

投资组合最优化

最近阅读了[《如何分配资金？组合优化的是是非非》](#)，手痒痒想做一下实证分析，但这篇文章并未涉及到技术性的细节，作为补充，写下这篇文章记录实证过程中遇到的技术性的问题。

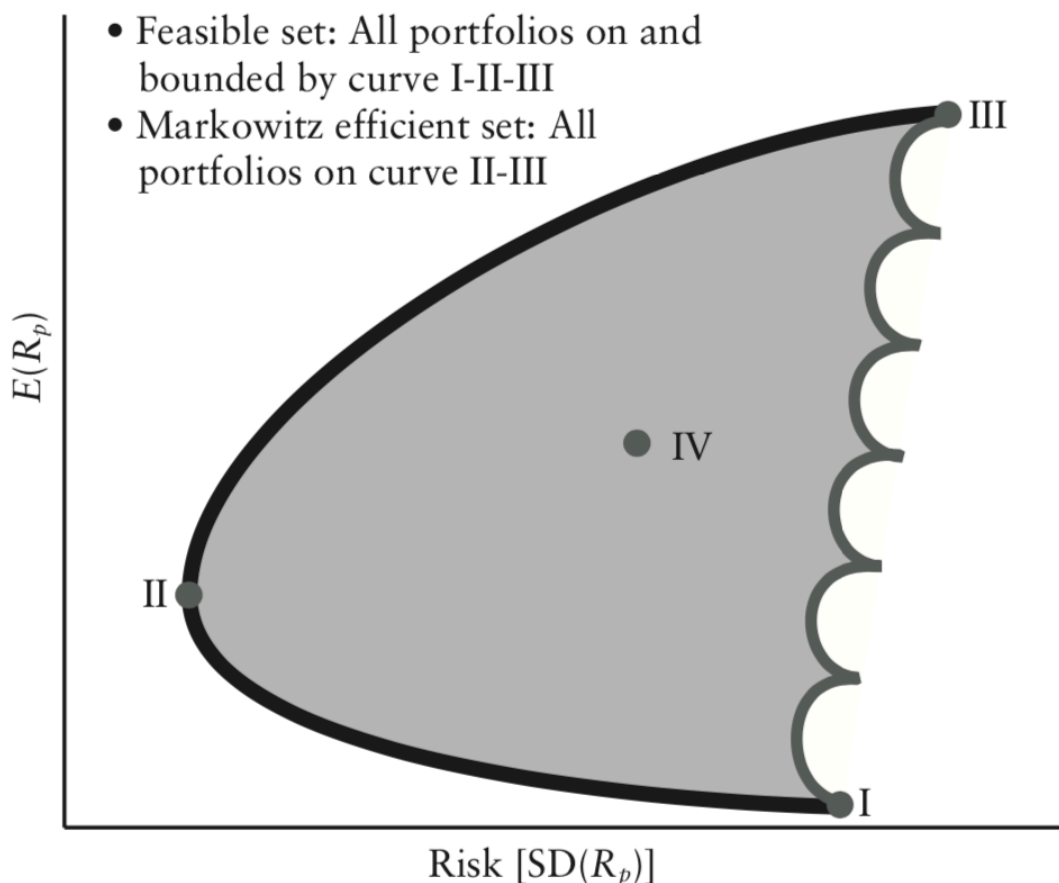
首先还是简单介绍一下啦，表达一下对大神的敬意。投资组合最优化是现代投资组合管理中一个极其重要的工具。Harry Markowitz 1952年的开创性的论文*Portfolio Selection*奠定了均值方差分析(mean-variance analysis), 均值方差最优化(mean-variance optimization), and 现代投资组合理论(Modern Portfolio Theory, MPT)的理论基础。惊叹一下，大神时年25岁！

作为基础，下面介绍一下均值方差分析。

1. 均值方差分析

Markowitz认为，投资者应该在风险和期望收益的权衡基础上做出选择。期望收益为所考察的时期内资产的预期收益率，风险则用资产的方差来衡量。需要指出的是，Markowitz的均值-方差模型并为假定资产的收益率服从联合正态分布。事实上，从以下两个不同的出发点出发所得到的结果是相同的：(1)在一定的假设下，期望效用最大化；(2)资产收益服从联合正态分布。

