

Utility Theory Violations by Multi-Criteria Hierarchical Weighting Methods

Jay R. Lund

Associate Professor

Department of Civil and Environmental Engineering

University of California

Davis, CA 95616

Abstract

This paper discusses two violations of utility theory that can arise with some common multi-criteria hierarchical weighting methods when employed for selection of the best of several alternatives considering multiple criteria. First, these hierarchical weighting methods are found to violate utility theory's principle of independence of irrelevant alternatives. Proof of this violation is made by counter-example. Second, a common violation of additive utility exists even when only two alternatives are considered, a case without irrelevant alternatives. The frequency of this incompatibility of hierarchical weighting with additive utility is explored through a numerical experiment and found to range between 4% and 13% for different problem sizes. These problems do not exist for application of weighting methods to single-objective problems.

INTRODUCTION

Over the last few decades, a number of multi-criteria weighting methods have become popular for evaluating alternative decisions. While these procedures are typically easy to implement, their compliance with fairly fundamental principles of rational decision-making, as represented by utility theory, has received little examination until recently. It seems appropriate to attempt to find the general forms of utility functions which would be implied by particular multi-criteria weighting methods.

It is often presumed that multi-criteria weighting methods conform to the conditions of utility theory, particularly additive utility (e.g., Palmer and Lund, 1985). While this has been shown to be false in earlier literature (Hobbs, 1980; deNeufville, 1990; Dyer 1990), this finding has not greatly deterred the use of multi-criteria weighting methods.

This short paper demonstrates the general incompatibility of hierarchical weighting methods with utility theory when used for the selection of a single best alternative. This incompatibility arises from violation of utility theory's principle of the independence of irrelevant alternatives. This is shown by counterexample.

A further incompatibility of hierarchical weighting methods with additive utility theory is also shown when both methods are used for selection of a single best alternative. This effect is distinct from violation of the independence axiom of utility theory. The results of numerical experiments are provided which show the frequency with which multi-criteria weighting methods lead to selection of a different alternative than would direct use of additive multi-criteria utility functions. Conditions of certainty are assumed throughout the paper.

HIERARCHICAL NORMALIZED WEIGHTING

Hierarchical normalized weighting schemes for decision-making determine an overall weight by summing the performance weights of each alternative weighted by an "objective weight" for each criterion,

$$(1) \quad \omega_i = \sum_{j=1}^m \beta_j \alpha_{ij}$$

where β_j is the weight of criterion j and α_{ij} is the performance weight of alternative i on criterion j . Final weights are often normalized, so that the sum of weights over the alternatives equals one,

$$(2) \quad \omega'_i = \frac{\omega_i}{\sum_{i=1}^n \omega_i} .$$

Typically, the α_{ij} are calculated as the ratio of the performance of alternative i on criterion j divided by the summed performance of all alternatives on criterion j . These ratios are normalized, so that the sum of α_{ij} for each j is one,

$$(3) \quad \alpha_{ij} = \frac{P_{ij}}{\sum_{i=1}^n P_{ij}} ,$$

where P_{ij} is the performance of alternative i on criterion j . Note that the α_{ij} are dimensionless. Sometimes the P_{ij} are not quantitative measures of alternative performance, but rather represent the utility of an alternative's performance on a single criterion, $U(P_{ij})$.

The objective weights β_j are commonly solicited from the decision-maker(s) and are considered the proportion of overall importance to be placed on the j -th criterion. These objective weights are also normalized to sum to one,

$$(4) \quad \beta_j = \frac{I_j}{\sum_{j=1}^m I_j} ,$$

where I_j is the subjective "importance" of objective j relative to all other objectives.

This general computational approach is used by a number of common multi-criteria weighting methods (Saaty, 1977; Hobbs, 1980).

