

基于 ADMM 的 L0 范数 短期稀疏投资组合优化算法复现与改进

摘要

本文提出了一种新的基于 0-范数的短期稀疏投资组合优化模型。与现有方法相比，该模型根据资产的短期增长潜力选择投资组合，并引入 0-范数约束直接控制所选投资组合中非零资产的最大数量。与基于 1-范数的方法不同，本文的模型中可以直接使用禁止卖空约束。此外，引入稀疏正则化项以消除 SSPO 系统中的琐碎交易。此外，为了求解包含的非凸优化系统，提出了一种基于交替方向乘子法 (ADMM) 概念的算法。文中还研究了算法的收敛性。最后，通过在四个真实数据集上的数值实验证明了该方法的有效性。

关键词：投资组合优化；稀疏正则化；ADMM

1 引言

投资组合优化是一种旨在从候选集合中选择投资资产的策略。在金融市场上，一个好的投资策略应该是高收益、低风险 [9]。1952 年，均值-方差 (MV) 优化模型被提出用于选择这类资产，为现代投资组合理论奠定了基础 [7]。为了进一步降低管理难度和交易成本，稀疏投资组合优化在过去的二十年里受到了极大的关注。总的来说，目前稀疏投资组合优化的研究可以分为两大类：长期稀疏投资组合优化和短期稀疏投资组合优化。

长期稀疏投资组合优化 (LSPO) 利用收益率的统计特征来选择资产和分配资本。由于短期数据的统计特征通常不可靠，LSPO 的组合重组时间间隔是一个月或一年甚至更长。值得注意的是，LSPO 注重选择资产，而非交易时机。

与 LSPO 相比，短期稀疏投资组合优化 (SSPO) 每天（甚至更短）改变一次投资组合。它根据资产的增长趋势而不是统计特征来选择资产。此外，不仅要选择资产，还要选择投资时机，以实现预期收益最大化。在 SSPO 中，不同的算法可以使用不同的标准。

系统在线惰性更新 (OLU) 在 [4,5] 中提出，其稀疏性反映在投资组合变化中，而不是投资组合向量本身。所以，严格来说，它不是一个稀疏的投资组合，它没有算法收敛的证明。在 [8] 中提出了一个基于 1-范数正则化的 SSPO 系统（称为 SSPO-L1）。然而，由于 1-范数项和自融资约束，不能将禁止卖空约束引入到这样的模型中。否则，禁止卖空和自我融资约束将共同导致 1-范数项失效。为了解决这一问题，SSPO-L1 引入了一个标准化的单纯形，它可以将投资组合向量投影到非卖空和自融资域。然而，在归一化之后，算法的输出可能不是原始问题的最优解。此外，SSPO-L1 [6] 不能直接控制选定投资组合中非零资产的最大数量。

由于 SSPO 的交易频率较高，相应的交易成本可能会对整体收益产生较大影响。在本文

中，使用基于 1-范数的稀疏正则化来消除 SSPO 系统中的琐碎交易。同时引入了一个 0-范数约束，直接控制所选投资组合中非零资产的最大数量（即投资组合的基数），提出了一种基于交替方向乘子法（ADMM）和硬阈值的算法。引文采用如下引用方式。

2 相关工作

本文所使用的方法与短期投资组合优化和 ADMM 密切相关，因此将在本节中简要介绍。

2.1 短期投资组合优化

在 SSPO 中，资产 i 的增长潜力可以通过对数回报 [1] 来评估，该对数回报由下式给出：

$$s_t^i = 1.1 \log \left(\frac{p_{max}^i}{p_t^i} \right) + 1 \quad (1)$$

SSPO 模型的目标是最大化投资组合的潜力 $\mathbf{w}^\top S_i$ 。对于 $\mathbf{w}^\top S_i$ ，[8] 中 SSPO-L1 的公式定义为：

$$\begin{aligned} \mathbf{w}_{t+1} &:= \min_{\mathbf{w}} \quad \mathbf{w}^\top \varphi_t + \theta \|\mathbf{w}\|_1, \\ \text{s.t.} \quad & \mathbf{w}^\top \mathbf{1} = 1 \end{aligned} \quad (2)$$

(2) 为引入 1-范数正则化生成稀疏投资组合方法。然而该方法模型不能包括禁止卖空的约束。否则约束 $\mathbf{w}^\top \mathbf{1} = 1$ 和 $\mathbf{w} \geq \mathbf{0}$ 将共同导致 $\|\mathbf{w}\|_1 = 1$ 失效，这意味着 1-范数正则化在这种情况下是无效的。由于卖空交易风险很大，在某些国家甚至是非法的，因此 SSPO-L1 引入了一个规范化程序，以确保没有卖空交易。但归一化后，算法的输出可能不是原问题的最优解，并且 SSPO-L1 不能直接控制稀疏性。

2.2 ADMM

ADMM 是一种求解优化问题的分布式优化方法。它通常将一个复杂的问题分解成几个容易解决的子问题 [2]。假设目标函数如下

$$\min_{\mathbf{u}} g(\mathbf{u}) + \mathbf{h}(\mathbf{u}) \quad (3)$$

其中 $u \in \mathbb{R}^N$ 包含决策变量。如果 $h(\cdot)$ 部分是非凸和非光滑的，那么很难直接获得最优解。为解决这一问题，引入了一个辅助向量 $v \in \mathbb{R}^N$ ，并将目标函数分解为 $g(\cdot)$ 和 $h(\cdot)$ 两部分。因此，将最小化问题重新定义为

$$\begin{aligned} \min_{\mathbf{u}, \mathbf{v}} \quad & g(\mathbf{u}) + \mathbf{h}(\mathbf{v}), \\ \text{s.t.} \quad & \mathbf{u} = \mathbf{v} \end{aligned} \quad (4)$$

