ESTIMATION OF PRIORITY WEIGHTS BASED ON A RESAMPLING TECHNIQUE AND A RANKING MRTHOD IN ANALYTIC HIERARCHY PROCESS

Indrani Basak

Penn State Altoona

Altoona, PA

ABSTRACT

Individual judgments are sought in order to elicit values of the entries of pairwise comparison matrices in Analytic Hierarchy Process (AHP). Some of these matrices are more consistent than others. But throwing out inconsistent matrices reduces the number of matrices. In this article, we propose a resampling technique to generate sets of pairwise comparison matrices which pass the consistency check. The advantage of the resampling technique is that one can generate as many sets of pairwise comparison matrices as needed to select the ones which satisfy the consistency requirement. Based on these selected matrices, the priority weights of the alternatives are then estimated. We propose rank-based statistical procedures to check the significance in the difference between estimated priorities of the alternatives to establish their most significant rank order.

Key Words: Bootstrap, Priority Weights, Kruskal-Wallis Statistic, Rank-Based Hypothesis Testing, Resampling Method.

1   Introduction

The Analytic Hierarchy Process (AHP) developed by Saaty (1977, 1980) is a popular and practical tool for dealing with complex decision problems. Often decision has to be made on the choice of one or few of the several alternatives. In other situations, priorities for different alternatives have to be sought. In all these decision situations, different alternatives need to be evaluated with respect to a number of criteria. In AHP, a hierarchic or network structure is used to simplify and represent the decision problem with its elements (criteria and alternatives). At each level of the hierarchy, pairwise comparison values of decision elements are used to arrive at priority scores of the elements under consideration. Expert opinion is often sought for these pairwise comparison values. Let the intensity of preference of alternative T_i over alternative T_j be denoted by a_{ij} for $i < j; i, j = 1, 2, \cdots, t$. These a_{ij} values are the entries of a matrix which is called pairwise comparison matrix. The elicitation of these preferences sometimes gives rise to the inconsistency condition $a_{ij}a_{ju} \neq a_{iu}$. Some pairwise comparison matrices are more inconsistent than others. But throwing out inconsistent matrices reduces the number of pairwise comparison matrices to be used for determining the priorities for different alternatives.

2   Literature Review

Saaty (1980) defines the Consistency Ratio (CR) $CR = \frac{CI}{RCI}$ as the measure of the inconsistency of a pairwise comparison matrix. CI is the Consistency Index of a pairwise comparison matrix and is given by $CI = \frac{\lambda \max - t}{t - 1}$, where $\lambda_{\max}$ is the largest eigenvalue of the pairwise comparison matrix. RCI is the Random Consistency Index and is derived by the average of CI values for randomly generated matrices of the same size. RCI values given by Saaty (2001) are provided in Table 1 for different values of t. In the AHP

pairwise comparison matrices are considered to be of acceptable consistency if CR ¡ 0.1. In some cases, Saaty (2001) requires that CR to be less than 0.05 for t=3 and less than 0.08 for t=4. However, in reality it is difficult to obtain matrices which satisfy these consistency criteria because of limited ability of human thinking to be consistent, particularly for a larger number of alternatives. Several authors have considered this issue of consistency in AHP either for individual or for the group decision making situation. Finan and Hurley (1996) proposed a procedure to construct a rank-order consistent matrix while Zeshui and Cuiping (1999) proposed a consistency improving method in AHP. Stein and Mizzi (2007) introduced a harmonic consistency index. Xu (2000), Lin et. al. (2008), Groselj and Stirn (2012) considered a weighted geometric mean complex judgment matrix for the consistency in the group decision making situation using AHP.

In practice, the issue of inconsistencies of these pairwise comparison matrices exists and needs to be dealt with. One possible way is returning the inconsistent matrices to the experts for reconsidering the entries of the matrix until the matrices with acceptable CR are obtained. But it is not practical because of the large amount of time and cost involved.

