文章编号: 1001-0920 (2014) 07-1181-06 DOI: 10.13195/j.kzyjc.2013.0597

基于灵敏度分析的含比例型手续费的投资组合优化

黄永皓, 陈 曦
(1. 清华信息科学与技术国家实验室 (筹)，北京 100084；2. 清华大学 自动化系，北京 100084)

摘 要: 研究含比例型手续费的离散时间投资组合优化问题. 基于马尔可夫决策过程模型和性能灵敏度分析方法, 推导两个不同投资策略之间的资产长期平均增值率的差分公式, 利用差分公式的结构特点, 证明了最优性方程, 并设计出可在线应用的策略迭代算法. 仿真实例验证了所提出算法的有效性.

关键词: 投资组合；马尔可夫决策过程；灵敏度分析；策略迭代
中图分类号: TP13 **文献标志码:** A

Sensitivity-based optimization approach for portfolio optimization with proportional transaction costs

HUANG Yong-hao, *CHEN Xi*
(1. Tsinghua National Laboratory for Information Science and Technology，Beijing 100084，China；2. Department of Automation，Tsinghua University，Beijing 100084，China．Correspondent：HUANG Yong-hao，E-mail：huangyonghao@tsinghua.org.cn)

Abstract: The discrete-time portfolio optimization problem with proportional transaction costs is considered. By formulating this problem as a Markov decision problem, the difference of the long-run average growth rate of total wealth of two different policies is obtained by using a sensitivity-based approach. By analyzing the structure of the difference formula, the optimality equation and the policy iteration algorithm for the optimal investment policy can be intuitively derived. Thereby the on-line implementation of the policy iteration algorithm is designed. A numerical example illustrates the effectiveness of the proposed algorithm.

Key words: portfolio optimization；Markov decision process；sensitivity-based approach；policy iteration

0 引 言

动态投资组合优化问题是金融领域的基本问题, 至今已有50余年的研究历史[1-5]. 在动态投资组合优化问题中, 投资者为了实现资产最大化, 需要适时地调整风险资产 (如股票等) 与无风险资产 (如债券等) 的比例. 在传统的研究中, 投资组合优化问题首先被描述为随机控制问题, 然后利用动态规划原理和哈密尔顿-雅可比-贝尔曼方程 (HJB 方程) 进行求解[1-4]. 该类方法常常以连续时间模型为基础, 允许无穷小时间内的无穷小量交易, 这在实际中往往不可行. 同时, 如果资产特性未知, 或者交易需付手续费, 则求解 HJB 方程会比较困难[6-7]. 针对这些情况, 需要考虑采用数值方法或近似方法进行求解, 如文献 [5, 8].

本文研究离散时间的动态投资组合优化, 并考虑含有比例型交易成本的问题, 投资优化的目标是使资产的长期平均增值率最大化. 本文采用基于性能灵敏度分析的优化方法, 首先分析了不同投资策略下, 资产长期平均增值率的差异, 并以此为基础, 推导出最优性条件, 开发出可以在线应用的基于样本路径的策略迭代优化算法.

相对于传统方法, 基于性能灵敏度分析的优化方法具有分析简洁直观、便于开发基于样本路径的在线算法、在资产特性的参数未知或者模型结构变化不大 (仍然满足马尔可夫性) 时仍然能够应用等特点. 同时, 与文献 [9] 基于灵敏度分析方法在线性系统连续状态问题的应用相比, 投资组合问题作为一般连续状态问题, 系统动态特性较为复杂. 本文通过利用有限状态拟合所有状态的性能势, 给出了连续状态问题性

收稿日期: 2013-05-09；**修回日期:** 2013-10-15.
基金项目: 国家自然科学基金项目(61203039)；高校学科创新引智计划项目(B06002)；清华信息科学与技术国家实验室 (筹).
作者简介: 黄永皓(1983−), 男, 博士生, 从事马尔可夫决策过程的研究；陈曦(1965−), 女, 副研究员, 从事随机最优控制、传感器网络等研究.

能势估计的通用公式. 仿真实例验证了在线策略迭代算法的有效性.

1 投资组合优化问题

考虑包含一个无风险资产 (如债券) 与 N 个风险资产 (如股票) 的投资组合. 对于增值率为常数 r_0 的无风险资产, 其价格 $p_0(t)$ 的变化满足 $p_0(t+1)/p_0(t) = \mathrm{e}^{r_0}$; 对于风险资产, 这里考虑基本的几何布朗运动模型, 第 i 支风险资产的价格 $p_i(t)$ 变化服从

$$p_i(t+1)/p_i(t) = \mathrm{e}^{r_i+\sigma_i w_{i,t}}.$$

其中: r_i 为增值率; σ_i 为波动率; $w_{i,t}$ 为服从标准正态分布的随机数, 满足 $E[w_{i,t}] = 0$, $E[w_{i,t}w_{j,t}] = \alpha_{i,j}$, $\alpha_{i,j}$ 描述了第 i 支风险资产与第 j 支风险资产之间的相关性.

