Stochastic Control of Mean Field Models
with Mixed Players

Minyi Huang ∗ Son Luu Nguyen ∗

∗ School of Mathematics and Statistics, Carleton University, Ottawa,
ON K1S 5B6, Canada (e-mail: {mhuang, snguyen}@math.carleton.ca)

Abstract: We consider mean field stochastic systems consisting of a major player and a large number of minor players. We study decentralized strategies for both game and social optimization problems. For the game problem, the objective is to obtain asymptotic Nash equilibria. For the social optimization problem, the objective is to nearly minimize a weighted sum of the individual costs. A very peculiar feature of the two problems is that the presence of the major player causes a lack of sufficient statistics for decentralized decision-making. This difficulty is overcome by a state space augmentation technique, which leads to stochastic, rather than deterministic, mean field approximations. For the social optimization problem, we apply a social cost perturbation argument to the augmented model to design decentralized strategies.

1. INTRODUCTION

Mean field dynamic decision problems have been extensively studied in recent years [1, 2, 6, 7, 8, 9, 11, 12, 13, 14, 16, 18, 20], and a central objective is to obtain decentralized strategies such that each player only uses local information. In a noncooperative game theoretic context, many authors have contributed to the development of decentralized solutions [1, 6, 9, 12, 18, 20]; a fundamental idea is to consider the large population limit and determine the mean field effect such that each agent optimally responds to that effect and such a behavior of all the agents also collectively replicates that mean field effect. This approach is called the Nash certainty equivalence (NCE) methodology [6, 9]. More recently it is shown that such a consistency based approach may be extended to social optimization problems where all the agents cooperatively minimize a social cost as the sum of their individual costs [7]. The decentralized control synthesis is achieved by calculation of the social cost change caused by the control perturbation of a single agent; the central result is the so-called social optimality theorem which states that the optimality loss of the obtained decentralized strategies tends to zero when the population size goes to infinity [8].

The recent work [5] investigates a mean field linear-quadratic-Gaussian (LQG) game model where the population consists of a major player and a large number of minor players, rather than just a population of comparably small agents (which may be called peers). Under this modeling, the dynamics and the cost of the major player receive an average effect of all minor players, but each minor player receives a significant impact from the major player, apart from the mean field effect of all the minor players. Traditionally, models differentiating the strength of players have been well studied in cooperative game theory, and they are customarily called mixed games with the players according called mixed players [3]. A surprising consequence of the modeling in [5] is that the set of decentralized Nash strategies must use additional information reflecting the stochastic evolution of the mean field. For instance, the strategy of the major player cannot just be a function of its own current state and time. This situation is termed as the lack of sufficient statistics. In this paper we first describe the state space augmentation approach developed in [5] for the construction of a complete set of sufficient statistics. Next, we consider the social optimization problem involving a major player and we combine the state space augmentation and the social cost perturbation to design decentralized strategies. For cooperative differential games using other optimality notions, see [4, 15, 19].

1.1 The mean field model

We consider the LQG mean field decision model with a major player $A_0$ and minor players $\{A_i, 1 \leq i \leq N\}$. At time $t \geq 0$, the states of $A_0$ and $A_i$, respectively, denoted by $x_0(t)$ and $x_i(t), 1 \leq i \leq N$. Let $(\Omega, \mathcal{F}, \mathcal{F}_t, t \geq 0, P)$ be the underlying filtration. The dynamics of the $N+1$ agents are given by a system of linear stochastic differential equations (SDE’s) with mean field coupling:

\[
\begin{align*}
    dx_0 &= [A_0 x_0 + B_0 v_0 + F_0 x^{(N)}]dt + D_0 dW_0, \\
    dx_i &= [A_i x_i + B_i u_i + F_i x^{(N)} + G_i x_0]dt + D dW_i,
\end{align*}
\]

