2014

The Analytic Hierarchy Process: A Mathematical Model for Decision Making Problems

Giang Huong Nguyen
The College of Wooster, giang.huong.nguyen92@gmail.com

The Analytic Hierarchy Process: A Mathematical Model for Decision Making Problems

Independent Study Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Bachelor of Arts in the Department of Mathematics and Computer Science at The College of Wooster

by
Giang Huong Nguyen

The College of Wooster
2014

Advised by:

Dr. John Ramsay

Abstract

The ability to make the right decision is an asset in many areas and lines of profession including social work, business, national economics, and international security. However, decision makers often have difficulty choosing the best option since they might not have a full understanding of their preferences, or lack a systematic approach to solve the decision making problems at hand. The Analytic Hierarchy Process (AHP) provides a mathematical model that helps the decision makers arrive at the most logical choice, based on their preferences. We investigate the theory of positive, reciprocal matrices, which provides the theoretical justification of the method of the AHP. At its heart, the AHP relies on three principles: Decomposition, Measurement of preferences, and Synthesis. Throughout the first five chapters of this thesis, we use a simple example to illustrate these principles. The last chapter presents a more sophisticated application of the AHP, which in turn illustrates the Analytic Network Process, a generalization of the AHP to systems with dependence and feedback.

Acknowledgements

I would like to thank Professor John Ramsay, my advisor, for his tremendous support with this thesis. His knowledge of the subject, passion for mathematics, and relentless encouragement were the main source of motivation that got me through a research project that at times seemed impossible. Being his student has been the most extraordinary experience in my sixteen years of schooling. I am most grateful to complete an Independent Study project under his guidance.

My sincere thanks goes to all of the professors in the Department of Mathematics and Computer Science. Each of them has offered advice that helped shape the outcome of this thesis.

I would also like to thank my friends, Dung Nguyen'14 and Chris and Christy Ranney, for their support. Without them, this thesis would not have been possible.

Last but not least, thank you, Mom and Dad, for having faith in my abilities and judgment.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

We make decisions every day. High school graduates decide which college to attend. Moms make up their minds as to how much time they should let their children watch TV each day. Chefs choose the spices to use in dishes. The Federal Reserve makes judgments about its monetary policy. Decision making is a part of everyday life, whether it be a personal choice, a decision with global impact, or anything in between. Being able to make right decisions can have a tremendously positive impact on individuals' lives.

However, making the right decision is often easier said than done. For starters, assuming that the decision makers are aware of their options, they need to identify their goals in choosing among these options. (As a simple example, consider a buyer, who is trying to decide between two kinds of tomatoes. He or she needs to know what qualities in tomatoes he or she values: the color, the shape, or the price.) Then, the decision makers need to determine how well each option satisfies their goals. Even when the decision makers have clear ideas of their goals and the qualities of their options, their judgments can still exhibit inconsistency, or can vary according to time. Variations can also occur due to the environment that surrounds the decision making process.

The Analytic Hierarchy Process (AHP), developed by Thomas Saaty [10] is a model that helps the decision makers arrive at the most logical choice. Using the AHP, we first determine the options available to the decision makers and their goals in making decisions. These goals and options will then be used to construct an analytic hierarchy, which reflects the various factors in the decision making process and their importance. The outcome of the AHP is a priority vector, which gives us an insight into the best option for the decision makers. In order to understand the strengths of the AHP, in the next section, we consider a simple problem of choosing universities.

1.2 The Initial Problem

Alice is in her last year of high school. She has been accepted to three major universities and is trying to decide which one to attend. Those universities are City University, Suburb University, and Town University. In evaluating a university, Alice considers three factors: the academic quality of the university, the financial aid package that the university offers her, and the quality of the town or city that surrounds the university. In our discussion of the AHP, we will refer to the universities that extend their acceptance to Alice as **alternatives** and the three factors as **objectives**.

One solution to Alice's problem is to directly assess her opinion of each university. For example, Alice can assign a score to each university, on a scale of 1–10, based on each of her objectives. Then the total score of each university is the sum of the scores of that university on the three objectives. The university with the highest total score is the one that Alice should attend. The solution to the problem seems easy enough.

However, what will happen if Alice has ten, instead of three, objectives? Then scoring the universities will be a more complicated and error-prone task. How about the possibility that each objective has a different level of importance to Alice? Then we have to take into account not only the score of each university on each objective, but also the weight of each objective to Alice's decision. The problem quickly escalates in terms of complexity if Alice also has different potential majors in mind, and her choice of university affects her choice of major. For example, one of the universities might not have a Pre-Health program, but it has a strong Music department (assuming that Alice is considering studying Pre-Health and Music). Another university has a prestigious Pre-Health curriculum, but it does not have a Music major. The third university has both Pre-Health and Music, allowing Alice more opportunities to explore her interest, but both of the programs are only average.

The AHP provides a systematic method of solving problems such as the one that Alice is facing. We mentioned in the last section that in order to apply the AHP, we first identify clearly the objectives and alternatives available to the decision makers. Then there are two steps that we need to carry out. First, we measure the importance of each objective to the decision maker, compared to the importance of the other objectives. This relative importance is called the **weight** of the objective. For ease of computing, we require that the measured weights sum to 1 [4, p. 30]. For this example, suppose that the AHP has found the weights for academic quality, financial aid, and quality of location to be 0.6, 0.3, and 0.1, respectively. According to these weights, Alice considers the quality of a university's education the most important factor in making her decision. Following in importance are the financial aid package that she is offered, and then the quality of the university's location.

In the next step, the AHP measures how well each of the decision maker's alternatives satisfies each of the objectives. The extent to which each alternative meets the decision maker's expectation in each objective is measured as a numerical value. This value is referred to as the **score** of the alternative on that objective. In this example, suppose that the AHP has found the scores for one of the universities, City University, on academic quality, financial aid, and quality of location, to be 5, 7.5, and 10, respectively. These scores indicate that Alice loves the setting of City University, while the financial aid that she is offered is moderately good, and the academic quality is only average. We noted earlier that a decision maker's preferences can exhibit inconsistency. The scores found by the AHP are not simply the decision maker's assessment of his or her own preferences. The AHP achieves these scores using a systematic approach that will be discussed in Chapter 4, helping the decision maker make the most logical choice given his or her preferences.

The reader might have guessed that in order for Alice to make a decision, both the weights of her objectives and the score of each alternative on those objectives must be taken into account. Indeed, we utilize both kinds of information to compute the total score for each alternative. For the first alternative, City University, we first take the products of each of City University's scores on an objective and the weight of that objective. Taking the sum of those products will give us the total score of City University. Using the values in this example, the total score of City University is computed as follows:

$$5 \times 0.6 + 7.5 \times 0.3 + 10 \times 0.1 = 6.25.$$

Given that we also know the score of the other alternatives on Alice's objectives, we can compute the total score of those alternatives using the same method explained above. The alternative with the highest total score should be chosen.

In subsequent chapters, it will be clear that the AHP involves more than just finding the weights of the objectives and the scores of the alternatives. In fact, the two steps outlined above are merely part of a hierarchy that is used to model the decision makers' preferences. The same idea of a hierarchy that is used to solve Alice's problem could be used to answer questions about promotion and tenure in higher education [10, p. 162], the optimum choice of coal plants [10, p. 156], and measuring the world influence of nations such as the U.S., China, France, the U.K., Japan, ... [10, p. 40]. For the purpose of understanding the foundations of the AHP, the next three chapters will investigate the two outlined tasks: finding the weights of the objectives and the scores of the alternatives. The discussion of the hierarchy will soon follow.

At this stage, the AHP problem is summarized in mathematical notations as follows: Suppose a decision maker has n objectives and m alternatives. In the first step of the AHP, for the i^{th} objective, the AHP generates a weight w_i. In the second step, for the i^{th} objective and the k^{th} alternative, the AHP obtains a score s_{ik} of the k^{th} alternative on the i^{th} objective. The total score of the k^{th} alternative is then computed by the following formula:

$$\sum_{i=1}^{n} w_i s_{ik}. \tag{1.1}$$

After the total scores of all alternatives have been calculated, the decision maker should choose the alternative that has the highest total score.

1.3 Research Outline

This thesis focuses on the method and mathematical reasoning of the AHP. The next two chapters cover the task of finding the weights of the objectives. Specifically, in Chapter 2, we will develop Alice's problem, introducing the basic terms and concepts of the AHP. As will be clear in this chapter, an essential concept of the AHP is the consistency of the decision maker's preferences. We will include the definition of consistency, and then focus on the consistent decision maker. In Chapter 3, we investigate the case of an inconsistent decision maker. We turn our attention to finding the scores of the alternatives in Chapter 4. The discussion of the scores naturally leads to the idea of a hierarchy. Having discussed the hierarchy, we apply it to extend our Alice's problem. In Chapter 5, we examine some measurement issues of the AHP. Specifically, we will provide an explanation for the measure of the decision maker's consistency, and comment on the scale that is used to measure the decision maker's pairwise preferences.

In the second-to-last chapter, we present a more sophisticated application of the AHP in medical diagnosis. The application in turn illustrates the Analytic Network Process, a generalization of the AHP to systems with dependence and feedback. Included in Chapter 7 are final remarks and ideas for future research.

The terminologies of the AHP are defined as they are first used in the subsequent chapters. Unless otherwise noted, these definitions apply throughout the thesis.

Chapter 2

The Consistent Decision Maker

2.1 The Eigenvalue Problem

In this chapter, we tackle the first task of the AHP: finding the weight of each objective of the decision maker. We continue with the Alice example presented in the last chapter, introducing terms and concepts that are fundamental to the AHP. One such concept is what is called the pairwise comparison matrix. This concept will naturally lead to the idea of the decision maker's consistency in preferences. The AHP problem will then be divided into two sub-problems: one for the consistent and the other for the inconsistent decision maker. In the rest of the chapter, we present the method of finding the weight vector for the consistent case. This method involves the eigenvalue problem, a concept central to the AHP. Finally, we provide an illustration of the eigenvalue problem for the Alice example.

It is the writer's intention to start the discussion of the AHP with a simple, instructive example such as the problem of choosing university for Alice.

However, the AHP has a wide array of applications in more complicated situations. The complexity of these situations is enhanced by the fact that

(i) the decision maker has a large number of objectives,

(ii) the objectives are divided into layers of importance,

(iii) there exists a relationship between objectives and alternatives,

or any combination of those mentioned above. For more sophisticated applications of the AHP, the interested reader is referred to [10, p. 91–163], which covers topics such as prediction of the number of children in a household, designing the transport system for the Sudan, and the future of higher education in the United States.

In decision making problems that involve a large number of objectives, it is difficult to compare each objective to the rest of the objectives. The process tends to result in error, much like the process of scoring each alternative on a scale of 1–10, compared to the other alternatives, that we mentioned in the introduction. A method that provides a more accurate assessment of the available objectives is to compare the importance of each objective to that of *each other* objective. This method results in the **pairwise comparisons of objectives**.

Returning to the Alice example, the AHP assumes that Alice knows the pairwise comparisons of her objectives. Suppose she values the quality of a university's education twice as much as the financial aid that the university offers, and six times as much as the quality of the university's location. Similarly, financial aid is three times as important to Alice as the quality of location. These numerical values that represent the pairwise comparisons of objectives are placed into a **pairwise comparison matrix** A:

$$A = \begin{bmatrix} 1 & 2 & 6 \\ \frac{1}{2} & 1 & 3 \\ \frac{1}{6} & \frac{1}{3} & 1 \end{bmatrix}.$$

The entry in the i^{th} row and the j^{th} column of matrix A is the importance of objective i compared to that of objective j. For example, $a_{13} = 6$ means that the first objective, academic quality, is six times as important as the third objective, location. It follows that the entries in the diagonal of matrix A are equal to 1. Moreover, if objective i is twice as important as objective j, then objective j must be half as important as objective i. In other words, $a_{ji} = \frac{1}{a_{ij}}$. We also note that if Alice is **consistent** in her preferences, then $a_{ij}a_{jk} = a_{ik}$. That is, if she prefers academic quality twice as much as financial aid, and financial aid three times as much as location, then she must prefer academic quality six times as much as location to be considered a consistent decision maker. According to this definition, Alice in our example is consistent. However, a decision maker in real life is rarely consistent in her preferences. In Chapter 3, we will discuss the process of obtaining the weights of the objectives in the inconsistent case. It is in this inconsistent case that the mathematical reasoning becomes more complicated. For now, we turn our attention to the consistent case.

The next definitions formalize our discussion of the pairwise comparison matrix.

Definition 2.1 (Pairwise Comparison Matrix)**.** The pairwise comparison matrix for a decision maker with n objectives is an $n \times n$ matrix $A = \left[a_{ij}\right]$ such that:

(i) $a_{ij} > 0$ for $i, j = 1, \ldots, n$, and

(ii) $a_{ji} = \frac{1}{a_{ij}}$ for $i, j = 1, \ldots, n$.

A matrix A that satisfies condition (i) is defined to be a **positive** matrix. If A satisfies condition (ii), then it is said to be a **reciprocal** matrix. In the next chapters, when we refer to the pairwise comparison matrix of a decision maker, the assumption is that the matrix is positive and reciprocal. Also note that conditions (i) and (ii) in Definition 2.1 imply that $a_{ii} = 1$ for $i = 1, \ldots, n$.

Definition 2.2 (Pairwise Comparison Matrix of a Consistent Decision Maker)**.** If a decision maker is consistent, then the pairwise comparison matrix A satisfies the conditions in Definition 2.1, and

(iii) $a_{ik} = a_{ij}a_{jk}$ for $i, j, k = 1, \ldots, n$.

Recall that the a_{ij} entry in A represents the importance of objective i, compared to that of objective j. Let w_i be the (unknown for the time being) weight of objective i. We assume each of the weights is positive, and the weights sum to 1. Then for a *consistent* decision maker, the ij entry of A can be written as:

$$a_{ij} = \frac{w_i}{w_j}.$$

We note that the above equality is guaranteed to be true only if the decision maker is consistent. Thus, we have an alternative definition of the pairwise comparison matrix for a consistent decision maker:

Definition 2.3. The pairwise comparison matrix A of a consistent decision maker has the following form:

$$A = \left[a_{ij}\right] = \begin{bmatrix} \frac{w_1}{w_1} & \frac{w_1}{w_2} & \cdots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & \frac{w_2}{w_2} & \cdots & \frac{w_2}{w_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \cdots & \frac{w_n}{w_n} \end{bmatrix},$$

where $w_i > 0$ and $\sum_{i=1}^{n} w_i = 1$.

Definition 2.3 leads to the formal definition of the **weight vector** of the decision maker:

Definition 2.4. The weight vector $\mathbf{w}$ of a decision maker is of the form:

$$\mathbf{w} = [w_i] = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix}^T,$$

where $w_i > 0$ and $\sum_{i=1}^{n} w_i = 1$. The weight vector $\mathbf{w}$ is also referred to as the **priority vector** of the decision maker.

As mentioned in the introduction, the goal of the AHP is to find $\mathbf{w}$. The next theorem guarantees that for a consistent decision maker, we can always obtain this weight vector from the pairwise comparison matrix A .

Theorem 2.1. *Suppose a decision maker is consistent and has n objectives. Let A be the corresponding pairwise comparison matrix, and $\mathbf{w}$ the weight vector. Then $\mathbf{w}$ is an eigenvector of A with corresponding eigenvalue $\lambda = n$.*

Proof. By Definition 2.3 and Definition 2.4

$$A\mathbf{w} = \begin{bmatrix} \frac{w_1}{w_1} & \frac{w_1}{w_2} & \cdots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & \frac{w_2}{w_2} & \cdots & \frac{w_2}{w_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \cdots & \frac{w_n}{w_n} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} \frac{w_1}{w_1}w_1 + \frac{w_1}{w_2}w_2 + \cdots + \frac{w_1}{w_n}w_n \\ \frac{w_2}{w_1}w_1 + \frac{w_2}{w_2}w_2 + \cdots + \frac{w_2}{w_n}w_n \\ \vdots \\ \frac{w_n}{w_1}w_1 + \frac{w_n}{w_2}w_2 + \cdots + \frac{w_n}{w_n}w_n \end{bmatrix}$$

$$= \begin{bmatrix} nw_1 \\ nw_2 \\ \vdots \\ nw_n \end{bmatrix} = n \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = n\mathbf{w}.$$

Thus, $\mathbf{w}$ is an eigenvector of A with corresponding eigenvalue n. □

From Theorem 2.1, we have a way to obtain the weight for each objective, given that we know the number of objectives n and the pairwise comparison matrix A. We know this by observing that:

$$A\mathbf{w} = n\mathbf{w}$$

$$\Leftrightarrow A\mathbf{w} - n\mathbf{w} = 0$$

$$\Leftrightarrow A\mathbf{w} - nI\mathbf{w} = 0$$

$$\Leftrightarrow (A - nI)\mathbf{w} = 0.$$

Thus, $\mathbf{w}$ is in null$(A - nI)$, where I is the identity matrix of an appropriate dimension.

The equation

$$A\mathbf{w} = n\mathbf{w} \tag{2.1}$$

proved in Theorem 2.1 is critical to the AHP. In Chapter 3, we will discuss certain conditions that allow us to apply Equation 2.1 to solve the decision making problem when the decision maker is inconsistent. In Chapter 5, this equation is once again important to our discussion of the decision maker's consistency. Throughout the rest of this thesis, we shall refer to Equation 2.1 as the **eigenvalue problem**.

The purpose of this section has been to present the eigenvalue problem in Equation 2.1. We close the section by the next theorem, which provides a quick and easy way to identify the weight vector when the decision maker is consistent.

Theorem 2.2. *The normalized form of any column of the matrix* $A = \left[\frac{w_i}{w_j}\right]$ *is a solution to the eigenvalue problem* $A\mathbf{w} = n\mathbf{w}$*, where* $\mathbf{w} = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix}^T$ *is the weight vector solution and n is the number of objectives.*

Proof. By Definition 2.3, the j column of A has the form

$$\begin{bmatrix} \frac{w_1}{w_j} & \frac{w_2}{w_j} & \cdots & \frac{w_n}{w_j} \end{bmatrix}^T,$$

for $j = 1, 2, \ldots, n$. Therefore, each column of A is simply a constant multiple of $\mathbf{w}$. It follows that the normalized form of any column of A is a solution to the eigenvalue problem. □

In the next section, we apply both Theorem 2.1 and Theorem 2.2 to find the weights of Alice's objectives. Our expectation is that the methods from both theorems give the same weight vectors.

2.2 Finding the Weight Vector for the Alice Example

In this section, we illustrate the process of finding the weight vectors for the Alice example, using both Theorem 2.1 and Theorem 2.2. By our definition of consistent pairwise comparison matrices, the matrix A in the Alice example, which was given in the last section, is consistent. This guarantees that we can apply both theorems to obtain the weight vectors to solve Alice's problem.

First, applying Theorem 2.1, we know that $\lambda = 3$ is an eigenvalue of A and that $\mathbf{w}$ is in the null space of $A - 3I$. In order to find the solutions of the homogeneous system $(A - 3I)\mathbf{w} = 0$, we first find:

$$A - nI = A - 3I = \begin{bmatrix} 1 & 2 & 6 \\ \frac{1}{2} & 1 & 3 \\ \frac{1}{6} & \frac{1}{3} & 1 \end{bmatrix} - \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} -2 & 2 & 6 \\ \frac{1}{2} & -2 & 3 \\ \frac{1}{6} & \frac{1}{3} & -2 \end{bmatrix}.$$

The reduced row echelon form of $A - 3I$ is: $\begin{bmatrix} 1 & 0 & -6 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{bmatrix}$. Using Gauss-Jordan elimination [8, p. 78], the eigenvectors of A with corresponding eigenvalue 3 have the form $\begin{bmatrix} 6s & 3s & s \end{bmatrix}^T$ $(s \in \mathbb{R})$.

It can be easily seen that an eigenvector of A with corresponding eigenvalue 3 is $\begin{bmatrix} 6 & 3 & 1 \end{bmatrix}^T$. Since the weights of all of the objectives sum to 1, the weight vector is

$$\mathbf{w} = \begin{bmatrix} \frac{6}{10} & \frac{3}{10} & \frac{1}{10} \end{bmatrix}^T.$$

According to this weight vector, the weights of the three objectives—academic quality, financial aid, and location—are 0.6, 0.3, and 0.1, respectively. This weight vector is indeed a constant multiple of any of the columns in A. So the methods of Theorem 2.1 and Theorem 2.2 yield the same result.

