Policy iteration for discrete-time systems with discounted costs: stability and near-optimality guarantees

Jonathan De Brusse[*1,2]

[1]Université de Lorraine – Centre de Recherche en Automatique de Nancy (CRAN) – France
[2]University of Melbourne – Australie

Résumé

Policy Iteration (PI) is a well-known algorithm in dynamic programming theory, commonly used to solve optimal control problems. The first goal of PI is to iteratively generate a policy (a sequence of control inputs) that minimizes a cost function representing the quality of the system's trajectories according to a chosen performance metric. This metric evaluates the system's state and control input at each time step.

In practice, PI is typically stopped after a finite number of iterations, making it essential to ensure that the difference between the cost induced by the final policy generated by PI and the optimal cost remains acceptably small. This concept is referred to as near-optimality guarantees. Additionally, it is crucial that the system controlled by PI meets certain stability requirements. Although many results on near-optimality and stability guarantees for systems controlled by dynamic programming algorithms exist in the literature, most are focused on undiscounted costs and alternative algorithms to PI. However, discounted costs are prevalent in dynamic programming and reinforcement learning due to their favorable properties and their relevance in some applications. Therefore, it is important to establish stability guarantees for systems controlled by dynamic programming algorithms like PI with discounted costs.

In this seminar, we will first demonstrate that, due to the presence of the discount factor, PI can be initialized with a non-stabilizing policy while still ensuring a finite induced cost. Assuming that the optimal closed-loop system satisfies a global asymptotic stability property, we then present a general near-optimality bound that is potentially tighter than those found in existing literature. Finally, we combine this near-optimality bound with the stability of the optimal closed-loop system to establish stability guarantees for systems controlled by PI after a sufficient number of iterations.

Mots-Clés: Optimal control, Stability of nonlinear systems, Lyapunov methods

---

[*]Intervenant