LEVEL *II*

R-2572-AF
May 1980

Analysis of Subjective Judgment Matrices

Cindy Williams, Gordon Crawford

A Project AIR FORCE report
prepared for the
United States Air Force

Rand
SANTA MONICA, CA. 90406

80 8 12 013

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER R-2572-AF	2. GOVT ACCESSION NO. AD-A087851	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Analysis of Subjective Judgment Matrices		5. TYPE OF REPORT & PERIOD COVERED Interim rept. 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) G. Williams, C. Crawford		8. CONTRACT OR GRANT NUMBER(s) F49620-77-C-0023
9. PERFORMING ORGANIZATION NAME AND ADDRESS The Rand Corporation 1700 Main Street Santa Monica, California 90401		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Requirements, Programs & Studies Group (AF/RDQM) Ofc, DCS/R&D & Acquisition Hq USAF, Washington, D. C. 20330		12. REPORT DATE May 1980 13. NUMBER OF PAGES 47
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release; Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

No Restrictions

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Eigenvector
Matrices (Mathematics)
Geometric Mean
Monte Carlo Method
Statistical Anlaysis

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

See Reverse Side

296600

A popular method for quantifying subjective
judgment utilizes the dominant eigenvector
of a matrix of paired comparisons. The
eigenvector yields a scale of the importance
of each element of a collection relative to
the others. The scale is based on a matrix
of subjective paired comparisons of elements
of the collection. Thomas Saaty has shown
this to be a useful tool for analyzing hier-
archical structures in many military and
industrial applications: by estimating the
scale at each level of a structured problem,
this procedure yields the relative importance
of the elements at the bottom level of the
hierarchy to the goals or output at the top
level. For this class of problems the geo-
metric mean vector is computationally easier
than statistically preferable to the eigen-
vector. Further, the geometric mean vector
is applicable to a wider class of problems
and has the advantage of arising from common
statistical and mathematical models. The
statistical advantages of the proprosed pro-
cedure are theoretically and empirically
demonstrated. (GC)

R-2572-AF
May 1980

Analysis of Subjective Judgment Matrices

Cindy Williams, Gordon Crawford

**A Project AIR FORCE report
prepared for the
United States Air Force**

Rand

SANTA MONICA, CA. 90406

iii

PREFACE

There is a growing literature and interest in methods for quantifying subjective judgments. Several ongoing Air Force efforts utilizing subjective judgment have come to the authors' attention. Mission Area Analysis requires subjective estimates of a large number of parameters. Long-range planning repeatedly draws on judgments about the future importance and worth of plans and geographical areas. The Constant Quest project, directed by the Readiness/NATO Coordination Board, highlighted the importance of subjective judgments in evaluating command and control systems.

Thomas Saaty of the University of Pennsylvania has advanced a popular tool for quantifying and scaling the worth of a set of objects or entities. For problems that fit the Saaty framework, this report details an improvement on Saaty's "eigenvector" technique (Refs. [19] to [30]) that is easier to use and more amenable to statistical inferences.

This report was prepared under the Project AIR FORCE research study effort, "Evolving Concepts for Long-Range Planning."

v

SUMMARY

The recurring need to utilize subjective information and
judgments in quantitative analyses has been widely recognized and has
been treated in a number of ways. This report considers the problem
of using subjective paired comparisons to estimate the relative worth
of each member of a collection of objects or alternatives.

Let $\{E_1, E_2, \ldots, E_n\}$ be a collection of objects or entities that
are in some sense comparable. The E_i may be alternative plans to
achieve some goal, alternative objects that have some comparable
utility, or generally a collection of entities that have varying
degrees of some common value. A vector $(u_1, u_2, \ldots, u_n)$ is called
a ratio scale for the collection if for each i and j, u_i/u_j is the
ratio of the value of E_i to the value of E_j. For example, u_2/u_5
is 4 if E_2 has four times the value of E_5. In this case E_5 has
1/4 the value of E_2 and u_5/u_2 is 1/4.

An important application of ratio scales is in the study of
hierarchies. A hierarchy is a collection of objects organized into
levels. Suppose that for each level there is a ratio scale for the
objects at that level relative to any object at the next level up.
The ratio scales for various levels can be combined multiplicatively
to give a view of the entire hierarchy. Because hierarchies are used
to model complex systems in many important military and industrial
applications, the estimation of ratio scales deserves considerable
attention.

Suppose that a ratio scale $(u_1, u_2, \ldots, u_n)$ for objects $E_1, E_2, \ldots, E_n$
exists but is not known. Let $a_{i,j}$, $i,j = 1, 2, \ldots, n$ be subjective
estimates of u_i/u_j made by a judge. In particular, $a_{i,i} = 1$ for
each i, and $a_{j,i} = 1/a_{i,j}$. The matrix $A = [a_{i,j}]$ of subjective
pairwise comparisons is called a judgment matrix.

If the judge is perfectly consistent in making estimates, then
the matrix A will satisfy the consistency criterion

$$a_{i,j}a_{j,k} = a_{i,k} \text{ for each } i,j,k \ .$$

If this condition is met, then any column of the matrix A gives a ratio scale for $\{E_1, E_2, \ldots E_n\}$. However, judgments are frequently inconsistent, and many judgment matrices do not satisfy the consistency criterion. A mathematical procedure is required for estimating an underlying ratio scale based on an inconsistent judgment matrix A.

A procedure proposed by Thomas Saaty has been used in a variety of applications calling for the estimation of ratio scales from judgment matrices. Saaty has shown that corresponding to any judgment matrix is a positive eigenvalue exceeding all the other eigenvalues in absolute value, and the normalized eigenvector corresponding to this maximal eigenvalue has strictly positive components. Because the normalized eigenvector corresponding to the unique nonzero eigenvalue of a consistent judgment matrix does give a ratio scale for the matrix, and because the components of the eigenvector depend continuously on the matrix entries, Saaty uses the normalized eigenvector corresponding to the maximal eigenvalue as an estimate for the ratio scale underlying any judgment matrix.

The normalized eigenvector corresponding to the maximal eigenvalue of a judgment matrix is not the only vector that gives the underlying ratio scale when the matrix is consistent. The underlying ratio scales in the consistent case are also given by any column of the matrix, by the vector of row sums, and by the vector of row geometric means (nth root of the product of row elements). In fact, if the judgment matrix is consistent, all these vectors are scalar multiples of the normalized eigenvector. Moreover, each of these vectors depends continuously on the matrix entries. Thus, the argument in favor of the eigenvector as a ratio scale estimator holds for several other vectors as well.

The problem of estimating ratio scales from judgment matrices can be viewed in a statistical framework as follows. Let $(u_1, u_2, \ldots, u_n)$ be a ratio scale for a collection $\{E_1, E_2, \ldots, E_n\}$, and let $A = [a_{i,j}]$ be a judgment matrix derived from pairwise comparisons of the E_i. The elements $a_{i,j}$ can be thought of as arising from ratios of the scale elements by multiplicative perturbations--i.e.,

$$a_{i,j} = (u_i/u_j)e_{i,j} ,\qquad (S.1)$$

where the $e_{i,j}$ are positive random variables. Taking logarithms one has

$$\ln a_{i,j} = b_i - b_j + d_{i,j} ,\qquad (S.2)$$

where $b_i = \ln u_i$ and $d_{i,j} = \ln e_{i,j}$. Assuming that the $d_{i,j}$ are independent random variables with zero means and equal variances, the best linear unbiased estimator of the b_i are the least squares estimators $\widehat{b}_i$ given by

$$\widehat{b}_i = \frac{1}{n} \sum_{j=1}^{n} \ln a_{i,j} .$$

Corresponding estimators of the u_i are given by

$$\widehat{u}_i = \prod_{j=1}^{n} a_{i,j}^{1/n} .$$

Under the additional assumption that the $d_{i,j}$ are normally distrib-
uted, the $\widehat{u_i}$ are maximum likelihood estimators of the u_i.

As an estimator of ratio scales, the geometric mean vector is
preferable to the dominant eigenvector in several respects. First,
the geometric mean vector is easy to calculate, even by hand; cal-
culation of the eigenvector usually requires an iterative procedure
that would be difficult to carry out by hand and can be time-consuming
even on a computer. Also, because the geometric mean vector arises
from a well-known statistical model, it lends itself to confidence
interval estimation and tests of hypotheses.

