Unit 23: Regression with ARMA Errors

Jeffrey Woo

Department of Statistics, University of Virginia

Spring 2016

Readings for Unit 23

Textbook chapter 5.5.

Last Unit

1. Linear filters to enhance/retain or dampen certain frequencies.

Motivation

One way to explore the relationship between two time series is via regression. One of the assumptions in linear regression is that the errors are independent. With time series, the errors are unlikely to be independent.

Regression Model with AR Errors

In linear regression, the error terms are assumed to be independent.

Question: What is a consequence of this assumption not being met?

We will see how we can adjust the linear regression model to allow for errors with an AR structure.

Regression Model with AR Errors

Suppose $\{y_t\}$ and $\{x_t\}$ are the time series variables. A simple linear regression model with AR errors can be written as

$$y_t = \beta_0 + \beta_1 x_t + \epsilon_t$$

where $\epsilon_t = \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \cdots + w_t$ and $w_t \sim iidN(0, \sigma_w^2)$.

Regression Model with AR Errors

Consider the usual AR polynomial, $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots$, the model for the errors can be written as

$$\phi(B)\epsilon_t = w_t.$$

Thus $\epsilon_t = \phi^{-1}(B) w_t$.

Regression Model with AR Errors

So the regression model with AR errors can be written as

$$y_t = \beta_0 + \beta_1 x_t + \phi^{-1}(B) w_t. \tag{1}$$

This can easily be extended to the multiple linear regression.

Regression Model with AR Errors

From (1), we also have

$$\phi(B)y_t = \phi(B)\beta_0 + \beta_1\phi(B)x_t + w_t.$$

Regression Model with AR Errors

Next, we let

Subbing these equations, our model is now

$$y_t^* = \beta_0^* + \beta_1 x_t^* + w_t. \tag{2}$$

Regression Model with AR Errors

Using model (2), we can adjust our estimates of the coefficients in model (1).

- $\hat{\beta}_1$ from model (2) will be the estimate of β_1 in model (1) (standard error as well).
- Since $\beta_0^* = \phi(B)\beta_0 = (1 - \phi_1 - \cdots \phi_p)\beta_0$, we have $\hat{\beta}_0 = \frac{\hat{\beta}_0^*}{1-\hat{\phi}_1-\cdots-\hat{\phi}_p}$.
- $s.e.(\hat{\beta}_0) = \frac{s.e.(\hat{\beta}_0^*)}{1-\hat{\phi}_1-\cdots-\hat{\phi}_p}$.

Building the Regression Model with AR Errors

The procedure in building the regression model with AR errors is

- Use ordinary least squares (OLS) to estimate $y_t = \beta_0 + \beta_1 x_t + \epsilon_t$.
- Examine the AR structure of the __________________.
- Estimate the coefficients $\phi_1, \cdots, \phi_p$ using ARIMA estimation.
- Use the estimated $\hat{\phi}_1, \cdots, \hat{\phi}_p$ to compute y_t^* and x_t^*.
- Use OLS to estimate $y_t^* = \beta_0^* + \beta_1 x_t^* + w_t$.

Worked Example: Sales

The file "company.txt" contains data for sales of a company. The company wishes to predict its sales by using industry sales as a predictor. The variables *company* and *industry* are the company sales in millions, and industry sales in millions. The data are collected over 20 quarters.

Worked Example: Sales

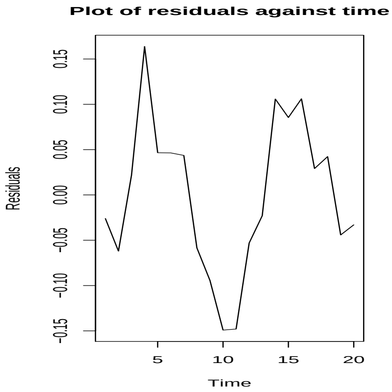
Step 1: Fit OLS and store residuals.

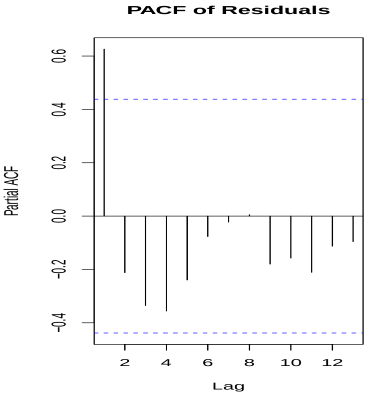
```
> result<-lm(company~industry)
> res<-result$residuals
```

Step 2: Plot time series for residuals and examine for structure. PACF also plotted.

```
> par(mfrow=c(1,2))
> plot(ts(res))
> pacf(res)
```

Worked Example: Sales





Residuals appear to have an AR(1) structure.

Worked Example: Sales

Step 3: Estimate ϕ_1.

```
> ar.1<-arima(res,order = c(1,0,0), include.mean = FALSE)
> phi.1<-ar.1$coef
> phi.1
      ar1
0.6052108
```

$\hat{\phi}_1 = 0.6052.$

Worked Example: Sales

Step 4: Calculate variables to use in the adjustment regression, y_t^* and x_t^*.

```
> y<-cbind(as.ts(company),lag(company,-1))
> ystar<-y[,1] - phi.1*y[,2]
> x<-cbind(as.ts(industry),lag(industry,-1))
> xstar<-x[,1] - phi.1*x[,2]
```

Worked Example: Sales

Step 5: Use OLS for y_t^* on x_t^* to estimate the model.

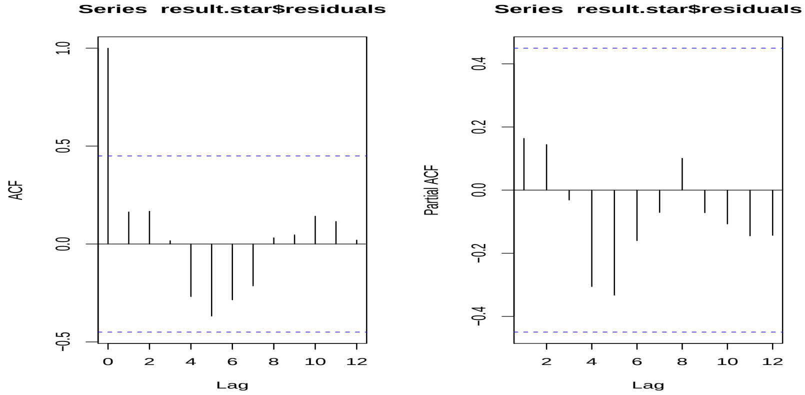
```
> result.star<-lm(ystar~xstar)
> summary(result.star)

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.440022   0.169408  -2.597   0.0188 *
xstar        0.174054   0.002806  62.025   <2e-16 ***

Residual standard error: 0.06738 on 17 degrees of freedom
  (2 observations deleted due to missingness)
Multiple R-squared:  0.9956,    Adjusted R-squared:  0.995
```

Worked Example: Sales

Check structure of residuals.



Worked Example: Sales

So the model is

Worked Example: Sales

Compare with OLS regression

```
> summary(result)

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.454750   0.214146  -6.793 2.31e-06 ***
industry     0.176283   0.001445 122.017  < 2e-16 ***

Residual standard error: 0.08606 on 18 degrees of freedom
Multiple R-squared:  0.9988,    Adjusted R-squared:  0.9987
F-statistic: 1.489e+04 on 1 and 18 DF,  p-value: < 2.2e-16
```

Extension to ARMA Errors

Suppose the errors follow an ARMA process such that

$$\epsilon_t = \phi^{-1}(B)\theta(B)w_t.$$

Then we have

$$\begin{aligned} & y_t = \beta_0 + \beta_1 x_t + \phi^{-1}(B)\theta(B)w_t \\ \Longrightarrow \quad & \phi(B)\theta^{-1}(B)y_t = \phi(B)\theta^{-1}(B)\beta_0 + \beta_1\phi(B)\theta^{-1}(B)x_t + w_t. \end{aligned}$$

Question: How to practically apply this transformation?