Barra 模型简介

Barra 因子是一组代表股票系统性风险来源（如规模、价值、动量、行业等）的量化指标，它通过衡量每只股票在这些因子上的暴露度，构建多因子模型来分解和预测组合的整体风险，并为风险归因、风格中性化与组合优化提供基础。

Barra 模型的求解是在截面上进行的，也就是已知因子暴露求因子收益率

因子暴露的定义：Barra 模型与 CAPM, Fama-French 等模型不同，它没有人为构造一个资产组合，**而是直接把标准化后的变量值作为因子暴露**。比如风格因子 S1 是市盈率，那么股票在市盈率上的暴露就是股票本身的市盈率的标准化。

注意到，国家因子与行业因子之间存在完全共线性，即 $1 = X_{I_1} + \ldots + X_{I_P}$，所以需要对矩阵 X 加一个约束，即：$f_{I_1}s_{I_1} + \ldots + f_{I_P}s_{I_P} = 0$，其中 s_{I_i} 代表行业 i 的流通市值。这个约束的意义是：行业因子收益率的市值加权为 0。将此式代入后可以消掉 f_{I_P}，解决完全共线性的问题

Barra 模型是一个典型的多因子模型。其创新点有四：

一是引入了国家因子，并解决了**国家因子与行业因子的多重共线性问题**。宏观经济因子全部被国家因子所代表，经济周期则被行业因子所代表

二是使用的是加权最小二乘，解决了回归中的异方差问题

三是直接使用标准化后的变量值作为因子暴露，而非人为构造资产组合，这种模型更加自由

四是标准化可以使得因子暴露的方差为 1，所以求出的因子收益率是可以比较的

一些比较重要的点：首先，风格因子中不应该有类似 CAPM 里面 $r_m - r_f$ 这种在截面上对于所有股票都相同的因子——**这些因子都被涵盖在国家因子中了**

其次，前二个模型均是当期与当期数据的回归，**Barra 模型中的 X 是上一期的，而 r 是当期的，也就是因变量与自变量之间隔了一期。**

Barra 模型不是用于预测股票收益率的！虽然 barra 模型中的 X 是上一期的，而 r 是当期的，看似可以用来预测股票收益率，但是求出的因子收益率向量 f 也是上一期的，预测股票收益率必须假设因子收益率向量 f 在时序上平稳，而这与实践不符

那么 Barra 模型是用来作什么的呢？是用来探究因子的性质的。在每个截面上都可以求得一个因子收益率 f，这样就有了每个因子的收益率序列，可以从中筛选出稳定的因子，也可以探究因子之间的相关性，进而探究股票之间的相关性，起到降维的作用

资料参考链接：https://blog.csdn.net/weixin_44607126/article/details/108275686

Barra CNE5 模型

在资料里的 Barra CNE5 模型中，行业因子共 32 个，采用 GICS（全球行业分类系统）作为分类标准。在 Barra CNE5 这个模型里，风格因子在按照定义计算出原始因子值后，需要进行标准化处理，需要注意的是，**μ是股票市值加权平均，而非简单平均；而标准差的计算仍使用等权标准差。**

同时，Barra 的各个因子还涉及了极值处理，在极值处理时，使用了缩尾 （Winsorize）的方法，即将均值两侧±n 倍标准差外的极值赋值为均值±n 倍标 准差，不同的风格因子 n 的取值可能为 2.5 或 3。这里需要注意的是，缩尾处理是，**无论是均值还是标准差都是使用等权的方法计算**，这与标准化是不同的。

最后，在 Barra 的模型中，**只对认为需要正交化的因子间做了正交化处理**， 比如非线性市值因子与市值因子，而非所有因子间都是正交的，因此部分因子间还有一定的线性相关性。

在计算 R^2 的时候,Barra 的 R^2 定义与传统定义不一样，**会偏大一些**。

Barra 模型的两大核心应用就是：

1. **风险归因（Risk Attribution）**：
   利用模型分解出的因子风险贡献和特质风险贡献，能够把组合的总风险拆解到各个因子（如规模、动量、估值、行业等）和个股特质上，量化每个因子对组合波动的边际影响，为风险监控和组合优化提供直观依据。
2. **因子中性化（Factor Neutralization）**：
   在构建或执行策略时，通过对策略持仓或信号向量相对于 Barra 因子暴露做回归、取残差的方式，将对指定因子的 β 敞口降至零，从而实现市场中性或风格中性，控制系统性风险，让策略收益更聚焦于选股 α 。

