Multivariate Volatility

Gerald P. Dwyer

Clemson University

February 2021

Outline

1. Multivariate GARCH
   - General
   - VECH GARCH
   - Diagonal VECH GARCH
   - BEKK GARCH
   - Constant conditional-correlation GARCH
   - Dynamic conditional-correlation GARCH
   - Summary

Multivariate GARCH

- Multivariate return series

$$\mathbf{y}_t = \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t$$

where the vectors are simple generalizations of a univariate process

- The vectors are m by one with m dependent variables, often asset returns
- $\boldsymbol{\mu}_t = \mathrm{E}\left[\boldsymbol{y}_t | F_{t-1}\right]$
- $\boldsymbol{\varepsilon}_t$ is the innovation in the returns in period t with $\mathrm{E}\,\boldsymbol{\varepsilon}_t = \mathbf{0}$ and

$$\boldsymbol{\Sigma}_t = \mathrm{Cov}\left[\boldsymbol{\varepsilon}_t | F_{t-1}\right]$$

  - $\boldsymbol{\Sigma}_t$ is m by m matrix with $m(m+1)/2$ distinct elements
- The number of distinct elements in $\boldsymbol{\Sigma}_t$ increases with the square of m because there are $\left(m^2 + m\right)/2$ distinct elements
  - The number of observations is mT and increases linearly with m

Curse of Dimensionality

Estimation of all parameters unconstrained not feasible for large systems

- The implication of the number of parameters in $\boldsymbol{\Sigma}_t$ increasing with the square of m
  - Five assets implies $m(m+1)/2 = 15$ distinct elements
  - Twenty assets implies $m(m+1)/2 = 210$ distinct elements
  - One hundred assets implies $m(m+1)/2 = 5,050$ distinct elements
- Suppose have 250 observations on each asset
  - Five assets implies $1,250$ observations
  - Twenty assets implies $5,000$ observations
  - One hundred assets implies $25,000$ observations
- Even one hundred assets is not a lot
- If have 1,000 assets
  - Then 500,500 distinct elements in the covariance matrix
  - Only 250,000 observations
  - May make more sense to look at volatility of portfolio in some circumstances, but not all

Constrained systems

- Want to find a way to allow for correlations but have sufficiently few parameters they can be estimated with reasonable precision

Constrained systems

- Want to find a way to allow for correlations but have sufficiently few parameters they can be estimated with reasonable precision
- And have nonlinear estimation converge

Strategies to estimate multivariate volatility models

- Exponential smoothing estimate
  - Includes all the parameters in the variance-covariance matrix
  - Adds only one additional parameter to get a time-varying variance-covariance matrix
- Multivariate GARCH models
  - VECH GARCH – Everything depends on everything
  - Diagonal VECH GARCH – GARCH(1,1) for each term in variance-covariance matrix
- Reparametrizations
  - Decomposition to ensure positivity of estimated variance-covariance matrix
- Restrictions to fewer parameters and limited interaction
  - Constant correlation model
  - Dynamic conditional correlation model

Exponential smoothing estimate

- Exponential smoothing forecasts are based on simple forecasts that smooth values of the data
  - E.g. the forecast at $t-1$ of the value of a series x_t is ${}_{t-1}f_t = (1-\lambda)\,x_{t-1} + \lambda\ {}_{t-2}f_{t-1}$
  - where ${}_{t-1}f_t$ is the forecast at $t-1$ for period t
- Apply this idea to the variance-covariance matrix
  - Estimate time-varying variance-covariance matrix as

$$\widehat{\boldsymbol{\Sigma}}_t = (1-\lambda)\,\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}'_{t-1} + \lambda\widehat{\boldsymbol{\Sigma}}_{t-1}$$

    with $\lambda < 1$
  - Pick a starting value and begin to forecast
  - Could pick initial value at $t=1$ of $\boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}'_1$ as starting $\widehat{\boldsymbol{\Sigma}}_1 = \boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}'_1$

VECH GARCH

- What do vec and vech mean?
- Stacking of matrix elements in a vector by stacking columns

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \text{vec} \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

- Symmetric matrix

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \text{vech} \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

VECH GARCH

- Conditional variances and conditional covariance depend on all innovations and cross-products
- GARCH two-variable example with one lag of innovations squared, one lagged covariance

