Master project

Controlling the unknown

end-to-end performance

Peyman Mohajerin Esfahani, Delft Center for Systems and Control, TU Delft
P.MohajerinEsfahani@tudelft.nl

Context In practise, systems become larger and more complex which makes modelling from first principles difficult and often intractable. ***Reinforcement learning*** and more classically, ***system identification*** and ***dynamic programming***, address these problems by supplying us with techniques to estimate and optimally control unknown dynamical systems.

There is one catch, the control laws are often designed as if the estimated model is the real model (*certainty equivalence*), which is rather unlikely. The question is, how optimal are these control laws based on identified models generated by ***finite data***? In other words, can we find theoretical bounds which relate data quality and quantity to controller performance? Moreover, can this performance be improved by ***adapting*** the control law and estimation scheme?

As a real world example consider the field of medicine. Models and time are scarce, but data is not. The overarching goal is to find optimal inputs which for example make some bacteria vanish in finite time, but such an input will of course only be applied if it is known with high probability that it will work.

Project goal—robust data-driven control To begin this project we start with linear dynamical systems and consider the linear-quadratic (LQ) infinite horizon optimal control problem, where the dynamics are however unknown. The reason is twofold, this LQ problem is well-studied and has a closed-form solution when the dynamics are known, but it is also notorious for being a linear optimal control problem for which self-tuning does not hold. This lack of self-tuning can be mainly attributed to the fact that the usual closed-loop matrix $(A + BK)$ does not have an unique representation in terms of A and B. Therefore it is widely believed that part of the solution should entail input design such that initially, or throughout, the system is perturbed in such a way that the open-loop system can be identified while the closed-loop system remains stable with high probability, see Figure 1.

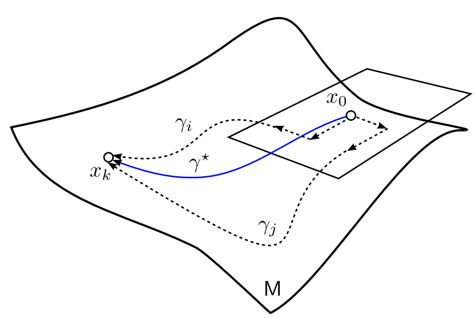

Figure 1: *Let the real, for us unknown, system be given by $x_{k+1} = A^\star x_k + B^\star u_k + M^\star v_k$, $v_k \in G(0, Q_v)$. Then with respect to the discounted cost $J = \mathbb{E}\left[\sum_{k=0}^{\infty} \gamma^k (x_k^\top Q x_k + u_k^\top R u_k) \mid \mathcal{F}^{x_0}\right]$ the optimal trajectory from x_0 to x_k is given by $\gamma^\star$. Now without knowing the dynamics, we would still like to steer x_0 to x_k. Say we have two strategies: i and j, then which trajectory, γ_i or γ_j is preferred? The cost in x might be lower for γ_i, but perhaps the initial input sequence is much more costly and statistically less safe.*

This problem is far from new, what would be new, are tractable and exact methods to attack such a problem, together with sharp theoretical performance bounds.

tags Stochastic Optimal Control, Robust Optimization, Linear Dynamical Systems.