Game Theory

Theory

Christian Rossi

Academic Year 2024-2025

Abstract

The theory begins by examining the main assumptions that distinguish decision theory from interactive decision theory. While decision theory focuses on individual decision-making in isolation, interactive decision theory explores how multiple decision-makers interact, considering each other's potential actions.

In the context of non-cooperative games, the discussion extends to games represented in extensive form, where players make decisions at various points, and games with perfect information, where all players are fully informed of prior moves. The technique of backward induction is key in solving such games. Additionally, combinatorial games are explored, emphasizing their strategic complexity.

Zero-sum games are analyzed in terms of conservative values, where each player seeks to minimize potential losses. The concept of equilibrium in pure strategies is introduced, and this is extended to mixed strategies in finite games, invoking von Neumann's theorem. Finding optimal strategies and determining the value of finite games is achieved through the use of linear programming techniques.

The Nash non-cooperative model plays a central role in understanding strategic interactions. Nash equilibrium is discussed, focusing on the existence of equilibria in both pure and mixed strategies within finite games. Examples of potential games are provided, along with methods for identifying potential functions. Notable examples include congestion games, routing games, network games, and location games. Concepts such as the price of stability, price of anarchy, and correlated equilibria are explored to analyze the efficiency and stability of these systems.

Finally, the discussion shifts to cooperative games, defining key concepts such as the core, nucleolus, Shapley value, and power indices. Examples of cooperative scenarios illustrate how these concepts help to determine fair outcomes and power distribution among players.

Contents

CHAPTER 1

Introduction

1.1 Games

Games provide valuable models for simulating a variety of real-world situations.

Definition (*Game*)**.** A game is a structured process that includes the following components:

- A group of participants, referred to as players, with at least two members.
- An initial state or starting condition.
- A set of rules that define how players can act.
- A range of possible outcomes or end states.
- The preferences of each player concerning these potential outcomes.

1.2 Game Theory Assumptions

Game theory operates under the following key assumptions about the players involved:

1. *Self-interested.*
2. *Rational.*

1.2.1 Self-interest

Players are assumed to focus solely on their own preferences concerning the outcomes of the game. This is a mathematical assumption, not an ethical judgment. In fact, it is essential for defining what constitutes a rational choice within the framework of game theory.

1.2.2 Rationality

Definition (*Preference relation*)**.** Let X be a set. A preference relation on X is a binary relation $\preceq$ that satisfies the following properties for all $x, y, z \in X$:

- *Reflexive*: $x \preceq x$ (every element is at least as preferred as itself).
- *Complete*: $x \preceq y$ or $y \preceq x$ (any two elements can be compared).
- *Transitive*: if $x \preceq y$ and $y \preceq z$, then $x \preceq z$ (preferences are consistent across comparisons).

The transitive property ensures that preferences can be consistently ranked.

Definition (*Utility function*)**.** Given a preference relation $\preceq$ over a set X, a utility function representing $\preceq$ is a function $u : X \to \mathbb{R}$ such that:

$$u(x) \geq u(y) \Leftrightarrow x \preceq y$$

While a utility function may not always exist in specific cases, it does exist in general settings, particularly when X is finite. If a utility function does exist, there are infinitely many such functions, differing by any strictly increasing transformation of the original function.

Each player i is assigned a set X_i, representing all the choices available to them. Therefore, the set $X = xX_i$ over which the utility function u is defined represents the combined choices of all players.

Rationality assumptions The following assumptions define the rational behavior of players:

1. *Consistent preferences*: players can establish a preference relation over the game's outcomes, and this ordering is consistent.
2. *Utility representation*: players can define a utility function that represents their preference relations when needed.
3. *Consistent use of probability*: players apply the laws of probability consistently, including computing expected utilities and updating probabilities according to Bayes' rule.
4. *Understanding consequences*: players comprehend the outcomes of their actions, the impact on other players, and the resulting chain of consequences.
5. *Application of decision theory*: players use decision theory to maximize their utility. Given a set of alternatives X and a utility function u, each player seeks $\bar{x} \in X$ such that:

$$u(\bar{x}) \geq u(x) \qquad \forall x \in X$$

One significant consequence of these axioms is the principle of eliminating strictly dominated strategies: a player will not choose an action a if there exists another action b that yields a strictly better outcome, regardless of the actions of other players.

Example:
Consider the following games:

Gain	**Probability**
2500	33%
2400	66%
0	1%

Table 1.1: Game A

Gain	Probability
2500	0%
2400	100%
0	0%

Table 1.2: Game B

In a sample of 72 participants, 82% chose to play Game B, indicating a preference for certainty—characteristic of risk-averse individuals. According to expected utility theory, this decision is rational if:

$$u(2400) > \frac{33}{100}u(2500) + \frac{66}{100}u(2400)$$

This simplifies to:

$$\frac{34}{100}u(2400) > \frac{33}{100}u(2500)$$

Now consider the following alternatives:

Gain	Probability
2500	33%
0	67%

Table 1.3: Game C

Gain	Probability
2400	34%
0	66%

Table 1.4: Game D

In this new setup, 83% of participants preferred Game C, reflecting a preference for a larger gain even with a lower probability of success. Rationality in this scenario requires:

$$\frac{34}{100}u(2400) < \frac{33}{100}u(2500)$$

This contradicts the earlier experiment, where the opposite preference was observed. Such behavior violates the independence axiom in expected utility theory, which states that consistent preferences should hold under similar probabilistic transformations.

This contradiction is known as the Allais Paradox, demonstrating that individuals do not always act as fully rational decision-makers.

Example:
A group of players is asked to choose an integer between 1 and 100. The mean of all chosen numbers, M is then calculated. The objective of the game is to select the number closest to qM, where $0 < q < 1$.

A purely rational player would conclude that the optimal number to choose is 1, regardless of the value of q. However, this player is likely to lose.

For example, let $q = \frac{1}{2}$. Since $M \leq 100$, in the first step, it seems irrational to choose a number greater than $\frac{1}{2} \cdot 100$, as this is the initial target value based on the game's rules.

However, in the second step, assuming all players are rational and recognize that others are also rational, each player would realize that others will also choose numbers below 50. Therefore, the new logical step would be to pick a number less than $\left(\frac{1}{2}\right)^2 \cdot 100$.

This reasoning continues iteratively: at step n, it becomes irrational to choose a number greater than $\left(\frac{1}{2}\right)^n \cdot 100$. Ultimately, after enough steps, the only rational choice would appear to be selecting the smallest possible number (1).

Despite this reasoning, experiments show that the actual winning number is far higher than 1. In fact, the winning number tends to increase as the value of q increases, revealing that real-life behavior often deviates from purely rational game theory predictions.

1.3 Bimatrices

Conventionally, Player 1 selects a row, while Player 2 selects a column. This results in a pair of values that represent the utilities for Player 1 and Player 2, respectively. These options can be conveniently displayed in a bimatrix.

Example:
Consider the following bimatrix:

$$\begin{pmatrix} \begin{pmatrix} 8 & 8 \end{pmatrix} & \begin{pmatrix} 2 & 7 \end{pmatrix} \\ \begin{pmatrix} 7 & 2 \end{pmatrix} & \begin{pmatrix} 0 & 0 \end{pmatrix} \end{pmatrix}$$

In this example, Player 1's utilities are given by:

$$\begin{pmatrix} 8 & 2 \\ 7 & 0 \end{pmatrix}$$

Since the second row is strictly dominated by the first (i.e., Player 1's utility in the first row is higher for any choice by Player 2), Player 1 will rationally choose the first row. Similarly, Player 2 will select the first column, as it strictly dominates the second column.

While the principle of eliminating strictly dominated strategies may seem simplistic, it can lead to surprisingly powerful insights and outcomes.

Example:
Consider the following two games:

$$\begin{pmatrix} \begin{pmatrix} 10 & 10 \end{pmatrix} & \begin{pmatrix} 3 & 15 \end{pmatrix} \\ \begin{pmatrix} 15 & 3 \end{pmatrix} & \begin{pmatrix} 5 & 5 \end{pmatrix} \end{pmatrix}$$

$$\begin{pmatrix} \begin{pmatrix} 8 & 8 \end{pmatrix} & \begin{pmatrix} 2 & 7 \end{pmatrix} \\ \begin{pmatrix} 7 & 2 \end{pmatrix} & \begin{pmatrix} 0 & 0 \end{pmatrix} \end{pmatrix}$$

Note that in the first game, players have outcomes like $\begin{pmatrix} 10 & 10 \end{pmatrix}$ and $\begin{pmatrix} 15 & 3 \end{pmatrix}$, which individually seem to offer higher utilities than most outcomes in the second game. However, applying rational decision-making principles leads to a surprising result.

According to the principle of elimination of dominated strategies, players will end up choosing the outcome pair $\begin{pmatrix} 8 & 8 \end{pmatrix}$ in the second game because it dominates other available outcomes. This leads them to prefer the second game over the first game, despite the fact that the first game contains outcomes with higher individual utilities, like $\begin{pmatrix}10 & 10\end{pmatrix}$ and $\begin{pmatrix}15 & 3\end{pmatrix}$.

Now, consider the expanded form of the first game, which contains even more outcomes:

$$\begin{pmatrix} \begin{pmatrix}1 & 1\end{pmatrix} & \begin{pmatrix}11 & 0\end{pmatrix} & \begin{pmatrix}4 & 0\end{pmatrix} \\ \begin{pmatrix}0 & 11\end{pmatrix} & \begin{pmatrix}8 & 8\end{pmatrix} & \begin{pmatrix}2 & 7\end{pmatrix} \\ \begin{pmatrix}0 & 4\end{pmatrix} & \begin{pmatrix}7 & 2\end{pmatrix} & \begin{pmatrix}0 & 0\end{pmatrix} \end{pmatrix}$$

This expanded version of the first game includes all the outcomes from the second game, plus some additional options. However, rationality axioms suggest that in the first game, players should choose the outcome $\begin{pmatrix}10 & 10\end{pmatrix}$, which dominates the other possibilities.

