Feature Selection by Singular Value Decomposition for Reinforcement Learning

TA: Aleksey Grinchuk

Dmitry Krylov

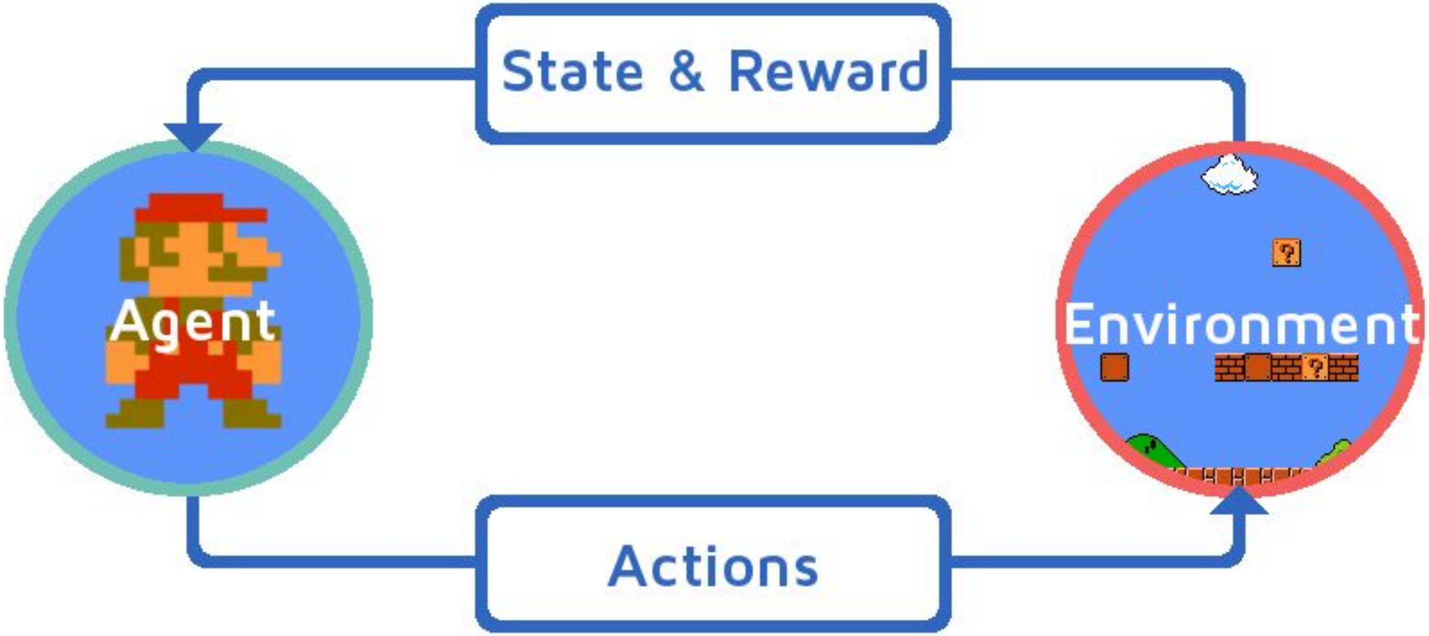
Nikita Kotelevskii

Ksenia Yagafarova

Anna Anikina

Introduction

Reinforcement learning with a large state space



Problem statement

- Choosing the correct number of features

  How to choose appropriate number of features

- Lack of theoretical explanation of LFD

  Why LFD works and how to explain it

- Time-consuming performance

  Train Estimator with LSPI

