

知乎

首发于
川流不息

协方差矩阵的 Newey-West 调整



石川

量化交易 话题的优秀回答者

已关注

编辑推荐

78 人赞同了该文章

1 引言

我们在《正确理解 Barra 的纯因子模型》介绍了 Barra 的多因子模型。该文讨论的重点在于从业务上说明国家、行业、风格纯因子投资组合的含义，而非具体的数学计算。

不过，后来我意识到我给自己挖了一个坑。因为有个小伙伴给我们留言询问在计算因子协方差矩阵时，Barra 使用的 Newey-West 调整是怎么一回事儿。所以今天就来填坑了。本文就来简单说说 Newey-West 调整对于协方差矩阵估计的重要性。

在我为了写作本文而查阅的相关资料中，除了文末参考文献中的几篇重要论文外，知乎上的两篇讨论也给我很多启发（见参考文献），特此感谢。

2 为什么关注因子协方差矩阵？

通过多因子模型，我们可以把个股的收益率表达为因子收益率和个股特异性收益率的形式：

$$\mathbf{r} = \mathbf{X}\mathbf{f} + \mathbf{u}$$

在上式中， \mathbf{r} 为 $N \times 1$ 维个股收益率向量（省略了时间下标，假设有 N 支股票）、 \mathbf{X} 为当期因子暴露矩阵（ $N \times K$ 矩阵， K 为因子个数）， \mathbf{f} 为 $K \times 1$ 维因子收益率向量， \mathbf{u} 为 $N \times 1$ 维个股



知乎

首发于
川流不息

使用因子模型的好处是可以用它来推算个股收益率之间的协方差矩阵。直接计算股票收益率协方差矩阵的问题是该矩阵有 $0.5 \times (N^2 + N)$ 个不同的参数需要估计。这意味着我们至少需要 N 个样本数据来计算它。由于 N 是个股的个数，通常很大，因此这几乎是不可能的任务。

多因子模型的好处是，它把股票的收益率转换为因子收益率的线性组合。因此股票的风险也转换为因子风险的组合。因为因子的个数远远小于股票的个数，估计因子收益率的协方差矩阵要容易的多。对 $\mathbf{r} = \mathbf{X}\mathbf{f} + \mathbf{u}$ 等号两边同时做协方差运算可得：

$$\mathbf{V} = \mathbf{X}\mathbf{V}_f\mathbf{X}' + \Delta$$

其中， \mathbf{V} ($N \times N$) 是股票收益率的协方差矩阵， \mathbf{V}_f ($K \times K$) 是因子收益率的协方差矩阵，而 Δ 为 $N \times N$ 对角阵，其对角线上的元素对应个股的特异性收益率的方差——多因子模型假设股票的特异性收益和因子解释的收益率之间是独立的，因此因子收益率和特异性收益率之间不存在协方差；此外，模型同时假设不同股票的特异性收益率是相互独立的，因此股票的特异性收益率的协方差也为 0。

可见，为了得到 \mathbf{V} ，对于 \mathbf{V}_f 的求解至关重要。Newey-West 调整就是为了更准确的计算出 \mathbf{V}_f 。

3 时序不相关条件下协方差矩阵求解

在介绍协方差矩阵的 Newey-West 调整前，我们首先看看当因子收益率在时序上没有相关性时的做法（通常的做法）。为了简化数学表达，在下面的推导中，假设收益率已经去均值化 (demean) 了。假设共有 K 个因子，令 \mathbf{F}_t 表示第 t 期这 K 个因子的收益率向量，它是一个 $K \times 1$ 向量：

$$\mathbf{F}_t = \begin{bmatrix} f_t^{(1)} \\ f_t^{(2)} \\ \vdots \\ f_t^{(K)} \end{bmatrix}$$

在上面的表达中，符号 $f_t^{(k)}$ 既有下标也有上标：下标 t 表示第 t 期，而上标 (k) 表示第 k 个因子，因此 k 的取值是从 1 到 K ，所以 $f_t^{(k)}$ 就代表第 t 期，因子 k 的收益率（按照本小节一开始的说明，所有的 $f_t^{(k)}$ 都已经 demean 了）。

