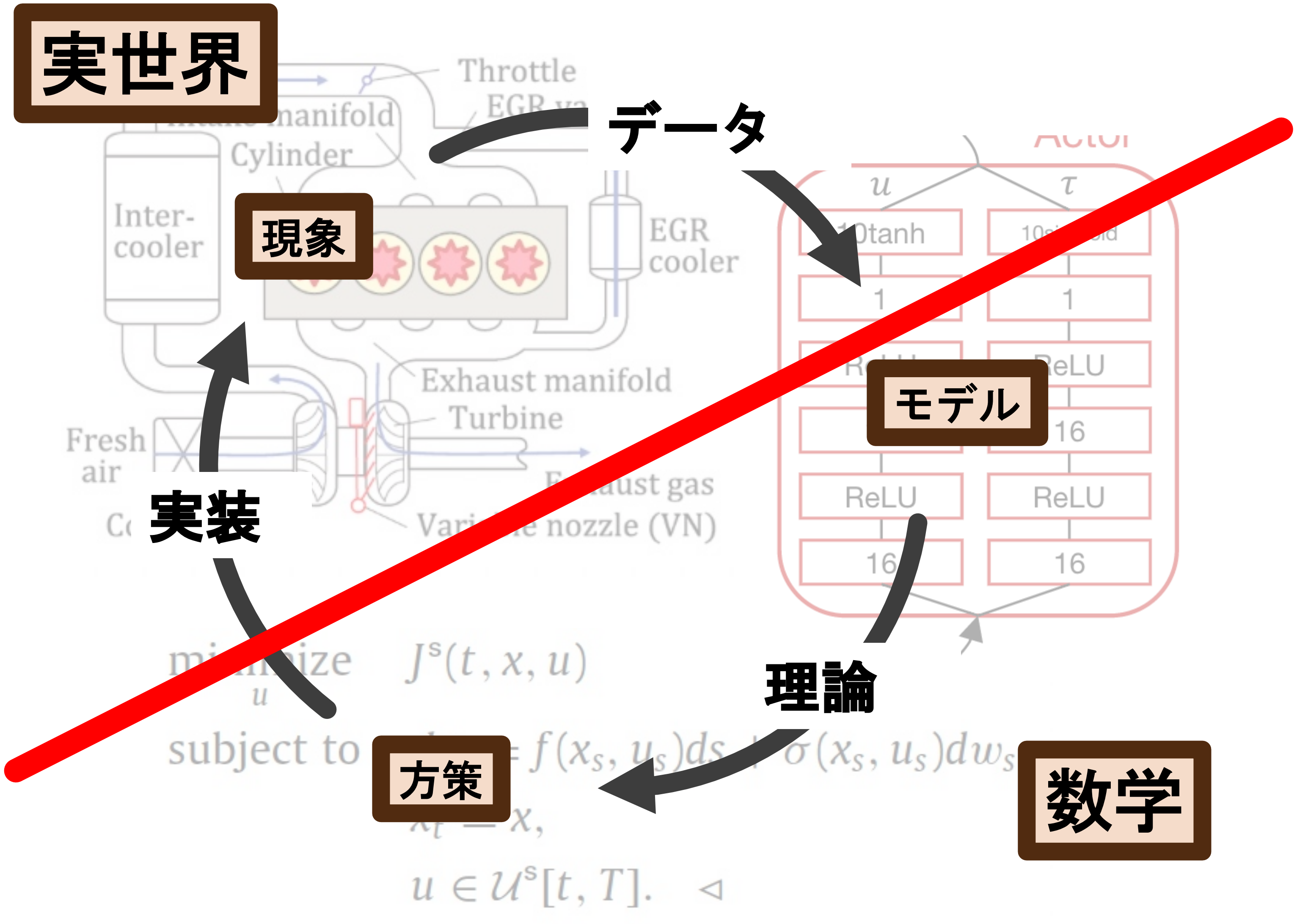
システム制御に使ってみたい統計的学習の理論と手法

京都大学 情報学研究科 数理工学専攻　加嶋 健司

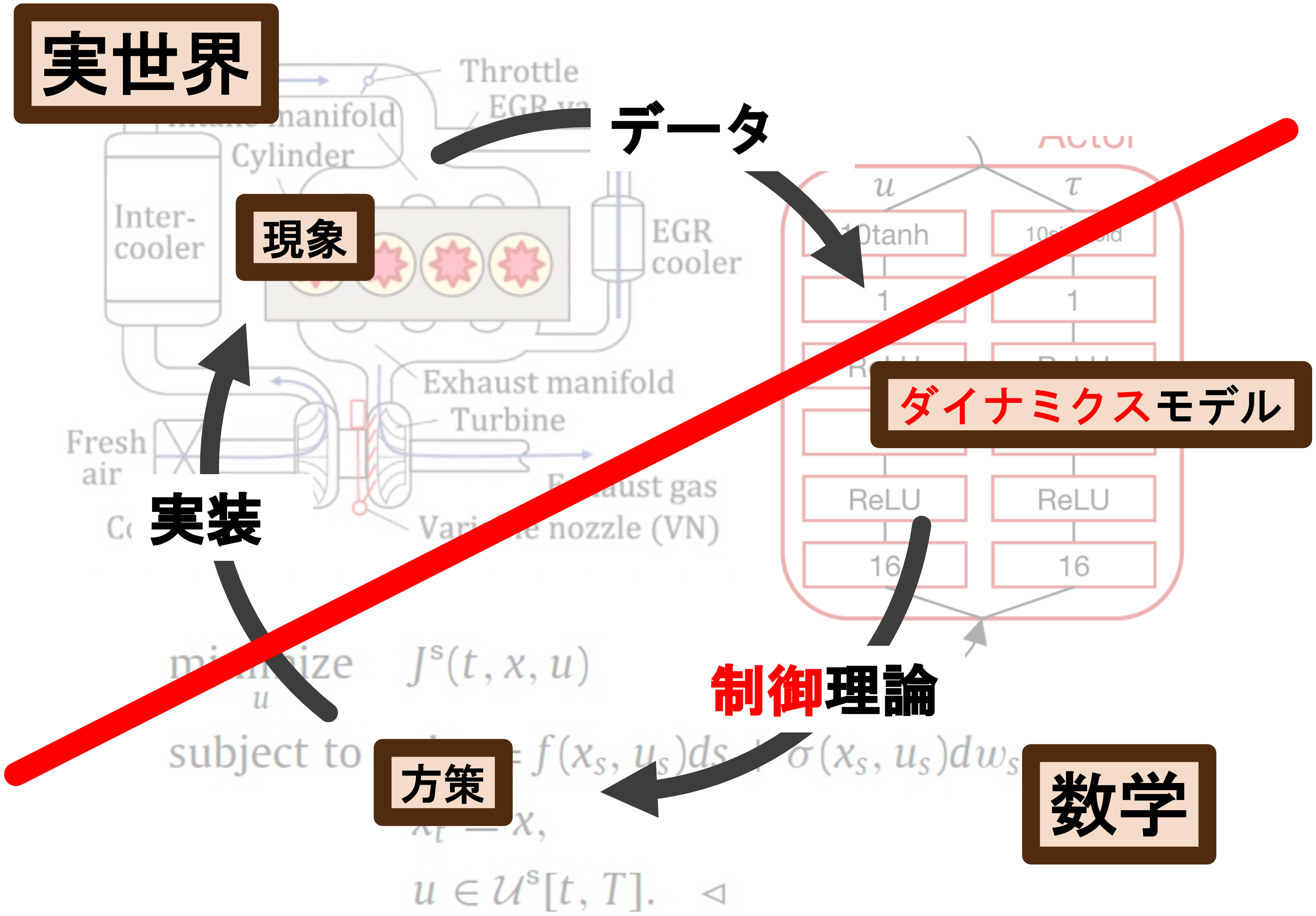
2022/12/10　Computational Intelligence フォーラム

数理工学


実世界
データ
現象
モデル
実装
理論
方策
数学
Throttle
Cylinder
Inter-
cooler
EGR
cooler
Exhaust manifold
Turbine
Fresh
air
ReLU
16

