DECISION-MAKING WITH THE AHP: WHY IS THE PRINCIPAL EIGENVECTOR NECESSARY ?

Thomas L. Saaty
University of Pittsburgh
Pittsburgh, PA 15260 – U.S.A.
saaty@katz.pitt.edu

Keywords: complex order**,** consistent, near consistent, priority, principal eigenvector

Summary*: We will show here that the principal eigenvector of a matrix is a necessary representation of the priorities derived from a positive reciprocal pairwise comparison consistent or near consistent matrix. A positive reciprocal n by n consistent matrix $W = (w_{ij})$ satisfies the relation $w_{ik} = w_{ij} w_{jk}$. If the principal eigenvector of W is $w=(w_1, \dots ,w_n)$ the entries of W may be written as $w_{ij} = w_i /w_j$. Matrices $A = (a_{ij})$ that are obtained from W by small positive reciprocal perturbations of w_{ij} , known as near consistent matrices, are pivotal in representing pairwise comparison numerical judgments in decision making. Since W is always of the form $w_{ij} = (w_i/w_j)$ a perturbation of W is given by $a_{ij} = (w_i/w_j) \varepsilon_{ij}$ and their corresponding reciprocals $a_{ji} = (w_j/w_i) (1/\varepsilon_{ij})$. A priority vector for use in decision making that captures a linear order for the n elements compared in the judgment matrix can be derived for both consistent and near consistent matrices, but it is meaningless for strongly inconsistent matrice except if they are the result of random situations that have associated numbers such as game matches where the outcomes do not depend on judgment. It is shown that the ratios w_i /w_j of the principal eigenvector of the perturbed matrix are close to a_{ij} , if and only if the principal eigenvalue of A is close to n. We then show that if in practice we can change some of the judgments in a judgment matrix, it is possible to transform that matrix to a near consistent one from which one can then derive a priority vector. The main reason why near consistent matrices are essential is that human judgment is of necessity inconsistent, which if controllable, is a good thing. In addition, judgment is much more sensitive and responsive to large than to small perturbations, and hence once near consistency is reached, it becomes uncertain which coefficients should be perturbed with small perturbations to transform a near consistent matrix to a consistent one. If such perturbations were forced, they would seem arbitrary and can distort the validity of the derived priority vector in representing the underlying real world problem.*

1. Introduction

In the field of decision-making, the concept of priority is of paramount importance and how priorities are derived can make a decision come out right or wrong. They must be unique and not one of many possibilities, they must also capture the order expressed in the judgments of the pairwise comparison matrix. The idea of a priority vector has much less validity for an arbitrary positive reciprocal matrix A than it does for a consistent and a near consistent matrix. It is always possible to use metric approximation to find a vector *w* the ratios of whose coefficients are according to some metric "close" to the coefficients of *A* without regard to consistency. The judgments in A may be given at random and may have nothing to do with understanding relations in a particular decision. Closeness to the numerical values of the coefficients says little about order. In this paper we show that the **principal eigenvector,** known in mathematics to be unique to within a positive multiplicative constant, is the **only possible candidate** in the quest **for deriving priorities**.

The fact that the AHP allows inconsistency is attributable to the fact that in making judgments people are naturally cardinally inconsistent and ordinally intransitive. For several reasons this is a good thing, otherwise people would be like robots unable to change their minds with new evidence and unable to look within for judgments which represent their thoughts and feelings.

Topology has at least two great branches, *metric topology* and *order topology* [9], (applied to both discrete and continuous topological spaces). In judgment matrices we deal with a discrete number of judgments that take on discrete scale values. In metric topology the central idea is that of a metric or distance usually minimized to find close approximations to solutions. One example is the method of least squares (LSM) which determines a priority vector by minimizing the Frobenius norm of the difference between the judgment matrix *A* and a positive rank one reciprocal matrix (y_i/y_j):

$$\min_{y \succ 0} \sum_{i,j=1}^{n} (a_{ij} - y_i/y_j)^2$$

Another is the method of logarithmic least squares (LLSM) which determines a priority vector by minimizing the Frobenius norm of $(\log a_{ij} x_j/x_i)$:

$$\min_{x \succ 0} \sum_{i,j=1}^{n} [\log a_{ij} - \log(x_i/x_j)]^2$$

In general, different methods give different priority vectors and different rankings. Here is an example:

	LLSM	LSM
$\begin{bmatrix} 1 & 2 & 5 \\ 1/2 & 1 & 7 \\ 1/5 & 1/7 & 1 \end{bmatrix}$	.541 .382 .077	.410 .514 .076

LSM does not always yield a unique priority vector. In this case, a second LSM solution is (.779, .097, .124); both solutions yield the desired minimum value 71.48. Here, both LSM and LLSM are square metrics and are examples of only one of an infinite number of potential metrics and an infinite number of possible (often contradictory) priority solutions resulting from these metrics. I considered both these methods in some detail in my first book on the AHP, *The Analytic Hierarchy Process*, and had to reject them. That was basically because they did not capture the idea of dominance and its intensity, that naturally emerge in the theory of graphs, their adjacency, path and circuit matrices and the corresponding fundamental applications in electrical engineering.

For a discussion of order, we need to introduce the concept of *simple or linear order*. A set is said to be simply ordered if its elements satisfy the two conditions: 1) For any two elements x and y exactly one of the relations x < y, x = y, y < x be valid (dominance), and 2) If x < y and y < z, then x < z (transitivity). Extracting a simple order from a matrix of judgments at first seems difficult because both conditions may be violated when that matrix is inconsistent. A priority vector $p = (p_1, \ldots, p_n)$ is a simple (cardinal) ordering of the elements, that is derived from a judgment comparison matrix *A* according to the order among its coefficients, and each of whose ratios p_i/p_j is close to its corresponding a_{ij}, the degree of closeness of all ratios together determined by the closeness of the principal eigenvalue to n, the order of the matrix. We call a priority vector *strong* when the two conditions hold, and *weak* when closeness does not hold. For judgment matrices, the priority vector must be strong; for matrices involving information on, for example, numbers of goals scored in a sport or the number of wins and losses in the games and sports of matches, a weak priority vector may be adequate. The principal eigenvector of an arbitrary (inconsistent) positive reciprocal matrix is a weak priority vector, that of a near consistent matrix, a strong priority vector and finally that of a matrix that is only positive, is unlikely to be a priority vector.

