Deep-Neuro Control with Contraction Theory

Myenogseok Ryu, Sesun You

20^{th} March 2025

Abstract

This project aims to develop control or estimator with deep neural network and contraction theory.

Contents

1 Introduction . . . **1**
1.1 Motivation . . . 1
1.2 Literature Review . . . 1
2 Notations and Preliminaries . . . **1**
3 Review of Contraction Theory . . . **2**
3.1 Basic Results of Contraction Theory for Deterministic Systems . . . 2
3.1.1 Notable Definitions . . . 2
3.1.2 Notable Theorems . . . 2
3.2 Basic Results of Contraction Theory for Stochastic Systems . . . 3
4 Deep-Neuro Control . . . **3**
A Notable Lemmas . . . **3**

1 Introduction

1.1 Motivation

1.2 Literature Review

2 Notations and Preliminaries

The following notations are used throughout this document:

- $:=$ denotes *defined as*.
- $(\cdot)^\top$ denotes the transpose of a matrix or a vector.
- $\boldsymbol{x} := [x_i]_{i\in\{1,\cdots,n\}} \in \mathbb{R}^n$ denotes the state vector.
- $\boldsymbol{A} := [a_{ij}]_{i,j\in\{1,\cdots,n\}} \in \mathbb{R}^{n\times n}$ denotes a matrix.
- $\lambda_i(\boldsymbol{A}),\ i \in \{\max, \min\}$ denotes the maximum and minimum singular value of $\boldsymbol{A}$, respectively.
- $\boldsymbol{I}_n$ denotes the identity matrix of size n and $\mathbf{0}_{n\times m}$ denotes the zero matrix of size $n \times m$.
- sym denotes the symmetric part of a matrix, i.e., $\mathrm{sym}(\boldsymbol{A}) := \boldsymbol{A} + \boldsymbol{A}^\top$ [1].

We introduce the following lemmas.

Lemma 2.1 (Comparison Lemma)**.** *Suppose that a continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ satisfies the following inequality:*

$$\tfrac{\mathrm{d}}{\mathrm{d}t} f(t) \le -af(t) + b, \quad \forall t \in \mathbb{R}_{\ge 0},$$

where $a, b > 0$. Then, the following inequality holds:

$$f(t) \le -af(0)e^{-at} + \tfrac{b}{a}(1 - e^{-at}), \quad \forall t \in \mathbb{R}_{\ge 0}$$

and remains in a compact set $f(t) \in \{\|f(t)\| \mid \|f(0)\| \le \frac{b}{a}\}$.

Proof. This is a simple special case of the comparison lemma [2, pp. 102-103]. See [2, pp. 659-660]. □

3 Review of Contraction Theory

For your smooth start, we recommend you to begin with [3].

3.1 Basic Results of Contraction Theory for Deterministic Systems

First, we start with the following deterministic systems:

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x}, t), \tag{1}$$

where $\boldsymbol{f}(\boldsymbol{x}, t)$ is an $n \times 1$ sufficiently smooth non-linear vector function and $\boldsymbol{x} \in \mathbb{R}^n$ is the state vector. The smooth property of $\boldsymbol{f}(\boldsymbol{x}, t)$ is essential to ensure the existence and uniqueness of the solution to (1) [2, see, pp. 88-89].

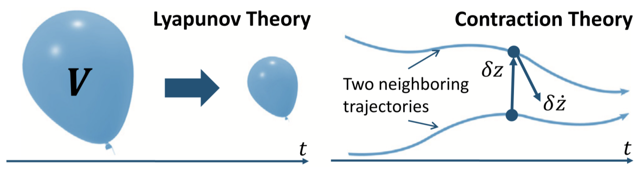

Figure 1: Difference between Lyapunov and contraction theory [4, Fig. 1]. The Lyapunov theory investigates the convergence to a single point and the contraction theory does regarding a single trajectory.

The biggest difference between the traditional Lyapunov theory and the contraction theory is that the contraction theory investigates the convergence of the state trajectory to a single trajectory (contraction behavior), while the Lyapunov theory focuses on the convergence of the state trajectory to a single point i.e., see, Fig. 1. For this, motivated by the calculus of variations [5, Chap. 4], (1) can be rewritten as differential dynamics using *differential displacement* $\delta\boldsymbol{x}$ as follows:

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\delta\boldsymbol{x} = \tfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}(\boldsymbol{x}, t)\delta\boldsymbol{x}..$$

For your information, $\delta\boldsymbol{x}$ is an infinitesimal displacement at *fixed time*.

3.1.1 Notable Definitions

Before we present the fundamental theorem of contraction theory, we introduce the following definitions. One can re-visit this section while reading further.

