信号理論

- No.11 不規則信号論 -

渡辺 裕

Signal Theory

- No.11 Stochastic Signal Theory -

Hiroshi Watanabe

確率時系列

- 時系列$\{x_n\}$が決定論的ではなく 確率的な場合，その特性は平均と自己相関関数で規定される

- 確率平均(期待値)

$$\bar{x}_n = E[x_n]$$

- 自己相関関数

$$r(n,m) = E[(x_n - \bar{x}_n)(x_{n+m} - \bar{x}_{n+m})]$$

Stochastic Series

- If series $\{x_n\}$ is not deterministic but stochastic, its characteristics are given by expectation and auto-correlation

- Stochastic mean (expectation)

$$\bar{x}_n = E[x_n]$$

- Auto-correlation function

$$r(n,m) = E[(x_n - \bar{x}_n)(x_{n+m} - \bar{x}_{n+m})]$$

定常過程

- $\{x_n\}$の平均と自己相関関数が時間nに依存しない場合を定常過程という．このとき平均値を0として一般性を失わない．
- 確率平均(期待値)

$$E[x_n] = 0$$

- 自己相関関数

$$r_x(m) = E[x_n x_{n+m}]$$

Stationary Process

- When expectation and auto-correlation function of $\{x_n\}$ do not depend on time n, it is called “stationary process.” Expectation can be set to 0 without loss of generality.

- Stochastic mean (Expectation)

$$E[x_n] = 0$$

- Auto-correlation function

$$r_x(m) = E[x_n x_{n+m}]$$

自己相関関数の性質 (1)

- 確率時系列$\{x_n\}$の自己相関関数は以下の性質を持つ.

$$r_x(m) = r_x(-m)$$

$$\left| r_x(m) \right| \le r_x(0)$$

- 証明 (コーシー-シュワルツの不等式)

$$E\left[x_n x_{n-m}\right] = E\left[x_{n+m} x_{n-m+m}\right] = E\left[x_{n+m} x_n\right] = E\left[x_n x_{n+m}\right]$$

$$\left| E\left[x_n x_{n+m}\right] \right| \le E\left[\left| x_n x_{n+m} \right|\right] \le \left(E\left[x_n{}^2\right]\right)^{1/2} \left(E\left[x_{n+m}{}^2\right]\right)^{1/2} = E\left[x_n{}^2\right]$$

Characteristics of Auto-correlation function (1)

- Auto-correlation function of stochastic series $\{x_n\}$ has the following relation.

$$r_x(m) = r_x(-m)$$
$$|r_x(m)| \le r_x(0)$$

- Proof (By Cauchy–Schwarz inequality)

$$E[x_n x_{n-m}] = E[x_{n+m} x_{n-m+m}] = E[x_{n+m} x_n] = E[x_n x_{n+m}]$$

$$|E[x_n x_{n+m}]| \le E[|x_n x_{n+m}|] \le \left(E[x_n{}^2]\right)^{1/2} \left(E[x_{n+m}{}^2]\right)^{1/2} = E[x_n{}^2]$$

自己相関関数の性質 (2)

- 二つの確率時系列$\{x_n\}$ $\{y_n\}$に対しては相互相関関数$c(m)$が定義される.

$$c_{xy}(m) \equiv E\left[x_n y_{n+m}\right]$$

- このとき以下の関係が成り立つ

$$c_{yx}(m) = c_{xy}(-m) \neq c_{xy}(m)$$

$$\left|c_{xy}(m)\right| \leq \sqrt{r_x(0) r_y(0)}$$

Characteristics of Auto-correlation function (2)

- Cross-correlation function $c(m)$ can be defined by two stochastic process $\{x_n\}$ $\{y_n\}$.

$$c_{xy}(m) \equiv E\left[x_n y_{n+m}\right]$$

- Here, two relations are hold.

$$c_{yx}(m) = c_{xy}(-m) \neq c_{xy}(m)$$

$$\left|c_{xy}(m)\right| \leq \sqrt{r_x(0) r_y(0)}$$

自己相関関数の性質 (3)

- 自己相関関数のフーリエ変換は時系列$\{x_n\}$のパワースペクトルに相当する.

$$P_x(e^{j\omega}) \equiv \sum_{n=-\infty}^{+\infty} r_x(n) e^{-jn\omega}$$

- 証明

$$P_x(e^{j\omega}) = \lim_{N\to\infty} \frac{1}{N+1} E\left[\left|\sum_{n=0}^{N} x(n)e^{-jn\omega}\right|^2\right]$$

$$= \lim_{N\to\infty} \frac{1}{N+1} E\left[\sum_{n=0}^{N} x(n)e^{-jn\omega} \sum_{m=0}^{N} x(m)e^{jm\omega}\right]$$

Characteristics of Auto-correlation function (3)

- Fourier transform of auto-correlation function is power spectrum of time series $\{x_n\}$

$$P_x(e^{j\omega}) \equiv \sum_{n=-\infty}^{+\infty} r_x(n) e^{-jn\omega}$$

- Proof

$$P_x(e^{j\omega}) = \lim_{N\to\infty} \frac{1}{N+1} E\left[\left| \sum_{n=0}^{N} x(n) e^{-jn\omega} \right|^2 \right]$$

$$= \lim_{N\to\infty} \frac{1}{N+1} E\left[\sum_{n=0}^{N} x(n) e^{-jn\omega} \sum_{m=0}^{N} x(m) e^{jm\omega} \right]$$

自己相関関数の性質 (4)

- 証明 (続き)

$$
\begin{aligned}
\frac{1}{N+1}E\left[\sum_{n=0}^{N}x(n)e^{-jn\omega}\sum_{m=0}^{N}x(m)e^{jm\omega}\right] &= \frac{1}{N+1}\sum_{n=0}^{N}\sum_{m=0}^{N}E\left[x(n)x(m)\right]e^{-j(n-m)\omega} \\
&= \frac{1}{N+1}\sum_{n=0}^{N}\sum_{m=0}^{N}r_x(n-m)e^{-j(n-m)\omega} \\
&= \frac{1}{N+1}\left(\sum_{k=0}^{N}\sum_{n=0}^{N-k}r_x(k)e^{-jk\omega}+\sum_{k=-1}^{-N}\sum_{m=0}^{N+k}r_x(k)e^{-jk\omega}\right) \\
&= \frac{1}{N+1}\left(\sum_{n=0}^{N}\sum_{k=0}^{N-n}r_x(k)e^{-jk\omega}+\sum_{m=0}^{N-1}\sum_{k=-1}^{-N+m}r_x(k)e^{-jk\omega}\right) \\
&= \frac{1}{N+1}\sum_{n=0}^{N}\sum_{k=-N+n}^{N-n}r_x(k)e^{-jk\omega}
\end{aligned}
$$

Characteristics of Auto-correlation function (4)

- Proof (Contd.)

$$
\begin{aligned}
\frac{1}{N+1} E\left[\sum_{n=0}^{N} x(n)e^{-jn\omega} \sum_{m=0}^{N} x(m)e^{jm\omega}\right] &= \frac{1}{N+1}\sum_{n=0}^{N}\sum_{m=0}^{N} E[x(n)x(m)]e^{-j(n-m)\omega} \\
&= \frac{1}{N+1}\sum_{n=0}^{N}\sum_{m=0}^{N} r_x(n-m)e^{-j(n-m)\omega} \\
&= \frac{1}{N+1}\left(\sum_{k=0}^{N}\sum_{n=0}^{N-k} r_x(k)e^{-jk\omega} + \sum_{k=-1}^{-N}\sum_{m=0}^{N+k} r_x(k)e^{-jk\omega}\right) \\
&= \frac{1}{N+1}\left(\sum_{n=0}^{N}\sum_{k=0}^{N-n} r_x(k)e^{-jk\omega} + \sum_{m=0}^{N-1}\sum_{k=-1}^{-N+m} r_x(k)e^{-jk\omega}\right) \\
&= \frac{1}{N+1}\sum_{n=0}^{N}\sum_{k=-N+n}^{N-n} r_x(k)e^{-jk\omega}
\end{aligned}
$$

自己相関関数の性質 (5)

- 証明 (続き2)

$$\lim_{N\to\infty}\frac{1}{N+1}\sum_{n=0}^{N}\sum_{k=-N+n}^{N-n} r_x(k)e^{-jk\omega} = \sum_{k=-\infty}^{\infty} r_x(k)e^{-jk\omega}$$

Characteristics of Auto-correlation function (5)

- Proof (Contd.2)

$$\lim_{N\to\infty} \frac{1}{N+1}\sum_{n=0}^{N}\sum_{k=-N+n}^{N-n} r_x(k)e^{-jk\omega} = \sum_{k=-\infty}^{\infty} r_x(k)e^{-jk\omega}$$

エルゴード性

- 集合平均(確率平均, 期待値)と時間平均が一致

$$E[x_m] \cong \frac{1}{N+1}\sum_{n=0}^{N} x(n)$$

$$E[x_n x_{n+m}] \cong \frac{1}{N+1}\sum_{n=0}^{N} x(n)x(n+m)$$

Ergodicity

- Stochastic mean (Expectation) equals average

$$E[x_m] \cong \frac{1}{N+1}\sum_{n=0}^{N} x(n)$$

$$E[x_n x_{n+m}] \cong \frac{1}{N+1}\sum_{n=0}^{N} x(n)x(n+m)$$

線形予測 (1)

- 自己回帰モデルによるシステムの記述
  - u_k:時刻kの入力で白色ガウス雑音, y_k:時刻kの出力, $a_i(i=1, \ldots, n)$: 線形システムのARパラメータ

$$y_k = \sum_{i=1}^{n} a_i y_{k-i} + u_k$$

$$= \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} y_{k-1} \\ y_{k-2} \\ \vdots \\ y_{k-n} \end{bmatrix} + u_k$$

$$= \boldsymbol{a}^t \boldsymbol{y} + u_k$$

Linear Prediction (1)

- System description by autoregressive model
  - u_k: input white gaussian at time k, y_k: output at time k, $a_i(i=1, ..., n)$: AR parameter of linear system

$$y_k = \sum_{i=1}^{n} a_i y_{k-i} + u_k$$

$$= \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} y_{k-1} \\ y_{k-2} \\ \vdots \\ y_{k-n} \end{bmatrix} + u_k$$

$$= \boldsymbol{a}^t \boldsymbol{y} + u_k$$

線形予測 (2)

- パラメータ予測モデル
  - $\hat{y}_k$: 予測値, $\hat{a}_i$: 予測パラメータ(予測AR係数)

$$\hat{y}_k = \sum_{i=1}^{n} \hat{a}_i y_{k-i}$$

$$= \begin{bmatrix} \hat{a}_1 & \hat{a}_2 & \cdots & \hat{a}_n \end{bmatrix} \begin{bmatrix} y_{k-1} \\ y_{k-2} \\ \vdots \\ y_{k-n} \end{bmatrix}$$

$$= \hat{\boldsymbol{a}}^t \boldsymbol{y}$$

Linear Prediction (2)

- Parameter prediction model
  - $\hat{y}_k$: prediction value, $\hat{a}_i$: prediction parameter (prediction AR parameter)

$$\hat{y}_k = \sum_{i=1}^{n} \hat{a}_i y_{k-i}$$

$$= \begin{bmatrix} \hat{a}_1 & \hat{a}_2 & \cdots & \hat{a}_n \end{bmatrix} \begin{bmatrix} y_{k-1} \\ y_{k-2} \\ \vdots \\ y_{k-n} \end{bmatrix}$$

$$= \boldsymbol{\hat{a}}^t \boldsymbol{y}$$

線形予測 (3)

- 予測誤差の評価関数J

$$J = E\left[(y_k - \hat{y}_k)^2\right] = E\left[\left(y_k - \sum_{i=1}^{n} \hat{a}_i y_{k-i}\right)^2\right]$$

$$= E\left[y_k^2\right] - 2E\left[y_k \sum_{i=1}^{n} \hat{a}_i y_{k-i}\right] + E\left[\left(\sum_{i=1}^{n} \hat{a}_i y_{k-i}\right)^2\right]$$

- ベクトル表現では

$$J = E\left[y_k^2\right] - 2\hat{\mathbf{a}}^t E\left[y_k \mathbf{y}\right] + E\left[\left(\hat{\mathbf{a}}^t \mathbf{y}\right)^2\right]$$

Linear Prediction (3)

- Evaluation functio for prediction error J

$$J = E\left[(y_k - \hat{y}_k)^2\right] = E\left[\left(y_k - \sum_{i=1}^{n} \hat{a}_i y_{k-i}\right)^2\right]$$

$$= E\left[y_k^2\right] - 2E\left[y_k \sum_{i=1}^{n} \hat{a}_i y_{k-i}\right] + E\left[\left(\sum_{i=1}^{n} \hat{a}_i y_{k-i}\right)^2\right]$$

- In vector format

$$J = E\left[y_k^2\right] - 2\hat{\mathbf{a}}^t E\left[y_k \mathbf{y}\right] + E\left[\left(\hat{\mathbf{a}}^t \mathbf{y}\right)^2\right]$$

線形予測 (4)

- 評価関数Jを最小とする予測係数を求めるために、$\hat{a}_i$で偏微分

$$\frac{\partial J}{\partial \hat{\mathbf{a}}} = -2E\left[y_k \mathbf{y}\right] + 2E\left[\mathbf{y}\mathbf{y}^t\right]\hat{\mathbf{a}}$$

- $\partial J / \partial \hat{\mathbf{a}} = 0$とおくと

$$E\left[\mathbf{y}\mathbf{y}^t\right]\hat{\mathbf{a}} = E\left[y_k \mathbf{y}\right]$$

$$\begin{bmatrix} r(0) & r(1) & \cdots & r(n) \\ r(1) & r(0) & \ddots & r(n-1) \\ \vdots & \ddots & \ddots & \vdots \\ r(n) & r(n-1) & \cdots & r(0) \end{bmatrix} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \vdots \\ \hat{a}_n \end{bmatrix} = \begin{bmatrix} r(1) \\ r(2) \\ \vdots \\ r(n) \end{bmatrix}$$

Linear Prediction (4)

- To obtain prediction parameters which minimize J, take partial derivatives by a_i

$$\frac{\partial J}{\partial \hat{\boldsymbol{a}}} = -2E\left[y_k \boldsymbol{y}\right] + 2E\left[\boldsymbol{y}\boldsymbol{y}^t\right]\hat{\boldsymbol{a}}$$

- Here, by setting $\partial J/\partial \boldsymbol{a}=0$

$$E\left[\boldsymbol{y}\boldsymbol{y}^t\right]\hat{\boldsymbol{a}} = E\left[y_k \boldsymbol{y}\right]$$

$$\begin{bmatrix} r(0) & r(1) & \cdots & r(n) \\ r(1) & r(0) & \ddots & r(n-1) \\ \vdots & \ddots & \ddots & \vdots \\ r(n) & r(n-1) & \cdots & r(0) \end{bmatrix} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \vdots \\ \hat{a}_n \end{bmatrix} = \begin{bmatrix} r(1) \\ r(2) \\ \vdots \\ r(n) \end{bmatrix}$$

線形予測 (5)

- 相関行列に予測係数をかけたものが，1個データをずらせた相関ベクトルとなっている．予測係数は

  $$\boldsymbol{R}\hat{\boldsymbol{a}} = \boldsymbol{r}$$

  から逆行列を解いて

  $$\hat{\boldsymbol{a}} = \boldsymbol{R}^{-1}\boldsymbol{r}$$

  として求めてもよいが，より簡単な手法が存在する．

- 相関行列はToeplitz行列の形式をしており，Cholesky分解と呼ばれるLDU分解を行って逆行列を解くことができる．

Linear Prediction (5)

- Vector by multiplying parameters to autocorrelation matrix equals correlation vector with one sample shift

$$\boxed{\boldsymbol{R}\hat{\boldsymbol{a}} = \boldsymbol{r}}$$

parameters can be obtained by inverse matrix

$$\boxed{\hat{\boldsymbol{a}} = \boldsymbol{R}^{-1}\boldsymbol{r}}$$

However, there are simpler solutions

- Autocorrelation matrix has a shape of Toeplitz matrix. Thus, Cholesky decomposition (LDU decomposition) can be used.

線形予測 (6)

- Levinson-Durbinの解法
  - 予測係数を求める行列方程式は次の形をもつ

$$\sum_{i=1}^{n} \hat{a}_i r(|k-i|) = r(k) \qquad (k = 1, \cdots, n)$$

  - 以下に示す再帰的手法により低次の係数から順に予測係数を求めることができる

$$E^{(0)} = r(0)$$

$$h_i = \left(r(i) - \sum_{j=1}^{i-1} a_j^{i-1} r(i-j) \right) \Big/ E^{(i-1)} \qquad (1 \le i \ge n)$$

Linear Prediction (6)

- Levinson-Durbin algorithm (another simple solution)
  - Matrix equation to obtain parameters

$$\sum_{i=1}^{n} \hat{a}_i r(|k-i|) = r(k) \qquad (k = 1, \cdots, n)$$

  - Recursive algorithm shown below can obtain parameters from low order parameters

$$E^{(0)} = r(0)$$

$$h_i = \left(r(i) - \sum_{j=1}^{i-1} a_j^{i-1} r(i-j) \right) \Big/ E^{(i-1)} \qquad (1 \leq i \geq n)$$

線形予測 (7)

- Levinson-Durbinの解法(続き)

$$a_i^{(i)} = h_i$$

$$a_j^{(i)} = a_j^{(i-1)} - h_i a_{i-j}^{(i-1)} \qquad (1 \leq j \geq i-1)$$

$$E^{(i)} = (1 - h_i^2) E^{(i-1)}$$

備考: $a_j^{(i)}$ は次数iの係数の予測におけるj番目の予測係数

Linear Prediction (7)

- Levinson-Durbin algorithm (Contd.)

$$a_i^{(i)} = h_i$$

$$a_j^{(i)} = a_j^{(i-1)} - h_i a_{i-j}^{(i-1)} \qquad (1 \leq j \geq i-1)$$

$$E^{(i)} = (1-h_i^2)E^{(i-1)}$$

Note: $a_j^{(i)}$ is a j-th parameter when obtain i-th parameter

線形予測 (8)

- 2次の行列方程式の場合

$$E^{(0)} = r(0)$$

$$h_1 = r(1)/r(0)$$

$$a_1^{(1)} = r(1)/r(0)$$

$$E^{(2)} = \left(r(0)^2 - r(1)^2\right)/r(0)$$

$$h_2 = \left(r(2)r(0) - r(1)^2\right)/\left(r(0)^2 - r(1)^2\right)$$

$$a_2^{(2)} = \left(r(2)r(0) - r(1)^2\right)/\left(r(0)^2 - r(1)^2\right)$$

$$a_1^{(2)} = \left(r(1)r(0) - r(1)r(2)\right)/\left(r(0)^2 - r(1)^2\right)$$

$$a_1 = a_1^{(2)}$$

$$a_2 = a_2^{(2)}$$

Linear Prediction (8)

- At 2nd order matrix equation

$$E^{(0)} = r(0)$$

$$h_1 = r(1)/r(0)$$

$$a_1^{(1)} = r(1)/r(0)$$

$$E^{(2)} = \left(r(0)^2 - r(1)^2\right)/r(0)$$

$$h_2 = \left(r(2)r(0) - r(1)^2\right)/\left(r(0)^2 - r(1)^2\right)$$

$$a_2^{(2)} = \left(r(2)r(0) - r(1)^2\right)/\left(r(0)^2 - r(1)^2\right)$$

$$a_1^{(2)} = \left(r(1)r(0) - r(1)r(2)\right)/\left(r(0)^2 - r(1)^2\right)$$

$$a_1 = a_1^{(2)}$$

$$a_2 = a_2^{(2)}$$

問題

- ベクトルの微分が以下のようになることを示せ

$$\frac{\partial(\boldsymbol{a}^t\boldsymbol{x})}{\partial\boldsymbol{x}}=\boldsymbol{a}$$

ただし $\boldsymbol{a}^t=\begin{bmatrix}a_1 & a_2 & \cdots & a_n\end{bmatrix},\quad \boldsymbol{x}^t=\begin{bmatrix}x_1 & x_2 & \cdots & x_n\end{bmatrix}$

Quiz

- Show the partial derivative of vector is given as

$$\frac{\partial(\boldsymbol{a}^t \boldsymbol{x})}{\partial \boldsymbol{x}} = \boldsymbol{a}$$

here $\boldsymbol{a}^t = [a_1 \quad a_2 \quad \cdots \quad a_n],\quad \boldsymbol{x}^t = [x_1 \quad x_2 \quad \cdots \quad x_n]$

問題

- 2行2列の相関行列を用いた予測係数2個を求める式を示せ.

Quiz

- Show an equation to obtain 2 parameters by using 2x2 auto-correlation matrix.