

Decision Methods and Models

Massimo Perego

Contents

I	Decision Problems	3
1	Introduction	4
1.1	Terminology	4
1.2	Modeling approach to decision	5
1.3	Why a formal approach?	6
1.4	Prescriptive and descriptive models	7
1.5	What makes a decision problem complicated?	8
1.5.1	Examples of complicated decision problems	9
2	Case studies	11
2.1	The tramway of Como	11
2.1.1	Context	11
2.1.2	The generation of the alternatives	13
2.1.3	The generation of scenarios	15
2.1.4	The definition and computation of the indicators	17
2.1.5	The definition of the stakeholders	19
2.1.6	The decision process	21
2.2	Reopening of the Navigli in Milan	22
2.2.1	The context	22
2.2.2	The definition of the alternatives	23
2.2.3	The definition of the scenarios	24
2.2.4	The definition and computation of the indicators	24
2.2.5	The time and space organization	25
2.2.6	The definition of the stakeholders	25
2.2.7	The decision process	25
3	Fundamental definitions and conceptual problems	27
3.1	Decision problem	27

3.1.1	Alternatives	28
3.1.2	Scenarios	29
3.1.3	Impacts	29
3.1.4	Impact function	30
3.1.5	Decision-makers or stakeholders	31
3.1.6	Preference relation	31
3.1.7	Property of binary relations	33

II Basic Decision Models 35

4 Structured preferences 36

4.1	Dominance relation	36
4.1.1	Decision-making on preorders	37
4.1.2	Decision-making on weak orders	38
4.2	Multi Attribute Utility Theory MAUT	41
4.2.1	Indifference Curves	42
4.2.2	Determining the utility function	44
4.2.3	Additive utility functions (game changer)	45
4.2.4	Preferential Independence	46
4.2.5	Marginal rate of substitution MRS	48
4.2.6	Additivity and MRS	50
4.2.7	Building an additive utility function	52
4.2.8	Determining the weights	54
4.2.9	The process in summary	55

5 Mathematical Programing 56

5.1	Basic concepts	57
5.1.1	Taylor's series expansion	58
5.1.2	Arcs	59
5.2	Necessary conditions for local optimality	60
5.2.1	A filtering approach	60
5.2.2	Feasibility condition	63
5.2.3	A first geometric interpretation	66
5.3	Karush-Kuhn-Tucker conditions	67
5.3.1	Farkas' Lemma	67
5.3.2	Second geometric interpretation	69
5.3.3	Standard form of Karush-Kuhn-Tucker conditions	69

Part I

Decision Problems

Chapter 1

Introduction

1.1 Terminology

A **system** is the portion of the world affected by the decision, it should allow different configurations, it should not be given once and for all, otherwise no decision would be possible.

A **solution** or **alternative** is the combination of all controllable aspects of the system, and **outcome** or **scenario** the combination of all its uncontrollable aspects.

A system combines controllable and uncontrollable aspects into a **configuration**. Each configuration is associated to an **impact**, that describes all aspects relevant for the decision.

Decision-maker or **stakeholder** refers to everybody who contributes to the choice of the alternative. The former indicates who takes part to the choice, while the latter also includes who does not participate but has interests at stake and could react to a disagreeable choice, exerting an indirect influence on the choice.

With **preference**, we denote the relative satisfaction between impacts.

A decision problem requires to *choose an alternative*:

- so as to *move the system into a configuration*
- such that *the decision-makers prefers its associated impact* to those of other configurations

- keeping into account that *the actual configuration depends on the alternative, but also on the scenario*

A **decision problem** implies two fundamental conditions:

- **Freedom**, i.e., availability of different choices (otherwise there is no decision)
- **Rationality**, i.e., the existence of preference criteria (otherwise the choice cannot be motivated)

This is a concept different from the definition of “decision problem” typically given in computer science (problem which admits only two possible solutions, *yes* or *no*). The decision problems considered here can be considered special cases of optimization/search problems, whose solution is an object with the maximum value (or minimum cost).

The focus is on practical decisions where a large amount of data must be taken into account, many choices are possible and the cost of a wrong choice is high.

We want to discuss *what make a decision complicated*, present the mathematical models to describe complicated situations and present the methods to deal with such situations, while recognizing limits and errors of this approach.

1.2 Modeling approach to decision

The modeling approach to decision requires a series of intermediate passages:

- Building a model of the problem
- Solving the model with algorithms, i.e., formal methods
- Interpreting the solution, with suitable methods

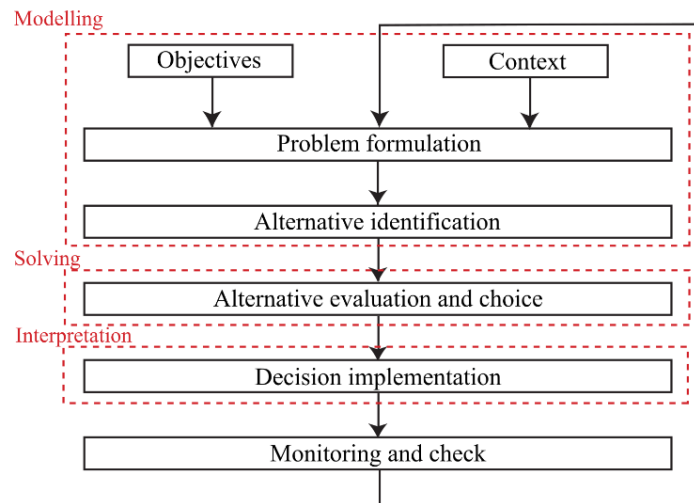
The strategy is to first **make a model**, then **compute**, and **finally decide**. From a problem to a model, solved by an algorithm, then actions are taken based on the solution.

The **decision process** occurs in an **iterative correction approach**:

1. *Problem formulation*: delineate the system, identifying impacts and preferences (Objectives) on one hand, decision-makers and scenarios

(Context) on the other

2. *Alternative identification*: define the set of feasible alternatives
3. *Alternative evaluation and choice*: evaluate the impact associated to each configuration (alternative and scenario) and choose an alternative based on the preferences of the decision-makers
4. *Decision implementation*: apply or simulate the alternative selected
5. *Monitoring and check*: observe the consequences of the decision; if unsatisfactory, make correction and repeat the process, introducing new scenarios, objectives, alternatives and evaluation methods



1.3 Why a formal approach?

A formal approach allows to:

- Predict in a more certain and precise way the impact of a decision, using descriptive models instead of intuition and experience
- Accelerate the decision process, using algorithms and information technology
- Consider a much larger number of possible alternatives
- Clarifying and certifying the decision process

- making explicit the assumptions made on alternatives, scenarios, preferences and decision-makers
- guaranteeing repeatability of the process
- allowing specific changes to the process, without starting from scratch

The formal approach is based on

- **Models**, to take decisions and predict their outcomes
- **Methods**, to build models, solve them, and interpret their results

1.4 Prescriptive and descriptive models

A decision model includes a number of sub-models, which can be classified based on the data they require and the result they produce. Decision models usually combine:

1. **Prescriptive models** that

- receive impacts and preferences in input
- return a suggested alternative in output

if this is the case, you should do that

2. **Descriptive/predictive** models

- receive the system, an alternative and a scenario in input
- return an impact in output

If you do this and if this happens, you will obtain that

The two families of models can have subtle and complex interactions. For example, any prescriptive model uses a set of descriptive models to obtain the impacts of possible alternatives and scenarios; on the other hand, some descriptive models can include prescriptive ones.

Example: a model prescribes a decision (*close or open streets*) based on models that describe a system (*amount of traffic*), including decisions prescribed by models (*satellite navigators*)

1.5 What makes a decision problem complicated?

A decision problem can be *complicated* due to:

- An **insufficient model** of the system
- **Complicating features** of the model, such as
 - complex preference structure, insufficient to define an optimum
 - uncertain environment, impact depends also on unknown scenario
 - multiple decision-makers, with potentially conflicting preferences
- A **computationally complex model**, everything is clearly defined, but no efficient algorithm is known to solve the problem

The three main complexity sources for decision problems are:

1. **Preference structure**: simple or complex
2. **Uncertainty**: a single scenario or many
3. **Decision-makers**: single or many

giving rise to $2^3 = 8$ families of decision problems.

We'll consider four basic families of prescriptive models:

1. Simple preference, a single scenario and a single decision maker:
 - mathematical programming
 - multi-attribute utility theory
2. Complex preference, a single scenario and a single decision maker:
 - paretian preferences
 - weak rationality models (*AHP* and *ELECTRE* methods)
3. Simple preference, multiple scenarios and a single decision-maker:
 - decisions in condition of ignorance (robust programming)
 - decisions in conditions of risk (stochastic programming)
4. Simple preference, a single scenario and multiple decision-makers:
 - independent decision-makers (game theory)
 - cooperating decision-makers (group decisions)

1.5.1 Examples of complicated decision problems

Before some large-size examples (Chapter 2), we present some realistic situations that might suggest what elements make a decision complicated in practice.

The search for parking: In a rather congested town, we are looking for a parking to leave out car and reach the place of an important meeting; we would prefer to park quickly and not walk for long to reach the destination.

- *System* is the local street network, with the set of all potential parking places
- *Alternative* is every possible trajectory of the car (path and time)
- *Scenario* is every possible distribution of the free parking spaces over space and time
- *Impact* are the driving time and walking time after parking
- *Decision-maker* is the driver

The impact reduces to the aspects which actually matter for the decision, but it depends on a complex combination (configuration) of every possible trajectory of the car and distribution of parking spaces (alternative and scenario).

Thermostat regulation: We want to tune the classroom's thermostat so that the temperature is pleasant for the teacher and students.

- *System* is the classroom
- *Alternative* is the position of the thermostat
- *Scenario* is the external temperature and the exposition of the classroom to the sun
- *Impact* is the internal temperature of the classroom (but probably, also the humidity)
- *Decision-makers* are all the people in the classroom (or just the teacher?)

This problem is easier than the others, but the definition of impact and decision-makers is not trivial.

Buying a car: We want to buy a car with good performance, comfort, design and low cost throughout its life cycle.

- *System* is the local market of cars, petrol, repairs, etc.
- *Alternative* is the car bought
- *Scenario* are the stock and prices of the car dealers, the occurrence of accidents, prices of petrol and repairs, etc.
- *Impact* are the characteristics of the car throughout its life cycle
- *Decision-maker* is the buyer (possibly some other family members/friends?)

A possible alternative, easy to forget, is to not buy a car, with an impact very bad for some aspects, but very good for others.

Risiko round: We want to play a round of Risiko, considering in turn all of the players.

- *System* is the map, with the distribution of territories, armies, cards
- *Alternative* are the territories from which and to which each attack is launched and the corresponding number of attacking and defending armies
- *Scenario* is the dices' outcome at each attack
- *Impact* is the number of armies destroyed for each player
- *Decision-makers* are the players

Here the presence of multiple decision-makers is intrinsic and unavoidable, moreover they all make their own choices autonomously.

Chapter 2

Case studies

We want to present two case studies with the aim of clarifying the concept of complicated decision problems and to understand the practical meaning of the terminology presented earlier.

Both case studies concern large public works and could undergo several variations and corrections, as they are not yet implemented.

2.1 The tramway of Como

2.1.1 Context

Como has been for centuries at the center of exchanges, connecting Italy to Switzerland and Germany, while being a connection to cities at the end of the alpine valleys (Varese, Lecco, Bergamo, etc.).

This favorable position has led to problems, mainly related to saturation of the access ways, with congestion, economic losses due to slow transfers, pollution and accidents as a result.

Private cars are the main responsible for these problems, they concentrate on a few relatively small streets, which can't be enlarged easily due to the orography of the site.

A possible improvement could be to use public transport, more efficient from a transportation capacity point of view. However, the fraction of transfers serviced by public transport is small, and very uneven. Buses service a minority of users and are subject to the same limitations as private cars.

As for train services, the state railway is rather unloaded (FS line), while the regional railway is nearly saturated (FNM line). How can the capacity of the FNM line be increased? Maybe while moving some of its passenger on the FS line.

Some critical points to consider while discussing interactions among the lines:

1. The crossing between the lines occurs with an acrobatic system of bridge: the Napoleona street overcomes the FNM line, which overcomes the FS line, which overcomes a small local street (Via dei Mulini)
2. Between the stations of Camerlata and Borghi, the FNM line has a single track, with steep slopes on both sides
3. Between the stations of Borghi and Lago, the FNM line has a single track, running along a narrow alley between houses
4. Joining the FS and FNM lines through the town plain would require first to cross either the walls (in a historical area) or along the lakefront (touristy, occupied by a main street and flooded from time to time), and then climb the steep slope on top of which lies the FS station of Como S. Giovanni

All these factors make a solution based on railways impossible to adopt, unless limited to the use of existing tracks. Increasing the frequency of the trains on the FNM line, however, has two problems:

- The single-track line forbids two trains to run along the Camerlata-Borghi-Lago track (unless perfect synchronization, difficult and prone to big problems)
- The levels crossing cut the town in two for several minutes every time a train passes along them, which could lead to significant traffic congestion

Hence the idea of a tramway line, which could:

- Have a double track, thanks to the shorter gauge
- Have fast crossing with traffic lights, instead of slow level crossings
- Be prolonged to the inside of town

But a classical tramway involves a big problem, as it requires to replace the whole track, therefore

- forcing train passengers from outside Como to change at an external station
- completely blocking traffic during the works

2.1.2 The generation of the alternatives

In this case, the set of alternatives is far from being clear. More correctly, it is *indefinite*, and must be built step by step by investigating the whole context.

We need to follow the iterative approach defined in Section 1.2. For instance, if one realizes that modifying the gauge of the tracks is too expensive, one can wonder if there exists tramways able to use standard train tracks. **They exist.**

The process of generating alternatives becomes easier if one decomposes the problem, identifying the basic *alternative elements*, which can be (at least partly) independent.

In this study, three fundamental elements have been identified, in order to base on them the generation of alternatives:

1. The technology used to implement the new line
2. The route of the new line
3. The management of the FNM trains with respect to the new line

Technology Previous studies on similar cases suggest three possible technologies:

1. Standard railway service: add a shuttle-train between Grandate e Como Lago, so as to increase the frequency of the current railway service
2. Tramway: replace the railway with a classic tramway
3. Interoperable: add to the current railway service a special tramway which is able to work with both systems (facing all the resulting problems)

Route The number of possible routes is huge and not well-defined *a priori*. To start, four possible routes have been identified:

1. Keeping the current FNM route, ending at the station of Como Lago
2. Linking the station of Como Lago (FNM) to Como S. Giovanni (FS) through new tracks crossing the town center
3. Linking the station of Como Lago (FNM) to Como S. Giovanni (FS) through new tracks passing along the lakeside
4. Building a ring around the town center through a new track from Como Lago (FNM) to Como S. Giovanni, plus a track from Como S. Giovanni (FS) to Como Borghi (FNM)

Management of the FNM trains This element of the alternatives describes the relation between the two rail transportation systems:

1. Keep the FNM service on its current route
2. Move the FNM service to the FS station of Como S. Giovanni, through the construction of an interchange between the FNM and FS lines (probably in the station of Camerlata)

These are extreme possibilities, with a range of intermediate solutions, either during the works or permanently.

Each of these aspects provides an *alternative element*. The single possible choices for an element can be associated to numerical values: sometimes a proper quantitative measure, other times simple arbitrary indices (e.g., 0 for the railway, 1 for the tramway and 2 for the interoperable). The purpose of this association, in general, is to describe effectively the set of alternatives, not to make computations.

Once the elements of the alternatives and the possible values for each element have been identified, one can proceed to enumerate their combinations, which in general are not all feasible. In the present case there are $3 \cdot 4 \cdot 2 = 24$ combinations, only 7 of which are feasible. The other ones can be excluded due to obvious contradictions, technical impossibility or common sense.

For example, the railway technology is incompatible with any route different from the present one (trains cannot go through towns).

A fundamental remark on the generation of alternatives is that **there always exists an alternative that consists in doing nothing** and keeping the current situation, this is conventionally denoted as **alternative zero**.

New alternatives should, at least, not worsen the current situation. This is far from trivial since “*worsen*” is a complicated concept.

In this discussion, we have neglected several other potential elements of alternative, such as

- The location of interchange parking (in various sites)
- The implementation of double-tracks rails along parts of the current single-track rail
- The construction of new stations along the current route and possible new routes
- Possible transfers of the current route
- The extension of the service out of Como, beyond the station of Granda
- The frequency of the new service
- The tariffs of the new service
- ...

In particular, some mobility studies suggest that there could be a strong request for a railway servicing parts of the province currently badly connected to Como, both eastward and westward. Other subsequent feasibility studies have focused on these aspects, rejecting the most extreme solutions (current route and town ring), merging the intermediate ones (center crossing and lakeside route) in a mixed variant.

We also neglected the possible *mitigation measures*, all those interventions that must be added to the alternatives in order to correct the negative impacts they produce, or to convince stakeholders strongly opposing the project to accept it (or at least not fight it too much), or to modify the impact of the work (e.g. limited access zones to discourage private cars).

All this confirms that even listing the possible solutions of a decision problem can be a preliminary problem in itself, that is not to be solved in a single phase, and actually is never ultimately solved.

2.1.3 The generation of scenarios

The generation of scenarios follows the same rationale as the generation of the alternatives, also requiring to predict events uncontrollable by the

decision-makers. First, one identifies the *scenario elements* and their possible values, then we consider only the feasible combinations. If new scenarios emerge, or old ones disappear, the study should be updated and corrected.

The scenario elements considered are the following ones:

1. Closure of the lakeside to private traffic: forbidding the access of vehicles to the street that runs between the city center and the lake, this interacts with the project as it frees up the space used by cars, making it easier to build a tramway along the lakeside
2. International Como-Chiasso station: it consists of merging the international railway stations of Como (Italy) and Chiasso (Switzerland) in a single stop, half-way, accelerating the trips between Italy and central Europe, this interacts with the project because the station of Como S. Giovanni would be reduced to the level of local station, introducing the need for a service taking travelers to Como S. Giovanni and possibly onward to the new international station
3. Anti-flood barriers: it consists of building barriers that protect the center of Como from lake floodings, this interacts because a tramway service passing along the lakeside would be blocked by floods
4. Borgovico tunnel: a toll tunnel that should run from the north-west to the south-west of Como, allowing to cross the town without congesting the western side
5. Underlake tunnel: a tunnel which should run under the lake to replace the street that currently connects the north-west and north-east of Como

In theory, each combination of these elements is possible, excluding combining the underlake tunnel with keeping the lakeside open (meaningless) and building both tunnels (too expensive).

This discussion neglects important aspects such as:

- The anticipated variations in the residential and economic structure of the area under study
- The variations in the *origin-destination matrix* (*O/D matrix*) of the potential transportation demand (i.e., the number of trips from place to place that will take place in the town and could be captured by the new service)

- The amount of European, state and region financing

Some of these scenario correspond to predictable numerical values, while others correspond to event which can either occur or not, or could occur in different ways.

For each numerical value, for each event or for each way in which an event can occur, it could be possible to estimate a probability, or at least provide a qualitative estimation of its likelihood. All this information concurs to describe the set of possible scenarios and its auxiliary information.

2.1.4 The definition and computation of the indicators

As the alternatives and the scenarios, also the impacts are combinations of different elementary quantities, which are usually named *indicators*. They characterize the satisfaction of the decision-makers for a configuration of the system, and determine their preferences. Indicators are typically much more numerous and diversified than the elements of the alternatives and the scenarios, so numerous that their generation adopts a hierarchical process that progressively details the sectors of the impact:

- Identify general macrosectors
- Each macrosector is decomposed into sectors, and possibly into sub-sectors, progressively more specific
- The elementary indicators are identified

This mechanism produces an *indicator tree*.

In this case, the three standard macrosectors used in Environmental Impact Assessment (EIA) have been adopted:

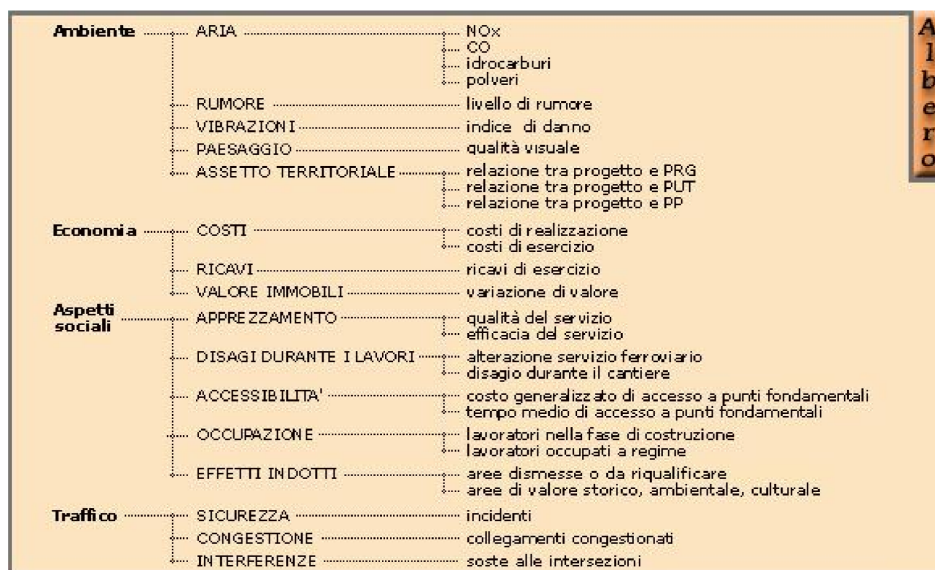
1. *Environment*, subdivided into Air pollution, Noise, Vibrations, Landscape and Territorial structure (that is, the compatibility with the current management plans of the area)
2. *Economy*, subdivided into Costs, Revenues and House values
3. *Society*, subdivided into Appreciation, Discomfort, Accessibility, Employment, Induced effects

but a fourth macrosector has been added

4. *Transportation*, subdivided into Security, Congestion and Interferences

This last one is usually included in the Society macrosector, but it has been attributed an autonomous role as the study evaluates a large transportation service.

This yields the indicator tree:



The hierarchical structure aims to organize the evaluation of the single indicators during the phase of data collection; experts in each sector will be assigned to lead the estimation, prediction and measurement of the values of each indicator.

Similarly, the construction of preferences in the decision phase will be organized based on the specific skills of experts of the various sectors; the “importance” of each indicator has to be determined.

Each indicator must be associated with a tool which will provide its values for each possible configuration. In most cases, these values will not be directly measurable, but are the result of estimates, given by a predictive model (e.g. CO_2 pollution can be measured now, but obviously not in all the hypothetical configurations).

The indicators can be numerical values that express physical measures, but also qualitative values on an ordinal scale. The purpose is not to make computations, just to describe the situation. In practice, some evaluation methods use such numerical values to make computations.

Time decomposition (phases) and space decomposition (zones)

Some indicators have an intrinsic *space* or *time* feature, which must be taken into account; in other words, the values of an indicator might be meaningless if referred to the whole territory or whole time horizon of the project.

To correctly describe the situation, the system has to be divided into distinct geographical zones, on which is meaningful to define a value of the considered indicator.

This increases the amount of information to manage and requires to evaluate the relative importance of the different zones on the decision, as if they were values of different indicators.

In this study, Como has been subdivided into 5 zones:

- Como centro
- Borghi
- Camerlata
- Lora
- Tavernola

Only the first three zones are directly affected by the project, whereas Lora and Tavernola have been maintained to check indirect effects on the whole town.

2.1.5 The definition of the stakeholders

We denote as *stakeholders* every person, organization, category or institution that is involved, directly or indirectly, in the project, even if it plays no role in the decision, since its interests are affected by the project.

The fundamental aspect in the description of stakeholders is to characterize their *preferential structure*. However, the stakeholders could also point out neglected alternatives, scenario elements and indicators.

For this case, four classes of stakeholders have been identified:

1. Institutional stakeholders (Municipality, Province, Region)
2. Society
 - Citizens

- Environmentalists associations
 - Category associations
3. Users
 - Regular (commuters, students, etc.)
 - Irregular
 4. Transportation companies
 - FNM, whose track between Grandate and Como Lago should be used by the new service
 - SPT, the local public transport company managing the bus lines, whose demand would be disrupted by the new service, requiring revision of routes and schedules
 - FS, whose station of Como S. Giovanni could host part or all the trains currently servicing the FNM line

Each stakeholder has limited interests and skills, often related to one or few subtrees of the indicator tree, hence it will usually be involved mainly in the phases of the decision process which concerns those specific indicators.

For this case, the relation between stakeholders and subsectors:

1. The local government stakeholders will be interested in the whole system of indicators
2. Concerning society stakeholders:
 - The citizen will probably focus their interest on the Environment, Society and transportation macrosectors
 - The environmental associations will focus on the environmental macrosector
 - The category associations (Camera di Commercio, Industria, Artigianato e Agricoltura, Confesercenti, ...), who support the economic interests of the traders and entrepreneurs, will focus on the Economy macrosector and on the discomfort introduced by the works
3. The users will focus on the society macrosector, in particular on quality and accessibility of the service

4. Concerning the transport companies

- FNM will focus on the Society (Employment and Discomfort) and Economy (Costs and Revenues) macrosectors
- SPT will focus on Society, Economy and Transportation macrosectors (Employment, Discomfort, Traffic congestion and its own Costs and Revenues)
- FS is interested in the management of the train service potentially moved to the station of Como S. Giovanni

Once stakeholders have been identified, a reference person (or group) should be defined for each stakeholder, and involved in all phases of the decision process and subsequent iterations.

The tools to interact with the person depend on the specific group of stakeholders considered; such tools could be talking with representatives of the group or polls, forms and interviews to the relevant subjects.

Some group of stakeholders can partly overlap (e.g. environmentalists and shopkeepers are also citizens), but the preference of common citizens will be different from those of a specific group, making it correct to represent them as different stakeholders.

2.1.6 The decision process

For each possible configuration of the system it's necessary to evaluate each indicator, sometimes requiring physical measurements, but most of the time a suitable descriptive model can be applied to compute the expected value. Speaking in qualitative terms:

- The alternatives exploiting the railway technology will have
 - low implementation costs and low discomfort
 - limited effect on the accessibility of sites not currently serviced
 - limited environmental improvements
 - nearly no impact on the employment
 - ...
- The alternatives exploiting the “interoperable” technology will have
 - higher economic costs

- stronger penetration of the service in the urban fabric, dependent on the chosen route
- intermediate discomfort
- ...
- The alternatives exploiting the classical tramway technology will have
 - much higher implementation costs and times
 - very detailed service
 - ...

