Symmetry in Decision Evaluation

Mats Danielson[1] and Love Ekenberg[2]

[1]DECIDE Research Group
Department of Computer and Systems Sciences
Royal Institute of Technology
Electrum 230, SE-164 40 Kista, SWEDEN

[2]Department of Information Technology
Mid Sweden University
SE-851 70 Sundsvall, SWEDEN

email: {mad,lovek}@dsv.su.se

Abstract

Several models for handling vague and imprecise information in decision situations have been suggested. In those contexts, various interval methods have prevailed, i.e. methods based on interval estimates of probabilities and, in some cases, interval utility estimates. Even if these approaches in general are well founded, little has been done for demonstrating whether the approaches are comprehensible for a decision maker. In particular, it is far from always clear what is actually expressed by a set of intervals where linear dependencies do occur. Furthermore, it is difficult to find reasonable decision rules that select an alternative out of a set of alternatives and correspond to the intuition of a decision maker. In this article, we investigate some problems that are inherent in interval approaches and suggest how the choice of particular evaluation rules might compensate for this.

Introduction

A quite widespread opinion is that the principle of maximising the expected utility captures the concept of rationality. However, the shortcomings of this principle, and of utility theory in general are severe, and have to be compensated for (Ekenberg, et al, 2001). One of the main problems with the principle is that it requires too hard aggregation of background data. Nevertheless, there is a need for efficient evaluation principles. A number of models with representations allowing imprecise probability statements have been suggested. Some of them are based on

- capacities (of order 2),
- evidence theory and belief functions,
- various kinds of logic,
- upper and lower probabilities, or
- sets of probability measures.

The common feature of the approaches is that they do not include the additivity axiom of probability theory and consequently do not require a decision maker to model and evaluate a decision situation using precise probability (and, in some cases, utility) estimates. The approaches are in general well founded, but quite little has been done in demonstrating whether they are understandable for a decision maker facing a real decision situation. In particular, it is not always clear what an agent expresses when providing, for instance, a set of intervals where linear dependencies do occur. The problem becomes of particular significance when evaluation models are considered. An advantage of approaches for upper and lower probabilities is that it is not necessary to take particular probability distributions into consideration. On the other hand, it is then difficult to find reasonable decision rules that chose an alternative out of a set of alternatives, and that really corresponds to the intuition of a decision maker. This problem is emphasised by the fact that no distributions over the intervals are taken into account to indicate values of most significance. The low-dimensional intuition of decision makers further adds to the problem. Effects of changes on input data are not always simple to communicate. We have found two complementary views on the input data to be particularly helpful to the decision maker – symmetry and quadracy.

Preliminaries

Interval approaches model decision situations, where numerically imprecise statements such as "the probability of consequence c is greater than 25%" or "the value of consequence c is between 100 and 300" occur. Furthermore, some such approaches also allow for comparative sentences such as "consequence c is preferred to consequence d". These statements are then represented in a numerical format.

Typically, the alternatives are represented by the consequences they might imply. Over these consequences, convex sets of candidates of possible probability and utility functions are defined. For instance, in (Danielson and Ekenberg, 1998) an approach is suggested, where the possible functions are expressed as vectors in polytopes that are solution sets to sets of probability and utility estimates. That the probability of c_{ij} lies between the numbers a_i and b_i is expressed as $p_{ij} \in [a,b]$, i.e. $p_{ij} \geq a$ and $p_{ij} \leq b$. Similarly, that the probability of c_{ij} is greater than the probability of c_{kl} is expressed by the inequality $p_{ij} \geq p_{kl}$.

Each statement is in this way represented by one or more constraints. The sets of probability estimates under consideration is the set of constraints of the types above, together with the equation $\sum_j p_{ij} = 1$, for each alternative involved. The utility estimates are represented by a set of constraints in a similar way.

Definition 1: A *decision frame* is a structure $\langle\{C_1,\ldots,C_m\},P,V\rangle$, where each C_i is a finite set of consequences $\{c_{i1},\ldots,c_{ih_i}\}$. P is a finite list of linear constraints in the probability variables and V is a finite list of linear constraints in the value variables.

