

工程管理的整体性原则对量化投资行业的启示

王武俊 肖俊涛* (通讯作者)

湖北汽车工业学院 442002

DOI:10.32629/ems.v8i5.20133

[摘要] 文章将工程管理系统观的整体性原则应用于量化投资行业,发现当前该行业的研发框架在时间维度上缺乏整体性,即投资组合优化仅聚焦在单期优化上而未能考虑多期优化,该整体性的缺失限制了量化投资行业的创新方向。通过进一步分析该行业中导致单期优化达不到全局最优的原因,文章发现量化投资行业应该在使用当期信息预测下期信息、使用当期信息预测下期信息与下期收益率的协方差以及做多期组合优化这些方向上加大研究力度。

[关键词] 整体性; 量化投资; 多期优化

引言

量化投资是指使用量化的指标或信息,按照预定的规则来决定对每种资产交易的数量,以期取得好的投资效果。近些年,量化投资行业飞速发展,截至2024年全球的量化基金管理的资产规模已经达到了2~3万亿美元^[1]。但是,该行业竞争非常激烈,即,当一个公司研发出好的投资策略,其不太可能长时间地使用该策略赚钱,因为别的公司也在研究金融市场,所以当别的公司也研发出该策略时,利润就会逐渐消失。所以,一个公司想要长久赚钱,就必须不断地研发出新的策略。

本文通过总结当前量化投资行业的研发框架的主要形式,结合工程管理系统观的整体性原则,发现其在多期优化方面研究不足,并为我们开发或改进量化投资策略指明了新的研究领域。

一、当前量化投资的主要框架及缺点

(一) 当前量化投资的主要框架

通过总结主要的量化投资书籍^{[2]-[5]},我们得到量化投资的主要模式如下。

在 t 时刻使用一些信息 \mathbf{x}_t (其为向量,表示各种信息)预测 t 到 $t+1$ 时刻的多种资产的收益率的 $\mathbf{r}_{t,t+1}$ (其为向量,表示多个资产的收益率),预测的主要是 $\mathbf{r}_{t,t+1}$ 的均值 $E(\mathbf{r}_{t,t+1}|\mathbf{x}_t)$ 和方差矩阵 $V(\mathbf{r}_{t,t+1}|\mathbf{x}_t)$ 。然后做组合优化,即选择每种资产上要投资的权重(即金额比例) \mathbf{w}_t (其各成分之和为1),使得投资组合的期望收益率进行风险调整后最大。所以组合优化问题可以表示为求解下述问题,

$$\max_{\mathbf{w}_t} \mathbf{w}_t' E(\mathbf{r}_{t,t+1}|\mathbf{x}_t) - \lambda \mathbf{w}_t' V(\mathbf{r}_{t,t+1}|\mathbf{x}_t) \mathbf{w}_t \quad (1.1)$$

其中 $\lambda > 0$ 为风险厌恶系数,求解该问题就能得到 t 时刻应该购买的各种资产的比例 \mathbf{w}_t 。

另外,现实中交易一个资产是有交易成本的。一般来说,

交易的量越大交易(冲击)成本越高,所以在决定各资产的投资比例时还要考虑交易成本。一般做法是对交易成本建模,即某种资产 i 在 t 到 $t+1$ 时刻之间的交易成本 TC_{it} (即实际购买的平均价格 \bar{P}_{it} 高出基准价格 P_{it} 的百分比 $\frac{\bar{P}_{it} - P_{it}}{P_{it}}$,或实际出售的平均价格 \bar{P}_{it} 低于基准价格 P_{it} 的百分比 $\frac{P_{it} - \bar{P}_{it}}{P_{it}}$)是交易量 X_{it} (如股票的手数)的某种函数形式 $TC_{it} = f(X_{it}, \gamma)$,然后使用历史交易数据估计其中的系数 γ 。然后在上述组合优化的过程中把交易成本也考虑在内,则组合优化问题为

$$\max_{\mathbf{w}_t} \mathbf{w}_t' E(\mathbf{r}_{t,t+1}|\mathbf{x}_t) - \lambda_1 \mathbf{w}_t' V(\mathbf{r}_{t,t+1}|\mathbf{x}_t) \mathbf{w}_t - \sum_{i=1}^N |w_{it} - w_{it}^b| \cdot f\left(gmv_t \frac{|w_{it} - w_{it}^b|}{P_{it}}, \gamma\right) \quad (1.2)$$

