Dynamic Conditional Correlation Models

I Background & Motivation

我们之前算的都是 single asset / portfolio 的 variance (GARCH)
这部分内容是关于为 comovement between asset 建模

为什么 correlation 这么重要呢？

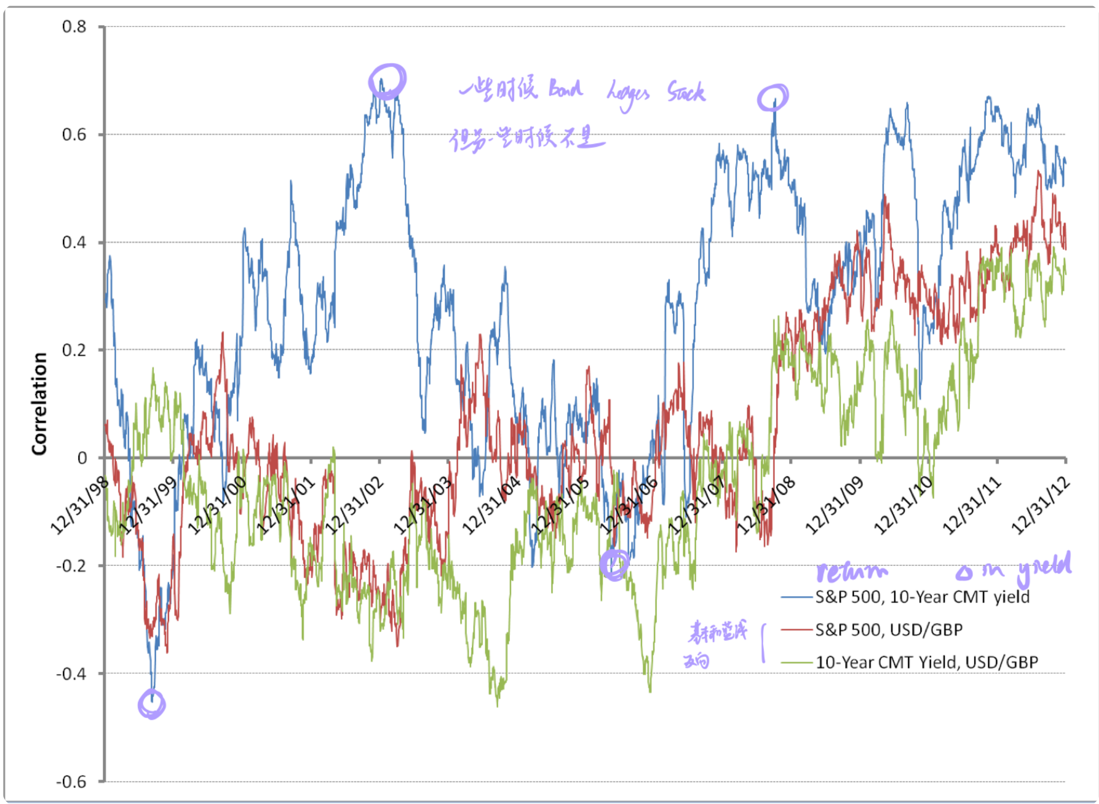
像方差一样，correlation 也随时间变化，而 correlation 和 s.d. 一起决定了 covariance matrix，协方差矩阵对 mean-variance optimization 和 VaR 计算至关重要

下面我们来看一下 correlation 变化很大的案子：

*10-year CMT yield 和 price 是反向的

- 这三条线变化都很大

Bond 和 Stock 一会同向变化一会反向变化，这个信息对我们很有用：决定了我们能否利用 Bond 来对冲 Stock 的 risk.



尽管一般的 assets 的 correlation 变化不会这么大

下面我们介绍一下 Background

先看一下 portfolio 的 return

$$r_{PF,t+1} = \sum_{i=1}^{n} w_{i,t}\, r_{i,t+1}$$

* r 是 simple return，不是 compounded return

$$\sigma^2_{PF,t+1} = w_t' \overset{t+1}{\Sigma} w_t$$

2×2 case：

$$\sigma^2_{PF,t+1} = \begin{bmatrix} w_{1,t} , & w_{2,t} \end{bmatrix} \begin{bmatrix} \sigma^2_{1,t+1} , & \sigma_{12,t+1} \\ \sigma_{12,t+1} , & \sigma^2_{2,t+1} \end{bmatrix} \begin{bmatrix} w_{1,t} \\ w_{2,t} \end{bmatrix}$$

$$= w_{1,t}^2 \sigma^2_{1,t+1} + w_{2,t}^2 \sigma^2_{2,t+1} + 2 w_{1,t} w_{2,t} \sigma_{12,t+1}$$

