Notes on Variance Reduction Techniques for Gradient Estimates in Reinforcement Learning

compiled by D.Gueorguiev, 4/16/2025

# Introductory Notes

Policy gradient methods for reinforcement learning avoid some of the undesirable properties of the value function approaches, such as policy degradation. However

# References

[1] [Variance Reduction Techniques for Gradient Estimates in Reinforcement Learning, Evan Greensmith, Peter L. Bartlett, Jonathan Baxter, JMLR, 2004](https://github.com/dimitarpg13/reinforcement_learning_and_game_theory/blob/main/articles/ReinforcementLearning/Variance_Reduction_Techniques_for_Gradient_Estimates_in_Reinforcement_Learning_Greensmith_2004.pdf)

[2] [Infinite-Horizon Policy-Gradient Estimation, J. Baxter, P. Bartlett, 2001](https://github.com/dimitarpg13/reinforcement_learning_and_game_theory/blob/main/articles/ReinforcementLearning/Infinite-Horizon_Policy-Gradient_Estimation_Baxter_2001.pdf)