Stochastic Differential Games in Finite and Infinite Population Regimes

Part I: *General introduction to stochastic differential games in the finite population regime. Information patterns, robustness, and equilibrium solution concepts*

Tamer Başar

Coordinated Science Laboratory
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
Urbana, IL - 61801

June 19-20, 2018

Outline

- Introduction to stochastic differential games: Players, dynamics, cost functions, information structures, equilibrium concepts
- Derivation of different types of equilibria
- Complexities associated with different information structures; informational non-uniqueness; time consistency (or inconsistency); strategic interactions
- Risk sensitivity, robustness, and elements of adversarial action
- Recap and final thoughts

Stochastic differential games with N players

Players, their controls, and state equation:

Player i with control $u_{i,t}$, $t \geq 0$, $i = 1, \ldots, N$; state equation as Itô SDE

$$dx_t = f(x_t, u_t)\, dt + \sqrt{\epsilon}\, D(x_t, u_t)\, db_t\,; \;\; x_{t=0} = x_0; \;\; u_t := \{u_{1,t}, \ldots, u_{N,t}\}$$

Cost (expected loss) function of Player i over the interval $[t, t_f]$:

$$J_i(\mu; t, x_t) := E\left[q_i(x_{t_f}) + \int_t^{t_f} g_i(s, x_s, u_s)\, ds\right], \;\; i = 1, \ldots, N, \;\; \mu := \{\mu_1, \ldots, \mu_N\}$$

which Player i wants to minimize using

$$u_{i,t} = \mu_i(t, x_{[0,t]}), \;\; i = 1, \ldots, N \;\; \text{– control law, closed-loop perfect state}$$

Other information structures (for each Player i) in combination:

- Open-loop (OL): $u_{i,t} = \mu_i(t, x_0)$
- Open-loop adapted (OLA) - adapted to $b_{[0,t)}$ at time t
- State feedback (SF): $u_{i,t} = \mu_i(t, x_t)$
- Delayed CLPS (D-CLPS): $u_{i,t} = \mu_i(t, x_{[0,t-\tau]})$, for $t > \tau$, otherwise, OL
- Sampled CLPS (s-CLPS): values of x at sampling times $t_1 < t_2, \ldots, t_s < t_f$
- Quantized state (Q-S)
- Noise-perturbed versions (e.g. $u_{i,t} = \mu_i(t, y^i_{[0,t]})$, $dy^i_t = h^i(t; x_t)dt + H^i_t db^i_t$)

Stochastic differential games with N players

Players, their controls, and state equation:

Player i with control $u_{i,t}$, $t \geq 0$, $i = 1, \dots, N$; state equation as Itô SDE

$$dx_t = f(x_t, u_t)\, dt + \sqrt{\epsilon}\, D(x_t, u_t)\, db_t\,; \;\; x_{t=0} = x_0; \;\; u_t := \{u_{1,t}, \dots, u_{N,t}\}$$

Cost (expected loss) function of Player i over the interval $[t, t_f]$:

$$J_i(\mu; t, x_t) := E\left[q_i(x_{t_f}) + \int_t^{t_f} g_i(s, x_s, u_s)\, ds\right], \;\; i = 1, \dots, N, \;\; \mu := \{\mu_1, \dots, \mu_N\}$$

which Player i wants to minimize using

$$u_{i,t} = \mu_i(t, x_{[0,t]})\,, \;\; i = 1, \dots, N \quad - \text{control law, closed-loop perfect state}$$

Other information structures (for each Player i) in combination:

- Open-loop (OL): $u_{i,t} = \mu_i(t, x_0)$
- Open-loop adapted (OLA) - adapted to $b_{[0,t)}$ at time t
- State feedback (SF): $u_{i,t} = \mu_i(t, x_t)$
- Delayed CLPS (D-CLPS): $u_{i,t} = \mu_i(t, x_{[0,t-\tau]})$, for $t > \tau$, otherwise, OL
- Sampled CLPS (s-CLPS): values of x at sampling times $t_1 < t_2, \dots, t_s < t_f$
- Quantized state (Q-S)
- Noise-perturbed versions (e.g. $u_{i,t} = \mu_i(t, y^i_{[0,t]})$, $dy^i_t = h^i(t; x_t)dt + H^i db^i_t$)

Solution Concepts

$$dx_t = f(x_t, u_t)\,dt + \sqrt{\epsilon}\,D(x_t, u_t)\,db_t\,;\quad x_{t=0} = x_0;\quad u_t := \{u_{1,t}, \ldots, u_{N,t}\}$$

Cost (expected loss) function of Player i over the interval $[t, t_f]$:

$$J_i(\mu; t, x_t) := E\Big[q_i(x_{t_f}) + \int_t^{t_f} g_i(s, x_s, u_s)\,ds\Big]\,,\quad i = 1, \ldots, N\,,\quad \mu := \{\mu_1, \ldots, \mu_N\}$$

which Player i wants to minimize with $\mu_i \in M_i$, policy space compatible with IS

If $N = 2$, and $J_1 + J_2 \equiv 0 \quad \Rightarrow$ **zero-sum SDG; saddle-point equilibrium**

$$J_1(\mu_1^*, \mu_2; t = 0, x_0) \le J_1(\mu_1^*, \mu_2^*; t = 0, x_0) \le J_1(\mu_1, \mu_2^*; t = 0, x_0)$$

for all $\mu_i \in M_i,\ i = 1, 2$, where (μ_1^*, μ_2^*) is the SP pair.

Otherwise, several possibilities exist, depending on whether players enter the decision making process symmetrically or not.

Possibilities are **Nash equilibrium**, **Stackelberg equilibrium**, or several variants of them depending on the levels of hierarchy, and tendencies of players to cooperate or not.

Solution Concepts

$$dx_t = f(x_t, u_t)\,dt + \sqrt{\epsilon}\,D(x_t, u_t)\,db_t\,; \quad x_{t=0} = x_0; \quad u_t := \{u_{1,t}, \dots, u_{N,t}\}$$

Cost (expected loss) function of Player i over the interval $[t, t_f]$:

$$J_i(\mu; t, x_t) := E\Big[q_i(x_{t_f}) + \int_t^{t_f} g_i(s, x_s, u_s)\,ds\Big]\,, \quad i = 1, \dots, N\,, \quad \mu := \{\mu_1, \dots, \mu_N\}$$

which Player i wants to minimize with $\mu_i \in M_i$, policy space compatible with IS

If $N = 2$, and $J_1 + J_2 \equiv 0 \quad \Rightarrow \quad$ **zero-sum SDG; saddle-point equilibrium**

$$J_1(\mu_1^*, \mu_2; t = 0, x_0) \le J_1(\mu_1^*, \mu_2^*; t = 0, x_0) \le J_1(\mu_1, \mu_2^*; t = 0, x_0)$$

for all $\mu_i \in M_i,\ i = 1, 2$, where (μ_1^*, μ_2^*) is the SP pair.

Otherwise, several possibilities exist, depending on whether players enter the decision making process symmetrically or not.

Possibilities are **Nash equilibrium**, **Stackelberg equilibrium**, or several variants of them depending on the levels of hierarchy, and tendencies of players to cooperate or not.

