The DURENO Techniques

Decision-Analytic Techniques by Prof. Mats Danielson, Dept. of Computer and Systems Sciences, Stockholm University, PO Box 7003, SE-164 07 Kista, Sweden, mats.danielson@su.se

Methods for multi-criteria decision analysis (MCDA) have been promising methods for decision support for quite a long time. However, many first-time users do not return to become regular users of those methods. While there is, of course, not one single reason that explains this fact, one reason stands out when discussing with decision-makers who have elected not to continue using the methods after a trial. It is articulated in many different ways but revolves around the inability to get to grips with and trust the criteria weights. And if those seem inaccurate or opaque, then there is seemingly no reason left to trust the outcome of the analysis.

This lack of trust in criteria weights reflects back to the nature of weights. They are, by their design, inherently relative and dependent on other information (such as value scale ranges), which makes them harder to understand from a layperson's perspective. The main factor of this mistrust is the **weight/scale duality**, which is an inescapable fact since the weights are formally scaling factors, and, as such, highly dependent on the underlying criteria scales. A couple of examples can illuminate this:

Imagine a situation where you are going to buy a flashlight. You consider three models, A, B, and C, from two perspectives (criteria): cost and brightness. Of course, you want it to be as cheap as possible, but since you will quite often use it at long distances, high brightness leading to a long beam distance is also important. Brightness (measured in lumen) drives production costs, so the most expensive flashlight has the longest beam distance. Now, you need to find a trade-off between the two criteria, i.e. assign weights to them so that their sum is 100%. Note, however, that a statement such as "*cost is the most important criterion*" is meaningless. Why? Assume that the models have the following costs: A: $70, B: $65, C: $60, and the following brightnesses: A: 2000 lumen, B: 500 lumen, C: 200 lumen. Since the difference in cost is so small, it is not very important in this case. If instead, we would have had A: $200, B: $100, C: $50, and A: 1000 lumen, B: 900 lumen, C: 800 lumen, then surely cost would have been the decisive criterion. Thus, the weights for the criteria cannot be set as if they were freestanding entities, they must always be set relative to the difference between the best and the worst alternative for each criterion. This illustrates the weight/scale-dualism and not accounting for it is a common mistake that decision-makers (and/or procedures) make.

Next, suppose a person says that "*price is more important than storage capacity*" about computer hard-disk drives. But if the prices of three disks under consideration are $50, $55, and $60 with storage capacities 1000 GB, 2000 GB, and 3000 GB, the decision is completely different from if prices were $50, $70, and $90 for hard drives with storage capacities 1300 GB, 1400 GB, and 1500 GB. Almost no matter how we weigh price in relation to capacity, we choose the third hard drive in the first of these two examples and the first one in the second. To sum up: **the key is to rank the criteria according to the differences between the options in each criterion.**

In the first example, only $10 distinguishes 2000 GB of storage capacity and in the second $40 distinguishes 200 GB. We must pitch these differences against each other, not the absolute values themselves. "*Price is more important than storage capacity*" is therefore a meaningless statement and constitutes insufficient information to proceed with in a decision analysis. Such a statement will lead you completely astray. When we have to rank the criteria, it is hence important to rank the respective ranges between the best and worst options under each criterion. It is precisely here that many decision-makers (or rather decision-analytic procedures) fail, so it deserves to be taken seriously.

The standard solution in literature, as well as in many computer tools, is to use a so-called **swing** technique. The main steps of a swing procedure are:

- Consider the difference between the worst and the best alternatives within each criterion
- Imagine a fictitious alternative (called the zero alternative) with the worst alternatives from the respective criteria or imagine a fictitious alternative (called the optimal alternative) with the best alternatives from the respective criteria
- For each criterion in turn, consider the improvement (swing) in the zero alternative by having the worst alternative in that criterion replaced by the best one or, conversely, consider the deterioration (swing) in the optimal alternative by having the best alternative in that criterion replaced by the worst one
- Assign importance numbers (or a ranking) to each criterion in such a way that they correspond to the assessed improvement/deterioration from having the criterion changed from the worst/best to the best/worst alternative

In this way weights are assigned, having the desired property of being relative, i.e. the weights reflect the underlying value scales. While this is a blessing for the correctness of the analysis, it is a burden for the decision-maker since the weights become invalid once at least one of the underlying criteria scales change, because the value of an alternative changed in such a way that the scale span is either enlarged or shrunk.