To evaluate a decision frame, it is important to find optima for given objective functions. The following two definitions are intended to simplify the procedures suggested in the following sections.

Definition 2: Given a consistent constraint set X in $\{x_i\}_{i\in I}$ and a function f,

$X_{max}(f(x)) =_{def} \sup(a \mid \{f(x) > a\} \cup X$ is consistent).

Similarly, $X_{min}(f(x)) =_{def} \inf(a \mid \{f(x) < a\} \cup X$ is consistent)

Definition 3: Given a consistent constraint set X in $\{x_i\}_{i\in I}$, the set of pairs $\{\langle X_{min}(x_i),X_{max}(x_i)\rangle\}_{i\in I}$ is the *orthogonal hull* of the set and is denoted $\langle X_{min}(x_i),X_{max}(x_i)\rangle_n$.[1]

The focal point is the most likely vector as perceived by the decision maker. It can be changed during the course of interaction, and varying focal points can be chosen by the decision maker at different times according to his appreciation of the current decision situation.

Definition 4: Given a constraint set X in $\{x_i\}_{i\in I}$ and the orthogonal hull $H = \langle a_i,b_i\rangle_n$ of X, a *focal point* is a solution vector $(r_1,\ldots,r_n)$ with $a_i \le r_i \le b_i$, $\forall i\in I$. The *hull midpoint* is $(m_1,\ldots,m_n)$ with $m_i = \frac{a_i + b_i}{2}$.

Further, an acceptable metric should be defined that complies with the decision maker's understanding of the decision problem. Thus, the standard concept of distance is introduced.

Definition 5: Given two vectors **a** and **b**, the *distance function* d is a function that satisfies

(i a) $d(\mathbf{a},\mathbf{b}) > 0$ if $\mathbf{a} \ne \mathbf{b}$
(i b) $d(\mathbf{a},\mathbf{a}) = 0$
(ii) $d(\mathbf{a},\mathbf{b}) = d(\mathbf{b},\mathbf{a})$
(iii) $d(\mathbf{a},\mathbf{b}) \le d(\mathbf{a},\mathbf{c}) + d(\mathbf{c},\mathbf{b})$ for all **c**.

For the definition to be meaningful in this context, the distance function must be reasonable, even though this does not follow directly from the definition. In general, the focal point does not need to coincide with the orthogonal hull midpoint. In fact, the hull midpoint need not even be consistent. In those cases, a set of constraints is said to be skewed, and the concept of skewness is introduced to describe this.

Definition 6: Given a constraint set X in $\{x_i\}_{i\in I}$, two real vectors $\mathbf{a} = (a_1,\ldots,a_n)$ and $\mathbf{b} = (b_1,\ldots,b_n)$ of the orthogonal hull $\langle a_i,b_i\rangle_n$ of X, a distance function d, a constant $k \in [0,1]$, a hull midpoint **m**, and a focal point **r**. The *skewness* of the base X with respect to **r** is $k \cdot \frac{d(\mathbf{r},\mathbf{m})}{d(\mathbf{a},\mathbf{b})}$.

As will be discussed in the section on symmetry, when a set of constraints is skewed, there exists a way of aiding the decision maker in avoiding this asymmetry by using the symmetric hull instead.

Definition 7: Given a constraint set X in $\{x_i\}_{i\in I}$, the orthogonal hull $\langle a_i,b_i\rangle_n$ of X, and a focal point $(r_1,\ldots,r_n)$. Let $d_i = \min(r_i-a_i, b_i-r_i)$, $\forall i\in I$. The *symmetric hull* is $\langle r_i-d_i,r_i+d_i\rangle_n$.[2]

Evaluation

Once the decision maker has entered his decision data into an evaluation tool, the evaluation phase commences. In any interesting decision problem stated by intervals, the solutions overlap in the sense that there exists no single course of action preferred regardless of which vectors are chosen from the polytope defined by the intervals. If there would be such a single course of action, then any experienced decision maker would be able to realize this without the aid of a decision support machinery.