然后，增广拉格朗日函数被定义为

$$L(u, v, b) = g(u) + h(v) + b^\top (u - v) + \frac{\rho}{2} \|u - v\|_2^2 \quad (5)$$

一般 ADMM 迭代方案通过不断迭代，从而能得到封闭形式的解。

3 本文方法

3.1 问题公式化

为了生成具有可控基数的稀疏投资组合，使用 0-范数项来代替 SSPO-L1 中的 1-范数正则化。因此，可以直接引入禁止卖空约束。此外，引入了一个基于 1-范数的正则化项来降低平凡交易和选择投资时机。因此，本文提出的模型的公式如下：

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbf{w}^\top \varphi_t + \lambda \|\mathbf{w} - \mathbf{w}_t\|_1, \\ & \mathbf{w}^\top \mathbf{1} = 1, \\ \text{s.t.} \quad & \|\mathbf{w}\|_0 \leq K, \\ & \mathbf{w} \geq 0. \end{aligned} \tag{6}$$

其中 \mathbf{w}_t 是最后一笔交易的投资组合向量， $\lambda > 0$ 是一个权衡参数， $\lambda \|\mathbf{w} - \mathbf{w}_t\|_1$ 可以将 $\mathbf{w} - \mathbf{w}_t$ 中的小元素精确地推到零，这意味着投资组合中几乎没有变化的交易将被减少， $K > 0$ 是一个预设的整数，表示投资组合中允许的最大资产数量。

3.2 SSPO-L0

在本节中称为 SSPO-L0 的基于 ADMM 的算法来求解模型 (6)。首先，式 (6) 中的模型可以重写为：

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbf{w}^\top \varphi_t + \lambda \|\mathbf{g}\|_1 + \mathcal{L}_A(\mathbf{z}) + \mathcal{L}_B(\mathbf{w}), \\ \text{s.t.} \quad & \mathbf{w}^\top \mathbf{1} = 1, \mathbf{w} = \mathbf{z}, \mathbf{w} - \mathbf{w}_t = \mathbf{g} \end{aligned} \tag{7}$$

式 (7) 中 $\mathbf{z} \in \mathbb{R}^N$ 和 $\mathbf{g} \in \mathbb{R}^N$ 是两个辅助变量， $\mathcal{L}_A(\mathbf{z})$ 和 $\mathcal{L}_B(\mathbf{w})$ 是定义的两个指示函数。应该注意的是，式 (6) 是非凸和非光滑的。式 (7) 中的模型将非凸项和非光滑项从凸项中分离出来，有利于后续计算。使用惩罚方法将式 (7) 改写为

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbf{w}^\top \varphi_t + \lambda \|\mathbf{g}\|_1 + \mathcal{L}_A(\mathbf{z}) + \mathcal{L}_B(\mathbf{w}) + \frac{C}{2} (\mathbf{w}^\top \mathbf{1} - 1)^2 + \frac{\lambda}{2\gamma} \|\mathbf{w} - \mathbf{w}_t - \mathbf{g}\|_2^2, \\ \text{s.t.} \quad & \mathbf{w} = \mathbf{z}. \end{aligned} \tag{8}$$

然后采用增广拉格朗日法求解

$$L(w, z, g, \beta) = \phi(w) + \lambda \|g\|_1 + Z_A(z) + Z_B(w) + \frac{\lambda}{2\gamma} \|w - w_t - g\|_2^2 + \beta^\top (w - z) + \frac{\rho}{2} \|w - z\|_2^2 \tag{9}$$

其中 $\beta \in \mathbb{R}^N$ 是拉格朗日乘子向量， $\rho > 0$ 是增广项的折衷参数。根据 ADMM 的概念和 [5] 中相应的 ADMM 方案，变量 z, g, w, β 迭代更新为

$$\begin{aligned} z^{k+1} &= \operatorname{argmin} L(w^k, z, g^k, \beta^k), \\ g^{k+1} &= \operatorname{argmin}_R \mathcal{L}(\mathbf{w}^k, z^{k+1}, g, \beta^k), \\ w^{k+1} &= \operatorname{argmin} L(w, z^{k+1}, g^{k+1}, \beta^k), \\ \beta^{k+1} &= \beta^k + \rho (w^{k+1} - z^{k+1}) \end{aligned} \tag{10}$$

式 (10) 中的更新步骤被称为 SSPO-L0。最后分别对变量 $\mathbf{w}, \mathbf{z}, \mathbf{g}, \beta$ 根据式 (10) 进行求解，得到各个变量的具体的迭代方程式。SSPO-L0 的算法框架如下

Algorithm 1 ADMM for SSPO-L0

Input: $\mathbf{w}_t, \varphi_t, K, C, \lambda, \gamma, \rho$.Initialize: $\mathbf{w}^0, \mathbf{z}^0, \mathbf{g}^0, \beta^0 = 0^N, k = 0$.**while** not converge **do**(1) **z**-update:

$$\mathbf{z}^{k+1} = \mathbf{H}_k(\mathbf{w}^k + \frac{\beta^k}{\rho}).$$

(2) **g**-update:

$$\mathbf{g}^{k+1} = \text{sign}(\mathbf{w}^k - \mathbf{w}_t)[|\mathbf{w}^k - \mathbf{w}_t| - \gamma \mathbf{1}]_+$$

