

インサイト・ファクトリー 社内セミナー
2020年4月

マーケティング・ミックス・モデリング

V. 時系列分析の基礎

小野 滋

統計学・データ解析	マーケティング・ミックス・モデリングへの適用
I. イントロダクション	
	II. 市場反応モデルとは
III. 回帰分析の基礎	
	IV. 静学的市場反応モデル
V. 時系列分析の基礎	
	VI. 動学的市場反応モデル(1)
VII. 状態空間モデルの基礎	
	VIII. 動学的市場反応モデル(2)

目次

この章の内容

1. なぜ時系列データを単純に回帰分析してはいけないのか
2. 基本的な統計量
3. 基本的な概念
4. 基本的な確率過程 I. 定常過程
5. 基本的な確率過程 II. 非定常過程
6. 1変量時系列のモデル
7. 時系列の回帰モデル
8. 季節変動

この章の引用文献

この章に登場したRの関数

Appendix 1. Rの日付クラス

Appendix 2. Rの時系列クラス

時系列データをモデル化する基礎的な方法について紹介します

本資料作成に用いたすべてのRコードを、以下で公開しています:
https://rpubs.com/shig_ono/MMM_5

1.なぜ時系列データを単純に回帰分析してはいけないのか

1.なぜ時系列データを単純に回帰分析してはいけないのか

Q: 下記のような時系列データがあります。

これまでの経験から、「ある月の売上は、その月の購入意向に影響され、それ以前の購入意向には影響されない」ことがわかっています。

購入意向が1ポイント上がったら、売上がどれだけ上がるかを知りたいです。

そこで、売上を目的変数、購入意向を説明変数として、単回帰分析をしようと思います。

	売上金額	購入意向
2018年1月	xxx	xxx
2018年2月	xxx	xxx
2018年3月	xxx	xxx
2018年4月	xxx	xxx
...		



A: だめです

このデータは、回帰分析の前提を満たしていません

回帰モデル

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_K X_{Ki} + U_i$$

[1] X_i は確率変数でないか
 U_i と統計的に独立

[2] $E(U_i) = 0$

[3] $\text{Var}(U_i) = \sigma^2$

[4] $\text{Cov}(U_i, U_j) = 0, i \neq j$

[5] U_i は正規分布に従う

[6] X の列ベクトルは一次独立。
列数は行数より小さい

市場反応の特性

時系列データ

- データは隣り合う2時点の間に相関を持つかもしれない→ [4]に反する
- 攪乱項の平均や分散が一定でないかもしれない→ [2][3]に反する
- ある期の購入意向は、その期よりも後の期の売上金額に効果を持つかもしれない→ [1]に反する

Q: なるほど、厳密に言えばだめなんですね。わかりました。

でも、科学研究とかじゃなくて、ビジネスですから... 統計学の専門家の方からみて
少しくらいまちがっていても、実務に役に立つことが大事ですから...

とりあえず試してみようと思います！

	売上金額	購入意向
2018年1月	xxx	xxx
2018年2月	xxx	xxx
2018年3月	xxx	xxx
2018年4月	xxx	xxx
...		



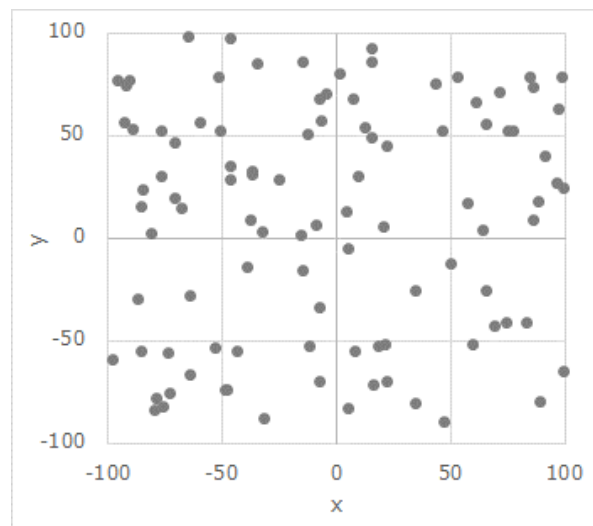
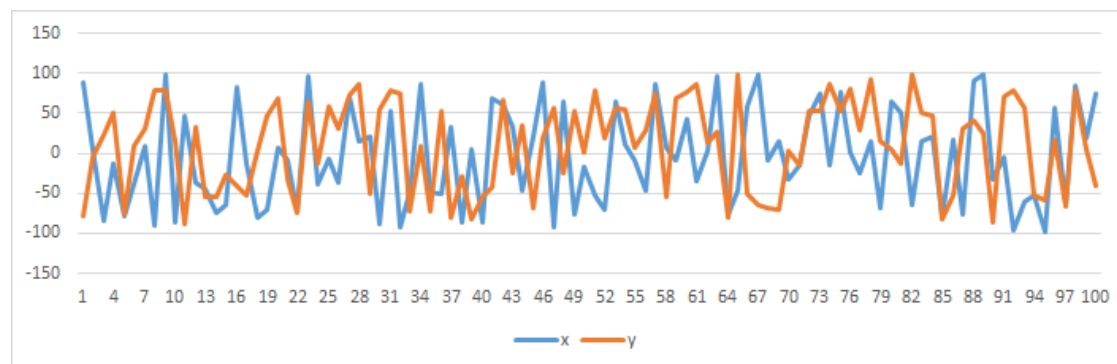
A: 絶対にだめです

時系列データの単純な回帰分析がいかに危険か、実例をごらんいただきましょう

例1) ランダムな変数の相関

-100~+100の一樣乱数

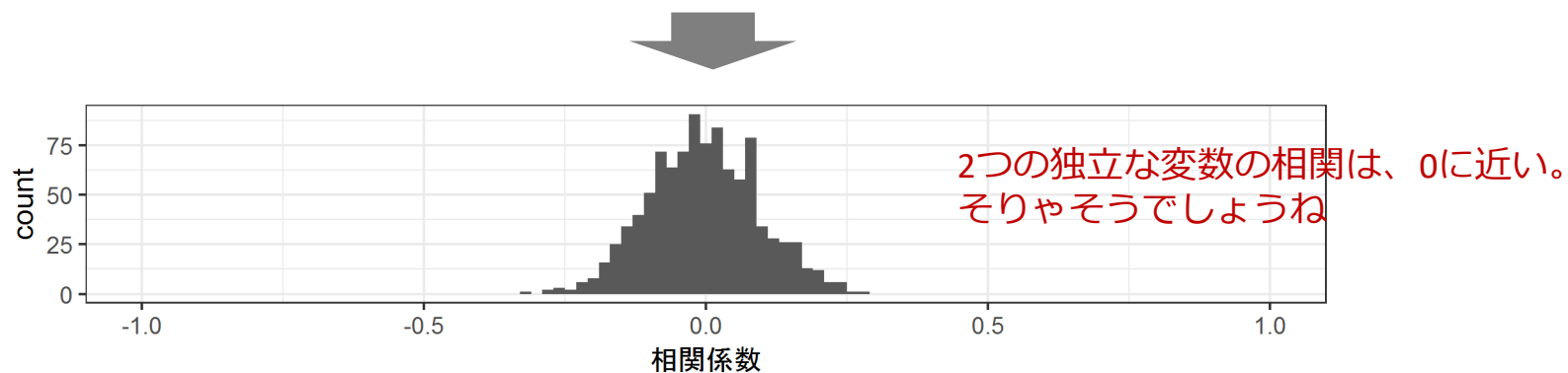
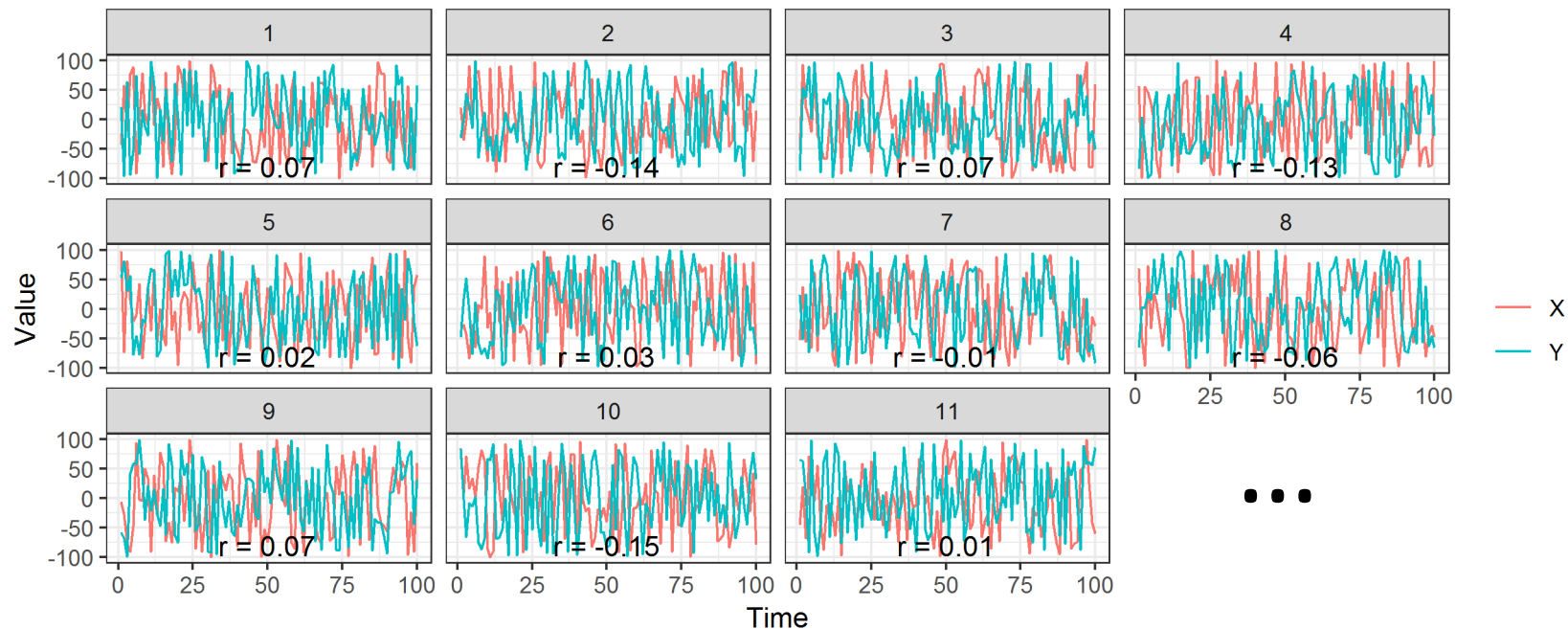
	x	y
1	89	-79
2	5	-4
3	-85	24
4	-13	51
5	-79	-77
6	-38	10
7	9	31
8	-91	78
9	98	79
10	-86	16
11	47	-89
12	-37	33
13	-44	-54
中略		
99	20	6
100	74	-40



相関: 0.07

1000回繰り返してみました

Code 1



例2) ランダムに変動する時系列の相関

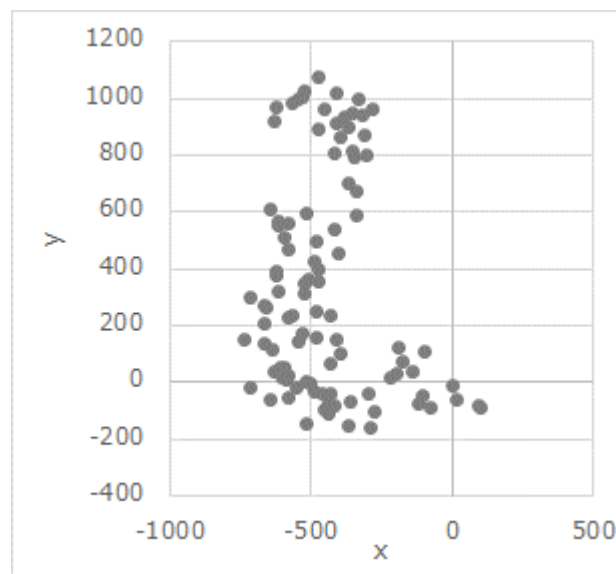
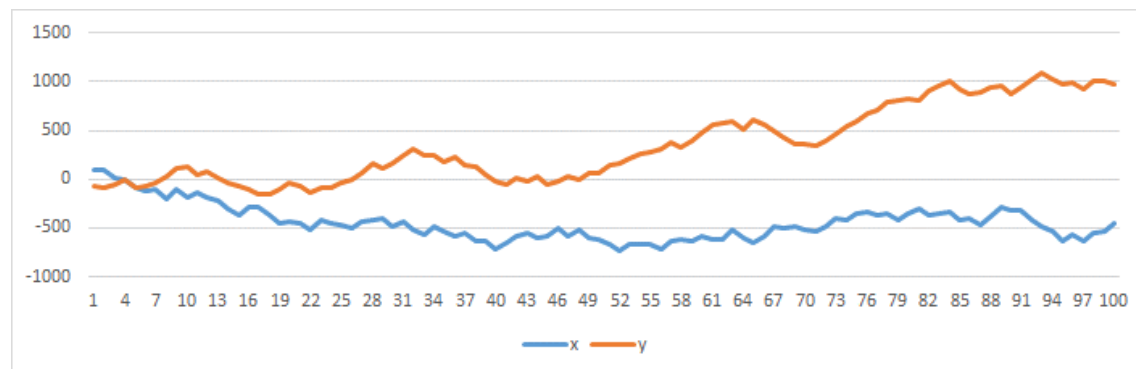
例1と同じ値

前期から Δx , Δy 変化する時系列

	Δx	Δy	x	y
1	89	-79	89	-79
2	5	-4	94	-83
3	-85	24	9	-59
4	-13	51	-4	-8
5	-79	-77	-83	-85
6	-38	10	-121	-75
7	9	31	-112	-44
8	-91	78	-203	34
9	98	79	-105	113
10	-86	16	-191	129
11	47	-89	-144	40
12	-37	33	-181	73
13	-44	-54	-225	19

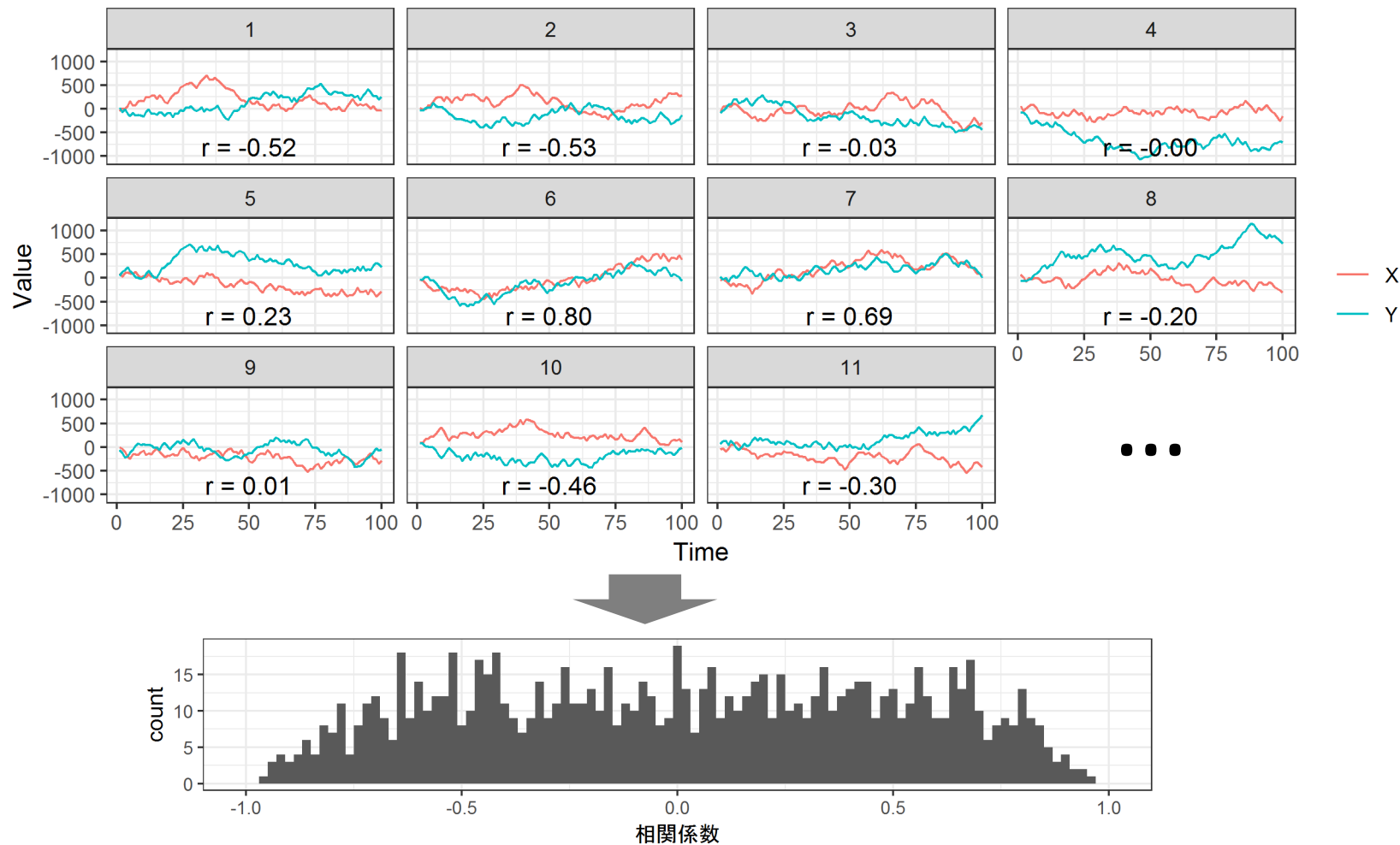
中略

99	20	6	-532	1007
100	74	-40	-458	967



相関: -0.15

1000回繰り返してみました



2つの時系列は、たとえ独立に変動していても、相関が0に近くなるとは限らない。
時系列データの分析には、時系列データのための分析手法が必要！

2. 基本的な統計量

2. 基本的な統計量

本章以降の説明で必要になる、基本的な統計量を定義する。

■ 期待値 (平均) $\mu_X = E(X_i)$

■ 分散 $\text{Var}(X_i) = E[(X_i - \mu_X)^2]$

■ 共分散 $\text{Cov}(X_i, Y_i) = E[(X_i - \mu_X)(Y_i - \mu_Y)]$

■ 相関
$$\text{Corr}(X_i, Y_i) = \frac{\text{Cov}(X_i, Y_i)}{\sqrt{\text{Var}(X_i)\text{Var}(Y_i)}}$$

(確率変数)



X_1 Y_1



X_2 Y_2

$\vdots$

$\vdots$

$\vdots$

標本統計量：

■ 平均

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

■ 分散

$$s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

■ 共分散

$$\text{cov}_{x,y} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

■ 相関

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

(観測値)



x_1 y_1



x_2 y_2

⋮

⋮

⋮



x_n y_n

■ 期待値 (平均)

$$\mu_t = E(Y_t)$$

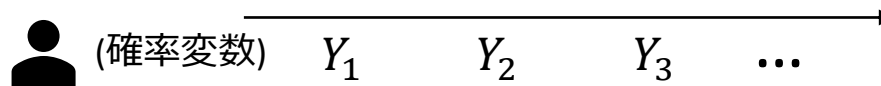
分散は γ_{0t}

■ k 次の自己共分散

$$\gamma_{kt} = Cov(Y_t, Y_{t-k}) = E[(Y_t - \mu_t)(Y_{t-k} - \mu_{t-k})]$$

■ k 次の自己相関

$$\rho_{kt} = Corr(Y_t, Y_{t-k}) = \frac{\gamma_{kt}}{\sqrt{\gamma_{0t} \times \gamma_{0,t-k}}}$$



標本統計量：

■ 標本平均

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

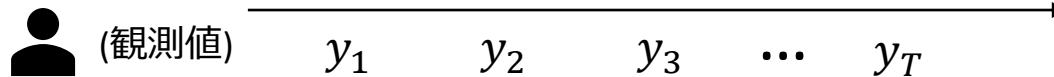
■ k 次の標本自己共分散

$$\hat{\gamma}_k = \frac{1}{T} \sum_{t=k+1}^T (y_t - \bar{y})(y_{t-k} - \bar{y}), \quad k = 0, 1, 2, \dots$$

長さTの時系列のk次の標本自己共分散を求める際、2時点の(T-k)個のペアについて偏差の積を求め、これを足し上げ、(T-kではなく)Tで割る、という点に注意

■ k 次の標本自己相関

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}$$



■ 1次の標本自己共分散・標本自己相関を算出する例

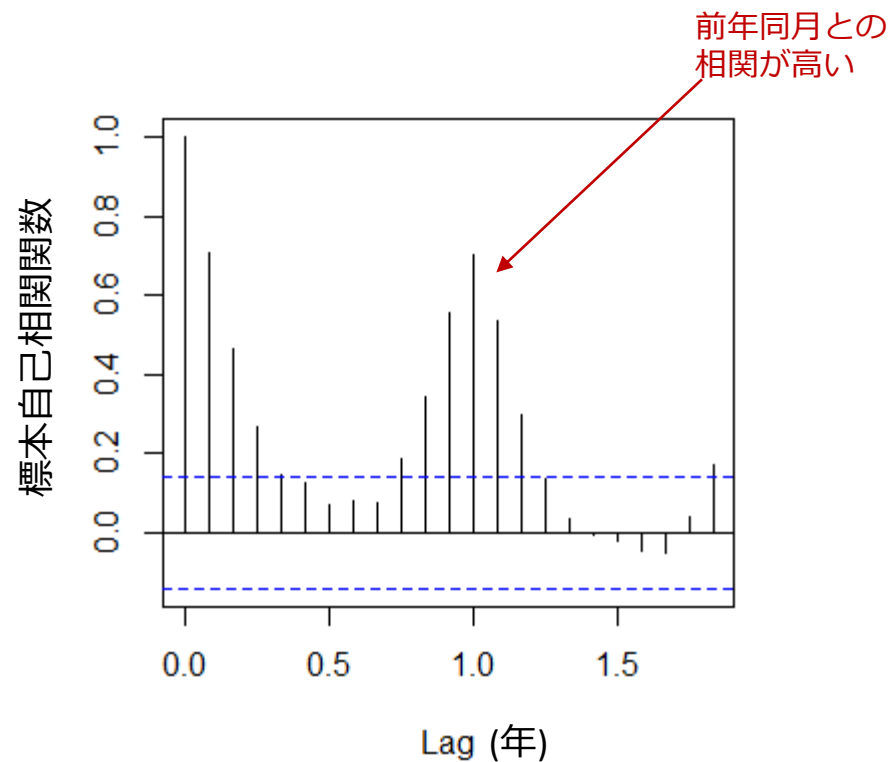
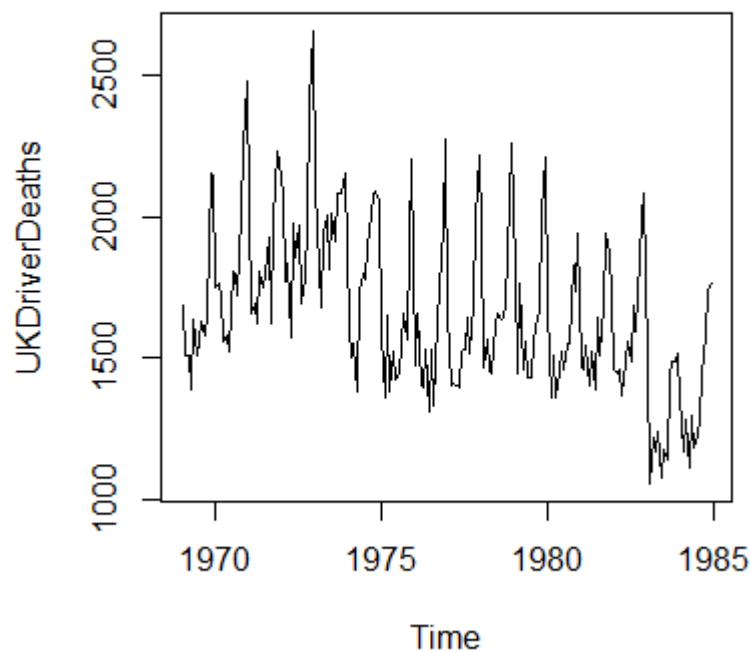
値			平均からの差		平均からの差の積	
時点	当期	1期前		当期	1期前	
1	1.41	1.48	→	1.17	1.25	→ 1.47
2	1.48	1.72	→	1.25	1.48	→ 1.85
3	1.72	1.04	→	1.48	0.81	→ 1.19
4	1.04	0.60	→	0.81	0.37	→ 0.30
5	0.60	-0.47	→	0.37	-0.71	→ -0.26
6	-0.47	-1.50	→	-0.71	-1.73	→ 1.23
7	-1.50	-1.05	→	-1.73	-1.28	→ 2.22
8	-1.05	-0.49	→	-1.28	-0.73	→ 0.93
9	-0.49	-0.39	→	-0.73	-0.63	→ 0.46
10	-0.39			-0.63		
合計				合計		
その1/10				その1/10		
↓				↓		
2.36				11.99		
0.24				1.20		
標本平均				標本分散		
				(0次の標本自己共分散)		
				合計		
				9.39		
				その1/10		
				0.94		
				1次の標本自己共分散		
				↓		
				0.78		
				1次の標本自己相関		

■ 標本自己相関関数 (標本コレログラム)

Code 2

k次の標本自己相関をkの関数と捉えたもの

英国自動車事故



英国の自動車運転者の死亡数(1969-84, 月次)

3. 基本的な概念

3. 基本的な概念

本章以降の説明で必要になる概念を紹介する。

1. 外見上の規則性 ... トレンド・循環変動・季節変動
2. ラグ演算子
3. 確率過程
4. 定常性

- **トレンド**

長期間にわたる一定方向への変化 (増大 or 減少)

- **確定的トレンド**

- 前の時点からの増分が確定的に決まっている

- **確率的トレンド**

- 前の時点からの増分が確率的に決まっている

} (後述)

