Equilibria and optima in some constrained stochastic games with independent state processes

N. Hemachandra

Industrial Engineering and Operations Research
Indian Institute of Technology Bombay

Joint work with Vikas Vikram Singh

NMI Workshop, IISc

January, 2016

Outline

Framework

- N-player/agent constrained stochastic games with independent state processes where all agents use expected average cost criterion. [2]
  - Belongs to the class of decentralized stochastic games.
  - Each agent controls its own Markov chain, whose transition probabilities do not depend on the actions of others. However, the expected average cost of each agent depends on the strategies of all the others.
  - Altman et al. [2] have shown that these games possess a stationary Nash equilibrium under strong Slater and unichain conditions.
- Altman and Shwartz [4] introduced N-player centralized constrained stochastic games with discounted and average cost criterion.
  - Stationary Nash equilibrium exist under strong Slater and unichain conditions.
  - Some *zero-sum* games can be solved by linear programs.

Notations I

N-player constrained stochastic game $\left(S^i, \gamma^i, A^i, c^i, d^i, p^i, \xi^i\right)_{i=1}^N$ for i, $i \in I = \{1, 2, \cdots, N\}$

- S^i is a finite state space of player i. Define, S:=$\times_{j=1}^N S^j$ and S^{-i}:=$\times_{j \neq i} S^j$. The element of S is denoted by s where $s = (s^1, s^2, \cdots, s^N)$ and $s^{-i} \in S^{-i}$ denotes a vector of states $s^j, j \neq i$.
- γ^i : probability distribution for the initial state of player i. Initial states of all the players are independent. Denote $\gamma = (\gamma^1, \gamma^2, \cdots, \gamma^N)$.
- A^i is a finite action set of player i. $A^i(s^i)$ denotes the set of all actions of player i at state s^i and $A^i = \bigcup_{s^i \in S^i} A^i(s^i)$. Define, $A(s)$=$\times_{j=1}^N A^j(s^j)$ for each $s \in S$.
- Define, $\mathcal{K}^i = \{(s^i, a^i) | s^i \in S^i, a^i \in A^i(s^i)\}$. Define, $\mathcal{K}$=$\times_{j=1}^N \mathcal{K}^j$ and $\mathcal{K}^{-i}$=$\times_{j \neq i} \mathcal{K}^j$.
- $c^i(s, a) : \mathcal{K} \to \mathbb{R}$ is immediate cost of player i.

Notations II

- $d^i = \left(d^{i,1}, d^{i,2}, \cdots, d^{i,n_i}\right)$, where $d^{i,k} : \mathcal{K} \to \mathbb{R}$ for all $k = 1, 2, \cdots, n_i$ are the immediate costs of player i. These are involved in the kth constraint, $k = 1, 2, \cdots, n_i$, on the expected average cost of player i.
- $p^i : \mathcal{K}^i \to \wp(S^i)$ is the transition law of player i, where $p^i(\bar{s}^i | s^i, a^i)$ is the probability that the state of player i moves from state s^i to $\bar{s}^i$ if he chooses an action $a^i \in A^i(s^i)$.
- $\xi^i = \left(\xi^i_1, \xi^i_2, \cdots, \xi^i_{n_i}\right)$ are the bounds defining the constraints of player i.
- Define, $h^i_t = \left(s^i_0, a^i_0, s^i_1, a^i_1, \cdots, s^i_{t-1}, a^i_{t-1}, s^i_t\right)$ a history of player i, $i \in I$, at time t where $s^i_m \in S^i$ for $m = 0, 1, \cdots, t$ and $a^i_m \in A^i(s^i_m)$ for $m = 0, 1, \cdots t - 1$. Let H^i_t denotes the set of all possible histories of length t of player i.
- A decision rule f^i_t of player i at time t assigns to each h^i_t with final state s^i_t a probability measure $f^i_t(h^i_t) \in \wp(A^i(s^i_t))$. These decision rules are called as history dependent decision rules.

