



システム制御理論と統計的機械学習

第 8 章：システム同定

加嶋 健司

October 10, 2025

京都大学 情報学研究科

本章の流れ

8.1 最小二乗法

最小二乗法を用いたシステム同定

8.2 スペクトル法

行列の固有値分解に基づいたシステム同定

最小二乘法

推定器としての最小二乗法 (1) 定式化

- 分布が既知の $p \sim \mathcal{N}(\bar{p}, \bar{\Sigma})$ および独立同分布の $w_k \sim \mathcal{N}(0, 1)$ と既知のベクトル q_k に対して,

$$y_k = p^\top q_k + w_k, \quad k = 0, 1, 2, \dots \quad (8.1)$$

が観測できるとする.

- $y_{0:k} = y_{0:k}$ のもとでの p の MAP 推定値は, 正則化付き二乗誤差

$$\sum_{i=0}^k (y_i - p^\top q_i)^2 + (p - \bar{p})^\top \bar{\Sigma}^{-1} (p - \bar{p})$$

を最小化する p と一致

- 観測値 y_0, y_1, \dots が順番に与えられる場合の推定値の逐次計算法

推定器としての最小二乗法 (2) 逐次最小二乗法 (recursive least square; RLS)

定理 8.1.1 – 正則化付き二乗誤差の最小化

$\bar{p} \in \mathbb{R}^n$, $\bar{\Sigma} \succ O$, $\alpha > 0$ および $q_k \in \mathbb{R}^n$, $y_k \in \mathbb{R}$, $k \in \mathbb{Z}_+$ に対して,

$$\check{p}_{-1} := \bar{p}, \quad \check{\Sigma}_{-1} := \alpha \bar{\Sigma}$$

$$\check{H}_k := \frac{\check{\Sigma}_{k-1} q_k}{\alpha + q_k^\top \check{\Sigma}_{k-1} q_k}, \quad \check{p}_k := \check{p}_{k-1} + \check{H}_k (y_k - q_k^\top \check{p}_{k-1}), \quad \check{\Sigma}_k := \frac{1}{\alpha} (I - \check{H}_k q_k^\top) \check{\Sigma}_{k-1}$$

とすると, \check{p}_k は正則化付き二乗誤差の最小化解

$$\sum_{i=0}^k \alpha^{k-i} (y_i - p^\top q_i)^2 + \alpha^k (p - \bar{p})^\top \bar{\Sigma}^{-1} (p - \bar{p}) \quad (8.6)$$

- 忘却係数 $\alpha (< 1)$: 小さい i (遠い過去) に対する誤差を小さく評価
- Σ_k を $1/\alpha (> 1)$ 倍し, 事後分布の信頼度をあえて下げる
- カルマンフィルタの特別な場合

推定器としての最小二乗法 (3) 一致性

$\alpha = 1$ かつ q_k が確率変数

定義 8.1.2 – \hat{p} : は p の一致推定量 (consistent estimator)

任意の $p \in \mathbb{R}^n$ に対して, $p = p$ のもとで \hat{p}_k が p に確率収束

定理 8.1.3 – 最小二乗推定の一貫性

$q_k \in \text{rv}(\mathbb{R}^n)$ および独立同分布の $w_k \in \text{rv}(\mathbb{R})$ に対して, $y_k = p^\top q_k + w_k$, $k \in \mathbb{Z}_+$ とし,

$$\mathbb{E}[w_k | q_k] = 0, \forall q_k, \quad \frac{1}{k} \sum_{i=0}^k q_i q_i^\top \xrightarrow{p} \Sigma_q \succ O, \quad \frac{1}{k} \sum_{i=0}^k q_i w_i \xrightarrow{p} 0$$

が成り立つと仮定する. このとき, 任意の $\bar{p} \in \mathbb{R}^n$, $\bar{\Sigma} \succ O$ に対して,

$$\check{p}_k := \left(\sum_{i=0}^k q_i q_i^\top + \bar{\Sigma}^{-1} \right)^{-1} \left(\sum_{i=0}^k q_i y_i + \bar{\Sigma}^{-1} \bar{p} \right) \quad (8.13)$$

は p の一致推定量である.

