Solving games

1 Modelling player behaviour
   • Solutions of zero-sum games
   • Best response
   • Repeated play; equilibria
   • Beliefs; rational solutions
   • Non-strictly-competitive games
   • Cooperation in games
Two-player zero-sum games: dominance

Example: simplify/reduce this two-player zero-sum game:

<table>
<thead>
<tr>
<th></th>
<th>b₁</th>
<th>b₂</th>
<th>b₃</th>
<th>b₄</th>
</tr>
</thead>
<tbody>
<tr>
<td>a₁</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>a₂</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>a₃</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>a₄</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

- Round 1
- Round 2
- Round 3
- Round 4

- Common knowledge: a player won’t play a dominated strategy; other players know this
- Game reduced (iterated dominance) to single strategy for each player: unique solution: \((a₂, b₂)\)
In games the uncertainty for each player includes the *behaviour* of other players; *i.e.*, which strategy they’ll choose.

This uncertainty can be reduced if players have *common knowledge* about the preferences and rationality of other players.

Dominance reduces *strategic uncertainty* about rational behaviour of other players (*e.g.*, rational players will never play dominated strategies).

General principle about rational behaviour: *best response* . . .

---

**Best response**

Re-visit previous zero-sum game:

<table>
<thead>
<tr>
<th></th>
<th>$b_1$</th>
<th>$b_2$</th>
<th>$b_3$</th>
<th>$b_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>$a_2$</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>$a_3$</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>$a_4$</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

- Play $(a_2, b_2)$ is maximal in its column and minimal in its row
- *i.e.*, if column player plays $b_2$, then $a_2$ gives best possible outcome for row player
- Conversely, if row player plays $a_2$, then $b_2$ gives best possible outcome for column player
**Best response: zero-sum games**

**Definition (Best response)**

A player’s strategy $s^*$ is a *best response* to another player’s strategy $s$ if $s^*$ gives a preference maximal outcome against $s$.

$$
\begin{array}{c|cc}
  & b_1 & b_2 \\
\hline
  a_1 & 2 & 0^* \\
  a_2 & 1^* & 3 \\
\end{array}
$$

In a zero-sum game:

- for any strategy of the column player, a best response of the row player is a strategy which maximises the column value ($\,*$)
- for any strategy of the row player, a best response of the column player is a strategy which minimises the row value ($\,*$)

---

Strictly dominated strategies are never best responses to any other player’s strategies

- Column player’s best responses are minimal in their row
- Every strategy has at least one best response; $a_2$ has two
- Row player’s are maximal in their column
- Multiple best responses must have same payoff
Best response: *Maximin*

<table>
<thead>
<tr>
<th></th>
<th>$b_1$</th>
<th>$b_2$</th>
<th>$b_3$</th>
<th>min</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>1*</td>
<td>2</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>$a_2$</td>
<td>7</td>
<td>3*</td>
<td>3*</td>
<td>3</td>
</tr>
<tr>
<td>$a_3$</td>
<td>3</td>
<td>2*</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>max</td>
<td>7</td>
<td>3</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

- Row player’s *Maximin* strategy is best strategy against ‘perfect play’ by opponent.
- Above, row player’s *Maximin* strategy is $a_2$; column player’s *Maximin* strategy (i.e., *miniMax* strategy) is $b_2$.
- *Maximin* is rational sometimes: e.g., if opponent can see your move.

Repeated play

<table>
<thead>
<tr>
<th></th>
<th>$b_1$</th>
<th>$b_2$</th>
<th>$b_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>1*</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>$a_2$</td>
<td>4</td>
<td>3*</td>
<td>4</td>
</tr>
<tr>
<td>$a_3$</td>
<td>7</td>
<td>2*</td>
<td>5</td>
</tr>
</tbody>
</table>

- Suppose initially row player plays $a_3$, hoping for best outcome; similarly column player plays $b_1$; play $(a_3, b_1)$.
- Row player happy (best response).
- Column player unhappy, so switches to best response $b_2$; in response row player plays $a_2$; ... 
- Play ‘stabilises’ at $(a_2, b_2)$. 
**Equilibrium**

In ‘stable’ play \((a_2, b_2)\) each strategy is a best response to the others.

\[
\begin{array}{ccc}
  b_1 & b_2 & b_3 \\
  a_1 & 1^* & 2 & 3^* \\
  a_2 & 4 & 3^* & 4 \\
  a_3 & 7 & 2^* & 5 \\
\end{array}
\]

**Definition (Nash equilibrium)**

A play is in *equilibrium* if each of its strategies is a best response to the others.

