资产定价

杨弘毅

创建：2021 年 6 月 10 日
修改：2022 年 1 月 26 日

目录

1 TODO

- campbell-shiller 分解

2 基础

2.1 CAPM 与 APT

资本资产定价模型（Capital Asset Pricing Model，CAPM）的诞生，才首次清晰的描绘出风险与收益率之间的关系。根据 CAPM 模型，资产的预期超额收益率由如下一元线性方程决定：

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

其中 R_i 为某资产 i 的收益率，R_f 为无风险收益率，$\mathbb{E}[R_M]$ 为市场组合的预期收益率，$\mathbb{E}[R_m] - R_f$ 为市场风险溢价（Market risk premium），也称为市场因子。其中有 $\beta_i = \frac{\mathrm{Cov}(R_i, R_M)}{\mathrm{Var}(R_M)}$，$\beta$ 刻画了该资产 i 收益对于市场收益的敏感程度，也被称为资产 i 对市场风险的暴露程度。

随后 Ross（1976）提出了著名的套利定价理论（Arbitrage Pricing Theory，APT），为多元线性模型:

$$\mathbb{E}[R_i^e] = \boldsymbol{\beta}_{\boldsymbol{i}}' \boldsymbol{\lambda}$$

同 CAPM 模型相同，$\boldsymbol{\beta}$ 为因子暴露（Factor exposure）或称为因子载荷（Factor loading），$\boldsymbol{\lambda}$ 是因子预期收益率（Factor expected return），或称为因子溢价（Factor risk premium）或因子风险溢酬。由此可见，资产 i 的预期超额收益率 $\mathbb{E}[R_i^e]$，为等式右侧一系列因子的预期收益率，以及该资产在这些因子上的暴露决定。

此时研究的是不同资产之间的预期超额收益率的差别，称为（横）截面（Cross-sectional）差异，而非时间序列（Time-series）或时序上的差异。因子代表了收益率的一种结构，给定了结构与因子预期收益率，不同资产预期超额收益率的差别，由其在这些因子上的暴露决定。那么多因子模型研究的核心问题，是找到一组能够解释股票预期收益率界面差异的因子。

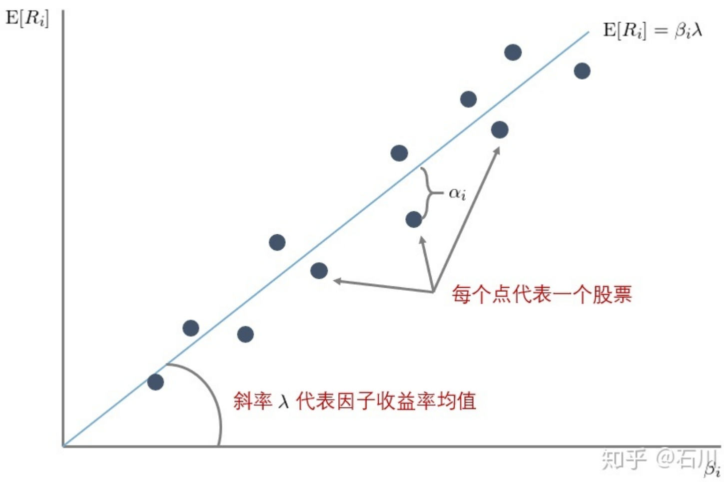

图 1: CAPM：Security Market Line

有几点需要注意，在这里使用 $\mathbb{E}[\boldsymbol{R}_i^e]$ 为代表资产的预期超额收益率，而非如 CAPM 中表示的 $\mathbb{E}[R_i] - R_f$，是因为在实证中，经常采用多空对冲建立投资组合，此时便无需再减去无风险收益率。另外，学界研究的对象始终为资产的预期草娥收益，因此有时将“超额”二字省略。

2.2 异象

而在实际过程中，等式两侧并不相等，而存在着定价误差 α_i（Pricing error）：

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

定价误差有可能由两方面产生：

- 模型设定偏误，即等式右侧遗漏了重要的因子，当被遗漏因子加入后，可消除定价误差
- 模型设定没有问题，但由于资产收益的实际数据只是总体的一个样本，那么误差总是存在的，此时需要通过统计的方法检验误差 α_i 是否显著不为零：
  - 若 α_i 并非显著偏离于零，则出现只是样本问题
  - 若 α_i **显著偏离零**，则说明了可以通过套利而获得超额收益的机会，市场对该资产出现错误定价（Mispricing），从而导致了实际预期收益率与多因子模型下的预期收益率出现偏离

假使我们根据基本面特征或量价指标等特征，挑选出一揽子股票并构建多空投资组合。若该组合的收益率无法被多因子模型（如 3 因子、4 因子、5 因子模型）解释，则称该特征为一个**异象**（Anomaly）。即该特征获得了多因子模型无法解释 $alpha$ 收益率，但从有效市场假说出发，市场中不应该存在很多异象。同时在学界不断的挖掘中，获得了 400+ 个异象，且在样本内都获得了很高的 t-statistics，这里可能存在两个原因：

- 数据挖掘，大量的异象在样本内被挖掘出，因此 Harvey，Liu 和 Zhu（2016）提出异象收益率的 t-statistic 至少要超过 3.0，而非传统的 5% 显著性对应的 2.0，才可能真正有效，而非来源于运气
- 模型相关，若以 $CAPM$ 为定价模型，那么许多异象都能获得 CAPM 无法解释的 α 收益率，同时随着定价模型中因子个数的增加，更多的异象变得不再显著，而真正的定价模型是未知的

