arXiv:2611.29552v1 [q-fin.PM] 22 Nov 2026

适用于含交易成本在线投资组合选择的门控收益加权净利润最大化

GRACE CHENGRONG XIE、ALVEY QIANLI LIN 和 WAI KI CHING

摘要。本文研究在线投资组合选择（OLPS）问题。基于自适应在线移动平均（AOLMA）的净利润最大化（NPM），即投资组合选择中的 AOLNPM 策略，会在预测收益超过交易成本时进行交易。从本质上说，这种方法高度依赖校准预测。事实上，我们表明，AOLMA 预测会高估实现的相对收益。因此，我们设计了一种门控机制，用于调整 AOLMA 预测的使用力度；对于每项资产，该机制根据相对于中性预测（不变）的预测似然，学习近期分配给 AOLMA 的混合比例。因此，该比例估计了预测实现的程度。随后，我们将所得预测输入 NPM 投资组合规则，其中门控机制充当针对特定资产的成本阈值。收益加权基金按财富比例组合采用固定衰减因子的门控 NPM 专家，且无需学习率。在交易成本率为 0。1% 的九个数据集上，我们的方法的几何平均财富是 AOLNPM 的 5。9 倍，并且在每个数据集上的表现均优于 AOLNPM；随着成本率上升，其优势也随之扩大。

1. 引言

在线投资组合选择（OLPS）是学术研究者和从业者都十分关注且具有重要意义的研究问题。在经典投资组合选择理论中， [17], 假设风险资产的收益服从某个分布函数。在这一假设下，可以计算收益的期望值和方差，从而得到预期收益和风险。随后，投资者将资金分配到不同资产中，以获取更高的投资收益或规避风险。在线投资组合选择与经典投资组合选择理论不同，它不对收益分布作任何假设，同时也是一种数据驱动的序列决策过程。在线投资组合选择仅利用历史价格，定期调整资产投资组合。文献中提出了许多不同的策略，例如， [8, 12, 14]。

自适应在线净利润最大化（AOLNPM） [7] 直接处理交易成本，并进一步结合状态依赖型价格估计器进行了扩展 [9]。该方法采用自适应在线移动平均（AOLMA）来预测每个价格相对值，并通过 NPM 选择投资组合，使扣除交易成本后的预测收益最大化。NPM 的解形式简单，且标准明确：卖出所有预测收益比最佳预测低出超过指定成本率的持仓 c，并将投资组合转移至最佳资产（命题 2.1）。然而，这条规则将预测结果视为确凿事实。只有当预期收益差超过交易成本时，从一种资产切换到另一种资产才会提高预期净收益；否则只会导致更多

日期：2026年11月24日。

不必要的交易变动。如果实际实现的预测差距仅占 β 预测差距的一部分，就应相应调整阈值为 c/β （命题 2.2）。我们的测试表明，AOLMA 确实存在这些局限。在四个基准数据集上， [7],预测值最高的资产相对于平均资产的预测超额收益实际仅为其预测值的14–30% 。在五个附加数据集上，该比例至多为7%（表 1）。因此，NPM 会针对不足以覆盖成本的差距进行交易。尽管 AOLNPM 会根据预测误差调整其衰减因子，但调整步长最终较小，且因数据而异。在我们的实验中，最终衰减因子介于0.51 和 0.72。

基于上述观察，本文提出两项改进措施。第一项是引入一个门控机制，用于调整对 AOLMA 预测的采纳程度。该方法根据每项资产相对于中性预测的预测似然，学习近期分配给 AOLMA 预测的混合比例。随后，将所得预测输入 NPM 投资组合规则。门控机制利用该比例估计预测实际实现的份额，从而在线校准每项预测（命题 4.1）。在 NPM 中，门控机制随后作为资产特定阈值，依据各资产信号的可靠性对资产重新排序（命题 4.2）。第二项改进是以收益加权基金取代自适应衰减因子。我们在一个网格上采用固定衰减因子运行门控 NPM 专家，并根据其赚取的财富为其分配权重。因此，该基金无需学习率。在独立账户中，最佳专家的财富至少有一个固定比例得以保留（命题 5.1）。我们将此方法称为门控收益加权 NPM（G-PWNPM）。

我们的数值实验在九个数据集上进行，交易成本率为 $c = 0.001$。就几何平均值而言，G-PWNPM 的终端财富最终达到 AOLNPM 的 5.92 倍，并且在全部九个数据集上均优于 AOLNPM，合并统计量为 $z = 4.68$。在其中四个基准数据集上，G-PWNPM 优于全部十三种基线方法。此外，在预测信息量较少的五个附加数据集上，它的表现优于 AOLNPM，但在基线方法中处于中游。我们的一项发现是，随着成本率提高，本方法相较于 AOLNPM 的优势也随之扩大，而且本方法的平均单向换手率更低。第一项改动，即门控，主要通过重新排序资产带来了大部分收益；而基金则确保了本方法在不同数据集上的表现一致。

本文结构如下：第 2 节回顾 AOLMA、NPM 和 AOLNPM 框架，并说明校准条件。第 3 节介绍两个预测模型及其似然。第 4 节介绍自适应混合权重，第 5节讨论收益加权基金。数值实验见第 6 节，第7 节总结全文。

2. 预备知识

2.1. 含交易成本的在线投资组合选择。 设 n 为资产数量， P_{ti} 为资产 i 在时期 t末的价格。其价格相对值为 $r_{ti} = P_{ti}/P_{t-1,\ i} > 0$。设 $\mathbf{b}_t$ 为所选资产权重向量，$\mathbf{h}_t$ 为从上一时期延续下来的交易前权重，且 c 为交易成本率。可行集为

$$\Delta_n = \{\mathbf{b} \geq 0 : \sum_{i=1}^{n} b_i = 1\}.$$

权重 $\mathbf{b}_t$ 是在观测到 $\mathbf{r}_t = (r_{t1}, \ldots, r_{tn})^\top$ 之前选定的，财富按照以下方式演变

$$W_t = W_{t-1}\left(\mathbf{b}_t^\top \mathbf{r}_t - \frac{c}{2}\|\mathbf{b}_t - \mathbf{h}_t\|_1\right), \qquad W_0 = 1. \tag{2.1}$$

这里 W_t 表示时期 t之后的财富，而 $\|\mathbf{b}_t - \mathbf{h}_t\|_1/2$ 表示单向换手率。每单位财富的买入或卖出成本为 $c/2$，因此将一单位财富从一种资产转移到另一种资产的成本为 c；这是 [7, 14] 的惯例。观察到收益后，下一个交易前的权重为$h_{t+1,\,i} = b_{ti} r_{ti}/(\mathbf{b}_t^\top \mathbf{r}_t)$，且 $\mathbf{h}_1 = \mathbf{1}/n$。

2.2. AOLMA 预测。 在线移动平均法通过对资产 i 的历史价格进行指数平均，来预测其下一期价格，

$$\widehat{P}_{ti} = \alpha P_{t-1,i} + (1-\alpha)\widehat{P}_{t-1,i}$$

^衰减因子为 $\alpha \in [0,1]$，其价格相对值为 $f_{ti} = P_{ti}/P_{t-1,\,i}$ [13]。将其除以$P_{t-1,\,i}$ 并利用

$$\widehat{P}_{t-1,i}/P_{t-1,i} = f_{t-1,i}/r_{t-1,i}$$

由此可得

$$y_{ti} = \frac{f_{t-1,i}}{r_{t-1,i}} \qquad \text{where} \qquad f_{ti} = \alpha + (1-\alpha)y_{ti}. \tag{2.2}$$