将 \mathbf{F}_t 和它的转置 \mathbf{F}_t' 相乘，利用线性代数的定义得到 $\mathbf{F}_t\mathbf{F}_t'$ ：



知乎

首发于
川流不息

$$\mathbf{F}_t \mathbf{F}_t' = \begin{bmatrix} f_t^{(2)} f_t^{(1)}, & f_t^{(2)} f_t^{(2)}, & \dots, & f_t^{(2)} f_t^{(K)} \\ \vdots & \vdots & \ddots & \vdots \\ f_t^{(K)} f_t^{(1)}, & f_t^{(K)} f_t^{(2)}, & \dots, & f_t^{(K)} f_t^{(K)} \end{bmatrix}$$

假设计算协方差矩阵的时间窗口为 T ，即 t 的取值为 1 到 T 。对 T 窗口内的所有 t 都进行上述运算并把不同 t 的 $\mathbf{F}_t \mathbf{F}_t'$ 相加得到 $\sum \mathbf{F}_t \mathbf{F}_t'$ ：

$$\sum_{t=1}^T \mathbf{F}_t \mathbf{F}_t' = \begin{bmatrix} \sum_{t=1}^T f_t^{(1)} f_t^{(1)}, & \sum_{t=1}^T f_t^{(1)} f_t^{(2)}, & \dots, & \sum_{t=1}^T f_t^{(1)} f_t^{(K)} \\ \sum_{t=1}^T f_t^{(2)} f_t^{(1)}, & \sum_{t=1}^T f_t^{(2)} f_t^{(2)}, & \dots, & \sum_{t=1}^T f_t^{(2)} f_t^{(K)} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{t=1}^T f_t^{(K)} f_t^{(1)}, & \sum_{t=1}^T f_t^{(K)} f_t^{(2)}, & \dots, & \sum_{t=1}^T f_t^{(K)} f_t^{(K)} \end{bmatrix}$$

最后，将 $\sum \mathbf{F}_t \mathbf{F}_t'$ 除以时间窗口长度 T 就得到 $(1/T) \sum \mathbf{F}_t \mathbf{F}_t'$ ，这正是以 T 窗口为长度计算出来的 K 个因子收益率的协方差矩阵 \mathbf{V}_f ：



知乎

首发于
川流不息

$$\mathbf{V}_f = \frac{1}{T} \sum_{t=1}^T \mathbf{F}_t \mathbf{F}_t' = \begin{bmatrix} \frac{1}{T} \sum_{t=1}^T f_t^{(2)} f_t^{(1)}, & \frac{1}{T} \sum_{t=1}^T f_t^{(2)} f_t^{(2)}, & \dots, & \frac{1}{T} \sum_{t=1}^T f_t^{(2)} f_t^{(K)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{T} \sum_{t=1}^T f_t^{(K)} f_t^{(1)}, & \frac{1}{T} \sum_{t=1}^T f_t^{(K)} f_t^{(2)}, & \dots, & \frac{1}{T} \sum_{t=1}^T f_t^{(K)} f_t^{(K)} \end{bmatrix}$$

写了这么“啰嗦”的推导，实在不是因为我有编辑数学公式的癖好，而是希望我们能够对如何从因子收益率向量求解协方差矩阵加深印象。这是因为 Newey and West (1987) 这篇提出 Newey-West 调整的论文行文风格非常干练、没有任何废话，上来就是矩阵和向量的运算，直接给出了计算向量 $h_t(\theta)$ 的协方差矩阵 S_T 的表达式（见下红框图划重点的部分）。我第一次读这篇论文的时候感到云里雾里，**难以把该文的推导和 Barra 文献中关于 Newey-West 调整的说明联系起来**（换句话说，看了 Barra 的文档说用了 Newey-West 调整，然后找到 1987 年的这篇论文一看，第一感觉却是“这俩有关系吗……”）。所以在我自己写作时，我花费了上面笔墨解释了协方差矩阵到底是怎么从单期收益率向量推导出来的，这能帮助我们更好的阅读 Newey and West (1987)。