In this chapter, we have discussed the components essential to the theoretical reasoning of the AHP. The discussion of the pairwise comparison matrix at the beginning of the chapter brought about the bifurcation of the AHP problem into the consistent and the inconsistent sub-problems. We also justified the existence of a solution to the AHP in the consistent case by the eigenvalue problem. In the next chapter, we shall see how this problem is utilized in solving for the weight vector in the inconsistent case.

Chapter 3

The Inconsistent Decision Maker

3.1 From Consistent to Inconsistent

In reality, decision makers are usually inconsistent. For example, if a prospective college student prefers academic quality twice as much as financial aid, and financial aid three times as much as location, it is unlikely that she will prefer academic quality six times as much as location. As a result, for the $n \times n$ pairwise comparison matrix $A = \left[a_{ij}\right]$, it no longer holds that $a_{ik} = a_{ij}a_{jk}$ for $i, j, k = 1, \ldots, n$. Therefore, we cannot directly apply the eigenvalue problem presented in the last chapter to the inconsistent case. (We note that A is still a positive, reciprocal matrix. This means that $a_{ij} > 0$ and $a_{ij} = \frac{1}{a_{ji}}$ for all i and j.)

Saaty contends that if the entries of a positive reciprocal matrix change by small amounts, then the eigenvalues of that matrix also change by small amounts [10, p. 51]. In addition, the corresponding eigenvectors do not vary by much. Provided that our decision maker is not too inconsistent, the inconsistent pairwise comparison matrix will not deviate much from the consistent matrix. We can find the weight vector for an inconsistent decision maker based on the weight vector in the *consistent* case. Thus, we want to further explore the eigenvalues and corresponding eigenvectors of a consistent pairwise matrix. The next theorem serves that purpose.

Theorem 3.1. *Suppose a decision maker is consistent and has n objectives. Then the pairwise comparison matrix has a unique largest eigenvalue n. All of the other eigenvalues are zero.*

Proof. In the last chapter, we proved that

$$A\mathbf{w} = n\mathbf{w}.$$

So $\mathbf{w}$ is an eigenvector of A with corresponding eigenvalue $\lambda = n$.

Consider:

$$A\begin{bmatrix} w_1 \\ 0 \\ \vdots \\ 0 \\ -w_n \end{bmatrix} = \begin{bmatrix} \frac{w_1}{w_1} & \frac{w_1}{w_2} & \cdots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & \frac{w_2}{w_2} & \cdots & \frac{w_2}{w_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{w_{n-1}}{w_1} & \frac{w_{n-1}}{w_2} & \cdots & \frac{w_{n-1}}{w_n} \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \cdots & \frac{w_n}{w_n} \end{bmatrix}\begin{bmatrix} w_1 \\ 0 \\ \vdots \\ 0 \\ -w_n \end{bmatrix} = \begin{bmatrix} w_1 - w_1 \\ w_2 - w_2 \\ \vdots \\ w_{n-1} - w_{n-1} \\ w_n - w_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} = 0\begin{bmatrix} w_1 \\ 0 \\ \vdots \\ 0 \\ -w_n \end{bmatrix}.$$

Thus, the vector $\begin{bmatrix} w_1 \\ 0 \\ \vdots \\ 0 \\ -w_n \end{bmatrix}$ is an eigenvector of A with corresponding eigenvalue $\lambda = 0$. Using the same approach, we can prove that the following vectors are the eigenvectors corresponding to the eigenvalue $\lambda = 0$ for A:

$$\underbrace{\begin{bmatrix} w_1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ -w_n \end{bmatrix}, \begin{bmatrix} 0 \\ w_2 \\ 0 \\ \vdots \\ 0 \\ -w_n \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ w_3 \\ \vdots \\ 0 \\ -w_n \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ w_{n-1} \\ -w_n \end{bmatrix}}_{n\text{-}1 \text{ vectors}}.$$

It can be easily seen that these $n-1$ vectors are linearly independent, since none of them could be written as a linear combination of the others. So the basis of the eigenspace associated with the eigenvalue $\lambda = 0$ has at least $n-1$ vectors. In other words, the geometric multiplicity of the eigenvalue $\lambda = 0$, which is the dimension of the eigenspace associated with that eigenvalue, is at least $n-1$. Since the algebraic multiplicity of the eigenvalue $\lambda = 0$, which is the number of times its factor occurs in the characteristic polynomial, is always bigger than or equal to its geometric multiplicity, the factor of the eigenvalue 0 occurs at least $n-1$ times in the characteristic polynomial.

The characteristic polynomial has a degree of n. In the proof of Theorem

2.1, we already proved that $\lambda = n$ is an eigenvalue of A. Therefore, the factor of the eigenvalue $\lambda = 0$ can only occur in the characteristic polynomial $n - 1$ times, since the factor of the eigenvalue $\lambda = n$ has to occur at least once in the characteristic polynomial with degree n. Therefore, the characteristic polynomial of A is $p(\lambda) = \lambda^{n-1}(\lambda - n)$.

Thus, the matrix A has a unique largest eigenvalue $\lambda = n$ and all of the other eigenvalues equal zero. □

From the proof above, when the eigenvalue of A is 0, the corresponding eigenvectors violate the assumption $w_i > 0$. Therefore, the only useful eigenvalue of A is $\lambda = n$. Provided that the decision maker is slightly inconsistent, we expect that A has a unique largest eigenvalue that approximates n. The corresponding eigenvector, denoted $\mathbf{w_0}$, approximates $\mathbf{w}$.

Our problem for the inconsistent case becomes: given a decision maker that is inconsistent in her preferences, find the weight vector $\mathbf{w_0}$ that satisfies

$$A\mathbf{w_0} = \lambda_{max}\mathbf{w_0}, \tag{3.1}$$

where λ_{max} is the unique largest eigenvalue for A. In Section 3.2, we will discuss Perron's theorem, which guarantees that Equation 3.1 always has a unique solution.

3.2 Positive Matrices and Their Eigenvalues

In this section, we discuss Perron's theorem for positive matrices. The proof of this theorem provides the theoretical foundation for the method of finding the weight vector in the inconsistent case. We begin by introducing several concepts and theorems about stochastic matrices. These results will be useful for our proof of Perron's theorem later. The statement and the proof of Perron's theorem are presented in the second half of the section. The idea of the proof of Perron's theorem in this section draws from an outline provided by Saaty [10, p. 170–176].

In Chapter 2, we mentioned the condition that makes a matrix positive. For the purpose of the theorems in this section, we provide a formal definition of the terms **non-negative matrix** and **positive matrix**.

Definition 3.1. A real matrix A is non-negative (or positive) if all of the entries of A are non-negative (or positive). We write $A \geq 0$ (or $A > 0$).

Definition 3.2. A non-negative matrix M is a stochastic matrix if each of the row sums equal 1 [16, p. 189].

In another common definition of stochastic matrix, the entries in each of the columns of M sum to 1. We can also say that the columns of M are probability vectors. An example of a column stochastic matrix is $M = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}$.

Theorem 3.2. *For a positive, $n \times n$, row stochastic matrix M*

$$\lim_{k\to\infty} M^k = \mathbf{e}\mathbf{v},$$

where $\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$ is a positive row vector, $\sum_{i=1}^{n} v_i = 1$, and $\mathbf{e} = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}^T$.

Proof. Theorem 3.2 states that for large k, M^k approaches a matrix with identical rows. In order to prove this theorem, we prove that each column of M^k approaches a column vector with identical entries. Let $\mathbf{y_0} = \begin{bmatrix} y_0^1 & y_0^2 & \cdots & y_0^n \end{bmatrix}^T$ be an arbitrary column vector in $\mathbb{R}^n$. Define $\mathbf{y_k} = M^k\mathbf{y_0}$, with $k = 0, 1, 2, \ldots$. Let a_k and b_k be the maximum and minimum components of $\mathbf{y_k}$, respectively. Further, let α be the minimum entry in M.

The outline of the proof is as follows: (i) first, we demonstrate that the sequences (a_k) and (b_k) are monotone, and (ii) bounded. As a result, (a_k) and (b_k) converge. (iii) Next, we prove that the difference between (a_k) and (b_k) approaches 0 as k approaches infinity. Therefore, (a_k) and (b_k) tend to a common limit, and all of the components in $\mathbf{y_k}$ also approach this limit. So $\mathbf{y_k}$ approaches a column vector with identical entries. (iv) Lastly, we choose $\mathbf{y_0}$ so that for each $\mathbf{y_0}$, $\mathbf{y_k}$ represent a column of M^k. Putting everything together, each column of M^k approaches a column vector whose entries are the same, and M^k approaches a matrix whose rows are identical.

(i) **The sequences (a_k) and (b_k) are monotone:**

We observe that for $k = 0, 1, 2, \ldots$, $\mathbf{y_{k+1}} = M^{k+1}\mathbf{y_0} = MM^k\mathbf{y_0} = M\mathbf{y_k}$. Let y_{k+1}^i be the i^{th} component of $\mathbf{y_{k+1}}$, y_k^j be the j^{th} component of $\mathbf{y_k}$, and $m_{i,j}$ be the entry in the i^{th} row and j^{th} column of M, we have

$$y_{k+1}^i = \sum_{j=1}^{n} m_{i,j}y_k^j = m_{i,1}y_k^1 + m_{i,2}y_k^2 + \ldots + m_{i,n}y_k^n. \tag{3.2}$$

Without loss of generality, assume that y_k^1 is the maximum component of

$\mathbf{y_k}$. Then $y_k^1 = a_k$. Equation 3.2 could be rewritten as

$$y_{k+1}^i = m_{i,1}a_k + m_{i,2}y_k^2 + \ldots + m_{i,n}y_k^n.$$

Since b_k is the minimum component of $\mathbf{y_k}$,

$$\begin{aligned} y_{k+1}^i &= m_{i,1}a_k + m_{i,2}y_k^2 + \ldots + m_{i,n}y_k^n \\ &\geq m_{i,1}a_k + m_{i,2}b_k + \ldots + m_{i,n}b_k \\ &= m_{i,1}a_k + (1 - m_{i,1})b_k. \end{aligned}$$

The last line was achieved because the entries in each row of M sum to unity. Further, since α is the minimum entry in M, $\alpha \leq m_{i,1}$ and $1 - \alpha \geq 1 - m_{i,1}$. Therefore

$$\begin{aligned} y_{k+1}^i &\geq m_{i,1}a_k + (1 - m_{i,1})b_k \\ &\geq \alpha a_k + (1 - \alpha)b_k. \end{aligned}$$

Similarly, without loss of generality, assume that y_k^2 is the minimum component of $\mathbf{y_k}$. Then $y_k^2 = b_k$. Equation 3.2 could be rewritten as

$$y_{k+1}^i = m_{i,1}y_k^1 + m_{i,2}b_k + m_{i,3}y_k^3 + \ldots + m_{i,n}y_k^n.$$

With the same reasoning as the case above, we could write

$$\begin{aligned} y^i_{k+1} &= m_{i,1}y^1_k + m_{i,2}b_k + \ldots + m_{i,n}y^n_k \\ &\leq m_{i,1}a_k + m_{i,2}b_k + \ldots + m_{i,n}a_k \\ &= m_{i,2}b_k + (1 - m_{i,2})a_k \\ &\leq \alpha b_k + (1 - \alpha)a_k. \end{aligned}$$

We just showed that an arbitrary component of $\mathbf{y_{k+1}}$ has the following bounds:

$$\alpha a_k + (1 - \alpha)b_k \leq y^i_{k+1} \leq \alpha b_k + (1 - \alpha)a_k. \tag{3.3}$$

Since the bounds hold for the largest component of $\mathbf{y_{k+1}}$

$$a_{k+1} \leq \alpha b_k + (1 - \alpha)a_k, \tag{3.4}$$

which is equivalent to

$$a_{k+1} - a_k \leq \alpha(b_k - a_k).$$

Because $b_k \leq a_k$ for $k = 0, 1, 2, \ldots$

$$a_{k+1} - a_k \leq \alpha(b_k - a_k) \leq 0$$

for $k = 0, 1, 2, \ldots$. Thus, (a_k) is decreasing.

Similarly, the bounds in Equation 3.3 hold for the smallest component of

$\mathbf{y_{k+1}}$. As a result

$$b_{k+1} \geq \alpha a_k + (1-\alpha)b_k, \tag{3.5}$$

which is equivalent to

$$b_{k+1} - b_k \geq \alpha(a_k - b_k) \geq 0,$$

for all $k = 0, 1, 2, \ldots$. Thus, (b_k) is increasing.

(ii) **The sequences (a_k) and (b_k) are bounded and convergent:**

We showed that (a_k) is decreasing and (b_k) is increasing. Thus, for $k = 0, 1, 2, \ldots$, $a_k \leq a_0$ and $b_k \geq b_0$.

For $k = 0, 1, 2, \ldots$, $b_k \leq a_k \leq a_0$. Thus, the increasing sequence (b_k) is bounded above by the number a_0.

Similarly, for $k = 0, 1, 2, \ldots$, $a_k \geq b_k \geq b_0$. Therefore, the decreasing sequence (a_k) is bounded below by the number b_0.

By the Monotone Convergence Theorem [1, p. 51], since (a_k) is monotone and bounded, it converges. Likewise, (b_k) converges.

(iii) **The vector $\mathbf{y_k}$ approaches a column vector with identical entries:**

Combining Equation 3.4 and Equation 3.5, we have

$$a_{k+1} - b_{k+1} \leq \alpha b_k + (1-\alpha)a_k - (\alpha a_k + (1-\alpha)b_k),$$

which is equivalent to

$$a_{k+1} - b_{k+1} \leq (1 - 2\alpha)(a_k - b_k). \tag{3.6}$$

Next, we prove that $(a_k - b_k) \leq (1 - 2\alpha)^k(a_0 - b_0)$ by induction.

Choosing $k = 0$ and applying Equation 3.6, we have

$$a_1 - b_1 \leq (1 - 2\alpha)^1(a_0 - b_0),$$

so the base case is satisfied.

Assume $a_n - b_n \leq (1 - 2\alpha)^n(a_0 - b_0)$. Then by Equation 3.6

$$a_{n+1} - b_{n+1} \leq (1 - 2\alpha)(a_n - b_n).$$

By our assumption of the induction hypothesis, the quantity on the left-hand side of the above inequality is less than or equal to $(1 - 2\alpha)(1 - 2\alpha)^n(a_0 - b_0)$. Thus

$$a_{n+1} - b_{n+1} \leq (1 - 2\alpha)^{n+1}(a_0 - b_0).$$

By induction,

$$a_k - b_k \leq (1 - 2\alpha)^k(a_0 - b_0). \tag{3.7}$$

We showed that (a_k) and (b_k) converge. Therefore, the Algebraic Limit Theorem [1, p. 45] implies that $(a_k - b_k)$ also converges. Next, we compute the limit of $(a_k - b_k)$.

Since M is a positive matrix whose row sums are ones, $0 < \alpha < 1$. Thus

$$0 < 2\alpha < 2 \Leftrightarrow -1 < 1 - 2\alpha < 1 \Leftrightarrow |1 - 2\alpha| < 1,$$

which means $\lim_{k\to\infty}(1 - 2\alpha)^k = 0$ [1, p. 56].

From Equation 3.7 and the way we defined a_k and b_k, we have

$$0 \leq a_k - b_k \leq (1 - 2\alpha)^k (a_0 - b_0).$$

We know $lim_{k\to\infty} 0 = 0$ and
$\lim_{k\to\infty}((1 - 2\alpha)^k (a_0 - b_0)) = (a_0 - b_0) \lim_{k\to\infty}((1 - 2\alpha)^k) = (a_0 - b_0)0 = 0.$

Using the Squeeze Theorem [1, p. 49], $\lim_{k\to\infty}(a_k - b_k) = 0$.

We showed that $\lim(a_k)$ and $\lim(b_k)$ exist. Again, by the Algebraic Limit Theorem

$$\lim_{k\to\infty}(a_k - b_k) = 0 = \lim_{k\to\infty}(a_k) - \lim_{k\to\infty}(b_k).$$

Thus, (a_k) and (b_k) approach a common limit. Since a_k and b_k are the maximum and minimum components of $\mathbf{y_k}$, respectively, all of the components in $\mathbf{y_k}$ also approach this limit. In other words, as k approaches infinity, the vector $\mathbf{y_k}$ approaches a column vector whose entries are all the same. Since the decreasing sequence (a_k) is bounded below by b_0 and the increasing sequence (b_k) is bounded above by a_0,

$$lim_{k\to\infty} y_k = C\mathbf{e},$$

where $b_0 \le C \le a_0$ and $\mathbf{e} = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}^T$.

(iv) M^k **approaches a matrix with identical rows:**

Let $\mathbf{y_0} = \begin{bmatrix} y_0^1 & y_0^2 & \cdots & y_0^n \end{bmatrix}^T$, with $y_0^i = 1$ and $y_0^j = 0$ for any $j \neq i$. Then $\mathbf{y_k} = M^k\mathbf{y_0}$ is the i^{th} column of M^k.

We showed that as k approaches infinity, the vector $\mathbf{y_k}$ approaches a column vector with identical entries. Therefore, the i^{th} column of M^k also approaches a column vector with identical entries, as k approaches infinity.

Let $C\mathbf{e} = \begin{bmatrix} c_i & c_i & \cdots & c_i \end{bmatrix}^T$. Since c_i is an entry in M^k and M is positive, c_i is also positive. Let $\mathbf{v}$ be a row vector such that $v_j = c_i$, we have

$$\lim_{k\to\infty} M^k = \begin{bmatrix} c_1 & c_2 & \cdots & c_n \\ c_1 & c_2 & \cdots & c_n \\ \vdots & & \ddots & \vdots \\ c_1 & c_2 & \cdots & c_n \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} = \mathbf{ev}.$$

Since c_i is positive, $\mathbf{v}$ is a positive row vector. The last item in the proof is to show that the sum of all of the components of $\mathbf{v}$ is 1.

Since the rows of M all sum to 1, $M\mathbf{e} = \mathbf{e}$, with $\mathbf{e} = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}^T$. We prove that M^k is also a stochastic matrix by induction.

First, since $M^1\mathbf{e} = \mathbf{e}$ as shown above, the base case is satisfied.

Assume that $M^n\mathbf{e} = \mathbf{e}$. Then $M^{n+1}\mathbf{e} = M(M^n\mathbf{e}) = M\mathbf{e} = \mathbf{e}$.

Therefore, $M^k\mathbf{e} = \mathbf{e}$, by induction. So each of the rows of M^k sums to unity, and therefore $\sum_{j=1}^{n} v_j = 1$.

□

We just proved that the powers of a positive, row stochastic matrix eventually reach a stable stage. This result will be important to the proof of the next lemma, which is in turn useful for the proof of Perron's theorem.

Lemma 3.1. *If A is a positive $n \times n$ matrix, then*

$$\lim_{k \to \infty} \frac{A^k}{\lambda^k} = \mathbf{wv},$$

where λ is a positive constant, $\mathbf{v}$ a positive row vector, and $\mathbf{w}$ is a positive column vector.

Proof. Let S be the set of all non-negative, column n-vectors such that the entries of each of those vectors sum to 1. We denote

$$S = \{\mathbf{x} | \mathbf{x} = [x_i]_{n \times 1}, x_i \geq 0, \sum_{i=1}^{n} x_i = 1, i = 1, 2, \ldots, n\}.$$

For any vector $\mathbf{y}$, define the function

$$f(\mathbf{y}) = \sum_{i=1}^{n} y_i,$$

and the matrix transformation

$$T(\mathbf{y}) = \frac{1}{f(A\mathbf{y})} A\mathbf{y},$$

where A is a positive, $n \times n$ matrix.

The outline of the proof is as follows: (i) first, we demonstrate that the transformation T is a continuous function, and $T(S) \subset S$. That is for an arbitrary $\mathbf{x} \in S$, $T(\mathbf{x}) \in S$. (ii) Next, using the Brouwer Fixed Point Theorem with the function T and the space S, we are able to find a positive fixed point $\mathbf{w}$. (iii) Finally, we apply Theorem 3.2 to prove the proposed limit.

(i) **T is a continuous function which transforms S to S:**

For an arbitrary $\mathbf{x} \in S$, the matrix transformation $A\mathbf{x}$ is continuous, since each component of $A\mathbf{x}$ is a linear function of $x_1, x_2, \ldots, x_n$ and therefore continuous. As a result, $T(\mathbf{x})$ is continuous.