Interestingly, in the second game, where fewer options are available, the players end up selecting $\begin{pmatrix}8 & 8\end{pmatrix}$. This leads to a paradoxical outcome: having fewer available actions can actually make players better off by simplifying the decision-making process and avoiding suboptimal choices.

Example:
Consider the rational outcomes of the following game.

$$\begin{pmatrix} \begin{pmatrix}0 & 0\end{pmatrix} & \begin{pmatrix}1 & 1\end{pmatrix} \\ \begin{pmatrix}1 & 1\end{pmatrix} & \begin{pmatrix}0 & 0\end{pmatrix} \end{pmatrix}$$

While we may not know the rational outcomes formally, it is clear that the preferred outcome for both players is $\begin{pmatrix}1 & 1\end{pmatrix}$. However, this leads to a coordination problem.

Both pairs of actions result in the same outcome $\begin{pmatrix}1 & 1\end{pmatrix}$, but there is no clear way for the players to distinguish between these two strategies. As a result, while the rational outcome is obvious, the players face difficulty coordinating on which specific actions to take to achieve it.

Example:
Consider a voting game with three players, each having the following preferences:

1. Player 1: $A \precneqq B \precneqq C$
2. Player 2: $B \precneqq C \precneqq A$
3. Player 3: $C \precneqq A \precneqq B$

Here, the notation $A \precneqq B$ indicates that Player 1 prefers $B \preceq A$, but not vice versa. The winner is determined by the alternative that receives the most votes. However, if there is a tie among three different votes, the alternative chosen by Player 1 will win.

Let's now analyse the rational outcome of the game through the elimination of dominated actions:

- Alternative A is a weakly dominant strategy for Player 1.
- Players 2 and 3 have their least preferred choice as a weakly dominated strategy.

To avoid their worst outcome, Player 2 retains options B and C (ordered in rows), while Player 3 keeps C and A (ordered in columns). Player 1 will consistently choose A. This simplifies the game to a 2×2 table with the following outcomes:

	C	**A**
B	A	A
C	C	A

Since $C \precneqq A$ for both Players 2 and 3, they will choose the outcome in the second row and first column, leading to the final result being C, which is the worst outcome for Player 1.

Example:
Consider the classic Chicken Game. Two cars are driving toward each other on a narrow road, and there isn't enough room for both cars to pass. If both cars keep going straight, they will crash, but if at least one deviates, the crash is avoided. The payoff bimatrix for the game is:

$$\begin{pmatrix} \begin{pmatrix}-1 & -1\end{pmatrix} & \begin{pmatrix}1 & 10\end{pmatrix} \\ \begin{pmatrix}10 & 1\end{pmatrix} & \begin{pmatrix}-10 & -10\end{pmatrix} \end{pmatrix}$$

The payoffs are as follows:

- Both players receive a utility of -1 if they both deviate.
- A player earns 10 for going straight when the other deviates, while the deviating player receives 1.
- If both go straight and crash, they both receive -10.

The best outcomes are either $\begin{pmatrix}10 & 1\end{pmatrix}$ and $\begin{pmatrix}1 & 10\end{pmatrix}$. However, there is no way to decisively determine which of these two outcomes will occur, as both are equally viable equilibria.

CHAPTER 2

Extensive form games

2.1 Games tree

Example:
Three politicians are tasked with deciding whether to raise their salaries. The voting is public and happens sequentially. Each politician prefers a salary increase but also wants to vote against it to maintain public support.

The optimal outcome for each politician is to receive a salary increase while voting against it. The game has the following characteristics:

1. The voting process is sequential, with politicians voting one after another.
2. Every possible situation is fully known to all players: they are aware of the entire history and all possible future actions.
3. The final outcome is determined by the majority vote.

To represent such a game we can use a tree, where each branch represents a player's vote: a YES vote goes left, and a NO vote goes right. The utilities depend on each player's vote and the final outcome:

1. YES, but no raise.
2. NO, and no raise.
3. YES, and a raise.
4. NO, and a raise.

The corresponding decision tree looks like this:

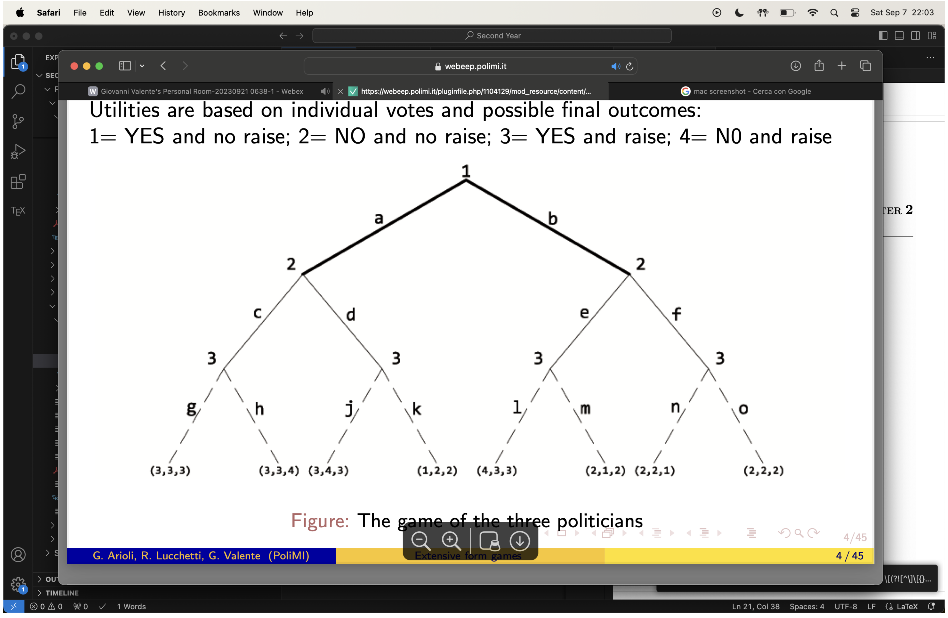


Figure 2.1: Voting game tree

This is an example of a game with perfect information, where each player is fully aware of all prior events.

Example:
Two players, Player 1 and Player 2, must decide sequentially whether to participate in a game. If both choose to play, a coin is flipped (random component R): Player 1 wins if it lands heads, and Player 2 wins if it lands tails. The corresponding decision tree is as follows:

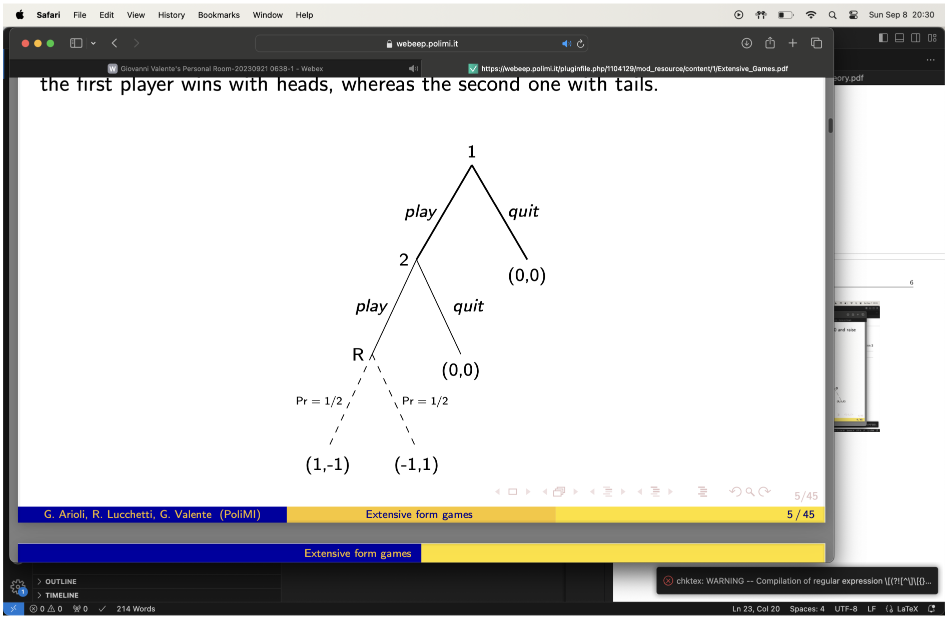


Figure 2.2: Chance game tree

Definition (*Finite directed graph*)**.** A finite directed graph is a pair (V, E) where:

- V is a finite set, called the set of vertices.
- $E \subset V \times V$ is a set of ordered pairs of vertices, called the directed edges.

Definition (*Path*)**.** A path from a vertex v_1 to a vertex v_{k+1} is a finite sequence of vertices and edges $v_1, e_1, v_2, \ldots, v_k, e_k, v_{k+1}$ such that $e_i \neq e_j$ if $i \neq j$ and $e_j = (v_j, v_{j+1})$.

The number k is called the length of the path.

Definition (*Oriented graph*)**.** An oriented graph is a finite directed graph with no bidirectional edges. That is, for all vertices v_j and v_k, at most one of (v_j, v_k) and (v_k, v_j) can be an edge in the graph.

Definition (*Tree*)**.** A tree is a triple (V, E, x_0) where (V, E) is an oriented graph, and x_0 is a vertex in V such that there is a unique path from x_0 to any other vertex $x \in V$.

Definition (*Child*)**.** A child of a vertex v is any vertex x such that $(v, x) \in E$.

Definition (*Leaf*)**.** A vertex is called a leaf if it has no children.

We say that the vertex x follows the vertex v if there is a path from v to x.

2.2 Extensive games

Definition (*Extensive form game with perfect information*)**.** An extensive form game with perfect information consists of:

1. A finite set $N = \{1, \ldots, n\}$ of players.
2. A game tree (V, E, x_0).
3. A partition of the non-leaf vertices into sets $P1, P2, \ldots, Pn + 1$.
4. A probability distribution for each vertex in P_{n+1}, defined on the edges from that vertex to its children.

