

時系列データ解析

R による時系列データの解析例 2 – 13

Excel による時系列データの解析例 14 – 23

R による時系列データの解析例

RとRパッケージについて	3
卸売り高データ	4
地震波データ	8
太陽黒点数データ	9
最高気温データ	10
日経225株価データ	11
船の多変量データ	12

R と R パッケージについて

Rはオープンソースのデータ解析環境です。Linux, Max OS X, Windowsなどで使えます。Rの本体及び多数のパッケージがCRAN (The Comprehensive R Archive Network) から取得でき、インストールは簡単です。

- Rの取得場所 (日本 (統計数理研究所) のミラーサイト)
<http://cran.ism.ac.jp>
- インストール方法 (解説もネットに多数あり。例えば)
<https://to-kei.net/sitemap/>

Rには多くのパッケージが用意されています。RパッケージTSSSを込みこんでおくとき系列解析ソフトのほかここで使うデータも読み込めます。

- TSSS (Time Series analysis with State Space model)
<http://cran.ism.ac.jp> で「packages」をクリックしてTSSSを探す。
参考文献2 (時系列解析プログラミング) に掲載されたプログラムを基に統計数理研究所が開発したR 関数群。教科書 (時系列解析入門) のほとんどの例はTSSSで再現できる。開発者：嵯峨優美, 中野純司

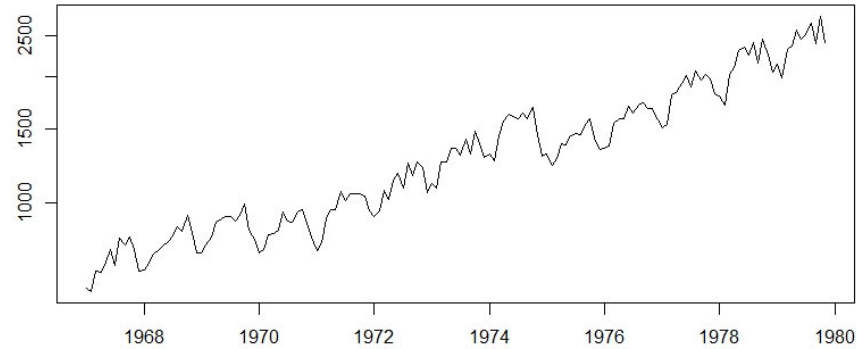
以下の解析例ではTSSSのデータを使っています

卸売り高データ

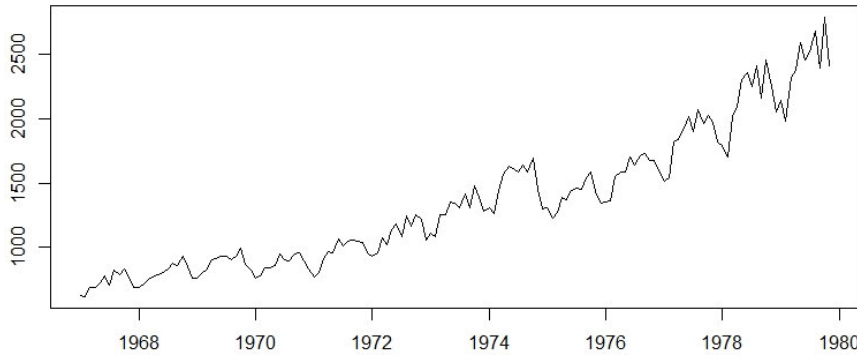
金額、人数、個数などの計数データは値が大きくなると変動も大きくなる場合が多い。そのような場合、対数変換すると変動がレベルによって変化しなくなることが多い。

データ名	WHARD
内容	あるハードウェアの毎月の卸売り高
観測間隔	一月
データ数	155
出展	合衆国Bureau of Labor Statistics

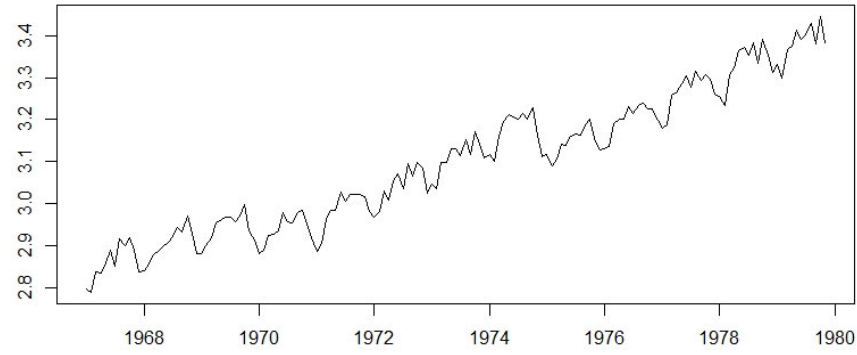
Y軸を対数目盛に変えて表示
`plot(WHARD, log="y")`



データを読み込む
`data(WHARD)`
データをプロットする
`plot(WHARD)`



データを対数変換してプロット
`plot(log10(WHARD))`



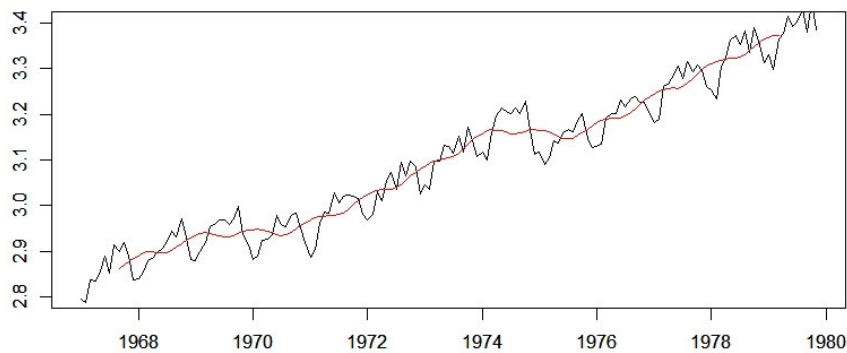
卸売り高データ(2)

- 移動平均で滑らかなトレンドが得られる
- 階差によってトレンドを除去できる
- 季節階差で周期を除去できる

移動平均を計算してプロットする

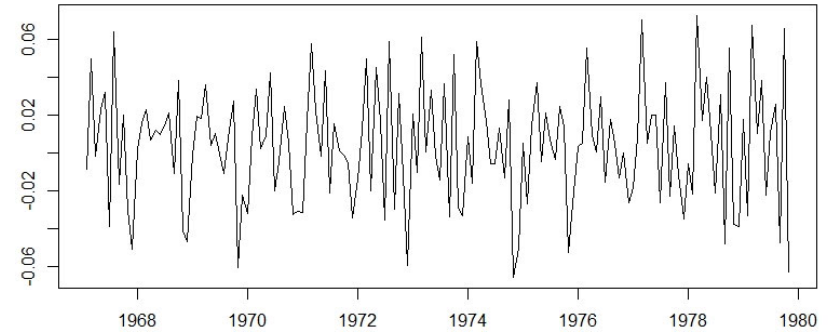
```
data( WHARD )  
x <- log10( WHARD )  
plot( x,type="l",ylim=c(2.8,3.4) )  
y <- x  
ndata <- length( x )  
y[1:ndata] <- NA  
kfilter <- 8  
n0 <- kfilter+1  
n1 <- ndata-kfilter  
for(i in n0:n1){  
  i0 <- i-kfilter  
  i1 <- i+kfilter  
  y[i] <- mean( x[i0:i1] )  
}  
par( new=T )  
plot( y,ylim=c(2.8,3.4),type="l" )  
lines(y,col=2,ylim=c(2.8,3.4))
```

この値を大きくすると滑らかになる

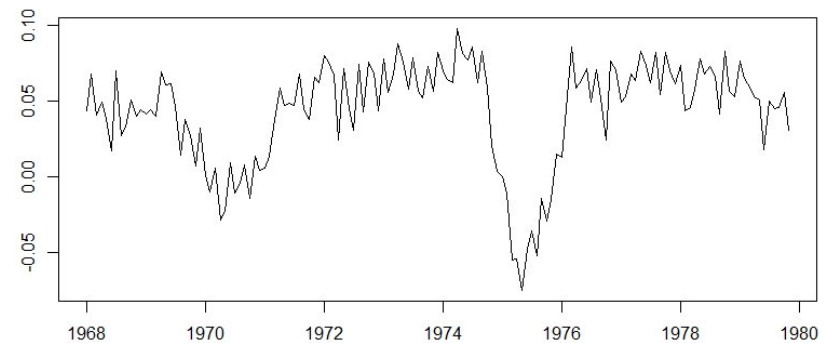


— WHARDの対数
— 移動平均

log10(WHARD)の階差を計算し表示する
plot(diff(log10(WHARD)))