Let $(A\mathbf{x})_i$ be the i^{th} component of $A\mathbf{x}$. Consider

$$\begin{aligned}
\sum_{i=1}^{n}(T(\mathbf{x}))_i &= \sum_{i=1}^{n}(\frac{1}{f(A\mathbf{x})}(A\mathbf{x}))_i \\
&= f(\frac{1}{f(A\mathbf{x})}A\mathbf{x}) \\
&= \frac{(A\mathbf{x})_1}{f(A\mathbf{x})} + \frac{(A\mathbf{x})_2}{f(A\mathbf{x})} + \ldots + \frac{(A\mathbf{x})_n}{f(A\mathbf{x})} \\
&= \frac{1}{f(A\mathbf{x})}\sum_{i=1}^{n}(A\mathbf{x})_i = \frac{1}{f(A\mathbf{x})}f(A\mathbf{x}) \\
&= 1.
\end{aligned}$$

Since A is $n \times n$ and $\mathbf{x}$ is $n \times 1$, $T(\mathbf{x})$ is $n \times 1$.

Further, since $\mathbf{x} \geq 0$ and $\sum_{i=1}^{n} x_i = 1$, $\mathbf{x}$ has at least one positive component. Since A is positive, $A\mathbf{x} > 0$, which leads to $f(A\mathbf{x}) = \sum_{i=1}^{n}(A\mathbf{x})_i > 0$. Therefore, $T(\mathbf{x}) > 0$.

To sum up, for $i = 1, 2, \ldots, n$, $T(\mathbf{x}) = [(T(\mathbf{x}))_i]_{n\times 1}$, $T(\mathbf{x}) > 0$, $\sum_{i=1}^{n}(T(\mathbf{x}))_i = 1$.

Therefore, $T(\mathbf{x}) \in S$. In other words, $T : S \to S$.

(ii) **Find $\mathbf{w}$ and λ such that $A\mathbf{w} = \lambda\mathbf{w}$:**

The space S is a closed bounded $n-1$ dimensional convex set in $\mathbb{R}^n$. Thus S is topologically equivalent to a closed disk in $\mathbb{R}^{n-1}$.

By the Brouwer Fixed Point Theorem [14, p. 277], there exists a fixed point $\mathbf{w} \in S$ such that

$$T(\mathbf{w}) = \mathbf{w}.$$

By our definition of T, the above equation could be written as

$$\frac{1}{f(A\mathbf{w})} A\mathbf{w} = \mathbf{w}, \tag{3.8}$$

which is equivalent to

$$A\mathbf{w} = f(A\mathbf{w})\mathbf{w}.$$

Like $\mathbf{x}$, $\mathbf{w}$ is non-negative, with entries that sum to 1, so $\mathbf{w}$ has at least one positive component. So $A\mathbf{w} > 0$. Therefore, the right-hand side of the above equation has to be positive. Since we know that $\mathbf{w}$ is non-negative, $f(A\mathbf{w}) > 0$ and therefore $\mathbf{w} > 0$.

Set $\lambda = f(A\mathbf{w})$. Then we have

$$A\mathbf{w} = \lambda\mathbf{w}, \tag{3.9}$$

where λ is a positive real number and $\mathbf{w}$ is a positive column n-vector.

(iii) **Proof of the proposed limit:**

Let D be a diagonal matrix, $d_{ii} = w_i$ and $d_{ij} = 0$ whenever $i \neq j$ (w_i is the i^{th} component of $\mathbf{w}$). Then it follows that $\mathbf{w} = D\mathbf{e}$, with $\mathbf{e} = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}^T$.

For an arbitrary column n-vector $\mathbf{z}$, $D\mathbf{z} = \mathbf{0}$ has only the trivial solution $\mathbf{z} = \mathbf{0}$. Therefore, D is invertible. D^{-1} is a diagonal matrix, with diagonal entries $\frac{1}{w_i}$. It follows that $D^{-1}\mathbf{w} = \mathbf{e}$.

We observe that

$$(D^{-1}(\frac{1}{\lambda})AD)\mathbf{e} = (D^{-1}(\frac{1}{\lambda})A)D\mathbf{e} = D^{-1}(\frac{1}{\lambda})A\mathbf{w} = D^{-1}\mathbf{w} = \mathbf{e}.$$

Thus, $D^{-1}(\frac{1}{\lambda})AD$ is a row stochastic matrix. Further, since D^{-1} and D are non-negative $n \times n$ matrices, A is a positive $n \times n$ matrix, and λ is positive, $D^{-1}(\frac{1}{\lambda})AD$ is a positive $n \times n$ matrix. Using Theorem 3.2,

$$\lim_{k \to \infty} (D^{-1}(\frac{1}{\lambda})AD)^k = \mathbf{e}\mathbf{v}^*,$$

where $\mathbf{v}^* = \begin{bmatrix} v_i^* \end{bmatrix}_{1 \times n} > 0$, and $\sum_{i=1}^{n} v_i^* = 1$.

On the other hand,

$$\begin{aligned} lim_{k \to \infty}(D^{-1}(\frac{1}{\lambda})AD)^k &= \lim_{k \to \infty} \underbrace{(D^{-1}(\frac{1}{\lambda})AD)(D^{-1}(\frac{1}{\lambda})AD) \ldots (D^{-1}(\frac{1}{\lambda})AD))}_{kterms} \\ &= \lim_{k \to \infty} D^{-1}(\frac{1}{\lambda}A)^k D \\ &= D^{-1}(\lim_{k \to \infty} \frac{1}{\lambda^k} A^k)D. \end{aligned}$$

Therefore,

$$D^{-1}(\lim_{k\to\infty} \frac{1}{\lambda^k} A^k)D = \mathbf{e}\mathbf{v}^*$$
$$\Leftrightarrow \lim_{k\to\infty} \frac{A^k}{\lambda^k} = D\mathbf{e}\mathbf{v}^* D^{-1} = \mathbf{w}\mathbf{v}^* D^{-1}.$$

Let $\mathbf{v} = \mathbf{v}^* D^{-1}$. Since $\mathbf{v}^*$ is a positive, row n-vector and D^{-1} is a non-negative, $n \times n$ matrix, $\mathbf{v}$ is a positive row n-vector.

We have shown that for a positive $n \times n$ matrix A,

$$\lim_{k\to\infty} \frac{A^k}{\lambda^k} = \mathbf{w}\mathbf{v},$$

where λ is a positive real number, $\mathbf{v}$ a positive row vector, and $\mathbf{w}$ is a positive column vector.

□

We are now ready to state and prove Perron's theorem, which justifies the existence of a unique largest eigenvalue and a corresponding eigenvector $\mathbf{w_0}$ that satisfies Equation 3.1.

Theorem 3.3 (Perron's Theorem)**.** *Let A be a positive matrix. Then*

1. *A has a real positive simple (i.e., not multiple) eigenvalue λ_{max}, whose modulus is larger than the modulus of any other eigenvalues.*

2. *Each of the right and left eigenvectors of A corresponding to the eigenvalue λ_{max} has positive components, and is essentially (to within multiplication by a constant) unique.*

3. *The number λ_{max} (also called the Perron root of A) is bounded above by the maximum row (or column) sum of A, and bounded below by the minimum row (or column) sum of A.*

Proof. We claim that for a matrix $A > 0$, the real number λ and the vectors $\mathbf{w}$ and $\mathbf{v}$ constructed in Lemma 3.1 satisfy Perron's theorem. We present the proof in the following steps: (i) We first prove that λ is a eigenvalue of A with corresponding right eigenvector $\mathbf{w}$ and left eigenvector $\mathbf{v}$, with λ, $\mathbf{w}$, and $\mathbf{v}$ constructed in Lemma 3.1. (ii) Next, we prove that λ is the unique, largest eigenvalue of A, which is the first item in Perron's theorem. (iii) The proof that $\mathbf{w}$ and $\mathbf{v}$ are unique will be presented next. (iv) Lastly, we prove that λ is bounded above by the maximum row (or column) sum, and bounded below by the minimum row (or column) sum of A.

(i) **The real number λ is an eigenvalue of A with corresponding right eigenvector $\mathbf{w}$ and left eigenvector $\mathbf{v}$:**

From Equation 3.9, we know that λ is an eigenvalue of A with corresponding right eigenvector $\mathbf{w}$.

Using the result of Lemma 3.1

$$\frac{1}{\lambda}\mathbf{w}\mathbf{v}A = \frac{1}{\lambda}\lim_{k\to\infty}\frac{A^k}{\lambda^k}A = \lim_{k\to\infty}\frac{A^{k+1}}{\lambda^{k+1}} = \mathbf{w}\mathbf{v}.$$

Therefore

$$\mathbf{w}\mathbf{v}A = \lambda\mathbf{w}\mathbf{v},$$

which is equivalent to

$$\mathbf{e}\mathbf{w}\mathbf{v}A = \lambda\mathbf{e}\mathbf{w}\mathbf{v},$$

in which $\mathbf{e} = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}$ is a row n-vector. Since $\mathbf{ew}$ is a constant, we have

$$\mathbf{v}A = \lambda \mathbf{v}, \tag{3.10}$$

which means that $\mathbf{v}$ is a left eigenvector of A with eigenvalue λ.

We just showed that $\mathbf{w}$ and $\mathbf{v}$ are the right and left eigenvectors of A, respectively, with eigenvalue λ. The vectors $\mathbf{w}$ and $\mathbf{v}$, as well as λ, are constructed in Lemma 3.1. In Lemma 3.1, we proved that λ is a positive real number, $\mathbf{w}$ is a positive column vector and $\mathbf{v}$ is a positive row vector.

(ii) **The eigenvalue λ is the unique, largest eigenvalue of A:**

Suppose h is another eigenvalue of A with corresponding eigenvector $\mathbf{u}$. Then $A\mathbf{u} = h\mathbf{u}$. For any positive integer k, $A^k\mathbf{u} = h^k\mathbf{u}$ [8, p. 307]. This means $\frac{1}{\lambda^k}A^k\mathbf{u} = (\frac{h}{\lambda})^k\mathbf{u}$.

Taking the limit of both sides of the above equality and using the result of Lemma 3.1, we have: $\mathbf{wvu} = \lim_{k\to\infty}(\frac{h}{\lambda})^k\mathbf{u}$.

Since $\mathbf{wvu}$ is an column n-vector, the limit on the right-hand side of the above equality has to exist, which means that $\lim_{k\to\infty}(\frac{h}{\lambda})^k$ has to exist. Therefore, $|\frac{h}{\lambda}| < 1$, which is equivalent to

$$|h| < |\lambda|,$$

in which case $\lim_{k\to\infty}(\frac{h}{\lambda})^k = 0$. Thus, λ is the unique, largest eigenvalue of A.

(iii) **The right and left eigenvectors $\mathbf{w}$ and $\mathbf{v}$ of A corresponding to λ are**

unique to within multiplication by a constant:

We first prove that $\mathbf{w}$ is unique to within multiplication by a constant.

Suppose $\mathbf{u}$ is another right eigenvector of A corresponding to the eigenvalue λ. Then $A\mathbf{u} = \lambda\mathbf{u}$. For any positive integer k, $A^k\mathbf{u} = \lambda^k\mathbf{u}$ [8, p. 307]. Therefore $\frac{1}{\lambda^k}A^k\mathbf{u} = \mathbf{u}$.

Taking the limit of both sides of the above equality and using the result of Lemma 3.1, we have: $\lim_{k\to\infty}(\frac{1}{\lambda^k}A^k\mathbf{u}) = \lim_{k\to\infty}\mathbf{u} \Leftrightarrow \mathbf{wvu} = \mathbf{u}$. Since $\mathbf{vu}$ is a constant, we can set $a = \mathbf{vu}$. Thus

$$a\mathbf{w} = \mathbf{u}.$$

Similarly, to prove that any other left eigenvector of A corresponding to λ is a constant multiple of $\mathbf{v}$, suppose $\mathbf{y}A = \lambda\mathbf{y}$. Then $\mathbf{y}A^k = \lambda^k\mathbf{y}$ (for any $k \in \mathbb{Z}, k > 0$)[8, p. 307]. Therefore $\frac{1}{\lambda^k}\mathbf{y}A^k = \mathbf{y}$, which means $\mathbf{y}\frac{1}{\lambda^k}A^k = \mathbf{y}$.

Taking the limit of both sides of the above equation and using the result of Lemma 3.1, we have $\mathbf{y}\lim_{k\to\infty}\frac{A^k}{\lambda^k} = \lim_{k\to\infty}\mathbf{y} \Leftrightarrow \mathbf{ywv} = \mathbf{y}$. Since $\mathbf{yw}$ is a constant, we can set $c = \mathbf{yw}$. Thus

$$c\mathbf{v} = \mathbf{y}.$$

λ is also called the **principal eigenvalue** of A, with corresponding **principal eigenvectors** $\mathbf{w}$ and $\mathbf{v}$. Up to this point, we have proved the first two items in Perron's theorem. What remains to be shown is the upper and lower bounds of λ.

(iv) **The eigenvalue λ is bounded above by the maximum row (or column) sum of A and bounded below by the minimum row (or column) sum of A:**

Let $\mathbf{e} = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}^T$ be a column n-vector. The row sums of A are given by the components of $A\mathbf{e}$. Let M be the maximum row sum and m be the minimum row sum of A. Then

$$m\mathbf{e} \leq A\mathbf{e} \leq M\mathbf{e}. \tag{3.11}$$

We proved that $\mathbf{v}$ is a left eigenvector of A with corresponding eigenvalue λ. From Equation 3.10, we have $\mathbf{v}A\mathbf{e} = \lambda\mathbf{v}\mathbf{e}$.

In addition, from 3.11, we have

$$\mathbf{v}m\mathbf{e} \leq \mathbf{v}A\mathbf{e} \leq \mathbf{v}M\mathbf{e}.$$

Therefore

$$\mathbf{v}m\mathbf{e} \leq \lambda\mathbf{v}\mathbf{e} \leq \mathbf{v}M\mathbf{e}. \tag{3.12}$$

Since $\mathbf{ve}$ is a positive real number (we proved that $\mathbf{v} > 0$), we can divide 3.12 by $\mathbf{ve}$, yielding:

$$m \leq \lambda \leq M.$$

Using similar techniques, we can prove that λ is bounded above by the maximum column sum and bounded below by the minimum column sum. This time, let $\mathbf{e} = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}$ be a row n-vector. Then the column

sums of A are given by the components of $\mathbf{e}A$. Let N be the maximum column sum and n be the minimum column sum of A. Then
$\mathbf{e}n \leq \mathbf{e}A \leq \mathbf{e}N$.

Using Equation 3.9, we have $\mathbf{e}A\mathbf{w} = \mathbf{e}\lambda\mathbf{w}$. Therefore, $\mathbf{e}n\mathbf{w} \leq \mathbf{e}\lambda\mathbf{w} \leq \mathbf{e}N\mathbf{w}$, which is equivalent to

$$n \leq \lambda \leq N.$$

□

Perron's theorem guarantees that in the inconsistent case, we can always find the weight vector from a positive, reciprocal pairwise comparison matrix A. (Equation 3.1 has a unique solution). The next theorem provides the justification for the method of finding the weight vector.

Theorem 3.4.

$$\lim_{k\to\infty} \frac{A^k\mathbf{e}}{\mathbf{e}^{\mathbf{T}}A^k\mathbf{e}} = \mathbf{w_1},$$

where $A > 0$, $\mathbf{w_1}$ *is its principal eigenvector corresponding to the maximum eigenvalue* λ_1, *such that* $\sum_{i=1}^{n}(\mathbf{w_1})_i = 1$.

For the proof of this theorem, the reader is referred to [10, p. 176]. Theorem 3.4 states that in order to compute the weight vector of an inconsistent pairwise comparison matrix, we raise the matrix to an arbitrarily large power, and then divide the sum of each row by the sum of the entries in the matrix .

In the next section, we illustrate this method of finding the weight vector for an inconsistent decision maker.

3.3 Finding the Weight Vector for the Inconsistent Alice Example

In this section, we illustrate the use of Theorem 3.4 with the inconsistent version of our Alice example. The consistent pairwise comparison matrix A is modified so that Alice is inconsistent in her preferences:

$$A = \begin{bmatrix} 1 & 2 & 5 \\ \frac{1}{2} & 1 & 3 \\ \frac{1}{5} & \frac{1}{3} & 1 \end{bmatrix}.$$

Since $a_{12}a_{23} = 6 \neq 5 = a_{13}$, A is not consistent. For $k = 1$:

$$A^1\mathbf{e} = A\mathbf{e} = \begin{bmatrix} 1 & 2 & 5 \\ \frac{1}{2} & 1 & 3 \\ \frac{1}{5} & \frac{1}{3} & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ \frac{9}{2} \\ \frac{23}{15} \end{bmatrix},$$

$$\mathbf{e^T}A^1\mathbf{e} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 5 \\ \frac{1}{2} & 1 & 3 \\ \frac{1}{5} & \frac{1}{3} & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{421}{30}.$$

Applying Theorem 3.4 and approximating the result to five decimal places, the eigenvector is

$$\mathbf{w^1} = \frac{A^1e}{e^TA^1e} = \frac{30}{421} \begin{bmatrix} 8 \\ \frac{9}{2} \\ \frac{23}{15} \end{bmatrix} = \begin{bmatrix} 0.57007 \\ 0.32066 \\ 0.10926 \end{bmatrix}.$$

Replicating this process for larger values of k yields:

$$\mathbf{w^2} = \frac{A^2e}{e^TA^2e} = \begin{bmatrix} 0.58176 \\ 0.30896 \\ 0.10928 \end{bmatrix}$$

$$\mathbf{w^3} = \frac{A^3e}{e^TA^3e} = \begin{bmatrix} 0.58157 \\ 0.30898 \\ 0.10945 \end{bmatrix}$$

$$\mathbf{w^4} = \frac{A^4e}{e^TA^4e} = \begin{bmatrix} 0.58155 \\ 0.30900 \\ 0.10945 \end{bmatrix}$$

$$\mathbf{w^5} = \frac{A^5e}{e^TA^5e} = \begin{bmatrix} 0.58155 \\ 0.30900 \\ 0.10945 \end{bmatrix}.$$

The values of the eigenvector have stabilized after five iterations. In general, we stop when

$$||\mathbf{w^i} - \mathbf{w^{i+1}}|| < \epsilon,$$

where $\epsilon > 0$ is predetermined. The weights of Alice's objectives are the entries of $\mathbf{w^5}$: the weight of the first objective is the first entry of $\mathbf{w^5}$, the weight of the second objective is the second entry, and so on. From $\mathbf{w^5}$, we can see that the weights of academic quality, financial aid, and location are 0.58155, 0.30900, and 0.10945, respectively. These weights are close to the weights in the consistent case, which were found in Chapter 2 to be $0.6, 0.3$ and 0.1. This indicates that Alice is not too inconsistent in her preferences. The method to measure the degree of inconsistency will be presented in Chapter 5.

Up until now, we have explained the methods used to find the weight vectors in both the consistent and the inconsistent sub-problems of the AHP. If a decision maker is consistent, Equation 2.1 states that the desired weight vector $\mathbf{w}$ is an eigenvector of A with corresponding eigenvalue n. By Theorem 2.2, we are assured that $\mathbf{w}$ is simply the normalized form of any column in A.

If the decision maker is inconsistent, we find the largest eigenvector of A corresponding to the principal eigenvalue λ_{max}. (In other words, we find a solution to Equation 3.1). Theorem 3.3 (Perron's theorem) guarantees the existence of a unique solution to Equation 3.1. In order to find this unique principal eigenvector, we apply Theorem 3.4: raising A to an arbitrarily large power, and then dividing each row sum by the sum of the entries in the matrix. The iterations are stopped when the difference between two resulting vectors is less than a prescribed value. Since the goal is for the resulting vectors to converge, a quick way to obtain the weight vector is to raise A to the $2 \cdot i$ power at the i iteration, for $i = 1, \ldots, n$ [10, p. 179].

In the next chapter, we turn our attention to the second task of the AHP: finding how well each alternative satisfies each objective.

Chapter 4

Finding the Score of an Alternative on an Objective

In this chapter, we present the method to determine the score of each alternative on each objective. After the scores have been obtained, we will be able to calculate the priority of each alternative. As it turns out, the method to compute the scores of the alternatives is by nature similar to the process used to obtain the weights of the objectives. This leads to the generalization of the AHP, which at its heart involves the construction of a **hierarchy** with different **levels** (or strata) in order to model the various elements in a decision maker's preferences.

