

クレジット:

UTokyo Online Education 数理手法Ⅶ 2019 北川源四郎

ライセンス:

利用者は、本講義資料を、教育的な目的に限ってページ単位で利用することができます。特に記載のない限り、本講義資料はページ単位でクリエイティブ・コモンズ 表示-非営利-改変禁止 ライセンスの下に提供されています。

<http://creativecommons.org/licenses/by-nc-nd/4.0/>

本講義資料内には、東京大学が第三者より許諾を得て利用している画像等や、各種ライセンスによって提供されている画像等が含まれています。個々の画像等を本講義資料から切り離して利用することはできません。個々の画像等の利用については、それぞれの権利者の定めるところに従ってください。



R のTSSSPッケージについて

- 前回講義で紹介した下記のサイトは復活しました。(2019/4/17)
<http://jasp.ism.ac.jp/ism/TSSS/>
- Rのサイトから新しいバージョンを公開しました。
Version 1.2.3 (2019/4/30まで)
<https://cran.r-project.org/package=TSSS>

Version 1.2.4 (2019/5/1以降)

時系列解析 (2)

2019/4/17

東京大学 数理・情報教育研究センター

北川源四郎

時系列の可視化(記述)

波形を見る 時系列プロット (前回)

相関をみる 相関関数

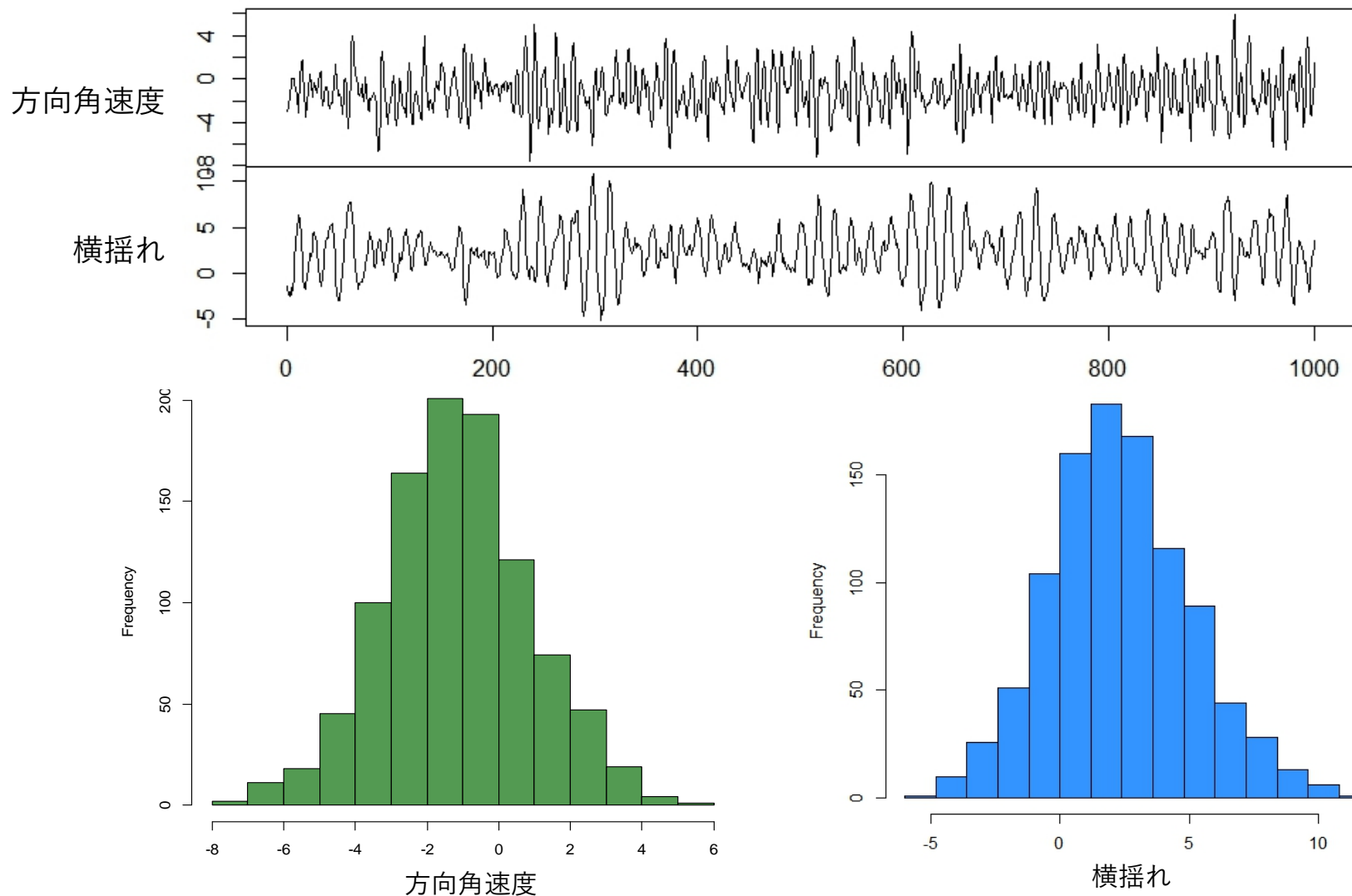
- 共分散関数
- 相関関数

周期性をみる スペクトル

- スペクトル
- ピリオドグラム

時系列とヒストグラム

```
plot(hakusan[,c(1,3)],main="")  
hist(hakusan[,1],main="方向角速度",col="#539952")  
hist(hakusan[,3],main="横揺れ",breaks=seq(-6,12,1.2), col="#3394FF")
```



y_n と y_{n-k} の散布図 ($k=2,4$)

```
y <- hakusan[,1] z<- hakusan[,3]
plot(y,lag(y,k=2),pch=19,col="#539952",ylab="y(n-2)",xlab="y(n)")
plot(z,lag(z,k=2),pch=19,col="#539952",ylab="y(n-2)",xlab="y(n)")
```

x と y の散布図

x を横軸, y を縦軸として
2次元データ(x,y)を点で
プロットしたもの

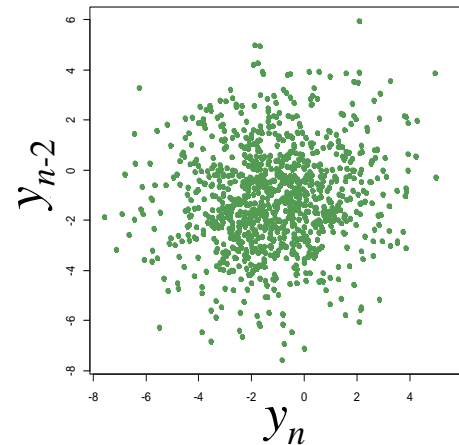
ここでは

$$x = y_n$$

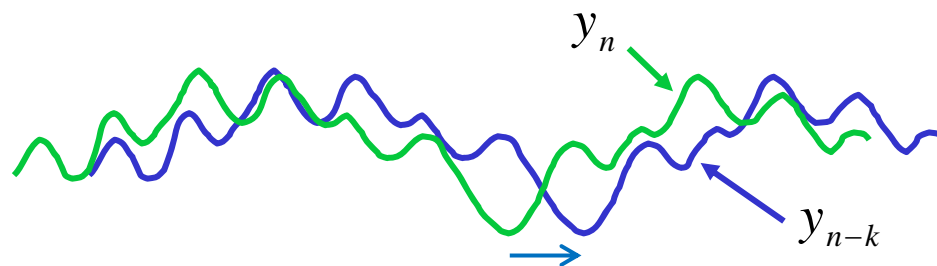
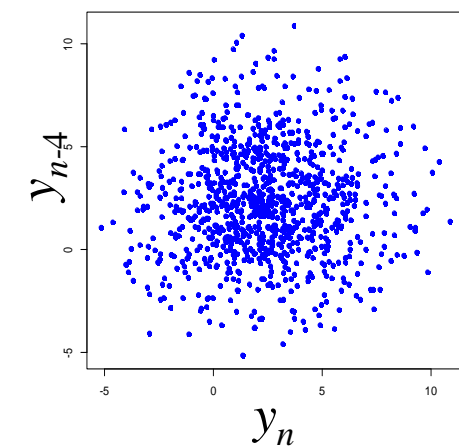
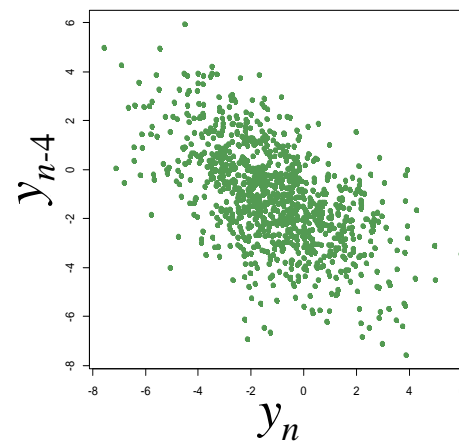
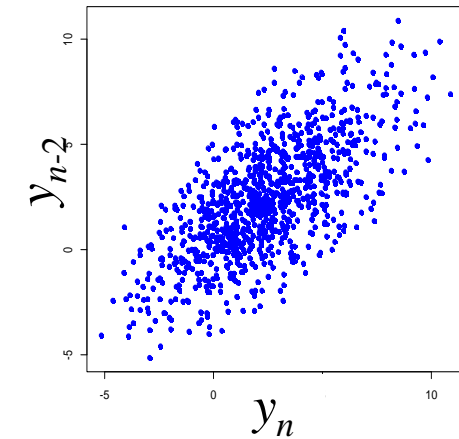
$$y = y_{n-k}$$

とする.

方向角速度

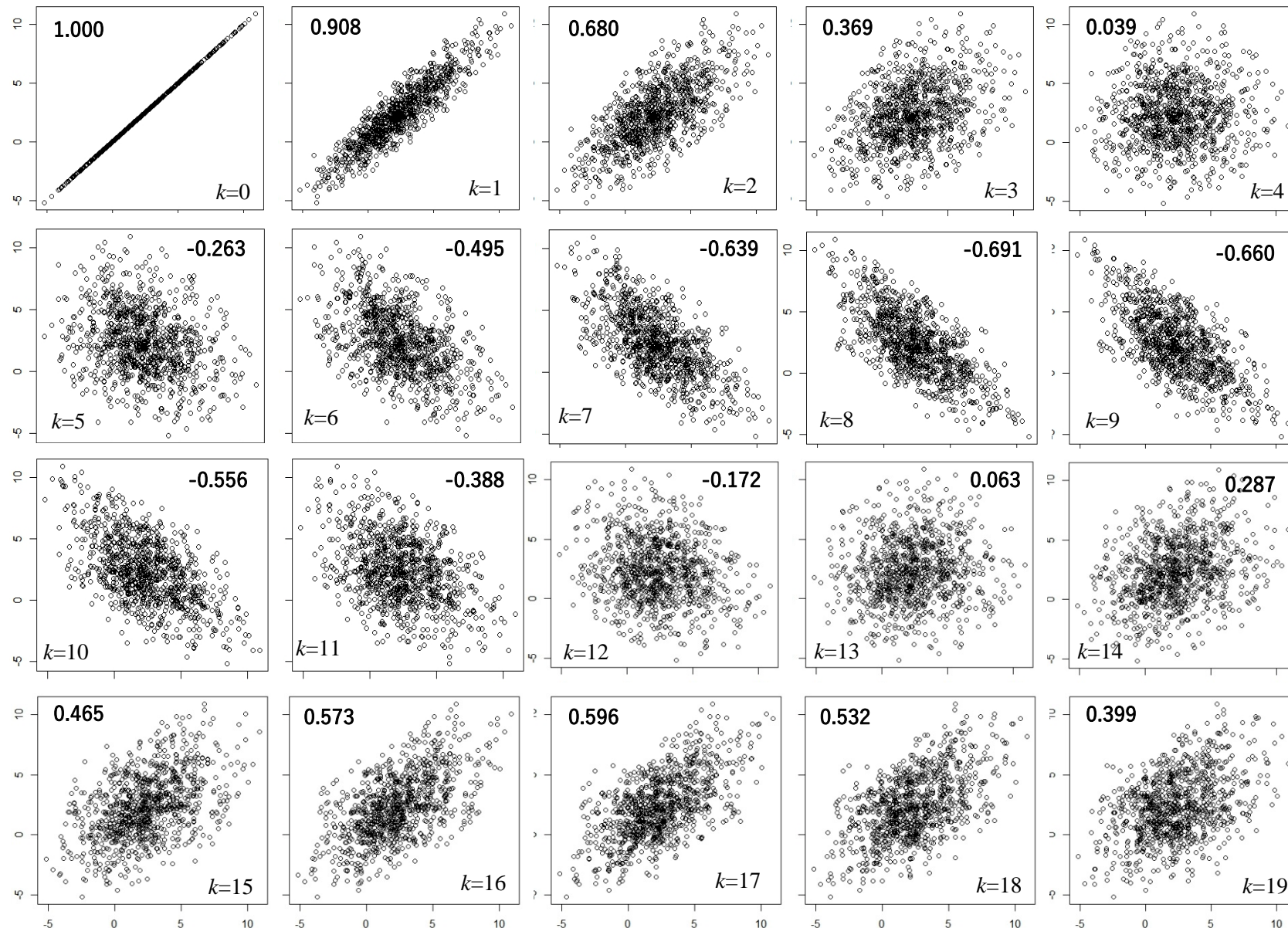


横揺れ



散布図と相関係数 ($k=0, \dots, 19$)

```
x <- hakusan[,3]  
plot(x,x)  
plot(x,lag(x,k=1))
```



1 変量時系列の平均, 分散, 自己共分散

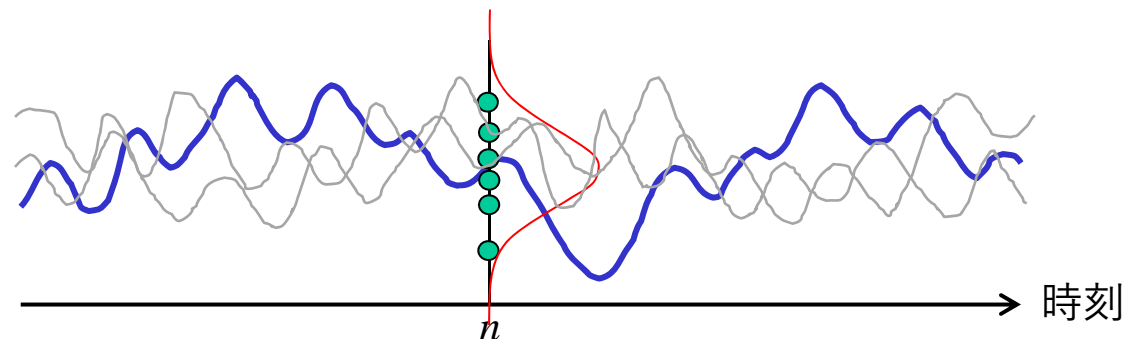
y_n : 時系列 (n : 時刻, N : データ数, y_1, \dots, y_N)

E : 期待値 (平均)

平均 $\mu_n = E[y_n]$

分散 $Var(y_n) = E[(y_n - \mu_n)^2]$

共分散 $Cov(y_n, y_m) = E[(y_n - \mu_n)(y_m - \mu_m)]$



定常性

確率構造が時間変化しないこと

弱定常： 平均・分散・共分散が（存在して）
変化しない（2次のモーメントまで
が不変）

強定常： 確率分布が時間変化しない

定常：通常は弱定常のこと

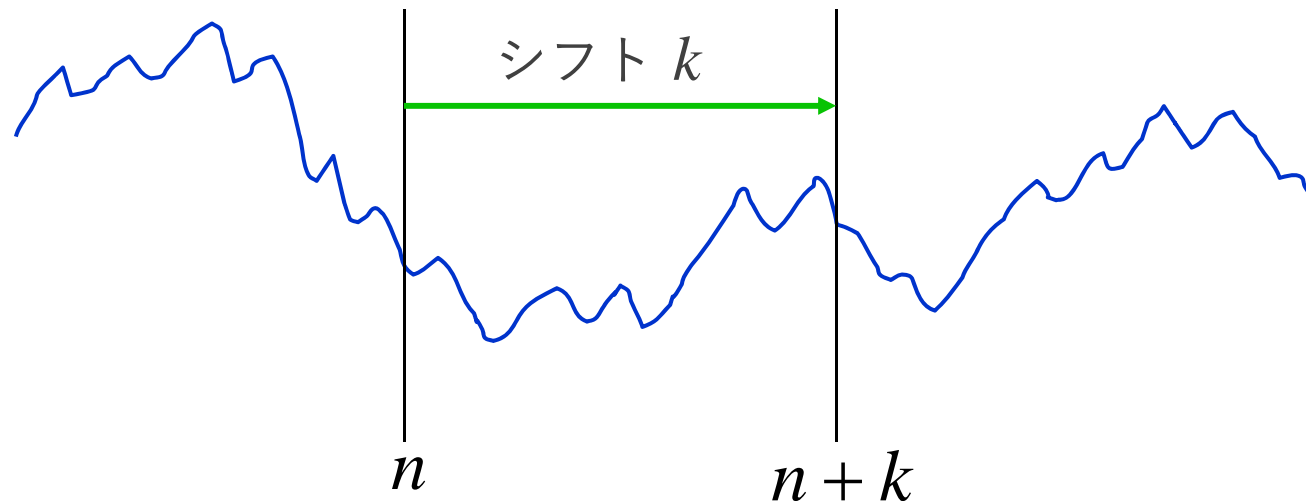
(弱) 定常性

任意のシフト量(ラグ), $k \in \{0, \pm 1, \pm 2, \dots\}$
に対して以下が成り立つ

$$E(y_n) = E(y_{n+k})$$

$$\text{Var}(y_n) = \text{Var}(y_{n+k})$$

$$\text{Cov}(y_n, y_m) = \text{Cov}(y_{n+k}, y_{m+k})$$



強定常性

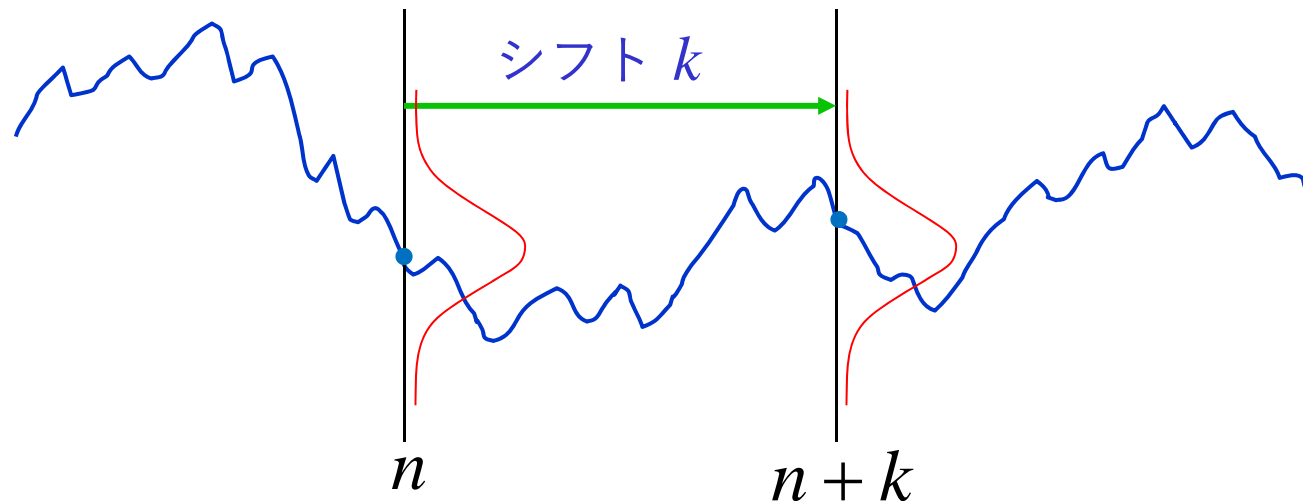
k : 任意のシフト量

m : 任意の次元 (時点数) ($=1, 2, \dots$)

n_1, \dots, n_m : 任意の時点

確率分布が時間変化しない

$$p(y_{n_1}, y_{n_2}, \dots, y_{n_m}) = p(y_{n_1+k}, y_{n_2+k}, \dots, y_{n_m+k})$$