(3) **w**-update:

$$\mathbf{w}^{k+1} = [(\rho + \frac{\lambda}{\gamma})\mathbf{I} + C\mathbf{1}\mathbf{1}^\top]^{-1}(\frac{\lambda}{\gamma}(\mathbf{g}^{k+1} + \mathbf{w}_t) + \rho\mathbf{z}^{k+1} + C\mathbf{1} - \beta^k - \varphi_t)_+.$$

(4) β -update:

$$\beta^{k+1} = \beta_k + \rho(\mathbf{w}^{k+1} - \mathbf{z}^{k+1}).$$

(5) $k = k + 1$.**end while**Output: $\mathbf{w}_{t+1} = \mathbf{w}^{k+1}$. =0

4 复现结果

在本节中，将进行几个数值实验来研究本文提出的算法的性能。实验的基本设置在 4.1 中给出。稀疏性、累积财富、夏普比率、信息比率和交易成本的实验结果在 4.2 中给出。

4.1 问题设置

数据集

实验是在四个来自真实世界股票市场的基准数据集上进行的。表 1 总结了这些数据集的相关信息。第一个数据集 TSE 由 Borodin 收集 [1]，其中包括多伦多证券交易所 (TSE) 1259 个交易日的 88 只股票。与 [1] 中的设置类似，第二个数据集 SP500 由 25 只股票组成，包括它们在 1276 个交易日期间的价格。第三个数据集 NYSE (N) 包括纽约证券交易所 (NYSE) 21 只股票 6431 个交易日价格数据 [9]。数据集 DJIA 由 30 只股票组成，包括 507 个交易日的每日价格 [1]。与原论文相比，本文扩大了 NYSE(N) 与 DJIA 数据集的大小。

表 1. 来自真实世界股票市场的四个基准数据集。

Dataset	Stocks	Time frame	Periods
TSE	88	4/1/1994-31/12/1984	1259
SP500	25	2/1/1998-31/1/2003	1276
NYSE(N)	23	1/1/1985-30/6/2010	6431
DJIA	30	14/1/2001-14/1/2003	507

性能测量

在实验中使用五个指标来评估所提出的方法的性能，分别是稀疏性、累积财富、信息比

率、夏普比率和交易成本。

1. 稀疏度：在最终投资组合中所选资产的数量。对于任何投资组合权重向量 \mathbf{w}_t ，其稀疏性由下式给出

$$Sparsity(\mathbf{w}_t) = \|\mathbf{w}_t\|_0. \quad (11)$$

2. 累积财富：它是评估投资组合投资业绩的主要指标。在第 t 个交易周期，设 S_t 表示投资组合权重向量 \mathbf{w}_t 的投资组合收益，即财富以 $S_t = \mathbf{w}_t^\top \mathbf{r}_t$ 的倍数增长。在不考虑交易成本的情况下， \mathbf{w}_t 从第 1 个交易周期到第 T 个交易周期的累积财富 S_T 定义为

$$S_T = S_0 \prod_{t=1}^T s_t = S_0 \prod_{t=1}^T \mathbf{w}_t^\top \mathbf{r}_t, \quad (12)$$

其中 S_0 表示初始财富。为了方便起见，在实验中将其设置为 1。

3. 夏普比率 (Sharpe Ratio, SR)：它可以衡量投资组合的平均风险调整收益。在相同的风险水平下，SR 越大，投资组合的收益越高。给予一个投资组合权重向量 \mathbf{w}_t ，从第 1 个交易周期到第 T 个交易周期的平均收益 u 和相应的标准差 σ 定义为

$$u = \frac{1}{T} \sum_{t=1}^T (\mathbf{w}_t^\top \mathbf{r}_t - 1) \quad (13)$$

$$\sigma = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (\mathbf{w}_t^\top \mathbf{r}_t - 1 - \mu)^2} \quad (14)$$

根据这些定义，夏普比率 S [12] 为：

$$S = \frac{\mu}{\sigma} \quad (15)$$

4. 信息比 (IR)：它是对投资组合的平均风险调整超额收益的度量，该收益超过相应市场给出的基准收益 [13]。较大的 IR 意味着在相同的风险水平下获得更多的超额收益。换句话说，在相同的回报率下，IR 越大，相应的风险就越低。假设基准是统一的买入并持有 (BAH) 策略，相应的投资组合权重向量由 \mathbf{b}_t 表示 [10]。因此，从第 1 个交易周期到第 T 个交易周期的超额平均收益 μ_e 和相应的标准差 σ_e 定义为：

$$u_e = \frac{1}{T} \sum_{t=1}^T (\mathbf{w}_t^\top \mathbf{r}_t - \mathbf{b}_t^\top \mathbf{r}_t) \quad (16)$$

$$\sigma_e = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (\mathbf{w}_t^\top \mathbf{r}_t - \mathbf{b}_t^\top \mathbf{r}_t - \mu_e)^2} \quad (17)$$

$$IR = \frac{\mu_e}{\sigma_e} \quad (18)$$

5. 交易成本：在实际交易中，交易成本是不可避免的更新投资组合 [14]。根据市场的实际情况，在实验中使用的交易费用率范围为 0 到 0.005。因此，在考虑交易成本的情况下，从第 1 个交易周期到第 T 个交易周期的 \mathbf{w}_t 的累积财富 S_t 被定义为

$$S_T^\eta = S_0 \prod_{t=1}^T \left[(\mathbf{w}_t^\top \mathbf{r}_t - \mathbf{b}_t^\top \mathbf{r}_t) \times \left(1 - \frac{\eta}{2} \sum_{i=1}^N |\mathbf{w}_t^{(i)} - \mathbf{w}_{t-1}^{(i)}| \right) \right] \quad (19)$$