BY WHITNEY K. NEWAY AND KENNETH D. WEST¹

MANY RECENT RATIONAL EXPECTATIONS MODELS have been estimated by the techniques developed by Hansen (1982), Hansen and Singleton (1982), Cumby, Huizinga, and Obstfeld (1983), and White and Domowitz (1984). These estimation techniques make use of an orthogonality condition $Eh_t(\theta^*) = 0$, where θ^* is a $(k \times 1)$ vector of unknown parameters and $h_t(\theta)$ is a $(r \times 1)$ vector of functions of the data and parameters, where $r \geq k$. This orthogonality condition can be employed to form a generalized method of moments (GMM, Hansen (1982)) estimator of θ^* by choosing $\hat{\theta}$ as the solution to

$$(1) \quad \min_{\theta} h_T(\theta)' \hat{W}_T h_T(\theta),$$

where $h_T(\theta) = \sum_{t=1}^T h_t(\theta)/T$ is the vector of sample moments of $h_t(\theta)$ and \hat{W}_T is a (possibly) random, symmetric weighting matrix.

As shown in Cumby, Huizinga, and Obstfeld (1983), Hansen (1982), and White and Domowitz (1984), the asymptotic covariance matrix of $\hat{\theta}$ is given by

$$(2) \quad V_T = (H_T' W_T H_T)^{-1} H_T' W_T S_T W_T H_T (H_T' W_T H_T)^{-1}$$

where $H_T = \sum_{t=1}^T E[h_{t\theta}(\theta^*)]/T$ and $h_{t\theta}(\theta)$ is the $(r \times k)$ matrix of partial derivatives of $h_t(\theta)$. W_T is a nonrandom matrix such that $\text{plim} (\hat{W}_T - W_T) = 0$, and $S_T = \sum_{s=1}^T \sum_{t=1}^T E[h_t(\theta^*) h_s(\theta^*)'] / T$. Consistent estimation of this asymptotic covariance matrix is essential for the construction of asymptotic confidence intervals and hypothesis tests.

Estimation of H_T and W_T is straightforward, since \hat{W}_T forms a natural estimator of W_T and under the regularity conditions in Hansen (1982) or White and Domowitz (1984) it will be the case that

在前文推导中， \mathbf{F}_t 可以被认为对应 Newey and West (1987) 中的 $h_t(\theta)$ ，而 \mathbf{V}_f 对应 Newey and West (1987) 中 S_T 的估计量。这样通过上面的推导就不难理解在 Newey and West (1987) 中 S_T 的估计量为什么会有和本文中的 \mathbf{V}_f 一样的表达式，这对于理解 Newey and West (1987) 很重要。

Newey and West (1987) 是严谨的数学论文，因此行文在总体和样本统计量之间切换。但在阅读本文时请暂时遗忘总体 vs 样本统计量。本文的所有 notation，比如 \mathbf{F}_t 、 \mathbf{V}_f 这些都是针对样本数据而言，正如 Barra 的模型一样——我们关注的是如何使用样本数据、通过 Newey-West 调整来对未知的协方差数据进行估计。在上文的数学推导中，因为目的是和 Newey and West (1987) 中的公式对应，所以求解协方差矩阵时除以 T ，而非 $T - 1$ 。在实际中为了得到无偏估计，应除以 $T - 1$ 。

在接下来的行文中，我们只要记住： \mathbf{F}_t 是一个 $K \times 1$ 的列向量，代表第 t 期 K 个因子的收益率向量（demean 之后的收益率）；而通过总共 T 期（sample size）的 \mathbf{F}_t ， $t = 1, 2, \dots, T$ 计算出来的因子收益率协方差矩阵 \mathbf{V}_f 为（请记住这个式子，下面讲 Newey-West 调整时还会用到）：

$$\mathbf{V}_f = \frac{1}{T} \sum_{t=1}^T \mathbf{F}_t \mathbf{F}_t'$$

知乎

首发于
川流不息

值得一提的是，在 Barra 的模型中，它们还对上式进行了一点点修改。上式中对时间窗口 T 内的各期收益率采用了简单平均，而 Barra 的模型采用了指数平均，目的是为了让更近期的数据有更高的权重，从而快速捕捉波动率的变化。使用指数平均对上式进行改进不是本文关注的重点，因此不再赘述。感兴趣的朋友请参考 Briner et al. (2009) 中的第 5.1 节（这个文献是 Barra 的 EUE3 模型——欧洲股权模型，它和 Barra 的 USE4 以及 CNE5 模型使用的方法相同）。

4 Newey-West 调整

上一节给出了当因子收益率在时序上不相关时求解协方差矩阵的方法。然而，当因子收益率在时序上有自相关性时，上节的计算方法就有问题了，它不是真实协方差矩阵的一个相合估计 (consistent estimation)。