Hou，Xue 和 Zhang（2017）长达 146 页对异象的研究中，复现了学术界提出的 447 个异象，涵盖动量（57 个）、价值/成长（68 个）、投资（38 个）、盈利（79 个）、无形资产（103 个）、以及交易摩擦（102 个）六大类。对于这 447 个异象，在排除了微小市值股票的影响后，其中 286 个（64%），在 5% 的显著性水平下不再显著（下同）。若按照 Harvey，Liu 和 Zhu（2016）的建议把 t-statistic 的阈值提升到 3.0，其中 380 个（85%）异象不再显著。最后，如果使用 Hou，Xue 和 Zhang（2015）提出的 4 因子模型作为定价模型，那么其中 436 个（98%）异象不再显著，仅有 11 个异象显著。

对于超额收益，学术界和业界主流的两种解释是错误定价和风险补偿，错误定价意味着投资者可以通过合理的策略获得潜在的超额收益；而风险补偿则意味着投资者获得的收益是以承担额外风险为代价的。

2.3 因子

异象有可能能成为优秀的因子，但不是所有异象都是因子。因为作为一个因子（Factor），需要能够解释资产（个股或投资组合）预期收益率截面上的差异，并有增量贡献。具体而言：

- 异象从方程的左侧，移动到右侧称为一个因子，称为解释变量，需要考察期是否能解释预期收益率截面上的差异
- 由于多个异象之间并不完全独立，需要排除相关性的影响，考察是否有增量贡献

如价值因子，也可以采用 E/P 或 B/P 构建 High-Minus-Low 组合，若同时使用，两者相关性必然很高，因此若使用其一作为价值因子，另一因子对资产预期收益率截面差异的解释能力的增量贡献将变得很低，无法称为因子。

对于因子模型接下去的问题就是，在构建多因子模型时：选取因子因子的数目；以及选取哪些因子。第一个问题，因遵循简约法则（The Law of Parsimony），或奥卡姆剃刀（Occam's razor）。若从 ICAPM（Intertemporal CAPM）的角度理解多因子模型，每个因子应代表某种状态变量（State variable），即为投资者想要对冲的某种风险。因此，因子的个数应该是有限的。目前主流的多因子模型如下：

- Fama-French 三因子模型（Fama and French 1993）：多因子模型的开山鼻祖，包括 MKT、HML 以及 SMB 三因子。其中包含了 MKT 市场因子，HML 价值因子，与 SMB 规模因子
- Carhart 四因子模型（Carhart 1997）：在 Fama-French 三因子模型上加上了动量 MOM 因子。
- Novy-Marx 四因子模型（Novy-Marx 2013）：包含 MKT，HML，MOM 以及 PMU 四个因子，其中 PMU 所用的财务指标是 Gross Profit-to-Asset，代表 Profitability 维度
- Fama-French 五因子模型（Fama and French 2015）：Fama 和 French 在其三因子模型的基础上加入了 CMA 和 RMW 两个因子，分别代表 Investment 和 Profitability 两个维度。
- Hou-Xue-Zhang 四因子模型（Hou, Xue and Zhang 2015）：包含 MKT，SMB，IVA 以及 ROE。其中 IVA 是 Total assets 的年增长率，代表 Investment 维度
- Stambaugh-Yuan 四因子模型（Stambaugh and Yuan 2016）：包含 MKT，SMB，MGMT 和 PERF 四个因子。MGMT 和 PERF 分别使用了 6 个和 5 个指标，代表 Management 以及 Performance 相关的两个 Mispricing 因子。虽然该模型只有四个因子，但它用到的基本面和量价指标多达 12 个。
- Daniel-Hirshleifer-Sun 三因子模型（Daniel, Hirshleifer and Sun 2018）：在 MKT 的基础上，使用 PEAD 和 FIN 两个指标作为短期和长期行为因子（Behavioral factors）的代理指标，构建了三因子模型。该模型由于包括了传统的 MKT 市场因子，又包括行为因子，故称为复合模型。

对于第二个问题，则涉及了不同多因子模型之间的比较。目前学界主要有三种方法：

- GRS tests
- Mean-Variance Spanning tests
- Bayesian approach

【待整理】

GRS tests（Gibbons，Ross 和 Shanken 1989）检验 n 个资产在给定因子模型下的定价错误 α，是否在统计上联合为零（jointly equal to zero）。在比较两个多因子模型时，使用两个模型的因子互为资产和定价模型进行检验。

Mean-Variance Spanning tests 考察 n 个已知资产构建的 mean-variance 有效前沿能否包含某个新资产（Huberman 和 Kandel 1987）。在比较两个多因子模型时，使用每个模型的因子构建有效前沿，并逐一检验其能否包含另一个模型中的因子。

在 Bayesian approach 中，假设比较两个多因子模型 M_1 和 M_2，数据集使用 D 表示。令 $P(M_1)$ 和 $P(M_2)$ 为这两个模型的先验概率，且有 $P(M_1) + prob(M_2) = 1$（这里假设把多个模型两两比较）。根据贝叶斯定理有：

$$P(M_i \,|\, D) = \frac{P(M_i)P(D \,|\, M_i)}{P(M_1)P(D \,|\, M_1) + P(M_2)P(D \,|\, M_2)}$$

