Editorial Policy

Aims and Scope

Journal of "Control and Optimization in Applied Mathematics (COAM)" is published twice a year (Spring-Autumn) by Payame Noor University (PNU). The COAM endeavors to publish significant research of broad interests in applied mathematics in the fields of Control and Optimization. For more information, one can see the Aims and Scope at the journal's website.

Peer Review

All submitted manuscripts will usually be peer reviewed critically by at least three referees. A manuscript can be sent to referees only when it meets the COAM's standard requirements; otherwise, it will be returned to the author for revision to meet journal requirement. The editor usually informs the authors regarding comments from the referees. If necessary, the authors would be asked to revise the manuscript according to the referees' suggestions. A cover letter that addresses the issues, point by point, raised in the suggestions should be included. This revision will be forwarded to the referee who will recheck the same. Revised manuscripts returned after two months will be considered as a new submission.

Submission

Papers are accepted for editorial consideration through the journal's website:
http://mathco.journals.pnu.ac.ir

Copyright

Indexing and Abstracting Services

The COAM is abstracted and indexed by indexed by "MSRT", Iranian Ministry of Science Research and Technology of IRAN (reference No.:3/155605 date:29 Sep 2018), Elmi-Pajooheshi in persian and also Islamic World Science Citation Center (ISC).

In the name of God

Control and Optimization in Applied Mathematics

(COAM)

Volume 4, Number 2, Autumn - Winter 2019

The journal of "Control and Optimization in Applied Mathematics-COAM" is authorized under the registration No. 91/35952 dated 1391/11/30 A.H.S. (Feb 18, 2013) by the Ministry of Culture and Islamic Guidance. This academic journal is indexed by "MSRT", Iranian Ministry of Science Research and Technology of IRAN (reference No.:3/155605 date:29 Sep 2018), Elmi-Pajooheshi in persian. This journal endeavours to publish significant research of board interests in applied mathematics in the fields of Control and Optimization. The COAM is published semi-annual, one volume is published each year, and each volume consists of two issues.

Period of publication:	Semi-annual
Subject :	Mathematics
p-ISSN:	2383-3130
e-ISSN:	2538-5615
Language:	English
Eighth Issue:	Autumn - Winter 2019
Owner :	Payame Noor University
Phone:	+9851-31837262
Fax:	+9851-38683001
Editorial Office:	Emam Khomeini BLV., Ave. 40, No. 9, Payame Noor University (PNU), Mashhad, Iran.
P.O. Box:	919
Website:	http://mathco.journals.pnu.ac.ir
Number of Offprints:	25 copies
E-mail Address:	mathco@pnu.ac.ir
Account Number:	2178609001007 Melli Bank (Branch of Banafshe)
Sheba Number:	IR 42 0170 0000 0217 8609 07
LaTeX Designer:	Sayyed Mahdi Moosavi

Payame Noor University Research Journals' Publication Ethics

This publication ethics is a commitment which draws up some moral limitations and responsibilities of research journals. The text is adapted according to the "Standard Ethics", approved by the Ministry of Science, Research and Technology, and the publication principles of Committee on Publication Ethics (COPE).

1 Introduction

Authors, Reviewers, editorial boards and editor-in-chiefs ought to know and commit all principles of research ethics and related responsibilities. Article submission, review of reviewers and editor-in-chief's acceptance or rejection, are considered as journals law compliance otherwise the journals have all the rights.

2 Authors Responsibilities

- Authors should present their works in accordance with journal's standards and title.
- Authors should ensure that they have written their original works/researches. Their works/researches should also provide accurate data, underlying other's references.
- Authors are responsible for their works' accuracy.

Note 1: Publishing an article is not known as acceptance of its contents by journal.

- Duplicate submission is not accepted. In other words, none of the article's' parts, should not carry on reviewing or publishing elsewhere.
- Overlapping publication, where the author uses his/her previous findings or published date with changes, is rejected.
- Authors are asked to have authors' permission for an accurate citation. When using ones direct speech, a quotation mark (" ") is necessary.
- Corresponding author should ensure that the complete information of all involved authors in the article.

Note 2: Do not write the statement of "Gift Authorship" and do not omit the statement of "Ghost Authorship".

- Corresponding author is responsible for the priorities of co-authors after their approval.
- Paper submission means that all of the authors have satisfied whole financial and local supports and have introduced them.

- Author(s) is/are responsible for any fault or inaccuracy of the article and in this case, journal's authorities should be informed immediately.
- Author(s) is/are asked to provide and reserve raw data one year after publication, in order to be able to respond journal audiences' questions.

3 Research and Publication Misconduct

Author(s) should avoid the research and publication misconduct. If some cases of research and publication misconduct occur within each steps of submission, review, edition or publication, journals have the right to legal action. The cases are listed as below:

- Fabrication: Fabrication is the practice of inventing data or results and reporting them in the research. Both of these misconducts are fraudulent and seriously alter the integrity of research. Therefore, articles must be written based on original data and use of falsified or fabricated data is strongly prohibited.
- Falsification: Falsification is the practice of omitting or altering research materials, equipment, data, or processes in such a way that the results of the research are no longer accurately reflected in the research record.
- Plagiarism: Plagiarism is the act of taking someone else's writing, conversation, idea, claims or even citations without any acknowledgment or explanation of the work producer or speaker.
- Wrongful Appropriation: Wrongful appropriation occurs when author(s) benefits another person's efforts and after a little change and manipulations in the research work, publish it on his/her own definitions
- False Attribution: It represents that a person is the author of a work but she/ he was not involved in the research.

4 Reviewers' Responsibility

Reviewers must consider the followings:

- Qualitative, contextual and scientific study in order to improve articles' quality and content.
- To inform editor-in-chief when accepts or reject the review and introduce an alternative.
- Should not accept the articles which consider the benefits of persons, organizations and companies or personal relationships; also the articles which she/he, own, contributed in its writing or analyze.
- The reviewing must be carried out upon scientific documents and any self, professional, religious and racial opinion is prohibited.

- Accurate review and declaration of the article's strengths and weaknesses through a clear, educational and constructive method.
- Responsibility, accountability, punctuality, interest, ethics adherence and respect to others' right.
- Not to rewrite or correct the article according to his/her personal interest.
- Be sure of accurate citations. Also reminding the cases which haven't been cited in the related published researches.
- Avoid of express the information and details of articles.
- Reviewers should not benefit new data or contents in favor of/against personal researches; even for criticism or discrediting the author(s). The reviewer is not permitted to reveal more details after a reviewed article being published.
- Reviewer is prohibited to deliver an article to another one for reviewing except with permission of editor-in-chief. Reviewer and co-reviewer's identification should be noted in each article's documents.
- Reviewer shouldn't contact with the author(s). Any contact with the authors should be made through the editorial office.
- Trying to report "research and publication misconduct" and submitting the related documents to editor-in-chief.

5 Editorial Board Responsibilities

- Journal maintenance and quality improvement are the main aims of editorial board.
- Editorial board should introduce the journal to universities and international communities and publish the articles of other universities and international societies on their priority.
- Editorial board must not have quota and excess of their personal article publishing.
- Editorial board is responsible for selecting the reviewers as well as accepting or rejecting on article after reviewers' comments.
- Editorial board should be well-known experts with several publications. They ought to be responsible, accountable, truth, adhere to professional ethics and contribute to improve journal aims.
- Editorial board is expected to have a database of suitable reviewers for journal and to update the information regularly.
- Editorial board should try to aggregate qualified moral, experienced and well-known reviewers

- Editorial board should welcome deep and reasonable reviews, and prevent superficial and poor reviews, and deal with one-sided and contemptuous reviews.
- Editorial board should record and archive the whole review's documents as scientific documents and to keep confidentially the reviewers' name.
- Editorial board must inform the final result of review to corresponding author immediately.
- Editorial board should keep the article's contents confidentially and do not disclose its information to others.
- Editorial board ought to prevent any conflict of interests due to any personal, commercial, academic and financial relations which may impact on accepting and publishing the presented articles.
- Editor-in-chief should check each type of research and publication misconduct which reviewers report seriously.
- If a research and publication misconduct occurs in an article, editor-in-chief should omit it immediately and inform indexing databases or audiences.
- In the case of being a research and publication misconduct, editorial board is responsible to represent a corrigendum to audiences rapidly.
- Editorial board must benefit of audiences' new ideas in order to improve publication policies, structure and content quality of articles.

References

[1] "Standard Ethics", approved by Vice-Presidency for Research & Technology, the Ministry of Science, Research and Technology.

[2] Committee on Publication Ethics, COPE Code of Conduct, www.publicationethics.org.

Control and Optimization in Applied Mathematics (COAM)

This journal is published under the auspices of the Payame Noor University (PNU).

*Full Professor
**Associated Professor

Control and Optimization in Applied Mathematics (COAM)

Editor in Chief's Letter

It would be our great honor to have you as the readers of Journal of "Control and Optimization in Applied Mathematics (COAM)". The present journal is published and supported by Payame Noor University (PNU) as a semi-annual journal. Our main objective is to facilitate scientific regional and global discussions and collaborations between specialists in different fields of applied mathematics, especially in the fields of control and optimization. We hope that scholars and experts of different fields of applied mathematics find our scientific journal a platform for international communications of insight and knowledge. To assure the respectful subscribers about high quality of the journal, each article is reviewed by subject-qualified referees, the same as any other well-known international journal of applied mathematics.We believe that by publishing high quality and creative researches, we will observe more collaborations with our journal. We kindly invite all applied mathematicians especially in the fields of control and optimization, to join us by submitting their original works to the Journal of "Control and Optimization in Applied Mathematics". I want to thank the respectful colleagues of COAM, as well as referees, reviewers, and editors for their kind dedication and vision.

Contents

Control and Optimization in Applied Mathematics (COAM)
DOI. 10.30473/coam.2020.54034.1143
Vol. 4, No. 2, Autumn - Winter 2019 (1-18), ©2016 Payame Noor University, Iran

Robust Switching Law Design for Uncertain Time-Delay Switched Linear Systems

M. R. Ramezani-al[1]*, S. Sabeti[2]

[1]Department of Electrical and Computer Engineering, Quchan University of Technology, P.O. Box. 19395-3697, Quchan, Iran,

[2]Department of Mathematics, Statistics and Computer Science, Semnan University, P.O. Box. 1355-16315, Semnan, Iran.

Received: July 12, 2020; **Accepted:** January 08, 2021.

Abstract. Guaranteed cost control (GCC) is an impressive method of controlling nonlinear systems, incredibly uncertain switched systems. Most of the recent studies of GCC on uncertain switched linear systems have been concerned with asymptotic stability analysis. In this paper, a new robust switching law for time-delay uncertain switched linear systems is designed. First, the switching law is designed, and second, a state-feedback controller based on Lyapunov-Krasovskii Functional (LKF) is designed. Also, using Linear Matrix Inequality (LMI) particular condition for the existence of a solution of obtained switching law and controller is achieved. Consequently, in the presented theorems, the exponential stability of the overall system under switching law and controller is analyzed. Finally, theoretical results are verified via presenting an example.

Keywords. Uncertain switched linear systems, Time-delay, Guaranteed cost control, LKF, LMI, Exponential stability.

MSC. 47N70.

* Corresponding author
m-ramezani@qiet.ac.ir, sabeti.samira@semnan.ac.ir
http://mathco.journals.pnu.ac.ir

1 Introduction

Switched systems, as a wide class of hybrid systems, are divided into two classes of systems: continuous and discrete-time subsystems. Generally, there is a switching strategy that selects a subsystem between other subsystems. In recent years, switching theory and its application have been extended to adaptive control to overcome disadvantages in the system's stability, where there are some difficulties in the proof of stability [1, 2, 3, 4, 5]. Some important problems in the concept of design procedures and stability analysis of switched systems have illustrated in [6].

There are many approaches in switched systems especially, looking for suitable switching; to stabilize the system even when the systems are unstable [7]. Also, dwell-time and its average concept have been studied for stabilization problems in the switched system with especial switching strategy has been performed [8, 6]. In recent past decades, time-delay systems have been concerned with expert researchers. These kinds of systems have many applications in electronics systems, transmission systems, chemical process systems, and power systems and, so on [9]. Delay mainly exists in some sensors and measurement units and frequently occurs in control systems [10]. Generally, since sensors and transducers are used in control systems to measure all or some important states, then, some delays may occur in these measurements. Switched systems with a time delay are a class of switched systems that has been focused on recent researches. In most studies on the time-delay switched systems, delay with a certain upper bound is assumed. Knowing such this upper bound can guarantee the stability of these kinds of systems. In this area, some rigorous researches have been achieved in recent years [11, 12, 13]. In [12], using Common Lyapunov Function (CLF), the stability of switching systems composed some finite linear subsystems which are described with time-delay differential equations has been performed. In [13], the Authors studied sufficient conditions for asymptotic stability analysis of a class of switched linear systems. Moreover, many types of research in the field of switched systems concentrate on the asymptotic behavior that reflects the system treatments in a limited interval time [14, 15]. In the concept of control a plant, designing a controller must guarantee not only the asymptotic stability of the system but also guarantee acceptable performance. Considering a quadratic performance index is a solution to formulate this problem. This method is named guaranteed cost control (GCC) [16]. In this approach, it is tried to provide an upper bound for a given cost function in the presence of uncertainties, and, based on this goal, the controller is designed [17, 18, 19, 20, 22]. Based on this approach, some significant researches have been reported on this topic in [20, 21, 23, 24, 25]. Some acceptable results have been reported for uncertain switched linear systems. In these studies, using CLF or Multiple Lyapunov Functions (MLFs), switching laws and state feedback controllers are designed. Moreover, for switching strategy design, a subsystem with minimum LF is chosen. When the switched system has only a switching signal to be designed, this approach provides asymptotic or exponential stability. Especially, to design switching laws using CLF, the designer must find some unknown matrices with solving some complex Linear Matrix Inequalities (LMIs) to be constructed via some theorems [26, 21, 23, 24, 25]. In more recent studies, some important researches on the exponential stability analysis and design of GCC for time delay switched systems has been performed [27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 41, 42, 43, 44]. In [31], using some extracted LMIs for switched time-delay systems, a sufficient condition for exponential stability analysis and GCC problem with the weighted form is obtained. Also, in [38] and based on dwell time and piecewise Lyapunov function approach exponential stability is studied, and its condition is derived. Besides, in [38], and based on the LKF method, to guarantee exponential stability and obtain the upper bound of the determined cost function, a new time delay condition is proposed. In this paper, by considering a complete form of uncertain time-delay switched systems containing delays both in states and control inputs, a new robust switching law is designed. To do this, motivated by the min-projection switching strategy [39] and Lyapunov-Krasovskii function (LKF), switching law and control are designed. The main contributions are listed in the following:

(i) Designing a new robust switching law to guarantee exponential stability of the switched system.

(ii) Proving that the proposed LKF satisfies the presented theorems.

Notation: Throughout the paper, m is an arbitrary positive integer that indicates the number of switched system's subsystems, and $\lambda(A)$ indicates eigenvalues of matrix A. the notation $P > 0$ denotes that P is a positive definite matrix.

2 Problem Formulation and Preparations

In this paper, the following general form of time-delay uncertain switched linear system is considered

$$\begin{aligned} \dot{x} &= (A_{\sigma(x,t)} + \Delta A_{\sigma(x,t)})x(t) + A_{d\sigma(x,t)}x(t-d) \\ &\quad + (B_{\sigma(x,t)} + \Delta B_{\sigma(x,t)})u(t) + W_{\sigma(x,t)}u(t-h), \\ x(t) &= \phi(t),\ t \in [-t_0,\ 0],\ t_0 \triangleq \max\{d,\ h\}, \end{aligned} \tag{1}$$

where, $x(t) \in R^n$ and $u(t) \in R^q$ are the state and control input vectors. $d > 0$ and $h > 0$ are delay constants in the states and inputs and $\sigma(x,t) \in \underline{m}$ is switching signal which is piecewise constant that determines the active subsystem. $A_i \in R^{n\times n}$, $B_i \in R^{n\times q}$, $A_{di} \in R^{n\times n}$ and $W_i \in R^{n\times n}$, $i \in \underline{m}$ are subsystem matrices and ΔA_i and ΔB_i, $i \in \underline{m}$, are additive uncertainties. The following notice shows the nature of uncertainties.

Notice 1. ΔA_i and ΔB_i in equation (1) are time-varying uncertain matrices and satisfy the following condition

$$[\Delta A_i \quad \Delta B_i] = N_i F_i [C_i \quad D_i], \qquad i \in \underline{m}, \tag{2}$$

where C_i, D_i and N_i are known matrices and F_i, $i \in \underline{m}$, are unknown matrices with Lebesgue measurable elements such that the following inequality holds

$$F_i^T(t) F_i(t) \le I, \qquad i \in \underline{m}. \tag{3}$$

Throughout the paper, our goal is to minimize the following performance index for the uncertain system (1)

$$J = \int_0^\infty (x^T Q x + u^T R u) dt, \tag{4}$$

where $Q \in R^{n\times n}$ and $R \in R^{q\times q}$ are symmetric positive definite matrices. The main goal of the paper is to find switching law $\sigma(x,t)$ and state-feedback controller $u = K_i x(t)$, where $K_i \in R^{q\times n}$, $i \in \underline{m}$ such that, the the system (1) to be exponential stable and the cost function (4) satisfies $J \le J^*$ where J^* is a guaranteed cost value, which is defined in Definition 1. Before presenting our main results, we introduce some necessary definitions, lemmas, and theorems.

Definition 1. [20] For all uncertainties satisfying (2) and (3), state-feedback control $u^*(t)$ and switching law $\sigma^*(x,t)$ are said to be guaranteed cost value (GCV) and guaranteed cost control law (GCCL), if the closed-loop system (1) to be asymptotic (or exponential) stable and the value of cost function (4) satisfies $J \le J^*$, where J^* is a positive scalar.

Definition 2. [24, 33] The system (1) under switching law $\sigma(x,t)$ and control $u = K_i x(t)$ is said to be exponential stable if the norm of state vector $x(t)$ satisfies (5)

$$\|x(t)\| \le k_1 e^{-k_2 t} \|x(0)\|, \tag{5}$$

where $k_1 > 0$ and $k_2 > 0$, and $\|x(0)\|$ is initial value at time $t = 0$.

Lemma 1. [26] For matrices L, P and $Q > 0$, the following inequality holds

$$\begin{bmatrix} P & L \\ L^T & -Q \end{bmatrix} < 0 \Longleftrightarrow P + LQ^{-1}L^T < 0. \tag{6}$$

Lemma 2. [33] Consider D, E, and F be real matrices, and matrix F satisfies $F^TF \leq I$. For any positive scalar ε, the following inequality holds

$$DFE + E^TF^TD^T \leq \varepsilon^{-1}DD^T + \varepsilon E^TE \tag{7}$$

Lemma 3. [25] For any symmetric matrix Y, arbitrary matrices M and N and for all F satisfying $F^TF \leq I, i \in \underline{m}$, the following inequality holds

$$Y + MFN + N^TF^TM^T < 0.$$

if and only if there exists positive scalar ε such that

$$Y + \varepsilon N^TN + \varepsilon^{-1}M^TM < 0,$$

Lemma 4. [40] For any real symmetric matrix $A \in R^{n\times n}$

$$\lambda_{\min}(A)\|x\|^2 \leq x^TAx \leq \lambda_{\max}(A)\|x\|^2, \tag{8}$$

where $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ are the smallest and largest eigenvalues of matrix A.

Theorem 1. For the system (1), if there exist matrices $P > 0$, $P_1 > 0$ and $P_2 > 0$, positive scalar α and positive definite scalar function $V(x(t))$ as a Lyapunov function for system (1) such that

$$\dot{V}(x(t)) \leq -\alpha\|x\|^2, \tag{9}$$

then, the switching law (10) can stabilize the switched system (1) exponentially.

$$\sigma(x,t) = arg \min_{i\in\underline{m}}\{x^TPf_i(x)\}. \tag{10}$$

Proof. In ([39]) using the min-projection switching strategy this theorem has been proved for nonlinear switched systems in the form of $\dot{x} = f_i(x)$, $i \in \underline{m}$ .To extend this theorem in switched systems (1), the following Lyapunov-Krasovskii function is proposed

$$V(x(t)) = x^T(t)Px(t) + \int_{-d}^{0} x^T(t+\tau)P_1x(t+\tau)d\tau$$
$$+ \int_{-h}^{0} x^T(t+\tau)P_2x(t+\tau)d\tau,$$

and it is proved to reach exponential stability, there exist positive scalars $k_1 = \sqrt{\frac{\lambda_{max}(P)}{\lambda_{min}(P)}}$ and $k_2 = \frac{\alpha}{2\lambda_{max}(P)}$ satisfy the exponential definition (5). □

3 Main Results

Theorem 2. System (1) under the following switching law is to be exponentially stable

$$\sigma(x,t) = arg\min_{i\in\underline{m}}\{\bar{x}^T Z_i \bar{x}\}, \tag{11}$$

where

$$Z_i = \begin{bmatrix} \theta_i & S_1 & PW_iK_i \\ S_1^T & -P_1 & 0 \\ K_i^T W_i^T P & 0 & -P_2 \end{bmatrix},$$
$$\bar{x} = \begin{bmatrix} x(t), x(t-d), x(t-h) \end{bmatrix}', \tag{12}$$

and

$$\chi_i = A_i + \Delta A_i + B_i K_i + \Delta B_i K_i,$$
$$\theta_i = \chi_i^T P + P\chi_i + P_1 + P_2 + Q + K_i^T R K_i,$$
$$S_1 = PA_{di},$$

if there exist symmetric positive-definite matrices P, P_1 and P_2, and matrices K_i, $i \in \underline{m}$, such that the following inequality holds:

$$\sum_{i=1}^{m}\Big[x^T(t)\theta_i x(t) + x^T(t)S_1 x(t-d) + x^T(t-d)S_1^T x(t)$$
$$+ x^T(t-h)K_i^T W_i^T P x(t) + x^T(t)PW_iK_i x(t-h)$$
$$- x^T(t-d)P_1 x(t-d) - x^T(t-h)P_2 x(t-h)\Big] < 0, \tag{13}$$

In addition, GSV is $J^* = \phi(0)^T P\phi(0) + \int_{-d}^{0}\phi^T(\tau)P_1\phi(\tau)d\tau + \int_{-h}^{0}\phi^T(\tau)P_2\phi(\tau)d\tau$.

Proof. Clearly from switching (11) and inequality (13), it is resulted that $\sum_{i=1}^{m} Z_i < 0$ and consequently, there exists an index $i \in \underline{m}$ such that $\bar{x}^T Z_i \bar{x} < 0$ for an augmented state vector $\bar{x} \in R^{3n}$, $\bar{x} \neq 0$. Now the following function is preposed as a Lyapunov-Krasovskii function, where P, P_1 and P_2 are symmetric positive definite matrices

$$V(x(t)) = x^T(t)Px(t) + \int_{-d}^{0} x^T(t+\tau)P_1 x(t+\tau)d\tau \tag{14}$$
$$+ \int_{-h}^{0} x^T(t+\tau)P_2 x(t+\tau)d\tau,$$

Time derivation of $V(x(t)$ and substituting $u(t) = K_i x(t)$ into system equations (1) and using Notice 1, results

$$\begin{aligned}\dot{V}(x(t)) &= \dot{x}^T(t)Px(t) + x^T(t)P\dot{x}(t) + x^T(t)P_1x(t) - x^T(t-d)P_1x(t-d)\\ &+ x^T(t)P_2x(t) - x^T(t-h)P_2x(t-h) = x^T(t)(A_i+\Delta A_i)^TPx(t)\\ &+ x^T(t)P(A_i+\Delta A_i)x(t) + x^T(t-d)A_{di}^TPx(t) + x^T(t)PA_{di}x(t-d)\\ &+ x^T(t)K_i^T(B_i+\Delta B_i)^TPx(t) + x^T(t)P(B_i+\Delta B_i)K_ix(t)\\ &+ x^T(t-h)K_i^TW_i^TPx(t) + x^T(t)PW_iK_ix(t-h) + x^T(t)P_1x(t)\\ &- x^T(t-d)P_1x(t-d) + x^T(t)P_2x(t) - x^T(t-h)P_2x(t-h)\\ &= x^T(t)\Big[P(A_i+B_iK_i) + (A_i+B_iK_i)^TP + PN_iF_i(C_i+D_iK_i)\\ &+ (C_i+D_iK_i)^TF_i^TN_i^TP + P_1 + P_2\Big]x(t) + x^T(t-d)A_{di}^TPx(t)\\ &+ x^T(t)PA_{di}x(t-d) + x^T(t-h)K_i^TW_i^TPx(t) + x^T(t)PW_iK_ix(t-h)\\ &- x^T(t-d)P_1x(t-d) - x^T(t-h)P_2x(t-h) \end{aligned} \tag{15}$$

Applying Lemma 2, we have

$$\begin{aligned} &PN_iF_i(C_i+D_iK_i) + (C_i+D_iK_i)^TF_i^TN_i^TP\\ &\quad \le \varepsilon_iPN_iN_i^TP + \varepsilon_i^{-1}(C_i+D_iK_i)(C_i+D_iK_i)^T. \end{aligned} \tag{16}$$

Rewritten equation (15) results

$$\begin{aligned}\dot{V}(x(t)) \le\ & x^T(t)\Big[P(A_i+B_iK_i) + (A_i+B_iK_i)^TP + \varepsilon_i^{-1}(C_i+D_iK_i)(C_i+D_iK_i)^T\\ &+ \varepsilon_iPN_iN_i^TP + P_1 + P_2\Big]x(t) + x^T(t-d)A_{di}^TPx(t) + x^T(t)PA_{di}x(t-d)\\ &+ x^T(t-h)K_i^TW_i^TPx(t) + x^T(t)PW_iK_ix(t-h) - x^T(t-d)P_1x(t-d)\\ &- x^T(t-h)P_2x(t-h). \end{aligned} \tag{17}$$

By defining

$$\begin{aligned} \theta_i &= P(A_i+B_iK_i) + (A_i+B_iK_i)^TP + \varepsilon_i^{-1}(C_i+D_iK_i)(C_i+D_iK_i)^T\\ &\quad + \varepsilon_iPN_iN_i^TP + P_1 + P_2 + Q + K_i^TRK_i\\ S_1 &= PA_{di}, \end{aligned}$$

and adding $x^T(t)(Q+K_i^TRK_i)x(t)$ to (17), results

$$\begin{aligned} \dot{V}(x(t)) + x^T(t)(Q+K_i^TRK_i)x(t) &\le x^T(t)\theta_ix(t) + x^T(t)S_1x(t-d)\\ &+ x^T(t-d)S_1^Tx(t) - x^T(t-d)P_1x(t-d) + x^T(t-h)K_i^TW_i^TPx(t)\\ &+ x^T(t)PW_iK_ix(t-h) - x^T(t-h)P_2x(t-h). \end{aligned} \tag{18}$$

Consequently inequality (18) can be written as

$$\begin{aligned} &\dot{V}(x(t)) + x^T(t)(Q+K_i^TRK_i)x(t)\\ &\le \begin{bmatrix} x(t)\\ x(t-d)\\ x(t-h)\end{bmatrix}^T \begin{bmatrix} \theta_i & S_1 & PW_iK_i\\ S_1^T & -P_1 & 0\\ K_i^TW_i^TP & 0 & -P_2\end{bmatrix} \begin{bmatrix} x(t)\\ x(t-d)\\ x(t-h)\end{bmatrix} \end{aligned}$$

$$= \bar{x}^T(t)\begin{bmatrix} \theta_i & S_1 & PW_iK_i \\ S_1^T & -P_1 & 0 \\ K_i^T W_i^T P & 0 & -P_2 \end{bmatrix}\bar{x}(t). \tag{19}$$

Now, it is concluded that there exist an $i \in \underline{m}$ such that $\bar{x}^T Z_i \bar{x} < 0$. Therefore, selecting switching law (11) for any time $t \in R$, results that $\bar{x}^T Z_i \bar{x} < 0$ and

$$\dot{V}(x(t)) + x^T(t)(Q + K_i^T RK_i)x(t) \le \bar{x}^T(t)Z_i\bar{x}(t) < 0. \tag{20}$$

So,

$$\begin{aligned} \dot{V}(x(t)) &< -x^T(t)Qx(t) - x^T(t)(K_i^T RK_i)x(t) \\ &= -x^T(t)(Q + K_i^T RK_i)x(t), \end{aligned} \tag{21}$$

Obviously, it is concluded that $G_i = Q+K_i^T RK_i$ is positive-definite matrix for any $i \in \underline{m}$. Therefore, using Lemma 4, $\forall x \in R^n, i \in \underline{m}$ the following inequality holds:

$$-\lambda_{\max}(G_i)\|x\|^2 \le -x^T G_i x \le -\lambda_{\min}(G_i)\|x\|^2. \tag{22}$$

Now, by choosing

$$\gamma = \lambda_{\min}(G) = \min_{i\in\underline{m}}(\lambda_{\min}(G_i)), \tag{23}$$

Then, applying Theorem 1 results that switched system (1) is exponentially stable. □

Remark 1. We need to find unknown matrices P, P_1 ,P_2 and control gains K_i to realize the switching law (11). Also, positive scalars ε_i are designing constants and can be selected by the designer arbitrarily or by some optimization methods. In the Theorem 2 and using Lemma 1 it is shown that (13) is equal to a set of LMIs (24).

