

多様体上の経路積分型制御

Path Integral Control on Manifold

植野 剛*1 河原 吉伸*1*2 鷺尾 隆*1*2
 Tsuyoshi Ueno Yoshinobu Kawahara Takashi Washio

*1 科学技術振興機構 *2 大阪大学 産業科学研究所
 Japan Science and Technology Agency The Institute of Scientific and Industrial Research, Osaka University

This paper presents a novel framework for solving large-scale and nonlinear stochastic optimal control problems by extracting an underlying low dimensional space for a system dynamics, and jointly constructing the optimal control strategy on the low dimensional space. In this framework, we characterize the target system as a probabilistic mixture of state-space models, and develop a new parameter estimation algorithm so as to represent the system dynamics on the low dimensional space by using a local linear function. Then, by following path integral control scheme, we derive the optimal controller by finding the optimal path distribution based on approximation inference techniques for dynamic bayesian networks. Finally, we demonstrate our proposed algorithm on a toy problem.

1. はじめに

確率最適制御における未解決問題の1つとして、大規模かつ非線形な動的システムでの最適制御問題を解くことが可能な枠組みの開発が挙げられる。一般に大規模なシステムでは、システムから得られる観測は高次元であっても、システムの本質的な次元数は観測に比べて小さい。従って、このような確率最適制御問題を解くためには、制御対象のシステムの低次元表現を抽出し、得られた低次元空間上で最適制御則を構築する方法論が必要となる。

本稿では機械学習分野で盛んに研究されている多様体学習の枠組み [Verbeek 06] を確率最適制御問題に応用し、システムの低次元表現の獲得と同時に、低次元空間で最適制御を実現する枠組みを提案する。本研究の新規性は次の2点に要約される。

1. 線形システム過程と線形観測過程の対からなる状態空間モデルを複数個、混合した**スイッチング状態空間モデル** (switching state-space model) [Murphy 02] を考える。そして、システム制御分野でよく知られる**部分空間法** [Katayama 05] を拡張して、各状態空間モデルのシステム過程が表す状態、ダイナミクスが大域的な空間での状態、ダイナミクスと一致するように、つまり、各状態空間モデルがアライメントするように、制約を課しながらシステムパラメータを推定するアルゴリズムを新たに開発する (3.2 節)。これにより、多様体上でのダイナミクスは最適制御則が考え易い局所線形関数として得ることができる。
2. **経路積分型制御** (path integral control) は、古典的な確率最適制御法と異なり、最小コストを実現する最適な経路分布を導出することで最適制御則を見つける [Kappen 09]。本研究では、この経路積分制御の枠組みを利用して、多様体上の最適経路分布を解析的に導出する (4.1 節)。そして最適経路分布の近似推論法を通して、多様体上での最適制御則の構築を実現する (4.2 節)。

最後に提案した枠組みを簡単なカオス制御問題の計算機実験に適用し、その有効性を示す。

連絡先: 植野 剛, 科学技術振興機構, 大阪府大阪市北区芝田 1 丁目 4-14 芝田町ビル 5 階 c 号室, ueno@ar.sanken.osaka-u.ac.jp

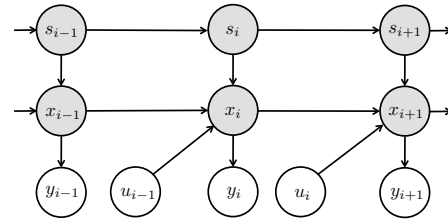


図 1: スwitching 状態空間モデル

2. 問題設定

本稿では、制御対象のシステムが図 1 に示すスイッチング状態空間モデルに従うとする [Murphy 02]。スイッチング状態空間モデルは各混合要素が状態空間モデルで与えられる確率混合モデルであり、 k 番目の混合要素は次のシステム過程 $p(x_i|x_{i-1}, u_{i-1}, s_i = k)$ と観測過程 $p(y_i|x_i, s_i = k)$ で定義される。