相对于最新已知价格表示的先前价格估计为 y_{ti}。展开（2.2），f_{ti} 是历史价格除以最新价格后的加权平均，因此当价格低于近期均值时，$f_{ti} > 1$。对收益作一阶近似，

$$f_{ti} - 1 \approx -\sum_{k\geq 1}(1-\alpha)^k (r_{t-k,i} - 1),$$

较小的 α 会带来较长的记忆和较大的反转信号，而 $\alpha \to 1$ 会给出中性预测 $\mathbf{1}$。

[7] 中提出的自适应在线移动平均（AOLMA）允许每个资产拥有自己的衰减因子 α_{ti}，初始值为 $\alpha_{1i} = 1/2$。在观测到 r_{ti} 之后，

$$\alpha_{t+1,i} = \alpha_{ti} + \gamma s_{ti}, \qquad s_{ti} = \begin{cases} 1, & \text{if } \mathbb{I}\{r_{ti} > f_{ti}\} = \mathbb{I}\{r_{t-1,i} > f_{t-1,i}\}, \\ -1, & \text{otherwise}, \end{cases} \tag{2.3}$$

其中，当 A 成立时，$\mathbb{I}\{A\}$ 为 $\mathbf{1}$，否则为 $\mathbf{0}$；如果 $\alpha_{t+1,i}$ 离开 $[0,1]$，则将其重置为 $1/2$。因此，当且仅当最近两个预测误差均为正或均为非正时，$s_{ti} = 1$。由于 $\partial f_{ti}/\partial \alpha_{ti} = 1 - y_{ti}$ 的符号与 r_{t-1}相同，$_i - f_{t-1,i}$，所以在两个预测误差均不为零时，(2.3) 是平方预测误差（$r_{ti} - f_{ti}$）上的带符号梯度步；这些情况对应于第 3 节的四种情形 [7,。步长 γ 针对每个数据集分别设定。

2.3. 净利润最大化。 给定预测 $\mathbf{f}_t = (f_{t1}, \ldots, f_{tn})^\top$，NPM 规则为

$$\mathbf{b}_t \in \arg\max_{\mathbf{b}\in\Delta_n}\left\{\mathbf{b}^\top \mathbf{f}_t - \frac{c}{2}\|\mathbf{b} - \mathbf{h}_t\|_1\right\}. \tag{2.4}$$

文献 [7] 提出的自适应在线净利润最大化（AOLNPM）策略，将2.4应用于 AOLMA 预测，并将其作为线性规划求解。该规划有一个闭式解。

命题 2.1。 设 $j^* \in \arg\max_i f_{ti}$ 且 $S = \{i \neq j^* : \quad f_{tj^*} - f_{ti} > c\}$。 (2.4) 的最优解为 $b_{ti} = 0$ 当 $i \in S$时， $b_{ti} = h_{ti}$ 当 $i \notin S \cup \{j^*\}$时，以及 $b_{tj^*} = h_{tj^*} + \sum_{l \in S} h_{tl}$。

证明。 令 $\mu = f_{tj^*} - c/2$ 和 $L(\mathbf{b}) = \mu + \sum_i \phi_i(b_i)$， 其中

$$\phi_i(b_i) = b_i(f_{ti} - \mu) - \frac{c}{2}|b_i - h_{ti}|.$$

在 Δ_n 上， L 与 (2.4) 的目标函数一致。每个 ϕ_i 都是凹的分段线性函数，斜率为 $f_{ti} - f_{tj^*} + c$ （当 $b_i < h_{ti}$ 时），为 $f_{ti} - f_{tj^*}$ （当 $b_i > h_{ti}$时）。因此，每个$b_{j^*} \geq h_{tj^*}$都能使 ϕ_{j^*} 达到最大值；对于 $i \neq j^*$，当 $f_{tj^*} - f_{ti} > c$ 时，ϕ_i 在 $b_i = 0$ 处达到最大值，否则在 $b_i = h_{ti}$ 处达到最大值。该命题中的向量使每个 ϕ_i 在 $b_i \geq 0$ 上达到最大值，且属于 Δ_n，因此它使L 在非负正交象限上达到最大值，并使式 (2.4) 中的目标函数在 Δ_n 上达到最大值。 □

简言之，NPM 会卖出所有预测值比最佳资产低出超过阈值 c 的持仓，并将更多投资组合配置到最佳资产上。将财富的 δ 从资产 i 转移到资产 j，会使目标函数改变 $\delta($ $f_{tj} - f_{ti} - c)$，因此阈值 c 就是这笔转移的成本。

2.4. NPM 阈值校准。

命题 2.2. 假设在给定过去信息的条件下， $\mathbb{E}[r_{tj} - r_{ti}] = \beta(f_{tj} - f_{ti})$ 对某个 $\beta > 0$成立。那么，当且仅当$f_{tj} - f_{ti} > c/\beta$时，将财富从资产 i 转移到资产 j 会提高预期净收益。

证明。 转移 δ 的财富，预期收益为 $\beta\delta(f_{tj} - f_{ti})$，成本为 δc。 □

因此，在命题 2.2中的比例校准假设下，预测差距的阈值为 c/β，当 $\beta = 1$时，它简化为 NPM阈值 c 。表 1估计了 β 在 $\alpha = 1/2$条件下 AOLMA 预测的情况：预测值最高的资产，其实际实现的相对于平均资产的预测超额收益仅占预测值的一小部分。这可能导致 NPM 根据预期收益不足以覆盖成本的预测差距进行交易。命题 2.2提出了两种补救办法：将阈值提高到约 c/β，或将每项预测重新缩放为 $1 + \beta_i(f_{ti} - 1)$，并采用资产特定的 β_i ，同时保持阈值为 c。第 3 节和第 4 节中的门控采用第二种办法，并在线学习 β_i 。

3. 两个预测模型及其似然

在时期 t之前，AOLMA预测器会提供背景预测 f_{ti}^B，对应 r_{ti}。上标 B 用于标识此背景预测。本节和下一节中的背景预测来自 AOLNPM [7]，第 5 节中的背景预测则来自固定衰减专家。其衰减因子为 $\alpha_{ti}^B \in [0, 1]$。相对于最新已知价格表示的先前价格估计为 y_{ti}^B。递推关系为

$$y_{ti}^B = \frac{f_{t-1,i}^B}{r_{t-1,i}}, \qquad f_{ti}^B = \alpha_{ti}^B + (1 - \alpha_{ti}^B)y_{ti}^B. \tag{3.1}$$

我们比较两个预测模型。模型1使用 f_{ti}^B。模型0使用中性预测1，对应于预期简单收益率为零。设 $w_{ti} \in [0, 1]$ 为交易前分配给模型1的权重。组合预测为

$$\widehat{f}_{ti} = w_{ti}f_{ti}^B + (1 - w_{ti}) \cdot 1 = 1 + w_{ti}(f_{ti}^B - 1). \tag{3.2}$$

数据集	预测超额收益	实现超额收益	β
TSE	0.0437	0.0064	0.15
MSCI	0.0131	0.0033	0.26
NYSE-N	0.0226	0.0031	0.14
NYSE-O	0.0243	0.0074	0.30
DJIA	0.0291	0.0010	0.04
标准普尔500指数	0.0266	0.0019	0.07
SZ50	0.0153	−0.0009	−0.06
COVID	0.0192	−0.0006	−0.03
new-DJIA	0.0170	0.0001	0.00

表1。使用 $\alpha = 1/2$ 对 AOLMA 预测进行校准。预测超额收益：预测值最高的资产的预测值减去横截面平均预测值。实现超额收益：该资产的价格相对值减去横截面平均值。两者均按周期取平均；β 是它们的比值。

