因子模型及相关计量经济学概述

1. 从CAPM出发

$$\mathbb{E}[R_i] - R_f = \beta_i \left(\mathbb{E}[R_M] - R_f\right)$$

$$\mathbb{E}[R_i - R_f] = \beta_i \left(\mathbb{E}[R_M - R_f]\right)$$

$$\mathbb{E}[R_i^e] = \beta_i \left(\mathbb{E}[R_M^e]\right)$$

模型假设

1. 相同预期
2. 无限制借贷
3. 所有人做单期财富最大化（只考虑期望和方差）
4. 流动性好，无交易成本
5. 所有人都是价格接受者
6. 资产数目确定

   模型实际需要关注的

- 没有截距
- expectation是什么意思？从实证的角度看，是时间还是横截面？
  - β_i 的下标i是什么意思？
  - i是研究的兴趣点所在。也就是，不同的β_i是否对应了不一样的收益？（横截面选股）
- 直觉上，CAPM说了件什么事情？
  - 对于证券i（或证券组合），超额收益率的期望值，等于，β_i 和 另一个东西的超额收益率的期望值的乘积
  - 这“另一个东西”，是因子(factor)。在CAPM里，“因子超额收益”是全市场超额收益
  - β_i，是证券i在这个因子上的“因子暴露”(factor exposure, factor loading)

套利定价理论 Arbitrage Pricing Theory

$$\mathbb{E}[R_i^e] = \boldsymbol{\beta_i'}\boldsymbol{\lambda}$$

- K个因子, $\boldsymbol{\lambda}$是K维向量
- APT的出发点是统计上的factor decomposition + 一价定律
- 因此，除了“一价定律”以外，APT实际上没有讨论“经济含义”
- $\boldsymbol{\lambda}$是因子的预期超额收益率向量。也常被称作 factor risk premium。这个收益率应当是正的，因为对于风险，需要给出报酬。
  - 以下$\boldsymbol{\lambda}$都作为“因子的预期超额收益率向量”，如果不是收益率向量，写作$\boldsymbol{f}$。
  - 有时候因子的收益率是不能简单看到的，因此，因子本身的 factor risk premium 的估计，是一个重要的问题
- 实际数据是否符合这个公式？未必。此时会有定价误差α_i出现：

$$\mathbb{E}[R_i^e] = \alpha_i + \boldsymbol{\beta_i'}\boldsymbol{\lambda}$$

- α_i如果不显著偏离0，可能是统计误差。α_i如果显著不为0，则可能是：
  - 某些因子还没有被完全找到，当加入新的因子以后，α_i不显著了
  - 即使因子都被完全找到了，模型本身也许有问题，例如非线性？
    - 从统计模型的角度说，这种分解总是能做的。但对于实际数据来讲，找到的因子数据是否能符合这条式子，不一定
    - 如果纯粹只做统计的分解(比如分解出latent factors)，对于实际的投资来讲，没有啥意义，因为我们总是希望找到**数据之间**的联系，而非仅仅是**数据本身**的某种分解
    - 如何检验模型本身有没有问题？样本内的拟合情况？样本外的预测情况？
  - 模型没错，市场也许对该资产出现了错误定价，形成了套利机会。此时这个资产就被称为一个“异象”（anomaly）
    - 若α_i为正，则买入i资产，卖出因子组合

- 这也就是常见的所谓α，某种资产的收益中不能被已有的共性的因子描述的部分。如果等式左边是基金经理的资产组合收益，则α可能显示了这个基金经理超出常人的能力
  总结：因为非系统性风险可以通过分散投资消除，因子描述了驱动个体资产预期收益的共性原因，也即某种系统性风险。因子收益率是这种系统性风险的风险溢价(risk premium)，是个体资产的共性收益。因为因子是系统性风险的衡量，其 预期 收益应当为正（统计意义上），作为对这种风险的补偿。

理论的争论

一个直觉上的疑问

- 因子模型的解释是：
  - 因子收益率 λ 是 **risk premium（风险补偿）**
  - 因子代表某种 **系统性风险**
- 但这里会产生一个直觉上的问题：
  - 如果某个因子 **长期平均收益更高**
  - 那么长期持有该因子组合即可获得更高回报
- 因此产生疑问：
  - **长期平均收益是否真的来自风险补偿?**
- 直觉上的逻辑是：
  - 风险意味着 **概率上的不确定性**
  - 但长期平均会使概率波动被“平均掉”
- 因此有人认为：
  - **长期稳定存在的因子收益**
  - 可能不是风险补偿
  - 而是某种 **阶段性的结构性优势或市场误定价**