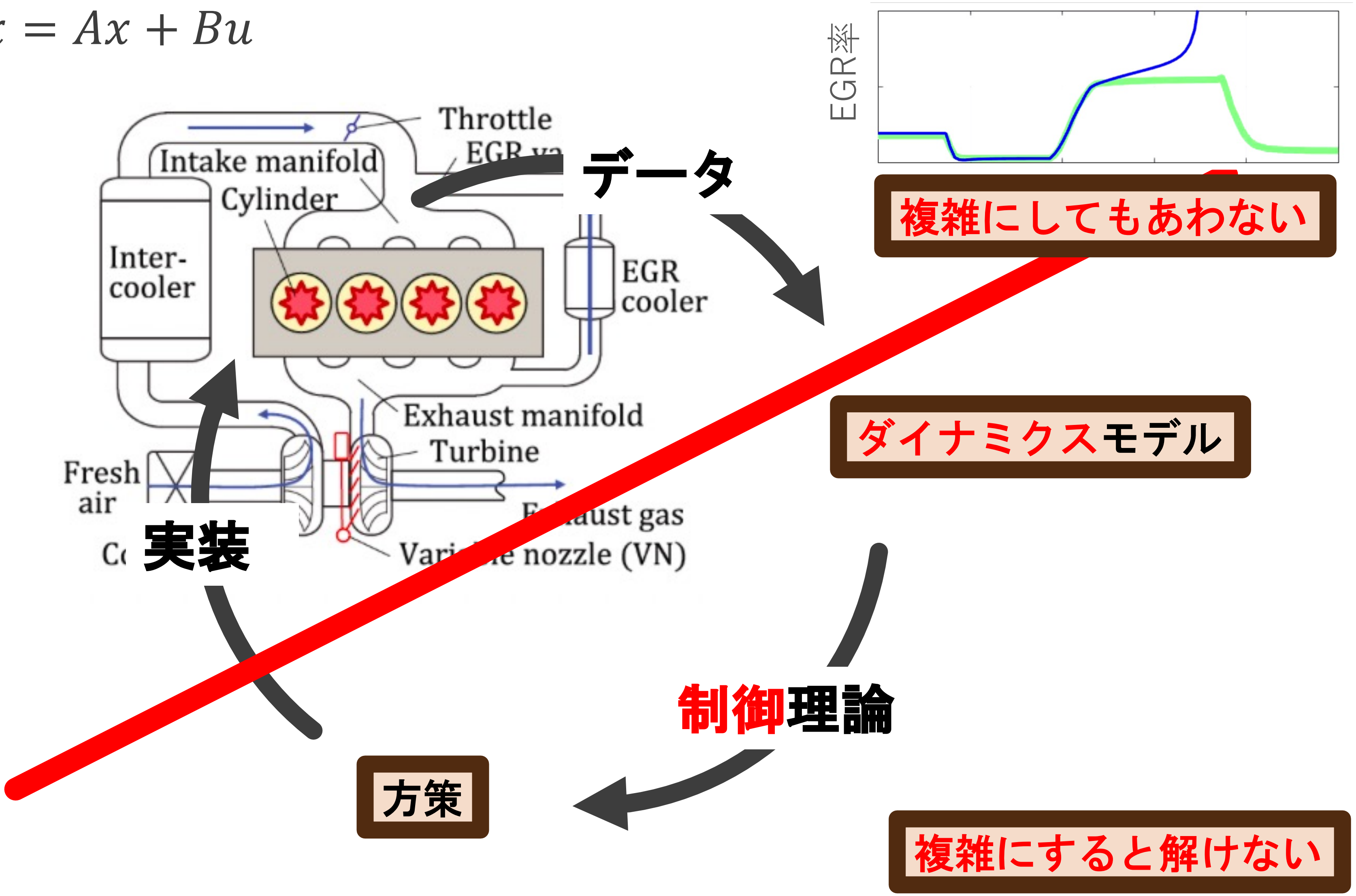
数理工学と制御工学

- 時々刻々と変化する現象



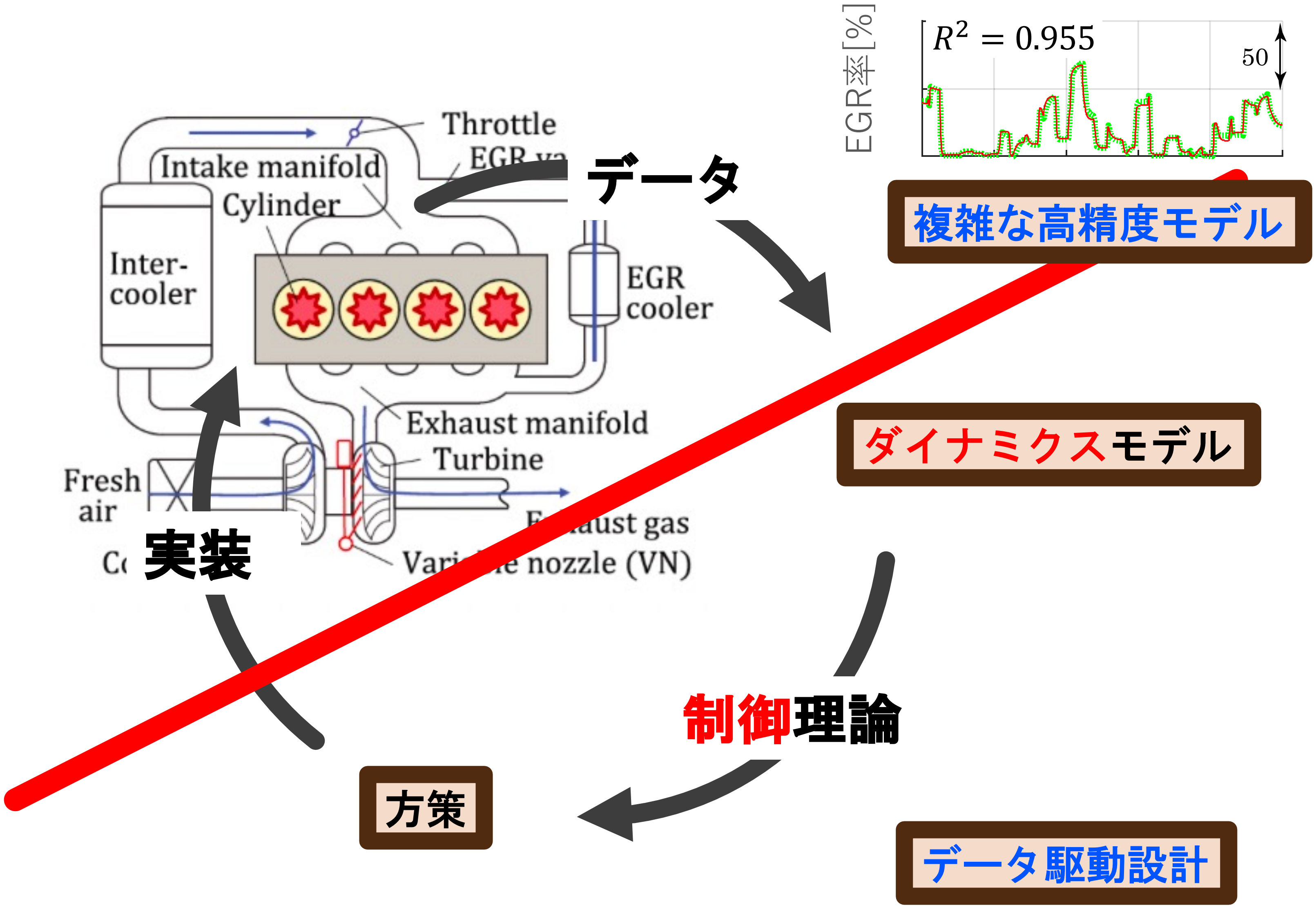
具体例（エンジン）

- 簡単なモデルに限らざるを得ないことが多い
  - $\dot{x} = Ax + Bu$



人工知能ならやってくれるはず…

- ビッグデータと計算機パワーが解決してくれる



現実

- モデリング
  - 誤差最小化問題を定式化し，GPUで勾配法を**実装できる**

- データ駆動設計
  - 最適制御問題を定式化し，GPUで勾配法（強化学習手法）を**実装できる**

- 個人的な感想
  - 期待する結果には程遠い
    - ✓ 対象が線形であっても難しい
    - ✓ 局所最適なのかデータ不足なのか不明
  - セイフティクリティカルな応用には受け入れられない
    - ✓ モデルが担保していた安定性・最適性

講演概要

- それでも私は肯定派です

- 新しい強力なツールとしての機械学習
  - 事前知識やモデルベースト設計の積極的活用
  - **構造付きダイナミクス学習**

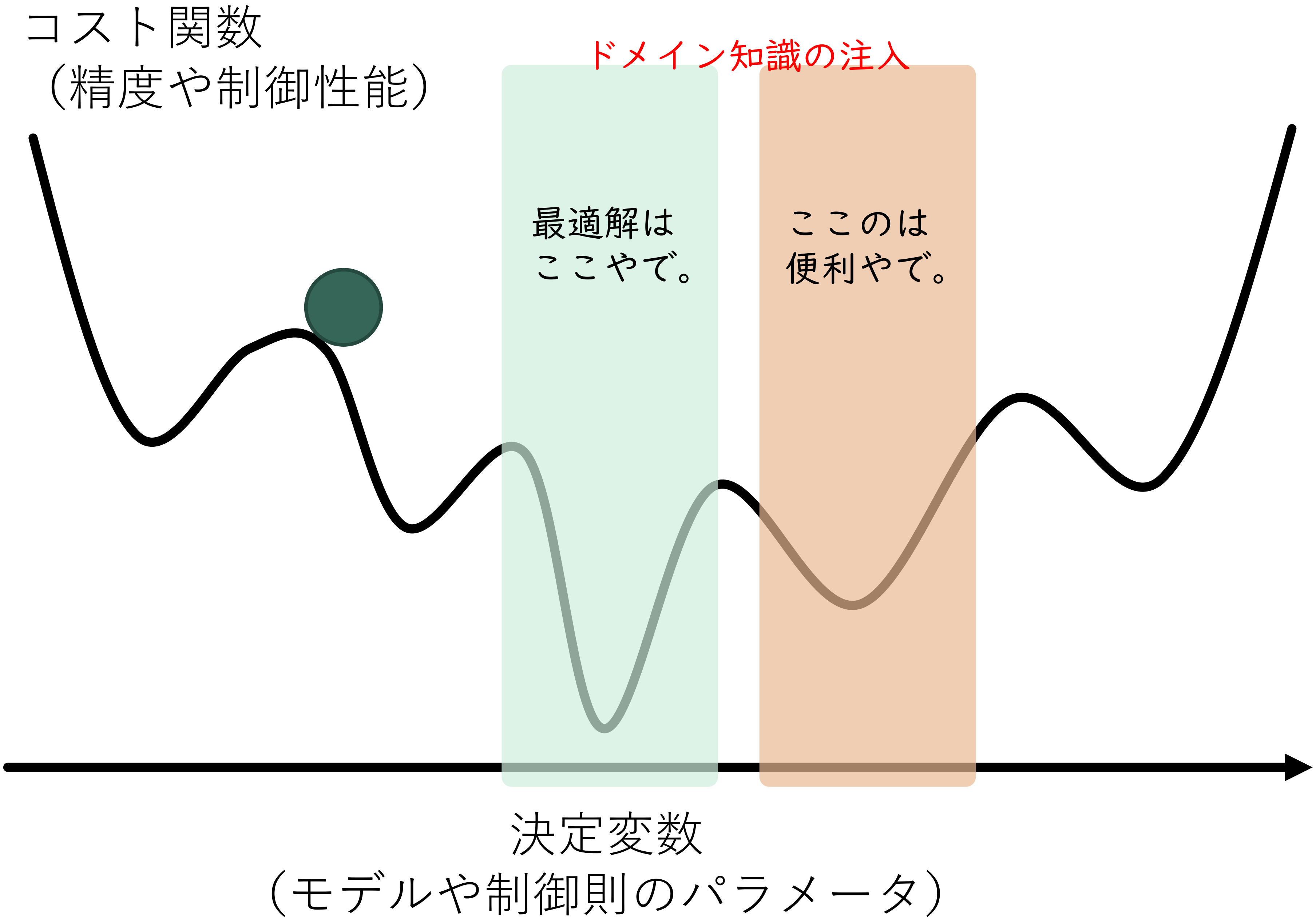
- 統計的学習と融合し深化する制御理論
  - 「情報」を陽に扱うことで生まれる異分野との接点
  - **最適輸送と最適制御**

- まとめ

データ以外の情報をいかに取り入れるか

- Physics-informed, Stability-aware, Optimization-oriented

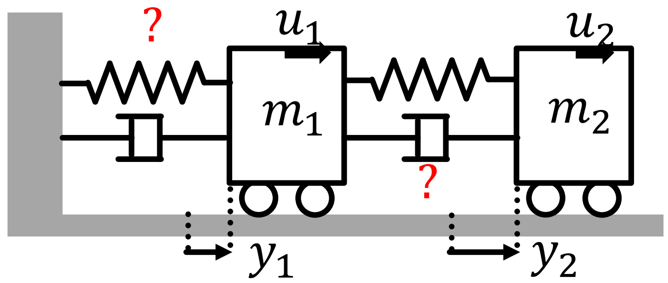
勾配法でぶん殴る

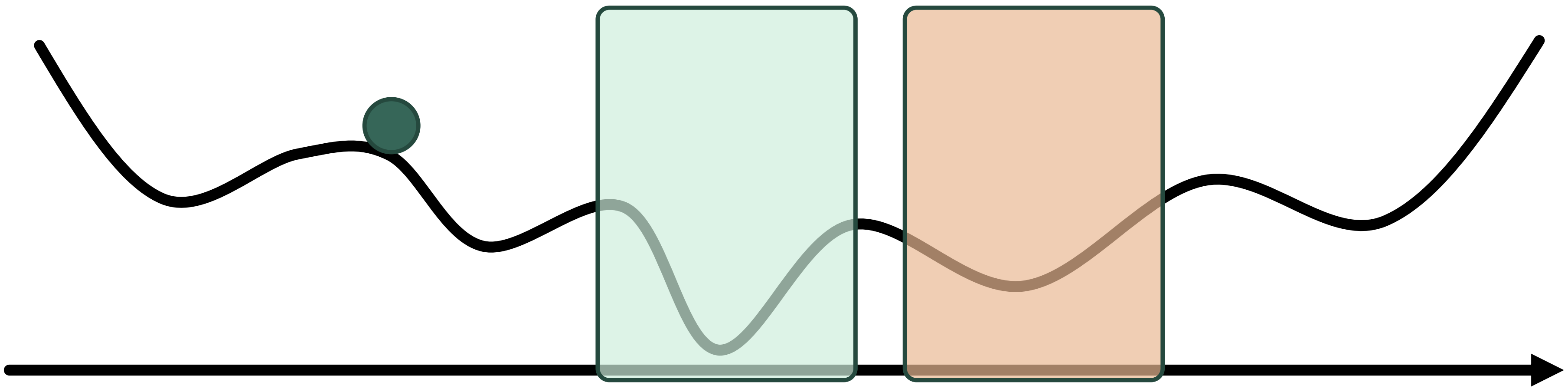

コスト関数
（精度や制御性能）
ドメイン知識の注入
最適解は
ここやで。
ここのは
便利やで。
決定変数
（モデルや制御則のパラメータ）

ドメイン知識の注入

- 探索範囲の限定
  - グレイボックスモデリング
  - 『動的な性質』：安定性、単調性
  - 『使い勝手』：凸性
  - 『陰関数』：模倣学習

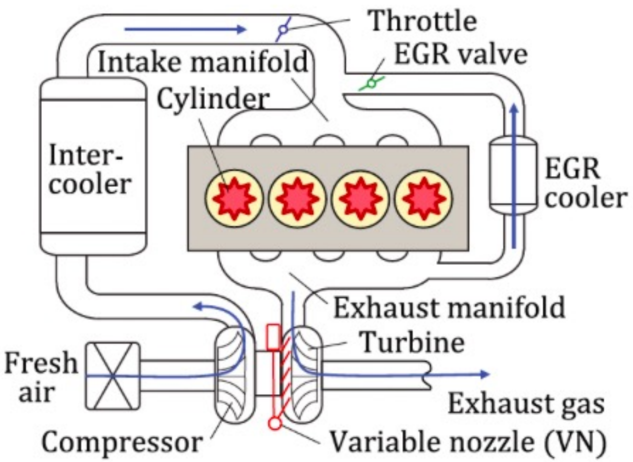
$$x_{k+1} = f(x_k, u_k), y_k = g(x_k)$$



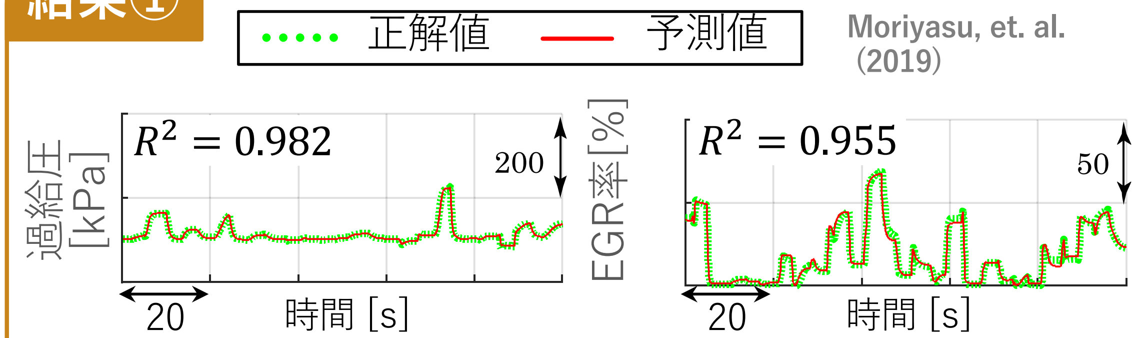


決定変数
（モデルや制御則のパラメータ）

具体例再訪



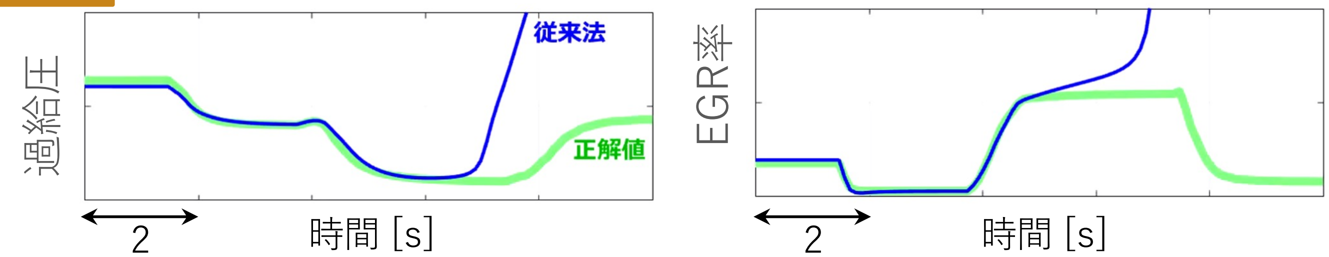
結果①



Moriyasu, et. al. (2019)

基本的には良好に予測できるが...