$$p(x_i|x_{i-1}, u_{i-1}, s_i = k) = N(x_i; A_k x_{i-1} + B_k u_{i-1}, \Gamma_k) \quad (1)$$

$$p(y_i|x_i, s_i = k) = N(y_i; C_k x_i, \Psi_k) \quad (2)$$

ここで、 $N(a; \mu, Q)$ は a 上に定義される平均ベクトル μ , 共分散行列 Q であるガウス分布である。 l 次元ベクトル y_i , h 次元ベクトル u_i はそれぞれ観測, 並びに制御入力である。 d 次元 ($d \ll l$) ベクトル x_i は潜在変数であり、システム過程の状態を表す。 s_i は混合要素のインデックスを表す 1 から K の離散変数であり、次の離散確率分布で表現される潜在ダイナミクスに従って遷移する。

$$p(s_{i+1} = k' | s_i = k) = \Pi_{k,k'} \quad (3)$$

ここで、スイッチング状態空間モデルのシステムパラメータは、 $d \times d$ 次元行列 A_k , $d \times h$ 次元行列 B_k , $l \times d$ 次元行列 C_k , $d \times d$ 次元正定値行列 Γ_k , $l \times l$ 次元正定値行列 Ψ_k , そして、遷移確率 $\Pi_{k,k'} \in [0, 1]$ である。

本研究の目的は、確率モデル (1)-(3) で特徴付けられるシステム上での最適制御則 $u_i^*(x_i)$ を導出することである*1。

*1 x_i は潜在変数であるため、実際に制御出力 u_t を生成する際は、

$$u_i^*(x_i) := \underset{u_{i:n-1}}{\operatorname{argmin}} J_i(x_i, u_{i:n-1}) \quad (4)$$

ここで、 $J_i(x_i, u_{i:n-1})$ は、終端時刻を n とする cost-to-go 関数で、以下で定義する。

$$J_i(x_i, u_{i:n-1}) = \left\langle \sum_{j=i}^{n-1} g_j(x_j, u_j, s_j, x_{j+1}, s_{j+1}) \right\rangle_p, \quad (5)$$

$g_i := g_i(x_i, u_i, s_i, x_{i+1}, s_{i+1})$ は時間依存の関数であり、各時刻でのコストである。 $\langle \cdot \rangle_p$ は確率分布 $p(x_{1:n}, s_{1:n} | x_0, u_{0:n-1})$ による期待値を表す。

本研究では、この目的を達成するために、1) 確率モデル (1)-(3) のシステムパラメータをどのように推定するか、2) どのようにして最適制御則 $u_i^*(x_i)$ を推定するのか、について議論する。これらは異なる2つの独立した問題に見えるが、実際には、互いに密接に関わりあっており、独立に考える事は好ましくない。なぜなら、最適制御則の設計は、同定したモデルに基づいて実行するため、最適制御の可否は同定するモデルの関数形に大きく依存するからである。従って、この2つの問題を別々に扱うのではなく、1) モデルの表現力が高く、かつ最適制御が簡単であるシステムパラメータの推定法、2) さまざまなシステムへの適用範囲が広い最適制御則の構築法を考える必要がある。

3. 部分空間法による多様体学習

本節では、スイッチング状態空間モデルのシステムパラメータの推定法について議論する。もし、スイッチング状態空間モデルの各要素(状態空間モデル)が、各時刻で1つのみ選択されるならば、各混合要素が表す局所的な状態、ダイナミクスは、大域的な状態、ダイナミクスと完全に一致する。つまり、システム過程の状態 x_i は低次元多様体上の座標空間とみなせ、多様体上のダイナミクスは局所線形で書けることを意味している。これは線形性を好む最適制御にとって非常に都合が良い。ゆえに、スイッチング状態空間モデルで多様体を表現するため、各要素ができるだけ離れるように(アライメントするように)、制約を課しながらシステムパラメータの推定を行うことが望ましい。

本研究では、状態空間モデルのシステムパラメータの推定法としてシステム制御分野で知られている部分空間法 [Katayama 05] を利用し、アライメントを考慮した新しいスイッチング状態空間モデルのパラメータ推定法を提案する。

