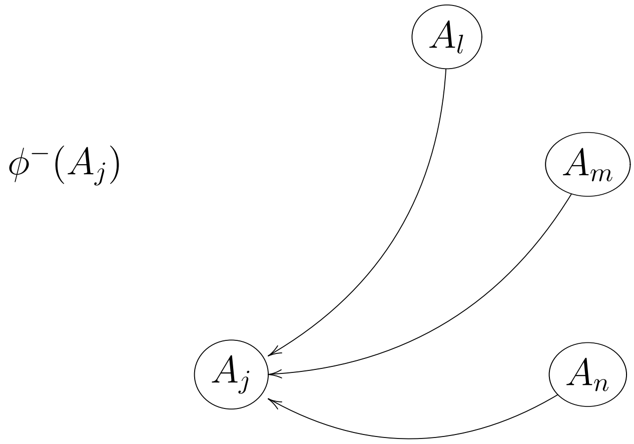


Definition (Negative outranking flow) The negative flow indicates the preference of the alternatives $A_k \in A$ versus alternative $A_j \neq A_k$. It is given by

$$\phi^-(A_j) = \frac{1}{m-1} \sum_{k=1}^{m} \pi(A_k, A_j) \qquad j = 1, \ldots, m \tag{5.12}$$

where m is the number of alternatives.

Graphically the negative outranking flow is represented as



Using the information provided by the positive and negative ranking we draw a graph according with a positive (leaving) and an negative (entering) flow for each alternative. Such flow is determined by the conditions established in Table 5.3.

Table 5.3: Relations between the alternatives for Partial Ranking in PROMETHEE

Preference relation	Cases	Representation in the graph
A_k is preferred to A_l	$\phi^+(A_k) > \phi^+(A_l)$ and $\phi^-(A_k) < \phi^-(A_l)$ $\phi^+(A_k) > \phi^+(A_l)$ and $\phi^-(A_k) = \phi^-(A_l)$ $\phi^+(A_k) = \phi^+(A_l)$ and $\phi^-(A_k) < \phi^-(A_l)$	$A_k \rightarrow A_l$
A_k is indifferent to A_l	$\phi^+(A_k) = \phi^+(A_l)$ and $\phi^-(A_k) = \phi^-(A_l)$	none
A_j is incomparable with A_k	otherwise	none

Having completed the graph of partial ranking a final decision is made by considering the alternative with more leaving flows which is the one that has been preferred more times when comparing with the other alternatives.

5.4.2 Complete ranking

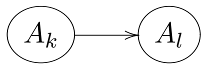
The complete ranking of the alternatives is given by the net flow.

Definition (Net flow) Given the positive ranking and the negative ranking of $A_j \in A$ the net flow is equal to

$$\phi(A_j) = \phi^+(A_j) - \phi^-(A_j) \qquad j = 1, \ldots, m \tag{5.13}$$

The complete ranking is then given by the conditions established in Table 5.4.

Table 5.4: Relations between the alternatives for Complete ranking in PROMETHEE

Preference relation	Cases	Representation in the graph
A_k is preferred to A_l	$\phi(A_k) > \phi(A_l)$	
A_k is indifferent to A_l	$\phi(A_k) = \phi(A_l)$	none

PART III

APPLICATION

CHAPTER 6

FORMULATING THE DECISION MODEL

6.1 Decision problem

We consider a situation where a decision has to be made to select a website which in terms of usability, provides better characteristics to the customers when buying books online. We can associate the situation with $f(\mathbf{x})$ where $\mathbf{x}$ represents the characteristics of a particular website. In this sense, a solution for $f(\mathbf{x})$ will be reached, when a possible solution $\mathbf{x}$, i.e. a particular website, provides the best possible characteristics. Therefore the goal of our decision problem is

$$\text{Maximise} \;\; f(\mathbf{x}) \tag{6.1}$$

However since usability, as explained in Chapter 3, comprises several criteria, the goal of the decision problem defined in (6.1) depends directly on several criteria that affect the performance of the website.

Let C_i be the ith usability criterion for $i = 1, \ldots, n$. Let us assume that every criterion C_i has an associated function f_i, for all $i = 1, \ldots, n$, and so we define

$$f(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_n(\mathbf{x}))^{\mathbf{T}} \tag{6.2}$$

In this situation, a solution for $f(\mathbf{x})$ will be reached if we can find an $\mathbf{x}$ that is the best possible for every $f_i(\mathbf{x})$ and so the problem defined in (6.1) can be considered as a *multicriteria optimisation problem.* If we can find a solution that optimises every criterion simultaneously that will be an ideal solution for this problem. However if we obtain instead a solution that let us make an informed decision then such solution will be satisfactory and therefore accepted.

6.2 Evaluation criteria

Considering those usability criteria explained in Section 3, for $n = 7$, we define the set of evaluation criteria C as

$$C = \left\{C_1, \ldots, C_{n=7}\right\} \tag{6.3}$$

where the C_i are defined as follows

C_1: Accessibility

C_2: Customisation and personalisation

C_3: Download speed

C_4: Ease of use

C_5: Errors

C_6: Navigation

C_7: Site content

Also note that each C_i is a criterion in the second level of the hierarchy and each of them is comprised by further subcriteria. Then let $C_j^{(i)}$ be the jth subcriterion of the criterion C_i, and so let the $C_j^{(i)}$ be defined as follows

C_1: Accessibility

- $C_1^{(1)}$: Availability to different agents
- $C_2^{(1)}$: Alternatives for multimedia presentations
- $C_3^{(1)}$: Readability
- $C_4^{(1)}$: Frames identification
- $C_5^{(1)}$: Skip-navigation links

C_2: Customisation and personalisation

- $C_1^{(2)}$: Possibility of connection with other people
- $C_2^{(2)}$: Personalisation
- $C_3^{(2)}$: Refinement and addition of content over time
- $C_4^{(2)}$: Market research

C_3: Download speed

- $C_1^{(3)}$: Simple and meaningful use of graphics and tables
- $C_2^{(3)}$: Limited use of animation
- $C_3^{(3)}$: Use of thumbnails

C_4: Ease of use

- $C_1^{(4)}$: Goals (prioritisation of the content)
- $C_2^{(4)}$: Structure of the website
- $C_3^{(4)}$: Feedback about the system status

C_5: Errors

- $C_1^{(5)}$: Number of errors
- $C_2^{(5)}$: Severity of the errors

$C_3^{(5)}$: Recovering from errors

C_6: Navigation

$C_1^{(6)}$: Organisation

$C_2^{(6)}$: Arrangement

$C_3^{(6)}$: Layout

$C_4^{(6)}$: Sequencing

C_7: Site content

$C_1^{(7)}$: Amount and variety of product information

$C_2^{(7)}$: Relevance of the content (useful)

$C_3^{(7)}$: Use of media (to make content attractive)

$C_4^{(7)}$: Appropriate content (depth and breadth)

$C_5^{(7)}$: Timely / current information

6.3 Alternatives

In this case study we will consider five different websites from which one can buy books online. We have selected the most commonly available websites. These websites correspond to the alternatives of our multicriteria decision problem. In Tables 6.1 to 6.5 the main characteristics of the alternatives are given.

Let us define A_j as the jth alternative for $j = 1, \ldots, m$. Then we have $m = 5$, the set of alternatives A in our case is given by

$$A = \Big\{A_1, \ldots, A_{m=5}\Big\} \tag{6.4}$$

where

A_1: Amazon (www.amazon.co.uk)

A_2: Blackwell (www.blackwell.co.uk)

A_3: Bookstore (www.bookstore.co.uk)

A_4: Borders (www.borders.com)

A_5: Waterstone's (www.waterstones.com)

Table 6.1: Alternative 1: Amazon

www.amazon.co.uk

General remarks	Amazon.co.uk is a trading name for Amazon EU S.à.r.l., Amazon Services Europe S.à.r.l. and Amazon Media EU S.à.r.l. All three are wholly owned subsidiaries of global online retailer Amazon.com.
Brief history	Amazon.co.uk has its origins in an independent online store, Bookpages, which was established in 1996 and acquired by Amazon.com in 1998. Amazon.co.uk opened its virtual doors in October 1998.
Other stores	The Amazon group also has online stores in the United States, Germany, France, Japan, China and Canada.
Source of information	(Amazon (1998)).

Table 6.2: Alternative 2: Blackwell

www.blackwell.co.uk

General remarks	Blackwell.co.uk is an academic bookseller.
Brief history	By the 1960s, Blackwell had built an international reputation for bookselling excellence with links to academic institutions and libraries around the globe. In 1995, www.blackwell.co.uk became the first transactional online bookstore in the UK, giving people across the world access to over 150,000 titles.
Other stores	Blackwell has over 60 outlets across England, Scotland and Wales.
Source of information	(Blackwell (2009)).

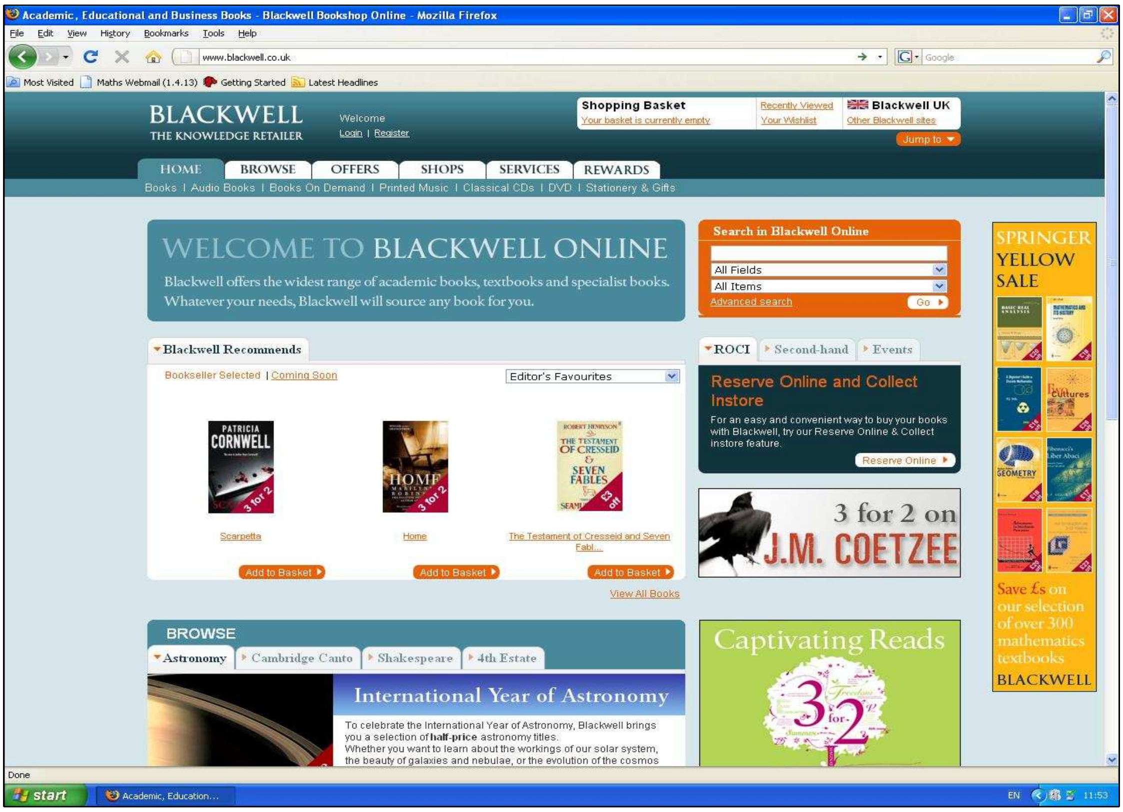


Table 6.3: Alternative 3: Bookstore

www.bookstore.co.uk

General remarks	Bookstore.co.uk Limited is an online retailer of books published in the UK and the USA.
Brief history	Bookstore have been delivering service to worldwide customers for over five years.
Source of information	(Bookstore (2009)).

