Financial Data Analysis

Multivariate GARCH

July 26, 2011
Multivariate GARCH

- Many problems in finance are inherently multivariate and require us to understand the dependence structure between assets.

- E.g.,
  - portfolio analysis,
  - volatility transmission: study of relations between the volatilities and covariances/correlations of several markets (e.g., emerging and developed markets, or different regions),
  - relation between correlations and volatilities in different market regimes (e.g., bull vs. bear markets),
  - tests of asset pricing models,
  - futures hedging.

- Multivariate GARCH: Models for the evolution of volatilities and covariances/correlations.
Consider a return vector $r_t$ consisting of $N$ components, i.e., $r_t = [r_{1t}, r_{2t}, \ldots, r_{Nt}]'$ (a column vector),

$$r_t = \mu_t + \epsilon_t \quad (1)$$

$$\mu_t = E(r_t|I_{t-1}) = E_{t-1}(r_t) \quad (2)$$

$$\epsilon_t|I_{t-1} \sim N(0, H_t) \quad (3)$$

$$H_t = \text{Var}(r_t|I_{t-1}) = \text{Var}_{t-1}(r_t) = \text{Var}_{t-1}(\epsilon_t), \quad (4)$$

where $I_t$ is the information available at time $t$, usually $I_t = \{r_t, r_{t-1}, \ldots\}$.

- The error term $\epsilon_t = [\epsilon_{1t}, \epsilon_{2t}, \ldots, \epsilon_{Nt}]'$.

- $H_t$ is the conditional covariance matrix of $r_t$. 


Covariance matrix

\[ H_t = \begin{bmatrix}
  h_{1t}^2 & h_{12,t} & h_{13,t} & \cdots & h_{1N,t} \\
  h_{12,t} & h_{2t}^2 & h_{23,t} & \cdots & h_{2N,t} \\
  h_{13,t} & h_{23,t} & h_{3t}^2 & \cdots & h_{3N,t} \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  h_{1N,t} & h_{2N,t} & h_{3N,t} & \cdots & h_{Nt}^2
\end{bmatrix}, \quad (5) \]

where

\[ h_{jt}^2 = \text{Var}_{t-1}(r_{jt}), \quad h_{ij,t} = \text{Cov}_{t-1}(r_{it}, r_{jt}), \quad (6) \]

is symmetric and positive definite:

- We know that for any linear combination (with weight vector \( w = [w_1, w_2, \ldots, w_N]' \)) of the elements of \( r_t \),

\[
0 < \text{Var}_{t-1}\left( \sum_i w_i r_{it} \right) = \sum_i w_i^2 h_{i,t}^2 + \sum_i \sum_{j \neq i} w_i w_j h_{ij,t} = w' H_t w.
\]

\(^{1}\)The variance may be zero if the components are linearly dependent.
• For example, with $N = 2$,

$$\text{Var}_{t-1}(w_1 r_{1t} + w_2 r_{2t}) = w_1^2 h_{1t}^2 + 2w_1 w_2 h_{12,t} + w_2^2 h_{2t}^2$$

$$= \begin{bmatrix} w_1 & w_2 \end{bmatrix} \begin{bmatrix} h_{1t}^2 & h_{12,t} \\ h_{12,t} & h_{2t}^2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}.$$  

• If the conditional distribution of $r_t$ is multivariate normal, then, for example, the conditional $100 \times \xi\%$ portfolio Value–at–Risk (VaR) for any portfolio combination $w$ can be calculated as

$$\text{VaR}_{t-1}(\xi) = w' \mu_t + \Phi^{-1}(\xi) \sqrt{w' H_t w}, \quad (7)$$

where $\Phi^{-1}(\xi)$ is the $\xi$–quantile of the standard normal distribution, e.g., $\Phi^{-1}(0.01) = -2.3263$ and $\Phi^{-1}(0.05) = -1.6449$. 