Finally, empirical evidence based on Monte Carlo experiments
indicates that the geometric mean vector is statistically better than
the dominant eigenvector under two quite different assumptions on the
distribution of the $d_{i,j}$. In both cases the relative efficiency of
the two procedures is nearly one when the variances of the $d_{i,j}$ are
small, and the geometric mean becomes preferable as the matrices
become increasingly inconsistent.

ACKNOWLEDGMENTS

CONTENTS

I. INTRODUCTION

Over the past three decades, military and industrial researchers have directed considerable effort to the quantitative analysis of subjective data. Analytic tools building upon the subjective judgments of experts have been used in such diverse fields as energy policy analysis, marketing research, economic forecasting, and military planning. Problems amenable to the analysis of subjective information abound, and numerous methods have been proposed for acquiring and treating judgmental data.

One question that arises in the treatment of subjective data is how to construct a scale of relative merit for a collection of objects or activities based upon subjective comparisons of each pair in the collection. For example, consider a collection of three objects, labeled A, B, and C. Suppose that an expert believes A has twice the merit of B, B has three times the merit of C, and A has six times the merit of C. It is natural to construct a scale of relative merit for A, B, and C as (1, 1/2, 1/6). However, suppose that the expert says A has twice the merit of B and B has three times the merit of C, but A has only four times the merit of C. In this case, it is not so easy to decide upon a scale for A, B, and C. This sort of inconsistency is common in human judgments, especially when complicated issues are involved.

Thomas Saaty of the University of Pennsylvania has developed a matrix eigenvector procedure for constructing scales of merit based on inconsistent pairwise comparisons. The method has been applied in a wide variety of planning and decision problems.

This report presents an alternative approach that is preferable to the eigenvector procedure in several important respects. The proposed procedure is derived within a statistical framework and is compared with the eigenvector method on the basis of theoretical and empirical considerations.

The remainder of this introduction discusses the motivation for dealing formally and quantitatively with subjective information and provides a brief review of some of the relevant literature. Section

II provides a short, nonrigorous discussion of the eigenvector method
and the proposed method for utilizing subjective judgments in quanti-
tative analysis. An example, illustrating the use of the two methods
as well as similarities and differences in their results, is intro-
duced in this section and examined throughout the report.

In Section III we give rigorous definitions and develop a frame-
work for treating the estimation problem with classical statistical
techniques. Section IV provides a mathematical treatment of the
eigenvector method. Section V deals with the application of subjec-
tive judgment methods to the study of hierarchical structures. The
example introduced in Section II is considered in further detail
there.

Section VI introduces the geometric mean vector and gives
theoretical justification for its use as an estimator of subjective
scales. In Section VII we define a statistical measure of consistency
for subjective judgment matrices. Section VIII presents results
of a Monte Carlo study comparing the two methods. Section IX con-
siders in greater detail the example introduced in Section II and
expanded on in Section V.

QUANTITATIVE ANALYSIS OF SUBJECTIVE DATA

The quantification of subjective data is essential for dealing
with a wide class of problems whose solution by other methods would be
extremely difficult or impossible. Such problems are often amorphous
and vaguely stated. They involve large, multifaceted issues of impor-
tance to decisionmakers and interest groups with diverse backgrounds
and biases. Their outcomes may determine the allocation of large
sums of public money and impinge critically on the public interest.
Moreover, some facets of the problems may lack any well-defined,
scalar-valued measures of merit. Even if there are appropriate mea-
sures, the collection of relevant objective data might be prohibitively
expensive or impossible.

Such problems frequently arise in the assessment of future needs
for large organizations. As an example, consider the problem of long-
range planning in the U.S. Air Force. This problem involves a great

many interrelated issues: the effects of political and economic
factors on national security, the importance of various geographic
regions to U.S. interests, the threat posed by conflicts of different
types in different regions, the current strength of forces to deal
with such conflicts, and so on. Although it might be possible to
define objective yardsticks to deal with some of these issues, it
certainly is not possible for all of them. For some issues, sub-
jective judgments of relative importance or value are the only
measures available.

In some problems the best information available is subjective, so
why is quantitative analysis desirable at all? Why not just ask the
experts to make plans and decisions based on an informal, intuitive
analysis? In fact, problems that are not amenable to hard analysis
are frequently resolved through intuitive analysis by experts and
decisionmakers. However, there are several good reasons for using a
formal, quantitative approach in these problems.

A formal analytic framework gives structure and definition to
an amorphous mass of data. It allows the decisionmaker to consider
relevant information systematically and to examine options and con-
sequences one at a time. In such a framework, the analyst can break
an unmanageable problem into manageable parts and then synthesize
information about the parts in a rational fashion.

A formal analytic framework also permits sensitivity analysis on
alternative judgments. When a problem is considered within a formal
framework, tradeoffs among alternative judgments can be spelled out
explicitly, and the effects of variations in subjective judgments on
outcomes can be studied. Sensitivity analysis may even provide a
basis for resolving different points of view.

Perhaps the greatest advantage of a formal analysis in government
and military applications is that it is repeatable. Formal analysis
provides the audit trail that is so important in matters involving
extensive allocation of public resources and impinging on the public
interest.

Research literature on the use of subjective information
emphasizes three major issues: how to elicit meaningful subjective

judgments from individuals or groups, how to synthesize subjective and objective data obtained from various facets of a large problem, and how to construct measurement scales based on subjective information. Following is a review of some of the literature related to each of these issues.

ELICITING SUBJECTIVE JUDGMENTS

Methods for eliciting subjective judgments have received considerable attention in operations research and forecasting literature. Two such methods are war gaming and scenario writing, both of which are used extensively in military planning to provide insights into possible future environments and needs.

Much of the literature on eliciting judgments deals with the problem of acquiring a collective expert opinion free from the usual negative effects of group pressure. An important method in this category is the Delphi technique, a controlled feedback procedure originally developed by researchers at The Rand Corporation [1]. In Delphi, a researcher interrogates a group of experts individually concerning their opinions on possible future events. The researcher assembles means and quartiles for quantitative data thus obtained and presents them individually to group members along with arguments and comments made by individuals. Group members can then revise their judgments. The procedure is repeated until the range of judgments narrows. The controlled feedback mechanism in Delphi makes it possible for a group of experts to avoid the usual social pressures of open discussion. The method has been used in many military and industrial applications (see, e.g., [2], [3]).

The Delphi technique has given rise to a number of modifications. The Probe method designed by researchers at TRW for forecasting technological events combines Delphi with a timing chart structure so that events can be considered in sequence [4]. The method of qualitative controlled feedback proposed by Press [5] is similar to Delphi in that it uses a controlled feedback loop to aid groups in arriving at judgments, but it differs in that at each iteration, members are supplied only arguments and comments from the group, with no information about the quantitative distribution of group answers.

SYNTHESIZING DATA IN LARGE PROBLEMS

Economists and statisticians have proposed a variety of methods for breaking large problems into smaller pieces and quantitatively synthesizing subjective and objective data from the pieces. One of the most popular of these methods is multi-attribute utility theory, which provides a framework for selecting an optimal decision from among multiple alternatives when some effects of the decision can be measured only subjectively. The expected value of each alternative is determined as a function of the decisionmaker's preferences for the possible consequences and the probabilities that the alternative will lead to those consequences. The probabilities are generally determined from subjective judgments. Some of the decisionmaker's preferences are determined on the basis of subjective indexes such as aesthetic appeal, and others are determined on the basis of objective measures such as cost. The alternative with maximum expected value is chosen as the optimum decision.

The mathematical foundation for multi-attribute utility theory was laid by von Neumann and Morgenstern [6]. Application of the theory to business problems was pioneered by Raiffa and extended by Keeney and others. The theory has been applied to many problems in industrial, government, and military settings (see [7], [8], [9], [10], [11]). The 1976 book by Keeney and Raiffa gives an excellent treatment of the subject [12].

A similar method was applied to military problems in a 1958 Master's thesis by Wells, who gives a detailed framework for assessing the relative desirability of existing or proposed weapon systems. System desirability is determined as a function of feasibility, cost, and an attribute Wells calls "military worth." Wherever possible, objective measures are used to evaluate these three factors, and expert judgments are used where there are no objective measures. In particular, military worth is an aggregate property evaluated by analyzing a complex hierarchy and subjective scales for a number of variables. The Honeywell Corporation used Wells's method in a military planning model called PATTERN [13]. A detailed description of the method can be found in Ref. [14].