在 n×n 的情况下有些难算，解决这个问题的一个方法是

$$r_{PF,t+1} = r_{MKT,t+1} + \underline{\varepsilon_{t+1}}$$

$$\Downarrow$$

ε_{t+1} 是 idiosyncratic risk term，独立于 market return

那么 $\sigma^2_{PF,t+1} = \sigma^2_{MKT,t+1} + \sigma^2_{\varepsilon}$

如果 portfolio is well-diversified，σ^2_{ε} will be very small，then

$$\sigma^2_{PF,t+1} = \sigma^2_{MKT,t+1}$$

但是上面的模型太简单了，我们需要更复杂的模型

$$r_{PF,t+1} = \beta\, r_{MKT,t+1} + \varepsilon_{t+1}$$

$\Downarrow$ 更多的 factor

$$r_{PF,t+1} = \beta_1 r_{F1,t+1} + \cdots + \beta_{10} r_{F10,t+1} + \varepsilon_{t+1}$$

美股有4000~5000支，算这么大的 correlation matrix 不现实，但是只有2000支股票会被机构投资者买卖，但2000也是个很大的数字

这就是为什么我们用 Factor Model

II Introducing the DCC Model

DCC 是预测 correlation 的标准模型，就像 GARCH 是方差的标准模型

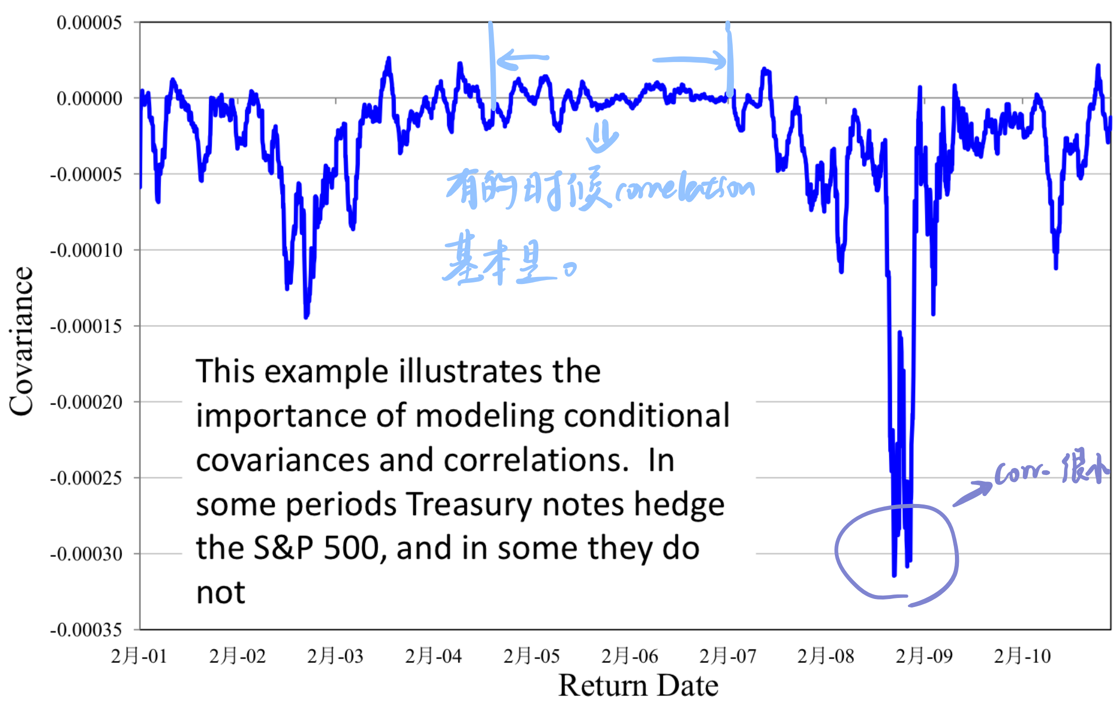
本部分要引入 DCC 模型

首先我们先讲一下以前我们是如何预测相关系数的呢？

Equally-weight estimator : $\sigma_{ij,t+1} = \frac{1}{m}\sum_{\tau=1}^{m} R_{i,t+1-\tau}R_{j,t+1-\tau}$

