6/25 兎澤

統計解析手法
時系列データの統計解析

2015/06/25
中西航
nakanishi [at] civil.t.u-tokyo.ac.jp

時系列解析

- (工学系や経済学系における)代表的な書籍
  - 北川源四郎「時系列解析入門」岩波書店
  - 沖本竜義「経済・ファイナンスデータの計量時系列分析」朝倉書店
  - 廣松毅・浪花貞夫・高岡慎「経済時系列分析」多賀出版
- 厚い本が1冊書けるようなトピック
- 学内にも時系列だけを1学期間扱う講義あり
- 数学的・情報学的な研究もいまだ盛ん
- 今日は本当の入門だけを紹介

時系列データとはなにか？

- 時系列データ：時間の推移と共に観測されるデータ
- しかし、ほとんどのデータは時間の推移と共に観測されるのではないか？
- 通常は観測される<u>順序</u>に意味があるもののことを時系列データという
- <u>順序に意味がある</u>とはどういうことか？

Aさんの身長 (6/24 PM 1:00で 172.1cm)
○意味がある ← Aさんの身長の時間変化を知りたいとき
○意味がない ← 20代の男性の平均身長を 〃

基本統計量

- 時間は離散化して考える

- データの記述方法

$y_t = y_1, y_2, \cdots, y_K, \cdots$

- 平均 mean

$\mu = E(y_t)$ ← $t = t_K$ での分布を見てる。
$[0, \infty]$ での y の平均を見てるわけではない

- 分散 variance

$V(y_t) = E(y_t - \mu_t)^2$

- 標準偏差(≒ボラティリティ) standard deviation

$\sqrt{V(y_t)}$

基本統計量

- 自己共分散 autocovariance

  同一の時系列における異時点間の共分散

  - 1次の自己共分散

    $\gamma_{1t} = Cov(y_t, y_{t-1}) = E[(y_t - \mu_t)(y_{t-1} - \mu_{t-1})]$

  - k次の自己共分散

    $\gamma_{kt} = Cov(y_t, y_{t-k})$

  - 自己共分散関数

    γ_{kt} を k の関数としてみたもの

基本統計量

- 自己相関(係数) autocorrelation coefficient

  統計の相関係数と意義は変わらない。

  - 自己相関関数 ← k の関数

    $\rho_{kt} = Corr(y_t, y_{t-k}) = \frac{Cov(y_t, y_{t-k})}{\sqrt{V(y_t)V(y_{t-k})}} = \frac{\gamma_{kt}}{\sqrt{\gamma_{0t}\gamma_{0\,t-k}}}$

    分散を0次の自己共分散とみる

    $\rho_{0t} = 1,\ |\rho_{kt}| \leq 1$

  - コレログラム

確率過程 Stochastic Process

- 時系列データはふつう1度しか観測できない
  - 無数の観測(試行)が可能であるという前提で平均や分散を定義することはできない
  - 2015年1月1日の気温から、2015年1月1日の気温の期待値や分散を求めることは統計的に不可能
- そこで、一連の観測を、確率変数の列からの実現値とみなす：**確率過程** stochastic process

$\{y_t\}_{t=-\infty}^{\infty}$ からの実現値の一つが

$\{y_t\}_{t=1}^{t} = y_1, y_2, \ldots, y_t$

定常性 Stationarity

- 確率過程に対して時間不変なものを想定する
- 何を不変とするかによって「弱定常性」と「強定常性」の2つがある
- 弱定常性：以下の(1)(2)が成り立つ場合
  - (1)平均は時間不変　$E(y_t) = \mu$
  - (2)自己共分散は時間差のみに依存

    $Cov(y_t, y_{t-k}) = \gamma_k$　t に依存しない!

  - このとき、自己相関も時間差のみに依存

    $Corr(y_t, y_{t-k}) = \frac{\gamma_k}{\gamma_0} = \rho_k$

定常性 Stationarity

10

- 確率過程に対して時間不変なものを想定する
- 何を不変とするかによって「弱定常性」と「強定常性」の2つがある

- 強定常性：以下が成り立つ場合
  - 同時分布が任意の時刻の組に対して同一

$(y_t, y_{t+1}, \cdots, y_{t+k})^T$ の同時分布が t,k について同一.

- 現実的な問題
  - 定常を仮定すると時系列解析が行いやすくなるので、いろいろな工夫をして定常を仮定することが多い
  - データが(本当に)定常かどうかを判定することは難しい

iid系列・ホワイトノイズ

11

- 各時点のデータが独立に同一の分布に従う系列：iid系列
  - 期待値0のiid系列はモデルの撹乱項(確率的変動を表現する部分：≒誤差項)として扱える

$y_t \sim iid(\mu, \sigma^2)$

- 独立・同一分布の仮定を少し緩める→ホワイトノイズ

一見 同じに 見えるが まあ スルーして

$E(\varepsilon_t) = 0$

$\gamma_k = (\varepsilon_t, \varepsilon_{t+k}) = \begin{cases} \sigma^2 & (k=0) \\ 0 & (k \neq 0) \end{cases}$

~~εt~~ $\varepsilon_t \sim wn(\sigma^2)$

平均との差

【参考】工学部の数学では何を行うのか

13

- 理論(モデル、解法)の勉強
  - モデルやそれが成立する条件、その数学的な解き方は数学者が示してくれている(と考えておく)

- 実践：現実の問題に対してどう適用するか
  - データの取得
    - どういうデータを集めるべきか
  - モデルの作成と決定
    - どういうモデルを当てはめるべきか、分析したいことはどういうモデルの成立条件を満たしているか
  - パラメータの推定と決定
    - 計算可能性を満たしているか
  - 結果の解釈・将来の予測
    - 妥当な結果か、新たな発見はあるか、役に立つか

時系列解析の代表的モデル

14

- 結局のところ自己相関のモデル化ができればよい
  - MA過程 (Moving Average: 移動平均)
  - AR過程 (Auto Regression: 自己回帰)
  - ARMA過程 (自己回帰移動平均)

- 今から説明すること
  - モデル自体の説明
  - パラメータ推定の方法
  - データを得たときのモデル化の方法
  - 複数のモデルが得られたときに良いものを選ぶ方法
  - 推定されたモデルを用いた予測の方法

MA過程

- 時系列データをホワイトノイズの線形和で表したもの
- 例：1次のMA過程(MA(1)と表記)

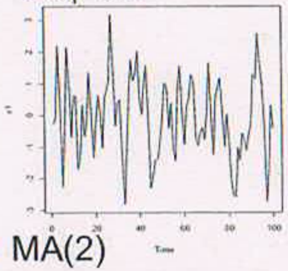
相関 $y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1}$: MA(1)
$y_{t-1} = \mu + \varepsilon_{t-1} + \theta_1 \varepsilon_{t-2}$

ε と θ が決まれば (推定する) y が決まる.

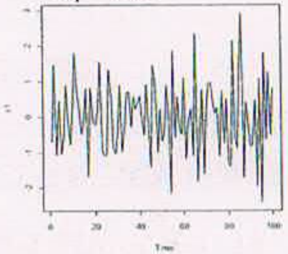
- 自己相関が発生する仕組み
- ホワイトノイズε_tの値が決まるとy_tの値も決まる

例

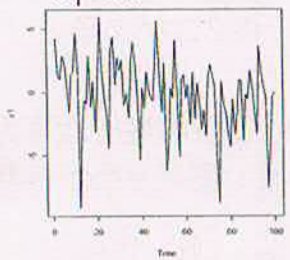
- MA(1)
  - θ_1=0.8　滑らか



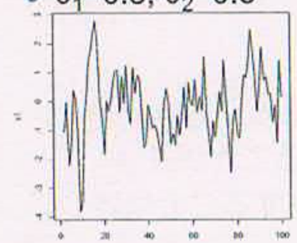
θ_1=-0.5　ギザギザ



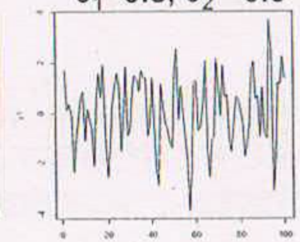
θ_1=2.7



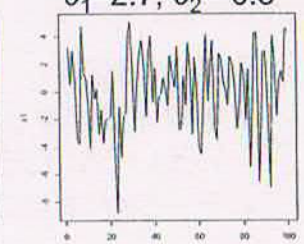
- MA(2)
  - θ_1=0.8, θ_2=0.5



θ_1=0.8, θ_2=-0.5



θ_1=2.7, θ_2=-0.5



MA(1)の統計量

- 期待値

$E(y_t) = E(\mu + \varepsilon_t + \theta_1 \underbrace{\varepsilon_{t-1}}_{E(\varepsilon)=0}) = \mu$

- 分散

$\gamma_{0t} = V(y_t) = V(\overset{定数}{\mu} + \varepsilon_t + \theta_1 \varepsilon_{t-1}) = V(\varepsilon_t) + \theta_1^2 V(\varepsilon_{t-1}) + 2\theta_1 Cov(\varepsilon_t, \varepsilon_{t-1})$

$= (1+\theta_1^2)\sigma^2 \geq \sigma^2$

- 共分散

$\gamma_{1t} = Cov(y_t, y_{t-1}) = Cov(\varepsilon_t, \varepsilon_{t-1}) + Cov(\varepsilon_t, \theta_1 \varepsilon_{t-2}) + Cov(\varepsilon_{t-1}, \theta_1 \varepsilon_{t-1}) + Cov(\theta_1 \varepsilon_{t-1}, \theta_1 \varepsilon_{t-2})$

$= \theta_1 \sigma^2$

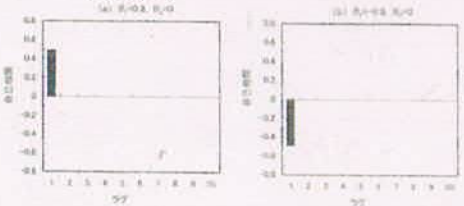
- 1次の自己相関　$\rho_1 = \frac{\theta_1}{1+\theta_1^2} \ (\leq \frac{1}{2})$
- 高次の自己相関　$k \geq 2$ のとき $\rho_k = 0$

MA(q)過程

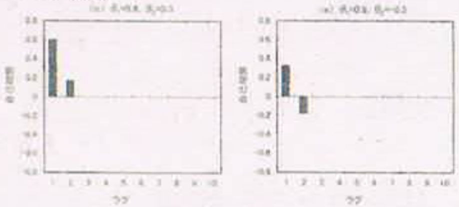
- 式　$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$
- 期待値　$E(y_t) = \mu$
- 分散　$V(y_t) = \gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)\sigma^2$
- 共分散　$\gamma_k = \begin{cases} (\theta_k + \theta_1\theta_{k+1} + \cdots + \theta_{q-k}\theta_q)\sigma^2 & (1 \leq k \leq q) \\ 0 & (それ以外) \end{cases}$
- 自己相関　$\rho_k = \begin{cases} \frac{\theta_k + \theta_1\theta_{k+1} + \cdots + \theta_{q-k}\theta_q}{1 + \theta_1^2 + \cdots + \theta_q^2} & (1 \leq k \leq q) \\ 0 & (それ以外) \end{cases}$
- パラメータの値によらず常に定常

MA(q)過程

- 例：パラメータの値とコレログラム
  - MA(1)



MA(2)



- 問題点
  - 自己相関が長い期間ある場合にはパラメータが膨大に必要
  - パラメータがホワイトノイズにかかっているので、得られたモデルや適用結果の意味解釈が難しい
  - 同一の期待値と自己相関を持つMA(q)過程が複数存在しうる

AR過程

- 過程が自身の過去に回帰された形で表現されるもの
- 例：1次のAR過程(AR(1)と表記)

$$y_t = c + \phi_1 y_{t-1} + \varepsilon_t$$

$$y_{t-1} = c + \phi_1 y_{t-2} + \varepsilon_{t-1}$$

- 自己相関が発生する仕組み

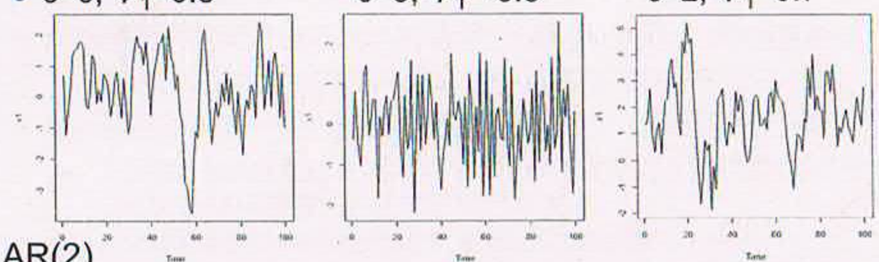
MAと同じ考え

- ホワイトノイズε_tの値が決まるとy_tの値も決まる
- MA過程と異なり、
  - AR過程は常に定常とは限らない
    - AR(1)が定常になるのは$|\Phi_1|<1$のとき
  - 期待値は定数項とは一致しない
  - 初期値を決定する必要がある

