A dynamical neural network approach for solving stochastic two-player zero-sum games

Dawen Wu

Université Paris Saclay, CNRS, CentraleSupelec
Laboratoire des Signaux et des Systèmes (L2S)

Joint work with Abdel Lisser

ECSO-CMS 2022

July 1st, 2022

Outline of the talk

1. Introduction
2. Stochastic two-player zero-sum games formulation
3. Neurodynamic optimization approach
4. Numerical experiments
5. Conclusion

Games:

Authors	Year	Games	Main results
von Neumann	1928	two-player zero-sum game	Saddle point existence
Nash	1950	n-players genral-sum game	Nash equilibrium existence
Charnes	1953	two-player zero-sum game with linear constraints	Saddle point existence
Singh & Lisser	2019	two-player zero-sum game with chance constraints	Saddle point existence

Neurodynamic optimization approach or **Dynamical neural network:**

Authors	Problems
Hopfield et al.(1986)	Linear programming problems (Hopfield network)
Kennedy et al.(1988)	Nonlinear programming problems based on the penalty method
Xia et al.(2007)	Nonlinear projection equations
Xu et al.(2020)	Constrained pseudoconvex programming problems

Charnes(1953) studied the two-player zero-sum game problem with linear constrained.

The LP formulation:

$$\begin{cases} \max_{x} x^T A y \\ \text{s.t.} \\ \quad Bx \leq b \\ \quad \mathbf{1}^T x = 1 \\ \quad x \geq 0, \end{cases} \qquad \begin{cases} \min_{y} x^T A y \\ \text{s.t.} \\ \quad Dy \geq d \\ \quad \mathbf{1}^T y = 1 \\ \quad y \geq 0, \end{cases}$$

$x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$ . . . decision vector
$x^* \in \mathbb{R}^n$, $y^* \in \mathbb{R}^m$ . . . Saddle point of the above linear programs
$B \in \mathbb{R}^{p\times n}, b \in \mathbb{R}^p$ . . . The linear constrant for player 1
$D \in \mathbb{R}^{q\times m}, d \in \mathbb{R}^q$ . . . The linear constrant for player 2

Charnes(1953) studied the two-player zero-sum game problem with linear constrained.

The LP formulation:

$$
\begin{cases}
\max_x x^T A y \\
\text{s.t.} \\
\quad Bx \leq b \\
\quad \mathbf{1}^T x = 1 \\
\quad x \geq 0,
\end{cases}
\qquad
\begin{cases}
\min_y x^T A y \\
\text{s.t.} \\
\quad Dy \geq d \\
\quad \mathbf{1}^T y = 1 \\
\quad y \geq 0,
\end{cases}
$$

$x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$ ... decision vector
$x^* \in \mathbb{R}^n$, $y^* \in \mathbb{R}^m$ ... Saddle point of the above linear programs
$B \in \mathbb{R}^{p \times n}, b \in \mathbb{R}^p$ ... The linear constrant for player 1
$D \in \mathbb{R}^{q \times m}, d \in \mathbb{R}^q$ ... The linear constrant for player 2

Consider the case of stocastic linear constraints

From linear constrained to linear **individual chance constraints:**

$$Bx \leq b \Longrightarrow P\left\{B_k^w x \leq b_k\right\} \geq \alpha_k^1, \quad \forall k \in \mathcal{J}_1$$

$$Dy \geq d \Longrightarrow P\left\{D_l^w y \geq d_l\right\} \geq \alpha_l^2, \quad \forall l \in \mathcal{J}_2$$

- $\mathcal{J}_1 = \{1, \ldots, p\}$, $\mathcal{J}_2 = \{1, \ldots, q\}$ are the index set of constraints.
- Now, each row vector B_k and D_l follows an elliptical distribution i.e. $B_k^w \sim \text{Ellip}_m\left(\mu_k^1, \Sigma_k^1, \varphi_k^1\right)$ and $D_l^w \sim \text{Ellip}_n\left(\mu_l^2, \Sigma_l^2, \varphi_l^2\right)$.
- $\alpha^1 = (\alpha_k^1)_{k \in \mathcal{J}_1}$ and $\alpha^2 = (\alpha_l^2)_{l \in \mathcal{J}_2}$ are confidence levels.