3   Hypotheses/Objectives

In this article, we propose a resampling technique called Bootstrap to generate sets of pairwise comparison matrices. Based on adequate number of Bootstrap generated matrices, one can ensure a fixed number of matrices with acceptable CR. These, in turn, generate replications of priority vectors of alternatives. Statistical significance tests of the difference in the estimated priorities of the decision alternatives would be possible by these repeated measurements. We propose rank-based statistical procedures to check the significance in the difference between estimated priorities of the alternatives and also to establish their most significant rank order.

4   Research Design/Methodology

Let a_{ij}^k be the intensity of preference of alternative T_i over alternative T_j given by the k-th expert for $i < j; i, j = 1, 2, \cdots, t$ and $k = 1, 2, \cdots, m$. Let A_k be the pairwise comparison matrix composed of these a_{ij}^k values given by the k-th expert, $k = 1, 2, \cdots, m$. A fixed large number of B sets of m matrices are chosen from original m matrices with replacement (i.e, matrix which is previously chosen can again appear). These B sets of m matrices are called Bootstrap samples. Making B large enough allows one to throw out highly inconsistent matrices. One can set a threshold value of the Consistency Index for the matrices to pass the consistency criteria.

Let n be the number of sets of m matrices out of the generated B sets of Bootstrap samples which satisfy the consistency criteria. For each of these n generated sets of matrices, the priority vector V can then be estimated by any group aggregation estimation method established in AHP. In this article, we consider the *logarithmic least squares technique* in which the priority vector V is estimated by the vector v which minimizes

$$\sum_{k=1}^{m}\sum_{i<j}(lna_{ij}^k - lnv_i + lnv_j)^2 \tag{4.1}$$

for each such generated matrix with acceptable CR. The associated set of normal equations can be written as

$$\sum_{k=1}^{m}\sum_{j=1}^{t}lna_{ij}^k = mt \ lnv_i - m\sum_{j=1}^{t}lnv_j; \quad i = 1, \cdots, t. \tag{4.2}$$

A solution to lnv_i is obtained by solving (4.2) with an additive degree of freedom. There is no unique solution to the system in (4.2). A particular solution is obtained by setting $\sum \ln v_j = 0$. An unnormalised solution to the logarithmic least squares problem in (4.1) can then be explicitly written as

$$v_i = exp\left[\frac{1}{mt}\sum_{k=1}^{m}\sum_{j=1}^{t}\ln a_{ij}^k\right] = (\prod_{k=1}^{m}\prod_{j=1}^{t}a_{ij}^k)^{1/mt}. \tag{4.3}$$

This means that the estimate v_i of the priority of T_i is found by taking the geometric mean of the a_{ij} over all $k = 1, 2, \cdots, m$ and $j = 1, 2, \cdots, t$. If desired, these v_is can be normalized. However, for the purpose of establishing a statistically significant rank order of alternatives the normalization is not necessary. Let the over-all estimate of the priority vector be given by $\bar{\boldsymbol{v}} = (\bar{v}_1, \bar{v}_2, \cdots, \bar{v}_t)$ in which

$$\bar{v}_i = \frac{\sum_{\alpha=1}^{n} v_{i\alpha}}{n}; i = 1, 2, \cdots, t. \tag{4.4}$$

Whether differences between the estimated priorities $\bar{v}_i$ in (4.4) are statistically significant can now be determined. This is very much needed when estimated priorities are very close. In this case, we propose rank-based statistical procedures to check the significance in the difference between estimated priorities of the alternatives and to establish their statistically most significant rank order. These rank-based methods do not require any assumptions on statistical distribution of pairwise comparison entries.