其中：

$$P(D \,|\, M_i) = \int_{\theta_i} P(\theta_i)P(D \,|\, \theta_i)d\theta_i$$

上式中，$P(\theta_i)$ 是模型 i 参数的先验分布，$P(D \,|\, \theta_i)$ 是模型 i 的似然函数。上述贝叶斯方法的核心在于确定 $P(\theta_i)$。根据 Pastor 和 Stambaugh（2000）以及 Barillas 和 Shanken（2018）的理论，它和以两个模型中的全部因子作为资产所构成的投资组合的预期最大夏普率的平方与市场夏普率的比值有关。

2.4 系统性风险

【待整理】

除了市场因子以外的风险都是可以被分散的，所以只要是超过市场组合收益的部分都叫做超额收益。

多因子模型的表达式同样强调，只有那些影响众多资产收益率共同运动的风险，而非资产的特质性风险（即可以通过分散化规避掉的风险），才是预期收益率的来源。

系统风险（Systematic risk）的暴露程度，即对于市场风险暴露的大小。即资产的预期超额收益率，由市场组合（市场因子）的预期超额收益率与该资产对市场风险的暴露大小决定。或可以理解为，单项资产的 β 系数是指资产预期超额收益率与市场组合预期超额收益率之间变动关系的敏感程度。

系统性风险（Systematic risk），又称市场风险或不可分散风险，是影响所有资产的、不能通过资产组合而消除的风险。这部分风险是由那些影响整个市场的风险所引起的，无论怎样分散投资，也不可能消除系统性风险。避免集中投资于单一市场可减少系统性风险。单项资产、证券资产组合或不同公司受系统性风险影响不一样，系统性风险的大小通常用 beta 系数（β 系数）来衡量。

3 多因子模型回归检验

对于多因子截面关系式：

$$\mathbb{E}[R_i] = \alpha_i + \boldsymbol{\beta}_i' \boldsymbol{\lambda}$$

在上述截面关系式中，α_i 代表股票 i 的定价错误。如果我们能够在统计上证明所有股票的 α_i 都很接近零，那么这个多因子模型就是很好的模型。因为这些因子能够较好的解释个股截面预期收益率的差别。因此，多因子模型的回归检验中的重中之重，就是所有这些 α_i 联合起来是否在统计上足够接近零。

因此，多因子模型的回归检验可以总结为：

- 挑选因子，计算个股在这些因子上的暴露 β_i
- 找到个股超额收益率均值 $\mathbb{E}[R_i]$ 和因子暴露 β_i 在截面上的关系
- 计算个股的定价误差 α_i，联合检验这些 α_i 是否在统计上为零

https://zhuanlan.zhihu.com/p/40984029

注意：超额收益率为 t+1 期，而因子为 t 期，由于对于所有股票，无风险利率相同，不会影

响 beta。并且向量都为列向量，其转置为行向量。

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \qquad \boldsymbol{a'} = \begin{bmatrix} a_1 & a_2 & \dots a_n \end{bmatrix}$$

3.1 时序回归

此时的因子为投资组合收益率，如经典的 HML，SMB 等，那么此时可以通过时序回归，来分析个股超额收益率 $\mathbb{E}[R_i]$ 与因子暴露 β_i 之间在截面上的关系。

对于个股 i，直接进行时间序列回归，即

$$R_{i,t} = \alpha_i + \boldsymbol{\beta_i'} \boldsymbol{f_t} + \varepsilon_{i,t} \ , \quad t = 1, 2, \dots, T$$

对 $R_{i,t}$ 与 f_t 在时序上取均值 $\mathbb{E}_T(\cdot)$，就得到了个股超额收益率与因子暴露在截面上的关系：

$$\mathbb{E}_T[R_i] = \alpha_i + \boldsymbol{\beta_i'} \mathbb{E}_T[\boldsymbol{f_t}] \ , \quad t = 1, 2, \dots, T$$

此时时序回归得到的截距 α_i 即为个股 i 的定价误差。Black，Jensen 和 Scholes（1972）基于如上的论述给出了时序回归法中求解因子预期收益率的简单方法，因子收益率 f_t，在时序上的均值就是因子的预期收益率：

$$\boldsymbol{\lambda} = \mathbb{E}_T[\boldsymbol{f}]$$

对于 N 支个股，每支个股 i 都有一组 $(\beta_i, \mathbb{E}[R_i])$，以 $\mathbb{E}[R_i] = \boldsymbol{\beta'} \mathbb{E}[\boldsymbol{f}]$ 作图，此时斜率为因子收益率 $\lambda = \mathbb{E}_T[\boldsymbol{f}]$。将当 $\boldsymbol{\beta_i} = 0$ 时与 $\boldsymbol{\beta_i} = 1$ 代入，可知该直线过 $(0, 0)$ 与 $(1, \mathbb{E}[\boldsymbol{f}])$ 两点。即对于因子投资组合，其对自身的因子暴露为 1。

