

NEURAL NETWORK SOLUTION FOR FIXED-FINAL TIME OPTIMAL
CONTROL OF NONLINEAR SYSTEMS

by

Tao Cheng

Presented to the Faculty of the Graduate School of
The University of Texas at Arlington in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

December 2006

ACKNOWLEDGEMENTS

I would first like to acknowledge the assistance and guidance, over the past four years, of my thesis advisor Professor Frank L. Lewis. In particular I am grateful that Professor Lewis allowed me the flexibility to search for a problem in which I was really interested. I am also indebted to Professor Lewis for his advice to pursue a curriculum loaded with mathematics and theory. I feel that the quality of my graduate education has been greatly enriched by this emphasis.

I would like to thank Dr. Murad Abu-Khalaf, his concrete theory background really impressed me, he gave me many helpful discussions, questions and critiques concerning my thesis work. I would also like to thank Professor Korzeniowski for his unselfish help in my math courses.

I would like to thank my colleagues at the Automation & Robotics Research Institute in particular those I worked with at the Advanced Controls & Sensors group.

This work was funded by the National Science Foundation ECS-0501451 and by the Army Research Office W91NF-05-1-0314 grant.

November 20, 2006

ABSTRACT

NEURAL NETWORK SOLUTION FOR FIXED-FINAL TIME OPTIMAL CONTROL OF NONLINEAR SYSTEMS

Publication No. _____

Tao Cheng

The University of Texas at Arlington, 2006

Supervising Professor: Frank L. Lewis

In this research, practical methods for the design of H_2 and H_∞ optimal state feedback controllers for unconstrained and constrained input systems are proposed. The dynamic programming principle is used along with special quasi-norms to derive the structure of both the saturated H_2 and H_∞ optimal controllers in feedback strategy form. The resulting Hamilton-Jacobi-Bellman (HJB) and Hamilton-Jacobi-Isaacs (HJI) equations are derived respectively.

Neural networks are used along with the least-squares method to solve the Hamilton-Jacobi differential equations in the H_2 case, and the cost and disturbance in the H_∞ case. The result is a neural network unconstrained or constrained feedback controller that has been tuned a priori offline with the training set selected using Monte

Carlo methods from a prescribed region of the state space which falls within the region of asymptotic stability.

The obtained algorithms are applied to different examples including the linear system, chained form nonholonomic system, and Nonlinear Benchmark Problem to reveal the power of the proposed method.

Finally, a certain time-folding method is applied to solve optimal control problem on chained form nonholonomic systems with above obtained algorithms. The result shows the approach can effectively provide controls for nonholonomic systems.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	ii
ABSTRACT	iii
LIST OF ILLUSTRATIONS.....	ix
NOMENCLATURE	xi
Chapter	
1 INTRODUCTION	1
1.1 Significance and Contribution of the Research.....	1
1.2 Approach	3
1.2.1 H_2 Optimal Control: Hamilton-Jacobi-Bellman (HJB) equation	4
1.2.2 H_∞ Optimal Control: Hamilton-Jacobi-Isaacs (HJI) equation	4
2 FIXED-FINAL TIME OPTIMAL CONTROL OF NONLINEAR SYSTEMS USING NEURAL NETWORK HJB APPROACH	6
2.1 Introduction	6
2.2 Background on Fixed-Final-Time HJB Optimal Control	7
2.3 Nonlinear Fixed-Final-Time HJB Solution by NN Least-Squares Approximation	10
2.3.1 NN Approximation of the Cost Function $V(x,t)$	10
2.3.2 Uniform Convergence in t For Time-Varying Function of the Method of Least-Squares	14
2.3.3 Optimal Algorithm Based on NN Approximation.....	28

2.3.4	Numerical Examples	29
2.3.4.1	Linear System	30
2.3.4.2	Nonlinear Chained System	32
2.4	Conclusion	35
3	NEURAL NETWORK SOLUTION FOR FINITE-FINAL TIME H-INFINITY STATE FEEDBACK CONTROL	37
3.1	Introduction	37
3.2	L_2 -gain and Dissipativity of Controlled Nonlinear Systems	37
3.3	NN Least-Squares Approximate HJI Solution	42
3.3.1	NN Approximation of $V(x)$	42
3.3.2	Convergence of the Method of Least-Squares	44
3.3.3	Optimal Algorithm Based on NN Approximation	51
3.4	Simulation-Benchmark Problem	52
3.5	Conclusion	56
4	NEURAL NETWORK SOLUTION FOR FIXED-FINAL TIME CONSTRAINED OPTIMAL CONTROL	57
4.1	Introduction	57
4.2	Background on Fixed-Final Time Constrained Optimal Control	57
4.2.1	HJB Case	58
4.2.2	HJI Case	59
4.3	Nonlinear Fixed-Final-Time Solution by NN Least-Squares Approximation	61
4.3.1	HJB Case	61

4.3.2	HJI Case	62
4.4	Numerical Examples	63
4.4.1	HJB Case	63
4.4.1.1	Linear System	64
4.4.1.2	Nonlinear Chained System	68
4.4.2	HJI Case	71
4.5	Conclusion.....	76
5	SUBOPTIMAL CONTROL OF CHAINED SYSTEM WITH TIME-FOLDING METHOD	77
5.1	Introduction	77
5.2	Problem Description.....	77
5.3	Neural Network Algorithm for Chained Form System with Time-Folding Method	78
5.3.1	Chained Form System Description	79
5.3.2	Dynamic Control Design	79
5.4	Simulation	80
5.5	Conclusion.....	84
6	CONTRIBUTIONS AND FUTURE WORK.....	85
6.1	Contributions	85
6.2	Future Work	86
	REFERENCES	87

BIOGRAPHICAL INFORMATION.....98

LIST OF ILLUSTRATIONS

Figure	Page
2-1 Linear System Weights	31
2-2 State Trajectory of Linear System.....	31
2-3 Optimal NN Control Law.....	32
2-4 Nonlinear System Weights.....	34
2-5 State Trajectory of Nonlinear System	34
2-6 Optimal NN Control Law.....	35
3-1 State feedback nonlinear H_∞ controller	38
3-2 Rotational actuator to control a translational oscillator	52
3-3 r, θ State Trajectories	54
3-4 $\dot{r}, \dot{\theta}$ State Trajectories	54
3-5 $u(t)$ Control Input	55
3-6 Disturbance Attenuation.....	55
4-1 Nonquadratic cost.....	65
4-2 Constrained Linear System Weights.....	66
4-3 State Trajectory of Linear System with Bounds	66
4-4 Optimal NN Control Law with Bounds	67
4-5 Unconstrained Control System Weights	68
4-6 Nonlinear System Weights.....	70
4-7 State Trajectory of Nonlinear System	70

4-8	Optimal NN Constrained Control Law	71
4-9	Rotational actuator to control a translational oscillator	72
4-10	r, θ State Trajectories	74
4-11	$\dot{r}, \dot{\theta}$ State Trajectories	74
4-12	$u(t)$ Control Input	75
4-13	Disturbance Attenuation.....	75
5-1	Nonlinear System Weights.....	82
5-2	State trajectories under the time folding control ($a = 0.5$)	82
5-3	State trajectories under the time folding control ($a = 0.61$)	83
5-4	Optimal NN Control Law.....	83

NOMENCLATURE

x	state vector of the dynamical system
$\ x\ $	the 2-norm of vector x
x'	transpose of the vector x
$V(x)$	value or cost of x
V_x	Jacobian of V with respect to x
H_2	2-norm on the Hardy space
H_∞	∞ -norm on the Hardy space
Ω	compact set of the state space
$C^m(\Omega)$	continuous and differentiable up to the m^{th} degree on Ω
w	neural network weight
\mathbf{w}	neural network weight vector
σ	neural network activation function
$\boldsymbol{\sigma}$	neural network activation functions vector
$\nabla \boldsymbol{\sigma}$	gradient of $\boldsymbol{\sigma}$ with respect to x
HJB	Hamilton-Jacobi-Bellman
HJI	Hamilton-Jacobi-Isaacs
DOV	Domain of Validity
\exists	there exists
$\sup_{x \in \Omega}$	supremum of a function with respect to x on Ω

\min_u minimum with respect to u

\max_d maximum with respect to d

$\langle a(x), b(x) \rangle$ integral $\int a(x)b(x)dx$ for scalar $a(x)$ and $b(x)$

CHAPTER 1

INTRODUCTION

1.1. Significance and Contribution of the Research

In this research, a practical design method to design H_2 and H_∞ optimal state feedback controllers for unconstrained and constrained input systems is proposed. The value function of the associated optimization problem is solved in a least-squares sense resulting in nearly optimal neural network state feedback controllers that are valid over a prescribed region of the state space. These feedback controllers are more appropriate for engineering applications. Hence, this work tries to bridge the gap between theoretical optimal control and practical implementations of optimal controllers. A unified framework for constructing neural network controllers that are nearly H_2 and H_∞ optimal for unconstrained and constrained input systems is provided.

The constrained input optimization of dynamical systems has been the focus of many papers during the last few years. Several methods for deriving constrained control laws are found in Saberi, Lin and Teel [76], Sussmann, Sontag and Yang [84] and Bernstein [15]. However, most of these methods do not consider optimal control laws for general constrained nonlinear systems. Constrained-input optimization possesses challenging problems, a great variety of versatile methods have been successfully applied in Athans [5], Bernstein [16], Dolphus [33] and Saberi [77]. Many problems can be formulated within the Hamilton-Jacobi-Bellman (HJB) and Lyapunov's frameworks,

but the resulting equations are difficult or impossible to solve, such as Lyshevski [60][61][62].

The optimal control of constrained input systems is theoretically well established. The controller can be found by applying the Pontryagin's minimum principle. This usually requires solving a split boundary differential equation and the result is an open-loop optimal control [53].

Optimal L_2 -gain disturbance attenuation controllers are also treated in this work. This comes under the framework of H_∞ optimal control. The H_∞ norm has played an important role in the study and analysis of robust optimal control theory since its original formulation in an input-output setting by Zames [91]. More insight into the problem was given after the H_∞ linear control problem was posed as a zero-sum two-person differential game by Başar [10]. The nonlinear counterpart of the H_∞ control theory was developed by Van der Schaft [87]. He utilized the notion of dissipativity, introduced by Willems [90][89], Hill and Moylan for nonlinear systems [41], to formulate the H_∞ control theory into a nonlinear L_2 -gain optimal control problem. He made use of the fact that the H_∞ norm in the frequency domain is nothing but the L_2 -induced norm from the input time-function to the output-time function for initial zero state. The L_2 -gain optimal control problem requires solving a Hamilton-Jacobi equation, namely the Hamilton-Jacobi-Isaacs (HJI) equation. Conditions for the existence of smooth solutions of the Hamilton-Jacobi equation were studied through invariant manifolds of Hamiltonian vector fields and the relation with

the Hamiltonian matrices of the corresponding Riccati equation for the linearized problem [87]. Later some of these conditions were relaxed by Isidori and Astolfi [45], into critical and noncritical cases. Viscosity solutions of the HJI equation were considered in [7][8].

Although the formulation of the nonlinear theory of H_∞ control has been well developed, solving the HJI equation remains a challenge. Several methods have been proposed to solve the HJI equation. In the work by Huang [44], the smooth solution is found by solving for the Taylor series expansion coefficients in a very efficient and organized manner. Another interesting method is by Beard and coworkers [13]. Beard proposed to iterate in policy space to solve the HJI successively, he then proposed a numerically efficient algorithm that solves the sequence of linear differential equations using Galerkin techniques which requires computing numerous integrals over a well valid region of the state space.

In this research, special nonquadratic performance functionals are used to encode the various constraints on the optimal control problem. Using the dynamic programming principle, the structure of the feedback strategy for the optimal control law is derived.

1.2. Approach

In this dissertation, fixed-final time constrained optimal control laws using neural networks to solve Hamilton-Jacobi equations for general affine in the unconstrained and constrained nonlinear systems are proposed. A neural network is used to approximate

the time-varying cost function using the method of least-squares on a pre-defined region. The result is a neural network nearly optimal constrained feedback controller that has time-varying coefficients found by a priori offline tuning.

1.2.1. H_2 Optimal Control: Hamilton-Jacobi-Bellman (HJB) equation

The approach here is based on HJB equation for the control input along with neural networks. In this case, the value function of the associated HJB equation is solved. As the order of the neural network is increased, the least-square solution of the HJB equation converges uniformly to the exact solution of the inherently nonlinear HJB equation. The result is a nearly optimal state feedback controller that has been tuned a priori off-line.

1.2.2. H_∞ Optimal Control: Hamilton-Jacobi-Isaacs (HJI) equation

The approach here is based on HJI equation on the input and the disturbance. Neural networks are used to approximately solve the finite-horizon optimal H_∞ state feedback control problem. The method is based on solving a related Hamilton-Jacobi-Isaacs equation of the corresponding finite-horizon zero-sum game. The neural network approximates the corresponding game value function on a certain domain of the state-space and results in a control computed as the output of a neural network. An H_∞ optimal control is obtained for the constrained input systems and the resulting available storage solves for the value function of the associated HJI equation of the associated zero-sum game. The saddle point strategy corresponding to the related

zero-sum differential game is derived, and shown to be the unique feedback saddle point.

CHAPTER 2

FIXED-FINAL TIME OPTIMAL CONTROL OF NONLINEAR SYSTEMS USING NEURAL NETWORK HJB APPROACH

2.1. Introduction

In many practical engineering problems, one is interested in finding finite-time optimal control laws for nonlinear systems. It is known that this optimization problem [53], requires solving a time-varying Hamilton-Jacobi-Bellman (HJB) equation that is hard to solve in most cases. Approximate HJB solutions for the infinite horizon time-invariant case have been found using many techniques such as those developed by Saridis and Lee [80], Beard et. Al [11][15], Beard, Bertsekas and Tsitsiklis [17], Munos et. al [65] and Kim, Lewis and Dawson [47]. Huang and Lin [44] provided a Taylor series expansion of the HJI equation which is closely related to the HJB equation. A local H_∞ controller is derived in [3] using perturbation methods.

Successful neural networks (NN) controllers not based on optimal techniques have been reported in Chen and Liu [26], Lewis, Jagannathan and Yesildirek [52], Ge [40]. It has been shown that NN can effectively extend adaptive control techniques to nonlinearly parameterized systems. NN applications to an optimal control via the HJB equation were first proposed by Werbos [63]. Parisini and Zoppoli [70] used NN to derive optimal control laws for discrete-time stochastic nonlinear systems.

In this chapter, we use NN to approximately solve the time-varying HJB equation in unconstrained and constrained cases. It is shown that using a NN approach, one can

simply transform the problem into solving an ordinary differential equation (ODE) equation backwards in time. The coefficients of this ODE are obtained by the weighted residuals method and a Kronecker product formulation [22].

We were motivated by the important results in [11]. However, in contrast to that work, we are able to approximately solve the time-varying HJB equation, and do not need to perform policy iteration using the so-called GHJB equation followed by control law updates. We accomplish this by using a neural network approximation for the value function which is based on a universal basis set, and by introduction of the Kronecker product to handle bilinear terms. The Galerkin integrals used in [11] are complicated to evaluate for bilinear terms. We also demonstrate uniform convergence results over a Sobolev space.

2.2. Background on Fixed-Final-Time HJB Optimal Control

Consider an affine in the control nonlinear dynamical system of the form

$$\dot{x} = f(x) + g(x)u(t), \quad (2-1)$$

where $x \in \mathbb{R}^n$, $f(x) \in \mathbb{R}^n$, $g(x) \in \mathbb{R}^{n \times m}$ and the input $u(t) \in \mathbb{R}^m$. The dynamics $f(x)$ and $g(x)$ are assumed to be known and $f(0) = 0$. Assume that $f(x) + g(x)u(t)$ is Lipschitz continuous on a set $\Omega \subseteq \mathbb{R}^n$ containing the origin, and that system (2-1) is stabilizable in the sense that there exists a continuous control on Ω that asymptotically stabilizes the system. It is desired to find the constrained input control $u(t)$ that minimizes a generalized functional

$$V(x(t_0), t_0) = \phi(x(t_f), t_f) + \int_{t_0}^{t_f} [Q(x) + W(u)] dt \quad (2-2)$$

with $Q(x)$, $W(u)$ positive definite on Ω , i.e. $\forall x \neq 0$, $x \in \Omega$, $Q(x) > 0$ and $x = 0 \Rightarrow Q(x) = 0$.

Definition 2.1. Admissible Controls.

A control $u(t)$ is defined to be admissible with respect to (2-2) on Ω , denoted by $u \in \Psi(\Omega)$, if $u(t)$ is continuous on Ω , $u(0) = 0$, $u(t)$ stabilizes (2-1) on Ω , and $\forall x_0 \in \Omega$, $V(x(t_0), t_0)$ is finite.

Under regularity assumptions, i.e. $V(x, t) \in C^1(\Omega)$, an infinitesimal equivalent to (2-2) is [53]

$$-\frac{\partial V(x, t)}{\partial t} = L + \left(\frac{\partial V(x, t)}{\partial x} \right)^T (f(x) + g(x)u(t)). \quad (2-3)$$

where $L = Q(x) + W(u)$. This is a time-varying partial differential equation with $V(x, t)$ the cost function for any given $u(t)$ and is solved backward in time from $t = t_f$. By setting $t_0 = t_f$ in (2-2) its boundary condition is seen to be

$$V(x(t_f), t_f) = \phi(x(t_f), t_f). \quad (2-4)$$

According to Bellman's optimality principle [53], the optimal cost is given by

$$-\frac{\partial V(x, t)^*}{\partial t} = \min_{u(t)} \left(L + \left(\frac{\partial V(x, t)^*}{\partial x} \right)^T (f(x) + g(x)u(t)) \right), \quad (2-5)$$

which yields the optimal control.

$$u^*(x, t) = -\frac{1}{2} R^{-1} g^T(x) \frac{\partial V(x, t)^*}{\partial x} \quad (2-6)$$

where $V^*(x, t)$ is the optimal value function, R is positive definite and assumed to be symmetric for simplicity of analysis.. Substituting (2-6) into (2-5) yields the well-known time-varying Hamilton-Jacobi-Bellman (HJB) equation [53]

$$\frac{\partial V(x, t)^*}{\partial t} + \frac{\partial V(x, t)^*}{\partial x} f(x) + Q(x) - \frac{1}{4} \frac{\partial V(x, t)^{*T}}{\partial x} g(x) R^{-1} g(x)^T \frac{\partial V(x, t)^*}{\partial x} = 0. \quad (2-7)$$

This equation and (2-6) provide the solution to fixed-final time optimal control for general nonlinear systems. However, close form solution for equation (2-7) is in general impossible to find. In [29][30][31][28][32], we showed how to approximately solve this equation using NN.

Remark 2.1. The HJB equation requires that $V(x, t)$ is continuously differentiable function. Usually, this requirement is not satisfied in constrained optimization because the control function is piecewise continuous. But control problems do not necessarily have smooth or even continuous value functions, (Huang [43], Bardi [8]). Lio [54] used the theory of viscosity solutions to show that for infinite horizon optimal control problems with unbounded cost functional, under certain continuity assumptions of the dynamics, the value function is continuous on some set Ω , $V^*(x, t) \in C(\Omega)$. Bardi [8] showed that if the Hamiltonian is strictly convex and if the continuous viscosity solution is semi-concave, then $V^*(x, t) \in C^1(\Omega)$ satisfying the HJB equation everywhere. In this chapter, all derivations are performed under the assumption of

smooth solutions to (2-7). A similar assumption was made by Van der schaft [87] and Isidori [45].

2.3. Nonlinear Fixed-Final-Time HJB Solution by NN Least-Squares Approximation

The HJB equation (2-11) is difficult to solve for the cost function $V(x, t)$. In this chapter, NN are used to solve approximately for the value function in (2-11) over Ω by approximating the cost function $V(x, t)$ uniformly in t . The result is an efficient, practical, and computationally tractable solution algorithm to find nearly optimal state feedback controllers for nonlinear systems.

2.3.1. NN Approximation of the Cost Function $V(x, t)$

It is well known that a NN can be used to approximate smooth time-invariant functions on prescribed compact sets (Hornik [42]). Since the analysis required here is restricted to the region of asymptotically stable (RAS) of some initial stabilizing controller, NN are natural for this application. In [78], it is shown that NNs with time-varying weights can be used to approximate uniformly continuous time-varying functions. We assume that $V(x, t)$ is smooth, and so uniformly continuous on a compact set. Therefore one can use the following equation to approximate $V(x, t)$ for $t \in [t_0, t_f]$ on a compact set

$$\Omega \subset \Re^n$$

$$V_L(x, t) = \sum_{j=1}^L w_j(t) \sigma_j(x) = w_L^T(t) \sigma_L(x). \quad (2-8)$$

This is a NN with activation functions $\sigma_j(x) \in C^1(\Omega)$, $\sigma_j(0) = 0$. The NN weights are $w_j(t)$ and L is the number of hidden-layer neurons. $\sigma_L(x) \equiv [\sigma_1(x)\sigma_2(x)...\sigma_L(x)]^T$ is the vector of activation function, $w_L(t) \equiv [w_1(t)w_2(t)...\sigma_L(t)]^T$ is the vector of NN weights.

The next result shows that initial conditions $x(t_0)$ can be selected to guarantee that $x(t) \in \Omega$ for $t \in [t_0, t_f]$.

Lemma 2.1 Let $\Omega \subset \mathbb{R}^n$ be a compact set. Then $\exists \Omega_0 \subset \Omega$, s. t, for system (2-1),

$$x(t) \in \Omega, \quad t \in [t_0, t_f], \quad \forall x(t_0) \in \Omega_0.$$

The set $\sigma_j(x)$ is selected to be independent. Then without loss of generality, they can be assumed to be orthonormal, i.e. select equivalent basis functions to $\sigma_j(x)$ that are also orthonormal [11]. The orthonormality of the set $\{\sigma_j(x)\}_{j=1}^\infty$ on Ω implies that if a function $\psi(x, t) \in L_2(\Omega)$ then

$$\psi(x, t) = \sum_{j=1}^{\infty} \langle \psi(x, t), \sigma_j(x) \rangle_{\Omega} \sigma_j(x),$$

where $\langle f, g \rangle_{\Omega} = \int_{\Omega} g \cdot f^T dx$ is inner product, and the series converges pointwise, i.e. for any $\varepsilon > 0$ and $x \in \Omega$, one can choose N sufficiently large to guarantee that

$$\left| \sum_{j=N+1}^{\infty} \langle \psi(x, t), \sigma_j(x) \rangle_{\Omega} \sigma_j(x) \right| < \varepsilon \quad \text{for all } t \in [t_0, t_f], \text{ see [12].}$$

Note that, since one requires $\partial V(x, t)/\partial t$ in (2-7), the NN weights are selected to be time-varying. This is similar to methods such as assumed mode shapes in the study of flexible mechanical systems [6]. However, here $\sigma_L(x)$ is a NN activation vector, not a set of eigenfunctions. That is, the NN approximation property significantly simplifies the specification of $\sigma_L(x)$. For the infinite final time case, the NN weights are constant [1]. The NN weights will be selected to minimize a residual error in a least-squares sense over a set of points sampled from a compact set Ω_0 inside the RAS of the initial stabilizing control [38].

Note that

$$\frac{\partial V_L(x, t)}{\partial x} = \frac{\partial \sigma_L^T(x)}{\partial x} \mathbf{w}_L(t) \equiv \nabla \sigma_L^T(x) \mathbf{w}_L(t), \quad (2-9)$$

where $\nabla \sigma_L(x)$ is the Jacobian $\partial \sigma_L(x)/\partial x$, and that

$$\frac{\partial V_L(x, t)}{\partial t} = \dot{\mathbf{w}}_L^T(t) \sigma_L(x). \quad (2-10)$$

Therefore approximating $V(x, t)$ by $V_L(x, t)$ uniformly in t in the HJB equation (2-7) results in

$$\begin{aligned} & -\dot{\mathbf{w}}_L^T(t) \sigma_L(x) - \mathbf{w}_L^T(t) \nabla \sigma_L(x) f(x) - 2 \int_0^u \Phi^{-T}(v) R dv \\ & + \mathbf{w}_L^T(t) \nabla \sigma_L(x) \cdot g(x) \cdot \Phi \left(\frac{1}{2} R^{-1} g^T(x) \nabla \sigma_L^T(x) \mathbf{w}_L(t) \right) - Q(x) \\ & = e_L(x, t) \end{aligned} \quad (2-11)$$

or

$$HJB \left(V_L(x, t) = \sum_{j=1}^L w_j(t) \sigma_j(x) \right) = e_L(x, t), \quad (2-12)$$

where $e_L(x, t)$ is a residual equation error. From (2-6) the corresponding optimal control input is

$$u_L(x, t) = -\frac{1}{2} R^{-1} g^T(x) \nabla \sigma_L^T(x) \mathbf{w}_L(t). \quad (2-13)$$

To find the least-squares solution for $\mathbf{w}_L(t)$, the method of weighted residuals is used [38]. The weight derivatives $\dot{\mathbf{w}}_L(t)$ are determined by projecting the residual error onto $\partial e_L(x, t) / \partial \dot{\mathbf{w}}_L(t)$ and setting the result to zero $\forall x \in \Omega_0$ and $\forall t \in [t_0, t_f]$ using the inner product, i.e.

$$\left\langle \frac{\partial e_L(x, t)}{\partial \dot{\mathbf{w}}_L(t)}, e_L(x, t) \right\rangle_{\Omega} = 0. \quad (2-14)$$

From (2-11) we can get

$$\frac{\partial e_L(x, t)}{\partial \dot{\mathbf{w}}_L} = \sigma_L(x). \quad (2-15)$$

Therefore we obtains

$$\begin{aligned} & \left\langle -\dot{\mathbf{w}}_L^T(t) \sigma_L(x), \sigma_L(x) \right\rangle_{\Omega} + \left\langle -\mathbf{w}_L^T(t) \nabla \sigma_L(x) f(x), \sigma_L(x) \right\rangle_{\Omega} \\ & + \left\langle \frac{1}{4} \mathbf{w}_L^T(t) \nabla \sigma_L(x) g(x) R^{-1} \cdot g^T(x) \nabla \sigma_L^T(x) \mathbf{w}_L(t), \sigma_L(x) \right\rangle_{\Omega} \\ & + \left\langle -Q(x), \sigma_L(x) \right\rangle_{\Omega} = 0 \end{aligned} \quad (2-16)$$

So that

$$\begin{aligned} & \dot{\mathbf{w}}_L(t) = \\ & -\left\langle \sigma_L(x), \sigma_L(x) \right\rangle_{\Omega}^{-1} \cdot \left\langle \nabla \sigma_L(x) f(x), \sigma_L(x) \right\rangle_{\Omega} \cdot \mathbf{w}_L(t) \\ & + \left\langle \sigma_L(x), \sigma_L(x) \right\rangle_{\Omega}^{-1} \cdot \left\langle \frac{1}{4} \mathbf{w}_L^T(t) \nabla \sigma_L(x) g(x) R^{-1} \cdot g^T(x) \nabla \sigma_L^T \mathbf{w}_L(t), \sigma_L(x) \right\rangle_{\Omega} \\ & - \left\langle \sigma_L(x), \sigma_L(x) \right\rangle_{\Omega}^{-1} \cdot \left\langle Q(x), \sigma_L(x) \right\rangle_{\Omega} \end{aligned} \quad (2-17)$$

with boundary condition $V(x(t_f), t_f) = \phi(x(t_f), t_f) = \mathbf{w}_L^T(t_f) \boldsymbol{\sigma}_L(x(t_f))$. Note that, given a mesh of $x(t_f)$ (see section 3.3), the boundary condition allows one to determine $\mathbf{w}_L(t_f)$.