相合估计大致可以理解为随着样本个数的增加，一个统计量的估计越来越逼近其真实值，实现在概率上收敛的效果。相合估计有助于我们计算统计量的估计误差，这对于后续使用该估计量至关重要（比如计算置信区间等）。

为了得到相合估计，必须考虑因子收益率之间的自相关性，从而在计算协方差矩阵时考虑自协方差的影响，这就是 Newey-West 调整的作用。

此外，Barra 的模型中必须要进行自协方差调整的另一个原因是，Barra 的多因子模型是日频的，因此每天都会有一期因子收益率，而它们的风险预测模型是月频的。这意味着 Barra 需要把日频的协方差矩阵通过尺度变换 (scaling) 变成月频的协方差矩阵。在这个过程中就必须考虑日频收益率之间的序列相关性。

*All EUE3 risk forecasts are **monthly** volatility estimates. The use of **daily** factor returns in (5.1) necessitates scaling the covariance matrices to monthly horizon. This scaling step needs to account for possible serial correlation in subsequent factor returns.*

假设单期的因子收益率 \mathbf{F}_t 满足一个 q 阶的序列相关性，即 \mathbf{F}_t 可以用 MA(q) 来刻画。则协方差矩阵的一个最简单的相合估计为（对应 Newey and West 1987 中的式 (4)）：



知乎

首发于
川流不息

i=1

$$\text{where } \mathbf{\Gamma}_0 = \frac{1}{T} \sum_{t=1}^T \mathbf{F}_t \mathbf{F}_t'$$

$$\mathbf{\Gamma}_i = \frac{1}{T} \sum_{t=1}^{T-i} \mathbf{F}_t \mathbf{F}_{t+i}'$$

在上式中， $\mathbf{\Gamma}_0$ 就是第三节中不考虑自相关性的协方差矩阵，而任何 $i \neq 0$ 对应的 $\mathbf{\Gamma}_i$ 代表着由 \mathbf{F}_t 和从时刻 t 滞后 i 期得到的 \mathbf{F}_{t+i} 计算出来的自协方差矩阵。举个例子，令 $i = 3$ 则 $\mathbf{\Gamma}_3$ 为：

$$\mathbf{\Gamma}_3 = \frac{1}{T} \sum_{t=1}^{T-3} \mathbf{F}_t \mathbf{F}_{t+3}' = \begin{bmatrix} \frac{1}{T} \sum_{t=1}^{T-3} f_t^{(1)} f_{t+3}^{(1)}, & \frac{1}{T} \sum_{t=1}^{T-3} f_t^{(1)} f_{t+3}^{(2)}, & \cdots, & \frac{1}{T} \sum_{t=1}^{T-3} f_t^{(1)} f_{t+3}^{(K)} \\ \frac{1}{T} \sum_{t=1}^{T-3} f_t^{(2)} f_{t+3}^{(1)}, & \frac{1}{T} \sum_{t=1}^{T-3} f_t^{(2)} f_{t+3}^{(2)}, & \cdots, & \frac{1}{T} \sum_{t=1}^{T-3} f_t^{(2)} f_{t+3}^{(K)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{T} \sum_{t=1}^{T-3} f_t^{(K)} f_{t+3}^{(1)}, & \frac{1}{T} \sum_{t=1}^{T-3} f_t^{(K)} f_{t+3}^{(2)}, & \cdots, & \frac{1}{T} \sum_{t=1}^{T-3} f_t^{(K)} f_{t+3}^{(K)} \end{bmatrix}$$

从这个例子中不难看出 $\mathbf{\Gamma}_3$ 本身不一定是对称的，因此在上述调整中，对于任何的滞后期 i ， $\mathbf{\Gamma}_i$ 和 $\mathbf{\Gamma}_i'$ 总是成对出现（ $\mathbf{\Gamma}_i + \mathbf{\Gamma}_i'$ 是对称的）。需要特别提醒的是，当计算滞后期为 i 的自协方差时，由于 \mathbf{F}_t 和 \mathbf{F}_{t+i} 之间有间隔 i ，因此在总共 T 长度时间窗口内，这二者的配对儿个数为 $T - i$ 、少于 T ，但是在计算 $\mathbf{\Gamma}_i$ 的表达式中，永远是除以 T ，而不是 $T - i$ 。