Theorem 3. If there exist invertible symmetric positive definite matrix X, P_2 and matrices M_i and V_i for some positive scalars ε_i, $i \in \underline{m}$, such that the following LMI to be satisfied

$$\begin{bmatrix} \Psi_i & A_{di} & W_iV_i & X^T & M_i^T & X^T & M_i^T & \bar{C}_i^T \\ A_{di}^T & -I & 0 & 0 & 0 & 0 & 0 & 0 \\ V_i^T W_i & 0 & -I & 0 & 0 & 0 & 0 & 0 \\ X & 0 & 0 & -I & 0 & 0 & 0 & 0 \\ M_i & 0 & 0 & 0 & -P_2^{-1} & 0 & 0 & 0 \\ X & 0 & 0 & 0 & 0 & -Q^{-1} & 0 & 0 \\ M_i & 0 & 0 & 0 & 0 & 0 & -R^{-1} & 0 \\ \bar{C}_i & 0 & 0 & 0 & 0 & 0 & 0 & -\varepsilon_i^{-1} I \end{bmatrix} < 0, \tag{24}$$

where

$$\Psi_i = (A_iX + B_iM_i)^T + A_iX + B_iM_i + \varepsilon_i^{-1} N_i N_i^T,$$
$$\bar{C}_i = C_iX + D_iM_i,$$

and then, inequality (13) holds and switching strategy (11) for the system (1) can be implemented.

Proof. Define the following matrix

$$Y = \begin{bmatrix} \bar{\Omega}_i & P^T A_{di} & PW_iV_i & I & K_i^T & I & K_i^T \\ A_{di}^T P & -P_1 & 0 & 0 & 0 & 0 & 0 \\ V_i^T W_i^T P & 0 & -I & 0 & 0 & 0 & 0 \\ I & 0 & 0 & -P_1^{-1} & 0 & 0 & 0 \\ K_i & 0 & 0 & 0 & -P_2^{-1} & 0 & 0 \\ I & 0 & 0 & 0 & 0 & -Q^{-1} & 0 \\ K_i & 0 & 0 & 0 & 0 & 0 & -R^{-1} \end{bmatrix}, \tag{25}$$

where

$$\bar{\Omega}_i = (A_i + B_iK_i)^T P + P^T (A_i + B_iK_i).$$

Using Lemma 2, the matrix inequality (13) is equal to the following

$$Y + \begin{bmatrix} \Phi_i & 0_{1\times 6} \end{bmatrix}^T F_i^T(t) \begin{bmatrix} N_i^T P & 0_{1\times 6} \end{bmatrix} + \begin{bmatrix} N_i^T P & 0_{1\times 6} \end{bmatrix}^T F_i(t) \begin{bmatrix} \Phi_i & 0_{1\times 6} \end{bmatrix} < 0 \tag{26}$$

where

$$\Phi_i = C_i + D_iK_i.$$

By rewriting inequality (26), we have

$$\begin{bmatrix} \psi_{1i} & P^T A_{di} & PW_iV_i & I & K_i^T & I & K_i^T \\ A_{di}^T P & -P_1 & 0 & 0 & 0 & 0 & 0 \\ V_i^T W_i^T P & 0 & -I & 0 & 0 & 0 & 0 \\ I & 0 & 0 & -P_1^{-1} & 0 & 0 & 0 \\ K_i & 0 & 0 & 0 & -P_2^{-1} & 0 & 0 \\ I & 0 & 0 & 0 & 0 & -Q^{-1} & 0 \\ K_i & 0 & 0 & 0 & 0 & 0 & -R^{-1} \end{bmatrix}$$
$$+ \begin{bmatrix} C_i^T + K_i^T D_i^T \\ 0_{6\times 1} \end{bmatrix} \begin{bmatrix} F_i^T(t) N_i^T P & 0_{1\times 6} \end{bmatrix}$$
$$+ \begin{bmatrix} P^T N_i \\ 0_{6\times 1} \end{bmatrix} \begin{bmatrix} F_i(t)C_i + F_i(t)D_iK_i & 0_{1\times 6} \end{bmatrix}$$

$$= \begin{bmatrix} \psi_{2i} & P^T A_{di} & PW_i V_i & I & K_i^T & I & K_i^T \\ A_{di}^T P & -P_1 & 0 & 0 & 0 & 0 & 0 \\ V_i^T W_i^T P & 0 & -I & 0 & 0 & 0 & 0 \\ I & 0 & 0 & -P_1^{-1} & 0 & 0 & 0 \\ K_i & 0 & 0 & 0 & -P_2^{-1} & 0 & 0 \\ I & 0 & 0 & 0 & 0 & -Q^{-1} & 0 \\ K_i & 0 & 0 & 0 & 0 & 0 & -R^{-1} \end{bmatrix} < 0, \quad (27)$$

where

$$\psi_{1i} = (A_i + B_i K_i)^T P + P^T (A_i + B_i K_i)$$
$$\psi_{2i} = (A_i + B_i K_i)^T P + P^T (A_i + B_i K_i) + C_i^T F_i^T(t) N_i^T P$$
$$+ K_i^T D_i^T F_i^T(t) N_i^T P + P^T N_i F_i(t) C_i + P^T N_i F_i(t) D_i K_i,$$

By simple calculations in

$$Y + \varepsilon_i \begin{bmatrix} \Phi_i & 0_{1\times 6} \end{bmatrix}^T \begin{bmatrix} \Phi_i & 0_{1\times 6} \end{bmatrix}$$
$$+ \varepsilon_i^{-1} \begin{bmatrix} N_i^T P & 0_{1\times 6} \end{bmatrix}^T \begin{bmatrix} N_i^T P & 0_{1\times 6} \end{bmatrix}$$
$$= \begin{bmatrix} \psi_{1i} & P^T A_{di} & PW_i V_i & I & K_i^T & I & K_i^T \\ A_{di}^T P & -P_1 & 0 & 0 & 0 & 0 & 0 \\ V_i^T W_i^T P & 0 & -I & 0 & 0 & 0 & 0 \\ I & 0 & 0 & -P_1^{-1} & 0 & 0 & 0 \\ K_i & 0 & 0 & 0 & -P_2^{-1} & 0 & 0 \\ I & 0 & 0 & 0 & 0 & -Q^{-1} & 0 \\ K_i & 0 & 0 & 0 & 0 & 0 & -R^{-1} \end{bmatrix}$$
$$+ \begin{bmatrix} \varepsilon_i (C_i + D_i K_i)^T (C_i + D_i K_i) & 0_{1\times 6} \\ 0_{6\times 1} & 0_{6\times 6} \end{bmatrix}$$
$$+ \begin{bmatrix} \varepsilon_i^{-1} P^T N_i N_i^T P & 0_{1\times 6} \\ 0_{6\times 1} & 0_{6\times 6} \end{bmatrix}$$
$$= \begin{bmatrix} \psi_{3i} & P^T A_{di} & PW_i V_i & I & K_i^T & I & K_i^T \\ A_{di}^T P & -P_1 & 0 & 0 & 0 & 0 & 0 \\ V_i^T W_i^T P & 0 & -I & 0 & 0 & 0 & 0 \\ I & 0 & 0 & -P_1^{-1} & 0 & 0 & 0 \\ K_i & 0 & 0 & 0 & -P_2^{-1} & 0 & 0 \\ I & 0 & 0 & 0 & 0 & -Q^{-1} & 0 \\ K_i & 0 & 0 & 0 & 0 & 0 & -R^{-1} \end{bmatrix} < 0, \quad (28)$$

where

$$\begin{aligned}\psi_{3i} &= (A_i + B_iK_i)^T P + P^T(A_i + B_iK_i)\\ &+ \varepsilon_i(C_i + D_iK_i)^T(C_i + D_iK_i) + \varepsilon_i^{-1}P^T N_i N_i^T P.\end{aligned}$$

Now, from the Lemma 1, inequality (28) is equal to (29)

$$\begin{bmatrix} \Omega_i & P^TA_{di} & P^TW_iV_i & I & K_i^T & I & K_i^T & \Phi^T\\ A_{di}^TP & -P_1 & 0 & 0 & 0 & 0 & 0 & 0\\ V_i^TW_i^TP & 0 & -I & 0 & 0 & 0 & 0 & 0\\ I & 0 & 0 & -P_1^{-1} & 0 & 0 & 0 & 0\\ K_i & 0 & 0 & 0 & -P_2^{-1} & 0 & 0 & 0\\ I & 0 & 0 & 0 & 0 & -Q^{-1} & 0 & 0\\ K_i & 0 & 0 & 0 & 0 & 0 & -R^{-1} & 0\\ \Phi_i & 0 & 0 & 0 & 0 & 0 & 0 & -\varepsilon_i^{-1}I \end{bmatrix}, \quad (29)$$

where

$$\Omega_i = \bar{\Omega}_i + \varepsilon_i^{-1}PN_iN_i^TP,$$

Multiplying both sides of (29) by $diag\{P^{-T}, P_1^{-1}, I, I, I, I, I, I\}$ and $diag\{P^{-1}, P_1^{-1}, I, I, I, I, I, I\}$ yields

$$\begin{aligned} &= \begin{bmatrix} P^{-T} & 0 & 0 & \dots & 0\\ 0 & P_1^{-1} & 0 & \dots & 0\\ 0 & 0 & I & \dots & 0\\ \vdots & \vdots & \vdots & \ddots & 0\\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}\\ &\times \begin{bmatrix} \varphi_i & P^TA_{di} & \omega_i & I & K_i^T & I & K_i^T & \Phi_i^T\\ A_{di}^TP & -P_1 & 0 & 0 & 0 & 0 & 0 & 0\\ \omega_i & 0 & -I & 0 & 0 & 0 & 0 & 0\\ I & 0 & 0 & -P_1^{-1} & 0 & 0 & 0 & 0\\ K_i & 0 & 0 & 0 & -P_2^{-1} & 0 & 0 & 0\\ I & 0 & 0 & 0 & 0 & -Q^{-1} & 0 & 0\\ K_i & 0 & 0 & 0 & 0 & 0 & -R^{-1} & 0\\ \Phi_i & 0 & 0 & 0 & 0 & 0 & 0 & -\varepsilon_i^{-1}I \end{bmatrix}\\ &\times \begin{bmatrix} P^{-1} & 0 & 0 & \dots & 0\\ 0 & P_1^{-1} & 0 & \dots & 0\\ 0 & 0 & I & \dots & 0\\ \vdots & \vdots & \vdots & \ddots & 0\\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}, \end{aligned}$$

where

$$\varphi_i = (A_i + B_iK_i)^T P + P^T(A_i + B_iK_i) + \varepsilon_i^{-1}P^TN_iN_i^TP,$$

$$\omega_i = P^T W_i V_i.$$

So we have

$$\begin{bmatrix} \begin{matrix} P^{-T}(A_i+B_iK_i)^T+(A_i+B_iK_i)P^{-1} \\ +\varepsilon_i^{-1}N_iN_i^T \end{matrix} & A_{di}T^{-1} & W_iV_i & I & P^{-T}K_i^T & P^{-T} & P^{-T}K_i^T & P^{-T}(C_i^T+K_i^TD_i^T) \\ T^{-1}A_{di} & -P_1^{-1} & 0 & 0 & 0 & 0 & 0 & 0 \\ V_i^TW_i^T & 0 & -I & 0 & 0 & 0 & 0 & 0 \\ P^{-1} & 0 & 0 & -P_1^{-1} & 0 & 0 & 0 & 0 \\ K_iP^{-1} & 0 & 0 & 0 & -P_2^{-1} & 0 & 0 & 0 \\ P^{-1} & 0 & 0 & 0 & 0 & -Q^{-1} & 0 & 0 \\ K_iP^{-1} & 0 & 0 & 0 & 0 & 0 & -R^{-1} & 0 \\ (C_i+D_iK_i)P^{-1} & 0 & 0 & 0 & 0 & 0 & 0 & -\varepsilon_i^{-1}I \end{bmatrix} < 0. \quad (30)$$

□

In summary, to obtain $\sigma(x,t)$, $u(t)$ and J^*, the following steps are required to perform.

Step 1: Select positive scalars ε_i, $i \in \underline{m}$.

Step 2: Solve LMIs (24) in Theorem 3 (Via LMI commands in the Matlab software or YALMIP toolbox) and obtain invertible symmetric positive-definite matrices X, P_2 and matrices M_i, $i \in \underline{m}$. Note that $X = P^{-1}$, $M_i = K_iX$ and consequently. $P = X^{-1}$, $K_i = M_iX^{-1}$. Positive definite matrix P_1 can be given from inequality $\bar{x}^T Z_i \bar{x} < 0$ in Theorem 2.

Step 3: Obtain State feedback $u(t) = K_i x(t)$.

Step 4: Calculate Z_i, $i \in \underline{m}$ in Theorem 2.

Step 5: Obtain switching law $\sigma(x,t) = arg\min_{i\in\underline{m}}\{\bar{x}^T Z_i \bar{x}\}$.

Step 6: Calculate guaranteed cost control J^*.

4 Illustrative Example

Example 1. Consider the following uncertain time-delay switched linear system with two subsystems.

$$\begin{aligned} \dot{x} &= (A_{\sigma(x,t)} + \Delta A_{\sigma(x,t)})x(t) + A_{d\sigma(x,t)}x(t-d), \\ &\quad + (B_{\sigma(x,t)} + \Delta B_{\sigma(x,t)})u(t) + W_{\sigma(x,t)}u(t-h), \\ x(t) &= \phi(t),\ t \in [-t_0,\ 0],\ t_0 \triangleq \max\{d,\ h\}, \end{aligned} \quad (31)$$

for $i = 1, 2$, and the following matrices

$$A_1 = \begin{bmatrix} -2 & -1 \\ 3 & -4 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} -4 & 1 \\ 0.5 & -1 \end{bmatrix},$$

$$B_1 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \qquad B2 = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix},$$
$$A_{d1} = \begin{bmatrix} 0 & 0.2 \\ 0.2 & 0 \end{bmatrix}, \qquad A_{d2} = \begin{bmatrix} 0.2 & 0 \\ 0 & 0.2 \end{bmatrix},$$
$$W_1 = \begin{bmatrix} 0.3 & 0.2 \\ 0.2 & 0.4 \end{bmatrix}, \qquad W_2 = \begin{bmatrix} 0.4 & 0.3 \\ 0.3 & 0.5 \end{bmatrix},$$
$$N_1 = \begin{bmatrix} 0.4 & 0.3 \\ 0.3 & 0.6 \end{bmatrix}, \qquad N_2 = \begin{bmatrix} 0.2 & 0.7 \\ 0.6 & 0.3 \end{bmatrix},$$
$$C_1 = \begin{bmatrix} 0.6 & 0.2 \\ 0.3 & 0.6 \end{bmatrix}, \qquad C_2 = \begin{bmatrix} 0.4 & 0.2 \\ 0.6 & 0.4 \end{bmatrix},$$
$$D_1 = \begin{bmatrix} 0 & 0.4 \\ 0.4 & 0 \end{bmatrix}, \qquad D_2 = \begin{bmatrix} 0 & 0.2 \\ 0.2 & 0 \end{bmatrix},$$

$d = 2$ and $h = 1$ and

$$x(t) = [e^t - e^t]^T, \qquad t \in [-2\ 0],$$

Also, weighted matrices Q and R are selected as

$$Q = R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{32}$$

Note that all subsystems of system (31) are stable and unknown matrices $F_i(t)$ in Notice 1 are considered as a diagonal random time-varying matrices such that $F_i^T(t)F_i(t) \leq I$. The aim is to find guaranteed cost controller $u{=}K_i x(\ t), i \in \{1,2\}$, switching signal $\sigma(x,t)$ and guaranteed cost J^* of the switched system (31) with weighted matrices (32). We perform the following steps.

step 1: Scalars ε_1 and ε_2 are selected as

$$\varepsilon_1 = 0.1,\ \varepsilon_2 = 0.1,$$

step 2: Solving LMIs (24) we obtain

$$X = \begin{bmatrix} 2.8359 & -0.9024 \\ -0.9024 & 1.2247 \end{bmatrix},$$
$$M_1 = \begin{bmatrix} -2.5784 & 0.0004 \\ -0.0003 & -2.5774 \end{bmatrix},$$
$$M_2 = \begin{bmatrix} -0.5288 & -2.0491 \\ -3.1068 & 0.5289 \end{bmatrix},$$

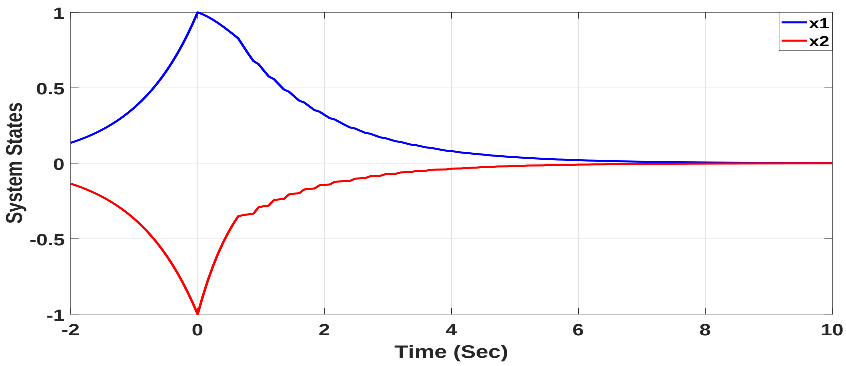


Figure 1: States $x_1(t)$ and $x_2(t)$.

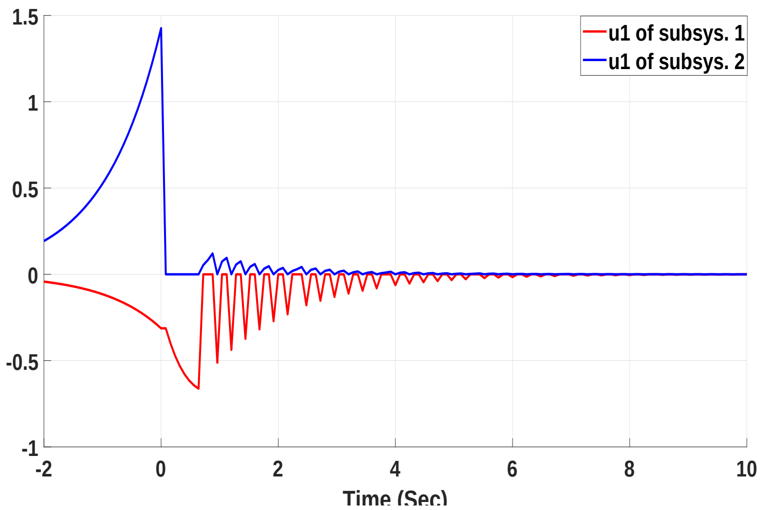


Figure 2: Control input $u_1(t)$ of each subsystem.

and thus

$$P = X^{-1} = \begin{bmatrix} 0.4606 & 0.3394 \\ 0.3394 & 1.0667 \end{bmatrix},$$

$$K_1 = \begin{bmatrix} -1.1876 & -0.8747 \\ -0.8750 & -2.7493 \end{bmatrix},$$

$$K_2 = \begin{bmatrix} -0.9391 & -2.3652 \\ -1.2516 & -0.4904 \end{bmatrix},$$

System states start from an initial condition x_0 and Figure 1 shows the state $x_1(t)$ and $x_2(t)$ and, Figure 2, Figure 3 and Figure 4 show control inputs $u_1(t)$ and $u_2(t)$ of each subsystem and switching signal $\sigma(x,t)$. It can be seen that theoretical results

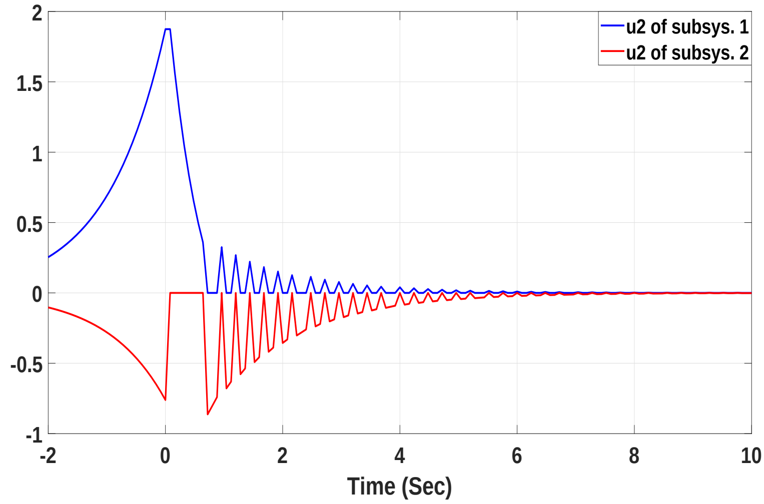


Figure 3: Control input $u_2(t)$ of each subsystem.

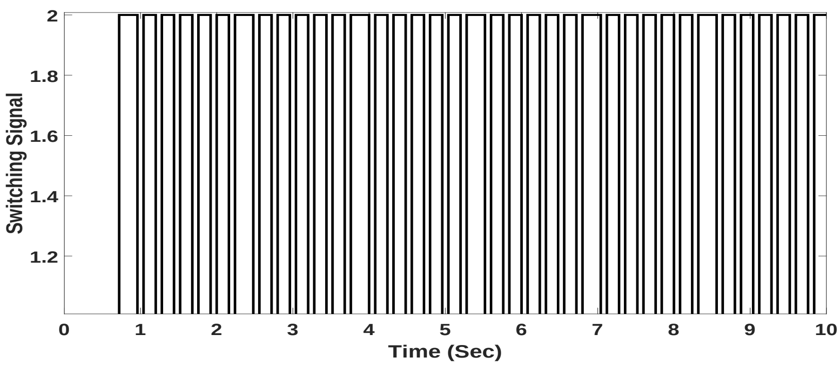


Figure 4: Switching signal $\sigma(x,t)$.

in the Theorem 2 and Theorem 3 which state that uncertain switched system (1) is exponentially stable under applying proposed switching strategy, are coincide with the simulation's results.

5 Conclusion

In this paper, a robust switching law for the GCC problem of a general form of uncertain time-delay switched system is designed. The presented method is based on using the LKF technique and extension of the min-projection switching strategy in this type of switched system. Also, uncertainties in each subsystem's dynamics are considered randomly and are additive. Besides switching law, guaranteed linear control is obtained via the solution of extracted LMIs in the presented theorems. Finally, simulation verifies the theorem's results.

References

[1] Liberzon D., 2003. "Switching in systems and control", Springer Science and Business Media.

[2] Sun Z., 2006. "Switched linear systems: control and design", Springer Science and Business Media.

[3] Branicky M.S. 1998. "Multiple Lyapunov functions and other analysis tools for switched and hybrid systems", IEEE Transactions on automatic control, 43(4), 475-482.

[4] DeCarlo R.A., Branicky M.S., Pettersson S., Lennartson B. 2000. "Perspectives and results on the stability and stabilizability of hybrid systems", Proceedings of the IEEE, 88(7), 1069-1082.

[5] Lin H., Antsaklis P. J. 2009. "Stability and stabilizability of switched linear systems: a survey of recent results", IEEE Transactions on automatic control, 54(2), 308-322.

[6] Liberzon D., Morse A.S. 1999. "Basic problems in stability and design of switched systems", IEEE control systems magazine, 19(5), 59-70.

[7] Peleties P., DeCarlo R. 1991. "Asymptotic stability of m-switched systems using Lyapunov-like functions", American Control Conference, IEEE, 1679-1684.

[8] Ishii H., Francis B.A. 2001. "Stabilizing a linear system by switching control with dwell time", In Proceedings of the 2001 American Control Conference, 3, 1876-1881.

[9] Xie Y., Gui W., Wang Y., Jiang Z., Ding S. 2009. "Memory state feedback controller design for singular systems with multiple internal constant point delays", IET Control Theory and Applications, 3(6), 631-641.

[10] Moon Y.S., Park P., Kwon W. H., Lee Y.S. 2001. "Delay-dependent robust stabilization of uncertain state-delayed systems", International Journal of control, 74(14), 1447-1455.

[11] Meyer C., Schroder S., De Doncker R.W. 2004. "Solid-state circuit breakers and current limiters for medium-voltage systems having distributed power systems", IEEE transactions on power electronics, 19(5), 1333-1340.

[12] Kim S., Campbell S.A., Liu X. 2006. "Stability of a class of linear switching systems with time delay", IEEE Transactions on Circuits and Systems I: Regular Papers, 53(2), 384-393.

[13] Xie G., Wang L. 2003. "Controllability of switched linear discrete-time systems with time delay", American Control Conference, IEEE, 4, 2797-2802.

[14] Lien X., Li S., Zou Y. 2017. "Finite-time stabilization of switched linear time-delay systems with saturating actuators", Applied Mathematics and Computation, 66-79.

[15] Mazenc F., Ahmed S., Ozbay H. 2017. "State Feedback Stabilization of Switched Systems with Delay: Trajectory Based Approach", American Control Conference, IEEE, 4540-4543.

[16] Chang S.S.S.L., Peng T. 1972. "Adaptive guaranteed cost control of systems with uncertain parameters", IEEE Transactions on Automatic Control, 17(4), 474-483.

[17] Hu H., Jiang B., Yang H. 2016. "Reliable guaranteed-cost control of delta operator switched systems with actuator faults: mode-dependent average dwell-time approach", IET Control Theory and Applications, 10(1), 17-23.

[18] Nian X., Sun Z., Wang H., Zhang H., Wang X. 2013. "Bilinear matrix inequality approaches to robust guaranteed cost control for uncertain discrete-time delay system", Optimal Control Applications and Methods, 34(4), 433-441.

[19] Wang Y., Wang Q., Zhou P., Duan D. 2012. "Robust guaranteed cost control for singular markovian jump systems with time-varying delay", ISA transactions, 51(5), 559-565.

[20] Wang R., Zhao J. 2006. "Non-fragile hybrid guaranteed cost control for a class of uncertain switched linear systems", Journal of Control Theory and Applications, 4(1), 32-37.

[21] Ramezani-al M., Kamyad A.V., Pariz N. 2015a. "A new switching strategy design for uncertain switched linear systems based on min–projection strategy in guaranteed cost control problem", IMA Journal of Mathematical Control and Information, 33(4), 1033-1049.

[22] Ramezani-al M., Kamyad A.V., Pariz N. 2015b. "A new switching strategy for exponential stabilization of uncertain discrete-time switched linear systems in guaranteed cost control problem", Iranian Journal of Electrical and Electronic Engineering, 11(2).

[23] Zhao C. 2011. "Memory hybrid guaranteed cost control for uncertain nonlinear time-delay systems", International Journal of Information Systems Science, 7, 42-51.

[24] Hien L.V., Phat V.N. 2009. "Exponential stabilization for a class of hybrid systems with mixed delays in state and control", Nonlinear Analysis: Hybrid Systems, 3(3), 259-265.

[25] Tian X., Li C.m., Chen S. 2012. "Guaranteed cost control of linear uncertain time-delay switched singular systems based on an LMI approach", International Journal of Modelling, Identification and Control, 17(2), 109-115.

[26] Boyd S., El Ghaoui L., Balakrishnan V. 1994. "Linear matrix inequalities in system and control theory", SIAM.

[27] Phat V., Pairote S. 2007. "Exponential stability of switched linear systems with time-varying delay", Electronic Journal of Differential Equations, 159, 1-10.

[28] Phat V., Botmart T., Niamsup P. 2009. "Switching design for exponential stability of a class of nonlinear hybrid time-delay systems", Nonlinear Analysis: Hybrid Systems, 3(1), 1-10.

[29] Sun X.M., Wang W., Liu G.P., Zhao J. 2008. "Stability analysis for linear switched systems with time-varying delay", IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 38(2), 528-533.

[30] Chen X., Du S.L., Wang L.D., Liu L.J. 2016. “Stabilization for linear uncertain systems with switched time-varying delays”, Neurocomputing, 191, 296-303.

[31] Wang R., Zhao J. 2007. “Guaranteed cost control a class of uncertain delay systems with actuator failures based on switching method”, International Journal of Control Automation and Systems, 5(5), 492-500.

[32] Tissir E.H. 2011. “Exponential stability and guaranteed cost of switched linear systems with mixed time-varying delays”, ISRN Applied Mathematics, Vol. 2011.

[33] Hien L.V., Phat V.N. 2009. “Exponential stability and stabilization of a class of uncertain linear time-delay systems”, Journal of the Franklin Institute, 346(6), 611-625.

[34] Lien C.H., Yu K.W., Chung Y.J., Lin Y.F., Chung L.Y., Chen J.D. 2009. “Exponential stability analysis for uncertain switched neutral systems with interval-time-varying state delay”, Nonlinear Analysis: Hybrid Systems, 3(3), 334-342.