所谓的中性化，**就是让组合在Barra因子上的因子暴露为0，从而这些风险因子的收益也为0**。由于Barra因子中含有行业因子和风格因子，因此中性化也分为行业中性化和风格中性化。具体而言，Barra中性化方法为，使用组合的个股持仓signal对行业哑变 量以及Barra风格因子暴露（即因子暴露矩阵**X**）回归取残差。这种回归 取残差的方法可以保证在行业和风格因子上的暴露为0，若回归时还带有国家因 子（即常数项），还可以保证因子的和为0（即零额投资组合）。由于Barra中的风格因子不存在稳定的alpha收益，因此属于风险因子。如果一个组合在Barra风格因子上有明显暴露，那么在进行Barra风格中性化后，组合的波动率都会有一定程度的降低。

以下是一些核心数学，如重要矩阵运算，系统风险，以及特异性风险归因的算法：

数学作业纸

金儿博士 Jiner Boshi

科目

Barra 模型.

班级: 姓名: 编号: 第 页

$$\Omega = K\begin{array}{c} \\ \left[\begin{array}{cccc} & s_1 & s_2 \cdots & s_N \\ a_1 & \Omega_{11} & & \\ a_2 & \Omega_{21} & & \\ \vdots & \vdots & & \\ a_K & \Omega_{K1} & \cdots & \Omega_{KN} \end{array}\right] \\ N \end{array}$$

Ω_{ij} : i-th alpha, j stock weight.

for factor i:

$$\sum_{j=1}^{N} \Omega_{ij} Z_{ij} = 1 \qquad \sum_{j=1}^{N} \Omega_{ij} Z_{lj} = 0, \ \forall l \neq j$$

单因子i的所有加权暴露.

对任意j个非目标因子,暴露组合都为零.

$$Z = K\begin{array}{c} \\ \left[\begin{array}{cccc} & s_1 & s_2 \cdots & s_N \\ a_1 & Z_{11} & & \\ a_2 & Z_{21} & & \\ \vdots & \vdots & & \\ a_K & Z_{K1} & \cdots & Z_{KN} \end{array}\right] \\ N \end{array}$$

单一因子收益: $\boxed{f_i = \sum_{n=1}^{N} w_{in} \cdot r_n .}$

$\boxed{f = \Omega R. \quad R: N\times 1}$, $\Omega = S(S^T X^T V X S)^{-1} S^T X^T V$

纯因子组合

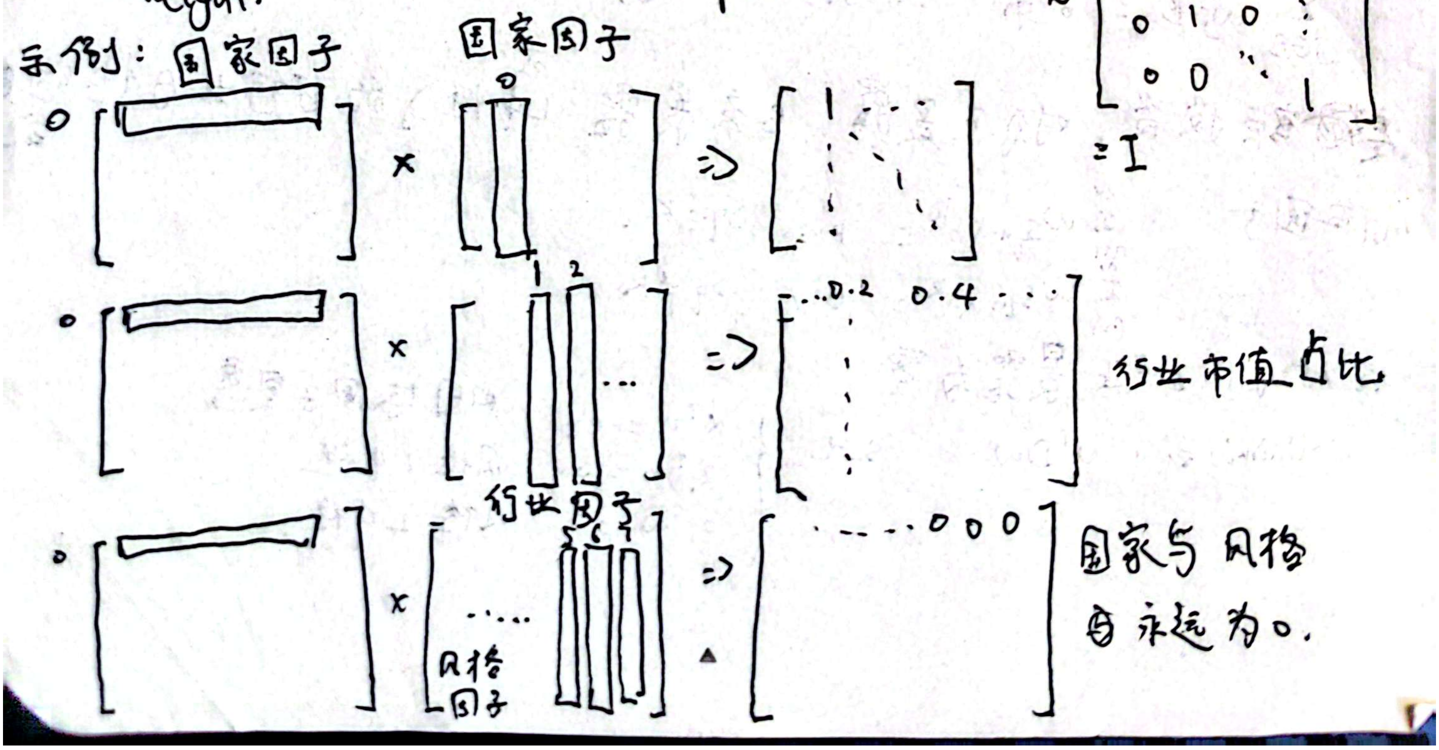
ΩX : $\Omega: K\times N$. $X: N\times K$. (Exposure). Exposure Matrix