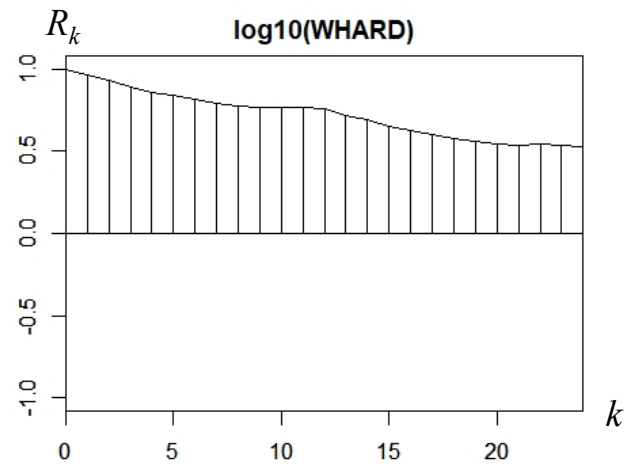
季節階差を計算し表示する
plot(diff(log10(WHARD), lag=12))



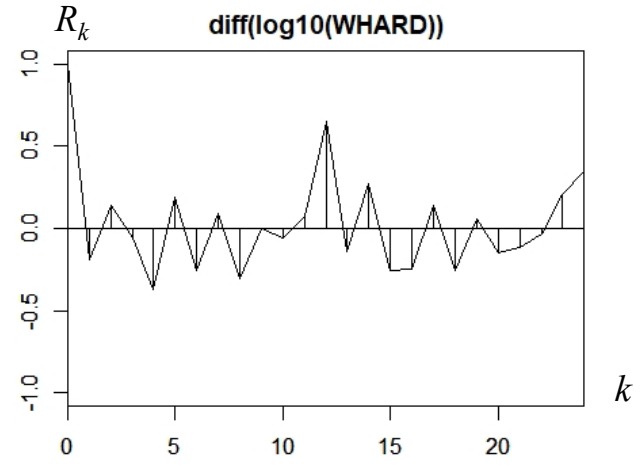
卸売り高データ(3)

- 自己相関関数で過去の値との相関の大きさがわかる。
- 階差や季節階差をとると自己相関関数も著しく変化する。

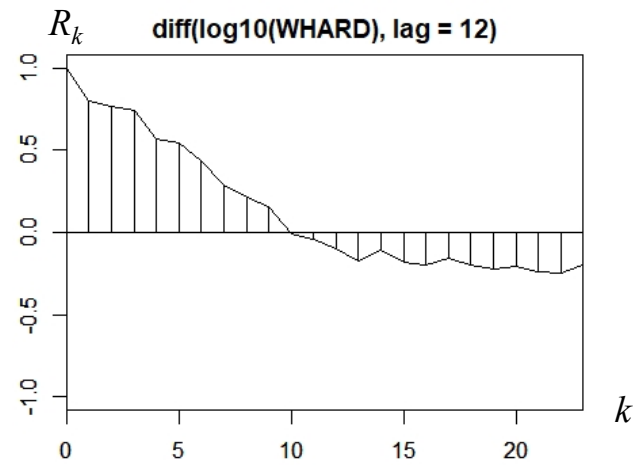
対数データの自己相関関数
`unicor(log10(WHARD))`



階差の自己相関関数を計算し表示する
`unicor(diff(log(WHARD)))`



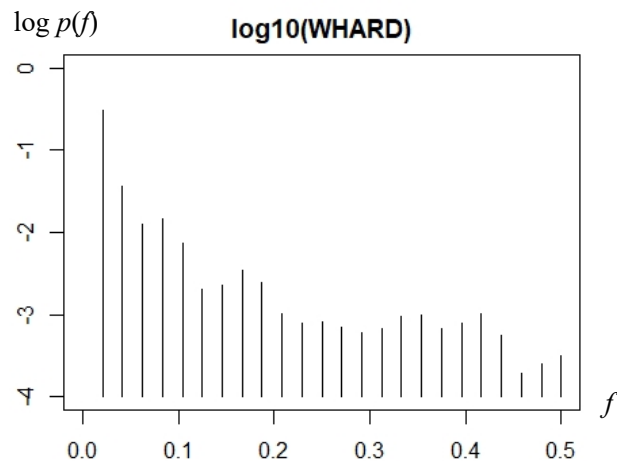
季節階差の自己相関関数を計算し表示する
`unicor(diff(log(WHARD), lag=12))`



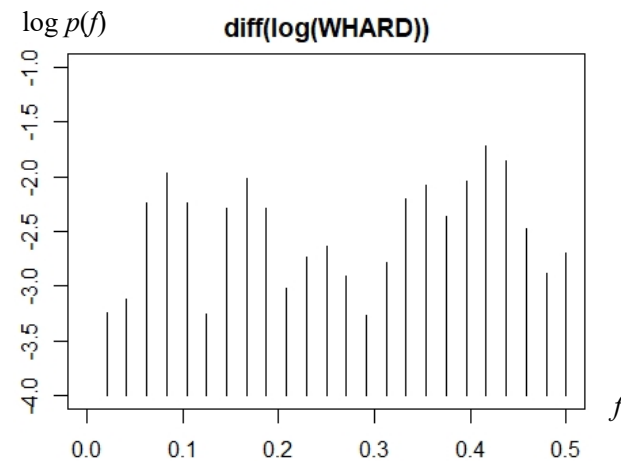
卸売り高データ(4)

- スペクトルで時系列の周期的特徴がわかる。
- 階差をとったデータのスペクトルは $f=0$ の付近が減少する。

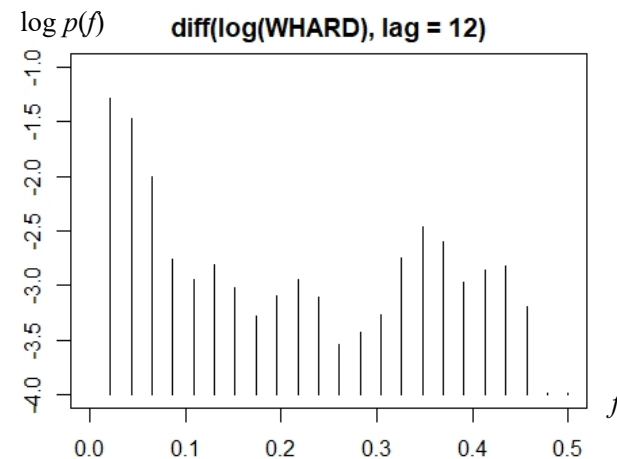
対数データのスペクトルを計算し表示する
`period(log10(WHARD))`



階差のスペクトルを計算し表示する
`period(diff(log(WHARD)))`



季節階差のスペクトルを計算し表示する
`period(diff(log(WHARD), lag=12))`



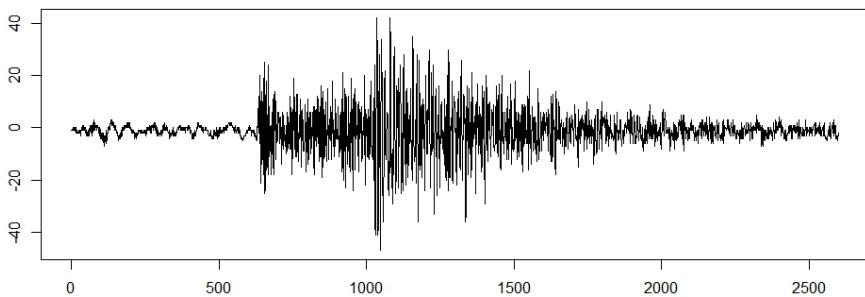
地震波データ

データ名	MYE1F
内容	地震波データ：東西方向
観測間隔	0.02秒
データ数	2600
出展	北海道大学 高波鐵夫氏 ISM data 43-3-01:, Ann. Inst. Statist. Math., 43, 605.

```

データを読み込む
data( MYE1F )

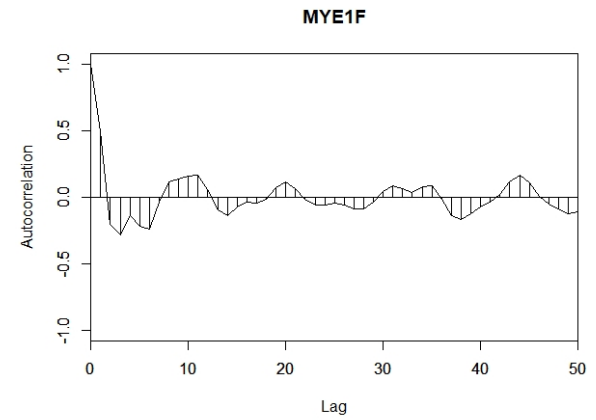
データをプロットする
plot( MYE1F )
    
```