- **季節変動**

- **循環変動**

- **不規則変動**

■ ラグ演算子

確率過程を簡単に表現するテクニック

$$LY_t = Y_{t-1}$$

「 L ほにゃらら」は
「1期前のほにゃらら」

$$L^k Y_t = Y_{t-k}$$

「 L^2 ほにゃらら」は
「2期前のほにゃらら」

$$L^0 Y_t = Y_t$$

「 L^0 ほにゃらら」は
「ほにゃらら」そのもの

$$L^k c = c$$

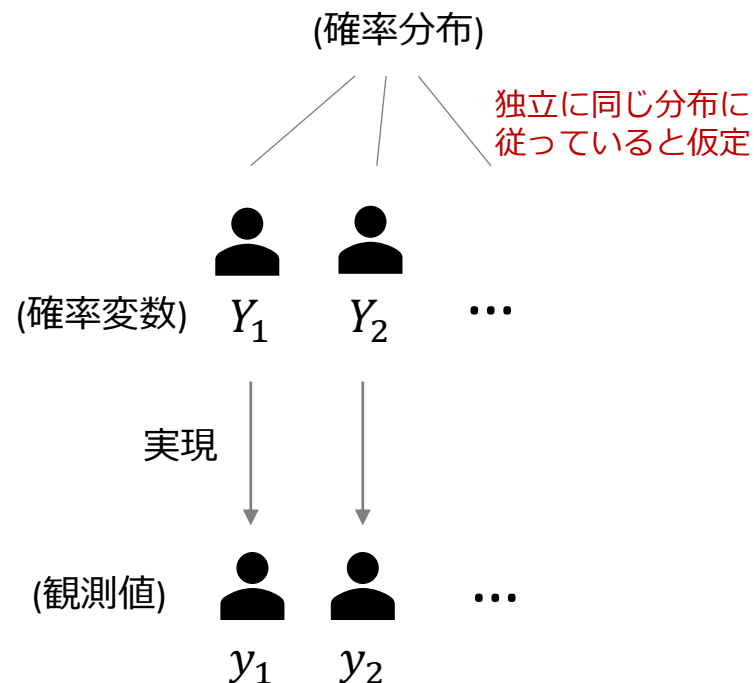
ラグ演算子は
定数に対しては無力

この研修資料では、ラグ演算子を使った数式を青色で表示します

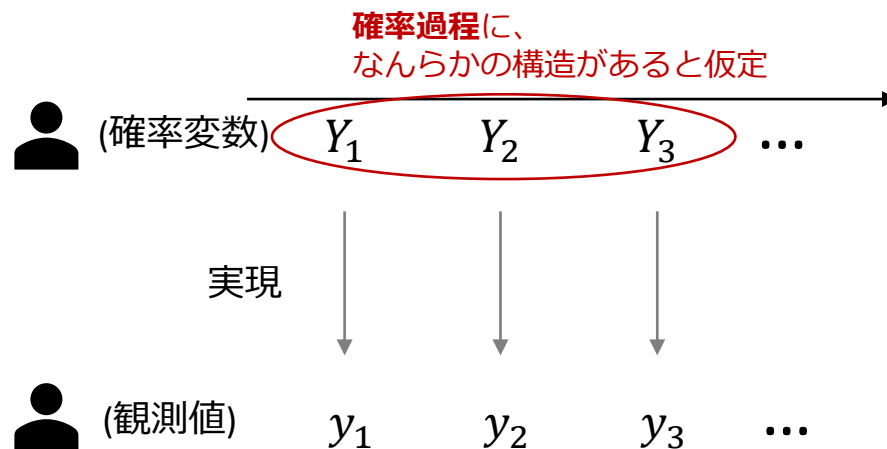
■ 確率過程

- データの背後にある確率変数の列のこと
- 以下では単に過程と呼ぶことが多い
 - データ生成過程 (DGP) の略と捉えてもよい

クロスセクション・データに対する考え方



時系列データに対する考え方



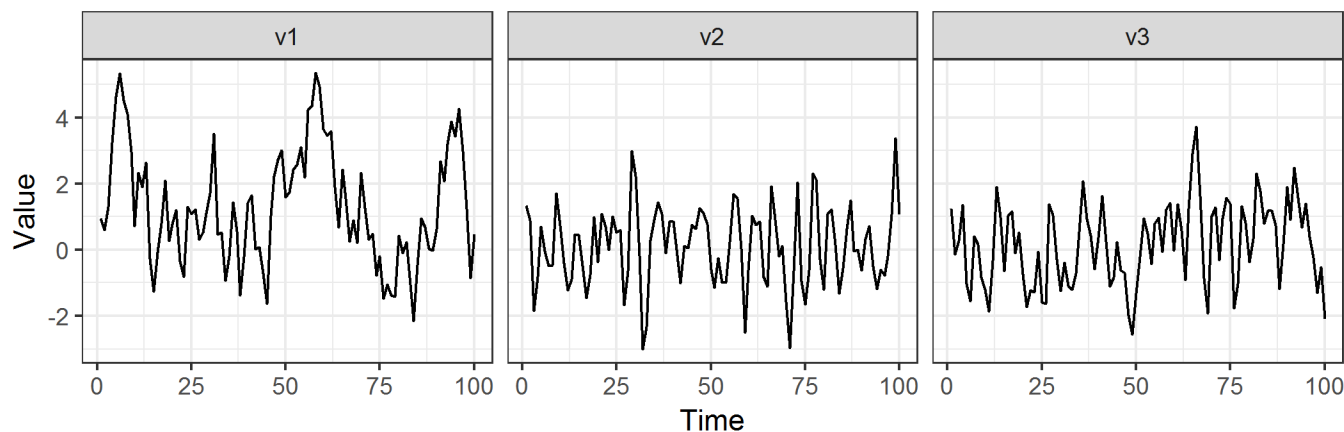
■ 定常性 (弱定常性)

- 確率過程の性質のひとつ。時系列分析にとって非常に重要な概念
- 次の2つが成り立つこと
 - 任意の t について $E(Y_t) = \mu$
 - 任意の t と k について $\text{Cov}(Y_t, Y_{t-k}) = \gamma_k$
- いいかえれば、
 - 期待値は時間を通じて一定
 - 2時点間の自己共分散・自己相関は、時間差が同じなら一定

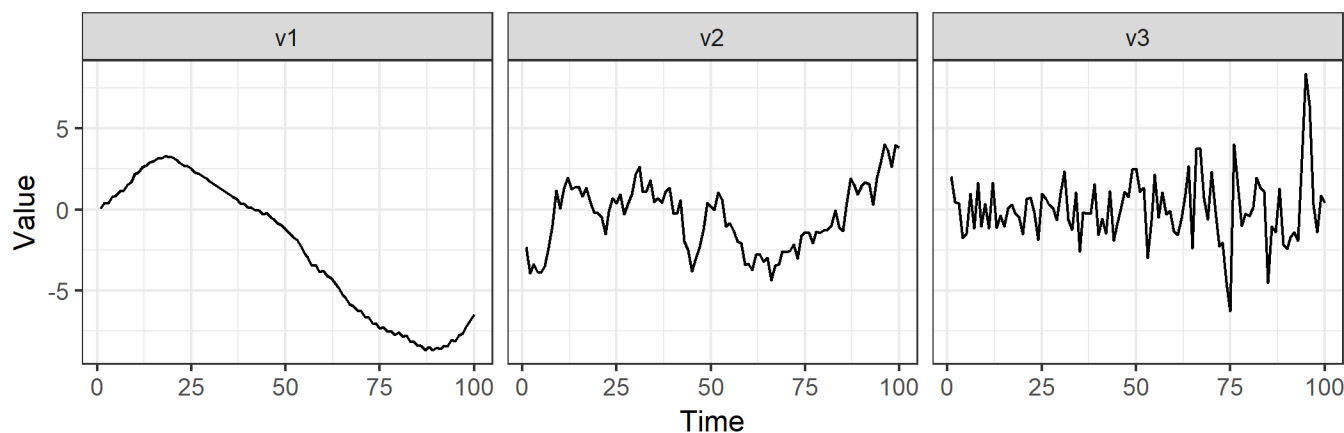
■ 強定常性

- 任意の t と k に対して $(Y_t, Y_{t+1}, \dots, Y_{t+k})'$ の同時分布が同一であること
- いいかえれば、
 - 2時点間の同時分布は、時間差が同じなら一定

背後にある過程が定常である(ような気がする)時系列データの例



背後にある過程が非定常である(ような気がする)時系列データの例



Q. 伝統的な時系列分析では、確率過程の定常性を仮定した枠組みのなかで時系列データを分析するのがふつうです。
しかし、実際の売上やマーケティング・ミックス変数は、ふつう時間を通じて一定ではありません。
定常性の仮定は現実的ではないと思うのですが...？

- A. たしかに、実際に観察される時系列は、定常的とはいいいがたいことが多いでしょう。しかし、**定常過程に基づく分析枠組みは、幅広い適用可能性を持っています**。なぜなら、
- 非定常的な挙動を示している時系列データも、適切な変換によって定常的な時系列データへと変換できることが多いです。
 - モデル構築を通じて、非定常的な時系列データをうまく扱えることがあります。

4. 基本的な確率過程 I. 定常過程

4. 基本的な確率過程 I. 定常過程

ここからは、以降の議論のために必要となる基本的な確率過程を導入する。
まず、いくつかの定常な確率過程を導入しよう。

4-1. ホワイトノイズ

4-2. ARMA過程

- 1) AR過程
- 2) MA過程
- 3) AR過程とMA過程の関係
- 4) ARMA過程

時系列モデルにおいて確率的変動を表現するために、よく用いられる確率過程。

■ ホワイトノイズ：

すべての t の下で以下の性質を持つ確率変数列のこと。

$$E(\varepsilon_t) = 0$$

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

つまり、どの時点でも期待値0、分散 σ^2 、自己相関0。もちろん定常。

■ 正規ホワイトノイズ：

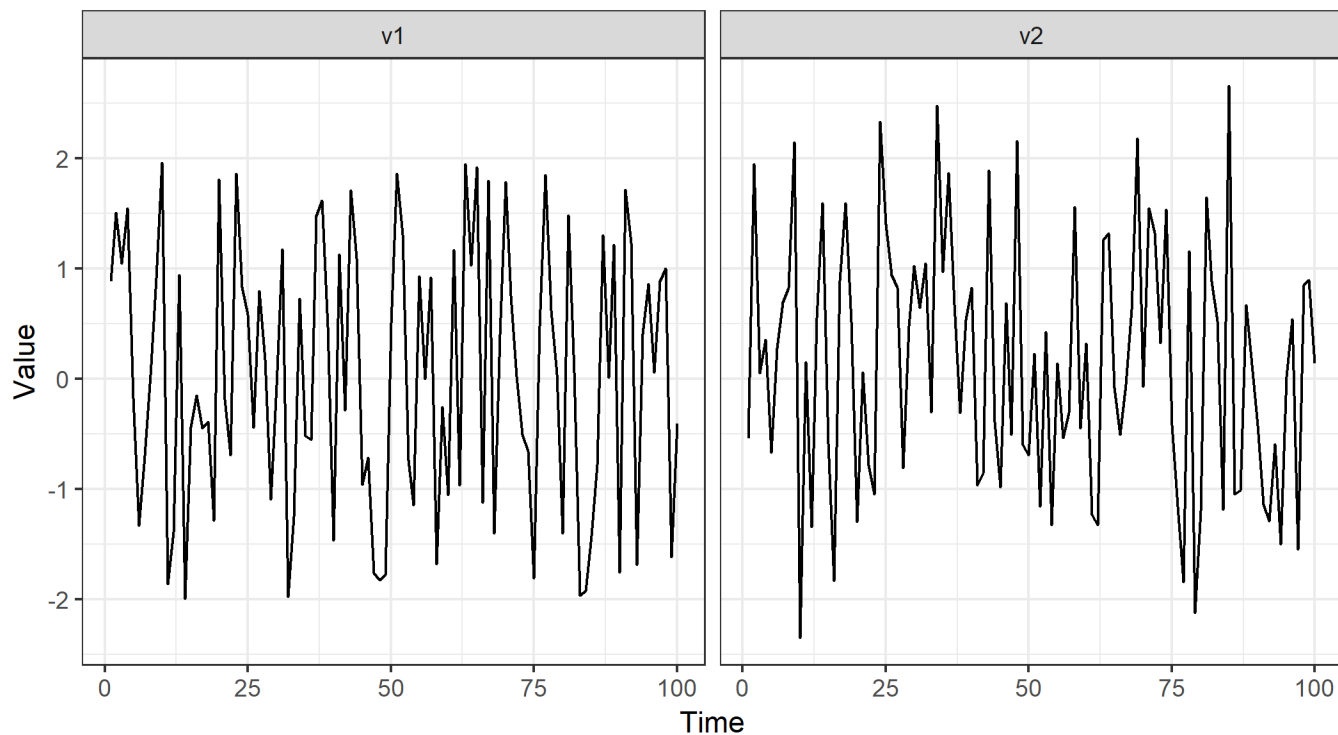
ホワイトノイズの特殊ケース。

$$\varepsilon_t \sim \text{iid } N(0, \sigma^2)$$

つまり、どの時点でも独立に同じ正規分布に従う。もちろん強定常。

ホワイトノイズに従う時系列の例

Code 4



$$\varepsilon_t \sim \text{iid Unif}(-2, +2)$$

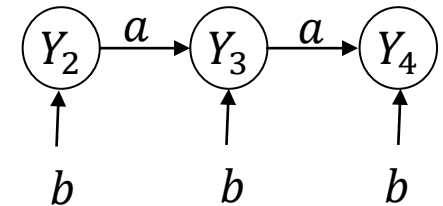
$$\varepsilon_t \sim \text{iid } N(0, 1)$$

自己相関を持つ時系列をモデル化する際、最も良く用いられる確率過程。

■ 自己相関をモデル化するには？

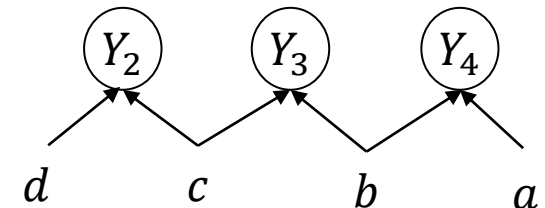
- ある時点のモデルに前の時点を含める → 自己回帰過程 (**AR過程**)

$$Y_t = aY_{t-1} + b$$



- 隣り合う時点のモデルに共通の成分を含める → 移動平均過程 (**MA過程**)

$$\begin{cases} Y_t = a + b \\ Y_{t-1} = b + c \end{cases}$$

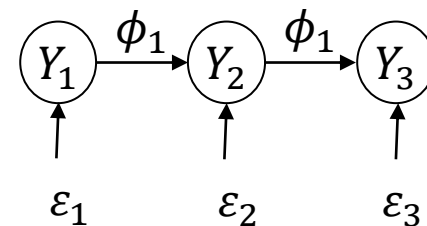


1) AR過程 (沖本, pp.26-34)

■ 1次の自己回帰過程 (AR(1)過程) :

$$Y_t = c + \phi_1 Y_{t-1} + \varepsilon_t$$

ただし、 ε_t はホワイトノイズ (分散 σ^2)。



- AR(1)に従う時系列を、Excelでつくってみよう！

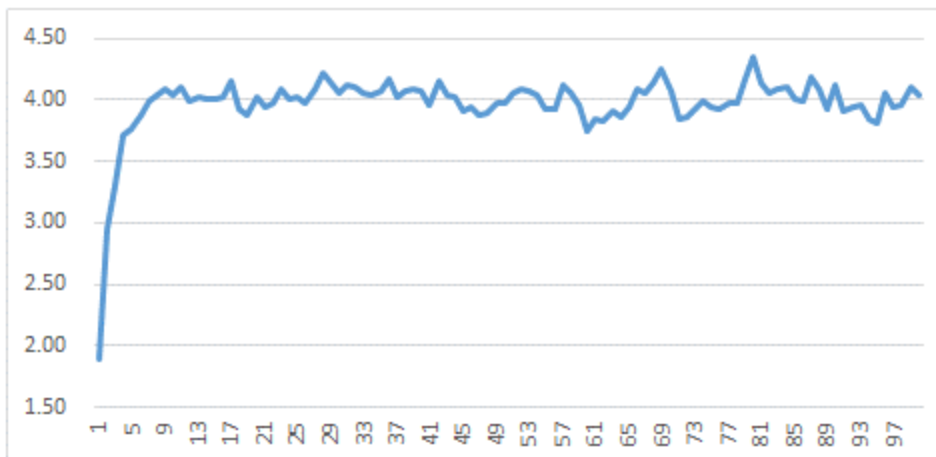
下の例では、 $Y_t = 2 + 0.5Y_{t-1} + \varepsilon_t$, $\varepsilon_t \sim N(0, 0.01)$ に従う時系列データを生成している。

このとき、 Y_t は4の近辺で変動することがみてとれる。

なお、データを生成する際には Y_0 が必要になるが、どんな値を選んでも、 Y_t はすみやかに4の近辺の値となる。

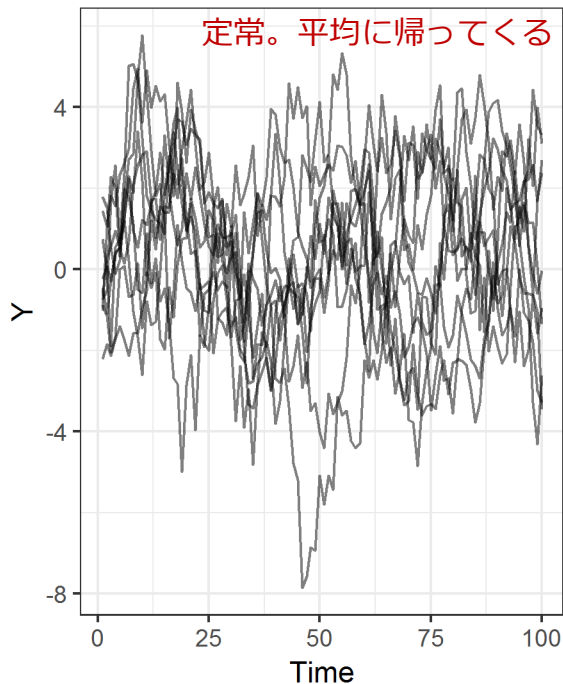
t	c	Phi1	Y[t-1]	E[t] (分散 0.01の正 規乱数)	Y[t]
1	2	0.5	0	-0.11	1.89
2	2	0.5	1.89	0.00	2.94
3	2	0.5	2.94	-0.17	3.30
4	2	0.5	3.30	0.06	3.71
5	2	0.5	3.71	-0.09	3.77
6	2	0.5	3.77	0.00	3.88
7	2	0.5	3.88	0.05	4.00
8	2	0.5	4.00	0.05	4.05
9	2	0.5	4.05	0.06	4.08
10	2	0.5	4.08	0.00	4.04
11	2	0.5	4.04	0.09	4.11
12	2	0.5	4.11	-0.06	4.00
13	2	0.5	4.00	0.02	4.02
14	2	0.5	4.02	-0.01	4.00
15	2	0.5	4.00	0.01	4.01
16	2	0.5	4.01	0.01	4.02

$$2 + 0.5 \times 2.94 - 0.17 = 3.30$$

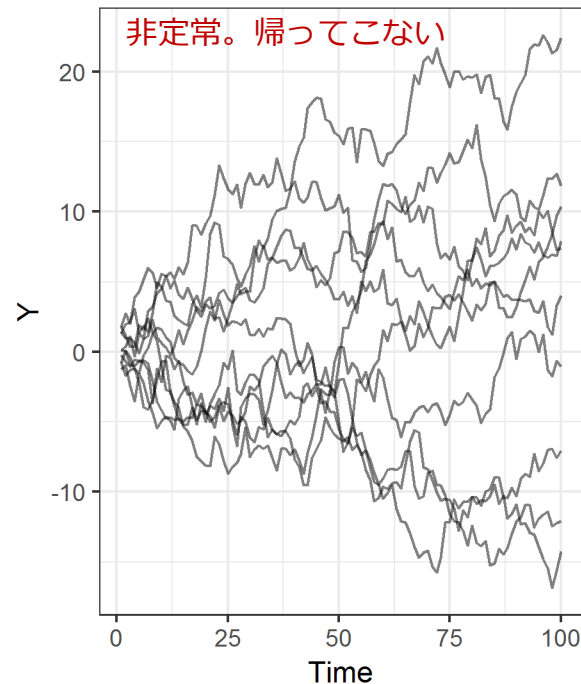


- AR(1) の定常性条件
 $|\phi_1| < 1$ のとき**定常**となる
 - 証明は略します。難しいからです、すいません
 - 下の例をみて納得して頂けると助かります

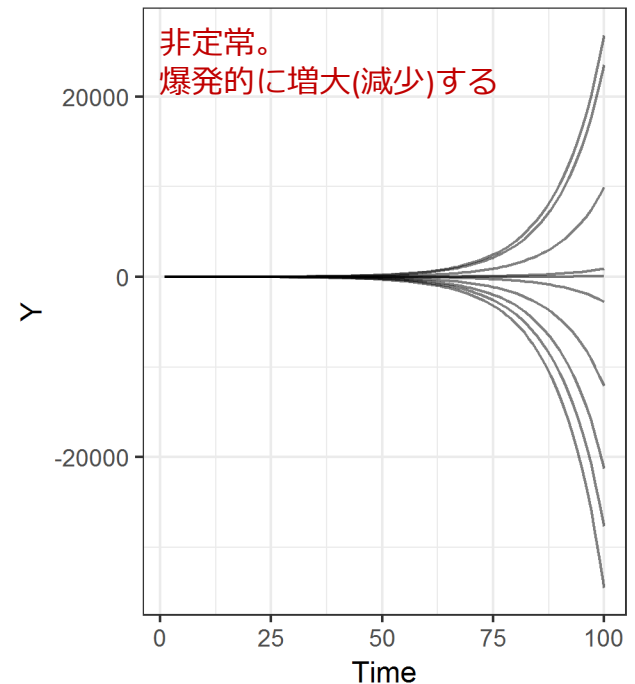
Code 5



$$Y_t = c + 0.9Y_{t-1} + \varepsilon_t$$



$$Y_t = c + Y_{t-1} + \varepsilon_t$$



$$Y_t = c + 1.1Y_{t-1} + \varepsilon_t$$

- 定常AR(1)過程を、ラグ演算子を使って書いてみよう

$$Y_t = c + \phi_1 Y_{t-1} + \varepsilon_t$$

ラグ演算子の登場！

$$Y_t = c + \phi_1 L Y_t + \varepsilon_t$$

第2項を左に移項

$$Y_t - \phi_1 L Y_t = c + \varepsilon_t$$

Y_t でくくろう

$$(1 - \phi_1 L) Y_t = c + \varepsilon_t$$

$$Y_t = \frac{1}{1 - \phi_1 L} c + \frac{1}{1 - \phi_1 L} \varepsilon_t$$

定常な場合、
両辺を $1 - \phi_1 L$ で割ってよい※

$$Y_t = \frac{1}{1 - \phi_1} c + \frac{1}{1 - \phi_1 L} \varepsilon_t$$

ラグ演算子は定数には無力だから

※「ラグ演算子で割る」ということがなにを意味するのか、訳が分からなくて気が狂いそうになりますが、あまり深く考えず、便利な略記方法として捉えてください。

$\frac{1}{1 - \phi_1 L} \varepsilon_t$ とは「 $(1 - \phi_1 L) X_t = \varepsilon_t$ である X_t 」という意味です。

- AR(1) 過程の性質

定常であるとき ($|\phi_1| < 1$ のとき)

- 平均 $\mu = \frac{c}{1-\phi_1}$
- 分散 $\gamma_0 = \frac{\sigma^2}{1-\phi_1^2}$
- 自己共分散 $\gamma_k = \phi_1 \gamma_{k-1}$
- 自己相関 $\rho_k = \phi_1 \rho_{k-1}$



なんだかキツネにつままれたような感じがするので、導出してみます

- 平均

モデルの両辺の期待値をとって

$$E(Y_t) = c + \phi_1 E(Y_{t-1})$$

定常であれば

$$\mu = c + \phi_1 \mu$$

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

- 分散

モデルの両辺の分散をとって

$$Var(Y_t) = Var(\phi_1 Y_{t-1} + \varepsilon_t)$$

攪乱項 ε_t と y_{t-1} との共分散は0なので、定常であれば

$$\gamma_0 = \phi_1^2 \gamma_0 + \sigma^2$$

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

- 自己共分散

$$\text{Cov}(Y_t, Y_{t-k}) = \text{Cov}(\phi_1 Y_{t-1} + \varepsilon_t, Y_{t-k})$$

攪乱項 ε_t と Y_{t-1} の共分散は0だから、分解して

$$\text{Cov}(Y_t, Y_{t-k}) = \phi_1 \text{Cov}(Y_{t-1}, Y_{t-k}) + \text{Cov}(\varepsilon_t, Y_{t-k})$$

攪乱項 ε_t と Y_{t-1} の共分散も0だから

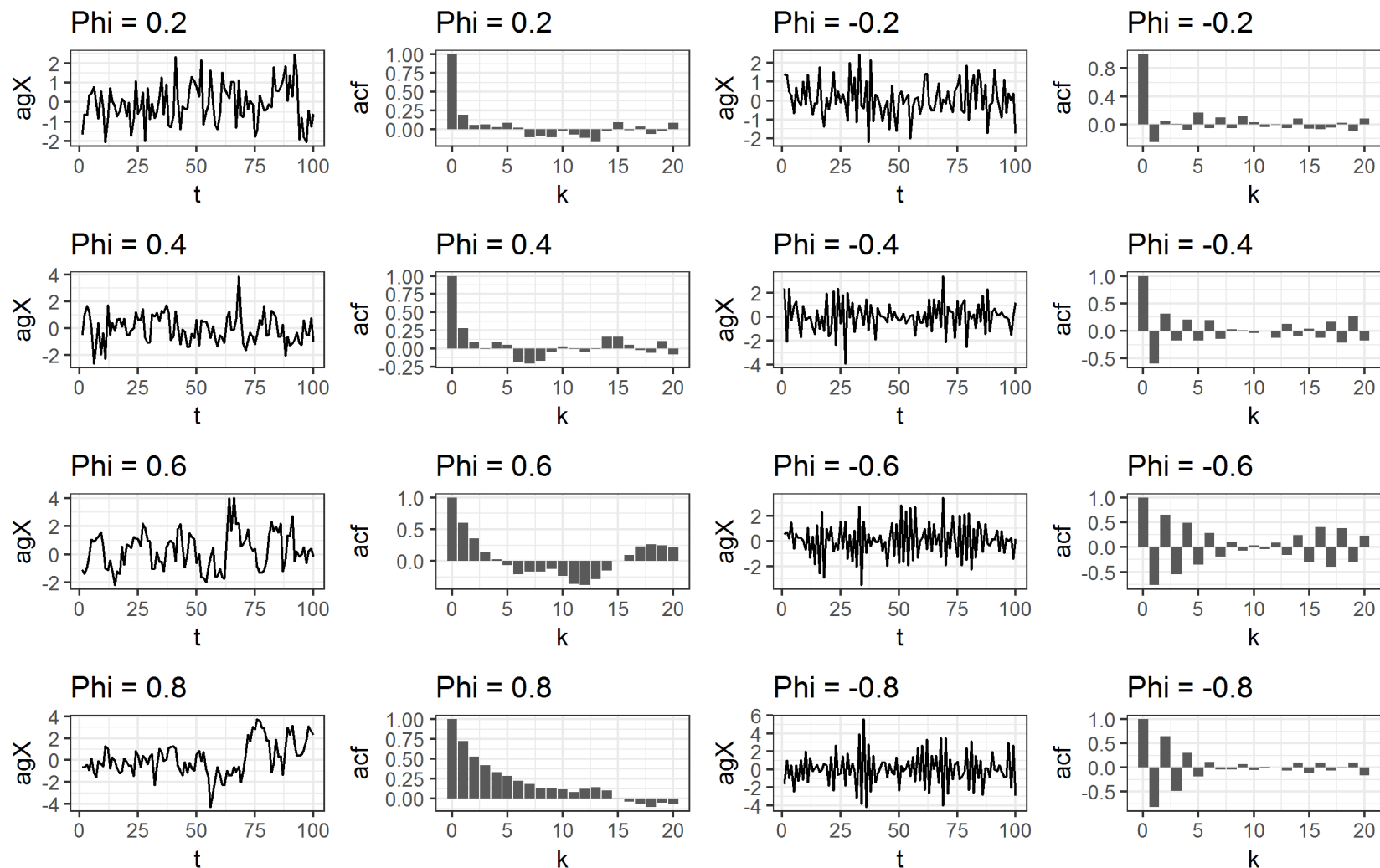
$$\gamma_k = \phi_1 \gamma_{k-1}$$

- 自己相関

上の式の両辺を γ_0 で割る。定常であれば $\rho_k = \gamma_k / \gamma_0$ なので

$$\rho_k = \phi_1 \rho_{k-1}$$

- AR(1) に従う時系列とその標本コレログラム



■ p次の自己回帰過程 (AR(p) 過程)

$$Y_t = c + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t$$

ただし、 ε_t はホワイトノイズ (分散 σ^2)。

• 性質

定常だったり非定常だったりする(次頁を参照)。

定常である場合は、

- ラグ演算子を使って書くと

$$y_t = \frac{c}{1 - \phi_1 - \dots - \phi_p} + \frac{\varepsilon_t}{1 - \phi_1 L - \dots - \phi_p L^p}$$

- 平均 $\mu = \frac{c}{1 - \phi_1 - \dots - \phi_p}$
- 分散, 自己共分散, 自己相関はちょっとややこしい。省略
- 自己相関は**指数的に減衰**する

- AR(p)が定常過程となる条件

- AR(p)過程 $Y_t = c + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t$ は、**AR特性方程式**とよばれるp次方程式

$$1 - \phi_1 z - \dots - \phi_p z^p = 0$$

のすべての解の絶対値が1より大きいときに定常である

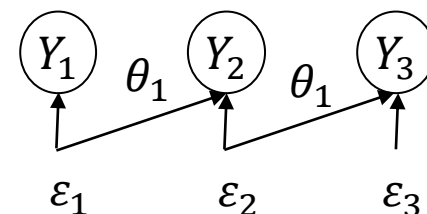
- 上の方程式を解くと...

