DMM Summary

Massimo Perego

Disclaimer This "summary" it's not supposed to be a standalone tool.

Contents

1 Decision Problems

1.1 Preference Relation

1.1.1 Properties

Considering impact set F and a binary relation Π, the main properties of said relation can be:

- **Reflexivity:** $\forall a \in F$, $(a, a) \in \Pi$; to check for it, on the graph representation every impact has to have a self-loop
- **Transitivity:** $\forall a, b, c \in F$, if $(a, b) \in \Pi$ and $(b, c) \in \Pi$ then $(a, c) \in \Pi$; to check for it, on the graph representation follow every arc and check that every possible "triangle" is complete
- **Antisymmetry:** $\forall a, b \in F$, if $(a, b) \in \Pi$ and $(b, a) \in \Pi$ then $a = b$; to check for it, on the graph representation there must be no impacts that point at each other
- **Completeness:** $\forall a, b \in F$, if $(a, b) \notin \Pi$ then $(b, a) \in \Pi$; to check for it, on the graph representation, from each node there must be an arc (either outgoing or incoming) connecting it to every other

1.1.2 Types of relations (orders)

Combining the properties, we can get different kinds of preferences:

- **Preorder:** reflexivity and transitivity. Guarantees that the set of nondominated alternatives is nonempty, even if not all alternatives are comparable, we can find at least one nondominated solution
- **Partial order:** reflexivity, transitivity and antisymmetry. This limits indifference, but still allows for incomparability, and thus not always leading to a definitive choice
- **Weak order:** reflexivity, transitivity and completeness. Guarantees that nondominated solutions exists and are all mutually indifferent, any of them can be chosen. Such orders admit representation by a value function (with ties), turning the decision problem into an optimization problem
- **Total order:** reflexivity, transitivity, antisymmetry and completeness. Provides a unique linear ranking of alternatives, there is always a unique best alternative

1.1.3 Derived relations

From the weak preference relation, one can derive:

- **Indifference relation:** Ind_Π

  $$(a,b),(b,a) \in \mathrm{Ind}_\Pi \Leftrightarrow (a,b) \in \Pi \wedge (b,a) \in \Pi$$

  To build it from the graph representation, add all self loops and each pair of arcs that point at each other

- **Strict preference relation:** Str_Π

  $$(a,b) \in \mathrm{Str}_\Pi \Leftrightarrow (a,b) \in \Pi \wedge (b,a) \notin \Pi$$

  To build it from the graph representation, add all arcs which do not have an equal one in the opposite direction (the set difference of the last one w.r.t. Π)

- **Incomparability relation:** Inc_Π

  $$(a,b),(b,a) \in \mathrm{Inc}_\Pi \Leftrightarrow (a,b) \notin \Pi \wedge (b,a) \notin \Pi$$

  To build it from the graph representation, add all arcs which are not present in either direction in the graph

2 Basic Decision Models

2.1 Fundamental definitions

A decision problem is defined by a 6-uple:

$$P = (X, \Omega, F, f, D, \Pi)$$

Where

- X is the feasible region, set of all alternatives
- Ω is the sample space, set of all possible scenarios
- F is the indicator space, set of all possible impacts
- $f : X \times \Omega \to F$ is the impact function, associates each configuration of the system to a impact
- D is the set of all decision-makers
- $\Pi : D \to 2^{F \times F}$ is the preference function, which associates to each decision maker a subset of impact pairs

The **alternatives** formally describe the events under the control of the decision-makers. X includes all possible choices.

The **scenarios** formally describe the events out of the control of the decision-makers. Ω defines all such events.

The *impacts* model all aspects relevant to the decision, they are described quantitatively as vector of real numbers.

The **impact function** is a vectorial function which associates each configuration to an impact.

A **decision-maker** is whoever takes part in the decision.

The decision-maker takes part directly in the choice of alternative, holding the power and responsibility to decide, while **stakeholders** are all the subjects who don't actively participate in the decision but whose interests are affected by its outcome.

In a decision problem, an **outperformer** $x \in X$ for $d \in D$ is an alternative which performs at least as good as any other $x' \in X$ in every scenario $\omega \in \Omega$

$$(f(x, \omega), f(x', \omega)) \in \Pi_d, \quad \forall \omega \in \Omega, x' \in X$$

and is strictly better for at least one comparison/scenario

$$\exists x^\circ \in X, \omega^\circ \in \Omega \quad (f(x,\omega), f(x^\circ,\omega)) \in \Pi_d \wedge (f(x^\circ),\omega), f(x,\omega)) \notin \Pi$$

2.2 Preference relation

For each decision maker, we need a relation between pair of impacts to determine which ones are preferable. $\Pi : D \to 2^{F \times F}$ associates each decision-maker to a subset of impact pairs.