Not only do the elements being compared have an order, but also the coefficients in the matrix, the judgments, have their order. We refer to that order as *complex order*. As we shall see below, the notion of complex order involves both dominance and transitivity, in compounded forms. Our problem, then, is to derive a simply ordered priority vector from the complex order of the coefficients. It is meaningless to speak of a priority vector for a very inconsistent matrix because even though an order may be obtained by

the principal eigenvector, or any other way, such as minimization used in the metric approach, the ratios of the solution vector would not be close to the a_{ij}.

It is known in graph theory that the number of paths of length k between two vertices of a graph can be obtained from the kth power of the adjacency matrix of that graph. A judgment matrix is an adjacency matrix whose entries represent strengths of dominance. The powers of a judgment matrix capture the transitivity of dominance or influence among the elements. To capture order in a priority vector, we need to consider all order transitions among the coefficients and their corresponding powers of A. First order transitions are given in the matrix itself, second order ones in its square and so on. Thus powers of A and their convergence become a central concern. The dominance in each of these matrices are obtained by the vector of normalized row sums, and the overall dominance is obtained by taking the average of these results, their Cesaro sum, known to converge to the same limit as the powers of the matrix. See next section. An upshot of the order approach is that it yields a unique outcome vector for the priorities that transforms the complex order among the coefficients to a linear order. The priority vector is given by the principal eigenvector of the judgment matrix. In a consistent matrix if $a_{ij} > a_{ik}$ for all j and k, that is if one row dominates another row, the priority of the ith element is greater than the priority of the kth element, which is to be expected. For an inconsistent matrix, it is inadequate to use the rows of the matrix, but it turns out that for arbitrarily large powers of the matrix [10], the rows acquire this characteristic, and thus in the limit yield the desired linear order.

In our case the existence of the eigenvector can be shown and its computation carried out in two ways: 1) by using the theory of Perron which applies to positive matrices (and our judgment matrices are positive), or 2) by using multiplicative perturbation theory applied to a positive reciprocal consistent matrix (and our judgment matrices are not only positive, but also reciprocal). A consistent matrix already has the Perron properties without need for the theorem of Perron.

Some advocates of the geometric mean (LLSM), which for a matrix of order n > 3 can give different rank order priorities than the eigenvector, have borrowed a failed concept from utility theory, that of always preserving rank, to justify using it as a priority vector. That there are many examples in the literature in which both preference and rank can reverse without introducing new criteria or changing priorities has not yet caught their attention. Numerous copies of an alternative and deliberately planted "phantom" alternatives among many others can cause rank reversals. If rank is to be preserved in some decisions and allowed to reverse in others, even when the problem and all its elements (criteria, alternatives, judgments) are the same, as has been convincingly argued in the AHP and other decision making approaches. Whether one wants to allow rank to reverse or not must depend on considerations outside decision theory. Thus we must have two ways to derive the priorities, one to preserve rank and one to make it possible for it to change [8] as the AHP does with its distributive and ideal modes. We know of no other decision theory that has this essential flexibility. To always preserve rank can lead to poor decisions as Luce and Raiffa [7] write in their book, after proposing to always preserve rank as a fourth variant of their axiom on rank. 'The all-or-none feature of the last form [i.e. always preserving rank] may seem a bit too stringent … **a severe criticism is that it yields unreasonable results**." Indeed it does and their warning is so far inadequately heeded by practitioners who seem oblivious to it and are willing to leave the world in the dark about this very crucial aspect of decision-making. A method that always preserves rank may not yield the correct rank-order as needed in many decision problems.

Other people have thought to create their own geometric mean version of the AHP by raising the priorities of the normalized alternatives to the power of the criteria multiplying across the criteria and taking the root to do hierarchic synthesis. A trivial example with dollars for two criteria (C_1,C_2) and three alternatives (A_1,A_2,A_3) with dollar values of (200,300,500) and (150,50,100) under C_1 and C_2 respectively defines the relative dollar importance of the criteria of (1000/1300,300/1300). Straight addition and normalization gives the correct answer (350/1300, 350/1300,600/1300) = (.269,.269,.462) that is also obtained by the usual AHP method of hierarchic composition by multiplying the normalized values of the alternatives by the priorities of the criteria and adding. However, the proposed procedure of raising the normalized priorities of the alternatives to the power of the criteria and taking the cubic root yields the wrong outcome of (.256,.272,.472). One must never use such a process because of several failures among which are its failure to give the correct decision [11], and its failure to be a meaningful way of synthesis in case of

dependence and feedback as in the Analytic Network Process (ANP) [12]. The survival of any idea in decision-making depends on what it does and how it performs when there is feedback.

Emphasis on the calculus and its methods of solution has made some people think that all real life problems can be solved through optimization methods that rely on metric approximation. But that is not true. By their definition, priorities require order in their construction.

Statistical methods are no more secure in dealing with questions of order and rank because they are fundamentally metric methods of closeness. With the one exception that follows the story of order and dominance told below, in their present state of development statistical methods cannot be counted on to correctly judge the effectiveness of ways to derive priorities from judgments.

2. The Priority Vector of a Consistent Matrix *A*

A consistent matrix is a positive reciprocal *n* by *n* matrix whose coefficients satisfy the relation $a_{ij} a_{jk} = a_{ik,}$, i,j,k = 1,…, n. We first show that it is necessary that the priority vector $w = (w_1,…, w_n)$ of a consistent matrix also be the principal eigenvector of that matrix. Before doing that we show (without using the theory of Perron) that a consistent matrix has a principal eigenvalue and a corresponding principal eigenvector. To prove that, we use the fact that a consistent matrix is always of the form $W = (w_i/w_j)$. It follows that $Ww = nw$ and because *W* has rank one, *n* is its largest eigenvalue and *w* is its corresponding eigenvector and the coefficients of *W* coincide with the coefficients of the matrix of ratios from the vector *w*. Actually we see in this case that the coefficients in the Hadamard product $WoW^\top$ are all equal to one and their sum is equal to n^2 and if averaged by dividing by the number of coefficients which is also n^2, the outcome is equal to one, and closeness applies perfectly in this case. The priority vector is the same for all metric and order methods in this case. It is the vector *w*.