Solution Concepts

$$dx_t = f(x_t, u_t)\,dt + \sqrt{\epsilon}\,D(x_t, u_t)\,db_t\,;\quad x_{t=0} = x_0;\quad u_t := \{u_{1,t}, \ldots, u_{N,t}\}$$

Cost (expected loss) function of Player i over the interval $[t, t_f]$:

$$J_i(\mu; t, x_t) := E\left[q_i(x_{t_f}) + \int_t^{t_f} g_i(s, x_s, u_s)\,ds\right],\quad i = 1, \ldots, N,\quad \mu := \{\mu_1, \ldots, \mu_N\}$$

which Player i wants to minimize with $\mu_i \in M_i$, policy space compatible with IS

If $N = 2$, and $J_1 + J_2 \equiv 0 \quad \Rightarrow \quad$ **zero-sum SDG; saddle-point equilibrium**

$$J_1(\mu_1^*, \mu_2; t = 0, x_0) \leq J_1(\mu_1^*, \mu_2^*; t = 0, x_0) \leq J_1(\mu_1, \mu_2^*; t = 0, x_0)$$

for all $\mu_i \in M_i,\ i = 1, 2$, where (μ_1^*, μ_2^*) is the SP pair.

Otherwise, several possibilities exist, depending on whether players enter the decision making process symmetrically or not.

Possibilities are **Nash equilibrium**, **Stackelberg equilibrium**, or several variants of them depending on the levels of hierarchy, and tendencies of players to cooperate or not.

Solution Concepts

$$dx_t = f(x_t, u_t)\,dt + \sqrt{\epsilon}\,D(x_t, u_t)\,db_t\,; \quad x_{t=0} = x_0; \quad u_t := \{u_{1,t}, \ldots, u_{N,t}\}$$

Cost (expected loss) function of Player i over the interval $[t, t_f]$:

$$J_i(\mu; t, x_t) := E\Big[q_i(x_{t_f}) + \int_t^{t_f} g_i(s, x_s, u_s)\,ds\Big]\,, \quad i = 1, \ldots, N\,, \quad \mu := \{\mu_1, \ldots, \mu_N\}$$

which Player i wants to minimize with $\mu_i \in M_i$, policy space compatible with IS

If $N = 2$, and $J_1 + J_2 \equiv 0 \quad \Rightarrow$ **zero-sum SDG; saddle-point equilibrium**

$$J_1(\mu_1^*, \mu_2; t = 0, x_0) \leq J_1(\mu_1^*, \mu_2^*; t = 0, x_0) \leq J_1(\mu_1, \mu_2^*; t = 0, x_0)$$

for all $\mu_i \in M_i,\ i = 1, 2$, where (μ_1^*, μ_2^*) is the SP pair.

Otherwise, several possibilities exist, depending on whether players enter the decision making process symmetrically or not.

Possibilities are **Nash equilibrium**, **Stackelberg equilibrium**, or several variants of them depending on the levels of hierarchy, and tendencies of players to cooperate or not.

Nash equilibrium

$$J_i(\mu_i^*, \mu_{-i}^*; t = 0, x_0) \leq J_i(\mu_i, \mu_{-i}^*; t = 0, x_0)$$

for all $\mu_i \in M_i, \; i = 1, 2, \ldots, N$, where (μ_i^*, μ_{-i}^*) is the NE N-tuple, and μ_{-i} is collection of all μ_j's except μ_i. Note *strategic and informational interaction*!

If compatible with the IS, such as CLPS, then a stronger (refined) NE is (*strongly time consistent*): **[Rec]**

$$J_i(\mu_i^*, \mu_{-i}^*; t, x) \leq J_i(\mu_i, \mu_{-i}^*; t, x)$$

for all $\mu_i \in M_i, \; i = 1, 2, \ldots, N$, and for all $t \in [0, t_f], \; x \in X$.

- Compatibility (and time consistency) holds if **all** players have CLPS or SF or s-CLPS (if t is restricted to sampling times), but not if they have mixed ISs.
- If all players have CLPS and D is full state rank, then every NE uses SF IS (that is only the current value of the state); if D is not full state rank, there exists a plethora of NE that depend on memory on the state–*informational non-uniqueness*; in this case the second set of inequalities do not apply. **[Rec]**
- If some players have CLPS and others have OL or D-CLPS or s-CLPS, the 2nd set of inequalities never apply, but the 1st set is still meaningful. E.g., if some players have CLPS and others have OL, then at NE the former's **[Rec]** policies depend on current state and x_0, while the latter's depend on only x_0.

Nash equilibrium

$$J_i(\mu_i^*, \mu_{-i}^*; t = 0, x_0) \leq J_i(\mu_i, \mu_{-i}^*; t = 0, x_0)$$

for all $\mu_i \in M_i,\ i = 1, 2, \ldots, N$, where (μ_i^*, μ_{-i}^*) is the NE N-tuple, and μ_{-i} is collection of all μ_j's except μ_i. Note *strategic and informational interaction*!
If compatible with the IS, such as CLPS, then a stronger (refined) NE is (*strongly time consistent*): **[Rec]**

$$J_i(\mu_i^*, \mu_{-i}^*; t, x) \leq J_i(\mu_i, \mu_{-i}^*; t, x)$$

for all $\mu_i \in M_i,\ i = 1, 2, \ldots, N$, and for all $t \in [0, t_f],\ x \in X$.

- Compatibility (and time consistency) holds if **all** players have CLPS or SF or s-CLPS (if t is restricted to sampling times), but not if they have mixed ISs.
- If all players have CLPS and D is full state rank, then every NE uses SF IS (that is only the current value of the state); if D is not full state rank, there exists a plethora of NE that depend on memory on the state–*informational non-uniqueness*; in this case the second set of inequalities do not apply. **[Rec]**
- If some players have CLPS and others have OL or D-CLPS or s-CLPS, the 2nd set of inequalities never apply, but the 1st set is still meaningful. E.g., if some players have CLPS and others have OL, then at NE the former's **[Rec]** policies depend on current state and x_0, while the latter's depend on only x_0.

Nash equilibrium

$$J_i(\mu_i^*, \mu_{-i}^*; t = 0, x_0) \leq J_i(\mu_i, \mu_{-i}^*; t = 0, x_0)$$

for all $\mu_i \in M_i,\ i = 1, 2, \ldots, N$, where (μ_i^*, μ_{-i}^*) is the NE N-tuple, and μ_{-i} is collection of all μ_j's except μ_i. Note *strategic and informational interaction*!
If compatible with the IS, such as CLPS, then a stronger (refined) NE is (*strongly time consistent*): **[Rec]**

$$J_i(\mu_i^*, \mu_{-i}^*; t, x) \leq J_i(\mu_i, \mu_{-i}^*; t, x)$$

for all $\mu_i \in M_i,\ i = 1, 2, \ldots, N$, and for all $t \in [0, t_f],\ x \in X$.

- Compatibility (and time consistency) holds if **all** players have CLPS or SF or s-CLPS (if t is restricted to sampling times), but not if they have mixed ISs.
- If all players have CLPS and D is full state rank, then every NE uses SF IS (that is only the current value of the state); if D is not full state rank, there exists a plethora of NE that depend on memory on the state–*informational non-uniqueness*; in this case the second set of inequalities do not apply. **[Rec]**
- If some players have CLPS and others have OL or D-CLPS or s-CLPS, the 2nd set of inequalities never apply, but the 1st set is still meaningful. E.g., if some players have CLPS and others have OL, then at NE the former's **[Rec]** policies depend on current state and x_0, while the latter's depend on only x_0.

Stackelberg Equilibrium

Consider the case of $N = 2$, Player 1 is leader, and Player 2 is follower. First determine the reaction set, $R(\mu_1)$ of Player 2 to policies μ_1 of Player 1:

$$R(\mu_1) := \{\nu_2 \in M_2 : J_2(\mu_1, \nu_2; 0, x_0) \leq J_2(\mu_1, \mu_2; 0, x_0)\,,\ \forall \mu_2 \in M_2\}$$

If R is a singleton for each μ_1, then μ_1^* is a Stackelberg policy for the leader (L) if it minimizes J_1 over M_1 with $\mu_2 = R(\mu_1)$, i.e. [Note, no worse than NE **[Rec]**]

$$J_1(\mu_1^*, R(\mu_1^*); 0, x_0) = \min_{\mu_1 \in M_1} J_1(\mu_1, R(\mu_1); 0, x_0)$$

and $\mu_2^* = R(\mu_1^*)$ is the corresponding optimum policy of Player 2.