Notations III

- Let $f^{ih} \in F^i$ where F^i denotes the set of all history dependent strategies of player i, $i \in I$ and $f^h = \left(f^{1h}, f^{2h}, \cdots, f^{Nh}\right) \in F$ where F=$\times_{i=1}^{N} F^i$ denotes the set of history dependent multi-strategies.
- Let $f^i = \left((f^i(1))^T, (f^i(2))^T, \cdots, (f^i(|S^i|))^T\right)^T$ denote a stationary strategy, where $f^i(s^i) \in \wp(A^i(s^i))$ for all $s^i \in S^i$. Let F_{S^i} and F_S=$\times_{i=1}^{N} F_{S^i}$ denotes the set of all stationary strategies of player i and multi-strategies respectively.
- Let $\{X_t, \mathbb{A}_t\}_{t=0}^{\infty}$, be a vector stochastic process where $X_t = \left(X_t^i\right)_{i=1}^N$, $\mathbb{A}_t = \left(\mathbb{A}_t^i\right)_{i=1}^N$.
- X_t^i denotes the state of player i and $\mathbb{A}_t^i$ denotes the action chosen by player i at time t.

Game dynamics

1. At time $t = 0$ the state of player i, $i \in I$, is $s_0^i \in S^i$ as chosen according to an initial distribution γ^i and player i chooses an action $a_0^i \in A^i(s^i)$ independent from other players.
2. Player i, $i \in I$ incurs an immediate cost $c^i(s_0, a_0)$.
3. Player i, $i \in I$ also incurs another n_i costs, $\left(d^{i,k}(s_0, a_0)\right)_{k=1}^{n_i}$.
4. State of player i, $i \in I$, switches to a new state s_1^i at time $t = 1$ with probability $p^i(s_1^i | s_0^i, a_0^i)$.
5. Dynamics of the Markov chains repeat at new state $s_1 = (s_1^1, \cdots, s_1^N)$.
6. This continues forever.

The expected average costs

- These are average functionals of the strategies of all the players.
- For a given initial distribution γ and a multi-strategy f^h the expected average cost of player i, $i \in I$ is defined as

$$C_{ea}^{i}(\gamma, f^h) = \limsup_{n \to \infty} \frac{1}{n} \sum_{t=0}^{n-1} \mathbb{E}_{f^h}^{\gamma} c^i(X_t, \mathbb{A}_t). \tag{1}$$

The expected average constraints

- For a given initial distribution γ and a multi-strategy f^h the expected average costs of player i, $i \in I$ are defined as

$$D_{ea}^{i,k}(\gamma, f^h) = \limsup_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} \mathbb{E}_{f^h}^{\gamma} d^{i,k}(X_t, \mathbb{A}_t),$$

  for all $k = 1, 2, \cdots, n_i$.

- $D_{ea}^{i,k}(\cdot,\cdot)$ can capture the average consumption of resource k, $k = 1, 2, \cdots, n_i$, by player i.
- The constraints of player i, $i \in I$ are given as

$$D_{ea}^{i,k}(\gamma, f^h) \leq \xi_k^i, \quad \forall\, k = 1, 2, \cdots, n_i. \tag{2}$$

  where $\{\xi_k^i\}$ are given reals.

Constrained stochastic games I

- G^c_{ea} : Constrained stochastic game where all the players choose their strategies independently and wish to minimize their expected average costs from (1) subject to their constraints from (2).
- f^h is i-feasible if it satisfies ith player's constraints from (2) and it is called feasible if it is i-feasible for every $i \in I$.
- F^ξ : set of all feasible history dependent multi-strategies for G^c_{ea}.
- F^ξ_S : set of all stationary feasible multi-strategies for G^c_{ea}.
- A multi-strategy $f^{h*} \in F^\xi$ is called a Nash equilibrium of G^c_{ea}, if for each player $i \in I$ and for any f^{ih} such that (f^{ih}, f^{-ih*}) is i-feasible, one has that

$$C^i_{ea}(\gamma, f^{h*}) \leq C^i_{ea}(\gamma, f^{ih}, f^{-ih*}). \quad (3)$$

Constrained stochastic games II

- Since player i faces a constrained Markov decision process (CMDP) f^* would still be a Nash equilibrium if we replace f^{h*} in (3) by $f^* \in F_S^{\xi}$ and f^{ih} by f^i for all $i \in I$.
- [1] showed that optimal strategy always exists in the space of stationary strategies.
- Characterization ?