Therefore, the NN weights are simply found by integrating this nonlinear ODE backwards in time.

We now show that this procedure provides a nearly optimal solution for the time-varying optimal control problem if L is selected large enough.

2.3.2. Uniform Convergence in t For Time-Varying Function of the Method of Least-Squares

In what follows, one shows convergence results as L increases for the method of least squares when NN are used to uniformly approximate the cost function in t . The following definitions and facts are required.

Let $F(t, x)$ be piecewise continuous in t and satisfy the Lipschitz condition

$$\|F(x, t) - F(y, t)\| \leq L\|x - y\|,$$

$\forall x, y \in B = \{x \in R^n \mid \|x - x_0\| \leq r\}, \forall t \in [t_0, t_1]$, where $\|F\|^2 = \langle F, F \rangle$. Then, there exists some $\delta > 0$ such that the state equation $\dot{x} = F(x, t)$ with $x(t_0) = x_0$ has a unique solution over $[t_0, t_0 + \delta]$. Provided the Lipschitz condition holds uniformly in t for all t in a given interval of time, function $F(x, t)$ is called globally Lipschitz if it is Lipschitz on R^n . (Khalil [46]).

Definition 2.2. Convergence in the Mean for Time-Varying Functions.

A sequence of functions $\{f_n(x, t)\}$ that is Lebesgue integrable on a set Ω , $L_2(\Omega)$, is said to converge (uniformly in t) in the mean to $f(x, t)$ on Ω if $\forall \varepsilon > 0$, $\forall t$, $\exists N(\varepsilon, t)$: $n > N \Rightarrow \|f_n(x, t) - f(x, t)\|_{L_2(\Omega)} < \varepsilon$.

Definition 2.3. Uniform Convergence for Time-Varying Functions.

A sequence of functions $\{f_n(x, t)\}$ converges to $f(x, t)$ (uniformly in t) on a set Ω if $\forall \varepsilon > 0$, $\forall t$, $\exists N(\varepsilon, t)$: $n > N \Rightarrow |f_n(x, t) - f(x, t)| < \varepsilon \quad \forall x \in \Omega$, or equivalently $\sup_{x \in \Omega} |f_n(x, t) - f(x, t)| < \varepsilon$.

Definition 2.4. Sobolev Space.

$H^{m,p}(\Omega)$: Let Ω be an open set in \mathbb{R}^n and let $u \in C^m(\Omega)$. Define a norm on u by

$$\|u\|_{m,p} = \sum_{0 \leq |\alpha| \leq m} \left(\int_{\Omega} |D^{\alpha} u(x)|^p dx \right)^{1/p}, \quad 1 \leq p < \infty.$$

This is the Sobolev norm in which the integration is Lebesgue. The completion of $u \in C^m(\Omega)$: $\|u\|_{m,p} < \infty$ with respect to $\|\cdot\|_{m,p}$ is the Sobolev space $H^{m,p}(\Omega)$. For $p = 2$, the Sobolev space is a Hilbert space.

The convergence proofs of the least-squares method are done in the Sobolev function space $H^{1,2}(\Omega)$ setting [2], since one requires to prove the convergence of both $V_L(x, t)$ and its gradient. The following Technical Lemmas are required.

Technical Lemma 2.1.

Given a linearly independent set of L functions $\{f_L\}$. Then for the series $a_L^T f_L$, it follows that $\|a_L^T f_L\|_{L_2(\Omega)}^2 \rightarrow 0 \Leftrightarrow \|a_L\|_{L_2(\Omega)}^2 \rightarrow 0$.

Proof: See [1]. ■

Technical Lemma 2.2.

Suppose that $\{\nabla \sigma_j(x)\}_1^L \neq 0$, then $\{\sigma_j(x)\}_1^L$ -linearly independent \Rightarrow $\{\nabla \sigma_j(x)\}_1^L$ -linearly independent.

Proof. See [11]. ■

Technical Lemma 2.3.

If $W(x) = \sum_{j=1}^{\infty} w_j(t) \phi_j(x)$ and $\phi_j(x)$ are continuous on Ω , then $\sum_{j=L+1}^{\infty} w_j(t) \phi_j(x)$ converges to zero uniformly in t on Ω iff

1) $W(x)$ is continuous on Ω .

2) $\sum_{j=1}^{\infty} w_j(t) \phi_j(x) \in PD(\Omega)$,

where $PD(\Omega)$ means pointwise decreasing on Ω .

Proof: See [11]. ■

The following assumptions are required.

Assumption 2.1. The system's dynamics and the performance integrands $Q(x)+W(u)$ are such that are solution of the cost function which is continuous and differentiable. Therefore, belonging to the Sobolev space $V \in H^{1,2}(\Omega)$. Here $Q(x)$ and $W(u)$ satisfy the requirement of existence of smooth solutions.

Assumption 2.2. We can choose a complete coordinate elements $\{\sigma_j(x)\}_1^\infty \in H^{1,2}(\Omega)$ such that the solution $V(x,t) \in H^{1,2}(\Omega)$ and $\{\partial V(x,t)/\partial x_1, \dots, \partial V(x,t)/\partial x_n\}$ can be uniformly approximated in t by the infinite series built from $\{\sigma_j(x)\}_1^\infty$.

Assumption 2.3. The coefficients $|w_j(t)|$ are uniformly bounded in t for all L .

The first two assumptions are standard in optimal control and Neural Networks control literature. Completeness follows from [42].

We now show the following convergence results.

Lemma 2.2 Convergence of Approximate HJB Equation.

Given $u \in \mathcal{U}(\Omega)$. Let $V_L(x,t) = \sum_{j=1}^L w_j^T(t) \sigma_j(x)$ satisfy $\langle HJB(V_L(x,t)), \sigma_L(x) \rangle_\Omega = 0$

and $\langle V_L(x(t_f), t_f), \sigma_L(x) \rangle_\Omega = 0$,

and let $V(x,t) = \sum_{j=1}^\infty c_j^T(t) \sigma_j(x)$ and $\mathbf{c}_L(t) \equiv [c_1(t) c_2(t) \dots c_L(t)]^T$ satisfy

$HJB(V(x,t)) = 0$ and $V(x(t_f), t_f) = \phi(x(t_f), t_f)$.

Then

$|HJB(V_L(x, t))| \rightarrow 0$ uniformly in t on Ω_0 as L increases.

Proof. The hypotheses imply that $HJB(V_L(x, t))$ are in $L_2(\Omega)$. Note that

$$\begin{aligned} \langle HJB(V_L(x, t)), \sigma_j(x) \rangle_\Omega &= \sum_{k=1}^L \dot{w}_k^T(t) \langle \sigma_k(x), \sigma_j(x) \rangle_\Omega \\ &+ \sum_{k=1}^L w_k^T(t) \langle \nabla \sigma_k(x) f(x), \sigma_j(x) \rangle_\Omega \\ &- \sum_{k=1}^L w_k^T(t) \cdot \left\langle \frac{1}{4} \nabla \sigma_k(x) g(x) R^{-1} \cdot g^T(x) \nabla \sigma_k^T(x), \sigma_j(x) \right\rangle_\Omega \cdot w_k(t), \\ &+ \langle Q(x), \sigma_j(x) \rangle_\Omega \end{aligned} \quad (2-18)$$

Since the set $\{\sigma_j(x)\}_1^\infty$ are orthogonal, $\langle \sigma_k(x), \sigma_j(x) \rangle_\Omega = 0$.

Then

$$\begin{aligned} |HJB(V_L(x, t))| &= \left| \sum_{j=1}^\infty \langle HJB(V_L(x, t)), \sigma_j(x) \rangle_\Omega \sigma_j(x) \right| \\ &= \left| \begin{aligned} &\sum_{j=L+1}^\infty \left(\sum_{k=1}^L \dot{w}_k^T(t) \langle \sigma_k(x), \sigma_j(x) \rangle_\Omega \right) \sigma_j(x) \\ &+ \sum_{j=L+1}^\infty \left(\sum_{k=1}^L w_k^T(t) \langle \nabla \sigma_k(x) f(x), \sigma_j(x) \rangle_\Omega \right) \sigma_j(x) \\ &- \sum_{j=L+1}^\infty \left(\sum_{k=1}^L w_k^T(t) \left\langle \frac{1}{4} \nabla \sigma_k(x) g(x) R^{-1} \cdot g^T(x) \nabla \sigma_k^T(x), \sigma_j(x) \right\rangle_\Omega w_k(t) \right) \sigma_j(x) \\ &+ \sum_{j=L+1}^\infty \langle Q(x), \sigma_j(x) \rangle_\Omega \sigma_j(x) \end{aligned} \right|, \end{aligned} \quad (2-19)$$

Since the set $\{\sigma_j(x)\}_1^\infty$ are orthogonal, $\langle \sigma_k(x), \sigma_j(x) \rangle_\Omega = 0$.

Therefore

$$\begin{aligned} |HJB(V_L(x, t))| &\leq \left| \left(\sum_{k=1}^L w_k(t) \sum_{j=L+1}^\infty \langle \nabla \sigma_k(x) f(x), \sigma_j(x) \rangle_\Omega \cdot \sigma_j(x) \right) \right| \\ &+ \left| \left(\sum_{k=1}^L w_k^2(t) \cdot \sum_{j=L+1}^\infty \left\langle \frac{1}{4} \nabla \sigma_k(x) g(x) R^{-1} g^T(x) \nabla \sigma_k^T(x), \sigma_j(x) \right\rangle_\Omega \cdot \sigma_j(x) \right) \right| \\ &+ \left| \sum_{j=L+1}^\infty \langle Q(x), \sigma_j(x) \rangle_\Omega \sigma_j(x) \right| \\ \therefore &\leq AB(x) + CD(x) + Vec \left(\left| \sum_{j=L+1}^\infty \langle Q(x), \sigma_j(x) \rangle_\Omega \sigma_j(x) \right| \right), \end{aligned} \quad (2-20)$$

where

$$A = \max_{1 \leq k \leq L} |w_k(t)|,$$

$$B(x) = \sup_{(t,x) \in [t_0, T] \times \Omega} \left| \sum_{k=1}^L \left(\sum_{j=L+1}^{\infty} \langle \nabla \sigma_k(x) f(x), \sigma_j(x) \rangle_{\Omega} \right) \sigma_j(x) \right|,$$

$$C = \max_{1 \leq k \leq L} |w_k^2(t)|,$$

$$D = \sup_{\substack{(t,x) \\ \in [t_0, T] \times \Omega}} \left\| \left(\sum_{k=1}^L \left(\sum_{j=L+1}^{\infty} \left\langle \frac{1}{4} \nabla \sigma_k(x) g(x) R^{-1} g^T(x) \nabla \sigma_k^T(x), \sigma_j(x) \right\rangle_{\Omega} \right) \sigma_j(x) \right) \right\|.$$

Suppose $\nabla \sigma_k(x) f(x)$, $\frac{1}{4} \nabla \sigma_k(x) g(x) R^{-1} g^T(x) \nabla \sigma_k^T(x)$ and $Q(x)$ are in $L_2(\Omega)$,

the orthonormality of the set $\{\sigma_j(x)\}_{j=1}^{\infty}$ implies that $B(x)$ and the second and third term on the right-hand side can be made arbitrarily small by an appropriate choice of L .

Therefore

$$A \cdot B(x) + C \cdot D(x) \rightarrow 0 \quad \text{and} \quad \left| \sum_{j=L+1}^{\infty} \langle Q(x), \sigma_j(x) \rangle_{\Omega} \sigma_j(x) \right| \rightarrow 0.$$

So $|HJB(V_L(x, t))| \rightarrow 0$ uniformly in t on Ω as L increases. ■

Lemma 2.3 Convergence of NN Weights

Given $u \in \Psi(\Omega)_0$ and suppose the hypotheses of Lemma 2.2 hold. Then

$$\|\mathbf{w}_L(t) - \mathbf{c}_L(t)\|_2 \rightarrow 0 \quad \text{uniformly in } t \text{ as } L \text{ increases.}$$

Proof: Define

$$e_L(x, t) = HJB(V_L(x, t)) \quad \text{and} \quad \hat{e}_L(x, t) = V_L(x(t_f), t_f) - \phi(x(t_f), t_f). \quad (2-21)$$

Then $\langle e_L(x, t), \sigma_L(x) \rangle_\Omega = \langle \hat{e}_L(x, t), \sigma_L(x) \rangle_\Omega = 0$. From the hypotheses one has that

$$\begin{aligned} HJB(V_L(x, t)) - HJB(V(x, t)) &= e_L(x, t) \\ V_L(x(t_f), t) - V(x(t_f), t) &= \hat{e}_L(x(t_f), t), \end{aligned} \quad (2-22)$$

substituting the series expansion for $V_L(x, t)$ and $V(x, t)$, and moving the terms in the series that are greater than L to the right-hand side one obtains

$$\begin{aligned} &(\dot{\mathbf{w}}_L(t) - \dot{\mathbf{c}}_L(t))^T \sigma_L(x) + (\mathbf{w}_L(t) - \mathbf{c}_L(t))^T \nabla \sigma_L(x) f(x) \\ &- (\mathbf{w}_L^T(t) \otimes \mathbf{w}_L^T(t) - \mathbf{c}_L^T(t) \otimes \mathbf{c}_L^T(t)) \cdot \text{Vec} \left(\frac{1}{4} \nabla \sigma_L(x) g(x) R^{-1} \cdot g^T(x) \nabla \sigma_L^T(x) \right) \\ &= e_L(x, t) + \sum_{j=L+1}^{\infty} \dot{c}_j^T(t) \sigma_j(x) + \sum_{j=L+1}^{\infty} c_j^T(t) \nabla \sigma_j(x) f(x) \\ &+ \sum_{j=L+1}^{\infty} c_j^2(t) \cdot \left(\frac{1}{4} \nabla \sigma_j(x) g(x) R^{-1} g^T(x) \nabla \sigma_j^T(x) \right) \end{aligned} \quad (2-23)$$

The final condition is

$$(\mathbf{w}_L^T(t_f) - \mathbf{c}_L^T(t_f)) \sigma_L(x) = \hat{e}_L(x, t) + \sum_{j=L+1}^{\infty} c_j(t_f) \sigma_j(x). \quad (2-24)$$

Taking the inner product of both sides over Ω_0 , and taking into account the orthonormality of the set $\{\sigma_j(x)\}_1^\infty$, one obtains

$$\begin{aligned} &(\dot{\mathbf{w}}_L(t) - \dot{\mathbf{c}}_L(t)) + \langle \nabla \sigma_L(x) f(x), \sigma_L(x) \rangle_\Omega^T (\mathbf{w}_L(t) - \mathbf{c}_L(t)) \\ &- \left\langle \frac{1}{4} \text{Vec}(\nabla \sigma_L(x) g(x) R^{-1} g^T(x) \nabla \sigma_L^T(x)), \sigma_L(x) \right\rangle_\Omega^T \cdot (\mathbf{w}_L(t) \otimes \mathbf{w}_L(t) - \mathbf{c}_L(t) \otimes \mathbf{c}_L(t)) \\ &= \sum_{j=L+1}^{\infty} c_j(t) \langle \nabla \sigma_j(x) f(x), \sigma_L(x) \rangle_\Omega^T \\ &+ \sum_{j=L+1}^{\infty} c_j^2(t) \left\langle \frac{1}{4} \nabla \sigma_j(x) g(x) R^{-1} g^T(x) \nabla \sigma_j^T(x), \sigma_L(x) \right\rangle_\Omega^T \end{aligned}$$

with final condition

$$\mathbf{w}_L(t_f) - \mathbf{c}_L(t_f) = 0. \quad (2-25)$$

Let $A = \langle \nabla \mathbf{\sigma}_L(x) f(x), \mathbf{\sigma}_L(x) \rangle_{\Omega}^T$, where A is scalar.

Define $\xi = \mathbf{w}_L(t) - \mathbf{c}_L(t)$, consider the equation

$$\begin{aligned} \dot{\xi} + A(t) \cdot \xi + f(\xi, t) &= 0 \\ \xi(t_f) &= 0 \end{aligned}, \quad (2-26)$$

where

$$\begin{aligned} f(\xi, t) &= - \left\langle \frac{1}{4} \text{Vec}(\nabla \mathbf{\sigma}_L(x) g(x) R^{-1} g^T(x) \nabla \mathbf{\sigma}_L^T(x)), \mathbf{\sigma}_L(x) \right\rangle_{\Omega}^T \\ &\quad \cdot (\mathbf{w}_L(x) \otimes \mathbf{w}_L(x) - \mathbf{c}_L(x) \otimes \mathbf{c}_L(x)) \end{aligned}$$

is continuously differentiable in a neighborhood of a point (ξ_0, t_0) . Since $A(t)$ is also piecewise continuous functions of t , over any finite interval of time $[t_0, t_f]$, the elements of $A(t)$ and $f(\xi, t)$ are bounded. Hence, $\|A(t)\| \leq a$, $\|f(\xi, t)\| \leq b$ and

$$\|f(x, t) - f(y, t)\| = \|A(t)(x - y)\| \leq \|A(t)\| \|x - y\| \leq a \|x - y\|,$$

$$\forall x, y \in R^n, \quad \forall t \in [t_0, t_f]$$

also

$$\|f(x_0, t)\| = \|A(t)x_0 + f(x, t)\| \leq a\|x_0\| + b \leq h,$$

for each finite x_0 , $\forall t \in [t_0, t_f]$.

Therefore, the system has a unique solution over $[t_0, t_f]$. Since t_f can be arbitrarily large, we can also conclude that if $A(t)$ and $f(x, t)$ are piecewise continuous $\forall t \geq t_0$, then the system has a unique solution $\forall t \geq t_0$, so (24) can satisfy a local Lipschitz condition [46].

Noting that

$$\sum_{j=L+1}^{\infty} c_j(t) \left\langle \nabla \sigma_j(x) f(x), \sigma_L(x) \right\rangle_{\Omega}^T + \sum_{j=L+1}^{\infty} c_j^2(t) \left\langle \frac{1}{4} \nabla \sigma_j(x) g(x) R^{-1} g^T(x) \nabla \sigma_j^T(x), \sigma_L(x) \right\rangle_{\Omega}^T$$

is continuous in t , we invoke the standard result from the theory of ordinary differential equations that a continuous perturbation in the system equations and the initial state implies a continuous perturbation of the solution (Arnold [4]). Note that

$$\begin{aligned} & \left\| \sum_{j=L+1}^{\infty} c_j(t) \left\langle \nabla \sigma_j(x) f(x), \sigma_L(x) \right\rangle_{\Omega}^T + \sum_{j=L+1}^{\infty} c_j^2(t) \left\langle \frac{1}{4} \nabla \sigma_j(x) g(x) R^{-1} g^T(x) \nabla \sigma_j^T(x), \sigma_L(x) \right\rangle_{\Omega} \right\|_{L_2(\Omega)} \\ & \leq \left\| \sum_{j=L+1}^{\infty} c_j(t) \cdot \left\langle \nabla \sigma_L(x) f(x), \sigma_L(x) \right\rangle_{\Omega}^T \right\|_{L_2(\Omega)} \\ & + \left\| \sum_{j=L+1}^{\infty} c_j^2(t) \left\langle \frac{1}{4} \nabla \sigma_j(x) g(x) R^{-1} g^T(x) \nabla \sigma_j^T(x), \sigma_j(x) \right\rangle_{\Omega} \right\|_{L_2(\Omega)} \\ & = \rho(t) \end{aligned}$$

here $\rho(t) \rightarrow 0$ as L increases.

This implies that for all $\varepsilon > 0$, there exists a $\rho(t) > 0$ such that $\forall t \in [t_0, t_f]$,

$$\|\mathbf{w}_L(t) - \mathbf{c}_L(t)\|_2 < \varepsilon. \quad (2-27)$$

So $\|\mathbf{w}_L(t) - \mathbf{c}_L(t)\|_2 \rightarrow 0$ uniformly in t on Ω as L increases. ■

Now we are in a position to prove our main results.

Lemma 2.4 Convergence of Approximate Value Function.

Under the hypotheses of Lemma 2.2, one has

$$\|V_L(x, t) - V(x, t)\|_{L_2(\Omega)} \rightarrow 0 \text{ uniformly in } t \text{ on } \Omega \text{ as } L \text{ increases.}$$

Proof. From Lemma 2.3, we have $\|\mathbf{w}_L(t) - \mathbf{c}_L(t)\|_2 \rightarrow 0$,

$$\begin{aligned}
& \|V_L(x, t) - V(x, t)\|_{L_2(\Omega)}^2 = \int_{\Omega} |V_L(x, t) - V(x, t)|^2 dx \\
& \leq \int_{\Omega} |(\mathbf{w}_L(t) - \mathbf{c}_L(t))^T \boldsymbol{\sigma}_L(x)|^2 dx + \int_{\Omega} \left| \sum_{j=L+1}^{\infty} c_j(t) \sigma_j(x) \right|^2 dx \\
& = (\mathbf{w}_L^T(t) - \mathbf{c}_L^T(t)) \left\langle \boldsymbol{\sigma}_L(x), \boldsymbol{\sigma}_L^T(x) \right\rangle_{\Omega} (\mathbf{w}_L(t) - \mathbf{c}_L(t)) \\
& \quad + \int_{\Omega} \left| \sum_{j=L+1}^{\infty} c_j(t) \sigma_j(x) \right|^2 dx
\end{aligned} \tag{2-28}$$

By the mean value theorem, Technical Lemmas 3.3, $\exists \xi \in \Omega$ such that

$$\|V_L(x, t) - V(x, t)\|_{L_2(\Omega)}^2 = \|\mathbf{w}_L(t) - \mathbf{c}_L(t)\|_2^2 + \lambda(\Omega) \cdot \left| \sum_{j=L+1}^{\infty} c_j(t) \sigma_j(\xi) \right|^2 \rightarrow 0$$

uniformly in t on Ω_0 as L increases. ■

Lemma 2.5 Convergence of Value Function Gradient.

Under the hypotheses of Lemma 2.2,

$$\left\| \frac{\partial V_L(x, t)}{\partial x} - \frac{\partial V(x, t)}{\partial x} \right\|_{L_2(\Omega)} \rightarrow 0 \text{ uniformly in } t \text{ on } \Omega_0 \text{ as } L \text{ increases.}$$

Proof. From Lemma 2.3, we have $\|\mathbf{w}_L(t) - \mathbf{c}_L(t)\|_2 \rightarrow 0$,

$$\begin{aligned}
& \left\| \frac{\partial V_L(x, t)}{\partial x} - \frac{\partial V(x, t)}{\partial x} \right\|_{L_2(\Omega)}^2 = \left\| \nabla \boldsymbol{\sigma}_L^T(x) (\mathbf{w}_L(t) - \mathbf{c}_L(t)) - \sum_{j=L+1}^{\infty} \nabla \sigma_j^T(x) c_j(t) \right\|_{L_2(\Omega)}^2 \\
& \leq \left\| \nabla \boldsymbol{\sigma}_L^T(x) (\mathbf{w}_L(t) - \mathbf{c}_L(t)) \right\|_{L_2(\Omega)}^2 + \left\| \sum_{j=L+1}^{\infty} \nabla \sigma_j^T(x) c_j(t) \right\|_{L_2(\Omega)}^2 \\
& = \left\| \nabla \boldsymbol{\sigma}_L^T(x) (\mathbf{w}_L(t) - \mathbf{c}_L(t)) \right\|_{L_2(\Omega)}^2 + \int_{\Omega} \left| \sum_{j=L+1}^{\infty} \nabla \sigma_j^T(x) c_j(t) \right|^2 dx
\end{aligned}$$

By the mean value theorem, Technical Lemmas 2.1, 2.2 and 2.3, $\exists \xi \in \Omega$ such that

$$\left\| \frac{\partial V_L(x, t)}{\partial x} - \frac{\partial V(x, t)}{\partial x} \right\|_{L_2(\Omega)}^2 = \left\| \nabla \boldsymbol{\sigma}_L^T(x) (\mathbf{w}_L(t) - \mathbf{c}_L(t)) \right\|_{L_2(\Omega)}^2 + \lambda(\Omega) \left| \sum_{j=L+1}^{\infty} \nabla \sigma_j^T(x) c_j(t) \right|^2.$$

Since $\nabla \sigma_L^T(x)$ is linearly independent and $\|\mathbf{w}_L(t) - \mathbf{c}_L(t)\|_2 \rightarrow 0$,

then

$$\left\| \frac{\partial V_L(x, t)}{\partial x} - \frac{\partial V(x, t)}{\partial x} \right\|_{L_2(\Omega)} \rightarrow 0 \text{ uniformly in } t \text{ on } \Omega_0 \text{ as } L \text{ increases.} \quad \blacksquare$$

Through Theorem 2.1 and 2.2 we have shown that the HJB approximating solution (2-12) guarantees convergence in Sobolev space $H^{1,2}$.

Lemma 2.6 Convergence of Control Inputs.

If the conditions of Lemma 2.2 are satisfied and

$$u_L(x, t) = -\frac{1}{2} R^{-1} g^T(x) \frac{\partial V_L(x, t)}{\partial x},$$

$$u(x, t) = -\frac{1}{2} R^{-1} g^T(x) \frac{\partial V(x, t)}{\partial x}.$$

Then $\|u_L(x, t) - u(x, t)\|_{L_2(\Omega)} \rightarrow 0$ in t on Ω_0 as L increases.

Proof.

Denote $\alpha_L(x) = -\frac{1}{2} R^{-1} g^T(x) \frac{\partial V_L(x, t)}{\partial x}$ and $\alpha(x) = -\frac{1}{2} R^{-1} g^T(x) \frac{\partial V(x, t)}{\partial x}$.

