Dynamic Incentive Design in Multi-Stage LQG Game with Asymmetric Information

Existence of Dynamic Incentive Scheme

Abhishek Gupta
Electrical Engineering Department
University of Southern California

November 20, 2015

Model

- $N + 1$-player T-horizon stochastic LQG game

$$\mathbf{X}_{t+1} = f_t(\mathbf{X}_t, \mathbf{U}_t^{0:N}, \mathbf{W}_t) \qquad \mathbf{Y}_t^i = h_t^i(\mathbf{X}_t, \mathbf{V}_t^i)$$

- $\{\mathbf{X}_1, \mathbf{W}_{1:T}, \mathbf{V}_{1:T}^{0:N}\}$ - mutually independent Gaussian random variables, f_t and h_t^i linear

MODEL

- $N + 1$-player T-horizon stochastic LQG game

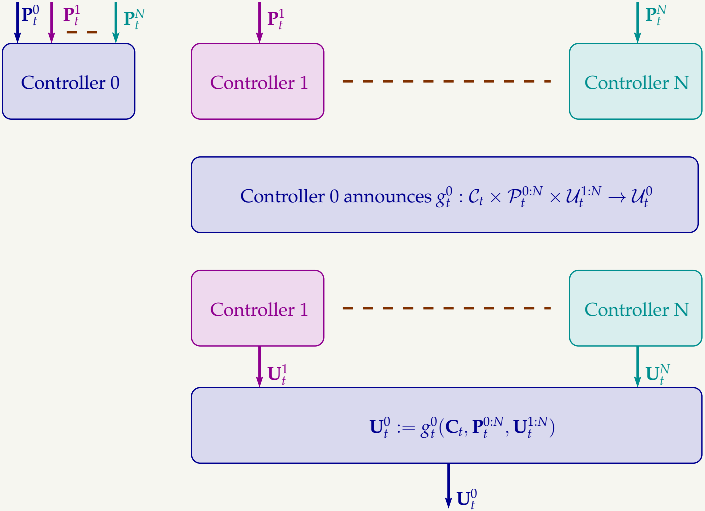
$$\mathbf{X}_{t+1} = f_t(\mathbf{X}_t, \mathbf{U}_t^{0:N}, \mathbf{W}_t) \qquad \mathbf{Y}_t^i = h_t^i(\mathbf{X}_t, \mathbf{V}_t^i)$$

- $\{\mathbf{X}_1, \mathbf{W}_{1:T}, \mathbf{V}_{1:T}^{0:N}\}$ - mutually independent Gaussian random variables, f_t and h_t^i linear

- $\mathbf{I}_t^i$ - information of DM i at time t

$$\mathbf{I}_t^i \subset \left\{\mathbf{Y}_{1:t}^{0:N}, \mathbf{U}_{1:t-1}^{0:N}\right\}, \qquad \mathbf{C}_t = \bigcap_{i=0}^{N} \mathbf{I}_t^i, \qquad \mathbf{P}_t^i = \mathbf{I}_t^i \setminus \mathbf{C}_t$$

Model


$\mathbf{P}_t^0$
$\mathbf{P}_t^1$
$\mathbf{P}_t^N$
Controller 0
$\mathbf{P}_t^1$
Controller 1
$\mathbf{P}_t^N$
Controller N
Controller 0 announces $g_t^0 : \mathcal{C}_t \times \mathcal{P}_t^{0:N} \times \mathcal{U}_t^{1:N} \to \mathcal{U}_t^0$
Controller 1
Controller N
$\mathbf{U}_t^1$
$\mathbf{U}_t^N$
$\mathbf{U}_t^0 := g_t^0(\mathbf{C}_t, \mathbf{P}_t^{0:N}, \mathbf{U}_t^{1:N})$
$\mathbf{U}_t^0$

Strategy and Cost Function

- Total cost to Controller i in the game is

$$J^i(\mathbf{g}^{0:N}) = \mathbb{E}\left[c_T^i(\mathbf{X}_T) + \sum_{t=1}^{T-1} c_t^i(\mathbf{X}_t, \mathbf{U}_t^0, \mathbf{U}_t^{1:N})\right]$$

- c_t^i– convex, subquadratic
- **MSID** - multi-stage incentive design problem

Goal

Admissible Incentive Policy

$\mathbf{g}^{0:N\star}$ is an admissible incentive strategy profile if

$$\begin{aligned} \text{Controller 0's optimization} \quad & J^0(\mathbf{g}^{0:N\star}) \leq J^0(\mathbf{g}^{0:N}) \\ \text{and Nash equilibrium} \quad & J^i(\mathbf{g}^{0:N\star}) \leq J^i(\mathbf{g}^i, \mathbf{g}^{-i\star}) \end{aligned}$$