[35] Dong Y., Li T., Mei S. 2016. “Exponential stabilization and L2-gain for uncertain switched nonlinear systems with interval time-varying delay”, Mathematical Methods in the Applied Sciences, 39(13), 3836-3854.

[36] Lien G., Li X., Zou H. 2017. “Robust Finite-time Guaranteed Cost Control for Impulsive Switched Systems with Time-varying Delay”, International Journal of Control, Automation and Systems, 15(1), 1-9.

[37] Cao X., Liu L., Fu Z., Song X., Song S. 2017. “Guaranteed cost finite-time control for positive switched linear systems with time-varying delays”, Journal of Control Science and Engineering, Vol. 2017.

[38] Wu L., Wang Z. 2009. “Guaranteed cost control of switched systems with neutral delay via dynamic output feedback”, International Journal of Systems Science, 40(7), 717-728.

[39] Pettersson S., Lennartson B. 2001. “Stabilization of hybrid systems using a min-projection strategy”, American Control Conference, IEEE, 1, 223-228.

[40] Meyer C. D. 2000. “Matrix analysis and applied linear algebra”, Siam, Vol. 71.

[41] Thuan M.V. 2019. “Robust finite-time guaranteed cost control for positive systems with multiple time delays”, Journal of System Science Complexity, 32, 496-509.

[42] Huanli G., Xueliang Liu., Fuchun Liu. 2019. “Robust guaranteed cost control for continuous-time uncertain Markov switching singular systems with mode-dependent time delays”, Neural Computing and Application, 31, 4479-4490.

[43] Pouralizadeh Moghaddam F., Gholizade Narm H. 2019. “Development of RMPC Algorithm for Compensation of Uncertain Time-Delay and Disturbance in NCS”, Control and Optimization in Applied Mathematics (COAM), 4(1), Payame Noor University.

[44] Amin Regaieg M., Kchaou M., El Hajjaji A., Chaabane M. 2019. “Robust H guaranteed cost control for discrete-time switched singular systems with time-varying delay”, Optim. Control Appl. Meth, 40(1), 19-140.

Control and Optimization in Applied Mathematics (COAM)
DOI. 10.30473/coam.2020.53510.1141
Vol. 4, No. 2, Autumn - Winter 2019 (19-37), ©2016 Payame Noor University, Iran

Enlarging the Region of Attraction for Nonlinear Systems through the Sum-of-Squares Programming

J. Alizadeh[1], H. Khaloozadeh[2*]

[1]Department of Mathematics, Payame Noor University (PNU), P.O. Box. 19395-3697, Tehran, Iran,
[2]Department of Systems and Control, K.N.Toosi University of Technology, P.O. Box. 16315-1355, Tehran, Iran.

Received: June 27, 2020; **Accepted:** November 14, 2020.

Abstract. In the present study, a novel methodology is developed to enlarge the Region of Attraction (ROA) at the point of equilibrium of an input-affine nonlinear control system. Enlarging the ROA for non-polynomial dynamical systems is developed by designing a nonlinear state feedback controller through the State-Dependent Riccati Equation (SDRE). Consequently, its process is defined in the form of a Sum-of-Squares (SOS) optimization problem with control and non-control constraints. Of note, the proposed technique is effective in estimating the ROA for a nonlinear system functioning on polynomial or non-polynomial dynamics. In the present study, the application of the proposed scheme is shown by numerical simulations.

Keywords. State-dependent Riccati equation, Region of attraction, Sum-of-squares programming, Lyapunov function.

MSC. 90C34; 90C40.

* Corresponding author
Jahangir_Alizadeh@yahoo.com, h_ khaloozadeh@kntu.ac.ir
http://mathco.journals.pnu.ac.ir

1 Introduction

Obtaining the ROA for a nonlinear system at a point of equilibrium is challenging for the control theory. The ROA at a point of equilibrium comprises a set of initial conditions the state trajectories of which converge on the equilibrium again. With a small ROA, a disturbance may quickly take the system out of the region and the system can not get back to the stable point of equilibrium. Therefore, the extent of the ROA is one of the criteria for the stability of nonlinear systems around the respective point of equilibrium. Computing the ROA is not an easy task and hence, its estimation has turned into a crucial problem.

In the present study, a new methodology for the enlargement of the ROA of input-affine nonlinear systems is developed by designing a nonlinear state feedback controller through the SDRE. The equation gives a shape factor that expands the ROA. Since the choice of the shape factor is affected by the system dynamics, a non-uniform extension of the ROA is established. The main advantage of the proposed approach is incorporating both polynomial and non-polynomial dynamics. Moreover, the proposed method deals with both bound and unbound ROAs. Accordingly, the ROA enlargement problem is defined as an SOS optimization problem. The present method is evaluated in comparison with the existing methods by its application to various nonlinear systems that have already been addressed in the previous studies. The methods proposed in the existing literature for the ROA estimation can be classified into categories, namely Lyapunov [1]-[9] and non-Lyapunov [10]-[13]. In the Lyapunov-based methods, a Lyapunov Function (LF) is sought that is positive definite with negative definite time derivative for its largest sublevel set. The mentioned largest sublevel set is considered as an estimate of the ROA. Reference [1] presented a new design to maximize the ROA for a saturated supercavitating vehicle. To cope with the immeasurability of the vertical speed, the development of the presented design was made possible through output feedback control schemes. The conditions required by the achieved controller to locally and asymptotically stabilize the closed-loop system were represented by LMI presentations.

Reference [3] developed a method for ROA enlargement in nonlinear systems by trajectory reversing. Numerical simulations were performed to validate the suggested method employing modeling systems of Van der Pol oscillator and Hahn.

In [5], stability analysis of polynomial systems was carried out based on LFs. For ROA enlargement, a region with variable size was used as a shape factor. The purpose was to achieve the biggest sublevel set of LFs with the biggest shape factor. In [6], similar to [5], by adopting a polynomial as a shape factor, the ROA was enlarged and the use of bilinear SOS programming with polynomial LFs was suggested. Also, due to their higher richness than the quadratic LFs, the sublevel sets of higher-order polynomial LFs were used. However, it should be noted that as soon as the degree of LF increases, the number of decision variables is considerably enlarged. Hence, to lower the number of decision variables, employing the maximum or a minimum number of a group of polynomial functions was suggested as a solution in [7]. Khodadadi et al. [9] proposed a numerical methodology for ROA estimating in nonlinear polynomial systems through SOS programming. On the other hand, for estimation enlargement, an SOS optimization problem was solved by defining a subset of the invariant set through a shape factor.

Reference [10] proposed a non-Lyapunov method for the estimation of the Robust Domain of Attraction (RDA) and directional enlargement of the Domain of Attraction (DOA) by using Markov chains and an invariant measure. Their methodology formulated the estimation of RDA and directional enlargement of DOA as an infinite-dimensional linear problem. A major shortcoming of Markov modeling in the estimation of DOA is that it incorporates the real DOA into the estimated one; as a result, in boundary partitioning, it does not ensure achieving stability. To tackle this problem, they suggested refining any partition set that had a large invariant measure. In recent studies, control parameters have been employed in the enlargement of the ROA [14]-[22]. Enlarging ROA in non-linear systems is a substantial issue for the designers of non-linear controllers. For the systems with large ROAs around a point of equilibrium, the problems concerning tracking and disturbance can be systematically tackled. In particular, it would be very useful to characterize the controllers that maximize the ROA. In [15], a formulation for Model Predictive Control (MPC) was presented in order to enlarge the DOA without the requirement to enlarge the horizon of prediction. An array of contractive control invariant sets was employed to substitute the MPC terminal region. Therefore, computing the contractive array of control invariant sets was essential in the developed formulation, which was solved only for linear systems.

Haghighatnia and Moghaddam [18] offered a method to enlarge the robust ROA in uncertain systems by designing linear controllers. The problem of enlarging the robust ROA was formulated as a novel optimization at three levels, which sought the optimal controlling parameters of the linear controller. The present study is arranged as follows. The ROA is estimated and then, enlarged; the pseudo-linearization is discussed, and the SDRE approach is developed in Section 2. The achieved results are illustrated in Section 3. In Section 4, by illustrative numerical examples, the effectiveness of the proposed method is proven and a comparison is made with other published works in the literature. Finally, a conclusion is briefly made in Section 5.

2 Preliminaries

In the following, the notation utilized in this study is provided.
R^n : an n-dimensional vector space over the field of the real numbers,
$\aleph_n$: the set of all polynomials with real coefficients in n variables,
Σ_n: the SOS polynomials set in n variables,
Ψ : the state set that is a bounded open subset of R^n Euclidean space and contains the origin,
$C^k(\Psi)$: tthe class of functions continuously differentiable for k times in Ψ.
Assume a system in the following form:

$$\dot{x} = f(x(t)), \tag{1}$$

where $x(t) \in R^n; x(0) = x_0$ is the initial state at $t = 0$ and f is an n-vector of elements of $\aleph_n$ with $f(0) = 0$.

Definition 1 (point of equilibrium)**.** $x_e \in R^n$ is a point of equilibrium in system (1) if $f(x_e) = 0$. The points of equilibrium in system (1) correlate with the intersection of its nullclines, i.e., the curves represented by $f(x) = 0$. Without loss of generality, we suppose that the point of equilibrium under investigation concurs with the origin of the state space of $R^n, (x_e = 0)$ in the sequel.

Definition 2 (ROA)**.** If the origin is a point of equilibrium for (1), the ROA of the origin is defined as follows:

$$\Omega = \{x_0 \in R^n \mid lim_{t\to 0}\phi(x_0, t) = 0\},$$

where $\phi(x_0, t)$ denotes the solution starting from the initial state.

Definition 3 (SOS polynomials)**.** A multivariate polynomial $P(x_1, x_2, ..., x_n) \triangleq P(x)$ is an SOS if polynomials $f_1(x), ..., f_m(x)$ exist such that

$$P(x) = \sum_{i=1}^{m} f_i^2(x). \tag{2}$$

From the definition, it can be deduced that the SOS polynomials set in n variables is a convex cone. It can be indicated that the existence of an SOS decomposition (2) is equivalent to the existence of a positive semidefinite matrix Q such that

$$P(x) = Z^T(x)QZ(x), \tag{3}$$

where $Z(x)$ is some properly chosen monomials vector. It is undisputed that an SOS polynomial is globally nonnegative. As a major characteristic of SOS polynomials, this is decisive for many applications in the field of control; in particular, the cases in which different polynomial inequalities are substituted with SOS conditions [23].

2.1 Estimating ROA

A Lyapunov-based method is developed for estimating the ROA. Enlarging the ROA provides more freedom in designing nonlinear controllers. Moreover, it can be considered as a way of improving the performance of nonlinear closed-loop systems. In this regard, we investigate a positive definite LF with a negative definite time derivative in the largest sublevel set. The mentioned largest sublevel set is an estimate of the ROA. We aim to achieve a provable ROA for the system such that all the points that start in this region will converge to the fixed point of origin.

Theorem 1 ([24])**.** If we have a function $V : R^n \to R$ that is continuously differentiable as follows:

$$V \quad is\ positive\ definite, \tag{4}$$

$$\Omega := \{x \in R^n \mid V(x) \le c\} \quad is \quad bounded, \quad and \tag{5}$$

$$\{x \in R^n \mid V(x) \le c\} \setminus \{0\} \subseteq \{x \in R^n \mid \frac{\partial V}{\partial t} f(x) < 0\}, \tag{6}$$

where c is a positive value. For all $x(0) \in \Omega$, a solution for (1) can be achieved and $lim_{t\to\infty} x(t) = 0$. Accordingly, Ω is the subset of the ROA for (1) and it is invariant. The continuously differentiable function $V(x)$ is called a local Lyapunov function. For arriving at a better estimate of the ROA, we need to acquire a $V(x)$ that leads to a larger Ω.

2.2 Enlarging ROA

One of the parameters that have a significant effect on the estimation of ROA is the shape factor. However, no systematic method for determining and selecting the shape factor has been presented so far. In the previous studies, since a fixed shape factor has always been used, the uniform extension of ROA has been addressed. In the present article, the shape factor is obtained through the SDRE approach and the dynamics of the system are effective in its selection. Therefore, the non-uniform extension of ROA is dealt with. With the aim of enlarging Ω, we consider a region with variable size $P_\beta = \{x \in R^n \mid P_0(x) \le \beta\}$, where $P_0(x)$ is a positive convex polynomial called the shape factor. Maximizing β, as long as $P_\beta \subseteq \Omega$ is established, leads to an estimation of the ROA. With the application of theorem 1, we can formulate the problem of estimating the ROA as an optimization problem in the following form [5]:

$$\begin{aligned}
&\max_{V\in R_n} \quad \beta \\
&s.t. \\
&\qquad V(x) > 0 \quad for\ all \quad x \in R^n \setminus \{0\} \ and \ V(0) = 0 \\
&\qquad the\ set\ \Omega \quad is \quad bounded, \\
&\qquad \{x \in R^n \mid P_0(x) \le \beta\} \subseteq \Omega, \\
&\qquad \{x \in R^n \mid V(x) \le c\} \setminus \{0\} \subseteq \{x \in R^n \mid \frac{\partial V}{\partial t} f(x) < 0\}.
\end{aligned} \tag{7}$$

Given that the choice of the shape factor in the estimation of the ROA plays a fundamental role, in this paper, the shape factor is obtained based on the SDRE idea.

2.3 Pseudo Linearization

The pseudo-linearization methodology represents the nonlinear system as a linear-like system for which the matrix is dependent on state variables. Assume an input-affine nonlinear system in the following form:

$$\dot{x} = f(x) + B(x)u. \tag{8}$$

where $x \in R^n$ is the state vector, $u \in R^m$ is the input vector, and $f : R^n \to R^n$, $B : R^n \to R^{n\times m}$. Here it is assumed that $f(0) = 0$. In short, the concept of pseudo-linearization for Eq. (8) can be formulated as follows:

$$\dot{x} = A(x)x + B(x)u. \tag{9}$$

Matrices $A(x)$ and $B(x)$ are called State-Dependent Coefficients (SDC) matrices [30]. Assuming $f(0) = 0$ and $f(.) \in C^1(\Psi)$, there is always a function $A(x)$ that is continuous nonlinear and matrix-valued as follows:

$$f(x) = A(x)x, \tag{10}$$

where $A : \Psi \to R^{n\times n}$ is achieved by mathematical factorization; it is obviously nonunique with $n > 1$. Of note, the above-mentioned assumptions for $f(x)$ ensure that a global SDC parameterization of $f(x)$ exists on Ψ. The theorem below indicates that $f(x)$ can be reformulated as given in (10).

Theorem 2 ([30]). Let us have $f : \Psi \to R^n$ in a way that $f(0) = 0$ and $f(.) \in C^k(\Psi), k \geq 1$. For all $x \in R^n$, there is an SDC parameterization (10) of $f(x)$ for some $A : R^n \to R^{n\times n}$. An instance of such parameterization ensured by the given conditions is the following

$$A(x) = \int_{t=0}^{1} \frac{\partial f(x)}{\partial x}|_{x=\lambda x} d\lambda, \tag{11}$$

where λ is a dummy variable introduced in the integration.

Proof. (11) can be validated by assuming the functions set: $\hat{f} : R \to R^n$ established by $\hat{f}(\lambda) \triangleq f(\lambda x)$. Then, for each $x \in R^n$,

$$f(x) = \hat{f}(1) = \hat{f}(0) + \int_{t=0}^{1} \frac{d\hat{f}(\lambda)}{d\lambda} d\lambda.$$

$\hat{f}(0) = 0$, as assumed, and $\frac{d\hat{f}(\lambda)}{d\lambda} = (\frac{\partial f}{\partial x}|_{x=\lambda x})x$, thus

$$f(x) = \int_{t=0}^{1} \frac{\partial f(x)}{\partial x}|_{x=\lambda x} d\lambda x. \tag{12}$$

Comparing (12) with (10) gives the desired result (11). Using extended linearization, any input-affine nonlinear system (8) that meets the conditions for $f(x)$ given in theorem (2) can always be formulated as an SDC (9). Of note, pseudo-linearization has many advantages. First, unlike Jacobian linearization, it retains all nonlinearity properties of the system and second, its non-uniqueness creates extra degrees of freedom that can be used to enhance controller performance [30]-[31].

2.4 SDRE Method

The SDRE method originates in extending the Linear Quadratic Regulator (LQR) problem into nonlinear systems by maintaining all the nonlinear properties. It functions on pseudo-linearization of the system and entails factorization (i.e., parameterization) [30] of the nonlinear dynamics to the state vector and a matrix-valued function dependent on the state. In this regard, the SDRE algorithm completely incorporates the nonlinearities of the system, turning it into a (non-unique) linear structure having SDC matrices and minimizing a nonlinear performance index with a quadratic-like structure. Consider the system (9) with cost function

$$J = \frac{1}{2}\int_{t=0}^{\infty}(x^T(t)Q(x)x(t) + u^T(t)R(x)u(t))dt, \tag{13}$$

where $Q(x)$ and $R(x)$ are state-dependent weighting matrices to meet $Q(x) \geq 0$ and $R(x) > 0$ for all the values of x. It is clear that the design of the SDRE controller is similar to the LQR one with minor variations in matrices A, B, Q and R, which are state-dependent. However, Q and R do not necessarily require to be state-dependent. Therefore, in order to minimize (13), the algebraic SDRE in the following should be solved [31]:

$$A^T(x)P(x) + P(x)A(x) - P(x)B(x)R^{-1}(x)B^T(x)P(x) + Q(x) = 0. \tag{14}$$

The optimal control for minimization of the cost function (13) is

$$u(t) = -R^{-1}(x)B^T(x)P(x)x, \tag{15}$$

where $P(x)$ is the unique, symmetric, positive definite solution for (14). The major privilege of the SDRE method may be the non-uniqueness of the pseudo-linear representation of the system, which gives the designer more freedom. By replacing delay in the system matrix, Batmani and Khaloozadeh introduced a method to find a sub-optimal solution for a class of nonlinear time-delayed systems using the SDRE method [25]-[29].

3 Main Results

In the following, first, the nonlinear optimal control system is written in the pseudo-linear form. Then, the nonlinear state feedback control is obtained through SDER. The controller, while stabilizing the system and minimizing the cost function, leads to the expansion of the ROA by using the shape factor. Since the choice of this shape factor is affected by the system dynamics, a non-uniform extension of the ROA will be established. Consequently, the problem of enlarging ROA is formulated in the form of an SOS optimization problem. Assume the following class of input-affine nonlinear systems:

$$\dot{x} = f(x) + B(x)u. \tag{16}$$

where $f : R^n \to R^n, \quad B : R^n \to R^{n\times m}$, with $x \in R^n$ representing the state and $u \in R^m$ indicating the control input under the component-wise saturation constraints as follows

$$u \in U : \quad | u_i | \leq \bar{u}_i \quad , i = 1, ..., m. \tag{17}$$

with U is a compact subset of R^m comprising the origin as an interior point. The origin is considered as a point of equilibrium for (16) when $u = 0$ i.e. $f(0) = 0$ ([32]). We seek for a strategy of state feedback regulation $u = K(x)$ that asymptotically stabilizes (16) into the origin under (17), has the largest estimate of the ROA, and minimizes the cost function (13). Reference [33] sought to obtain an LF $V(x)$ which was an upper bound for the cost function: $J(K(x), x) \leq V(x)$. According to [5], when an SOS $V(x)$ and a polynomial state-dependent control law $u = K(x)$ can be achieved at a considered time instant t that are consistent with (17) such that $J(K(x), x) \leq V(x)$, the set $\Omega = \{x \in R^n | V(x) \leq c\}$ is a positive invariant region for the regulated input constrained plant. First, we write the system (16) in the following pseudo-linear form:

$$\dot{x} = A(x)x + B(x)u. \tag{18}$$

Applying the SDRE method, we calculate the feedback control $u = K(x)$. Replacing it in (18) we obtain:

$$\begin{cases} \dot{x} = A(x)x + B(x)K(x) = F(x) \\ K(x) \in U \end{cases} \tag{19}$$

Now the problem of estimating the ROA of (19) is written as the following optimization problem:

$$\begin{aligned} &\max_{V \in \mathcal{R}_n} \quad \beta \\ &s.t. \\ &\quad V(0) = 0, V(x) > 0, \quad x \in R^n \\ &\quad \{x \in R^n \mid P_0(x) \leq \beta\} \subseteq \{x \in R^n \mid V(x) \leq c\}, \\ &\quad \{x \in R^n \mid V(x) \leq c\} \setminus \{0\} \subseteq \{x \in R^n \mid \frac{\partial V}{\partial x} F(x) < 0\} \\ &\quad \{x \in R^n \mid V(x) \leq c\} \subseteq \{x \in R^n \mid \ \mid K(x) \mid \leq \bar{u}_i\}. \end{aligned} \tag{20}$$

In order to enlarging the ROA, using the Positive Stellensatz theorem [34], the optimization problem (20) is converted to the SOS programming problem as follows:

$$\begin{aligned} &\max_{V \in \mathcal{R}_n, s_1, s_2, s_3, s_4 \in \Sigma_n} \quad \beta \\ &\qquad s.t. \\ &\qquad\quad V - l_1 \in \Sigma_n \\ &\qquad\quad (c - V - s_1(\beta - P_0)) \in \Sigma_n \\ &\qquad\quad -\left(\frac{\partial V}{\partial x} F(x) + l_2 + s_2(c - V)\right) \in \Sigma_n \\ &\qquad\quad (\bar{u}_i - K - s_3(c - V)) \in \Sigma_n \\ &\qquad\quad (\bar{u}_i + K - s_4(c - V)) \in \Sigma_n \end{aligned} \tag{21}$$

where s_1, s_2, s_3 and s_4, are SOS polynomials. Also, $l_i(x)$ is a positive definite polynomial as $l_i(x) = \sum_{j=1}^{n} \varepsilon_{ij} x_j^2$ for $i = 1, 2$, in which ε_{ij} are numbers with positive values.

The previous description is summarized in the form of an algorithm below:

Step 1: linearize the system and compute the SDRE control.

Writing Equation (16) in the form of (18) and then, designing a nonlinear state feedback controller according to the SDRE and replacing it in (18), we obtain (19).

Step 2: To calculate the initial shape factor and LF, do:

$$A = \frac{\partial F}{\partial x}\Big|_{x=0}. \tag{22}$$

Then, solve equation (23)

$$A^T P_1 + P_1 A = -I. \tag{23}$$

Using the matrix $P_1 \geq 0$, we define the initial shape factor and LF as follows:

$$P_0(x) = x^T P_1 x, \tag{24}$$

$$V = P_0(x). \tag{25}$$

Step 3: Set V to a fixed value and perform the SOS optimization for s_2, s_3, s_4

$$\begin{aligned} &\max_{c \in R, s_2, s_3, s_4 \in \Sigma_n} \quad c^* \\ &s.t. \\ &\quad -\left(\frac{\partial V}{\partial x} F(x) + l_2 + s_2 (c - V)\right) \in \Sigma_n, \\ &\quad (\bar{u}_i - K - s_3 (c - V)) \in \Sigma_n, \\ &\quad (\bar{u}_i + K - s_4 (c - V)) \in \Sigma_n. \end{aligned} \tag{26}$$

Step 4: Set V and $P0$ to fixed values and perform the SOS optimization for s_1

$$\begin{aligned} &\max_{\beta \in R, s_1 \in \Sigma_n} \quad \beta^* \\ &s.t. \\ &\quad (c - V - s_1 (\beta - P_0)) \in \Sigma_n. \end{aligned} \tag{27}$$

Step 5: Using $s_1, s_2, s_3, s_4, c^*, \beta^*, P_0$ from the previous steps, compute V as following:

$$\begin{aligned} &V - l_1 \in \Sigma_n \\ &(c - V - s_1 (\beta - P_0)) \in \Sigma_n, \\ &-\left(\frac{\partial V}{\partial x} F(x) + l_2 + s_2 (c - V)\right) \in \Sigma_n, \\ &(\bar{u}_i - K - s_3 (c - V)) \in \Sigma_n, \\ &(\bar{u}_i + K - s_4 (c - V)) \in \Sigma_n. \end{aligned} \tag{28}$$

Step 6: Use the quadratic part of the new LF as new P_0 and replace $V \ with \dfrac{V}{c^*}$
Using the newly achieved LF and shape factor, make repetitions to get convenient LF.

4 Simulation Results

The following instances are given to demonstrate capability of the method developed in the present study in ROA enlargement for the nonlinear optimal control system.

Example 1. Consider the following input-affine nonlinear system examined in reference [33]

$$\begin{cases} \dot{x}_1 = x_2(t), \\ \dot{x}_2 = -x_1(t) - \left(1 - x_1^2(t)\right) x_2(t) + u(t). \end{cases} \tag{29}$$

The weighting matrices $Q = diag([0.01 \;\; 0.01]), R = 1$ and input saturation constraint $|u(t)| \leq 0.2.$ are assumed. In [33], the system's ROA was estimated by the fixed shape factor $p_0(x) = x^T x = x_1^2 + x_2^2$, as Figure 1.

In this paper, we search for the feedback control $u(t) = K(x(t))$ that stabilizes the system asymptotically, minimizes the cost function, and has the largest estimation of the ROA. We use the proposed method for Example 1.

Step 1: Note that $l_1 = l_2 = 10^{-6}(x_1^2 + x_2^2)$.

First, we reformulate the system (29) in the pseudo-linear form below:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -(1 - x_1^2(t)) \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t). \tag{30}$$

Then, by applying the SDRE method (14), the matrix $P \geq 0$ is obtained. Matrix-array P is state-dependent:

$$P = \begin{bmatrix} 0.005x_1^2 + 0.015 & 0.005 \\ 0.005 & 0.01x_1^2 + 0.01 \end{bmatrix}, \tag{31}$$

$$u = -Kx = -R^{-1}B^T Px = -(0.005x_1 + 0.01x_1^2 x_2 + 0.01x_2). \tag{32}$$

The control obtained by the SDRE method leads to the minimum cost function. By replacing (32) in the control system (30) we obtain:

$$\begin{cases} \dot{x}_1 = x_2(t) \\ \dot{x}_2 = -1.005x_1(t) + 0.99x_1^2(t)x_2(t) - 1.01x_2(t) \end{cases} \Rightarrow \dot{x} = F(x). \tag{33}$$

Step 2: by using (22) and solving Equation (23), obtain:

$$P_1 = \begin{bmatrix} 1.4951 & 0.4975 \\ 0.4975 & 0.9876 \end{bmatrix}. \tag{34}$$

Using the matrix $P_1 \geq 0$, we define the initial shape factor as follows:

$$P_0(x) = x^T P_1 x = 1.4951x_1^2 + 0.99502x_1x_2 + 0.98764x_2^2, \tag{35}$$

$$V(x) = P_0(x), \tag{36}$$

Step 3: by using (26), obtain:

$$s_2 = 0.14130x_1^4 - 0.05192x_1^3x_2 - 0.08065x_1^2x_2^2 + 0.33568x_1^2$$
$$0.04097x_1x_2^3 + 0.09480x_1x_2 + 0.11980x_2^4 + 0.31168x_2^2$$

$$s_3 = 0.04746x_1^4 + 0.01754x_1^3x_2 + 0.05305x_1^2x_2^2 + 0.01933x_1^2$$
$$+ 0.01792x_1x_2^3 + 0.03783x_1x_2 + 0.03009x_2^4 + 0.00299x_2^2 + 0.05463$$

$$s_4 = 0.04746x_1^4 + 0.01754x_1^3x_2 + 0.05305x_1^2x_2^2 + 0.01933x_1^2$$

$$+0.01792x_1x_2^3+0.03783x_1x_2+0.03009x_2^4+0.00299x_2^2+0.05463$$

Step 4: by using (27), obtain:

$$\begin{aligned}s_1 &= 0.18938x_1^4+0.08673x_1^3x_2+0.27538x_1^2x_1^2-0.16025x_1^2\\ &+0.09146x_1x_2^3+0.15174x_1x_2+0.11653x_2^4-0.18607x_2^2+0.78816\end{aligned}$$

Step 5: by using (28), obtain:

$$\begin{aligned}V &= 0.03230x_1^6-0.00278x_1^5x_2+0.00654x_1^4x_2^2-0.08683x_1^4\\ &-0.01464x_1^3x_2^3+0.00108x_1^3x_2+0.00576x_1^2x_2^4\\ &+0.04102x_1^2x_2^2+0.59692x_1^2+0.00581x_1x_2^54+0.02795x_1x_2^3\\ &+0.38961x_1x_2+0.01006x_2^6-0.03219x_2^4+0.40588x_2^2\end{aligned}$$

Step 6: The new shape factor is: $P_0 = 0.59692x_1^2+0.38961x_1x_2+0.40588x_2^2$ Repeat the process in 35 iterations to obtain:

$$\begin{aligned}V &= 0.02717x_1^6-0.02229x_1^5x_2+0.03442x_1^4x_2^2-0.04244x_1^4\\ &+0.05791x_1^3x_2^3+0.09611x_1^3x_2+0.05382x_2^4+0.16735x_1^2x_2^2\\ &+0.16528x_1^2+0.01949x_1x_2^5+0.08147x_1x_2^3+0.10276x_1x_2\\ &+0.00546x_2^6+0.01261x_2^4+0.09084x_2^2.\end{aligned} \tag{37}$$

The level set of V in (37) is the estimated ROA for System (33), as illustrated in Figure 1. It has been compared with the estimates derived by the WK-SOS method in [33].