因此，$w_{ti} = 1$ 保留 AOLMA 预测，并且 $w_{ti} = 0$ 给出中性预测。组合预测代入 NPM 规则（2.4）。

为了使用观测收益评估每个模型，令隐含标签 $Z_{ti} \in \{0,1\}$ 标识其成分分布。记 $f_{1ti} = f_{ti}^{B}$ 和 $f_{0ti} = 1$。令 $v_{ti} > 0$ 为对数收益的共享方差代理量。其平方根表示单期噪声尺度。条件工作模型为

$$\log r_{ti} \mid Z_{ti} = k \sim N\left(\log f_{kti} - \frac{v_{ti}}{2}, v_{ti}\right), \qquad k \in \{0,1\}, \tag{3.3}$$

其均值为 $\log f_{kti} - v_{ti}/2$，方差为 v_{ti}。修正项 $-v_{ti}/2$ 使平均价格相对值等于 f_{kti}。我们通过指数平均估计对数收益的近期二阶矩：

$$v_{t+1,i} = \rho_v v_{ti} + (1-\rho_v)(\log r_{ti})^2, \qquad \rho_v = 0.98. \tag{3.4}$$

每次更新都将先前的估计值与新观测到的对数收益平方结合起来。经过反复更新，较早的观测值会获得指数递减的权重。我们在 (3.3) 中采用这一二阶矩作为通用的方差代理。这里的恒等式 $E[X^2] = \mathrm{Var}(X) + (E[X])^2$ 说明了这一近似：当局部均值较小时，二阶矩趋近于方差。前六个观测值用于初始化 v，并设置数值下限10^{-12}。

对于模型 k，对数收益服从由 (3.3) 指定的高斯密度。将变量从 $\log r_{ti}$ 变换为 r_{ti} 会引入因子 $1/r_{ti}$。在观测到的收益处计算所得密度，即可得到模型似然：

$$\ell_{kti} = \frac{1}{r_{ti}\sqrt{2\pi v_{ti}}} \exp\left[-\frac{(\log r_{ti} - \log f_{kti} + v_{ti}/2)^2}{2v_{ti}}\right]. \tag{3.5}$$

更高的 ℓ_{kti} 意味着模型会为接近观测值的收益赋予更高的概率密度。两个模型使用相同的噪声尺度，因此下文只需关注两个似然的比值。

对于 $x_{ti} = \log r_{ti}$ 和 $d_{ti} = \log f^B_{ti}$,

$$u_{ti} = \log \frac{\ell_{1ti}}{\ell_{0ti}} = \frac{(x_{ti} + v_{ti}/2)^2 - (x_{ti} - d_{ti} + v_{ti}/2)^2}{2v_{ti}} = \frac{d_{ti}(x_{ti} + v_{ti}/2) - d_{ti}^2/2}{v_{ti}}. \tag{3.6}$$

对数似然比是两个模型的平方误差之差，并按方差代理进行缩放；当 AOLMA 预测对观测的描述更准确时，该比值为正。工作模型仅用于为这两个预测评分：它决定 w_{ti}，而投资组合由 NPM 选择。每日对数收益的尾部比（3.3）更厚；第 6.6节将高斯密度替换为 Student-t 密度。

4. 自适应混合权重

对于每项资产 i，我们会随时间追踪其隐藏模型标签 Z_{ti}。标签 1 表示 AOLMA 预测分布，标签 0 表示（3.3）中的中性预测分布。每项资产都有各自的混合比例和软计数。基准策略比例 π_{ti}概括了累积的预测证据。

4.1. 学习基准混合比例。 在时期 t之前，令 M_{1ti} 和 M_{0ti} 分别表示资产 i中分配给 AOLMA 模型和中性模型的折扣软计数。它们的相对大小决定了基准 AOLMA 比例:

$$\pi_{ti} = \frac{M_{1ti}}{M_{1ti} + M_{0ti}}. \tag{4.1}$$

令 t_0 表示其收益首次纳入门控校准的时期。在观察到该收益之前，初始化 $M_{1,\,t_0,\,i} = M_{0,\,t_0,\,i} = 1$，从而得到 $\pi_{t_0,\,i} = 1/2$。这些相等的伪计数为两个模型赋予相等的初始权重。校准日期见第 6.1节。

观察到 r_{ti}后，使用公式（3.5）计算两个似然 ℓ_{1ti} 和 ℓ_{0ti}。在基准策略混合下，分配给模型 1 的后验概率为

$$z^{\mathrm{M}}_{ti} = \frac{\pi_{ti}\ell_{1ti}}{\pi_{ti}\ell_{1ti} + (1 - \pi_{ti})\ell_{0ti}} = \frac{\pi_{ti}e^{u_{ti}}}{\pi_{ti}e^{u_{ti}} + 1 - \pi_{ti}}. \tag{4.2}$$

上标 M 表示自适应混合。该后验概率也称为模型对该观测的责任度。该观测使 AOLMA 计数增加 z^{M}_{ti}，使中性计数增加 $1 - z^{\mathrm{M}}_{ti}$。给定遗忘因子 $\rho_M = 0.98$，更新

$$\begin{aligned} M_{1,t+1,i} &= \rho_M M_{1ti} + z^{\mathrm{M}}_{ti}, \\ M_{0,t+1,i} &= \rho_M M_{0ti} + 1 - z^{\mathrm{M}}_{ti}. \end{aligned} \tag{4.3}$$

这里将后验责任度视为分数计数，正如在线混合估计中所采用的 [4]。在时期 t之前，在时期 $s < t$ 计算出的责任度剩余权重为 ρ_M^{t-1-s}。每个初始伪计数的剩余权重为 $\rho_M^{t-t_0}$。因此，π_{ti} 通过对过去观测进行折扣，随时间自适应；总计数趋于 $1/(1 - \rho_M) = 50$，且长期来看，每次观测使π_{ti} 的变化幅度至多约为 $1/50$。自适应混合门控在公式（3.2）中使用 $w_{ti} = \pi_{ti}$。动态模型平均同样使用预测似然和遗忘因子来更新模型权重 [19]。

4.2. **混合比例估计**。

命题 4.1。 假设给定过去信息后，对数收益 x_{ti} 服从均值为$\beta d_{ti} - v_{ti}/2$ 、方差为 v_{ti}的高斯分布。设 $\varepsilon_{ti} = d_{ti}/\sqrt{v_{ti}}$ 为信噪比，且 $n_{ti} = M_{1ti} + M_{0ti}$。那么，我们有

$$\mathbb{E}[\pi_{t+1,i} - \pi_{ti} \mid past] = \frac{\pi_{ti}(1-\pi_{ti})(\beta - \pi_{ti})}{\rho_M n_{ti} + 1}\,\varepsilon_{ti}^2 + O(\varepsilon_{ti}^4).$$

在首阶近似下， π_{ti} 因而趋向于 β （若 $\beta \le 0$ 则趋向0 ，若 $\beta \ge 1$则趋向1 ）。在$\pi_{ti} = \beta$ 时，组合预测 $\widehat{f}_{ti} = 1 + \beta(f_{ti}^B - 1)$ 等于过去的 $\mathbb{E}[r_{ti} \mid] = (f_{ti}^B)^\beta$，误差至多为 $O(d_{ti}^2)$。

证明见附录 A。因此，自适应混合是对已实现预测份额进行的在线资产特定估计，而公式（3.2）据此重新缩放每个预测：这是第 2 节提出的第二种补救方法。由于每个责任值都落在 $[0,1],$ ，单个极端收益至多使计数改变 1，而 π_{ti} 长期来看至多约为 $1/50$ 。

4.3. **NPM 规则中的组合预测**。 将背景递推关系（3.1）代入（3.2）， $\widehat{f}_{ti} = 1 + w_{ti}(1 - \alpha_{ti}^B)($ $y_{ti}^B - 1)$，可得

$$\widehat{f}_{ti} = \alpha_{ti}^{\text{eff}} + (1 - \alpha_{ti}^{\text{eff}})y_{ti}^B, \qquad \alpha_{ti}^{\text{eff}} = w_{ti}\alpha_{ti}^B + (1 - w_{ti}). \tag{4.4}$$

因此，门控通过其预测模型权重改变有效衰减因子。由于y_{ti}^B 仍遵循背景递推，门控保留了移动平均的记忆，只缩小其信号的幅度。

命题 4.2。 将 *NPM* 规则 (2.4) 应用于组合预测 (3.2).