where $x^{(N)} = (1/N) \sum_{i=1}^{N} x_i$ is the mean field term. The initial states are measurable on $\mathcal{F}_0$ and are, respectively, $x_0(0)$ and $x_i(0), 1 \leq i \leq N$. We may choose $\mathcal{F}_t$ as the $\sigma$-algebra $\mathcal{F}_t^{x_i(0), W_j} = \sigma(x_j(0), W_j(\tau), 0 \leq j \leq N, \tau \leq t)$. In (2), $x_0$ has a constant coefficient $G$ while each $x_j$ as a component in $x^{(N)}$, $0 < j \neq i$, is associated with a factor $1/N$. This modeling feature indicates that $A_0$ has a significant influence on others while, in contrast, each minor player only has a negligible impact on others for large $N$. Our modeling may be generalized to multiple major players and multiple classes of minor players, but for simplicity of analysis we will focus on the above model.
The states \( x_0, x_1 \) and controls \( u_0, u_1 \) are, respectively, \( n \) and \( n_1 \) dimensional vectors. The noise processes \( W_0, W_1 \) are \( n_2 \) dimensional independent standard Brownian motions adapted to \( F_1 \), which are also independent of \( (x_j(0), 0 \leq j \leq N) \). The constant matrices \( A_0, B_0, F_0, D_0, A, B, F, G \) and \( D \) have compatible dimensions.

Given a matrix \( M \geq 0 \), the quadratic form \( z^T M z \) may be denoted as \( |z|^2_M \). For \( 0 \leq j \leq N \), denote \( u_{-j} = (u_0, \ldots, u_{j-1}, u_{j+1}, \ldots, u_N) \). The cost for \( A_0 \) is given by

\[
J_0(u_0, u_{-0}) = E \int_0^\infty e^{-\rho t} \{ |x_0 - \Phi(x^{(N)})|^2_Q + u_0^T R_0 u_0 \} dt,
\]

where \( \Phi(x^{(N)}) = H x_0 + \bar{H} x^{(N)} + \eta \). The constant matrices or vectors \( H, \bar{H}, Q_0 \geq 0, Q \geq 0, R_0 > 0, R > 0, \eta, \eta_0 \) and \( \eta_1 \) have compatible dimensions. The cost (4) contains the term \( H x_0 \) to capture the strong influence of the major player.

(A1) The initial states \( x_j(0), 0 \leq j \leq N \), are independent, \( \bar{E} x_j(0) = 0 \) for each \( i \geq 1 \), and there exists \( c_0 < \infty \) independent of \( N \) such that \( \sup_{j \geq 0} E \{ x_j(0) \} \leq c_0 \).</p>

For simplicity, in (A1) it is assumed that all minor players have zero initial mean. It is possible to generalize our analysis to different initial means.

Let \( k \geq 1 \) be an integer. Define the function class \( C_{\rho/2}([0, \infty), \mathbb{R}^k) = \{ f | f \in C([0, \infty), \mathbb{R}^k), \sup_{t \geq 0} \{ f(t) e^{-\rho t/2} \} < \infty \) for some \( \rho' \in [0, \rho) \} \).

Notice that \( \rho' \) may vary with each \( f \) within the above set.

2. THE GAME PROBLEM

Within the Nash game formulation, intuition might suggest that in the large population limit, the set of equilibrium strategies takes a form such that at time \( t \), the major player and each minor player only need to know \( x_0(t) \) and \( x_0(t), x_1(t) \), respectively. But, as a surprise, those variables do not form a set of sufficient statistics for making the decision. Below we first briefly review the state space augmentation based approach developed in [5], which Markovianizes the limiting decision problem.

2.1 The limiting two-player model

We approximate \( x^{(N)} \) by the following process

\[
\dot{z}(t) = \bar{A} z(t) dt + \bar{C} \bar{e}_0(t) dt + \bar{m}(t) dt,
\]

where \( \bar{A} \in \mathbb{R}^{n \times n}, \bar{C} \in \mathbb{R}^{n \times n} \) are constant matrices, and \( \bar{m}(t) \in C([0, \infty), \mathbb{R}^n) \); see [5] for the motivation of (6). We take \( \bar{z}(0) = 0 \) due to the zero initial mean assumption in (A1). The equation of \( x_0(t) \) is given below.