Saaty proposes that complex decision problems be viewed in
terms of hierarchies of objects or properties. At each level of a
hierarchy, Saaty uses subjective judgments to estimate a merit scale
of the objects. Scales from all the levels are combined mathemati-
cally to provide quantitative information about the whole problem.
The results in this report are applied within the framework of hier-
archical analysis proposed by Saaty. Hierarchies are discussed in
Section V.

CONSTRUCTING MEASUREMENT SCALES

Many methods have been developed for constructing scales of
measurement based on subjective data. Several books and hundreds of
articles have been written about these methods. A classic reference
for early contributions, especially for work on psychophysical scales,
is Torgerson's 1958 book [15].

Churchman and Ackoff did pioneering work in the area of esti-
mating scales of values for decision problems in 1954. They used a
criterion of additive order consistency to estimate scales from suc-
cessive subjective judgments. Their 1954 paper described a number
of applications to industrial problems [16]. Wells and others later
applied the Churchman and Ackoff method in military decision problems
[14].

Much of the literature on subjective scales concerns the esti-
mation of scales from pairwise comparison data. A good deal of sta-
tistical work in this area goes under the name "paired comparisons."
In the simplest paired comparison experiment, each of several judges
examines a number of objects two at a time and states which of the
two objects is preferred. No indication of strength of preference is
given. Data from these paired comparisons are then used in a statis-
tical model to estimate a scale of preference for the objects. Such
an experiment might be used by marketing researchers to determine the
relative taste appeals of several new food items.

A good reference for the statistical theory of paired comparisons
is David's 1963 book [17]. A bibliography of recent articles on the
subject was compiled in 1976 by Davidson and Farquhar [18].

Saaty has proposed another method for estimating subjective scales using pairwise comparisons in which a single judge makes pairwise comparisons of a number of objects. For each pair, the judge states not only which object is preferred, but to what degree that object is preferred over the other. A preference scale is determined for the objects based on an eigenvector analysis of the matrix of pairwise comparisons.

Saaty has published a number of articles (Refs. [19] to [30]) describing the eigenvector procedure for estimating subjective scales and illustrating the usefulness of this procedure in analyzing complex hierarchical structures. He has applied the procedure in a broad range of problems in the social sciences ([19], [21], [28]). The procedure has also gained acceptance in military applications and is currently being used as a tool in Air Force long range planning.

II. PAIRWISE COMPARISONS, THE JUDGMENT MATRIX, AND
THE ESTIMATION PROBLEM

Consider the problem of purchasing a new car.*   Suppose that a
preliminary investigation yields five specific makes that seem appro-
priate. The price of each make is known, and although some other
measures of merit may have been quantified (principally performance
measures), the important subjective question of how much each car
satisfies the overall needs is difficult to quantify.

We will attempt to assign to each make of automobile an estimate
of utility in such a way that if u_i is the utility of the ith make,
then u_i/u_j is a measure of the preference of the ith make to the jth
make. The vector $(u_1, u_2,...,u_5)$ will be called a __ratio scale__.

Some aspects of the usefulness of such a ratio scale are imme-
diately apparent. We could, in this example, choose between the cars
on the basis of utility per dollar of initial cost, or with more fore-
sight, on the basis of utility per dollar of expected life cycle cost.

To estimate the vector of utilities $(u_1, u_2,...,u_5)$ Saaty has
suggested the following procedure ([19] to [30], esp. [28]): We
construct a matrix composed of our subjective estimates of the ratios
of the utilities of all possible pairwise combinations, so that the
elements $a_{i,j}$ of the matrix A are our estimates of u_i/u_j. Thus we
know that the diagonal elements are given by $a_{i,i} = 1$, $i = 1,...,5$.
Additionally, the lower off-diagonal elements are determined by
the upper off-diagonal elements: $a_{j,i} = 1/a_{i,j}$.

Saaty [28] proves that in this case the matrix A has a maximal
eigenvalue and a corresponding eigenvector (the dominant eigenvector)
all of whose components are positive. Saaty proposes, primarily with

---

*This example, due to Capt. Jordan Kreindler and Major Michael
Parmentier of the U.S. Air Force Directorate for Programs Evaluation,
Systems Analysis Division, is treated in further detail in Sections V
and IX.

empirical justification, that this dominant eigenvector be used as an estimate of the ratio scale.

Suppose that when we form our estimates, the relative utility of Make 1 to Make 2 is 2, of Make 1 to Make 3 is 1/9, of Make 1 to Make 4 is 1/6 and of Make 1 to Make 5 is 1/7. Then the first row of our judgment matrix has the form

1, 2, 1/9, 1/6, 1/7 .

Continuing, suppose that we have filled in the upper off-diagonal of our judgment matrix:

$$
\begin{array}{ccccc}
1 & 2 & 1/9 & 1/6 & 1/7 \\
  & 1 & 1/9 & 1/6 & 1/7 \\
  &   & 1   & 6   & 4 \\
  &   &   & 1   & 1/7 \\
  &   &   &   & 1
\end{array}
$$

Then, in view of reciprocal symmetry we have

$$
A = \begin{array}{ccccc}
1   & 2 & 1/9 & 1/6 & 1/7 \\
1/2 & 1 & 1/9 & 1/6 & 1/7 \\
9   & 9 & 1   & 6   & 4 \\
6   & 6 & 1/6 & 1   & 1/7 \\
7   & 7 & 1/4 & 7   & 1
\end{array}
$$

Continuing with this example we compute the dominant eigenvector w of this matrix and get:

$$
w = \begin{array}{c}
.0378 \\
.0294 \\
.5239 \\
.1131 \\
.2958
\end{array}
$$

Thus, in this case, our estimate of the utility of the first make is .0378 and of the third make is .5239. For a detailed treatment of this procedure see Refs. [19] and [28].

This example is discussed in more detail in Sections V and IX, where it is expanded to illustrate the value of ratio scales in analyzing hierarchical structures. In application to hierarchies it is assumed that the objects at each level of the hierarchy depend on the objects of the next lower level in some way. The procedure enables the user to estimate the influence each object in a level has on all the objects or goals in superior levels.

For problems where the eigenvector procedure is useful there is another estimation procedure that is preferable in several respects.

Where Saaty would estimate the utility of the ith object with the ith component of the dominant eigenvector, we give arguments that a better estimator is given by the vector $v = v_1, v_2, \ldots, v_n$, where

$$v_i = \prod_{j=1}^{n} a_{i,j}^{1/n}$$

is the geometric mean of the elements in the ith row of A. In the example above this yields the following estimates:

Object	Eigenvector	Geometric Mean Vector
1	.0378	.0409
2	.0294	.0310
3	.5239	.5307
4	.1131	.1132
5	.2958	.2842

Compared with the dominant eigenvector, the geometric mean vector

1. Is statistically better;
2. Is easier and faster to calculate on a hand calculator or computer;

3. Gives rise to a measure of consistency that is more meaning-
    ful than the eigenvector-eigenvalue measure;

4. Gives rise to a measure of consistency that has known statis-
    tical properties, allowing tests of hypotheses, confidence
    interval estimation, etc.;

5. Gives rise to estimates of utility with known statistical
    properties, allowing tests of hypotheses, confidence interval
    estimation, etc.;

6. Is supported by statistical literature describing methods of
    handling a wealth of variations of the problem;

7. Is rooted in a mathematical approach to estimation that pro-
    vides an intuitive understanding of the problem and a means
    for assessing suitability of the method.

III. CONSISTENT MATRICES AND RATIO SCALES

Consider a set of n activities or objects E_1, E_2,...,E_n, which
contribute to some objective. Suppose the activities can be ranked
on a ratio scale $(u_1, u_2,...,u_n)$, $u_i > 0$, so that u_i/u_j measures the
degree to which E_i is more important than E_j relative to the objective.
In particular, $u_i/u_j > 1$ if E_i is more important than E_j. Let A =
$[a_{i,j}]$ be the n × n matrix of pairwise comparisons of E_1, E_2,...,E_n
given by

$$a_{i,j} = u_i/u_j , \quad i,j = 1,2,...,n .\tag{3.1}$$

Then A has the property that

$$a_{j,i} = 1/a_{i,j}, \quad i,j = 1,2,...,n ,\tag{3.2}$$

and in particular

$$a_{i,i} = 1, \quad i = 1,2,...,n .$$

A square matrix A with positive entries satisfying (3.2) will be called
a _judgment matrix_.

