

確率的 Slow Feature Analysis における 観測ノイズの影響

関口 智樹^{†1} 大森 敏明^{†1,†2} 岡田 真人^{†1,†2}

Slow Feature Analysis(SFA) は時系列データからゆっくりと変化する情報を抽出する数理モデルであり、神経システムのモデルなどに応用されている。近年、SFA の確率モデルが提案されているが、先行研究における SFA の確率モデルでは、データに加わる観測ノイズに関する近似を行っており、その影響についての定量的な議論が行われていなかった。本論文で我々は、パラメータ推定の精度や推定される slow feature のダイナミクスの振る舞いを調べることで、SFA の確率モデルにおける観測ノイズの影響を明らかにし、最もゆっくりと変化する成分が観測ノイズの影響を強く受けることを示す。

The Effect of Observation Noise in Probabilistic Slow Feature Analysis

TOMOKI SEKIGUCHI,^{†1} TOSHIAKI OMORI^{†1,†2}
and MASATO OKADA^{†1,†2}

The slow feature analysis (SFA) is a mathematical model that extracts slowly varying features from time series data. For example, the SFA has been applied for neural systems. Recently, a probabilistic version of SFA was proposed. This probabilistic SFA includes approximation on observation noise. However, quantitative evaluation on the effect of the observation noise in the probabilistic SFA has not been investigated in the previous study, and thus it remains unclear how the observation noise affects the performance in the probabilistic SFA. In this paper, we investigate the effect of observation noise in the probabilistic SFA by evaluating the accuracy of estimated parameters including slow feature dynamics. We show that the most slowly varying feature suffers from strong effect of the observation noise.

1. はじめに

我々は自分が見ている視覚情報から物体の位置、形、明るさなど様々な情報を得ることができる。時々刻々と変化する視覚情報の中で、ゆっくりと変化する情報が物体や空間を認識するために重要な役割を担っていることが近年注目されている¹⁾⁻³⁾。Slow Feature Analysis(SFA) は時系列データから最もゆっくりと変化する成分 (slow feature) を抽出する数理モデルであり、以上のような視覚情報からの情報抽出を行う神経科学の数理モデルとして、2002 年に Wiskott らによって提案された²⁾。SFA は視覚野における複雑型細胞 (complex cell) や海馬における空間認知のための場所細胞 (place cell) などの神経システムの生成モデルに応用されてきた⁴⁾⁻⁶⁾。

近年、Turner らは SFA の確率的な定式化を行った³⁾。確率的な定式化を行うことで、通常の SFA に比べていくつかの利点が得られる。例えば、ノイズ除去に関する確率的画像処理の枠組みを利用することが可能になる⁷⁾。また、他の確率モデルとの比較や、欠損データの取り扱いが可能になることなども挙げられる。確率的な定式化は他のデータ解析のモデルでもなされており⁸⁾⁻¹⁰⁾、確率的な主成分分析はその代表例といえる¹¹⁾。

Turner らは SFA を確率的に解釈することに成功したが³⁾、定式化の際にデータに加わる観測ノイズに関する近似を行っており、slow feature 抽出が実際に可能かどうかについて、機械学習的な観点からの定量的な議論は与えられていなかった。実データには少なからず観測ノイズが加わっていることが予想され、そのような状況下での情報抽出性能を議論することは重要な課題である。

本論文では、従来の研究において、主に神経科学分野での応用が考えられてきた SFA を機械学習的な観点から議論する。我々は、パラメータ推定の精度や推定される slow feature のダイナミクスの振る舞いを調べることで、観測ノイズが強い状況においては slow feature を正確に推定することができないこと、そして、潜在変数が複数ある場合、最もゆっくりと変化する成分が観測ノイズの影響を強く受けることを示す。すなわち、本研究で行った機械学習的な観点からの議論によって、確率的 SFA による情報抽出に対する観測ノイズの重要性を明らかにし、観測ノイズの効果を取り入れた厳密な取り扱いが必要であることを示す。

^{†1} 東京大学大学院 新領域創成科学研究科
Graduate School of Frontier Sciences, The University of Tokyo

^{†2} 理化学研究所 脳科学総合研究センター
RIKEN Brain Science Institute

2. 理論

本章では、決定論的な SFA のモデルと 2007 年に Turner らによって提案された確率的 SFA のモデルを紹介し、最後に確率的 SFA に含まれている観測ノイズに関する近似から、モデルの問題点について議論する。