After replacing \( x^{(N)} \) in (1)-(2) by \( \bar{z} \), the dynamics of the limiting two-player game are given by

\[
\dot{x}_0 = (A_0 x_0 + B_0 u_0 + F_0 \bar{z}) dt + D_0 dW_0,
\]

\[
\dot{x}_1 = (A x_1 + B u_1 + F \bar{z} + G \bar{e}_0) dt + D dW_1,
\]

where \( \bar{x}_0(0) = x_0(0) \) and \( \bar{x}_1(0) = x_1(0) \). To distinguish from the original model of \( N + 1 \) players, we use the new state variables \( \bar{x}_0 \) and \( \bar{x}_1 \). But we still use \( u_0, u_1, W_0 \) and \( W_1 \) in this population limit model and this reuse of notation should cause no risk of confusion. Let \( A_0 \) and \( \bar{A} \) stand for the two players described by (6)-(8), still called the major player and the minor player, respectively. The costs for \( A_0 \) and \( \bar{A} \), respectively, are given by

\[
J_0(u_0) = E \int_0^\infty e^{-\rho t} \{ |\bar{x}_0 - \Phi(\bar{x}^{(N)})|^2_{Q_0} + u_0^T R_0 u_0 \} dt,
\]

\[
J_1(u_1, u_0) = E \int_0^\infty e^{-\rho t} \{ |\bar{x}_1 - \Phi(\bar{x}^{(N)})|^2_{Q_1} + u_1^T R_1 u_1 \} dt.
\]
\[ \begin{align*}
\dot{\bar{x}}_i &= A \bar{x}_i + G \bar{x}_0 + F \bar{u}_i + D \bar{d}W_i, & i = 1, \ldots, N, \\
\dot{\bar{x}}_0 &= A \bar{x}_0 + G \bar{x}_i + F \bar{u}_i + D \bar{d}W_i, & i = 1, \ldots, N.
\end{align*} \]

where \( \bar{x}_i(0) = x_i(0), \bar{x}_0(0) = x_0(0), \bar{z}(0) = 0. \) Denote

\[ \bar{A} = \begin{bmatrix} A & G \\ 0 & A_0 - B_0 R^{-1} B_0^T P_0 \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} B \\ 0 \end{bmatrix}, \quad \bar{M} = \begin{bmatrix} 0_{n \times 1} \\ M_0 - B_0 R^{-1} B_0 \end{bmatrix}. \]

The optimal control problem is given by

\[ \begin{align*}
\min_{\bar{u}_0} & \quad J_{\text{soc}}(\bar{u}_0, \hat{u}_1, \ldots, \hat{u}_N), \\
\text{s.t.} & \quad \bar{A} \bar{x}_0 = A_0 \bar{x}_0 + G \bar{x}_0 + F \bar{u}_0 + D \bar{d}W_0,
\end{align*} \]

where \( \bar{u}_0 \) is an \( F \times (0), W \)-adapted control.

For decentralized social optimization, our approach consists of the following steps:

- Following Section 2, the system is augmented by introducing the stochastic mean field dynamics.
- The limiting decision problems for the major player and a representative minor player are constructed, and the optimal control laws are obtained.
- The consistency condition is imposed for the stochastic ODE of the mean field.

Compared with Section 2, the construction of the limiting control problem is different since now in the population limit each player cannot simply minimize its cost by optimally responding to the mean field (and the major player if it is a minor player).