The decision situation can be evaluated by calculating the expected value for each alternative, but which of the infinite number of vectors should be chosen as representative for the alternative in the calculation? And what if comparisons exist between consequences in the two alternatives? There is a strong element of comparison inherent in a decision procedure. The evaluation results are interesting in comparison to the results of the other consequence sets. Hence, it is reasonable to consider the differences in expected value (strength) as well. Then it makes sense to evaluate the relative strength of C_i compared to C_j in addition to the strengths themselves, since such strength values are compared to some other strengths anyway in order to rank the consequence sets.

[1] I is an index set, i.e. a set of integers.

[2] If the symmetric hull coincides with the orthogonal hull, then the skewness is zero. This follows from $d(\mathbf{r},\mathbf{m}) = 0$ if the midpoint **m** is equal to the focal point **r**.

Definition 8: Given a decision frame $\langle\{\{C_{ik}\}_{m_i}\}_m, P_0, V_0\rangle$, δ_{ij} denotes the expression $\Sigma_k\, p_{ik}\cdot v_{ik} - \Sigma_k\, p_{jk}\cdot v_{jk} = p_{i1}\cdot v_{i1} + p_{i2}\cdot v_{i2} + \ldots + p_{im_i}\cdot v_{im_i} - p_{j1}\cdot v_{j1} - p_{j2}\cdot v_{j2} - \ldots - p_{jm_j}\cdot v_{jm_j}$ over all consequences in the consequence sets C_i and C_j.

This is, however, not enough. Sometimes, the decision maker wants to put more emphasis on the maximal difference (displaying a difference-prone behavior). At other times, the minimal difference is of more importance. This is captured in the medium difference.

Definition 9: Given a decision frame $\langle C,P,V\rangle$, let $\alpha \in [0,1]$ be real number. The *α-medium difference of* δ_{ij} in the frame is ${}^{PV}[\alpha]\text{mid}(\delta_{ij}) = \alpha\cdot{}^{PV}\max(\delta_{ij}) + (1-\alpha)\cdot{}^{PV}\min(\delta_{ij})$. The *average difference of* δ_{ij} in the frame is ${}^{PV}\text{avg}(\delta_{ij}) = {}^{PV}[0.5]\text{mid}(\delta_{ij})$.

The α can be considered a precedence parameter that indicates if one boundary should be given more weight than the other. The average is also the relative strength, i.e. the difference in maximal δ-values when the frame is considered from the viewpoint of each consequence set respectively. Thus, it is a measure of difference in strength between the consequence sets. This view duality is a key to understanding the selection process proposed later.

Definition 10: The *relative strength* of C_i compared to C_j in a decision frame is

$$^{PV}mid(\delta_{ij}) = \frac{^{PV}\max(\delta_{ij}) - {}^{PV}\max(\delta_{ji})}{2}$$

Dominance

The selection procedure suggested in this paper is based on the contraction principle as introduced below and on the concepts of strong, marked, and weak dominance as introduced in definition 11.

Definition 11: Given a decision frame $\langle C,P,V\rangle$, C_i *strongly dominates* C_j iff

$$^{PV}\min\left(\sum p_{ik}\cdot v_{ik} - \sum p_{ik}\cdot v_{jk}\right) \geq 0\,.$$

C_i *markedly dominates* C_j iff

$$^{PV}avg\left(\sum p_{ik}\cdot v_{ik} - \sum p_{jk}\cdot v_{jk}\right) \geq 0\,.$$

C_i *weakly dominates* C_j iff

$$^{PV}\max\left(\sum p_{ik}\cdot v_{ik} - \sum p_{jk}\cdot v_{jk}\right) \geq 0\,.$$

The decision maker's selection procedure then proceeds as follows. Usually a number of consequence sets are still being considered. An example shows the use of dominance.