By Theorem 2.2 and the fact that $g(x)$ is continuous and therefore bounded on Ω ,

hence

$$\begin{aligned} & \left\| -\frac{1}{2} R^{-1} g^T(x) \frac{\partial V_L(x, t)}{\partial x} + \frac{1}{2} R^{-1} g^T(x) \frac{\partial V(x, t)}{\partial x} \right\|_{L_2(\Omega)}^2 \\ & \leq \left\| -\frac{1}{2} R^{-1} g^T(x) \right\|_{L_2(\Omega)}^2 \left\| \frac{\partial V_L(x, t)}{\partial x} - \frac{\partial V(x, t)}{\partial x} \right\|_{L_2(\Omega)}^2 \rightarrow 0 \\ & \Rightarrow \alpha_L(x) - \alpha(x) \rightarrow 0. \end{aligned}$$

Because $\varphi(\cdot)$ is smooth, and under the assumption that its first derivative is bounded, then we have $\|\varphi(\alpha_L(x)) - \varphi(\alpha(x))\| \leq M\|(\alpha_L(x) - \alpha(x))\|$. Therefore

$$\begin{aligned} \|\alpha_L(x) - \alpha(x)\|_{L_2(\Omega)} &\rightarrow 0 \\ \Rightarrow \|\varphi(\alpha_L(x)) - \varphi(\alpha(x))\|_{L_2(\Omega)} &\rightarrow 0, \end{aligned}$$

hence $\|u_L(x, t) - u(x, t)\|_{L_2(\Omega)} \rightarrow 0$ in t on Ω_0 as L increases. ■

At this point we have proven uniform convergence in t in the mean of the approximate HJB equation, the NN weights, the approximate value function and the value function gradient. This demonstrates uniform convergence in t in the mean in Sobolev space $H^{1,2}(\Omega)$. In fact, the next result shows even stronger convergence properties, namely uniform convergence in both x and t .

Lemma 2.7 Convergence of State Trajectory.

Let $x_L(t)$ be the state using control (2-13), suppose the hypotheses of Lemma 2.2 hold.

Then

$x(t) - x_L(t) \rightarrow 0$ uniformly in t on Ω as L increases.

Proof:

$$\begin{aligned} \dot{x}(t) &= f(x) + g(x)u(t) = f(x) - \frac{1}{2}g(x)R^{-1}g^T(x)\frac{\partial V(x, t)}{\partial x} \\ \dot{x}_L(t) &= f(x_L) + g(x_L)u(t) = f(x_L) - \frac{1}{2}g(x_L)R^{-1}g^T(x_L)\frac{\partial V(x_L, t)}{\partial x_L} \\ x_L(t_0) &= x(t_0) \end{aligned}$$

Since $f(x) - f(x_L) \leq L\|x - x_L\|$

$$\begin{aligned} & \dot{x}(t) - \dot{x}_L(t) \\ &= f(x) - f(x_L) - \left(\frac{1}{2} g(x) R^{-1} g^T(x) \frac{\partial V(x, t)}{\partial x} - \frac{1}{2} g(x_L) R^{-1} g^T(x_L) \frac{\partial V_L(x_L, t)}{\partial x_L} \right) \\ &\leq L\|x - x_L\| - \left(\frac{1}{2} \|R^{-1}\| \cdot (\|g(x)\|_2^2 - \|g(x_L)\|_2^2) \frac{\partial V(x, t)}{\partial x} \right. \\ &\quad \left. + \frac{1}{2} \left(g(x_L) R^{-1} g^T(x_L) \left(\frac{\partial V(x, t)}{\partial x} - \frac{\partial V(x_L, t)}{\partial x_L} \right) \right) \right) \end{aligned}$$

Define

$$\tilde{x}(t) = x(t) - x_L(t),$$

Consider the equation

$$\dot{\tilde{x}} - L\|\tilde{x}\| + h(\tilde{x}, t) = \rho(x)$$

$$\tilde{x}(t_0) = 0$$

where

$$h(\tilde{x}, t) = - \left(\frac{1}{2} \|R^{-1}\| \cdot (\|g(x)\|_2^2 - \|g(x_L)\|_2^2) \frac{\partial V(x, t)}{\partial x} \right)$$

$$\rho(x) = - \frac{1}{2} \left(g(x_L) R^{-1} g^T(x_L) \left(\frac{\partial V(x, t)}{\partial x} - \frac{\partial V(x_L, t)}{\partial x_L} \right) \right)$$

are continuously differentiable in a neighborhood of a point (\tilde{x}_0, t_0) . Over any finite interval of time $[t_0, t_f]$, the elements of $h(\tilde{x}, t)$ are bounded. Therefore (26) has a unique solution. From Lemma 2.5, $\rho(x) \rightarrow 0$ as L increases. We invoke the standard result from the theory of ordinary differential equations, as in Lemma 2.3 proof, so that $\|\tilde{x}\| \rightarrow 0$ uniformly in t on Ω as L increases. ■

Lemma 2.8 Uniform Convergence.

Since a local Lipschitz condition holds on (2-29), then

$$\sup_{x \in \Omega} |V_L(x, t) - V(x, t)| \rightarrow 0, \quad \sup_{x \in \Omega} \left| \frac{\partial V_L(x, t)}{\partial x} - \frac{\partial V(x, t)}{\partial x} \right| \rightarrow 0$$

$$\text{and } \sup_{x \in \Omega} |u_L(x, t) - u(x, t)| \rightarrow 0.$$

Proof. This follows by noticing that $\|\mathbf{w}_L(t) - \mathbf{c}_L(t)\|_2^2 \rightarrow 0$ uniformly in t and the series with c_j is uniformly convergent in t , and Technical Lemma 2.1. ■

The final result shows that if the number L of hidden layer units is large enough, the proposed solution method yields an admissible control.

Lemma 2.9 Admissibility of $u_L(x, t)$.

If the conditions of Lemma 2.2 are satisfied, then $\exists L_0 : L \geq L_0, u_L \in \Psi(\Omega_0)$.

Proof. Define

$$V(x, u) = \phi(t_0, t_f, x(t_f), u) + \int_{t_0}^{t_f} [Q(x) + W(u)] dt$$

We must show that for L sufficiently large, $V(x, u_L) < \infty$ when $V(x, u) < \infty$. But the solution of (2-1) depends continuously on u , i.e., small variations in u result in small variations in solution of (2-1). Also since $\|u_L(\cdot)\|_{L_2(\Omega_0)}^2$ can be made arbitrarily close to $\|u(\cdot)\|_{L_2(\Omega_0)}^2$, $V(x, u_L)$ can be made arbitrarily close to $V(x, u)$. Therefore for L sufficiently large, $V(x, u_L) < \infty$ and hence $u_L(x, t)$ is admissible. ■

2.3.3. Optimal Algorithm Based on NN Approximation

Solving the integration in (2-20) is expensive computationally, since evaluation of the L_2 inner product over Ω_0 is required. This can be addressed using the collocation method [38]. The integrals can be well approximated by discretization. A mesh of points over the integration region can be introduced on Ω_0 of size Δx . The terms of (2-21) can be rewritten as follows

$$A = [\boldsymbol{\sigma}_L(x)|_{x_1} \dots \boldsymbol{\sigma}_L(x)|_{x_p}]^T, \quad (2-29)$$

$$B = [\boldsymbol{\sigma}_L(x)f(x)|_{x_1} \dots \boldsymbol{\sigma}_L(x)f(x)|_{x_p}]^T, \quad (2-30)$$

$$C = \begin{bmatrix} \frac{1}{4}(\nabla \boldsymbol{\sigma}_L(x)g(x)R^{-1}g^T(x)\nabla \boldsymbol{\sigma}_L^T(x))|_{x_1} \\ \dots \frac{1}{4}(\nabla \boldsymbol{\sigma}_L(x)g(x)R^{-1}g^T(x)\nabla \boldsymbol{\sigma}_L^T(x))|_{x_p} \end{bmatrix}^T, \quad (2-31)$$

$$D = [Q(x)|_{x_1} \dots Q(x)|_{x_p}]^T, \quad (2-32)$$

where p in x_p represents the number of points of the mesh. Reducing the mesh size, we have

$$\langle -\dot{\mathbf{w}}_L^T(t)\boldsymbol{\sigma}_L(x), \boldsymbol{\sigma}_L(x) \rangle_\Omega = \lim_{\|\Delta x\| \rightarrow 0} -(A^T A) \cdot \dot{\mathbf{w}}_L(t) \cdot \Delta x, \quad (2-33)$$

$$\langle -\mathbf{w}_L^T(t)\nabla \boldsymbol{\sigma}_L(x)f(x), \boldsymbol{\sigma}_L(x) \rangle_\Omega = \lim_{\|\Delta x\| \rightarrow 0} -(A^T B) \cdot \mathbf{w}_L(t) \cdot \Delta x, \quad (2-34)$$

$$\begin{aligned} & \left\langle \frac{1}{4}\mathbf{w}_L^T(t)\nabla \boldsymbol{\sigma}_L(x)g(x)R^{-1} \cdot g^T(x)\nabla \boldsymbol{\sigma}_L^T\mathbf{w}_L(t), \boldsymbol{\sigma}_L(x) \right\rangle_\Omega, \\ & = \lim_{\|\Delta x\| \rightarrow 0} A^T \mathbf{w}_L^T(t)C\mathbf{w}_L(t) \cdot \Delta x \end{aligned} \quad (2-35)$$

$$\langle -Q(x), \boldsymbol{\sigma}_L(x) \rangle_\Omega = \lim_{\|\Delta x\| \rightarrow 0} -(A^T \cdot D) \cdot \Delta x. \quad (2-36)$$

This implies that (2-16) can be converted to

$$\dot{\mathbf{w}}_L(t) = -(A^T A)^{-1} \mathbf{w}_L(t) A^T B + (A^T A)^{-1} A^T \mathbf{w}_L^T(t) C \mathbf{w}_L(t) - (A^T A)^{-1} A^T D. \quad (2-37)$$

This is a nonlinear ODE that can easily be integrated backwards using final condition $\mathbf{w}_L(t_f)$ to find the least-squares optimal NN weights. Then, the nearly optimal value function is given by

$$V_L(x, t) = \mathbf{w}_L^T(t) \boldsymbol{\sigma}_L(x),$$

and the nearly optimal control by

$$u_L(t) = -\frac{1}{2} R^{-1} g^T(x) \nabla \boldsymbol{\sigma}_L^T(x) \mathbf{w}_L(t). \quad (2-38)$$

Note that in practice, we use a numerically efficient least-squares relative to solve (2-37) without matrix inversion.

Remark 2.2. The closed-loop Neural Network least-squares policy gives correct answer as long as $x \in \Omega$, this control policy would be valid as long as $x(t)$ remains in Ω for all t . This means the set of initial condition Ω , which guarantees that $x(t) \in \Omega$ for all $x(t)$ is smaller than Ω itself. This can be enlarged by carefully selecting larger size of Neural Network.

2.3.4. Numerical Examples

We now show the power of our NN control technique for finding nearly optimal fixed-final time constrained controllers. Two examples are presented.

2.3.4.1. Linear System

a) We start by applying the algorithm obtained above for the linear system

$$\begin{aligned}\dot{x}_1 &= 2x_1 + 3x_2 + u_1 \\ \dot{x}_2 &= 5x_1 + 6x_2 + 2u_2\end{aligned}\quad (2-39)$$

Define performance index

$$V(x(t_0), t_0) = \frac{1}{2} x(t_f)^T S(t_f) x(t_f) + \frac{1}{2} \int_{t_0}^T (x^T Q x + u^T R u) dt,$$

Here Q and R are chosen as identity matrices. The steady-state solution of the Riccati equation can be obtained by solving the algebraic Riccati equation (ARE). The result is $\begin{bmatrix} 3.1610 & 2.8234 \\ 2.8234 & 3.6777 \end{bmatrix}$. Our algorithm should give the same steady-state value.

To find a nearly optimal time-varying controller, the following smooth function is used to approximate the value function of the system

$$V(x_1, x_2) = w_1 x_1^2 + w_2 x_1 x_2 + w_3 x_2^2.$$

This is a NN with polynomial activation functions, and hence $V(0) = 0$.

Note that if $V = x^T P x$, then $P = \begin{bmatrix} w_1 & w_2/2 \\ w_2/2 & w_3 \end{bmatrix}$.

In this example, three neurons are chosen and $\mathbf{w}_L(t_f) = [10, 10, 0]$. Our algorithm was used to determine the nearly optimal time-varying control law by backwards integrating to solve (2-37). A least-square algorithm was used to compute $\dot{\mathbf{w}}_L(t)$ at each integration time. From Figure 2-1 it is obvious that about six seconds from t_f , the weights converge to the solution of the algebraic Riccati equation. The control signal is

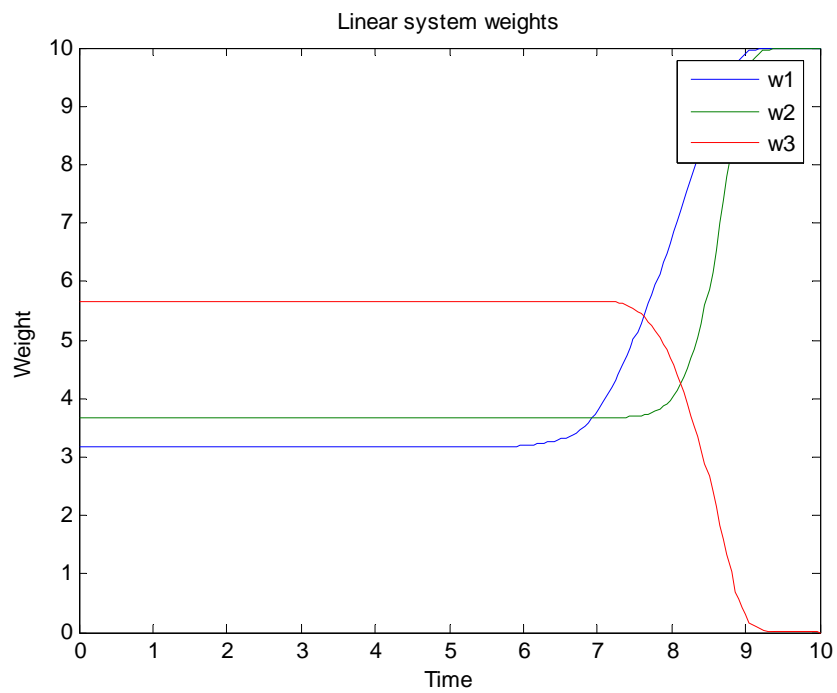


Figure 2-1 Linear System Weights

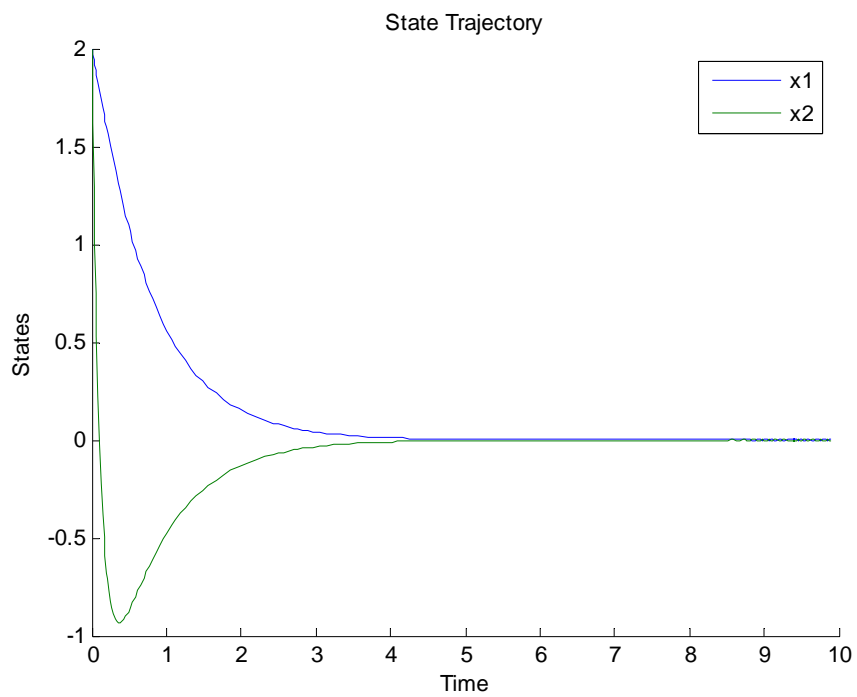


Figure 2-2 State Trajectory of Linear System

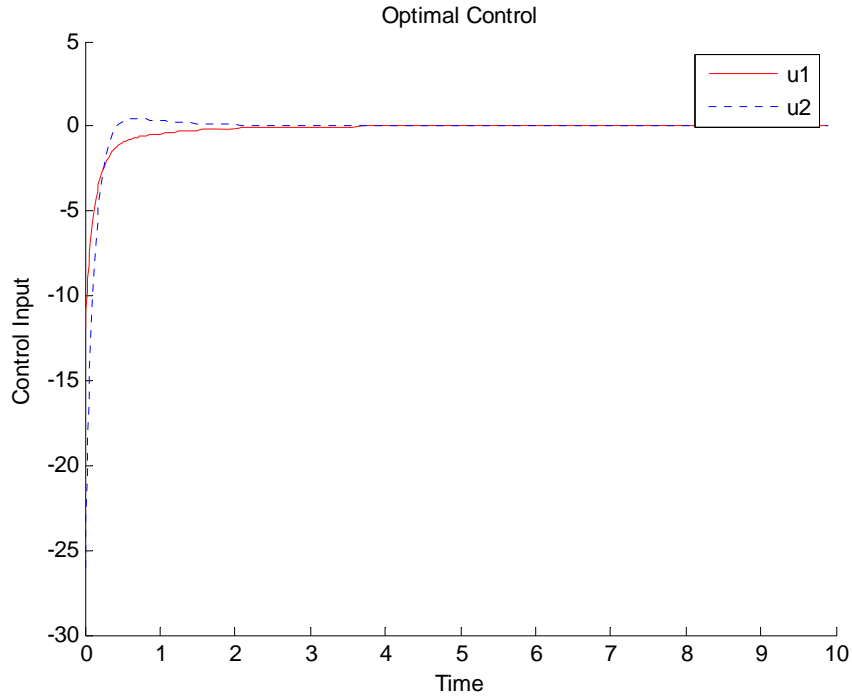


Figure 2-3 Optimal NN Control Law

$$u = -\frac{1}{2}R^{-1}g^T Px. \quad (2-40)$$

The states and control signal are shown in Figures 2-2 and 2-3.

2.3.4.2. Nonlinear Chained System

One can apply the results of this chapter to a mobile robot, which is a nonholonomic system [48]. It is known [23] that there does not exist a continuous time-invariant feedback control law that minimizes the cost. Some methods for deriving stable controls of nonholonomic systems are found in Bloch [18][19], Egeland [35], Escobar [36], Fierro and Lewis [37], Murray [66][67], Pomet [72] and Sordalen [81]. Our method will yield a time-varying gain. From Moylan [32], under some sufficient conditions, a nonholonomic system can be converted to chained form as

$$\begin{aligned}
\dot{x}_1 &= u \\
\dot{x}_2 &= v \\
\dot{x}_3 &= x_1 v
\end{aligned} \tag{2-41}$$

Define performance index

$$V(x(t_0), t_0) = \phi(x(t_f), t_f) + \int_{t_0}^t (Q(x) + W(u)) dt.$$

Here Q and R are chosen as identity matrices. To solve for the value function of the related optimal control problem, we selected the smooth approximating function

$$\begin{aligned}
V(x_1, x_2, x_3) &= w_1 x_1^2 + w_2 x_2^2 + w_3 x_3^2 + w_4 x_1 x_2 + w_5 x_1 x_3 + w_6 x_2 x_3 + w_7 x_1^4 + w_8 x_2^4 \\
&+ w_9 x_3^4 + w_{10} x_1^2 x_2^2 + w_{11} x_1^2 x_3^2 + w_{12} x_2^2 x_3^2 + w_{13} x_1^2 x_2 x_3 + w_{14} x_1 x_2^2 x_3 + w_{15} x_1 x_2 x_3^2 \\
&+ w_{16} x_1^3 x_2 + w_{17} x_1^3 x_3 + w_{18} x_1 x_2^3 + w_{19} x_1 x_3^3 + w_{20} x_2 x_3^3 + w_{21} x_2^3 x_3
\end{aligned} \tag{2-42}$$

The selection of the NN is usually a natural choice guided by engineering experience and intuition. This is a NN with polynomial activation functions, and hence $V(0) = 0$.

This is a power series NN with 21 activation functions containing powers of the state variable of the system up to the fourth order. Convergence was not observed using a NN with only second-order powers of the states. The number of neurons required is chosen to guarantee the uniform convergence of the algorithm. In this example,

$$\mathbf{w}_L(t_f) = [10; 10; 10; 0; 0; 0; 10; 10; 10; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0]$$

and $t_f = 30$ seconds.

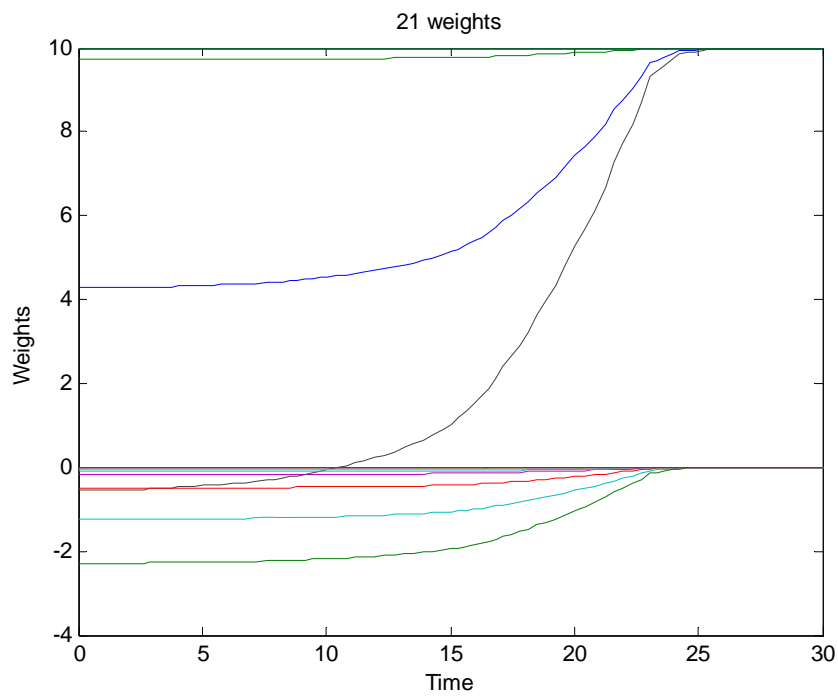


Figure 2-4 Nonlinear System Weights

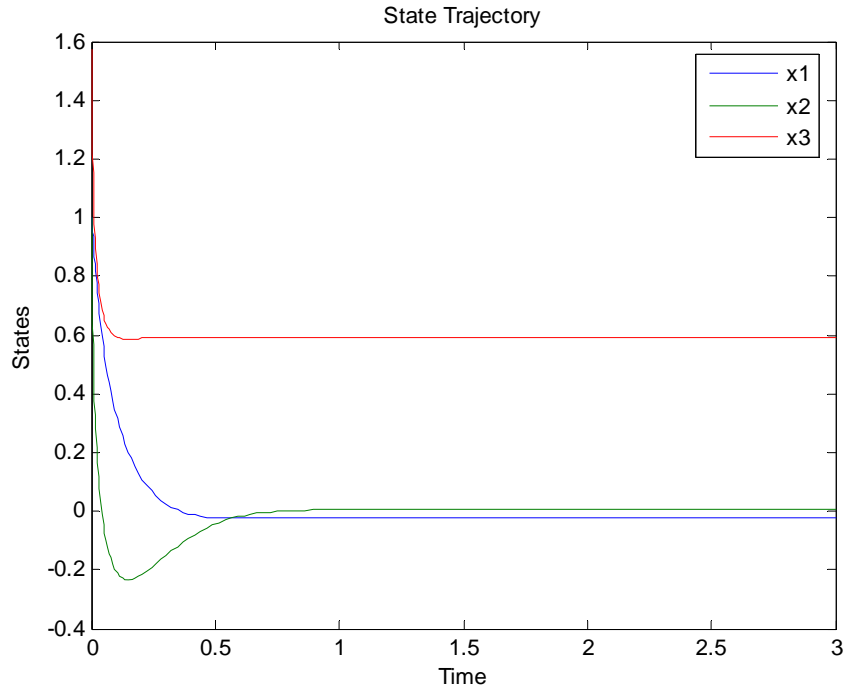


Figure 2-5 State Trajectory of Nonlinear System

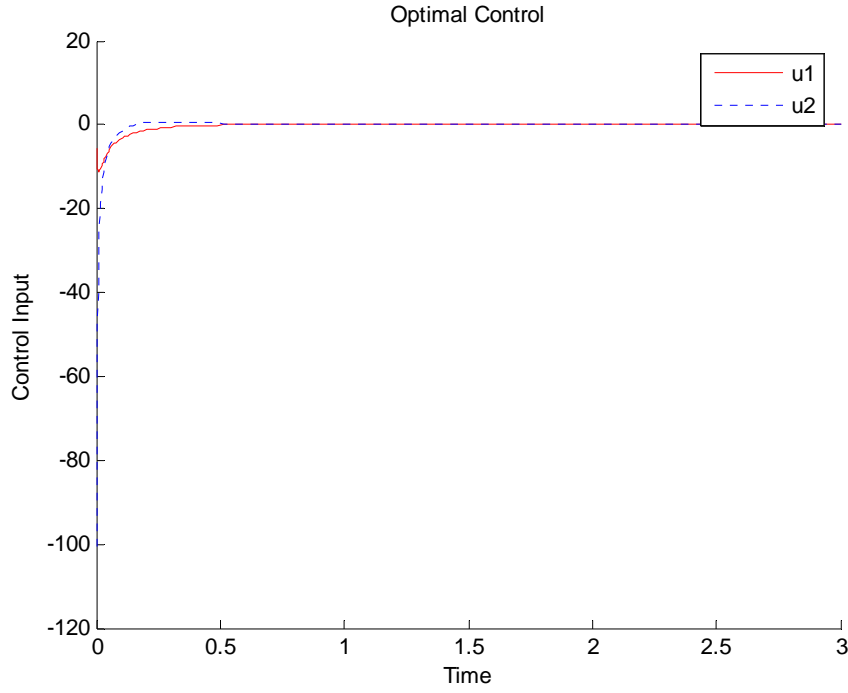


Figure 2-6 Optimal NN Control Law

Figure 2-5 indicates that weights converge to constant when they are integrated backwards. The time-varying controller (2-38) is then applied using interpolation. Figure 2-5 shows that the systems' states response, including x_1 , x_2 and x_3 are all bounded. It can be seen that the states do converge to a value close to the origin. Figure 2-6 shows the optimal control converges to zero.

2.4. Conclusion

In this chapter, optimal control of unconstrained input systems is discussed, a neural network approximation of the value function is introduced, and the method is employed in a least-squares sense over a mesh with certain size on Ω . We are able to approximately solve the time-varying HJB equation, and do not need to perform policy

iteration using the so-called GHJB equation followed by control law updates. The Galerkin integrals used in [11] are complicated to evaluate for bilinear terms.

