Filter-Based Adaptive Optimal Control Designs for Uncertain Linear Time Invariant Systems

SUMIT KUMAR JHA

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY DELHI
MARCH 2019

© Indian Institute of Technology Delhi (IITD), New Delhi, 2019

Filter-Based Adaptive Optimal Control Designs for Uncertain Linear Time Invariant Systems

by

SUMIT KUMAR JHA
Department of Electrical Engineering

Submitted
in fulfilment of the requirements of the degree of **Doctor of Philosophy**

to the

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY DELHI
MARCH 2019

CERTIFICATE

This is to certify that the thesis entitled **Filter-Based Adaptive Optimal Control Designs for Uncertain Linear Time Invariant Systems** submitted by **Sumit Kumar Jha** to the Indian Institute of Technology Delhi, for the award of the Degree of **Doctor of Philosophy**, is a record of the bonafide research work carried out by him under my supervision and guidance. The thesis has reached the standards fulfilling the requirements of the regulations relating to the degree.

The results contained in this thesis have not been submitted either in part or in full to any other University or Institute for the award of any degree or diploma to the best of my knowledge.

Dr. Shubhendu Bhasin

Department of Electrical Engineering,

Indian Institute of Technology Delhi.

(Supervisor)

ABSTRACT

The theory of optimal control, an important part of modern control theory, has found multifarious applications in diverse field as engineering, aerospace, modern industries, medicine, social and economic development etc, due to its inherent ability to find the best possible control action which optimizes the performance of the system under consideration in presence of certain predefined physical constraints. A typical optimal control design requires complete system information to provide an offline solution. In contrast, adaptive optimal control (AOC) methodologies can be used for dynamic controller synthesis for uncertain systems by estimating/approximating system and control parameters online.

The focus of this thesis is to investigate the problem of AOC design for continuous-time (CT) linear time invariant (LTI) systems with uncertain dynamics. Most of the traditional AOC schemes for CT LTI systems require knowledge of system dynamics, restrictive condition of persistence of excitation (PE), finite-window integrals (FWI). This research intends to address some of the aforementioned limitations by proposing various novel filter-based AOC algorithms. The contributions of the thesis can be divided into two major parts.

- A novel continuous-time adaptive linear quadratic regulator (LQR) is proposed for uncertain LTI systems, which guarantees exponential stability of closed-loop system under PE condition. The proposed work uses gradient-based update laws for online adaptation of control policy without using an initial stabilizing control policy. Since the PE condition lacks in its applicability for practical scenarios, filter-based data-driven direct and indirect AOC methodologies are presented next, which utilize past data stored along the system trajectory for estimating the unknown controller parameters. The proposed schemes use past data-driven parameter tuning laws, simultaneously with the low-pass filters based state derivative estimator, to update the control policy online, while guaranteeing parameter convergence without requiring the PE condition.

- Further, this thesis proposes two-tier filter-based iterative AOC methods to solve adaptive LQR problems for uncertain LTI systems. The proposed algorithms guarantee global parameter convergence under excitation assumption, unlike the AOC schemes previously proposed in this thesis. The crucial features of these methods are the strategic design of two-layered low pass filters that obviate the need of FWIs, which are either memory-intensive or possibly delay the fulfillment of the full rank condition, and requirement of intelligent data-storage. The design of novel PI-based controller for partially known systems is presented first, in which parameter convergence is achieved under excitation assumption on the regressor in continuous time-intervals, a weaker assumption than the classical PE condition. This work is further extended for LTI systems with unknown system and control matrices by introducing a novel filter-based iterative algorithm without using the restrictive PE condition. The convergence of control parameters are assured under a milder condition of initial excitation (IE), which is argued to be a necessary condition for the full rank condition, typically required for past data storage based AOC methods.

सार

इष्टतम नियंत्रण का सिद्धांत, आधुनिक नियंत्रण सिद्धांत का एक महत्वपूर्ण हिस्सा है, विभिन्न क्षेत्रों में इंजीनियरिंग, एयरोस्पेस, आधुनिक उद्योग, चिकित्सा, सामाजिक और आर्थिक विकास आदि के रूप में बहुपक्षीय अनुप्रयोगों को पाया गया है, जो कि सर्वोत्तम संभव नियंत्रण कार्रवाई खोजने की अपनी अंतर्निहित क्षमता के कारण है। कुछ पूर्वनिर्धारित भौतिक बाधाओं की उपस्थिति में विचाराधीन प्रणाली के प्रदर्शन का अनुकूलन करता है। एक विशिष्ट इष्टतम नियंत्रण डिज़ाइन को ऑफ़लाइन समाधान प्रदान करने के लिए संपूर्ण सिस्टम जानकारी की आवश्यकता होती है। इसके विपरीत, अनुकूली इष्टतम नियंत्रण (एओसी) कार्यप्रणाली का उपयोग ऑनलाइन प्रणाली का अनुमान लगाने और अनुमान लगाने के लिए अनिश्चित प्रणाली के लिए गतिशील नियंत्रक संश्लेषण के लिए किया जा सकता है।