A **weak preference** $f \preceq_d f'$ is when the decision-maker accepts the exchange of f for f'.

Two impacts are **indifferent** if the decision-maker considers both equally as satisfactory, i.e., is willing to exchange one for the other in both directions.

Two impacts are **incomparable** if the decision-maker is unable or unwilling to choose between them, rejecting the exchange in both directions.

2.2.1 Property of binary relations

Some preference relations enjoy special properties:

- **Reflexivity:** each impact is in relation with itself

$$\forall f \in F, \quad (f,f) \in \Pi$$

- **Transitivity:** every "triangle" of preference is complete

$$\forall f,g,h \in F, \quad \text{if } (f,g) \in \Pi \wedge (g,h) \in \Pi \implies (f,h) \in \Pi$$

- **Antisymmetry:** two impacts are indifferent only if they're equal

$$\forall f,g \in F, \quad (f,g) \in \Pi \wedge (g,f) \in \Pi \implies f = g$$

- **Completeness:** every impact is in relation with every other, one way or another

$$\forall f,g \in F, \quad (f,g) \in \Pi \vee (g,f) \in \Pi$$

Kinds of relation These properties allow to identify some main kinds of relations:

- **Preorder:** reflexivity and transitivity
- **Partial order:** reflexivity, transitivity and antisymmetry
- **Weak order:** reflexivity, transitivity and completeness
- **Total order:** ALL OF THEM

2.3 Dominance relation

2.3.1 Decision making on weak orders

If a preference relation is a weak order

- the induced dominance is a weak order
- and the solution set is finite and nonempty, nondominated solutions exists and are all mutually indifferent

A **value function** $f : F \to \mathbb{R}$ associates each impact to a real value. A value function is **consistent** with preference relation Π when

$$f \preceq f' \Leftrightarrow v(f) \geq v(f'), \quad \forall f, f' \in F$$

If a preference relation Π admits a consistent value function, then Π it's a weak order.

2.3.2 Weak order preference models

Some ways to sort *the stuff*.

Lexicographic order Order the alternatives w.r.t. the value of the first indicator (for some ordering of the indicators) and break ties with the subsequent one. This yields a total order.

The variant **with aspiration levels** introduces a "minimum requirement" ϵ_i, rejecting all alternatives with indicator f_i worse than ϵ_i (higher or lower depending on whether it's a benefit or cost). It applies the lexicographic order on a restricted feasible region.

Utopia point Identify an ideal impact with the best possible value for each indicator (optimize them independently) and evaluate all alternatives with the distance from the ideal impact.

Different definitions of distance yield different results. Some definitions:

- L_1 Manhattan distance

$$d(f, f') = \sum_{l \in P} |f_l - f'_l|$$

- L_2 Euclidean distance

$$d(f, f') = \sqrt{\sum_{l \in P} \left(f_l - f'_l\right)^2}$$

- L_∞ Chebyshev distance/maximum norm

$$d(f, f') = \max_{l \in P} |f_l - f'_l|$$

Borda count In the case of finite alternatives, they can be sorted by counting how many alternatives are worse than each one

$$B(f) = \left|\left\{f' \in F \mid f \preceq f'\right\}\right|$$

2.4 MAUT

Multi Attribute Utility theory assumes the preference relation of the decision-maker is a weak order, admitting a consistent value function. We need to derive such value function from the preference relation.

2.4.1 Indifference curves

An *indifference curve* is a subset of the impact space $I \subseteq F$ of reciprocally indifferent impacts. By definition:

- The curves cover F
- Any two curves have empty intersection
- Weak order on impacts maps to total order on curves

Usually, **continuity** is assumed (they are mathematical objects and not a general set of points), and each indifference curve is expressed in the implicit form $u(f) = c$, each c identifies a curve.

2.4.2 Determining the utility function

The general process requires sampling and guessing a utility function, with consistency checks. This is complex and error prone.

Some properties which allow for easier estimates of $u(f)$:

- **Invertibility:** $u(f)$ can be solved w.r.t. each f_l
- **Monotony:** to compensate the variation of an indicator, the others must vary in a well-defined direction
- **Convexity** or **concavity:** the indifference curves compensate for the increase of an indicator by a certain amount with variations of the other ones that increase (or decrease) with the value of the first indicator

2.4.3 Additive utility function

A utility function is additive when it can be expressed as the sum of functions of the single indicators. If the utility function is additive, the problem to estimate it can be reduced to the estimation of p single-variable functions, simplifying the process.

