In this paper, we propose an algorithm to solve the Nash equilibrium solution for an n-person noncooperative dynamic game by the extremum seeking control approach with sliding mode. For each player, a switching function is defined as the difference between the player's cost function and a reference signal. The extremum seeking controller for each player is designed so that the system converges to a sliding boundary layer defined in the vicinity of a sliding mode corresponding to the switching function and inside the boundary layer, the cost function tracks the reference signal and converges to the Nash equilibrium solution.
Nash Solution by Extremum Seeking Control Approach

Yaodong Pan2    Tankut Acarman3    Úmit Özgüner4

Abstract

In this paper, we propose an algorithm to solve the Nash equilibrium solution for an n-person noncooperative dynamic game by the extremum seeking control approach with sliding mode. For each player, a switching function is defined as the difference between the player's cost function and a reference signal. The extremum seeking controller for each player is designed so that the system converges to a sliding boundary layer defined in the vicinity of a sliding mode corresponding to the switching function and inside the boundary layer, the cost function tracks the reference signal and converges to the Nash equilibrium solution.

Keyword: Noncooperative Game, Nash Equilibrium Solution, Extremum Seeking, Sliding Model

1 Introduction

For an n-person noncooperative dynamic game, each player defines a cost function and adjusts some of the control parameters to minimize his own cost function [1][2] to find a Nash equilibrium solution. When the cost function as a measurable variable or a combination of some measurable variables can not be exactly formulated, i.e. when the form of the cost function is not given mathematically although it is measurable, extremum seeking control with sliding mode can be used to solve for the Nash solution.

Extremum seeking control approaches have been proposed to find a setpoint and/or track a varying setpoint so that a cost function (which may be unknown) of the system reaches the extremum[3][4][5][6]. The extremum seeking controllers with sliding mode have been proposed[7][8][9][10], and can be explained by the configuration in Figure 1. The switching function s(t) is defined as

\[ s(t) = y(t) - g(t) \]

where g(t) is a reference signal. The setpoint for the minimum (or maximum) can be reached no matter how the plant changes. With this control method, the sliding mode s(t) = 0 happens, the system oscillates in the vicinity of the sliding mode s(t) = 0, i.e. oscillates inside a sliding boundary layer \( |s(t)| \leq \epsilon \), and a minimum or maximum point can be reached in the sliding mode, as shown in Figure 1. Designing an extremum seeking controller for each player ensures that the dynamic game system converges to the Nash equilibrium point.

The arrangement of the paper is as follows. Section 2 describes the problem formulation; Section 3 proposes the extremum seeking algorithm using sliding mode to calculate the Nash equilibrium solution; and Section 4 gives simulation results.

2 Problem Formulation

Consider an n-person noncooperative dynamic game described by a nonlinear system

\[ \frac{d}{dt}x(t) = f(x(t), u_1(t), u_2(t), \ldots, u_n(t)) \quad (1) \]

with cost function for i-th player

\[ J_i(t) = J_i(x(t)), \quad (i \in N) \quad (2) \]

where N is the index set of player defined as

\[ N = \{1, 2, \ldots, n\} \]

\[ x(t) \in R^n, \quad u_i(t) \in R \quad (i \in N), \quad \text{and} \quad J_i(t) \in R \quad (i \in N) \]

are the state variable, the i-th player's control input,
and the i-th player's cost function, respectively. The functions, \( f(x(t), u_1(t), u_2(t), \ldots, u_n(t)) \) and \( J_i(x) \) (i ∈ N) are assumed to be smooth.

**Assumption 1** There exist smooth control laws

\[
u_i(t) = \alpha_i(x(t), \theta_i), \quad (i \in N)
\]

for all players to stabilize the above nonlinear system (1), where \( \theta_i \in \Theta_i \) (i ∈ N) is a control parameter.