Similar to the univariate GARCH,

\[ r_t = \mu_t + \epsilon_t, \quad \epsilon_t = \sigma_t \eta_t, \quad \eta_t \iid \sim \text{N}(0, 1), \]

(3) is often written as

\[ \epsilon_t = H_t^{1/2} z_t, \quad z_t \iid \sim \text{N}(0, I), \]

where \( \text{N}(0, I) \) denotes the \( N \)-dimensional normal distribution with a mean vector of zeros and identity covariance matrix, i.e., the \( N \)-dimensional standard normal.

- \( H_t^{1/2} \) is an \( N \times N \) matrix such that \( H_t^{1/2}(H_t^{1/2})' = H_t \) (matrix square root).

- As \( H_t \) is a covariance matrix, such a factorization exists, e.g., the Cholesky decomposition.
• A symmetric positive definite matrix $A$ can be factored as $A = LL'$, where $L$ is lower triangular with positive diagonal elements (the Cholesky factorization of $A$).\(^2\)

• For example, if $N = 2$ (bivariate case), where

$$H_t = \begin{bmatrix} h_{1t}^2 & h_{12,t} \\ h_{12,t} & h_{2t}^2 \end{bmatrix},$$

the Cholesky factorization is

$$L = \begin{bmatrix} \sqrt{h_{1t}^2} & 0 \\ h_{12,t}/\sqrt{h_{1t}^2} & \sqrt{h_{2,t}^2 - h_{12}^2/h_{1t}^2} \end{bmatrix}.$$ 

• $LL' = H_t$ is easily checked, and $h_{2,t}^2 - h_{12}^2/h_{1t}^2 = (h_{1t}^2 h_{2,t}^2 - h_{12}^2)/h_{1t}^2 = (\det H_t)/h_{1t}^2 > 0$ since $H_t$ is positive definite.

\(^2\)Other factorizations exist.
• It then follows from (8) that

\[
\text{Var}_{t-1}(r_t) = \text{Var}_{t-1}(\epsilon_t) = \text{E}_{t-1}(\epsilon_t \epsilon'_t) - \underbrace{\text{E}_{t-1}(\epsilon_t) \text{E}_{t-1}(\epsilon_t)'}_{=0} = \text{E}_{t-1}(H_t^{1/2} z_t z'_t (H_t^{1/2})') = H_t^{1/2} \begin{pmatrix} H_t^{1/2} \\ 0 \end{pmatrix} = H_t^{1/2} (H_t^{1/2})' = H_t.
\]

(9) (10) (11) (12) (13)
Main Problems

• There are two main problems when it comes to the specification of multivariate GARCH models:

  (i) To keep estimation feasible, we need parsimonious models (i.e., models with a moderate number of parameters) which are still flexible enough to capture the most important aspects of the volatility/covariance dynamics.
  (ii) We have to make sure that the conditional covariance matrix will remain positive definite at each point of time.

• For the sake of illustration, consider a bivariate GARCH(1,1) of the general vec–type.

• The covariance matrix is then given by

\[ H_t = \begin{bmatrix} h_{1t}^2 & h_{12,t} \\ h_{12,t} & h_{2t}^2 \end{bmatrix}, \]
where, in the most general case

\[
\begin{align*}
    h_{1t}^2 &= c_1 + a_{11} \epsilon_{1,t-1}^2 + a_{12} \epsilon_{1,t-1} \epsilon_{2,t-1} + a_{13} \epsilon_{2,t-1}^2 \\
    &\quad + b_{11} h_{1,t-1}^2 + b_{12} h_{12,t-1} + b_{13} h_{2,t-1}^2 \\
    h_{12,t} &= c_2 + a_{21} \epsilon_{1,t-1}^2 + a_{22} \epsilon_{1,t-1} \epsilon_{2,t-1} + a_{23} \epsilon_{2,t-1}^2 \\
    &\quad + b_{21} h_{1,t-1}^2 + b_{22} h_{12,t-1} + b_{23} h_{2,t-1}^2 \\
    h_{2t}^2 &= c_3 + a_{31} \epsilon_{1,t-1}^2 + a_{32} \epsilon_{1,t-1} \epsilon_{2,t-1} + a_{33} \epsilon_{2,t-1}^2 \\
    &\quad + b_{31} h_{1,t-1}^2 + b_{32} h_{12,t-1} + b_{33} h_{2,t-1}^2,
\end{align*}
\]

