An Introduction to Noncooperative Games

Alberto Bressan

Department of Mathematics, Penn State University
Optimal decision problem

maximize: \( \Phi(x, y) \)

The choice \((x^*, y^*) \in \mathbb{R}^2\) yields the maximum payoff
A game for two players

- Player A wishes to maximize his payoff $\Phi^A(a, b)$
- Player B wishes to maximize his payoff $\Phi^B(a, b)$

Player A chooses the value of $a \in A$
Player B chooses the value of $b \in B$
A cooperative solution

- maximize the sum of payoffs $\Phi^A(a, b) + \Phi^B(a, b)$
- split the total payoff *fairly* among the two players (how ???)
The best reply map

If Player A adopts the strategy $a$, the set of best replies for Player B is

$$R^B(a) = \left\{ b; \quad \Phi^B(a, b) = \max_{s \in B} \Phi^B(a, s) \right\}$$

If Player B adopts the strategy $b$, the set of best replies for Player A is

$$R^A(b) = \left\{ a; \quad \Phi^A(a, b) = \max_{s \in A} \Phi^A(s, b) \right\}$$
Nash equilibrium solutions

A couple of strategies \((a^*, b^*)\) is a **Nash equilibrium** if

\[ a^* \in R^A(b^*) \quad \text{and} \quad b^* \in R^B(a^*) \]

Antoin Augustin Cournot (1838)
John Nash (1950)
Existence of Nash equilibria

**Theorem.** Assume

- **Sets of available strategies for the two players:** $A, B \subset \mathbb{R}^n$ are compact and convex
- **Payoff functions:** $\Phi^A, \Phi^B : A \times B \mapsto \mathbb{R}$ are continuous
- For each $a \in A$, the map $b \mapsto \Phi^B(a, b)$ is concave
- For each $b \in B$, the map $a \mapsto \Phi^A(a, b)$ is concave

Then the game admits at least one Nash equilibrium.

For each $a \in A$, the set of best replies $R^B(a) \subseteq B$ is compact, convex

For each $b \in B$, the set of best replies $R^A(b) \subseteq A$ is compact, convex
Proof. If all functions $a \mapsto \Phi^A(a, b)$ and $b \mapsto \Phi^B(a, b)$ are strictly concave, then the best reply maps $R^A, R^B$ are single valued and continuous.

Hence the map

$$(a, b) \mapsto (R^A(b), R^B(a))$$

is a continuous map from the compact convex set $A \times B$ into itself.

By Brouwer’s fixed theorem, it has a fixed point $(a^*, b^*)$. This is a Nash equilibrium.

In the general case, the best replies $R^A, R^B$ are multifunctions with closed graph and convex values.

By Kakutani’s fixed point theorem there exists $(a^*, b^*) \in (R^A(b^*), R^B(a^*))$
One-dimensional version of Brouwer’s and Kakutani’s theorems

Brouwer 1910
\[ x^* = f(x^*) \]

Kakutani 1941
\[ x^* \in F(x^*) \]

no fixed point

Luitzen Egbertus Jan Brouwer (1910)
Shizuo Kakutani (1941)
Arrigo Cellina (1969)
Stackelberg equilibrium

- Player A (the leader) announces his strategy \( a \in A \) in advance
- Player B (the follower) adopts his best reply: \( b \in R^B(a) \subseteq B \)

What is the best strategy for the leader? \( \max_{a \in A} \Phi^A(a, R^B(a)) \)

A couple of strategies \((a^*, b^*)\) is a **Stackelberg equilibrium** if \( b^* \in R^B(a^*) \) and
\[
\Phi^A(a^*, b^*) \geq \Phi^A(a, b) \quad \text{for all } a \in A, \ b \in R^B(a)
\]
Existence of Stackelberg equilibria

**Theorem.** Assume

- Sets of available strategies for the two players: \( A, B \subset \mathbb{R}^n \) are compact
- Payoff functions: \( \Phi^A, \Phi^B : A \times B \mapsto \mathbb{R} \) are continuous

Then the game admits at least one Stackelberg equilibrium.