強定常と弱定常の関係

1. 弱定常、正規分布 → 強定常
2. 平均，分散，共分散が存在するとき
強定常 → 弱定常
3. 強定常だが弱定常でないことがある

例：コーシー分布に従う時系列

$$p(y_n) = \frac{1}{\pi} \frac{1}{y_n^2 + \tau^2}$$

平均，分散が存在しない

定常時系列の自己共分散関数

y_n : (弱)定常時系列

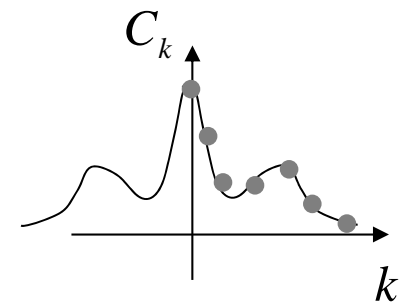
$$\mu_n, \text{Cov}(y_n, y_{n-k})$$

- 時刻 n に依存しない
- 時間差 k だけに依存

$$\mu \equiv E[y_n]$$

$$C_k \equiv \text{Cov}(y_n, y_{n-k}) = E[(y_n - \mu)(y_{n-k} - \mu)]$$

- C_k 自己共分散
- $C_k, k = 0, \pm 1, \dots, \pm L$ 自己共分散関数
- k : ラグ (lag)
- コレログラム : 自己共分散関数を図示したもの



自己共分散関数の性質

1. 自己共分散関数は偶関数

$$C_{-k} = C_k$$

2. 絶対値の最大は C_0

$$|C_k| \leq C_0$$

3. 半正定値

$$\sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j C_{i-j} \geq 0, \quad \forall \alpha_1, \dots, \alpha_m$$

自己共分散関数の性質 (1)

定常時系列の自己共分散関数は偶関数

$$C_{-k} = C_k$$

$$C_k = \text{Cov}(y_n, y_{n-k})$$

$$\begin{aligned} C_{-k} &= \text{Cov}(y_n, y_{n+k}) \\ &= \text{Cov}(y_{n+k}, y_n) \\ &= \text{Cov}(y_n, y_{n-k}) = C_k \end{aligned}$$

↑
定常性

- 1 変量だけに成り立つ性質
- コレログラムは $k \geq 0$ だけ表示すればよい
- ARモデルが簡単に推定できる理由

自己共分散関数の性質 (2)

$$|C_k| \leq C_0$$

$$E(a^2)E(b^2) \geq E(ab)^2$$

$$a = y_n - \mu, \quad b = y_{n-k} - \mu$$

$$E(a^2) = E(y_n - \mu)^2 = C_0$$

$$E(b^2) = E(y_{n-k} - \mu)^2 = C_0$$

$$E(ab) = E(y_n - \mu)(y_{n-k} - \mu) = C_k$$

$$E(\lambda a + b)^2 \geq 0$$

$$= \lambda^2 E(a^2) + 2\lambda E(ab) + E(b^2)$$

$$= E(a^2) \left(\lambda - \frac{E(ab)}{E(a^2)} \right)^2 + E(b^2) - \left(\frac{E(ab)}{E(a^2)} \right)^2$$

$$E(b^2) - \left(\frac{E(ab)}{E(a^2)} \right)^2 \geq 0$$

$$E(a^2)E(b^2) \geq E(ab)^2$$

$$C_0 C_0 \geq C_k^2 \quad \Rightarrow \quad C_0 \geq |C_k|$$

自己共分散関数の性質 (3)

$$\sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j C_{i-j} \geq 0$$

非負定値性

$$\begin{aligned} E \left(\sum_{i=1}^m \alpha_i y_{n-i} \right)^2 &= \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j E[y_{n-i} y_{n-j}] \\ &= \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j C_{i-j} \geq 0 \end{aligned}$$

- Yule-Walker方程式が解を持つ

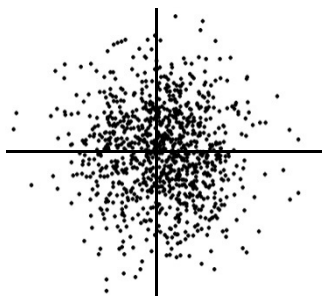
$$\begin{bmatrix} C_0 & C_1 & \cdots & C_{m-1} \\ C_1 & C_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & C_1 \\ C_{m-1} & \cdots & C_1 & C_0 \end{bmatrix}$$

自己相関関数

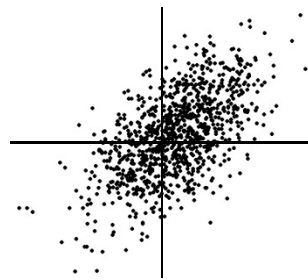
$$R_k \equiv \text{Cor}(y_n, y_{n-k}) = \frac{\text{Cov}(y_n, y_{n-k})}{\sqrt{\text{Var}(y_n)\text{Var}(y_{n-k})}}$$

- y_n と y_{n-k} の相関係数
- 定常時系列の場合

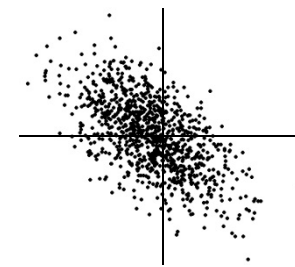
$$R_k = \frac{\text{Cov}(y_n, y_{n-k})}{\sqrt{\text{Var}(y_n)\text{Var}(y_{n-k})}} = \frac{C_k}{\sqrt{C_0 C_0}} = \frac{C_k}{C_0}$$



$$R_k = 0$$



$$R_k > 0$$

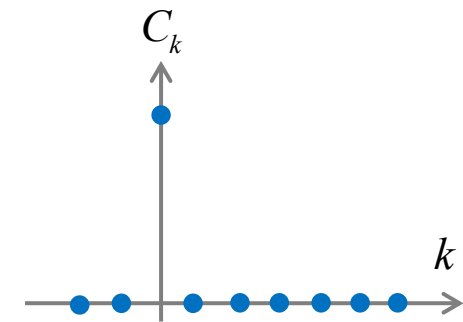


$$R_k < 0$$

白色雑音

時間的に相関がない時系列

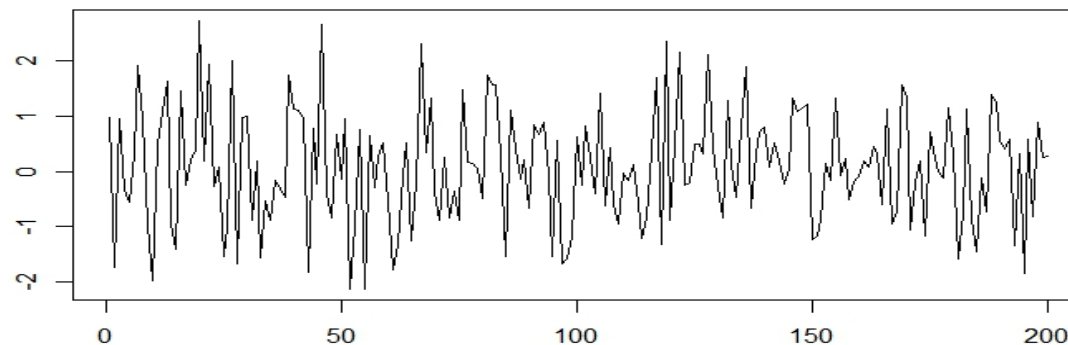
$$C_k = 0, \quad R_k = 0, \quad k \neq 0$$



白色雑音の自己共分散関数

モデリングにおいて重要

- 決定論的モデル $f(y_n, y_{n-1} \dots) = 0$
- 時系列モデル $f(y_n, y_{n-1} \dots) = w_n, \quad w_n: \text{白色雑音}$
- (例) 独立に生成した正規乱数列 `plot(as.ts(rnorm(200)))`



【例】簡単な予測の問題

$$\bar{y}_n = \alpha y_{n-k} + \beta$$

予測誤差

$$\begin{aligned} E(y_n - \bar{y}_n)^2 &= E(y_n - \alpha y_{n-k} - \beta)^2 \\ &= E(y_n^2 + \alpha^2 y_{n-k}^2 + \beta^2 - 2\alpha y_n y_{n-k} - 2\alpha\beta y_{n-k} - 2\beta y_n) \\ &= C_0 + \alpha^2 C_0 + \beta^2 - 2\alpha C_k \equiv J \end{aligned}$$

最小二乗法

$$\min_{\alpha, \beta} J = \min_{\alpha, \beta} E(y_n - \bar{y}_n)^2$$

$$\frac{\partial J}{\partial \beta} = 0 \quad \Rightarrow \quad \beta = 0$$

$$\frac{\partial J}{\partial \alpha} = 0 \quad \Rightarrow \quad 2\alpha C_0 - 2C_k = 0$$

$$\begin{aligned} \alpha &= \frac{C_k}{C_0} = R_k \\ \beta &= 0 \end{aligned}$$

時系列における推定について

独立なデータの場合 y_1, \dots, y_N

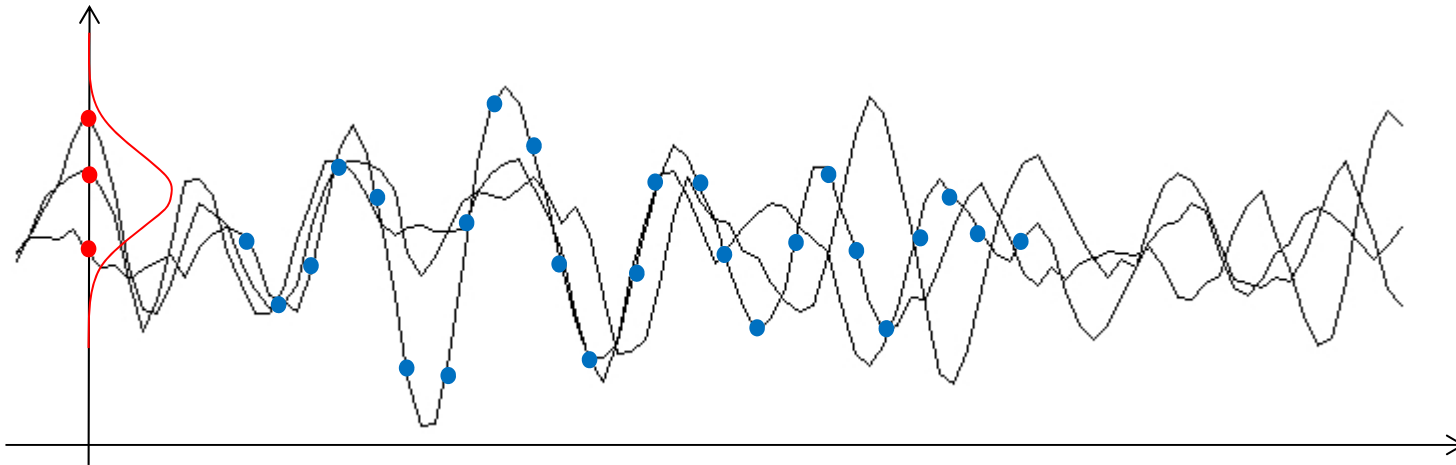
大数の法則
標本平均 \longrightarrow 平均

$$\hat{\mu} = \frac{1}{N} \sum_{n=1}^N y_n \Rightarrow \mu = E[y_n]$$

$$\hat{\sigma}^2 = \frac{1}{N-1} \sum_{n=1}^N (y_n - \hat{\mu})^2 \Rightarrow \sigma^2 = E(y_n - \mu)^2$$

定常時系列の場合

期待値(空間平均) = 時間平均 (エルゴード性)

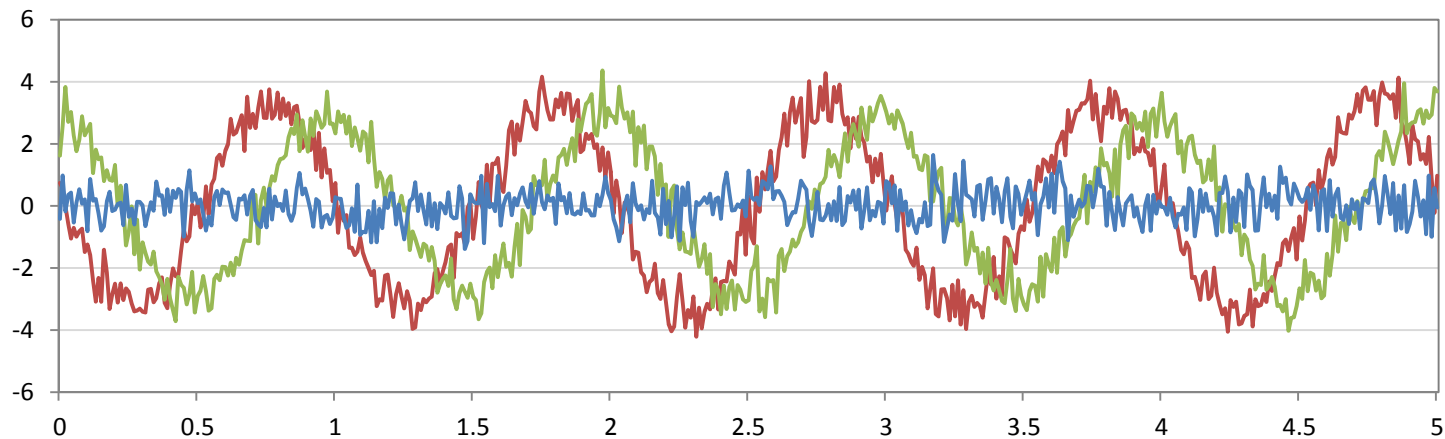


エルゴード性が成り立たない定常時系列

$X(\omega), \theta(\omega)$: 確率変数, $X \sim N(0, 25)$, $\theta = U[0, 2\pi)$

$$y_n = X(\omega) \sin(2\pi fn + \theta(\omega)) + w_n$$

y_n : 定常, $w_n \sim N(0, 0.25)$



空間平均 \neq 時間平均

$$E[y_n] = 0$$

$$E[y_n y_{n+k}] = E[X(\omega)^2 \sin(2\pi fn + \theta(\omega)) \sin(2\pi f(n+k) + \theta(\omega))] + 0.25$$

$$= E[X(\omega)^2] \frac{1}{2} E[\cos(2\pi fk) - \cos(4\pi fn + 2\theta(\omega) + 2\pi fk)] + 0.25 = \frac{25}{2} \cos(2\pi fk) + 0.25$$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N y_n y_{n+k} = \frac{X^2(\omega)}{2} \cos 2\pi fk + 0.25$$

自己共分散関数の推定

定常時系列 y_1, \dots, y_N

標本平均

$$\hat{\mu} = \frac{1}{N} \sum_{n=1}^N y_n$$

標本自己共分散関数

$$\hat{C}_k = \frac{1}{N} \sum_{n=k+1}^N (y_n - \hat{\mu})(y_{n-k} - \hat{\mu})$$

標本自己相関関数

$$\hat{R}_k = \frac{\hat{C}_k}{\hat{C}_0}$$

\hat{C}_k の平均

$$\hat{C}_k = \frac{1}{N} \sum_{n=k+1}^N (y_n - \hat{\mu})(y_{n-k} - \hat{\mu}) \quad \text{簡単のために } \mu=0 \text{ と仮定}$$

$$E[\hat{C}_k] = E\left[\frac{1}{N} \sum_{i=k+1}^n y_i y_{i-k}\right] = \frac{1}{N} \sum_{i=k+1}^n E[y_i y_{i-k}] = \frac{N-k}{N} C_k$$

- \hat{C}_k は不偏推定量ではない
- $1/N$ を $1/(N-k)$ に替えると不偏になる
- しかし、時系列解析では通常 N が使われる
 - ✓ 最尤推定に対応
 - ✓ 非負定値性を満たす
 - ✓ 常に定常モデルが推定される
 - ✓ ソフトを使うときは定義に注意

\hat{C}_k の分散

$$\text{Var}(\hat{C}_k) = \frac{1}{N^2} \sum_{i=k+1}^N \sum_{j=k+1}^N (C_{i-j}^2 + C_{i-j-k} C_{i-j+k}) \quad \text{正規分布のとき}$$

$$\begin{aligned} \text{Var}(\hat{C}_k) &= E \left[\left(\hat{C}_k - E[\hat{C}_k] \right)^2 \right] \\ &= E \left[\left(\frac{1}{N} \sum_{i=k+1}^N y_i y_{i-k} - \frac{N-k}{N} C_k \right)^2 \right] = \frac{1}{N^2} E \left[\left(\sum_{i=k+1}^N y_i y_{i-k} - (N-k) C_k \right)^2 \right] \\ &= \frac{1}{N^2} E \left[\left(\sum_{i=k+1}^N (y_i y_{i-k} - C_k) \right)^2 \right] = \frac{1}{N^2} \sum_{i=k+1}^N \sum_{j=k+1}^N E \left[(y_i y_{i-k} - C_k)(y_j y_{j-k} - C_k) \right] \\ &= \frac{1}{N^2} \sum_{i=k+1}^N \sum_{j=k+1}^N \left\{ E \left[(y_i y_{i-k} y_j y_{j-k}) \right] - C_k E \left[y_i y_{i-k} \right] - C_k E \left[y_j y_{j-k} \right] + C_k^2 \right\} \\ &= \frac{1}{N^2} \sum_{i=k+1}^N \sum_{j=k+1}^N (C_{i-j}^2 + C_{i-j-k} C_{i-j+k}) \end{aligned}$$

正規確率変数

$$\begin{aligned} &E[y_i y_{i-k} y_j y_{j-k}] \\ &= E[y_i y_{i-k}] E[y_j y_{j-k}] + E[y_i y_j] E[y_{i-k} y_{j-k}] + E[y_i y_{j-k}] E[y_{i-k} y_j] \\ &= C_k^2 + C_{i-j}^2 + C_{i-j+k} C_{i-j-k} \end{aligned}$$

推定量の性質(白色雑音の場合)

$$\hat{C}_0 \sim N(C_0, \frac{2\sigma^4}{N})$$

$$\hat{C}_k \sim N(0, \frac{\sigma^4}{N}) \quad (k \neq 0)$$

$$\hat{C}_k = \frac{1}{N} \sum_{i=k+1}^N y_i y_{i-k}$$

$$E[\hat{C}_k] = \frac{N-k}{N} C_k = 0$$

$$\text{Var}(C_k) = \frac{1}{N^2} \sum_{i=k+1}^N \sum_{j=k+1}^N (C_{i-j}^2 + C_{i-j-k} C_{i-j+k})$$

$$\text{Var}(C_0) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (C_{i-j}^2 + C_{i-j}^2) = \frac{2}{N} C_0^2 = \frac{2\sigma^4}{N}$$

$$\text{Var}(C_k) = \frac{1}{N^2} \sum_{i=k+1}^N \sum_{j=k+1}^N (C_{i-j}^2 + 0) = \frac{N-k}{N^2} C_0^2 \cong \frac{\sigma^4}{N}, \quad (k \neq 0)$$

