Distributional Reinforcement Learning

Eighth week of machine learning workshop

Outline

- Overview of Classical RL
  - Value function
  - Action-value function
  - Bellman and Bellman optimality update
  - What are the sources of randomness in evaluation return?
  - The limitation of Value (taking expectation of return and reporting just a number as a value of the state or value of the state-action)
- Value distribution
  - Why considering distribution for value or action-value?
  - Risk aware RL
  - Bellman update equation in the case of value distribution
  - Bellman update equation for both policy evaluation and control settings
- Distance between distributions
  - KL distance
  - Total Variation
  - Wasserstein distance
  - Why we are using Wasserstein distance in policy evaluation setting instead of KL and other distances?
- Contraction Mapping
  - Proving the convergence of Q-Learning in classical RL (In the expectation scenario)

- Contraction of the policy evaluation Bellman operator in distributional RL by applying Wasserstein distance
- Instability in the control setting
- Approximate Distributional Learning
    - Projected Bellman update
    - C51 Algorithm
    - Experimental results on Atari games

References

1. Bellemare, Marc G., Will Dabney, and Rémi Munos. 2017. "A Distributional Perspective on Reinforcement Learning." CoRR. 2017
2. Dabney, W., M. Rowland, M. G. Bellemare, and R. Munos. 2017. "Distributional Reinforcement Learning with Quantile Regression." ArXiv E-Prints, October.
3. Mavrin, Borislav, et al. "Distributional Reinforcement Learning for Efficient Exploration." arXiv preprint arXiv:1905.06125 (2019).
4. Mnih, Volodymyr, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, et al. 2015. "Human-Level Control Through Deep Reinforcement Learning." Nature 518 (7540). Nature Research:529–33.
5. Yang, Derek, et al. "Fully Parameterized Quantile Function for Distributional Reinforcement Learning." Advances in Neural Information Processing Systems. 2019.