(a) 若对所有 i 都有 $w_{ti} = w > 0$ ，则该解等于在$\mathbf{f}_t^B$上使用阈值 c/w 的 *NPM* 解。*(b)* 一般而言， j^*最大化 $w_{ti}(f_{ti}^B - 1)$；当且仅当 $w_{tj^*}(f_{tj^*}^B - 1) - w_{ti}(f_{ti}^B - 1) > c$ 时，资产 i 会被卖出。*(c)* 若对所有 i 都有 $w_{ti} = 0$ ，则 *NPM* 不进行交易。

证明。 (a) 由于 $\mathbf{b}^\mathsf{T}\mathbf{1} = 1$，目标函数等于

$$1 + w[\mathbf{b}^\mathsf{T}(\mathbf{f}_t^B - \mathbf{1}) - \frac{c}{2w}\|\mathbf{b} - \mathbf{h}_t\|_1],$$

且命题 2.1 适用于成本 c/w。(b) 和 (c) 由命题 2.1推出，因为$\widehat{f}_{tj} - \widehat{f}_{ti} = w_{tj}(f_{tj}^B - 1) - w_{ti}$ $(f_{ti}^B - 1)$。 □

因此，共同权重 w 会将阈值提高到 c/w，而资产特定权重还会改变哪项资产看起来最优：门控会根据反转信号的可靠性对资产重新排序。

5. 收益加权门控专家集成

AOLNPM 使用数据集特定步长，根据预测误差调整其衰减因子。我们则在网格上运行采用固定衰减因子的 NPM 专家，并根据它们所累积的财富为其分配权重。

5.1. **专家。** 令 $\mathscr{A} = \{0.05, 0.10, \ldots, 0.95\}$。对于每个 $\alpha \in \mathscr{A}$，专家 α 使用固定衰减因子 $\alpha_{ti}^{B} = \alpha$运行递推关系（3.1），采用第 4 节的自适应混合及其自身的软计数，并对组合预测 $\hat{\mathbf{f}}_t^{(\alpha)}$ 应用 NPM 规则。该专家保留自己的交易前权重 $\mathbf{h}_t^{(\alpha)}$ 和财富

$$S_t^{(\alpha)} = S_{t-1}^{(\alpha)} \left(\mathbf{b}_t^{(\alpha)\mathsf{T}} \mathbf{r}_t - \frac{c}{2} \|\mathbf{b}_t^{(\alpha)} - \mathbf{h}_t^{(\alpha)}\|_1 \right), \qquad S_0^{(\alpha)} = 1. \tag{5.1}$$

方差代理（3.4）仅取决于收益，并由各专家共享。

5.2. **基金。** 基金根据各专家迄今积累的财富为其分配权重，

$$\lambda_t^{(\alpha)} = \frac{S_{t-1}^{(\alpha)}}{\sum_{\alpha' \in \mathscr{A}} S_{t-1}^{(\alpha')}}, \qquad \mathbf{b}_t = \sum_{\alpha \in \mathscr{A}} \lambda_t^{(\alpha)} \mathbf{b}_t^{(\alpha)}, \tag{5.2}$$

并根据自身的交易前权重 $\mathbf{h}_t$进行交易，支付 $\frac{c}{2}\|\mathbf{b}_t - \mathbf{h}_t\|_1$，如（2.1）所示。不同专家方向相反的交易会相互抵消。权重（5.2）是专家的后验分布，其中对数财富发挥对数似然的作用，这与由有限个策略构成的通用投资组合相同 [2,5,6]；因此不需要学习率。我们将这一方法称为门控收益加权 NPM（G-PWNPM）。不使用门控（$w_{ti}^{(\alpha)} = 1$）时，它就是收益加权NPM（PWNPM）。

命题5.1。 设 $W_t^{\mathrm{sep}} = |\mathscr{A}|^{-1} \sum_\alpha S_t^{(\alpha)}$ 为一位投资者的财富，该投资者将初始财富平均分配给各专家，并将其置于独立账户中。那么

(i) 对数 $W_T^{\mathrm{sep}} \geq$ 最大值$_\alpha$ 对数 $S_T^{(\alpha)} -$ 对数 $|\mathscr{A}|$；***(ii)*** 在时期 t 该投资者持有 $\mathbf{b}_t$ 份 (5.2) 总计，交易前总权重$\sum_\alpha \lambda_t^{(\alpha)}$ $\mathbf{h}_t^{(\alpha)}$ ；***(iii)*** 仅执行各账户交易的净额，其成本至多为各账户单独成本之和。

证明。**(i)** 非负项之和至少不小于其中最大的一项。
(ii) 账户 α 投资 $|\mathscr{A}|^{-1} S_{t-1}^{(\alpha)} \mathbf{b}_t^{(\alpha)}$；在上一时期结束后，其持有 $|\mathscr{A}|^{-1} S_{t-1}^{(\alpha)} \mathbf{h}_t^{(\alpha)}$，因为成本按持仓比例扣除。除以 $\sum_{\alpha'} |\mathscr{A}|^{-1} S^{(\alpha')}{}_{t-1}$ 即可得到这两个结论。**(iii)**

$$\frac{c}{2} \| \sum_\alpha \lambda_t^{(\alpha)} (\mathbf{b}_t^{(\alpha)} - \mathbf{h}_t^{(\alpha)}) \|_1 \leq \frac{c}{2} \sum_\alpha \lambda_t^{(\alpha)} \|\mathbf{b}_t^{(\alpha)} - \mathbf{h}_t^{(\alpha)}\|_1.$$

□

基金对独立账户的交易进行净额结算。基金自身的交易前权重与 **(ii)** 中的总权重之间的差异，仅来自专家在上一时期产生的不同成本，因此 **(i)** 对基金而言在 c 阶的项以内成立；在我们的实验中，净额结算后的基金在每个数据集上的表现都优于独立账户。

5.3. **算法。** 算法 1 总结了 G-PWNPM。每次决策仅使用 $r_1, \ldots, r_{t-1}$。每个时期的成本为 $O(|\mathscr{A}|\, n)$ 次运算，而我们的 Python 实现在每个数据集上的运行时间均不到一秒。

算法1 门控收益加权 NPM（G-PWNPM）

要求： 成本率 c；网格 $\mathscr{A}$；$\rho_v = \rho_M = 0.98$；时期 t_0 的校准结果；使用的门控来自时期 t_g

$S^{(\alpha)} \leftarrow 1$, $\mathbf{h}^{(\alpha)} \leftarrow \mathbf{1}/n$, $M_{1i}^{(\alpha)} = M_{0i}^{(\alpha)} \leftarrow 1$) 对所有 α, i; $\mathbf{h} \leftarrow \mathbf{1}/n$, $W \leftarrow 1$

对于 $t = 1, \ldots, T$ **do**

  $f_{ti}^{(\alpha)} \leftarrow$ 递推关系（3.1），衰减 α（前六个周期中的预测 1）

  $w_{ti}^{(\alpha)} \leftarrow M_{1i}^{(\alpha)} / (M_{(\alpha)1i} + M_{0i}^{(\alpha)})$，如果 $t \geq t_g$，否则为 $\mathbf{1}$；$\hat{f}_{ti}^{(\alpha)} \leftarrow 1 + w_{ti}^{(\alpha))} \left(f_{(\alpha)ti} - 1\right)$