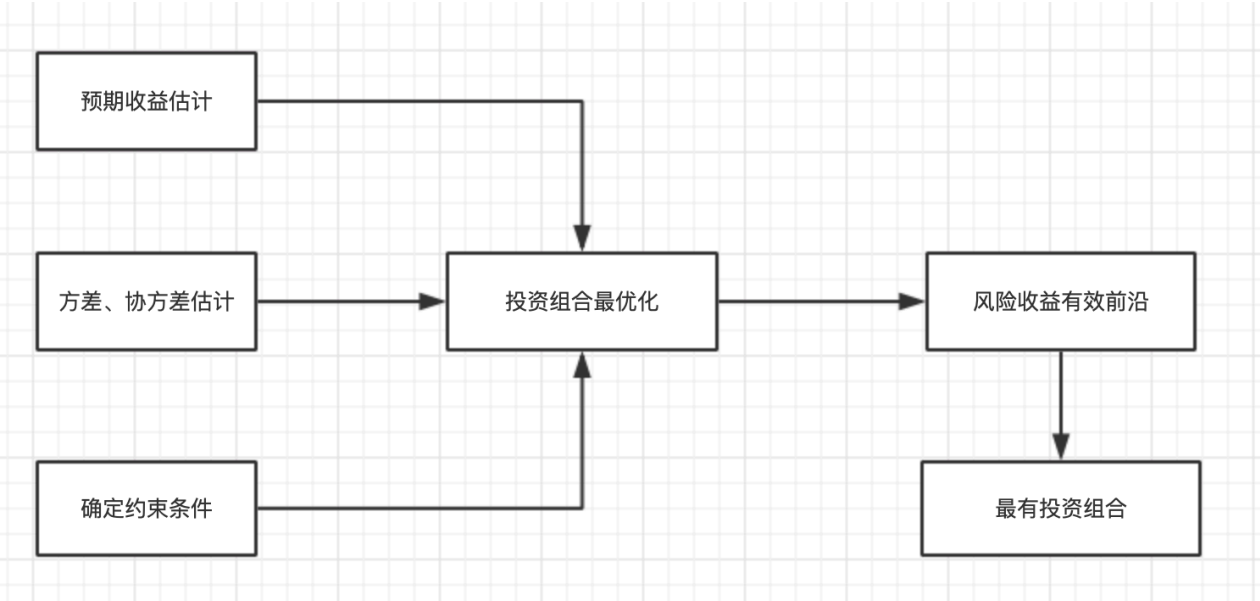
Markowitz主张在期望收益给定的前提下，一个理性的投资者应当选择所有可行的投资组合中方差最小的组合。如下图所示，阴影部分为可行集(feasible set)，可行集边界I-II-III为有效前沿(efficient frontier)，II-III为有效集(Markowitz efficient set)，II为全局最小方差组合(GMVP)。



(图片来自 *Quantitative Equity Investing : Techniques and Strategies*)

2. 均值方差最优化

下图从现代投资组合理论角度给出了投资过程，这个过程也被称为均值-方差最优化或者投资组合的选择理论。



下面从数学角度说明该模型。

首先介绍一下符号标记：

符号	含义
N	组合内资产数量
w	权重向量($N \times 1$)
w_i	资产i的权重
μ	预期收益率($N \times 1$)
σ	资产波动率($N \times 1$)
σ_i	资产i的波动率
σ_{ij}	资产i 和资产j的协方差
Σ	方差协方差矩阵 $\Sigma(N \times N)$
σ_p	组合波动率
μ_p	组合预期收益
λ	风险厌恶系数
r_f	无风险利率

2.1 基本模型

均值-方差优化问题的目标函数为

$$\begin{aligned} \min_w \quad & \frac{1}{2} w' \Sigma w \\ \text{s.t.} \quad & w' \mu = \mu_0 \\ & \sum w_i = 1 \end{aligned}$$

上面方程表达的含义是：对于给定的期望收益，最小化组合的波动，因此称之为风险最小化模型。该问题具有解析解，且有效投资组合在 (μ_0, σ_0) 坐标系中为一条双曲线，即上文提到的有效边界。另外，对于不等式约束，比如 $w' \mu \leq \mu_0$ 上述优化问题则需要使用数值方法求解。

