INDIAN INSTITUTE OF TECHNOLOGY MADRAS

सिद्धिर्भवति कर्मजा

Time Series Analysis

Kaustabh Ganguly

ch23m514 roll no.

Leaked paper solution

1. **Answer the following:**

Section A

Let $Y = aX + b$.

(a) Find the covariance of X and Y.

(b) Find the correlation coefficient of X and Y.

Section B

Let $X_1, \ldots X_n$ be n random variables, Find the variance $Var(\sum a_i x_i)$ (for I = 1 to n) in terms of covariance.

Covariance:

covariance measures how much two random variables vary together

$$\text{cov}(X, Y) = E[(X - \mu_x)(Y - \mu_y)]$$

$E[\cdot]$ is the expected value & μ_x & μ_y are means of X & Y.

covariance of X & Y when $Y = aX + b$

$$\text{cov}(X, Y) = E[(X - \mu_x)(aX + b - \mu_y)]$$

$\{\because Y = aX + b, \ \mu_y = a\mu_x + b\}$

$$\text{cov}(X, Y) = E[(X - \mu_x)(aX + \cancel{b} - a\mu_x - \cancel{b})]$$

$$\Rightarrow \text{cov}(X, Y) = aE[(X - \mu_x)(X - \mu_x)]$$

$$= a\,Var(X)$$

Correlation coefficient:

The correlation coefficient, denoted ρ is the measure of strength & direction of the relationship between two variables.

Defined as:

$$\rho_{x,y} = \frac{Cov(X,Y)}{\sigma_x \sigma_y}$$ [σ_x & σ_y are s.d. of X & Y]

If $Y = aX + b$, then $\sigma_y = |a| \sigma_x$

$$\rho_{x,y} = \frac{a Var(X)}{\sigma_x |a| \sigma_x} = sgn(a)$$

where, sgn(a) is the signum function, which extracts the sign of a.

Variance of a Linear Combination of Random Variable

Finding $Var\left(\sum_{i=1}^{n} a_i X_i\right)$:

Let, $Z = \sum_{i=1}^{n} a_i X_i$

$$\mu_Z = E\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i E[X_i]$$

$$Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + 2ab\, Cov(X,Y)$$

$$Var(Z) = Var\left(\sum_{i=1}^{n} a_i X_i\right)$$

$$= Var(a_1 X_1 + a_2 X_2 + \cdots + a_n X_n)$$

$$= \sum_{i=1}^{n} a_i^2 Var(X_i) + 2 \sum_{i<j} a_i a_j Cov(X_i, X_j)$$

2. Answer the following:

Section A

Consider a random process $X(t)$ defined by

$$X(t) = U \cos t + V \sin t \qquad -\infty < t < \infty$$

where U and V are independent r.v.'s, each of which assumes the values -2 and 1 with the probabilities $\frac{1}{3}$ and $\frac{2}{3}$, respectively. Show that $X(t)$ is WSS but not strict-sense stationary.

Section B

Consider a random process given by:

X(t) = Y Coswt for t>= 0. Here w is a constant and Y is a uniformly distributed random variable in the interval [0 1]. Find

a) E(X(t))
b) Autocovariance function at any lag l of X(t)

→ Wide Sense Stationarity:

A stochastic process $X(t)$ is WSS when:

a) The mean $E[X(t)]$ is invariant over time.

b) The autocorrelation function $R_{xx}(\tau) = E[X(t)X(t+\tau)]$ relies solely on the time difference τ and not on the absolute time instants t & $t+\tau$.

Mathematically:

$$E[X(t)] = \mu \;\; \forall t$$

$$R_{xx}(\tau) = E(X(t)X(t+\tau)] = R_{xx}(t_2 - t_1)$$

Strict Sense Stationarity:

A process $X(t)$ is SSS if every finite dimensional distribution remained unchanged when shifted in time. Formally, for every positive integer n, for all point $t_1, t_2, \ldots, t_n$ and for all time shifts τ, the joint distribution of $[X(t_1), X(t_2) \ldots X(t_n)]$ is same as $[X(t_1+\tau), X(t_2+\tau), \ldots, X(t_n+\tau)]$

So for section 1:

$$X(t) = U\cos t + V\sin t \quad \forall\ t \in (-\infty, \infty)$$

$$E[X(t)] = \cos t\ E[U] + \sin t\ E[V]$$

$\{E[\cdot]$ is a linear operator$\}$

$$= \cos t \times \left(-2 \times \frac{1}{3} + 1 \times \frac{2}{3}\right) + \sin t \times \left(-2 \times \frac{1}{3} + 1 \times \frac{2}{3}\right) = 0$$

So, it satisfies first condition of WSS.

$$X(t)\,X(t+\tau)$$

$$= (U\cos t + V\sin t)\ (U\cos(t+\tau) + V\sin(t+\tau))$$

$$= U^2\cos t\cos(t+\tau) + UV\cos t\sin(t+\tau) + VU\sin t\cos(t+\tau) + V^2\sin t\sin(t+\tau)$$

$$E[X(t)X(t+\tau)] = E[U^2]\cos t\cos(t+\tau) + E[V^2]\sin t\sin(t+\tau)$$

$[\because U$ & V are independent $E[UV] = E[U]E[V] = 0]$

$$= 2 \times \{\cos t\cos(t+\tau) + \sin t\sin(t+\tau)\}$$

$$\left\{\because E[U^2] = E[V^2] = (-2)^2 \times \frac{1}{3} + 1^2 \times \frac{2}{3} = 2\right\}$$

$$= 2\cos(t - (t+\tau))$$

$$= 2\cos(\tau) \quad [\because \cos(-\theta) = \cos\theta]$$

$\therefore$ The term is only dependent on τ. (Ans)

So $X(t)$ is definitely WSS

Checking SSS:

We prove X(t) is not SSS by showing it's joint CDF changes with time shifts.

$$X(t) = U\cos(t) + V\sin(t)$$

Let, consider 2 arbitrary time points t & t+τ.

The joint CDF at t & t+τ:

$$F_{X(t),X(t+\tau)}(x,y) = P(X(t) \leq x, X(t+\tau) \leq y)$$

At $t=0$ & $\tau = \pi/2$

$X(0) = U$ $\quad X(\pi/2) = V$

$X(\pi) = -U$ $\quad (T+2\tau)$

So joint distribution of t & t+τ is different than t+τ & t+2τ. So it's not SSS.

Section B

Consider a random process given by:

X(t) = Y Coswt for t>= 0. Here w is a constant and Y is a uniformly distributed random variable in the interval [0 1]. Find

a) E(X(t))
b) Autocovariance function at any lag l of X(t)

→ $X(t) = Y\cos(\omega t)$ for $t \geq 0$, ω is constant, Y is uniform in interval [0,1]

$$E[x] = \sum_i x_i P(X = x_i) \rightarrow \text{For discrete}$$

$$E[x] = \int_{-\infty}^{\infty} x f(x)\,dx \rightarrow \text{For continuous}$$

$\{f(x) = \text{pdf}\}$

Autocovariance:

The autocovariance function provides a measure of the linear dependence of a random process with itself at two points in time. For a random process $X(t)$, the autocovariance function is given by:

$$R_x(t_1, t_2) = E\left[\left(X(t_1) - E[X(t_1)]\right)\left(X(t_2) - E[X(t_2)]\right)\right]$$

where $R_x(t_1, t_2)$ is the autocovariance function, & $E[\cdot]$ denotes expectations.
For stationary process, the autocovariance function depends only on the difference between t_1 & t_2 and can be written as:

$$R_x(\tau) = E\left[\left(X(t) - E[X(t)]\right)\left(X(t+\tau) - E[X(t+\tau)]\right)\right]$$

$$\{\tau = t_1 - t_2\}$$

$E[X(t)]$:

$X(t) = Y\cos(\omega t)$ for $t \geq 0$, and Y is uniformly distributed in the interval $[0, 1]$

$$E[X(t)] = E[Y\cos\omega t]$$
$$= E[Y] \times \cos\omega t$$

The expected value of an uniformly distributed random variable over $[a, b]$

is $\frac{a+b}{2}$. $E[Y] = \frac{0+1}{2} = 1/2$

$\therefore E[X(t)] = \frac{1}{2}\cos\omega t$. (a). Ans

b) Autocovariance function at any lag ℓ of $X(t)$:

$$R_x(\ell) = E\left[\left(X(t) - E[X(t)]\right)\left(X(t+\ell) - E[X(t+\ell)]\right)\right]$$

$$= E\left[(Y-1/2)\cos\omega t \times (Y-1/2)\cos\omega(t+\ell)\right]$$

$$= E\left[(Y-1/2)^2\right]\cos\omega t\cos\omega(t+\ell)$$

$$= \left(E[Y^2] - E[Y] + \frac{1}{4}\right)\cos\omega t\cos\omega(t+\ell)$$

$$= \left(\int_0^1 y^2\,dy - \frac{1}{2} + \frac{1}{4}\right)\cos\omega t\cos\omega(t+\ell)$$

$$= E\left[1/12\cos\omega t\cos\omega(t+\ell)\right]$$

$$= E\left[1/24\left[\cos\omega\ell + \cos(2\omega t + \ell)\right]\right]$$

$$\left\{\cos A\cos B = \frac{1}{2}\cos(A-B) + \cos(A+B)\right\}$$

$$= E\left[1/24\left\{\cos\omega\ell + \left[\cos(2\omega t + \omega\ell)\right]\right\}\right]$$

$$= 1/24\left\{E[\cos\omega\ell] + E\left[\cos(2\omega t + \omega\ell)\right]\right\}$$

$$= \frac{1}{24}\cos\omega\ell \quad \left\{\text{As} \int_0^{2\pi}\cos(2\omega t + \omega\ell)\,dt = 0\right\}$$

$\therefore R_x(\ell) = \frac{1}{24}\cos\omega\ell$. (Ans)

3. Answer the following:

A time-series is modelled using the following equation:

$$v[k] + d_1 v[k-1] + \cdots + d_N v[k-N] = e[k]$$

where $e[k]$ is the white-noise sequence.

Devise a **theoretical** method to estimate the N parameters $\{d_k\}$ that involves the use of ACF of $v[k]$. Verify your solution for the case when $N = 3$ and $d_1 = 0.7$, $d_2 = 0.3$ and $d_3 = 0.5$.

In the above question white-noise sequence is simply an i.i.d sequence (like sequence of tossing of coins). You can assume its mean to be 0 and variance to be ����. Hint is to compute Autocorrelation function and find a procedure to estimate the unknown parameter values. Verify the method using the example system for N = 3.

A time sereies is a sequence of data points typically consisting of successive measurements made over a time interval. White noise is a random signal having equal intensity at different frequencies. It refers to a sequence that is uncorrelated over time.

Autocorrelation function (ACF)

The ACF measures the linear relationship b/w an observation at time t and the observations at previous times. It provides insight into the repeating patterns or seasonality of a time series. ACF at lag l:

$$R_v(l) = E[V[K] \times V[K-l]]$$

For a linear time series model, the relationship between the time series value at time t & it's past values can be represented by a linear combination of coefficients and past values. The ACF can be used to estimate these coefficients and past values. The ACF can be used to estimate these coefficients by setting up a system of equations based on the ACF values at different lags

$$v[k] + d_1 v[k-1] + \dots + d_N v[k-N] = e[k]$$

N is the order of the model.

This is a linear difference equation.

i) $v[k]$ is the value of time series at k

ii) d_i are the co-efficient or parameters.

iii) $e[k]$ is the white noise.

Multiplying both sides with $v[k-l]$ & taking expectations:

$$E\big[v[k]v[k-l]\big] + d_1 E\big[v[k-1]v[k-l]\big] + \dots + d_N E\big[v[k-N]v[k-l]\big] = E\big[e[k]v[k-l]\big]$$

$$R_v(l) + d_1 R_v(l-1) + \dots + d_N R_v(l-N) = \sigma^2 \delta(l)$$

$\delta(l) = 1$ if $l = 0$ & 0 otherwise

$v[k-l] = 0$ for $l > 0$

For $N=3$, when

$$R_v(0) + d_1 R_v(-1) + d_2 R_v(-2) + d_3 R_v(-3) = \sigma^2$$

$$R_v(1) + d_1 R_v(0) + d_2 R_v(-1) + d_3 R_v(-2) = 0$$

$$R_v(2) + d_1 R_v(1) + d_2 R_v(0) + d_3 R_v(-1) = 0$$

$$R_v(3) + d_1 R_v(2) + d_2 R_v(1) + d_3 R_v(0) = 0$$

As autocorrelation function is symmetric
$R_v(l) = R_v(-l)$

$d_1 = 0.7$ $d_2 = 0.3$ & $d_3 = 0.5$

Yule Walker equation can be written as Matrix form:

$$Rd = b$$

R is toeplitz matrix of autocorrelation.
d is the vector of parameters $\{d_1, d_2, d_3\}$.
b is the vector on right hand side.

$$\begin{bmatrix} R_v(0) & R_v(1) & R_v(2) & R_v(3) \\ R_v(1) & R_v(0) & R_v(1) & R_v(2) \\ R_v(2) & R_v(1) & R_v(0) & R_v(1) \\ R_v(3) & R_v(2) & R_v(1) & R_v(0) \end{bmatrix} \begin{bmatrix} 1 \\ d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} \sigma^2 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

We can also write the following:

$$R_v(0) + 0.7R_v(1) + 0.3R_v(2) + 0.5R_v(3) = \sigma^2$$
$$R_v(1) + 0.7R_v(0) + 0.3R_v(1) + 0.5R_v(2) = 0$$
$$R_v(2) + 0.7R_v(1) + 0.3R_v(0) + 0.5R_v(1) = 0$$
$$R_v(3) + 0.7R_v(2) + 0.3R_v(1) + 0.5R_v(0) = 0$$

$$\begin{bmatrix} 1 & 0.7 & 0.3 & 0.5 \\ 0.7 & 1.3 & 0.5 & 0 \\ 0.3 & 1.2 & 1 & 0 \\ 0.5 & 0.3 & 0.7 & 1 \end{bmatrix} \begin{bmatrix} R_v(0) \\ R_v(1) \\ R_v(2) \\ R_v(3) \end{bmatrix} = \begin{bmatrix} \sigma^2 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

```python
import numpy as np
from sympy import symbols, Eq, solve, Matrix

# Step 1: Define the system parameters and symbols
d1, d2, d3, R0, R1, R2, R3, sigma2 = symbols('d1 d2 d3 R0 R1 R2 R3 sigma2', real=True)

# Given values for the coefficients
d1_val = 0.7
d2_val = 0.3
d3_val = 0.5

# Step 2: Formulate the Yule-Walker equations
# These equations relate the ACF values at different lags to the model coefficients
eq1 = Eq(R0 + d1*R0 + d2*R1 + d3*R2, sigma2)
eq2 = Eq(R1 + d1*R0 + d2*R1 + d3*R2, 0)
eq3 = Eq(R2 + d1*R1 + d2*R0 + d3*R1, 0)
eq4 = Eq(R3 + d1*R2 + d2*R1 + d3*R0, 0)

# Step 3: Express the Yule-Walker equations in matrix form (Ax = b)
# Define the coefficient matrix A
A = Matrix([[1 + d1_val, d2_val, d3_val, 0],
            [d1_val, 1 + d2_val, d3_val, 0],
            [d2_val, d1_val + d3_val, 1, 0],
            [d3_val, d2_val, d1_val, 1]])

# Define the constant vector b
b = Matrix([sigma2, 0, 0, 0])

# Step 4: Solve the matrix equation for x
x = A.LUsolve(b)

# Print the result
print(x)
```

```
Matrix([[0.56*sigma2], [-0.44*sigma2], [0.36*sigma2], [-0.4*sigma2]])
```

$R_v(0) = 0.56\sigma^2$

$R_v(1) = -0.44\sigma^2$

$R_v(2) = 0.36\sigma^2$

$R_v(3) = -0.4\sigma^2$. (Ans)

4. Answer the following:

Consider the analog signal $x(t) = 3\cos(100\pi t)$

(a) Determine the minimum sampling rate required to avoid aliasing

(b) Suppose that the signal is sampled at the rate $F_s = 200$ Hz. What is the discrete-time signal obtained after sampling?

(c) Suppose that the signal is sampled at the rate $F_s = 75$ Hz.. What is the discrete-time signal obtained in practice after sampling? What is the frequency of the continuous-time signal obtained upon reconstruction from this d.t. signal?

$$x(t) = 3\cos(100\pi t)$$

a) $\omega = 2\pi F = 100\pi$

$F = 50$

$\therefore$ Nyquist rate $2 \times 50 = 100$ Hz.

b) $x[n] = x(nT_s)$ $T_s = \frac{1}{F_s}$

$x[n] = 3\cos(100\pi n T_s)$

$F_s = 200$ Hz

$x[n] = 3\cos\left(\frac{100}{200}\pi n\right) = 3\cos(0.5\pi n)$

c) $F_s = 75$ Hz

$F_{alias} = |F - K \times F_s|$ where K minimises F_{alias}

& $F_{alias} < F_s/2$

$F = 50$ Hz $F_s = 75$ Hz

$K = 0 :\rightarrow$ $F_{alias} = |50 - 0| = 50$ Hz

$K = 1 :\rightarrow$ $F_{alias} = |50 - 75| = 25$ Hz

$k=2 : \rightarrow F_{alias} = |50-150| = 100\,Hz$

$k=3 : \rightarrow F_{alias} = |50-225| = 175\,Hz.$

so, at $k=1$ it is minimum.

$25\,Hz < \frac{F_s}{2}$

$25 < 37.5$ so this is valid.

$x[n] = 3\cos\left(2\pi \times 25 \times \frac{1}{75} n\right)$

$= 3\cos\left(\frac{2\pi}{3} n\right).$

5. **Answer the following:**

Part A

The joint pdf of a bivariate r.v. (X, Y) is given by

$$f_{XY}(x, y) = \begin{cases} kxy & 0 < x < 1, 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

where k is a constant.

(*a*) Find the value of k.

(*b*) Are X and Y independent?

Part B

Show that summation (until time t) of a stationary i.i.d noise e(t) with distribution N(0,��[��]**) is non stationary.**

$\rightarrow$

$$f_{XY}(x,y) = \begin{cases} kxy & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 \end{cases}$$

$$\int_0^1 \int_0^1 kxy \, dx\, dy = 1$$

$\Rightarrow K \times \frac{1}{4} = 1 \Rightarrow K = 4$ a. (Ans)

b) Marginal pdf for X & Y :

$$f_X(x) = \int_0^1 f_{XY}(x,y)\,dy = \frac{4}{2} \times x$$

$$f_Y(y) = \int_0^1 f_{XY}(x,y)\,dx = \frac{4}{2} \times y$$

We can say

$$f_{XY}(x,y) = f_X(x) \cdot f_Y(y)$$

LHS : $4xy$ RHS : $2x \times 2y = 4xy$

LHS = RHS

so, X & Y are independent.

Part B:

given stationary i.i.d noise $e(t)$ with distribution $\mathcal{N}(0, \sigma^2)$, summation:

$$s(t) = \sum_{i=1}^{t} e(i)$$

The mean of $S(t)$ is :

$$E[S(t)] = E\left[\sum_{i=1}^{t} e(i)\right] = \sum_{i=1}^{t} E[e(i)] = 0 \;\forall t$$

Variance:

$$Var[S(t)] = E\left[\left\{\sum_{i=1}^{t} e(i)\right\}^2\right] \quad \text{as mean}=0$$

$$= E\left[\sum_{i=1}^{t} e(i)^2 + \sum_{i=1}^{t} \sum_{i \neq j} e(i)e(j)\right]$$

as $e(i)$ & $e(j)$ are independent & have 0 mean

$$E[e(i)e(j)] = E[e(i)]E[e(j)] = 0$$

$$Var[S(t)] = \sum_{i=1}^{t} E[e(i)^2] = t\sigma^2$$

For a stationary process mean & var shouldn't change with time. but here it is dependent on time. So, it's non-stationary.