CHAPTER 3

NEURAL NETWORK SOLUTION FOR FINITE-FINAL TIME H-INFINITY STATE FEEDBACK CONTROL

3.1. Introduction

This chapter is an extension to chapter 2, where it is shown how to use NN to approximately solve the time-varying HJB equation arising in optimal control without policy iterations. In this chapter, we present the main algorithm for the approximate solution of the HJI equation for H_∞ controllers and provide uniform convergence results and stabilities results over a Sobolev space. Finally, the resulting approach is simulated on a Rotational/Translational Actuator (RTAC) nonlinear benchmark system [85] with the relevant simulation results demonstrated. The simulation results are effective.

3.2. L_2 -gain and Dissipativity of Controlled Nonlinear Systems

Consider the following controlled nonlinear system with disturbance,

$$\begin{aligned}\dot{x} &= f(x) + g(x)u(t) + k(x)d(t) \\ y &= x \\ z &= \psi(x, u)\end{aligned}\tag{3-1}$$

$$x \in R^n, u \in U, y \in R^p, d \in R^q.$$

where $\psi(x, u)$ is such that $\|z(t)\|^2 = h(x)^T h(x) + \|u(t)\|^2$ for some function $h(x)$. This is equivalent to the absence of cross terms of x and u in other H_∞ formulations.

We further assume that $f(0)=0, \psi(0,0)=0$. Here $y(t)$ is the measured output, which we assume to be the full state vector of the system. Moreover $x=0$ is assumed to be an equilibrium point. The penalty output $z(t)$ is a function of the state and the control input $u(t)$. Note that we require the assumption that there are no cross terms of the state and the control as far as calculating $\|z(t)\|$ is concerned. The dynamics (1) are depicted in Figure 3-1.

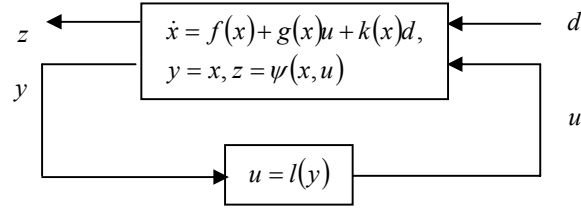


Figure 3-1. State feedback nonlinear H_∞ controller

We require the following background.

Definition 1. A closed-loop system, i.e. for some $u(t)$, system (1) has an L_2 -gain $\leq \gamma$, where $\gamma \geq 0$, if

$$\int_0^{t_f} \|z(t)\|^2 dt \leq \gamma^2 \int_0^{t_f} \|d(t)\|^2 dt$$

for all $t_f \geq 0$ and all $d \in L_2(0, t_f)$. ■

For linear systems, there are explicit formulae to compute γ^* (Chen [25]).

Throughout this chapter we shall assume that γ is fixed.

Definition 3.1. System (3-1) with supply rate $w(t) = \gamma^2 \|d(t)\|^2 - \|z(t)\|^2$ is said to be

dissipative if there exists $V \geq 0$, called the storage function, such that

$$V(x_0) + \int_{t_0}^{t_f} w(t) dt \geq V(x(t_f)) \quad \blacksquare$$

We are interested in determining a control u which under the worst of uncertainty, or disturbance d , renders the performance functional

$$V(x(t_0), t_0) = \phi(x(t_f), t_f) + \int_{t_0}^{t_f} (h^T(x)h(x) + \|u(t)\|^2 - \gamma^2 \|d(t)\|^2) dt \quad (3-2)$$

nonpositive for all $d(t) \in L_2(0, \infty)$. Note that $\phi(x(t_f), t_f) = V(x(t_f), t_f)$. In other words,

L_2 -gain $\leq \gamma$ for some prescribed γ . In terms of the storage function of the system, (3-2) becomes

$$V(x(t_f), t_f) - V(x(t_0), t_0) + \int_{t_0}^{t_f} (h^T(x)h(x) + \|u(t)\|^2 - \gamma^2 \|d(t)\|^2) dt \leq 0. \quad (3-2)$$

This also has an infinitesimal equivalence which is

$$\frac{\partial V(x, t)}{\partial t} + \frac{\partial V^T(x, t)}{\partial x} (f(x) + g(x)u + k(x)d) + \|h(t)\|^2 + \|u(t)\|^2 - \gamma^2 \|d(t)\|^2 \leq 0. \quad (3-4)$$

Definition 3.2. Admissible Controls.

A control $u(t)$ is defined to be admissible with respect to (3-3) on Ω , denoted by

$u \in \Psi(\Omega)$, if

- $u(t)$ is continuously differentiable on Ω .
- $u(0) = 0$, $u(t)$ stabilizes (1) on Ω .
- $\phi(x(t_f), t_f) + \int_{t_0}^{t_f} (h^T(x)h(x) + \|u(t)\|^2 - \gamma^2 \|d(t)\|^2) dt < \infty \quad \forall x_0 \in \Omega$.
- The L_2 gain is bounded by a prescribed value $\gamma > \gamma^*$. ■

The available storage is a result of the following optimal control problem

$$V^*(u(t), d(t)) = \min_u \max_d \left(\phi(x(t_f), t_f) + \int_{t_0}^{t_f} (h^T(x)h(x) + \|u(t)\|^2 - \gamma^2 \|d(t)\|^2) dt \right), \quad (3-5)$$

Equation (3-5) is a two-variable optimal control problem. The uniqueness of the game value of (3-5) has been demonstrated [45]. For that one uses the well-established theory of zero-sum differential games which can be interpreted as either minimax or maximin optimization respectively

$$\min_u \max_d V(x_0, u, d) \quad \text{or} \quad \max_d \min_u V(x_0, u, d). \quad (3-6)$$

For such strategy needs to be a unique saddle point of the Hamiltonian of the optimization, and the corresponding upper and lower game value needs to satisfy the same HJI equation. The optimal control solution is unique if

$$V(x_0, u^*, d) \leq V(x_0, u^*, d^*) \leq V(x_0, u, d^*). \quad (3-7)$$

This is equivalent to

$$\max_d \min_u V(x_0, u, d) = \min_u \max_d V(x_0, u, d). \quad (3-8)$$

The pair (u^*, d^*) that satisfies (3-7) is called a game-theoretic saddle point.

Define the Hamiltonian function

$$\begin{aligned} H(x, p, u, d) \\ = \frac{\partial V^T(x, t)}{\partial x} (f(x) + g(x)u(t) + k(x)d(t)) + h^T(x)h(x) + \|u(t)\|^2 - \gamma^2 \|d(t)\|^2. \end{aligned} \quad (3-9)$$

The first-order necessary conditions that follow from stationarity for this optimization problem are

$$\frac{\partial H}{\partial u} = 0, \quad \frac{\partial H}{\partial d} = 0. \quad (3-10)$$

Minimizing the Hamiltonian of the optimal control problem with regard to u and d gives

$$g^T(x) \frac{\partial V(x,t)}{\partial x} + 2u^*(t) = 0, \quad (3-11)$$

$$k^T(x) \frac{\partial V(x,t)}{\partial x} - 2\gamma^2 d^*(t) = 0, \quad (3-12)$$

so

$$u^*(t) = -\frac{1}{2} g(x)^T \frac{\partial V(x,t)}{\partial x}, \quad (3-13)$$

$$d^*(t) = \frac{1}{2\gamma^2} k(x)^T \frac{\partial V(x,t)}{\partial x}. \quad (3-14)$$

In order to achieve (3-7), we need to have $H_1^* = H_2^*$, where $H_1^* = \max_d \min_u H$, and

$$H_2^* = \min_u \max_d H.$$

Second-order necessary conditions are

$$\frac{\partial^2 H}{\partial u^2} \geq 0, \quad \frac{\partial^2 H}{\partial d^2} \leq 0. \quad (3-15)$$

These conditions become sufficient when they are replaced with strict inequalities. This is equivalent to

$$H(x_0, u^*, d) \leq H(x_0, u^*, d^*) \leq H(x_0, u, d^*). \quad (3-16)$$

For finite time problem, a saddle point in the Hamiltonian H implies a saddle point in the performance $V(x, t)$.

When $H^*(x, p, u^*, d^*) = 0$, we have the Hamilton-Jacobi-Isaacs equation

$$\begin{aligned} & \frac{\partial V(x,t)}{\partial t} + \frac{\partial V^T(x,t)}{\partial x} (f(x) + g(x)u^*(t) + k(x)d^*(t)) \\ & + h^T(x)h(x) + \|u^*(t)\|^2 - \gamma^2 \|d^*(t)\|^2 = 0 \end{aligned} \quad (3-17)$$

From (3-13) and (3-14), (3-17) becomes

$$\begin{aligned} HJI(V^*(x,t)) &= \frac{\partial V^*(x,t)}{\partial t} + \frac{\partial V^{*T}(x,t)}{\partial x} f(x) \\ & - \frac{\partial V^{*T}(x,t)}{\partial x} \hat{g}(x)\hat{g}(x)^T \frac{\partial V^*(x,t)}{\partial x} + h^T(x)h(x) = 0 \end{aligned} \quad (3-18)$$

with boundary condition $V(x(t_f), t_f) = \phi(x(t_f), t_f)$.

Here $\hat{g}(x)\hat{g}(x)^T = \frac{1}{4} g(x)g(x)^T - \frac{1}{4\gamma^2} k(x)k(x)^T$.

Equations (3-13) (3-14) and (3-18) provide the solution to finite-horizon optimal control for general nonlinear systems. However, equation (3-18) is generally impossible to solve for nonlinear systems.

3.3. NN Least-Squares Approximate HJI Solution

Now we use unconstrained case for NN approximation. The HJI equation (3-18) is difficult to solve for the cost function $V(x,t)$. In this section, NNs are used to solve approximately for the value function in (3-18) over Ω by approximating the cost function to find nearly optimal H_∞ state feedback controllers for nonlinear systems.

3.3.1. NN Approximation of $V(x,t)$

In chapter 2 it is noted that

$$\frac{\partial V_L(x,t)}{\partial x} = \frac{\partial \sigma_L^T(x)}{\partial x} \mathbf{w}_L(t) \equiv \nabla \sigma_L^T(x) \mathbf{w}_L(t), \quad (3-19)$$

where $\nabla \boldsymbol{\sigma}_L(x)$ is the Jacobian $\frac{\partial \boldsymbol{\sigma}_L(x)}{\partial x}$, and that $\frac{\partial V_L(x,t)}{\partial t} = \dot{\mathbf{w}}_L^T(t) \boldsymbol{\sigma}_L(x)$.

For the $HJI(V(x,t), d(t)) = 0$, the solution $V(x,t)$ is replaced with $V_L(x,t)$ having a residual error

$$\begin{aligned} & \dot{\mathbf{w}}_L^T(t) \boldsymbol{\sigma}_L(x) + \mathbf{w}_L^T(t) \nabla \boldsymbol{\sigma}_L(x) f(x) \\ & - \mathbf{w}_L^T(t) \boldsymbol{\sigma}_L(x) \hat{g}(x) \hat{g}^T(x) \boldsymbol{\sigma}_L^T(x) \mathbf{w}_L(t) + h^T(x) h(x), \\ & = e_L(x, t) \end{aligned} \quad (3-20)$$

or

$$HJI\left(V_L(x, t) = \sum_{j=1}^L w_j(t) \sigma_j(x), d(t)\right) = e_L(x, t), \quad (3-21)$$

where $e_L(x, t)$ is a residual equation error. The weight derivatives $\dot{\mathbf{w}}_L(t)$ are determined by projecting the residual error onto $\frac{\partial e_L(x, t)}{\partial \dot{\mathbf{w}}_L(t)}$ and setting the result to

zero $\forall x \in \Omega$ using the inner product, i.e.

$$\left\langle \frac{\partial e_L(x, t)}{\partial \dot{\mathbf{w}}_L(t)}, e_L(x, t) \right\rangle_{\Omega} = 0. \quad (3-22)$$

From (3-20) we can get

$$\frac{\partial e_L(x, t)}{\partial \dot{\mathbf{w}}_L} = \boldsymbol{\sigma}_L(x). \quad (3-23)$$

Therefore one obtains

$$\begin{aligned} & \left\langle \dot{\mathbf{w}}_L^T(t) \boldsymbol{\sigma}_L(x), \boldsymbol{\sigma}_L(x) \right\rangle_{\Omega} + \left\langle \mathbf{w}_L^T(t) \nabla \boldsymbol{\sigma}_L(x) f(x), \boldsymbol{\sigma}_L(x) \right\rangle_{\Omega} \\ & - \left\langle \mathbf{w}_L^T(t) \nabla \boldsymbol{\sigma}_L(x) \hat{g}(x) \hat{g}^T(x) \nabla \boldsymbol{\sigma}_L^T(x) \mathbf{w}_L(t), \boldsymbol{\sigma}_L(x) \right\rangle_{\Omega} \\ & + \left\langle h^T(x) h(x), \boldsymbol{\sigma}_L(x) \right\rangle_{\Omega} = 0 \end{aligned} \quad (3-24)$$

So that

$$\begin{aligned}
\dot{\mathbf{w}}_L(t) = & \\
& - \langle \sigma_L(x), \sigma_L(x) \rangle_{\Omega}^{-1} \cdot \langle \nabla \sigma_L(x) f(x), \sigma_L(x) \rangle_{\Omega} \cdot \mathbf{w}_L(t) \\
& + \langle \sigma_L(x), \sigma_L(x) \rangle_{\Omega}^{-1} \cdot \langle \mathbf{w}_L^T(t) \nabla \sigma_L(x) \hat{g}(x) \hat{g}^T(x) \nabla \sigma_L^T \mathbf{w}_L(t), \sigma_L(x) \rangle_{\Omega} \cdot \\
& - \langle \sigma_L(x), \sigma_L(x) \rangle_{\Omega}^{-1} \cdot \langle h^T(x) h(x), \sigma_L(x) \rangle_{\Omega} \cdot \mathbf{w}_L(t)
\end{aligned} \tag{3-25}$$

with boundary condition $V(x(t_f), t_f) = \phi(x(t_f), t_f)$.

Therefore, the NN weights are simply found by integrating this nonlinear ODE backwards in time. In practice, one does not invest $\langle \sigma_L(x), \sigma_L(x) \rangle_{\Omega}$, but uses efficient least-square methods to solve (3-24).

We now show that this procedure provides a nearly optimal solution for the time-varying optimal control problem if time-varying L is selected large enough.

3.3.2. Convergence of the Method of Least-Squares

In what follows, we show convergence results as L increases for the method of least-squares when NN are used to approximate the cost of function.

Lemma 3.1. Convergence of Approximate HJI Equation.

Given $u \in \mathcal{U}(\Omega)$. Let $V_L(x, t) = \sum_{j=1}^L w_j^T(t) \sigma_j(x)$ satisfy

$$\langle HJI(V_L(x, t)), \sigma_L(x) \rangle_{\Omega} = 0 \quad \text{and} \quad \langle V_L(t_f), \sigma_L(x) \rangle_{\Omega} = 0, \quad \text{let} \quad V(x, t) = \sum_{j=1}^{\infty} c_j^T(t) \sigma_j(x)$$

and note that $HJI(V(x, t)) = 0$ and $V(x(t_f), t_f) = \phi(x(t_f), t_f)$.

Define $\mathbf{c}_L(t) \equiv [c_1(t)c_2(t) \dots c_L(t)]^T$. If Ω is compact, $Q(x)$ are continuous on Ω and are in the space $\text{span}\{\sigma_j(x)\}_{j=1}^\infty$, and if the coefficients $|w_j(t)|$ are uniformly bounded for all L , then

$\|HJI(V_L(x, t))\| \rightarrow 0$ on Ω as L increases.

Proof. The hypothesis implies that $HJI(V_L(x, t)) \in \text{span}\{\sigma_j(x)\}_{j=1}^\infty$. Note that

$$\begin{aligned} \langle HJI(V_L(x, t)), \sigma_j(x) \rangle_\Omega = & \sum_{k=1}^L \dot{w}_k^T(t) \langle \sigma_k(x), \sigma_j(x) \rangle_\Omega + \sum_{k=1}^L w_k^T(t) \langle \nabla \sigma_k(x) f(x), \sigma_j(x) \rangle_\Omega \\ & - \sum_{k=1}^L w_k^T(t) \cdot \langle \nabla \sigma_k(x) \hat{g}(x) \hat{g}^T(x) \nabla \sigma_k^T(x), \sigma_j(x) \rangle_\Omega \cdot w_k(t) + \langle h^T(x) h(x), \sigma_j(x) \rangle_\Omega \end{aligned} \quad (3-26)$$

Then

$$\begin{aligned} \|HJI(V_L(x, t))\| &= \left\| \sum_{j=1}^\infty \langle HJI(V_L(x, t)), \sigma_j(x) \rangle_\Omega \sigma_j(x) \right\| \\ &= \left\| \sum_{j=L+1}^\infty \left(\sum_{k=1}^L \dot{w}_k^T(t) \langle \sigma_k(x), \sigma_j(x) \rangle_\Omega \right) \cdot \sigma_j(x) \right. \\ &\quad + \sum_{j=L+1}^\infty \left(\sum_{k=1}^L w_k^T(t) \langle \nabla \sigma_k(x) f(x), \sigma_j(x) \rangle_\Omega \right) \cdot \sigma_j(x) \\ &\quad - \sum_{j=L+1}^\infty \left(\sum_{k=1}^L w_k^T(t) \cdot \langle \nabla \sigma_k(x) \hat{g}(x) \hat{g}^T(x) \nabla \sigma_k^T(x), \sigma_j(x) \rangle_\Omega \cdot w_k(t) \right) \cdot \sigma_j(x) \\ &\quad \left. + \sum_{j=L+1}^\infty \langle h^T(x) h(x), \sigma_j(x) \rangle_\Omega \cdot \sigma_j(x) \right\|. \end{aligned} \quad (3-27)$$

Since the set $\{\sigma_j(x)\}_{j=1}^\infty$ are orthogonal, $\langle \sigma_k(x), \sigma_j(x) \rangle_\Omega = 0$.

Therefore

$$\begin{aligned} &= \left\| \sum_{k=1}^L w_k(t) \sum_{j=L+1}^\infty \langle \nabla \sigma_k(x) f(x), \sigma_j(x) \rangle_\Omega \cdot \sigma_j(x) \right\| \\ &+ \left\| \sum_{k=1}^L w_k^2(t) \cdot \sum_{j=L+1}^\infty \langle \nabla \sigma_k(x) \hat{g}(x) \cdot \hat{g}^T(x) \nabla \sigma_k^T(x), \sigma_j(x) \rangle_\Omega \cdot \sigma_j(x) \right\| \\ &+ \left\| \sum_{j=L+1}^\infty \langle h^T(x) h(x), \sigma_j(x) \rangle_\Omega \sigma_j(x) \right\| \end{aligned}$$

$$\leq AB(x) + CD(x) + Vec \left(\sum_{j=L+1}^{\infty} \langle h^T(x)h(x), \sigma_j(x) \rangle_{\Omega} \sigma_j(x) \right), \quad (3-28)$$

where

$$A = \max_{1 \leq k \leq L} |w_k(t)|,$$

$$B(x) = \sup_{(t,x) \in [t_0, T] \times \Omega} \left| \sum_{k=1}^L \left(\sum_{j=L+1}^{\infty} \langle \nabla \sigma_k(x) f(x), \sigma_j(x) \rangle_{\Omega} \right) \cdot \sigma_j(x) \right|,$$

$$C = \max_{1 \leq k \leq L} |w_k^2(t)|,$$

$$D = \sup_{(t,x) \in [t_0, T] \times \Omega} \left| \left(\sum_{k=1}^L \left(\sum_{j=L+1}^{\infty} \langle \nabla \sigma_k(x) \hat{g}(x) \hat{g}^T(x) \nabla \sigma_k^T(x), \sigma_j(x) \rangle_{\Omega} \right) \cdot \sigma_j(x) \right) \right|$$

Suppose $\nabla \sigma_k(x) f(x)$, $\nabla \sigma_k(x) \hat{g}(x) \hat{g}^T(x) \nabla \sigma_k^T(x)$ and $h^T(x)h(x)$ are in $L_2(\Omega)$, the orthonormality of the set $\{\sigma_j(x)\}_1^{\infty}$ implies that $B(x)$, $D(x)$ and the third term on the right-hand side can be made arbitrarily small by an appropriate choice of L . Therefore

$$A \cdot B(x) + C \cdot D(x) \rightarrow 0 \quad \text{and} \quad \left| \sum_{j=L+1}^{\infty} \langle h^T(x)h(x), \sigma_j(x) \rangle_{\Omega} \sigma_j(x) \right| \rightarrow 0.$$

So $|HJI(V_L(x, t))| \rightarrow 0$. ■

Lemma 3.2. Convergence of NN Weights.

Given $u \in \Psi(\Omega)$ and suppose the hypothesis of Lemma 3.1 hold. Then

$$\|\mathbf{w}_L(t) - \mathbf{c}_L(t)\|_2 \rightarrow 0 \quad \text{as } L \text{ increases.}$$

Proof: Define $e_L(x, t) = HJI(V_L(x, t))$ and

$$\hat{e}_L(x, t) = V_L(x(t_f), t_f) - \phi(x(t_f), t_f). \quad (3-29)$$

Then $\langle e_L(x, t), \sigma_L(x) \rangle_{\Omega} = \langle \hat{e}_L(x, t), \sigma_L(x) \rangle_{\Omega} = 0$. From the hypothesis we have that

$$\begin{aligned} HJI(V_L(x, t)) - HJI(V(x, t)) &= e_L(x, t) \\ V_L(x(t_f), t_f) - V(x(t_f), t_f) &= \hat{e}_L(x, t) \end{aligned} \quad (3-30)$$

substituting the series expansion for $V_L(x, t)$ and $V(x, t)$, and moving the terms in the series that are greater than L to the right-hand side we obtain

$$\begin{aligned} & (\dot{\mathbf{w}}_L(t) - \dot{\mathbf{c}}_L(t))^T \boldsymbol{\sigma}_L(x) + (\mathbf{w}_L(t) - \mathbf{c}_L(t))^T \nabla \boldsymbol{\sigma}_L(x) f(x) \\ & - (\mathbf{w}_L^T(t) \otimes \mathbf{w}_L^T(t) - \mathbf{c}_L^T(t) \otimes \mathbf{c}_L^T(t)) \cdot \text{Vec}(\nabla \boldsymbol{\sigma}_L(x) \hat{g}(x) \hat{g}^T(x) \nabla \boldsymbol{\sigma}_L^T(x)) \\ & = e_L(x, t) + \sum_{j=L+1}^{\infty} \dot{c}_j^T(t) \sigma_j(x) + \sum_{j=L+1}^{\infty} c_j^T(t) \nabla \sigma_j(x) f(x) \\ & - \sum_{j=L+1}^{\infty} c_j^2(t) (\nabla \boldsymbol{\sigma}_j(x) \hat{g}(x) \hat{g}^T(x) \nabla \boldsymbol{\sigma}_j^T(x)) \\ & (\mathbf{w}_L^T(t_f) - \mathbf{c}_L^T(t_f)) \boldsymbol{\sigma}_L = \hat{e}_L(x, t) + \sum_{j=L+1}^{\infty} c_j(t_f) \sigma_j(x). \end{aligned} \quad (3-31)$$

Taking the inner product of both sides over Ω , and taking into account the orthonormality of the set $\{\sigma_j(x)\}_1^{\infty}$, we obtain [22]

$$\begin{aligned} & (\dot{\mathbf{w}}_L(t) - \dot{\mathbf{c}}_L(t)) + \langle \nabla \boldsymbol{\sigma}_L(x) f(x), \boldsymbol{\sigma}_L(x) \rangle_{\Omega}^T (\mathbf{w}_L(t) - \mathbf{c}_L(t)) \\ & - \langle \text{Vec}(\nabla \boldsymbol{\sigma}_L(x) \hat{g}(x) \hat{g}^T(x) \nabla \boldsymbol{\sigma}_L^T(x)), \boldsymbol{\sigma}_L(x) \rangle_{\Omega}^T (\mathbf{w}_L(t) \otimes \mathbf{w}_L(t) - \mathbf{c}_L(t) \otimes \mathbf{c}_L(t)) \\ & = \sum_{j=L+1}^{\infty} c_j(t) \cdot \langle \nabla \sigma_j(x) f(x), \boldsymbol{\sigma}_L(x) \rangle_{\Omega}^T - \sum_{j=L+1}^{\infty} c_j^2(t) \cdot \langle \nabla \sigma_j(x) \hat{g}(x) \hat{g}^T(x) \nabla \boldsymbol{\sigma}_j^T(x), \boldsymbol{\sigma}_L(x) \rangle_{\Omega}^T \\ & \mathbf{w}_L(t_f) - \mathbf{c}_L(t_f) = 0. \end{aligned} \quad (3-32)$$

Let $A = \langle \nabla \boldsymbol{\sigma}_L(x) f(x), \boldsymbol{\sigma}_L(x) \rangle_{\Omega}^T$, where A is scalar.