Consider the case of stocastic linear constraints

From linear constrained to linear **individual chance constraints:**

$$Bx \leq b \Longrightarrow P\left\{B_k^w x \leq b_k\right\} \geq \alpha_k^1, \quad \forall k \in \mathcal{J}_1$$

$$Dy \geq d \Longrightarrow P\left\{D_l^w y \geq d_l\right\} \geq \alpha_l^2, \quad \forall l \in \mathcal{J}_2$$

- $\mathcal{J}_1 = \{1, \ldots, p\}$, $\mathcal{J}_2 = \{1, \ldots, q\}$ are the index set of constraints.
- Now, each row vector B_k and D_l follows an elliptical distribution i.e. $B_k^w \sim \text{Ellip}_m\left(\mu_k^1, \Sigma_k^1, \varphi_k^1\right)$ and $D_l^w \sim \text{Ellip}_n\left(\mu_l^2, \Sigma_l^2, \varphi_l^2\right)$.
- $\alpha^1 = (\alpha_k^1)_{k \in \mathcal{J}_1}$ and $\alpha^2 = (\alpha_l^2)_{l \in \mathcal{J}_2}$ are confidence levels.

Consider the case of stocastic linear constraints

From linear constrained to linear **individual chance constraints:**

$$Bx \leq b \Longrightarrow P\left\{B_k^w x \leq b_k\right\} \geq \alpha_k^1, \quad \forall k \in \mathcal{J}_1$$

$$Dy \geq d \Longrightarrow P\left\{D_l^w y \geq d_l\right\} \geq \alpha_l^2, \quad \forall l \in \mathcal{J}_2$$

- $\mathcal{J}_1 = \{1, \dots, p\}$, $\mathcal{J}_2 = \{1, \dots, q\}$ are the index set of constraints.
- Now, each row vector B_k and D_l follows an elliptical distribution i.e. $B_k^w \sim \text{Ellip}_m\left(\mu_k^1, \Sigma_k^1, \varphi_k^1\right)$ and $D_l^w \sim \text{Ellip}_n\left(\mu_l^2, \Sigma_l^2, \varphi_l^2\right)$.
- $\alpha^1 = (\alpha_k^1)_{k\in\mathcal{J}_1}$ and $\alpha^2 = (\alpha_l^2)_{l\in\mathcal{J}_2}$ are confidence levels.

From two-player zero-sum game with linear constraint to the **stochastic two-player zero-sum game problem.**

Here is the stochastic optimization problem we are going to solve, denoted as $G(\alpha)$:

$$\begin{cases} \max_x x^T A y \\ \text{s.t.} \\ \quad P\left\{B_k^w x \leq b_k\right\} \geq \alpha_k^1, \quad \forall k \in \mathcal{J}_1 \\ \quad \mathbf{1}^T x = 1 \\ \quad x \geq 0, \end{cases}$$

$$\begin{cases} \min_y x^T A y \\ \text{s.t.} \\ \quad P\left\{D_l^w y \geq d_l\right\} \geq \alpha_l^2, \quad \forall l \in \mathcal{J}_2 \\ \quad \mathbf{1}^T y = 1 \\ \quad y \geq 0. \end{cases}$$

Denote the feasible strategy sets of the two player as $S_1\left(\alpha^1\right)$ and $S_2\left(\alpha^2\right)$, respectivly.