The game includes the following:

1. The set P_i, for $i \leq n$, consists of nodes where Player i must choose a child of v, representing a possible move by that player.
2. P_{n+1} is the set of nodes where chance plays a role (i.e., random events). Here, $n+1$ refers to the players plus the random component. P_{n+1} can be empty, indicating no random events in the game.
3. When P_{n+1} is empty, the n players only have preferences over the leaves, meaning no random component influences the outcome, so a utility function isn't required.

2.2.1 Extensive game solution

To determine the optimal outcome, we apply the axioms of rationality.

> **Example:**
> For the voting game described earlier:

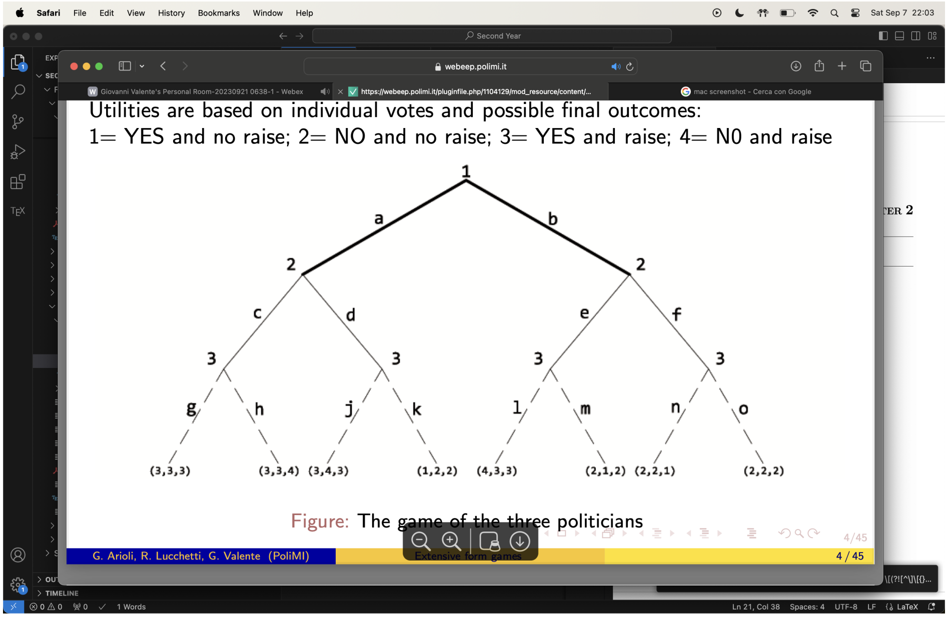


Figure 2.3: Voting game tree

We can identify the preferences as follows:

- For Player 3:
  - $h = \begin{pmatrix} 3 & 3 & 4 \end{pmatrix}$ over $g = \begin{pmatrix} 3 & 3 & 3 \end{pmatrix}$.
  - $j = \begin{pmatrix} 3 & 4 & 3 \end{pmatrix}$ over $k = \begin{pmatrix} 1 & 2 & 2 \end{pmatrix}$.
  - $l = \begin{pmatrix} 4 & 3 & 3 \end{pmatrix}$ over $m = \begin{pmatrix} 2 & 1 & 2 \end{pmatrix}$.
  - $o = \begin{pmatrix} 2 & 2 & 2 \end{pmatrix}$ over $n = \begin{pmatrix} 2 & 2 & 1 \end{pmatrix}$.
- For Player 2:
  - $d = \begin{pmatrix} 3 & 4 & 3 \end{pmatrix}$ over $c = \begin{pmatrix} 3 & 3 & 4 \end{pmatrix}$.
  - $e = \begin{pmatrix} 4 & 3 & 3 \end{pmatrix}$ over $f = \begin{pmatrix} 2 & 2 & 2 \end{pmatrix}$.
- For Player 1:
  - $b = \begin{pmatrix} 4 & 3 & 3 \end{pmatrix}$ over $a = \begin{pmatrix} 3 & 4 & 3 \end{pmatrix}$.

Thus, the optimal outcome is found.

Definition (*Length*)**.** The length of a game is defined as the length of the longest path in the game tree.

Using decision theory and the assumption of rationality:

- Rationality assumption 5 allows us to solve games of length 1.
- Rationality assumption 4 allows us to solve games of length $i + 1$ if all games of length at most i have already been solved.

This iterative process is called backward induction, where we work backwards from the leaves of the tree to the root to determine the optimal sequence of actions.

Theorem 2.2.1 (First rationality theorem)**.** *The rational outcomes of a finite game with perfect information are those determined by the backward induction procedure.*

Backward induction can be applied because every vertex v in the game is the root of a subgame consisting of all the vertices that follow v. These subgames are derived from the original game.

Example:
For the chance game:

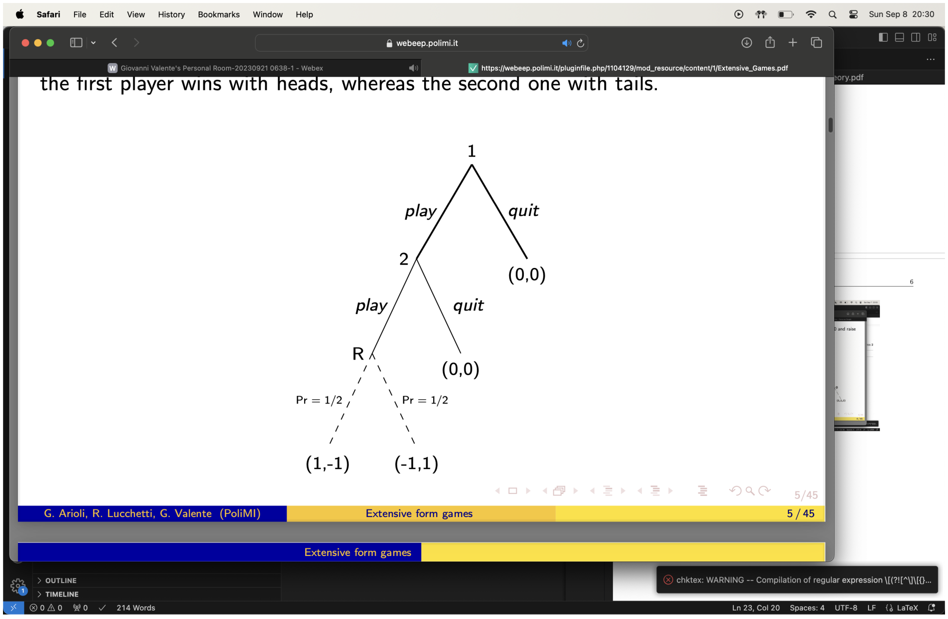


Figure 2.4: Chance game tree

The outcomes obtained through backward induction are $\begin{pmatrix}4 & 3\end{pmatrix}$ and $\begin{pmatrix}3 & 4\end{pmatrix}$. Player 2 has no strict preference between $\begin{pmatrix}4 & 3\end{pmatrix}$ and $\begin{pmatrix}0 & 3\end{pmatrix}$, indicating that in general, the solutions may not be unique.

2.3 Chess theorem

Theorem 2.3.1 (Von Neumann)**.** *In the game of chess, one and only one of the following alternatives holds:*

1. *The White has a way to win, no matter what the Black does.*
2. *White has a strategy to guarantee a win, regardless of what Black does.*
3. *Both White and Black can force at least a draw, regardless of the opponent's actions.*

Proof. Assume the game has a finite length of $2K$ moves, where each player makes K moves. Let a_i represent White's move at their i-th stage, and b_i represent Black's corresponding move.
The first possibility in the theorem can be formulated as follows:

$$\exists a_1 : \forall b_1 \exists a_2 : \forall b_2 \dots \exists a_K : \forall b_K \implies \text{white wins}$$

Now, suppose this is false. Then the negation is:

$$\forall a_1 \exists b_1 : \forall a_2 : \exists b_2 : \dots \forall a_K : \exists b_K \implies \text{white does not win}$$

This means Black has the possibility to prevent White from winning, ensuring at least a draw.

If White does not have a winning strategy, then Black can secure at least a draw. Similarly, if Black does not have a winning strategy, then White can secure at least a draw. Therefore, if neither of the first two possibilities holds, the third one must be true. □

2.3.1 Extension

Von Neumann's theorem can be extended to any finite game of perfect information where the possible outcomes are either a win for one player or a tie.

Corollary 2.3.1.1. *Consider a finite, perfect information game with two players, where the only possible outcomes are a win for one of the players or a tie. Then, exactly one of the following holds:*

1. *Player 1 has a winning strategy, no matter what the second player does.*

2. *Player 2 has a winning strategy, no matter what the second player does.*

The possible solutions for a game are classified as follows:

- *Very weak solution*: the game has a rational outcome, but it is not accessible in practice, as with chess.

- *Weak solution*: the outcome of the game is known, but the method to achieve it is not generally understood.

- *Solution*: there exists an algorithm that can determine the outcome.

Example:
In the game of Chomp, players take turns removing tiles from a rectangular grid, where removing a tile also removes all tiles to its right and above. The game ends when a player is unable to make a move, and the last player to play wins.

In this scenario, we can distinguish between two types of solutions: a definite solution occurs when the grid is square, while a weak solution arises with a rectangular grid. Let's consider a specific configuration, illustrated in the provided figure.

