Institute for Design and Control of Mechatronical Systems

Diploma thesis

Approximation of Optimal Control of Continuous Control Affine Planar Nonlinear System

Author	**Thomas Stanger**
Supervisors	**Prof. Dr. Luigi del Re**
	DI Thomas Passenbrunner
Finished	**October 2011**

Abstract

For nonlinear systems the solution of the optimal control problem is given by the solution of the Hamilton Jacobi Bellman Equation which has no general solution. The method proposed in this thesis obtains a solution by successive approximation due to the solution of the Generalized Hamilton Jacobi Bellman Equation. Successive improvement of the control law leads to an approximation of the optimal control, which is optimal in a bounded region around the origin. Application of Policy Iteration to an example, an instable, nonlinear, inverse pendulum will demonstrate the capabilities of the whole approach.

Two different implementations of Policy Iteration have been applied to this example. One uses simulation to approximate the solution of the Generalized Hamilton Jacobi Bellman Equation and the other one is based on a numerical solution. While the first realization requires a long runtime, but only little theoretical knowledge, the second one is much faster. For this example the improvement of the control law gained by this approach is up to 30% with respect to the LQR.

Introduction

Optimal control of dynamic systems is perhaps one of the most challenging tasks in modern control engineering. Although a lot of research has been done, a general solution to this problem has not been found yet. As a consequence optimal control is often reduced to the optimal control of linearized systems and nonlinear optimal control remains an attractive discipline.

For linear systems the optimal control problem is solved as soon as a solution to the *Continuous Algebraic Riccati Equation* (CARE) is found. For the nonlinear counterpart, the *Hamilton Jacobi Bellman Equation* (HJB-equation), a quadratic, partial, first order differential equation, no general solution exists. Since the solution of the CARE is the first order approximation of the solution to the HJB-equation attempts have been made to obtain higher order approximation. As the resulting *Albrecht's* Method is not guaranteed to converge for a larger region around the origin, piecewise extensions of the solution have been investigated by Navasca.

In parallel to the optimal control problem, the related *Inverse Optimal Control Problems* have been studied. A simple converse approach is also part of this thesis. A parameterized value function will be optimized such that the solved optimal control problem is approximately the considered optimal control problem. Because of the nonlinear nature of the HJB-equation nonlinear optimization is required. This restricts the optimizable number of parameters and therefore also the complexity of the value function.

All these approaches either give an unrewarding approximation or are very complex and difficult to understand. Consequently optimal nonlinear control is still very unpopular in engineering. With *Policy Iteration* (PI) a largely unfamiliar approach to the optimal control problem of planar, control affine systems is given. This method is an intuitive and simple applicable procedure and yields a numerical approximation of the optimal control policy within a bounded region around the origin.

The optimal control problem

Given a control affine system

$$\dot{x} = f(x) + g(x)u$$

and quadratic cost function

$$J(x(t), u(t)) = \int_0^\infty \left(x(t)^T Q x(t) + u(t)^T R u(t) \right) dt$$

with $Q > 0$ and $R > 0$, the optimal control problem is given by

$$\min_u \int_0^\infty J(x(t), u(t))\, dt$$
$$\text{s.t.} \quad \dot{x}(t) = f(x(t)) + g(x(t))u(t)$$
$$x(0) = x_0$$

The minimal costs $V^*(x_0)$ for any initial state x_0 are given by the solution $V^*(x) > 0$ of the *Hamilton Jacobi Bellman Equation*, a quadratic, partial, first order differential equation

$$x^T Q x \; - \; \frac{1}{4} \frac{\partial V^*(x)}{\partial x} g(x) R^{-1} g(x)^T \left(\frac{\partial V^*(x)}{\partial x} \right)^T \; + \; \frac{\partial V^*(x)}{\partial x} f(x) \; = \; 0$$

with $V^*(0) = 0$ being an essential condition. The HJB also features a second solution $V^*(x) < 0$ which is not of interest to optimal control. The optimal control which achieves the optimal costs along the optimal trajectory is related to the optimal costs by

$$u^* \; = \; -\frac{1}{2} R^{-1} g(x)^T \left(\frac{\partial V^*(x)}{\partial x} \right)^T$$

The HJB is difficult to solve and a general solution only exist for linear systems where the HJB becomes the CARE.

