

Latent Dynamics 研究会資料

次元削減と動的システムの学習 について

東京大学先端科学技術研究センター

矢入健久

2010年6月16日

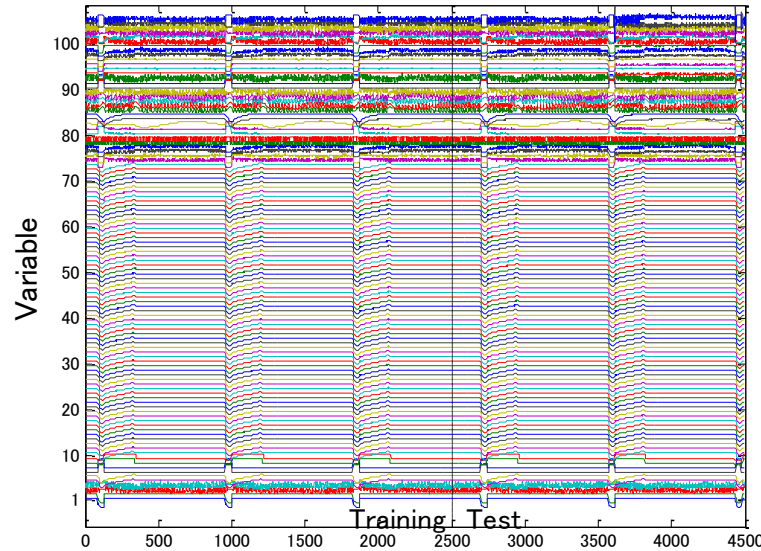
本日の話題

1. はじめに
 - 動機例題
2. 動的システムの学習の研究動向
 - 機械学習分野での動向
 - システム同定分野での動向
3. 部分空間同定法の整列による非線形動的システムの学習
4. まとめ

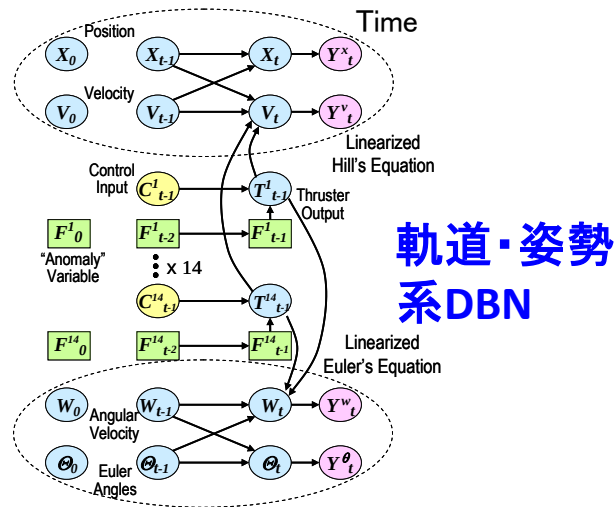
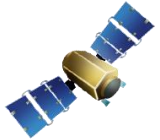
1. はじめに

動機例題1:

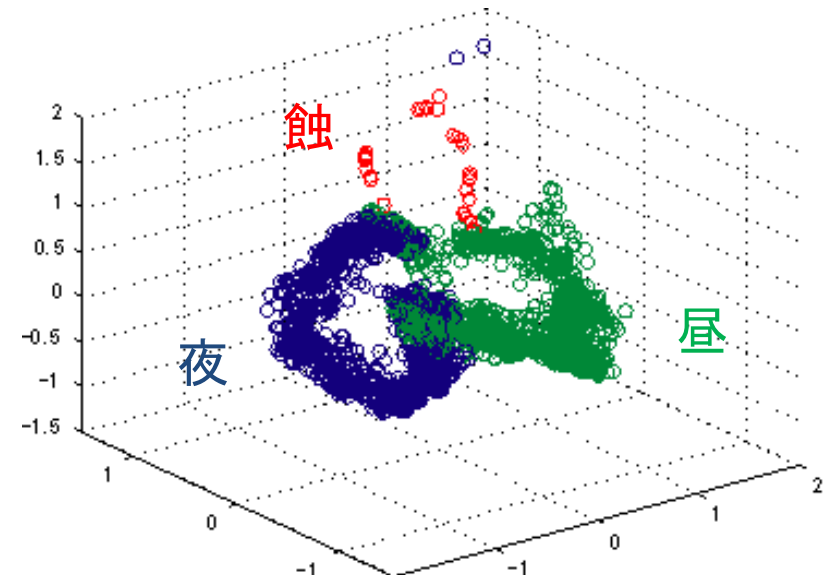
次元削減による人工衛星テレメトリ監視



- 数百次元の時系列
- モデル化は限定的
- 本質的次元は低い
- 非線形かつモード間遷移



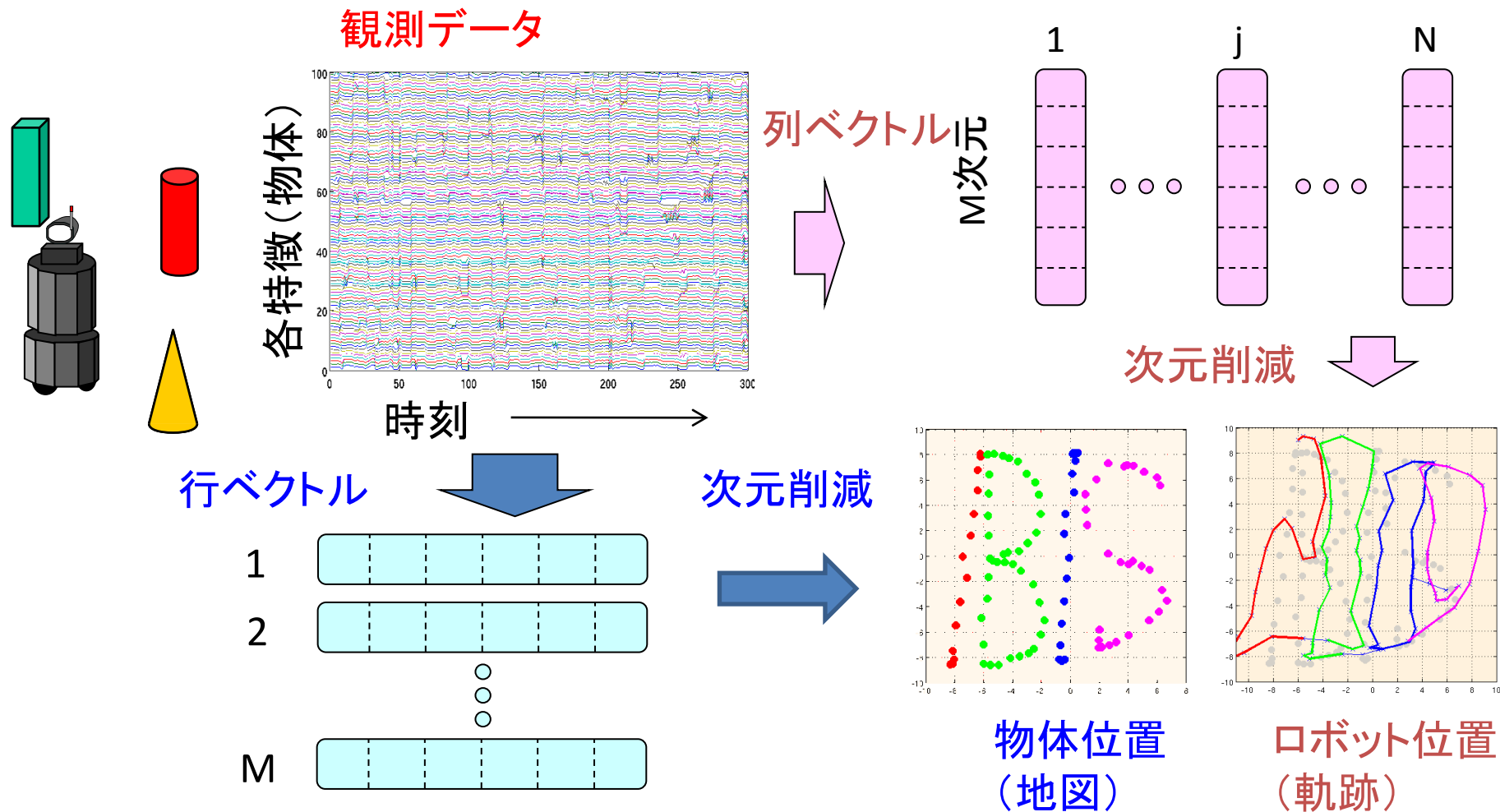
非線形
次元削減
(Isomap)



データ＝潜在構造＋潜在状態遷移

動機的例題2:

次元削減によるロボット位置推定・地図作成



データ = 潜在構造 (地図) + 潜在状態遷移 (軌跡)

共通問題: 動的システムの学習

- 状態空間モデルやDBNで記述可能
- モデル既知の場合: Bayesian filtering問題
- モデル(パラメータ、構造)が未知: モデルと状態の同時推定
- Latent Dynamics 的には?

状態方程式

$$\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t) + \mathbf{w}_t$$

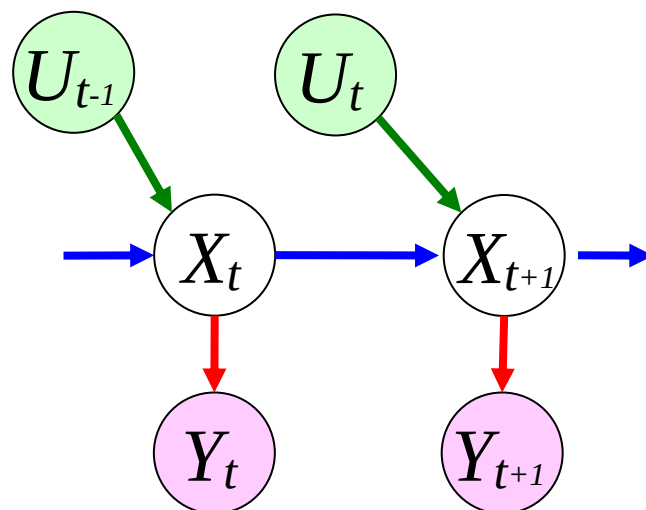
出力方程式

$$\mathbf{y}_t = g(\mathbf{x}_t) + \mathbf{v}_t$$

入力

隠れ
状態

観測



Latent Dynamics

(LD研究会HPより抜粋)

「第1のLD」: 今回のスコープ

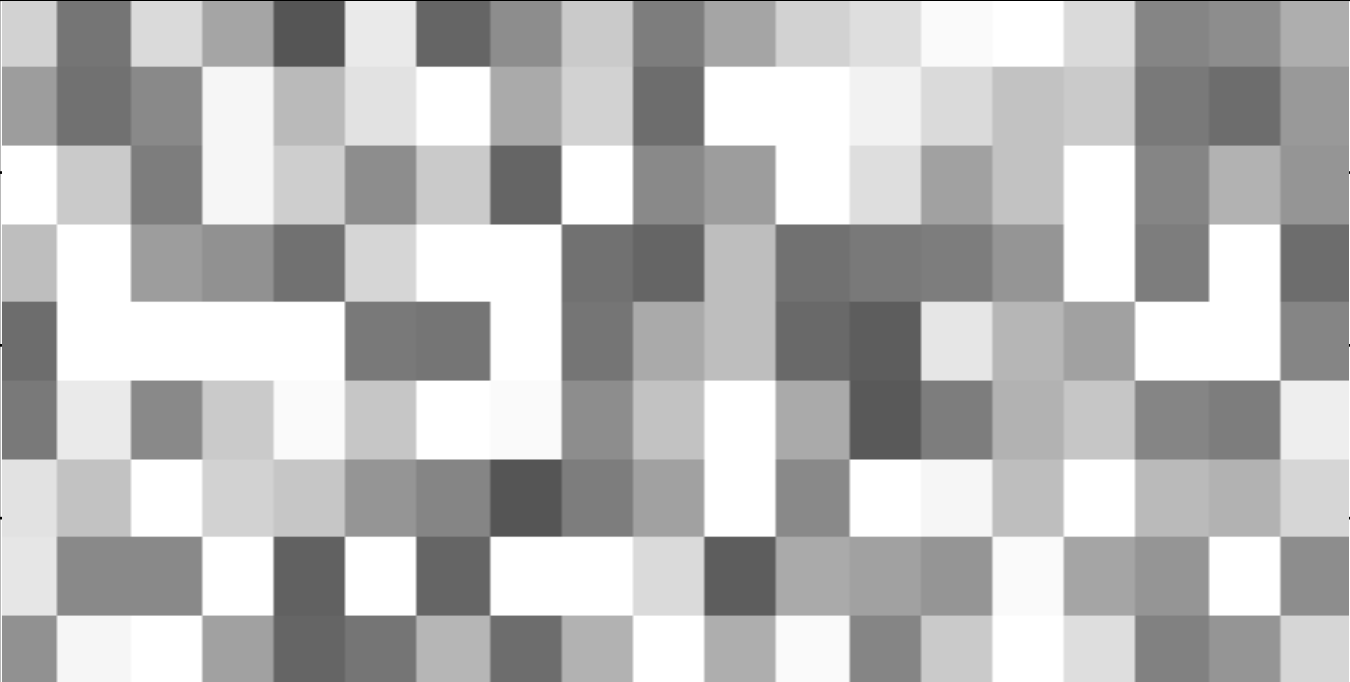
従来の隠れマルコフモデルや状態空間モデルでも状態変数のダイナミックスが扱われていたが、そこでは扱われていなかった構造的な変化(潜在変数を規定する、よりマクロな構造的な情報の変化)をも問題にしていきたい。