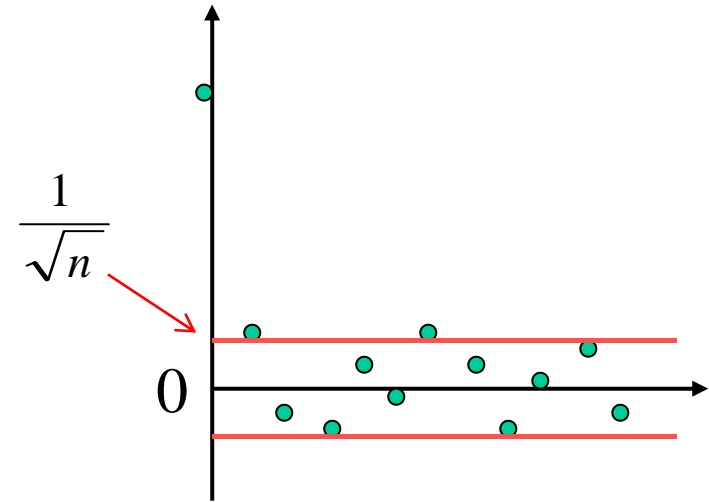
推定量の性質(白色雑音の場合)

$$C_0 = \sigma^2, \quad R_0 = 1, \quad C_k = R_k = 0, \quad N : \text{データ数}$$

$$\hat{C}_0 \sim N\left(C_0, \frac{2\sigma^4}{N}\right)$$

$$\hat{C}_k \sim N\left(0, \frac{\sigma^4}{N}\right)$$

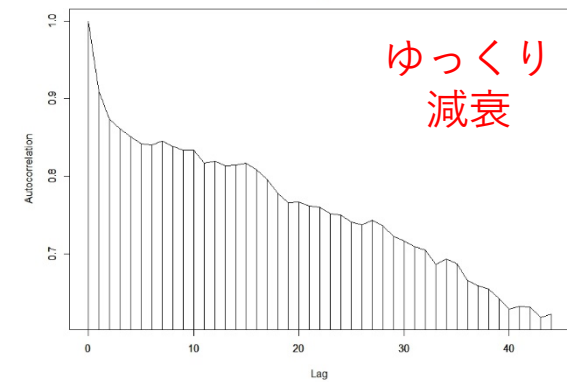
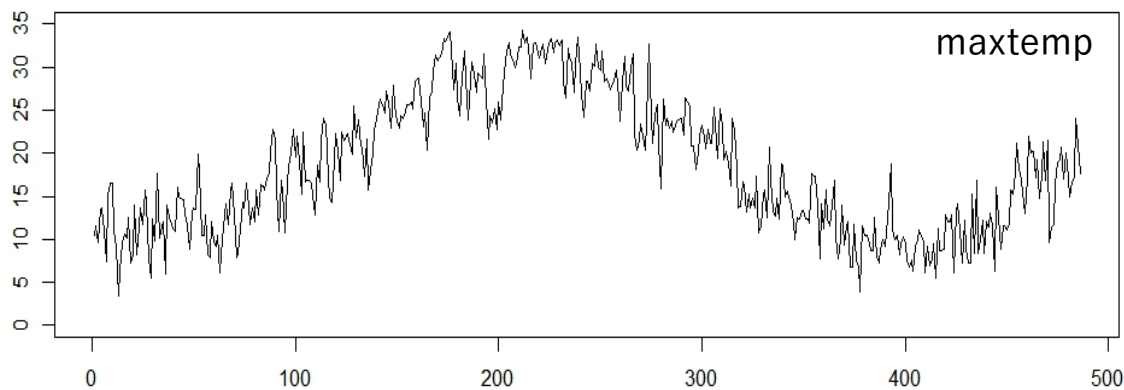
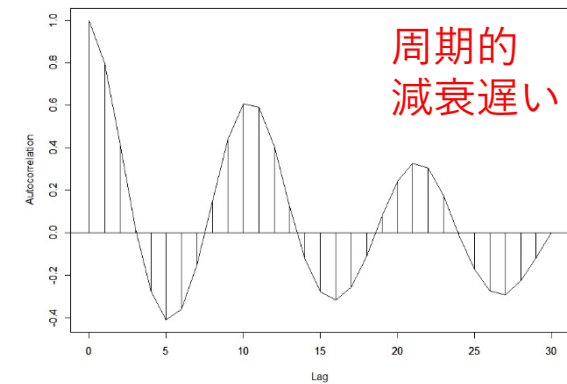
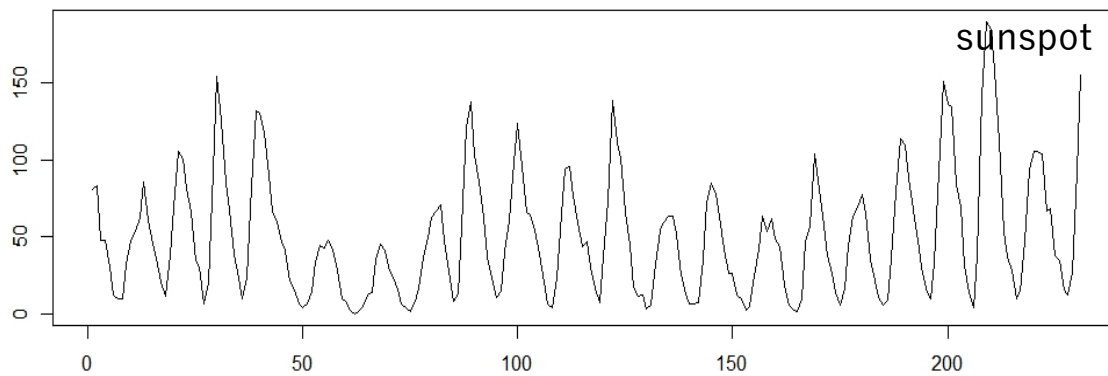
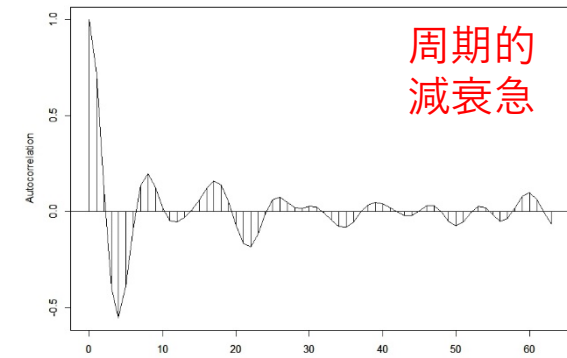
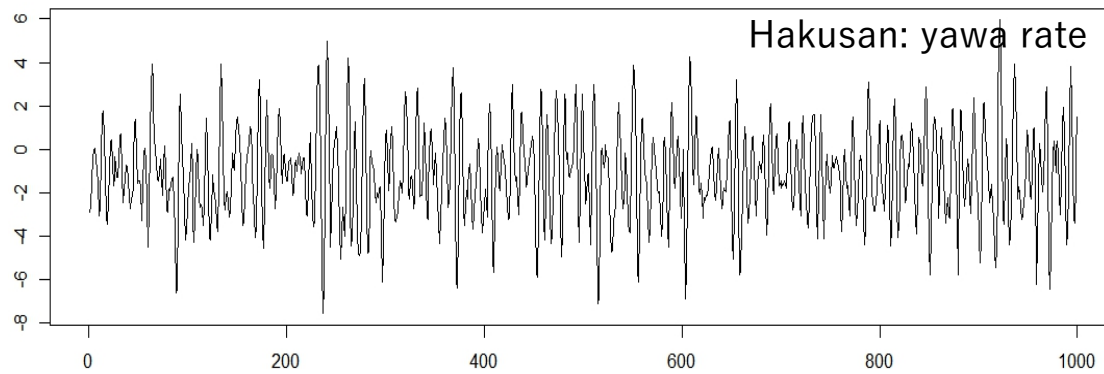
$$\hat{R}_k \sim N\left(0, \frac{1}{N}\right)$$



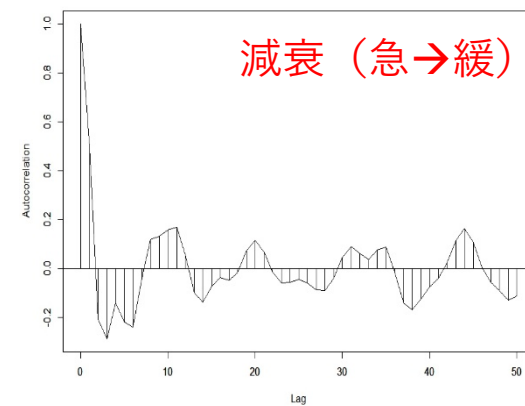
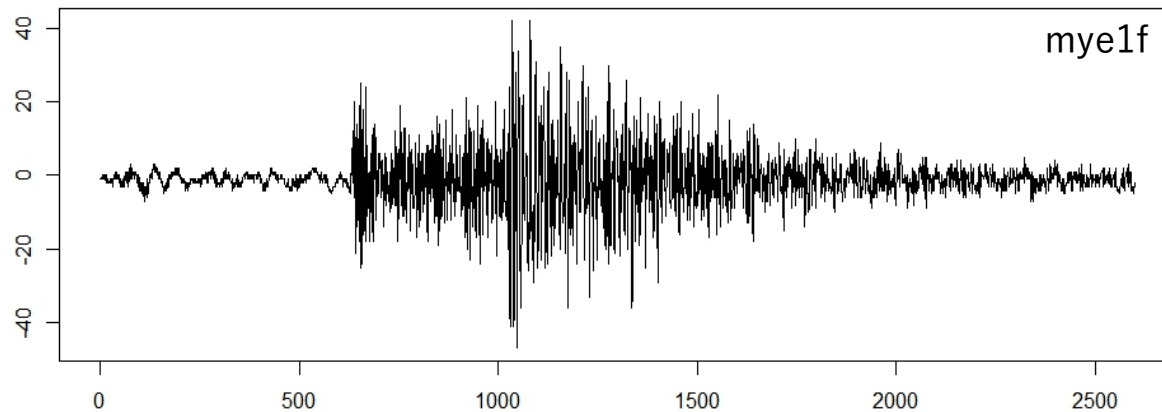
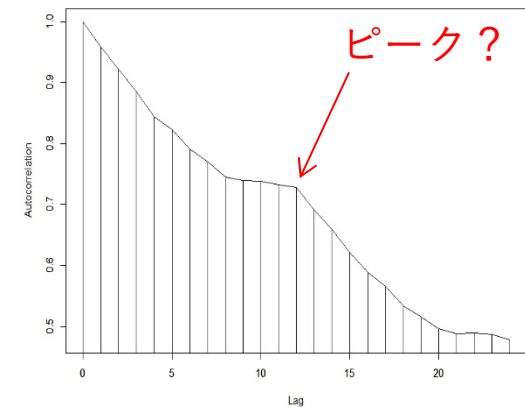
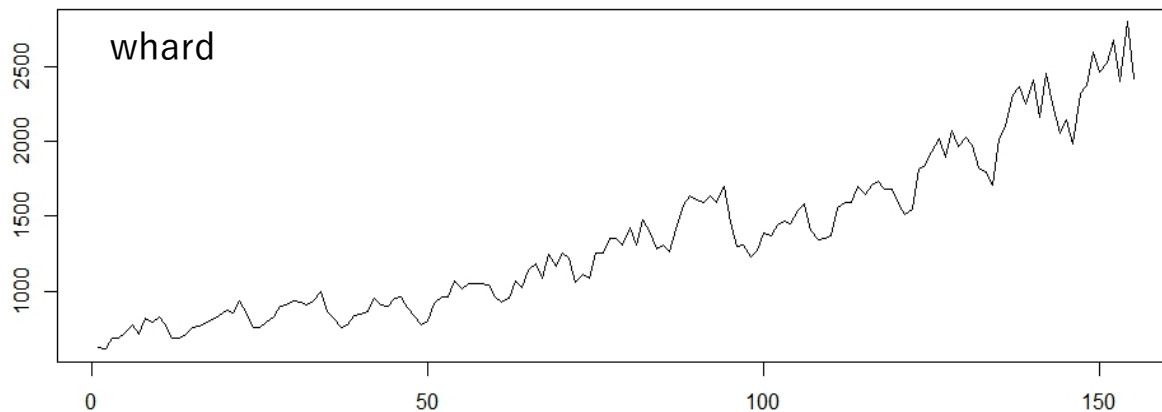
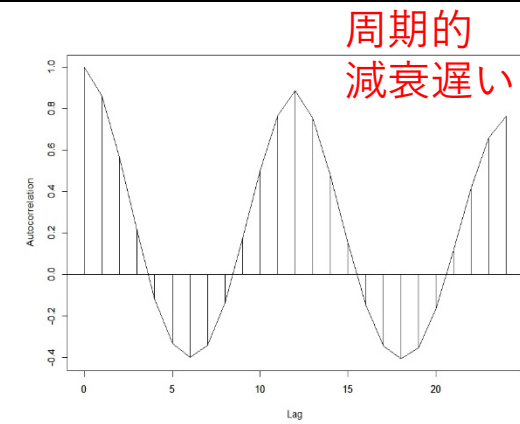
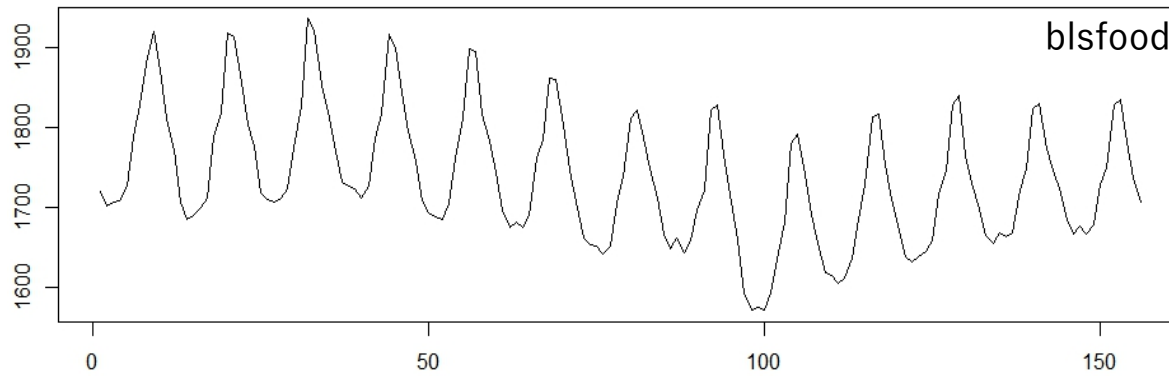
N	25	100	1000	10000
$\frac{1}{\sqrt{N}}$	0.2	0.1	0.03	0.01

- 簡易な白色性評価に用いられる

自己相関関数



自己相関関数



推定量の性質(一般の場合)

$$E[\hat{C}_k] = \frac{N-k}{N} C_k$$

$$\text{Var}(\hat{C}_k) \sim \frac{1}{N} \sum_{j=-\infty}^{\infty} (C_j^2 + C_{j-k} C_{j+k})$$

$$\text{Var}(\hat{R}_k) \sim \frac{1}{N} \sum_{j=-\infty}^{\infty} (R_j^2 + R_{j-k} R_{j+k} - 4R_j R_k R_{j-k} + 2R_j^2 R_k^2)$$

μ_n 既知と仮定

一般の場合:

自己相関関数の分散を一般的に求めるのは難しい。

Bartlett(1946)

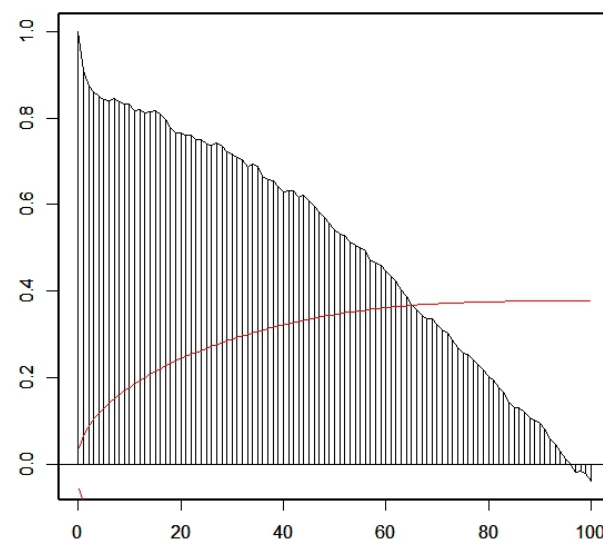
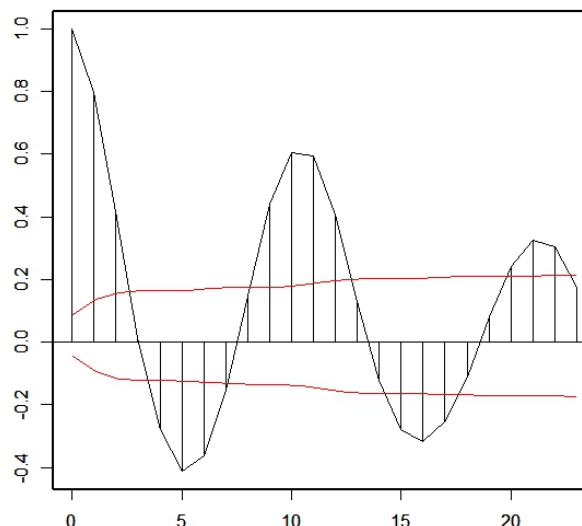
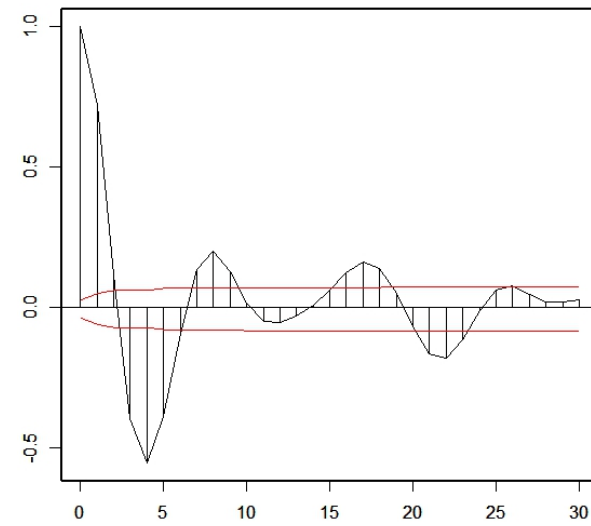
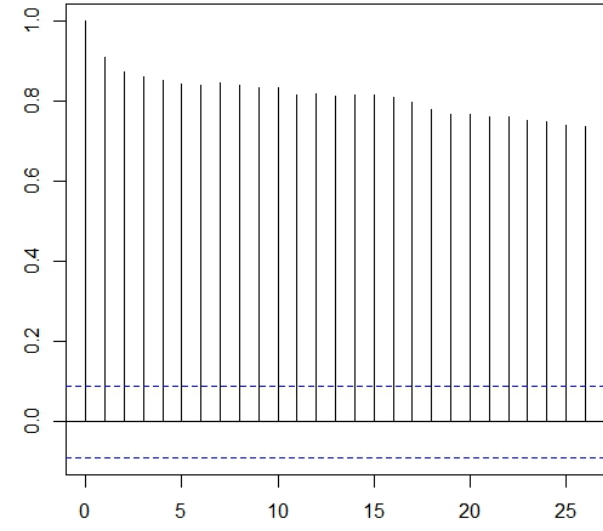
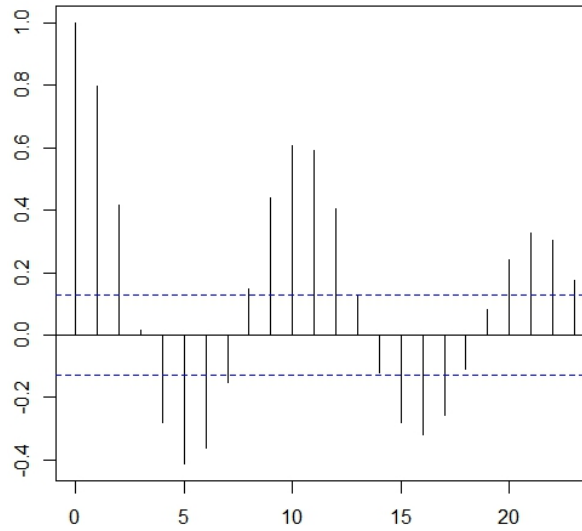
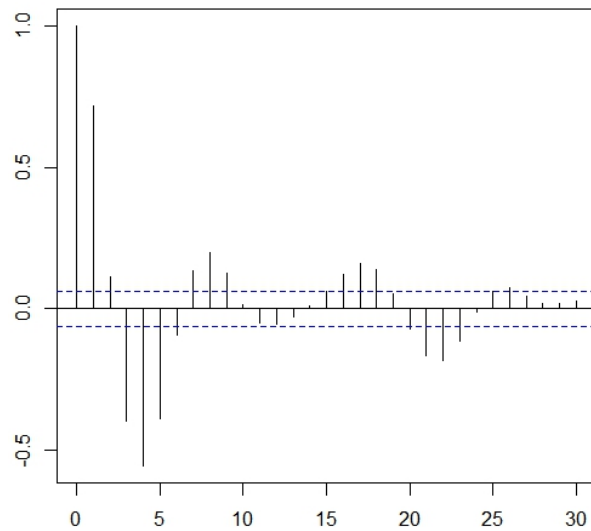
$$y_n = w_n + b_1 w_{n-1} + \cdots + b_\ell w_{n-\ell} \quad \text{の場合}$$
$$C_k = (1 + b_1^2 + \cdots + b_\ell^2) \sigma^2$$

$$\text{Var}(\hat{R}_k) \cong \frac{1}{N} \left(1 + 2 \sum_{j=1}^{k-1} \hat{R}_j^2 \right)$$