其中 w_{it}^b 为 t 时刻调仓之前已经持有的资产 i 的比例, gmv_t 为 t 时刻调仓之前持有的所有资产的总市值。求解上述最优化问题,就能得到 t 时刻应该购买的各种资产的比例 \mathbf{w}_t 。

在上述框架下研发或改进量化投资策略有三个方向。

第一,寻找各种信号预测资产的期望收益率 $E(\mathbf{r}_{t,t+1}|\mathbf{x}_t)$,称为收益模型。量化投资策略创新的主要努力集中在这里。例如,从行为金融或心理学的理论出发找一些信号来预测期望收益率^[6];使用机器学习或深度学习用量价信息来预测资产收益率^[7];使用文本等非结构化数据预测资产收益率等等^[8]。总之,有什么新的信息变量或新的预测方法,量化投资公司会立即尝试。

第二,寻找各种信号预测资产的方差矩阵 $V(\mathbf{r}_{t,t+1}|\mathbf{x}_t)$,称为风险模型。关于方差矩阵的预测,有专业公司专注于生产风险模型,如MSCI的Barra风险模型^[9],其目的是让量化投资公司专注于收益模型。当然,量化投资公司基本都会开发自己的风险模型:其中,有简单的,如只是在组合优化的时

候控制在市场风险因子上暴露为零;也有复杂的,如在组合优化时,控制在多个风险因子上的暴露为零并使得投资组合方差的小于某一阈值。但是,总体来讲量化公司在风险模型上的投入没有收益模型上的大。

第三,想各种办法减少交易成本,称为算法交易。算法交易与上述组合优化问题一般而言是不同的。组合优化的目的是,决定接下来一段时间(如1周),投资人要持有的资产数量,称为目标持有量;而算法交易,是由组合优化决定的目标持有量跟已经持有量计算差额从而得到要交易的数量,然后在相对较短的时间内(如1日),用最低的交易成本完成这交易数量。算法交易中也要使用一些信号做预测,目的是使得交易成本最小甚至还有额外收益。证券公司一般都有部门研发算法交易,但是随着量化投资扩展到了较小的时间尺度上,如高频交易,不少量化投资公司也自己研发算法交易。

(二) 当前量化投资框架的缺点

工程管理系统观的整体性原则指出,任何一个工程或项目都由多个子系统构成,这些子系统之间存在紧密联系和相互依赖的关系;管理者必须从整体的角度出发统筹规划,综合考虑系统各部分之间的相互作用及对整体目标的贡献,而不应仅优化某个局部子系统的性能或目标,否则很可能达不到全局最优^[10]。

将该原则与上述量化投资的框架相比较,可以看到上述组合最优化问题中,优化的目标是单期的,每一期与下一期的优化是没有关联的,从而可能导致虽然局部(一期)达到最优,但是没有达到全局(多期)最优。下面我们就单期优化达不到全局最优的原因进行详细分析,以便我们改进之。单期优化达不到全局最优的原因至少有两个,分别称为基于交易成本的原因,和基于对冲状态变量风险的原因。

对基于交易成本的原因,我们通过下述具体例子来说明。假设除了现金之外,只有一个风险资产,如果在 t 时刻预测其预期收益率 $E(r_{t+1}|x_t)$ 为3%,且风险为0,但是其交易成本为5%,则在上述组合优化框架下我们不应该购买该资产。但是,如果在 t 时刻,我们通过 t 时刻信息 x_t 还能够预测到 $t+1$ 时刻信息 x_{t+1} 很可能保持不变,即该资产在 $t+1$ 到 $t+2$ 时刻的收益率很可能也是3%,那么这个资产在 t 到 $t+2$ 的时间段里是值得持有的。所以,由于当前的量化投资框架(尤其是组合优化)只考虑了单期最优而没有考虑多期最优,从而错过了很好的投资机会。

对于基于对冲状态变量风险的原因我们也通过一个例子来说明。具体来说,假设投资人的目标是,通过在每一期交易资产使得 T 期(简单起见,以2期为例)后的资产总值对应的效用的期望最大,即

$$\begin{aligned} & \max_{\mathbf{w}_0, \mathbf{w}_1} E_0(U(W_2)) \\ & s.t. \\ & W_1 = W_0(1 + \mathbf{w}'_0 \mathbf{r}_{0,1}) \\ & W_2 = W_1(1 + \mathbf{w}'_1 \mathbf{r}_{1,2}) \end{aligned} \quad (1.3)$$