or

\[
\begin{bmatrix}
    h_{1t}^2 \\
    h_{12,t} \\
    h_{2t}^2
\end{bmatrix}
= h_t =
\begin{bmatrix}
    c_1 \\
    c_2 \\
    c_3
\end{bmatrix}
+ \begin{bmatrix}
    a_{11} & a_{12} & a_{13} \\
    a_{21} & a_{22} & a_{23} \\
    a_{31} & a_{32} & a_{33}
\end{bmatrix}
\begin{bmatrix}
    \epsilon_{1,t-1}^2 \\
    \epsilon_{1,t-1} \epsilon_{2,t-1} \\
    \epsilon_{2,t-1}^2
\end{bmatrix}
+ \begin{bmatrix}
    b_{11} & b_{12} & b_{13} \\
    b_{21} & b_{22} & b_{23} \\
    b_{31} & b_{32} & b_{33}
\end{bmatrix}
\begin{bmatrix}
    h_{1,t-1}^2 \\
    h_{12,t-1} \\
    h_{2,t-1}^2
\end{bmatrix}.
\]
In this specification, both conditional variances, $h_{1t}^2$ and $h_{2t}^2$, and the conditional covariance, $h_{12,t}$, may depend on all lagged squared returns and variances and all lagged cross–products $\epsilon_{1,t-1}\epsilon_{2,t-1}$ and covariances.

Although flexible, this model is difficult to handle in practice, since it requires estimation of 21 parameters (and this is for the bivariate case).

Moreover, without further restrictions, there is no guarantee that the sequence of covariance matrices implied by an estimated process will be positive definite for all $t$.

Such conditions are very tedious to work out and to impose in estimation.

The system above is a bivariate version of the vec model, which is a straightforward generalization of univariate GARCH.

The general case is still useful, as it nests many more practicable specifications.
• The name derives from the fact that it uses the *vech operator*.

• As the $N \times N$ matrix $H_t$ is symmetric, it contains only $N(N+1)/2$ independent elements, which may be obtained, for example, by excluding the upper triangular (redundant) part.

• The *vech* operator then stacks the remaining elements columnwise into an $N(N+1)/2$ column vector, e.g.,

$$
\text{vech} \left( \begin{bmatrix} h_{1t}^2 & h_{12,t} \\ h_{12,t} & h_{2t}^2 \end{bmatrix} \right) = \begin{bmatrix} h_{1t}^2 \\ h_{12,t} \\ h_{2t}^2 \end{bmatrix}
$$

$$
\text{vech}(\epsilon_t \epsilon_t') = \text{vech} \left( \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix} \begin{bmatrix} \epsilon_{1t} & \epsilon_{2t} \end{bmatrix} \right)
$$

$$
= \text{vech} \left( \begin{bmatrix} \epsilon_{1t}^2 & \epsilon_{1t} \epsilon_{2t} \\ \epsilon_{1t} \epsilon_{2t} & \epsilon_{2t}^2 \end{bmatrix} \right) = \begin{bmatrix} \epsilon_{1t}^2 \\ \epsilon_{1t} \epsilon_{2t} \\ \epsilon_{2t}^2 \end{bmatrix}
$$

• The *vec* operator is similar, but without excluding the upper triangular part.
• Then the vec(1,1) model can be written

\[ h_t = c + A\eta_{t-1} + Bh_{t-1}, \]  \hspace{1cm} (14)

where

\[ h_t = \text{vech} H_t \]  \hspace{1cm} (15)

\[ \eta_t = \text{vech}(\epsilon_t\epsilon'_t). \]  \hspace{1cm} (16)

• Without restrictions, there are

- \( N(N + 1)/2 \) parameters in \( c \)
- \( N^2(N + 1)^2/4 \) parameters in \( A \)
- \( N^2(N + 1)^2/4 \) parameters in \( B \).
- With \( N = 2, 3, 5, 10 \) assets, we have 21, 78, 465, 6105 parameters.
Stationarity and Unconditional Variance

- The covariance stationarity for the vec(1,1) model (14),

\[ h_t = c + A\eta_{t-1} + Bh_{t-1}, \]  

(17)

requires the eigenvalues of matrix

\[ Q = A + B \]

to be inside the unit circle.

