Time Series Analysis - STAT 478 - Final Exam - Part 1 Q3

Alex Towell (atowell@siue.edu)

Part 1: Problem 3

Consider the model $Y_t = \beta_1 t + X_t$.

Part (a)

$\{X_t\}$ is a zero mean white noise process with $\text{Var}(X_t) = \sigma^2$. Find the least square estimator of β_1.

Proof. Y_t is a random variable defined by

$$Y_t = \beta_1 t + X_t$$

where X_t is zero mean white noise. Then, $X_t = Y_t - \beta_1 t$ and we wish to find a value for β_1 that minimizes

$$\text{Q}(\beta_1) = \sum_{t=1}^{T} X_t^2 = \sum_{t=1}^{T} (Y_t - \beta_1 t)^2.$$

since X_t has an expectation of zero mean. We observe that Q is convex with a global minimum where its derivative is zero, thus we solve for $\hat{\beta}_1$ in

$$\begin{aligned}
\left.\frac{\mathrm{d}Q}{\mathrm{d}\beta}\right|_{\hat{\beta}_1} &= 0 \\
-2\sum (Y_t - \hat{\beta}_1 t)t &= 0 \\
\sum tY_t &= \hat{\beta}_1 \sum t^2 \\
\hat{\beta}_1 &= \frac{\sum tY_t}{\sum t^2}.
\end{aligned}$$

Note that $\sum_{t=1}^{T} t^2 = T(T+1)(2T+1)/6$ and thus

$$\hat{\beta}_1 = \frac{6\sum tY_t}{T(T+1)(2T+1)}. \tag{1}$$

□

This is the BLUE estimator of β_1.

Part (b)

Suppose $\{X_t\}$ is a process of the form

$$X_t = X_{t-1} + e_t - \theta e_{t-1}.$$

Derive the ACF for $\{\nabla Y_t\}$ and show that $\{\nabla Y_t\}$ is stationary. What is the name of the process identified by $\{\nabla Y_t\}$?

First, we derive an explicit equation for ∇Y_t,

$$\begin{aligned}\nabla Y_t &= Y_t - Y_{t-1}\\ &= (\beta_1 t + X_t) - (\beta_1(t-1) + X_{t-1})\\ &= \beta_1 + X_t - X_{t-1}.\end{aligned}$$

We rewrite the above by replacing X_t with its definition,

$$\begin{aligned}\nabla Y_t &= \beta_1 + (X_{t-1} + e_t - \theta e_{t-1}) - X_{t-1}\\ &= \beta_1 + e_t - \theta e_{t-1}.\end{aligned}$$

This is the form of an MA(1) with $\mu = \beta_1$. It is known that MA(1) is always stationary and has an ACF

$$\rho_k = \begin{cases} 1 & \text{if } |k| = 0\\ \frac{-\theta}{1+\theta^2} & \text{if } |k| = 1\\ 0 & \text{if } |k| > 1,\end{cases}$$

but we will derive the result from first principles.

The covariance of ∇Y_t and ∇Y_{t-k} is given by

$$\begin{aligned}\mathrm{Cov}(\nabla Y_t, \nabla Y_{t-k}) &= \mathrm{Cov}(\beta_1 + e_t - \theta e_{t-1}, \beta_1 + e_{t-k} - \theta e_{t-k-1})\\ &= \mathrm{Cov}(e_t - \theta e_{t-1}, e_{t-k} - \theta e_{t-k-1})\\ &= \mathrm{Cov}(e_t, e_{t-k}) + \mathrm{Cov}(e_t, -\theta e_{t-k-1}) + \mathrm{Cov}(-\theta e_{t-1}, e_{t-k}) + \mathrm{Cov}(-\theta e_{t-1}, -\theta e_{t-k-1}).\end{aligned}$$

Now we do a case analysis on values of k, but first note that by symmetry

$$\mathrm{Cov}(\nabla Y_{t-k}, \nabla Y_t) = \mathrm{Cov}(\nabla Y_t, \nabla Y_{t-k})$$

and so we only consider non-negative values of k.

Case 1: $k = 0$

The covariance of ∇Y_t and ∇Y_t is given by

$$\begin{aligned}\mathrm{Cov}(\nabla Y_t, \nabla Y_t) &= \mathrm{Cov}(e_t, e_t) + \mathrm{Cov}(e_t, -\theta e_{t-1}) + \mathrm{Cov}(-\theta e_{t-1}, e_t) + \mathrm{Cov}(-\theta e_{t-1}, -\theta e_{t-1})\\ &= \mathrm{Cov}(e_t, e_t) + \mathrm{Cov}(-\theta e_{t-1}, -\theta e_{t-1})\\ &= \sigma^2 + \theta^2\sigma^2\\ &= \sigma^2(1+\theta^2).\end{aligned}$$

By definition, $\mathrm{Var}(\nabla Y_t) = \mathrm{Cov}(\nabla Y_t, \nabla Y_t)$.

Case 2: $k = 1$

The covariance of ∇Y_t and ∇Y_{t-1} is given by

$$\begin{aligned}\mathrm{Cov}(\nabla Y_t, \nabla Y_{t-1}) &= \mathrm{Cov}(e_t, e_{t-1}) + \mathrm{Cov}(e_t, -\theta e_{t-2}) + \mathrm{Cov}(-\theta e_{t-1}, e_{t-1}) + \mathrm{Cov}(-\theta e_{t-1}, -\theta e_{t-2})\\ &= \mathrm{Cov}(-\theta e_{t-1}, e_{t-1})\\ &= -\theta\,\mathrm{Cov}(e_{t-1}, e_{t-1})\\ &= -\theta\sigma^2.\end{aligned}$$

Case 3: $k \geq 2$

First, we consider $k = 2$ and generalize the result. The covariance of ∇Y_t and ∇Y_{t-2} is given by

$$\begin{aligned}\text{Cov}(\nabla Y_t, \nabla Y_{t-2}) &= \text{Cov}(e_t, e_{t-2}) + \text{Cov}(e_t, -\theta e_{t-3}) + \text{Cov}(-\theta e_{t-1}, e_{t-2}) + \text{Cov}(-\theta e_{t-1}, -\theta e_{t-3}) \\ &= 0.\end{aligned}$$

Generalizing the result, we see that for $k > 2$, every pair of errors have 0 covariance. Thus,

$$\text{Cov}(\nabla Y_t, \nabla Y_{t-k}) = 0$$

if $k > 0$.

Autocorrelation function

We see that the covariances are independent of time t and are only a function of lag k, and thus the *autocorrelation* function is given by

$$\begin{aligned}\rho_k &= \frac{\text{Cov}(\nabla Y_t, \nabla Y_{t-k})}{\sqrt{\text{Var}(\nabla Y_t)}\sqrt{\text{Var}(\nabla Y_{t-k})}} \\ &= \frac{\text{Cov}(\nabla Y_t, \nabla Y_{t-k})}{\sigma^2(1+\theta^2)},\end{aligned}$$

which simplifies to

$$\rho_k = \begin{cases} 1 & \text{if } |k| = 0 \\ \frac{-\theta}{1+\theta^2} & \text{if } |k| = 1 \\ 0 & \text{if } |k| > 1. \end{cases}$$

Stationarity conditions

We see that its mean is constant $\text{E}(\nabla Y_t) = \beta_1$, its variance is constant, and its autocorrelation function is strictly a function of lag k. Thus, ∇Y_t is stationary.