Learned system dynamics for adaptive optimal feedback control

Djordje Mitrovic, Stefan Klanke, and Sethu Vijayakumar
{d.mitrovic, s.klanke, sethu.vijayakumar}@ed.ac.uk
School of Informatics, University of Edinburgh, United Kingdom

Optimal feedback control (OFC) has been proposed as an attractive movement generation strategy in goal reaching tasks for anthropomorphic manipulator systems. In contrast to classic open loop optimizers that produce "just" a minimal-cost trajectory with implicit resolution of kinematic and dynamic redundancies, the OFC framework additionally yields a feedback control law which corrects errors *only* if they adversely affect the task performance (minimum intervention principle).

Locally, the optimal feedback control law for systems with non-linear dynamics and non-quadratic costs can be found by iterative methods, such as the recently introduced iterative Linear Quadratic Gaussian (iLQG) algorithm [1]. So far this framework relied on an analytic form of the system dynamics, which may often be unknown, difficult to estimate for more realistic control systems or may be subject to frequent systematic changes.

We present a novel combination of learning a forward dynamics model within the iLQG framework, for which we employ Locally Weighted Projection Regression (LWPR) [2]. Utilising such adaptive internal models can compensate for complex dynamic perturbations of the controlled system in an online fashion. Moreover, through the availability of analytic derivatives of the learned model, the adaptive iLQG–LD framework we introduce lends itself to a computationally more efficient implementation of the iLQG optimization without sacrificing control accuracy, allowing the method to scale to large DoF systems.

Up to now, we studied iLQG–LD on two different joint torque controlled manipulators as simulated by the Matlab Robotics Toolbox: i) a planar 2 DoF manipulator, which is ideal for performing extensive (quantitative) comparison studies and to test the manipulator under controlled perturbations and force fields during planar motion, and ii) a 6 DoF manipulator with realistic physical parameters. We successfully tested our iLQG–LD framework with both stationary and non-stationary dynamics, simulating constant or velocity-dependent force fields.

Our current work concentrates on implementing the iLQG–LD framework on the 7-DOF robot arm hardware of the German Aerospace Centre (DLR), which raises challenges such as a very high-dimensional input space (7 commands + 14 motor states + 14 joint states) and a high sampling rate of 1kHz. In the future, we will aim for the biomorphic *variable stiffness* based highly redundant actuation system that is currently developed at DLR – this will not only explore an alternative control paradigm, but will also provide the only viable and principled control strategy for such a system.

References

[1] E. Todorov and W. Li. A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems. In *Proc. of the American Control Conference*, 2005.

[2] S. Vijayakumar, A. D'Souza, and S. Schaal. Incremental online learning in high dimensions. *Neural Computation*, 17:2602–2634, 2005.