ADLR Project Draft Proposal

Learning to Optimize Motion Planning

Objective: The main objective for this topic is to evaluate the performance of using reinforcement learning to solve a trajectory optimization problem. Given a robot and an environment, an objective function for trajectory optimization can be generated according to [1]. [2] takes varies similar objective functions (from similar robot and environment) as input training samples and outputs an optimization algorithm that is suitable for those training objective functions. The final evaluation of such an optimization algorithm is performed on some unseen objective functions (unseen robots and environments). The performance can be defined from two perspectives: 1. If a test objective is successfully optimized by the learned algorithm, how unsimilar it could be to training objectives? 2. Is the learned optimization algorithm faster than manually crafted optimization algorithm like SGD?

Related work:

The work of [1] formulated the trajectory optimization problem for a certain robot and environment as a general unconstraint optimization problem. They proposed a novel norm so that the objective function is *manifest invariant* to the choice of parametrization and the gradient is *covariant* to re-parametrization [1]. This ensures that any first order method applied to solve the optimization problem results in same trajectory update. The work of [2] proposed to use reinforcement learning (guided policy search) to develop an optimization algorithm for a certain class of objective functions. They yielded 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. As mentioned above, we also need to generate varies robot and environment pairs as training samples for the reinforcement learning algorithm. This requires parametrization of our environment which is discussed neither by [1] nor [2].

Technical outline:

To achieve our goal, we first need to formulate the robots and environments that we are interested in as two classes of functions, namely the forward kinematics for robot and the cost function for environment. Let $\mathrm{B} \subset R^3$ be the body points of the robot, and Q the configuration space, the forward kinematics $x: Q \times \mathrm{B} \rightarrow R^3$ maps a robot configuration q and a body point u to a point $x(q, u)$ in 3D workspace. A cost function $c: R^3 \rightarrow R$ maps a 3D point in workspace to its cost. Both the forward kinematics x and cost function c are parametrized by w to consider varying environment and robot. By changing w, multiple training objectives can be generated and sent to reinforcement learning algorithm. The reinforcement learning algorithm outputs an optimization algorithm to solve similar trajectory planning tasks. The main contribution of our project may include:

1. Compare the learned optimization algorithm with handcrafted optimization algorithm on robot trajectory planning tasks.
2. Evaluate the generalization power 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*.