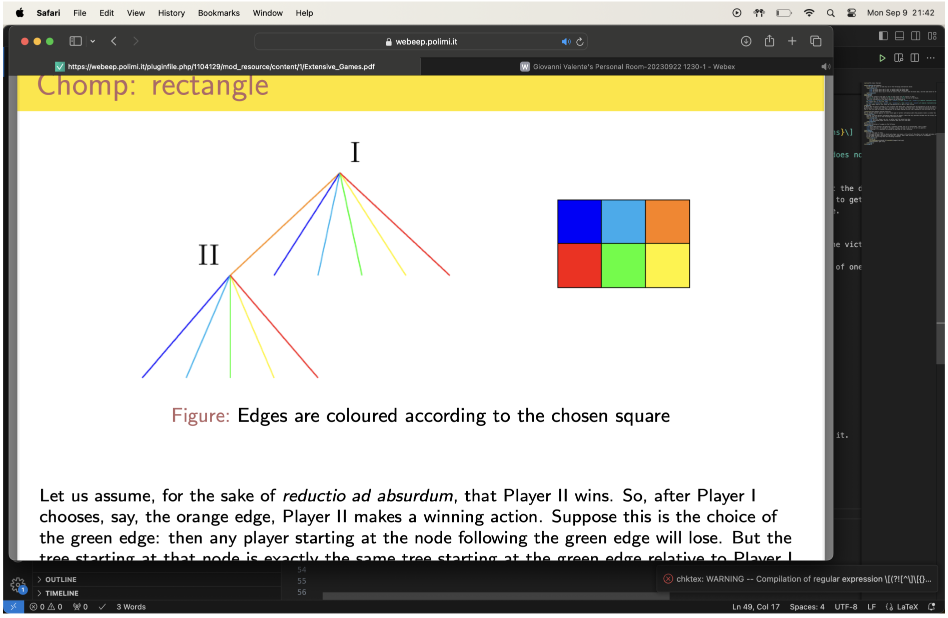


Figure 2.5: Rectangular chomp

To explore the outcome of the game, let's assume for contradiction that Player 2 is the winner. Suppose Player 1 chooses a particular edge (represented in orange), and then Player 2 makes a winning move by selecting another edge (the green edge). If we examine the game state following this move, any player starting from the node that follows the green edge is destined to lose.

However, this situation mirrors the tree of outcomes that begins at the green edge from Player 1's perspective. This means that Player 1 has a move that guarantees a victory against Player 2's position. Hence, the assumption that Player 2 can win leads to a contradiction, implying that Player 1 must be the one who wins the game.

Thus, we conclude that Player 1 has a winning strategy in this instance of the game of Chomp.

2.4 Impartial combinatorial games

Definition (*Impartial combinatorial game*)**.** An impartial combinatorial game is defined by the following characteristics:

1. There are two players who alternate turns.
2. The game consists of a finite number of positions.
3. Both players adhere to the same set of rules.
4. The game concludes when no further moves can be made.
5. The outcome of the game is not influenced by chance.
6. In the classical version of the game, the winner is the player who leaves the opponent with no available moves; in the misère version, the objective is reversed.

Example:
Several examples illustrate impartial combinatorial games:

- k piles of cards: each player, on their turn, can take any number of cards (at least one) from a single pile.
- k piles of cards with restrictions: each player can take any number of cards (at least one) from no more than $j < k$ piles during their turn.
- k cards in a row: players can take j_l cards on their turn.

In all these variations, the player who is left without cards loses. In the first two examples, the positions can be represented as $(n_1, \cdots, n_k)$, where each n_i is a non-negative integer corresponding to the number of cards in each pile. In the third example, the positions are characterized by all non-negative integers less than or equal to k.

To solve impartial combinatorial games, we begin by partitioning the set of all possible positions (which are finite in number) into two distinct categories:

1. P-positions: these are positions where the previous player has a winning strategy, meaning they are losing positions for the player who is about to move.
2. N-positions: these are positions where the next player has a winning strategy, indicating they are winning positions for the player who is about to move.

It is important to note that the current state of the game is what matters, rather than which player is designated to move.

Partition rules The rules for the partitions are:

- The terminal position $(0, 0, \cdots, 0)$ is classified as a P-position. This is a losing position because the player has no cards left to play.
- From any P-position, only N-positions can be reached. This means that the next player is guaranteed to have a winning strategy.
- From any N-position it is possible (but not obligatory) to move to a P-position. The player in an N-position can make a move that leads their opponent to a losing position.

Therefore, the player who starts from an N-position is assured of a victory, given that they play optimally.

2.4.1 Nim game

The Nim game is characterized by a tuple $(n_1, \cdots, n_k)$, where each n_i is a positive integer. During their turn, each player must choose one pile n_i and replace it with $\hat{n}_i$, ensuring that $\hat{n}_i < n_i$. The player who reduces the position to $(0, \ldots, 0)$ wins.

Therefore, each player's action involves removing cards from a single pile with the objective of clearing the entire table.

Theorem 2.4.1 (Bouton)**.** *A position $(n_1, n_2, \ldots, n_k)$ in the Nim game is a P-position if and only if:*

$$n_1 \oplus n_2 \oplus \cdots \oplus n_k = 0$$

Proof. The terminal position $(0, 0, \ldots, 0)$ is a P-position corresponding to a Nim-sum of zero.

If the Nim-sum $n_1 \oplus n_2 \oplus \cdots \oplus n_k = 0$, any subsequent position will have a non-zero Nim-sum. Assume the next position is $(\hat{n}_1, n_2, \ldots, n_k)$ such that $\hat{n}_1 \oplus n_2 \oplus \cdots \oplus n_k = 0$. Then we would have:

$$n_1 \oplus n_2 \oplus \cdots \oplus n_k = 0$$

Which, by the cancellation law, implies $\hat{n}_1 = n_1$. This is a contradiction, as the game rules stipulate that $\hat{n}_1 < n_1$.

Conversely, if $n_1 \oplus n_2 \oplus \cdots \oplus n_k \neq 0$, it is possible to move to a position with a zero Nim-sum. Let $z = n_1 \oplus n_2 \oplus \cdots \oplus n_k \neq 0$. Identify a pile where the binary representation of z has a 1 in its leftmost column. Change that digit to 0 and adjust the digits to the right, leaving unchanged the digits that correspond to 0. This operation produces a new number that is smaller than the original. □

Example:
Consider the configuration with a non-zero Nim-sum: $4, 6, 5$. By removing a card from the first pile, we can reach the configuration $3, 6, 5$, which has a zero Nim-sum. Thus, there are three initial advantageous moves, one available for each pile.

2.4.2 Conclusions

Games with perfect information can typically be resolved through backward induction. However, this method is primarily effective for relatively simple games due to the constraints of limited rationality. Depending on the specifics of the game, we may arrive at varying degrees of solutions.

2.5 Strategies

In backward induction, a specific move must be identified at every node. Let P_i denote the set of all nodes at which player i is required to make a decision.

Definition (*Pure strategy*)**.** A pure strategy for player i is defined as a function on the set P_i, which associates each node v in P_i with a child node x, or equivalently, an edge (v, x).

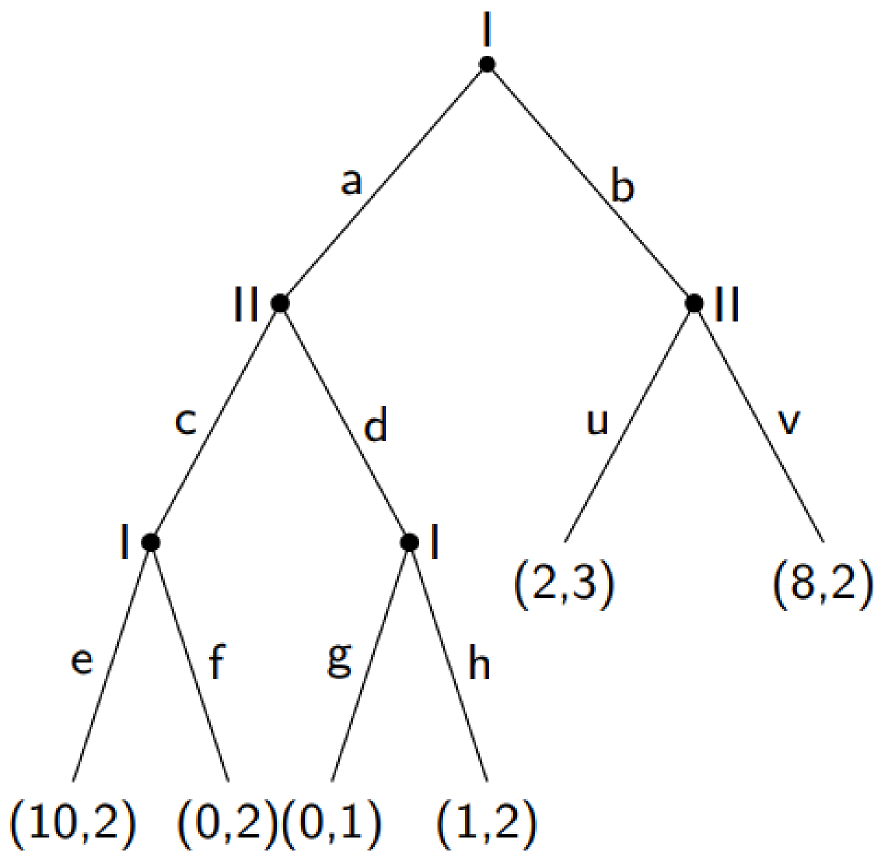
Definition (*Mixed strategy*)**.** A mixed strategy refers to a probability distribution over the set of pure strategies.

When a player possesses n pure strategies, the collection of their mixed strategies is represented as:

$$\sum_n = \left\{ p = (p_1, \ldots, p_n) | p_i \geq 0 \text{ and } \sum p_i = 1 \right\}$$

Here, $\sum_n$ forms the fundamental simplex in n-dimensional space.

Example:
Consider the following tree:



The strategies depicted in the tree are shown below:

	cu	cv	du	dv
aeg	(10,2)	(10,2)	(0,1)	(0,1)
aeh	(10,2)	(10,2)	(1,2)	(1,2)
afg	(0,2)	(0,2)	(0,1)	(0,1)
afh	(0,2)	(0,2)	(1,2)	(1,2)
beg	(2,3)	(8,2)	(2,3)	(8,2)
beh	(2,3)	(8,2)	(2,3)	(8,2)
bfg	(2,3)	(8,2)	(2,3)	(8,2)
bfh	(2,3)	(8,2)	(2,3)	(8,2)

Note that Player 1's strategies are listed in the rows, while Player 2's are in the columns. All combinations are included, even if they are equivalent (e.g., strategies b– for Player 2). The table may contain repeated pairs, as different strategies can lead to the same outcomes.

- *Extensive form*: the various moves of the players are presented sequentially.

- *Strategic form*: all players' strategies are presented simultaneously.