另外，均值-方差优化问题还有其他一些等价的形式。

比如期望收益最大化模型：

$$\begin{aligned} \max_w \quad & w' \mu \\ \text{s.t.} \quad & w' \Sigma w = \sigma_0^2 \\ & \sum w_i = 1 \end{aligned}$$

风险厌恶模型：

$$\begin{aligned} \max_w \quad & w' \mu - \lambda \frac{1}{2} w' \Sigma w \\ \text{s.t.} \quad & \sum w_i = 1 \end{aligned}$$

其中， λ 为风险厌恶系数，作为对风险项的惩罚。风险厌恶系数因人而异，因资产类别而已，在实际应用中较难确定。

2.2 考虑无风险资产

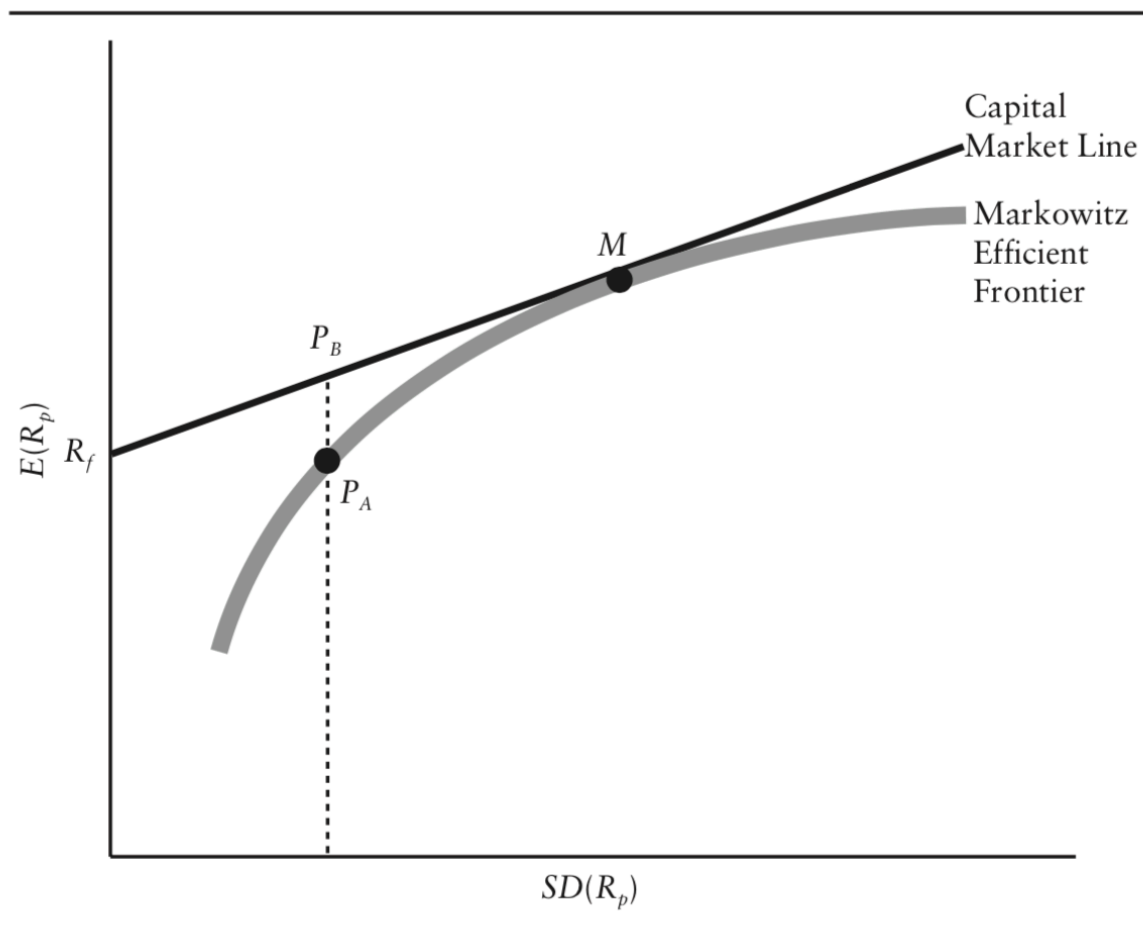
上面几种模型未考虑无风险资产，下面说明一下考虑无风险资产之后的优化问题。

可以证明，考虑无风险资产之后，采用均值方差优化框架求解出来的有效集要优于不考虑无风险资产的结果（William Sharpe（1964），James Tobin（1958），and John Lintner（1965））。