Constrained Markov Decision Processes: Formulation I

A finite state-action constrained Markov decision process is a $5-$tuple $\{S, A, P, c, d\}$ [1] where,

- S : finite state space
- A : finite set of actions. $\mathcal{K} = \{(s, a) : s \in S, a \in A(s)\}$ to be the set of state-action pairs
- $P(a)$: $\{P_{ij}(a)\}$ is the transition matrix when action a is taken
- c : $c(s, a)$ is the immediate cost at state s using action a

Constrained Markov Decision Processes: Formulation II

- $\bar{d} : \bar{d}(s,a) = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_k \end{bmatrix}$ is a vector of immediate costs, related to constraints, when at state s and using action a

- For any policy u and initial distribution β, the finite horizon cost for a horizon n is defined as

$$C^n(\beta, u) = \sum_{t=1}^{n} E^u_\beta c(S_t, A_t) \tag{4}$$

The expected average cost (with finite and infinite horizons, respectively) is defined as

$$C^n_{ea}(\beta, u) = \frac{\sum_{t=1}^{n} E^u_\beta c(S_t, A_t)}{n}, \quad C_{ea}(\beta, u) = \limsup_{n\to\infty} \sum_{t=1}^{n} E^u_\beta c(S_t, A_t)$$

Constrained Markov Decision Processes: Formulation III

- The cost functions related to the immediate costs d are defined similarly; e.g., the finite horizon cost related to d_k, $k = 1, \ldots, K$, is

$$D^{n,k}(\beta, u) = \sum_{t=1}^{n} E_\beta^u d^k(S_t, A_t)$$

- For a fixed vector $V = (V_1, \ldots, V_k)$ of real numbers, we define the constrained control problem **COP** as:

  Find a policy that minimizes $C(\beta, u)$ *subject to* $D(\beta, u) \leq V$

LPs for CMDPs

- **LP(β)**: Find the infimum C^* of $C(x) := \langle \rho, c \rangle$ subject to:

$$D^k(x) := \langle x, d^k \rangle \leq V_k, k = 1, \ldots, K, \quad x \in Q_{ea}(\beta),$$

where $Q_{ea}(\beta)$ is the set of vectors $x \in \mathbb{R}^{|\mathcal{K}|}$ satisfying

$$\sum_{y \in S} \sum_{a \in A(s)} x(y,a)(\delta_s(y) - P_{ys}(a)) = 0 \;\; s \in S$$

$$\sum_{y \in S} \sum_{a \in A(s)} x(y,a) = 1 \quad x(y,a) \geq 0 \;\; \forall y, a$$

- $f(x)$ is defined to be any stationary policy such that

$$f_{x(y,a)}(y) = x(y,a) \left(\sum_{a \in A(y)} x(y,a) \right)^{-1}$$

whenever the denominator is non-zero.

Theorem

*Equivalence between **COP** and **LP**(β)*

i $C^* = C_{ea}(\beta)$.

ii *For any* $u' \in U$, *there exists a dominating stationary policy* $u \in U_S$ *such that* $x(u) := g_{ea}(\beta, u) \in Q_{ea}(\beta), C_{ea}(\beta, u) = C(x(u))$ *and* $D_{ea}(\beta, u) = D(x(u))$; *conversely, for any* $x \in Q_{ea}(\beta)$, *the stationary policy* $f(x)$ *satisfies* $C_{ea}(\beta, f(x)) = C(x)$ *and* $D_{ea}(\beta, f(x)) = D(x)$.

iii ***LP***(β) *is feasible if and only if* ***COP*** *is feasible. Assume that* ***COP*** *is feasible. Then there exists an optimal solution* x^* *for* ***LP***(β), *and the stationary policy* $f(x^*)$ *is optimal for* ***COP***.

Assumptions (back to stochastic games)

As in [2] we also assume:

(A1) Ergodicity: For each agent i, $i \in I$, and for any stationary strategy f^i the state process of the agent i is an irreducible Markov chain with one ergodic class (and possibly some transient states).