  $\mathbf{b}_t^{(\alpha)} \leftarrow$ 命题 2.1 应用于 $\hat{\mathbf{f}}_t^{(\alpha)}$ 和 $\mathbf{h}^{(\alpha)}$（前六个周期中的等权重）

  $\lambda_t^{(\alpha)} \leftarrow S^{(\alpha)} / \sum_{\alpha'} S^{(\alpha')}$; $\mathbf{b}_t \leftarrow \sum_\alpha \lambda_t^{(\alpha)} \mathbf{b}_t^{(\alpha)}$

  观测 $\mathbf{r}_t$; $W \leftarrow W(\mathbf{b}_t^\intercal \mathbf{r}_t - \frac{c}{2}\|\mathbf{b}_t - \mathbf{h}\|_1)$; $\mathbf{h} \leftarrow \mathbf{b}_t \odot \mathbf{r}_t / \mathbf{b}_t^\intercal \mathbf{r}_t$

  更新 $S^{(\alpha)}$ 通过（5.1）和 $\mathbf{h}^{(\alpha)} \leftarrow \mathbf{b}_t^{(\alpha)} \odot \mathbf{r}_t / \mathbf{b}_t^{(\alpha)\intercal} \mathbf{r}_t$

  如果 $t \geq t_0$：更新 $M_{1i}^{(\alpha)}, M_{0i}^{(\alpha)}$ 通过（4.2）–（4.3）与 $f_{1ti} = f_{(\alpha)ti}$

  更新 v_{ti} 通过（3.4）

结束循环

6. 数值结果

6.1. **数据与设置。** 表 2展示了九个数据集。前四个数据集 TSE、MSCI、NYSE-N 和 NYSE-O 分别以收益期和资产计量，其规模为 1259×88, 1043×24, 6431×23 和 5651×36。它们也是 [7]中使用的基准数据集。它们的背景学习率分别为 5×10^{-5}（TSE）、6×10^{-4}（MSCI）、10^{-5}（NYSE-N）和 2×10^{-5}（NYSE-O），与 [7]中相同。额外的完整面板数据集 DJIA、SP500、SZ50、COVID 和 new-DJIA 使用相同的门控设置和背景学习率5×10^{-5}。背景因子初始值为 0。5，从第 2 个时期开始更新，越过边界时重置为 0。5。此外，所有 NPM 策略在前六个时期均使用预测1 和等权重投资组合。

数据集	描述	资产	周期	γ
TSE	多伦多证券交易所股票，1994–1998年	88	1259	5×10^{-5}
MSCI	MSCI世界指数成分指数，2006–2010年	24	1043	6×10^{-4}
NYSE-N	纽约证券交易所股票，1985–2010年	23	6431	1×10^{-5}
NYSE-O	纽约证券交易所股票，1962–1984年	36	5651	2×10^{-5}
DJIA	道琼斯工业平均指数成分股，2001–2003年	30	507	5×10^{-5}
标准普尔500指数	标准普尔500指数成分股，1998–2003年	25	1276	5×10^{-5}
SZ50	上证50成分股，2014–2018年	26	1219	5×10^{-5}
COVID	道琼斯工业平均指数成分股，2019–2023年	30	1189	5×10^{-5}
new-DJIA	道琼斯工业平均指数成分股，2008–2016年	30	2115	5×10^{-5}

表2。数据集。周期表示价格相对数的数量；γ 是 AOLNPM 的背景学习率。

周期从 1 开始编号。算法 1 中的所有 G-PWNPM 专家采用相同的时序。门控校准从第 21 期观测到的收益开始，因此 $t_0 = 21$。在此之前

对于每项资产和每位专家，这两个软计数在该收益期之前均初始化为 1。学习得到的门控权重首次用于调整第 102 期交易的预测，因此 $t_g = 102$。除非另有说明，交易成本率为 $c = 0.001$ 。

6.2. **比较与评估**。 我们比较了十三种基线方法。基准策略包括市场（买入并持有均匀投资组合）、均匀常数再平衡投资组合（UCRP）以及事后最佳常数再平衡投资组合（BCRP）[6]；BCRP 使用未来价格，仅作为参考。跟随赢家策略包括指数梯度（EG） [10]和 ONS [1]。跟随输家策略包括 Anticor [3], PAMR [16], OLMAR-1 和 OLMAR-2 [13], 以及 RMR [11]。成本感知策略包括 TCO1 和 TCO2 [15] 以及 AOLNPM [7]。所有策略均需支付（2.1）中的成本。附录 B 给出了参数，并将我们的结果与 [7, 表 2]进行比较。

作为我们的比较贡献，共涉及三种策略。将第 4 节中的门控直接应用于 AOLNPM 的预测，即得到 AOLNPM-AM。PWNPM 是第5 节中不带门控的基金。最后，我们提出的方法是 G-PWNPM，其网格包含 $\mathscr{A}$ 个19 种衰减因子。

对于日对数增长分别为 $g_t = \log(W_t/W_{t-1})$和 g_t'的两种策略，其终端财富之比为$\exp \sum_t(g_t - g_t')$。我们使用基于具有20个滞后阶数的Newey–West标准误 [18]的t统计量，检验 $g_t - g_t'$ 的均值是否为正。在九个数据集上，我们报告比值的几何平均值、比值大于1的数据集数量以及合并统计量 $z = \sum_d t_d/3$。如果各数据集相互独立，则在原假设下， z 将近似服从标准正态分布；但由于部分数据集在时间上存在重叠，因此 z 仅具有指示性。我们还报告预热期后的平均单向换手率 $\|\mathbf{b}_t - \mathbf{h}_t\|_1/2$ 。

6.3. **终端财富**。 表 3和 4报告了初始财富为1时的终端财富，图 1 展示了基准数据集上的财富路径。在四个基准数据集上，G-PWNPM的最终表现超过全部十三种基线方法，其终端财富分别是TSE、MSCI、NYSE-N和NYSE-O上AOLNPM终端财富的2。7、1。2、15 和 4，567 倍。仅门控策略在这些数据集上的表现也优于AOLNPM，而AOLNPM-AM在TSE和MSCI上的终端财富最高。

在五个附加数据集上，AOLMA 预测所含信息很少（表 1），除了 Anticor 之外，所有跟随输家策略在上证50和 COVID 上都亏损。G-PWNPM 在这五个数据集上的最终表现仍优于 AOLNPM，优势倍数介于 1.17 （道琼斯工业平均指数）和 5.75 （new-DJIA）之间，但其排名处于基线方法的中游。在道琼斯工业平均指数和标准普尔500指数上，TCO2 的最终表现最佳。在市场上涨的上证50、COVID 和 new-DJIA 上，基准策略、跟随赢家策略、Anticor 和 TCO2 的最终表现都优于 G-PWNPM。

6.4. 显著性和换手率。表5比较了 G-PWNPM 与 AOLNPM 在各数据集上的表现。G-PWNPM 在全部九个数据集上的最终表现都优于 AOLNPM，几何平均比率为 5.92，合并 $z = 4.68$。在单个数据集上，每日差异波动较大；在 NYSE-O、上证50 和 new-DJIA 上，差异在 5% 显著性水平上显著。与各基准数据集上的最佳基线策略（TSE 和 NYSE-O 上的 OLMAR-2、MSCI 上的 AOLNPM，以及 NYSE-N 上的 OLMAR-1）相比，优势倍数分别为 2.60、1.20、10.3 和 3,878，对应 $t = 0.89$、0.66、1.17 和 3.80。