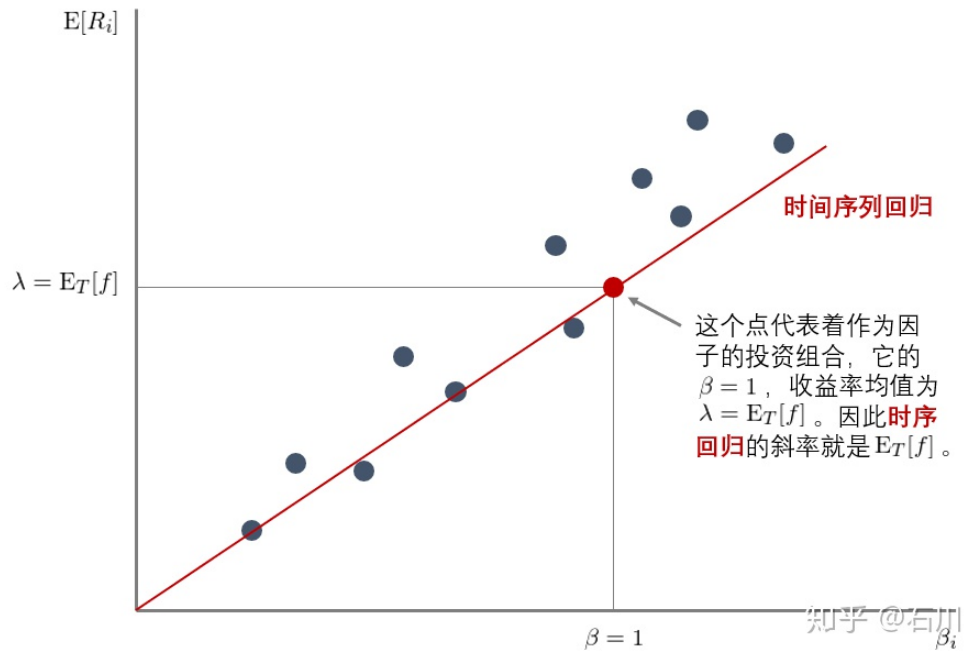

图 2: 时序回归

对于检验 α_i 联合起来是否统计上为零，若残差不相关和同方差，标准误可以由 OLS 标准公式计算。若残差满足独立同分布且为正态分布，可以使用 GRS 检验。而当残差之间存在相关性或者异方差，则需要使用 GMM。

3.2 截面回归

相比于时序回归，使用截面回归的优势是，因子的选择范围更广。可以为 GDP、CPI、利率等宏观经济指标，而不仅投资组合收益率。进行截面回归，首先需先进行时序回归，用于确定因子暴露 $\boldsymbol{\beta}$，即个股收益率对这些因子在时序上的敏感程度。对于个股 i，应有:

$$R_{i,t} = a_i + \boldsymbol{\beta_i'} \boldsymbol{f_t} + \boldsymbol{\varepsilon_{i,t}}\,, \quad t = 1, 2, \ldots, T$$

由于此时的 $\boldsymbol{f_t}$ 并非个股收益率，阶矩 a_i 并非个股定价误差。得到了个股因子暴露 β 之后，再进行截面回归。此时等式左边为 $\mathbb{E}_T[R_i]$ 为整个 T 其的收益率均值，或上述 $\mathbb{E}_T[R_i]$，而右侧为时序

回归得到的 $\boldsymbol{\beta_i}$。对于 n 支个股，每支个股都有一组 $(\mathbb{E}[R_i], \boldsymbol{\beta_i})$，则截面回归的表达式为：

$$\mathbb{E}[R_i] = \alpha_i + \boldsymbol{\beta}_i'\boldsymbol{\lambda}$$

此时的截距即为定价误差，但由于只进行了一次回归，因此只能得到单个定价误差 α_i 与单个因子预期收益率 $\boldsymbol{\lambda}$。

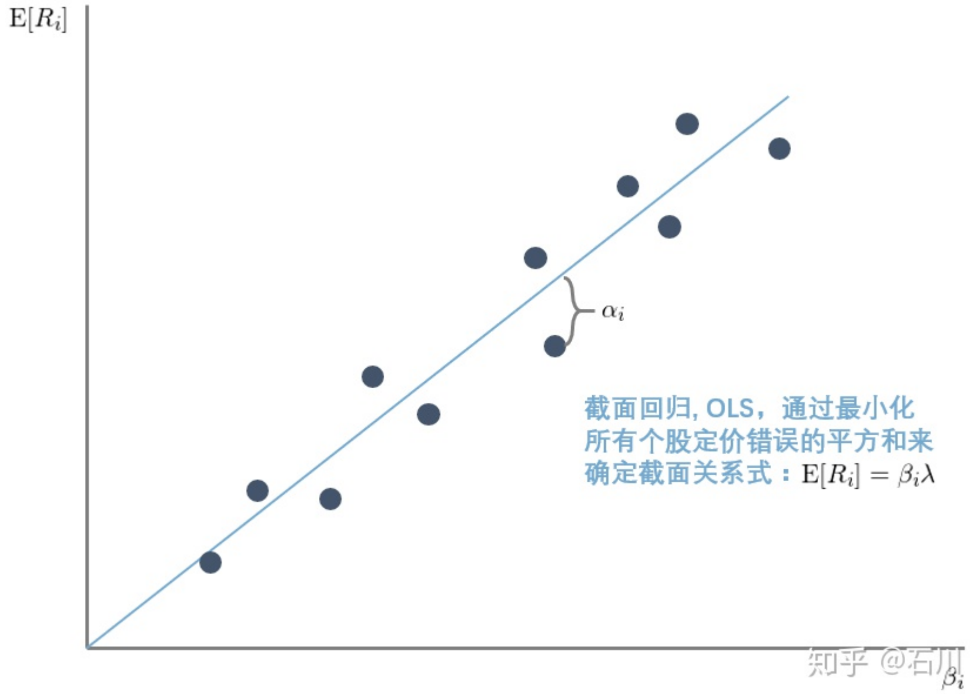

图 3: 截面回归

因此截面回归，也称为两步回归估计（Two-pass regression estimate）。首先，通过时序回归首先得到个股对因子的暴露。其次，对超额收益率取均值，即个股转换为点 $(\beta_i, \mathbb{E}[R_i])$，并对 n 支个股同时进行截面回归，最终的到定价误差与因子收益率。

