LSTD with Random Projections

Mohammad Ghavamzadeh

joint work with Alessandro Lazaric, Odalric Maillard, Rémi Munos

INRIA Lille – Nord Europe, Team SequeL

Poster Number: T59

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE

INRIA

centre de recherche
LILLE - NORD EUROPE

Problem

Setting: RL in *high-dimensional spaces* — the number of features is bigger than the number of samples.

Problem: *Overfitting* and poor *prediction* performance.

Solutions

- **Regularization:** adding ℓ_1 and ℓ_2 regularizations to value-function approximation algorithms.
- **Random projections: this work.**

LSTD: A RL alg. for learning the value function of a given policy.

Algorithm – LSTD-RP: LSTD in a *low-dim* space generated by a **random projection** from the *high-dim* space.

Results

- We present the LSTD-RP algorithm and discuss its computational complexity.

less expensive than performing LSTD in the high-dim space.

- We provide finite-sample performance bounds for LSTD-RP.
  - **Markov design:** performance on the training samples.
  - **Random design:** generalization over the entire state space

avoid overfitting – better prediction: smaller **estimation error** *and not much larger* **approximation error**.

Results

- **Uniqueness** of the LSTD-RP solution.
  - When the *model-based* LSTD solution in *high-dim* exists,
  - How many *samples* are needed in order for LSTD-RP (*low-dim* random space) to have a unique solution w.h.p.?

LSTD-RP is more **stable** *and needs* **less samples** *to have a unique solution, compared to LSTD in the high-dim space*.

- We show how the error of LSTD-RP is propagated through the iterations of LSPI and present a bound for LSPI-RP.