**Equilibrium: belief interpretation**

- If row player believes column player will play \(b_2\), then row player cannot improve outcome by switching, and vice versa.
- More generally, if each player believes the other will play according to their equilibrium strategy, then neither can improve their outcome by deviating from their equilibrium strategy.
**Equilibrium: existence and uniqueness**

- Not all games have an equilibrium . . . in pure strategies
  
  \[
  \begin{array}{c|cc}
  & b_1 & b_2 \\
  a_1 & 2 & 0^* \\
  a_2 & 1^* & 3 \\
  \end{array}
  \]

- Some games have multiple equilibria:
  
  \[
  \begin{array}{c|cccc}
  & b_1 & b_2 & b_3 & b_4 \\
  a_1 & 4 & 2^* & 5 & 2^* \\
  a_2 & 2 & 1 & -1 & -2 \\
  a_3 & 3 & 2^* & 4 & 2^* \\
  a_4 & -1 & 0 & 6 & 1 \\
  \end{array}
  \]

**Zero-sum games: finding equilibria**

**Definition (Saddle point)**

An entry in a matrix is called a *saddle point* iff it is minimal in its row and maximal in its column.

\[
\begin{array}{c|ccc}
  & b_1 & b_2 & b_3 \\
  a_1 & 1^* & 3 & 4 \\
  a_2 & 7 & 5^* & 6^* \\
  a_3 & 3^* & 4 & 8 \\
  \end{array}
\]

**Theorem (Minimax)**

*In zero-sum games, saddle points represent equilibria.*
Zero-sum games: solutions

**Theorem**

*If a zero-sum game has an equilibrium, then it corresponds to the players playing Maximin strategies.*

\[
\begin{array}{c|ccc|c}
& b_1 & b_2 & b_3 & \text{min} \\
\hline
a_1 & 1 & 3 & 4 & 1 \\
a_2 & 7 & \star & 6 & 5 \\
a_3 & 3 & 4 & 8 & 3 \\
\hline
\text{max} & 7 & 5 & 8 & \\
\end{array}
\]

Because matrix entries are payoffs for row player, the column player’s Maximin strategy translates to a miniMax strategy.

---

Zero-sum games: equilibrium

**Theorem (Unique value)**

*All equilibria in a zero-sum game yield the same payoff. This payoff is said to be the value of the game.*

- The value of the game above is 5
- Equilibria in zero-sum games are paired Maximin strategies (miniMax for column player)
Zero-sum games: finding equilibria

<table>
<thead>
<tr>
<th></th>
<th>$b_1$</th>
<th>$b_2$</th>
<th>$b_3$</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>$a_2$</td>
<td>7</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>$a_3$</td>
<td>3</td>
<td>4</td>
<td>8</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

- Saddle points are equilibria
- To find equilibria:
  - Use *Maximin* to evaluate each of the players’ strategies (i.e., *miniMax* for column player)
  - If the *Maximin* values agree (e.g., 5 above), then that play is a saddle point of the game

Behaviour and beliefs

- A game matrix includes all possible strategies and outcomes, but says nothing about the players’ preferences or *behaviour*, i.e., which strategies the players should play
- Dominance and best response are principles about preference and rational *behaviour*
- An agent’s behaviour should depend on its *beliefs* about the other players’ behaviour (including likelihoods)
- In order to better explain behaviour we must formulate an agent’s beliefs
Beliefs and behaviour

- Beliefs about the other players’ play can be represented by a mixture of the other players’ pure strategies.
- Player A assigns to player B’s strategy $b_j$ a ‘proportion’ $p_j$ if A’s belief in the ‘degree of likelihood’ that B will play $b_j$ is $p_j$.
- Recall that utilities encode preferences in the presence of uncertainty (risk).