4.1 Finding the Score of Each Alternative in the Alice Example

The computation of the scores starts with the construction of a pairwise comparison matrix. In Chapter 2 and Chapter 3, the pairwise comparison matrix has been used to compare each pair of the objectives in terms of their importance to the goal. In this chapter, we construct a pairwise comparison matrix for each objective, assessing how well each alternative satisfies that objective, compared to each other alternative. In other words, we are now interested in the **pairwise comparisons of alternatives** on each objective.

Not surprisingly, we aim to obtain the **score vector s** on an objective from the pairwise comparison matrix for that objective. If the matrix is consistent, we apply Theorem 2.2 to find **s**. If the matrix is inconsistent, we find **s** by using Theorem 3.4. In a nutshell, finding the scores of the alternatives on an objective uses the same process as finding the weights of the objectives on the decision maker's final choice. The only difference is that we have to repeat this process for each objective. We illustrate this process with the Alice example.

Recall that Alice has been accepted into three universities: City University, Suburb University, and Town University. Suppose further that we know how well each university satisfies each objective, compared to how well *each other* university satisfies the same objective. The extent to which alternative i satisfies an objective, compared to the extent to which alternative j satisfies that same objective, is measured on an integer-valued, 1–9 scale.[1] These

[1] This scale is the same scale that was used to measure the pairwise comparisons of objectives in Chapter 2. This scale will be further discussed in Chapter 5.

pairwise comparative scores are placed into a pairwise comparison matrix for each objective. For example, for the first objective, academic quality, the pairwise comparison matrix comparing each pair of the three universities is

$$B_1 = \begin{bmatrix} 1 & 4 & 8 \\ \frac{1}{4} & 1 & 3 \\ \frac{1}{8} & \frac{1}{3} & 1 \end{bmatrix}.$$

The ij entry in B_1 reflects the score of university i on academic quality, compared to the score of university j on the same objective. In this example, suppose we refer to City University, Suburb University, and Town University as the first, second, and third university, respectively. Then the entry in the second row and the first column of B_1, which is $\frac{1}{4}$, means that Suburb University scores one fourth as well as City University on academic quality.

Though B_1 is a positive, reciprocal matrix as required, it is inconsistent, since $b_{12}b_{23} = 4 \cdot 3 = 12 \neq 8 = b_{13}$. Applying Theorem 3.4, we find the score vector $\mathbf{s_1}$ for B_1:

$$\mathbf{s_1^1} = \frac{B_1^1 e}{e^T B_1^1 e} = \begin{bmatrix} 0.69488 \\ 0.22717 \\ 0.07795 \end{bmatrix}$$

$$\mathbf{s_1^2} = \frac{B_1^2 e}{e^T B_1^2 e} = \begin{bmatrix} 0.71788 \\ 0.20459 \\ 0.07753 \end{bmatrix}$$

$$\mathbf{s_1^3} = \frac{B_1^3 e}{e^T B_1^3 e} = \begin{bmatrix} 0.71677 \\ 0.20496 \\ 0.07826 \end{bmatrix}$$

$$\mathbf{s_1^4} = \frac{B_1^4 e}{e^T B_1^4 e} = \begin{bmatrix} 0.71664 \\ 0.20509 \\ 0.07826 \end{bmatrix}$$

$$\mathbf{s_1^5} = \frac{B_1^5 e}{e^T B_1^5 e} = \begin{bmatrix} 0.71665 \\ 0.20509 \\ 0.07826 \end{bmatrix}$$

$$\mathbf{s_1^6} = \frac{B_1^6 e}{e^T B_1^6 e} = \begin{bmatrix} 0.71665 \\ 0.20509 \\ 0.07826 \end{bmatrix}.$$

The values of $\mathbf{s_1}$ are within our convergence tolerance $\epsilon = 0.00001$ after the sixth iteration. The first entry of this vector is the score of the first alternative on the objective academic quality, the second entry corresponds to the score of the second alternative on that objective, and so on. Based on $\mathbf{s_1^6}$, we know that the scores of City University, Suburb University, and Town University on the objective academic quality are $0.71665, 0.20509$, and 0.07826, respectively. Note that just like the weights of the objectives, these scores sum to 1.

We can use the same process to compute the score vectors of the three universities for the other two objectives. Suppose the pairwise comparison matrix comparing each pair of the three universities on the objective financial aid is

$$B_2 = \begin{bmatrix} 1 & \frac{1}{7} & \frac{1}{4} \\ 7 & 1 & 2 \\ 4 & \frac{1}{2} & 1 \end{bmatrix}.$$

Like B_1, B_2 is also inconsistent. Applying Theorem 3.4, we find that the values of the score vector $\mathbf{s_2}$ are within our convergence tolerance $\epsilon = 0.00001$ after the fifth iteration. The score vector for B_2 is

$$\mathbf{s_2^5} = \frac{B_2^5 e}{e^T B_2^5 e} = \begin{bmatrix} 0.08234 \\ 0.60263 \\ 0.31503 \end{bmatrix}.$$

Based on this vector, the scores of City University, Suburb University, and Town University on financial aid are 0.08234, 0.60263, and 0.31503, respectively.

Finally, the pairwise comparison matrix comparing each pair of the three universities on location is

$$B_3 = \begin{bmatrix} 1 & \frac{1}{2} & 4 \\ 2 & 1 & 9 \\ \frac{1}{4} & \frac{1}{9} & 1 \end{bmatrix}.$$

The score vector $\mathbf{s_3}$, whose values are within our convergence tolerance $\epsilon = 0.00001$ after five iterations of Theorem 3.4, is

$$\mathbf{s_3^5} = \frac{B_3^5 e}{e^T B_3^5 e} = \begin{bmatrix} 0.30116 \\ 0.62644 \\ 0.07239 \end{bmatrix}.$$

The interpretation of this score vector is the same as the interpretations of the two previous score vectors. For the objective location, the scores of the City, Suburb, and Town University are 0.30116, 0.62644, and 0.07239, in that order.

Now that we have calculated the weights of the objectives, as well as the scores of the alternatives on the objectives, we can compute the priorities of the three alternatives. Applying Equation 1.1, we find the total score of City University:

$$0.71665 \times 0.58155 + 0.08234 \times 0.30900 + 0.30116 \times 0.10945 = 0.4751728295,$$

the total score of Suburb University:

$$0.20509 \times 0.58155 + 0.60263 \times 0.30900 + 0.62644 \times 0.10945 = 0.3740466175,$$

and the total score of Town University:

$$0.07826 \times 0.58155 + 0.31503 \times 0.30900 + 0.07239 \times 0.10945 = 0.1507794585.$$

We observe that these total scores can be computed by matrix multiplication. We put the three column score vectors into a 3×3 score matrix called S, and then right-multiply this matrix by the weight vector $\mathbf{w}$, computed in Chapter 3. The result is the column 3-vector, whose entries are the priorities of the alternatives. This method provides a short computational way for problems in which the decision makers have a large number of objectives.

$$S\mathbf{w} = \begin{bmatrix} 0.71665 & 0.08234 & 0.30116 \\ 0.20509 & 0.60263 & 0.62644 \\ 0.07826 & 0.31503 & 0.07239 \end{bmatrix} \begin{bmatrix} 0.58155 \\ 0.30900 \\ 0.10945 \end{bmatrix} = \begin{bmatrix} 0.475172829500000 \\ 0.374046617500000 \\ 0.150779458500000 \end{bmatrix}.$$

One way to think of the computation $S\mathbf{w}$ is that we are weighing the scores of each university on the objectives by the importance of those objectives. From the obtained priority matrix, the priorities (or total scores) of City University, Suburb University, and Town University, are 0.48, 0.37, and 0.15, respectively. These scores are the same as the scores calculated separately for each alternative. Based on these scores, Alice should choose to go to City University, as it has the highest total score.

4.2 The Hierarchy

The similarity of the methods used to compute the weights of the objectives and the scores of the alternatives might have led the reader to guess that there is a connection between the roles of the objectives and the alternatives in the AHP. Indeed, this similarity results from the fundamental idea of the AHP: the construction of a **hierarchy**, which consists of different **levels**, in order to reflect the various layers of factors that affect the decision making process.

There is no set of rules that prescribes the construction of a hierarchy. However, it is ideal to choose layers that reflect the decision maker's preferences as well as the complexity of the decision making problem [10, p. 14]. Figure 4.1 shows the hierarchy for the Alice example. The first level is

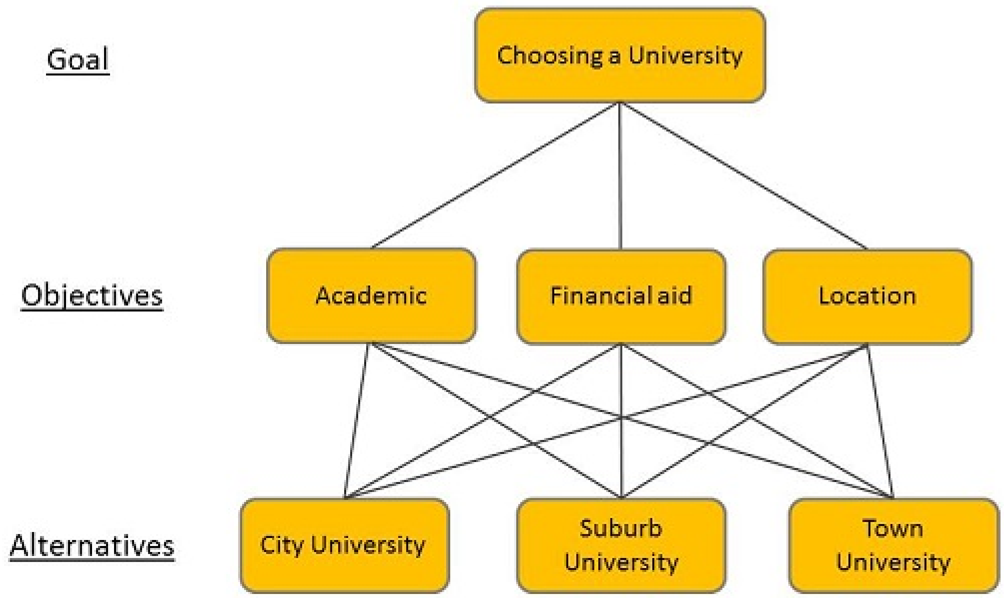

Figure 4.1: The Hierarchy for the Alice Example

the goal of the decision making process, choosing a university. The second level consists of the objectives used when choosing a university, academic quality, financial aid, and location. The third level specifies the alternatives available to Alice: City University, Suburb University, and Town University.

In Figure 4.1, each line segment represents the **influence** (or **impact**) that an **element** of a higher level has on an element of a lower level. Through the intermediary level (which is Objectives in this case), the goal has an impact on the alternatives. In our Alice example, the influence of the second level on the lowest level is how well each alternative satisfies each objective. This influence is expressed mathematically by the score matrix S. The influence of the first level on the second level is the weights of the objectives, represented by the weight vector $\mathbf{w}$. Finally, the overall influence of the first level on the lowest level is the values of the alternatives to Alice. These values are computed by the matrix multiplication $S\mathbf{w}$.

The hierarchy in Figure 4.1 is of the simplest form. It is a linear system, extending from one level down to the next. A more complicated form of the hierarchy is a system with feedback, in which case the levels in the hierarchy interact with one another in a nonlinear manner. An example of this nonlinear system will be presented in Chapter 6. In the next section, we perform a simple extension of the linear hierarchy for the Alice example.

4.3 Extension of the Hierarchy in the Alice Example

The hierarchy for the Alice example can be easily extended to solve a slightly different problem. Suppose Alice is trying to decide between studying Engineering and studying Medicine in college. All of the three universities that she considers offer these two programs, and her choice of major will affect her choice of university. The two majors add a fourth layer to the hierarchy called the Sub-alternatives. Figure 4.2 shows this extended hierarchy.

In Figure 4.2, each alternative has an influence on each of the sub-alternatives. In this example, this influence is interpreted as the academic strength and reputation of each major in each university. Each of the objectives also has an influence on each major through the alternatives in the third level. Likewise, the overall goal has an influence on the majors through the

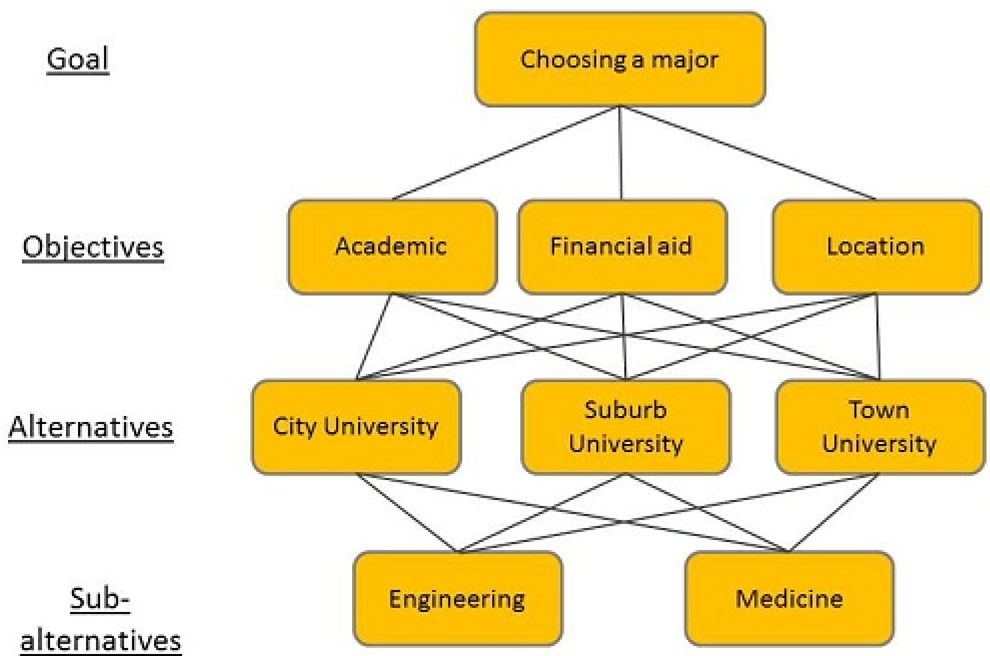

Figure 4.2: The Extended Hierarchy for the Alice Example

objectives in the second level *and* the alternatives in the third level.

We aim to find the score of the sub-alternatives on each alternative. The computation of the score vectors introduced earlier in this chapter is easily applicable to this new layer of sub-alternatives. We begin with a pairwise comparison matrix for each university. The ij entry in each matrix represents the strength of program i, compared to the strength of program j, at the university. In this example, let Engineering be the first, and Medicine be the second program. Suppose that the pairwise comparison matrices for the City, Suburb and Town Universities are C_1, C_2, and C_3, respectively:

$$C_1 = \begin{bmatrix} 1 & 2 \\ \frac{1}{2} & 1 \end{bmatrix} \qquad C_2 = \begin{bmatrix} 1 & 4 \\ \frac{1}{4} & 1 \end{bmatrix} \qquad C_3 = \begin{bmatrix} 1 & \frac{1}{3} \\ 3 & 1 \end{bmatrix}.$$

We observe that all of the three matrices are consistent. This is in fact the case for every 2×2 positive, reciprocal matrix, since it always holds that $a_{ik} = a_{ij}a_{jk}$. The corresponding eigenvectors of these matrices are:

$$\mathbf{s_1} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix} \qquad \mathbf{s_2} = \begin{bmatrix} \frac{4}{5} \\ \frac{1}{5} \end{bmatrix} \qquad \mathbf{s_3} = \begin{bmatrix} \frac{1}{4} \\ \frac{3}{4} \end{bmatrix}.$$

In order to compute the priority vector of the two majors, we first put $\mathbf{s_1}$, $\mathbf{s_2}$, and $\mathbf{s_3}$ into a 2×3 matrix S_0. Right-multiplying S_0 by $S\mathbf{w}$, we obtain the desired priority vector.

$$S_0 S\mathbf{w} = \begin{bmatrix} \frac{2}{3} & \frac{4}{5} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{5} & \frac{3}{4} \end{bmatrix} \begin{bmatrix} 0.71665 & 0.08234 & 0.30116 \\ 0.20509 & 0.60263 & 0.62644 \\ 0.07826 & 0.31503 & 0.07239 \end{bmatrix} \begin{bmatrix} 0.58155 \\ 0.30900 \\ 0.10945 \end{bmatrix} = \begin{bmatrix} 0.65371 \\ 0.34628 \end{bmatrix}.$$

According to this result, the priority of studying Engineering is 0.65, while the priority of studying Medicine is 0.35. Therefore, studying Engineering is the most logical choice for Alice, given her objectives and alternatives.

This example serves its purpose to illustrate an extension of the AHP. However, the example does not provide a good model for a real-world problem. The sub-alternatives have been chosen as if Alice is "partially" going to all of the three universities. Since the calculation of the priorities of the sub-alternatives takes into account the priorities of the alternatives, the priorities of Engineering and of Medicine have been weighed accordingly to the weights of the three Universities. Specifically, the priority of Engineering has been calculated as if Alice would divide her Engineering curriculum into 47% at City University, 37% at Suburb University, and 15% at Town University. The same argument can be made for the calculation of the priority of Medicine.

Further, since we have already determined that City University is the most logical choice for Alice, among the three universities, and Engineering is twice as strong as Medicine at City University (the entry in the first row and the second column of C_1 is 2), we can easily conclude that Alice should study Engineering at City University, without using the pairwise comparison matrices to find S_0. A reasonable argument would be that if the strength of

Medicine greatly exceeds the strength of Engineering at a university that has a lower score than City University, that would tilt the odds in this lower-scored university's favor. We changed the pairwise comparison matrix C_2 for the next best university, Suburb University, so that Medicine is 9 times as strong as Engineering at this university. The AHP recommends that Alice should study Medicine (priority of 0.61, as opposed to 0.39 for Engineering). Again, this recommendation assumes that Alice studies Medicine 47% at City University, 37% at Suburb University, and 15% at Town University, which is a condition that cannot be satisfied in real life.

It is interesting that changing the relative strength of Medicine from $\frac{1}{4}$ to 9 times of Engineering in C_2 does not affect the AHP's recommendation that Alice goes to City University. This strange result is due to the fact that the influence among levels in a hierarchy extends *downward.* In Figure 4.2, the choice of universities affects the choice of majors, but not vice versa. A problem in which the choice of majors affects the choice of universities requires a system with feedback, a generalization of the hierarchy. In Chapter 6, after our understanding of the fundamental components of the AHP has been complete, we will present an example that illustrates this kind of system.

In this section, we have discussed the hierarchy, an essential component of the AHP. As mentioned in the introduction, the hierarchy has been applied to solve complicated problems in the social sciences. For a more detailed analysis of the advantages of hierarchies, we refer the reader to [10, p. 14]. The construction of a hierarchy for each decision making problem requires an in-depth understanding of the various factors of that problem. Saaty proposes a few suggestions on the construction a hierarchy. These suggestions are not based on theories in mathematics and are therefore beyond the scope of this thesis. For more information, the interested reader can refer to [10, p. 14–16].

At this stage, we have investigated most of the foundations of the AHP, including the idea of hierarchy, and the justification and computation of the weights of the elements in a hierarchy. However, we still need to fill some remaining gaps in the theory of the AHP. These are two of the metrics utilized in the AHP: measure of a decision maker's consistency, and the integer-valued 1–9 scale that was used to quantify the pairwise comparisons. We will spend more time developing both of these metrics in the next chapter.

Chapter 5

Metrics

In Chapter 3, we presented the argument made by Saaty [10, p. 179] that when the entries of a positive reciprocal matrix change by small amounts, then the eigenvalues change by small amounts. This argument provides the foundation necessary for us to move from an eigenvalue problem in the consistent case to the same problem in the inconsistent case. In this chapter, we shall provide the theoretical justification for our argument. The central content of this justification is the first metric utilized in the AHP to measure a decision maker's consistency.

As we shall see, a positive, reciprocal matrix A is consistent if and only if the principal eigenvalue λ_{max} equals the number of objectives n. In addition, if the perturbation of the entries a_{ij} is small, and the number of objectives n is also small, then the principal eigenvalue λ_{max} does not deviate much from its original value n. This leads to a method to measure a decision maker's consistency. In the second half of the chapter, we discuss a second metric used in the AHP: the integer-valued 1–9 scale that was used to measure the decision maker's pairwise preferences.