风险补偿解释（主流资产定价理论）

- 标准资产定价理论认为：

$$\mathbb{E}[R_i] - R_f$$

- 本身就是 **风险补偿**
- 因子收益高意味着：
  - 该因子在某些 **坏状态（bad states）** 下表现更差

    - 投资者厌恶这种风险
    - 因此要求更高期望收益
- 在这一框架下：
    - 风险不是简单的 **波动率**
    - 而是资产收益与 **宏观状态或边际效用** 的协方差

行为金融 / 误定价解释

- 另一种观点认为：
    - 因子收益可能来自 **市场误定价**
- 可能来源包括：
    - 投资者行为偏差
    - 机构约束（如杠杆限制）
    - 套利成本
    - 职业风险
- 在这种解释下：
    - 因子收益不是风险补偿
    - 而是 **市场效率不完全**

难题：风险补偿难以直接检验

- 现代资产定价理论的核心定价关系：

$$E[mR_i] = 1$$

- 其中：
    - m 是 **随机贴现因子（stochastic discount factor, SDF）**
- 问题在于：
    - m 在现实中 **不可直接观察**
- 因此：
    - 我们可以观察 **资产收益**
    - 却无法直接观察 **真实的风险价格**
- 这导致一个识别问题：
    - 因子收益高
    - 是 **风险补偿**
    - 还是 **市场误定价**
    数据本身很难区分。

对因子模型的启示

- 因子模型可以写成：

$$\mathbb{E}[R_i^e] = \boldsymbol{\beta_i'\lambda}$$

- 但在实证研究中：
  - 任何稳定存在的收益模式
  - 都可能被解释为 **新的风险因子**
- 因此资产定价文献中出现了所谓的
  - **factor zoo**
- 一句吐槽：Asset pricing is a theory that looking for risks
  - **资产定价理论在某种程度上是“不可证伪”的。** 只要我们没法准确定义和观测投资者的边际效用 m，任何关于“高收益是因为高风险”的结论，都可以通过更换 m 的定义来强行拟合。这种“寻找 m 的游戏”在已经玩了半个世纪。

更现实的理解

- 因子模型可以有两种理解：
  理论解释
  - 因子 = 系统性风险
  - 因子收益 = 风险补偿
    统计解释
  - 因子 = 描述收益共变结构的统计变量
  - 因子收益 = 数据中的稳定收益模式

总结

- 因子模型既是：
  - **资产定价理论**
  - 也是 **统计描述工具**
- 因子收益的来源可能是：
  - **风险补偿**
  - **市场误定价**
- 在现实研究中，两种机制 **可能同时存在。**

如何看待经济学、金融学中的模型?

> the only relevant test of the validity of a hypothesis is the comparison of its predictions with experience. The hypothesis is rejected if its predictions are contradicted; it is accepted if its predictions are not contradicted.
> --Milton Friedman

> But a theory is not like an airline or bus timetable . We are not interested simply in the accuracy of its predictions. A theory also serves as a base for thinking. It helps us to understand what is going on by enabling us to organize our thoughts. Faced with a choice between a theory which predicts well but gives us little insight into how the system works and one which gives us this insight but predicts badly, I would choose the latter, and I am inclined to think that most economists would do the same.
> -- Ronald Coase

> Everything should be made as simple as possible, but no simpler.
> (It can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience.)
> -- Albert Einstein

我的观点:

- Ockham's Razor (奥卡姆剃刀)是提出理论的原则
- 实证检验（某种意义上的prediction，不管是明显的还是不明显的，可以做的还是不能做的）是判断理论是否“正确”（有用）的必要条件
- 没有完美判断理论正确与否的充分条件

2. 如何检验?

时间序列回归

如果:

① $\mathbb{E}(\lambda) > 0$,

- $\boldsymbol{\lambda}$是收益率序列，且已知
- 已知 R_{it}^e的时间序列
  则$\boldsymbol{\beta}_i$需要估计。检验目标:

② $\alpha_1 = \alpha_2 = \cdots = \alpha_N = 0$?