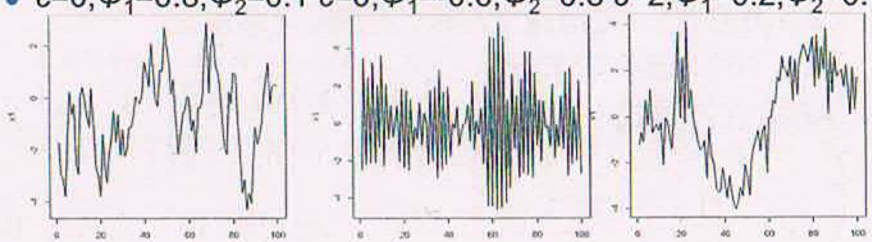
例

定常の例

- AR(1)
  - c=0, Φ_1=0.8　　c=0, Φ_1=-0.5　　c=2, Φ_1=0.7



- AR(2)
  - c=0,Φ_1=0.8,Φ_2=0.1　c=0,Φ_1=-0.6,Φ_2=0.3　c=2,Φ_1=0.2,Φ_2=0.7



AR(1)過程の統計量

- 期待値

$$E(y_t) = E(c + \phi_1 y_{t-1} + \varepsilon_t) = c + \phi_1 E(y_{t-1})$$

$$\Leftrightarrow \mu = c + \phi_1 \mu \Leftrightarrow \mu = \frac{c}{1-\phi_1}$$

- 分散

$$\gamma_{0t} = V(y_t) = V(c + \phi_1 y_{t-1} + \varepsilon_t) = \phi_1^2 V(y_{t-1}) + V(\varepsilon_t) + 2\phi_1 \underbrace{cov(y_{t-1}, \ldots)}_{0}$$

$$= \phi_1^2 \underbrace{V(y_{t-1})}_{\gamma_0} + \sigma^2$$

$$\therefore \gamma_0 = \frac{\sigma^2}{1-\phi_1^2}$$

AR(1)過程の統計量

- 共分散

$\gamma_k = Cov(y_t, y_{t-k}) = Cov(y_t, y_{t-k}) = Cov(\phi_1 y_{t-1} + \varepsilon_t, y_{t-k})$

$= Cov(\phi_1 y_{t-1}, y_{t-k}) = \phi_1 \gamma_{k-1} \iff \gamma_k = \phi_1 \gamma_{k-1}$

- 自己相関
  - ユール・ウォーカー方程式 Yule-Walker equation

$\rho_k = \phi_1 \rho_{k-1}$　自己相関の漸化式

AR(p)過程[定常な場合]

- 式　$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t$
- 期待値　$E(y_t) = \dfrac{c}{1-\phi_1-\cdots-\phi_p}$
- 分散　$V(y_t) = \gamma_0 = \dfrac{\sigma^2}{1-\phi_1\rho_1-\cdots-\phi_p\rho_p}$
- 共分散と自己相関

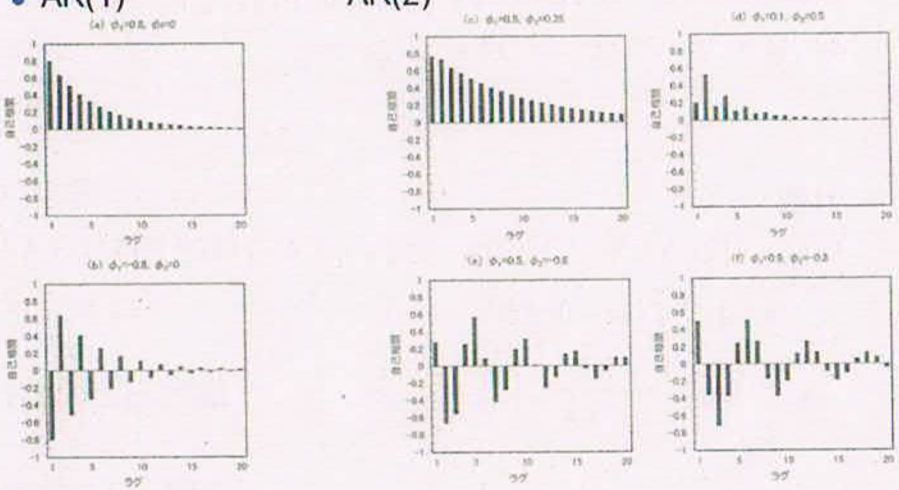
$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \cdots + \phi_p \gamma_{k-p} \quad (k \geq 1)$

$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p}$

- 自己相関は指数的に減少する

AR(p)過程

- 例：パラメータの値とコレログラム
  - AR(1)　　AR(2)



AR過程の定常性

- AR(2)過程のユール・ウォーカー方程式

$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}$

$k=1$のとき、$\rho_1 = \phi_1 \underset{1}{\rho_0} + \phi_2 \underset{\rho_1\text{∵弱定常性}}{\rho_{-1}} \iff \rho_1 = \dfrac{\phi_1}{1-\phi_2}$

$k=2$のとき $\rho_2 = \phi_1\rho_1 + \phi_2\rho_0 = \dfrac{\phi_1^2 + \phi_2 - \phi_2^2}{1-\phi_2}$

- AR(p)過程のユール・ウォーカー方程式[省略]
  - ρ_1からρ_{p-1}までは連立方程式を解くと求まる
  - ρ_p以降は逐次求まる
- AR特性方程式と定常性

$\rho_k = 1 - \phi_1 z - \cdots - \phi_p z^p = 0$

p個の解の絶対値が全て1より小さければ定常。

$1 - \phi_1 z - \phi_2 z^2 = 0$ の解の逆数を λ_1, λ_2 とすると、

$\rho_k = \dfrac{(1-\lambda_2^2)\lambda_1^{k+1} - (1-\lambda_1^2)\lambda_2^{k+1}}{(\lambda_1-\lambda_2)(1+\lambda_1\lambda_2)}$ ← 解けるということだけ示される

ARMA(p, q)過程[定常な場合]

- p次のAR過程とq次のMA過程を組み合わせたもの
- ARMA過程(ARMA(p, q)と表記)

$y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$

- 平均 $E(y_t) = \dfrac{c}{1-\phi_1-\cdots-\phi_p}$ ← ARと同じ
- q+1次以降の自己共分散・自己相関

$\gamma_k = \phi_1 \gamma_{k-1} + \cdots + \phi_p \gamma_{k-p}$

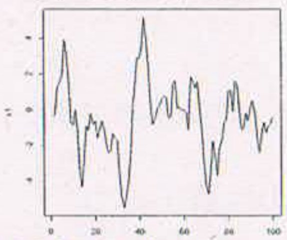
$\rho_k = \phi_1 \rho_{k-1} + \cdots + \phi_p \rho_{k-p}$

その方色々表現できる。

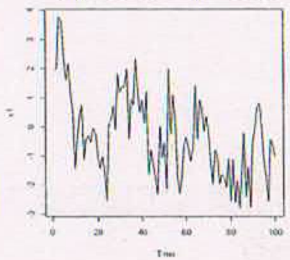
  - q次まではMAの性質があるため表現は難しい
- 自己相関は指数的に減衰

例

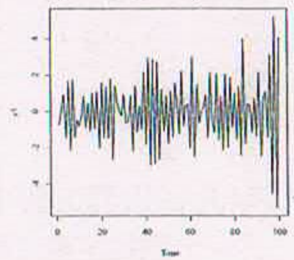
- ARMA(1,1), c=0
  - Φ_1=0.8, Θ_1=0.4



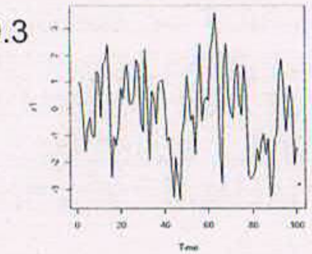
Φ_1=0.8, Θ_1=-0.4



Φ_1=-0.8, Θ_1=-0.4



- ARMA(2,2), c=0
  - Φ_1=0.2, Φ_2=0.1, Θ_1=0.8, Θ_2=-0.3



ARMA過程の定常性

- ARMA過程のAR部分が定常⇒そのARMA過程は定常 （MA定常より）
- いくつかの重要なこと
  - AR過程が定常である⇔AR過程がMA(∞)過程に書き直せる
    - 従って、定常なARMA過程はMA(∞)過程に書き直せる
  - MA過程がAR(∞)過程に書き直せるとき、反転可能とよぶ
    - 【復習】MA過程は定常だが、同一の期待値と自己相関を持つ異なるMA過程が存在する場合がある
→反転可能なMA過程を選択することが望ましい
    - MA特性方程式と反転可能な条件

$1 + \theta_1 z + \cdots + \theta_q z^q = 0$　MA特性方程式

反転可能：解の絶対値が全て1未満

    - これを満たすとき、ARMA過程全体も反転可能

ARMAモデルの推定[最小二乗法]

残差二乗和が最小

- AR(1)を例に、誤差項はiid正規分布に従うと仮定

$y_t = c + \phi_1 y_{t-1} + \varepsilon_t$, $\varepsilon_t \sim iid\, N(0, \sigma^2)$

残差 $e_t = y_t - \hat{c} - \hat{\phi}_1 y_{t-1}$

$E = \sum_{t=1}^{T} e_t^2 = \sum_{t=1}^{T} (y_t - \hat{c} - \hat{\phi}_1 y_{t-1})^2 \rightarrow \min$

正規方程式 $\begin{cases} \dfrac{\partial E}{\partial c} = \cdots\cdots = 0 \\ \dfrac{\partial E}{\partial \phi_1} = \cdots\cdots = 0 \end{cases}$ ← 板書探して

自分で微分する、覚えるだけだとまちがえるよー。

$\hat{c} = \overline{y_t} - \hat{\phi}_1 \overline{y_{t-1}}$

$\hat{\phi}_1 = \dfrac{\sum_{t=1}^{T} (y_t - \overline{y_t})(y_{t-1} - \overline{y_{t-1}})}{\sum_{t=1}^{T} (y_{t-1} - \overline{y_{t-1}})^2}$

ただし $\overline{y_t} = \frac{1}{T} \sum_{t=1}^{T} y_t$.

ARMAモデルの推定[最小二乗法]

- AR(1)を例に、誤差項はiid正規分布に従うと仮定

- AR(p)過程の最小二乗推定量は以下の性質を持つ
  - 一致推定量である
  - 基準化すると漸近的に正規分布に従う
  - 誤差項がiid正規分布に従うとき、一致推定量のなかで、漸近的に最小の分散共分散行列を持つ
  - しかし、BLUEではない(不偏ではない)
- ARMAは最小二乗法では推定できない

ARMAモデルの推定[最尤法]

- AR(1)を例に、誤差項はiid正規分布に従うと仮定
- 未知パラメータ(c, Φ_1, σ^2)
- 同時密度$f\left(y_T, y_{T-1}, \dots, y_1 \mid y_0; c, \phi_1, \sigma^2\right) = \prod_{t=1}^{T} \underline{f\left(y_t \mid y_{t-1}; c, \phi_1, \sigma^2\right)} \to \max.$

AR(1)だから

- ここで $f(y_t \mid y_{t-1}) \sim N\left(c + \phi_1 y_{t-1}, \sigma^2\right)$ を用いて計算すると

$$\tilde{c} = \bar{y}_t - \tilde{\phi}_1 \bar{y}_{t-1},\quad \tilde{\phi}_1 = \frac{\sum_{t=1}^{T}(y_t - \bar{y}_t)(y_{t-1} - \bar{y}_{t-1})}{\sum_{t=1}^{T}(y_{t-1} - \bar{y}_{t-1})^2},\quad \tilde{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}\left(y_t - \tilde{c} - \tilde{\phi}_1 y_{t-1}\right)^2$$

ただし、

$$\bar{y}_t = \frac{1}{T}\sum_{t=1}^{T} y_t$$

答え一致（最小二乗法と）

- AR(p)過程の最尤推定量は以下の性質を持つ
  - 一致推定量である
  - 基準化すると漸近的に正規分布に従う
  - 一致推定量のなかで漸近的に最小の分散共分散行列を持つ

ARMAモデルの選択

- ARMA(p, q)のpとqを決めないとパラメトリックにモデルが記述できないので、推定もできない
- 自己相関の様子をみながら、モデルの候補を列挙する
  - 人間が見て判断するand/orコンピュータに任せる
  - **あくまでも参考**になる情報としては以下：
    - MA(q)は自己相関関数が最大となるqを選ぶとよい**かも**
    - AR(p)は偏自己相関関数が最大となるpを選ぶとよい**かも**
  - 必ずいくつかの候補で推定を試すこと！
- 偏自己相関 Partial autocorrelation
  - y_tとy_{t-k}の間だけの相関

