Example: Learning to Control the Learning Rate with Reinforcement Learning

Carolin Benjamins[1,2]

[1]Leibniz University Hannover
[2]`https://www.AutoML.org`

Method. Here the learning rate of the `sgd_env` is dynamically adapted with a learned Reinforcement Learning agent. The assumption is that the state and the reward are expressive enough to guide an RL agent to learn how to adapt the learning rate.

The learning rate $\gamma \in [0.00001, 1]$ is adjusted in the log-space. We used a SAC Haarnoja et al. (2018) agent from the library `stable-baselines3` Raffin et al. (2021) and trained it for 1000000 steps. For training we used $n_{instances} = 1000$ and a policy architecture of two hidden layers with 64 neurons each. We switch the instance each reset in a round-robin manner.

Limitations. There are few possible limitations. First, no hyperparameters are tuned which means that the selected agent might perform better. Second, no information about instances is leveraged during training making the agent oblivious of the cMDP. For example, instance features could be included in the state or be used to build a curriculum. Third, the state is a simple concatenation of the observation dictionary content. The state features do not necessarily lie on the same scale, maybe they can be scaled to the unit interval or preprocessed and/or condensed in any other way.

Reproducibility. We provide the code and the command to reproduce the model in `README.md`. The training took approximately 3.5 hours on an Intel Core i9-9900K CPU with 16 cores.

Hyperparameters. This method introduces the type of RL agent and its hyperparameters. The RL agent type was manually set, and the default hyperparameters are used.

References

Haarnoja, T., Zhou, A., Abbeel, P., and Levine, S. (2018). Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In Dy, J. G. and Krause, A., editors, *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018*, volume 80 of *Proceedings of Machine Learning Research*, pages 1856–1865. PMLR.

Raffin, A., Hill, A., Gleave, A., Kanervisto, A., Ernestus, M., and Dormann, N. (2021). Stable-baselines3: Reliable reinforcement learning implementations. *Journal of Machine Learning Research*, 22(268):1–8.