Let $R_{i\alpha}$ denote the overall rank of $v_{i\alpha}; i = 1, \cdots, t, \alpha = 1, \cdots, n$ among all $N = tn$ entries. Use of these overall ranks enable one to test whether the difference in ranks of any two alternatives is statistically significant or not. Evaluate the average ranks of the alternatives as

$$\bar{R}_{i.} = \frac{\sum_{\alpha=1}^{n} R_{i\alpha}}{n}; i = 1, 2, \cdots, t. \tag{4.5}$$

The belief that priorities of the alternatives are same can be formulated as the statistical null hypothesis $H_0 : V_1 = \cdots = V_t$. As the first step, the alternative hypothesis of interest in this situation may be formulated as $H_a^1 : V_1, \cdots, V_t$ not all equal. The test statistic for testing this hypothesis is known as Kruskal-Wallis (1952) statistic and is given as

$$H_1^* = \sum_{i=1}^{t} (1 - \frac{1}{t}) \left[\frac{\bar{R}_{i.} - \frac{N+1}{2}}{\sqrt{(N-n)(N+1)/12n}} \right]^2. \tag{4.6}$$

Under H_0, H_1^* has an asymptotic χ^2 distribution with $(t-1)$ degrees of freedom. When the Kruskal-Wallis statistic rejects H_0 implying that not all $V_1, \cdots, V_t$ are equal, pairwise multiple comparisons may be carried out to locate the source of significance. The following procedure was first suggested by Dunn (1964) for this purpose and is discussed in Hettmansperger (1984) in details. Note that there are $t(t-1)/2$ pairwise comparisons in all. Based on the difference in the average ranks of the alternatives, let us define

$$D_{ij} = \frac{1}{\sqrt{N}} (\bar{R}_{i.} - \bar{R}_{j.}); i, j = 1, 2, \cdots, t. \tag{4.7}$$

The expected value and variance of D_{ij} will be denoted by $E(D_{ij})$ and $Var(D_{ij})$ respectively. Under the null hypothesis H_0 of equal priorities, the $E(D_{ij})$ is zero and $Var(D_{ij})$ is calculated as $(Var(\bar{R}_{i.}) + Var(\bar{R}_{j.}) - 2Cov(\bar{R}_{i.}, \bar{R}_{j.}))/N$ in which Var and Cov stand for the variance and covariance respectively. Simple calculations show that $Var(D_{ij})$ reduces to $(N+1)/6n$ since $Var(\bar{R}_{i.}) = (N-n)(N+1)/12n$ and $Cov(\bar{R}_{i.}, \bar{R}_{j.}) = -(N+1)/12$. $Var(D_{ij})$ has an asymptotic (as n increases) value of $t/6$. For a prescribed overall error rate of α, the pairwise comparison error rate is $\alpha' = 2\alpha/t(t-1)$. V_i and V_j will be declared significantly different at overall error rate α if

$$|D_{ij}| \geq z_{\alpha'/2} \sqrt{\frac{t}{6}}, \tag{4.8}$$

in which D_{ij} is as in (4.7). These pairwise multiple comparisons, however, do not detect any increasing (or decreasing) order of priorities. In some situations, it is more appropriate to tailor a test which establishes statistically most significant order of priorities. Following Hettmansperger (1984), we provide a rank-based testing procedure for this purpose. Suppose we wish to test the null hypothesis H_0 versus the alternative hypothesis of a specific order of (say) $H_a^2 : V_1 \geq \cdots \geq V_t$ with at least one strict inequality. The statistic assessing the degree of agreement between the observed average ranks $\bar{R}_{i.}; i = 1, \cdots, t$ and the ordering i as specified in H_a^2 is given by

$$L = \frac{1}{\sqrt{N}} \sum_{i=1}^{t} (i - \frac{t+1}{2})(\bar{R}_{i.} - \frac{N+1}{2}). \tag{4.9}$$