Contraction

The contraction is a generalized sensitivity analysis to be carried out in a large number of dimensions. In non-trivial decision situations, when a decision frame contains numerically imprecise information, the domination principles suggested above are often too weak to yield a conclusive result by themselves. Thus, after the elimination of undesirable consequence sets, the decision maker could still find that no conclusive decision has been made. One way to proceed could be to determine the stability of the relation between the consequence sets under consideration. A natural way to investigate this is to consider values near the boundaries of the constraint intervals as being less reliable than the core due to the former being deliberately imprecise. This is taken into account by measuring the dominated regions indirectly using the concept of contraction.

Definition 12: Given a decision frame X with the variables $x_1,\ldots,x_n$, $\pi \in [0,1]$ is a real number, and $\{\pi_i \in [0,1] : i = 1,\ldots,n\}$ is a set of real numbers. $[a_i, b_i]$ is the interval corresponding to the variable x_i in the solution set of the system of constraints, and $(k_1,\ldots,k_n)$ is a focal point in X. A *π-contraction* of X is to add the interval statements $\{x_i \in [a_i+\pi\cdot\pi_i\cdot(k_i-a_i),\ b_i-\pi\cdot\pi_i\cdot(b_i-k_i)] : i = 1,\ldots,n\}$ to the frame X.

Contrary to volume estimates, contractions are not measures of sizes of solution sets but rather of the strength of statements when the original solution sets are modified in controlled ways. Both the set of intervals under investigation and the scale of individual contractions can be controlled. Consequently, an expansion can be regarded as a focus parameter that zooms out from central sub-intervals (the core) to the full statement intervals. This is not a precise selection procedure, and it is not meant to be. It should be guided by the aims and investigation patterns of the decision maker. Its particular instantiation depends on the decision situation, whether the decision maker is a human or a machine, and whether the goal is to make an ultimate decision or (very common for humans) to gain a better understanding of the decision problem.

Frame Symmetry

A couple of frame views may affect the evaluation. They can be regarded as views taken by the decision maker on the input data. The first view is the symmetry of the decision frame as introduced above. A frame is skewed (or asymmetric) if the focal point does not coincide with the hull midpoint. For skewed frames, the symmetric hull is meaningful and constructed by adjusting the interval of each hull dimension from one side so that the focal and midpoints coincide. The idea is that considerably skewed bases might be the result of misconceptions or mistakes from the decision maker. This is a view taken on the input material, and it might be changed during the evaluation to appreciate its effect on the outcome of the selection process.

Example 1: Consider a decision situation involving two consequence sets C_1 and C_2. C_1 has ten consequences while C_2 has only one. There are no probability constraints and the probability focal point for C_1 is 0.1 for each consequence. While the orthogonal hull covers all consistent probability assignments, i.e. [0,1] for each probability variable, the symmetric hull is symmetric around the focal point, i.e. [0,0.2]. The value base contains.

$v_{11} \in [1.00, 1.00]$
$v_{1i} \in [0.00, 0.00]$ for i=2..10
$v_{21} \in [0.10, 0.20]$

An evaluation using the orthogonal hull results in the consequence set C_1 being the preferred one for almost all contractions. This is counter to many decision maker's appreciation of the example. An evaluation using the symmetric hull, on the other hand, yields that consequence set C_2 is the preferred one for all contractions and the result is stable. This result is perceived to be more indicative by many decision makers.

Quadratic Compensation

Another frame view is the quadracy of the expected value. The idea is that the decision maker works in a very low-dimensional fashion, considering only a few statements at the same time. Then it is sometimes perceived as an unexpected effect when wider intervals centered around the same focal point evaluates to better numeric values. One candidate to alleviate this effect is the *quadratic compensation*. Perhaps it is easiest pointed out by some examples.

Example 2: Consider a decision situation involving two consequence sets C_1 and C_2 that have two consequences each. The corresponding decision frame contains the following statements.