- If this holds, the unconditional covariance matrix (its vech) can be obtained by taking expectations on both sides of (17),

\[
\begin{align*}
\mathbb{E}(h_t) &= c + A\mathbb{E}(\eta_{t-1}) + B\mathbb{E}(h_{t-1}) \\
&= c + A\mathbb{E}(h_{t-1}) + B\mathbb{E}(h_{t-1}) \\
&= c + (A + B)\mathbb{E}(h_t),
\end{align*}
\]
hence
\[ E(\text{vech } H_t) = E(h_t) = (I - A - B)^{-1} c. \]

- Covariance matrix forecasts:

\[
\begin{align*}
    h_{t+1} &= c + A\eta_t + B h_t \\
    E_t(h_{t+2}) &= c + A E_t\eta_{t+1} + B h_{t+1} = c + (A + B)h_{t+1} \\
    E_t(h_{t+3}) &= c + A E_t\eta_{t+2} + B E_t h_{t+2} \\
    &= c + (A + B)E_t h_{t+2} = c + (A + B)c + (A + B)^2 h_{t+1} \\
    &\vdots \\
    E_t(h_{t+\tau}) &= \sum_{i=0}^{\tau-2} (A + B)^i c + (A + B)^{\tau-1} h_{t+1} \\
    &= E(h_t) + (A + B)^{\tau-1}(h_{t+1} - E(h_t)),
\end{align*}
\]

using
\[
\sum_{i=0}^{\tau-2} (A + B)^i = [I - (A + B)^{\tau-1}](I - A - B)^{-1}.
\]
• $E_t(h_{t+\tau})$ converges to the unconditional covariance matrix provided the covariance stationarity condition is satisfied.

• Calculation of higher moments of the vec model is considerably more involved than in the univariate GARCH model.\(^3\)

---

Special Case I: Diagonal VEC

• To reduce the number of parameters, this restricts the matrices $A$ and $B$ in (14) to be diagonal.

• This means that
  
  – each variance $h_{ii,t}^2$ depends only on its own past squared error $\epsilon_{i,t-1}^2$ and its own lag (as in the univariate case)
    \[ h_{ii,t}^2 = c_{ii} + a_{ii}\epsilon_{i,t-1}^2 + b_{ii}h_{i,i,t-1}^2, \quad i = 1, \ldots, N, \]  
    (18)
  
  – each covariance $h_{ij,t}$ depends only on its own past cross–product of errors $\epsilon_{i,t-1}\epsilon_{j,t-1}$ and its own lag,
    \[ h_{ij,t} = c_{ij} + a_{ij}\epsilon_{i,t-1}\epsilon_{j,t-1} + b_{ij}h_{ij,t-1}, \quad i, j = 1, \ldots, N. \]  
    (19)

• Often this specification is sufficient to represent the dynamics of variances and covariances.
• However, it does not allow for volatility transmissions, so not suitable for this kind of application.

• With $N = 2, 3, 5, 10$ assets, we have 9, 18, 45, 165 parameters.

• Even in the diagonal vec model, conditions for positive definiteness are difficult to check and impose in estimation.

• Methods for doing so and applying the model to a large number of assets are discussed in Ledoit et al. (2003).\textsuperscript{4} and Ding and Engle (2001).\textsuperscript{5}


Special Case II: BEKK

- BEKK (Baba, Engle, Kraft, and Kroner) was suggested by Engle and Kroner (1995).\(^6\)

- This specifies, in its simplest form,

\[
H_t = \tilde{C}^* \tilde{C}^* + A^* \epsilon_{t-1} \epsilon_{t-1}' A^* + B^* H_{t-1} B^* ',
\]

where \( \tilde{C} \) is a triangular matrix and \( A^* \) and \( B^* \) are \( N \times N \) parameter matrices.