Preferential independence A subset of indicators L is **preferentially independent** from the complementary subset $P \setminus L$ when given two impacts with identical values of $P \setminus L$ the preference relation does not depend on such values.

Mutual preferential independence A problem enjoys **mutual preferential independence** when every proper subset of indicator is independent from its complement.

Mutual preferential independence is necessary for additivity. Intuitively if subsets of indicators depend on each other there can't be a way to express the utility function in an additive way by summing such indicators (or value of the subsets).

For a decision problem with $p \geq 3$ indicators mutual preferential independence is a sufficient condition for additivity.

2.4.4 MRS

The Marginal Rate of Substitution λ_{12} between two indicators f_1 and f_2 represents how much of f_1 are we willing to *"give up"* for a unit of f_2; e.g., if we're willing to give 4 units of f_1 for a unit of f_2 then $\lambda_{12} = 1/4$.

It's the ratio of the partial derivatives of the utility function w.r.t. f_1 and f_2

$$\lambda_{12}(f) = \frac{\frac{\partial u}{\partial f_1}}{\frac{\partial u}{\partial f_2}}$$

A uniform MRS corresponds to a linear utility function $u(f) = w_1 f_1 + w_2 f_2$, and as such:

$$\lambda_{12}(f) = \frac{w_1}{w_2}$$

It represents the steepness of the indifference curve (slope).

Corresponding trade-off condition We denote as **corresponding trade-off condition the property**

$$\lambda_{12}(f_1', f_2')\lambda_{12}(f_1'', f_2'') = \lambda_{12}(f_1'', f_2')\lambda_{12}(f_1', f_2'')$$

Multiplying two MRSs on a diagonal equate to multiplying the values on the opposite diagonal.

A preference relation Π admits an additive utility function if and only if it enjoys both mutual preferential independence and the corresponding trade-off condition.

2.5 Mathematical Programming: *How to*

The general process for solving MP problems is:

1. Draw a graphical representation of the feasible region
2. Find nonregular points
3. Write the generalized Lagrangian function
4. Write the KKT conditions
5. Solve the system of conditions to reject candidate points, hoping that few remain (add nonregular points after this)

6. Evaluate the function in all the remaining points, choosing the optimum

Easy enough right? (It's not)

Let's use an example:

$$\begin{aligned} \min f(x) &= (x_1 - 1)^2 + x_2^2 \\ g_1(x) &= -x_1^2 - x_2^2 + 4 \leq 0 \\ g_2(x) &= x_1 - 3/2 \leq 0 \end{aligned}$$

Nonregular points All points in which the gradients of the active constraints are linearly independent are regular. Only active constraints must be considered, the whole feasible region is composed of regular points. A constraint is active when $= 0$.

Why are active constraints zero? The constraints are active only on the borders of the feasible region, since a optimal solution can only be found on that border, "pushing the boundaries" of the problem. Each point strictly inside the feasible region is next to another point, slightly better, slightly more towards the border.

Calculate the gradients of each constraint and check wether they can be zero or not. If the gradient can be zero and in such point:

- The constraint is inactive: business as usual
- The constraint is active: nonregular point, has to be added to the candidate set

In our example, the gradients are:

$$\begin{aligned} \nabla g_1(x) &= [-2x_1 \ -2x_2] \\ \nabla g_2(x) &= [1 \ 0] \end{aligned}$$

And:

- *The first one is 0 only in the origin, point in which the constraint is nonactive ($g_1(0,0) = 4$)*
- *The second one is never zero*

Then check points in which pairs of constraints are active, i.e., make a system in which both are zero (I think equality constraints always have to be added? I'll get back to you on that, not sure).

In our example:

$$\begin{cases} -x_1^2 - x_2^2 + 4 = 0 \\ x_1 - 3/2 = 0 \end{cases} \implies \begin{cases} x_1 = 3/2 \\ x_2^2 = 7/4 \end{cases}$$

From which we get the points

$$A = \left(\frac{3}{2}, \frac{\sqrt{7}}{2}\right), \quad B\left(\frac{3}{2}, -\frac{\sqrt{7}}{2}\right)$$

Check wether the gradients are linearly independent or not in the points found. To verify this the simplest way is to compose a 2×2 matrix with the values of the gradients considered in each point, if the determinant of such matrix is nonzero the gradients are linearly independent in the point considered.