3.1 部分空間法

まず最初に、 $K = 1$ のとき、つまり単一の状態空間モデルのシステムパラメータの推定問題を考え、部分空間法について簡潔に説明する。まず、 ρl 次元ベクトル f_i^+ 、 $\rho(l+h)$ 次元ベクトル w_i^- 、そして ρh 次元ベクトル u_i^+ を以下で定義する。

$$\begin{aligned} f_i^+ &= [y_i^\top, y_{i+1}^\top, \dots, y_{i+\rho-1}^\top]^\top \\ w_i^- &= [y_{i-\rho}^\top, u_{i-\rho}^\top, y_{i-\rho+1}^\top, u_{i-\rho+1}^\top, \dots, y_{i-1}^\top, u_{i-1}^\top]^\top \\ u_i^+ &= [u_i^\top, u_{i+1}^\top, \dots, u_{i+\rho-1}^\top]^\top \end{aligned}$$

f_i^+ は未来の観測の ρ ステップの部分系列、 w_i^- は過去の観測・入力の ρ ステップの部分系列、そして u_i^+ は未来の入力の ρ ステップの部分系列を表す。ここでは簡単のため、 f^+ 、 w^- 並びに u^+ の平均は 0 と仮定する。このとき、過去の観測・入力の

x_t は事後確率 $q(x_i | y_{0:i})$ の期待値 $\int x_i q(x_i | y_{0:i}) dx_i$ で置き換える。これは cost-to-go 関数 (5) を最小とする観測から出力への写像 $u_i^*(y_{0:i})$ で最適であることが知られている。詳細は [Bertsekas 07] を参照。

部分系列 w_i^- と未来の入力部分系列 u_i^+ を利用して、未来の部分系列 f_i^+ を予測する問題を考える。幾何学的な考察から、未来の観測 f_i^+ の最適な予測 \hat{f}_i^+ は、過去の観測・入力部分 w_i^- と未来の入力 u_i^+ によって張る空間への直交射影で得ることができる。

$$\hat{f}_i^+ = \hat{\mathbb{E}}_{\|u_i^+\} \{f_i^+ | W_i^-\} + \hat{\mathbb{E}}_{\|w_i^-} \{f_i^+ | U_i^+\}$$

ここで W_i^- と U_i^+ はベクトル w_i^- と u_i^+ で張るヒルベルト空間である。 $\hat{\mathbb{E}}_{\|u_i^+\} \{f_i^+ | W_i^-\}$ 、 $\hat{\mathbb{E}}_{\|w_i^-} \{f_i^+ | U_i^+\}$ は f_i^+ の U_i^+ に沿った W_i^- への平行射影、 f_i^+ の W_i^- に沿った U_i^+ への平行射影であり、それぞれ $\hat{\mathbb{E}}_{\|u_i^+\} \{f_i^+ | W_i^-\} = \Sigma_{fw|u} \Sigma_{ww|u}^{-1}$ 、 $\hat{\mathbb{E}}_{\|w_i^-} \{f_i^+ | U_i^+\} = \Sigma_{fu|w} \Sigma_{uu|w}^{-1}$ で計算できる。ここで、 $\Sigma_{ab|c} := \Sigma_{ab} - \Sigma_{ac} \Sigma_{cc}^{-1} \Sigma_{cb}$ であり、 Σ_{ab} は a と b に関する共分散行列を表す。

ここで重要な考察として、任意の $\tau \geq 0$ に関して条件付き独立性 $p(y_{i:i+\tau} | x_i, y_{0:i-1}, u_{0:i+\tau}) = p(y_{i:i+\tau} | x_i, u_{i:i+\tau})$ が成り立つ。このことは状態ベクトル x_i は、未来の出力 f_i^+ を予測するための、過去の観測・入力に関する全情報を保持していることを示している。従って、平行射影 $\hat{\mathbb{E}}_{\|u_i^+\} \{f_i^+ | W_i^-\}$ には状態 x_i が基底となる d 次元の有効部分空間が存在すると考えられる。部分空間法は、この有効部分空間を観測から復元することで状態 x_i を推定する。