To simplify the analysis, we introduce the assumption:

(A2) \( F = F^0 = G = 0, Q_0 \geq 0, Q > 0. \)

Let \( J_{\text{soc}}^{(N)} \) be minimized by \( \bar{u} = (\bar{u}_0, \bar{u}_1, \ldots, \bar{u}_N) \). We regard the individual controls \( u_j(t) \) (or \( \bar{u}_j(t) \)) as functionals of the initial states \( (x_0(0), \ldots, x_N(0)) \) and the Brownian motions \( (W_0(s), \ldots, W_N(s)), s \leq t, \) which will be called \( F^x(t,0), W \)-adapted controls. Note that if the controls are given in a linear feedback form, they may be realized as \( F^x(t,0), W \)-adapted controls by running the closed-loop system of the \( N + 1 \) agents. For the analysis below, when player \( i \) is picked out to perturb its control, we suppose that the controls of all other players are set as \( \bar{u}_j \). Let \( \bar{x}_j \) be associated with \( \bar{u}_j \). Denote \( x_{-i}^{(N)} = 1/N \sum_{j \neq i} x_j \) and \( \bar{x}_{-i} \) as the unique optimal control for the optimal control problem

(PP0) \( \begin{align*}
\min_{\bar{u}_0} & \quad J_{\text{soc}}^{(N)}(\bar{u}_0, \bar{u}_1, \ldots, \bar{u}_N), \\
\text{s.t.} & \quad \bar{A} \bar{x}_0 = A_0 \bar{x}_0 + G \bar{x}_0 + F \bar{u}_0 + D \bar{d}W_0,
\end{align*} \]

where \( \bar{u}_0 \) is an \( F \times (0), W \)-adapted control.

**Proposition 1.** Suppose that \( \bar{u} \) minimizes \( J_{\text{soc}}^{(N)} \). Then

- i) \( \bar{u}_0 \) is the unique optimal control for the optimal control problem

\[ \begin{align*}
\min_{\bar{u}_0} & \quad J_{\text{soc}}^{(N)}(\bar{u}_0, \bar{u}_1, \ldots, \bar{u}_N), \\
\text{s.t.} & \quad \bar{A} \bar{x}_0 = A_0 \bar{x}_0 + G \bar{x}_0 + F \bar{u}_0 + D \bar{d}W_0,
\end{align*} \]

where \( \bar{u}_0 \) is an \( F \times (0), W \)-adapted control;
ii) $\tilde{u}_i$ is the unique optimal control for the optimal control problem

$$(Pn0)\quad dx_i = Ax_i dt + Bu_idt + Ddw_i,$$

minimize $u_i \quad J^{(N)}_{soc}(\ldots, u_{i-1}, u_i, u_{i+1}, \ldots)$,

where $u_i$ is an $F_i(t)$-adapted control. $\square$

3.1 The major player

We consider the control perturbation of the major player while the controls of the minor players are $(\tilde{u}_1, \ldots, \tilde{u}_N)$. Lemma 2. For finding $\tilde{u}_0$, Problem (PM0) is equivalent to (PM1)

$$dx_0 = A_0x_0 dt + B_0u_0dt + D_0dw_0,$$

minimize $u_0 \quad \bar{J}_0 = \int_0^\infty e^{-\rho t} \tilde{L}_0(x_0, \tilde{x}^{(N)}, u_0) dt$,

where

$$\tilde{L}_0 = |x_0 - H_0\tilde{x}^{(N)}|_Q^2 - 2\eta_0 Q_0[x_0 - H_0\tilde{x}^{(N)}] + u_0^T R_0 u_0$$

$$+ \alpha \{ x_0^T H^T Q H x_0 - 2x_0^T H^T Q(\tilde{H} - \tilde{Z}^{(N)}) - \eta \}.$$

Proof: For $i \geq 1$, denote $J_i = E \int_0^\infty e^{-\rho t} (Z_i + \tilde{Z}_i) dt$, where

$$Z_i = x_i^T H^T Q H x_0 - 2x_0^T H^T Q(\tilde{x}_i - \tilde{H} \tilde{Z}^{(N)} - \eta),$$

$$\tilde{Z}_i = |\tilde{x}_i - \tilde{H} \tilde{Z}^{(N)} - \eta|_Q^2 + \tilde{u}_i^T R_0 \tilde{u}_i.$$