【待整理】

由于截面个股残差的相关性，虽然不会影响 OLS 估计，但会导致 OLS 给出的标准误存在较大误差，因此可以使用 GLS 取代 OLS，由于 GLS 考虑了残差的协方差因此可以得到准确的标准误。但估计残差的协方差矩阵在现实中有较大的障碍。此时则可使用 GMM，由于截面回归的 $\boldsymbol{beta_i}$ 并不是使用真实的因子收益率回归得到，而是从时间序列上回归出的估计值，称为 generated regressors，存在误差。Shanken（1992）给出了修正方法称为 Shanken correction。同时利用 Shanken correction 与 GMM，即可检验 α_i 是否联合为 0。

3.3 时序回归 vs 截面回归

对于时序回归而言，因子收益率为在时序上的均值，即 $\boldsymbol{\lambda} = \mathbb{E}_T[\boldsymbol{f}]$，来得到隐含的截面关系。因此 $\mathbb{E}[R_i] = \boldsymbol{\beta_i}\boldsymbol{\lambda}$ 必然过原点，以及因子收益率的投资组合 $(1, \mathbb{E}_T[\boldsymbol{f}])$（$\boldsymbol{\beta_i} = 2$）。而在截面回归中，因子暴露已经通过时序回归确定，而在第二次进行截面回归时，充分使用了个股数据，若使用 OLS，以最小化个股定价误差 α_i 的平方和为目标。因子收益率与时序中直接通过计算均值的方式不同，更加合理。

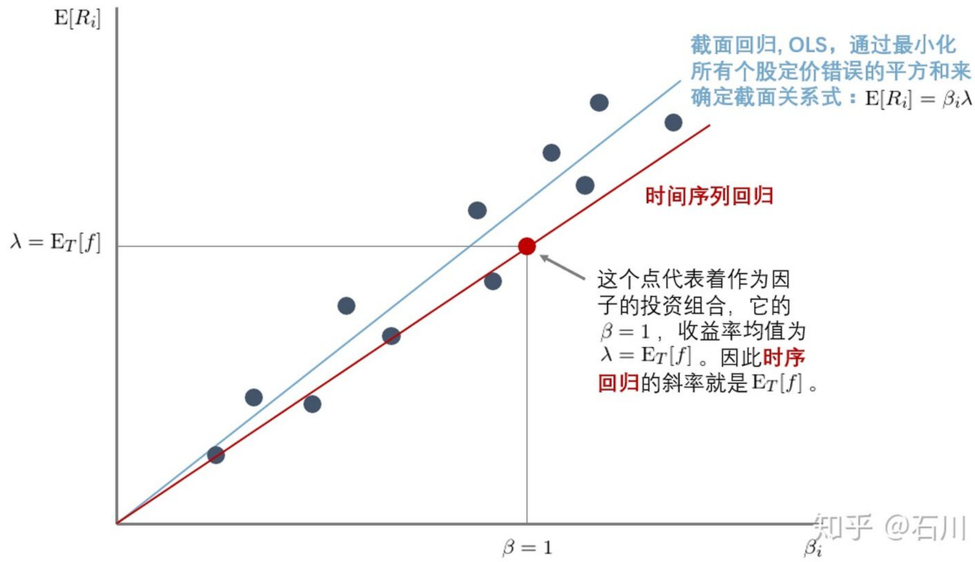

图 4: 时序回归与截面回归

3.4 Fama-MacBeth 回归

3.4.1 主要思想

在 Fama 和 MacBeth（1973）中，提出了 Fama-MacBeth Regression（Fama and MacBeth 1973），目的是为了检验 CAPM。Fama-MacBeth 回归也是两步回归，但与传统截面回归有所不同。

首先，与截面回归相同的是也需要先进行时序回归，得到资产 i 超额收益率（R_i 或 R_i^e）在因

子上的暴露 $\boldsymbol{\beta}_i$。此时可对 $t=1,\ldots,T$ 整体时序样本，只进行一次时序回归，得到一个不随时间改变的因子暴露。此时由于 $\boldsymbol{\beta_i}$ 不变，那么对于传统截面回归的先均值再回归，与 FM 截面回归的先回归再均值，两种方式在所有 T 期上得到的估计是相同的（但 FM 截面回归在检验上一定的优势）。或采用在 Fama 和 MacBeth（1973）中，滚动窗口的方法估计 $\boldsymbol{\beta_i}$，因此对于不同的 t，$\boldsymbol{\beta_i}$ 发生改变。

$$\boldsymbol{R}_{i,t}=\boldsymbol{a}_i+\boldsymbol{\beta_i}\boldsymbol{F_t}+\boldsymbol{\varepsilon_{i,t}}\ ,\quad \forall\, t=1,2,\ldots,T$$