「第2のLD」: 今回は触れません

2. 動的システム学習研究の動向

～ 機械学習・システム同定両分野のサーベイ ～

両分野のスタンスの比較

	機械学習 (動的システム学習)	システム同定 (部分空間同定法)
共通問題	観測列 $Y_{1:T}$ と入力列 $U_{1:T}$ から、状態列 $X_{1:T}$ および、状態遷移モデル f , 観測モデル g を推定	
対象系		
興味		
モデル化		
手法		

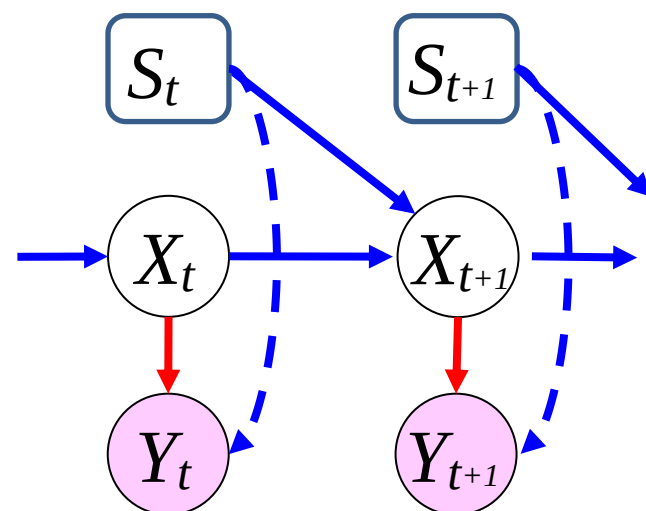
混合(スイッチング)線形モデル学習

- Switching Linear Dynamical Systems (SLDS)
(a.k.a. SSSM, SKF)
- 初期の研究: [Ghahramani 98], [Pavlovic 01]
 - LDSの混合モデルをEMアルゴリズムで推定
 - E-step: 状態 x_t および混合要素 s_t の推定。様々な近似法 (Viterbi-like, 変分近似, etc.)
 - 入力無し, 状態モデルのみスイッチング
 - 混合要素数などのパラメータはgiven
 - HMCデータなど

離散
状態

連続
状態

観測



混合（スイッチング）線形モデル学習（続）

- 近年、局所モデル数の自動決定等で進歩
- [Fox 2008] “Nonparametric Bayesian Learning of Switching Linear Dynamical Systems”, NIPS-2008
 - Dirichlet Process クラスターリングをSLDSに応用（HDP-HMM: 階層ディリクレ過程HMM）
 - 全パラメータをMCMCで推定
- [Chiappa 2008] “Using Bayesian Dynamical Systems for Motion Template Libraries”, NIPS-2008
 - 変分ベイズでパラメータ値を決定
 - シーケンス分類が主目的（≠ スwitchingモデル）

RBFN状態空間モデル

- [Roweis & Ghahramani 2000] “Learning Nonlinear Dynamical Systems using the EM Algorithm”
- 状態遷移・観測モデル f, g をRBFNで近似

$$\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t) + \mathbf{w}_t \quad \text{状態遷移モデル}$$

$$\approx \sum_{i=1}^I h_i \rho_i(\mathbf{x}_t, \mathbf{u}_t) + \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t + \mathbf{w}_t$$

Gaussian RBF

- EM で隠れ状態列とモデルを交互に更新
 - E-step : 拡張Kalman平滑化 (EKS)
- 入力 $\mathbf{u}_t$ ありだが、小規模問題のみ

ガウス過程状態空間モデル

GPDM [Wang 05] “Gaussian Process Dynamical Models”, NIPS-2005

- GPLVM[Lawrence 04] の拡張
- モデル f, g をGPで表現・学習
 - f, g は積分消去
 - 状態列 x とハイパーパラメータについて周辺尤度を最大化 (PCAで初期化)
 - 入力なし
- GPIL [Turner 10] : 疑似入力(pseudo inputs)によるスパース化, 状態列 x の分布考慮, EM

ガウス過程状態空間モデル(続)

- GP-Bayesfilter[Ko 09], GP-ADF[Deisenroth 09]
 - GPDM同様に、 f と g をGP回帰(Semi-parametric なバージョンもあり)で学習
 - 入力 $\mathbf{u}_t$ あり
 - ただし、状態列 $X=\{x_1, \dots, x_N\}$ が直接入手可能という前提
 - ⇒ 単なる2つの教師あり回帰問題に帰着
 - スパースGPも検討
 - 飛行船のシステム同定等、実用的問題に適用

カーネル状態空間モデル

Kernel Kalman Filter (KKF) [Ralaivola 05]

- カーネルによる非線形化
 - 特徴空間では線形モデルが成り立つと仮定
 - 特徴空間内の基底はKPCAで求める
- モデル学習はEM algorithmで
- Preimage (特徴空間 $\Rightarrow$ 観測空間)を別途学習
- 入力なし

Kernel 部分空間同定法 [Kawahara 06]