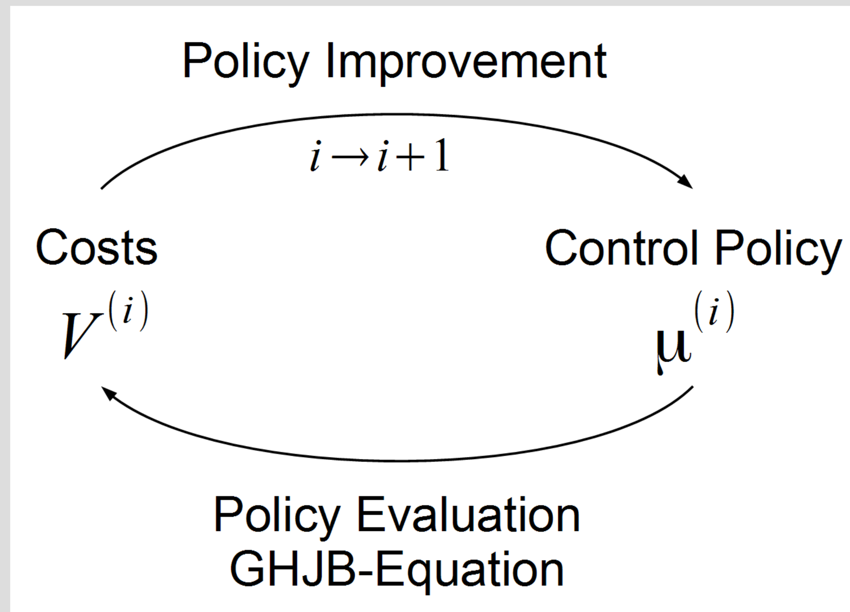
$$PA \; + \; A^T P \; - \; PBR^{-1} B^T P \; + \; Q \; = \; 0$$

with $V^*(x) = x^T P x > 0$ and $\frac{\partial V^*(x)}{\partial x} = 2 x^T P$.

An iterative solution to the HJB and two different implementations for general, control affine, nonlinear systems is proposed. This iterative solution is based on *Policy Iteration* and the approximative solution of the so called *Generalized Hamilton Jacobi Bellman Equation.*

Policy Iteration

Once having obtained an admissible control, e.g. the LQR of the linearized system, this control can be improved by *Policy Iteration*. *Policy Iteration* can be divided into *Policy Evaluation* and *Policy Improvement*. Given an admissible control $u = \mu^{(i)}$ the related costs $V^{(i)}$ are determined by *Policy Evaluation*. By *Policy Improvement* an improved control policy is derived from $V^{(i)}$. By iteratively performing *Policy Evaluation* and *Policy Improvement* the series of $V^{(i)}$ converges to the optimal costs V^* thus $\mu^{(i)}$ must converge to the optimal control μ^*.



The related costs $V^{(i)}(x)$ satisfy

$$\frac{dV^{(i)}(x)}{dt} \; = \; \frac{\partial V^{(i)}(x)}{\partial x} \left(f(x) + g(x) \mu^{(i)}(x) \right)$$

As a consequence, the costs are given by the solution of the *Generalized Hamilton Jacobi Bellman Equation*, a linear, partial, first order differential equation.

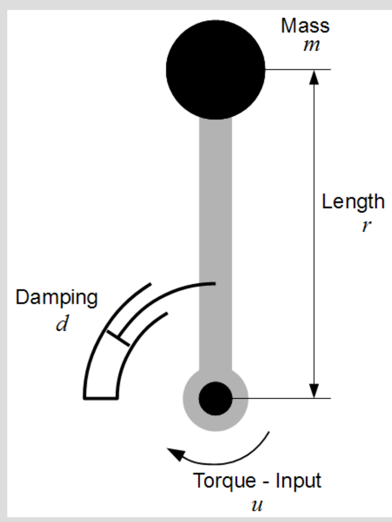
$$\frac{\partial V^{(i)}(x)}{\partial x} \left(f(x) + g(x) \mu^{(i)}(x) \right) \; + \; x^T Q x \; + \; \mu^{(i)T} R \mu^{(i)} \; = \; 0$$