The resulting numerical and qualitative values will have to be aggregated depending on the preferences of the various stakeholders. It's necessary to model the preferential structure of each stakeholder, to determine preference among different impacts.

It is also necessary to aggregate the preferences of different stakeholders in a final ordering of the alternatives, or at least in the selection of one alternative. The *sensitivity* of the final ordering should also be taken into account, that is the boundaries within which it remains valid when the data changes.

2.2 Reopening of the Navigli in Milan

2.2.1 The context

Milan was born as a river town, rich in natural rivers, later further enriched by a network of canals. Modernity led to covering the whole network in favor of traffic and, more recently, to a periodic proposal and discussion of project to reopen part of it.

The main motivations for the project are:

- The *hydraulic continuity*, that is the possibility to better control and limit the flooding of water streams in the northern part of Milan, currently flowing through pipes, limiting the capacity and making cleansing more difficult
- The *touristic and commercial navigability*, with their social, cultural and economic outcomes

- The *production of electricity*, not by directly installing turbines in the Navigli, but with the increase of water flow in the Naviglio Pavese, which has several small hydroelectric plants south of Milan
- The *feeding of heat-pump plants* in the area of Darsena using nappe (underground) water
- The *decrease in the nappe level*, thanks to the extraction of water from wells which are currently closed, reducing the use of pumps in several underground stations that defend the tunnels from flooding

2.2.2 The definition of the alternatives

The definition of alternatives seems easier in this case, given the Navigli's well defined historical route. In practice, things are not easy.

Starting from a basic dichotomy, one can distinguish:

- “Virtual” alternatives, in which the Navigli are not actually reopened, but made enjoyable with public visual aids (sings, pavings, ...)
- “Physical” alternatives, in which the Navigli are actually reopened

All physical alternatives do not share the same route, there are many alternatives:

- A partial reopening, keeping some covered stretches
- A complete reopening of the classic circle north-east-south, which includes a number of minor variations concerning
 - Porta Nuova park
 - Cavour Square
 - the Vallone Naviglio
 - the Viarenna basin
- The reopening of the whole inner circle, also on the western side: some technical reasons make this alternative quite impractical (mainly, it interferes with underground line 2)
- Additional works with respect to the historical route:
 - the Vettabbia canal

- the “Darsena 2” project, a brand new basin near the railway station Porta Genova

Another fundamental element of the alternatives is navigability: the new canal could be open or closed to boat trips. The elements of alternative (route, navigability and other possible ones) must then be combined to build the alternatives, removing impossible combinations (virtual alternatives obviously forbid navigability).

2.2.3 The definition of the scenarios

This point is not developed in the project described, we therefore make a superficial analysis. The fundamental aspects of the scenario are:

- Availability of public funding to build and manage the canal
- The (more or less) strict closure of the city center to private cars
- The construction of underground line 4 (completed as of writing) which would provide easy access on foot to most of the city center, even if private cars and some bus lines could not access it

2.2.4 The definition and computation of the indicators

We, once again, proceed in a hierarchical way, to identify large sectors and then dividing them. To give a non-exhaustive list:

- Impact on pollution
- Variation of travel times to various parts of town
- Impact on traffic congestion
- Impact on public transport
- Construction costs
- The hedonic impact on the real estate prices
- Impact on commercial activities

Some of these impacts could be so negative as to introduce a feedback loop on the definition of the alternatives, a correction and update of the alternatives through the introduction of devices to reduce those impacts. E.g. bridges used by tramway can't have slopes, which implies a water level low enough and boat suitably designed in order to allow navigability.

2.2.5 The time and space organization

Several indicators are naturally related to geographical positions. Some indicators refer to zones, whereas others to linear stretches of the Naviglio. The detailed project divides the route of the canal into 16 stretches.

Concerning the time phases, some of them correspond, somehow, to alternatives, meaning it's possible to implement a specific phase then leave the following "frozen" for a long time, waiting for favorable conditions to complete them. For example, virtual alternatives can be seen as a preliminary phase to a global reopening. They could achieve a partial result at low cost, however they could increase the overall cost.

The construction of a double Naviglio, with an underground pipe running under the canal is an additional cost, but it also allows to reach in short time the result of hydraulic continuity and to spread over time the following phases, dividing the expense.

2.2.6 The definition of the stakeholders

The definition of the stakeholders is similar to what described in for the tramway of Como (2.1.5). The mayor of Milan can be considered as the decision-maker, but other subjects must be taken into account, such as transportation companies (mainly ATM, railway lines are not affected), the environmental and category associations, citizens. Among the citizens, the owners of real estate along the route of the canal could be considered specifically.

2.2.7 The decision process

Once again, it's necessary to evaluate each indicator in each possible configuration. This implies the application of suitable descriptive models. One could expect that

- The "virtual" alternatives will be particularly cheap and have negligible impact on traffic and pollution, but will not offer any advantage with respect to hydraulic continuity (unless combined with the construction of an underground pipe), to the production of electricity and to the management of water nappe, and they have limited impact on tourism and commerce; the aesthetic impact is harder to evaluate
- The "physical" alternatives will be much more expensive, reaching the objective of hydraulic continuity, electricity production, management

of water nappe, while favoring tourism, commerce, real estate prices and aesthetic impact

Moreover

- The alternatives involving a navigable canal will have touristic, commercial and accessibility advantages, but at a higher cost, requiring to manage a system of gatehouses to get over the height difference between subsequent stretches of the canal
- The alternatives involving a non-navigable canal will be cheaper, but less advantageous for accessibility, commerce and tourism, even if they could have a better aesthetic impact, since a navigable canal requires a relatively low water level to let boats pass under bridges

After all indicators have been evaluated, the aggregation phase can start, taking into account the preferential structure of all the stakeholders. This can be done in several ways and should conclude with an ordering of the alternatives, or at least the choice of one, also with the information on the sensitivity on such an ordering or choice.

Chapter 3

Fundamental definitions and conceptual problems

While analyzing the two case studies, we remarked that

- The *alternatives* are generated combining elements of alternative and there is always an alternative zero (leaving the system as is)
- The *scenarios* are generated combining elements of scenario
- The *impacts* are generated combining indicators and indicators are organized into a hierarchy
- Each configuration generates an impact through descriptive models
- Decision-makers should never ignore stakeholders (people affected by the decision and can react to it)
- Preferences are expressed on impacts (not on configurations)

3.1 Decision problem

Definition 1. A *decision problem* is defined by a 6-uple

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

where:

- X is the **feasible region**, the set of all alternatives;

- Ω is the **sample space**, the set of all possible scenarios, or outcomes
- F is the **indicator space**, the set of all possible impacts
- $f : X \times \Omega \rightarrow F$ is the **impact function**
 - every pair $(x, \omega) \in X \times \Omega$ describes a configuration of the system on which a decision should be taken
 - f associates to each configuration (x, ω) of the system an impact $f(x, \omega) \in F$
 - impact $f(x, \omega)$ describes all the features of the configuration (x, ω) that are relevant for the decision (costs, profit, quality levels, ...)
- D is the **set of all decision-makers**, or stakeholders
- $\Pi : D \rightarrow 2^{F \times F}$ is the **preference function**, a function that associates to each decision maker $d \in D$ a subset of impact pairs $\Pi(d) \subseteq F \times F$; this is interpreted as a binary relation representing the preference of decision maker d

The aim of the problem is to identify a solution $x^* \in X$ ir a subset of solutions $X^* \subseteq X$ that the decision-makers consider satisfactory based on their preferences between the impacts $f(x^*, \omega)$ with $\omega \in \Omega$ and the other impacts $f \in F$

3.1.1 Alternatives

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

We'll assume that

$$X \subseteq \mathbb{R}^n \Leftrightarrow x = [x_1 \dots x_n]^T \text{ with } x_i \in \mathbb{R} \quad \forall i \in \{1, \dots, n\}$$

with a finite number n of elements of alternative or decision variables. Real number can model several different situations (e.g. binary, enumerative, qualitative, quantitative, functions within families).

Several **kinds of feasible regions** X exist

- Continuous
- Discrete
 - Infinite

– Finite

- * Combinatorial (typically $|X| \in \Omega(d^n)$ for some $d > 1$)
- * “Finite” (typically $|X| \in O(1)$)

Focusing on a specific kind allows algorithms which are more effective and less general.

The notation is, so far, purely descriptive and not used to make computations.

3.1.2 Scenarios

The scenarios formally describe the events out of the control of the decision-makers that have non-indifferent effect on the system. The set Ω defines all such events during the time horizon relevant to the decision. Scenarios also describe modeling errors (disturbances).

We’ll assume

$$\Omega \subseteq \mathbb{R}^r \Leftrightarrow \omega = [\omega_1 \dots \omega_r]^T \text{ with } \omega_k \in \mathbb{R} \quad \forall k \in \{1, \dots, r\}$$

with a finite number r of elements of scenario or exogenous variables.

3.1.3 Impacts

The impacts formally describe all aspects that are relevant for the decision, i.e., decisions are taken based on the impact.

The impacts are described quantitatively as vectors of real numbers. We’ll assume

$$F \subseteq \mathbb{R}^p \Leftrightarrow f = [f_1 \dots f_p] \text{ with } f_I \in \mathbb{R}$$

with a finite number p of

- indicators
- criteria
- attributes
- objectives

3.1.4 Impact function

The impact function $f : X \times \Omega \rightarrow F$ is a vectorial function that associates each configuration (x, ω) to an impact $f(x, \omega)$.

Trivial example

- buy quantities x_i of products (n decision variables)
- Pay costs ω_i ($r = n$ exogenous variables)
- The total (mono-dimensional $p = 1$) cost is $f(x, \omega) = \sum_{i=1}^n \omega_i x_i$

The impact function can be

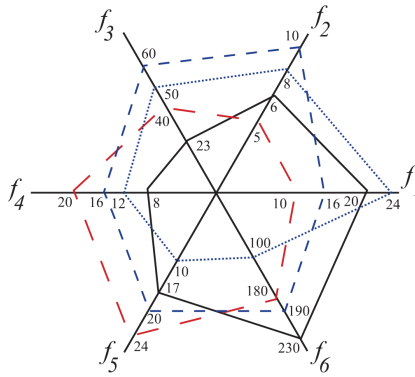
- a mathematical expression that describes a computation
- a simulation or integration software
- an empirically generated graph or table

Two classical representations for finite problems are

1. The **evaluation matrix**

$f(x, \omega)$	f_1	f_2	f_3	f_4	f_5	f_6
(x', ω')	10	5	40	20	24	180
(x', ω'')	16	10	60	16	20	190
(x'', ω')	20	6	23	8	17	230
(x'', ω'')	24	8	50	12	10	100

2. The **radar chart**



3.1.5 Decision-makers or stakeholders

A decision-maker is whoever takes part in a decision. The subjects who do not directly contribute to a decision, but play a role in it with their preferences are often denoted as stakeholders.

D is a finite set which includes all stakeholders. They (directly or indirectly) set the value of the decision variables x

- independently

$$X = X^{(1)} \times \dots \times X^{(d)}$$

- cooperatively

$$X = X^{(1)} \cap \dots \cap X^{(d)}$$

3.1.6 Preference relation

For each decision-maker, there is the need to distinguish *which impact is preferred*, i.e., preferring $f \in F$ to $f' \in F$ means *considering acceptable the replacement of f' with f* . This corresponds to establishing a relation between pairs of impacts.

$\Pi : D \rightarrow 2^{F \times F}$ associates each decision-maker $d \in D$ to a subset of impact pairs that represents a binary relation $\Pi(d)$ on F .

Subset $\Pi(d) \in 2^{F \times F} \Leftrightarrow \Pi(d) \subseteq F \times F$ collects the specific pairs of impacts between which d has a weak preference

$$\Pi_d = \{(f, f') \in F \times F \mid d \text{ weakly prefers } f \text{ to } f'\}$$

Definition 2. A *weak preference* is when d accepts to cede f' for f (not strict)

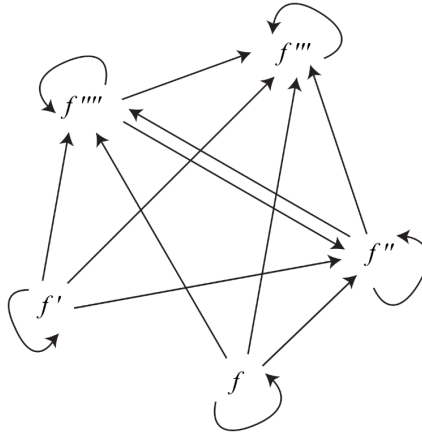
$$f \preceq_d f' \Leftrightarrow (f, f') \in \Pi_d$$

In the finite case, a preference relation can be represented by

1. an **incidence matrix** (rows and columns are impacts, the ones are preferences)

	f	f'	f''	f'''	f''''
f	1	0	1	1	1
f'	0	1	1	1	1
f''	0	0	1	1	1
f'''	0	0	0	1	0
f''''	0	0	1	1	1

2. a preference graph



Derived relations

From the weak preference relation, one can derive

- **Indifference** relation ($\text{Ind}_\Pi = \Pi \cap \Pi^{-1}$), definition:

$$f \sim f' \Leftrightarrow f \preceq f' \text{ and } f' \preceq f$$

The decision-maker accepts the exchange in both directions

- **Strict preference** relation ($\text{Str}_\Pi = \Pi \setminus \Pi^{-1}$), definition

$$f \prec f' \Leftrightarrow f \preceq f' \text{ and } f' \not\preceq f$$

The decision-maker accepts the exchange only in the given direction

- **Incomparability** relation ($\text{Inc}_\Pi = \bar{\Pi} \cap \bar{\Pi}^{-1}$), definition

$$f \bowtie f' \Leftrightarrow f \not\preceq f' \text{ and } f' \not\preceq f$$

The decision-maker rejects the exchange in both directions (the decision-maker is not indifferent between the impacts, but is unable or unwilling to chose between them, refusing both exchanges)

3.1.7 Property of binary relations

Not all binary relations express reasonable preference relations for a decision-maker. In principle, the preference relation should be

- *Realistic*, it models correctly the aims of the decision-maker
- *Effective*, it allows algorithms that yield a satisfactory choice

The two are, surprisingly, often conflicting.