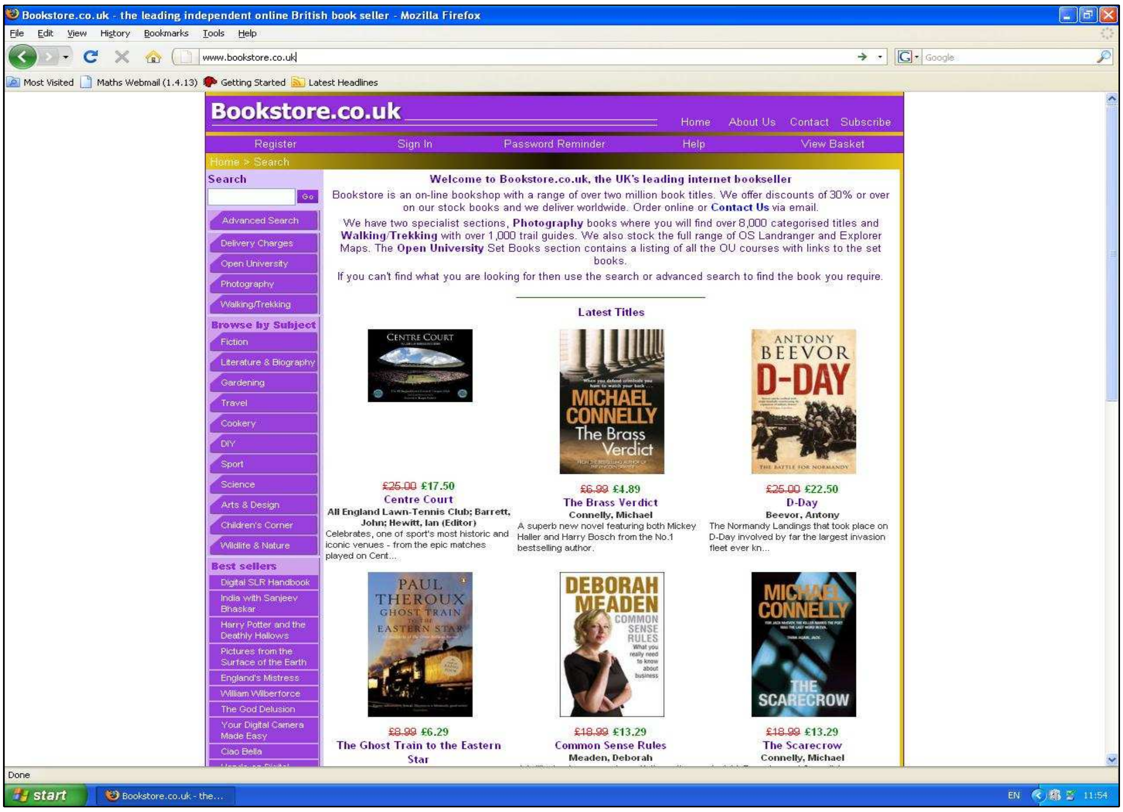


Table 6.4: Alternative 4: Borders

www.borders.com

General remarks	Throughout more than 1,000 stores, headquartered in Ann Arbor, Michigan, Borders Group, Inc., is a publicly held company.
Brief history	In 2008 the company launched Borders.com for online shopping.
Other stores	Borders Group operates over 515 Borders superstores in the U.S.; three stores in Puerto Rico; and approximately 377 stores in the Waldenbooks Specialty Retail segment, including Waldenbooks, Borders Express, Borders airport stores, and Borders Outlet.
Source of information	(Borders (2009)).

Table 6.5: Alternative 5: Waterstone's

www.waterstones.com

General remarks	Part of HMV Group.
Brief history	Waterstone's was founded by Tim Waterstone in 1982. WHSmith took a share in Waterstone's in 1989, and in 1998 HMV Media (now HMV Group plc) acquired Waterstone's, having already acquired the Dillon's chain. In 1999 HMV re-branded its portfolio of Dillon's stores to Waterstone's.
Other stores	Currently trades from more than 300 stores in the UK, Republic of Ireland and continental Europe (Brussels and Amsterdam) as well as on the Isle of Man, Jersey and the Isle of Wight.
Source of information	(Waterstone's (2009)).

CHAPTER 7

SOLVING THE DECISION MODEL

7.1 Using Analytic Hierarchy Process

We will use the methodology that was explained in Chapter 4.

7.1.1 Step 1: Decision hierarchy

Having defined the goal of the decision model and selected the objective functions, the construction of the hierarchy structure is straightforward. One starts by placing the goal f in the highest level of a multilevel tree, then the objective functions C_i are placed in the second level, and the subcriteria $C_j^{(i)}$ are situated in lowest level of the tree, for all i, j. The hierarchy for the decision model studied here is represented in Figure 7.1.

Nevertheless, the pairwise comparison will take place between criteria on the second level of the hierarchy only. The alternatives will be evaluated at every criterion in the third level of the hierarchy. The values of the alternatives with respect to the second level of the hierarchy will be computed from the values of the third level and their weights. In this situation the hierarchy of the decision model, including the alternatives can be

Table 7.1: Hierarchy of the decision model

Level 1	*Usability of websites*						
Level 2	*Accessibility*	*Customisation/ personalisation*	*Download speed*	*Ease of use*	*Errors*	*Navigation*	*Site content*
Level 3	Availability	Connection with people	Graphics and tables	Goals	Number of errors	Organisation	Amount and variety
	Multimedia presentations	Personalisation	Use of animation	Structure	Severity of errors	Arrangement	Relevance of the content
	Readability	Refinement	Use of thumbnails	Feedback	Recovering of errors	Layout	Use of media
	Frames identification	Market research				Sequencing	Appropriate content
	Skip-Navigation links						Timely/current information

represented as shown in Figure 7.2.

7.1.2 Step 2: Weights of the criteria

In order to find a reasonable value to represent the weights of the criteria based on their importance, the opinion of 6 anonymous experts in the area of website design, employees of worldwide publicity agency, was taken under consideration.[1] Thus let $\mathbf{D}_k$ be the kth decision maker for $k = 1, \ldots, l$ and $l = 6$.

Now recall the definition of practical pairwise comparison matrices as stated in Section 4.2.2 and let $\mathbf{P}$ be an $n \times n$ matrix, with the entries p_{ij} representing the subjective judgement given by the experts when considering the criterion C_i and the criteria C_j, for $i, j = 1, \ldots, n$, according to

$$p_{ij} = \begin{cases} \omega_{ij} & \text{if } C_i \text{ is more important than } C_j, \\ 1/\omega_{ij} & \text{if } C_j \text{ is more important than } C_i. \end{cases}$$

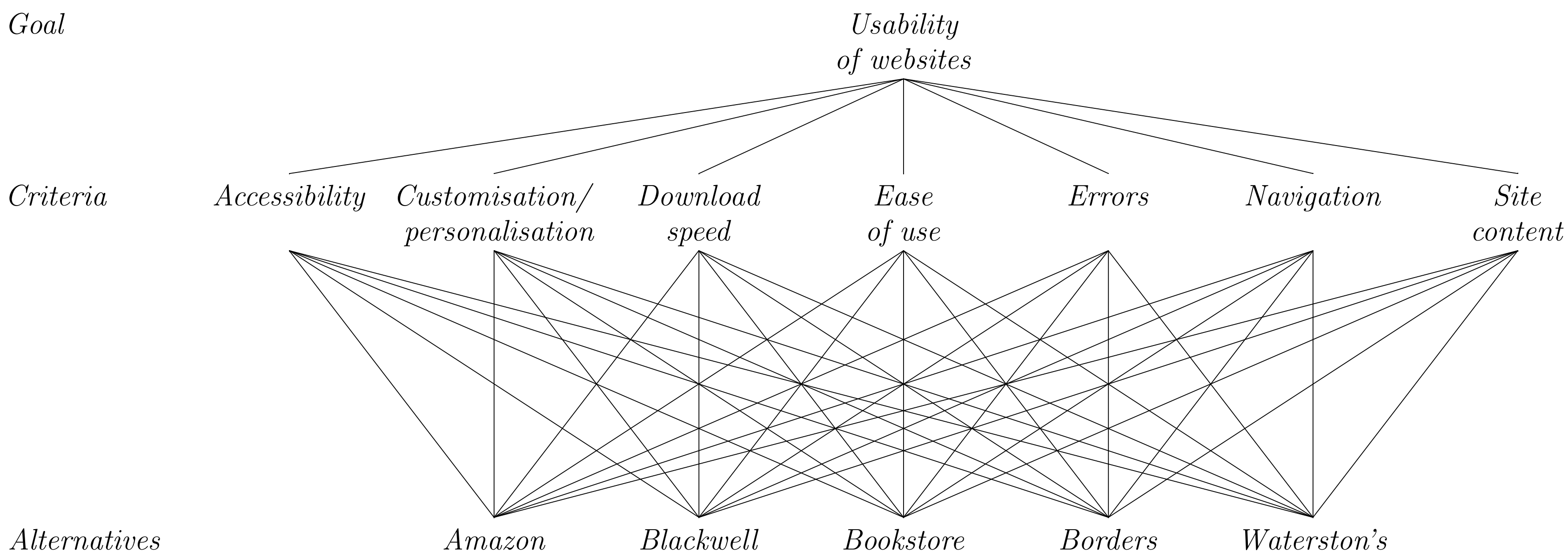
where ω_{ij} is the intensity of importance evaluated on a scale of $1 - 9$, with 1 being equal importance and 9 indicating that one is much more important than the other.[2]

Thus, based on their experience, the selected experts compared every criterion with each other and from this comparison we construct $l = 6$ pairwise comparison matrices, one for each expert $\mathbf{D}_k$, $k = 1, \ldots, l = 6$ as indicated in the matrices (7.1) to (7.6).