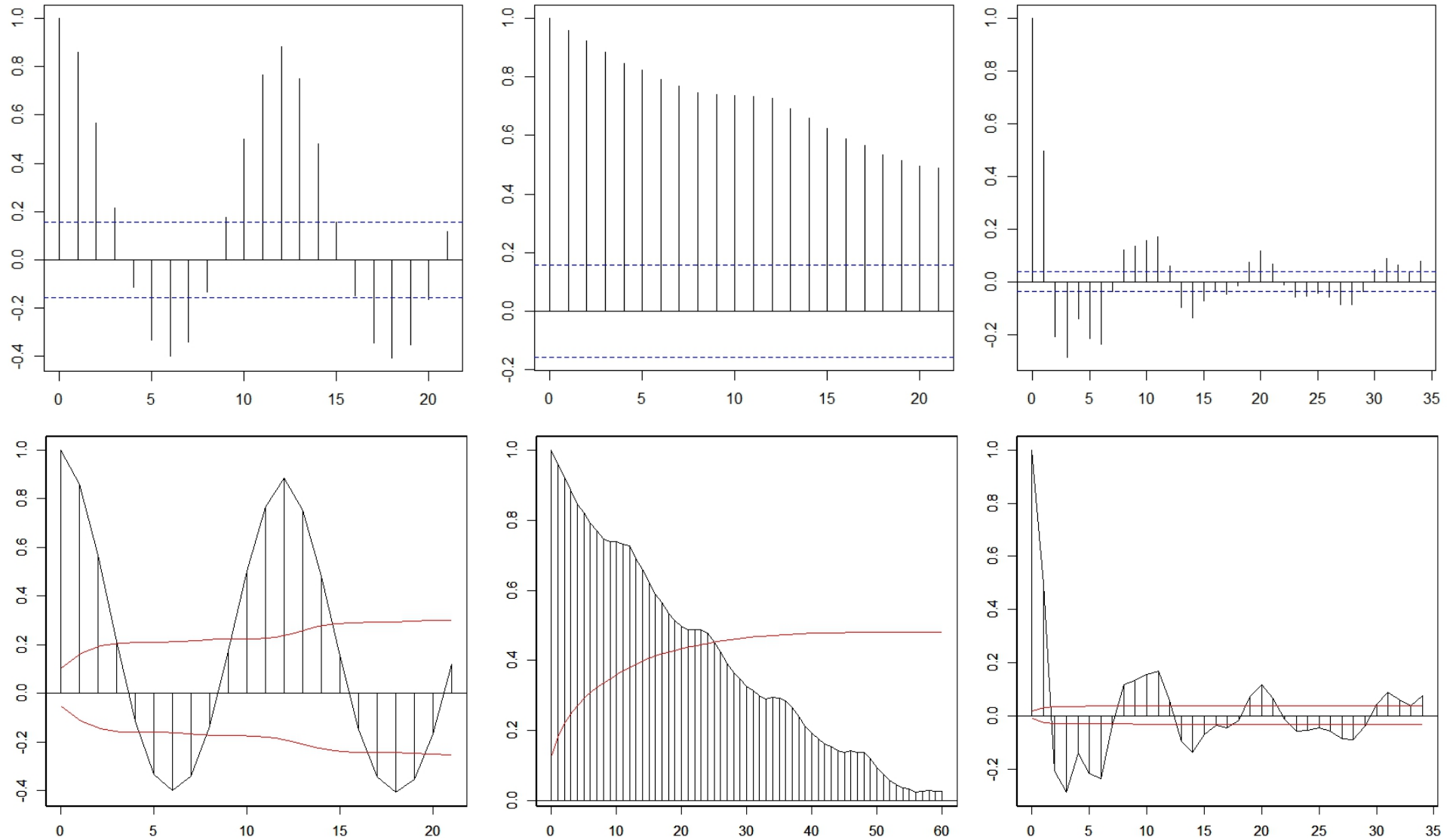
Box-Jenkins(1976)

$$\text{Var}(\hat{R}_k) \cong \frac{1}{N} \left(\frac{N-k}{N+2} \right)$$

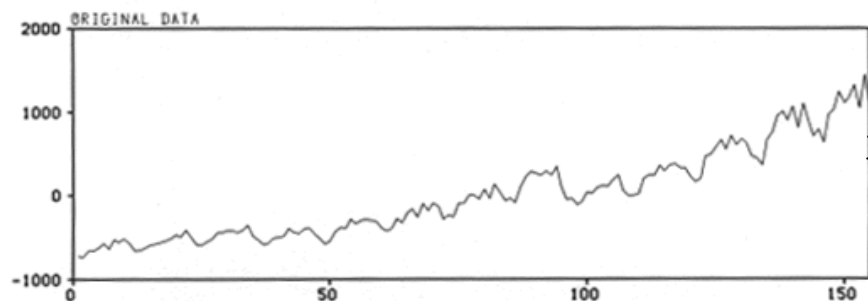
標準誤差（白色雑音と一般の場合）



標準誤差（白色雑音と一般の場合）



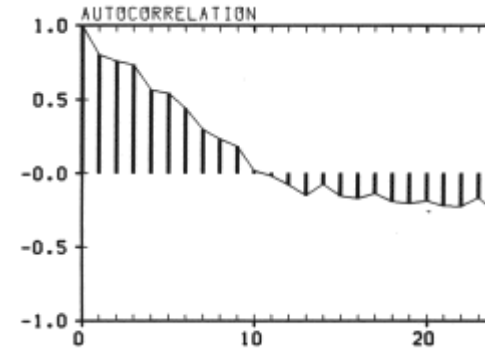
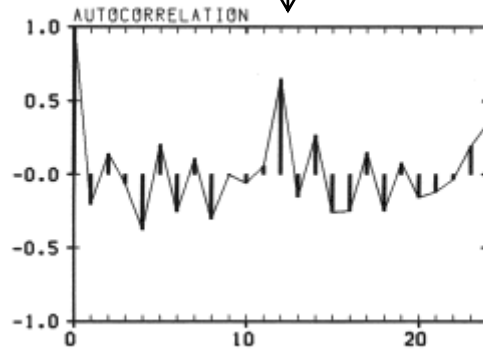
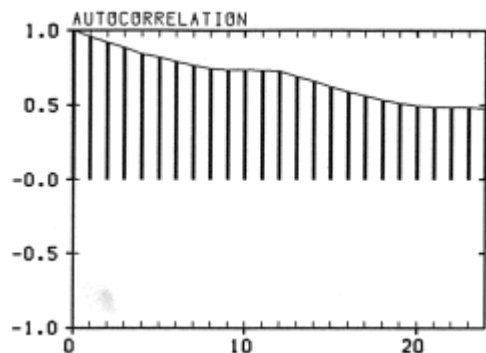
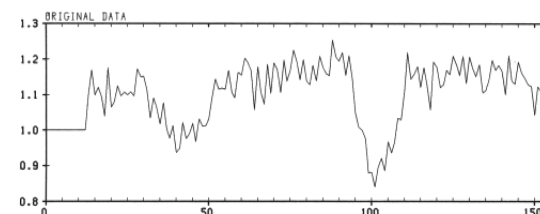
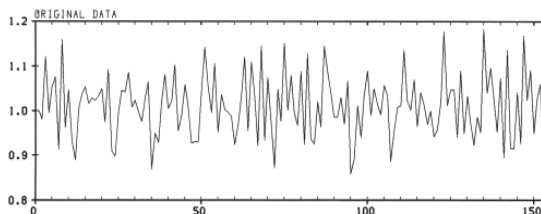
前処理の影響 (WHARD データ)



原データ

前期比

前年同期比



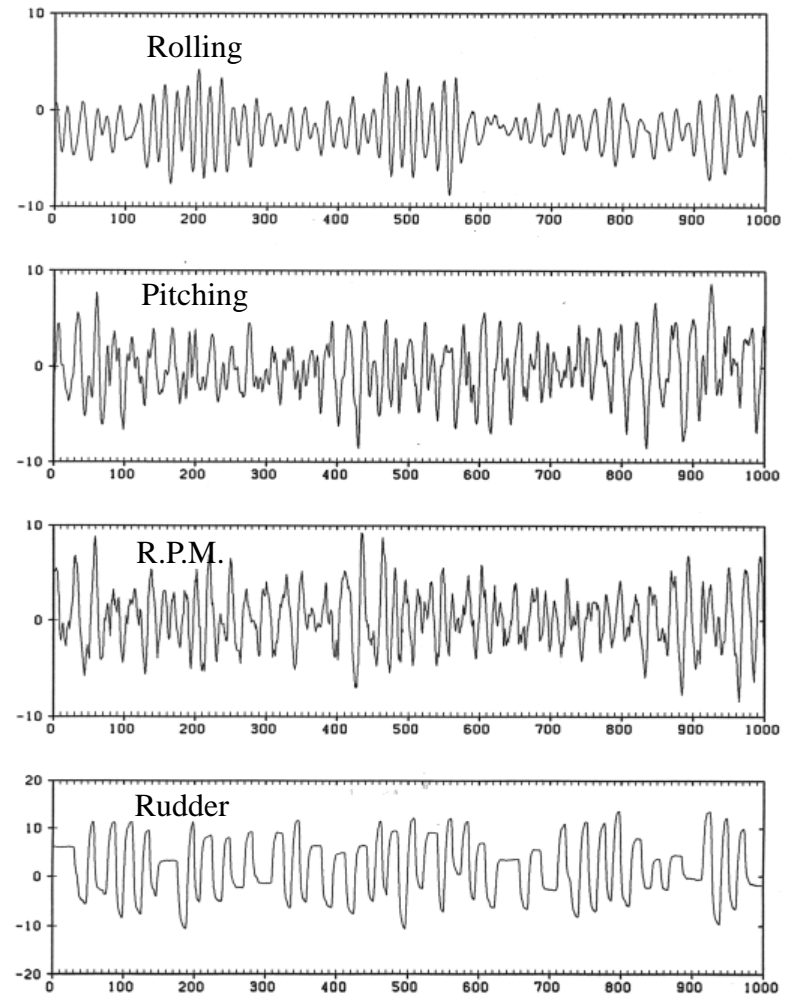
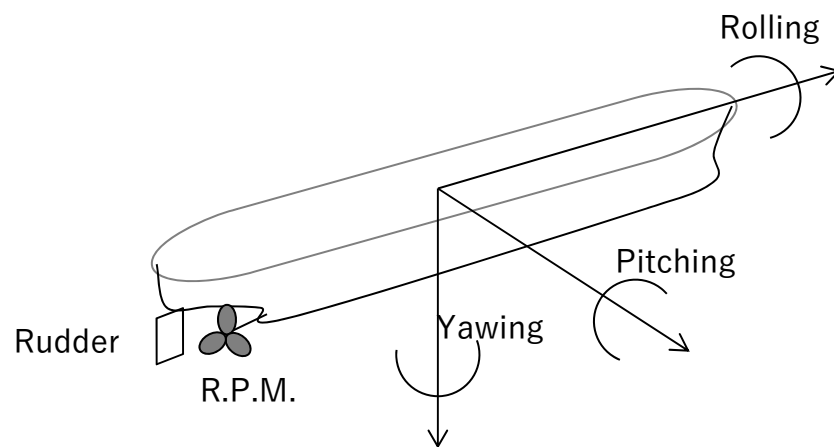
多変量時系列

$$y_n \in R^m, n = 1, \dots, N$$

多変量 (m 変量) 時系列

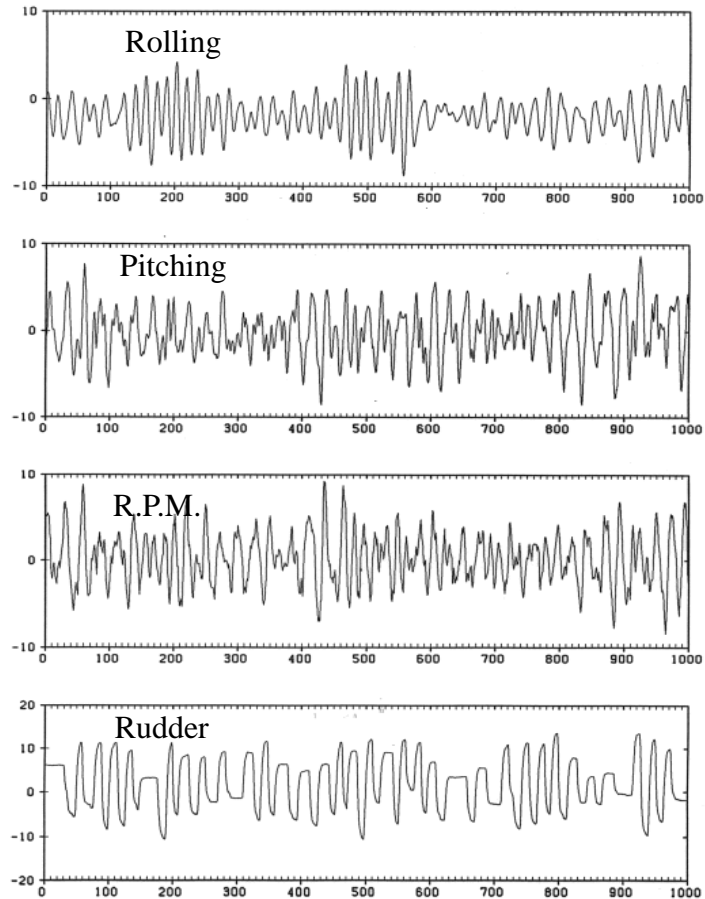
$$y_n = \begin{bmatrix} y_n(1) \\ y_n(2) \\ \vdots \\ y_n(m) \end{bmatrix}$$

関連する複数の時系列を
同時に記録したもの

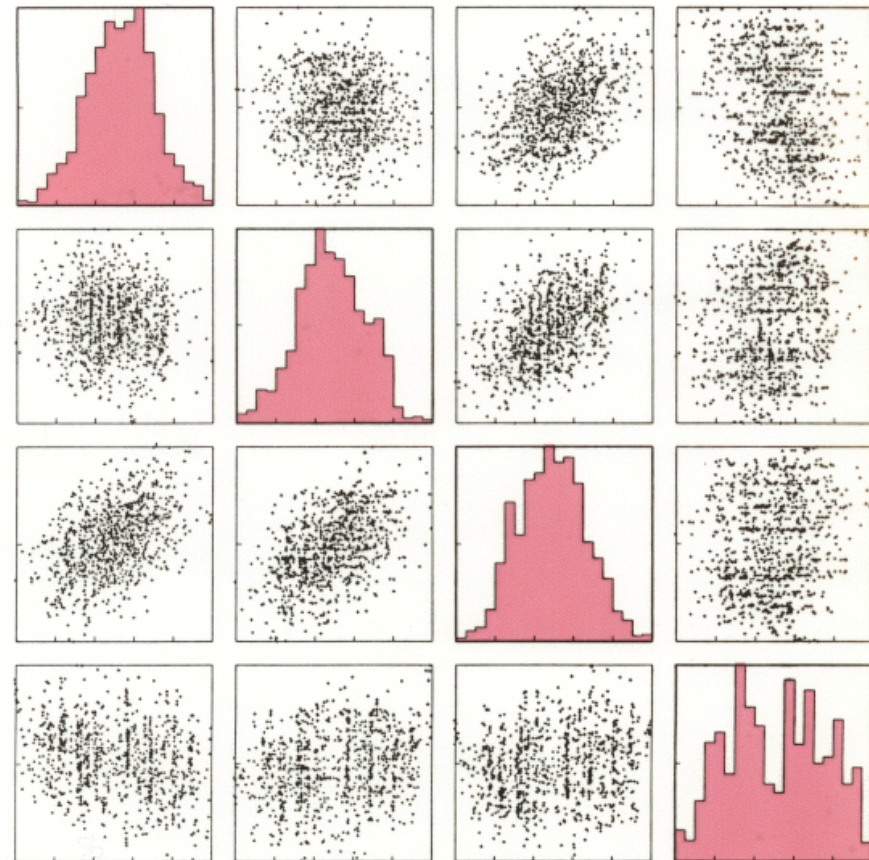


多変量時系列

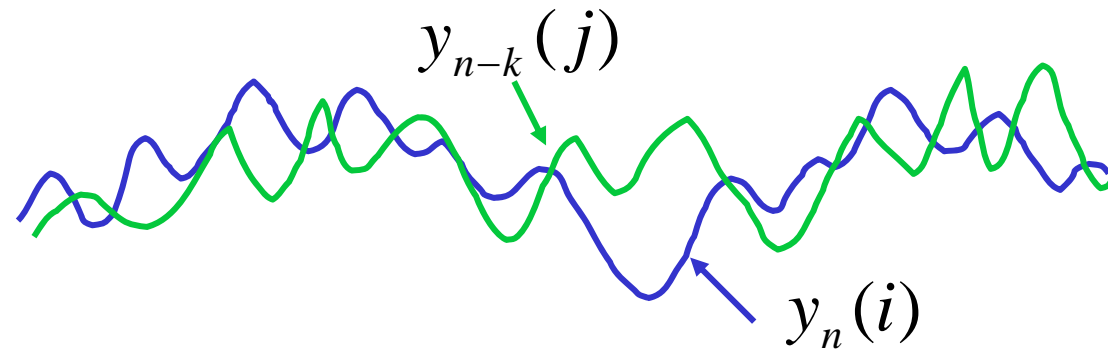
時系列プロット



ヒストグラムと散布図



定常時系列の相互共分散関数



$$\begin{aligned} C_k(i, j) &= \text{Cov}(y_n(i), y_{n-k}(j)) \\ &= E[(y_n(i) - \mu_i)(y_{n-k}(j) - \mu_j)] \end{aligned}$$

$$C_k = \begin{bmatrix} C_k(1,1) & \cdots & C_k(1,m) \\ \vdots & \ddots & \vdots \\ C_k(m,1) & \cdots & C_k(m,m) \end{bmatrix}$$

相互相関関数

$$R_k(i, j) = \frac{C_k(i, j)}{\sqrt{C_0(i, i)C_0(j, j)}}$$

$$R_k = \begin{bmatrix} R_k(1,1) & \cdots & R_k(1,m) \\ \vdots & \ddots & \vdots \\ R_k(m,1) & \cdots & R_k(m,m) \end{bmatrix}$$

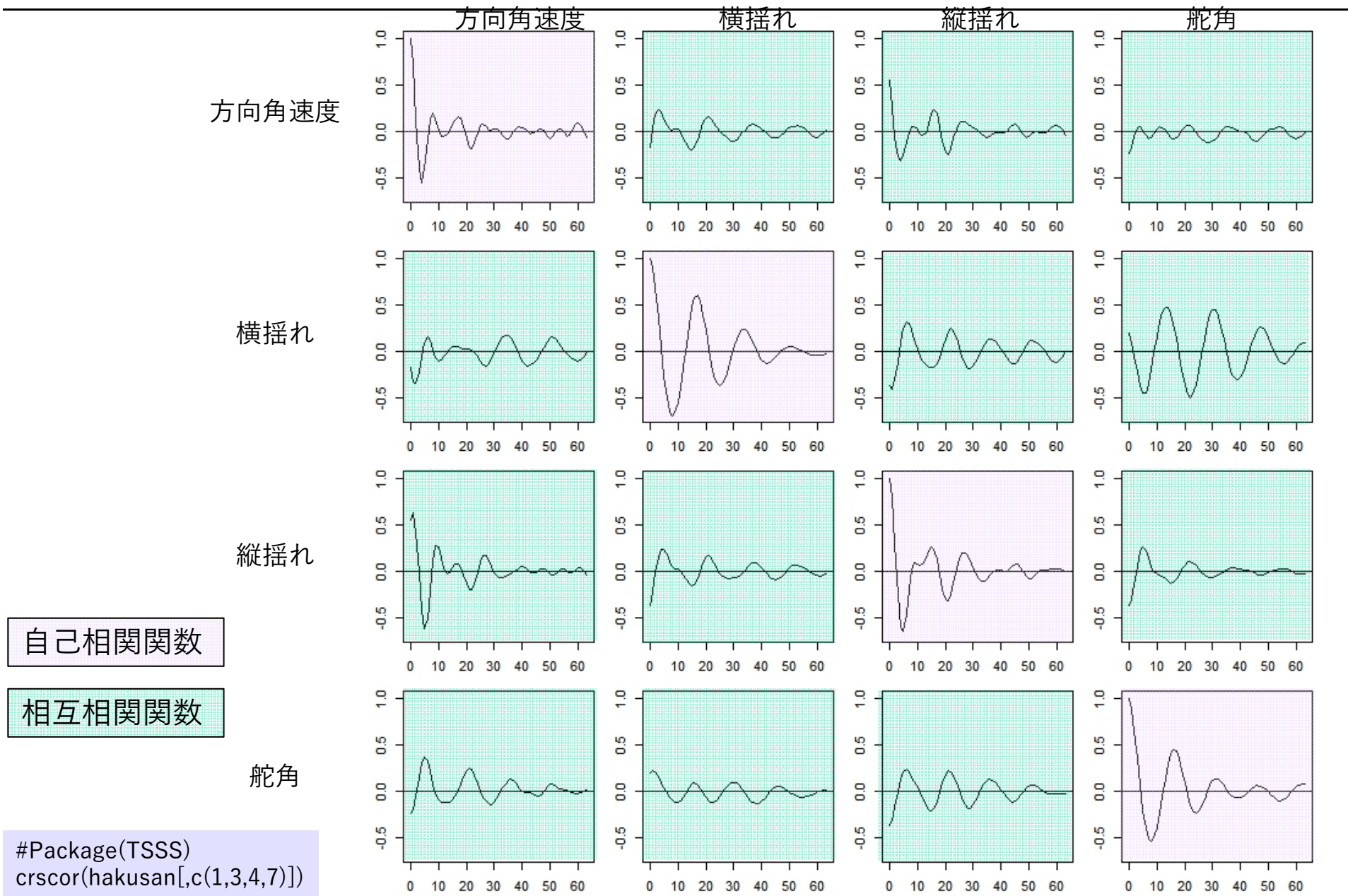
相互共分散（相関）関数の性質

$$C_{-k} = C_k^T, \quad R_{-k} = R_k^T$$

$$\begin{aligned} C_{-k}(i, j) &= \text{Cov}(y_n(i), y_{n+k}(j)) \\ &= \text{Cov}(y_{n+k}(j), y_n(i)) \\ &= \text{Cov}(y_n(j), y_{n-k}(i)) \leftarrow \text{定常性} \\ &= C_k(j, i) \end{aligned}$$

