[Definition]Definition
[Corollary]Corollary
[Theorem]Theroem

Two-player zero-sum games

Guangyu Hou

2024/09/07

Outline

Static games

Static games

- Action sets A and B, objective function $F : A \times B \to \mathbb{R}$
- *zero-sum* game: the objective of player A is to maximize F and the one of player B is to maximize F (or equivalently minimize F), the two players' rewards sum to zero
- Here the Nash equilibrium is a pair $(a^*, b^*) \in A \times B$ such that

$$\begin{cases} F(a^*, b^*) = \sup_{a \in A} F(a, b^*) \\ -F(a^*, b^*) = \sup_{b \in B} -F(a^*, b). \end{cases}$$

- Equivalently, $\inf_{b \in B} F(a^*, b) = F(a^*, b^*) = \sup_{a \in A} F(a, b^*)$

Notes.

If there exists (a^*, b^*) satisfies $F(a^*, b^*) = \sup_{a \in A} \inf_{b \in B} F(a, b) = \inf_{b \in B} \sup_{a \in A} F(a, b)$ (saddle point of F), we reach the Nash equilibrium and the common value is the *value of the game*.

Stochastic differential games

Stochastic differential games

We suppose that two players each control a common d-dimensional state process X, which evolves according to

$$dX_t = b\left(X_t, \alpha_t, \beta_t\right) dt + \sigma\left(X_t, \alpha_t, \beta_t\right) dW_t$$

where $X_0 = x$ and W is an m-dimensional Brownian motion. Similarly to the stochastic control framework, the objective function takes the form

$$J(\alpha, \beta) = \mathbb{E}\left[\int_0^T f\left(X_t, \alpha_t, \beta_t\right) dt + g\left(X_T\right)\right]$$

Definition

The value of a two-player game

The game has value if

$$\sup_{\alpha \in \mathcal{A}} \inf_{\beta \in \mathcal{B}} J(\alpha, \beta) = \inf_{\beta \in \mathcal{B}} \sup_{\alpha \in \mathcal{A}} J(\alpha, \beta)$$

Admissible Control

The three most common choices are:
1. Open loop: Let $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ be the filtration generated by the Brownian motion. Player A chooses an $\mathcal{A}$-valued $\mathbb{F}$-adapted process $\alpha = (\alpha_t(\omega))$. Similarly for player B.
2. Closed loop (Markovian): Player A chooses a (measurable) function $\alpha : [0, T] \times \mathbb{R}^d \to A$. In this case $\alpha\,(t, X_t(\omega))$ is the control process. The function $\alpha(t, x)$ is called the *feedback function*. Similarly for player B.
3. Closed loop (path dependent): Player A chooses a (measurable) function $\alpha : [0, T] \times C\left([0, T]; \mathbb{R}^d\right) \to A$ with the following adaptedness property: For each $t \in [0, T]$ and each $x, y \in C\left([0, T]; \mathbb{R}^d\right)$ satisfying $x_s = y_s$ for all $s \leq t$, we have $\alpha(t, x) = \alpha(t, y)$. Similarly for player B.

Relationship

Note that a closed loop Markovian control (and similarly for a closed loop path dependent control) always gives rise to an open loop control. Indeed, if $\alpha : [0, T] \times \mathbb{R}^d \to A$ and $\beta : [0, T] \times \mathbb{R}^d \to B$ are closed loop Markovian controls, then the state process is determined by solving the SDE

$$dX_t = b\left(X_t, \alpha\left(t, X_t\right), \beta\left(t, X_t\right)\right) dt + \sigma\left(X_t, \alpha\left(t, X_t\right), \beta\left(t, X_t\right)\right) dW_t, \quad X_0 = x$$

and let us assume for this discussion that the SDE is well-posed. Then $\tilde{\alpha}_t(\omega) := \alpha\left(t, X_t(\omega)\right)$. and $\tilde{\beta}_t(\omega) := \beta\left(t, X_t(\omega)\right)$. both define open loop controls. In games, the choice of admissibility class influences the equilibrium outcome. Previously, the control $\tilde{\alpha}$ as a process depends on the choice of the other player! If Player B switches to a different closed loop control $\beta' : [0, T] \times \mathbb{R}^d \to B$ while player A keeps the same control $\alpha : [0, T] \times \mathbb{R}^d \to A$, then we **must** resolve the state equation

Relationship

$$dX'_t = b\left(X'_t, \alpha\left(t, X'_t\right), \beta'\left(t, X'_t\right)\right) dt + \sigma\left(X'_t, \alpha\left(t, X'_t\right), \beta'\left(t, X'_t\right)\right) dW_t, \quad X'_0 = x$$

The equation gives rise a different state process, then the control process of player A becomes $\tilde{\alpha}'_t(\omega) = \alpha(t, X'_t(\omega))$. However, in the open loop regime, the process has no relation to states, then as one of the players switches controls, the other doesn't react to the change.

