6.6: Regression with ARIMA errors

Taylor

University of Virginia

Motivation

These are useful when you want to run a regression, but your errors are correlated.

Motivation

Consider the model

$$Y_t = x_t'\beta + W_t$$

$$\phi(B)W_t = \theta(B)Z_t, \quad \{Z_t\} \sim WN(0, \sigma^2)$$

when $t = 1, \ldots, n$

The linear regression you're used to assumes $W_t \sim WN(0, \sigma^2)$.

Ordinary Least Squares (the old way)

Let's write this in matrix form: $\mathbf{Y} = \mathbf{X}\beta + \mathbf{W}$. When $\text{Var}(W) = \sigma^2 I$, we have

$$(\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta) = \sum_{t=1}^{n}(y_t - x_t'\beta)^2.$$

Minimizing this gives you the **ordinary least squares (OLS)** estimates of β.

1. $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$
2. $E\hat{\beta} = \beta$
3. $\text{Var}(\hat{\beta}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$

Generalized Least Squares (the new way)

$\mathbf{Y} = \mathbf{X}\beta + \mathbf{W}$. When $\text{Var}(W) = \Gamma$, our loss function is

$$(\mathbf{Y} - \mathbf{X}\beta)'\Gamma^{-1}(\mathbf{Y} - \mathbf{X}\beta).$$

Minimizing this gives you the **generalized least squares (GLS)** estimates of β.

1. $\hat{\beta} = (\mathbf{X}'\Gamma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Gamma^{-1}\mathbf{y}$
2. $E\hat{\beta} = \beta$
3. $\text{Var}(\hat{\beta}) = (\mathbf{X}'\Gamma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Gamma^{-1}\mathbf{X}(\mathbf{X}'\Gamma^{-1}\mathbf{X})^{-1} = (\mathbf{X}'\Gamma^{-1}\mathbf{X})^{-1}$

Problem: Γ is a function of ϕ and θ, and so is unknown. The $\hat{\beta}$ depends on ϕ and θ, which means it's usually unavailable.

Generalized Least Squares (the new way)

If $\mathbf{Y} = \mathbf{X}\beta + \mathbf{W}$, $\text{Var}(W) = \Gamma$, then let $T'T = \sigma^2\Gamma^{-1}$. Then

$$\tilde{\mathbf{Y}} = T'\mathbf{Y} = T'\mathbf{X}\beta + T'\mathbf{W} = \tilde{\mathbf{X}}\beta + \mathbf{e}$$

and we can use OLS (homework question).

Cochran and Orcutt: ultiplying our data by T' is similar to applying $\phi(B)$ to our data if W_t is AR(p). We don't explore this further.

Generalized Least Squares (the new way) MLE

Now pretend that we don't know the parameters ϕ and θ. We look at a more complicated loss function.

We take negative twice the log of the following:

$$L(\beta, \phi, \theta, \sigma^2) = (2\pi)^{-n/2}(\det \Gamma)^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{Y} - \mathbf{X}\beta)'\Gamma^{-1}(\mathbf{Y} - \mathbf{X}\beta)\right\}$$

See how the previous procedure was a special case of this?

Generalized Least Squares (the new way) MLE

- Example 6.6.1 fits the model: $Y_t = \beta + W_t$ where W_t is an MA(1) process.
- Example 6.6.2 fits the model $Y_t = \beta_0 + \beta_1 t + W_t$ where W_t is AR(2).
- Example 6.6.3 fits the model $Y_t = \beta_0 + \beta_1 1(t \geq 99) + W_t$ where W_t is MA(12)

Example 6.6.3

If $Y_t = \beta_0 + \beta_1 1(t \geq 99) + W_t$ then

$$Y_t - Y_{t-12} = \beta_1[1(t \geq 99) - 1(t - 12 \geq 99)] + N_t$$

where $N_t = W_t - W_{t-12}$. So let's difference our data and fit the model:

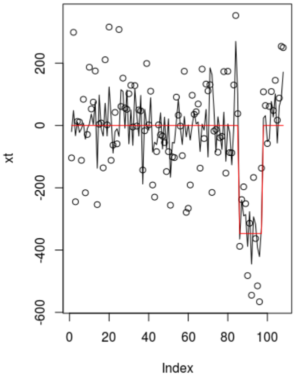
model

$$X_t = \beta g_t + N_t$$

where $X_t = Y_t - Y_{t-12}$ and
$g_t = 1(t \geq 99) - 1(t \geq 99 + 12) = 1(99 \leq t \leq 110)$.

Example 6.6.3

$X_t = \beta g_t + N_t$
Because we difference with lag 12, our dataset starts at $t = 13$, meaning we only have $120 - 12 = 108$ datapoints. See 6.6.R.



Red lines are the naive regression fitted values.
The black lines are the better model's fitted values.