风险资产组合的权重向量为 $w_R = (w_{R1}, w_{R2}, \dots, w_{RN})$ ，权重加总和可以不为1，且如果允许无风险借贷，则权重和可以为正也可以为负。无风险资产的权重为 $1 - \sum w_{Ri}$ 。优化问题为：

$$\begin{aligned} \min_{w_R} \quad & w_R' \Sigma w_R \\ \text{s.t.} \quad & w_R' \mu + (1 - \sum w_{Ri}) R_f = \mu_0 \end{aligned}$$

而该问题的解位于下图CML线上。



从上图可以看到，无论投资者的风险偏好如何，所选择的资产组合都为切点组合M。由于风险偏好的不同造成的选择不同体现在资金在风险资产和无风险资产的配置比例上。这便是分离定理。这一性质在实际中也有很重要的意义，具体地说，投资过程可以分解成两步：风险组合的构建以及资金在风险、无风险资产上的分配。

2.3 其他衍生模型

Strategy	Objective	Assumptions for Mean-Variance Optimality
Minimum Variance Portfolio	$\text{Min } \sigma_p^2 = w' \Omega w$	Expected returns are equal
Maximum Decorrelation Portfolio	$\text{Min } \sigma_p^2 = w' \Omega w$	Expected returns and volatilities are equal
Maximum Diversification Portfolio	$\text{Max } D_p = \frac{w' \sigma}{\sqrt{w' \Omega w}}$	Expected returns are proportional to volatilities
Risk-Parity Portfolio/Equal Risk Contribution	$\frac{\partial \sigma_p}{\partial w_i} w_i = \frac{\partial \sigma_p}{\partial w_j} w_j \text{ for any } i, j$	Expected returns are proportional to volatilities; correlations must be constant
Maximum Deconcentration	Maximizes $1/w'w$	Expected returns and volatilities are equal; correlations are constant
Maximum Sharpe Ratio	$\text{Max } SR_p = \frac{w' r}{\sqrt{w' \Omega w}}^*$	No assumptions needed; however, expected returns must be specified, either as an additional input or assumed to be proportional to risk

关于这些优化策略的介绍，详见一开始提到的那篇文章。（图片中的Q为方差-协方差矩阵）

3. 现实中的问题

现实中，问题不可能像上述模型如此简化，目标函数、约束条件均可能存在变化，模型中输入参数的估计、模型的求解也可能存在困难。下面具体说说存在的一些问题，当然，这些必然不全面。

3.1 约束条件和优化目标

制度特征和投资决策往往会导致比均值-方差问题的原始模型所具有的更加复杂的约束条件和投资组合管理目标。下面介绍一些约束条件和优化目标。

- 做空限制: $w \geq 0$;
- 换手率约束: $|w - w_0| \leq U$ (单个资产约束), or $\sum |w_i - w_{0i}| \leq U_p$ (整个投资组合约束);
- 仓位约束: 单个资产的约束 $L_i \leq w_i \leq U_i$, 资产类别的约束, 比如行业仓位的约束, $L_j \leq \sum w_i \leq U_j$ (属于行业 j 的权重和的限制);
- 风险因子约束: 对于某因子的暴露权重的约束 $\sum_{i=1}^N \beta_{ik} w_i \leq U_k$, 其中 β 为对因子 k 的暴露; 比如, 要求市值中性时, 上述条件可写成 $\sum_{i=1}^N \beta_i w_i = 0$, β 为对市值因子的暴露;
- 基准暴露: $\|w - w_b\| \leq M$; 跟踪误差: $(w - w_b)' \Sigma (w - w_b) \leq \sigma_{TE}^2$; 需要注意, 这两种约束仅为相对基准 (某指数) 的风险, 而未考虑组合的绝对风险。
- 最小交易规模约束条件: 经典的均值方差最优化问题经常产生一些大的与许多小的头寸。实践中, 由于存在交易成本和其他票据收费, 因此应尽量避免小额的资产交易。实践中, 常用的处理办法是: 先不考虑该约束, 当求解出个股票的权重之后, 去掉小于某阈值的股票, 剩下的股票再按其最初权重比例再分配新的权重。