Some preference relations enjoy some other special properties

- **Reflexivity:** $f \preceq f$, $\forall f \in F$; it would be meaningless to consider a decision-maker unable to compare an impact to itself
- **Antisymmetry:** if $f \preceq f'$ and $f' \preceq f$, then $f = f'$ for all $f, f' \in F$; two impacts are indifferent only if they are identical
- **Completeness:** if $f \not\preceq f'$, then $f' \preceq f$ for all $f, f' \in F$; it requires the decision-maker to be able to always sort the impacts, though allowing ties; this excludes incomparability
- **Transitivity:** if $f \preceq f'$ and $f' \preceq f''$, then $f \preceq f''$, for all $f, f', f'' \in F$; discussed later

While building the model, we must investigate whether they hold or not, because they could be

- useful to simplify the computation
- but unrealistic for the given problem

Main kinds of preference relation Combining these properties, we can identify different kinds of preferences:

- A **preorder** enjoys reflexivity and transitivity
- a **partial order** enjoys reflexivity, transitivity and antisymmetry; example: subset inclusion
- a **weak order** enjoys reflexivity, transitivity and completeness; example: rankings in sport
- a **total order** enjoys reflexivity, transitivity, antisymmetry and completeness; example: number sizes

Conceptual problems on transitivity Transitivity is the most common assumption for a preference relation, and is often interpreted as an equivalent for rationality.

Rational decision-makers are transitive, humans are not. There are two main approaches to deal with this point:

1. *humans should be transitive*, i.e., use thought experiments to show the evil consequences of giving up transitivity
2. *let humans be non-transitive*, i.e., appeal to practical and thought experiments and adopt weaker models of rationality

Part II

Basic Decision Models

Chapter 4

Structured preferences

How can we use the concept introduced to choose a solution for a decision problem?

We'll assume:

- A simple preference relation Π
- A certain environment $|\Omega| = 1 \implies f(x, \bar{\omega})$ reduces to $f(x)$
- A single decision-maker $|D| = 1 \implies \Pi_d$ reduces to Π

Single scenario and single decision-maker. Also, the preference relation of the decision-maker will be a weak order (reflexive, transitive and complete).

The first two conditions allow a well-posed definition of solutions that can be justifiably selected and the third allows the choice of a single solution.

4.1 Dominance relation

The preference relation between impacts ($\Pi \subseteq F \times F$) projects onto an **induced relation between solutions**. A solution *dominates* another when the impact of the former is preferable to the impact of the latter.

$$x \preceq x' \Leftrightarrow f(x) \preceq f(x') \quad \forall x, x' \in X$$

Definition 3. We denote as **dominated solution** a solution $x \in X$ such that $\exists x' \in X : f(x') \prec f(x)$, **nondominated solution**. We denote as $X^* \subseteq X$ the set of **all nondominated solutions**.

This implies a partition of the feasible region into

- dominated solutions: $x \in X$ such that $\exists x' \in X : x' \prec x$
- nondominated solutions: the other ones

Reflexivity looks natural in a preference relation. When solving a decision problem, it is also rather natural to

- **exclude dominated solutions**, that is choose $x^* \in X^*$
- Choose an **arbitrary solution from a set of mutually indifferent ones**

But this conflicts with some possible situations:

- All solutions in a strict dominance circuit would be removed
- Two solutions might be indifferent with respect to a third one, but incomparable with each other

Transitivity solves both problems. **Preorders** are strong candidates to be **preference relations**.

4.1.1 Decision-making on preorders

Theorem 1. *If the preference relation Π is a preorder, the induced dominance is a preorder.*

Theorem 2. *If the preference relation Π is a preorder (reflexive and transitive) and the solution set is finite and nonempty ($X \neq \emptyset$), the nondominated solution set X^* is nonempty.*

This guarantees that decision problems admit reasonable solutions. In some cases, X^* will contain a single solution, or several mutually indifferent, in other cases it will include incomparable solutions. Therefore, determining X^* simplifies the problem, but doesn't solve it completely.

Theorem 3. *If preference Π is a preorder and X^* is nonempty, the non-dominated solutions partition into disjoint components*

- *they are mutually indifferent within each component*
- *they are mutually incomparable between different components*

\implies *if there is only one component, the problem is solved (requires completeness).*

Identification of the nondominated solutions

If X is a finite set, it's possible to build the **strict preference graph**, whose nodes correspond to solutions, while the arcs correspond to solution pairs whose impacts are related by a strict preference. In this graph there are no indifferent pairs.

The nondominated solutions correspond to *nodes with no ingoing arc* (from each node there's an arrow towards dominated solutions), to identify them it's sufficient to scan each node of the graph and search for nodes with zero indegree.

Let $O(\gamma)$ be the complexity of computing the preference between two given impacts (non necessarily constant, they must be computed and compared), then the overall complexity for the search is $O(\gamma|X|^2)$ (for each node, compare its impact to the one of every other in time γ).

This obviously can't be applied to infinite sets and could be impractical even in the case of combinatorial sets.

4.1.2 Decision-making on weak orders

Theorem 4. *If preference Π is a weak order (reflexive, transitive and complete), the induced dominance is a weak order.*

Theorem 5. *If preference Π is a weak order (reflexive, transitive and complete) and X is finite and nonempty, nondominated solutions exist and are all mutually indifferent.*

This allows to choose any of such solutions as the overall solution of the problem. This is good, the aim is to “make the right choice”. Weak orders allow to sort impact on a line, with possible ties, as if associating a degree to each impact.

Definition 4. A **value function** is a function $v : F \rightarrow \mathbb{R}$ that associates a real value to each impact (also called *utility functions in economics*). Function v is **consistent** with preference relation Π when

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

or, equivalently

$$\Pi = \{(f, f') \in F \times F \mid v(f) \geq v(f')\}$$

This offers a compact way to represent preference relations, that is also good for computation

$$\max_{x \in X} \{v(f(x))\}$$

if we have analytic expressions for X and $v(f(\cdot))$ and a solving algorithm.

Value functions are **not univocal** (infinite equivalent ones always exist).

If a preference relation admits a consistent value function, the derived relations of indifference and strict preference correspond to identity and strict inequality between the values of the value function.

Theorem 6. *If a preference relation Π admits a consistent value function $v(f)$, then Π is a weak order (reflexive, transitive and complete).*

The proof is simple, the point is to show that the preference enjoys reflexivity (because Π includes pair (f, f) for all $f \in F$), transitivity (because for all triplets $f, f', f'' \in F$ such that relation Π includes pairs (f, f') and (f', f'') , it also includes pair (f, f'')), and completeness (because for each pair (f, f') not included in Π , pair (f', f) belongs to it).

In practice, we start from a preference relation, not from a value function, the inverse would be more useful (the decision problem could be reduced to a maximization of the value function), but it's not always true.

Weak orders not reducible to a consistent value function

Lexicographic order The main example of weak order relation (actually, strong order relation), which doesn't admit a consistent value function.

Considering the simplest two-dimensional case, with real components ($F = \mathbb{R}^2$), the preference relation is defined as

$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \preceq \begin{bmatrix} f'_1 \\ f'_2 \end{bmatrix} \Leftrightarrow f_1 < f'_1 \text{ or } f_1 = f'_1 \wedge f_2 \leq f'_2$$

The decision-maker prefers the smaller impact of the first one, for any value of the second, preferring the smaller value of the second only in the case of a tie for the first.

It can be proven that this relation doesn't admit any consistent value function $v(f)$, that is assigning to each impact in F a real value such that the

preference between two impacts correspond to an inequality between the function values.

The intuitive reason is that no improvement of the second can compensate for a worsening, even very small, of the first.

Multiplicity of the consistent value functions

The existence of a way to build a consistent value function does not imply its uniqueness.

Theorem 7. *Given a value function $v : F \rightarrow \mathbb{R}$, consistent with a preference relation Π on F , for any strictly increasing function $\phi : \mathbb{R} \rightarrow \mathbb{R}$ the composite function $\phi(v(\cdot))$ is also consistent with Π .*

If a weak order admits a consistent value function, it admits infinite equivalent ones, that associate to the same impacts different values, but sorted in the same way.

Weak order preference models

Scalar impact When the impact is one-dimensional, it's often easy (though not always) to turn it into a value function. E.g., if the impact

- is a benefit, set $v(f) = f$
- is a cost, set $v(f) = -f$
- has a target value \bar{f} , set $v(f) = -\text{dist}(f, \bar{f})$

Borda count In the finite case, every weak order admits a value function (Borda count)

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

Basically, the Borda count $B(f)$ is the number of solutions dominated by f .

Lexicographic order If the indicators are all costs (or benefits) and are sorted by importance ($P = (\pi_1, \dots, \pi_p)$), the preference relation Π is a total

order

$$f \preceq f' \Leftrightarrow f_{\pi_1} < f'_{\pi_1} \text{ or } \left\{ \begin{array}{l} f_{\pi_1} = f'_{\pi_1} \\ f_{\pi_2} < f'_{\pi_2} \end{array} \right\} \text{ or } \dots \text{ or } \left\{ \begin{array}{l} f_{\pi_1} = f'_{\pi_1} \\ f_{\pi_2} = f'_{\pi_2} \\ \dots \\ f_{\pi_p} \leq f'_{\pi_p} \end{array} \right\}$$

It does not admit value functions, but can be solved as follows

- find the whole set $X_{\pi_1}^*$ of optimal solutions for $\min_{x \in X} f_{\pi_1}(x)$
- find the whole set $X_{\pi_2}^*$ of optimal solutions for $\min_{x \in X_{\pi_1}^*} f_{\pi_2}(x)$
- ...
- find a single optimal solution $x_{\pi_p}^*$ for $\min_{x \in X_{\pi_{p-1}}^*} f_{\pi_p}(x)$

Utopia point This model of preference:

1. Identifies an ideal impact f^* independently optimizing each indicator

$$f_l^* = \min_{x \in X} f_l(x)$$

and combining the optimal values in a vector $f^* = [f_1^* \dots f_p^*]^T$. Determining such value is a classical optimization problem, sometimes hard, but generally possible

2. Finds a solution with impact having minimum “distance” from f^*

$$\min_{x \in X} \text{dist}(f(x), f^*)$$

Different definitions of distance imply different results (Manhattan distance, Euclidean distance, ...; infinite different distances can be defined), and the choice is arbitrary.

If the indicators are heterogeneous, the units of measure have an influence and conversion coefficients are required to standardize them. The choice of coefficient is complex and, at least partly, arbitrary.

4.2 Multi Attribute Utility Theory MAUT

The **Multi Attribute Utility Theory (MAUT)** assumes that the preference relation of the decision-maker is a weak order, admitting a consistent

value function, but the decision-maker is unable to make the value function explicit without help.

It poses the problem to derive from the preference relation Π the consistent value function $u : F \rightarrow \mathbb{R}$. We'll denote $u(f)$ as *utility function* (economical notation).

We assume:

- a preference relation Π with a consistent utility function $u(f)$
- a certain environment: $|\Omega| = 1 \implies f(x, \bar{\omega})$ reduces to $f(x)$
- a single decision maker: $|D| = 1 \implies \Pi_d$ reduces to Π

And we know the preference Π , but not the utility function $u(f)$. *How to build it?*

Specific models of preference have specific applications, *case by case, they might work or not*. We want a **general way to derive $u(f)$ from Π**

1. Introduce a graphical tool (indifference map)
2. Turn the graph into a function, with a complex error-prone process
3. Define a special case with a simpler process (additive value functions)
4. Characterize the preference relations falling within the special case

4.2.1 Indifference Curves

Definition 5. *Given an impact set F in the space of the indicators \mathbb{R}^p and a preference relation Π which is a weak order, we denote as **indifference curve** every subset of impacts that are reciprocally different.*

An indifference curve is a set $I \subseteq F$ of **reciprocally indifferent impacts**. They always enjoy the following properties:

- The **curves covers F** : every impact belongs to a curve, due to the completeness of the relation
- Any two indifference curves I and I' have **empty intersection** (transitivity would merge them)
- The weak order on impacts maps onto a **total order on curves (antisymmetry)**; there are no incomparable impacts and all indifferent

impacts belong to the same curve, so impacts on different curves are linked by strict preference

Usually, technical assumptions are made:

- **Continuity** implies that the curves are *regular mathematical objects* and not completely general set of points
- The utility function $u(f)$ has a **continuous infinity of values** $c \in \mathbb{R}$
- Each indifference curve is expressed in **implicit form**

$$u(f) = c$$

- When the implicit form $u(f) = c$ can be turned into an **explicit form**

$$f_l = f_l(c, f_1, \dots, f_{l-1}, f_{l+1}, \dots, f_p)$$

an indifference curve is a **hypersurface of $p - 1$ dimensions in \mathbb{R}^p** (when $p = 2$, it's a line in the plane)

If a value function is known, the family of all indifference curves admits an analytic representation in the indicator space F through the equation $u(f) = c$, where c is a constant parameter that identifies each single curve. This corresponds to an analytic representation of the solution space X through the parametric equation $u(f(x)) = c$.