有効部分空間の復元方法はこれまで多数提案されているが、単純であるが、幅広く使われている特異値分解を利用したアルゴリズムを紹介する。

アルゴリズム 1. [特異値分解による部分空間法]

1. $\hat{\Sigma}_{fw|u} \hat{\Sigma}_{ww|u}^{-1}$ 、並びに $\hat{\Sigma}_{fu|w} \hat{\Sigma}_{uu|w}^{-1}$ を計算する。ここで $\hat{\Sigma}_{ab|c} = \hat{\Sigma}_{ab} - \hat{\Sigma}_{ac} \hat{\Sigma}_{cc}^{-1} \hat{\Sigma}_{cb}$ であり、 $\hat{\Sigma}_{ab}$ は a と b のサンプル共分散行列である。
2. $L^{-1} \hat{\Sigma}_{fw|u} (M^{-1})^\top$ に対して特異値分解を行い、 d 次元特異値行列 S_d 、並びに、 S_d の各特異値に対応する左右特異値ベクトルで構成される左特異値行列 U_{1d} 、右特異値行列 U_{2d} を計算する。ここで L と M はそれぞれ、 $\hat{\Sigma}_{ff|u}$ と $\hat{\Sigma}_{ww|u}$ のコレスキー分解である。
3. 状態ベクトルの推定値 $\hat{x}_i = S_d^{1/2} U_{2d}^\top M^{-1} w_i^-$ と設定する。
4. 状態ベクトルの推定値 \hat{x}_i を利用して、モデル (1)-(2) に基づき、最小二乗法でシステムパラメータの推定値 \hat{A} , \hat{B} , \hat{C} , $\hat{\Gamma}$ 、そして $\hat{\Psi}$ を計算する。

アルゴリズム 1 は部分系列長 $\rho \rightarrow \infty$ のとき真のシステムパラメータへ一致することが保証される。詳細な理論、並びにアルゴリズムの導出は文献 [Katayama 05] を参照されたい。

3.2 部分空間法によるアライメントを考慮したスイッチング状態空間モデルのシステムパラメータ推定

本節では、前節で紹介した部分空間法を拡張して、スイッチング状態空間モデルのアライメントを考慮したシステムパラメータの推定法を提案する。もし真のシステムのダイナミクスが観測空間に埋め込まれた多様体上の局所線形関数で書ける、つまり、真のシステムがアライメントを考慮したスイッチング状態空間モデルで記述されるならば、各状態空間モデルの担当領域内で部分空間法を個別に適用することで、正しくシステムパラメータを推定することができる。本稿では、この考察に基づき、過去の観測・入力 w_i^- と未来の入力 u_i^+ による未来の観測 f_i^+ の予測問題に対して、局所線形関数に基づく確率混合モデルを導入し、観測空間を K 個の部分空間に分割し、その分割

した空間内で部分空間法を適用する 2 段階アルゴリズムを提案する。

提案手法では過去の観測・入力 w_i^- と未来の入力 u_i^+ を説明変数、未来の観測 f_i^+ を目的変数とした局所線形回帰モデルを混合要素とする混合回帰モデルについて考える。

$$p(f_i^+ | w_i^-, u_i^+, z_i = k) = N(L_{1k} w_i^- + L_{2k} u_i^+ + \mu_k, R_k) \quad (6)$$