VIOLATION OF UTILITY THEORY'S INDEPENDENCE AXIOM

Von Neumann and Morgenstern's utility theory is derived from three axioms: order, independence, and continuity (Von Neumann and Morgenstern, 1944). The independence axiom can be stated: If an alternative is non-optimal for a decision problem, it cannot be made optimal by adding new alternatives to the problem (Luce and Raiffa, 1957). This axiom is sometimes called the principle of the independence of irrelevant alternatives.

The violation of the independence axiom by hierarchical weighting methods applied to selection of a single best alternative is demonstrated by counter-example. Consider the following applications of a hierarchical weighting method for a bi-criteria problem. The objective weights have values $\beta_1 = 0.25$ and $\beta_2 = 0.75$.

First, consider the case of three alternatives with performances, P_{ij} , given by:

	j=1	j=2
i=1	1.0	5.0
i=2	2.0	4.0
i=3	2.0	2.0

Note that the first two alternatives are members of a non-dominated solution set, with the third alternative being clearly inferior (Cohon, 1978). For this problem, the final decision weights given by applying Equation 1 appear in the Case A column of Table 1. For this case, alternative 1 is the optimal choice.

Next, consider the same problem, except that the performance of the third (inferior) alternative has been changed, but the alternative remains inferior. The performances, P_{ij} , are now:

	j=1	j=2
i=1	1.0	5.0
i=2	2.0	4.0
i=3	0.0	4.0

The final hierarchical decision weights for this case are given in the Case B column of Table 1. Alternative 2 now becomes the optimal choice.

With the modification of a third irrelevant alternative, the choice of the optimal decision is changed. The hierarchical method therefore can violate the principle of the independence of irrelevant alternatives. This behavior should caution the use of hierarchical weighting methods for resolving multi-objective problems.

Table 1: Final Hierarchical Weights (ω'_i) for the First Counterexample Problem
(* marks optimal choice)

Alternative	Case A	Case B
i=1	0.39*	0.37
i=2	0.37	0.40*
i=3	0.24	0.23

ADDITIVE UTILITY

The general normative approach to evaluating alternatives is through the use of a utility function which evaluates the utility of the performance of each alternative on each objective. This general utility function has the form: $U_i(P_{i1}, \dots, P_{ij}, P_{im})$, where P_{ij} represents the performance of alternative i on objective j . Keeney and Raiffa (1976) discuss a special case of this utility function, the *additive utility function*,

$$(5) \quad U_i = \sum_{j=1}^m c_j U(P_{ij}),$$

where c_j is a constant. For cases where the choice of an alternative is not of great consequence, as with the selection of a design for a small water resource project being evaluated from a national perspective, the utility functions can be treated as linear, in a way similar to the approach of Arrow and Lind (1970). This results in the *additive linear utility function*:

$$(6) \quad U_i = \sum_{j=1}^m c_j P_{ij} .$$

Using a Taylor series expansion, it can be shown that this additive linear utility model is the first-order approximation of the general utility function in the vicinity of the utility of the present condition (Appendix II). In this case all P_{ij} are measured relative to performance on each objective under current conditions and the constants c_j represent $\partial U_i / \partial P_{ij}$ evaluated under present pre-decision conditions. Such utility functions should therefore be commonplace for many forms of decision making within large public and private bureaucracies. Schoemaker and Waid (1982) discuss establishing decision weights for additive utility models.

DISAGREEMENTS WITH ADDITIVE UTILITY

There are two major areas of application for the weighting methods described above. The first is for selecting a single alternative from several candidate alternatives. Examples of this type of application are the selection of a preferred location from several proposed locations, selection of an employee from a pool of applicants, or selection of a specific computer system from several proposals.

When additive utility and hierarchical weighting methods are applied to problems where selection of a single alternative is desired, the two methods frequently will disagree. This disagreement is shown by counter-example below and the frequency of this effect demonstrated through numerical experiments. The problem arises from the standardization of alternative performance weights within each objective independent of the standardization of the relative weights of objectives. In human decision-making experiments, additive utility theory and Saaty's hierarchical weighting method, the Analytic Hierarchy Process (AHP), have been shown to yield frequently different final alternative scores (Belton, 1986). Here, the focus of the numerical experiments will be on the frequency of disagreement in alternative selection outcome between hierarchical weighting and additive linear utility approaches.

