

結合カオス系における隠れた同期の検出

カーネル正準相関分析に基づくアプローチ

JST ERATO 合原複雑数理モデルプロジェクト 末谷 大道¹
 情報・システム研究機構 統計数理研究所 伊庭 幸人
 東京大学 生産技術研究所, JST ERATO 合原 一幸

1 はじめに

結合カオス系において、2つの素子のダイナミクスが非線形関数 Ψ によって互いに結ばれ、系全体の運動の実効自由度が制限される現象は、一般化同期 (Generalized Synchronization: GS)[1] と呼ばれている。結合カオス系における GS は、現実の複雑系において見られる要素間の隠れた相関に対する一つのメタファーと捉えられるだけでなく、癲癇などの脳の活動について GS の観点からの解析が行われるなど [2]、実際の応用面からの関心も持たれている。

実験や観察で得られたデータから GS を解析するためには、2つのダイナミクス間の非線形の関連性の深さを適切な尺度で定量化する必要がある。最近、データに内在する非線形構造を抽出するデータ解析の一般的な手法として、(再生) カーネル法と呼ばれる新たな機械学習の手法が注目を集めている [3]。現在までにサポートベクトルマシンに代表される様々なカーネル学習法が提案されているが、その中の1つに、2つの多変量データ間の相関の度合いを調べる多変量解析手法である正準相関分析 (Canonical Correlation Analysis: CCA)[4] にカーネル法の概念を適用したカーネル CCA が赤穂らによって提案されている [5, 6]。

本報告では、GS を示す力学系モデルに対するカーネル CCA の適用を通じて、2つのダイナミクスの間の非線形の関連性を解析する手法としてのカーネル CCA の有用性について論じる。

2 一般化同期とは

一般化同期の具体例として、パラメタの異なる2つの Hénon 写像が結合した次の4次元の離散時間・連続状態力学系を考える：

$$X : \begin{cases} x_1(t+1) = 1.4 - x_1(t)^2 + 0.3x_2(t) \\ x_2(t+1) = x_1(t), \end{cases} \quad (1)$$

$$Y : \begin{cases} y_1(t+1) = 1.4 - \{\gamma x_1(t)y_1(t) + (1-\gamma)y_1(t)^2\} + 0.1y_2(t) \\ y_2(t+1) = y_1(t). \end{cases} \quad (2)$$

ここで γ は結合係数を表す。図 1 (a) は結合が全くない場合 ($\gamma = 0$)、図 1 (b) は結合が十分強い場合 ($\gamma = 0.25$)、のカオスアトラクタの (x_1, y_1) 平面への射影である。結合がなければ個々の素子は独立に振る舞い何の関連性も見られないが (図 1(a))、十分な結合の強さの下では2つの素子の間に強い非線形の関連性が見られる (図 1(b))。実際に、図 1(b) の状態では、2つの力学系の軌道 $x(t) = (x_1(t), x_2(t))$, $y(t) = (y_1(t), y_2(t))$ の間に関数関係 $y(t) = \Psi(x(t))$ が存在して、お互いの運動が束縛されてる (このことは、方程式の形が具体的に分かっている場合は、条件付き Lyapunov 指数という量を計算してその符号を調べることでより確認することが出来る [7])。このような、結合力学系のアトラクタが時間に依存しない関数 Ψ によって $\mathcal{M} = \{(x, y) : y = \Psi(x)\}$ に閉じ込められる現象を一般化同期 (Generalized Synchronization: GS) という [1]。