$$K\left[\begin{array}{cccc} & s_1 & s_2 \cdots & s_N \\ a_1 & \Omega_{11} & \cdots & \Omega_{1N} \\ a_2 & \Omega_{21} & \cdots & \\ \vdots & \vdots & & \\ a_K & \Omega_{K1} & \cdots & \Omega_{KN} \end{array}\right]_{N} \times N\left[\begin{array}{cccc} & a_1 & a_2 \cdots & a_K \\ s_1 & X_{11} & X_{21} \cdots & X_{1K} \\ s_2 & X_{21} & & \\ \vdots & \vdots & & \\ s_N & X_{N1} & \cdots & X_{NK} \end{array}\right]_{K} \Rightarrow K\left[\begin{array}{cccc} & a_1 & a_2 \cdots & a_K \\ a_1 & \xi_{11} & \xi_{12} \cdots & \xi_{1K} \\ a_2 & \xi_{21} & & \\ \vdots & \vdots & & \\ a_K & \xi_{K1} & \cdots & \xi_{KK} \end{array}\right]$$

weight. Exposure Value

$$\approx \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \vdots \\ 0 & 0 & \ddots & \\ & & & 1 \end{array}\right]_{K} = I$$

示例: 国家因子



定义：暴露 = 组和权重向量与该因子在各只股票上的暴露值做内积。 $Exposure_A = w^T z(A)$. （行业，$\sum w \cdot 1 = \sum w$.）

行业纯因子组合：

1. 对本行业有正暴露.
2. 对国家因子（市场整体）中性（不受市场涨跌影响）.
3. 对其它所有因子中性.

做多行业A，做空国家纯因子组合，中和市场的整体影响.

例: $S = [\underbrace{0.4, 0.1}_{A}, \underbrace{0.3, 0.2}_{B}]$.

令 $w_1^A = w_2^A = 0.5$　　$\sum_{j\in A} w_j^A = 1$. ⇐ 行业A暴露.

$\sum_j s_j w_j^A = 0.4\times 0.5 + 0.1\times 0.5 = 0.25 > 0$.

=> 冲掉这个暴露，做空市场.

$\Rightarrow -0.25\times[0.4, 0.1, 0.3, 0.2] + [0.5, 0.5, 0, 0]$

$= [0.4, 0.475, -0.075, -0.05]$.

$\sum_j s_j w_j^A = 0.4\cdot 0.4 + 0.475\cdot 0.1 + 0.3\times(-0.075) + (0.2)\times(-0.05) = 0$

$\sum_{j\in A} w_j^A = 0.4 + 0.475 = 0.875 > 0$

$\sum_{j\in B} w_j^B = \cancel{0.5}\ -0.075 + (-0.05) = -0.08 < 0$.

零额投资，对A正暴露，B负暴露（或中性）的行业A纯组合.

风格因子：$\sum_{n=1}^{N} w_{s_q n} x_n^k = 1,\quad s_q = k$.

$\sum_{n=1}^{N} w_{s_q n} x_n^k = 0,\quad s_q \neq k$

在别的因子上暴露为零.

Minimize: $w^T D w$　s.t. $\begin{cases} X^T w = e_k & \text{对目标因子暴露.} \\ z_1^T w = 0 & \text{风格1中性} \\ z_2^T w = 0 & \text{风格2中性.} \\ \vdots \end{cases}$

数学作业纸 科目

金儿博士 Jinor Boshi

班级： 姓名： 编号： 第 页

风格因子处理

1. 标准化： $d_{nl} = \frac{d_{nl}^{raw} - \mu}{\sigma}$ ， μ：市值加权的平均
σ：等权标准差.