- This guarantees positive definiteness if the initialization of \( H_t \) is positive definite.

- So the number of parameters is \( N(5N + 1)/2 \), i.e., for \( N = 2, 3, 5, 10 \) assets, we have 11, 24, 65, 255 parameters.

• To see that this is a restricted vec model, consider the case $N = 2$, where

$$
\begin{bmatrix}
  h_{1t}^2 & h_{12,t} \\
  h_{12,t} & h_{2,t}^2
\end{bmatrix} = \begin{bmatrix}
  c_{11}^* & 0 \\
  c_{21}^* & c_{22}^*
\end{bmatrix} \begin{bmatrix}
  c_{11} & c_{12}^* \\
  0 & c_{22}^*
\end{bmatrix}
$$

$$
\begin{bmatrix}
  a_{11}^* & a_{12}^* \\
  a_{21}^* & a_{22}^*
\end{bmatrix} \begin{bmatrix}
  \epsilon_{1,t-1}^2 & \epsilon_{1,t-1} \epsilon_{2,t-1} \\
  \epsilon_{1,t-1} \epsilon_{2,t-1} & \epsilon_{2,t-1}^2
\end{bmatrix} \begin{bmatrix}
  a_{11} & a_{12}^* \\
  a_{21}^* & a_{22}^*
\end{bmatrix}
$$

$$
\begin{bmatrix}
  b_{11}^* & b_{12}^* \\
  b_{21}^* & b_{22}^*
\end{bmatrix} \begin{bmatrix}
  h_{1,t-1}^2 & h_{12,t-1} \\
  h_{12,t-1} & h_{2,t-1}^2
\end{bmatrix} \begin{bmatrix}
  b_{11} & b_{12} \\
  b_{21}^* & b_{22}^*
\end{bmatrix}
$$

or

$$
\begin{align*}
  h_{1,t}^2 &= c_1 + a_{11}^* \epsilon_{1,t-1}^2 + 2a_{11}^* a_{12}^* \epsilon_{1,t-1} \epsilon_{2,t-1} + a_{12}^* \epsilon_{2,t-1}^2 \\
  &\quad + b_{11}^* h_{1,t-1}^2 + 2b_{11}^* b_{12}^* h_{12,t-1} + b_{12}^* h_{2,t-1}^2 \\
  h_{12,t} &= c_2 + a_{11}^* a_{21}^* \epsilon_{1,t-1}^2 + (a_{11}^* a_{22}^* + a_{21}^* a_{12}^*) \epsilon_{1,t-1} \epsilon_{2,t-1} + a_{22}^* a_{12}^* \epsilon_{2,t-1}^2 \\
  &\quad + b_{11}^* b_{21}^* h_{1,t-1}^2 + (b_{11}^* b_{22}^* + b_{12}^* b_{21}^*) h_{12,t-1} + b_{22}^* b_{12}^* h_{2,t-1}^2.
\end{align*}
$$
• For the general relation between the models, the Kronecker product $\otimes$ turns out to be useful.

• For an $m \times n$ matrix $A$ and an $p \times q$ matrix $B$, this is defined as the $mp \times nq$ matrix

\[
A \otimes B = \begin{bmatrix}
a_{11}B & a_{12}B & \cdots & a_{1n}B \\
a_{21}B & a_{22}B & \cdots & a_{2n}B \\
\vdots & \vdots & \ddots & \vdots \\
a_{m1}B & a_{m2}B & \cdots & a_{mn}B
\end{bmatrix}.
\]

• Important rule in time series analysis:

\[
\text{vec}(ABC') = (C' \otimes A)\text{vec}(B).
\]

• Then (20) can be written as

\[
\text{vec}(H_t) = \text{vec}(\tilde{C}^*\tilde{C}^*) + (A^* \otimes A^*)\text{vec}(\epsilon_{t-1}\epsilon_{t-1}') + (B^* \otimes B^*)\text{vec}(H_{t-1}).
\] (21)
• Representation (21) directly leads to stationarity conditions and covariance matrix forecasts for the BEKK model. E.g., covariance stationarity requires the eigenvalues of

\[ A^* \otimes A^* + B^* \otimes B^* \]  \hspace{1cm} (22)

to be smaller than one in magnitude.