$\alpha_i = 0$?

- 是否只用$\boldsymbol{\beta}_i'\boldsymbol{\lambda}$就能解释$\mathbb{E}[R_i^e]$，也即**截距是否等于0**
- $\boldsymbol{\lambda}$是否预期收益大于0
  步骤:

1. 时间序列回归(对每个i)：

$$R^e_{i(t)} = \alpha_i + \boldsymbol{\beta}'_i \boldsymbol{\lambda}_{(t)} + \epsilon_{i(t)}$$

   注：回归的样本用括号()表示，比如(t)表示在时间序列上回归
   得到估计量$\hat{\alpha}_i$，$\hat{\boldsymbol{\beta}}_i$
2. 因子模型的估计为：

$$\frac{1}{T}\sum_{t=1}^{T} R^e_{it} = \hat{\alpha}_i + \hat{\boldsymbol{\beta}}'_i\left(\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{\lambda}_t\right)$$

   记：$\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{\lambda}_t \equiv \hat{\mathbb{E}}[\boldsymbol{\lambda}]$，$\frac{1}{T}\sum_{t=1}^{T}R^e_{it} \equiv \hat{\mathbb{E}}[R^e_i]$。此二者即为因子预期收益、证券$i$预期收益的估计
3. 对于因子λ_k，感兴趣其期望收益是否不为0（风险因子是否给了正回报），单独做t检验即可。
4. 检验$\hat{\alpha}_i$是否为0。单个$\hat{\alpha}_i$只需做t检验即可。
5. 检验 $\hat{\alpha}_i$ 是否联合为0。Gibbons-Ross-Shanken (GRS) 检验：

$$\frac{T-N-K}{N}\left(1+\hat{\mathbb{E}}[\boldsymbol{\lambda}]'\hat{\Sigma}_\lambda^{-1}\hat{\mathbb{E}}[\boldsymbol{\lambda}]\right)^{-1}\hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha} \sim F_{N,T-N-K}$$

χ^2

- T: 时间序列的样本数；N：横截面样本数；K：因子数
- $\hat{\Sigma}_\lambda = \frac{1}{T}\sum_{t=1}^{T}\left[\boldsymbol{\lambda}_t - \hat{\mathbb{E}}[\boldsymbol{\lambda}_t]\right]\left[\boldsymbol{\lambda}_t - \hat{\mathbb{E}}[\boldsymbol{\lambda}_t]\right]'$， $K\times K$矩阵
- $\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}\hat{\boldsymbol{\epsilon}}_t\hat{\boldsymbol{\epsilon}}'_t$， $N\times N$矩阵

  理解：

1. $\hat{\alpha}^\top\hat{\Sigma}^{-1}\hat{\alpha}$
   - α 的加权平方和
     - $\hat{\alpha}$ 是一个 $N\times 1$ 的向量，包含每个资产的估计超额回报（即 α_i）
     - $\hat{\Sigma}$ 是一个 $N\times N$ 的残差协方差矩阵，反映资产回报中无法被因子解释的部分的相关性
     - $\hat{\alpha}^\top\hat{\Sigma}^{-1}\hat{\alpha}$ 是一个二次型，类似于“标准化后的 α 平方和”(把无法被因子解释部分的相关性考虑进来)
2. $\left(1+\mathbb{E}[\lambda]^\top\Sigma^{-1}\mathbb{E}[\lambda]\right)^{-1}$
   - 调整因子，反映了因子本身的“波动性”或“信号强度”
     - $\mathbb{E}[\lambda]$ 是因子回报的期望值
     - Σ 是因子回报的协方差矩阵，反映因子的波动性和相关性
     - $\mathbb{E}[\lambda]^\top\Sigma^{-1}\mathbb{E}[\lambda]$ 类似于因子回报的“夏普比率平方”（Sharpe Ratio squared）。
       - 在单因子模型（如 CAPM）中，$\mathbb{E}[\lambda]^\top\Sigma^{-1}\mathbb{E}[\lambda] = \frac{(\mathbb{E}[\lambda])^2}{\mathrm{Var}(\lambda)}$

