

9.3 Autocorrelation

MBC 638

Data Analysis and Decision Making

Regression on Time Series Data

Regression on Time Series Data

- Modeling trend and seasonal components may not generate random residuals

Regression on Time Series Data

- Modeling trend and seasonal components may not generate random residuals
- Residual plots help assess the fit of a regression line

Autocorrelation: Definition

Autocorrelation: Definition

- Relationships between neighboring points

Autocorrelation: Definition

- Relationships between neighboring points
 - E.g., January data affects February, February affects March, etc.

Autocorrelation: Definition

- Relationships between neighboring points
 - E.g., January data affects February, February affects March, etc.
 - Can cause lack of randomness in data

Autocorrelation: Definition

- Relationships between neighboring points
 - E.g., January data affects February, February affects March, etc.
 - Can cause lack of randomness in data
- **Autocorrelation:** correlation between successive values

Autocorrelation: Why Do We Care?

Autocorrelation: Why Do We Care?

- Can't use regression model if data violates assumption of independent residuals

Autocorrelation: Why Do We Care?

- Can't use regression model if data violates assumption of independent residuals
- Autocorrelation in residuals indicates opportunity to improve fit

Autocorrelation: Why Do We Care?

- Can't use regression model if data violates assumption of independent residuals
- Autocorrelation in residuals indicates opportunity to improve fit
 - Add elements to model to increase predictive power

Autocorrelation: How Can We Tell?

Autocorrelation: How Can We Tell?

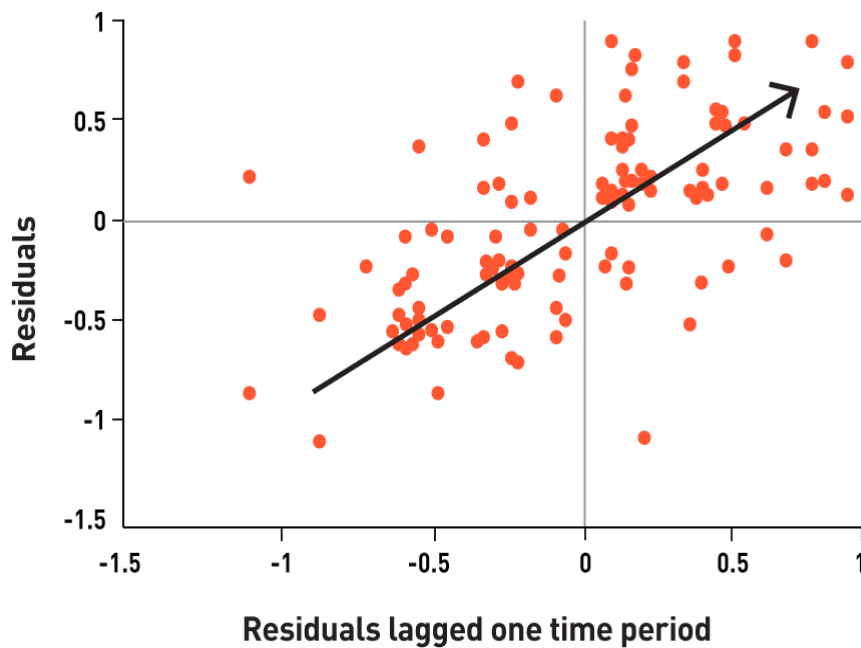
- Test residuals by lagging, moving one time period
 - Residual = $e = y_{\text{actual}} - y_{\text{predicted}}$

Autocorrelation: How Can We Tell?

- Test residuals by lagging, moving one time period
 - Residual = $e = y_{\text{actual}} - y_{\text{predicted}}$
 - Lagged residual plot = $(e_1, e_2), (e_2, e_3), (e_3, e_4)$
 $\dots(e_{n-1}, e_n)$

Autocorrelation: How Can We Tell?

- Test residuals by lagging, moving one time period
 - Residual = $e = y_{\text{actual}} - y_{\text{predicted}}$
 - Lagged residual plot = $(e_1, e_2), (e_2, e_3), (e_3, e_4)$
 $\dots(e_{n-1}, e_n)$
- Plot residuals, look for pattern



Autocorrelation: How Do We Improve Our Model?

Autocorrelation: How Do We Improve Our Model?

- Add other x inputs to model

Autocorrelation: How Do We Improve Our Model?

- Add other x inputs to model
 - Past values of time series

Autocorrelation: How Do We Improve Our Model?

- Add other x inputs to model
 - Past values of time series
 - Lagged values of residuals

Summary

- We manipulate the y as input variable

Summary

- We manipulate the y as input variable
- Work with limited existing data to better predict future through manipulation

Summary

- We manipulate the y as input variable
- Work with limited existing data to better predict future through manipulation
- Look for autocorrelation, which indicates manipulation is required

Summary

- We manipulate the y as input variable
- Work with limited existing data to better predict future through manipulation
- Look for autocorrelation, which indicates manipulation is required
 - Autocorrelation: relationship between neighboring points