其次，在传统截面回归中，第二步只进行一次截面回归。即对资产 i，在时序 $t=1,2,\ldots,T$ 上对 $R_{i,t}$ 取均值，得到 $\mathbb{E}_T[R_i]$。与该资产 i 的因子暴露 $\boldsymbol{\beta}_i$ 进行回归，只进行了 <u>1 次回归</u>：

$$\mathbb{E}_T[\boldsymbol{R}_i]=\boldsymbol{\beta_i}\boldsymbol{\lambda_i}+\boldsymbol{\alpha}_i\ ,\quad \forall\, i=1,2,\ldots,N$$

而在 Fama-MacBeth 回归中，第二步并不对 $R_{i,t}$ 在时序上取均值，需要对每个 t 时间上进行一次截面回归，一共进行 <u>T 次回归</u>：

$$\boldsymbol{R}_{i,t}=\boldsymbol{\beta}_i\boldsymbol{\lambda}_{i,t}+\boldsymbol{\alpha}_{i,t}\ ,\quad \forall\, i=1,2,\ldots,N\ \forall\, t=1,2,\ldots,T$$

Fama-MacBeth 截面回归和传统截面回归的相同点和区别是：

- FM 回归与截面回归，都需先进行时序回归，以确定个股的因子暴露 $\boldsymbol{\beta_i}$
- 传统截面回归将 $R_{i,t}$ 在时序上取均值得到 $\mathbb{E}[R_i]$，再进行 1 次截面回归，得到 α_i 和 λ
- 对于 FM 回归，在各个不同的 t 时刻使用 $R_{i,t}$ 与 $\boldsymbol{\beta_i}$ 进行回归，再将回归结果在时序上取均值得到 $\alpha_i=\mathbb{E}[\alpha_{i,t}]$ 与 $\lambda=\mathbb{E}[\lambda_t]$???
- 即传统回归为先取均值，后回归。而 FM 回归为先回归，后取均值

3.4.2 具体回归

假设共有 $t=1,\ldots,T$ 期数据，并且有 $i=1,\ldots,N$ 个资产与 $j=1,\ldots,K$ 个因子，资产 i 的超额收益率为 $R_{i,t}$。第一步，通过对每个资产 i 进行回归，得到每个资产 i 对不同因子的风险暴露 $\boldsymbol{\beta}_i$。注意此时获得的截距项，用 a_i 表示，而并非定价误差 α，因为此时的解释变量并非因子收益率。注意，如上所述 T 可为全样本，也可以为滚动窗口。

对各个资产进行多元 OLS 回归时，在因子 $\boldsymbol{F}$ 前需加入常数列 $\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}$，得到截距项，因此有 $m+1$ 列。回归得到 i 资产的因子风险暴露 $\boldsymbol{\beta}_i$，写为矩阵形式有：

$$\underset{T\times N}{\boldsymbol{R}} = \underset{T\times(K+1)}{\boldsymbol{F}}\underset{(K+1)\times N}{\hat{\boldsymbol{\beta}}} + \underset{T\times N}{\boldsymbol{\epsilon}}$$

其中：

$$\boldsymbol{R} = \begin{bmatrix}\boldsymbol{R}_1 & \dots & \boldsymbol{R}_N\end{bmatrix} = \begin{bmatrix}R_{1,1} & \dots & R_{1,N}\\ \vdots & \vdots & \ddots\\ R_{T,1} & \dots & R_{T,N}\end{bmatrix}$$

$$\boldsymbol{F} = \begin{bmatrix}\boldsymbol{F}_1^{'}\\ \vdots\\ \boldsymbol{F}_T^{'}\end{bmatrix} = \begin{bmatrix}1 & F_{1,1} & \dots & F_{1,K}\\ \vdots & \vdots & \ddots & \vdots\\ 1 & F_{T,1} & \dots & F_{T,K}\end{bmatrix}$$

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix}\hat{\boldsymbol{\beta}}_0^{'}\\ \hat{\boldsymbol{\beta}}_1^{'}\\ \vdots\\ \hat{\boldsymbol{\beta}}_K^{'}\end{bmatrix} = \begin{bmatrix}\hat{\beta}_{0,1} & \dots & \hat{\beta}_{0,N}\\ \hat{\beta}_{1,1} & \dots & \hat{\beta}_{1,N}\\ \vdots & \ddots & \vdots\\ \hat{\beta}_{K,1} & \dots & \hat{\beta}_{K,N}\end{bmatrix}$$

$$\boldsymbol{\epsilon} = \begin{bmatrix}\boldsymbol{\epsilon}_1 & \dots & \boldsymbol{\epsilon}_N\end{bmatrix} = \begin{bmatrix}\epsilon_{1,1} & \dots & \epsilon_{1,N}\\ \vdots & \vdots & \ddots\\ \epsilon_{T,1} & \dots & \epsilon_{T,N}\end{bmatrix}$$

第二步，在 $t=1,\dots,T$ 期，进行 T 次截面回归。即对于每个资产 i，都需进行 T 次回归，此时需注意 $\hat{\boldsymbol{\beta}}$ 由上一步得到，为估计值。对于 t 期而言：

$$\underset{N\times 1}{\boldsymbol{R}_t} = \underset{N\times K}{\hat{\boldsymbol{\beta}}}\underset{K\times 1}{\boldsymbol{\lambda}_t} + \underset{N\times 1}{\boldsymbol{\alpha}_t}$$