Theorem 2.5.1 (Von Neumann on strategies)**.** *In the game of chess, one of the following scenarios must hold:*

1. *White has a winning strategy.*
2. *Black has a winning strategy.*
3. *Both players possess a strategy that guarantees at least a tie.*

The first outcome occurs when there exists a row containing all winning elements. The second outcome arises when there is a column consisting of all winning elements. The third outcome features mixed results, including ties, but does not encompass all three outcomes in a single row or column.

If $P_i = \{v_1, \ldots, v_k\}$ and v_j has n_j children, then the total number of strategies available to Player i is $n_1 \cdot n_2 \cdot \cdots \cdot n_k$. This illustrates that the number of strategies, even in short games, is typically quite substantial.

Example:
In the game of Tic-Tac-Toe, if the game is halted after three moves, the first player has $9 \cdot 7^{(8\times 9)}$ strategies available (not accounting for symmetrical configurations).

2.6 Games with imperfect information

In certain scenarios, players must make their moves simultaneously, which prevents them from having complete knowledge of each other's actions. This situation can still be represented using a game tree.

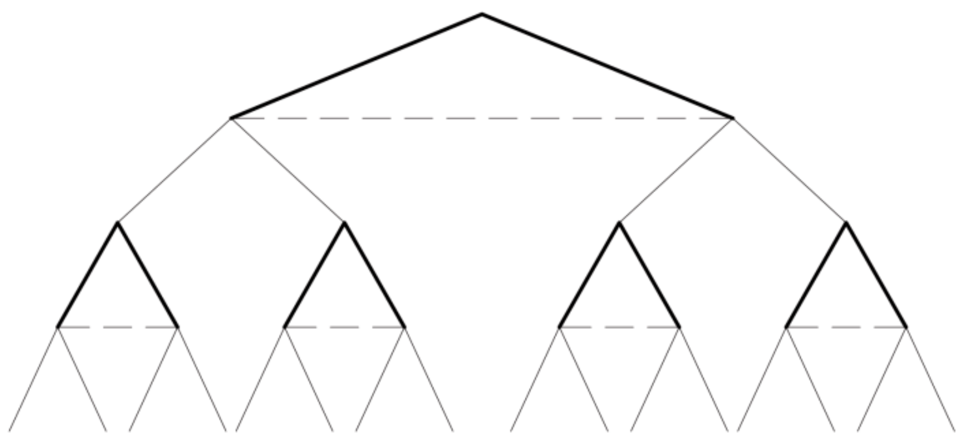


Figure 2.6: Tree with imperfect information

The dashed lines in the figure indicate that a player does not know exactly which vertex they occupy.

Definition (*Information set*)**.** An information set for Player i is a pair $(U_i, A(U_i))$ satisfying the following conditions:

1. $U_i \subset P_i$ is a non-empty set of vertices $v_1, \cdots, v_k$.

2. Each vertex $v_j \in U_i$ has the same number of children.

3. $A_i(U_i)$ is a partition of the children of $v_1 \cup \cdots \cup v_k$ such that each element of the partition contains exactly one child from each vertex v_j.

Thus, Player i knows they are in U_i but cannot determine the exact vertex. The partition defines the choice function, indicating that each set in $A_i(U_i)$ corresponds to an available move for the player (graphically, this represents the same choice, or edge, emanating from different vertices).

Definition (*Extensive form game with imperfect information*)**.** An extensive form game with imperfect information is characterized by the following components:

1. A finite set $N = \{1, \ldots, n\}$ of players.

2. A game tree (V, E, x_0).

3. A partition comprising sets $P_1, P_2, \ldots, P_{n+1}$ of the non-leaf vertices.

4. A partition $(U_i^j), j = 1, \ldots, ki$ of the set P_i, for all i, with (U_i^j, A_i^j) being the information set for all players i at all vertices j (having the same number of children).

5. A probability distribution defined for each vertex in P_{n+1} on the edges leading to its children.

6. An n-dimensional vector assigned to each leaf.

It is important to note that if the partition consists of only a single vertex, then a game with imperfect information effectively becomes a game with perfect information.

Definition (*Pure strategy*)**.** A pure strategy for player i in an imperfect information game is a function defined over the collection $\mathcal{U}$ of their information sets, assigning to each $U_i \in \mathcal{U}$ an element from the partition $A(U_i)$.

Definition (*Mixed strategy*)**.** A mixed strategy is defined as a probability distribution over the pure strategies.

A game of perfect information is a specific type of imperfect information game where all information sets for all players are singletons.

CHAPTER 3

Zero sum games

3.1 Introduction

Definition (*Zero sum game*)**.** A two-player zero-sum game in strategic form can be described as a triplet $(X, Y, f : X \times Y \to \mathbb{R})$, where:

- X is the strategy space of Player 1.
- Y is the strategy space of Player 2.
- $f(x, y)$ represents the payoff Player 1 receives from Player 2 when they play strategies x and y, respectively.

Since this is a zero-sum game, Player 2's utility function g is defined as the negative of Player 1's utility function:

$$g = -f$$

In the case where the strategy spaces are finite, i.e., $X = \{1, 2, \ldots, n\}$, $Y = \{1, 2, \ldots, m\}$, the game can be represented by a payoff matrix P. In this matrix, Player 1 chooses a row i, and Player 2 chooses a column j:

$$\begin{pmatrix} p_{11} & \cdots & p_{1m} \\ \cdots & p_{ij} & \cdots \\ p_{n1} & \cdots & p_{nm} \end{pmatrix}$$

Here, p_{ij} denotes the payment Player 2 makes to Player 1 when they select strategies i and j, respectively.

To determine the optimal strategy, both players can employ conservative reasoning:

- Player 1 can ensure a minimum payoff of $v_1 = \max_i \min_j p_{ij}$.
- Player 2 can limit their losses to at most $v_2 = \min_j \max_i p_{ij}$.

These values, v_1 and v_2, are known as the conservative values for Player 1 and Player 2, respectively.

Example:
Consider the following game with the payoff matrix:

$$\begin{pmatrix} 4 & 3 & 1 \\ 7 & 5 & 8 \\ 8 & 2 & 0 \end{pmatrix}$$

For Player 1, the minimum values for each row are $\begin{pmatrix} 1 & 5 & 0 \end{pmatrix}$ and then we choose the maximum between them. The maximum of these is $v_1 = 5$, so the conservative value for Player 1 is 5.

For Player 2, the maximum values for each column are $\begin{pmatrix} 8 & 5 & 8 \end{pmatrix}$ and then we choose the mimimum between them. The minimum of these is $v_2 = 5$, so the conservative value for Player 2 is also 5.

Thus, the rational outcome of the game is 5, with Player 1 choosing row $\bar{i} = 2$ and Player 2 choosing column $\bar{j} = 2$.

3.1.1 Generalization

In more general cases where the strategy spaces X and Y are not finite, a similar reasoning applies. Let (X, Y, f) describe the game, where X and Y are arbitrary strategy sets. The conservative values can be defined as follows:

- Player 1: $v_1 = \sup_x \inf_y f(x, y)$.
- Player 2: $v_2 = \inf_y \sup_x f(x, y)$.

If $v_1 = v_2$, the game has a value $v = v_1 = v_2$.

3.2 Rationality

Now, suppose the following holds:

- $v_1 = v_2 = v$.
- There exists a row $\bar{i}$ such that $p_{\bar{i}\bar{j}} \geq v_1 = v$ for all j.
- There exists a column $\bar{j}$ such that $p_{\bar{i}\bar{j}} \leq v_2 = v$ for all i.

In this case, $p_{\bar{i}\bar{j}} = v$, and this value represents the rational outcome of the game.

The strategies $\bar{i}$ and $\bar{j}$ are optimal because:

- Player 1 cannot guarantee more than v_2, the conservative value of Player 2.
- Player 2 cannot pay less than v_1, the conservative value of Player 1.

Thus, $\bar{i}$ maximizes the function $\alpha(i) = \min_j p_{ij}$, and $\bar{j}$ minimizes the function $\beta(j) = \max_i p_{ij}$.

3.2.1 Existence of a rational outcome

To demonstrate the existence of a rational outcome in a zero-sum game, we need to establish the following:

1. *Equality of conservative values*: the conservative values of both players agree, i.e., $v_1 = v_2$.
2. *Existence of an optimal strategy for Player 1*: there exists a strategy $\bar{x}$ such that:
$$v_1 = \inf_y f(\bar{x}, y)$$
This ensures that $\bar{x}$ is an optimal strategy for Player 1.
3. *Existence of an optimal strategy for Player 2*: there exists a strategy $\bar{y}$ such that:
$$v_2 = \sup_x f(x, \bar{y})$$
This ensures that $\bar{y}$ is an optimal strategy for Player 2.

In the case where the strategy spaces are finite, such optimal strategies $\bar{x}$ and $\bar{y}$ always exist. Therefore, proving the existence of a rational outcome is equivalent to demonstrating the equality of the conservative values, i.e., $v_1 = v_2$.

Theorem 3.2.1 (Von Neumann)**.** *There always exists a rational outcome for a finite two-player zero-sum game, as described by its payoff matrix P.*

This fundamental result, known as the Minimax theorem, guarantees that in every finite zero-sum game, the conservative values for both players coincide, and optimal strategies exist for both players, leading to a rational outcome.

Proof. Suppose, without loss of generality, that all p_{ij} in the matrix P are positive. Consider the column vectors $p_1, \ldots, p_m \in \mathbb{R}^n$, and let C denote their convex hull. Define the set

$$Q_t = \{x \in \mathbb{R}^n : x_i \leq t\}$$

and

$$v = \sup\{t \geq 0 : Q_t \cap C = \emptyset\}$$

Since $\operatorname{int} Q_v \cap C = \emptyset$, the sets Q_v and C can be separated by a hyperplane. Hence, there exist coefficients $\bar{x}_1, \ldots, \bar{x}_n$, with some $\bar{x}_i \neq 0$, and $b \in \mathbb{R}$ such that:

$$(\bar{x}, u) = \sum_{i=1}^{n} \bar{x}_i u_i \leq b \leq \sum_{i=1}^{n} \bar{x}_i w_i = (\bar{x}, w)$$

for all $u = (u_1, \ldots, u_n) \in Q_v$ and $w = (w_1, \ldots, w_n) \in C$.

