MODELING AND IDENTIFICATION OF BIOMOLECULAR SYSTEMS

ABHISHEK DEY

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY DELHI
MAY 2019

MODELING AND IDENTIFICATION OF BIOMOLECULAR SYSTEMS

by

ABHISHEK DEY

DEPARTMENT OF ELECTRICAL ENGINEERING

Submitted

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

to the

INDIAN INSTITUTE OF TECHNOLOGY DELHI

MAY 2019

Certificate

This is to certify that the thesis entitled **"Modeling and Identification of Biomolecular Systems"**, submitted by **Abhishek Dey** to the Indian Institute of Technology Delhi, for the award of the degree of **Doctor of Philosophy** in Electrical Engineering, is a record of the original, bonafide research work carried out by him under my supervision and guidance. The thesis has reached the standards fulfilling the requirements of the regulations related to the award of 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. Shaunak Sen

Associate Professor,

Department of Electrical Engineering,

Indian Institute of Technology Delhi.

(Thesis Supervisor)

Acknowledgements

My life at IIT Delhi has been a true learning experience and I am grateful to many people for being a part of this.

First, I would like to express my heartiest gratitude for my supervisor, Prof. Shaunak Sen for his valuable guidance and support. He helped come up with the problems addressed in this thesis and always provided valuable feedback for the research work. He also gave me numerous important advises to develop skills such as writing and presenting. My research committee members, Prof. Indra Narayan Kar, Prof. S. Janardhanan and Prof. James Gomes have been immensely helpful giving insightful feedback over the course of last five years.

I acknowledge Atul Agarwal, Kushal Chakrabarti and Krishan Kumar Gola who did their M. Tech in Control and Automation group, IIT Delhi for collaborating in my work. I am grateful for having collaborated with Dr. Rishi Relan of Technical University of Denmark. I thank Venkat and Abhilash for critical reading of manuscripts and providing important feedback. I especially thank Joyjit, Nalin and all the former and current members of control and automation group for making my stay in IIT Delhi memorable. I thank Mr. Sunil and Virender of Control laboratory for their help in doing official works on so many occasions.

Finally, I thank my family for their eternal support. I always had the blessings of my late grandmother in whatever I did. My parents have always encouraged and supported me to pursue research, no words are enough to express my gratitude for them. My wife Thumree, being a researcher, always could relate to my worries and was still able to bring out the positive invariably. Without her love and encouragement my journey as a Ph.D. student would have been far more difficult.

Abhishek Dey

Abstract

Modeling and identification are integral in understanding any natural system and in designing new ones. In case of biomolecular systems, challenges arise due to their nonlinear, complex, and stochastic nature. Given these challenges, the general problems of estimating parameters from noisy measurements and obtaining approximate models in biomolecular systems are relatively unclear. Particularly, we answer three questions in this thesis relating to modeling and identification of biomolecular systems. Firstly, we address the problem of predicting bacterial growth response to dynamic environmental conditions that is important in designing robust synthetic circuits. Secondly, we address the challenge of parameter estimation in biomolecular systems from noisy data investigating the dependence of noise on system states and parameters. Thirdly, we address the problem of approximating nonlinear and complex biomolecular systems to useful linear models. To address these problems, we combine theoretical, computational, and experimental approaches. More specifically, to predict bacterial growth when temperature is varying dynamically, we estimate parameters in phenomenological nonlinear growth models at different fixed temperatures from experimental data and use these estimates for prediction. Next, we use the information of process noise dependence on states and parameters in the Langevin framework to adapt the standard Extended Kalman filtering algorithm by updating the process noise covariance at each time step. We find that the updated filter can achieve a balance between mean square estimation error and parameter convergence time and we discuss its optimality. Finally, we use the describing function-based technique to approximate nonlinear input-output maps in biomolecular systems using finite amplitude inputs. We find that the approximation error using describing function technique is lesser than standard linearisation and propose an analytical bound for this approximation error. The describing function-based approximation provides useful insight for the biomolecular systems investigated such as an account of the pure delay involved in a positive transcriptional feedback circuit, a bound of oscillation amplitude for biomolecular oscillators. Overall, the modeling and identification problems addressed in this thesis will aid in the analysis and design of biomolecular networks.