2.1 Slow Feature Analysis

Slow Feature Analysis(SFA) は、複数の時系列データの中から最もゆっくりと変化する成分 (slow feature) を抽出する数理モデルであり、2002 年に Wiskott らによって提案された²⁾。この数理モデルはデータに応じた特徴抽出が可能となっている教師無し学習である。ここでは SFA の決定論的なアルゴリズムについて述べる。

入力として、連続な時系列データを $x(t)$ とすると、SFA における出力 $y_j(t)$ は、スカラー関数 $g_j(x)$ による変換 $y_j(t) = g_j(x(t))$ を行うことで得られる。この変換 $g_j(x)$ は次に示す量 (Δ -value) を最小化するように決められる。

$$\Delta(y_j) := \langle \dot{y}_j^2 \rangle_t \quad (1)$$

つまり、出力 y の時刻に関する微分係数が最小になるような変換 $g_j(x)$ を行うことになる。 Δ -value が最小になる出力関数 $y_j(t)$ を slow feature と呼ぶ。さらに、これに加えて次のような拘束条件を加える。

$$\langle y_j \rangle_t = 0 \quad (\text{zero mean}) \quad (2)$$

$$\langle y_j^2 \rangle_t = 1 \quad (\text{unit variance}) \quad (3)$$

$$\langle y_i y_j \rangle_t = 0 \quad (\text{decorrelation}) \quad (4)$$

式 (2) は計算を簡単にするため、式 (3) は $g_j(x) = 0$ という解を避けるため、式 (4) は出力の重複を避けるために設定されている。

2.2 確率的 Slow Feature Analysis

2007 年には、Turner らが確率的な枠組みによって SFA を提案している。観測される M 次元の値 (観測変数) を x_t 、SFA によって抽出される N 次元の成分 (潜在変数) を y_t というベクトルでそれぞれ表す。確率的な枠組みにおいては、与えられた観測変数 x_t から、決定論的な SFA の性質を取り入れた確率分布に基づいて潜在変数 y_t を推定する。図 1 に示しているように、決定論的な SFA における時系列データが観測変数 x_t に、出力関数が潜在変数 y_t にそれぞれ対応している。また、決定論的な SFA では連続時間における時系列データを扱っていたが、確率的 SFA では離散的な時間における時系列データとしてそれぞ

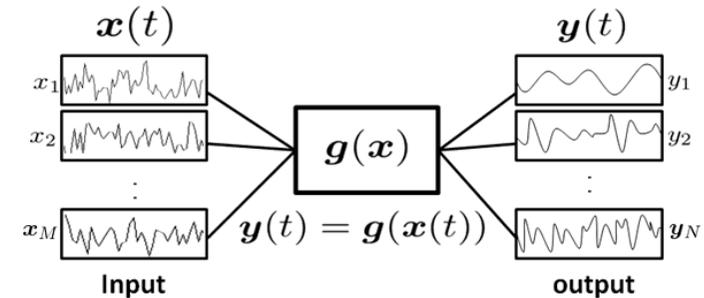


図 1 Slow Feature Analysis の概念図。複数の時系列データを入力 x_t とし、その中に含まれるゆっくりと変化する成分を出力 y_t として得る。

Fig. 1 Schematic diagram of slow feature analysis. SFA extracts slow varying features y_t from time series input data x_t .

れの変数を扱う。

slow feature を含む潜在変数 y_t を特徴付ける式はシステムモデルといい、次のように表される³⁾。

$$y_t = \lambda y_{t-1} + \eta_t \quad (5)$$

潜在変数 y_t は 1 時刻前の潜在変数 y_{t-1} に依存する形で書かれている。 λ はその依存度を定めるパラメータであり、各 $y_{n,t}$ に対する値 λ_n を並べた $N \times N$ の対角行列となっている。ここで加わるノイズ η_t はシステムノイズといい、平均 0、共分散行列 Σ に従う正規白色雑音である。共分散行列 Σ は対角要素 $\sigma_{1:N}^2$ をもつ $N \times N$ 行列であり、各 $y_{n,t}$ には異なる分散のノイズが加わることになる。このように、潜在変数 y_t は 2 つのパラメータ λ 、 Σ によって特徴付けられている。 λ_n は 0 から 1 の範囲で、大きければ前時刻の影響を受けやすくなるので、 $y_{n,t}$ のダイナミクスはゆっくりとしたものになる。逆に λ_n が小さければノイズの影響が大きくなり、ダイナミクスは速い変化になる。したがって、確率的 SFA では λ_n の中で最も大きい値をもつ潜在変数 y_t が slow feature に対応することになる。