With the control input (3), the closed-loop system of the nonlinear system (1) is determined by

\[
\frac{d}{dt} x(t) = f(x(t), \alpha_1(x(t), \theta_1), \alpha_2(x(t), \theta_2), \ldots, \alpha_n(x(t), \theta_n))
\]

**Assumption 2** There exist a smooth function \( x_e : \mathbb{R} \rightarrow \mathbb{R}^n \) such that

\[
f(x(t), \alpha_1(x(t), \theta_1), \alpha_2(x(t), \theta_2), \ldots, \alpha_n(x(t), \theta_n)) = 0
\]

\[
x = x_e(\theta_1, \theta_2, \ldots, \theta_n)
\]

i.e., every n-tuple of the control parameters \( \theta_i \in \Theta_i \) (i ∈ N) determines a unique equilibrium point \( x_e(\theta_1, \theta_2, \ldots, \theta_n) \).

**Assumption 3** The static performance map at the equilibrium point \( x_e(\theta_1, \theta_2, \ldots, \theta_n) \) from a n-tuple of \( \theta_i \in \Theta_i \) (i ∈ N) to \( J_i(t) \) is smooth and has a unique Nash equilibrium solution

\[
J^*(\theta_1, \theta_2, \ldots, \theta_n)
\]

at point \( (\theta_1^*, \theta_2^*, \ldots, \theta_n^*) \) such that

\[
J_i^*(\theta_1^*, \theta_2^*, \ldots, \theta_i^*, \ldots, \theta_n^*) \leq J_i(\theta_1, \theta_2, \ldots, \theta_n), \quad (i \in N)
\]

\[
\forall \theta_i \in \Theta_i, (i \in N)
\]

**Assumption 4** The partial derivative of the static performance map \( J_i^* \) (i ∈ N) satisfies

\[
\frac{\partial}{\partial \theta_i} J_i^* \leq \frac{\partial}{\partial \theta_j} J_i^*, \quad \forall j \neq i \quad (i \in N)
\]

The control objective is to solve the Nash equilibrium solution \( J^*(\theta_1^*, \theta_2^*, \ldots, \theta_n^*) \) by adjusting the parameters \( \theta_i \) by each player (i ∈ N) separately.

### 3 Extremum Seeking with Sliding Mode

To design an extremum seeking controller with sliding mode for the i-th player (i ∈ N), a switching function is defined as

\[
s_i(t) = J_i(t) - g_i(t)
\]

where the reference signal \( g_i(t) \in \mathbb{R} \) is determined by

\[
g_i(t) = \beta_i(t)
\]

where the time-varying parameter \( \beta_i(t) \) will be given later. Then a sliding boundary layer based on the above switch function is defined as

\[
|s_i(t)| \leq \varepsilon_i
\]

where \( \varepsilon_i > 0 \) is a small positive constant.

Let the variable structure control law be

\[
v_i(t) = -k_i \text{sgn}(s_i(t))
\]

and the parameter \( \theta_i \) satisfy

\[
\dot{\theta}_i(t) = v_i(t)
\]

where \( k_i \) is a small enough positive constant.

**Assumption 5** The dynamic system given in (1) is much faster than the one of the parameter \( \theta_i \)'s adjusting process, i.e.

\[
\frac{d}{dt} x(t) \gg \frac{d}{dt} \theta_i.
\]

Therefore in the design of the extremum seeking controller, the cost function \( J_i(t) \) can be replaced by the static performance map

\[
J_i^* = J_i(\theta_1, \theta_2, \ldots, \theta_n).
\]

Assumption 5 is reasonable once \( k_i \) is small enough.