3.2 期望收益和风险的估计问题

统计估计是存在误差的, 并且依赖于数据的质量和所使用的统计技术。一般地, 期望收益和风险的好的估计量, 需要有如下性质:

- 能提供一定的预测能力而不仅仅是对过去表现的事后性的历史总结;
- 计算成本合理;
- 所使用的技术不会放大该技术中所使用的数据中的误差;
- 预测应当符合直觉。

众所周知, 期望收益具有显著的时序变化性 (非平稳), 因而使用历史数据计算的历史收益对未来收益的预期能力很弱, 即 "过去可以预测未来" 的假设并不一定成立, 事实上, 很多情况下并不成立。

因此, 基于历史数据计算的样本均值和样本方差协方差矩阵往往存在很多问题。在资产收益服从独立同分布的假设前提下, 样本方差协方差矩阵是总体协方差矩阵的 MLE 估计量, 且服从自由度为 $N-1$ 的 Wishart 分布。可以看到, 样本方差协方差的性质还算比较不错的。对于期望收益而言, 样本均值则是期望收益的一个很差的估计量, 原因在于收益序列往往具有后尾分布, 这导致样本均值不再是总体期望的最佳线性无偏估计量 (BLUE)。另外, 金融时间序列的不平稳特性, 使得样本均值也不再是一个好的预测值。

对于均值方差优化问题, 模型对于输入非常敏感, 期望收益的小小的改变往往导致组合权重的大幅变化。同样, 方差协方差矩阵的估计误差也会造成这样的结果 (程度要弱于期望收益)。因此, 我们需要采用更好的估计量作为模型的输入。

期望收益估计的解决途径包括: 对估计量施加更多的结构化的限制 (比如因子模型), 贝叶斯方法 (Black-Littleman 模型) 或者压缩估计。后面会详细介绍。

方差协方差估计的解决途径包括: 衰减权重加权 (近期数据权重更大)、结构化模型、压缩估计、估计量组合等等。

值得一提的是，针对金融时间序列常出现的自相关和异方差的问题，常用的解决方案是进行"Newey-West修正"。具体方法参考《[协方差矩阵的 Newey-West 调整](#)》。

3.3 缺失和截短数据（Missing and Truncated Data）的处理

实践中，可能存在数据缺失、数据错误、数据长度不一致这样的问题，如果不考虑这些数据质量问题，模型的估计结果的准确性将得不到保证。下面说上述各种问题常见的解决方法。

缺失值问题可以采用所谓的**期望最大化算法(EM)**来解决，参考See Roderick J. A. Little and Donald B. Rubin, *Statistical Analysis with Missing Data*; and Joe L. Schafer, *Analysis of Incomplete Multivariate Data*。

数据长度不一致的问题，比如有些公司上市10年，有些则上市3年，我们可以考虑截短数据的方法，使用3年的数据计算方差协方差矩阵；另外，Stambaugh提出的方法比上述截短方法更好，该方法从截短数据计算的方差协方差矩阵出发，然后利用额外的数据对其进行调整，参考*Analyzing Investments Whose Histories Differ in Length*。

3.4 数据频率

Merton(*On Estimating the Expected Return on the Market: An Exploratory Investigation*)指出，对于预期收益而言，技术预期收益为常量，为提高估计量的准确性，还是需要长期的历史数据；而对于方差协方差矩阵而言，在一定的条件下，可以增加数据的频率来提高估计量的准确性。

利用日内高频数据或者日度数据可以获得一个改善的波动率估计量，参考*On the Estimation of Security Price Volatilities from Historical Data*。

3.5 风险的度量方式

上面的叙述中，均采用方差来衡量风险，此外还存在其他风险的测度方式，主要有离差法和下行风险度量。不再详细介绍。

离差法同等程度的看待上行和下行变化，视其均为风险，事实上，离差法度量的是收益的不确定性，上行的可不叫风险。注意，方差只是离差法的二阶形式。

The mean-absolute moment (MAM_q) of order q is defined by

$$MAM_q(R_p) = \left(E \left(|R_p - E(R_p)|^q \right) \right)^{1/q}, q \geq 1$$

and is a straightforward generalization of the mean-standard deviation ($q = 2$) and the mean absolute deviation ($q = 1$) approaches.