Define $\xi = \mathbf{w}_L(t) - \mathbf{c}_L(t)$, consider the equation

$$\begin{aligned} \dot{\xi} + A \cdot \xi + f(\xi, t) &= 0 \\ \xi(t_f) &= 0 \end{aligned} \quad (3-33)$$

where

$$f(\xi, t) = -\langle \text{Vec}(\nabla \boldsymbol{\sigma}_L(x) \hat{g}(x) \hat{g}^T(x) \nabla \boldsymbol{\sigma}_L^T(x)), \boldsymbol{\sigma}_L(x) \rangle_{\Omega}^T (\mathbf{w}_L(t) \otimes \mathbf{w}_L(t) - \mathbf{c}_L(t) \otimes \mathbf{c}_L(t))$$

is continuously differentiable in a neighborhood of a point (ξ_0, t_0) . Since this is an ordinary differential equation, satisfying a local Lipschitz condition [46], it has a unique solution, namely $\xi(t) = 0, \forall t \in [t_0, t_f]$. Noting that

$$\sum_{j=L+1}^{\infty} c_j(t) \cdot \langle \nabla \sigma_j(x) f(x), \sigma_L(x) \rangle_{\Omega}^T - \sum_{j=L+1}^{\infty} c_j^2(t) \cdot \langle \nabla \sigma_j(x) \hat{g}(x) \hat{g}^T(x) \nabla \sigma_j^T(x), \sigma_L(x) \rangle_{\Omega}^T$$

is continuous in t , we invoke the standard result from the theory of ordinary differential equations that a continuous perturbation in the system equations and the initial state implies a continuous perturbation of the solution (Arnold [4]). This implies that for all $\varepsilon > 0$, there exists a $\rho(t) > 0$ such that $\forall t \in [t_0, t_f]$,

$$\begin{aligned} & \left\| \sum_{j=L+1}^{\infty} c_j(t) \cdot \langle \nabla \sigma_j(x) f(x), \sigma_L(x) \rangle_{\Omega}^T - \sum_{j=L+1}^{\infty} c_j^2(t) \cdot \langle \nabla \sigma_j(x) \hat{g}(x) \hat{g}^T(x) \nabla \sigma_j^T(x), \sigma_L(x) \rangle_{\Omega}^T \right\|_{L_2(\Omega)} \\ & \leq \left\| \sum_{j=L+1}^{\infty} c_j(t) \cdot \langle \nabla \sigma_L(x) f(x), \sigma_L(x) \rangle_{\Omega}^T \right\|_{L_2(\Omega)} + \left\| \sum_{j=L+1}^{\infty} c_j^2(t) \cdot \langle \nabla \sigma_j(x) \hat{g}(x) \hat{g}^T(x) \nabla \sigma_j^T(x), \sigma_j(x) \rangle_{\Omega} \right\|_{L_2(\Omega)} \\ & \leq \rho(t) \\ & \Rightarrow \|\mathbf{w}_L(t) - \mathbf{c}_L(t)\|_2 < \varepsilon. \end{aligned} \tag{3-34}$$

So $\|\mathbf{w}_L(t) - \mathbf{c}_L(t)\|_2 \rightarrow 0$. ■

Lemma 3.3. Convergence of Disturbance

If the conditions of Lemma 3.1 are satisfied and

$$d_L(t) = \frac{1}{2\gamma^2} k(x)^T \frac{\partial V_L(x, t)}{\partial x},$$

$$d(t) = \frac{1}{2\gamma^2} k(x)^T \frac{\partial V(x, t)}{\partial x}.$$

Then $\|d_L(x) - d(x)\|_{L_2(\Omega)} \rightarrow 0$ on Ω as L increases.

Proof.

$$\begin{aligned} & \|d_L(t) - d(t)\|_{L_2(\Omega)} \\ & \leq \left\| \frac{1}{2\gamma^2} k^T(x) \nabla \sigma_L^T(x) (\mathbf{w}_L(t) - \mathbf{c}_L(t)) \right\|_{L_2(\Omega)} + \left\| \frac{1}{2\gamma^2} \sum_{j=L+1}^{\infty} c_j(t) k^T(x) \nabla \sigma_j(x) \right\|_{L_2(\Omega)}, \end{aligned}$$

So $d(t) = \frac{1}{2\gamma^2} \sum_{j=1}^{\infty} c_j(t) k^T(x) \nabla \sigma_j(x)$ implies that the second term on the right-hand side

converges to 0. By Lemma 3.2 and 4, we know that

$$\left\| \nabla \sigma_L^T(x) (\mathbf{w}_L(t) - \mathbf{c}_L(t)) \right\|_{L_2(\Omega)} \rightarrow 0.$$

Since $k^T(x)$ is continuous on $\Omega \times [t_0, t_f]$ and hence uniformly bounded, we have that

$$\left\| k^T(x) \nabla \sigma_L^T(x) (\mathbf{w}_L(t) - \mathbf{c}_L(t)) \right\|_{L_2(\Omega)} \rightarrow 0. \quad \blacksquare$$

At this point we have proven convergence in the mean of the approximate HJI equation, the NN weights, the approximate value function, the value function gradient and control inputs are proved in chapter 2. This demonstrates convergence in the mean in Sobolev space $H^{1,2}(\Omega)$. In fact, the next result shows even stronger convergence properties.

Lemma 3.4. Uniform Convergence.

Since a local Lipschitz condition holds on (3-25), then

$$\sup_{x \in \Omega} |V_L(x, t) - V(x, t)| \rightarrow 0, \quad \sup_{x \in \Omega} \left| \frac{\partial V_L(x, t)}{\partial x} - \frac{\partial V(x, t)}{\partial x} \right| \rightarrow 0,$$

$$\sup_{x \in \Omega} |u_L(t) - u(t)| \rightarrow 0, \quad \sup_{x \in \Omega} |d_L(t) - d(t)| \rightarrow 0$$

Proof. This follows by noticing that $\|\mathbf{w}_L(t) - c_L(t)\|_2^2 \rightarrow 0$ and the series with $c_j(t)$ is uniformly convergent, and [42]. ■

The final result shows that if the number L of hidden layer units is large enough, the proposed solution method yields an admissible control.

Lemma 3.5. Admissibility of $u_L(t)$ and $d_L(t)$

If the conditions of Lemma 3.1 are satisfied, then $\exists L_0 : L \geq L_0, u_L(t) \in \Psi(\Omega), d_L(t) \in \Psi(\Omega)$.

Proof. Define

$$V(x(t_0), t_0) = \phi(x(t_f), t_f) + \int_{t_0}^{t_f} (h^T(x)h(x) + \|u(t)\|^2 - \gamma^2 \|d(t)\|^2) dt.$$

We must show that for L sufficiently large, $V(x, u_L) < \infty$ when $V(x, u) < \infty$. But $\phi(x(t_f), t_f)$ depends continuously on W , i.e., small variations in W result in small variations in ϕ . Also since $\|u_L(\cdot)\|_{L_2(\Omega)}^2$ can be made arbitrarily close to $\|u(\cdot)\|_{L_2(\Omega)}^2$ and $\|d_L(\cdot)\|_{L_2(\Omega)}^2$ can be made arbitrarily close to $\|d(\cdot)\|_{L_2(\Omega)}^2$, $V(x, u_L)$ can be made arbitrarily close to $V(x, u)$. Therefore for L sufficiently large, $V(x, u_L) < \infty$ and hence $u_L(t)$ and $d_L(t)$ are admissible. ■

3.3.3. Optimal Algorithm Based on NN Approximation

Solving the integration in (3-24) is expensive computationally. Since evaluation of the L_2 inner product over Ω is required. This can be addressed using the collocation method [38]. The integrals can be well approximated by discretization. A mesh of points over the integration region can be introduced on Ω of size Δx . The terms of (3-24) can be rewritten as follows

$$A = [\sigma_L(x)|_{x_1} \dots \sigma_L(x)|_{x_p}]^T, \quad (3-35)$$

$$B = [\sigma_L(x)f(x)|_{x_1} \dots \sigma_L(x)f(x)|_{x_p}]^T, \quad (3-36)$$

$$C = [(\nabla \sigma_L(x) \hat{g}(x) \hat{g}^T(x) \nabla \sigma_L^T(x))|_{x_1} \dots (\nabla \sigma_L(x) \hat{g}(x) \hat{g}^T(x) \nabla \sigma_L^T(x))|_{x_p}]^T, \quad (3-37)$$

$$D = [h^T(x)h(x)|_{x_1} \dots h^T(x)h(x)|_{x_p}]^T \quad (3-38)$$

where p in x_p represents the number of points of the mesh. Reducing the mesh size, we have

$$\langle -\dot{\mathbf{w}}_L^T(t) \sigma_L(x), \sigma_L(x) \rangle_\Omega = \lim_{\|\Delta x\| \rightarrow 0} - (A^T A) \cdot \dot{\mathbf{w}}_L(t) \cdot \Delta x, \quad (3-39)$$

$$\langle -\mathbf{w}_L^T(t) \nabla \sigma_L(x) f(x), \sigma_L(x) \rangle_\Omega = \lim_{\|\Delta x\| \rightarrow 0} - (A^T B) \cdot \mathbf{w}_L(t) \cdot \Delta x, \quad (3-40)$$

$$\langle \mathbf{w}_L^T(t) \nabla \sigma_L(x) \hat{g}(x) \hat{g}^T(x) \nabla \sigma_L^T(t) \mathbf{w}_L(t), \sigma_L(x) \rangle_\Omega = \lim_{\|\Delta x\| \rightarrow 0} A^T \mathbf{w}_L^T(t) C \mathbf{w}_L(t) \cdot \Delta x, \quad (3-41)$$

$$\langle -h^T(x)h(x), \sigma_L(x) \rangle_\Omega = \lim_{\|\Delta x\| \rightarrow 0} - (A^T \cdot D) \cdot \Delta x. \quad (3-42)$$

This implies that (3-24) can be converted to

$$-A^T A \cdot \dot{\mathbf{w}}_L(t) - A^T B \cdot \mathbf{w}_L(t) + A^T \mathbf{w}_L^T(t) C \mathbf{w}_L(t) - A^T D = 0, \quad (3-43)$$

$$\dot{\mathbf{w}}_L(t) = -\left(A^T A\right)^{-1} \mathbf{w}_L(t) A^T B + \left(A^T A\right)^{-1} A^T \mathbf{w}_L^T(t) C \mathbf{w}_L(t) - \left(A^T A\right)^{-1} A^T D. \quad (3-44)$$

This is a nonlinear ODE that can easily be integrated backwards using final condition $\mathbf{w}_L(t_f)$ to find the least-squares optimal NN weights.

3.4. Simulation-Benchmark Problem

In this example, we will show the power of our NN control technique for finding nearly optimal finite-horizon H_∞ state feedback controller for the Rotational/Translational Actuator shown in Figure 3-2. This was defined as benchmark problem in [24].

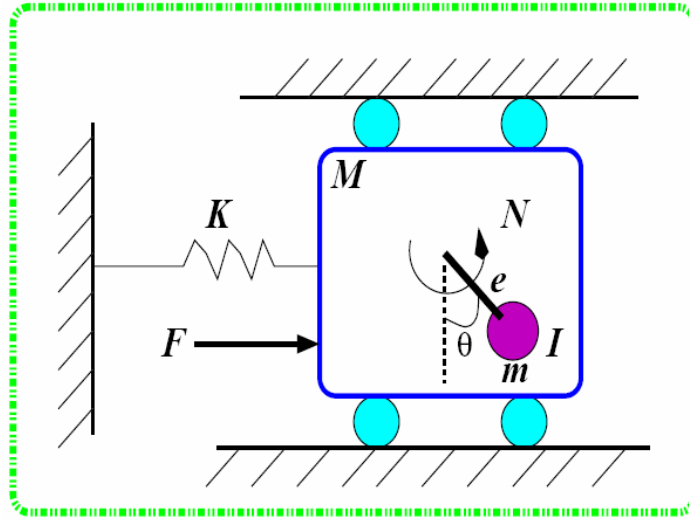


Figure 3-2 Rotational actuator to control a translational oscillator

$$\dot{x} = f(x) + g(x)u + k(x)d(t), \quad |u| \leq 2$$

$$z^T z = x_1^2 + 0.1x_2^2 + 0.1x_3^2 + 0.1x_4^2 + \|u(t)\|_q^2,$$

$$\varepsilon = \frac{me}{\sqrt{(I + me^2)(M + m)}} = 0.2, \quad \gamma = 10,$$

$$f = \begin{bmatrix} x_2 & \frac{-x_1 + \varepsilon x_4^2 \sin x_3}{1 - \varepsilon^2 \cos^2 x_3} & x_4 & \frac{\varepsilon \cos x_3 (x_1 - \varepsilon x_4^2 \sin x_3)}{1 - \varepsilon^2 \cos^2 x_3} \end{bmatrix}^T$$

$$g = \begin{bmatrix} 0 & \frac{-\varepsilon \cos x_3}{1 - \varepsilon^2 \cos^2 x_3} & 0 & \frac{1}{1 - \varepsilon^2 \cos^2 x_3} \end{bmatrix}^T,$$

$$k = \begin{bmatrix} 0 & \frac{1}{1 - \varepsilon^2 \cos^2 x_3} & 0 & \frac{-\varepsilon \cos x_3}{1 - \varepsilon^2 \cos^2 x_3} \end{bmatrix}^T.$$

Here the state x_1 and x_2 are the normalized distance r and velocity of the cart \dot{r} ,
 $x_3 = \theta$, $x_4 = \dot{\theta}$.

Define performance index

$$V(x(t_0), t_0) = \phi(x(t_f), t_f) + \int_{t_0}^{t_f} (h^T(x)h(x) + \|u(t)\|^2 - \gamma^2 \|d(t)\|^2) dt.$$

To solve for the value function of the related optimal control problem, we selected the smooth approximating function

$$\begin{aligned} V(x_1, x_2, x_3, x_4) = & w_1 x_1^2 + w_2 x_1 x_2 + w_3 x_1 x_3 + w_4 x_1 x_4 + w_5 x_2^2 + w_6 x_2 x_3 \\ & + w_7 x_2 x_4 + w_8 x_3^2 + w_9 x_3 x_4 + w_{10} x_4^2 + w_{11} x_1^4 + w_{12} x_1^3 x_2 + w_{13} x_1^3 x_3 \\ & + w_{14} x_1^3 x_4 + w_{15} x_1^2 x_2^2 + w_{16} x_1^2 x_2 x_3 + w_{17} x_1^2 x_2 x_4 + w_{18} x_1^2 x_3^2 + w_{19} x_1^2 x_3 x_4 \\ & + w_{20} x_1^4 x_2^2 + w_{21} x_1 x_2^3 + w_{22} x_1 x_2^2 x_3 + w_{23} x_1 x_2^2 x_4 + w_{24} x_1 x_2 x_3^2 + w_{25} x_1 x_2 x_3 x_4 \\ & + w_{26} x_1 x_2 x_4^2 + w_{27} x_1 x_3^3 + w_{28} x_1 x_3^2 x_4 + w_{29} x_1 x_3 x_4^2 + w_{30} x_1 x_4^3 w_{31} x_2^4 + w_{32} x_2^3 x_3 \\ & + w_{33} x_2^3 x_4 + w_{34} x_2^2 x_3^2 + w_{35} x_2^2 x_3 x_4 w_{36} x_2^2 x_4^2 + w_{37} x_2 x_3^3 + w_{38} x_2 x_3^2 x_4 \\ & + w_{39} x_2 x_3 x_4^2 + w_{40} x_2 x_4^3 w_{41} x_2^2 x_4^2 + w_{42} x_2 x_3^3 + w_{43} x_2 x_3^2 x_4 + w_{44} x_2 x_3 x_4^2 + w_{45} x_2 x_4^3 \end{aligned} \quad (3-45)$$

The selection of the NN is usually a natural choice guided by engineering experience and intuition. This is a NN with polynomial activation functions, and hence $V(0) = 0$.

This is a power series NN with 45 activation functions containing powers of the state variable of the system up to the fourth order. The number of neurons required is chosen

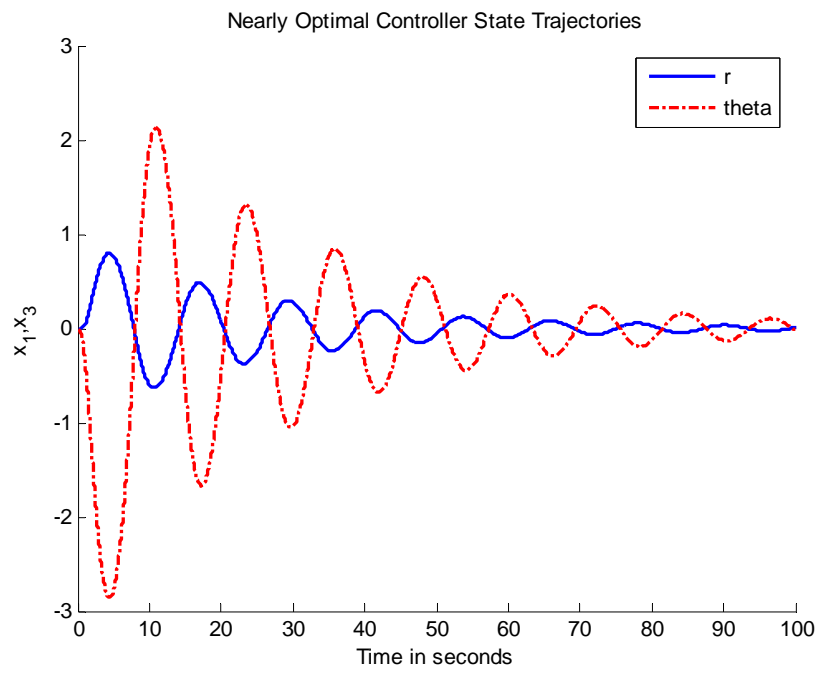


Figure 3-3 r, θ State Trajectories

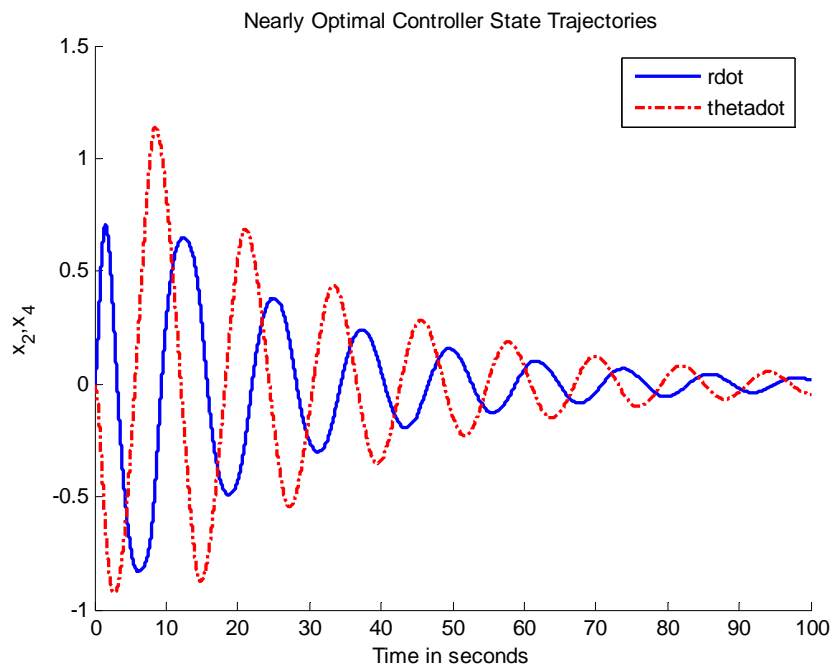


Figure 3-4 $\dot{r}, \dot{\theta}$ State Trajectories

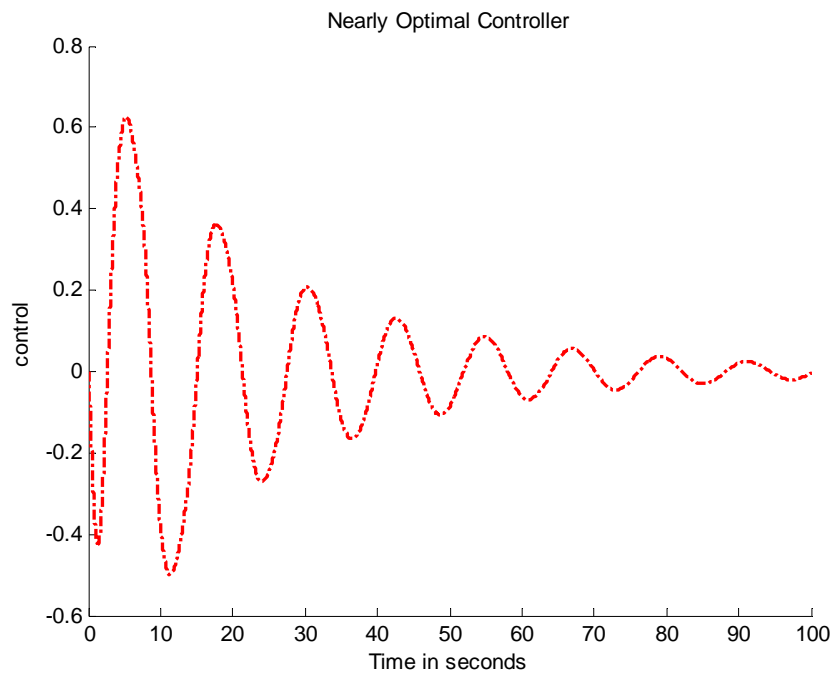


Figure 3-5 $u(t)$ Control Input

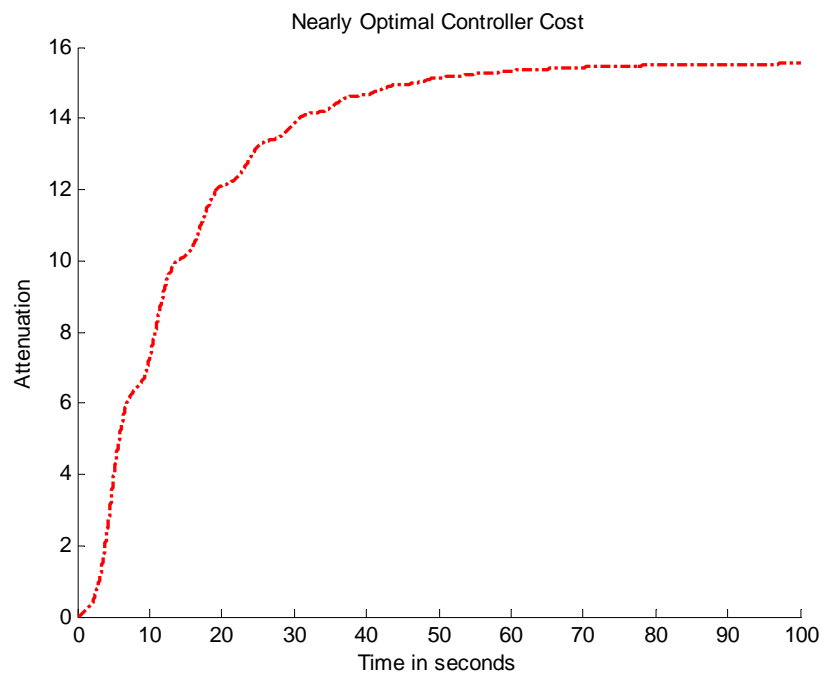


Figure 3-6 Disturbance Attenuation

to guarantee the uniform convergence of the algorithm. In this example,

$$\mathbf{w}_L(t_f) = [10; 10; 10; 10; 0]$$

and $t_f = 100$ seconds.

Figure 3-3 and 3-4 shows the states trajectories when the system is at rest and experiencing a disturbance $d(t) = 5 \sin(t)e^{-t}$. Figure 3-5 and 3-6 shows the control signal and attenuation respectively. The graphs imply it is L_2 -gain bounded by γ^2 .

3.5. Conclusion

In this chapter, neural networks are used to approximately solve the finite-horizon optimal H_∞ state feedback control problem. The method is based on solving a related Hamilton-Jacobi-Isaacs equation of the corresponding finite-horizon zero-sum game. The neural network approximates the corresponding game value function on a certain domain of the state-space and results in a control computed as the output of a neural network. It is shown that the neural network approximation converges uniformly to the game-value function and the resulting controller provides closed-loop stability and bounded L_2 gain. The result is a nearly exact H_∞ feedback controller with time-varying coefficients that is solved a priori offline.

CHAPTER 4

NEURAL NETWORK SOLUTION FOR FIXED-FINAL TIME CONSTRAINED OPTIMAL CONTROL

4.1. Introduction

This chapter is an extension to chapter 2 and 3. The constrained input optimization of dynamical systems has been the focus of many papers during the last few years. Several methods for deriving constrained control laws are found in Saberi and Bernstein [15]. However, most of these methods do not consider optimal control laws for general constrained nonlinear systems. Constrained-input optimization possesses challenging problems, a great variety of versatile methods have been successfully applied in Athans [5], Bernstein [16], Dolphus [33] and Saberi [77]. Many problems can be formulated within the Hamilton-Jacobi (HJ) and Lyapunov's frameworks, but the resulting equations are difficult or impossible to solve, such as Lyshevski [60][61][62]. In this chapter, we use NN to approximately solve the time-varying HJ equation for constrained control nonlinear systems. It is shown that using a NN approach, one can simply transform the problem into solving a nonlinear (ODE) backwards in time.

4.2. Background on Fixed-Final Time Constrained Optimal Control

Consider now the case when the control input is constrained by a saturated function $\phi(\cdot)$, e.g. \tanh , etc. To guarantee bounded controls, [1][56] introduced a generalized nonquadratic functional

$$W(u) = 2 \int_0^u \Phi^{-T}(v) R dv,$$

$$\Phi(v) = [\phi(v_1) \cdots \phi(v_m)]^T,$$

$$\Phi^{-1}(u) = [\phi^{-1}(u_1) \cdots \phi^{-1}(u_m)],$$

where $v \in \Re^m$, $\Phi \in \Re^m$, and $\Phi(\cdot)$ is a bounded one-to-one function that belongs to $C^p(p \geq 1)$ and $L_2(\Omega)$. Moreover, it is a monotonic odd function with its first derivative bounded by a constant M . Note that $W(u)$ is positive definite because $\Phi^{-1}(u)$ is monotonic odd and R is positive definite.

4.2.1. HJB Case

When (4-1) is used, (2-2) becomes

$$V(x(t_0), t_0) = \phi(x(t_f), t_f) + \int_{t_0}^{t_f} \left[Q(x) + 2 \int_0^u \Phi^{-T}(v) R dv \right] dt. \quad (4-1)$$

and (2-5) becomes

$$-\frac{\partial V(x, t)^*}{\partial t} = \min_{u(t)} \left(Q(x) + 2 \int_0^u \Phi^{-T}(v) R dv + \frac{\partial V(x, t)^*}{\partial x} (f(x) + g(x)u(t)) \right). \quad (4-2)$$

Minimizing the Hamiltonian of the optimal control problem with regard to u gives

$$g^T(x) \frac{\partial V(x, t)^*}{\partial x} + 2\Phi^{-1}(u^*) = 0,$$

so

$$u(t)^* = -\Phi \left(\frac{1}{2} R^{-1} g(x)^T \frac{\partial V(x, t)^*}{\partial x} \right), \quad u \in U \subset \Re^m. \quad (4-3)$$

This is constrained as required.

When (4-3) is used, (4-2) becomes

$$\begin{aligned} HJB(V^*(x,t)) = & \frac{\partial V^*(x,t)}{\partial t} + \frac{\partial V^{*T}(x,t)}{\partial x} f + 2 \int_0^u \phi^{-T}(v) R dv \\ & - \frac{\partial V^{*T}(x,t)}{\partial x} \cdot g \cdot \Phi \left(\frac{1}{2} R^{-1} g^T(x) \frac{\partial V^*(x,t)}{\partial x} \right) + Q(x) = 0 \end{aligned} \quad (4-4)$$

If this HJB equation can be solved for the value function $V(x,t)$, then (4-3) gives the optimal constrained control. This HJB equation cannot generally be solved. There is currently no method for rigorously solving for the value function of this constrained optimal control problem.

Lemma 2.1 The smooth bounded control law (4-3) guarantees at least a strong relative minimum for the performance cost (4-1) for all $x \in X \subset \Re$ on $[t_0, t_f]$. Moreover, if an optimal control exists, it is unique and represented by (4-3).