**Proof.**

- Consider the graph of the best reply map for player \( B \):
  \[ \Gamma \triangleq \{(a, b) \in A \times B ; \quad b \in R^B(a)\} \]
- \( \Gamma \) is a closed, hence compact subset of \( A \times B \).
- The optimization problem
  \[ \max_{(a,b) \in \Gamma} \Phi^A(a, b) \]
  has a solution \( (a^*, b^*) \). This is a Stackelberg equilibrium.
Zero-sum games: \( \Phi^A(a, b) + \Phi^B(a, b) \equiv 0 \)

- If \( \Phi^A(a, b) = -\Phi^B(a, b) \equiv \Phi(a, b) \), the game is called **zero-sum**.
- Player A seeks to maximize \( \Phi(a, b) \), while player B seeks to minimize \( \Phi(a, b) \).
- A Nash equilibrium exists if
  \[
  a \mapsto \Phi(a, b) \quad \text{is concave for every} \quad b \in B
  \]
  \[
  b \mapsto \Phi(a, b) \quad \text{is convex for every} \quad a \in A
  \]
  This corresponds to a **saddle point** of the function \( \Phi \)
Example: the rock-paper-scissors game

Sets of strategies for the two players: \( A = B = \{1, 2, 3\} \), with

1 - rock, 2 - paper, 3 - scissors

\[
\begin{array}{ccc}
R & P & S \\
R & 0 & -1 & 1 \\
P & 1 & 0 & -1 \\
S & -1 & 1 & 0 \\
\end{array}
\]

\[
\begin{array}{ccc}
\Phi_{11} & \Phi_{12} & \Phi_{13} \\
\Phi_{21} & \Phi_{22} & \Phi_{23} \\
\Phi_{31} & \Phi_{32} & \Phi_{33} \\
\end{array}
\]
Randomized strategies

Sets of random strategies

\[ \mathcal{A} = \mathcal{B} = \{ (\theta_1, \theta_2, \theta_3); \theta_i \geq 0, \theta_1 + \theta_2 + \theta_3 = 1 \} \]

If the players choose

\[ \theta^A = (\theta^A_1, \theta^A_2, \theta^A_3), \quad \theta^B = (\theta^B_1, \theta^B_2, \theta^B_3) \]

the payoff is

\[ \Phi(\theta^A, \theta^B) = \sum_{i,j=1}^{3} \Phi_{ij} \theta^A_i \theta^B_j \]

An equilibrium solution exists because

- The sets \( \mathcal{A}, \mathcal{B} \) are compact, convex
- \( \Phi(\cdot, \theta^B) \) is linear (hence concave) as a function of \( \theta^A \)
- \( \Phi(\theta^A, \cdot) \) is linear (hence convex) as a function of \( \theta^B \)
Game theoretical models in Economics and Finance

- Sellers (choosing prices charged) vs. buyers (choosing quantities bought)

- Companies competing for market share (choosing production level, prices, amount spent on research & development or advertising)

- Auctions, bidding games

- Economic growth. Leading player: central bank (choosing prime rate) followers: private companies (choosing investment levels)

- Debt management. Lenders (choosing interest rate) vs. borrower (choosing repayment strategy)

- Harvesting of natural resources. Competing companies choosing extraction rates, prices
Differential games in finite time horizon

\[ x(t) \in \mathbb{R}^n = \text{state of the system} \]

Dynamics:
\[
\dot{x}(t) = f(x(t), u_1(t), u_2(t)), \quad x(t_0) = x_0
\]
\[ u_1(\cdot), u_2(\cdot) = \text{controls implemented by the two players} \]

Goal of \( i \)-th player:

Maximize:
\[
J_i = \psi_i(x(T)) - \int_{t_0}^{T} L_i(x(t), u_1(t), u_2(t)) \, dt
\]
\[ = \text{[terminal payoff]} - \text{[running cost]} \]
Differential games in infinite time horizon

Dynamics: \[ \dot{x} = f(x, u_1, u_2), \quad x(0) = x_0 \]

Goal of \(i\)-th player:

minimize: \[ J_i = \int_{0}^{+\infty} e^{-\gamma t} L_i(x(t), u_1(t), u_2(t)) \, dt \]