Since all $\bar{x}_i$'s must be non-negative, we can assume $\sum \bar{x}_i = 1$. Additionally, $b = v$, since $\bar{v} := (v, \ldots, v) \in Q_v$, and

$$(\bar{x}, \bar{v}) = \sum_i \bar{x}_i v = v \sum_i \bar{x}_i = v$$

Therefore, $b \geq v$. If $b > v$, by choosing a small $a > 0$ such that $b \geq v + a$, we would have

$$\sup\left\{\sum_{i=1}^{n} \bar{x}_i u_i : u \in Q_{v+a}\right\} < b$$

which would imply $Q_{v+a} \cap C = \emptyset$, contradicting the definition of v.

Next, since $Q_v \cap C \neq \emptyset$, let $\bar{w} = \sum_{j=1}^{m} \bar{y}_j p_j$ (as C is convex) for some $\bar{y} = (\bar{y}_1, \ldots, \bar{y}_m) \in \Sigma_m$. Since $\bar{w} \in Q_v$, we have $\bar{w}_i \leq v$ for all i.

We now show that $\bar{x}$ is optimal for Player 1, $\bar{y}$ is optimal for Player 2, and v is the value of the game.

For Player 1, since $(\bar{x}, w) \geq v$ for every $w \in C$ by the separation result, and since each column $p_{\cdot j} \in C$, we have

$$(\bar{x}, p_{\cdot j}) \geq v, \quad \text{for all } j$$

For Player 2, consider $w = \sum_{j=1}^{m} \bar{y}_j p_j \in Q_v \cap C$ as before. Then, $w_i = \bar{y} p_{i\cdot}$, and since $w \in Q_v$, it follows that $w_i \leq v$ for every i. Hence, we have:

$$v \geq w_i = \bar{y} p_{i\cdot}$$

$\square$

Von Neumann's theorem guarantees that even when a finite zero-sum game has no solutions in pure strategies, the following holds:

- For Player 1, there exists a mixed strategy, represented as a probability distribution $\mathbf{x} = \begin{pmatrix} x_1 & \ldots & x_n \end{pmatrix}$, over her pure strategies. For every column j:

$$(x, p_{\cdot j}) = \sum_{i=1}^{n} x_i p_{ij} = x_1 p_{1j} + x_2 p_{2j} + \cdots + x_n p_{nj} \geq v$$

- For Player 2, there exists a mixed strategy, represented as a probability distribution $\mathbf{y} = \begin{pmatrix} y_1 & \ldots & y_n \end{pmatrix}$, over her pure strategies. For every row i:

$$(y, p_{i\cdot}) = \sum_{j=1}^{n} y_j p_{ij} = y_j p_{i1} + y_2 p_{i2} + \cdots + y_m p_{im} \leq v$$

The constant v is the value of the game under mixed strategies. Player 1 aims to maximize v, while Player 2 seeks to minimize it.

3.3 Optimality

Let X and Y be arbitrary sets. Suppose:

1. $v_1 = v_2 := v$.
2. There exists a strategy $\bar{x}$ such that $f(\bar{x}, y) \geq v$ for all $y \in Y$.
3. There exists a strategy $\bar{y}$ such that $f(x, \bar{y}) \leq v$ for all $x \in X$.

Then:

- v is the rational outcome of the game.
- $\bar{x}$ is an optimal strategy for Player 1.
- $\bar{y}$ is an optimal strategy for Player 2.

It follows that $\bar{x}$ is optimal for Player 1 since it maximizes $\alpha(x) = \inf_y f(x, y)$, while $\bar{y}$ is optimal for Player 2 since it minimizes $\beta(y) = \sup_x f(x, y)$. The values $\alpha(x)$ and $\beta(y)$ represent the best responses for the players if they knew the opponent's strategy.

3.3.1 Conservative values different or equal

Proposition. Let X and Y be nonempty sets, and let $f : X \times Y \to \mathbb{R}$ be an arbitrary real-valued function. Then:

$$v_1 = \sup_x \inf_y f(x,y) \leq \inf_y \sup_x f(x,y) = v_2$$

Proof. By definition, for all $x \in X$ and $y \in Y$:

$$\inf_y f(x,y) \leq f(x,y) \leq \sup_x f(x,y)$$

Thus, for all x and y, it holds that:

$$\alpha(x) = \inf_y f(x,y) \leq \sup_x f(x,y) = \beta(y)$$

Taking the supremum over x and the infimum over y, we conclude:

$$\sup_x \alpha(x) \leq \inf_y \beta(y)$$

□

As a result, it follows that for any game, $v_1 \leq v_2$.

Example:
Consider the game of rock-paper-scissors, represented by the following matrix:

$$\begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$$

The conservative values are not the same: in fact, $v_1 = -1$ and $v_2 = 1$.

Here, $v_1 = -1$ and $v_2 = 1$, indicating the conservative values are not equal. Therefore, no single deterministic strategy guarantees a win. However, in a repeated game with mixed strategies, both players should play each option with equal probability (one-third of the time), resulting in an expected utility of zero for both players.

3.3.2 Conservative values not equal

When the conservative values differ, mixed strategies must be considered. In this case, the strategy spaces for both players are probability distributions:

$$\sum_k = \left\{ x = (x_1, \ldots, x_k) | x_i \geq 0 \text{ and } \sum_{i=1}^{k} x_i = 1 \right\}$$

Here, $k = n$ for Player 1 and $k = m$ for Player 2. The utility function is extended to:

$$f(x,y) = \sum_{i=1,\ldots,n, j=i,\ldots,m} x_i y_j p_{ij} = (x, Py)$$

Thus, the mixed extension of the original game is given by:

$$\left(\sum_n, \sum_m, f(x,y) = (x, Py) \right)$$

3.3.3 Pure strategies optimality

Theorem 3.3.1. *If a player knows the strategy being used by the opposing player, they can always adopt a pure strategy to achieve the best possible outcome.*

This means that once one player's choice is fixed, the optimization problem reduces to a linear problem over a simplex, given that the utility function in such a game is bilinear.

Proof. Consider Player 2, who knows that Player 1 is using a mixed strategy $\bar{x}$. Player 2's task is then to minimize the function:

$$f(\bar{x}, y) = (\bar{x}, Py)$$

over the simplex $\sum_m$ (the set of mixed strategies for Player 2). The optimal value will be attained at one of the vertices e_j of the simplex, which corresponds to a pure strategy. Thus, Player 2 can use a pure strategy to achieve the optimal outcome. □

Given a payoff matrix P, let the column vector corresponding to the j-th pure strategy be denoted as $p_{\cdot j}$, and the row vector corresponding to the i-th pure strategy as $p_{i\cdot}$, respectively. The payoff of the first player in the mixed extension of the game is given by:

$$f(x, y) = (x, Py)$$

The previous theorem implies that, to verify the existence of a rational outcome for the game, we need to show the existence of mixed strategies $\bar{x}$ and $\bar{y}$, as well as a value v, such that:

- $(\bar{x}, P_{e_j}) = (\bar{x}, p_{\cdot j})$ for every column j.
- $(e_i, p_{i\cdot}\bar{y}) \leq v$ for every row i.

Here, e_j is the j-th strategy of Player 2, and e_i is the i-th strategy of Player 1.

Example:
Consider the following payoff matrix of a game:

$$P = \begin{pmatrix} 7 & 1 & 4 & 9 \\ 3 & 10 & 6 & 2 \\ 4 & 5 & 3 & 0 \end{pmatrix}$$

The third row is trictly dominated by a convex combination of the first two. Thus, the payoff matrix can be reduced to:

$$P = \begin{pmatrix} 7 & 1 & 4 & 9 \\ 3 & 10 & 6 & 2 \end{pmatrix}$$

Now the set C is the polygon with the following vertices:

$$\begin{pmatrix} 7 & 3 \end{pmatrix} \qquad \begin{pmatrix} 1 & 10 \end{pmatrix} \qquad \begin{pmatrix} 4 & 6 \end{pmatrix} \qquad \begin{pmatrix} 9 & 2 \end{pmatrix}$$

The separating hyperplane is $\frac{1}{2}p_{1j} + \frac{1}{2}p_{2j} = 5$, computed on $j = 1$ or $j = 3$. This means that $\mathbf{x} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$ is the optimal strategy of Player 1.

Moreover, this indicates that the value of the game is $v = 5$.

For the other player, instead, we need to obtain $\begin{pmatrix} 5 & 5 \end{pmatrix}$ as the result of a convex combination of the vectors in the first and third columns, while the probability of the second and fouth columns is $y_2 = 0$ and $y_4 = 0$, one shows that $\begin{pmatrix} \frac{1}{2} & 0 & \frac{2}{3} & 0 \end{pmatrix}$ is the optimal strategy for Player 2.

3.3.4 General case otimality

Von Neumann proof can be efficiently used to find rational outcome of payoff matrices that can be reduced to matrices where ine player has only two strategies. However, in higher dimensions this procedure becomes more complicateds, since it is not clear when and where the set Q_t meets C. Therefore, we need to use Linear Programmming.