It follows immediately from (3.1) that

$$a_{i,j}a_{j,k} = a_{i,k} .\tag{3.3}$$

A matrix with positive entries satisfying (3.3) is said to be _consistent_.
It is easy to see that every consistent matrix is a judgment matrix.

Let A be an arbitrary consistent matrix. Because

$$a_{j,k} = a_{1,k}/a_{1,j} \quad \text{for any } j,k,$$

every element of A can be determined from the first row of A. It
follows that A is a matrix of rank one with exactly one nonzero
eigenvalue. Moreover, it follows from (3.3) that

$$A^2 = nA .$$

Thus any column of A is an eigenvector of A, and the single nonzero
eigenvalue of A is n.

Let $w = w_1, w_2, \ldots, w_n$ be any eigenvector corresponding to
the eigenvalue n. For any k, the kth column of A is an eigenvector
corresponding to the same eigenvalue; therefore for each i and j,

$$w_i = c a_{i,k}$$

$$w_j = c a_{j,k}$$

for some $c \neq 0$, and therefore

$$\frac{w_i}{w_j} = a_{i,j} .$$

Thus, w is a ratio scale for A. In fact, it is clear that there are
infinitely many such scales, each one corresponding to a different
scalar multiple of the kth column of A.

The normalized eigenvector with components

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

is the particular scale that Saaty [28] associates with the consistent
matrix A.

A ratio scale corresponding to a consistent matrix A can be derived in several ways. Any column of A is such a scale. The vector of reciprocals of elements of an arbitrary row of A is also a ratio scale for A. It is easy to see that the vector r of row sums defined by

$$r_i = \sum_{j=1}^{n} a_{i,j}$$

and the geometric mean vector v defined by

$$v_i = \prod_{j=1}^{n} a_{i,j}^{1/n}$$

also provide ratio scales for A. When these scales are normalized, they are equal to the normalized eigenvector scale for a consistent matrix.

IV. SAATY'S NORMALIZED EIGENVECTOR SCALE

Consider again the activities E_1, E_2,...,E_n that contribute to some objective. Suppose a judge makes pairwise comparisons on some scale of the relative importance of each pair of activities with respect to the underlying objective. If $a_{i,j}$ represents the relative importance of E_i over E_j, so that $a_{i,j} > 1$ if and only if E_i is more important than E_j, it is then natural to insist that the judge make comparisons in such a way that

$$a_{i,j} = 1/a_{j,i} \text{ for each } i,j .$$

In other words, such a pairwise comparison matrix is a judgment matrix. The ideal pairwise comparison matrix would also be consistent. For example, if E_i is twice as important as E_j and E_j is three times as important as E_k, one would expect E_i to be six times as important as E_k. However, human judgment is often inconsistent, and it is not likely that a judge making pairwise comparisons will construct a consistent matrix except in cases where the dimension is small. A simple example in which pairwise comparisons do not result in a consistent matrix is that of a tournament: X may win against Y and Y against Z, but X may lose to Z.

The problem we consider is this: Given an inconsistent judgment matrix A, how can we construct a ratio scale that in some sense best reflects the information in the matrix? Saaty proposes that the appropriate scale is the normalized eigenvector corresponding to the maximal eigenvalue of A.

Saaty [28] argues as follows. If the judgment matrix A is consistent, then the normalized eigenvector corresponding to the single nonzero eigenvalue n does give the underlying ratio scale: A theorem of Frobenius for matrices with positive entries [31] guarantees that any judgment matrix has a positive eigenvalue L that exceeds all the other eigenvalues in absolute value. This maximal eigenvalue has an associated eigenvector that is positive in all its components.

Now an inconsistent judgment matrix can be viewed as having been derived from a consistent one by perturbation of some or all of the matrix components. Because the eigenvalues and eigenvectors of a matrix depend continuously on its components, small perturbations in the components will result in small changes in the eigenvalues and eigenvectors. Thus when the perturbations of the components are small, the maximal eigenvalue is close to n, and the corresponding normalized eigenvector is close to the normalized eigenvector of the unperturbed consistent matrix. Therefore, Saaty selects the suitably normalized eigenvector associated with the maximal eigenvalue as the ratio scale corresponding to the judgment matrix.

Saaty also proposes an index of consistency for judgment matrices. He shows that an n × n judgment matrix whose only nonzero eigenvalue is n must be consistent, and that the maximal eigenvalue L for an inconsistent judgment matrix is strictly greater than n. Therefore he uses the normalized difference

$$\mu = \frac{L-n}{n-1}$$

as the index of consistency of an n × n judgment matrix with maximal eigenvalue L. Notice that the index is zero for consistent matrices and positive for inconsistent ones. Saaty also shows that the index increases as perturbations of the components away from consistency increase.

Unfortunately, the question of how small the perturbations of matrix components must be to give rise to a given deviation in the maximal eigenvalue is a delicate one. Saaty [28] describes an empirical investigation of this question in which he determines the consistency indices corresponding to randomly generated judgment matrices of different dimensions. However, because the eigenvector does not fit into any standard statistical framework, there is no readily available device against which deviations from consistency can be measured.

Saaty does show ([28, p. 238]) that the consistency index μ reflects the variance in judgmental errors for an inconsistent matrix, in the following sense. Suppose that the pairwise comparisons $a_{i,j}$ in the judgment matrix A actually arise from perturbations of the ratios of components of some underlying scale $(u_1, u_2, \ldots, u_n)$, i.e.,

$$a_{i,j} = \frac{u_i}{u_j} e_{i,j} \, ,$$

$$e_{i,j} = 1 + d_{i,j} \, .$$

Saaty shows that for small $d_{i,j}$, 2μ is an estimate of the variance of the $d_{i,j}$. Starting from this estimate, Saaty develops a test of the hypothesis of consistency for a judgment matrix ([28], p. 238).

The choice of a scale to be used in filling in a pairwise comparison matrix is somewhat arbitrary. Because humans find it difficult to rank more than about seven objects at a time, Saaty recommends a subjective pairwise comparison scale consisting of the integers from one to nine together with their reciprocals. In this scale, a value of 1 is assigned to pairs of objects that are equally important. The integers 3, 5, 7, and 9 are associated with descriptive words (9 means "absolute importance," 5 means "essential or strong importance"), and the integers 2, 4, 6, and 8 are used for intermediate values. Reciprocals of integers are used so that the matrix of pairwise comparisons is a judgment matrix--that is, is reciprocal symmetric.

V. APPLICATIONS OF THE RATIO SCALE

Saaty presents numerous applications requiring the estimation of ratio scales from pairwise comparison information ([19] to [30]). He cites examples in economics, political science, and transportation planning, as well as in personal planning areas such as choosing a school or a vacation spot.

One of the most interesting and useful applications of the ratio scale is in the study of hierarchical systems. A hierarchy is a collection of objects grouped according to levels. Objects at a given level of the hierarchy depend on objects at lower levels. The objects at one level may be ranked on a ratio scale according to their importance relative to a given object at the next higher level. Thus one may construct a system of ratio scales, one scale for each level relative to every object in the next level up.

Once such a system of ratio scales has been constructed, it can be used to study interactions among all levels of the hierarchy. For example, in a hierarchy consisting of three levels, we may determine the ranked importance of objects on the lowest level relative to each object in the highest level. Suppose for simplicity that the highest level consists of a single object, that the second level has n objects, and the third has m objects. Let $(w_1, w_2, \ldots, w_n)$ be the ratio scale that reflects the importance of objects in the second level relative to the single object in the first level. Now the objects in the third (lowest) level may be ranked on a ratio scale relative to each object of the second level. Let

$$(u_{1,j}, u_{2,j}, \ldots, u_{m,j}), \ j = 1, 2, \ldots, n$$

be the ratio scale for the third level relative to the jth object of the second level. Then according to Saaty the importance of objects in the lowest level relative to the highest level may be measured by the vector $(v_1, v_2, \ldots, v_m)$, where

$$v_i = \sum_{j=1}^{n} u_{i,j}\, w_j, \quad i = 1,2,\ldots,m \; .$$

In matrix notation, if $w = (w_1, w_2, \ldots, w_n)$ is the ratio scale for the second level relative to the single object in the first level, and if

$$U = \begin{bmatrix} u_{1,1} & u_{1,2} & \cdots & u_{1,n} \\ u_{2,1} & u_{2,2} & \cdots & u_{2,n} \\ \vdots & \vdots & & \\ u_{m,1} & u_{m,2} & \cdots & u_{m,n} \end{bmatrix}$$

is the matrix whose jth column is the ratio scale for the objects in the third level relative to the jth object in the second level, then

$$v = Uw$$

gives a scale of importance of objects in the third level relative to the first.