Indifference curves and utility function (Indifference map)

If a preference relation admits a consistent value function, it admits infinitely many (as seen in 4.1.2); the indifference curves of such functions, however, are always the same.

The correspondence between preference relations and indifference curves is one-to-one, the one between preference relations and value functions is one-to-many.

Given a weak order preference relation Π on F , its **indifference map** \mathcal{I}_Π is the ordered family of indifference curves covering F . The correspondence between Π and \mathcal{I}_Π is one to one

- Π identifies all groups of indifferent impacts (curves) and their order
- \mathcal{I}_Π identifies the preference between all pairs of impacts

The indifference map \mathcal{I}_Π corresponds to infinite utility functions $u(f)$.

4.2.2 Determining the utility function

Given a preference relation Π on F :

1. Extract a sample \tilde{F} from F
2. Ask the decision-maker to
 - sort the sampled impacts
 - Identify their equivalence classes
3. Draw an interpolating curve for each equivalence class
4. Guess a parametric utility function family from their shape

$$u = u_{\alpha}(f) \text{ with } \alpha = [\alpha_1 \dots \alpha_p]^T$$

5. Each pair of indifferent impacts implies an equation on α

$$f \sim f' \Leftrightarrow u(f(\alpha)) = u(f'(\alpha))$$

6. Add a normalization condition to select one of the equivalent utilities
7. Make consistency checks by comparing other pairs of impacts; if they fail, go back to 4 and change the parametric utility

The process is, in general, very complex and error prone

- Large samples are costly
- The sample must include at least $p - 1$ pairs of indifferent impacts found by trial and error
- Small samples ($\approx p$) lead to likely incorrect curves; if the samples is too large, the workload for the decision maker becomes huge (if k different values for each indicator leads to k^p different impacts to evaluate)
- Numerical errors over many equations combine in cascade
- Mutually dependent pairs are useless
- High dimensional spaces make it hard to draw curves and guess u_{α}

Some properties (not guaranteed) which allow to make easier estimates, helping to draw curves and guessing $u(f)$:

- **Invertibility:** $u(f) = c$ can be solved with respect to each f_i

- always verified when the indicators are costs or benefits
- basically, when we can get each indicator f_l from the function $u(f)$ (analytically, the function can be inverted)

Under this assumption, an indifference curve $u(f) = c$ can be written in explicit form as $f_l = f_l(c, f_1, \dots, f_p)$

- **Monotony:** strictly decreasing or increasing difference curves (*an increase in one is balanced by a decrease in another*)
 - always verified when the indicators are all costs or benefits

In order to compensate for the variations of an indicator, the other ones must vary in a well-determined direction

- **Convexity or concavity** (“law of diminishing marginal utility”):
 - costs lead to concave curves, benefits to convex ones

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; basically, if a resource (indicator) is scarce, increasing it brings a large utility, compensated by a strong decrease of other resources, if a resource is abundant the same increase brings a small utility, compensated by a weak decrease of other resources (this is the case of convex curves, they would be concave in the case of indicators expressing cost)

4.2.3 Additive utility functions (game changer)

Definition 6. We denote a utility function as **additive** when it can be expressed as the sum of functions of the single indicators:

$$u(f_1, \dots, f_p) = \sum_{l=1}^p u_l(f_l)$$

It’s a specific case that brings many simplifications:

- Ask different decision-makers for each indicator f_l (*split the work*)
- Ask decision-makers with experience in the sector (*more reliable*)
- Compare scalar values f_l instead of vectors f (*easier and better*)
- Build functions $u_l(f_l)$ with one argument (*easier and better*)

Since a utility function has infinitely many forms, **a nonadditive function can have an additive equivalent form**, but *that is not guaranteed, how can we know?*

Example (Cobb-Douglas functions):

$$u(f) = \prod_{l=1}^p f_l^{\alpha_l} \text{ is equivalent to } u'(f) = \log u(f) = \sum_{l=1}^p \alpha_l f_l$$

If the utility function is additive, the problem of estimating it can be reduced to the estimation of p single-variable functions, making the process easier.

4.2.4 Preferential Independence

How to know that Π admits an additive consistent utility function? Given the set of attribute indices $P = \{1, \dots, p\}$, focus on subset $L \subset P$; we can write that

$$f = \begin{bmatrix} f_L \\ f_{P \setminus L} \end{bmatrix}$$

where f_L and $f_{P \setminus L}$ are the subvectors of impact f corresponding, respectively, to the indicators of L and $P \setminus L$ (we just divided the whole thing in two parts).

Definition 7. *A proper subset of indicators L is **preferentially independent** from the complementary subset $P \setminus L$ when, given two impacts with identical values of the indicators in $P \setminus L$, the preference relation between them does not depend on such values:*

$$\begin{bmatrix} f_L \\ \phi \end{bmatrix} \preceq \begin{bmatrix} f'_L \\ \phi \end{bmatrix} \Leftrightarrow \begin{bmatrix} f_L \\ \psi \end{bmatrix} \preceq \begin{bmatrix} f_L \\ \psi \end{bmatrix}$$

for all subvectors ϕ, ψ, f_L, f'_L such that the four impacts are in F :

$$\begin{bmatrix} f_L \\ \phi \end{bmatrix}, \begin{bmatrix} f'_L \\ \phi \end{bmatrix}, \begin{bmatrix} f_L \\ \psi \end{bmatrix}, \begin{bmatrix} f'_L \\ \psi \end{bmatrix} \in F$$

Preferences between values in L do not depend on the values out of L .

For example: a cost is preferentially independent from all other indicators (the lower the better), but this is not true for a thermostat, at a certain

humidity a lower temperature is better, at other ones, a higher temperature might be better.

Preferential independence is **not symmetric**:

$$L \text{ independent from } P \setminus L \not\Rightarrow P \setminus L \text{ independent from } L$$

Preferential independence **on single indicators does not imply independence on larger subsets**:

$$\{l\} \text{ independent from } P \setminus \{l\}, \forall l \in P \not\Rightarrow L \text{ independent from } P \setminus L, \forall L \subseteq P$$

Mutual preferential independence

Preferential independence of relation Π and additivity of function $u(f)$ are strictly connected, even if not rigorously equivalent.

Definition 8. *We say that a problem enjoys **mutual preferential independence** when every proper subset of indicators $L \subset P$ is independent from its complement $P \setminus L$.*

Mutual preferential independence is a necessary condition for additivity.

The above definition requires to check every nonempty proper subset $P \subset L$:

- $2^p - 2$ subsets
- infinite 4-tuples of subimpacts (to sample) for each subset

Theorem 8. *Mutual preferential independence holds if and only if, given $\bar{l} \in P$, every pair $L = \{l, \bar{l}\}$ is preferentially independent from $P \setminus L$ for all $l \in P \setminus \{\bar{l}\}$*

We need to check only $p - 1$ pairs (*but single indicators are not enough*).

Theorem 9. *If Π admits a consistent additive utility function $u(f)$, then Π enjoys mutual preferential independence.*

To verify this is enough to apply the definition (for a more detailed proof, p.92 of the notes).

The problem is that the opposite is needed, we can verify mutual preferential independence and not additivity, as the utility function is always unknown,

we need a sufficient condition.

Theorem 10. *A decision problem with $p \geq 3$ indicators enjoying mutual preferential independence admits an additive utility function $u(f)$.*

Unfortunately, when $p = 2$, mutual preferential independence is necessary but not sufficient for additivity, as it reduces to checking each indicator with respect to the other one.

4.2.5 Marginal rate of substitution MRS

To summarize, we know the preference Π , not the utility function $u(f)$ and the general process to find it is complex and error-prone. For additive functions it's much simpler, but is $u(f)$ additive?

- for $p \geq 3$, mutual preferential independence \Leftrightarrow additivity
- for $p = 2$, mutual preferential independence \Leftarrow additivity

The problem lies in how the indifference curves behave in F ; the missing condition to guarantee additivity concerns the steepness of the indifference curve. We need a measure of steepness.

The behavior of a curve is described by the relation between f_1 and f_2 : from f , vary f_1 and update f_2 so as to remain on the indifference curve.

Definition 9. *We denote as **marginal rate of substitution (MRS)** of f_1 with f_2 in a given impact f the limit*

$$\lambda_{12}(f) = \lim_{\delta f_1 \rightarrow 0} - \frac{\delta f_2(f, \delta f_1)}{\delta f_1}$$

where $\delta f_2(f, \delta f_1)$ is such that $f + \begin{bmatrix} \delta f_1 \\ \delta f_2(f, \delta f_1) \end{bmatrix} \sim f$

In other words, the MRS is the limits, for infinitesimal variations of f_1 , of the ration between the variations of f_2 and f_1 that produces impacts indifferent with respect to f . The sign of the limit is reversed because it's often negative (e.g. when the two indicators are both costs or benefits).

In problems with more than two indicators, a marginal rate of substitution is defined for each pair of indicators, and is computed assuming that all other indicators remain constant.

The definition assumes to start from an impact f , slightly modifying the value of indicator f_1 and determining the corresponding variation of f_2 that allows to stay on the initial indifference curve.

Assuming that the indifference curves are regular arcs and representing them in parametric form (two functions of parameter α , continuous up to the first derivative):

$$\begin{cases} f_1 = f_1(\alpha) \\ f_2 = f_2(\alpha) \end{cases}$$

The variations δf_1 and δf_2 used in the definition of $\lambda_{12}(f)$ correspond to a variation $\delta\alpha$ of the parameter, guaranteeing that the impact remains on the indifference curve and allowing us to express the MRS as

$$\lambda_{12}(f)? \lim_{\delta f_1 \rightarrow 0} -\frac{\delta f_2(f, \delta f_1)}{\delta f_1} = \lim_{\delta\alpha \rightarrow 0} -\frac{f_2(\alpha + \delta\alpha) - f_2(\alpha)}{f_1(\alpha + \delta\alpha) - f_1(\alpha)} = -\frac{\frac{df_2}{d\alpha}}{\frac{df_1}{d\alpha}}$$

This form is useful to prove **reciprocity**:

$$\lambda_{12}(f) = \frac{1}{\lambda_{21}(f)}$$

MRS and utility function A second expression shows the relation between MRS and $u(f)$; by changing α we move along the indifference curve, but the utility function remains constant: $u(f(\alpha)) = c$ for all α , therefore having a first derivative equal to zero with respect to α :

$$\frac{du(f_1(\alpha), f_2(\alpha))}{d\alpha} = 0 \implies \frac{\partial u}{\partial f_1} \frac{df_1}{d\alpha} + \frac{\partial u}{\partial f_2} \frac{df_2}{d\alpha} = 0 \implies -\frac{\frac{df_2}{d\alpha}}{\frac{df_1}{d\alpha}} = \frac{\frac{\partial u}{\partial f_1}}{\frac{\partial u}{\partial f_2}}$$

And this implies that

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

The MRS of f_1 with f_2 is the ratio of the partial derivatives of the utility function with respect to f_1 and f_2 . If the utility depends strongly on f_1 and weakly on f_2 , the MRS is large, i.e., a large variation of f_2 is required to compensate for a small variation of f_1 .

This form also holds for $p \geq 3$ indicators, given that the rate of substitution is defined keeping all indicators constant, except for two.

MRS and indifference curves If the indifference curves are invertible, each value of parameter α corresponds to a different value of f_1 and f_2 , one can define a function $\alpha = \alpha(f_1)$ which implies an explicit expression for the indifference curve: $f_2 = f_2(\alpha(f_1))$. The first derivative of that expression is

$$\frac{df_2}{df_1} = \frac{df_2}{d\alpha} \frac{d\alpha}{df_1} = \frac{\frac{df_2}{d\alpha}}{\frac{df_1}{d\alpha}}$$

From which

$$\lambda_{12}(f) = -\frac{df_2}{df_1}$$

The MRS is the *negative slope of the tangent to the indifference curve*; the steepness of the indifference curves, with a reverse sign.

In simple terms, the MRS is how much of f_2 are you willing to give away for a small increase of f_1 .

4.2.6 Additivity and MRS

By combining two distinct values for f_1 (f'_1 and f''_1) and f_2 (f'_2 and f''_2) one can build four different impacts. In general, the MRS in these four are fully independents, however, they can be related.