在每个数据集上，G-PWNPM 的交易量都低于 AOLNPM，平均单向换手率比 AOLNPM 低 15–55%。不过，换手率仍然较高：每个时期的投资组合换手率为 0.31–0.66，而 AOLNPM 为 0.69–0.78，因为 NPM 会将所有卖出的持仓转入一个

策略	TSE MSCI		NYSE-N	NYSE-O
市场	1.613	0.906	18.06	14.50
UCRP	1.581	0.923	30.35	26.20
BCRP	6.722	1.503	115.7	235.0
EG	1.580	0.922	29.88	26.25
ONS	1.524	0.826	18.95	94.00
Anticor	33.48	2.448	61,449	1.59×10^{7}
PAMR	101.4	6.045	5,689	4.43×10^{13}
OLMAR-1	27.06	7.334	7.69×10^{6}	1.77×10^{15}
OLMAR-2	262.9	9.992	4.10×10^{6}	1.21×10^{16}
RMR	81.20	8.060	5.24×10^{6}	3.68×10^{15}
TCO1	60.90	6.004	4.96×10^{5}	5.98×10^{12}
TCO2	67.89	3.982	3.64×10^{6}	9.08×10^{11}
AOLNPM	251.4	12.73	5.23×10^{6}	1.03×10^{16}
PWNPM	251.0	9.498	3.36×10^{7}	1.07×10^{16}
AOLNPM-AM	**1,642**	**16.47**	7.23×10^{7}	3.64×10^{19}
G-PWNPM	683.0	15.21	$\mathbf{7.90\times10^{7}}$	$\mathbf{4.69\times10^{19}}$

表3. 基准数据集上的终端财富， $W_0 = 1$ 和 $c = 0.001$。粗体表示每列中在线策略的最高终端财富；BCRP 使用未来价格。

策略	道琼斯工业平均指数 标准普尔500指数		上证50指数 COVID		new-DJIA
市场	0.764	1.342	2.037	1.418	2.192
UCRP	0.810	1.633	2.137	1.427	2.368
BCRP	1.235	4.033	6.230	4.009	5.484
EG	0.807	1.618	2.131	1.427	2.360
ONS	1.481	3.162	**2.713**	**1.471**	**4.762**
Anticor	1.506	4.411	1.913	1.198	2.667
PAMR	0.451	1.805	0.101	0.081	0.101
OLMAR-1	1.593	7.416	0.449	0.536	0.893
OLMAR-2	0.790	3.557	0.254	0.235	0.274
RMR	1.708	4.026	0.331	0.365	1.013
TCO1	1.166	2.725	0.475	0.360	1.231
TCO2	**1.741**	**8.184**	1.005	0.741	2.502
AOLNPM	0.901	2.671	0.257	0.259	0.250
PWNPM	1.109	3.750	0.292	0.356	0.592
AOLNPM-AM	0.994	3.488	0.377	0.216	0.893
G-PWNPM	1.058	3.426	0.770	0.396	1.437

表4. 附加数据集 $W_0 = 1$ 和 $c = 0.001$上的终端财富。粗体表示每列中在线策略的最高终端财富。

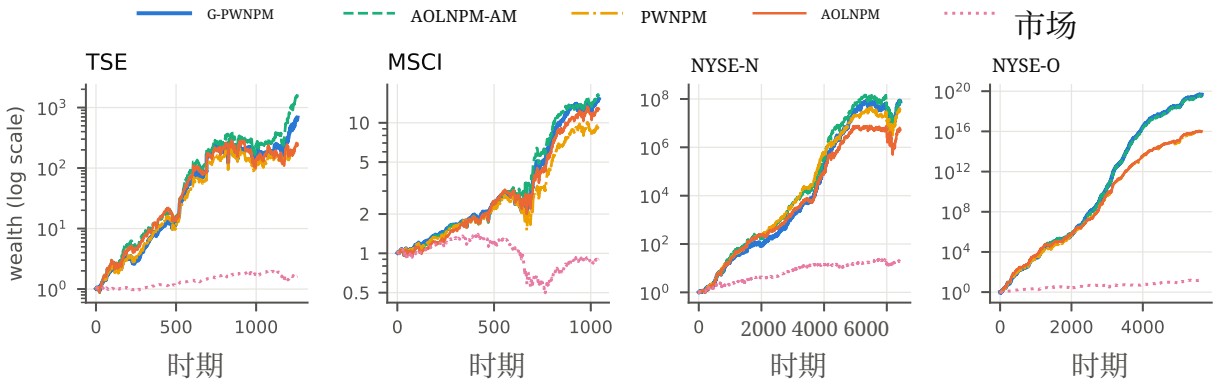

图1. 基准数据集上从初始财富 1 开始的财富路径，$c = 0.001$处采用对数纵轴。

	TSE	MSCI	NYSE-N	NYSE-O	道琼斯工业平均指数	标准普尔500指数	上证50指数	COVID	new-DJIA
财富比率	2.72	1.20	15.1	4,567	1.17	1.28	2.99	1.53	5.75
t-统计量	0.92	0.66	1.40	3.91	0.57	0.33	2.14	1.09	3.01
换手率，AOLNPM	0.78	0.78	0.72	0.78	0.74	0.75	0.69	0.70	0.70
换手率，G-PWNPM	0.66	0.49	0.40	0.48	0.53	0.52	0.31	0.46	0.32

表5. 在 $c = 0.001$ 下，G-PWNPM 与 AOLNPM 的比较。财富比率：G-PWNPM 的终端财富除以 AOLNPM 的终端财富；t-统计量：日对数增长差值的 Newey–West t-统计量；换手率：每个时期的平均单向换手率。

单一资产。G-PWNPM 在全部九个数据集上的年化夏普比率也都高于 AOLNPM，并且在其中八个数据集上的最大回撤更小。

6.5. 交易成本与消融。表6将成本率从 0 调整到 0.7%，涵盖了 [7] 的 0.1–0.7% 范围。随着成本率上升，G-PWNPM 相较于 AOLNPM 的优势不断扩大：从无成本时的 3.02，增至成本率为 $c = 0.1\%$ 时的 5.92，以及成本率为 $c = 0.7\%$ 时的 131（几何平均值）；当成本率为 $c \geq 0.1\%$ 时，G-PWNPM 在每个数据集上的表现都优于 AOLNPM。无成本时，NPM 仅持有预测值最高的资产，因此 AOLNPM-AM 在 $c = 0$ （2.97）取得的收益完全来自命题4.2(b)的重新排序。随着成本率上升，仅持有基金也比 AOLNPM 获益更多（PWNPM 从 1.11 增至 14.16）。在各基线方法中，OLMAR-2 和 RMR 在高成本率下远远落后于 AOLNPM。TCO2 同样会控制成本，从 $c = 0.2\%$ 起超过 AOLNPM，但在所有成本率下，其几何平均值仍低于 G-PWNPM。

表6 中的 PWNPM 和 AOLNPM-AM 两行分别展示基金和门控的作用。收益主要来自门控：当 $c = 0.1\%$的门控仅应用于 AOLNPM 的预测（AOLNPM-AM）时，终端财富变为原来的5.16 倍（合并后的几何平均值为 $z = 3.53$），而基金单独运作（PWNPM， $z = 1.72$）时为1.47 倍。两者结合后，G-PWNPM 达到5.92；它是在全部九个数据集上表现都超过 AOLNPM 的唯一策略，且合并后的 z最大。