(A2) Strong Slater condition: Every player i, $i \in I$ has some strategy g^i such that for any multi-strategy f^{-i} of other players,

$$D_{ea}^{i,k}(\gamma, (f^{-i}, g^i)) < \xi_k^i, \ \ \forall\, k = 1, 2, \cdots, n_i.$$

(A3) Information: The agents do not observe their costs, i.e., the strategy chosen by any agent does not depend on the realization of the cost.

Under (A1)-(A3) Altman et al. [2] show the existence of a stationary Nash equilibrium for the game G_{ea}^c.

Definition

- For an initial distribution γ^i and a stationary strategy f^i of player i, $i \in I$, define the occupation measure as

$$\pi^i_{ea}(\gamma^i, f^i) := \left\{\pi^i_{ea}(\gamma^i, f^i; s^i, a^i) : s^i \in S^i, a^i \in A^i(s^i)\right\}.$$

- For all $s^i \in S^i$, $a^i \in A^i(s^i)$, we have

$$\pi^i_{ea}(\gamma^i, f^i; s^i, a^i) = \pi^{f^i}(s^i) f^i(s^i, a^i), \tag{5}$$

  where $\pi^{f^i} = \left(\pi^{f^i}(1), \pi^{f^i}(2), \cdots, \pi^{f^i}(|S^i|)\right)$ is the unique steady state distribution of Markov chain induced by strategy f^i of player i, which exists under (A1).

- Occupation measure in (5) is unique and independent from initial distribution γ^i.

Expected average costs in terms of occupation measure I

- For any multi-strategy $f \in F_S$

$$C^i_{ea}(f) = \sum_{(s,a)\in\mathcal{K}} \left[\prod_{j=1}^{N} \pi^j_{ea}(f^j; s^j, a^j)\right] c^i(s,a).$$

$$D^{i,k}_{ea}(f) = \sum_{(s,a)\in\mathcal{K}} \left[\prod_{j=1}^{N} \pi^j_{ea}(f^j; s^j, a^j)\right] d^{i,k}(s,a),$$

for all $k = 1, 2, \cdots, n_i$.

Expected average costs in terms of occupation measure II

- Let Q^i_{ea}, $i \in I$, be the set of vectors $x^i \in \mathbb{R}^{|\mathcal{K}^i|}$ satisfying

$$\sum_{(s^i,a^i)\in\mathcal{K}^i} \left(\delta(s^i,\bar{s}^i) - p^i(\bar{s}^i|s^i,a^i)\right) x^i(s^i,a^i) = 0, \ \forall\, \bar{s}^i \in S^i$$

$$\sum_{(s^i,a^i)\in\mathcal{K}^i} x^i(s^i,a^i) = 1$$

$$x^i(s^i,a^i) \geq 0, \ \ \forall\, s^i \in S^i, a^i \in A^i(s^i).$$

  $\delta(\cdot,\cdot)$ is a Kronecker delta.

- Completeness : For each player i, $i \in I$, the set of occupation measures achieved by history dependent strategies equals to those achieved by stationary strategies and further equals to the set Q^i_{ea} [1].

Expected average costs in terms of occupation measure III

- For any $x^i \in Q^i_{ea}$, we have for each $(s^i, a^i) \in \mathcal{K}^i$, $x^i(s^i, a^i) = \pi^i_{ea}(f^i; s^i, a^i)$ where strategy f^i is such that

$$f^i(s^i, a^i) = \frac{x^i(s^i, a^i)}{\sum_{b^i \in A^i(s^i)} x^i(s^i, b^i)}, \ \forall\, s^i \in S^i, a^i \in A^i(s^i), \tag{6}$$

- The immediate costs of player i, $i \in I$, when he uses action a^i atstate s^i and other players use f^{-i} are defined as in [2],

$$c^i(f^{-i}; s^i, a^i) = \sum_{(s,a)^{-i} \in \mathcal{K}^{-i}} \Big[\prod_{j=1; j\neq i}^{N} \pi^j_{ea}(f^j; s^j, a^j) \Big] c^i(s, a).$$

$$d^{i,k}(f^{-i}; s^i, a^i) = \sum_{(s,a)^{-i} \in \mathcal{K}^{-i}} \Big[\prod_{j=1; j\neq i}^{N} \pi^j_{ea}(f^j; s^j, a^j) \Big] d^{i,k}(s, a),$$

for all $k = 1, 2, \cdots, n_i$.