其中， η 表示交易成本率。

对比算法

本文将提出的算法与五个投资组合优化方法比较。将比较算法分为两类：基准测试，优化方法。第一个是一个简单但广泛使用的基准算法：相等权重（EW），即统一的买入并持有（BAH）策略 [10]。第二种是事后诸葛亮的方法：最佳股票，它可以在候选资产中选择最好的单一股票 [10]。既然不知道哪只股票表现最好，这个算法只能作为业绩比较的参考。接下来的三种基于优化方法的策略分别是在线移动平均回归算法（OLMAR） [11]，SSPO-L1 算法，趋势促进价格跟踪算法（TPPT） [3]。原文中将 SSPO-L0 算法与其他三种基于深度学习策略的方法作比较，但由于在原文中这类方法的表现与其他两类相比并不是很好，因此复现工作中忽略该类。为了比较所提出的算法在不同稀疏度下的性能，在实验中设置 $K = 3$ $K = 5$ 和 $K = 7$ 。相应的算法分别称为 SSPO-L0-K3、SSPO-L0-K5 和 SSPO-L0-K7。

4.2 性能比较

稀疏性

EW、Best stock、OLMAR、TPPT 策略不是稀疏投资组合优化算法。因此没有必要比较这些方法的稀疏性。在复现实验中，只比较了 SSPO-L1 和论文提出的算法 SSPO-L0-K3，SSPO-L0-K5 和 SSPO-L0-K7 的稀疏性。4 个数据集的比较结果如图 1 所示。第一行到第四行的数字分别是 TSE、SP500、NYSE (N) 和 DJIA 的相应结果。第一列中的数字显示了不同算法在不同交易时间的输出的稀疏性，其中这些数字的 x 轴表示交易时间，相应的 y 轴表示稀疏性。第二列中的数字是稀疏度的统计直方图，其中 x 轴是稀疏度，y 轴是不同情况下不同稀疏度的统计百分比。

对于所有的数据集，可以观察到，所有的稀疏投资组合优化方法给出的投资组合权重向量是稀疏的。SSPO-L1 的稀疏度在 1 到 2 之间，在大多数情况下为 1，而本文提出的算法 SSPO-L0-K3、SSPO-L0-K5 和 SSPO-L0-K7 的稀疏性范围从 1 到 K，部分情况下是 K。这表明 SSPO-L0 可以控制其稀疏性。换句话说，所提出的算法可以根据投资需求较灵活地控制选择资产的数量。

表 2. 该表显示了比较算法和本文算法在四个数据集上的中值返回值。

Method	TSE	SP500	NYSE(N)	DJIA
EW	1.002425	1.000636	1.000756	0.999131
Best stock	0.975904	1.001129	1.000330	1.000000
OLMAR	1.037040	1.001986	1.000000	1.002047
TPPT	1.060610	1.001862	1.000000	1.002130
SSPO-L1	1.075340	1.002043	1.000000	1.002028
SSPO-L0-K3	0.965812	1.001225	1.000550	1.000401
SSPO-L0-K5	0.992252	1.001431	1.000847	1.000737
SSPO-L0-K7	1.038170	1.001406	1.000833	1.000739

表 3. 通过投资组合优化算法在四个数据集上实现累积财富。

Method	TSE	SP500	NYSE(N)	DJIA
EW	1.596410	1.332759	18.230572	0.763539
Best stock	6.000153	3.792372	84.821430	1.194302
OLMAR	54.938655	15.829293	430278654.281261	2.250202
TPPT	177.211151	10.292425	140640193.257118	2.952519
SSPO-L1	87.638508	14.974413	1746919582.692650	2.957883
SSPO-L0-K3	235.669150	14.391491	55299733.808087	2.187436
SSPO-L0-K5	87.377398	16.187662	53351921.928973	2.179208
SSPO-L0-K7	135.110039	16.253020	52677250.218009	2.186119

累积财富

图 2 中显示了所有方法在四个数据集上的投资组合相对收益的分布，并且每个方法的相应中位数收益在表 2 中给出。在 TSE 和 NYSE (N) 数据集上，本文算法 (SSPO-L0-K3、SSPO-L0-K5 和 SSPO-L0-K7) 的中值收益率高于所有其他比较算法。在 DJIA 上，虽然本文提出的方法的中位数收益率低于 SSPO-L1 和 TPPT。值得注意的是，与本文提出的方法相比，SSPO-L1 和 TPPT 的相对收益更加分散，特别是当值小于 1 时。这意味着 SSPO-L1 和 TPPT 的风险高于我们提出的方法。

不同算法在不同数据集上的最终累积财富在表 3 中给出。根据公式 (12) 计算累积财富。从表 3 中的结果来看，SSPO-L0 的累积财富在 TSE 和 SP500 数据集上高于其他方法。图 3 显示了不同算法的累积财富随时间的波动。可以看出，SSPO-L0 和 SSPO-L1 的累计财富总体趋势相似。在大多数情况下，SSPO-L0 的累积财富高于其他方法。因此，本文提出的算法是一定程度上优于其他的算法的累积财富。