Proof. See [60]. ■

4.2.2. HJI Case

Define the Hamiltonian function

$$\begin{aligned} H(x, p, u, d) = & \frac{\partial V^T(x,t)}{\partial x} (f(x) + g(x)u(t) + k(x)d(t)) \\ & + h^T(x)h(x) + 2 \int_0^u \phi^{-T}(v) dv - \gamma^2 \|d(t)\|^2 \end{aligned} \quad (4-5)$$

Minimizing the Hamiltonian of the optimal control problem with regard to u and d gives

$$\Phi \left(g^T(x) \frac{\partial V(x,t)}{\partial x} \right) + 2u^*(t) = 0, \quad (4-6)$$

$$k^T(x) \frac{\partial V(x,t)}{\partial x} - 2\gamma^2 d^*(t) = 0, \quad (4-7)$$

so

$$u^*(t) = -\Phi \left(\frac{1}{2} g^T(x) \frac{\partial V(x,t)}{\partial x} \right), \quad (4-8)$$

$$d^*(t) = \frac{1}{2\gamma^2} k^T(x) \frac{\partial V(x,t)}{\partial x}. \quad (4-9)$$

When $H^*(x, p, u^*, d^*) = 0$, we have the Hamilton-Jacobi-Isaacs equation

$$\begin{aligned} & \frac{\partial V(x,t)}{\partial t} + \frac{\partial V^T(x,t)}{\partial x} (f(x) + g(x)u^*(t) + k(x)d^*(t)) \\ & + h^T(x)h(x) + 2 \int_0^u \Phi^{-T}(v) dv - \gamma^2 \|d^*(t)\|^2 = 0 \end{aligned} \quad (4-10)$$

From (4-8) and (4-9), (4-10) becomes

$$\begin{aligned} HJI(V^*(x,t)) &= \frac{\partial V^*(x,t)}{\partial t} + \frac{\partial V^{*T}(x,t)}{\partial x} f(x) \\ &+ 2 \int_0^{u^*} \Phi^{-T}(v) dv - \frac{\partial V^{*T}(x,t)}{\partial x} g(x) \cdot \Phi \left(\frac{1}{2} g^T(x) \frac{\partial V^*(x,t)}{\partial x} \right), \\ &+ \frac{1}{4\gamma^2} \frac{\partial V^{*T}(x,t)}{\partial t} k(x) k^T(x) \frac{\partial V^*(x,t)}{\partial t} + h^T(x)h(x) = 0 \end{aligned} \quad (4-11)$$

with boundary condition $V(x(t_f), t_f) = \phi(x(t_f), t_f)$.

Equations (4-8), (4-9) and (4-10) provide the solution to finite-horizon optimal control for general nonlinear systems. However, equation (4-11) is generally impossible to solve for nonlinear systems. There is currently no method for rigorously solving for the value function of this constrained optimal control problem.

It can be easily shown that

$$\int_0^{t_f} \left(h^T(x)h(x) + 2 \int_0^{u^*} \Phi^{-T}(v)dv \right) dt \leq \gamma^2 \int_0^{t_f} \|d(t)\|^2 dt, \quad (4-12)$$

When $x(0)=0$, therefore the quasi- L_2 -gain $\leq \gamma$. Controllers derived using (4-11) are suboptimal H_∞ controllers. If the suboptimal controller is found for the smallest γ , then it is called the optimal H_∞ controller.

4.3. Nonlinear Fixed-Final-Time Solution by NN Least-Squares Approximation

Like in unconstrained case

$$\frac{\partial V_L(x)}{\partial x} = \frac{\partial \sigma_L^T}{\partial x} \mathbf{w}_L(t) \equiv \nabla \sigma_L^T \mathbf{w}_L(t), \quad (4-13)$$

where $\nabla \sigma_L(x)$ is the Jacobian $\frac{\partial \sigma_L(x)}{\partial x}$, and that

$$\frac{\partial V_L(x)}{\partial t} = \dot{\mathbf{w}}_L^T(t) \sigma_L(x). \quad (4-14)$$

4.3.1. HJB Case

Therefore approximating $V(x, t)$ by $V_L(x, t)$ in the HJB equation (4-4) results in

$$\begin{aligned} & -\dot{\mathbf{w}}_L^T(t) \sigma_L(x) - \mathbf{w}_L^T(t) \nabla \sigma_L(x) f(x) - 2 \int_0^u \Phi^{-T}(v) R dv \\ & + \mathbf{w}_L^T(t) \nabla \sigma_L(x) \cdot g(x) \cdot \Phi \left(\frac{1}{2} R^{-1} g^T(x) \nabla \sigma_L^T(x) \mathbf{w}_L(t) \right) - Q(x), \\ & = e_L(x, t) \end{aligned} \quad (4-15)$$

or

$$HJB \left(V_L(x, t) = \sum_{j=1}^L w_j(t) \sigma_j(x), u(t) \right) = e_L(x, t), \quad (4-16)$$

where $e_L(x, t)$ is a residual equation error. From (4-3) the corresponding constrained

optimal control input is

$$u_L(t) = -\Phi \left(\frac{1}{2} R^{-1} g^T(x) \nabla \sigma_L^T(x) \mathbf{w}_L(t) \right). \quad (4-17)$$

Therefore one obtains

$$\begin{aligned} & \left\langle -\dot{\mathbf{w}}_L^T(t) \sigma_L(x), \sigma_L(x) \right\rangle_{\Omega} + \left\langle -\mathbf{w}_L^T(t) \nabla \sigma_L(x) f(x), \sigma_L(x) \right\rangle_{\Omega} + \left\langle -2 \int_0^u \Phi^{-T}(v) R dv, \sigma_L(x) \right\rangle_{\Omega} \\ & + \left\langle \mathbf{w}_L^T(t) \nabla \sigma_L(x) \cdot g(x) \Phi \left(\frac{1}{2} R^{-1} g^T(x) \nabla \sigma_L^T(x) \mathbf{w}_L(t) \right), \sigma_L(x) \right\rangle_{\Omega} \\ & + \left\langle -Q(x), \sigma_L(x) \right\rangle_{\Omega} = 0 \end{aligned} \quad (4-18)$$

So that

$$\begin{aligned} \dot{\mathbf{w}}_L(t) = & - \left\langle \sigma_L(x), \sigma_L(x) \right\rangle_{\Omega}^{-1} \cdot \left\langle \nabla \sigma_L(x) f(x), \sigma_L(x) \right\rangle_{\Omega} \cdot \mathbf{w}_L(t) \\ & - \left\langle \sigma_L(x), \sigma_L(x) \right\rangle_{\Omega}^{-1} \left\langle 2 \int_0^u \Phi^{-T}(v) R dv, \sigma_L(x) \right\rangle_{\Omega} \\ & + \left\langle \sigma_L(x), \sigma_L(x) \right\rangle_{\Omega}^{-1} \cdot \left\langle \mathbf{w}_L^T(t) \nabla \sigma_L(x) \cdot g(x) \Phi \left(\frac{1}{2} R^{-1} g^T(x) \nabla \sigma_L^T(x) \mathbf{w}_L(t) \right), \sigma_L(x) \right\rangle_{\Omega} \\ & - \left\langle \sigma_L(x), \sigma_L(x) \right\rangle_{\Omega}^{-1} \cdot \left\langle Q(x), \sigma_L(x) \right\rangle_{\Omega} \end{aligned} \quad (4-19)$$

with boundary condition $V(x(t_f), t_f) = \phi(x(t_f), t_f)$.

4.3.2. HJI Case

For the $HJI(V(x, t), d(t)) = 0$, the solution $V(x, t)$ is replaced with $V_L(x, t)$ having a residual error

$$\begin{aligned} & \dot{\mathbf{w}}_L^T(t) \sigma_L(x) + \mathbf{w}_L^T(t) \nabla \sigma_L(x) f(x) + 2 \int_0^u \Phi^{-T}(v) R dv \\ & - \mathbf{w}_L^T(t) \nabla \sigma_L(x) \cdot g(x) \cdot \Phi \left(\frac{1}{2} R^{-1} g^T(x) \nabla \sigma_L^T(x) \mathbf{w}_L(t) \right) \\ & + \frac{1}{4\gamma^2} \mathbf{w}_L^T(t) \nabla \sigma_L(x) k(x) k^T(x) \nabla \sigma_L^T(x) \mathbf{w}_L(t) + h^T(x) h(x) = e_L(x, t) \end{aligned} \quad (4-20)$$

or

$$HJI\left(V_L(x,t)=\sum_{i=1}^L w_i(t)\sigma_i(x),d(t)\right)=e_L(x,t), \quad (4-21)$$

Therefore one obtains

$$\begin{aligned} \dot{\mathbf{w}}_L(t) = & -\langle \boldsymbol{\sigma}_L(x), \boldsymbol{\sigma}_L(x) \rangle_{\Omega}^{-1} \cdot \langle \nabla \boldsymbol{\sigma}_L(x) f(x), \boldsymbol{\sigma}_L(x) \rangle_{\Omega} \cdot \mathbf{w}_L(t) \\ & -\langle \boldsymbol{\sigma}_L(x), \boldsymbol{\sigma}_L(x) \rangle_{\Omega}^{-1} \left\langle 2 \int_0^u \boldsymbol{\varphi}^{-T}(v) R dv, \boldsymbol{\sigma}_L(x) \right\rangle_{\Omega} \\ & + \langle \boldsymbol{\sigma}_L(x), \boldsymbol{\sigma}_L(x) \rangle_{\Omega}^{-1} \cdot \left\langle \mathbf{w}_L^T(t) \nabla \boldsymbol{\sigma}_L(x) g(x) \boldsymbol{\varphi} \left(\frac{1}{2} R^{-1} g^T(x) \nabla \boldsymbol{\sigma}_L^T(x) \mathbf{w}_L(t) \right), \boldsymbol{\sigma}_L(x) \right\rangle_{\Omega}, \quad (4-22) \\ & -\langle \boldsymbol{\sigma}_L(x), \boldsymbol{\sigma}_L(x) \rangle_{\Omega}^{-1} \cdot \left\langle \frac{1}{4\gamma^2} \mathbf{w}_L^T(t) \nabla \boldsymbol{\sigma}_L(x) k(x) k^T(x) \nabla \boldsymbol{\sigma}_L^T(x) \mathbf{w}_L(t), \boldsymbol{\sigma}_L(x) \right\rangle_{\Omega} \\ & -\langle \boldsymbol{\sigma}_L(x), \boldsymbol{\sigma}_L(x) \rangle_{\Omega}^{-1} \cdot \langle h^T(x) h(x), \boldsymbol{\sigma}_L(x) \rangle_{\Omega} \end{aligned}$$

with boundary condition $V(x(t_f), t_f) = \phi(x(t_f), t_f)$.

Also we get

$$u^*(t) = -\boldsymbol{\varphi} \left(\frac{1}{2} g^T(x) \nabla \boldsymbol{\sigma}_L^T \mathbf{w}_L(t) \right) \quad \text{and} \quad d^*(t) = \frac{1}{2\gamma^2} k^T(x) \nabla \boldsymbol{\sigma}_L^T \mathbf{w}_L(t).$$

This yields a feedback controller that is formulated from a neural network.

4.4. Numerical Examples

We now show the power of our NN control technique for finding nearly optimal fixed-final time constrained controllers. Two examples are presented.

4.4.1. HJB Case

In this section, two examples are shown to illustrate the algorithm, both of them applies constrained case.

4.4.1.1. Linear System

a) We start by applying the algorithm obtained above for the linear system

$$\begin{aligned}\dot{x}_1 &= 2x_1 + x_2 + x_3 \\ \dot{x}_2 &= x_1 - x_2 + u_2 \quad . \\ \dot{x}_3 &= x_3 + u_1\end{aligned}\tag{4-23}$$

Define performance index

$$V(x(t_0), t_0) = \phi(x(t_f), t_f) + \int_{t_0}^T \left(Q(x) + 2 \int_0^u \phi^{-T}(v) R dv \right) dt .\tag{4-24}$$

Here $Q = 10 * I_{3 \times 3}$ and $R = I_{2 \times 2}$, where I is identity matrices. It is desired to control the system with input constraints $|u_1| \leq 5$, $|u_2| \leq 20$. In order to ensure the constrained control, a nonquadratic cost performance term (4-24) is used. To show how to do this for the general case of $|u| \leq 5$, we use $A \cdot \tanh(1/A \cdots)$ for $\phi(\cdots)$. Hence the nonquadratic cost is

$$W(u) = 2 \int_0^u A \cdot \tanh^{-T}(v/A) R dv .$$

The plot is shown in Figure 4-1. This nonquadratic cost performance is used in the algorithm to calculate the optimal constrained controller. The algorithm is run over the region Ω_0 defined by $-2 \leq x_1 \leq 2$, $-2 \leq x_2 \leq 2$, $-2 \leq x_3 \leq 2$. To find a nearly optimal time-varying controller, the following smooth function is used to approximate the value function of the system

$$V(x_1, x_2) = w_1 x_1^2 + w_2 x_2^2 + w_3 x_3^2 + w_4 x_1 x_2 + w_5 x_1 x_3 + w_6 x_2 x_3 .$$

This is a NN with polynomial activation functions, and hence $V(0) = 0$.

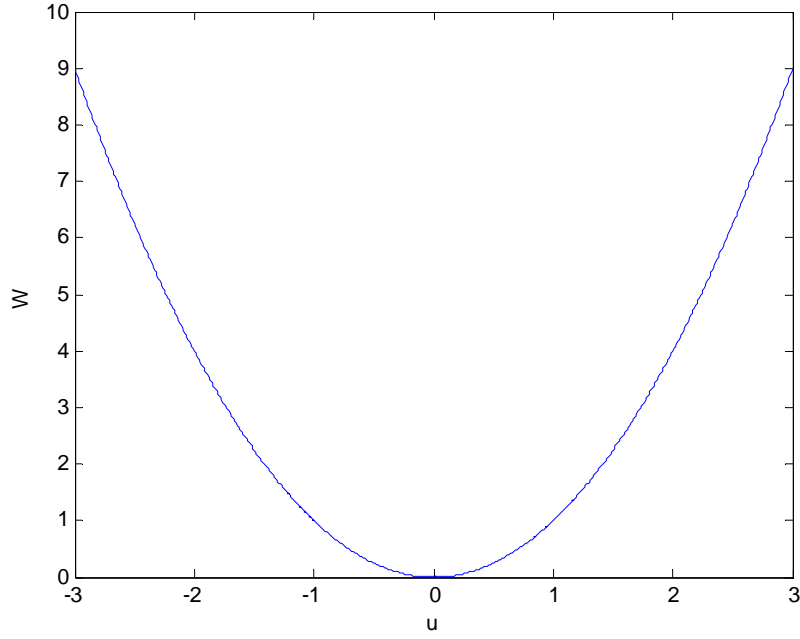


Figure 4-1 Nonquadratic cost

In this example, six neurons are chosen and $\mathbf{w}_L(t_f) = [10, 10, 10, 0, 0, 0]$. Our algorithm was used to determine the nearly optimal time-varying constrained control law by backwards integrating to solve (2-37). The required quantities A, B, C, D, E in (2-37) were evaluated for 5000 points in Ω_0 . A least-square algorithm from MATLAB was used to compute $\dot{\mathbf{w}}_L(t)$ at each integration time. The solution was obtained in 30 seconds. From Figure 4-2 it is obvious that about 25 seconds from t_f , the weights converge to constant. The states and control signal obtained by a forward integration are shown in Figures 4-3 and 4-4. The control is bounded as required.

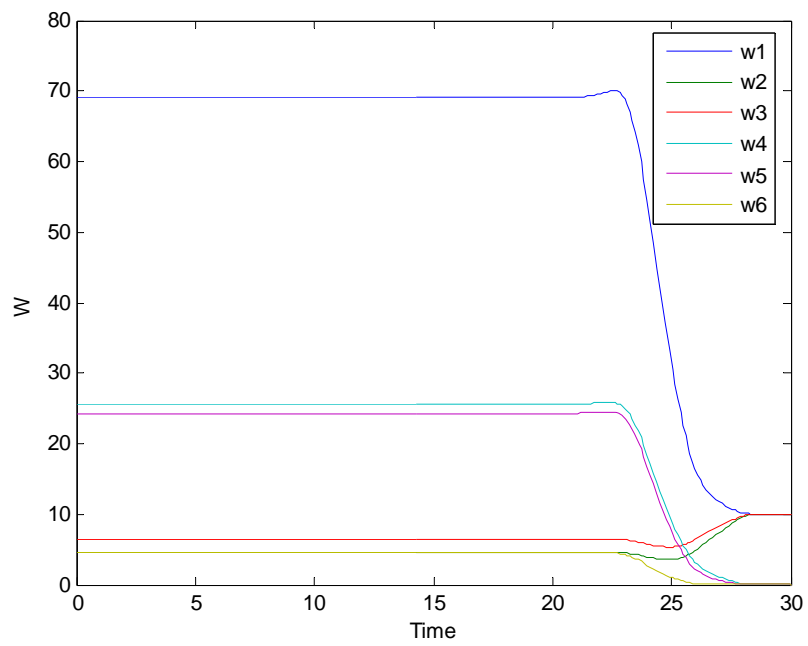


Figure 4-2 Constrained Linear System Weights

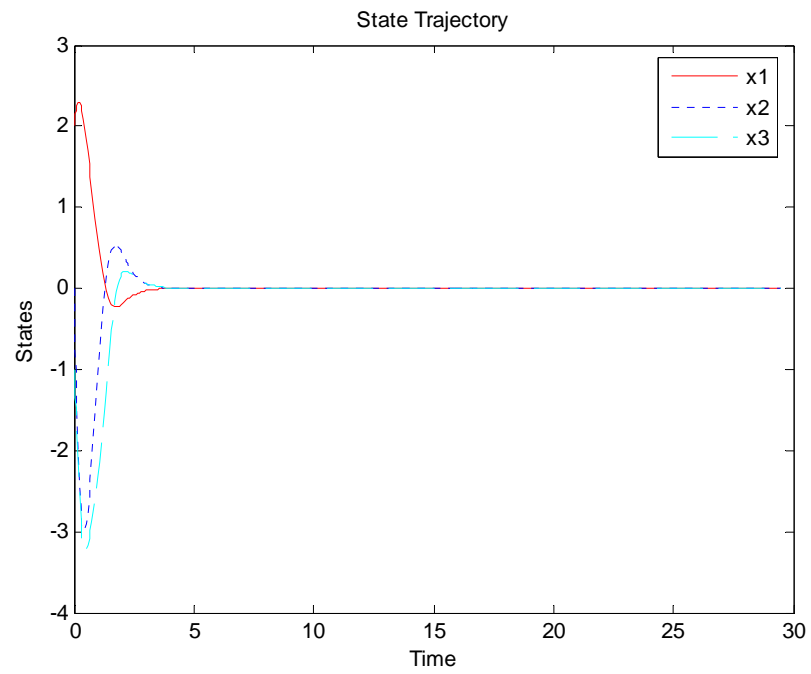


Figure 4-3 State Trajectory of Linear System with Bounds

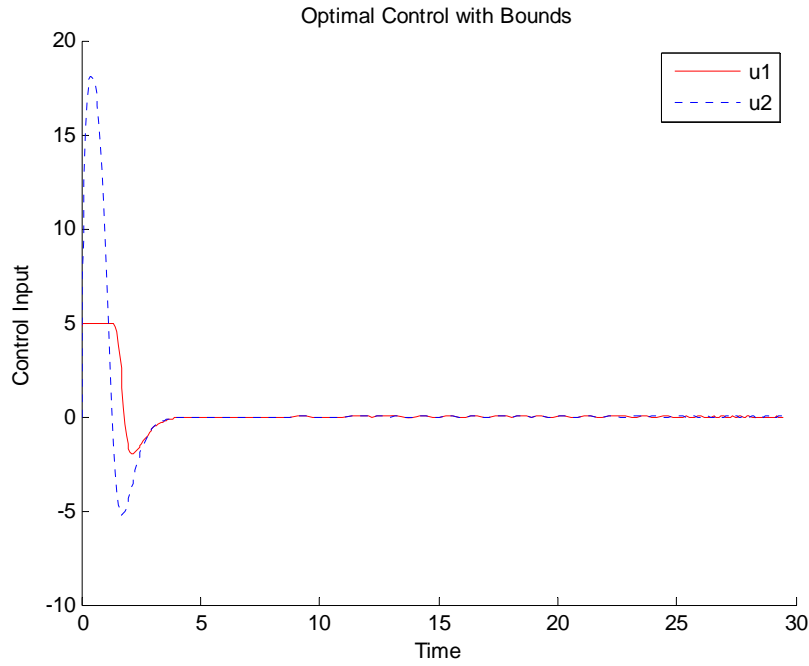


Figure 4-4 Optimal NN Control Law with Bounds

b) Now let $A=100$ so that the control constraints are effectively removed. The algorithm is run and the plots of P_{11} , P_{22} , P_{33} and function of time are shown in Figure 4-5. These plots converge to steady state values of $P_{11} = 69.0573$, $P_{22} = 4.6208$, $P_{33} = 6.5008$. These correspond exactly to the algebraic Riccati equation solution obtained by standard optimal control methods [53], which is

$$P = \begin{bmatrix} 69.0573 & 12.8164 & 12.1491 \\ 12.8164 & 4.6208 & 2.2448 \\ 12.1491 & 2.2448 & 6.5008 \end{bmatrix}.$$

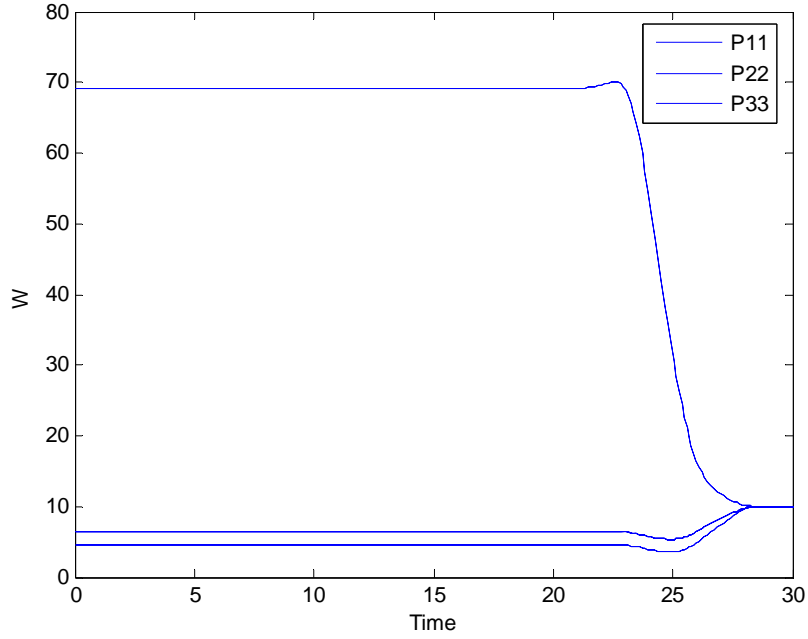


Figure 4-5 Unconstrained Control System Weights

4.4.1.2. Nonlinear Chained System

One can apply the results of this chapter to a mobile robot, which is a nonholonomic system [48]. It is known [23] that there does not exist a continuous time-invariant feedback control law that minimizes the cost. Some methods for deriving stable controls of nonholonomic systems are found in Bloch [18][19], Egeland [35], Escobar [36], Fierro and Lewis [37], Murray [66][67], Pomet [72] and Sordalen [81]. Our method will yield a time-varying gain. From Murray [32], under some sufficient conditions, a nonholonomic system can be converted to chained form as

$$\begin{aligned}
 \dot{x}_1 &= u_1 \\
 \dot{x}_2 &= u_2 \\
 \dot{x}_3 &= x_1 u_2
 \end{aligned} \tag{4-25}$$

Define performance index (4-24). Here Q and R are chosen as identity matrices. It is desired to control the system with control limits of $|u_1| \leq 1$, $|u_2| \leq 2$. A similar nonquadratic cost performance term is used as in the last example. Here the region Ω_0 is defined by $-2 \leq x_1 \leq 2$, $-2 \leq x_2 \leq 2$, $-2 \leq x_3 \leq 2$. To solve for the value function of the related optimal control problem, we selected the smooth approximating function

$$\begin{aligned}
V(x_1, x_2, x_3) = & w_1 x_1^2 + w_2 x_2^2 + w_3 x_3^2 + w_4 x_1 x_2 + w_5 x_1 x_3 + w_6 x_2 x_3 \\
& + w_7 x_1^4 + w_8 x_2^4 + w_9 x_3^4 + w_{10} x_1^2 x_2^2 + w_{11} x_1^2 x_3^2 + w_{12} x_2^2 x_3^2 + w_{13} x_1^2 x_2 x_3 \\
& + w_{14} x_1 x_2^2 x_3 + w_{15} x_1 x_2 x_3^2 + w_{16} x_1^3 x_2 + w_{17} x_1^3 x_3 + w_{18} x_1 x_2^3 + w_{19} x_1 x_3^3 \\
& + w_{20} x_2 x_3^3 + w_{21} x_2^3 x_3
\end{aligned} \tag{4-26}$$

The selection of the NN is usually a natural choice guided by engineering experience and intuition. This is a NN with polynomial activation functions, and hence $V(0) = 0$. This is a power series NN with 21 activation functions containing powers of the state variable of the system up to the fourth order. Convergence was not observed using a NN with only second-order powers of the states. The number of neurons required is chosen to guarantee the uniform convergence of the algorithm. In this example, $\mathbf{w}_L(t_f) = [10; 10; 10; 0; 0; 0; 10; 10; 10; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0]$ and $t_f = 30$ seconds. The required quantities A, B, C, D, E in (2-37) were evaluated for 5000 points in Ω_0 . Figure 4-6 indicates that the weights converge to constants when they are integrated backwards. Figure 4-7 shows that the systems' states response, including x_1, x_2 and x_3 are all bounded. It can be seen that the states do converge to a value close to the origin. Figure 4-8 shows the optimal control is constrained as required and converges to zero.