(running cost, exponentially discounted in time)
Pursuit games

Dynamics: \[ \dot{x} = f(x, u_1, u_2), \quad x(0) = x_0 \]

Target set: \( S \subset \mathbb{R}^n \)

- Player 1 is the pursuer, Player 2 is the quarry
- \( S \) is a set of positions where capture occurs

Time of capture: \( T = \min\{t \geq 0; \ x(t) \in S\} \)

- Player 1 seeks to minimize \( T \)
- Player 2 seeks to maximize \( T \)
Lion and Man in a circular arena

State: \( x = (l_1, l_2, m_1, m_2) = (L, M) \) = positions of lion and man

Dynamics: \( \dot{L} = u_1(t), \quad \dot{M} = \sigma u_2(t) \quad u_1, u_2 \in B(0, 1) \)

speed of lion = 1, speed of man = \( \sigma \leq 1 \)

Target set: \( T = \{(t, x); \quad l_1 = m_1, \quad l_2 = m_2\} \) (lion captures man)

Payoff: \( \psi(t, x) = t = \) [survival time of man] = [waiting time for lion’s meal]
Homicidal chauffeur

State: \( x = (m_1, m_2, c_1, c_2, \alpha) = \text{positions of man and car} \)

- man can move in any direction with speed \( \leq \sigma \)
- car can move forward with speed \( 0 \leq s(t) \leq 1 \), orientation angle satisfies \( |\dot{\alpha}| \leq \kappa \cdot s(t) \) \( \Rightarrow \) there is a minimum radius of curvature

Target set: \( T = \{ \text{man lies within the region occupied by the car} \} \)

Payoff: \( \psi(t, x) = t = [\text{survival time of man}] = +\infty \) if man is never hit
Lady in the lake

State: \( x = (r, \alpha; \beta) = \text{positions of swimmer and runner} \)

- swimmer can move in any direction with speed \( \leq \sigma \)
- runner can move along the shore with speed \( \leq 1 \)

Target set: \( \mathcal{T} = \{ r = 1 \} \) (swimmer reaches the shore)

Payoff: \( \psi(t, x) = |\alpha - \beta| = \text{distance between swimmer and runner when shore is reached} \)
Example 1: an advertising game

- Two companies, competing for market share

  state variable: \( x(t) \in [0, 1] \) = market share of company 1, at time \( t \)

  \[
  \dot{x} = (1 - x) u_1 - x u_2
  \]

  controls: \( u_1, u_2 = \) advertising rates

  payoffs: \( J_i = N x_i(T) p_i - \int_0^T c_i u_i(t) \, dt \) \quad i = 1, 2

\( N = \) expected number of items purchased by consumers

\( p_i = \) profit made by player \( i \) on each sale

\( c_i = \) advertising cost

\( x_i = \) market share of player \( i \) \quad (x_1 = x, \quad x_2 = 1 - x) \)
Example 2: harvesting of marine resources

\[ x(t) = \text{amount of fish in a lake, at time } t \]
\[ \dot{x} = \alpha x (M - x) - xu_1 - xu_2 \]

controls: \( u_1, u_2 = \text{harvesting efforts by the two players} \)

payoffs:
\[ J_i = \int_0^{+\infty} e^{-\gamma t} (p xu_i - c_i u_i) \, dt \]

\( p = \text{selling price of fish} \)
\( c_i = \text{harvesting cost} \)
Example 3: a producer vs. consumer game

State variables: \( \begin{align*}
    p &= \text{price} \\
    q &= \text{size of the inventory}
\end{align*} \)

Controls: \( \begin{align*}
    a(t) &= \text{production rate} \\
    b(t) &= \text{consumption rate}
\end{align*} \)

The system evolves in time according to \( \begin{align*}
    \dot{p} &= p \ln\left(\frac{q_0}{q}\right) \\
    \dot{q} &= a - b
\end{align*} \)

Here \( q_0 \) is an “appropriate” inventory level

Payoffs:
\( \begin{align*}
    J_{\text{producer}} &= \int_0^{+\infty} e^{-\gamma t} \left[ p(t) \cdot b(t) - c\left(a(t)\right) \right] \, dt \\
    J_{\text{consumer}} &= \int_0^{+\infty} e^{-\gamma t} \left[ \phi(b(t)) - p(t)b(t) \right] \, dt
\end{align*} \)

