Proposal - Learning to Optimize Motion Planning

Yufan Zhao, Mingyang Wang

1 Objective

The main objective of this topic is to utilize the reinforcement learning technique to solve a trajectory optimization problem and evaluate the performance of such a method as well. Given the robot and environment, a trajectory planning problem can be formulated as an unconstrained optimization problem, which can further be solved by some traditional optimization algorithms like gradient descent. However, these traditional optimization algorithms are generally defined. They do not capture and take advantage of the certain features of objective functions that they are optimizing on to achieve faster convergence and local optima avoidance. With the help of reinforcement learning, a learned optimization algorithm can be generated for a certain class of similar objective functions. This optimization algorithm utilizes features of those similar objectives that it was trained on and can be generalized to some unseen but similar objectives. This notion is valuable for robot motion planning problems because we can use task-specific, robust (to environment change) optimization algorithms to achieve a feasible solution with a higher speed.

2 Related work

The work of [1] formulates the trajectory optimization problem for a certain robot and environment as a general unconstrained optimization problem. The work of [2] proposes to use reinforcement learning (guided policy search) to develop an optimization algorithm for a certain class of objective functions. They yield good results on logistic regression, robust linear regression and neural networks [3]. However, their idea was not evaluated on objective function originated from trajectory planning. Thus, we need to test the idea on trajectory planning problems.

3 Technical outline

As the first attempt to test this idea, we need to start simply. We are going to work on a point robot in a 2D environment with multiple rectangle obstacles. The start and endpoint of the robot are fixed, while the location, size are changing. By varying the environment, we can generate different objective functions based on the principle that the trajectory should be as short and smooth as possible and can successfully avoid obstacles (we may also include dynamic property in a later stage). These objectives are sent to the reinforcement learning algorithm as input training data. Then through learning, the reinforcement learning framework outputs an optimal policy, which refers to an optimization algorithm with best performance. The reinforcement learning method used in this context is guided policy search, which works in a supervised fashion. It uses updates of traditional optimization methods to guide the policy update. The policy can further be improved by the reinforcement learning mechanism to go beyond traditional methods. If time allows, we may also extend our environment to the 2D robot arm and a 3D robot arm cases.

The main contribution of our project may include:

- Compare the performance of the learned optimization algorithm with handcrafted optimization algorithms on robot trajectory planning tasks.
- Evaluate the generalization power or robustness of the learned optimization algorithm.

References:

[1] Zucker, M., Ratliff, N., Dragan, A. D., Pivtoraiko, M., Klingensmith, M., Dellin, C. M., Bagnell, J. A., and Srinivasa, S. S. (2013). Chomp: Covariant hamiltonian optimization for motion planning. International Journal of Robotics Research, 32(9-10):1164-1193.
[2] Li, K. and Malik, J. (2016). Learning to optimize. CoRR, abs/1606.01885.
[3] Li, K., & Malik, J. (2017). Learning to optimize neural nets. arXiv preprint arXiv:1703.00441.