# Unit 23: Regression with ARMA Errors

Taylor R. Brown

Department of Statistics, University of Virginia

Fall 2020

# Readings for Unit 23

Textbook chapter 3.8.

# 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.

1. Regression Model with Correlated Errors
2. Regression Model with AR Errors
3. Worked Example

## Regression Model with Correlated Errors

In linear regression, the error terms are assumed to be i.i.d. $N(0, \sigma_w^2)$.

**Question**: What is a consequence if the errors are not independent?

We will see how we can adjust the linear regression model to allow for correlated errors.

# Regression Model with Correlated Errors

Consider the classical regression model:

$$y_t = \sum_{j=1}^{r} \beta_j z_{tj} + x_t \tag{1}$$

where

- $y_t$: response variable at time $t$,
- $z_{tj}$: the $j$th predictor at time $t$,
- $\beta_j$: the $j$th coefficient,
- $x_t$: error term that is Gaussian white noise.

# Regression Model with Correlated Errors

Using matrix and vector notation, (1) can be written as

$$\boldsymbol{y} = \boldsymbol{Z}\boldsymbol{\beta} + \boldsymbol{x} \tag{2}$$

where

- $\boldsymbol{y} = (y_1, \cdots, y_n)'$,
- $\boldsymbol{Z}$ is the $n \times r$ design matrix,
- $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_r)'$,
- $\boldsymbol{x} = (x_1, \cdots, x_n)'$
- $\text{Var}(\boldsymbol{x}) = \{\gamma_x(s, t)\}_{s,t} = \Gamma$

## Regression Model with Correlated Errors

- In (1) and (2), ordinary least squares (OLS) regression is used to estimate the coefficients.
- If $x_t$ are correlated with some covariance $\gamma_x$, then **weighted** least squares (WLS) should be used instead.

## Weighted Least Squares

Let $\boldsymbol{\Gamma}$ denote the covariance matrix for $\boldsymbol{x}$, then multiplying (2) by $\boldsymbol{\Gamma}^{-1/2}$, we obtain

$$\boldsymbol{y}^* = \boldsymbol{Z}^* \boldsymbol{\beta} + \boldsymbol{\delta} \tag{3}$$

where

# Weighted Least Squares

So (3) can be viewed as a classical regression model and solved using "ordinary least squares"':

$$\begin{aligned}\hat{\beta} &= \left[\boldsymbol{Z}^{*'}\boldsymbol{Z}^{*}\right]^{-1}\boldsymbol{Z}^{*'}\boldsymbol{y}^{*}\\ &= \left[\left(\boldsymbol{\Gamma}^{-1/2}\boldsymbol{Z}\right)'\boldsymbol{\Gamma}^{-1/2}\boldsymbol{Z}\right]^{-1}\left(\boldsymbol{\Gamma}^{-1/2}\boldsymbol{Z}\right)'\boldsymbol{\Gamma}^{-1/2}\boldsymbol{y}\\ &= \left[\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Z}\right]^{-1}\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{y}\end{aligned}$$

$$\begin{aligned}\mathrm{Var}\left(\hat{\beta}\right) &= \left[\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Z}\right]^{-1}\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\mathrm{Var}(\boldsymbol{y})\left[\left[\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Z}\right]^{-1}\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\right]'\\ &= \left[\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Z}\right]^{-1}\boldsymbol{Z}'\left[\left[\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Z}\right]^{-1}\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\right]'\\ &= \left[\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Z}\right]^{-1}\end{aligned}$$

# Weighted Least Squares

Using the **covariance structure** of the error terms, we **transform** the variables so that classical methods can still be used.

1. Regression Model with Correlated Errors
2. Regression Model with AR Errors
3. Worked Example

# Regression Model with AR Errors

Consider the regression framework presented in (1) and (2). Instead of $\boldsymbol{x}$ being uncorrelated errors, $\boldsymbol{x}$ is now an AR(p) error where

$$\phi(B)x_t = w_t$$

where $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ and $w_t$ is Gaussian white noise.

# Regression Model with AR Errors

Multiplying both sides of (1) by $\phi(B)$, we obtain

$$\phi(B)y_t = \sum_{j=1}^{r} \beta_j \phi(B) z_{tj} + \phi(B) x_t. \tag{4}$$

Note the last term in (4) is now Gaussian and **no longer an AR(p)**.

# Regression Model with AR Errors

Letting $y_t^* = \phi(B)y_t$ and $z_{tj}^* = \phi(B)z_{tj}$ in (4), we obtain

$$y_t^* = \sum_{j=1}^{r} \beta_j z_{tj}^* + w_t. \tag{5}$$

# Regression Model with AR Errors

We are back to working in the regression framework, after transforming the response variable and predictors using $y_t^* = \phi(B)y_t$ and $z_{tj}^* = \phi(B)z_{tj}$. For example, if $p = 1$, then

# Building the Regression Model with AR Errors: The Cochrane/Orcutt procedure

The CO procedure in building the regression model with AR errors is

1. Use ordinary least squares (OLS) to estimate (1).
2. Examine the AR structure of the **sample residuals** from step 1: $\hat{x}_t = y_t - \sum_{j=1}^{r} \hat{\beta}_j z_{tj}$.
3. Estimate the coefficients $\phi_1, \cdots, \phi_p$ using ARIMA estimation.
4. Use the estimated $\hat{\phi}_1, \cdots, \hat{\phi}_p$ to compute $y_t^* = \hat{\phi}(B) y_t$ and $z_{tj}^* = \hat{\phi}(B) z_{tj}$.
5. Use OLS to estimate (5).

If the steps above are done properly, the residuals at the end of step 5 should be **white noise**.

# Building the Regression Model with AR Errors: Our Procedure

Our procedure is slightly different:

1. Use ordinary least squares (OLS) to estimate (1).
2. Examine the AR structure of the **sample residuals** from step 1: $\hat{x}_t = y_t - \sum_{j=1}^{r} \hat{\beta}_j z_{tj}$.
3. Figure out which $p$ would be suitable to model these residuals but don't estimate the coefficients $\phi_1, \cdots, \phi_p$.
4. Use MLE to estimate the overall model's parameters: $\beta_1, \ldots, \beta_r$, $\sigma^2$, and $\phi_1, \ldots, \phi_p$.

If the steps above are done properly, the residuals at the end of step 4 should be **white noise**. We can also try several models, and compare their information criteria.

## 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.

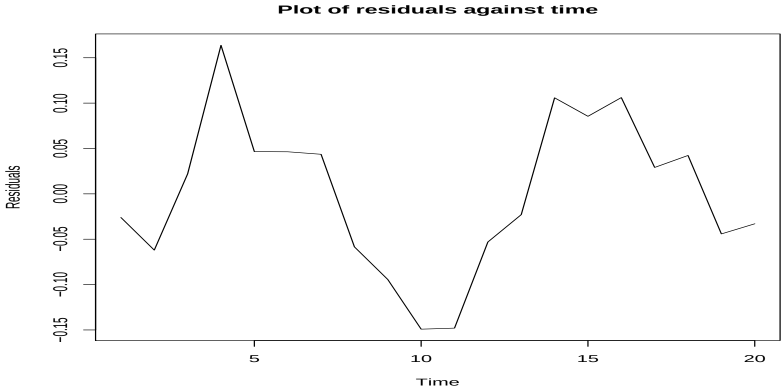
```
> result<-lm(company~industry)
> res<-result$residuals
```

# Worked Example: Sales

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

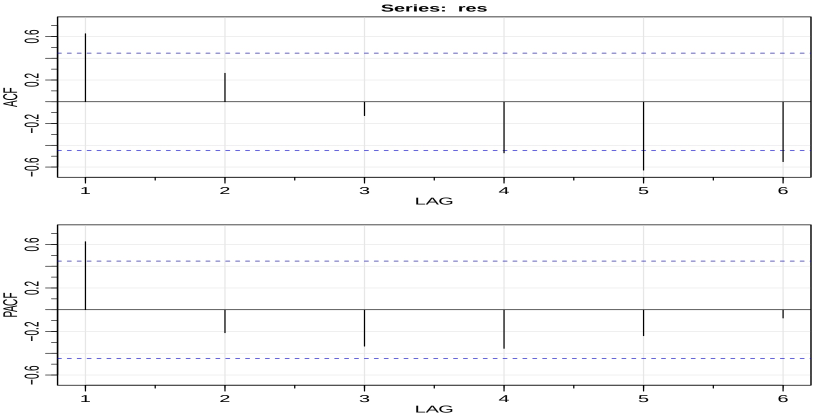
```
> plot(ts(res))
> acf2(res)
```

# Worked Example: Sales




Residuals appear to be curved (keep in mind). A quadratic term for the predictor should be considered.

# Worked Example: Sales




Residuals appear to have an AR(1) structure.

## Worked Example: Sales

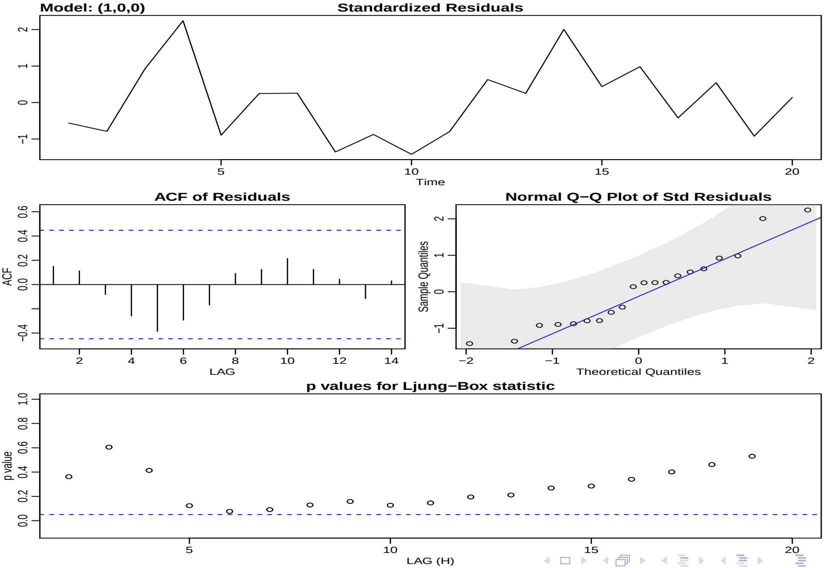
`sarima()` function in R from `astsa` package performs maximum likelihood estimation, but you have to pick which model to estimate:

```
> sarima(company, 1, 0, 0, xreg=cbind(industry))
```

# Worked Example: Sales

| | Estimate | SE | t.value | p.value |
|---|---|---|---|---|
| ar1 | 0.6295 | 0.1772 | 3.5531 | 0.0024 |
| intercept | -1.2876 | 0.3523 | -3.6549 | 0.0020 |
| industry | 0.1751 | 0.0024 | 73.3658 | 0.0000 |

# Worked Example: Sales



# Worked Example: Sales

So the model is

## Worked Example: Sales

Should try a model with a quadratic term for the predictor. Turns out the quadratic term is insignificant.

# Extension to ARMA Errors

Suppose the errors follow an ARMA process such that

$$\phi(B)x_t = \theta(B)w_t.$$

**Question**: How do we transform (1)?