我们来看一下图：

有的时候 treasury note 和 stock 是不相关 (uncorrelated)，有时却可以用 treasury note 来 hedge stock



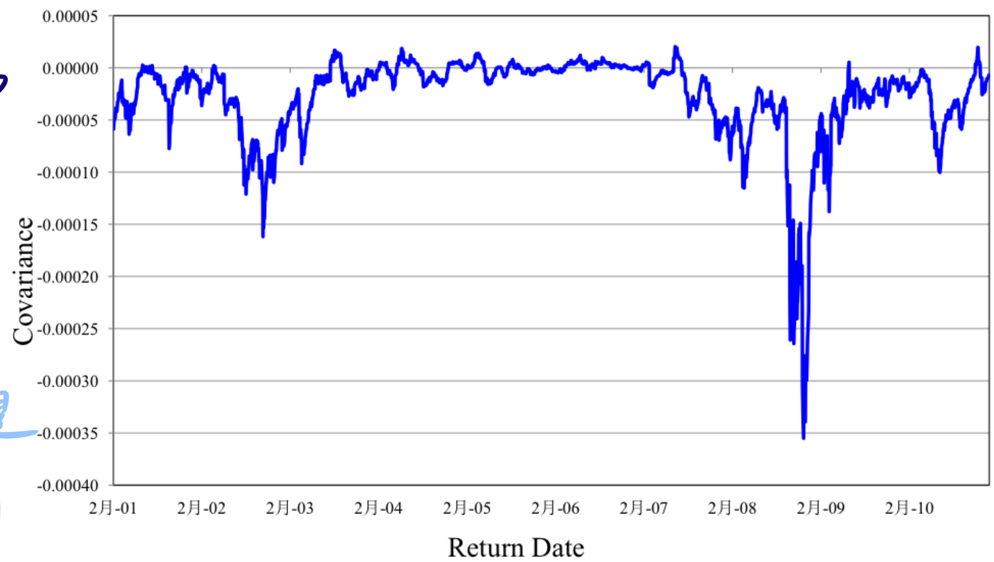
这个做法有什么坏处呢？

在 estimation window 中我们就有 weight，一出 window 立刻就是 0 了 — 不平滑

另一种方法

Exponentially-weighted estimator $\sigma_{ij,t+1} = (1-\lambda)\, R_{i,t}R_{j,t} + \lambda\sigma_{ij,t}$

在这个图中极端值更极端一些，因为极端值被赋予了更高的权重



但是！* $1-\lambda + \lambda = 1$. 这不是我们想要的，因为这样就没有 mean-reversion in covariance 了 为什么？

如果明天预测的covariance很高，它会一直都高 for all future horizons, rather than revert back to its mean

我们可以用 GARCH Model 解决上面的 limitation

$$\sigma_{ij,t+1} = \omega_{ij} + \alpha R_{i,t}R_{j,t} + \beta\sigma_{ij,t}$$
$$\sigma_{ij} = \omega_{ij}/(1-\alpha-\beta)$$

但是这样的模型也是有其局限性的

α. β 没有下标 → same for every pair of stocks

为什么要这样呢？确保了 covariance matrix 是半正定的. variance of portfolio ≥ 0

之前的模型（equally-weighted / exponential-weighted 都没有这个问题

我们进行的这个假设（α.β 都一样）是很强的限制条件. 不太现实

所以我们怎么办呢？我们直接为correlation建模：（DCC Model））

$\rho_{ij,t+1} = \sigma_{ij,t+1} / (\sigma_{i,t+1}\,\sigma_{j,t+1})$

所以我们用 { GARCH 算出来的方差 进行组合，得到 covariance matrix
DCC 算出来的 ρ

$$\Sigma_{t+1} = D_{t+1} \Upsilon_{t+1} D_{t+1}$$
$$\begin{bmatrix} \sigma_1 & \\ & \sigma_2 \end{bmatrix} \begin{bmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & \sigma_2 \end{bmatrix}$$

