中国股市FF三因子模型实证检验

蒋志强

zqjiang.ecust@qq.com

常用的多因子模型

Fama-French三因子模型(Fama and French, 1996)

FF模型之前的实证结果

- 规模效应:小市值公司较大市值公司有更高的平均收益率(Banz,1981)
- 财务杠杆:负债与股票市值之比可解释股票平均收益率(Bhandari, 1988)
- BE/ME:账面市值比(B/M)与美国股票平均收益率正相关(Stattman 1980,Lanstein 1985)
- E/P:盈余与股票市值之比(市盈率倒数)有助于解释股票平均收益率(Basu, 1977)

FF模型

风险因素

- 市场风险溢价R_m-R_f
- 账面市值比（B/M）溢价HML
- 规模溢价SMB

FF模型

规模组合构建：

1. 每年6月底计算各上市股票的市值（规模）
2. 根据市值中位数将股票分成两类

   规模小的所有股票称为S类

   规模大的所有股票称为B类

FF模型

账面市值比组合构建：

每年(t年)6月基于账面市值比B/M分类股票

1. 计算 t-1 年末股票账面值 B
2. 计算 t-1 年末股票最后一个交易日的股价×发行股数 M
3. 计算B/M并分成三组：

   最低30%的股票记为L类

   中间40%的股票记为M类

   最高30%股票记为H类

FF模型

规模（S、B）和B/M（H、M、L）

两两组合构成6个投资组合

S/L、S/M、S/H、B/L、B/M、B/H

规模溢价SMB = 公司规模最小的投资组合收益率—公司规模最大的投资组合收益率

$$SMB = \frac{S/L + S/M + S/H}{3} - \frac{B/L + B/M + B/H}{3}$$
$$= \frac{(S/L - B/L) + (S/M - B/M) + (S/H - B/H)}{3}$$

账面市值比溢价HML = 高账面市值比投资组合收益率 - 低账面市值比投资组合收益率

$$HML = \frac{B/H + S/H}{2} - \frac{B/L + S/L}{2}$$

FF模型

市场风险溢酬为R_m - R_f

- ➢ R_f为无风险利率
- ➢ R_m所有股票的市值加权投资组合收益率(市场指数的收益率替代)

资产组合 i 的 FF 模型为

$$r_i = \alpha_i + \beta_{mi} r_{mt} + \beta_{si} SMB + \beta_{hi} HML + \xi_i$$

FF模型的时间序列估计与检验

单资产多因子模型检验步骤:

1. 用OLS估计参数, 得到 α_i 的估计值
2. 计算 $\alpha_i = 0$ 的 t 检验统计量
3. 确定显著性水平, 比较分位数或计算 p值, 作出统计推断

FF模型的时间序列估计与检验

假设有T个样本，如何进行参数估计和检验？

$$r_{it} = \hat{\alpha}_i + \hat{\beta}_{mi} r_{mt} + \hat{\beta}_{si} SMB_t + \hat{\beta}_{hi} HML_t + \xi_{it}$$

- ✓ **Fama-French三因子模型**

$$r_i = \alpha_i + \beta_{mi} r_{mt} + \beta_{si} SMB + \beta_{hi} HML + \varepsilon i$$

- ✓ 抽象为多元线性回归模型

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

- ✓ 最小二乘线性拟合进行参数估计

$$\min Q = \sum (Y_i - \beta_0 - \beta_1 X_1 - \beta_2 X_2 - \beta_3 X_3)^2$$

FF模型的时间序列估计与检验

整理后：

$$\begin{cases} n\hat{\beta}_0 + \sum X_{1i}\hat{\beta}_1 + \sum X_{2i}\hat{\beta}_2 + \sum X_{3i}\hat{\beta}_3 = \sum Y_i \\ \sum X_{1i}\hat{\beta}_0 + \sum X_{1i}^2\hat{\beta}_1 + \sum X_{1i}X_{2i}\hat{\beta}_2 + \sum X_{1i}X_{3i}\hat{\beta}_3 = \sum X_{1i}Y_i \\ \sum X_{2i}\hat{\beta}_0 + \sum X_{2i}X_{1i}\hat{\beta}_1 + \sum X_{2i}^2\hat{\beta}_2 + \sum X_{2i}X_{3i}\hat{\beta}_3 = \sum X_{2i}Y_i \\ \sum X_{3i}\hat{\beta}_0 + \sum X_{3i}X_{1i}\hat{\beta}_1 + \sum X_{3i}X_{2i}\hat{\beta}_2 + \sum X_{3i}^2\hat{\beta}_3 = \sum X_{3i}Y_i \end{cases}$$