Next we show that order holds for a consistent matrix *A*. Element A_i is said to *dominate* element A_j in one step, if the sum of the entries in row *i* of *A* is greater than the sum of the entries in row *j*. It is convenient to use the vector $e = (1,…,1)^T$ to express this dominance: Element A_i *dominates* element A_j in one step if $(Ae)_i > (Ae)_j$. An element can dominate another element in more than one step by dominating other criteria that in turn dominate the second criterion. Two-step dominance is identified by squaring the matrix and summing its rows, three-step dominance by cubing it, and so on. Thus, A_i dominates A_j in k steps if $(A^k e)_i > (A^k e)_j$. Element A_i is said simply to dominate A_j if entry *i* of the vector obtained by averaging over the one step dominance vector, two step dominance vector, k step dominance vector and passing to the limit:

$$\lim_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} A^k e / e^T A^k e$$

is greater than its entry *j*. But this limit of weighted averages (the Cesaro sum) can be evaluated: We have for an n by n consistent matrix A: $A^k = n^{k-1} A$, $A = (w_i/w_j)$ and the foregoing limit is simply the eigenvector *w* normalized. In general, it is also true that the Cesaro sum converges to the same limit as does its kth term $A^k e / e^T A^k e$ that yields k step dominance.

Here we see that the requirement for rank takes on the particular form of the principal eigenvector. We will not assume it for the inconsistent case but prove its necessity again for that more general case.

We have shown that for a consistent matrix the principal eigenvector both yields ratios close to the coefficients of the matrix and captures the rank order information defined by them. It is thus the priority vector. The foregoing proof is also valid for all methods proposed so far for obtaining the priority vector. They break down in the next stage, when *A* is inconsistent, but the eigenvector approach can be generalized as we see next.

3. A Near Consistent Matrix and its Priority Vector

A near consistent matrix is a matrix that is a small reciprocal (multiplicative) perturbation of a consistent matrix. It is given by the Hadamard product: $A = WoE$ where $W = (w_i/w_j)$ and $\boldsymbol{E} \equiv (\varepsilon_{ij}), \varepsilon_{ji} = \varepsilon_{ij}^{-1}$.

Small means ε_{ij} is close to one. Although we make use of perturbation theory here with proofs using additive perturbations, for small perturbations, we do not distinguish as to whether they are additive or multiplicative. With slight alteration, the same proofs apply for both so long as they are small.

We need near consistent pairwise comparison matrices for decision-making. The reason is that when human judgment is represented numerically, it is inevitably inconsistent. That is a good thing because judgment is approximate and uncertain, and forcing it to be consistent can distort relations that need not be as precise as required by consistency. Further, judgment is more sensitive and responsive to large changes than small ones. Once near consistency is reached, it becomes difficult to decide on which coefficients to perturb by a small amount to transform a near consistent matrix to one that is even closer to consistency. If very small perturbations are forced, they would seem arbitrary and if carried out, can distort the validity of the derived priority vector in representing the underlying real world problem. In other words, the priority vector of a near consistent decision matrix can be a more accurate representation of the understanding reflected in the judgments than the priority vector of some consistent matrix to which it is transformed by a sequence of small perturbations.

First we apply perturbation theory to show that a matrix that is a small perturbation of a consistent matrix also has a simple positive and real principal eigenvalue and corresponding eigenvector. We then show that the resulting eigenvector is the priority vector. We note that because a near consistent matrix has a principal eigenvector, it can be written as a reciprocal perturbation of the matrix of ratios W corresponding to that vector. Also there can be many near consistent matrices that have that same vector as an eigenvector. Thus all near consistent matrices, and in fact even more generally, all positive reciprocal matrices can be put into equivalence classes according to the eigenvector they have in common. Beyond near consistency, positive reciprocal matrices only interest us in so far as we can transform them by voluntary, not artificially forced, and plausible changes in judgments that lead to the perturbations that concern us here.

Remark. Perron's theorem [4,6] says that a positive matrix always has a single positive real eigenvalue that dominates all other eigenvalues in modulus and to which corresponds an eigenvector with positive real entries that is unique to within multiplication by a positive constant. To this maximum eigenvalue and eigenvector are often attached two names, "principal" and "Perron" mostly the latter. What do these facts and names have to do with a positive reciprocal matrix A? Clearly, because A is positive, it has a Perron eigenvalue and eigenvector. What does the fact that A is reciprocal add to the characterization of its eigenstructure? It turns out that for the special class of near consistent matrices, one can keep the label "principal," but can drop the label "Perron" because the principal eigenvalue and eigenvector are derived by perturbation from a consistent matrix that already has them. We do not need the theorem of Perron to do that.

Here we provide a different proof for the validity of this for near consistent matrices without using Perron's arguments. We very much admire and value the theory of Perron, and can use it directly here. But, given that we have Perron's results automatically for a consistent matrix without Perron's proofs, we believe that we must also derive its conclusions, rather than invoking them, on near consistent positive reciprocal matrices.

4. Background on Perturbation

We wish to prove, **without** assuming the theorem of Perron, that a positive n by n reciprocal matrix $A=(a_{ij})$, $a_{ji} = a_{ij}^{-1}$ that is near consistent (is a "small perturbation" of a consistent matrix: $a_{ij}\,a_{jk} = a_{ik}$, i,j,k = 1,...,n), has a simple, real and positive eigenvalue that dominates all other eigenvalues in modulus and to which corresponds a positive eigenvector that is unique to within a multiplicative constant.

From $a_{ij} = a_{ik} / a_{jk}$ we have $a_{ji} = a_{jk} / a_{ik} = a_{ij}^{-1}$ and a consistent matrix is reciprocal. From the validity of the statement of the theorem for consistent matrices we would like to show that it is also true for near consistent positive reciprocal matrices.