一方、観測変数 x_t を特徴付ける式は観測モデルといい、次のように表される。

$$x_t = W^{-1} y_t + \varepsilon_t \quad (6)$$

観測変数 x_t は潜在変数 y_t に $M \times N$ 行列による変換 W^{-1} を施し、ノイズ ε_t を加えることで生成される。このノイズは観測ノイズと呼ばれ、平均 0、共分散行列 $\sigma_x^2 I$ に従う正規白色雑音である (I は単位行列)。ここまで説明したように、確率的 SFA では式 (5) (6)

に従って観測変数 x_t と潜在変数 y_t を記述している．式 (5) (6) をまとめて状態空間モデルと呼ぶ．

確率的 SFA の状態空間モデルは次のようなガウス分布でも書くことができる．

$$p(x_t|y_t, \mathbf{W}, \sigma_x^2) = \mathcal{N}(x_t|\mathbf{W}^{-1}y_t, \sigma_x^2 \mathbf{I}) \quad (7)$$

$$p(y_t|y_{t-1}, \lambda, \Sigma) = \mathcal{N}(y_t|\lambda y_{t-1}, \Sigma) \quad (8)$$

$$p(y_1|P_0) = \mathcal{N}(y_1|0, P_0) \quad (9)$$

式 (9) は時刻 $t = 1$ における潜在変数 y_t の分布である．また， P_0 はシステムノイズの初期条件を表す共分散行列である．これらの分布をもとに，潜在変数 y_t の推定を進める．

2.3 最尤推定

ここまで述べてきた確率的 SFA のモデルをもとに，最尤法によって潜在変数 y_t の推定を行う．最尤法では尤度関数が最大となる点におけるパラメータの値を尤（もっと）もらしい値として考え推定を行う．確率的 SFA における尤度関数は次のように表される．

$$p(\mathbf{x}_{1:T}|\theta) = \int d\mathbf{y}_{1:T} \left(\prod_{t=1}^T p(x_t|y_t, \mathbf{W}, \sigma_x^2) \right) \times \left(\prod_{n=1}^N p(y_{n,1}|\sigma_{n,1}^2) \prod_{t=2}^T p(y_{n,t}|y_{n,t-1}, \lambda_n, \sigma_n^2) \right) \quad (10)$$

本論文で注目する Turner らのモデルでは，この尤度関数を計算する際に $\sigma_x^2 \rightarrow 0$ とする近似を行い，決定論的な写像に戻す．この近似を行うことで，観測モデルのガウス分布は以下の通りデルタ関数に置きかえられることになり，尤度関数を求めるための積分計算が容易に行えるようになる．

$$p(\mathbf{x}_{1:T}|\theta) \approx \int d\mathbf{y}_{1:T} \left(\prod_{t=1}^T \delta(x_t - \mathbf{W}^{-1}y_t) \right) \times \frac{1}{Z} \exp \left(- \sum_{n=1}^N \left(\frac{1}{2\sigma_{n,1}^2} y_{n,1}^2 + \frac{1}{2\sigma_n^2} \sum_{t=2}^T (y_{n,t} - \lambda_n y_{n,t-1})^2 \right) \right) \quad (11)$$

この近似によって計算された尤度関数の対数をとると次式のようになる．

$$L(\theta) = \log p(\mathbf{x}_{1:T}|\theta) = c + T \log |\det \mathbf{W}| - \sum_{n=1}^N \left(\frac{1}{2\sigma_{n,1}^2} (\mathbf{w}_n^T \mathbf{x}_1)^2 + \frac{1}{2\sigma_n^2} \sum_{t=2}^T (\mathbf{w}_n^T \mathbf{x}_t - \lambda_n \mathbf{w}_n^T \mathbf{x}_{t-1})^2 \right) \approx c + T \log |\det \mathbf{W}| - \frac{T}{2} \text{tr} (\mathbf{W} \mathbf{B} \mathbf{W}^T \Lambda^{(2)} + \mathbf{W} \mathbf{A} \mathbf{W}^T \Lambda^{(1)}) \quad (12)$$