The same procedure may be extended to hierarchies with more than three levels. Thus knowing only the measures of importance of objects in each level relative to individual objects in the adjacent higher level, we may deduce their ranked importance relative to objects at all higher levels. In particular, objects at the lowest level can be ranked according to their importance relative to the object (or objects) at the highest level.

As an example, consider the problem of selecting an automobile. (This example is treated in more detail in Section IX.) The problem can be viewed in terms of the hierarchical structure shown in Fig. 1. The highest level of the hierarchy is the final selection of the automobile. On the second level are attributes of the automobiles--namely, status, cost, economy, and size. The third level consists of the automobile makes to be considered. The automobiles themselves are
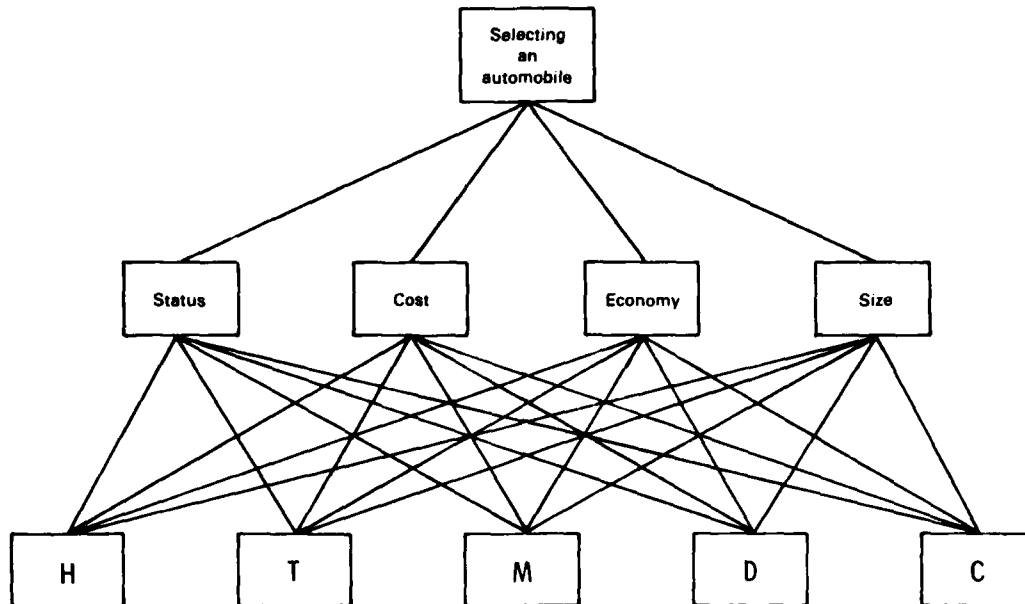
Fig. 1--Hierarchy for selecting an automobile

ranked according to each of the attributes, and the attributes are ranked according to their importance relative to the overall objective of selecting a car.

Table 1 gives the ratio scales determined from judgments made by one prospective buyer. The order of automobile preference for this buyer is then given by the product of the matrix and the vector in Table 1:

$$
\begin{array}{ccccc}
.0378 & .2703 & .3357 & .3143 & .0885 \\
.0294 & .4196 & .2929 & .4630 & .2425 \\
.5239 & .0554 & .1071 & .0596 & .2579 \\
.1131 & .1571 & .1500 & .1288 & .4112 \\
.2958 & .0976 & .1143 & .0343 &
\end{array}
\cdot
\begin{array}{c}
.2847 \; H \\
.3703 \; T \\
= \quad .1119 \; M \\
.1397 \; D \\
.0934 \; C
\end{array}
$$

Table 1

RATIO SCALES

Ratio Scale of Second Level Relative to First Level	
Status	.0885
Cost	.2425
Economy	.2579
Size	.4112

Ratio Scales of Third Level Relative to Attributes at the Second Level

	Status	Cost	Economy	Size
H	.0378	.2703	.3357	.3143
T	.0294	.4196	.2929	.4630
M	.5239	.0554	.1071	.0596
D	.1131	.1571	.1500	.1288
C	.2958	.0976	.1143	.0343

The final ranking reflects the buyer's perception of the relative status, cost, economy, and size of the five automobiles considered as well as his judgment of the relative importance of these four attributes in the selection of an automobile. This buyer's first choice should be T and his last choice C.

VI. THE GEOMETRIC MEAN SCALE

It is clear from the example in the last section that the study of interactions among various levels of a hierarchy depends heavily upon our assessment of the ranked importance of objects at each level relative to objects in the level above. The basic building blocks in a hierarchical study are the ratio scales measuring the relative importance of objects at a given level. Therefore one would like to know that the estimates of the ratio scales are well-grounded in statistical theory, that they work well empirically, and that they can be calculated quickly. In this section, we propose a method for constructing a ratio scale based on pairwise comparisons that is superior to the eigenvector procedure when judged according to each of these three criteria.

For $n \times n$ judgment matrices $A = [a_{i,j}]$ and $C = [c_{i,j}]$, define

$$m(A,C) = \left(\sum_{i=1}^{n} \sum_{j>i} [\ln a_{i,j} - \ln c_{i,j}]^2 \right)^{1/2} .$$

It is not difficult to verify that m is a metric on the space of $n \times n$ judgment matrices. Theorem 3 will show that for any $n \times n$ judgment matrix A, there is a consistent matrix C that is m-closest to A. In fact, such a consistent matrix is given by

$$c_{i,j} = v_i/v_j ,$$

where

$$v_i = \prod_{j=1}^{n} a_{i,j}^{1/n} ,$$

that is, v_i is the geometric mean of the elements of the $\underline{i}$th row of A. We will use the vector v, suitably normalized, as the estimate of the ratio scale corresponding to A.

The following two invariance properties show that m is a suitable choice of metric for the space of judgment matrices. Their proofs are straightforward and will not be provided here.

Theorem 1 (Invariance under Transpose). (i) Let $A = [a_{i,j}]$ and $C = [c_{i,j}]$ be n × n judgment matrices. Then A^T and C^T are also judgment matrices, and

$$m(A^T, C^T) = m(A,C)$$

(ii) Let $A = [a_{i,j}]$ be an n × n judgment matrix, and suppose that $C = [c_{i,j}]$ is the* consistent matrix that is m-closest to A. Then C^T is the consistent matrix that is m-closest to A^T.

Theorem 2 (Invariance under Change of Scale). (i) Let $A = [a_{i,j}]$ and $C = [c_{i,j}]$ be n × n judgment matrices and $(w_1, w_2, \ldots, w_n)$ a ratio scale. Define

$$A' = [a_{i,j} w_i / w_j]$$

$$C' = [c_{i,j} w_i / w_j] .$$

Then A', C' are judgment matrices, and

$$m(A',C') = m(A,C)$$

(ii) Let A, C, A', C', be as in (i), and suppose that C is the m-closest consistent matrix to A. Then C' is the m-closest consistent matrix to A'.

---

*The existence and uniqueness of C are guaranteed by Theorem 3.

Recall that we seek a procedure for associating ratio scales to judgment matrices in such a way that the ratio scales capture the subjective information inherent in the corresponding matrices. Let A be an $n \times n$ judgment matrix. Let $C = [c_{i,j}]$ be a consistent matrix that is m-closest to A, and suppose that $v = (v_1, v_2, \ldots, v_n)$ is a ratio scale for B; i.e.,

$$c_{i,j} = v_i/v_j \; .$$

We choose $(v_1, v_2, \ldots, v_n)$ as the estimator of the ratio scale corresponding to A.

Under this association, Theorem 1 guarantees that the scale $(1/v_1, 1/v_2, \ldots, 1/v_n)$ is the estimator of the scale corresponding to A^T. This is a natural requirement, because if v_i/v_j estimates $a_{i,j}$, then $(1/v_i/1/v_j)$ should estimate $a_{j,i}$. Theorem 2 guarantees that our choice of ratio scale is invariant under a scale change in the judgment matrix. The eigenvector scale does not satisfy either of these requirements.