Best response to beliefs: zero-sum games

- Suppose player A believes that player B is twice as likely to play $b_2$ as $b_1$; i.e., B will play $b_1$ with probability $\frac{1}{3}$ and $b_2$ with probability $\frac{2}{3}$.
- Let $\beta \sim \left(\frac{1}{3}, \frac{2}{3}\right)$ represent A’s ‘belief’ about B’s behaviour.
- For belief $\beta$ calculate the Bayes values of A’s strategies:
  \[
  V_B^\beta(a_1) = \frac{1}{3}(2) + \frac{2}{3}(0) = \frac{2}{3},
  \]
  \[
  V_B^\beta(a_2) = \frac{1}{3}(1) + \frac{2}{3}(3) = \frac{7}{3}.
  \]
- Therefore, A’s best response given belief $\beta$ about B is $a_2$. 

Victor Jauregui  Engineering Decisions
Rationalisation of behaviour and belief

- Any strategy by another player which will not be played should receive degree of belief (i.e., probability) 0
- In general, a strategy which isn’t Bayes for some belief $\beta$ can be eliminated; compare with **admissibility**
- In a zero-sum game, rational strategies (strategies which aren’t eliminated) must be on the player’s ‘admissibility frontier’

Non-zero-sum games: best response

- If Alice were to wait, then Bob’s best counter-move would be to climb
- Conversely, if Bob were to climb, then Alice’s best counter-move would be to wait below
Solving games

- What if Alice moves first?

Exercises
- What is Bob’s best response to Alice waiting? To Alice Climbing?
- Are there any equilibrium pairs/points? If so, which are they?

Equilibrium and solutions

Exercise
For the problems above, find all the equilibrium plays.

- In games that aren’t strictly competitive, solution are less clear, because opportunities for co-operation arise
- Other considerations include: group benefit (Pareto optimality), initial tendencies (equilibrium), etc.
**Non-strictly-competitive games**

**Example (The Prisoner’s Dilemma)**

Alice and Bob are suspects in a joint crime. The police doesn’t yet have enough evidence to convict both/either, so it is trying to get either to implicate the other. The police inspector offers each separately a reduced sentence if they ‘defect’ (D) by implicating their accomplice.

If both suspects defect they will get a moderate sentence each (2 years). If only one defects he will get immunity, and the other will get the full sentence (3 years). If neither defects—i.e., they both cooperate (C) with each other—both will be charged for only a minor offence (1 year).

<table>
<thead>
<tr>
<th></th>
<th>d</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>1,1</td>
<td>3,0</td>
</tr>
<tr>
<td>C</td>
<td>0,3</td>
<td>2,2</td>
</tr>
</tbody>
</table>

The payoff is the reduction in the player’s sentence: \(3 - s\), where \(s \in \{0, 1, 2, 3\}\) is the length of the sentence.

**Cooperation in non-zero-sum games**

<table>
<thead>
<tr>
<th></th>
<th>d</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1,1</td>
<td>3,0</td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0,3</td>
<td>2,2</td>
</tr>
</tbody>
</table>

- Individual rationality (dominance) suggests both should defect (Dd); but mutual cooperation (Cc) better outcome for both
- In games which aren’t strictly competitive cooperation may be possible
- What’s best individually (individual rationality) may not best collectively, and vice versa
- Here Cc gives each player a better payoff than the individually rational play Dd
The Prisoner’s Dilemma

Definition (Pareto optimality)

An outcome is *Pareto optimal* iff there is no other outcome which is at least as good or better for all the agents.

Pareto principle

Pareto optimal outcomes are optimal for a group.

Two-player *play diagram*:

- $x$-value (abscissa) is A’s payoff
- $y$-value (ordinate) is B’s payoff

Pareto optimal outcomes represented by points on solid line

The equilibrium is Dd (circled)

The Pareto optimal outcomes are: Cc, Cd, Dc

Play Cc, which is Pareto optimal, is better than Dd for both players

Conclusion

In two-player non-strictly-competitive games, what’s best for the individual may not be best for the group; *i.e.*, cooperation desirable.
Can regarded single-agent decisions as games against a neutral player called ‘Nature’, or ‘Chance’, who has no preferences.

Games in which some of the players’ preferences are unknown are said to have *incomplete information*—as opposed to *imperfect information*, in which information sets may have multiple nodes.

In extensive form, Nature’s moves take place at chance nodes, and they correspond to chance events.

---

**Summary**

- Best response strategies
- Equilibrium in zero-sum games: saddle-points
- Solutions: rational reduction
- Non-zero-sum games: group preference and Pareto optimality; cooperation
- Single agent decisions are ‘games against nature’