上述调整（请注意，我没有称该调整为 Newey-West 调整！下面马上就会解释）的本质是使用最大到 q 阶的自协方差 $\mathbf{\Gamma}_i$ 对 $\mathbf{\Gamma}_0$ 进行修正，从而得到调整后的因子收益率协方差矩阵 \mathbf{V}_f 。

上面这个调整有一个小问题，就是如此得到的协方差矩阵 \mathbf{V}_f 不一定是半正定（positive semi-definite）的，而协方差矩阵必须是半正定的。为了解决这个问题，大名鼎鼎的 Newey-West 调整出场。它在上述调整的思想，对 $\mathbf{\Gamma}_i$ 的修正加入了 Bartlett 权重系数 $1 - i/(1 + q)$ 。可以看到，该系数和滞后期 i 成反比，说明两个收益率向量 \mathbf{F}_t 和 \mathbf{F}_{t+i} 的间隔越大， $\mathbf{\Gamma}_i$ 的贡献越小。最终，协方差矩阵的 Newey-West 调整为（对应 Newey and West 1987 中的式 (5)）



$$\text{where } \Gamma_0 = \frac{1}{T} \sum_{t=1}^T \mathbf{F}_t \mathbf{F}_t'$$

$$\Gamma_i = \frac{1}{T} \sum_{t=1}^{T-i} \mathbf{F}_t \mathbf{F}_{t+i}'$$

$$\omega_i = 1 - \frac{i}{1+q}$$

Newey and West (1987) 证明了上面这个协方差矩阵是一个相合估计，而且它是半正定的。

这就是 Barra 采用的 Newey-West 调整。在 Barra 的 EUE3 研究报告 (Briner et al. 2009) 中的第五节对此有简单的说明。此外，前文提到，Barra 需要用日频的协方差矩阵通过 scaling 转换成月频的。为此，Barra 的做法是对 Newey-West 调整后的日频协方差矩阵乘以一个月内的交易天数，即 22，这就得到了最终的因子收益率协方差矩阵（下图摘自 Briner et al. 2009，在 EUE3 中 Barra 采用的最大滞后期为 15）。

Specifically, the monthly covariance matrices, $C^{(m)}$, are calculated as:

$$C^{(m)} = 22 \left[C^{(d)} + \sum_{\Delta=1}^D \left(1 - \frac{\Delta}{D+1} \right) (C_{+\Delta}^{(d)} + C_{-\Delta}^{(d)}) \right] \quad (5.2)$$

Here $C^{(d)}$ denotes the daily contemporaneous covariance matrix introduced in (5.1), and $C_{+\Delta}^{(d)}$ is a Δ days lagged daily covariance matrix. EUE3 includes correction terms up to a maximum lag of $D = 15$.

最后想要指出的是，在计算股票特异性收益率的协方差矩阵上，Barra 同样采用了 Newey-West 调整，只不过对于个股特异型收益率，Barra EUE3 模型选择的最大滞后期为 10。

5 结语

好了，终于把之前的坑填上了。学习大概就是不断的挖坑然后再填坑的过程。

从 Barra 自己的论述来看，它们在因子收益率协方差矩阵以及股票特异性收益率的方差矩阵上²都进行了 Newey-West 调整，且这一调整被沿用到了最新版的模型中，足见这一步的重要程

知乎

首发于
川流不息

2011)。因此，它们对因子的波动率和相关系数矩阵分别进行 Newey-West 调整。Barra 的 USE4 模型并没有披露具体细节，但**万变不离其宗**。我猜想应该和 EUE3 模型中的处理方法（即本文介绍的方法）一致，**区别就是我们使用不同因子的收益率序列 $\{f_t^{(k)}\}$ 计算出类似于本文中的 Γ_i ，即利用 K 个因子的 $\{f_t^{(k)}\}$ 序列求出不同滞后期 i 下相关系数的矩阵以及方差的对角阵，然后用这个矩阵替换 Γ_i 套入 Newey-West 调整的表达式即可；核心是用 $\{f_t^{(k)}\}$ 找到正确的矩阵。**