Let us explain the computation of L with the help of the following example. Suppose $H_a^2 : V_4 \geq V_1 \geq V_2 \geq V_3$. Then $L = \frac{1}{\sqrt{N}} \left[(4 - \frac{5}{2})(\bar{R}_4. - \frac{N+1}{2}) + (3 - \frac{5}{2})(\bar{R}_1. - \frac{N+1}{2}) + (2 - \frac{5}{2})(\bar{R}_2. - \frac{N+1}{2}) + (1 - \frac{5}{2})(\bar{R}_3. - \frac{N+1}{2}) \right]$.
Large values of L support the alternative hypothesis H_a^2. The expected value and variance of L will be denoted by $E(L)$ and $Var(L)$ respectively. Under the null hypothesis H_0, L has an asymptotic normal distribution with $E(L) = 0$ and

$$Var(L) = (t^2 - 1)(nt + 1)/144n \tag{4.10}$$

(see page 191 of Hettmansperger(1984)). Therefore, to test H_0 against H_a^2, the test statistic could be used as

$$H_2^* = \frac{L}{\sqrt{Var(L)}}. \tag{4.11}$$

Reject H_0 against H_a^2 if $H_2^* \geq z_\alpha$ for an approximate α-level test where z_α is the upper α percentile of the standard normal distribution.

For a set of t alternatives, the alternative hypothesis H_a^2 could be formed in t! ways. The test statistic value H_2^* is calculated for each of these t! ways. One possible way of comparing these t! order of alternatives is to compare their p-values (probability that H_2^* statistic exceeds its computed value). The most significant order of the alternatives is the one with smallest p-value. According to the fact that largest calculated value of the test statistic yield the smallest p-value, first step to establish the ordering in the alternatives would be to pick up that particular order for which the calculated value of test statistic is the largest. Against that selected order, reject the null hypothesis of equality of priorities among alternatives if $H_2^* \geq z_\alpha$.

We illustrate procedures proposed in this article providing a numerical example in the next section.

5   Data/Model Analysis

Based on the Bootstrap resampling procedure, $n = 1000$ replications of pairwise comparison matrices of acceptable consistency were generated. These generated matrices have maximum inconsistency percentage of 10%. For each generated matrix, the priority vector $\boldsymbol{V}$ is estimated by the logarithmic least squares regression technique as discussed in section 2 yielding estimate vectors $\boldsymbol{v}_\alpha; \alpha = 1, \cdots, 1000$ as well as the overall estimate vector $\bar{\boldsymbol{v}}$ to be $(0.284, 0.206, 0.221, 0.289)^T$. We now apply the methodologies suggested in section 4.

First the null hypothesis of equal priorities $H_0 : V_1 = V_2 = V_3 = V_4$ is tested against the alternative hypothesis H_a^1 that they are not all equal. Computations using (4.5) yield $\bar{R}_1. = 2200.4, \bar{R}_2. = 1821.7, \bar{R}_3. = 1789.4$ and $\bar{R}_4. = 2190.5$ Here $t = 4, n = 1000, N = 4000$. The computed value of test statistic H_1^* as specified in (4.6) is 114.42. The critical value of this statistic at α as small as .01 is 11.34. Clearly the null hypothesis that all alternatives have same priorities is rejected at 1% level of significance.

Now we employ the pairwise multiple comparisons to detect which alternatives differ in their priorities. As discussed in section 4, the criterion that V_i and V_j will be declared significantly different is given by (4.8). If we take an overall error rate of $\alpha = .01$ then $\alpha' = .002, z_{\alpha'/2} = 3.08$ and we would use $|\bar{R}_i. - \bar{R}_j.| \geq 50.29$ to be the criterion to say that V_i and V_j are significantly different. Use of this criterion yields that alternatives T_1 and T_4 do not have significant difference in their priorities, $\bar{R}_1. - \bar{R}_4.$ being 9.9. Similarly alternatives T_2 and T_3 have non-significant difference because $\bar{R}_2. - \bar{R}_3.$ equals 32.3. All other alternatives have pairwise significant difference in priorities. This indicates that in the estimated priority vector $\bar{\boldsymbol{v}} = (0.284, 0.206, 0.221, 0.289)^T$, the difference between 0.289 and 0.284 as well as the difference between 0.221 and 0.206 are not statistically significant while all other differences are.