- AR(1)は $1 - \phi_1 z = 0$ である z の絶対値が1より大きいときに定常
 - すなわち、 $|\phi_1| < 1$ のときに定常
- AR(2)は $1 - \phi_1 z - \phi_2 z^2 = 0$ の2つの解の絶対値がともに1より大きいときに定常
 - すなわち、 $\phi_2 > -1$ かつ $\phi_2 < 1 + \phi_1$ かつ $\phi_2 < 1 - \phi_1$ のときに定常

■ 1次の移動平均過程 (MA(1)過程) :

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

ただし、 ε_t はホワイトノイズ (分散 σ^2)。



• 性質

ラグ演算子を使って書くと

$$Y_t = \mu + (1 + \theta_1 L)\varepsilon_t$$

常に定常。

- 平均 μ
- 分散 $\gamma_0 = (1 + \theta_1^2)\sigma^2$
- 1次の自己共分散 $\gamma_1 = \theta_1 \sigma^2$
- 1次の自己相関 $\rho_1 = \frac{\theta_1}{1 + \theta_1^2}$
- 2次以降の自己共分散・自己相関は0

導出してみます

- 平均

モデルの両辺の期待値をとる

$$E(Y_t) = \mu + \theta_1 E(\varepsilon_{t-1}) + E(\varepsilon_t)$$

ホワイトノイズの期待値は0なので、平均は μ

- 分散

モデルの両辺の分散をとる。異なるホワイトノイズの共分散は0なので

$$Var(Y_t) = \theta_1^2 Var(\varepsilon_{t-1}) + Var(\varepsilon_t)$$

ホワイトノイズの分散は σ^2 なので

$$\gamma_0 = (1 + \theta_1^2)\sigma^2$$

- 1次自己共分散

$$Cov(Y_t, Y_{t-1}) = Cov(\theta_1 \varepsilon_{t-1} + \varepsilon_t, \theta_1 \varepsilon_{t-1} + \varepsilon_{t-2})$$

異なるホワイトノイズの共分散は0なので

$$Cov(Y_t, Y_{t-1}) = \theta_1 Cov(\varepsilon_{t-1}, \varepsilon_{t-1}) = \theta_1 \sigma^2$$

- k次自己共分散(k>1)

$$Cov(Y_t, Y_{t-k}) = Cov(\theta_1 \varepsilon_{t-1} + \varepsilon_t, \theta_1 \varepsilon_{t-k} + \varepsilon_{t-k})$$

異なるホワイトノイズの共分散は0なので, 0

- 定常性

- 平均、自己共分散のいずれも時点に依存していない。つまり定常

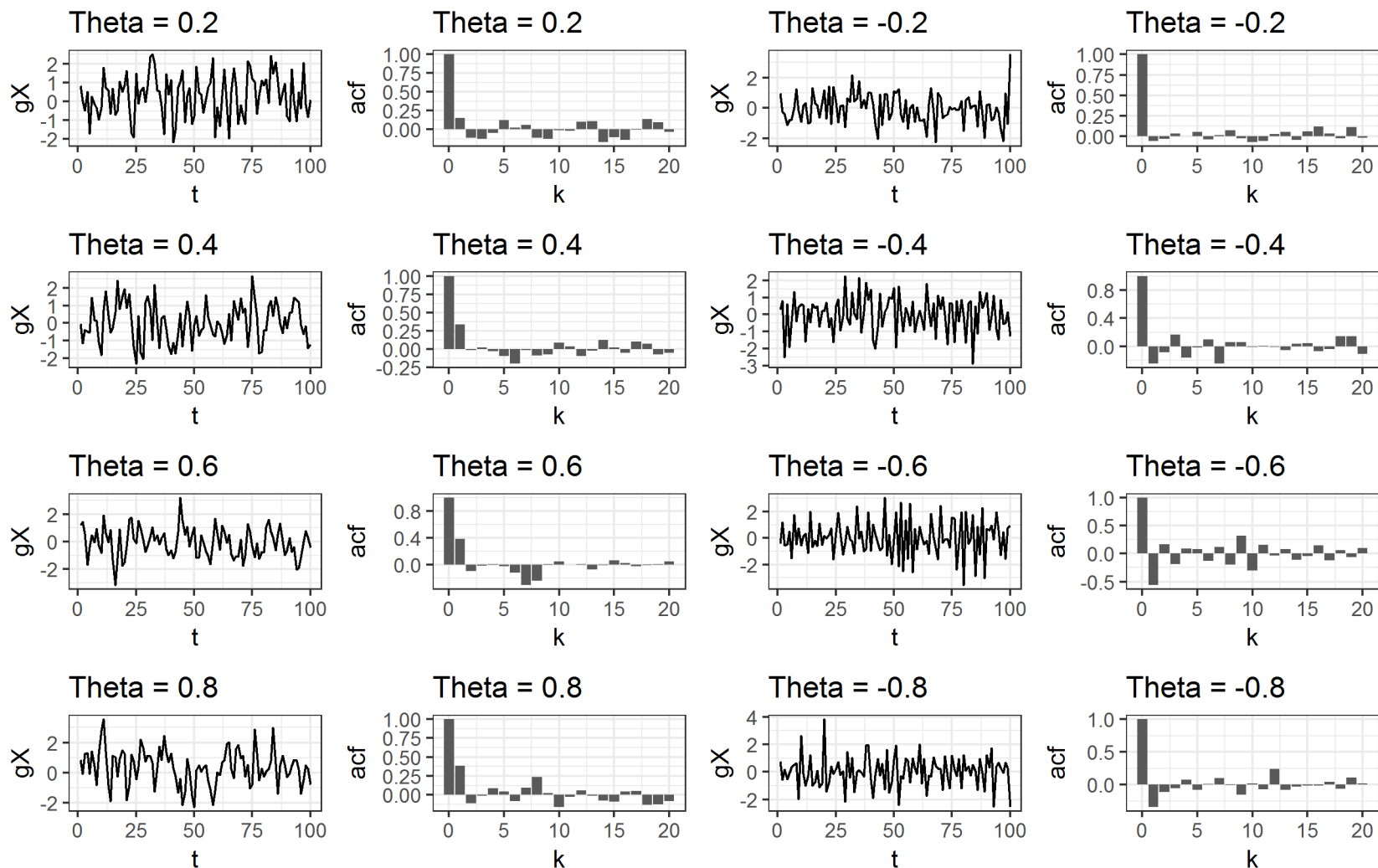
- 1次自己相関

1次自己共分散の式の両辺を $\gamma_0 = (1 + \theta_1^2)\sigma^2$ で割る。

定常だから $\rho_1 = \gamma_1/\gamma_0$ なので

$$\rho_1 = \frac{\theta_1}{1 + \theta_1^2}$$

- MA(1) に従う時系列とその標本コレログラム



■ q次の移動平均過程 (MA(q)過程) :

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

ただし、 ε_t はホワイトノイズ (分散 σ^2)。

• 性質

ラグ演算子を使って書くと

$$Y_t = \mu + (1 + \theta_1 L + \cdots + \theta_q L^q) \varepsilon_t$$

常に定常。

- 平均 μ
- 分散 $\gamma_0 = (1 + \theta_1^2 + \cdots + \theta_q^2) \sigma^2$
- 1次～q次の自己共分散は0でない。ややこしいので省略
- q+1次以降の自己共分散・自己相関は0

3) AR過程とMA過程の関係

- AR(p)は、定常であれば、MA(∞)過程に書き換えることができる

例) AR(1)過程

$$\begin{aligned}
 Y_t &= \phi_1 Y_{t-1} + \varepsilon_t \\
 &= \phi_1^2 Y_{t-2} + \phi_1 \varepsilon_{t-1} + \varepsilon_t \\
 &= \phi_1^3 Y_{t-3} + \phi_1^2 \varepsilon_{t-2} + \phi_1 \varepsilon_{t-1} + \varepsilon_t \\
 &\dots \\
 &= \phi_1^m Y_{t-m} + \phi_1^{m-1} \varepsilon_{t-(m-1)} + \phi_1^{m-2} \varepsilon_{t-(m-2)} + \dots + \varepsilon_t
 \end{aligned}$$

$Y_{t-1} = \phi_1 Y_{t-2} + \varepsilon_{t-1}$ を代入
 $Y_{t-2} = \phi_1 Y_{t-3} + \varepsilon_{t-2}$ を代入

よく見ると MA(2)になっている

よく見ると MA(m)になっている

$|\phi_1| < 1$ であれば、mが
 無限大に近づくと
 0に収束する

■ MA(q)は、**反転可能**であれば、AR(∞)過程に書き換えることができる

例) MA(1)過程 $y_t = \theta_1 \varepsilon_{t-1} + \varepsilon_t$ 。 $|\theta| < 1$ のときに反転可能である。

ε_t

$$= -\theta_1 \varepsilon_{t-1} + Y_t$$

$$= -\theta_1^2 \varepsilon_{t-2} - \theta_1 Y_{t-1} + Y_t$$

$$= -\theta_1^3 \varepsilon_{t-3} - \theta_1^2 Y_{t-2} - \theta_1 Y_{t-1} + Y_t$$

...

$$= \underbrace{-\theta_1^m \varepsilon_{t-m} - \theta_1^{m-1} Y_{t-(m-1)} + \theta_1^{m-2} Y_{t-(m-2)} + \cdots + Y_t}_{\text{よく見ると AR(m)になっている}}$$

$\varepsilon_{t-1} = -\theta_1 \varepsilon_{t-2} + Y_{t-1}$ を代入

$\varepsilon_{t-2} = -\theta_1 \varepsilon_{t-3} + Y_{t-2}$ を代入

よく見ると AR(2)になっている

よく見ると AR(m)になっている

$|\theta_1| < 1$ であれば、 m が
無限大に近づくと
0に収束する

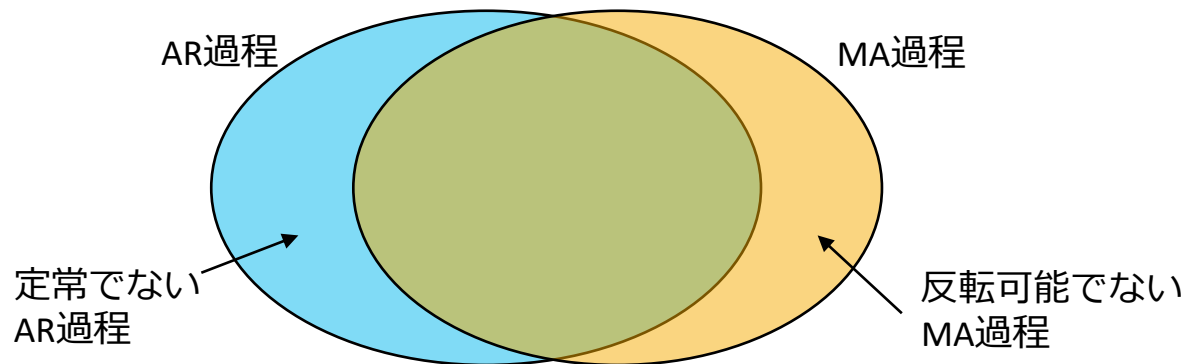
■ まとめ：AR過程とMA過程の関係

AR過程は、定常であつたりなかつたりする。

- AR(1) は、定常 ($|\phi| < 1$) であれば、MA(∞) に書き換えられる
- AR(p) は、定常であれば、MA(∞) に書き換えられる

MA過程は常に定常。反転可能であつたりなかつたりする。

- MA(1) は、反転可能 ($|\theta| < 1$) であれば、AR(∞) に書き換えられる
- MA(p) は、反転可能であれば、AR(∞) に書き換えられる



4) ARMA過程 (沖本, pp.34-47)

■ (p,q)次の自己回帰移動平均過程 (ARMA(p,q)過程) :

$$Y_t = c + \underbrace{\phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p}}_{\text{AR(p)の部分}} + \underbrace{\varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}}_{\text{MA(q)の部分}}$$

ただし、 ε_t はホワイトノイズ (分散 σ^2)。

• 性質

ARの部分が定常であれば、定常であり、

- ラグ演算子を使って書くと

$$Y_t = \frac{c}{1 - \phi_1 L - \dots - \phi_p L^p} + \frac{1 + \theta_1 L + \dots + \theta_q L^q}{1 - \phi_1 L - \dots - \phi_p L^p} \varepsilon_t$$

← MA(q)の部分
← AR(p)の部分

- 平均 $\mu = \frac{c}{1 - \phi_1 - \dots - \phi_p}$
- 分散, 自己共分散, 自己相関は滅茶苦茶にややこしい
- 自己相関は**指数的に減衰**する

■ ARMA(p,q) 過程の意義

p,qを変えるだけで、複雑な自己相関を表現できる

しかも、p, q は0~3くらいで十分だといわれている

→ 便利！

5. 基本的な確率過程 II. 非定常過程

どのような意味で定常でないか？

- 期待値が時間とともにゆるやかに変動する (トレンドを持っている)
 - 5.2 トrend定常過程
 - 5.3 単位根過程

- 分散が時間とともに変動する
- ある時点を境に、確率過程の性質が変わる
- ...

} このセミナーでは扱わない

5-2.トレンド定常過程

定常過程に、**確定的トレンド**が加わっている過程

- 確定的トレンド：前の時点からの増分が確定的に決まっているトレンド

■ さまざまなトレンド定常過程

- トレンドつきホワイトノイズ

$$Y_t = \beta_0 + \beta_1 t + \varepsilon_t$$

確定的トレンド

- トレンドつき定常AR(1)過程

$$Y_t = \beta_0 + \beta_1 t + \frac{1}{1 - \phi_1 L} \varepsilon_t$$

確定的トレンド

ラグ演算子を使わずに書くと、

$$Y_t = \beta_0 + \beta_1 t + X_t$$

$$X_t = \phi_1 X_{t-1} + \varepsilon_t$$

- トレンドつき定常ARMA(p,q)過程

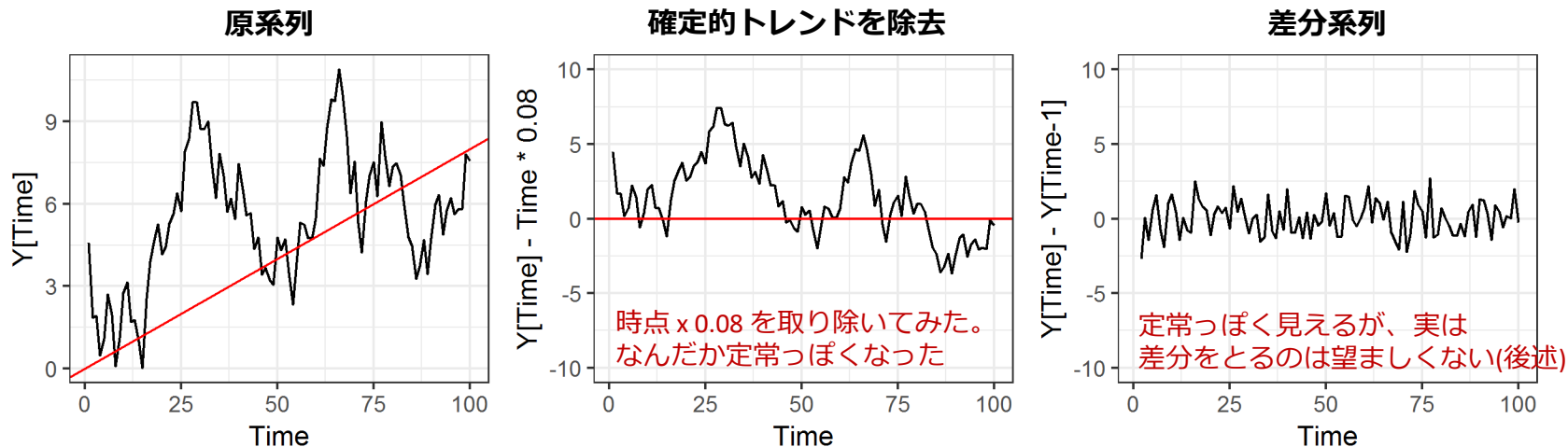
$$Y_t = \beta_0 + \beta_1 t + \frac{1 + \theta_1 L + \dots + \theta_q L^q}{1 - \phi_1 L - \dots - \phi_p L^p} \varepsilon_t$$

確定的トレンド

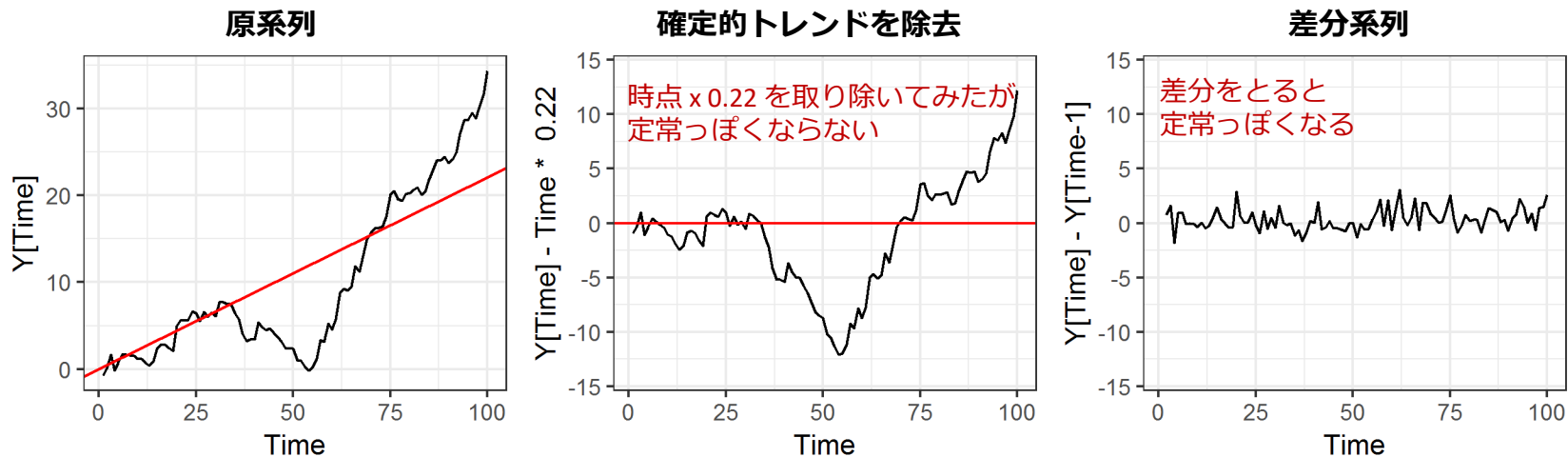
差分系列が定常である非定常過程 (差分定常過程)

- 差分系列の平均が0でないとき、**確率的トレンド**を持つ
 - 確率的トレンド：前の時点からの増分が確率的に決まるトレンド
- さまざまな非定常過程のなかでも、特に重要な概念

- 背後にある過程がトレンド定常である(ような気がする)時系列データの例



- 背後にある過程が差分定常である(ような気がする)時系列データの例



■ 「単位根過程」という名前の由来は？

- おさらい: AR(p)過程 $Y_t = c + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t$ は、**AR特性方程式**

$$1 - \phi_1 z - \dots - \phi_p z^p = 0$$

のすべての解の絶対値が1より大きいときに定常

- さらに、上の方程式を解いたとき、解のなかに $z = 1$ という解がひとつ含まれていたなら、それは差分定常過程であることがわかっている
- 昔は方程式の解のことを「根」と呼んだ。また、1である解を「単位根」と呼んだ
- ここから、差分定常過程のことを「AR過程でいえばAR特性方程式が単位根を持っている過程」、略して「単位根過程」と呼ぶ
- また、単位根過程を「単位根を持っている」と表現することが多い

■ ランダムウォーク

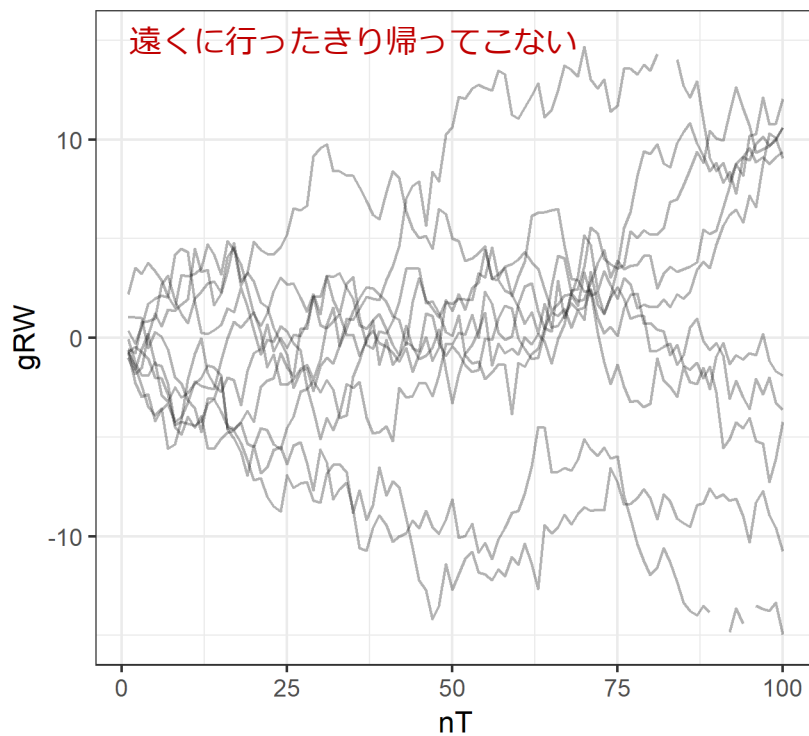
単位根過程の代表例。係数 1 で定数項がないAR(1)過程のこと。

$$Y_t = Y_{t-1} + \varepsilon_t$$

ただし、 ε_t はホワイトノイズ (分散 σ^2)。

• 性質

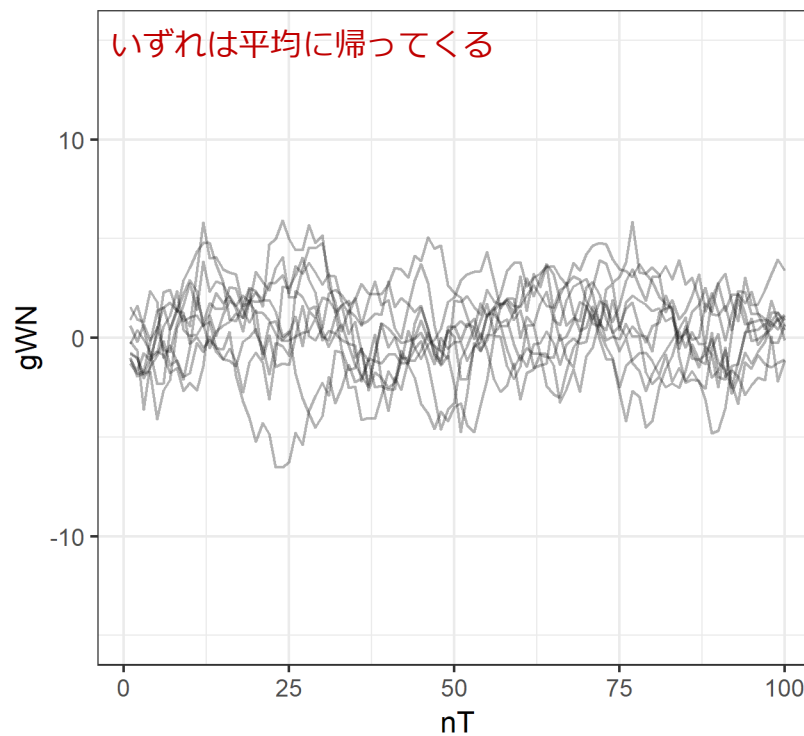
- 差分系列は $\Delta Y_t = Y_t - Y_{t-1} = \varepsilon_t$ (ホワイトノイズ)
- 期待値 $E(Y_t) = Y_0$
- 分散 $Var(Y_t) = \sigma^2 t$ ← 分散はどんどん大きくなる



ランダムウォーク

$$Y_t = Y_{t-1} + \varepsilon_t$$

ただし $Y_0 = 0, \varepsilon_t \sim N(0, 1)$



参考：定常AR(1)

$$Y_t = 0.9Y_{t-1} + \varepsilon_t$$

ただし $Y_0 = 0, \varepsilon_t \sim N(0, 1)$

いずれも10本の時系列を示している

■ ドリフトつきランダムウォーク

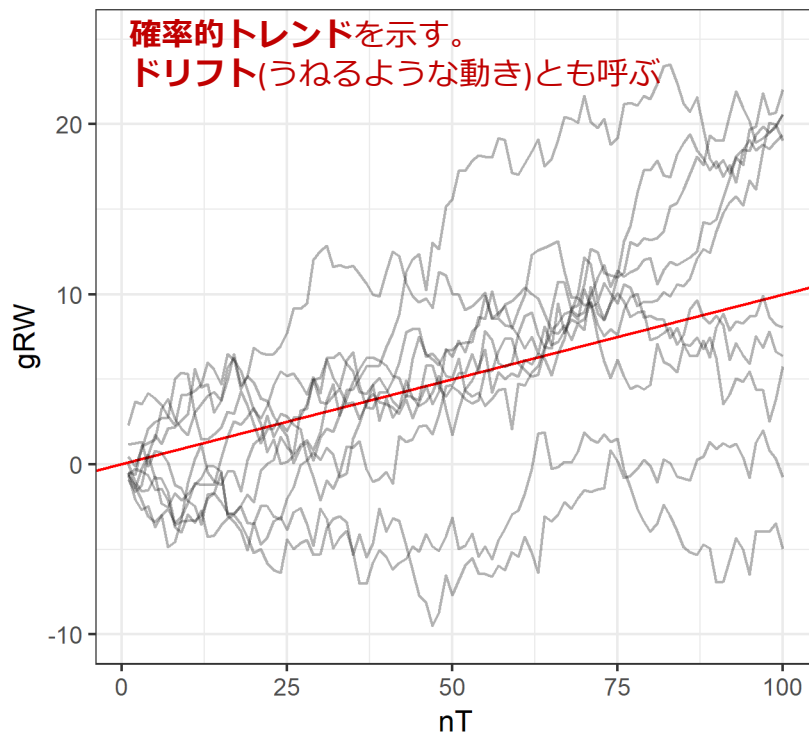
係数が 1 で定数項があるAR(1)過程のこと。

$$Y_t = \beta_1 + Y_{t-1} + \varepsilon_t \quad \leftarrow \beta_1 \text{を「ドリフト率」と呼ぶことがある}$$

ただし、 ε_t はホワイトノイズ (分散 σ^2)。

• 性質

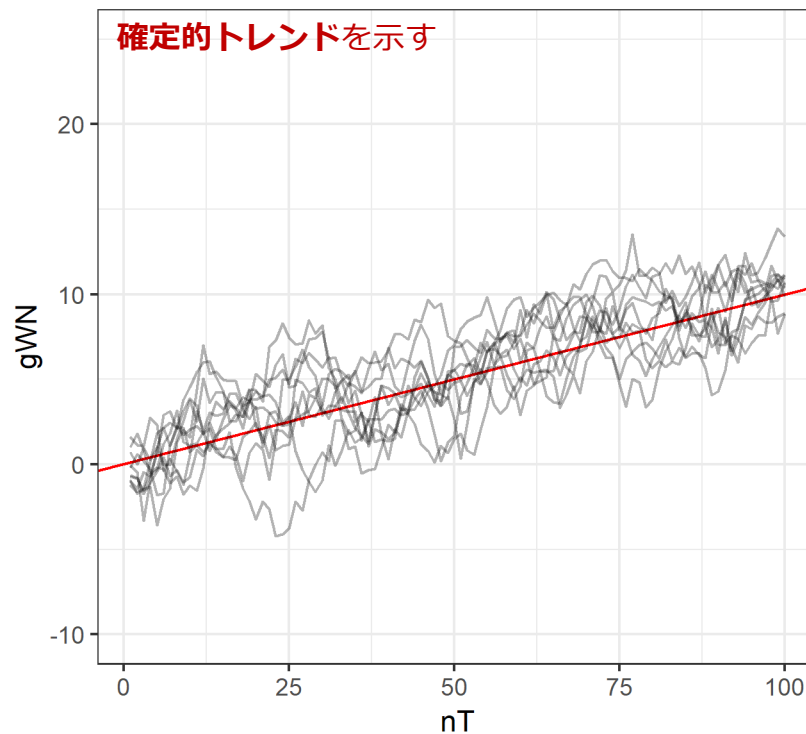
- 差分系列は $\Delta Y_t = \beta_1 + \varepsilon_t$ (定常)
- 期待値 $E(Y_t) = Y_0 + \beta_1 t$ $\leftarrow$ 期待値は時間とともに直線的に増える(減る)
- 分散 $Var(Y_t) = \sigma^2 t$ $\leftarrow$ 分散はどんどん大きくなる