If R is **not** a singleton for each μ_1, then one possibility is to introduce a worst-case Stackelberg policy for the leader, μ_1^*, defined by

$$\sup_{\mu_2 \in R(\mu_1^*)} J_1(\mu_1^*, \mu_2; 0, x_0) = \min_{\mu_1 \in M_1} \sup_{\mu_2 \in R(\mu_1)} J_1(\mu_1, \mu_2; 0, x_0)$$

- If L has dynamic information, derivation of SE is very complex due to structural characterization of $R(\mu_1)$ **[Rec]**; general theory does not exist.
- If L has OL information, a set of necessary conditions can be obtained. **[Rec]**
- If both players have CLPS, then **feedback SE** can be introduced. **[Rec]**

Stackelberg Equilibrium

Consider the case of $N = 2$, Player 1 is leader, and Player 2 is follower. First determine the reaction set, $R(\mu_1)$ of Player 2 to policies μ_1 of Player 1:

$$R(\mu_1) := \{\nu_2 \in M_2 : J_2(\mu_1, \nu_2; 0, x_0) \leq J_2(\mu_1, \mu_2; 0, x_0)\,,\ \forall \mu_2 \in M_2\}$$

If R is a singleton for each μ_1, then μ_1^* is a Stackelberg policy for the leader (L) if it minimizes J_1 over M_1 with $\mu_2 = R(\mu_1)$, i.e. [Note, no worse than NE **[Rec]**]

$$J_1(\mu_1^*, R(\mu_1^*); 0, x_0) = \min_{\mu_1 \in M_1} J_1(\mu_1, R(\mu_1); 0, x_0)$$

and $\mu_2^* = R(\mu_1^*)$ is the corresponding optimum policy of Player 2.

If R is **not** a singleton for each μ_1, then one possibility is to introduce a worst-case Stackelberg policy for the leader, μ_1^*, defined by

$$\sup_{\mu_2 \in R(\mu_1^*)} J_1(\mu_1^*, \mu_2; 0, x_0) = \min_{\mu_1 \in M_1} \sup_{\mu_2 \in R(\mu_1)} J_1(\mu_1, \mu_2; 0, x_0)$$

- If L has dynamic information, derivation of SE is very complex due to structural characterization of $R(\mu_1)$ **[Rec]**; general theory does not exist.
- If L has OL information, a set of necessary conditions can be obtained. **[Rec]**
- If both players have CLPS, then **feedback SE** can be introduced. **[Rec]**

Stackelberg Equilibrium

Consider the case of $N = 2$, Player 1 is leader, and Player 2 is follower. First determine the reaction set, $R(\mu_1)$ of Player 2 to policies μ_1 of Player 1:

$$R(\mu_1) := \{\nu_2 \in M_2 : J_2(\mu_1, \nu_2; 0, x_0) \leq J_2(\mu_1, \mu_2; 0, x_0)\,,\ \forall \mu_2 \in M_2\}$$

If R is a singleton for each μ_1, then μ_1^* is a Stackelberg policy for the leader (L) if it minimizes J_1 over M_1 with $\mu_2 = R(\mu_1)$, i.e. [Note, no worse than NE **[Rec]**]

$$J_1(\mu_1^*, R(\mu_1^*); 0, x_0) = \min_{\mu_1 \in M_1} J_1(\mu_1, R(\mu_1); 0, x_0)$$

and $\mu_2^* = R(\mu_1^*)$ is the corresponding optimum policy of Player 2.

If R is **not** a singleton for each μ_1, then one possibility is to introduce a worst-case Stackelberg policy for the leader, μ_1^*, defined by

$$\sup_{\mu_2 \in R(\mu_1^*)} J_1(\mu_1^*, \mu_2; 0, x_0) = \min_{\mu_1 \in M_1} \sup_{\mu_2 \in R(\mu_1)} J_1(\mu_1, \mu_2; 0, x_0)$$

- If L has dynamic information, derivation of SE is very complex due to structural characterization of $R(\mu_1)$ **[Rec]**; general theory does not exist.
- If L has OL information, a set of necessary conditions can be obtained. **[Rec]**
- If both players have CLPS, then **feedback SE** can be introduced. **[Rec]**

Stackelberg Equilibrium

Consider the case of $N = 2$, Player 1 is leader, and Player 2 is follower. First determine the reaction set, $R(\mu_1)$ of Player 2 to policies μ_1 of Player 1:

$$R(\mu_1) := \{\nu_2 \in M_2 : J_2(\mu_1, \nu_2; 0, x_0) \leq J_2(\mu_1, \mu_2; 0, x_0)\,,\ \forall \mu_2 \in M_2\}$$

If R is a singleton for each μ_1, then μ_1^* is a Stackelberg policy for the leader (L) if it minimizes J_1 over M_1 with $\mu_2 = R(\mu_1)$, i.e. [Note, no worse than NE **[Rec]**]

$$J_1(\mu_1^*, R(\mu_1^*); 0, x_0) = \min_{\mu_1 \in M_1} J_1(\mu_1, R(\mu_1); 0, x_0)$$

and $\mu_2^* = R(\mu_1^*)$ is the corresponding optimum policy of Player 2.

If R is **not** a singleton for each μ_1, then one possibility is to introduce a worst-case Stackelberg policy for the leader, μ_1^*, defined by

$$\sup_{\mu_2 \in R(\mu_1^*)} J_1(\mu_1^*, \mu_2; 0, x_0) = \min_{\mu_1 \in M_1} \sup_{\mu_2 \in R(\mu_1)} J_1(\mu_1, \mu_2; 0, x_0)$$

- If L has dynamic information, derivation of SE is very complex due to structural characterization of $R(\mu_1)$ **[Rec]**; general theory does not exist.
- If L has OL information, a set of necessary conditions can be obtained. **[Rec]**
- If both players have CLPS, then **feedback SE** can be introduced. **[Rec]**

NE under CLPS for all players

$$dx_t = f(x_t, u_t)\, dt + \sqrt{\epsilon}\, D\, db_t\,; \quad x_{t=0} = x_0; \quad u_t := \{u_{1,t}, \ldots, u_{N,t}\}$$

Cost (expected loss) function of Player i over the interval $[t, t_f]$:

$$J_i(\mu; t, x_t) := E\Big[q_i(x_{t_f}) + \int_t^{t_f} g_i(s, x_s, u_s)\, ds\Big]\,, \quad i = 1, \ldots, N\,, \quad \mu := \{\mu_1, \ldots, \mu_N\}$$

which Player i wants to minimize with $\mu_i \in M_i$, policy space compatible with IS

If the N-tuple μ^* is in NE, and $V^i(t,x) := J_i(\mu^*; t, x)$ is C_1 in t and C_2 in x, for each i, then V_i satisfies the coupled HJB equations **[Rec]**:

$$-V_t^i(t,x) = \min_{v \in M_i} \big\{ V_x^i(t,x) f(x, v, \mu_{-i}^*(t,x)) + g_i(t, x, v, \mu_{-i}^*(t,x)) \big\} + \frac{\epsilon}{2} \mathrm{Tr}[V_{xx}^i DD']$$

subject to $V^i(t_f, x) \equiv q_i(x)$, and the minimizing v is $\mu_i^*(t,x)$, $i = 1, \ldots, N$.