- Kernel CCA, 入力あり

多様体学習の確率モデル化

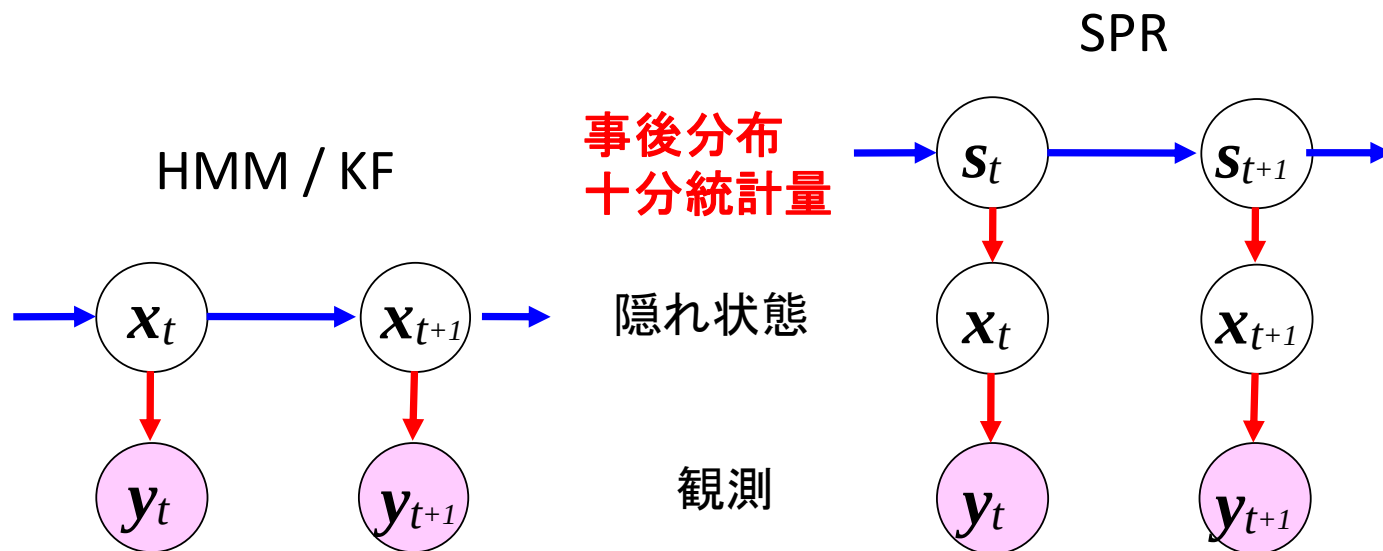
LELVM (Laplacian Eigenmap Latent Variable Model)
[Lu, et.al. NIPS07]

- Laplacian Eigenmap の拡張
 - 確率モデル化
 - out-of-sampleに対応
- 高次元観測データ(HMCなど)を次元削減して観測モデルを構築
- 状態遷移はランダムウォークでモデル化
 - 観測モデルのみの学習という点では素朴
- Gaussian mixture Sigma-point particle filter
- 入力なし

Langfordの方法

"Learning Nonlinear Dynamic Models", Langford ,
Salakhutdinov, Zhang, ICML-2009

- 確率的な動的システムを、「非確率的」な任意の回帰学習アルゴリズムで学習する方法
- 確率的状態 $\mathbf{x}_t$ を、決定論的な変数 s_t に変換
 - Sufficient Posterior Representation: SPR

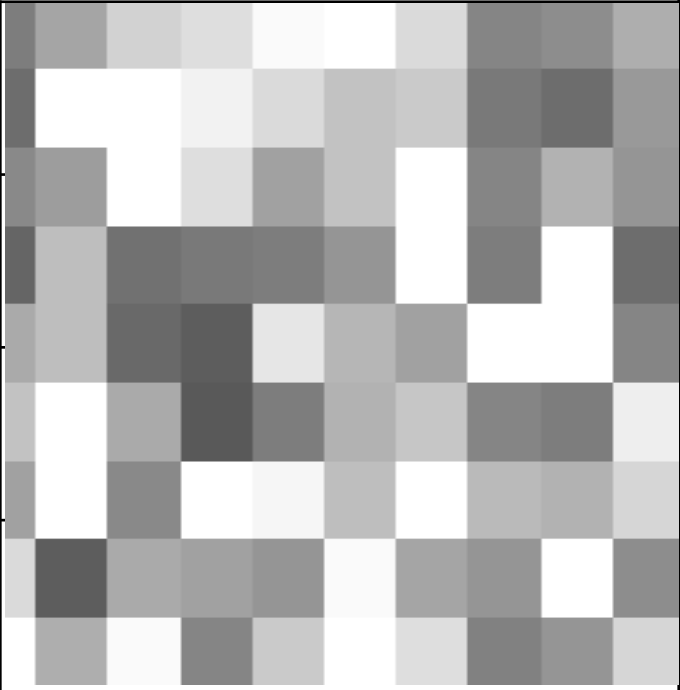


非定常DBN(その他の話題)

時間とともに構造が変化していくDBNの学習

- “Non-stationary dynamic Bayesian networks”,
Robinson and Hartemink, NIPS-2008
 - MCMCで推定
- “Time-Varying Dynamic Bayesian Networks”,
Song, et.al., NIPS-2009
 - 忘却率とL1正則化を導入したベクトル自己回帰
- Latent Dynamicsの趣旨に近い(?)

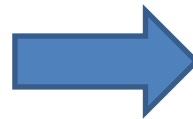
両分野のスタンスの比較

	機械学習 (動的システム学習)	システム同定 (部分空間同定法)
共通 問題	観測列 $Y_{1:T}$ と入力列 $U_{1:T}$ から、状態列 $X_{1:T}$ および、状態遷移モデル f , 観測モデル g を推定	
対象系	非線形一般、高次元出力、入力無し	
興味	観測の次元削減(可視化)、学習アルゴリズム	
モデル化	離散・連続変数区別無し(DBN)	
手法	EM algorithm による状態・モデル並行推定	

拡張状態ベクトル法

- 最も素朴な方法
- 未知モデルパラメータを状態変数ベクトルに含めてしまう
- あとは、EKF, PFなどでオンライン推定
- 当然、使えるケースは限られる

$$\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t, \boldsymbol{\theta}) + \mathbf{w}_t$$



$$\mathbf{y}_t = g(\mathbf{x}_t, \boldsymbol{\theta}) + \mathbf{v}_t$$

$$\mathbf{z}_t \equiv \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\theta} \end{bmatrix}$$

