arXiv:2303.03157v2 [eess.SY] 4 Apr 2023

数据驱动控制与内在李雅普诺夫稳定性

Youngjae Min[1], Spencer M. Richards[2],Navid Azizan[1]

摘要——**基于学习的控制近期进展利用神经网络等深度函数逼近器，对受控动力系统随时间的演化进行建模。然而，学习动力学模型和稳定控制器的问题依然存在，因为即便对于已知的非线性系统，合成稳定反馈律也是一项艰巨任务，更不用说还要拟合数据的复杂参数化表示。为此，我们提出**内在李雅普诺夫稳定性控制（ColLS），**一种从数据中联合学习非线性动力学模型的参数化表示和稳定控制器的方法。为此，我们的方法同时学习一个参数化李雅普诺夫函数，从而内在地约束动力学模型，使其能够由学习得到的控制器稳定。除了通过我们新颖的构造保证学习得到的动力学具有可稳定性外，我们还证明，在对学习得到的动力学保真度作出某些假设的条件下，学习得到的控制器能够稳定真实动力学。最后，我们在多种仿真非线性动力系统上展示了ColLS的有效性。**

I. 引言

数据驱动方法在解决机器人学和自主系统领域复杂的非线性控制问题方面取得了显著成功，例如自主导航、多智能体控制以及物体抓取与操作 [8], [11], [12]。在基于学习的控制问题中，由于缺乏对系统动力学的了解，或对其存在不确定性，控制器综合任务往往更加复杂。为此，人们广泛使用神经网络，在综合控制器之前根据数据进行系统辨识 [4], [7], [14], [15], [24]。然而，使用此类复杂的参数化表示来建模动力学时，很难对学习得到的系统的行为提供任何保证，尤其是难以保证系统在闭环反馈下的稳定性。

稳定非线性动力系统的一种传统方法，是在平衡点附近对系统动力学进行线性化，并使用线性二次调节器（LQR）技术来最小化相对于该平衡点的偏差。LQR 方法能够在一个较小区域内实现闭环稳定性，在该区域中，线性动力学近似是准确的；但一旦远离这一区域，它们就可能彻底失效，尤其是在执行敏捷机动的高度非线性系统中 [22]。如果已知系统的控制李雅普诺夫函数（CLF），则可以综合出经认证具有全局稳定性的非线性控制器 [20]。然而，即使已知动力学，精确构造控制李雅普诺夫函数也可能十分困难，因此近似方法颇受欢迎。例如，对动力学进行多项式近似后，可以通过半定规划（SDP）搜索作为李雅普诺夫函数的平方和（SOS）多项式[23]。然而，多项式近似会对所用函数逼近器的类别施加显著限制。

A. 相关工作

为此，关于动力系统稳定性证书的数据驱动学习，尤其是针对采用复杂参数化表示的动力系统，相关研究文献已大幅增加。此类研究中最早的一项工作，通过惩罚违反相应李雅普诺夫下降条件的情况，为已知的无控制动力系统拟合李雅普诺夫函数 [16]。对于动力系统，利用采样 点上的惩罚项或对稳定性违约施加约束，来学习李雅普诺夫函数、障碍函数和收缩度量函数等证书函数，是文献中普遍采用的思路 [2], [5], [6], [10], [18]。然而，当这类方法与对未知动力学进行回归相结合时，甚至无法保证学习得到的模型是稳定的。为解决这一问题，[13]近期提出了一种联合学习稳定的无控制动力学模型和李雅普诺夫函数的方法。该方法将学习得到的动力学模型限制在由李雅普诺夫下降条件描述的稳定半空间内，从而保证模型的稳定性。然而，将其简单扩展到受控动力学时，会将学习得到的模型限制为可由任意控制输入稳定化，从而妨碍其对一般受控非线性系统进行建模。

对于受控动力学，文献中仍大多沿用基于采样点约束的范式。 [19] 通过采样点处的线性矩阵不等式（LMI）约束，共同学习动力学模型和证书，以正则化动力学模型，使其在较长时间范围内表现良好，但并未学习任何特定的控制器。 [21] 假设动力学已知，并共同学习控制器和收缩度量，通过与稳定性不等式采样点违例相对应的损失项，来证明闭环系统的稳定性。对于基于李雅普诺夫的方法，[3] 针对已知动力学共同学习控制器和李雅普诺夫函数，并利用反例搜索器主动添加违反李雅普诺夫下降条件的训练样本。[25] 将该方法扩展到具有一定稳定性保证的未知控制系统，但其方法要求掌握系统的某些信息，例如其利普希茨常数以及原点附近的线性化模型。此外，其学习得到的动力学无法保证李雅普诺夫函数的存在。 [9] 直接将 [13] 的方法应用于

[1]Youngjae Min 和 Navid Azizan 任职于麻省理工学院信息与决策系统实验室，美国马萨诸塞州剑桥市，邮编 02139。 {yjm, azizan}@mit.edu[2]Spencer M. Richards 任职于斯坦福大学自主系统实验室（ASL），美国加利福尼亚州斯坦福市，邮编 94305。 spenrich@stanford.edu

受控情形，尽管仅适用于执行器矩阵已知的控制仿射系统。总而言之，针对动力学完全未知的系统，高效地同时学习李雅普诺夫函数和非线性控制器，仍是控制与机器人学领域尚未解决的问题。

B. 贡献

在本研究中，我们致力于解决一项艰巨任务：为未知的非线性动力系统学习具有已证明保证的稳定化反馈控制器。为此，我们提出了内在李雅普诺夫稳定性控制（CoILS），这是一种从数据中联合学习受控动力系统模型和反馈控制器的新方法，使得该模型通过构造即可保证能够在学习得到的控制器作用下实现闭环稳定。CoILS 通过同时学习一个参数化李雅普诺夫函数来实现这一点，并利用该函数将开环动力学约束在由学习得到的控制器可实现闭环稳定的动力学子空间内。我们进一步证明，在对学习得到的动力学模型的保真度满足某些假设的条件下，学习得到的控制器也能够保证真实动力学稳定。最后，我们在若干受控非线性动力系统中展示了这种联合学习方法的性能。