2. winsorize：将均值 $\pm n$ 倍标准差以外的极值赋值为均值 $\pm n$ 倍标准差.
$n = 3.5$ or 3.

3. 正交化：对需要的因子做正交化，如非线性市值，市值因子.
注意：非所有因子都正交化.

Barra 默认大市值公司对整体风险与风险暴露的贡献更大.

最终因子归因：

步骤： 1. 在 t 截面上，得到标准化暴露值 $X(t)$.

$$X(t) = N\begin{bmatrix} & q_1, q_2 \cdots q_K \\ S_1 & X_{11} & \vdots \\ S_2 & X_{21} & \vdots \\ \vdots \\ S_N & X_{K1} \cdots X_{NK} \end{bmatrix}_K$$

2. 计算 Ω 矩阵（权重）. $\Omega = S(S^T X^T V X S)^{-1} S^T X^T V$

$$\Omega = K\begin{bmatrix} & s_1 \cdots s_N \\ q_1 & w_{11} \\ \vdots \\ q_K & w_{K1} \cdots w_{KN} \end{bmatrix}_N, \quad V = \mathrm{diag}(v_1, \ldots, v_N), \quad v_n = \sqrt{\text{市值}_n}$$

3. 计算因子收益 $f(t)$ $f(t) = \Omega \cdot (R(t)) \in \mathbb{R}^K$.

$$f(t) = K\begin{bmatrix} f_1 \\ \vdots \\ f_K \end{bmatrix}_1 \qquad R(t) = N\begin{bmatrix} r_1 \\ \vdots \\ r_N \end{bmatrix}_1$$

4. 计算残差 $R(t) = X(t) f(t) + u(t)$.
$u(t) = R(t) - X(t) f(t)$.

$$u(t) = N\begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix}_1$$

5. 总体残差贡献.

$res_t = w(t)^T u(t)$, $w(t)$: 股票持仓权重.

$$w(t) = \begin{bmatrix} w_1 \\ \vdots \\ w_N \end{bmatrix}$$

6. 每日平均因子暴露.

在时间t，敞口记作　　$\beta(t) = w(t)^T X(t)$.

$w(t) \in \mathbb{R}^{N\times 1}$　　，　$X(t) \in \mathbb{R}^{N\times K}$

股票权重.　　　　暴露值矩阵.

即 $\beta(t) = [\underbrace{w(t)^T X_{\cdot,1}(t)}_{\beta_1(t)}, \underbrace{w(t)^T X_{\cdot,2}(t)}_{\beta_2(t)}, \ldots, \underbrace{w(t)^T X_{\cdot,K}(t)}_{\beta_K(t)}]$.

$$B = \begin{array}{c} \\ T \end{array} \begin{bmatrix} \beta_1(1), \beta_2(1) & \cdots & \beta_K(1) \\ \beta_1(2), & \cdots & \vdots \\ \vdots & & \vdots \\ \beta_1(T) & \cdots & \beta_K(T) \end{bmatrix}$$
$$K$$

把所有t连起来.

$\overline{\beta_K} = \frac{1}{T}\sum_{t=1}^{T} \beta_K(t)$. ← 每个因子每日平均暴露.

总结：Barra 风格/风险因子，用于拆分并解释股票收益来源于已知风格.

①对101个 alpha 因子信号做 Barra 中性化，确保 alpha 表现不靠系统性风格，而是真正选股能力.

②做2套回归：　1. Barra 风格回归，看多少收益来自已知风格.
2. Alpha 信号回归，看哪些 alpha 信号在剔除风格后仍有效，贡献了多少剩余收益.

对每个 alpha 信号做截面回归.

$s_j = X_{Barra}\,\beta + \varepsilon_j$，取残差 ε_j 作为系统性风格剔后信号

蒙卡 CNE5 特征值偏差校正.

Recall: $\Sigma_t = B_t F_t B_t^T + D_t$.

目标，校正 F_t. 矩阵.

1. 取 $\underbrace{\{f_{t-T+1}, f_{t-T+2}, \dots, f_t\}}_{\text{length } T}$, $f_\tau \in \mathbb{R}^K$.

2. Compute F_t. $\hat{F} = \frac{1}{T-1} \sum_{\tau=t-T+1}^{t} (f_\tau - \bar{f})(f_\tau - \bar{f})^T$, $\bar{f} = \frac{1}{T} \sum_\tau f_\tau$.