Theorem 1 is old, relates to the roots of polynomials and was known in the 19th century [3] (small perturbation of the coefficients of a polynomial equation that leads to a small perturbation of its roots).

Theorem 1: If an arbitrary matrix $A=(a_{ij})$ has the eigenvalues $\mu_1,...,\mu_s$, where the multiplicity of μ_h is m_h with $\sum_{h=1}^{s} m_h = n$, then given $\varepsilon > 0$, there is a $\delta=\delta(\varepsilon) > 0$ such that if $\left| a_{ij} + \gamma_{ij} - a_{ij} \right| = \left| \gamma_{ij} \right| \leq \delta$, for all i and j, the matrix $B = (a_{ij} + \gamma_{ij})$ has exactly m_h eigenvalues in the circle $\left| \lambda - \mu_h \right| < \varepsilon$ for each $h = 1,...,s$.

<u>Proof [3]</u>: Define $f(\lambda,B) = \det(\lambda I - B)$. Let $\varepsilon_o = \frac{1}{2}\min\left| \mu_i - \mu_j \right|$, $1 \leq i < j \leq s$ and let $\varepsilon < \varepsilon_o$. The circles C_h: $\left| \lambda - \mu_h \right| = \varepsilon$, h=1,...,s are disjoint. Let $r_h = \min_{\lambda \in C_h} | f(\lambda,B)|$; r_h is well defined since f is a continuous function of λ, and $r_h > 0$ because the roots of $f(\lambda,B) = 0$ are the centers of the circles. The function $f(\lambda,B)$ is continuous in the $1+n^2$ variables λ and $a_{ij} + \gamma_{ij}$, i,j=1,...,n and hence for some $\delta > 0$, $f(\lambda,B) \neq 0$ for λ on any C_h, h=1,...,s if $\left| \gamma_{ij} \right| \leq \delta$, i,j=1,...,n.

From the theory of functions of a complex variable, the number m_h of roots λ of $f(\lambda,B)=0$ that lie inside C_h is given by $m_h(B) = \frac{1}{2\pi i} \int_{C_h} \frac{f'(\lambda,B)}{f(\lambda,B)} d\lambda$, h=1,...,s which is also a continuous function of the $1+n^2$ variables in C_h and $\left| \gamma_{ij} \right| \leq \delta$, i,j=1,...,n. In particular it is a continuous function of $a_{ij} + \gamma_{ij}$ with $\left| \gamma_{ij} \right| \leq \delta$.

For $B = A$ we have $n_h(A) = m_h$, h=1,...,s. Since the integral is continuous it cannot jump from $n_h(A)$ to $n_h(B)$ and the two must be equal and have the common value m_h, h=1,...,s for all B with $\left| \gamma_{ij} \right| \leq \delta$, i,j=1,...,n.

Corollary : If $W = (w_{ij} \equiv \frac{w_i}{w_j})$ is consistent, then given $\varepsilon > 0$, there is a $\delta=\delta(\varepsilon) > 0$ such that if $\left| \gamma_{ij} \right| \leq \delta$, uniformly for all i and j, the matrix $B = (W_{ij} + \gamma_{ij})$ has exactly one positive real eigenvalue in the circle $\left| z - n \right| < \varepsilon$ that is largest in modulus, and all other eigenvalues of B lie within the circle $\left| z \right| < \varepsilon$.

<u>Proof</u>: The only non-zero eigenvalue of a consistent matrix is n = Trace (A). Theorem 1 implies that B has exactly one eigenvalue, λ_1, in the circle $\left| \lambda - n \right| < \varepsilon$. It is real, for otherwise, if it were complex both it and its complex conjugate would have the same modulus and both would fall in the circle $\left| \lambda - n \right| < \varepsilon$

contradicting the fact that $|\lambda - n| < \varepsilon$ contains only one eigenvalue of B. Further, $|\lambda_1 - n| < \varepsilon$ implies $n - \varepsilon < \lambda_1 < n + \varepsilon$ and hence $\lambda_1 > 0$. By Theorem 1, all other eigenvalues must be in the circle $|\lambda| < \varepsilon$. Thus $\lambda_1 \equiv \lambda_{max}$ is real and positive and dominates all other eigenvalues in modulus.

Theorem 2, stronger than theorem 1, is important in practical applications and is more recent. Its proof is found in [5]. In the reference the authors also show how to obtain the coefficients in the series expansions.

Theorem 2: If 8 is an eigenvalue of A of multiplicity 1, then for sufficiently small $*t*$, there is an eigenvalue 8(t) of A(t) such that 8(t) = 8 + t $8^{(1)}$ + $t^2 8^{(2)}$ + ... Also there are right and left eigenvectors w(t) and v(t), respectively, associated with 8(t) for which $v^{(T)}$(t) w(t) = 1, and w(t) = w + t $w^{(1)}$ + $t^2 w^{(2)}$ + ... , v(t) = v + t $v^{(1)}$ + $t^2 v^{(2)}$ + ...

The following simple and elegant proof of this theorem was communicated to me by Professor Jerry L. Kazdan of the University of Pennsylvania. It can be extended to the left eigenvector without difficulty.

Given a polynomial $p(x,t) = x^n + a_{n-1}(t)x^{n-1} + \cdots + a_1(t)x + a_0(t)$ whose coefficients depend smoothly on a parameter t. Assume at $t = 0$ the number $s = c$ is a simple root of the polynomial $p(c,0) = 0$. Show that for all t sufficiently near 0 there is a unique root $x(t)$ with $x(0) = c$ that depends smoothly on t. To see this note that given $p(c,0) = 0$ we want to solve $p(x,t) = 0$ for $x(0) = c$. This is immediate from the implicit function theorem. Since $x(0) = $ c is a simple zero of $p(c,0) = 0$, the derivative $p_x(c,0)$ is not zero.

The example $p(x,t) := x^3 - t = 0$, so $(x,t) = t^{1/3}$, shows this may not be true at a multiple root.

Given a square matrix $A(t)$ whose elements depend smoothly on a real parameter t, if $8 = 8_0$ is a simple eigenvalue at $t = 0$, show that for all t near 0 there is a corresponding eigenvalue and (normalized) eigenvector that depend smoothly on t.

