時系列の状態空間モデル

樋口さぶろお https://hig3.net

龍谷大学大学院 理工学研究科 数理情報学専攻

理論物理学特論　L11 (2022-12-07 Wed)

最終更新: Time-stamp: "2022-12-13 Tue 10:19 JST hig"

今日の目標

- 確率過程の自己回帰移動平均モデル ARMA(m, k) を説明できる.
- 時系列モデルの状態空間モデルを説明できる

L10-Q1

Quiz 解答: 移動平均 実は問題の $x(t)$ はこうやって作ってました．周期が 4 だから，4 次の移動平均でトレンドをだいたい回復できるでしょ．

t	0	1	2	3	4	5	6	7	8	
トレンド	0.0	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	
周期	2.0	−2.0	2.0	−2.0	2.0	−2.0	2.0	−2.0	2.0	−
ノイズ	−0.2	0.2	0.2	0.2	0.2	−0.2	0.2	−0.2	0.2	
x	1.8	−1.6	2.6	−1.2	3.0	−1.2	3.4	−0.8	3.8	
y_3		0.9	−0.1	1.5	0.2	1.7	0.5	2.1	1.0	
y_4			0.0	0.8	0.9	1.1	1.2	1.5		

L10-Q2

Quiz 解答: コレログラム 1-a, 2-c,3-b

L10-Q3

Quiz 解答: 自己回帰モデル AR(2) の例

1. モデルの定義より $X_2 = \phi_1 X_1 + \phi_2 X_0 + \epsilon_2 = \epsilon_2$.

2. モデルの定義より
$X_3 = \phi_1 X_2 + \phi_2 X_1 + \epsilon_3 \phi_1(\phi_1 X_1 + \phi_2 X_0 + \epsilon_2) + \phi_2 X_1 + \epsilon_3 = \phi_1\epsilon_2 + \epsilon_3$.
$\mathrm{E}[(\phi_1\epsilon_2 + \epsilon_3)^2] = \phi_1^2\sigma^2 + 2 \cdot \phi_1 \cdot 0 + \sigma^2$.
3. モデルの定義より
$X_4 = \phi_1 X_3 + \phi_2 X_2 + \epsilon_4 = \phi_1^2\epsilon_2 + \phi_1\epsilon_3 + \phi_2\epsilon_2 + \epsilon_4$.
$\mathrm{E}[X_3] = \mathrm{E}[X_4] = 0$ より,
$\mathrm{Cov}[X_3, X_4] = \mathrm{E}[X_3 X_4] = \phi_1^3\mathrm{E}[\epsilon_2^2] + \phi_1\phi_2\mathrm{E}[\epsilon_2^2] + \phi_1\mathrm{E}[\epsilon_3^2] =$
$(\phi_1^3 + \phi_1\phi_2 + \phi_1)\sigma^2$.

ここまで来たよ

10 時系列の MA モデル
- k 次の移動平均モデル MA(k)

11 時系列の状態空間モデル
- 状態空間モデル
- 状態空間モデルの例

移動平均の定義

定義 (移動平均)

時系列 $\{X_t\}$.
k 次の後方移動平均 $Y_{t,k} = \frac{1}{k}(X_{t-k+1} + \cdots + X_{t-1} + X_t)$.

移動平均モデルの定義のアイデア

$$X_t = \phi_1 Y_{t-1,1} + \phi_2 Y_{t-1,2} + \cdots + \phi_k Y_{t-1,k} + \epsilon_t$$

としては?
漸化式を解いていくと, けっきょくすべて $\epsilon_{t'}$ $t-k-1 \leq t' \leq t-1$ に帰着する.

k 次の移動平均モデル MA(k)

定義 (k 次の移動平均モデル MA(k))

$t = 0, 1, 2, \ldots$: 時刻
$\{\epsilon_t\}_{t=1,\ldots}$,$\{X_t\}_{t=0,1,\ldots}$: 連続型確率変数の列
$\theta_k \in \mathbb{R}$: パラメタ
$X_t (t \geq 1)$ は次のように定まる.

$$X_t = \beta_0 + \theta_0 \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_k \epsilon_{t-k}. \quad (t \geq k)$$

ただし $\theta_0 = 1$. ϵ_t は次を満たす.

$$\mathrm{E}[\epsilon_t] = 0 \quad \text{(WN1)}$$
$$\mathrm{E}[\epsilon_t \epsilon_s] = \sigma^2 \delta_{ts} \quad \text{(WN2)}$$
$$\mathrm{E}[X_t \epsilon_s] = 0 \quad (t < s) \quad \text{(PI)}$$

移動平均モデル MA(k) の母平均値・母自己共分散

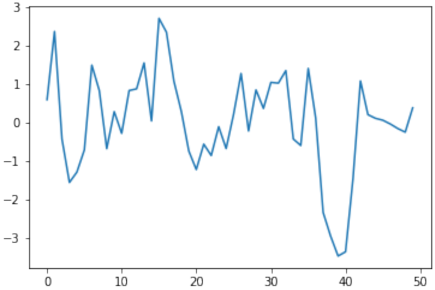
$$\mathrm{E}[X_t] = \beta_0.$$

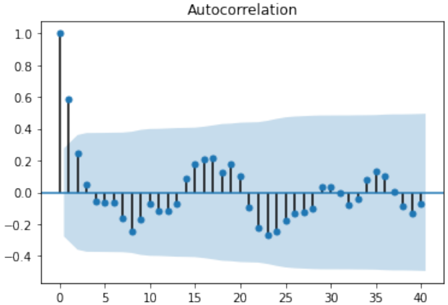
$$\begin{aligned}\mathrm{V}[X_t] =& \mathrm{V}[\theta_0\epsilon_t + \theta_1\epsilon_{t-1} + \cdots + \theta_k\epsilon_{t-k}] \\ =& \theta_0^2\mathrm{V}[\epsilon_t] + \theta_1^2\mathrm{V}[\epsilon_{t-1}] + \cdots + \theta_k^2\mathrm{V}[\epsilon_{t-k}] \\ =& (\theta_0^2 + \theta_1^2 + \cdots + \theta_{t-k}^2)\sigma^2.\end{aligned}$$

$$\begin{aligned}\mathrm{Cov}[X_t, X_{t-s}] =& \mathrm{E}[(\theta_0\epsilon_t + \theta_1\epsilon_{t-1} + \cdots + \theta_k\epsilon_{t-k}) \\ & \times (\theta_0\epsilon_{t-s} + \theta_1\epsilon_{t-s-1} + \cdots + \theta_k\epsilon_{t-s-k})] \\ =& \cdots \\ =& \begin{cases} 0 & (k < s) \\ (\theta_0\theta_s + \theta_1\theta_{s+1} + \cdots + \theta_{k-s}\theta_k)\sigma^2 & (k \geq s) \end{cases}.\end{aligned}$$

移動平均モデル MA(k) のサンプルパスとコレログラム

MA(1), $\theta_1 = 0.9$.





青網掛けは, Bartlett’s formula による信頼区間. この外に出ると, 帰無仮説 $r = 0$ が有意水準 $\alpha = 0.05$ で棄却される.

L10-Q4

Quiz(移動平均モデル MA(1) の例)

1 次の移動平均モデル MA(1) モデルを考える (授業中に使った記号で).
$\beta_0 = 0$ とする.

1. 確率変数 X_2 を, ϵ_t, θ_1 で表そう.
2. 母期待値 $\mathrm{E}[X_2^2]$ を σ と θ_1 で表そう.
3. 母自己共分散 $\mathrm{Cov}[X_2, X_3]$ を σ と θ_1 で表そう.
4. 母自己共分散 $\mathrm{Cov}[X_2, X_4]$ を求めよう.

m, k 次の自己回帰移動平均モデル ARMA(m, k)

X_t の漸化式の右辺が, ϵ_t に加え, AR(m) と MA(k) の和であるモデル.

階差数列

時系列の階差を繰り返しとると, 定常になって, ARMA で解析できることがある.

階差 difference $\leftrightarrow$ 和分 (離散和のこと) integrated

m, d, k 次の自己回帰和分移動平均モデル ARIMA(m, d, k)

d 階階差が ARMA(m, k) にしたがう時系列, ARMA を和分した時系列.

Code example for ARMA(2,2)

Listing 1: arma

```
from statsmodels.tsa.arima_process import ArmaProcess
ar=[1,-phi1,-phi2]            # 1は必ず含める マイナスをつける
ma=[1,theta1,theta2]          # [theta_0,theta_1,theta_2] 1は必ず
MA=ArmaProcess(ar,ma) # 自己回帰移動平均モデルを計算するオブジ
x=MA.generate_sample(nsample=50) # 時間の長さ=50
```

ar, ma は, 漸化式の L のラグ多項式 の係数のリスト.