下行风险度量包括Roy's Safety-First、半方差、下偏矩、在险价值、条件在险价值等。需要说明的是，即使这些方法在理论上很有吸引力，但在实际应用中，下跌风险的计算方法更加复杂，单个资产的下跌风险不能很容易的加总到资产组合的下跌风险度量当中，因为它要求整个证券收益的联合分布的信息。因此，常采用非参数估计、模拟和最优化技术来解决这类问题。

另外，下跌风险度量的估计风险通常比标准均值-方差方法更大，原因在于我们只使用一部分数据来计算下跌风险（也许甚至只是经验分布的尾部）。

4. Bayes技术和Black-Littleman模型

上文中提到，诸如采用样本均值和协方差矩阵这样的方法构建投资策略，一般在实际中表现均很差。这样的现象并非是组合优化技术的失效，更可能是投资组合优化框架对输入参数十分敏感，不准确的参数估计使模型的结果表现糟糕。

4.1 常见问题

- 对估计误差的敏感性；
- 优化过程中输入不确定性的影响；
- 准确估计投资组合优化框架的输入参数所必需的大量数据要求。

4.1.1 对估计误差的敏感性

在资产组合优化问题中，具有较高预期收益和较低标准差的证券往往会被过度的提高权重。所以，对预期收益和方差协方差矩阵的估计误差会导致最优权重出现大的误差。相对而言，预期收益的估计误差较之协方差矩阵的估计误差对最优权重的影响更大。相对重要性取决于风险厌恶程度，但根据一般的经验法则，预期收益率估计误差的影响大约是方差协方差矩阵影响的10倍，且方差的估计误差的影响大约是协方差影响的2倍。随着风险承受能力的提升，预期收益率的估计误差的相对影响越来越重要。

根据这个简单的规则，我们的研究重点应该放在预期收益率的估计上，其次才是方差协方差矩阵的估计。

约束组合权重

一些研究表明，在均值-方差框架中加入约束，可以带来更好的样本外表现。实际中，常利用不可卖空约束或规定持有每种证券的上下限以避免权重过分集中于某些资产。另外，约束权重有助于抑制波动性。例如，无卖空限制（下限约束）与降低方差协方差效果相同，上限约束与提高方差协方差效果相同。原因在于，组合优化倾向于给方差协方差大的资产较小的权重，因此，大方差的资产往往被赋予负的权重，限制卖空就意味着选择方差协方差较小的资产。

但要注意的是，强加约束时需要特别小心，如果约束条件太严格，他们将完全决定组合配置一而不是预测了。

敏感性分析的重要性

在实际中，为了最小化由于估计误差导致的巨大变化，进行敏感性分析是必要的。例如，对输入参数进行一个较小的改变（小的冲击），观察其对最优化结果的影响。这一冲击可能需要分别对于每一种资产均进行测试，找出最敏感的证券。

高度相关资产的问题

包含高度相关资产是造成均值-方差框架不稳定性另一个主要原因。当相关矩阵采用历史数据估计得到时，高度相关问题往往会恶化。特别的，当相关性矩阵利用稍微不同时期的数据重新估计时，相关性可能因此发生改变。这是，利用收缩估计量或者采用因子模型可能更好。

4.1.2 输入中融入不确定性

由于预期收益率和方差协方差矩阵是估计值，存在一定的估计误差。若仅将其作为确定值代入均值方差最优化框架中，则会存在一定的风险。因此，将估计值的不确定性纳入到优化过程中，更符合实际。