सार

मॉडलिंग और पहचान किसी भी प्राकृतिक प्रणाली को समझने और नए डिजाइन करने में अभिन्न हैं। बायोमोलेक्युलर सिस्टम के मामले में, उनके अरेखीय, जटिल और स्टोचैस्टिक प्रकृति के कारण चुनौतियां पैदा होती हैं। इन चुनौतियों को देखते हुए, नॉइसि डेटा से पैरामीटर्स का अनुमान लगाने और बायोमोलेक्युलर सिस्टम में अनुमानित मॉडल प्राप्त करने की सामान्य समस्याएं अपेक्षाकृत अस्पष्ट हैं। विशेष रूप से, हम इस थीसिस में तीन सवालों के जवाब मॉडलिंग और बायोमोलेक्युलर सिस्टम की पहचान से संबंधित हैं। सबसे पहले, हम गतिशील पर्यावरणीय परिस्थितियों के लिए बैक्टीरिया के विकास की प्रतिक्रिया की भविष्यवाणी की समस्या को संबोधित करते हैं जो मजबूत सिंथेटिक सर्किट को डिजाइन करने में महत्वपूर्ण है। दूसरे, हम सिस्टम स्टेटस और पैरामीटर्स पर नॉइज़ की निर्भरता की जांच नॉइसि डेटा से बायोमोलेक्युलर सिस्टम में पैरामीटर आकलन की चुनौती को संबोधित करते हैं। तीसरा, हम उपयोगी रैखिक मॉडल के लिए गैर-रेखीय और जटिल बायोमोलेक्युलर सिस्टम को अनुमानित करने की समस्या को संबोधित करते हैं। इन समस्याओं को दूर करने के लिए, हम सैद्धांतिक, कम्प्यूटेशनल और प्रयोगात्मक दृष्टिकोणों को जोड़ते हैं। अधिक विशेष रूप से, बैक्टीरिया की वृद्धि की भविष्यवाणी करने के लिए जब तापमान गतिशील रूप से भिन्न हो रहा है, हम प्रायोगिक डेटा से अलग-अलग निश्चित तापमान पर घटनात्मक अरेखीय विकास मॉडल में मापदंडों का अनुमान लगाते हैं और भविष्यवाणी के लिए इन अनुमानों का उपयोग करते हैं। इसके बाद, हम हर समय कदम पर प्रक्रिया नॉइज़ सहसंयोजक को अद्यतन करके मानक विस्तारित कलामन फ़िल्टरिंग एल्गोरिदम को अनुकूलित करने के लिए लैंगविन फ्रेमवर्क में स्टेटस और पैरामीटर्स पर नॉइज़ की निर्भरता की जानकारी का उपयोग करते हैं। हम पाते हैं कि अपडेट किया गया फ़िल्टर माध्य वर्ग अनुमान त्रुटि और पैरामीटर अभिसरण समय के बीच एक संतुलन प्राप्त कर सकता है और हम इसकी इष्टतमता पर चर्चा करते हैं। अंत में, हम परिमित आयाम इनपुट का उपयोग करते हुए बायोमोलेकुलर सिस्टम में नॉनलाइनियर इनपुट-आउटपुट मैप्स के अनुमानित डिस्क्रिबिंग फ़ंक्शन-आधारित तकनीक का उपयोग करते हैं। हम पाते हैं कि डिस्क्रिबिंग फंक्शन तकनीक का वर्णन करते हुए सन्निकटन त्रुटि मानक रैखिककरण की तुलना में कम है और इस सन्निकटन त्रुटि के लिए एक विश्लेषणात्मक बाउंड प्रस्तावित करता है। डिस्क्रिबिंग फ़ंक्शन-आधारित सन्निकटन जैव-आणविक प्रणालियों के लिए उपयोगी अंतर्दृष्टि प्रदान करता है, जिसकी जांच एक सकारात्मक ट्रांसक्रिप्शनल फीडबैक सर्किट में शामिल शुद्ध विलंब के एक खाते के रूप में की जाती है, जो बायोमोलेक्यूलर ऑसिलेटर्स के लिए दोलन आयाम का एक समूह है। कुल मिलाकर, इस थीसिस में संबोधित मॉडलिंग और पहचान की समस्याएं बायोमोलेक्यूलर नेटवर्क के विश्लेषण और डिजाइन में सहायता करेगी।

Contents

List of Figures

List of Tables