    - 夏普比率衡量了因子的风险调整后回报，反映了因子的“信号强度”。如果因子的夏普比率高（即 $\mathbb{E}[\lambda]$ 大，Σ 小），说明因子很“强”，它的波动较小，回报较高。
    - 整个项 $\left(1+\mathbb{E}[\lambda]^{\top}\Sigma^{-1}\mathbb{E}[\lambda]\right)^{-1}$ 是一个标准化因子，用于调整 $\hat{\alpha}^{\top}\hat{\Sigma}^{-1}\hat{\alpha}$ 。
  - 如果因子的夏普比率很高（即因子很强），$\mathbb{E}[\lambda]^{\top}\Sigma^{-1}\mathbb{E}[\lambda]$ 很大，这一项会变小，意味着对 α 的检验更严格。因为一个“强因子”应该能解释更多的回报，任何非零的 α 都会被认为是更严重的模型失效。
  - 如果因子的夏普比率很低（即因子很弱），这一项会接近 1，意味着检验对 α 的容忍度更高。

OLS GRS检验的假设：

- 时间序列上ϵ_{it}同方差，无自相关，ϵ_{it}和λ_{tk}独立，也即ϵ_{it} i.i.d. 有限样本下，F分布的结果显然也是假设了ϵ_{it}时间序列上正态分布
- 大样本下，无需正态分布假设，但对应有限样本的χ^2统计量仍然假设了同方差和无自相关
- 若直接从矩条件出发:

$$\begin{cases}\mathbb{E}[\boldsymbol{\epsilon}_t]=0, & N\text{个矩条件}\\ \mathbb{E}[\boldsymbol{\lambda}_t\otimes\boldsymbol{\epsilon}_t]=0, & K\times N\text{个矩条件}\end{cases}$$

MM

Methods Moments.

  给出GMM估计量，则可以给出Robust的error，无需假设同方差和自相关。这里矩条件个数和参数个数 ($\boldsymbol{\alpha}$有N个，$\boldsymbol{\beta}$有$N\times K$个)一样，恰好识别

两步法横截面回归(two-pass)

如果：

- $\boldsymbol{\lambda}_t$未知，也即因子收益率不知道，此时要想办法先估计$\boldsymbol{\lambda}_t$
- 然后检验$\hat{\alpha}_i$是否联合为0

因子值.

1. 时间序列回归(对每个i)：

$$R^e_{i(t)}=a_i+\boldsymbol{\beta}_i'\boldsymbol{f}_{(t)}+u_{i(t)}$$

$\times\ \alpha.$ $\beta(\lambda)$

   注意：

   - $\boldsymbol{f}_{(t)}$不是**因子的收益率**，而是因子的某个代理变量，或者就是因子值
   - 此时截距项a_i不是定价误差
   - $\boldsymbol{\beta}_i$ 是**因子暴露的估计值**。换句话说，因为不能直接观察到和因子直接联系的**因子收益率**，只能观察到**因子值或因子的性质的某个代理变量**，我们用这个时间序列回归来估计证券i的因子暴露。

$E(X\varepsilon)=0$. Normal Equation? $\frac{1}{N}\sum_{i=1}^{N}\alpha_i\cdot 1=0$

Linear Regression?

1. 先求R^e_{it}时间上的均值，作为资产i的预期收益率的估计值，$\frac{1}{T}\sum_{t=1}^{T} R^e_{it} \equiv \hat{\mathbb{E}}[R^e_i]$。接着做如下横截面回归：

OLS? $E(\hat{x}^2) = x^2$ Jensen b 一个撇 $\Sigma \alpha_{(i)} = 0$

$$\hat{\mathbb{E}}[R^e_{(i)}] = \boldsymbol{\lambda}'\hat{\boldsymbol{\beta}}_{(i)} + \alpha_{(i)}$$

无偏性

此时$\boldsymbol{\lambda}$是因子预期收益率的估计值。α_i是回归的残差。因为是OLS，这个估计本身就是从"最小化α_i的平方和"出发来计算的。

有效性 Best Linear Unbiased Estimator

- 这个回归也可以加截距。不加截距的回归，是理论；加了截距，是考虑**理论也有可能是错的**。因此，计量里常见的efficiency和robustness的讨论在这里也成立。也即：

vs 一致性

  - 如果理论($\alpha_i = 0$)真的是对的，不加截距的回归能增强参数估计的efficiency，也即更有可能拒绝错误的假设($\alpha_i \neq 0$)。加了截距项回归，则更robust，但牺牲了efficiency。也即此时可能会增大接受实际错误的假设($\alpha_i \neq 0$)的可能性。如果此时也能拒绝假设，那么更放心了。
  - 如果理论是错的，因子没找全从而应当加上截距项。此时若回归没加截距项，则遗漏变量，可能估计不一致。