VECH GARCH

- VECH GARCH with two variables
- Mean equations

$$y_{1,t} = \alpha_{10} + \varepsilon_{1,t}, \quad \varepsilon_{1,t} = \sigma_{1,t}\nu_{1,t}, \quad \nu_{1,t} \sim \text{iid}\,(0,1)$$
$$y_{2,t} = \alpha_{20} + \varepsilon_{2,t}, \quad \varepsilon_{2,t} = \sigma_{1,t}\nu_{2,t}, \quad \nu_{2,t} \sim \text{iid}\,(0,1)$$

- where αs are parameters, $\sigma_{1,t} = \text{SD}\left[\varepsilon_{1,t}|F_{t-1}\right]$ and $\sigma_{2,t} = \text{SD}\left[\varepsilon_{2,t}|F_{t-1}\right]$

VECH GARCH

- Conditional variances and conditional covariance depend on all innovations and cross-products
- GARCH two-variable example with one lag of innovations squared, one lagged covariance

VECH GARCH

- VECH GARCH with two variables

$$\begin{aligned}
\sigma_{11,t} &= \gamma_{10} + \gamma_{11}\varepsilon^2_{1,t-1} + \gamma_{12}\varepsilon^2_{2,t-1} + \gamma_{13}\varepsilon_{1,t-1}\varepsilon_{2,t-1} \\
&\quad + \delta_{11}\sigma_{11,t-1} + \delta_{12}\sigma_{22,t-1} + \delta_{13}\sigma_{12,t-1} \\
\sigma_{22,t} &= \gamma_{20} + \gamma_{21}\varepsilon^2_{1,t-1} + \gamma_{22}\varepsilon^2_{2,t-1} + \gamma_{23}\varepsilon_{1,t-1}\varepsilon_{2,t-1} \\
&\quad + \delta_{21}\sigma_{11,t-1} + \delta_{22}\sigma_{22,t-1} + \delta_{23}\sigma_{12,t-1} \\
\sigma_{12,t} &= \gamma_{30} + \gamma_{31}\varepsilon^2_{1,t-1} + \gamma_{32}\varepsilon^2_{2,t-1} + \gamma_{33}\varepsilon_{1,t-1}\varepsilon_{2,t-1} \\
&\quad + \delta_{31}\sigma_{11,t-1} + \delta_{32}\sigma_{22,t-1} + \delta_{33}\sigma_{12,t-1}
\end{aligned}$$

  - where γs and δs are parameters
  - $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are the innovations in equations
  - $\sigma_{11,t} = \text{Var}\left[\varepsilon_{1,t} | F_{t-1}\right]$, $\sigma_{22,t} = \text{Var}\left[\varepsilon_{2,t} | F_{t-1}\right]$ and $\sigma_{12,t} = \text{Cov}\left[\varepsilon_{1,t}, \varepsilon_{2,t} | F_{t-1}\right]$
- Everything depends on everything

Problems with VECH GARCH

- VECH GARCH

$$\begin{aligned}
\sigma_{11,t} &= \gamma_{10} + \gamma_{11}\varepsilon^2_{1,t-1} + \gamma_{12}\varepsilon^2_{2,t-1} + \gamma_{13}\varepsilon_{1,t-1}\varepsilon_{2,t-1} \\
&\quad + \delta_{11}\sigma_{11,t-1} + \delta_{12}\sigma_{22,t-1} + \delta_{13}\sigma_{12,t-1} \\
\sigma_{22,t} &= \gamma_{20} + \gamma_{21}\varepsilon^2_{1,t-1} + \gamma_{22}\varepsilon^2_{2,t-1} + \gamma_{23}\varepsilon_{1,t-1}\varepsilon_{2,t-1} \\
&\quad + \delta_{21}\sigma_{11,t-1} + \delta_{22}\sigma_{22,t-1} + \delta_{23}\sigma_{12,t-1} \\
\sigma_{12,t} &= \gamma_{30} + \gamma_{31}\varepsilon^2_{1,t-1} + \gamma_{32}\varepsilon^2_{2,t-1} + \gamma_{33}\varepsilon_{1,t-1}\varepsilon_{2,t-1} \\
&\quad + \delta_{31}\sigma_{11,t-1} + \delta_{32}\sigma_{22,t-1} + \delta_{33}\sigma_{12,t-1}
\end{aligned}$$

- Curse of dimensionality will overtake this quickly
- No guarantee of estimated positive variance-covariance matrix every period

