

Deep-Neuro Control with Contraction Theory

Myenogseok Ryu, Sesun You

20th 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.
- $\mathbf{x} := [x_i]_{i \in \{1, \dots, n\}} \in \mathbb{R}^n$ denotes the state vector.
- $\mathbf{A} := [a_{ij}]_{i,j \in \{1, \dots, n\}} \in \mathbb{R}^{n \times n}$ denotes a matrix.
- $\lambda_i(\mathbf{A})$, $i \in \{\max, \min\}$ denotes the maximum and minimum singular value of \mathbf{A} , respectively.
- \mathbf{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., $\text{sym}(\mathbf{A}) := \mathbf{A} + \mathbf{A}^\top$ [1].

We introduce the following lemmas.

Lemma 1 (Comparison Lemma). *Suppose that a continuously differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ satisfies the following inequality:*

$$\frac{d}{dt}f(t) \leq -af(t) + b, \quad \forall t \in \mathbb{R}_{\geq 0},$$

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

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

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

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

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:

$$\frac{d}{dt}\mathbf{x} = \mathbf{f}(\mathbf{x}, t), \quad (1)$$

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

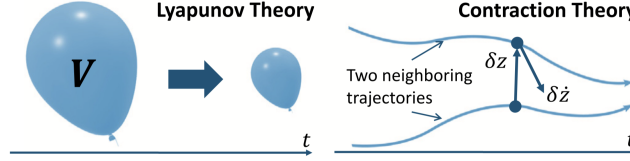


Figure 1: Difference between Lyapunov and contraction theory [4, Fig. 1]. The Lyapunov theory investigates the convergence to a single point and the contraction theory does regarding a single trajectory.

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* $\delta\mathbf{x}$ as follows:

$$\frac{d}{dt}\delta\mathbf{x} = \frac{\partial\mathbf{f}}{\partial\mathbf{x}}(\mathbf{x}, t)\delta\mathbf{x}..$$

For your information, $\delta\mathbf{x}$ 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 1 (see, Def. 2.2 [4]). If any two trajectories $\xi_1(t)$ and $\xi_2(t)$ 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:

$$\|\xi_1(t) - \xi_2(t)\| \leq C\|\xi_1(0) - \xi_2(0)\| \exp^{-\alpha t}, \quad \forall t \in \mathbb{R}_{\geq 0}.$$

The result of Theorem 1 equivalently implies the incremental exponential stability, since we have $\|\xi_1(t) - \xi_2(t)\| = \|\int_{\xi_1(t)}^{\xi_2(t)} \delta\mathbf{x}(t)\|$.

Definition 2. Let $\Theta(\mathbf{x}, t)$ be a smooth coordinate transformation of $\delta\mathbf{x}$ to $\delta\mathbf{z}$, i.e., $\delta\mathbf{z} = \Theta(\mathbf{x}, t)\delta\mathbf{x}$. Then, a symmetric continuously differentiable matrix $\mathbf{M}(\mathbf{x}, t) := \Theta(\mathbf{x}, t)^\top \Theta(\mathbf{x}, t)$ is said to be a *metric* of the system (1).

Definition 3. The covariant derivative of $\mathbf{f}(\mathbf{x}, t)$ in $\delta\mathbf{x}$ coordinate is represented as

$$\mathbf{F} := \left(\frac{d}{dt}\Theta + \Theta \frac{\partial\mathbf{f}}{\partial\mathbf{x}} \right) \Theta^{-1},$$

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

3.1.2 Notable Theorems

The following theorem is the fundamental theorem of contraction theory.

Theorem 1 (see, T. 2.1 [4]). *If $\exists \mathbf{M}(\mathbf{x}, t) = \mathbf{\Theta}(\mathbf{x}, t)^\top \mathbf{\Theta}(\mathbf{x}, t) > 0, \forall \mathbf{x}, t$ where $\mathbf{\Theta}(\mathbf{x}, t)$ defines a smooth coordinate transformation of $\delta \mathbf{x}$ to $\delta \mathbf{z}$, i.e., $\delta \mathbf{z} = \mathbf{\Theta}(\mathbf{x}, t) \delta \mathbf{x}$, subject to the following equivalent conditions holds for $\exists \alpha \in \mathbb{R}_{>0}, \forall \mathbf{x}, t$:*

$$\begin{aligned} \lambda_{\max}(\mathbf{F}(\mathbf{x}, t)) &= \lambda_{\max} \left(\left(\frac{d}{dt} \mathbf{\Theta} + \mathbf{\Theta} \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right) \mathbf{\Theta}^{-1} \right) \leq -\alpha, \\ \frac{d}{dt} \mathbf{M} + \text{sym}(\mathbf{M} \frac{\partial \mathbf{f}}{\partial \mathbf{x}}) &\leq -2\alpha \mathbf{M}, \end{aligned}$$

where the arguments of $\mathbf{M}(\mathbf{x}, t)$ and $\mathbf{F}(\mathbf{x}, t)$ 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.