結果②



中には
発散する
条件もある

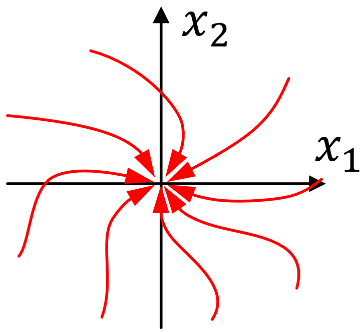
エンジンは **物理的には安定** だが，その **性質が保存されない**

工夫１：ドメイン知識チート

- 線形モデル＆大変な努力で高精度モデル

- 軌道の位相的性質
  - データから獲得する必要はない

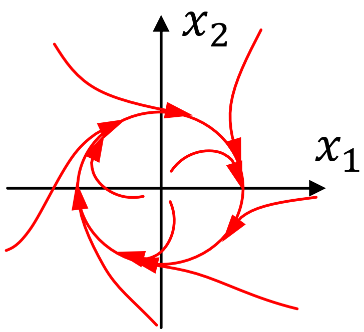
■漸近安定（単一平衡点）



エンジン

その他
大部分の
工業製品

■周期変動（リミットサイクル）

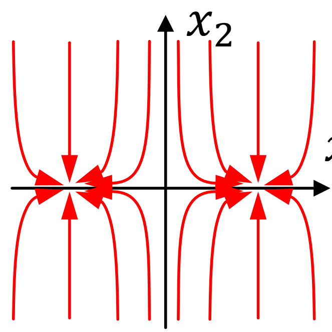


メトロノーム　**生物リズム**

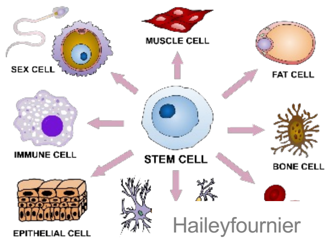
Badajoz, España

@yb_woodstock

■多安定（複数平衡点）

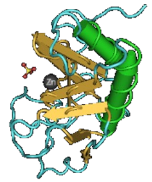


細胞分化



Haileyfournier

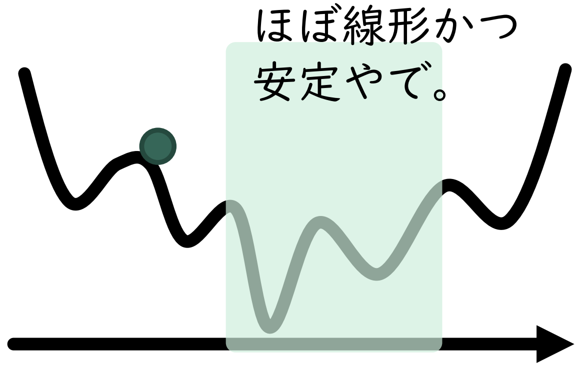
遺伝子スイッチ

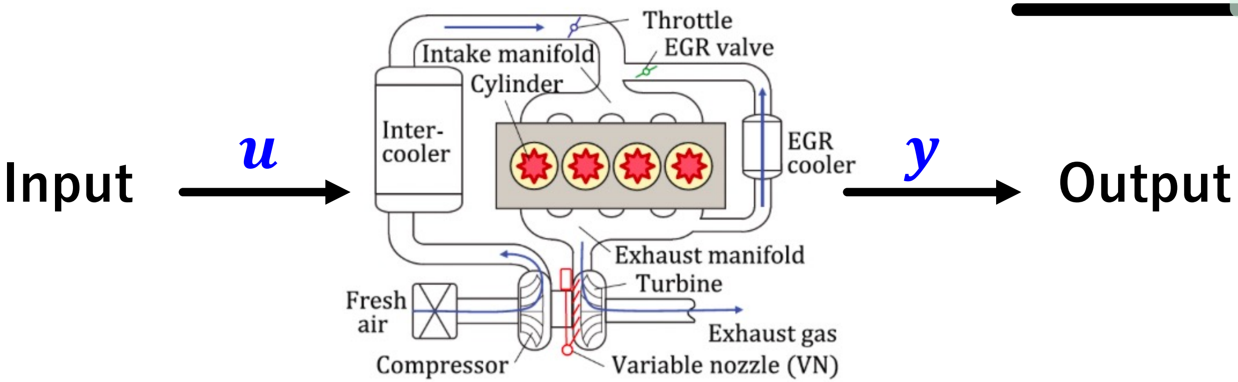


Hall, T.M., et. al. (1995)

工夫１：ドメイン知識チート

- 線形モデル＆座標変換
  - 漸近安定性は検証不要，多少の外挿には耐える
- 微分同相写像のみ深層学習
  - Normalizing flowなど





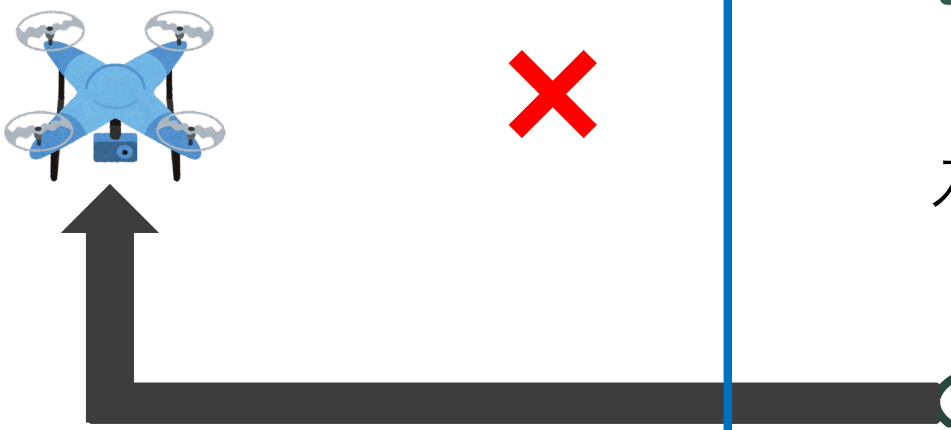
$$\text{Input} \xrightarrow{\boldsymbol{u}} \boxed{\boldsymbol{\Psi}} \xrightarrow{\boldsymbol{v}} \boxed{\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{v}} \xrightarrow{\boldsymbol{x}} \boxed{\boldsymbol{\Phi}^{-1}} \xrightarrow{\boldsymbol{y}} \text{Output}$$

モデル予測制御による実時間最適制御

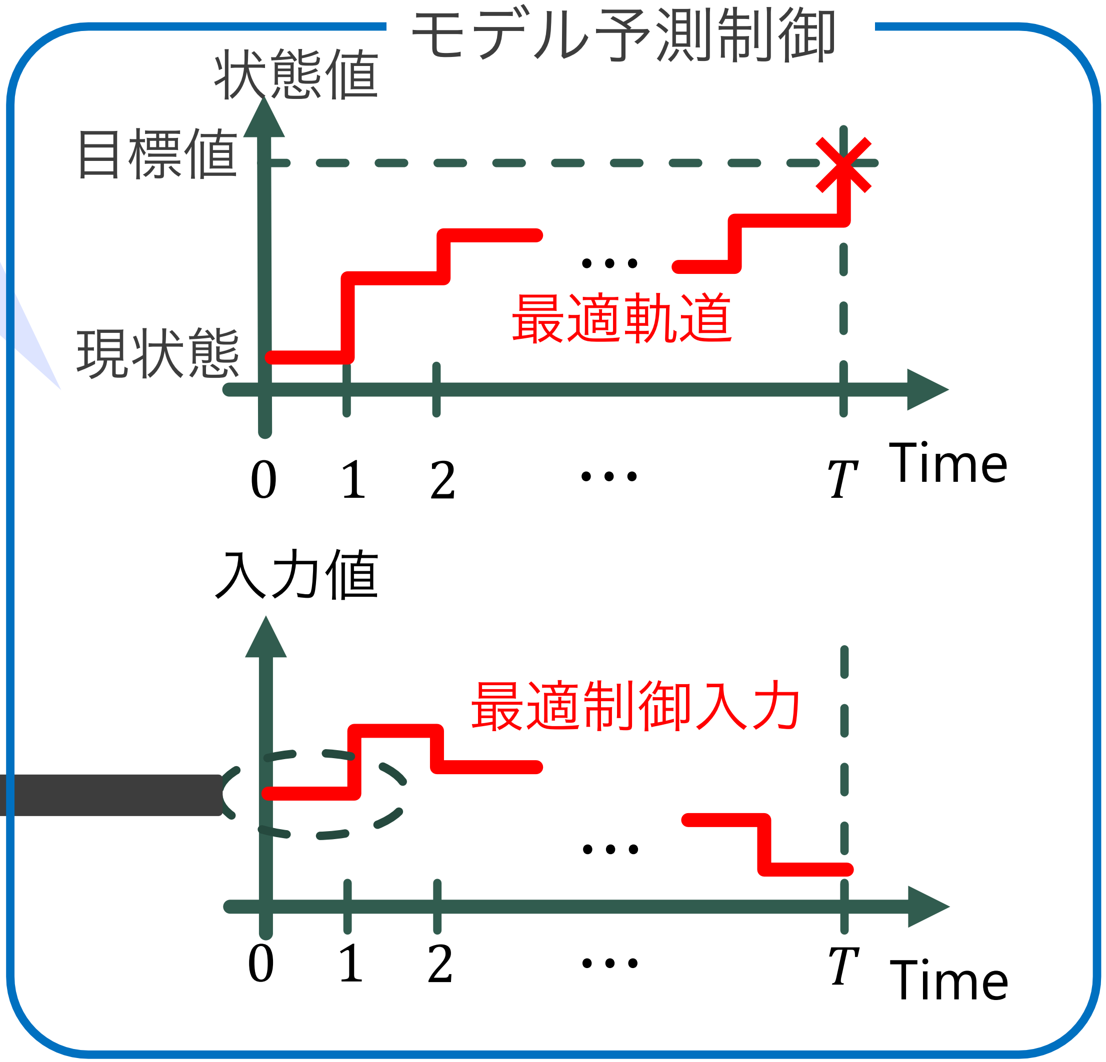
モデル予測制御（Model predictive control; MPC）

- 「有限時間先の未来までの最適化」を毎時刻行う

$$
\begin{aligned}
\min_{u} \quad & \sum_{k=0}^{T} \ell(x(k), u(k)) \\
\text{s.t.} \quad & x(0) = x^0, \\
& x(k+1) = Ax(k) + Bu(k) \\
& x \in \mathbb{X}, u \in \mathbb{U}
\end{aligned}
$$



最適系列の初期値を入力

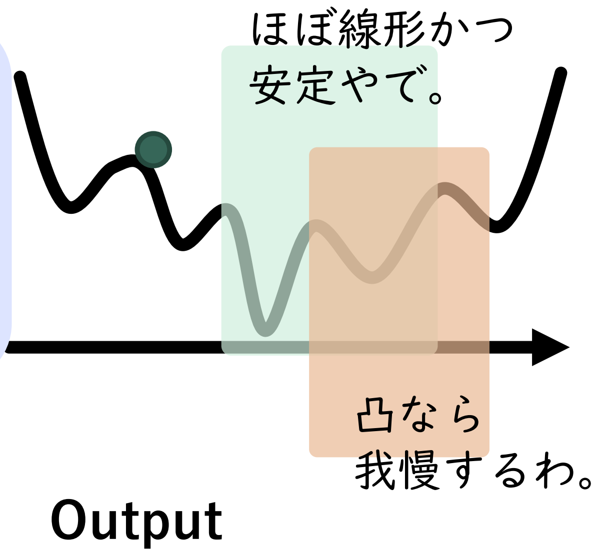


工夫２：モデル予測制御への援護射撃

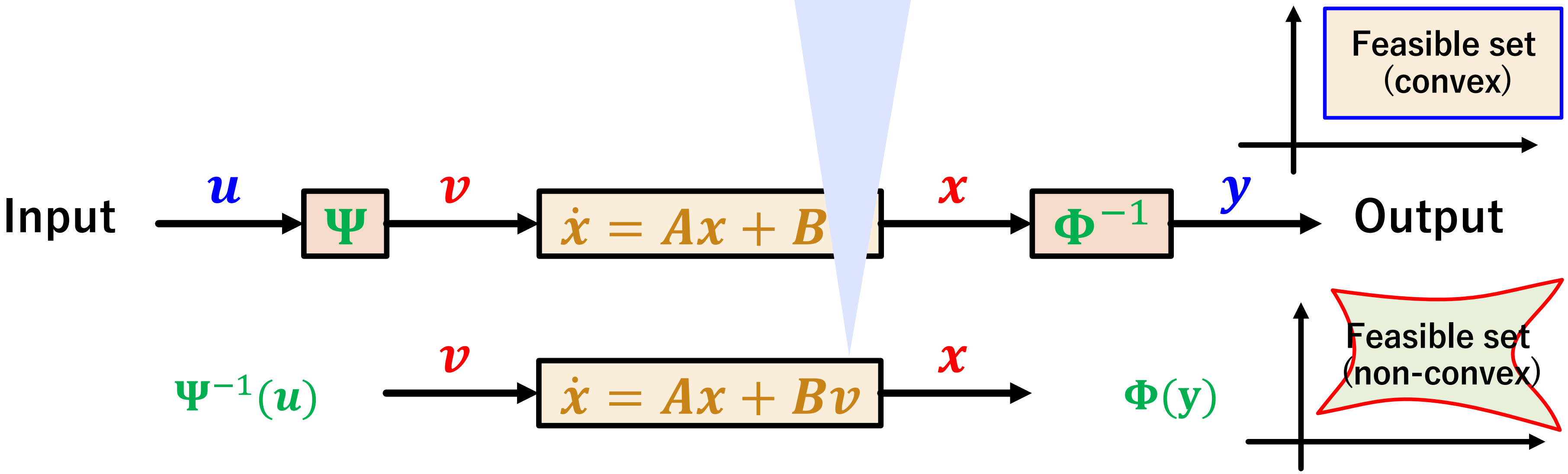
- 線形システムの制御問題への変換
  - 凸制約であればオンライン最適化可（数ms）