\( c(a) = \text{production cost}, \quad \phi(b) = \text{utility to the consumer} \)
Solution concepts

- No outcome can be optimal simultaneously for all players

Different outcomes may arise, depending on

- information available to the players
- their ability and willingness to cooperate
Feedback Nash equilibria (in infinite time horizon)

Seek: feedback strategies: $u_1^*(x)$, $u_2^*(x)$ with the following properties

- Given the strategy $u_2 = u_2^*(x)$ adopted by the second player, for every initial data $x(0) = y$, the assignment $u_1 = u_1^*(x)$ provides a solution to the optimal control problem for the first player:

$$
\min_{u_1(\cdot)} \int_0^\infty e^{-\gamma t} L_1(x, u_1, u_2^*(x)) \, dt
$$

subject to

$$
\dot{x} = f(x, u_1, u_2^*(x)), \quad x(0) = y
$$

- Similarly, given the strategy $u_1 = u_1^*(x)$ adopted by the first player, the feedback control $u_2 = u_2^*(x)$ provides a solution to the optimal control problem for the second player.
Solving a differential game by PDE methods

Given the feedback $u_2 = u_2^*(x)$ adopted by the second player, Player 1 must solve the optimal control problem

$$\max_{u_1(\cdot)} \int_0^\infty e^{-\gamma t} L_1(x, u_1, u_2^*(x)) \, dt$$

subject to

$$\dot{x} = f(x, u_1, u_2^*(x)), \quad x(0) = y$$

The value function $V_1$ for Player 1 satisfies the PDE

$$\gamma V_1 = H_1(x, \nabla V_1, u_2^*(x)) \doteq \min_\omega \left\{ L_1(x, \omega, u_2^*(x)) + \nabla V_1(x) \cdot f(x, \omega, u_2^*(x)) \right\}$$

Optimal feedback: $u_1^*(x) = \arg\min_\omega \left\{ L_1(x, \omega, u_2^*(x)) + \nabla V_1(x) \cdot f(x, \omega, u_2^*(x)) \right\}$

Similarly, given $u_1 = u_1^*(x)$, the value function $V_2$ for Player 2 satisfies

$$\gamma V_2(x) = H_2(x, u_1^*(x), \nabla V_1) \doteq \min_\omega \left\{ L_2(x, u_1^*(x), \omega) + \nabla V_2(x) \cdot f(x, u_1^*(x), \omega) \right\}$$
A class of games with separated dynamics and costs

**Dynamics:** \( \dot{x} = f_1(x, u_1) + f_2(x, u_2) \)

Cost to player 1: \( J_1 = \int_0^{+\infty} e^{-\gamma t} [L_{11}(x, u_1) + L_{12}(x, u_2)] \, dt \)

Cost to player 2: \( J_2 = \int_0^{+\infty} e^{-\gamma t} [L_{21}(x, u_1) + L_{22}(x, u_2)] \, dt \)

**Value functions:** \( V_1, V_2 \)

Optimal feedback controls:

\[ u_1 = u_1^*(x, \nabla V_1(x)) = \arg \min_{\omega \in U_1} \left\{ L_{11}(x, \omega) + \nabla V_1 \cdot f_1(x, \omega) \right\} \]

\[ u_2^* = u_2^*(x, \nabla V_2(x)) = \arg \min_{\omega \in U_2} \left\{ L_{22}(x, \omega) + \nabla V_2 \cdot f_2(x, \omega) \right\} \]

\[
\begin{cases}
\gamma V_1(x) = L_{11}(x, u_1^*) + L_{12}(x, u_2^*) + \nabla V_1(x) \cdot [f_1(x, u_1^*) + f_2(x, u_2^*)] \\
\gamma V_2(x) = L_{21}(x, u_1^*) + L_{22}(x, u_2^*) + \nabla V_2(x) \cdot [f_1(x, u_1^*) + f_2(x, u_2^*)]
\end{cases}
\]