逐次システム同定 (1) システム同定とは

確率システムの同定問題

u_{\cdot}, y_{\cdot} の実現値のデータから

$$x_{k+1} = f(x_k, u_k, v_k), y_k = h(x_k, w_k) \quad (8.16)$$

が成り立つ状態 x , 雑音 v, w および f, h を決定すること

- ・ 状態変数 x の次元や取り方も決める必要がある.
- ・ 天下りの的に, 状態変数を n ステップの出力履歴, m ステップの入力履歴からなる

$$x_k := \begin{bmatrix} y_k & \dots & y_{k-n+1} & u_{k-1} & \dots & u_{k-m+1} \end{bmatrix}^{\top} \in \mathbb{R}^{n+(m-1)}$$

(出力関数は $h(x) := \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}$)

逐次システム同定 (2) ARX モデル

ARX モデル

$$y_{k+1} = \sum_{i=1}^n a_i y_{k+1-i} + \sum_{i=1}^m b_i u_{k+1-i} + v_k, \quad v_k : \text{i.i.d.} \quad (8.17)$$

ARX モデルは $w_{k+1} := v_k$ とおくと,

$$q_{k+1} := \begin{bmatrix} y_k & \dots & y_{k-n+1} & u_k & \dots & u_{k-m+1} \end{bmatrix}^\top \quad (8.19)$$

$$y_{k+1} = p^\top q_{k+1} + w_{k+1} \quad (8.20)$$

と表せる。ここで,

$$p \in \mathcal{N}(\bar{p}, \bar{\Sigma}), \quad v_k \sim \mathcal{N}(0, \sigma^2), \text{ i.i.d.}$$

とすると, 前項で考えた推定問題と同一であるため, 結果を直接用いることができる。

逐次システム同定 (3) RLS フィルタ

定理 8.1.1 – 正則化付き二乗誤差の最小化 (再掲)

$$\check{p}_{-1} := \bar{p}, \check{\Sigma}_{-1} := \alpha \bar{\Sigma}$$

$$\check{H}_k := \frac{\check{\Sigma}_{k-1} q_k}{\alpha + q_k^\top \check{\Sigma}_{k-1} q_k}$$

$$\check{p}_k := \check{p}_{k-1} + \check{H}_k (y_k - q_k^\top \check{p}_{k-1})$$

$$\check{\Sigma}_k := \frac{1}{\alpha} (I - \check{H}_k q_k^\top) \check{\Sigma}_{k-1}$$

とすると, \check{p}_k は正則化付き二乗誤差

$$\sum_{i=0}^k \alpha^{k-i} (y_i - p^\top q_i)^2 + \alpha^k \|p - \bar{p}\|_{\bar{\Sigma}}^2$$

の最小化解である.

アルゴリズム 3 忘却係数付き RLS フィルタ

- 1: **Input:** $\tilde{p}, \tilde{\Sigma}, q, y, u$
- 2: **Output:** p, Σ, q_+
- 3: $H \leftarrow \frac{\tilde{\Sigma} q}{\alpha \sigma^2 + q^\top \tilde{\Sigma} q}$
- 4: $p \leftarrow \tilde{p} + \tilde{\Sigma} (y - \tilde{p}^\top q)$
- 5: $\Sigma \leftarrow \frac{1}{\alpha} (I - H q^\top) \tilde{\Sigma}$
- 6: $q_+ \leftarrow [y, (q)_{1:n-1}, u, (q)_{n+1:n+m-1}]$

逐次システム同定 (4) 同定の例 – 入出力信号

ARX モデル

$$y_{k+1} = p^\top q_{k+1} + w_{k+1}$$

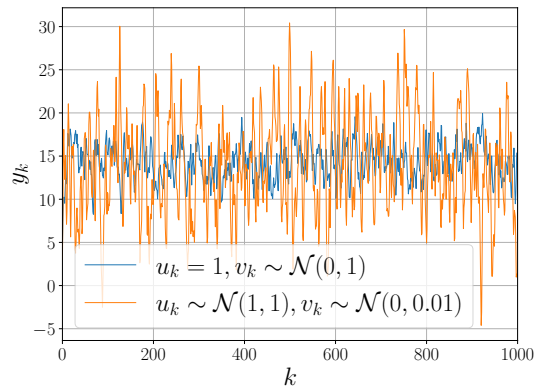
$$p \in \mathcal{N}(\bar{p}, \bar{\Sigma}), v_k \sim \mathcal{N}(0, \sigma^2), \text{ i.i.d.}$$