The following theorem guarantees that the geometric mean scale gives the m-closest consistent matrix to any judgment matrix.

Theorem 3. Let $A = [a_{i,j}]$ be an $n \times n$ judgment matrix. Let $C = [c_{i,j}]$ be the consistent matrix given by

$$c_{i,j} = v_i/v_j,$$

where v_i is the geometric mean of the elements of the ith row of A; i.e.,

$$v_i = \prod_{j=1}^{n} a_{i,j}^{1/n}, \quad i = 1, 2, \ldots, n \; .$$

Then $m(A,C)$ is the minimal m-distance from A to any $n \times n$ consistent matrix.

   Proof. For any consistent matrix $C = [c_{i,j}]$, we can write

$$c_{i,j} = w_i/w_j \ ,$$

where $w = (w_1, w_2, \ldots, w_n)$ is a ratio scale. Thus we seek a scale that minimizes the expression

$$\sum_{i=1}^{n} \sum_{j>i} [\ln a_{i,j} - (\ln w_i - \ln w_j)]^2.$$

Because the estimating scale need be known only up to a scale factor, we may normalize by imposing the side condition

$$\prod_{i=1}^{n} w_i = 1 \ .$$

Let

$$y_{i,j} = \ln a_{i,j}, \ i,j = 1,2,\ldots,n \ ,$$

$$b_i = \ln w_i, \ i = 1,2,\ldots,n \ .$$

Then the problem is to minimize

$$\sum_{i=1}^{n} \sum_{j>i} [y_{i,j} - (b_i - b_j)]^2$$

under the side condition

$$\sum_{i=1}^{n} b_i = 0 .$$

Since

$$y_{j,i} = -y_{i,j}, \quad i,j = 1,2,\ldots,n$$

and

$$y_{i,i} = 0 ,$$

this is equivalent to minimizing

$$S = \sum_{i=1}^{n} \sum_{j=1}^{n} [y_{i,j} - (b_i - b_j)]^2$$

under the side condition

$$\sum_{i=1}^{n} b_i = 0 .$$

Now S is convex in the differences $b_i - b_j$ and therefore convex in the vector b, so its minimum occurs at the point where

$$\frac{\partial S}{\partial b_i} = 0 \text{ for } i = 1,2,\ldots,n .$$

Setting these partial derivatives equal to zero, we have for $k = 1, 2, \ldots, n$

$$\frac{\partial S}{\partial b_k} = -2 \sum_{j=1}^{n} [y_{k,j} - b_k + b_j]$$

$$= -2 \left[\sum_{j=1}^{n} y_{k,j} - nb_k + \sum_{j=1}^{n} b_j \right] = 0$$

and therefore, since $\sum_{j=1}^{n} b_j = 0$,

$$\sum_{j=1}^{n} y_{k,j} = nb_k .$$

Thus S is minimized by

$$b_k = \frac{\sum_{j=1}^{n} y_{k,j}}{n} ,$$

i.e.,

$$\ln w_k = \frac{\sum_{j=1}^{n} \ln a_{k,j}}{n} , \quad k = 1, 2, \ldots, n .$$

Consequently the m-distance from A to C is minimized by the vector v given by

$$v_k = \prod_{j=1}^{n} a_{k,j}^{1/n}$$

This completes the proof of Theorem 3.

Recall from Section III that if the matrix A is consistent, then the normalized geometric mean scale is equal to the normalized eigenvector scale. Interestingly, the two scales are always the same for judgment matrices of dimension less than or equal to three. To see this in the case n = 3, let

$$A = \begin{array}{ccc} 1 & a & b \\ 1/a & 1 & c \\ 1/b & 1/c & 1 \end{array}.$$

Then the geometric mean vector for A,

$$v = \begin{array}{c} (ab)^{1/3} \\ (c/a)^{1/3} \\ (1/bc)^{1/3} \end{array}$$

is an eigenvector for A corresponding to the eigenvalue

$$L = 1 + (ac/b)^{1/3} + (b/ac)^{1/3}.$$

Since L is of the form

$$1 + x + 1/x ,$$

its value is no less than 3, the dimension of A (with equality in the consistent case). Therefore w is an eigenvector corresponding to the

maximal eigenvalue for A. Hence in the case n = 3, the normalized geometric mean vector and the normalized eigenvector are the same. This result does not hold in general for inconsistent matrices with n greater than 3.

We have shown that given an arbitrary judgment matrix A, the geometric mean vector gives rise to the m-closest consistent matrix to A. The problem of representing a judgment matrix by a ratio scale can also be cast in the framework of the general linear model. We will show that under suitable assumptions the geometric mean vector is the maximum likelihood estimator for the ratio scale corresponding to a judgment matrix.

Let $A = [a_{i,j}]$ be an n × n judgment matrix. We assume that there is an underlying scale $(w_1, w_2, \ldots, w_n)$ whose ratios are perturbed to give the elements of A; namely,

$$a_{i,j} = \frac{w_i}{w_j} \, e_{i,j} \, , \tag{6.1}$$

and thus

$$\ln a_{i,j} = \ln w_i - \ln w_j + \ln e_{i,j} \, ,$$

$$i = 1,2,\ldots,n; \; j > i \, . \tag{6.2}$$

Suppose that the errors $e_{i,j}$ are independent and lognormally distributed with means 0 and variances σ^2. Making the substitutions

$$Y = \begin{matrix} \ln a_{1,1} \\ \ln a_{1,2} \\ . \\ . \\ . \\ \ln a_{n-1,n} \end{matrix} \, , \qquad B = \begin{matrix} \ln w_1 \\ . \\ . \\ \ln w_n \end{matrix} \, , \qquad E = \begin{matrix} \ln e_{1,1} \\ \ln e_{1,2} \\ . \\ . \\ . \\ \ln e_{n-1,n} \end{matrix} \, ,$$

we have the general linear equation

$$Y = XB + E$$

where the matrix X has components -1, 0, +1 determined by the
equations (6.2). In this framework it is well known [32] that the
maximum likelihood estimate for $B = [\ln w_i]$ is the least-squares
estimate given by

$$\widehat{b_i} = \frac{1}{n} \sum_{j=1}^{n} \ln a_{i,j} \; ,$$

and that the estimate has all of the usual desirable properties of
least-squares estimates under the general linear hypothesis. Taking
exponentials we obtain the maximum likelihood estimate of w_i:

$$\widehat{w_i} = \exp(b_i) = \prod_{j=1}^{n} a_{i,j}^{1/n} \; ,$$

the same estimate derived above from the metric m on the space of
judgment matrices.

The procedure outlined above can be modified to solve more
general estimation problems. For example, suppose that instead of
a single comparison for each pair of objects E_i and E_j we have
$n_{i,j}$ comparisons, where $n_{i,j}$ may be zero (reflecting missing data)
or greater than one (reflecting multiple comparisons, say by different
judges). The problem is then to find a w that minimizes the sum of
squares:

$$S = \sum_{i=1}^{n} \sum_{j>i} \sum_{k=1}^{n_{i,j}} [\ln a_{i,j,k} - (\ln w_i - \ln w_j)]^2$$

This generalization does not yield a simple closed-form solution such as the geometric mean vector, but in practice S can be minimized using standard least-squares regression packages.

VII. THE STATISTICAL MEASURE OF CONSISTENCY

The geometric mean estimation procedure outlined in Section VI leads to a natural measure of consistency for judgment matrices that is well-grounded in statistical theory and suited for use in statistical hypothesis testing.

Let A be an $n \times n$ judgment matrix whose components arise from multiplicative perturbations of ratios of an underlying scale w--that is

$$a_{i,j} = \frac{w_i}{w_j} e_{i,j} \quad ,$$

where $\ln e_{i,j}$, $i < j$ are independent random variables with mean zero and variance σ^2. It was shown in Section VI that the maximum likelihood estimate of w_i is the geometric mean of the elements of the $\underline{i}$th row,

$$\widehat{w_i} = \prod_{j=1}^{n} a_{i,j}^{1/n} \quad ,$$

Let s^2 be the residual mean square

$$s^2 = \frac{SS^2}{d.f.}$$

where

$$SS^2 = \sum_{i=1}^{n} \sum_{j>i} \ln a_{i,j} - \ln (w_i/w_j)^2 ,$$

$$d.f. = \frac{n(n-1)}{2} - (n-1) = \frac{(n-1)(n-2)}{2} .$$

Then s^2 is an unbiased estimator of σ^2 and is a natural measure of consistency of A.