$y_{t-1}, y_{t-2}, \dots y_{t-(k-1)}$は入らな…

  - 具体的には次式の解Φ_{kk}のこと

$$\begin{pmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{pmatrix}$$

ARMAモデルの選択

- 複数のARMA(p, q)が推定されたとき、どれを選ぶか？
- 「情報量規準 Information criterion が最小のものを選ぶ」
  - 直感的な説明：当てはまりがよくシンプルなモデルが良い
    - pやqを大きくすると…
      - いま持っているデータへの当てはまりが良くなる
      - パラメータの推定精度が下がる
      - 推定されたモデルによる未知のデータの予測力が落ちる
  - 赤池情報量規準 AIC= -2L(θ) + 2k
  - ベイズ情報量規準 BIC= -2L(θ) + $k$ln(T)
    - L(θ)：最大対数尤度、k：パラメータ数、T：標本数
    - 第1項が当てはまりの良さ、第2項が複雑さを表す
  - 【参考】情報量規準自体もいまなお数学の研究対象
- いろいろ試しても当てはまりが悪いときは、定常性を疑うなど、より原点に立ち返る必要がある

モデルを用いた予測

残差じゃないよ−

- (例)時刻tまでの結果から、時刻t+1以降の値を予測したい
- 予測誤差： e_{t+1} = (y_{t+1}の観測値) − (y_{t+1}の予測値)

AR(1)：$y_{t+1} = c + \phi_1 y_t + \varepsilon_{t+1}$ まだ未知 → y_{t+1}はε_{t+1} (ε_t) に従う分布
→ e_{t+1}も分布.

- 平均2乗誤差Mean Squared Error
  - いま持っている情報からy_{t+1}の分布を得る
  - 書き方：$y_{t+1|t}$：tまでの情報から得た$t+1$の予測値

$MSE = E(\hat{y}_{t+1} - y_{t+1})^2 = E(e_{t+1})^2 \rightarrow \min$

- 最適予測optimal forecast: 平均2乗誤差が最小となる予測

定常AR過程の予測

- AR(1)過程$y_t = c + \phi_1 y_{t-1} + \varepsilon_t, \varepsilon_t \sim \mathrm{iid}N(0, \sigma^2)$ の場合
  - 1期先最適予測

$y_{t+1} = c + \phi_1 y_t + \varepsilon_{t+1}$　それぞれ一かな一程度で. $E(\varepsilon_{t+1}) =$

$\hat{y}_{t+1|t} = c + \phi_1 y_t$, $e_{t+1} = \varepsilon_{t+1}$, $MSE(\hat{y}_{t+1|t}) = \sigma^2$

  - 2期先最適予測

$y_{t+2} = (1+\phi_1)c + \phi_1^2 y_t + \varepsilon_{t+2} + \phi_1 \varepsilon_{t+1}$

$\hat{y}_{t+2|t} = (1+\phi_1)c + \phi_1^2 y_t$, $MSE(\hat{y}_{t+2|t}) = (1+\phi_1^2)\sigma^2$

  - h期先最適予測

$y_{t+h} = (1+\phi_1+\cdots+\phi_1^{h-1})c + \phi_1^h y_t + \varepsilon_{t+h} + \phi_1 \varepsilon_{t+h-1} + \cdots + \phi_1^{h-1} \varepsilon_{t+1}$

$= c\sum_{k=1}^{h} \phi_1^{k-1} + \phi_1^h y_t + \sum_{k=0}^{h-1} \phi_1^{k} \varepsilon_{t+h-k}$

$= \underbrace{\frac{1-\phi_1^h}{1-\phi_1} c + \phi_1^h y_t}_{\hat{y}_{t+h|t}} + \underbrace{\frac{1-\phi_1^{2h}}{1-\phi_1^2} \sigma^2}_{MSE}$

定常AR過程の予測

- AR(p)過程の場合

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t, \varepsilon_t \sim \mathrm{iid}N(0, \sigma^2)$$

  - 1期先予測

$$\hat{y}_{t+1|t} = c + \phi_1 y_t + \cdots + \phi_p y_{t-p+1}, \mathrm{MSE}(\hat{y}_{t+1|t}) = E(\varepsilon_{t+1}^2) = \sigma^2$$

