

量化投资--技术篇(5) 投资组合策略

Portfolio Policy

一. 前言

投资界基本上认可多元化投资是一种有效规避投资风险的一个技术手段，但是如何构建一个合理的多元化投资组合、具体技术方案和策略都是投资者们非常关注的问题。在本人的博客《机器学习与金融》系列中曾空提到投资组合策略的基本组成方式以及智能投顾所广泛采用的方案。本文就是对这些策略中关于资产配置的进一步的介绍。

二. 市场收益特性

首先，我们来总结一下市场上的资产收益有哪些特性。

- 金融市场收益基本上都不符合正态分布。
 - 资产收益呈现序列相关性。
 - 概率分布上显示左厚尾现象，也就是负偏度。
 - 高波动性下，金融资产的相关性并不稳定。
- 风格特性
 - 波动性聚簇
 - 自回归行为
 - 厚尾部特征
 - 尾部厚度变化基于不同频率数据
- 时序特性
 - 尖峰厚尾
 - 异常值明显
 - 单变量时序数据，特别是日数据，并不是独立同分布的。
 - 收益时序数据不具备方差齐性。
 - 绝对收益或这平方收益显示出较强的自相关性。
 - 同时期相关性并不稳定
 - 收益序列极值的出现往往伴随着其他收益序列极值的出现。

对于资产配置模型或者风险模型来说，导致了如下影响：

1. 假设独立同分布的模型，对于所有市场环境来说是不够的。
2. 假设正态分布的模型，会低估极端事件出现的频率。
3. 模型不得不考虑不同的波动性时期。
4. 模型应该考虑资产之间的依赖性。

三. 主要投资组合策略

1. Modern Portfolio Theory

这个是我们开篇就介绍的理论，可采用二次规划来求解均值方差最优，来得到在同样波动性下期望收益最大或者在同样期望收益下波动性最小。关于这套方法存在非常多的变种，笔者会在《实践篇》中详细介绍，并试图克服MVO存在的固有缺陷。

2. Black and Litterman Model

对于克服MVO中对输入太过敏感的问题，BL模型大概是理论上最理想的方案，它巧妙地使用贝叶斯分析方法来结合其他的输入数据。BL模型可以被看成是不同方法的结合提：MVO做组合优化，CAPM理论，并引入资产分配约束。它主要使用两类信息来预测收益：均衡收益和投资者观点。

分离定理告诉我们最优的资产分配方案应该是与可选资产的市值是成比例的。均衡期望收益可以从资产价格中导出，然后贝叶斯方法混合投资者关于资产未来价格的特有观点。这套理论我们也会在《实践篇》中加以详述。

BL模型的主要优点如下：

- 考虑了不确定性
- 平滑化投资者观点
- 收缩法的运用使其更具鲁棒性
- 给与投资者观点极大的自由度。

同时，BL模型的实践中，也有不少先驱者认为其

- BL模型只能提供便利性，在资产管理上并没有比其他方案显得更有效。
- BL依赖的参数较多，这意味着其错误的可能性更大。
- 同时采用均衡模型和投资者的先验知识，这本身存在一定矛盾。

3. Simple return-agnostic strategies

很多时候，模型的误差来自于参数太多。有鉴于此，一个自然而然的思路是构建一个需要参数较少的模型，这样可以减少对参数估算错误的机会。其实最难估计也是偏差最大的参数是对预期收益的估算，所以最多采用的收益无关策略都不需要预期收益作为输入参数。典型的两个方案为最小方差策略和等权重策略。

4. Strategies based on higher order moments

MVO模型在资产非正态分布或者投资者效用函数不是二次函数的时候，其表现往往差强人意。在这种情况下，一个可能的增强手段是结合高阶距，如偏度和峰度。有学者曾经采用一个两阶段的贝叶斯框架来解决高阶距和风险估计的问题。

当需要估算投资组合的高阶距时，通常会遇到两个挑战：第一，估算误差会被放大。第二，所需估算的参数数量比观察数据还要大。距离来说，例如有20个资产，协方差、协偏度和协峰度分别需要估算210，1540，8555个参数。