Simultaneity
① 反身性
② omit variables
③ measurement error

1. 检验$\hat{\alpha}_i$是否联合为0。这里的矩阵和向量都没有加粗，请留意。
记

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_{11} & \cdots & \hat{\beta}_{1K} \\ \vdots & \ddots & \vdots \\ \hat{\beta}_{N1} & \cdots & \hat{\beta}_{NK} \end{bmatrix}, \quad \hat{\alpha} = \begin{bmatrix} \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_N \end{bmatrix}, \quad u_t = \begin{bmatrix} u_{1t} \\ \vdots \\ u_{Nt} \end{bmatrix}, \quad \hat{\mathbb{E}}\left[R^e\right] = \begin{bmatrix} \hat{\mathbb{E}}\left[R^e_1\right] \\ \vdots \\ \hat{\mathbb{E}}\left[R^e_N\right] \end{bmatrix}, \quad \hat{\lambda} = \begin{bmatrix} \hat{\lambda}_1 \\ \vdots \\ \hat{\lambda}_K \end{bmatrix}$$

则参数估计为：

$$\hat{\lambda} = \left(\hat{\beta}'\hat{\beta}\right)^{-1}\hat{\beta}'\hat{\mathbb{E}}\left[R^e\right]$$ y ☆

$$\hat{\alpha} = \hat{\mathbb{E}}\left[R^e\right] - \hat{\beta}\hat{\lambda}$$

记$\Sigma_f \equiv \operatorname{cov}\left(f_t\right)$，$(K \times K$矩阵$)$，$\Sigma \equiv \operatorname{cov}(u_t)$，$(N \times N$矩阵$)$
则

$$\operatorname{cov}(\hat{\lambda}) = \frac{1}{T}\left[\left(\hat{\beta}'\hat{\beta}\right)^{-1}\hat{\beta}'\Sigma\hat{\beta}\left(\hat{\beta}'\hat{\beta}\right)^{-1}\left(1 + \lambda'\Sigma_f^{-1}\lambda\right) + \Sigma_f\right]$$

$$\operatorname{cov}(\hat{\alpha}) = \frac{1}{T}\left[I - \hat{\beta}\left(\hat{\beta}'\hat{\beta}\right)^{-1}\hat{\beta}'\right]\Sigma\left[I - \hat{\beta}\left(\hat{\beta}'\hat{\beta}\right)^{-1}\hat{\beta}'\right]' \cdot \left(1 + \lambda'\Sigma_f^{-1}\lambda\right)$$

统计量：

$$\hat{\alpha}' \operatorname{cov}(\hat{\alpha})^{-1}\hat{\alpha} \sim \chi^2_{N-K}$$

因子预期收益率$\boldsymbol{\lambda}$的检验：从 $\operatorname{cov}(\hat{\lambda})$取出对角线上的元素，开方即是标准误$\sigma_{\lambda_k}$。用$\hat{\lambda}_k/\sigma_{\lambda_k}$做自由度为T-1的t检验即可。

- $\left(1+\lambda' \Sigma_f^{-1} \lambda\right)$是Shanken(1992)的调整，因为第一步的因子暴露β是估计出来的，要把估计误差考虑进来。
- 在月度数据上检验模型，这个误差实际影响很小
- 也有对应的GLS版本和GMM版本，具体可查阅 Cochrane (2005)。

β_i^1 β_i^2

Fama-MacBeth 回归

"1973"

0 λ_1 λ_2 λ_3 T

1. 时间序列回归(对每个i)：

$$R_{i(t)}^e=a_i+\boldsymbol{\beta}_i' \boldsymbol{f}_{(t)}+u_{i(t)}$$

rolling window.