比如因子 (1) 和 (2) 之间的滞后期为 i 的相关系数可以通过下式计算，对所有因子和所有最大滞后期 q 以内的 i 计算就可以求出类似于本文中 Γ_i 的相关系数矩阵，然后就可以运用 Newey-West 调整得到相关系数矩阵的相合估计。

$$r_i^{(1)(2)} = \frac{\sum_{t=1}^{T-i} f_t^{(1)} f_{t+i}^{(2)}}{\sqrt{\sum_{t=1}^T (f_t^{(1)})^2} \sqrt{\sum_{t=1}^{T-i} (f_{t+i}^{(2)})^2}}$$

我自己尚未对 Newey-West 调整在 A 股上的有效性进行验证，但是国内一些券商的金工团队在这方面已经有了不少的探索。在这方面，天风证券应该算是走在了前列（天风直接对协方差矩阵调整，类似 EUE3 模型）。它应该是我最早看到将 Barra 这一套系统应用于国内 A 股市场上的（至少是 1 年以前），并且还非常有创造性的利用了最优化的手段配合 Barra 的体系来进行选股。进行最优化的前提条件当然是各种输入要尽可能准确，这就能体现出 Newey-West 调整的重要性了。在今后我们进行实证之后，如果有新的发现，也会及时和各位分享。

参考文献

- Briner, B. G., R. C. Smith, and P. Ward (2009). *The Barra Europe Equity Model (EUE3)*. MSCI Barra Research Notes.
- Menchero, J., D. J. Orr, and J. Wang (2011). *The Barra US Equity Model (USE4)*. MSCI Barra Research Notes.
- Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, Vol. 55(3), 703 – 708.
- zhuanlan.zhihu.com/p/27...
- zhihu.com/question/5735...

免责声明：文章内容不可视为投资意见。市场有风险，入市需谨慎。





编辑于 2019-07-02

BARRA模型 多因子模型 量化

▲ 赞同 78 ▼ 24 条评论 分享 收藏 ...

文章被以下专栏收录



川流不息

北京量信投资管理有限公司是一家在中国基金业协会备案登记的专业私募基金管理人...

关注专栏

推荐阅读

ARCH
ED

indexes are
established
research.
dings have
identification
it of the MSCI
dexes.

40+ Years of
Factor
Experience

MSCI FACTOR
INDEXES

EMPI
EVID

MSCI, beginn
invented a cor
to explain ri
through the
Barra factor n
to create
MSCI Fac

Barra 模型的填坑、讨论帖

石川

什么是多因子量化选股模型？

引言 量化投资中经常听到的“多因子模型”是个什么鬼？因子是影响因素的简称，或简单理解成指标。我们都知道股票收益受到多重因素的影响，比如宏观、行业、流动性、公司基本面、交易情绪等...

CuteH... 发表于Pytho...

【多因子模型 (2)

感到不爽

24 条评论

⇌ 切换为时间排序

写下你的评论...



lyzheng

1 年前

分别对相关系数和波动率做调整的关键是要保证对两者做完nw调整后的协方差矩阵仍然是正定的

👍 2



石川 (作者)

1 年前

知乎

首发于
川流不息

1 页



Chi Franklin

1 年前

您好！能不能接着做一下Barra USE4 Newey-West之后的特征值调整？那一段比较复杂并且看得懂运算过程不理解为什么要这么做，有什么实际含义，非常感谢！

赞



石川 (作者) 回复 Chi Franklin

1 年前

好 我回头看看 接着填坑☹

赞



Chi Franklin 回复 石川 (作者)

1 年前

非常感谢！

赞

展开其他 1 条回复



顾故如卿

1 年前

相关系数有两种求法，比如因子1和因子2，可以因子1取t期，因子2取t+i期，也可以反过来，所以要把两种方法求到的矩阵相加，再用NW权重加权得到最后结果吗？

赞



石川 (作者) 回复 顾故如卿

1 年前

对，应该和文中那个求滞后期 = 3 的协方差矩阵的例子相同。单一个相关系数矩阵不是对称的，对角线两边对应位置上就是你说的谁在前谁在后的问题。然后把这个矩阵和它的转置相加，作为滞后 i 期的调整，再乘以权重。

赞



li lu

1 年前

请问一下 应该先对日频率的协方差矩阵做Newey-West调整 然后再对月频率的协方差矩阵做特征值调整 还是应该先对日频率的协方差矩阵做特征值调整 然后再进行Newey-West调整？谢谢