[1] The name of the company is omitted as a request of the experts.

[2] This is the scale comparison used in AHP models (Saaty (1980)), and it is justified by human psychological behaviour when taking subjective decisions.

Table 7.2: Hierarchy of the decision model including the alternatives



$$
\mathbf{D}_1 = \begin{array}{c} \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \end{array}
\begin{array}{c} \begin{array}{ccccccc} C_1 & C_2 & C_3 & C_4 & C_5 & C_6 & C_7 \end{array} \\
\begin{pmatrix}
1 & \frac{1}{9} & \frac{1}{5} & \frac{1}{9} & \frac{1}{7} & \frac{1}{5} & \frac{1}{5} \\
9 & 1 & 7 & 5 & 1 & 3 & 1 \\
5 & \frac{1}{7} & 1 & 3 & \frac{1}{3} & 1 & \frac{1}{7} \\
9 & \frac{1}{5} & \frac{1}{3} & 1 & \frac{1}{5} & 1 & \frac{1}{3} \\
7 & 1 & 3 & 5 & 1 & 3 & 1 \\
5 & \frac{1}{3} & 1 & 1 & \frac{1}{3} & 1 & \frac{1}{3} \\
5 & 1 & 7 & 3 & 1 & 3 & 1
\end{pmatrix} \end{array} \tag{7.1}
$$

$$
\mathbf{D}_2 = \begin{array}{c} \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \end{array}
\begin{array}{c} \begin{array}{ccccccc} C_1 & C_2 & C_3 & C_4 & C_5 & C_6 & C_7 \end{array} \\
\begin{pmatrix}
1 & \frac{1}{3} & 2 & 3 & 7 & 5 & 7 \\
3 & 1 & 1 & 3 & 9 & 9 & 9 \\
\frac{1}{2} & 1 & 1 & 3 & 7 & 7 & 5 \\
\frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 1 & 7 & 5 & 7 \\
\frac{1}{7} & \frac{1}{9} & \frac{1}{7} & \frac{1}{7} & 1 & 1 & \frac{1}{2} \\
\frac{1}{5} & \frac{1}{9} & \frac{1}{7} & \frac{1}{5} & 1 & 1 & 3 \\
\frac{1}{7} & \frac{1}{9} & \frac{1}{5} & \frac{1}{7} & 2 & \frac{1}{3} & 1
\end{pmatrix} \end{array} \tag{7.2}
$$

$$
\mathbf{D}_3 = \begin{array}{c} \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \end{array}
\begin{array}{c} \begin{array}{ccccccc} C_1 & C_2 & C_3 & C_4 & C_5 & C_6 & C_7 \end{array} \\
\begin{pmatrix}
1 & \frac{1}{3} & 2 & 3 & 7 & 5 & 7 \\
3 & 1 & 3 & 5 & 9 & 9 & 9 \\
\frac{1}{2} & \frac{1}{3} & 1 & 3 & 7 & 7 & 5 \\
\frac{1}{3} & \frac{1}{5} & \frac{1}{3} & 1 & 7 & 5 & 7 \\
\frac{1}{7} & \frac{1}{9} & \frac{1}{7} & \frac{1}{7} & 1 & 1 & \frac{1}{2} \\
\frac{1}{5} & \frac{1}{9} & \frac{1}{7} & \frac{1}{5} & 1 & 1 & 3 \\
\frac{1}{7} & \frac{1}{9} & \frac{1}{5} & \frac{1}{7} & 2 & \frac{1}{3} & 1
\end{pmatrix} \end{array} \tag{7.3}
$$

$$
\mathbf{D}_4 = \begin{array}{c} \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \end{array}
\begin{array}{c} \begin{array}{ccccccc} C_1 & C_2 & C_3 & C_4 & C_5 & C_6 & C_7 \end{array} \\
\begin{pmatrix}
1 & 5 & 3 & 3 & 5 & 3 & 3 \\
\frac{1}{5} & 1 & 2 & 1 & \frac{1}{2} & \frac{1}{5} & 1 \\
\frac{1}{3} & \frac{1}{2} & 1 & 3 & 1 & \frac{1}{3} & \frac{1}{3} \\
\frac{1}{3} & 1 & \frac{1}{3} & 1 & 2 & \frac{1}{3} & \frac{1}{3} \\
\frac{1}{5} & 2 & 1 & \frac{1}{2} & 1 & \frac{1}{3} & \frac{1}{3} \\
\frac{1}{3} & 5 & 3 & 3 & 3 & 1 & 1 \\
\frac{1}{3} & 1 & 3 & 3 & 3 & 1 & 1
\end{pmatrix} \end{array} \tag{7.4}
$$

$$\mathbf{D}_5 = \begin{array}{c} \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \end{array} \begin{array}{c} \begin{array}{ccccccc} C_1 & C_2 & C_3 & C_4 & C_5 & C_6 & C_7 \end{array} \\ \begin{pmatrix} 1 & 3 & 1 & 3 & 3 & 2 & 2 \\ \frac{1}{3} & 1 & \frac{1}{3} & \frac{1}{2} & 2 & \frac{1}{3} & \frac{1}{3} \\ 1 & 3 & 1 & 3 & 3 & 1 & 1 \\ \frac{1}{3} & 2 & \frac{1}{3} & 1 & 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & 3 & \frac{1}{3} & 1 & 1 & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{2} & 3 & 1 & 2 & 5 & 1 & 1 \\ \frac{1}{2} & 3 & \frac{1}{3} & 1 & 3 & \frac{1}{2} & 1 \end{pmatrix} \end{array} \tag{7.5}$$

$$\mathbf{D}_6 = \begin{array}{c} \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \end{array} \begin{array}{c} \begin{array}{ccccccc} C_1 & C_2 & C_3 & C_4 & C_5 & C_6 & C_7 \end{array} \\ \begin{pmatrix} 1 & 7 & 1 & 2 & 9 & 1 & 1 \\ \frac{1}{7} & 1 & \frac{1}{5} & \frac{1}{5} & 2 & \frac{1}{5} & \frac{1}{7} \\ 1 & 5 & 1 & 2 & 9 & 1 & 1 \\ \frac{1}{2} & 5 & \frac{1}{2} & 1 & 7 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{9} & \frac{1}{2} & \frac{1}{9} & \frac{1}{7} & 1 & \frac{1}{9} & \frac{1}{5} \\ 1 & 5 & 1 & 2 & 9 & 1 & 1 \\ 1 & 7 & 1 & 2 & 5 & 1 & 1 \end{pmatrix} \end{array} \tag{7.6}$$

So far we have defined $l = 6$ practical pairwise comparison matrices $\mathbf{P}_{n\times n}$, composed of the elements p_{ij} for $i, j = 1, \ldots, n$ and $n = 7$. Now we need to aggregate somehow those individual judgements (i.e., p_{ij}), corresponding to each pairwise comparison matrix, into a single judgement.

To do this we use the methodology presented by Aczél and Saaty (1983). In their work they proved that, given $l \geqslant 2$ decision makers with individual judgements p_{ij}^k for $k = 1, \ldots, l$, we can find a synthesising function $f(p_{ij}^1, \ldots, p_{ij}^l)$, corresponding to the aggregate value of the single judgements, by using the geometric mean

$$f(p_{ij}^1, \ldots, p_{ij}^l) = \prod_{k=1}^{l} (p_{ij}^k)^{1/l} \qquad i, j = 1, \ldots, n \tag{7.7}$$

Since we have $l = 6 > 2$ decision makers, then we can use equation (7.7). Let $\mathbf{A}$ be an $n \times n$ aggregate matrix, with the entries a_{ij} representing the function $f(p_{ij}^1, \ldots, p_{ij}^l)$. The resultant matrix, using (7.7), is represented in (7.8). Observe that the values given in the matrix (7.8) are only accurate to 2 decimal positions.

$$\mathbf{A} = \begin{array}{c} \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \end{array} \begin{array}{c} \begin{array}{ccccccc} C_1 & C_2 & C_3 & C_4 & C_5 & C_6 & C_7 \end{array} \\ \left(\begin{array}{ccccccc} 1 & 1.25 & 1.51 & 1.79 & 3.61 & 3.01 & 2.15 \\ 0.80 & 1 & 1.03 & 0.89 & 2.33 & 1.40 & 1.25 \\ 0.66 & 0.97 & 1 & 1.79 & 2.76 & 2.08 & 1.42 \\ 0.56 & 1.12 & 0.56 & 1 & 3.57 & 1.63 & 1.70 \\ 0.28 & 0.43 & 0.36 & 0.28 & 1 & 0.58 & 0.42 \\ 0.33 & 0.71 & 0.48 & 0.61 & 1.73 & 1 & 1.24 \\ 0.47 & 0.80 & 0.70 & 0.59 & 2.38 & 0.81 & 1 \end{array} \right) \end{array} \tag{7.8}$$

Now let $\mathbf{w}$ be the vector representing the weights of the criteria. Then $\mathbf{w}$ is calculated by using Theorem 4.2.39, specifically using

$$\lim_{k\to\infty} \frac{\mathbf{A}^k\mathbf{e}}{\mathbf{e}^{\mathbf{T}}\mathbf{A}^k\mathbf{e}} = c\mathbf{w}$$

where $\mathbf{e} = (1, \ldots, 1)^{\mathbf{T}}$, $\mathbf{w}$ is the principal eigenvector corresponding to the largest eigenvalue λ_{max}, and c is a constant.

Having calculated the principal eigenvector $\mathbf{w}$, the largest eigenvalue λ_{max} is calculated by using (4.76), i.e.

$$\mathbf{A}\mathbf{w} = \lambda_{max}\mathbf{w}$$

The consistency index CI of the generated matrix is then calculated using (4.42), and so

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

where λ_{max} is the largest eigenvalue of $\mathbf{A}$ and n is the number of rows.

Having calculated a consistency index CI by using Equation (4.42) Saaty (1990) suggests that it needs to be compared with a random consistency index RI. This index is calculated by generating reciprocal matrices randomly using a scale from 1 to 9 and with reciprocity forced. The average consistency index for a sample of 500 random matrices according with Saaty (1990) can be seen in Table 7.3.

The proposed comparison of consistent indexes is called *consistency ratio* CR and is calculated by using

$$CR = \frac{CI}{RI} \tag{7.9}$$

Table 7.3: Random Consistency Index RI

n	1	2	3	4	5	6	7	8	9	10
RI	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

According with Saaty (1990) if CR is less than or equal to 10% then the practical pairwise comparison matrix is consistent enough. Otherwise the judgements need to be revised. In our case we use (7.9) with $RI = 1.32$ according with Table (7.3) for $n = 7$.

The results, including the weights of the criteria $\mathbf{w}$, largest eigenvalue λ_{max}, consistency index CI and consistency ratio CR, are presented in Table 7.4.