5.1 Measure of Consistency

5.1.1 Consistency of a Pairwise Reciprocal Matrix

In Chapter 2, we defined the pairwise comparison matrix $A = \left[a_{ij}\right]$ to be a positive, reciprocal matrix. The weight vector $\mathbf{w} = \left[w_i\right]$ is a positive, column n-vector, where w_i is the weight of objective i, and $\sum_{i=1}^{n} w_i = 1$.

Recall that in the consistent case, $A_0 = \left[\frac{w_i}{w_j}\right]$, and the weight vector $\mathbf{w_0}$ is the solution to the following equation:

$$A_0\mathbf{w_0} = n\mathbf{w_0},$$

where n is the number of objectives.

For an inconsistent pairwise comparison matrix A, the priority vector $\mathbf{w}$ is the solution to the following equation:

$$A\mathbf{w} = \lambda_{max}\mathbf{w}. \tag{5.1}$$

We found that λ_{max} is the principal eigenvalue of A corresponding to the principal eigenvector $\mathbf{w}$.

The objective of this subsection is to show that if the perturbation of the entries in A from those in A_0 is small, then λ_{max} does not deviate much from n, provided that n is small. (Saaty contends that n should be less than 10 [10, p. 181]). For a discussion of the sensitivity of the eigenvector, see section 7-7 of

[10, p. 192].

We assume that all perturbations of the entries in the original pairwise comparison matrix A can be represented by $a_{ij} = \frac{w_i}{w_j}\epsilon_{ij}$. Specifically, $\frac{w_i}{w_j} + \alpha_{ij} = \frac{w_i}{w_j}(1 + \frac{w_j}{w_i}\alpha_{ij})$. Then A is consistent when $\epsilon_{ij} = 1$. We note that since A is a positive matrix in both the consistent and the inconsistent cases, $\epsilon_{ij} > 0$. Furthermore, since A is reciprocal, $\epsilon_{ji} = \frac{1}{\epsilon_{ij}}$ for all i and j.

From Equation 5.1, for each $i = 1, 2, \ldots, n$:

$$\sum_{j=1}^{n} a_{ij} w_j = \lambda_{max} w_i.$$

Therefore

$$\begin{aligned}
\lambda_{max} &= \sum_{j=1}^{n} a_{ij} \frac{w_j}{w_i} \\
&= a_{i1}\frac{w_1}{w_i} + a_{i2}\frac{w_2}{w_i} + \ldots + a_{ii}\frac{w_i}{w_i} + \ldots + a_{in}\frac{w_n}{w_i} \\
&= \sum_{j \neq i} a_{ij}\frac{w_j}{w_i} + 1,
\end{aligned}$$

for each $i = 1, 2, \ldots, n$.

The last line was achieved because by construction, the entries on the diagonal of the pairwise matrix A are ones. Taking the sum of λ_{max} over n rows yield:

$$n\lambda_{max} = \sum_{j \neq 1} a_{1j}\frac{w_j}{w_1} + 1 + \sum_{j \neq 2} a_{2j}\frac{w_j}{w_2} + 1 + \ldots + \sum_{j \neq n} a_{nj}\frac{w_j}{w_n} + 1.$$

We can rewrite the above equation as:

$$n\lambda_{max} = \sum_{1\leq i<j\leq n} a_{ij}\frac{w_j}{w_i} + \sum_{1\leq j<i\leq n} a_{ij}\frac{w_j}{w_i} + n. \tag{5.2}$$

The first quantity on the right-hand side of Equation 5.2 corresponds to the entries above the diagonal of A, and the second quantity corresponds to the entries below the diagonal. Combining these two quantities, we have

$$n\lambda_{max} = \sum_{1\leq i<j\leq n} (a_{ij}\frac{w_j}{w_i} + a_{ji}\frac{w_i}{w_j}) + n,$$

which is equivalent to

$$n\lambda_{max} - n = \sum_{1\leq i<j\leq n} (a_{ij}\frac{w_j}{w_i} + a_{ji}\frac{w_i}{w_j}). \tag{5.3}$$

The eigenvalue of the consistent pairwise comparison matrix is n, and the eigenvalue of the inconsistent pairwise comparison matrix is λ_{max}. Therefore, $\lambda_{max} - n$ gives an intuitively reasonable measure of consistency. As it turns out, the measure of consistency depends on both $\lambda_{max} - n$ and the number of objectives n.

Define

$$\mu = \frac{\lambda_{max} - n}{n-1}. \tag{5.4}$$

We shall prove later that μ is a measure of consistency.

We have

$$\mu = \frac{\lambda_{max} - 1 + 1 - n}{n-1} = \frac{n(\lambda_{max} - 1)}{n(n-1)} + \frac{1-n}{n-1},$$

and by Equation 5.3

$$\mu = -1 + \frac{1}{n(n-1)} \sum_{1 \le i < j \le n} (a_{ij}\frac{w_j}{w_i} + a_{ji}\frac{w_i}{w_j}). \tag{5.5}$$

We want to express μ in terms of ϵ_{ij} so that μ is tractable for our discussion of the inconsistent case. Since the relation $a_{ji} = \frac{1}{a_{ij}}$ still holds when the decision maker is inconsistent, and by assumption

$$a_{ij} = \frac{w_i}{w_j}\epsilon_{ij}, \tag{5.6}$$

we have

$$a_{ji} = \frac{w_j}{w_i}\frac{1}{\epsilon_{ij}}. \tag{5.7}$$

Substituting Equation 5.6 and Equation 5.7 into Equation 5.5, we have

$$\mu = -1 + \frac{1}{n(n-1)} \sum_{1 \le i < j \le n} (\frac{w_i}{w_j}\epsilon_{ij}\frac{w_j}{w_i} + \frac{w_j}{w_i}\frac{1}{\epsilon_{ij}}\frac{w_i}{w_j}).$$

Thus

$$\mu = -1 + \frac{1}{n(n-1)} \sum_{1 \le i < j \le n} (\epsilon_{ij} + \frac{1}{\epsilon_{ij}}). \tag{5.8}$$

We note that

$$
\begin{aligned}
\lim_{\epsilon_{ij}\to 1} \mu &= -1 + \frac{1}{n(n-1)} \sum_{1\le i<j\le n} \lim_{\epsilon_{ij}\to 1} (\epsilon_{ij} + \frac{1}{\epsilon_{ij}}) \\
&= -1 + \frac{1}{n(n-1)} \sum_{1\le i<j\le n} 2 \\
&= -1 + \frac{1}{n(n-1)} \frac{2n(n-1)}{2} \\
&= 0.
\end{aligned}
$$

The second to last line was achieved because the number of entries that are above the diagonal of A is $\sum_{i=1}^{n-1} i = \frac{n(n-1)}{2}$.

The above limit states that as consistency is approached (ϵ_{ij} approaches 1), μ approaches 0. This suggests that we choose μ as a measure of the decision maker's inconsistency. We will show that the general form of μ in Equation 5.4 suffices as a measure of consistency. In order to accomplish this goal, we need to introduce the next three theorems.

First, we define

$$
\delta_{ij} = \begin{cases} \epsilon_{ij} - 1 & \epsilon_{ij} \ge 1 \\ \epsilon_{ji} - 1 & \epsilon_{ij} < 1. \end{cases}
$$

Since A is reciprocal, if $\epsilon_{ij} < 1$, then $\epsilon_{ji} = \frac{1}{\epsilon_{ij}} > 1$. Thus,

$$
\delta_{ij} \ge 0,
$$

for all i and j. We note that $\delta_{ij} = 0$ when A is consistent (because then

$\epsilon_{ij} = \epsilon_{ji} = 1$). Developing from Equation 5.8 we have:

$$\begin{aligned}
\mu &= \frac{1}{n(n-1)} \sum_{1 \le i < j \le n} (\epsilon_{ij} + \frac{1}{\epsilon_{ij}}) - 1 \\
&= \frac{1}{n(n-1)} (\sum_{1 \le i < j \le n} (\epsilon_{ij} + \frac{1}{\epsilon_{ij}}) - \frac{2n(n-1)}{2}).
\end{aligned}$$

In computing the limit of μ, we found that $\frac{2n(n-1)}{2} = \sum_{1 \le i < j \le n} 2$. Therefore we can write μ as

$$\mu = \frac{1}{n(n-1)} \sum_{1 \le i < j \le n} (\epsilon_{ij} + \frac{1}{\epsilon_{ij}} - 2).$$

If $\epsilon_{ij} > 1$, then $\epsilon_{ij} + \frac{1}{\epsilon_{ij}} = 1 + \delta_{ij} + \frac{1}{1+\delta_{ij}}$.

If $\epsilon_{ij} < 1$, and consequently $\epsilon_{ji} > 1$, then $\epsilon_{ij} + \frac{1}{\epsilon_{ij}} = \frac{1}{\epsilon_{ji}} + \epsilon_{ji} = \frac{1}{1+\delta_{ij}} + 1 + \delta_{ij}$.

Therefore

$$\mu = \frac{1}{n(n-1)} \sum_{1 \le i < j \le n} (1 + \delta_{ij} + \frac{1}{1 + \delta_{ij}} - 2).$$

Further simplification yields

$$\mu = \frac{1}{n(n-1)} \sum_{1 \le i < j \le n} \frac{\delta_{ij}^2}{1 + \delta_{ij}}. \tag{5.9}$$

With this new notion of μ, we can proceed to prove the next theorems, which will be helpful to our discussion of the measure of consistency.

Theorem 5.1. *For an $n \times n$ positive, reciprocal matrix A, the principal eigenvalue $\lambda_{max} \ge n$.*

Proof. From Equation 5.4 and Equation 5.9

$$\frac{\lambda_{max} - n}{n-1} = \mu = \frac{1}{n(n-1)} \sum_{1 \le i < j \le n} \frac{\delta_{ij}^2}{1+\delta_{ij}},$$

which is non-negative, since $\delta_{ij} \geq 0$ by its construction.

Thus,

$$\lambda_{max} \geq n. \tag{5.10}$$

□

Theorem 5.2. *An $n \times n$ positive, reciprocal matrix A is consistent if and only if $\lambda_{max} = n$.*

Proof. If A is consistent, then $\delta_{ij} = 0$ (by construction of δ_{ij}). As a result, $\mu = 0$, by Equation 5.9. This means that $\frac{\lambda_{max}-n}{n-1} = \mu = 0$, by Equation 5.4. Thus, $\lambda_{max} = n$.

Conversely, if $\lambda_{max} = n$, by Equation 5.4, $\mu = \frac{\lambda_{max}-n}{n-1} = 0$. Then, by Equation 5.9, $\frac{1}{n(n-1)} \sum_{1 \le i < j \le n} \frac{\delta_{ij}^2}{1+\delta_{ij}} = \mu = 0$. This implies that $\delta_{ij} = 0$ for each choice of i and j. As a result, A is consistent. □

Theorem 5.3. *Let*

$$\delta = \max_{i,j} \delta_{ij}.$$

Then

$$\lambda_{max} - n < \frac{n-1}{2}\delta^2.$$

Proof. From Equation 5.4 and Equation 5.9

$$
\begin{aligned}
\lambda_{max} - n &= \frac{1}{n} \sum_{1 \le i < j \le n} \frac{\delta_{ij}^2}{1 + \delta_{ij}} \\
&< \frac{1}{n} \sum_{1 \le i < j \le n} \delta_{ij}^2 \\
&\le \frac{1}{n} \sum_{1 \le i < j \le n} \delta^2 \\
&= \frac{1}{n} \frac{n(n-1)}{2} \delta^2 \\
&= \frac{n-1}{2} \delta^2.
\end{aligned}
$$

To arrive at the conclusion above, we have used the fact that $\delta_{ij} > 0$ for some i, j (by construction of δ_{ij}), assuming that A is inconsistent. Consequently, $1 + \delta_{ij} \ge 1$ for all i, j and $1 + \delta_{ij} > 1$ for at least one i, j. Therefore, $\frac{\delta_{ij}^2}{1+\delta_{ij}} < \delta_{ij}^2$. This justifies the second step of our argument. To arrive at the fourth step, we used the fact that the number of the entries above the diagonal of a matrix is $\sum_{i=1}^{n-1} i = \frac{n(n-1)}{2}$. □

Theorem 5.3 states that the deviation of λ_{max} from n depends on the maximum perturbation δ of the entries in A and the number of objectives n. Small perturbations in the entries of A would cause λ_{max} to deviate from n by a small amount. Therefore, we can find the weight vector of the inconsistent case using the same eigenvalue problem in the consistent case, which is $A_0 \mathbf{w_0} = n \mathbf{w_0}$. Theorem 5.3 thus provides the theoretical reasoning for our method in Chapter 3.

However, one issue remains: a decision maker can be extremely inconsistent. In this case, the entries of A will deviate from the entries of a consistent matrix by *large* amounts. The weight vector that results from our approach will be meaningless. As such, it is necessary that we develop a measure of consistency. This measure will provide some insight into the accuracy of the weight vector obtained from our eigenvalue approach. In the next subsection, we will present this measure of consistency.

5.1.2 Consistency Index

We have argued that an $n \times n$ positive, reciprocal matrix A is consistent if and only if its principal eigenvalue λ_{max} equals the number of objectives n. Theorem 5.3 shows that the difference $\lambda_{max} - n$ depends on the magnitude of the maximum perturbation of the entries in A and $n - 1$. For this reason,

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

is used as a measure of the closeness of A to consistency. The quantity μ is called the **consistency index** (C.I.) [10, p. 21].

From Theorem 5.3, we have

$$\mu = \frac{\lambda_{max} - n}{n - 1} < \frac{\delta^2}{2}.$$

Therefore, $\frac{\delta^2}{2}$ provides an upper bound for our measure of the consistency index. However, we note that if n is large, then $\mu < \frac{\delta^2}{2}$ even if λ_{max} is far away from n. Therefore, for a large number of objectives, μ might not provide a meaningful measure of consistency. For n small, μ provides a reasonable measure of how far λ_{max} is from n, and, consequently, how far the decision maker is from consistency. Saaty suggests that n should be less than 10 [10, p. 181].

In order to check for consistency, Saaty uses both the consistency index and another measure called the **random index**. A random sample of 500 pairwise reciprocal matrices is constructed. Each matrix is generated randomly, and its entries are subject to the constraints that

(i) a_{ij} are values from the integer scale 1–9, and

(ii) $a_{ij} = 1$ if $i = j$, and

(iii) $a_{ij} = \frac{1}{a_{ji}}$ for $i, j = 1, 2, \ldots, n$.

The average value of the consistency indexes of these 500 matrices is called the **random index** (R.I.). Table 5.1 gives the random indexes and the corresponding matrix orders [10, p. 21]. Since Saaty suggests using the AHP when the number of objectives is less than 10, this table only lists the R.I. for matrices up to order 10.

Table 5.1: Values of the Random Index (R.I.)

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

The **consistency ratio** (C.R.) of a matrix is the ratio of the C.I. of that matrix to the R.I. for the same matrix order. Thus, for a specific decision maker, the consistency ratio takes into account both the measure of consistency for that specific decision maker (through the consistency index), *and* the measure of consistency of a random sample of 500 other decision makers (through the random index). Therefore, the consistency ratio provides a reasonable measure of the consistency of the decision maker in question.

If the consistency ratio is 0.10 or less, the decision maker is not too inconsistent and the result obtained by the AHP is acceptable. However, if the C.R. is larger than 0.10, more serious inconsistency exists and the priority vector might not provide an accurate solution to the decision making process [17, p. 788]. We illustrate the use of the consistency ratio with our Alice example.

5.1.3 Finding the Consistency Ratio for the Alice Example

In Chapter 3, we discussed the eigenvalue problem of the Alice example:

$$A\mathbf{w_0} = \lambda_{max}\mathbf{w_0}, \tag{5.11}$$

where

$$A = \begin{bmatrix} 1 & 2 & 5 \\ \frac{1}{2} & 1 & 3 \\ \frac{1}{5} & \frac{1}{3} & 1 \end{bmatrix}.$$

We found the weight vector to be

$$\mathbf{w_0} = \frac{A^5 e}{e^T A^5 e} = \begin{bmatrix} 0.58155 \\ 0.30900 \\ 0.10945 \end{bmatrix}.$$

From Equation 5.11 and the rule of matrix multiplication, we know that

$$\lambda_{max} = \sum_{j=1}^{3} a_{ij} \frac{w_j}{w_i},$$

for $i = 1, 2, 3$, $A = \left[a_{ij}\right]$, and $\mathbf{w_0} = [w_i]$. Thus, λ_{max} can be computed as the ratio of any component in $A\mathbf{w_0}$ to the corresponding component in $\mathbf{w_0}$. Computing these ratios for all pairs of components in $A\mathbf{w_0}$ and $\mathbf{w_0}$ yield

$$\sum_{j=1}^{3} a_{1j} \frac{w_j}{w_1} = \frac{1(0.58155) + 2(0.30900) + 5(0.10945)}{0.58155} = 3.003694708,$$

$$\sum_{j=1}^{3} a_{2j} \frac{w_j}{w_2} = \frac{\frac{1}{2}(0.58155) + 1(0.30900) + 3(0.10945)}{0.30900} = 3.003694450,$$

$$\sum_{j=1}^{3} a_{3j} \frac{w_j}{w_3} = \frac{\frac{1}{5}(0.58155) + \frac{1}{3}(0.30900) + 1(0.10945)}{0.10945} = 3.003694636.$$

Taking $\lambda_{max} = 3.003694$, we have the consistency index

$$\frac{\lambda_{max} - n}{n - 1} = \frac{3.003694 - 3}{3 - 1} = 0.001847.$$

For a matrix of order 3, the random index (as in Table 5.1) is 0.58. Therefore, the consistency ratio is

$$\frac{0.001847}{0.58} = 0.003184 < 0.1.$$

So the inconsistency is acceptable and the weight vector $\mathbf{w_0}$ provides an accurate solution to the Alice example.

In this section, we have investigated the theories that allow us to apply the same eigenvalue approach in the consistent case to the inconsistent case. In the process, we also derived the consistency index $\mu = \frac{\lambda_{max}-n}{n-1}$, which is used together with the random index to measure the decision maker's inconsistency. In the next section, we will discuss the integer-valued scale 1–9 that was used to represent the decision maker's preferences for each pair of objectives.

5.2 Measure of Pairwise Preferences

In Chapter 2, we introduced the pairwise comparison matrix A. Recall that the a_{ij} entry in A represents the importance of objective i for the decision maker, compared to that of objective j. In this section, we explain the metric that is used to measure the decision maker's pairwise preferences. We will not aim to investigate the derivation of this metric, since this derivation is based on theories in the social sciences.

The values of these pairwise comparisons are drawn from a scale of integers ranging from 1 to 9. The values of this scale and their interpretations are presented in Table 5.2 [17, p. 787]. The interpretations of the values in this table are modified in wording so that they fit the context of our decision making problem. For the original table by Saaty, see [10, p. 54].

Table 5.2: Interpretations of Entries in a Pairwise Comparison Matrix

Value of a_{ij}	**Interpretations**
1	Objectives i and j are of equal importance.
3	Objective i is weakly more important than objective j.
5	Experience and judgment indicate that objective i is strongly more important than objective j.
7	Objective i is very strongly or demonstrably more important than objective j.
9	Objective i is absolutely more important than objective j.
2, 4, 6, 8	Intermediate values. For example, a value of 8 means that objective i is midway between strongly and absolutely more important than objective j.

Since A is reciprocal, if a_{ij} is assigned one of the values in the above table, a_{ji} is simply the reciprocal of that value. The derivation of Table 5.2 is based on the theories of stimulus and response in psychology. For a discussion of this derivation, we refer the interested reader to the section on scale comparison written by Saaty [10, p. 53–64].

The discussion in this chapter has completed our understanding of the fundamental components of the AHP. At this stage, we have presented the method and theories of the AHP. The next chapter will focus on an application of the AHP in medical diagnosis. As mentioned in Chapter 4, this application illustrates a system with feedback, a generalization of the hierarchy to include feedback among levels of the hierarchy.

Chapter 6

An Application in Medical Diagnosis

In the previous chapters, we have used the simple Alice example to illustrate the method of the AHP. The objective of this chapter is to present a more sophisticated application of the AHP in medical diagnosis. We first introduce a few concepts that are useful to the execution of this application. Such concepts, e.g. the supermatrix, are not fundamental elements of the AHP and therefore are presented in this chapter as extensions of the AHP. The second half of the chapter is dedicated to the application of the AHP in medical diagnosis.

