Volatility and Correlation Modeling

This lecture agenda

- Models for time-varying volatility
- Models for time-varying correlations

Modeling returns distribution

- Volatilities (and correlations) are key ingredients for calculating parametric VaR.
- If the returns process was truly i.i.d. normal our job would be very easy.
- First, calculate the sample standard deviation:

$$\widehat{\sigma} = \sqrt{\frac{1}{N-1}\sum_{t=1}^{N}\left(r_t - \frac{1}{N}\sum_{t=1}^{N} r_t\right)^2}$$

- This is our estimate of the *daily* volatility.
- VaR calculation is straightforward:

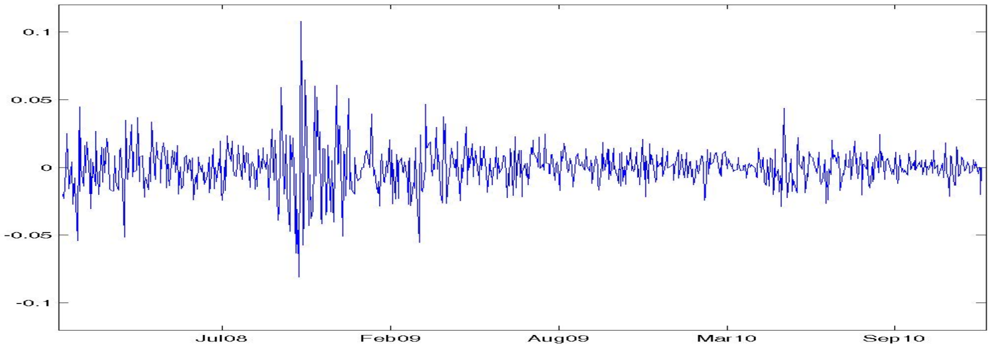
$$VaR(1-p;T) = W\alpha\widehat{\sigma}\sqrt{T}$$

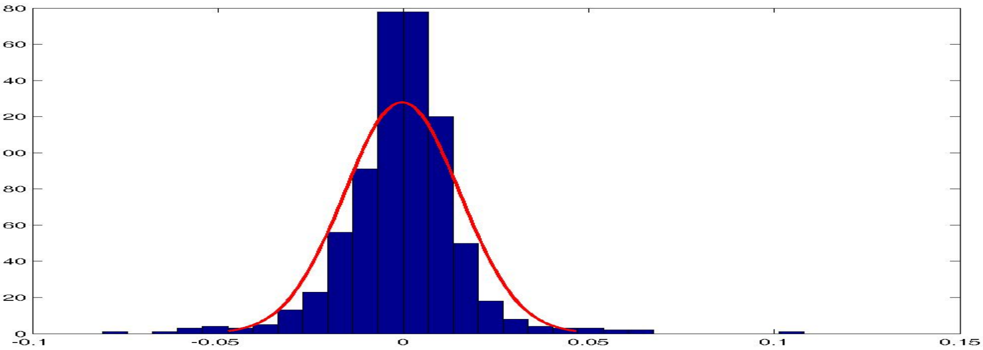
where T is the number of working days

Time-variation in risk

- There is mounting evidence that both volatilities and correlations are time-varying.
- This fact characterizes virtually all financial series, such as individual stocks, aggregate stock markets, currencies, interest rates, commodities.
- We need to move beyond the simple i.i.d. world and develop some new tools.
- Long-horizon volatility forecasts do not obtain by simply multiplying by $\sqrt{T}$!
- Volatility at high frequencies tends to cluster: periods of high (low) volatility tend to be followed by periods of high (low) volatility.

SMI during 2008–2010





Univariate distribution: fat tails

Various explanations to fatter tails:

1. Returns are i.i.d., but their *unconditional* distribution is not normal (say, it is a t distribution)
2. The *conditional* distribution is indeed normal, but with time-varying conditional volatility. Thus, the unconditional distribution is the effect of a mixture of normals with different volatility, hence the "outliers"
3. A combination of the two: the conditional distribution changes over time, but is not normal.

We look at the most successful models in the spirit of approach #2

Part I

Models for time-varying volatility

Rolling window or MA

- The crudest way of measuring time-variation in volatility is by looking at the moving average (MA) of squared returns over a window of length M:

$$\sigma_t = \sqrt{\sum_{i=1}^{M} \frac{1}{M} r_{t-i}^2}$$

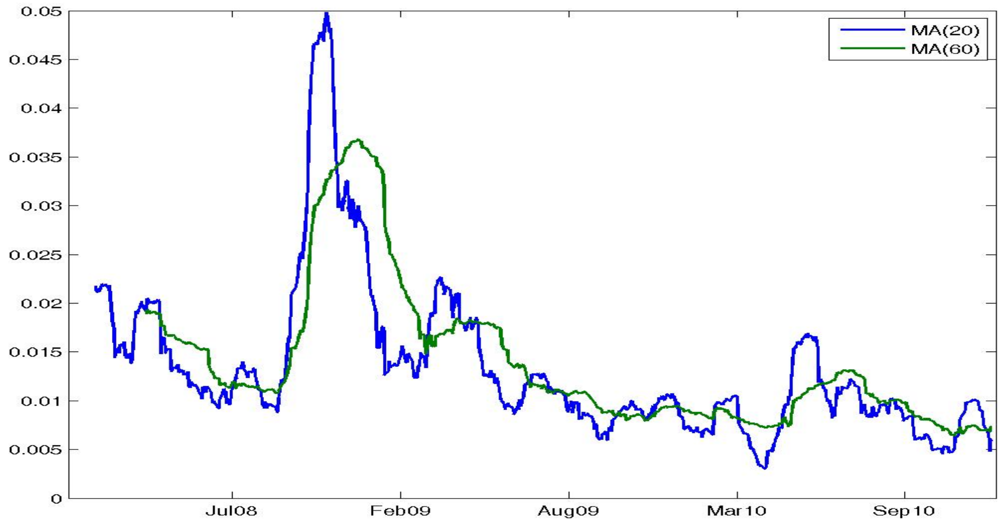
- Two differences from the formula above:
  - The mean effect is ignored ($\mu = 0$) as it is of second-order effect for daily data.
  - The denominator is M and not $M - 1$.

  Both of these adjustments make life easier at very little cost.

- IMPORTANT notation to keep in mind henceforth: σ_t is the conditional volatility for day t. It is calculated up to closing (or midnight) of day $t - 1$ and refers to the change in value during day t.

- The choice of M is somewhat arbitrary. Say, $M = 20$ or $M = 60$.

SMI during 2008–2010, MA(20) and MA(60)



Different weights

- Idea: give more recent observations a higher weight than older data.
- Weight differently the observations

$$\sigma_t^2 = \sum_{i=1}^{M} \alpha_i r_{t-i}^2 \quad \text{where} \quad \alpha_i < \alpha_j \;\; i > j \quad \text{and} \quad \sum_{i=1}^{M} \alpha_i = 1 \quad (*)$$

- Engle(1982) - 2003 Nobel Prize in Economics - suggested putting some weight on a long-run average variance V_L toward which the current variance reverts:

$$\sigma_t^2 = \gamma V_L + \sum_{i=1}^{M} \alpha_i r_{t-i}^2 \quad \text{where } \gamma + \sum_{i=1}^{M} \alpha_i = 1$$

- Define $\omega = \gamma V_L$ and obtain the ARCH(M) model:

$$\sigma_t^2 = \omega + \sum_{i=1}^{M} \alpha_i r_{t-i}^2 \quad (**)$$

- A popular extension of $(*)$ is RiskMetrics EWMA, a popular extension of $(**)$ is the class of GARCH models.

EWMA model – I

- The exponentially weighted moving average (EWMA) model is a simple yet powerful weighting scheme for the variance σ_t^2.
- Start the series from σ_0^2. This can be r_0^2, or historical variance.
- Then, revise the variance estimate based on current information as:

$$\boxed{\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1-\lambda)r_{t-1}^2 \qquad 0 < \lambda < 1}$$

- Substituting backward, one sees that the equation above becomes a special case of $(*)$ where $\alpha_i = (1-\lambda)\lambda^{i-1}$.
- λ governs the decay factor of the weights: $\alpha_{i+1} = \lambda\alpha_i$. Higher values of λ imply more persistent, smooth σ_t^2; lower λ produces volatile σ_t^2.
- Riskmetrics (1994) uses an EWMA model with $\lambda = 0.94$ for daily volatilities. Data more than 100 days old are essentially given no weight.
- Pros: simplicity, few data to store, easy estimation (λ). Cons: no mean-reverting behavior.

GARCH

- GARCH stands for Generalized AutoRegressive Conditional Heteroskedasticity models. Proposed by Bollerslev in 1986.
- *Generalized*: with respect to the ARCH. *AutoRegressive*: the variance depends on its lagged value. *Conditional Heteroskedasticity*: models time-variation in conditional variance.
- The GARCH family has then grown exponentially yielding a flurry of variants and more sophisticated models - TARCH, EGARCH, GARCH-M, QGARCH, AVGARCH, NGARCH, ...
- Yet, the basic GARCH(1,1) remains a useful, parsimonious, and hard-to-beat benchmark for the modeling of most financial time-series.
- In a GARCH(1,1) model, the conditional variance depends on a long-run mean, on its own lagged value, and on the most recent squared innovation:

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

The GARCH(1,1) Model - I

- In equation
$$\sigma_t^2 = \gamma V_L + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2$$
we want $\gamma + \alpha + \beta = 1$.
- γ pulls the variance back to its long-term mean; α defines the weight to the most recent observation; β governs the persistence of the variance process.
- As before, set $\omega = \gamma V_L$ to obtain
$$\boxed{\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2}$$
If we have estimates of (α, β) we can recover γ from the constraint, or $\gamma = 1 - \alpha - \beta$.
- The implied long-run mean is then
$$V_L = \frac{\omega}{1 - \alpha - \beta}$$
For $V_L > 0$, we must have $\alpha + \beta < 1$. This is a stable GARCH(1,1) process. Empirically, that's not always the case!

The GARCH(1,1) Model - Estimation

- Think of $r_t = \mu + u_t$, where $u_t \sim N(0, \sigma_t^2)$ and σ_t^2 follows a GARCH(1,1) process.
- Issue: we need the best estimates of $\Theta = (\alpha, \beta, \omega)$.
- The model is estimated via maximum likelihood. We find the values of Θ such that is maximized the probability that the data were drawn by a process with these characteristics.
- Joint probability:
$$\mathcal{L}(\omega, \alpha, \beta | r) = \Pi_{t=1}^{N} \phi(r_t | \sigma_t^2) = \Pi_{t=1}^{N} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{(r_t - \mu)^2}{2\sigma_t^2}\right)$$
where $\phi(.)$ is the normal pdf.
- From a practical point of view, it is easier to maximize the natural logarithm of the likelihood:
$$\arg\max_{\Theta} \sum_{t=1}^{N} -0.5 \left(\ln(2\pi) + \ln(\sigma_t^2) + r_t^2 / \sigma_t^2\right)$$

The GARCH(1,1) Model - Likelihood

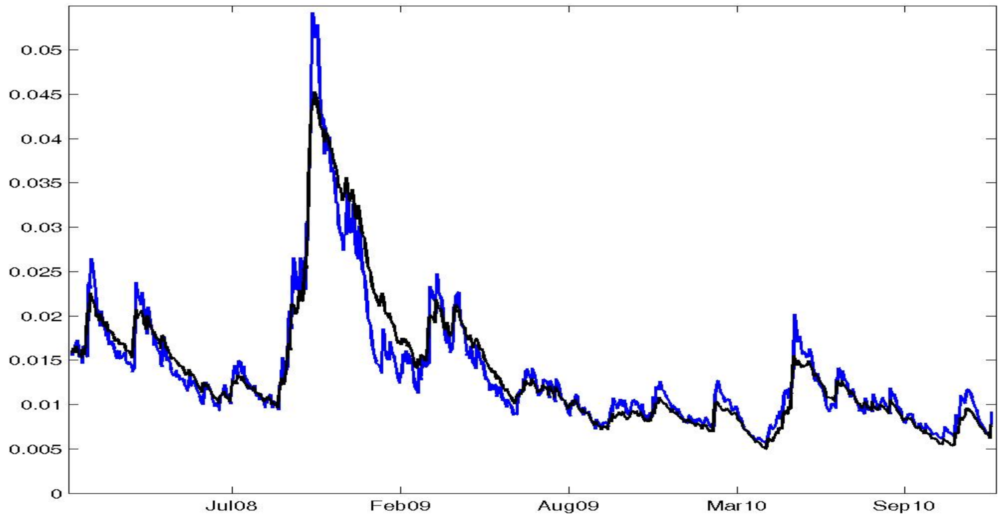
The estimation strategy in Excel proceeds as follows:

- create the time-series of conditional variances σ_t^2 using the GARCH formula and some initial (ω, α, β)
- calculate the log-likelihood for each observations
- calculate the sum of these individual likelihoods
- use Excel SOLVER to maximize it with respect to (ω, α, β)
- use the estimates of (ω, α, β) to create the predictions of future variance (volatility) to be used in the VaR.

GARCH(1,1) Model – Variance Forecasts

- After fitting a GARCH, we can construct forecasts of the expected return variance
- "If my best (today's) estimate of tomorrow's variance is σ_t^2, what is my best estimate for τ-day ahead variance, under the assumption the data truly follow a GARCH process?"
$$E[\sigma_{t+\tau}^2] = V_L + (\alpha+\beta)^\tau [\sigma_t^2 - V_L]$$
- For VaR Computation: "If my best (today's) estimate of tomorrow's variance is σ_t^2, what is my best estimate for the variance of T-day ahead (cumulative) return, under the assumption the data truly follow a GARCH process?"
$$E[\sigma_{t,T}^2] = V_L \left[T - 1 - (\alpha+\beta)\frac{1-(\alpha+\beta)^{T-1}}{1-(\alpha+\beta)}\right] + \frac{1-(\alpha+\beta)^T}{1-(\alpha+\beta)}\sigma_t^2$$
- Suppose $(\alpha+\beta) = 0.9935$, $V_L = 0.0002075 \cong 0.0144^2$, and $\sigma_t^2 = 0.00003 \cong 0.01732^2$. The expected variance rate in 10 days is:
$$E[\sigma_{t+10}^2] = 0.0002075 + 0.9935^{10}(0.0003 - 0.0002075) = 0.00002942 \cong 0.0172^2$$
still above the long-term vol, but in 500 days (2 years)
$$E[\sigma_{t+500}^2] = 0.0002075 + 0.9935^{500}(0.0003 - 0.0002075) = 0.00002110 \cong 0.0145^2$$
very close to the long-term vol

SMI during 2008–2010, GARCH(1,1) and RiskMetrics



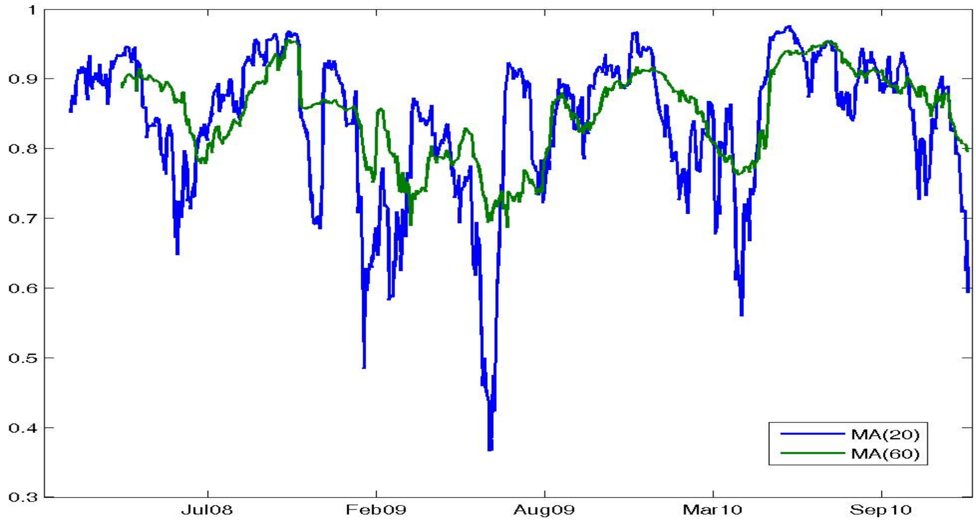
Part II

Models for time-varying correlations

Modeling correlations

- Correlations are equally important inputs in VaR calculations, especially for large portfolios.
- Changes in correlations may arise due to institutional reforms (liberalizations), improved technology, increased trading across countries, ...
- Correlations across stock markets tend to be higher during periods of market turmoil.
- Main issues in modeling correlations: (i) the number of the models' parameters grows exponentially as we consider more and more assets (ii) positive definiteness of the VCV matrix must be ensured.

SMI and DAX during 2008–2010, MA(20) and MA(60)



Multivariate GARCH

- There are multivariate extensions of GARCH, such as the VEC(1,1) model

$$\boldsymbol{\sigma_t^2} = w + A\boldsymbol{\xi_{t-1}} + B\boldsymbol{\sigma_{t-1}^2}$$

where $\boldsymbol{\sigma_t^2} = (\sigma_{1,t}^2 \ \ \sigma_{12,t} \ \ \sigma_{2,t}^2)'$ and $\boldsymbol{\xi_{t-1}} = (r_{1,t-1}^2 \ \ r_{1,t-1}r_{2,t-1} \ \ r_{2,t-1}^2)'$

- For $N = 2$, it means

$$\begin{bmatrix} \sigma_{1,t}^2 \\ \sigma_{12,t} \\ \sigma_{2,t}^2 \end{bmatrix} = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} r_{1,t-1}^2 \\ r_{1,t-1}r_{2,t-1} \\ r_{2,t-1}^2 \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} \sigma_{1,t-1}^2 \\ \sigma_{12,t-1} \\ \sigma_{2,t-1}^2 \end{bmatrix}$$

- For $N = 2$, the model involves already 21 parameters. For $N = 3$, we have 78! Moreover, restrictions among them should be imposed in order to guarantee positive definiteness.

RiskMetrics

- RiskMetrics is a special case, where $A = \lambda$ and $B = (1 - \lambda)$ so that

$$\begin{aligned} \sigma_{i,t}^2 &= (1-\lambda)r_{i,t-1}^2 + \lambda\sigma_{i,t-1}^2 \ \ \forall i = 1, ..., N \\ \sigma_{ij,t} &= (1-\lambda)r_{i,t-1}r_{j,t-1} + \lambda\sigma_{ij,t-1} \ \ \forall i,j = 1, ..., N \ j > i \end{aligned}$$

- Need the same λ to ensure correlations are well behaved (between -1 and 1). But this imposes the restriction that all correlations and volatilities are affected by more recent history in the same way.

- More realistic models allow correlations to vary independently of volatilities – Dynamic Conditional Correlation model (DCC) of Engle (2002)