Table 7.4: Weights of the criterion $C_i \in C$ using AHP methodology with $l = 6$ decision makers.

$l = 6$	C_1	C_2	C_3	C_4	C_5	C_6	C_7	Weights* $\mathbf{w}$
C_1	1.00	1.25	1.51	1.79	3.61	3.01	2.15	0.24
C_2	0.80	1.00	1.03	0.89	2.33	1.40	1.25	0.15
C_3	0.66	0.97	1.00	1.79	2.76	2.08	1.42	0.18
C_4	0.56	1.12	0.56	1.00	3.57	1.63	1.70	0.16
C_5	0.28	0.43	0.36	0.28	1.00	0.58	0.42	0.06
C_6	0.33	0.71	0.48	0.61	1.73	1.00	1.24	0.10
C_7	0.47	0.80	0.70	0.59	2.38	0.81	1.00	0.11

$\lambda_{max} = 7.11$ $CI = 0.02$ $CR = 0.01$

*The values are only accurate to 2 decimal positions.

Therefore the vector $\mathbf{w}$ representing the weights of the criteria in the second level of the hierarchy is given by

$$\mathbf{w} = \begin{pmatrix} 0.24 & 0.15 & 0.18 & 0.16 & 0.06 & 0.10 & 0.11 \end{pmatrix}^{\mathbf{T}} \tag{7.10}$$

Graphically the weights of the criteria are presented in Figure 7.1.

Now to find the weights of the criteria in the third level of the hierarchy we could have

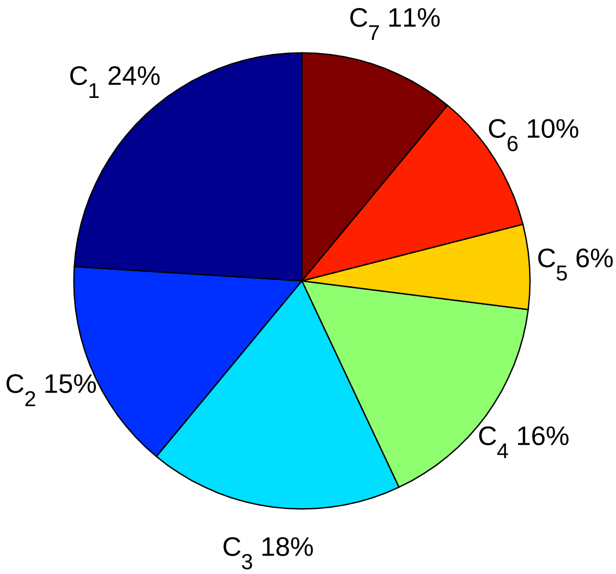


Figure 7.1: Weights of the criteria using AHP methodology

compared every subcriterion with each other, and applied the same methodology as in the level of the criteria C_i. However, as mentioned in Section 7.1, the pairwise comparison of the criteria was carried out at the second level of the hierarchy (i.e., major criteria) only.

Thus the relation between the second and third level of the tree (i.e., the weights of the subcriteria) is obtained by assigning a value to each subcriteria $C_j^{(i)}$, between 1 and 100, according to their importance within the corresponding major criterion C_i, as an indicator of the level of influence inside the criterion. The sum of the values of the subcriteria $C_j^{(i)}$ corresponding to the ith criterion is equal to 100 for each i. The values assigned for $l = 6$ decision makers, as a relation between the second and third level of the hierarchy, are presented in Table 7.5.

Having assigned the values for the decision makers we calculate the final weight of each subcriterion $C_j^{(i)}$ by taking the average. Observe that we do not use the methodology presented by Aczél and Saaty (1983) because that methodology is used when dealing with pairwise comparison matrices. The weights of the criteria are summarised in Table 7.6.

Table 7.5: Importance of the subcriterion $C_j^{(i)}$ within the criterion $C_i \in C$

1. Accessibility	1	2	3	4	5	6
Availability to different agents	10	15	30	20	15	20
Alternatives for multimedia presentations	30	35	30	20	20	20
Readability	40	30	20	30	25	30
Frames identification	15	15	10	10	15	20
Skip-Navigation Links	5	5	10	20	25	10
	100	100	100	100	100	100

2. Customisation and Personalisation	1	2	3	4	5	6
Possibility of connection with other people	30	15	20	40	30	15
Personalisation	30	25	30	20	20	25
Refinement and addition of content over time	30	40	40	20	35	35
Market research	10	20	10	20	15	25
	100	100	100	100	100	100

3. Download speed	1	2	3	4	5	6
Simple and meaningful use of graphics and tables	20	35	50	60	40	40
Limited use of animation	60	35	20	30	40	30
Use of thumbnails	20	30	30	10	20	30
	100	100	100	100	100	100

4. Ease of use	1	2	3	4	5	6
Goals (prioritisation of the content)	40	30	30	20	20	30
Structure of the website	30	40	50	60	45	40
Feedback about the system status items	30	30	20	20	35	30
	100	100	100	100	100	100

5. Errors	1	2	3	4	5	6
Number of errors	30	10	30	30	30	0
Severity of the errors	40	10	40	30	30	0
Recovering from errors	30	80	30	40	40	100
	100	100	100	100	100	100

6. Navigation	1	2	3	4	5	6
Organisation	30	40	20	30	20	30
Arrangement	15	10	40	20	20	20
Layout	40	30	20	30	35	30
Sequencing	15	20	20	20	25	20
	100	100	100	100	100	100

7. Site Content	1	2	3	4	5	6
Amount and variety of product information	10	15	10	10	20	15
Relevance of the content	10	30	30	10	20	25
Use of media	30	20	30	20	10	20
Appropriate content	20	15	10	30	30	15
Timely / current information	30	20	20	30	20	25
	100	100	100	100	100	100

Table 7.6: Weights of the criteria at all levels of the hierarchy

C_1: Accessibility		24%
$C_1^{(1)}$: Availability to different agents	18%	
$C_2^{(1)}$: Alternatives for multimedia presentations	26%	
$C_3^{(1)}$: Readability	29%	
$C_4^{(1)}$: Frames identification	14%	
$C_5^{(1)}$: Skip-Navigation Links	13%	
C_2: Customisation and Personalisation		15%
$C_1^{(2)}$: Possibility of connection with other people	25%	
$C_2^{(2)}$: Personalisation	25%	
$C_3^{(2)}$: Refinement and addition of content over time	33%	
$C_4^{(2)}$: Market research	17%	
C_3: Download speed		18%
$C_1^{(3)}$: Simple and meaningful use of graphics and tables	41%	
$C_2^{(3)}$: Limited use of animation	36%	
$C_3^{(3)}$: Use of thumbnails	23%	
C_4: Ease of use		16%
$C_1^{(4)}$: Goals (prioritisation of the content)	28%	
$C_2^{(4)}$: Structure of the website	44%	
$C_3^{(4)}$: Feedback about the system status items	28%	
C_5: Errors		6%
$C_1^{(5)}$: Number of errors	22%	
$C_2^{(5)}$: Severity of the errors	25%	
$C_3^{(5)}$: Recovering from errors	53%	
C_6: Navigation		10%
$C_1^{(6)}$: Organisation	28%	
$C_2^{(6)}$: Arrangement	21%	
$C_3^{(6)}$: Layout	31%	
$C_4^{(6)}$: Sequencing	20%	
C_7: Site Content		11%
$C_1^{(7)}$: Amount and variety of product information	13%	
$C_2^{(7)}$: Relevance of the content	21%	
$C_3^{(7)}$: Use of media	22%	
$C_4^{(7)}$: Appropriate content	20%	
$C_5^{(7)}$: Timely / current information	24%	

7.1.3 Step 3: Evaluation of alternatives

The selected alternatives A_j, $j = 1, \ldots, m$, are evaluated at every criterion in the lowest level of the hierarchy (i.e., $C_j^{(i)}$) by using the eigenvector method. Since the choice of considering one or more decision makers depends on their experience, if one of the experts has more experience and knowledge than the others then it is acceptable to have only one decision maker. To evaluate the alternatives we consider the opinion of one anonymous expert employee of the Leo Burnett company. The expert evaluated the selected alternatives by comparisons. The resultant pairwise comparison matrices, including the eigenvector, the largest eigenvalue, the consistency index and the consistency ratio are summarised in Tables 7.7 to 7.13.

Observe that in the Tables 7.7 to 7.13, the alternatives have been evaluated only on each criterion in the third level of the hierarchy (i.e., $C_j^{(i)}$). To calculate the values of the alternatives with respect to the criteria in the second level of the hierarchy (i.e. the criteria C_i) we use the data collected in Table 7.6 which provides a relation between the second and third level of the hierarchy. And so we apply the equation (4.79), i.e.

$$x_{ij} = \sum_{k=1}^{n_i} x_{kj}^{(i)} w_k^{(i)} \qquad j = 1, \ldots, n$$

Finally, arranging the existing data (i.e., the weights of the criteria and the alternatives) in a decision matrix, the decision table represented in the matrix (7.11) is obtained.

Table 7.7: Evaluation of the alternative $A_j \in A$ on each subcriterion of Accessibility

Availability	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	5	7	3	1/3	0.26
A_2	1/5	1	3	1/3	1/7	0.06
A_3	1/7	1/3	1	1/5	1/9	0.03
A_4	1/3	3	5	1	1/5	0.13
A_5	3	7	9	5	1	0.51
	$\lambda_{max} = 5.24$		$CI = 0.06$		$CR = 0.05$	

Multimedia presentations	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	3	5	2	1/2	0.27
A_2	1/3	1	3	1/3	1/3	0.10
A_3	1/5	1/3	1	1/5	1/7	0.04
A_4	1/2	3	5	1	1/2	0.21
A_5	2	3	7	2	1	0.38
	$\lambda_{max} = 5.14$		$CI = 0.03$		$CR = 0.03$	

Readability	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	1/5	3	1	1/5	0.09
A_2	5	1	7	5	1	0.39
A_3	1/3	1/7	1	1/3	1/7	0.04
A_4	1	1/5	3	1	1/5	0.09
A_5	5	1	7	5	1	0.39
	$\lambda_{max} = 5.09$		$CI = 0.02$		$CR = 0.02$	

Frames identification	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	1/4	1/2	1/3	1/5	0.06
A_2	4	1	3	3	2	0.38
A_3	2	1/3	1	1/2	1/3	0.10
A_4	3	1/3	2	1	1/4	0.14
A_5	5	1/2	3	4	1	0.32
	$\lambda_{max} = 5.22$		$CI = 0.06$		$CR = 0.05$	

Skip-Navigation links	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	2	3	5	1	0.32
A_2	1/2	1	2	5	1/2	0.20
A_3	1/3	1/2	1	3	1/3	0.12
A_4	1/5	1/5	1/3	1	1/3	0.06
A_5	1	2	3	3	1	0.30
	$\lambda_{max} = 5.16$		$CI = 0.04$		$CR = 0.04$	

* The values are only accurate to 2 decimal positions.