6.1 The Supermatrix

6.1.1 The Supermatrix Approach for the Alice Example

We make use of this section to develop several key ideas for the medical application of the AHP. In Chapter 4, we mentioned that the hierarchy for the Alice example was a linear system: we start with the highest level, and then extend downward from one level to the next. In terms of the relationship between two consecutive levels, the higher has an influence on the lower, as was pointed out in the extended Alice example (where the sub-alternatives were added). However, in a real-world problem, it is possible that the lower level of a hierarchy has an influence on a higher level as well, or the elements in a level have dependent relationships. In fact, the hierarchy for the medical application that will be presented in the next section exhibits both of those characteristics. A hierarchic structure with such nonlinear relationship between layers, or between elements in a layer, is called a **system with feedback**. We will soon demonstrate how a linear hierarchy, such as that for the Alice example, is modified to display the feedback relationship. In order to do that, we first represent the hierarchy in the form of a **network**, which shows clearly the relationship between levels in the hierarchy.

Figure 6.1 shows the network for the Alice example. The network has three **nodes**, which correspond to the three levels in the hierarchy given in Figure 4.1. Each node in turn includes one or more **elements**. For example, the node Goal consists of only one element: the goal of choosing a university, while the node Objectives has three elements: academic quality, financial aid,

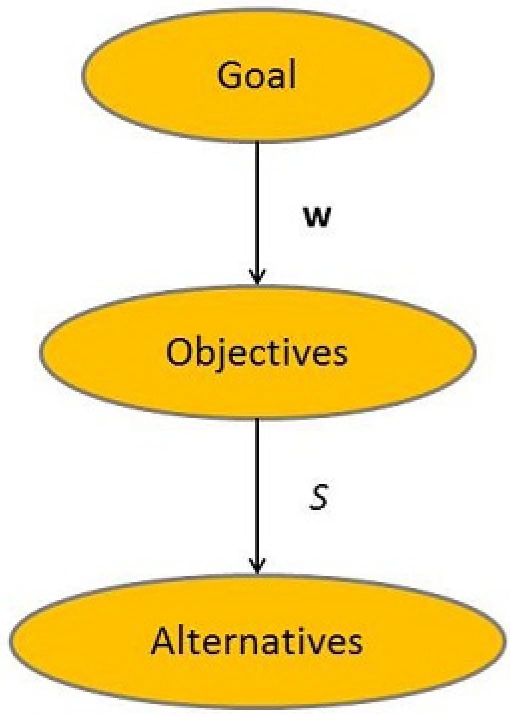

Figure 6.1: Network for the Alice Example

and location. Each arrow in the network represents an influence that the elements in the higher node has on the elements in a lower node. In this example, the influence of the elements in the node Objectives on the elements in the node Alternatives is represented by the scores of the alternatives on the objectives. Similarly, the influence of the element in the node Goal on the elements in the node Objectives corresponds to the weights of the objectives on the goal of choosing a university. In Chapter 4, we expressed these scores and weights as the score matrix S and the weight vector $\mathbf{w}$, respectively.

In order to obtain the priorities of the three alternatives on Alice's goal of choosing a university, we enter S and $\mathbf{w}$ into a matrix W that displays the interactions among elements in the nodes of a network.

$$W = \begin{array}{l} \\ \text{Goal} \\ \text{Objectives} \\ \text{Alternatives} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Goal} & \text{Objectives} & \text{Alternatives} \end{array} \\ \begin{bmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \\ W_{31} & W_{32} & W_{33} \end{bmatrix} \end{array}. \tag{6.1}$$

Equation 6.1 shows the matrix W for a network with three nodes. The construction of W is as follows. Let the first, second, and third node of the network correspond to the first, second, and third row (column) of W, respectively. Then the ij component in W reflects the influence of the elements in the j node on the elements in the i node of the network. By influence, we mean the *priorities* of the elements in the i node, with respect to the elements in the j node of the network.

Each ij component in W is itself a matrix. For this reason, W is referred to as a **supermatrix**. (In the rest of this thesis, whenever we refer to the **components** of W, we mean the matrices that reflect the relationship between elements in the nodes in a network, as opposed to the entries in W.) If there is no dependent relationship between the i and the j nodes of the network, then the ij component in W is the zero matrix. If dependent relationship exists, W_{ij} is nonzero. In the Alice example, the weight vector **w** reflects the influence of the goal on the objectives. Therefore, **w** is entered in the 2,1 position of W. Using the same reasoning, the score matrix S is entered into the 3,2 position of W. The form of the supermatrix for the Alice example is as follows:

$$W = \begin{bmatrix} 0 & 0 & 0 \\ \mathbf{w} & 0 & 0 \\ 0 & S & I \end{bmatrix}. \tag{6.2}$$

Since $\mathbf{w}$ is a weight vector, the entries in $\mathbf{w}$ are positive and sum to 1. Each of the columns in S is itself a weight vector, so S is a column stochastic matrix. By putting the identity matrix in the 3,3 position of W, we have made W a column stochastic matrix. Saaty contends that the powers of W will eventually reach a stable stage, denoted W^k, and that by raising W to powers, we will obtain the desired priorities [13, p. 494]. Since the purpose of this chapter is to illustrate a sophisticated application of the AHP, we will not investigate the theoretical justification of this claim. However, for more theoretical discussion, we refer the reader to [10, p. 206–214] and [11].

The target priority vector can be found as a column in W^k or as a component in W^k, depending on the form of W. Thus, we need several definitions about certain types of matrix that will help us categorize W.

Definition 6.1. A square matrix is *irreducible* (by permutations) if it cannot be decomposed into the form $\begin{bmatrix} A_1 & 0 \\ A_2 & A_3 \end{bmatrix}$, where A_1 and A_3 are square matrices and 0 is the zero matrix. Otherwise the matrix is said to be *reducible* [10, p. 168].

Definition 6.2. A non-negative, irreducible matrix A is *primitive* if and only if there is an integer $m \geq 1$ such that $A^m > 0$. Otherwise A is called *imprimitive* [10, p. 176].

Thus, W can be reducible, irreducible and primitive, or irreducible and imprimitive. If W is reducible, the priorities of the elements in the n node of the network, with respect to the elements in the j node, is given by the component at the n,j position of W^k, where n is the number of nodes in the hierarchy. Thus, the priorities of the elements in the lowest node, with respect to the elements in the remaining nodes, can always be read from the last row of components in W. We note that by component of W^k, we are also referring to a matrix, since W^k is the result of raising the supermatrix W to powers.

If W is primitive, W^k has identical columns. Each of these columns gives the desired priorities. For the theoretical justification of the cases when W is reducible, or primitive, the reader is referred to [13, p. 494] and [10, p. 208–214]. If W is imprimitive, Saaty shows that W can always be made primitive by substituting arbitrarily small positive numbers for the zero entries in W, subject to the condition that W remains column stochastic. More detailed discussion on this topic is available at [11], [2], and [12]. In the remaining of this chapter, we will apply the supermatrix approach to obtain the target priorities both when W is reducible (the Alice example) and when it is primitive (the application in medical diagnosis).

Substituting the entries of S and $\mathbf{w}$ into W in Equation 6.2, we have the supermatrix for the Alice example, where the entries in $\mathbf{w}$ are in bold, and the entries in S are italicized:

$$W = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0.58155} & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0.30900} & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0.10945} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathit{0.71665} & \mathit{0.08234} & \mathit{0.30116} & 1 & 0 & 0 \\ 0 & \mathit{0.20509} & \mathit{0.60263} & \mathit{0.62644} & 0 & 1 & 0 \\ 0 & \mathit{0.07826} & \mathit{0.31503} & \mathit{0.07239} & 0 & 0 & 1 \end{bmatrix}.$$

It is obvious that W is reducible. W can be decomposed into the form $\begin{bmatrix} A_1 & 0 \\ A_2 & A_3 \end{bmatrix}$, where $A_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0.58155 & 0 & 0 & 0 \\ 0.30900 & 0 & 0 & 0 \\ 0.10945 & 0 & 0 & 0 \end{bmatrix}$ is a square matrix, and A_3 is the 3×3 identity matrix.

Raising W to powers, we find that the entries in W stabilize after three iterations. In other words, W^3 is the stable form of W:

$$W^3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0.47517282950} & 0.71665 & 0.08234 & 0.30116 & 1 & 0 & 0 \\ \mathbf{0.37404661750} & 0.20509 & 0.60263 & 0.62644 & 0 & 1 & 0 \\ \mathbf{0.15077945850} & 0.07826 & 0.31503 & 0.07239 & 0 & 0 & 1 \end{bmatrix}.$$

The component in the 3,1 position of W^3 (in bold) gives the priorities of the elements in the last node (Alternatives) of the network, with respect to the element in the first node (Goal). These priorities are precisely what we found by multiplying $S\mathbf{w}$ in Chapter 4. The component in the 3,2 position of W^3 gives the priorities of the alternatives with respective to the objectives. This component is exactly the score matrix S that we found in Chapter 4. For the Alice example, raising the supermatrix W to powers yields the same result as successively weighing the elements in a hierarchy.

The purpose of our discussion thus far has been to introduce the network and the supermatrix, so that we can utilize them to solve a decision making problem in a more complex system. The next subsection will present the supermatrix approach in light of a system with feedback.

6.1.2 The Supermatrix Approach for a System with Feedback

We open this section with a brief discussion of the usefulness of systems with feedback, which itself necessitates the supermatrix approach to obtain the desired priority vector. Saaty contends that many problems in the social sciences have such complex situations that the linear form of a hierarchy fails to capture. An example is the various forms of organizations that cannot be put in a hierarchic structure [10, p. 199].

In both the initial and the extended versions of the Alice example, the desired priority vector was obtained by matrix multiplication. In other words, the priorities of the elements in the lowest level, with respect to the element in the highest level of the hierarchy, were obtained by successively weighing the priorities of the elements in each level with respect to the elements in the level immediately above. In the initial Alice example, the priorities of the three alternatives on the goal were obtained by weighing the priorities of the alternatives on the objectives, and then the priorities of the objectives on the goal. In terms of matrix multiplication, this successive weighing process is represented by $S\mathbf{w}$. However, this approach is impossible to execute in a system with feedback, since we no longer have a hierarchy extending from one level down to the next. We now have a network, in which it is no longer clear which node is lower and which is higher. Furthermore, there might be more than one interaction between the elements in two nodes. The supermatrix comes to our rescue. A supermatrix enables us to express and compute two-way interactions between elements in different nodes, as well as the relationship among elements in the same node.

Figure 6.2 illustrates the network for a system with feedback. As in the network for the Alice example, each arrow represents the influence that the elements in the node at the starting point of the arrow has on the elements in the node at the ending point of the arrow. In this network, the elements in the node Objectives exhibit dependent relationship. The matrix multiplication approach used in Chapter 4 has no procedure to take into account this relationship while deriving the priorities of the alternatives. The supermatrix approach, however, offers a solution to this problem. Let us take a look at the corresponding supermatrix for this system with feedback:

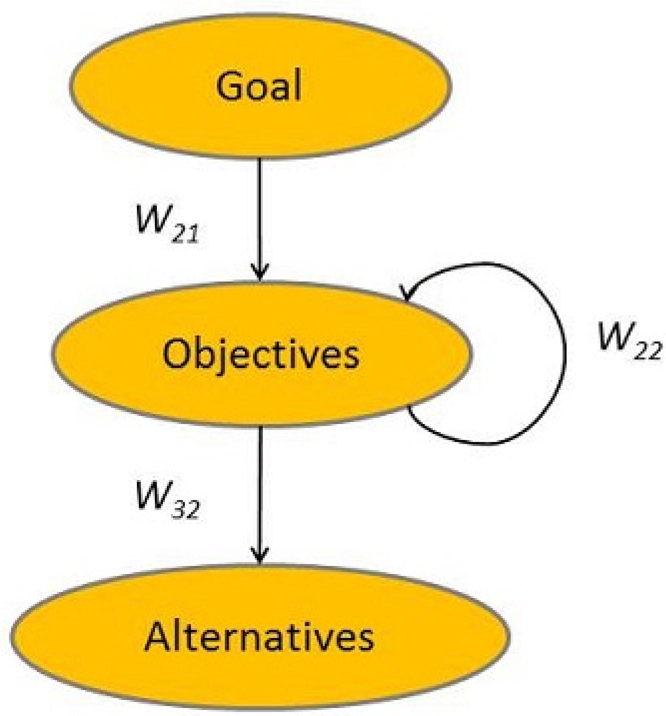

Figure 6.2: Network for a System with Feedback

$$W_0 = \begin{array}{r} \\ \text{Goal} \\ \text{Objectives} \\ \text{Alternatives} \end{array} \!\!\! \begin{array}{c} \begin{array}{ccc} \text{Goal} & \text{Objectives} & \text{Alternatives} \end{array} \\ \left[\begin{array}{ccc} 0 & 0 & 0 \\ W_{21} & W_{22} & 0 \\ 0 & W_{32} & I \end{array}\right] \end{array}.$$

The construction of W_0 follows the same rule that we used to construct W in Equation 6.1: the component at the i,j position of W_0 reflects the priorities of the elements in the i node of the network, with respect to the elements in the j node. To express the dependent relationship among elements in the i node, we use the component W_{ii}. A zero matrix in the ij component of W_0 signifies that there is no relationship between the elements in the i node and those in the j node. We note that the 3,3 position of W is the identity matrix, instead of the zero matrix, even though the alternatives are independent of one another. The purpose is that the transformed form of W will be column stochastic. From here on, any recurrence of the placement of I in the component in the last row and the last column of the supermatrix will serve the same purpose.

This rule for construction of the supermatrix extends to the general case, as given in the next definition.

Definition 6.3. Let N be a network with n nodes, denoted $N_1, N_2, \ldots, N_n$. The supermatrix W for the network N is:

$$W = \begin{array}{c} \\ N_1 \\ N_2 \\ \vdots \\ N_n \end{array} \begin{array}{c} \begin{array}{cccc} N_1 & N_2 & \cdots & N_n \end{array} \\ \begin{bmatrix} W_{11} & W_{12} & \cdots & W_{1n} \\ W_{21} & W_{22} & \cdots & W_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ W_{n1} & W_{n2} & \cdots & W_{nn} \end{bmatrix} \end{array},$$

where each W_{ij} is a matrix [10, p. 207].

We recall that the network for a hierarchy is of a linear form, in which there are only one-way interactions, extending from the highest node down to the lowest node. Thus, the supermatrix for a hierarchy always has the form

$$W^H = \begin{array}{c} \\ N_1 \\ N_2 \\ N_3 \\ \vdots \\ N_{n-1} \\ N_n \end{array} \begin{array}{c} \begin{array}{ccccccc} N_1 & N_2 & N_3 & \cdots & N_{n-2} & N_{n-1} & N_n \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ W_{21} & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & W_{32} & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & W_{n-1,n-2} & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & W_{n,n-1} & I \end{bmatrix} \end{array},$$

where each W_{ij}^H is a matrix [10, p. 209].

As in the case with W, each component of W^H is a matrix. We can obtain the target priorities by looking at the last row of components in the stable form of W^H, as illustrated in the Alice example. We observe that W^H is simply a special case of W defined in Definition 6.3. More generally, our method of using the hierarchy to find the desired priorities is a special case of the supermatrix approach.

In this section, we have introduced the network, as well as the important supermatrix approach, which allows us to solve problems in systems with feedback. We also noted that the matrix multiplication method to obtain target priorities is a special case of the supermatrix approach. Before ending this section, we make an important observation that W might not be column stochastic, even though each of its components is a column stochastic matrix. An example is the matrix W_0 introduced in this section. In order to apply the method of raising W to powers to obtain the target priorities, we need to transform W into a column stochastic matrix. This is achieved by deriving the pairwise priorities of the appropriate *nodes*, and then weighing each component in those nodes by the priorities. The execution of this idea will be illustrated in the case study in medical diagnosis in the next section.

6.2 A Case Study in Medical Diagnosis

6.2.1 Preliminary Analysis of the Case Study

In this section, we apply the ideas developed in the first half of the chapter to solve a problem in medical diagnosis. The case study in this section, as well as the data on the pairwise comparison matrices and the supermatrix, is drawn from an article by Saaty [13].

The medical case is described as follows.

Case Study: A woman in her second trimester of pregnancy was admitted to a local hospital. The tests at the hospital revealed the following seven symptoms:

- Anemia (An),
- Low Platelets (LP),
- Abnormal Liver (ABL),
- Blood Clotting (BC),
- High Activated Partial ThromboPlastin Time (APTT-H),
- High AntiNuclear Antibody (ANA-H), and
- High AntiCardiolipin Antibody (ACA-H).

At this point the physicians considered four possible diseases that could cause the symptoms:

- Lupus,
- Thrombotic Thrombocytopenic Purpura (TTP),
- Hemolysis, Elevated Liver function, and Low Platelets (HELLP), and
- AntiCardiolipin Antibody Syndrome (ACA Syn).

For more information on these medical terms, the reader can consult the encyclopedic reference [3].

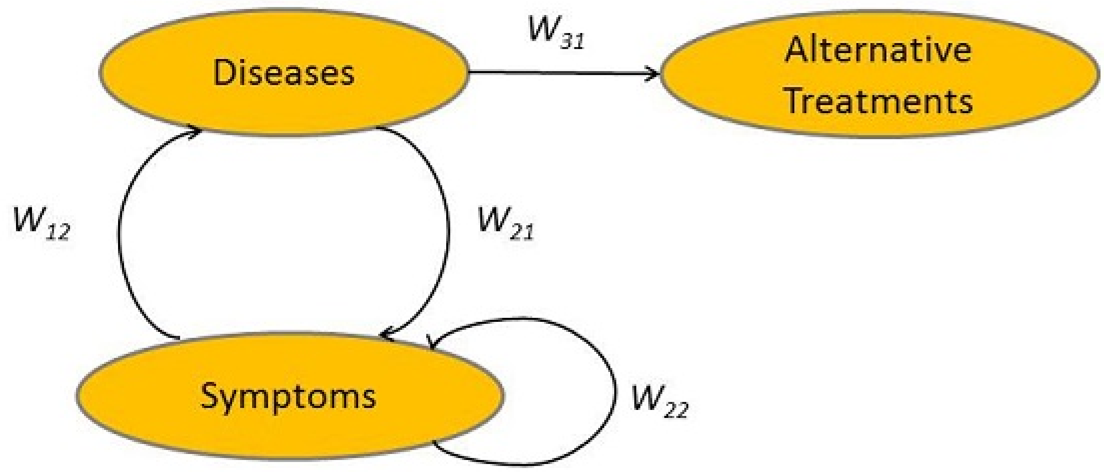

Figure 6.3: Network for the Case Study in Medical Diagnosis

Given the condition of the patient, the physicians need to make a decision between two alternatives: to terminate the pregnancy, or to treat the patient for her symptoms and let her proceed with the pregnancy. The network for this case study is given in Figure 6.3.

As in the previous networks, the arrow reflects the influence of the elements in one node on the elements in another, or the same, node. As we can see in Figure 6.3, the diseases and the symptoms have two-way interactions, which would be impossible to depict a linear hierarchic structure. Furthermore, each symptom also has an influence on other symptoms. This dependent relationship among symptoms is in fact the reason for the desirability of the AHP in solving this problem. An alternative approach, Bayes' theorem, has been particularly popular in medical diagnostics. (See, for example, [5], [15], and [9].) However, the use of Bayes' theorem in diagnosing diseases requires an assumption that is seldom satisfied in real life: the symptoms are independent of one another [6].

Saaty suggests that the reason for this assumption is the lack of information on the relationship of symptoms, as well as the daunting task of conducting a sufficient number of experiments to obtain the necessary statistical data to apply Bayes' theorem [10, p. 492]. Using the AHP, on the other hand, allows the incorporation of physicians' judgment in order to take into account the relationship among symptoms.

From Figure 6.3, the interactions among the elements in the nodes of the network for the case study are described as follows.

(i) Each symptom observed in the patient has an influence on the possible diseases. This influence is interpreted as the likelihood of the diseases, given the symptom;

(ii) Each possible disease has an influence on the observed symptoms. This influence means the extent to which the symptoms are characteristic of the disease;

(iii) Each observed symptom has an influence on other observed symptoms. This influence means the likelihood that a given symptom is associated with or occurs jointly with other symptoms;

(iv) Each possible disease has an influence on the alternative treatments. This influence is precisely the priority of the alternative treatments, given the possible diseases.