Recall from Section VI that SS^2 can be viewed as the squared distance from A to the m-closest consistent matrix. Therefore s^2 is zero when A is consistent, is close to zero when A is close to consistent, and is far from zero when A is far from consistent. Moreover, since s^2 depends entirely on ratios of elements of A, it is invariant under scale changes.

VIII. EMPIRICAL COMPARISON OF THE TWO METHODS: A MONTE CARLO STUDY

The geometric mean vector gives an estimate for ratio scales
based on pairwise comparison data that is easy to calculate, satis-
fies the theoretical requirements of invariance under scale change
and transpose, and is well-grounded in statistical theory. It follows
from Section VI that under certain assumptions on the distribution of
the perturbations, the logarithms of the geometric mean vector compo-
nents are the *minimum variance unbiased estimators* of the logarithms
of the underlying ratio scale factors. In addition, there is empirical
evidence that the geometric mean vector is a better estimator of ratio
scales than the eigenvector when perturbations are not lognormally dis-
tributed, particularly in cases of extreme data inconsistency. This
section describes results of a Monte Carlo study comparing the geometric
mean vector and the eigenvector as estimators of ratio scales.

Suppose we start with an n $\times$ n consistent matrix A with elements

$$a_{i,j} = u_i/u_j .$$

Let $(e_{i,j}, i = 1,\ldots,n, j > i)$ be a collection of positive independent
random variables drawn from a suitable population, and construct a
perturbed matrix D with elements

$$d_{i,j} = \begin{cases} \dfrac{u_i}{u_j}\, e_{i,j} & \text{for } i = 1,2,\ldots,n, \ j > i \\[2ex] \dfrac{1}{d_{j,i}} & \text{for } i = 1,2,\ldots,n, \ j < i \\[2ex] 1 & \text{for } i = 1,2,\ldots,n, \ j = i \ . \end{cases}$$

Then D is a judgment matrix, but because of the perturbations it is
not necessarily consistent. Now one would hope that any estimate of

an underlying ratio scale for D would give rise to a consistent matrix
that is in some sense close to A. Two measures of closeness are used
in the Monte Carlo study described here: The sum of squares of errors

$$SSE = \sum_{i=1}^{n} [\widehat{u_i} - u_i]^2$$

and the sum of squares of the errors of logarithms.

$$SSL = \sum_{i=1}^{n} [\ln \widehat{u_i} - \ln u_i]^2 \quad ,$$

where

$$(u_i, \ i = 1,2,\ldots,n)$$

is the actual normalized ratio scale, and

$$(\widehat{u_i}, \ i = 1,2,\ldots,n)$$

is the estimated ratio scale.

The choice of population from which the perturbing random vari-
ables $e_{i,j}$ are drawn should reflect the fact that in a judgment matrix
constructed by an actual judge, errors are likely to be reciprocal
symmetric. That is, for any a,b > 0, the random variables $e_{i,j}$ should
satisfy:

$$P(a < e_{i,j} < b) = P(a < \frac{1}{e_{i,j}} < b) \quad .$$

Two convenient and quite dissimilar distributions satisfying this
property are a lognormal distribution whose underlying normal distri-

bution has mean zero and a distribution obtained from the ratio of
two independent, uniformly distributed random variables. We consider
perturbations $e_{i,j}$ drawn from populations with distributions of both
of these types.

Table 2 gives results of the Monte Carlo comparison of the geo-
metric mean vector and the eigenvector estimates when the underlying
ratio scale is perturbed by lognormal factors. For the first row,
ratios of a normalized scale consisting of five elements were per-
turbed by $e_{i,j}$ drawn from a lognormal population with $\sigma^2 = .01$. Then
the normalized maximal eigenvector and the normalized geometric mean
vector for the resulting judgment matrix were computed. This proce-
dure was repeated in 1000 similar trials, each time with the same
matrix dimension and population variance.

Table 2

MONTE CARLO COMPARISON OF GEOMETRIC MEAN VECTOR AND EIGENVECTOR
FOR RATIO SCALES PERTURBED BY LOGNORMAL ERRORS

Perturbation Variance	Scale Dimension	Sum of Squared Error Geometric Mean (1000 trials)	Sum of Squared Error Eigenvector (1000 trials)	Percentage of trials in which Geometric Mean Is smaller (Scales better)	Sum of Squared Error of Logs Geometric Mean (1000 trials)	Sum of Squared Error of Logs Eigenvector (1000 trials)	Percentage of trials in which Geometric Mean is smaller (Logs better)

The sums of squares of errors and sums of squares of errors of
logarithms were computed for individual trials and then totaled over
1000 trials for both the geometric mean vector and the eigenvector
estimates.

Columns 3 and 4 in the table give the total sum of squared errors
in 1000 trials. Column 5 gives the percentage of trials in which the

geometric mean vector is a better estimator in a least squares sense then the eigenvector. Columns 6, 7, and 8 present the same information for the sum of squares of errors of logarithms.

Comparisons based on 1000 trials were made for scales consisting of 5, 7, and 10 elements with perturbations drawn from lognormal populations with underlying normal population variances of .01, .16, .64, 1.0, and 2.0. The results in Table 2 indicate that the eigenvector and geometric mean vector give very close results when the variance is small--that is, when perturbations away from consistency are minimal. However, for larger perturbations, the geometric mean vector deviates less from its underlying ratio scale.

Table 3 gives results of the Monte Carlo comparison of the two methods when the underlying ratio scale is perturbed by random variables drawn from a population of ratios of uniform random variables. As in Table 2, underlying scales with 5, 7, and 10 elements are considered along with population variances for $\ln e_{i,j}$ of .01, .16, .64, 1.0, and 2.0. Again, there is very good agreement between the two estimators when variances are small. But for larger variances, the geometric mean vector is closer to the underlying scale in both the least squares and the least squares of logarithms sense.

Table 3

MONTE CARLO COMPARISON OF GEOMETRIC MEAN VECTOR AND EIGENVECTOR
FOR RATIO SCALES PERTURBED BY RATIO OF UNIFORM ERRORS

We conclude that in the presence of marked inconsistency, the geometric mean vector provides an estimate of underlying ratio scales that is empirically better than the maximal eigenvector.

IX.   EXAMPLE

In this section we discuss the use of the eigenvector and the geometric mean vector in a specific subjective judgment situation. Estimates of underlying utility vectors and consistency values derived from the two methods are compared.

Consider the automobile selection problem introduced in Sections II and V. The hierarchical structure for this problem is given in Fig. 1.

Subjective judgment data for this example were obtained from one prospective buyer. The buyer made pairwise comparisons reflecting his perceptions of the relative importance of the attributes of status, cost, economy, and size in selecting an automobile. Judgments were made based on the subjective judgment scale developed by Saaty (see p. 17). The buyer made comparisons in such a way that the resulting pairwise comparison matrix would be reciprocal symmetric (i.e., a judgment matrix). The resulting judgment matrix A is:

	Status	Cost	Economy	Size	
Status	1	1/5	1/5	1/2	
Cost	5	1	1	1/3	= A
Economy	5	1	1	1/2	
Size	2	3	2	1	

Next, judgment matrices were constructed from the buyer's pairwise comparisons of the five types of automobiles relative to each of the two attributes of status and size:

Subjective Comparison Relative to Status

	H	T	M	D	C	
H	1	2	1/9	1/6	1/7	
T	1/2	1	1/9	1/6	1/7	
M	9	9	1	6	4	= B
D	6	6	1/6	1	1/7	
C	7	7	1/4	7	1	

Subjective Comparison Relative to Size

	H	T	M	D	C	
H	1	1/2	7	3	8	
T	2	1	8	4	9	
M	1/7	1/8	1	1/3	3	= C
D	1/3	1/4	3	1	4	
C	1/8	1/9	1/3	1/4	1	

The eigenvector and the geometric mean vector were calculated separately for each of the three subjective judgment matrices above. Resulting scale estimates for the two techniques are given in Table 4.

Notice that the scales determined by the two methods are very close in value. This is as expected from theoretical considerations, because the two methods give the same results for consistent matrices and should agree closely for nearly consistent ones.