$$\min_{u} \sum_{k=0}^{T} \ell\big(x(k), v(k)\big)$$

$$\text{s.t. } x(0) = x^0, x \in \mathbb{X}, v \in \mathbb{V}$$



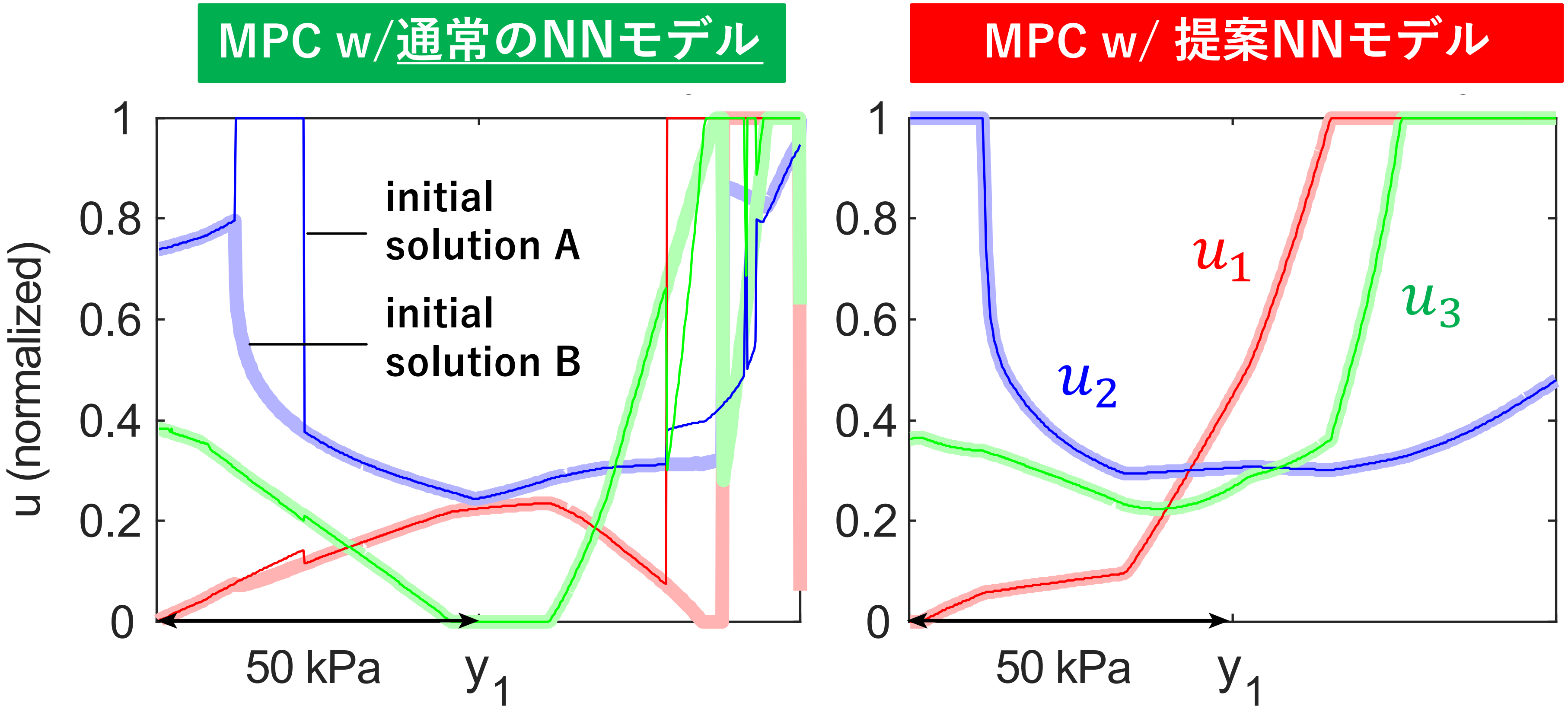
- 同相写像を凸関数に限定
  - **Partially Input Convex NN**



定理

■ 各状態に対して最適制御入力値は一意かつ連続

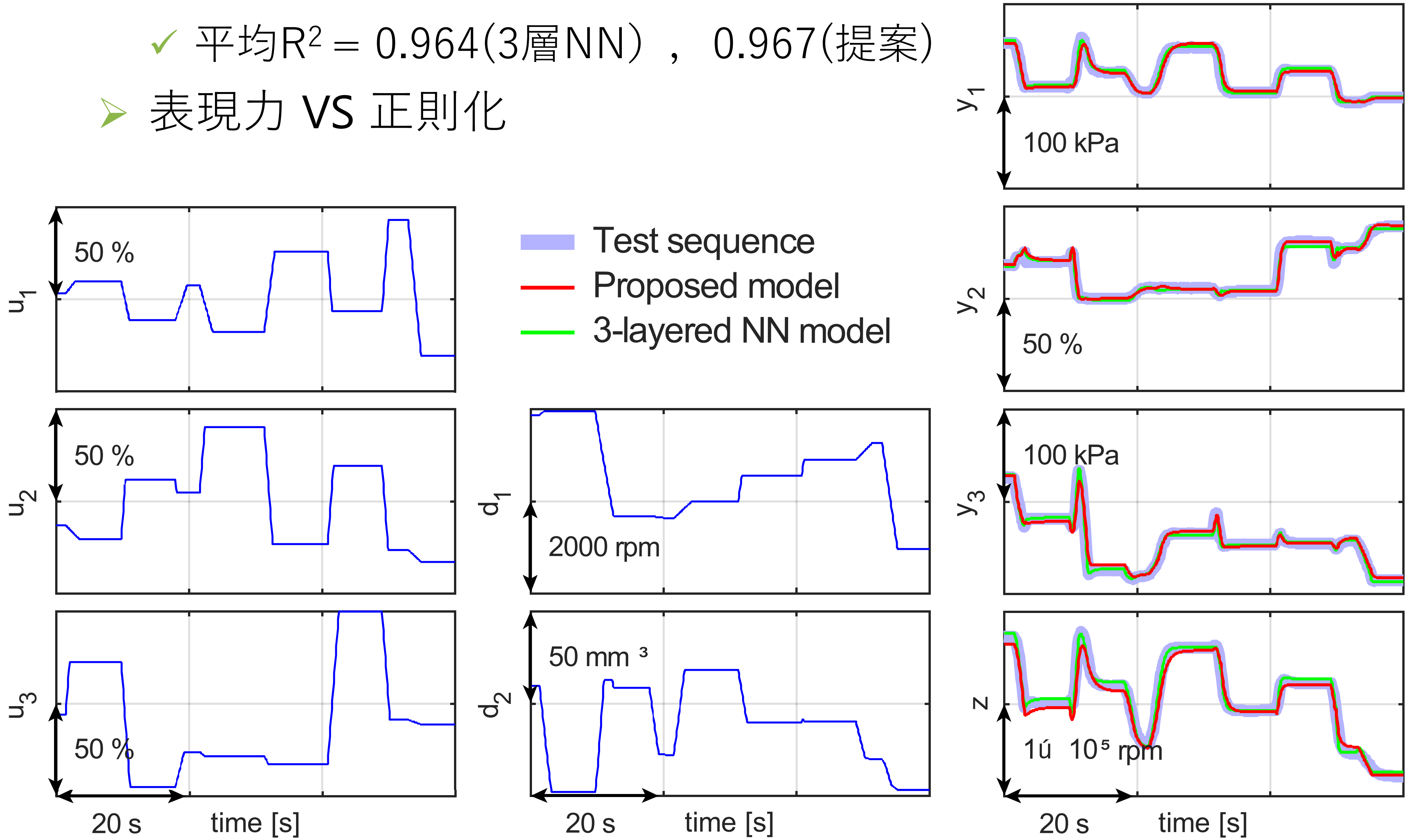
現時刻の状態がyのときに、入力すべきuの値



R. Moriyasu, et al., "Structured Hammerstein-Wiener Model Learning for MPC," *IEEE L-CSS* (2021)

検討例

- モデル精度
  - おおむね同等
    - ✓ 平均R^2 = 0.964(3層NN)， 0.967(提案)
  - 表現力 VS 正則化



検討例

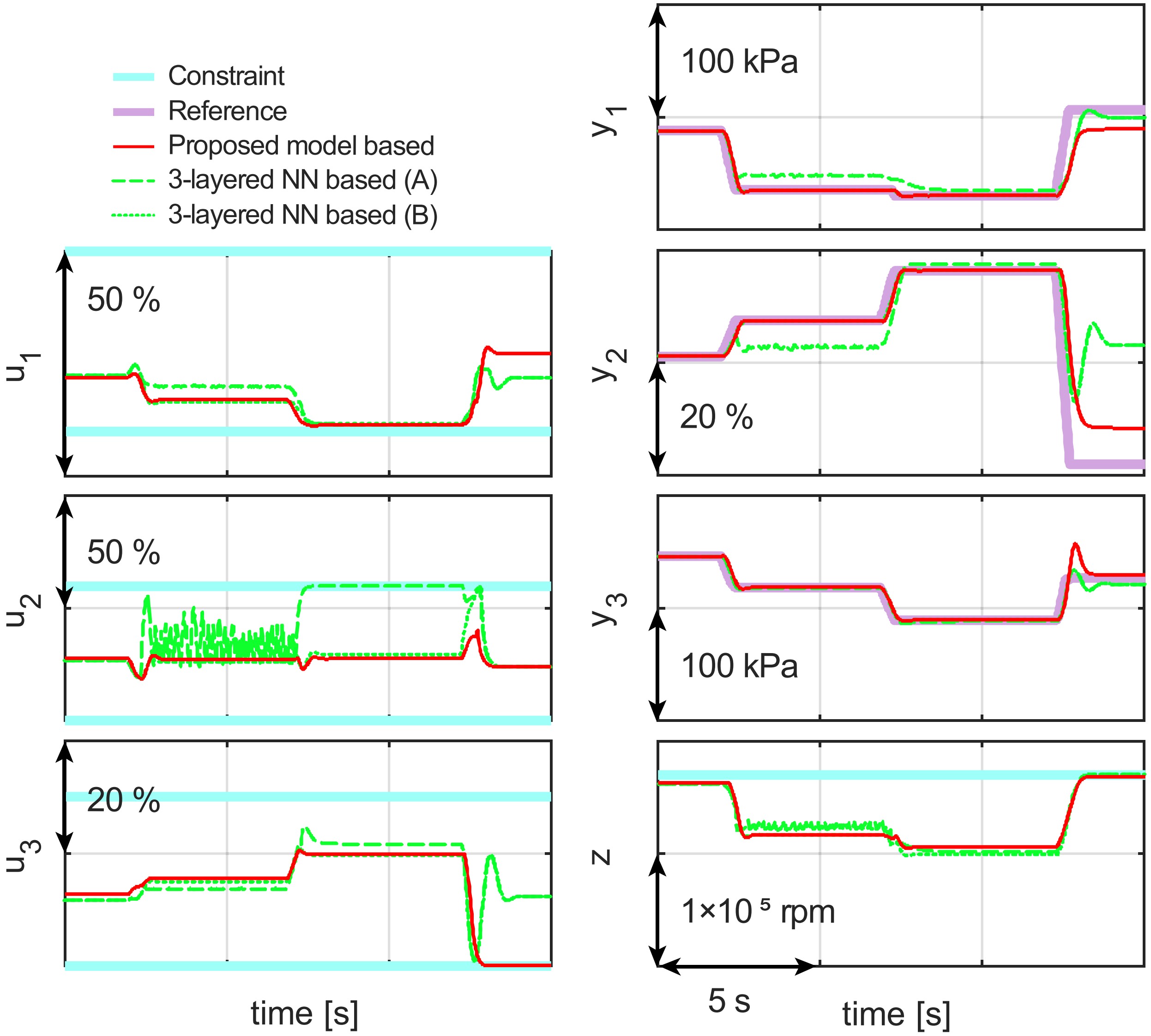
■ 制御シミュレーション

MPC w/ 通常のNNモデル

・毎時刻解く問題に解の一意性や滑らかさが保証されていないためバタつく

MPC w/ 提案NNモデル

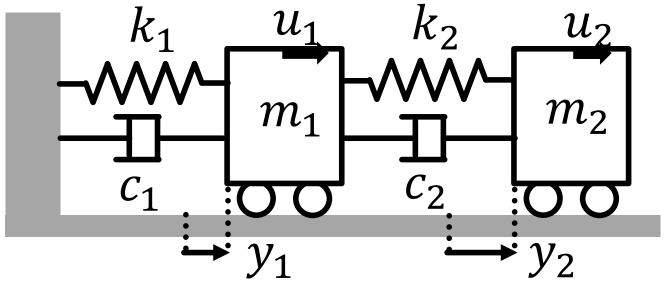
・



陰関数と模倣学習

■ 最適制御則からモデルとコストを決定する

線形プラント: 直列MCK系 (2入力2出力)



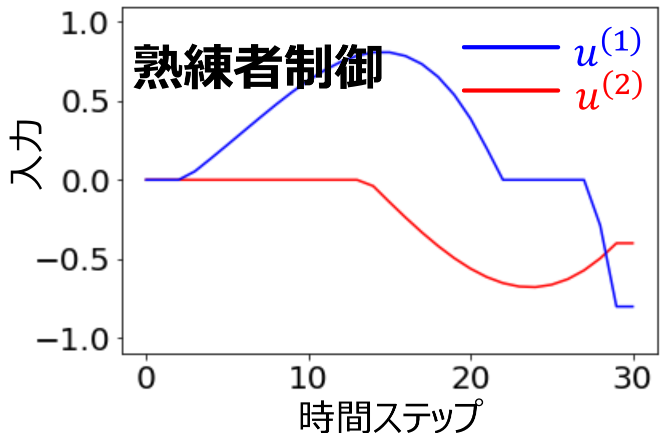
熟練者

$$U^{\text{expert}} = \operatorname*{argmin}_U \sum_{i=1}^{2} \left(q_i^{\text{e}} \overbrace{\left(y_{\text{final}}^{(i)}\right)^2}^{\text{最終時刻での原点からの偏差}} + r_i^{\text{e}} \overbrace{\sum_k \left(u_k^{(i)}\right)^2}^{\text{入力コスト}} \right)$$