$p_{11} \in [0.00, 1.00]$	$v_{11} \in [0.30, 0.70]$
$p_{12} \in [0.00, 1.00]$	$v_{12} = 0.00$
$p_{21} \in [0.40, 0.60]$	$v_{21} \in [0.30, 0.70]$
$p_{22} \in [0.00, 1.00]$	$v_{22} = 0.00$

The orthogonal hull for the probability base is

$p_{1i} \in [0.00, 1.00]$ for i=1..2
$p_{2i} \in [0.40, 0.60]$ for i=1..2

For a π-contraction, the probability variables for $\pi \in [0,1]$ become

$p_{1i} \in [\pi/2, 1-\pi/2]$ for i=1..2
$p_{2i} \in [0.40+\pi/10, 0.60-\pi/10]$ for i=1..2

Since the example is simple, the calculations can be done directly.

$^{PV}avg(\delta_{12}) = ((1-\pi/2)\cdot 0.7 - (0.4+\pi/10)\cdot 0.3 - ((0.6-\pi/10)\cdot 0.7 - \pi/2\cdot 0.3))/2 = 0.08 - 0.08\cdot\pi$

According to the rule, consequence set C_1 is the preferred one for all contractions $\pi \in [0,1[$. ■

The only difference between the two consequence sets in the example is the width of the probability statement for the first consequence, reflecting greater uncertainty. This translates into higher δ-values because of the quadracy of the expected value. While this is mathematically correct, it is perceived by some decision makers as a deficiency. It does not fit into their view of the data, and the evaluation of dominance should allow for this to be compensated. For small, local changes to single intervals, the effect on the expected value of adjusting both probability and value statements may be linearized by using quadratic compensation. The quadratically compensated medium δ_{ij} for the expected value becomes the following expression.[3]

Definition 13: Given a decision frame ⟨C,P,V⟩, the *quadratic average difference of* δ_{ij} in the frame is

$$^{PV}qavg(\delta_{ij}) = \frac{^{PV}\max(\delta_{ij}) + {}^{P}\max{}^{V}\min(\delta_{ij}) + {}^{P}\min{}^{V}\max(\delta_{ij}) + {}^{PV}\min(\delta_{ij})}{4}$$

The quadratically compensated average concept leads to a modification of the rule for marked dominance.

Definition 14: Given a decision frame ⟨C,P,V⟩, C_i *q-markedly dominates* C_j iff

$$^{PV}qavg\left(\sum p_{ik}\cdot v_{ik} - \sum p_{jk}\cdot v_{jk}\right) \geq 0.$$

It is not the case that either the standard or the quad-compensated marked dominance is the "natural" one, even though formally the case is clear. They represent different views on the input data, preferably to be considered together, depending on the type of application at hand.

Definition 15: Given a decision frame ⟨C,P,V⟩, the *medium difference* δ_{ij} in the frame is

$$^{PV}mid(\delta_{ij}) = \frac{^{P}\max{}^{V}\min(\delta_{ij}) + {}^{P}\min{}^{V}\max(\delta_{ij})}{2}.$$

Then it follows that

---

3 $P_{max}V_{min}(\sum_k p_{ik}\cdot v_{ik})$ should be interpreted as $P_{max}(V_{E_i}^{min})$ with $V_{E_i}^{min}$ as an optimizing procedure to find the local minimum in the value constraint set..

$$^{PV}qavg(\delta_{ij}) = \frac{^{PV}avg(\delta_{ij}) + {}^{PV}mid(\delta_{ij})}{2}.$$

This simplifies the calculations in the example.

Example 2 (cont'd): Let the example above continue by applying the newly defined functions. $^{PV}qavg(\delta_{ij})$ is calculated as follows.

$^{PV}mid(\delta_{12}) = ((1-\pi/2)\cdot 0.3 - (0.4+\pi/10)\cdot 0.7 - ((0.6-\pi/10)\cdot 0.3 - \pi/2\cdot 0.7))/2 = -0.08 + 0.08\cdot\pi$

$^{PV}qavg(\delta_{12}) = (0.08 - 0.08\cdot\pi - 0.08 + 0.08\cdot\pi)/2 = 0.$

The $^{PV}mid(\delta_{12})$ is the same as $^{PV}avg(\delta_{12})$ but with opposite signs. In a sense, $^{PV}mid(\delta_{12})$ balances out $^{PV}avg(\delta_{12})$. Since the only difference between the consequence sets is the width of the intervals, the $^{PV}qavg(\delta_{12})$ difference is zero.

