

未来予測に基づく 行動決定手法の提案

情報電子工学系専攻 認知ロボティクス研究室 杉本 大志

1. 本研究の目的

- ☞ ロボットの行動獲得において、自身の過去の経験や内部状態を基に未知の環境を含む変化を予測し、この予測結果に従って最適な行動を決定することは重要である。
- ☞ このような行動獲得手法をロボットに実装し、未来の予測に基づき最適な行動をオンライン学習で決定できる手法を提案したい。
- ☞ Online SVR の手法を用いて、内部状態とロボットへの入力から、1 時刻先のデータの予測、この結果から現在の最適な行動を決定する。

2. Online SVR の予測機能の拡張

- ☞ これまでは、1次元データに対する予測を目標とした。
- ☞ ロボット制御に適用した場合、内部状態を取り扱う上で、 n 次元データに対する予測を可能となることが望ましい。

そこで

- ☞ これまでの提案手法を拡張し、 n 次元データに対する予測が可能となる予測式を立てた。

- ☞ これまで： x は 1 次元

$$\hat{y}_{N+1} = \begin{cases} 0 & \text{if } N = 0 \\ \Delta\theta & \text{if } N = 1 \\ \mathbf{k}_{sv}(x_{N+1})^T (\mathbf{K}_{sv} + \lambda \mathbf{I})^{-1} \mathbf{y}_{sv} + b' & \text{otherwise} \end{cases}$$

- ☞ 今回： x は n 次元 (任意の次元)

【アプローチ】入力データが多次元化しても、そのデータの個数によって、特徴空間における超平面の構成の可否が決まる！

$$\hat{y}_{N+1}^i = \begin{cases} 0 & \text{if } N = 0 \\ \Delta\theta & \text{if } N = 1 \\ \mathbf{k}_{sv}(\mathbf{x}_{N+1}) (\mathbf{K}_{sv} + \lambda \mathbf{I})^{-1} \mathbf{y}_{sv}^i + b & \text{otherwise} \end{cases}$$

