Newsletter of the European Working Group
"Multicriteria Aid for Decisions"

Bulletin du Groupe de Travail Européen
"Aide Multicritère à la Décision"

Groupe de Travail Européen "Aide Multicritère à la Décision"
Série 3, nº16, automne 2007.

European Working Group "Multiple Criteria Decision Aiding"
Series 3, nº 16, Fall 2007.

Forum

(Robustness Analysis)

Robustness analysis in project prioritisation

Alec Morton
London School of Economics and Political Science

Background

A common approach to prioritising investment opportunities (for example, scientific projects) is to sort them in some sort of value for money ordering, that is to say in terms of a ratio v_i / c_i where v_i is a measure of value of a project, and c_i is a measure of the input cost, with i a project index from $K=\{1,\ldots, k\}$. In this note, I'll suppose for simplicity that v and c are elicited directly, although they could and often would arise from some transformation and combination of more disaggregate judgements. This priority ordering may be used by the organisation for a one-time allocation of some fixed budget as part of a planning process. Alternatively, it may be kept in a manager's desk drawer, in case a sudden cash or capacity crunch forces the organisation to disinvest from marginal projects: having already thought through priorities can mean that the organisation can respond to this sort of challenge in a relatively co-ordinated way.

Costing and valuation are however, difficult tasks: few assessors can assert their complete confidence in the figures they supply. Accordingly, it is helpful to have a sense of the robustness of this priority ordering. Some people, including this author, find it natural to ascribe a probabilistic meaning to robustness (Butler, Jia et al. 1997; Lahdelma, Hokkanen et al. 1998; Jiménez, Mateos et al. 2005; Morton 2007; Tervonen and Figueira forthcoming). The question which one asks if one takes this view is: if assessments are subject to error, how confident can one be in the priority order delivered by the model?

I shall formalise this notion of error in judgement. To start of with, suppose that we have elicited vectors of cardinal valuations $v=(v_i)$ and of costs $c=(c_i)$ for the projects. Suppose further that there are true costs and valuations (in a sense to be discussed subsequently). As our state of knowledge of these is uncertain, think of these as random vectors and write them as $V=(V_i)$ and $C=(C_i)$ respectively. The expected values of V and C may or may not be equal in value to v and c: in the former case, we will say the assessments are unbiased, in the latter that they are biased.

Given v and c, it should be possible to put the k projects in a value for money order: I will write the rank of an individual project i in this ordering as $o_i(v, c)$, thus defining a vector-valued function $o(v, c)$. Similarly, the (random) vector of project ranks according to their true costs and valuations is $o(V, C)$. With modern software it is easy to find a simulated distribution for $o(V, C)$. One possible way of presenting the information back is just to display box plots (Butler, Jia et al. 1997) of $o_i(V, C)$ for each i. However, in an application setting, the number of projects can be quite large and so there may be an interest in having some sort of summary measure. I propose as a measure of the robustness a function of the form:

$$G(o(v, c), \boldsymbol{E}(g(o_1(v, c)- o_1(V, C))), \ldots, \\ \boldsymbol{E}(g(o_k(v, c)- o_k(V, C)))).$$

where $\boldsymbol{E}$ is the expectation operator. For convenience, I'll call the terms $\boldsymbol{E}(g(o_i(v, c)- o_i(V, C)))$ the Expected Transformed Rank Differences (ETRDs). Somewhat similar rank-oriented summary statistics have been proposed by Lahdelma and Salminen (2001) but their interest is primarily in identifying attractive compromise solutions in 1-of-n choice tasks rather than in prioritisation tasks and so their development is somewhat different from that here.

This set-up raises some interesting questions of which I now consider three.

Three questions

1. What distributional forms are appropriate for C?

If the approach outlined here was to be built into software, I envisage that users could paramterise distributions selected from a menu based on the degree of uncertainty that they feel about *v* and *c*. However, the question arises of what should be included in this menu of distributional forms? Answering this question convincingly would require some sort of empirical data gathering exercise. If we were measuring the error properties associated with the measurement of some physical property (mass, volume, etc), what one should do is apply the instrument repeatedly to establish the spread of measurements, and validate against some more accurate instrument to identify whether bias exists. Some authors (e.g. Kleinmuntz 1990) have argued that this sort of reasoning is applicable in the case of modelling error in judgement. However, it is hard to see how one could operationally establish a distribution under this interpretation, as unlike physical instruments, people have memory, with the consequence that successive elicitations can hardly be said to be under the same circumstances.