$$p(z_i = k) = \pi_k \quad (7)$$

z_i は潜在変数であり、混合要素のインデックスを表す 1 から K の離散変数である。ここで、サンプル

$\{w_\rho^-, u_\rho^+, f_\rho^+, \dots, w_{n-\rho+1}^-, u_{n-\rho+1}^+, f_{n-\rho+1}^+\}$ が独立に確率モデル (6)-(7) に従って生成されると仮定し、パラメータである $\rho l \times \rho(l+h)$ 次元行列 L_{1k} , $\rho l \times \rho h$ 次元行列 L_{2k} , $\rho l \times \rho l$ 次元正定値行列 R_k , ρl 次元ベクトル μ_k , そして混合確率 π_k をサンプルから推定する。提案手法では、パラメータ推定に際して、確率モデル (6)-(7) による未来の観測の予測結果が大域的に局所線形関数となるように制約を課すことで、アライメントを考慮したスイッチング状態空間モデルのシステムパラメータの推定を実現する。混合分布のアライメントを考慮する方法は、これまで多数提案されているが [Verbeek 06], 本研究ではベイズ推定の枠組みを利用し、行列 L_{1k} , L_{2k} , ベクトル μ_k , 正定値行列 R_k にガウス・ウィシャート分布を事前分布として導入し、変分ベイズ推定を行う。このとき、混合比 π_k を変分自由エネルギーに対して点推定することで、関連度自動決定を行い、冗長な混合要素の混合比が 0 に縮減し、混合要素のアライメントを促進させる [Bishop 06]。

確率モデル (6)-(7) のパラメータが決定したら、 z_i の事後分布に基づき、観測空間を分解し、それぞれに部分空間法 (アルゴリズム 1) を適用することで、システムパラメータを推定することができる。以下に提案アルゴリズムをまとめる。

アルゴリズム 2. [部分空間法による多様体学習]

1. 変分ベイズ法により、確率モデル (6)-(7) のパラメータを推定する。
2. 1. で求めたパラメータを用いて、混合要素に関する事後確率 $q(z_i | f_i^+, w_i^-, u_i^+)$ を計算する。ここで、 $s_i = \operatorname{argmax}_k q(z_i = k | f_i^+, w_i^-, u_i^+)$ と設定し、遷移確率 $\Pi_{k,k'}$ の推定値 $\hat{\Pi}_{k,k'}$ を計算する。
3. 各混合要素内でアルゴリズム 1 を適用し、システムパラメータの推定値 $\hat{A}_k, \hat{B}_k, \hat{C}_k, \hat{\Gamma}_k$, そして $\hat{\Psi}_k$ を求める。部分空間法の適用に必要な統計量は事後確率 $q(z_i | f_i^+, w_i^-, u_i^+)$ の重み付けを利用して計算する。

4. 多様体上の確率最適制御

本節では、スイッチング状態空間モデル上での最適制御の構築法について議論する。古典的な確率最適制御では、関数近似を用いて最適 cost-to-go 関数 $J_i^*(x_i) = \min_{u_{i:n}} J_i(x_i)$ をベルマン方程式を通して推定する。しかし、予め未知である多様体上で関数近似器の設計を行うのは容易ではない。本稿では、経路積分型制御の枠組み [Kappen 09] を踏襲し、最適 cost-to-go 関数の推定を直接行わず、最適制御下での軌道の確率分布 (最適経路分布) を求めることで、最適制御則を導く。

4.1 最適経路分布の導出

本稿では、コスト関数を以下の関数に制限した確率最適制御問題を考える。

$$g_i := \ell_i + \ln \frac{p(x_{i+1} | x_i, s_{i+1}, u_i)}{m(x_{i+1} | x_i, s_{i+1})}, \quad (8)$$

ここで、関数 $\ell_i := \ell(x_i, s_i)$ は任意の関数である。 $m(x_{i+1} | x_i, s_{i+1})$ は無制御時のダイナミクスであり、 $m(x_{i+1} | x_i, s_{i+1}) := p(x_{i+1} | x_i, u_i = 0, s_{i+1})$ を意味する。コスト (8) を (5) に代入すると cost-to-go 関数は次の式で書き直すことができる。

$$J_i(x_i, u_{i:n-1}) = \left\langle \sum_{j=i}^n \ell_j \right\rangle_{p_i} + D(p_i \| m_i), \quad (9)$$