令 $S_0(t)$ 表示在时刻 t 投资组合中无风险资产的市值, $S_i(t)$ 表示在时刻 t 投资组合中第 i 支风险资产的市值, $i = 1, 2, \cdots, N$. 在每个时刻, 投资者均会根据既定的投资策略决定如何分配组合中无风险资产与风险资产的比例. 令 $U_i(t)$ 表示在时刻 t 从无风险资产流向第 i 支风险资产的资金量 (即出售无风险资产, 购买第 i 支风险资产), $V_i(t)$ 表示在时刻 t 从第 i 支风险资产流向无风险资产的资金量 (即出售第 i 支风险资产, 购买无风险资产), $i = 1, 2, \cdots, N$. 这里不考虑融资或融券交易, 即买卖风险资产不存在借贷资金或资产的情况. 假设交易成本与实际交易资金量成正比, 买卖第 i 支风险资产的手续费比例分别记为 $\lambda_i \in [0, 1]$ 和 $\mu_i \in [0, 1]$, $i = 1, 2, \cdots, N$. 根据以上描述, 资产组合的动态变化满足如下方程:

$$S_0(t+1) = \mathrm{e}^{r_0}\Big[S_0(t) - \sum_{i=1}^{N}(1+\lambda_i)U_i(t) + \sum_{i=1}^{N}(1-\mu_i)V_i(t)\Big], \tag{1}$$

$$S_i(t+1) = \mathrm{e}^{r_i+\sigma_i w_{i,t}}[S_i(t)+U_i(t)-V_i(t)],\quad i = 1, 2, \cdots, N. \tag{2}$$

其中 $w_{i,t}$ 为服从标准正态分布的随机数, $i = 1, 2, \cdots, N$. 在时刻 t, 资产组合的总价值记为

$$M(t) = S_0(t) + \sum_{i=1}^{N} S_i(t).$$

投资者的目标是实现资产组合总价值的长期平均增值率 η 最大, 即

$$\eta = \lim_{T\to\infty}\frac{1}{T}E[\ln M(T)]. \tag{3}$$

以上是投资组合优化问题的基本描述, 但从控制的角度看, 该随机过程的状态 $\{S_0(t), S_1(t), \cdots, S_N(t), t \geqslant 0\}$ 随着时间推移可能会变得无穷大, 因此是不稳定的. 定义新的状态变量

$$s_0(t) := \frac{S_0(t)}{M(t)},\ s_i(t) := \frac{S_i(t)}{M(t)},\ i = 1, 2, \cdots, N.$$

新的控制泛函

$$u_i(t) := \frac{U_i(t)}{M(t)},\ v_i(t) := \frac{V_i(t)}{M(t)},\ i = 1, 2, \cdots, N.$$

基于新定义状态变量和控制泛函, 资产总值满足

$$M(t+1) = M(t)\Big\{\mathrm{e}^{r_0}\Big[s_0(t) - \sum_{i=1}^{N}(1+\lambda_i)u_i(t) + \sum_{i=1}^{N}(1-\mu_i)v_i(t)\Big] + \sum_{i=1}^{N}\mathrm{e}^{r_i+\sigma_i w_{i,t}}(s_i(t)+u_i(t)-v_i(t))\Big\}. \tag{4}$$

资产总值的长期平均增值率可以改写为

$$\eta = \lim_{T\to\infty}\frac{1}{T}E\Big\{\sum_{t=0}^{T-1}\ln\Big[\mathrm{e}^{r_0}\Big(s_0(t) - \sum_{i=1}^{N}(1+\lambda_i)u_i(t) + \sum_{i=1}^{N}(1-\mu_i)v_i(t)\Big) + \sum_{i=1}^{N}\mathrm{e}^{r_i+\sigma_i w_{i,t}}(s_i(t)+u_i(t)-v_i(t))\Big]\Big\}. \tag{5}$$