If $f = Ax + \sum_{i=1}^N B_i u_i$, $q_i = x' Q_i^f x$, $g_i = x' Q_i x + \sum_{j=1}^N u_j' R_{ij} u_j$, then **[Rec]** $V^i(t,x) = x' Z_i(t) x$ + constant, where $\{Z_i\}$ satisfy the coupled Riccati equations:

$$\dot{Z}_i + Z_i F + F' Z_i + Q_i + \sum_j Z_j B_j R_{jj}^{-1} R_{ij} R_{jj}^{-1} B_j' Z_j = 0\,, \; Z_i(t_f) = Q_i^f$$

where $F := A - \sum_i B_i R_{ii}^{-1} B_i' Z_i$. And the CLPS NE is $\mu_i^*(t,x) = -R_{ii}^{-1} B_i' Z_i(t) x_t$.

NE under CLPS for all players

$$dx_t = f(x_t, u_t)\,dt + \sqrt{\epsilon}\,D\,db_t\,; \quad x_{t=0} = x_0; \quad u_t := \{u_{1,t}, \ldots, u_{N,t}\}$$

Cost (expected loss) function of Player i over the interval $[t, t_f]$:

$$J_i(\mu; t, x_t) := E\left[q_i(x_{t_f}) + \int_t^{t_f} g_i(s, x_s, u_s)\,ds\right], \quad i = 1, \ldots, N, \quad \mu := \{\mu_1, \ldots, \mu_N\}$$

which Player i wants to minimize with $\mu_i \in M_i$, policy space compatible with IS

If the N-tuple μ^* is in NE, and $V^i(t,x) := J_i(\mu^*; t, x)$ is C_1 in t and C_2 in x, for each i, then V_i satisfies the coupled HJB equations **[Rec]**:

$$-V_t^i(t,x) = \min_{v \in M_i}\left\{V_x^i(t,x) f(x, v, \mu_{-i}^*(t,x)) + g_i(t, x, v, \mu_{-i}^*(t,x))\right\} + \frac{\epsilon}{2}\mathrm{Tr}[V_{xx}^i DD']$$

subject to $V^i(t_f, x) \equiv q_i(x)$, and the minimizing v is $\mu_i^*(t,x)$, $i = 1, \ldots, N$.

If $f = Ax + \sum_{i=1}^N B_i u_i$, $q_i = x' Q_i^f x$, $g_i = x' Q_i x + \sum_{j=1}^N u_j' R_{ij} u_j$, then **[Rec]** $V^i(t,x) = x' Z_i(t) x$ + constant, where $\{Z_i\}$ satisfy the coupled Riccati equations:

$$\dot{Z}_i + Z_i F + F' Z_i + Q_i + \sum_j Z_j B_j R_{jj}^{-1} R_{ij} R_{jj}^{-1} B_j' Z_j = 0\,, \; Z_i(t_f) = Q_i^f$$

where $F := A - \sum_i B_i R_{ii}^{-1} B_i' Z_i$. And the CLPS NE is $\mu_i^*(t,x) = -R_{ii}^{-1} B_i' Z_i(t) x_t$.

From risk neutrality to risk sensitivity

Loss function: $L(u,\xi)$; Measurement: y; Decision rule: $u = \mu(y)$

Risk-neutral cost: $\min_\mu \; E_{\xi|y} L(\mu(y),\xi)$

Risk-sensitive cost:

$J_\theta(\mu, y) = \frac{1}{\theta} \ln E_{\xi|y}\{\exp \theta L(\mu(y),\xi)\}$ $\Rightarrow$ minimize over decision rules μ

θ : risk sensitivity parameter

Around $\theta = 0$: $J_\theta(\mu, y) \sim E_{\xi|y} L + \frac{\theta}{2} \mathrm{var} L + O(\theta^2)$

$\theta = 0$: risk-neutral

$\theta < 0$: risk-seeking (optimistic)

$\theta > 0$: risk-averse (pessimistic)

If (ξ, y) is jointly Gaussian distributed :

$$\theta \exp J_\theta \sim \int e^{\theta L(\mu(y),\xi)} e^{-\frac{1}{2\sigma^2}(\xi - \alpha y)^2} d\xi$$

For finiteness, we need :

$$\theta L(\mu(y),\xi) < \frac{1}{2\sigma^2}\xi^2 \quad \text{as } |\xi| \to \infty$$

From risk neutrality to risk sensitivity

Loss function: $L(u,\xi)$; Measurement: y; Decision rule: $u = \mu(y)$

Risk-neutral cost: $\min_\mu E_{\xi|y} L(\mu(y),\xi)$

Risk-sensitive cost:

$J_\theta(\mu,y) = \frac{1}{\theta} \ln E_{\xi|y}\{\exp \theta L(\mu(y),\xi)\} \quad \Rightarrow \quad$ minimize over decision rules μ

θ : risk sensitivity parameter

Around $\theta = 0$: $J_\theta(\mu,y) \sim E_{\xi|y} L + \frac{\theta}{2} \mathrm{var} L + O(\theta^2)$

$\theta = 0$: risk-neutral
$\theta < 0$: risk-seeking (optimistic)
$\theta > 0$: risk-averse (pessimistic)

If (ξ, y) is jointly Gaussian distributed :

$$\theta \exp J_\theta \sim \int e^{\theta L(\mu(y),\xi)} e^{-\frac{1}{2\sigma^2}(\xi - \alpha y)^2} d\xi$$

For finiteness, we need :

$$\theta L(\mu(y),\xi) < \frac{1}{2\sigma^2}\xi^2 \quad \text{as } |\xi| \to \infty$$

From risk neutrality to risk sensitivity

Loss function: $L(u, \xi)$; Measurement: y; Decision rule: $u = \mu(y)$

Risk-neutral cost: $\min_\mu E_{\xi|y} L(\mu(y), \xi)$

Risk-sensitive cost:

$J_\theta(\mu, y) = \frac{1}{\theta} \ln E_{\xi|y}\{\exp \theta L(\mu(y), \xi)\}$ $\Rightarrow$ minimize over decision rules μ

θ : risk sensitivity parameter

Around $\theta = 0$: $J_\theta(\mu, y) \sim E_{\xi|y} L + \frac{\theta}{2} \mathrm{var} L + O(\theta^2)$

$\theta = 0$: risk-neutral
$\theta < 0$: risk-seeking (optimistic)
$\theta > 0$: risk-averse (pessimistic)

If (ξ, y) is jointly Gaussian distributed :

$$\theta \exp J_\theta \sim \int e^{\theta L(\mu(y), \xi)} e^{-\frac{1}{2\sigma^2}(\xi - \alpha y)^2} d\xi$$

For finiteness, we need :

$$\theta L(\mu(y), \xi) < \frac{1}{2\sigma^2} \xi^2 \quad \text{as } |\xi| \to \infty$$

An example

$$L(u,\xi) = (u-\xi)^2 + u^2\,; \quad \xi \sim N(1,1)$$

No measurement. Threshold on θ is $1/2$.

$J_\theta(u)$ unbounded for all u if $\theta \geq 1/2$

For $\theta < 1/2$, unique minimum is:

$$u_\theta = \frac{1}{2(1-\theta)}\,; \quad J_\theta(u_\theta) = \frac{1}{2(1-\theta)} - \frac{\ln(1-2\theta)}{2\theta}$$

$\Rightarrow$ $J_\theta(u_\theta)$ monotonically increasing in θ

Consider the stochastic zero-sum game ($\theta > 0$):

$$L_\theta(u,w,\xi) = (u+w-\xi)^2 + u^2 - \frac{1}{2\theta}w^2$$

$E[L_\theta(u,w,\xi)]$: Min wrt u and max wrt w

$$\Rightarrow \quad u_\theta = \frac{1}{2(1-\theta)}\,; \quad w_\theta = -\frac{\theta}{1-\theta}\,; \qquad \theta < 1/2$$

$$\Rightarrow E\left[(u_\theta + w - \xi)^2 + {u_\theta}^2\right] \leq \frac{1}{2\theta}w^2 + \frac{3-2\theta}{2(1-\theta)} \quad \forall\, w$$

THE TWO FORMULATIONS ARE EQUIVALENT FOR u_θ !