II. 问题陈述

本文关注如何控制未知的非线性动力系统

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

其状态为 $x(t) \in \mathcal{X} \subset \mathbb{R}^n$，控制输入为 $u(t) \in \mathcal{U} \subset \mathbb{R}^m$，时刻为 $t \in \mathbb{R}$。尽管我们不知道其动力学 $f : \mathcal{X} \times \mathcal{U} \to \mathbb{R}^n$，但假设可以获取一个有限大小的数据集 $\mathcal{D} = \{(x_i, u_i, \dot{x}_i)\}_{i=1}^N$，其中包含该系统的输入输出测量（1）。

我们希望共同学习动力学 f 以及如何对其进行控制。具体而言，我们希望稳定系统在某个平衡点附近。点 $x_e \in \mathcal{X}$ 是闭环系统 $f_{u^*}(x) := f(x, u^*(x))$ 在反馈控制器 $u^* : \mathcal{X} \to \mathcal{U}$ 作用下的一个平衡点，如果

$$f_{u^*}(x_e) = f(x_e, u^*(x_e)) = 0. \tag{2}$$

稳定性有多种类型；下面我们总结无控制系统和闭环系统中相关的几种定义，也就是说，动力学仅是状态x 的函数。

定义 II.1。对于 $x \in \mathcal{X}$，系统 $\dot{x} = f(x)$ 在其平衡点 $x_e \in \mathcal{X}$ 处是稳定的，如果对于任意 $\epsilon > 0$，都存在$\delta_\epsilon > 0$，使得 $\|x(0) - x_e\|_2 < \delta_\epsilon$ 蕴含 $\|x(t) - x_e\|_2 < \epsilon$对所有 $t \geq 0$均成立。如果系统在 x_e 处相对于$B \subset \mathcal{X}$是渐近稳定的，则它在 x_e处稳定，并且对所有 $x(0) \in B$，$\lim_{t\to\infty} \|x(t) - x_e\|_2 = 0$。如果存在 m，$\alpha > 0$，使得对所有 $x(0) \in B$，$\|x(t) - x_e\|_2 \leq m\|x(0) - x_e\|_2 e^{-\alpha t}$，则系统在 x_e 处相对于$B \subset \mathcal{X}$是指数稳定的。

对于受控系统（1），我们假设存在反馈控制器 $u^* : \mathcal{X} \to \mathcal{U}$，使得由此产生的闭环系统 $\dot{x} = f_{u^*}(x)$ 在x_e 相对于 $\mathcal{X}$指数稳定。不失一般性，我们假设 $x_e = 0$。因此，我们的总体目标是联合学习动力学 f以及一个使系统指数稳定的反馈控制器 u^*。正如我们在第III节所讨论的，我们的方法依靠通过构造在 u^* 中将这种可稳定性编码到 f 中。然而，为此，直接使用定义II.1中的稳定性定义并不方便，因此我们借助李雅普诺夫稳定性理论，引入能够概括动力系统稳定性性质的标量值函数 [1]，[17]。

命题 II.2。考虑系统 $\dot{x} = f(x)$，其中 $x \in \mathcal{X}$。假设存在连续可微函数V：$\mathcal{X} \to \mathbb{R}$，其为正定函数（即，对于 $x \neq 0$，$V(x) > 0$，且 $V(0) = 0$），并满足

$$\nabla_f V(x) := \nabla V(x)^\top f(x) < 0, \tag{3}$$

对于所有 $x \in \mathcal{X} \setminus \{0\}$。则该系统在 $x = 0$ 处关于 $\mathcal{X}$渐近稳定。

这样的函数 V 称为 李雅普诺夫函数。关键思想是，根据条件（3），V 沿着由 f 生成的任意轨迹递减，并最终收敛到 0，这意味着 $x = 0$。具有附加性质的李雅普诺夫函数的存在性，也是指数稳定性的充要条件，具体如下。

命题 II.3。考虑系统 $\dot{x} = f(x)$，其中 $x \in \mathcal{X}$。该系统在 $x = 0$ 处关于 $\mathcal{X}$ 指数稳定，当且仅当存在连续可微函数V：$\mathcal{X} \to \mathbb{R}$，使得

$$c_1\|x\|_2^2 \leq V(x) \leq c_2\|x\|_2^2, \quad \nabla_f V(x) \leq -\alpha V(x), \tag{4}$$

对于所有 $x \in \mathcal{X} \setminus \{0\}$ 以及某些常数 α、c_1、$c_2 > 0$

当命题II.2 和II.3 针对具有控制器 $\dot{x} = f_u(x)$ 的闭环系统 $u : \mathcal{X} \to \mathcal{U}$重新表述时，随附的李雅普诺夫函数 V 也称为控制李雅普诺夫函数（*CLF*）。在下一节中，我们使用命题II.3 来约束 f 的模型，使其能够由参数化控制器 u^* 实现指数稳定化，这由随附的参数化李雅普诺夫函数 V保证。

III. 动力学、控制器和李雅普诺夫函数的联合学习

本文提出了一种如图1 所示的新型架构，其在构造上满足李雅普诺夫稳定性条件。鉴于为单独学习得到的动力学模型构造李雅普诺夫函数十分困难，甚至不可行，我们转而考虑对它们进行联合学习。具体而言，我们提出了一种动力学模型的参数化方法，该方法融合了给定的参数化李雅普诺夫函数和控制器。这种关联使我们能够通过一个损失函数应用自动微分，从而联合优化它们以拟合数据集。

更具体地说，我们通过根据参数化李雅普诺夫函数 V：$\mathcal{X} \to \mathbb{R}$ 对参数化名义模型 f 进行投影，构建动力学模型 $f^* \hat{}$：$\mathcal{X} \times \mathcal{U} \to \mathbb{R}^n$，并

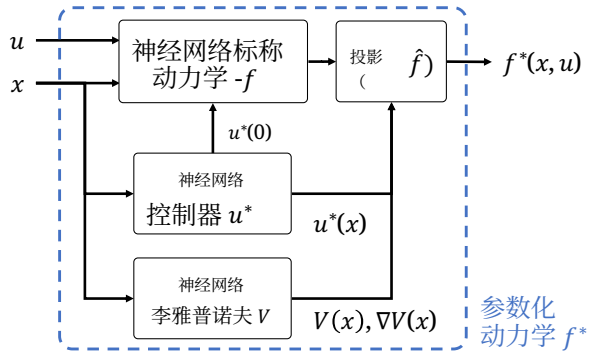


图 1. CoILS 示意图。它通过在架构中融入反馈控制器 u^* 和李雅普诺夫函数 V，对名义动力学模型 f 进行投影，从而在构造时确保稳定性。

构建参数化反馈控制器 $u^*:\ \mathcal{X}\to\mathcal{U}$。该投影约束动力学模型 f^* 满足 Lyapunov稳定性条件

$$\nabla_{f^*_{u^*}}V(x)=\nabla V(x)^\top f^*(x,u^*(x))\le -\alpha V(x) \tag{5}$$