夏普比率

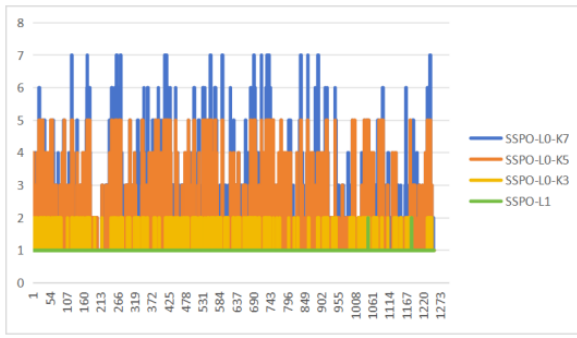
SR 评估绝对收益与总风险之间的关系，根据 (15) 计算。不同投资组合优化算法的 SR 如表 4 所示。可以看出，SSPO-L0 在 TSE 和 SP500 数据集上高于其他方法。研究表明，在一定程度上，SSPO-L0 实现了绝对收益与总风险的良好权衡。

信息比率

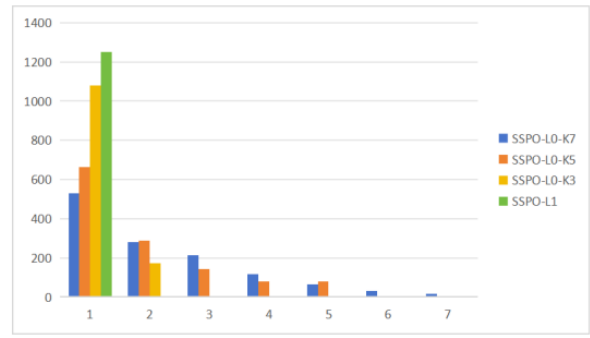
IR 的计算可参考公式 (18)。我们在表 4 中显示了不同投资组合优化算法的投资回报率。显然，表 4 中的实验结果表明，SSPO-L0 可以在部分数据集上实现平均超额收益和风险之间的良好权衡。例如，在不同 K 下，SSPO-L0 在 SP500 上的 IR 分别为 1.3239, 1.3221 和 1.2546，而 OLMAR, SSPO-L1 和 TPPT 的相应结果分别为 1.2773, 1.2293 和 1.1006。

表 4. 四个数据集上的投资组合优化算法的信息比率 (IR) 和夏普比率 (SR)。

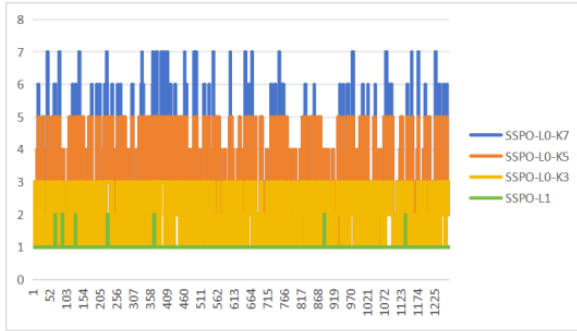
Method	TSE		SP500		NYSE(N)		DJIA	
	IR	SR	IR	SR	IR	SR	IR	SR
EW	0.0000	0.7904	0.0000	0.3554	0.0000	0.7270	0.0000	(0.4326)
Best stock	0.7672	0.9184	0.6206	0.7689	0.2365	0.7496	0.6691	0.4744
OLMAR	1.2211	1.2905	1.2505	1.2773	1.5332	1.6497	1.4695	1.0464
TPPT	1.4659	1.5297	1.0508	1.1006	1.4362	1.5648	1.8136	1.2981
SSPO-L1	1.3195	1.3853	1.2059	1.2293	1.5787	1.6885	1.7509	1.2609
SSPO-L0-K3	1.5319	1.2936	1.2654	1.2546	1.3715	1.4971	1.5018	1.0116
SSPO-L0-K5	1.3301	1.3964	1.3612	1.3221	1.3878	1.5126	1.5626	1.0348
SSPO-L0-K7	1.4351	1.5013	1.3633	1.3239	1.3871	1.5119	1.5682	1.0389



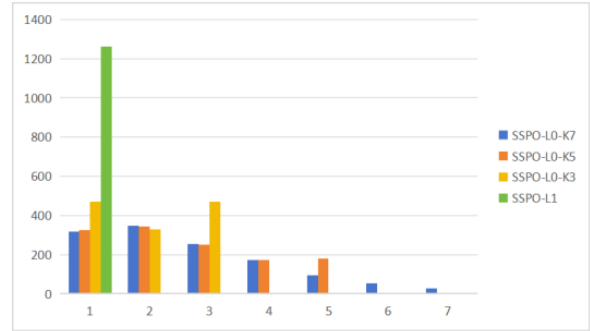
(a)



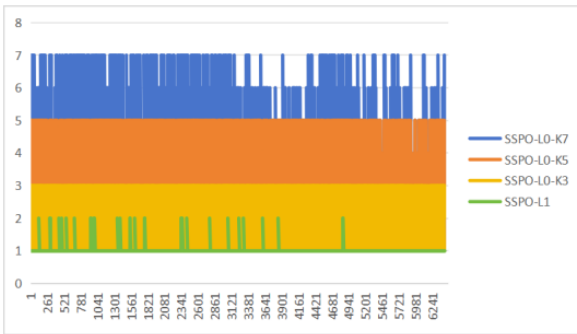
(b)