不难看出, 新的状态变量 $s_0(t) \in [0, 1]$ 和 $s_i(t) \in [0, 1]$ $(i = 1, 2, \cdots, N)$ 分别为无风险资产和第 i 支风险资产占资产总值的比例, 满足

$$s_0(t) + \sum_{i=1}^{N} s_i(t) = 1,\ \forall t \geqslant 0.$$

类似地, 新的控制泛函 $u_i(t)$ 和 $v_i(t)$ 分别为时刻 t 买卖第 i 支风险资产占资产总值的比例, 满足

$$0 \leqslant \sum_{i=1}^{N}(1+\lambda_i)u_i(t) \leqslant s_0(t),\ 0 \leqslant v_i(t) \leqslant s_i(t).$$

根据新定义的状态变量和控制泛函, 投资组合优化问题可以看作是随机过程 $s(t) = \{s_0(t), s_1(t), \cdots, s_N(t)\}$ 上的最优控制问题, 控制策略 (即投资策略) 为 $d = \{u_i(t, s(t)), v_i(t, s(t)), i = 1, 2, \cdots, N\}$. 投资者的目标是寻找最优的投资策略, 使得其对应的资产总值长期平均增值率最大. 最优资产增值率记为 $\eta^* = \max_d \eta^d$, 其中上标 d 为策略的相关性.

2 基于性能灵敏度分析的优化方法

本节采用基于性能灵敏度分析的方法分析和求解投资组合优化问题. 首先定义与投资组合优化问题等价的马尔可夫决策过程, 然后基于性能势的核心概念, 推导不同投资策略的资产长期增值率的差分公式, 并设计策略迭代算法.

2.1 等价马尔可夫决策过程

考虑以 $s(t) = \{s_0(t), s_1(t), \cdots, s_N(t),\ t \geqslant 0\}$ 为状态变量的投资组合优化问题, 由于在任何时刻, 均

有 $s_0(t)+\sum_{i=1}^{N}s_i(t)=1$, 只需取 $X(t)=(s_1(t),s_2(t),\cdots,s_N(t))^{\mathrm{T}}$作为马尔可夫决策过程的状态变量, 即以各支风险资产占资产总值的比例为状态变量, 状态空间为 $\boldsymbol{S}=[0,1]^N$. 文献 [3] 已证明, 在风险资产增值率与波动性都时不变的情况下, 连续时间投资组合优化问题的最优投资策略是平稳的, 因此只考虑平稳策略, 即投资策略 $d=\{u_i(s(t)),v_i(s(t)),\ i=1,2,\cdots,N\}$ 是从状态空间到行动空间的映射.

令 $\boldsymbol{B}$ 为状态空间 $\boldsymbol{S}$ 的 σ 代数. 为了简单, 假设本文所讨论的集合和函数均可测. 当时刻 t 各支风险资产所占总资产的比例为 $X(t)=x\in[0,1]^N$ 时, 由于各种资产的价格变化和投资者根据某既定策略 d 进行交易, 在时刻 $t+k$, 状态变量 $X(t+k)$ 处于子集 $B\in\boldsymbol{B}$ 的概率可以表示为 $P^{d,k}(B|x)$, $k\geqslant 1$, 上标 d 为其策略相关性. 集函数 $P^{d,k}(B|x)$ 称为 k 步状态转移函数, 满足 $P^{d,k}(\boldsymbol{S}|x)=1$, $\forall x\in\boldsymbol{S}$. 当 $k=1$ 时, 一步状态转移函数 $P^{d,1}$ 通常简记为 P^d. 为了一致性, 记 $P^{d,0}(B|x)=I(B|x)$, I 为示性函数, 满足 $I(B|x)=1$ 当且仅当 $x\in B$. 两个转移函数 $P^{d,k_1}(B|x)$ 和 $P^{d,k_2}(B|x)$ 的乘积定义为

$$(P^{d,k_1}P^{d,k_2})(B|x):=\int_{\boldsymbol{S}}P^{d,k_2}(B|y)P^{d,k_1}(\mathrm{d}y|x). \tag{6}$$

进而有 $P^{d,k}=P^{d,k-1}P^d$. 对于任何转移函数 $P^{d,k}(B|x)(k\geqslant 0)$ 和任何函数 $q(x)$, 定义对应的转移算子为

$$(\boldsymbol{P}^{d,k}q)(x):=\int_{\boldsymbol{S}}q(y)P^{d,k}(\mathrm{d}y|x). \tag{7}$$

同时, 可知 $(\boldsymbol{P}^{d,k}q)(x)=E[q(X(k))|X(0)=x]$.