¹E-mail: suetani@aihara.jst.go.jp

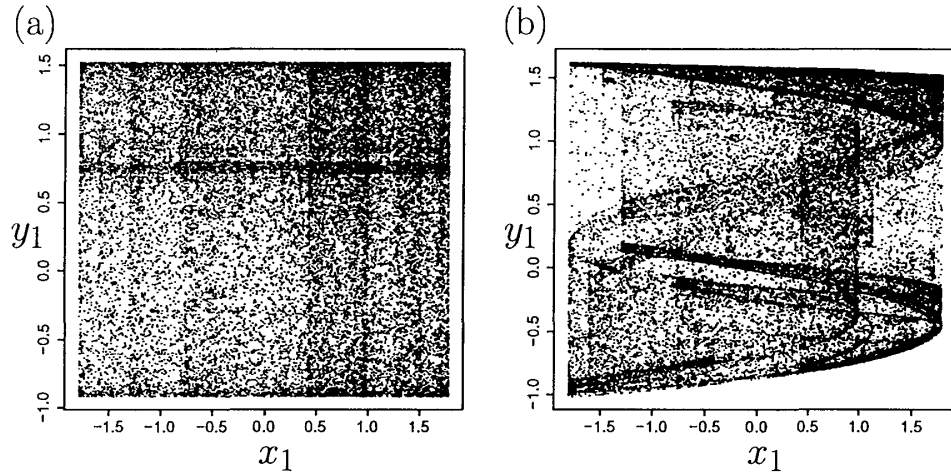


図 1: (a) $\gamma = 0$, (b) $\gamma = 0.25$ の場合における結合 Hénon 写像系のアトラクタの (x_1, y_1) 平面への射影.

3 カーネル正準相関分析

実験や観測で得られた時系列データだけを用いて GS を解析しようとする際、同期多様体 $\mathcal{M}$ は強い非線形性を有するため、相関係数のような変数間の線形的な関連性を測る指標では GS を検出することが困難であり、何らかの非線形変換をデータに施す必要がある。

カーネル法では、データを元の空間から高次元空間へ一旦写してから線形の高次元解析手法の適用を試みる。この際、高次元空間における 2 点間の内積を、元のデータ空間上の 2 点で定義されるカーネル関数によって評価する（カーネルトリック）ことで、非線形に拡張しながらも、計算量の逓減化やニューラルネットワーク学習で見られる局所最小解のような問題の回避を実現している [3]。本節では、種々のカーネル法の中でも、2 つのデータ間の非線形の関連性を調べることを目的として提案されたカーネル CCA [5, 6] の直観的な導入を行う。

いま、非線形の関連性を持つ 1 組の変数のペア $x \in \mathbf{R}^p, y \in \mathbf{R}^q$ を、あるクラス $\mathcal{F}$ に属する非線形関数 $f: \mathbf{R}^p \rightarrow \mathbf{R}, g: \mathbf{R}^q \rightarrow \mathbf{R}$ によって各々変換したときに、変数変換後の相関係数：

$$\rho_{\mathcal{F}} = \frac{\text{cov}(f(x), g(y))}{\sqrt{\text{var}(f(x))} \sqrt{\text{var}(g(y))}}, \quad (3)$$

が最大になるような f, g を求める問題を考えよう。 N 個のサンプルデータの組 $\{(x_i, y_i)\}_{i=1}^N$ が得られているとき、非線形関数 $f(\cdot), g(\cdot)$ として、以下の様に定義されるカーネル関数の重み付き線形和

$$f(x) = \sum_{i=1}^N \alpha_i k(x, x_i), \quad g(y) = \sum_{j=1}^N \beta_j k(y, y_j), \quad (4)$$

を用いる。例えば、 $k(x, x') = \exp(-\|x - x'\|^2 / 2\sigma^2)$ (Gaussian), $((x \cdot x') + \theta)^d$ (Polynomial), $\tanh(\eta(x \cdot x') + \theta)$ (Sigmoidal) などがカーネル関数として広く用いられている。重み係数 $\alpha = \{\alpha_i\}, \beta = \{\beta_j\}$ やカーネル関数のパラメタを調節することによって、(4) 式は様々な非線形関数を表現することが出来る。このような関数を考えることによって、(3) 式の最大化は以下の最適化問題となる。

$$\rho_{\mathcal{F}} = \max_{\alpha, \beta \in \mathbf{R}^N} \frac{\alpha' K_X K_Y \beta}{\sqrt{\alpha' K_X^2 \alpha + \kappa \|w_X\|^2} \sqrt{\beta' K_Y^2 \beta + \kappa \|w_Y\|^2}}. \quad (5)$$

ここで K_X, K_Y は, Gram 行列と呼ばれる $k(x_n, x_{n'})$ や $k(y_n, y_{n'})$ を第 nn' 成分とする $N \times N$ の行列であり, また, 分母の $\|\mathbf{w}_X\|^2 = \sum_{n=1}^N \alpha_n f(x_n) = \alpha' K_X \alpha$, $\|\mathbf{w}_Y\|^2 = \sum_{n=1}^N \beta_n g(y_n) = \beta' K_Y \beta$ はオーバーフィットを防止するために導入された正則化因子で, κ はそのパラメタである.