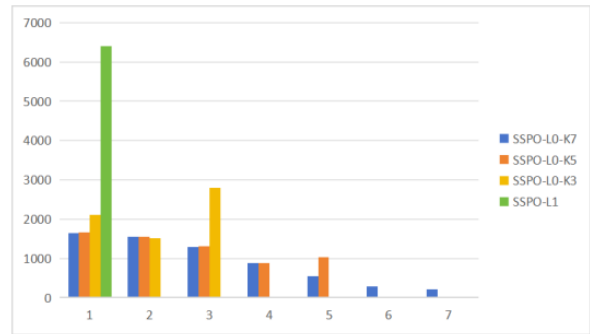
(c)



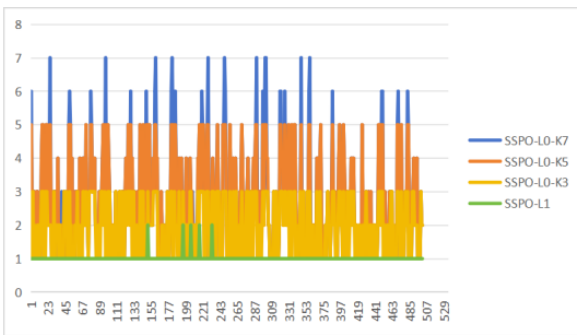
(d)



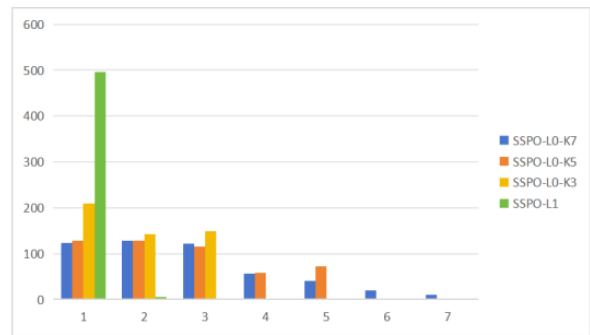
(e)



(f)



(g)



(h)