इस थीसिस का फोकस निरंतर डायनामिक्स के साथ निरंतर-समय (सीटी) रैखिक समय अपरिवर्तनीय (एलटीआई) सिस्टम के लिए एओसी डिजाइन की समस्या की जांच करना है। सीटी एलटीआई सिस्टम के लिए अधिकांश पारंपरिक एओसी योजनाओं में सिस्टम की गतिशीलता, उत्तेजना की दृढ़ता (पीई), परिमित-विंडो इंटीग्रल्स (एफडब्ल्यूआई) के ज्ञान की आवश्यकता होती है। यह शोध विभिन्न उपन्यास फ़िल्टर-आधारित AOC एल्गोरिदम का प्रस्ताव करके कुछ पूर्वोक्त सीमाओं को संबोधित करने का इरादा रखता है। थीसिस के योगदान को दो प्रमुख भागों में विभाजित किया जा सकता है।

- अनिश्चित एलटीआई सिस्टम के लिए एक उपन्यास निरंतर-समय अनुकूली रैखिक द्विघात नियामक (एलक्यूआर) प्रस्तावित है, जो पीई स्थिति के तहत बंद लूप सिस्टम की घातीय स्थिरता की गारंटी देता है। प्रस्तावित कार्य एक प्रारंभिक स्थिरीकरण नियंत्रण नीति का उपयोग किए बिना नियंत्रण नीति के ऑनलाइन अनुकूलन के लिए ढाल-आधारित अद्यतन कानूनों का उपयोग करता है। चूंकि पीई स्थिति व्यावहारिक परिदृश्यों के लिए इसकी प्रयोज्यता में कमी है, इसलिए फ़िल्टर-आधारित डेटा-चालित प्रत्यक्ष और अप्रत्यक्ष एओसी कार्यप्रणाली अगले प्रस्तुत की जाती हैं, जो अज्ञात नियंत्रक मापदंडों का अनुमान लगाने के लिए सिस्टम प्रक्षेपवक्र में संग्रहीत पिछले डेटा का उपयोग करते हैं। प्रस्तावित योजनाएँ पीई-शर्त की आवश्यकता के बिना पैरामीटर अभिसरण की गारंटी देते हुए नियंत्रण नीति को ऑनलाइन अपडेट करने के लिए कम-पास फिल्टर आधारित राज्य व्युत्पन्न अनुमानक के साथ-साथ पिछले डेटा-संचालित पैरामीटर ट्यूनिंग कानूनों का उपयोग करती हैं।
- इसके अलावा, यह थीसिस अनिश्चित एलटीटी प्रणालियों के लिए अनुकूली LQR समस्याओं को हल करने के लिए दो-स्तरीय फिल्टर-आधारित पुनरावृत्त AOC तरीकों का प्रस्ताव करती है। प्रस्तावित

एल्गोरिदम, इस धारणा में प्रस्तावित AOC योजनाओं के विपरीत, उत्तेजना धारणा के तहत वैश्विक पैरामीटर अभिसरण की गारंटी देता है। इन तरीकों की महत्वपूर्ण विशेषताएं दो-स्तरित कम पास फिल्टर की रणनीतिक डिजाइन हैं जो एफडब्ल्यूआई की आवश्यकता को कम करती हैं, जो या तो मेमोरी-इंटेंसिव हैं या संभवत: पूर्ण रैंक की स्थिति की पूर्ति में देरी, और बुद्धिमान डेटा-स्टोरेज की आवश्यकता। आंशिक रूप से ज्ञात प्रणालियों के लिए उपन्यास पीआई-आधारित नियंत्रक का डिज़ाइन पहले प्रस्तुत किया गया है, जिसमें पैरामीटर अभिसरण निरंतर समय-अंतराल में प्रतिगमन पर उत्तेजना धारणा के तहत प्राप्त किया जाता है, शास्त्रीय पीई स्थिति की तुलना में एक कमजोर धारणा है। प्रतिबंधात्मक पीई स्थिति का उपयोग किए बिना एक उपन्यास फिल्टर-आधारित पुनरावृत्ति एल्गोरिथ्म की शुरुआत करके इस कार्य को पूरी तरह से अज्ञात एलटीआई प्रणालियों के लिए आगे बढ़ाया गया है। नियंत्रण मापदंडों के अभिसरण को प्रारंभिक उत्तेजना (IE) की एक मामूली स्थिति के तहत आश्वासन दिया जाता है, जो पूर्ण रैंक की स्थिति के लिए एक आवश्यक शर्त होने का तर्क देता है, आमतौर पर पिछले डेटा भंडारण आधारित एओसी विधियों के लिए आवश्यक है।