誤差最小化

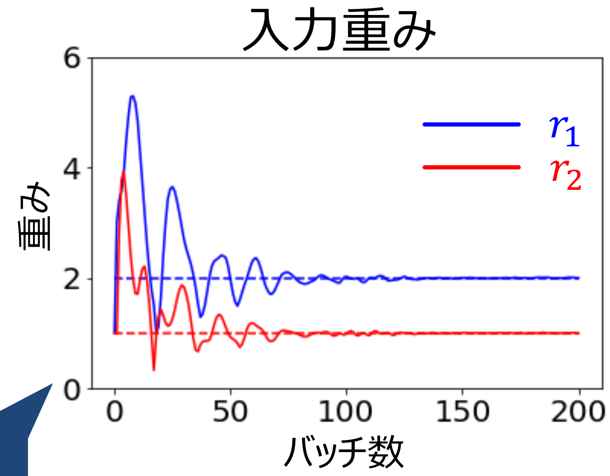
最適性条件
（リッカチ・KKT）

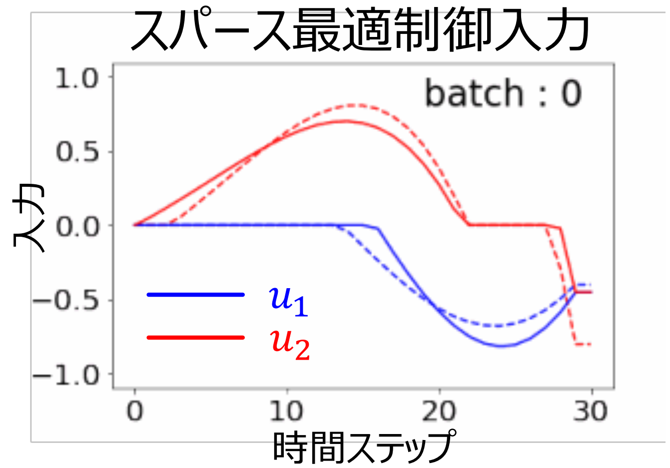
入出力データ → **モデル・コスト** → **最適制御則**



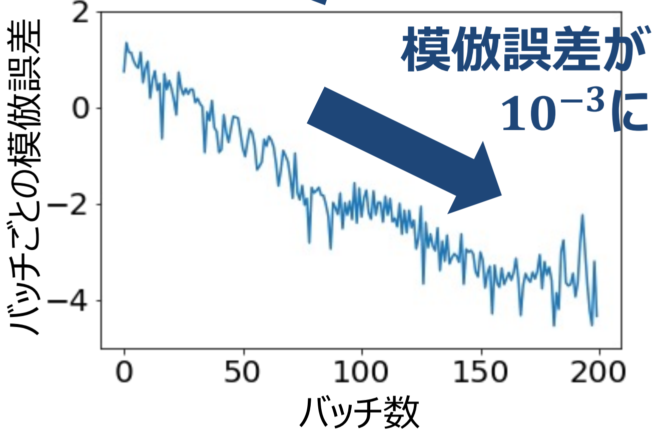
数値例

- 教師データ100個
  （初期状態をランダム生成）
- Adamにより最適化
- バッチサイズ10, エポック数20
- 初期値　$[10;\ 10;\ 1;\ 1]$
  （教師：$[60;\ 30;\ 2;\ 1]$）

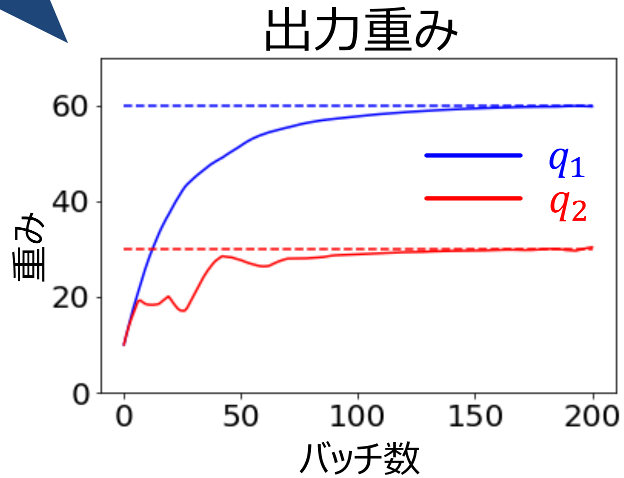


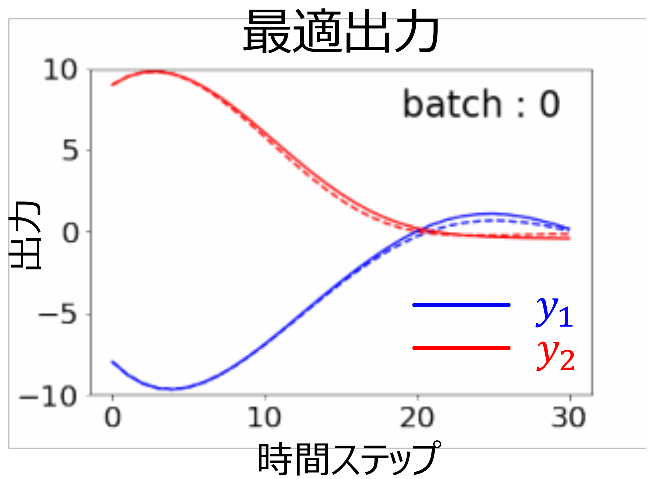


教師制御を模倣できている



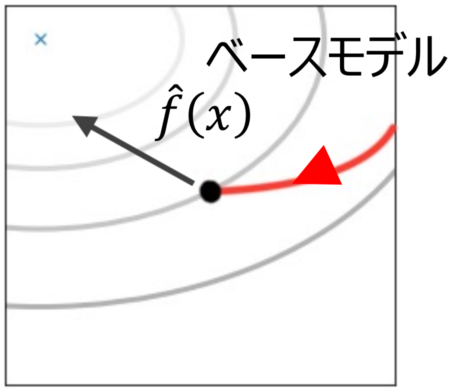
模倣誤差が 10^{-3}に

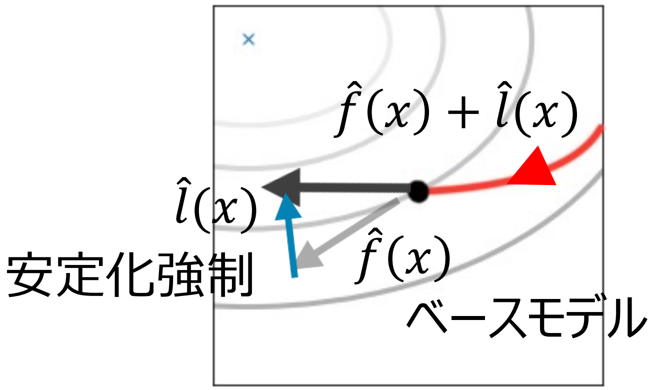




【玄人向け】

- リアプノフ関数でダイナミクスをパラメトライズ
  - $\dot{x} = f_{\mathrm{model}}(x) \coloneqq \hat{f}(x) + \hat{l}(x)$
    - ✓ $\hat{l}(x) \coloneqq \begin{cases} 0, & \text{if} \quad \nabla V^\top \hat{f} < 0 \\ -\frac{\nabla V^\top \hat{f}}{\nabla V^\top \nabla V} \nabla V, & \text{otherwise} \end{cases}$
  - Vはいつでもリアプノフ関数（モデルに沿って$\dot{V} \leq 0$）
  - $\min_{V, \hat{f}} \sum_t \| x^*(t+\tau) - x^*(t) - \tau f_{\mathrm{model}}(x^*(t)) \|^2$
    - ✓ $x^*(t)$：学習用軌道データ



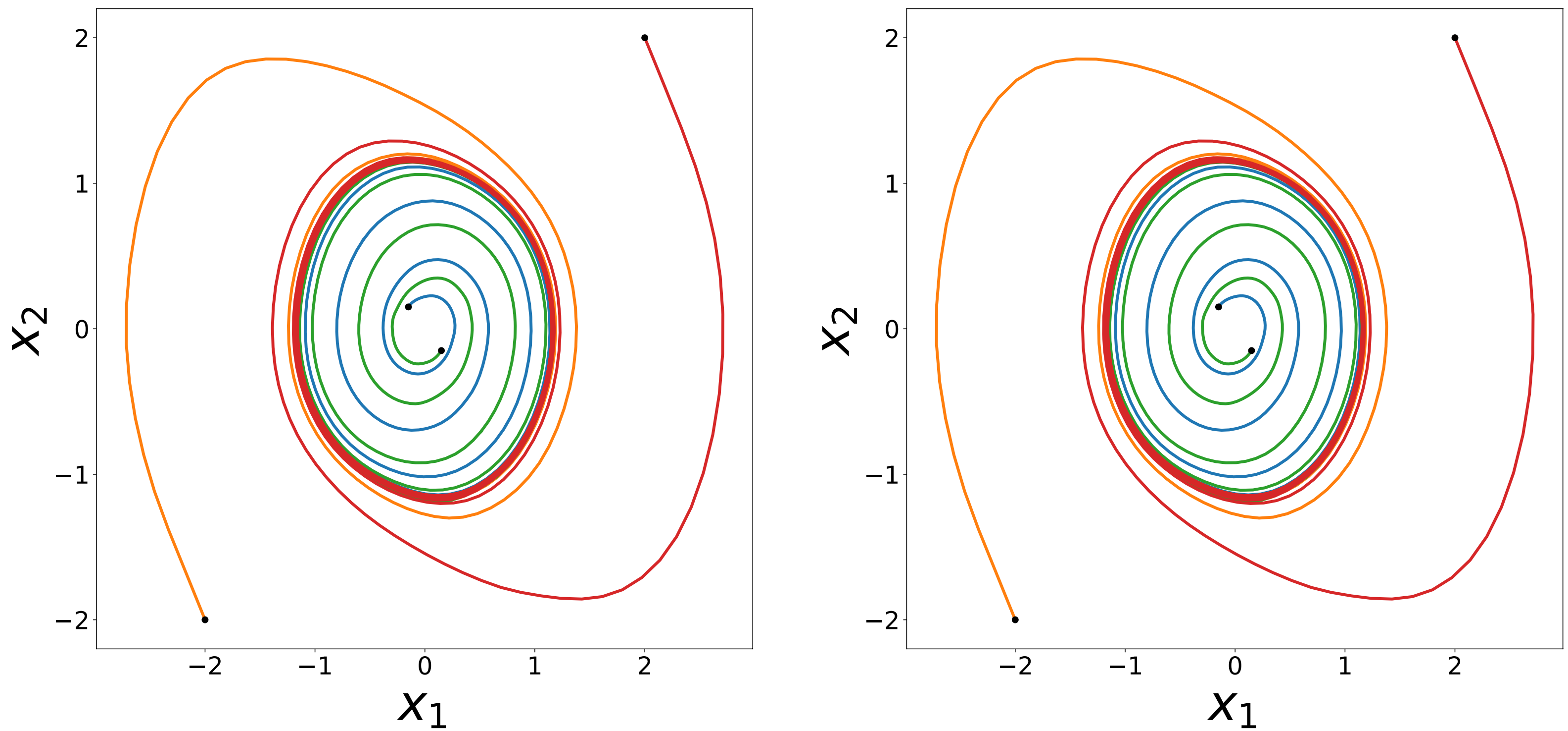


Kolter, J. Zico, and Gaurav Manek. "Learning Stable Deep Dynamics Models." *NeurIPS* (2019).

【玄人向け】

- モデルとリアプノフ関数が同時に求まる
- 不安定システムへの拡張
  - 可安定性の仮定のもと、モデル・制御リアプノフ関数・安定化制御則が求まる

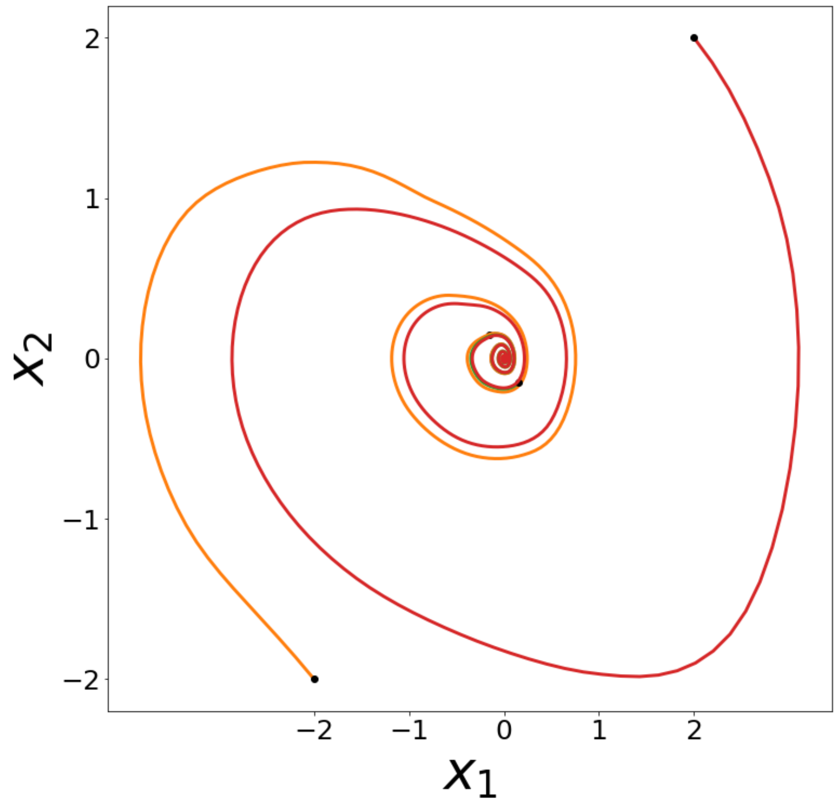
$$\dot{x} = f^*(x) \qquad \dot{x} = f_{\mathrm{model}}(x) \coloneqq \hat{f}(x) + \hat{l}(x)$$

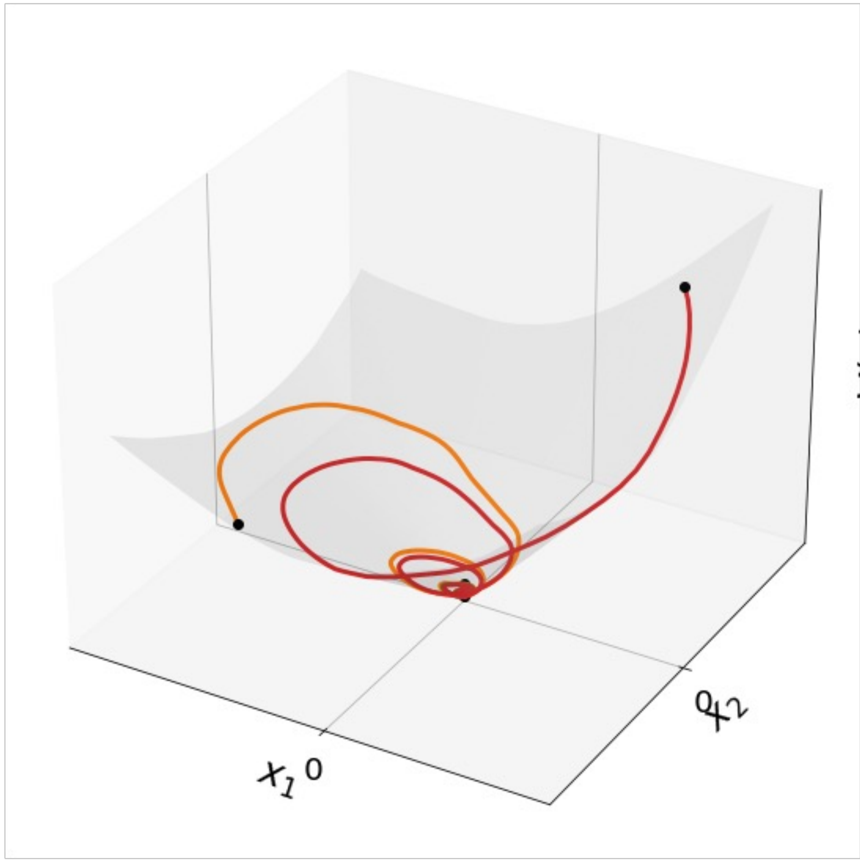


Kashima et al. "Learning Stabilizable Deep Dynamics Models." *arXiv* (2022).