2. **在每个时间t上** 做横截面回归

$$R_{(i)t}^e=\lambda_t' \hat{\beta}_{(i)}+\alpha_{(i)t}$$

- 也可以有截距项
- 和两步法横截面回归的区别：
  - 两步法的第二步先取时间平均，然后做一次横截面回归
  - FM 在每个时间t上都做横截面回归

    好处：相当于增加了样本。

1. 假设误差在时间上独立，则有以下统计量：

$$\hat{\lambda}=\frac{1}{T} \sum_{t=1}^T \hat{\lambda}_t \quad \hat{\alpha}_i=\frac{1}{T} \sum_{t=1}^T \hat{\alpha}_{it}$$

$$\sigma^2(\hat{\lambda})=\frac{1}{T^2} \sum_{t=1}^T\left(\hat{\lambda}_t-\hat{\lambda}\right)^2, \quad \sigma^2\left(\hat{\alpha}_i\right)=\frac{1}{T^2} \sum_{t=1}^T\left(\hat{\alpha}_{it}-\hat{\alpha}_i\right)^2$$

- 检验因子预期收益率是否为0，用$\hat{\lambda}_t$序列做t检验即可
- 联合检验$\hat{\alpha}_i$是否为0:

  记$\hat{\alpha}=\frac{1}{T} \sum_{t=1}^T \hat{\alpha}_t$, 则可用

$$\operatorname{cov}(\hat{\alpha})=\frac{1}{T^2} \sum_{t=1}^T\left(\hat{\alpha}_t-\hat{\alpha}\right)\left(\hat{\alpha}_t-\hat{\alpha}\right)'$$

作为方差协方差矩阵的估计（横截面上的相关性，和两步法类似），检验：

"p" $\leftarrow \frac{p-1}{p}, \frac{p-2}{p}$

reg y x_1, x_2, r

$$\hat{\alpha}' \operatorname{cov}(\hat{\alpha})^{-1} \hat{\alpha} \sim \chi^2_{N-K}$$

Robust Error

- 时间独立假设在资产定价问题里可以接受，因为收益率在时间上相关性很弱。公司金融里则未必。
- 即使时间上不独立，也可以很容易把估计量更改为调整了时间相关性的估计量，例如 $\sigma^2(\hat{\lambda}) = \frac{1}{T}\sum_{j=-\infty}^{\infty} \operatorname{cov}_T\left(\hat{\lambda}_t, \hat{\lambda}_{t-j}\right)$，当然可以用Newey-West调整
- 另外，第一步的时间序列回归，可以是每个t上得到不一样的β_{it}。比如，对应于t期，用rolling window做回归，用的数据是t-h到t-1期的(FM原文用的是5年的月收益率)，得到$\hat{\beta}_{it-1}$。接下来在第二步时就用这些每一期不一样的$\hat{\beta}_{it-1}$对R^e_{it}做横截面回归
- 第二步的回归就可以这样理解：在每个时间点上，$\hat{\beta}_{it-1}$是这个N个资产在横截面上的差异，这些差异能否带来第t期的R^e_{it}的不一样

3. 方法差异（用什么作为因子收益率？） λ_t $\frac{1}{T}\sum\lambda_t$

问题：如果因子收益已知（也即这类因子的特征可以直接从可交易的证券中观察到，例如size, momentum，而非GDP），那么时间序列上直接求平均就是因子预期收益的估计。此时如果仍然做两步法横截面回归或者FM回归，又得到一个从OLS而来的所谓的因子预期收益的估计，这两者有什么不同？该用哪个？

差异之处：
横截面回归时，

K×1 回归 $\hat{\lambda}$

$$\hat{\lambda} = \left(\hat{\beta}'\hat{\beta}\right)^{-1}\hat{\beta}'\,\boxed{\mathbb{E}\left[R^e\right]}$$

N×1

注意 $\hat{\beta} = \begin{bmatrix} \hat{\beta}_{11} & \cdots & \hat{\beta}_{1K} \\ \vdots & \ddots & \vdots \\ \hat{\beta}_{N1} & \cdots & \hat{\beta}_{NK} \end{bmatrix}$ 是因子暴露矩阵，$\hat{\beta}_{ij}$是第i个资产在第j个因子上的暴露