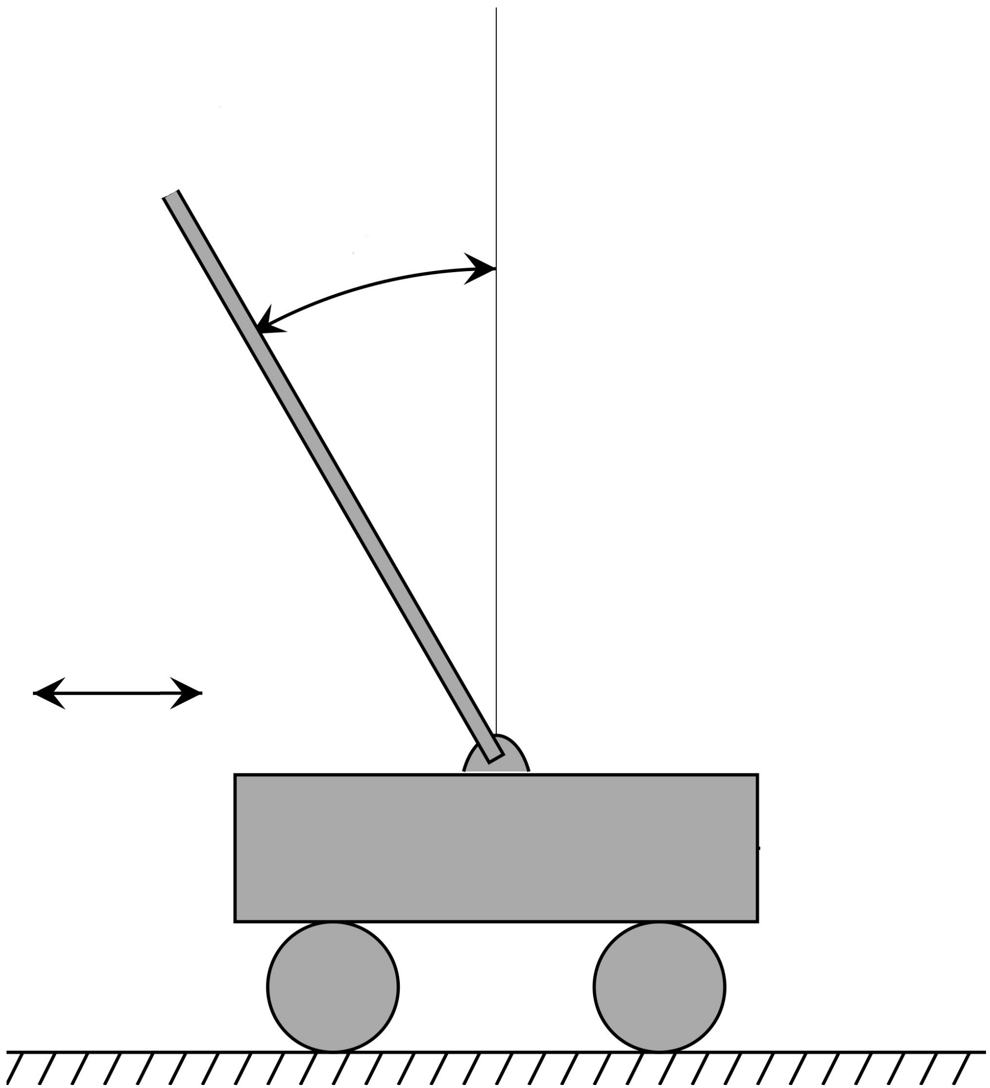
- Assumption about transition matrix

  Low-rank Transition Matrix

Main ideas from paper (Behzadian, Petrik 2018)

- LFD is a low-rank approximation to the matrix of transition probabilities
- We can obtain good feature encoder using SVD

Formulation of the problem



Methodology

Markov Decision Process:
S - states, A - actions, P - transition probabilities matrix, r - rewards

Value function $$\boldsymbol{v}^{\pi} = \boldsymbol{r}^{\pi} + \gamma P^{\pi} \boldsymbol{v}^{\pi}$$