ここで $D(p_i \| m_i)$ は分布 $p(x_{i+1:n}, s_{i+1:n} | x_i, s_i, u_{i:n-1})$ と分布 $\prod_{j=i}^{n-1} m(x_{j+1} | x_j, s_{j+1})$ のカルバック・ライブラー (KL) ダイバージェンスである。式 (9) の右辺の第 2 項は、制御時のダイナミクスが無制御時のダイナミクスから離れるほど大きくなるため、制御入力に対する罰則項と解釈できる。これは入力エネルギー最小化の観点から妥当な制約と考えられるため、コスト (8) が導く最適制御問題のクラスに限定して議論することは、一般性を大きく損うことはない*2。

経路型積分制御は cost-to-go 関数 (9) を $u_{i:n-1}$ の関数としてでなく、経路分布に関する関数と考え、cost-to-go 関数 (9) を経路分布に関して最適化する。

定理 1. Cost-to-go 関数 (9) を最小とする最適経路分布 $p^*(x_{1:n}, s_{1:n} | x_0, s_0)$ は次の式で計算できる。

$$p^*(x_{1:n}, s_{1:n} | x_0, s_0) \propto \prod_{i=0}^{n-1} m(x_{i+1} | x_i, s_{i+1}) p(s_{i+1} | s_i) \exp \left[- \sum_{i=0}^n \ell_i \right] \quad (10)$$

また最適な制御入力 $u_i^*(x_i)$ は以下の式で計算できる。

$$u_i^*(x_i) = \langle \bar{u}_i^*(x_i, x_{i+1}, s_{i+1}) \rangle_{p_i^*} \quad (11)$$

$\langle \cdot \rangle_{p_i^*}$ は $p^*(x_{i+1}, s_{i+1} | x_i)$ による期待値である。また

$$\begin{aligned} \bar{u}_i^*(x_i, x_{i+1}, s_{i+1} = k') \\ := \left(B_{k'}^\top \Gamma_{k'}^{-1} B_{k'} \right)^{-1} B_{k'}^\top \Gamma_{k'}^{-1} (x_{i+1} - A_{k'} x_i) \end{aligned}$$

注目すべき点として、最適制御入力 $u_i^*(x_i)$ の計算に必要な周辺分布 $p^*(x_{i+1}, s_{i+1} | x_i, s_i)$ は以下の方程式から再帰的に計算可能である。

$$p^*(x_{i+1}, s_{i+1} | x_i, s_i) = \sum_{s_{i+2}} \int m(x_{i+1} | x_i, s_{i+1}) p(s_{i+1} | s_i) \cdot \exp[-\ell_{i+1}] p^*(x_{i+2}, s_{i+2} | x_{i+1}, s_{i+1}) dx_{i+2}$$

この再帰方程式は、ポテンシャル関数を

$m(x_{i+1} | x_i, s_{i+1}) p(s_{i+1} | s_i) \exp[-\ell_{i+1}]$ と定義したマルコフ鎖における後ろ向きメッセージ伝搬と見ることができる。つまり、コスト関数 (9) で特徴付けられる最適制御問題のクラスは、**動的ベイジアンネットワークの確率推論問題**と解釈できる。

しかし、最適経路分布 (10) は混合分布であるため、混合要素数が後ろ向きメッセージ伝搬の回数に対して指数的に増大するため、解析的に計算することは困難である。したがって、本稿では前向き-後ろ向き型メッセージ伝搬を利用した最適経路分布の近似法を提案する。

4.2 最適経路分布の近似推論

最初に時刻 i から時刻 $i+1$ への前向きメッセージ伝搬について考える。時刻 i での前向きメッセージ $\alpha_i := \alpha_i(x_i, s_i)$ を以下で定義する。

*2 実際には、経路積分型制御以外に Linear solvable Markov decision process (LSMDP) な線形 2 次ガウスレギュレータ等有名な確率最適制御法もこのクラスに含まれる [Kappen 09]。