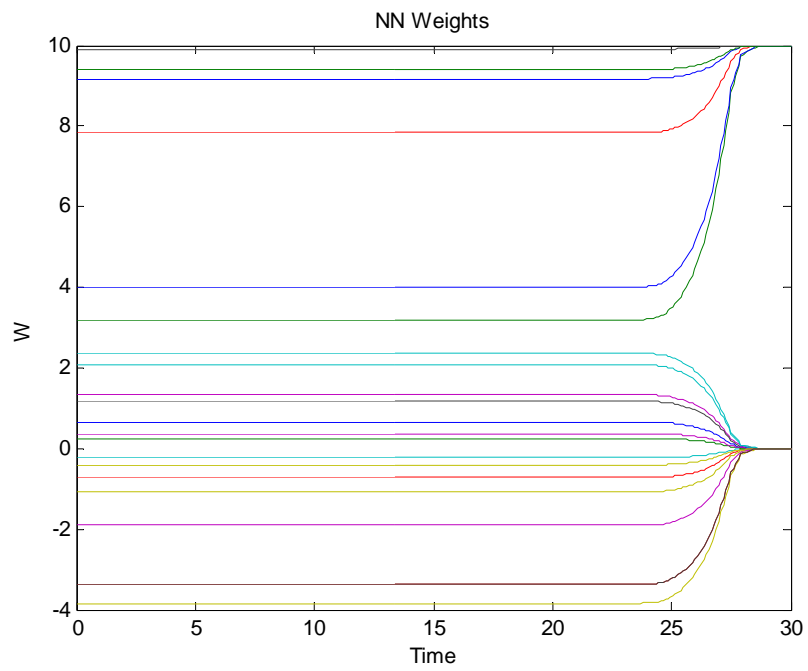


Figure 4-6 Nonlinear System Weights

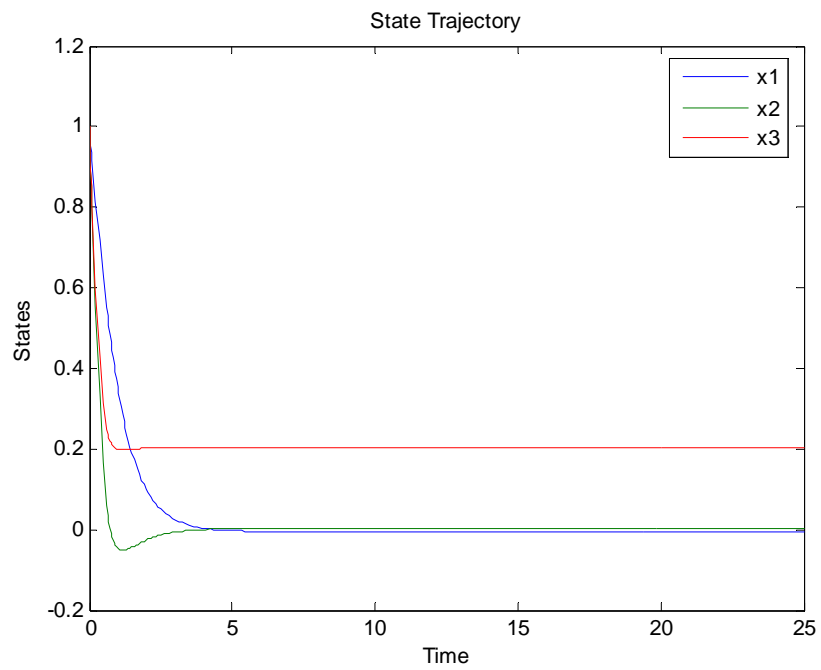


Figure 4-7 State Trajectory of Nonlinear System

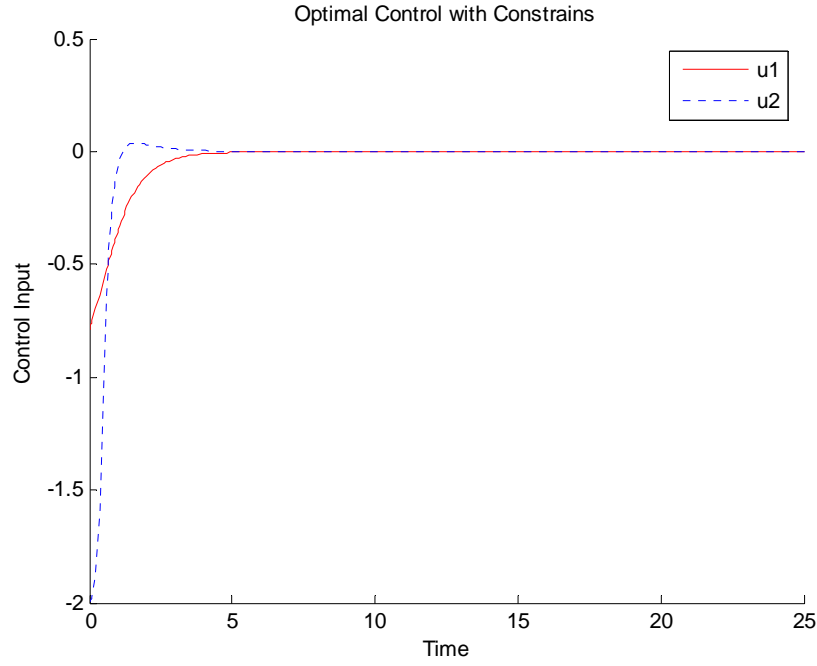


Figure 4-8 Optimal NN Constrained Control Law

4.4.2. HJI Case

In this example, we will show the power of our NN control technique for finding nearly optimal finite-horizon H_∞ state feedback controller for the Rotational/Translational Actuator shown in Figure 4-9. This was defined as benchmark problem in [24].

$$\dot{x} = f(x) + g(x)u(t) + k(x)d(t), \quad |u(t)| \leq 2,$$

$$z^T z = x_1^2 + 0.1x_2^2 + 0.1x_3^2 + 0.1x_4^2 + \|u(t)\|_q^2,$$

$$\varepsilon = \frac{me}{\sqrt{(I + me^2)(M + m)}} = 0.2, \quad \gamma = 10,$$

$$f = \begin{bmatrix} x_2 & \frac{-x_1 + \varepsilon x_4^2 \sin x_3}{1 - \varepsilon^2 \cos^2 x_3} & x_4 & \frac{\varepsilon \cos x_3 (x_1 - \varepsilon x_4^2 \sin x_3)}{1 - \varepsilon^2 \cos^2 x_3} \end{bmatrix}^T,$$

$$g = \begin{bmatrix} 0 & \frac{-\varepsilon \cos x_3}{1 - \varepsilon^2 \cos^2 x_3} & 0 & \frac{1}{1 - \varepsilon^2 \cos^2 x_3} \end{bmatrix}^T,$$

$$k = \begin{bmatrix} 0 & \frac{1}{1 - \varepsilon^2 \cos^2 x_3} & 0 & \frac{-\varepsilon \cos x_3}{1 - \varepsilon^2 \cos^2 x_3} \end{bmatrix}^T.$$

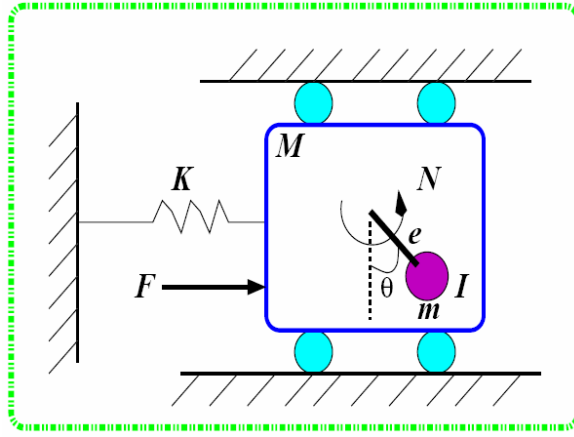


Figure 4-9 Rotational actuator to control a translational oscillator.

Here the state x_1 and x_2 are the normalized distance r and velocity of the cart \dot{r} ,

$$x_3 = \theta, \quad x_4 = \dot{\theta}.$$

Define performance index

$$V(x(t_0), t_0) = \phi(x(t_f), t_f) + \int_{t_0}^{t_f} \left(h^T(x)h(x) + 2 \int_0^u \phi^{-T}(v)dv - \gamma^2 \|d(t)\|^2 \right) dt.$$

It is desired to control the system with input constraints $|u(t)| \leq 0.5$. Here we use

$A \tanh(1/A \cdots)$ for $\phi(\cdots)$, hence the nonquadratic functional is

$$W(u) = 2 \int_0^u A \tanh^{-T}(v/A) R dv.$$

This nonquadratic cost performance is used in the algorithm to calculate the optimal constrained controller. The algorithm is run over the region Ω defined by $-2 \leq x \leq 2$.

To find a nearly optimal time-varying controller, the following smooth function is used to approximate the value function of the system

$$\begin{aligned}
V(x_1, x_2, x_3, x_4) = & w_1 x_1^2 + w_2 x_1 x_2 + w_3 x_1 x_3 + w_4 x_1 x_4 + w_5 x_2^2 + w_6 x_2 x_3 + w_7 x_2 x_4 + w_8 x_3^2 \\
& + w_9 x_3 x_4 + w_{10} x_4^2 + w_{11} x_1^4 + w_{12} x_1^3 x_2 + w_{13} x_1^3 x_3 + w_{14} x_1^3 x_4 + w_{15} x_1^2 x_2^2 + w_{16} x_1^2 x_2 x_3 \\
& + w_{17} x_1^2 x_2 x_4 + w_{18} x_1^2 x_3^2 + w_{19} x_1^2 x_3 x_4 + w_{20} x_1^4 x_4^2 + w_{21} x_1 x_2^3 + w_{22} x_1 x_2^2 x_3 + w_{23} x_1 x_2^2 x_4 \\
& + w_{24} x_1 x_2 x_3^2 + w_{25} x_1 x_2 x_3 x_4 + w_{26} x_1 x_2 x_4^2 + w_{27} x_1 x_3^3 + w_{28} x_1 x_3^2 x_4 + w_{29} x_1 x_3 x_4^2 \\
& + w_{30} x_1 x_4^3 + w_{31} x_2^4 + w_{32} x_2^3 x_3 + w_{33} x_2^3 x_4 + w_{34} x_2^2 x_3^2 + w_{35} x_2^2 x_3 x_4 + w_{36} x_2^2 x_4^2 + w_{37} x_2 x_3^3 \\
& + w_{38} x_2 x_3^2 x_4 + w_{39} x_2 x_3 x_4^2 + w_{40} x_2 x_4^3 + w_{41} x_2^2 x_4^2 + w_{42} x_2 x_3^3 \\
& + w_{43} x_2 x_3^2 x_4 + w_{44} x_2 x_3 x_4^2 + w_{45} x_2 x_4^3
\end{aligned}$$

The selection of the NN is usually a natural choice guided by engineering experience and intuition. This is a NN with polynomial activation functions, and hence $V(0)=0$. This is a power series NN with 45 activation functions containing powers of the state variable of the system up to the fourth order. The number of neurons required is chosen to guarantee the uniform convergence of the algorithm. In this example,

[illegible]

and $t_f = 100$ seconds.

Figure 4-10 and 4-11 shows the states trajectories when the system is at rest and experiencing a disturbance $d(t) = 5 \sin(t)e^{-t}$. Figure 4-12 and 4-13 shows the control signal and attenuation respectively. The graphs imply it is L_2 -gain bounded by γ^2 .

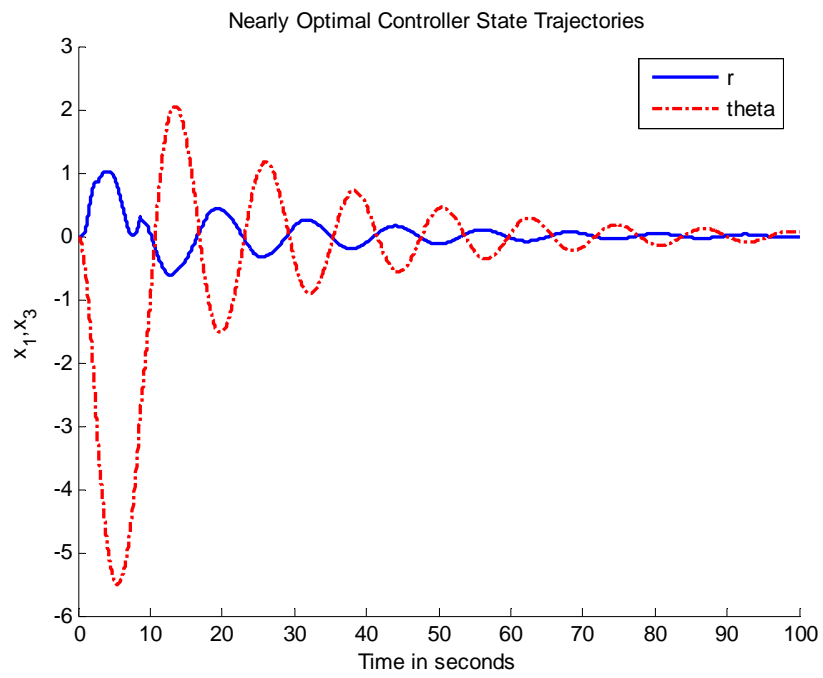


Figure 4-10 r, θ State Trajectories

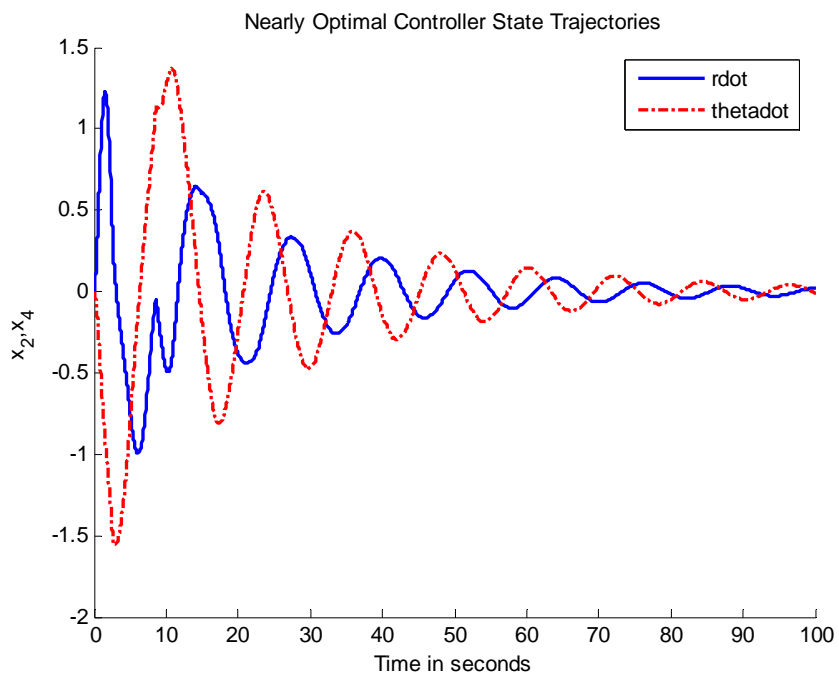


Figure 4-11 $\dot{r}, \dot{\theta}$ State Trajectories

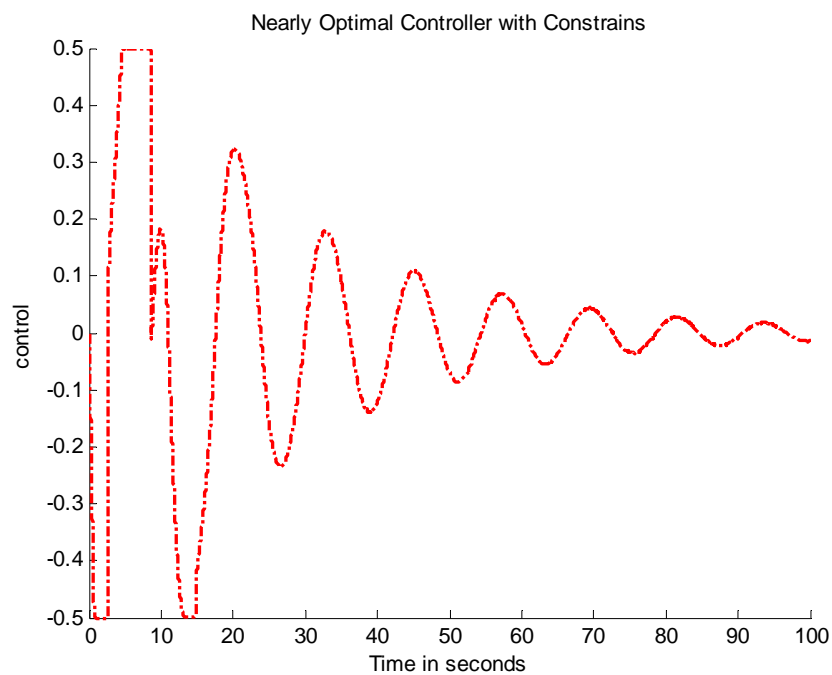


Figure 4-12 $u(t)$ Control Input

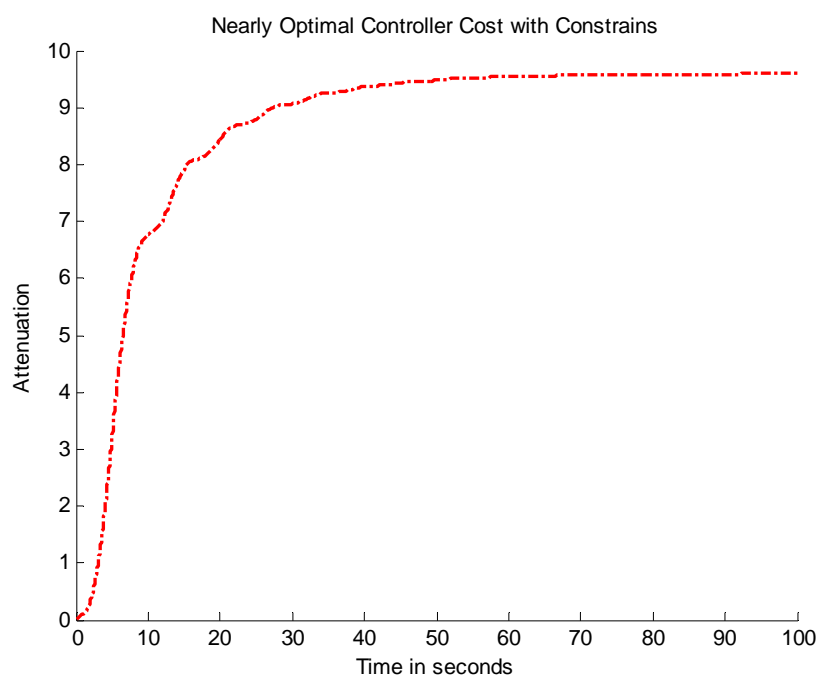


Figure 4-13 Disturbance Attenuation

4.5. Conclusion

In this chapter, optimal control of constrained input systems is discussed, a neural network approximation of the value function is introduced, and the method is employed in a least-squares sense over a mesh with certain size on Ω . Linear and chained form system examples are shown.

CHAPTER 5

SUBOPTIMAL CONTROL OF CHAINED SYSTEM WITH TIME-FOLDING METHOD

5.1. Introduction

In this chapter, we develop fixed-final time nearly optimal control laws for a class of nonholonomic chained form systems by using neural networks to approximately solve an HJB equation. A certain time-folding method is applied to recover uniform complete controllability for the chained form system. This method requires an innovative design of a certain dynamic control component. Using this time-folding method, the chained form system is mapped into a controllable linear system for which controllers can systematically be designed to ensure exponential or asymptotic stability as well as nearly optimal performance. The result is a neural network feedback controller that has time-varying coefficients found by a priori offline tuning. The results of this chapter are demonstrated on an example.

5.2. Problem Description

Stabilization of chained system remains to be a difficult and interesting problem because of the following technical issues:

- (1) Topologically, the chained system cannot be stabilized under any continuous control $u = u(x)$ due to its nonlinear characteristics.

- (2) While the system is nonlinearly controllable everywhere, the system as it is not globally feedback linearizable (though local feedback linearizable is possible as shown by the σ -process but singularity manifold remains in all the neighborhoods around the origin), and nonlinear controllability does not necessarily translate into systematic control design.
- (3) Chained system is not linearly controllable around the origin.

Above three issues make the chained system complex, the main problem is the product term of the chained system can't converge to zero. In this chapter, we use time-folding method to solve this problem.

Using Time-folding method, the chained form system is mapped into a controllable linear time-varying system for which control can systematically be designed to ensure exponential or asymptotic stability as well as optimal performance. Simulations show that the method is feasible.

5.3. Neural Network Algorithm for Chained Form System with Time-Folding Method

Brockett's theorem indicates that nonholonomic systems cannot be asymptotically stabilized around a fixed point under any smooth (or even continuous) time-independent state feedback control law. In this section, a smooth nearly-optimal time-varying control is designed to stabilize the chained form system using a time-folding method [73][74], With a new dynamic control design, a global nonlinear time transformation is found to transform the chained form system into a controllable linear time-varying system.

5.3.1. Chained Form System Description

Consider the following 2-input 3-dimensional nonholonomic chained form system:

$$\begin{aligned}\dot{x}_1 &= u_1 \\ \dot{x}_2 &= x_3 u_1, \\ \dot{x}_3 &= u_2\end{aligned}\tag{5-1}$$

where $x = [x_1 \dots x_n]^T \in \mathfrak{R}^n$ is the state, $u = [u_1 \ u_2]^T \in \mathfrak{R}^2$ is the control input. The objective of this chapter is to present time-varying and continuous feedback controls that globally stabilize system (5-1) and are optimal with respect to certain performance indices. It is straightforward to extend the proposed results to m -input Nonholonomic systems that can be transformed into the chained form.

The chained form system (5-1) can be decomposed into the following two interconnected subsystems:

$$\dot{x}_1 = u_1,\tag{5-2}$$

$$\dot{z} = u_2 A_1 z + B_1 u_2,\tag{5-3}$$

Where $z = [z_1 \ z_2]^T = [x_2 \ x_3]^T$, and $A_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

5.3.2. Dynamic Control Design

In this subsection, two dynamic feedback control components u_1 and u_2 will sequentially be designed to form the proposed asymptotically stabilizing control. As the first step, dynamic feedback control u_1 is chosen to be of the following form:

$$\hat{u}_1 = \lambda(t) u_1,\tag{5-4}$$

$$\hat{u}_2 = \lambda(t)u_2 = (t+a)u_2, \quad (5-5)$$

$$\lambda(t) = t+a, \quad (5-6)$$

where \hat{u}_1 , \hat{u}_2 are transformed controls, and a is constant. From (39), letting $\tau = \ln(t+a)$, then

$$\frac{d\hat{z}}{d\tau} = \gamma(\tau, a)A_1\hat{z} + B_1\hat{u}_2, \quad (5-7)$$

Where $\hat{z}(\tau) = z(t)$, $\gamma(\tau, a)$ is scale factor.

With the above transformation, the control should be changed to:

$$u(t) = -\frac{1}{2}\Phi(\lambda(t)R^{-1}g^T\nabla\sigma_L^T(x)\mathbf{w}_L(t)). \quad (5-8)$$

5.4. Simulation

We now show the power of our NN control technique using time-folding method for finding nearly optimal fixed-final time controllers to a mobile robot, which is a nonholonomic system [48]. Its kinematics model can be transformed into chained form (37) with $n = 3$. It is known [23] that there does not exist a continuous time-invariant feedback control law that minimizes the cost. Our method will yield a time-varying gain.

For a nonholonomic system, define performance index

$$V(x(t_0), t_0) = \phi(x(t_f), t_f) + \int_{t_0}^t (Q(x) + W(u))dt.$$

Here Q and R are chosen as identity matrices. To solve for the value function of the related optimal control problem, we selected the smooth approximating function

$$\begin{aligned}
V(x_1, x_2, x_3) = & w_1 x_1^2 + w_2 x_2^2 + w_3 x_3^2 + w_4 x_1 x_2 + w_5 x_1 x_3 + w_6 x_2 x_3 + w_7 x_1^4 \\
& + w_8 x_2^4 + w_9 x_3^4 + w_{10} x_1^2 x_2^2 + w_{11} x_1^2 x_3^2 + w_{12} x_2^2 x_3^2 + w_{13} x_1^2 x_2 x_3 + w_{14} x_1 x_2^2 x_3 \\
& + w_{15} x_1 x_2 x_3^2 + w_{16} x_1^3 x_2 + w_{17} x_1^3 x_3 + w_{18} x_1 x_2^3 + w_{19} x_1 x_3^3 + w_{20} x_2 x_3^3 + w_{21} x_2^3 x_3
\end{aligned} \quad (5-9)$$

The selection of the NN is usually a natural choice guided by engineering experience and intuition. This is a NN with polynomial activation functions, and hence $V(0) = 0$. This is a power series NN with 21 activation functions containing powers of the state variable of the system up to the fourth order. The number of neurons required is chosen to guarantee the uniform convergence of the algorithm. In this example,

$$\mathbf{w}_L(t_f) = [10; 10; 10; 0; 0; 0; 10; 10; 10; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0]$$

and $t_f = 30$ seconds.

In the simulation, initial condition of the state is set to be $x(t_0) = [1 \quad -1 \quad \pi/2]^T$. Figure 5-1 indicates that weights converge to constants when they are integrated backwards. Figure 5-2 and 5-3 show that the system's state response, including x_1 , x_2 and x_3 , are all bounded. It can be seen that the state x_3 's steady value can be controlled by changing a in eq. (42). When $a = 0.61$, x_3 does converge to the origin. Figure 5-4 shows the nearly-optimal control converges to zero.