- P波とS波の到着によって振幅が2回大きく変化している。
- 自己相関関数は最初急激に減少するが、その後は小さな変動が続く。
- スペクトルは $f=0.1$ (0.2秒周期) と0.07付近に顕著なピークがある。

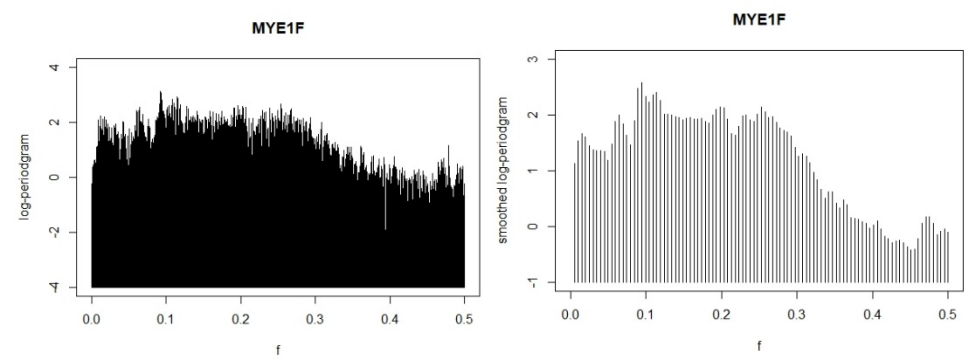
```

自己相関関数を計算しプロット
unicor( MYE1F, lag=50 )
    
```



```

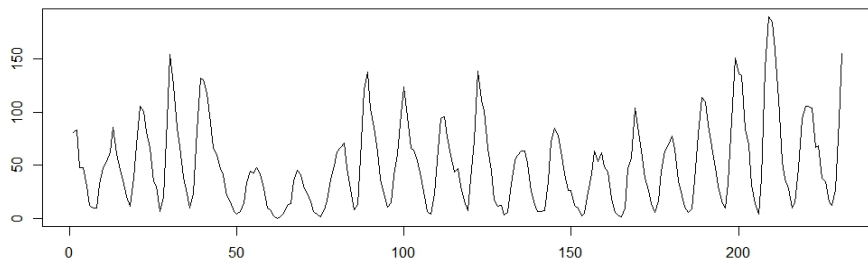
スペクトル (ピリオドグラム) を計算しプロット
period( MYE1F, window=0 )
period( MYE1F )
    
```



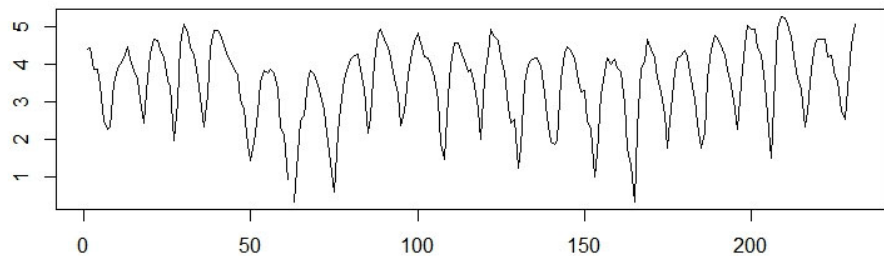
太陽黒点数データ

データ名	Sunspot
内容	Wolfers太陽黒点数
観測間隔	一月
データ数	231
出展	合衆国Bureau of Labor Statistics

データを読み込む
`data(Sunspot)`
 データをプロットする
`plot(Sunspot)`

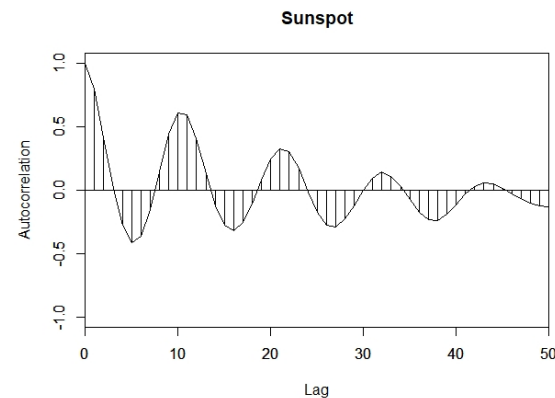


データの対数値を計算しプロット
`plot(log(Sunspot))`

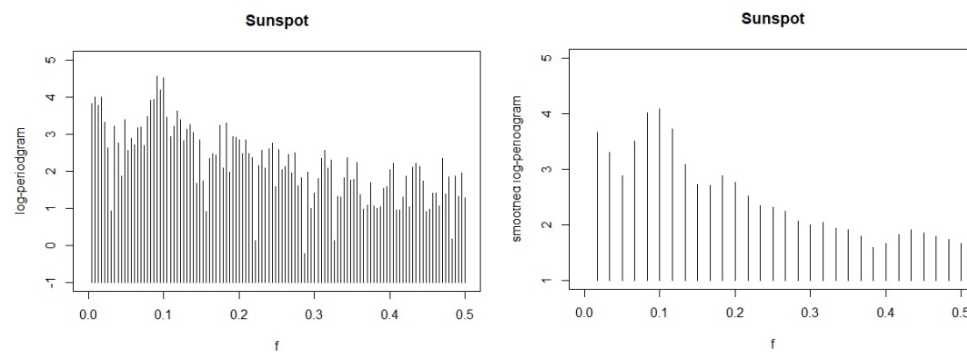


- 原系列は上に尖って非対称だが、対数をとると逆に下に尖る。
- 自己相関関数は11年程度の周期に対応して変動し、減衰も遅い。
- スペクトルは $f=0.09$ (約11年周期) に顕著なピークがある。

自己相関関数を計算しプロット
`unicor(Sunspot, lag=50)`



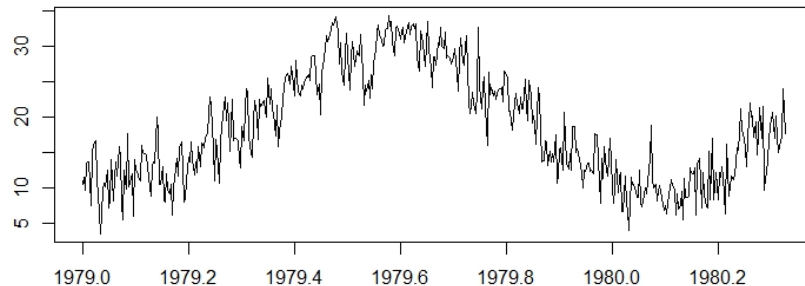
スペクトル (ピリオドグラム) を計算しプロット
`period(Sunspot, window=0)`
`period(Sunspot)`



