BAAI

智源学者成果展示——人工智能的数理基础

作者　戴彧虹（中国科学院数学与系统科学研究院研究员、智源研究员）

2020年6月

GANs 相关的约束极小极大问题理论

中国科学院数学与系统科学研究院研究员、智源研究员戴彧虹等研究了来源于生成对抗网络(Generative Adversarial Networks)、对抗训练(Adversarial Training)和多智能体强化学习(Multi-Agent Reinforcement Learning)中的约束极小极大问题，并在局部极小极大点(Local Minimax Point)的意义下给出了最优性理论。 特别地，在内层满足 Jacobian 唯一性假设下，证明了一阶和二阶必要性最优条件和二阶充分性最优条件；同时在对内层满足强正则假设下，证明了一阶必要性最优条件。

Yu-Hong Dai and Liwei Zhang, Optimality Conditions for Constrained Minimax Optimization, arXiv:2004.09730v1 (Accepted by CSAM), 2020.

Theorem 3.1 *(Necessary Optimality Conditions) Let $(x^*, y^*) \in \Re^n \times \Re^m$ be a point around which f, h, g are twice continuously differentiable and H, G are twice continuously differentiable around x^*. Let (x^*, y^*) be a local minimax point of Problem (1.1). Assume that the linear independence constraint qualification holds at y^* for constraint set $Y(x^*)$. Then there exists a unique vector $(\mu^*, \lambda^*) \in \Re^{m_1} \times \Re^{m_2}$ such that*

$$\begin{aligned} &\nabla_y \mathcal{L}(x^*; y^*, \mu^*, \lambda^*) = 0, \\ &h(x^*, y^*) = 0, \\ &0 \geq \lambda^* \perp g(x^*, y^*) \leq 0. \end{aligned} \tag{3.4}$$

For any $d_y \in C_{x^*}(y^*)$, we have that

$$\langle \nabla^2_{yy} \mathcal{L}(x^*; y^*, \mu^*, \lambda^*) d_y, d_y \rangle \leq 0. \tag{3.5}$$

Assuming Problem (P_{x^*}) satisfies Jacobian uniqueness conditions at (y^*, μ^*, λ^*) and the Mangasarian-Fromovitz constraint qualification holds at x^* for the constraint set Φ, there exists $(u^*, v^*) \in \Re^{n_1} \times \Re^{n_2}$ such that

$$\begin{aligned} &\nabla_x \mathcal{L}(x^*; y^*, \mu^*, \lambda^*) + \mathcal{J}H(x^*)^T u^* + \mathcal{J}G(x^*)^T v^* = 0, \\ &H(x^*) = 0, \\ &0 \leq v^* \perp G(x^*) \leq 0. \end{aligned} \tag{3.6}$$

The set of all (u^*, v^*) satisfying (3.6), denoted by $\Lambda(x^*)$, is nonempty compact convex set. Furthermore, for every $d_x \in C(x^*)$ where $C(x^*)$ is defined by (3.3),

$$\begin{aligned} &\max_{(u,v)\in\Lambda(x^*)} \left\{ \left\langle \left[\sum_{j=1}^{n_1} u_i \nabla^2_{xx} H_j(x^*) + \sum_{i=1}^{n_2} v_i \nabla^2_{xx} G_i(x^*) \right] d_x, d_x \right\rangle \right\} \\ &\quad + \left\langle \left[\nabla^2_{xx} \mathcal{L}(x^*; y^*, \mu^*, \lambda^*) - N(x^*)^T K(x^*)^{-1} N(x^*) \right] d_x, d_x \right\rangle \geq 0, \end{aligned} \tag{3.7}$$

where $K(x)$ is defined by (2.6) and $N(x)$ is defined by

$$N(x) = \begin{bmatrix} \nabla^2_{x,y} \mathcal{L}(x; y(x)\mu(x), \lambda(x)) \\ 0 \\ \mathcal{J}_x h(x, y(x)) \\ \mathcal{J}_x g(x, y(x)) \end{bmatrix}. \tag{3.8}$$

Beijing Academy of
Artificial Intelligence

微信关注
北京智源人工智能研究院