Linear value function $$\tilde{\boldsymbol{v}}^{\pi} = \Phi \, \boldsymbol{w},$$

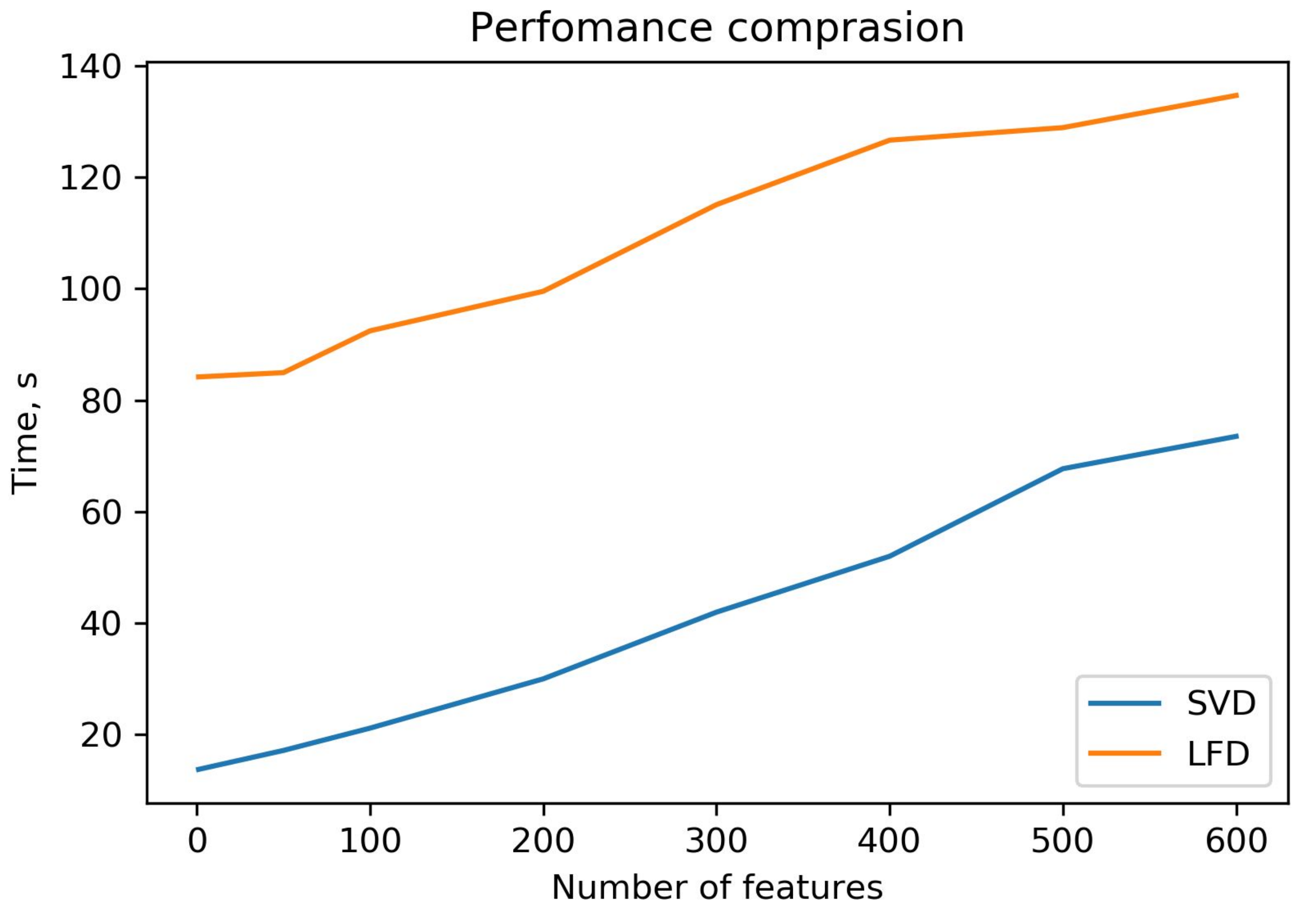
$\boldsymbol{\Phi}$ is the feature matrix of dimensions |S| × k ; the columns of this matrix are the features ϕ_i

$$P_{\Phi}^{\pi} = (\Phi^{\top}\Phi)^{-1}\Phi^{\top}P^{\pi}\Phi \qquad \boldsymbol{r}_{\Phi}^{\pi} = (\Phi^{\top}\Phi)^{-1}\Phi^{\top}\boldsymbol{r}^{\pi}$$