Player one Player 1 must choose a probability distribution $\mathbf{x} = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \in \sum_n$ in order to maximize v with the following constraints:

$$\begin{cases} (x, p_{\cdot,1}) = x_1 p_{11} + \cdots + x_n p_{n1} \geq v \\ \cdots \\ (x, p_{\cdot,j}) = x_1 p_{1j} + \cdots + x_n p_{nj} \geq v \\ \cdots \\ (x, p_{\cdot,m}) = x_1 p_{1m} + \cdots + x_n p_{nm} \geq v \end{cases}$$

It is a linear maximization problem where we need to find the value v and we do not know the vector $\mathbf{x}$. In matrices, we have:

$$\begin{cases} \min_{\mathbf{x},v} v : \\ P^T \mathbf{x} \geq v \mathbf{1}_m \\ \mathbf{x} \geq 0 \qquad \begin{pmatrix} 1 & \mathbf{x} \end{pmatrix} = 1 \end{cases}$$

Player two Player 2 must choose a probability distribution $\mathbf{y} = \begin{pmatrix} y_1 & \cdots & y_m \end{pmatrix} \in \sum_m$ in order to maximize w with the following constraints:

$$\begin{cases} (x, p_{1,\cdot}) = x_1 p_{11} + \cdots + x_m p_{1m} \leq w \\ \cdots \\ (x, p_{i,\cdot}) = x_1 p_{i1} + \cdots + x_m p_{im} \leq w \\ \cdots \\ (x, p_{n,\cdot}) = x_1 p_{n1} + \cdots + x_m p_{nm} \leq w \end{cases}$$

It is a linear maximization problem where we need to find the value w and we do not know the vector $\mathbf{y}$. In matrices, we have:

$$\begin{cases} \min_{\mathbf{y},w} w : \\ P \mathbf{y} \leq w \mathbf{1}_n \\ \mathbf{y} \geq 0 \qquad \begin{pmatrix} 1 & \mathbf{y} \end{pmatrix} = 1 \end{cases}$$

Here, $\mathbf{1}$ is a vector of right dimensions whose componenst are all 1's. Ideally, the maximum value for v is equal to the minimal value for w, so as to yield the value of the game.

CHAPTER 4

Mathematical concepts

4.1 Binary sum

We define a binary operation $\oplus$ on the set of natural numbers $\mathbb{N} = \{0, 1, 2, \ldots, n, \ldots\}$ as follows. For any two natural numbers $n_1, n_2 \in \mathbb{N}$:

1. Convert n_1 and ${}_n2$ into their binary representations, denoted as $[n_1]_2$ and $[n_2]_2$.
2. Perform the binary addition of $[n_1]_2$ and $[n_2]_2$ using the standard addition method, but without carrying over. This means if the addition of two bits results in 2 (i.e., $1 + 1$), it should be represented as 0 in that position with no carry to the next higher bit.
3. The result of the operation $\oplus$ is then represented in binary form, which corresponds to the sum computed in step 2.

Example:
Let's apply the $\oplus$ operation to the numbers $1, 2, 4$, and 1:

1. Convert the numbers to binary:
   - 1 in binary: $[1]_2 = 001$
   - 2 in binary: $[2]_2 = 010$
   - 4 in binary: $[4]_2 = 100$
   - 1 in binary: $[1]_2 = 001$
2. Perform the $\oplus$ operation:

$$\begin{array}{c} [1]_2 = 001 + \\ [2]_2 = 010 + \\ [4]_2 = 100 + \\ [1]_2 = 001 = \\ \hline [6]_2 = 110 \end{array}$$

3. The result of the operation $1 \oplus 2 \oplus 4 \oplus 1$ in decimal form is 6.

4.2 Group

Definition (*Group*)**.** A group is defined as a nonempty set A equipped with a binary operation $\cdot$ such that the following conditions hold:

1. *Closure*: for any elements $a, b \in A$, the result of the operation $a \cdot b$ is also an element of A.
2. *Associativity*: the operation $\cdot$ is associative, meaning that for all $a, b, c \in A$, it holds that $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
3. *Identity element*: there exists a unique element e known as the identity element, such that for every $a \in A$, the following holds: $a \cdot e = e \cdot a = a$.
4. *Inverse element*: for every element $a \in A$, there exists a unique element $b \in A$ (denoted as a^{-1}) such that $a \cdot b = b \cdot a = e$. This element b is called the inverse of a.

Definition (*Abelian group*)**.** A group A is termed an abelian group (or commutative group) if the operation is commutative; that is, for all $a, b \in A$, the equation $\cdot$ $a \cdot b = b \cdot a$ holds true.

Example:
Examples of abelian groups:

1. *The integers* $\mathbb{Z}$: the set of integers, equipped with the usual addition operation $(+)$, forms an abelian group. This group satisfies all the group properties:
   - *Closure*: the sum of any two integers is an integer.
   - *Associativity*: addition is associative, i.e., $(a + b) + c = a + (b + c)$.
   - *Identity*: the identity element is 0 since $a + 0 = a$ for any integer a.
   - *Inverses*: for every integer a, the inverse is $-a$ because $a + (-a) = 0$.
2. *The non-zero real numbers* $\mathbb{R}^*$: the set of all real numbers except 0, equipped with the usual multiplication operation $(\times)$, is an abelian group. It fulfills the following criteria:
   - *Closure*: the product of any two non-zero real numbers is also a non-zero real number.
   - *Associativity*: multiplication is associative, i.e., $(a \times b) \times c = a \times (b \times c)$.
   - *Identity*: the identity element is 1 because $a \times 1 = a$ for any non-zero real number a.
   - *Inverses*: for every non-zero real number a, the inverse is $\frac{1}{a}$ since $a \times \frac{1}{a} = 1$.

Examples of non abelian groups:

1. *The group of $n \times n$ matrices with non-zero determinant*: the set of all $n \times n$ matrices with a non-zero determinant, equipped with the usual matrix multiplication, is a non-abelian group (often denoted as $\mathrm{GL}(n, \mathbb{R})$). This group satisfies the group properties as follows:
   - *Closure*: the product of two invertible matrices is invertible, thus remaining in the group.

- *Associativity*: matrix multiplication is associative, i.e., $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.
- *Identity*: the identity matrix serves as the identity element.
- *Inverses*: each invertible matrix has an inverse that is also an invertible matrix.
- *Non-abelian*: for matrices A and B, it is generally true that $A \cdot B \neq B \cdot A$.

Proposition. Let $(A, \cdot)$ be a group. Then the cancellation law holds:

$$a \cdot b = a \cdot c \implies b = c$$

Proof. To demonstrate the cancellation law, we start with the equation $a \cdot b = a \cdot c$.

By multiplying both sides of this equation by the inverse of a, denoted as a^{-1}, we obtain:

$$a^{-1}a \cdot b = a^{-1}a \cdot c$$

Utilizing the property of inverses, this simplifies to:

$$e \cdot b = e \cdot c$$

Here, e is the identity element of the group. By the definition of the identity, we have $b = c$. □

Proposition. The set of natural numbers with the operation $\oplus$ forms an abelian group.

Proof. We verify the group properties. The identity element is 0, since $n \oplus 0 = n$ for any natural number n. For any natural number n, the inverse with respect to $\oplus$ is n itself. However, since we consider the natural numbers as a set starting from 0, the formal definition of inverses in this context might not strictly apply, but 0 serves as an absorbing element for addition. The operation $\oplus$ is associative, as $(n_1 \oplus n_2) \oplus n_3 = n_1 \oplus (n_2 \oplus n_3)$ holds for all natural numbers n_1, n_2, n_3. The operation $\oplus$ is commutative since $n_1 \oplus n_2 = n_2 \oplus n_1$ for any natural numbers n_1 and n_2.

Therefore, since all group properties are satisfied, $(\mathbb{N}, \oplus)$ is an abelian group. Consequently, the cancellation law holds:

$$n_1 \oplus n_2 = n_1 \oplus n_3 \implies n_2 = n_3$$

□

4.3 Convexity

Definition (*Convex set*)**.** A set $C \subset \mathbb{R}^n$ is called convex if for any points $x, y \in C$ and for any $\lambda \in [0, 1]$ the point $\lambda x + (1 - \lambda)y \in C$.

This means that the line segment connecting any two points in C is entirely contained within C. The properties of a convex set are:

- The intersection of an arbitrary family of convex sets is convex.
- A closed convex set with a nonempty interior coincides with the closure of its internal points.

Definition (*Convex combination*)**.** A convex combination of elements $x_1, \ldots, x_n$ is any vector x of the form:

$$x = \lambda_1 x_1 + \cdots + \lambda_n x_n$$

where $\lambda_1 \geq 0, \ldots, \lambda_n \geq 0$ and $\sum_{i=1}^{n} \lambda_i = 1$.

Proposition. A set C is convex if and only if for every $\lambda_1 \geq 0, \ldots, \lambda_n \geq 0$ such that $\sum_{i=1}^n \lambda_i = 1$, and for every $c_1, \ldots, c_n \in C$, we have

$$\sum_{i=1}^n \lambda_i c_i \in C$$

Definition (*Convex hull*)**.** The convex hull of a set C, denoted by co C, is the smallest convex set containing C. It is defined as:

$$\text{co } C = \bigcap_{A \in \mathcal{C}} A$$

where $\mathcal{C} = \{A | C \subset A \text{ and } A \text{ is convex}\}$.

Proposition. The convex hull of a set C can be expressed as:

$$\text{co } C = \left\{ \sum_{i=1}^n \lambda_i c_i | \lambda_i \geq 0, \sum_{i=1}^n \lambda_i = 1, c_i \in C \quad \forall i, n \in \mathbb{N} \right\}$$

The convex hull of a set C consists of all convex combinations of points in C. When C is a finite set, the convex hull is called a polytope.

Theorem 4.3.1. *Given a closed convex set C and a point x outside C, there exists a unique point $p \in C$ such that for all $c \in C$:*

$$\|p - x\| \leq \|c - x\|$$

The projection p is the point in C closest to x and satisfies the following:

1. $p \in C$.
2. $(x - p, c - p) \leq 0$ for all $c \in C$.

Theorem 4.3.2. *Let C be a convex subset of $\mathbb{R}^l$, and assume $\bar{x} \in$ cl C^c (the closure of the complement of C). Then, there exists a nonzero $x^* \in \mathbb{R}^l$ such that for all $c \in C$:*

$$(x^*, c) \geq (x^*, \bar{x})$$

This result provides a criterion to distinguish points outside of C from those inside.