ARX モデルで与えられる確率システムにおいて, $n = 3, m = 2$ および

$$\bar{p} = 0, \bar{\Sigma} = 10^4 \times I,$$

$$(z - 0.5)(z - 0.4)(z - 0.3)$$

$$= z^3 - a_1 z^2 - a_2 z - a_3, b_1 = 2, b_2 = 1$$



標本経路

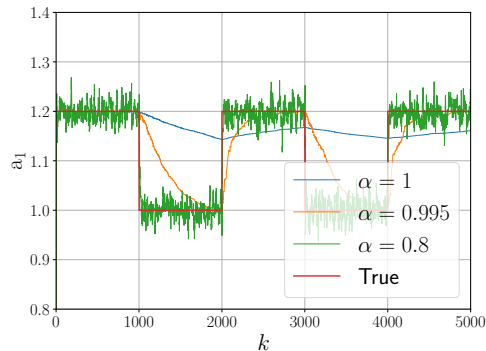
逐次システム同定 (5) 同定の例 – 忘却係数の影響

$$u_k \sim \mathcal{N}(0, 1), v_k \sim \mathcal{N}(0, 0.01)$$

a_1 のみ時間的に変化

忘却係数の影響

- 忘却係数 α が 1 に近ければ長い区間での誤差を小さくすることを要求するため、突発的な外乱の影響に対して頑健であるが、実応用の動的な環境で用いた場合は係数列 p の変化が推定結果になかなか反映されない。
- α が小さくなるほど直近の誤差を重点的に評価するため推定の速応性がよいが、観測雑音の影響を受けやすい。

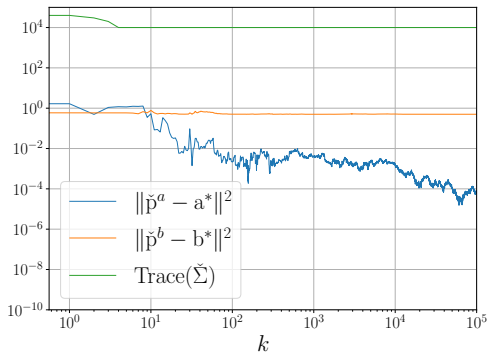


忘却係数の影響

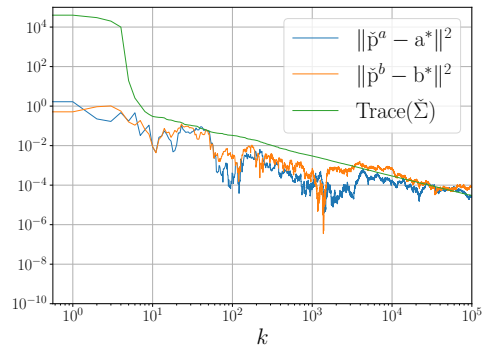
逐次システム同定 (6) 同定の例 – PE 性

$\alpha = 1$, 雑音に関して $v_k \sim \mathcal{N}(0, 1)$

$u_k = 1$ のとき, $\mathbb{E}[q_k q_k^\top]$ は非正則



$u_k = \mathcal{N}(0, 1)$



スペクトル法

モデル低次元化 (1)

モデル低次元化 (model reduction)

システムの性質を保持し，状態変数の次元を削減

- ・ 流体のように空間的に分布した現象を離散化した場合などの高次元モデル
- ・ 解析や設計に用いるには，非現実的な計算量が必要となることが多い

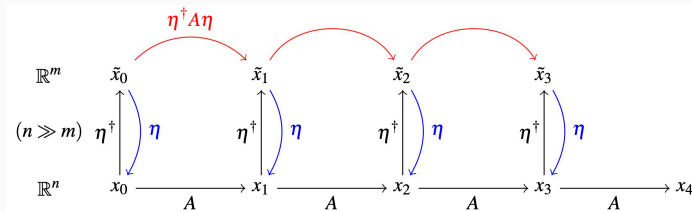
モデル低次元化 (2) 射影による低次元化

ペトロフ・ガレルキン射影

- 列フルランク行列 $\eta := [\eta_1, \dots, \eta_m] \in \mathbb{R}^{n \times m}$, $m \ll n$ に対して, 近似的に $x_k \in \text{Im } \eta$
- $\tilde{x}_k := \eta^\dagger x_k \in \mathbb{R}^m$, $\eta^\dagger := (\eta^\top \eta)^{-1} \eta^\top$ が $x_k \approx \eta \tilde{x}_k$ を満たす ($x_k \in \text{Im } \eta$ ならば等号)
- $x_{k+1} = Ax_k + Bu_k$ に左から η^\dagger を掛けて得られる

$$\tilde{x}_{k+1} = (\eta^\dagger A \eta) \tilde{x}_k + \eta^\dagger Bu_k \quad (8.25)$$

にしたがう \tilde{x}_k から, $x_k = \eta \tilde{x}_k$ により x_k が復元できる



モデル低次元化 (3) データ駆動型手法

大規模行列 $A \in \mathbb{R}^{n \times n}$ ではなく $x_{k+1} = Ax_k$ の実現値の時系列 $\{x_k\}_{k=0}^{\bar{k}}$ が与えられたときに，低次元モデルを求める（外生入力なし， $n \gg \bar{k}$ を想定）

定理 8.2.1 – 遷移行列の二乗誤差最小化解

$$\sum_{k=0}^{\bar{k}-1} \|x_{k+1} - Ax_k\|^2 (= \|Y - AX\|^2) \quad (8.28)$$

を最小化する A は

$$A_{\text{LS}} := YX^\dagger \in \mathbb{R}^{n \times n}, \quad X^\dagger := VD^{-1}U^\top \quad (8.29)$$

$$X := \begin{bmatrix} x_0 & x_1 & \dots & x_{\bar{k}-1} \end{bmatrix}, \quad Y := \begin{bmatrix} x_1 & x_2 & \dots & x_{\bar{k}} \end{bmatrix} \quad (8.26)$$

$$X = UDV^\top, \quad U^\top U = V^\top V = I, \quad D = \text{diag}(\sigma_1, \dots, \sigma_r) \succ O \quad (8.27)$$

モデル低次元化 (4) データが偏る部分空間への射影

$$X := \begin{bmatrix} x_0 & x_1 & \dots & x_{\bar{k}-1} \end{bmatrix}, Y := \begin{bmatrix} x_1 & x_2 & \dots & x_{\bar{k}} \end{bmatrix}$$

$$X = UDV^\top, U^\top U = V^\top V = I, D = \text{diag}(\sigma_1, \dots, \sigma_r) \succ O$$

- 流体力学などの実応用例においては $r \ll n$ であることが多く, $x_k \in \text{Im}U$ とする**固有直交分解** (proper orthogonal decomposition; POD) とよばれる低次元化手法が有効
- U は大規模密行列 $XX^\top \in \mathbb{R}^{n \times n}$ の非零固有値に対する固有ベクトルを並べた行列であるが, 補題 7.2.3 (行列の固有値分解に関する補題) と $X^\top X \in \mathbb{R}^{\bar{k} \times \bar{k}}$ の固有値分解により比較的効率よく計算できる.

モデル低次元化 (5) 動的モード分解

$\{x_k\}$ のデータ順序の情報も活用して, A_{LS} の固有空間を求める

定理 8.2.2 – 動的モード

$\tilde{A} := U^\top YVD^{-1} \in \mathbb{R}^{r \times r}$ の 0 でない固有値および対応する固有ベクトルを λ_i, ζ_i とする
このとき, A_{LS} の 0 でないすべての固有値・固有ベクトルは $\lambda_i, \eta_i := YVD^{-1}\zeta_i \in \mathbb{R}^n$.

- 動的モード分解 (dynamic mode decomposition; DMD) : 高次元システム
 $x_{k+1} = A_{LS}x_k$ を, 遷移行列が対角化された m 次元モデルに

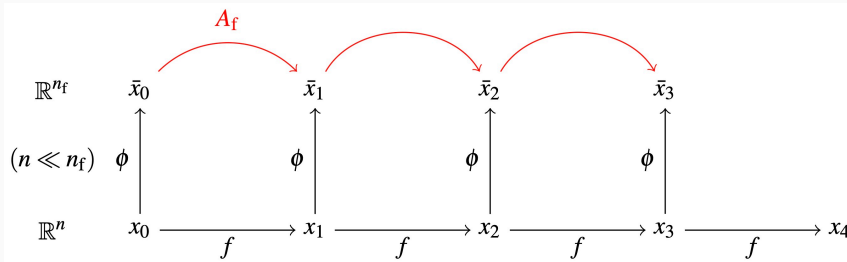
$$\eta := \begin{bmatrix} \eta_1 & \dots & \eta_m \end{bmatrix} \in \mathbb{R}^{n \times m}, \tilde{x}_0 := \eta^\dagger x_0 \quad (8.32)$$

$$\tilde{x}_{k+1} = \text{diag}(\lambda_1, \dots, \lambda_m) \tilde{x}_k, x_k = \eta \tilde{x}_k \quad (8.33)$$

モデル低次元化 (6) 特徴量の動的モード分解

非線形システム $x_{k+1} = f(x_k)$ の場合

- 既知の特徴写像 $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^{n_f}$ を用いて (ϕ により高次元空間 \mathbb{R}^{n_f} にリフトする), $\phi(x_{k+1}) = A_f \phi(x_k)$ が成り立つと仮定
- データ $\{\phi(x_k)\}_k$ に DMD を適用して低次元表現が求められる



クープマン作用素 (1) 観測行為の時間発展

特徴写像を用いた DMD の計算に「クープマン作用素とよばれる作用素の固有値の近似計算」という別解釈を与える

- 特徴写像 ϕ の選び方や DMD 以上の計算効率化を与えるものではない

例 8.2.3 – 線形システムの観測作用素の時間発展

初期状態 x_0 と時刻 k での観測値 y_k の関係が $y_k = CA^k x_0$

- 状態が時間発展 $x_{k+1} = Ax_k$, 観測は時不変行列 C の乗算
- 状態は x_0 から不変, 時刻 k における観測操作は $C_k := CA^k$ の乗算
 - 観測行列が $\mathcal{K}[C] := CA$ および $C_{k+1} = \mathcal{K}[C_k]$ にしたがって時間発展し, $y_k = \mathcal{K}^k[C]x_0$

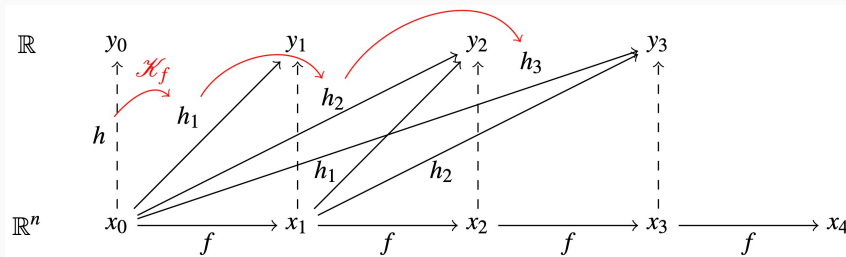
クープマン作用素 (2) 定義

定義 8.2.4 – クープマン作用素

関数 $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ に対して, \mathbb{R}^n から \mathbb{R} への有界な関数の集合 $L^\infty(\mathbb{R}^n)$ 上の汎関数

$$(\mathcal{K}_f[h])(\cdot) := h(f(\cdot)) \quad (8.37)$$

を f のクープマン作用素 (Koopman operator) とよぶ.