$$\mathbf{z}_{t+1} = f^+(\mathbf{z}_t, \mathbf{u}_t) + \mathbf{w}_t^+$$

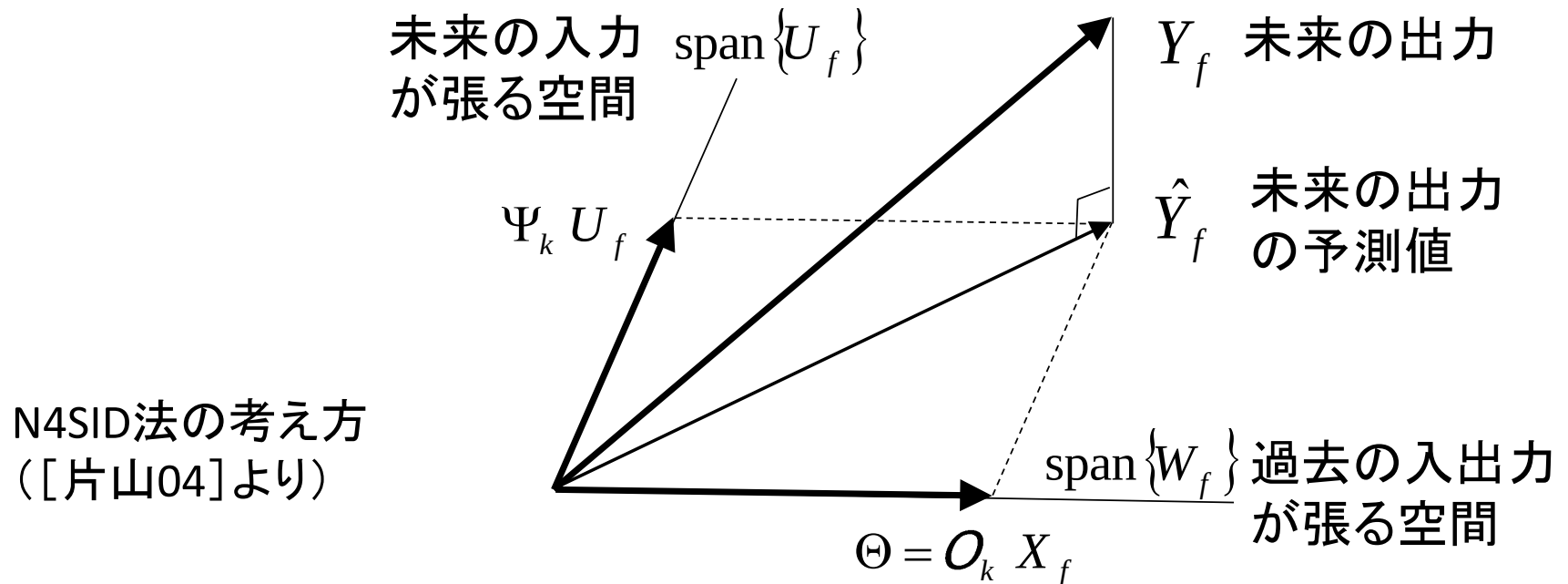
$$\mathbf{y}_t = g(\mathbf{z}_t) + \mathbf{v}_t$$

予測誤差法 (PEM)

- (iterative) Prediction-Error Minimization Method
 - 部分空間同定法以前の主流
 - 予測誤差を最小化するパラメータを求める (≡ 最尤法)
 - Levenberg-Marquardt 法など
 - 状態列 $X = [x_1, \dots, x_N]$ は点推定(?)
 - 実は、EM algorithmは近年まであまり知られてなかった!? [Ljung 08][Schön 06]
 - 部分空間同定法の「後処理」としても利用

部分空間同定法

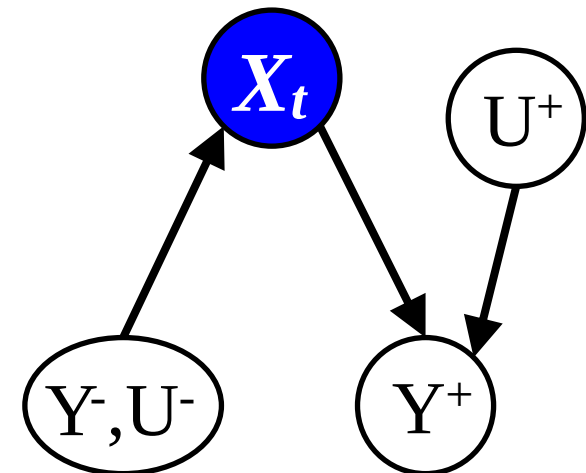
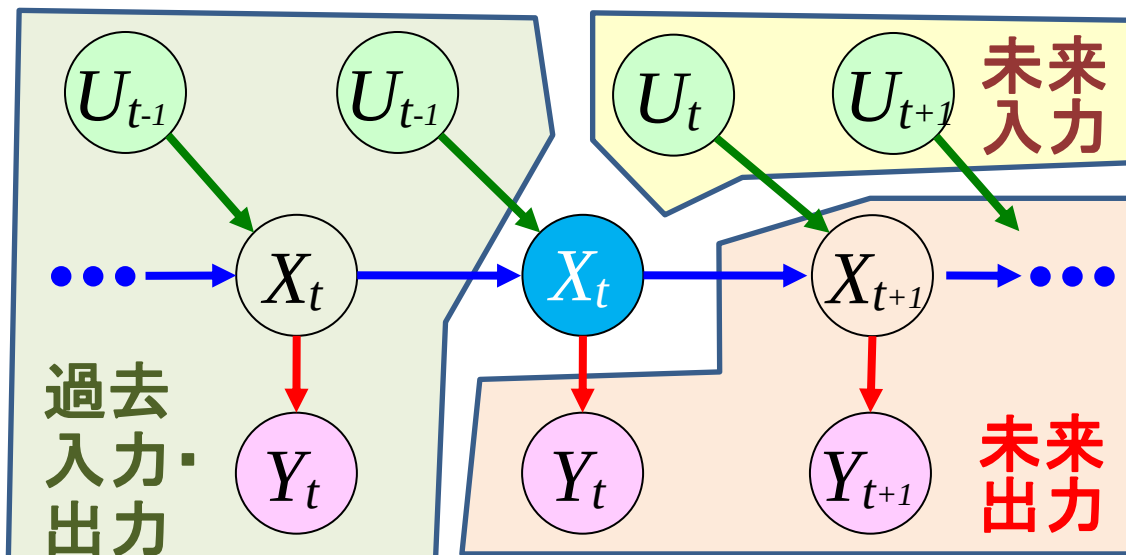
- 入出力データが張る部分空間上における幾何学的演算により, データを生成する状態空間モデル, 状態ベクトルを推定する方法
- MOESP, N4SID, CVA(CCA)など(統一理論あり)
- 詳細は[片山04]



N4SID法の考え方 ([片山04]より)

部分空間同定法(続)

- 正準相関分析(CCA)による同定[Akaike76]
[Larimore83][片山04]
- 「状態とは、未来(過去)を予測するために必要な過去(未来)の情報を縮約したもの」