对于所有给定 $\alpha>0$ 的 $x\in\mathcal{X}\setminus\{0\}$。动力学模型的构造如下。

$$f^*(x,u)=\begin{cases}\hat f(x,u) & \text{if } x=0\\ \mathsf{Proj}_{u^*(x),V(x),\nabla V(x)}\left(\hat f(x,\cdot)\right)(u) & \text{o.w.}\end{cases} \tag{6}$$

其中

$$\begin{aligned}&\mathsf{Proj}_{u^*(x),V(x),\nabla V(x)}\left(\hat f(x,\cdot)\right)\\ &:=\hat f(x,\cdot)+\underset{\Delta f\in\mathbb{R}^n}{\arg\min}\quad \|\Delta f\|_2\\ &\qquad\text{s.t.}\quad f^*(x,\cdot)=\hat f(x,\cdot)+\Delta f\\ &\qquad\qquad \nabla V(x)^\top f^*(x,u^*(x))\le-\alpha V(x)\\ &=\hat f(x,\cdot)-\nabla V(x)\frac{\mathsf{ReLU}\left(\nabla V(x)^\top\hat f(x,u^*(x))+\alpha V(x)\right)}{\|\nabla V(x)\|_2^2}\end{aligned} \tag{7}$$

给定 $\nabla V(x)\neq 0$，对于 $x\neq 0$。

对投影的解释如下。对于所有 $\hat{}x\in\mathcal{X}\setminus\{0\}$，如果名义模型 f 满足条件（5），则无需对其进行修改，因此 $f^*(x,\ u)=f(x,\ u)$ 对所有$u\in\mathcal{U}$均成立；另一方面，如果它违反该条件，我们会对其进行最小幅度的调整（按 ℓ^2 范数衡量），使投影模型 f^* 恰好以等号满足条件（5）。如第 IV 节所述，通过构造，该投影模型与反馈控制器 u^* 配合时，可保证在原点全局指数稳定。

我们使用神经网络来表示标称动力学模型 f、反馈控制器 u^* 和李雅普诺夫函数 V。为确保这些组件各自满足所需条件，我们谨慎地选择了相应神经网络的架构，如下所述。

A. 标称动力学模型与控制器

首先，我们应确保原点是闭环系统 $f^*_{u^*}$ 的平衡点，即 $f^*(0,\ u^*(0))=\hat{}f(0,\ u^*(0))=0$。这可以通过将标称动力学模型设为以下形式来实现：

$$\hat f(x,u)=g_f(x,u)-g_f(0,u^*(0)) \tag{8}$$

其中 $g_f:\ \mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ 是一个架构任意的无约束神经网络。

对于控制器，我们要求候选控制器的取值范围限制在预先指定的集合 $\mathcal{U}$ 内。在大多数机器人应用中，控制输入通常由电机产生，因此会受到物理限制而饱和。为处理这类常见情况，我们考虑 $\mathcal{U}=\{u\in\mathbb{R}^m:-u_{\lim}\le u\le u_{\lim}\}$，其中 $u_{\lim}\ge 0$。然后，我们通过设置以下内容来限制反馈控制器的取值范围

$$u^*(x)=\mathsf{diag}(u_{\lim})\tanh(g_u(x)), \tag{9}$$

其中，$\mathsf{diag}(u_{\lim})$ 是一个对角线元素为 $u_{\lim}$ 的对角矩阵，而 $g_u:\mathbb{R}^n\to\mathbb{R}^m$ 是一个无约束神经网络。

B. 李雅普诺夫函数

李雅普诺夫函数的适用条件有三项。首先，V应为正定，即对于 $x\neq 0$，$V(x)>0$，且 $V(0)=0$。我们通过将V定义为以下形式来确保这一条件

$$V(x)=\sigma(g_V(x)-g_V(0))+\epsilon_{\mathrm{pd}}\|x\|^2 \tag{10}$$

对于某个 $\epsilon_{\mathrm{pd}}>0$，其中 $g_V:\ \mathbb{R}^n\to\mathbb{R}$ 是一个神经网络，$\sigma:\ \mathbb{R}\to\mathbb{R}$ 是一个函数，且对于所有 $w\le 0$，$\sigma(w)=0$。

此外，V应当连续可微。我们通过使用平滑ReLU函数来确保这一性质

$$\sigma(z)=\begin{cases}0 & \text{if } x\le 0\\ x^2/2d & \text{if } 0<x<d\\ x-d/2 & \text{otherwise}\end{cases} \tag{11}$$

在（10）中，并将其用作神经网络 g_V 的激活函数。

最后，V不应存在非零临界点，因为对于任意非零状态，其时间导数都严格为负。确保这一性质的一种方法是使用 g_V 的输入凸神经网络（ICNN），如 [13] 所介绍。结合严格递增的 σ，V严格凸且不存在非零临界点。然而，李雅普诺夫函数不一定是凸函数，使用 ICNN 会限制 V 的模型容量。因此，我们考虑对 g_V 使用无约束神经网络，并将式（7）中的投影近似为

$$\hat f(x,u)-\nabla V(x)\frac{\mathsf{ReLU}\left(\nabla_{\hat f}V(x,u^*(x))+\alpha V(x)\right)}{\max(\|\nabla V(x)\|_2^2,\epsilon_{\mathrm{proj}})} \tag{12}$$

对于某个较小的 $\epsilon_{\mathrm{proj}}>0$。网络训练完成后，若 (5) 中的 $V(x)$ 是正定的，则意味着 $\nabla V(x)$ 必定非零，因为 $\nabla_{f^*_{u^*}}V(x)$ 严格为负。那么，近似式 (12) 也将恢复原始投影。