Table 7.8: Evaluation of the alternative $A_j \in A$ on each subcriterion of Customisation and Personalisation

Connection with people	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	3	3	7	1/3	0.25
A_2	1/3	1	2	5	1/5	0.13
A_3	1/3	1/2	1	3	1/5	0.08
A_4	1/7	1/5	1/3	1	1/9	0.03
A_5	3	5	5	9	1	0.51

$\lambda_{max} = 5.18$ $CI = 0.04$ $CR = 0.04$

Personalisation	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	1	9	2	2	0.31
A_2	1	1	9	3	2	0.34
A_3	1/9	1/9	1	1/5	1/7	0.03
A_4	1/2	1/3	5	1	2	0.17
A_5	1/2	1/2	7	1/2	1	0.15

$\lambda_{max} = 5.15$ $CI = 0.04$ $CR = 0.03$

Refinement	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	1	5	3	2	0.32
A_2	1	1	3	2	3	0.30
A_3	1/5	1/3	1	1/5	1/5	0.05
A_4	1/3	1/2	5	1	1/2	0.14
A_5	1/2	1/3	5	2	1	0.19

$\lambda_{max} = 5.34$ $CI = 0.08$ $CR = 0.08$

Market research	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	1/5	7	1/3	3	0.15
A_2	5	1	9	1	5	0.40
A_3	1/7	1/9	1	1/9	1/5	0.03
A_4	3	1	9	1	5	0.35
A_5	1/3	1/5	5	1/5	1	0.08

$\lambda_{max} = 5.29$ $CI = 0.07$ $CR = 0.06$

* The values are only accurate to 2 decimal positions.

Table 7.9: Evaluation of the alternative $A_j \in A$ on each subcriterion of Download speed

Simplicity and graphics and tables	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	1/3	1/3	1	2	0.12
A_2	3	1	2	3	5	0.40
A_3	3	1/2	1	3	5	0.30
A_4	1	1/3	1/3	1	3	0.13
A_5	1/2	1/5	1/5	1/3	1	0.06

$\lambda_{max} = 5.10$ $CI = 0.02$ $CR = 0.02$

Use of Animation	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	1/3	1/7	1/5	2	0.06
A_2	3	1	1/5	1/3	3	0.13
A_3	7	5	1	1	7	0.43
A_4	5	3	1	1	5	0.33
A_5	1/2	1/3	1/7	1/5	1	0.05

$\lambda_{max} = 5.15$ $CI = 0.04$ $CR = 0.03$

Use of Thumbnails	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	1	5	1/2	1	0.20
A_2	1	1	3	1/3	1/3	0.13
A_3	1/5	1/3	1	1/5	1/5	0.05
A_4	2	3	5	1	2	0.37
A_5	1	3	5	1/2	1	0.25

$\lambda_{max} = 5.15$ $CI = 0.04$ $CR = 0.03$

* The values are only accurate to 2 decimal positions.

Table 7.10: Evaluation of the alternative $A_j \in A$ on each subcriterion of Ease of use

Goals	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	2	2	3	1	0.30
A_2	1/2	1	1/2	1	1/3	0.11
A_3	1/2	2	1	1	1/2	0.16
A_4	1/3	1	1	1	1/3	0.12
A_5	1	3	2	3	1	0.32

$\lambda_{max} = 5.07$ $CI = 0.02$ $CR = 0.02$

Structure	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	3	2	3	1	0.30
A_2	1/3	1	1/3	1	1/3	0.09
A_3	1/2	3	1	3	1/2	0.20
A_4	1/3	1	1/3	1	1/5	0.08
A_5	1	3	2	5	1	0.33

$\lambda_{max} = 5.08$ $CI = 0.02$ $CR = 0.02$

Feedback	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	3	3	2	2	0.36
A_2	1/3	1	2	1	1/3	0.12
A_3	1/3	1/2	1	1/2	1/5	0.08
A_4	1/2	1	2	1	1/2	0.15
A_5	1/2	3	5	2	1	0.30

$\lambda_{max} = 5.14$ $CI = 0.03$ $CR = 0.03$

* The values are only accurate to 2 decimal positions.

Table 7.11: Evaluation of the alternative $A_j \in A$ on each subcriterion of Errors

Number of errors	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	2	3	1	1	0.27
A_2	1/2	1	2	3	1	0.25
A_3	1/3	1/2	1	1/2	1/2	0.10
A_4	1	1/3	2	1	1	0.18
A_5	1	1	2	1	1	0.21

$\lambda_{max} = 5.28$ $CI = 0.07$ $CR = 0.06$

Severity of errors	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	1	5	3	1	0.30
A_2	1	1	2	2	1	0.23
A_3	1/5	1/2	1	1/3	1/5	0.07
A_4	1/3	1/2	3	1	1/2	0.13
A_5	1	1	5	2	1	0.27

$\lambda_{max} = 5.15$ $CI = 0.04$ $CR = 0.03$

Recovering from errors	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	1	2	3	2	0.30
A_2	1	1	3	2	1	0.26
A_3	1/2	1/3	1	2	1/2	0.13
A_4	1/3	1/2	1/2	1	1	0.12
A_5	1/2	1	2	1	1	0.19

$\lambda_{max} = 5.22$ $CI = 0.05$ $CR = 0.05$

* The values are only accurate to 2 decimal positions.

Table 7.12: Evaluation of the alternative $A_j \in A$ on each subcriterion of Navigation

Organisation	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	1/5	1/3	1/3	1/7	0.05
A_2	5	1	3	3	1/3	0.26
A_3	3	1/3	1	1	1/5	0.11
A_4	3	1/3	1	1	1/3	0.12
A_5	7	3	5	3	1	0.47

$\lambda_{max} = 5.15$ $CI = 0.04$ $CR = 0.03$

Arrangement	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	1/5	1/7	1/3	1/9	0.03
A_2	5	1	1/3	1/3	1/7	0.08
A_3	7	3	1	3	1/5	0.22
A_4	3	3	1/3	1	1/3	0.13
A_5	9	7	5	3	1	0.53

$\lambda_{max} = 5.47$ $CI = 0.12$ $CR = 0.11$

Layout	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	1/5	1/3	1/5	1/9	0.04
A_2	5	1	5	5	1/3	0.30
A_3	3	1/5	1	1	1/3	0.10
A_4	5	1/5	1	1	1/5	0.10
A_5	9	3	3	5	1	0.46

$\lambda_{max} = 5.39$ $CI = 0.10$ $CR = 0.09$

Sequencing	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	7	3	5	2	0.40
A_2	1/7	1	1/3	1/3	1/5	0.04
A_3	1/3	3	1	3	1/5	0.13
A_4	1/5	3	1/3	1	1/7	0.07
A_5	1/2	5	5	7	1	0.36

$\lambda_{max} = 5.36$ $CI = 0.09$ $CR = 0.08$

* The values are only accurate to 2 decimal positions.

Table 7.13: Evaluation of the alternative $A_j \in A$ on each subcriterion of Site content

Amount and variety	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	9	7	5	3	0.51
A_2	1/9	1	1/5	1	1/9	0.04
A_3	1/7	5	1	3	1/5	0.11
A_4	1/5	1	1/3	1	1/5	0.05
A_5	1/3	9	5	5	1	0.30
	$\lambda_{max} = 5.41$		$CI = 0.10$		$CR = 0.09$	

Relevance of the content	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	9	5	5	1	0.39
A_2	1/9	1	1/3	1/3	1/7	0.04
A_3	1/5	3	1	3	1/5	0.12
A_4	1/5	3	1/3	1	1/5	0.07
A_5	1	7	5	5	1	0.38
	$\lambda_{max} = 5.22$		$CI = 0.06$		$CR = 0.05$	

Use of media	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	1/3	5	1	1/7	0.10
A_2	3	1	5	3	1/5	0.20
A_3	1/5	1/5	1	1/5	1/9	0.03
A_4	1	1/3	5	1	1/5	0.10
A_5	7	5	9	5	1	0.57
	$\lambda_{max} = 5.34$		$CI = 0.08$		$CR = 0.08$	

Appropriate content	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	7	5	3	1/3	0.29
A_2	1/7	1	1/2	1/3	1/7	0.04
A_3	1/5	2	1	1	1/5	0.08
A_4	1/3	3	1	1	1/3	0.11
A_5	3	7	5	3	1	0.47
	$\lambda_{max} = 5.19$		$CI = 0.05$		$CR = 0.04$	

Timely/Current information	A_1	A_2	A_3	A_4	A_5	Priority vector*
A_1	1	3	9	3	3	0.44
A_2	1/3	1	5	3	1	0.20
A_3	1/9	1/5	1	1/5	1/7	0.03
A_4	1/3	1/3	5	1	1/3	0.11
A_5	1/3	1	7	3	1	0.21
	$\lambda_{max} = 5.23$		$CI = 0.06$		$CR = 0.05$	

* The values are only accurate to 2 decimal positions.

$$
\begin{array}{c} \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \end{array}
\begin{array}{c} \begin{array}{ccccc} A_1 & A_2 & A_3 & A_4 & A_5 \end{array} \\
\begin{pmatrix}
19 & 23 & 6 & 13 & 39 \\
27 & 28 & 5 & 16 & 24 \\
12 & 24 & 29 & 26 & 10 \\
32 & 10 & 15 & 11 & 32 \\
29 & 25 & 11 & 13 & 21 \\
11 & 19 & 13 & 11 & 46 \\
34 & 11 & 7 & 9 & 39
\end{pmatrix} \end{array} \tag{7.11}
$$

The results presented in the matrix (7.11) are shown graphically in Figure 7.2.

7.1.4 Step 4: Ranking of the alternatives

To rank the alternatives we are going to use the distributive method presented in Section 4.4. By using this method we get the ranking vector

$$
\mathbf{x} = \begin{pmatrix} 0.22 & 0.20 & 0.12 & 0.15 & 0.30 \end{pmatrix}^{\mathbf{T}} \tag{7.12}
$$

And so using AHP methodology we get the ranking presented in Table 7.14. This result is shown graphically in Figure 7.3.