In our example, $\nabla g_1(A) = \begin{bmatrix} -3 & -\sqrt{7} \end{bmatrix}$ *and* $\nabla g_2(A) = [1\ 0]$*, the resulting matrix being*

$$M = \begin{bmatrix} -3 & 1 \\ -\sqrt{7} & 0 \end{bmatrix}$$

Whose determinant is:

$$det(M) = (-3 \cdot 0) - (-\sqrt{7} \cdot 1) = \sqrt{7} \neq 0$$

So the gradients are linearly independent.

Then check points in which triples of constraints are active, similarly to earlier. You can guess how this goes on.

In our example there are no more constraints, but you would simply check that the gradients are linearly independent in the points resulting from the system given by $g_1(x) = 0$*,* $g_2(x) = 0$*,* $g_3(x) = 0$*.*

The aim of this phase is only to find nonregular points. Points with not linearly independent gradients are nonregular and as such have to be added to the candidate set after "sifting" with the KKT conditions.

Generalized Lagrangian function The generalized Lagrangian function is defined as:

$$\ell(x) = f(x) + \sum_{i=1}^{s} \lambda_i h_i(x) + \sum_{j=1}^{m} \mu_j g_j(x)$$

With λ_i free multipliers and $h_i(x)$ equality constraints (always active).

In our example, the function becomes:

$$\begin{aligned}\ell(x) &= f(x) + \mu_1 g_1(x) + \mu_2 g_2(x) \\ &= (x_1 - 1)^2 + x_2^2 + \mu_1 \left(-x_1^2 - x_2^2 + 4\right) + \mu_2(x_1 - 3/2)\end{aligned}$$

KKT Conditions The KKT conditions state that if a point is regular and locally minimal:

1. The partial derivatives of the Lagrangian function w.r.t. the x variables are equal to zero ($\partial\ell/\partial x_i = 0$)
2. The partial derivatives of the Lagrangian function w.r.t. the λ multipliers are equal to zero ($\partial\ell/\partial\lambda_i = h_j = 0$), that is, the equality constraints are respected
3. The product of the functions expressing the inequality constraints, times the corresponding multipliers are equal to zero ($\mu_k g_k = 0$)
4. All inequalities constraints are satisfied
5. All multipliers of the inequality constraints are nonnegative ($\mu_k \geq 0$)

These conditions allow to restrict the number of candidate points from *all regular and nonregular points* to *nonregular and a few regular.*

In our example, the conditions become

$$\begin{aligned}\partial\ell/\partial x_1 &= 2(x_1 - 1) - 2\mu_1 x_1 + \mu_2 = 0 \\ \partial\ell/\partial x_2 &= 2x_2 - 2\mu_1 x_2 = 0 \\ \mu_1 g_1 &= \mu_1(-x_1^2 - x_2^2 + 4) = 0 \\ \mu_2 g_2 &= \mu_2(x_1 - 3/2) = 0 \\ g_1 &= -x_1^2 - x_2^2 + 4 \leq 0 \\ g_2 &= x_1 - 3/2 \leq 0 \\ \mu_1 &\geq 0 \\ \mu_2 &\geq 0\end{aligned}$$

Solving conditions To solve the system without exhaustively exploring all possible cases one can use a search tree, whose nodes divide the feasible region in disjoint parts. To do that we build on the stricter conditions, i.e., the products $\mu_k g_k = 0$ (at first choose the simplest one). Given a constraint, we can distinguish two cases:

1. $\mu_k = 0$ and $g_k \leq 0$
2. $\mu_k > 0$ and $g_k = 0$

These assumption simplify the system, allowing it to be analyzed in an easier way. The process can be repeated on the resulting sub-problems, if necessary.

The idea is to restrict the possible solutions and find them a little at a time. Divide in "easy" sub-problems and all solutions to such problems are candidate points.