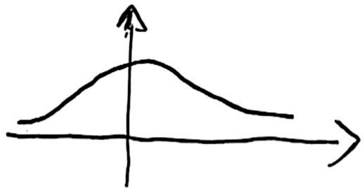
3. Eigen-Decomposition: $\hat{F} = Q \hat{\Lambda} Q^T$, $\hat{\Lambda} = \operatorname{diag}(\hat{\lambda}_1, \dots, \hat{\lambda}_K)$, $Q^T Q = I$

Q: E-vector, $\hat{\Lambda}$: E-value

4. Monte Carlo. $(M \geq 1000)$. 对每次模拟 j, $j = 1, \dots, M$

(1). 视 $\hat{F}$ 为真协方差，生成长度 T 的多元正态序列.
$\tilde{f}_\tau^{(j)} \sim N(0, \hat{F})$, $\tau = t-T+1, \dots, t$

(2). 重估样本协方差.
$\hat{F}^{(j)} = \frac{1}{T-1} \sum_\tau \tilde{f}_\tau^{(j)} \tilde{f}_\tau^{(j)T}$



(3) Eigen-Decomposition.
$\hat{F}^{(j)} = Q^{(j)} \Lambda^{(j)} Q^{(j)T}$, $\Lambda^{(j)} = \operatorname{diag}(\hat{\lambda}_1^{(j)}, \dots, \hat{\lambda}_K^{(j)})$

5. 计算期望与校正系数.
对于每个特征方向 $i = 1, \dots, K$.

(1). 模拟均值. $\mu_i = \frac{1}{M} \sum_{j=1}^{M} \hat{\lambda}_i^{(j)}$

(2). $c_i = \frac{\mu_i}{\hat{\lambda}_i}$ 校正系数. 注: $\hat{\lambda}_i$ 为原 eigen-value.

6. $F_{adj} = Q \operatorname{diag}(c_1 \hat{\lambda}_1, \dots, c_K \hat{\lambda}_K) Q^T = Q \operatorname{diag}(\mu_1, \dots, \mu_K) Q^T$

$\Rightarrow \boxed{\Sigma_t^{adj} = B_t F_t^{adj} B^T + D_t}$, 组合风险 $w^T \Sigma_t^{adj} w_t$.

解释： $F = Q\Lambda Q^T$， $\Lambda = diag(\lambda_1, \dots, \lambda_K)$

λ_i 的意义：如果把组合投到方向 q_i 上，它的波动率（风险）有多大。每个 λ_i 对应第 i 条正交风险方向，对应特征向量上 q_i 的方差。

* DEWIV 和 Garch 缩放.

DEWIV： $\boxed{\sigma^2_{r,t} = 21(1-\lambda)\sum_{s=1}^{\infty}\lambda^{s-1}(r_{t-s}-\bar{r})^2}$

r_{t-s}： $t-s-1$ 到 $t-s$ 收益.

$\bar{r}$：滚动均值或 0-样均值

λ：半衰期： $\lambda = 0.5^{\frac{1}{\text{Half-Life}}}$

Garch(1,1)： $r_t = E[r_t] + \varepsilon_t$

$E[r_t]$：在 t-1 时对 r_t 的条件预期（模型给出）.

ε_t：冲击（不可预见）. $E[\varepsilon_t] = 0$.

$$\boxed{\sigma^2_{r,t} = \omega + \alpha\varepsilon^2_{t-1} + \beta\sigma^2_{r,t-1} + \theta\varepsilon_{t-1}}$$

α：冲击效应： ε_{t-1} 大，$\uparrow \sigma^2_{r,t}$（很大涨，大跌）.

β：波动率惯性. $\alpha+\beta<1$， $\Rightarrow \sigma^2_{r,t}$ mean revert.

ω：长期波动平均. $\Rightarrow \frac{\omega}{1-\alpha-\beta}$

θ：非对称波动反应.

$\theta < 0$，坏消息（大跌）比好消息，波动率更大.

Recall： $\Sigma_t = \underset{\uparrow\ \text{系统性风险}}{X_t F_t X_t^T} + \underset{\uparrow\ \text{特异性风险}}{\Delta_t}$.

加市场组合权重： $\alpha = \frac{V_{target}}{V_{orig}}$

$\sigma^2 = \underset{\uparrow\ \text{原始系统风险}\ V_{orig}}{h_m^T X_t F_t X_t^T h_m} + \underset{\uparrow\ \text{特异风险}}{h_m^T \Delta_t h_m}$. s.t.

系统预测风险 $\Rightarrow V_{target} = \underset{\uparrow\ \text{DEWIV 总风险 (GARCH 预测)}}{\sigma_s^2} - \underset{\uparrow\ \text{特异风险}}{h_m^T \Delta_t h_m}$.