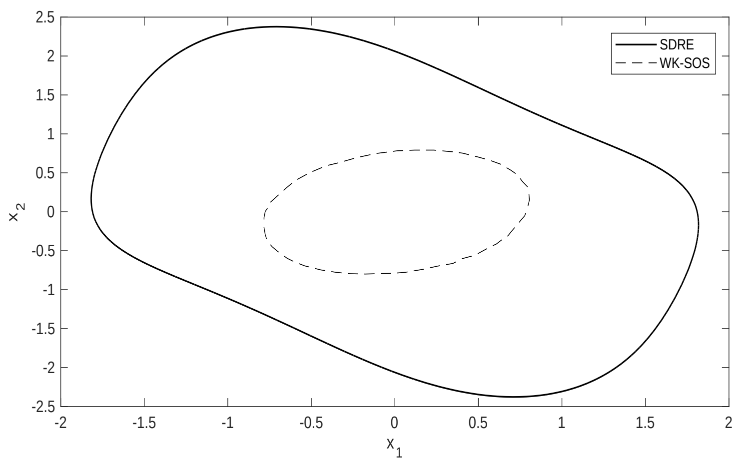


Figure 1: AR estimation provided in[33] (Dashed line) and estimated AR using proposed method (Continuous line).

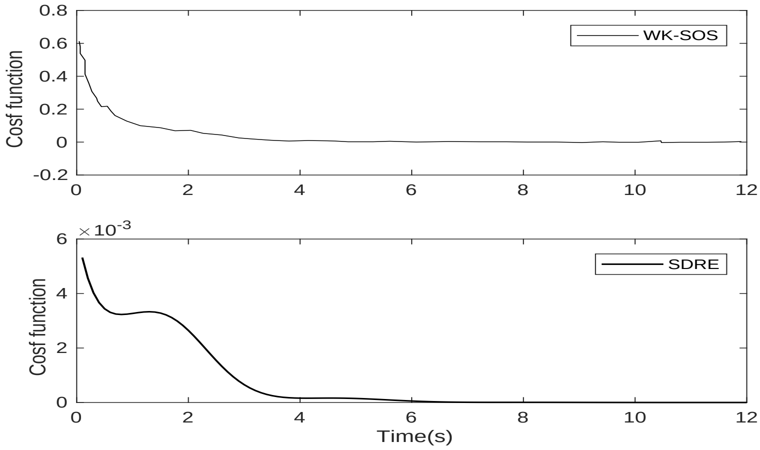


Figure 2: Cost function in the proposed method and reference [33] .

As shown in Figure 2, the cost function value in reference [33] starts from 0.7 and gradually decreases. In our method, this value is much better, starting at 0.004 and converging at a good speed. Compared to the method proposed in [33], which uses an uninspired shape factor, our method uses the shape factor by the selection of $P_0(x)$, which expands the ROA. Also, Figure 2 shows that the cost function of the proposed SDRE method is nonlinear, which indicates the efficiency of the proposed method. In Figure 3, the control function graph and state is shown.

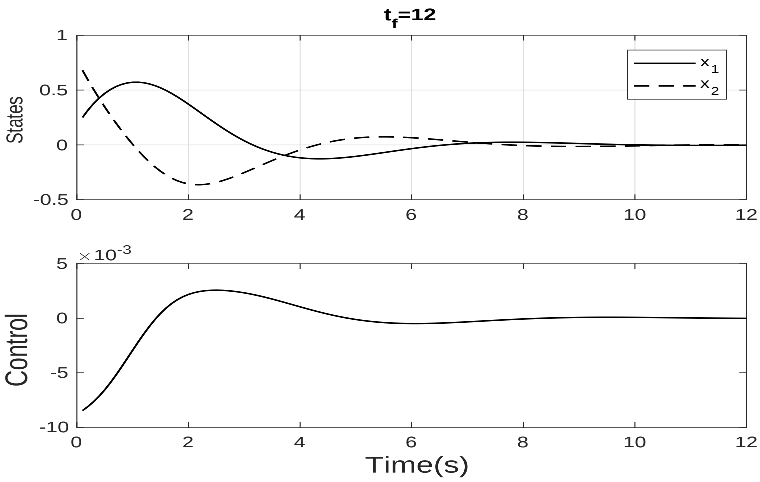


Figure 3: State trajectories and control signal by proposed method.

Remark 1. Given that SOS works with polynomials, if a system has non-polynomial dynamics, we need to find a polynomial approximation of the non-polynomial term. Then, we apply the method to the estimation of the ROA.

Example 2. For the non-linear system in [35] as given below:

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = -sin(x_1) - 0.5x_2(t), \end{cases} \tag{38}$$

by applying the Taylor series: $sin(x_1) = x_1 - x_1^3/6$ and replace it in (38) we obtain:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_1^3/6 - x_1 - 0.5x_2(t). \end{cases} \tag{39}$$

Now, we are looking for an estimation of the ROA of the system (39). By using the proposed method, the estimated ROA is plotted in Figure 4. The ROA computed in [35] is shown in Figure 5. By comparing the Figures 4 and 5, it is observed that the estimation of ROA by the method developed in this study is considerably more efficient than the estimations carried out by the method in reference [35]. As shown in Example 2, it is evident that our method is appropriate for DOA estimation in systems of both polynomial and non-polynomial types.

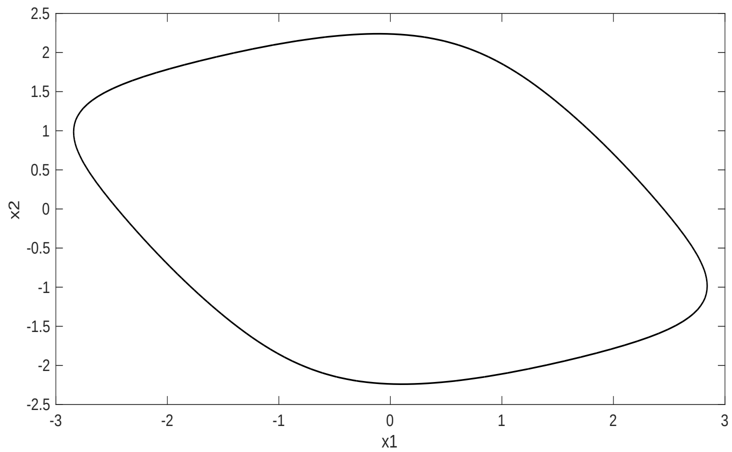


Figure 4: Estimated AR by using proposed method for Example 2.

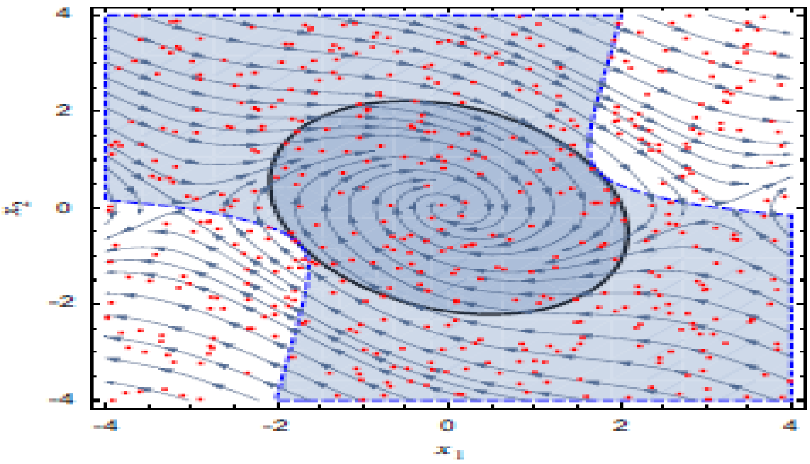


Figure 5: The black ellipsoid represents the RA estimated in [35] for Example 2. the dashed blue line (the boundary of the light blue area) demonstrates the $\dot{V}(x) < 0$, region; system trajectories are illustrated by the arrows; and the points in red indicate the randomly selected sampling states.

Example 3. Consider the following input-affine nonlinear system in [3]:

$$\begin{cases} \dot{x}_1 = -x_1 + 2x_1^2 x_2 + u, \\ \dot{x}_2 = -x_2. \end{cases} \tag{40}$$

It is further assumed that $Q = diag([1 \quad 1]), R = 1$. First, we formulate the system in the pseudo-linear form below:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & 2x_1^2 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t), \tag{41}$$

Then, we find the feedback control $u(t) = K(x(t))$ that asymptotically stabilizes the system, minimizes the cost function , and has the largest estimate of the ROA. By applying the SDRE method, the matrix P is obtained. The matrix-array is state-dependent: Then we find the feedback control $u(t) = K(x(t))$ that asymptotically stabilizes the system, the cost function minimum, and have the largest estimate of the AR. By applying the SDRE method, the matrix P is obtained. Matrix-array are state-dependent:

$$P = \begin{bmatrix} 0.414 & 0.343x_1^2 \\ 0.343x_1^2 & 0.627x_1^4 + 0.5 \end{bmatrix}, \tag{42}$$

$$u = -Kx = -R^{-1}B^T Px = -(0.414x_1 + 0.343x_1^2). \tag{43}$$

The control obtained by the SDRE method leads to the minimum cost function. By replacing (43) in the control system (41), we obtain:

$$\begin{cases} \dot{x}_1 = -1.414x_1 + 2x_1^2 x_2 - 0.343x_1^2, \\ \dot{x}_2 = -x_2. \end{cases} \tag{44}$$

Now, we are looking for an estimation of the ROA of the system (44). By solving the above problem, the estimated ROA is plotted in Figure 6.

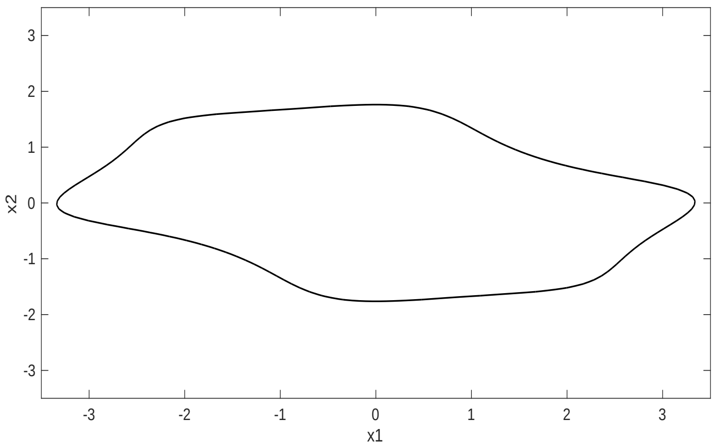


Figure 6: Estimated AR by using proposed method for Example 3.

The ROA computed in [3] is shown in Figure 7.

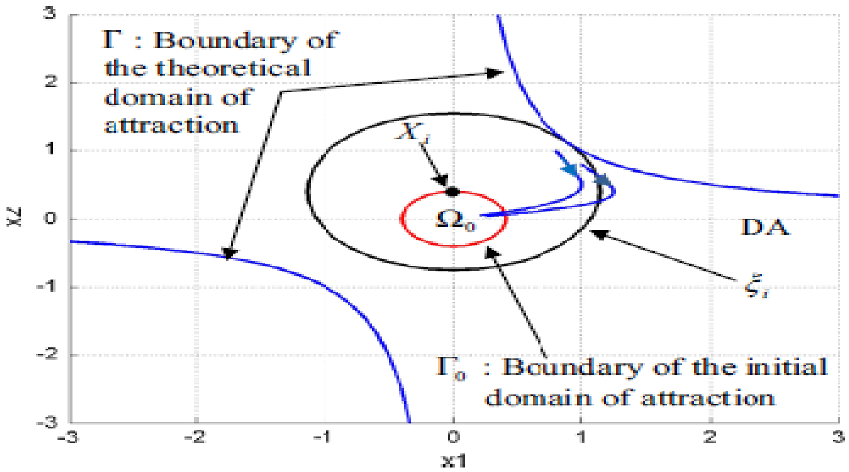


Figure 7: the RA estimated in [3] for Example 3 .

The value of the cost function converts to a constant number and equals zero (see Figures 8 and 9).

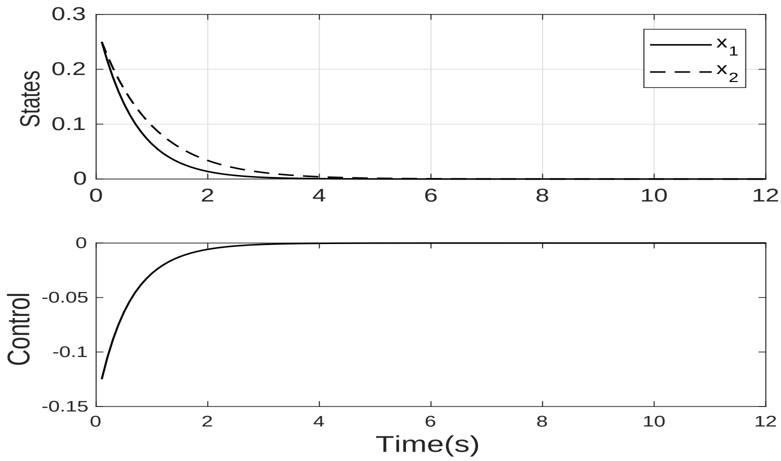


Figure 8: States trajectories and control signal by proposed method for Example 3.

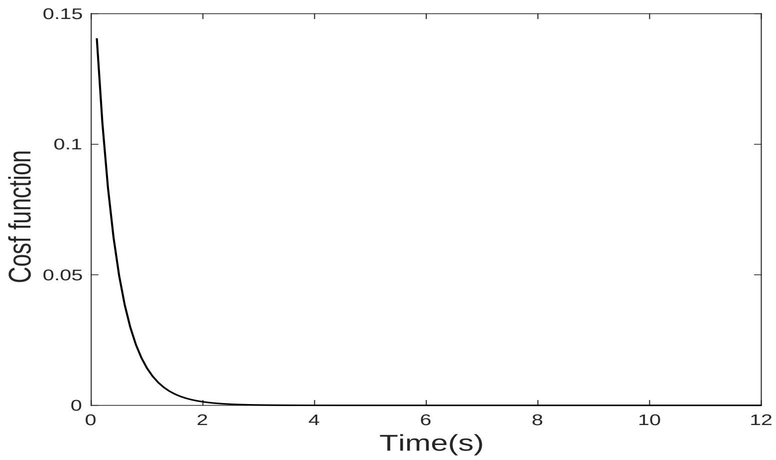


Figure 9: Cost function by proposed method for Example 3.

Figure 9 shows that the cost function of the presented methodology is nonlinear, which indicates the efficiency of the proposed method.

5 Conclusion

In the present research study, a new methodology was developed to estimate the ROA of nonlinear systems. Moreover, in order to enlarge the ROA, a combination of pseudo-linearization and SOS programming methods was employed. By designing a nonlinear state feedback controller through the SDRE, a shape factor was obtained that expanded the ROA. Since the choice of this shape factor was influenced by the dynamics of the system, we had a non-uniform extension of the ROA. The major privilege of the proposed approach was the incorporation of both the polynomial and non-polynomial dynamics. Moreover, the proposed method could deal with both the bound and unbound ROAs. Finally, to support the applicability as well as the superiority of the proposed method, numerical simulations were provided and the results were compared with other studies in the literature.

References

[1] Qiang B., Zhang L. (2018). “Output feedback control design to enlarge the domain of attraction of a super cavitating vehicle subject to actuator saturation”, Transactions of the Institute of Measurement and Control, 40 (10), 3189-3200.

[2] Azizi S., Torres L. A. B., Palhares R. M. (2018). “Regional Robust Stabilization and Domain of Attraction Estimation for MIMO Uncertain Nonlinear Systems with Input Saturation”, International Journal of Control, 91 (1), 215-229.

[3] Jerbi H. (2017). “Estimations of the Domains of Attraction for Classes of Nonlinear Continuous Polynomial Systems”, Arabian Journal for Science and Engineering ,42(7), 2829-2837.

[4] Amato F., Calabrese F. et al. (2011). “Stability analysis of nonlinear quadratic systems via polyhedral Lyapunov finctions”, Automatica, 47, 614-617.

[5] Jarvis-Wloszek Z. W., (2003). “Lyapunov Based Analysis and Controller Synthesis for Polynomial Systems using Sum-of-Squares Optimization”, Berkeley, University of California.

[6] Tan W., Packard A. (2006). “Stability Region Analysis using Sum of Squares Programming”, In American Control Conference, 2297-2302.

[7] Tan W. (2006). “Nonlinear Control Analysis and Synthesis using Sum-of-Squares Programming, Dissertation (Ph.D.)”, University of California, Berkeley.

[8] Rapoport L. B., Morozov Yu. V. (2008). “Numerical methods for estimation of the attraction domain in the problem of control of the wheeled robot”, Autom. Remote Control, 69(1), 13-26.

[9] Khodadadi L., Samadi B., Khaloozadeh H. (2013). “Estimation of a region of attraction for polynomial nonlinear systems: A numerical method”, ISA Transactions, 53, 25-32.

[10] Moghaddam R. Pariz K., N., Shanechi H. M., Kamyad A. V. (2014). “Estimating the robust domain of attraction and directional enlargement of attraction domain via Markov models”, Optim. Control Appl. Meth., 35, 21-40.

[11] Korda M., Henrion D., Jones C. N. (2014). “Controller design and region of attraction estimation for nonlinear dynamical systems”. arXiv: 1310.2213.

[12] Najafi E., Lopes G. A., Babuska R. (2014). “Balancing a legged robot using state-dependent Riccati equation control”, In: Proceedings of the 19th IFAC World Congress, pp. 2177-2182.

[13] Baier R., Gerdts M. (2009). “A computational method for nonconvex reachable sets using optimal control”, In Proceedings of the European Control Conference, 23-26.

[14] Khlebnikov M. V. (2015). “Estimates of the Attraction Domain of Linear Systems under L2-Bounded Control”, Autom. Remote Control, 76 (3), 369-376.

[15] Lim´on D., Alamo T., Camacho E. F. (2005). “Enlarging the domain of attraction of MPC controllers”, Automatica, 41 (4), 629 – 635.

[16] Pesterev A. V. (2017). “Attraction Domain Estimate for Single-Input Affine Systems with Constrained Control”, Automation and Remote Control, 78 (4), 581-594.

[17] Pesterev A. V. (2019). “Estimation of the Attraction Domain for an Affine System with Constrained Vector Control Closed by the Linearizing Feedback”, Automation and Remote Control, 80 (5), 840-855.

[18] Haghighatnia S., Moghaddam R. K. (2012) . “Directional Extension of the Domain of Attraction to Increase Critical clearing time of nonlinear systems”, Journal of American Science, 8 444-449.

[19] Chesi G. (2004). “Design of controllers enlarging the domain of attraction: a quasi-convex approach”, IFAC Symposium on Nonlinear Control Systems, Stuttgart, Germany, September.

[20] Hashemzadeh F., Yazdanpanah M. J. (2006). “Semi-Global Enlargement of Domain of Attraction for a Class of Affine Nonlinear Systems”, In Proceedings of the IEEE International Conference on Control Applications, 1-6.

[21] Chen X. Y., Li C. J., Lu J. F., Jing Y. W. (2012). “The domain of attraction for an SEIR epidemic model based on sum of square optimization”, Bulletin of the Korean Mathematical Society, 49 (3), 517-528.

[22] Li C. , Ryoo C., Li N., Cao L. (2009). “Estimating the domain of attraction via moment matrices”, Bull. Korean Math. Soc. 46 (6), 1237–1248.

[23] Prajna S., Papachristodoulou A., Seiler P., Parrilo P.A. (2004). “SOSTOOLS: Sum of Squares Optimization Toolbox for MATLAB, User’s guide”, Dynamical Systems.

[24] Topcu, U., Packard, A. K., Seiler, P., Balas, G. J. (2010). “Robust region-of attraction estimation”, IEEE Transactions on Automatic Control, 55 (1), 137–142.

[25] Batmani Y., Khaloozadeh H. (2013). “Optimal chemotherapy in cancer treatment: state-dependent Riccati equation control and extended Kalman filter”, Optimal Control, Applications and Methods, 34 (5), 562-577.

[26] Batmani Y., Khaloozadeh H. (2013). “On the design of human immunodeficiency virus treatment based on a non-linear time-delay model”, IET Systems Biology, 10.1049/iet-syb.2013.0012.

[27] Batmani Y., Khaloozadeh H. (2013). “On the design of suboptimal sliding manifold for a class of nonlinear uncertain time-delay systems”, International Journal of Systems Science.

[28] Batmani Y., Khaloozadeh H. (2013). "Optimal drug regimens in cancer chemotherapy: a multi-objective approach", Computers in Biology and Medicine, 43, 2089-2095.

[29] Batmani Y., Khaloozadeh H. (2013). "On the design of observer for nonlinear time-delay systems", Asian Journal of Control, 10.1002/asjc.795.

[30] Çimen T. (2010). "Systematic and effective design of nonlinear feedback controllers via state dependent Riccati equation (SDRE) method", Annual Reviews in control, 34 (1), 3-5.

[31] Çimen T. (2012). "Survey of state-dependent Riccati equation in nonlinear optimal feedback control synthesis", Journal of Guidance, Control, and Dynamics, 35 (4), 1025-1047.

[32] Chen H., Allgöwer F. (1998). "A Quasi-Infinite Horizon Nonlinear Model Predictive Control Scheme with Guaranteed Stability", Automatica, 34 (10), 1205-1217.

[33] Franze G., Casavola A., Famularo D., Garone E. (2008). "An off-line MPC strategy for nonlinear systems based on SOS programming", IFAC Proceedings Volumes, 41 (2), 10063-10068.

[34] Cox D. A., Little J., O'Shea D. (2004). "Using algebraic geometry", Little, Springer Verlag, Berlin.

[35] Najafi E., Babuska R., Lopes G. A. (2016). "A fast sampling method for estimating the domain of attraction", Nonlinear Dyn, 86, 823–834.

Control and Optimization in Applied Mathematics (COAM)
DOI. 10.30473/coam.2020.49118.1127
Vol. 4, No. 2, Autumn - Winter 2019 (39-48), ©2016 Payame Noor University, Iran

Discover the Maximum Descriptive User Groups on the Social Web

Z. Abbasi[1], N. Akhoundi[2*]
[1,2]School of Mathematics and Computer Science, Damghan University, Damghan, Iran,

Received: December 12, 2019; **Accepted:** December 11, 2020.

Abstract. Product reviews in E-commerce websites such as restaurants, movies, E-commerce products, etc., are essential resources for consumers to make purchasing decisions on various items. In this paper, we model discovering groups with maximum descriptively from E-commerce website of the form $< i, u, s >$, where $i \in \mathcal{I}$ (the set of items or products), $u \in \mathcal{U}$ (the set of users) and s is the integer rating that user u has assigned to the item i. Labeled groups from user attributes are found by solving an optimization problem. The performance of the approach is examined by some experiments on real data-sets.

Keywords. Maximum descriptively, Optimization, User group discovery, Rating record.

MSC. 90C26.

* Corresponding author
zahra.abbasi@yahoo.com, akhoundi@du.ac.ir
http://mathco.journals.pnu.ac.ir

1 Introduction

Today, collaborative rating sites drive numerous decisions. For example, online shoppers rely on ratings on E-commerce websites to purchase a variety of goods. Typically, the number of ratings (such as user comments and star rating) associated with an item (a set of items) can easily achieve hundreds or thousands, thus deciding over such a huge amount of data can be cumbersome. Before making an informed decision, a user can either spend lots of time researching dozens of ratings and reviews or can be satisfied only on an average overall rating, associated with an item. There is no surprise that most users choose the second option because of lacking time. For example, Digikala is an E-commerce site that sells a large variety of products from electronic products such as mobiles, notebooks, etc., to apparel accessories. For each item, users can leave their reviews, feedback about the item, their rating, and their experience with that item. In the category browsing page, Digikala shows the number of reviews and average ratings that users assigned to that item (e.g., see the figure 1). Users should read all reviews or trust on average ratings while trying to decide which item is better to buy. Some useful reviewers' attributes such as gender, age, location, occupation, etc., can be useful for making a better decision. Analysis of such data enables innovative insights in various scenarios such as population studies [1], online recommendation [2], and targeted advertisement [3]. In this paper, we propose an algorithm that groups reviewers based on their attribute values, and the result will be shown by some description short sentences such as "Young female students rate this item 4". For this end, we define reviewer groups by users attributes, such as reviewer group $\{\langle$ gender,female$\rangle, \langle$age,young$\rangle\}$. We aim to find reviewers group sets by maximum descriptively for a rating of records. We define the set of items by $\mathcal{I}$, and define the set of users (reviewers)

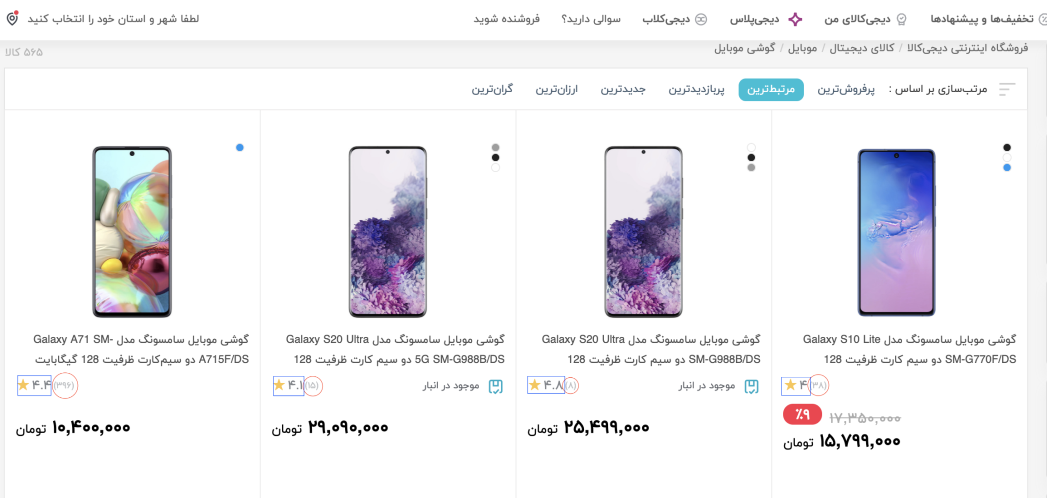


Figure 1: Mobile category browsing page, number of reviews are indicated by red circles, and the total average of ratings are indicated by blue rectangles.

by $\mathcal{U}$. For given datasets of rating records in the form $\langle i, u, s\rangle$, where $i \in \mathcal{I}$ (the set of items

Digikala is an Iranian E-commerce company based in Tehran or products), $u \in \mathcal{U}$ (the set of users) and s is the integer rating that user u has assigned to item i. An user group is defined by conjunction of users' demographic attributes over rating records, such as *male teachers* or *young students who live in Tehran*. The problem of user group discovering is to find group set of G including user groups such that some objectives are optimized. In [5], the problem of user group discovery is modeled as the following constrained optimization:

$$\begin{aligned} \text{Min} \quad & \text{error}(G), \\ \text{S.T.} \quad & \text{covarage}(G) \geq \alpha, \\ & |G| \leq k. \end{aligned} \tag{1}$$

where G is taken over all user group sets. $\text{error}(G)$ is a function that computes the sum of total distance between mean scores in each group and mean scores of rating records. Function $\text{coverage}(G)$ computes the percentage of covering the rating record I, by G. In [4], in order to solve the problem of user group discovery, the constrained multi-objective optimization problem is defined as follows: for a given set of rating records R and integer constants σ and k, the problem is to identify all group-sets, such that each group-set G satisfies:

$$\begin{aligned} \text{Max} \quad & \text{Coverage}(G), \\ \text{Max} \quad & \text{Diversity}(G), \\ \text{Opt.} \quad & \text{rDistb}(G), \\ \text{S.T.} \quad & |G| \leq k, \\ & \forall g \in G : |g| \geq \sigma. \end{aligned} \tag{2}$$

where $\text{Diversity}(G)$ measures how distinct groups are in group-set G. The last constraint states that a group g should contain at least σ rating records, an application-defined threshold. Each group in G is a description of its attributes. For example, if a group G is $G = \{\langle \text{gender,female} \rangle, \langle \text{age,young} \rangle\}$, then G can be described as *young female* group. We would like to focus on discovering user groups with more accurate descriptions. To achieve this aim, our strategy is to find user groups with maximum descriptive attributes. In both optimization problems (1.1) and (1.2), the number of attributes of the user group is not considered. Sometimes, returned groups have their minimum number of attributes (only one attribute). However, groups with more attributes have better descriptions. So, it can be a good idea to find the optimal group with the maximum number of attributes. Maximizing the number attributes (maximum descriptively) should be considered as an objective of the user group discovery model. The rest of paper is organized as follows: Section 2 describes problem definitions. Section 3 encompasses the basic definitions and concepts of user group discovery problem, and our proposed algorithm. Finally some experiments are shown in section 4.