Finally, we wish to establish the most statistically significant order of priorities of the alternatives. Since the priority difference between alternatives T_1 and T_4 as well as the priority difference between alternatives T_2 and T_3 are not statistically significant, we compute the value of test statistic H_2^* as given in (4.11) for four probable order of preference $V_4 \geq V_1 \geq V_2 \geq V_3, V_4 \geq V_1 \geq V_3 \geq V_2, V_1 \geq V_4 \geq V_2 \geq V_3$ and $V_1 \geq V_4 \geq V_3 \geq V_2$. $Var(L)$ as appears in expression (4.10) is calculated to be 0.4168. The statistic L in (4.9) measuring the degree of assessment between the observed average ranks and the ranking of the alternatives specified by the afore-mentioned four probable order are 12.51, 11.99, 12.66 and 12.15

respectively. These, in turn, yield the test statistic value of H_2^* to be 19.38, 18.58, 19.61 and 18.82 respectively. 19.61 being largest, the most significant order of preference is $V_1 \geq V_4 \geq V_2 \geq V_3$ with at least one strict inequality. Since 19.61 is larger than the critical value z_α of 2.33 at 1% level of significance, the null hypothesis H_0 that all alternatives have same priority is rejected against the most significant order of preference of $H_a^2 : V_1 \geq V_4 \geq V_2 \geq V_3$ at this significance level. Finally, the over-all estimate vector is modified to be $(0.2865, 0.2135, 0.2135, 0.2865)^T$ in which $v_1 = v_4 > v_2 = v_3$. This relation is in compliance with the most significant order of preference established above.

6 Limitations

The methodology developed in the article involves a computationally involved technology of resampling method. In order to get a pre-fixed number of sets of pairwise comparison matrices with acceptable CR, one might need to run the computer program little longer in some situations. For that purpose, it is better to start the resampling method with initial m matrices which are not highly inconsistent. Some Measures (like returning the matrices to the experts asking them to reconsider the entrees of the matrix to improve consistency somewhat) can be used for improving the consistencies of the initial matrices.

7 Conclusions

In this article, we proposed a statistical methodology in AHP in which the entries of the pairwise comparison matrices are considered to be random variables. A statistical resampling method is applied to generate pairwise comparison matrices. This resampling method yields desired large number of pairwise comparison matrices whose consistencies can be controlled. Finally, rank-based statistical methodologies are applied in order to test whether differences in the estimated priorities are statistically significant and to establish the most statistically significant priority order of the alternatives.

8 Key References

Bradley, E. (1982) The Jackknife, the Bootstrap and Other Resampling Plans. Society for Industrial and Applied Mathematics. Philadelphia, Peensylvania.

Finan, J. S. and Hurley, W. J. (1996) A note on a method ensuring rank-order consistency in the analytic hierarchy process, International Transactions of Operational Research 3, 99-103.

Groselj, P. and Stirn, L. Z. (2012) Acceptable consistency of aggregated comparison matrices in analytic hierarchy process, European Journal of Operational Research 223, 417- 420.

Hettmansperger, T.P. (1991). *Statistical Inference Based on Ranks.* Krieger Publishing Company, Florida.

Lin, R., Lin J. S-J, Chang, J., Tang, D. and Julian, P. C. (2008) Note on group consistency in analytic hierarchy process, European Journal of Operational Research 190, 672-678.

Murphy, C. K. (1993) Limits on the analytic hierarchy process from its consistency index, European Journal of Operational Research 65, 138-139.

Stein, W. E. and Mizzi, P. J. (2007) The harmonic consistency index for the analytic hierarchy process, European Journal of Operational Research 177, 488-497.

Xu, Z. (2000) On consistency of the weighted geometric mean complex judgment matrix in AHP, European Journal of Operational Research 126, 683-687.

Zeshui, X. and Cuiping, W. (1999) A consistency improving method in the analytic hierarchy process, European Journal of Operational Research 116, 443-449.