定義 (ラグ演算子)

ラグ演算子 L を $LX_t = X_{t-1}$ で定める.

$$
\begin{aligned}
X_t =& \phi_1 X_{t-1} \phi_2 X_{t-2} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} \\
(1 - \phi_1 L^1 - \phi_2 L^2) X_t =& (1 + \theta_1 L^1 + \theta_2 L^2) \epsilon_t
\end{aligned}
$$

ここまで来たよ

10 時系列の MA モデル
- k 次の移動平均モデル MA(k)

11 時系列の状態空間モデル
- 状態空間モデル
- 状態空間モデルの例

時系列の隠れた変数

琵琶湖の漁の毎年の漁獲量は時系列. 過去の漁獲量 (kg) のデータを見て, AR(2) っぽい, とか, ϕ_1, ϕ_2, σ^2 はいくつ? とか推定することはできる. しか〜し.

- 実は背後に, 琵琶湖の生息するナマズの総数とか, モロコの総数とか, もっと言えば, 水温とか水草の量とか, 観測できないより本質的な時系列があるのでは?
  - 被食者捕食者の増減の話とかやったよね?$/$やります.　微分方程式, 数理モデル
- ホワイトノイズ ϵ_t っていうけど, 今年はたまたま稚魚が育たなかったっていうノイズと, 今年はコロナで漁船の整備ができなかった (?) っていうノイズは別なのでは?

システムと観測を分けて, もっと精密に考えたい!

定義 (状態空間モデル (の簡単な場合))

- 観測できない量
  - $\mathbf{X}_t$: 状態ベクトル (2 次元「状態空間」の元)
  - $\mathbf{v}_t$: システムノイズベクトル (2 次元)
  - w_t: 観測ノイズ (1 次元) $\mathrm{V}[w_t] = \sigma_w^2$.
  - $\mathbf{v}_t, w_t$ は互いに独立なホワイトノイズ
- 観測できる量
  - Y_t: 観測値 (1 次元)
- 定数
  - F, G: 2×2 行列
  - H: 1×2 行列

$$
\begin{aligned}
\text{システム (状態) モデル}\quad \mathbf{X}_t &= F\mathbf{X}_{t-1} + G\mathbf{v}_t \\
\text{観測モデル}\quad Y_t &= H\mathbf{X}_t + w_t
\end{aligned}
$$

状態空間 $= \mathbf{X}_t$ の観測できない空間
状態ベクトルは,

$$
\begin{aligned}
\mathbf{X}_t =& F\mathbf{X}_{t-1} + G\mathbf{v}_t \\
=& F(F\mathbf{X}_{t-2} + G\mathbf{v}_{t-1}) + G\mathbf{v}_{t-1} \\
=& \cdots = F^t\mathbf{X}_0 + \sum_{t'=0}^{t} F^{t-t'} G\mathbf{v}_{t'}
\end{aligned}
$$

のように漸化式で計算される.
観測値 Y_t は漸化式で結ばれておらず, 各時刻ごとに $\mathbf{X}_t$ から

$$
Y_t = H\mathbf{X}_t + w_t
$$

で計算される.

ここまで来たよ

10 時系列の MA モデル
- k 次の移動平均モデル MA(k)

11 時系列の状態空間モデル
- 状態空間モデル
- 状態空間モデルの例

例: 琵琶湖の漁獲量

$$F = \begin{pmatrix} 1.01 & 0.3 \\ -0.5 & 0.98 \end{pmatrix}, \mathbf{X}_t = \begin{pmatrix} \text{なまずの個体数} \\ \text{モロコの個体数} \end{pmatrix},$$

$$G = \begin{pmatrix} 1 & 2 \\ 100 & 200 \end{pmatrix}, \mathbf{v}_t = \begin{pmatrix} \text{正規分布 (水草の量)} \\ \text{正規分布 (水温)} \end{pmatrix}$$

$$Y_t = \text{漁獲量 (kg)}, H = \begin{pmatrix} 1000 & 10 \end{pmatrix}, w_t = \text{正規分布 (漁船の数)}$$

例:AR(m)

AR(2)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t.$$

は 2 次元の状態空間で表せる.