【玄人向け】

- モデルとリアプノフ関数が同時に求まる
- 不安定システムへの拡張
  - 可安定性の仮定のもと、モデル・制御リアプノフ関数・安定化制御則が求まる





真のシステム・モデルともに漸近安定化される

講演概要

- 新しい強力なツールとしての機械学習
  - 事前知識やモデルベースト設計の積極的活用
  - **構造付きダイナミクス学習**

- 統計的学習と融合し深化する制御理論
  - 「情報」を陽に扱うことで生まれる異分野との接点
  - **最適輸送と最適制御**

- まとめ

制御と学習のギャップを埋める「分布の制御」

- Optimal transport, Maximum-Entropy

モビリティシェアリング

■ モビリティ基盤数理研究ユニット

- ➢ これからの移動と新たな社会のあり方を、
数理の視点から構想する

クルマや移動そのものの概念が多様化してきているなか、未来のモビリティ社会の本質をつかみ、次世代の基盤をつくるために、すべてのモノや情報が動くシステムの根本にある「数理の力」は不可欠です。

世界的に活躍する数理研究者とモビリティ・カンパニーが手を取りあい、数理の視点から、これからの移動と新たな社会のあり方を構想します。

https://mobility.amp.i.kyoto-u.ac.jp/

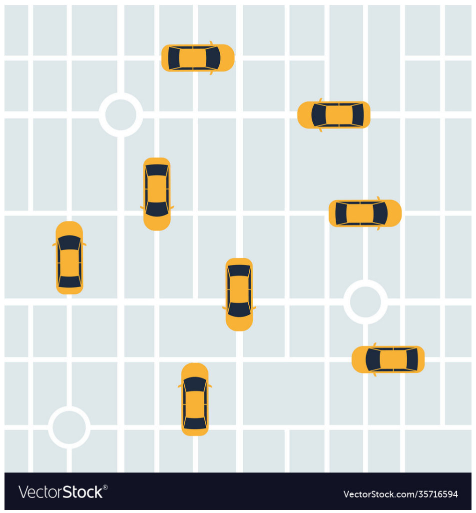
研究背景

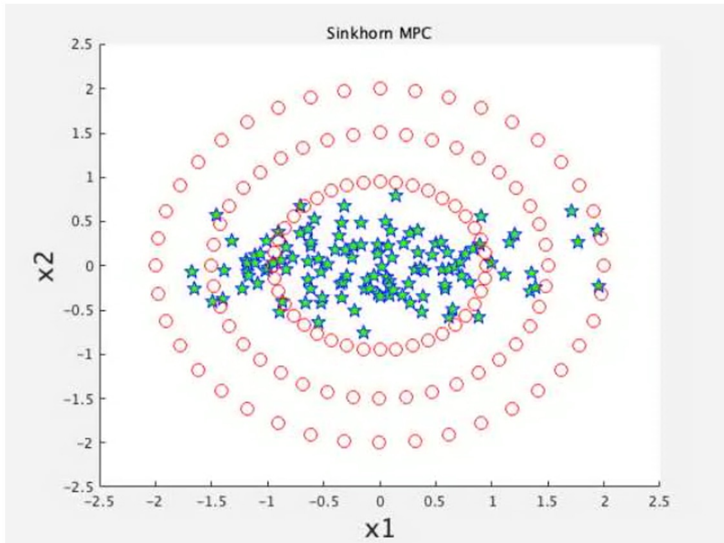
IoT，センシング技術の発展

- 超スマート社会実現にむけた大規模システムへの期待
  （群ロボット，交通システム，電力ネットワーク…）

大規模システムの制御

■ 例：群ロボットの協調運搬，効率的な配車





★：エージェント，○：目標状態

最適制御＆最適輸送

■ 最適制御：最小コストで制御目標を達成せよ

コスト $\sum_{k=0}^{N-1} L(x_k, u_k) \rightarrow \{u_k\}$ について最小化

x_N =(目標状態)

$$x_{k+1} = f(x_k, u_k)$$

x_k ：状態変数，u_k ：制御入力

■ 最適輸送：物資を目標分布に最小コストで輸送せよ

最適制御＆最適輸送

- 最適制御：最小コストで制御目標を達成せよ

  コスト $\sum_{k=0}^{N-1} L(x_k, u_k) \to \{u_k\}$ について最小化

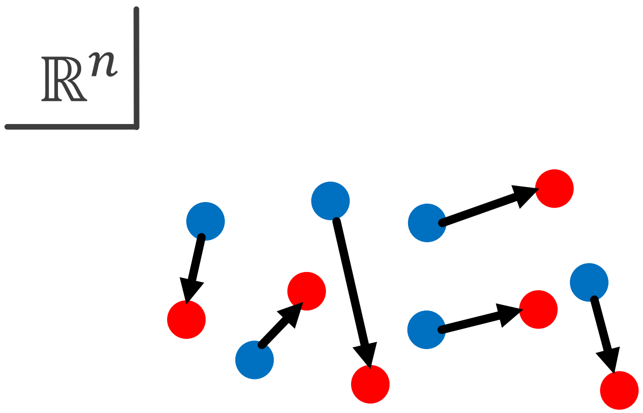
x_N =(目標状態)

$$x_{k+1} = f(x_k, u_k)$$

x_k：状態変数，u_k：制御入力

- 最適輸送：物資を目標分布に最小コストで輸送せよ

  総輸送コスト $\underset{\sigma}{\text{minimize}} \sum_{i=1}^{n} c\left(x_i, y_{\sigma(i)}\right)$, 置換 σ



$C_{ij} := c(x_i, y_j) : x_i \to y_j$ の単位輸送コスト

例）$c(x_i, y_j) = \left\| x_i - y_j \right\|^2$

輸送(制御)方法

所与の目標値に効率よく収束(安定化)させる

最適制御問題（エージェントの安定化）

$$\underset{u_i}{\text{minimize}} \quad \sum_{k=0}^{\infty} \ell\big(x_i(k), u_i(k)\big)$$

$$\text{subject to} \quad x_i(k+1) = Ax_i(k) + Bu_i(k), \quad k \geq 0,$$

$$x_i(0) = x_i^0, \quad \lim_{k\to\infty} x_i(k) = y_j$$

（$x_i(k)$：時刻 k での状態，$u_i(k)$：制御入力）

最適制御が輸送方法，最小値が輸送コスト

動的な最適輸送の定式化

初期状態 $\{x_i^0\}_{i=1}^N$, 目標状態 $\{y_j\}_{j=1}^N$

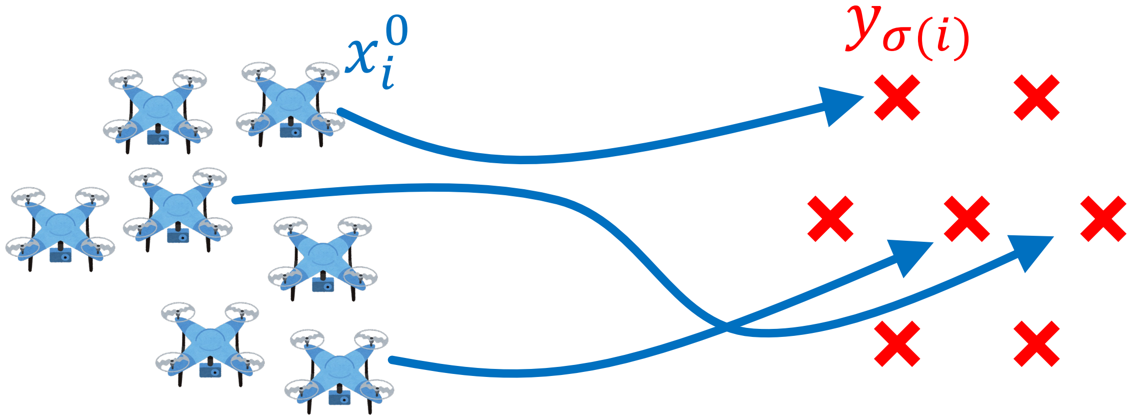
動的システム上の最適輸送（エージェント分布の安定化）

$$\underset{\{u_i\}_i,\ \sigma}{\text{minimize}} \quad \sum_{i=1}^{N}\sum_{k=0}^{\infty} \ell(x_i(k), u_i(k)) \quad \text{（総輸送コスト）}$$

$$\text{subject to} \quad x_i(k+1) = Ax_i(k) + Bu_i(k), \quad k \geq 0, \quad \forall i$$

$$x_i(0) = x_i^0, \quad x_i(\infty) = y_{\sigma(i)}.$$

$x_i \in \mathbb{R}^n$: 状態, $u_i \in \mathbb{R}^m$: 制御入力, σ : 置換



MPCをそのまま適用

$c(x_i(k), y_j)$

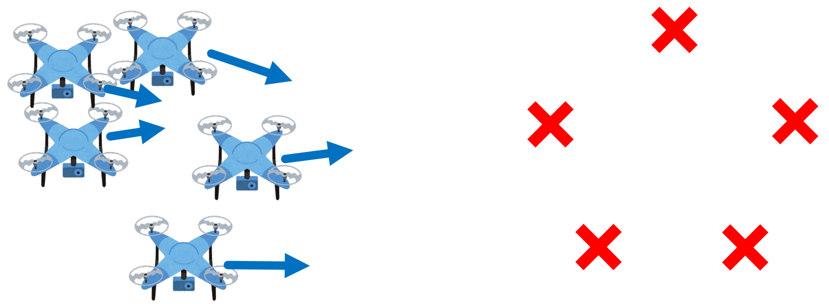
：現状態$x_i(k)$と目標y_jの輸送コスト

$$\underset{\sigma}{\text{minimize}} \sum_{i=1}^{N} c\left(x_i(k), y_{\sigma(i)}\right),$$

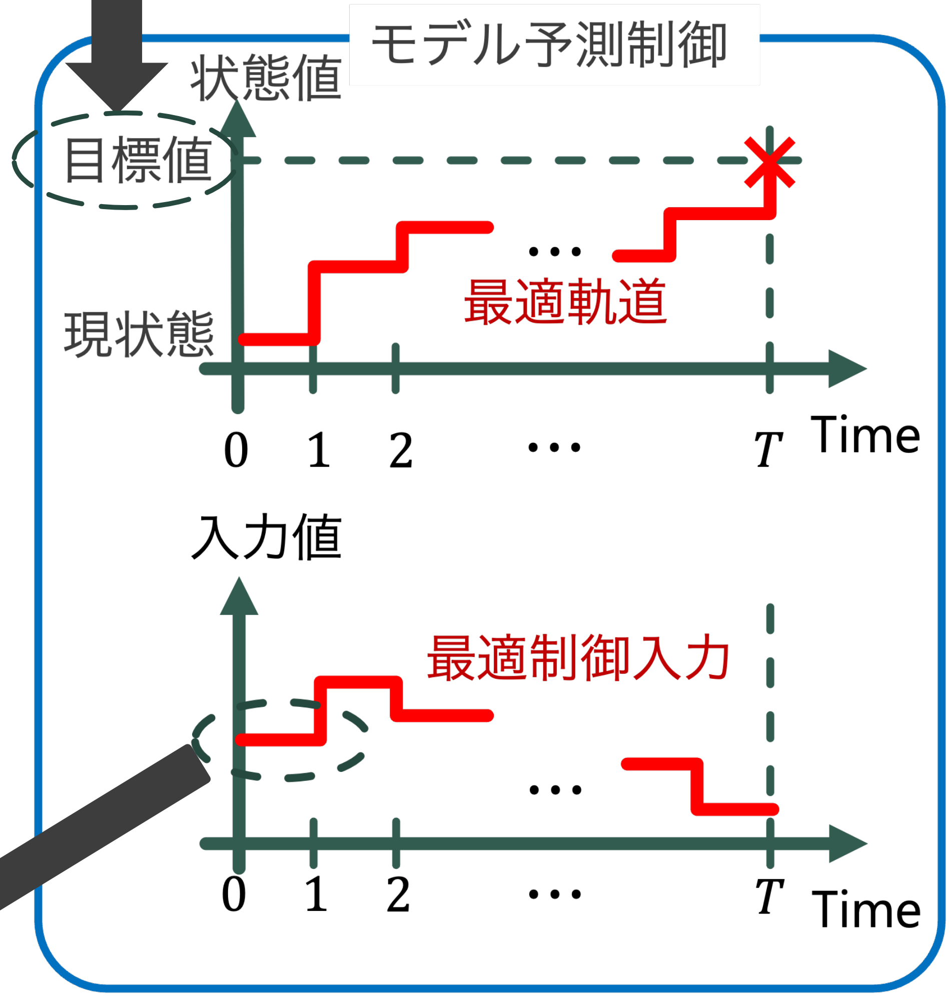
を解き，現時刻での輸送先(**暫定**)を決定 ← 計算コスト大！

σ^*から暫定目標値決定

現状態$\{x_i(k)\}$



制御入力印加



アイデア：MPC+エントロピー正則化

$C(x(k))$: $\{x_i(k)\}$と目標間の輸送コスト行列

エントロピー正則化最適輸送

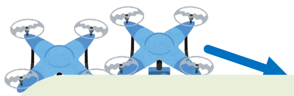
$\sum_{i,j} C_{ij}(x(k))P_{ij} - \varepsilon$(エントロピー)