N×K

$\left(\hat{\beta}'\hat{\beta}\right)^{-1}\hat{\beta}'$是一个$K \times N$矩阵，第$k$行可以看成是第$k$个因子的投资组合，所以$\left(\hat{\beta}'\hat{\beta}\right)^{-1}\hat{\beta}'$的第$j$行、第$i$列的值是第$j$个因子在第$i$个资产上的权重
此时，

$$\left[\left(\hat{\beta}'\hat{\beta}\right)^{-1}\hat{\beta}'\right]\hat{\beta} = \left(\hat{\beta}'\hat{\beta}\right)^{-1}\left(\hat{\beta}'\hat{\beta}\right) = \boxed{I_K}$$

因此：
$\left(\hat{\beta}'\hat{\beta}\right)^{-1}\hat{\beta}'$这个因子的投资组合权重矩阵有如下性质：

- 第k个因子在所有资产上的投资组合($\left(\hat{\beta}'\hat{\beta}\right)^{-1}\hat{\beta}'$的第$k$行给出的权重)，在第k个因子上的暴露是1

- 第k个因子在所有资产上的投资组合($\left(\hat{\beta}'\hat{\beta}\right)^{-1}\hat{\beta}'$的第$k$行给出的权重)，在其他因子上的暴露是0
- 因此，横截面回归得到的因子预期收益率的估计，更“纯净”
- 有点耍赖的感觉，因为$\hat{\beta}$是前一步算出来的，又用到后面作为回归变量。万一这个$\hat{\beta}$也是错的呢?

排序法

小结前面内容:
因子的概念是抽象的:

- 因子收益率是什么? 如何估计?
- 因子暴露是什么? 如何估计?
  实践中，只能考虑用可以交易的证券的组合，来模拟某个因子收益率。理想的模拟组合，应当满足:

1. 只在这个被模拟的因子上暴露，在其他因子上不暴露
2. 如果有多种满足条件1的组合，应当选择特质性风险较小的那个组合

- 条件1很好理解：如果同时在其他因子上有暴露，则这个投资组合的收益率不能仅仅归因于这个因子，同其他因子也有关系。
- 条件2说的是：$R_{it}^{e}=\alpha_i+\beta_{ki}\lambda_{kt}+0*\lambda_{[k]t}+\epsilon_{it}$ 中的 ϵ_{it}的变化要比较小，否则也会干扰这个投资组合的模拟程度。中括号表示除k之外的其他因子，其β为0

鸡生蛋还是蛋生鸡?

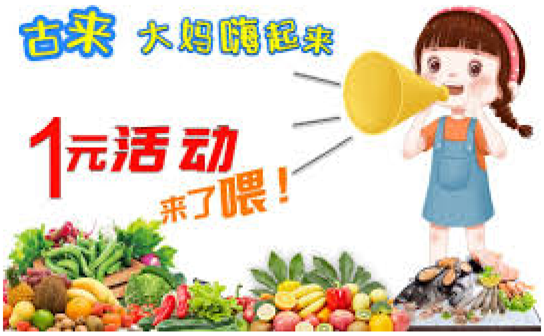
- 按照理想中的模型，估计因子暴露需要知道因子收益率
- 要知道因子收益率，需要知道因子模拟组合，而构建因子模拟组合，需要知道因子暴露
  如果因子是可以交易的证券的某个特征，比如Book/Market ratio，则利用排序法可以模拟构建因子收益率。

size. 市值.

排序法: f
别估计因子暴露了，就用这个想法：如果证券对应的这个特征的指标高，则它在这个因子上的因子暴露就高

- 这个特征并不一定是因子暴露。也没有什么模型
- 因此，如果因子不是可以交易的证券的特征，比如 GDP，那么得回到两步回归的方法来估计因子暴露和因子收益率
  抛开所有理论的直觉：拥有某个特征的某种证券(因子暴露更高)，在未来的收益会更好(因子预期收益为正)。

换句话说，横向比较，某些证券未来更好，并且可以在现在通过某些指标看出来。
量化交易野路子：菜场大妈选股策略（10年100倍）



单排

1. 确定特征，选择股票池
2. 按照特征的取值大小分组，例如10组。构建多空组合：做多最高组的股票，做空最低组的股票。通常要求多空两组的金额相同，无需自有资金。由于多空两组通常包含多只证券，还需考虑单只证券的权重，例如按市值加权
3. 定时更新(rebalance)

双排

双排可以控制另一个特征(因子)的影响
一般常见的 double sorting 有两种，一种是 independent sort, 一种是 dependent sort.