C. 损失函数

我们通过最小化以下损失函数，联合学习这三个组成部分：

$$\mathcal{L}(\theta)=\frac{1}{N}\sum_{(x,u,\dot{x})\in\mathcal{D}}k(u,u^*(x))\|\dot{x}-f^*(x,u)\|^2+\lambda\|\theta\|_2^2 \tag{13}$$

其中，θ、$k(\cdot,\cdot)$和λ分别表示神经网络、核函数和正则化系数的参数。核函数$k:\mathcal{U}\times\mathcal{U}\to\mathbb{R}$在两个参数越接近时取值越大。通过对$\ell^2$差异加权，我们鼓励$u^*$学习能够生成模型误差更小的轨迹的策略。相反，我们也鼓励f^*在可行轨迹附近学习得更好，即$(x,u^*(x))$。在本文中，我们使用以下核函数：

$$k(u,u')=1+\exp(-\beta\|u-u'\|_2^2). \tag{14}$$

IV. 稳定性保证

本节将严格分析学习得到的模型的稳定性性质。我们首先证明所构造模型f^*与反馈控制器u^*的固有稳定性（定理 IV.1）。随后，我们根据学习性能，将固有稳定性与真实系统f的行为联系起来（定理 IV.2）；证明见附录 B-C 。

定理 IV.1（固有稳定性）。考虑由（6）定义的投影动力学f^*的闭环系统$f_{u^*}^*$，其中$\hat{}f$见（8），u^*见（9），且V见（10）。对于任意$r_2\geq r_1>0$，令$B_{r_1,\,r_2}:=\{x\in\mathcal{X}:\ r_1\leq\|x\|_2\leq r_2\}$。则，

1）$f_{u^*}^*$关于$B_{r_1,\,r_2}$在原点处指数稳定。*2）*若存在$c>0$使得对所有$x\in\mathcal{X}$，$V(x)\leq c\|x\|_2^2$，则$f_{u^*}^*$关于$\mathcal{X}$在原点处指数稳定。

CoILS 如定理 IV.1 所证明，固有地保证学习得到的模型的指数稳定性。请注意，该性质即使对于模型的任意初始化也成立。接下来，我们进一步讨论其与真实动力学的联系。当学习性能达到令人满意的水平时，我们可以预期这一稳定性性质能够迁移至真实动力学。我们将在下一个定理中详细阐述这一关系。

定理 *IV.2.* 考虑式 *(1)* 中的动力系统f及其在式 *(6)* 中的学习得到的模型$\hat{}f^*$，其中f、u^*和V分别见式 *(8)*、*(9)* 和 *(10)*。假设f和f^*分别是利普希茨连续的，且其利普希茨常数分别为L_f和L_{f^*}。定义如下，令δ和e分别表示f^*学习得到的密集程度和准确度：

$$\begin{aligned}\delta&:=\sup_{x\in\mathcal{X}}\min_{(y,v)\in\mathcal{D}}\|(x,u^*(x))-(y,v)\|_2,\\ e&:=\max_{(y,v)\in\mathcal{D}}\|f(y,v)-f^*(y,v)\|_2.\end{aligned} \tag{15}$$

对于任意$r>0$，令原点邻域$N_r:=\{x\in\mathcal{X}:\ \|x\|_2<r\}$且$M_r:=\sup_{x\in\mathcal{X}\setminus N_r}\|\nabla V(x)\|_2$。那么，闭环系统$f_{u^*}$总会到达$N_r$，也就是说，对于由$f_{u^*}$从$x(0)\in\mathcal{X}$生成的任意轨迹$x(t)$，

表 I 实验的超参数

参数	值
α	1.0 或 5.0
β	$5/u_{\lim}$
λ	0.0
ϵ_{pd}	0.5
ϵ_{proj}	0.001
d	0.005

存在$T\geq 0$使得$x(T)\in N_r$，如果δ和e足够小，使得

$$(L_f+L_{f^*})\delta+e<\frac{\alpha\epsilon_{\mathrm{pd}}r^2}{M_r}. \tag{16}$$

V. 实验

在本节中，我们展示CoILS在稳定未知非线性控制系统方面的有效性。我们首先验证所提出架构所获得的投影模型的稳定性性质（第V-A节）。然后，我们分别在三个不同的控制问题中展示学习得到的控制器和动力学的性能与准确性（第V-B-V-D节）。

对于每种情形，我们使用$N=10^5$个数据元组训练模型，这些数据元组在$\mathcal{X}=\{x\in\mathbb{R}^n:\ x_{lb}\leq x\leq x_{ub}\}$和$\mathcal{U}=\{u\in\mathbb{R}^m:\ -u_{\lim}\leq u\leq u_{\lim}\}$范围内均匀采样。模型采用神经网络构建，如第III节所述。我们对g_f、g_u和g_V使用三层全连接神经网络（FCN）。它们的每个隐藏层分别包含100、50和50个隐藏神经元。对于g_V，我们在输出层添加乘以10的 tanh激活函数，以限制李雅普诺夫函数的尺度。模型使用带梯度裁剪的小批量梯度下降优化器进行训练，学习率为0.0001。实验中使用的其他超参数见表I。

A. 随机网络

在本实验中，我们研究了对随机初始化模型进行投影的有效性。正如定理 IV.1所证明的，我们验证了投影模型f^*在（6）中即使在随机初始化的情况下，与反馈控制器u^*构成闭环时也具有固有稳定性。我们对$\mathcal{X}\in\mathbb{R}^2$和$\mathcal{U}\in\mathbb{R}$的神经网络进行初始化。结果如图2所示。投影模型$f^*$在控制器$u^*$的作用下稳定于原点，而名义模型$\hat{}\hat{}f$则不会。根据（8），名义模型$f$和投影模型$f^*$在与$u^*$构成闭环时，均在原点处取零值。这确保了原点是闭环系统的平衡点。有趣的是，我们观察到，尽管在（V中使用通用全连接神经网络作为g_V（10）中的候选李雅普诺夫函数，如图所示，它并没有任何临界点。这种有利的构造在大多数情况下都能观察到。