The second application of these weighting methods is for resource allocation. Here, a limited resource budget must be allocated among a number of activities. The design of an investment portfolio, the allocation of a budget among several departments, or the distribution of aid among numerous potential recipients are some examples.

For allocation-type applications, additive linear utility and hierarchical weighting solutions will almost always disagree. Unconstrained maximization of additive linear utility will always assign the entire resource budget to a single "best" alternative. Weighting methods will typically distribute the limited resource among activities in proportion to their final decision weights.

Where maximization of simple additive utility is seen as appropriate, the utility and weighting solutions will still disagree. The solution to the problem of maximizing additive (but non-linear) utility becomes a non-linear programming problem whose solution can vary with the level of the resource budget. This contrasts with the weighting solutions which result from a simple computation and whose allocative proportions will not vary with the level of the resource budget.

PROOF BY COUNTER-EXAMPLE

The potential disagreement between linear multi-criteria utility models and multi-criteria weighting methods is illustrated by the following example. Consider a bi-criteria decision-making problem with three alternatives. The additive linear utility function, Equation 6, is assumed to measure the multi-criteria utility of each alternative. Let us further suppose that $c_1 = c_2 = 0.5$ for the example.

This implies that for the hierarchical weighting method, the ratios of overall importance for each criterion are also 0.5, $\beta_1 = \beta_2 = 0.5$, using Equation 4.

The performance of each alternative on each criterion and the consequent ratio-weights for each objective are given in Table 2. The ratio-weights α_{ij} are found for each criterion from the performance values by Equation 3.

Table 2: Performance Data for Counter-Example

Alternative	P_{i1}	P_{i2}	α_{i1}	α_{i2}
A	1.0	5.0	0.33	0.63
B	2.0	3.0	0.67	0.37

Most hierarchical weighting methods would develop final decision weights for each alternative using Equations 1 and 2.

The comparable decision weights derived from the additive linear utility function are found using Equation 6 and the normalizing equation,

$$U'_i = \frac{U_i}{\sum_{i=1}^2 U_i} .$$

This expression gives final utility-based decision weights which sum to one.

For this example, pairwise comparison and utility-derived intermediate and final weights appear in Table 3. The starred alternative is the best, as indicated by each of the four sets of decision weights.

Here, the pairwise-comparison-based decision weights are clearly different from those arising from direct use of linear utility functions. Indeed, for this example, the two approaches choose different solutions to the problem.

Note that there would still be a disagreement between the utility model and the weighting model if the P_{ij} were re-defined as the value or utility of performance on the j -th criterion by the i -th alternative. This implies that this class of weighting methods can yield incorrect decisions, not only for additive linear utility functions, but for the larger class of additive utility functions.

This particular disagreement with additive utility does not require a third, irrelevant alternative, and so is a different problem than the violation of the independence axiom presented above.

Table 3: Final Raw and Normalized Ratio and Additive Linear Utility Decision Weights

Alternative	Raw Utility U_i	Raw Weight ω_i	Normalized Weight ω'_i	Normalized Utility U'_i
A	3.0*	0.48	0.48	0.55*
B	2.5	0.52*	0.52*	0.45
Total:	5.5	1.0	1.0	1.0

FREQUENCY OF DISAGREEMENT

To estimate the frequency with which additive linear utility methods disagree with multi-criteria weighting methods, a series of randomized numerical experiments were undertaken. These experiments were based on general form of the example above.

The frequency of disagreements between the two methods was found for several different combinations of numbers of alternatives, numbers of criteria, and ranges over which alternative performances (P_{ij}) and utility values of a unit of performance (c_j) could vary.