The need for improvement in precision differs between various input modes, viz. whether the utility values are expressed as numbers or rankings and whether the criteria weights are expressed as numbers or rankings. This leads to the following table:

Technique selection		**Criteria weights**	
		Numerical	*Ranking*
Utility values	*Numerical*	DURENO-I	DURENO-II
	Ranking	No need	No need

Table 1. *Selection of DURENO procedure based on input formats*

If the utility values are entered by a ranking scale, the underlying scale will always be automatically normalised, say a $[0,1]$ scale. Thus, there is no need for a rescaling/renormalisation.

The DURENO-I Procedure

Fortunately, if the decision-maker is assisted by a computer tool, it is possible to readjust the numerical weights by a procedure called **duality renormalisation** (hence the name DURENO). The algorithm for such a renormalisation can be described in pseudo-code as follows:

```
/* Procedure DURENO-I: Assume that there are m criteria indexed 1..m */
/* The k:th criterion has had its scale changed by new information */
/* The values for the consequences reside in c_vals[k] for criterion k */
new_min[k] = get_min(c_vals[k]);
new_max[k] = get_max(c_vals[k]);
/* The old scale endpoints are in old_min[k] and old_max[k] */
scale_factor = (new_max[k]-new_min[k]) / (old_max[k]-old_min[k]);
/* The criteria weights are stored in w_crit */
/* Calculate the new weight norm */
norm = (scale_factor-1.0)* w_crit[k] + 1.0;
/* Rescale the affected criterion by the scaling factor */
w_crit[k] *= scale_factor;
/* Renormalise all criteria */
for (i=1; i<=m; i++)
      w_crit[k] /= norm;
```

This process is straightforward to carry out on a computer. And while the algorithm description only deals with pointwise weights, it is easy to extend it to work with intervals as well. The same duality renormalisation occurs for the upper and lower bounds as for the mean (or only) weight per criterion.

The DURENO-I procedure solves the most problematic issue with using MCDA weights incorrectly for numerical criteria weights. But there remains a related problem for ranked criteria weights. Those weights are most often arrived at by ordering them from most to least important and letting a computer tool assign weights based on the ranking information. The precision in such an assignment process can, by the very nature of the procedure, be lower compared to the utility values that constitute the other large information source in an MCDA analysis. A procedure that increases the ranked weights' precision (and thus quality) is considered next..

The DURENO-II Procedure

The DURENO-II technique consists of an amended swing-type technique at its core. But while a traditional swing session embraces only from-worst-to-best swings, DURENO-II also employs intermediate comparisons. This will, at a fast rate, aid the convergence of the weights for the criteria. Second, there is no use for zero alternatives or similar synthetic constructs, instead making use of many more real data points. We also introduce intervals around the surrogate weights to enable a sensitivity analysis during the evaluation phase.

Assume that values for each attribute A_i under each criterion C_j have been elicited. The ensuing step will be to assign weights to the criteria such that $\sum_j w_j = 1$. The DURENO-II procedure is then carried out in two steps as follows. The basic idea is that after an ordinary weight assignment has been carried out, another step is added for the purpose of verifying that the initial weighting statements made are still valid, i.e. an indication that the decision-maker is aware of what he or she is expressing. Another important feature is the possibility of increasing the precision in the estimates by comparing sub-scales with each other. The DURENO-II procedure steps are:

a) In a traditional swing-type session, the decision-maker is asked to compare the swings between the endpoints (best and worst outcome) on the criteria's respective value scales. The criteria weights are ranked using an ordinal ranking function amended with '='. Questions asked are of the type "*Which is the most important to you: the difference between the scale endpoints in criterion C_i or in criterion C_j?*" The result of this step might, for instance, be the ranking $w_1 > w_2 = w_3 > w_4 = w_5$, which is represented by a weight ranking together with a suitable interval around them.

   Note that if we assume that v_{i0} and v_{i1} are the endpoints of the value scale for criterion C_i, the comparisons are then of the type $(v_{i1}-v_{i0})\cdot w_i > (v_{j1}-v_{j0})\cdot w_j$, i.e. of the character of standard comparisons $w_j > w_i$.

b) The next step consists of fractions of the criteria's respective value scales being compared. Questions asked are now of the type *"Which is the most important to you: the difference between the values α_2 and α_1 in criterion C_i or between the values α_4 and α_3 in criterion C_j?"* This step thus introduces a new feature: the possibility of comparing parts of scales with each other.

   The statements then consequently become of the type $(\alpha_1\cdot v_{i1}-\alpha_2\cdot v_{i0})\cdot w_i > (\alpha_3\cdot v_{j1}- \alpha_4\cdot v_{j0})\cdot w_2$ for real value statements α_1 to α_4 in [0,1], where $\alpha_m\cdot v_{i1}-\alpha_n\cdot v_{i0} > 0$, for all *i, n, m*.