根据总资产长期平均增值率的表达式 (5), 可以定义单阶段总资产的期望增值率. 当风险资产的比例为 $(s_1,s_2,\cdots,s_N)$ 时, 定义

$$f^d(s_1,s_2,\cdots,s_N)=$$
$$E\Big\{\ln\Big[\mathrm{e}^{r_0}\Big(1-\sum_{i=1}^{N}s_i-\sum_{i=1}^{N}(1+\lambda_i)u_i+$$
$$\sum_{i=1}^{N}(1-\mu_i)v_i\Big)+\sum_{i=1}^{N}\mathrm{e}^{r_i+\sigma_iw_i}(s_i+u_i-v_i)\Big]\Big\}. \tag{8}$$

其中: 上标 $d=\{u_i,v_i,i=1,2,\cdots,N\}$ 为策略相关性, f^d 为关于 n 维正态随机变量 $(w_1,w_2,\cdots,w_N)$ 求期望. 给定某初始资产比例 $x=(s_1,s_2,\cdots,s_N)$, 在策略 d 下, 总资产长期平均增值率为

$$\eta^d(x)=$$
$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}E[f^d(X(t))\Big|X(0)=x]=$$
$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}(\boldsymbol{P}^{d,t}f^d)(x)=(\boldsymbol{P}^{d*}f^d)(x), \tag{9}$$

其中 $\boldsymbol{P}^{d*}$ 为样本路径平均算子, 定义为

$$\boldsymbol{P}^{d*}:=\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\boldsymbol{P}^{d,t}.$$

在马尔可夫决策过程中, η^d 称为策略 d 的性能指标, f^d 称为单步收益函数. 因为风险资产的波动性满足高斯分布, 而高斯分布在状态空间的各个子集均有支撑, 所以假设对于任意函数 $q(x)$, 当 $t\to\infty$ 时, 有如下遍历性:

$$\lim_{t\to\infty}(\boldsymbol{P}^{d,t}q)(x)=(\pi^dq)e(x). \tag{10}$$

其中: π^d 为概率测度; $\pi^dh:=\int_{\boldsymbol{S}}q(x)\pi^d(\mathrm{d}x)$ 为在测度 π^d 下函数 h 的均值; $e(x)\equiv 1,\forall x\in\boldsymbol{S}$. π^d 通常称为转移函数 $P^{d,k}$ 对应的稳态分布 (或不变分布), 它可以看作是由投资策略 d 经过长期交易使得资产分布稳定后, 各支风险资产占总资产的比例, 因此有

$$\eta^d(x)=\lim_{k\to\infty}(\boldsymbol{P}^{d,k}f^d)(x)=(\pi^df^d)e(x). \tag{11}$$

在不产生歧义的情况下, 本文也用 η^d 表示标量 π^df^d, 有 $\eta^d(x)=\eta^de(x)$.

对于任意 $x\in\boldsymbol{S}$, 定义性能势 $g(x)$ 为

$$g^d(x)=$$
$$\lim_{T\to\infty}E\Big\{\sum_{t=0}^{T-1}[f^d(X(t))-\eta^d]\Big|X(0)=x\Big\}. \tag{12}$$

直观上说, $g^d(x)$ 度量了某种初始风险资产比例对总资产长期平均增值率的潜在贡献. 可以证明, $g^d(x)$ 满足如下泊松方程:

$$(\boldsymbol{I}-\boldsymbol{P}^d)g^d(x)=f^d(x)-\eta^d(x),\ x\in\boldsymbol{S}. \tag{13}$$

泊松方程的解相差一个常数, 即如果 g^d 是式 (13) 的解, 则对于任意常数 c, g^d+ce 也是它的解.

2.2 投资策略优化

基于灵敏度分析方法的核心思想是从两个策略的性能差异入手评价和改进策略. 首先讨论两个不同投资策略下资产增值率的差分公式. 记两个投资策略 d 和 h 对应的马尔可夫过程分别为 $\{X^d(t),t\geqslant 0\}$ 和 $\{X^h(t),t\geqslant 0\}$, 同理, 其他与策略相关的量也以不同上标 d 和 h 区分. 根据泊松方程 (13), 两个投资策略下总资产的长期平均增值率满足如下引理.