ただし，2 式から 3 式の導出の際には，データ点が十分多いとして計算している．ここで，

$$\mathbf{A}_{mn} = \langle \Delta x_m \Delta x_n \rangle = \frac{1}{T-1} \sum_{t=1}^{T-1} (x_{m,t+1} - x_{m,t})(x_{n,t+1} - x_{n,t}) \quad (13)$$

$$\mathbf{B}_{mn} = \langle x_m x_n \rangle = \frac{1}{T} \sum_{t=1}^T x_{m,t} x_{n,t} \quad (14)$$

$$\Lambda^{(1)} = \delta_{mn} \frac{\lambda_n}{\sigma_n^2} \quad (15)$$

$$\Lambda^{(2)} = \delta_{mn} \frac{(1 - \lambda_n)^2}{\sigma_n^2} \quad (16)$$

とされている．最尤法ではこの対数尤度 $L(\theta)$ を最大化するパラメータの値を推定値とする．最大値となる点を探るために，対数尤度 $L(\theta)$ を各パラメータで偏微分していく．観測ノイズ σ_x^2 は 0 と近似しているため，推定するパラメータは $\theta = \{\mathbf{W}, \lambda, \Sigma\}$ の 3 つである．偏微分の結果は次のようになる．

$$\frac{\partial L}{\partial \mathbf{W}} = T \left((\mathbf{W}^{-1})^T - \Lambda^{(1)} \mathbf{W} \mathbf{A} - \Lambda^{(2)} \mathbf{W} \mathbf{B} \right) \quad (17)$$

$$\frac{\partial L}{\partial \lambda} = -\frac{T}{2} \left(\mathbf{W} \mathbf{A} \mathbf{W}^T \Sigma^{-1} - 2 \mathbf{W} \mathbf{B} \mathbf{W}^T \Sigma^{-1} + (\mathbf{W} \mathbf{B} \mathbf{W}^T \Sigma^{-1} + \Sigma^{-1} \mathbf{W} \mathbf{B} \mathbf{W}^T) \lambda \right) \quad (18)$$

$$\frac{\partial L}{\partial \Sigma} = -\frac{1}{2} \left((T-1)(\Sigma^{-1})^T - T(\Sigma^{-1})^T \mathbf{W} \mathbf{B} \mathbf{W}^T (\mathbf{I} - \lambda)^2 (\Sigma^{-1})^T - T(\Sigma^{-1})^T \mathbf{W} \mathbf{A} \mathbf{W}^T \lambda (\Sigma^{-1})^T \right) \quad (19)$$

これらをそれぞれ 0 とおくことで推定値の計算を行う．

3. 数値計算による評価

本章では，前章で述べた確率的 SFA のモデルにおける観測ノイズの影響を定量的に議

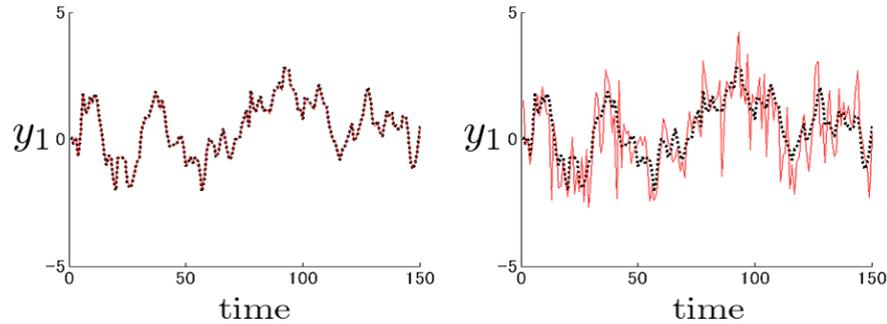


図2 確率的 SFA による潜在変数 y_1 の推定結果．点線が真の値，実線が推定値である．左側が観測ノイズ $\sigma_x^2 = 0$ ，右側が観測ノイズ $\sigma_x^2 = 1.0$ の場合の結果をそれぞれ表している．

Fig.2 Estimated slow feature y_1 using probabilistic SFA. Dotted line shows true value, whereas solid line shows estimated value. (Left) the case of observation noise $\sigma_x^2 = 0$. (Right) the case of observation noise $\sigma_x^2 = 1.0$.