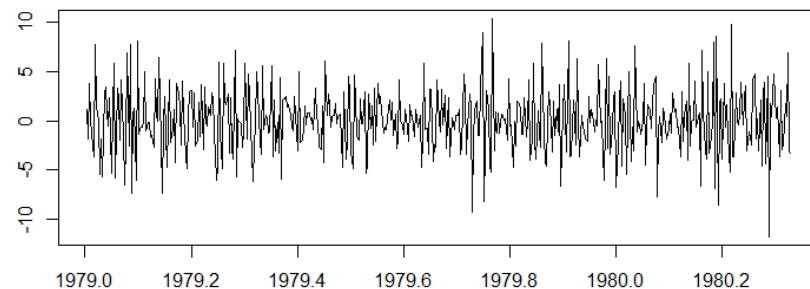
最高気温データ

データ名	Temperature
内容	東京の日最高気温
観測間隔	1日
データ数	486
出展	気象庁

データを読み込む
`data(Temperature)`
 データをプロットする
`plot(Temperature)`

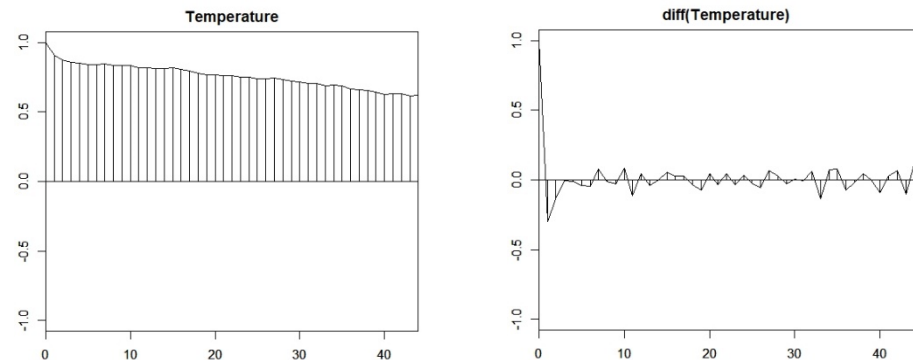


データの階差を計算しプロットする
`plot(diff(Temperature))`

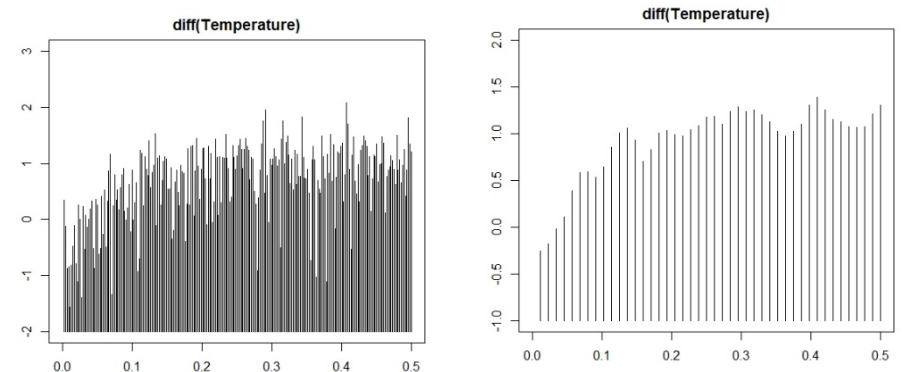


- 原データには顕著な1年周期がある。階差で周期は除去できる
- 原系列の自己相関関数は非常に減衰が遅いが、対数をとると急激に減少に無相関に近くなる。
- 階差系列のスペクトルは原点で落ち込み、顕著な周期は見えない。

原データの自己相関関数を計算しプロット
`unicor(Temperature)`
 階差データの自己相関関数を計算しプロット
`unicor(diff(Temperature))`



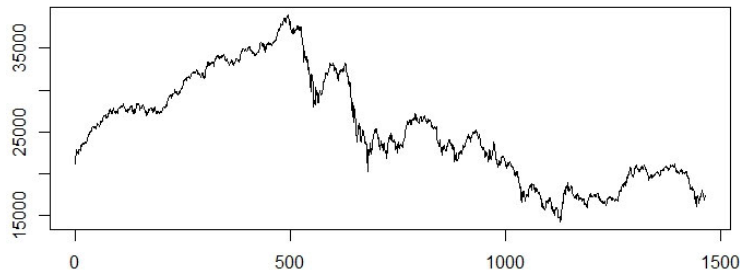
階差データのスペクトル (ピリオドグラム) を計算しプロット
`period(diff(Temperature), window=0)`
 階差データのスペクトル (平滑化ピリオドグラム) を計算しプロット
`period(diff(Temperature))`



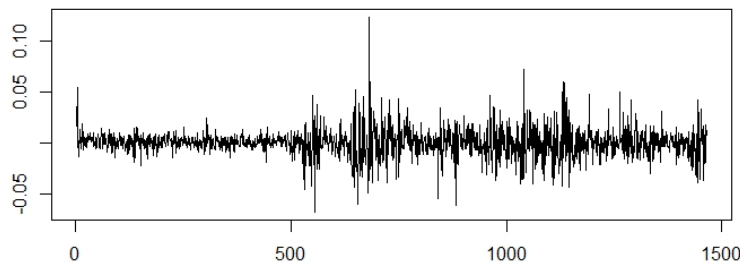
日経225株価データ

データ名	Nikkei225
内容	日経225平均株価(1988/1/4-1993/12/30)
観測間隔	1日
データ数	1480
出展	https://indexes.nikkei.co.jp/nkave/archives/data

データを読み込む
`data(Nikkei225)`
 データをプロットする
`plot(Nikkei225)`

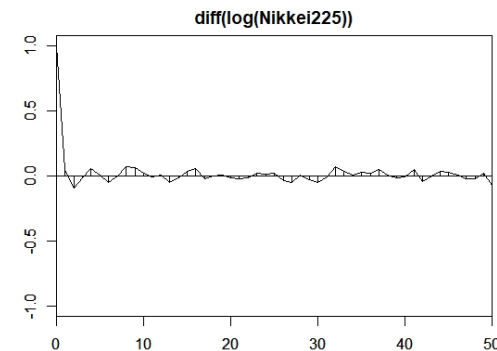


データの対数の階差を計算しプロットする
`plot(diff(log(Temperature)))`

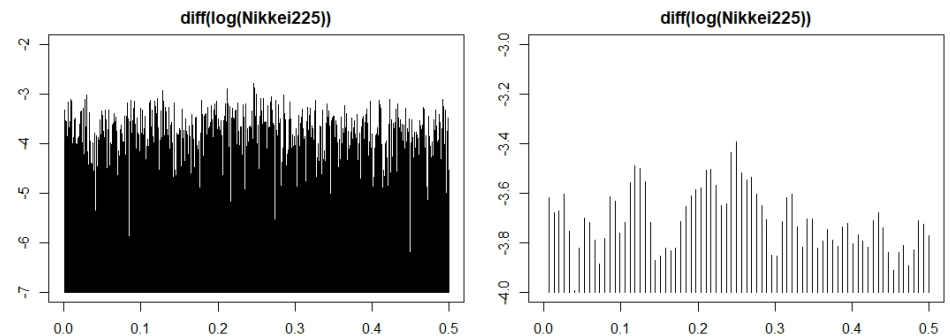


- トレンドがあり、時々急激な下落が見られる。対数差分をとると分散の変化が明確になる。
- 対数差分系列の自己相関関数は白色雑音に近い。
- スペクトルもフラットに近い。

階差データの自己相関関数を計算しプロット
`unicor(diff(log(Nikkei225)), lag=50)`



階差データのスペクトル (ピリオドグラム) を計算しプロット
`period(diff(log(Nikkei225)), window=0)`
 階差データのスペクトル (平滑化ピリオドグラム) を計算しプロット
`period(diff(log(Nikkei225)))`