	成本率 c						
策略	0	0.05%	0.1%	0.2%	0.3%	0.5%	0.7%
OLMAR-2	1.09 (7)	1.04 (7)	0.98 (4)	0.81 (3)	0.57 (2)	0.16 (0)	0.03 (0)
RMR	0.91 (4)	0.98 (6)	1.05 (6)	1.11 (5)	1.00 (5)	0.47 (2)	0.13 (2)
TCO2	0.21 (5)	0.34 (5)	0.53 (5)	1.26 (6)	2.62 (6)	7.03 (8)	12.40 (8)
PWNPM	1.11 (5)	1.28 (6)	1.47 (7)	2.21 (7)	3.19 (8)	7.21 (9)	14.16 (9)
AOLNPM-AM	2.97 (7)	3.98（8）	5.16（8）	12.93（8）	25.63（8）	47.84（9）	89.52（9）
G-PWNPM	3.02 (8)	4.41 (8)	5.92 (9)	14.80 (9)	30.70 (9)	71.41（9）	131（9）

表6。在相同成本率下相对于 AOLNPM 的终端财富：九个数据集上的几何平均值，以及括号中财富比率高于 1 的数据集数量
比率大于 1。

与 PWNPM 相比，门控使终端财富变为原来的 4.03 倍（$z = 4.23$）；与 AOLNPM-AM 相比，基金带来的平均增益较小（1.15， $z = 1.16$）。基金主要提升了表现的一致性：在五个附加数据集上，几何平均值从 AOLNPM-AM 的 1.44 升至 2.09；而在基准数据集上，AOLNPM-AM 表现更高（25.36，对比 21.75）。去除收益最大的 NYSE-O 数据集后，G-PWNPM 仍能获得 2.58 倍的收益。基金的最终表现与事后最佳专家相差不超过 1.85–5.81 倍，远低于命题5.1(i)中的因子 $|\mathscr{A}| = 19$ 。

命题 4.2 将门控的两种影响区分开来：共同权重会提高 NPM阈值，而资产特定权重会对资产重新排序。用第 102 个时期 AOLNPM-AM 的权重替换权重，可以将这两种影响分离开来。共同常数权重 w 等价于在原始预测上使用阈值 c/w 。在九个数据集上均采用 $w = 1/8$ 时，相对于 AOLNPM 的几何平均终端财富比率为 2.15。事后从 $\{1/10, 1/8, 1/6, 1/5, 1/4, 1/3, 1/2, 1\}$ 中为每个数据集选择表现最佳的权重，可将该比率提高到 2.33。在全样本均值水平保持 π_{ti} 的横截面变异，可以恢复自适应混合的大部分收益（4.42 对 5.16）；而让每个资产采用另一个资产的比例则会损失财富（0.13）。因此，按照信号的可靠性对资产进行重新排序是主要影响。

6.6. 稳健性。每次只更改 G-PWNPM 的一个设置，即计数记忆$\rho_M \in \{0.95, 0.97, 0.99, 0.995\}$、方差记忆 $\rho_v \in \{0.94, 0.97, 0.99\}$、门控的起始时期（从第 52 或第 202 期开始交易，或从 r_{51} 开始计数），或网格（包含 9、13 或 39 个衰减因子），其相对于 AOLNPM 的几何平均比率仍保持在 5.44 到 6.48 之间，且 G-PWNPM 在至少七个数据集上的表现仍优于 AOLNPM。自由度为 $\nu \in \{3, 4\}$、方差为 v_{ti} 的 Student-t 似然可将几何平均比率提高至 7.00–7.72，但此时 G-PWNPM 在标准普尔 500 指数上的表现略逊于 AOLNPM；对于 $\nu = 3$ ，在 MSCI 上也略逊于 AOLNPM，因此我们保留第 3 节中的高斯似然。G-PWNPM 的持仓较为集中：第 101 期之后，其持仓最多的股票的平均权重介于 8%（new-DJIA）和 42%（NYSE-O）之间。从每个数据集中移除该股票并重新运行两种策略后，G-PWNPM 在八个数据集上的最终表现仍优于 AOLNPM，几何平均比率为 2.88；在 NYSE-O 上，该比率从 4,567 降至 16.8；在 TSE 上则降至 0.60。因此，收益并非依赖于单只股票，尽管 NYSE-O 上的大部分收益确实来自这只股票。

7. 结论

NPM 会在预测差距超过成本时进行交易，因此其成本控制效果取决于预测的校准程度，而 AOLMA 预测远未校准。对于每项资产，我们根据 AOLMA 预测和中性预测的预测似然，在线学习二者自适应混合中的占比，并在运行 NPM 前按该占比缩减预测。该占比估计预测中实际实现的部分。对于 NPM，门控起到资产特定阈值的作用，而它对资产的重新排序是收益的主要来源。固定衰减因子的收益加权门控专家集成取代了 AOLNPM 的数据集特定步长。在九个数据集上，G-PWNPM 最终在所有九个数据集上的表现均优于 AOLNPM，换手率更低，且在大多数数据集上回撤更小；其优势还会随着成本率的增加而扩大。

门控学习得很慢，需要经过数百个周期（附录 A），而在 AOLMA 预测能力较弱的一些数据集上，G-PWNPM 会降低 NPM 的损失，但不会将损失转化为收益。更快地学习各比例的方法（例如通过汇集不同资产之间的信息），以及同样持有追随赢家专家的基金，都是自然的扩展方向。

附录A. 命题 4.1 的证明

证明。 省略下标 t 和 i，并将 $x + v/2 = \beta d + \sqrt{v}\,\xi$ 写作 $\xi \sim N(0,\ 1)$。根据式（3.6），

$$u = \frac{d(\beta d + \sqrt{v}\,\xi) - d^2/2}{v} = \varepsilon\xi + \left(\beta - \tfrac{1}{2}\right)\varepsilon^2,$$

因此 $\mathbb{E}\,u = (\beta - \frac{1}{2})\varepsilon^2$、$\mathbb{E}\,u^2 = \varepsilon^2 + O(\varepsilon^4)$、$\mathbb{E}\,u^3 = O(\varepsilon^4)$ 和 $\mathbb{E}\,u^4 = O(\varepsilon^4)$。根据（4.2），$z = g(u)$，其中 $g(u) = \pi e^u/(\pi e^u + 1 - \pi)$ 是 $u + \log\frac{\pi}{1-\pi}$ 的逻辑函数，且其导数有界，$g(0) = \pi$，$g'(0) = \pi(1-\pi)$，以及 $g''(0) = \pi(1-\pi)(1-2\pi)$。根据带四阶余项的泰勒定理可得

$$\mathbb{E}\,z - \pi = g'(0)\,\mathbb{E}\,u + \tfrac{1}{2}g''(0)\,\mathbb{E}\,u^2 + \tfrac{1}{6}g'''(0)\,\mathbb{E}\,u^3 + O(\mathbb{E}\,u^4) = \pi(1-\pi)(\beta - \pi)\varepsilon^2 + O(\varepsilon^4).$$

由 (4.3)，$\pi_{t+1} = (\rho_M n\pi + z)/(\rho_M n + 1)$，因此 $\pi_{t+1} - \pi = (z - \pi)/(\rho_M n + 1)$。由于 n 由过去的信息决定，命题 4.1 中的期望结论随之成立。对于 $\pi < \beta$，首项为正；对于 $\pi > \beta$，首项为负；若 $\beta \le 0$，则该项在 (0, 1) 上为负；若 $\beta \ge 1$，则为正。最后，$\mathbb{E}[r \mid \text{过去}] = \mathbb{E}\,e^x = e^{\beta d} = (f^B)^\beta$，且 $1 + \beta(e^d - 1) - e^{\beta d} = \frac{1}{2}\beta(1 - \beta)d^2 + O(d^3)$。□

由于 $n_{ti} \to 1/(1 - \rho_M) = 50$，首项约为每个周期 $\pi_{ti}(1 - \pi_{ti})(\beta - \pi_{ti})\varepsilon_{ti}^2/50$。对于 AOLNPM 预测，在我们的数据集中，$\varepsilon_{ti}^2$ 的平均值为 0.21–0.32，因此 π_{ti} 经过数百个周期后趋近于 β。