Contents

Acknowledgements **iii**

Abstract **v**

List of Figures **xi**

List of Abbreviations **xiii**

List of symbols **xv**

1 Introduction and Preliminaries **1**

1.1 Introduction 1

1.2 Relevant Literature and Motivation 2

1.3 Preliminaries 5

1.3.1 Optimal Control Problem Formulation 5

1.3.2 Policy Iteration 6

1.3.3 Persistence of Excitation and Initial Excitation 8

1.3.4 Concurrent Learning 9

1.4 Adaptive Linear Quadratic Regulator 9

1.4.1 Iterative AOC for CT LTI systems 10

1.4.2 Continuous AOC for CT LTI systems 11

1.5 Contribution of the Thesis 13

1.5.1 Filter-based AOC Designs . . . 13
1.5.2 Chapter-wise Contribution . . . 14

2 Continuous Adaptive LQR Designs 19

2.1 Introduction . . . 19
2.2 Q-learning Based Adaptive LQR Design . . . 20
2.2.1 Q-learning . . . 20
2.2.2 Problem formulation and methodology . . . 21
2.2.3 State derivative estimation . . . 24
2.2.4 Simulation . . . 26
2.3 Adaptive LQR for Continuous-Time Systems with Uncertain Dynamics . . . 31
2.3.1 Preliminaries and problem formulation . . . 31
2.3.2 Adaptive LQR design for unknown LTI systems . . . 33
2.3.3 Development of controller parameter estimation error dynamics . . . 36
2.3.4 Lyapunov stability analysis . . . 38
2.3.5 Comparison with existing literature . . . 42
2.3.6 Simulation . . . 43
2.4 Summary . . . 47

3 Data-Driven Filter-based Continuous Adaptive LQR 49

3.1 Introduction . . . 49
3.2 Data-Driven Filter-based Indirect Adaptive LQR for Uncertain Systems . . . 50
3.2.1 Problem formulation and methodology . . . 51
3.2.2 Filter-based system parameter estimation . . . 51
3.2.3 Data-driven indirect adaptive optimal control design . . . 54
3.2.4 Simulation . . . 61

3.3 Data-Driven Filter-Based Direct Adaptive LQR for Uncertain Systems 63
3.3.1 Data-driven filter-based direct LQR design 64
3.3.2 Simulation 71
3.4 Summary 73

4 Two-Tier Filter-Based Iterative Adaptive LQR 75

4.1 Introduction 75
4.2 Filter-Based Policy Iteration Technique for Adaptive LQR 77
4.2.1 First-layer filtering 79
4.2.2 Second-layer filtering 80
4.2.3 Convergence analysis 85
4.2.4 Closed-loop stability analysis 86
4.2.5 Boundedness of auxiliary signals 89
4.2.6 Simulation 89
4.3 Filter-Based Explorized Policy Iteration Algorithm for Adaptive LQR 92
4.3.1 First-layer filtering 94
4.3.2 Second-layer filtering 95
4.3.3 Convergence analysis 97
4.3.4 Closed-loop stability analysis 99
4.3.5 Boundedness of auxiliary signals 102
4.3.6 Simulation 103
4.4 Analytical Comparison with Existing Data-Storage Based PI Algorithms 105
4.4.1 FWIs vs first-layer filters 105
4.4.2 Rank condition vs excitation condition 107
4.5 Summary 109

5 Two-Tier Filter-Based Iterative Adaptive LQR with Initial Excitation for Unknown LTI Systems **111**

5.1 Introduction . . . 111

5.2 Preliminary Definitions . . . 113

5.3 Filter-Based AOC Algorithm Using IE Condition (FAOC-IE) . . . 114

5.3.1 First-layer filtering . . . 116

5.3.2 Second-layer filtering . . . 117

5.4 Convergence Analysis . . . 124

5.5 Boundedness of Auxiliary Signals . . . 125

5.6 Analytical Comparison with Existing Data-Storage Based PI Algorithms . . . 125

5.7 On-Policy Implementation of FAOC-IE . . . 126

5.8 Simulation . . . 127

5.9 Summary . . . 128

6 Conclusion and Future Directions of Research **133**

6.1 Conclusion . . . 133

6.2 Future Work . . . 135

Bibliography **137**

Appendix **147**

A Evaluation of bound for $d_4 \left\| vec(\tilde{K}) \right\| \Sigma_{per}$ . . . 147