Best response linear programs

- The best response strategy of each player can be obtained by solving a CMDP.
- By using these best response linear programs for each player, we obtain an Optimization Problem which characterizes the stationary Nash equilibria of G^c_{ea} via its global minimizers.
- The best response of player i, $i \in I$, against the fixed stationary strategies f^{-i} of other players:

$$\min_{x^i} \sum_{(s^i,a^i)\in\mathcal{K}^i} c^i(f^{-i}; s^i, a^i) x^i(s^i, a^i)$$

$$\text{s.t.}$$

$$(i) \sum_{(s^i,a^i)\in\mathcal{K}^i} d^{i,k}(f^{-i}; s^i, a^i) x^i(s^i, a^i) \leq \xi^i_k, \ \forall\, k = 1, 2, \cdots, n_i$$

$$(ii)\ x^i \in Q^i_{ea}. \tag{7}$$

Dual linear programs

- If x^{i*} is the optimal solution of the linear program (7), then, by using x^{i*} the best response f^{i*} of player i can be obtained from (6).
- The dual of (7) is

$$\max_{v^i, u^i, \lambda^i} \left[v^i - \sum_{k=1}^{n_i} \lambda_k^i \xi_k^i \right]$$

s.t.

$$(i)\ v^i + u^i(s^i) \leq c^i(f^{-i}; s^i, a^i) + \sum_{k=1}^{n_i} d^{i,k}(f^{-i}; s^i, a^i)\lambda_k^i + \sum_{\bar{s}^i \in S^i} p^i(\bar{s}^i | s^i, a^i) u^i(\bar{s}^i),$$
$$\forall\ s^i \in S^i, a^i \in A^i(s^i)$$
$$(ii)\ \lambda_k^i \geq 0, \ \ \forall\ k = 1, 2, \cdots, n_i. \tag{8}$$

Optimization Problem

Let $\zeta^T := \left(v^i, (u^i)^T, (x^i)^T, (\lambda^i)^T\right)_{i=1}^N$ and $\psi(\zeta)$ denote the decision variables and the objective function of [OP] respectively. By using N primal-dual linear programs given by (7) and (8), we have the following result.

Theorem

(a) If $(f^{i*})_{i=1}^N$ is a stationary Nash equilibrium of G_{ea}^c, then, there exists a vector $\zeta^{*T} = \left(v^{i*}, (u^{i*})^T, (x^{i*})^T, (\lambda^{i*})^T\right)_{i=1}^N$ such that it is a global minimum of an Optimization Problem [OP] *given below*

$$\text{[OP]} \quad \min_{\zeta} \sum_{i=1}^{N} \left[\sum_{(s,a)\in\mathcal{K}} \left(\prod_{j=1}^{N} x^j(s^j, a^j) \right) c^i(s,a) - \left(v^i - \sum_{k=1}^{n_i} \lambda_k^i \xi_k^i \right) \right]$$

Optimization Problem

s.t.

$$(i)\ v^i + u^i(s^i) \le \sum_{(s,a)^{-i}\in\mathcal{K}^{-i}} \left(\prod_{j=1;j\neq i}^{N} x^j(s^j,a^j)\right) c^i(s,a) + \sum_{k=1}^{n_i} \lambda_k^i \left[\sum_{(s,a)^{-i}\in\mathcal{K}^{-i}} \left(\prod_{j=1;j\neq i}^{N} x^j(s^j,a^j)\right) d^{i,k}(s,a)\right] + \sum_{\bar{s}^i\in S^i} p^i(\bar{s}^i|s^i,a^i)u^i(\bar{s}^i), \forall\ s^i\in S^i, a^i\in A^i(s^i), i\in I$$

$$(ii)\ \sum_{(s,a)\in\mathcal{K}} \left(\prod_{j=1}^{N} x^j(s^j,a^j)\right) d^{i,k}(s,a) \le \xi_k^i,\ \forall\ k=1,2,\cdots,n_i, i\in I$$

$$(iii)\ x^i\in Q_{ea}^i,\ \forall\ i\in I$$

$$(iv)\ \lambda_k^i \ge 0,\ \ \forall\ k=1,2,\cdots,n_i,\ i\in I\ \text{ with}\psi(\zeta^*)=0.$$

