Nonlinear adaptive hybrid control by combining Gaussian process system identification with classical control laws

Jan-P. Calliess, Mike Osborne and Stephen J. Roberts

Introduction: In recent years there has been much research into the use of Gaussian processes (GPs) for online identification of value functions and discrete-time transitions and their use in adaptive model-predictive control. These methods typically require significant computation for policy evaluation. This limits their applicability in settings where robust continuous control signal is required. Furthermore, there is little understanding of modelling the continuous system closed-loop dynamics resulting from application of control at discrete times. On the other hand, classical control theory provides the necessary machinery to derive continuous closed-form control policies in settings where the underlying plant dynamics are sufficiently simple and known. Unfortunately, in many situations these dynamics are too complex, poorly known or unknown.

In this work, we propose a Gaussian process to learn the non-linearity of the uncontrolled state-evolution. To enable the controlled plant to exhibit desirable target behaviour, the control is set to the negative prediction output (to eliminate the non-linearity) plus a feedback term of our choice enabling desirable target behaviour. As a starting point, we focus on adaptive control where we wish to drive the state towards a reference signal. The feedback term could be implemented in analogue hardware (e.g. a PID circuit) whereas the GP output might come from a digital device connected to a sample-hold circuit. Open questions under investigation are whether it is possible to choose the feedback term in closed form, such that the overall control law incurs low control cost.

Method: We consider a plant with dynamics[1] $\dot{x} = a(x) + u(x)$ and require it to follow a prescribed reference trajectory $\xi : I = [t_0, \infty) \to \mathbb{R}^D$ in D-dimensional state space. We assume the non-linear function $a : \mathbb{R}^D \to \mathbb{R}^D$ is unknown at start time t_0 but that we can obtain noisy samples from it from time to time (e.g. by observing uncontrolled velocities).

Our approach is to set $u(x) := -\hat{a}(x) + \phi(x; w, \xi)$. Here, $\hat{a}$ is the posterior mean of a Gaussian process conditioned on the observations; $\phi(x; w, \xi)$ is a fixed feedback control law (with parameter w) designed to drive the state towards reference ξ.

The resulting closed-loop dynamics can be modelled as the SDE $\dot{x} = \phi(x; w, \xi) + \nu(x)$ where $\nu(x)$ captures the prediction uncertainty. Hence, in explored regions of state space the dynamics is could be modelled as an ODE $\dot{x} = \phi(x; w, \xi)$.

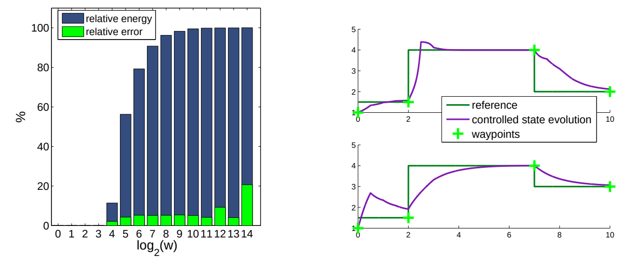

Fig. 1. **Left**: Comparison our adaptive control law with fixed proportional control law $u(x) = w(\xi - x)$ as a function of proportional gain factor w. Bars depict the percentage of the control error of our method compared to the fixed controller (green) and the percentage of control energy expended (blue). Our GP method outperforms the fixed gain controller both in terms of control energy and error. Notice, for $w = 2^0, ..., 2^3$ the fixed proportional controller was unable to stabilize the system whereas our adpative control succeeded. An example for $w = 1$ is depicted **right**: GP adaptive control with two nonlinear dynamics vs. time [sec]. (top-right: quadratic, bottom-left: sinusoidal). Training examples were added online every 0.5 seconds.

The choice of control law ϕ may depend on the control objective. If it merely is to drive the state to ξ, we could consider defining $\phi(x; w, \xi) := w(\xi - x)$ where $w > 0$. If a is bounded, we can guarantee stability. If not, we can establish convergence success (in expectation) if a is drawn from our GP prior. Even if this assumption is violated we may be able to stabilize the system sufficiently. As an illustration, consider Fig. 1 (right). Starting with an unknown function $a(x) = (\sin(x_1) + 3, x_2^2 - 1)^T$, our adaptive control rule learns to drive the prescribed reference signal to follow the reference. We see that our combined adaptive method outperforms a static linear feedback controller both in terms of control energy and control success (Fig. 1 (left)).

Ongoing work: Ongoing work investigates suitable choices of feedback component ϕ in the presence of control cost functions. In optimal control, we wish to find a control solution to the variational problem $\min_u J(u)$, st:$[\dot{x} = a(x) + u(x)]$, where J is a cost functional chosen to penalize the control error as well as control energy. Our current work explores the choice of ϕ such that our adaptive control mechanism yields low overall cost whenever our GP model has made enough observations. As an example we presently investigate the performance of setting ϕ to the optimal feedback law of $\min_u J(u)$, st:$[\dot{x} = u(x)]$ (which, due to the simplicity of the learned dynamics, $\dot{x} = u(x)$) can invariably be found via the Hamilton-Jacobi-Bellman equation.

The authours gratefully acknowledge funds via EPSRC EP/I011587.

The authors are with the Machine Learning Research Group, Information Engineering, University of Oxford, UK.

[1]Our method can be easily generalized to work on $\dot{x} = a(x) + Bu(x)$ where B is a matrix.