$$A = \begin{bmatrix} n & \sum X_{1i} & \sum X_{2i} & \sum X_{3i} \\ \sum X_{1i} & \sum X_{1i}^2 & \sum X_{1i}X_{2i} & \sum X_{1i}X_{3i} \\ \sum X_{2i} & \sum X_{2i}X_{1i} & \sum X_{2i}^2 & \sum X_{2i}X_{3i} \\ \sum X_{3i} & \sum X_{3i}X_{1i} & \sum X_{3i}X_{2i} & \sum X_{3i}^2 \end{bmatrix} \qquad b = \begin{bmatrix} \sum Y_i \\ \sum X_{1i}Y_i \\ \sum X_{2i}Y_i \\ \sum X_{3i}Y_i \end{bmatrix}$$

FF模型的时间序列估计与检验

最小二乘估计量为: $\hat{\beta}_i = [\hat{\alpha}_i, \hat{\beta}_{mi}, \hat{\beta}_{si}, \hat{\beta}_{hi}]$

$$\hat{\beta}_i = A \backslash b$$

检验H_0: $\alpha_i = 0$ 的统计量（t 检验）

$$t_{\hat{\beta}} = \hat{\beta}_i./\sqrt{s^2\{\mathrm{diag}[(X^T X)^{-1}]\}}$$

$$s^2 = \frac{\sum e_t^2}{T - K - 1}$$

$$X = [\mathbf{1}, X_1, X_2, X_3]$$

FF模型的时间序列估计与检验

Fama-French三因子模型

$$r_i - r_f = \alpha_i + \beta_i r_{mt} + s_i SMB + h_i HML + \varepsilon i$$

数据（锐思数据库, 2009-2017年, 月度）：

行业指数（行业组合）、市场溢酬因子 $(r_m - r_f)$、
公司规模因子 (SMB)、公司价值因子(HML)、无风险收益 r_f

**行业组合：上证能源、上证金融、上证消费、上证材料、
上证工业、上证医药、上证信息**

FF模型的时间序列估计与检验

行业组合：上证能源、上证金融、上证消费、上证材料、上证工业、上证医药、上证信息

	α	β	s	h
材料	-0.0019	1.1989***	0.0432	-0.1363
工业	-0.0025	1.1745***	-0.0527	0.1447
金融	0.0038	1.1065***	-0.2955***	0.3742***
信息	0.0022	1.0917***	0.3180**	-1.0102***
能源	-0.0068	1.0711***	-0.037	0.2413*
医药	0.0087*	0.6921***	0.0546	-0.9498***
消费	0.0067*	0.7794***	0.0604	-0.5464***

- ✓ 材料和工业组合可用CAPM模型解释。不能拒绝$\alpha = 0$。
- ✓ 金融和信息组合可用FF三因子模型解释。不能拒绝$\alpha = 0$。
- ✓ 医药和消费组合不能完全被FF三因子模型解释。10%显著性拒绝 $\alpha=0$。
- ✓ 一般 $s > 0, h < 0$，金融和能源与此相反。通常剔除金融；能源金融属性强。

FF模型的时间序列估计与检验

多资产多因子模型检验步骤：

1. 用OLS估计参数，得到 α_i 的估计值
2. 联合检验 H_0: $\alpha_1 = \alpha_2 =, \ldots, \alpha_N = 0$，计算统计量
3. 确定显著性水平，比较分位数或计算 p值，作出统计推断