An example

$$L(u,\xi) = (u-\xi)^2 + u^2\,; \quad \xi \sim N(1,1)$$

No measurement. Threshold on θ is $1/2$.

$J_\theta(u)$ unbounded for all u if $\theta \geq 1/2$

For $\theta < 1/2$, unique minimum is:

$$u_\theta = \frac{1}{2(1-\theta)}\,; \quad J_\theta(u_\theta) = \frac{1}{2(1-\theta)} - \frac{\ln(1-2\theta)}{2\theta}$$

$\Rightarrow$ $J_\theta(u_\theta)$ monotonically increasing in θ

Consider the stochastic zero-sum game ($\theta > 0$):

$$L_\theta(u,w,\xi) = (u+w-\xi)^2 + u^2 - \frac{1}{2\theta}w^2$$

$E[L_\theta(u,w,\xi)]$: Min wrt u and max wrt w

$$\Rightarrow \quad u_\theta = \frac{1}{2(1-\theta)}\,; \quad w_\theta = -\frac{\theta}{1-\theta}\,; \qquad \theta < 1/2$$

$$\Rightarrow E\left[\,(u_\theta + w - \xi)^2 + {u_\theta}^2\,\right] \leq \frac{1}{2\theta}\,w^2 + \frac{3-2\theta}{2(1-\theta)} \quad \forall\, w$$

THE TWO FORMULATIONS ARE EQUIVALENT FOR u_θ !

An example

$$L(u,\xi) = (u-\xi)^2 + u^2\,; \quad \xi \sim N(1,1)$$

No measurement. Threshold on θ is $1/2$.

$J_\theta(u)$ unbounded for all u if $\theta \geq 1/2$

For $\theta < 1/2$, unique minimum is:

$$u_\theta = \frac{1}{2(1-\theta)}\,; \quad J_\theta(u_\theta) = \frac{1}{2(1-\theta)} - \frac{\ln(1-2\theta)}{2\theta}$$

$\Rightarrow$ $J_\theta(u_\theta)$ monotonically increasing in θ

Consider the stochastic zero-sum game ($\theta > 0$):

$$L_\theta(u,w,\xi) = (u+w-\xi)^2 + u^2 - \frac{1}{2\theta}w^2$$

$E[L_\theta(u,w,\xi)]$: Min wrt u and max wrt w

$$\Rightarrow \quad u_\theta = \frac{1}{2(1-\theta)}\,; \quad w_\theta = -\frac{\theta}{1-\theta}\,; \qquad \theta < 1/2$$

$$\Rightarrow E\left[\,(u_\theta + w - \xi)^2 + {u_\theta}^2\,\right] \leq \frac{1}{2\theta}\,w^2 + \frac{3-2\theta}{2(1-\theta)} \quad \forall\, w$$

THE TWO FORMULATIONS ARE EQUIVALENT FOR u_θ !

Now consider the case $\theta < 0$:

Risk-seeking

$$L_\theta(u, w, \xi) = (u + w - \xi)^2 + u^2 - \frac{1}{2\theta} w^2$$

$E[L_\theta(u, w, \xi)]$: Minimize wrt u and w

$$\Rightarrow \quad u_\theta = \frac{1}{2(1-\theta)}\,; \quad w_\theta = -\frac{\theta}{1-\theta}\,; \qquad \theta < 0$$

AGAIN THE TWO FORMULATIONS ARE EQUIVALENT FOR u_θ

Risk-Neutral Case $\theta = 0$

$$u = \frac{1}{2}\,; \qquad w = 0$$

Now consider the case $\theta < 0$:

Risk-seeking

$$L_\theta(u, w, \xi) = (u + w - \xi)^2 + u^2 - \frac{1}{2\theta} w^2$$

$E[L_\theta(u, w, \xi)]$: Minimize wrt u and w

$$\Rightarrow \quad u_\theta = \frac{1}{2(1-\theta)}\,; \quad w_\theta = -\frac{\theta}{1-\theta}\,; \qquad \theta < 0$$

AGAIN THE TWO FORMULATIONS ARE EQUIVALENT FOR u_θ

Risk-Neutral Case $\theta = 0$

$$u = \frac{1}{2}\,; \qquad w = 0$$

Now consider the case $\theta < 0$:

Risk-seeking

$$L_\theta(u, w, \xi) = (u + w - \xi)^2 + u^2 - \frac{1}{2\theta} w^2$$

$E[L_\theta(u, w, \xi)]$: Minimize wrt u and w

$$\Rightarrow \quad u_\theta = \frac{1}{2(1-\theta)}\,; \quad w_\theta = -\frac{\theta}{1-\theta}\,; \qquad \theta < 0$$

AGAIN THE TWO FORMULATIONS ARE EQUIVALENT FOR u_θ

Risk-Neutral Case $\quad \theta = 0$

$$u = \frac{1}{2}\,; \qquad w = 0$$

Same example with noisy information [Rec]

$$L(u,\xi) = (u-\xi)^2 + u^2\,; \quad \xi \sim N(1,1)\,, \ \ y-\xi \sim N(0,1)$$

Measurement y, with $y-\xi$ independent of ξ. Threshold on θ is 1.

For $\theta < 1$, unique minimizing policy is:

$$u_\theta = \mu_\theta(y) = \frac{1}{2(2-\theta)}(y+1)\,; \quad J_\theta(u_\theta) = \frac{1}{2(1-\theta)} - \frac{1}{2\theta}\ln\left[\frac{2(1-\theta)^2}{2-\theta}\right]$$

$\Rightarrow$ Again $J_\theta(u_\theta)$ monotonically increasing in θ

Consider the stochastic zero-sum game ($\theta > 0$):

$$L_\theta(u,w,\xi) = (u+w-\xi)^2 + u^2 - \frac{1}{\theta}w^2$$

$E[L_\theta(\mu(y),\nu(y),\xi)|y]$: Minimize wrt u and maximize wrt w

$$\Rightarrow \quad u_\theta = \mu_\theta(y) = \frac{1}{2(2-\theta)}(y+1)\,; \quad w_\theta = \nu_\theta(y) = \frac{\theta}{2(2-\theta)}(y+1)\ ; \qquad \theta < 1$$

THE TWO FORMULATIONS ARE EQUIVALENT FOR u_θ !

Same example with noisy information [Rec]

$$L(u,\xi) = (u-\xi)^2 + u^2\,; \quad \xi \sim N(1,1)\,, \;\; y-\xi \sim N(0,1)$$

Measurement y, with $y-\xi$ independent of ξ. Threshold on θ is 1.

For $\theta < 1$, unique minimizing policy is:

$$u_\theta = \mu_\theta(y) = \frac{1}{2(2-\theta)}(y+1)\,; \quad J_\theta(u_\theta) = \frac{1}{2(1-\theta)} - \frac{1}{2\theta}\ln\left[\frac{2(1-\theta)^2}{2-\theta}\right]$$

$\Rightarrow$ Again $J_\theta(u_\theta)$ monotonically increasing in θ

Consider the stochastic zero-sum game ($\theta > 0$):

$$L_\theta(u,w,\xi) = (u+w-\xi)^2 + u^2 - \frac{1}{\theta}w^2$$

$E[L_\theta(\mu(y),\nu(y),\xi)|y]$: Minimize wrt u and maximize wrt w

$$\Rightarrow \quad u_\theta = \mu_\theta(y) = \frac{1}{2(2-\theta)}(y+1)\,; \quad w_\theta = \nu_\theta(y) = \frac{\theta}{2(2-\theta)}(y+1)\;; \qquad \theta < 1$$

THE TWO FORMULATIONS ARE EQUIVALENT FOR u_θ !