ドリフトつきランダムウォーク

$$Y_t = 0.1 + Y_{t-1} + \varepsilon_t$$

ただし $Y_0 = 0, \varepsilon_t \sim N(0, 1)$



参考：トレンドつき定常AR(1)

$$Y_t = 0.1t + \frac{1}{1 - 0.9L} \varepsilon_t$$

ただし $Y_0 = 0, \varepsilon_t \sim N(0, 1)$

いずれも10本の時系列を示している

- ドリフトつきランダムウォークとランダムウォークの関係

$$Y_t = \beta_1 + Y_{t-1} + \varepsilon_t$$

ラグ演算子を使って書くと

$$Y_t = \beta_1 + LY_t + \varepsilon_t$$

$$(1 - L)Y_t = \beta_1 + \varepsilon_t$$

$\beta_1 = (1 - L)A_t, \varepsilon_t = (1 - L)E_t$ と書き換えることができるとして、

$$\cancel{(1 - L)}Y_t = \cancel{(1 - L)}A_t + \cancel{(1 - L)}E_t$$

$$Y_t = A_t + E_t$$

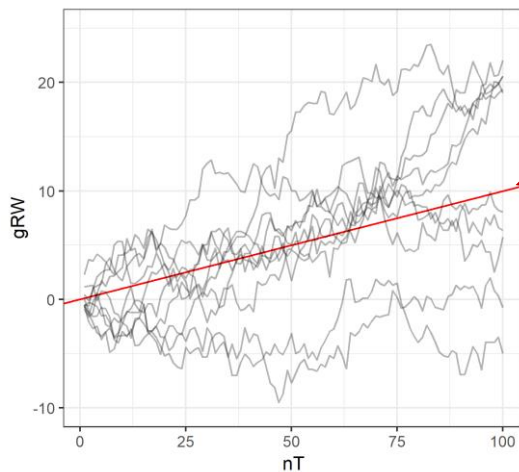
$\beta_1 = (1 - L)A_t$ である A_t とは？

$$A_t - A_{t-1} = \beta_1$$

だから、 A_t は「時点ごとに β_1 が加えられていく時系列」である。そこで

$$A_t = \beta_0 + \beta_1 t$$

と書き換えよう。ドリフトつきランダムウォークは下式となる：



$$Y_t = \beta_0 + \beta_1 t + E_t$$

$$(1 - L)E_t = \varepsilon_t$$

ホワイトノイズの累積
(ランダムウォーク)

■ ドリフトとトレンドつきランダムウォーク

$$Y_t = \beta_1 + \beta_2 t + Y_{t-1} + \varepsilon_t$$

← β_1 をドリフト、 $\beta_2 t$ をトレンドと呼ぶことがある

ただし、 ε_t はホワイトノイズ (分散 σ^2)。

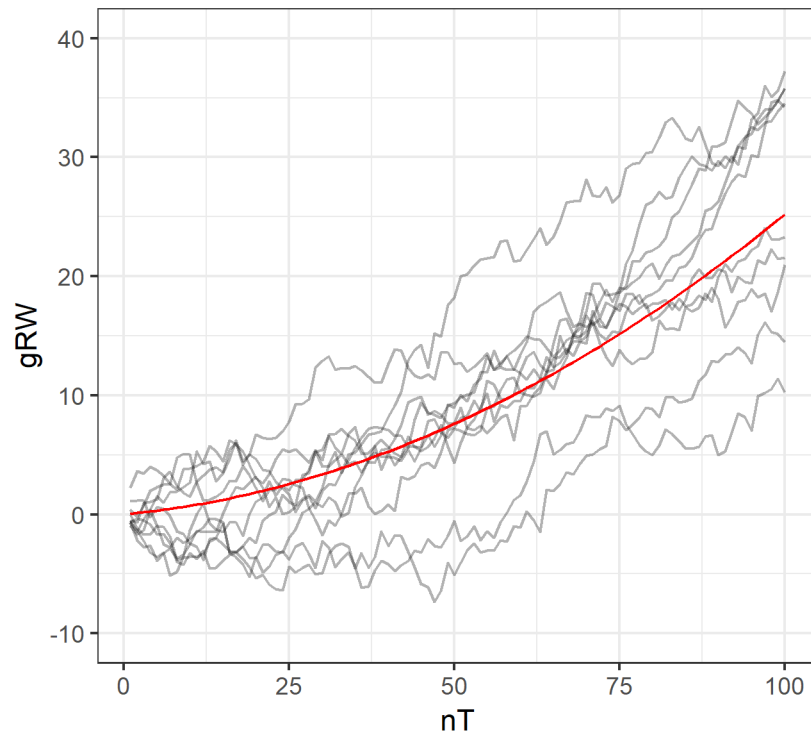
• 性質

- 差分系列は $\Delta Y_t = \beta_1 + \beta_2 t + \varepsilon_t$ (非定常)

- 2階差分系列は $\Delta^2 Y_t = \beta_2 + \varepsilon_t - \varepsilon_{t-1}$ (定常)

- 期待値 $E(y_t) = Y_0 + \left(\beta_1 + \frac{1}{2}\beta_2\right)t + \frac{1}{2}\beta_2 t^2$ ← 時間とともに二次曲線的に増える(減る)

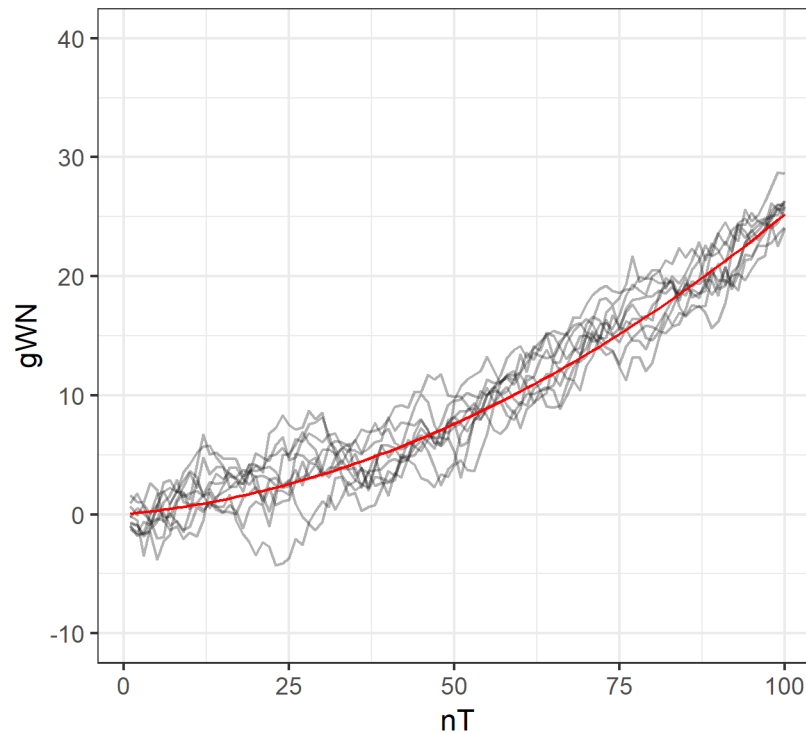
- 分散 $Var(Y_t) = \sigma^2 t$ ← 分散はどんどん大きくなる



ドリフトとトレンドつきランダム
ウォーク

$$Y_t = 0.05 + 0.004t + Y_{t-1} + \varepsilon_t$$

ただし $Y_0 = 0, \varepsilon_t \sim N(0, 1)$



参考：二次トレンドつき定常AR(1)

$$Y_t = 0.05t + 0.004t^2 + \frac{1}{1 - 0.9L} \varepsilon_t$$

ただし $Y_0 = 0, \varepsilon_t \sim N(0, 1)$

いずれも10本の時系列を示している

■ 市場反応モデルにおける実質的意義 (HPS pp.260-262)

- 単位根過程の性質
 - 平均に回帰しない。外生的ショックの影響は**永続的**
- (トレンドつき) 定常過程の性質
 - 平均に回帰する。外生的ショックの影響は**一時的**

↓

- 市場反応モデルにおいては：
 - 単位根過程
 - **進化的**市場を表現する
 - i.e. 「長期的なマーケティング効果」がありうる市場
 - (トレンドつき)定常過程
 - **定常的**市場を表現する
- 実際の市場は進化的？定常的？
 - これまでの1変量時系列分析の研究では...
 - 売上時系列の68%が進化的
 - シェア時系列の78%が定常的

6. 1変量時系列のモデル

本節では、**ある時系列のモデルを構築する**という問題を考えます

実際には、ある売上時系列のモデルを構築したところで、マーケティング活動の効果はわかりません

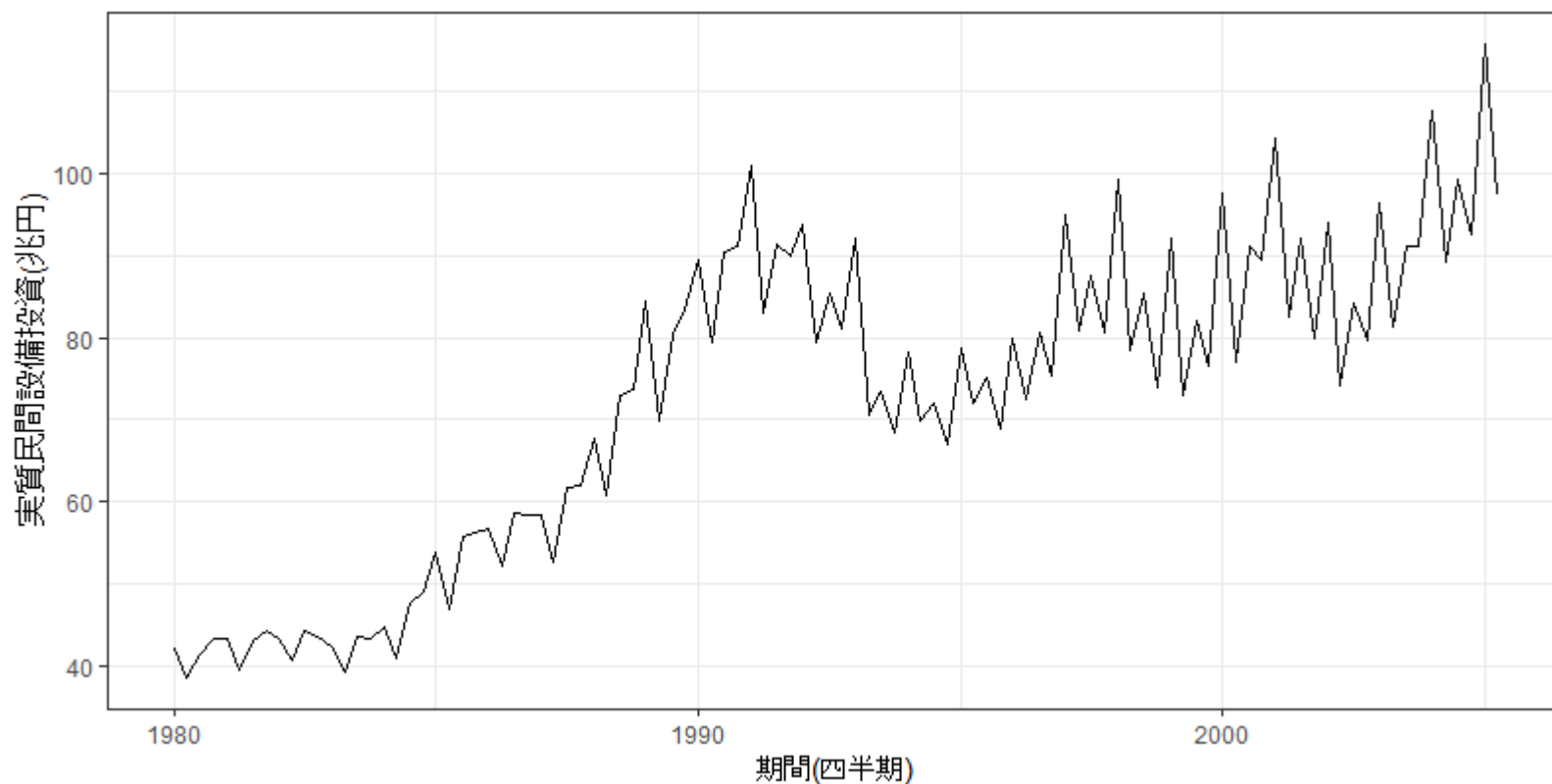
しかし、ここからの議論はとても重要です。なぜなら ... (HPS p.233)

- 時系列データに基づく市場反応モデルの基礎になるから
- 近い将来の値を説明変数なしで予測する際の手法として有用だから
- ある時系列から自己相関では説明できない変化を取り出す手法として有用だから

■ サンプルデータ: 民間設備投資

Code 11

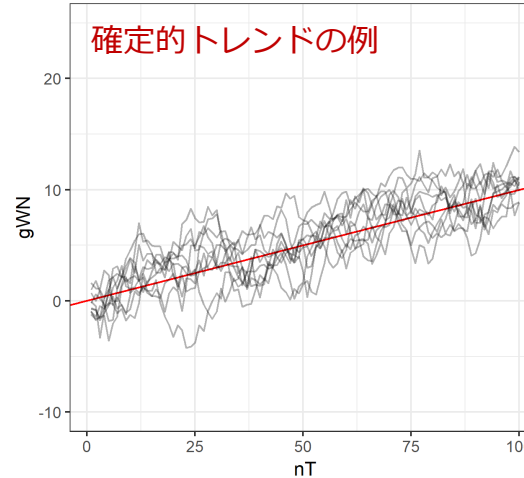
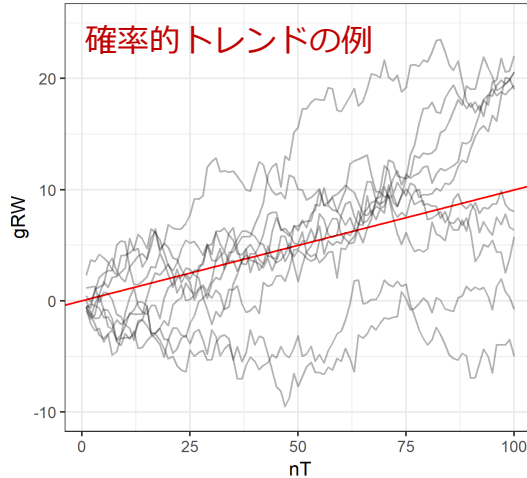
1980年第1四半期から2005年第2四半期までの日本の実質民間設備投資
(GDPの構成要素となる時系列)



■ モデル構築の手順

1. 時系列の観察
2. 時系列の変換
3. 定常性のチェック
4. モデルのあてはめ
5. 短期的予測

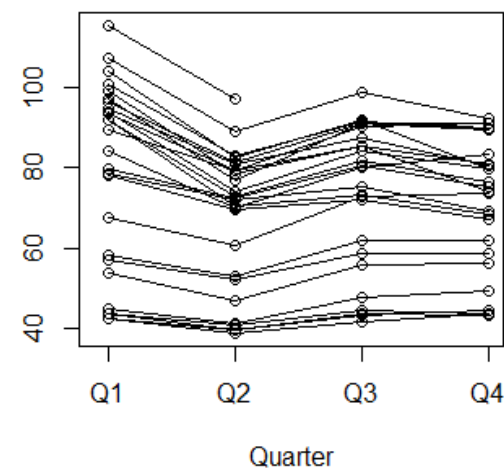
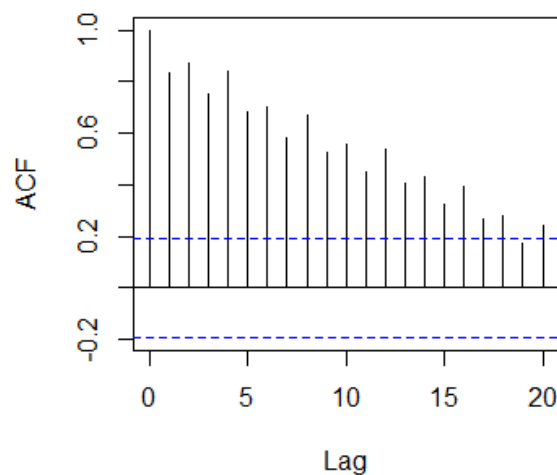
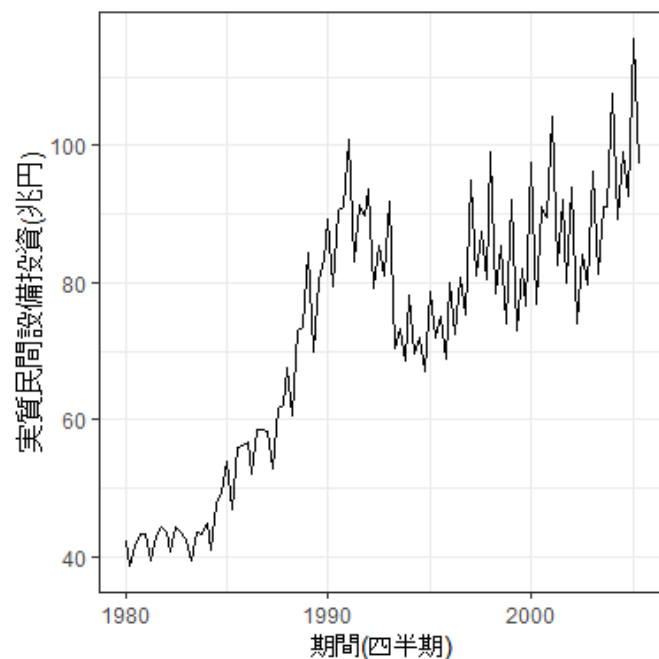
- トレンドはあるか？
 - 確率的トレンド
 - 時点ごとに増える(減る)が、量は確率的に変動する
 - 確定的トレンド
 - 時点ごとに一定の量だけ増える(減る)



Code 9

- 季節変動は？
- なんらかの説明が可能な変動は？
- 分散は均一か？

民間設備投資

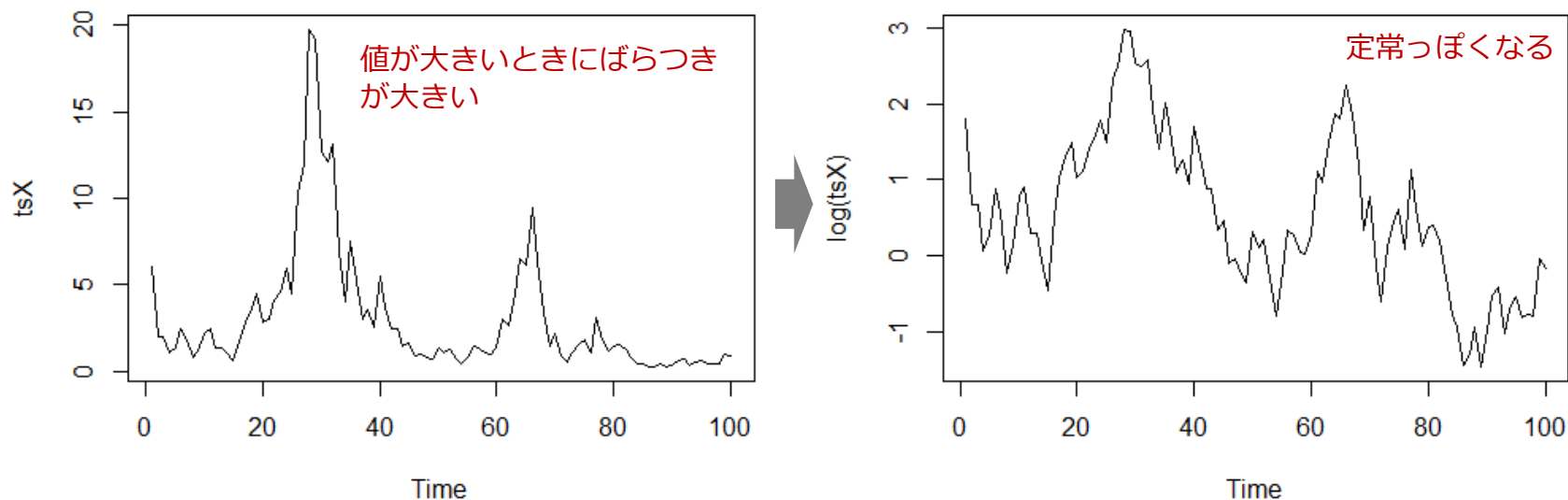


- トrendはある。確率的か確定的かはわからない
- 季節変動もありそう
- 分散は均一でない(高い金額で分散が大きい)

- 時系列分析では、もとの時系列データ (原系列) をなんらかのかたちで変換して分析することが多い
 - 実質的な観点から
 - 定常な時系列に近づけるため

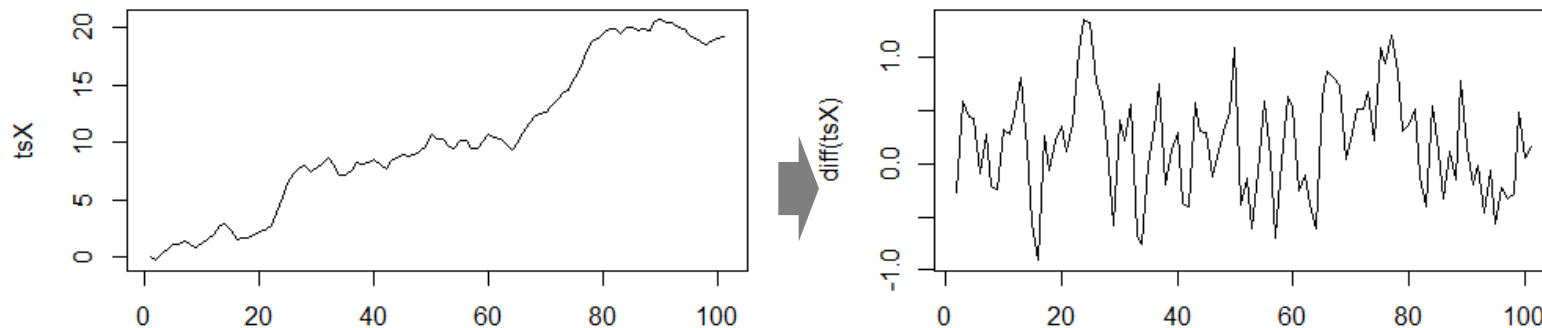
- **対数系列** $\log y_t$
分散均一性を得るために多用される

Code 12

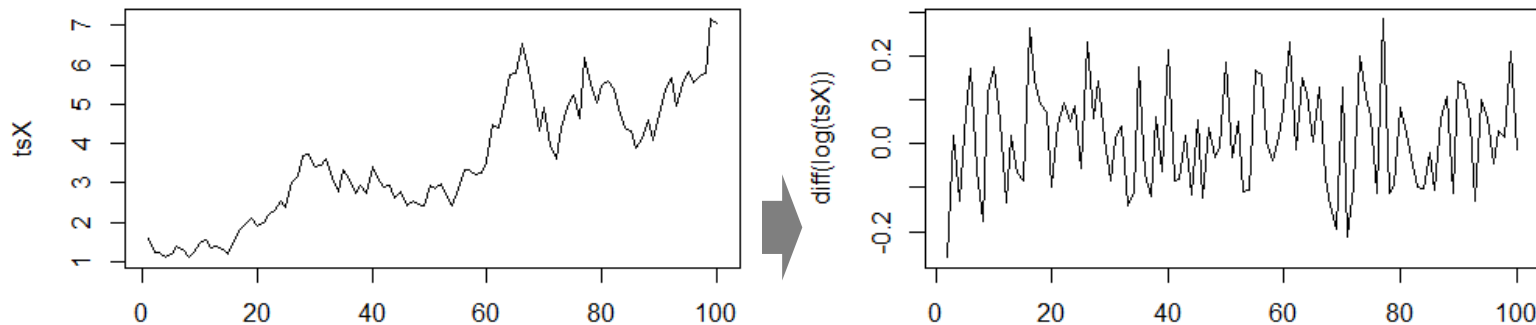


- 差分系列 $\Delta y_t = y_t - y_{t-1}$ (前期からの差)
確率的トレンドを除去する

Code 12



- 対数差分系列 $\Delta \log y_t = \log y_t - \log y_{t-1} \approx \frac{y_t - y_{t-1}}{y_{t-1}}$ (前期を1とした変化)



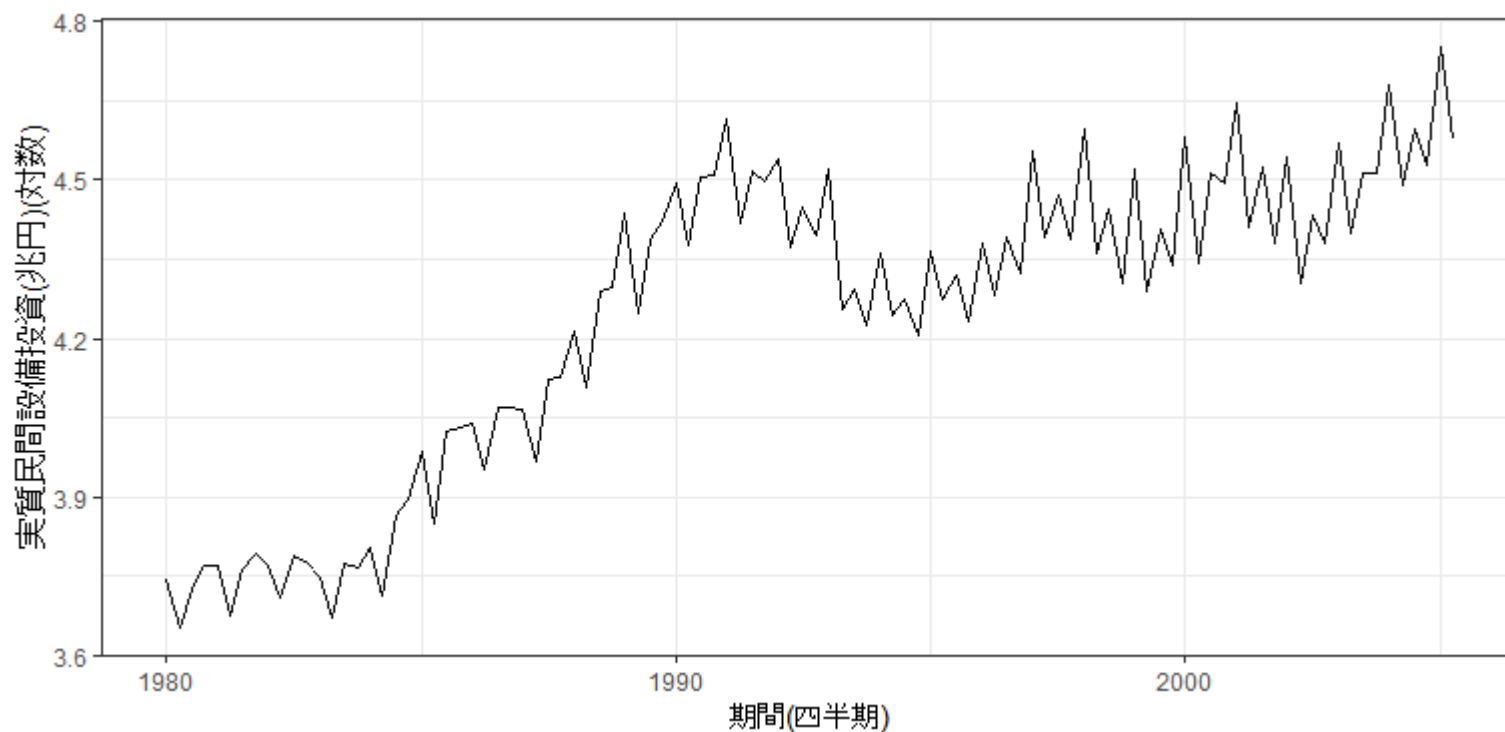
- 差分をとる変換は、単位根検定 (後述) を踏まえて行ったほうがよい

民間設備投資

分散均一性を得るため、対数系列をとることにしよう

ここでは差分はとらないことにしよう

季節調整もしないことにしよう



■ 単位根検定

- 時系列の背後にあるのが、(トレンドつき)定常過程なのか、単位根過程なのかを判断する方法
 - いいかえると：「(確定的トレンドをとり除けば)定常なのか、差分をとれば定常なのか」

■ 背景

- 非定常過程に従っている時系列に、定常過程を仮定したモデルをあてはめると、誤った推定を行うことになる
- 単位根過程は非定常過程の代表例
 - 外見では(トレンド)定常過程と見分けにくい
- 適した処理が異なる
 - **トレンド定常過程**
 - 確定的トレンドを取り除けば定常過程になる
 - 誤って差分をとると、偽りの自己相関が生じることが知られている(**過剰差分**)
 - **単位根過程**
 - 差分をとれば定常過程となる
 - 誤って確定的トレンドを取り除いても定常過程にはならない

■ 単位根検定の方法

- いろいろある
- 「単位根過程である」という帰無仮説を立てる手法
 - 「単位根過程でない」ことを積極的に示すのに適している
 - DF検定 (Dickey-Fuller検定)
 - **ADF検定** (拡張Dickey-Fuller検定)
 - ADF-GLS検定 ... ADF検定より優れている (黒住, 2008)
 - PP検定 (Phillips-Perron検定)
- 「確定的トレンド+定常過程である」という帰無仮説を立てる手法
 - 「単位根過程である」ことを積極的に示すのに適している
 - KPSS検定
- ここでは、ADF検定について紹介する
- Rの関数もいろいろある。ここでは次の2つを紹介する
 - `urca::ur.df()`
 - `forecast::ndiff()`