B. 范德波尔振荡器

我们在三个不同的控制问题中测试所提出的方法。首先，我们尝试稳定一个系统——范德波尔振荡器。该系统即使没有任何

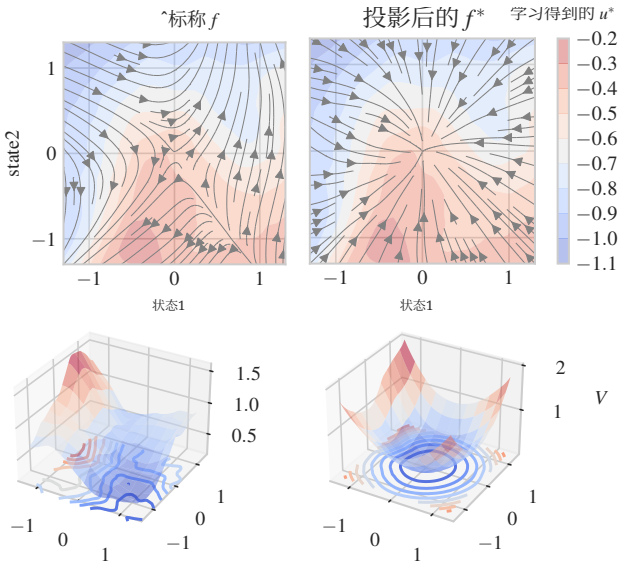


图 2. 随机初始化模型的闭环动力学和李雅普诺夫函数。左上：带控制器 u^* 的名义模型 $\hat{}\, f$。右上：带控制器 u^* 的投影模型 f^*。左下：神经网络 g_V。右下：候选李雅普诺夫函数 V。

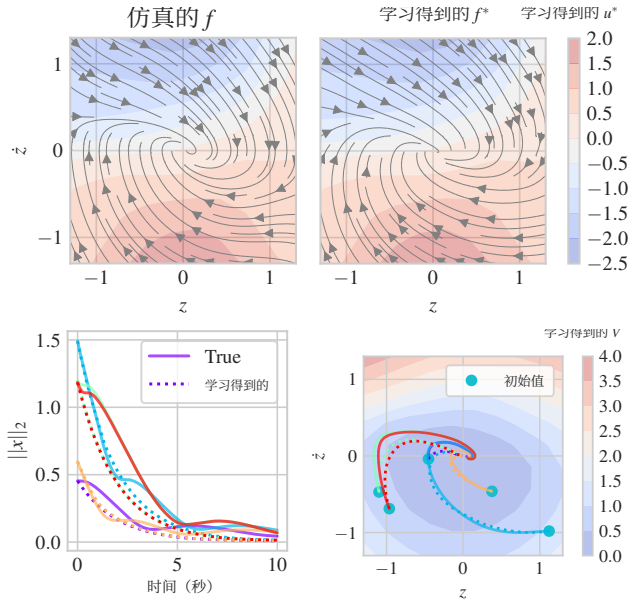


图 3. 范德波尔振荡器的仿真与学习得到的模型对比。左上：真实模型 f 与学习得到的控制器 u^*的闭环动力学。右上：学习得到的模型 f^* 与学习得到的控制器 u^*的闭环动力学。左下：真实系统与学习得到的系统中 5 条随机初始化轨迹的对比。右下：状态空间中的轨迹图，以及学习得到的李雅普诺夫函数的等高线图。

控制输入，也已在原点处稳定。自治范德波尔振荡器是一种众所周知的非线性系统，会以极限环的形式呈现稳定振荡。通过添加控制输入 $u \in \mathbb{R}$，我们将真实动力学视为

$$\ddot{z} = u - z + \mu(1 - z^2)\dot{z} \tag{17}$$

对于状态 $x = [z, \dot{z}] \in \mathbb{R}^2$ 和参数 μ=1。数据集以 x_{lb}=[−1.3, −1.3]，x_{ub}=[1.3, 1.3]，$u_{\lim}$=5进行采样。结果如图3所示。真实系统 f_{u^*} 和学习得到的系统 $f^*_{u^*}$ 的闭环动力学表现出相似的行为。由于存在一些模型误差，轨迹并未完全重合，但学习得到的控制器成功地使真实系统和学习得到的系统都稳定在原点。

C. 倒立摆

接下来，我们展示 CoILS 在一个更具挑战性的控制问题中的表现：稳定倒立摆。图 4所示的倒立摆若没有适当的控制，会因重力作用而轻易偏离原点。对于倒立位置的角位置θ 及其角速度 θ，我们考虑状态 $x = [\theta, \dot{\theta}] \in \mathbb{R}^2$。通过在支点施加扭矩 $u \in \mathbb{R}$作为控制输入，我们将真实动力学表示为

$$\ddot{\theta} = \frac{mgl \sin\theta + u - b\dot{\theta}}{ml^2}. \tag{18}$$

我们设定参数 $m = 0.15$、$g = 9.81$、$l = 0.5$、$b = 0.1$，其中 $x_{lb} = [-4, -4]$、$x_{ub} = [4, 4]$、$u_{\lim} = 5$。结果如图5所示。真实系统 f_{u^*} 与学习得到的系统 $f^*_{u^*}$的闭环动力学总体行为相似ℓ^2 ，但轨迹呈现出十分不同的范数演化。这种差异源于 $\theta = 0$周围小区域内的 f_{u^*} 模型误差。这些区域内 f_{u^*} 的幅值较小，减缓了状态

尽管两个系统的运动方向相似，但运动的方向相似。这些模型误差可以通过利用状态元素之间的物理关系得到改善，即第二个元素是第一个元素的导数。

D. 自行车路径跟踪

前面的两个例子都是控制仿射系统，而 CoILS 适用于不具备此类结构的一般控制系统。在本实验中，我们评估 CoILS 在一种控制输入会产生非线性影响的控制系统中的表现。我们考虑控制一辆恒速自行车沿单位圆路径行驶的问题，如图 4所示。我们的目标是通过控制转向角 d_e ，使距离误差 θ_e和角度误差 $u \in \mathbb{R}$均为零。对于状态 $x = [d_e, \theta_e] \in \mathbb{R}^2$，其动力学表示为

$$\begin{aligned} \dot{d_e} &= v \sin\theta_e, \\ \dot{\theta_e} &= \frac{v \tan u}{L} - \frac{v \cos\theta_e}{1 - d_e}. \end{aligned} \tag{19}$$