Although we will not use it, the eigenvalue part is immediate from the previous problem. It is the eigenvector aspect that takes a bit more work.

Given $A(0)X_0 = 8_0$ for some vector X_0 with $\|X_0\| = 1$, we want a function $8(t)$ and a vector $X(t)$ that depend smoothly on t with the properties

$$A(t)X(t) = 8(t)X(t), \langle X_0, X(t)\rangle = 1, \text{ and } 8(0) = 8_0, X(0) = X_0.$$

We could also have used the slightly more complicated normalization $\|X(t)\|^2 = 1$.

Of course the proof uses the implicit function theorem.

SOME BACKGROUND ON THE IMPLICIT FUNCTION THEOREM. If $H : \;|^{\rceil} H \;|\rightarrow |^{\rceil}$, say we want to solve the equations $H(Z,t) = 0$ for $Z = Z(t)$. These are N equations for the N unknowns $Z(t)$. Assume that $Z = Z_0$ is a solution at $t = 0$, so $H(Z_0,0) = 0$. Expanding H in a Taylor series in the variable Z near $Z = Z_0$ we get

$$H(Z,t) = H(Z_0,t) + H_Z(Z_0,t)(Z - Z_0) + \cdots,$$

Where H_Z is the derivative matrix and $\cdots$ represent higher order terms. If these higher order terms were missing then the solution of $H(Z,t) = 0$ is simply $Z - Z_0 = - [H_Z(Z_0,t)]^{-1} H(Z_0,t)$, that is $Z = Z_0 - [H_Z(Z_0,t)]^{-1} H(Z_0,t)$. This assumes that the first derivative matrix $H_Z(Z_0,0)$ is invertible for all t near zero). If there are higher order terms, the *implicit function theorem* says that there is still a solution $Z(t)$. The key assumption is that the first derivative matrix $H_Z(Z_0,0)$ is invertible. Although we assume $t\ 0\ |$, the identical ideas work if the parameter $t\ 0\ |^k$.

We may assume that $8(0) = 0$. Write

$$F(X,\lambda,t) = \left.\begin{pmatrix} A(t) - \lambda X \\ \langle X_0, X \rangle - 1 \end{pmatrix}\right|_{t=0}$$

To solve: $F(X, \mathscr{S}, t) = 0$ for both $X(t)$ and $\mathscr{S}(t)$ near $t = 0$. In the notation of the previous paragraph, $Z = (X, \mathscr{S})$ and $H(Z,t) = F(X, \mathscr{S}, t)$. Thus the derivative matrix H_Z involves differentiation with respect to both X and $\mathscr{S}$.

We now compute the derivate matrix at $t = 0$ with respect to the parameters X and $\mathscr{S}$. The directional derivative with respect to X in the direction of a vector V is

$$\left.\frac{d}{ds} F(X + sV, \lambda, t)\right|_{s=0} = \begin{pmatrix} A(t) - \lambda V \\ \langle X_0, V \rangle \end{pmatrix}$$

while the directional derivative with respect to $\mathscr{S}$ in the direction of a scalar r is

$$\left.\frac{d}{ds} F(X, \lambda + sr, t)\right|_{s=0} = \begin{pmatrix} -rX \\ 0 \end{pmatrix}$$

Thus at $t = 0$ the derivative with respect to X and $\mathscr{S}$ is the partitioned matrix

$$F'(X,\lambda,t) \begin{pmatrix} V \\ r \end{pmatrix} = \begin{pmatrix} A(0) & -X_0 \\ X_0 & 0 \end{pmatrix} \begin{pmatrix} V \\ r \end{pmatrix}$$

For the implicit function theorem we need the matrix on the right to be invertible. It is enough to show its kernel is zero. Thus, say $A(0) V - X_0 r = 0$ and $\langle X_0, V \rangle = 0$. From the first equation we find $A(0)^2 V = rA(0)X_0 = 0$.

By assumption, the eigenvalue $\mathscr{S}_0 = 0$ has multiplicity one so $V = $ (const) X_0. But then $\langle X_0, V \rangle = 0$ gives $V = 0$. Consequently also $r = 0$.

Since the derivative matrix $F''(X, \mathscr{S}, 0)$ is invertible, by the implicit function theorem the equation $F(X, \mathscr{S}, t) = 0$ has the desired smooth solution near $t = 0$.

Remark: We note that a small perturbation of a nonnegative not necessarily reciprocal matrix can give rise to a principal eigenvalue that is not near the original eigenvalue. The nonnegative matrix $A = (a_{ij})$ with $a_{i,i+1} = 1$, $a_{ij} = 0$ otherwise, has all its eigenvalues equal to zero. The same matrix with a_{n1} replaced by ε, where $\varepsilon > 0$ is small, has the maximum eigenvalue $\lambda_{max} = \varepsilon^{1/n}$ which tends to one with increasing n. Although λ_{max} changes continuously with the coefficient ε, its value becomes large for small ε. Yet this example does not violate the perturbation theorems above.

Given the positive vector $w = (w_1, \ldots, w_n)$ and the matrix $W = (\dfrac{w_i}{w_j})$ derived from w, a reciprocal perturbation of W is a matrix $A = (a_{ij})$ where $a_{ij} = \dfrac{w_i}{w_j} \varepsilon_{ij}$, $\varepsilon_{ji} = \varepsilon_{ij}^{-1}$, thus the matrix $E \equiv (\varepsilon_{ij})$ is positive and reciprocal . A reciprocal perturbation is said to be small if ε_{ij} is close to one for all i and j. We have:

Theorem 3: Let $A(t)$ be a positive reciprocal matrix that is a small reciprocal perturbation of a consistent matrix W, then A has a simple positive real eigenvalue λ_{max} that dominates all other eigenvalues of A in modulus and to which correspond right and left eigenvectors w and v, respectively.

Thus for a near consistent matrix we have a principal eigenvalue and eigenvector without use of the theorem of Perron. The problem remains as to how to bring an arbitrary judgment matrix that is a positive reciprocal matrix to have a status of near consistency and what theory justifies it.