部分空間同定法の実線形化

- 非線形性のタイプを細かく分類し対処
 - Hammerstein型[Goethals 05], Wiener型, など
 - 固定基底群の利用が基本
- 区分線形(PWL)システム [Verdult04]
 - SLDSに似た考えだが、 s_t に関して教師あり
 - 切替点での整合性を考慮して各局所座標を変換
- Kernel特徴空間での部分空間同定法 [Kawahara06]
- 非線形性とは別に、フィードバックありシステムへの拡張も盛んに議論

システム同定分野から見た機械学習(余談)

- Lennart Ljung, “Perspectives on System Identification”, IFAC Congress, (keynote), 2008
 - システム同定の大御所
- システム同定との交流が期待される分野として、
 - Statistical Learning Theory: kernel の利用
 - Machine Learning : DT, GP, etc.。SLTと接近
 - Manifold Learning : LLE, グラフ正則化[Ohlsson08]
 - Data Mining : データの大規模化を挙げている

両分野のスタンスの比較

	機械学習 (動的システム学習)	システム同定 (部分空間同定法)
共通 問題	観測列 $Y_{1:T}$ と入力列 $U_{1:T}$ から、状態列 $X_{1:T}$ および、状態遷移モデル f , 観測モデル g を推定	
対象系	非線形一般、高次元出力、入力無し	線形が基本(非線形拡張あり)、入力あり
興味	観測の次元削減(可視化)、学習アルゴリズム	ダイナミクス、制御、系自体の性質
モデル化	離散・連続変数区別無し(DBN)	基本的に連続変数(Hybrid系もあり)
手法	EM algorithm による状態・モデル同時推定	SVD, QR等で状態、モデルの一方を先に推定

3. 局所線形モデルの整列による 非線形システムの学習法(*)

Learning Non-linear Dynamical Systems
by Alignment of Local Linear Models(†)

*上甲, 河原, 矢入, 人工知能学会全国大会 2010, 長崎

† Joko, Kawahara and Yairi, ICPR-2010, Istanbul, (To appear)

研究の目的

- 動的システム学習法

- 制御(システム同定)分野では、代表的な学習法の一つに部分空間同定法がある

- ⇒ただし、主に線形システム

- 一方、機械学習分野では、混合モデルや多様体学習などによって非線形な動的システムに関する研究が発達してきた

- ⇒ただし、主に反復法、入力なし



両者の融合(良いところ取り)

発表の構成

1. 研究の背景・目的
2. アプローチ
3. シミュレーションと結果
4. 結論と今後の課題

準備 部分空間同定法とは

- 線形モデルの仮定
$$\begin{cases} x(t+1) = Ax(t) + w(t) \\ y(t) = Cx(t) + v(t) \end{cases}$$

観測ベクトル 直接観測されない状態ベクトル

- 過去と未来の観測列で張られる部分空間上における幾何学的演算により状態空間を構築する手法

$$\begin{cases} x^f(t) \equiv S^{1/2} V^T \Sigma_{pp}^{-1/2} \check{y}_k(t) \\ x^b(t) \equiv S^{1/2} U^T \Sigma_{ff}^{-1/2} y_k(t) \end{cases} \quad \left(\begin{array}{l} \check{y}_k(t) \equiv [y(t-1)^T, y(t-2)^T, \dots, y(t-k)^T]^T \\ y_k(t) \equiv [y(t)^T, y(t+1)^T, \dots, y(t+k-1)^T]^T \\ \Sigma_{ff} \equiv E\{(y_k(t) - \mu)(y_k(t) - \mu)^T\} \\ \Sigma_{ff}^{-1/2} \Sigma_{fp} \Sigma_{pp}^{-T/2} = U S V^T \end{array} \right)$$

- 部分空間同定法を用いた学習アルゴリズムの枠組み
 - 部分空間同定法によって状態ベクトル $x(t)$ を得る
 - 最小自乗法を適用し、システム行列等 (A, C) を得る

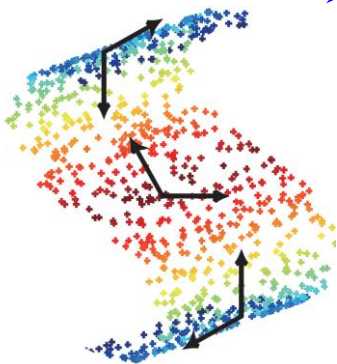
アプローチ (2/6) 学習アルゴリズム

- 提案する学習アルゴリズムの枠組み

1. 局所線形アライメントアプローチ (提案手法) によって状態ベクトルを得る
2. 各部分空間において最小自乗法を適用し、システム行列を求める

$$\begin{cases} x(t+1) = Ax(t) + w(t) \\ y(t) = Cx(t) + v(t) \end{cases} \quad \Rightarrow \quad \begin{cases} p(x(t+1) | x(t), s) = N(\mu_s^t + A_s x(t), \Sigma_s^t) \\ p(y(t) | x(t), s) = N(\mu_s^o + C_s x(t), \Sigma_s^o) \end{cases}$$

(拡張手順)



1. CCAに基づく部分空間同定法を確率的に解釈する
2. 多様体学習のひとつであるアライメントベースの混合確率モデル学習アプローチを適用し、非線形に拡張する

拡張手順1: 部分空間同定法の確率的解釈

- 時刻 t における未来の出力、過去の出力は $x^f(t)$ 、 $x^b(t)$ を用いて、以下のように表される

$$\begin{aligned} y_k(t) &= O_k^f x^f(t) + v_k \\ \check{y}_k(t) &= O_k^b x^b(t) + \check{v}_k \end{aligned} \quad (*) \quad \left[\begin{array}{l} \text{ただし } O_k^f, O_k^b \text{ は拡大} \\ \text{可観測行列} \end{array} \right]$$

- 逆に上式から、 $x^f(t)$ 、 $x^b(t)$ はそれぞれ過去と未来の観測列の十分統計であることがわかる。つまり、

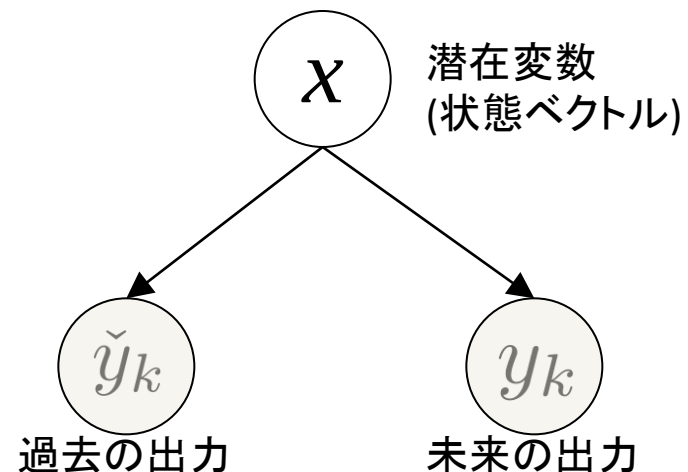
$$\begin{cases} p(x^f) = p(x | y_k) \\ p(x^b) = p(x | \check{y}_k) \end{cases}$$

