HW3

Jiseon Yang

2024-11-10

## Question 1

Compute the coherence (generalized correlation), A black text on a white background

Description automatically generated

1. for the cross-covariance A black and white image of a square and a square

   Description automatically generated with medium confidence.

A close-up of a paper

Description automatically generated

1. for the cross-covariance A black and white symbol

   Description automatically generated.

A close-up of math equations

Description automatically generated

## Question 2

Let be a bivariate process with a stationary cross-covariance matrix function

A close-up of a mathematical equation

Description automatically generated

and a set of covariates **x(s)**. Let A black and white text

Description automatically generatedbe the data vector, with A black text with a white background

Description automatically generated, and .

1. Show that the cokriging predictor has the for , with appropriate definitions of and .

A paper with math equations

Description automatically generated

1. Show further that if  is a site where is observed, then for , if and only if .

A close-up of a paper

Description automatically generated

## Question 3

For a moving average process of the form

where are independent with zero means and variance , determine the autocovariance and autocorrelation functions as a function of lag and plot the ACF as a function of .

A math equations on a notebook

Description automatically generated

A math equations on a notebook

Description automatically generated

A graph on a piece of paper

Description automatically generated

# Define the parameters  
sigma\_w2 <- 1 # Variance of w\_t  
  
# Function to calculate autocovariance for a given lag h  
calc\_autocovariance <- function(h) {  
 if (h == 0) {  
 # Variance at lag 0  
 return(6 \* sigma\_w2)  
 } else if (h == 1) {  
 # Covariance at lag 1  
 return(4 \* sigma\_w2)  
 } else if (h == 2) {  
 # Covariance at lag 2  
 return(1 \* sigma\_w2)  
 } else {  
 # Covariance for h >= 3  
 return(0)  
 }  
}  
  
# Calculate autocovariance values for lags 0 to 4  
lags <- 0:4  
gamma\_values <- sapply(lags, calc\_autocovariance)  
  
# Calculate autocorrelation values  
gamma\_0 <- gamma\_values[1] # Autocovariance at lag 0  
rho\_values <- gamma\_values / gamma\_0  
  
# Print autocovariance and autocorrelation values  
cat("Autocovariance values:\n")

## Autocovariance values:

print(gamma\_values)

## [1] 6 4 1 0 0

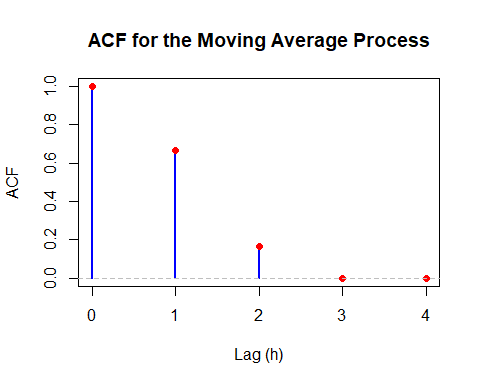
cat("Autocorrelation values:\n")

## Autocorrelation values:

print(rho\_values)

## [1] 1.0000000 0.6666667 0.1666667 0.0000000 0.0000000

# Plot the ACF  
plot(lags, rho\_values, type = "h", main = "ACF for the Moving Average Process",  
 xlab = "Lag (h)", ylab = "ACF", col = "blue", lwd = 2)  
points(lags, rho\_values, col = "red", pch = 16)  
abline(h = 0, col = "gray", lty = 2)



## Question 4

In this problem, we explore the difference between a random walk and a trend stationary process.

1. Generate four series that are random walk with drift, , of length with and . Call the data for . Fit the regression using least squares. Plot the data, the true mean function (i.e. ) and the fitted line, , on the same graph. Hint: The following R code may be useful.

par(mfrow=c(2,2), mar=c(2.5,2.5,0,0)+.5, mgp=c(1.6,.6,0)) # set up   
for (i in 1:4){  
 x = ts(cumsum(rnorm(100,.01,1))) # data  
 regx = lm(x~0+time(x), na.action=NULL) # regression  
 plot(x, ylab='Random Walk w Drift') # plots  
 abline(a=0, b=.01, col=2, lty=2) # true mean (red - dashed)  
 abline(regx, col=4) # fitted line (blue - solid)  
}

A group of graphs showing different types of walking

Description automatically generated

* The plot shows each random walk series (black line), the true trend (red dashed line), and the fitted regression line (blue solid line)
* Black: generated data for the random walk with drift,

xt – δt + Σwj, δ= 0.01, wj ~ N(0,1)

* Red: True mean function μt= 0.01t. Expected linear drift over time with slop δ = 0.01.
* Blue: Fitted regression line fitted from the least squares estimation. While it approximates the trend in random walk, it is influenced by the variability inherent in the random walk itself, which may cause deviations from the true mean.
* Plots show how random walks can significantly **diverge** from the true mean function over time **due to the accumulation of random terms**. This variability reflects the unpredictable nature of random walks and how they can draft away from the expected trend, even when a constant drift component δ is present.

1. Generate four series of length that are linear trend plus noise, say , where and are as in part (a). Fit the regression using least squares. Plot the data, the true mean function (i.e. ) and the fitted line, , on the same graph.

set.seed(123) # For reproducibility  
par(mfrow = c(2, 2), mar = c(2.5, 2.5, 0, 0) + 0.5, mgp = c(1.6, 0.6, 0))  
  
# Generate 4 series of linear trend plus noise and plot  
for (i in 1:4) {  
 t = 1:100 # Time index  
 y = 0.01 \* t + rnorm(100, 0, 1) # Generate data: linear trend + noise  
 regy = lm(y ~ 0 + t, na.action = NULL) # Fit regression: y\_t = beta \* t + w\_t  
 plot(y, type = "l", ylab = 'Linear Trend + Noise', main = paste("Series", i))  
 abline(a = 0, b = 0.01, col = 2, lty = 2) # True mean function (red, dashed)  
 abline(regy, col = 4) # Fitted line (blue, solid)  
}

A group of graphs showing different types of noise

Description automatically generated

* Black: the generated data series with a linear trend plus noise, showing the random fluctuations around the trend line.
* Red: true mean function, the expected trend μt=0.01t.
* Blue: fitted regression line obtained from least squares estimation, the estimated trend based on the noisy data.
* The slop (0.01t) is visible but but the noise (wt) introduces variability around the trend which makes the data fluctuate around the trend line while still following an underlying linear trend.
* **The fitted line (blue) and the true mean function (red) are generally aligned**, indicating the estimated beta\_hat is close to the true slope of 0.01. Due to the noise term wt, the fitted lines and true mean do not perfectly match as the noise influences the data and the regression fit.

1. Comment (what did you learn from this assignment).

* Randone Walk with Drift: The fitted regression line diverges from the true mean function. It is due to the process *accumulating random noise over time*. xt = δt +Σwj includes a cumulative sum of random noise with a draft term δ. Thus, deviations from the expected path can become significant over time, leading substantial divergence from the true mean function due to the accumulated effect of random noise (compounding effect of random fluctuations) 🡺 Characteristic of random walks
* Trend Stationary Process: The fitted regression line closely matches the true mean function because the noise term wt is NOT accumulated over time. xt = δt + wj has a fixed linear trend with the noise fluctuating around the linear trend and noise does not change the slope over time (not compounding). Thus, unlike a random walk, the series maintains a stationary variance over time with fluctuation around the true mean 🡺 centered around its trend, making it easier to predict.