論する．観測変数 x_t と潜在変数 y_t の次元はともに3次元 ($N = M = 3$)，データの個数は $T = 10000$ とした．また，システムモデルのパラメータを $\lambda_1 = 0.9, \lambda_2 = 0.5, \lambda_3 = 0.00001, \sigma_1^2 = 0.19, \sigma_2^2 = 0.75, \sigma_3^2 = 1.0$ とし，観測モデルにおける行列 W には正規直交行列を用いている．この場合， $\lambda_1 = 0.9$ に対応する潜在変数 y_1 が最もゆっくりと変化する slow feature に対応する．一方で， $\lambda_3 = 0.00001$ に対応する潜在変数 y_3 はノイズのように乱雑なダイナミクスをもつ．前章で得られたパラメータ推定の式 (17)–(19) に従って3つのパラメータ $\theta = \{W, \lambda, \Sigma\}$ を最急上昇法によって決定し，得られたパラメータの値から潜在変数 y_t をカルマンフィルタによって推定する．

3.1 潜在変数の推定結果

前章で述べた観測ノイズに関する近似について，まず議論すべきは，観測ノイズが加わったデータに対しても slow feature の推定を正確に行うことができるのか，ということである．この検証を行うために，観測モデルに従って，観測ノイズを加えたデータを生成し，そのデータから潜在変数 y_t を推定する計算を行った．図2は slow feature に対応する潜在変数 y_1 の推定結果を示したものである．観測ノイズが加わっていない場合 (左図) は推定値が真の値に一致する正確な結果を与えているが，観測ノイズが加わった場合 (右図) は真の値に比べて乱雑なダイナミクスが推定されていることが分かる．潜在変数の他の成分 y_2, y_3 についても同様の傾向が見られた (図省略)．観測ノイズに関するデルタ関数近似に

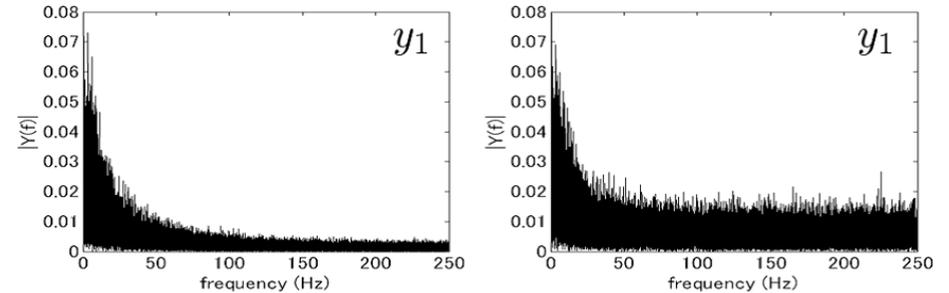


図3 確率的 SFA による潜在変数 y_1 に対するフーリエ変換の結果．左側が観測ノイズ $\sigma_x^2 = 0$ ，右側が観測ノイズ $\sigma_x^2 = 1.0$ の場合の結果をそれぞれ表している．

Fig.3 The result of Fourier transform for estimated slow feature y_1 using probabilistic SFA. (Left) the case of observation noise $\sigma_x^2 = 0$. (Right) the case of observation noise $\sigma_x^2 = 1.0$.

よって，観測ノイズのパラメータ σ_x^2 を推定する枠組みが失われたために，このような影響が現れたと考えられる．

3.2 フーリエ変換による解析

推定された潜在変数 y_t のダイナミクスの振る舞いをより詳しく知るために，フーリエ変換による解析を行った．図3は slow feature に対応する潜在変数 y_1 の推定結果に対するフーリエ変換の結果を示したものである．サンプリング周波数を 500Hz とした．左図が観測ノイズ $\sigma_x^2 = 0$ ，右図が観測ノイズ $\sigma_x^2 = 1.0$ の場合の結果をそれぞれ表している．slow feature に対応する潜在変数 y_1 はゆっくりと変化するダイナミクスを持つため，低周波側にピークが現れている．観測ノイズ $\sigma_x^2 = 0$ と $\sigma_x^2 = 1.0$ の場合を比較してみると， $\sigma_x^2 = 1.0$ の場合のスペクトルは $\sigma_x^2 = 0$ の場合のスペクトルに比べて振幅が増加していることが分かる．特に，高周波成分の増加が目立つ．これは正規白色雑音の性質をもつ観測ノイズ σ_x^2 の影響が潜在変数 y_t に転嫁されたことによるもので，図2の潜在変数 y_t の推定結果のズレを説明するものになっているといえる．潜在変数の他の成分 y_2, y_3 についても同様に，周波数スペクトルの増加が見られた (図省略)．