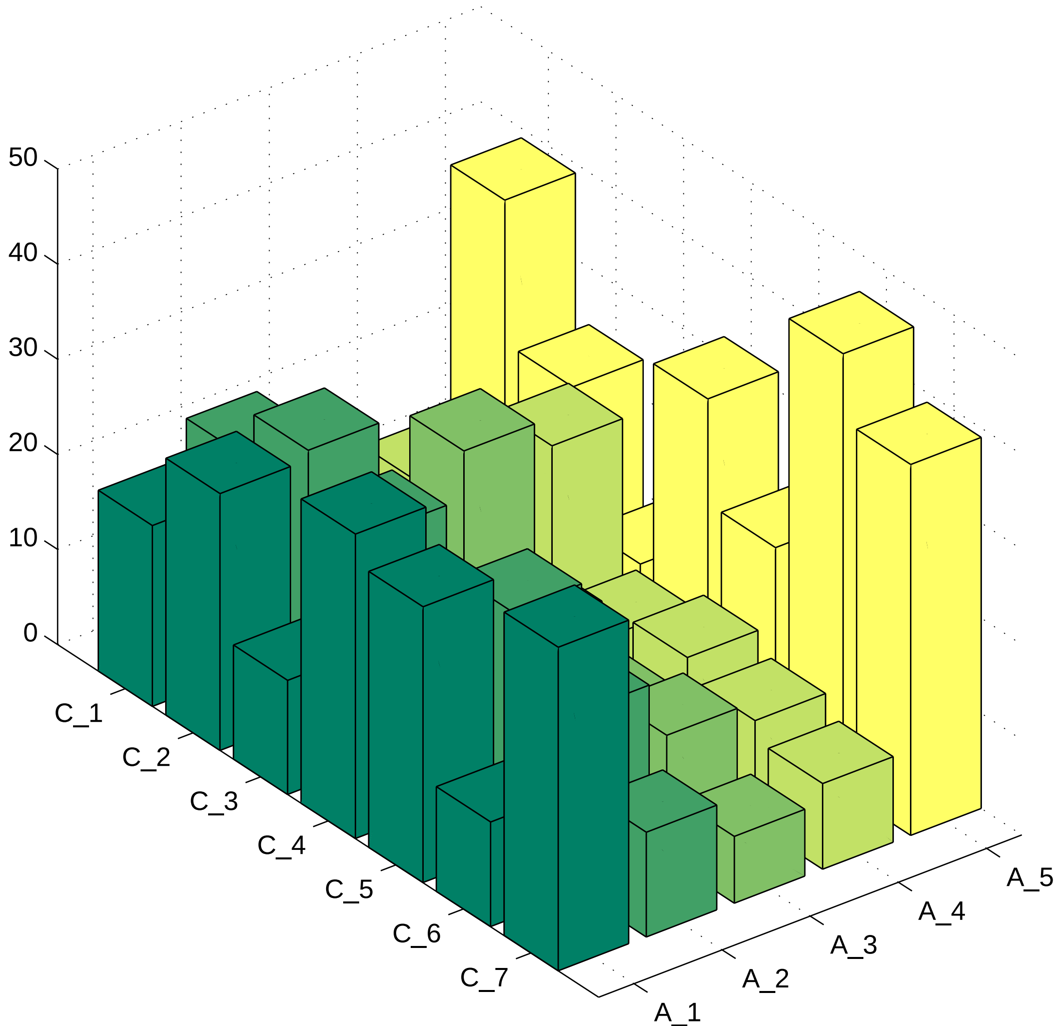


Figure 7.2: Evaluation of the alternatives with respect to the criteria by using AHP methodology

Table 7.14: Ranking of the alternatives using AHP methodology.

Alternative	Score*	Ranking position
A_1: Amazon	0.22	2
A_2: Blackwell	0.20	3
A_3: Bookstore	0.12	5
A_4: Borders	0.17	4
A_5: Waterstones	0.30	1

*The values are only accurate to 2 decimal positions.

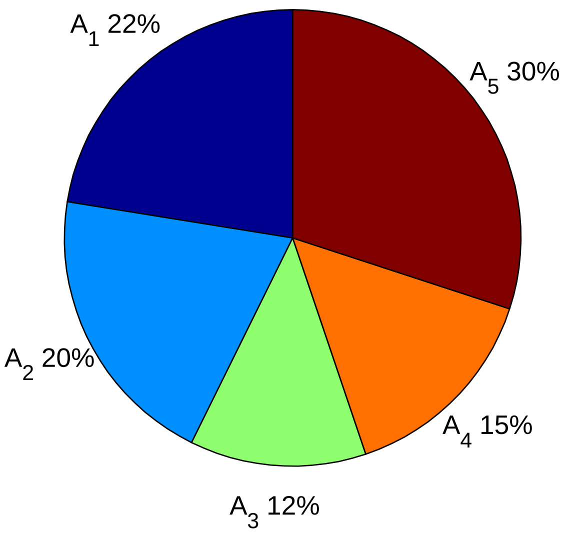


Figure 7.3: Ranking of the alternatives using AHP methodology

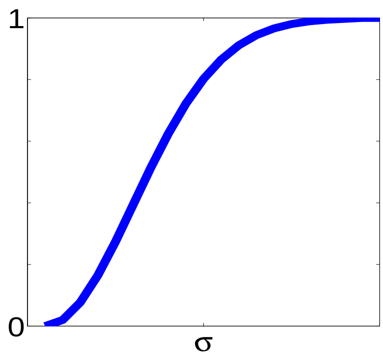
7.2 Using PROMETHEE

For the application of PROMETHEE methodology, as explained in Chapter 5, recall the evaluation criteria $C_i \in C$, for $i = 1, \ldots, 7$ defined in Section 6.2, the alternatives $A_j \in A$, for $j = 1, \ldots, 5$ defined in Section 6.3, and the decision table given in the matrix (7.11).

7.2.1 Step 1: Criterion function

Since each criterion C_i is the result of the preference of the decision maker, for all $i = 1, \ldots, n$, as a result, the criteria are said to be subjective. Therefore we use the Gaussian criterion function for $i = 1, \ldots, n$ which is normally used for continue data when using PROMETHEE methodology (Brans (1984)). We have seen that the Gaussian criterion function is given by

$$P(d) = \begin{cases} 0 & d \leqslant 0, \\ 1 - e^{-d^2/2\sigma^2} & d \geqslant 0. \end{cases}$$



Thus we need to define an appropriate value for the parameter σ. To determine the standard deviation we consider the data in the decision table of the matrix (7.11), and we use the equation (5.5), i.e.

$$\sigma_i^2 = \frac{\sum_{j=1}^{n}(x_{ij} - \mu)^2}{n-1} \qquad i = 1, \ldots, n \tag{7.13}$$

where μ corresponds to the mean of the data, i.e.

$$\mu_i = \frac{1}{n}\sum_{j=1}^{n} x_{ij} \qquad i = 1,\ldots,n \tag{7.14}$$

The results of σ_i corresponding to each criterion are given in Table 7.15

Table 7.15: Standard deviation of the data in the decision table on each criterion $C_i \in C$.

Criterion	Mean μ	Standard deviation* σ
C_1	20	12
C_2	20	10
C_3	20	9
C_4	20	11
C_5	20	8
C_6	20	15
C_7	20	15

* The values are only accurate to 2 decimal positions.

7.2.2 Step 2: Criterion value

To calculate the criterion value it is necessary first to calculate the deviation $d_i(A_k, A_l)$ based on pairwise comparison for expressing the evaluation of the alternatives A_k and A_l according to the criterion i. To do this, recall equation (5.8), i.e.

$$d_i(A_k, A_l) = x_{ik} - x_{il}$$

where x_{ij} represents the values given in the decision table represented in the matrix (7.11). The results of the deviations are presented in Table 7.16.

Table 7.16: Deviations $d_i(A_k - A_l)$ in the evaluations between the alternatives $A_k \in A$ and $A_l \in A$ on each criterion $C_i \in C$.*

C_1	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$
$k=1$	0	-4	14	6	-19
$k=2$	4	0	17	10	-16
$k=3$	-14	-17	0	-7	33
$k=4$	-6	-10	7	0	-26
$k=5$	19	16	33	26	0
C_2	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$
$k=1$	0	-2	22	11	3
$k=2$	2	0	23	13	4
$k=3$	-22	-23	0	-11	-19
$k=4$	-11	-13	11	0	-8
$k=5$	-3	-4	19	8	0
C_3	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$
$k=1$	0	-12	-17	-14	2
$k=2$	12	0	-5	-2	14
$k=3$	17	5	0	3	19
$k=4$	14	2	-3	0	16
$k=5$	-2	-14	-19	-16	0
C_4	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$
$k=1$	0	21	16	21	0
$k=2$	-21	0	-5	0	-21
$k=3$	-16	5	0	5	-16
$k=4$	-21	0	-5	0	-21
$k=5$	0	21	16	21	0
C_5	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$
$k=1$	0	4	18	16	8
$k=2$	-4	0	14	12	4
$k=3$	-18	-14	0	-2	-11
$k=4$	-16	-12	2	0	-8
$k=5$	-8	-4	11	8	0
C_6	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$
$k=1$	0	-8	-2	1	-35
$k=2$	8	0	6	9	-27
$k=3$	2	-6	0	3	-33
$k=4$	-1	-9	-3	0	-35
$k=5$	35	27	33	35	0
C_7	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$
$k=1$	0	22	27	24	-5
$k=2$	-22	0	4	2	-27
$k=3$	-27	4	0	-2	-32
$k=4$	-24	-2	2	0	-29
$k=5$	5	27	32	29	0

*The values are only accurate to 2 decimal positions.

Having calculated the deviations based on pairwise comparisons, the values of $P_i(d)$ are calculated by

$$P_i(A_k, A_l) = \begin{cases} 0 & d \leqslant 0, \\ 1 - e^{-d^2/2\sigma^2} & d \geqslant 0. \end{cases}$$

where the values of d corresponds to the value $d_i(A_k, A_l)$ as presented in Table 7.16. The values of $P_i(A_k, A_l)$ are summarised in Table 7.17.