一般情况下，我们使用的仅是预期收益和方差协方差的点估计，这不符合谨慎投资者的行为方式。投资者可能更倾向于选择在不同情景下都具有良好表现的投资组合，这也在一定程度上规避了估计风险和模型风险。投资者为规避这些风险，自愿放弃在某些情景下能获得的收益，寻求稳健的投资组合，也就是能保证在最差情景下的表现的投资组合。相当于考虑参数的区间估计或者分布估计。可以通过常用的稳健估计技术Bayes估计量和压缩估计来解决（这里不再具体说明，参考《数量化股票投资：技术与策略》第九章）

4.1.3 大量的数据要求

经典均值方差优化框架中，需要为所考虑的投资领域的所有证券提供预期收益和协方差的估计。然而现实中，我们通常仅能对部分证券的收益做出可靠的预测，毕竟得到所有参数的良好的估计不太现实。

4.2 收缩估计

Stein(1956)发现，有偏估计通常会得到比无偏估计更好的估计量。例如James-Stein收缩估计量通常是比样本均值更好的估计量，因为其二次损失更小。

For example, the so-called *James-Stein shrinkage estimator*

$$\hat{\mu}_{JS} = (1 - w)\hat{\mu} + w\mu_0 \mathbf{1}$$

has a lower quadratic loss than the sample mean, where

$$w = \min\left(1, \frac{N - 2}{T(\hat{\mu} - \mu_0 \mathbf{1})' \Sigma^{-1} (\hat{\mu} - \mu_0 \mathbf{1})}\right)$$

实际上，收缩是对不同的估计量进行平均的一种形式。收缩估计量一般包括三个部分：（1）无结构或者具有简单结构的估计量，如样本均值；（2）具有复杂结构的估计量，如收缩目标；（3）收缩强度。收缩目标的选择有以下两种要求：（1）应具有少量的参数；（2）具有同被估计的未知量相同的一些基本特征。收缩强度可以根据理论性质或数值模拟选取。

金融文献中最著名的收缩估计量也许就是Jorion提出的：

Probably the most well-known shrinkage estimator¹⁷ used to estimate expected returns in the financial literature is the one proposed by Jorion,¹⁸ where the shrinkage target is given by $\mu_g \mathbf{1}$ with

$$\mu_g = \frac{\mathbf{1}' \Sigma^{-1} \hat{\mu}}{\mathbf{1}' \Sigma^{-1} \mathbf{1}}$$

and

$$w = \frac{N + 2}{N + 2 + T(\hat{\mu} - \mu_g \mathbf{1})' \Sigma^{-1} (\hat{\mu} - \mu_g \mathbf{1})}$$

We note that μ_g is the return on the GMV portfolio discussed in Chapter 8. Several studies document that for the mean-variance framework: (1) the

我们也可以对协方差矩阵的估计采用压缩技术，这涉及到将一个非结构化的协方差矩阵估计量缩减成一个更具结构化的协方差矩阵估计量。例如利用单指数模型计算方差协方差矩阵或者常相关系数模型等。另外常用的也可以使用多因素模型来进行结构化。

4.3 Black-Littleman模型

BL模型的基本特征是将市场均衡同投资者的看法结合起来。

BL模型的使用过程可以简单的分成三步：

- 基本假设和出发点：假设之一便是证券的预期收益率应该和市场均衡保持一致，即在没有额外的观点的情况下，投资者应持有市场组合。
- 表达投资者看法：

Formally, K views in the Black-Litterman model are expressed as a K -dimensional vector \mathbf{q} with

$$\mathbf{q} = \mathbf{P}\boldsymbol{\mu} + \boldsymbol{\epsilon}_q, \boldsymbol{\epsilon}_q \sim N(0, \boldsymbol{\Omega})$$

where \mathbf{P} is a $K \times N$ matrix (explained in the following example) and $\boldsymbol{\Omega}$ is a $K \times K$ matrix expressing the confidence in the views. In order to understand this mathematical specification better, let us take a look at an example.