$\Rightarrow \boxed{\begin{aligned} &h_m^T X_t(\alpha F_t)X_t^T h_m \\ &= \alpha V_{orig} \\ &= \sigma_s^2 - h_m^T \Delta_t h_m. \end{aligned}}$

$\boxed{F_s = \alpha F_t}$， $\Sigma_t^{cal} = X_t F_s X_t^T + \Delta_t$.

Specific Risk Modeling.

$r_i = \sum_{k=1}^{K} x_{ik} f_k + u_i \Leftarrow$ idiosyncratic return.

一、原始数据目标： i：股票 k：因子 s：时间。

1. $\{u_{i,s}\}$ $\forall i = 1, \dots, N$, $\forall s = 1, \dots, t-1$

2. 基本面，因子暴露 $\{z_{i,k,s}\}$ $\forall i$, $k = 1, \dots K$, $s = 1, \dots, t-1$.

Goal: predict specific risk （月末 t-1 到月末 t）的特异风险.

二、平均绝对残差时序模型.

1. $a_{i,s} = |u_{i,s}|$ 绝对值.

$S_s = \frac{1}{N} \sum_{i=1}^{N} a_{i,s}$, $s = 1, \dots, t-1$

2. AR(p) $\boxed{S_s = \alpha + \sum_{j=1}^{p} \beta_j S_{s-j} + \eta_s}$ white noise. , $s = 1, \dots, t-1$

let $Y = \begin{pmatrix} S_{p+1} \\ S_{p+2} \\ \vdots \\ S_{t-1} \end{pmatrix}$ $X = \begin{pmatrix} 1 & S_p & \dots & S_1 \\ 1 & S_{p+1} & \dots & S_2 \\ \vdots & \vdots & & \vdots \\ 1 & S_{t-2} & \dots & S_{t-1-p} \end{pmatrix}$

solve $(X^T X)^{-1} X^T Y = \begin{pmatrix} \hat{\alpha} \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_p \end{pmatrix}$

Compute: $\boxed{\hat{S}_t = \hat{\alpha} + \sum_{j=1}^{p} \hat{\beta}_j S_{t-j}}$, for known S_{t-1}, $S_{t-2}, \dots, S_{t-p}$

三、相对绝对残差截面模型.

1. for $s \le t-1$, $v_{i,s} = \frac{a_{i,s}}{S_s} - 1$, $i = 1, \dots, N$

2. Regression. $\boxed{v_{i,s} = \sum_{k=1}^{K} \gamma_k z_{i,k,s} + \varepsilon_{i,s}}$, $\varepsilon_{i,s} \sim$ white noise

设 $s = t-1$，拟合： $v_{i,t-1} = \sum_{k=1}^{K} \gamma_k z_{i,k,t-1} + \varepsilon_{i,t-1}$.

则得到 $\hat{\gamma}_1^{(t-1)}, \dots, \hat{\gamma}_K^{(t-1)}$，计算： $\boxed{\hat{v}_{i,t} = \sum_{k=1}^{K} \hat{\gamma}_k^{(t-1)} z_{i,k,t-1}}$

四. 量纲系数 K。

1. $\sigma_{i,s}^{real} = Std(u_{i,s-L+1}, \ldots, u_{i,s})$, for $s=1,\ldots,t-1$

yield $\{\sigma_{i,1}^{real}, \ldots, \sigma_{i,t-1}^{real}\}$, L 为自定义滑动窗口.

2. 预测绝对残差: $\boxed{X_{i,s} = (1+v_{i,s}) S_s}$, for $s=1,\ldots,t-1$

yield: $\{X_{i,1}, \ldots, X_{i,t-1}\}$.

minimize: $\sum_{s\le t-1} \sum_{i=1}^{N} (K X_{i,s} - \sigma_{i,s}^{real})^2$ $\Rightarrow$ $\boxed{\hat{K} = \dfrac{\sum_{s\le t-1} \sum_{i=1}^{N} X_{i,s} \sigma_{i,s}^{real}}{\sum_{s\le t-1} \sum_{i=1}^{N} X_{i,s}^2}}$

五. 合成, $\boxed{\hat{\sigma}_{i,t} = \hat{K}(1+\hat{v}_{i,t})\hat{S}_t}$

作为估算的特异风险

总结. 结合:

$D_{t,ii} = \hat{\sigma}_{i,t}^2$, $i=1,\ldots,N$

$D_t = diag(\hat{\sigma}_{1,t}^2, \ldots, \sigma_{N,t}^2)$

假设不同股票特异性残差不相关.