For each combination, 1,000 sets of performance values (P_{ij}) and utility values function coefficients (c_j) were generated randomly. For each set, values of P_{ij} and c_j were drawn from a uniform frequency distribution extending from zero to a specified common maximum.

Utility and decision weight values were calculated for each set, and the best alternative was found by each method. The number of disagreements between these two methods was accumulated. These results are presented in Table 4.

The range over which performance and utility coefficient values were allowed to vary had no effect on the frequency of disagreement, and is therefore not included in the table. Ranges of 0-50, 0-100, and 0-1,000 all resulted in exactly the same frequency of failure, and the same sets failing to agree. This lack of effect stems from the range maximum acting as a constant in each method which does not affect the ordering of alternatives.

The rate of disagreement was always greater than 4% and tended to be on the order of 10%. Disagreement rates tended to be greater with more criteria, and were also usually greater for cases with fewer alternatives. While this rate of disagreement is unsettlingly high, Table 4 illustrates that the error rate of hierarchical weighting methods is substantially lower than the error rate for random selection of alternatives.

This measure of disagreement, where the two methods failed to agree on which alternative was the best, seems quite appropriate for cases where the final decision weights are employed to select a single best alternative.

Table 4: Frequency of Disagreement Over Best Alternative Between Additive Linear Utility and Hierarchical Weighting Decision Models (1,000 random experiments for each cell)

	All Parameter Values Vary Randomly			Guessing Disagreement
	2 Criteria	5 Criteria	10 Criteria	Frequency
2 Alternatives	5.9%	9.7%	11.1%	50%
5 Alternatives	5.8%	12.3%	13.3%	80%
10 Alternatives	6.8%	10.6%	12.7%	90%
20 Alternatives	4.3%	8.8%	9.9%	95%

Weighting methods also commonly are employed to allocate resources, typically in proportion to their final decision weight, ω_i . Some such uses of multi-criteria weighting methods suggested in the literature include producing a diversified portfolio, allocating a limited budget to a number of alternative activities, and other allocation problems for scarce resources (Saaty and Mariano, 1979; Cook et al, 1984). This use of this type of weighting method would also violate linear utility theory. The utility maximizing allocation of resources among uses or investments where the value of each use was judged by multi-linear utility theory under certainty should always result in all the resources being devoted to the use with the greatest unit multi-utility.

MODIFICATIONS TO IMPROVE WEIGHTING METHODS

Schoner et al (1993) suggest improvements to Saaty's AHP to avoid violating utility theory's independence axiom. These approaches all involve soliciting information from decision makers comparing the performance of one alternative on one objective with the performance of another alternative on another objective. These alternative comparisons across objectives replace direct comparisons of objectives (implied in Equation 4). Thus, for the case where two alternatives are being compared with two objectives, the decision maker might be asked, "How much better is Alternative 1's performance on objective 2 than Alternative 2's performance on objective 1?" This approach can be used in a way compatible with additive linear utility (Equation 6). Let the answer to the preceding question be $a = U(P_{12})/U(P_{21})$. Assuming additive utility, $a = (c_2P_{12})/(c_1P_{21})$, or, $c_2/c_1 = a P_{21}/P_{12}$.

A conceptually similar approach would establish the alternative weights on each criterion based on the performance of an alternative relative to the best or worst alternative on the criterion:

$$\alpha'_{ij} = \frac{P_{ij}}{P_{\min j}},$$

and the importance weights of each criterion based on the ratio of each criterion's minimum (or maximum) performance compared with the minimum (or maximum) performance on a selected criterion,

$$\beta_j = \frac{U(P_{\min j})}{U(P_{\min \text{ std}})} .$$

Both approaches result in weighting schemes compatible with additive linear utility assumptions. But such approaches, eliminating independent objective and alternative performance weighting, are no longer hierarchical weighting schemes. Implementing such unusual cross-criterion questions to real decision-makers may pose practical or behavioral problems, such inconsistent responses.