附录B. 基线方法及与 [7] 的比较

所有基线方法均采用 [14], 中的默认参数， [7], 并按（2.1）支付成本；参数采用所引论文中的记号。市场持有均匀买入并持有投资组合，UCRP 每个时期都再平衡至 $\mathbf{1}/n$。BCRP 是不计成本的对数最优常数再平衡投资组合，通过 Cover 的乘法更新在事后计算；随后，它支付再平衡至该投资组合的成本。指数梯度采用 $\eta = 0.05$，ONS 采用 $\eta = 0$、$\beta = 1$ 和 $\delta = 1/8$。Anticor 是窗口为2、.....、30的 Anticor 专家所构成的买入并持有组合；我们根据专家扣除成本后的财富为其赋权，并且仅进行交易

净额，如（5.2）所示。PAMR 采用 $\varepsilon = 0.5$。OLMAR-1 采用 5 期简单移动平均，OLMAR-2 采用参数为 $\alpha = 0.5$的指数移动平均，两者均使用 $\varepsilon = 10$。RMR 使用过去 5 个价格的 L_1中位数，并通过 Weiszfeld算法计算，参数为 $\varepsilon = 5$。TCO1（上一期价格相对值反转）和 TCO2（5 期移动平均反转）采用 $\eta = 10$和 $\lambda = 10c$ [15]。

表7将我们的终端财富与 [7, 表 2]进行了比较。对于基准策略和跟随赢家策略，两者的差异在0.1%以内。对于PAMR、OLMAR-1、OLMAR-2和RMR，两者的对数财富差异最多为1.5%。Anticor的差异更大，我们认为这是其专家组合造成的。我们的AOLNPM按照 [7, 第 3]节所述，采用表2中的步长，但在全部四个数据集上的结果均低于已发表的数值；我们尚未找到差异的来源。本文所有NPM策略都采用我们对AOLMA和NPM的实现，因此第6节中与AOLNPM的比较是在相同约定下进行的。G-PWNPM在全部四个数据集上的结果也高于已发表的AOLNPM数值：TSE上为683.0，而AOLNPM为529.6；MSCI上为15.21，而AOLNPM为14.94；NYSE-N上为 7.90×10^7，而AOLNPM为 7.73×10^6；NYSE-O上为 4.69×10^{19}，而AOLNPM为 3.45×10^{16}。

	MSCI		NYSE-O		NYSE-N		TSE	
策略	Ours	[7]	Ours	[7]	Ours	[7]	Ours	[7]
市场	0.906	0.906	14.50	14.50	18.06	18.06	1.613	1.613
UCRP	0.923	0.922	26.20	26.19	30.35	30.33	1.581	1.580
BCRP	1.503	1.503	235.0	235.1	115.7	115.7	6.722	6.726
EG	0.922	0.922	26.25	26.24	29.88	29.87	1.580	1.579
ONS	0.826	0.826	94.00	93.94	18.95	18.94	1.524	1.524
Anticor	2.448	3.000	1.59×10^7	4.92×10^6	61,449	43,957	33.48	21.23
PAMR	6.045	6.030	4.43×10^{13}	4.31×10^{13}	5,689	5,633	101.4	100.8
OLMAR-1	7.334	7.407	1.77×10^{15}	1.59×10^{15}	7.69×10^6	7.51×10^6	27.06	26.29
OLMAR-2	9.992	10.03	1.21×10^{16}	1.17×10^{16}	4.10×10^6	3.97×10^6	262.9	274.5
RMR	8.060	8.123	3.68×10^{15}	3.48×10^{15}	5.24×10^6	4.18×10^6	81.20	81.86
AOLNPM	12.73	14.94	1.03×10^{16}	3.45×10^{16}	5.23×10^6	7.73×10^6	251.4	529.6

表7. $c = 0.001$时的终端财富：我们的实现与 [7, 表 2]。

致谢： 本研究工作获香港研究资助局资助，资助编号17309522。

参考文献

[1] Amit Agarwal、Elad Hazan、Satyen Kale 和 Robert E Schapire，《基于牛顿法的投资组合管理算法》，第23届国际机器学习会议论文集，2006年，第9–16页。[2] Avrim Blum 和 Adam Kalai，《有交易成本和无交易成本的通用投资组合》，《机器学习》**35** （1999），第3期，193–205。[3] Allan Borodin、Ran El-Yaniv 和 Vincent Gogan，《我们能否学会击败最佳股票》，人工智能研究杂志 **21** （2004），579–594。[4] Olivier Cappé 和 Eric Moulines，《潜在数据模型的在线期望最大化算法》，皇家统计学会杂志B辑：统计方法论 **71** （2009），第3期，593–613。

[5] Nicolò Cesa-Bianchi 和 Gábor Lugosi，《预测、学习与博弈》，剑桥大学出版社，剑桥，2006年。[6] Thomas M Cover，《通用投资组合》，《数学金融》1（1991），第1期，1–29。[7] Sini Guo、Jia-Wen Gu 和 Wai-Ki Ching，《带交易成本的自适应在线投资组合选择》，《欧洲运筹学杂志》295（2021），第3期，1074–1086。[8] Sini Guo、Jia-Wen Gu 和 Wai-Ki Ching，《数据驱动的智能金融投资组合选择》，施普林格，2026年。[9] Sini Guo、Jia-Wen Gu、Christopher H Fok 和 Wai-Ki Ching，《采用状态依赖价格估计器和交易成本的在线投资组合选择》，《欧洲运筹学杂志》311（2023），第1期，333–353。[10] David P Helmbold、Robert E Schapire、Yoram Singer 和 Manfred K Warmuth，《使用乘法更新的在线投资组合选择》，《数学金融》8（1998），第4期，325–347。[11] Dingjiang Huang、Junlong Zhou、Bin Li、Steven CH Hoi 和 Shuigeng Zhou，《用于在线投资组合选择的稳健中位数反转策略》，《IEEE知识与数据工程汇刊》28（2016），第9期，2480–2493。[12] Bin Li 和 Steven CH Hoi，《在线投资组合选择：综述》，《ACM计算调查》（CSUR）46（2014），第3期，1–36。[13] Bin Li、Steven CH Hoi、Doyen Sahoo 和 Zhi-Yong Liu，《用于在线投资组合选择的移动平均反转策略》，人工智能222（2015），104–123。[14] Bin Li 和 Steven Chu Hong Hoi，《在线投资组合选择：原理与算法》，CRC出版社，2015年。[15] Bin Li、Jialei Wang、Dingjiang Huang 和 Steven CH Hoi，《在线投资组合选择的交易成本优化》，《定量金融》18（2018），第8期，1411–1424。[16] Bin Li、Peilin Zhao、Steven CH Hoi 和 Vivekanand Gopalkrishnan，《PAMR：用于投资组合选择的被动激进均值反转策略》，《机器学习》87（2012），第2期，221–258。[17] Harry Markowitz，《投资组合选择》，《金融学杂志》7（1952），第1期，77–91。[18] Whitney K Newey 和 Kenneth D West，《一个简单的正半定异方差和自相关一致协方差矩阵》，《计量经济学》55（1987），第3期，703–708。[19] Adrian E Raftery、Miroslav Kárný 和 Pavel Ettler，《模型不确定性下通过动态模型平均进行在线预测：在冷轧机上的应用》，《技术计量学》52（2010），第1期，52–66。

香港大学数学系，薄扶林，香港电子邮件地址： gracexiecr@connect.hku.hk

香港大学数学系，薄扶林，香港电子邮件地址： alveyqlin@connect.hku.hk

香港大学数学系，薄扶林，香港电子邮件地址： wching@hku.hk