- independent sort 就是对于 variable 1, 计算 quantile，按照 quantile 分组。对于 variable 2，计算 quantile，再分组。取之间的交集。
    - independent sort 的好处是对于 variable 1，2 的顺序没有要求。坏处是如果 variabl 1 和 variable 2 高度相关，则可能导致某些交叉组里的成员很多，某些交叉组里的成员又很少，甚至没有。
- dependent sort 就是对于 variable 1, 计算 quantile，分组。然后在每组内部，按照 variable 2 计算 quantile，再分成小组。
    - dependent sort 相当于是一个 "conditional on"，也就是 “基于” variable 1 的关于 variable 2 的分组。看的是在 variable 1 各个组内部 variable 2的情况。
    - 因此，dependent sort的顺序很重要

4. 截面多因子模型 vs. 时序多因子模型（用什么作为因子暴露？）

估计因子暴露似乎有很多种套路：

- 最传统的Fama-French方法：先按照特征(特征并不是因子暴露)排序得到因子收益率，时序回归得到因子暴露。这种传统方法被称作“时序多因子模型”
- 就把特征作为因子暴露，不做时序回归了。直接开始做第二步在每个时间t上做横截面回归，然后考察λ_t的均值是否不为0。这种被称作“横截面多因子模型”。

FM

- 甚至还可以这么做：
  - 先用公司特征和R^e_{it}做每个t上的横截面回归，得到λ_t
  - 再用λ_t和R^e_{it}做时序回归，又得到$\hat{\beta}_{it}$
- Fama French (2020)给出了这几种方法的比较。结论：横截面多因子模型的解释力比较好。

5. 检验异象

排序法检验

如果一个因子和另一个因子的相关性很强，则按照一个因子进行排序可能捕捉的是另一个因子的风险。此时可以用双排法：
同时用两个因子分别独立排序，把交集拿出来检验。
多个因子时，多重排序法存在问题：交集会越来越少，甚至没有交集。

时序检验

1. 已知一个异象的收益率序列$R^e_{Anomaly,t}$（例如，可以是用exposure构建的排序股票组合）。已知其他因子收益率序列$\boldsymbol{\lambda}_t$。把$R^e_{Anomaly,t}$放等式左边，$\boldsymbol{\lambda}_t$放等式右边，做时间序列回归
2. 看截距项是否显著不为0 (NW 调整)
   例：

'α'

收益率

β时变

$$R^e_{Anomaly,t} = \alpha + \underbrace{\beta_1 R^e_{Mt} + \beta_2 SMB_t + \beta_3 HML_t}_{\times} + \epsilon_t$$

因子择时、版块轮动

FM回归检验

1. 已知异象的指示变量(exposure)，已知其他因子的factor exposure。用下一期个股收益放等式左边，factor exposures 放等式右边
2. 在每个时间t上做横截面回归，得到参数估计$\hat{\lambda}_t$的时间序列
3. 看$\hat{\lambda}_t$是否平均不为0 (NW调整)
   例：

$$R_{it+1}^e = \gamma_t + \lambda_{0t} Anomaly_{it} + \lambda_{1t}\beta_{it} + \lambda_{2t}Size_{it} + \lambda_{3t}BM_{it} + u_{it}$$

Book/Market

细节：求简单均值时，如何做 NW 调整来计算标准误？

解答：做单变量回归。$\lambda_{0t} = \text{intercept} + u_t$，intercept的估计值就是$\lambda_{0t}$的均值，同时，软件可以轻松计算NW调整，无需自己手动计算。

6. 总结

个人经验：

- 各种方法差异不大。一种方法给出了显著的结果，另一种方法通常也不会不显著。
- 为了稳健起见，学术论文通常考察多种方法，各种检验。
- 然而，p-hacking、数据挖掘出来的假象难以避免。
- 另一方面，市场本身不是平稳的，风险源一直都在变，样本外的表现不稳定。
- 实践中，应当仔细思索因子的经济含义，避免陷入数据挖掘的狂热。

> The cost of computing has dropped exponentially, but the cost of thinking is what it always was. That is why we see so many articles with so many regressions and so little thought.
> --Zvi Griliches

参考文献

Cochrane, John H., 2005, *Asset Pricing*. Revised edition. (Princeton University Press, Princeton, N.J).