- ✓ 相互相関関数は偶関数ではない
- ✓ C_0, C_1, \dots は十分な情報を持つ

相互共分散関数 (船舶データ)



自己相関関数

相互相関関数

```
#Package(TSSS)
crscor(hakusan[,c(1,3,4,7)])
```


周期性の可視化(検出)

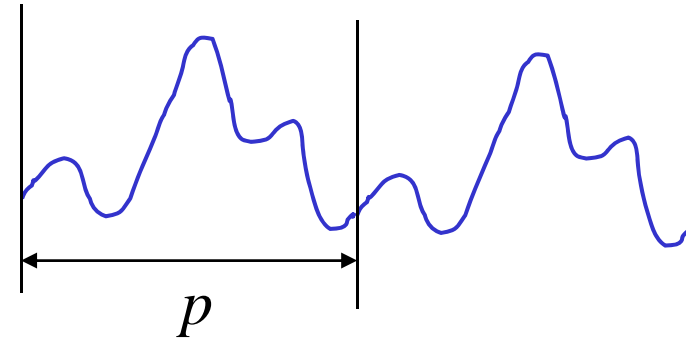
周期性とは

$$y_t = y_{t-p} \quad \begin{array}{l} \text{がなりたつ} \\ p \quad \text{周期} \end{array}$$

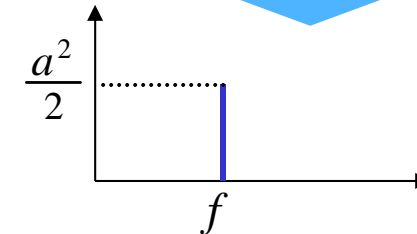
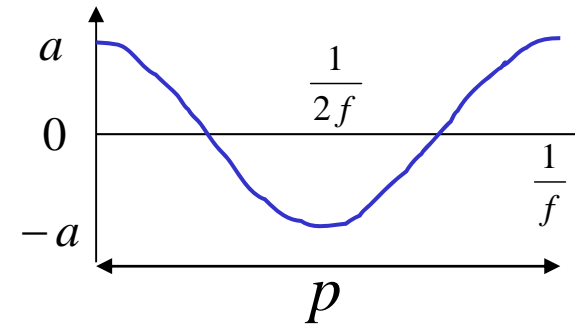
(例) 三角関数

$$y_n = a \cos(2\pi fn)$$
$$C_k = \frac{a^2}{2} \cos(2\pi fk)$$

(f, a) 振動数と振幅



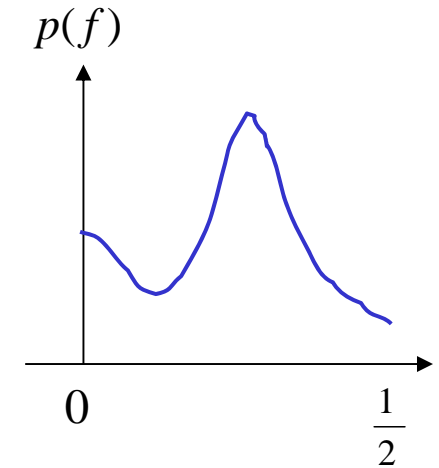
三角関数



(パワー) スペクトル

$$C_k \quad \text{自己共分散関数} \quad \sum_{k=-\infty}^{\infty} |C_k| < \infty$$

$$p(f) = \sum_{k=-\infty}^{\infty} C_k e^{-2\pi i k f} \quad -\frac{1}{2} \leq f \leq \frac{1}{2}$$



$$C_{-k} = C_k$$

$$\begin{aligned} p(f) &= C_0 + \sum_{k=1}^{\infty} C_k (e^{-2\pi i k f} + e^{2\pi i k f}) \\ &= C_0 + 2 \sum_{k=1}^{\infty} C_k \cos 2\pi k f \end{aligned}$$

$$e^{2\pi i k f} = \cos 2\pi k f + i \sin 2\pi k f$$

$$e^{-2\pi i k f} = \cos 2\pi k f - i \sin 2\pi k f$$

$$e^{2\pi i k f} + e^{-2\pi i k f} = 2 \cos 2\pi k f$$

パワースペクトルの性質

- 偶関数 $p(-f) = p(f)$

$$p(-f) = \sum_{k=-\infty}^{\infty} C_k e^{2\pi i k f} = \sum_{k=-\infty}^{\infty} C_{-k} e^{-2\pi i k f} = \sum_{k=-\infty}^{\infty} C_k e^{-2\pi i k f} = p(f)$$

$C_k = C_{-k}$

- コサイン変換 $p(f) = C_0 + 2 \sum_{k=1}^{\infty} C_k \cos(2\pi k f)$

$$\begin{aligned} p(f) &= \sum_{k=-\infty}^{\infty} C_k e^{-2\pi i k f} \\ &= \sum_{k=-\infty}^{\infty} C_k \{ \cos(2\pi k f) + i \sin(2\pi k f) \} \\ &= \sum_{k=-\infty}^{\infty} C_k \cos(2\pi k f) \\ &= C_0 + 2 \sum_{k=1}^{\infty} C_k \cos(2\pi k f) \end{aligned}$$

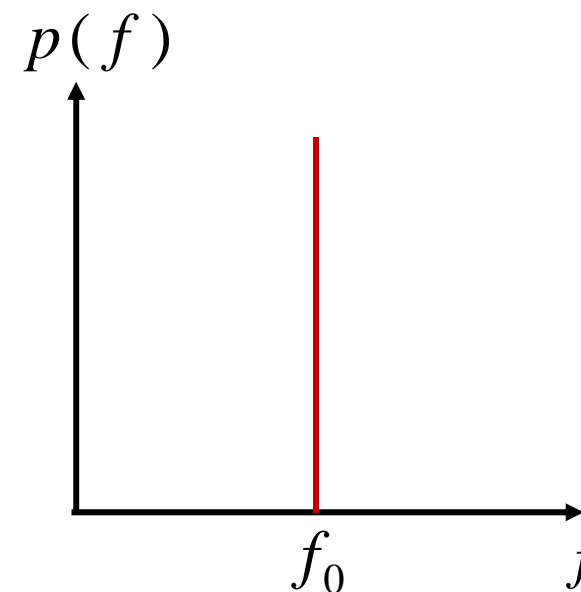
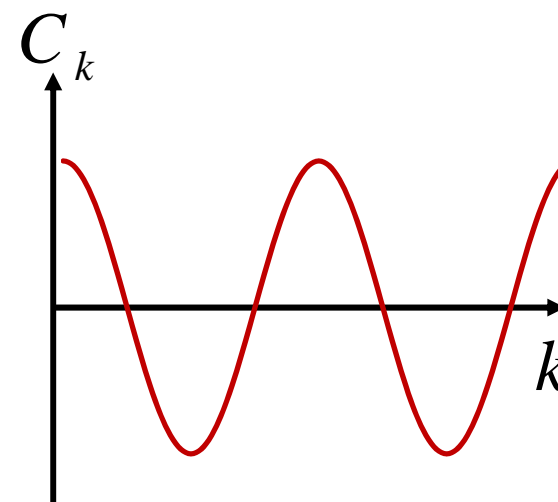
- 逆変換 $C_k = \int_{-\frac{1}{2}}^{\frac{1}{2}} p(f) e^{2\pi i k f} df = \int_{-\frac{1}{2}}^{\frac{1}{2}} p(f) \cos(2\pi k f) df$

(例 1) 単振動

$$C_k = \frac{a^2}{2} \cos(2\pi f_0 k)$$

$$\begin{aligned} p(f) &= \sum_{k=-\infty}^{\infty} C_k \cos(2\pi kf) \\ &= \sum_{k=-\infty}^{\infty} \cos(2\pi kf_0) \cos(2\pi kf) \\ &= \frac{1}{2} \sum_{k=-\infty}^{\infty} \cos(2\pi k(f_0 - f)) - \cos(2\pi k(f_0 + f)) \\ &= \begin{cases} \frac{1}{2} \delta & \text{if } f = f_0 \\ 0 & \text{if } f \neq f_0 \end{cases} \end{aligned}$$

- 単色光のイメージ
- 完全な周期関数
- ラインスペクトル (デルタ関数)

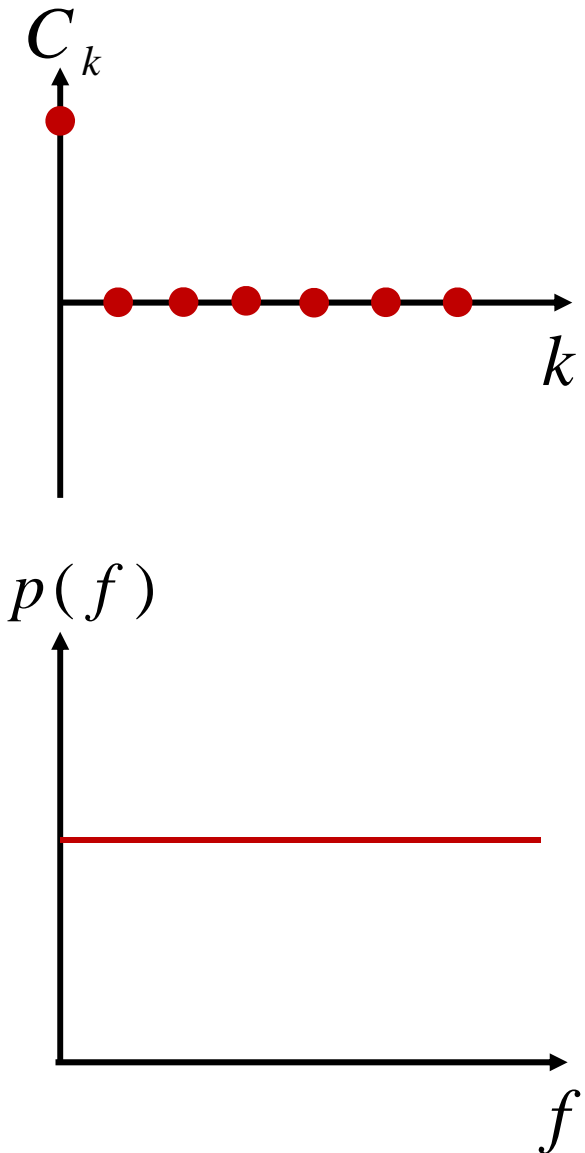


(例2) 白色雑音

$$C_0 = \sigma^2, \quad C_k = 0 \quad k \neq 0$$

$$\begin{aligned} p(f) &= \sum_{k=-\infty}^{\infty} C_k \cos(2\pi kf) \\ &= C_0 \\ &= \sigma^2 \end{aligned}$$

- 白色光のイメージ
- 無相関な時系列
- スペクトルはフラット (あらゆる周期が等量)



自己回帰モデルAR(1)のスペクトル

$$y_n = ay_{n-1} + w_n \quad w_n \sim N(0, \sigma^2)$$

$$C_k = \sigma^2 (1 - a^2)^{-1} a^{|k|}$$

$$C_0 = aC_1 + \sigma^2$$

$$C_k = aC_{k-1}$$

$$C_k = \sigma^2 (1 - a^2)^{-1} a^{|k|}$$

$$p(f) = \frac{\sigma^2}{|1 - ae^{-2\pi if}|^2} = \frac{\sigma^2}{1 - 2a \cos 2\pi f + a^2}$$

$$\begin{aligned} p(f) &= \frac{\sigma^2}{1 - a^2} \left\{ 1 + \sum_{k=1}^{\infty} a^k e^{-2\pi ifk} + \sum_{k=1}^{\infty} a^k e^{2\pi ifk} \right\} \\ &= \frac{\sigma^2}{1 - a^2} \left\{ 1 + \frac{ae^{-2\pi if}}{1 - ae^{-2\pi if}} + \frac{ae^{2\pi if}}{1 - ae^{2\pi if}} \right\} \\ &= \frac{\sigma^2}{|1 - ae^{-2\pi if}|^2} = \frac{\sigma^2}{1 - 2a \cos 2\pi f + a^2} \end{aligned}$$

AR(2)のスペクトル

$$y_n = a_1 y_{n-1} + a_2 y_{n-2} + w_n \quad w_n \sim N(0, \sigma^2)$$

$$C_0 = \frac{(1-a_2)\sigma^2}{(1+a_2)(1-2a_2+a_2^2-a_1^2)}, \quad C_1 = \frac{a_1}{(1-a_2)}C_0, \quad C_k = a_1 C_{k-1} + a_2 C_{k-2}$$

$$R_1 = \frac{a_1}{1-a_2}, \quad R_k = a_1 R_{k-1} + a_2 R_{k-2} \quad (k = 2, 3, \dots)$$

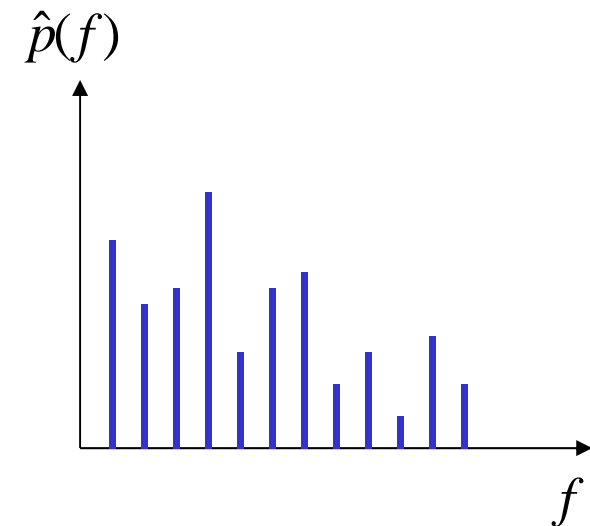
$$\begin{aligned} p(f) &= \frac{\sigma^2}{\left|1 - a_1 e^{-2\pi i f} - a_2 e^{-4\pi i f}\right|^2} \\ &= \frac{\sigma^2}{1 - 2a_1(1-a_2)\cos 2\pi f - 2a_2 \cos 4\pi f + a_1^2 + a_2^2} \end{aligned}$$

ピリオドグラム

$$y_1, \dots, y_N \quad \longrightarrow \quad \hat{C}_0, \dots, \hat{C}_{N-1}$$

$$\begin{aligned} \hat{p}(f) &= \sum_{k=-N+1}^{N-1} \hat{C}_k e^{-2\pi i k f} \\ &= \hat{C}_0 + 2 \sum_{k=1}^{N-1} \hat{C}_k \cos(2\pi k f) \end{aligned}$$
$$f = 0, \frac{1}{N}, \frac{2}{N}, \dots, \frac{1}{2}$$

$\hat{C}_0, \dots, \hat{C}_{N-1}$ のフーリエ変換

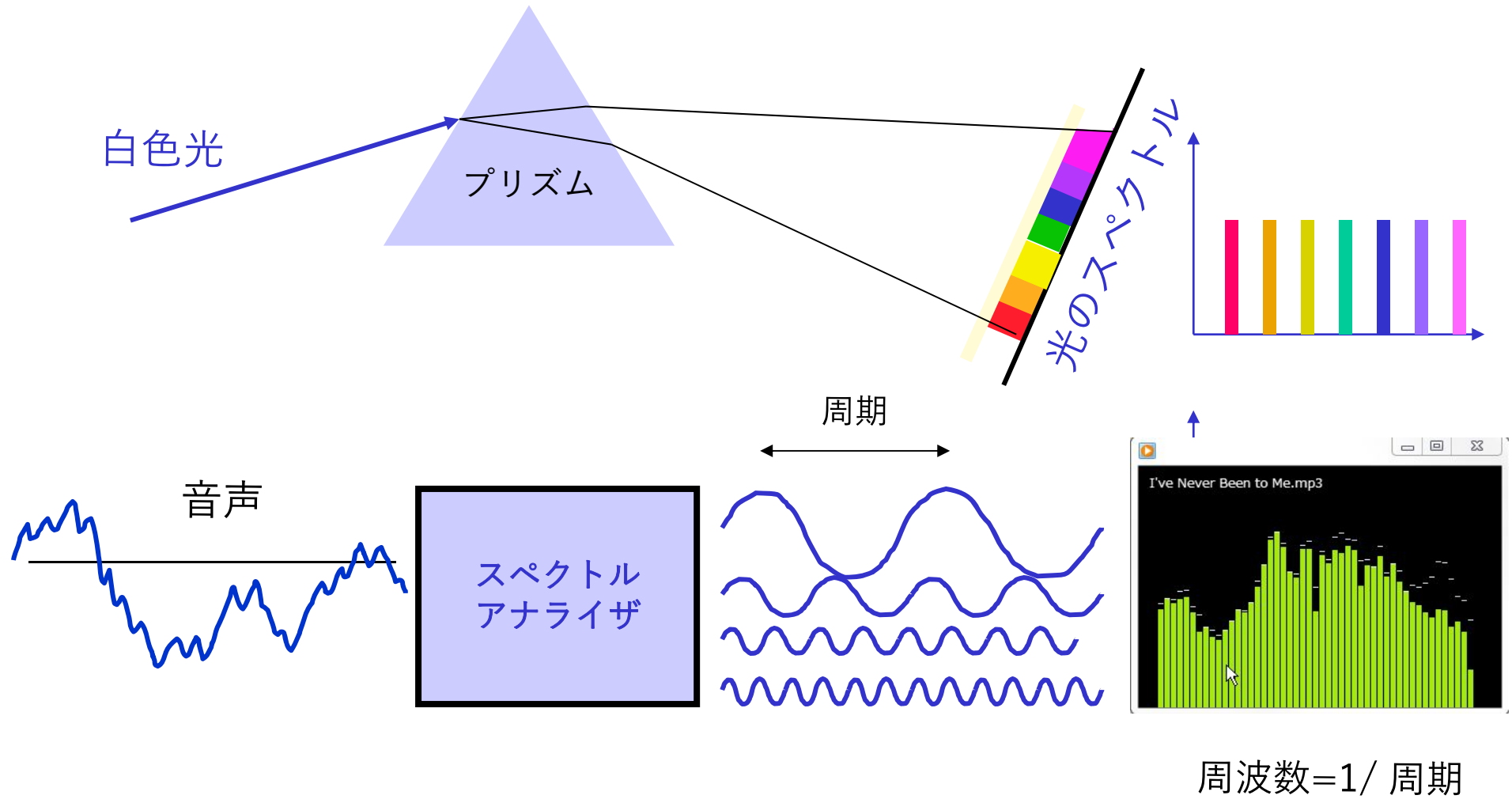


$$\hat{C}_k = \int_{-\frac{1}{2}}^{\frac{1}{2}} \hat{p}(f) e^{2\pi i k f} df \quad k = 0, 1, \dots, N-1$$

$$\hat{p}(f) = \sum_{k=-N+1}^{N-1} \hat{C}_k e^{-2\pi i k f} \quad -\frac{1}{2} \leq f \leq \frac{1}{2}$$

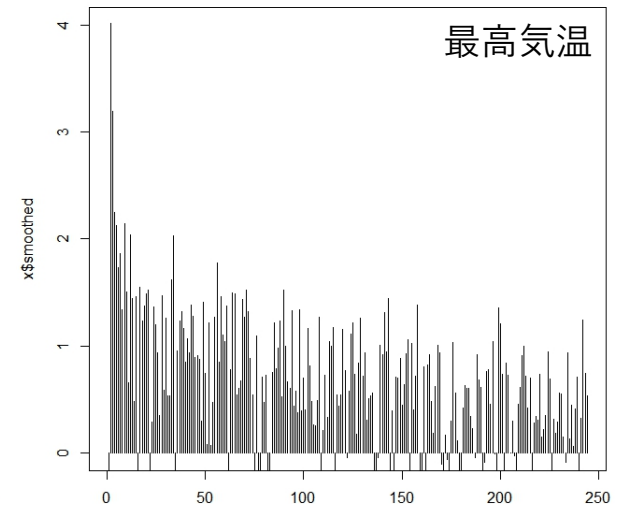
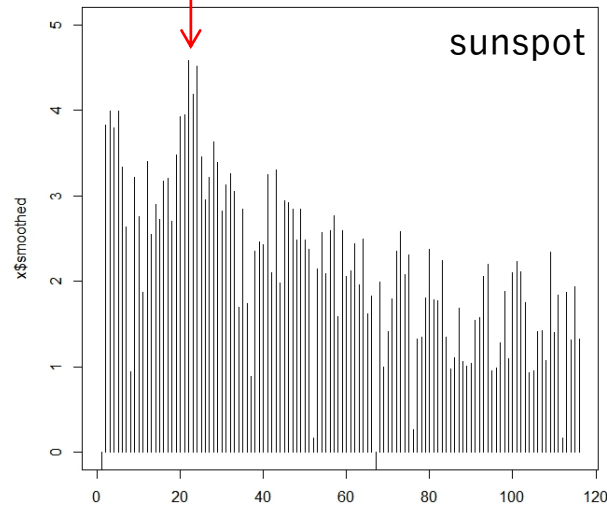
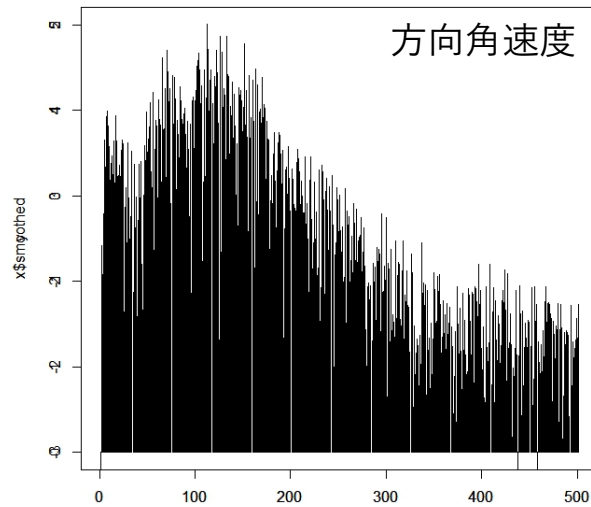
標本スペクトル

スペクトル分解

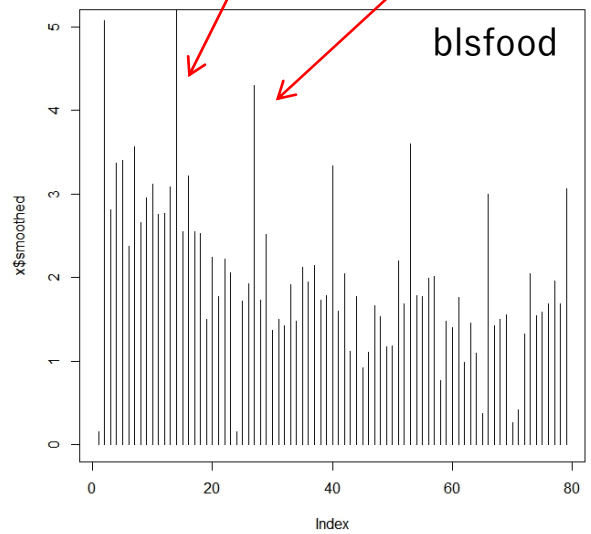


ピリオドグラム

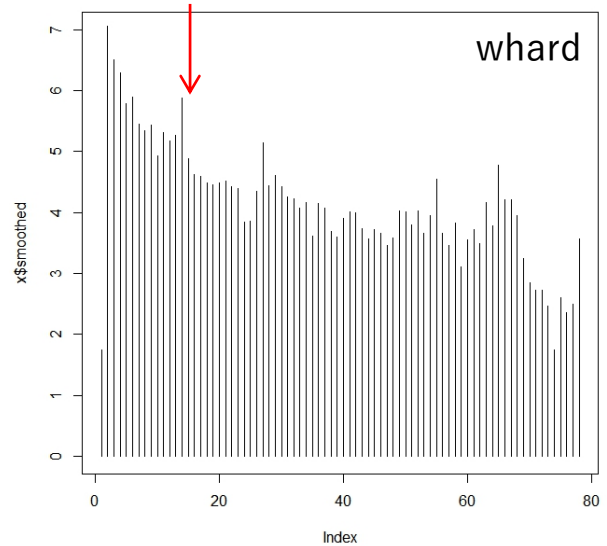
11年周期



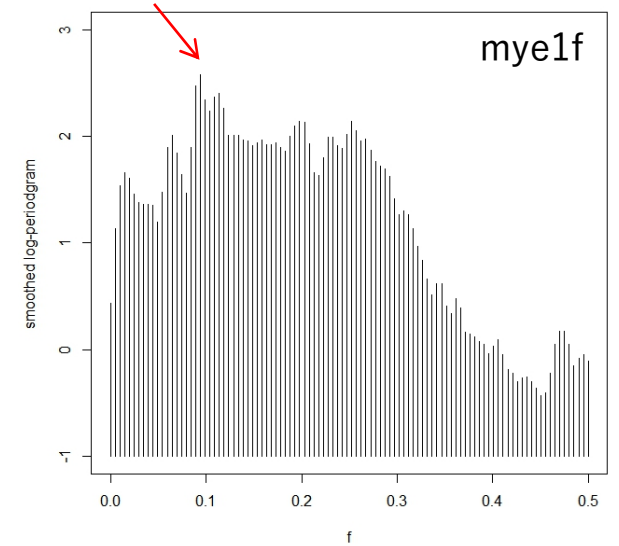
12か月周期, 6か月周期



12か月周期



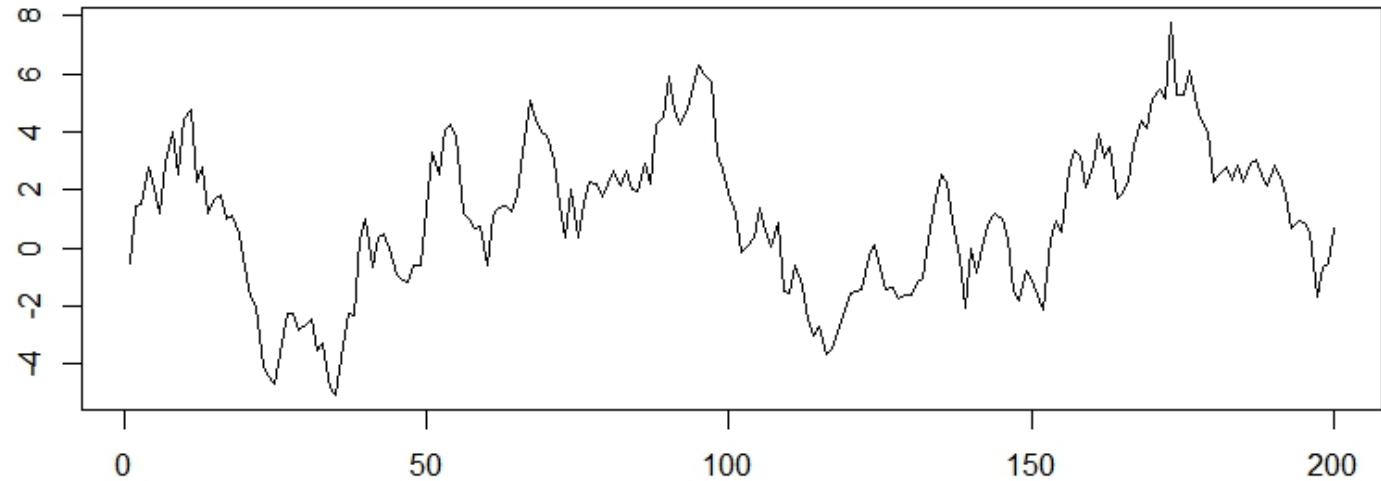
約10Hz



AR(1)過程 ($a=0.9$)

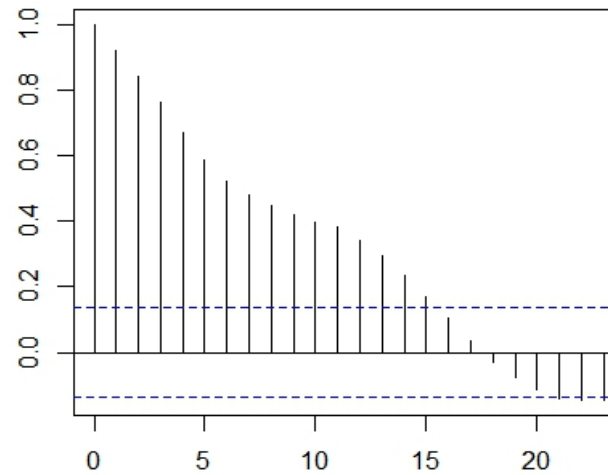
```
# AR(1)の生成
t <- 2:300
r <- as.ts(rnorm(300))
x <- rep(0,300)
for (i in t) x[i] = 0.9*x[i-1] + r[i]
plot(as.ts(x[101:300]))
#
# 自己相関関数
acf(x[101:300])
#
# ピリオドグラム
period(x[101:300])
```

時系列

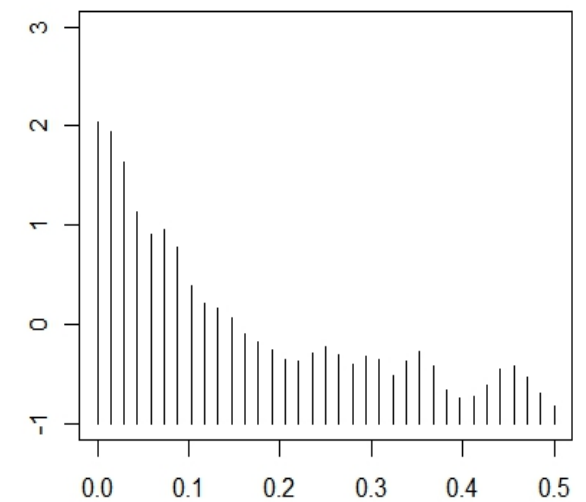


$$y_n = ay_{n-1} + w_n$$
$$a = 0.9$$

自己相関関数



ピリオドグラム

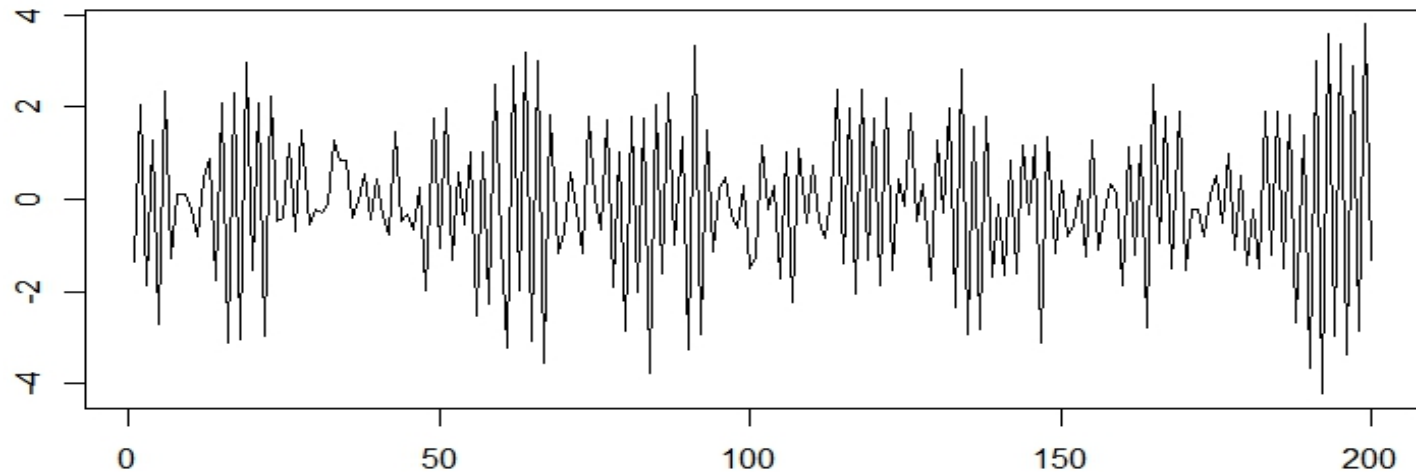


AR(1)過程 ($a = -0.9$)

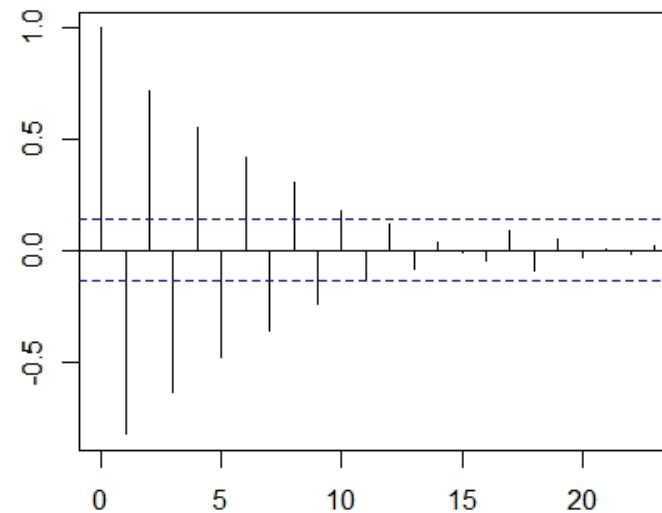
$$y_n = ay_{n-1} + w_n$$

$$a = -0.9$$

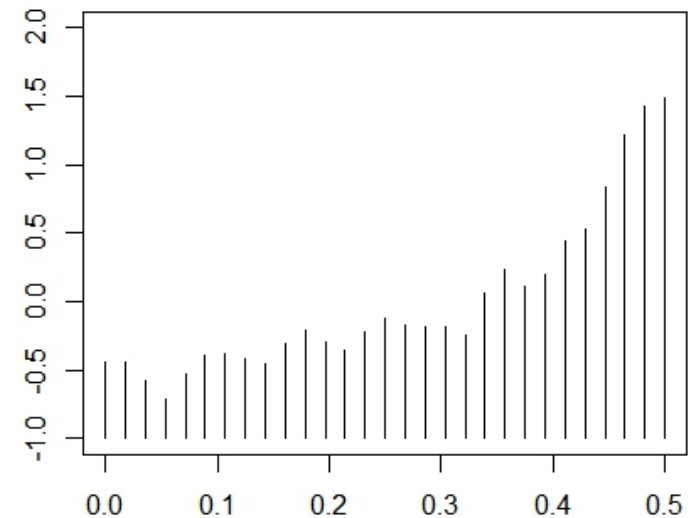
時系列



自己相関関数

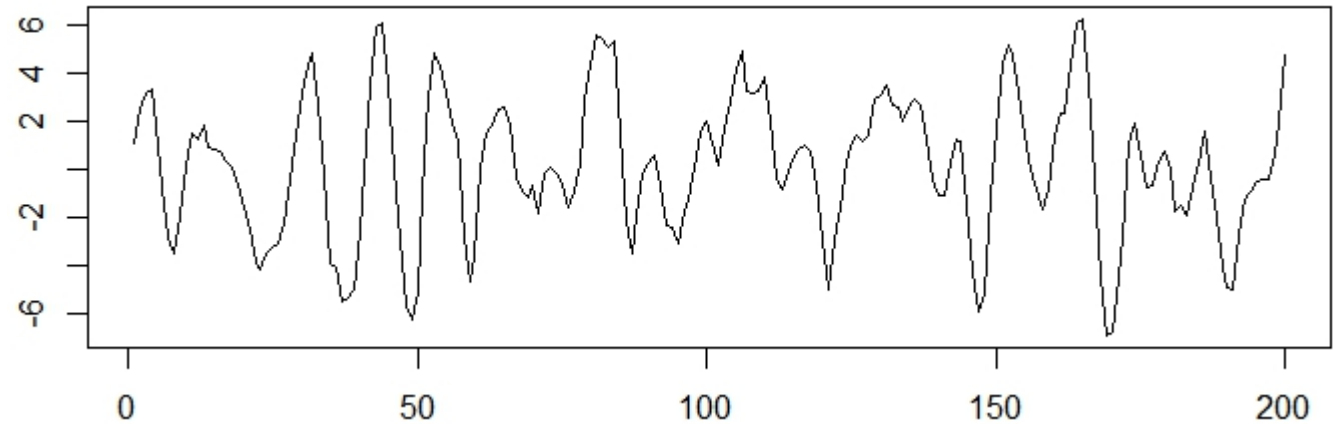


ピリオドグラム



AR(2)過程

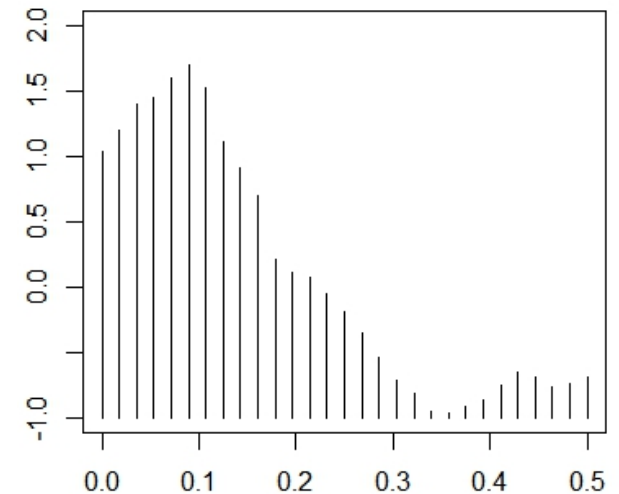
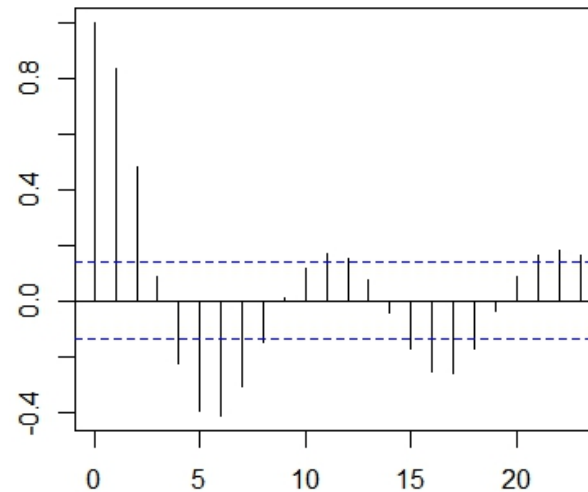
```
t <- 3:300
r <- as.ts(rnorm(300))
a1 <- 0.9*sqrt(3)
a2 <- -0.81
x <- rep(0,300)
for (i in t)
x[i] = a1*x[i-1] + a2*x[i-2] + r[i]
plot(as.ts(x[101:300]))
acf(x[101:300])
perid(x[101:300])
```



