Email Address:

zp2130@caa.columbia.edu (mailto:zp2130@caa.columbia.edu)

p@caa.columbia.edu (mailto:p@caa.columbia.edu)

Neural-network-based decentralized control of continuous-time nonlinear interconnected systems with unknown dynamics

Neural-network-based decentralized control of continuous-time nonlinear interconnected systems with unknown dynamics (https://blogs.cuit.columbia.edu/zp2130/files/2019/03/Neural-network-based-decentralized-control-of-continuous-time-nonlinear-interconnected-systems-with-unknown-dynamics.pdf)

– Math and Optimal Control (https://blogs.cuit.columbia.edu/zp2130/files/2019/03/IMG_20190311_0004_NEW.pdf)

Problem formulation

Consider a continuous-time nonlinear large-scale system ∑ composed of N interconnected subsystems described by

$$\sum : \dot{x}_i(t) = f_i[x_i(t)] + g_i[x_i(t)]\{u_i[x_i(t)] + Z_i[x(t)]\} \quad (1)$$
$$i = 1, 2, ..., N$$

where

$x_i(t) \in R^{n_i}$: state.

The overall state of the large-scale system ∑ is denoted by $x = [x_1^T x_2^T ... x_N^T]^T \in \mathbb{R}^n, where\ n = \sum_{i=1}^{N} n_i$

$u_i[\,x_i(t)\,] \in R^{m_i}$: control input vector of the ith subsystem.

f_i : continuous nonlinear internal dynamics function. f_i (0)=0. $\mathbb{R}^{n_i} \to \mathbb{R}^{n_i}$

g_i[x_i(t)] : input gain function $\mathbb{R}^{n_i} \to \mathbb{R}^{n_i \times m_i}$

Z_i [x(t)] : interconnected term for the ith subsystem.

The ith isolated subsystem

$$\sum_i : \dot{x}_i(t) = f_i[x_i(t)] + g_i[x_i(t)]\{u_i[x_i(t)]\} \quad (2)$$
$$i = 1, 2, ..., N$$

Decentralized control law

Optimal control

Reinforcement Learning and Optimal Control Methods for Uncertain Nonlinear Systems (https://blogs.cuit.columbia.edu/zp2130/files/2019/03/Reinforcement-Learning-and-Optimal-Control-Methods-for-Uncertain-Nonlinear-Systems.pdf)

Page 27-29 2.3 Infinite Horizon Optimal Control Problem is the same as Definition 1.

Notation:

$x(t) \in \chi \subseteq \mathbb{R}^n$: state.

$u(t) \in U \subseteq \mathbb{R}^m$: control input.

$$\dot{x} = F(x, u) \quad \text{(2-5)}$$

Cost function for the system Eq. 2-5:

$$J(x(t), u(\tau)_{t \le \tau < \infty}) = \int_t^\infty r(x(s), u(s))ds \quad \text{(2-6)}$$

where t : initial time.

r(x,u) $\in R$: immediate or local cost for the state and control.

$$r(x, u) = Q(x) + u^T Ru \quad \text{(2-7)}$$

where Q(x) $\in R$ continuously differentiable and positive definite.

$R \in R^{m \times m}$: positive-definite symmetric matrix.

Optimal value function:

$$V^*(x(t)) = min_{u(\tau) \in \Psi(\chi),\ t \le \tau < \infty} \int_t^\infty r\{x(s), u[x(s)]\}ds \quad \text{(2-8)}$$

where

$\Psi(\cdot)$: set of admissible controls.

Bellman's principle of optimally can be used to derive the following optimality condition

$$0 = min_{u(t) \in \Psi(\chi)} \left[r(x, u) + \frac{\partial V^*(x)}{\partial x} F(x, u) \right] \quad \textbf{(2-9)}$$

which is a nonlinear partial differential equation (PDE), also called the **HJB equation**.

Optimal control: (using convex local cost in Eqs. 2-7 and 2-9.)

$$u^*(x) = -\frac{1}{2} R^{-1} \frac{\partial F(x, u)^T}{\partial u} \frac{\partial V^*(x)^T}{\partial x} \quad \text{(2-10)}$$

For the control-affine dynamics of the form

$$\dot{x} = f(x) + g(x)u = F(x, u) \quad \text{(2-11)}$$

Eq. 2-10 -> in terms of the system state

$$u^*(x) = -\frac{1}{2}R^{-1}g^T(x)\frac{\partial V^*(x)^T}{\partial x} \quad \text{(2-12)}$$

The **HJB** in **Eq. 2-9** can be rewritten in terms of the optimal value function by substituting for the local cost in Eq. 2-7, the system in Eq. 2-11 and the optimal control in Eq. 2-12, as

$$\begin{aligned} 0 &= min_{u(t)\in\Psi(\chi)}\left[r(x,u) + \frac{\partial V^*(x)}{\partial x}F(x,u)\right] \\ &= min_{u(t)\in\Psi(\chi)}\left[Q(x) + u^T Ru + \frac{\partial V^*(x)}{\partial x}\left[f(x)+g(x)u\right]\right] \\ &= Q(x) + {u^*}^T Ru^* + \frac{\partial V^*(x)}{\partial x}\left[f(x)+g(x)u^*\right] \\ &= Q(x) + \left[-\frac{1}{2}R^{-1}g^T(x)\frac{\partial V^*(x)^T}{\partial x}\right]^T R\left[-\frac{1}{2}R^{-1}g^T(x)\frac{\partial V^*(x)^T}{\partial x}\right] + \frac{\partial V^*(x)}{\partial x}\left\{f(x)+g(x)\left[-\frac{1}{2}R^{-1}g^T(x)\frac{\partial V^*(x)^T}{\partial x}\right]\right\} \end{aligned}$$