Given the costs $V^{(i)}(x)$ caused by an applied control the improved control is given by

$$\mu^{(i+1)} \; = \; -\frac{1}{2} R^{-1} g(x)^T \left(\frac{\partial V^{(i)}(x)}{\partial x} \right)^T$$

Since an algebraic solution for the *GHJB-equation* is hardly obtainable a numerical approximation is used. Therefore $V^{(i)}(x)$ is approximated only for some discrete states and the gradient is approximated by finite difference quotient. The values of $V^{(i)}(x)$ can be determined either by integration of the *GHJB-equation* along the trajectory of the system controlled by $\mu^{(i)}(x)$ (simulation) or a finite difference approximation of the solution.

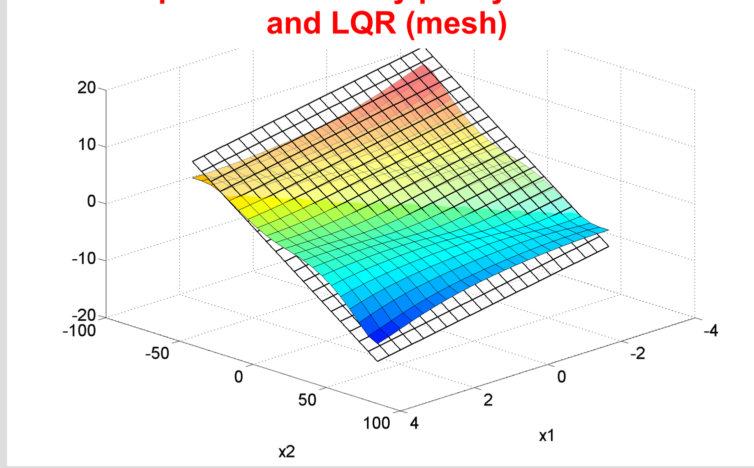
Example: Inverse Pendulum



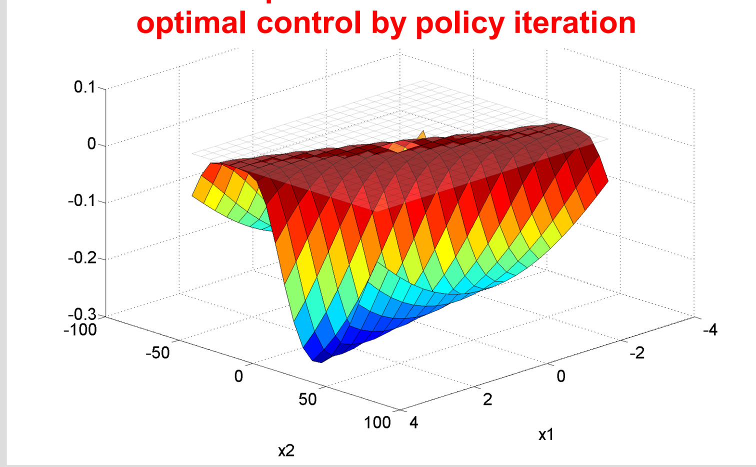
The plant is an inverse pendulum with damping d and the input torque u directly acting at the hinge.

A comparison of the LQR and an approximation of the optimal control $\mu^{(7)}(x)$ after seven policy iterations as shown in the right Figure.