**Assumption 6** The setpoint \( (\theta_1^*, \theta_2^*, \ldots, \theta_n^*) \) corresponding to the Nash equilibrium solution is in the vicinity of the initial n-tuple of \( \theta_i(0) \) (i ∈ N). Thus the partial derivative of the cost function \( J_i(t) \) on \( \theta_i \) is bounded by a positive constant \( \gamma_i \), i.e.,

\[
|\frac{\partial}{\partial \theta_i} J_i(\theta_1, \theta_2, \ldots, \theta_i, \ldots, \theta_n)| \leq \gamma_i
\]

Based on the above assumptions, the derivative of the switching function \( s_i(t) \) is given by

\[
\frac{d}{dt} s_i(t) = \sum_{j=1}^{n} \frac{\partial}{\partial \theta_j} J_i(\theta_1, \theta_2, \ldots, \theta_n) \dot{\theta}_j(t) - \dot{g}_i(t)
\]

\[
= -W_i(\theta_1, \theta_2, \ldots, \theta_n) k_i \text{sgn}(s_i(t)) - \beta_i(t)
\]
where $W(\theta_1, \theta_2, \cdots, \theta_n)$ is determined by

$$W_i(\theta_1, \theta_2, \cdots, \theta_n) = \frac{\partial}{\partial \theta_j} J_i(\theta_1, \theta_2, \cdots, \theta_n) k_j \operatorname{sgn}(s_j(t))$$

According to Assumptions 4 and 6 and if the positive constant $K_i (i \in N)$ is bounded by some constant, it can be shown that $W_i(\theta_1, \theta_2, \cdots, \theta_n)$ is bounded. To simplify the notations, it is assumed to be bounded by $\gamma_i (i \in N)$, i.e.

$$|W_i(\theta_1, \theta_2, \cdots, \theta_n)| \leq \gamma_i \quad (i \in N)$$

Define a Lyapunov function as

$$V_i(t) = \frac{1}{2} s_i^2(t)$$

Then

$$\frac{d}{dt} V_i(t) = s_i(t) \frac{d}{dt} s_i(t) = -W_i(\theta_1, \theta_2, \cdots, \theta_n) k_i |s_i(t)| - s_i(t) \beta_i(t)$$

According to sliding mode control theory, to ensure the convergence to the sliding mode $s_i(t) = 0$ or the sliding boundary layer $|s_i(t)| \leq \epsilon_i$, the above derivative must be negative. Therefore the time-varying parameter $\beta_i(t)$ outside the sliding boundary layer (6) is chosen as

$$\beta_i(t) = \begin{cases} -\tilde{\beta}_i, & s_i(t) < -\epsilon_i \\ \beta_i, & s_i(t) > \epsilon_i \end{cases}$$

where $\tilde{\beta}_i$ and $\beta_i$ are positive constants satisfying

$$\tilde{\beta}_i > \gamma_i k_i + \epsilon_i \quad \beta_i > \gamma_i k_i + \epsilon_i$$

$\epsilon_i$ is a positive constant. Thus the following holds.

$$\frac{d}{dt} V_i(t) \leq -\epsilon_i |s_i(t)|, \quad |s_i(t)| > \epsilon_i \quad (11)$$

which means that the system will enter the sliding boundary layer $|s_i(t)| \leq \epsilon_i$ in a finite time and stay there since then.

Inside the sliding boundary layer $|s_i(t)| \leq \epsilon_i$, the time-varying parameter $\beta_i(t)$ is chosen as

$$\beta_i(t) = \begin{cases} -\tilde{\beta}_i, & s_i(t) < -\epsilon_i \\ 2\epsilon_i \delta(t - t_0), & s_i(t) = \epsilon_i \end{cases}$$

where $\tilde{\beta}_i$ is a positive constant satisfying

$$\tilde{\beta}_i > \gamma_i k_i + \epsilon_i$$

$\delta(t - t_0)$ is the impulse function defined as

$$\delta(t - t_0) = \begin{cases} 0, & t \neq t_0 \\ \infty, & t = t_0 \end{cases}$$

$$\int_{t_0}^{t_1} \delta(t - t_0) dt = 1$$

t_0 is the time instant when $s_i(t)|t=t_0 = \epsilon_i$.

