Volatility Modeling

Phat Aphiwatanakoon

20 September 2024

Outline

Stylized Facts for Financial Returns

- The distribution of returns is not normal, but it
  - is approximately symmetric,
  - has fat tails, and
  - has a high peak;
- There is almost no correlation between returns for different days;
- **There is a positive dependence between absolute returns on nearby days, and likewise for squared returns.**

Source: Taylor (2011)

Stylized Facts for Financial Volatility

- **Volatility Clustering:** Volatility tends to cluster over time, i.e., periods of high volatility tend to be followed by more periods of high volatility, and periods of low volatility tend to be followed by more periods of low volatility.
- **Heavy Tails:** Return empirical distributions have fatter tails than what is predicted by the normal distribution. Extreme events occur more frequently than would be expected under a normal distribution.
- **Leverage Effect:** Volatility tends to increase when asset prices fall and decrease when asset prices rise.
- Further extensions also include mean reversion (Stochastic volatility models, e.g. Heston) and long memory (FIGARCH).

Models Overview

Stylized Facts	Hist-orical	ARCH -type	GBM	PJD	SVM
Vol. Clus.	✗	✓	✗	✗	✓
Heavy Tail	✗	(✓)	✗	✓	✓
Lev. Eff.	✗	(✓)	✗	✗	✓

Simulation is best with SDE-type models, i.e., GBM (Gaussian Brownian Motion), PJD (Poisson Jump Diffusion), and SVMs (Stochastic Volatility Models).

Prediction Methods Based on Historical Volatility

- Historical Average (all available data)

$$\tilde{\sigma}^2_{t+1} = \frac{1}{t}\sum_{j=1}^{t} \hat{\sigma}^2_j$$

- Simple Moving Average (last m single-period sample estimates)

$$\tilde{\sigma}^2_{t+1} = \frac{1}{m}\sum_{j=0}^{m-1} \hat{\sigma}^2_{t-j}$$

- Exponential Moving Average (all available data)

$$\tilde{\sigma}^2_{t+1} = (1-\beta)\hat{\sigma}^2_t + \beta\tilde{\sigma}^2_t, \quad 0 \le \beta \le 1$$

- Exponential Weighted Moving Average (last m)

$$\tilde{\sigma}^2_{t+1} = \frac{\sum_{j=0}^{m-1}\left(\beta^j \hat{\sigma}^2_{t-j}\right)}{\sum_{j=0}^{m-1}\beta^j}$$

Source: MIT 18.S096 Dr. Kempthorne Fall 2013

ARCH: The Model

Let $y_t = \log(S_t/S_{t-1})$. The standard ARCH model is given by:

$$
\begin{aligned}
y_t &= \mu + \epsilon_t \\
\epsilon_t &= \sigma_t z_t \\
z_t &\sim i.i.d.N(0,1)
\end{aligned}
$$

where

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-2}^2 + \cdots + \alpha_p \epsilon_{t-p}^2$$

$$\alpha_i \geq 0$$

with the conditional heteroscedasticity being

$$\sigma_t^2 = VAR(y_t | \mathcal{F}_{t-1})$$

ARCH: More about ARCH

- Note the constant return. We typically just approximate with zero in applications.
- ARCH(p) implies an AR(p) model in ϵ_t^2
- The first equation is called the mean model. Call be other models such as ARIMA or one with exogenous variables.

ARCH: Estimation

The joint probability density function for the observed series is

$$\mathcal{L}(\theta|y) = \prod_{t=1}^{n} f(y_t|\mathcal{F}_{t-1};\theta)$$

where f is

$$f(y_t|\mathcal{F}_{t-1};\theta) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{(y_t-\mu)^2}{2\sigma_t^2}\right)$$

The loglikelihood function is

$$\log \mathcal{L}(\theta|y) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{n}\left(\log \sigma_t^2 + \frac{(y_t-\mu)^2}{\sigma_t^2}\right)$$

GARCH: The Model

Instead of ARCH's

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

we have

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

where

$$\alpha_i \geq 0, \beta_j \geq 0$$

GARCH: More ARCH-type Models

Exponential GARCH (EGARCH)

- Asymmetric effects of shocks on volatility. Also ensure positive variance.

$$\log(\sigma_t^2) = \omega + \sum_{i=1}^{q} \beta_i \log(\sigma_{t-i}^2) + \sum_{i=1}^{p} \alpha_i \left(\frac{|\epsilon_{t-i}|}{\sigma_{t-i}} - \mathbb{E}\left[\frac{|\epsilon_{t-i}|}{\sigma_{t-i}} \right] \right) + \sum_{i=1}^{p} \gamma_i \frac{\epsilon_{t-i}}{\sigma_{t-i}}$$

Glosten-Jagannathan-Runkle GARCH (GJR-GARCH)

- Indicator function distinguishes between positive and negative shocks. Leverage effect.

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2 + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \gamma_i \epsilon_{t-i}^2 \mathbb{I}(\epsilon_{t-i} < 0)$$

GBM: The Model

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

where

- S_t is the asset price at time t,
- dS_t is the infinitesimal change in S_t,
- μ is the drift rate (representing the expected rate of return of the asset),
- σ is the volatility (which is constant),
- W_t is a Wiener process (standard Brownian motion), which satisfies
    - Normal increments: $W_{t'} - W_t \sim N(0, t' - t)$
    - Independent disjoint increments

GBM: Estimation

Brief maths. The SDE is

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

Ito's Lemma, with $X_t = \log S_t$ yields

$$dX_t = \left(\mu - \sigma^2/2\right) dt + \sigma dW_t$$

We observe

$$y_t = \Delta X_t = \left(\mu - \sigma^2/2\right) \Delta t + \sigma \epsilon_t \sqrt{\Delta t}$$

$$y_t \sim N\left(\left(\mu - \sigma^2/2\right) \Delta t, \sigma^2 \Delta t\right)$$

GBM: Estimation

The likelihood of observing a sequence of log returns under the normal distribution, given the parameters, can be written as:

$$\mathcal{L}(\mu, \sigma^2) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma^2\Delta t}} \exp - \frac{(y_t - (\mu - \sigma^2/2)\Delta t)^2}{2\sigma^2\Delta t}$$

We typically maximize the log-likelihood instead, which is

$$\log \mathcal{L}(\mu, \sigma^2) = -\frac{T}{2}\log(2\pi\sigma^2\Delta t) - \frac{1}{2\sigma^2\Delta t}\sum_{t=1}^{T}(y_t - (\mu - \sigma^2/2)\Delta t)^2$$

GBM: Implementation

In passing, note that the MLE estimators have closed-form formulas, i.e.

$$\hat{\sigma}^2 = \frac{1}{\Delta t} \frac{1}{T-1} \sum_{t=1}^{T} \left(y_t - \frac{1}{T} \sum_{t=1}^{T} y_t \right)^2$$

$$\hat{\mu} = \frac{1}{\Delta t} \frac{1}{T} \sum_{t=1}^{T} y_t + \frac{\hat{\sigma}^2}{2}$$

Familiar? What does this mean?

Now, let's implement in Python!

PJD: The Model

GBM with an extra component.

$$dS_t = \mu S_t dt + \sigma S_t dW_t + J S_t dN_t$$

where

- N_t is a Poisson process with intensity λ, modeling the jump process,
- $J \sim N(\alpha, \delta^2)$ is the jump size, following a normal distribution

One can also augment with a correlation between dW_t and the jump size.

PJD: Estimation

The likelihood is formed by considering that jumps occur with probability $\exp \lambda dt$:
The likelihood function is:

$$\mathcal{L}(\mu, \sigma^2, \lambda, \gamma, \delta) = \prod_{t=1}^{T} (P(\text{no jump}) f_{\text{diffusion}}(y_t) \\ + P(\text{jump}) f_{\text{jump}}(y_t))$$

where

- $P(\text{no jump}) = e^{-\lambda \Delta t}$,
- $P(\text{jump}) = 1 - P(\text{no jump}) = 1 - e^{-\lambda \Delta t}$,
- $\log \mathcal{L}_{\text{diffusion}}$ we consider $y_t \sim N\left(\left(\mu - \sigma^2/2\right) \Delta t, \sigma^2 \Delta t\right)$
- $\log \mathcal{L}_{\text{jump}}$ we consider $y_t \sim N\left(\left(\mu - \sigma^2/2\right) \Delta t + \alpha, \sigma^2 \Delta t + \delta^2\right)$

Going forward

Heston Model Volatility is now time-varying:

$$dS_t = \mu S_t \, dt + \sqrt{v_t} S_t \, dW_t^S$$

The variance follows a mean-reverting square-root process (Cox-Ingersoll-Ross (CIR) process):

$$dv_t = \kappa(\theta - v_t) \, dt + \sigma_v \sqrt{v_t} \, dW_t^v$$

Correlation between price and variance processes:

$$dW_t^S \cdot dW_t^v = \rho \, dt$$

- Joint Likelihood. Complexity from correlation.
- Nonlinearity in the variance process.
- Volatility of volatility.
- No closed-form solutions.