(\text{HJ})
Games with terminal costs

Dynamics: \[ \dot{x} = f_1(x, u_1) + f_2(x, u_2) \]

Cost to player 1: \[ J_1 = \int_0^T [L_{11}(x, u_1) + L_{12}(x, u_2)] \, dt + \psi_1(x(T)) \]

Cost to player 2: \[ J_2 = \int_0^T [L_{21}(x, u_1) + L_{22}(x, u_2)] \, dt + \psi_2(x(T)) \]

Value functions: \[ V_1, V_2 \] defined for \((t, x) \in [0, T] \times \mathbb{R}^n\)

Optimal feedback controls:

\[ u_1 = u_1^*(x, \nabla V_1(t, x)) = \arg\min_{\omega \in U_1} \left\{ L_{11}(x, \omega) + \nabla V_1 \cdot f_1(x, \omega) \right\} \]

\[ u_2^* = u_2^*(x, \nabla V_2(t, x)) = \arg\min_{\omega \in U_2} \left\{ L_{22}(x, \omega) + \nabla V_2 \cdot f_2(x, \omega) \right\} \]
A system of H-J equations for the value functions

\[
\begin{align*}
\partial_t V_1(t, x) + L_{11}(x, u_1^*) + L_{12}(x, u_2^*) + \nabla V_1(t, x) \cdot [f_1(x, u_1^*) + f_2(x, u_2^*)] &= 0 \\
\partial_t V_2(t, x) + L_{21}(x, u_1^*) + L_{22}(x, u_2^*) + \nabla V_2(t, x) \cdot [f_1(x, u_1^*) + f_2(x, u_2^*)] &= 0
\end{align*}
\]  

\[(HJ)\]

\[
\begin{align*}
V_1(T, x) &= \psi_1(x) \\
V_2(T, x) &= \psi_2(x)
\end{align*}
\]

(terminal conditions)

Optimal feedback controls:

\[
\begin{align*}
u_1 &= u_1^*(x, \nabla V_1(t, x)) = \arg\min_{\omega} \left\{ L_{11}(x, \omega) + \nabla V_1 \cdot f_1(x, \omega) \right\} \\
u_2^* &= u_2^*(x, \nabla V_2(t, x)) = \arg\min_{\omega} \left\{ L_{22}(x, \omega) + \nabla V_2 \cdot f_2(x, \omega) \right\}
\end{align*}
\]
An example

Dynamics: \[ \dot{x} = f(x) + g_1(x)u_1 + g_2(x)u_2 \]

Player \( i \) seeks to minimize: \[ J_i = \int_0^\infty e^{-\gamma t} \left( \phi_i(x(t)) + \frac{u_i^2(t)}{2} \right) \, dt \]

Given the strategy \( u_2^*(x) \) of Player 2, the optimal control problem for Player 1 is:

\[
\begin{align*}
\text{minimize } J_1 \quad \text{subject to: } \quad \dot{x} &= f(x) + g_1(x)u_1 + g_2(x)u_2^*(x)
\end{align*}
\]

PDE for the Value Function \( V_1(\bar{x}) = \text{minimum cost starting at } \bar{x} \)

\[
\gamma V_1 = \min_{\omega} \left\{ \left( \phi_1(x) + \frac{\omega^2}{2} \right) + \nabla V_1 \cdot \left( f(x) + g_1(x)\omega + g_2(x)u_2^*(x) \right) \right\}
\]

\[
= \phi_1(x) + \nabla V_1 \cdot \left( f(x) + \nabla V_1 \cdot g_2(x)u_2^*(x) \right) - \frac{1}{2} \left( \nabla V_1 \cdot g_1(x) \right)^2
\]

Optimal feedback control for Player 1

\[ u_1^*(x) = -\nabla V_1(x) \cdot g_1(x) \]
A system of PDEs for the value functions

The value functions $V_1, V_2$ for the two players satisfy the system of H-J equations

$$
\begin{align*}
\gamma V_1 &= (f \cdot \nabla V_1) - \frac{1}{2}(g_1 \cdot \nabla V_1)^2 - (g_2 \cdot \nabla V_1)(g_2 \cdot \nabla V_2) + \phi_1 \\
\gamma V_2 &= (f \cdot \nabla V_2) - \frac{1}{2}(g_2 \cdot \nabla V_2)^2 - (g_1 \cdot \nabla V_1)(g_1 \cdot \nabla V_2) + \phi_2
\end{align*}
$$

Optimal feedback controls: $u_i^*(x) = -\nabla V_i(x) \cdot g_i(x)$, $i = 1, 2$

highly nonlinear, implicit!