2 Problem Definitions

An E-commerce website like Digikala, comprises three main part. The first part is the set of items (the set of all products) denoted by $\mathcal{I}$. The second part is the set of users, we denote by

$\mathcal{U}$. We denote by $\mathcal{R}$ the set of rating records. It is the third part of E-commerce website. Each rating record $r \in \mathcal{R}$ is itself a triple $\langle i, u, s \rangle$, where $i \in \mathcal{I}$, $u \in \mathcal{U}$ and s is the integer rating that user u has associated to item i. In this paper this three parts are modeled as a triple $\langle \mathcal{I}, \mathcal{U}, \mathcal{R} \rangle$. The set of items $\mathcal{I}$ is associated with a set of attributes, denoted as $\mathcal{I}_A = \{ia_1, ia_2, \ldots\}$, where each item $i \in \mathcal{I}$ is a tuple with $\mathcal{I}_A$ as its schema. In other words, $i = \langle iv_1, iv_2, \ldots \rangle$, where each iv_j is a set of values for attribute ia_j. The schema for the reviewers is $\mathcal{U}_A = \{ua_1, ua_2, \ldots\}$, i.e., $u =< uv_1, uv_2, \ldots > \in \mathcal{U}$, where each uv_j is a value for attribute ua_j. As a result, the tuple for i, the tuple for u, and the numerical rating score s are joint by $r = \langle i, u, s \rangle$ which itself is a tuple in the form $\langle iv_1, iv_2, \ldots, uv_1, uv_2, \ldots, s \rangle$. The set of all attributes is denoted as $A = \{a_1, a_2, \ldots\}$.

Definition 1. We define a group g as a set of $\{\langle a_1, v_1 \rangle, \langle a_2, v_2 \rangle, \cdots\}$ where each $a_i \in A$ (set of all attributes) and each v_i is a set of values for a_i.

The set of attributes g are denoted by $A(g)$, and the number of rating records in g is denoted by $|g|$.

For example, in MovieLens datasets, the group

$$g = \{\langle \text{gender,female} \rangle, \langle \text{location,DC} \rangle, \langle \text{genre,romance} \rangle\}$$

contain rating records for romance movies whose reviewers are all female in DC. We note that $A(g) = \{\text{gender,location,genre}\}$.

Definition 2. Given a rating record $r = \langle v_1, v_2, \ldots, v_k, s \rangle$, where each v_i is a set of values for its corresponding attribute in the schema A, and a group

$$g = \{\langle a_1, v_1 \rangle, \langle a_2, v_2 \rangle, \ldots, \langle a_n, v_n \rangle\}, n \leq k,$$

we say that g covers r, and denote by $r < g$, if and only if $\forall i \in [1, n], \exists r \cdot v_j$ such that $r \cdot v_j$ is a subset of values for attribute $g \cdot a_i$ i.e., $r \cdot v_j \subset g \cdot v_i$.

For example, based on Definition 2, the rating $\langle$female,DC,student,4$\rangle$ is covered by the group $\{\langle$gender,female$\rangle, \langle$location,DC$\rangle\}$. The set of all possible groups form a lattice where nodes correspond to groups and edges correspond to parent/child and ancestor/descendant relationships. In Figure 2 a partial lattice for rating records of the movie *Toy Story*(1995) is illustrated. We have four reviewer attributes **gender, age, location** (CA stands for California) and **occupation** to analysis. For simplicity, exactly one distinct value per attribute is shown. The complete lattice contains 15582 attribute-value combinations, see for example [4].

Definition 3. We say that two groups g_1 and g_2 are similar and denoted by $g_1 \sim g_2$, if and only if $A(g_1) = A(g_2)$.

We have the following two lemmas, proof of them is straightforward, so we omit them.

Lemma 1. Let g_1 and g_2 be two groups, define $B_1 = \{r \in R | r < g_1\}, B_2 = \{r \in R | r < g_2\}$. If $g_2 \subset g_1$, then we have $B_1 \subseteq B_2$.

Lemma 2. Let g_1 and g_2 be two groups and h is some arbitrary group differ from g_1 and g_2. Define $C_1 = \{r \in R | r < h \wedge r < g_1\}$, $C_2 = \{r \in R | r < h \wedge r < g_2\}$. If $g_2 \subset g_1$ then $C_1 \subseteq C_2$.

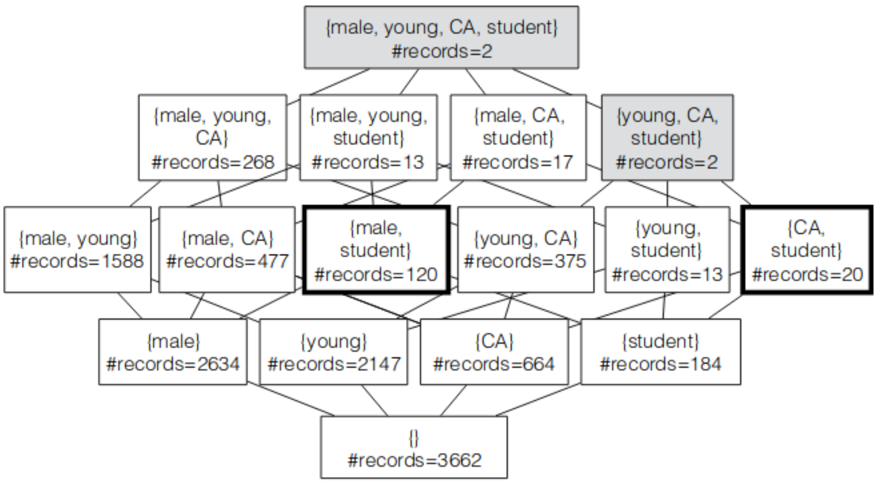


Figure 2: Partial lattice for movie *Toy Story*.

Before formalizing the mining problem, quality dimensions should be defined for groups. For a set of rating records $R \subseteq \mathcal{R}$ and a group-set G, the percentage of rating records in R contained in groups in G is measured by a quality dimension called coverage. Coverage is a value between 0 and 1 and it is defined as follows.

$$\text{coverage}(G,R) = \frac{|\cup_{g\in G}\{r \in R, r < g\}|}{|R|}. \tag{3}$$

Another quality dimension is called diversity. Diversity of G is a value between 0 and 1 that measures how distinct groups in group-set G are from each other, is defined as follows

$$\text{diversity}(G,R) = \frac{1}{1+\sum\limits_{g_1,g_2\in G}|\{r\in R, r<g_1 \wedge r<g_2\}|}. \tag{4}$$

For a group set G, we define the number of attributes as following,

$$\text{attributes}(G) = |\cup_{g\in G} A(g)|, \tag{5}$$

for example number of attributes $G = \{g_1, g_2\}$ where

$$g_1 = \{\langle\text{gender,male}\rangle\}, g_2 = \{\langle\text{gender,male}\rangle, \langle\text{occupation,student}\rangle\},$$

is

$$|\{\text{gender}\} \cup \{\text{gender,occupation}\}| = 2.$$

Group sets that have more attributes provide users with more information to make their decisions.

3 Maximum Description Optimization

We define our constrained optimization problem as follows: for a given set of rating records R, the problem is to identify all group sets, such that each group set satisfies:

$$\begin{array}{ll} \text{Max} & \mathsf{attributes}(G), \\ \text{S.T.} & \mathsf{coverage}(G,R) \geq \alpha \\ & \mathsf{diversity}(G,R) \geq \beta \\ & |G| \geq k. \end{array} \tag{6}$$

Definition 4. Let g is a group and G is a group set, we say $g < G$, if and only if $\forall h \in G, h \nprec g$, and there are $\tilde{g} \in G$ such that $\tilde{g} \subset g$. we denote by $G_g^{-\tilde{g}}$ a group set that was constructed by replacing g with $\tilde{g}$ in G, i.e., $G_g^{-\tilde{g}} = G - \{\tilde{g}\} \cup \{g\}$.

Theorem 1. If a group set G has the following two properties

$$\forall g_1, g_2 \in G, \quad g_1 \nprec g_2, \tag{7}$$

$$\text{There is no two groups,} \quad g_1, g_2 \in G \ \ \text{such that} \ \ g_1 \subset g_2, \tag{8}$$

then for some group g such that $g < G$, the following statements are holds.

1. $\mathsf{attributes}(G_g^{-\tilde{g}}) > \mathsf{attributes}(G)$.
2. $\mathsf{coverage}(G_g^{-\tilde{g}}) \leq \mathsf{coverage}(G)$.
3. $\mathsf{diversity}(G_g^{-\tilde{g}}) \geq \mathsf{diversity}(G)$.

Proof. By definition 4 there exist, $\tilde{g} \in G$ such that $\tilde{g} \subset g$. We have $A(\tilde{g}) \subset A(g)$, hence $\bigcup_{h \in G} A(h) \subset \bigcup_{h \in G_g^{-\tilde{g}}} A(h)$ hence

$$\mathsf{attributes}(G_g^{-\tilde{g}}) = |\bigcup_{h \in G_g^{-\tilde{g}}} A(h)| > |\bigcup_{h \in G} A(h)| = \mathsf{attributes}(G).$$

This is complete the proof of the first part. For section 2, by lemma 1 $\bigcup_{h \in G} \{r \in R | r < h\} \subseteq \bigcup_{h \in G_g^{-\tilde{g}}} \{r \in R | r < h\}$ hence $\mathsf{coverage}(G_g^{-\tilde{g}}) \leq \mathsf{coverage}(G)$. Lastly by lemma 2 we have

$$\sum_{g_1, g_2 \in G_g^{-\tilde{g}}} |\{r \in R, r < g_1 \wedge r < g_2\}| \leq \sum_{g_1, g_2 \in G} |\{r \in R, r < g_1 \wedge r < g_2\}|,$$

hence

$$\begin{aligned} \mathsf{diversity}(G_g^{-\tilde{g}}) &= \frac{1}{1 + \sum_{g_1, g_2 \in G_g^{-\tilde{g}}} |\{r \in R, r < g_1 \wedge r < g_2\}|} \\ &\geq \frac{1}{1 + \sum_{g_1, g_2 \in G} |\{r \in R, r < g_1 \wedge r < g_2\}|} = \mathsf{diversity}(G). \end{aligned}$$

□

Remark 1. The group sets that contain one group with one attribute satisfies in (7) and (8).

Based on theorem 1, we can develop an algorithm with smart search to find local maximum of optimization problem (6). For example in partial lattice of Toy Story movie (see figure 2), we can set $G = \{g_1\}$ as an initial solution, where $g_1 = \{\langle\text{gender,male}\rangle\}$ (g_1 is a group with one attributes and maximum coverage). Three parent nodes of g_1, are as following:

- $h_1 = \{\langle\text{gender,male}\rangle, \langle\text{age,young}\rangle\}$
- $h_2 = \{\langle\text{gender,male}\rangle, \langle\text{location,CA}\rangle\}$
- $h_3 = \{\langle\text{gender,male}\rangle, \langle\text{occuption,student}\rangle\}$.

We see that $h_i < G$ for $i = 1,2,3$. Hence based on Theorem 1, for each h_i where $\text{coverage}(G_{h_i}^{-g_1}) \geq \alpha$, new group set $G_{h_i}^{-g_1}$ is better than G. If there is no such group, then G is a local optimal solution of optimization problem 6, otherwise we can choose $h_k = \arg\max_i\{\text{coverage}(G_{h_i}^{-g_1})\}$ and substitute G with a better solution $G_{h_k}^{-g_1}$.

3.1 Algorithm

The algorithm is started with initial groups that have one attribute. These groups are chosen based on the best coverage. Let the groups by one attribute (penultimate level of the lattice, e.g., see Fig. 2) are ordered as follows

$$\text{coverage}(g_1) \geq \text{coverage}(g_2) \geq \ldots \geq \text{coverage}(g_n). \tag{9}$$

Based on Theorem 1, we search in parent lattice of groups to increase the number of attributes as long as coverage condition is satisfied. Our algorithm is described in details in **Algorithm 1**. The pseudo-code of algorithm 1 works as follows: In line 1, the parameters of α, m are given. In line 3, the initial group sets are generated as defined in (9). In lines 4-15, the search step is performed over the m initial group sets to find the local optimal solution. We know that in group set G, if we let h is a parent of some group $g \in G$, then it is easy to show that $h < G$. The search procedure is based on parent searching because of several reasons. The first reason, since the initial group set G satisfies the requirements of (7) and (8), and h is a parent of $g \in G$, the group set G_h^{-g} satisfies these requirements too. Secondly based on Theorem 1, we see that $attributes(G_h^{-g}) > attributes(G)$, so if $\text{coverage}(G_h^{-g}) \geq \alpha$ (Line 9 algorithm), then G_h^{-g} is better solution than G. Finally, we can describe the parent-based search procedure in detail, as follows. The search procedure was performed for each initial solution in the for-loop in line 4. For each initial solution, until convergence occur, we find a group h in parent lattice-based of $G_i^{(k)}$ in line 8 that satisfies in $\text{coverage}(G_h^{-g}) \geq \alpha$. If there is no such group, then $G_i^{(k)}$ is a local optimal solution, and in line 14 we add it into the local optimal solution set G_{opt}. Otherwise the solution $G_i^{(k)}$ is replaced by the current better solution G_h^{-g}. Finally, in line 16, the best local optimal solution inside G_{opt} is selected and return in line 17.

Algorithm 1 Lattice search algorithm

Data: α, m
$G_{opt} \leftarrow \emptyset$;
Initialization :
For $i = 1, \ldots, m$ Choose initial group sets $G_i^{(0)} = \{g_i\}$ (as defined in (12));
Search Step :
for $i = 1, ..., m$ **do**
  for $k = 0, 1, 2, 3, ...$ **do**
    $G = G_i^{(k)}$;
    opt = true;
    for *$g \in G$ and $\forall h$ in parent lattice-based of g* **do**
      if $coverage(G_h^{-g}) \geq \alpha$ **then**
        `// Replace` $G_i^{(k)}$ `with better group set` G_h^{-g}
        $G_i^{(k+1)} \leftarrow G_h^{-g}$;
        `//` G `isn't local optimal solution`
        opt = false;
        break;
    `// Check if` $G_i^{(k)}$ `is optimal solution`
    if *opt == true* **then**
      $G_{opt}.add(G_i^{(k)}, attributes(G_i^{(k)}))$;
      break;
let $(G', attributes(G'))$ be the pair with maximum number attributes in G_{opt};
return G';

4 Experiments

Real datasets, MovieLens, have been used for our experiments. For each user, gender, age-group, occupation, and zip code are provided. The MovieLens 1M datasets contain 100000 ratings of 3952 movies by 6040 users. The attribute of gender takes two distinct values: male or female. The numeric age is converted into categorical attribute values, namely teen-aged, young, middle-aged, and old. 21 occupations such as student, doctor, lawyer, etc are also listed. Finally, zip codes are converted into the USA states (http://zip.usps.com). Thus, 52 distinct values can be taken for the attribute location [3]. Five items are selected randomly and then, the groups are provided by our algorithm (Table 1) that we assume $\alpha = 0.8, \beta = 0.8, k = m = 2$ and DEM method [5] (Table 2). In Table 1-2 the column **Cov**, **Natt**, and **Div** denote coverage, number attributes, and diversity respectively. The algorithm was written in PHP and Laravel. The algorithm is freely available as a Laravel package in https://github.com/NARooshnavand/user-group-discovery.

Table 1: Our Algorithm

Id	Cov	Natt	Div	Optimal group set
73	0.804	4	1	First group={ young student women in California} Second group={men}
200	0.801	3	1	First group={ young student women} Second group={men}
500	0.806	3	1	First group={ young student women} Second group={men}
600	1	4	1	First group={ old administer men in California} Second group={old educator men in Seattle }
821	0.818	4	1	First group={ old educator men in Texas } Second group={ young }

Table 2: DEM Algorithm

Id	Cov	Natt	Div	Optimal group set
73	0.812	4	1	First group={young women in California} Second group={men}
200	0.908	2	1	First group={men} Second group={young women}
500	1	1	1	First group={men} Second group={women}
600	1	4	1	First group={old administer men in California} Second group={old educator men in Seattle}
821	0.812	3	1	First group={middle-aged men} Second group={young}

References

[1] Amer-Yahia, S., Kleisarchaki, S., Kolloju, N. K., Lakshmanan, L. V., Zamar, R. H. (2017). “Exploring rated datasets with rating maps”, In Proceedings of the 26th International Conference on World Wide Web (pp. 1411-1419).

[2] Amer-Yahia, S., Omidvar-Tehrani, B., Basu, S., Shabib, N. (2015). “Group recommendation with temporal affinities”, EDBT.

[3] Omidvar-Tehrani, B., Amer-Yahia, S., Termier, A. (2015). ”Interactive user group analysis”, In Proceedings of the 24th ACM International on Conference on Information and Knowledge Management (pp. 403-412).

[4] Omidvar-Tehrani B., Amer-Yahia S., Dutot P.F., Trystram D. (2016). “Multi-Objective Group Discovery on the Social Web” ECML PKDD 2016, part I,LNAI 9851, 296-312

[5] Das, M., Amer-Yahia, S., Das, G., Yu, C. (2011). "Mri: Meaningful interpretations of collaborative ratings", Proceedings of the VLDB Endowment, Vol. 4, No. 11.

Control and Optimization in Applied Mathematics (COAM)
DOI. 10.30473/coam.2020.55039.1148
Vol. 4, No. 2, Autumn - Winter 2019 (49-60), ©2016 Payame Noor University, Iran

Analysis of Students' Mistakes in Solving Integrals to Minimize their Mistakes

M. Darvishzadeh[1], A. Shahvarani Semnani[2*], H. Alamolhodaei[3], M. H. Behzadi[4]

[1,2]Department of Mathematics, Science and Research Branch, Islamic Azad University, Tehran, Iran,
[3]Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran,
[4]Department of Statistics, Science and Research Branch, Islamic Azad University, Tehran, Iran,

Received: September 13, 2020; **Accepted:** December 11, 2020.

Abstract. Experiences of teaching Integral have indicated that the vast majorities of Iranian university students commit numerous errors while solving integral problems and have weak skills in this field; we might even say that they hide away from integral and consider it the nightmare of mathematics. On the other hand, Integral is the base of pure and applied mathematics for all students of science, especially engineering, which some of their lessons are dependent on it directly or indirectly, so it is important to pay attention to it. Through descriptive method-exposed factor, an exam has been conducted in the form of three questions, the first of which is consisted of 4 sections on fifty students from different fields, and then interviews were conducted with a few of those students about their answers in order to study the students' behaviors when solving integral problems and to determine the type of their errors. By analyzing the performance of students in this test, we can see that students often struggle with integral and mostly have a feeble performance in solving trigonometric integrals. They want to learn computational integral instead of how to conceptualize integral in their minds correctly. The error most committed by university students was procedural errors, which arise from using derivative instead of integral. Besides most of the mistakes happen in solving definite integrals, and calculating finite areas between two curves. This is due to a lack of understanding of integrals and a lack of information in other areas of mathematics.

Keywords. Calculus, Integral, Student understanding, Understanding mathematical education, Initial function concept, Riemann concept, Teaching, Challenge.

MSC. 97C70.

---

* Corresponding author
darvish.math@gmail.com, shahvarani1393@gmail.com,
alamolhodaei@yahoo.com, alamolhodaei@yahoo.com
http://mathco.journals.pnu.ac.ir

1 Introduction

The mathematical concepts, according to Dane, Faruk Cetin, Ba & Özturan Saçgırulilı (2016), have complex, abstract, and hierarchical levels. In formal education, we face an increase in these aforementioned levels of mathematical concepts and also the class level(Cetin, Dane & Bekdemir (2012)). Also, The fragmentation of the students' thinking structure had been shown by Adi Wibawa, Nusantara, Subanji & Parta (2017) in solving the problems of the application of the definite integral in area. For all students and especially university students, the integral and its related concepts are one of the important, fundamental, and necessary topics in learning basic mathematics. On the one hand, the integral is an Irrefutable tool in solving applied problems for university students, especially in the majors of basic sciences and engineering (Dancis, 2001). The mathematics education system in Iran is planned in such a way that in the last years of high school and the first year of university and after learning the concept of derivative and differential, students get acquainted with the concept of integral with an overemphasis on a symbolic form of the primary function (anti-derivative), in contrast, they should be first introduced to integral with the conceptualization of perimeter and area symbolic form, the Riemann symbolic form or by the adding up pieces symbolic form. Therefore, considering the approach taken by teachers and professors, the ability of students to solve routing integral problems is far greater than their ability to solve practical issues, and even these routine integral problems are solved with numerous procedural errors. Now, we look to analyze these problems and identify their roots and offer solutions where it is possible. Experience has shown that Iranian students learn integral superficially and like a parrot; their goal is to learn how to calculate an integral. However, the teaching method of professors is also useful in this regard. According to the above issues, several studies have been conducted worldwide regarding the issues students face in integral problems; one of the best studies has been done by Seah Eng Kia (2005), also Avital & Libeskind (1978), Ronaldson (1963), Chou (2002), and Everton (1983) have done some research and presented their results for their own countries and with different samples; but unfortunately, nothing fundamental has been done in this regard in Iran. With a thorough study of these researches, it can be said that students have a fundamental problem in conceptualizing integrals using Riemann sum or the sum of infinite rectangles. For example, Thomas & Yi (1996) sought to examine the misconceptions of learners when confronted with the Riemann integral, and they found that learners do not have a precise understanding of the Riemann integral and use an algorithm-based learning process to solve problems. We are now looking to localize the work of previous people, especially S.E.Kiat, in Iran and implement it following the Iranian educational system and reference books. Therefore, we have modified the test he has designed with a little reduction and change in order to analyze the students' problems and students' mistakes for solving integrals. In order to achieve this goal, we implement the conceptual framework of S.E.Kiat, which itself was developed from the interrogative framework of Ronaldson (1963), Libeskind (1978), Everton (1983). This developed interrogative framework is given in Appendix (I).

2 Research Methodology and Participants

Students participating in this research have selected from 280 engineering students who study at Islamic Azad University of Mashhad, such that at first, these students were trained for two semesters in pre-university mathematics and general Math I and II, (their training in these lessons based on the proposed exam include all details and preparations). All have participated in two midterm and final exams for each lesson, and after reviewing the results of these exams, and examining their high school education records, fifty students were selected who the selected students are among the average students in the field of mathematics. The present study was conducted after the passing of these students in General Mathematics 2 and was carried out in the second semester of the 2018-19 academic years in the Islamic Azad University of Mashhad such that fifty students majoring in engineering and sciences (physics and chemistry) and of which ten students are electrical students, ten students are civil engineering students, ten students are mechanical students, and the rest are physics and chemistry students. There were ten students in each field and a total of 30 male and 20 female students. All these fifty people have been trained with the same teaching method in the calculus 1 and 2 and have an average score of 13 in the calculus 1 and 13.5 in the calculus 2 (Scores in Iran's universities are calculated from 20). They have also taken introductory physics courses. Since the research follows the responses from the number of students in terms of the number and type of errors that are made in integral problems. In order to study the above topics carefully, we based our research on the observation and through a descriptive method-exposed factor of this fifty students who are currently studying the major. On the other hand, we have proposed an exam with eight problems in the first phase. In the second phase, eight problems are studied in terms of content validity by six professors of pure mathematics and mathematics education in the Islamic Azad University of Mashhad. With regard to views and content validity indexes, three questions are accepted finally (see Appendix II). Note that the purpose of each question is different, and some questions are designed for different purposes. Nevertheless, it can be said that the main primary purpose of the questions is to examine the level of students' capability in solving integrals, which are based on the primary function and Riemann or infinite sum of rectangles. Other minor goals include the ability of students to solve in solving integrals that have different powers of x. (Objective I), which is also pursued in the second and third questions, and another goal is to study the ability of students in solving integrals of functions as $(a+bx)^n$, as well as trigonometric and exponential functions (Objective II) and the study of students' ability in using certain integrals to calculate the area of the next objective (Objective III). We note that while pursuing the above goals, the basic necessary skills of students are also examined. Also, all problems have taken from Apostol (1967). The reliability of all six problems is proved on the same samples of students then its reliability is determined that was more than 0.75. Participants were asked to present their responses along with the complete explanation of their responses' details in (written manners).

3 Data Collection

The above test consists of two stages. One is a written test in which all students have participated, and the other is an interview with some of them based on their written answers. The students interviewed often fall into one of three categories:

- Students who have given complete answers to the questions.
- Students who have performed poorly in solving questions.
- Students who have an excellent educational background but had a disappointing performance in answering questions.

Each session held in presence of four persons that one of them is a student participating in the interview and three interviewers who are experienced and successful faculty members in the department of mathematics, Islamic Azad University of Mashhad. It is important to mention that the students' oral responses to the problems were much longer than what would come, but we tried to convey their meaning as much as possible. The interview has done through questions that proposed by one of the authors about solving integral's. The interview was purposeful, i.e. the interviewer presented questions according to their possible responses to get nearer to the purpose of the interview. The following items were considered in the test:

- The test was taken simultaneously from all 50 people.
- Before the exam, students are informed to have enough time to study and prepare for the exam.
- The duration of the written test is 1 hour.
- The interviews were filmed for accuracy in and later analysis.
- Before the interview, each student is allowed to think about their answers.
- An attempt has been made to conduct a fully structured interview, i.e., the questions have been designed according to the written answers of students and their assumed answer to the interviewer's questions.

4 Results and Related Discussion

A summary of the written test results is given in Table 1.
Note that the score of each question is calculated according to the types of error that students may make for solving it, which is listed in Table 1. By studying the above table, we see that 20 students have solved question 1a entirely and accurately, and 10 have solved more than half of it. In other words, 60% of students answered question 1a, so it can be said that students can solve integrals of the functions $(a+bx)^n$ even though they have difficulty writing the details. Also, regarding the integration of trigonometric functions that are included in the first question of part 1bi and 1bii, only 32 people were able to solve it, among which only 20 people solved

Table 1: Test results summary

Percentage pass	Number of students who scored (in%)			S.D.	Main	Total mark	Q
	100	50-99	0-49				
60%	20	10	20	1.92	2.66	5	1a
44%	15	7	28	1.48	2.52	5	1bi
20%	5	5	40	1.28	1.8	5	1bii
50%	15	10	25	1.59	2.95	5	1biii
46%	5	18	27	2.6	3.72	7	2
20%	3	7	40	1.8	1.64	7	3

these two questions accurately, and 12 people were able to solve more than half of it. In other words, in these two questions, they performed much worse than in question 1a. In fact, by examining their answers, we find that in integrating trigonometric functions, many students have used derivatives instead of integrals (the most common procedural mistake), which is due to the lack of accurate and complete conceptualization of integrals and their incomplete understanding of functions; and that is why some students have written $2\sin(2x-1)$ in response to question 1bi. In question 1bii, they have not been able to use the trigonometric formulas to convert the function under integral and then integrate. They have made numerous mistakes in solving it, and therefore have the lowest efficiency among the questions. Only five people have been able to solve this question completely. On the other hand, we note that in this question, how to perform calculations is also essential, and students have answered this part of the question almost acceptably. In question 1biii, in which the integration of exponential functions is included and aims to examine the performance of students, we see that only 15 people were able to thoroughly answer it and 10 people managed to solve more than half of it. They had a more satisfying performance compared to questions 1bi and 1bii. We should note that some people have made some procedural severe mistakes (using the derivative instead of the integral). In general, and by analyzing the first question, we can say that the performance of students in integrating functions of $(a+bx)^n$ is much better than integrating exponential functions, and their performance in integrating exponential functions is much better than that of trigonometric functions and they make fewer mistakes. However, there have been interrogative and technical errors in doing all four sections of this question. In the second and third questions, students' skills in solving definite integrals and mainly calculating the area are considered. According to the given function, integration of different powers x is also considered (Objective I). By studying the table, we see that due to the closeness of these two questions, the way students have answered them is so different; such that students answering the third question are less than half of the students who were able to answer the second question, i.e., $42\%, 20\%$. Only 10% of students were able to answer the second question thoroughly, and only 6% of them managed to answer the third question accurately. In fact, in the second question, the analysis of the answers shows that they did not understand the need to draw a graph of the function and even marking it, and went straight to integrating the function (x^2-4x) at a distance of $x=0$ to $x=5$ and have not recognized that part of the function graph is above and a part of it is below the x-axis. Regarding the third question, they had to calculate the ratio of the two hatched areas, and again, there is a part above and a part below the x-axis. Even the diagram of this question is presented, and yet again, students had many problems. They had the weakest efficiency in this question, such that some were not able to calculate the area under the x-axis, some had difficulty in determining the integral boundaries, and some did not know from which function they should integrate, which all three refer to fundamental weaknesses in the conceptual understanding of integrals. In general, it can be said that students perform much better in problems that directly require integral computation than problems in which the application of integrals is considered. In fact, in performing direct definite and indefinite integrals, they have a relatively better ability than using integrals to calculate the area. Nevertheless, the exciting thing is that they easily integrate different x functions.