3.3 パラメータ推定値の変化

最後に，システムモデルのパラメータが観測ノイズに対してどのように変化しているのかを調べた結果を紹介する．これまで述べてきた潜在変数の推定結果は，システムモデルのパラメータ λ_i, σ_i^2 によって表現されているため，観測ノイズが加わった状況におけるパラメータの変化を知ることは重要であるといえる．また，異なる潜在変数のパラメータの変化

を比較することで、複数ある潜在変数に観測ノイズの影響がどのような割合で配分されるのかも調べる。

図4に観測ノイズに対するシステムモデルのパラメータ変化を示す。左図がパラメータ λ_i 、右図がシステムノイズ σ_i^2 の変化をそれぞれ表している。 $\sigma_x^2 = 0$ の点がそれぞれの真の値であり、その点から遠ざかるほど正確に推定できていないことになる。パラメータ λ_i は観測ノイズの増加とともに減少、システムノイズ σ_i^2 は増加している。システムモデルの式(5)より、システムモデルのパラメータの変化は観測ノイズの増加とともに、潜在変数 y_t のダイナミクスを速く・乱雑にする変化をしているといえる。

また、図4における各パラメータの変化の様子を比べてみると、パラメータ λ_i 、システムノイズ σ_i^2 とともに slow feature に対応する潜在変数 y_1 に関するものの変化幅が大きくなっていることが分かる。一方で、ノイズのような乱雑なダイナミクスをもつ潜在変数 y_3 に関するパラメータの変化幅は他の2つの成分に比べて小さく、特に、 λ_3 については影響をほとんど受けていない。つまり、観測ノイズの影響が強く現れているのは slow feature に対応する潜在変数 y_1 であり、 λ_1 の値が大きいくほど、すなわち、ゆっくりと変化する成分ほど観測ノイズの影響を受けやすく、誤った値が推定されてしまうという結果になっている。

4. おわりに

本論文では、確率的 SFA による情報抽出性能に対する観測ノイズの影響を定量的に議論し、観測ノイズが強い状況においては slow feature を含む潜在変数 y_t を正確に推定することができないこと、そして、潜在変数 y_t が複数ある場合、ゆっくりと変化する成分ほど観測ノイズの影響を受けやすいという性質を明らかにした。

まず、観測ノイズが加わった状況における潜在変数 y_t の推定結果から、近似によって推定の枠組みから外れた観測ノイズの影響が潜在変数 y_t のダイナミクスを乱雑に変化させることを示した。観測ノイズが潜在変数 y_t の推定結果へと転嫁される様子は、フーリエ変換によって得られた周波数スペクトルからも確かめられる。

さらに、システムモデルのパラメータ λ_i, σ_i^2 が潜在変数 y_t のダイナミクスを速く・乱雑にする方向へ変化することを示した。しかし、1つ1つのパラメータの変化幅は一様ではなく、パラメータ λ_i が大きいゆっくりと変化する成分ほど強く観測ノイズの影響を受ける。すなわち、意味のある情報を抽出することが難しくなる。ゆっくりと変化する成分の抽出を目的とする SFA において、この結果は情報抽出性能の低下を意味している。現実的な状況を考えて、実データには少なからず観測ノイズが加わっていると考えられる。そのため、