Accordingly, a more appropriate strategy of investigation may to be identify a number of qualified and reasonably homogeneous cost assessors, and invite them each to assess the costs of the list of projects. The existence or otherwise of systematic bias could be established by investigating the relationship between the assessed judgements and the actual experienced cost of the actual delivered project. The reader will note that the above procedure yields *not* the distribution of true scores given a judgement, but the distribution of judgements given some true score. With suitable supplementary data gathering, however, Bayes Theorem would enable us to deduce the former from the latter.

2. What distributional forms are appropriate for *V*?

Establishing an appropriate distributional form for *V* poses a parallel but more tricky problem to establishing a distributional form for *C*, since judgements of value, by their nature, are not "right" or "wrong" in any absolute sense, and so it is hard to think of them as deviating from some underlying true value.

This doesn't give us any difficulty in establishing a distribution of assessments, which is as easy or difficult as in the case of cost assessments discussed above. However, it does make it hard to say how we should establish "true" values. One option (which Kleinmuntz seems to suggest) is to take the mean of the distribution of assessed values as the true value. This has the consequence that value judgements cannot be biased: they must, definitionally, be on average correct.

An alternative view would be to try to develop a parallel approach to that suggested above for dealing with cost. In this case, one would contrast the *ex ante* judgements of value of projects prior to sign-off with *ex post* judgements of value subsequent to delivery. The analogy with cost is not complete, since organisations have to arrive an agreed definition of what things cost for financial reporting purposes, but not of what things are worth. Thus, this approach would pose some tricky methodological challenges, but could – perhaps – be doable.

3. How should we select function*s* *g* and *G*?

g's purpose is to transform the rank difference between the ranks according to the elicited values and the ranks according to the true values. The simplest option is to take $g(x)=|x|$. This gives twice the weight to a movement of two places in the ranking to one place. However, one can imagine cases where decision makers might feel either more than or less than twice as bad as a consequence, suggesting *g* should take be a convex or concave increasing function of $|x|$ respectively. One could also imagine cases (for example, where judgements are systematically biased in some way) where a decision maker may be interested in knowing which options tend to move up and which tend to move down the ranking. In such cases, it might be useful to make *g* a vector-valued which splits a variable into its positive and negative parts, *i.e.* of the form $g(x)=(max(x,0), min(x,0))$.

In some circumstances, it may be that one can simply take *G* as the identity function, and produce a vector of ETRDs for each project *i*. This could give quite a bit of insight – for example, if the options have a single peaked distribution of values/ cost, one would expect that those in the centre of the distribution would have higher ETRDs for comparable levels of error, since those projects would have more options in their immediate neighbourhood.

Let us suppose however that we are interested in producing a single "headline" statistic which synthesises all the ETRDs. One possibility is the average ETRD,

$$\frac{\sum_{i=1}^{k} E\Big(g\big(o_i(v,c)-o_i(V,C)\big)\Big)}{k}.$$

However, often, when priority orderings are used to support decision making, a triage line of reasoning is relevant: options high in the priority ordering will probably be done, and options low in the priority ordering will probably not be done, and so the really critical ranks are those of the options in the centre of the ordering. In this case, one might be interested in some sort of weighted average ETRD,

$$\frac{\sum_{i=1}^{k} w\left(o_i\left(v,c\right)\right) g\left(o_i\left(v,c\right)-o_i\left(V,C\right)\right)}{\sum_{i=1}^{k} w\left(o_i\left(v,c\right)\right)}$$

where $w\left(o_i\left(v,c\right)\right)$ is some sort of concave function which peaks in the middle of the range.

References

Butler J, Jia J and Dyer J (1997). Simulation techniques for the sensitivity analysis of multi-criteria decision models. European Journal of Operational Research **103**: 531-546.

Jiménez A, Mateos A and Ríos-Insua S (2005). Monte Carlo Simulation Techniques in a Decision Support System for Group Decision Making. Group Decision and Negotiation **14**(2): 109-130.

Kleinmuntz D N (1990). Decomposition and the control of error in decision analytic models. Insights in decision making: a tribute to Hillel J Einhorn. R. M. Hogarth. Chicago, University of Chicago Press**:** 107-126.

Lahdelma R, Hokkanen J and Salminen P (1998). SMAA - stochastic multiobjective acceptability analysis. European Journal of Operational Research **106**: 137-143.

Lahdelma R and Salminen P (2001). SMAA-2: stochastic multicriteria acceptability analysis for group decision making. Operations Research **49**(3): 444-454.

Morton A (2007). Probabilistic Sensitivity Analysis for Project Prioritisation with Multi-Attribute Value Theory London School of Economics Operational Research Group working paper 07.95: London

Tervonen T and Figueira J (forthcoming). A survey on stochastic multicriteria acceptability analysis. Journal of Multi-Criteria Decision Analysis.