其中：

$$\boldsymbol{R}_t = \begin{bmatrix} R_{1,t} \\ \vdots \\ R_{N,t} \end{bmatrix} \qquad \hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_{1,1} & \dots & \hat{\beta}_{1,K} \\ \vdots & \ddots & \vdots \\ \hat{\beta}_{N,1} & \dots & \hat{\beta}_{N,K} \end{bmatrix}$$

$$\boldsymbol{\lambda}_t = \begin{bmatrix} \lambda_{1,t} \\ \vdots \\ \lambda_{K,t} \end{bmatrix} \qquad \boldsymbol{\alpha}_t = \begin{bmatrix} \alpha_{1,t} \\ \vdots \\ \alpha_{N,t} \end{bmatrix}$$

如上所述，若使用滚动方法。即对于 t 期，使用截止 $t-1$ 期的一段给定窗口的历史数据进行第一步的时序回归，估计 $\hat{\boldsymbol{\beta}}_{i,t-1}$。并且使用其作为第二步 t 期截面回归的解释变量：

$$\boldsymbol{R}^e_{i,t} = \hat{\boldsymbol{\beta}}_{i,t-1}\boldsymbol{\lambda}_t + \boldsymbol{\alpha}_{i,t}\ , \quad \forall\, i = 1, 2, \dots, N$$

或如 Cochrane（2005）中，在截面回归中加入截距项 $\boldsymbol{\gamma}_t$：

$$\boldsymbol{R}^e_{i,t} = \boldsymbol{\gamma}_t + \hat{\boldsymbol{\beta}}_{i,t-1}\boldsymbol{\lambda}_t + \boldsymbol{\alpha}_{i,t}\ , \quad \forall\, i = 1, 2, \dots, N$$

3.4.3 检验

那么此时对于 $\alpha_{i,t}$ 与 λ_t 都有 T 次回归结果。对于单个资产 i，从 $t=1$ 时刻到 $t=T$ 时刻，有 $(\boldsymbol{\beta_{i,1}}, \mathbb{E}[R_{i,1}])\,, (\boldsymbol{\beta_{i,2}}, \mathbb{E}[R_{i,2}])\,, \dots, (\boldsymbol{\beta_{i,T}}, \mathbb{E}[R_{i,T}])$。即可以以此估计定价误差以及因子收益率：

$$\hat{\lambda} = \frac{1}{T}\sum_{t=1}^{T}\lambda_t$$
$$\hat{\alpha}_i = \frac{1}{T}\sum_{t=1}^{T}\alpha_{i,t}$$

不同于传统截面回归，只得到 λ 和 α 的一个样本估计，在 FM 截面回归中得到 T 个 λ 和 α

的样本估计，这样就可以求出两者标准误为：

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

【待整理】由上面的介绍可知，Fama-MacBeth 回归的最大优点是它排除了残差截面相关性对标准误的影响。股票的残差收益率在截面上具有很高的相关性，因此该修正对于准确计算标准误至关重要。下面来说说它的不足。首先，Fama-MacBeth 回归对于残差在时序上的相关性无能为力。如果残差在时序上存在相关性，则需要对 Fama-MacBeth 回归得到的标准误进一步修正。Petersen (2009) 分析了不同的回归技术在分析面板数据（panel data）时由于忽略残差的时序或截面相关性而导致不准确的标准误（低估了其真实值）。这篇文章非常值得一读。其次，上文提到，在截面回归中用到的 $\boldsymbol{\beta}_i$ 并不是已知的，而是通过时间序列得到的估计值（generated regressors），因此存在误差。Fama-MacBeth 回归对此也无能为力，需要 Shanken correction。

3.5 Barra 多因子模型

Barra 多因子模型也是截面回归模型，其考虑了行业因子和来自基本面和技术面的风格因子。但与传统的截面回归模型不同的是在 Barra 模型中，因子暴露并非来自时间序列回归，而是直接来自基本面或者技术面数据本身。

例如 Book-to-Market ratio，在 Fama-French 三因子模型中，其被用来构建 HML（High-Minus-Low）投资组合，投资组合的收益率作为因子。由时序回归决定个股在这个因子上的暴露，与个股实际的 BM 无关。但在 Barra 模型中，BM 直接用于确定因子暴露，但需要进行标准化。有了因子暴露，Barra 截面模型与传统截面模型相同，都是通过截面回归确定因子收益率。因此，Barra 模型（业界代表）和学术界流行的因子模型最大的不同就是因子暴露 $\boldsymbol{\beta}_i$ 的确定。

对于两种确定因子暴露的方法，通过时间序列得到的 $\boldsymbol{\beta}$，经过平滑，变化更加缓慢。而直接使用基本面数据获得 $\boldsymbol{\beta}$，可以更快的捕捉公司的变化。需要注意的是，这些因子都需要进行标准化，不能直接使用原始数据。例如，公司市值不经过标准化而作为因子暴露，公司市值差异巨大。假设 A 公司市值为 B 公司的 100 倍，此时显然不能说 A 公司的超额收益率对于市值因子的暴露为 B 公司的 100 倍。所以对于市值因子或其他风格因子，常见的是首先取对数，然后再进行标准化。对