■ ADF検定 (拡張Dickey-Fuller検定)

2次トレンドつき定常AR(p)過程とドリフト・トレンドつき単位根AR(p)過程は、下の式に書き換えることができる。

$$\Delta Y_t = \beta_1 + \beta_2 t + \pi Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \cdots + \gamma_p \Delta Y_{t-p} + \varepsilon_t$$

さらに、 Y_t が単位根を持つとき、 $\pi = 0$ となることが知られている。

そこで、上のモデルをデータにあてはめ、 $\pi = 0$ について検定する

■ 単位根検定の手順 (田中, 2008)

Step 1. $\Delta Y_t = \beta_1 + \beta_2 t + \pi Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \dots + \gamma_p \Delta Y_{t-p} + \varepsilon_t$

検定1. 帰無仮説: $\pi = 0$ (=単位根を持つ)

検定2. 帰無仮説: $\pi = 0$ かつ $\beta_2 = 0$

検定3. 帰無仮説: $\pi = 0$ ($\beta_2 \neq 0$ と前提)

民間設備投資

Code 13

```
dfIP <- read_csv("./IPO1980.csv")
LIP <- log(dfIP$IP)

library(urca)
ourDF1 <- ur.df(LIP, type = "trend", lags = 4)
summary(ourDF1)
```

ある程度大きな値

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression trend

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-0.109590 -0.024968 -0.008654  0.031423  0.105416

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.3201651  0.1243460   2.575   0.0117 *
z.lag.1      -0.0818769  0.0326016  -2.511   0.0138 *
tt           0.0006361  0.0003163   2.011   0.0473 *
z.diff.lag1 -0.0963497  0.0722731  -1.333   0.1859
z.diff.lag2 -0.0365374  0.0705355  -0.518   0.6057
z.diff.lag3 -0.1393002  0.0696887  -1.999   0.0486 *
z.diff.lag4  0.7690449  0.0687855  11.180 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0441 on 90 degrees of freedom
Multiple R-squared:  0.8901,    Adjusted R-squared:  0.8828
F-statistic: 121.5 on 6 and 90 DF,  p-value: < 2.2e-16

value of test-statistic is: -2.5114 2.6964 3.1949

Critical values for test statistics:
      1pct  5pct 10pct
tau3  -3.99 -3.43 -3.13
phi2   6.22  4.75  4.07
phi3   8.43  6.49  5.47
```

z.diff.lag4が有意でなかったら
 (ここに*がなかったら)、
 lags 引数を1つ減らしてやり直す



```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression trend

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-0.109590 -0.024968 -0.008654  0.031423  0.105416

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.3201651  0.1243460   2.575   0.0117 *
z.lag.1      -0.0818769  0.0326016  -2.511   0.0138 *
tt           0.0006361  0.0003163   2.011   0.0473 *
z.diff.lag1 -0.0963497  0.0722731  -1.333   0.1859
z.diff.lag2 -0.0365374  0.0705355  -0.518   0.6057
z.diff.lag3 -0.1393002  0.0696887  -1.999   0.0486 *
z.diff.lag4  0.7690449  0.0687855  11.180 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0441 on 90 degrees of freedom
Multiple R-squared:  0.8901,    Adjusted R-squared:  0.8828
F-statistic: 121.5 on 6 and 90 DF,  p-value: < 2.2e-16

Value of test-statistic is: -2.5114 2.6964 3.1949

Critical values for test statistics:
      1pct  5pct 10pct
tau3  -3.99 -3.43 -3.13
phi2   6.22  4.75  4.07
phi3   8.43  6.49  5.47
```

検定1.
この値が下の丸印より小さかったら、
帰無仮説を棄却する



```
#####  
# Augmented Dickey-Fuller Test Unit Root Test #  
#####  
  
Test regression trend  
  
Call:  
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)  
  
Residuals:  
    Min       1Q   Median       3Q      Max   
-0.109590 -0.024968 -0.008654  0.031423  0.105416  
  
Coefficients:  
              Estimate Std. Error t value Pr(>|t|)      
(Intercept)  0.3201651  0.1243460   2.575   0.0117 *      
z.lag.1      -0.0818769  0.0326016  -2.511   0.0138 *      
tt           0.0006361  0.0003163   2.011   0.0473 *      
z.diff.lag1  -0.0963497  0.0722731  -1.333   0.1859        
z.diff.lag2  -0.0365374  0.0705355  -0.518   0.6057        
z.diff.lag3  -0.1393002  0.0696887  -1.999   0.0486 *      
z.diff.lag4   0.7690449  0.0687855  11.180  <2e-16 ***    
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.0441 on 90 degrees of freedom  
Multiple R-squared:  0.8901,    Adjusted R-squared:  0.8828  
F-statistic: 121.5 on 6 and 90 DF,  p-value: < 2.2e-16  
  
Value of test-statistic is: -2.5114 2.6964 3.1949  
  
Critical values for test statistics:  
    1pct  5pct 10pct  
tau3  -3.99 -3.43 -3.13  
phi2   6.22  4.75  4.07  
phi3   8.43  6.49  5.47
```

検定2.
この値が下の丸印より大きかったら、
帰無仮説を棄却する

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression trend

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-0.109590 -0.024968 -0.008654  0.031423  0.105416

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.3201651  0.1243460   2.575   0.0117 *
z.lag.1      -0.0818769  0.0326016  -2.511   0.0138 *
tt           0.0006361  0.0003163   2.011   0.0473 *
z.diff.lag1 -0.0963497  0.0722731  -1.333   0.1859
z.diff.lag2 -0.0365374  0.0705355  -0.518   0.6057
z.diff.lag3 -0.1393002  0.0696887  -1.999   0.0486 *
z.diff.lag4  0.7690449  0.0687855  11.180 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0441 on 90 degrees of freedom
Multiple R-squared:  0.8901,    Adjusted R-squared:  0.8828
F-statistic: 121.5 on 6 and 90 DF,  p-value: < 2.2e-16

Value of test-statistic is: -2.5114 2.6964 3.1949

Critical values for test statistics:
      1pct  5pct 10pct
tau3  -3.99 -3.43 -3.13
phi2   6.22  4.75  4.07
phi3   8.43  6.49  5.47
```

検定3.
この値が±1.96の外側だったら、
帰無仮説を棄却する

検定1.	検定2	検定3		判断
棄却			→	単位根なし
棄却されない	棄却	棄却されない	→	単位根あり
棄却されない	棄却	棄却	→	単位根なし
棄却されない	棄却されない		→	Step 2に進む

民間設備投資(対数)
はここ

Step 2. $\Delta Y_t = \beta_1 + \pi Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \dots + \gamma_p \Delta Y_{t-p} + \varepsilon_t$

検定1. 帰無仮説: $\pi = 0$ (=単位根を持つ)

検定2. 帰無仮説: $\pi = 0$ かつ $\beta_1 = 0$

検定3. 帰無仮説: $\pi = 0$ ($\beta_1 \neq 0$ と前提)

$\beta_2 t$ を削除

民間設備投資

Code 13

```
oURDF2 <- ur.df(LIP, type = "drift", lags = 4)
summary(oURDF2)
```

Step 1 と同様。
ある程度大きな値から始め、もしその
z.diff.lagが有意にならなかったら一つ減らす



```
#####  
# Augmented Dickey-Fuller Test Unit Root Test #  
#####  
  
Test regression drift  
  
Call:  
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)  
  
Residuals:  
      Min       1Q   Median       3Q      Max   
-0.113547 -0.027109 -0.003352  0.031242  0.101719  
  
Coefficients:  
              Estimate Std. Error t value Pr(>|t|)      
(Intercept)   0.11514     0.07237   1.591   0.1151      
z.lag.1       -0.02554     0.01696  -1.507   0.1354      
z.diff.lag1   -0.14707     0.06885  -2.136   0.0354 *     
z.diff.lag2   -0.07528     0.06898  -1.091   0.2780      
z.diff.lag3   -0.17159     0.06894  -2.489   0.0146 *     
z.diff.lag4    0.75120     0.06934  10.833  <2e-16 ***  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.04484 on 91 degrees of freedom  
Multiple R-squared:  0.8852,    Adjusted R-squared:  0.8789  
F-statistic: 140.3 on 5 and 91 DF,  p-value: < 2.2e-16  
  
Value of test-statistic is: -1.5065 1.957  
  
Critical values for test statistics:  
      1pct   5pct  10pct  
tau2 -3.46 -2.88 -2.57  
phi1  6.52  4.63  3.81
```

検定1.
この値が下の丸印より小さかったら、
帰無仮説を棄却する



```
#####  
# Augmented Dickey-Fuller Test Unit Root Test #  
#####  
  
Test regression drift  
  
Call:  
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)  
  
Residuals:  
      Min       1Q   Median       3Q      Max   
-0.113547 -0.027109 -0.003352  0.031242  0.101719  
  
Coefficients:  
              Estimate Std. Error t value Pr(>|t|)      
(Intercept)  0.11514     0.07237   1.591   0.1151      
z.lag.1      -0.02554     0.01696  -1.507   0.1354      
z.diff.lag1  -0.14707     0.06885  -2.136   0.0354 *     
z.diff.lag2  -0.07528     0.06898  -1.091   0.2780      
z.diff.lag3  -0.17159     0.06894  -2.489   0.0146 *     
z.diff.lag4   0.75120     0.06934  10.833  <2e-16 ***  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.04484 on 91 degrees of freedom  
Multiple R-squared:  0.8852,    Adjusted R-squared:  0.8789  
F-statistic: 140.3 on 5 and 91 DF,  p-value: < 2.2e-16  
  
Value of test-statistic is: -1.5066 1.957  
  
Critical values for test statistics:  
      1pct   5pct 10pct  
tau2 -3.46 -2.88 -2.57  
phi1  6.52 4.63  3.81
```

検定2.
この値が下の丸印より大きかったら、
帰無仮説を棄却する

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression drift

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-0.113547 -0.027109 -0.003352  0.031242  0.101719

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.11514    0.07237   1.591  0.1151
z.lag.1      -0.02554    0.01696  -1.507  0.1354
z.diff.lag1  -0.14707    0.06885  -2.136  0.0354 *
z.diff.lag2  -0.07528    0.06898  -1.091  0.2780
z.diff.lag3  -0.17159    0.06894  -2.489  0.0146 *
z.diff.lag4   0.75120    0.06934  10.833 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.04484 on 91 degrees of freedom
Multiple R-squared:  0.8852,    Adjusted R-squared:  0.8789
F-statistic: 140.3 on 5 and 91 DF,  p-value: < 2.2e-16

Value of test-statistic is: -1.5065 1.957

Critical values for test statistics:
      1pct  5pct 10pct
tau2  -3.46 -2.88 -2.57
phi1   6.52  4.63  3.81
```

検定3.
この値が±1.96の外側にあったら、
帰無仮説を棄却する

検定1.	検定2	検定3		判断
棄却			→	単位根なし
棄却されない	棄却	棄却されない	→	単位根あり
棄却されない	棄却	棄却	→	単位根なし
棄却されない	棄却されない		→	Step 3に進む

民間設備投資(対数)
はここ

Step 3. $\Delta Y_t = \pi Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \cdots + \gamma_p \Delta Y_{t-p} + \varepsilon_t$

検定1. 帰無仮説: $\pi = 0$ (=単位根を持つ)

β_1 を削除

民間設備投資

Code 13

```
oURDF3 <- ur.df(LIP, type = "none", lags = 4)
summary(oURDF3)
```

Step 1 と同様。
ある程度大きな値から始め、もしその
z.diff.lagが有意にならなかったら一つ減らす



```
#####  
# Augmented Dickey-Fuller Test Unit Root Test #  
#####
```

Test regression none

Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:

Min	1Q	Median	3Q	Max
-0.117731	-0.025814	-0.005051	0.032132	0.103038

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
z.lag.1	0.001368	0.001173	1.166	0.2465
z.diff.lag1	-0.156547	0.069164	-2.263	0.0260 *
z.diff.lag2	-0.075600	0.069551	-1.087	0.2799
z.diff.lag3	-0.166829	0.069445	-2.402	0.0183 *
z.diff.lag4	0.760465	0.069670	10.915	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.04521 on 92 degrees of freedom
Multiple R-squared: 0.8825, Adjusted R-squared: 0.8761
F-statistic: 138.2 on 5 and 92 DF, p-value: < 2.2e-16

value of test-statistic is: 1.1662

Critical values for test statistics:

	1pct	5pct	10pct
tau1	-2.58	-1.95	-1.62

検定1.
この値が下の丸印より小さかったら、
帰無仮説を棄却する

検定1.		判断
棄却	→	単位根なし
棄却されない	→	単位根あり

民間設備投資(対数)
はここ

- 「単位根なし」と判断されたら、次のステップに進む
- 「単位根あり」と判断されたら？
 - 差分系列をとって、「単位根なし」と判断されるまで繰り返す

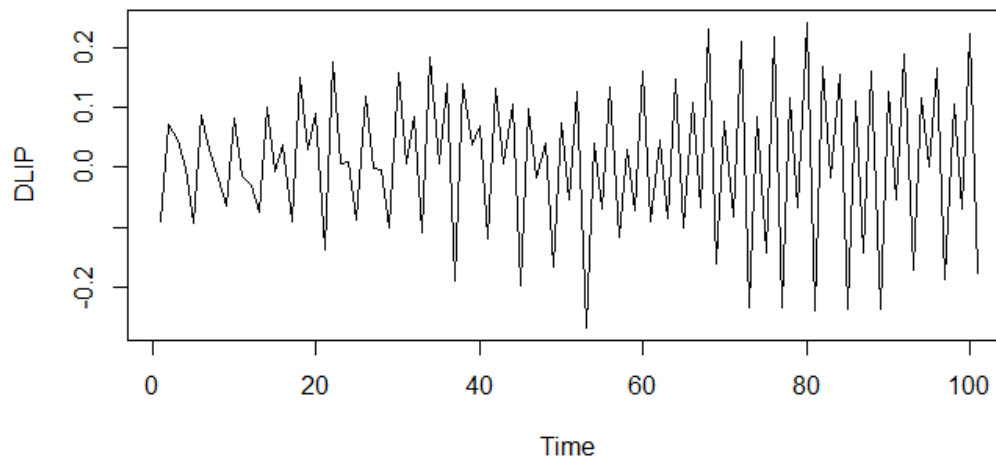
民間設備投資

「単位根あり」と判断されたので、差分をとってやり直しましょう...

Code 13

```
DLIP <- diff(LIP)
ts.plot(DLIP, type = "l")
OURDF1 <- ur.df(DLIP, type = "trend", lags=3)
summary(OURDF1)
OURDF2 <- ur.df(DLIP, type = "drift", lags=3)
summary(OURDF2)
OURDF3 <- ur.df(DLIP, type = "none", lags=3)
summary(OURDF3)
```

lags = 4 では有意にならなかった
ので、3に減らした





```
#####  
# Augmented Dickey-Fuller Test Unit Root Test #  
#####  
  
Test regression trend  
  
Call:  
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)  
  
Residuals:  
    Min       1Q   Median       3Q      Max   
-0.118232 -0.027526 -0.004958  0.031374  0.101757  
  
Coefficients:  
            Estimate Std. Error t value Pr(>|t|)      
(Intercept)  8.868e-03  1.027e-02   0.864  0.39011      
z.lag.1      -6.573e-01  2.272e-01  -2.893  0.00477 **     
tt           -4.645e-05  1.665e-04  -0.279  0.78084      
z.diff.lag1  -5.030e-01  1.759e-01  -2.859  0.00527 **     
z.diff.lag2  -5.833e-01  1.273e-01  -4.583  1.46e-05 ***    
z.diff.lag3  -7.554e-01  7.054e-02 -10.708 < 2e-16 ***  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.04537 on 91 degrees of freedom  
Multiple R-squared:  0.9659,    Adjusted R-squared:  0.964  
F-statistic: 515.2 on 5 and 91 DF,  p-value: < 2.2e-16  
  
Value of test-statistic is: -2.8932 2.8028 4.1964  
  
Critical values for test statistics:  
      1pct  5pct 10pct  
tau3 -3.99 -3.43 -3.13  
phi2  6.22  4.75  4.07  
phi3  8.43  6.49  5.47
```

Step 1
検定1も検定2も棄却されない

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression drift

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-0.118073 -0.027069 -0.005819  0.031957  0.102086

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.006369   0.005001   1.274  0.20603
z.lag.1      -0.647927   0.223566  -2.898  0.00469 **
z.diff.lag1  -0.510441   0.173022  -2.950  0.00403 **
z.diff.lag2  -0.588340   0.125337  -4.694  9.31e-06 ***
z.diff.lag3  -0.757703   0.069683 -10.874 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.04514 on 92 degrees of freedom
Multiple R-squared:  0.9658,    Adjusted R-squared:  0.9644
F-statistic: 650.5 on 4 and 92 DF,  p-value: < 2.2e-16

Value of test-statistic is: -2.8981 4.2074

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.46 -2.88 -2.57
phi1  6.52  4.63  3.81
```

Step 2
検定1で棄却された！

**1回差分をとれば定常時系列になる
 ということがわかった。**

以上の手続きにより、差分をとるべき回数があきらかになった。

頑張ってご説明してまいりましたが...

■ 単位根検定の手順 (手っ取り早いバージョン)

forecast::ndiffs (時系列)

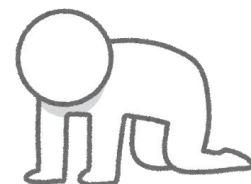
「何回差分をとれば定常時系列になるのか」を示す

民間設備投資

Code 13

```
library(forecast)
ndiffs(LIP, test = "adf", type = "trend")
```

```
[1] 1
```



■ ARMA(p,q)モデル

$$Y_t = \alpha + \underbrace{\phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p}}_{\text{AR(p)の部分}} + \underbrace{\varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}}_{\text{MA(q)の部分}}$$

- ラグ演算子を使って書くと

$$Y_t = \alpha + \boxed{\frac{1 + \theta_1 L + \dots + \theta_q L^q}{1 - \phi_1 L - \dots - \phi_p L^p}} \varepsilon_t$$

この部分を
ARMA攪乱項と呼ぼう

- なぜARMA(p, q) モデル？
 - 少数のパラメータ (p+q個) で、自己相関を柔軟に表現できるから
 - パラメータ推定値にはあまり関心がない
- 推定の方法は？
 - 通常は最尤法を使う (後述)

■ ARIMA (p,d,q)モデル (自己回帰和分移動平均モデル)

- 「 Y_t についてd階差分をとってARMA(p,q)モデルをあてはめる」こと
- d=1の場合について、ラグ演算子を使って書くと...

$$Y_t = \alpha_0 + \alpha t + \boxed{E_t} \quad \leftarrow \begin{array}{l} \text{ARMA撹乱項の累積。} \\ \text{ARIMA撹乱項と呼ぼう} \end{array}$$

ただし

$$(1 - L)E_t = \boxed{\frac{1 + \theta_1 L + \dots + \theta_q L^q}{1 - \phi_1 L - \dots - \phi_p L^p} \varepsilon_t} \quad \text{ARMA撹乱項}$$

注) 前頁の式の導出:

ARIMA(p,1,q)モデルは、差分系列について

$$Y_t - Y_{t-1} = \alpha + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

差分をとっている

$$(1 - L)Y_t = \alpha + \frac{1 + \theta_1 L + \dots + \theta_q L^q}{1 - \phi_1 L - \dots - \phi_p L^p} \varepsilon_t$$

$$\alpha = (1 - L)A_t, \quad \frac{1 + \theta_1 L + \dots + \theta_q L^q}{1 - \phi_1 L - \dots - \phi_p L^p} \varepsilon_t = (1 - L)E_t \text{ とすれば}$$

$$\cancel{(1 - L)}Y_t = \cancel{(1 - L)}A_t + \cancel{(1 - L)}E_t$$

$$Y_t = A_t + E_t$$

$\alpha = (1 - L)A_t$ より、 A_t は每期 α ずつ増大する時系列だから

$$Y_t = \alpha_0 + \alpha t + E_t$$

パラメータ推定の方法

■ 最小二乗法

- AR(p) モデルの場合
 - 不偏性はない
 - 説明変数 Y_{t-1} は一期前の攪乱項 ε_{t-1} と独立でない
 - 従って、重回帰の古典的仮定のうち [1] が満たされていない
 - しかし一貫性を持つ
 - 攪乱項が正規ホワイトノイズであれば、一致推定量の中で分散が漸近的に最小
- MA(q), ARMA(p, q) モデルの場合
 - 不偏性も一貫性もない

■ 最尤法

- 攪乱項が正規ホワイトノイズであることを仮定
- AR(p) モデルの場合、最小二乗法の解と同一
- 不偏性はない
- しかし一貫性を持つ
- 一致推定量の中で分散が漸近的に最小

モデル選択の手順

■ 差分をとることなく「単位根なし」と判断された場合

- トрендがありそうな場合は、
 - $Y_t = b_1 t + Y'_t$ ないし $Y_t = b_1 t + b_2 t^2 + Y'_t$ を最小二乗法で推定し、残差 Y'_t を求め...
 - Y'_t に ARMA(p, q) モデルをあてはめる
- トрендがなさそうな場合は
 - ARMA(p, q) モデルをあてはめる

■ d回差分をとって「単位根なし」と判断された場合

- ARIMA(p, d, q) モデルをあてはめる

■ 次数 (p, q) をどのように決めるか？

- さまざまな (p, q) についてモデルを推定し比較する

■ Rパッケージ

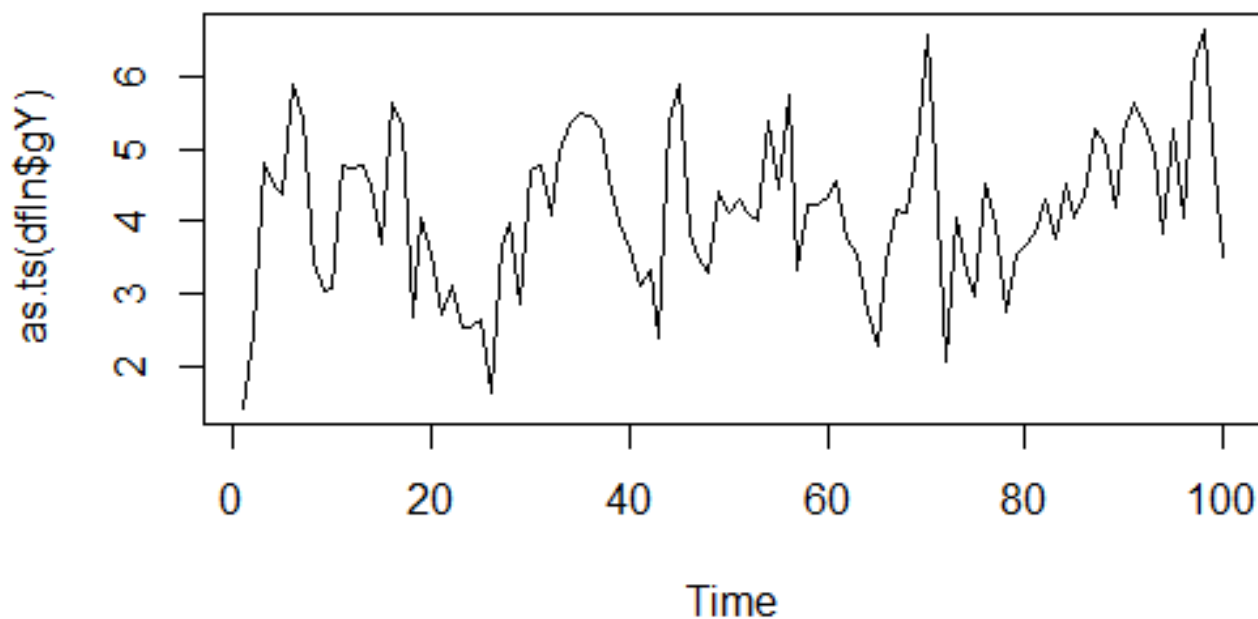
いろいろある。ここでは次の2つを紹介する

- `forecast::Arima()`
- `forecast::auto.arima()`

■ データ例

Code 14

```
# dfIn$gYに時系列がはいっている。単位根を持たないことがわかっている  
plot(as.ts(dfIn$gY))
```



■ AR(p)の推定

最小二乗推定が可能

```
# forecast::Arima() で AR(1)モデルを当てはめる  
library(forecast)  
Arima(dfIn$gY, order = c(1, 0, 0), method = "CSS")
```

```
Series: dfIn$gY  
ARIMA(1,0,0) with non-zero mean
```

```
Coefficients:
```

	ar1	mean
	0.4487	4.1901
s.e.	0.0858	0.1649

```
sigma^2 estimated as 0.8344: part log likelihood=-132.33
```

モデルを $Y_t = c + \phi_1 Y_{t-1} + \varepsilon_t$, $\mu = c/(1 - \phi_1)$ として
 $\hat{\phi}_1 = 0.4487, \hat{\mu} = 4.19$

なお $\hat{c} = 4.19(1 - 0.4487) = 2.31$

```
# 参考: lm()で推定した場合
# dfIn$gLagYに前期の値が入っているとして
summary(lm(gY ~ gLagY, data = dfIn))
```

```
Call:
lm(formula = gY ~ gLagY, data = dfIn)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-2.37379	-0.58777	0.00698	0.57049	2.08449

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.31016	0.37320	6.190	1.44e-08 ***
gLagY	0.44866	0.08709	5.151	1.36e-06 ***

SEの推定は正しくない

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9181 on 97 degrees of freedom
(1 observation deleted due to missingness)

Multiple R-squared: 0.2148, Adjusted R-squared: 0.2067

F-statistic: 26.54 on 1 and 97 DF, p-value: 1.358e-06

■ ARMA(p,q)の推定

p, qを指定して最尤推定

```
# ARMA(2, 0)
Arima(dfIn$gY, order = c(2, 0, 0))
```

```
Series: dfIn$gY
ARIMA(2,0,0) with non-zero mean
```

```
Coefficients:
```

	ar1	ar2	mean
	0.5053	-0.0659	4.1207
s.e.	0.1020	0.1034	0.1659

```
sigma^2 estimated as 0.8989: log likelihood=-135.17
AIC=278.34   AICc=278.76   BIC=288.76
```

```
# ARMA(1, 1)
Arima(dfIn$gY, order = c(1, 0, 1))
```

ARIMA(1,0,1) with non-zero mean

Coefficients:

	ar1	ma1	mean
	0.1169	0.4254	4.1293
s.e.	0.3476	0.3473	0.1498

sigma^2 estimated as 0.8917: log likelihood=-134.8
AIC=277.59 AICc=278.01 BIC=288.01

AICが小さくなった。

AR(2)よりもARMA(1,1)のほうがよさそうだ...

しかし、ARの係数が小さい。むしろMA(0,1)のほうがよいかも...

→ p, qが0~4くらいまでの値について、片っ端から推定する

■ ARMA(p,q)の推定 (手っ取り早いバージョン)

```
forecast::auto.arima()
```

さまざまなモデルを片っ端から試し、よさそうなモデルを返す

```
auto.arima(  
  dfIn$gY, d = 0,  
  trace=T, seasonal = F, ic = "aic", stepwise=F, approximation = F  
)
```

```
ARIMA(0,0,0) with zero mean      : 576.4883  
ARIMA(0,0,0) with non-zero mean : 298.7064  
ARIMA(0,0,1) with zero mean     : 467.3743  
(中略)  
ARIMA(5,0,0) with non-zero mean : 280.2263
```

Best model: ARIMA(0,0,1) with non-zero mean

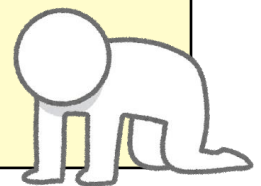
結局, MA(1)がお勧め

```
Series: dfIn$gY  
ARIMA(0,0,1) with non-zero mean
```

Coefficients:

	ma1	mean
	0.5313	4.1324
s.e.	0.0925	0.1420

```
sigma^2 estimated as 0.8835: log likelihood=-134.86  
AIC=275.72   AICc=275.97   BIC=283.53
```



■ ARIMA(p,d,q)の推定

p, d, qを指定して最尤推定

民間設備投資

```
# ARIMA(2, 1, 2)の推定
Arima(LIP, order = c(2, 1, 2), method = "ML")
```

```
Series: LIP
ARIMA(2,1,2)
```

```
Coefficients:
```

	ar1	ar2	ma1	ma2
	-0.6612	0.3365	0.4000	-0.4750
s.e.	0.2055	0.2052	0.1734	0.1542