With the parameter $\beta_i(t)$ designed above, inside the sliding boundary $|s_i(t)| \leq \epsilon_i$ except for one of the boundaries, $s_i(t) = \epsilon_i$, the reference signal $g_i(t)$ keeps decreasing with

$$\dot{g}_i(t) = -\tilde{\beta}_i.$$

At the same time, the cost function $J_i(t)$ may increase or decrease but the absolute value of the change rate of $J_i(t)$ is less than the one of $g_i(t)$ as

$$|J_i(t)| = |W_i(\theta_1, \theta_2, \cdots, \theta_n)| k_i \leq \gamma_i k_i < \beta_i - \sigma_i < \beta_i = |g_i(t)|$$

Therefore $s_i(t) = J_i(t) - g_i(t)$ will increase, i.e., the system will move toward the sliding boundary $s_i(t) = \epsilon_i$ and reach the boundary at some time instant $t_0$. Then with the adjusting rule $\dot{g}_i(t) = -2\epsilon_i \delta(t - t_0)$, the system will move to the another boundary $s_i(t) = -\epsilon_i$ as

$$s_i(t_0) = J_i(t_0) - g_i(t_0) = \epsilon_i$$

$$s_i(t_0^+) = J_i(t_0^+) - g_i(t_0^+)$$

$$= (J_i(t_0) - g_i(t_0)) - 2\epsilon_i \int_{t_0}^{t_0^+} \delta(t - t_0) dt$$

$$= -\epsilon_i$$

After then, the system will move from the boundary $s_i(t) = -\epsilon_i$ to another one $s_i(t) = \epsilon_i$ again while the reference signal $g_i(t)$ keeps decreasing. In this way, the system will vibrate inside the sliding boundary layer $|s_i(t)| \leq \epsilon_i$.

It is assumed that the system reaches the boundary $s_i(t) = \epsilon_i$ at time instants $t = t_0, t_1, t_2, \cdots$ as shown in Figure 2. If the sliding boundary layer is chosen to be narrow enough, then it is reasonable to assume that every period $[t_i, t_{i+1})$ $i = 0, 1, 2, \cdots$ is very short so that the function $W_i(\theta_1(t), \theta_2(t), \cdots, \theta_n(t))$, denoted as

$$\alpha_i(t) = W_i(\theta_1(t), \theta_2(t), \cdots, \theta_n(t))|_{t \in [t_j, t_{j+1})}, (j = 0, 1, 2, \cdots)$$

be narrow enough, then it is reasonable to assume that every period $[t_i, t_{i+1})$ $i = 0, 1, 2, \cdots$ is very short so that the function $W_i(\theta_1(t), \theta_2(t), \cdots, \theta_n(t))$, denoted as

$$\alpha_i(t) = W_i(\theta_1(t), \theta_2(t), \cdots, \theta_n(t))|_{t \in [t_j, t_{j+1})}, (j = 0, 1, 2, \cdots)$$

$\alpha_i(t)$ is a positive constant and monotonically decreasing, then

$$\alpha_i(t) \leq \alpha_i(t_0), \quad t = t_0, t_1, t_2, \cdots$$

Theorem 1: Extremum Seeking Control with Sliding Mode

Figure 2: Extremum Seeking Control with Sliding Mode
is a constant in each period and satisfies
\[ |\alpha_{i,j}| \leq \gamma_i \quad (j = 0, 1, 2, \cdots) \]

Now let's show that the cost function \( J_i(t) \) will decrease while the system vibrates inside the sliding boundary layer.