when $i = 1, \dots, \dim(\mathbf{x})$

3. 検証実験

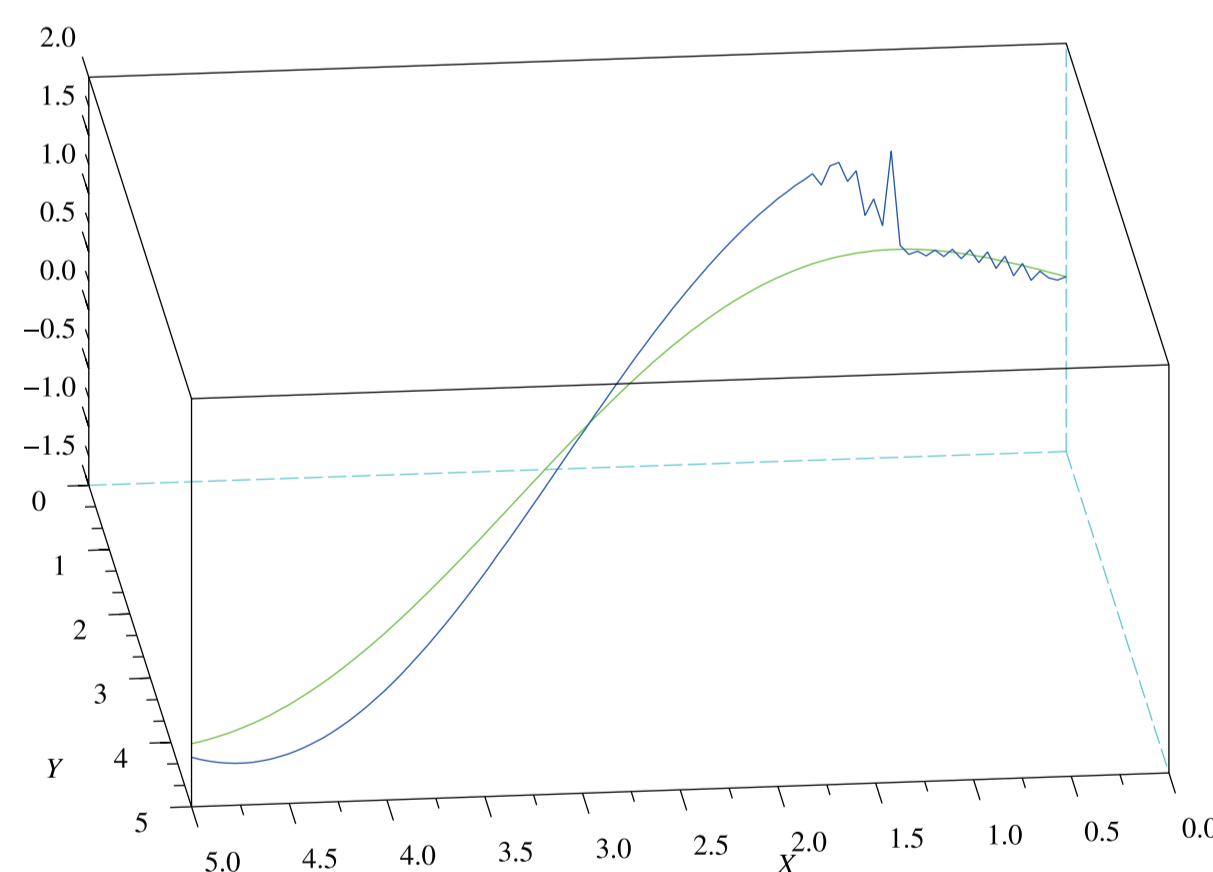
- ☞ 3次元データを与え、それぞれのデータに対する予測が出来る事を確認する。

- RBF カーネルを適用する。
- 3 時刻先まで予測を行う。
- シミュレーション時間 $0 \leq t \leq 5[\text{sec.}]$

- 与式

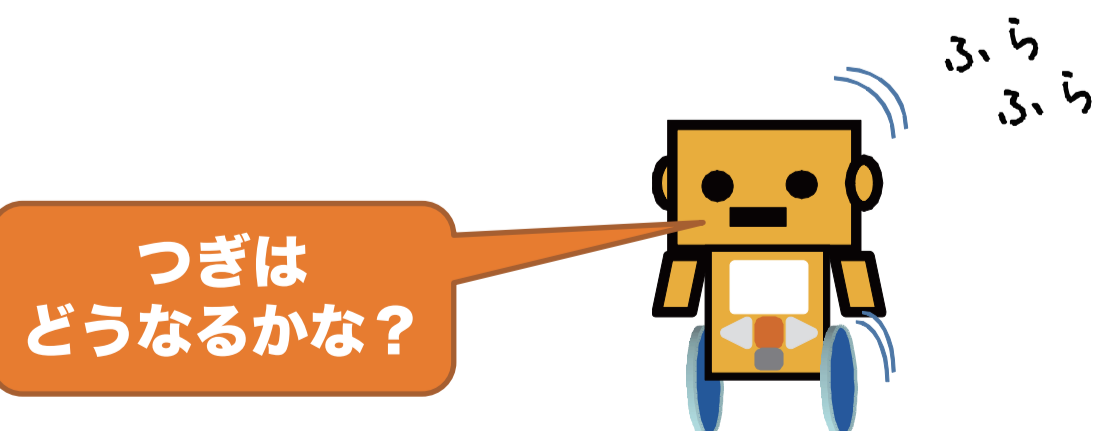
$$\mathbf{x}(t) = [\sin(t) \quad \cos(t)]$$