Same example with noisy information [Rec]

$$L(u,\xi) = (u-\xi)^2 + u^2\,; \quad \xi \sim N(1,1)\,, \;\; y-\xi \sim N(0,1)$$

Measurement y, with $y-\xi$ independent of ξ. Threshold on θ is 1.

For $\theta < 1$, unique minimizing policy is:

$$u_\theta = \mu_\theta(y) = \frac{1}{2(2-\theta)}(y+1)\,; \quad J_\theta(u_\theta) = \frac{1}{2(1-\theta)} - \frac{1}{2\theta}\ln\left[\frac{2(1-\theta)^2}{2-\theta}\right]$$

$\Rightarrow$ Again $J_\theta(u_\theta)$ monotonically increasing in θ

Consider the stochastic zero-sum game ($\theta > 0$):

$$L_\theta(u,w,\xi) = (u+w-\xi)^2 + u^2 - \frac{1}{\theta}w^2$$

$E[L_\theta(\mu(y),\nu(y),\xi)|y]$: Minimize wrt u and maximize wrt w

$$\Rightarrow \quad u_\theta = \mu_\theta(y) = \frac{1}{2(2-\theta)}(y+1)\,; \quad w_\theta = \nu_\theta(y) = \frac{\theta}{2(2-\theta)}(y+1)\;; \qquad \theta < 1$$

THE TWO FORMULATIONS ARE EQUIVALENT FOR u_θ !

R-S stochastic control

State dynamics :

$$dx_t = f(t, x_t, u_t)\, dt + \sqrt{\epsilon}\, D\, db_t\,; \quad x_{t|t=0} = x_0$$

$b_t, t \geq 0$, standard Wiener process; $\epsilon > 0$;

$u_t \in U,\ t \geq 0$ (**state FB control law** $\mu \in \mathcal{M}$)

Objective : Choose μ **to minimize :** ($\theta > 0$)

$$J(\mu; t, x_t) = \frac{2\epsilon}{\theta} \ln E\left\{ \exp \frac{\theta}{2\epsilon} L(x_{[t,t_f]}, u_{[t,t_f]}) \right\}$$

$$L(x_{[t,t_f]}, u_{[t,t_f]}) := q(x_{t_f}) + \int_t^{t_f} g(s, x_s, u_s) ds$$

$\psi(t; x)$ – **value function associated with**

$$E\left\{ \exp \frac{\theta}{2\epsilon} \left[q(x_{t_f}) + \int_t^{t_f} g(s, x_s, u_s)\, ds \right] \right\}$$

$$\Rightarrow \qquad V(t; x) := \inf_{\mu \in \mathcal{M}} J(\mu; t, x) =: \frac{2\epsilon}{\theta} \ln \psi(t; x)\,,$$

R-S stochastic control

State dynamics :

$$dx_t = f(t, x_t, u_t)\, dt + \sqrt{\epsilon}\, D\, db_t\, ; \quad x_{t|t=0} = x_0$$

$b_t, t \geq 0$, standard Wiener process; $\epsilon > 0$;

$u_t \in U, \ t \geq 0$ (**state FB control law** $\mu \in \mathcal{M}$)

Objective : Choose μ to minimize : ($\theta > 0$)

$$J(\mu; t, x_t) = \frac{2\epsilon}{\theta} \ln E\Big\{ \exp \frac{\theta}{2\epsilon} L(x_{[t,t_f]}, u_{[t,t_f]})\Big\}$$

$$L(x_{[t,t_f]}, u_{[t,t_f]}) := q(x_{t_f}) + \int_t^{t_f} g(s, x_s, u_s) ds$$

$\psi(t; x)$ – **value function associated with**

$$E\Big\{ \exp \frac{\theta}{2\epsilon} \Big[q(x_{t_f}) + \int_t^{t_f} g(s, x_s, u_s)\, ds \Big] \Big\}$$

$$\Rightarrow \qquad V(t; x) := \inf_{\mu \in \mathcal{M}} J(\mu; t, x) =: \frac{2\epsilon}{\theta} \ln \psi(t; x)\, ,$$

R-S stochastic control

State dynamics :

$$dx_t = f(t, x_t, u_t)\, dt + \sqrt{\epsilon}\, D\, db_t\,; \quad x_{t|t=0} = x_0$$

$b_t, t \geq 0$, standard Wiener process; $\epsilon > 0$;

$u_t \in U, \;\; t \geq 0$ (**state FB control law** $\mu \in \mathcal{M}$)

Objective : Choose μ to minimize :   ($\theta > 0$)

$$J(\mu; t, x_t) = \frac{2\epsilon}{\theta} \ln E\{ \exp \frac{\theta}{2\epsilon} L(x_{[t,t_f]}, u_{[t,t_f]}) \}$$

$$L(x_{[t,t_f]}, u_{[t,t_f]}) := q(x_{t_f}) + \int_t^{t_f} g(s, x_s, u_s) ds$$

$\psi(t; x)$ – **value function associated with**

$$E\{ \exp \frac{\theta}{2\epsilon} \left[q(x_{t_f}) + \int_t^{t_f} g(s, x_s, u_s)\, ds \right] \}$$

$$\Rightarrow \qquad V(t; x) := \inf_{\mu \in \mathcal{M}} J(\mu; t, x) =: \frac{2\epsilon}{\theta} \ln \psi(t; x)\,,$$

DP and Itô differentiation rule [Rec] $\Rightarrow$

$$-V_t(t;x) = \inf_{u \in U} \left\{ V_x(t;x)\, f(t,x,u) + g(t,x,u) \right\} + \frac{1}{4\gamma^2} |DV_x'(t;x)|^2 + \frac{\epsilon}{2} \mathrm{Tr}\left[V_{xx}\, DD' \right]$$

$$V(t_f;x) \equiv q(x) \qquad \left(\gamma^{-2} := \theta\right)$$

If $U = \mathbf{R}^{m_1}$, f linear in u, and g quadratic in u :

$$f(x,u) = f_0(t,x) + B(t,x)u\,;\; g(t,x,u) = g_0(t,x) + |u|^2$$

Optimal control law: [Rec]

$$u^*(t) = \mu^*(t,x) = -\frac{1}{2} B'(t,x)\, V_x'(t;x)\,, \quad 0 \le t \le t_f$$

$\Rightarrow$ **HJB equation :**

$$-V_t = V_x\, f_0(t,x) + g_0(t,x) - \frac{1}{4}\left[|BV_x'|^2 - \gamma^{-2} |DV_x'|^2 \right] + \frac{\epsilon}{2} \mathrm{Tr}\left[V_{xx}(t;x) DD' \right]; \qquad V(t_f;x) \equiv q(x)$$

DP and Itô differentiation rule [Rec] $\Rightarrow$

$$-V_t(t;x) = \inf_{u \in U} \left\{ V_x(t;x)\, f(t,x,u) + g(t,x,u) \right\}$$

$$+\frac{1}{4\gamma^2}|DV_x'(t;x)|^2 + \frac{\epsilon}{2}\mathrm{Tr}\left[V_{xx}\, DD'\right]$$

$$V(t_f;x) \equiv q(x) \qquad \left(\gamma^{-2} := \theta\right)$$

If $U = \mathbf{R}^{m_1}$, f linear in u, and g quadratic in u :

$$f(x,u) = f_0(t,x) + B(t,x)u\,;\; g(t,x,u) = g_0(t,x) + |u|^2$$

Optimal control law: [Rec]

$$u^*(t) = \mu^*(t,x) = -\frac{1}{2}\, B'(t,x)\, V_x'(t;x)\,, \quad 0 \le t \le t_f$$