图 1. SSPO-L1 和 SSPO-L0 在四个数据集上的稀疏性。

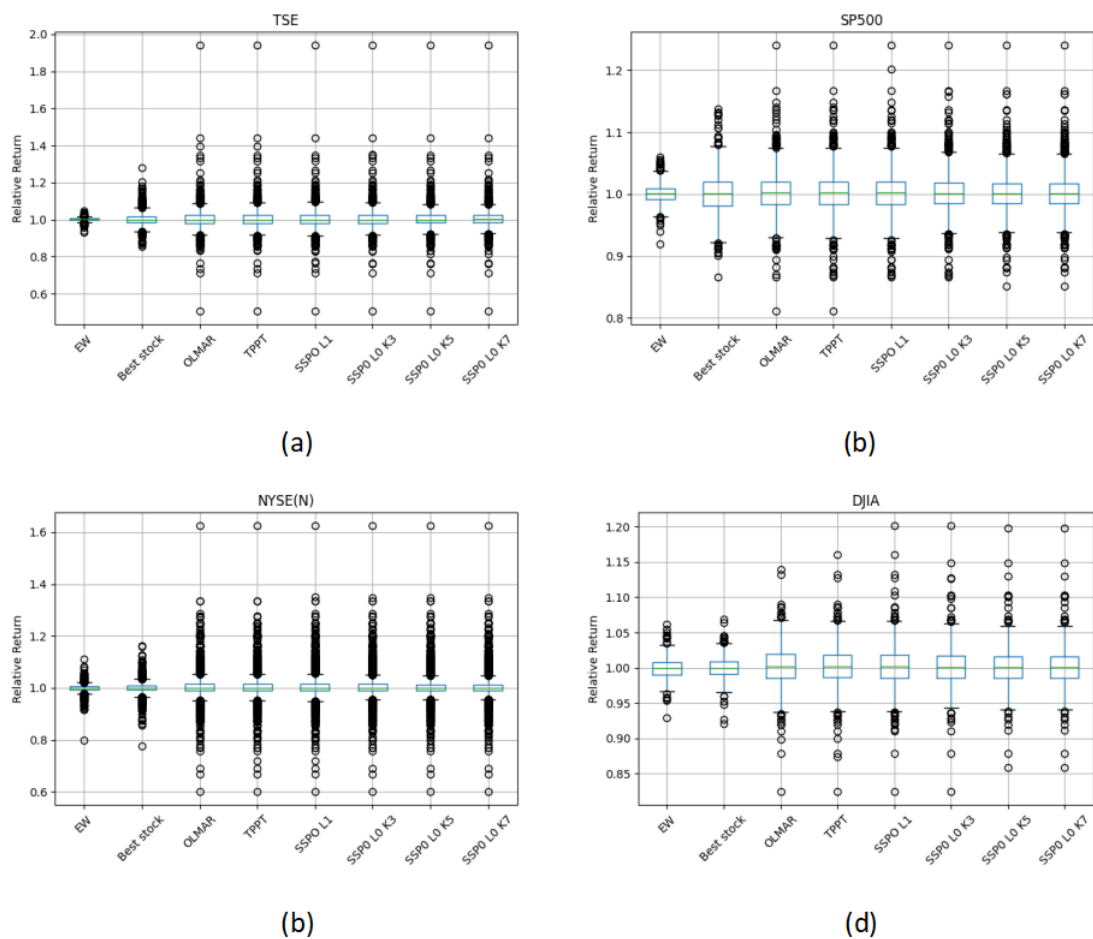


图 2. 四个数据集上投资组合相对收益的箱形图。

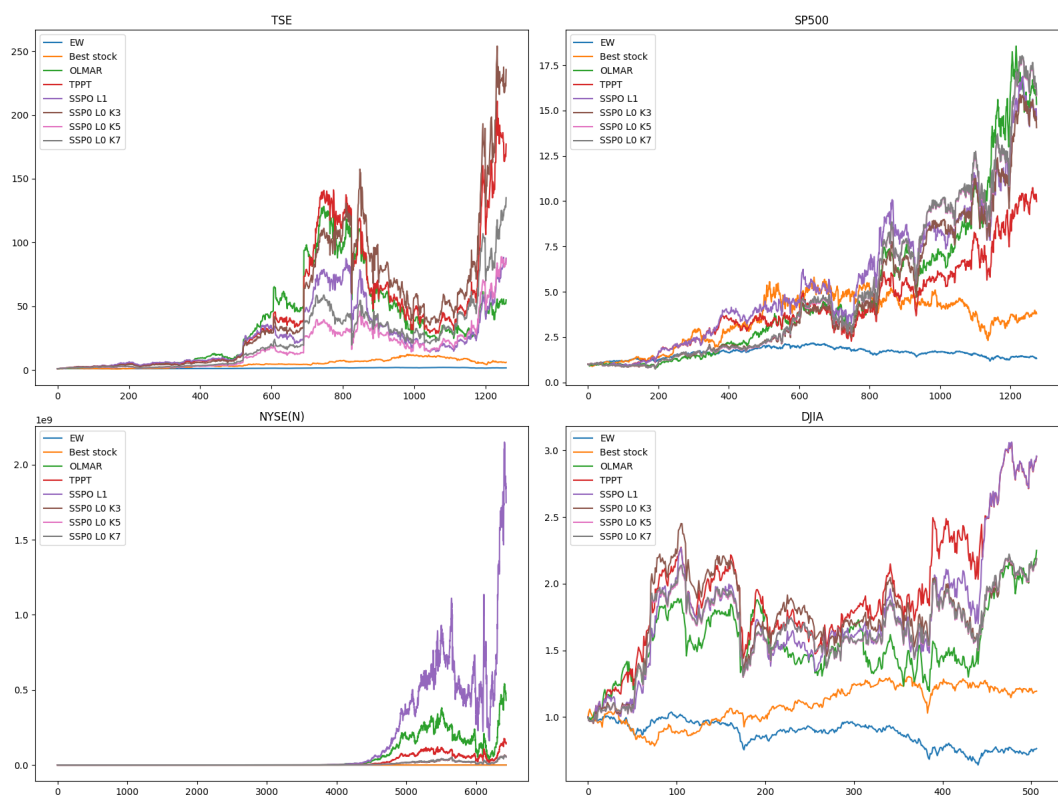


图 3. 四个数据集上的累积财富演化路径。

交易成本

由于 EW 和 Best stock 的投资组合权重向量保持不变，因此交易成本不会影响它们。因此，我们仅保留 EW 作为参考。此外，对于本文提出的算法，我们只考虑 SSPO-L0-K5，因为 SSPO-L0-K3 和 SSPO-L0-K7 的相应结果与 SSPO-L0-K5 非常相似。为了说明交易成本的影响，我们比较了不同交易成本率下的累积财富波动。在四个数据集上的实验结果如图 4 所示。可以看出，随着交易费用率的增加，所有方法的累积财富都在减少。不过，SSPO-L0-K5 累计财富的下降速度（47.4%）低于除基于深度学习的方法之外的其他比较算法。根据图 4 中的结果，基于深度学习的方法的累积财富受交易成本的影响较小。原因可能是这些方法选择的投资组合并不稀疏，而且这些方法更倾向于长期收益，因此它们的投资组合权重向量在不同交易时段变化较小。在所有方法中，SSPO-L1 的下降率最大（76.1%）。基于累积财富的 SSPO-L0 方法通常上级其他方法，并且由于引入了投资时间选择模块，受交易成本的影响较小。