Table 4

COMPARISON OF SCALE ESTIMATES FOR THE AUTOMOBILE EXAMPLE

Matrix					Normalized Geometric Mean	Consistency s^2	Normalized Eigenvector	Consistency μ
Matrix A								
Status	1	1/5	1/5	1/2	.0812	.5680	.0885	.1513
Cost		1	1	1/3	.2453		.2425	
Economy			1	1/2	.2715		.2579	
Size				1	.4019		.4112	
Matrix B **Status**								
H	1 2	1/9	1/6	1/7	.0409	.5550	.0378	.1770
T	1	1/9	1/6	1/7	.0310		.0294	
M		1	6	4	.5307		.5239	
D			1	1/7	.1132		.1131	
C				1	.2842		.2958	
Matrix C **Size**								
H	1 1/2	7	3	9	.3152	.1150	.3143	.0417
T	1	8	4	9	.4632		.4630	
M		1	1/3	3	.0581		.0596	
D			1	4	.1299		.1288	
C				1	.0336		.0343	

The results in Table 4 indicate that the prospective buyer considers automobile size to be considerably more important than status, cost, or economy. He ranks T as having the best size of any of the cars under consideration.

Since exact values were available for cost and economy of the five automobile types, it was not necessary to compute scales for them from pairwise comparisons. The normalized scale values for the automobiles relative to these two attributes were determined to be:

	Cost Ratings	Economy Ratings
H	.2703	.3357
T	.4196	.2929
M	.0554	.1071
D	.1571	.1500
C	.0976	.1143

From each method, we now have scale estimates for the five automobile types relative to each of the four automobile attributes, as well as a scale of importance of the four attributes. Scale estimates for the five cars relative to each attribute are used as columns of a 5 × 4 matrix. This matrix is multiplied by the 4-dimensional vector of importance of attributes. The resulting 5-dimensional vector reflects the prospective buyer's ranking of the five automobiles. The calculations for both methods are carried out in Table 5.

Table 5
DETERMINATION OF RANKINGS OF FIVE AUTOMOBILES

Geometric Mean Vector Method

Make of Car	Status	Cost	Economy	Size	Criterion Scale	Final Scale
H	.0409	.2703	.3357	.3152	.0812	.2874
T	.0310	.4196	.2929	.4632	.2453	.3711
M	.5307	.0554	.1071	.0581	.2715	.1091
D	.1132	.1571	.1500	.1299	.4019	.1407
C	.2842	.0976	.1143	.0336		.0916

Columns combined as: (matrix) × (criterion scale) = (final scale)

Eigenvector Method

Make of Car	Status	Cost	Economy	Size	Criterion Scale	Final Scale
H	.0378	.2703	.3357	.3143	.0885	.2847
T	.0294	.4196	.2929	.4630	.2425	.3703
M	.5239	.0554	.1071	.0596	.2579	.1120
D	.1131	.1571	.1500	.1288	.4112	.1398
C	.2958	.0976	.1143	.0343		.0934

Columns combined as: (matrix) × (criterion scale) = (final scale)

REFERENCES

1. T. J. Gordon and O. Helmer, "Reporting on a Long-range Forecasting Study," The Rand Corporation, P-2982, September 1964.

2. R. U. Ayres, Technological Forecasting and Long-range Planning, McGraw-Hill, New York, 1969.

3. H. A. Linstone and M. Turoff, The Delphi Method--Techniques and Applications, Addison-Wesley Publishing Company, Inc., Reading, Mass., 1975.

4. H. Q. North and D. L. Pyke, "Technology, The Chicken--Corporate Goals, the Egg," 1st Annual Technology and Management Conference, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1968.

5. S. J. Press, "Qualitative Controlled Feedback for Forming Group Judgments and Making Decisions," The Rand Corporation, P-6290, January 1979.

6. J. von Neumann and O. Morgenstern, Theory of Games and Economic Behavior, 2nd ed., Princeton University Press, Princeton, New Jersey, 1947.

7. R. L. Keeney and K. Nair, "Decision Analysis for Siting of Nuclear Plants: The Relevance of Multiattribute Utility Theory," Technical Report No. 96, Operations Research Center, MIT, 1974.

8. R. L. Keeney, "A Decision Analysis with Multiple Objectives: The Mexico City Airport," Bell Journal of Economics and Management Science, Vol. 4, 1973, pp. 101-117.

9. C. W. Kelly, III, "Decision Analysis and Decision Aids," Proceedings of the Workshop on Decision Information for Tactical Command and Control, Airlie, Virginia, September 22-25, 1976, pp. 287-301.

10. K. R. MacCrimmon, "Improving the System Design and Evaluation Process by the Use of Trade-off Information: An Application to Northeast Corridor Transportation Planning," The Rand Corporation, RM-5877-DOT, April 1969.

11. H. Raiffa, "Preferences for Multiattributed Alternatives," The Rand Corporation, RM-5868-DOT/RC, April 1969.

12. R. L. Keeney and H. Raiffa, Decisions with Multiple Objectives: Preferences and Value Tradeoffs, John Wiley and Sons, New York, 1976.

13. J. V. Sigford and R. H. Parvin, "Project PATTERN: A Methodology for Determining Relevance in Complex Decision-Making," IEEE Transactions on Engineering Management, Vol. EM-12, March 1965.

14. H. A. Wells, "Weapon System Planner's Guide," IEEE Transactions on Engineering Management, Vol. EM-14, March 1967, pp. 14-16.

15. W. S. Torgerson, Theory and Methods of Scaling, John Wiley and Sons, New York, 1958.

16. C. W. Churchman and R. L. Ackoff, "An Approximate Measure of Value," Journal of the Operations Research Society of America, Vol. 2, No. 2, May 1954, pp. 172-187.

17. H. A. David, The Method of Paired Comparisons, Hafner Press, Macmillan, New York, 1963.

18. R. R. Davidson and P. H. Farquhar, "A Bibliography on the Method of Paired Comparisons," Biometrics, Vol. 32, June 1976, pp. 241-252.

19. T. L. Saaty and J. P. Bennett, "A Theory of Analytical Hierarchies Applied to Political Candidacy," Behavioral Science, Vol. 22, 1977, pp. 237-245.

20. T. L. Saaty, "The Sudan Transport Study," Interfaces, Vol. 8, No. 1, Part 2, November 1977, pp. 37-57.

21. J. M. Alexander and T. L. Saaty, "The Forward and Backward Processes of Conflict Analysis," Behavioral Science, Vol. 22, 1977, pp. 87-98.

22. T. L. Saaty, "Scenarios and Priorities in Transport Planning: Application to the Sudan," Transportation Research, Vol. 11, 1977, pp. 343-350.

23. T. L. Saaty, Hierarchies and Priorities--Eigenvalue Analysis, a paper dedicated to the Memory of J.E.I. Heller, University of Pennsylvania, 1977 (draft).

24. T. L. Saaty and M. W. Khouja, "A Measure of World Influence," Journal of Peace Science, Vol. 2, No. 1, Spring 1976, pp. 31-48.

25. T. L. Saaty, "Estimating Technological Coefficients by Hierarchical Measurement," The Wharton School, University of Pennsylvania, 1977 (draft).

26. "Decision Information for Tactical Command and Control," Proceedings of the Workshop on Decision Information for Tactical Command and Control, Airlie, Virginia, September 22-25, 1976.

27. T. L. Saaty and P. C. Rogers, "Scenario Construction Using a Hierarchical Framework with Eigenvector Weighting," *Socio-Economic Planning Sciences*, Vol. 10, Pergamon Press, Great Britain, 1976, pp. 251-263.

28. T. L. Saaty, "A Scaling Method for Priorities in Hierarchical Structures," *Journal of Mathematical Psychology*, Vol. 15, pp. 234-281.

29. T. L. Saaty and P. C. Rogers, "Higher Education in the United States (1985-2000)," *Socio-Economic Planning Sciences*, Vol. 10, No. 6, 1976, pp. 251-263.

30. T. L. Saaty, "Operation Gaming for Energy-Policy Analysis," *Energy Policy*, Vol. 5, No. 1, 1977, pp. 63-75.

31. J. N. Franklin, *Matrix Theory*, Prentice-Hall, Englewood Cliffs, New Jersey, 1968.

32. H. Scheffe, *The Analysis of Variance*, John Wiley and Sons, New York, 1959.