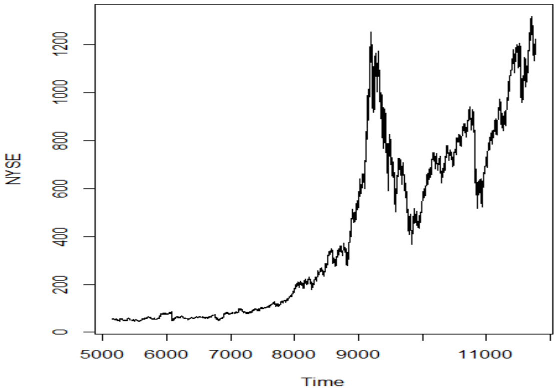
NYSE data Analysis

Shuoli Chen

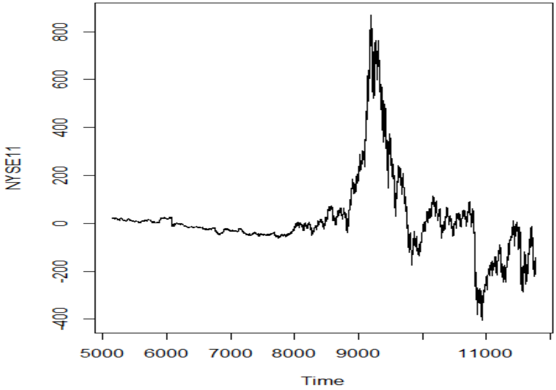
Institute of Statistics & Big Data
Renmin University of China

- First we plot the data with timing diagram:



- Intuitively, the data is not stationary, not to mention white noise.

- Make the data be stationary with curve-fittin. We use $e^{a_0+a_1x+a_2x_2}$ to fit the data, and get the coefficients (3.482e+00, 6.837e-04, -1.793e-08).
- We plot the stationary timing diagram:



- We conduct Ljung-Box white noise test to the transformed data and the p-value is 2.2e-16, so we can see that the transformed data is significantly not white noise.
- we check if the transformed data is stationary with adf test, the p-value is 0.5184, so the transformed data is stationary
- Obviously there is ARCH effect in the data. First, we can routinely build a linear time series model, for example the ARMA model. Then we consider to construct an appropriate ARCH model for the residue.

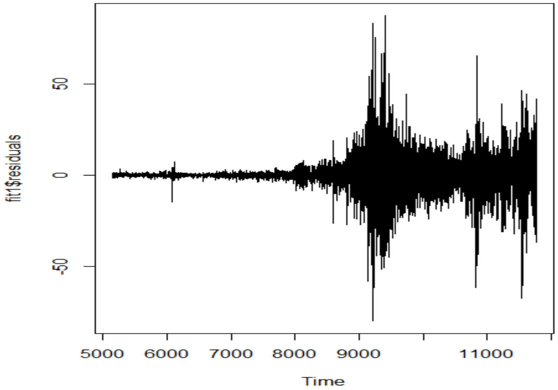
- Build an ARMA model.
  The notation ARMA (p, q) refers to the model with p autoregressive terms and q moving-average terms.
- This model contains the $AR(p)$ and $MA(q)$ models

$$X_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}$$

  μ is the expectation of X_t (often assumed to equal 0), and the $\{\epsilon_t\}$ are white noise error terms.

- ARMA(5,4) is the best model according to the corresponding AIC values. The coefficients are $\psi = (-0.48, -0.58, 0.605, 0.495), \theta = (0.952, 1.491, 2.05, 1.43, 0.945), c = 14.71.$
- Box-Ljung test shows the residual of ARMA(5,4) model is white noise, significance test shows the coefficients of ARMA(5,4) are significant.

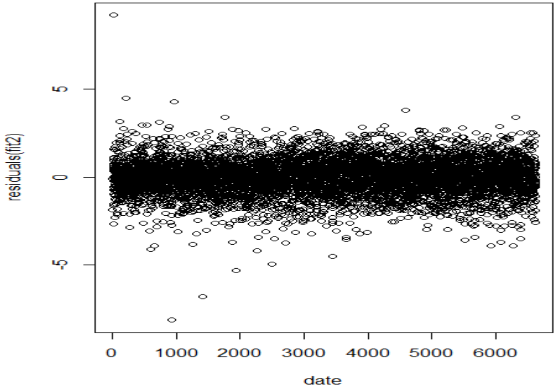
- We plot the ARMA residuals with timing diagram:



- Obviously we can observe the ARCH effect, so we need to build a model for the residuals of ARMA(5,4).

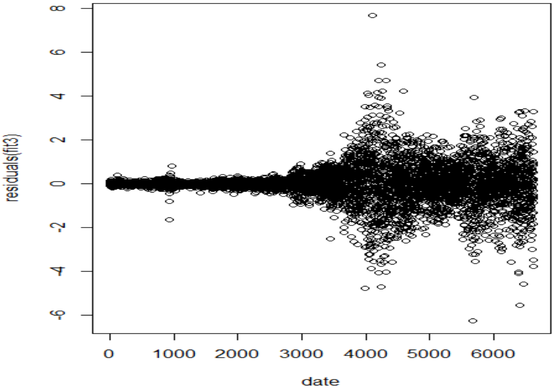
- ARCH model:
  Assume $\epsilon_t = \sigma_t z_t$, where $z_t \overset{iid}{\sim} N(0,1)$, The model of σ_t is
  $\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_p \varepsilon_{t-p}^2.$
- If we apply the idea of ARMA to model σ_t^2, we get GARCH model:
  $\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_q \varepsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_p \sigma_{t-p}^2.$
- We consider GARCH model. In fact, GARCH(1,1) is simple and efficient.
  The cofficients of GARCH(1,1) are
  $\alpha_0 = 0.007149, \alpha_1 = 0.099418, \beta_1 = 0.907365.$

- Box-Ljung test shows the residual of ARMA(5,4) model is white noise, significance test shows the coefficients of ARMA(5,4) are significant.
- We plot the GARCH residuals:



- According to the figure, we can see that there is no ARCH effect in GARCH residual. Hence GARCH(1,1) model is predictive to the risk of stock market. With GARCH(1,1) we can predict when high risk happens and then avoid it.

- In fact, it is not every GARCH(p,q) that works well, for example, the residuals of GARCH(2,2) model is



- It is obvious that there is ARCH effect in residual of GARCH(2,2), hence GARCH(2,2) can not help us to avoid high risk in stock exchange.