7.2.3 Step 3: Preference index

Recall that a preference index is calculated by

$$\pi(A_k, A_l) = \sum_{i=1}^{n} w_i P_i(A_k, A_l) \qquad k, l = 1, \ldots, m$$

where w_i is the weight corresponding to the criterion $C_i \in C$ for $i = 1, \ldots, n$. The values of $\pi(A_k, A_l)$ corresponding to the alternatives given have been summarised in (7.15)

$$\pi(A_k, A_l) : \quad \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \end{array} \begin{array}{c} \begin{array}{ccccc} A_1 & A_2 & A_3 & A_4 & A_5 \end{array} \\ \begin{pmatrix} 0 & 0.21 & 0.49 & 0.37 & 0.03 \\ 0.14 & 0 & 0.35 & 0.21 & 0.16 \\ 0.16 & 0.04 & 0 & 0.02 & 0.16 \\ 0.13 & 0.00 & 0.11 & 0 & 0.15 \\ 0.27 & 0.44 & 0.70 & 0.61 & 0 \end{pmatrix} \end{array} \qquad (7.15)$$

Table 7.17: Criterion values $P_i(A_k - A_l)$ between the alternatives $A_k \in A$ and $A_l \in A$ on each criterion $C_i \in C$.*

C_1	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$
$k=1$	0.00	0.00	0.45	0.12	0.00
$k=2$	0.04	0.00	0.62	0.28	0.00
$k=3$	0.00	0.00	0.00	0.00	0.00
$k=4$	0.00	0.00	0.15	0.00	0.00
$k=5$	0.71	0.56	0.97	0.89	0.00
C_2	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$
$k=1$	0.00	0.00	0.92	0.48	0.04
$k=2$	0.01	0.00	0.95	0.58	0.10
$k=3$	0.00	0.00	0.00	0.00	0.00
$k=4$	0.00	0.00	0.45	0.00	0.00
$k=5$	0.00	0.00	0.85	0.30	0.00
C_3	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$
$k=1$	0.00	0.00	0.00	0.00	0.02
$k=2$	0.64	0.00	0.00	0.00	0.74
$k=3$	0.86	0.14	0.00	0.05	0.91
$k=4$	0.74	0.02	0.00	0.00	0.82
$k=5$	0.00	0.00	0.00	0.00	0.00
C_4	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$
$k=1$	0.00	0.85	0.67	0.84	0.00
$k=2$	0.00	0.00	0.00	0.00	0.00
$k=3$	0.00	0.10	0.00	0.09	0.00
$k=4$	0.00	0.00	0.00	0.00	0.00
$k=5$	0.00	0.86	0.68	0.85	0.00
C_5	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$
$k=1$	0.00	0.12	0.94	0.87	0.40
$k=2$	0.00	0.00	0.82	0.69	0.12
$k=3$	0.00	0.00	0.00	0.00	0.00
$k=4$	0.00	0.00	0.05	0.00	0.00
$k=5$	0.00	0.00	0.60	0.41	0.00
C_6	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$
$k=1$	0.00	0.00	0.00	0.00	0.00
$k=2$	0.13	0.00	0.08	0.15	0.00
$k=3$	0.01	0.00	0.00	0.01	0.00
$k=4$	0.00	0.00	0.00	0.00	0.00
$k=5$	0.93	0.80	0.91	0.94	0.00
C_7	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$
$k=1$	0.00	0.67	0.79	0.74	0.00
$k=2$	0.00	0.00	0.04	0.01	0.00
$k=3$	0.00	0.00	0.00	0.00	0.00
$k=4$	0.00	0.00	0.01	0.00	0.00
$k=5$	0.06	0.81	0.89	0.86	0.00

*The values are only accurate to 2 decimal positions.

7.2.4 Step 4a: Partial ranking

To determine the partial ranking of the given alternatives we use the following equations.

$$\phi^{+}(A_i) = \frac{1}{n-1} \sum_{k=1}^{l=5} \pi(A_i, A_k) \qquad i = 1, \ldots, n = 5 \tag{7.16}$$

$$\phi^{-}(A_i) = \frac{1}{n-1} \sum_{k=1}^{l=5} \pi(A_k, A_i) \qquad i = 1, \ldots, n = 5 \tag{7.17}$$

The results for the defined problem in this case study are presented in (7.18). Note that the values presented in (7.18) are only accurate to 2 decimal positions.

(7.18)

	A_1	A_2	A_3	A_4	A_5	$\phi^{+}(A_i)$
A_1	0	0.21	0.49	0.37	0.03	0.28
A_2	0.14	0	0.35	0.21	0.16	0.22
A_3	0.16	0.04	0	0.02	0.16	0.10
A_4	0.13	0.00	0.11	0	0.15	0.10
A_5	0.27	0.44	0.70	0.61	0	0.50
$\phi^{-}(A_i)$	0.18	0.17	0.41	0.30	0.13	

The partial ranking of the alternatives can be represented graphically by

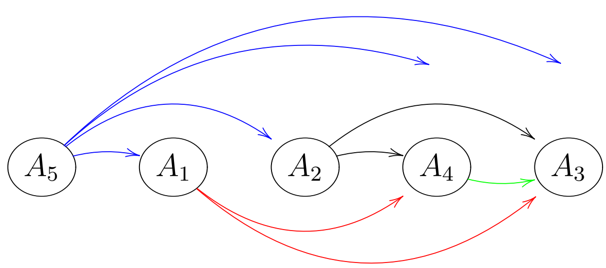


Table 7.18: Partial ranking between the alternatives A_k and $A_l \in A$.

A_k vs A_l	Preference $\phi^+(A_k) > \phi^+(A_l)$ and $\phi^-(A_k) < \phi^-(A_l)$ $\phi^+(A_k) > \phi^+(A_l)$ and $\phi^-(A_k) = \phi^-(A_l)$ $\phi^+(A_k) = \phi^+(A_l)$ and $\phi^-(A_k) < \phi^-(A_l)$	Indifference $\phi^+(A_k) = \phi^+(A_l)$ $\phi^-(A_k) = \phi^-(A_l)$	Not comparable otherwise	Graph
A_1 vs A_2	-	-	✓	-
A_1 vs A_3	✓	-	-	$A_1 \rightarrow A_3$
A_1 vs A_4	✓	-	-	$A_1 \rightarrow A_4$
A_1 vs A_5	-	-	✓	-
A_2 vs A_1	-	-	✓	-
A_2 vs A_3	✓	-	-	$A_2 \rightarrow A_3$
A_2 vs A_4	✓	-	-	$A_2 \rightarrow A_4$
A_2 vs A_5	-	-	✓	-
A_3 vs A_1	-	-	✓	-
A_3 vs A_2	-	-	✓	-
A_3 vs A_4	-	-	✓	-
A_3 vs A_5	-	-	✓	-
A_4 vs A_1	-	-	✓	-
A_4 vs A_2	-	-	✓	-
A_4 vs A_3	✓	-	-	$A_4 \rightarrow A_3$
A_4 vs A_5	-	-	✓	-
A_5 vs A_1	✓	-	-	$A_5 \rightarrow A_1$
A_5 vs A_2	✓	-	-	$A_5 \rightarrow A_2$
A_5 vs A_3	✓	-	-	$A_5 \rightarrow A_3$
A_5 vs A_4	✓	-	-	$A_5 \rightarrow A_4$

Observe that since we do not have information between the alternatives A_1 and A_2 then it is necessary to do the complete ranking as well.

7.2.5 Step 4b: Complete ranking

To determine the complete ranking, the net outranking flow for each alternative is required. The latter is obtained by using

$$\phi(A_i) = \phi^+(A_i) - \phi^-(A_i) \tag{7.19}$$

The net flow in our case is presented in Table 7.19.

Table 7.19: Net flow of the alternatives using PROMETHEE.

A_i	$\phi^+(A_i)$	$\phi^-(A_i)$	$\phi(A_i)$
A_1	0.28	0.18	0.10
A_2	0.22	0.17	0.04
A_3	0.10	0.41	-0.32
A_4	0.10	0.30	-0.20
A_5	0.50	0.13	0.38

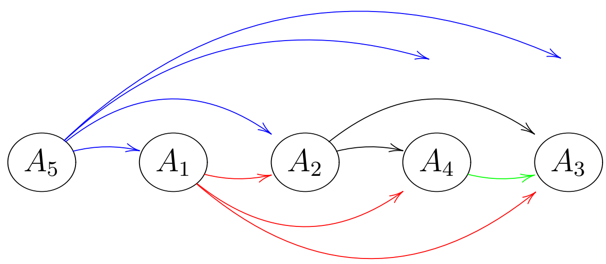
*The values are only accurate to 2 decimal positions.

We analyse the conditions to establish the complete ranking in Table 7.20.

Table 7.20: Complete ranking between the alternatives $A_k \in A$ and $A_l \in A$.

A_k vs A_l	Preference $\phi(A_k) > \phi(A_l)$	Indifference $\phi(A_k) = \phi(A_l)$	Graph
A_1 vs A_2	✓	-	$A_1 \rightarrow A_2$
A_1 vs A_3	✓	-	$A_1 \rightarrow A_3$
A_1 vs A_4	✓	-	$A_1 \rightarrow A_4$
A_1 vs A_5	-	-	-
A_2 vs A_1	-	-	-
A_2 vs A_3	✓	-	$A_2 \rightarrow A_3$
A_2 vs A_4	✓	-	$A_2 \rightarrow A_4$
A_2 vs A_5	-	-	-
A_3 vs A_1	-	-	-
A_3 vs A_2	-	-	-
A_3 vs A_4	-	-	-
A_3 vs A_5	-	-	-
A_4 vs A_1	-	-	-
A_4 vs A_2	-	-	-
A_4 vs A_3	✓	-	$A_4 \rightarrow A_3$
A_4 vs A_5	-	-	-
A_5 vs A_1	✓	-	$A_5 \rightarrow A_1$
A_5 vs A_2	✓	-	$A_5 \rightarrow A_2$
A_5 vs A_3	✓	-	$A_5 \rightarrow A_3$
A_5 vs A_4	✓	-	$A_5 \rightarrow A_4$

The complete ranking of the alternatives according with the information of Table 7.20 can be represented graphically by



And so by using PROMETHEE methodology we get the results shown in Table 7.21.

Table 7.21: Ranking of the alternatives using PROMETHEE methodology.

Alternative	Ranking position
A_1: Amazon	2
A_2: Blackwell	3
A_3: Bookstore	5
A_4: Borders	4
A_5: Waterstones	1

CHAPTER 8

RESULTS

8.1 Multiple criteria optimisation

Using both AHP and PROMETHEE methodologies we found the same ranking of the alternatives based on usability criteria, as shown in Table 8.1.

Table 8.1: Ranking of the alternatives using AHP methodology.

Ranking position	AHP Alternative	Score	PROMETHEE
1	A_5: Waterstones	0.30	A_5: Waterstones
2	A_1: Amazon	0.22	A_1: Amazon
3	A_2: Blackwell	0.20	A_2: Blackwell
4	A_4: Borders	0.17	A_4: Borders
5	A_3: Bookstore	0.12	A_3: Bookstore

Observe that AHP additionally gives us information about the weights of the criteria whereas PROMETHEE only provides us with a ranking relation.

8.2 One criterion optimisation

If we analyse each criterion one at a time and try to find an optimal solution this will reduce the complexity of the problem, because instead of having multiple criteria we will have to optimise only one criterion. In this situation to attain an optimal solution is a straightforward procedure. Recall the decision table presented in the matrix (7.11).

$$\begin{array}{c} \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \end{array} \begin{array}{c} \begin{array}{ccccc} A_1 & A_2 & A_3 & A_4 & A_5 \end{array} \\ \left(\begin{array}{ccccc} 19 & 23 & 6 & 13 & 39 \\ 27 & 28 & 5 & 16 & 24 \\ 12 & 24 & 29 & 26 & 10 \\ 32 & 10 & 15 & 11 & 32 \\ 29 & 25 & 11 & 13 & 21 \\ 11 & 19 & 13 & 11 & 46 \\ 34 & 11 & 7 & 9 & 39 \end{array} \right) \end{array} \tag{8.1}$$

We could rank the alternatives for each criterion according with the data in the decision table. In this case our results are presented in Table 8.2.