5 Error Analysis

The types of mistakes made by students in solving the test questions is shown in Table 2. Examining Table 2, we see that the most technical errors occurred in the first question, However,

Table 2: Type and quantity of errors made by students

ZERO	No Errors	Technical Errors	Procedural Errors	Conceptual Errors	Q
0	20	7	20	3	1a
0	15	6	24	5	1bi
1	5	30	11	4	1bii
3	15	6	22	4	1biii
0	5	4	8	32	2
1	3	13	5	28	3
5	63	66	90	76	Total

the interrogative(Conceptual) errors occurred in the second and third questions, and there are procedural errors in all the questions. Column zero in Table 2 also means that students did not answer the question. The table above shows that students committed 76 interrogative errors, 90 procedural errors, and 66 technical errors. For example, the first interrogative error occurred in question 1a, where students had to calculate the integral of function $2(3+4x)^4$, and three persons made this error by using 3 in the answer instead of 5, for example. Hirad replied that:

$$\int 2\,(3+4x)^4 dx = \frac{2}{4}(3+4x)^3 + C,$$

Hirad combined the derivative and the integral and provided this answer, but since only three people made such a mistake, it can be said that most of them understood this question. Table 2 also shows that most of the interrogative errors occurred in the second and third questions. Even though the graph of the function and the regions are presented, most of the students have had problems calculating the area under the x-axis and have used either the wrong bounds or the wrong functions to solve these questions. For example, Negin wrote this in response to the third question:

$$
\begin{aligned}
S_A &= \int_0^2 \left(x^2 - 6x + 8\right) dx + \int_2^4 (8 - 2x)\, dx \\
&= \left(\frac{x^3}{3} - 3x^2 + 8x\right)_0^2 - \left(8x - x^2\right) \\
&= \left(\frac{64}{3} - 12 + 32\right) + (16 - 12) = \cdots, \\
S_B &= \int_2^4 \left(x^2 - 6x + 8\right) dx = \left(\frac{x^3}{3} - 3x^2 + 8x\right)_2^4 \\
&= \left(\frac{64}{3} - 12 + 32\right) - \left(\frac{8}{3} - 12 + 16\right) = \cdots .
\end{aligned}
$$

When we ask Negin why she used these two integrals to calculate zone A, she points to the area between the diagonal line and the curve and says that we have to calculate this part, and therefore, we have to take integrals from these two functions. Hence, this mistake is due to the lack of precise and complete conceptualization of definite integral in her mind. Even though we continue to guide her on how to calculate the area of zone A, she insists on her answer. Hooman also wrote this in response to this question:

$$
\begin{aligned}
S_A &= \int_0^4 (8 + 2x)\, dx = \left(8x + x^2\right)_0^4 = 32 + 16 = 48 \\
S_B &= \int_2^4 \left(x^2 - 6x + 8\right) dx = \left(\frac{x^3}{3} - 3x^2 + 8x\right)_2^4 = \cdots .
\end{aligned}
$$

Hooman's answer in calculating area B is the same as Negin's answer. Both made the same mistake in not understanding that the function negative and under the x-axis, but in calculating area A, Hooman acted very differently and ignored the curve function; when discussing with him how to calculate the area of A, the extreme lack of conceptualization of the integral is evident in his mind, whereas we note that if only the linear function $y = 8 + 2x$ was involved in calculating the area of A, he could have easily calculated it. Hooman: To calculate the area A, it is sufficient to integrate the function $y + 2x = 8$ at a distance of 0 to 4. (points to the hatched part between the line and the x-axis.) Interviewer: But how do we calculate the part between the function $y + 2x = 8$ and the curve $y^2 - 6x + 8$: Hooman? After about 3 minutes of thought and silence, Hooman only states that it is no different from before and that we must integrate directly from the linear function. Mahshid's answer to the second question is as follows:

$$
S = \int_0^5 \left(x^2 - 4x\right) dx = \left(\frac{x^3}{3} - 2x^2\right)_0^5 = \frac{125}{3} - 50 = \cdots .
$$

Mahshid did not understand that part of the $x^2 - 4x$ function is below the x-axis and that she has to integrate it twice, and when we ask her the reason for this action, she says: "There is no need to mark the function or even draw it, and I did not think at all that such a thing could happen, but now I remember a similar problem that I solved before and I should have done this." We find that if Mahshid had marked the integral, she could have solved the question entirely. Table 2 tells us that the highest number of errors was the procedural error, which is 90 times. For example, Negar wrote this in response to question 1bi:

$$\int \mathrm{Cos}\,(2x-1)\,dx = \frac{1}{2}\mathrm{Sin}(2x-1).$$

Besides Negar, 20 other people have made the same mistake and have not written the constant value of C, and in response, they only say that they forgot; in fact, it can be said that they do not have an exact concept of this constant C in their minds that will help them not forget. In the same question, Hamin has written this:

$$\int \mathrm{Cos}\,(2x-1)\,dx = -2\mathrm{Sin}\,(2x-1)\ \ +C.$$

Although she wrote C, she has used the derivative instead of the integral (this is the most common type of mistake) Negin has given the following answer to question 1biii:

$$\int e^{(2x+3)}dx = 2e^{2x+3}.$$

Here, in addition to using the derivative instead of the integral, Negin did not write down the constant C and made two mistakes at the same time. and In response to why she used 2 instead of 1/2, she immediately admitted her mistake and said that she had mistaken it for the derivative. Examining all the answers, we find that students often have problems in the process of calculating integrations in a direct way, which depends on the conceptualization of integrals based on the primary forms of the functions such as trigonometric or exponential ones. Hence, and they use the derivative instead of the integral. Regarding the third type of error, i.e., technical errors that have occurred 66 times in the all test, we can say that the main reason is the lack of sufficient information in other mathematical fields such as algebra, geometry, trigonometry, or even carelessness. These errors have nothing to do with the correct and accurate conceptualization of integrals by students. Whether the conceptualization is complete or not, these errors may occur, while it can be said that students with fewer conceptual and procedural errors certainly make fewer technical errors. However, because technical errors make it impossible to answer the questions related to the integral fully, they are worth reviewing. Hesam has written this in response to question 1a:

$$\int 2(3+4x)^4dx = 2\int \left(3^4+(4x)^4\right)dx = 2\left(3^4x+\ \frac{4}{5}x^5\right)+C.$$

insufficient knowledge of algebra has caused this error. In response to question 1biii and due to weakness in trigonometry and lack of sufficient information in this field, Kimia did not know that $\mathrm{Sec}^2\ x = 1+\mathrm{Tan}^2x$ and used $1+\mathrm{Sec}^2\ x = \mathrm{Tan}^2x$ and has written the following:

$$\int_0^{\frac{\pi}{2}} \mathrm{Tan}^2\ 2x\ dx =\ \int_0^{\frac{\pi}{2}} \left(1+\mathrm{Sec}^2\,2x\right)dx = (x+\tan 2x)_0^{\frac{\pi}{2}} = \frac{\pi}{2}.$$

(Note that there is also a procedural error in this question) In response to the second question, Zohreh did not pay enough attention as well and has written the following:

$$S = \int_0^5 \left(x^2-4x\right)dx = \frac{x^3}{3}+4x^2 = \frac{125}{3}+100.$$

She also forgot to divide 4 by 2. When we ask students about these issues, they immediately realize their mistakes and correct them. Finally, Table 3 provides a list of all the mistakes

Table 3: Types of mistakes

Errors	Description	Numbers
Conceptual) (Interrogative)	Combining derivatives and integrals	**3**
	Not recognizing the need to mark the function	**29**
	Failure to select the correct function	**25**
	Failure to select the correct bounds	**19**
Procedural (Executive)	Failure to write the constant C	**24**
	Mistake between derivative and integral	**66**
Technical	Coordinate geometry	**18**
	Computational Algebra	**16**
	Trigonometry	**15**
	Inaccuracy	**17**

students made during the test and the interview that followed. Examining Table 3, we find that most errors are procedural ones. One of the most important reasons for this error is the use of derivatives instead of integrals, which can be related to lack of experience in this field, which itself is because of not having enough practice. The type of error lease made by the students was a technical error, which means that if students were taught to conceptualize integrals more wholly and accurately, they would be more able to solve problems. Finally, we point out that although students are among the average students at the university under study and even though they were given enough time to study before the exam, they are still weak in the field of integral and have much work to do.

6 Discussion

Considering the conditions and questions of the exam and the interviews conducted and also considering that only 50 people have participated in this exam and they have also been selected from average students, we cannot say that this exam is a complete, accurate and perfect test and therefore performing similar tests in this subject is always recommended. We also need to look at how integrals should be taught and how students should learn them. However, it can be said that in general, students have a fundamental problem in understanding integrals. They should be re-taught the basic concepts related to it, and its prerequisites, especially the derivative, before the subject of integral is brought up. Although the fact that teaching hours in Iran are limited in these subjects, and it is not possible to emphasize the prerequisite topics should be considered in this regard, one solution can be to hold extracurricular or problem-solving classes in this field because basic concepts like algebra, geometry, and trigonometry are important concepts for understanding integrals correctly and the large number of technical errors made by students shows the importance of these concepts. On the other hand, teaching mathematics in Iran, especially in the first year of university, is done without the necessary and sufficient emphasis on drawing diagrams and various graphs of functions and this subject is discussed in a maximum of 2 hours and is very vital and essential for learning the concept of integral and determining and calculating different areas of different functions, and is taught to strengthen students' visual perspectives so they would not make many mistakes in solving relevant integrals. In other words, useful emphasis on visual explanations by teachers and strengthening students' visual perspectives facilitates the interrogative comprehension of integrals. However, if all these subjects are completed, the test will show us that students have difficulty in conceptualizing and understanding integrals, the many procedural errors made prove this, and in other words, their interrogative understanding of integrals is incomplete, especially since in Iranian universities, and especially for engineering students, integrals are often conceptualized through the primary function, and the emphasis is on calculating the integral rather than understanding its concept, so terrible mistakes have occurred in calculating different areas and even in choosing the appropriate bounds for the integral or in the selection of the appropriate function to calculate the area using integral. Therefore, teaching methods and how to teach students is effective in reducing students' mistakes. For example, although the function diagram for the desired area is given in the third question, students still have many interrogative and procedural problems in solving the problem. Therefore, we emphasize that it is better to conceptualize the integral with the help of Riemann's concept or the summation of infinite areas. With the help of the primary function, it was found in the interviews that many students, unfortunately, read mathematics to write, so we emphasize the importance of strengthening the communication skills of writing compared to the communication skills of reading in students. One suggested solution is to write an example of calculating a finite area for students and force them to provide a written answer, and then ask them to discuss their answers so that they can understand the integral. Writing mathematics greatly reduces a variety of errors, primarily procedural and technical ones. Note, however, that the most massive largest error made by students participating in the test is procedural errors that are caused by mistakes in using integrals and derivatives, which requires much emphasis when teaching this subject. The main suggestion is to use derivatives and integrals parallelly so they will become engraved in the minds of students. Finally, although integral is one of the most essential concepts in mathematics for students, especially in the field of engineering, on one hand, many of them are dealing with the fundamental difficulties with this concept to understand. On the other hand, they do not have any outstanding effort to learn it (why? Of course, there are many reasons for this issue that are beyond the scope of this research article).

Appendix I: Classification of students' errors

Appendix II: Test on Integration

First question:

a) Solve this integral: $\int 2(3+4x)^4 dx$.

b) Evaluate the following integrals:

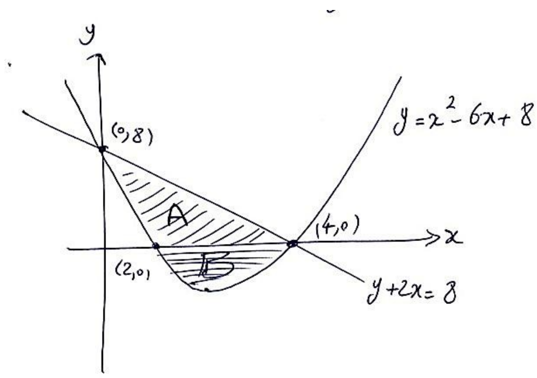
Error type:	Description of the error with an example
Conceptual error: Errors due to inability of learners to understand the relations in the question or due to their inability to understand its meaning.	Calculate the area between the curve y $= x^2 - 4x$ and the x-axis from $x = 0$ to $x = 5$ $\int_0^5 \left(x^2 - 4x\right) dx = \left(\frac{x^3}{3} - 2x^2\right)_0^5 = \frac{-25}{3}$ They did not realize that part of the curve is below the x-axis.
Procedural error: Error due to inability to perform calculations or algorithms despite learners' understanding of the concepts in the question.	$\int \tan^2 2x dx = \int (\text{Sec}^2 2x - 1) dx$ $\tan^2 2x - x + C$ Not writing $\frac{1}{2}$ behind the $\tan^2 2x$ function
Technical error: Error due to carelessness or lack of sufficient information on other issues.	$\int 2(3+4x)^4 d = \int (6+8x)^4 dx$ $= \frac{(6+8x)^5}{5\times 8} + C$ Wrong multiplication of $(3+4x)$

i) $\int \cos(2x-1) dx$,

ii) $\int_0^{\pi/2} \tan^2 2x dx$,

iii) $\int e^{(2x+3)} dx$.

Second question: Find the area bounded by the curve $y = x^2 - 4x$ and the x-axis from $x = 0$ to $x = 5$. **Third question:** The following diagram shows a part of the line $y + 2x = 8$ and the curve of the function $y = x^2 - 6x + 8$. Find the ratio of the area of zone A to the area of zone B.



References

[1] Adi Wibawa K., Nusantara T., Subanji Parta N. (2017). "Fragmentation of Thinking Structure's Students to Solving the Problem of Application Definite Integral in Area", International Education Studies, 10(5), 48-60.

[2] Allen D. (2001). "Learning integral calculus tee rough non-template", Problem-solving, XI(2), 147-160.

[3] Apostol T.M. (1967). "Calculus. Xerox Corporation", Waltham: Massachusetts.

[4] Baki M. Cekmez E. (2012). "Prospective elementary mathematics teachers' understandings about the formal definition of limit", Turkish Journal of Computer and Mathematics Education, 3(2), 81-98.

[5] Biber A. ,. Argün Z. (2015). "The relations between concept knowledge related to the limits concepts in one and two variables functions of mathematics teachers candidates", Bartin University Journal of Faculty of Education, 4(2), 501-515.

[6] Cetin F.Ö., Dane A., Bekdemir M. (2012). "A concept of accumulation point and its usage", Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education, 6(2), 217-233.

[7] Dan A., Faruk Cetin O., Ba F., Özturan Saçgırulılı , M. (2016). "A Conceptual and Procedural Research on the Hierarchical Structure of Mathematics Emerging in the Minds of University Students: An Example of Limit-Continuity-Integral-Derivative", International Journal of Higher Education, 5(2), 82-91.

[8] Dancis J.(2001). "Student modeling of physical phenomena as they derive standard integral formulae: One way to reduce proof phobia", PRIMUS, XI(2), 124-146.

[9] Kula S., Bukova Güzel E. (2015). "Reflections of Mathematics Student Teachers' Knowledge Related to the Purposes of the Curriculum on Their Limit Teaching", Journal of Theoretical Educational Science, 8(1), 28-49.

[10] Jonse S.R. (2013). "Understanding the integral: Students symbolic forms", Journal of mathematical behavior, 32, 122-141.

[11] Kiat S.E. (2005) "Analysis of student difficulties in solving integration problems", 9(1), 39-59.

[12] Orton A. (1983). "Student's understanding of integration", Educational studies in mathematics, 14, 1-18.

[13] Özkaya M., Isık A., Konyalıoçglu A. (2014). "Proofing and counter-exampling performances of students in the elementary mathematics education department for continuous functions", Middle Eastern & African Journal of Educational Research, 11, 26-42.

[14] Thomas M., & Hong Y.Y. (1996). "The Riemann Integral in Calculus: Student's Processes and Concepts, In P. C. Clarkson (Ed.), Proceeding of the 19^{th} Annual Conference of the Mathematics Education Research Group of Australia", Melbourne, Victoria, Australia.

[15] Ubuz B. (2011). "Factors associated with success in a calculus course: An examination of personal variables", International journal of mathematical education in science and technology, 42(1), 1-12.

Control and Optimization in Applied Mathematics (COAM)
DOI. 10.30473/coam.2021.55089.1149
Vol. 4, No. 2, Autumn - Winter 2019 (61-67), ©2016 Payame Noor University, Iran

Weakly Perfect Graphs of Modules

M. Nouri Jouybari[1], Y. Talebi[2], S. Firouzian[3]*

[1,2]Department of Mathematics, University of Mazandaran, Babolsar, Iran
[3]Department of Mathematics, Payame Noor University (PNU),
P.O. Box. 19395-3697, Tehran, Iran.

Received: September 13, 2020; **Accepted:** January 28, 2021.

Abstract. In this study, R and M are assumed to be a commutative ring with non-zero identity M and an R-module, respectively. Scalar Product Graph of M, denoted by $G_R(M)$, is a graph with the vertex-set M and two different vertices a and b in M are connected if and only if there exists r belong to R such that $a = rb$ or $b = ra$. This paper studies some properties of such weakly perfect graphs.

Keywords. Scalar Product, Graph join, Weakly Perfect, Module.

MSC. 05C25, 13CXX.

* Corresponding author
mostafa.umz@gmail.com, talebi@umz.ac.ir, siamfirouzian@pnu.ac.ir.
http://mathco.journals.pnu.ac.ir

1 Introduction

Throughout this paper, let G be a simple graph with a set of vertices $V = V(G)$ and a set of edges $E = E(G)$. Let $u_1, u_j \in V(G)$. If u_i is adjacent to u_j, we have $u_i \sim u_j$. The k-coloring of G is an appropriate form of k colors to $V(G)$ such that no two adjacent vertices have the same color. The smallest number k with this property, denoted by $\chi(G)$, is called the chromatic number of G. A clique of G is a complete subgraph of G and the cardinality of the largest clique of G is called the clique number of G and denoted by $\omega(G)$. A graph G is called weakly perfect if $\chi(G) = \omega(G)$. Maimani and et al. [4] introduced a class of such graphs. Nikandish and et al. [5] presented a graph of ideals that are weakly perfect. The graph is perfect if every induced subgraph is weakly perfect. Hence, every perfect graph is weakly perfect and there are several classes that indicate that the converse may not hold in general. Fander [3] introduced a new class of perfect graphs.
Let $S \subseteq V(G)$. Herein, S is an independent set if the maximum degree of the subgraph induced by $V(G)$ is zero. Independent number, denoted by $\alpha(G)$, is the maximum cardinality of any independent set. It is trivial that vertex S is a clique of G if and only if it is an independent set of $\bar{G}$. Thus, $\alpha(G) = \omega(\bar{G})$.
A topological index is a numerical quantity that is invariant under automorphisms of the graph. The topological index based on the distance function was first used by H. Wiener [7]. If $u, v \in V(G)$ are two different vertices, then $d(u, v)$ is the length of the shortest path between u and v. Therefore, the Wiener index of G is:

$$W(G) = \frac{1}{2} \sum_{u,v \in V(G)} d(u, v) \tag{1}$$

Suppose that R is a commutative ring with identity and $W(R)$ is a set of non-unit elements of R. Afkhami et al. [1] defined the Cozero-divisor graph of R, denoted by $\Gamma'(R)$, with vertices $W(R)^* = W(R) \setminus \{0\}$ and $x, y \in W(R)^* (x \neq y)$; then, $x \sim y$ if and only if $x \notin Ry$ and $y \notin Rx$ where Rc is an ideal generated by $c \in R$. Suppose that M is an R-module and $W_R(M) = \{x \in M \mid Rm \neq M\}$. With R as R-module, $W_R(R)$ is a set of all non-unit elements of R. Alibemani et al. [2] introduced Cozero-divisor graphs in relation to R-module M in which vertices are $W_R(M)^* = W_R(M) \setminus \{0\}$ and $m, n \in W_R(M)^* (m \neq n)$ and then, $m \sim n$ if and only if $m \notin Rn$ and $n \notin Rm$. The mentioned authors studied the properties of this graph.
The next section introduces a new class of graphs arising from weakly perfect modules. Moreover, a formula is presented for $\chi(G)$ and $\omega(G)$ of such graphs. In Section 3, The Wiener index of such graphs is calculated.

2 Weakly Perfect Graphs of Modules

This section defines a scalar product graph of the module and shows that it is weakly perfect in some cases. The definition of the join of two graphs needs to be noted here. Suppose that X and Y are two separate graphs. $X+Y$ is join of X and Y with a set of vertices $V(X+Y) = V(X) \cup V(Y)$

and a set of edges $E(X+Y) = E(X) \cup E(Y) \cup \{xy : x \in V(X), y \in V(Y)\}$.
In addition, $|V(X+Y)| = |V(X)| + |V(Y)|$ and $|E(X+Y)| = |E(X)| + |E(Y)| + |V(X)||V(Y)|$.
Also, for two graphs X and Y we have $\chi(X+Y) = \chi(X) + \chi(Y)$.

Lemma 1. Let G and H be separate graphs. Then $\alpha(G+H) = max\{\alpha(G), \alpha(H)\}$ and $\omega(G+H) = \omega(G) + \omega(H)$.

Proof. Suppose that $max\{\alpha(G), \alpha(H)\} = \alpha(G)$ and $S = \{u_1, u_2, \ldots, u_{\alpha(G)}\}$ is the maximum independent number of G. For any (u_i, u_j), $1 \leq i, j \leq \alpha(G)$, $i \neq j$, and the edge $u_i u_j$ is not in $E(G)$; thus, $u_i u_j \notin E(G+H)$. It is implied that S is an independent set of $G+H$. Indeed, $\alpha(G+H) \geq max\{\alpha(G), \alpha(H)\}$. Now, for the converse, suppose that S' is the maximum independent number and the sum of graphs G and H. Then, S' is not the subset of $V(G)$ and $V(H)$ contemporary. Suppose that $S' \subset V(G)$ and therefore, $\alpha(G+H) \leq \alpha(G)$ and $\alpha(G+H) \leq \alpha(H)$. Hence, $\alpha(G+H) \leq max\{\alpha(G), \alpha(H)\}$. Suppose that C is an arbitrary clique of $G+H$. It can be assumed that $C = C_1 \cup C_2$ in which $C_1 \subseteq V(G)$ and $C_2 \subseteq V(H)$. It is quite trivial that $|C1| \leq \omega(G)$ and $|C2| \leq \omega(H)$. Therefore, $\omega(G+H) \leq \omega(G) + \omega(H)$. Thus, We have $\omega(G+H) \geq \omega(G) + \omega(H)$. □

Definition 1. [6] Suppose that R is a commutative ring with non-zero identity and M be an R-module. We define the Scalar-product graph of R-module M, namely $G_R(M)$, in which the vertices of $G_R(M)$ are elements of M and $x, y \in M (x \neq y)$ then, $x \sim y$ is adjacent if and only if there exists r belonging to R such that $x = ry$ or $y = rx$.

Remark 1. Let $G_R(M)$ be a Scalar-product graph of R-module M. If $x, y \in M$ then x is adjacent to y if and only if $Rx \subseteq Ry$ or $Ry \subseteq Rx$.

Remark 2. According to the definition of the cozero-divisor graph over modules, we have the followings:
(1) If M is an R-module, the subgraph of $G_R(M)$ in which vertices are $W_R(M)^*$ is the complement of the cozero-divisors graph of M.
(2) We have $G_R(M) = G_1 + G_2$ where G_1 is a complete graph with $|W_R(M)^*|$ vertices and G_2 is the complement of the the cozero-divisor graph of M.

In the following, if $G_R(M)$ is the scalar product graph of some R- module M, we compute $\chi(G_R(M))$ and $\omega(G_R(M))$.

Lemma 2. Suppose that M is an R-module. Then, the scalar product graph $G_R(M)$ is complete if and only if the cyclic submodules of M are linearly ordered by inclusion relation.

Proof. Let M be an R-module and $N_1 = < a >, N_2 = < b >$ be two cyclic submodules of M in which $a \neq b$ in M. Since the scalar product graph $G_R(M)$ is complete, a and b are adjacent. We have $< a > \subseteq < b >$ or $< b > \subseteq < a >$ and $N_1 \subseteq N_2$ or $N_2 \subseteq N_1$. Conversely, Let M be an R-module in which the cyclic submodules are linearly ordered by inclusion relation. If $a \neq b$ represents two vertices of $G_R(M)$ then $< a > \subseteq < b >$ or $< b > \subseteq < a >$. Therefore, a and b are adjacent in $G_R(M)$. Hence, $G_R(M)$ is complete. □

Suppose that M is R-module and A, B are two non-zero submodules of M.Then, M is called uniserial if $A \subseteq B$ or $B \subseteq A$. Clearly, the valuation ring R is uniserial as a module over itself. Also, submodules and quotients of uniserial modules are again uniserial.

Lemma 3. Let $\mathbb{Z}_n$ be a $\mathbb{Z}$-module. If p, m are prime and positive integer numbers, then for $n = 1, p, p^m$, the scalar product graph $G_{\mathbb{Z}}(\mathbb{Z}_n)$ will be complete.

Proof. Let M be a simple module. Then, submodules of M are linearly ordered by inclusion. Hence, submodules of $\mathbb{Z}_p$ are uniserial. Through Lemma 2.5, the scalar product graph of $\mathbb{Z}_p$ is complete.
Also, $\mathbb{Z}_{p^n} = \frac{1\mathbb{Z}}{p^n\mathbb{Z}} \supset \frac{p\mathbb{Z}}{p^n\mathbb{Z}} \supset \frac{p^2\mathbb{Z}}{p^n\mathbb{Z}} \supset \ldots \supset \frac{p^{n-1}\mathbb{Z}}{p^n\mathbb{Z}} \supset \frac{p^n\mathbb{Z}}{p^n\mathbb{Z}} = 0$, here $\mathbb{Z}_{p^n}$ is uniserial. Therefore its scalar product graph is complete. □

Theorem 1. Suppose that p is a prime number. Then, the edge number of $G_{\mathbb{Z}}(\mathbb{Z}_{2p})$ is $2p^2 - 2p + 1$.

Proof. In Remark 2.5, we have $G_{\mathbb{Z}}(\mathbb{Z}_{2p}) = K_p + G_2$ such that K_p is a complete graph with p vertices and G_2 is the complement of the cozero-divisor graph of $\mathbb{Z}_{2p}$ which is $\overline{K_{1,p-1}}$. By definition 2.1, we have:
$| E(G_{\mathbb{Z}}(\mathbb{Z}_{2p})) |= \frac{p(p-1)}{2} + \frac{(p-1)(p-2)}{2} + p^2 = 2p^2 - 2p + 1$ □

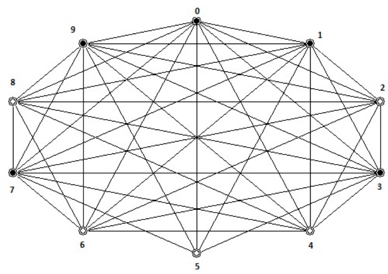


Figure 1: Scalar Product of $\mathbb{Z}$-module $\mathbb{Z}_{10}$.

Theorem 2. Let $\mathbb{Z}_n$ be a $\mathbb{Z}$-module. If $n = 1, p, p^m$ and $n = 2p$, then the graph $G_{\mathbb{Z}}(\mathbb{Z}_n)$ is weakly perfect. Also, if $n = 2p$, we have $\chi(G_{\mathbb{Z}}(\mathbb{Z}_n)) = \omega(G_{\mathbb{Z}}(\mathbb{Z}_n)) = 2p - 1$.

Proof. By Lemma 2.7, $G_{\mathbb{Z}}(\mathbb{Z}_n)$ is a complete graph with n vertices. Hence, It is weakly perfect. If $n = 2p$, then by Remark 2.5, we have $G_{\mathbb{Z}}(\mathbb{Z}_{2p}) = K_p + G_2$ such that K_p is a complete graph with p vertices and G_2 is the complement of cozero-divisor graph of $\mathbb{Z}_{2p}$ which is $\overline{K_{1,p-1}}$. Also, $\chi(K_p) = \omega(K_p) = p$ and $\chi(G_2) = \omega(G_2) = p - 1$. Therefore, by Lemma 2.2, we have $\chi(G_{\mathbb{Z}}(\mathbb{Z}_{2p})) = \omega(G_{\mathbb{Z}}(\mathbb{Z}_{2p})) = 2p - 1$. □

Table 1 show clique, chromatic and edge number of the scalar-product graph of $\mathbb{Z}_{2p}$:

Theorem 3. Suppose that p is a prime number. Then, the edge number of $G_{\mathbb{Z}}(\mathbb{Z}_{3p})$ is $\frac{9}{2}p^2 - \frac{7}{2}p + 2$.