Definition 3.1 (see, Def. 2.2 [4])**.** If any two trajectories $\boldsymbol{\xi}_1(t)$ and $\boldsymbol{\xi}_2(t)$ of (1) converge to a single trajectory, then the system (1) is said to be *incrementally exponentially stable*, if $\exists C, \alpha > 0$, subject to the following holds:

$$\|\boldsymbol{\xi}_1(t) - \boldsymbol{\xi}_2(t)\| \leq C\|\boldsymbol{\xi}_1(0) - \boldsymbol{\xi}_2(0)\| \exp^{-\alpha t},\ \forall t \in \mathbb{R}_{\geq 0}.$$

The result of Theorem 3.1 equivalently implies the incremental exponential stability, since we have $\|\boldsymbol{\xi}_1(t) - \boldsymbol{\xi}_2(t)\| = \|\int_{\boldsymbol{\xi}_1(t)}^{\boldsymbol{\xi}_2(t)} \delta\boldsymbol{x}(t)\|$.

Definition 3.2. Let $\boldsymbol{\Theta}(\boldsymbol{x}, t)$ be a smooth coordinate transformation of $\delta\boldsymbol{x}$ to $\delta\boldsymbol{z}$, i.e., $\delta\boldsymbol{z} = \Theta(\boldsymbol{x}, t)\delta\boldsymbol{x}$. Then, a symmetric continuously differentiable matrix $\boldsymbol{M}(\boldsymbol{x}, t) := \boldsymbol{\Theta}(\boldsymbol{x}, t)^\top\boldsymbol{\Theta}(\boldsymbol{x}, t)$ is said to be a *metric* of the system (1).

Definition 3.3. The covariant derivative of $\boldsymbol{f}(\boldsymbol{x}, t)$ in $\delta\boldsymbol{x}$ coordinate is represented as

$$\boldsymbol{F} := \left(\tfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\Theta} + \boldsymbol{\Theta}\tfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right)\boldsymbol{\Theta}^{-1},$$

and is called the *generalized Jacobian*. This can be easily derived by differentiating $\delta\boldsymbol{z} = \Theta(\boldsymbol{x}, t)\delta\boldsymbol{x}$ with respect to t.

3.1.2 Notable Theorems

The following theorem is the fundamental theorem of contraction theory.

Theorem 3.1 (see, T. 2.1 [4]). *If* $\exists \boldsymbol{M}(\boldsymbol{x},t) = \boldsymbol{\Theta}(\boldsymbol{x},t)^\top \boldsymbol{\Theta}(\boldsymbol{x},t) > 0, \forall \boldsymbol{x}, t$ *where* $\boldsymbol{\Theta}(\boldsymbol{x},t)$ *defines a smooth coordinate transformation of* $\delta \boldsymbol{x}$ *to* $\delta \boldsymbol{z}$*, i.e.,* $\delta \boldsymbol{z} = \Theta(\boldsymbol{x},t)\delta \boldsymbol{x}$*, subject to the following equivalent conditions holds for* $\exists \alpha \in \mathbb{R}_{>0},\ \forall \boldsymbol{x}, t$*:*

$$\lambda_{\max}(\boldsymbol{F}(\boldsymbol{x},t)) = \lambda_{\max}\left(\left(\tfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\Theta} + \boldsymbol{\Theta}\tfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right)\boldsymbol{\Theta}^{-1}\right) \leq -\alpha,$$
$$\tfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{M} + \mathrm{sym}(\boldsymbol{M}\tfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}) \leq -2\alpha \boldsymbol{M},$$

where the arguments of $\boldsymbol{M}(\boldsymbol{x},t)$ *and* $\boldsymbol{F}(\boldsymbol{x},t)$ *are omitted for simplicity, then, the system* (1) *is said to be contracting with an exponential rate* α*, i.e., all trajectories of* (1) *converge to a single trajectory. The converse is also true.*

Proof. □

3.2 Basic Results of Contraction Theory for Stochastic Systems

4 Deep-Neuro Control

A Notable Lemmas

References

[1] H. Tsukamoto, S.-J. Chung, and J.-J. E. Slotine, "Neural stochastic contraction metrics for learning-based control and estimation," *IEEE Control Systems Letters*, vol. 5, no. 5, pp. 1825–1830, 2021.

[2] H. K. Khalil, *Nonlinear systems; 3rd ed.* Upper Saddle River, NJ: Prentice-Hall, 2002. The book can be consulted by contacting: PH-AID: Wallet, Lionel.

[3] W. LOHMILLER and J.-J. E. SLOTINE, "On contraction analysis for non-linear systems," *Automatica*, vol. 34, no. 6, pp. 683–696, 1998.

[4] H. Tsukamoto, S.-J. Chung, and J.-J. E. Slotine, "Contraction theory for nonlinear stability analysis and learning-based control: A tutorial overview," *Annual Reviews in Control*, vol. 52, pp. 135–169, 2021.

[5] D. Kirk, *Optimal Control Theory: An Introduction.* Dover Books on Electrical Engineering Series, Dover Publications, 2004.