Derivative-free optimization

A survey on random search for reinforcement learning

Thomas Lautenschläger

Abstract Insert your abstract here. Include keywords, PACS and mathematical subject classification numbers as needed.

Address(es) of author(s) should be given

1 Introduction

References: [7] [3] [1] [2] [12] [8] [6] [10] [5] [4] [13] [11] [9] [8]

2 Derivative-free and derivative-based optimization

3 Random search

3.1 Train a linear policy

3.2 Train a stochastic policy

4 Sample efficiency

5 Performance repeatability

6 Critical view

7 Outlook

8 Conclusion

References

1. Barmi, Z.A., Ebrahimi, A.H., Feldt, R.: Alignment of requirements specification and testing: A systematic mapping study. Proceedings - 4th IEEE International Conference on Software Testing, Verification, and Validation Workshops, ICSTW 2011 pp. 476–485 (2011). DOI 10.1109/ICSTW.2011.58
2. Bogolubsky, L., Dvurechensky, P., Gasnikov, A., Gusev, G., Nesterov, Y., Raigorodskii, A., Tikhonov, A., Zhukovskii, M.: Learning Supervised PageRank with Gradient-Based and Gradient-Free Optimization Methods (Nips 2016), 1–9 (2016). DOI 10.2307/3833729. URL http://arxiv.org/abs/1603.00717
3. Flaxman, A.D., Kalai, A.T., McMahan, H.B.: Online convex optimization in the bandit setting: gradient descent without a gradient pp. 385–394 (2004). URL http://arxiv.org/abs/cs/0408007
4. Hu, J., Fu, M.C., Ramezani, V.R., Marcus, S.I.: An evolutionary random policy search algorithm for solving markov decision processes. INFORMS Journal on Computing **19**(2), 161–174 (2007). DOI 10.1287/ijoc.1050.0155
5. Mania, H., Guy, A., Recht, B.: Simple random search provides a competitive approach to reinforcement learning pp. 1–22 (2018). URL http://arxiv.org/abs/1803.07055
6. Me107485: Title of the 107485th document pp. 1–32 (1995). DOI 10.1145/1389095.1389254
7. Miguel, L., Nikolaos, R.: Derivative-free optimization: A review of algorithms and comparison of software implementations. Journal of Global Optimization **56**, 1247–1293 (2013). DOI 10.1007/s10898-012-9951-y
8. Nesterov: Random Gradient-Free Minimization of Convex Functions Yurii NESTEROV (2011)
9. Osband, I., Aslanides, J., Cassirer, A.: Randomized Prior Functions for Deep Reinforcement Learning (Nips) (2018). DOI 10.5465/amj.2012.4005. URL http://arxiv.org/abs/1806.03335

10. Peker, M.: Comparison of the Complex Valued and Real Valued Neural Networks Trained with Gradient Descent and Random Search Algorithms. Esann **03**(April), 1121–1128 (2013). URL https://www.elen.ucl.ac.be/Proceedings/esann/esannpdf/es2011-42.pdf
11. Tang, Y.: Implicit Policy for Reinforcement Learning pp. 1–21
12. Tetris, I., Nowell, L., Method, C.e., Function, V., Selection, A.: Learning Tetris Using the Noisy Cross-Entropy Method **2941**(1996), 2936–2941 (2006)
13. Tu, S., Recht, B.: The Gap Between Model-Based and Model-Free Methods on the Linear Quadratic Regulator : An Asymptotic Viewpoint pp. 1–45 (2018)