Day 15 - Time Series Regression

Introduction

Today we learn how to formally fit a time series regression model by combining a regression model with a serially correlated error process.

Review

The code below creates a time series regression model

```
# data
data("AirPassengers")
ap <- AirPassengers
ap_tbl <- tibble(
  ap = c(ap), year = rep(1949:1960, each = 12),
  month = rep(1:12, 12) %>% factor()
) %>% mutate(t = 1:n(), t2 = t^2)  %>%
  mutate(t_scaled = c(scale(t)), t2_scaled = c(scale(t2))) %>%
  mutate(log_ap = log(ap))
ap_sub_tbl <- ap_tbl %>% filter(year < 1960)

# fit model
ols_fit <- lm(log_ap ~ t + t2 + month, ap_sub_tbl)
```

Generalized least squares overview

Linear model theory

A ________________________________ is a model of the form

$$y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \dots + \beta_p x_{i,p} + \epsilon_i$$

where ϵ_i is ____ distributed $N(0, \sigma^2)$. In matrix notation, the above model is equivalent to

$$y = X\beta + \epsilon$$

where

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}, \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

By properties of normal distributions, the above model is equivalent to

$$y \sim \mathcal{N}(X\beta, \Sigma)$$

where $\Sigma = \sigma^2 \mathcal{I}_n$ and $\mathcal{I}_n$ is an $n \times n$ identity matrix. To fit the model, we must estimate ________________ and ___________________. The estimated regression equation is written as:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i,1} + \hat{\beta}_2 x_{i,2} + \dots + \hat{\beta}_p x_{i,p}$$

Parameter estimates are obtained by minimizing the ___________________________________. The error in regression is called a ______________________,

$$e_i = y_i - \hat{y}_i$$

The ____, which minimizes the ______________, is

$$\hat{\beta} = (X^\top X)^{-1} X^\top y$$

Problems

- Estimating GLS models is *very* hard
- No closed for expression for the standard error of the prediction
- Solutions:
  - Delta method
  - Bootstrapping
  - Use Bayes

State-space model using `arima`