Recall: $F_s = \alpha F_t$, $\Sigma_t^{cal} = X_t F_s X_t^T + \Delta_t$.

substitude D_t, $\Rightarrow$ $\boxed{\Sigma_t^{cal} = X_t F_s X_t^T + D_t.}$

应用: 如果投资组合权重: $\vec{w} = \langle w_1, \ldots, w_N \rangle$.

$Var(R_p)_t = w^T \Sigma_t w = \boxed{w^T (X_t F_s X_t^T) w + D_t}$

1. 因子协方差矩阵 F.

($f_{t,j}$).

$$
\begin{array}{c|cccc|c}
 & s_1 & s_2 & \cdots & s_N & \\
a_1 & \beta_{11} & \beta_{12} & \cdots & \beta_{1N} & f_1 \\
a_2 & \cdots & & & & f_2 \\
\vdots & & & & & \\
a_K & & & & & f_K
\end{array}
\quad K\times N
$$

$\Downarrow$

$$
\begin{bmatrix} f_{t1} \\ f_{t2} \\ f_{t3} \\ \vdots \\ f_{tK} \end{bmatrix}
$$

at t

expand $\Rightarrow$

$$
T\;
\begin{array}{c|cccc|}
 & a_1 & a_2 & \cdots & a_K \\
t_1 & f_{t_1 a_1} & f_{t_1 a_2} & \cdots & f_{t_1 a_K} \\
t_2 & f_{t_2 a_1} & \cdots & & \\
\vdots & \vdots & & & \vdots \\
t_T & f_{t_T a_1} & \cdots & & f_{t_T a_K}
\end{array}
= \{f_{t,j}\}.
$$

K

multiple period.

$$\boxed{r_{t,i} = \beta_i f_t + \varepsilon_{t,i}}$$

$$\boxed{\beta_i : \frac{x_i - \mu}{\sigma}}$$

factor exposure.

covariance: $F_{jl} = \frac{1}{T-1} \sum_{t=1}^{T} (f_{t,j} - \bar{f}_j)(f_{t,l} - \bar{f}_l)$.

$$
K\;
\begin{array}{c|cccc|}
 & a_1 & a_2 & \cdots & a_K \\
a_1 & F_{11} & F_{12} & \cdots & F_{1K} \\
a_2 & F_{21} & F_{22} & & F_{2K} \\
\vdots & & & \ddots & \vdots \\
a_K & F_{K1} & F_{K2} & \cdots & F_{KK}
\end{array}
= F_{(K\times K)}
$$

K

3. 因子暴露矩阵. at t.

$$
\underset{(N\times K)}{B_t} = N\;
\begin{array}{c|cccc|}
 & a_1 & a_2 & \cdots & a_K \\
s_1 & \beta_{11} & \beta_{12} & \cdots & \beta_{1K} \\
s_2 & \beta_{21} & \beta_{22} & \cdots & \beta_{2K} \\
\vdots & & & & \\
s_N & \beta_{N1} & \beta_{N2} & \cdots & \beta_{NK}
\end{array}
$$

K

$$\boxed{\Sigma_t = B_t F B_t^T + D_t}$$

可选滚动动态 F_t.

2. 特异性方差矩阵 D ;在t这天.

$r_{t,i} = \sum_{j=1}^{K} \beta_{ij} f_{t,j} + \varepsilon_{t,i}$

i: stock

j: factor.

$\Rightarrow [\varepsilon_{t,1}, \varepsilon_{t,2}, \ldots, \varepsilon_{t,N}]$. at t

S

$$
T\;
\begin{array}{c|cccc|}
 & S_1 & S_2 & \cdots & S_N \\
t_1 & \varepsilon_{t_1,1} & \varepsilon_{t_1,2} & \cdots & \varepsilon_{t_1,N} \\
t_2 & \varepsilon_{t_2,1} & \varepsilon_{t_2,2} & \cdots & \\
\vdots & \vdots & \vdots & & \\
t_T & \varepsilon_{t_T,1} & \cdots & & \varepsilon_{t_T,N}
\end{array}
$$

N

$$\boxed{\sigma^2_{\varepsilon_i} = \frac{1}{\underbrace{T-K-1}_{T.df}} \sum_{t=1}^{T} (\varepsilon_{t,i} - \bar{\varepsilon}_i)^2}$$

$$
\Rightarrow N\;
\begin{array}{c|cccc|}
 & S_1 & S_2 & \cdots & S_N \\
S_1 & \sigma^2_{\varepsilon_1} & 0 & & \\
S_2 & 0 & \sigma^2_{\varepsilon_2} & & \\
\vdots & & & \ddots & \\
S_N & 0 & & & \sigma^2_{\varepsilon_N}
\end{array}
$$

