An Introduction to Noncooperative Games

Alberto Bressan

Department of Mathematics, Penn State University
• introduction to non-cooperative games: solution concepts

• differential games in continuous time

• applications to economic models
Optimal decision problem

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

\[ \Phi = \text{constant} \]

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 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 $b \in B$
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

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

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

$$R^A(b) = \left\{ a ; \Phi^A(a, b) = \max_{s \in A} \Phi^A(s, b) \right\}$$
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)
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 set of best replies $R^B(a) \subset B$ is convex
- For each $b \in B$, the set of best replies $R^A(b) \subset A$ is convex

Then the game admits at least one Nash equilibrium.

Proof. If the best reply maps are single valued, 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 $(a^*, b^*)$.

If $R^A, R^B$ are convex-valued, 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
\[ x^* \in F(x^*) \]

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

$$\Phi^A(a^*, b^*) \geq \Phi^A(a, b) \quad \text{for all } a \in A, \; b \in R^B(a)$$
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)
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:

\[ \begin{align*}
\text{maximize:} \quad J_i &= \psi_i(x(T)) - \int_{t_0}^{T} L_i(x(t), u_1(t), u_2(t)) \, dt \\
&= [\text{terminal payoff}] - [\text{running cost}] 
\end{align*} \]
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:

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

(running payoff, exponentially discounted in time)
Example 1: an advertising game

- Two companies, competing for market share

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

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

  controls: \( u_1, u_2 = \text{advertising rates} \)

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

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

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

\( c_i = \text{advertising cost} \)

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

\[ x(t) = \text{amount of fish in a lake, at time } t \]

dynamics: \[ \dot{x} = \alpha x (M - x) - xu_1 - xu_2 \]

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

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

\[ p = \text{selling price of fish} \]

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

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

\[
\max_{u_1(\cdot)} \int_0^\infty e^{-\gamma t} \psi_1(x, u_1, u_2^*(x)) \, dt
\]

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 an optimal control problem by PDE methods

\[ V(y) = \inf_{u(\cdot)} \int_{0}^{+\infty} e^{-\gamma t} L(x(t), u(t)) \, dt \]

subject to:

\[ \dot{x}(t) = f(x(t), u(t)) \quad x(0) = y \quad u(t) \in U \]

\[ V(y) = \text{minimum cost, if the system is initially at } y \]
A PDE for the value function

\[ y + \varepsilon f(y, \omega) \]

If we use the constant control \( u(t) = \omega \) for \( t \in [0, \varepsilon] \)
then we play optimally for \( t \in [\varepsilon, \infty[ \), the total cost is

\[
J^{\varepsilon, \omega} = \left( \int_{0}^{\varepsilon} + \int_{\varepsilon}^{\infty} \right) e^{-\gamma t} L(x(t), u(t)) \, dt
\]

\[
= \varepsilon L(y, \omega) + e^{-\gamma \varepsilon} V\left(y + \varepsilon f(y, \omega)\right) + o(\varepsilon)
\]

\[
= \varepsilon L(y, \omega) + (1 - \gamma \varepsilon) V(y) + \nabla V(y) \cdot \varepsilon f(y, \omega) + o(\varepsilon)
\]

\[
\geq V(y)
\]

Since this is true for every \( \omega \in U \),

\[
V(y) \leq V(y) - \gamma \varepsilon V(y) + \varepsilon \cdot \min_{\omega \in U} \left\{ L(y, \omega) + \nabla V(y) \cdot f(y, \omega) \right\} + o(\varepsilon)
\]
Choosing $\omega = u^*(y)$ the optimal control, we should have equality. Hence

$$V(y) = V(y) - \gamma \varepsilon V(y) + \varepsilon \cdot \min_{\omega \in U} \left\{ L(y, \omega) + \nabla V(y) \cdot f(y, \omega) \right\} + o(\varepsilon)$$

Letting $\varepsilon \to 0$ we obtain

$$\gamma V(y) = \min_{\omega \in U} \left\{ L(y, \omega) + \nabla V(y) \cdot f(y, \omega) \right\} = H(y, \nabla V(y))$$

If $V(\cdot)$ is known, the optimal feedback control can be recovered by

$$u^*(y) = \arg\min_{u \in U} \left\{ L(y, u) + \nabla V(y) \cdot f(y, u) \right\}$$
An example

minimize: \( \int_0^\infty e^{-\gamma t} \left( \phi(x(t)) + \frac{u^2(t)}{2} \right) dt \)

subject to: \( \dot{x} = f(x) + g(x) u, \quad x(0) = y, \quad u(t) \in \mathbb{R} \)

The value function \( V(y) = \text{minimum cost, starting at} \ y \) satisfies the PDE

\[
\gamma V(x) = \min_{\omega \in \mathbb{R}} \left\{ \phi(x) + \frac{\omega^2}{2} + \nabla V(x) \cdot \left( f(x) + g(x) \omega \right) \right\} \\
= \phi(x) + \nabla V(x) \cdot f(x) - \frac{1}{2} (\nabla V(x) \cdot g(x))^2
\]
Finding the optimal feedback control

\[ \dot{x} = f(x) + g(x) u \]

If \( V(\cdot) \) is known, the optimal control can be recovered by

\[ u^*(x) = \arg\min_{u \in \mathbb{R}} \left\{ \nabla V(x) \cdot g(x) u + \frac{u^2}{2} \right\} = -\nabla V(x) \cdot g(x) \]
Solving a differential game by PDE methods

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:

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

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

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

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

highly nonlinear, implicit!
Differential games in finite time horizon

Player $i$ seeks to maximize: $J_i \equiv \psi_i(x(T)) - \int_{t_0}^{T} \left( \phi_i(x(t)) + \frac{u_i^2(t)}{2} \right) dt$

subject to $\dot{x} = f(x) + g_1(x)u_1 + g_2(x)u_2, \quad x(t_0) = x_0.$

The value functions satisfy the system of H-J equations

$$\begin{aligned}
V_{1,t} &= \phi_1 - (f \cdot \nabla V_1) - \frac{1}{2}(g_1 \cdot \nabla V_1)^2 - (g_2 \cdot \nabla V_1)(g_2 \cdot \nabla V_2) \\
V_{2,t} &= \phi_2 - (f \cdot \nabla V_2) - \frac{1}{2}(g_2 \cdot \nabla V_2)^2 - (g_1 \cdot \nabla V_1)(g_1 \cdot \nabla V_2)
\end{aligned}$$

Terminal conditions: $V_1(T,x) = \psi_1(x), \quad V_1(T,x) = \psi_1(x)$

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

a hard PDE problem to solve!
Linear - Quadratic games

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

\[
\dot{x} = (Ax + b_0) + b_1u_1 + b_2u_2, \quad x(0) = y
\]

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

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

\[
\dot{x} = f_0(x) + g_1(x)u_1 + g_2(x)u_2 \approx (Ax + b_0) + b_1u_1 + b_2u_2, \quad x(0) = y
\]

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?
References

On optimal control:


On differential games:

