時系列データ

$y_t = \beta_0 + \beta_1 x_{t1} + . . . + \beta_k x_{tk} + u_t$

1. Basic Analysis

Ch.10　時系列データの基本

1. 時系列データの性質
2. 時系列回帰の例
3. OLS推定量の有限標本特性
4. 関数形、ダミー変数、指数
5. トレンドと季節性

10.1　時系列データの性質

◆時系列データの性質

- 観測値の時間的順序：任意の並べ替え不可
- 典型的な特徴：観測値の系列相関/非独立性
- 時系列データのランダム性をどう考えるべきか？
  - 経済変数（GDPや株価等）の値は不確定
    - したがって確率変数としてモデル化されるべき
  - 時系列は確率変数の連続（=確率過程）
  - ランダム性は母集団からのサンプリングに依らない
  - 「サンプル」：確率過程がとりうる多くの経路のうち、時系列の実現経路の1つ

10.1　時系列データの性質

◆例：米国のインフレと失業率1948-2003

- 右表は2系列のみ表示しているが、実際には時間的経路が同時観測されるような多くの変数が存在しうる
- 時系列分析は、ある変数の過去の変数および他の変数の現在と過去の値への依存性をモデル化することに焦点

$y_t = f(y_{t-1}, x_t, x_{t-1}, \cdots)$

TABLE 10.1 Partial Listing of Data on U.S. Inflation and Unemployment Rates, 1948–2003

Year	Inflation	Unemployment
1948	8.1	3.8
1949	−1.2	5.9
1950	1.3	5.3
1951	7.9	3.3
.	.	.
.	.	.
.	.	.
1998	1.6	4.5
1999	2.2	4.2
2000	3.4	4.0
2001	2.8	4.7
2002	1.6	5.8
2003	2.3	6.0

10.2　時系列回帰の例

◆10-2a 静学モデルStatic models

- 静学時系列モデルでは、変数の現在値は説明変数の現在値の結果としてモデル化
- 静学モデルの例
  - $inf_t = \beta_0 + \beta_1 unem_t + u_t$
    - 同時点での失業とインフレの関係（=Phillips curve）
  - $mrdrte_t = \beta_0 + \beta_1 convrte_t + \beta_2 unem_t + \beta_3 yngmle_t + u_t$
    - **現在**の殺人率は、**現在**の有罪判決率や失業率、人口における若年男性の割合により決定

10.2　時系列回帰の例

◆10-2b 有限分布ラグ（FDL）モデル

- 有限分布ラグモデルでは、説明変数はタイムラグを伴い被説明変数に影響を及ぼすことを許容
- 有限分布ラグモデルの例
  - 出生率は子供の税額に依存しうるが、生物学的および行動上の理由から、効果には遅れが生じうる

$$gfr_t = \alpha_0 + \delta_0 \boxed{pe_t} + \delta_1 \boxed{pe_{t-1}} + \delta_2 \boxed{pe_{t-2}} + u_t$$

t期の女性千人当り出産数

t期の税控除

t - 1期の税控除

t - 2期の税控除

10.2 時系列回帰の例

- 有限分布ラグモデルの解釈

  $y_t = \alpha_0 + \delta_0 z_t + \delta_1 z_{t-1} + \ldots + \delta_q z_{t-q} + u_t$

  - 過去のショックによるyの現在値への効果
    - 一時的ショックtransitory shockの影響

      $\boxed{\frac{\Delta y_t}{\Delta z_{t-s}} = \delta_s}$ 過去に1回のショックがあった場合、被説明変数は対応する遅れの係数によって示される量だけ一時的に変化

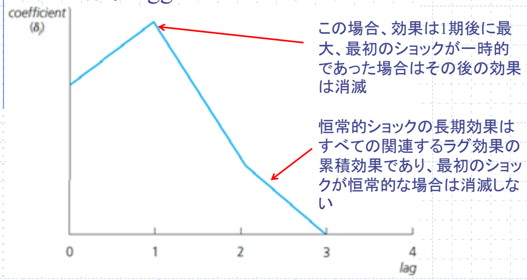
    - 恒常的ショックpermanent shockの影響

      $\boxed{\frac{\Delta y_t}{\Delta z_{t-q}} + \ldots + \frac{\Delta y_t}{\Delta z_t} = \delta_0 + \ldots + \delta_q}$