List of Figures

1.1 Typical architecture of iterative AOC. 10

1.2 Typical architecture of continuous AOC. 12

2.1 System states trajectories ($\gamma = 0$). 27

2.2 Parameter estimate $\hat{P}_A$ for $\gamma = 0$ ('✡' denotes the optimal values). 27

2.3 Parameter estimate $\hat{P}_B$ for $\gamma = 0$ ('✡' denotes the optimal values). 28

2.4 State derivative estimation error $\dot{\tilde{x}}$ ($\gamma = 0$). 28

2.5 System states trajectories ($\gamma = 0.4$). 28

2.6 Parameter estimate $\hat{P}_A$ for $\gamma = 0.4$ ('✡' denotes the optimal values). 29

2.7 Parameter estimate $\hat{P}_B$ for $\gamma = 0.4$ ('✡' denotes the optimal values). 29

2.8 State derivative error estimation $\dot{\tilde{x}}$ ($\gamma = 0.4$). 29

2.9 PE signal. 30

2.10 PE signal. 44

2.11 The evolution of parameter estimate $\hat{K}(t)$ for the proposed method. 44

2.12 System state trajectories for the proposed method. 45

2.13 Comparison of the parameter estimation error norms between [1] and the proposed method. 46

2.14 Comparison of the control inputs between [1] and the proposed method. 46

3.1 The evolution of system parameter estimates $\hat{\theta}(t)$. 61

3.2 The evolution of the system states $x(t)$. 62

3.3 The evolution of optimal control parameter estimates $\hat{P}(t)$. 62

3.4 The evolution of system states $x(t)$. 71

3.6 The evolution of Riccati matrix estimate $\hat{P}(t)$. 72

3.5 The evolution of optimal control gain estimate $\hat{K}(t)$. 72

4.1 Flowchart of Algorithm 4.1. 84

4.2 Architecture of the proposed two-tier filter-based PI algorithm. 85

4.3 Evolution pattern of Lyapunov functions $V_k(t)$, $\forall k \in \mathbb{W}$. 88

4.4 The evolution of system states $x(t)$. 90

4.5 The evolution of optimal control gain K_k. 91

4.6 The evolution of parameter matrix P_k ($P_k(1,1)$, $P_k(1,2)$, $P_k(1,3)$). 91

4.7 The evolution of parameter matrix P_k ($P_k(2,2)$, $P_k(2,3)$, $P_k(3,3)$). 92

4.8 Architecture of the proposed two-tier filter-based explorized PI algorithm 98

4.9 The evolution of system states $x(t)$. 103

4.10 The evolution of optimal control gain K_k. 104

4.11 The evolution of parameter matrix P_k ($P_k(1,1)$, $P_k(1,2)$, $P_k(1,3)$). 104

4.12 The evolution of parameter matrix P_k ($P_k(2,2)$, $P_k(2,3)$, $P_k(3,3)$). 104

5.1 The evolution of system states $x(t)$. 128

5.2 The evolution of optimal control gain matrix K_k ($K_k(1,1)$, $K_k(1,2)$, $K_k(1,3)$), $K_k(1,4)$). 129

5.3 The evolution of optimal control gain matrix K_k ($K_k(2,1)$, $K_k(2,2)$, $K_k(2,3)$), $K_k(2,4)$). 129

5.4 The evolution of parameter matrix P_k ($P_k(1,1)$, $P_k(1,2)$, $P_k(1,3)$), $P_k(1,4)$). . . 130

5.5 The evolution of parameter matrix P_k ($P_k(2,2)$, $P_k(2,3)$, $P_k(2,4)$). 130

5.6 The evolution of parameter matrix P_k ($P_k(3,3)$, $P_k(3,4)$, $P_k(4,4)$). 131

List of Abbreviations

ADP: Approximate Dynamic Programming

AOC: Adaptive Optimal Control/Controller

ARE: Algebraic Riccati Equation

CT: Continuous-Time

FWI: Finite Window Integral

HJB: Hamilton-Jacobii-Bellman

IE: Initial Excitation

LQR: Linear Quadratic Regulator

LTI: Linear Time Invariant

PE: Persistence of Excitation

PI: Policy Iteration

RL: Reinforcement Learning

UUB: Uniformly Ultimately Bounded

VI: Value Iteration

List of Symbols

$\mathbb{R}$	Real Space
$\mathbb{R}^n$	Real Vector Space of dimension n
$\mathbb{R}^{n\times m}$	Real Matrix Space of dimension $n \times m$
$\mathbb{W}$	Set of whole numbers
$\mathbf{I}$	Identity Matrix with appropriate dimension
$\lambda_{\min}(\mathbf{\Psi})$	Minimum eigenvalue of the matrix $\mathbf{\Psi}$
$\lambda_{\max}(\mathbf{\Psi})$	Maximum eigenvalue of the matrix $\mathbf{\Psi}$