Optimization Problem

(b) If $\zeta^{*T} = \left(v^{i*}, (u^{i*})^T, (x^{i*})^T, (\lambda^{i*})^T\right)_{i=1}^N$ is a global minimum of [OP] with $\psi(\zeta^*) = 0$, then, $(f^{i*})_{i=1}^N$ is a stationary Nash equilibrium of G_{ea}^c where, strategy f^{i*}, $i \in I$, is such that

$$f^{i*}(s^i, a^i) = \frac{x^{i*}(s^i, a^i)}{\sum_{b^i \in A^i(s^i)} x^{i*}(s^i, b^i)},$$

for all $s^i \in S^i$, $a^i \in A^i(s^i)$ whenever the denominator is non-zero (when it is zero $f^{i*}(s^i)$ is chosen arbitrarily from $\wp(A^i(s^i))$).

Optimization Problem

Remark

The objective function of [OP] is non-convex because the diagonal elements of its Hessian matrix are zero. So it will have some positive as well as some negative eigenvalues.
As there are some non-convex constraints, the feasible region is also not a convex set. So, [OP] is a non-convex constrained optimization problem.

Additive Cost Decoupled Constraints Cost (AC-DCC) games

- We consider the situation when
  - The immediate costs of each agent which correspond to its expected average costs as defined in (1) are additive over players and
  - the immediate costs of each player which correspond to its expected average constraints as defined in (2) are decoupled, i.e., these costs do not depend on the states and actions of other players.
- This class is characterized by the following additional assumptions:

  (a) $c^i(s,a) = \sum_{j=1}^{N} c^i_j(s^j, a^j), \ \forall\, s \in S, \ a \in A(s), \ i \in I.$

  (b) $d^{i,k}(s,a) = d^{i,k}(s^i, a^i), \ \forall\, s \in S, a \in A(s), k = 1, 2, \cdots, n_i, \ i \in I.$
- The condition (b) corresponds to the situation when each player has its own resources that are not shared by other players.

Additive Cost Decoupled Constraints Cost (AC-DCC) games I

- Under the assumptions (a) and (b) the Optimization Problem [OP] reduces to a LP as given below:

$$[\text{LP}] \quad \min_{\zeta} \sum_{i=1}^{N} \left[\sum_{j=1}^{N} \sum_{(s^j,a^j)\in\mathcal{K}^j} c_j^i(s^j,a^j)x^j(s^j,a^j) - \left(v^i - \sum_{k=1}^{n_i} \lambda_k^i \xi_k^i \right) \right]$$

s.t.

$$(i)\ v^i + u^i(s^i) \le c_i^i(s^i,a^i) + \sum_{j=1;j\neq i}^{N} \sum_{(s^j,a^j)\in\mathcal{K}^j} c_j^i(s^j,a^j)x^j(s^j,a^j) + \sum_{k=1}^{n_i} d^{i,k}(s^i,a^i)\lambda_k^i$$

$$+ \sum_{\bar{s}^i\in S^i} p^i(\bar{s}^i|s^i,a^i)u^i(\bar{s}^i),\ \forall\ s^i \in S^i, a^i \in A^i(s^i), i \in I$$

$$(ii)\ \sum_{(s^i,a^i)\in\mathcal{K}^i} d^{i,k}(s^i,a^i)x^i(s^i,a^i) \le \xi_k^i,\ \forall\ k = 1,2,\cdots,n_i, i \in I$$

Additive Cost Decoupled Constraints Cost (AC-DCC) games II

$$(iii)\ x^i \in Q^i_{ea},\ \forall\, i \in I$$
$$(iv)\ \lambda^i_k \geq 0,\ \ \forall\, k = 1, 2, \cdots, n_i,\ \ i \in I.$$

Corollary

Let the conditions (a) and (b) hold for a constrained stochastic game G^c_{ea}. Then a stationary Nash equilibrium of G^c_{ea} can be obtained from an optimal solution of a linear program [LP].