  - この先を書くのは複雑でたいへん

- AR(p)過程の最適予測の性質
  - 最適予測は直近のp期のyの値にのみ依存する
  - 予測期間が長くなると、
    - 最適予測における過去の観測値の影響は指数的に減衰する
    - 最適予測は過程の期待値に収束する
  - 予測期間が長くなるにつれて、
    - MSEは単調に増大する
    - 過程の分散に収束する

反転可能MA過程の予測

- 無限個のyの観測値が利用可能だと考える

→AR(∞)として $y_t = \sum_{k=1}^{\infty} \eta_k y_{t-k} + \varepsilon_t$ と書ける

→ε_tの値は $\varepsilon_t = y_t - \sum_{k=1}^{\infty} \eta_k y_{t-k}$ として求まる

- MA(2)過程 $y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}, \varepsilon_t \sim \mathrm{iid}N(0, \sigma^2)$ の場合
  - 1期先最適予測

$y_{t+1} = \mu + \varepsilon_{t+1} + \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1}$

$\hat{y}_{t+1|t} = \mu + \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1}$　$MSE = \sigma^2$

  - 2期先最適予測

$\hat{y}_{t+2|t} = \mu + \theta_2 \varepsilon_t$　$MSE = (1+\theta_1^2)\sigma^2$.

反転可能MA過程の予測

- 3期先最適予測

  ry

- h(>3)期先最適予測

  $\hat{y}_{t+h|t} = \mu$ 　$MSE = (1+\theta_1^2+\theta_2^2)\sigma^2 = \gamma_0$

  先がよく分からない。

- 反転可能なMA(q)過程の最適予測の性質
  - q期までの最適予測はすべての観測値に依存する
  - q+1期以降の最適予測は過程の期待値と等しい
  - q期までのMSEは予測期間が長くなるにつれて単調に増加
  - q+1期以降のMSEは過程の分散と等しい

反転可能MA過程の予測

- 有限個のyしか観測できない場合はどうするか？
  - q+1期以降の予測は $\hat{y}_{t+q+1|t} = \mu, \mathrm{MSE}\left(\hat{y}_{t+q+1|t}\right) = \gamma_0$
  - q期以前は具体的にε_tの近似値を求めていく
  - たとえばMA(2)でε_0=0とすれば以下のような手順

$$\hat{\varepsilon}_1 = y_1 - \mu$$
$$\hat{\varepsilon}_2 = y_2 - \mu - \theta_1\hat{\varepsilon}_1$$
$$\hat{\varepsilon}_3 = y_3 - \mu - \theta_1\hat{\varepsilon}_2 - \theta_2\hat{\varepsilon}_1$$
$$\vdots$$

(定常)ARMA過程の予測

- たとえば、有限個のy、ARMA(1,1)、ε_0=ε_1=0とすると、

$$y_t = c + \phi_1 y_{t-1} + \varepsilon_t + \theta_1\varepsilon_{t-1}, \varepsilon_t \sim \mathrm{iid}N\left(0,\sigma^2\right)$$

- ここで、

$$\varepsilon_t = y_t - c - \phi_1 y_{t-1} - \theta_1\varepsilon_{t-1}$$

だから、

$$\hat{\varepsilon}_2 = y_2 - c - \phi_1 y_1$$
$$\hat{\varepsilon}_3 = y_3 - c - \phi_1 y_2 - \theta_1\hat{\varepsilon}_2$$
$$\hat{\varepsilon}_4 = y_4 - c - \phi_1 y_3 - \theta_1\hat{\varepsilon}_3$$

- これを使うと、最適予測

$$\hat{y}_{t+1|t} = c + \phi_1 y_t + \theta_1\hat{\varepsilon}_t$$
$$\mathrm{MSE}\left(\hat{y}_{t+1|t}\right) \to \sigma^2$$

【参考】応用・発展的な事項

- 枠組みの拡張
  - 平均や分散が時間変化する場合
  - 被説明変数が2変数以上の場合(ベクトル自己回帰)
  - 季節変動などが含まれる場合
  - 観測値ではなく潜在状態が時系列に従う場合(状態空間)
  - 時刻が連続値をとる場合(確率微分方程式)
- 推定方法・計算方法
  - 逐次推定・ベイズ推定
  - モンテカルロシミュレーション
- 応用分野
  - モデル特定
  - モデルスイッチング
  - データマイニング・異常検出