引理 1　增值率差分公式为

$$\eta^h-\eta^d=\pi^h[(f^h+\boldsymbol{P}^hg^d)-(f^d+\boldsymbol{P}^dg^d)]. \tag{14}$$

证明　由式 (9)、(10)和泊松方程 (13) 可得

$$\eta^h(x)-\eta^d(x)=(\boldsymbol{P}^{h*}f^h)(x)-\eta^d(x)=$$
$$(\boldsymbol{P}^{h*}f^h)(x)+(\boldsymbol{P}^{h*}g^d)(x)-(\boldsymbol{P}^{h*}g^d)(x)-$$
$$(\boldsymbol{P}^{h*}\eta^d)(x)+(\boldsymbol{P}^{h*}\eta^d)(x)-\eta^d(x)=$$
$$[\boldsymbol{P}^{h*}(f^h+\boldsymbol{P}^hg^d)](x)-[\boldsymbol{P}^{h*}(g^d+\eta^d)](x)+$$
$$\eta^d(\boldsymbol{P}^{h*}e)(x)-\eta^d(x)=$$

$\{\pi^h[(f^h+\boldsymbol{P}^h g^d)-(f^d+\boldsymbol{P}^d g^d)]\}e(x)$. □

上式最后一步利用了变形之后的泊松方程 $g^d+\eta^d=\boldsymbol{P}^d g^d+f^d$. 该性能差分公式的直观解释为: 在两个不同投资策略下, 总资产的长期平均增值率由两部分决定, 一部分是当前状态两个策略单阶段资产增值率的差, 即 f^h-f^d; 另一部分是当前交易对总资产长期平均增值率在未来的长期影响, 由 $(\boldsymbol{P}^h g^d-\boldsymbol{P}^d g^d)$ 度量.

根据差分公式 (14), 可以设计出求解最优投资策略的迭代算法. 首先, 需要重新定义两个函数之间的关系 $=$、$\geqslant$、$>$ 和 $\succeq$. 给定状态空间 $\boldsymbol{S}$ 上的某个概率测度 ν, 对于函数 $q(x)$ 和 $q'(x)$, 除了某个测度为零 (满足 $\nu(B)=0$) 的集合 B 外, 对于所有 $x\in\boldsymbol{S}$, 有:

1) 如果 $q'(x)=q(x)$, 则记 $q'=_\nu q$;

2) 如果 $q'(x)\geqslant q(x)$, 则记 $q'\geqslant_\nu q$;

3) 如果 $q'(x)>q(x)$, 则记 $q>_\nu q$;

4) 如果 $q'(x)\geqslant q(x)$, 且在某个测度为正 $(\nu(B)>0)$ 的集合 B 内有 $q'(x)>q(x)$, 则记 $q\succeq_\nu q$.

类似地, 还可以定义关系 $\leqslant_\nu$、$<_\nu$ 和 $\prec_\nu$. 由上述定义的关系符号和性能差分公式 (14), 得到如下引理.

引理 2 (比较引理) 如果 $f^h+\boldsymbol{P}^h g^d\succeq_{\pi^h} f^d+\boldsymbol{P}^d g^d$, 则 $\eta^h>\eta^d$.

通常, 由于不知道 π^h 的分布, 验证关系 $\succ_{\pi^h}$ 较为困难. 然而, 在投资组合优化问题中, 因为风险资产的波动性由高斯分布描述, 高斯分布在状态空间的任何子集上均有支撑, 对于所有平稳策略 d, 其对应的稳态概率测度满足 $\pi^d(B)>0, \forall B\in\boldsymbol{B}$, 所以可以在关系符号中略去下标 π^d. 由引理 2, 得到如下最优条件和最优方程.

定理 1 (最优条件与最优方程) 策略 $\hat{d}$ 是最优的, 当且仅当

$$f^{\hat{d}}+\boldsymbol{P}^{\hat{d}}g^{\hat{d}}\succeq f^d+\boldsymbol{P}^d g^{\hat{d}},\ \forall d, \tag{15}$$

即

$$f^{\hat{d}}+\boldsymbol{P}^{\hat{d}}g^{\hat{d}}=\max_d\{f^d+\boldsymbol{P}^d g^{\hat{d}}\}. \tag{16}$$

利用引理 2, 可以设计求解最优投资策略的迭代算法, 其思路为, 从任意策略 d_0 出发, 在第 k 步迭代, 根据策略 d_k, 按如下方式构造改进策略:

$$d_{k+1}=\arg\max_d\{f^d+\boldsymbol{P}^d g^{d_k}\}. \tag{17}$$

如果 d_{k+1} 和 d_k 只在测度为零的集合上有差别, 则迭代停止. 由引理 2 可知, 在每步迭代中, 改进策略的资产增值率均增加, 定理 1 保证了算法停止时得到的策略是最优的.

3 基于样本路径的策略迭代算法

根据式 (17), 在每一步迭代中, 需要求出泊松方程的解 g^{d_k}. 由于投资组合优化问题的非线性性质, 求解泊松方程的闭式解并不容易, 需要利用样本路径上以往的交易和资产价格变化信息, 由式 (12) 估计得到性能势. 除了性能势外, 基于样本路径上的历史信息还可以估计出转移函数 P^d 和单阶段资产增值率 f^d. 换言之, 即使不知道资产的增值率和波动性参数, 仍可以采用基于样本路径的策略迭代算法. 本节将详细讨论基于样本路径的策略迭代算法的设计和实现.

