

MGMT 675

AI-ASSISTED FINANCIAL ANALYSIS



RICE | BUSINESS
Jones Graduate School of Business

AUTOCORRELATION

- Correlation of a variable with its own lagged value
- First order autocorrelation = first lag
- Second order autocorrelation = second lag

EXAMPLE

Time	Variable
1	2.1
2	4.5
3	5.3
4	3.2
5	1.6

EXAMPLE

Time	Variable	Lag
1	2.1	
2	4.5	2.1
3	5.3	4.5
4	3.2	5.3
5	1.6	3.2

EXAMPLE

Time	Variable	Lag
1	2.1	
2	4.5	2.1
3	5.3	4.5
4	3.2	5.3
5	1.6	3.2

Correlation of a variable with its lag is first-order autocorrelation.

SECOND-ORDER AUTOCORRELATION

Time	Variable	2nd Lag
1	2.1	
2	4.5	
3	5.3	2.1
4	3.2	4.5
5	1.6	5.3

Correlation of a variable with its 2nd lag is second-order autocorrelation.

AUTOCORRELATION FUNCTION

- The autocorrelation function (acf) shows autocorrelation at multiple lags.
- Lag 0 is usually presented too, even though the correlation at lag 0 is always 100%.
- The band is a confidence interval under the null hypothesis that the autocorrelation is zero.
- Estimates outside the band are statistically significant.

FAMA-FRENCH FACTORS

- Ask Julius to use pandas datareader to get the monthly Fama-French factors for the maximum history available.
- Ask Julius to plot the acf's of Mkt-RF, HML, and SMB.

AUTOREGRESSIONS

FIRST-ORDER AUTOREGRESSION

- A first-order autoregression is the equation

$$x_t = \alpha + \beta x_{t-1} + \varepsilon_t$$

- Positive $\beta \Leftrightarrow$ positive first-order autocorrelation.
- Positive beta \Rightarrow momentum.
- Negative beta \Rightarrow reversals.

HIGHER ORDER AUTOREGRESSIONS

- A p th order autoregression is the equation

$$x_t = \alpha + \beta_1 x_{t-1} + \cdots + \beta_p x_{t-p} + \varepsilon_t$$

- There are standard methods for choosing the optimal p , trading off goodness of fit and parsimony.
- Ask Julius to fit an AR(p) to HML and find the optimal p .

LEVELS OR CHANGES?

WHAT SHOULD WE TRY TO FORECAST?

- Price of AAPL?
- Change in price of AAPL?
- Percent change in price of AAPL (return)?
- A forecast of changes or percent changes implies a forecast of the price and vice versa, so only need to forecast one of them directly.
- What variable should we use in an autoregression?

STATIONARITY

- One issue is that autocorrelation or autoregression estimates are reliable only for stationary variables.
- Prices grow over time (unstationary).
- Changes in prices become larger over time (in absolute value)
- Returns are the right thing to use in an autoregression.

OTHER EXAMPLES

- Autoregression for crude oil price or change in crude price or percent change in crude price?
- Autoregression for interest rate or change in interest rate or percent change in interest rate?

AUTOREGRESSION AND Δx

Rearrange

$$x_t = \alpha + \beta x_{t-1} + \varepsilon_t$$

as

$$\Delta x_t = \alpha + (\beta - 1)x_{t-1} + \varepsilon_t$$

$$\Delta x_t = (\beta - 1) \left(x_{t-1} - \frac{\alpha}{1 - \beta} \right) + \varepsilon_t$$

MEAN REVERSION

- An AR(1) is

$$\Delta x_t = (\beta - 1) \left(x_{t-1} - \frac{\alpha}{1 - \beta} \right) + \varepsilon_t$$

- $\beta < 1$ implies regression towards the mean.
- The mean is $\alpha / (1 - \beta)$.
- $|\beta - 1|$ is called the rate of mean reversion.
- $\beta > 1$ implies nonstationary.

EXAMPLE

Ask Julius to simulate the process $x_t = 1 + 0.5 * x_{t-1} + e_t$ by drawing 1,000 standard normals for e_t starting at $x_0=1$. Ask Julius to plot the process.

FORECASTS

- Given today's value x_t , the AR(1) forecast for the next period's value is

$$\hat{x}_{t+1} = \alpha + \beta x_t$$

- The forecast for the period after that is

$$\hat{x}_{t+2} = \alpha + \beta \hat{x}_{t+1}$$

- Etc.

FORECAST CONVERGENCE

- The forecasts will converge to $\alpha/(1 - \beta)$ as we look further out into the future.
- They will converge quickly if β is small.

$$\beta \text{ small} \Rightarrow 1 - \beta \text{ large}$$

INTEREST RATES

- Ask Julius to use pandas-datareader to download 10-year Treasury yields starting in 1980 from FRED.
- Ask Julius to fit an AR(1) and to use it to forecast the yield for the next 20 days.

VECTOR AUTOREGRESSION

- Multiple variables (vector)
- One equation for each variable. Lags of all variables are used as predictors, including the variable's own lag.
- Can answer questions like: does this month's SMB return forecast next month's HML return?

EXAMPLE

- Ask Julius to use pandas datareader to download the monthly 5 Fama-French factors since 1970 from French's data library.
- Ask Julius to fit a VAR(1) to MKT-RF, SMB, HML, RMW, and CMA and to print the summary to a text file.