クーブマン作用素 (3) 固有関数

クーブマン作用素は線形作用素

- $(\mathcal{K}_f[\lambda h])(x) = \lambda \cdot h(f(x)) = \lambda(\mathcal{K}_f[h])(x)$
- $(\mathcal{K}_f[h + g])(x) = h(f(x)) + g(f(x)) = (\mathcal{K}_f[h] + \mathcal{K}_f[g])(x)$

$\mathcal{K}_f[h_i] = \lambda_i h_i$ の解として与えられる固有値 $\lambda_i \in \mathbb{C}$, 固有関数 $h_i \in L^\infty(\mathbb{R}^n)$

- $h(x_k) := \sum_i c_i h_i(x_k)$ の挙動に興味がある場合,

$$h(x_k) = (\mathcal{K}_f^k[h])(x_0) = \sum_i c_i \lambda_i^k h_i(x_0) \quad (8.40)$$

- 0 に近い λ_i に対応する固有関数 h_i で与えられる x_k の特徴量は, 時刻 k とともに直ちに 0 に収束
- 絶対値が大きい固有値に対応する固有関数は, 長期的に x_k の挙動を説明する特徴量

クーブマン作用素 (4) 動的モード分解による近似計算

$x_{k+1} = f(x_k)$ にしたがう x_i の時系列データから, f を具体的に求めることなく \mathcal{K}_f の固有値・固有関数を計算

定理 8.2.5 – クーブマンモード

関数 $\phi : \mathbb{R}^n \rightarrow \mathbb{R}_f^n$ の各要素は一次独立, A_f が $\phi(f(x)) = A_f \phi(x), \forall x \in \mathbb{R}^n$ を満たす

このとき, $\lambda, \bar{\eta}$ が A_f^\top の固有値および対応する固有ベクトルであることと, $\lambda, \bar{\eta}^\top \phi$ が定義域を $\mathcal{F}_\phi := \{p^\top \phi : p \in \mathbb{R}^{n_f}\}$ に限定した f のクーブマン作用素 \mathcal{K}_f の固有値と対応する固有関数であることは等価である.

- $\phi(x_{k+1}) \approx A_{LS} \phi(x_k)$ が成り立つ A_{LS} の固有値 (DMD により効率的に計算できる) は, \mathcal{K}_f の近似固有値であると解釈できる.

正準相関解析 (1) 定式化

$y_+ \in \text{rv}(\mathbb{R}^{\bar{n}})$, $y_- \in \text{rv}(\mathbb{R}^n)$ の平均が 0, 共分散行列が

$$\mathbb{E} \left[\begin{bmatrix} y_+ \\ y_- \end{bmatrix} \begin{bmatrix} y_+^\top & y_-^\top \end{bmatrix} \right] = \begin{bmatrix} \Sigma_+ & \Sigma_\pm \\ \Sigma_\pm^\top & \Sigma_- \end{bmatrix}, \Sigma_+ \succ 0, \Sigma_- \succ 0 \quad (8.42)$$

正準相関解析 (canonical correlation analysis; CCA) : 相関係数を用いて, y_- が有する y_+ に関する情報は w_- に圧縮し, y_- から得られる y_+ に関する情報は w_+ に集約する.

$$\rho := \frac{\text{Cov}[p_+^\top y_+, p_-^\top y_-]}{\sqrt{\text{Var}[p_+^\top y_+]} \sqrt{\text{Var}[p_-^\top y_-]}} = \frac{p_+^\top \Sigma_\pm p_-}{\sqrt{(p_+^\top \Sigma_+ p_+)(p_-^\top \Sigma_- p_-)}} \quad (8.43)$$

を最大にする $p_+ \in \mathbb{R}^{\bar{n}}$, $p_- \in \mathbb{R}^n$ を見つける

正準相関解析 (2) 一般化固有値問題への帰着

定理 8.2.6 – 正準相関解析

$$\begin{bmatrix} 0 & \Sigma_{\pm} \\ \Sigma_{\pm}^{\top} & 0 \end{bmatrix} \begin{bmatrix} p_{+} \\ p_{-} \end{bmatrix} = \rho \begin{bmatrix} \Sigma_{+} & 0 \\ 0 & \Sigma_{-} \end{bmatrix} \begin{bmatrix} p_{+} \\ p_{-} \end{bmatrix} \quad (8.45)$$

を満たす $\begin{bmatrix} p_{+}^{\top} & p_{-}^{\top} \end{bmatrix}^{\top} \neq 0$ が存在する ρ のうち、非負のものを大きい順に $\rho_1 \geq \rho_2 \geq \dots$ のように並べる ($\bar{n} \leq \underline{n}$ とすると $\{\rho_1, -\rho_1, \dots, \rho_{\bar{n}}, -\rho_{\bar{n}}, 0, \dots, 0\}$ の $\bar{n} + \underline{n}$ 個ある)。

このとき、規格化条件 $p_{+}^{\top} \Sigma_{+} p_{+} = p_{-}^{\top} \Sigma_{-} p_{-} = 1$ のもとで

$$\rho := \frac{\text{Cov}[p_{+}^{\top} y_{+}, p_{-}^{\top} y_{-}]}{\sqrt{\text{Var}[p_{+}^{\top} y_{+}]} \sqrt{\text{Var}[p_{-}^{\top} y_{-}]}} = \frac{p_{+}^{\top} \Sigma_{\pm} p_{-}}{\sqrt{(p_{+}^{\top} \Sigma_{+} p_{+})(p_{-}^{\top} \Sigma_{-} p_{-})}}$$

の最大値は ρ_1 である。

- ρ_i を第 i 正準相関係数とよぶ。

正準相関解析 (3) 特異値分解による計算

正定値行列 Σ_+ , Σ_- を下三角行列とその転置行列の積

$$\Sigma_+ = LL^\top, \Sigma_- = MM^\top \quad (8.48)$$

に分解（コレスキー分解）し、さらに $L^{-1}\Sigma_\pm M^{-\top}$ の特異値分解

$$L^{-1}\Sigma_\pm M^{-\top} = UDV^\top, U^\top U = V^\top V = I_n, n := \text{rank } \Sigma_\pm \quad (8.49)$$

$$D = \text{diag}(\rho_i), \rho_1 \geq \rho_2 \geq \dots \geq \rho_n > 0 \quad (8.50)$$

により ρ_i を定めればよく, $L^{-\top}U$, $M^{-\top}V$ の各列が定理の p_+ , p_- を与える

正準相関解析 (4) 正準ベクトル

正準ベクトルとよばれる

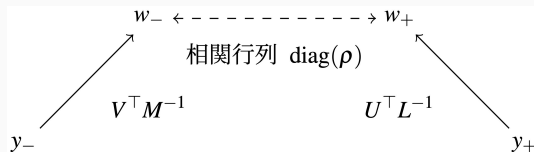
$$w_+ := U^\top L^{-1} y_+, \quad w_- := V^\top M^{-1} y_- \in \text{rv}(\mathbb{R}^n) \quad (8.51)$$

は

$$\mathbb{E}[w_+ w_+^\top] = \mathbb{E}[w_- w_-^\top] = I_n, \quad (8.52)$$

$$\mathbb{E}[w_+ w_-^\top] = D \quad (8.53)$$

を満たす.



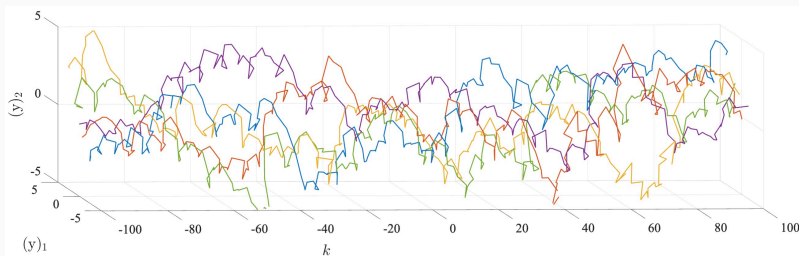
確率実現 (1) マルコフモデル

マルコフモデルとよばれる状態空間モデル

$$x_{k+1} = Ax_k + v_k, y_k = Cx_k + w_k, k \in \mathbb{Z} \quad (8.54)$$

で表現できることだけが既知の \mathbb{R}^{n_y} -値確率過程 y を考える.