$$y = x_1 + x_2$$



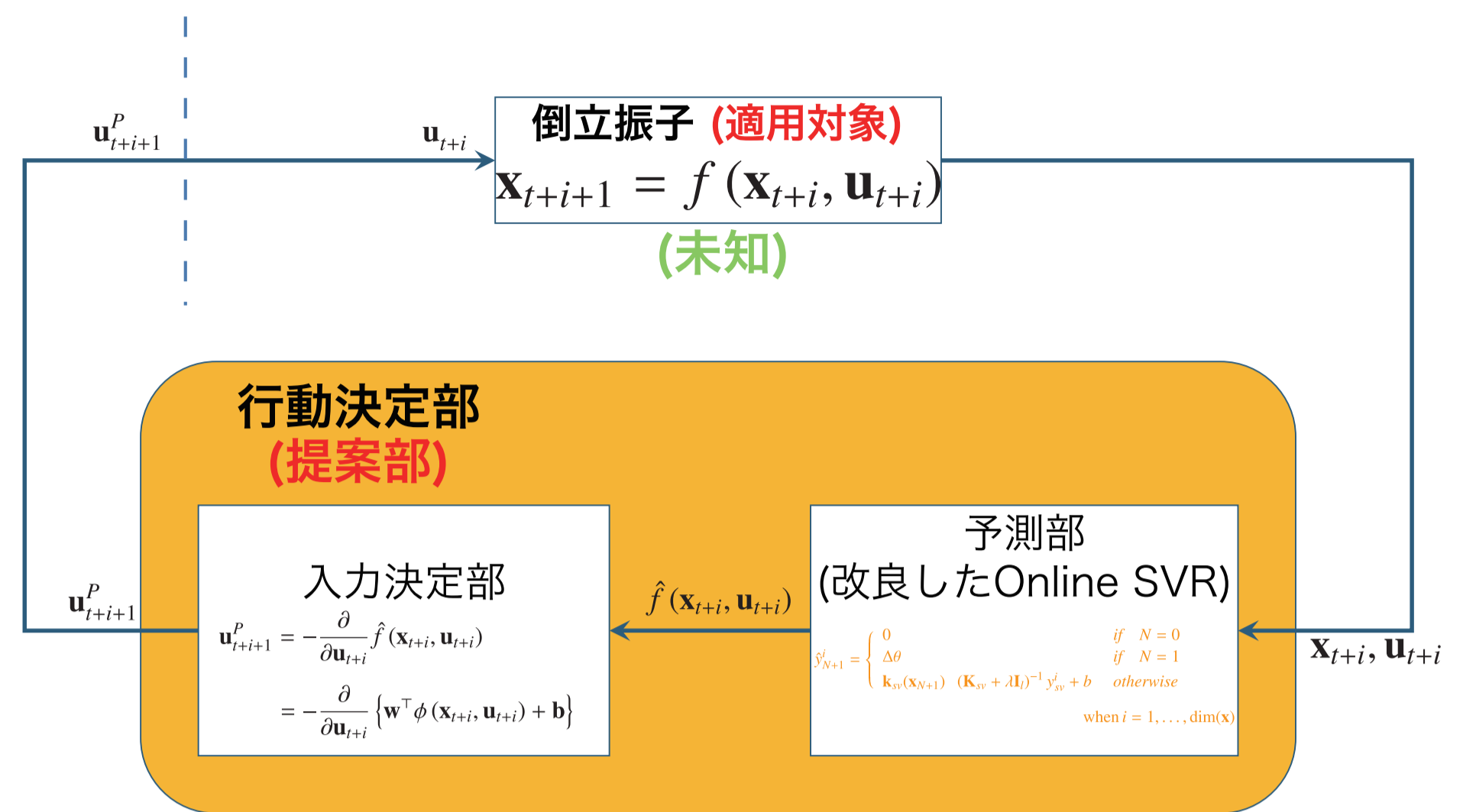
4. Online SVR を用いた行動決定の考え方

- ☞ Online SVR を用い、各時刻でリアルタイムに予測した次状態を用いている事で、ロボット自身が制御における操作量の決定に用いる。
- 次状態を目標状態に近付ける為に、現在与えた入力と 1 時刻先の予測値を用い、次時刻で与えるべき行動 (制御入力) を決定する。



適用対象を倒立振子の制御として系の次状態を予測し
最適な行動を選択するシステムを構築する

5. システム構成と処理の流れ



- ☞ 倒立振子の非線型モデルは $\mathbf{x}_{t+i+1} = f(\mathbf{x}_{t+i}, \mathbf{u}_{t+i})$ で与えられるものとする。

- ☞ 倒立振子が次時刻において取るべき行動 (入力すべき目標値) は

$$\mathbf{u}_{t+i+1}^p = -\frac{\partial}{\partial \mathbf{u}_{t+i}} f(\mathbf{x}_{t+i}, \mathbf{u}_{t+i})$$

で与えられるものとする。

- 倒立振子における目標状態は $\mathbf{x}^* \rightarrow 0$ (鉛直近傍で倒立状態を保つ) である。

- ☞ Online SVR を用い、倒立振子の内部状態を推測し、この結果を用いて行動を決定する場合は

$$\mathbf{u}_{t+i+1}^p = -\frac{\partial}{\partial \mathbf{u}_{t+i}} \hat{f}(\mathbf{x}_{t+i}, \mathbf{u}_{t+i}) = -\frac{\partial}{\partial \mathbf{u}_{t+i}} \{ \mathbf{w}^T \phi(\mathbf{x}_{t+i}, \mathbf{u}_{t+i}) + \mathbf{b} \}$$