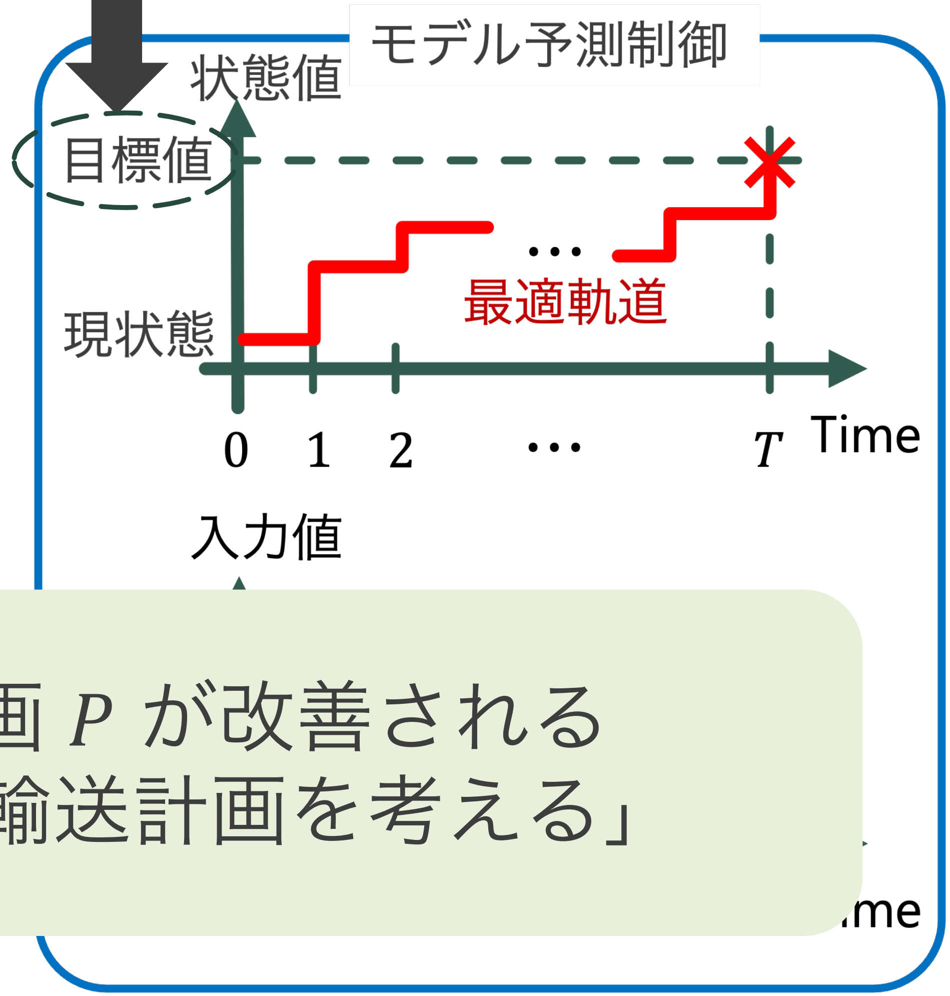
→ minimize

の最適解$P^*(x(k))$を計算

$P^*(x(k))$から暫定目標値決定

現状態$\{x_i(k)\}$





制御と共に輸送計画 P が改善される

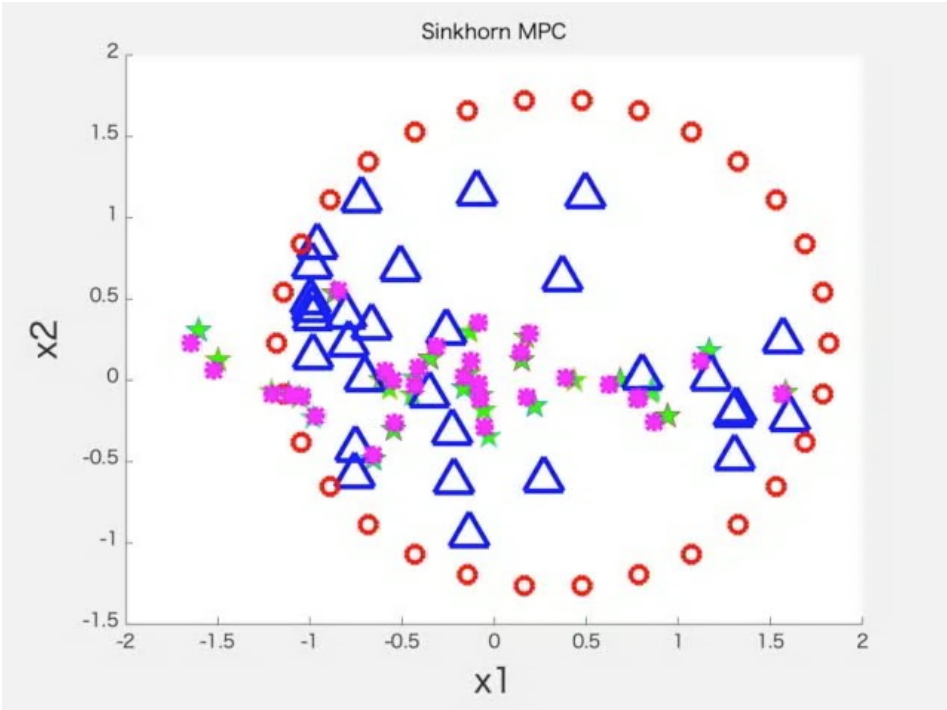
「制御しながら最適輸送計画を考える」

定理

■ 安定性などを数学的に証明

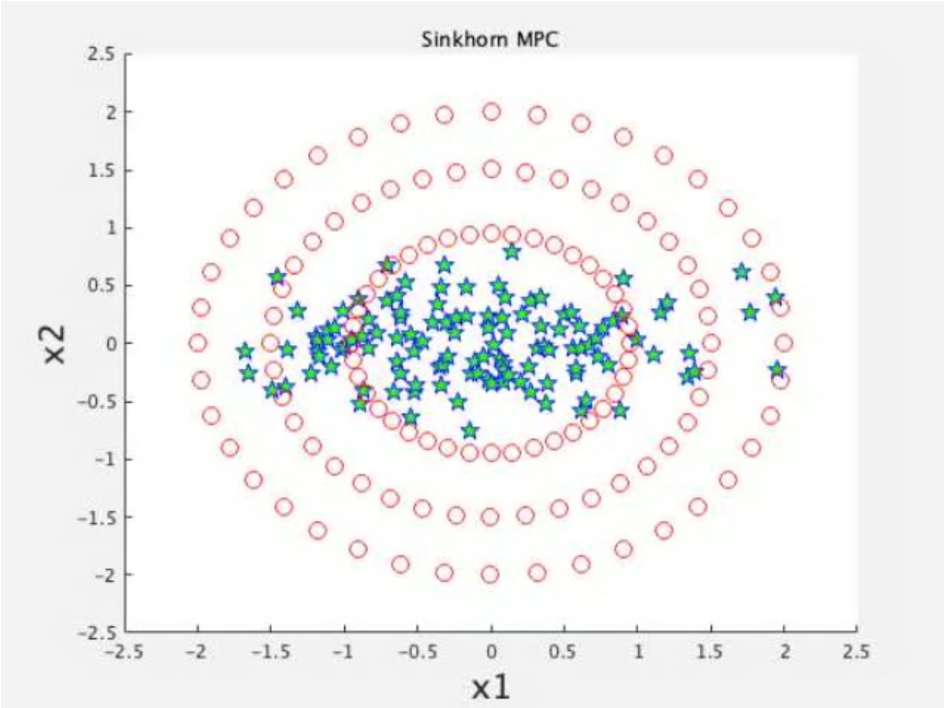
➢ 条件を満たさない場合

Magenta: エージェント初期配置, Green: エージェント
Red: ターゲット, Blue: 暫定ターゲット



Ito and Kashima, "Sinkhorn MPC: Model predictive optimal transport over dynamical systems," arXiv

数値例



★：エージェント，◯：目標状態

- $A = \begin{bmatrix} 1.2 & 0.13 \\ -0.05 & 1.1 \end{bmatrix}$, $B = 0.1I$
- 正則化の強さ： $\varepsilon = 0.5$
- MPCの予測時間： $T = 50$
- 各時刻のSinkhorn反復回数 $S_{\mathrm{iter}} = 15$

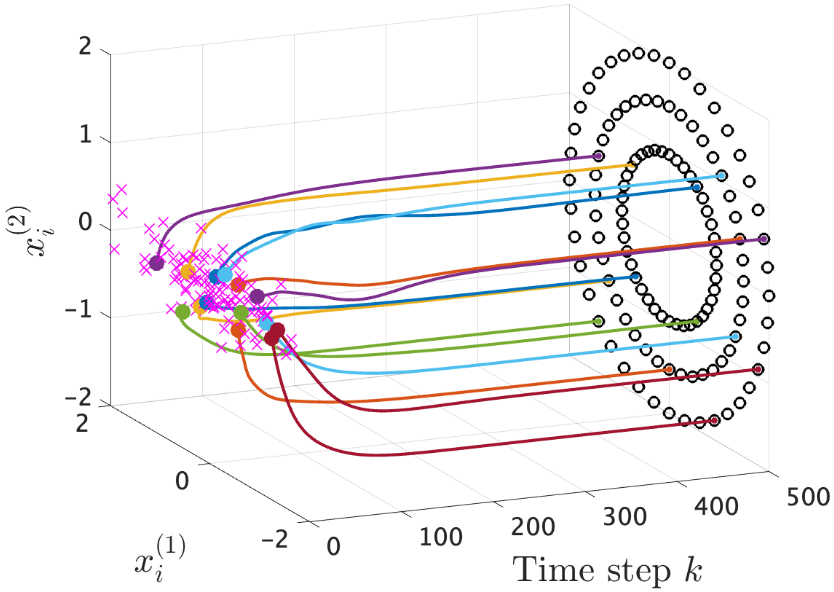
Sinkhorn1反復の計算時間		ハンガリアン法
0.21 ms (エージェント数 $N = 500$)	vs.	0.6 s ($N = 500$)
0.83 ms ($N = 1000$)		3.0 s ($N = 1000$)
3.0 ms ($N = 3000$)		57.0 s ($N = 3000$)