Note that with a reciprocal perturbation we ensure that $\lambda_{max} \geq n$ which helps determine the validity of w as a priority vector. We have

$$\sum_{j=1}^{n} \varepsilon_{ij} = \sum_{j} a_{ij} w_j / w_i = [Aw]_i / w_i = \lambda_{max} w_i / w_i = \lambda_{max}.$$

The computation

$$n\lambda_{max} = \sum_{i=1}^{n}(\sum_{j=1}^{n} \varepsilon_{ij}) = \sum_{i=1}^{n} \varepsilon_{ii} + \sum_{\substack{i,j=1 \\ i \neq j}}^{n} (\varepsilon_{ij} + \varepsilon_{ji}) = n + \sum_{\substack{i,j=1 \\ i \neq j}}^{n} (\varepsilon_{ij} + \varepsilon_{ij}^{-1}) \geq n + (n^2 - n)/2 = n^2 \text{ reveals}$$

that $\lambda_{max} \geq n$. Moreover, since $x + 1/x \geq 2$ for all x > 0, with equality if and only if $x = 1$, we see that $\lambda_{max} = n$ if and only if all $\gamma_{ij} = 1$, which is equivalent to having all $a_{ij} = w_i / w_j$. The foregoing arguments show that a positive reciprocal matrix A has $\lambda_{max} \geq n$, with equality if and only if A is consistent.

5. The Priority Vector of a Near Consistent Matrix

Let us prove that the principal eigenvector is the priority vector of a near consistent matrix without using Perron's proof. We use [2, p. 175] for a part of our proof.

Note that the Hadamard product $A = Wo\,E$ can be written as $A = D_v\,E\,D_v^{-1}$ where D_v and D_v^{-1} are diagonal matrices with the vectors w and w^{-1} on their diagonals respectively.

Theorem 4: An arbitrary near consistent positive reciprocal matrix A satisfies $\lim_{k\to\infty} \dfrac{A^k e}{e^T A^k e} = cw$, where c is an arbitrary constant.

<u>Proof</u>: $a_{ij} = \dfrac{w_i}{w_j} \varepsilon_{ij}$ implies $\sum_{j=1}^{n} \varepsilon_{ij} = \sum_{j=1}^{n} a_{ij} \dfrac{w_j}{w_i} = \lambda_{max}$, and $\frac{1}{\lambda_{max}(E)} E$, $E \equiv (\varepsilon_{ij}), \varepsilon_{ji} = \varepsilon_{ij}^{-1}$ is a

$\frac{1}{\lambda_{max}(E)} Ee = 1$ (E is row stochastic). Thus, $\lim_{k\to\infty} \left(\frac{1}{\lambda_{max}(E)} E\right)^k = V \equiv \begin{pmatrix} v_1 & v_2 & \cdots & v_n \\ v_1 & v_2 & \cdots & v_n \\ \vdots & \vdots & \ddots & \vdots \\ v_1 & v_2 & \cdots & v_n \end{pmatrix}$ with

$v_j > 0, j = 1,...,n$, and we have

$$\lim_{k \to \infty} \frac{A^k e}{e^T A^k e} = \lim_{k \to \infty} \frac{D_w \left(\frac{1}{\lambda_{\max}(E)} \right)^k E^k D_w^{-1} e}{e^T D_w \left(\frac{1}{\lambda_{\max}(E)} \right)^k E^k D_w^{-1} e}$$

$$= \frac{D_w V D_w^{-1} e}{e^T D_w V D_w^{-1} e} = \frac{w \left(\sum_j v_j / w_j \right)}{\left(\sum_j v_j / w_j \right) \left(\sum_j w_j \right)} = cw, \quad c = \left(\sum_j w_j \right)^{-1}.$$

Why is the eigenvector close to the priority vector? We have

$$\frac{1}{n^2} \sum_{i=1}^{n} (\sum_{j=1}^{n} a_{ij} w_j / w_i) = \frac{\lambda_{\max}}{n}$$

This says that the matrix $W = (\dfrac{w_i}{w_j})$ formed from the principal eigenvector w is as close to A as $\lambda_{\max}$ is close to n or as near consistency is close to consistency. Closeness may also be studied as a nonlinear programming minimization problem: Find the vector $w > 0$ that minimizes λ subject to the constraints $a_{ij} = \dfrac{w_i}{w_j} \varepsilon_{ij}$, $\varepsilon_{ji} = \varepsilon_{ij}^{-1}$.

We show below that the perturbation argument extends to any reciprocal matrix even if it is not near consistent. Briefly, the argument consists in establishing that the space of reciprocal matrices is strongly connected in the sense that given any two reciprocal matrices one matrix can be reached from the other by continuous Hadamard product reciprocal perturbations.

6. Structural properties of Positive Reciprocal Matrices

We make the following observations on the structure of reciprocal matrices. The elementwise product of two n by n reciprocal matrices is a reciprocal matrix. It follows that the set of reciprocal matrices is closed under the operation Hadamard product. The matrix $e^T e$ is the identity: $e^T e = e^T e o e^T e = e e^T$ and A^T is the inverse of A, $A o A^\top = A^\top o A = e^T e$. Thus the set G of n by n reciprocal matrices is an abelian group. Because every subgroup of an abelian group is normal, in particular the set of n by n consistent matrices is a normal subgroup ($E o W o E^T = W$) of the group of positive reciprocal matrices.

Two matrices A and B are R-equivalent ($A \, R \, B$) if, and only if, there are a vector w and positive constants a and b such that $(1/a) \, Aw = (1/b) \, Bw$. The set of all consistent matrices can be partitioned into disjoint equivalence classes. Given a consistent matrix W and a perturbation matrix E such that $Ee = ae$, $a > 0$ a constant, we use the Hadamard product to define $A = WoE$ such that $(1/a) \, A \, w = (1/n) \, Ww$. A and W are R-equivalent. There is a 1-1 correspondence between the set of all consistent matrices and the set of all matrices A defined by such Hadamard products. An R-equivalence class $Q(W)$ is the set of all A such that $A \, R \, W$. The set of equivalence classes $Q(W)$ forms a partition of the set of reciprocal matrices. It is known that all the elements in $Q(W)$ are connected by perturbations $E, EN, EO,...$, corresponding to a fixed value of $a > 0$ such that $(EoENoEO...)e = ae$. Thus given an arbitrary reciprocal matrix A, there exists an equivalence class to which A belongs.