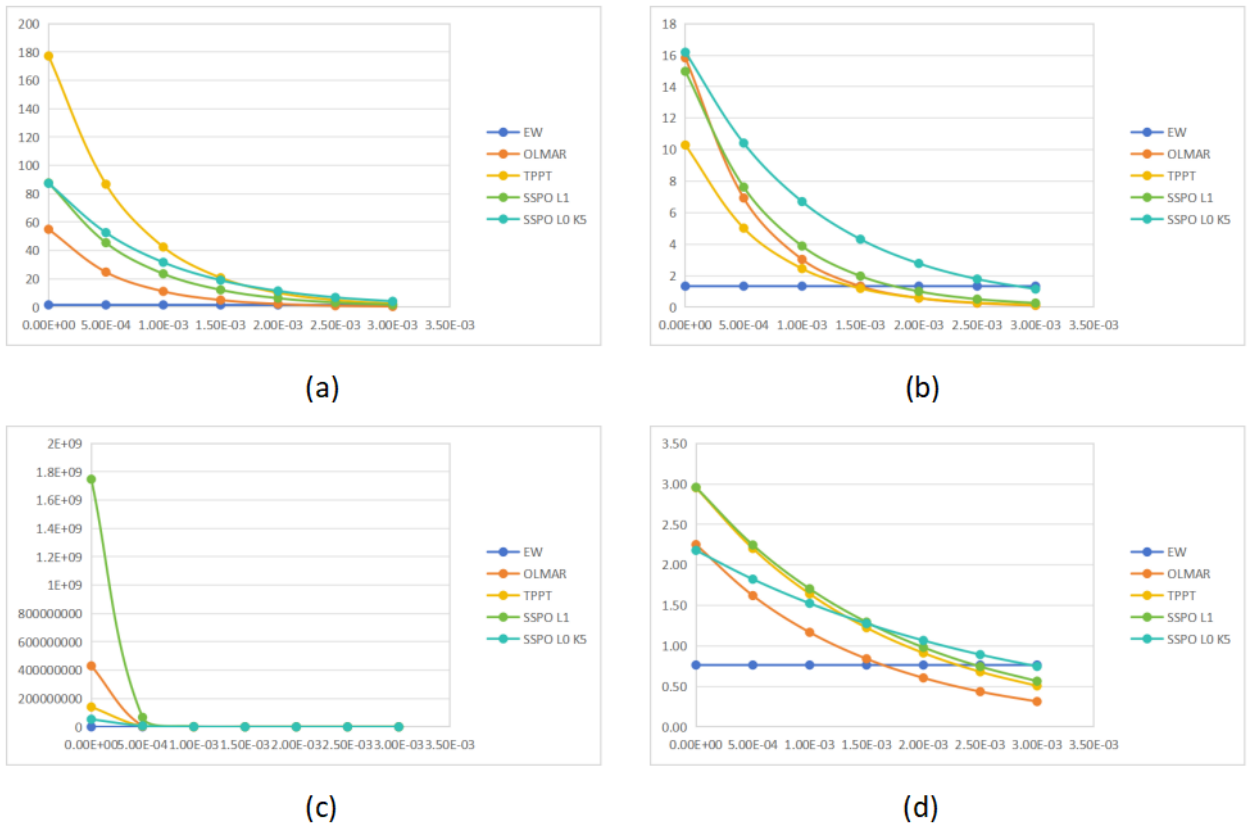


图 4. 四个数据集上的累积财富演化路径。

5 总结与展望

本文提出了一种基于 ADMM 的无卖空约束 SSPO 问题的算法。首先，我们在 SSPO 中引入了一个 0-范数正则化，它可以用来控制所得到的投资组合中非零资产的最大数量。其次，与 SSPO-L1 方法不同的是，该模型可以直接引入禁止卖空的约束条件。第三，使用 1-范数正则化项来消除 SSPO 系统中的琐碎交易。结果，选择模型的事务处理时间。由于 0-范数约束，该模型是非凸的，非光滑的。我们开发了一个算法的基础上的 ADMM 框架来解决问题，它分为四个子问题的原始模型。此外，所有这些子问题都有封闭形式的解决方案。作为一个数值迭代算法，我们提出的算法的收敛性进行了研究。最后，在 4 个包含不同指标的

真实数据集上进行了实验，结果表明，该方法在稀疏投资组合优化中的性能优于现有的几种方法。未来可通过优化对不同资产的增长动力的预测，来更加精准捕捉各资产的增长趋势。

参考文献

- [1] Allan Borodin, Ran El-Yaniv, and Vincent Gogan. Can we learn to beat the best stock. *Advances in Neural Information Processing Systems*, 16, 2003.
- [2] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, Jonathan Eckstein, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, 3(1):1–122, 2011.
- [3] Hong-Liang Dai, Chu-Xin Liang, Hong-Ming Dai, Cui-Yin Huang, and Rana Muhammad Adnan. An online portfolio strategy based on trend promote price tracing ensemble learning algorithm. *Knowledge-Based Systems*, 239:107957, 2022.
- [4] Puja Das, Nicholas Johnson, and Arindam Banerjee. Online lazy updates for portfolio selection with transaction costs. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 27, pages 202–208, 2013.
- [5] Puja Das, Nicholas Johnson, and Arindam Banerjee. Online portfolio selection with group sparsity. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 28, 2014.
- [6] John Duchi, Shai Shalev-Shwartz, Yoram Singer, and Tushar Chandra. Efficient projections onto the l_1 -ball for learning in high dimensions. In *Proceedings of the 25th international conference on Machine learning*, pages 272–279, 2008.
- [7] Frank J Fabozzi, Harry M Markowitz, and Francis Gupta. Portfolio selection. *Handbook of finance*, 2, 2008.
- [8] Zhao-Rong Lai, Pei-Yi Yang, Liangda Fang, and Xiaotian Wu. Short-term sparse portfolio optimization based on alternating direction method of multipliers. *The Journal of Machine Learning Research*, 19(1):2547–2574, 2018.
- [9] Bin Li, Steven C. H. Hoi, Peilin Zhao, and Vivekanand Gopalkrishnan. Confidence weighted mean reversion strategy for online portfolio selection. *ACM Trans. Knowl. Discov. Data*, 7(1):38, 2013.
- [10] Bin Li and Steven CH Hoi. Online portfolio selection: A survey. *ACM Computing Surveys (CSUR)*, 46(3):1–36, 2014.
- [11] Bin Li, Steven CH Hoi, Doyen Sahoo, and Zhi-Yong Liu. Moving average reversion strategy for on-line portfolio selection. *Artificial Intelligence*, 222:104–123, 2015.

- [12] William F Sharpe. Mutual fund performance. *The Journal of business*, 39(1):119–138, 1966.
- [13] Jack L Treynor and Fischer Black. How to use security analysis to improve portfolio selection. *The journal of business*, 46(1):66–86, 1973.
- [14] Linlong Wu, Yiyong Feng, and Daniel P Palomar. General sparse risk parity portfolio design via successive convex optimization. *Signal Processing*, 170:107433, 2020.