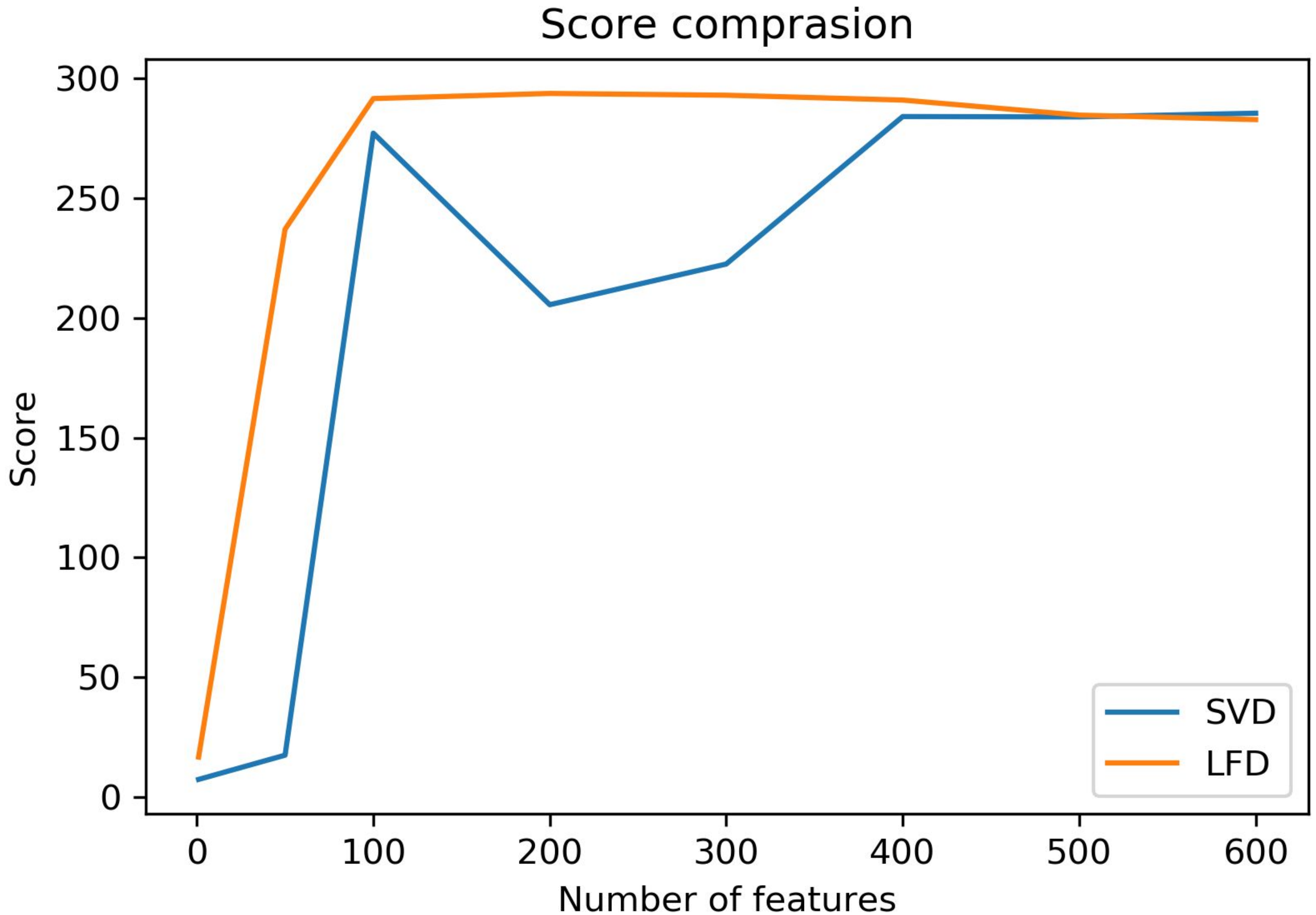
$$\boldsymbol{w}_{\Phi}^{\pi} = \boldsymbol{r}_{\Phi}^{\pi} + \gamma P_{\Phi}^{\pi} \boldsymbol{w}_{\Phi}^{\pi}$$

$$P_{A}^{\pi} = \sum_{i=1}^{m} \sigma_i u_i v_i^{\top} = U \Sigma V^{\top} \qquad \begin{aligned} E^{\pi} &= U_1, \\ D^{\pi} &= \Sigma_1 V_1^{\top} \end{aligned}$$

Results obtained



Results obtained



Conclusion

- No need to know number of features to select in advance
- Easier to implement
- Less-time consuming
- Better performance

References

- Song, Z.; Parr, R. E.; Liao, X.; and Carin, L. 2016. Linear feature encoding for reinforcement learning. In Advances in Neural Information Processing Systems, 4224–4232.
- Behzadian, B. and Petrik M. 2018. Low-rank feature selection for reinforcement learning, ISAIM (2018).
- Eckart, G., and Young, G. 1936. The approximation of one matrix by another of lower rank. Psychometrika 1:211218.
- Hrinchuk O. 2017. Feature space perspectives for Reinforcement learning.

Thank you for your attention!

Questions?