$\Rightarrow$ **HJB equation :**

$$-V_t = V_x\, f_0(t,x) + g_0(t,x) - \frac{1}{4}\left[|BV_x'|^2 - \gamma^{-2}|DV_x'|^2\right]$$

$$+\frac{\epsilon}{2}\mathrm{Tr}\left[V_{xx}(t;x)DD'\right]; \qquad V(t_f;x) \equiv q(x)$$

A further special case : LEQG Problem

$$f_0(t,x) = A(t)x\,, \quad g_0(t,x) = \frac{1}{2}\,x'Qx\,,\ Q \geq 0$$

$$q(x) = (1/2)\,x'Q_f x$$

⇒ **Explicit solution:**

$$V(t;x) = \frac{1}{2}\,x'Z(t)x + \ell^{\epsilon}(t)\,, \quad t \geq 0$$

$$\dot{Z} + A'Z + ZA + Q - Z\,(\,BB' - \gamma^{-2}DD'\,)\,Z = 0$$

$$\ell^{\epsilon}(t) = \frac{\epsilon}{2}\int_t^{t_f} \mathrm{Tr}\,[\,Z(s)\,D(s)D'(s)\,]\,ds$$

$$\Rightarrow \quad u^*(t) = \mu^*(t,x) = -B'(t)\,Z(t)\,x\,, \quad 0 \leq t \leq t_f$$

A class of stochastic differential games **[Rec]**

Two Players : **Player 1:** u_t ; **Player 2:** w_t

$$dx_t = f(x_t, u_t)\,dt + D(x_t) w_t\,dt + \sqrt{\epsilon}\,D\,db_t\,; \quad x_0$$

$$J(\mu,\nu;t,x_t) := E\Big\{q(x_{t_f}) + \int_t^{t_f} g(s,x_s,u_s)\,ds - \gamma^2 \int_t^{t_f} |w_s|^2\,ds\Big\}$$

Upper-Value (UV) Function :

$$\bar{W}(t;x) = \inf_{\mu} \sup_{\nu} J(\mu,\nu;t,x)$$

HJI UV equation :

$$\inf_{u\in U} \sup_{w\in \mathbf{R}^{m_2}} \Big\{\bar{W}_t + \bar{W}_x\,(\,f + Dw\,) + g - \gamma^2|w|^2 + \frac{\epsilon}{2}\mathrm{Tr}\big[\bar{W}_{xx}DD'\big]\Big\} = 0$$

A class of stochastic differential games **[Rec]**

Two Players : **Player 1:** u_t **;** **Player 2:** w_t

$$dx_t = f(x_t, u_t)\,dt + D(x_t)w_t\,dt + \sqrt{\epsilon}\,D\,db_t\,; \quad x_0$$

$$J(\mu, \nu; t, x_t) := E\Big\{q(x_{t_f}) + \int_t^{t_f} g(s, x_s, u_s)\,ds - \gamma^2 \int_t^{t_f} |w_s|^2\,ds\Big\}$$

Upper-Value (UV) Function :

$$\bar{W}(t; x) = \inf_{\mu} \sup_{\nu} J(\mu, \nu; t, x)$$

HJI UV equation :

$$\inf_{u \in U} \sup_{w \in \mathbf{R}^{m_2}} \Big\{\bar{W}_t + \bar{W}_x\,(\,f + Dw\,) + g - \gamma^2|w|^2 + \frac{\epsilon}{2}\mathrm{Tr}\big[\bar{W}_{xx}DD'\big]\Big\} = 0$$

HJI UV equation :

$$\inf_{u \in U} \sup_{w \in \mathbf{R}^{m_2}} \left\{ \bar{W}_t + \bar{W}_x \,(f + Dw \,) + g - \gamma^2 |w|^2 \right.$$

$$\left. + \frac{\epsilon}{2} \mathrm{Tr}\left[\bar{W}_{xx} DD' \right] \right\} = 0$$

Isaacs condition holds $\Rightarrow$ **Value Function :**

$$-W_t(t;x) = \inf_{u \in U} \left\{ W_x(t;x)\, f(t,x,u) + g(t,x,u) \right\}$$

$$+ \frac{1}{4\gamma^2} |DW_x'(t;x)|^2 + \frac{\epsilon}{2} \mathrm{Tr}\left[W_{xx}(t;x) DD' \right];$$

$$W(t_f;x) \equiv q(x)$$

- IDENTICAL with V for all permissible ϵ, γ

HJI UV equation :

$$\inf_{u\in U}\ \sup_{w\in \mathbf{R}^{m_2}} \left\{\bar{W}_t + \bar{W}_x\,(\,f + Dw\,) + g - \gamma^2|w|^2\right.$$

$$\left. + \frac{\epsilon}{2}\mathrm{Tr}\left[\bar{W}_{xx}DD'\right]\right\} = 0$$

Isaacs condition holds $\Rightarrow$ **Value Function :**

$$-W_t(t;x) = \inf_{u\in U}\left\{W_x(t;x)\,f(t,x,u) + g(t,x,u)\right\}$$

$$+\frac{1}{4\gamma^2}|DW_x'(t;x)|^2 + \frac{\epsilon}{2}\,\mathrm{Tr}\left[W_{xx}(t;x)DD'\right];$$

$$W(t_f;x) \equiv q(x)$$

- IDENTICAL with V for all permissible ϵ, γ

Implication

Original stochastic dynamics

$$dx_t = f(x_t, u_t)\,dt + \sqrt{\epsilon}\,D\,db_t\,; \quad x_0$$

Optimum RS FB control and value:

$$u_t = \mu^*(t, x_t)\,; \quad V(t; x_t; t_f)\,, \quad t \geq t_0$$

Under μ^*, and for the perturbed dynamics

$$d\tilde{x}_t = f(\tilde{x}_t, u_t)\,dt + D(\tilde{x}_t)w_t\,dt + \sqrt{\epsilon}\,D\,db_t$$

$$E\left\{q(\tilde{x}_{t_f}) + \int_t^{t_f} g(s, \tilde{x}_s, \mu^*(\tilde{x}_s, s))\,ds\right\}$$

$$\leq \gamma^2 \int_t^{t_f} |w_s|^2\,ds + V(t; \tilde{x}_t; t_f)$$

In particular, if $\tilde{z}$ is controlled output:

$$\frac{1}{t_f}\int_0^{t_f} E\{|\tilde{z}_s|^2\} \leq \frac{\gamma^2}{t_f}\int_0^{t_f} |w_s|^2\,ds + \frac{1}{t_f}V(0; x_0; t_f)$$

A STOCHASTIC DISSIPATION INEQUALITY !

ROBUSTNESS TO MODELING ERROR !!

Implication

Original stochastic dynamics

$$dx_t = f(x_t, u_t)\, dt + \sqrt{\epsilon}\, D\, db_t\,; \quad x_0$$

Optimum RS FB control and value:

$$u_t = \mu^*(t, x_t)\,; \quad V(t; x_t; t_f)\,, \quad t \geq t_0$$

Under μ^*, and for the perturbed dynamics

$$d\tilde{x}_t = f(\tilde{x}_t, u_t)\, dt + D(\tilde{x}_t) w_t\, dt + \sqrt{\epsilon}\, D\, db_t$$

$$E\Big\{q(\tilde{x}_{t_f}) + \int_t^{t_f} g(s, \tilde{x}_s, \mu^*(\tilde{x}_s, s))\, ds\Big\}$$

$$\leq \gamma^2 \int_t^{t_f} |w_s|^2\, ds + V(t; \tilde{x}_t; t_f)$$

In particular, if $\tilde{z}$ is controlled output:

$$\frac{1}{t_f} \int_0^{t_f} E\{|\tilde{z}_s|^2\} \leq \frac{\gamma^2}{t_f} \int_0^{t_f} |w_s|^2\, ds + \frac{1}{t_f} V(0; x_0; t_f)$$

A STOCHASTIC DISSIPATION INEQUALITY !