在时刻 t, 记各支资产的价格为 $p(t)=(p_0(t), p_1(t),\cdots,p_N(t))^{\mathrm{T}}$. 为了一致性, 将无风险资产看作波动性为 0 的风险资产, 即 $\sigma_0=0$. 对于单阶段资产增值率 f^d, 根据式 (8), 如果资产的增值率 r_i 和波动性 σ_i 已知, 则可以采用数值积分的方法直接计算 f^d; 如果资产的参数未知, 则需要从样本路径上估计得到 f^d. 传统方法首先估计 r_i 和 σ_i, 然后计算 f^d. 然而, 这种方法只适用于风险资产满足高斯模型的情况. 本节采用更一般的对风险资产模型依赖较少的方法, 从样本路径上的交易和价格变化历史直接估计 f^d, 有

$$f^d(x)=\frac{\displaystyle\sum_{t=0}^{T-1}[I(x|X(t))f^d(X(t),p(t),p(t+1))]}{\displaystyle\sum_{t=0}^{T-1}I(x|X(t))}. \tag{18}$$

其中: $I(x|X(t))$ 为示性函数, $f^d(X(t),p(t),p(t+1))$ 定义为

$$\begin{aligned}&f^d(X(t),p(t),p(t+1))=\\&\ln\Big[\frac{p_0(t+1)}{p_0(t)}\Big(s_0(t)-\sum_{i=1}^N(1+\lambda_i)u_i(t)+\\&\sum_{i=1}^N(1-\mu_i)v_i(t)\Big)+\sum_{i=1}^N\frac{p_i(t+1)}{p_i(t)}(s_i(t)+\\&u_i(t)-v_i(t))\Big].\end{aligned} \tag{19}$$

在策略迭代算法中, 最重要的一步是估计性能势 g^{d_k}. 因为资产组合问题是连续状态空间问题, 所以只能先估计有限个状态的性能势, 然后利用它们拟合其他状态的性能势. 令 $\boldsymbol{S}_M=\{X_1,X_2,\cdots,X_M\}\subset\boldsymbol{S}$ 表示状态空间中由 M 个不同状态组成的子集, 称为基集, $X_m\in\boldsymbol{S}_M$ 为基状态, $m=1,2,\cdots,M$. 根据性能势定义 (12), 对于基状态 $X_m\in\boldsymbol{S}_M$, 其性能势可以利用下式估计:

$$g^d(X_m)\approx\frac{\displaystyle\sum_{t=0}^{T-L+1}\Big[\rho(X_m,X(t))\sum_{l=0}^{L-1}f^d(X(t+l))\Big]}{\displaystyle\sum_{t=0}^{T-L+1}\rho(X_m,X(t))}, \tag{20}$$

其中 $\rho(X_m,X(t))$ 为权重系数. 事实上, 式 (20) 估计了

泊松方程的某个特解, 它满足规范化条件

$$g^d(x)=\lim_{T\to\infty}E\Big\{\sum_{t=0}^{T-1}f^d(X(t))\Big|X(0)=x\Big\}.$$

式 (20) 可以看作多种方法的通用表达式. 如果将状态空间离散为 M 个子集, 然后在每个子集中选择一个代表状态作为基状态 X_m, 并将示性函数作为权重函数, 即 $\rho(X_m,X(t))=1$ 当且仅当 X_m 和 $X(t)$ 在同一子集中, 则式 (20) 可以看作是离散化方法. 基状态的选择会影响策略迭代算法的收敛速度, 理论上, 基状态个数越多, 算法收敛越快; 同时, 基状态在空间分布比较均匀合理时, 算法的收敛也会较快. 本文选取样本路径上的 M 个不同状态作为基状态, 为了操作简单, 可以选择随机投资策略下的样本路径的前 M 个状态, 这样产生的基状态空间分布比较均匀. 同时, 本文选择与距离相关的权重函数进行拟合, 如 $\rho(X_m,X(t))=\|X_m,X(t)\|^2$, 那么, 根据基状态的性能势, 其他非基状态的性能势可以按下式进行估计:

$$g^d(x)\approx\sum_{m=1}^{M}\kappa(X_m,x)g^d(X_m),\ x\in\boldsymbol{S}-\boldsymbol{S}_M,\qquad(21)$$

其中 $\kappa(X_m,x)$ 为权重系数, 满足 $\sum_{m=1}^{M}\kappa(X_m,x)=1$.