Observe that we have different alternatives as optimal solutions (alternatives in the first column) if we analyse each criterion separately. Therefore there is no alternative that optimises each criterion at the same time. However by using AHP and PROMETHEE methodologies a satisfactory solution, in the sense that we cannot get a better solution, is attained.

Table 8.2: Ranking of the alternatives for each criterion C_i, for $i = 1, \ldots, n$.

Criterion	Ranking 1		Ranking 2		Ranking 3		Ranking 4		Ranking 5	
C_1	A_5	39%	A_2	23%	A_1	19%	A_4	13%	A_3	6%
C_2	A_2	28%	A_1	27%	A_5	24%	A_4	16%	A_3	5%
C_3	A_3	29%	A_4	26%	A_2	24%	A_1	12%	A_5	10%
C_4	A_1	32%	A_5	32%	A_3	15%	A_4	11%	A_2	10%
C_5	A_1	39%	A_2	25%	A_5	21%	A_4	13%	A_3	11%
C_6	A_5	46%	A_2	19%	A_3	13%	A_1	11%	A_4	11%
C_7	A_5	39%	A_1	34%	A_2	11%	A_4	9%	A_3	7%

BIBLIOGRAPHY

Aczél, J. and Saaty, T. L. (1983). Procedures for synthesising ratio judgements. *Journal of Mathematical Psychology*, 27:93–102.

Agarwal, R. and Venkatesh, V. (2002). Assessing a firm's web presence: A heuristic evaluation procedure for the measurement of usability. *Information Systems Research*, 13(2):168–186.

Amazon (1998). Company information. Retrieved on June 22, 2009 from http://www.amazon.co.uk/gp/press/info/home/ref=cs_nav_bn_ab.

Andrilli, S. F. and Hecker, D. (2003). *Elementary linear algebra*. Gulf Professional Publishing, 3 edition.

Ballestero, E. and Romero, C. (1998). *Multiple Criteria Decision Making and Its Applications to Economic Problems*. Springer.

Bapat, R. B. and Raghavan, T. E. S. (1997). *Nonnegative Matrices And Applications*. Cambridge University Press.

Blackwell (2009). The history of blackwell. Retrieved on June 22, 2009 from http://bookshop.blackwell.co.uk/jsp/editorial/services/about_blackwells.

Bookstore (2009). About us. Retrieved on June 22, 2009 from http://www.bookstore.co.uk/TBP.Direct/CustomerAccessControl/AboutUs.

Borders (2009). About us. Retrieved on June 22, 2009 from http://www.borders.com/online/store/BGIView_aboutus.

Brans, J. P. (1982). L'Ingéniérie de la décision. Elaboration d'instrument d'aide à la décision. Méthode PROMETHEE. In *Colloque d'Aide à la Décision*, pages 183–213, Québec, Canada. Université Laval.

Brans, J.-P. and Mareschal, B. (2005). *Multiple Criteria Decision Analysis: State of the Art Surveys*, volume 78 of *International Series In Operations Research & Management Science*, chapter Promethee Methods, pages 163–186. Springer New York.

Brans, J.-P., Vincke, P., and Mareschal, B. (1986). How to select and how to rank projects: The promethee method. *European Journal of Operational Research*, 24:228–238.

Brans, J.P., M. B. V. P. (1984). Promethee: a new family of outranking methods. In *IFORS*, volume 84 of *MCDM*, page 477490, North Holland.

Card, S., Moran, T. P., and Newell., A. (1983). *The Psychology of Human Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cauchy, A.-L. (1821). *Cours d'analyse de l'cole Royale Polytechnique*, volume premire partie. Analyse algbrique, Paris.

Claiborne, S. (2005). Effective site navigation. Retrieved on October 20, 2008 from http://www.webcredible.co.uk/user-friendly-resources/web-usability/effective-site-navigation.shtml.

Dyer, J. S., Fishburn, P. C., Steuer, R. E., Wallenius, J., and Zionts, S. (1992). Multiple criteria decision making, multiattribute utility theory: the next ten years. *Management Science*, 38(5):645–654.

Erica S. Davis, D. A. H. (2001). The effects of download delay on performance and end-user satisfaction in an internet tutorial. *Computers in Human Behavior*, 17:249–268.

Franklin, J. N. (1968). *Matrix Theory*. Prentice-Hall, New Jersey.

Frobenius, F. G. (1912). Über matrizen aus nicht negativen elementen. *Sitzungsberichte der Königlich Preusischen Akademie der Wissenschaften zu Berlin*, 4:456–477.

Goumas, M. and Lygerou, V. (2000). An extension of the promethee method for decision making in fuzzy environment: Ranking of alternative energy exploitation projects. *European Journal of Operational Research*, 123:606–613.

Horn, R. A. and Johnson, C. R. (1999). *Matrix analysis*. Cambridge University Press.

Institute of Electrical and Electronics Engineers (1990). 610.12-1990. *IEEE Standard Glossary of Software Engineering Terminology*. Los Alamitos, CA: Author.

International Organization for Standardization (1998). ISO 9241-11. *Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs), Part 11: Guidance on Usability*. Geneva: Author.

Jacobs, I. (2008). World Wide Web Consortium (W3C). "Facts about W3C". Retrieved on October 18, 2008 from http://www.w3.org/Consortium/facts.

Johnson, C. R. (1980). Constructive critique of a hierarchical prioritization scheme employing paired comparisons. In *Proceedings of the International Conference of Cybernetics and Society of the IEEE*, pages 373–378, Cambridge, Massachusetts. Institute of Electrical Engineers.

Keeker, K. (1997). Improving web site usability and appeal: Guidelines compiled by msn usability research,. Retrieved on October 21, 2008 from http://msdn.microsoft.com/en-us/library/cc889361(office.11).aspx.

Koyani, S. J., Bailey, R. W., and Nall, J. R. (2004). *Research-Based Web Design & Usability Guidelines*. U.S. Department of Health and Human Services.

Larichev, O. (2002). Properties of methods of acceptance of decisions in multicriteria problems of an individual choice. *Avtomatika i Telemechanika (Automatics and Telemechanics)*, 2:146–158. (in Russian).

Meyer, C. D. (2000). *Matrix Analysis and applied linear algebra*. siam.

Nielsen, J. (1993). *Usability Engineerin*. Morgan Kaufmann.

Nielsen, J. (1997). The need for speed. Retrieved on October 19, 2008 from http://www.useit.com/alertbox/9703a.html.

Nielsen, J. (2000). *Designing Web Usability*. New Riders, Indianapolis, IN, Nunnally, J, C, 1978, *Psychometric Theory*, McGraw Hill, New York.

Nielsen, J. (2003). Usability 101: Introduction to usability. Retrieved on October 19, 2008 from http://www.useit.com/alertbox/20030825.html.

Norros, L. and Savioja, P. (2004). Usability evaluation of complex systems: A literature review. Technical Report STUK-YTO-TR 204, STUK Radiation and Nuclear Safety Authority, Helsinki.

Palmer, J. W. (2002). Web site usability, design, and performance metrics. *Information System Research*, 13(2):151–167.

Park, K. S. and Lim, C. H. (1999). A structured methodology for comparative evaluation of user interface designs using usability criteria and measures. *International Journal of Industrial Ergonomics*, 23:379–389.

Pearson, J. M. and Pearson, A. M. (2008). An exploratory into determining the relative importance of key criteria in web usability: A multicriteria approach. *Journal of Computer Information Systems*, pages 115–127.

Perron, O. (1907). Zur theorie der matrices. *Mathematische Annalen*, 64:248–263.

Porter, M. E. (2001). Strategy and the internet. *Harvard Business Review*, 79(3):62–78.

Ranganathan, C. and Ganapathy, S. (2002). Key dimensions of business-to-consumers web sites. *Information & Management*, 39:457–465.

Rapcsák, T. (2007). Többszempontú döntési problémák. Egyetemi oktatáshoz segédanyag, Budapesti Corvinus Egyetem MTA Számítástechnikai és Automatizálási Kutató Intézetébe kihelyezett Gazdasági Döntések Tanszék.

Rose, G., Khoo, H., and Strub, D. (1999). Current technological impediments to business-to-consumer electronic commerce. *Communications of AIS*, 16:1–74.

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

Saaty, T. L. (1990). *Multicriteria Decision Making - The Analytic Hierarchy Process*, volume I of *the AHP series*. University of Pittsburgh, Pittsburgh, 2nd edition.

Seffah, A., Donyaee, M., Kline, R. B., and Padda, H. K. (2006). Usability measurement and metrics: A consolidated model. *Software Quality Journal*, 14:159–178.

Shneiderman, B. (1998). *Designing the User Interface*. Addison-Wesley Longman, Inc, Boston, MA,, 3rd edition.

Steuer, R. E. and Na, P. (2003). Multiple criteria decision making combined with finance: A categorised bibliography. *European Journal of Operational Research*, 150(3):496–515.

Texas A&M University (2004). Web accessibility and usability procedures. Retrieved on November 15, 2008 from http://www.tamug.edu/computing/accessibility/procedures.htm.

Triantaphyllou, E. (2000). *Multi-Criteria Decision Making Methods: A comparative Study*, volume 44 of *Applied Optimization*. Kluwer Academic Publishers.

Turban, E. and Gehrke, D. (2000). Determinants of e-commerce web site. *Human Systems Management*, 19(2):111–120.

Vaidya, O. S. and Kumar, S. (2006). Analytic hierarchy process: An overview of applications. *European Journal of Operational Research*, 169:1–29.

Varga, R. S. (1962). *Matrix Iterative Analysis*. Prentice-Hall, New Jersey.

Venkatesh, V. (2000). Determinants of perceived ease of use: Integrating perceived behavioral control, computer anxiety and enjoyment into the technology acceptance model. *Information Systems Research*, 11:342–365.

Venkatesh, V., Morris, M., Davis, G., and Davis, F. (2003). User acceptance of infonnation technology: Toward a unified view. *MIS Quarterly*, 27(3):425–478.

Vincke, J. and Brans, P. (1985). A preference ranking organization method. The PROMETHEE method for MCDM. *Management Science*, 31:641656.

Waterstone's (2009). About waterstone's. Retrieved on June 22, 2009 from http://www.waterstones.com/waterstonesweb/navigate.do?pPageID=100000002.

Wim De Keyser, P. P. (1996). A note on the use of promethee multicriteria methods. *European Journal of Operational Research*, 89:457–461.

World Wide Web Consortium (WC3) (1999). Web content accessibility guidelines 1.0. Retrieved on Octuber 19, 2008 from http://www.w3.org/TR/1999/WAI-WEBCONTENT-19990505/wai-pageauth.pdf.

Zeleny, M. (1998). Multiple criteria decision making: Eight concepts of optimality. *Human Systems Management*, 17(2).