Invalid Equation

$$\xrightarrow[R:symmetric]{R^T=R} {R^{-1}}^T = R^{-1}$$

$$\begin{aligned} 0 &= min_{u(t)\in\Psi(\chi)}\left[r(x,u) + \frac{\partial V^*(x)}{\partial x}F(x,u)\right] \\ &= min_{u(t)\in\Psi(\chi)}\left[Q(x) + u^T Ru + \frac{\partial V^*(x)}{\partial x}\left[f(x)+g(x)u\right]\right] \\ &= Q(x) + {u^*}^T Ru^* + \frac{\partial V^*(x)}{\partial x}\left[f(x)+g(x)u^*\right] \\ &= Q(x) + \frac{1}{4}\frac{\partial V^*(x)}{\partial x}g(x){R^{-1}}^T g^T(x)\frac{\partial V^*(x)^T}{\partial x} + \frac{\partial V^*(x)}{\partial x}f(x) - \frac{1}{2}\frac{\partial V^*(x)}{\partial x}g(x)R^{-1}g^T(x)\frac{\partial V^*(x)^T}{\partial x} \\ &= Q(x) + \frac{1}{4}\frac{\partial V^*(x)}{\partial x}g(x)R^{-1}g^T(x)\frac{\partial V^*(x)^T}{\partial x} + \frac{\partial V^*(x)}{\partial x}f(x) - \frac{1}{2}\frac{\partial V^*(x)}{\partial x}g(x)R^{-1}g^T(x)\frac{\partial V^*(x)^T}{\partial x} \end{aligned}$$

$$\begin{aligned} 0 &= Q(x) + \frac{\partial V^*(x)}{\partial x}f(x) - \frac{1}{4}\frac{\partial V^*(x)}{\partial x}g(x)R^{-1}g^T(x)\frac{\partial V^*(x)^T}{\partial x} \quad \textbf{(2-13)} \\ 0 &= V^*(0) \end{aligned}$$

---

edit (https://blogs.cuit.columbia.edu/zp2130/wp-admin/post.php?post=4956&action=edit)

Author: Z Pei (https://blogs.cuit.columbia.edu/zp2130/author/zp2130/) on March 4, 2019

Categories: AI (https://blogs.cuit.columbia.edu/zp2130/category/ai/), Control (https://blogs.cuit.columbia.edu/zp2130/category/control/), Reinforcement Learning (https://blogs.cuit.columbia.edu/zp2130/category/reinforcement-learning/), RL (https://blogs.cuit.columbia.edu/zp2130/category/rl/)

Tags: AI (https://blogs.cuit.columbia.edu/zp2130/tag/ai/), Control (https://blogs.cuit.columbia.edu/zp2130/tag/control/), Optimal Control (https://blogs.cuit.columbia.edu/zp2130/tag/optimal-control/), Reinforcement Learning (https://blogs.cuit.columbia.edu/zp2130/tag/reinforcement-learning/), RL (https://blogs.cuit.columbia.edu/zp2130/tag/rl/)

Other posts

Reinforcement Learning is Direct Adaptive Optimal Control (https://blogs.cuit.columbia.edu/zp2130/reinforcement_learning_is_direct_adaptive_optimal_control/) «» Decentralized Optimal Control of Distributed Interdependent Automata With Priority Structure (https://blogs.cuit.columbia.edu/zp2130/decentralized_optimal_control_of_distributed_interdependent_automata_with_priority_structure/)

Last posts

- Symbolic Netlist to Innovus-friendly Netlist (https://blogs.cuit.columbia.edu/zp2130/symbolic_netlist_to_innovus-friendly_netlist/)
- Finite-Sample Convergence Rates for Q-Learning and Indirect Algorithms (https://blogs.cuit.columbia.edu/zp2130/finite-sample_convergence_rates_for_q-learning_and_indirect_algorithms/)
- Solving H-horizon, Stationary Markov Decision Problems In Time Proportional To Log(H) (https://blogs.cuit.columbia.edu/zp2130/paul_tseng_1990/)
- Randomized Linear Programming Solves the Discounted Markov Decision Problem In Nearly-Linear (Sometimes Sublinear) Run Time (https://blogs.cuit.columbia.edu/zp2130/randomized_linear_programming_solves_the_discounted_markov_decision_problem_in_nearly-

linear_sometimes_sublinear_run_time/)

- KL Divergence (https://blogs.cuit.columbia.edu/zp2130/kl_divergence/)
- The Asymptotic Convergence-Rate of Q-learning (https://blogs.cuit.columbia.edu/zp2130/the_asymptotic_convergence-rate_of_q-learning/)
- Hierarchical Apprenticeship Learning, with Application to Quadruped Locomotion (https://blogs.cuit.columbia.edu/zp2130/hierarchical_apprenticeship_learning_with_application_to_quadruped_locomotion/)
- Policy Gradient Methods (https://blogs.cuit.columbia.edu/zp2130/policy_gradient_methods/)
- Actor-Critic Algorithms for Hierarchical Markov Decision Processes (https://blogs.cuit.columbia.edu/zp2130/actor-critic_algorithms_for_hierarchical_markov_decision_processes/)
- Hierarchical Deep Reinforcement Learning: Integrating Temporal Abstraction and Intrinsic Motivation (https://blogs.cuit.columbia.edu/zp2130/hierarchical_deep_reinforcement_learning_integrating_temporal_abstraction_and_intrinsic_motivation/)