那么 Υ 怎么求呢？

$$z_{i,t+1} = \frac{R_{i,t+1}}{\sigma_{i,t+1}}$$

$E_t(z_{i,t+1} z_{j,t+1}) = \rho_{ij,t+1}$ for all i, j

↳ 它们的 covariance 就是 correlation

那么我们如何估计 $z_{i,t+1}$ $z_{j,t+1}$ 呢？我们引入中间量 q

$$q_{ij,t+1} = \rho_{ij} + \alpha\left(z_{i,t} z_{j,t} - \rho_{ij}\right) + \beta\left(q_{ij,t} - \rho_{ij}\right)$$

q 几乎是 correlation. 但它的范围有问题，($|q|$ 可能 > 1)，需要 rescale

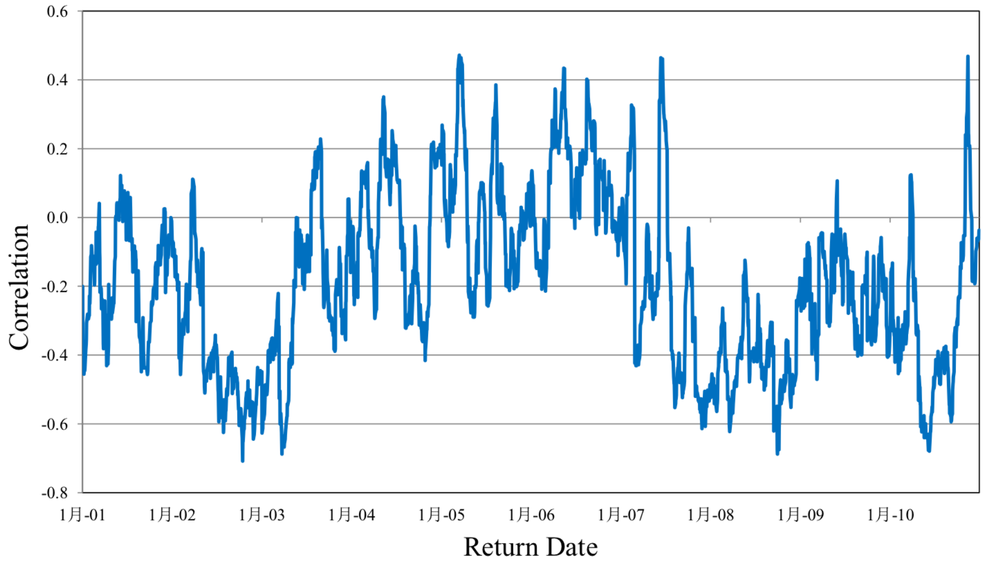
$$\rho_{ij,t+1} = \frac{q_{ij,t+1}}{\sqrt{q_{ii,t+1} q_{jj,t+1}}}$$

这是 DCC (1, 1) Model

注意这里的 α 和 β 还是一样的，但是 ρ_{ij} 不一样

→ does imply the persistence in correlation is constant across assets correlation 就是随机变化的

DCC correlation between S&P 500 and 10-Year Treasury note :



上面的模型要估的系数太多了,那我们怎么办呢?

$$\bar{\rho}_{ij} = \frac{1}{T}\sum_{t=1}^{T} z_{i,t} z_{j,t}$$

$$q_{ij,t+1} = \bar{\rho}_{ij} + \alpha\left(z_{i,t} z_{j,t} - \bar{\rho}_{ij}\right) + \beta\left(q_{ij,t} - \bar{\rho}_{ij}\right)$$

只有两个未知系数

III Estimation

下面介绍如何估计 DCC 模型的系数:

尽管我们直觉上想用MLE, 但是MLE是不行的

We will take two shortcuts

- The first shortcut is to do the estimation sequentially
  - First estimate a set of GARCH models for variances 先用GARCH估计方差
  - Compute normalized returns → z_i / z_j
  - Estimate the DCC model on the normalized returns
- The second shortcut is the replace the log-likelihood with a composite likelihood formed from bivariate log likelihoods

用新方法重新估计

我们先看：MLE有什么问题：

假设 returns 是 conditionally normally distributed

(当然别的分布也可以)

用MLE需要估计 DCC 的 α 和 β，每一个 asset 的 α_i，β_i，σ_i^2，σ_{1i}^2。

要估太多系数了！☹

The conditional density of r_{t+1} is

$$f(r_{t+1}) = \frac{1}{\sqrt{(2\pi)^k |\Sigma_{t+1}(\theta)|}} \exp\left[-\frac{1}{2} r'_{t+1} (\Sigma_{t+1}^{-1}(\theta)) r_{t+1}\right]$$

The log likelihood is obtained by taking the log and summing over the observations

