

МЕТОДЫ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ В ЗАДАЧЕ УПРАВЛЕНИЯ ОДНИМ КЛАССОМ МНОГОСВЯЗНЫХ СИСТЕМ

Яковлева Маргарита Александровна

Студент

Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия

E-mail: myalexq42@gmail.com

Научный руководитель — *Фомичев Василий Владимирович*

В теории автоматического управления широко известна задача стабилизации перевернутого маятника в вертикальном положении. Эта известность и популярность объясняется тем, что математическая модель такого объекта управления описывается системой дифференциальных уравнений, которые довольно просто выводятся и, при этом, являются нелинейными, а положение равновесия объекта является неустойчивым. Эта классическая задача также широко используется в качестве эталона для тестирования алгоритмов управления, так как модель маятника может быть промоделирована с достаточной точностью и в большинстве случаев для тестирования достаточно компьютерного моделирования. Ее решение находит практическое применение в самых разных областях, таких как системы ракетного управления и наведения, робототехника, управление строительными кранами, самобалансирующиеся транспортные устройства и т. д.

В данной работе рассматривается класс многозвенных перевернутых маятников на тележке, в частности - двухзвенных. Ставится задача стабилизации маятника в вертикальном положении с использованием алгоритмов обучения с подкреплением.

Математическая модель двухзвенного перевернутого маятника описывается следующей системой дифференциальных уравнений:

$$\left\{ \begin{array}{l} (M + m_1 + m_2)\ddot{x} + (m_1 + m_2)l_1(\cos\theta_1\ddot{\theta}_1 - \sin(\theta_1)\dot{\theta}_1^2) + m_2l_2(\cos(\theta_2)\ddot{\theta}_2 - \sin(\theta_2)\dot{\theta}_2^2) = F \\ (m_1 + m_2)l_1\cos\theta_1\ddot{x} + (I_1 + (m_1 + m_2)l_1^2)\ddot{\theta}_1 + m_2l_1l_2(\cos(\theta_1 - \theta_2)\ddot{\theta}_2 + \sin(\theta_1 - \theta_2)\dot{\theta}_2^2) - (m_1 + m_2)l_1g\sin\theta_1 = 0 \\ m_2l_2\cos\theta_2\ddot{x} + m_2l_1l_2(\cos(\theta_1 - \theta_2)\ddot{\theta}_1 - \sin(\theta_1 - \theta_2)\dot{\theta}_1^2) + (m_2l_2^2 + I_2)\ddot{\theta}_2 - m_2l_2g\sin\theta_2 = 0 \end{array} \right.$$

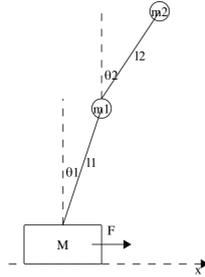
Иллюстрации

Рис. 1. Двухзвенный перевернутый маятник

Методы обучения с подкреплением разделяются на множество конкретных алгоритмов. После анализа литературы на данную тему были выделены несколько наиболее перспективных для решения задачи стабилизации двухзвенного перевернутого маятника алгоритмов, в частности это - Deep Deterministic Policy Gradient [3], Deep Q-learning [3], Proximal Policy Optimization [1].

В рамках данной работы были реализованы отмеченные алгоритмы, а также проведены эксперименты по стабилизации двухзвенного перевернутого маятника.

Литература

1. Lillicrap T., Hunt J., Pritzel A., Heess N., Erez T., Tassa Y., Silver D., Wierstra D. Continuous control with deep reinforcement learning. arXiv preprint arXiv:1509.02971
2. Gustafsson F. Control of Inverted Double Pendulum using Reinforcement Learning. IEEE 2022 8th International Conference on Control, Instrumentation and Automation (ICCIA)
3. Bates D. Virtual Reinforcement Learning for Balancing an Inverted Pendulum in Real Time.