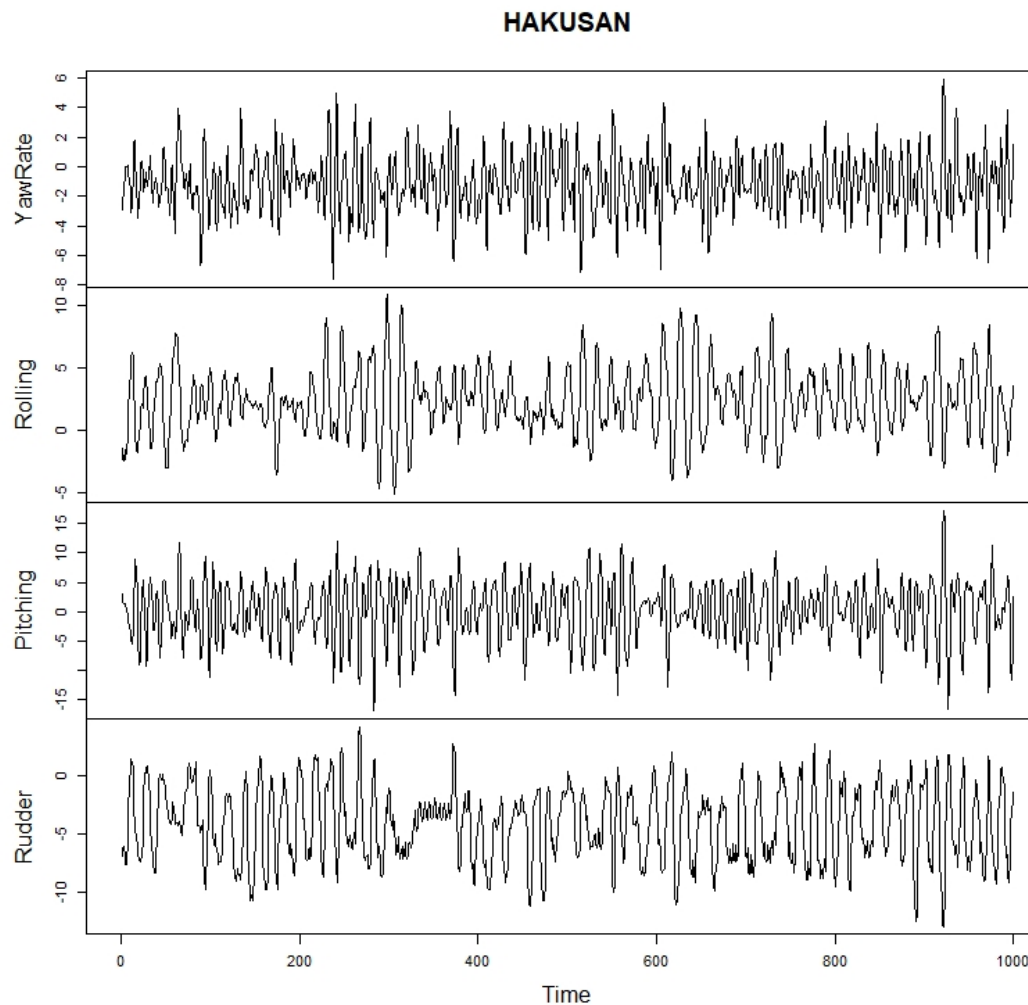


船の多変量データ

• 4変量の時系列横揺れと舵角は他の2つよりゆっくり変動している。

データ名	HAKUSAN
内容	船の航海中のデータ (4チャンネル) 1: 方向角速度 2: 横揺れ 3: 縦揺れ 4: 舵角
観測間隔	1秒
データ数	1000
出展	東京海洋大学 大津浩平名誉教授

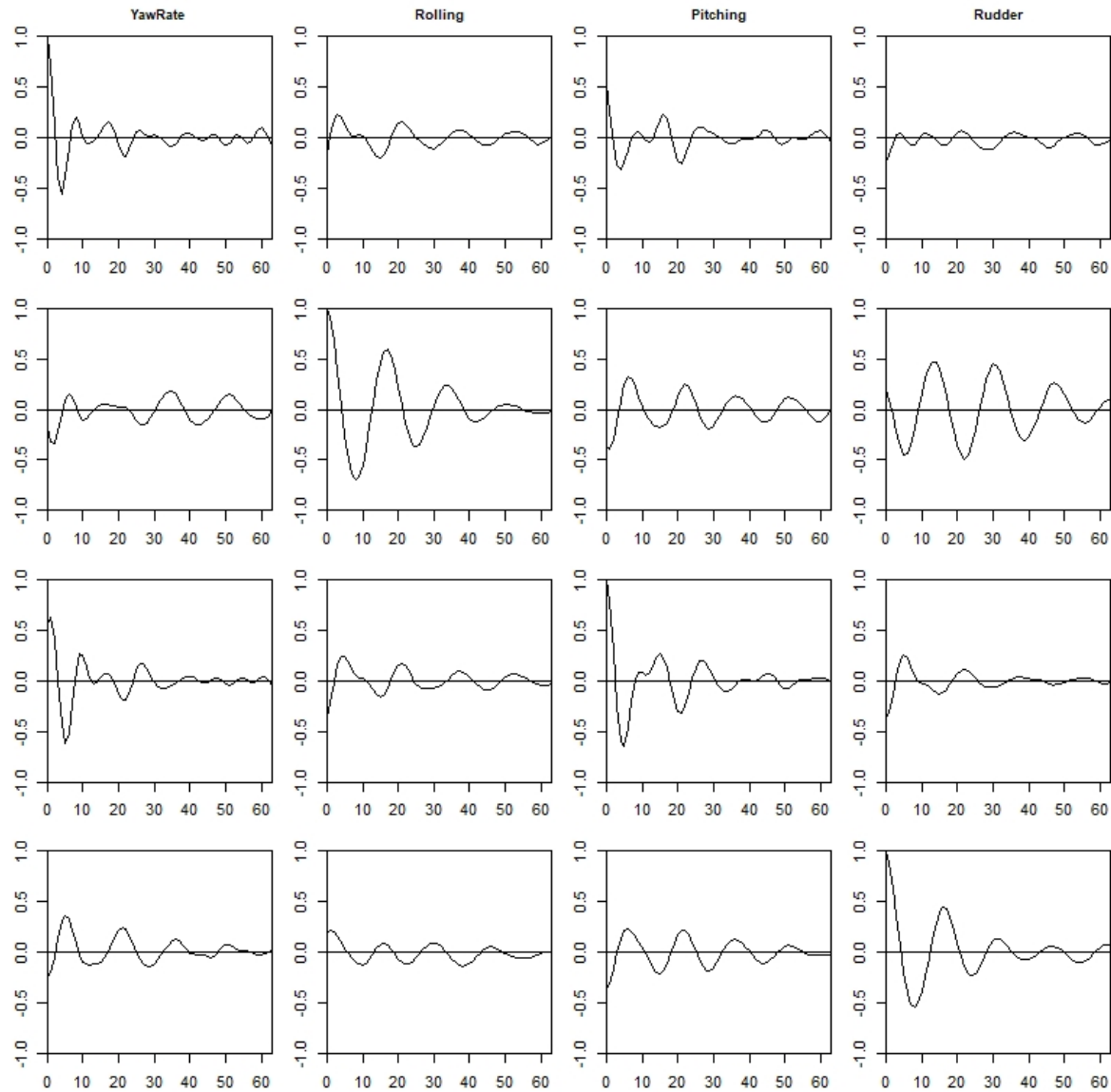
データを読み込む
`data(HAKUSAN)`
データをプロットする
`plot(HAKUSAN)`



船の多変量データ(2)

4つの時系列の相互相関関数
crscor(HAKUSAN, lag=30)

- 対角成分の4つが自己相関関数。非対角成分が相互相関関数。
- 横揺れと舵角の自己相関関数はゆっくりした変動をしている。
- 相互相関関数から舵角→横揺れ、方向角速度→縦揺れ、舵角への影響が見られる。



Excelによる時系列データの解析例

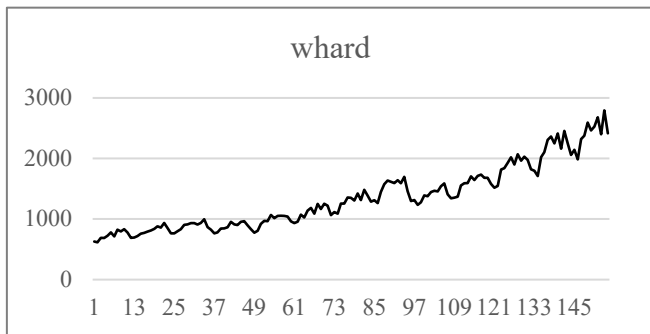
卸売り高データ	15
地震波データ	18
太陽黒点数データ	19
最高気温データ	20
日経225株価データ	21
船の多変量データ	22

卸売り高データ

データ名	whard.xlsx
内容	あるハードウェアの毎月の卸売り高
観測間隔	一月
データ数	155
出展	合衆国Bureau of Labor Statistics