最后, 需要估计 $\boldsymbol{P}^dg^d$. 由于资产组合问题的非线性性质, 估计状态转移概率并计算 $\boldsymbol{P}^dg^d$ 并不容易, 但是可以将 $\boldsymbol{P}^dg^d$ 作为整体直接估计. 根据定义

$$\boldsymbol{P}^dg^d(x)=\int_{\boldsymbol{S}}g^d(y)P^d(\mathrm{d}y|x)=E[g^d(X(t{+}1))|X(t)=x],\qquad(22)$$

利用 $g^d(X(t+1))$ 作为 $\boldsymbol{P}^dg^d(X(t))$ 的估计式, 得到

$$\boldsymbol{P}^dg^d(x)\approx\frac{\sum_{t=0}^{T-1}[I(x|X(t))g^d(X(t+1))]}{\sum_{t=0}^{T-1}I(x|X(t))},\qquad(23)$$

其中 $I(x|X(t))$ 为示性函数. 类似地, 将式 (23) 中的策略 d 换为 h 可以估计得到 $\boldsymbol{P}^hg^d$.

根据上述各个估计式, 基于样本路径的策略迭代算法如下.

算法 1　基于样本路径的策略迭代算法.

Step 1: 初始化. 选择算法中的参数, 包括 T、M 等, 任选一个投资策略 d_0, 并令 $k=0$.

Step 2: 策略评价. 由策略 d_k 进行交易, 得到样本路径 $\{X^k(0),X^k(1),\cdots,X^k(T-1)\}$ 和价格变化历史 $\{p^k(0),p^k(1),\cdots,p^k(T-1)\}$. 基于历史信息和式 (18)、(20)、(21) 和 (23), 估计式 (16) 中的各个参量.

Step 3: 策略改进. 根据式 (16) 或 (17) 构造改进策略 d_{k+1}. 如果 $d_{k+1}=d_k$ (或满足其他停止条件), 则算法停止, 最优投资策略为 d_k; 否则, 令 $k=k+1$, 返回 Step 2.

文献 [9] 证明了机器学习算法的收敛性. 在实际应用中, 由于归一化后状态和行动变量都在 [0,1] 之间, 可以采取离散化的方法将状态与行动离散化为有限个值, 这样根据实际可用样本的数量选择离散化的粗细程度, 使得算法的收敛满足要求. 下面通过一个实例验证基于样本路径的策略迭代算法的效果.

例 1　考虑只包含一支风险资产和一支无风险资产的投资组合, 无风险资产的增值率为 r_0, 风险资产的增值率为 r_1, 波动率为 σ_1. 由于含有手续费的离散时间投资组合优化问题的最优解不易求得, 这里使用连续时间问题的最优解进行比较.

文献 [3] 已经证明, 在满足 $|r_1-r_0|<\sigma_1^2/2$ 的一般情况下, 含有比例型手续费的连续时间投资组合优化问题的最优策略是采用最少的交易操作保持风险资产的比例位于特定的区间; 在满足 $|r_1-r_0|\geqslant\sigma_1^2/2$ 的特殊情况下, 最优策略是全部投资于风险资产. 本例中取 $r_0=6\%$, $r_1=12\%$, $\sigma_1=0.6$. 根据上述结论, 当没有交易手续费时, 连续时间情况的最优策略是保持风险资产的比例为

$$s_1^*=(r_1-r_0)/\sigma_1^2+1/2\approx0.67;$$

当交易费用存在时, 最优策略是保持风险资产的比例位于以 $s_1^*\approx0.67$ 为中心的特定区间内, 并且随着交易费用的增加, 区间也会扩大.

采用基于样本路径的策略迭代算法对以下 3 种不同手续费的情况进行求解: 1) $\lambda_1=\mu_1=0$, 即无手续费的情况; 2) 手续费比例为 $\lambda_1=\mu_1=0.01$; 3) 手续费比例为 $\lambda_1=\mu_1=0.03$.

以每周作为交易周期, 在策略迭代算法中, 首先任意选择初始投资策略, 如随机策略. 在初始随机策略下, 在每个交易周期开始时, 随机买卖一定数量的风险资产. 经过一定时间, 根据已有的交易、历史价格的变化信息、式(18)、(20)、(21) 和 (23), 计算出策略迭代算法中所需要的相关参量, 并构造更好的改进策略. 在改进策略下, 重复上述过程, 进行策略迭代. 经过 3 次迭代, 得到如图 1 所示的策略.