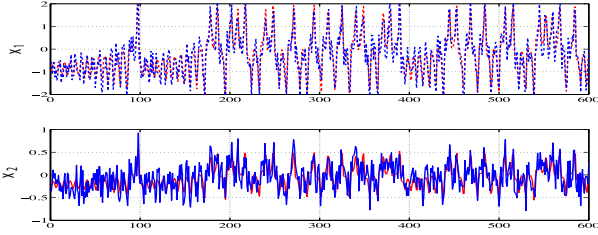


図 2: 予測結果. 破線が真のダイナミクス, 実線が推定したモデルによる予測結果.

$$\alpha_i := \sum_{s_{0:i-1}} \int p^*(x_{1:i}, s_{1:i}|x_0, s_0) p(x_0, s_0) dx_{0:i-1}$$

このとき, 時刻 $i+1$ の前向きメッセージ α_{i+1} は, 時刻 i の前向きメッセージ α_i を利用して次の再帰方程式から計算することができる.

$$\alpha_{i+1} = \sum_{s_i} \int \alpha_i \exp[-\ell_i] m(x_{i+1}|x_i, s_{i+1}) p(s_{i+1}|s_i) dx_i$$

もし α_i がガウス分布であり, ℓ_i が x_i の 2 次関数で書けるならば, GPB2 (generalized pseudo-Bayesian of order 2) アルゴリズムに代表される collapse 演算を用いた近似アルゴリズムを適用できる. そうでない場合, ラオ・ブラックウェル粒子フィルタなどサンプリング法を駆使して数値的にメッセージの計算を行うことになる.

同様に時刻 $i+1$ から時刻 i への後ろ向きメッセージ伝搬について考える. 時刻 i の後ろ向きメッセージ $\gamma_i := \gamma_i(x_i, s_i)$ は次の式で定義する.

$$\gamma_i := \sum_{s_{i+1:n}} \int p^*(x_{i:n}, s_{i:n}) dx_{i+1:n}$$

時刻 $i+1$ の後ろ向きメッセージ γ_{i+1} と時刻 i の前向きメッセージ α_i を利用して次の再帰方程式で計算できる.

$$\gamma_i = \sum_{s_i} \int \frac{\alpha_i \exp[-\ell_{i+1}] m(x_{i+1}|x_i, s_{i+1}) p(s_{i+1}|s_i) \gamma_{i+1}}{\sum_{s_{i+1}} \int \alpha_i \exp[-\ell_{i+1}] m(x_{i+1}|x_i, s_{i+1}) p(s_{i+1}|s_i) dx_i} dx_{i+1}$$

後ろ向きメッセージに関しても前向きメッセージ同様の近似法を用いて, メッセージを計算できる.

時刻 i の前向きメッセージの近似 $\hat{\alpha}_i$, 時刻 $i+1$ の後ろ向きメッセージの近似 $\hat{\gamma}_{i+1}$ を用いて, 周辺分布 $p^*(x_{i+1}, s_{i+1}|x_i, s_i)$ の近似分布 $\hat{p}(x_{i+1}, s_{i+1}|x_i, s_i)$ をベイズ則より導出できる.

$$\hat{p}(x_{i+1}, s_{i+1}|x_i, s_i) = \frac{\hat{p}(s_i, s_{i+1}, x_i, x_{i+1})}{\hat{p}(x_i, s_i)}$$

ただし,

$$\hat{p}(s_i, s_{i+1}, x_i, x_{i+1}) = \hat{\alpha}_i \exp[-\ell_{i+1}] m(x_{i+1}|x_i, s_{i+1}) p(s_{i+1}|s_i) \hat{\gamma}_{i+1}$$

$$\hat{p}(x_i, s_i) = \sum_{s_{i+1}} \int \hat{p}(s_i, s_{i+1}, x_i, x_{i+1}) dx_{i+1}$$

である. これを用いて式 (11) より最適制御入力を計算することができる.