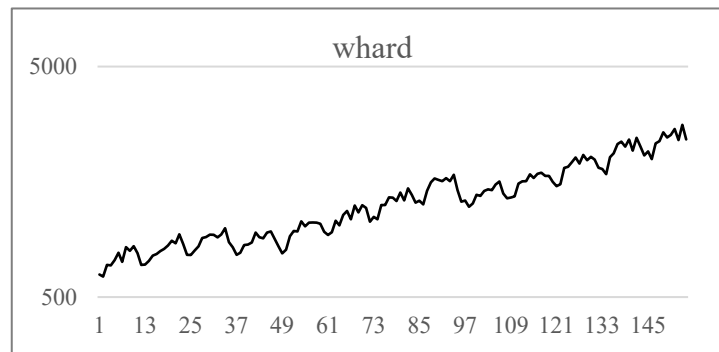
データを表示する

- ① A列を選択
- ② 挿入 > 2-D折れ線



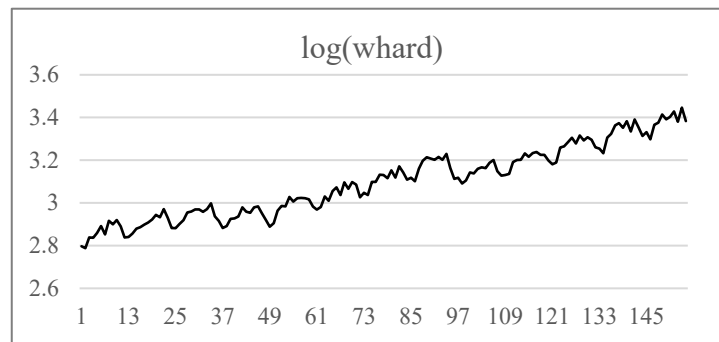
Y軸を対数目盛に変える

- ① 図の y 軸の数値を右クリック >> 軸の書式設定
- ② 軸の書式設定で「対数目盛を表示する」をチェック



B列をデータの対数変換としてプロット

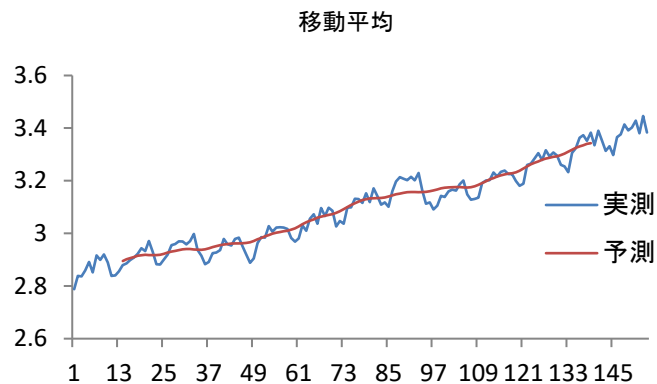
- ① セルB2に「=LOG(A2)」と入力し、B156までドラッグ
- ② 挿入 > 2-D折れ線



卸売り高データ(2)

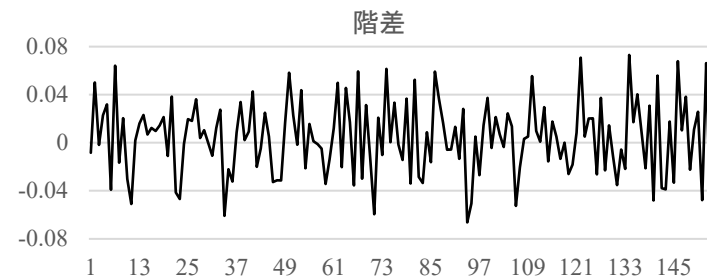
移動平均を計算する

- ① データ > データ分析 > 移動平均を選択して「OK」
- ② 入力範囲： \$B\$1:\$B\$156
「先頭行をラベルとして使用」にチェックを入れる
区間： 29
出力先： \$G\$2:\$G\$156
OK とすると移動平均が赤線で表示される
注意： 移動平均値は8時点だけ遅れて表示される $8=(17-1)/2$
- ③ C30:C15611を切り取ってC15:C140にペーストする



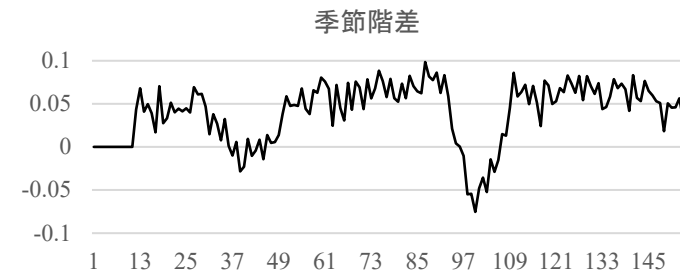
階差を計算し表示する

- ① セルD3に「=B3-B2」と入力し、D156までドラッグ
注意： セルD2は計算できない
- ② 挿入 > 2-D折れ線で階差データを表示する



季節階差を計算し表示する

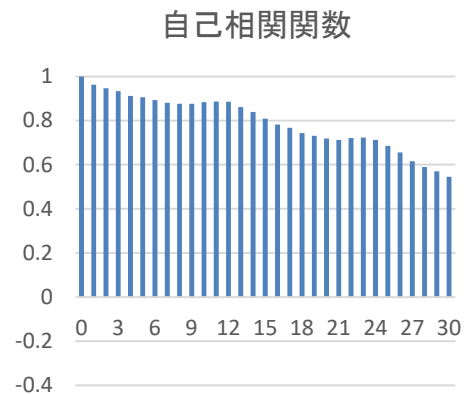
- ① セルE3に「=B14-B2」と入力し、E156までドラッグ
注意： セルE2-E13は計算できない
- ② 挿入 > 2-D折れ線で階差データを表示する



卸売り高データ(3)

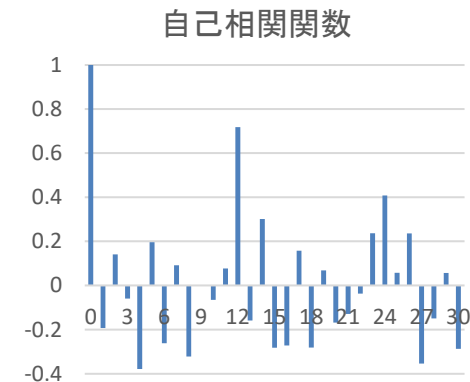
対数データの自己相関関数

- ① B2-AVERAGE(\$B\$2:\$B\$156)によって対数データから平均を除去したデータを計算し、F2~F156に保存する
- ② ラグを生成する
セルG2~G32にk = 0, 1, 2, . . . ,30 を入れておく
- ③ 自己共分散関数の計算
 $C(k) = \text{SUMPRODUCT}(F3:F156 * F2:\text{OFFSET}(F156, -G3, 0))$
- ④ 自己相関関数の計算
 $R(k) = C(k) / C(0)$
- ⑤ R(k)をプロットする



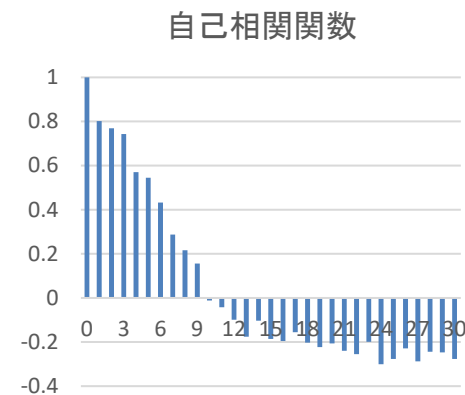
階差の自己相関関数を計算し表示する

データB2~B156の代わりに階差データD3~D156を使って同様に計算する



季節階差の自己相関関数を計算し表示する

データB2~B156の代わりに季節階差データE14~E156を使って同様に計算する



地震波データ

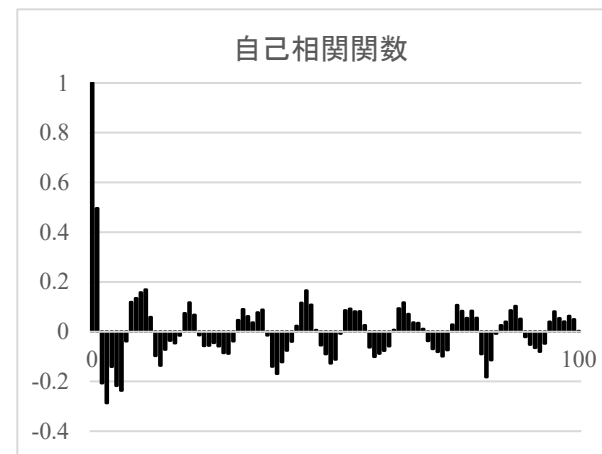
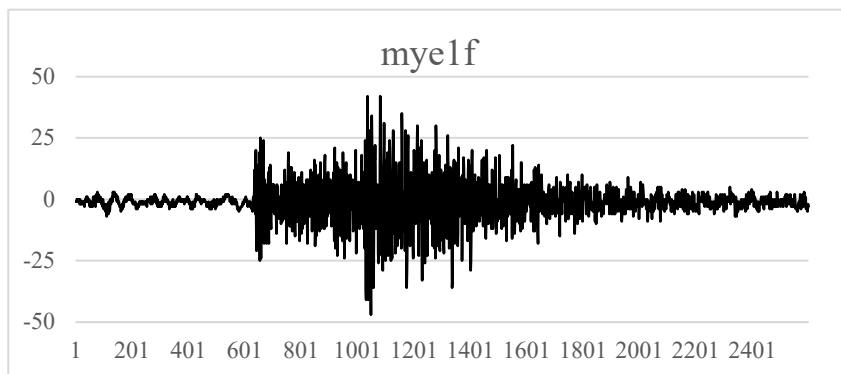
データ名	myelf.xlsx
内容	地震波データ：東西方向
観測間隔	0.02秒
データ数	2600
出展	北海道大学 高波鐵夫氏 ISM data 43-3-01:, Ann. Inst. Statist. Math., 43, 605.