We have

$$J^{(N)}_{soc}(u_0, \tilde{u}_1, \ldots, u_N) = E \int_0^\infty e^{-\rho t} [x_0 - H_0\tilde{x}^{(N)}]_Q^2$$

$$- 2\eta_0 Q_0[x_0 - H_0\tilde{x}^{(N)}] + u_0^T R_0 u_0 + (\alpha/N) \int_1^\infty e^{-\rho t} \sum_{i=1}^N Z_i dt + (\alpha/N) \int_0^\infty e^{-\rho t} \sum_{i=1}^N \tilde{Z}_i dt.$$

The lemma follows since $Z_i'$ does not change with $u_0$. $\square$

3.2 The minor player

We pick out player $i, 1 \leq i \leq N$, and perturb its control. We write $J_0$ in the form $J_0 = E \int_0^\infty e^{-\rho t} (Z_0 + \tilde{Z}_0) dt$, where

$$Z_0 = (1/N^2) x_i^T H_i^T Q_i H_i x_i$$

$$- (2/N)[\tilde{x}_0 - H_0 \tilde{x}^{(N)} - \eta]_Q T Q_0 H_i x_i,$$

$$\tilde{Z}_0 = |\tilde{x}_0 - H_0 \tilde{x}^{(N)} - \eta|_Q^2 + \tilde{u}_0^T R_0 \tilde{u}_0.$$

For player $i$, denote $J_i = E \int_0^\infty e^{-\rho t} (Z_i + \tilde{Z}_i) dt$, where

$$Z_i = x_i^T (I - \tilde{H} \tilde{N}) Q (I - \tilde{H} \tilde{N}) x_i + u_i^T R_i$$

$$- 2[H_0 \tilde{x}_0 + H \tilde{x}^{(N)} + \eta] T Q (I - \tilde{H} \tilde{N}) x_i,$$

$$\tilde{Z}_i = |H_0 \tilde{x}_0 + H \tilde{x}^{(N)} + \eta|_Q^2.$$

For player $0 < k \neq i$, $J_k = E \int_0^\infty e^{-\rho t} (Z_k + \tilde{Z}_k) dt$, where

$$Z_k = (1/N^2) x_k^T H_k^T Q H_k x_k$$

$$- (2/N)[\tilde{x}_k - H \tilde{x}^{(N)} - \eta] T Q (I - \tilde{H} \tilde{N}) x_k,$$

$$\tilde{Z}_k = |\tilde{x}_k - H \tilde{x}^{(N)} - \eta|_Q^2 + \tilde{u}_k^T R_0 \tilde{u}_k.$$

Summarizing the above calculation, we obtain Lemma 3. For finding $\tilde{u}_i$, Problem (PM0) is equivalent to (PM1)

$$dx_i = Ax_i dt + Bu_i dt + Ddw_i,$$

minimize $u_i \quad \bar{J}_i = E \int_0^\infty e^{-\rho t} (Z_0 + (\alpha/N) \sum_{k=1}^N Z_k) dt$. $\square$

4. THE LIMITING OPTIMAL CONTROL PROBLEMS FOR SOCIAL OPTIMIZATION

4.1 The limiting control problem of the major player

Following Section 2, we approximate $\tilde{Z}^{(N)}$ and $\tilde{z}^{(N)}$ in Problems (PM1) and (PM1) by $\tilde{z}$ as follows

$$d\tilde{z}(t) = \tilde{z}(t) dt + \tilde{G}\tilde{x}(t) dt + \tilde{m}(t) dt,$$

