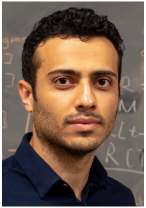
Learning the model-free LQR via random search

Mihailo Jovanović

viterbi-web.usc.edu/∼mihailo/



Hesameddin Mohammadi   Mahdi Soltanolkotabi

Learning for Dynamics and Control; June 11, 2020

Linear Quadratic Regulator problem

$$\begin{aligned} &\text{minimize} && \text{quadratic cost} \\ &\text{subject to} && \text{linear dynamics} \end{aligned}$$

Linear Quadratic Regulator problem

$$\begin{array}{ll} \text{minimize} & \text{quadratic cost} \\ \text{subject to} & \text{linear dynamics} \end{array}$$

LTI dynamics: $$\frac{\mathrm{d}x(t)}{\mathrm{d}t} \;=\; A\,x(t) \,+\, B\,u(t), \;\; x(0) \;=\; x_0$$

quadratic cost: $$\mathbb{E}_{x_0}\left[\int_0^\infty \left(x^T(t)\,Q\,x(t) \,+\, u^T(t)\,R\,u(t)\right)\mathrm{d}t\right]$$

$Q,\, R \succ 0$ – *state and control weights*

- OPTIMAL SOLUTION

$$u(t) \;=\; -K^{\star}x(t)$$

Riccati-based characterization

$$\begin{aligned} A^T P \;+\; PA \;-\; P\,B\,R^{-1}B^T P \;+\; Q &\;=\; 0 \\ K^{\star} &\;=\; R^{-1}B^T P \end{aligned}$$

Model-free reinforcement learning

- LEARN FEEDBACK POLICY BY OPTIMIZING OVER K

$$\underset{K}{\text{minimize}} \quad f(K)$$

- QUESTIONS

  - convergence properties (lack of convexity)
  - statistical properties (lack of models)

Our contribution

- RANDOM SEARCH METHOD FOR LQR

convergence and sample complexity

linear convergence:	over the set of stabilizing K
logarithmic complexity:	wrt the reciprocal of accuracy

with high probability

Linear convergence of gradient descent

SMOOTHNESS

$$K \in \mathcal{S}_a \;\;\Rightarrow\;\; \|\nabla^2 f(K)\|_2 \;\leq\; L$$

GRADIENT DOMINANCE

$$K \in \mathcal{S}_a \;\;\Rightarrow\;\; \|\nabla f(K)\|_F^2 \;\geq\; \mu\,(f(K) \,-\, f(K^\star))$$

- SUB-LEVEL SET OF THE LQR COST

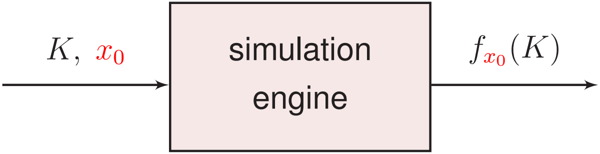
$$\mathcal{S}_a \;:=\; \{\text{stabilizing } K \mid f(K) \leq a\}$$

Fazel, Ge, Kakade, Mesbahi, ICML '18

Mohammadi, Zare, Soltanolkotabi, Jovanović, IEEE CDC '19

Model-free setup

- GRADIENTS NOT AVAILABLE
  - ⋆ can only access random function values



  - ⋆ objective function

$$f(K) \;=\; \mathbb{E}_{x_0}[f_{x_0}(K)]$$

x_0 – *sub-Gaussian*

Random search for model-free LQR

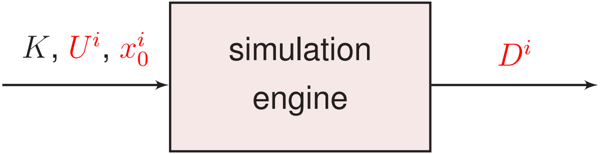
- EMULATE GRADIENT DESCENT

$$K^+ = K - \alpha D$$

  - gradient estimation using random function values

$$D = \begin{cases} \dfrac{f_{x_0}(K + rU) - f_{x_0}(K - rU)}{2r} \, U & \text{two-point} \\ \\ \dfrac{f_{x_0}(K + rU) - f_{\hat{x}_0}(K - rU)}{2r} \, U & \text{one-point} \end{cases}$$

U – *random matrices*



⋆ mini-batch gradient estimate

$$D = \frac{1}{N} \sum_{i=1}^{N} D^i$$

N – *number of samples*

Main result: linear rate; log complexity

- THEOREM

  Random search LQR achieves a desired accuracy ϵ, i.e.,

$$f(K^T) \; - \; f(K^\star) \; \leq \; \epsilon$$

with high probability, if

$$\textbf{number of iterations} \quad T \; = \; O\left(\log\left(1/\epsilon\right)\right)$$

$$\textbf{number of samples (per iteration)} \quad N \; = \; c\, n \left(\log n\right)^6$$

n – number of states

two-point estimate

Mohammadi, Zare, Soltanolkotabi, Jovanović, arXiv:1912.11899

State-of-the-art for random search LQR

polynomial sample complexity

sub-linear convergence rate

decreasing stepsize

for ϵ-accuracy	*one-point*	*two-point*
iterations T	$O(\epsilon^{-2})$	$O(\epsilon^{-1})$
samples N	1	1
stepsize α	$O(\epsilon^2)$	$O(\epsilon)$