At the time instant \( t_0 \),
\[ s_i(t_0) = J_i(t_0) - g_i(t_0) = \epsilon_i. \]

Then at the time instant \( t_0^+ \),
\[ J_i(t_0^+) = J_i(t_0) \quad g_i(t_0^+) = g_i(t_0) + 2\epsilon_i, \quad s_i(t_0^+) = -\epsilon_i. \]

Let \( t_0 \) denote the time instant when \( s_i(t)|_{t=t_0} = 0 \) \((t_0 < t_0 < t_1)\). For \( t_0 \leq t < t_0 \), the followings hold.
\[ s_i(t) = J_i(t) - g_i(t) < 0, \quad J_i(t) = -\alpha_i k_i \text{sgn}(s_i(t)) = \alpha_i k_i, \quad g_i(t) = -\beta_i < 0 \]
which yield
\[ J_i(t) = J_i(t_0) + \alpha_i k_i(t - t_0) \quad g_i(t) = g_i(t_0) + 2\epsilon_i - \beta_i(t - t_0). \]

According to
\[ s_i(t_0) = J_i(t_0) - g_i(t_0) = 0 \]
the time instant \( t_0 \) can be found as
\[ t_0 = t_0 + \frac{\epsilon_i}{\beta_i + \alpha_i k_i} \]
and \( J_i(t) \) and \( g_i(t) \) at \( t = t_0 \) are given by
\[ J_i(t_0) = J_i(t_0) + \frac{\epsilon_i \alpha_i k_i}{\beta_i + \alpha_i k_i} \quad g_i(t_0) = g_i(t_0) + 2\epsilon_i - \frac{\epsilon_i \beta_i}{\beta_i + \alpha_i k_i} \]

For \( t_0 < t < t_0 \), the followings hold.
\[ s_i(t) = J_i(t) - g_i(t) < 0, \quad J_i(t) = -\alpha_i k_i, \quad g_i(t) = -\beta_i < 0 \]
which yield
\[ J_i(t) = J_i(t_0) - \alpha_i k_i(t - t_0) \quad g_i(t) = g_i(t_0) - \beta_i(t - t_0) \]

According to
\[ s_i(t_1) = J_i(t_1) - g_i(t_1) = \epsilon_i \]
the time instant \( t_1 \) can be found as
\[ t_1 = t_0 + \frac{\epsilon_i}{\beta_i - \alpha_i k_i} \]
and \( J_i(t) \) and \( g_i(t) \) at \( t = t_1 \) are determined by
\[ J_i(t_1) = J_i(t_0) - 2\epsilon_i \frac{\alpha_i k_i^2}{\beta_i - \alpha_i k_i^2} \quad g_i(t_1) = g_i(t_0) - 2\epsilon_i \frac{\alpha_i k_i^2}{\beta_i - \alpha_i k_i^2} \]
i.e., the cost function \( J_i(t) \) and the reference signal \( g_i(t) \) decrease in the period \([t_0, t_1]\) as
\[ J_i(t_1) - J_i(t_0) = -2\epsilon_i \frac{\alpha_i k_i^2}{\beta_i - \alpha_i k_i^2} < 0 \]
\[ g_i(t_1) - g_i(t_0) = -2\epsilon_i \frac{\alpha_i k_i^2}{\beta_i - \alpha_i k_i^2} < 0. \]

In a similar way, it can be shown that before the Nash equilibrium solution is reached, the followings hold.
\[ J_i(t_0) > J_i(t_1) > J_i(t_2) > \cdots = 0 \quad g_i(t_2) > g_i(t_1) > g_i(t_2) > \cdots = 0 \]

When the Nash Solution is reached at a time instant \( t_m \), i.e. when \( \alpha_{i,j} = 0 \) \((j = m, m + 1, m + 2, \cdots)\), \( J_i(t) \) and \( g_i(t) \) will keep to be a constant.

**Theorem 1** Consider the dynamic noncooperative game described by the state equation in (1) with the control input in (3), the sliding mode controller with extremum seeking control approach for the \( i \)-th player \((i \in N)\) designed as
\[ \dot{s}_i(t) = -\alpha_i k_i \text{sgn}(s_i(t)) \quad \dot{g}_i(t) = \beta_i \]
ensures that the cost functions \( J_i(t) \) \((i \in N)\) are minimized to get the Nash equilibrium solution \( \theta^*(\theta_1^*, \theta_2^*, \cdots, \theta_n^*) \).