最適化問題 (5) 式の解は, 以下の一般化固有値問題の解として与えられる:

$$\begin{bmatrix} 0 & K_X K_Y \\ K_Y K_X & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \rho \begin{bmatrix} K_X(K_X + \kappa I) & 0 \\ 0 & K_Y(K_Y + \kappa I) \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}. \quad (6)$$

サンプルデータ数 N が大きいときは Gram 行列も巨大になるため, 実際の数値計算では不完全 Cholesky 分解などによって Gram 行列の次元を逡減させてから縮約した行列に関する一般化固有値問題を解くことになる. (6) 式では, 特に, 最大固有値が x, y 間の非線形に関連性の深さを定量的に表す指標として重要な意味を持つ. また, (4) 式における係数 (α, β) が固有ベクトルとして与えられ, その結果, 非線形関数 $f(\cdot), g(\cdot)$ が求まる. このとき, ρ_F を (第一) 正準相関係数 (canonical correlation coefficient), 非線形関数にデータ x, y を入れたときの変数 $u = f(x), v = g(y)$ を (第一) 正準変数 (canonical variates) という.

以上は重心がゼロ ($(1/N) \sum_{i=1}^N f(x_i) = (1/N) \sum_{i=1}^N g(y_i) = 0$) という仮定の下での定式化であるが, そうでない場合でも, Gram 行列に対して以下の操作を行うことで同じ固有値問題になる.

$$\hat{K} = K - \frac{1}{N}(\mathbf{j}\mathbf{j}')K - \frac{1}{N}K(\mathbf{j}\mathbf{j}') + \frac{1}{N^2}(\mathbf{j}\mathbf{j}')K(\mathbf{j}\mathbf{j}'). \quad (7)$$

ここで $\mathbf{j}$ は成分が全て 1 の N 次元ベクトルである.

4 数値実験の結果

結合 Hénon 写像 (1)・(2) 式に対してカーネル CCA を適用した結果について述べる. 図 2 は, 図 1 のときと同じ $\gamma = 0$ 及び $\gamma = 0.25$ の場合での長さ $N = 2 \times 10^3$ の時系列データに対してカーネル CCA を行ったときの第一正準変数空間での散布図 ($u_i = f(x_i), v_i = g(y_i), i = 1, 2, \dots, N$ の同時プロット) である. カーネル関数として幅 $\sigma = 0.1$ の Gauss カーネルを使用し, κ は 0.3 に設定した. また, データは立方体 $[0, 1]^4$ の範囲に収まるように規格化してある. 結合が無く, 各々のダイナミクスが独立に振る舞う場合は, 変数変換してもやはり何の相関も見られないが (図 2 (a)), GS が起きている場合は, 正準変数間の高い線形相関として GS が明瞭に検出されていることが分かる (図 2 (b)).

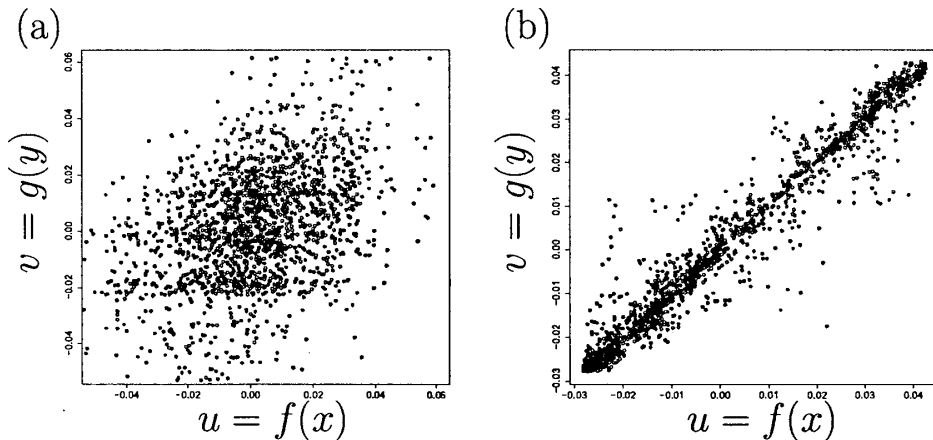


図 2: サンプルデータの第一正準変数空間への射影. (a) $\gamma = 0$, (b) $\gamma = 0.25$, の場合.