      - 過去に恒常的ショックがある場合（例：説明変数が恒常的に1単位増加）、従属変数への影響は関連するすべてのラグの累積効果、つまり従属変数に対する長期的な影響

10.2 時系列回帰の例

- ラグ効果lagged effects



10.3 OLSの有限標本特性

- 10-3a OLSの不偏性
  - 仮定TS.1 (Linear in parameters)

    $y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \ldots + \beta_k x_{tk} + u_t$

    - 関係する時系列は線形関係に従う
    - 確率過程y_t, x_{t1},..., x_{tk}が観測的、誤差過程u_tは非観測的
    - 説明変数の定義は一般的で、ラグまたは他の説明変数の関数であってもよい
  - 仮定TS.2 (No perfect collinearity)
    - 標本（したがって、基礎的な時系列過程）では、説明変数は一定でも他の変数との完全な線形結合でもない

10.3 OLSの有限標本特性

- 仮定TS.3 (Zero conditional mean)

  $E(u_t|\mathrm{X}) = 0, t = 1, 2, \ldots, n$

  - 観測されなかった要因の平均値は、**すべての期間における**説明変数の値と無相関
  - Notation

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ \vdots & \vdots & & \vdots \\ x_{t1} & x_{t2} & \cdots & x_{tk} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix}$$

期間tの全説明変数の値

  - この行列は全説明変数の完全な時間経路に関するすべての情報を収集

10.3 OLSの有限標本特性

- 仮定TS.3について
  - 強外生性Strict exogeneity　$E(u_t|\mathbf{X}) = 0$
    - 誤差項の平均値は**全期間**の説明変数の値と無相関
    - TS.3は被説明変数y_t ($\leftarrow u_t$)から説明変数の将来値x_{t+1}へのフィードバックを排除
      - 被説明変数の過去の変化に対応して、説明変数を「調整」している場合には疑問（例：y-殺人率、x-警官数）
  - 同時外生性Contemporaneously exogeneity　$E(u_t|\mathbf{x}_t) = 0$
    - 誤差項の平均値は**同じ期間**の説明変数と無相関
    - 誤差項が説明変数の過去の値に相関する場合、これらの値を同時外生性変数として含める

10.3 OLSの有限標本特性

- 定理10.1 (Unbiasedness of OLS)

  $TS.1 - TS.3 \quad \Rightarrow \quad E(\hat{\beta}_j) = \beta_j, \quad j = 0, 1, \ldots, k$

- 10-3b 仮定TS.4 (Homoskedasticity)

  $Var(u_t|\mathrm{X}) = Var(u_t) = \sigma^2, t = 1, 2, \ldots, n$

  - TS.4の十分条件：誤差項の分散が説明変数と独立的&経時的に一定
    - 誤差項の分散は**全期間**の説明変数と相関してはいけない
  - 時系列の文脈では均一分散は容易に抵触する可能性
    - 例：従属変数を国債金利とすると、その（欠落要因の）変動は政策変更（→財政赤字&インフレに影響）に依存→[10.11]

10.3 OLSの有限標本特性

◆仮定TS.5 (No serial correlation)

$Corr(u_t, u_s|\mathbf{X}) = 0,\ t \neq s$

- Xの条件下で観測できない要因は経時的に相関しない

- 仮定TS.5について
  - 任意の説明変数値の下で、欠落要因が経時的に相関する場合、この仮定は成立しない
  - なぜこの仮定が横断的データにはないのか？
  - この仮定は、横断面抽出が完全に無作為でない場合の無作為抽出仮定の代用としても役立ち得る
    - この場合、任意の説明変数値の下で、誤差項は横断面ユニット（例：州）において無相関である必要

10.3 OLSの有限標本特性

◆定理10.2 (OLS sampling variances)

- 仮定TS.1 - TS.5の下で

$$Var(\hat{\beta}_j|\mathbf{X}) = \frac{\sigma^2}{SST_j(1-R_j^2)}, \quad j = 1, \ldots, k$$

  - 横断面データのケースと同じ式
    - 説明変数値Xの条件の理解は容易ではない
    - これは主に、有限標本では説明変数の無作為性による抽出変動を無視しうることを意味
    - この種の抽出変動は（合計のため）通常大きくない

◆定理10.3（誤差分散の不偏推定）

$$TS.1 - TS.5 \quad \Rightarrow \quad E(\hat{\sigma}^2) = \sigma^2$$

10.3 OLSの有限標本特性

◆定理10.4 (Gauss-Markov Theorem)

- 仮定TS.1-TS.5の下で、OLS推定量はすべての線形不偏推定量のうち最小分散推定量
  - 説明変数の条件の有無に関わらず有効

◆10-3c 仮定TS.6 (Normality)