We proceed to construct a supermatrix, whose components represent these interactions among the elements in the nodes of the network. Let the nodes Diseases, Symptoms and Alternative Treatments correspond to the first, second, and third row (column) of W. Then using Definition 6.3, the supermatrix for this case study is:

$$W = \begin{array}{l} \\ \text{Diseases} \\ \text{Symptoms} \\ \text{Alternatives} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Diseases} & \text{Symptoms} & \text{Alternatives} \end{array} \\ \begin{bmatrix} 0 & W_{12} & 0 \\ W_{21} & W_{22} & 0 \\ W_{31} & 0 & I \end{bmatrix} \end{array}.$$

We would like to obtain the priorities of the alternative treatments from W. Since W is not column stochastic, we cannot directly apply the method of raising W to powers introduced in the last section. However, we can transform W into a column stochastic matrix by the following method.

Consider the first column of components in W, which consists of the zero matrix, W_{21}, and W_{31}. Let α_1 be the weight of the node Symptoms with respect to the node Diseases, and α_2 the weight of Alternatives with respect to Diseases, such that α_1 and α_2 are positive real numbers, and $\alpha_1 + \alpha_2 = 1$. We obtain α_1 and α_2, then multiply each entry in W_{21} by α_1, and each entry in W_{31} by α_2. Recall that each of W_{21} and W_{31} is a matrix of weight vectors and is therefore column stochastic. By multiplying the entries in W_{21} and W_{31} by α_1

and α_2, we will make the block $\begin{bmatrix} 0 \\ W_{21} \\ W_{31} \end{bmatrix}$ in W column stochastic.

To obtain α_1 and α_2, we construct a pairwise comparison matrix M for Diseases:

$$M = \begin{array}{c} \text{Diseases} \\ \text{Symptoms} \\ \text{Alternatives} \end{array} \begin{array}{c} \begin{array}{cc} \text{Symptoms} & \text{Alternatives} \end{array} \\ \begin{bmatrix} 1 & m_{12} \\ m_{21} & 1 \end{bmatrix} \end{array},$$

where the ij entry in M is the weight of the i node with respect to Diseases, compared to the weight of the j node with respect to Diseases. In particular, the entry m_{21} reflects the importance of knowing about the alternatives, *compared to the importance of knowing about the symptoms*, in diagnosing the disease(s). Then the eigenvector $\mathbf{u} = \begin{bmatrix} u_1 & u_2 \end{bmatrix}^T$ of M such that $u_1 + u_2 = 1$ gives the weights of the nodes Symptoms and Alternatives on the node Diseases. By our construction of M, we have that $u_1 = \alpha_1$ and $u_2 = \alpha_2$.

The reader must have realized that this process of finding the weight vector for the nodes Symptoms and Alternatives, with respect to Diseases, resembles the process of finding the priority vector that we discussed in Chapter 2. Indeed, these two procedures are the same. In Chapter 2 and Chapter 3, we discussed the method to find the weights of the objectives on the goal, and of the alternatives on an objective. In light of our recent discussion of network, what we did in those chapters was finding the weights of a group of elements (in a node) on another element (in another node). We now apply the same principles to find the weights of a group of *nodes* on another *node*. This approach yields the desired priorities in this case as long as we keep in mind that we are finding the weights of the *nodes*, not of the elements in the nodes. As long as this condition is satisfied, the discussion in Chapter 2 through Chapter 5 applies.

Let β_1 and β_2 be the weights of Diseases and Symptoms, respectively, with respect to Symptoms. Using the same process outlined above, we can obtain β_1 ad β_2. By multiplying each entry in W_{12} by β_1, and each entry in W_{22} by β_2, we will make the block $\begin{bmatrix} W_{12} \\ W_{22} \\ 0 \end{bmatrix}$ in W column stochastic. Thus, the transformed form of W:

$$T = \begin{bmatrix} 0 & \beta_1 W_{12} & 0 \\ \alpha_1 W_{21} & \beta_2 W_{22} & 0 \\ \alpha_2 W_{31} & 0 & I \end{bmatrix}$$

is column stochastic. Each of W_{12}, W_{21}, W_{22}, and W_{31} in T reflects the weights of the elements in a node, with respect to the elements in another node. Each of the coefficients α_1, α_2, β_1, and β_2 reflects the weight of a node, with respect to another node. Therefore, T represents all of the interactions of the factors in this case study. Raising T to powers would give us the target priorities of the alternatives.

Before proceeding to find T, we make two important observations:

First, in finding the weights of Symptoms and Alternatives with respect to Diseases, we used the pairwise comparison matrix approach. Recall that in the initial Alice example, the merit of using a pairwise comparison matrix lies in the fact that this matrix allows utilization of the information on all the objectives. Further, comparing the weight of each objective to the weights of *all other* objectives is an error-prone task. However, when we have to find the weights of only two nodes on another node, as in the supermatrix for this case study, it is quicker to directly compare the importance of the two nodes, instead of using the pairwise comparison matrix approach.

In particular, to find α_1 and α_2, we ask the question: Is the diagnosis of a disease more a direct result of knowing about the symptoms, or of knowing about the alternatives, and how much more so? If the answer is that knowing about the symptoms contributes 80% to the diagnosis of a disease, then $\alpha_1 = 0.8$, and $\alpha_2 = 0.2$. To estimate β_1 and β_2, we ask: Is knowledge of a symptom and its usefulness in making the diagnosis more a direct result of knowing about the diseases, or of knowing about the other symptoms, and how much more so? If the answer is that knowing about the disease contributes 40% to the knowledge of the symptom, then $\beta_1 = 0.4$, and $\beta_2 = 0.6$. These two questions are answered by physicians, taking into account their knowledge of the diseases, the symptoms and the medical history of the patient.

Second, we note that the question asked in finding α_1 and α_2 does not make much sense. Certainly, knowing about the alternatives, which can only be "terminate" or "not terminate the pregnancy", does not have any contribution whatsoever to the diagnosis of a disease. Thus, α_2 must be 0,

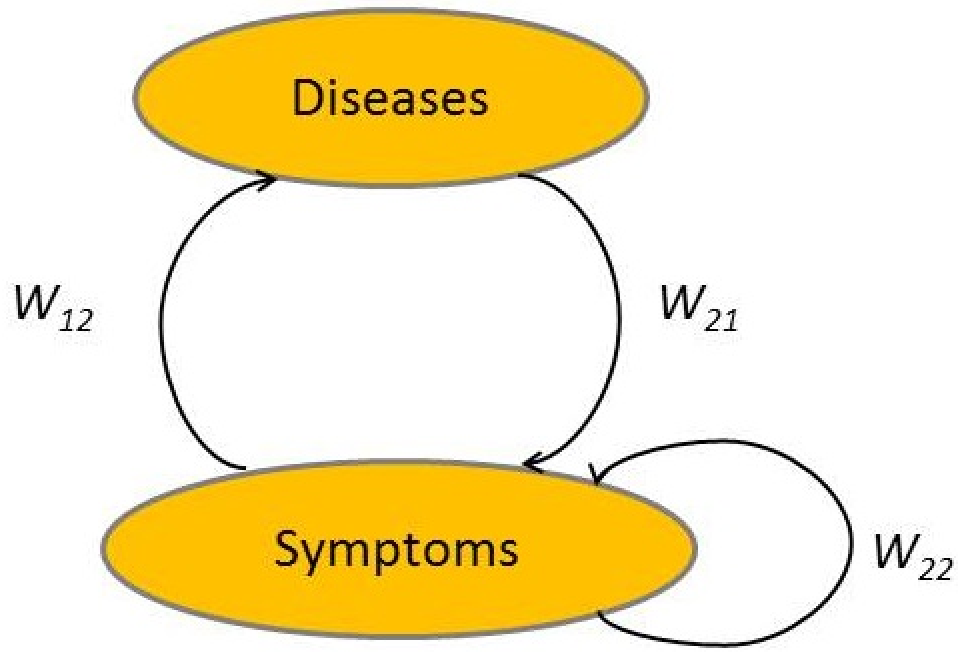

Figure 6.4: Reduced Network for the Case Study in Medical Diagnosis

which makes $\alpha_2 W_{31}$ the zero matrix. T is reduced to

$$Q = \begin{array}{r} \\ \text{Diseases} \\ \text{Symptoms} \end{array} \begin{array}{c} \begin{array}{cc} \text{Diseases} & \text{Symptoms} \end{array} \\ \begin{bmatrix} 0 & \beta_1 W_{12} \\ W_{21} & \beta_2 W_{22} \end{bmatrix} \end{array}, \tag{6.3}$$

which corresponds to the network in Figure 6.4.

By raising Q to powers, we will obtain the priorities of the possible diseases, given the symptoms. However, the ultimate goal is to find the priorities of the alternatives, given the diseases. We overcome this issue by observing that the problem in this case study can be presented in the form of a

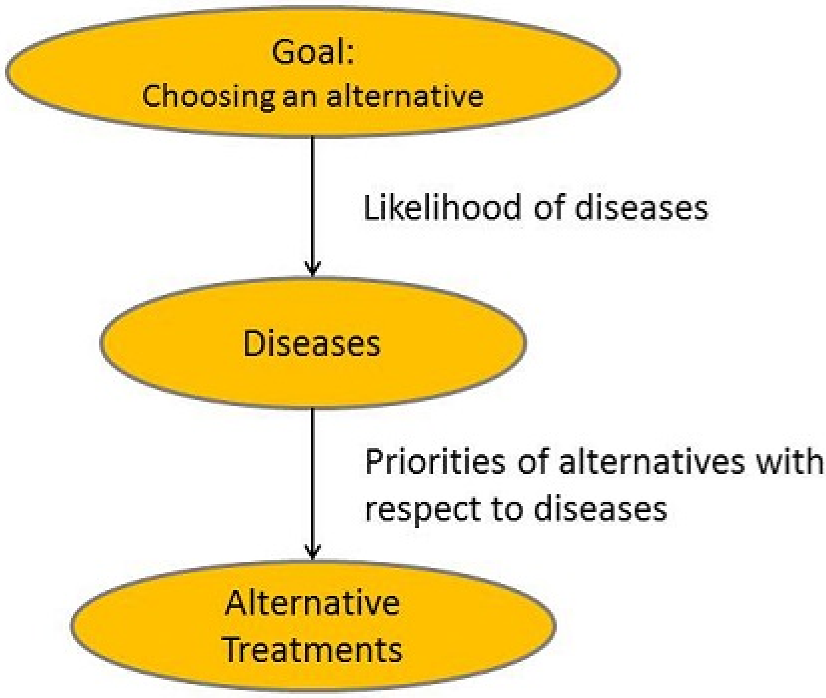

Figure 6.5: Linear Network for the Case Study in Medical Diagnosis

linear network as in Figure 6.5, *provided that we have obtained the priorities of the possible diseases*.

The case study becomes a problem that can be solved in three steps:

1. **Finding the likelihood of the diseases**. In Figure 6.5, this is represented as the influence of the element in the node Goal on the elements in the node Diseases. We achieve this likelihood by applying the supermatrix approach to the network in Figure 6.4. The supermatrix approach in turn has the following five steps:

    (a) Finding the priorities of the observed symptoms with respect to the possible diseases. This is W_{21} in Q.

(b) Finding the priorities of the possible diseases with respect to the observed symptoms. This is W_{12} in Q.

(c) Finding the priorities of the observed symptoms with respect to one another, which is W_{22} in Q.

(d) Finding the weights β_1 and β_2 of Diseases on Symptoms and of Symptoms on Symptoms, respectively.

(e) Putting W_{21}, W_{12}, W_{22}, β_1, and β_2 into Q. Raising Q to powers to obtain the priorities of the possible diseases, given the observed symptoms.

2. **Finding the priorities of the alternative treatments with respect to the possible diseases**. This is represented in Figure 6.5 as the influence of the elements in the node Diseases on the elements in the node Alternative Treatments. The priorities are obtained as principal eigenvectors of the pairwise comparison matrices that compare the alternatives for each disease.

3. **Finding the priorities of the alternative treatments for the patient**, which is the influence of the element in the node Goal on the elements in the node Alternative Treatments in Figure 6.5. We achieve this by the matrix multiplication approach presented in Chapter 4. Specifically, we right-multiply the matrix of priorities obtained in step 2 by the priority matrix obtained in step 1.

The information on the relationship among diseases, symptoms, and alternatives for the patient is obtained based on the physicians' answers. The rest of this chapter will be devoted to finding the priorities of the alternatives using the presented approach.

6.2.2 Finding the Likelihood of the Diseases

In this subsection, we explain how Saaty found the likelihood of the possible diseases. The method is to apply the supermatrix approach to the network in Figure 6.4. Recall that the supermatrix associated with this network is Q given in Equation 6.3. Each of the nonzero components in Q is a matrix of priorities, comprised of principal eigenvectors of pairwise comparison matrices. We illustrate the process to find these components below.

Finding the priorities of the observed symptoms with respect to the possible diseases (W_{21}**)**

For each of the diseases, a pairwise comparison matrix is constructed. The entry in the ij position of this matrix reflects the extent to which symptom i is characteristic of the disease, compared to the extent to which symptom j is characteristic of the disease, as judged by the physicians. In order to help obtain the entries in the pairwise comparison matrix, the physicians answered the following question:

For (the given) disease and for two symptoms, which symptom is more characteristic of the disease, and how much more is it?

The judgments are provided verbally as equally, weakly, strongly, very strongly, or absolutely (more characteristic of the disease). The judgments are interpreted into numerical values according to Table 5.2, and these numerical values are placed into the pairwise comparison matrix for the disease. The method presented in Chapter 2 and Chapter 3 are used to obtain the principal eigenvector of the pairwise comparison matrix, which gives the priorities of symptoms with respect to the disease.

The process outlined above follows closely what we have described in the first five chapters. In the rest of this chapter, we will encounter more opportunities in which our knowledge of the pairwise comparison matrix and the procedure to find the priority vector is utilized. In such situations, our discussion in Chapter 2 through Chapter 5 applies.

For the disease Lupus, the pairwise comparison matrix A that provides comparisons on the pairs of symptoms is

$$A = \begin{array}{l} \text{Lupus} \\ \text{An} \\ \text{LP} \\ \text{BC} \\ \text{APTT-H} \\ \text{ANA-H} \\ \text{ACA-H} \end{array} \begin{array}{c} \begin{array}{cccccc} \text{An} & \text{LP} & \text{BC} & \text{APTT-H} & \text{ANA-H} & \text{ACA-H} \end{array} \\ \begin{bmatrix} 1 & 5 & 4 & 4 & \frac{1}{9} & 2 \\ \frac{1}{5} & 1 & 1 & 1 & \frac{1}{9} & 1 \\ \frac{1}{4} & 1 & 1 & 1 & \frac{1}{9} & \frac{1}{2} \\ \frac{1}{4} & 1 & 1 & 1 & \frac{1}{9} & \frac{1}{2} \\ 9 & 9 & 9 & 9 & 1 & 9 \\ \frac{1}{2} & 1 & 2 & 2 & \frac{1}{9} & 1 \end{bmatrix} \end{array}.$$

For example, the entry in the fifth row and the sixth column in A is 9, so ANA-H is absolutely more characteristic of Lupus, compared to ACA-H. The symptom ABL was not included in this matrix because ABL is not characteristic of the disease Lupus. The priority of ABL with respect to Lupus is therefore 0. The weight vector $\mathbf{w}$, which is the principal eigenvector of A, gives the priorities of all of the symptoms, except for ABL, with respect to Lupus.

$$\mathbf{w_1} = \begin{bmatrix} 0.156 & 0.050 & 0.046 & 0.046 & 0.630 & 0.072 \end{bmatrix}^T .$$

For the other diseases, the original article does not provide the pairwise comparison matrices, but information about the corresponding eigenvectors is available. Next, the priority vectors are placed into Table 6.1. The entries of $\mathbf{w_1}$ are placed in the first column of the table.

Table 6.1: Priorities of Symptoms with Respect to Diseases

	Lupus	TTP	HELLP	ACA Syn
An	0.156	0.133	0.313	0.053
LP	0.050	0.789	0.313	0.158
ABL	0.000	0.000	0.313	0.000
BC	0.046	0.026	0.000	0.263
APTT-H	0.046	0.026	0.061	0.263
ANA-H	0.630	0.000	0.000	0.000
ACA-H	0.072	0.026	0.000	0.263

In Table 6.1, the entries in the j column give the priorities of the symptoms, with respect to the j disease. In terms of our case study, these entries represent the *relative* probabilities that the symptoms are observed, given the j disease. By relative probabilities, we mean, for example, the probability that the symptom An, out of all of the other symptoms, is observed in the patient if she has Lupus. According to the first column of the table, the probability that An is observed in the patient if she has Lupus is 0.156. Thus, by looking at all of the entries in a column, we can identify the symptom that is most prevalent for a given disease. From the first column of the table, given that the disease is Lupus, ANA-H is the most likely symptom to be exhibited, with a probability of 0.63.

We note that Table 6.1 gives the entries in W_{21}. Since the table is obtained from the eigenvectors of the pairwise comparison matrices, such that the entries of each vector sum to 1, each column sum of the table is 1, and W_{21} is column stochastic.

Before moving on to the next step, we note that the probabilities in Table 6.1 are not for a general, hypothetical pair of disease-symptom, but for the pair of disease-symptom pertinent to the patient in this case study. This means that the physicians whose judgments allowed the derivation of these probabilities gave their answers based on their knowledge *both* of the disease-symptom pair, *and* of the patient's medical history and current condition. This subtle difference also applies to the probabilities found in the next two steps.

Finding the priorities of the possible diseases with respect to the observed symptoms (W_{12})

For each of the symptoms, a pairwise comparison matrix is constructed. The ij entry in this matrix represents the likelihood that the i disease exhibits the given symptom, compared to the likelihood that the j disease exhibits the symptom. By likelihood, we mean the chance that the disease causes the symptom, as judged by the physicians. The following question was answered by the physicians:

For (the given) symptom and two diseases, which disease is more likely to exhibit this symptom, and how much more likely is it?

For the symptom An, the pairwise comparison matrix B that provides paired comparisons of all the diseases, with respect to An, is

$$B = \begin{array}{l} \text{An} \\ \text{Lupus} \\ \text{TTP} \\ \text{HELLP} \\ \text{ACA Syn} \end{array} \begin{array}{c} \begin{array}{cccc} \text{Lupus} & \text{TTP} & \text{HELLP} & \text{ACA Syn} \end{array} \\ \begin{bmatrix} 1 & \frac{1}{5} & \frac{1}{9} & \frac{1}{5} \\ 5 & 1 & 1 & 1 \\ 9 & 1 & 1 & 1 \\ 5 & 1 & 1 & 1 \end{bmatrix} \end{array}.$$

For example, the entry in the 2,1 position of B is 5. This means that TTP is strongly more likely than Lupus to exhibit An as a symptom. The principal eigenvector of B gives the priorities of all of the diseases, with respect to An:

$$\mathbf{w_2} = \begin{bmatrix} 0.052 & 0.299 & 0.350 & 0.299 \end{bmatrix}^T.$$

The pairwise comparison matrices for the other symptoms are not available in the original article. The priority vectors are available and were placed into Table 6.2. The entries of $\mathbf{w_2}$ are placed in the first column of the table.

Table 6.2: Priorities of Diseases with Respect to Symptoms

	An	LP	ABL	BC	APTT-H	ANA-H	ACA-H
Lupus	0.052	0.231	0.000	0.279	0.222	0.706	0.119
TTP	0.299	0.461	0.000	0.070	0.056	0.088	0.030
HELLP	0.350	0.231	1.000	0.093	0.056	0.088	0.020
ACA Syn	0.299	0.077	0.000	0.558	0.666	0.118	0.831

In Table 6.2, the j column gives the priorities of the possible diseases with respect to the j symptom. In other words, the entries in the j column represents the *relative* probabilities that the possible diseases cause the j symptom. Relative probabilities mean, for example, the probability that the disease HELLP, among all of the other possible diseases, causes the symptom LP. From the table, the probabilities that HELLP causes LP is 0.231, which corresponds to the entry in the 3,2 position in the table. Therefore, by looking at all of the entries in the j column, we can identify the disease that is most likely to cause the symptom j. For example, by looking at the second column in the table, we know that the disease that is most likely to cause LP is TTP, with a probability of 0.461.