**Remark 1** The variable structure control rule for the \( i \)-th player \((i \in N)\) may be replaced by
\[ \dot{u}_i(t) = -k_i \text{sgn}(\sin(\omega_i s_i(t) \tau/2\epsilon_i)), \quad (i \in N) \]
where \( \omega_i \geq 1 \) is a positive number, then the amplitude of the vibration on the cost function \( J_i(t) \) becomes smaller for larger \( \omega_i \), which still results in a stable extremum seeking control.
Remark 2 To implement the proposed algorithm for the sampled-data system and also to simplify the controller, the reference signal $g_i(t)$ ($i \in N$) can be modified as

$$
\dot{\tilde{g}}_i(t) = \begin{cases}
-\tilde{\beta}_i, & s_i(t) < \varepsilon_i \\
\tilde{\beta}_i, & s_i(t) \geq \varepsilon_i
\end{cases} \quad (i \in N) \tag{14}
$$

where $\tilde{\beta}_i$ satisfies

$$
l_i\tilde{\beta}_iT_i = 2\varepsilon_i
$$

$T_i$ is the sampling interval and $l_i$ is a positive constant which indicates the number of the sampling intervals when $\dot{\tilde{g}}_i(t) = \tilde{\beta}_i$ ($i \in N$).

4 Examples

Consider a two-person noncooperative dynamic game described by a second-order linear system with unknown parameters.

$$
\dot{x}(t) = 
\begin{bmatrix}
-1 & 0.2 \\ 0.3 & -1
\end{bmatrix} x(t) 
+ \begin{bmatrix}
0.5(u_1(t) - 2 - 0.1u_2(t))^2 + 1.0 \\
0.7(u_2(t) - 1 - 0.2u_1(t))^2 + 0.5
\end{bmatrix}
$$

The cost functions for two players are respectively defined as

$$
J_1(t) = x_1(t) \\
J_2(t) = x_2(t)
$$

The control input is chosen to be the control parameter, i.e.

$$
u_i(t) = \theta_i(t), \quad (i = 1, 2)
$$

Then it is clear that the Nash equilibrium point is given by

$$
\theta_i^* = \left\{ \begin{array}{ll}
1.4286 & (i = 1) \\
2.1429 & (i = 2)
\end{array} \right.
$$

The proposed extremum seeking control algorithm is implemented for the above system with sampling interval as $T = 0.01$ second and other parameters as

$$
\tilde{\beta}_i = 0.005, \quad \tilde{\beta}_i = 5.0, \quad \varepsilon_i = 0.05, \quad k_i = 0.01, \quad (i = 1, 2)
$$

The simulation results are given in Figure 3, which shows that the system enters the sliding boundary layer in a finite time and then oscillates inside the layer while the cost function keeps decreasing with oscillation until the Nash equilibrium point is reached. The amplitude of the vibration can be reduced by choosing a smaller boundary layer $\varepsilon$. Figure 4 are simulation results with

$$
\varepsilon_i = 0.01, \quad \tilde{\beta}_i = 1.0 (i = 1, 2)
$$

Using the control laws given in Remark 1, as shown in Figure 5, results in higher control accuracy with a larger constants $\omega_i$ ($i = 1, 2$).

5 Conclusion

The extremum seeking control approach with sliding mode proposed in [7][8][9][10] was implemented in an n-person noncooperative dynamic game to calculate the Nash equilibrium solution. With the designed controller for each player in the game, the system enters a sliding boundary layer and stays there while the cost function decreases with oscillating behavior, until the Nash equilibrium solution is reached. The simulation result show the effectiveness.

References

Figure 3: Nash Solution by Extremum Seeking Control (ε₁ = ε₂ = 0.05)

Figure 4: Nash Solution by Extremum Seeking Control (ε₁ = ε₂ = 0.01)

Figure 5: Nash Solution by Extremum Seeking Control (ε₁ = ε₂ = 0.05, ω₁ = ω₂ = 10)