$$y_n = a_1 y_{n-1} + a_2 y_{n-2} + w_n$$

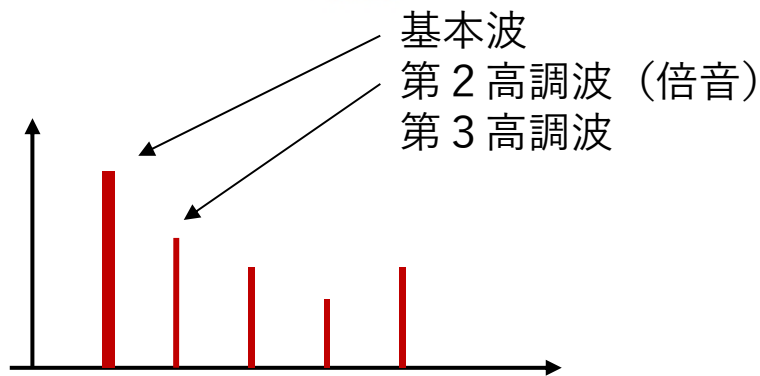
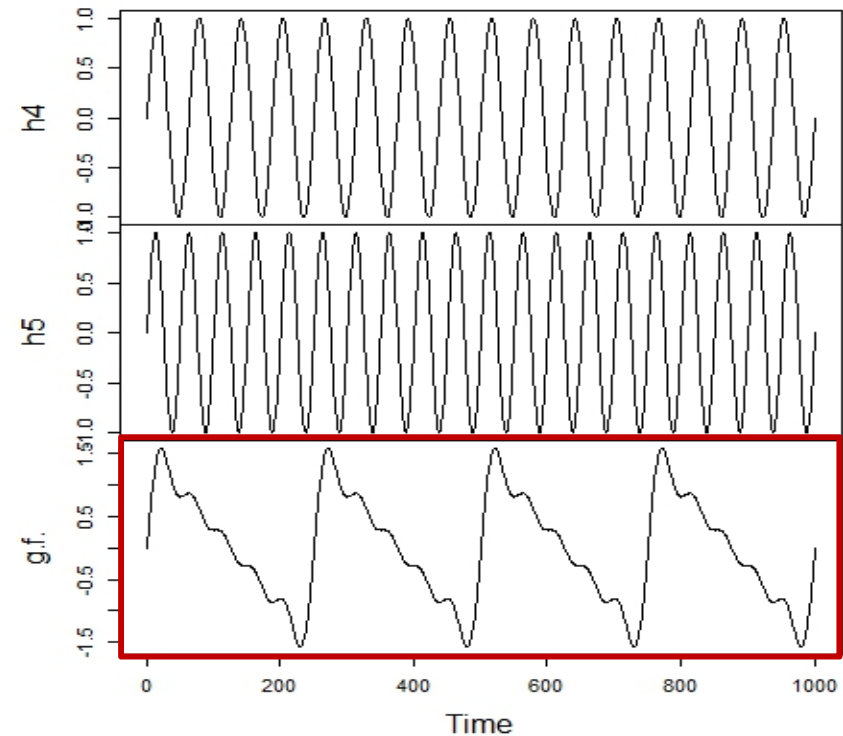
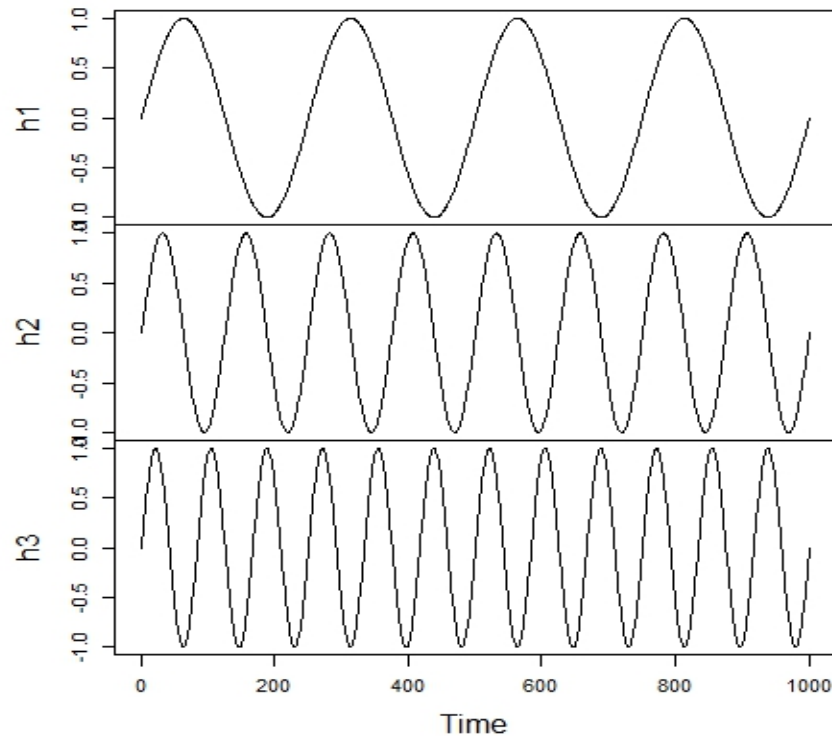
$$a_1 = 0.9\sqrt{3}$$

$$a_2 = -0.81$$



高調波の例

higher harmonics



Blfoodデータ, whardデータ

ピリオドグラムの性質

(1) 漸近的に不偏

$$\lim_{n \rightarrow \infty} E[\hat{p}(f)] = p(f)$$

(2) 一致推定量でない

$$\lim_{n \rightarrow \infty} \hat{p}(f) \neq p(f)$$

$$\frac{2\hat{p}(f_1)}{p(f_1)}, \dots, \frac{2\hat{p}(f_m)}{p(f_m)} \sim \chi_2^2$$

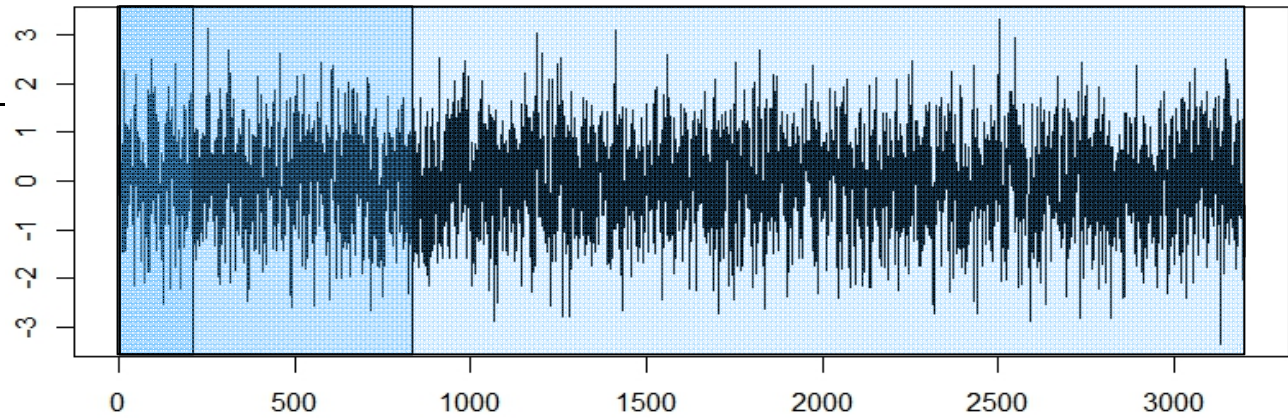
$$m = \left[\frac{N}{2} \right] - 1$$

$$\frac{\hat{p}(0)}{p(0)}, \frac{\hat{p}(\frac{1}{2})}{p(\frac{1}{2})} \sim \chi_1^2$$

データ数に関わらず同じ分布!

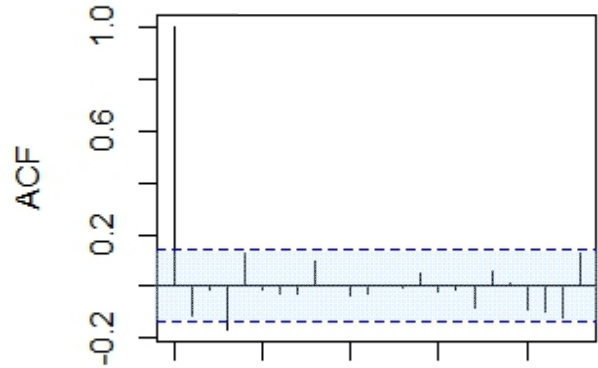
白色雑音

```
r <- rnorm(3200)
```

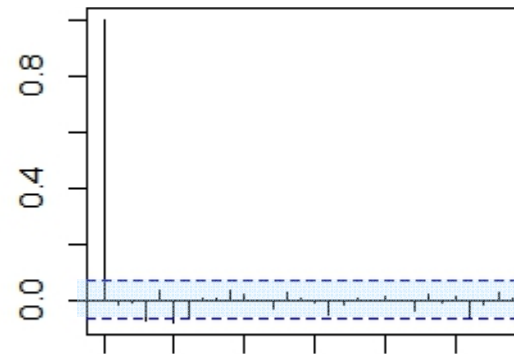


自己相関関数

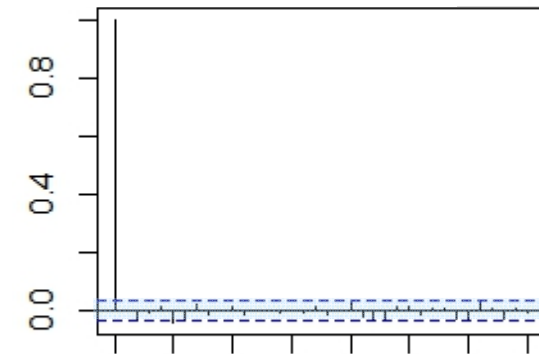
r[1:200]



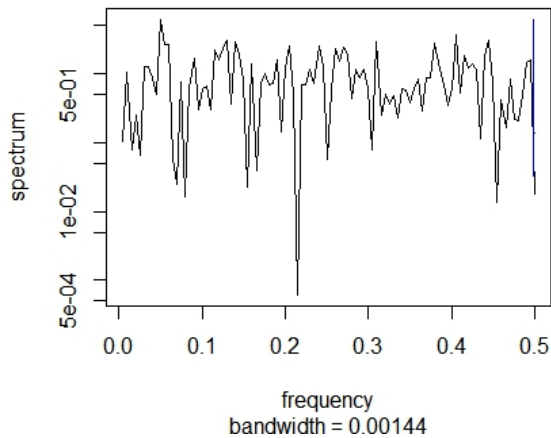
r[1:800]



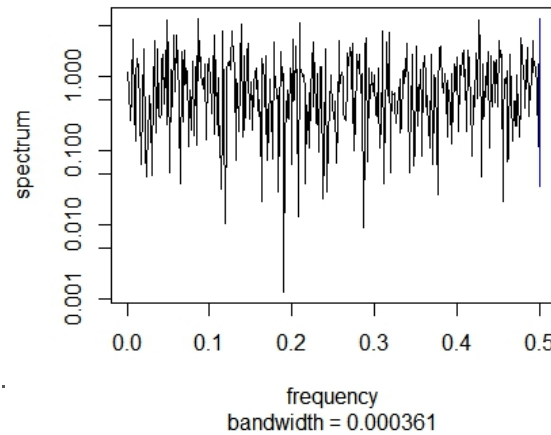
r[1:3200]



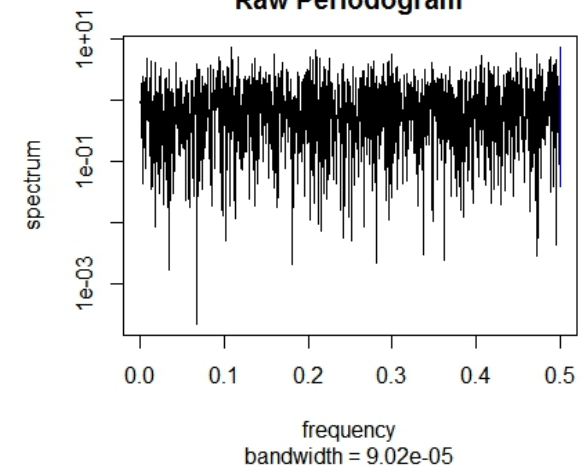
ピリオドグラム Series: r[1:200]
Raw Periodogram



Series: r[1:800]
Raw Periodogram

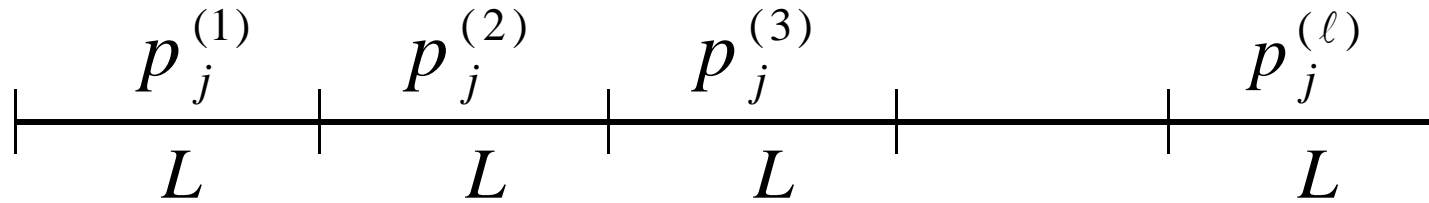


Series: r
Raw Periodogram



ピリオドグラム の平均

$$\ell = \frac{N}{L}$$



$$\hat{p}_j = \frac{1}{\ell} \sum_{k=1}^{\ell} p_j^{(k)}$$

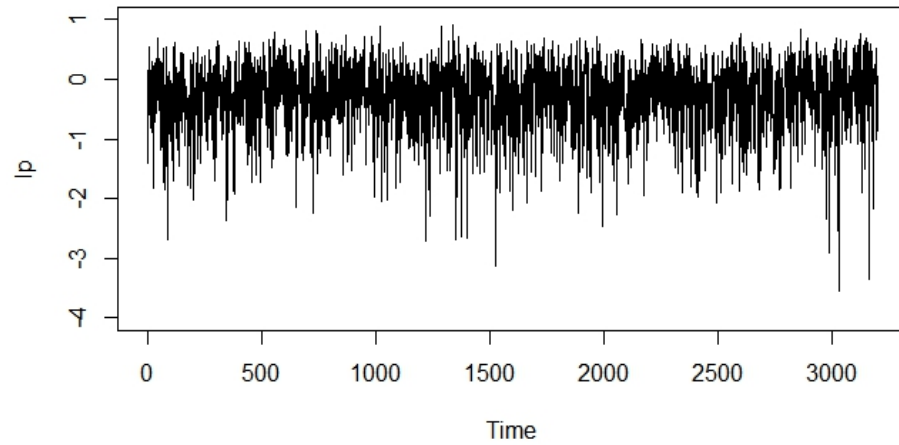
分散は $1/\ell$ に減少

実はラグ $L-1$ までのフーリエ変換でよい

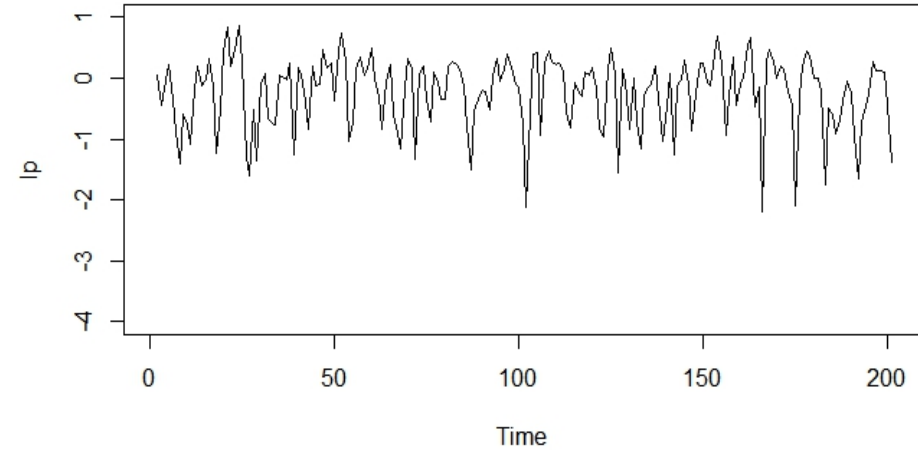
$$p_j = \hat{C}_0 + 2 \sum_{k=1}^{L-1} \hat{C}_k \cos(2\pi k f_j)$$