Proof. By Remark 2.5, we have $G_{\mathbb{Z}}(\mathbb{Z}_{3p}) = K_{2p-1} + G_3$ such that K_{2p-1} is a complete graph with p vertices and G_3 is the complement of cozero-divisor graph of $\mathbb{Z}_{3p}$ which is $\overline{K_{2,p-1}}$. By Definition 2.1, we have:
$| E(G_{\mathbb{Z}}(\mathbb{Z}_{3p})) |= \frac{(2p-1)(2p-2)}{2} + 1 + \frac{(p-1)(p-2)}{2} + (2p-1).(p+1) = \frac{9}{2}p^2 - \frac{7}{2}p + 2$ □

Table 1: clique, chromatic and edge number of $G_{\mathbb{Z}}(\mathbb{Z}_{2p})$

p	$\chi(G)$	$\omega(G)$	$\mid E(G) \mid$
3	5	5	13
5	9	9	41
7	13	13	85
11	21	21	145

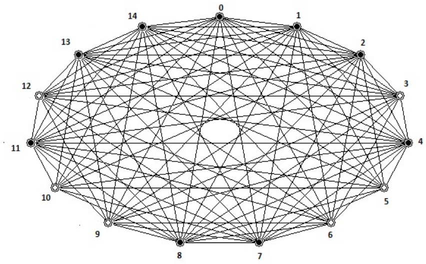


Figure 2: Scalar Product of $\mathbb{Z}$-module $\mathbb{Z}_{15}$.

Theorem 4. Let $\mathbb{Z}_n$ be a $\mathbb{Z}$-module. If $n = 3p$, then the graph $G_{\mathbb{Z}}(\mathbb{Z}_n)$ is weakly perfect. Also $\chi(G_{\mathbb{Z}}(\mathbb{Z}_n)) = \omega(G_{\mathbb{Z}}(\mathbb{Z}_n)) = 3p - 2$.

Proof. If $n = 3p$, then by Remark 2.5, we have $G_{\mathbb{Z}}(\mathbb{Z}_{3p}) = K_{2p-1} + G_3$ where K_{2p-1} is a complete graph with $2p-1$ vertices and G_3 is the complement of the cozero-divisor graph of $\mathbb{Z}_{3p}$ which is $\overline{K_{2,p-1}}$. Also, $\chi(K_{2p-1}) = \omega(K_{2p-1}) = 2p - 1$ and $\chi(G_3) = \omega(G_3) = p - 1$. Therefore, by Lemma 2.2, we have $\chi(G_{\mathbb{Z}}(\mathbb{Z}_{3p})) = \omega(G_{\mathbb{Z}}(\mathbb{Z}_{3p})) = 3p - 2$. □

Table 2 shows the clique, chromatic and edge number of the scalar-product graph of $\mathbb{Z}_{3p}$:

Table 2: clique, chromatic, and edge number of $G_{\mathbb{Z}}(\mathbb{Z}_{3p})$

p	$\chi(G)$	$\omega(G)$	$\mid E(G) \mid$
5	13	13	97
7	19	19	198
11	31	31	508
13	37	37	717

Theorem 5. Suppose that p is a prime number. Then, the edge number of $G_{\mathbb{Z}}(\mathbb{Z}_{5p})$ is $\frac{25}{2}p^2 - \frac{13}{2}p + 4$.

Proof. By Remark 2.5, we have $G_{\mathbb{Z}}(\mathbb{Z}_{5p}) = K_{4p-3} + G_5$ such that K_{4p-3} is a complete graph with $4p-3$ vertices and G_3 is the complement of the cozero-divisor graph of $\mathbb{Z}_{5p}$ which is $\overline{K_{4,p-1}}$. By Definition 2.1, we have:
$\mid E(G_{\mathbb{Z}}(\mathbb{Z}_{5p})) \mid = \frac{(4p-3)(4p-4)}{2} + 6 + \frac{(p-1)(p-2)}{2} + (4p-3).(p+3) = \frac{25}{2}p^2 - \frac{13}{2}p + 4.$ □

Theorem 6. Let $\mathbb{Z}_n$ be a $\mathbb{Z}$-module. If $n = 5p$, then the graph $G_{\mathbb{Z}}(\mathbb{Z}_n)$ is weakly perfect. Also, $\chi(G_{\mathbb{Z}}(\mathbb{Z}_n)) = \omega(G_{\mathbb{Z}}(\mathbb{Z}_n)) = 5p-4$.

Proof. If $n = 5p$, then by Remark 2.5, we have $G_{\mathbb{Z}}(\mathbb{Z}_{5p}) = K_{4p-3} + G_5$ such that K_{4p-3} is a complete graph with $4p-3$ vertices and G_5 is the complement of the cozero-divisor graph of $\mathbb{Z}_{5p}$ which is $\overline{K_{4,p-1}}$. Also, $\chi(K_{4p-3}) = \omega(K_{4p-3}) = 4p-3$ and $\chi(G_5) = \omega(G_5) = p-1$. Therefore, by Lemma 2.2, we have $\chi(G_{\mathbb{Z}}(\mathbb{Z}_{5p})) = \omega(G_{\mathbb{Z}}(\mathbb{Z}_{5p})) = 5p-4$. □

3 Wiener Index of $G_R(M)$

Suppose that G is a graph. The Wiener index of G is half of the sum of the distance between two distinct vertices. For example, we have $W(K_n) = \frac{1}{2}n(n-1)$ and $W(K_{1,n-1}) = (n-1)^2$. This section computes Wiener indices of $G_{\mathbb{Z}}(\mathbb{Z}_{2p})$ and $G_{\mathbb{Z}}(\mathbb{Z}_{3p})$ for some prime p. Similar to what we had before, the Scalar product graphs of $\mathbb{Z}$-module $\mathbb{Z}_{2p}$ and $\mathbb{Z}_{3p}$ are the join of complete graph and complement of a cozero-divisor graph. Therefore, we seek a formula for the Wiener index of the join of two graphs.

Theorem 7. [8] For any two graphs X_1 and X_2, we have:

$$\begin{aligned} W(X_1+X_2) = &\mid V(X_1)\mid^2 - \mid V(X_1)\mid + \mid V(X_2)\mid^2 - \mid V(X_2)\mid \\ &+ \mid V(X_1)\mid\mid V(X_2)\mid - \mid E(X_1)\mid - \mid E(X_2)\mid. \end{aligned}$$

Now, we have the following propositions.

Proposition 1. Suppose that p is a prime number. Then, we have $W(G_{\mathbb{Z}}(\mathbb{Z}_{2p})) = 2p^2-1$.

Proof. By Proof 1, the scalar product graph of $\mathbb{Z}_{2p}$ is the join of K_p and $\overline{K_{1,p-1}}$. Thus, from Theorem 7, we have

$$\begin{aligned} W(K_p+\overline{K_{1,p-1}}) = &\mid V(K_p)\mid^2 - \mid V(K_p)\mid + \mid V(\overline{K_{1,p-1}})\mid^2 - \mid V(\overline{K_{1,p-1}})\mid \\ &+ \mid V(K_p)\mid\mid V(\overline{K_{1,p-1}})\mid - \mid E(K_p)\mid - \mid E(\overline{K_{1,p-1}})\mid \\ =&p^2-p+p^2-p+p^2-\frac{1}{2}p(p-1)-\frac{1}{2}(p-1)(p-2) \\ =&2p^2-1. \end{aligned}$$

□

Proposition 2. Suppose that p is a prime number. Then, we have $W(G_{\mathbb{Z}}(\mathbb{Z}_{3p})) = \frac{9}{2}p^2+\frac{1}{2}p-2$.

Proof. By Proof 3, the scalar product graph of $\mathbb{Z}_{3p}$ is the join of K_{2p-1} and $\overline{K_{2,p-1}}$. Thus, according to Theorem 7, we have:

$$\begin{aligned} W(K_{2p-1}+\overline{K_{2,p-1}}) = &\mid V(K_{2p-1})\mid^2 - \mid V(K_{2p-1})\mid + \mid V(\overline{K_{2,p-1}})\mid^2 - \mid V(\overline{K_{2,p-1}})\mid \\ &+ \mid V(K_{2p-1})\mid\mid V(\overline{K_{2,p-1}})\mid - \mid E(K_{2p-1})\mid - \mid E(\overline{K_{2,p-1}})\mid \\ =&(2p-1)^2-(2p-1)+(p+1)^2-(p+1) \end{aligned}$$

$$
\begin{aligned}
&+(2p-1)(p+1)-\frac{1}{2}(2p-1)(2p-2)-[1+\frac{1}{2}(p-1)(p-2)] \\
&=\frac{9}{2}p^2+\frac{1}{2}p-2.
\end{aligned}
$$

□

Proposition 3. Suppose that p is a prime number. Then, we can have $W(G_{\mathbb{Z}}(\mathbb{Z}_{5p}))=\frac{25}{2}p^2+\frac{3}{2}p-4$.

Proof. By Proof 5, the scalar product graph of $\mathbb{Z}_{5p}$ is the join of K_{4p-3} and $\overline{K_{4,p-1}}$. Thus, from Theorem 7, we have:

$$
\begin{aligned}
W(K_{4p-3}+\overline{K_{4,p-1}}) =& |V(K_{4p-3})|^2-|V(K_{4p-3})|+|V(\overline{K_{4,p-1}})|^2-|V(\overline{K_{4,p-1}})| \\
&+|V(K_{4p-3})||V(\overline{K_{4,p-1}})|-|E(K_{4p-3})|-|E(\overline{K_{4,p-1}})| \\
=&(4p-3)^2-(4p-3)+(p+3)^2-(p+3) \\
&+(4p-3)(p+3)-\frac{1}{2}(4p-3)(4p-4)-[6+\frac{1}{2}(p-1)(p-2)] \\
=&\frac{25}{2}p^2+\frac{3}{2}p-4.
\end{aligned}
$$

□

References

[1] Afkhami M., Khashyarmanesh K. (2011). "The cozero-divisor graph of a commutative ring", Southeast Asian Bulletin of Mathematics, 35, 753-762.

[2] Alibemani A., Hashemi E., Alhevaz A. (2018). "The cozero-divisor graph of a module", Asian-European Journal of Mathematics, 11, DOI: 10.1142/S1793557118500924.

[3] Fander M.R. (2015). "Chromatic and clique number of a class of perfect graph", Transactions on Combinatorics, 4, 1-4.

[4] Maimani H.R., Pournaki M.R., Yassemi S. (2010). "A class of weakly perfect graphs", Czechoslovak Mathematical Journal, 60, 1037-1041.

[5] Nikandish R., Maimani H.R., Izanloo H. (2016). "The annihilating-ideal graph of Z_n is weakly perfect", Contributions to Discrete Mathematics 11, 16-21.

[6] Nouri Jouybari M., Talebi Y., Firouzian S. (2019). "Scalar Product Graphs of Modules", International Journal of Nonlinear Analysis and Applications, 10, Special Issue (Nonlinear Analysis in Engineering and Sciences), 75-82.

[7] Wiener H. (1947). "Structural Determination of Paraffin Boiling Points", Journal American Chemistry Society, 69, 17-20.

[8] Yeh Y., Gutman I. (1994). "On the some of all Distances in Composite Graphs", Discrete Mathematics, 135, 359-365.

Control and Optimization in Applied Mathematics (COAM)
DOI. 10.30473/coam.2021.50933.1134
Vol. 4, No. 2, Autumn - Winter 2019 (69-80), ©2016 Payame Noor University, Iran

Model Predictive Control for a 3D Pendulum on $SO(3)$ Manifold Using Convex Optimization

S. Mansourinasab[1], M. Sojoodi[2*], S. R. Moghadasi[3]

[1,2] Control Engineering Department, Electrical and Computer Engineering Faculty, Tarbiat Modares University, Tehran, Iran,

[3]Department of Mathematical Science, Sharif University of Technology, Tehran, Iran.

Received: January 25, 2020; **Accepted:** January 28, 2021.

Abstract. Conventional model predictive control (MPC) methods are usually implemented to systems with discrete-time dynamics laying on smooth vector space $\mathbf{R}^n$. In contrast, the configuration space of the majority of mechanical systems is not expressed as Euclidean space. Therefore, the MPC method in this paper has developed on a smooth manifold as the configuration space of the attitude control of a 3D pendulum. The Lie Group Variational Integrator (LGVI) equations of motion of the 3D pendulum have been considered as the discrete-time update equations since the LGVI equations preserve the group structure and conserve quantities of motion. The MPC algorithm is applied to the linearized dynamics of the 3D pendulum according to its LGVI equations around the equilibrium using diffeomorphism. Also, as in standard MPC algorithms, convex optimization is solved at each iteration to compute the control law. In this paper, the linear matrix inequality (LMI) is used to solve the convex optimization problem under constraints. A numerical example illustrates the design procedure.

Keywords. Model predictive control, Convex Optimization, Linear matrix inequality, Lie group variational integrator.

MSC. 90C34; 90C40.

* Corresponding author
sara.mansouri@modares.ac.ir, sojoodi@modares.ac.ir, moghadasi@math.sharif.edu
http://mathco.journals.pnu.ac.ir

1 Introduction

Model Predictive Control (MPC) is based on predicting the behavior of a system using its dynamical model and optimizing the prediction to have the best decision. Due to this reason, the dynamic models of the system play an essential role in solving MPC problems. A control system is considered as a family of vector fields, and the dynamical system is the flow generated by this vector field [1, 2]. Differential geometry is a new language of Lagrangian and Hamiltonian Mechanics. In differential geometry, the state space of the system is modeled as manifolds, which are locally diffeomorphic with Euclidean space. As a result, this method is a coordinate-free method that applies to infinite-dimensional systems. Since mechanical systems are symmetric, the states of the system do not change under a certain transformation; this criterion is expressed by Lie group actions [3]. A Lie group that is a smooth manifold with a group structure, is a mathematical concept appropriate for describing continuously varying groups of transformation [3]. In [4], geometric mechanics of rigid bodies on a Lie group is expressed based on the Euler/Lagrange equation of mechanical systems that are developed according to Hamilton's principle. A so-called Lie Group Variational Integrator (LGVI) method has been produced for systems with a Lie group configuration space. The main target of LGVI is implementing an exponential map representing the variation of a curve on a Lie group in terms of Lie algebra element. This method is developed to acquire the discrete-time dynamic equations of the system that preserves the Lie group structure. The main result of this method is that the achieved update discrete-time equations are coordinate-free, namely, there is no need to choose a specific local coordinate. This totally avoids ambiguity and singularity associated with local coordinates [4, 5]. Considering the LGVI method to model the discrete-time update equations of motions, preserves the conserved quantities of motion and therefore provides a more realistic prediction model. As other standard integrating methods such as Runge-Kutta do not use the group structure so that they are deprived of this property. The LGVI method updates the rotational matrix by multiplying two matrices in $SO(3)$, which guarantees the rotational matrix still remains on $SO(3)$ and preserves the conservative motion. [7, 6] provide other types of variational integrator methods. The conventional MPC methods are usually applied to systems with discrete dynamics on $\mathbf{R}^n$ vector space. However, the configuration space of the majority of systems is smooth manifolds which are not diffeomorphic to $\mathbf{R}^n$. For designing the predictive dynamics of such systems, the manifolds with limited dimensions are embedded in $\mathbf{R}^n$, then standard integrating methods are implemented until the discrete updating equations are achieved. Different methods of integrating system dynamics on manifolds have been developed, see [7, 8, 9] for example. Development of the model predictive control design for dynamics evolving on smooth manifolds is considered in [10, 11]. The method of linearizing and embedding the system in $\mathbf{R}^n$ has been used in these papers. Implementing the MPC-based LMI approach is a technique for controlling plants with uncertainties. Since the optimization-based LMI method can be solved in polynomial time, it is applicable to implement it in on-line optimization problems [12]. Solving an optimization problem at each sampling time within a receding horizon is the main contribution of MPC algorithms so that the development of optimization methods improves the ability to solve MPC problems. In such issues, an optimization problem that is efficiently solvable via linear matrix inequality (LMI) can be extended to MPC. In this case, a min-max optimization is solved, which computes the control law by minimizing a quadratic cost function subject to constraints in worst-case at each time step. The issue of solving a min-max optimization can be considered as a convex optimization with linear matrix inequality. Besides, using the LMI optimization scheme with MPC at each time instant can incorporate uncertainties as input and output constraints, and guarantees the robustness properties of the system at the same time. Since Lie group variational integrator is one of the rigid body computational methods that maintain Lagrangian/Hamilton structures as well as the structure of rigid body configuration group, in this paper, the rigid body dynamics are implemented considering its exact geometric properties using the LGVI method. As a result, the classical model predictive control is generalized to the LGVI model of the system. The proposed method of applying convex optimization for solving MPC problems using geometric considerations is applied to a 3D pendulum, which is a rigid body supported at a frictionless pivot acting under the influence of uniform gravity with substantial invariant properties [13]. The novelty of this paper is using the LMI approach for solving the MPC control of the 3D pendulum with a variational model. This paper is organized as follows. Section II is devoted to the problem statement. Firstly, the dynamics of a 3D pendulum using LGVI are expressed, and the linearized state-space model of 3D pendulum dynamics is extracted using infinitesimal variations of parameters evolving on a manifold. Then, based on this linearized model of the system, a quadratic objective function is introduced. Section III extends the standard MPC problem on Euclidean state space to smooth manifold based on convex optimization using LMI. Simulation results are presented in section IV to prove the efficiency of using LMI in solving MPC algorithms on smooth manifolds. A comparison with the non-LMI method is also mentioned in this section. Finally, concluding remarks are presented in section V.

2 Problem Statement

2.1 3D Pendulum Dynamics

The configuration space in a 3D pendulum is a SO(3) manifold. Geometric forms of Hamilton's equations of a 3D pendulum on the configuration manifold SO(3) using LGVI method have been expressed in [4] as discrete-time forced Hamilton's equation as follows

$$\hat{\Pi}_k = \frac{1}{h}(F_k J_d - J_d F_k^T), \tag{1}$$

$$\Pi_{k+1} = F_k^T \Pi_k + h\mathcal{M}_{k+1} + hBu_{k+1}, \tag{2}$$

$$R_{k+1} = R_k F_k, \tag{3}$$

where $R_k \in \mathrm{SO}(3)$ is a rotation matrix from the body-fixed frame to the initial frame denotes the attitude of the rigid body at time k, $\Pi_k \in \mathbf{R}^3$ is the angular momentum of the pendulum expressed in the body-fixed frame, $F_k \in \mathrm{SO}(3)$ is a one-step change in R_k, J_d is a non-standard moment of inertia matrix and $J_d = \frac{1}{2}trace(J)I - J$ where J is the standard inertia matrix. $"h"$

is the time step for the discrete system. In a 3D pendulum $\mathcal{M}_k = mg\rho \times R_k^T e_3$. The hat map denotes $\hat{\cdot} : \mathbf{R} \rightarrow \mathfrak{so}(3)$ that for a given vector $\omega = \begin{pmatrix} \omega_1, \omega_2, \omega_3 \end{pmatrix}^T$ represents the skew-symmetric matrix

$$\hat{\omega} = \begin{pmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{pmatrix}.$$

As in Eucliedaian spaces, a linear vector field can estimate the Hamiltonian vector field on TSO(3) locally in an open subset of TSO(3), which is the tangent space to SO(3). Using local coordinates in the neighboring of the equilibrium point is a method of linearizing the vector field. To extract the linearized discrete Hamilton's equation for 3D pendulum as discussed in [14], a local exponential coordinate is introduced as local coordinates. The variation of the rotational matrix R is an ϵ-parameterized differentiable curve $R_{k,\epsilon}$ that takes value in SO(3), is given by [4]

$$R_{k,\epsilon} = R_k\, exp(\epsilon\, \hat{\eta}_k). \tag{4}$$

The variation of matrix R is expressed as an exponential map that returns the variations across the rotational axis η with angle ϵ. $\eta(t)$ is a differentiable curve that has value on Lie groups of rotational matrices and is identity in t_0 and t_f. The exp map is a local diffeomorphism between Lie algebra and Lie group. Other parameters' variations are also formulated as

$$F_{k,\epsilon} = F_k\, exp(\epsilon\, \hat{\xi}_k),$$

$$\Pi_{k,\epsilon} = \Pi_k + \epsilon\, \delta\Pi_k,$$

where $\delta\Pi_k$ is an infinitesimal variation of Π_k. Infinitesimal variations of the motion can be shown to be

$$\delta R_k = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} R_{k,\epsilon} = R_k\, \hat{\eta}_k,$$

$$\delta F_k = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} F_{k,\epsilon} = F_k\, \hat{\xi}_k,$$

$$\delta \Pi_k = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \Pi_{k,\epsilon} = \delta\Pi_k.$$

The infinitesimal variation of δR_{k+1} can be expressed from two points of view. On the one hand, the variation of R_{k+1} is calculated as infinitesimal variations of its parameters R_k, F_k as follows

$$\delta R_{k+1} = \delta R_k F_k + R_k \delta F_k = R_k\, \hat{\eta}_k\, F_k + R_k F_k \hat{\xi}_k\,, \tag{5}$$

on the other hand, its infinitesimal variation is calculated directly as

$$\begin{aligned} \delta R_{k+1} &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} R_{k+1,\epsilon} = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} R_{k+1} exp(\epsilon\, \hat{\eta}_{k+1}) \\ &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} R_k F_k\, exp(\epsilon\, \hat{\eta}_{k+1}) = R_k F_k \hat{\eta}_{k+1}\,. \end{aligned} \tag{6}$$

Then, the last parts of both equations (5),(6) can be equated, and the parameter η_{k+1} can be extracted as

$$\eta_{k+1} = F_k^T \eta_k + \xi_k. \tag{7}$$

Note: The relation $R^T \hat{x} R = \widehat{R^T x}$ is used. In fact, (7) is the constrained variation of the equation $R_{k+1} = R_k F_k$. Since the linearized system should be in terms of the state variables $\left[\ \eta_k,\ \delta\Pi_k\ \right]^T$, ξ_k should be replaced in terms of η_k and $\delta\Pi_k$. Due to this reason, ξ_k is calculated from (1). Firstly, the variations of (1) are obtained as follows, since $\delta\Pi_k, \delta F_k$ are not independent

$$\begin{aligned}
\delta\hat{\Pi}_k &= \frac{1}{h}(\delta F_k J_d - J_d \delta F_k^T) \\
&= \frac{1}{h}(F_k \hat{\xi}_k J_d + J_d \hat{\xi}_k F_k^T) \\
&= \frac{1}{h}(\widehat{F_k \xi_k} F_k J_d + J_d F_k^T \widehat{F_k \xi_k}).
\end{aligned}$$

Note: $\hat{x}A + A^T \hat{x} = \widehat{(trace[A] I_{3\times 3} - A)x)}$

$$\begin{aligned}
\delta\hat{\Pi}_k &= \frac{1}{h}\widehat{((trace(F_k J_d) I_{3\times 3} - F_k J_d) F_k \xi_k)}, \\
\delta\Pi_k &= \frac{1}{h}(trace(F_k J_d) I_{3\times 3} - F_k J_d) F_k \xi_k.
\end{aligned}$$

As a result,

$$\xi_k = \beta_k \delta\Pi_k, \tag{8}$$

where

$$\beta_k = h F_k^T (trace(F_k J_d) I_{3\times 3} - F_k J_d)^{-1} \in \mathbf{R}^{3\times 3}.$$

By replacing (8) in (7) η_{k+1} is extrapolated as

$$\eta_{k+1} = F_k^T \eta_k + \beta_k \delta\Pi_k, \tag{9}$$

which gives the linearized rotation matrix R in terms of its rotation axis η. The dynamical equation (2) is linearized by substituting the variations of M_k. Since the torque M_k is related to the attitude of a rigid body, its variation δM_k is written as a variation of the rotational matrix

$$\delta M_k = \mathcal{M}_k \eta_k,$$

while $M_k \in \mathbf{R}^{3\times 3}$ is expressed as the attitude of the rigid body and is attained by the potential field. Using (8), (9), variations of M_{k+1} is equal to

$$\delta M_{k+1} = \mathcal{M}_{k+1} \eta_{k+1} = \mathcal{M}_{k+1} F_k^T \eta_k + \mathcal{M}_{k+1} \beta_k \delta\Pi_k.$$

As a result,

$$\begin{aligned}
\delta\Pi_{k+1} =& \delta F_k^T \Pi_k + F_k^T \delta\Pi_k + h\delta\mathcal{M}_{k+1} + hB\delta u_{k+1} \\
=& -\xi_k^T F_k^T \Pi_k + F_k^T \delta\Pi_k + h\mathcal{M}_{k+1} F_k^T \eta_k \\
& + h\mathcal{M}_{k+1} \beta_k \delta\Pi_k + hB\delta u_k \\
=& -(\beta_k \delta\Pi_k\hat{)} F_k^T \Pi_k + F_k^T \delta\Pi_k + h\mathcal{M}_{k+1} F_k^T \eta_k \\
& + h\mathcal{M}_{k+1} \beta_k \delta\Pi_k + hB\delta u_k,
\end{aligned}$$

$$\begin{aligned}\delta\Pi_{k+1} =&((F_k^T\Pi_k)\hat{}\beta_k + F_k^T + h\mathcal{M}_{k+1}\beta_k)\delta\Pi_k \\ &+ h\mathcal{M}_{k+1}F_k^T\eta_k + hB\delta u_k.\end{aligned} \tag{10}$$

Consequently, the linearized system (1-3) is summarized as follows

$$\begin{pmatrix} \eta_{k+1} \\ \delta\Pi_{k+1} \end{pmatrix} = \begin{pmatrix} \mathcal{A}_k & \mathcal{B}_k \\ \mathcal{C}_k & \mathcal{D}_k \end{pmatrix}\begin{pmatrix} \eta_k \\ \delta\Pi_k \end{pmatrix} + \begin{pmatrix} 0 \\ hB \end{pmatrix}\delta u_k, \tag{11}$$

while,

$$\begin{aligned}\mathcal{A}_k &= F_k^T, \\ \mathcal{B}_k &= \beta_k, \\ \mathcal{C}_k &= h\mathcal{M}_{k+1}F_k^T, \\ \mathcal{D}_k &= F_k^T + (F_k^T\Pi_k)\hat{}\mathcal{B}_k + h\mathcal{M}_{k+1}\mathcal{B}.\end{aligned}$$

and, in a 3D pendulum

$$\begin{aligned}\delta M_k &= \mathcal{M}_k\eta_k \\ \mathcal{M}_k &= mg\rho \times R_k^T e_3 = mg\hat{\rho}(R_k e_3)\hat{}.\end{aligned}$$

2.2 Optimization Problem

As in MPC problems, a convex optimization should be solved, a quadratic cost function for the linearized system (11) is introduced as follows to minimize cost as well as energy [15, 16, 17]

$$\mathcal{J} = \mathcal{F}(R_N, \hat{\Pi}_N) + \sum_{i=0}^{N} \mathcal{L}(R_k, \hat{\Pi}_k, \hat{u}_k)$$

such that

$$\begin{aligned}\mathcal{F} &= trace(P_1(I_{3\times3} - R_N)) + trace(P_2\hat{\Pi}_N) \\ &= \frac{1}{2}\|P_1^{1/2}(I_{3\times3} - R_N)\|_F^2 + \frac{1}{2}\|P_2^{1/2}\hat{\Pi}_N\|_F^2,\end{aligned}$$

$$\begin{aligned}\mathcal{L}(R_k, \pi_k, u_k) = &\, tr(Q_1(I - R_k)) + \frac{1}{h^2}trace(Q_2(I - F_k)) \\ &+ trace(u_k^T W_1 u_k) \\ = &\, \frac{1}{2}\|Q_1^{1/2}(I - R_k)\|_F^2 + \frac{1}{2h^2}\|Q_2^{1/2}(I - F_k)\|_F^2 \\ &+ \frac{1}{2}\|W_1^{1/2}u_k\|_F^2.\end{aligned}$$