5. 数値実験

提案手法の有効性を検証するため, 次のカオス制御問題を考える.

$$x_{1,i+1} = 1.9x_{1,i} - x_{1,i}^3 + x_{2,i} + u_i + \epsilon_{x,i}$$

$$x_{2,i+1} = 0.5x_{1,i} + \epsilon_{x,i}$$

$$y_{1,i} = 0.2x_{1,i} + 0.3x_{2,i} + \epsilon_{y,i}$$

$$y_{2,i} = 0.6x_{1,i} - 0.5x_{2,i} + \epsilon_{y,i}$$

$$y_{3,i} = -0.1x_{1,i} + 1.2x_{2,i} + \epsilon_{y,i}$$

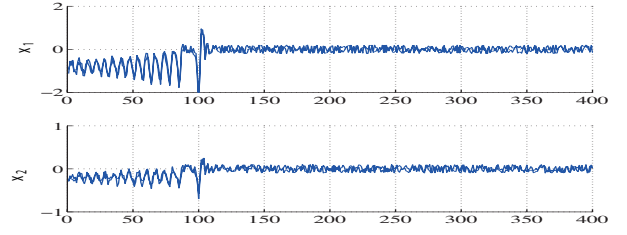


図 3: 最適制御の結果.

ただし, ϵ_x, ϵ_y はそれぞれガウス分布 $N(\epsilon_x|0, 0.05^2)$, $N(\epsilon_y|0, 0.1^2)$ とする. $a_{i,b}$ はベクトル a_i の第 b 成分を意味する. 制御コスト g_i は式 (8) の関数形をとるとし, $\ell(x_i) := 0.2 \cdot x_i^\top x_i$ と設定する. このとき, cost-to-go 関数 (9) を最小とする最適制御則 u_i^* を導出する.

本実験では, まず最初に -1.5 から 1.5 のランダムな制御入力を加え, 時刻 $i=0$ から時刻 $i=600$ までサンプル系列を 2 つ生成し, そのうち 1 つを用いて, スwitching 状態空間モデルのシステムパラメータの推定を行い, もう一方を利用して, 推定したモデルに基づく予測の性能を評価した. そして, 推定したモデルを元に最適制御則を構築し, 最適制御入力を加えたとき, システムダイナミクスの遷移を考察した.

図 2 は検証用サンプル下での, 推定したモデルに基づく x_1, x_2 のフィルタリング結果 (実線) と, 実際のダイナミクス (破線) を示している. 図より, 提案手法が正確に実際のダイナミクスを再現している様子が見て取れる. 図 3 は時刻 $i=100$ まででは非制御とし, 時刻 $i=101$ から時刻 400 まで制御を行った時の x_1, x_2 の遷移を描いている. 時刻 $i=101$ 以降は x_1, x_2 ともに高速に 0 に近づき, 安定している様子が見てとれる. 以上より, 我々の提案手法は一定の有効性を示すことを確認できた.

6. 結論

確率最適制御では, システムのモデリング, 最適制御則の導出を常に対として考慮する必要がある. 本稿では部分空間法を利用して, 最適制御に適した多様体構造を導く多様体学習法を導出した. また, 最適経路分布の近似推論を通じて多様体上での最適制御則の構築法を開発した. 最後に簡単なシミュレーション実験により提案アルゴリズムの性能を検証し, 適切な制御則を導くことが確認できた.

参考文献

- [Bertsekas 07] Bertsekas, D. P.: *Dynamic Programming and Optimal Control*, Athena Scientific (2007)
- [Bishop 06] Bishop, C. M.: *Pattern Recognition and Machine Learning*, Springer New York (2006)
- [Kappen 09] Kappen, B., Gomez, V., and Opper, M.: Optimal control as a graphical model inference problem, Technical Report arXiv:0901.0633, Arxiv preprint (2009)
- [Katayama 05] Katayama, T.: *Subspace methods for system identification*, Springer Verlag (2005)
- [Murphy 02] Murphy, K. P.: *Dynamic bayesian networks: representation, inference and learning*, PhD thesis, University of California (2002)
- [Verbeek 06] Verbeek, J.: Learning nonlinear image manifolds by global alignment of local linear models, *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 28, No. 8, pp. 1236–1250 (2006)