In our example, the constraint chosen is $\mu_2 g_2 = 0$. We now consider the case P^1, with $\mu_2 > 0$ and $g_2(x) = 0$, and thus $x_1 = 3/2$. The constraints now become:

$$\begin{aligned} 1 - 3\mu_1 + \mu_2 &= 0 \\ x_2(1 - \mu_1) &= 0 \\ \mu_1(7/4 - x_2^2) &= 0 \\ 7/4 - x_2^2 &= 0 \\ \mu_1 &\geq 0 \end{aligned}$$

We can now say that $\mu_1 = (\mu_2 + 1)/3 > 0$ and $x_2^2 = 7/4$. This yields candidate points:

$$A = \left(\frac{3}{2}, \frac{\sqrt{7}}{2}\right), \quad B\left(\frac{3}{2}, -\frac{\sqrt{7}}{2}\right)$$

Now consider P^2, with $\mu_2 = 0$ and $g_2 \leq 0$, the constraints become:

$$\begin{aligned} 2(x_1 - 1) - 2\mu_1 x_1 &= 0 && \implies x_1(1 - \mu_1) = 1 \implies \mu_1 \neq 1 \\ x_2(1 - \mu_1) &= 0 && \implies x_2 = 0 \\ \mu_1(-x_1^2 - x_2^2 + 4) &= 0 && \implies \mu_1(4 - x_1^2) = 0 \\ x_1^2 + x_2^2 &\geq 4 && \implies x_1^2 \geq 4 \\ x_1 &\leq 3/2 \\ \mu_1 &\geq 0 \\ \implies x_1 &= -2 \end{aligned}$$

This yields candidate point:

$$C = (-2, 0)$$

Choose optimum We now want to evaluate the function in all the candidate points (remember to consider nonregular points) and choose the best solution.

In our example:

$$\begin{cases} f(A) = 2 \\ f(B) = 2 \\ f(C) = 9 \end{cases}$$

Which implies that both A and B are globally optimal points.

3 Models with complex preferences

The preference relation now is not a weak order.

We denote as **Paretian preference** (in the case of costs) the relation

$$\Pi = \left\{(f, f') \mid f_l \leq f'_l, \text{ for each } l \in \{1, \ldots, p\}\right\}$$

Which is a **partial order**.

We denote as **dominated solution** a solution which is not better than w.r.t. each indicator of another solution, and strictly worse in at least one indicator.

We denote as **Paretian solution** every solution such that no other solution dominates it. We call **Paretian region** the set of all Paretian solutions.

Paretian solutions are not preferable to all other solutions and are not all reciprocally indifferent. Thus we want to identify the whole Paretian region

3.1 Identifying the Paretian region

3.1.1 Applying the definition

In the finite case, the Paretian region can be found by applying the definition through pairwise comparisons.

This is exact but slow.

3.1.2 Inverse transformation method

If the solution can be graphically represented, compute the inverse function $\phi : F \to X$ of $f : X \to F$, build the image of X in F through f, find graphically the nondominated impacts (lower left quadrant empty), find a parametric way to describe such impacts, transform them back using the inverse function.

This is exact, but human intervention is required and is limited to 2 indicators (*maybe* 3).

3.1.3 KKT conditions

KKT conditions can be extended to Paretian preference, by repeating the derivation with minor changes, obtaining a set that is usually larger than the Paretian region (finds an overestimate).

Not usable in discrete problems, as always.

3.1.4 Weighted sum method

Consists in building a linear combination of the indicators and optimizing it. The result is sufficient conditions for a point to be Paretian (underestimate of the region).

3.1.5 ϵ-constraint method

Replace all indicators but one with constraints that require the solution to respect a quality threshold and solve the auxiliary problem. The result is a necessary condition for a point to be Paretian.

It's needed to find parametrically how the Paretian region is described w.r.t. the variation of ϵ. It's just a constraint, so: when it varies, which solutions are feasible? Get the intervals for which the Paretian region doesn't change and optimize the non-replaced indicator for the feasible solutions.

This can be applied to any problem and provides an overestimate of the Paretian region, but it requires to consider all possible values of ϵ and find all globally optimal solutions, increasing the complexity of the problem.

3.2 Weak rationality methods

Decision-makers often can't estimate correctly, so let's just *embrace* that the pairwise comparison matrix is incorrect and the normalized utilities are incorrect.

3.2.1 AHP

The **Analytic Hierarchy Process** was introduced in 1980 based on the following criticisms:

1. The reconstruction of the single-variable normalized utility functions is subject to strong approximation errors

2. The estimation of the weights is subject to strong approximation errors when the number of attributes p is large

3. The various approximation errors combine in cascade

The method replaces absolute measures with relative ones, and quantitative ratios with qualitative scales, building a hierarchy of indicators to compare only conceptually similar quantities.

The preference among impacts is measured with an arbitrary qualitative scale, allowing to compare heterogeneous quantities, translating verbal judgments and building an evaluation matrix.

Humans find it difficult to compare nonhomogeneous things, so build an indicator tree and compare only siblings: leaves include elementary attributes, upper levels summarize them, getting progressively more general.