where $\tilde{A} \in \mathbb{R}^{nxn}$, $\tilde{G} \in \mathbb{R}^{nxm}$ are constant matrices, and $\tilde{m} \in C_{p/2}(0, \infty, \mathbb{R}^m)$. We still use $(\tilde{A}, \tilde{G}, \tilde{m})$ to describe $\tilde{z}$ although the triple will differ from that in Section 2. To approximate (PM1), we construct the auxiliary optimal control problem with dynamics

$$d\tilde{x}_0(x_0, \tilde{x}, u_0) dt$$

$$+ \left[ \begin{array}{c} \tilde{Q}_0 \tilde{z} \\ \tilde{z} \end{array} \right] dt + \left[ \begin{array}{c} B_0 \\ 0_{n \times 1} \end{array} \right] u_0 dt$$

where

$$\tilde{Q}_0 = [I, -H_0] T Q[I, -H_0] + \alpha [H, \tilde{H}, \tilde{I}] T Q[H, \tilde{H}, \tilde{I}].$$

By noting that in (30)-(31), $I = I_{nxn}$ and $0 = 0_{nxn}$. We introduce the following assumption.

(A3) The matrix $\tilde{Q}_0$ is positive definite. $\Diamond$

Example: If $n = 1$ and we take $H_0 = h, H = 1/2, \tilde{H} = 1/2$ and $Q_0 = Q = 1$, then

$$\tilde{Q}_0 = \begin{bmatrix} 1 + \frac{\alpha}{4} & -h - \frac{\alpha}{4} \\ -h - \frac{\alpha}{4} & h^2 \end{bmatrix}$$

is positive definite if $4h^2\alpha - 8h\alpha - \alpha^2 > 0$. $\Diamond$

Define

$$\tilde{A}_0 = \left[ \begin{array}{c} \tilde{A}_0 \\ 0_{nx1} \end{array} \right], \quad \tilde{B}_0 = \left[ \begin{array}{c} B_0 \\ 0_{nx1} \end{array} \right], \quad \tilde{M}_0 = \left[ \begin{array}{c} 0_{nx1} \\ \end{array} \right].$$

We introduce the ARE and ODE

$$\tilde{P}_0 = \tilde{P}_0 \tilde{A}_0 + \tilde{A}_0^T \tilde{P}_0 - \tilde{P}_0 \tilde{B}_0 \tilde{R}_0^{-1} \tilde{B}_0^T \tilde{P}_0 + \tilde{Q}_0,$$

$$\tilde{P}_0 = \tilde{P}_0 \tilde{B}_0 \tilde{R}_0^{-1} \tilde{B}_0^T \tilde{P}_0 + \tilde{M}_0,$$

where $\tilde{s}_0$ is to be sought within the set $C_{p/2}(0, \infty, \mathbb{R}^n)$. The optimal control law for $\tilde{A}_0$ is given as

$$\tilde{u}_0 = -R_0^{-1} \tilde{B}_0^T \tilde{P}_0 \left[ \begin{array}{c} \tilde{x}_0 \\ \tilde{z} \end{array} \right] + \tilde{s}_0.$$
Remark: $A_0$ and $M_0$ (and also $\hat{\Lambda}, \hat{M}$ below) are generally different from these matrices in Section 2. The reuse of notation should cause no risk of confusion. \hfill \Diamond

4.2 The limiting control problem of the minor player

As an approximation to Problem (Pm1) for player $i$, we construct the optimal control problem with dynamics

$$
d \begin{bmatrix} \dot{\bar{x}}_i \\ \bar{z} \end{bmatrix} = A \begin{bmatrix} 0 & 0 \\ 0 & A_0 - B_0 R_0^{-1} B_0^T P_0 \end{bmatrix} \begin{bmatrix} \bar{x}_i \\ \bar{z} \end{bmatrix} dt + \begin{bmatrix} B \\ 0_{2n \times n} \end{bmatrix} u_idt + \begin{bmatrix} 0_{n \times 1} \\ M_0 - B_0 R_0^{-1} B_0^T \bar{s}_0 \end{bmatrix} dt + \begin{bmatrix} DdW_i \\ DdW_0 \end{bmatrix},
$$