```
sigma^2 estimated as 0.004752: log likelihood=127.56
AIC=-245.12   AICc=-244.49   BIC=-232.05
```


■ ARIMA(p,d,q)の推定 (手っ取り早いバージョン)

民間設備投資

```
auto.arima(  
  LIP, d = 1,  
  trace=T, seasonal = F, ic = "aic", stepwise=F, approximation = F  
)
```

```
ARIMA(0,1,0) : -129.0637  
ARIMA(0,1,0) with drift : -127.4935  
(中略)  
ARIMA(5,1,0) with drift : Inf
```

Best model: ARIMA(3,1,2) with drift

ARIMA(3,1,2)がお勧め

Series: LIP
ARIMA(3,1,2) with drift

Coefficients:

	ar1	ar2	ar3	ma1	ma2	drift
	-0.9764	-0.8642	-0.8481	0.6576	0.5076	0.0092
s.e.	0.0590	0.0884	0.0614	0.1201	0.1349	0.0029

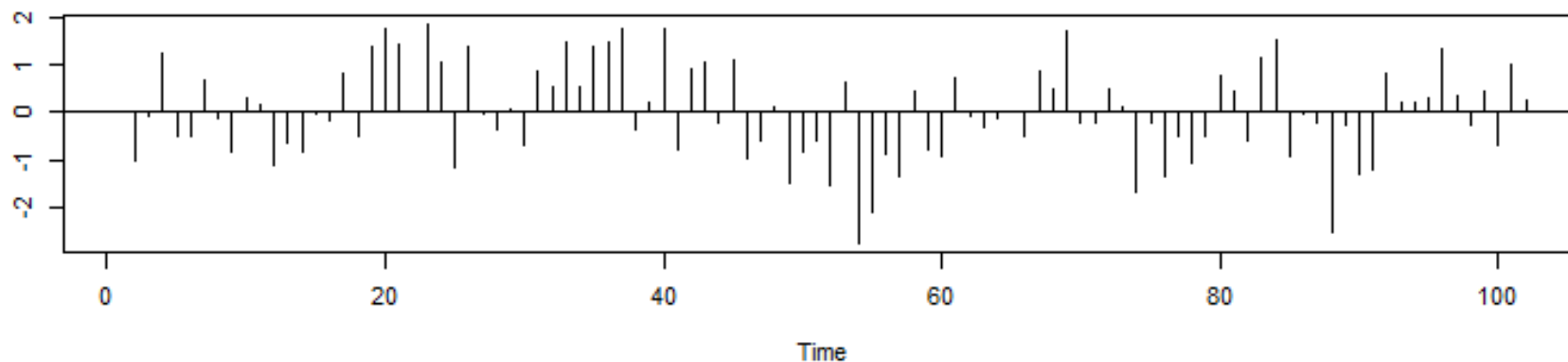
sigma^2 estimated as 0.00262: log likelihood=157.76
AIC=-301.52 AICc=-300.32 BIC=-283.22

標準化残差。

ホワイトノイズになっていることが期待される。

この例では周期性がみられる

Standardized Residuals

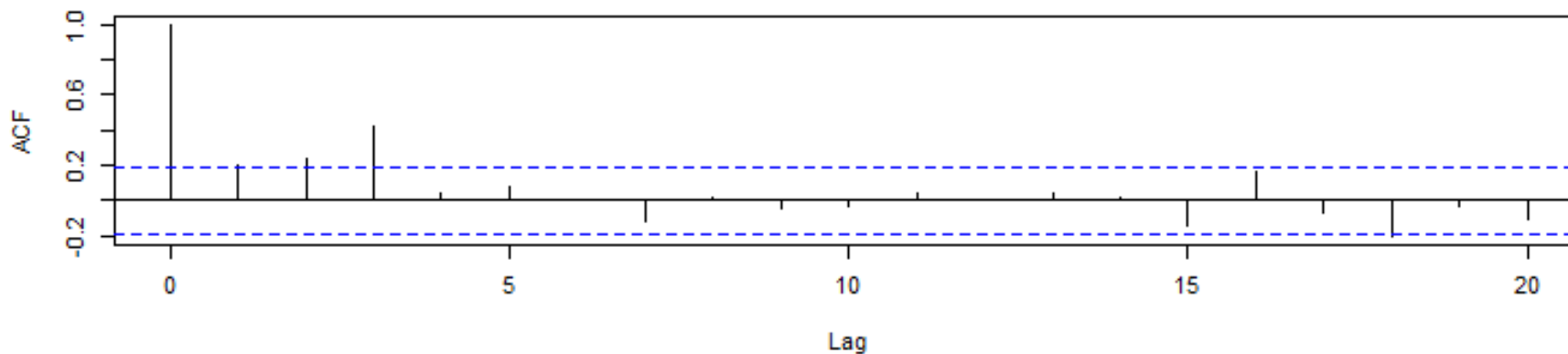


残差コレログラム。

ラグ1以降が0に近いことが期待される。

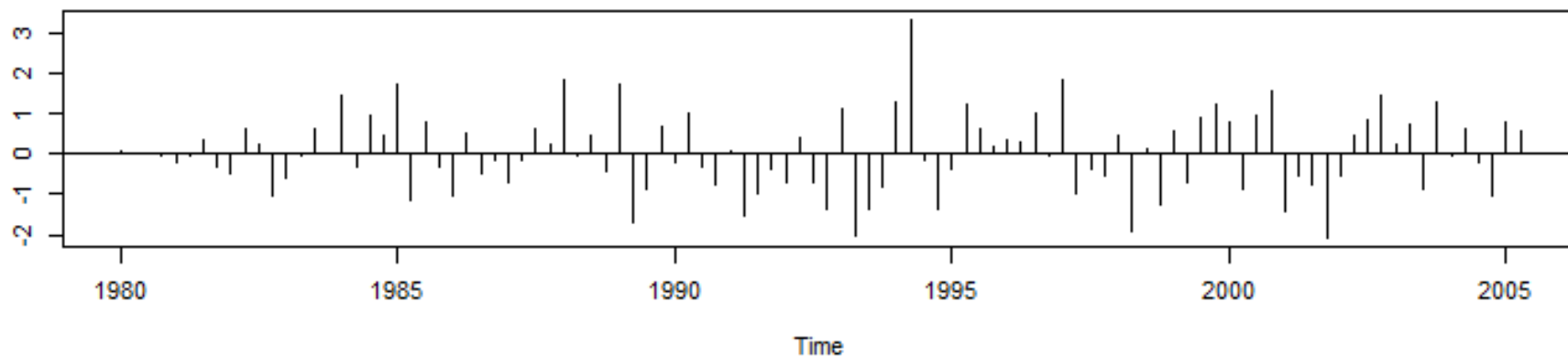
この例ではラグ4で高い。季節変動が示唆される

ACF of Residuals



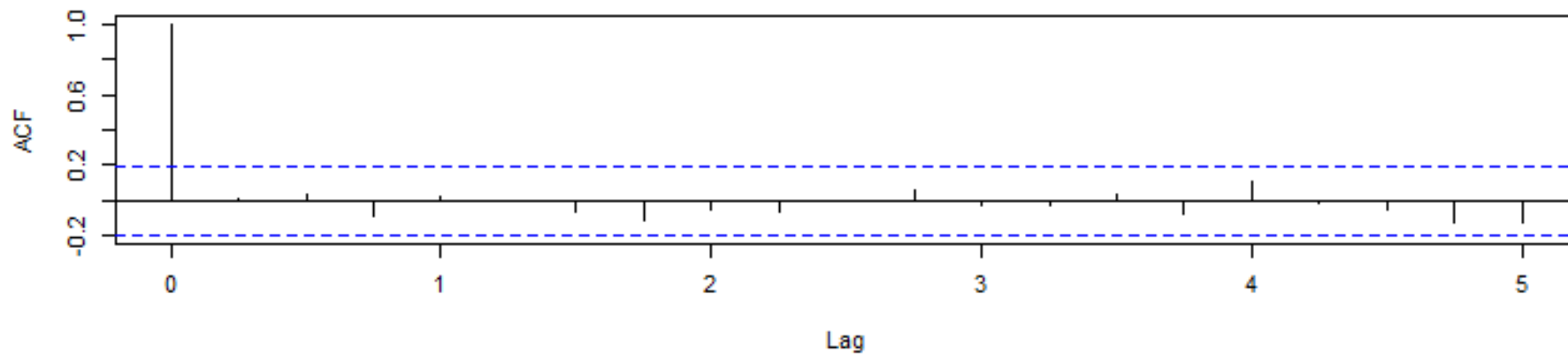
標準化残差。ホワイトノイズになっていることが期待される

Standardized Residuals



残差コレログラム。ラグ1以降が0に近いことが期待される

ACF of Residuals



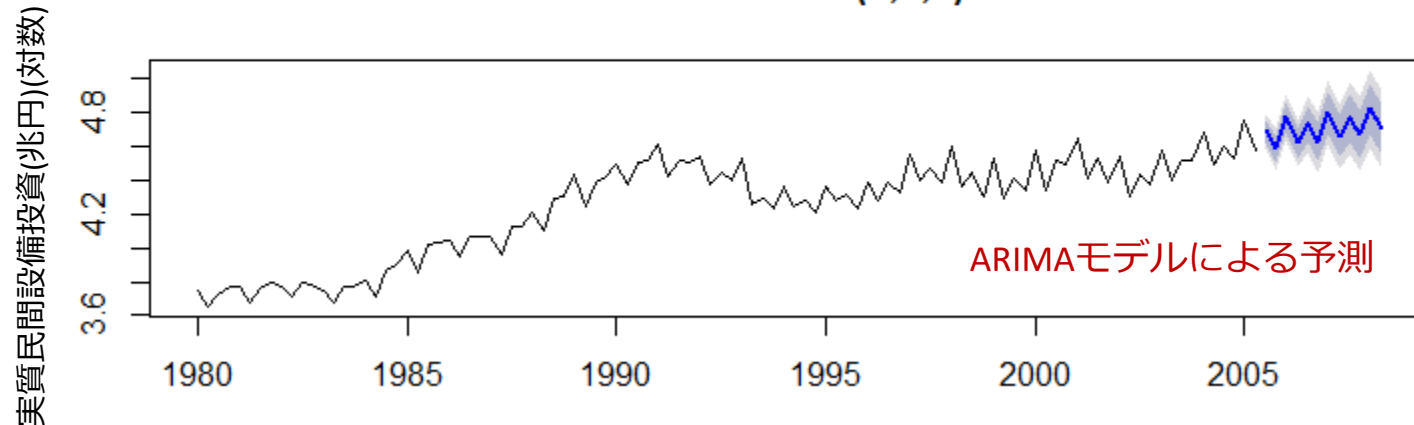
6-5. 短期的予測 (沖本 pp.56-72)

ARMAモデル(ARIMAモデル) によって、短期的予測が可能

民間設備投資

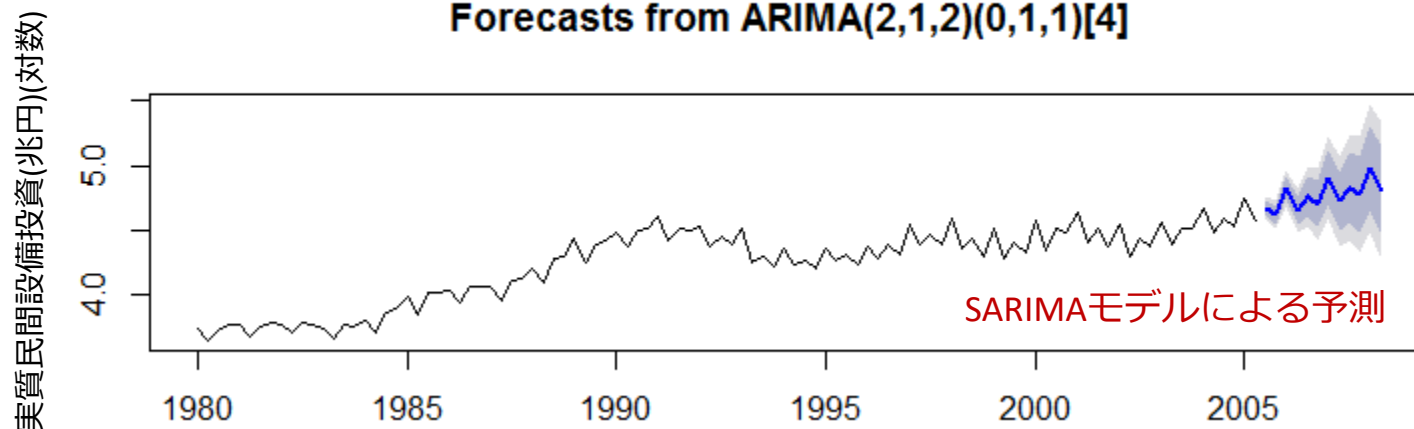
Forecasts from ARIMA(3,1,2) with drift

Code 15



Forecasts from ARIMA(2,1,2)(0,1,1)[4]

Code 16



■ まとめ：1 変量時系列のモデル構築

1. 時系列の観察
2. 時系列の変換
3. 定常性のチェック
4. モデルのあてはめ
5. 短期的予測

7. 時系列の回帰モデル

本節では、**時系列どうしの回帰モデルを構築する**という問題を考えます

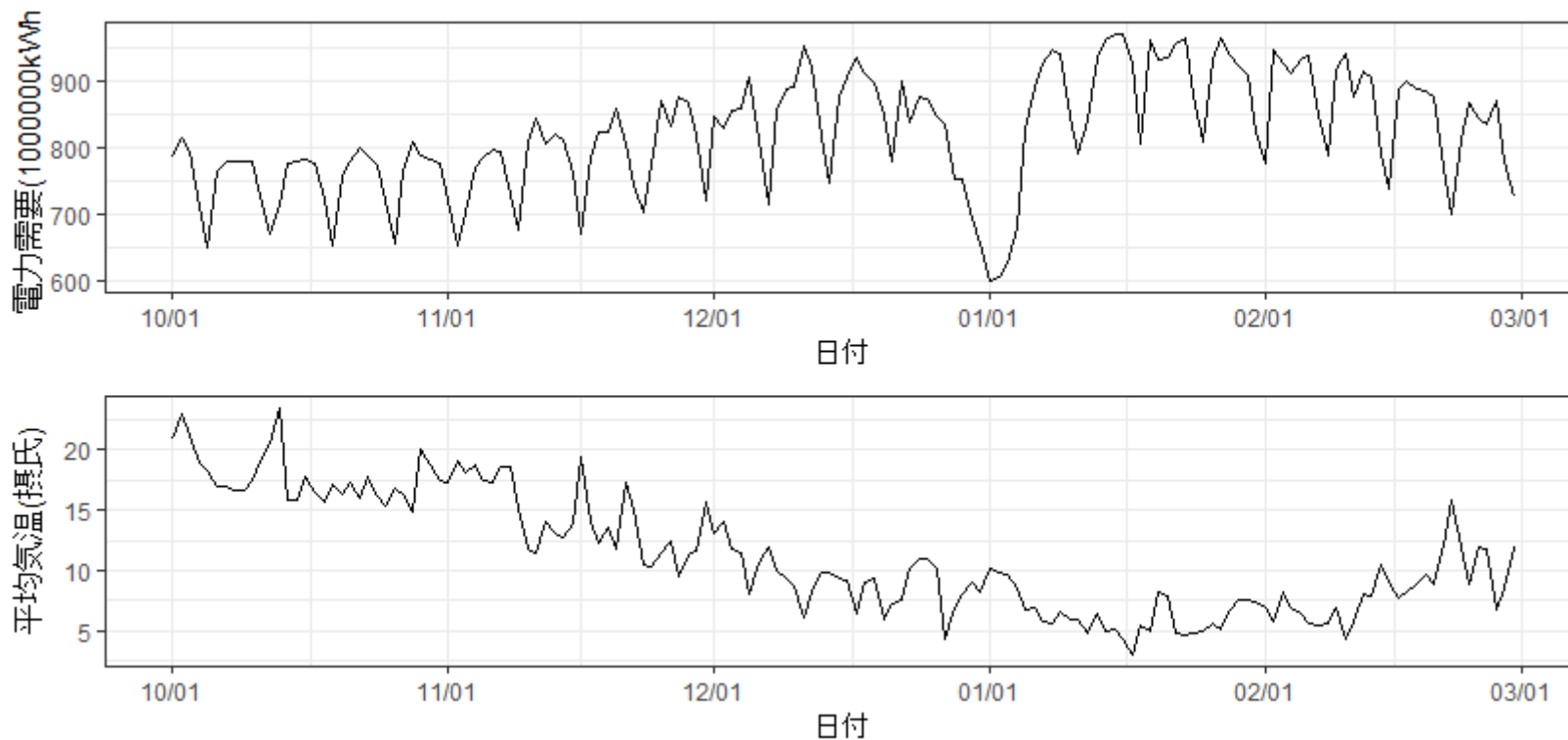
ただし、

- 説明変数の効果が即時的である場合について考えます
 - ある説明変数の値は、目的変数の同期の値にのみ影響する
- 説明変数の効果が時間を通じて一定である場合について考えます
 - (偏) 回帰係数は定数

■ サンプルデータ: 電力需要

Code 17

東京電力がカバーしているエリアの電力需要と、東京の気温のデータ。
2003/10/01-2004/02/28 の日時データ。



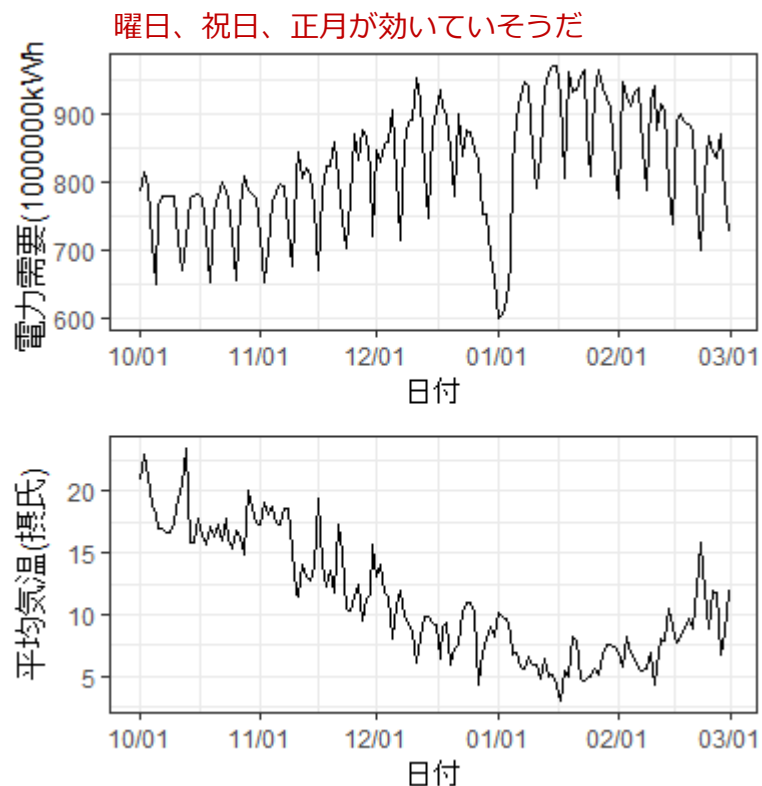
■ 時系列どうしの回帰分析の手順

1. 時系列の観察
2. 時系列の変換
3. 定常性のチェック
4. 回帰モデルの構築
5. 自己相関のチェック
6. 自己相関のモデル化
7. 解釈

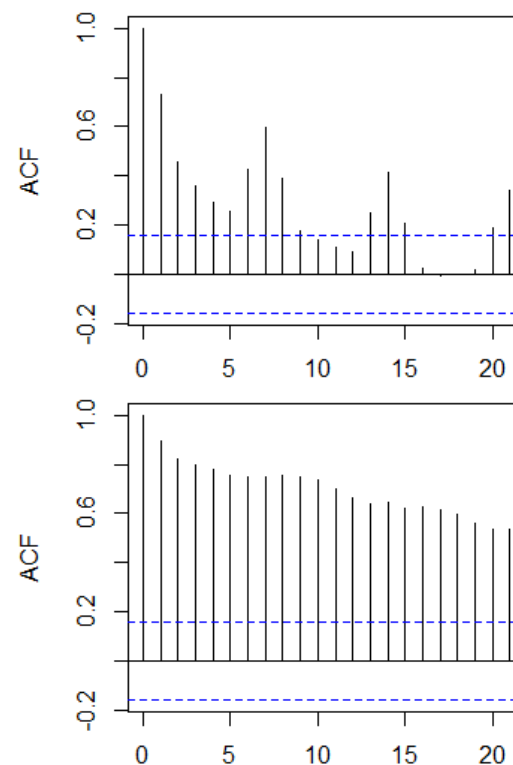
7-1. 時系列の観察

- トレンドはあるか？
- 季節変動は？
- なんらかの説明が可能な変動は？
- 分散は均一か？

電力需要



Code 17



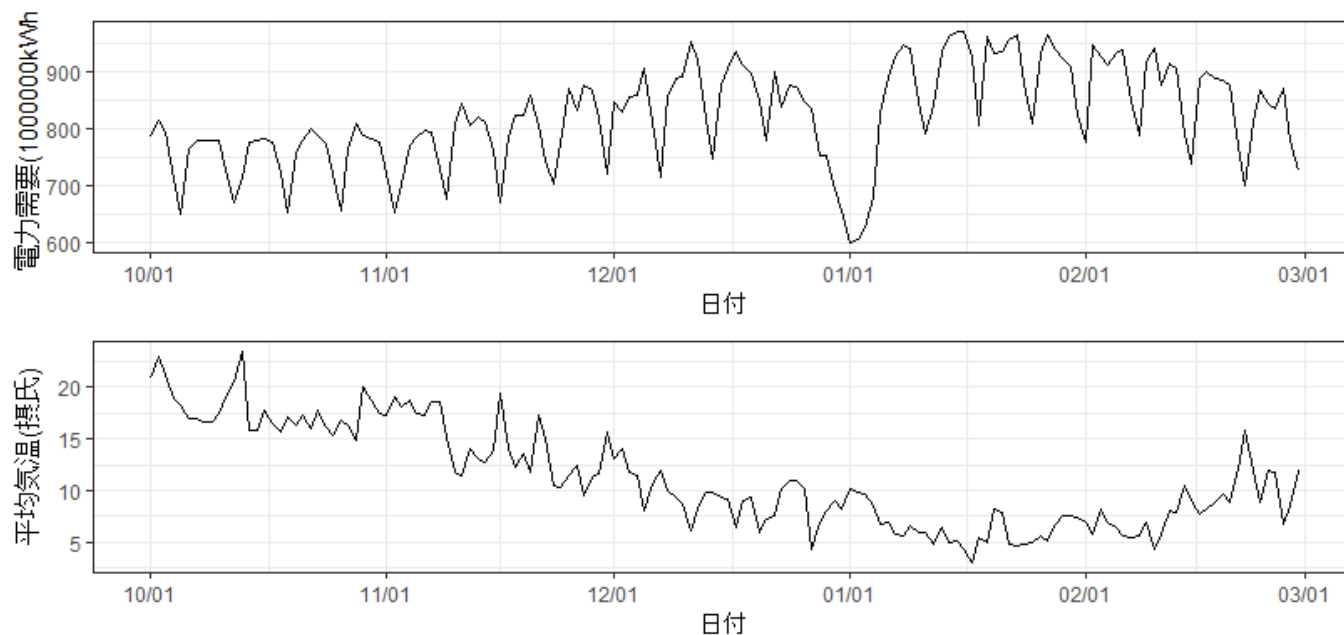
7-2. 時系列の変換

■ 実質的な観点から
IV章を参照

■ 定常な時系列に近づけるため
本章6節を参照

Code 17

電力需要



ま、原系列でいいか...

■ すべての時系列が単位根検定を通過することを確認する

■ なぜ？

- 目的変数が非定常である場合
 - 攪乱項の平均・分散が一定でないかもしれない
→ 回帰モデルの仮定[2][3]に反する

- 説明変数が非定常である場合
 - 回帰モデルのパラメータ推定量が
一致性を持たない
 - 実は、回帰モデルのパラメータ
推定量が一致性を持つ条件の
ひとつは「説明変数が定常
であること」





■ 最小二乗推定量 $\hat{\alpha}, \hat{\beta}$ の性質 (穀谷, pp.27-37)

- Y_i の線形関数である (=データ Y_i の重みつき合計として表現できる)

$$v_i = \sum_{i=1}^n \left(\frac{1}{n} - \bar{X} w_i \right), \quad w_i = \frac{(X_i - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2},$$

として

$$\hat{\alpha} = \sum_{i=1}^n v_i Y_i, \quad \hat{\beta} = \sum_{i=1}^n w_i Y_i$$

- 仮定[1][2]のもとで、 α, β の不偏推定量である
 - 推定量 $\hat{\alpha}, \hat{\beta}$ の期待値は α, β である
- 仮定[1][2] (と X に関するある仮定)の下で、 α, β の一致推定量である
 - 推定量 $\hat{\alpha}, \hat{\beta}$ は、観測点の数が無限に大きければ、 α, β に一致する

Insight Factory Inc. 41

- 特に、目的変数と説明変数の両方が単位根を持つときは、独立であっても高い相関を示す (「**見せかけの回帰**」と呼ばれる現象)
 - 本章1節で紹介した

■ 単位根を持つ時系列があった場合の処置

- すべての時系列について1階差分をとり、単位根検定をやりなおす
 - 次頁を参照
- 共和分分析 (沖本 pp.129-134; HPS pp.283-285)
 - 単位根を持つ時系列同士がある条件を満たしているとき、その関係をモデリングする手法
 - 本セミナーでは割愛
- 単位根過程を扱える分析手法に乗り換える → VII章

■ 目的変数と説明変数のうちどちらかだけが単位根を持っていたら、どうするか？

- すべて差分をとった場合: ARIMA誤差回帰モデル (後述)

$$Y_t = \alpha + \beta X_t + E_t$$

- X_t が1増えると、 Y_t は β 増える
- 解釈の際には、差分をとったことは気にしなくてよい
- 弊害：定常時系列についても差分をとることになる

- 説明変数のみ差分をとった場合

$$Y_t = \alpha + \beta(X_t - X_{t-1}) + E_t$$

- X_t が1増えると、 Y_t は β 増え、 Y_{t+1} は β 減る
- Y_t が売上ならば、「 X_t を増やすと、当期の売上は増えるが、次期の売上はその分減る」ことを意味する
- X_t がマーケティング・ミックス変数の場合、ふつう不適切
 - 実売価格ならありうるかもしれないが...

- 目的変数のみ差分をとった場合

$$Y_t - Y_{t-1} = \alpha + \beta X_t + E_t$$

- X_t が1増えると、 $Y_t, Y_{t+1}, Y_{t+2}, \dots$ がすべて β 増える
- Y_t が売上ならば、「 X_t を増やすと、当期以降の売上が一律に増える」ことを意味する
- X_t がマーケティング・ミックス変数の場合、ふつう不適切
 - 新ブランド上市時のブランド広告ならありうるかもしれないが...

電力需要

いずれの時系列も単位根検定を通過した

Code 18

```
> ndiffs(dfDenryoku$gDemand, test = "adf", type = "trend")  
[1] 0  
  
> ndiffs(dfDenryoku$gTemper, test = "adf", type = "trend")  
[1] 0
```



■ 時系列データであることを気にせずに、時系列同士の回帰モデルを構築する

■ なぜ？

- おさらい：時系列同士の回帰分析を行ってはいけない理由
 - 時系列は隣り合う2時点の間で相関を持つかもしれない
→ 回帰モデルの仮定 [4] に反する
 - 時系列は定常でないかもしれない
→ 回帰モデルの仮定 [2][3] に反する
 - ある時点の説明変数が、その時点よりも後の時点の目的変数に対し
て効果を持つかもしれない
→ 回帰モデルの仮定 [1] に反する
- ← 仮定[4]が失われても、推定量は一致性を持つ
- ← チェック済
- ← 本節では即時的効果しかないと仮定している

電力需要

モデル0: (福地・伊藤, 2011) :

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \beta_5 X_{5t} + \beta_6 X_{6t} + \beta_7 X_{7t} + U_t$$

Y_t : 電力需要(百万kWh)

X_{2t} : 平均気温(摂氏)

X_{3t} : 月曜ダミー (月曜日にのみ1となるダミー変数)

X_{4t} : 土曜ダミー

X_{5t} : 日曜ダミー

X_{6t} : 祝日ダミー

X_{7t} : 正月ダミー (12/30-1/3)

※福地・伊藤(2011) では12/31-1/3を正月としていますが、一日伸ばしました

Call:

```
lm(formula = gDemand ~ gTemper + bMon + bSat + bSun + bHoliday +  
    bShogatu, data = dfDenryoku)
```

Residuals:

Min	1Q	Median	3Q	Max
-116.434	-15.659	-0.808	16.562	83.491

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	994.3924	6.0448	164.504	< 2e-16 ***
gTemper	-11.5438	0.4639	-24.885	< 2e-16 ***
bMon	-32.9638	6.9163	-4.766	4.52e-06 ***
bSat	-71.9321	6.5546	-10.974	< 2e-16 ***
bSun	-134.4875	6.5975	-20.384	< 2e-16 ***
bHoliday	-35.9926	12.9406	-2.781	0.00613 **
bShogatu	-236.1801	12.6029	-18.740	< 2e-16 ***

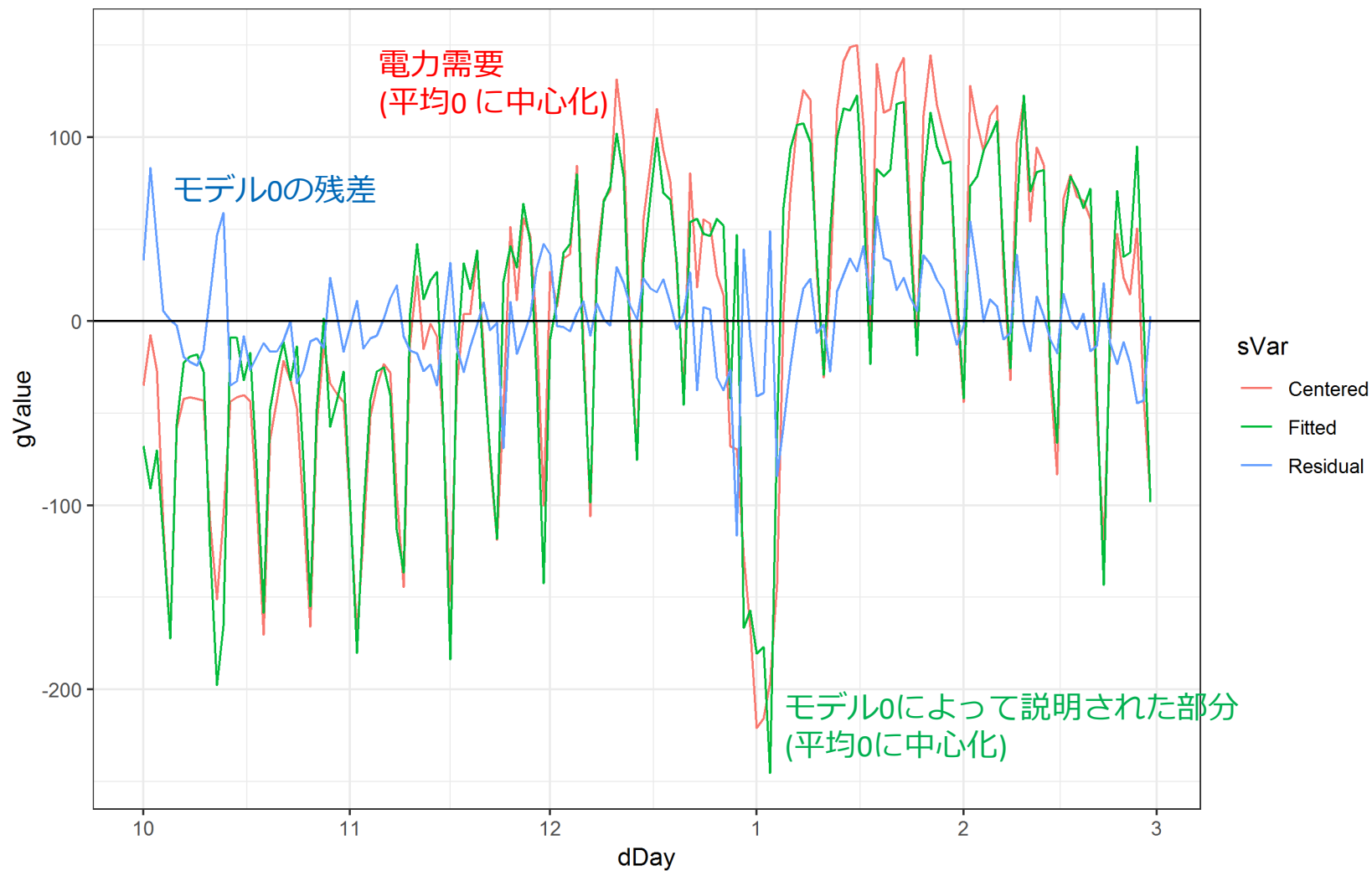
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 27.46 on 145 degrees of freedom

Multiple R-squared: 0.9052, Adjusted R-squared: 0.9012

F-statistic: 230.6 on 6 and 145 DF, p-value: < 2.2e-16

平均気温が1度上がると、
電力需要が11.54 百万kWh下がる



7-5. 自己相関のチェック

■ 残差の自己相関をチェックする

■ なぜ？

- おさらい：時系列同士の回帰分析を行ってはいけない理由
 - 時系列は隣り合う2時点の間で相関を持つかもしれない
→ 回帰モデルの仮定 [4] に反する

仮定[4]が失われているか
どうかを知りたい
 - 時系列は定常でないかもしれない
→ 回帰モデルの仮定 [2][3] に反する

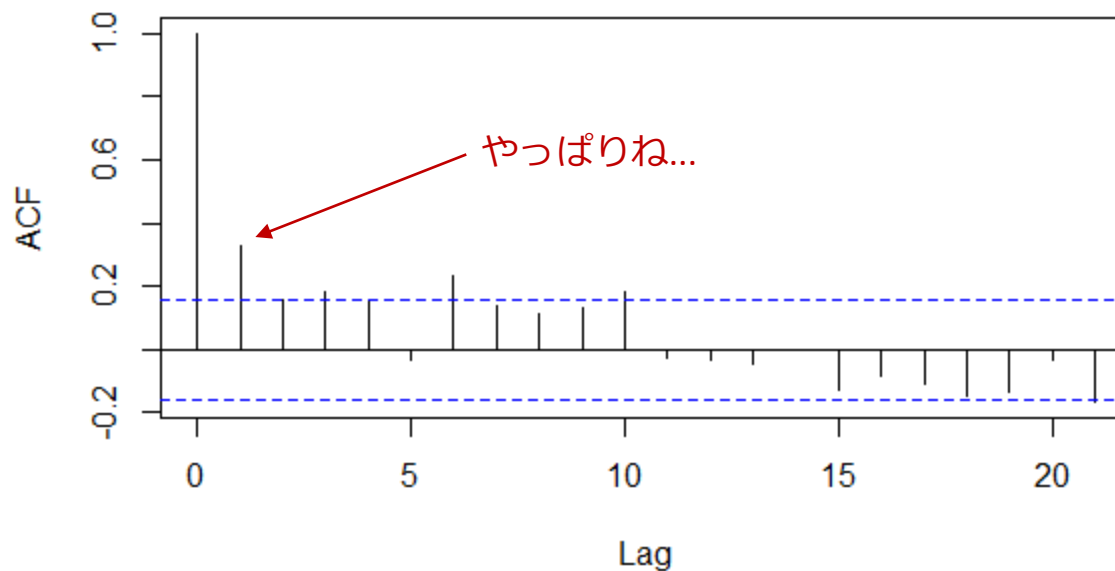
チェック済
 - ある時点の説明変数が、その時点よりも後の時点の目的変数に対し
て効果を持つかもしれない
→ 回帰モデルの仮定 [1] に反する

本節では
即時的効果しかないと
仮定している

■ まずは残差コレログラムを確認しておこう

電力需要

Code 19



■ 自己相関の検定

- Durbin-Watson 検定 ... 自己相関の検定の古典
- Durbinのh検定
- Breusch-Godfrey検定
 - 帰無仮説：「誤差項が $AR(p)$ の形の自己相関を持っていない」

電力需要

1次のBreusch-Godfrey検定を試してみると...

Code 19

```
> bgtest(oModel)

Breusch-Godfrey test for serial correlation of order up to 1

data:  oModel
LM test = 17.533, df = 1, p-value = 2.824e-05
```

有意。そりゃそうでしょうね

7-6. 自己相関のモデル化

■ 回帰モデルに攪乱項の自己相関を組み込む

■ 方法

(1) ARMA誤差回帰モデル

(2) ARIMA誤差回帰モデル

(3) 実行可能一般化最小二乗法 (FGLS)

(1) ARMA誤差回帰モデル

■ おさらい：ARMAモデルとは

$$Y_t = \alpha + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

ラグ演算子を使って書きなおすと

$$Y_t = \alpha + \frac{1 + \theta_1 L + \dots + \theta_q L^q}{1 - \phi_1 L - \dots - \phi_p L^p} \varepsilon_t$$

ARMA撓乱項

■ ARMA誤差回帰モデル

$$Y_t = \beta_1 + \beta_2 X_{2t} + \dots + \beta_K X_{Kt} + \frac{1 + \theta_1 L + \dots + \theta_q L^q}{1 - \phi_1 L - \dots - \phi_p L^p} \varepsilon_t$$

通常の重回帰モデル

ARMA撓乱項

ARMAモデルと同様に、最尤法で推定する

- (ARMAの次数の指定が正しければ) 一致性と漸近的有効性を持つ

(2) ARIMA誤差回帰モデル

■ ARIMA誤差回帰モデル

- すべての変数($Y_t, X_{2t}, \dots, X_{Kt}$)についてd回差分をとってARMA誤差回帰モデルを適用すること
- 差分系列の回帰モデルを構築していた場合に便利
- ARIMA(p,1,q) 誤差回帰モデルは

$$Y_t = \beta_0 + \beta_1 t + \beta_2 X_{2t} + \dots + \beta_K X_{Kt} + \boxed{E_t} \text{ ARIMA撹乱項}$$

ただし

$$(1 - L)E_t = \frac{1 + \theta_1 L + \dots + \theta_q L^q}{1 - \phi_1 L - \dots - \phi_p L^p} \varepsilon_t \text{ ARMA撹乱項}$$

- ARMA誤差回帰モデルとARIMA誤差回帰モデルのちがいは
 - 「値の関係についてのモデル」と「変化の関係についてのモデル」のちがいであるともいえるし、
 - 単に撹乱項がちがうだけだともいえる
- 偏回帰係数を解釈する際には、ARMA誤差回帰モデルとARIMA誤差回帰モデルのちがいは気にしなくてよい

(3) 実行可能一般化最小二乗法 (FGLS)

- 最小二乗法を拡張した方法。自己相関があるデータについて推定できる
 - 攪乱項の自己相関の構造を指定する
 - 構造としてARMAを指定すれば、ARMA誤差回帰モデルを推定したことになる
 - (ARMAの次数の指定が正しければ) 一致性と有効性を持つ

電力需要

Code 19

モデル1: AR(1) の場合

Series: dfDenryoku\$gDemand
Regression with ARIMA(1,0,0) errors

Coefficients:

	ar1	intercept	gTemper	bMon	bSat	bSun	bHoliday	bShogatu
	0.7827	952.1809	-8.3642	-24.4706	-69.1524	-128.9732	-36.4822	-107.1146
s.e.	0.0619	14.5001	1.0067	5.1014	4.7706	5.6495	8.5043	20.4405

平均気温が1度上がると、
電力需要が8.36 百万kWh下がる

sigma^2 estimated as 572.5: log likelihood=-694.64
AIC=1407.28 AICc=1408.55 BIC=1434.5

モデル2: forecast::auto.arima()で自動選択したところ、ARMA(5, 1)が選択された

Series: dfDenryoku\$gDemand
Regression with ARIMA(5,0,1) errors

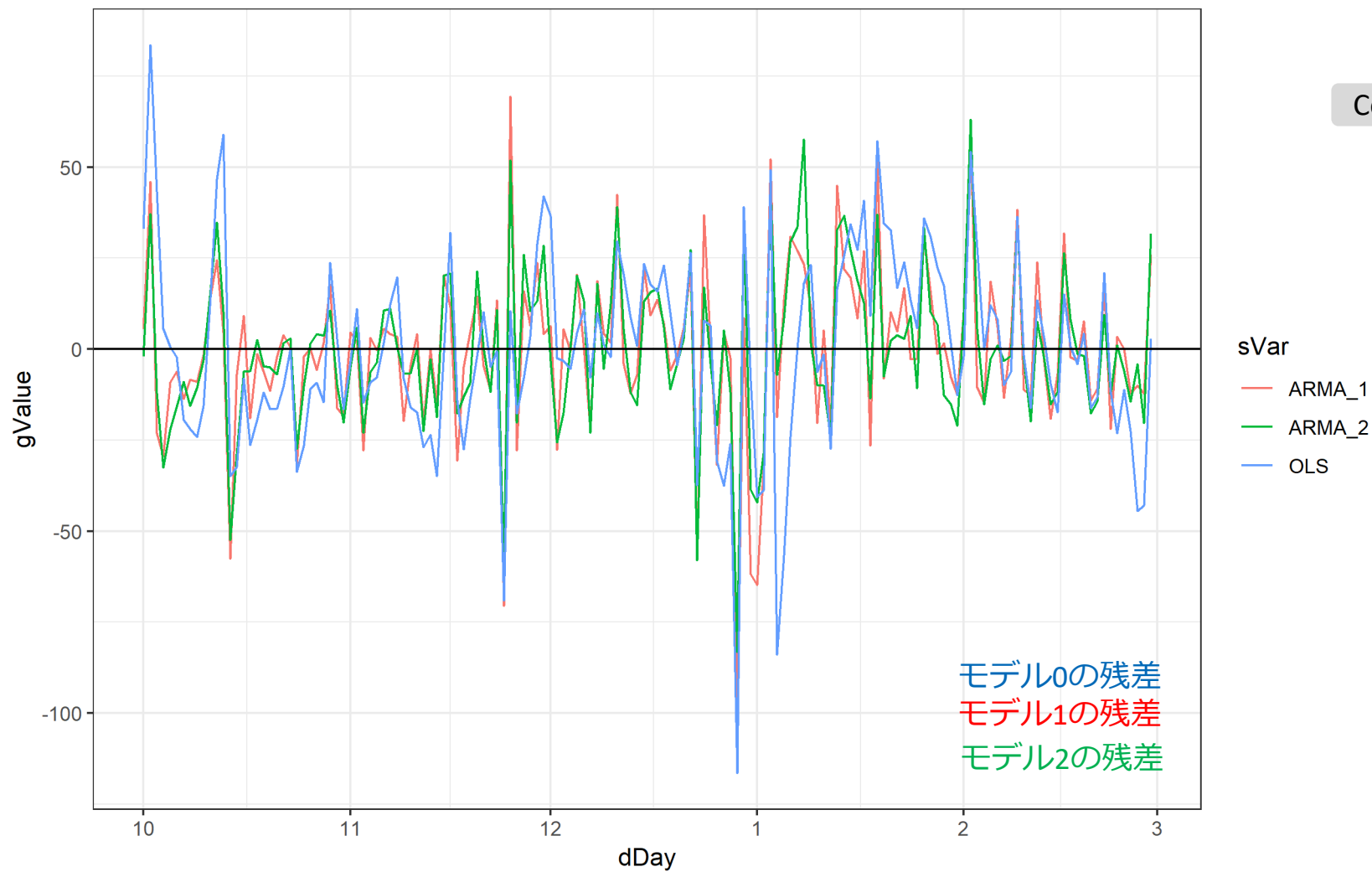
Coefficients:

	ar1	ar2	ar3	ar4	ar5	ma1	intercept	gTemper	bMon
	-0.2618	0.4900	0.3610	0.2796	-0.1552	0.7538	945.8745	-7.4875	-30.7042
s.e.	0.1565	0.1054	0.0944	0.1024	0.0997	0.1255	15.5885	1.0253	4.5078

	bSat	bSun	bHoliday	bShogatu
	-68.4723	-135.3686	-30.1991	-169.6477
s.e.	4.0671	4.6051	8.9254	20.6859

平均気温が1度上がると、
電力需要が7.49 百万kWh下がる

sigma^2 estimated as 512.5: log likelihood=-683.83
AIC=1395.65 AICc=1398.72 BIC=1437.99



モデル3: FGLS, 撓乱項の共分散としてAR(1) を指定

Generalized least squares fit by REML

Model: gDemand ~ gTemper + bMon + bSat + bSun + bHoliday + bShogatu

Data: dfDenryoku

AIC	BIC	logLik
1371.166	1397.957	-676.583

Correlation Structure: AR(1)

Formula: ~1

Parameter estimate(s):

Phi
0.7961851

平均気温が1度上がると、
電力需要が8.27 百万kWh下がる

Coefficients:

	Value	Std. Error	t-value	p-value
(Intercept)	951.0448	14.253143	66.72527	0
gTemper	-8.2712	0.939645	-8.80253	0
bMon	-24.3504	5.162340	-4.71693	0
bSat	-69.1323	4.856087	-14.23621	0
bSun	-128.9388	5.757543	-22.39475	0
bHoliday	-36.5359	8.646249	-4.22564	0
bShogatu	-104.9998	18.241606	-5.75606	0

(中略)

Standardized residuals:

Min	Q1	Med	Q3	Max
-4.10772640	-0.54330393	0.04209911	0.62534388	1.91621487

Residual standard error: 39.40679

Degrees of freedom: 152 total; 145 residual

モデル4: FGLS, 撓乱項の共分散としてARMA(5, 1) を指定

Generalized least squares fit by REML

Model: gDemand ~ gTemper + bMon + bSat + bSun + bHoliday + bShogatu

Data: dfDenryoku

AIC	BIC	logLik
1360.33	1402.004	-666.1648

Correlation Structure: ARMA(5,1)

Formula: ~1

Parameter estimate(s):

Phi1	Phi2	Phi3	Phi4	Phi5	Theta1
-0.2681756	0.5042322	0.3692546	0.2879284	-0.1367366	0.7624480

Coefficients:

	Value	Std.Error	t-value	p-value
(Intercept)	943.7072	16.192804	58.27942	0e+00
gTemper	-7.3042	0.905353	-8.06781	0e+00
bMon	-30.5827	4.398100	-6.95361	0e+00
bSat	-68.5679	4.067206	-16.85872	0e+00
bSun	-135.4897	4.556282	-29.73691	0e+00
bHoliday	-30.6078	8.725392	-3.50790	6e-04
bShogatu	-170.6390	15.137568	-11.27255	0e+00

平均気温が1度上がると、
電力需要が7.30 百万kWh下がる

(中略)

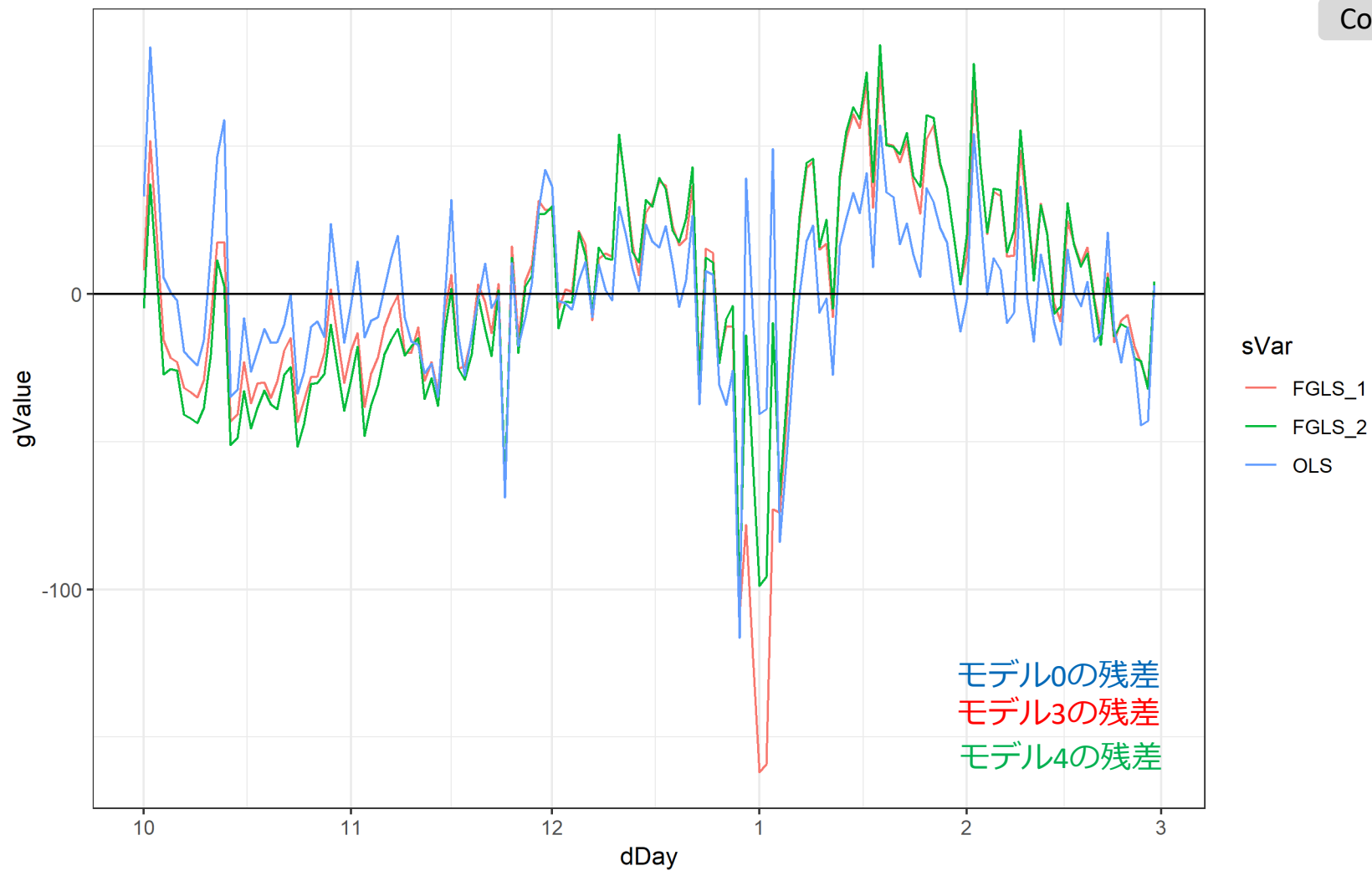
Standardized residuals:

Min	Q1	Med	Q3	Max
-2.76668703	-0.69624219	-0.07090714	0.71660764	2.27464780

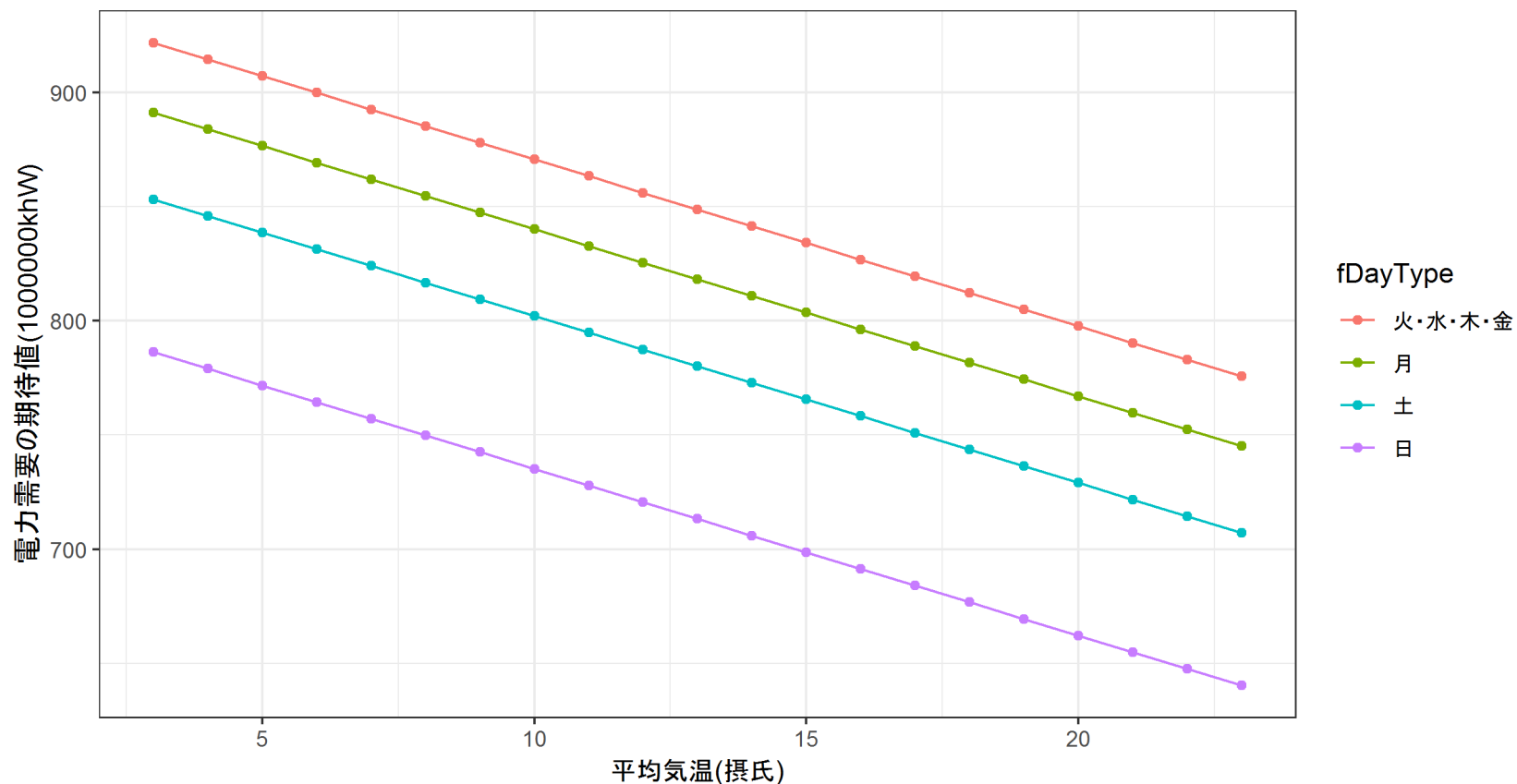
Residual standard error: 37.0372

Degrees of freedom: 152 total; 145 residual

Code 19



モデル4によれば、
他の変数が同じならば、**平均気温が1度上がると、電力需要が7.30 百万kWh下がる**
(95%信頼区間: 5.53 – 9.08)



(祝日・正月を除く期待値を示す)

■ まとめ：時系列どうしの回帰分析

1. 時系列の観察
2. 時系列の変換
3. 定常性のチェック
4. 回帰モデルの構築
5. 自己相関のチェック
6. 自己相関のモデル化
7. 解釈

8. 季節変動の扱い

8. 季節変動の扱い

目的変数に季節変動がみられるとき、どうするか？

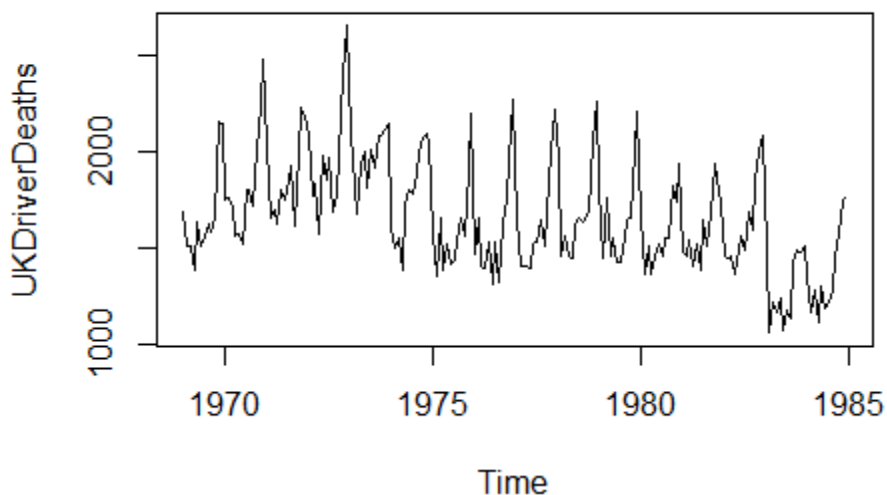
- **変数変換**によって除去する
- なんらかの季節調整法によって除去する
 - お勧めできない。このセミナーでは割愛
- 季節の効果をモデルに組み込む
 - **季節ダミー変数**
 - **季節ARIMAモデル (SARIMAモデル)**
 - **調和回帰**

8-1. 変数変換による除去

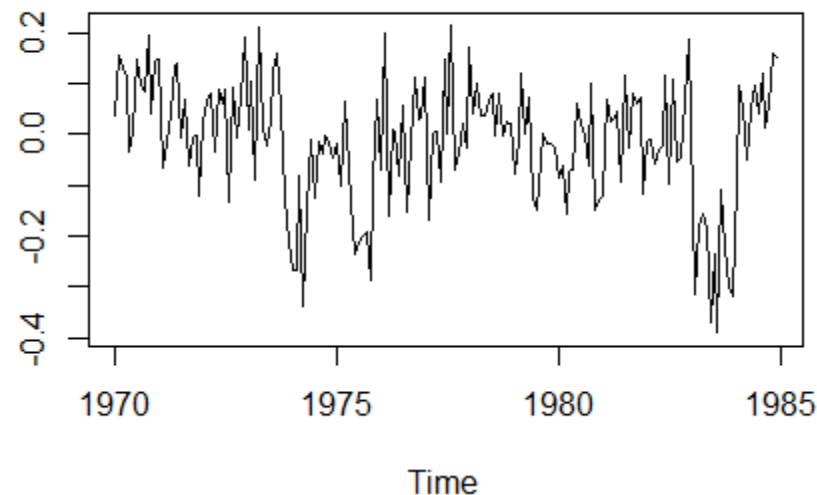
■ 前年同期比

$$\frac{y_t - y_{t-12}}{y_{t-12}} \approx \log y_t - \log y_{t-12} \text{ (月次の場合)}$$

英国自動車事故



diff(log(UKDriverDeaths), lag = 12)



Code 12

■ 季節ダミー変数

- 四半期データの場合
 - {Q2において1となるダミー変数、Q3において1となるダミー変数、Q4において1となるダミー変数}を説明変数として投入する
- 月次データの場合
 - 単純に考えると11個のダミー変数が必要
 - パラメータ数が増えてしまう...
 - 目的変数の時系列を観察し、適当な数のダミー変数にまとめる
 - 例) 「7,8,9月に1となるダミー変数」
- 週次データ、日時データの場合
 - ダミー変数による表現は困難

■ 季節ARIMAモデル (SARIMAモデル)

- ARIMA攪乱項と、「前期」を「前年同期」に読み替えたARIMA攪乱項を合体させる
- ARMA(1,1) 攪乱項 (=ARIMA(1,0,1) 攪乱項)

$$E_t = \varepsilon_t + \phi E_{t-1} + \theta \varepsilon_{t-1}$$

$$E_t = \frac{1 + \theta L}{1 - \phi L} \varepsilon_t$$

- SARIMA(1,0,1, 1,0,1) 攪乱項 ※月次の場合

$$E_t = \frac{(1 + \theta_s L^{12}) (1 + \theta L)}{(1 - \phi_s L^{12}) (1 - \phi L)} \varepsilon_t$$

- 特徴
 - 季節効果を確認的に捉えている。季節効果そのものは推定しない
 - パラメータ数が少ない
 - 週次データには向かない。「前年同期」の定義が難しいから

■ SARIMA(p,d,q,k,l,m)の推定

民間設備投資