   This also means that the questions only focus on real alternatives that exist in the current decision context. This way, a revised system of inequalities (and equalities) is formed, and if this system has a solution, it is consistent, i.e. the decision-maker has made a consistent assessment of how important different criteria are compared to each other. The weights are adjusted in accordance with the new system.

After a session, two sets of linear constraints are produced, one containing the values of the alternatives under the respective criterion and one containing the weight statements.

Example

Assume a procurement process for new premises is initiated because the existing premises have become inadequate. The choice is between four office space providers, *A*, *B*, *C* and *D*, to realise this project. The criteria emphasised are *functionality* (basically the degree of adequacy of the new premises), *localisation* (geographical and infrastructural), *opportunities for interaction with the surrounding society,* and *price.*

Elicitation

To begin with, the values for the alternative providers are summarised below. The qualitative scales are in the range [0,1] and the scale for prices is the actual price.

Functionality	**Localisation**	**Opportunities**	**Price**
A is better than B B is slightly better than C C is better than D	B is slightly better than C C is better than A A is better than D	B is better than A A is better than C C is better than D	A costs 5.5 MEUR B costs 6.0 MEUR C costs 5.0 MEUR D costs 4.0 MEUR

That results in the following value statements:

$v_F(A) >_2 v_F(B)$	$v_L(B) >_1 v_L(C)$	$v_O(B) >_2 v_O(A)$	$v_P(A) = 5.5$
$v_F(B) >_1 v_F(C)$	$v_L(C) >_2 v_L(A)$	$v_O(A) >_2 v_O(C)$	$v_P(B) = 6.0$
$v_F(C) >_2 v_F(D)$	$v_L(A) >_2 v_L(D)$	$v_O(C) >_2 v_O(D)$	$v_P(C) = 5.0$
			$v_P(D) = 4.0$

Following the process described above, and assuming that there are no immediate conflicts in the initial preferences, they make an initial ranking that results in Functionality being the most important criterion, followed by Localisation. Thereafter follows Opportunities and lastly Price.

Considering the scale endpoints, assume that the participants provide the following statements as a result of step (i), yielding the following initial ranking:

Functionality is slightly more important than Localisation, which is more important than opportunities. Finally, Opportunities is more important than Price. This is translated into the following cardinal ranking order.

- $w(F) >_1 w(L)$
- $w(L) >_2 w(O)$
- $w(O) >_2 w(P)$

In step (ii), having seen this, the participants provide the following supplementary statements for the criteria:

- The difference between B and C in *Functionality* is more important than B and A in *Localisation*.
- The difference between C and D in *Functionality* is more important than A and D in *Opportunity*.
- The difference between C and A in *Localisation* is more important than B and D in *Opportunity*.
- The difference between B and C in *Localisation* is more important than a *price* difference of 1 MEUR.

Evaluation

The problem is structurally simple, but it is still relatively difficult to give an absolute recommendation by just considering the problem intuitively. Thus, first, we determine the value scales, resulting in the following values:

Criterion Functionality:

A	0.950
B	0.600
C	0.400
D	0.050

Criterion Localisation:

A	0.400
B	0.950
C	0.800
D	0.050

Criterion Opportunities:

A	0.667
B	0.958
C	0.333
D	0.042

Criterion Price:

A	0.250
B	0.063
C	0.500
D	0.938

Thereafter, we calculate the criteria weights to be the following:

w(F)	0.453
w(L)	0.302
w(O)	0.170
w(P)	0.075

The supplementary statements arrived at in step (ii) in the process, are translated as:

- $0.4w(F) > 0.6w(L)$
- $0.4w(L) > w(O)$
- $0.4w(F) > 2/3\ w(O)$
- $0.2w(L) > 0.5w(P)$

After the statements in step (ii) have been considered, the weight have been adjusted but are still consistent, which *indicates* that the decision-makers have understood the relativeness of the criteria weights.

The modified weight intervals and adjusted midpoint are then the following:

w(F)	0.480
w(L)	0.283
w(O)	0.163
w(P)	0.074

This new set of weights is a better one in the sense that it more closely mirrors the criteria priorities that the decision-maker has expressed during the process.

Together with DURENO-I, this procedure makes up a pair that considerably enhances the quality and precision of the decision analysis by eliminating the two must troubling sources of error and mistrust in the analysis results. In essence, the DURENO techniques make MCDA much more useful in real-life decision situations.

Note: An earlier version of the DURENO-II procedure is known under the name P-SWING (where P stood for “precision”, later “partial”). However, no DURENO-I procedure accompanies P-SWING. Therefore, the complete DURENO set of techniques is to prefer in order to make the weight elicitation process as versatile as possible since it is also applicable to numerical weights, not only ranking weights.