DeTurck [1] has proved that: The structure group G of the set of positive reciprocal $n \, x \, n$ matrices has $2n!$ connected components. It consists of nonnegative matrices that have exactly one nonzero entry in each row and column. These matrices can be written as $D \, S$, where D is a diagonal matrix with positive diagonal entries and S is a permutation matrix, and the negatives of such matrices. The connected component G_0 of

the identity consists of diagonal matrices with positive entries on the diagonal. If A is a positive reciprocal matrix with principal right eigenvector $w = (w_1, w_2,...,w_n)^T$ and D, G_0 is a diagonal matrix with positive diagonal entries $d_1, d_2,...d_n$ then $I_D(A) = DAD^{-1}$ is a positive reciprocal matrix with principal eigenvector $w = (d_1 w_1,..., d_n w_n)^T$. The principal eigenvalue is the same for both matrices. If $v = (v_1, v_2,...,v_n)^T$ and $w = (w_1, w_2,...,w_n)^T$ are two positive column vectors, then conjugation by the diagonal matrix D_{vw} with entries $v_1/w_1,...,v_n/w_n$ on the diagonal maps A_w onto A_v. The corresponding diagonal matrix D_{wv} provides the inverse map. Moreover, D_{wv} maps the consistent matrix of A_w to the consistent matrix of A_v.

7. The General Case: How to Transform a Positive Reciprocal Matrix to a Near Consistent Matrix

To improve the validity of the priority vector, we must transform the matrix to a near consistent one. In practice, the judgments available to make the comparisons may not be sufficient to bring the matrix to near consistency. In that case we abandon making a decision based on the information we have, and must seek additional knowledge to modify the judgments.

The final question then is how to construct the γ perturbations in a general reciprocal matrix. Inconsistency in a matrix can be due to a single entry (entering the reciprocal value by mistake) which if appropriately changed the matrix becomes near consistent. A second observation is that a judgment matrix already has some built in consistency; it is not an arbitrary reciprocal matrix. It should be meaningful to try and improve its consistency. Because the eigenvector is necessary for representing dominance, we must use an algorithm based on the eigenvector, whose existence is assured by Perron's theory for positive matrices, to improve the consistency of a matrix until it is near consistent.

A metric space X is well chained if, for any two points we have a sequence of points of distance $d(x_i, x_j)$ < γ joining them with γ > 0 arbitrarily small, uniformly chosen for all i and j. This is the property we need in order to bring an arbitrary positive reciprocal matrix close to a consistent matrix that then has its eigenvalue and eigenvector as perturbations of those of the consistent matrix to which it is close. We have the following theorem from topology:

Theorem 5: If X is connected, X is well chained.

Basically we have an existence theorem that ensures that we can transform an arbitrary positive reciprocal matrix to a near consistent matrix. Now we show how to make this transformation.

For a given positive reciprocal matrix $A = [a_{ij}]$ and a given pair of distinct indices $k > l$, define $A(t) = [a_{ij}(t)]$ by $a_{kl}(t) = a_{kl} + t$, $a_{lk}(t) = (a_{lk} + t)^{-1}$, and $a_{ij}(t) = a_{ij}$ for all $i > k, j > l$, so $A(0) = A$. Let $\lambda_{max}(t)$ denote the Perron eigenvalue of $A(t)$ for all t in a neighborhood of $t = 0$ that is small enough to ensure that all entries of the reciprocal matrix $A(t)$ are positive there. Finally, let $v = [v_i]$ be the unique positive eigenvector of the positive matrix A^T that is normalized so that $v^T w = 1$. Then a classical perturbation formula [4, theorem 6.3.12] tells us that

$$\left. \frac{d\lambda_{max}(t)}{dt} \right|_{t=0} = \frac{v^T A'(0)w}{v^T w} = v^T A'(0)w = v_k w_l - \frac{1}{a_{kl}^2} v_l w_k.$$

We conclude that

$$\frac{\partial \lambda_{max}}{\partial a_{ij}} = v_i w_j - a_{ji}^2 v_j w_i \text{ for all } i,j = 1,...,n.$$

Because we are operating within the set of positive reciprocal matrices, $\frac{\partial \lambda_{max}}{\partial a_{ji}} = -\frac{\partial \lambda_{max}}{\partial a_{ij}}$ for all i and j.

Thus, to identify an entry of A whose adjustment within the class of reciprocal matrices would result in the largest rate of change in $\lambda_{\max}$ we should examine the n(n-1)/2 values $\{v_i w_j - a_{ji}^2 v_j w_i\}, i > j$ and select (any) one of largest absolute value. For the pairwise comparison matrix in Table 1, $v = (.042, .114, .063, .368, .194, .168, .030, .021)^T$.

Table 1 A Family's House Buying Pairwise Comparison Matrix for the Criteria

	Size	Trans	Nbrhd	Age	Yard	Modern	Cond	Finance	Normalized Priority Vector w
Size	1	5	3	7	6	6	1/3	1/4	.175
Trans.	1/5	1	1/3	5	3	3	1/5	1/7	.062
Nbrhd.	1/3	3	1	6	3	4	1/2	1/5	.103
Age	1/7	1/5	1/6	1	1/3	1/4	1/7	1/8	.019
Yard	1/6	1/3	1/3	3	1	1/2	1/5	1/6	.034
Modern	1/6	1/3	1/4	4	2	1	1/5	1/6	.041
Cond.	3	5	2	7	5	5	1	1/2	.221
Finance	4	7	5	8	6	6	2	1	.345

$\lambda_{\max} = 8.811$,   Consistency Ratio (C.R.) =.083

Table 2 gives the array of partial derivatives (3) for the matrix of criteria in Table 1 with $a_{37} = 6$ which was its original value later improved to ½ as it is now.