Assumption 1

1. *The set $S_1(\alpha^1)$ is strictly feasible, i.e., there exists an $x \in \mathbb{R}^n$ which is a feasible point of $S_1(\alpha^1)$ and the inequality constraints of $S_1(\alpha^1)$ are strictly satisfied by x.*
2. *The set $S_2(\alpha^2)$ is strictly feasible, i.e., there exists an $y \in \mathbb{R}^m$ which is a feasible point of $S_2(\alpha^2)$ and the inequality constraints of $S_2(\alpha^2)$ are strictly satisfied by y.*

The saddle point existence theorem of $G(\alpha)$:

Theorem 1 (Singh & Lisser (2019))

Consider a stochastic two-player zero-sum game $G(\alpha)$. Let the row vectors $B_k^w \sim Ellip_m\left(\mu_k^1, \Sigma_k^1, \varphi_k^1\right), k \in \mathcal{J}_1$, and $D_l^w \sim \text{Ellip}_n\left(\mu_l^2, \Sigma_l^2, \varphi_l^2\right), l \in \mathcal{J}_2$. For all k and l, $\Sigma_k^1 \succ 0$ and $\Sigma_l^2 \succ 0$. Then, there exists a saddle point equilibrium for the game $G(\alpha)$ for all $\alpha \in (0.5, 1]^{J_1} \times (0.5, 1]^{J_2}$.

Obtaining the optimal mixed strategy y^* of Player 2 can be reformulated as the following SOCP:

$$
(\mathcal{P}) \begin{cases}
\min_{y, v^1, (\delta_k^1)_{k\in\mathcal{J}_1}, \lambda^1} v^1 + \sum_{k\in\mathcal{J}_1} \lambda_k^1 b_k \\
\text{s.t.} \\
(1) Ay - \sum_{k\in\mathcal{J}_1} \lambda_k^1 \mu_k^1 - \sum_{k\in\mathcal{J}_1} \left(\Sigma_k^1\right)^{\frac{1}{2}} \delta_k^1 \leq v^1 \mathbf{1}_m \\
(2) - y^T \mu_l^2 + \Psi_{\xi_l^2}^{-1}\left(\alpha_l^2\right) \left\| \left(\Sigma_l^2\right)^{\frac{1}{2}} y \right\| \leq -d_l, \quad \forall l \in \mathcal{J}_2 \\
(3) \left\| \delta_k^1 \right\| \leq \lambda_k^1 \Psi_{\xi_k^1}^{-1}\left(\alpha_k^1\right), \quad \forall k \in \mathcal{J}_1 \\
(4) \mathbf{1}^T y = 1 \\
(5) y \geq 0 \\
(6) \lambda_k^1 \geq 0, \quad \forall k \in \mathcal{J}_1
\end{cases}
$$

Obtaining the optimal mixed strategy x^* of Player 1 can be reformulated as the following SOCP:

$$
(\mathcal{D}) \begin{cases}
\max_{x, v^2, (\delta_l^2)_{l \in \mathcal{J}_2}, \lambda^2} v^2 + \sum_{l \in \mathcal{J}_2} \lambda_l^2 d_l \\
\text{s.t.} \\
(1) A^T x - \sum_{l \in \mathcal{J}_2} \lambda_l^2 \mu_l^2 - \sum_{l \in \mathcal{J}_2} \left(\Sigma_l^2\right)^{\frac{1}{2}} \delta_l^2 \geq v^2 \mathbf{1}_n \\
(2) x^T \mu_k^1 + \Psi_{\xi_k^1}^{-1}\left(\alpha_k^1\right) \left\| \left(\Sigma_k^1\right)^{\frac{1}{2}} x \right\| \leq b_k, \quad \forall k \in \mathcal{J}_1 \\
(3) \left\| \delta_l^2 \right\| \leq \lambda_l^2 \Psi_{\xi_l^2}^{-1}\left(\alpha_l^2\right), \quad \forall l \in \mathcal{J}_2 \\
(4) \mathbf{1}^T x = 1 \\
(5) x \geq 0 \\
(6) \lambda_l^2 \geq 0, \quad \forall l \in \mathcal{J}_2
\end{cases}
$$