次に、正準相関係数 ρ_F の γ 依存性の結果を図 3 に示す。図 3 では、各 γ について、長さ $N = 10^3$ の時系列データに対してカーネル CCA を適用して最大固有値を求め、それを 10 回繰り返して平均したものを ρ_F として表示している。ここで、 κ は 0.1 に設定し、 $\sigma = 0.1, 0.2, 0.4$ の 3 通りについて調べた。また、条件付き Lyapunov 指数 λ の γ 依存性の結果も同時に示す。 $\gamma \sim 0.17$ までに ρ_F の急激な上昇が見られるが、これは λ の値が γ の増加に対して急激に減少して負に転移する領域と対応している。また、 ρ_F vs. γ のグラフでは $\gamma \sim 0.35$ 付近で少し窪んだ領域も見られるが、これも λ の γ の非単調な変化と対応しているなど、 ρ_F が GS を定量的に捉えるための良い指標であることを示している。一方、図 3 では通常の（線形変換の）CCA による結果も示しているが、CCA はこのような GS 転移を全く捉えていない。

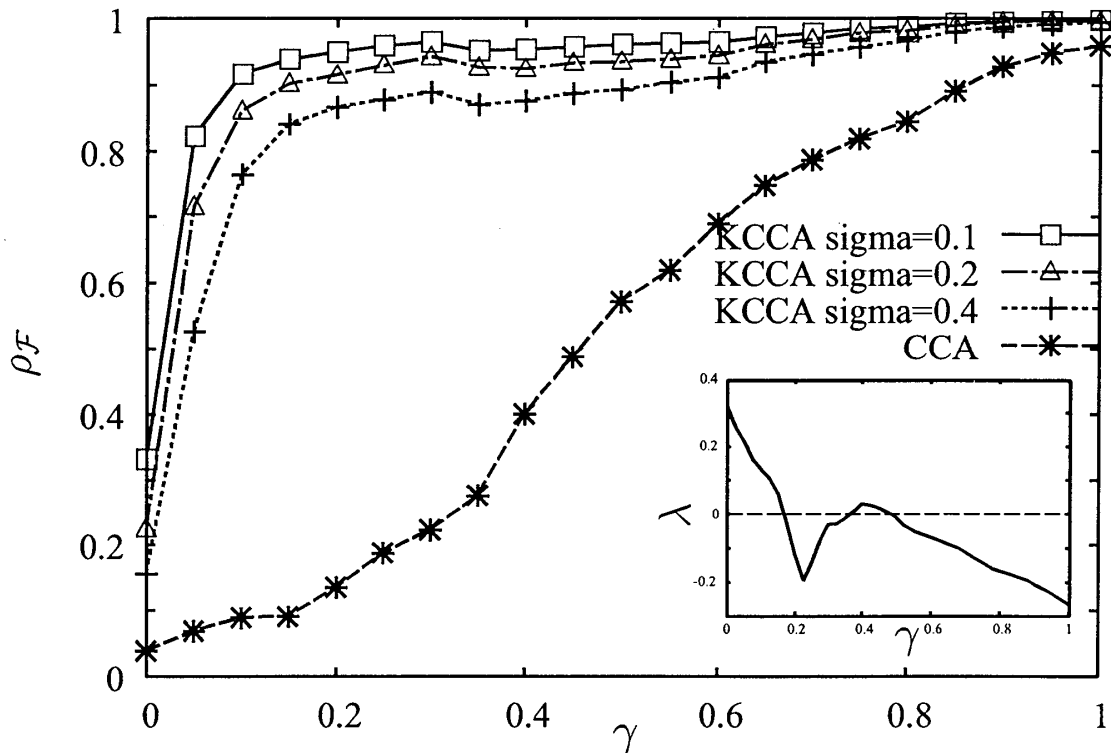


図 3: 第一正準相関係数 ρ_F の γ 依存性。Inset: 条件付き Lyapunov 指数の γ 依存性。

カーネル CCA では、固有値だけでなく固有ベクトルも同時に求まるので、それを活用した解析も行うことが出来る。ここでは、結合 Hénon 写像の結合係数 γ の値が時間と共に変化するとき、その時間変化を観測データから如何に検出するか、という問題を考えてみよう。図 4 (a) は、 γ の値が図 4 (b) のように $\gamma_0 = 0.6$ を基準に時間変化するときの力学変数 x_1, y_1 の時系列である。 $\gamma(t) \neq \gamma_0$ なる時間領域では、軌道は $\gamma = \gamma_0$ における同期多様体 M から外れているが、生の時系列を見てもそのような γ の非定常な変化を捉えることは出来ない。もし、非線形関数 Ψ の形が分かっていたら、 $\|y - \Psi(x)\|$ の時間発展を見ることによって γ の時間変化も分かる。しかしながら、図 1 (b) を見れば分かるように、通常 Ψ は簡単な式の形で与えることは出来ない。そこで、 Ψ そのものの代わりにカーネル CCA で求まる非線形変換を利用することを考える。もしカーネル CCA によって Ψ の構造が良く捉えられているならば、2つの正準変数の差の時間発展を見ることによって γ の時間変化を検出することが出来る。

図 4 (c) は、 $\gamma(t) = \gamma_0$ (一定) の下で生成した時系列を学習データとして予め用意しておいてから重み係数ベクトル α, β を推定した場合、図 4 (d) は、 $\gamma(t)$ が図 4 (b) の様に変化する

るときの時系列をそのまま使って α, β を推定した場合、の2つの正準変数の差の時系列である。図4(c)の結果から、同期多様体 M 上についているデータだけを学習に利用できた場合は、 M からの乖離は $u-v$ の時系列における連続的なバースト相として観測されることが分かる。また、図4(d)で示されているように、例えば M 上にはないデータも学習サンプルの中に混じっていても、 $1600 < t \leq 1800$ の時間領域を除いてやはりバースト相として γ の時間変化が検出できている。 $\gamma(t)$ が γ_0 から変化している時間領域でのデータは、学習を行う際の一種のノイズとも言え、図4(d)の結果はカーネルCCAがある程度のノイズがあってもロバストに働くことを示唆している。

(a)