ROBUSTNESS TO MODELING ERROR !!

RS stochastic differential games with two players **[Rec]**

Two Players : **Player 1:** u_t **;** **Player 2:** v_t

$$dx_t = f(x_t, u_t, v_t)\, dt + \sqrt{\epsilon}\, D(x_t, u_t, v_t)\, db_t\, ; \quad x_0$$

$$J_i(\mu, \nu; t, x_t) := \frac{2\epsilon}{\theta} \ln E\Big\{ \exp \frac{\theta}{2\epsilon} \Big[q_i(x_{t_f}) + \int_t^{t_f} g_i(s, x_s, u_s, v_s)\, ds\Big]\, , \quad i = 1, 2$$

$J_1 + J_2 \equiv 0 \quad \Rightarrow$ **RSZSDG; saddle-point equilibria**

Otherwise, **RSNZS SDG; Nash equilibrium**

- In both cases, the equilibrium solutions admit robustness interpretation, with the players having virtual adversaries, making the underlying game a neutral 4-player NZS SDG, with Nash equilibrium as the solution concept—obtained from coupled HJI equations.
- This naturally extends to N players.

RS stochastic differential games with two players **[Rec]**

Two Players : **Player 1:** u_t **;** **Player 2:** v_t

$$dx_t = f(x_t, u_t, v_t)\, dt + \sqrt{\epsilon}\, D(x_t, u_t, v_t)\, db_t\,; \quad x_0$$

$$J_i(\mu, \nu; t, x_t) := \frac{2\epsilon}{\theta} \ln E\Big\{ \exp \frac{\theta}{2\epsilon} \Big[q_i(x_{t_f}) + \int_t^{t_f} g_i(s, x_s, u_s, v_s)\, ds \Big]\,, \quad i = 1, 2$$

$J_1 + J_2 \equiv 0 \quad \Rightarrow$ **RSZSDG; saddle-point equilibria**

Otherwise, **RSNZS SDG; Nash equilibrium**

- In both cases, the equilibrium solutions admit robustness interpretation, with the players having virtual adversaries, making the underlying game a neutral 4-player NZS SDG, with Nash equilibrium as the solution concept—obtained from coupled HJI equations.
- This naturally extends to N players.

RS stochastic differential games with two players [Rec]

Two Players : **Player 1:** u_t ; **Player 2:** v_t

$$dx_t = f(x_t, u_t, v_t)\, dt + \sqrt{\epsilon}\, D(x_t, u_t, v_t)\, db_t\,; \quad x_0$$

$$J_i(\mu, \nu; t, x_t) := \frac{2\epsilon}{\theta} \ln E\Big\{ \exp \frac{\theta}{2\epsilon} \Big[q_i(x_{t_f}) + \int_t^{t_f} g_i(s, x_s, u_s, v_s)\, ds \Big], \quad i = 1, 2$$

$J_1 + J_2 \equiv 0 \quad \Rightarrow$ **RSZSDG; saddle-point equilibria**

Otherwise, **RSNZS SDG; Nash equilibrium**

- In both cases, the equilibrium solutions admit robustness interpretation, with the players having virtual adversaries, making the underlying game a neutral 4-player NZS SDG, with Nash equilibrium as the solution concept—obtained from coupled HJI equations.
- This naturally extends to N players.

Recap and final thoughts

- Complexity of equilibrium solutions in SDGs (existence, uniqueness, characterization) depends very much on the IS of the game, with elevated difficulty if IS is asymmetric, particularly with noisy information. **[Rec]**
- Generally no relationship between NE under different ISs (this applies also to deterministic DGs, unless DG is ZS). **[Rec]**
- Stackelberg SDGs with dynamic information for leader(s) lead to even more complex optimization problems–connections to mechanism design. **[Rec]**
- Some special structures as in MFGs, with infinite population of players, make otherwise unsolvable problems manageable, as next two lectures will show.

An insightful statement on games with a large number of agents:

- Theory of Games and Economic Behavior (1944, pp.13-14)
- "When the number of participants becomes really great, some hope emerges that the influence of every particular participant will become negligible and that the above difficulties may recede and a more conventional theory becomes possible."
- "It is well known phenomenon in many branches of the exact and physical sciences that very great numbers are often easier to handle than those of medium size... This is of course, due to the excellent possibility of applying the laws of statistics and probabilities in the first case"

Recap and final thoughts

- Complexity of equilibrium solutions in SDGs (existence, uniqueness, characterization) depends very much on the IS of the game, with elevated difficulty if IS is asymmetric, particularly with noisy information. **[Rec]**
- Generally no relationship between NE under different ISs (this applies also to deterministic DGs, unless DG is ZS). **[Rec]**
- Stackelberg SDGs with dynamic information for leader(s) lead to even more complex optimization problems–connections to mechanism design. **[Rec]**
- Some special structures as in MFGs, with infinite population of players, make otherwise unsolvable problems manageable, as next two lectures will show.

An insightful statement on games with a large number of agents:

- Theory of Games and Economic Behavior (1944, pp.13-14)
- "When the number of participants becomes really great, some hope emerges that the influence of every particular participant will become negligible and that the above difficulties may recede and a more conventional theory becomes possible."
- "It is well known phenomenon in many branches of the exact and physical sciences that very great numbers are often easier to handle than those of medium size... This is of course, due to the excellent possibility of applying the laws of statistics and probabilities in the first case"

Recap and final thoughts

- Complexity of equilibrium solutions in SDGs (existence, uniqueness, characterization) depends very much on the IS of the game, with elevated difficulty if IS is asymmetric, particularly with noisy information. **[Rec]**
- Generally no relationship between NE under different ISs (this applies also to deterministic DGs, unless DG is ZS). **[Rec]**
- Stackelberg SDGs with dynamic information for leader(s) lead to even more complex optimization problems–connections to mechanism design. **[Rec]**
- Some special structures as in MFGs, with infinite population of players, make otherwise unsolvable problems manageable, as next two lectures will show.

An insightful statement on games with a large number of agents:

- Theory of Games and Economic Behavior (1944, pp.13-14)
- "When the number of participants becomes really great, some hope emerges that the influence of every particular participant will become negligible and that the above difficulties may recede and a more conventional theory becomes possible."
- "It is well known phenomenon in many branches of the exact and physical sciences that very great numbers are often easier to handle than those of medium size... This is of course, due to the excellent possibility of applying the laws of statistics and probabilities in the first case"

Selected references for Lecture I

- T. Başar and G.J. Olsder, Dynamic Noncooperative Game Theory, SIAM, 1999.
- T. Başar and P. Bernhard, H^∞ Optimal Control and Related Minimax Design Problems, Birkhäuser, 1995.
- T. Başar, "Nash equilibria of risk-sensitive nonlinear stochastic differential games," JOTA, 100(3): 479-498, 1999.
- T. Başar, "Risk-averse designs: From exponential cost to stochastic games," in T.E. Djaferis and I.C. Schick, editors, System Theory: Modeling, Analysis and Control, Kluwer, pp. 131-144, 2000.
- T. Başar, "Stochastic differential games and intricacy of information structures," in J. Haunschmied, V.M. Veliov, and S. Wrzaczek (Eds), Dynamic Games in Economics, Springer, 2014, pp. 23-49.
- T. Başar, "Introduction to the theory of games," in T. Başar and G. Zaccour (Eds.), Handbook of Dynamic Game Theory, Springer International Publishing, 2017.
- T. Başar, A. Haurie, and G. Zaccour, "Nonzero-sum differential games," in T. Başar and G. Zaccour (Eds.), Handbook of Dynamic Game Theory, Springer International Publishing, pp. 1-49, 2016.
- P. Whittle, Risk-Sensitive Optimal Control, Wiley, 1990