$$\begin{pmatrix} X_t \\ X_{t-1} \end{pmatrix} = \begin{pmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} X_{t-1} \\ X_{t-2} \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \text{正規ノイズ} \\ \text{正規ノイズ} \end{pmatrix},$$

$$Y_t = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} X_t \\ X_{t-1} \end{pmatrix} + 0 \cdot \text{正規ノイズ}$$

AR(3) 3 次元の状態空間で表せる.

$$\begin{aligned} X_t &= \phi_1 X_{t-1} && + \phi_2 X_{t-2} + \phi_3 X_{t-3} && + \epsilon_t, \\ X_{t-1} &= 1 \cdot X_{t-1}, \\ X_{t-2} &= && + 1 \cdot X_{t-2}, \\ Y_t &= X_t && && + 0 \cdot \epsilon'_t \end{aligned}$$

例: ローカルレベルモデル (ランダムウォークプラスノイズモデル)

$$X_t = X_{t-1} + \text{乱数},$$
$$Y_t = X_t + \text{別の乱数}$$

ランダムウォークだが, 観測者によるウォーカーの位置の測定値 Y_t は正確ではない. その前提で, 正確な位置 X_t を求めたい.

L11-Q1

Quiz(AR(4) の状態空間モデル)

4 次の自己回帰モデル AR(4) モデルを状態空間モデルとして考える (授業の記号で).

1. 状態ベクトルを書こう.
2. システムモデル, 観測モデルに現れる行列の, F, G, H を書こう.

状態空間モデルの特徴

- パラメタが多い
- $\mathrm{AR}(m)$ や親戚の MA, ARMA, ARIMA, SARIMA などのモデルを統一的に表せる.
- これでないと表せないモデルもある.
- 制御工学でよく扱われて, 解法が整備されている. $Y_1, \ldots, Y_t$ から状態推定する.
- 問題の分類
  - 平滑化 過去の状態 $X_{t'}$ $(t' < t)$ を推定する
  - フィルタ 現在の状態 $X_{t'}$ $(t' = t)$ を推定する
  - 予測 将来の状態 $X_{t'}$ $(t' > t)$ を推定する
- 効率のよい解法カルマンフィルタがある
- HMM=Hidden Markov model 隠れマルコフモデルは, 状態空間ベクトル $\mathbf{X}_t$ を離散値にしたバージョン

状態空間モデルを使う手続き

1. 与えられた Y_t を見る
2. (原理から考えて) モデルを書く
3. F, G, H が未知ならそれをパラメタ推定する
4. 与えられた学習データ $\{Y_t\}_{t=0,\ldots,T}$ から $\mathbf{X}_t$ を状態推定 (予測, フィルタ, 平滑化) する.
5. テストデータ Y_t と, $\mathbf{X}_t$ から求めた $\hat{y}_t$ を比較して, 性能を評価する

カルマンフィルタを備える, 状態空間モデルのライブラリ
https://www.statsmodels.org/dev/generated/statsmodels.tsa.statespace.structural.UnobservedComponents.html

状態空間モデルの扱いやすい場合 1: 正規ノイズ

- $\mathbf{v}_t, \mathbf{w}_t$ が正規分布にしたがうと仮定する．このとき，正規分布の加法性から，$\mathbf{X}_t, Y_t$ などがすべて正規分布にしたがい，それらの母平均値，母共分散の間の関係式をつかって解析できる．

状態空間モデルの扱いやすい場合 2: 隠れマルコフモデル

- 「$\mathbf{X}_t$ が離散型確率変数の場合」を考えたい. $X_t = 1, 2, \cdots, n$.
- システムモデルは「 $\mathbf{X}_t$ が確率的に時間発展する」ことを表すので, マルコフ連鎖で書くのが自然. M は推移確率行列.　計算科学 B, 確率モデル及び実習

$$P(X_t = 1) = M_{11}P(X_{t-1} = 1) + M_{12}P(X_{t-1} = 2) + \cdots + M_{1n}P(X_{t-1} = n)$$
$$P(X_t = 2) = M_{21}P(X_{t-1} = 1) + M_{22}P(X_{t-1} = 2) + \cdots + M_{2n}P(X_{t-1} = n)$$
$$\vdots$$
$$P(X_t = n) = M_{n1}P(X_{t-1} = 1) + M_{n2}P(X_{t-1} = 2) + \cdots + M_{nn}P(X_{t-1} = n)$$

- 観測できない (隠れた)$X = 1, 2, \cdots, n$ という状態があり, ある法則で時間発展する. 観測される Y_t の値は X から決まる.