拡張手順1: 部分空間同定法の確率的解釈

- (＊)の立式は正準相関分析の潜在変数モデルと一致する[Bach and Jordan 06]。よって、部分空間同定法の潜在変数モデルは以下のように書ける

$$x \sim N(0, I_d), y_k | x \sim N(F_1 x, \Psi_1), \check{y}_k | x \sim N(F_2 x, \Psi_2)$$

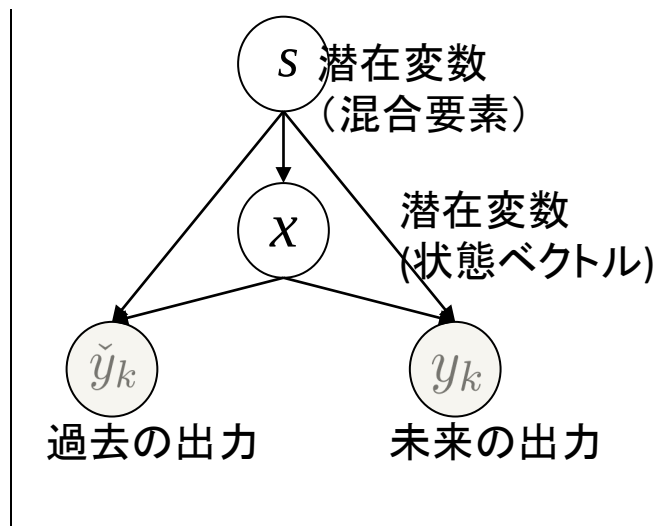
正準相関分析は、2組の集合の間に存在する線形関係を明らかにする多変量解析の手法であり、[Bach and Jordan 06]によって確率的な解釈がなされた。



拡張手順2: 多様体学習法の適用

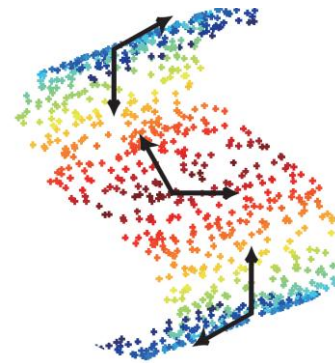
- 確率的に解釈された部分空間同定法に対して、**局所線形モデルの整列**を考え、大域的な状態空間を構築する(手順は[Verbeek 06]に従う)

グラフィカルモデル

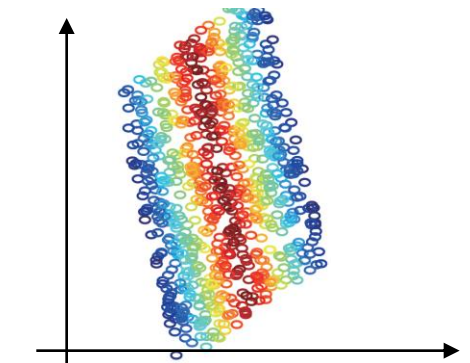
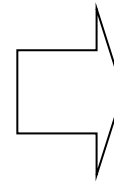


イメージ図

(混合要素数 $C=3$, 低次元空間の次元数 $d=2$)



観測空間



大域的な状態空間

ペナルティ付きの
対数尤度目的関数

$$L' = \sum_{n=1}^{n=N} [\log p(\check{y}_k(n), y_k(n)) - D(q_n(x) \| p(x | \check{y}_k(n), y_k(n)))]$$

ペナルティ項: 個々の混合分布の整合性をとる制約

アプローチ (6/6)

学習アルゴリズムまとめ

- 提案する学習アルゴリズムの枠組み
 1. 局所線形アライメントアプローチ(提案手法)によって状態ベクトルを得る
 2. 各部分空間において最小自乗法を適用し、システム行列を求める

$$\begin{cases} x(t+1) = Ax(t) + w(t) \\ y(t) = Cx(t) + v(t) \end{cases} \Rightarrow \begin{cases} p(x(t+1) | x(t), s) = N(\mu_s^t + A_s x(t), \Sigma_s^t) \\ p(y(t) | x(t), s) = N(\mu_s^o + C_s x(t), \Sigma_s^o) \end{cases}$$

発表の構成

1. 研究の背景・目的
2. アプローチ
3. シミュレーションと結果
4. 結論と今後の課題

実験1 従来手法との予測性能の比較

- 対象とするシステム

- ローレンツアトラクタ

$$\begin{cases} y_1(t) = x_1(t) + \varepsilon_1(t) \\ y_2(t) = x_2(t) + \varepsilon_2(t) \\ y_3(t) = x_3(t) + \varepsilon_3(t) \end{cases} \begin{cases} \dot{x}_1(t) = -ax_1(t) + ax_2(t) \\ \dot{x}_2(t) = -x_1(t)x_3(t) + rx_1(t) - x_2(t) \\ \dot{x}_3(t) = x_1(t)x_2(t) - bx_3(t) \end{cases}$$

- 比較指標

- 予測対数尤度

- 比較手法（線形手法のみ）

- カルマンフィルタベーストアルゴリズム (KFS)

- 直交分解に基づく部分空間同定法 (ORT)