自己相関関数を計算しプロット

- ① A2-AVERAGE(\$A\$2:\$A\$2601)によって対数データから平均を除去したデータを計算し、B2~B2601に保存する
- ② ラグを生成する
セルC2~C102に $k = 0, 1, 2, \dots, 100$ を入れておく
- ③ 自己共分散関数の計算
 $C(k) = \text{SUMPRODUCT}(B2:B2601 * B2:\text{OFFSET}(B2601, -C3, 0))$
- ④ 自己相関関数の計算
 $R(k) = C(k)/C(0)$ を計算し、E2 ~E102に入れる
- ⑤ R(k)をプロットする

データを表示する

- ① カラムAを選択
- ② 挿入 > 2-D折れ線

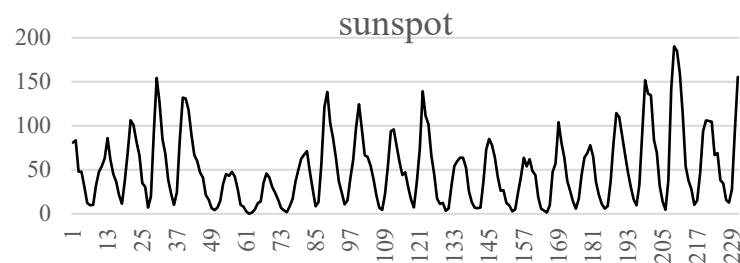


太陽黒点数データ

データ名	Sunspot.xlsx
内容	Wolfert太陽黒点数
観測間隔	一月
データ数	231
出展	合衆国Bureau of Labor Statistics

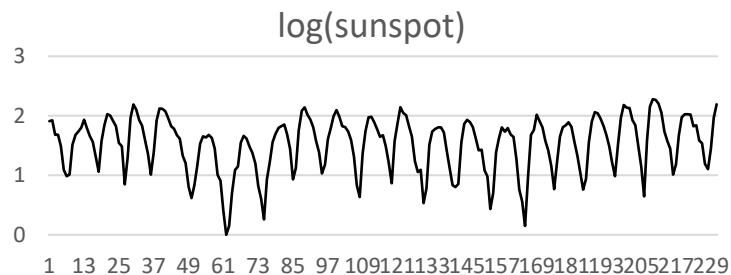
データを表示する

- ① カラムAを選択
- ② 挿入 > 2-D折れ線



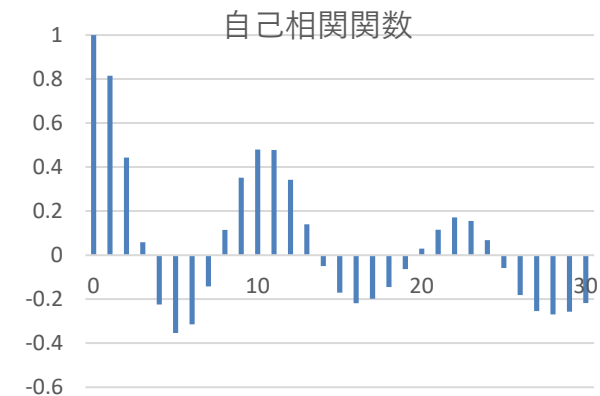
データの数値を計算する

- ① セルB2に「=LOG(A2)」と入力し、B2601までドラッグ



対数データの自己相関関数

- ① $A2-AVERAGE(\$A\$2:\$B\$2601)$ によって対数データから平均を除去したデータを計算し、B2~B2601に保存する
- ② ラグを生成する
セルC2~C102に $k = 0, 1, 2, \dots, 100$ を入れておく
- ③ 自己共分散関数の計算
 $C(k) = \text{SUMPRODUCT}(B3:B\$2601 * B\$2:\text{OFFSET}(B\$2601, -C3, 0))$
- ④ 自己相関関数の計算
 $R(k) = C(k)/C(0)$ を計算し、D2 ~D102に入れる
- ⑤ $R(k)$ をプロットする

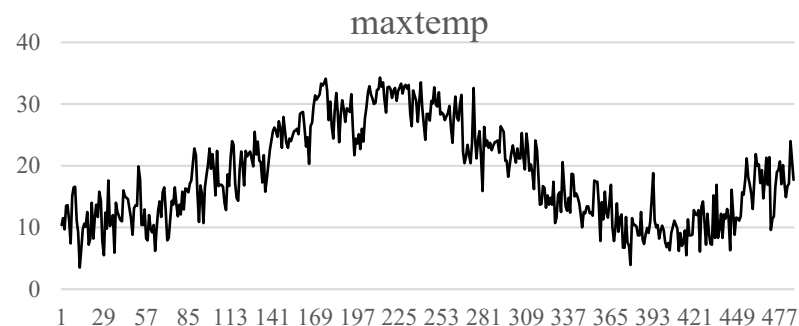


最高気温データ

データ名	Temperature
内容	東京の日最高気温
観測間隔	1日
データ数	486
出展	気象庁

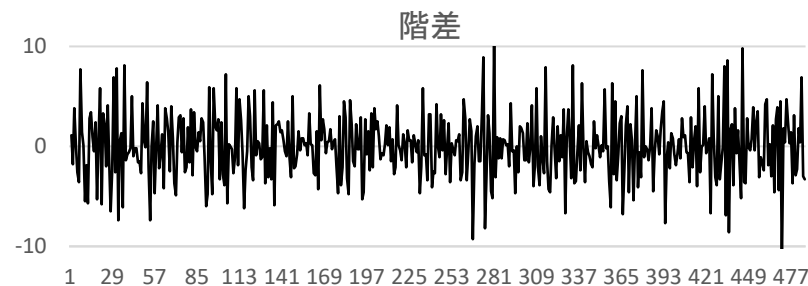
データを表示する

- ① カラムAを選択
- ② 挿入 > 2-D折れ線



データの階差を計算する

- ① セルB3に「=A3-A2」と入力し、B487までドラッグ

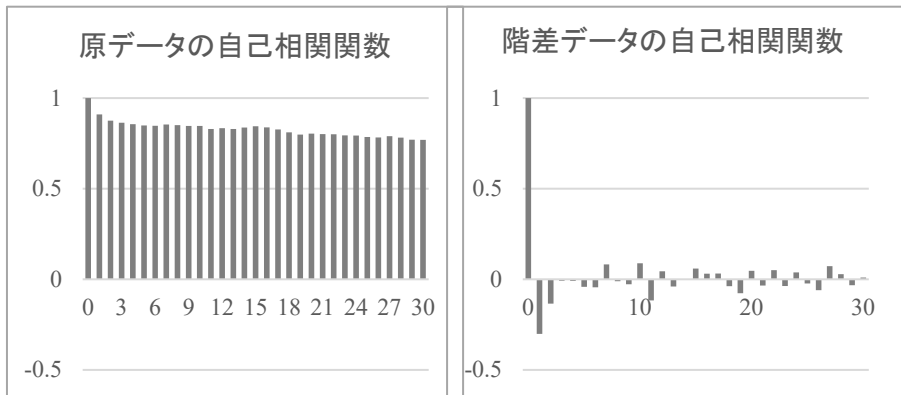


原データの自己相関関数

- ① $A2-AVERAGE(\$A\$2:\$A\$487)$ によって対数データから平均を除去したデータを計算し、C2~C487に保存する
- ② ラグを生成する
セルD2~D32に $k = 0, 1, 2, \dots, 30$ を入れておく
- ③ 自己共分散関数の計算
 $C(K)=SUMPRODUCT(C3:C487*C\$2:OFFSET(C\$487-D3,0))$
- ④ 自己相関関数の計算
 $R(k) = C(k)/C(0)$ を計算し、E2 ~E32に入れる
- ⑤ $R(k)$ をプロットする