- 大量かつ十分に長い標本経路が入手可能であれば, このモデル (状態空間の次元 n_x の選択を含む) を求めることはできるだろうか.
- 「過去の軌道が有する未来の軌道に関する情報を集約した変数」は何か?



確率実現 (2) 定式化

問題 8.2.8 – 確率実現問題

$\sum_{t=-\infty}^{\infty} \|\Lambda(t)\| < \infty$ を満たす $\{\Lambda(t)\} \in \mathbb{R}^{n_y \times n_y}$, $t = 0, \pm 1, \dots$ が与えられるとき, 任意の $k \in \mathbb{Z}$ に対して, 平均値が $\mathbb{E}[y_k] = 0$, 共分散行列が

$$\mathbb{E}[y_{k+t} y_k^\top] = \Lambda(t), \quad t = 0, \pm 1, \dots$$

を満たす y_\cdot のマルコフモデルを求めよ.

時刻 k を基準に, 未来・過去から作られる無限次元ベクトル

$$y_k^+ := \begin{bmatrix} y_k \\ y_{k+1} \\ \vdots \end{bmatrix}, \quad y_k^- := \begin{bmatrix} y_{k-1} \\ y_{k-2} \\ \vdots \end{bmatrix} \quad (8.55)$$

正準ベクトル w_- を状態変数 x_k に

確率実現 (3) 共分散行列

y_k^+, y_k^- の相互共分散行列, 自己共分散行列であるため, 正準相関分析がおこなえる

$$\Sigma_{\pm} := \mathbb{E}[y_k^+(y_k^-)^{\top}] = \begin{bmatrix} \Lambda(1) & \Lambda(2) & \Lambda(3) & \cdots \\ \Lambda(2) & \Lambda(3) & \Lambda(4) & \cdots \\ \Lambda(3) & \Lambda(4) & \Lambda(5) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad \Sigma_+ := \mathbb{E}[y_k^+(y_k^+)^{\top}] = \begin{bmatrix} \Lambda(0) & \Lambda(1)^{\top} & \Lambda(2)^{\top} & \cdots \\ \Lambda(1) & \Lambda(0) & \Lambda(1)^{\top} & \cdots \\ \Lambda(2) & \Lambda(1) & \Lambda(0) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

$$\Sigma_- := \mathbb{E}[y_k^-(y_k^-)^{\top}] = \begin{bmatrix} \Lambda(0) & \Lambda(1) & \Lambda(2) & \cdots \\ \Lambda(1)^{\top} & \Lambda(0) & \Lambda(1) & \cdots \\ \Lambda(2)^{\top} & \Lambda(1)^{\top} & \Lambda(0) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

特異値分解 (再掲): $n := \text{rank}(\Sigma_{\pm}) < \infty$ を仮定

$$L^{-1}\Sigma_{\pm}M^{-\top} = UDV^{\top}, \quad U^{\top}U = V^{\top}V = I_n, \quad n := \text{rank } \Sigma_{\pm}$$

$$D = \text{diag}(\rho_i), \quad \rho_1 \geq \rho_2 \geq \cdots \geq \rho_n > 0$$

確率実現 (4) 具体的な計算

正準相関解析の場合と同様に，正準ベクトルを

$$w_k^+ := U^\top L^{-1} y_k^+, w_k^- = V^\top M^{-1} y_k^- \quad (8.59)$$

- 過去の全情報 y_k^- が未来 y_k^+ に関して有する情報は w_k^- に集約されていると考える
- 便宜上，スケーリングして状態変数に採用

$$x_k := D^{1/2} w_k^- = D^{1/2} V^\top M^{-1} y_k^- \quad (8.60)$$

- y_i の時系列データの標本が得られれば， x_i の時系列データの標本も得られることになるので，最小二乗法などで $x_{k+1} - Ax_k$ や $y_k - Cx_k$ が適切な意味で小さくなるよう A, C を求めればよい。

実際，この考え方は有効であり，**部分空間同定**（subspace identification）とよばれる一連の手法につながっていく

第 8 章

1. 最小二乗法

推定器としての最小二乗法

逐次システム同定

2. スペクトル法

モデル低次元化

クーブマン作用素

正準相関解析

確率実現