Proof. Assume that $\bar{x} \notin$ cl C and let p be its projection onto cl C. By the previous theorem, we have:

$$(\bar{x} - p, c - p) \leq 0 \qquad \forall c \in C$$

Now, define $x^* = p - \bar{x} \neq 0$, then the inequality becomes:

$$(-x^*, c - \bar{x} - x^*) = (-x^*, -x^*) + (-x^*, c - \bar{x}) \leq 0$$

This implies that:

$$(x^*, c - \bar{x}) \geq \|x^*\|^2$$

Since $\|x^*\|^2 > 0$, by linearity we obtain:

$$(x^*, c) \geq (x^*, \bar{x}) \qquad \forall c \in C$$

Since x^* appears on both sides of the inequality, we can renormalize and choose $\|x^*\| = 1$.

If $\bar{x} \in \text{cl } C \setminus C$, take a sequence $\{x_n\} \subset C^c$ such that $x_n \to \bar{x}$. From the first part of the proof, there exists a sequence of vectors x_n^*, each of norm 1, such that:

$$(x_n^*, c) \geq (x_n^*, x_n) \qquad \forall c \in C$$

By taking the limit along a subsequence where $x_n^* \to x^*$, we obtain:

$$(x^*, c) \geq (x^*, \bar{x}) \qquad \forall c \in C$$

□

Corollary 4.3.2.1. *For any closed convex set C in Euclidean space, and any point x on the boundary of C, there exists a hyperplane that contains x and leaves all points in C on one side of the hyperplane.*

This hyperplane is called a supporting hyperplane for C at x.

Corollary 4.3.2.2. *Any closed convex set C in Euclidean space can be represented as the intersection of all half-spaces that contain it.*

Theorem 4.3.3. *Let A and C be closed convex subsets of $\mathbb{R}^l$, with int $A \neq \varnothing$ and int $A \cap C = \varnothing$. Then, there exists a nonzero vector x^* and a scalar $b \in \mathbb{R}$ such that for all $a \in A$ and $c \in C$:*

$$(x^*, a) \geq b \geq (x^*, c)$$

This provides a criterion to determine whether a point lies in A or C.

Proof. Since $0 = \bar{x} \in (\text{int}A - C)^c$, by the previous separation theorem, there exists $x^* \neq 0$ such that:

$$(x^*, x) \geq 0 \qquad \forall x \in \text{int}A - C$$

By linearity, for $x = a - c$, we obtain:

$$(x^*, a) \geq (x^*, c) \qquad \forall a \in \text{int } A, \forall c \in C$$

Extending this inequality to the closure of int A, we have:

$$(x^*, a) \geq (x^*, c) \qquad \forall a \in \text{cl int } A = A, \forall c \in C$$

□

The hyperplane $H = \{x : (x^*, x) = b\}$ is the separating hyperplane, with A and C located in different half-spaces defined by H.

4.4 Linear programming

Definition (*Duality first form*)**.** The following two linear programs are said to be in duality:

$$\begin{cases} \min(\mathbf{c}, \mathbf{x}) \\ \mathbf{A}\mathbf{x} \geq \mathbf{b} \\ \mathbf{x} \geq 0 \end{cases} \qquad \begin{cases} \max(\mathbf{b}, \mathbf{y}) \\ \mathbf{A}^T\mathbf{y} \leq \mathbf{c} \\ \mathbf{y} \geq 0 \end{cases}$$

Here the matrix $A \in \mathbb{R}^{\gtrdot \times \ltimes}$ and the vectors $\mathbf{c}, \mathbf{x} \in \mathbb{R}^n$, $\mathbf{b}, \mathbf{y} \in \mathbb{R}^m$.

The minimization problem is called primal problem and the maximization is called dual problem.

Definition (*Duality second form*)**.** The following two linear programs are said to be in duality:

$$\begin{cases} \min(\mathbf{c}, \mathbf{x}) \\ \mathbf{A}\mathbf{x} \geq \mathbf{b} \end{cases} \qquad \begin{cases} \max(\mathbf{b}, \mathbf{y}) \\ \mathbf{A}^T\mathbf{y} \leq \mathbf{c} \\ \mathbf{y} \geq 0 \end{cases}$$

The minimization problem in the second form can be written in an equivalent way in the first form; dualizing this shows that the dual is equivalent to the dueal of the second form, in the sense that the solution is the same.

Given two problems in duality, there are three options:

1. Both can be feasible.
2. Only one can be feasible.
3. They can both be infeasible.

Example:
Consider the following problem:

$$\begin{cases} \min x_1 + x_2 \\ x_1 + 2x_2 \geq 1 \\ x_1, x_2 \geq 0 \end{cases}$$

In this case we have:

$$\mathbf{c} = \begin{bmatrix} 1 & 1 \end{bmatrix} \qquad \mathbf{A} = \begin{bmatrix} 1 & 2 \end{bmatrix} \qquad b = 1$$

The corresponding dual problem is:

$$\begin{cases} \max y \\ y \leq 1 \\ 2y \leq 1 \\ y \geq 0 \end{cases}$$

The solution for the first problem is $\begin{pmatrix} 0 & \frac{1}{2} \end{pmatrix}$, and for the second is $\frac{1}{2}$. Therefore, they are both feasible.

Example:
Consider the following problem:

$$\begin{cases} \min x_1 - x_2 \\ x_1 + x_2 \geq 2 \\ -x_1 - x_2 \geq -1 \\ x_1, x_2 \geq 0 \end{cases}$$

In this case we have:

$$\mathbf{c} = \begin{bmatrix} 1 & -1 \end{bmatrix} \qquad \mathbf{A} = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \qquad b = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

The corresponding dual problem is:

$$\begin{cases} \max 2y_1 - y_2 \\ y_1 - y_2 \leq 1 \\ y_1 - y_2 \leq -1 \\ y_1, y_2 \geq 0 \end{cases}$$

The primal problem is infeasible, while the dual problem has a solution $\begin{pmatrix} 0 & 1 \end{pmatrix}$

4.4.1 Duality theorems

Theorem 4.4.1 (*Weak duality*)**.** *Let v be the value of the primal minimization problem and V the value of the dual maximization problem. Then:*

$$v \geq V$$

Proof. In the first form we have:

$$(\mathbf{c}, \mathbf{x}) \geq (\mathbf{A}^T\mathbf{y}, \mathbf{x}) = (\mathbf{y}, \mathbf{A}\mathbf{x}) \geq (\mathbf{y}, \mathbf{b})$$

In the second form we have:

$$(\mathbf{c}, \mathbf{x}) = (\mathbf{A}^T\mathbf{y}, \mathbf{x}) = (\mathbf{y}, \mathbf{A}\mathbf{x}) \geq (\mathbf{y}, \mathbf{b})$$

□

Theorem 4.4.2 (*Strong duality*)**.** *If the primal and the dual problems are feasible, then both problems have optimal solutions $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ and the optimal values coincide, thai is:*

$$v = (\mathbf{c}, \bar{\mathbf{x}}) = (\mathbf{b}, \bar{\mathbf{y}}) = V$$

If the primal is feasible and the dual is infeasible, then $v = V = -\infty$.
If the primal is infeasible and the dual is feasible, then $v = V = +\infty$.
If both the primal and the dual are infeasible, then $v = +\infty > V = -\infty$.

Corollary 4.4.2.1. *If one problem is feasible and has an optimal solution, then also the dual problem is feasible and has solution. Morovere, there is no duality gap.*

4.4.2 Complementarity

Theorem 4.4.3 (Complementarity condition first form)**.** *Let $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ be primal and dual feasible. Then $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ are simultaneously optimal if and only if:*

$$\begin{cases} \forall i \bar{x}_i > 0 \implies \sum_{k=1}^{m} a_{ik}\bar{y}_k = c_i \\ \forall i \bar{y}_i > 0 \implies \sum_{k=1}^{n} a_{kj}\bar{x}_k = b_i \end{cases}$$

Proof. Recall that $(\mathbf{c}, \mathbf{x}) \geq (\mathbf{A}^T\mathbf{y}, \mathbf{x}) \geq (\mathbf{y}, \mathbf{A}\mathbf{x}) \geq (\mathbf{b}, \mathbf{y})$. So, $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ are optimal if and only if:

$$(\mathbf{c}, \bar{\mathbf{x}}) = (\mathbf{A}^T\bar{\mathbf{y}}, \bar{\mathbf{x}}) = (\bar{\mathbf{y}}, \mathbf{A}\bar{\mathbf{x}}) = (\mathbf{b}, \bar{\mathbf{y}})$$

This is equivalent to:

$$(\mathbf{A}^T\bar{\mathbf{y}} - \mathbf{c}, \bar{\mathbf{x}}) = 0 \qquad (\mathbf{A}\bar{\mathbf{y}} - \mathbf{b}, \bar{\mathbf{y}}) = 0$$

Since $\bar{\mathbf{x}}, \bar{\mathbf{y}} \geq 0$, $\mathbf{A}\bar{\mathbf{x}} \geq \mathbf{b}$, and $\mathbf{A}^T\bar{\mathbf{y}} \leq \mathbf{c}$ the latter equations are equivalent to the complementary conditions states by the theorem. □

Example:
Consider the following linear programming problem:

$$\begin{cases} \min x_1 + x_2 \\ 2x_1 + x_2 \geq 2 \\ x_1 + 2x_2 \leq 2 \\ x_1, x_2 \geq 0 \end{cases}$$

The corresponding dual is:

$$\begin{cases} \max 2y_1 - 2y_2 \\ 2y_1 - y_2 \leq 1 \\ y_1 - 2y_2 \leq 1 \\ y_1, y_2 \geq 0 \end{cases}$$

We have that $v = 1$, and $\bar{\mathbf{x}} = \begin{bmatrix} 1 & 0 \end{bmatrix}$. We have that $V = 1$, and $\bar{\mathbf{x}} = \begin{bmatrix} \frac{1}{2} & 1 \end{bmatrix}$.

We may now check for the complementary conditions:

$$\bar{y}_1 = \frac{1}{2} > 0 \implies 2\bar{x}_1 - \bar{x}_2 = 2 \qquad \bar{x}_1 = 1 > 0 \implies 2y_1 + y_2 = 1$$