Lecture 9

Regression with Time Series Errors

Readings: Brockwell & Davis Ch 6.6; Shumway & Stoffer Ch 3.8

MATH 8090 Time Series Analysis
October 14, 2021

Whitney Huang
Clemson University

Agenda

1. **Time Series Regression Models**

2. **Generalized Least Squares Regression**

3. **Lake Huron Example**

Time Series Regression

Suppose we have the following time series model for $\{Y_t\}$:

$$Y_t = m_t + \eta_t,$$

where

- m_t captures the mean of $\{Y_t\}$, i.e., $\mathbb{E}(Y_t) = m_t$
- $\{\eta_t\}$ is a zero mean stationary process with ACVF $\gamma_\eta(\cdot)$

The component $\{m_t\}$ may depend on time t, or possibly on other explanatory series

Example Models for m_t: Trends and Seasonality

- Constant trend model: For each t let $m_t = \beta_0$ for some unknown parameter β_0
- Simple linear regression: For unknown parameters β_0 and β_1,

$$m_t = \beta_0 + \beta_1 x_t,$$

where $\{x_t\}$ is some explanatory variable indexed in time (may just be a function of time or could be other series)

- Harmonic regression: For each t let

$$m_t = A\cos(2\pi f t + \phi),$$

where $A > 0$ is the amplitude (an unknown parameter), $f > 0$ is the frequency of the sinusoid (usually known), and $\phi \in (-\pi, \pi]$ is the phase (usually unknown). We can rewrite this model as

$$m_t = \beta_0 x_{1,t} + \beta_1 x_{2,t},$$

where $x_{1,t} = \cos(2\pi f t)$ and $x_{2,t} = \sin(2\pi f t)$

The Multiple Linear Regression Model

Suppose there are p explanatory series $\{x_{j,t}\}_{j=1}^{p}$, the time series model for $\{Y_t\}$ is

$$Y_t = m_t + \eta_t,$$

where

$$m_t = \beta_0 + \sum_{j=1}^{p} \beta_j x_{j,t},$$

and $\{\eta_t\}$ is a mean zero stationary process with ACVF $\gamma_\eta(\cdot)$
We can write the linear model in matrix notation:

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\eta},$$

where $\boldsymbol{Y} = (Y_1, \cdots, Y_n)^T$ is the observation vector, the coefficient vector is $\boldsymbol{\beta} = (\beta_0, \beta_1, \cdots, \beta_p)^T$, $\boldsymbol{\eta} = (\eta_1, \cdots, \eta_n)^T$ is the error vector, and the design matrix is

$$\boldsymbol{X} = \begin{bmatrix} 1 & x_{1,1} & x_{2,1} & \cdots & x_{p,1} \\ 1 & x_{1,2} & x_{2,2} & \cdots & x_{p,2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{1,n} & x_{2,n} & \cdots & x_{p,n} \end{bmatrix}$$

The Model Estimates and Distributional Results for i.i.d. Errors Case

Suppose $\{\eta_t\}$ is i.i.d. $\mathrm{N}(0, \sigma^2)$. Then the ordinary least squares (OLS) estimate of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}}_{\text{OLS}} = \left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T\boldsymbol{Y},$$

with

$$\hat{\sigma}^2 = \frac{\left(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}_{\text{OLS}}\right)^T\left(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}_{\text{OLS}}\right)}{n - (p+1)}$$

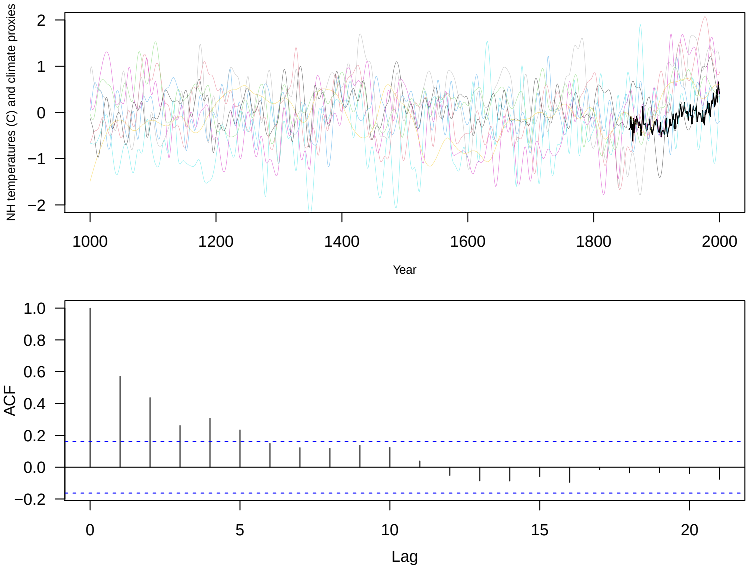
- Gauss-Markov theorem: $\hat{\boldsymbol{\beta}}_{\text{OLS}}$ is the best linear unbiased estimator (BLUE) of $\boldsymbol{\beta}$
- We have

  $$\hat{\boldsymbol{\beta}}_{\text{OLS}} \sim \mathrm{N}(\boldsymbol{\beta}, \sigma^2\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1})$$

  is independent of

  $$\frac{(n-(p+1))\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-(p+1)}$$

Climate Over Past Millennia [Jones & Mann, 2004]



Residuals from a linear regression fit are correlated in time

Generalized Least Squares Regression

When dealing with time series the errors $\{\eta_t\}$ are typically correlated in time

- Assuming the errors $\{\eta_t\}$ are a stationary Gaussian process, consider the model

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\eta},$$

  where $\boldsymbol{\eta}$ has a multivariate normal distribution, i.e., $\boldsymbol{\eta} \sim \mathrm{N}(\mathbf{0}, \Sigma)$

- The generalized least squares (GLS) estimate of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}}_{\mathrm{GLS}} = \left(\boldsymbol{X}^T \Sigma^{-1} \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \Sigma^{-1} \boldsymbol{Y},$$

  with

$$\sigma^2 = \frac{\left(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}_{\mathrm{GLS}}\right)^T \left(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}_{\mathrm{GLS}}\right)}{n - (p+1)}$$

Distributional Properties of Estimators

Gauss-Markov theorem: $\boldsymbol{\beta}_{\mathrm{GLS}}$ is the best linear unbiased estimator (BLUE) of $\boldsymbol{\beta}$

- We have
$$\hat{\boldsymbol{\beta}}_{\mathrm{GLS}} \sim \mathrm{N}(\boldsymbol{\beta}, \sigma^2 \left(\boldsymbol{X}^T \Sigma^{-1} \boldsymbol{X}\right)^T)$$

- The variance of linear combinations of $\hat{\boldsymbol{\beta}}_{\mathrm{GLS}}$ is less than or equal to the variance of linear combinations of $\hat{\boldsymbol{\beta}}_{\mathrm{OLS}}$

Applying GLS In Practice

The main problem in applying GLS in practice is that Σ depends on ϕ, $\boldsymbol{\theta}$, and σ^2 and we have to estimate these

- A two-step procedure

  1. Estimate $\boldsymbol{\beta}$ by OLS, calculating the residuals $\hat{\boldsymbol{\eta}} = \boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}_{\mathrm{OLS}}$, and fit a ARMA to $\hat{\eta}$ to get Σ
  2. Re-estimate $\boldsymbol{\beta}$ using GLS

- Alternatively, we can consider one-shot maximum likelihood methods

Likelihood-based Regression Methods

Model:

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\eta},$$

where $\boldsymbol{\eta} \sim \mathrm{N}(\mathbf{0}, \Sigma)$

$$\Rightarrow \boldsymbol{Y} \sim \mathrm{N}(\boldsymbol{X}\boldsymbol{\beta}, \Sigma)$$

- We maximum the Gaussian likelihood

$$L_n(\boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\theta}, \sigma^2) = (2\pi)^{-n/2}|\Sigma|^{-1/2} \exp\left[-\frac{1}{2}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta})^T \Sigma^{-1} (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta})\right]$$

  with respect to the regression parameters $\boldsymbol{\beta}$ and ARMA parameters $\boldsymbol{\phi}$, $\boldsymbol{\theta}$, σ^2 simultaneously

- As before, we can re-express the likelihoods using the one-step-ahead predictions

An Example: Lake Huron Levels

Model:

$$Y_t = m_t + \eta_t$$

where

$m_t = \beta_0 + \beta_1 t$

$\{\eta_t\}$ is some ARMA(p, q) process

- Scientific Question: Is there evidence that the lake level has been changing steadily over the years 1875-1972?
- Statistical Hypothesis:

Fitting Result form the Two-Step Procedure

```
> lm <- lm(LakeHuron ~ years)
> lm$coefficients
 (Intercept)        years
625.55491791  -0.02420111
> (MLE_est1 <- arima(lm$residuals, order = c(2, 0, 0),
+                    include.mean = FALSE))

Call:
arima(x = lm$residuals, order = c(2, 0, 0), include.mean = FALSE)

Coefficients:
         ar1      ar2
      1.0050  -0.2925
s.e.  0.0976   0.1002

sigma^2 estimated as 0.4572:  log likelihood = -101.26,  aic = 208.51
```

Fitting Result from One-Step MLE

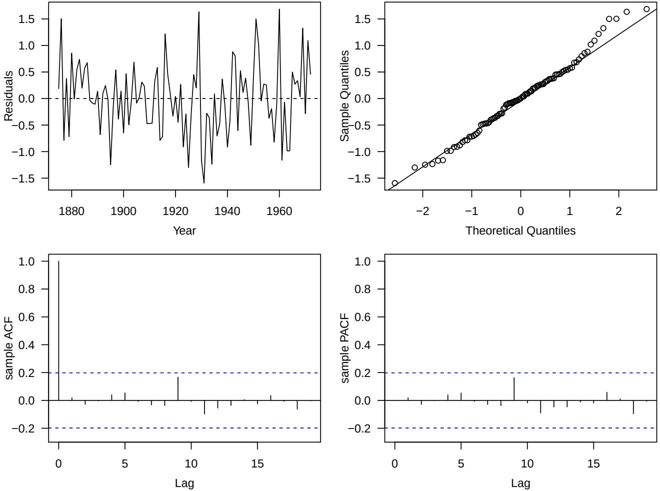

```
> mle <- arima(LakeHuron, order = c(2, 0, 0),
+              xreg = cbind(rep(1,length(LakeHuron)), years),
+              include.mean = FALSE)
> mle

Call:
arima(x = LakeHuron, order = c(2, 0, 0), xreg = cbind(rep(1, length(LakeHuron)),
    years), include.mean = FALSE)

Coefficients:
         ar1      ar2  rep(1, length(LakeHuron))
      1.0048  -0.2913                   620.5115
s.e.  0.0976   0.1004                    15.5771
        years
      -0.0216
s.e.   0.0081

sigma^2 estimated as 0.4566:  log likelihood = -101.2,  aic = 212.4
```

MLE Fit Diagnostics




```
> plot.residuals(years, resid(mle), xlab = "Year", ylab = "Residuals")

        Box-Ljung test

data:  y
X-squared = 6.2088, df = 19, p-value = 0.9974
```

Comparing Confidence Intervals

```
> confint(lm)
                   2.5 %      97.5 %
(Intercept) 610.14291793 640.9669179
years        -0.03221272  -0.0161895
> confint(MLE_est1)
         2.5 %      97.5 %
ar1  0.8137180  1.19630830
ar2 -0.4888881 -0.09606208
> confint(mle)
                                  2.5 %       97.5 %
ar1                          0.81348340  1.196124084
ar2                         -0.48806617 -0.094573470
rep(1, length(LakeHuron)) 589.98093574 651.042054268
years                       -0.03744268 -0.005694972
```