我们设置参数 $v = 6$，$L = 1$ 其中 $x_{lb} = [-0.8, -0.8]$，$x_{ub} = [0.8, 0.8]$，$u_{\lim} = 0.4\pi$。结果如图6所示。真实系统 f_{u^*}与学习得到的系统 $f^*_{u^*}$ 的闭环动力学彼此相似。此外，学习得到的控制器在两个系统上生成的轨迹都很接近，并成功将两个系统稳定在原点。

VI. 结论

我们提出了一种新的数据驱动方法 CoILS，用于稳定未知受控动力系统。我们联合学习动力学模型和带有李雅普诺夫

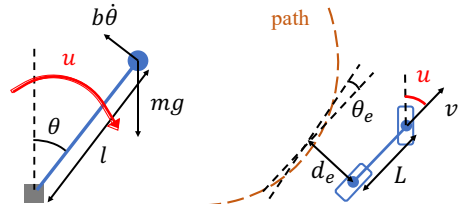


图 4. 倒立摆（左）和自行车路径跟踪系统（右）的示意图。

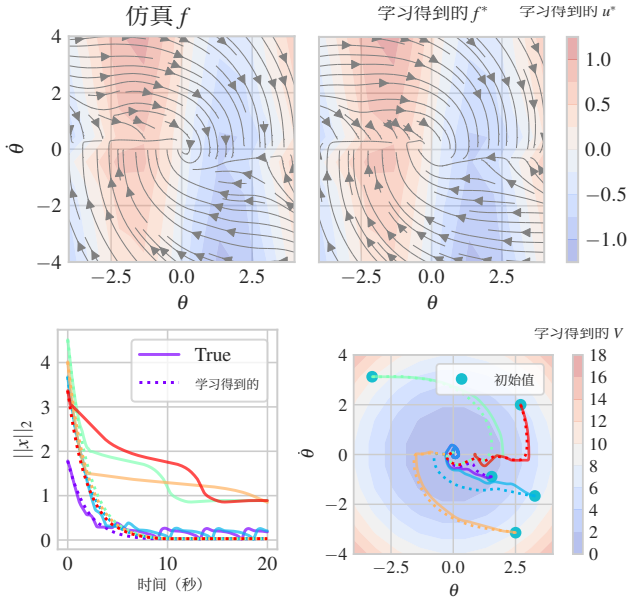


图 5. 倒立摆系统的仿真结果与学习得到的模型的比较。左上：真实模型 f 的闭环动力学。右上：学习得到的模型 f^* 的闭环动力学。左下：5 条随机初始化轨迹的比较。右下：状态空间中的轨迹图，以及学习得到的李雅普诺夫函数的等高线图。

函数的反馈控制器，从而通过构造保证投影模型能够在闭环中由学习得到的控制器实现稳定。我们进一步证明，在对学习得到的动力学模型保真度作出一定假设的情况下，学习得到的控制器也能保证稳定真实动力学。我们通过对多个受控非线性动力系统进行仿真，展示了该方法的性能。

未来有几个值得探索的方向。首先，可以探索不同的损失函数，例如纳入面向控制的指标。本文重点在于学习动力学模型，而没有显式优化控制器的性能。由于实践中存在多种能够稳定系统的控制器，因此引导学习过程找到性能更优的控制器将是一个有趣的方向。此外，探索本方法的不同超参数将有助于进一步优化我们的方法。还可以考虑将本方法扩展到部分可观测系统。寻找能够稳定未知动力系统的输出反馈控制器虽有挑战，但值得尝试。

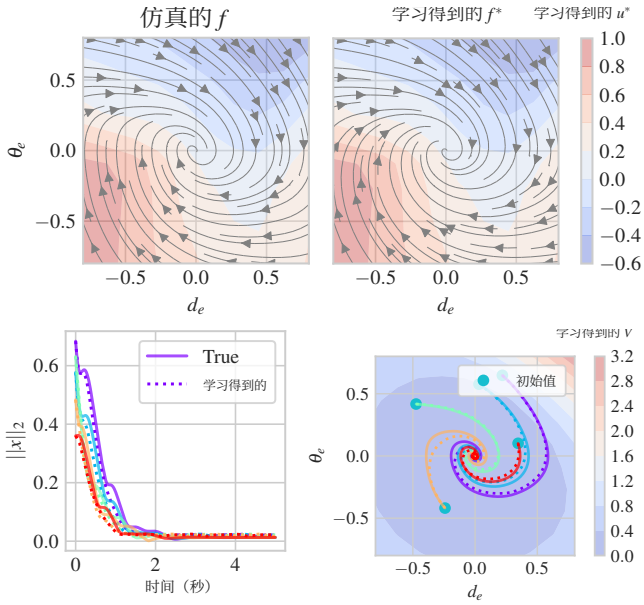


图 6. 自行车路径跟踪系统的仿真与学习得到的模型的比较。左上：真实模型 f 与学习得到的控制器 u^*的闭环动力学。右上：学习得到的模型 f^* 与学习得到的控制器 u^*的闭环动力学。左下：真实系统与学习得到的系统中 5 条随机初始化轨迹的比较。右下：状态空间中的轨迹图，以及学习得到的李雅普诺夫函数的等高线图。

参考文献

[1] F. Blanchini 和 S. Miani。《控制中的集合论方法》，第 78 卷。Springer，2008年。[2] N. M. Boffi、S. Tu、N. Matni、J.-J. E. Slotine 和 V. Sindhwani。从数据中学习稳定性证书。载于《机器人学习会议》，2020年。[3] Y.-C. Chang、N. Roohi 和 S. Gao。神经李雅普诺夫控制。载于 Advances in Neural Information Processing Systems，第 32 卷，2019年。[4] S. Chen、S. A. Billings 和 P. Grant。使用神经网络进行非线性系统辨识。《国际控制杂志》，51(6):1191–1214，1990年。[5] C. Dawson、S. Gao 和 C. Fan。基于学习证书的安全控制：神经李雅普诺夫方法、障碍函数方法和收缩方法综述。可从 https://arxiv.org/abs/2202.11762 获取，2022年。[6] B. El Khadir、J. Varley 和 V. Sindhwani。具有连续时间安全性的遥操作器模仿。载于 Robotics: Science and Systems，2019年。[7] J. K. Gupta、K. Menda、Z. Manchester 和 M. J. Kochenderfer。用于机器人学习与控制的结构化机械模型。载于动力学与控制学习会议，2020年。[8] L. Hewing、K. P. Wabersich、M. Menner 和 M. N. Zeilinger。基于学习的模型预测控制：迈向控制中的安全学习。《控制、机器人与自主系统年度评论》，3:269–296，2020年。[9] K. Kashima、R. Yoshiuchi 和 Y. Kawano。学习可稳定化深度动力学模型。可从 https://arxiv.org/abs/2203.09710 获取，2022年。[10] M. Khosravi 和 R. S. Smith。基于吸引域先验知识的非线性系统辨识。《IEEE控制系统快报》，5(3):1091–1096，2021年。[11] S. Levine、P. Pastor、A. Krizhevsky、J. Ibarz 和 D. Quillen。结合深度学习和大规模数据采集，学习机器人抓取中的手眼协调。《国际机器人研究杂志》，37(4-5):421–436，2018年。[12] S. Liu、G. Lever、Z. Wang、J. Merel、S. A. Eslami、D. Hennes、W. M. Czarnecki、Y. Tassa、S. Omidshafiei、A. Abdolmaleki 等。从运动控制到模拟人形机器人足球中的团队协作。《科学·机器人学》，7(69):eabo0235，2022年。[13] G. Manek 和 J. Z. Kolter。学习稳定深度动力学模型。载于神经信息处理系统会议，2019年。

[14] S. M. Richards、N. Azizan、J.-J. E. Slotine 和 M. Pavone。面向非线性系统的自适应控制导向元学习。载于 Robotics: Science and Systems，2021年。[15] S. M. Richards、N. Azizan、J.-J. E. Slotine 和 M. Pavone。面向控制的元学习。《国际机器人研究杂志》，2023年。即将刊出。[16] S. M. Richards、F. Berkenkamp 和 A. Krause。李雅普诺夫神经网络：用于安全学习动力系统的自适应稳定性认证。载于《机器人学习会议》，2018年。[17] S. Sastry。《非线性系统：分析、稳定性与控制》，第 10 卷。Springer Science & Business Media，2013年。[18] V. Sindhwani、S. Tu 和 M. Khansari。学习收缩向量场，实现稳定模仿学习。可从 https://arxiv.org/abs/1804.04878 获取，2018年。[19] S. Singh、S. M. Richards、V. Sindhwani、J.-J. E. Slotine 和 M. Pavone。通过基于收缩的正则化学习可稳定化非线性动力学。《国际机器人研究杂志》，2020年。[20] J.-J. E. Slotine 和 W. Li。《应用非线性控制》。Prentice Hall，1991年。[21] D. Sun、S. Jha 和 C. Fan。使用收缩度量学习经认证的控制。载于《机器人学习会议》，2020年。[22] S. Sun、A. Romero、P. Foehn、E. Kaufmann 和 D. Scaramuzza。四旋翼敏捷飞行中非线性模型预测控制与基于微分平坦性的控制的比较研究。IEEE Transactions on Robotics，38(6):3357–3373，2022年。[23] R. Tedrake、I. R. Manchester、M. Tobenkin 和 J. W. Roberts。LQR 树：通过平方和验证进行反馈运动规划。《国际机器人研究杂志》，29(8):1038–1052，2010年。[24] Y. Wang。一种利用 LSTM 神经网络进行动态系统辨识的新方法。载于 2017年 American Control Conference (ACC)，第 5324–5329 页。IEEE，2017年。[25] R. Zhou、T. Quartz、H. De Sterck 和 J. Liu。对具有稳定性保证的未知非线性系统进行神经李雅普诺夫控制。载于 Advances in Neural Information Processing Systems，2022年。

附录

A. 特例：控制仿射系统

当已知真实动力学（1）具有控制仿射结构时，我们无需显式学习反馈控制器，即可学习其李雅普诺夫控制。考虑标准形式的控制仿射系统

$$\dot{x}(t) = f_1(x(t)) + f_2(x(t))u(t) \tag{20}$$

对于 $x(t) \in \mathcal{X} \subset \mathbb{R}^n$，$u(t) \in \mathcal{U} \subset \mathbb{R}^m$。利用这一结构知识，我们也可以构造控制仿射的投影模型f^*，即 $f^*(x, u) = f_1^*(x) + f_2^*(x)u$。于是，李雅普诺夫稳定性条件（5）变为

$$\nabla V(x)^\top (f_1^*(x) + f_2^*(x)u^*(x)) \leq -\alpha V(x) \tag{21}$$

对于所有 $x \in \mathcal{X} \setminus \{0\}$，给定 $\alpha > 0$。由于具有控制仿射结构，我们可以轻松计算出使（21）左侧最小化的控制器，而无需直接学习它：

$$u^*(x) = \arg\min_{u \in \mathcal{U}} \nabla V(x)^\top f_2^*(x)u. \tag{22}$$

考虑 $\mathcal{U} = \{u \in \mathbb{R}^m : -u_{\lim} \leq u \leq u_{\lim}\}$ 对于某个 $u_{\lim} \geq 0$，（22）可简化为

$$u^{*(i)}(x) = -\text{sign}(\{\nabla V(x)^\top f_2^*(x)\}^{(i)})u_{\lim}^{(i)} \tag{23}$$

对于每个 $i = 1, 2, \ldots, m$ 其中$(\cdot)^{(i)}$表示 i-th 分量。然后，我们可以像之前一样，仅优化动力学模型和李雅普诺夫函数的神经网络，并采用诱导的反馈控制器（23）。

只需将名义模型设为 f^*，即可轻松构造出控制仿射的投影模型

$$\hat{f}(x, u) = \hat{f}_1(x) + \hat{f}_2(x)u$$

会诱导出$\hat{f}_2^* = f_2$。为确保在该结构下原点是平衡点（如（8）所示），我们可以设为

$$\begin{aligned} \hat{f}_2(x) &= g_{f_2}(x), \\ \hat{f}_1(x) &= g_{f_1}(x) - g_{f_1}(0) - \hat{f}_2(0)u^*(0), \end{aligned} \tag{24}$$

其中 $g_{f_1} : \mathbb{R}^n \to \mathbb{R}^n$ 和 $g_{f_2} : \mathbb{R}^n \to \mathbb{R}^{n \times m}$ 是无约束神经网络。

定理*IV.1* 的证明

证明。我们通过根据李雅普诺夫函数$V(x(t))$的指数衰减上界，推导出 $\|x(t)\|_2$ 的指数衰减上界，来证明闭环系统 $f_{u^*}^*$ 的指数稳定性。