階差データの自己相関関数を計算しプロット

- ① 自己共分散関数の計算
 $C(K)=SUMPRODUCT(B2:B487*B\$2:OFFSET(B\$487,-D3,0))$
- ④ 自己相関関数の計算
 $R(k) = C(k)/C(0)$ を計算し、F2 ~F32に入れる
- ⑤ $R(k)$ をプロットする

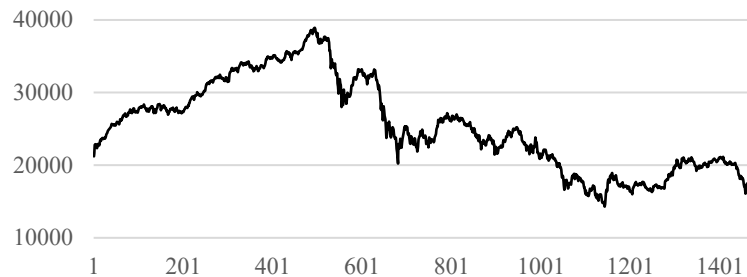


日経225株価データ

データ名	nikkei225
内容	日経225平均株価(1988/1/4-1993/12/30)
観測間隔	1日
データ数	1480
出展	https://indexes.nikkei.co.jp/nkave/archives/data

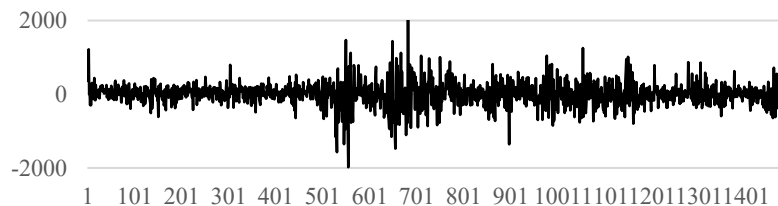
データを表示する

- ① カラムAを選択
- ② 挿入 > 2-D折れ線



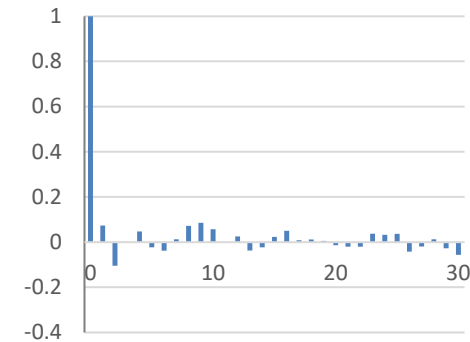
データの階差を計算する

- ① セルB3に「=A3-A2」と入力し、B487までドラッグ



差分データの自己相関関数

- ① ラグを生成する
セルC2~C32に $k = 0, 1, 2, \dots, 30$ を入れておく
- ② 自己共分散関数の計算
・セルD2に計算「=SUMPRODUCT(B2:b\$1481*B\$2:OFFSET(B\$1481, -\$C2,0)/SUMPRODUCT(B\$2:B\$1481*B\$2:B\$1481))」と入力する。
・セルD2を選択して下にドラッグしD32まで計算する。
- ④ D列をプロットする。



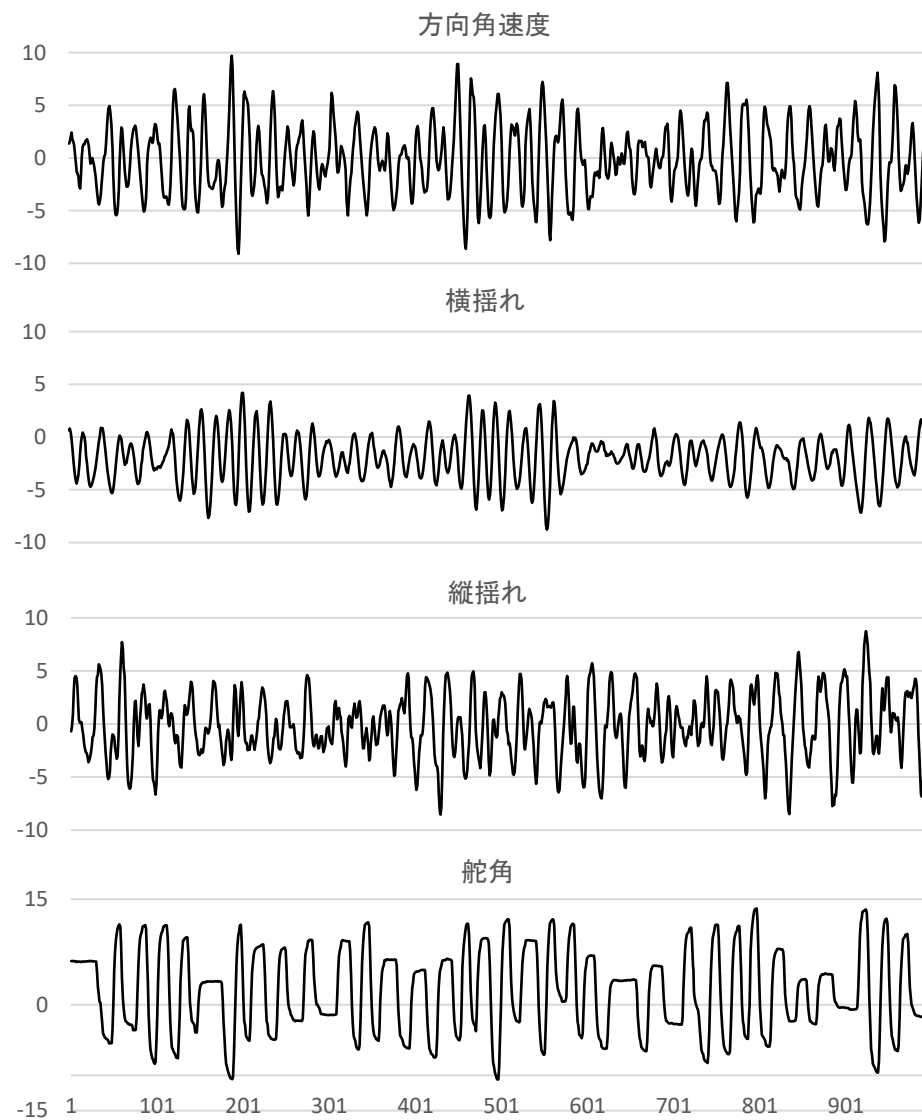
船の多変量データ

データ名	HAKUSAN
内容	船の航海中のデータ (4チャンネル) 1: 方向角速度 2: 横揺れ 3: 縦揺れ 4: 舵角
観測間隔	1秒
データ数	1000
出展	東京海洋大学 大津浩平名誉教授

方向角速度 (1チャンネル) データを表示する場合

- ① カラムAを選択
- ② 挿入 > 2-D折れ線

横揺れ, 縦揺れ, 舵角も同様



船の多変量データ(2)

4つの時系列の自己相関関数

- ① ラグを生成する
セルE2~E32に $k = 0, 1, 2, \dots, 30$ を入れておく。
- ② 各時系列から平均を除去したデータをF~I列に作成する。
 - ・セルF2に計算式「 $=A2-AVERAGE(\$A\$2:\$A\$1001)$ 」を入力して時系列データから平均を除去したデータを計算する。
 - ・セルF2を選択して横にドラッグし、G2, H2, I2 を計算する。
 - ・F2~I2を選択して、1001まで下にドラッグする。
- ③ 自己相関関数を計算する
 - ・セルJ2に計算式「 $=SUMPRODUCT(F2:F\$1001*F\$2:OFFSET(F\$1001,-\$E2,0)/SUMPRODUCT(F\$2:F41001*F\$2:F\$1001))$ 」と入力する。
 - ・セルJ2を選択して横にドラッグしてK2,L2,M2を計算する。
 - ・セルJ2~M2を選択して下にドラッグK32~M32まで計算する。
- ④ J列~M列をプロットする。