Goal

Admissible Incentive Policy

$\mathbf{g}^{0:N\star}$ is an admissible incentive strategy profile if

$$\begin{array}{rlcl} \text{Controller 0's optimization} & J^0(\mathbf{g}^{0:N\star}) & \leq & J^0(\mathbf{g}^{0:N}) \\ \text{and Nash equilibrium} & J^i(\mathbf{g}^{0:N\star}) & \leq & J^i(\mathbf{g}^i, \mathbf{g}^{-i\star}) \end{array}$$

Common Information Based Subgame-Perfect Incentive Scheme

1. $\mathbf{g}^{0:N\star}$ is an admissible incentive strategy profile

   Let $\mathbf{G}_t(\mathbf{c}_t)$ denote the incentive design subgame with $\mathbf{c}_t$

2. For every $t \in [T-1]$, $\{g^{0\star}_{t:T-1}, \ldots, g^{N\star}_{t:T-1}\}$ is an admissible incentive strategy profile for subgame $\mathbf{G}_t(\cdot)$

Assumptions on Information Structures

1. Common information increases $\mathbf{C}_t \subset \mathbf{C}_{t+1}$ and

$$\mathbf{P}^i_{t+1} = \xi^i_{t+1}([\mathbf{P}^{i\intercal}_t, \mathbf{U}^{i\intercal}_t, \mathbf{Y}^{i\intercal}_{t+1}]^\intercal)$$

2. Define $\mathcal{S}_t := \mathcal{X}_t \times \mathcal{P}^{0:N}_t$ and *common information based conditional belief* Π_t as

$$\Pi_t(d\mathbf{s}_t) = \mathbb{P}^{g^{0:N}_{1:t-1}} \left\{ d\mathbf{s}_t \middle| \mathbf{C}_t \right\}$$

   Π_t is independent of $\mathbf{g}^{0:N}$ at each time t

3. Information structure is partially nested
4. c^0_t is quadratic

Main Results

- Assumptions 3 and 4 $\implies$ Controller 0's optimization problem admits affine optimal solution

$$J^0(\alpha_{1:T-1}^{0:N\star}) \leq J^0(g_{1:T-1}^{0:N})$$

- Incentive strategy of Controller 0 is affine
- Best response of followers are also affine (even with non-quadratic cost)
- in the class of all Borel strategies

SSID Problem

- State X
- Controller i observes Y^i and takes an action U^i
- Cost $i = b^i$ (convex, differentiable and subquaratic)

SSID PROBLEM

- State X
- Controller i observes Y^i and takes an action U^i
- Cost $i = b^i$ (convex, differentiable and subquaratic)
- $\exists\ \alpha^{0:N\star}$ – solution to Controller 0's optimization problem
- **SSID** - single-stage incentive design problem
- Use supporting hyperplane theorem to design incentive scheme for Controller 0

$$P^i(\mathsf{y}^{0:N}) \;:=\; \mathbb{E}\left[\frac{\partial b^i}{\partial \mathsf{u}^0}(\mathsf{X}, \mathsf{Y}^{0:N}, \mathsf{U}^{0:N})\Big|_{\mathsf{U}^0 := \alpha^{0\star}(\mathsf{Y}^{0:N}), \mathsf{U}^i = \alpha^{i\star}(\mathsf{Y}^i)} \Big| \mathsf{y}^{0:N} \right]$$

$$M^i(\mathsf{y}^i) \;:=\; \mathbb{E}\left[\frac{\partial b^i}{\partial \mathsf{u}^i}(\mathsf{X}, \mathsf{Y}^{0:N}, \mathsf{U}^{0:N})\Big|_{\mathsf{U}^0 := \alpha^{0\star}(\mathsf{Y}^{0:N}), \mathsf{U}^i = \alpha^{i\star}(\mathsf{Y}^i)} \Big| \mathsf{y}^i \right]$$