Definition 10. We denote as **corresponding trade-off condition** the following property

$$\lambda_{12}(f'_1, f'_2) \lambda_{12}(f''_1, f''_2) = \lambda_{12}(f''_1, f'_2) \lambda_{12}(f'_1, f''_2)$$

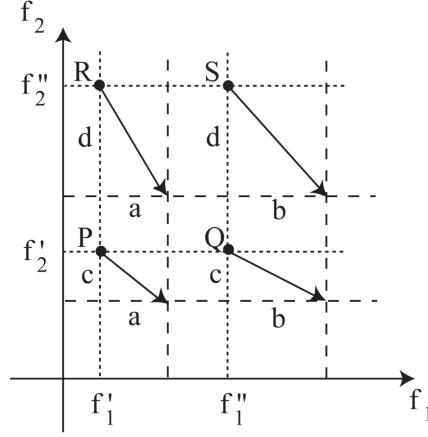
for every quadruplets of impacts $(f'_1, f'_2), (f''_1, f''_2), (f''_1, f'_2), (f'_1, f''_2) \in F$.

Some indifference maps enjoy such property, and it is **global**: it relates far away impacts.

The corresponding trade-off condition is easier to interpret if rewritten as

$$\frac{\lambda_{12}(f'_1, f'_2)}{\lambda_{12}(f''_1, f'_2)} = \frac{\lambda_{12}(f'_1, f''_2)}{\lambda_{12}(f''_1, f''_2)} \quad \text{or} \quad \frac{\lambda_{12}(f'_1, f'_2)}{\lambda_{12}(f'_1, f''_2)} = \frac{\lambda_{12}(f''_1, f'_2)}{\lambda_{12}(f''_1, f''_2)}$$

Basically, multiplying two MRSs on the diagonal of a rectangle (composed by 2 values for each of the 2 indicators) is the same as multiplying the two MRSs for the values on the opposite diagonal.



In the image it can be seen that in the four points P , Q , R and S the MRSs are, respectively, equal to $-a/c$, $-b/c$, $-a/d$, $-b/d$. The product of the rates in P and S is ab/cd , and it coincides with the product of the rates in Q and R .

The ratio of λ between points with the same abscissa does not depend on the abscissa. In other words, even if the MRS is nonuniform, it changes by the same factor moving between the same coordinates.

Theorem 11. *A preference relation Π admits an additive utility function $u(f)$ if and only if it enjoys both*

1. *mutual preferential independence*
2. *the corresponding trade-off condition*

The “only-if part” it’s easy to prove, the “if” was needed in order to assume additivity.

Proof. Once again, it’s a matter of applying the definition:

$$\lambda_{12}(f'_1, f'_2) \lambda_{12}(f''_1, f''_2) = \frac{\frac{\partial u}{\partial f'_1}}{\frac{\partial u}{\partial f'_2}} \frac{\frac{\partial u}{\partial f''_1}}{\frac{\partial u}{\partial f''_2}}$$

Since $u(f_1, f_2) = u_1(f_1) + u_2(f_2)$:

$$\lambda_{12}(f'_1, f'_2) \lambda_{12}(f''_1, f''_2) = \frac{\frac{\partial u}{\partial f'_1}}{\frac{\partial u}{\partial f'_2}} \cdot \frac{\frac{\partial u}{\partial f''_1}}{\frac{\partial u}{\partial f''_2}} = \frac{\frac{\partial u}{\partial f'_1}}{\frac{\partial u}{\partial f'_2}} \frac{\frac{\partial u}{\partial f''_1}}{\frac{\partial u}{\partial f''_2}}$$

and getting back to function $u(f_1, f_2)$:

$$\lambda_{12}(f'_1, f'_2) \lambda_{12}(f''_1, f''_2) = \frac{\frac{\partial u}{\partial f'_1}}{\frac{\partial u}{\partial f'_2}} \frac{\frac{\partial u}{\partial f''_1}}{\frac{\partial u}{\partial f''_2}} = \lambda_{12}(f'_1, f'_2) \lambda_{12}(f''_1, f''_2)$$

□

The corresponding trade-off condition is thus necessary for additivity (also sufficient, under suitable conditions).

To assume additivity:

- When $p \geq 3$, check mutual preferential independence for $p-1$ indicator pairs
- When $p = 2$ check
 - independence for the single indicators
 - the corresponding trade-off condition

4.2.7 Building an additive utility function

The expression for an additive utility function

$$u(f) = \sum_{l=1}^p u_l(f_l)$$

assumes the same measure unit for all u_l function. In practice, different fields use different units, therefore:

1. Adopt normalized utilities: pure numbers $\tilde{u}_l(f_l)$ instead of $u_l(f_l)$, usually within the interval $[0, 1]$
2. Introduce weights w_l to combine them

$$u(f) = \sum_{l=1}^p w_l \tilde{u}_l(f_l)$$

Intuitively, we are splitting the task into “rescaling” indicators into utilities, removing nonlinearities, and “combining” heterogeneous utilities into a single one.

The normalized utility expression is a **linear** and **convex** combination:

- **Conic:** all coefficients are nonnegative ($w_l \geq 0$ for all $l \in P$)
- **Affine:** the coefficients have unitary sum ($\sum_{l=1}^p w_l = 1$)

The bisection method

Building a utility function for a one-dimensional impact ($F \subseteq \mathbb{R}$) is much easier than for a multidimensional one, it only takes comparing pairs of numbers and wondering which one of the two is better.

However, the function we want to build must respect a condition stronger than just sorting all impacts, we want to measure the strength of relative preference: nearly indifferent impacts should have similar utility values and impacts well distinguished from a preferential point of view should have very different utility values.

The bisection methods builds such function with a dichotomic approach. For example, to build the normalized utility $\tilde{u}_C(C)$ for the daily calorie intake C , given $F_C = [0, 10000]$, interview an expert about a specific individual

1. Ask the expert for the worst values in F_C (for example, $C_1^\dagger \leq 1000$ and $C_2^\dagger \geq 6000$) and set $\tilde{u}_C(C) = 0$ for such values
2. Ask the expert for the best values in F_C (for example, $2200 \leq C^* \leq 2600$) and set $\tilde{u}_T(22) = 1$ for such values
3. Ask the expert for values of exactly intermediate utility between T^\dagger and T^* (for example, $C = 1800$ and 3000) and set $\tilde{u}_C(C) = 1/2$
4. Go on, asking for values of intermediate utility between the fixed ones and set \tilde{u}_T accordingly
5. Guess an interpolating function

When the indicator represents a cost or benefit, sometimes it's possible to assume the utility proportional to the indicator f_l . We can easily generate a normalized utility function:

$$\tilde{u}_l(f_l) = \frac{f_l - \min_{x \in X} f_l(x)}{\max_{x \in X} f_l(x) - \min_{x \in X} f_l(x)}$$

Switch them for costs.

If $\min_{x \in X} f_l(x)$ or $\max_{x \in X} f_l(x)$ are unknown or hard to compute, an underestimate and overestimate can be used instead, yielding a normalized utility function, different from above mentioned one. This can lead to a

“compression” of the value range if the estimates are loose (it’s not using the whole interval $[0, 1]$).

4.2.8 Determining the weights

We still need to determine the weights with which to combine the components of the additive utility function. This requires to identify a sufficient number of pairs of indifferent impacts, equaling the utilities of the two impacts imposes a constraint on the weights of w_l , allowing us to find a single solution, with enough constraints (linear equation system).

As in the general case, $p - 1$ independent pairs of indifferent impacts are needed:

- The equations are linear in w

$$f \sim f' \Leftrightarrow \tilde{u}(f) = \tilde{u}(f') \Leftrightarrow \sum_{l=1}^p \tilde{u}_l(f_l) w_l = \sum_{l=1}^p \tilde{u}_l(f'_l) w_l$$

- The normalization condition imposes convexity

$$\sum_{l=1}^p w_l = 1$$

This process works correctly in the ideal case, but if indifference is imprecise, the pairs give wrong weights. The solution is to build a complete pairwise comparison of all indicators and analyze its consistency.

The pairwise comparison matrix

Select a pair of indifferent impacts (f, f') with:

- Different values for f_l and f_m
- Identical values for all other indicators, $f_n = f'_n$ for all $n \in P \setminus \{l, m\}$

The equation reduces to

$$w_l \tilde{u}_l(f_l) + w_m \tilde{u}_m(f_m) = w_l \tilde{u}_l(f'_l) + w_m \tilde{u}_m(f'_m)$$

That simply becomes

$$\tilde{\lambda}_{lm} = \frac{w_l}{w_m} = - \frac{\tilde{u}_m(f'_m) - \tilde{u}_m(f_m)}{\tilde{u}_l(f'_l) - \tilde{u}_l(f_l)}$$

Definition 11. We denote as *pairwise comparison matrix* $\tilde{\Lambda} = \{\tilde{\lambda}_{lm}\}$ the matrix reporting all rates of substitution between the normalized utilities.

It contains all the weight ratios

$$\tilde{\Lambda} = \left\{ \frac{w_l}{w_m} \right\}$$

expressing the relative weights between the single normalized utilities, that is, indicators (once nonlinearities and units of measure are removed).

A correct pairwise comparison matrix $\tilde{\Lambda}$ enjoys the following properties:

- **Positivity:** $\tilde{\lambda}_{lm} > 0$ for all $l, m \in P$
- **Reciprocity:** $\tilde{\lambda}_{lm} = \frac{1}{\tilde{\lambda}_{ml}}$ for all $l, m \in P$
- **Consistency:** $\tilde{\lambda}_{ln} = \tilde{\lambda}_{lm}\tilde{\lambda}_{mn}$ for all $l, m, n \in P$

We can check these properties on $\tilde{\Lambda}$ to be sure that $u(f)$ makes sense.

4.2.9 The process in summary

In summary, the process to determine the utility function is composed of the following steps:

1. Interview the decision-maker in order to understand whether mutual preferential independence holds (ask to compare different pairs of indicators with the complementary subset)
2. In the positive case, focus on each attribute, fixing the other ones to any value, and build a single-variable component of the utility function with a method respecting the entity of the relative preference gaps
3. Determine the weights with which to combine such functions, sampling impacts until a sufficient number of pairs of indifferent impacts has been found
4. Check *a posteriori* that the obtained utility function is valid, testing it on impacts not used before

In any moment, it can be necessary to go back to correct the single components of the utility or even rejecting the additivity assumption, in which case the whole process is invalidated.

Chapter 5

Mathematical Programing

This chapter will deal with the simplest class of decision problems. We assume:

- A preference relation Π with a known consistent utility function $u(f)$

$$\exists u : F \rightarrow \mathbb{R} : \Pi = \{(f, f') \in F \times F \mid u(f) \geq u(f')\}$$

- A certain environment: $|\Omega| = 1 \implies f(x, \bar{\omega})$ reduces to $f(x)$; a single scenario, the impact f depends only on the alternative x
- A single decision-maker: $|D| = 1 \implies \Pi_d$ reduces to Π

This class of problems reduces the decision problem to classical optimization:

$$\max_{x \in X} u(f)$$

We'll discuss a solving technique that is very general but complex and inefficient.

This class of problems is mainly treated in mathematical literature, where preference is expressed as a cost function, so the most common form is

$$\min_{x \in X} f(x)$$

where $f(x)$ replaces $-u(f(x))$ (not the original f).

We also assume regularity for the objective and feasible region:

- $f(x) \in C^1(X)$; the cost function is continuous with its first derivative

- $X = \{x \in \mathbb{R}^n \mid g_j(x) \leq 0, j = 1, \dots, m\}$ with $g_j(x) \in C^1(X)$; the solution set X can be described through a finite system of inequalities, where the constraint functions $g_j(x)$ are real-valued continuous functions with a continuous first derivative in the feasible solution set of the problem

These are very general assumptions. This class of problem (with assumptions more or less strict with respect to the continuity of functions) are denoted under the label of **Mathematical Programming** problems, describing both preference and feasible region through real-valued functions allows to apply mathematical analysis to these decision problems.

5.1 Basic concepts

Definition 12. Given a set $X \subseteq \mathbb{R}^n$ and a function $f : X \rightarrow \mathbb{R}$ we denote as **global optimum point** a point $x^\circ \in X$ such that

$$f(x^\circ) \leq f(x) \quad \forall x \in X$$

Definition 13. Given a set $X \subseteq \mathbb{R}^n$ and a function $f : X \rightarrow \mathbb{R}$, we denote as **local optimum point** a point $x^* \in X$ such that

$$\exists \epsilon > 0 : f(x^*) \leq f(x), \quad \forall x \in X \cap \mathcal{U}_{x^*, \epsilon}$$

where

$$\mathcal{U}_{x^*, \epsilon} = \{x \in \mathbb{R}^n \mid \|x - x^*\| < \epsilon\}$$

is a neighborhood of x^* of width ϵ ($\|x - x^*\|$ is the norm of vector $x - x^*$).

It can easily be seen that a global optimum is, by definition, also a local optimum:

$$X^\circ \subseteq X^*$$

There is no process to find X° or X^* directly, so we want to search the conditions necessary for local optimality. These conditions aim to identify points “candidate” to being local optimum, weakening twice our request:

$$\begin{array}{ccccc} \text{Global optimum} & \implies & \text{Local optimum} & \implies & \text{KKT conditions} \\ X^\circ & \subseteq & X^* & \subseteq & X^{KKT} \end{array}$$

The conditions are known as **Karush-Kuhn-Tucker conditions**, from the names of their discoverers.

