Financial Risk Forecasting
Chapter 3
Multivariate volatility models

Jon Danielsson ©2020
London School of Economics

To accompany
Financial Risk Forecasting
www.financialriskforecasting.com
Published by Wiley 2011
Version 5.0, August 2020
Volatility

• The previous chapter focused on the volatility of a single asset
• In most we hold a portfolio of assets
• And therefore need to estimate both the volatility of each asset in the portfolio
• And the correlations between all the assets
• This means that it is much more complicated to estimate multivariate volatility models than univariate models
The focus of this chapter is on

- EWMA
- Orthogonal GARCH
- CCC and DCC models
- Estimation comparison
- Multivariate extensions of GARCH (MV GARCH and BEKK)
Notation

\[ \Sigma_t \]  Conditional covariance matrix
\[ Y_{t,k} \]  Return on asset \( k \) at time \( t \)
\[ y_{t,k} \]  Sample return on asset \( k \) at time \( t \)
\[ y_t = \{ y_{t,k} \} \]  Vector of sample returns on all assets at time \( t \)
\[ y = \{ y_t \} \]  Matrix of sample returns on all assets and dates
\[ A \text{ and } B \]  Matrices of parameters
\[ R \]  Correlation matrix
\[ D_t \]  Conditional variance forecast
Data

- Amazon and Google daily stock price from 2006-2015
  www.financialriskforecasting.com/data/amzn-goog.csv

```r
p = read.csv("http://www.financialriskforecasting.com/data/amzn-goog.csv")
y = apply(log(p[,2:3]),2,diff)
print(head(y))
```
R estimation

- It is easy to implement EWMA directly in R
- But no package exists for estimating all the models discussed below
- We mostly use Alexios Ghalanos’s rmgarch
Other reading

- Many surveys, e.g.
- For large (more than 25 assets) this paper has a proposal
Multivariate Volatility Forecasting
• Consider the univariate volatility model:

\[ Y_t = \sigma_t Z_t \]

where \( Y_t \) are returns, \( \sigma_t \) is conditional volatility and \( Z_t \) are random shocks

• If there are \( K > 1 \) assets under consideration, it is necessary to indicate which asset and parameters are being referred to, so the notation becomes more cluttered:

\[ Y_{t,i} = \sigma_{t,i} Z_{t,i} \]

where the first subscript indicates the date and the second subscript the asset
Conditional covariance matrix $\Sigma_t$

- The conditional covariance between two assets $i$ and $j$ is indicated by:
  \[
  \text{Cov}(Y_{t,i}, Y_{t,j}) \equiv \sigma_{t,ij}
  \]

- In the three-asset case (note that $\sigma_{t,ij} = \sigma_{t,ji}$):
  \[
  \Sigma_t = \begin{pmatrix}
  \sigma_{t,11} & \sigma_{t,12} & \sigma_{t,13} \\
  \sigma_{t,12} & \sigma_{t,22} & \sigma_{t,23} \\
  \sigma_{t,13} & \sigma_{t,23} & \sigma_{t,33}
  \end{pmatrix}
  \]
Portfolios

- If $w$ is the vector of portfolio weights
- The portfolio variance is

$$\sigma^2_{\text{portfolio}} = w'\Sigma w$$
The curse of dimensionality

- Number of diagonal elements is $K$ and off diagonal elements $K(K - 1)/2$ so in all

$$K + K(K - 1)/2$$

- For two assets it is 2+1, for 3 assets 3+4, for 4 10, etc.
- The explosion in the number of variance and especially covariance terms, as the number of assets increases, is known as the *curse of dimensionality*
- This is one reason why it is more difficult to estimate the covariance matrix
Positive semi-definiteness

- For univariate volatility we need to ensure that the variance is not negative \( \sigma^2 \geq 0 \)
- And for a portfolio
  \[
  \sigma_{\text{portfolio}}^2 = w' \Sigma w \geq 0
  \]
- So a covariance matrix should be positive semi-definite:
  \[
  |\Sigma| \geq 0
  \]
- This can be difficult to ensure
What about multivariate (MV) GARCH?

- For one asset
  \[ \sigma_{t+1}^2 = \omega + \alpha y_t^2 + \beta \sigma_t^2 \]
  
- For two
  \[ \sigma_{t+1,1}^2 = \omega_1 + \alpha_1 y_{t,1}^2 + \beta_2 \sigma_{t,1}^2 + \alpha_2 y_{t,2}^2 + \beta_2 \sigma_{t,2}^2 + \delta_1 \sigma_{t+1,1,2}^2 + \gamma_1 y_{t,1} y_{t,2} \]
  \[ \sigma_{t+1,2}^2 = \omega_2 + \alpha_3 y_{t,1}^2 + \beta_3 \sigma_{t,1}^2 + \alpha_4 y_{t,2}^2 + \beta_4 \sigma_{t,2}^2 + \delta_2 \sigma_{t+1,1,2}^2 + \gamma_2 y_{t,1} y_{t,2} \]
  \[ \sigma_{t+1,1,2}^2 = \omega_3 + \alpha_5 y_{t,1}^2 + \beta_5 \sigma_{t,1}^2 + \alpha_6 y_{t,2}^2 + \beta_6 \sigma_{t,2}^2 + \delta_3 \sigma_{t+1,1,2}^2 + \gamma_3 y_{t,1} y_{t,2} \]

- Or 21 parameters to estimate
- Almost impossible in practice
Numerical issues

• Stationarity is more important for multivariate volatility
• For univariate GARCH model, violation of covariance stationarity doesn’t matter for many applications and
• Does not hinder the estimation process with numerical problems
• A univariate volatility forecast is still obtained even if $\alpha + \beta > 1$
• This is generally not the case for MV GARCH models
• A parameter set resulting in violation of covariance stationarity might also lead to unpleasant numerical problems
• Numerical algorithms need to address these problems, thus complicating the programming process considerably
• Problems with multiple local minima, flat surfaces and other pathologies discussed in the last chapter
Instead

- Use some simplification approaches
- Unfortunately they come with significant trade-offs
- So MV estimation is much harder and much less accurate than univariate estimation

1. EWMA
2. CCC
3. DCC (perhaps most widely used for portfolios)
4. OGARCH (perhaps most widely used for combining portfolios)
5. BEKK (not to be recommended except for very small problems, $K = 2$, perhaps $K = 3$)
6. MV GARCH (practically impossible except maybe when $K = 2$)
7. Composite likelihood
EWMA
**EWMA model**

- **Univariate:**
  \[
  \hat{\sigma}_t^2 = \lambda \hat{\sigma}_{t-1}^2 + (1 - \lambda) y_{t-1}^2
  \]

- A vector of returns is
  \[
  \mathbf{y}_t = [y_{t,1}, y_{t,2}, \ldots, y_{t,K}]
  \]

- The multivariate EWMA is:
  \[
  \hat{\Sigma}_t = \lambda \hat{\Sigma}_{t-1} + (1 - \lambda) \mathbf{y}'_{t-1} \mathbf{y}_{t-1}
  \]

  with an individual element given by:
  \[
  \hat{\sigma}_{t,ij} = \lambda \hat{\sigma}_{t-1,ij} + (1 - \lambda) y_{t-1,i} y_{t-1,j}
  \]
Properties

- The same weight, $\lambda$, is used for all assets
- It is pre-specified and not estimated
- The variance of any particular asset only depends on its own lags
Pros and cons of multivariate EWMA model

Usefulness:

• Straightforward implementation, even for a large number of assets
• Covariance matrix is guaranteed to be positive semi-definite

Drawbacks:

• Restrictiveness:
  • Simple structure
  • The assumption of a single and usually non-estimated $\lambda$
**R estimation**

\[ \text{EWMA} = \text{matrix}(\text{nrow=dim}(y)[1], \text{ncol}=3) \]

\[ \lambda = 0.94 \]

\[ S = \text{cov}(y) \]

\[ \text{EWMA}[1,] = c(S)[c(1,4,2)] \]

\[
\text{for } (i \text{ in } 2: \text{dim}(y)[1]) \{
    S = \lambda \ast S + (1 - \lambda) \ast \ y[i-1,] \text{t}(y[i-1,])
    \text{EWMA}[i,] = c(S)[c(1,4,2)]
\}
\]

\[ \text{EWMArho} = \text{EWMA}[3] / \text{sqrt}(\text{EWMA}[1] \ast \text{EWMA}[2]) \]

Be careful with the matrix multiplications, getting the transposes right \((y[i-1,] \text{t}(y[i-1,])\) and not \(\text{t}(y[i-1]) \text{t}(y[i-1])\).
Constant Conditional Correlation Models
Steps

• We usually start by removing the mean — de-mean returns

• Separate out correlation modelling from volatility modelling
  1. correlation matrix
  2. variances

• Model volatilities with GARCH or some standard method

• The correlation matrix can be static (\textit{CCC}) or dynamic (\textit{DCC})
Definitions

• Let $D_t$ be a diagonal matrix where each element is the *volatility* of each asset

\[
D_{t,ii} = \sigma_{t,i}, \quad i = 1, \ldots, K
\]

\[
D_{t,ij} = 0, \quad i \neq j
\]

• Let $R$ or $R_t$ be the constant or time varying correlation matrix, respectively

• We will employ a two-step procedure and estimate $D_t$ and $R_t$ separately

• Then combine them to get the conditional covariance matrix, $\Sigma_t$
Conditional variance

• Use univariate GARCH (or some method) to estimate the variance of each asset separately, so for asset $i$

$$\sigma_{t,i}^2 = \omega_i + \alpha_i y_{t-1,i}^2 + \beta_i \sigma_{t-1,i}^2$$

• Note the parameters will be different for each asset
• $D_t$ is then created by putting these into the diagonal elements

$$D_{t,ii} = \sigma_{t,i}, \quad i = 1, \ldots, K$$
Getting the correlations

- After fitting the volatilities, divide them into the returns to get residuals

\[ \epsilon_{t,i} = \frac{y_{t,i}}{D_{t,i}} \]

- The vector of residuals at time \( t \) is then

\[ \epsilon_t = D_t^{-1}y_t \]

- And the matrix of residuals for all assets and time is

\[ \epsilon \]

\[ T \times K \]
Correlations of the residuals

• We want the correlations of the residuals
• For two different assets they are

\[ \text{Corr}(\epsilon_{i,t}, \epsilon_{j,t}) \]

• Since mean is zero and variance one, the \( K \times K \) correlation matrix is simply

\[ \hat{R} := \text{Cov}(\epsilon) = \frac{\epsilon \epsilon'}{T} \]
Constant conditional correlations (CCC)

- We now have the constant matrix $R$ and the time varying matrix $D_t$
- Combine these two to create the conditional variance matrix

$$\hat{\Sigma}_t = \hat{D}_t \hat{R} \hat{D}_t$$
Pros and cons

Pros

• Guarantees the positive definiteness of $\hat{\Sigma}_t$ if $\hat{R}$ is positive definite, as it is by definition
• Simple model, easy to implement
• Since matrix $\hat{D}_t$ has only diagonal elements, we can estimate each volatility separately

Cons

• The assumption of correlations being constant over time is at odds with the vast amount of empirical evidence supporting nonlinear dependence
Dynamic Conditional Correlation Models
Dynamic conditional correlations (DCC)

- DCC model is an extension of the CCC
- Let the correlation matrix $\hat{R}_t$ be *time dependent*
- The obvious way to do that is to do it like GARCH model

$$R_t = a + b\epsilon_{t-1}\epsilon'_{t-1} + cR_{t-1}$$

- That will not work since it does not ensure $|\Sigma_t| > 0$, i.e. positive definite
Conditions for positive definiteness

- $D_t$ is positive definite by construction since all elements in $D_t$ are positive or zero
- We can then ensure $|\Sigma_t| > 0$ if
- All elements in $R_t$ are $\leq 1$ and $\geq -1$
So need more steps

• We decompose $R_t$ into two parts
• One drives the dynamics — call that $Q_t$
• The other re-scales the dynamics to ensure each element is between -1 and +1 — call that $Z_t$
• So

$$R_t = Z_t Q_t Z_t$$
• Decompose $R_t$ into
• $Q_t$ is a positive definite matrix that dives the dynamics
• $Z_t$ re-scales $Q_t$ to ensure each element in it, the $|q_{t,ij}| < 1$
• All that is needed is to make each diagonal element of $Z_t$ be one over the corresponding diagonal element in $Q_t$

\[ Z_t = \begin{pmatrix}
\frac{1}{\sqrt{q_{t,11}}} & 0 & \cdots & 0 \\
0 & \frac{1}{\sqrt{q_{t,22}}} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & \frac{1}{\sqrt{q_{t, kk}}} \\
\end{pmatrix} \]
The $Q_t$ follow a GARCH type process (ARMA) type process

$$Q_t = (1 - \zeta - \xi)\bar{R} + \zeta \epsilon_{t-1}\epsilon'_{t-1} + \xi Q_{t-1}$$

- $\bar{R}$ is the $(K \times K)$ unconditional covariance matrix of $\epsilon$
- $\zeta$ and $\xi$ are parameters
- Parameter restrictions:
  - Positive definiteness $\zeta, \xi > 0$
  - Stationarity $\zeta + \xi < 1$
Pros and cons of DCC

Pros

- Large covariance matrices can be easily be estimated

Cons

- Parameters $\zeta$ and $\xi$ are constants
- So the conditional correlations of all assets are driven by the same underlying dynamics as the two parameters are same for all assets
Large problems
Large problems

- Even a medium-sized financial institutions will have hundreds of thousands or millions of types of assets — what is known as \textit{risk factors}
- Estimating the covariance matrix for the entire institution is effectively impossible using methods like EWMA or DCC
- Instead, we can split the risk factors up into subcomponents
- Estimate the covariance of each
- And then combine them back
For example

- US
  - Bonds
  - Commodities
    - S&P 100 Index
  - Equities
    - S&P 400 Mid Cap Index
- China
  - Bonds
  - Equities
- Brazil
  - Bonds
  - Equities
- Bank
Simplifying the problem

• Suppose you are willing to assume a single correlation between all Brazilian assets and all US assets

• And you are willing to assume that the correlation between US bonds and US equities is the same for all bonds and equities

• And you are willing to assume that the correlation between US large-cap and small-cap stocks is the same for all stocks

• Then it is easy to calculate the covariance matrix for the entire bank

• Use perhaps DCC for US small caps and other portfolios

• And combine them by the constant correlation
Consider an example...

- Your financial institution has the following securities:
Consider an example...

- Your financial institution has the following securities:

\[ \sigma_E^2 \]
Consider an example...

- Your financial institution has the following securities:

\[ \sigma^2_E \quad \sigma^2_B \]
Consider an example...

• Your financial institution has the following securities:

\[ \sigma_E^2 \quad \sigma_B^2 \quad \sigma_B^2 \]
Consider an example...

• Your financial institution has the following securities:

\[
\begin{align*}
\sigma_E^2 & \quad \sigma_B^2 \\
\sigma_E^2 & \quad \sigma_B^2
\end{align*}
\]
Consider an example...

- Your financial institution has the following securities:

\[
\sigma_E^2 \quad \sigma_B^2 \quad \sigma_E^2 \quad \sigma_B^2
\]

where red is UK, blue is Germany, E denotes EQUITY and B is BOND
Consider an example...

- Your financial institution has the following securities:

\[
\begin{bmatrix}
\sigma_E^2 & \sigma_B^2 \\
\sigma_E^2 & \sigma_B^2 \\
\end{bmatrix}
\]

where red is UK, blue is Germany, E denotes EQUITY and B is BOND

- Individual covariance matrices of your securities can be combined into a covariance matrix for the entire financial institution:
Consider an example...

- Your financial institution has the following securities:

\[
\begin{bmatrix}
\sigma^2_E \\
\sigma^2_B \\
\sigma^2_E \\
\sigma^2_B \\
\end{bmatrix}
\]

where red is UK, blue is Germany, E denotes EQUITY and B is BOND

- Individual covariance matrices of your securities can be combined into a covariance matrix for the entire financial institution:

\[
\sigma^2_E
\]
Consider an example...

- Your financial institution has the following securities:
  \[
  \sigma_E^2 \quad \sigma_B^2 \quad \sigma_E^2 \quad \sigma_B^2
  \]
  where red is UK, blue is Germany, E denotes EQUITY and B is BOND

- Individual covariance matrices of your securities can be combined into a covariance matrix for the entire financial institution:
  \[
  \sigma_E^2 \quad \sigma_B^2
  \]
Consider an example...

- Your financial institution has the following securities:
  
  \[
  \begin{pmatrix}
  \sigma_E^2 & \sigma_B^2 \\
  \sigma_E^2 & \sigma_B^2 \\
  \end{pmatrix}
  \]

  where red is UK, blue is Germany, E denotes EQUITY and B is BOND.

- Individual covariance matrices of your securities can be combined into a covariance matrix for the entire financial institution:

  \[
  \begin{pmatrix}
  \sigma_E^2 \\
  \rho_{E,B} \sigma_B^2 \\
  \end{pmatrix}
  \]
Consider an example...

• Your financial institution has the following securities:

\[
\begin{array}{cccc}
\sigma^2_E & \sigma^2_B & \sigma^2_E & \sigma^2_B \\
\end{array}
\]

where red is UK, blue is Germany, E denotes EQUITY and B is BOND

• Individual covariance matrices of your securities can be combined into a covariance matrix for the entire financial institution:

\[
\begin{array}{ccc}
\sigma^2_E & \rho_{E,B} & \sigma^2_B \\
\end{array}
\]
Consider an example...

• Your financial institution has the following securities:

\[
\begin{bmatrix}
\sigma_E^2 & \sigma_B^2 \\
\rho_{E,B} & \sigma_E^2 & \sigma_B^2 \\
\rho_{E,E} & \rho_{B,E} & \sigma_E^2 \\
\end{bmatrix}
\]

where red is UK, blue is Germany, E denotes EQUITY and B is BOND

• Individual covariance matrices of your securities can be combined into a covariance matrix for the entire financial institution:
Consider an example...

- Your financial institution has the following securities:
  \[ \sigma_E^2 \quad \sigma_B^2 \quad \sigma_E^2 \quad \sigma_B^2 \]
  where red is UK, blue is Germany, E denotes EQUITY and B is BOND

- Individual covariance matrices of your securities can be combined into a covariance matrix for the entire financial institution:
  \[ \sigma_E^2 \quad \rho_{E,B} \quad \sigma_B^2 \quad \rho_{E,E} \quad \rho_{B,E} \quad \sigma_E^2 \quad \sigma_B^2 \]
Consider an example...

- Your financial institution has the following securities:
  \[
  \begin{array}{cccc}
  \sigma^2_E & \sigma^2_B & \sigma^2_E & \sigma^2_B \\
  \end{array}
  \]
  where red is **UK**, blue is **Germany**, E denotes EQUITY and B is BOND.

- Individual covariance matrices of your securities can be combined into a covariance matrix for the entire financial institution:
  \[
  \begin{array}{cccc}
  \sigma^2_E & \rho_{E,B} & \sigma^2_B \\
  \rho_{E,E} & \rho_{B,E} & \sigma^2_E \\
  \rho_{B,B} & \rho_{E,B} & \sigma^2_B \\
  \end{array}
  \]
Consider an example...

• Your financial institution has the following securities:

\[
\begin{pmatrix}
\sigma_E^2 & \sigma_B^2 \\
\sigma_E^2 & \sigma_B^2
\end{pmatrix}
\]

where red is UK, blue is Germany, E denotes EQUITY and B is BOND

• Individual covariance matrices of your securities can be combined into a covariance matrix for the entire financial institution:
Consider an example...

- Your financial institution has the following securities:
  \[
  \begin{pmatrix}
  \sigma_E^2 & \sigma_B^2 & \sigma_E^2 & \sigma_B^2 \\
  \sigma_E^2 & \rho_{E,B} & \rho_{E,E} & \rho_{E,B} \\
  \rho_{B,E} & \sigma_B^2 & \sigma_E^2 & \sigma_B^2 \\
  \rho_{B,B} & \rho_{E,B} & \sigma_B^2 & \sigma_B^2
  \end{pmatrix}
  \]
  where red is UK, blue is Germany, E denotes EQUITY and B is BOND

- Individual covariance matrices of your securities can be combined into a covariance matrix for the entire financial institution:
Consider an example...

- Your financial institution has the following securities:

\[
\begin{array}{cccc}
\sigma^2_E & \sigma^2_B & \sigma^2_E & \sigma^2_B \\
\rho_{E,B} & \sigma^2_B & \rho_{E,E} & \rho_{B,E} \\
\rho_{B,B} & \rho_{B,E} & \sigma^2_E & \rho_{E,B} \\
\sigma^2_B & \rho_{E,B} & \rho_{B,E} & \sigma^2_B \\
\end{array}
\]

where red is UK, blue is Germany, E denotes EQUITY and B is BOND.

- Individual covariance matrices of your securities can be combined into a covariance matrix for the entire financial institution:
Estimation Comparison
Prices

- MSFT
- IBM
(b) Returns

- MSFT
- IBM
Correlation estimates

with average correlation 49%
Correlation estimates
with average correlation 49%
Let’s focus on 2008, the midst of the 2007-2009 crisis, when the correlations of all stocks increased dramatically...
Returns

MSFT

IBM

Correlations

with average correlation 49%
Correlations

with average correlation 49%
BEKK
The BEKK model

- An alternative to the MV-GARCH models
- The matrix of conditional covariances is $\Sigma_t$
- A function of the outer product of lagged returns and lagged conditional covariances
- Each pre-multiplied and post-multiplied by a parameter matrix
- Results in a quadratic function that is guaranteed to be positive semi-definite
The two-asset, one-lag BEKK(1,1,2) model is defined as:

$$\Sigma_t = \Omega \Omega' + A' Y_{t-1} Y_{t-1} A + B' \Sigma_{t-1} B$$

or:

$$\Sigma_t = \begin{pmatrix} \sigma_{t,11} & \sigma_{t,12} \\ \sigma_{t,12} & \sigma_{t,22} \end{pmatrix} = \begin{pmatrix} \omega_{11} & 0 \\ \omega_{21} & \omega_{22} \end{pmatrix} \begin{pmatrix} \omega_{11} & 0 \\ \omega_{21} & \omega_{22} \end{pmatrix}'$$

$$+ \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}' \begin{pmatrix} Y_{t-1,1}^2 & Y_{t-1,1} Y_{t-1,2} \\ Y_{t-1,2} Y_{t-1,1} & Y_{t-1,2}^2 \end{pmatrix} \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}'$$

$$+ \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix}' \begin{pmatrix} \sigma_{t-1,11} & \sigma_{t-1,12} \\ \sigma_{t-1,12} & \sigma_{t-1,22} \end{pmatrix} \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix}$$
• The general $\text{BEKK}(L_1, L_2, K)$ model is given by:

$$
\Sigma_t = \Omega \Omega' + \sum_{k=1}^{K} \sum_{i=1}^{L_1} A'_{i,k} Y'_{t-i} Y_{t-i} A_{i,k} + \sum_{k=1}^{K} \sum_{j=1}^{L_2} B'_{j,k} \Sigma_{t-j} B_{j,k}
$$

• The number of parameters in the $\text{BEKK}(1,1,2)$ model is $K(5K + 1)/2$
  • 11 in two asset case
  • 24 in three asset case
  • 42 in four asset case
Pros and cons

Pros

• Allows for interactions between different asset returns and volatilities
• Relatively parsimonious

Cons

• Parameters hard to interpret
• Many parameters are often found to be statistically insignificant, which suggests the model may be overparametrized
• Can only handle a small number of assets
Multivariate volatility analysis

- Pick 7 stocks
- Some that did quite well in the crisis
- Others that performed really badly
- And yet others that performed average
- Use DCC to get the conditional correlations
- Normalize the prices of all to start at one at the beginning of the sample
Amazon

- Amazon prices jumped by 8.5% at the end of January
- Because of really good fourth-quarter results
- Note how that affects the conditional correlations