※収束には400反復

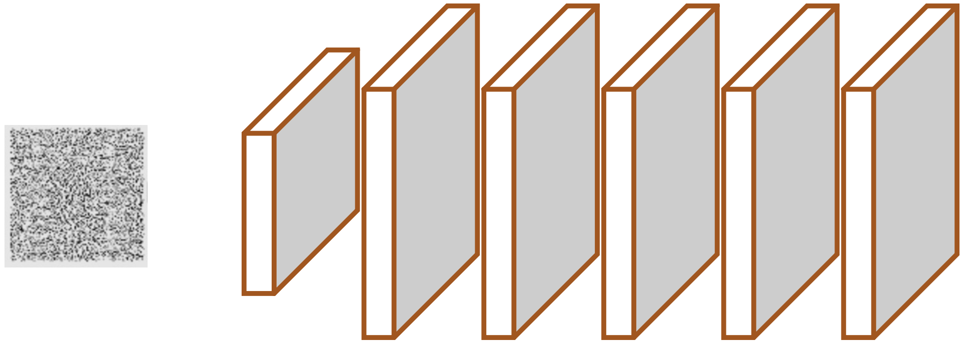
急激に増大

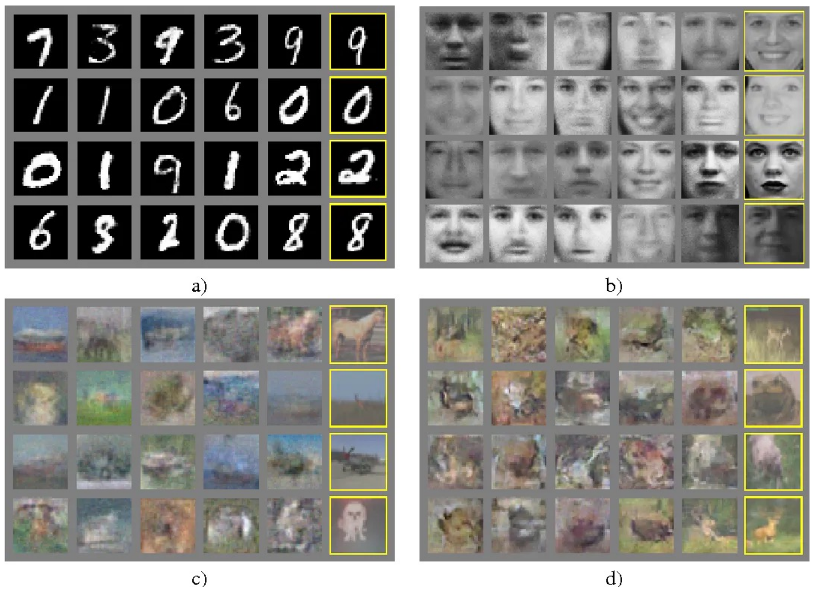
分布制御＝深層生成モデル

- 分布の制御



- 深層生成モデル





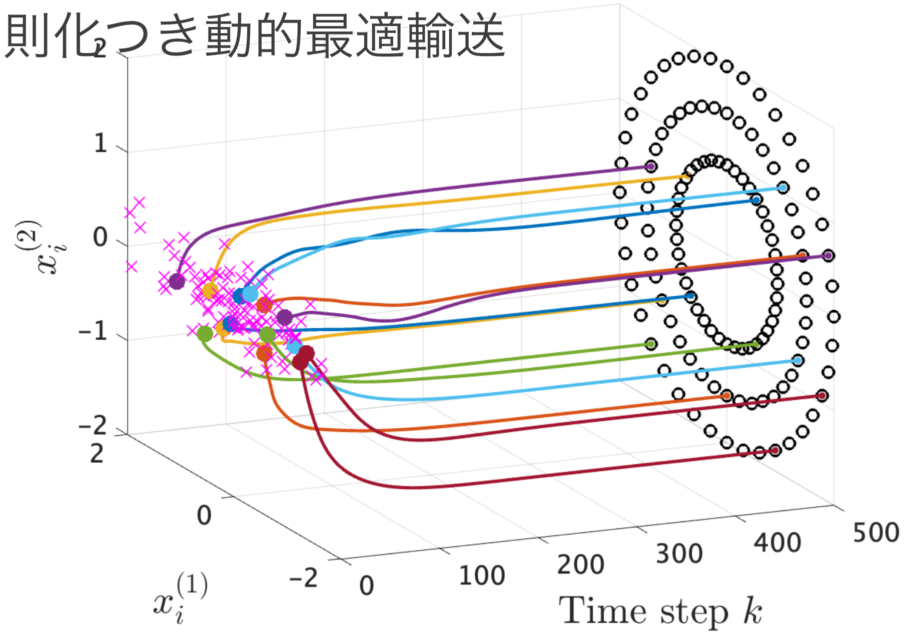
Generative adversarial nets, NeurIPS 2017

ダークサイドに堕ちた？

- せっかく綺麗な理論をやってたのに、
  とうとう**そっち側**にいってしまったんですね。。

最大エントロピー分布制御

エントロピー正則化つき動的最適輸送

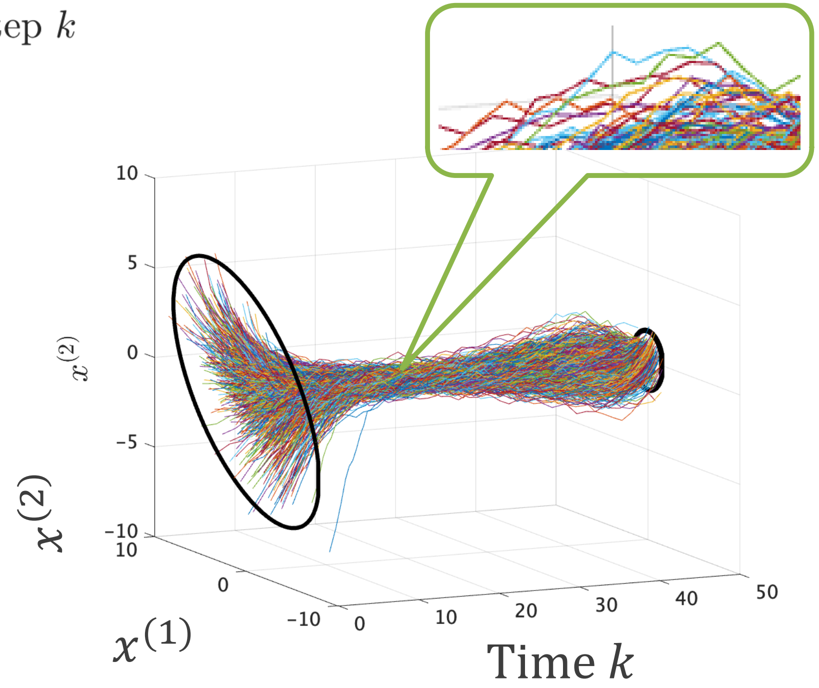


入力に雑音入れたら
ボーナスあげるで。

MaxEnt分布制御：

$$\underset{\{\pi_k\}_{k=0}^{N-1}}{\text{minimize}}\ \mathbb{E}\left[\sum_{k=0}^{N-1}\left(L(x_k,u_k)-\varepsilon \mathrm{H}(\pi_k(\cdot\,|x_k))\right)\right]$$

$$\text{subject to}\quad x_{k+1}=Ax_k+Bu_k,\quad u_k\sim\pi_k(\cdot\,|x_k)$$
$$x_0\sim\mathcal{N}(0,\bar{\Sigma}_0),\ x_N\sim\mathcal{N}(0,\bar{\Sigma}_N)$$



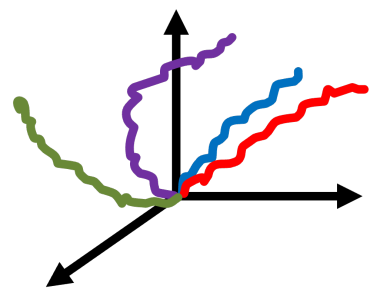
制御とエントロピー

最大エントロピー(MaxEnt)制御

$$\mathbb{E}\left[\sum_{k=0}^{N-1} L(x_k, u_k) - \varepsilon \mathrm{H}\left(\pi_k(\cdot\,|x_k)\right)\right]$$

→ 制御方策 $\{\pi_k\}$ について最小化

- なぜエントロピー正則化？
  - 強化学習
    - ✓ 学習用データのランダム探索効果[TH]
    - ✓ 雑音へのロバスト性
  - プライバシー保護
    - ✓ 信号のランダム性を増大させて推定を困難に
  - 通常の最適制御の近似解法
    - ✓ きれいな解が得られる
    - ✓ モデルフリー設計との相性もよい



状態空間の探索

[TH] T. Haarnoja, *et al.*, “Reinforcement learning with deep energy-based policies,” *ICML*, 2017.

数値例

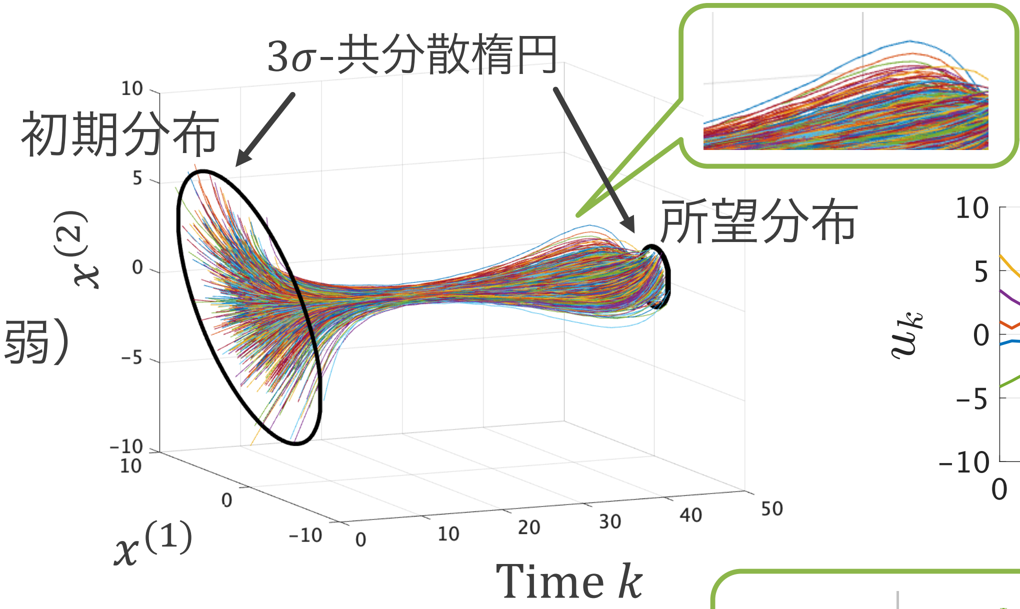
$$A = \begin{bmatrix} 0.9 & 0.1 \\ 0.05 & 1.2 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 0.22 \end{bmatrix}$$

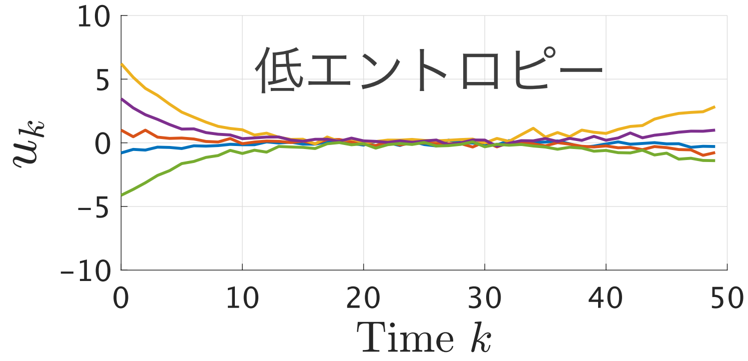
最適状態軌道の1000サンプル

制御過程のサンプル

$\varepsilon = 0.02$

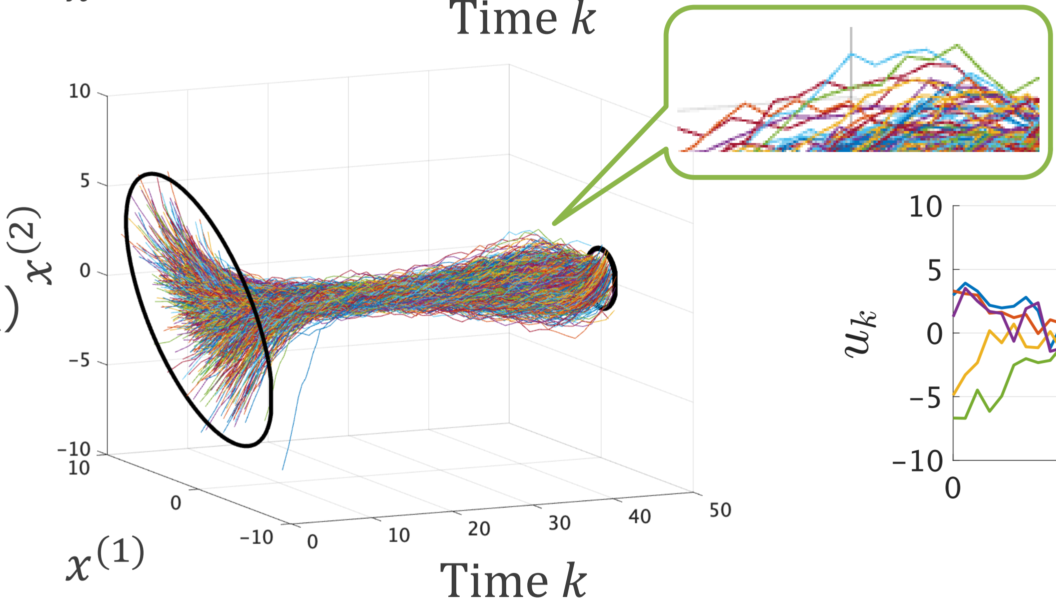
(正則化：弱)

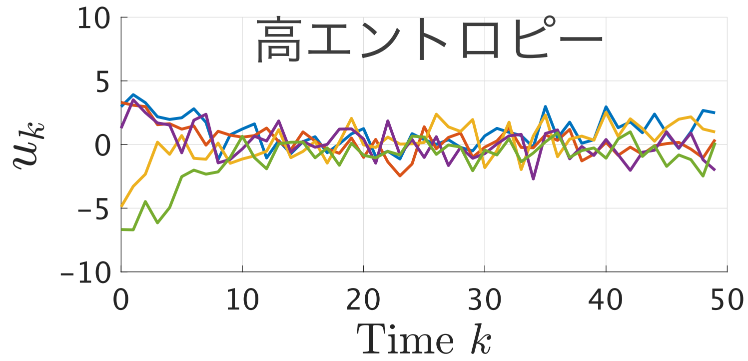




$\varepsilon = 1.0$

(正則化：強)





参考文献（2022）

- Structured Hammerstein-Wiener model learning for model predictive control,
  - Moriyasu, Ikeda, Kawaguchi, Kashima, IEEE L-CSS
- スパース最適制御問題をモデルとする模倣学習,
  - 島, 森安, 川口, 加嶋, 自動制御連合講演会
- Learning stabilizable deep dynamics,
  - Kashima, Yoshiuchi, Kawano, arXiv
- Bayesian differential privacy for linear dynamical systems,
  - Sugiura, Ito, Kashima, L-CSS
- Multiple sparsity constrained control node scheduling with application to rebalancing of mobility networks,
  - Ikeda, Sakurama, Kashima, IEEE TAC
- Observability Gramian for Bayesian inference in nonlinear systems with its industrial application,
  - Lee, Umezu, Konnno, Kashima, IEEE L-CSS, Early Access
- Entropic model predictive optimal transport over dynamical systems,
  - Ito, Kashima, submitted to Automatica
- Resilience evaluation of entropy regularized logistic networks with probabilistic cost,
  - Oishi, Hashizume, Jimbo, Kaji, Kashima, arXiv

まとめ

- 新しい強力なツールとしての機械学習
  - 事前知識やモデルベースト設計の積極的活用
  - **構造付きダイナミクス学習**

- 統計的学習と融合した深化した制御理論
  - 「情報」を陽に扱うことで生まれる異分野との接点
  - **最大エントロピー分布制御**

- 謝辞
  - 森安竜大様（豊田中研），伊藤海斗君（東工大）
  - 科研費 21H04875「情報の取得を包含した制御理論と統計的学習理論の融合数理基盤」

ご意見、雑談、講演、共同研究、入学相談、なんでも歓迎いたします。

kk@i.kyoto-u.ac.jp