Curse of Dimensionality with GARCH

- In a VECH GARCH(1,1) in which all variances and covariances depend on all lagged variances and cross-products of errors
  - The number of parameters in each equation is $3m + 1$
  - The number of equations is $m(m+1)/2$
  - The total number of parameters is $[m(m+1)/2](3m+1)$
  - Two assets implies 21 parameters
  - Five assets implies 240 parameters
  - Ten assets implies 1705 parameters

Diagonal VECH GARCH

- The Diagonal VECH GARCH model has each term in the variance-covariance matrix evolve independently according to a GARCH(1,1)
- Two-variable example

$$
\begin{aligned}
\sigma_{11,t} &= \gamma_{10} + \gamma_{11}\varepsilon_{1,t-1}^2 + \delta_{11}\sigma_{11,t-1} \\
\sigma_{22,t} &= \gamma_{20} + \gamma_{22}\varepsilon_{2,t-1}^2 + \delta_{22}\sigma_{22,t-1} \\
\sigma_{12,t} &= \gamma_{30} + \gamma_{33}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + \delta_{33}\sigma_{12,t-1}
\end{aligned}
$$

- Far fewer parameters but no interaction of volatility across assets
- Still no guarantee that the estimated variance-covariance matrix will be positive definite in every period

Use of squares of underlying matrices in computations can guarantee positive definiteness

- Baba-Engle-Kraft-Kroner (BEKK) GARCH Model
- We want the estimated variance-covariance matrix to be positive definite every period
  - Models such as the Diagonal VECH GARCH model with each term in the variance-covariance matrix evolving independently according to a GARCH(1,1) will not necessarily satisfy positive definiteness
- BEKK: Ensure positive definiteness of the covariance matrix every period by
  - Restricting the estimation problem using a diagonal matrix which is squared to generate constant terms
  - Restricting coefficient matrices to be squares of underlying matrices estimated

Constant conditional-correlation GARCH model

- Constant conditional-correlation GARCH

$$\rho = \frac{\text{Cov}\left[\varepsilon_{1t}, \varepsilon_{2t}\right]}{\text{SD}\left[\varepsilon_{1t}\right]\text{SD}\left[\varepsilon_{2t}\right]} = \frac{\sigma_{12,t}}{\sqrt{\sigma_{11,t}\sigma_{22,t}}}$$

  - A constant correlation is not an obvious constraint to impose with time-varying variances
    - Implies that the covariance varies proportionately to the standard deviations of the innovations
    - Does allow for multivariate aspect of process

- A two-variable Constant conditional-correlation GARCH model in matrix form is

$$\begin{bmatrix} \sigma_{11,t} \\ \sigma_{22,t} \end{bmatrix} = \begin{bmatrix} \gamma_{10} \\ \gamma_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} \varepsilon^2_{1,t-1} \\ \varepsilon^2_{2,t-1} \end{bmatrix} + \begin{bmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{bmatrix} \begin{bmatrix} \sigma_{11,t-1} \\ \sigma_{22,t-1} \end{bmatrix}$$

Constant conditional-correlation GARCH model

- The Constant conditional-correlation GARCH model

$$
\begin{aligned}
\sigma_{11,t} &= \gamma_{10} + \gamma_{11}\varepsilon^2_{1,t-1} + \gamma_{12}\varepsilon^2_{2,t-1} + \delta_{11}\sigma_{11,t-1} + \delta_{12}\sigma_{22,t-1} \\
\sigma_{22,t} &= \gamma_{20} + \gamma_{21}\varepsilon^2_{1,t-1} + \gamma_{22}\varepsilon^2_{2,t-1} + \delta_{21}\sigma_{11,t-1} + \delta_{22}\sigma_{22,t-1} \\
\sigma_{12,t} &= \rho\left(\sigma_{11,t}\sigma_{22,t}\right)^{1/2}
\end{aligned}
$$

- Do not need an equation for covariance because its value each period is implied by the estimated constant correlation
- Bollerslev determined the conditions necessary for this model to be stationary
- Quasi-maximum likelihood estimator is consistent under certain conditions and the estimators are asymptotically normally distributed