Since $\hat{\eta}_k = \log R_k$; besides, in the neighborhood of (I,I) it is true that $\delta\Pi_k \approx \Pi_k$ [17]:

$$trace(Q_1(I - R_k)) = \frac{1}{2}\|Q_1^{1/2}(I - R_k)\|_F^2$$

$$= trace(Q_1(I - exp(\hat{\eta}_k)))$$
$$, = \frac{1}{2}\eta_k^T \tilde{Q}_1 \eta_k$$

$$\begin{aligned} trace(Q_2(I - F_k)) &= \frac{1}{2h^2}\|Q_2^{1/2}(I - F_k)\|_F^2 \\ &= \frac{1}{2h^2}\|Q_2^{1/2} h(J^{-1}\delta\Pi_k)\hat{}\|_F^2 \\ &= \frac{1}{2}\delta\Pi_k^T J^{-1}\tilde{Q}_{22}J^{-1}\delta\Pi_k, \end{aligned}$$

where $\tilde{Q}_{1,22} = trace(Q_{1,22})I_{3\times3} - Q_{1,22}$ for symmetric positive definite $Q_{1,2} \in \mathbf{R}^{3\times3}$, and $\tilde{Q}_2 = J^{-1}\tilde{Q}_{22}J^{-1}$. These matrices are evaluated from the property discussed in [15], which implies that for any positive semi-definite symmetric matrix B and $c \in \mathbf{R}^3$, $\frac{1}{2}\|B^{\frac{1}{2}}\hat{c}\|_F^2 = \frac{1}{2}c^T\tilde{B}c$, where $\tilde{B} = tr(B)I_{3\times3} - B$. As a result,

$$\begin{aligned} \acute{\mathcal{L}}(\eta_k, \delta\Pi_k, \tau_k) &= \frac{1}{2}\eta_k^T\tilde{Q}_1\eta_k \\ &+ \frac{1}{2}\delta\Pi_k^T\tilde{Q}_2\delta\Pi_k + \frac{1}{2}\tau_k^T W\tau_k, \\ \acute{\mathcal{F}}(\eta_N, \delta\Pi_N) &= \frac{1}{2}\eta_N^T\tilde{P}_1\eta_N + \frac{1}{2}\delta\Pi_N^T\tilde{P}_2\delta\Pi_N, \end{aligned}$$

where $W = trace(W_1)I - W_1$, and $u_k \in \mathfrak{so}(3)^*$ is expressed in terms of the applied torque τ_k as $u_k = \hat{\tau}_k$ where $\mathfrak{so}(3)^*$ is the dual of $\mathfrak{so}(3)$. We rewrite the linearized system and cost function as follows

$$\zeta_{k+1} = \bar{A}\zeta_k + \bar{B}\delta u_k, \tag{12}$$

$$\mathcal{J}_N = \frac{1}{2}\zeta_N^T P\zeta_N + \frac{1}{2}\sum_{i=1}^{N}\zeta_i^T Q\zeta_i + \tau_i^T W\tau_i, \tag{13}$$

where

$$\zeta_k = \begin{pmatrix} \eta_k \\ \delta\Pi_k \end{pmatrix}, \bar{A} = \begin{pmatrix} \mathcal{A}_k & \mathcal{B}_k \\ \mathcal{C}_k & \mathcal{D}_k \end{pmatrix}, \bar{B} = \begin{pmatrix} 0 \\ hB \end{pmatrix}$$
$$Q = \begin{pmatrix} \tilde{Q}_1 & 0 \\ 0 & \tilde{Q}_2 \end{pmatrix}, P = \begin{pmatrix} \tilde{P}_1 & 0 \\ 0 & \tilde{P}_2 \end{pmatrix}.$$

3 LMI Based Model Predictive Control

No control action is applied to the system after the instant $k+m-i$; namely, $u(k+i|k) = 0$ for $i \geq m$. From the viewpoint of the receding horizon, only the first calculated control is implemented to the system. In the next sampling time, the optimization problem $min\ \mathcal{J}$ is solved using the new measurement of the system. As a result, both m and p go one step ahead. Considering system (8), the minimization problem of the cost function is replaced by a worst-case minimization problem in each sampling time, namely the following $min - max$ problem

$$\min_{u(k+i|k)} \max_{[\mathcal{A}(k)\mathcal{B}(k)]\in\Omega} \mathcal{J}(k) \tag{14}$$

$$\begin{aligned}\mathcal{J}_N(k) &= \frac{1}{2}\zeta_N^T P\zeta_N + \frac{1}{2}\sum_{i=1}^{N}\zeta(k+i|k)^T Q\zeta(k+i|k)\\ &+ \tau(k+i|k)^T W\tau(k+i|k).\end{aligned} \tag{15}$$

Maximization is on the set $\Omega = ConvexHull[\bar{A}\ \bar{B}]$, arising from stability considerations, which gives an upper bound for the Lyapunov function and leads to a robust performance objective. Then, this bound should be minimized by the use of the state feedback control law $u(k+i|k) = K\zeta(k+i|k)\, i \geq 0$. The quadratic Lyapunov function is introduced in the form of

$$V(\zeta(k|k)) = \zeta(k|k)^T P\zeta(k|k), \quad P > 0. \tag{16}$$

According to the Lyapunov stability theorem, variations of $V(\zeta(k|k))$ should be negative in order to guarantee the stability of the system. Suppose that for any $\zeta(k+i|k), u(k+i|k)$ and $i \leq 0$, the variation of Lyapunov function is smaller than a negative quadratic function, which considers being the summand of cost function such as[12]

$$\begin{aligned}\Delta V(\zeta(k+i)) &= V(\zeta(k+i+1|k)) - V(\zeta(k+i|k))\\ &= \zeta(k+i+1|k)^T P_{k+1}\zeta(k+i+1|k) - \zeta(k+i|k)^T P_k\zeta(k+i|k)\\ &\leq -(\zeta(k+i|k)^T Q_k\zeta(k+i|k) + \tau(k+i|k)^T W\tau(k+i|k)).\end{aligned} \tag{17}$$

Let us calculate the sum of both sides of (17). Firstly, the sum of the first side of it is calculated as follows [18]

$$\begin{aligned}&\sum_{i=0}^{N-1}\Delta\Big[\zeta(k+i+1|k)^T P_{k+1}\zeta(k+i+1|k)\Big]\\ &= \sum_{i=0}^{N-1}\Big[\zeta(k+i+1|k)^T P_{k+1}\zeta(k+i+1|k)\ -\zeta(k+i|k)^T P_k\zeta(k+i|k)\Big]\\ &= \zeta^T(k+N|k)P_{k+1}\zeta(k+N|k) - \zeta^T(k|k)P_k\zeta(k|k).\end{aligned}$$

Note: $\zeta(k+N|k)$ is summerized as ζ_N. Finally, it gives

$$-\zeta(k|k)^T P_k\zeta(k|k) \leq -\zeta_N^T P\zeta_N - \sum_{i=0}^{N-1}\Big[\zeta(k+i|k)^T Q_k\zeta(k+i|k)\ +\tau(k+i|k)^T W\tau(k+i|k)\Big],$$

as a result

$$\begin{aligned}-V(\zeta(k|k)) &\leq -\mathcal{J}_N(k)\\ \mathcal{J}_N(k) &\leq \zeta(k|k)^T P_k\zeta(k|k).\end{aligned} \tag{18}$$

So (18) is an upper bound for the cost function $\mathcal{J}$; namely, the problem of $max\mathcal{J}$ gives the upper bound $V(\zeta(k|k))$ for $\mathcal{J}$. It is clear that the minimization problem has been changed to determining the state feedback control gain K of $\tau(k+i|k) = K\zeta(k+i|k),\ i \geq 0$ in each sampling time k for the minimization of this upper bound of $V(\zeta(k|k))$. It means we should minimize the upper bound of $\mathcal{J}$, which is the function V. Similar to the problem of standard MPC, firstly, the first calculated input $\tau(k|k) = K\zeta(k|k)$ should be implemented to the system. In the next sampling time, the state ζ_{k+1} is measured, and optimization is repeated to compute F again. Since the linearized system has been embed on Euclidean space using the exponential map, it is a convex optimization problem and can be solved under linear matrix inequality conditions. The following theorem expresses LMI conditions on which the gain of controller K is going to calculate

Theorem 1. Let $\zeta_k = \zeta(k|k)$ be the state of the system (12) measured at the sampling time k. There are no constraints on inputs and outputs of the system. Then, the feedback matrix K in control law $\tau(k+i|k) = K\zeta(k+i|k), i \leq 0$ which minimizes the upper bound $V(\zeta(k|k))$ on the robust performance of cost function in sampling time k can be computed as follows[12]

$$K = LE^{-1}, \tag{19}$$

where the matrices $E > 0, L$ (if it exists) are obtained from the following linear minimization problem:

$$\min_{\gamma,E,L} \quad \gamma \tag{20}$$

$$subject.to.:$$

$$\begin{pmatrix} 1 & \zeta(k|k)^T \\ \zeta(k|k) & E \end{pmatrix} \geq 0, \tag{21}$$

$$\begin{pmatrix} E & EA^T + L^T B^T & EQ^{1/2} & L^T W_1^{1/2} \\ AE + BL & E & 0 & 0 \\ Q^{1/2}E & 0 & \gamma I & 0 \\ W_1^{1/2}E & 0 & 0 & \gamma I \end{pmatrix}. \geq 0 \tag{22}$$

Proof: See Appendix. A in [12]. □

4 Simulation Results

In this section, a numerical simulation is presented in order to analyze the effectiveness of the proposed method. Standard inertial matrix is chosen as $J = diag(1, 2.8, 2)$. The discretization time-step parameter $h = 0.2$. The initial angular velocity in the body-fixed frame is considered as $\Omega_0 = [0, 0, 1]$ while $\Pi = J.\Omega$, and $\eta_0 = [0, 0, 1.5]$. Solving LMI, the control gain K is computed in each iteration with control horizon 2, and only the first parameter of K is implemented to the system. The simulation results are depicted on figures (1,2). Figures show that only 7 seconds takes for the pendulum to reach its equilibrium. Simulation repeated for more complicated initial conditions, which expressed complicated starting point of the pendulum, result with initial conditions as $\Omega = [0, 1, 1]$ and $\eta_0 = [-0.5, 0, 1.5]$ is illustrated on figure (3). Using the LMI method for solving the MPC problem on the manifold is compared with the standard

MPC method on a manifold without using the LMI. Results are depicted on figures (4,5), which demonstrate fewer control efforts. This is a rest-to-rest initial condition simulation.

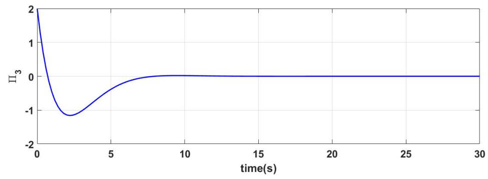


Figure 1: Angular Momentum Π_3.

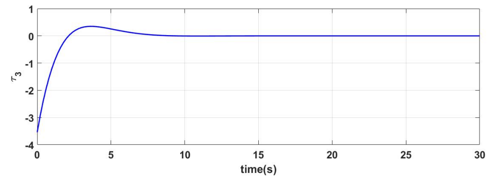


Figure 2: Input torque τ_3.

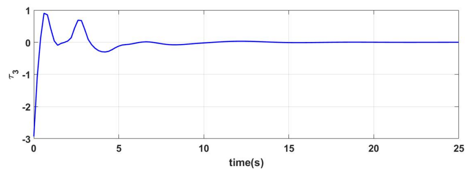


Figure 3: Input torque τ_3.

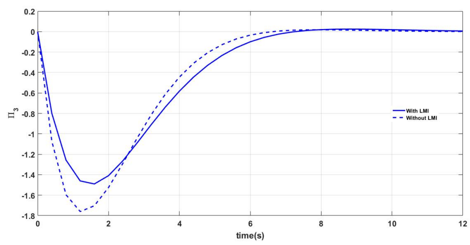


Figure 4: Comparing MPC based LMI on manifold method with regular MPC on manifold method for the parameter Π_3.

5 Conclusions

This paper formulated a model predictive method for the 3D pendulum, in which its configuration space is expressed as a manifold. Its dynamics are used as LGVI equations, and a

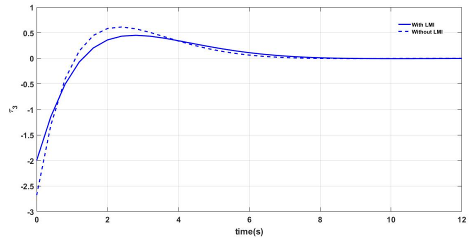


Figure 5: Comparing MPC based LMI on manifold method with regular MPC on manifold method for input torque parameter τ_3.

linearization method on manifolds has been used in order to generalize the conventional MPC methods from Euclidean spaces to manifolds. Solving MPC and calculating control gain is achieved using LMI conditions.

References

[1] Agrachev A., Sachkov Y. (2013). "Control theory from the geometric viewpoint", Springer Science & Business Media, 87.

[2] Holm D., Schmah T., Stoica C. (2009). "Geometric mechanics and symmetry: From finite to infinite dimensions", Oxford University Press, 12.

[3] Kizil E. (2020). "Control Homotopy of Trajectories", Journal of Dynamical and Control Systems.

[4] Lee T. (2008). "Computational Geometric Mechanics and Control of Rigid Bodies", Ph.D. thesis, Ann Arbor: University of Michigan.

[5] Lee T., Leok M., Harris McClamroch N. (2017). "Global Formulations of Lagrangian and Hamiltonian Dynamics on Manifolds", Springer.

[6] Colombo L.J., Marina H.G. (2018). "A Variational Integrator for the distance-based formation control of multi-agent systems", IFAC-papers online, 51, Issue 23, 76-81.

[7] Man S., Gao Q., Zhong W. (2020). "Variational Integrators in Holonomic Mechanics", Mathematics 8, 1358.

[8] Iserles A., Munthe-Kaas H.Z., Nørsett S.P., Zanna A. (2000). "Lie-group methods", Acta numerica, 9, 215-365.

[9] Hairer E., Lubich C., Wanner G. (2006). "Geometric numerical integration: structure-preserving algorithms for ordinary differential equations", 31. Springer Science and Business Media.

[10] Kalabić U.V., Gupta R., Cairano S.Di, Boloch A.M., Kolmanovsky I.V. (2017). "MPC on manifolds with an application to the control of spacecraft attitude on SO(3)", Automatica, 76, 293-300.

[11] Kalabić U.V., Gupta R., Cairano S.D., Bloch A.M., Kolmanovsky I.V. (2016). "MPC on manifolds with applications to the control of systems on matrix Lie groups", Automatica.

[12] Kothare, M.V., Venkataramanan B., Manfred M. (1996). "Robust constrained model predictive control using linear matrix inequalities", Automatica, 32.10, 1361-1379.

[13] Lee T., McClamroch N. H., Leok M. (2005). "A Lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3D pendulum", IEEE Conference on Control Applications, 962-967.

[14] Lee T., Leok M., Harris McClamroch N. (2008). "Optimal attitude control of a rigid body using geometrically exact computations on SO(3)", Dynamical and Control Systems. J, 14(4), 465-487.

[15] Kalabic U., Gupta R., Di Cairano S., Bloch A., Kolmanovsky I. (2014). "Constrained spacecraft attitude control on SO(3) using reference governors and nonlinear model predictive control", American Control Conference (ACC), 5586-5593.

[16] Saccon A., Hauser J., Aguiar P. (2010). "Exploration of kinematic optimal control on the Lie group SO(3)", 8th IFAC Symposium on Nonlinear Control Systems, 1302-1307.

[17] Gupta R., Kalabíc U., Di Cairano S., Bloch A., Kolmanovsky I. (2015). "Constrained spacecraft attitude control on SO(3) using fast nonlinear model predictive control", American Control Conference (ACC), 2980-2986.

[18] Strejc V. (1972). "State-space synthesis of discrete linear systems", Kybernetika, 8(2).

[19] Gupta R. (2016). "Analytical and Numerical Methods for Optimal Control Problems on Manifolds and Lie Groups".

[20] Lee T., Leok M., Harris McClamroch N. (2011). "Stable manifolds of saddle equilibria for pendulum dynamics on S^2 and $SO(3)$", in 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC), 3915-3921.

[21] Poursafar N., Taghirad H.D., Haeri M. (2010). "Model predictive control of non-linear discrete-time systems: a linear matrix inequality approach", IET Control Theory and Applications, 4(10) 1922-1932.

Persian Abstracts

طراحی یک قانون کلیدزنی مقاوم برای سیستم‌های کلیدزنی خطی دارای عدم قطعیت و تاخیر زمانی

رمضانی آل، م. ر. - نویسنده مسئول
دانشکده مهندسی برق و کامپیوتر، گروه مهندسی کنترل، دانشگاه صنعتی قوچان،
قوچان، ایران، صندوق پستی ۳۶۹۷-۱۹۳۹۵،
m-ramezani@qiet.ac.ir

ثابتی، س.
دانشکده ریاضی، گروه آمار و علوم کامپیوتر، دانشگاه سمنان،
سمنان، ایران، صندوق پستی ۱۳۵۵-۱۶۳۱۵۰.
sabeti.samira@semnan.ac.ir

تاریخ دریافت: ۲۲ تیر ۱۳۹۹ تاریخ پذیرش: ۱۹ دي ۱۳۹۹

چکیده

کنترل هزینه تضمینی یکی از روش‌های موثر درکنترل سیستم‌های غیرخطی به ویژه سیستم‌های کلیدزنی دارای عدم قطعیت است. بسیاری از تحقیقات اخیر در مسیله کنترل هزینه تضمینی سیستم‌های کلیدزنی دارای عدم قطعیت بر روی تحلیل پایداری مجانبی تمرکزیافته است. در این مقاله یک قانون سوییچ مقاوم جدید برای کنترل سیستم‌های کلیدزنی دارای عدم قطعیت و تاخیر زمانی ارایه می‌شود. در ابتدا قانون کلیدزنی و سپس کنترل کننده فیدبک حالت خطی بر اساس تابع لیاپانوف-کراسوفسکی طراحی می‌گردد. همچنین با استفاده از نامساوی‌های خطی ماتریسی، شرایط ویژه برای وجود جواب در قانون کلیدزنی و کنترل کننده خطی به دست می‌آید. همزمان و بر اساس قضایای ارایه شده، پایداری نمایی کل سیستم تحت اعمال قانون کلیدزنی و کنترل فیدبک حالت تحلیل و اثبات می‌گردد. در انتها و با شبیه‌سازی، نتایج تحلیلی پایداری نمایی نشان داده می‌شوند.

کلمات کلیدی

سیستم‌های کلیدزنی دارای عدم قطعیت، تاخیر زمانی، کنترل هزینه تضمینی، تابع لیاپانوف-کراسوفسکی، نامساوی خطی ماتریسی، پایداری نمایی.

توسیع ناحیه جذب سیستم‌های غیرخطی با استفاده از برنامه نویسی مجموع مربعات

علیزاده، ج.
ایران، تهران، دانشگاه پیام نور، گروه ریاضی، صندوق پستی ۳۶۹۷-۱۹۳۹۵
Jahangir_Alizadeh@yahoo.com ،

خالوزاده، ح. – نویسنده مسئول
ایران، تهران، دانشگاه صنعتی خواجه نصیرالدین طوسی، دانشکده مهندسی برق، گروه کنترل و سیستم، صندوق پستی ۱۳۵۵-۱۶۳۱۵.
h_khaloozadeh@kntu.ac.ir

تاریخ دریافت: ۷ تیر ۱۳۹۹ تاریخ پذیرش: ۲۴ آبان ۱۳۹۹

چکیده

در این پژوهش روشی جدید برای توسیع ناحیه جذب نقطه تعادل یک سیستم کنترل غیرخطی آفین ارائه شده است. برای توسیع ناحیه جذب سیستم‌های دینامیکی غیر چندجمله‌ای به کمک طراحی کنترل‌کننده فیدبک حالت غیرخطی از معادله ریکاتی وابسته به حالت استفاده شده است. فرایند بدست آوردن ناحیه جذب باعث ایجاد یک مسئله بهینه سازی مجموع مربعات با محدودیت‌های کنترلی و غیر کنترلی می‌شود. نکته قابل توجه این است که روش پیشنهادی برای تخمین ناحیه جذب سیستم‌های غیرخطی چندجمله‌ای و غیر چندجمله‌ای کارا است. همچنین در این مطالعه کاربرد روش پیشنهادی با شبیه سازی‌های عددی نشان داده شده است.

کلمات کلیدی

معادله ریکاتی وابسته به حالت، ناحیه جذب، فاکتور شکل، برنامه‌نویسی مجموع مربعات، تابع لیاپانوف.

کشف گروه‌های کاربری با بیشترین توصیف‌پذیری در شبکه‌های اجتماعی

عباسی، ز.
ایران- دامغان – دانشگاه دامغان- دانشکده ریاضی و علوم کامپیوتر،
zahra.abbasi@yahoo.com

آخوندی روشناوند، ن. – نویسنده مسئول
ایران – دامغان – دانشگاه دامغان- دانشکده ریاضی و علوم کامپیوتر
akhoundi@du.ac.ir

تاریخ دریافت: ۲۱ آذر ۱۳۹۸ تاریخ پذیرش: ۲۱ آذر ۱۳۹۹

چکیده

نظرات و قضاوت کاربران در وب‌سایت‌های تجارت الکترونیک مانند وب‌سایت‌های رستوران‌ها، فروش فیلم، فروش محصولات و غیره، برای مشتری‌ها در خصوص اخذ تصمیم خرید بسیار اهمیت دارند. در این مقاله، ما کشف گروه‌های کاربری با بیشترین توصیف‌پذیری را از وب‌سایت‌های تجارت الکترونیک به صورت $\langle i,u,s\rangle$، که $i \in I$ (مجموعه آیتم‌های محصولات)، $u \in U$ (مجموعه کاربران) و s عدد صحیح که امتیاز کاربر u به آیتم i اختصاص داده است. گروه‌های برچسب‌دار از صفات کاربری با حل مساله بهینه‌سازی به دست می‌آیند. کارایی روش با برخی آزمون‌ها بر روی مجموعه داده‌های واقعی مورد ارزیابی قرار می‌گیرد.

کلمات کلیدی

بیشترین توصیف‌پذیری، بهینه‌سازی، کشف گروه کاربری، داده‌های ارزیابی.

تجزیه و تحلیل اشتباهات دانش‌آموزان در حل انتگرال برای به حداقل رساندن اشتباهات آن‌ها

درویش‌زاده، م.
ایران، تهران، دانشگاه آزاد اسلامی، گروه ریاضی،
Darvish.math@gmail.com

شهورانی سمنانی، ا. – نویسنده مسئول
ایران، تهران، دانشگاه آزاد اسلامی، گروه ریاضی،
Shahverani.1393@gmail.com

علم‌الهدایی، ح.
ایران، مشهد، دانشگاه فردوسی مشهد، گروه ریاضی،
alamolhodaei@yahoo.com

بهزادی، م. ح.
ایران، تهران، دانشگاه آزاد اسلامی، گروه آمار،
behzadi@srbiau.ac.ir

تاریخ دریافت: ۲۳ شهریور ۱۳۹۹ تاریخ پذیرش: ۲۱ آذر ۱۳۹۹

چکیده

اکثر دانشجویان ایرانی در حل مسائل مربوط به انتگرال دارای مشکلات عدیده‌ای می‌باشند و در این مبحث ضعیف‌اند و حتی می‌توان گفت که از انتگرال فراری هستند و یا به عبارتی، انتگرال را کابوس ریاضیات خود می‌دانند. لذا می‌خواهیم با برگزاری آزمون و سپس مصاحبه، مشکلات پیش‌روی دانشجویان مهندسی و علوم پایه درحل مسایل انتگرالی را تجزیه و تحلیل کرده، و سپس در حد امکان راه حل‌هایی برای رفع آنها ارائه دهیم . این تحقیق شامل سه سوال می‌باشد که سوال اول خود از ۴ بخش تشکیل شده است، که از دانشجویان منتخب گرفته شده و سپس مصاحبه‌ای کوتاه با چند نفر از آنها در مورد پاسخ هایشان انجام شده است. با بررسی عملکرد دانشجویان در این آزمون، مشاهده می‌شود که دانشجویان غالباً با مباحث انتگرال مشکل داشته، به‌ویژه در حل مسائل انتگرال‌های مثلثاتی بسیار ضعیف عمل می‌کنند، در واقع آنها به جاي مفهوم‌سازی کامل و بی‌عیب انتگرال بیشتر به دنبال یادگیری انتگرال محاسبه‌ای هستند. بیشترین خطای انجام شده توسط دانشجویان خطای رویه‌ای بوده است که غالباً این نوع خطاها ناشی از استفاده مشتق به‌جای انتگرال می‌باشد. همچنین اکثر اشتباهات دانشجویان در حل انتگرال‌های معین و محاسبه مساحت‌های محدود بین دو منحنی می‌باشد که این هم ناشی از عدم درک کافی انتگرال و هم ناشی از نداشتن اطلاعات لازم در سایر زمینه‌های ریاضی می‌باشد.

کلمات کلیدی

انتگرال، درک دانشجو، درک آموزش ریاضی، مفهوم تابع اولیه، مفهوم ریمان.

گراف‌های به طور ضعیف تام روی مدول‌ها

نوری جویباری، م.
دانشجوی دکتری دانشگاه مازندران،
دانشگاه مازندران، دانشکده علوم ریاضی ، گروه ریاضی محض، بابلسر، ایران،
mostafa.umz@gmail.com

طالبی رستمی، ی.
دانشیار ریاضی دانشگاه مازندران
دانشگاه مازندران، دانشکده علوم ریاضی، گروه ریاضی محض، بابلسر، ایران،
talebi@umz.ac.ir

فیروزیان، س. - نویسنده مسئول
استادیار ریاضی دانشگاه پیام نور
دانشگاه پیام نور، گروه ریاضی، تهران، ایران، صندوق پستی ۳۶۹۷-۱۹۳۹۵.
siamfirouzian@pnu.ac.ir

تاریخ دریافت: ۲۳ شهریور ۱۳۹۹ تاریخ پذیرش: ۹ بهمن ۱۳۹۹

چکیده

در این مطالعه، R و M به ترتیب یک حلقه‌ی جابجایی با همانی غیر صفر و یک R-مدول است. گراف ضرب اسکاری روی M، را که با $G_R(M)$ نشان می‌دهیم، گرافی با مجموعه رئوس M است و دو راس متمایز a و b در M مجاورند اگر و تنها اگر r متعلق به R وجود داشته باشد به طوری که $a = rb$ یا $b = ra$. این مقاله برخی از خواص این گراف‌های به طور ضعیف تام را مطالعه می‌کند.

کلمات کلیدی

ضرب اسکالری، گراف متصل، به طور ضعیف تام، مدول.

طراحی کنترل پیش‌بین مدل برای کنترل پاندول سه بعدی روی منیفلد SO(۳) با استفاده از بهینه‌سازی محدب

منصوری نسب، س.
دانشگاه تربیت مدرس، دانشکده مهندسی برق و کامپیوتر، گروه کنترل، تهران، ایران
sara.mansouri@modares.ac.ir

سجودی، م. - نویسنده مسئول
دانشگاه تربیت مدرس، دانشکده مهندسی برق و کامپیوتر، گروه کنترل، تهران، ایران
sojoodi@modares.ac.ir

مقدسی، س.ر.
دانشگاه صنعتی شریف، دانشکده علوم ریاضی، تهران، ایران.
moghadasi@math.sharif.edu

تاریخ دریافت: ۵ بهمن ۱۳۹۸ تاریخ پذیرش: ۹ بهمن ۱۳۹۹

چکیده

روش‌های کنترل پیش‌بین بر پایه مدل (MPC) غالباً به سیستم‌های با دینامیک گسسته که در فضای برداری هموار $\mathbf{R}^n$ قرار دارند اعمال می‌شود. در حالی‌که فضای پیکره‌بندی بسیاری از سیستم‌های مکانیکی نمی‌تواند به صورت فضای اقلیدسی سراسری بیان شود. به همین جهت، روش MPC در این مقاله روی منیفلد هموار به عنوان فضای پیکره‌بندی برای کنترل وضعیت پاندول سه بعدی اعمال می‌شود. از انتگرال وردشی گروه لی (LGVI) معادلات حرکت پاندول سه بعدی به عنوان معادلات زمان گسسته استفاده می‌شود، زیرا معادلات LGVI ساختار گروه را حفظ می‌کنند. معادلات LGVI پاندول سه بعدی حول نقطه تعادل با استفاده از دیفئومورفیسم روی $\mathbf{R}^n$ خطی‌شده، سپس الگوریتم MPC به معادلات خطی‌شده اعمال می‌گردد. همچنین، مشابه روش‌های معمول MPC، مسئله بهینه‌سازی محدب حل شده و قانون کنترل در هر تکرار بدست می‌آید. در این مقاله، از نامساوی ماتریس خطی برای حل مسئله بهینه‌سازی محدب استفاده می‌شود. از یک مثال عددی برای نشان دادن عملکرد روش طراحی استفاده شده است.

کلمات کلیدی

کنترل پیش‌بین مدل، بهینه‌سازی محدب، نامساوی ماتریسی خطی، انتگرال وردشی گروه لی.

فرم اشتراک

علاقه‌مندان به اشتراک نشریه

Control and Optimization in Applied Mathematics-COAM

می توانند فرم ذیل را تکمیل نمایند و به همراه فیش بانکی به مبلغ ۵۰۰۰۰ ریال (پنجاه هزار ریال) به شماره حساب ۲۱۷۸۶۰۹۰۰۱۰۰۷: و معادل شبای متمرکز IR ۴۲ ۰۱۷۰ ۰۰۰۰ ۰۲۱۷ ۸۶۰۹ ۰۷ ، نزد بانک ملی ایران شعبه بنفشه، کد: ۱۵۰۸ به دبیرخانه مجله ارسال دارند تا مجله برای آنان فرستاده شود.

نام:
نام موسسه یا مرکز:
آخرین مدرک تحصیلی:

نام خانوادگی:
رشته تحصیلی:
شماره اشتراک:

اشتراک از شماره............. تا شماره .... و تعداد مورد نیاز از هر شماره .... نسخه
نشانی کامل پستی:

کدپستی:
تلفن تماس:
دورنگار:
تاریخ:

تلفن همراه:
پست الکترونیکی:
امضاء: