Multivariate volatility modeling

Daniele Bianchi[1]

whitesphd.com

[1]School of Economics and Finance
Queen Mary, University of London

Summary

This week we introduce models for multivariate volatility and correlations within the context of financial returns.

Contents

Introduction

Volatility

- The previous lectures focused on the volatility of a single asset
- In most situations we hold more than a one asset, i.e., a portfolio.
- ... and therefore need to estimate both the volatility of each asset in the portfolio...
- ... as well as the correlations among all assets.
- This means that it is much more complicated to estimate multivariate volatility models than univariate models.

Notation

- Σ_t Conditional covariance matrix
- $y_{t,k}$ Return on asset k at time t
- $\mathbf{y}_t$ Vector of sample returns on all assets at time t
- $\mathbf{y}$ Matrix of sample returns on all assets and dates
- A and B Matrices of parameters
- R Correlation matrix
- D_t Conditional variance forecast

Multivariate volatility models

Consider the univariate volatility model:

$$y_t = \sigma_t Z_t,$$

where y_t are returns, σ_t is the conditional volatility and Z_t are the 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;

$$y_{i,t} = \sigma_{i,t} Z_{i,t},$$

where the first subscript indicates the date and the second subscript the asset.

Conditional covariance matrix

The conditional covariance between two assets i and j is indicated by:

$$\text{Cov}\left(y_{i,t}, y_{j,t}\right) \equiv \sigma_{ij,t}$$

In the three-asset case (note that $\sigma_{ij,t} = \sigma_{ji,t}$):

$$\Sigma_t = \begin{pmatrix} \sigma_{t,11} & & \\ \sigma_{t,12} & \sigma_{t,22} & \\ \sigma_{t,13} & \sigma_{t,23} & \sigma_{t,33} \end{pmatrix}$$

The conditional covariance matrix has direct implications for portfolio modeling. If w is the portfolio weights, then the portfolio variance is:

$$\sigma_p^2 = w' \Sigma w$$

Issue 1: The curse of dimensionality

Number of diagonal elements is K and off diagonal elements $K + K(K-1)/2$ so in all we have to estimate $K + K(K-1)/2$ parameters

For two assets it is $2+1$, for 3 assets $3+4$, for 4 is 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

Issue 2: Positive semi-definiteness

For univariate volatility we need to ensure that the variance is not negative $\left(\sigma^2 > 0\right)$

And for a portfolio

$$\sigma_p^2 = w'\Sigma w \geq 0$$

So a covariance matrix should be positive semi-definite:

$$\|\Sigma\| \geq 0$$

This somewhat 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 assets:

$$\begin{aligned}
\sigma_{1,t+1}^2 &= \omega_1 + \alpha_1 y_{1,t}^2 + \beta_1 \sigma_{1,t}^2 + \alpha_2 y_{2,t}^2 + \beta_2 \sigma_{2,t}^2 + \delta_1 \sigma_{1,2,t+1} + \gamma_1 y_{1,t} y_{2,t} \\
\sigma_{2,t+1}^2 &= \omega_2 + \alpha_3 y_{1,t}^2 + \beta_3 \sigma_{1,t}^2 + \alpha_4 y_{2,t}^2 + \beta_4 \sigma_{2,t}^2 + \delta_2 \sigma_{1,2,t+1} + \gamma_2 y_{1,t} y_{2,t} \\
\sigma_{1,2,t+1} &= \omega_3 + \alpha_5 y_{1,t}^2 + \beta_5 \sigma_{1,t}^2 + \alpha_6 y_{2,t}^2 + \beta_6 \sigma_{2,t}^2 + \delta_3 \sigma_{1,2,t+1} + \gamma_3 y_{1,t} y_{2,t}
\end{aligned}$$

Or 21 parameters to estimate; and this is just for two assets!!!
Almost impossible for reasonable portfolios

Multivariate volatility models

EWMA

Recall the univariate case:

$$\sigma_t^2 = \lambda \sigma_{t-1}^2 + (1-\lambda)\, y_{t-1}^2$$

Define a vector of returns as:

$$\mathbf{y}_t = (y_{1,t}, y_{2,t}, \ldots, y_{K,t})$$

The multivariate version of the EWMA is then:

$$\Sigma_t^2 = \lambda \Sigma_{t-1}^2 + (1-\lambda)\, \mathbf{y}_{t-1}' \mathbf{y}_{t-1}$$

with an individual ij element given by:

$$\sigma_{ij,t}^2 = \lambda \sigma_{ij,t-1}^2 + (1-\lambda)\, y_{i,t-1} y_{j,t-1}$$

EWMA Properties

Properties:

- The same weight λ is used for all assets
- It is pre-specified and not estimated
- The variance of an asset only depends on its own lags.

Pros:

- Easy to implement, even for a large number of assets
- The matrix Σ_t is guaranteed to be positive semi-definite

Cons:

- The simple structure perhaps too restrictive
- For instance, the assumption of a single and usually calibrated λ may be unrealistic.

Constant Conditional Correlation

The Constant Conditional Correlation (CCC) model estimation proceeds in steps:

Step 1 De-mean returns

Step 2 Separate out correlation modelling from volatility modelling:

- Correlation matrix
- Variances

Step 3 Model volatilities with GARCH or some standard method

Step 4 The correlation matrix is assumed static in the CCC model

CCC Estimation

Let D_t be a diagonal matrix where each element is the volatility of each asset:

$$\begin{aligned} D_{ii,t} &= \sigma_{i,t}, \qquad i = 1, \ldots, K \\ D_{ij,t} &= 0, \qquad i \neq j \end{aligned}$$

Use univariate GARCH (or some method) to estimate the variance of each asset separately, i.e., to get D_t

$$D_{ii,t} = \sigma_{i,t} = \sqrt{\omega_i + \alpha_i y_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2}, \qquad i = 1, \ldots K$$

We want the correlations of the residuals:

$$\epsilon_{i,t} = \frac{y_{i,t}}{\sigma_{i,t}}, \qquad \text{or in matrix form} \qquad \epsilon_t = D_t^{-1} \mathbf{y}_t$$

CCC Estimation (Cont'd)

Define $\epsilon = (\epsilon_1, \ldots, \epsilon_T)$. Then the correlations are

$$R = \frac{\epsilon\epsilon'}{T}$$

Then combine these elements to get

$$\Sigma_t = D_t R D_t$$

Note that while D_t is time dependent, R is constant over time.

CCC: Pros and Cons

Pros:

- Guarantees the positive definiteness of Σ_t if R is positive.
- Simple model, easy to implement
- Since matrix 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

DCC model is an extension of the CCC

Let the correlation matrix R_t be time-varying

While one might propose a model like

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

That will not work since we have to ensure that $|\Sigma_t| > 0$, i.e., positive definite. For that:

- D_t is positive definite by construction since all elements D_t are positive
- All elements in R_t need to be ≤ 1 and ≥ -1.

Dynamic Conditional Correlation (DCC)

So one need more steps. Decompose R_t into

$$R_t = Q_t^{*\prime} Q_t Q_t^*$$

Q_t is a positive definite matrix that drives the dynamics

Q_t^* re-scales Q_t to ensure each element $|q_{ij,t}| < 1$ s.t.

$$Q_t^{*\prime} = \begin{pmatrix} 1/\sqrt{q_{11,t}} & 0 & \dots & 0 \\ 0 & 1/\sqrt{q_{22,t}} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 1/\sqrt{q_{KK,t}} \end{pmatrix}$$

Dynamic Conditional Correlation (DCC)

Then have Q_t follow and ARMA type process

$$Q_t = (1 - \alpha - \beta)\,\overline{Q} + \alpha\epsilon_{t-1}\epsilon_{t-1}' + \beta Q_{t-1}$$

where:

- $\overline{Q}$ is the $(K \times K)$ unconditional covariance matrix of ϵ
- α and β are parameters with some restriction:
  - Positive definiteness $\alpha, \beta > 0$
  - Stationarity $\alpha + \beta < 1$

Dynamic Conditional Correlation (DCC)

Pros:

- Relatively large covariance matrixes can be easily estimated

Cons:

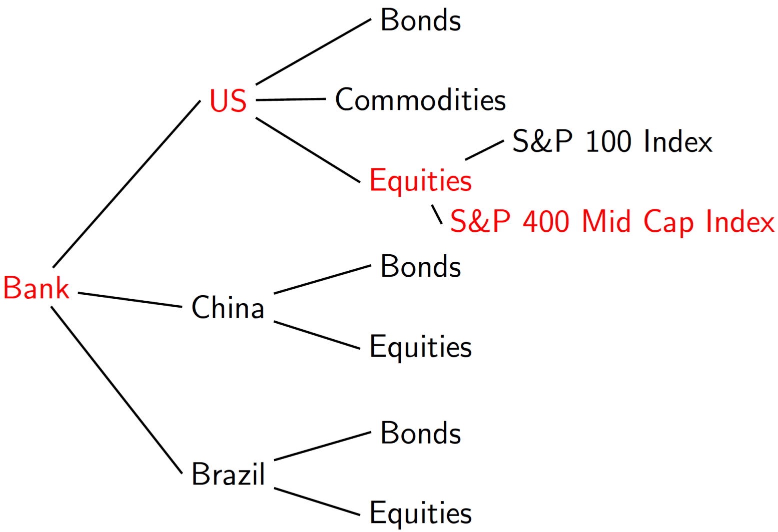
- The parameters α and β are the same across assets.
- So the conditional correlations of all assets are driven by the same underlying dynamics, parameters are same for all assets.

Large-dimensional problems

Large problems

- Even a medium-sized fund have hundreds of assets to trade and risks to monitor
- Estimating the covariance matrix for the entire universe of assets is practically impossible using methods like the EWMA or the DCC
- Instead, we can split the assets/risk factors up into subcomponents
- Estimate the covariance of each
- And then combine them back

Large problems (cont'd)



Large problems (cont'd)

- The orthogonal approach transforms the observed returns matrix into a set of portfolios with the key property that they are uncorrelated
- We can forecast their volatilities separately
- Orthogonalization based on Principal Component Analysis (PCA)
- Known as orthogonal GARCH, or O-GARCH
- This is because it involves transforming correlated returns into uncorrelated portfolios and then using GARCH to forecast the volatilities of each uncorrelated portfolio separately.

Large problems (cont'd)

The idea is simple: Take a matrix of returns $\underset{T\times K}{\mathbf{Y}}$ which has covariance $\underset{T\times K}{\Sigma}$

Let

$$D = \sqrt{\text{diag}\left(\Sigma\right)}$$

Then the correlations are

$$R = D^{-1}\Sigma D^{-1}$$

Calculate the $K \times K$ matrix Λ of eigenvectors, then define

$$\mathbf{U} = \Lambda \times \mathbf{Y}$$

which transform the matrix $\mathbf{Y}$ into the uncorrelated portfolios $\mathbf{U}$.

Large problems (cont'd)

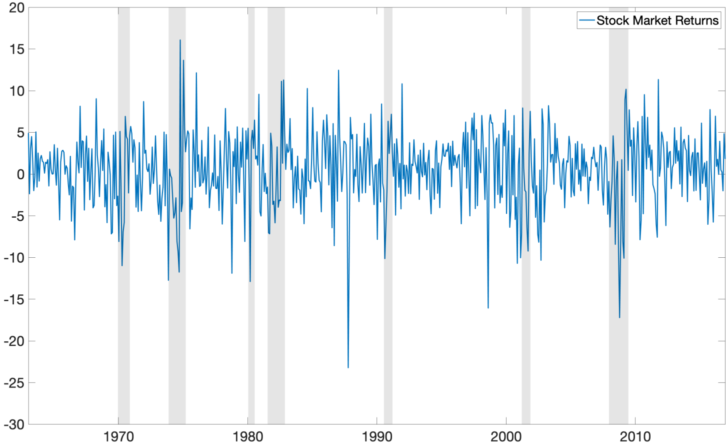
Main advantages:

- Large number of assets
- The method also allows estimates for volatilities and correlations of variables to be generated even when data are sparse (e.g., in illiquid markets)
- The use of PCA guarantees the positive definiteness of the covariance matrix
- PCA also facilitates building a covariance matrix for an entire financial institution by iteratively combining the covariance matrices of the various trading desks, simply by using one or perhaps two PCs

Empirical Application

US Stocks and Bonds

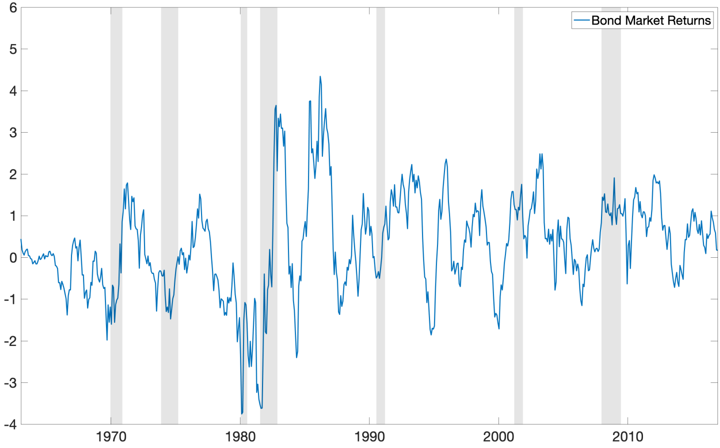
Figure 1: US Aggregate Stock Market Returns (monthly)



This figure shows the monthly returns on the US aggregate stock market index which consists of the AMEX/NYSE (see CRSP) from January 1963 to December 2016.

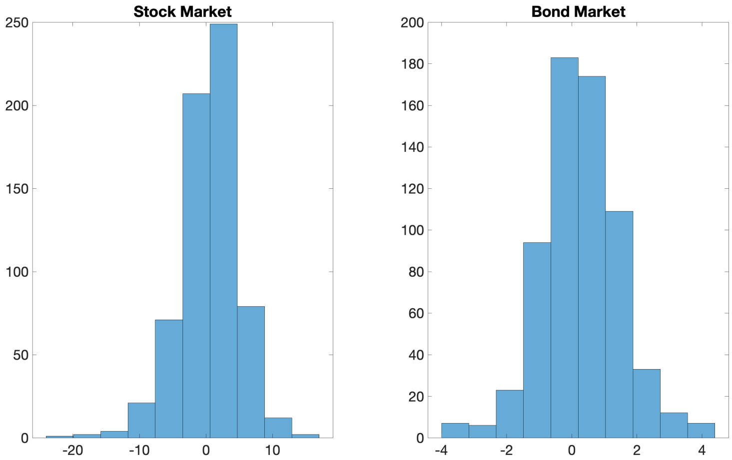
US Stocks and Bonds

Figure 2: US Aggregate Bond Market Returns (monthly)



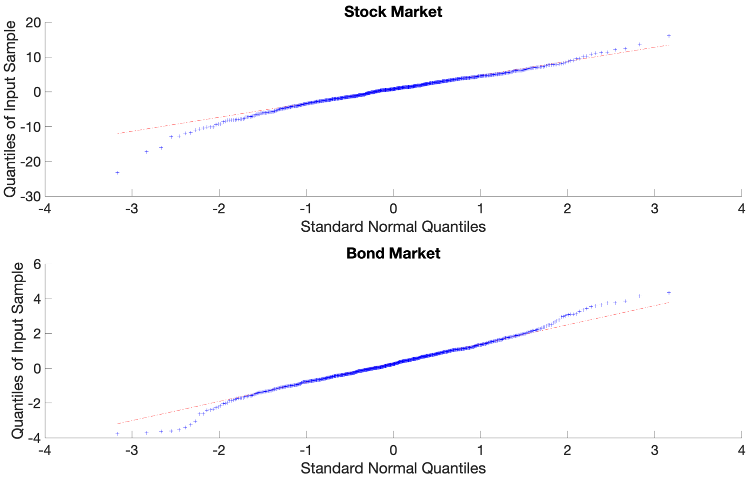
This figure shows the monthly returns on the US aggregate bond market index which from January 1963 to December 2016.

US Stocks and Bonds



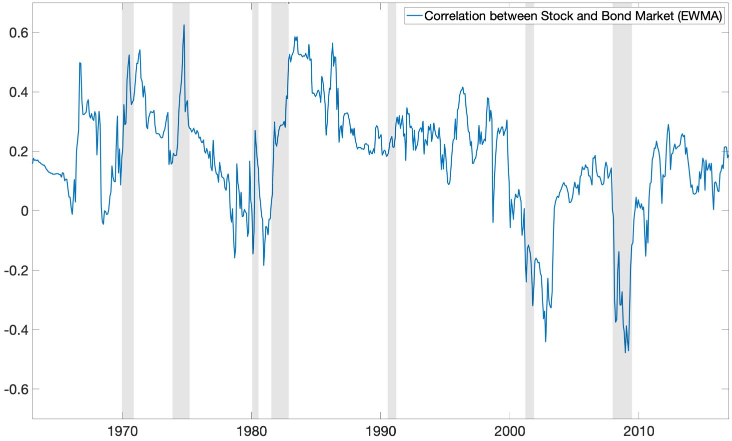
This figure shows the unconditional distribution of both monthly returns on the US aggregate bond market (right panel) and the monthly returns on the US aggregate stock market (left panel). The sample is from January 1963 to December 2016.

US Stocks and Bonds



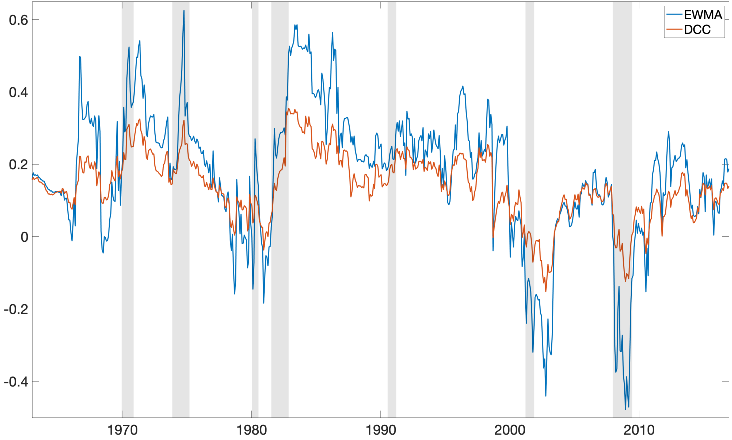
This figure shows the qq-plot both monthly returns on the US aggregate bond market (bottom panel) and the monthly returns on the US aggregate stock market (top panel). The sample is from January 1963 to December 2016.

US Stocks and Bonds



This figure shows the correlation from an EWMA with $\lambda = 0.4$ of monthly returns on the US aggregate bond market and the US aggregate stock market. The sample is from January 1963 to December 2016.

US Stocks and Bonds



This figure shows the correlation from an EWMA with $\lambda = 0.4$ of monthly returns on the US aggregate bond market and the US aggregate stock market. The sample is from January 1963 to December 2016.

The BEKK model

- An alternative to the MV-GARCH models.
- The matrix of conditional covariances is Σ_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 BEKK model

The two-asset one-lag BEKK(1,1,2). model is defined as:

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

The general BEKK(L_1, L_2, K) model is given by:

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

The number of parameters in the BEKK(1,1,2) model is $K(5K+1)/2$, e.g., 42 in four asset case

The BEKK model

Pros:

- Allows for interactions between different asset returns and volatilities.
- Relatively parsimonious.

Cons:

- Parameters hard to interpret.
- Parameters hard to interpret statistically insignificant, which suggests the model may be overparametrized.
- Can only handle a small number of assets