Dynamic conditional-correlation GARCH model

- A constant conditional-correlation GARCH model

$$\sigma_{11,t} = \gamma_{10} + \gamma_{11}\varepsilon^2_{1,t-1} + \gamma_{12}\varepsilon^2_{2,t-1} + \delta_{11}\sigma_{11,t-1} + \delta_{12}\sigma_{22,t-1}$$
$$\sigma_{22,t} = \gamma_{20} + \gamma_{21}\varepsilon^2_{1,t-1} + \gamma_{22}\varepsilon^2_{2,t-1} + \delta_{21}\sigma_{11,t-1} + \delta_{22}\sigma_{22,t-1}$$
$$\sigma_{12,t} = \rho\left(\sigma_{11,t}\sigma_{22,t}\right)^{1/2}$$

- A simple dynamic conditional-correlation GARCH model (Engle 2002)

$$\sigma_{11,t} = \gamma_{10} + \gamma_{11}\varepsilon^2_{1,t-1} + \delta_{11}\sigma_{11,t-1}$$
$$\sigma_{22,t} = \gamma_{20} + \gamma_{22}\varepsilon^2_{2,t-1} + \delta_{22}\sigma_{22,t-1}$$
$$\sigma_{12,t} = \rho_{12,t}\left(\sigma_{11,t}\sigma_{22,t}\right)^{1/2}$$

- Will show that the conditional dynamic correlation is the covariance of the standardized innovations
- Might specify a GARCH process for this covariance

Standardized innovations and the relationship between conditional correlation and conditional covariance I

- As above, examine the problem with two series, $i = 1, 2$

$$
\begin{aligned}
y_{i,t} &= \mu_{i,t} + \varepsilon_{i,t} \\
\varepsilon_{i,t} &= \sigma_{i,t}\nu_{i,t} \\
\nu_{i,t} &\sim \text{iid}\,(0,1) \\
\mathrm{E}_{t-1}\,\nu_{i,t} &= 0 \\
\mathrm{E}_{t-1}\,\nu_{i,t}^2 &= 1 \\
\mathrm{E}\,\nu_{1,t}\nu_{1,t} &= \rho_{12}^{\nu}
\end{aligned}
$$

- To simplify notation, I will use the convention that $\sigma_{i,t} = \sigma_{ii,t}^{(1/2)}$
- I also will use $\mathrm{E}\left[x_t | F_{t-1}\right] = \mathrm{E}_{t-1}\,x_{t-1}$
- The conditional correlation of the two series is

$$
\rho_{12,t} = \frac{\mathrm{Cov}_{t-1}\left[y_{1,t}, y_{2,t}\right]}{\mathrm{SD}_{t-1}\left[y_{1,t}\right]\mathrm{SD}_{t-1}\left[y_{2,t}\right]}
$$

Standardized innovations and the relationship between conditional correlation and conditional covariance II

- For simplicity, suppose that

$$\mu_{i,t} = 0$$

which implies

$$y_{i,t} = \varepsilon_{i,t}$$

- The conditional correlation of the two series is

$$\rho_{12,t} = \frac{\mathrm{Cov}_{t-1}\left[\varepsilon_{1,t}, \varepsilon_{2,t}\right]}{\mathrm{SD}_{t-1}\left[\varepsilon_{1,t}\right]\mathrm{SD}_{t-1}\left[\varepsilon_{2,t}\right]}$$

which, because all innovations have zero mean, is

$$\rho_{12,t} = \frac{\mathrm{E}_{t-1}\left[\varepsilon_{1,t}\varepsilon_{2,t}\right]}{\left(\mathrm{E}_{t-1}\left[\varepsilon_{1,t}^2\right]\mathrm{E}_{t-1}\left[\varepsilon_{2,t}^2\right]\right)^{1/2}}$$

Standardized innovations and the relationship between conditional correlation and conditional covariance III