于其他的风格因子，也需要采用相应的标准化处理。在 Barra 的文档中对如何标准化因子暴露有详细的说明。对于行业因子，Barra 将因子暴露处理为 Binary 变量或虚拟变量，例如工商银行在银行业的暴露为 1，而在其他行业的暴露为 0。

3.6 GRS

3.7 GMM

GMM，它可以轻松的求出我们需要的各种量（Hansen 功不可没啊）。另外值得一提的是，在截面回归时用到的 β_i 并不是已知、真实的，而是从时间序列回归得出的估计值，它们称为 generated regressors，存在误差。Shanken (1992) 给出了解决该问题的修正方法，称为 Shanken correction。利用 Shanken correction 和 GMM，就可以检验 α 是否为零了。

如今我们有了 GMM 这样的大杀器，能够方便的处理残差的各种相关性。

附录 A APT 推导

第一步，假设资产的收益率，在单因子情形下，满足如下线性模型：

$$R_i = \mu_i + \beta_i f + \varepsilon_i \ , \quad i = 1, \ldots, n$$

此时 R_i 为资产收益率，μ_i 资产 i 的预期收益率，β_i 是资产在因子上的暴露，而 f 是因子取值，而非因子风险溢酬，ε_i 为资产 i 收益率的随机扰动或特质性收益率，并满足 $\mathbb{E}[f] = \mathbb{E}[\varepsilon_i] = 0$，若改写为向量形式，为 Ross 在 APT 中使用的收益率模型，有：

$$\boldsymbol{R} = \boldsymbol{\mu} + \boldsymbol{f}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

第二步，构建一个 arbitrage portfolio，这个投资组合中资产的权重 ω 满足如下两点特性。首先，该投资组合是零额投资的，即：

$$\boldsymbol{w'1 = 0}$$

其中有 $\mathbf{1}$ 为全为 1 的向量。其次，并且 $\boldsymbol{w'\beta} = 0$，即该投资组合在该因子上的暴露为零。此

时这个投资组合的收益率应为：

$$R_p = \boldsymbol{w}'\boldsymbol{R} = \boldsymbol{w}'\boldsymbol{\mu} + \boldsymbol{w}'\boldsymbol{\beta}\boldsymbol{f} + \boldsymbol{w}'\boldsymbol{\varepsilon}$$

由上可知，这个投资组合的收益率可化简为：

$$R_p = \boldsymbol{w}'\boldsymbol{R} = \boldsymbol{w}'\boldsymbol{\mu}$$

第三步，运用无套利约束，可知根据 $\boldsymbol{w}$ 构建的投资组合有如下性质：零额投资；对因子的暴露为零（因此该组合没有系统性风险）；没有特质性风险暴露（因为组合中特质性收益率为零）。换言之，这样一个投资组合，既没有资金投入又没有风险暴露，因此根据无套利约束条件，它的收益率必须为零，即：

$$R_p = \boldsymbol{w}'\boldsymbol{\mu} = 0$$

根据几何可知，$w'1 = w'\beta = w'\mu = 0$，说明 $\boldsymbol{w}'$ 与 $\mathbf{1}$，$\boldsymbol{\beta}$ 和 $\boldsymbol{\mu}$，都相互垂直。因此 $\boldsymbol{\mu}$ 必然在 $\mathbf{1}$ 和 $\boldsymbol{\beta}$ 构成的平面内。

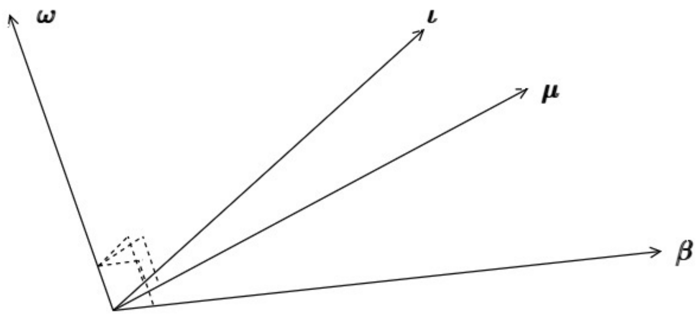

图 5: APT

因此在数学上，资产预期收益率 $\boldsymbol{\mu}$ 可以写成 $\mathbf{1}$ 与 $\boldsymbol{\beta}$ 的线性组合，即：

$$\boldsymbol{\mu} = \gamma_1\mathbf{1} + \gamma_2\boldsymbol{\beta}$$

那么此时上式应对任何资产都成立，为了求解 γ_1 与 γ_2 可代入特殊资产，无风险资产（R_f），

与市场组合 (R_M)。由于无风险资产的因子暴露为零，代入上式可得：

$$R_f = \gamma_1$$

对于市场组合，其 $beta = 1$，并将 $\gamma_1 = R_f$ 代入，可得：

$$R_M = R_f + \gamma_2 \times \mathbf{1}$$

因此有：

$$\gamma = R_M - R_f$$

代入原式，最终可得到 CAPM 的表达式：

$$\mu_i = R_f + \beta_i(R_M - R_f)$$

在此基础上，扩展至多因子，即得到 APT 模型：

$$\mathbb{E}[R_i^e] = \mu_i - R_f = \beta_{i,1}\lambda_1 + \beta_{i,2}\lambda_2 + \cdots + \beta_{i,K}\lambda_K$$