SOCP reformulation theorem:

Theorem 2 (Singh & Lisser (2019))

Consider a stochastic two-player zero-sum game $G(\alpha)$. Let the row vector $B^w \sim$ Ellip $_m\left(\mu_k^1, \Sigma_k^1, \varphi_k^1\right), k \in \mathcal{J}_1$, where $\Sigma_k^1 \succ 0$, and the row vector $D_l^w \sim \text{Ellip}_n\left(\mu_l^2, \Sigma_l^2, \varphi_l^2\right), l \in \mathcal{J}_2$ where $\Sigma_l^2 \succ 0$. Let Assumption 1 holds. Then, for a given $\alpha \in (0.5, 1]^p \times (0.5, 1]^q$, (x^*, y^*) is a saddle point equilibrium of the game $G(\alpha)$ if and only if $\left(y^*, v^{1*}, \left(\delta_k^{1*}\right)_{k\in J_1}, \lambda^{1*}\right)$ and $\left(x^*, v^{2*}, \left(\delta_l^{2*}\right)_{l\in \mathcal{J}_2}, \lambda^{2*}\right)$ are optimal solutions of primal-dual pair of SOCPs $(\mathcal{P})$ and $(\mathcal{D})$, respectively.

- The saddle point (x^*, y^*) of the game $G(\alpha)$ is contained in the optimal solution of the SOCPs $(\mathcal{P})$ and $(\mathcal{D})$.
- $(\mathcal{P})$ and $(\mathcal{D})$ are a primal-dual pair of SOCPs.

The SOCP $(\mathcal{P})$ can be written as:

$$\begin{aligned} &\min_s f(s) \\ &\text{s.t.} \\ &\quad g(s) \leq 0, \end{aligned} \tag{1}$$

The related KKT conditions are:

$$\begin{aligned} \nabla f(s) + \nabla g(s)^T u = 0 \\ g(s) \leq 0, \quad u^T \geq 0, \quad u^T g(s) = 0 \end{aligned} \tag{2}$$

- $s = \left(y, v^1, (\delta_k^1)_{k \in \mathcal{J}_1}, \lambda^1\right)$ is the primal variable.
- $u = \left(x, v^2, (\delta_l^2)_{l \in \mathcal{J}_2}, \lambda^2\right)$ is the dual variable.
- The saddle point (x^*, y^*) can be obtained by solving the KKT conditions (2).

Here, we introduce the time variable t, and $s(t)$ and $u(t)$ become time dependent functions.

Let $r(t)$ contain $s(t), u(t)$, i.e.,
$r(t) = (s(t), u(t)) = (y(t), v(t), \delta(t), \lambda(t), u(t))^T$.

We use this following ODE system, $\frac{dr}{dt} = \Phi(r)$, to solve the primal-dual pair of SOCPs, $(\mathcal{P})$ and $(\mathcal{D})$.

$$\frac{dr}{dt} = \begin{bmatrix} \frac{dy}{dt} \\ \frac{dv}{dt} \\ \frac{d\delta}{dt} \\ \frac{d\lambda}{dt} \\ \frac{du}{dt} \end{bmatrix} = \begin{bmatrix} -\left(\nabla f_y + \nabla g_y^T (u+g)^+\right) \\ -\left(\nabla f_v + \nabla g_v^T (u+g)^+\right) \\ -\left(\nabla f_\delta + \nabla g_\delta^T (u+g)^+\right) \\ -\left(\nabla f_\lambda + \nabla g_{\lambda^1}^T (u+g)^+\right) \\ (u+g)^+ - u \end{bmatrix}, \tag{3}$$

Theorem 3

The point $r^* = (y^*, v^*, \delta^*, \lambda^*, u^*)^T$ is the equilibrium point of the ODE system (3) if and only if it is also the KKT point of the SOCP problem.