Assumption: $\|P^i\|^2_{\mathfrak{H}_{\mathsf{y}^i}(\mathcal{Y}^{-i} \to \mathcal{U}^0)} \neq 0$ for μ-almost every y^i

$$Q^i(\mathsf{y}^{0:N})[\mathsf{u}^i] = \frac{P^i(\mathsf{y}^{0:N}) M^i(\mathsf{y}^i)}{\|P^i\|^2_{\mathfrak{H}_{\mathsf{y}^i}(\mathcal{Y}^{-i} \to \mathcal{U}^0)}} \mathsf{u}^i.$$

$$P^i(\mathsf{y}^{0:N}) := \mathbb{E}\left[\frac{\partial b^i}{\partial \mathsf{u}^0}(\mathsf{X}, \mathsf{Y}^{0:N}, \mathsf{U}^{0:N})\Big|_{\mathsf{U}^0:=\alpha^{0\star}(\mathsf{Y}^{0:N}),\mathsf{U}^i=\alpha^{i\star}(\mathsf{Y}^i)}\middle|\mathsf{y}^{0:N}\right]$$

$$M^i(\mathsf{y}^i) := \mathbb{E}\left[\frac{\partial b^i}{\partial \mathsf{u}^i}(\mathsf{X}, \mathsf{Y}^{0:N}, \mathsf{U}^{0:N})\Big|_{\mathsf{U}^0:=\alpha^{0\star}(\mathsf{Y}^{0:N}),\mathsf{U}^i=\alpha^{i\star}(\mathsf{Y}^i)}\middle|\mathsf{y}^i\right]$$

Assumption: $\|P^i\|^2_{\mathfrak{H}_{\mathsf{y}^i}(\mathcal{Y}^{-i}\to\mathcal{U}^0)} \neq 0$ for μ-almost every y^i

$$Q^i(\mathsf{y}^{0:N})[\mathsf{u}^i] = \frac{P^i(\mathsf{y}^{0:N})M^i(\mathsf{y}^i)}{\|P^i\|^2_{\mathfrak{H}_{\mathsf{y}^i}(\mathcal{Y}^{-i}\to\mathcal{U}^0)}}\mathsf{u}^i.$$

Theorem

Incentive strategy of Controller 0 given by

$$g^{0\star}(\mathsf{y}^{0:N}, \mathsf{u}^{1:N}) = \alpha^{0\star}(\mathsf{y}^{0:N}) - \left(\sum_{i=1}^{N}\left(Q^i(\mathsf{y}^{0:N})[\mathsf{u}^i - \alpha^{i\star}(\mathsf{y}^i)]\right)\right)$$

MSID Problem

Let $\alpha_{1:T-1}^{0:N\star}$ = Controller 0's optimal solution

$$V_t^i(\mathbf{c}_t) := \mathbb{E}\Bigg[c_t^i\Big(\mathbf{X}_t, \alpha_t^{0\star}(\mathbf{c}_t, \mathbf{P}_t^{0:N}), \alpha_t^{1\star}(\mathbf{c}_t, \mathbf{P}_t^1), \ldots, \alpha_t^{N\star}(\mathbf{c}_t, \mathbf{P}_t^N)\Big) + V_{t+1}^i(\mathbf{C}_{t+1})\Bigg|\mathbf{c}_t\Bigg]$$

1. $\mathbf{C}_{t+1}$ is affine in $\mathbf{U}_t^{0:N}$ for all $t \in [T-2]$
2. V_t^i is convex and differentiable in $\mathbf{C}_t$ for all $t \in [T-1]$
3. V_{t+1}^i is convex and differentiable in $\mathbf{U}_t^i$ for all $t \in [T-2]$
4. V_t^i is subquadratic

MSID Problem

$$b^i_{T-1}(\mathbf{c}_{T-1}; \mathbf{x}_{T-1}, \mathbf{p}^{0:N}_{T-1}, \mathbf{u}^{0:N}_{T-1}) = c^i_{T-1}(\mathbf{x}_{T-1}, \mathbf{u}^{0:N}_{T-1}) + \mathbb{E}\left[c^i_T(\mathbf{X}_T) \middle| \mathbf{c}_{T-1}, \mathbf{x}_{T-1}, \mathbf{p}^{0:N}_{T-1}, \mathbf{u}^{0:N}_{T-1} \right]$$

$$b^i_s(\mathbf{c}_s; \mathbf{x}_s, \mathbf{p}^{0:N}_s, \mathbf{u}^{0:N}_s) = c^i_s(\mathbf{x}_s, \mathbf{u}^{0:N}_s) + \mathbb{E}\left[V^i_{t+1}(\mathbf{C}_{t+1}) \middle| \mathbf{c}_s, \mathbf{x}_s, \mathbf{p}^{0:N}_s, \mathbf{u}^{0:N}_s \right]$$

Decompose MSID into a sequence of SSID$(t, \mathbf{c}_t)$ problems

1. The state of the system is $\mathbf{X}_t$
2. The observation of Controller i is $\mathbf{P}^i_t$
3. The probability measure over $\mathcal{X}_t \times \mathcal{P}^{0:N}_t$ is π_t
4. The control action of Controller i is $\mathbf{U}^i_t$
5. The cost functions of Controller i is $b^i_t(\mathbf{c}_t; \cdot)$