。其中 W_0 已知,为0时刻的资产总额; $U(\cdot)$ 为效用函数。投资者需要求解得到0时刻和1时刻的最优投资组合 $\mathbf{w}_0, \mathbf{w}_1$ 。这是一个动态规划问题,所以我们使用动态规划的逆推法原理解。首先,对于1时刻的投资组合,投资者根据 \mathbf{x}_1 得到 $E(\mathbf{r}_{1,2}|\mathbf{x}_1)$ 和 $V(\mathbf{r}_{1,2}|\mathbf{x}_1)$,然后根据(1.1)式求解最优组合,可以看到1时刻的最优组合求解方法与前述单期求解方法相同。有了1时刻的最优组合,我们就要求解0时刻的最优组合,在求解0时刻的最优组合前,根据逆推法原理,我们要先得到1时刻的价值函数,根据上述求解过程,可以看到该价值函数的状态变量为 \mathbf{x}_1, W_1 ,将该价值函数记为 $V_1(\mathbf{x}_1, W_1)$ 。然后,根据逆推法原理,0时刻的组合求解问题为

$$\begin{aligned} & \max_{\mathbf{w}_0} E_0(V_1(\mathbf{x}_1, W_1)) \\ & s.t. \\ & W_1 = W_0(1 + \mathbf{w}'_0 \mathbf{r}_{0,1}) \end{aligned} \quad (1.4)$$

对于该问题的求解,如果价值函数 V_1 只是 W_1 的函数,则其求解仍然等同于前面的单期最优化,但是正是由于 V_1 不仅是 W_1 而且是 \mathbf{x}_1 的函数,所以该问题求解不同于前面的单期最优化,即该问题的最优组合不同于单期均值-方差最优化得到的组合。

根据多期组合优化理论^[11],上述问题的最优化的组合是在均值方差最优化组合的基础上加上一定数量的一个对冲组合,该对冲组合的目的通过购买一定数量的该对冲组合来减少 \mathbf{x}_1 带来的风险。要求解出该对冲组合的具体成分数量,必须在0时刻知道1时刻的资产收益率和状态变量的协方差矩阵,更一般地,需要知道 $Cov(\mathbf{x}_{t+1}, \mathbf{r}_{t+1}|\mathbf{x}_t)$ 。

二、改进方向

工程管理系统观的整体性原则要求我们最做组合最优化时,不应该把每一期割裂起来,仅仅单独考虑每一期的最优化,还应该考虑每一期之间的关联,应该关注多期的最优化。

动态规划的框架正好就是考虑了当期与下期之间关联而且进行多期最优化的最成熟的框架^[12]。具体来说,动态规划问题中,系统 $P(\mathbf{r}_{t,t+1}, \mathbf{x}_{t+1}|\mathbf{x}_t)$ 给出了条件在信息 \mathbf{x}_t 上的资产收益率 $\mathbf{r}_{t,t+1}$ 和下期状态 \mathbf{x}_{t+1} 的条件概率,然后将期望多期收益作为目标函数,就可以求解出多期最优的策略 $\mathbf{w}_t(\mathbf{x}_t)$ 。将该动态规划框架与上述单期最优化框架相比较,可见单期最优化框架基于当期信息 \mathbf{x}_t ,只预测了下期收益率 $\mathbf{r}_{t,t+1}$ 的均值和

方差矩阵,而没有基于当期信息 \mathbf{x}_t 预测下一期信息 \mathbf{x}_{t+1} ,也没有基于当期信息 \mathbf{x}_t 预测下一期信息 \mathbf{x}_{t+1} 和下期收益率 $\mathbf{r}_{t,t+1}$ 之间的协方差矩阵,也从而无法考虑当期与下期之间的关联,从而无法进行多期最优化。

基于前述基于单期最优化的框架,我们已经知道当前量化投资行业主要通过三个方向进行研发和创新。而基于动态规划框架,我们得出两个新的创新方向。一个是使用当期信息 \mathbf{x}_t 预测下一期信息 \mathbf{x}_{t+1} ,另一个是使用当期信息 \mathbf{x}_t 预测下一期信息 \mathbf{x}_{t+1} 和下期收益率 $\mathbf{r}_{t,t+1}$ 之间的协方差矩阵。如果这两方面的内容预测得好,我们就可以将其应用在动态规划框架中,从而提升我们的量化投资策略的效果。