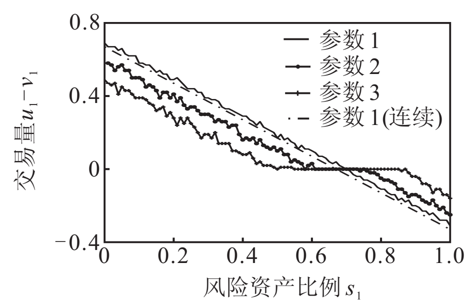

图 1　**基于样本路径的策略迭代算法结果**

图 1 中, x 轴为风险资产占总资产的比例, 即状

态变量 $X = s_1$; y 轴为不同状态下所采取的交易, 以 $u_1 - v_1$ 的形式表示. 即在某个状态:

1) 若 $u_1 - v_1 > 0$, 则买入风险资产, 数量为 $|u_1 - v_1|$;

2) 若 $u_1 - v_1 < 0$, 则卖出风险资产, 数量为 $|u_1 - v_1|$;

3) 若 $u_1 - v_1 = 0$, 则不进行任何交易.

对于第 1 种情况 ($\lambda_1 = \mu_1 = 0$), 连续时间情况下的最优策略应该是一条经过点 $(0, 0.67)$、$(0.67, 0)$ 和 $(1, -0.33)$ 的直线. 由图 1 可见, 基于样本路径的策略迭代算法给出的策略与连续时间情况下的最优策略非常接近. 其中的差别主要源于本文考虑的是离散时间情况的问题. 对于其他两种情况, 迭代算法给出的策略形式是合理的, 与连续时间情况下的策略形式相符, 其中一些误差主要来自估计相关参量时的误差. 即使这些误差存在, 对策略性能的影响也比较微弱, 这是因为在上述形式的策略下, 系统的稳定状态以大概率位于以 0.67 为中心的无交易区间, 上述误差只是影响过渡状态的策略性能, 对于稳定状态下的系统性能影响比较微弱.

4 结 论

本文采用基于性能灵敏度分析的优化方法分析和求解金融领域的动态投资组合优化问题, 推导出不同投资策略下资产长期平均增值率的差分公式, 以此为基础证明了最优性条件, 并开发出可以在线应用的基于样本路径的策略迭代优化算法. 与传统基于动态规划的方法相比, 基于灵敏度分析的方法具有以下特点: 1) 分析简洁直观, 对投资策略收益的比较、最优性方程的推导、策略迭代算法的设计都是基于投资策略增值率的差分公式; 2) 性能势等核心概念都是用样本路径定义, 因此可以直接开发基于样本路径的在线算法, 即根据历史投资信息对投资策略进行学习优化; 3) 所提出的方法适用性较强, 在资产特性参数未知或模型结构变化不大 (满足马尔可夫性) 时, 仍然能够应用. 例如, 考虑风险资产增值率存在跳变的情况, 如果跳变具有马尔可夫性, 则本文分析框架和方法仍然适用. 如何挖掘问题特点, 设计更加高效的在线算法是未来的研究重点.

参考文献(References)

[1] Merton R C. Lifetime portfolio selection under uncertainty: The continuous-time case[J]. Review of Economics and Statistics, 1969, 51(3): 247-257.

[2] Samuelson P. A lifetime portfolio selection by dynamic stochastic programming[J]. Review of Economics and Statistics, 1969, 51(3): 239-246.

[3] Taksar M, Klass M J, Assaf D. A diffusion model for optimal portfolio selection in the presence of brokerage fees[J]. Mathematics of Operations Research, 1988, 13(2): 277-294.

[4] Akian N, Sulem A, Taksar M. Dynamic optimization of long term growth rate for a portfolio with transaction costs and logarithmic utility[J]. Mathematical Finance, 2001, 11(2): 153-188.

[5] Muthuraman K, Zha H. Simulation-based portfolio optimization for large portfolios with transaction costs[J]. Mathematical Finance, 2008, 18(1): 115-134.

[6] Davis M H A, Norman A R. Portfolio selection with transaction costs[J]. Mathematics of Operations Research, 1990, 15(4): 676-713.

[7] Cover T M. Universal portfolios[J]. Mathematical Finance, 1991, 1(1): 1-29.

[8] Ormoneit D, Glynn P. Kernel-based reinforcement learning in average-cost problems[J]. IEEE Trans on Automatic Control, 2002, 47(10): 1624-1636.

[9] Cao X R. Stochastic learning and optimization: A sensitivity-based approach[M]. New York: Springer, 2007: 341-380.

(责任编辑: 郑晓蕾)