After finding the “weakened” set, we enumerate X^{KKT} exhaustively to find X° .

The KKT conditions identify candidate points as follows:

1. Lay out the conditions as a system of equalities and inequalities
2. Solve the system, finding X^{KKT}
3. Evaluate every point in X^{KKT} , keeping only the best ones (which will compose X°)

The method requires X^{KKT} to be a finite set, or at least a set which can be analytically describes, in order to determine X° .

The basic tool will be linear approximation in small neighborhoods (hence the false positives).

5.1.1 Taylor’s series expansion

Every sufficiently regular function (i.e., continuous function with continuous derivatives up to a suitable order k) can be locally approximated with a polynomial of degree k .

Theorem 12. *Let $f \in C^k(\mathcal{U}_{\tilde{x},\epsilon})$ be a function of a real variable $x \in \mathbb{R}$. For all $x \in \mathcal{U}_{\tilde{x},\epsilon}$, Taylor’s series expansion holds:*

$$f(x) = \sum_{i=0}^k \frac{f^{(i)}(\tilde{x})}{i!} (x - \tilde{x})^i$$

$$\text{with } f^{(0)} = f, \quad f^{(i)} = \frac{d^i f}{dx^i} \quad \forall i \in \mathbb{N}^+, \text{ and } \lim_{x \rightarrow \tilde{x}} \frac{R_k(x - \tilde{x})}{\|x - \tilde{x}\|} = 0$$

Any regular function can be locally approximated in \tilde{x} by its tangent line, terms with exponents higher than 1 improve the approximation (which we’ll not consider). Any function regular up to the first order ($f : \mathbb{R} \rightarrow \mathbb{R}$ and $f \in C^1(\mathcal{U}_{\tilde{x},\epsilon})$) admits a linear approximation

$$f(x) = f(\tilde{x}) + f'(\tilde{x})(x - \tilde{x}) + R_1(|x - \tilde{x}|)$$

which can be generalized to multiple-variable functions ($x \in \mathbb{R}^n$) by

$$f(x) = f(\tilde{x}) + (\nabla f(\tilde{x}))^T (x - \tilde{x}) + R_1(\|x - \tilde{x}\|)$$

where

$$\lim_{x \rightarrow \tilde{x}} \frac{R_1(\|x - \tilde{x}\|)}{\|x - \tilde{x}\|} = 0$$

and $\nabla f(x)$ is the gradient vector

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \dots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

It's the direction of quickest increase for $f(\cdot)$.

5.1.2 Arcs

\mathbb{R}^n offers many ways to move away from \tilde{x} .

Definition 14. Given a point $\tilde{x} \in \mathbb{R}^n$, an **arc** in \tilde{x} is a parametric curve

$\xi : \mathbb{R}^+ \rightarrow \mathbb{R}^n$, that is $\xi(\alpha) = \begin{bmatrix} \xi_1(\alpha) \\ \dots \\ \xi_n(\alpha) \end{bmatrix}$, such that $\xi(0) = \tilde{x}$ and $\xi_1(\alpha) \in C^1(\mathbb{R}^+)$.

Basically, a function that starts at \tilde{x} (since $\xi(0) = \tilde{x}$) and then gets further from it, with each component continuously differentiable for $\alpha > 0$.

Definition 15. An arc $\xi(\alpha)$ is **feasible** for a given region $X \subseteq \mathbb{R}^n$ when the curve remains in X for small α

$$\exists \bar{\alpha}_f > 0 : \xi(\alpha) \in X, \quad \forall \alpha \in [0; \bar{\alpha}_f)$$

Given a region X (our function), we can travel “a little” (a small enough interval, as defined) along the arc without leaving X .

Definition 16. An arc $\xi(\alpha)$ is **improving** for a given function $f : X \rightarrow \mathbb{R}$ when f is strictly better in $\xi(\alpha)$ than in \tilde{x} for all small positive α

$$\exists \bar{\alpha}_i > 0 : f(\xi(\alpha)) < f(\tilde{x}), \quad \forall \alpha \in (0; \bar{\alpha}_i)$$

Along the curve, but close enough to \tilde{x} , the value of the function f is strictly better than in \tilde{x} .

5.2 Necessary conditions for local optimality

Theorem 13. *If $\tilde{x} \in X \subseteq \mathbb{R}^n$ (a point in a region), $f(\cdot) \in C^1(X)$ (continuously differentiable function) and $\xi(\alpha)$ is an arc in \tilde{x} , feasible for X and improving for $f(\cdot)$, then \tilde{x} is not locally optimal for $f(\cdot)$ in X .*

If a feasible improving arc in \tilde{x} exists, then \tilde{x} can't be a local optima; no improving arc can exist as a local optima.

Proof. By assumption, for suitable values $\bar{\alpha}_f > 0$ and $\bar{\alpha}_i > 0$, we have:

- $\xi(\alpha)$ feasible: $\xi(\alpha) \in X$ for all $\alpha \in [0, \bar{\alpha}_f]$
- $\xi(\alpha)$ improving: $f(\xi(\alpha)) < f(\tilde{x})$ for all $\alpha \in (0, \bar{\alpha}_i)$

Since $\xi(\alpha)$ is a continuous arc

$$\lim_{\alpha \rightarrow 0} \xi(\alpha) = \tilde{x} \Leftrightarrow \forall \epsilon > 0, \exists \bar{\alpha}_\epsilon : \|\xi(\alpha) - \tilde{x}\| < \epsilon, \forall \alpha \in (0, \bar{\alpha}_\epsilon)$$

that is, $\xi(\alpha) \in \mathcal{U}_{\tilde{x}, \epsilon}$, $\forall \alpha \in (0, \bar{\alpha}_\epsilon)$.

Now $\alpha = \frac{1}{2} \min(\bar{\alpha}_f, \bar{\alpha}_i \bar{\alpha}_\epsilon)$ satisfies all three conditions:

- $\alpha < \bar{\alpha}_f \implies \xi(\alpha) \in X$
- $\alpha < \bar{\alpha}_i \implies f(\xi(\alpha)) < f(\tilde{x})$
- $\alpha < \bar{\alpha}_\epsilon \implies \xi(\alpha) \in \mathcal{U}_{x^*, \epsilon}$

But this contradicts local optimality

$$f(x) \geq f(\tilde{x}), \quad \forall x \in \mathcal{U}_{\tilde{x}, \epsilon} \cap X$$

□

5.2.1 A filtering approach

This suggests an approach to find candidate points: remove from X all the points that are provably nonoptimal:

1. Consider all feasible points as candidates: $C := X$
2. Scan the set of candidate points C
3. For each candidate point \tilde{x} , scan all feasible arcs
4. For each arc d feasible in \tilde{x} , check whether the arc is improving for $f(\cdot)$: if it is, remove \tilde{x} from the candidate set

5. In the end, scan the remaining candidate points, keeping only the best ones

The pseudo-code of which being:

```

 $X^{KKT} := X;$ 
// (continuous set for  $x$ )
for each  $x \in X^{KKT}$  do
  // (continuous set for  $\xi$ , interval for  $\alpha$ )
  for each arc  $\xi(\alpha)$  in  $x$  feasible for  $X$  do
    // (interval for  $\alpha$ )
    if  $\xi(\alpha)$  is improving in  $x$  for  $f(\cdot)$  then
       $X^{KKT} := X^{KKT} \setminus \{x\};$ 
return  $X^{KKT};$ 

```

Problem: the set of feasible points and directions is, generally, infinite. We can replace the infinite loops with more efficient analytic conditions, all based on first-order approximations.

Definition 17. We denote as **tangent direction** of an arc $\xi(\alpha)$ feasible in \tilde{x} the vector composed of the first derivatives of the components ξ_i with respect to parameter α , evaluated in the starting point of the arc

$$p_\xi = \begin{bmatrix} \xi'_1(0) \\ \vdots \\ \xi'_n(0) \end{bmatrix}$$

Straight lines $\xi(\alpha) = \tilde{x} + \alpha d$ have tangent direction d (*arcs generalize directions*).

Example: the arc in $\tilde{x} = (2, 0)$

$$\xi(\alpha) = \begin{bmatrix} 2 \cos \alpha \\ 2 \sin \alpha \end{bmatrix}$$

describes the circumference with center in the origin and radius 2. Its tangent direction is

$$p_\xi = \begin{bmatrix} -2 \sin 0 \\ 2 \cos 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

Theorem 14. If x^* is locally optimal in X for $f(\cdot)$ and $\xi(\alpha)$ is a feasible arc in x^* for X , then

$$\nabla f(\tilde{x})^T p_\xi \geq 0$$

Proof. If $\xi(\alpha)$ is a feasible arc, there exists a coefficient $\bar{\alpha}$ such that $\xi(\alpha) \in X$ for all $\alpha \in [0; \bar{\alpha})$ (remains in X for a small α) and since x^* is a local optima, for small α it holds that $f(\xi(\alpha)) \geq f(x^*)$.

Apply Taylor's expansion to $f(\xi(\alpha))$ in $\alpha = 0$

$$\begin{aligned} f(\xi(\alpha)) &\geq f(x^*) \implies \\ \implies f(\xi(0)) + \left. \frac{df}{d\alpha} \right|_{\alpha=0} \alpha + R_1(\|\xi(\alpha) - \xi(0)\|) &\geq f(x^*) \implies \\ \implies (\nabla f(x^*))^T p_\xi + \frac{R_1(\|\xi(\alpha) - \xi(0)\|)}{\alpha} &\geq 0 \end{aligned}$$

If α converges to 0, by continuity the inequality is preserved

$$\begin{aligned} \lim_{\alpha \rightarrow 0} \left((\nabla f(\tilde{x}))^T p_\xi + \frac{R_1(\|\xi(\alpha) - \xi(0)\|)}{\|\xi(\alpha) - \xi(0)\|} \frac{\|\xi(\alpha) - \xi(0)\|}{\alpha} \right) &\geq 0 \implies \\ \implies (\nabla f(\tilde{x}))^T p_\xi &\geq 0 \end{aligned}$$

□

Example:

$$\begin{aligned} \min f(x) &= x_2 \\ g_1(x) &= x_1^2 + x_2^2 \leq 4 \end{aligned}$$

with $\nabla f^T = [0 \ 1]$

- Arc $\xi(\alpha) = \tilde{x} + \alpha[1 \ -1]^T$ is improving in $\tilde{x} = (-2, 0)$: therefore \tilde{x} is not locally optimal

$$\nabla f(-2, 0)^T p_\xi = [0 \ 1] \cdot [1 \ -1]^T = -1 < 0$$

- Arc $\xi(\alpha) = \tilde{x} + \alpha[1 \ 1]^T$ is nonimproving in $\tilde{x} = (0, -2)$: \tilde{x} could be locally optimal (remains candidate until disproval)

$$\nabla f(0, -2)^T p_\xi = [0 \ 1] \cdot [1 \ 1]^T = 1 \geq 0$$

The earlier code can be simplified (possibly missing some removals) to

```

 $X^{KKT} := X;$ 
// (continuous set for  $x$ )
for each  $x \in X^{KKT}$  do
    // (continuous set for  $\xi$ , interval for  $\alpha$ )
    for each arc  $\xi(\alpha)$  in  $x$  feasible for  $X$  do
        // ( $\xi(\alpha)$  is improving in  $x$  for  $f(\cdot)$ )
        if  $\nabla f(\tilde{x})^T p_\xi < 0$  then
             $X^{KKT} := X^{KKT} \setminus \{x\};$ 
return  $X^{KKT};$ 

```

5.2.2 Feasibility condition

Given the analytic description of the feasible region

$$X = \{x \in \mathbb{R}^n \mid g_j(x) \leq 0 \text{ for } j = 1, \dots, m\}$$

we approximate each function $g_j(\cdot)$ with Taylor's expansion.

However, feasibility differs from improvement in two regards:

- It involves many inequalities, instead of a single objective
- It requires weak conditions, instead of a strict one

Definition 18. We denote as **active constraint** in a point $\tilde{x} \in X$ any constraint $g_j(x) \leq 0$ such that $g_j(\tilde{x}) = 0$. We'll indicate by $J_a(x) = \{j \in \{1, \dots, m\} : g_j(x) = 0\}$ the set of indices of the active constraints.

Basically, the constraints exactly equal to zero when considered in point \tilde{x} , and J_a is the set of indices for which this condition holds.