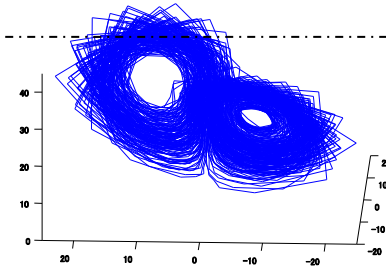


表: 予測対数尤度の比較

KFS	ORT	MIX (提案手法)
-9.73	-9.28	-6.69

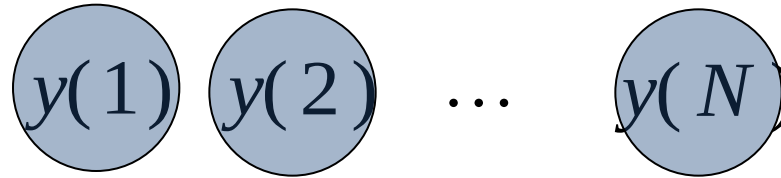
非線形手法や
SLDSとの比較
は未実施

線形の部分空間同定法よりは優れた予測精度を示している

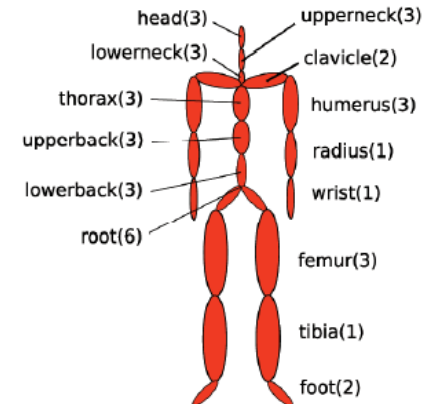
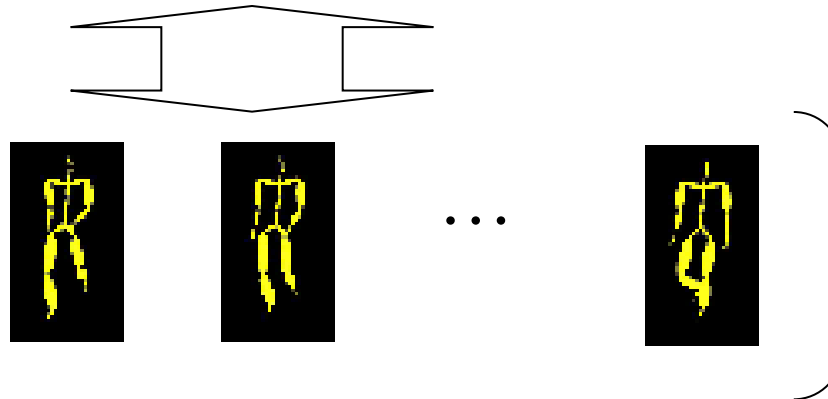
実験2 モーションキャプチャへの適用

- 高次元複雑システムの代表例であるHMCデータ(CMU mocap)に本手法を適用する
- データセット $\{y(t), t = 0, 1, 2, \dots, N\}$ ($y \in R^{50}$)
- 設定: 混合要素数 $C=4$, 低次元 $d=5$

観測
(関節データ)



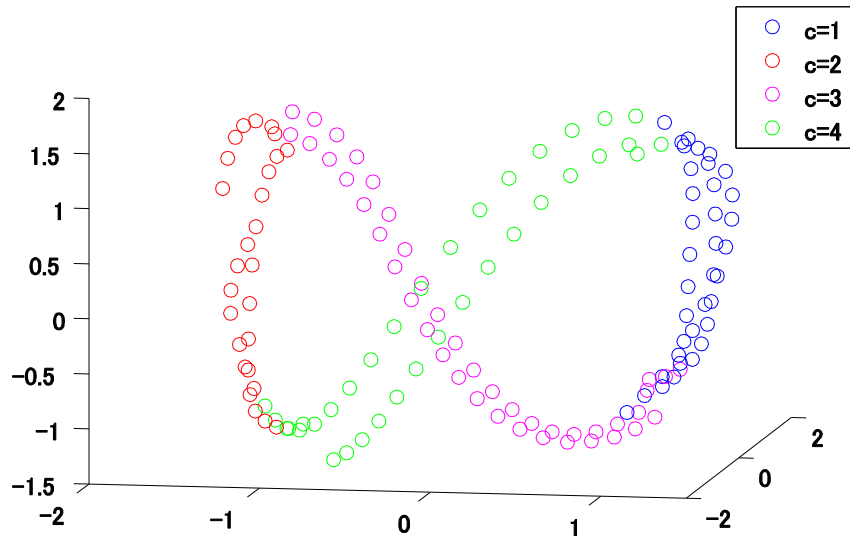
再現データ



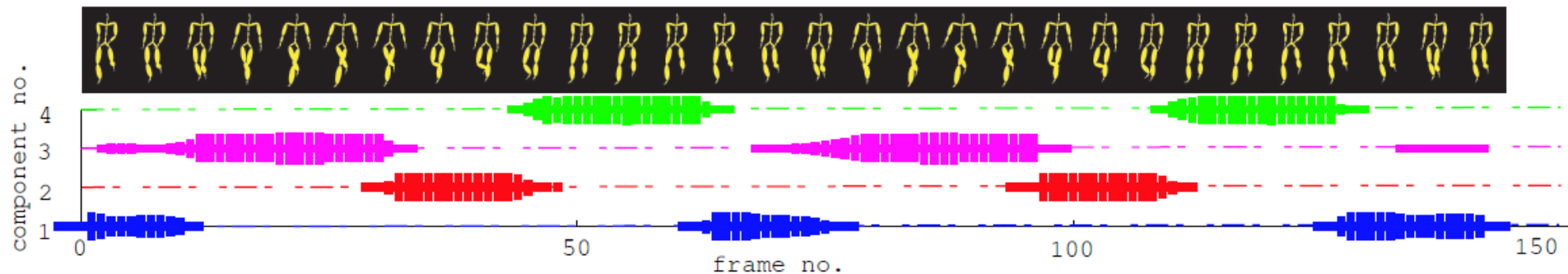
シミュレーションと結果 (3/3)

実験2 モーションキャプチャへの適用

• 低次元状態空間での軌跡と各混合要素



- GPDM[Wang05] の結果と類似
- 局所線形なので計算量が少
- 整列により、局所モデルのつながりが滑らか
- 動作の分節化
- 異常姿勢の検知



4. おわりに

まとめ

- 観測データからの動的システム学習に関する機械学習分野、および、システム同定分野のトレンドを概観した
 - 着眼点、興味、文化の違い
- 両者の「得意技」を組み合わせた提案手法「局所線形モデルの整列による非線形システムの学習法」を紹介した
 - ただし、入力項 u_t なし
 - 従来手法（非線形学習法、SLDS、等）との比較は未実施

謝辞

- 動的システム学習に関する調査は、井手剛氏 (IBM) との共著解説「"機械学習技術の最近の発展とシステムモデリングへの応用", 計測と制御, 2010年7月号」に基づきます
- 部分空間同定法と機械学習との関連性については、主に、河原吉伸氏 (大阪大) から得た情報に基づきます
- 「局所線形モデルの整列による非線形システムの学習法」は、上甲昌郎氏 (東大院)、河原氏との共同研究です
- 関係者に厚くお礼申し上げます