The current mathematical theory mainly covers:

- Zero-sum games
- Linear-quadratic games
Linear - Quadratic games

Assume that the dynamics is linear and the cost functions are quadratic:

\[ \dot{x} = (Ax + b_0) + b_1 u_1 + b_2 u_2, \quad x(0) = y \]

\[ J_i = \int_0^{+\infty} e^{-\gamma t} \left( a_i \cdot x + x^T P_i x + c_i u_i + \frac{u_i^2}{2} \right) dt \]

Then the system of PDEs has a special solution of the form

**quadratic polynomial:**

\[ V_i(x) = \alpha_i + \beta_i \cdot x + x^T \Gamma_i x \quad i = 1, 2 \]

**optimal controls:**

\[ u_i^*(x) = \arg\min_{\omega \in \mathbb{R}} \left\{ c_i \omega + \frac{\omega^2}{2} + (\beta_i + 2x^T \Gamma_i) b_i \omega \right\} \]

\[ = -c_i - (\beta_i + 2x^T \Gamma_i) \cdot b_i \]

To find this solution, it suffices to determine the coefficients \( \alpha_i, \beta_i, \Gamma_i \) by solving a system of algebraic equations.
Validity of linear-quadratic approximations?

Assume the dynamics is almost linear

\[ \dot{x} = f_0(x) + g_1(x)u_1 + g_2(x)u_2 \approx (Ax + b_0) + b_1 u_1 + b_2 u_2, \quad x(0) = y, \]

and the cost functions are almost quadratic

\[ J_i = \int_0^{+\infty} e^{-\gamma t} \left( \phi_i(x) + \frac{u_i^2}{2} \right) dt \approx \int_0^{+\infty} e^{-\gamma t} \left( a_i \cdot x + x^T P_i x + \frac{u_i^2}{2} \right) dt \]

Is it true that the nonlinear game has a feedback solution close to the one for linear-quadratic game?
Zero sum games

**Dynamics:** \[ \dot{x} = f_1(x, u_1) + f_2(x, u_2) \]

**Cost:** \[ J_1 = - J_2 = \int_0^T [L_1(x, u_1) + L_2(x, u_2)] \, dt + \psi(x(T)) \]

**Value functions:** \[ V_1 = - V_2 = V \] defined for \((t, x) \in [0, T] \times \mathbb{R}^n\)

Optimal feedback controls:

\[ u_1 = u_1^*(x, \nabla V(t, x)) = \arg\min_{\omega \in U_1} \left\{ L_1(x, \omega) + \nabla V \cdot f_1(x, \omega) \right\} \]

\[ u_2^* = u_2^*(x, \nabla V(t, x)) = \arg\max_{\omega \in U_2} \left\{ L_2(x, \omega) + \nabla V \cdot f_2(x, \omega) \right\} \]

The value function satisfies the backward PDE

\[ \partial_t V(t, x) + L_1(x, u_1^*) + L_2(x, u_2^*) + \nabla V(t, x) \cdot [f_1(x, u_1^*) + f_2(x, u_2^*)] = 0 \quad (HJ) \]

\[ V(T, x) = \psi(x) \] (terminal conditions)
Hamiltonian function:

\[ H(x, p) = \min_{u_1 \in U_1} \max_{u_2 \in U_2} \left\{ L_1(x, u_1) + L_2(x, u_2) + p \cdot [f_1(x, u_1) + f_2(x, u_2)] \right\} \]

The value function \( V = V(t, x) \) is a viscosity solution to the scalar H-J equation

\[ -\left[ \partial_t V + H(x, \nabla V) \right] = 0 \]

\[ V(T, x) = \psi(x) \]
References

On optimal control:


On differential games:

