北京航空航天大學
BEIHANG UNIVERSITY

LITERATURE READING

Lyapunov-stable neural-network control

Hongkai Dai, Benoit Landry, Lujie Yang, Marco Pavone, and Russ Tedrake

Robotics: Science and Systems (RSS) 2021

Jinjie LI

School of Automation Science and Electrical Engineering
Beihang University

October 16, 2021

TOYOTA RESEARCH INSTITUTE

Stanford University

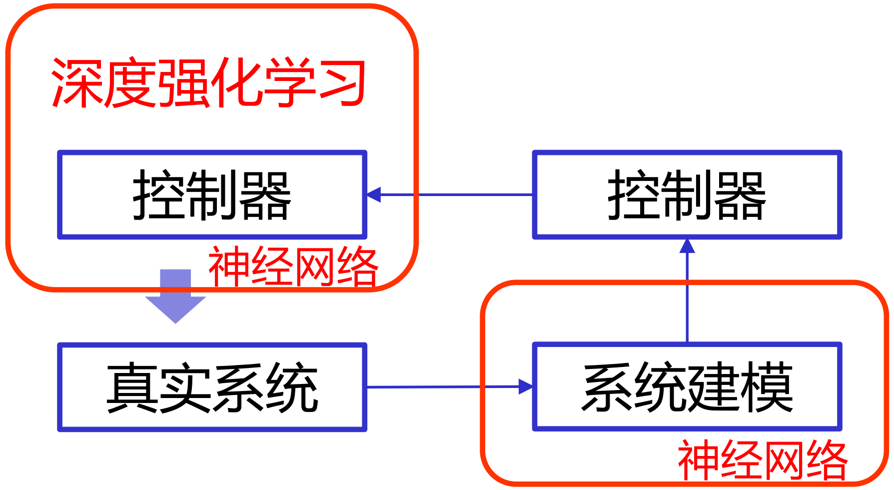
Massachusetts Institute of Technology

目录

- ❑ 背景
- ❑ 挑战
- ❑ 方法
- ❑ 结果
- ❑ 总结
- ❑ 推荐

背景

Hwangbo et al., 2017

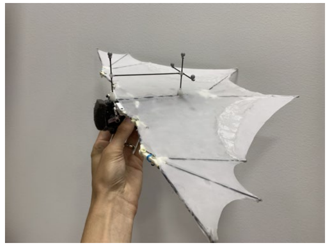


Kabzan et al., 2021

Lee et al., 2020

Finn et al., 2016

Torrente et al., 2021



Hoff et al., 2021

- ❑ **Q：神经网络可否被应用于航空航天领域？**

安全性

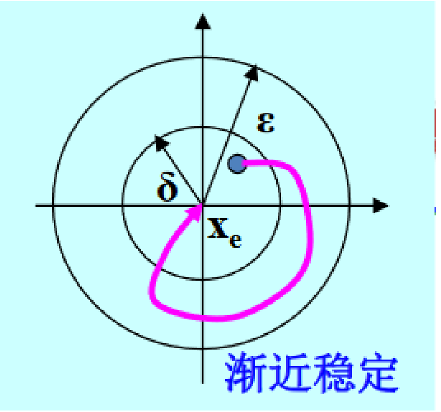
背景

- ❑ **Q：神经网络可否被应用于航空航天领域?**

  系统的安全性 = 工程意义下的稳定

  闭环系统 → 李雅普诺夫意义下渐近稳定

  常用工具：李雅普诺夫第二定理

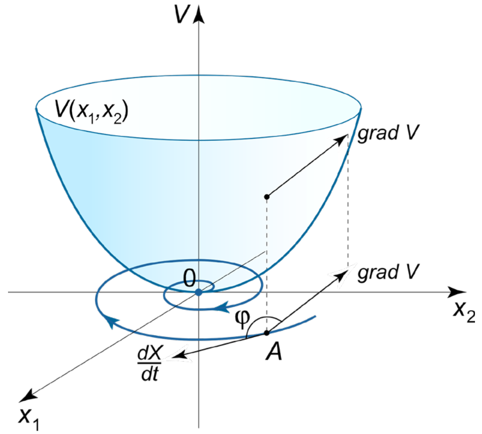


结论 5.13［小范围渐近稳定性定理］ 对连续时间非线性时变自治系统(5.46)，若可构造对 x 和 t 具有连续一阶偏导数的一个标量函数 $V(x,t)$，$V(0,t)=0$，以及围绕状态空间原点的一个吸引区 Ω，使对所有非零状态 $x\in\Omega$ 和所有 $t\in[t_0,\infty)$ 满足如下条件：

(i) $V(x,t)$为正定且有界；

(ii) $\dot{V}(x,t)\triangleq dV(x,t)/dt$ 为负定且有界；

则系统原点平衡状态 $x=0$ 在 Ω 域内为一致渐近稳定。

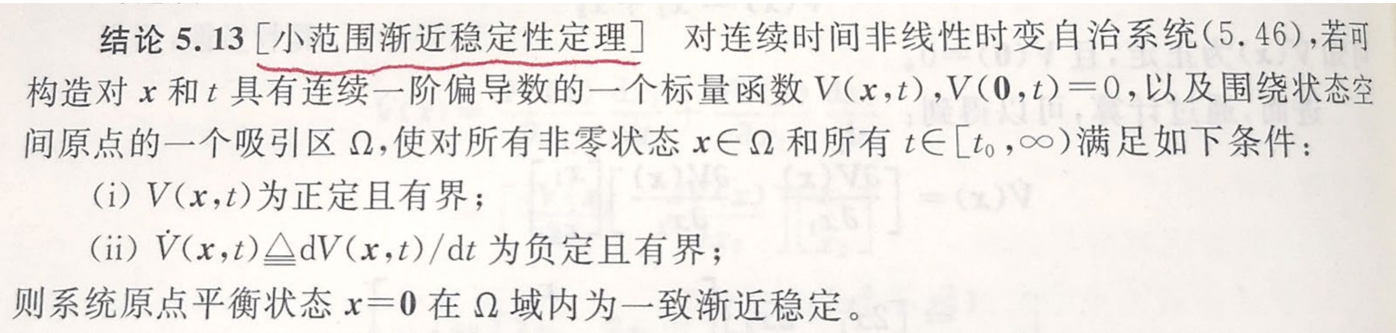


- ❑ **Challenge：是否存在神经网络控制器，满足李雅普诺夫第二定理?**

挑战

- ❑ **Challenge：是否能找到神经网络控制器，满足李雅普诺夫第二定理?**

难点：李雅普诺夫函数的选取



结论 5.13 [小范围渐近稳定性定理] 对连续时间非线性时变自治系统(5.46)，若可构造对 $\boldsymbol{x}$ 和 t 具有连续一阶偏导数的一个标量函数 $V(\boldsymbol{x},t)$，$V(\boldsymbol{0},t)=0$，以及围绕状态空间原点的一个吸引区 Ω，使对所有非零状态 $\boldsymbol{x}\in\Omega$ 和所有 $t\in[t_0,\infty)$ 满足如下条件：

(i) $V(\boldsymbol{x},t)$ 为正定且有界；

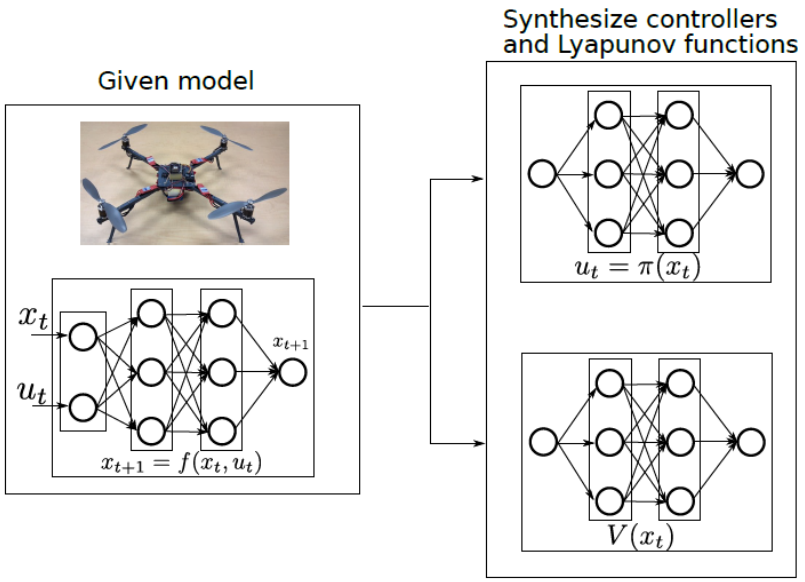
(ii) $\dot{V}(\boldsymbol{x},t)\triangleq \mathrm{d}V(\boldsymbol{x},t)/\mathrm{d}t$ 为负定且有界；

则系统原点平衡状态 $\boldsymbol{x}=\boldsymbol{0}$ 在 Ω 域内为一致渐近稳定。

"Lyapunov-stable neural-network control"

只有魔法才能打败魔法

- ➢ 同时优化神经网络Lyapunov函数和神经网络控制器
- ➢ 计算闭环系统的吸引域（region of attraction, ROA）
- ➢ 在倒立摆、2D、3D无人机上进行仿真验证



方法

结论 5.13［小范围渐近稳定性定理］ 对连续时间非线性时变自治系统(5.46)，若可构造对 $\boldsymbol{x}$ 和 t 具有连续一阶偏导数的一个标量函数 $V(\boldsymbol{x},t)$，$V(\boldsymbol{0},t)=0$，以及围绕状态空间原点的一个吸引区 Ω，使对所有非零状态 $\boldsymbol{x}\in\Omega$ 和所有 $t\in[t_0,\infty)$ 满足如下条件：

(i) $V(\boldsymbol{x},t)$为正定且有界；

(ii) $\dot{V}(\boldsymbol{x},t)\triangleq \mathrm{d}V(\boldsymbol{x},t)/\mathrm{d}t$ 为负定且有界；

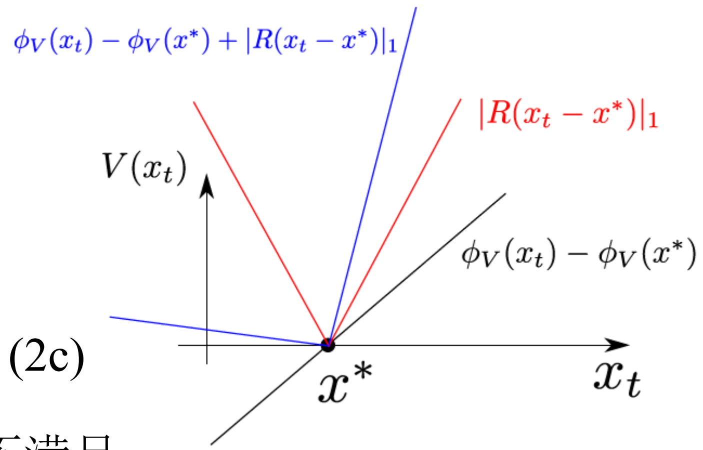
则系统原点平衡状态 $\boldsymbol{x}=\boldsymbol{0}$ 在 Ω 域内为一致渐近稳定。

离散系统：

$$V(x_t) > 0 \ \forall x_t \in \mathcal{S}, x_t \neq x^* \qquad (2a)$$

$$V(x_{t+1}) - V(x_t) \leq -\epsilon_2 V(x_t) \ \forall x_t \in \mathcal{S}, x_t \neq x^* \qquad (2b)$$

$$V(x^*) = 0 \qquad (2c)$$



神经网络形式V函数： $V(x_t) = \phi_V(x_t) - \phi_V(x^*) + |R(x_t - x^*)|_1,$ 满足(2a), (2c)

满足**所有**：给定x的定义域，求值域的**最小值**，如果最小值大于0，则满足；否则不满足

因此，条件(2a) ⟶ $V(x_t) \geq \epsilon_1 |R(x_t - x^*)|_1 \ \forall x \in \mathcal{S}, \ 0 < \epsilon_1 < 1$ ⟶ $\max\limits_{x_t \in \mathcal{B}} \epsilon_1 |R(x_t - x^*)|_1 - V(x_t)$

条件(2b) ⟶ $\max\limits_{x_t \in \mathcal{B}} V(x_{t+1}) - V(x_t) + \epsilon_2 V(x_t),$

类似的，$u_t = \pi(x_t) = \mathrm{clamp}\left(\phi_\pi(x_t) - \phi_\pi(x^*) + u^*, u_{\min}, u_{\max}\right)$

注意： $x_{t+1} = f(x_t, \pi(x_t)),$

$$\max_{x_t \in \mathcal{B}} \epsilon_1 |R(x_t - x^*)|_1 - V(x_t)$$

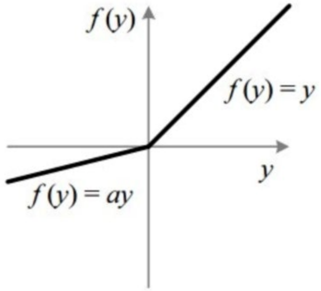
$$\max_{x_t \in \mathcal{B}} V(x_{t+1}) - V(x_t) + \epsilon_2 V(x_t),$$

求解这个优化问题的 方法

❑ Leaky ReLU函数

一种分段线性函数，

NN中最常用的函数



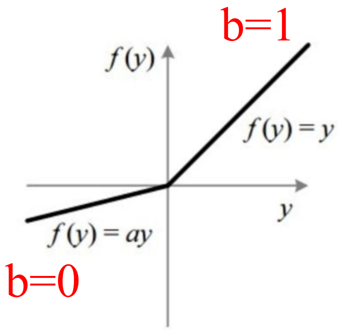
❑ Mixed-Integer Programs (MIP)

问题定义：

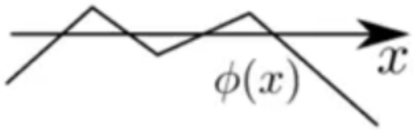
$$\max_{x,b} c^T x + d^T b$$
$$s.t.\ Ax + Fb \le g$$
$$b \in \{0,1\}$$

❑ 对于一个确定的采用ReLU函数的神经网络，求解输出的最值问题就是一个MIP问题

$$z_i = \sigma(W_i z_{i-1} + b_i), i = 1, \ldots, n-1$$
$$z_n = W_n z_{n-1} + b_n, z_0 = x,$$



分段线性函数



对$\forall x$，

求$\max\phi(x)$

$\max\phi(x) < 0$？

$$\max_{x,b} c^T x + d^T b$$
$$s.t.\ Ax + Fb \le g$$
$$b \in \{0,1\}$$

MIP求解器

Tjeng, V., Xiao, K., & Tedrake, R. (2017). Evaluating robustness of neural networks with mixed integer programming. *arXiv preprint arXiv:1711.07356*.

方法

❑ 训练流程

方法一：对抗性训练

$$\max_{x_t \in \mathcal{B}} \epsilon_1 |R(x_t - x^*)|_1 - V(x_t)$$

$$\max_{x_t \in \mathcal{B}} V(x_{t+1}) - V(x_t) + \epsilon_2 V(x_t),$$

候选控制器$u_t = \pi(x_t)$，
候选Lyapunov函数$V(x_t)$

MIP检测器检查是否
满足$V(x_t) > 0$和
$V(x_{t+1}) - V(x_t) < 0, \forall x_t \in \beta$

Yes

成功！

No

将不满足的x_{adv}
加入训练集χ

通过训练集χ训练**控制器**和
V函数的神经网络，以最小
化损失函数$loss_\theta(\chi_1, \chi_2)$

梯度下降

新$u_t{}' = \pi(x_t)$
新$V'(x_t), R$

$$\eta_1(x_1^i) = \max(\epsilon_1 |R(x_1^i - x^*)|_1 - V(x_1^i), 0) \tag{11a}$$

$$\eta_2(x_2^i) = \max(V(f(x_2^i, \pi(x_2^i))) - V(x_2^i) + \epsilon_2 V(x_2^i), 0), \tag{11b}$$

$$\text{loss}_\theta(\mathcal{X}_1, \mathcal{X}_2) = |\eta_1(\mathcal{X}_1)|_p + |\eta_2(\mathcal{X}_2)|_p,$$

特点：训练速度快，但容易过拟合

方法

❑ 训练流程

方法二：min-max优化

$$\max_{x_t \in \mathcal{B}} \epsilon_1 |R(x_t - x^*)|_1 - V(x_t)$$

$$\max_{x_t \in \mathcal{B}} V(x_{t+1}) - V(x_t) + \epsilon_2 V(x_t),$$

候选控制器$u_t = \pi(x_t)$，
候选Lyapunov函数$V(x_t)$

MIP检测器检查是否满足$V(x_t) > 0$和$V(x_{t+1}) - V(x_t) < 0, \forall x_t \in \beta$

Yes

成功！

No

直接计算MIP相对于参数θ的梯度，并记为$\frac{\partial \gamma}{\partial \theta}$

使用梯度下降公式更新神经网络以实现外层的最小优化：

$$\theta = \theta - \text{StepSize} \times \frac{\partial \gamma}{\partial \theta}$$

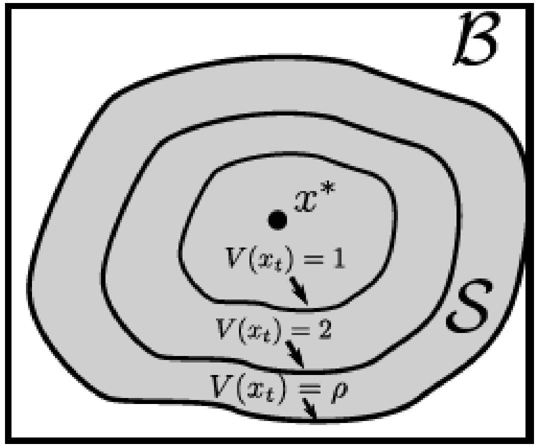
$$\min_{\theta} \left(\underbrace{\max_{x_t \in \mathcal{B}} \epsilon_1 |R(x_t - x^*)|_1 - V(x_t)}_{\text{MIP (10a)}} + \underbrace{\max_{x_t \in \mathcal{B}} V(x_{t+1}) - V(x_t) + \epsilon_2 V(x_t)}_{\text{MIP (10b)}} \right)$$

目标函数

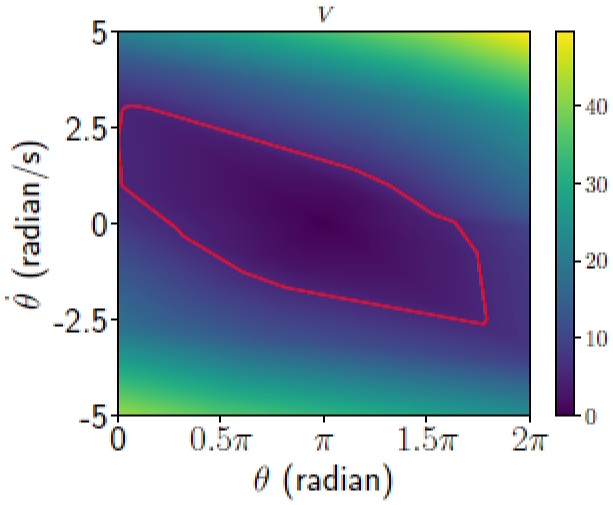
特点：训练速度慢，但更易收敛

方法

❑ 求解吸引域



已经获得V函数了



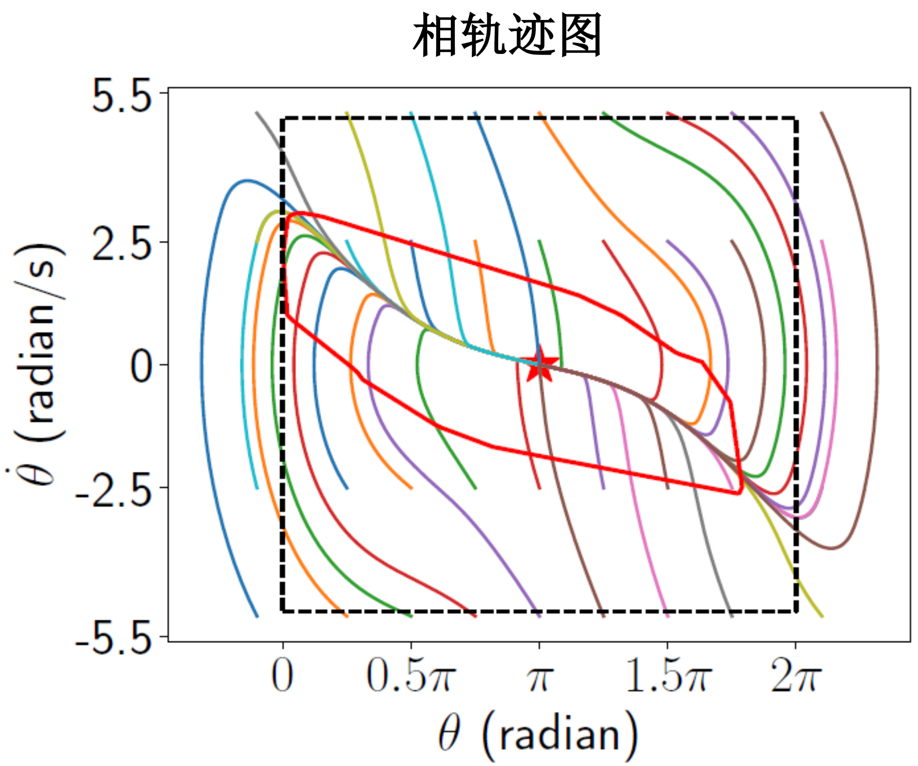
倒立摆问题的吸引域

步骤1： $\rho = \min_{x_t \in \partial \mathcal{B}} V(x_t),$ 同样是MIP问题

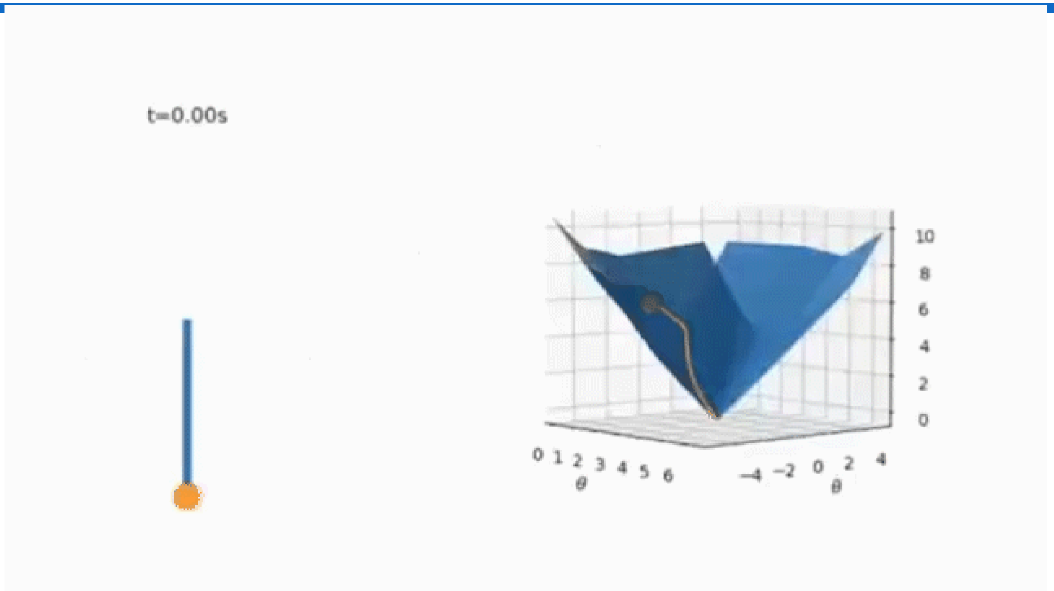
步骤2： 在B集中搜索 $V(x_t) = \rho$

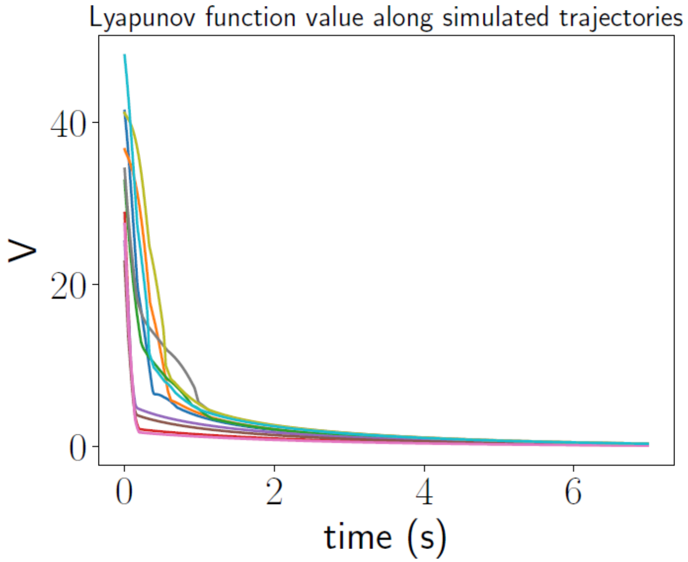
结果

❑ 倒立摆



红心：平衡点
红线：吸引域
黑色虚线：设定的搜索域



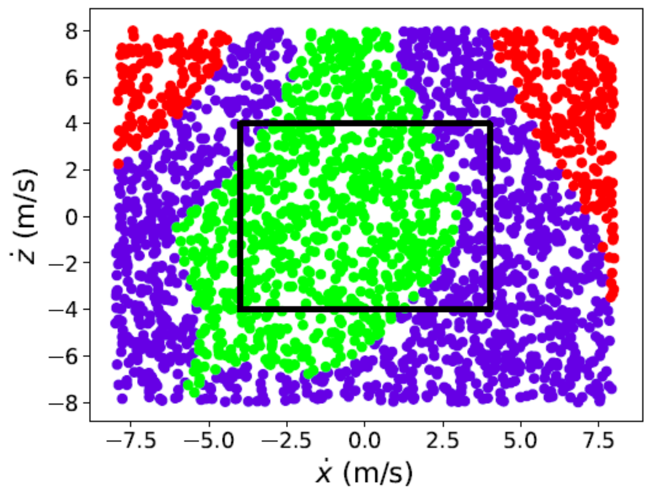


结果

❑ 2D无人机

	LQR成功	LQR失败
NN成功	8078	1918
NN失败	0	4

使用拉格朗日动力学，选取10000个随机初始化位置，在LQR控制和NN控制后的结果。NN效果更好。

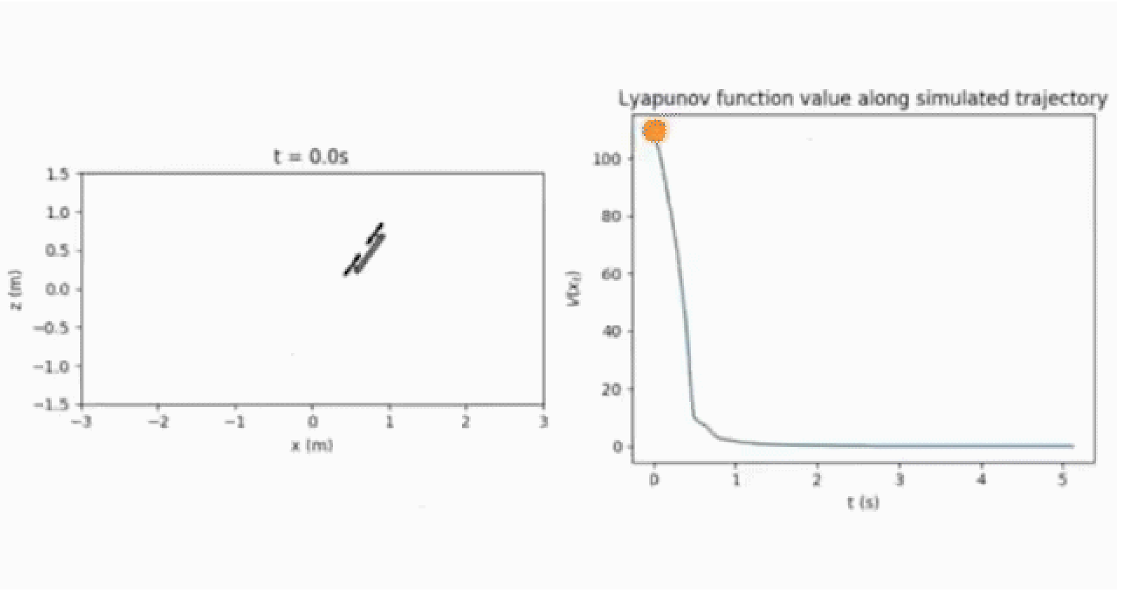


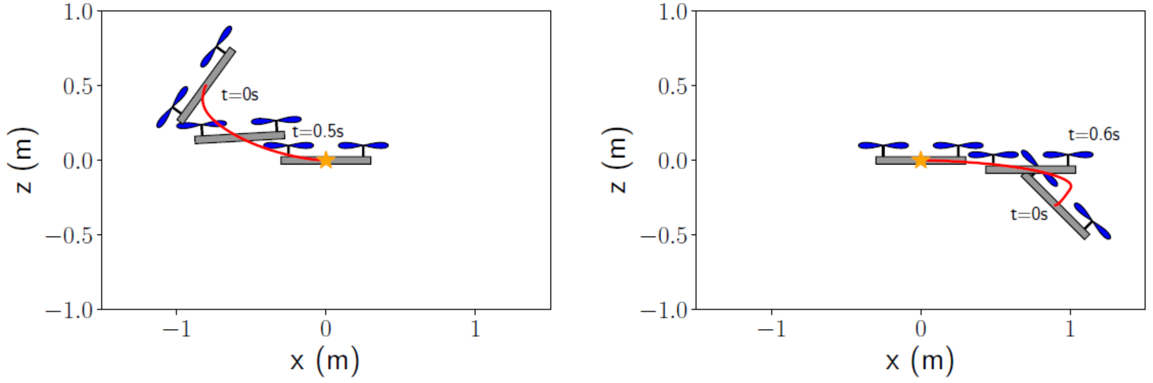
紫色：NN成功但LQR失败

绿色：NN和LQR都成功

红色：NN和LQR都失败

通过采样进一步验证稳定性。





2D无人机飞行轨迹

总结

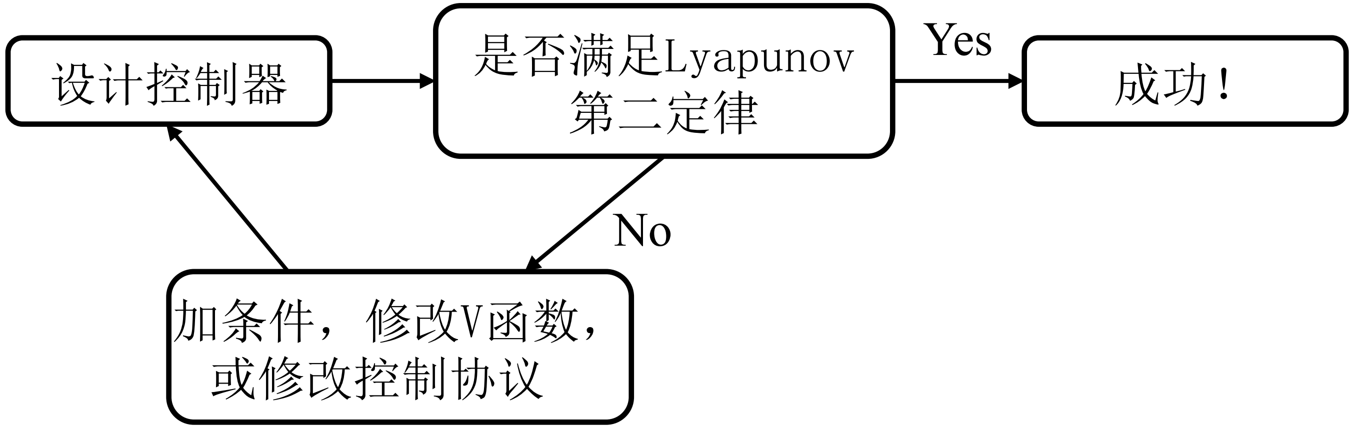
❑ 总结