Optimal control by policy iteration and LQR (mesh)



Relative improvement w.r.t. the LQR of optimal control by policy iteration



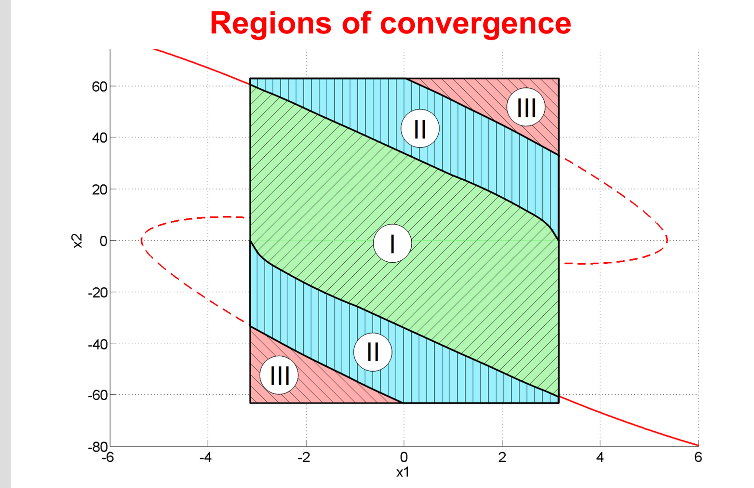
The performance of the nonlinear control and the initial control is compared by means of

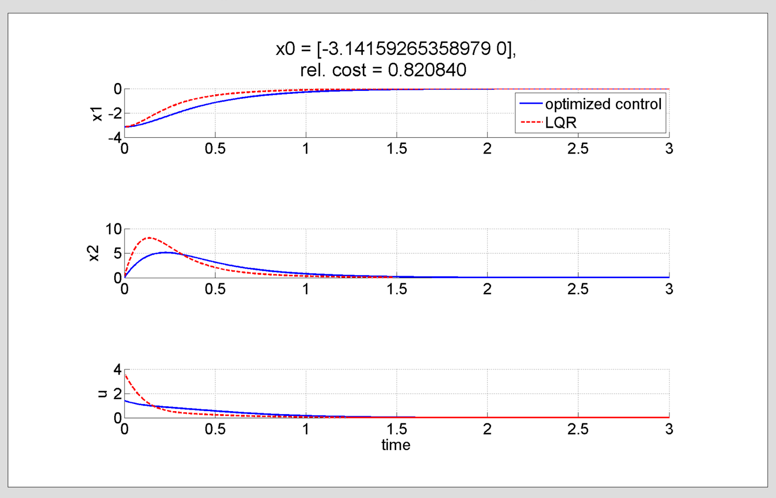
$$\frac{V^{(7)}(x) - V^{(0)}(x)}{V^{(0)}(x)}$$

The Figure on the left-hand side shows, an improvement of up to 30%. Near the origin no improvement with respect to the LQR is possible since the LQR is the local optimal control at the origin.

Policy Iteration has been evaluated by simulation on a discrete grid within the region $[-\pi \ldots \pi] \times [-20\pi \ldots 20\pi]$ and the final control is optimal within the greatest forward invariant region I. By extrapolation $\mu^{(7)}(x)$ becomes almost optimal for region II but not for region III.

Regions of convergence





Some signal curves, comparing the action of the LQR and $\mu^{(7)}(x)$, moving the pendulum from the lower stable to the upper unstable equilibrium, are shown on the left side. While the optimal control is more conservative than the LQR, the integral costs are 18% less.

Conclusion and Outlook

Optimal control of nonlinear systems is not a simple task but also with quit simple nonlinear approaches significant improvements can be achieved. Policy iteration is a very general principle, therefore it admits a wide variety of procedures. In addition to the proposed methods, using simulations or numerical solutions of the *GHJB-equation*, combinations are possible. Thus the advantages of both methods could be combined and give an even more powerful design framework. Policy Iteration using simulations can be performed with little mathematical effort. The major drawback of this method is the long runtime of the simulations and the consequential limitations according to the accuracy. When performing policy iteration by the numerical solution of the *GHJB-equation* more accurate controls are obtained in less time but the mathematical effort is higher.

A restriction to optimal control problems, exclusively optimizing a pure quadratic cost function is not essential. Utilizing generalized, non quadratic cost function will give the possibility of considering input and state boundaries by barrier or penalty functions. Further *Policy Iteration* is also applicable to systems with multiple inputs and even the restriction to planar systems is not essential. While the application to systems with more than one input is straight forward, a higher number of state variables will be more challenging.