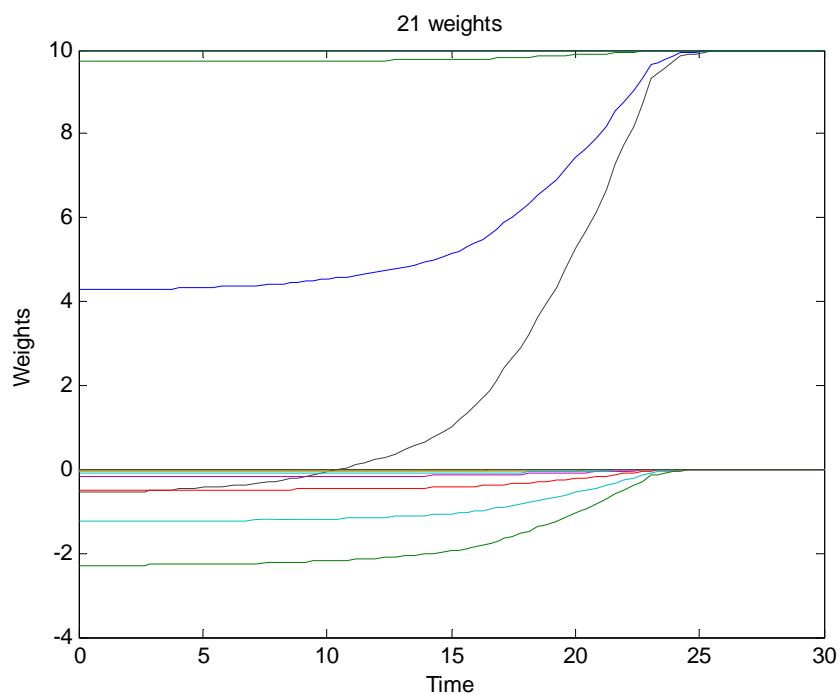


Figure 5-1 Nonlinear System Weights

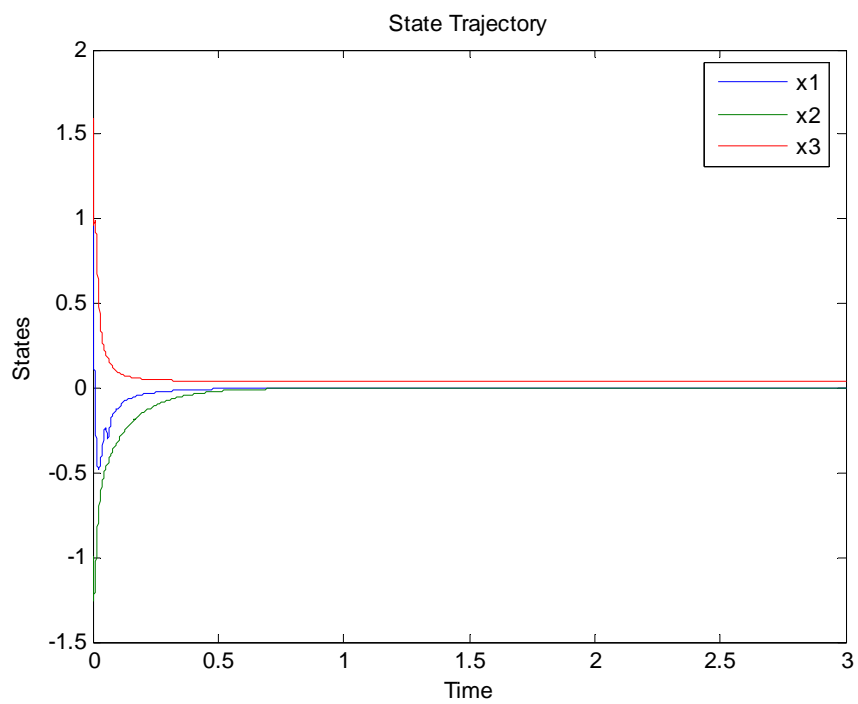


Figure 5-2 State trajectories under the time folding control ($a = 0.5$)

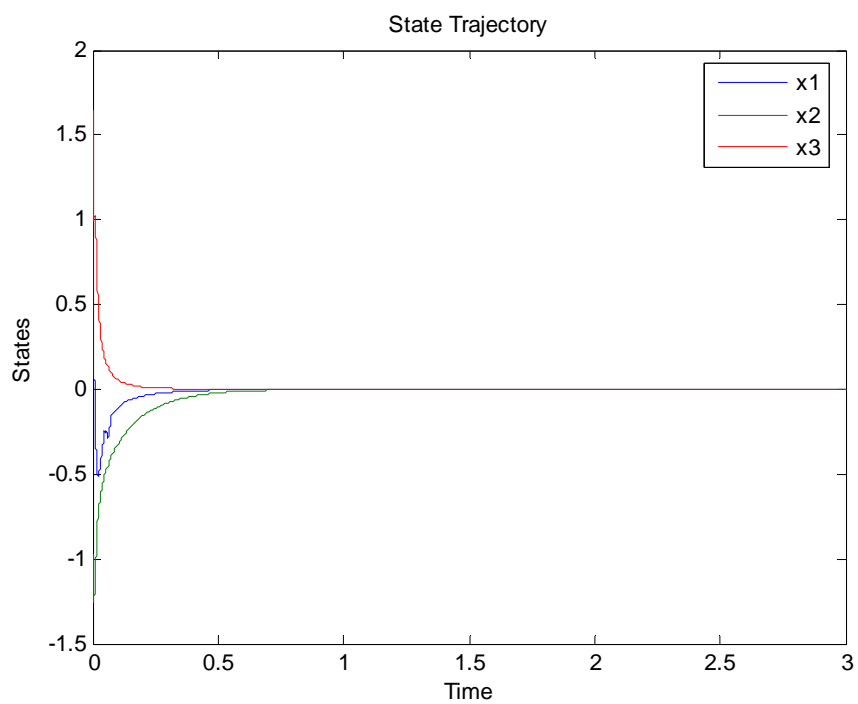


Figure 5-3 State trajectories under the time folding control ($a = 0.61$)

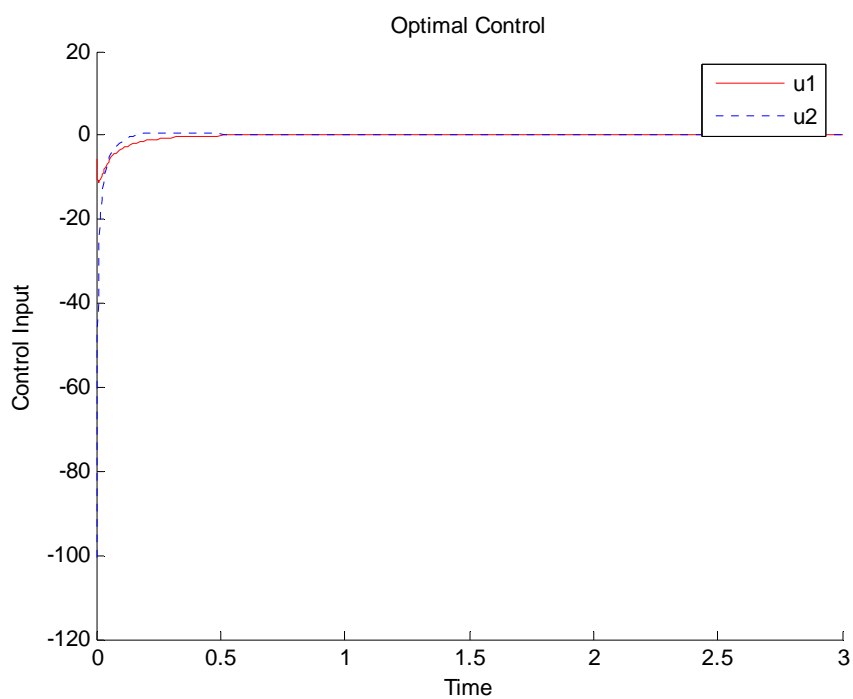


Figure 5-4 Optimal NN Control Law

5.5. Conclusion

In this chapter, nonholonomic chained systems are solved by investigating uniform complete controllability and developing relevant results. Illustrative example shows that linear controllability does not hold for stabilization of the chained system but can be recovered under time scaling transformation. The time-folding method yields a continuous asymptotically-stabilizing control without the need of using any state transformation.

CHAPTER 6

CONTRIBUTIONS AND FUTURE WORK

6.1. Contributions

In this dissertation, neural networks are used to obtain optimal control with unconstrained and constrained control. The main theme of this research is based on solving a related Hamilton-Jacobi-Bellman or Hamilton-Jacobi-Isaacs equation of the corresponding finite-horizon zero-sum game. It is shown that the neural network approximation converges uniformly to the game-value function and the resulting nearly optimal feedback controller provides closed-loop stability. The result is a nearly optimal controller with time-varying coefficients that is solved a priori offline.

The contribution of this research can be summarized in the following points:

1. In chapter two, it is shown that the HJB equation can be solved by using neural networks, fixed-final time optimal control laws are achieved. The result is a neural network feedback controller that has time-varying coefficients found by a priori offline tuning. Convergence results are shown.
2. In chapter three, neural networks are used to approximately solve the finite-horizon optimal H_∞ state feedback control problem. The method is based on solving a related Hamilton-Jacobi-Isaacs equation of the corresponding finite-horizon zero-sum game. The neural network approximates the corresponding game value function on a certain domain of the state-space

and results in a control computed as the output of a neural network. The results of this chapter are applied to the Rotational/Translational Actuator benchmark nonlinear control problem.

3. In chapter four, we use NN to approximately solve the time-varying HJ equation for constrained control nonlinear systems. It is shown that using a NN approach, one can simply transform the problem into solving a nonlinear ordinary differential equation (ODE) backwards in time.
4. In chapter five, time-folding method is introduced to solve chained system problem. Regularly, the state with product term can't converge to zero perfectly, with this method, the issue was solved, the state can converge to any value as we like.

6.2. Future Work

In future, stability, controllability and optimality of chained system with time-folding method need to be proved.

Further more, one can consider the case of online training of the neural network. So far, the algorithms considered in this dissertation were offline techniques. Also, it would be interesting to apply the algorithm to discrete-time nonlinear system.

Also, nonlinear control system in discrete case can be studied. We can consider the use of nonlinear networks towards obtaining nearly optimal solutions to the control of nonlinear discrete-time systems. The method can be based on least-squares approximation solution of HJB equation..

REFERENCES

- [1] Abu-Khalaf, M., Lewis, F. L., “Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach”, *Automatica*, Vol41, pp. 779-791, 2005.
- [2] Adams, R., Fournier, J., *Sobolev spaces*, 2nd ed., New York: Academic Press, 2003.
- [3] Aguilar, L. T., Orlov, Y., Acho, L., “Nonlinear H_∞ -control of nonsmooth time-varying systems with application to friction mechanical manipulators”, *Automatica*, 39, pp. 1531-1542, 2003.
- [4] Arnold, V. I., *Ordinary Differential Equations*. MIT Press. 1973.
- [5] Athans, M., Falb, P., L., *Optimal Control: An Introduction to the Theory and its Applications*, McGraw-Hill Book Company, New York, 1966.
- [6] Balas, M. J., “Modal control of certain flexible dynamic systems”, *SIAM Journal of Control and Optimization*, Vol 16, No. 3, pp. 450-462, 1978.
- [7] Ball, J., W. Helton, “Viscosity Solutions of Hamilton-Jacobi Equations Arising in Nonlinear H_∞ -Control”, *Journal of Mathematical Systems, Estimation, and Control*, vol. 6, no. 1, pp. 1-22, 1996.
- [8] Bardi, M., Capuzzo-Dolcetta, I., *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, Boston, MA: Birkhauser, 1997.
- [9] Bartle, R., G., *The elements of real analysis*, 2nd edition, New York: Wiley, 1976.

- [10]Başar, T., P. Bernard, *H_∞ Optimal Control and Related Minimax Design Problems*, Birkhäuser, 1995.
- [11]Beard, R., *Improving the Closed-Loop Performance of Nonlinear Systems*, PhD thesis, Rensselaer Polytechnic Institute, Troy, NY 12180, 1995.
- [12]Beard, R., Saridis, G., Wen, J., “Galerkin Approximations of the Generalized Hamilton-Jacobi-Bellman Equation”, *Automatica*, Vol. 33, December, pp. 2159-2177, 1997.
- [13]Beard, R., T. McLain, Successive Galerkin approximation algorithms for nonlinear optimal and robust control. *International Journal of Control*, vol 71. no. 5, pp 717-743, 1998.
- [14]Berkemeier, M. D., and Fearing, R. S., “Tracking fast inverted trajectories of the Underactuated Acrobot, *IEEE Transactions on Robotics and Automation*, vol. 15, no. 4, pp. 740-750, 1999.
- [15]Bernstein, D. S., “Optimal nonlinear, but continuous, feedback control of systems with saturating actuators”, *International Journal of Control*, Vol 62, No. 5, pp. 1209-1216, 1995.
- [16]Bernstein, D., S., “Nonquadratic cost and nonlinear feedback control”, *International Journal of Robust and Nonlinear Control*, Vol. 3, pp. 211-229, 1993.
- [17]Bertsekas, D. P., Tsitsiklis, J. N., “Neuro-Dynamic Programming: An Overview”, *Proceedings of the IEEE Conference on. Decision & Control*, pp. 560-564, December 1995.
- [18]Bloch, A. M., Mahmut, Reyhanoglu, McClamroch, N. Harris, “Control and

- Stabilization of Nonholonomic Dynamic Systems”, *IEEE Transactions on automatic control*, Vol. 37 No. 11, November 1992.
- [19] Bloch, A. M., McClamroch, N. H., Mahmut, R., “Controllability and Stabilizability Properties of a Nonholonomic Control System”, *Proceedings of the 29th Conference on Decision & Control*, Honolulu, Hawaii, December 1999.
- [20] Bortoff, S. A., *Pseudolinearization using Spline Functions with Application to the Acrobot*, PhD Thesis, Dept. of Electrical and Computer Engineering, University of Illinois at UrbanaChampaign, 1992.
- [21] Bortoff, S., and Spong, M. W., “Pseudolinearization of the Acrobot using Spline Functions”, *IEEE Conf. on Decision and Control*, Tucson, AZ, pp. 593-598, Dec 1992.
- [22] Brewer, J. W., “Kronecker Products and Matrix Calculus in System Theory”, *IEEE Transactions on Circuits and Systems*, Vol. Cas-25, No. 9, September, 1978.
- [23] Brockett, R. W., “Asymptotic stability and feedback stabilization”, *Differential Geometric Control Theory*, R. W. Brockett, R. S. Millman, and H. J. Sussmann, Eds. Boston, MA: Birkhauser, 1983.
- [24] Bupp, R., Bernstein, D., Coppola, V., “A benchmark problem for nonlinear control design”, *International Journal of Robust and Nonlinear Control*, Vol 8, 307-310, 1998.
- [25] Chen, B. M., *Robust and H_∞ Control*, Springer, Berlin, 2000.
- [26] Chen, F. C. and Liu, C. C., “Adaptively controlling nonlinear continuous-time systems using multiplayer neural networks”, *IEEE Transactions on Automatic*

Control, Vol. 39, No. 6, pp. 1306-1310, June 1994.

- [27]Chen, Z. and Jagannathan, S., “Neural Network-based Nearly Optimal Hamilton-Jacobi-Bellman Solution for Affine Nonlinear Discrete-Time Systems”, *Proceeding of the 44th IEEE Conference on Decision and Control*, and the European Control Conference 2005, Seville, Spain, December 12-15, 2005.
- [28]Cheng, T., Lewis, F. L., “Neural Network Solution for Finite-Horizon H-Infinity Constrained Optimal Control of Nonlinear Control”, Submitted to *Journal of Control Theory and Applications*. 2006.
- [29]Cheng, T., Lewis, F. L., Abu-Khalaf, M., “A neural network solution for fixed-final time optimal control of nonlinear systems”, submitted, 2005.
- [30]Cheng, T., Lewis, F. L., Abu-Khalaf, M., “Fixed-Final Time Constrained Optimal Control of Nonlinear Systems Using Neural Network HJB Approach”, Submitted to *IEEE Transaction on Neural Networks*, 2005.
- [31]Cheng, T., Lewis, F. L., Abu-Khalaf, M., “Neural Network Solution for Finite-Horizon H_{∞} State Feedback Control of Nonlinear Systems”, Submitted to *International Journal of System Science*, 2006.
- [32]Cheng, T., Qu, Z. and Lewis, F. L., “Neural Network Solution for Suboptimal Control of Nonholonomic Chained Form System”, Submitted to *Transaction of the Institute of Measurement and Control*, Special Issue, 2006.
- [33]Dolphus, R. M., Schmitendorf, W., E., “Stability analysis for a class of linear controllers under control constraints”, *Dynamics and Control*, Vol. 5. No. 4, pp. 191-203, 1995.

- [34]Doyle, J. H., K. Glover, P. Khargonekar, B. Francis, "State-Space Solutions to Standard H_2 and H_∞ Control Problems," *IEEE Trans. Automat. Control*, vol. 34, no. 8, pp. 831-847, 1989.
- [35]Egeland, O., Berglund, E., and Sordalen, O. J., "Exponential stabilization of a nonholonomic underwater vehicle with constant desired configuration", *IEEE International Conference on Robotics*, Aug, 1994.
- [36]Escobar, G., Ortega, R., Reyhanoglu, M., "Regulation and Tracking of the Nonholonomic Double Integrator: A Field-oriented Control Approach", *Automatics*, Vol. 34. No. 1, pp. 125-131, 1998.
- [37]Fierro, R., Lewis, F. L., "Robust Practical Point Stabilization of a Nonholonomic Mobile Robot Using Neural Networks", *Journal of Intelligent and Robotic Systems* 20: 295-317, 1997.
- [38]Finlayson, B. A., *The method of weighted residuals and variational principles*, New York: Academic Press, 1972
- [39]Furuta, K., and Yamakita, M., "Swing up Control of Inverted Pendulum", *In IECON'91*, pp. 2193-2198, 1991.
- [40]Ge, S. S., "Robust adaptive NN feedback linearization control of nonlinear systems", *International Journal of Systems Science*, Vol. 27, No. 12, pp. 1327-1338, 1996.
- [41]Hill, D., P. Moylan, "The Stability of Nonlinear Dissipative Systems," *IEEE Trans. Automatic Control*, vol. 21, pp. 708-711, 1976.
- [42]Hornik, K., Stinchcombe, M., White, H., "Universal approximation of an unknown

- mapping and its derivatives using multilayer feedforward networks”. *Neural Networks*, 3, 551-560, 1990.
- [43]Huang, C. S., Wang, S., Teo, K. L., “Solving Hamilton-Jacobi-Bellman equations by a modified method of characteristics”, *Nonlinear Analysis*, 40, 279-293, 2000.
- [44]Huang, J., Lin, C. F., “Numerical approach to computing nonlinear H_∞ control laws”, *Journal of Guidance, Control, and Dynamics*, 18 (5), 989-994, 1995.
- [45]Isidori, A., Astolfi, A., “Disturbance Attenuation and H_∞ -Control via Measurement Feedback in Nonlinear Systems”, *IEEE Transactions on Automatic Control*, Vol. 37, No. 9, pp. 1283-1293, 1992.
- [46]Khalil, H. K, *Nonlinear Systems*, Prentice-Hall, Inc. Upper Saddle River, 2002.
- [47]Kim, Y. H., Lewis, F. L., Dawson. D., “Intelligent optimal control of robotic manipulators using neural networks”, *Automatic* 36, pp. 1355-1364,2000.
- [48]Kolmanovsky, I., McClamroch, N. H., “Developments in Nonholonomic Control Problems”, *IEEE Control Systems*, pp. 20-36, December 1995.
- [49]Kratochwil, K., Engelbrecht, R., and Jorgl, H., “A Reward/Punishment Learning Method to Swing up a Pendulum into its Upright Position”, *Proc. IFAC Symposium*, Sydney, Australia, 1993.
- [50]Lafferriere, G., Sussmann, H., “Motion planning for controllable systems”, *Proceedings of the 1991 IEEE, International conference on Robotics and Automation*, Sacramento, California, April 1991.
- [51]Leitmann, G., *The Calculus of Variations and Optimal Control*, Plenum Press, New York, 1981.

- [52] Lewis, F. L., Jagannathan, S., Yesildirek, A., *Neural Network Control of Robot Manipulators and nonlinear Systems*, Taylor & Francis, 1999.
- [53] Lewis, F. L., Syrmos, V. L., *Optimal Control*, John Wiley & Sons, Inc. New York, NY, 1995.
- [54] Lio, F. D., “On the Bellman equation for infinite horizon problems with unbounded cost functional”, *Applied Mathematics and Optimization*, 41, 171-197, 2000.
- [55] Liu, X., Balakrishnan, S. N., “Adaptive Critic Based Neuro-Observer”, *Proceedings of the American Control Conference*, June. 2001, pp. 1616-1621
- [56] Lyshevski, S. E., “Optimal Control of Nonlinear Continuous-Time Systems: Design of Bounded Controllers Via Generalized Nonquadratic Functionals”, *Proceedings of the American Control Conference*, pp. 205-209, June. 1998.
- [57] Lyshevski, S. E., “Optimization of a Class of Nonholonomic Dynamic Systems”, *Proceedings of the American Control Conference*, San Diego, California, June 1999.
- [58] Lyshevski, S. E., *Control Systems Theory with Engineering Applications*, Birkhauser, Boston, MA, 2001.
- [59] Lyshevski, S. E., Meyer, A. U., “Control System Analysis and Design Upon the Lyapunov Method”, *Proceedings of the American Control Conference*, pp. 3219-3223, June. 1995.
- [60] Lyshevski, S., E., “Constrained Optimization and Control of Nonlinear Systems: New Results in Optimal Control”, *Proceedings of the 35th Conference on Decision & Control*, Kobe, Japan, December 1996.

- [61]Lyshevski, S., E., “Optimal Tracking Control of Nonlinear Dynamic Systems With Control Bounds”, *Proceedings of the 38th Conference on Decision & Control*, Phoenix, Arizona USA, December 1999.
- [62]Lyshevski, S., E., “Robust Nonlinear Control of Uncertain Systems with State and Control Constraints”, *Proceedings of the 34th Conference on Decision & Control*, New Orleans, La, December 1995.
- [63]Miller, W. T., Sutton, R., Werbos, P., *Neural Networks for Control*, The MIT Press, Cambridge, Massachusetts, 1990.
- [64]Mori, S., Nishihara, H., and Furuta, K., “Control of Unstable Mechanical Systems: Control of Pendulum”, *Int. J. Control*, V. 23, pp. 673-693, 1976.
- [65]Munos, R., Baird, L. C., Moore, A., “Gradient Descent Approaches to Neural-Net-Based Solutions of the Hamilton-Jacobi-Bellman Equation”, *International Joint Conference on Neural Networks IJCNN*, Vol. 3, pp. 2152-2157, 1999.
- [66]Murray, R. M., Sastry, S. S., “Steering nonholonomic systems in chained form”, *Proceedings of the 30th Conference on Decision & Control*, Brighton, England, December 1991.
- [67]Murray, R., Sastry, S., “Steering nonholonomic systems using sinusoids”, *Proceedings of the IEEE Conference on Decision & Control*, Honolulu, HI, pp. 2097-2101, 1990.
- [68]Nakanishi, J., Fukuda, T. and Koditschek, D. E., “A brachiating robot controller, *IEEE Transactions on Robotics and Automation*, vol. 16, no. 2, pp. 109-123, 1999.

- [69]Olfati-Saber, R. and Megretski, A., “Controller design for a class of Underactuated nonlinear systems, *Proceedings of 37th IEEE Conference of Decision and Control*, pp. 4182-4187, 1998.
- [70]Parisini, T., Zoppoli, R. “Neural Approximations for Infinite-Horizon Optimal Control of Nonlinear Stochastic Systems”, *IEEE Transactions on Neural Networks*, Vol. 9, No. 6, November 1998, pp. 1388-1408.
- [71]Polycarpou, M. M., “Stable adaptive neural control scheme for nonlinear systems”, *IEEE Transactions on Automatic Control*, Vol. 41, No. 3, pp. 447-451, Mar. 1996.
- [72]Pomet, J. B., Huilot, B., Bastin, C. G., “A hybrid strategy for the feedback stabilization of nonholonomic mobile robots”, *Proceedings of the 1992 IEEE International Conference on Robotics and Automatics*, Nice, France, May 1992.
- [73]Qu, Z, Wang, J, Hull, R. A. and Martin, J, “Continuous Control Designs for Stabilizing Chained Systems: A Global State Scaling Transformation and A Time Scaling Method”, *45th IEEE Conference on Decision and Control*, San Diego, CA, USA, December 2006.
- [74]Qu, Z., Wang, J., Plaisted, C. E., and Hull, R. A., “Global-stabilizing near-optimal control design for nonholonomic chained systems”, *IEEE Transactions on Automatic Control*, Accepted as a regular paper, 2006.
- [75]Rovithakis, G. A. and Christodoulou, M. A., “Adaptive control of unknown plants using dynamical neural networks”, *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 24, No. 3, pp. 400-412, 1994.
- [76]Saber, A, Lin, Z., Teel, A., “Control of Linear Systems with Saturating Actuators”,

- IEEE Transactions on Automatic Control*, Vol 41, NO. 3, pp. 368-378, March 1996.
- [77]Saber, A., Lin, Z., Teel, A., R., “Control of linear systems with saturating actuators”, *IEEE Transaction on Automatic Control*, Vol. AC-41, No. 3, pp. 368-378, 1996.
- [78]Sandberg, Erwin W., “Notes on Uniform Approximation of Time-Varying Systems on Finite Time Intervals”, *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications*, Vol. 45, No. 8, August 1998.
- [79]Sanner, R. M. and Slotine, J. J. E., “Stable adaptive control and recursive identification using radial gaussian networks”, *Proceedings of the IEEE Conference on Decision & Control*, pp. 2116-2123, Brighton, 1991.
- [80]Saridis, G., Lee, C. S., “An Approximation Theory of optimal Control for Trainable Manipulators”, *IEEE Transactions on Systems, Man, Cybernetics*, Vol. 9, No. 3, pp. 152-159, March 1979.
- [81]Sordalen. O. J., Egeland. O., “Exponential Stabilization of Nonholonomic Zchained Systems”, *IEEE Transactions on Automatic Control*, Vol. 40, No. 1, January 1995.
- [82]Spong, M. W., “The swing up control problem for the Acrobot”, *IEEE Control Systems Magazine*, vol. 15, no. 1, pp. 49-55, 1995.
- [83]Spong, M. W., and Vidyasagar, M., *Robot Dynamics and Control*, John Wiley & Sons, Inc., New York, 1989.
- [84]Sussmann, H., Sontag, E. D., Yang, Y., “A general result on the stabilization of linear systems using bounded controls”, *IEEE Transactions on Automatic Control*,

- 39(12), pp. 2411-2425, 1994.
- [85]Tsiotras, P., Corless, M., Rotea, M., “An L_2 disturbance attenuations solution to the nonlinear benchmark problem”, *International Journal of Robust and Nonlinear Control*, Vol 8, 311-330, 1998.
- [86]Van Der Schaft, A. J., L_2 -Gain and Passivity Techniques in Nonlinear Control. London, U.K.: Springer-Verlag, 1999.
- [87]Van der Schaft, A., J., “ L_2 -Gain Analysis of Nonlinear Systems and Nonlinear State Feedback H_∞ Control”, *IEEE Transaction on Automatic Control*, Vol. 37, No. 6, June 1992.
- [88]Wiklund, M., Kristenson, A., and Astrom, K. J., “A New Strategy for Swinging up and Inverted Pendulum”, *Proc. IFAC Symposium*, Sydney, Australia, 1993.
- [89]Willems, J. C. , “Dissipative Dynamical Systems Part II: Linear Systems with Quadratic Supplies,” *Archive for Rational Mechanics and Analysis*, vol 45, no.1, pp. 352-393, 1972.
- [90]Willems, J. C., “Dissipative Dynamical Systems Part I: General Theory,” *Archive for Rational Mechanics and Analysis*, vol 45, no.1, pp. 321-351, 1972.
- [91]Zames, G., “Feedback and Optimal Sensitivity: Model Reference Transformations, Multiplicative Seminorms, and Approximate Inverses,” *IEEE Trans. Automat. Control*, vol. 26 no. 2, pp. 301-320, 1981.

BIOGRAPHICAL INFORMATION

Tao Cheng was born in P. R. China in 1976. He received his Bachelor's degree in Electrical Engineering from Hubei Institute of Technology in 1998. He then joined Beijing Polytechnic University from which he received the Master's of Science in Electrical Engineering in 2001. Currently he is working on his PH.D. degree at Automation and Robotics Research Institute at The University of Texas at Arlington. His research interest is in time-varying optimal nonlinear systems, nonholonomic vehicle systems.