- Xと独立のもとで（TS.3-TS.5より） $u_t \sim \text{Normal}(0, \sigma^2)$

◆定理10.5 (Normal sampling distributions)

- 仮定TS.1-TS.6のもとで、OLS推定量はXの条件下で正規分布を持つ→F & t 検定も有効

10.3 OLSの有限標本特性

◆例10.1：Static Phillips curve

$$\widehat{inf}_t = \underset{(1.72)}{1.42} + \underset{(.289)}{.468}\ unem_t \qquad n = 49, R^2 = .053, \bar{R}^2 = .033$$

- 理論に反して、推定されたフィリップスカーブはインフレと失業のトレードオフを示していない

- CLM仮定について
  - TS.1： $inf_t = \beta_0 + \beta_1 unem_t + u_t$
    - 誤差項は、通貨、所得/需要、原油価格、供給、為替の各ショック等の要因を含む
    - 線形関係は制約的だが、良い近似である必要
  - TS.2：
    - 失業率が経時的に変化する限り、完全共線性問題はない

10.3 OLSの有限標本特性

- CLM仮定について（続き）
  - TS.3： $E(u_t|unem_1, \ldots, unem_n) = 0$
    - 反証例1： $unem_{t-1} \uparrow \rightarrow u_t \downarrow$
      過去の失業が将来需要に影響し、インフレを弱める可能性
    - 反証例2： $u_{t-1} \uparrow \rightarrow unem_t \uparrow$
      原油価格変化がインフレを招き、将来の失業率上昇に影響
  - TS.4： $Var(u_t|unem_1, \ldots, unem_n) = \sigma^2$
    - 高失業率時に金融政策がより「緊張」するなら、不成立
  - TS.5： $Corr(u_t, u_s|unem_1, \ldots, unem_n) = 0$
    - インフレ率の影響が経時的に持続するなら、不成立
  - TS.6： $u_t \sim \text{Normal}(0, \sigma^2)$
    - 仮定の妥当性は疑わしい

10.3 OLSの有限標本特性

◆例10.2：インフレ&財政赤字の国債への影響

$$\widehat{i3}_t = \underset{(0.43)}{1.73} + \underset{(.082)}{.606}\ inf_t + \underset{(.118)}{.513}\ def_t$$

3ヶ月米国債金利　政府赤字（対GDP）

$$n = 56, R^2 = .602, \bar{R}^2 = .587$$

- CLM仮定について
  - TS.1： $i3_t = \beta_0 + \beta_1 inf_t + \beta_2 def_t + u_t$
    - 誤差項は一般に金利を決定する他の要因を表す（例：ビジネスサイクル効果等）
    - 近似的に線形関係で表現
  - TS.2：
    - 完全共線性は実際にはほぼ問題なし

10.3 OLSの有限標本特性

- CLM仮定について（続き）
  - TS.3： $E(u_t|inf_1,\ldots,inf_n,def_1,\ldots,def_n)=0$
    - 反証例1： $def_{t-1}\uparrow\rightarrow\ u_t\uparrow$
      過去の財政支出が経済を浮揚させ、金利を高める可能性
    - 反証例2： $u_{t-1}\uparrow\rightarrow\ inf_t\uparrow$
      需要ショックが金利を上昇させ、将来のインフレ上昇に影響
  - TS.4：$Var(u_t|inf_1,\ldots,def_n)=\sigma^2$
    - 赤字増大が財政不安と急激な金利変動を招くなら、不成立
  - TS.5：$Corr(u_t,u_s|inf_1,\ldots,def_n)=0$
    - 景気循環の影響が続くならば、不成立
  - TS.6： $u_t\sim\text{Normal}(0,\sigma^2)$
    - 仮定の妥当性は疑わしい

10.4 関数形、ダミー変数、指数

◆例10.4 時系列でのダミー変数

t年の合計特殊出生率　年tの税控除　第2次世界大戦期(1941-45)　避妊薬導入(1963-)

$$\widehat{gfrt}=98.68+.083\,pe_t-24.24\,\boxed{ww2_t}-31.59\,\boxed{pill_t}$$
$$(3.21)\quad(.030)\quad(7.46)\quad(4.08)$$
$$n=72,R^2=.473,\bar{R}^2=.450$$