Malik, Pananjady, Bhatia, Khamaru, Bartlett, Wainwright, JMLR '20

- ONE-POINT MINI-BATCH ESTIMATE
  - ⋆ polynomial complexity, linear rate, fixed stepsize

Fazel, Ge, Kakade, Mesbahi, ICML ’18

- ONE-POINT MINI-BATCH ESTIMATE
  - ⋆ polynomial complexity, linear rate, fixed stepsize

Fazel, Ge, Kakade, Mesbahi, ICML '18

- OUR APPROACH
  - ⋆ DON'T control gradient estimation error

key in our proof

gradient estimate concentrates with high probability

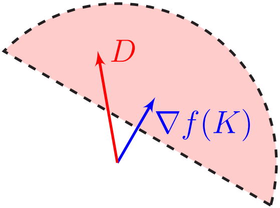
(when projected to the direction of the gradient)

two-point estimate

Approximate gradient descent

$$\langle D, \nabla f(K) \rangle \;\geq\; \theta_1 \, \|\nabla f(K)\|_F^2$$

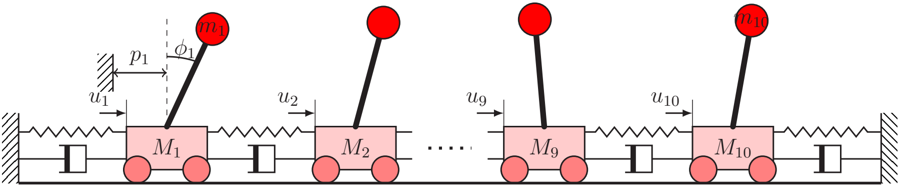
$$\|D\|_F^2 \;\leq\; \theta_2 \, \|\nabla f(K)\|_F^2$$



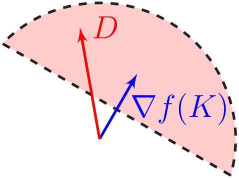
Linear convergence with fixed stepsize

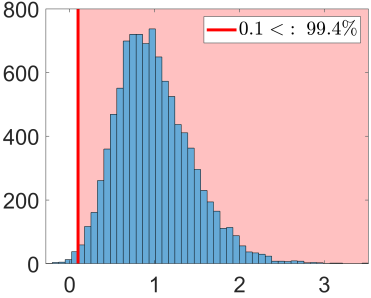
$$f(K^k) \, - \, f(K^\star) \;\leq\; \left(1 \, - \, \frac{\theta_1^2 \, \mu}{\theta_2 \, L}\right)^k \big(f(K^0) \, - \, f(K^\star)\big)$$

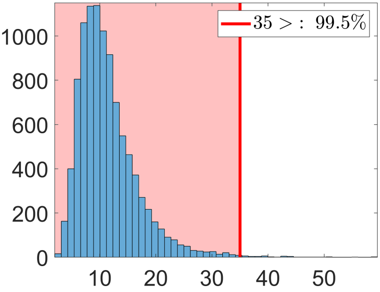
An example



Occurrence with high probability



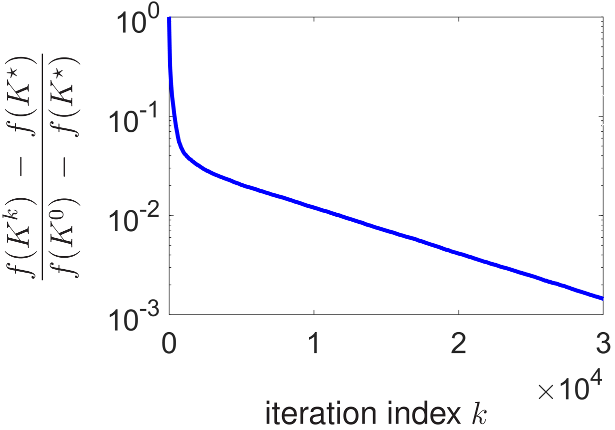




$$\frac{\langle D, \nabla f(K)\rangle}{\|\nabla f(K)\|_F^2} \geq \underbrace{0.1}_{\theta_1}$$

$$\frac{\|D\|_F^2}{\|\nabla f(K)\|_F^2} \leq \underbrace{35}_{\theta_2}$$

Linear convergence

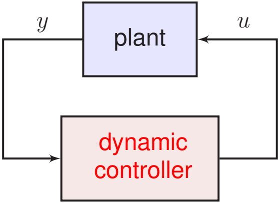


Ongoing effort

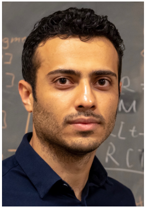
- Sparsity-promoting RL

$$\begin{bmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \\ u_4(t) \\ u_5(t) \end{bmatrix} = -\underbrace{\begin{bmatrix} * & * & & & \\ * & * & & & * \\ & * & * & * & \\ & & * & * & \\ & * & & * & * \end{bmatrix}}_{K} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \\ x_5(t) \end{bmatrix}$$

- Output-feedback RL



Acknowledgements



Hesam

Mahdi

- SUPPORT
  - ⋆ AFOSR FA9550-16-1-0009 (Program manager: Fred Leve)
  - ⋆ NSF ECCS-1809833 (Program manager: Kishan Baheti)