where $\bar{x}_i(0) = x_i(0)$, $\bar{z}(0) = 0$. Note that the SDE (35) is obtained by combining (34) with (28). The cost is given by

$$
J_i^* = \int_0^\infty e^{-\rho t} \bar{L}_i^*(\bar{x}_i, \bar{z}, \bar{u}_i) dt,
$$

where

$$
\bar{L}_i^* = \begin{bmatrix} \bar{x}_i \\ \bar{z} \end{bmatrix}^T \bar{Q} \begin{bmatrix} \bar{x}_i \\ \bar{z} \end{bmatrix} + 2 \begin{bmatrix} \bar{x}_i \\ \bar{z} \end{bmatrix}^T \bar{\eta}^* + \alpha \bar{u}_i^T R \bar{u}_i,
$$

and

$$
\bar{Q} = \begin{bmatrix} H_0, 0, 0 \end{bmatrix}^T \begin{bmatrix} H_0, 0, 0 \end{bmatrix} - \begin{bmatrix} I, -H_0 \end{bmatrix}^T \begin{bmatrix} Q_0, 0, 0 \end{bmatrix} - \begin{bmatrix} H_0, 0, 0 \end{bmatrix}^T \begin{bmatrix} I, 0, 0 \end{bmatrix} + \alpha \begin{bmatrix} 1, 0, 0 \end{bmatrix}^T \begin{bmatrix} 0, 0, 0 \end{bmatrix}.
$$

The following proposition means that we may compute a well-defined optimal closed-loop (not depending on the choice of $c$) when $\hat{Q}^* \geq 0$ does not hold.

Proposition 4. Let $c$ be a positive constant such that $\hat{Q}^{c,*} > 0$ and $\hat{P}^c$ the solution to (43). Then the solution $\bar{x}_i$ to the closed-loop equation

$$
d\bar{x}_i = A \bar{x}_i dt - BR^{-1} B^T \hat{P}^c (\bar{x}_i, \bar{z}_i, \bar{z}_i^T) dt - BR^{-1} B^T \bar{z}_i dt + DdW_i,
$$

does not depend on the choice of $c$. \hfill \Box

5. THE CONSISTENT SOLUTION FOR SOCIAL OPTIMIZATION

For the matrices $\hat{P}$ and $\bar{P}_0$, we introduce the partition

$$
\hat{P} = \begin{bmatrix} \hat{P}_{11} & \hat{P}_{12} & \hat{P}_{13} \\ \hat{P}_{21} & \hat{P}_{22} & \hat{P}_{23} \\ \hat{P}_{31} & \hat{P}_{32} & \hat{P}_{33} \end{bmatrix},
\bar{P}_0 = \begin{bmatrix} \bar{P}_{0,11} & \bar{P}_{0,12} \\ \bar{P}_{0,21} & \bar{P}_{0,22} \end{bmatrix},
$$

where each submatrix $\hat{P}_{ij}$ or $\bar{P}_{0,ij} \in \mathbb{R}^{n \times n}$.

When $N \rightarrow \infty$, we obtain from (42) the equation for $\bar{z}$

$$
d\bar{z} = (A - BR^{-1} B^T \hat{P}_{11} - BR^{-1} B^T \hat{P}_{13}) \bar{z} dt - BR^{-1} B^T \bar{P}_{12} \bar{x}_i dt - BR^{-1} B^T \bar{z} dt,
$$

where $\bar{z}(0) = 0$. Now the consistency condition on (28) and (45) translates into the the constraints:

$$
\bar{A} = A - BR^{-1} B^T \hat{P}_{11} - BR^{-1} B^T \hat{P}_{13},
\bar{G} = -BR^{-1} B^T \hat{P}_{12},
\bar{m} = -BR^{-1} B^T \hat{s},
$$

Definition 5. The set of constant matrices $(\hat{P}, \bar{A}, \bar{G}, \bar{P})$ is said to be a consistent solution to the equation system of (32), (39), (46) and (47), if $\hat{P}_0 \geq 0$, $\hat{P}$ is symmetric and $\hat{A} - BR^{-1} B^T \hat{P} = (\rho/2) I$ is Hurwitz. \hfill \Diamond

We do not require $\hat{P} \geq 0$. We will refer to (32), (39), (46) and (47) as the constrained algebraic Riccati equations (cAREs).