三、障碍和破解途径

事实上,上述动态规划的多期模型在学术界很早就广为人知^[11],但是量化投资实业界为什么却没有采用该模型而采取了单期最优化模型呢?可能的原因有两个。

第一个原因是,动态规划求解原理虽然简单,但是当 \mathbf{x}_t 的维度很高的时候很难求解。但是随着近些年强化学习技术的突飞猛进,求解上述高纬度问题变得比以前容易多了。例如,相比于以前的列表形式的强化学习方法,近似强化学习方法(将价值函数或行动函数设定成状态变量的参数化函数来学习这些参数从而获得近似的最优策略)的计算成本要小多了,从而能求解高维问题^[12]。

第二个原因是,金融系统的信噪比一般低很,仅是估计 \mathbf{x}_t 与 $\mathbf{r}_{t,t+1}$ 的关系就很可能导致估计出来的系数方差较大,从而导致组合优化出来的组合很可能不是最优的(即回测过拟合)^[14]。如果还要去预测 \mathbf{x}_t 和 \mathbf{x}_{t+1} 的关系以及 \mathbf{x}_t 和 \mathbf{x}_{t+1} 与 $\mathbf{r}_{t,t+1}$ 协方差的关系,那么整个系统的方差就会更大,很可能使得过拟合更严重。但是,第一,在量化投资实务中,到底是加入这些额外的关系带来的好处(多期最优)大,还是坏处(更大的系统方差)大,还是要(使用样本外数据)回测了才知道。第二,我们也可以从经济逻辑出发,对于 \mathbf{x}_t 和 \mathbf{x}_{t+1} 的关系以及 \mathbf{x}_t 和 \mathbf{x}_{t+1} 与 $\mathbf{r}_{t,t+1}$ 协方差的关系推导出一些有用的结论,从而在做多期最优化的时候使用,所以很可能不会增加整个系统的方差。

四、结论

工程管理的系统观的整体性原则,启示我们当前量化投资行业存在只考虑单期优化而缺乏多期优化的弊端,启示我们一方面要加强用当期信息预测下期信息方面的研发力度,另一方面要加强用当期信息预测下期信息与下期收益率的协方差方面的研发力度,并做多期最优化。在当前量化投资行

业日益竞争激烈的背景下,这为我们提升量化投资策略的整体效果提供了新思路,开辟了新赛道。

[参考文献]

- [1]Wikipedia. Quantitative fund[EB/OL]. (2025-09-24) [2025-09-24]. https://en.wikipedia.org/wiki/Quantitative_fund.
- [2]Pedersen, L. H. Efficiently inefficient: How smart money invests and market prices are determined[M]. Princeton University Press, 2015.
- [3]Grinold, R. C., & Kahn, R. N. Active portfolio management[M]. 2nd ed. McGraw-Hill, 2000.
- [4]Zhou, X., & Jain, S. Active equity management[M]. Wiley, 2014.
- [5]Narang, R. K. Inside the black box: A simple guide to quantitative and high-frequency trading[M]. 2nd ed. Wiley, 2013.
- [6]Lee, C. M., & So, E. C. Alphanomics: The informational underpinnings of market efficiency[J]. Foundations and Trends in Accounting, 2015, 9(2-3): 59-258.
- [7]Zhang, Z., Zohren, S., & Roberts, S. DeepLOB: Deep convolutional neural networks for limit order books[J]. IEEE Transactions on Signal Processing, 2019, 67(11): 3001-3012.
- [8]Ke, Z. T., Kelly, B. T., & Xiu, D. Predicting returns with text data[C]//National Bureau of Economic Research. Working Paper No. 26186. Cambridge, MA: NBER, 2019.
- [9]MSCI. The Barra US Equity Model (USE4) [EB/OL]. (2011) [2024-12-20]. Retrieved from MSCI website.
- [10]汪应洛. 工程管理概论 [M]. 西安: 西安交通大学出版社, 2013.
- [11]Back, K. E. Asset pricing and portfolio choice theory[M]. 2nd ed. Oxford, UK: Oxford University Press, 2017.
- [12]Bertsekas, D. P. Dynamic programming and optimal control[M]. Vol. 1, 3rd ed. Athena Scientific, 2005.
- [13]Sutton, R. S., & Barto, A. G. Reinforcement learning: An introduction[M]. 2nd ed. MIT Press, 2020.
- [14]Bailey, D. H., Borwein, J. M., Salehipour, A., López de Prado, M., & Zhu, Q. Backtest overfitting in financial markets[J]. Journal of Computational Finance, 2016, 20(4): 39-69.