• In practice, the diagonal BEKK model is sometimes used to further reduce the number of parameters to be estimated, where the parameter matrices \( A^* \) and \( B^* \) are diagonal.
Factor Models

• Basic idea: Co-movements of returns are driven by a small number of (observable or unobservable) common underlying variables, which are called factors.

• For example, as an observable factor, the return of a market index may be used as a proxy for the general tendency of the stock market.

• Consider the simplest case of just a single observable factor.

• Think of this as the market return, denoted by $r_{Mt}$.

• In portfolio analysis, where factor models are often used to structure covariance matrices, the model is also known as single index model (SIM).
• The return of asset $i$, $i = 1, \ldots, N$, is described by

\begin{align*}
  r_{it} &= \alpha_i + \beta_i r_{Mt} + \epsilon_{it}, \quad i = 1, \ldots, N; \\
  E(\epsilon_{it}) &= 0, \quad \text{Var}_{t-1}(\epsilon_{it}) = \sigma_{\epsilon_i}^2, \quad i = 1, \ldots, N; \\
  \text{Cov}_{t-1}(\epsilon_{it}, \epsilon_{jt}) &= 0, \quad i \neq j.
\end{align*}

(23) (24) (25)

• Expected return and variance of the market return will be denoted by $E_{t-1}(r_{Mt}) = \mu_{Mt}$ and $\text{Var}_{t-1}(r_{Mt}) = \sigma_{Mt}^2$, and we assume

\begin{align*}
  \text{Cov}_{t-1}(r_{Mt}, \epsilon_{it}) &= 0, \quad i = 1, \ldots, N.
\end{align*}

(26)

• This structure implies that

\begin{align*}
  E_{t-1}(r_{it}) &= \alpha_i + \beta_i \mu_{Mt}, \quad i = \ldots, N, \\
  \text{Var}_{t-1}(r_{it}) &= \beta_i^2 \sigma_{Mt}^2 + \sigma_{\epsilon_i}^2, \quad i = 1, \ldots, N, \\
  \text{Cov}_{t-1}(r_{it}, r_{jt}) &= \beta_i \beta_j \sigma_{Mt}^2, \quad i, j = 1, \ldots, N, \quad i \neq j.
\end{align*}

(27) (28) (29)
• For the covariance structure of the returns, given by (29), Assumption (25) is crucial, as it implies that the only reason for asset $i$ and asset $j$ moving together is their joint dependence on the market return $r_{Mt}$.

• The first part of (28) is also often referred to as the *systematic* risk (which is related to the general tendency of the market), whereas the second part is the *unsystematic* (idiosyncratic, specific) risk, which is not related to market factors.
• In contrast to the market–related, systematic risk, the specific risk can be diversified away.

• Consider an equally, weighted portfolio, i.e., a portfolio with weights \( w_i = 1/N, \ i = 1, \ldots, N \).

• Then the portfolio variance is, assuming the SIM correctly describes the covariance structure,

\[
\sigma^2_{pt} = \frac{1}{N^2} \sum_{i=1}^{N} (\beta_i^2 \sigma^2_{Mt} + \sigma^2_{\epsilon_i}) + \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j \neq i} \beta_i \beta_j \sigma^2_{Mt} \\
= \left( \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \beta_i \beta_j \right) \sigma^2_{Mt} + \frac{1}{N^2} \sum_{i=1}^{N} \sigma^2_{\epsilon_i} \\
= \left( \frac{1}{N} \sum_{i=1}^{N} \beta_i \right)^2 \sigma^2_{Mt} + \frac{1}{N^2} \sum_{i=1}^{N} \sigma^2_{\epsilon_i}.
\]
• Now
\[
\frac{1}{N^2} \sum_{i=1}^{N} \sigma_{\epsilon_i}^2 \leq \frac{\max\{\sigma_{\epsilon_1}^2, \ldots, \sigma_{\epsilon_N}^2\}}{N} \rightarrow \infty 0,
\]
provided the variances of the unsystematic risks are bounded.