Table 2 Partial Derivatives for the House Example

	Size	Trans.	Nbrhd.	Age	Yard	Modern	Cond.	Finance
Size	-	0.001721	0.007814	-0.00041	0.00054	0.000906	-0.08415	-0.03911
Trans.	-	-	-0.00331	0.001291	0.002485	0.003249	-0.06021	-0.01336
Nbrhd.	-	-	-	-0.00091	-0.00236	-5.7E-05	0.008376	-0.07561
Age	-	-	-	-	-0.01913	-0.03372	0.007638	**0.094293**
Yard	-	-	-	-	-	-0.01366	-0.01409	0.041309
Modern	-	-	-	-	-	-	-0.02599	0.029355
Cond.	-	-	-	-	-	-	-	0.006487
Finance	-	-	-	-	-	-	-	-

The (4,8) entry in Table 2 (in bold print) is largest in absolute value. Thus, the family could be asked to reconsider their judgment (4,8) of Age vs. Finance. One can then repeat this process with the goal of bringing the C.R. within the desired range. If the indicated judgments cannot be changed fully according to one's understanding, they can be changed partially. Failing the attainment of a consistency level with justifiable judgments, one needs to learn more before proceeding with the decision.

Three other methods, presented here in order of increasing observed efficiency in practice, are conceptually different. They are based on the fact that

$$n\lambda_{\max} - n = \sum_{\substack{i,j=1 \\ i \neq j}}^{n} (\varepsilon_{ij} + \varepsilon_{ij}^{-1}).$$

This suggests that we examine the judgment for which γ_{ij} is farthest from one, that is, an entry a_{ij} for which $a_{ij} w_j / w_i$ is the largest, and see if this entry can reasonably be made smaller. We hope that such a change of a_{ij} also results in a new comparison matrix with that has a smaller Perron eigenvalue. To demonstrate how improving judgments works, let us return the house example matrix in Table 1. The family first gave a judgment of 6 for the a_{37} entry. This caused the matrix to have C.R. = .17, which is high. To identify an entry ripe for consideration, construct the matrix γ_{ij} (Table 3). The largest value in Table 3 is 5.32156, which focuses attention on $a_{37} = 6$.

Table 3: $\gamma_{ij} = a_{ij} \ w_j/w_i$

1.00000	1.55965	3.26120	0.70829	1.07648	1.25947	0.32138	0.48143
0.64117	1.00000	1.16165	1.62191	1.72551	2.01882	0.61818	0.88194
0.30664	0.86084	1.00000	0.55848	0.49513	0.77239	**5.32156**	0.35430
1.41185	0.61656	1.79056	1.00000	0.59104	0.51863	1.36123	2.37899
0.92895	0.57954	2.01967	1.69193	1.00000	0.58499	1.07478	1.78893
0.79399	0.49534	1.29467	1.92815	1.70942	1.00000	0.91862	1.52901
3.11156	1.61765	2.25498	0.73463	0.93042	1.08858	1.00000	0.99868
2.07712	1.13386	2.82246	0.42035	0.55899	0.65402	1.00133	1.00000

How does one determine the most consistent entry for the (3,7) position? When we compute the new eigenvector w after changing the (3,7) entry, we want the new (3,7) entry to be w_3 / w_7 and the new (7,3) to be w_7 / w_3. On replacing a_{37} by w_3 / w_7 and a_{73} by w_7 / w_3 and multiplying by the vector w one obtains the same product as one would by replacing a_{37} and a_{73} by zeros and the two corresponding diagonal entries by two. We take the Perron vector of the latter matrix to be our w and use the now-known values of w_3 / w_7 and w_7 / w_3 to replace a_{37} and a_{73} in the original matrix. The family is now invited to change their judgment towards this new value of a_{37} as much as they can. Here the value was $a_{37} = 1/ 2.2$, approximated by $1/ 2$ from the AHP integer valued scale and we hypothetically changed it to 1/2 to illustrate the procedure. If the family does not wish to change the original value of a_{37}, one considers the second most inconsistent judgment and repeats the process. The procedure just described is used in the AHP software *Expert Choice*.

A refinement of this approach is due to P. Harker. One by one, each reciprocal pair a_{ij} and a_{ji} in the matrix is replaced by zero and the corresponding diagonal entries a_{ii} and a_{jj} are replaced by 2, the Perron eigenvalue λ_{max} is computed. The entry with the largest resulting λ_{max} is identified for change as described above. This method, unpublished, is in use in the Analytic Network Process (ANP) software program Super Decisions [12].

References

1. DeTurck, D.M., "The Approach to Inconsistency in the Analytic Hierarchy Process," Mathematical Modeling, 9/3-5, pp. 345-352, 1987.

2. Doob, J.L*., Stochastic Processes*, John Wiley and Sons, Inc., New York, 1953.

3. Franklin, J.L., *Matrix Theory*, Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1968.

4. Horn, R.A. and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.

5. Lancaster, P. and M. Tismenetsky, *The Theory of Matrice*s, Second Edition, Academic Press, New York, 1985.

6. Lax,P., *Linear Algebra*, John Wiley and Sons, Inc., 1997.

7. Luce, R. D. and H. Raiffa, *Games and Decisions*, Dover Publications, Inc., New York, original copyright John Wiley and Sons, Inc., 1957; copyright renewed by Luce and Raiffa, 1985.

8. Millet, I. and T. Saaty, "On the relativity of relative measures – accommodating both rank preservation and rank reversals in the AHP," *European Journal of Operational Research*, 121 (2000), pp. 205-212.

9. Nachbin, L., *Topology and Order*, D. Van Nostrand Company, Inc., 1965. See also Encyclopedic Dictionary of Mathematics, MIT Press, 1987, p 1607.

10. Saaty, T., and L. Vargas, "Inconsistency and Rank Preservation," " *Journal of Mathematical Psychology*, Vol. 28, No. 2, June 1984.

11. Saaty,T., and G. Hu, "Ranking by the Eigenvector Versus Other Methods in the Analytic Hierarchy Process," *Appl. Math. Lett*. Vol. 11, No.4, pp. 121-125, 1998.

12. Saaty, T. L., *Decision Making with Dependence and Feedback: The Analytic Network Proc*ess, RWS Publications, 4922 Ellsworth Ave. Pittsburgh, PA. 15213, 2001.