- Note that

$$\varepsilon_{i,t} = \sigma_{i,t} \nu_{i,t}$$

which implies

$$\begin{aligned}
\rho_{12,t} &= \frac{\mathrm{E}_{t-1}\left[\varepsilon_{1,t}\varepsilon_{2,t}\right]}{\left(\mathrm{E}_{t-1}\left[\varepsilon_{1,t}^2\right]\mathrm{E}_{t-1}\left[\varepsilon_{2,t}^2\right]\right)^{1/2}} \\
&= \frac{\mathrm{E}_{t-1}\left[\sigma_{1,t}\nu_{1,t}\sigma_{2,t}\nu_{2,t}\right]}{\left(\mathrm{E}_{t-1}\left[\sigma_{11,t}\nu_{1,t}^2\right]\mathrm{E}_{t-1}\left[\sigma_{22,t}\nu_{2,t}^2\right]\right)^{1/2}} \\
&= \frac{\mathrm{E}_{t-1}\left[\nu_{1,t}\nu_{2,t}\right]}{\left(\mathrm{E}_{t-1}\left[\nu_{1,t}^2\right]\mathrm{E}_{t-1}\left[\nu_{2,t}^2\right]\right)^{1/2}} \\
&= \mathrm{E}_{t-1}\left[\nu_{1,t}\nu_{2,t}\right]
\end{aligned}$$

Standardized innovations and the relationship between conditional correlation and conditional covariance IV

- This means that the conditional correlation between the unstandardized innovations is the same as the conditional covariance of underlying zero-mean unit-variance innovations
- It proves that $\rho_{12,t} = \sigma^{\nu}_{12,t}$ where $\sigma^{\nu}_{12,t}$ is the covariance of the standardized innovations ν
- The constant-conditional correlation GARCH model assumes that the underlying innovations have a constant correlation
- This need not be the case though
- One easily can imagine the correlation of the shocks changing over time
- The expression above shows that allowing the conditional correlation of the ε's to change is the same as allowing the conditional correlations between the ν's to change over time
- There are various ways of specifying this correlation and covariance

Standardized innovations and the relationship between conditional correlation and conditional covariance V

- Engle (2002) suggests specifying simple GARCH processes for the variances and the correlation
- For example, specify GARCH(1,1) models for the underlying variances and conditional correlation
- A simple GARCH process that could work for many assets is

$$\begin{aligned}
\sigma_{11,t} &= \gamma_{1,0} + \gamma_{1,1}\varepsilon^2_{1,t-1} + \delta_{1,1}\sigma_{11,t-1} \\
\sigma_{22,t} &= \gamma_{2,0} + \gamma_{2,1}\varepsilon^2_{2,t-1} + \delta_{2,1}\sigma_{22,t-1} \\
\sigma^{\nu}_{12,t} &= \gamma_{3,0} + \gamma_{3,1}\nu_{1,t-1}\nu_{2,t-1} + \delta_{3,1}\sigma^{\nu}_{12,t-1}
\end{aligned}$$

- The last equation is the same as the correlation because the variances of the standardized innovations are unity

Restriction of unconditional correlation to sample correlation

- Rather than let the implied unconditional (also "'long run") correlation float, the equation can be constrained so that the correlation in the data $\bar{\rho}_{12}$ is implied
- This requires modifying the third equation to

$$\sigma^{\nu}_{12,t} = \bar{\rho}_{12} + \gamma_{3,1}\left(\nu_{1,t-1}\nu_{2,t-1} - \bar{\rho}_{12}\right) + \gamma_{3,2}\left(\sigma^{\nu}_{12,t-1} - \bar{\rho}_{12}\right)$$

- This estimator has very nice properties
- In particular, a positive definite variance-covariance matrix
  - More complicated models are possible
  - The key is to keep the unconditional correlation equal to the sample correlation

Estimation of Dynamic Conditional Correlation GARCH model

- Engle shows that the model can be estimated consistently by maximum likelihood
- Other estimation methods are possible and relatively simple
- For example, estimate the GARCH equations for the variances first and then estimate the equation for the correlation
    - First estimate GARCH equations for variances by maximum likelihood
    - Then estimate GARCH equation for correlation by maximum likelihood conditional on the estimated GARCH equations for the variances
    - Engle (2002) shows that this estimator is consistent under standard conditions although it is inefficient
    - With this simple structure, the process is not subject to the curse of dimensionality that affects a general process

Summary I

- Multivariate GARCH is useful for a few assets but is problematic for very many assets without restrictions
- Restrictions for multivariate GARCH
- Diagonal process
- BEKK with a diagonal process to ensure positive definiteness
- Constant conditional-correlation multivariate GARCH
  - Issues with a constant correlation:
  - The correlations across assets change in financial crises for example
  - Common factors become and less more important
- Engle's dynamic conditional-correlation GARCH
  - Suppose that the changing correlation does not affect the variances
  - This makes it possible to estimate the GARCH equations for the variances first
  - Then estimate the GARCH equations for the correlations