• Hence, for large \(N\),
\[
\sigma_{pt}^2 \approx \left( \frac{1}{N} \sum_{i=1}^{N} \beta_i \right)^2 \sigma_{Mt}^2 = \overline{\beta}_p^2 \sigma_{Mt}^2,
\]
where
\[
\overline{\beta}_p = \frac{1}{N} \sum_{i=1}^{N} \beta_i
\]
is the portfolio’s \(\beta\).

• That is, the market risk cannot be diversified away.
• The conditional variance of the market factor can be modeled by means of a univariate (asymmetric) (E)GARCH model, e.g.,

\[
\sigma_{Mt}^2 = c + a\epsilon_{M,t-1}^2 + b\sigma_{Mt-1}^2,
\]

where

\[
\epsilon_{Mt} = r_{Mt} - \mu_{Mt}.
\]

• Equation (28) implies that the GARCH effects in the market will be transferred to all the assets’ variances.
• Defining

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_N \end{bmatrix}, \quad \Sigma_\epsilon = \begin{bmatrix} \sigma_{\epsilon_1}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{\epsilon_2}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{\epsilon_N}^2 \end{bmatrix},$$

the conditional covariance matrix of the $N$–dimensional $r_t = [r_{1t}, r_{2t}, \ldots, r_{Nt}]'$ can be written as

$$\text{Cov}_{t-1}(r_t) = \begin{bmatrix} \beta_1^2 \sigma_{Mt}^2 + \sigma_{\epsilon_1}^2 & \beta_1 \beta_2 \sigma_{Mt}^2 & \cdots & \beta_1 \beta_N \sigma_{Mt}^2 \\ \beta_1 \beta_2 \sigma_{Mt}^2 & \beta_2^2 \sigma_{Mt}^2 + \sigma_{\epsilon_2}^2 & \cdots & \beta_2 \beta_N \sigma_{Mt}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1 \beta_N \sigma_{Mt}^2 & \beta_2 \beta_N \sigma_{Mt}^2 & \cdots & \beta_N^2 \sigma_{Mt}^2 + \sigma_{\epsilon_N}^2 \end{bmatrix} = \beta \beta' \sigma_{Mt}^2 + \Sigma_\epsilon.$$
The single factor model can be written as

\[ r_t = \alpha + \beta f_t + \epsilon_t, \]

where \( f_t \) is the factor.

In the \( k \)-factor case, \( f_t = [f_{1t}, f_{2t}, \ldots, f_{kt}]' \), and

\[ r_t = \alpha + B f_t + \epsilon_t, \]

where \( B \) is a \( N \times k \) matrix of factor loadings.

The conditional covariance matrix of the return vector is

\[ \text{Cov}_{t-1}(r_t) = B \Sigma_{f_t} B' + \Sigma_{\epsilon}, \]

where \( \Sigma_{f_t} \) is the conditional covariance matrix of the risk factors, which may be specified as a low-dimensional multivariate GARCH process.

The BEKK or diagonal vec may be appropriate in this framework.
Modeling Conditional Correlations

- The models considered so far specified models for the conditional covariances, in addition to the variances.

- Another approach is to model the variances and the conditional correlations.

- One advantage is that conditional variances (or standard deviations) and conditional correlations can be modeled separately, which often allows for consistent two–step model estimation, thus reducing the computational burden.