- 解釈
  - 第2次世界大戦中、出生率は一時的に低い
  - 1963年避妊薬の導入以来、恒常的に低い

10.5 トレンドと季節性

◆10-5a 線形時間トレンドのモデル化

$y_t=\alpha_0+\alpha_1 t+e_t\quad\Leftrightarrow\quad E(\Delta y_t)=E(y_t-y_{t-1})=\alpha_1$

- $\Delta y_t/\Delta t=\alpha_1$
  - 他の条件を一定として、従属変数は時間単位ごとに一定量だけ増加
- $E(y_t)=\alpha_0+\alpha_1 t$
  - 他の解釈として、従属変数の期待値は時間の線形関数

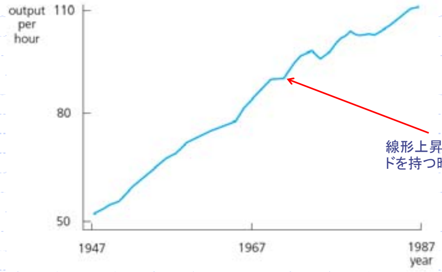
◆指数時間トレンドのモデル化

$\log(y_t)=\alpha_0+\alpha_1 t+e_t\quad\Leftrightarrow\quad E(\Delta\log(y_t))=\alpha_1$

- $(\Delta y_t/y_t)/\Delta t=\alpha_1$
  - 従属変数は時間単位ごとに一定の割合で増加

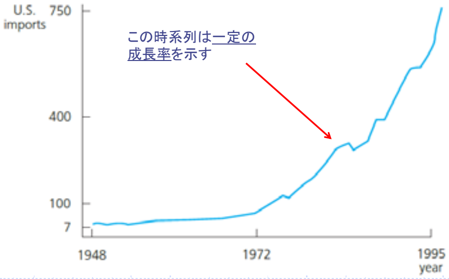
10.5 トレンドと季節性

◆トレンドを持つ時系列



10.5 トレンドと季節性

◆図10.3 指数トレンドを持つ時系列



10.5 トレンドと季節性

◆10-5b 回帰分析でのトレンド変数

- 独立&従属変数が共通トレンドから成る場合、偽の関係***spurious relationship***が生じる可能性
  - この場合、回帰にトレンドを含めることが重要
- 例10.7：住宅投資と価格

$$\widehat{\log}(invpc)=-\ .550\ +\ 1.241\ \log(price)$$
$$(.043)\qquad(.382)$$

1人当り住宅投資　住宅価格指数

$$n=42,R^2=.208,\bar{R}^2=.189$$

- この回帰結果からは、一見、住宅投資と価格の間に、正の関係性があるように見える

10.5 トレンドと季節性

- 例10.7：住宅投資と価格（続き）

$$\widehat{\log(invpc)} = \underset{(1.36)}{-.913} - \underset{(.679)}{.381}\log(price) + \boxed{\underset{(.0035)}{.0098\,t}}$$

$$n = 42,\ R^2 = .341,\ \bar{R}^2 = .307$$

  - 住宅投資と価格に有意な関係はない
- トレンドはいつ含めるべきか？
  - 従属変数が明らかなトレンドを示す場合
  - 従属変数と独立変数の両方にトレンドがある場合
  - 独立変数のいくつかだけが傾向を持っている場合
    - 独立変数の従属変数への（真の）影響は、トレンドが除去された後にのみ表示

10.5 トレンドと季節性

◆10-5c 時間トレンドを伴う回帰の解釈
- トレンドを含む回帰におけるOLS係数は、事前にトレンド除去した回帰係数と同じ
  - これは多重回帰の一般的解釈

◆10-5d 従属変数がトレンドを持つR^2
- トレンド除去後の回帰のR^2がより適切
  - トレンドにより従属変数の変動は過大になる
  - トレンドを除去した従属変数を全独立変数
    - 説明変数がトレンドを持つならば、トレンド変数も加えて回帰

10.5 トレンドと季節性

◆10-5e 時系列での季節性のモデル化
- 季節ダミーの付加

$$y_t = \beta_0 + \delta_1 \boxed{feb_t} + \delta_2 \boxed{mar_t} + \delta_3 \boxed{apr_t} + \ldots + \delta_{11} \boxed{dec_t} + \beta_1 x_{t1} + \beta_2 x_{t2} + \ldots + \beta_k x_{tk} + u_t$$

12月であれば1
それ以外は0

- 時間トレンドの場合と同様の解釈
  - 説明変数の回帰係数は、従属変数と説明変数の季節性を除去した結果として見ることができる
  - 従属変数の季節性を除去した回帰のR^2は、説明変数の説明力をより反映する可能性