- ❑ 本文将神经网络控制器和Lyapunov稳定性约束进行组合，通过将神经网络求极值问题转化为求解MIP问题，以迭代的方式同时训练出神经网络Lyapunov函数和神经网络控制器，实现了神经网络的稳定控制。

❑ 思考

- ❑ 与深度强化学习对比，提供了另外一种训练神经网络控制器的思路——满足Lyapunov条件。
- ❑ 与控制论文的推导过程类似

  手工业时代→自动化时代?



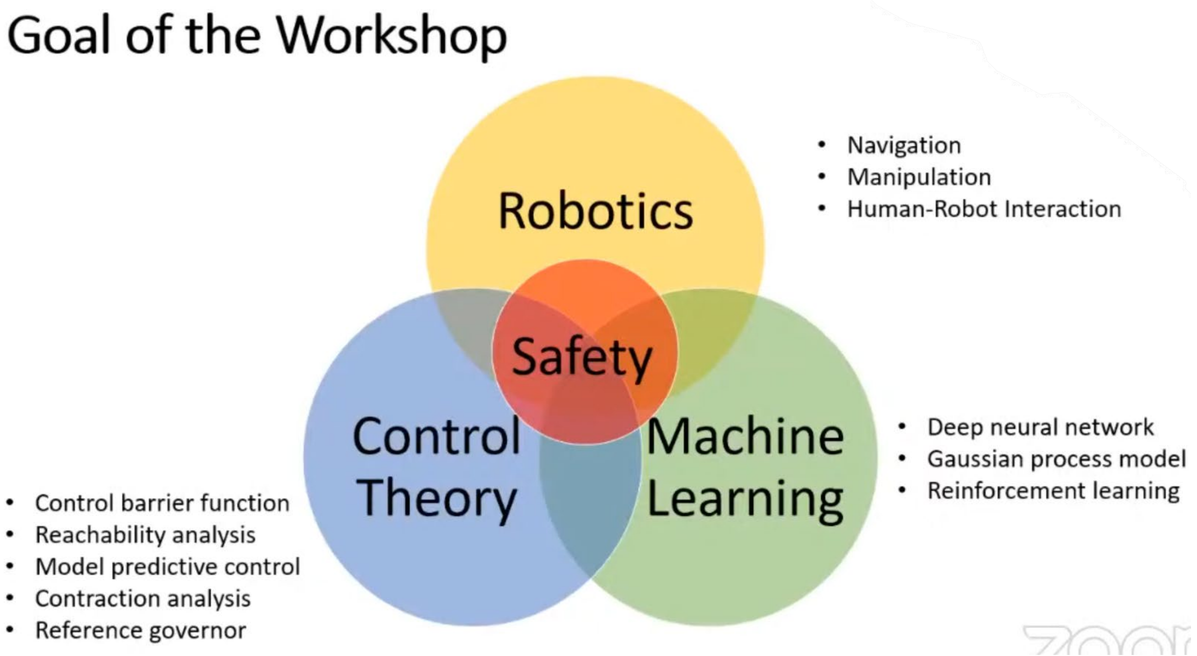
总结

- ❑ 存在问题与未来工作
  - ❑ 如果网络规模较大，则MIP求解较慢。加速训练。
  - ❑ 现有理论针对的是离散时间系统，希望拓展到连续时间系统。
  - ❑ 希望可以综合Barrier functions，在训练中设定动作不可到达的区域。

资料推荐

- ICRA 2021 Workshop on Safe Robot Control with Learned Motion and Environment Models, 2021.06.04

  网址：https://scl-icra2021.github.io/



- Control Meets Learning Seminar

  Caltech主办, 网址：https://sites.google.com/view/control-meets-learning/

- 相关论文：

Chang, Ya-Chien, Nima Roohi, and Sicun Gao. "Neural lyapunov control." *arXiv preprint arXiv:2005.00611* (2020).

谢谢聆听!

Q&A