Under the setting of Closed Loop Controls

First, suppose that player B chooses $\beta(t, x)$. Then player A solves the following problem:

$$\begin{cases} \sup_\alpha \mathbb{E}\left[\int_0^T f\left(X_t, \alpha_t, \beta\left(t, X_t\right)\right) dt + g\left(X_T\right)\right] \\ dX_t = b\left(X_t, \alpha_t, \beta\left(t, X_t\right)\right) + \sigma\left(X_t, \alpha_t, \beta\left(t, X_t\right)\right) dW_t \end{cases}$$

Given β, let $v^\beta(t, x)$ be the value function of player A. In this case, the HJB equation that $v^\beta(t, x)$ should solve is

$$\partial_t v^\beta(t, x) + \sup_{\alpha \in A} h\left(x, \nabla v^\beta(t, x), \nabla^2 v^\beta(t, x), \alpha, \beta(t, x)\right) = 0$$

with terminal $v^\beta(T, x) = g(x)$ and where

$$h(x, y, z, \alpha, \beta) = b(x, \alpha, \beta) \cdot y + \frac{1}{2} \operatorname{Tr}\left[\sigma\sigma^\top(x, \alpha, \beta) z\right] + f(x, \alpha, \beta)$$

The HJB equation

We can then find the optimal $\alpha(t,x)$ by maximizing h pointwise above. Similarly, if player A chooses $\alpha(t,x)$, denote the value function of player B as $v_\alpha(t,x)$. The HJB equation for $v_\alpha(t,x)$ is

$$\partial_t v_\alpha(t,x) + \inf_{\beta \in B} h\left(x, \nabla v_\alpha(t,x), \nabla^2 v_\alpha(t,x), \alpha(t,x), \beta\right) = 0$$

with $v_\alpha(T,x) = g(x)$. Again, the optimal $\beta(t,x)$ is the pointwise minimizer of h. Now suppose that the pair (α, β) is Nash. In that case, both $v^\beta(t,x)$ and $v_\alpha(t,x)$ satisfy the same PDE

$$\partial_t v(t,x) + h\left(x, \nabla v(t,x), \nabla^2 v(t,x), \alpha(t,x), \beta(t,x)\right) = 0$$

and thus, by the Feynman-Kac representation, we must have $v \equiv v^\beta \equiv v_\alpha$. We must then have

$$\sup_{\alpha \in A} h\left(x, \nabla v(t,x), \nabla^2 v(t,x), \alpha, \beta(t,x)\right) = \inf_{\beta \in B} h\left(x, \nabla v(t,x), \nabla^2 v(t,x), \alpha(t,x), \beta\right),$$

which in turn implies that $(\alpha(t,x), \beta(t,x))$ is a saddle point for the function $(\alpha, \beta) \to h\left(x, \nabla v(t,x), \nabla^2 v(t,x), \alpha, \beta\right)$.

Isaacs's condition

Before stating the verification theorem, we introduce some notation. Define the functions $H^+$and H^-as

$$H^+(x, y, z) = \inf_{\beta \in B} \sup_{a \in A} h(x, y, z, \alpha, \beta)$$
$$H^-(x, y, z) = \sup_{\alpha \in A} \inf_{b \in B} h(x, y, z, \alpha, \beta)$$

Suppose also that there are value functions $v^+$and v^-that solve

$$\partial_t v^\pm(t, x) + H^\pm\left(x, \nabla v^\pm(t, x), \nabla^2 v^\pm(t, x)\right) = 0$$

Definition

Isaacs ' condition. We say Isaacs ' condition holds if $\mathrm{H}^+ \equiv \mathrm{H}^-$.

Verification Theorem

Theorem

Verification theorem

Assume that Isaacs' condition holds. Assume also that there is a v which is a smooth solution of

$$\partial_t v(t,x) + H\left(x, \nabla v(t,x), \nabla^2 v(t,x)\right) = 0$$

with terminal condition $v(T,x) = g(x)$. Suppose α and β are measurable functions from $[0,T] \times \mathbb{R}^d$ into A and B, respectively, and that $(\alpha(t,x), \beta(t,x))$ is a saddle point for the function $(\alpha,\beta) \to h\left(x, \nabla v(t,x), \nabla^2 v(t,x), \alpha, \beta\right)$, for each $(t,x) \in [0,T] \times \mathbb{R}^d$. If the state equation

$$dX_t = b\left(X_t, \alpha\left(t, X_t\right), \beta\left(t, X_t\right)\right) dt + \sigma\left(X_t, \alpha\left(t, X_t\right), \beta\left(t, X_t\right)\right) dW_t$$

is well-posed, then (α,β) is a closed loop Nash equilibrium.

Reference

Olivier Guéant, Jean-Michel Lasry, Pierre-Louis Lions. Mean field games and applications

Thank you!