Example 3: Next consider a similar decision situation involving two consequence sets C_1 and C_2 having two consequences each. In this case, C_1 has wider statements but C_2 has the most favorable focal point. The corresponding decision frame contains the following constraints:

$p_{11} \in [0.00, 1.00]$	$v_{11} \in [0.00, 1.00]$
$p_{12} \in [0.00, 1.00]$	$v_{12} = 0.00$
$p_{21} \in [0.50, 0.70]$	$v_{21} \in [0.50, 0.70]$
$p_{22} \in [0.00, 1.00]$	$v_{22} = 0.00$

The orthogonal hull for the probability base is

$p_{1i} \in [0.00, 1.00]$ for i=1..2
$p_{21} \in [0.50, 0.70]$
$p_{22} \in [0.30, 0.50]$

For a π-contraction the probability variables for $\pi \in [0,1]$ become

$p_{1i} \in [\pi/2, 1-\pi/2]$ for i=1..2
$p_{21} \in [0.50+\pi/10, 0.70-\pi/10]$
$p_{22} \in [0.30+\pi/10, 0.50-\pi/10]$

Then $^{PV}avg(\delta_{12})$ and $^{PV}mid(\delta_{12})$ becomes:

$^{PV}avg(\delta_{12}) = ((1-\pi/2)\cdot 1 - (0.5+\pi/10)\cdot 0.5 - ((0.7-\pi/10)\cdot 0.7))/2 = 0.13 - 0.24\cdot\pi$

$^{PV}mid(\delta_{12}) = (-(0.5+\pi/10)\cdot 0.7 - ((0.7-\pi/10)\cdot 0.5 - \pi/2\cdot 1))/2 = -0.35 + 0.24\cdot\pi$

$^{PV}qavg(\delta_{12}) = (0.13 - 0.24\cdot\pi - 0.35 + 0.24\cdot\pi)/2 = -0.11.$

According to $^{PV}avg(\delta_{12})$, consequence set C_1 is the preferred one for contractions π up to 54%, after which C_2 is to prefer. According to $^{PV}mid(\delta_{12})$, consequence set C_2 is again to prefer for all π. The contraction dependent term in $^{PV}avg(\delta_{12})$ is the same as in $^{PV}mid(\delta_{12})$ but with opposite signs. Thus, they cancel out in $^{PV}qavg(\delta_{12})$, but the constant -0.11 remains. This is in accordance with many decision maker's understanding of the input data.

Conclusions

Using interval approaches, it is difficult to find reasonable decision rules that choose an alternative out of a set of alternatives, and that really corresponds to the intuition of a decision maker. We have found two complementary views on the input data to be particularly helpful to the decision maker – symmetry and quadracy. The idea of symmetry is that considerably skewed bases might be the result of misconceptions or mistakes from the decision maker. The idea of quadracy of the expected value is that the decision maker works in a very low-dimensional fashion, considering only a few statements at the same time. Then it is sometimes perceived as an unexpected effect when wider intervals centered around the same focal point evaluates to better numeric values. By offering the decision maker alternative views on the data, he is in a better position to appreciate the situation. These concepts scale well, especially when there are no comparisons of probability values between different consequence sets. Then ordinary linear programming methods can be employed (Danielson and Ekenberg, 1998).

References

M. Danielson and L. Ekenberg, "A Framework for Analysing Decisions under Risk," European Journal of Operational Research, vol. 104/3, pp. 474-484, 1998.

L. Ekenberg, M. Boman, and J. Linneroth-Bayer, "General Risk Constraints," Journal of Risk Research 4 (1), pp.31-47, 2001.