Lemma 4

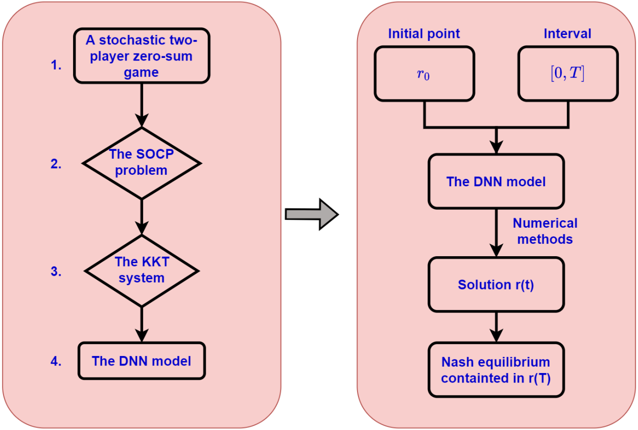
The equilibrium point of the proposed ODE system (3) is unique.

Theorem 5

The equilibrium point $r^* = (y^*, v^*, \delta^*, \lambda^*, u^*)$ of the proposed ODE system (3) is globally asymptotically stable.

- Theorem 3: The equilibrium point of the ODE system coincides with the optimal solution of the SOCPs.
- Theorem 5: Starting with any initial point (t_0, r_0), the state solution $r(t)$ converges to the optimal solution r^* as the time go to infinity, i.e., $\lim_{t \to \infty} r(t) = r^*$.

The flowchart summarize how we obtain the saddle point (x^*, y^*) of a stochastic two-player zero-sum game using neurodynamic optimization approach.



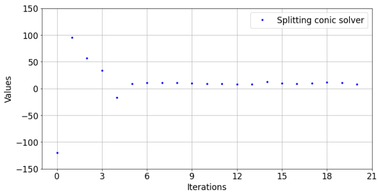
Formation to DNN models

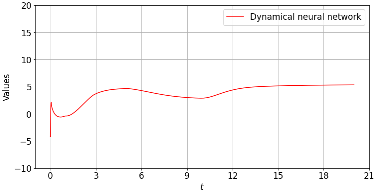
Solution of DNN models

We compare our neurodynamic optimization approach with the conventional approach, the SCS method.

Left: The SCS method. The objective value with respect to iterations.

Right: The neurodynamic approach. The objective value with respect to time t of the ODE system.





The SCS method solve the problem in a Discrete manner. The neurodynamic approach solve the problem in a Time-continuous manner.

The point (s, u) is an approximate KKT point with ϵ-error if it satisfies

$$\begin{aligned} \left\|\nabla f(s) + \nabla g(s)^T u\right\| &\leq \epsilon, \\ \|u \circ g(s)\| &\leq \epsilon, \\ \left\|g(s)_+\right\| &\leq \epsilon, \\ \|u_-\| &\leq \epsilon, \end{aligned} \tag{4}$$

ϵ-error represents how well the solution (s, u) satisfy the KKT conditions.

The ϵ-error comparison between the SCS method and our neurodynamic optimization approach:

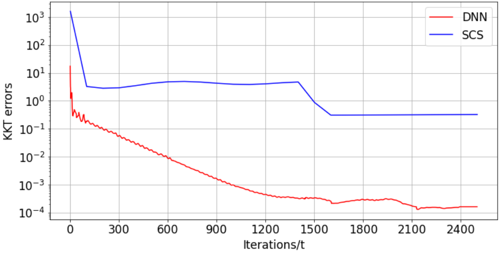


Figure: The x-axis represents the number of iterations (for the SCS method) or the time t (for the Neurodynamic approach). The y-axis represents the ϵ error.

The SCS method may stop or become slow after some iterations. The neurodynamic approach keep minimizing the ϵ-error thanks to the global convergence theorem.