- 投资者看法同市场均衡相结合：

Let us first recall the specification of market equilibrium

$$\boldsymbol{\Pi} = \boldsymbol{\mu} + \boldsymbol{\epsilon}_\Pi, \boldsymbol{\epsilon}_\Pi \sim N(0, \tau\boldsymbol{\Sigma})$$

and the one for the investor's views

$$\mathbf{q} = \mathbf{P}\boldsymbol{\mu} + \boldsymbol{\epsilon}_q, \boldsymbol{\epsilon}_q \sim N(0, \boldsymbol{\Omega})$$

We can stack these two equations together in the form

$$\mathbf{y} = \mathbf{X}\boldsymbol{\mu} + \boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} \sim N(0, \mathbf{V})$$

where

$$\mathbf{y} = \begin{bmatrix} \boldsymbol{\Pi} \\ \mathbf{q} \end{bmatrix}, \mathbf{X} = \begin{bmatrix} \mathbf{I} \\ \mathbf{P} \end{bmatrix}, \mathbf{V} = \begin{bmatrix} \boldsymbol{\tau}\boldsymbol{\Sigma} & \\ & \boldsymbol{\Omega} \end{bmatrix}$$

with \mathbf{I} denoting the $N \times N$ identity matrix. We observe that this is just a standard linear model for the expected returns $\boldsymbol{\mu}$. Calculating the Generalized Least Squares (GLS) estimator for $\boldsymbol{\mu}$, we obtain

$$\begin{aligned} \hat{\boldsymbol{\mu}}_{BL} &= (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \\ &= \left(\begin{bmatrix} \mathbf{I} & \mathbf{P}' \end{bmatrix} \begin{bmatrix} (\boldsymbol{\tau}\boldsymbol{\Sigma})^{-1} & \\ & \boldsymbol{\Omega}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I} \\ \boldsymbol{\Pi} \end{bmatrix} \right)^{-1} \begin{bmatrix} \mathbf{I} & \mathbf{P}' \end{bmatrix} \begin{bmatrix} (\boldsymbol{\tau}\boldsymbol{\Sigma})^{-1} & \\ & \boldsymbol{\Omega}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Pi} \\ \mathbf{q} \end{bmatrix} \\ &= \left(\begin{bmatrix} \mathbf{I} & \mathbf{P}' \end{bmatrix} \begin{bmatrix} (\boldsymbol{\tau}\boldsymbol{\Sigma})^{-1} \\ \boldsymbol{\Omega}^{-1}\mathbf{P} \end{bmatrix} \right)^{-1} \begin{bmatrix} \mathbf{I} & \mathbf{P}' \end{bmatrix} \begin{bmatrix} (\boldsymbol{\tau}\boldsymbol{\Sigma})^{-1}\boldsymbol{\Pi} \\ \boldsymbol{\Omega}^{-1}\mathbf{q} \end{bmatrix} \\ &= [(\boldsymbol{\tau}\boldsymbol{\Sigma})^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}]^{-1} [(\boldsymbol{\tau}\boldsymbol{\Sigma})^{-1}\boldsymbol{\Pi} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{q}] \end{aligned}$$

The last line in the above formula is the Black-Litterman expected returns that *blend* the market equilibrium with the investor's views.

Note: 我们可以将多因子模型和BL模型结合起来, 将其作为表达投资者观点的方法:

It is straightforward to incorporate factor models in the Black-Litterman framework. Let us assume we have a factor representation of the returns of some of the assets, that is

$$R_i = \alpha_i + \mathbf{F}\boldsymbol{\beta}_i + \varepsilon_i, i \in I$$

where $I \subset \{1, 2, \dots, N\}$. Typically, from a factor model it is easy to obtain an estimate of the residual variance, $\text{var}(\varepsilon_i)$. In this case, we set

$$q_i = \begin{cases} \alpha_i + \mathbf{F}\boldsymbol{\beta}_i, & i \in I \\ 0, & \text{otherwise} \end{cases}$$

and the corresponding confidence

$$\omega_{ii}^2 = \begin{cases} \text{var}(\epsilon_i), i \in I \\ 0, \text{ otherwise} \end{cases}$$

The **P** matrix is defined by

$$p_{ii} = \begin{cases} 1, i \in I \\ 0, \text{ otherwise} \end{cases}$$
$$p_{ij} = 0, i \neq j$$

需要注意的是，BL模型更多的用于大类资产的配置，因为资产类别相对较少，研究者可以将观点融入到预期收益和方差协方差矩阵的估计中去。