N

$= D_t\ (N\times N)$

Barra 模型常见用法： $1\times N$

1. 组合风险：

$w=(w_1, w_2, \dots, w_N)^T$

全市场，无权重则为0，归一化

$$Var_t = w_t^T \Sigma_t w_t$$

$$\sigma_t = \sqrt{w_t^T \Sigma_t w_t}$$

$$\sigma_t^{ann} = \sigma_t \times \sqrt{\text{total trading days}}$$

年化化组合波动，

$$\Sigma_t = \begin{array}{c c} & \begin{array}{cccc} S_1 & S_2 & \cdots & S_N \end{array} \\ \begin{array}{c} S_1 \\ S_2 \\ \vdots \\ S_N \end{array} & \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \\ \Sigma_{21} & \Sigma_{22} & \cdots & \\ & & \ddots & \\ & & & \Sigma_{NN} \end{bmatrix} \end{array}$$

$N \times N$

2. 单股票风险归因

* 边际风险贡献

$$MRC_{t,i} = \cancel{\frac{\sum \Sigma_{t,ij} w_{t,j}}{\sqrt{w_t^T \Sigma_t w_t}}} \boxed{\frac{\sum_{j=1}^{N} \Sigma_{t,ij}\, w_{t,j}}{\sqrt{w_t^T \Sigma_t w_t}}}$$

$$= \frac{[\Sigma_t w_t]_i}{\sigma_t} = \boxed{\frac{\partial \sigma_t}{\partial w_{t,i}}}$$

波动随权重（个股）的变化率..

* $RC_{t,i} = \boxed{w_{t,i}\, MRC} = \boxed{w_{t,i} \frac{\partial \sigma_t}{\partial w_{t,i}}}$, $\boxed{\sum_{i=1}^{N} RC_{t,i} = \sigma_t}$

绝对风险贡献：股票i对组合波动的绝对贡献

3. 因子风险归因

组合因子暴露：$\beta_{port\,t,j} = \boxed{\sum_{i=1}^{N} \beta_{t,ij}\, w_{t,i}}$ （因子j）

$$\boxed{\beta_{port} = B_t^T w_t}$$

* 因子风险贡献：

$$RC^{fac}_{t,j} = \boxed{\beta_{port\,t,j}\,[F_t \beta_{port}]_j} = \cancel{\beta_{port\,t,j} \sum_{l=1}^{K} F_{t,ij}\, \beta_{port,l}}$$

$$\beta_{port\,t,j} \sum_{l=1}^{k} F_{t,jl}\, \beta_{port,t,l}$$

$$\Rightarrow RC^{fac}_{t,j} = \boxed{\sum_{i=1}^{N} B_{t,ij}\, w_{t,i} \times \left(\sum_{l=1}^{K} F_{t,jl} \sum_{k=1}^{N} B_{t,kl} \cdot w_{t,k}\right)}$$

，且 $\sum_{j=1}^{K} RC^{fac}_{t,j} + \sum_{i=1}^{N} w_{t,i}^2 D_{t,ii} = \sigma_t^2$

4. 风险限额与优化

* 最小方差组合：$\boxed{\min_{w_t} w_t^T \Sigma_t w_t}$ s.t. $\sum_{i=1}^{N} w_{t,i} = 1$, $w_{t,i} \ge 0$.

* 风险平价：目标：$\boxed{w_{t,i}\,[\Sigma_t w_t]_i = w_{t,j}\,[\Sigma_t w_t]_j}$ $\forall i,j$.

$$w_{t,i} \sum_{k=1}^{N} \Sigma_{ik}\, w_{t,k} = w_{t,j} \sum_{k=1}^{N} \Sigma_{t,jk}\, w_{t,k}.$$

s.t. $\sum_{i=1}^{N} w_{t,i} = 1$. $w_{t,i} \ge 0$.

* 因子敞口约束：$\boxed{-c_j \le \beta_{port,t,j} \le c_j}$, for $j = 1, \dots, K$.

5、业绩归因：

note $\beta_{port} = B_t^T w_t$

$$\boxed{R_{t,port} = \beta_{port}^T f_t + \varepsilon_{t,port}}$$

组合当周实际收益　　因子贡献　　特异性收益．

注意、$R_{t,port} = \sum_{i=1}^{N} w_{t,i} r_{t,i}$ ，$\beta_{port}^T f_t = \sum_{j=1}^{K} \beta_{port,t,j} f_{t,j}$

组合收益（加权）　　因子暴露 × 贡献

$\varepsilon_{t,port}$ 为剩下的残差收益．

$\varepsilon_{t,port}$ 为正，⇒ 特异性收益．

为负，⇒ alpha 失效．