FF模型的时间序列估计与检验

假设有 N 个资产，每个资产有 T 个样本，如何进行参数估计和检验？

- ✓ Fama-French三因子模型写成向量形式

$$\mathbf{r}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}_m r_{mt} + \boldsymbol{\beta}_s SMB_t + \boldsymbol{\beta}_h HML_t + \boldsymbol{\varepsilon}_t$$

$$\mathbf{r}_t = [r_{1t}, r_{2t}, r_{3t}, \cdots, r_{Nt}]^T; \boldsymbol{\alpha} = [\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_t]^T$$

$$\boldsymbol{\beta}_m = [\beta_{m1}, \beta_{m2}, \beta_{m3}, \cdots, \beta_{mN}]^T; \boldsymbol{\beta}_s = [\beta_{s1}, \beta_{s2}, \beta_{s3}, \cdots, \beta_{sN}]^T$$

$$\boldsymbol{\beta}_h = [\beta_{h1}, \beta_{h2}, \beta_{h3}, \cdots, \beta_{hN}]^T; \boldsymbol{\varepsilon}_t = [\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t}, \cdots, \varepsilon_{Nt}]^T$$

- ✓ 最小二乘线性拟合进行参数估计

$$\min Q = \sum \boldsymbol{\varepsilon}_t^{\mathrm{T}} \boldsymbol{\varepsilon}_t$$

FF模型的时间序列估计与检验

假设有N个资产，每个资产有T个样本，如何进行参数估计和检验？

用OLS估计参数

$$\mathbf{r}_t = \alpha + \beta_m r_{mt} + \beta_s SMB_t + \beta_h HML_t + \varepsilon t$$

$$[\alpha, \beta_m, \beta_s, \beta_h] = \left[\sum_{t=1}^{T} \mathbf{r}_t, \sum_{t=1}^{T} \mathbf{r}_t r_{mt}, \sum_{t=1}^{T} \mathbf{r}_t SMB_t, \sum_{t=1}^{T} \mathbf{r}_t HML_t\right] \times$$

$$\begin{bmatrix} T & \sum_{t=1}^{T} r_{mt} & \sum_{t=1}^{T} SMB_t & \sum_{t=1}^{T} HML_t \\ \sum_{t=1}^{T} r_{mt} & \sum_{t=1}^{T} r_{mt}^2 & \sum_{t=1}^{T} SMB_t r_{mt} & \sum_{t=1}^{T} HML_t r_{mt} \\ \sum_{t=1}^{T} SMB_t & \sum_{t=1}^{T} r_{mt} SMB_t & \sum_{t=1}^{T} SMB_t^2 & \sum_{t=1}^{T} HML_t SMB_t \\ \sum_{t=1}^{T} HML_t & \sum_{t=1}^{T} r_{mt} HML_t & \sum_{t=1}^{T} SMB_t HML_t & \sum_{t=1}^{T} HML_t^2 \end{bmatrix}^{-1}$$

FF模型的时间序列估计与检验

假设有N个资产，每个资产有T个样本，如何进行参数估计和检验？

联合检验 H_0: $\alpha_1 = \alpha_2 =, \ldots, \alpha_N = 0$（GRS检验）

$$S_{\text{GRS}} = \frac{T - N - K}{N}\left[1 + \hat{\mu}_K^T \hat{\Omega}^{-1} \hat{\mu}_K\right]^{-1} \hat{\alpha}^T \hat{\Sigma}^{-1} \hat{\alpha} \sim F(N, T - N - K)$$

$$\hat{\mu}_K = \frac{1}{T}\sum_{t=1}^{T} \mathbf{f}_t, \mathbf{f}_t = [f_{1t}, \cdots, f_{Kt}]^T$$

$$\hat{\Omega} = \frac{1}{T}\sum_{t=1}^{T} (\mathbf{f}_t - \hat{\mu}_K)(\mathbf{f}_t - \hat{\mu}_K)^T$$

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{e}_t \mathbf{e}_t^T, \mathbf{e}_t = [e_{1t}, \cdots, e_{Nt}]^T$$

FF模型的时间序列估计与检验

Fama-French三因子模型

$$r_i - r_f = \alpha_i + \beta_i r_{mt} + s_i SMB + h_i HML + \varepsilon i$$

数据（锐思数据库，2009-2017年，月度）：

行业指数（行业组合）、市场溢酬因子 $(r_m - r_f)$、公司规模因子 (SMB)、公司价值因子(HML)、无风险收益 r_f

行业组合：上证能源、上证金融、上证消费、上证材料、上证工业、上证医药、上证信息

alpha1	alpha2	alpha3	alpha4	alpha5	alpha6	alpha7	GRS	pvalue
-0.0034	0.0074	0.0091	0.0020	0.0013	0.0109	0.0057	3.0412	0.0022