に従うものとする。

- 実機を用いた場合は、内部状態は入出力から推定するより他ない。

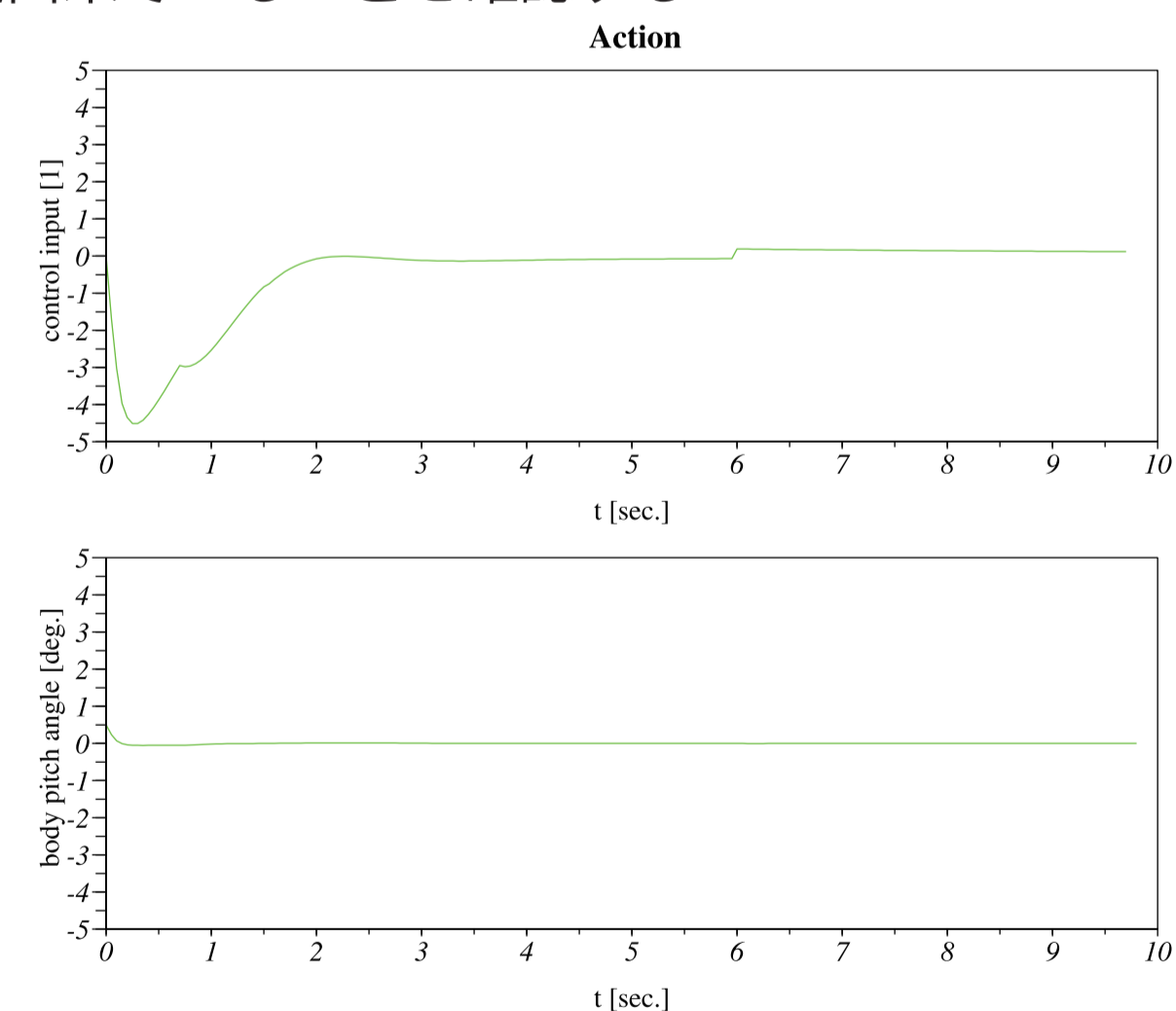
6. シミュレーション (実験結果)

- ☞ 倒立振子を対象として提案手法を適用し、その有用性を確認する。

- RBF カーネルを使う。

- 初期振り傾斜角度を 0.5[deg.] とする。

- 1 時刻先における内部状態を予測し、その結果に対する最適な行動を決定出来ていることを確認する。



7. 今後の展望

- ☞ 提案手法を洗練する。

- 提案部におけるさらなる改善。

- ☞ 長期的な予測が可能となるよう、提案手法を拡張する。

- 今回は、1 時刻先の予測と、これに対する最適行動の選択である。

- 予測に確率を導入することで、将来における予測状態の発生確率を提示し、この確率に従い最適な行動を決定することを考えてゆく。

- 【アプローチ】行動決定部において、遠い未来まで \mathbf{x} と \mathbf{u} を繰り返し予測して、それぞれの \mathbf{x} に対して \mathbf{u} を個別に導き、その収束性を見る。

【対外発表】