Financial Econometrics

Multivariate Volatility

Gerald P. Dwyer, Jr.

Trinity College, Dublin

February 2009

Multivariate GARCH

- Multivariate return series

$$\mathbf{r}_t = \boldsymbol{\mu}_t + \mathbf{a}_t$$

where the vectors are simple generalizations of a univariate process

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

$$\Sigma_t = \mathrm{Cov}\left[\mathbf{a}_t | F_{t-1}\right]$$

  - Σ_t is m by m with $m(m+1)/2$ distinct elements

- The number of parameters in Σ_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 Σ_t increasing with the square of m
  - Five assets implies $m(m+1)/2 = 15$ parameters
  - Twenty assets implies $m(m+1)/2 = 210$ parameters
  - One hundred assets implies $m(m+1)/2 = 5,050$ parameters
- 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 parameters in the covariance matrix
  - Only 250,000 observations
  - Probably makes more sense to look at volatility of portfolio anyway in many circumstances

Constrained systems

- Want to find a way to allow for correlations but not have too many parameters estimated with poor estimates

Outline of chapter 10

- Exponentially 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
  - Diagonal VEC model
    - GARCH(1,1) for each term in variance-covariance matrix
- Reparameterizations
  - Replace covariances by correlations
  - Cholesky decomposition
- GARCH models for bivariate returns
  - Constant correlation models
  - Time-varying correlation models
- Higher-dimensional volatility models
- Factor volatility models

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

$$\widehat{\Sigma}_t = \frac{1-\lambda}{1-\lambda^{t-1}} \sum_{j=1}^{t-1} \lambda^{j-1} \mathbf{a}_{t-j}\mathbf{a}'_{t-j}, \quad 0<\lambda<1$$

  - The weights satisfy $\sum_{j=1}^{t-1} \frac{(1-\lambda)\lambda^{j-1}}{1-\lambda^{t-1}} = 1$ by construction
  - As $t \to \infty$, $\widehat{\Sigma}_t \to (1-\lambda)\,\mathbf{a}_{t-1}\mathbf{a}'_{t-1} + \lambda\widehat{\Sigma}_{t-1}$

Estimation of exponential smoothing estimate

- Can estimate parameters by maximum likelihood
  - Likelihood function assuming an underlying normal distribution is

$$\ln L\left(\Theta, \boldsymbol{\lambda}\right) \propto -\frac{1}{2}\sum_{t=1}^{T}\ln\left|\Sigma_t\right| - \frac{1}{2}\sum_{t=1}^{T}\left(\mathbf{r}_t - \boldsymbol{\mu}_t\right)\Sigma_t^{-1}\left(\mathbf{r}_t - \boldsymbol{\mu}_t\right)'$$

  - Don't know value of Σ_t
  - Know value of $\widehat{\Sigma}_t$
  - Substitute $\widehat{\Sigma}_t$ for Σ_t as compute sums
    - Given the tentative estimate of λ in the maximization
    - Given an initial value of the variance-covariance matrix, $\widehat{\Sigma}_0$

Diagonal VEC GARCH model

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

$$\Sigma_t = \begin{bmatrix} \sigma_{11,t} & \sigma_{12,t} \\ \sigma_{21,t} & \sigma_{22,t} \end{bmatrix}$$

  - with the covariance terms $\sigma_{12} = \sigma_{21}$
  - Let As and Bs represent parameters

$$\begin{aligned} \sigma_{11,t} &= A_{11,0} + A_{11,1} a^2_{1,t-1} + B_{11,1}\sigma_{11,t-1} \\ \sigma_{12,t} &= A_{12,0} + A_{12,1} a_{1,t-1} a_{2,t-1} + B_{12,1}\sigma_{12,t-1} \\ \sigma_{22,t} &= A_{22,0} + A_{22,1} a^2_{2,t-1} + B_{22,1}\sigma_{22,t-1} \end{aligned}$$

- Problem with this model: There is no guarantee that the estimated variance-covariance matrix will be positive definite in every period

Use of Cholesky decomposition in computations can guarantee positive definiteness

- We want the estimated Σ_t to be positive definite every period
  - Models such as the Diagonal VEC GARCH model with each term in the variance-covariance matrix evolving independently according to a GARCH(1,1) will not necessarily satisfy positive definiteness
- Transforming the estimation problem using a Cholesky decomposition ensures positive definiteness of the covariance matrix every period
  - Interesting, illuminating discussion in Tsay
  - Not sure it matters unless you are writing a program to estimate GARCH models or using a program that doesn't use this algorithm

Constant-correlation GARCH model

- Constant-correlation model
  - 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
    - This implication can be seen from

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

    - Does allow for multivariate aspect of process
  - A two-variable GARCH model is

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

  - Do not need a term for covariance because its value each period is implied by the estimated constant correlation

Constant-correlation GARCH model

- In matrix notation, the constant correlation GARCH model is

$$\Xi_t^* = \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1 \mathbf{a}_{t-1}^2 + \boldsymbol{\beta}_1 \Xi_{t-1}^*$$

where $\Xi_t^{*\prime} = \left[\begin{array}{cc} \sigma_{11,t} & \sigma_{22,t} \end{array} \right]$

- In non-matrix notation, the constant-correlation model can be written

$$\begin{aligned} \sigma_{11,t} &= \alpha_{10} + \alpha_{11} a_{1,t-1}^2 + \alpha_{12} a_{2,t-1}^2 + \beta_{11} \sigma_{11,t-1} + \beta_{12} \sigma_{22,t-1} \\ \sigma_{22,t} &= \alpha_{20} + \alpha_{21} a_{1,t-1}^2 + \alpha_{22} a_{2,t-1}^2 + \beta_{21} \sigma_{11,t-1} + \beta_{22} \sigma_{22,t-1} \end{aligned}$$

- Bollerslev determined the conditions necessary for this model to be well behaved
    - Eigenvalues of $\boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1$ must be less than one for covariance stationarity

Time-varying correlation GARCH

- Can have a time-varying correlation
  - One representation suggested based on a positive correlation is a GARCH(1,1) model for the variances combined with

$$\begin{aligned} \rho_{21,t} &= \frac{\exp\left(q_t\right)}{1+\exp\left(q_t\right)} \\ q_t &= \overline{\omega}_0 + \overline{\omega}_1 \rho_{21,t-1} + \overline{\omega}_2 \frac{a_{1,t-1} a_{2,t-1}}{\sqrt{\sigma_{11,t-1}}\sqrt{\sigma_{22,t-1}}} \end{aligned}$$

    where $\overline{\omega}_0$, $\overline{\omega}_1$ and $\overline{\omega}_2$ are parameters to be estimated
  - If the correlation can be negative, a better choice of function for the correlation would be

$$\rho_{21,t} = \frac{\exp\left(q_t\right) - 1}{\exp\left(q_t\right) + 1}$$