Definition 6. Suppose that $(\hat{P}_0, \bar{A}, \bar{G}, \bar{P})$ is a consistent solution to the cAREs. The triple $(\bar{s}_0, \bar{s}, \bar{m})$ is said to be a consistent solution to the equation system (33), (40) and (48), if the three vector functions are all in $C_{\rho/2}([0, \infty), \mathbb{R}^n)$, with $k = 2n, 3n, n$, respectively, and they satisfy the three equations. \hfill \Diamond

In fact the definition of a consistent solution to the cAREs implicitly requires that the optimal control problem of the major player has a stable closed-loop after the exponential discount of $e^{-\rho/2t}$.

Proposition 7. If $(\hat{P}_0, \bar{A}, \bar{G}, \bar{P})$ is a consistent solution to the cAREs, then $A_0 - BR^{-1} B^T \hat{P}_0 - (\rho/2) I$ is Hurwitz.
Proof: Following the argument in [5, Proposition 2], we may write $\bar{A} - B\bar{R}^{-1}B^T \bar{P}$ in an upper block-triangular form. The proposition follows. □

Next we check whether the limiting optimal control problem of the minor player is well-defined if $(\bar{A}, \bar{C})$ ensures that $A_0 - B_0 R_0^{-1} B_0^T P_0 - (\rho/2)I$ is Hurwitz. Here $(\bar{A}, \bar{C})$ is not necessarily specified via a consistent solution.

Proposition 8. Suppose (i) $(\bar{A}, \bar{C})$ is such that $A_0 - B_0 R_0^{-1} B_0^T P_0 - (\rho/2)I$ is Hurwitz for the limiting control problem of the major player, (ii) $Q > 0$ and $R > 0$, (iii) $[A - (\rho/2)I, B]$ is stabilizable. Then the limiting optimal control problem for the minor player is uniquely solvable.

Proof: First, we can see that the 3n-dimensional system for the minor player is stabilizable since $(\bar{x}_0, \bar{z})$ is not affected by $u_i$ and hence may be treated as an exogenous signal. Then by stabilizability, it is easy to show that $J^*_i$ can attain a finite value and its infimum is bounded from below by a finite constant. So the limiting optimal control problem of the minor player is well-defined. Next, we fix $(\bar{x}_0, \bar{z})$; then the optimal control problem has a cost integrand convex in $(\bar{x}_i, u_i)$ and strictly convex in $u_i$. So the optimal control is unique. □

Remark: We may write the dynamic programming (DP) equation for the limiting optimal control problem of the minor player and assume a quadratic form of the value function. By using this DP equation, we can derive (39) and (40). This provides a justification for introducing these equations even if $\bar{Q}^*$ may not be positive semi-definite. ◊

Remark: After the two limiting control problems are solved, decentralized control laws may be constructed for the original social optimization problem by (i) letting $\bar{z}$ in (28) be driven by the state $\bar{x}_0$ of the major player, and (ii) setting $\bar{x}_0 = x_0$ and $\bar{x}_i = x_i$ for $\bar{u}_0$ and $\bar{u}_i$. ◊

6. CONCLUDING REMARKS

We consider stochastic control problems involving a major player and a large number of minor players. We first review the state space augmentation approach for the game problem, and next combine it with the calculation of social cost variation for decentralized control in social optimization problems. In future work, it will be of interest to obtain performance estimates and develop numerical methods for the resulting decentralized strategies in the social optimization problem.

REFERENCES