首先，我们根据投影模型所满足的李雅普诺夫条件（5），计算 V的上界。令 $x(t)$为由闭环系统 $f_{u^*}^*$ 生成的轨迹，其初始状态为 $x(0) \in \mathcal{X}$。那么，对于所有 $t \geq 0$,

$$\begin{aligned} \frac{d}{dt}V(x(t)) &= \nabla V(x(t))^\top f^*(x(t), u^*(x(t))) \\ &\leq -\alpha V(x(t)). \end{aligned} \tag{25}$$

对该不等式进行积分，可得上界为

$$V(x(t)) \leq V(x(0))e^{-\alpha t}. \tag{26}$$

接下来，我们根据 $\|x(t)\|_2$ 由（26）推导其指数衰减上界。由V在（10）中的正定结构可知，$V(x) \geq \epsilon_{\text{pd}}\|x\|_2^2$，这意味着

$$\|x(t)\|_2 \leq \sqrt{\frac{V(x)}{\epsilon_{\text{pd}}}} \leq \sqrt{\frac{V(x(0))e^{-\alpha t}}{\epsilon_{\text{pd}}}}. \tag{27}$$

然后，剩下的步骤是将上界表示为 $\|x(0)\|_2$ 的形式。给定 $r_2 \geq r_1 > 0$，我们定义

$$M = \sup_{x \in \mathbb{R}^n} \frac{V(x)}{\|x\|_2^2} \quad \text{s.t. } r_1 \leq \|x\|_2 \leq r_2. \tag{28}$$

（28）中的优化可行集是闭且有界的，因此是紧集。于是，上确界 M 在该可行集上可以取得，因此是有限的。由于可行集包含B_{r_1, r_2}，（28）意味着对于任意 $x(0) \in B_{r_1, r_2}$，$V(x(0)) \leq M\|x(0)\|_2^2$。结合（27），

$$\|x(t)\|_2 \leq \sqrt{\frac{V(x(0))e^{-\alpha t}}{\epsilon_{\text{pd}}}} \leq \sqrt{\frac{M}{\epsilon_{\text{pd}}}}\|x(0)\|_2 e^{-\alpha t/2}. \tag{29}$$

因此，系统 $f_{u^*}^*$ 在原点处相对于 B_{r_1, r_2} 指数稳定。

如果存在 $c > 0$ 使得 $V(x) \leq c\|x\|_2^2$ 对所有$x \in \mathcal{X}$均成立，则同样地（或根据命题 II.3），系统$f_{u^*}^*$关于 $\mathcal{X}$在原点处指数稳定。 □

C. 定理证明IV.2

证明。假设（16）成立，并表示差分 e as

$$\Delta := \frac{\alpha\epsilon_{\mathrm{pd}} r^2}{M_r} - (L_f + L_{f^*})\delta - e > 0. \tag{30}$$

我们首先证明，对于真实闭环系统 f_{u^*}，沿着 V 在 N_r 之外的任意轨迹， V 都严格递减。然后，我们用反证法证明它们会到达 N_r 。

对于所有 $x \in \mathcal{X} \setminus N_r$，由于下式成立， V在闭环系统 $f^*_{u^*}$ 中严格递减

$$\nabla_{f^*_{u^*}} V(x) \le -\alpha V(x) \le -\alpha\epsilon_{\mathrm{pd}}\|x\|^2 \le -\alpha\epsilon_{\mathrm{pd}} r^2 < 0. \tag{31}$$

然后，我们通过对差值进行界定，证明 V 在真实系统f_{u^*} 中也严格递减

$$\begin{aligned} &\nabla_{f_{u^*}} V(x) - \nabla_{f^*_{u^*}} V(x) \\ &\quad \le \|\nabla V(x)\|_2 \|f(x, u^*(x)) - f^*(x, u^*(x))\|_2. \end{aligned} \tag{32}$$

学习得到的闭环系统的泛化误差可界定为

$$\begin{aligned} &\|f(x, u^*(x)) - f^*(x, u^*(x))\|_2 \\ &\le \|f(x, u^*(x)) - f(y, v)\|_2 + \|f(y, v) - f^*(y, v)\|_2 \\ &\qquad + \|f^*(y, v) - f^*(x, u^*(x))\|_2 \\ &\le L_f\delta + e + L_{f^*}\delta = (L_f + L_{f^*})\delta + e \end{aligned} \tag{33}$$

其中 $(y,\ v)$ 是 $\mathcal{D}$ 中与 $(x,\ u^*(x))$最近的数据样本。第一个不等式由三角不等式得到，第二个不等式则由 L_f、 L_{f^*}、 δ、 e 的定义得到。将上述不等式与（30）结合，可得

$$\begin{aligned} \nabla_{f_{u^*}} V(x) &\le \nabla_{f^*_{u^*}} V(x) \\ &\quad + \|\nabla V(x)\|_2 \|f(x, u^*(x)) - f^*(x, u^*(x))\|_2 \\ &\le -\alpha\epsilon_{\mathrm{pd}} r^2 + M_r((L_f + L_{f^*})\delta + e) = -M_r\Delta. \end{aligned} \tag{34}$$

现在，我们用反证法证明所有轨迹都会到达 N_r 。假设存在 $x(0) \in \mathcal{X} \setminus N_r$ ，使得闭环系统 f_{u^*} 对所有 $t \ge 0$ 都生成一条轨迹 $x(t) \notin N_r$。那么， $\|x(t)\|_2 \ge r$ 意味着

$$V(x(t)) \ge \alpha\epsilon_{\mathrm{pd}} r^2. \tag{35}$$

然而，由于 $x(t) \notin N_r$，对于所有$t \ge 0$，都有 $\dot{V}(x(t)) \le -M_r\Delta < 0$ 。从 $V(x(0))$ 积分可得，

$$\begin{aligned} V(x(t)) &= V(x(0)) + \int_0^t \dot{V}(x(\tau))d\tau \\ &\le V(x(0)) - M_r\Delta t. \end{aligned} \tag{36}$$

这意味着存在足够大的 T使得 $V(x(T)) < \alpha\epsilon_{\mathrm{pd}} r^2$ ，这与（35）矛盾。因此， $x(t)$ 到达 N_r。

□