CONCLUSIONS

The short paper has demonstrated that a common class of hierarchical weighting methods for multi-criteria decision-making does not accommodate the independence axiom of utility theory and is also not consistent with assumption of additive utility, even in the absence of irrelevant alternatives. The choice of alternatives resulting from these weighting methods can differ from the choice made using utility theory. For conditions where additive utility or additive linear utility are likely to apply, such weighting methods frequently suggest a different alternative than would be suggested by direct use of utility functions, as found by numerical experiment. The conditions when additive or additive linear utility are most likely to apply are when the decision is of relatively small consequence relative to the overall utility of the status quo.

Another common application of hierarchical weighting methods is for resource allocation, such as allocating a limited budget to an array of activities. Where additive linear utility applies to these problems (a rare circumstance), an expected utility criterion would always allocate the entire budget to a single activity. This too is at variance with the allocations resulting from hierarchical weighting methods.

The conditions under which hierarchical weighting methods can be rigorously applied to decision-making problems has yet to be demonstrated. Weighting methods can be made compatible with simple utility theory precepts by replacing the hierarchical weighting structure with weightings of performance across objectives.

ACKNOWLEDGMENTS

Morris Israel and Robert P. Anex are thanked for their comments on an earlier version of this paper.

REFERENCES

- Arrow, K and R. Lind (1970), "Uncertainty and the Evaluation of Public Investment Decisions," *American Economic Review*, Vol. 60, June.
- Belton, V. (1986), "A Comparison of the Analytic Hierarchy Process and a Simple Multi-Attribute Value Function," *European Journal of Operational Research*, Vol. 26, No. 1, July, pp. 7-21.
- Cohon, J.L. (1978), *Multiobjective Programming and Planning*, Academic Press, N.Y.
- Cook, T., P. Falcho, and R. Mariano (1984), "An Urban Allocation Model Combining Time Series and Analytical Hierarchical Methods," *Management Science*, Vol. 30, No. 2, pp. 198-208.
- de Neufville, Richard (1990), *Applied Systems Analysis: Engineering Planning and Technology Management*, McGraw-Hill Publishing Co., N.Y.
- Dyer, J.S. (1990), "Remarks on the Analytic Hierarchy Process," *Management Science*, Vol. 36, No. 3, pp. 249-258.
- Hobbs, B.F. (1980), "A Comparison of Weighting Methods in Power Plant Siting," *Decision Sciences*, Vol. 11, No. 4, October, pp. 725-737.
- Keeney, R.L. and H. Raiffa (1976), *Decisions with Multiple Objectives: Preferences and Value Trade-Offs*, Wiley and Sons, N.Y., 569 pp.
- Luce, R.D. and H. Raiffa (1957), *Games and Decisions: Introduction and Critical Survey*, John Wiley & Sons, N.Y.
- Palmer, R.N. and J.R. Lund (1985), "Multiobjective Evaluation with Subjective Information," *Journal of Water Resources Planning and Management*, ASCE, Vol. 111, No. 4, October, pp. 399-416.
- Saaty, T.L. (1977), "A Scaling Method for Priorities in Hierarchical Structures," *Journal of Mathematical Psychology*, Vol. 15, pp. 234-281.
- Saaty, T.L. and R. Mariano (1979), "Rationing Energy to Industries; Priorities and Input-Output Independence," *Energy Systems and Policy* (Winter).
- Schoemaker, P.J.H. and C.C. Waid (1982), "An Experimental Comparison of Different Approaches to Determining Weights for Additive Utility Models," *Management Science*, Vol. 28, No. 2, February, pp. 182-196.
- Schoner, B., W.C. Wedley, and E.U. Choo (1993), "A unified approach to AHP with linking pins," *European Journal of Operations Research*, Vol. 64, pp. 384-392.
- Von Neumann, J. and Morgenstern (1944), *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ.