赞



石川 (作者) 回复 li lu

1 年前

你的前一种理解是正确的。首先计算出日频协方差矩阵，然后通过 NW 调整找到月频的。根据Barra USE4 的模型文档，bias statistics 是按照月频计算的，而它是进行

eigenfactor 调整的目标，所以协方差矩阵的 eigenfactor 调整应该是针对月频协方差矩阵进行的，也就是在 NW 之后。





首发于
川流不息



ll lu 回复 石川 (作者)

1 年前

谢谢。另外 还有个问题 下式中

$$V_t = \frac{1}{T} \sum_{i=1}^T F_i F_i' = \begin{bmatrix} \frac{1}{T} \sum_{i=1}^T f_i^{(1)} f_i^{(1)} & \frac{1}{T} \sum_{i=1}^T f_i^{(1)} f_i^{(2)} & \dots & \frac{1}{T} \sum_{i=1}^T f_i^{(1)} f_i^{(K)} \\ \frac{1}{T} \sum_{i=1}^T f_i^{(2)} f_i^{(1)} & \frac{1}{T} \sum_{i=1}^T f_i^{(2)} f_i^{(2)} & \dots & \frac{1}{T} \sum_{i=1}^T f_i^{(2)} f_i^{(K)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{T} \sum_{i=1}^T f_i^{(K)} f_i^{(1)} & \frac{1}{T} \sum_{i=1}^T f_i^{(K)} f_i^{(2)} & \dots & \frac{1}{T} \sum_{i=1}^T f_i^{(K)} f_i^{(K)} \end{bmatrix}$$

取的是1/T, 无偏估计似乎是1/(T-1) 为什么不去1/(T-1)呢 谢谢

👍 赞

展开其他 2 条回复



tonyhfj0613

1 年前

您好, barra use4 在对特质波动率进行估计时也用到了newey west调整, 但是相比对因子协方差的newey west调整多了一个参数 “Newey-West Auto-Corr. Half-Life”, 这个half life和specific volatility half life不一样, 请问您是怎么理解这个half life的?

👍 赞



YauChunfung

8 个月前

你好, 很感谢你做的研究, 这个结论应该是有错误的, 在newey-west论文中, St指的是正交表达是的协方差矩阵, 而并非待估计参数的协方差矩阵, 因子收益率的协方差矩阵对应的应是文中的Vt,如有错误望纠正, 谢谢

👍 赞



石川 (作者) 回复 YauChunfung

8 个月前

您好, 感谢留言。我的理解是 F_t 对应了 NW 原文中的 $h(\theta)$, 而非 θ 。因此根据原文中 St 的定义, 本文中的 Vt 对应了 NW 原文中的 St。谢谢。

👍 赞



杰杆

7 个月前

您好, 请问下有些因子收益率指数加权算出的协方差矩阵不正定了, 是怎么回事呢

👍 赞



空鱼

7 个月前

请问下后面相关系数公司除数那里是T应该是T-1是笔误还是本来如此?

👍 赞



石川 (作者) 回复 空鱼

7 个月前

不够严谨, 实际中我用的T-1。谢谢

👍 赞

知乎

首发于
川流不息

石老师您好，协方差矩阵和相关系数的Newey-West调整我理解了，但是方差的Newey-West我不太理解，因为协方差和相关系数是两个因子之间的特性，方差是单个因子的特性，请问具体如何做呢？

赞



0FsN

6 个月前

您好，如果 \mathbf{F}_t 中出现空值，那么 \mathbf{T}_i 计算出来为空，有比较合适的处理方式吗，烦请指点，非常感谢！

赞



追风

5 个月前

东方证券也有一系列研究报告，LW压缩效果不比多因子模型差

赞



土十方

4 个月前

多谢~

赞



水逊

25 天前

请教一下，你提到“Barra 的模型采用了指数平均，目的是为了更近期的数据有更高的权重，从而快速捕捉波动率的变化”。如果同时考虑指数平均和Newey-West调整的话，Gamma_0是同期的收益率相乘的历史平均，按时间正常加权平均就行了，但后面Gamma_i是不同期的收益率相乘，该用哪个时间的权重做平均呢？还是有什么不一样的处理方法？

赞 回复 踩 举报