ピリオドグラムの平均による分散減少

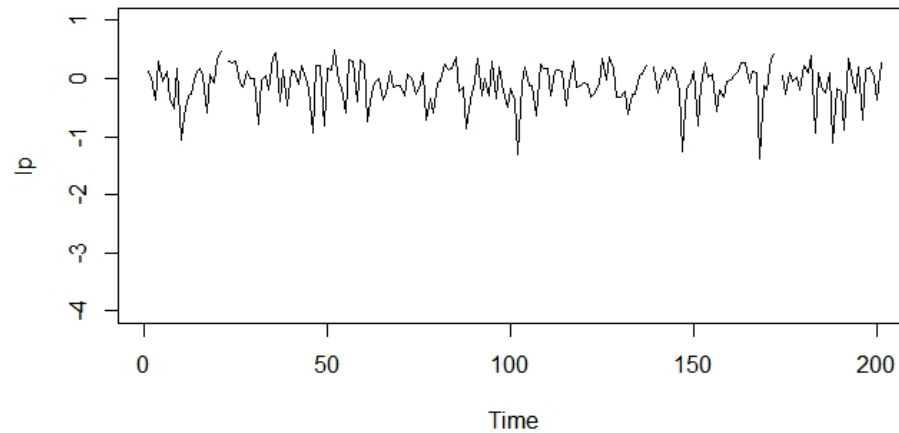
N=3200, lag=3200



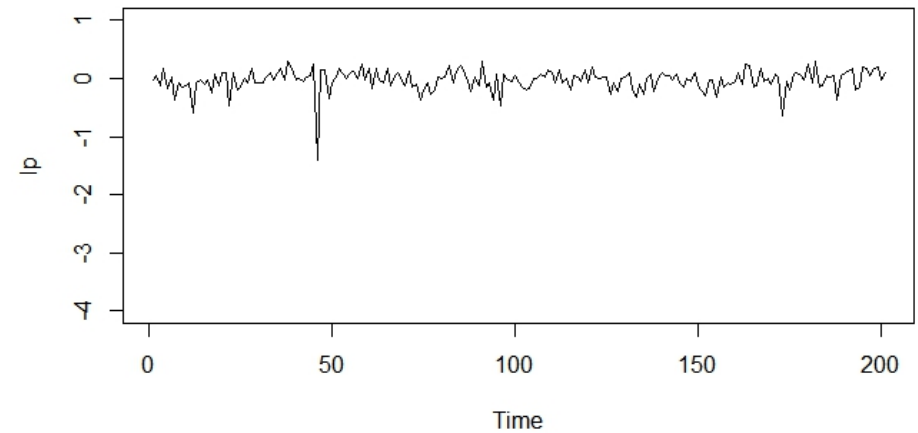
N=200, lag=200



N=800, lag=200



N=3200, lag=200



ピリオドグラムの平滑化

スペクトルは比較的滑らかという前提

$$\tilde{p}_j = \sum_{i=-m}^m W_i \hat{p}_{j-i}$$

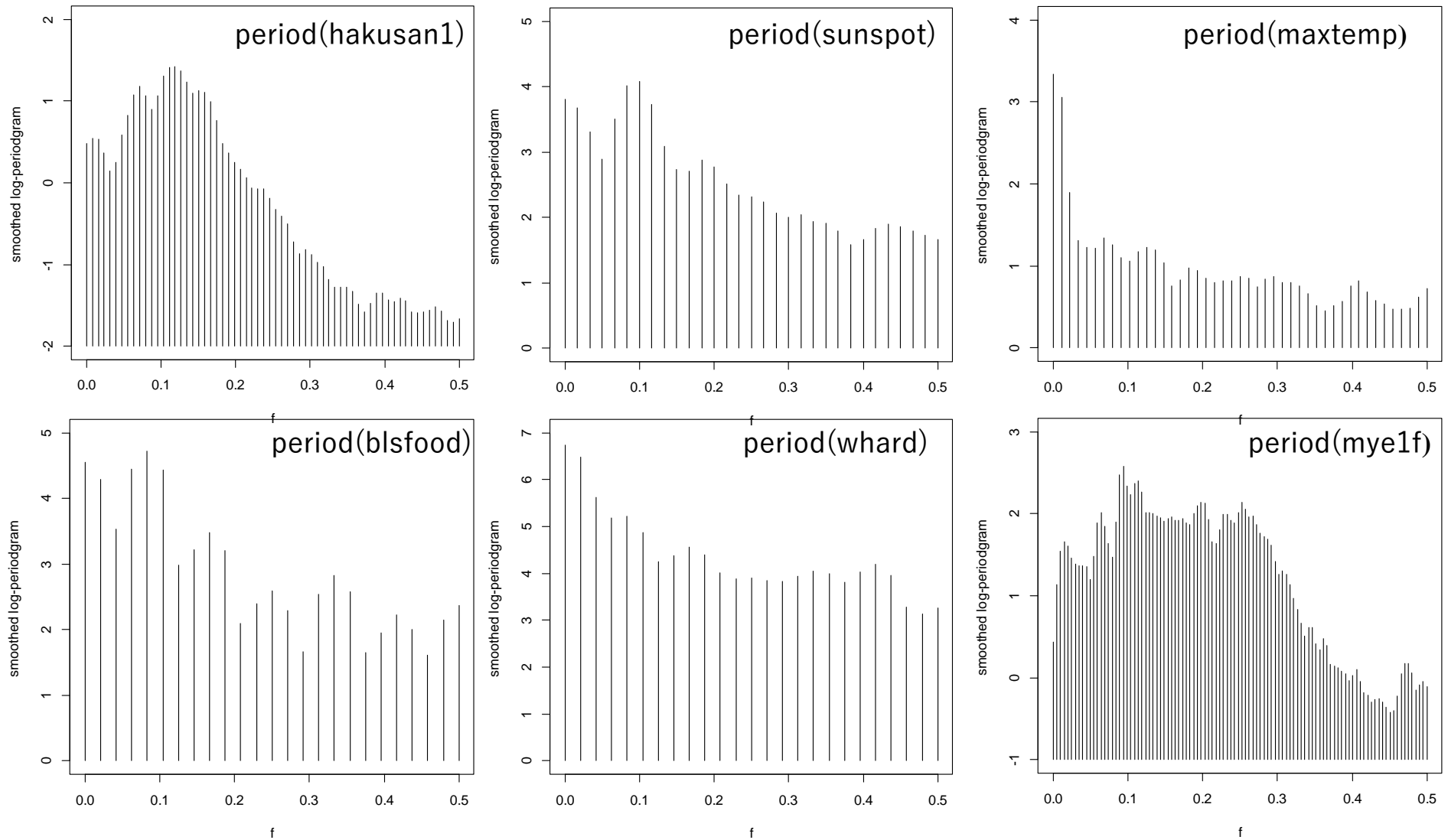
W_i : Window

Window	m	W_0	W_1
Hanning	1	0.50	0.25
Hamming	1	0.54	-0.23

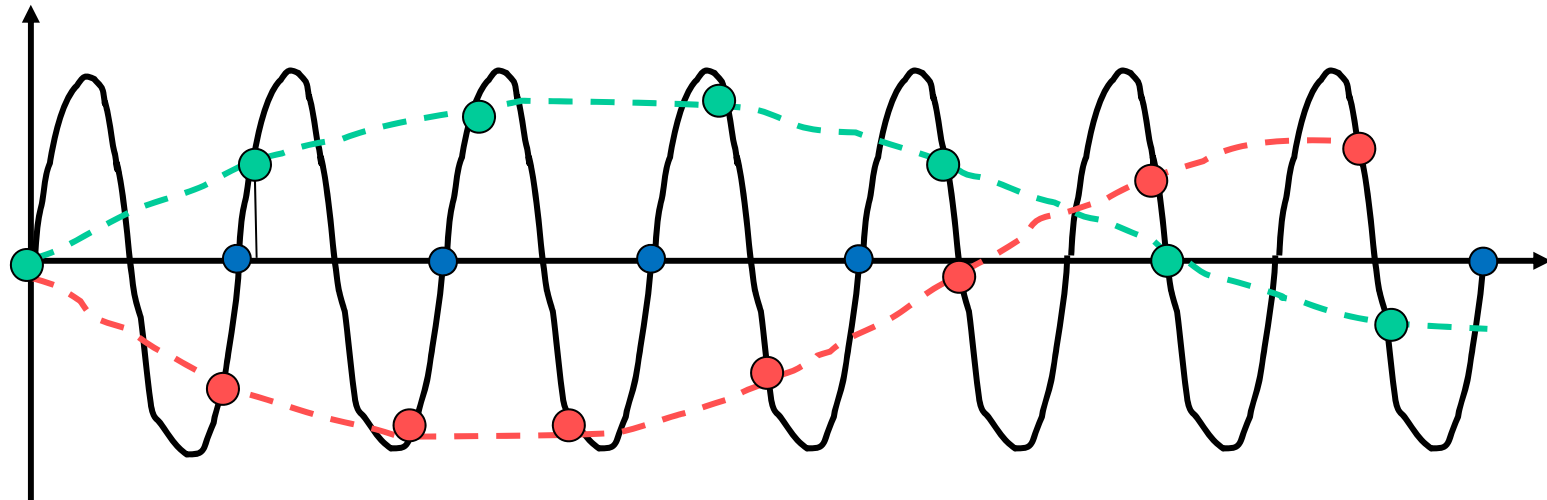
Blackman-Tukey 法

Windowは多数提案されている (Parzen, Bartlett, 赤池, ...)

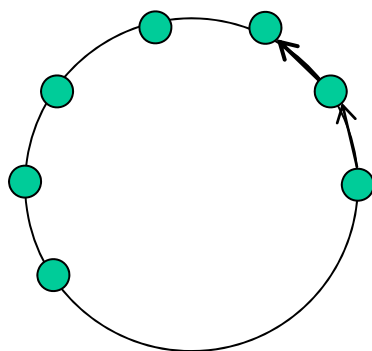
ピリオドグラムの平滑化



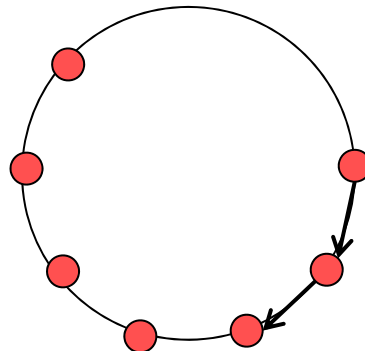
エイリアシング効果



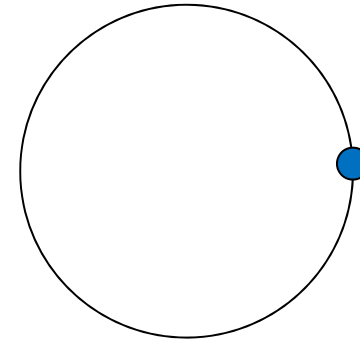
連続的な変動を離散間隔でサンプリングしたときの現象



ゆっくり回転



ゆっくり逆回転

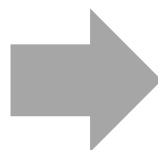
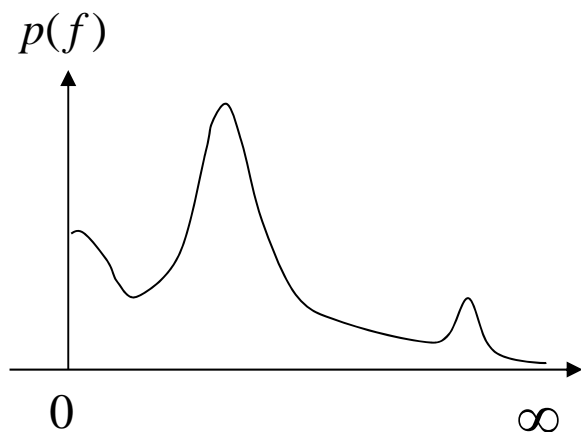


無回転

スペクトルのフォールディング

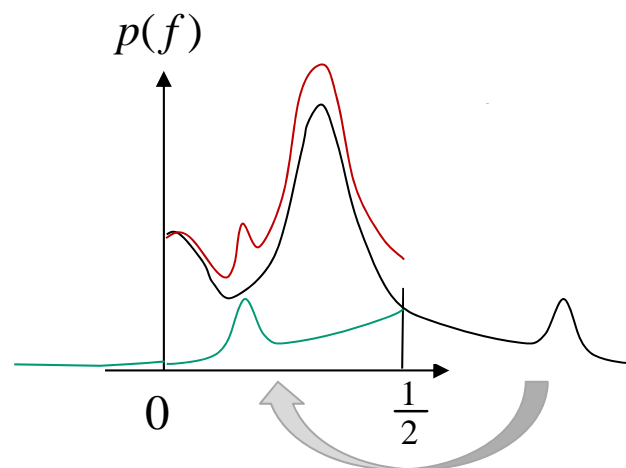
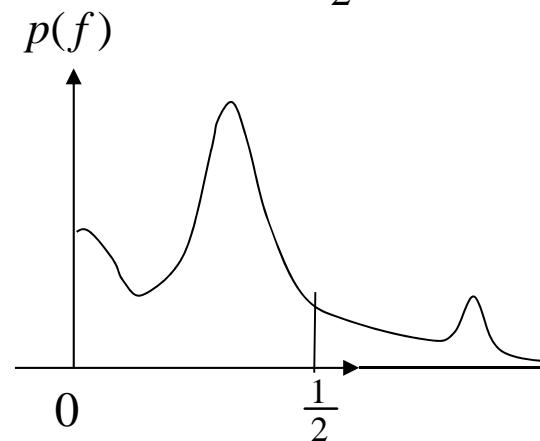
連続時間時系列のスペクトル

$$p(f), \quad -\infty < f < \infty$$



離散時間時系列のスペクトル

$$p(f), \quad -\frac{1}{2} < f < \frac{1}{2}$$



$p(f), -\infty < f < \infty$ を $0 \leq f < \frac{1}{2}$
に繰り返し折り返して加算する

FFT (Fast Fourier Transform)

(原理的には任意の素因数分解が利用できる)

$N = p^\ell$ のとき, 計算量 $N^2 \rightarrow Np\ell$

$$\frac{Np\ell}{N^2} = \frac{p\ell}{N}$$

$$N = 1024 = 2^{10} \quad \frac{2 \times 10}{1024} \sim \frac{1}{50}$$

$$N = 4096 = 2^{12} \quad \frac{2 \times 12}{4096} \sim \frac{1}{170}$$

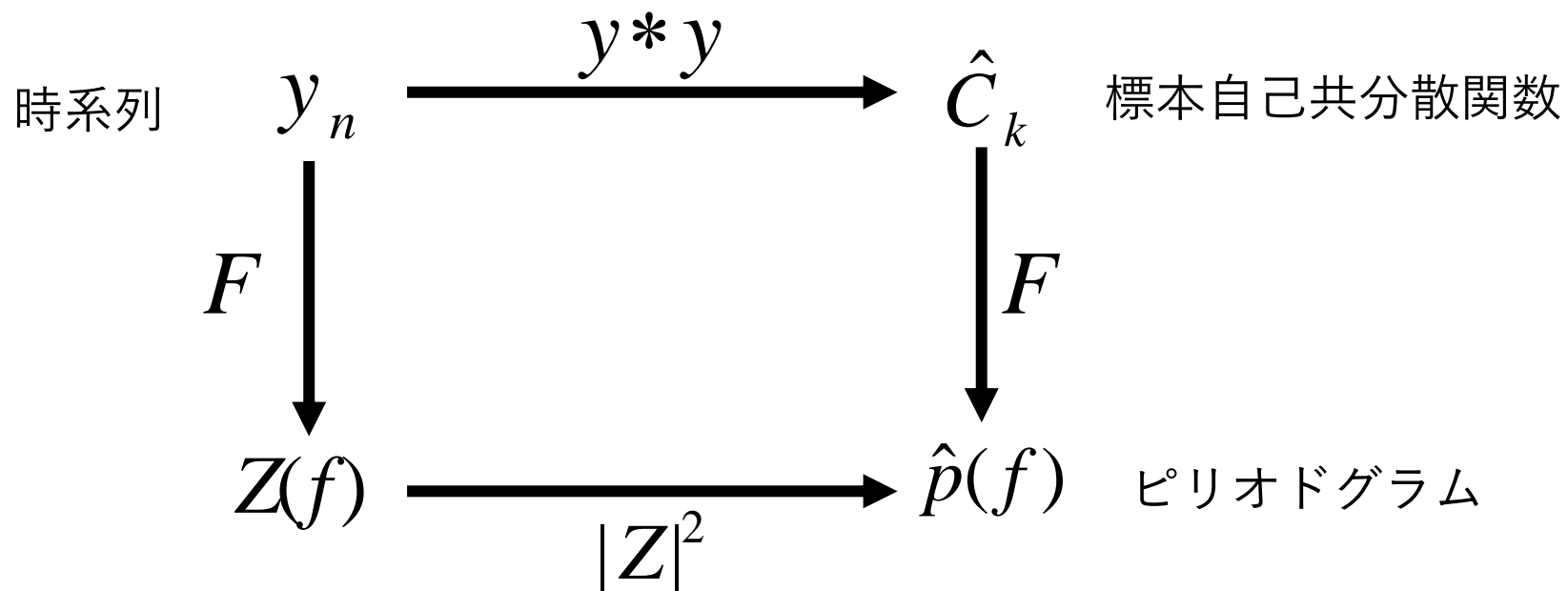
ただし \hat{C}_k の計算に $O(N^2)$ 必要

FFTによるピリオドグラムの計算

$$\begin{aligned} Z_j &= \sum_{n=1}^N y_n e^{-\frac{2\pi i(n-1)j}{N}} \\ &= \sum_{n=1}^N y_n \cos\left(\frac{2\pi(n-1)j}{N}\right) - i \sum_{n=1}^N y_n \sin\left(\frac{2\pi(n-1)j}{N}\right) \\ &= FC_j - iFS_j \end{aligned}$$

$$\begin{aligned} P_j &= \frac{|Z_j|^2}{N} = \frac{1}{N} \left| \sum_{n=1}^N y_n e^{-2\pi i(n-1)j/N} \right|^2 \\ &= \frac{1}{N} (FC_j^2 + FS_j^2) \end{aligned}$$

データ，共分散関数，ピリオドグラムの関係



ピリオドグラムとFFT

$$\hat{p}(f) = \hat{C}_0 + 2 \sum_{k=1}^{N-1} \hat{C}_k \cos(2\pi kf) \quad f = 0, \frac{1}{N}, \frac{2}{N}, \dots, \frac{1}{2}$$

$$p_j = \frac{1}{N} (FC_j^2 + FS_j^2)$$

N が 2^k でないとき

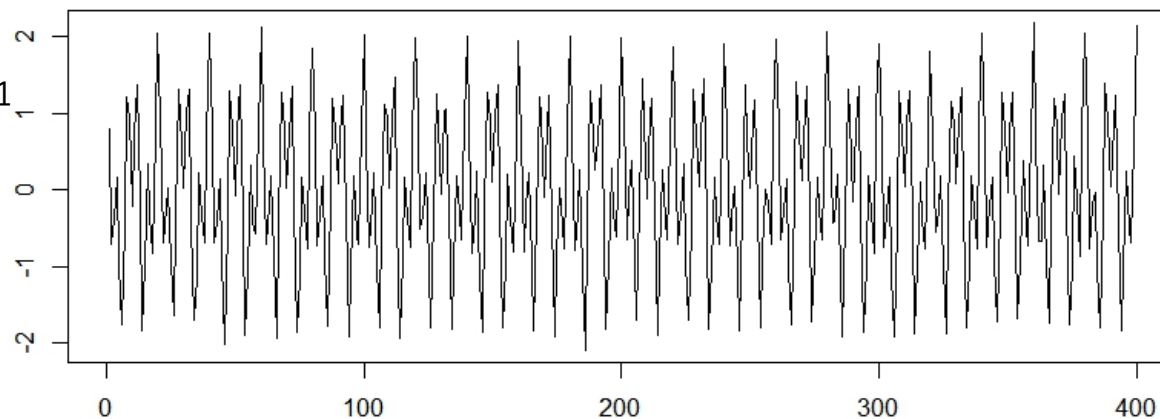
- $N = p_1^{k_1} \times \dots \times p_m^{k_m}$
- $\{y_1, \dots, y_N, 0, \dots, 0\}$ として 2^k の長さの時系列を作ってFFT
 - ・ $\hat{p}(f)$ と同じ値が得られる
 - ・ ただし, 周波数が $f = \frac{j}{N}$ に限定されることに注意

(例) 2つのcos関数+ノイズ

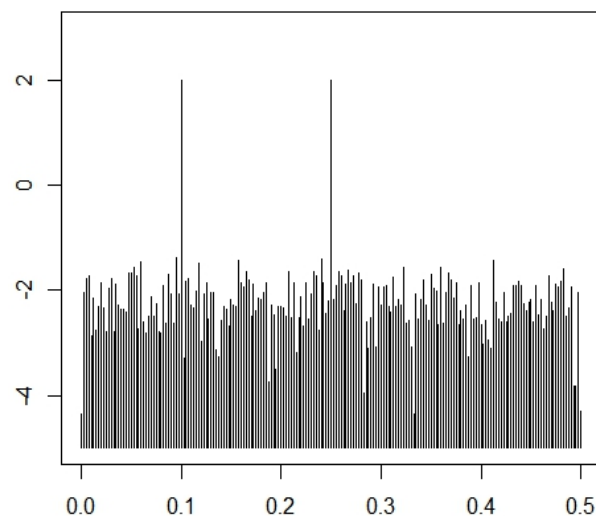
```
# simulation of 2 cosine function + noise
t <- 1:400
r <- rnorm(400)
y <- rep(0,400)
for (i in t) {
y[i] <- cos(2*pi*i/10) + cos(2*pi*i/4) + r[i]*0.1
}
y <- as.ts(y)
plot(y)
```

```
period(y>window=0)
```

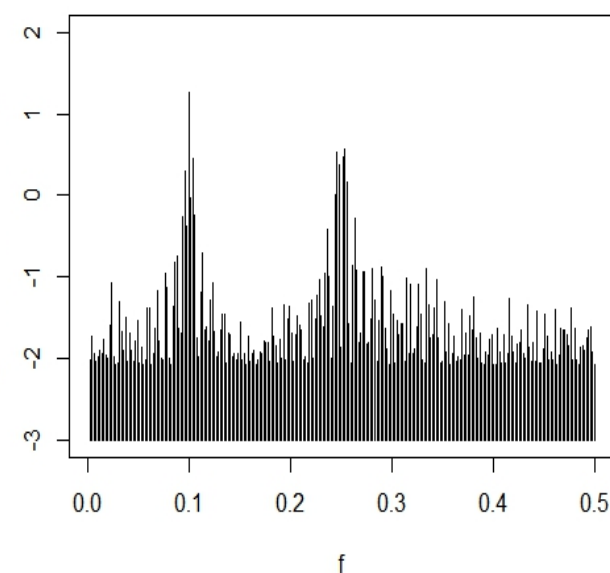
```
fftper(y>window=0)
```



ピリオドグラム

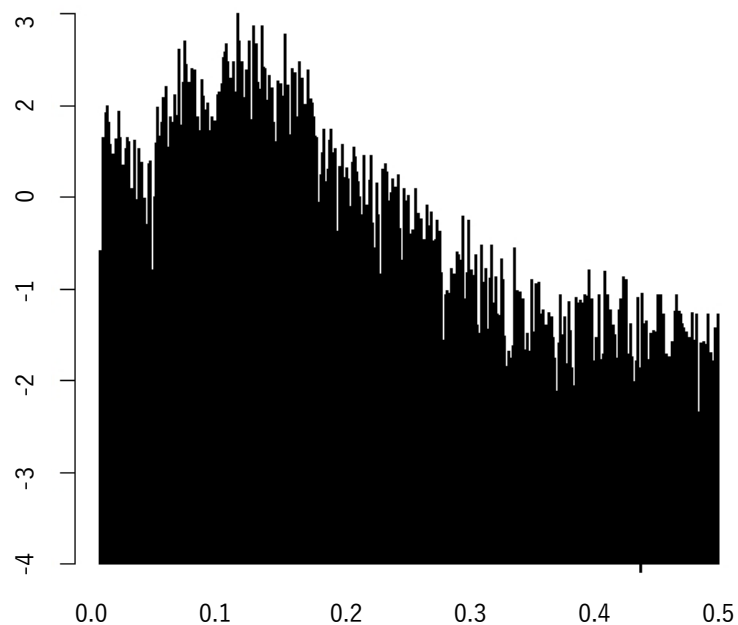


FFTピリオドグラム



FFTピリオドグラム (船舶方向角データ)

ピリオドグラム



FFTピリオドグラム