Given point \tilde{x} we partition the constraints into two classes

1. The **active constraints** ($J_a(\tilde{x})$) are exactly satisfied: $g_j(\tilde{x}) = 0$
2. The nonactive constraints are largely satisfied: $g_j(\tilde{x}) < 0$

Example:

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

Active constraints:

- for $x = (-2, -2)$, no active constraints: $J_a(-2, -2) = \emptyset$
- for $x = (3/2, 0)$, one active constraint: $J_a(3/2, 0) = \{2\}$
- for $x = (3/2, \sqrt{7}/2)$, two active constraints: $J_a(3/2, \sqrt{7}/2) = \{1, 2\}$

Theorem 15. *If $\xi(\alpha)$ is a feasible arc in $\tilde{x} \in$ for X , p_ξ its tangent vector. Then*

$$\nabla g_j(\tilde{x})^T p_\xi \leq 0, \quad \forall j \in J_a(\tilde{x})$$

Proof. If $\xi(\alpha)$ is an arc feasible in \tilde{x} , there exists a coefficient $\bar{\alpha}$ such that

$$g_j(\xi(\alpha)) \leq 0, \quad \forall \alpha \in [0; \bar{\alpha}) \text{ and } j = 1, \dots, m$$

that is

$$\begin{aligned}g_j(\xi(\alpha)) &= g_j(\xi(0)) + \left. \frac{dg_j}{d\alpha} \right|_0 \alpha + R_1(\|\xi(\alpha) - \xi(0)\|) = \\ &= g_j(\tilde{x}) + \alpha(\nabla g_j(\tilde{x}))^T p_\xi + R_1(\|\xi(\alpha) - \xi(0)\|) \leq 0\end{aligned}$$

The inequality is obvious for all constraints not active in \tilde{x} and the other two terms converge to zero as $\alpha \rightarrow 0$. Therefore, by continuity, when α is sufficiently small, they cannot reverse the inequality.

For the constraints which are active in \tilde{x} , on the contrary, the inequality introduces a strict condition, in fact, for such constraints $g_j(\tilde{x}) = 0$ so that

$$g_j(\xi(\alpha)) = \alpha(\nabla g_j(\tilde{x}))^T p_\xi + R_1(\|\xi(\alpha) - \xi(0)\|) \leq 0$$

Dividing both terms by α and considering the limit, by continuity the thesis follows

$$\lim_{\alpha \rightarrow 0} \left[(\nabla g_j(\tilde{x}))^T p_\xi + \frac{R_1(\|\xi(\alpha) - \xi(0)\|)}{\|\xi(\alpha) - \xi(0)\|} \frac{\|\xi(\alpha) - \xi(0)\|}{\alpha} \right] = (\Delta g_j(\tilde{x}))^T p_\xi \leq 0$$

□

For any feasible arc $\xi(\alpha)$, vector p_ξ satisfies the conditions above, but a vector p that satisfies them is not always tangent to a feasible arc. Luckily, the problem concerns only some degenerate points.

Regular points

Definition 19. We denote as **regular point** for a given system of constraints $g_j(x) \leq 0$ a point in which all active constraint have linearly independent gradients (**constraint qualification condition**).

In a regular point, the conditions on the gradients of the active constraints are not only necessary, but also sufficient to guarantee that there exists a feasible arc with the given tangent vector.

Theorem 16. If \tilde{x} is a regular point, then $\nabla g_j(\tilde{x})^T p \leq 0$ for all $j \in J_a(\tilde{x})$ if and only if there exists an arc $\xi(\alpha)$ in \tilde{x} feasible for X with tangent direction $p_\xi = p$.

The necessary conditions for feasibility are also sufficient in regular points. The analytical conditions on the gradients can be used to check feasibility in all regular points, nonregular points must be explicitly tested (candidates by default).

Corollary 1. Let $f \in C^1(X)$, $\tilde{x} \in X$ be a regular and local optimum point and p a vector such that

$$(\nabla g_j(\tilde{x}))^T p \leq 0, \quad \forall j \in J_a(\tilde{x})$$

Then

$$(\nabla f(\tilde{x}))^T p \geq 0$$

The code now becomes:

```

 $X^{KKT} := X \setminus \text{NonRegular}(g, X);$ 
// (continuous set for  $x$ )
for each  $x \in X^{KKT}$  do
    // (arc  $\xi(\alpha)$  in  $x$  feasible for  $X$ )
    for each  $p \in \mathbb{R}^n : \nabla g_j(x)^T p \leq 0, \forall j \in J_a(x)$  do
        // ( $\xi(\alpha)$  is improving in  $x$  for  $f(\cdot)$ )
        if  $\nabla f(\tilde{x})^T p_\xi < 0$  then
             $X^{KKT} := X^{KKT} \setminus \{x\};$ 
 $X^{KKT} := X^{KKT} \cup \text{NonRegular}(g, X);$ 
return  $X^{KKT};$ 

```

Special case: equality constraints

Special case: equality constraints $h_i(x) = 0$ are always active and can be treated as pairs of active inequalities: $h_i(x) \leq 0$ and $-h_i(x) \leq 0$

$$\begin{cases} \nabla h_i(\tilde{x})^T p_\xi & \leq 0 \\ -\nabla h_i(\tilde{x})^T p_\xi & \leq 0 \end{cases} \implies \nabla h_i(\tilde{x})^T p_\xi = 0$$

5.2.3 A first geometric interpretation

The property of Corollary 1 has a geometric interpretation that makes it quite intuitive. For each point x taken into account, two sets of vectors are considered:

1. The **feasible cone** $C_{feas}(x)$ is the set of vectors that form angles $\geq 90^\circ$ with the gradients of all active constraints

$$(\nabla g_j(x))^T p \leq 0, \quad \forall j \in J_a(x)$$

It's a cone because each active constraint identifies a half-space of directions pointing "on the opposite side" of its gradient, and intersecting them yields a cone. In regular points, these are the directions tangent to feasible arcs.

2. The **improving half-space** $C_{impr}(x)$: it's the set of vectors that form angles $> 90^\circ$ with the gradient of the objective function

$$(\nabla f(x))^T p < 0$$

Corollary 1 states *the feasible cone must not intersect the improving half-space*. Notice that the former is closed, whereas the latter is open, so that the two sets can touch, but not share internal points. The reason is that the directions in which the two cones touch have zero scalar product with the gradient of the objective and with the gradient of the active constraints; in that case, the first order information is insufficient to determine whether the direction is actually improving or not and actually feasible or not. Therefore, such points cannot be rejected and must remain candidate.

5.3 Karush-Kuhn-Tucker conditions

We want to turn the conditions of Corollary 1 into computationally tractable equivalent conditions, that is into solution of a system of equalities and inequalities.

We have modified the filtering approach using analytic conditions:

- nonregular points are identified and considered as candidates
- in regular points, analytic conditions identify feasible arcs
- analytic conditions identify improving arcs

5.3.1 Farkas' Lemma

Theorem 17 (Farkas' Lemma). *Let $f \in \mathbb{R}^n$ and $g_j \in \mathbb{R}^n$ with $j = 1, \dots, m$ be a vector and a family of vectors in the Euclidean n -dimensional space. Then*

$$\exists \mu_j \geq 0 : f = \sum_{j=1}^m \mu_j g_j \Leftrightarrow p^T f \leq 0, \quad \forall p : p^T g_j \leq 0, \quad \text{with } j = 1, \dots, m$$

In words, vector f is a conical combination of vectors g_j if and only if all vectors pointing on the opposite side of all g_j also point on the opposite side of f .

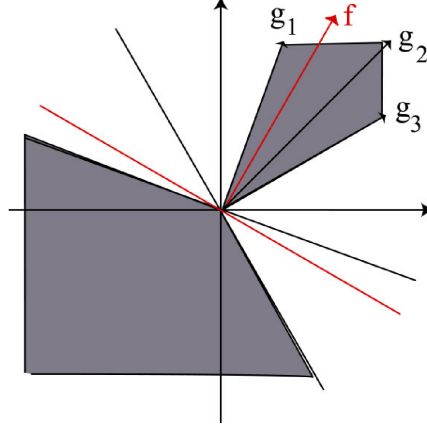
The direct implication is trivial, the converse will be given for granted (hard).

Farkas' lemma states that:

- If f falls within the cone of vectors g , then the cone that is opposite to vectors g_j falls in the half-plane opposite to f

- If the cone that is opposite to vectors g_j falls in the half-plane opposite to f then f falls within the cone of vectors g

The image below provides an example of geometric interpretation of this lemma, with $m = 3$ vectors g_j in a $n = 2$ -dimensional space.



Farkas' lemma and local optimality

We want to apply Farkas' lemma to sift candidate points. Let us map:

- The vectors g_j onto the gradients of the active constraints $\nabla g_j(\tilde{x})$

$$g_j = \nabla g_j(\tilde{x})$$

- Vector f onto the *antigradient* (opposite of the gradient) of the objective function $-\nabla f(\tilde{x})$

$$f = -\nabla f(\tilde{x})$$

The statement of Farkas' lemma becomes

$$\exists \mu_j \geq 0 : \nabla f(\tilde{x}) + \sum_{j \in J_a(\tilde{x})} \mu_j \nabla g_j(\tilde{x}) = 0$$

if and only if

$$(\nabla f(\tilde{x}))^T p \geq 0, \quad \forall p : (\nabla g_j(\tilde{x}))^T p \leq 0, \quad \forall j \in J_a(\tilde{x})$$

A vector p with $p^T \nabla g_j(x) \leq 0, \forall j \in J_a(x)$ is tangent to a feasible arc. We remove point x when $\nabla f(x)^T p < 0$ for all such vectors. Therefore, x is

candidate when $\nabla f(x)^T p \geq 0$ for all feasible tangents, i.e., the second expression in Farkas' lemma, as stated above.

Replacing it with the equivalent first expression, we can tell when a regular point is candidate for local optimality (the first expression is the condition). Don't loop on all x and p : solve a system of equations in μ and x .

5.3.2 Second geometric interpretation

The antigradient $-\nabla f(x)$ is the direction of quickest improvement. The gradients of the active constraint $\nabla g_j(x)$ are the directions of quickest violation.

Multipliers $\mu_j \geq 0$ let combination $\sum_{j \in J_a(x)} \mu_j \nabla g_j(x)$ describe a cone, denoted as C_g , the cone of the gradients of the active constraints.

In candidate points, the antigradient of the objective falls in C_g

$$-\nabla f(x) = \sum_{j \in J_a(x)} \mu_j \nabla g_j(x)$$

When $-\nabla f(x) \in C_g(x)$ the objective improves only violating constraints (*no vector is tangent to a feasible and improving arc*).

5.3.3 Standard form of Karush-Kuhn-Tucker conditions

In general, KKT-conditions are reported in an equivalent standard form, that has two small modifications.

The first problem of the analytic conditions is that

$$\nabla f(x) + \sum_{j \in J_a(x)} \mu_j \nabla g_j(x) = 0$$

The sum applies only to the constraints active in an unknown point x . *How to find them?*

The conditions are reformulated

1. Adding the gradients of the nonactive constraints

$$\nabla f(x) + \sum_{j=1}^m \mu_j \nabla g_j(x) = 0$$

2. Forcing to zero the multipliers of the nonactive constraints

$$\mu_j g_j(x) = 0, \quad \text{for } j = 1, \dots, m$$

These equalities are called the **complementarity conditions**, as they are equivalent to $g_j(x) < 0 \implies \mu_j = 0$.

Second, we have already seen how equality constraints can be replaced by pairs of inequalities (see 5.2.2), but it's simpler to

1. Use a single multiplier λ_i
2. Relax the nonnegativity condition on the multiplier: $\lambda_i \in \mathbb{R}$

Two opposite gradients with multipliers ≥ 0 can always be replaced by a single gradient with a multiplier equal to the difference of the original two. In fact

$$\begin{cases} \nabla g_{j'_i}(x) &= \nabla h_i(x) \\ \nabla g_{j''_i}(x) &= -\nabla h_i(x) \end{cases}$$

and

$$\mu_{j'_i} \nabla g_{j'_i}(x) + \mu_{j''_i} \nabla g_{j''_i}(x) = (\mu_{j'_i} - \mu_{j''_i}) \nabla h_i(x) = \lambda_i \nabla h_i(x)$$

with $\mu_{j'_i} \geq 0$, $\mu_{j''_i} \geq 0$ and $\mu_{j'_i} - \mu_{j''_i} = \lambda_i$ free in \mathbb{R} .

Theorem 18 (KKT-conditions). *Let $X = \{x \in \mathbb{R}^n \mid h_i(x) = 0, g_j(x) \leq 0\}$, with $i = 1, \dots, s$ and $j = 1, \dots, m$ and $f, h_i, g_j \in C^1(X)$ for $i = 1, \dots, s$ and $j = 1, \dots, m$. If x^* is a regular point in X and a local optimum point for f in X , then there exists free multipliers λ_i and nonnegative multipliers μ_j such that:*

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^s \lambda_i \nabla h_i(x^*) + \sum_{j=1}^m \mu_j \nabla g_j(x^*) &= 0 \\ \mu_j g_j(x^*) &= 0 \quad j = 1, \dots, m \\ h_i(x^*) &= 0 \quad i = 1, \dots, s \\ g_j(x^*) &\leq 0 \quad j = 1, \dots, m \\ \mu_j &\geq 0 \quad j = 1, \dots, m \end{aligned}$$

The “sifting”, therefore, does not consist in scanning all points and directions, but it consists of solving a system of equalities and inequalities, identifying all points that satisfy it.

The system consists of $n + s + m$ equations:

- n equations from Farkas' lemma
- s equations from equality constraints
- m equations from the complementarity conditions

imposed on $n + s + m$ variables:

- n variables to determine the solution x
- s variables to determine the multiplier vector λ
- m variables to determine the multiplier vector μ

and is therefore balanced.

The $2m$ inequalities

- m constraints on the solution
- m nonnegativity conditions on the multipliers

remove solutions, but do not decrease freedom degrees (*the number of solutions is probably finite*).