定義

確率過程

確率変数 X_t が順序付けされてまとめられたものを**時系列データ**といい、$X = \{X_t \mid t \in \mathbb{Z}\}$ のように表す。

時系列 X の確率変数 X_i, X_k の共分散 $\mathrm{Cov}(X_i, X_k)$ を $|i-k|$ 次の**自己共分散**という。この $|i-k|$ のことをラグと呼ぶ。

${}^\forall X_i \in X$ に対して $E[X_i] = \mu$ となる値 μ が 1 つ存在し、${}^\forall X_i, X_{i+k} \in X$ に対して $Cov(X_i, X_{i+k}) = \sigma_k$ となる値 σ_k がラグ k についてのみ依存して定まる時、時系列 X は弱定常性を持つという。

ホワイトノイズ

ホワイトノイズとは、弱定常性を持つ時系列 X で次の性質をもつものをいう。

$$ {}^\forall X_i, X_k \in X\ (i \neq k)\ \text{に対し} \quad E[X_i] = 0,\ V[X_i] = \sigma^2,\ Cov(X_i, X_k) = 0 \tag{1} $$

性質

$$
\begin{aligned}
E[aX+b] =& aE[X] + b & (2)\\
E[X_1 + X_2] =& E[X_1] + E[X_2] & (3)\\
V[X] =& E[(X - E[X])^2] & (4)\\
V[X] =& E[X^2] - (E[X])^2 & (5)\\
Cov(X_1, X_2) =& E[(X_1 - E[X_1])(X_2 - E[X_2])] & (6)\\
V[X] =& Cov(X, X) & (7)\\
Cov(X_1, X_2) =& E[X_1 X_2] - E[X_1]E[X_2] & (8)\\
V[X_1 + X_2] =& V[X_1] + V[X_2] + 2Cov(X_1, X_2) & (9)\\
Cov(X_1 + X_2, X_3) =& Cov(X_1, X_3) + Cov(X_2, X_3) & (10)
\end{aligned}
$$

問題

1. $X = \{X_t; t \in \mathbb{Z}\}$ を実数値弱定常時系列で、平均関数が 0 であるものとし、X の自己共分散関数を γ とする。
   (a) ${}^\forall h \in \mathbb{Z}$ に対して、$\gamma(-h) = \gamma(h)$ が成り立つことを示せ。

   自己共分散関数 γ は $\gamma(h) = Cov(X_{i+h}, X_i)$ であるとすると $\gamma(-h) = Cov(X_{i+h}, X_i)$ となる。

共分散の定義から次の性質が得られる。

$$Cov(X_{i+h}, X_i) = E[(X_{i+h} - E[X_{i+h}])(X_i - E[X_i])] \tag{11}$$
$$= E[(X_i - E[X_i])(X_{i+h} - E[X_{i+h}])] = Cov(X_i, X_{i+h}) \tag{12}$$

よって、$\gamma(-h) = \gamma(h)$ である。

(b) $n \in \mathbb{N}$ とするとき、次の n 次正方行列 Γ_n の固有値が全て 0 以上であることを示せ。

$$\Gamma_n = \begin{pmatrix} \gamma(0) & \gamma(1) & \gamma(2) & \cdots & \gamma(n-1) \\ \gamma(1) & \gamma(0) & \gamma(1) & \cdots & \gamma(n-2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma(n-1) & \gamma(n-2) & \gamma(n-3) & \cdots & \gamma(0) \end{pmatrix} \tag{13}$$

……………………………………………………………………

時系列 X より連続した n 個の確率変数 $X_1, \cdots, X_n$ を取り出し分散共分散行列 Σ を作る。

$$\Sigma = \begin{pmatrix} Cov(X_1, X_1) & Cov(X_1, X_2) & \cdots & Cov(X_1, X_n) \\ Cov(X_2, X_1) & Cov(X_2, X_2) & \cdots & Cov(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_n, X_1) & Cov(X_n, X_2) & \cdots & Cov(X_n, X_n) \end{pmatrix} \tag{14}$$

取り出した確率変数は連続していれば全て同じ行列となり $\Gamma_n = \Sigma$ である。
分散共分散行列は半正定値行列であるので固有値は全て 0 以上である。

……………………………………………………………………

共分散は偏差の積の平均である。

$$Cov(X_i, X_k) = E[(X_i - E[X_i])(X_k - E[X_k])] \tag{15}$$
$$= \frac{1}{N} \sum_{j=1}^{N} (X_i^{(j)} - E[X_i])(X_k^{(j)} - E[X_k]) \tag{16}$$

(ただし、X_i は N 個のデータから成る確率変数で、$E[X_i] = \frac{1}{N} \sum_{j=1}^{N} X_I^{(j)}$ とする)

この為、共分散 $Cov(X_i, X_k)$ はベクトルの内積を用いて次のように書ける。

$$Cov(X_i, X_k) \tag{17}$$

$$=\frac{1}{N}\begin{pmatrix} X_i^{(1)} - E[X_i] & X_i^{(2)} - E[X_i] & \cdots & X_i^{(N)} - E[X_i] \end{pmatrix} \begin{pmatrix} X_k^{(1)} - E[X_k] \\ X_k^{(2)} - E[X_k] \\ \vdots \\ X_k^{(N)} - E[X_k] \end{pmatrix} \tag{18}$$

そこで $n \times N$ 行列 C を次のように作る。

$$C = \begin{pmatrix} X_1^{(1)} - E[X_1] & \cdots & X_1^{(N)} - E[X_1] \\ \vdots & \ddots & \vdots \\ X_n^{(1)} - E[X_n] & \cdots & X_n^{(N)} - E[X_n] \end{pmatrix} \tag{19}$$

これにより行列 C とその転置行列 ${}^tC$ を用いて分散共分散行列を次のように書ける。

$$\Sigma = \frac{1}{N} C\, {}^tC \tag{20}$$

任意のベクトル $\boldsymbol{x}$ を用いて ${}^t\boldsymbol{x}\Sigma\boldsymbol{x}$ を計算する。

$${}^t\boldsymbol{x}\Sigma\boldsymbol{x} = \frac{1}{N}{}^t\boldsymbol{x}C\,{}^tC\boldsymbol{x} = \frac{1}{N}{}^t({}^tC\boldsymbol{x}){}^tC\boldsymbol{x} \tag{21}$$

ここで $\boldsymbol{y} = {}^tC\boldsymbol{x}$ とすると上の式は次のようになる。

$${}^t\boldsymbol{x}\Sigma\boldsymbol{x} = \frac{1}{N}{}^t\boldsymbol{y}\boldsymbol{y} = \frac{1}{N}\sum_{j=1}^{N} y_j^2 \geq 0 \tag{22}$$

よって、分散共分散行列 Σ は半正定値行列である。

....

半正定値

対称行列 A が次を満たす時、半正定値という。

$$\text{任意のベクトル } \boldsymbol{x} \text{ に対して } {}^t\boldsymbol{x}A\boldsymbol{x} \geq 0 \tag{23}$$

半正定値の性質

次は同値な条件である。

- A は半正定値
- A の固有値は全て非負

---

2. 実数値確率変数 X, Y は次を満たすものとする。

$$E[X] = E[Y] = 0,\ E[X^2] < \infty,\ E[Y^2] < \infty \tag{24}$$

この時、次の不等式が成り立つことを示せ。

$$(Cov(X, Y))^2 \leq V[X]V[Y] \tag{25}$$

....

確率変数 XY の分散は次の式で求められる。

$$V[XY] = E[(XY)^2] - (E[XY])^2 \tag{26}$$

分散は定義から常に 0 以上の値を取るので、上の式から次が得られる。

$$E[(XY)^2] \geq (E[XY])^2 \tag{27}$$

条件 $(E[X] = E[Y] = 0)$ より $Cov(X, Y), V[X], V[Y]$ は次のように変形できる。

$$Cov(X, Y) = E[XY] - E[X]E[Y] = E[XY] \tag{28}$$
$$V[X] = E[X^2] - (E[X])^2 = E[X^2] \tag{29}$$
$$V[Y] = E[Y^2] - (E[Y])^2 = E[Y^2] \tag{30}$$

$$E[X^2Y^2] \geq (Cov(X, Y))^2 \tag{31}$$

X, Y の相関係数を ρ とすると $-1 \leq \rho \leq 1$ であり、$0 \leq \rho^2 \leq 1$ である。
ρ は次の式で求められる。

$$\rho = \frac{Cov(X, Y)}{\sqrt{V[X]}\sqrt{V[Y]}} \tag{32}$$

よって、

$$\rho^2 = \frac{(Cov(X, Y))^2}{V[X]V[Y]} \leq 1 \tag{33}$$

であり、分母を払うことで次が得られる。

$$(Cov(X, Y))^2 \leq V[X]V[Y] \tag{34}$$

---

3. $\sigma > 0$ を定数とし、弱定常時系列 $X = \{X_t; t \in \mathbb{Z}\}$ はホワイトノイズ $\mathrm{WN}(0, \sigma^2)$ であるとする。時系列 $S = \{S_t; t \in \mathbb{N} \cup \{0\}\}$ を次のように定義する。

$$S_t = \begin{cases} 0, & (t = 0) \\ \sum_{j=1}^{t} X_j, & (t \in \mathbb{N}) \end{cases} \tag{35}$$

……………………………………………………………………………

ホワイトノイズであるということなので、各 $X_i \in X$ に対して $E[X_i]=0,\ V[X_i]=\sigma^2$ であり、$X_i, X_k \in X\ (i \neq k)$ に対して $Cov(X_i, X_k)=0$ である。

(a) 各 $t \in \mathbb{N}$ に対して、S_t の平均を求めよ。

……………………………………………………………………………

$t=0$ の時、$E[S_0]=E[0]=0$ である。
$t>0$ の場合を考える。
次のように平均を計算できる。

$$E[S_t]=E\left[\sum_{j=1}^{t} X_j\right]=\sum_{j=1}^{t} E[X_j]=\sum_{j=1}^{t} 0=0 \tag{36}$$

よって、${}^\forall S_t \in S$ に対して $E[S_t]=0$ である。

(b) 各 $t \in \mathbb{N}$ に対して、S_t の分散を求めよ。

……………………………………………………………………………

$t=0$ の時、$V[S_0]=V[0]=0$ である。
$t>0$ の場合を考える。
S_t の分散 $V[S_t]$ を計算する。

$$V[S_t]=V[S_{t-1}+X_t]=V[S_{t-1}]+V[X_t]+2Cov(S_{t-1},X_t) \tag{37}$$

ここで、$Cov(S_{t-1},X_t)$ は

$$Cov(S_{t-1},X_t)=Cov(X_1,X_t)+Cov(X_2,X_t)+\cdots+Cov(X_{t-1},X_t) \tag{38}$$

と分解できるが、ホワイトノイズであるので共分散は全て 0 となる。
これより、式 (37) は $V[S_t]=V[S_{t-1}]+V[X_t]$ となる。これを繰り返し行うと次の式が得られる。

$$V[S_t]=\sum_{j=1}^{t} V[X_j] \tag{39}$$

ホワイトノイズであるので全ての分散は $V[X_j]=\sigma^2$ である。これにより分散は $V[S_t]=t\sigma^2$ となる。

(c) 各 $t,k \in \mathbb{N}$ に対して、S_{t+k} と S_t の共分散を求めよ。

……………………………………………………………………………

$k=0$ の時、$Cov(S_t,S_t)=V[S_t]=t\sigma^2$ である。

$k>0$ について考える。

$k=1$ の場合次のような変形が行える。

$$Cov(S_{t+1},S_t)=Cov(S_t+X_{t+1},S_t)=Cov(S_t,S_t)+Cov(X_{t+1},S_t) \tag{40}$$

$Cov(X_{t+1},S_t)$ は

$$Cov(X_{t+1},S_t)=Cov(X_{t+1},X_1)+\cdots+Cov(X_{t+1},X_t)=0 \tag{41}$$

となる。その為、次のように共分散が求まる。

$$Cov(S_{t+1},S_t)=Cov(S_t,S_t)=t\sigma^2 \tag{42}$$

$k>1$ の場合、次のように S_{t+k} を S_{t+k-1} にすることができる。

$$\begin{aligned}Cov(S_{t+k},S_t)=&Cov(S_{t+k-1}+X_{t+k},S_t) &(43)\\ =&Cov(S_{t+k-1},S_t)+Cov(X_{t+k},S_t) &(44)\\ =&Cov(S_{t+k-1},S_t) &(45)\end{aligned}$$

よって、S_{t+k} と S_t の共分散は $Cov(S_{t+k},S_t)=t\sigma^2$ となる。

(d) 時系列 $\{\nabla S_t; t\in\mathbb{N}\cup\{0\}\}$ が弱定常性を持つかどうか説明せよ。

. . . .

---

4. p を 0 以上の整数とし、時系列 $X=\{X_t; t\in\mathbb{Z}\}$ のトレンド成分が次のように与えられているとする。

$$m(t)=\sum_{r=0}^{p} b_r t^r \tag{46}$$

各 $a_{-2},\ldots,a_2$ を次のようにする。

$$a_0=\frac{1}{4},\ a_{-1}=a_1=\frac{1}{4},\ a_{-2}=a_2=\frac{1}{8} \tag{47}$$

この時、次の式が成り立つような p を求めよ。

$$\sum_{k=-2}^{2} a_k m(t+k)=m(t) \tag{48}$$

. . . .

$\sum$ 記号を展開し分数をまとめると次のような形になる。

$$\sum_{k=-2}^{2} a_k m(t+k) = \frac{1}{8}(m(t-2)+2m(t-1)+2m(t)+2m(t+1)+m(t+2)) \tag{49}$$

これが $m(t)$ と等しいため、次のように分母を払い移行し式を整理する。

$$\begin{aligned}
m(t) =& \frac{1}{8}(m(t-2)+2m(t-1)+2m(t)+2m(t+1)+m(t+2)) && (50)\\
0 =& m(t-2)+2m(t-1)-6m(t)+2m(t+1)+m(t+2) && (51)\\
=& \sum_{r=0}^{p} b_r(t-2)^r + 2\sum_{r=0}^{p} b_r(t-1)^r && (52)\\
& -6\sum_{r=0}^{p} b_r t^r + 2\sum_{r=0}^{p} b_r(t+1)^r + \sum_{r=0}^{p} b_r(t+2)^r && (53)\\
=& \sum_{r=0}^{p} b_r((t-2)^r + 2(t-1)^r - 6t^r + 2(t+1)^r + (t+2)^r) && (54)
\end{aligned}$$

b_r が定まっていないためこれの係数 $((t-2)^r+2(t-1)^r-6t^r+2(t+1)^r+(t+2)^r)$ が 0 となるような r を求める。

$$\begin{aligned}
(t-2)^0+2(t-1)^0-6t^0+2(t+1)^0+(t+2)^0 =& 0 && (55)\\
(t-2)^1+2(t-1)^1-6t^1+2(t+1)^1+(t+2)^1 =& 0 && (56)\\
(t-2)^2+2(t-1)^2-6t^2+2(t+1)^2+(t+2)^2 =& 12 && (57)\\
(t-2)^3+2(t-1)^3-6t^3+2(t+1)^3+(t+2)^3 =& 36t && (58)\\
(t-2)^4+2(t-1)^4-6t^4+2(t+1)^4+(t+2)^4 =& 72t^2+36 && (59)
\end{aligned}$$

これにより $p=0,1$ の時 $\sum_{k=-2}^{2} a_k m(t+k) - m(t) = 0$ であり、$p=2$ のとき、$\sum_{k=-2}^{2} a_k m(t+k) - m(t) = 12b_2$ である。
$p>2$ の時 $\sum_{k=-2}^{2} a_k m(t+k) - m(t)$ は t の多項式となり 0 にならない。
よって、$\sum_{k=-2}^{2} a_k m(t+k) = m(t)$ を満たすのは $p=0,\ p=1$ の時となる。

---

sagemath 用コード

https://sagecell.sagemath.org/

```
x = var('x')   # 変数x の定義
t = var('t')   # 変数t の定義

```

```
# 変数x,t の多項式を定義
pol(x,t) = (t-2)^x + 2*(t-1)^x - 6*t^x + 2*(t+1)^x + (t+2)^x

# 変数x を0から増やしながら式を展開し出力
for i in range(7):
    print("r=", i, " ",expand(pol(i,t)))
```