Time Series Analysis

Modeling Heteroskedasticity

Nicoleta Serban, Ph.D.
Professor
Stewart School of Industrial and Systems Engineering

Basic Concepts

Georgia Tech

About This Lesson

Independent VS Uncorrelated data

If $Y_1, \ldots, Y_t$ are:

1. <u>Independent</u>: $f(Y_1), \ldots, f(Y_T)$ are also independent for any non-linear transformation f.
2. <u>Uncorrelated</u>: $cov(Y_t, Y_{t-h})=0$ for $h \neq 0$ but $f(Y_1), \ldots, f(Y_T)$ are NOT necessarily independent for any non-linear transformation f; therefore, if higher-order dependent $cov(Y_t, Y_{t-h})=0$ for $h \neq 0$

Example of lack of independence in uncorrelated data:

- $Y_1, \ldots, Y_t$ are uncorrelated, i.e. $cov(Y_t, Y_{t-h})=0$ but $cov\left(Y_t^2, Y_{t-h}^2\right) \neq 0$
- This will be a characteristic of many of the time series when modeling heteroskedasticity

Heteroskedasticity in Time Series

Given a time series $\{Y_t, t = 1, \ldots, T\}$:

- The *conditional variance or volatility* is $\mathbb{V}(Y_t | Y_{t-1}, \ldots, Y_1)$.
- The *unconditional variance* is $\mathbb{V}(Y_t)$.

<u>Independent data</u>: $\mathbb{V}(Y_t | Y_{t-1}, \ldots, Y_1) = \mathbb{V}(Y_t)$

<u>Uncorrelated data</u>: $\mathbb{V}(Y_t | Y_{t-1}, \ldots, Y_1) \neq \mathbb{V}(Y_t)$

- When $\mathbb{V}(Y_t | Y_{t-1}, \ldots, Y_1)$ depends on time $\Rightarrow$ *Heteroskedasticity* in the time series
- Modeling Heteroskedasticity: Model conditional variance for weakly stationary time series

Diagnosis: Heteroskedasticity in Time Series

Diagnostics for independent and uncorrelated data:

1. **Independent:** ACF of the time series $Y_1, \ldots, Y_t$ as well as of its non-linear transformations $f(Y_1), \ldots, f(Y_T)$ resemble the ACF of white noise.
2. **Uncorrelated but Dependent:** ACF of the time series $Y_1, \ldots, Y_t$ resemble the ACF of white noise but the ACF of at least one non-linear transformation of the time series (e.g. $Y_1^2, \ldots, Y_T^2$) does not resemble the ACF of white noise.

Modeling Heteroskedasticity

The class of heteroskedastic models is a generalization of the models so far:

- ARIMA models: estimation and inference on the conditional expectation: $E(Y_t|Y_{t-1},\ldots,Y_1)$
- Heteroskedasticity models: estimation and inference on the conditional variance: $\mathbb{V}(Y_t|Y_{t-1},\ldots,Y_1)$
- General model: $Y_t = \mathrm{f}(Y_{t-1},\ldots,Y_1)+\sigma(Y_{t-1},\ldots,Y_1)\,\varepsilon_t$

Example: Modeling stock returns or other financial indicators that display volatility.

Model Structure

In a generic ARMA model, the model is described by $\phi(B)Y_t = \theta(B)Z_t$ with

$$\mathrm{E}[Z_t] = 0, \qquad \mathrm{E}[Z_t^2] = \sigma^2, \qquad \mathrm{E}[Z_t Z_r] = 0, \qquad \text{for } r \neq t.$$

- Under the ARMA modeling the unconditional variance of Z_t is assumed constant. However, in most real data studies, the conditional variance of Z_t could change with time.
- We can rewrite the error term as

$$Z_t = \sigma_t R_t \text{ with } \mathrm{E}[R_t] = 0,\ \mathrm{E}[R_t^2] = 1$$

where $\sigma_t = \sigma_{t|t-1,t-2,\ldots,1}$ can be a deterministic and R_t is a sequence of iid random variables.

Characteristics of Conditional Variance

Some common characteristics of the function σ_t, commonly referred in financial literature as *volatility* are:

1. The function σ_t may be high for certain time periods and low for other periods - clusters of variability
2. The function σ_t varies with time in a continuous manner (i.e. there are no discontinuities, anomalies in the volatility)
3. The function σ_t varies within a specific range.
4. The function σ_t commonly is assumed to have the so called *leverage property*, i.e. it reacts differently to large amplitude changes from the small amplitude changes in the time process.

Nonparametric Estimation of the Conditional Variance

1. Fit ARIMA model and obtain estimates of the coefficients of the polynomials $\phi(z)$ and $\theta(z)$
2. Approximate Z_t by the residuals of the ARIMA model
3. Since $Z_t = \sigma_t R_t$ take the log of the squared residuals: $\log(Z_t^2) = \log(\sigma_t^2) + \varepsilon_t$ where the error term is $\varepsilon_t = \log(R_t^2)$
4. Fit a nonparametric regression to estimate $\log(\sigma_t^2)$
5. Re-scale back by taking the exponential to obtain an estimate for σ_t^2

Georgia Tech

Why important to model heteroskedasticity?

- If we can effectively forecast volatility then we will be able to price more accurately, create more sophisticated risk management tools, or come up with new strategies that take advantage of the volatility;
- **Application in finance**
  - Options Pricing - The Black-Scholes model for options prices is dependent upon the volatility of the underlying instrument
  - Tradeable Securities - Volatility can now be traded directly by the introduction of the CBOE Volatility Index (VIX)
- **Other Applications**
  - Volatility $\approx$ Risk in management and investment
  - Volatility $\approx$ Uncertainty in decision making

Summary