Table 6.2 gives the entries in W_{12}. Since each column in Table 6.2 is obtained from a priority vector, the sum of the entries in each column is 1, and W_{12} is column stochastic.

Finding the priorities of the observed symptoms with respect to one another (W_{22})

For each of the symptoms, a pairwise comparison matrix is constructed.

The entry in the ij position of this matrix reflects the likelihood that the i symptom occurs at the same time as the given symptom, compared to the likelihood that the j symptom occurs jointly with the given symptom. Likelihood means the chance that a symptom occurs at the same time as the given symptom, as judged by the physicians. The physicians answered the following question in order to help derive the pairwise comparison matrix:

Given a symptom, e.g., ANA-H, and two other symptoms that may be related to it, e.g., An and LP, which of the two latter symptoms is more likely to be associated with, or occur jointly with, the given symptom? How much more likely is it?

The pairwise comparison matrix C that compare pairs of symptoms, with respect to the symptom ANA-H is

$$C = \begin{array}{l} \text{ANA-H} \\ \text{An} \\ \text{LP} \\ \text{BC} \\ \text{ACA-H} \end{array} \begin{array}{c} \begin{array}{cccc} \text{An} & \text{LP} & \text{BC} & \text{ACA-H} \end{array} \\ \begin{bmatrix} 1 & 1 & 4 & 4 \\ 1 & 1 & 1 & 1 \\ \frac{1}{4} & 1 & 1 & 1 \\ \frac{1}{4} & 1 & 1 & 1 \end{bmatrix} \end{array}.$$

For example, the entry in the first row and the second column in C is 1, which means that An is as equally likely as LP to occur jointly with ANA-H. Note that ABL and APTT-H were omitted from C because these symptoms are not related to ANA-H, as judged by the physicians. ANA-H is also skipped, since the question asks about a symptom and two other symptoms that are different from the original one. The priorities of ABL, APTT-H, and ANA-H, with respect to ANA-H, are 0. The principal eigenvector $\mathbf{w_3}$ of C gives the priorities of all of the other symptoms with respect to ANA-H:

$$\mathbf{w_3} = \begin{bmatrix} 0.455 & 0.235 & 0.155 & 0.155 \end{bmatrix}^T .$$

The pairwise comparison matrices for the other symptoms are not available from the original articles. The priority vectors are available and placed into Table 6.3. The entries of $\mathbf{w_3}$ are placed in the sixth column of this table.

Table 6.3: Priorities of Symptoms with Respect to Symptoms

	An	LP	ABL	BC	APTT-H	ANA-H	ACA-H
An	1.000	0.000	0.250	0.000	0.000	0.455	0.095
LP	0.000	0.000	0.750	0.105	0.106	0.235	0.048
ABL	0.000	0.500	0.000	0.000	0.000	0.000	0.000
BC	0.000	0.000	0.000	0.000	0.429	0.155	0.381
APTT-H	0.000	0.000	0.000	0.421	0.000	0.000	0.381
ANA-H	0.000	0.500	0.000	0.053	0.036	0.000	0.095
ACA-H	0.000	0.000	0.000	0.421	0.429	0.155	0.000

The j column in Table 6.3 gives the priorities of the observed symptoms with respect to the j symptom. These priorities mean the *relative* probabilities that given symptom j, the other symptoms that are observed in the patient are also present. In particular, the entry in the i row and the j column of the table represents the relative probability that the i symptom occurs jointly with the j

symptom. Relative probability means the probability that symptom i, among all of the other symptoms, occurs jointly with symptom j. For example, among all of the symptoms, the probability that An occurs jointly with ABL is 0.25, and the probability that LP occurs jointly with ABL is 0.75. As a special case, the ii entry of the table, which technically represents the probability that symptom i occurs jointly with itself, is 0 if at least one of the other entries in the i column is nonzero. The ii entry is 1 if all of the other entries in the i column is 0. The purpose is that each of the column vectors in W_{22}, which are precisely the columns from Table 6.3, sums to 1. As a result, W_{22} is column stochastic.

Finding the weights of Diseases and Symptoms on Symptoms

Recall that β_1 is the weight of the node Diseases on the node Symptoms, and β_2 is the weight of the node Symptoms on the node Symptoms. To estimate β_1 and β_2, we ask: *Is knowledge of a symptom and its usefulness in making the diagnosis more a direct result of knowing about the diseases, or of knowing about the other symptoms, and how much more so?*

The original article solves the case study problem with $\beta_1 = \beta_2 = 0.5$, which implies that knowledge of the diseases and knowledge of the observed symptoms both contribute equally to knowledge of a symptom and its usefulness in making the diagnosis. The article then solves the problem with $\beta_1 = 0.99$, and $\beta_2 = 0.01$, which means that knowledge of the dependent relationships among symptoms contributes only 1 percent to the final diagnosis. The results in both cases vary by an insignificant amount. (The priority for one of the alternative treatment decreases by 0.03 if β_2 decreases from 0.5 to 0.01.) Since our purpose is solely to illustrate an application of the extensions of the AHP, we choose to pursue only the case when $\beta_1 = \beta_2 = 0.5$.

Obtaining the priorities of the possible diseases

The idea of this subsection has been to apply the supermatrix approach to the network in Figure 6.4. The supermatrix associated with this network is Q in Equation 6.3. In this step, the entries of W_{21}, W_{12}, W_{22}, β_1, and β_2 are placed into Q. Note that each of the entries in W_{12} and in W_{22} is multiplied by $\beta_1 = \beta_2 = 0.5$. We have the following matrix:

$$Q = \begin{array}{l} \\ \text{Lupus} \\ \text{TTP} \\ \text{HELLP} \\ \text{ACA Syn} \\ \text{An} \\ \text{LP} \\ \text{ABL} \\ \text{BC} \\ \text{APTT-H} \\ \text{ANA-H} \\ \text{ACA-H} \end{array} \begin{array}{c} \begin{array}{ccccccccccc} \text{Lupus} & \text{TTP} & \text{HELLP} & \text{ACA Syn} & \text{An} & \text{LP} & \text{ABL} & \text{BC} & \text{APTT-H} & \text{ANA-H} & \text{ACA-H} \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 0.025 & 0.115 & 0 & 0.140 & 0.111 & 0.353 & 0.059 \\ 0 & 0 & 0 & 0 & 0.150 & 0.231 & 0 & 0.035 & 0.028 & 0.044 & 0.015 \\ 0 & 0 & 0 & 0 & 0.175 & 0.115 & 0.500 & 0.047 & 0.028 & 0.044 & 0.010 \\ 0 & 0 & 0 & 0 & 0.150 & 0.038 & 0 & 0.279 & 0.333 & 0.059 & 0.416 \\ 0.156 & 0.133 & 0.313 & 0.053 & 0.500 & 0 & 0.125 & 0 & 0 & 0.227 & 0.048 \\ 0.050 & 0.789 & 0.313 & 0.158 & 0 & 0 & 0.375 & 0.053 & 0.053 & 0.117 & 0.024 \\ 0 & 0 & 0.313 & 0 & 0 & 0.250 & 0 & 0 & 0 & 0 & 0 \\ 0.046 & 0.026 & 0 & 0.263 & 0 & 0 & 0 & 0 & 0.214 & 0.078 & 0.190 \\ 0.046 & 0.026 & 0.061 & 0.263 & 0 & 0 & 0 & 0.211 & 0 & 0 & 0.190 \\ 0.630 & 0 & 0 & 0 & 0 & 0.250 & 0 & 0.026 & 0.018 & 0 & 0.048 \\ 0.072 & 0.026 & 0 & 0.263 & 0 & 0 & 0 & 0.211 & 0.214 & 0.078 & 0 \end{bmatrix} \end{array}.$$

When Q is raised to powers, the entries in Q reach a stable stage, in which Q has identical columns. A column $\mathbf{w}$ of Q is:

$$\mathbf{w} = \begin{bmatrix} 0.073 & 0.068 & 0.087 & 0.106 & 0.169 & 0.144 & 0.063 & 0.067 & 0.066 & 0.088 & 0.070 \end{bmatrix}^T.$$

Since $\mathbf{w}$ is positive, the stable form of Q is positive, and we know that Q is primitive. The first four entries in $\mathbf{w}$ give the priorities of the diseases, with respect to the symptoms. The last seven entries in $\mathbf{w}$ give the priorities of the symptoms, with respect to the diseases. In order to obtain the relative likelihood of the diseases, the first four entries are normalized. The normalized form of these four entries is given in $\mathbf{w_0}$

$$\mathbf{w_0} = \begin{bmatrix} 0.218 & 0.203 & 0.262 & 0.317 \end{bmatrix}^T.$$

The first, second, third, and fourth entry in $\mathbf{w_0}$ represents the relative likelihood of Lupus, TTP, HELLP, and ACA Syn, respectively. This is the target result for the first step in the three-step solution to the case study. The second step is explained in the next subsection.

6.2.3 Finding the Priorities of the Alternatives with Respect to the Diseases

In this subsection, we illustrate the process to find the priorities of the alternative treatments with respect to the possible diseases. Recall that these priorities correspond to the influence of the elements in the node Diseases on the elements in the node Alternative Treatments in Figure 6.5. The method used to find these priorities is the same as the method that allowed us to find the priorities (or scores) of the alternatives on the objectives in Chapter 4.

In general, i.e. regardless of the number of alternatives, for each disease, a pairwise comparison matrix is constructed. This matrix provides the paired comparisons of the alternatives. The principal eigenvector of this matrix gives the priorities of the alternatives, with respect to the disease.

There are two alternative treatments for the patient in this case study: to terminate the pregnancy (denoted alternative T), and to proceed with the pregnancy (denoted alternative NT). Observe that the goal in this step is to find the weights of these two alternatives on each disease. As in the case of finding β_1 and β_2, it is quicker to directly compare the importance of the two alternatives, instead of using the pairwise comparison matrix approach. Specifically, the physicians assess the priorities of the two alternative treatments with respect to each of the diseases. Those priorities are given in Table 6.4.

Table 6.4: Priorities of Alternatives with Respect to Diseases

	Lupus	TTP	HELLP	ACA Syn
T	0.200	0.800	0.800	0.833
NT	0.800	0.200	0.200	0.167

The j column of Table 6.4 gives the priorities of the alternatives with respect to the j disease. These priorities are interpreted as the extent to which the alternative is appropriate for the given disease. For example, from the first column of the table, the appropriate treatment for Lupus is to not terminate the pregnancy (NT), with priority 0.8.

We note that the entries in Table 6.4 give the entries in W_{31}. Even though W_{31} was not entered into the supermatrix Q, it is the priority matrix desired in step 2 of the solution to the case study. In the next subsection, we illustrate the final step of the solution.

6.2.4 Finding the Priorities of the Alternatives

This is the last step in the solution to this case study. It uses the results from the first two steps. The method is to apply the matrix multiplication approach presented in Chapter 4 to the linear hierarchy in Figure 6.5.

Recall that to obtain the priorities of the elements in the lowest level of a hierarchy, with respect to the element in the highest level, we successively weigh the priorities of the elements in each level with respect to the elements in the level immediately above. In the initial Alice example in Chapter 4, we weighed the priorities of the alternatives on the objectives, and then the priorities of the objectives on the goal. This successive weighing process is represented by right-multiplying the score matrix S by the weight vector $\mathbf{w}$. The result was the priorities of the alternatives with respect to the goal.

Using the same approach, we weigh the priorities of the alternative treatments with respect to the diseases, and then the priorities of the diseases with respect to the goal. The corresponding matrices are W_{31} (found in step 2), and $\mathbf{w_0}$ (found in step 1). In terms of matrix multiplication, we right-multiply W_{31} by $\mathbf{w_0}$. This multiplication is shown below:

$$W_{31}\mathbf{w_0} = \begin{bmatrix} 0.200 & 0.800 & 0.800 & 0.833 \\ 0.800 & 0.200 & 0.200 & 0.167 \end{bmatrix} \begin{bmatrix} 0.218 \\ 0.203 \\ 0.262 \\ 0.317 \end{bmatrix} = \begin{bmatrix} 0.68 \\ 0.32 \end{bmatrix}.$$

The priority of T is 0.68, and the priority of NT is 0.32. Thus, the result of the AHP recommends that the pregnancy be terminated. Saaty noted that this recommendation was in agreement with the decision made by a doctor of the patient [13, p. 500]. Even though this fact does not guarantee the infallibility of the AHP, it gives some insight into its validity as a mathematical model for solving decision making problems.

Chapter 7

Conclusion

The case study in the last chapter is in fact an illustration of the **Analytic Network Process** (ANP). The ANP is a generalization of the AHP to the case in which there exists dependence and feedback among factors in decision making problems. In particular, the supermatrix approach is the generalization of the hierarchy approach. Regardless, both the AHP and the ANP rely on these three steps: **Decomposition**, **Measurement of preferences**, and **Synthesis** [p. 492][10].

Decomposition is the process of breaking the problems into elements, grouping these elements into levels, and representing those levels in such a way that it reflects various factors in the decision making problems. The result of this process is either a hierarchy (in which case we have the AHP), or a network (in which case we have the ANP). We briefly discussed the construction of a hierarchy in Chapter 4. Saaty's suggestions for building a hierarchy more or less stem from the social sciences [10, p. 14–16]. The construction of a network, however, is based on graph theory, which is explained in [10, p. 200–204].

In the second step, Measurement of preferences, pairwise comparisons are made about elements, which allows the derivation of pairwise comparison matrices. For the AHP, each of the principal eigenvectors of these matrices gives the priorities of the elements in a lower level with respect to an element in a higher level. By putting these eigenvectors into a matrix, we have a priority matrix, which represents the priorities of the elements in a lower level with respect to the elements in a higher level. For the ANP, each principal eigenvector gives the priorities of the elements in a node with respect to an element in a different or the same node. Each of the priority matrices represents the priorities of the elements in a node with respect to the elements in a different or the same node.

The last step, Synthesis, occurs after we have obtained the priority matrices for all of the interactions in the hierarchy (or network). The supermatrix approach is the synthesizing step for the ANP, while for the AHP, we successively weigh the priorities of the levels, from the top to the bottom of the hierarchy.

Recall that the AHP yields meaningful results only when the number of objectives is less than 10. Thus, the AHP is not particularly advantageous when the decision making problem involves a very large number of objectives. Moreover, the priority vector provides a meaningful result only when the decision maker is not too inconsistent. Specifically, we found in Chapter 5 that the consistency ratio needs to be 0.10 or less in order for the result of the AHP to be acceptable.

The strength of the AHP, as well as the ANP, lies precisely in the three steps that we outlined above. By decomposing and grouping the elements of the decision making problems, the decision makers gain a better understanding of the problems and their preferences. The hierarchy and the network allow the decision makers to look at the problems at hand in an analytical manner. Finally, the synthesis of the priority matrices offers a systematic approach to arrive at the best solution for the decision makers.

Due to time constraints, this thesis has not investigated in depth the theoretical justification of the ANP. Researchers who wish to pursue further work on the AHP might find it worthwhile to further explore the theory behind the ANP. So far, two issues have emerged as worthy of future research.

The first issue is the explanation of the stable stage of the supermatrix. An interesting fact, which is perhaps also useful for future research, is that the method of the ANP parallels that of Markov chains, as Saaty himself notes in [10, p. 206]. Specifically, the nodes in a network correspond to the states in Markov chains, and the influence of the i node on the j node at time k corresponds to the transition from state i to state j at time k. Further, the concepts "priority" in the ANP and "probability" in Markov Chains coincide. For a more comprehensive list of the correspondence of the terminologies between the two systems, the reader is referred to the cited source.

Second, recall that the case study was solved by replacing the original network (Figure 6.3) with a linear hierarchic structure (Figure 6.5), through the reduction of the original network to a simpler one (Figure 6.4). The author of this thesis based this approach on the discussion of the AHP in the thesis. This can be viewed as an alternative explanation to the approach used by Saaty, which is the same as the three-step solution presented in the last chapter.

However, Saaty attributes the raising of the reduced supermatrix to powers to the claim that the components in this supermatrix are the essential components of the network. These components are called the **sources** or **impact-priority-diffusing components**. By raising the reduced supermatrix Q to powers, we obtain $\mathbf{w}$, which is the **limiting impact priorities** of the network. Further, by right-multiplying W_{31} by $\mathbf{w_0}$, we obtain the **limiting absolute priorities** of the network. For discussion of these topics by Saaty, the reader is referred to [10, p.213] and [13, p. 498–500].

The AHP, as well its generalization, the ANP, is at heart a mathematical model for solving decision making problems. The mathematics in this thesis, especially in Chapter 2, Chapter 3, and Chapter 5 provide the theories underpinning the method of these models. The case study in the last chapter gives only a glimpse of the application of the ANP. In fact, these models have a vast array of applications in diverse fields, including national security (e.g., an analysis of terrorism for the Arms Control and Disarmament Agency [7]), international peace (e.g., a study of conflict in Northern Ireland [4, p. 225–241]), business (e.g., applications of the models in a consulting environment [4, p. 192–212]), and development (selection of research projects about surface water resources for South Africa [4, p. 122–137]). This list by no means exhausts the possibilities of applications of the models. It will also be beneficial for future research projects to devise creative applications of the models to solve real-world problems.

Bibliography

[1] Stephen Abbott. *Understanding analysis*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 2001.

[2] Andrew R. Blair, Robert Nachtmann, and Thomas L. Saaty. Incorporating expert judgment in economic forecasts: The case of the us economy in 1992. *Fundamentals of the Analytic Hierarchy Process*, 4:175–192, 1993.

[3] Robert W. Colman. *Hemostasis and thrombosis: basic principles and clinical practice*. Lippincott Williams & Wilkins, Philadelphia, PA, 2006.

[4] Bruce L. Golden, Edward A. Wasil, and Patrick T. (Eds.) Harker. *The Analytic hierarchy process: applications and studies*. Springer-Verlag, Berlin, 1989.

[5] Glen W. Hamilton, Gene B. Trobaugh, James L. Ritchie, K. Lance Gould, Timothy A. DeRouen, and David L. Williams. Myocardial imaging with 201thallium: An analysis of clinical usefulness based on bayes' theorem. *Seminars in Nuclear Medicine*, 8(4):358 – 364, 1978.

[6] N.E.G. Jonson. Everyday diagnostics—a critique of the bayesian model. *Medical Hypotheses*, 34(4):289–295, 1991.

[7] Robert H. Kupperman and Darrell M. Trent. *Terrorism: threat, reality, response*. Publication Series. Hoover Institution Press, Stanford University, 1979.

[8] David Poole. *Linear algebra: a modern introduction*. Available 2011 Titles Enhanced Web Assign Series. Cengage Learning, Boston, MA, 2010.

[9] Martin J. Prince. Predicting the onset of alzheimer's disease using bayes' theorem. *American Journal of Epidemiology*, 143(3):301–308, 1996.

[10] Thomas L. Saaty. *The analytic hierarchy process: planing, priority setting, resource allocation*. Analytic Hierarchy Process Series. RWS Publ., Pittsburgh, PA, 1990.

[11] Thomas L Saaty. *The analytic network process: decision making with dependence and feedback*. RWS Publ., Pittsburgh, PA, 2001.

[12] Thomas L. Saaty and David S. Turner. Prediction of the 1996 super bowl: An application of the ahp with feedback (the supermatrix approach). In *Fourth International Symposium on The Analytic Hierarchy Process*, pages 12–15, 1996.

[13] Thomas L. Saaty and Luis G. Vargas. Diagnosis with dependent symptoms: Bayes theorem and the analytic hierarchy process. *Operations Research*, 46(4):491–502, 1998.

[14] Allan J. Sieradski. *An introduction to topology and homotopy*. The Prindle, Weber & Schmidt Series in Advanced Mathematics. PWS-KENT Publishing Co., Boston, MA, 1992.

[15] Steven J. Skates, Feng-Ji Xu, Yin-Hua Yu, Kerstin Sjövall, Nina Einhorn, Yuchiao Chang, Robert C. Bast, and Robert C. Knapp. Toward an optimal algorithm for ovarian cancer screening with longitudinal tumor markers. *Cancer*, 76(S10):2004–2010, 1995.

[16] Shlomo Sternberg. *Dynamical systems*. Dover Publications Inc., Mineola, NY, 2010.

[17] Wayne L. Winston. *Operations research: applications and algorithms*. Thomson/Brooks/Cole, 2004.