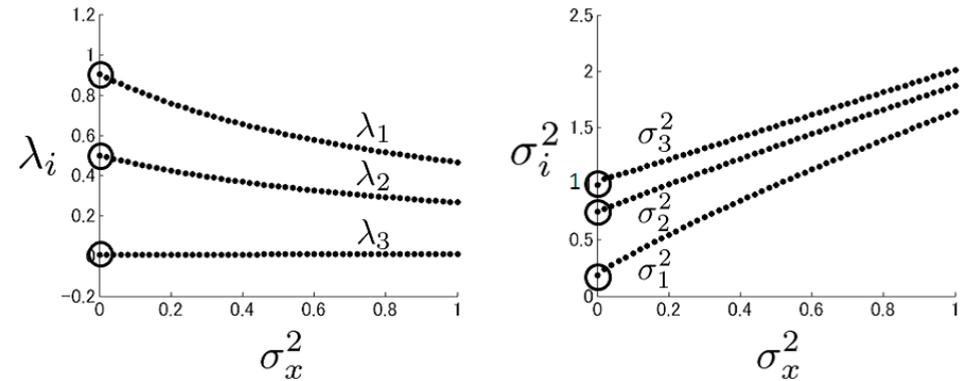


図4 観測ノイズ σ_x^2 に対するシステムモデルのパラメータ λ_i, σ_i^2 の変化。横軸に観測ノイズ σ_x^2 、縦軸に各パラメータの推定値をプロットしている。黒丸(●)は各パラメータの推定値を、 $\sigma_x^2 = 0$ における白丸(○)は各パラメータの真の値を示している。左側がパラメータ λ_i 、右側がシステムノイズ σ_i^2 の変化をそれぞれ表している。

Fig.4 Changes of system model parameters λ_i, σ_i^2 for observation noise σ_x^2 . The horizontal axis shows the observation noise σ_x^2 , and the vertical axis shows estimated value of system model parameters. Filled circle shows estimated value. Open circle in $\sigma_x^2 = 0$ shows true value of each parameter.(Left) the result of parameter λ_i . (Right) the result of system noise σ_i^2 .

確率的 SFA によって信頼性の高い結果を得るためには、観測ノイズを考慮したモデルの構築が必要である。

謝辞 本研究の一部は、文部科学省科学研究費補助金新学術領域研究「ヘテロ複雑システムによるコミュニケーション理解のための神経機構の解明」(領域番号 4103)[課題番号 22120506(大森)],「学際的研究による顔認知メカニズムの解明」(領域番号 4002)[課題番号 23119708(岡田)], 基盤研究(C)[課題番号 20509001(大森)], 基盤研究(A)[課題番号 20240020(岡田)], 挑戦的萌芽研究[課題番号 22650041(岡田)]の援助の下行われた。

参考文献

- 1) Hinton, G. E.: Connectionist learning procedures, *Artif. Intell.*, vol. 40, No.1-3, pp.185-234 (1989)

- 2) Wiskott, L. and Sejnowski, T. J.: Slow feature analysis: unsupervised learning of invariances, *Neural Comput.*, vol. 14, No.4, pp.715-770 (2002)
- 3) Turner, R. and Sahani, M.: A maximum-likelihood interpretation for slow feature analysis, *Neural Comput.*, vol. 19, No.4, pp.1022-1038 (2007)
- 4) Berkes, P. and Wiskott, L.: Slow feature analysis yields a rich repertoire of complex cell properties, *J. Vis.*, vol. 5, No.6, pp.579-602 (2005)
- 5) Franzius, M., Sprekeler, H. and Wiskott, L.: Slowness and sparseness lead to place, head-direction, and spatial-view cells, *PLoS Comput. Biol.*, vol. 3, No.8, pp.1605-1622 (2007)
- 6) Sprekeler, H. and Wiskott, L.: A theory of slow feature analysis for transformation-based input signals with an application to complex cells, *Neural Comput.*, vol. 23, No.2, pp.303-335 (2011)
- 7) 田中和之: 確率モデルによる画像処理技術入門, 森北出版 (2006)
- 8) Tanaka, K. and Inoue, J.: Maximum likelihood hyperparameter estimation for solvable markov random field model in image restoration, *IEICE on Trans. Inf. Syst.*, vol. E85-D, No.3, pp.546-557 (2002)
- 9) Tanaka, K., Shouno, H., Okada, M., and Titterington, D. M.: Accuracy of the Bethe approximation for hyperparameter estimation in probabilistic image processing, *J. Phys. A: Math. Gen.*, vol. 37, No.36, pp.8675-8695 (2004)
- 10) Hasegawa, R., Okada, M., Miyoshi, S.: Image segmentation using region-based latent variables and belief propagation, *J. Phys. Soc. Jpn.*, vol. 80, No.9, pp.093802-093802-4 (2011)
- 11) Tipping, M. E. and Bishop, C. M.: Probabilistic principal component analysis, *J. R. Stat. Soc.*, vol. 61, No.3, pp.611-622 (1999)