Here $\Sigma_{t+1}^{-1}(\theta)$ is a highly nonlinear function of the parameters, and when n is large matrix inversion is slow

In addition, $\Sigma_{t+1}(\theta)$ is often close to singular (stock returns tend to be highly correlated), making the situation worse → 接近不可逆

Maximization of the log likelihood of the DCC model is a challenging problem 都不可逆了还求个锤子！

所以我们怎么办呢？我们走两个捷径：

第一个捷径： sequentially

一步一步算，先把GARCH都搞出来，再算 $z_{i,t+1} = \frac{r_{i,t+1}}{\sigma_{i,t+1}}$

现在MLE只有两个系数了： DCC 的 α 和 β

但ppt里的正定矩阵问题还是有的：

The conditional density of the vector z_{t+1} is

$$f(z_{t+1}) = \frac{1}{\sqrt{(2\pi)^k |\Upsilon_{t+1}(\alpha,\beta)|}} \exp\left[-\frac{1}{2} z_{t+1}{}' (\Upsilon_{t+1}^{-1}(\alpha,\beta)) z_{t+1}\right]$$

Remember that Υ_{t+1} is the correlation matrix

The log likelihood is again obtained by taking the log and summing over the observations

- Now there are only two parameters, α and β
- $\Upsilon_{t+1}^{-1}(\alpha,\beta,)$ is still a highly nonlinear function of the parameters, and when n is large matrix inversion is slow
- In addition, $\Upsilon_{t+1}(\alpha,\beta)$ is often close to singular (stock returns tend to be highly correlated), making the situation worse
- Maximization of the log likelihood is a less challenging numerical problem than before, but still can be a problem

第二个捷径: Composite Likelihood

我们把 n 维的矩阵换成 $z_{i,t+1}$ 和 $z_{j,t+1}$ 的 joint density

$$f(z_{i,t+1}, z_{j,t+1}) = \frac{1}{2\pi\sqrt{1-\rho_{ij,t+1}^2}} \exp\left[-\frac{1}{2}\frac{z_{i,t+1}^2 - 2\rho_{ij,t+1}z_{i,t+1}z_{j,t+1} + z_{j,t+1}^2}{1-\rho_{ij,t+1}^2}\right]$$

$$\rho_{ij,t} = \frac{q_{ij,t}}{\sqrt{q_{ii,t}q_{jj,t}}}$$

where

$$q_{ii,t+1} = 1 + \alpha(z_{i,t}{}^2 - 1) + \beta(q_{ii,t} - 1)$$
$$q_{ij,t+1} = \bar{\rho}_{ij} + \alpha(z_{i,t}z_{j,t}{}^2 - \bar{\rho}_{ij}) + \beta(q_{ij,t} - \bar{\rho}_{ij})$$
$$q_{jj,t+1} = 1 + \alpha(z_{j,t}{}^2 - 1) + \beta(q_{jj,t} - 1)$$

$q_{ii,1}=1$ $\quad q_{jj,1}=1$ $\quad q_{ij,1}=\bar{\rho}_{ij}$ $\quad \bar{\rho}_{ij} = \frac{1}{T}\sum_{t=1}^{T} z_{i,t}z_{j,t}$

$$\ln(CL_c) = -\frac{1}{2}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j>i}\left(\ln\left(1-\rho_{ij,t}^2\right) + \frac{\left(z_{i,t}^2 + z_{j,t}^2 - 2\rho_{ij,t}z_{i,t}z_{j,t}\right)}{\left(1-\rho_{ij,t}^2\right)}\right)$$

接下来讲一个 Generalization: Asymmetric correlation model

$$Q_{t+1} = (1-\alpha-\beta)\,E\left[z_t z_t'\right] + \alpha\left(z_t z_t'\right) + \beta Q_t + \gamma\left(\eta_t\eta_t' - E\left[\eta_t\eta_t'\right]\right)$$

where the $\eta_{i,t}$ for asset i is defined as the negative part of $z_{i,t}$ as follows:

$$\eta_{i,t} = \begin{cases} z_{i,t}, \text{ if } z_{i,t} < 0 \\ 0, \text{ if } z_{i,t} > 0 \end{cases} \quad \text{for all } i$$

Note that $E[\eta_t\eta_t']$ is a matrix; it can be estimated as the sample average of $\eta_t\eta_t'$.

Similarly, $E[z_t z_t']$ can be estimated as the sample average of $z_t z_t'$.

correlation tends to be higher in down markets