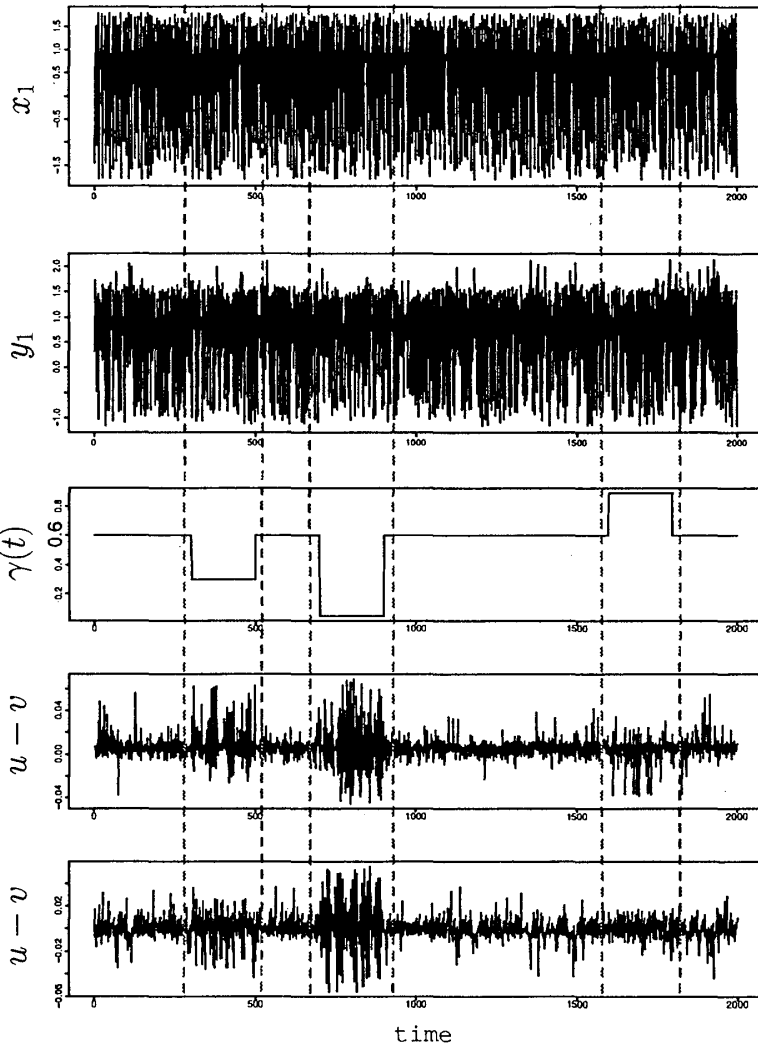


図4: (a) γ の値が (b) のように変化するときの x_1, y_1 の時系列. (c) $\gamma = 0.6$ (一定) の下で生成したデータから非線形変換 $f(\cdot), g(\cdot)$ を推定した場合 ($\sigma = 0.1, \kappa = 0.1$), (d) (b) のように γ が変化するときのデータから $f(\cdot), g(\cdot)$ を推定したとき ($\sigma = 0.2, \kappa = 0.1$) の、2つの正準変数の差の時系列.

5 まとめ

本報告では、カーネル学習法の一つであるカーネル CCA を導入し、GS を示す単純な写像モデルに対して適用した結果について紹介した。カーネル CCA から得られる固有値や固有ベクトルを利用することにより、GS の様な強い非線形構造を持つデータを解析できることを示した。

最後に幾つかの課題や将来への展望について述べる。今回の解析では、 σ や κ の値は実験者の側で適当に決めた。本報告では結果の図を示していないが、 σ の値を小さくし過ぎたときの影響について調べた所、 $f(\cdot), g(\cdot)$ が大域的な非線形変換とならず、少数のサンプルに重みが局在化してしまう現象などを観察している。カーネル関数や正則化因子のパラメタの最適値は、本来はデータから決めるべきものである。その方法の一つとして、例えばクロスバリデーション（交差検証法）が考えられる。このようなハイパーパラメタの自動的な決定については今後の課題である。少数のサンプルへの偏りの回避については、例えば、重み係数 α の分布のエントロピーを出来るだけ大きくするような制約項を付加することなどが挙げられる [5]。また、今回の解析ではサンプルの順序を入れ替えても結果は変わらないのでその意味では時系列解析にはなっていない。状態空間でのサンプル点同士の類似度を表すカーネル関数に、時間方向の情報も取り込むことによって本当の意味でのダイナミクスの解析への拡張を考えて行きたい。

参考文献

- [1] N.F. Rulkov et al., Phys. Rev. E **51**, 980 (1995).
- [2] S.J. Schiff et al., Phys. Rev. E **54**, 6708 (1996); M. Le Van Quyen et al., Brain Res. **792**, 24 (1998); R. Quian Quiroga, J. Arnhold, and P. Grassberger, Phys. Rev. E **61**, 5142 (2000).
- [3] J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Recognition*, (Cambridge University Press, Cambridge, 2004); B. Schölkopf and A.J. Smola, *Learning with Kernels, Support Vector Machines, Regularization, Optimization, and Beyond*, (The MIT Press, Cambridge, 2002); K.-R. Müller et al., IEEE Trans. on Neural Networks **12**, 181 (2001).
- [4] H. Hotelling, Biometrika **28**, 321 (1936).
- [5] 赤穂昭太郎, 情報論的学習理論ワークショップ (IBIS2000) 予稿集, 123 (2000); S. Akaho, In *Proceedings of the International Meeting of the Psychometric Society (IMPS2001)*, (Springer, Tokyo, 2001).
- [6] P.L. Lai and C. Fyfe, Int. J. of Neural Sys. **10**, 365 (2000); F.R. Bach and M.I. Jordan, J. of Machine Learning Res. **3**, 1 (2002).
- [7] L. Kocarev and U. Parlitz, Phys. Rev. Lett. **76**, 1816 (1996).