为了解决上述问题，有学者提出了采用单因子或多因子模型来导出协偏度和协峰度，另外，还有学者结合期权定价模型来融入前向信息。

5. Factor-based strategies

投资组合的收益是投资多个资产所得的收益的合成，我们也可以理解是市场的系统性根源因子的收益的组合。基于这种思想，一个值得探索的方法就是基于因子的组合策略。

经验表明大部分常见的风险因子几乎都是不相关的，即使实在市场危机爆发的时候。所以这种策略可以有效的把风险分散化。从前人的实践上来看，三个主要的因子类别是：

☐ 特征因子

属于该资产的可观察特征，例如股票所属企业的大小，B/P ratio，所属行业等。

☐ 宏观因子

通货膨胀，失业率等

☐ 统计因子

协方差等

三个最有价值的因子是

❑ 价值因子

衡量资产收益率和资产长期价值的关系。

❑ 动量因子

追逐成功者而淘汰失败者，主要看投资者使用公开信息的时序特性的能力。

❑ 质量因子

可以辨别企业或公司更能为股票持有人带来价值，而有效避免代理机构的陷阱。

学术界在这方面的研究非常多，另外还有一些商业应用，包括：

1) BIRR

因子包括投资者信心、利率、通货膨胀、商业活动和市场指数。

2) RAM

因子包括长期经济增长的变化的期望、短期商业周期、长期债券收益变化、短期国债变化、通胀抖动和利率。

3) BARRA

BARRA的方法论是基于横向研究的，它采用基本面分析结合风险因子。

6. Risk-based strategies

投资者早已意识到估算资产收益率有诸多缺点，所以基于风险的策略已称为投资组合策略中非常重要的一个分支。其主要的两类方法包括最大多元化和风险平价模型。我们会在《实践篇》中详细讨论。

7. Combining strategies

确定一个合适的方法来处理模型的不确定性不是一个简单的事情。单一的模型总是存在这样和那样的问题，所以可能采用模型的联合会取得更好的效果，从某种意义上讲，联合模型天生具备多元化属性。这里介绍三种思路做模型联合：

- 多个策略的简单平均
- 平均使样本外数据的方差最小
- 平均使样本外的收益率最大

8. "Online" strategies

在线策略与前面所述的策略有很大不同，主要是基于资产池顺序选择投资组合以达到一个投资目标。下表给出了前人所做出的分类和所尝试的方法。

- Follow The Winner
 - Universal Portfolios
 - Exponential Gradient
 - Follow the Leader
 - Follow the Regularized Leader
 - Aggregating-type Algorithms
- Follow The Loser
 - Anti Correlation
 - Passive Aggressive Mean Reversion
 - Confidence Weighted Mean Reversion
 - Online Moving Average Reversion

Robust Median Reversion

- Pattern Matching Approaches

Nonparametric Histogram Log-optimal Strategy

Nonparametric Kernel-based Log-optimal Strategy

Nonparametric Nearest Neighbor Log-optimal Strategy

Correlation-driven Nonparametric Learning Strategy

Nonparametric Kernel-based Semi-log-optimal Strategy

Nonparametric Kernel-based Markowitz-type Strategy

Nonparametric Kernel-based GV-type Strategy

- Meta-learning Algorithms

Aggregating Algorithm

Fast Universalization Algorithm

Online Gradient Updates

Online Newton Updates

Follow the Leading History

四. 总结

本文总结了学术界和工业界常用的资产配置和投资组合策略的主要方法，其中的部分方法会在《实践篇》中深入讨论。总的来书，这些方法各有千秋，笔者认为没有一种方法可以在任何情况下都优于另外一种方法。曾静有学者用13年的数据比较了多个方法，得出结论：没有一个方法可以一直击败最简单的 $1/n$ ，即等权重策略。

除了前面介绍的主流策略以外，还有一类比较影响力的策略是增长优化策略(Growth Optimal Portfolio)，该策略是按照期望增长优化方法渐进地追求长期复合收益最大化。但是在实践中，这种方法风险非常大，所以需要其他辅助方法控制风险，并把理论上地无限投资时间转换为有限时间。

实际操作中，我们往往需要结合约束。约束就是限制资产地投资比重不能超过或者不能少于一个约束限制条件。约束条件从某种意义上来说反映了投资者地一种观点，一般来说通过约束条件限制可以得到较好地样本外回测结果。我们可以从贝叶斯分析中地先验概率来理解约束，或者从机器学习中地过拟合限制条件来解释约束。

投资组合策略是一个非常复杂地问题，本文所介绍地只不过是个开始，实际中还要配合杠杆操作和调仓策略等，我们会在后续系列和《实践篇》中逐步深入。