Numerical results

Comparasion between Neurodynamic, SCS and CVXOPT

Game size	Probability distribution	α α_1	α_2	Neurodynamic CPU time	Value	ϵ-Error	SCS CPU time	Value	ϵ-Error	CVXOPT CPU time	Value	ϵ-Error
(4, 4)	Normal	0.8	0.8	1.79	4.59	0.12	0.015	5.03	8.20	0.031	4.39	0.00
		0.9	0.9	1.79	4.30	0.05	0.015	4.75	3.23	0.015	4.39	0.00
	Laplace	0.8	0.8	1.66	4.40	0.11	0.015	4.74	4.63	0.015	4.39	0.00
		0.9	0.9	1.28	4.67	0.02	0.015	4.80	2.86	0.015	4.64	0.00
(10, 10)	Normal	0.8	0.8	2.39	5.51	0.01	0.015	5.64	0.45	0.015	5.54	0.00
		0.9	0.9	2.67	5.78	0.09	0.015	5.80	0.25	0.015	5.88	0.00
	Laplace	0.8	0.8	2.61	5.53	0.03	0.015	5.66	1.00	0.015	5.59	0.00
		0.9	0.9	2.65	6.08	0.05	0.015	6.15	0.03	0.031	6.14	0.00
(50, 50)	Normal	0.8	0.8	56.33	5.26	0.18	0.063	4.71	4.59	0.271	5.15	0.00
		0.9	0.9	68.78	5.17	0.09	0.062	4.87	4.71	0.249	5.16	0.00
	Laplace	0.8	0.8	58.37	5.25	0.15	0.055	5.04	7.33	0.257	5.15	0.00
		0.9	0.9	47.07	5.16	0.08	0.046	5.17	2.18	0.249	5.18	0.00
(100, 100)	Normal	0.8	0.8	381.33	5.02	0.02	0.111	4.69	1.77	1.48	5.00	0.00
		0.9	0.9	369.88	4.99	0.04	0.105	4.97	1.56	1.59	5.00	0.00
	Laplace	0.8	0.8	319.04	5.00	0.02	0.109	4.82	1.63	Failed	Failed	Failed
		0.9	0.9	359.95	5.04	0.04	0.109	4.99	1.84	Failed	Failed	Failed
(200, 200)	Normal	0.8	0.8	8984.51	4.97	0.03	0.283	4.75	5.14	10.45	4.96	0.00
		0.9	0.9	8862.24	4.99	0.04	0.281	4.90	2.06	Failed	Failed	Failed
	Laplace	0.8	0.8	8381.08	4.98	0.04	0.252	4.77	4.18	Failed	Failed	Failed
		0.9	0.9	11218.82	5.00	0.05	0.265	4.96	0.94	Failed	Failed	Failed

- **Pros: 1. Achieve a lower ϵ-error 2. Solve the game of large size.**
- **Cons: Time consuming.**

Conclusion

- In this paper, we studied a neurodynamic approach to solve a two-player zero-sum game with stochastic linear constraints.
- We show that the equilibrium point of the ODE system is the saddle point for the game.
- We show that the ODE system can converge to the saddle point of the game.
- We use this neurodynamic approach to solve the game of size up to $(200, 200)$.

Reference

Thank you for your attention

Reference:

1. Wu, D., Lisser, A. (2022). A dynamical neural network approach for solving stochastic two-player zero-sum games. Neural Networks 152: 140-149 (2022).

2. Singh, V. V., Lisser, A. (2019). A second-order cone programming formulation for two player zero-sum games with chance constraints. European Journal of Operational Research, 275 , 839–845.

3. Y. Xia, G. Feng, A new neural network for solving nonlinear projection equations, Neural Networks 20 (5) (2007) 577–589.

4. C. Xu, Y. Chai, S. Qin, Z. Wang, J. Feng, A neurodynamic approach to nonsmooth constrained pseudo-convex optimization problem, Neural Networks (2020) 180–192