APPENDIX I: NOTATION

- c_j = the unit increase of utility per increase in performance on criterion j
- P_{ij} = the performance of alternative i on criterion j
- U_i = the multi-criteria utility of alternative i
- U'_i = the relative, normalized multi-criteria utility of alternative i
- α_{ij} = the relative performance weight of alternative i on criterion j (dimensionless)
- β_j = the relative decision weight of criterion j (dimensionless)
- ω_i = the raw decision weight of alternative i (dimensionless)
- ω'_i = the final, normalized decision weight of alternative i (dimensionless)

APPENDIX II: DERIVATION OF LINEAR MULTI-LINEAR UTILITY

Given the general utility function $U(X_1, X_2)$, dependent on performance on two criteria, let X_{10} and X_{20} represent the present, pre-decision values of performance on each criterion. And let $\Delta X_1 = X_1 - X_{10}$ and $\Delta X_2 = X_2 - X_{20}$, the improvement on current performance achieved by some alternative.

Taking the Taylor series expansion of $U(X_1, X_2)$ about the present condition, X_{10}, X_{20} , yields the following approximation.

$$\begin{aligned} U(X_1, X_2) = & U(X_{10}, X_{20}) + \frac{\partial U}{\partial X_1} \Delta X_1 + \frac{\partial U}{\partial X_2} \Delta X_2 \\ & + \frac{1}{2!} \left[\frac{\partial^2 U}{\partial X_1^2} (\Delta X_1)^2 + 2 \frac{\partial^2 U}{\partial X_1 \partial X_2} (\Delta X_1 \Delta X_2) + \frac{\partial^2 U}{\partial X_2^2} (\Delta X_2)^2 \right] \\ & + \dots + \frac{1}{n!} \left[\frac{\partial^n U}{\partial X_1^n} (\Delta X_1)^n + \dots \right] + \dots \end{aligned}$$

The first line of this expression represents the first-order approximation of the general utility function. The first and second lines together represent the second-order approximation of $U(X_1, X_2)$, etc. Since the term $U(X_{10}, X_{20})$ will be a constant for all decision-making alternatives, this term can essentially be neglected, or subtracted from the entire expression.

This leaves the following first-order approximation for the general utility function:

$$U(X_1, X_2) \sim \frac{\partial U}{\partial X_1} \Delta X_1 + \frac{\partial U}{\partial X_2} \Delta X_2.$$

This is the *additive linear utility* model and can be simplified to:

$$U(X_1, X_2) \sim c_1 \Delta X_1 + c_2 \Delta X_2,$$

where c_1 and c_2 are constants. This expression should be adequate for analysis where any alternative decision would only result in a small change in performance, $\Delta X_1 \ll X_{10}$ and $\Delta X_2 \ll X_{20}$. In this case, higher order terms will usually drop out quickly.

If the utility function $U(X_1, X_2)$ is assumed to be linear in X_1 and X_2 , independently, then a slightly more complex approximation of the general utility function can be obtained for alternatives that represent somewhat larger improvements in total performance. The performance of each alternative still is assumed to be small enough that the third-order and higher terms of the expansion can be neglected. With these two assumptions, the second derivatives of each single variable vanish, leaving the second-order approximation:

$$U(X_1, X_2) \approx \frac{\partial U}{\partial X_1} \Delta X_1 + \frac{\partial U}{\partial X_2} \Delta X_2 + \frac{\partial^2 U}{\partial X_1 \partial X_2} (\Delta X_1 \Delta X_2).$$

This is a form of Keeney and Raiffa's (1976) multi-linear utility model which assumes linear utility functions for each objective. Substituting constant parameters for each derivative, this *linear multi-linear utility* model can be simplified to:

$$U(X_1, X_2) \approx c_1 \Delta X_1 + c_2 \Delta X_2 + c_3 (\Delta X_1 \Delta X_2).$$