Zero sum constrained stochastic game [3]

- Special case of constrained stochastic game G^c_{ea} where we consider two player zero sum games with decoupled constraints [3].
- Setting $c^1(s^1, s^2, a^1, a^2) = -c^2(s^1, s^2, a^1, a^2) = c(s^1, s^2, a^1, a^2)$ for all $s^1 \in S^1$, $s^2 \in S^2$, $a^1 \in A^1(s^1)$, $a^2 \in A^2(s^2)$ and under assumption (b) from, the OP results in a primal-dual linear programs which are same as given in [3].

Contributions

- Stationary Nash equilibria of N-player constrained stochastic games, can be obtained from the global minimizers of a certain non-convex optimization problem.
- Converse is also true , i.e., from a stationary Nash equilibrium of a given such game we can construct a point which is a global minimum of the corresponding optimization problem.
- Identified a subclass of these N-player games, called Additive Cost - Decoupled Constraints Cost games (AC-DCC games).
  - The OP reduces into a LP, and hence, the Nash equilibrium of these games can be obtained from an optimal solution of a LP.
  - For two player zero sum game, our OP results in a primal-dual LPs.

Related and ongoing work

- Charnes [5]: LPs for zero-sum matrix (one shot) games
- Mangasarian and Stone [7]: Quadratic programs for bi-matrix games
- Parthasarathy et al [9]: for SIR-SET games
- Parthasarthy and Raghavan [8]: Order field property for some stochastic games
- Filar et al [6]: Math programs for stochastic games
- Vikas et al [11], etc.: Blackwell optimality for stochastic games
- V and H [10]: OPs for constrained single controller games
- Etc.

References I

E. Altman.
Constrained Markov Decision Processes.
Chapman and Hall/CRC, London, 1999.

E. Altman, K. Avrachenkov, N. Bonneau, M. Debbah, R. El-Azouzi, and D. Sadoc Menasche.
Constrained cost-coupled stochastic games with independent state processes.
Operations Research Letters, 36:160–164, 2008.

E. Altman, K. Avrachenkov, R. Marquez, and G. Miller.
Zero-sum constrained stochastic games with independent state processes.
Mathematical Methods of Operations Research, 62:375–386, 2005.

References II

E. Altman and A. Shwartz.
Constrained Markov games: Nash equilibria.
In V. Gaitsgory, J. Filar, and K. Mizukami, editors, *Annals of International Society of Dynamic games*, volume 5, pages 303–323. Birkhauser, 2000.

A. Charnes.
Constrained games and linear programming.
Proceedings of National Academy of Sciences of the USA, 39:639–641, 1953.

J. A. Filar, T. A. Schultz, F. Thuijsman, and O. J. Vrieze.
Nonlinear programming and stationary equilibria in stochastic games.
Mathematical Programming, 50:227–237, 1991.

References III

O. L. Mangasarian and H. Stone.
Two-person nonzero-sum games and quadratic programming.
Journal of Mathematical Analysis and Applications, 9:348–355, 1964.

T. Parthasarathy and T. E. S. Raghavan.
An order field property for stochastic games when one player controls transition probabilities.
Journal of Optimization Theory and Applications, 33:375–392, 1981.

T. Parthasarathy, S. H. Tijs, and O. J. Vrieze.
Stochastic games with state independent transitions and separable rewards.
In G. Hammer and D. Pallaschke, editors, *Selected Topics in OR and Mathematical Economics.*, Lecture Notes Series 226. Springer, 1984.

References IV

Vikas V. Singh and N. Hemachandra.
A characterization of stationary nash equilibria of single controller constrained stochastic games.
International Game Theory Review, 17:1540018–22, 2015.

Vikas V. Singh, N. Hemachandra, and K. S. Mallikarjuna Rao.
Blackwell optimality in stochastic games.
International Game Theory Review, 15:1340025–18, 2013.