```
# SARIMA(2, 1, 2, 1, 1, 1)の推定
Arima(
  LIP,
  order = c(2, 1, 2),
  seasonal = list(order = c(1,1,1), period = 4),
  method = "ML"
)
```

```
Series: LIP
ARIMA(2,1,2)(1,1,1)[4]

Coefficients:
          ar1          ar2          ma1          ma2          sar1          sma1
      0.9742   -0.6440   -1.1279    0.9367    0.1518   -0.5818
s.e.  0.1685    0.1064    0.0915    0.0644    0.2397    0.1980

sigma^2 estimated as 0.00177:  log likelihood=171.75
AIC=-329.49   AICc=-328.24   BIC=-311.47
```

■ SARIMA(p,d,q,k,l,m)の推定 (手っ取り早いバージョン)

民間設備投資

```
auto.arima(  
  LIP, d = 1,  
  trace=T, seasonal = T, ic = "aic", stepwise=F, approximation = F  
)
```

(中略)

Best model: ARIMA(2,1,2)(0,1,1)[4]

Series: LIP

ARIMA(2,1,2)(0,1,1)[4]

季節を取り除いた部分についてARIMA(2,1,2),
季節についてARIMA(0, 1, 1)がお勧め

Coefficients:

	ar1	ar2	ma1	ma2	sma1
	1.0031	-0.6347	-1.1523	0.9315	-0.4625
s.e.	0.1762	0.1055	0.0961	0.0536	0.1053

sigma^2 estimated as 0.001761: log likelihood=171.56

AIC=-331.12 AICc=-330.19 BIC=-315.67

■ 調和回帰 harmonic regression

- モデルにフーリエ級数の項を含める：

$$Y_t = \sum_{i=1}^S \left(a_i \cos \frac{2\pi i t}{n} + b_i \sin \frac{2\pi i t}{n} \right) + \dots$$

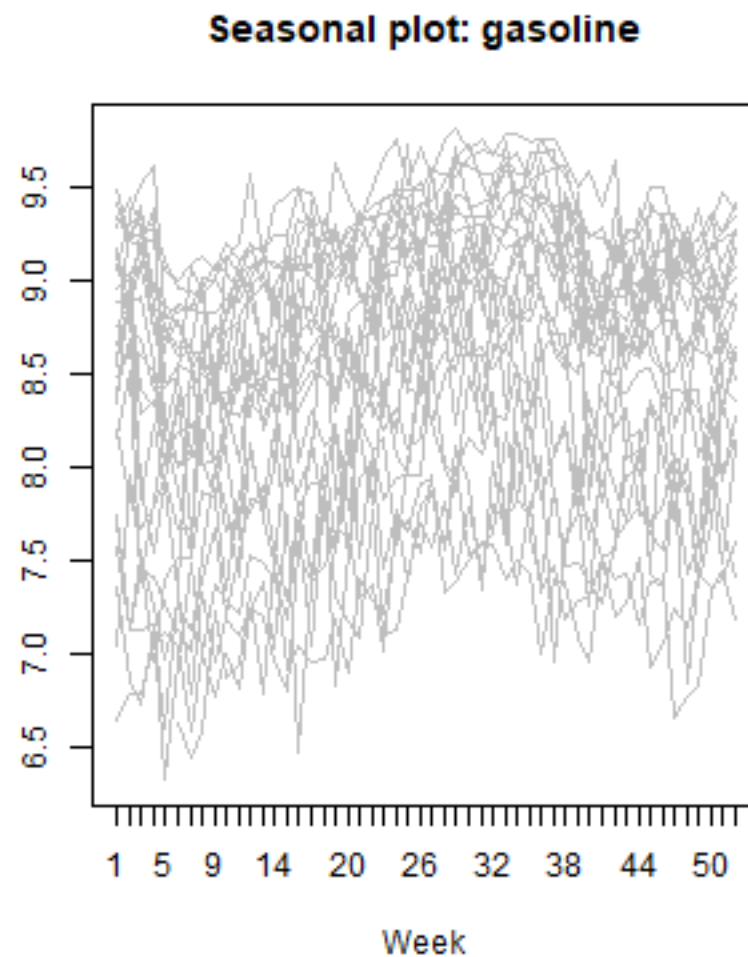
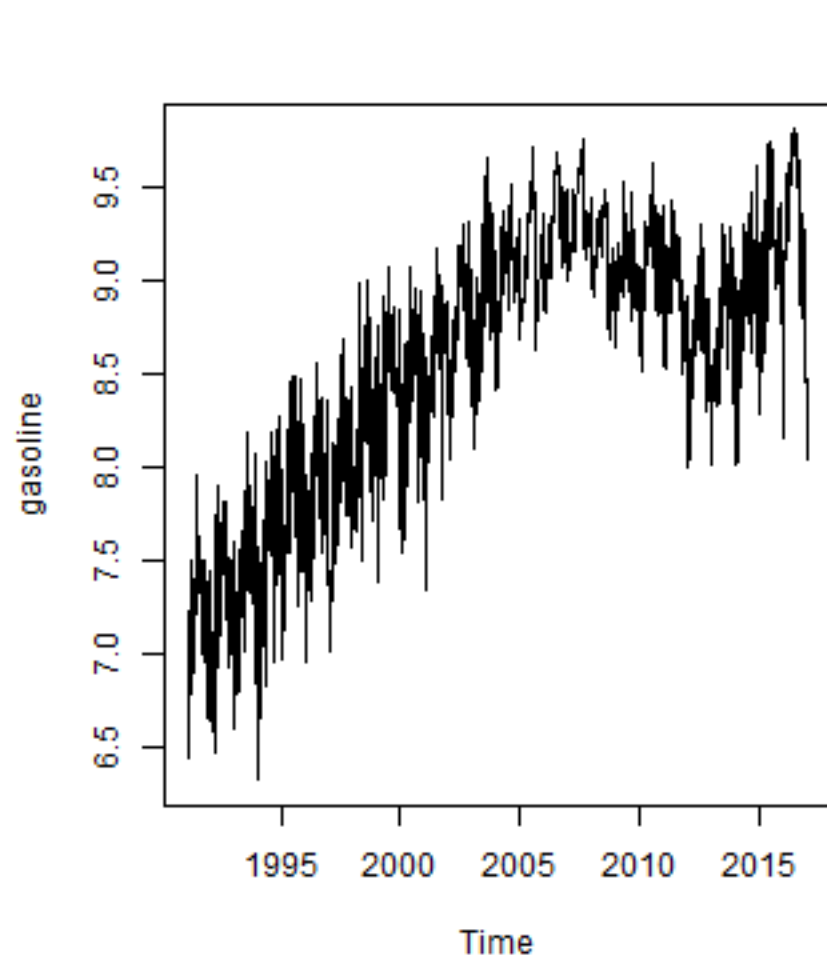
ARIMAモデルや
ARIMA誤差回帰モデル

ただし n は年あたり時点数(週次データの場合 52.18)。

- S を大きくすると、季節変動をより細やかに表現できる
- $2S$ 個のパラメータが追加される

- 例: usガソリン供給量

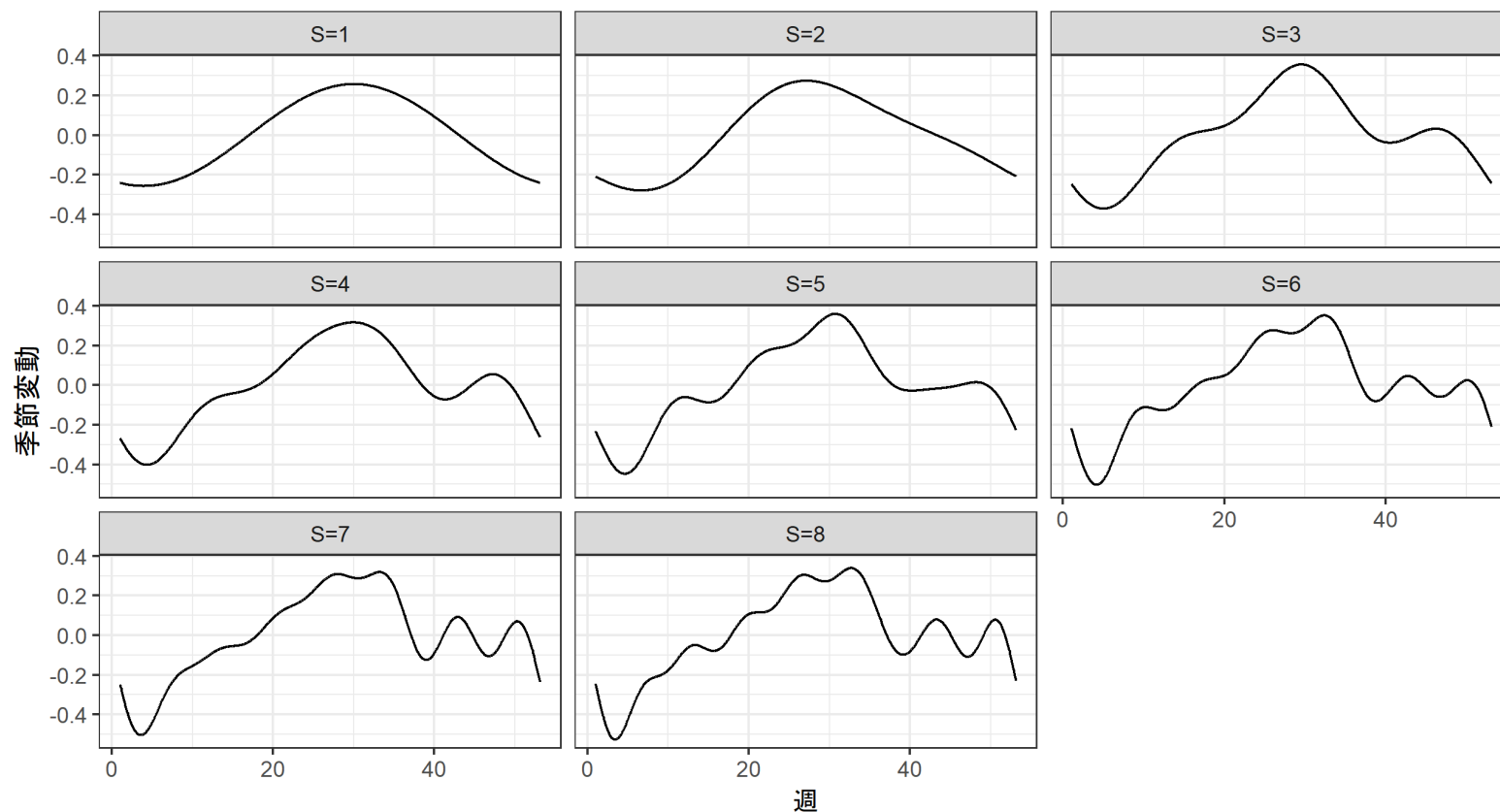
Code 20



フーリエ級数つきARIMA(1,1,1)モデル

$$Y_t = \sum_{i=1}^S \left(a_i \cos \frac{2\pi i t}{n} + b_i \sin \frac{2\pi i t}{n} \right) + \mu_0 + \mu_1 t + E_t, \quad (1-L)E_t = \frac{1-\theta_1 L}{1-\phi_1 L} \varepsilon$$

をあてはめた。季節効果の推定値を示す。



田中孝文(2008)「Rによる時系列分析入門」,シーエーピー出版.

福地純一郎・伊藤有希(2011)「Rによる計量経済分析」,朝倉書店,

黒住英司(2008) 経済時系列分析と単位根検定：これまでの発展と今後の展望. 日本統計学会誌, 38(1), 39-57.

この章に登場したRの関数

- `acf`(ベクトル)
 - 標本コレログラムを表示する
- `forecast::seasonplot`(tsオブジェクト)
 - 季節を横軸に取ったチャートを表示する
- `arma.sim`(model = list(AR=c(AR係数), MA=c(MA係数)), n=長さ)
 - ARMA過程に従う時系列データを生成する
- `urca::ur.df`(ベクトル, type=モデルのタイプ, lags=ARモデルの次数)
 - ADF検定
- `forecast::ndiffs`(ベクトル, test=検定のタイプ, type=モデルのタイプ)
 - 単位根検定によって何階差分が定常になるかを判定して表示する
- `forecast::Arima`(ベクトル, order=c(AR次数, 差分の階数, MA次数), xreg=説明変数行列)
 - ARIMAモデルをあてはめ、Arimaオブジェクトを返す
- `forecast::auto.arima`(ベクトル, , xreg=説明変数行列)
 - ARIMAモデルを自動的にあてはめ、Arimaオブジェクトを返す
- `forecast::forecast`(Arimaオブジェクト)
 - ARIMAモデルによる将来予測
- `lmtest::bgtest`(lmオブジェクト)
 - Breusch-Godfrey検定
- `tsdiag`(時系列モデルオブジェクト)
 - 時系列モデルの診断
- `nlme::gls`(フォーミュラ, corr=corARMA(p=AR次数, q = MA次数), data=データフレーム)
 - ARMA誤差を仮定した回帰のFGLS推定

Appendix 1. Rの日付クラス

- 日付クラスとは
 - 日付ベクトルを表すデータ構造
 - 日付データは、文字列 (例, “2019-03-01”) や数値 (例, 20190301) ではなく、日付クラスのオブジェクトにしたほうがはるかに便利
- Rの日付クラス
 - **Date クラス**
 - 標準的な日付クラス
- Rの時間クラス
 - POSIXctクラス, POSIXltクラス
 - 年月日時分秒を表す
 - 説明は割愛します

- Dateクラスへの変換
 - as.Date (文字列ベクトル)
 - 日付を表すたいていの文字列ベクトルをうまくこと変換してくれる

```
> x <- as.Date(c("2019/03/01", "2019/04/01"))
```

yyyy/mm/dd 形式

```
> x  
[1] "2019-03-01" "2019-04-01"
```

いっけん
文字列ベクトルに
みえるが...

```
> str(x)  
Date[1:2], format: "2019-03-01" "2019-04-01"
```

実はDate型

```
> as.Date(c("2019-03-01", "2019-04-01"))  
[1] "2019-03-01" "2019-04-01"
```

yyyy/mm/dd 形式でも
うまくいく

- lubridate::ymd (文字列ベクトルないし数値ベクトル)
 - 年(y), 月(m), 日(d) をこの順で表している文字列ベクトル・数値ベクトルなら、たいていうまく変換してくれる
 - 他に、ydm(), mdy(), myd(), dmy(), dym(), yq() など、さまざまな関数がある

```
> library(lubridate)

> x <- ymd(c("2019/03/01", "2019/04-01"))

> x
[1] "2019-03-01" "2019-04-01"

> str(x)
Date[1:2], format: "2019-03-01" "2019-04-01"

> ymd(c(20190301, 20190401))
[1] "2019-03-01" "2019-04-01"
```

← yyyy/mm/dd 形式

← いっけん
文字列ベクトルに
みえるが...

← 実はDate型

← なんと、こういう数値でも、
なんとかしてくれる。すごい

- Dateベクトルの演算・加工
 - たいていの処理は、標準的関数、ないしlubridateパッケージの関数として用意されている
 - 自力で書こうと思わず、まずは検索！

```
> library(lubridate)
> x <- ymd(c("2019/03/01", "2019/04-01"))

> year(x) # 年を取り出す
[1] 2019 2019

> month(x) # 月を取り出す
[1] 3 4

> day(x) # 日を取り出す
[1] 1 1

> wday(x) # 曜日を取り出す
[1] 6 2

> x[2] - x[1] # 差を求める
Time difference of 31 days
```

Appendix 2. Rの時系列クラス

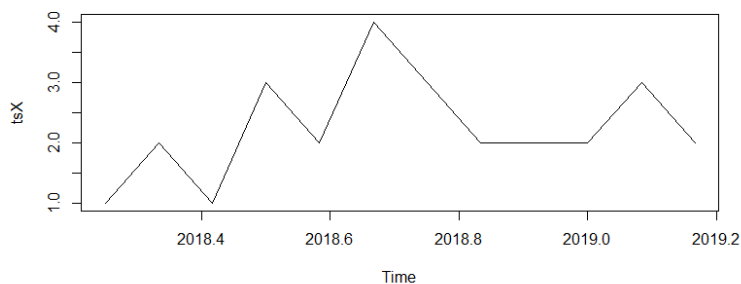
- 時系列クラスとは
 - 時系列専用のデータ構造
 - 時系列データは、通常の数値ベクトル or データフレームではなく、時系列クラスのオブジェクトにしたほうが便利 (かもしれない)
- 現在、Rには時系列クラスを提供するパッケージが乱立している
 - それぞれに特色があり、どれが良いともいいにくい
- 以下では主なものを紹介します
 - 以下の説明では、連続した月次時系列を仮定します

```
## データ例
dfIn <- tibble(
  sMonth = c(
    "2018-04", "2018-05", "2018-06", "2018-07", "2018-08", "2018-09",
    "2018-10", "2018-11", "2018-12", "2019-01", "2019-02", "2019-03"
  ),
  gX = c(1, 2, 1, 3, 2, 4, 3, 2, 2, 2, 3, 2)
)
```

- tsクラス (標準機能)
 - Rの時系列クラスの標準
 - 時系列分析のための多くの関数がtsクラスのオブジェクトを返す
 - ちょっと使いにくいかも...
 - 作成

ts (数値ベクトル, start=(開始年, 開始月), frequency=12)

```
> tsX <- ts(dfIn$gX, start = c(2018, 4), frequency = 12)
> print(tsX)
      Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
2018          1  2  1  3  2  4  3  2  2
2019  2  3  2
> plot(tsX)
```

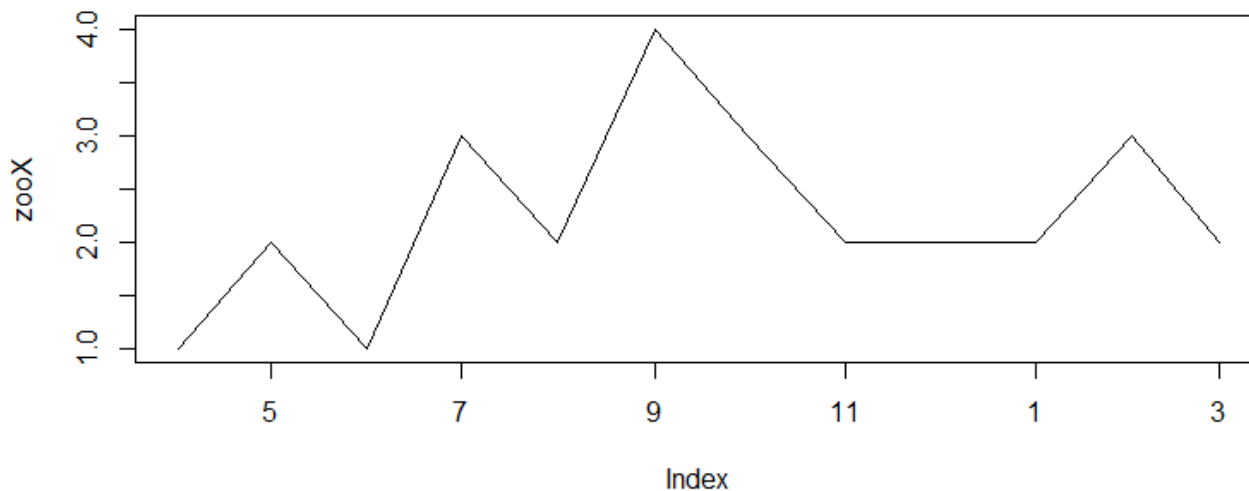


← 軸の目盛がずれてみえる...

- zooクラス (zooパッケージ)
 - tsクラスの機能拡張を目指す
 - 作成
zoo(数値ベクトル, 日付ベクトル)

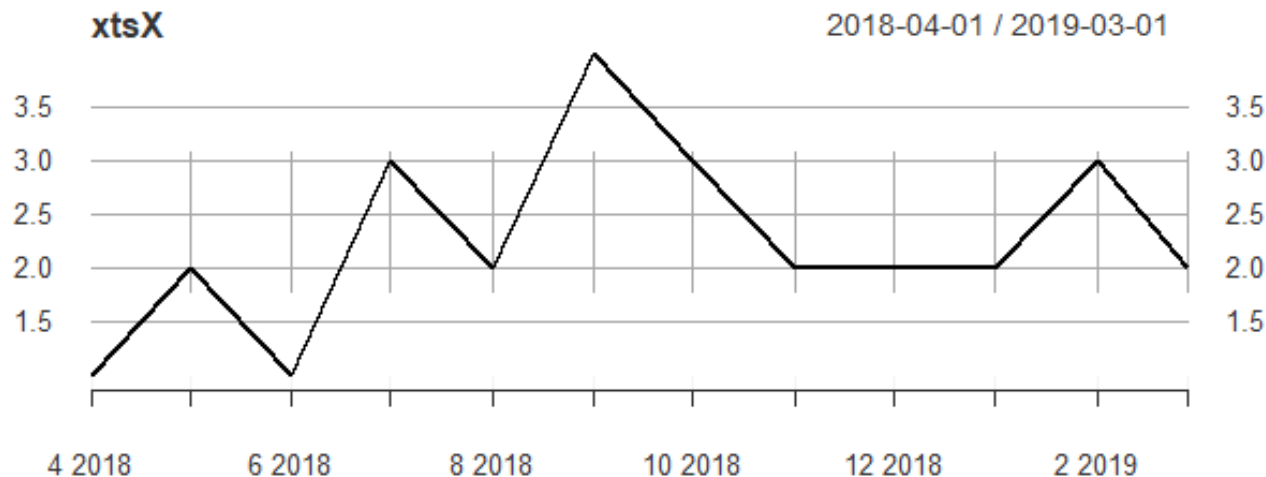
文字列ベクトルの末尾に"-01"をつけて、
日付らしく見える文字列にしている

```
library(zoo)
zooX <- zoo(dfIn[, -1], as.Date(paste0(dfIn$sMonth, "-01")))
plot(zooX)
```



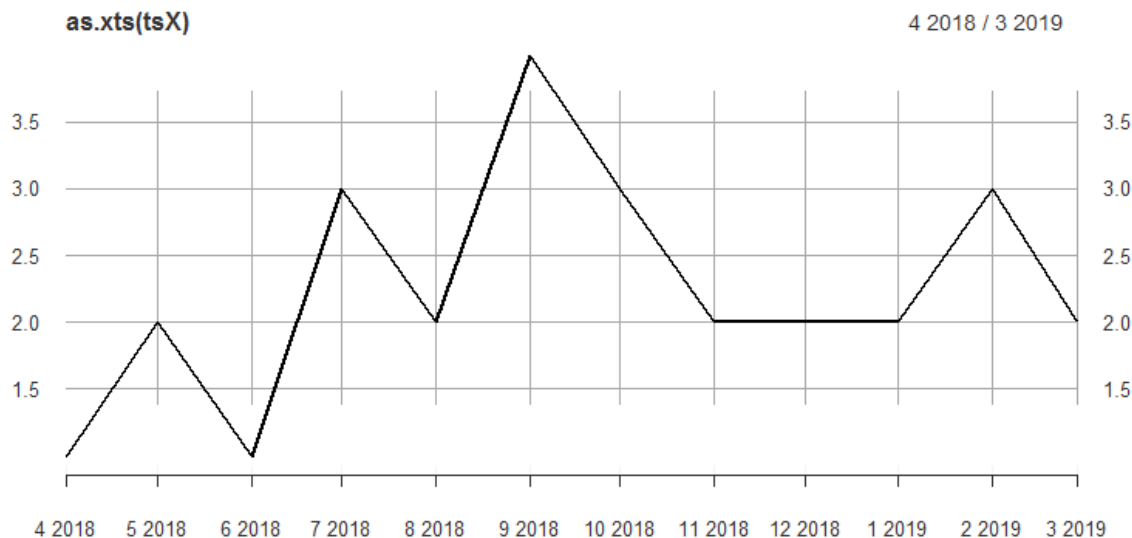
- xtsクラス (xtsパッケージ)
 - zooクラスの機能拡張を目指す
 - 作成
 - xts(数値ベクトル, 日付ベクトル)

```
library(xts)
xtsX <- xts(dfIn$gX, as.Date(paste0(dfIn$sMonth, "-01")))
plot(xtsX)
```



- xts クラスへの変換
as.xts(オブジェクト)
オブジェクト: 数値ベクトル, tsクラス, zooクラス, データフレームetc.
- ts クラスや zoo クラスの時系列をplotする際には、xtsクラスに変換してplotすると、多少わかりやすくなるような気がします

```
# さきほど作成した ts クラスの時系列オブジェクトを素早くプロット  
plot(as.xts(tsX))
```



- `tbl_ts` クラス (`tsibble` パッケージ)
 - `tidyverse` に準拠した時系列クラス。tsibbleはツィブルと読むそうです
 - 作成

データフレーム %>% `as_tsibble(index = 日付変数名)`

```
library(tsibble)
tsIn <- dfIn %>%
  mutate( dMonth = as.Date(paste0(sMonth, "-01")) ) %>%
  dplyr::select(dMonth, gX) %>%
  as_tsibble( index = dMonth )
```

- `tbl_time` クラス (`tibbletime` パッケージ)
 - こちらも `tidyverse` に準拠した時系列クラス
 - 作成

データフレーム %>% `as_tbl_time(index = 日付変数名)`

```
library(tibbletime)
tsIn <- dfIn %>%
  mutate( dMonth = as.Date(paste0(sMonth, "-01")) ) %>%
  dplyr::select(dMonth, gX) %>%
  as_tbl_time( index = dMonth )
```