Financial Time Series

Lecture 5

Volatility modeling

- Volatility is the conditional standard deviation in an asset return

- Important in risk management and option pricing

- We denote volatility at time t by σ_t

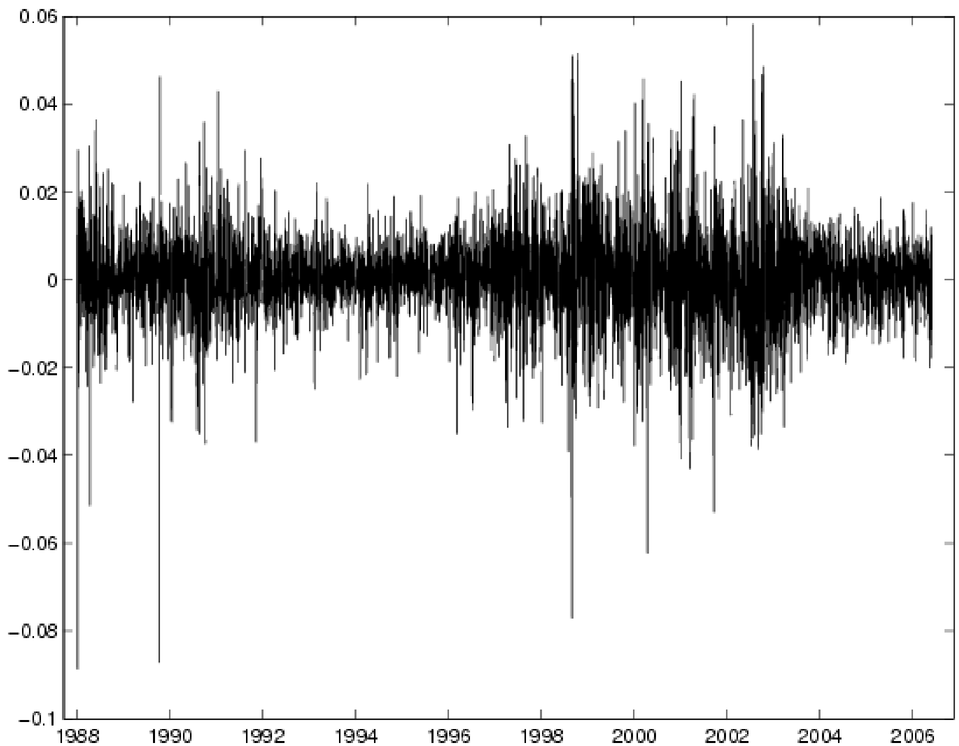
Historic and implied volatility

- By historic volatility we mean volatility estimates derived from time series of returns

- Implied volatility is a volatility estimate given by equating model (typically B-S) option prices and market option prices

Clustering

- Volatility is non-constant over time and typically comes in clusters



In the B-S

- In the original and famous Black-Scholes model for option pricing, volatitity is assumed to be constant over time since asset prices are modeled by

$$P_t = P_0 exp\{\mu t + \sigma W_t\}$$

where, $\{W_t\}$ is standard Wiener

Adjsuting the B-S

- Of course one-may adjust the model to include non-constant volatility but option pricing will be messy as there typically will be an infinite number of risk-neutral measures instead of one unique risk-neutral measure

- An infinite number of risk-neutral measures implies that there are infinitely many possible prices...

Implied volatility

- In the B-S framework, we get the price of a European call option as

$$C(P_t, \sigma_t, T, K) = \Phi(d_1)P_t - \Phi(d_2)Ke^{-r(T-t)}$$

$$d_1 = \frac{ln(P_t/K) + (r + {\sigma_t}^2/2)(T - t)}{\sigma_t\sqrt{T - t}}$$

$$d_2 = d_1 - \sigma_t\sqrt{T - t}$$

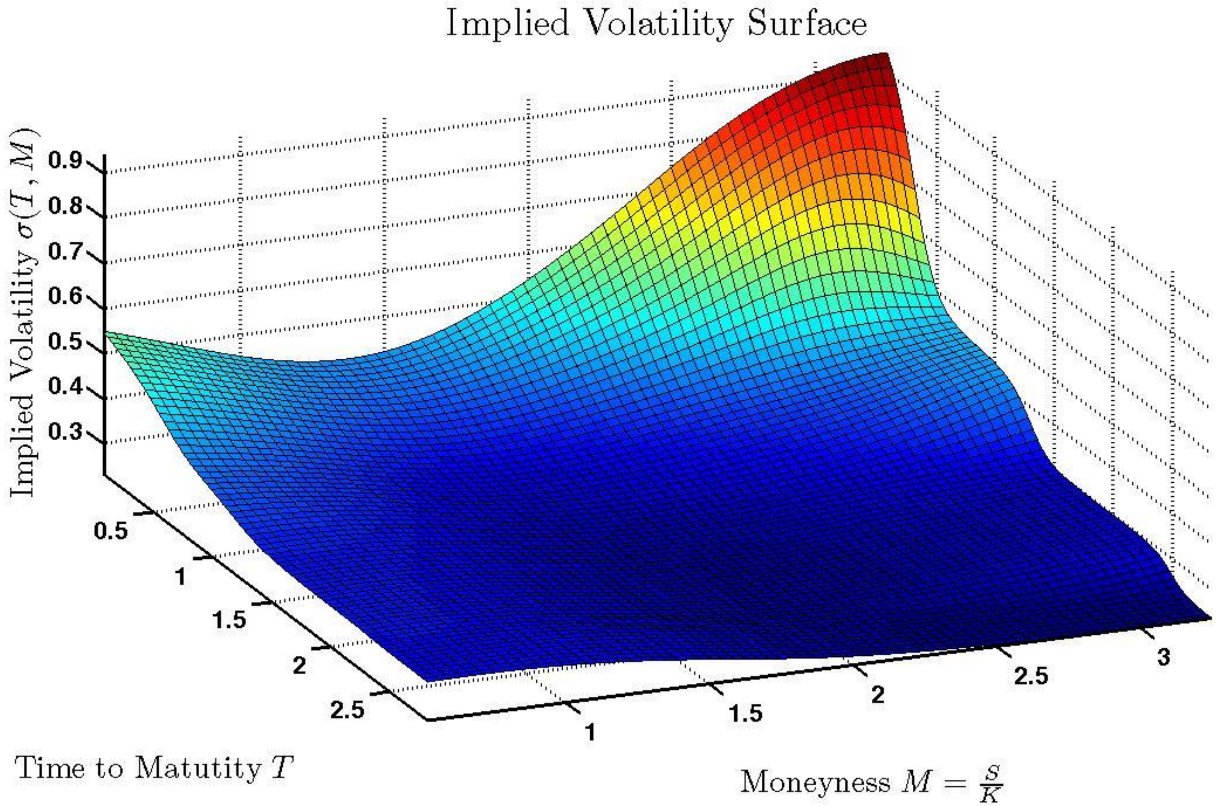
Implied volatility

- For a given market price C of the option, knowing the price of the underlying asset, P_t, the risk-free interest rate, r, the strike price, K and the maturity time, T, we may obtain σ_t from numerical methods

- The obtained value of σ_t is the volatility implied by the market, i.e., the implied volatility

Characteristic feature of implied volatility

- Smile ☺



VIX

- One of the most famous volatility indices is the VIX, which gives estimates on the 30-day volatility for the S&P 500
- It’s considered ”the benchmark” for stock market volatility
- Sometimes referred to as ”the fear index”
- It is a model free or non-parametric approach in the sense that it does not assume anything about distributions of data

VIX

- To compute the VIX one uses

$$\hat{\sigma}_t{}^2 = \frac{2}{T}\sum_i \frac{\Delta K_i}{K_i{}^2} e^{rT} Q(K_i) - \frac{1}{T}\left(\frac{F}{K_0} - 1\right)^2$$

where $VIX = 100\hat{\sigma}_t$

VIX components

- T is the time to maturity expressed as a year-based percentage

- F is the forward index level (given from the put-call parity using call and put options on the S&P 500 that are close in price)

- K_0 is the first strike price below F

VIX components

- K_i is the strike price of the ith out-of-the-money option, which means that if $K_i < K_0$ the option is a call, that if $K_i > K_0$ the option is a put, and both call and put if "$=$"
- $\Delta K_i = (K_{i+1} - K_{i-1})/2$
- r is the risk-free interest rate
- $Q(K_i)$ is the mid-point of the bid-ask spread for each option with strike price K_i

For more on the VIX

- ”A tale of two indices” Carr and Wu (2006)

- *”VIX white paper”*

  www.cboe.com/micro/vix/vixwhite.pdf

How do we measure historic volatility?

- Let $r_1, \dots, r_T$ be observed returns

- Crude (variance window):

$$\hat{\sigma}_t{}^2 = \frac{1}{\tau - 1} \sum_{i=t-\tau}^{t} (r_i - \bar{r}_t)^2$$

where $\bar{r}_t = \frac{1}{\tau} \sum_{i=t-\tau}^{t} r_i$

Realized volatility

- The volatility estimate given by the variance window is sometimes referred to as realized volatility
- May be used as a benchmark for other volatility models once one has the decided on a time frame
- The width of the time frame when modeling volatility is related to the availability of data, if intra-day data is available the time frame may be one day, if only daily data is available, the time frame is typically 30 days.

Conditional heteroskedasticity

- As volatility often comes in clusters, we want models that can capture this behaviour

- We want the distribution of returns to depend on past values and volatility levels

- We let F_t denote the information generated up to and at time t

Conditional heteroskedasticity

- For de-meaned (log) returns $r_1, \ldots, r_T$ without serial correlation, we assume that

$$r_t = \sigma_t z_t$$

where $\{z_t\}$ is WN with variance one

- By conditional heteroskedasticity, we mean, if z_t is Gaussian, that $r_t | F_{t-1} \sim N(0, \sigma_t)$, where F_{t-1} is the information available at time $t - 1$

Conditional heteroskedasticity

- In an ARCH(p)-model we assume that

$$\sigma_t{}^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{t-i}{}^2$$

- A special case is the EWMA where $\alpha_i = \lambda^i$ for $i \geq 1$ and often $\lambda = 0.94$, as suggested by RiskMetrics

ARCH

- We see that, using an ARCH model, high volatility is like the be followed by high volatility, i.e. clustering effects

ARCH

- To identify ARCH behavior we consider the multivariate linear regression

$$r_t{}^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{t-i}{}^2 + e_t$$

for $t = p + 1, \ldots, T$

Test for ARCH effects

- The null hypothesis is $\alpha_1 = \cdots = \alpha_p = 0$
- Let $SSR_0 = \sum_{t=p+1}^{T} \left({r_t}^2 - \overline{{r_T}^2} \right)^2$, where $\overline{{r_T}^2} = \frac{1}{T} \sum_{t=1}^{T} {r_t}^2$
- Let $SSR_1 = \sum_{t=p+1}^{T} {\hat{e}_t}^2$ where $\hat{e}_t$ is the residual from the least squares estimation of the regression model

The test statistic

- The statistic for the test of ARCH effects is

$$F = \frac{(SSR_0 - SSR_1)/p}{SSR_1/(T - 2p - 1)}$$

Which is asymptotically ${\chi^2}_p$

- Available in matlab econometrics toolbox as ”archtest”

Parameter estimation

- Parameter estimation is often made using ML
- For Gaussian noise, the log-likelihood function is given by

$$l(r_{p+1}, \dots, r_T | \alpha_0, \dots, \alpha_p, r_1, \dots, r_p) \\ = -\frac{1}{2} \sum_{t=p+1}^{T} \left(ln({\sigma_t}^2) + \frac{{r_t}^2}{{\sigma_t}^2} \right)$$

Parameter estimation

- The log likelihood may be maximized in, e.g., matlab, using fminsearch for $-l$

- One may also use other noise distributions, such as t or a generalized error distribution

- If one uses a t distribution the degrees of fredoom may be specified á priori or estimated in the ML routine

Properties of ARCH-models

- The shed some light on the properties of the ARCH-models we consider the ARCH(1)
- For the returns, we have that $r_t = \sigma_t z_t$
- For the volatility, we have that

$$\sigma_t{}^2 = \alpha_0 + \alpha_1 r_{t-1}{}^2$$

Properties of ARCH-models

- Letting F_{t-1} be the information set available at time $t-1$, the unconditional mean of r_t is

$$E[r_t] = E\big[E[r_t|F_{t-1}]\big] = E[\sigma_t E[z_t]] = 0$$

- The unconditional variance is

$$\begin{aligned} Var(r_t) &= E[{r_t}^2] = E\big[E[{r_t}^2|F_{t-1}]\big] \\ &= E[\alpha_0 + \alpha_1 {r_{t-1}}^2] = \alpha_0 + \alpha_1 E[{r_{t-1}}^2] \end{aligned}$$

Properties of ARCH-models

- Since $\{r_t\}$ is stationary, we get

$$Var(r_t) = \frac{\alpha_0}{1 - \alpha_1}$$

- Note that we must have $0 \leq \alpha_1 < 1$

Properties of ARCH-models

- One may also show that the unconditional kurtosis is

$$K(r_t) = 3\frac{1-{\alpha_1}^2}{1-3{\alpha_1}^2} > 3$$

- This means that return distributions are leptokurtic
- Also, we must have that ${\alpha_1}^2 < 1/3$

Conditional heteroskedasticity

- In the famous GARCH(1,1) the evolution of the volatility is governed by

$$\sigma_t{}^2 = \omega + \alpha r_{t-1}{}^2 + \beta \sigma_{t-1}{}^2$$

- May be considered as an ARCH(∞)

GARCH

- It is possible to fit a GARCH(p,q) but it turns out that in many applications a GARCH(1,1) is sufficient

- As for the ARCH, we assume that

$$r_t = \sigma_t z_t$$

where is WN, typically $N(0,1)$ or t with degrees of freedom between 3 and 6

GARCH properties

- The GARCH(1,1) is (weakly) stationary with $Cov(r_s, r_t) = 0$ for $s \neq t$ iff $\alpha + \beta < 1$ (proof in Bollerslev 1986)

- The $2m$-th unconditional moments of r_t exist iff

$$\sum_{j=0}^{m} \binom{m}{j} a_j \alpha^j \beta^{m-j} < 1$$

where $a_0 = 1, \;\; a_j = \prod_{i=1}^{j}(2j-1), \;\; j = 1, \ldots$

Given existence

- The unconditional mean of r_t is zero (same proof as for ARCH)

- The unconditional variance of r_t is (same proof as for ARCH)

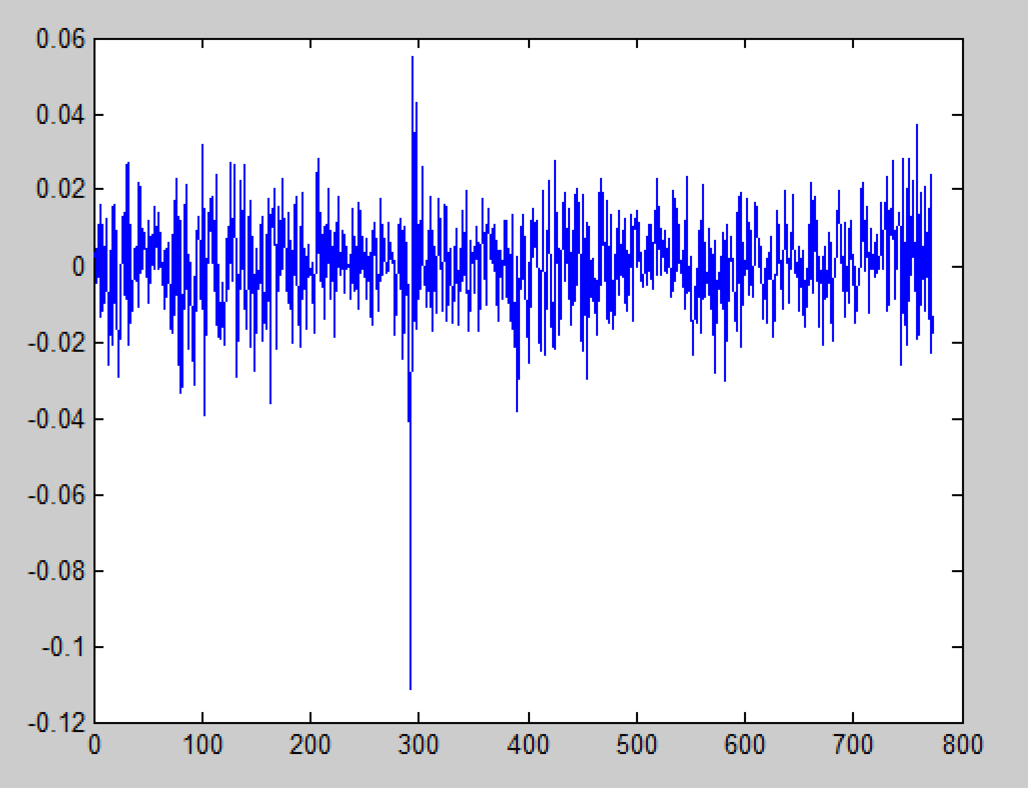
$$\frac{\omega}{1-\alpha-\beta}$$

- The unconditional kurtosis is

$$\frac{3(1-(\alpha+\beta)^2)}{1-\beta^2-2\alpha\beta-3\alpha^2} > 3$$

Example, using garchfit in matlab

- Returns from N225 (note ”Tsunami/Fukushima extreme event”)



Example, using garchfit in matlab

- Is it ”correct” to fit a model to this data and to use it to predict values ”now”?

- To what extent does the extreme event affect the parameter estimates?

- We will assume $N(0,1)$ noise and use garchfit in matlab which is an ML based method

Example, using garchfit in matlab

- If we use the whole data set ”as is”, we get

$$\hat{\sigma}_t{}^2 = 1.63 \cdot 10^{-5} + 0.14 r_{t-1}{}^2 + 0.76 \hat{\sigma}_{t-1}{}^2$$

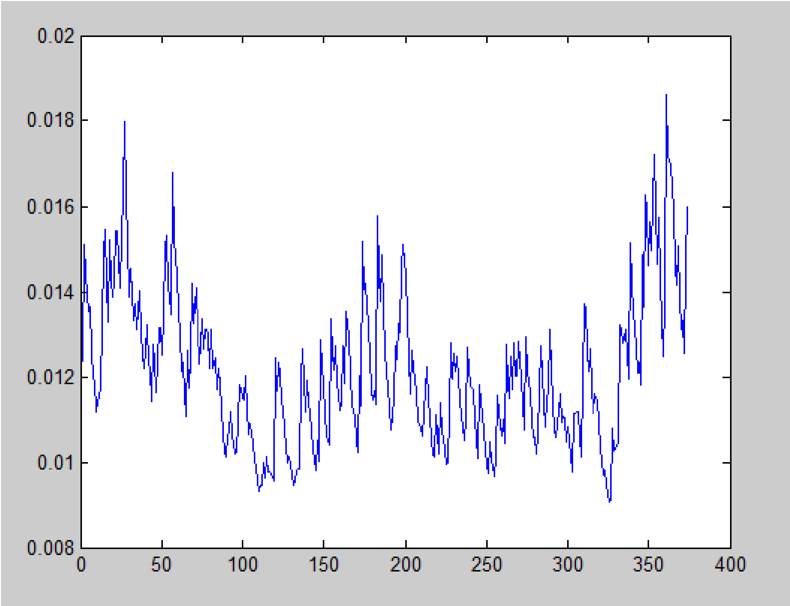
- If we instead use data from obs 400 (which is after the extreme event) and forward, we get

$$\hat{\sigma}_t{}^2 = 8.00 \cdot 10^{-4} + 0.044 r_{t-1}{}^2 + 0.93 \hat{\sigma}_{t-1}{}^2$$

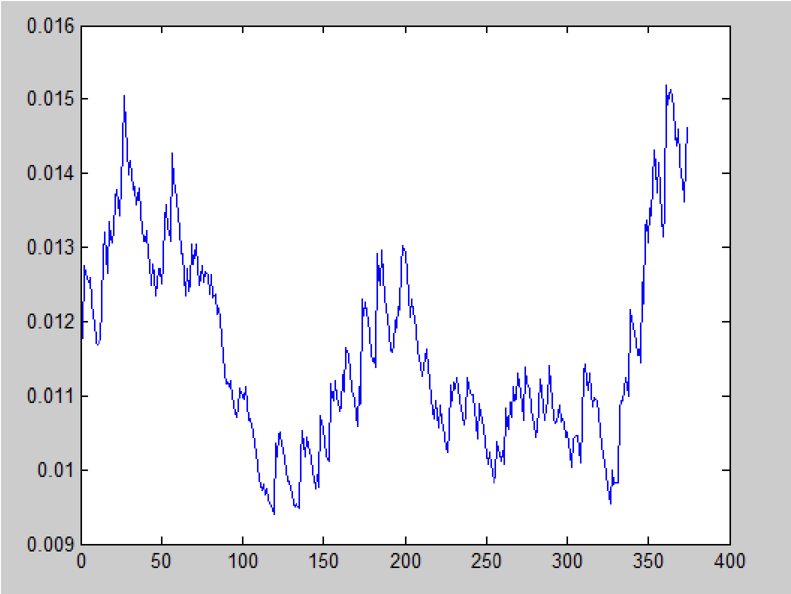
Example, using garchfit in matlab

- So, we see that the extreme event from more than two years ago greatly affects the parameter estimates
- We may use the output from garchfit for further analysis

Volatility fits with/without the extreme event



”With”, jumpy, nervous

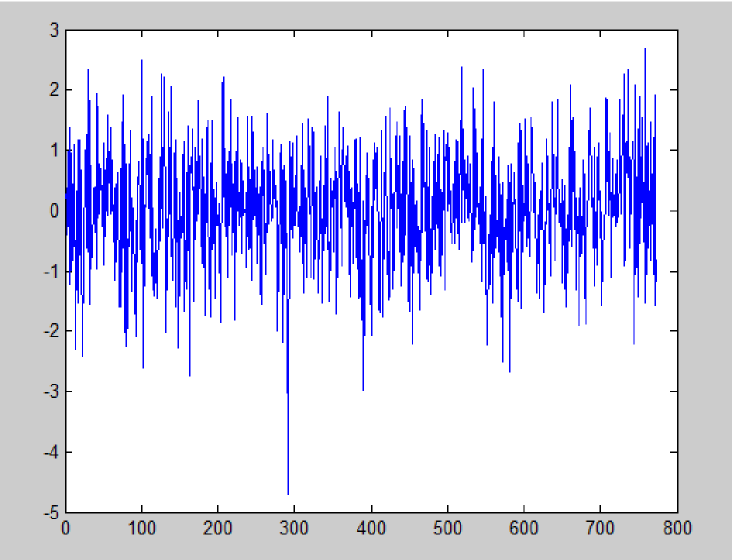


”Without”, more calm

Devolatized returns

- If the model fit is ok, we want devolatized returns $\hat{z}_t = r_t/\hat{\sigma}_t$ to act like white noise

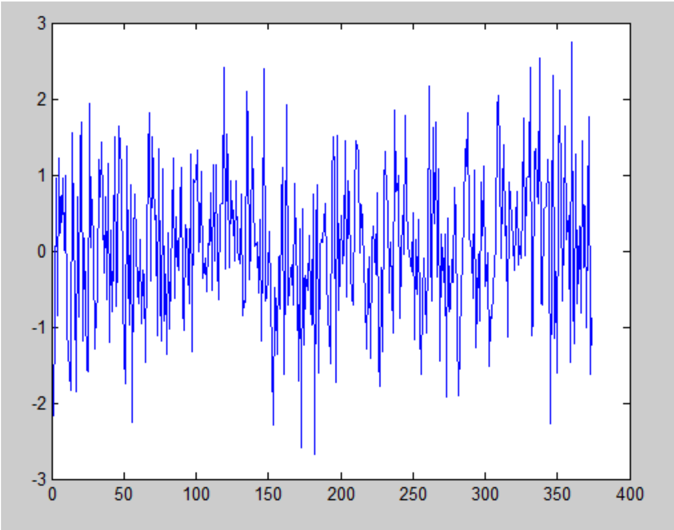
- For the series N225 with the extreme event, we get



Devolatized returns

- We see that the extreme event is ”still extreme”

- For the series without the extreme event, we get



Devolatized returns

- We may use Ljung-Box, lbqtest in matlab

- Gives p-value 0.4497, which is satisfactory