- For these models, we write $H_t$ as

\[
H_t = D_t R_t D_t \tag{32}
\]

\[
D_t = \begin{bmatrix}
\sqrt{h_{1t}^2} & 0 & \cdots & 0 \\
0 & \sqrt{h_{2t}^2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \sqrt{h_{Nt}^2}
\end{bmatrix} \tag{33}
\]
i.e., $H_t$ is a diagonal matrix with the conditional standard deviations on its main diagonal, and

$$
R_t = \begin{bmatrix}
1 & \rho_{12,t} & \cdots & \rho_{1N,t} \\
\rho_{12,t} & 1 & \cdots & \rho_{2N,t} \\
\vdots & \vdots & \ddots & \vdots \\
\rho_{1N,t} & \rho_{2N,t} & \cdots & 1
\end{bmatrix}
$$

(34)

is the conditional correlation matrix, i.e.,

$$
\rho_{ij,t} = \text{Corr}_{t-1}(\epsilon_{it}, \epsilon_{jt}), \quad i, j = 1, \ldots, N, \quad i \neq j,
$$

is the conditional correlation between assets $i$ and $j$.

- The conditional covariances are

$$
h_{ij,t} = \rho_{ij,t} \sqrt{h_{it}^2 h_{jt}^2}, \quad i \neq j.
$$

- Positive definiteness of $H_t$ follows from that of $R_t$ and the positivity of the conditional standard deviations in $D_t$. 
Constant Conditional Correlations (CCC)

- One of the first multivariate GARCH models (Bollerslev, 1990).\(^7\)
- In this model \(R_t = R\) is constant in (32), i.e., the conditional correlations are constant.
- We may write this as
  \[
  \epsilon_t = D_t z_t, \tag{35}
  \]
  where \(\{ (z_{1t}, \ldots, z_{Nt})' \}\) is an iid series of (e.g., normally distributed) innovations with mean zero and covariance matrix \(R\), i.e.,
  \[
  z_t \sim \mathcal{N}(0, R). \tag{36}
  \]
- For some time, this has been the most popular multivariate GARCH model due to the fact that it can easily be estimated even for high-dimensional time series.

• Note that $R$ is the constant *conditional* correlation matrix (i.e., the correlation matrix of the innovations), not the unconditional correlation matrix of the returns.

• Consistent two–step estimation for high–dimensional time series feasible:

• First estimate univariate GARCH models for each series.

• This allows for flexible specification of the univariate processes. For example, we may specify a standard GARCH for one component, AGARCH or EGARCH for another...

• Calculate the standardized residuals,

$$
\hat{z}_{it} = \frac{\epsilon_{it}}{\sqrt{\hat{h}_{it}^2}}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T. \quad (37)
$$

• Then, in view of (35), estimate $R$ as the correlation matrix of the standardized residuals (37).
Dynamic Conditional Correlation (DCC) Models

- The two-step estimation procedure makes application of the CCC to high-dimensional systems feasible, but more often than not the hypothesis of constant conditional correlations is rejected.

- For example, it is often observed that correlations between financial time series increase in turbulent periods, and are very high in crash situations.

- Thus models for dynamic conditional correlations (DCC) have been proposed.

- As an example, consider the model proposed by Engle (2002).\(^8\)

---

• In its simplest (scalar) form, this can be written as

\[
\epsilon_t \sim N(0, D_t R_t D_t), \quad (38)
\]

\[
D_t \sim \text{GARCH} \quad (39)
\]

\[
z_t = D^{-1} \epsilon_t \quad (\text{produces standardized residuals (37)})
\]

\[
Q_t = (1 - a - b)S + az_t z_t'_{t-1} + bQ_{t-1}, \quad (40)
\]

\[
a, b \geq 0, \quad a + b < 1,
\]

\[
R_t = \{\text{diag}(Q_t)\}^{-1/2} Q_t \{\text{diag}(Q_t)\}^{-1/2}. \quad (41)
\]

• In (40), \(S\) is the unconditional correlation matrix of the standardized residuals \(z_t\).

• If the starting value for \(Q_t\) in (40) is positive definite, then \(Q_t\) is positive definite, but will not in general be a valid correlation matrix (i.e., with ones on the diagonal).

• Thus, the rescaling in (41) is necessary.