

求解大规模稀疏 Min-CVaR 投资组合模型的对偶方法

王云龙, 沈春根
(上海理工大学 理学院, 上海 200093)

摘要: 投资组合管理需要在资产池中挑选合适的标的资产并确定投资组合配资比例, 相关计算的时效性非常重要。基于 L_1 范数的稀疏 Min-CVaR (Minimum Conditional Value at Risk) 模型可以同步完成挑选标的资产和确定配资比例, 由于增加稀疏项, 相比于标准 Min-CVaR 模型求解更复杂。为了有效求解大规模问题, 基于原始模型中约束的结构特征, 构造了其拉格朗日对偶模型, 通过最新的商业求解器 GUROBI 12.8 求解对偶模型完成计算工作。使用基于三因子模型模拟的大规模情景数据 (50 000 行 200 列) 和标普 500 高频交易数据 (28 805 行 483 列) 进行了数值实验, 结果显示对偶模型可以数倍提升计算效率, 甚至比直接求解不具有稀疏性的标准 Min-CVaR 模型更快。

关键词: 条件在险价值; 投资组合模型; 风险管理; 稀疏优化; 对偶模型
中图分类号: O 022 **文献标志码:** A

A dual method for solving large-scale sparse Min-CVaR portfolio selection model

WANG Yunlong, SHEN Chungen
(College of Science, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: Portfolio management requires selecting the right underlying asset in the asset pool and determining the proportion of the portfolio. The timeliness of calculations of the portfolio models is extremely important. The L_1 norm sparse Min-CVaR (Minimum-CVaR, Conditional Value at Risk) model is able to simultaneously select the target assets and determine the capital allocation. Meanwhile, the L_1 norm sparse Min-CVaR is harder to handle owing to the L_1 norm regularization. In order to solve large-scale problems effectively, a Lagrangian dual model based on the structural features of constraints in the original model was constructed and solved by the state-of-the-art commercial solver GUROBI 12.8. Numerical experiments were conducted on the simulated data (50 000 rows and 200 columns) and real-world data (the S&P 500: 28 805 rows and 483 columns). The numerical results show that, compared to the primal model, the dual method is several times faster than the primal method, and it is even faster than the standard Min-CVaR portfolio selection model without regularization.

Keywords: conditional value at risk; portfolio selection model; risk management; sparse optimization; dual model

1 背景与现状

随着我国金融市场的形成和发展壮大, 金融投资已成为我国居民日常生活中的一个热门话题。金融投资组合可以通过挑选特定的投资标的资产, 在有限的投资资本下, 降低风险并提高期望收益。这种在风险与收益之间的权衡, 关键在于不同投资标的之间配资比例的确定。经典的马科维兹均值-方差模型 (Mean-Variance, MV) 将投资组合中标的资产的配资问题转化为一个带约束的二次规划问题^[1], 通过求解此最优化问题, 确定在获得期望收益时使风险最小或在可承受风险内期望收益最大的投资配比^[2]。后续的研究中, 不少学者在此基础上继续探索如何使风险最小并且收益最大^[3-6]。在 MV 模型中, 风险通过方差定义, 并且假设标的资产的收益率服从正态分布或椭圆分布, 这种假设在金融投资领域应用中并不完全可行。首先, 由于方差的对称性, 无论正向偏离还是负向偏离均会被同等考虑进模型中, 即如果减少方差会带来正向偏离和负向偏离的同步减小, 投资组合的收益同时被对风险的管理所影响, 因而以方差作为风险度量的 MV 模型一直广受诟病^[7]。其次, MV 模型需要计算所有标的资产收益率的协方差矩阵, 这在大规模场景下会使计算成本急剧增加, 求解极其困难。

现实中的风险度量倾向于区分正向偏离和负向偏离^[8]。Markowitz^[9]曾提出半方差 (Semi-Variance) 模型度量低于平均水平的收益波动性 (风险)。在险价值 (Value-at-Risk, VaR)^[10]和期望损失 (Expected Shortfall, ES, 或称条件在险价值, Conditional Value-at-Risk, CVaR)^[11]是目前为流行的风险度量, 它们注重于对尾部损失 (风险) 的度量^[12]。对 VaR 和 CVaR 数学性质的分析和比较可见 Rockafellar 和 Uryasev 的文章^[13]及 Pflug 的相关书籍^[14]。虽然 CVaR 作为 VaR 的补充统计量而非可以替代 VaR, 但值得指出的是, CVaR 具有一致性风险度量的良好性质, 它考虑了超 VaR 部分的损失, 对尾部风险的衡量更准确和稳定^[14-16]。求解 Min-CVaR 模型实际上是一个线性规划问题, 而当输入情景数据 (行数) 过多时, 此线性规划需要增加等量级的辅

助变量, 且约束的数目也急剧增加, 这对一般线性规划求解而言极具困难。实际上, 使用最新的商业求解器 (如 CPLEX、GUROBI) 求解, 当情景数据在十万量级, 资产个数在百千量级时, 需要花费数个小时。为了避免这种情况, 不少学者设计了各种解决思路, 如情景分解方法^[17-18]、二阶段方法^[19]、切平面法^[20]。上述研究均没有考虑解的稀疏性要求。稀疏性目标可以利用一系列稀疏方法优化实现, 通过添加惩罚项或者增加约束, 使优化问题的解具有较少的非零元素。常见的具有稀疏诱导性质的惩罚项有 L_1 范数和 L_0 范数。 L_0 可以直接衡量解中非零元的个数, 稀疏效果最好, 而实际上是 NP-hard 的问题, 求解不够高效。文献 [21] 研究以 L_1 范数为惩罚项的稀疏马科维兹投资组合模型, 并分析了引入 L_1 范数带来的理论影响和实际意义。本文选择 L_1 范数作为惩罚项加入 Min-CVaR 模型的目标函数中, L_1 可以取得较少的非零元, 相当于挑选较少的标的资产。相比于 L_0 范数, L_1 范数会惩罚较大的元素, 相当于分散配资比例避免极端头寸。

稀疏模型虽然具有以上优点, 但增加稀疏约束或惩罚项会使模型更加复杂, 尤其在数据规模很大的情况下。目前国内外对稀疏投资组合模型的研究^[22-24]中, 情景数据的规模仍为几百个或一千多个, 在大数据时代, 尤其是高频交易场景下, 利用超大规模历史情境数据建立投资模型是一种趋势。文献 [25] 曾利用对偶方法求解投资组合优化问题, 文献 [6] 利用 SCAD 函数作为惩罚项并利用 QR 技术求解稀疏的投资组合优化问题, 但文献 [6] 中的原始模型很难通过对偶模型加速。本文在文献 [25] 的基础上改进了求解稀疏投资组合问题的优化模型。相比之下, 本文引入 L_1 范数作为稀疏项有利于对偶模型求解。 L_1 范数作为惩罚项, 会使原本投资组合优化模型变得复杂, 如 L_1 范数具有非光滑性。关于带 L_1 范数的优化问题一直被广泛关注^[26-28]。此外, 不同于文献 [25], 本文增加了允许卖空的市场环境。因而, 利用对偶思想求解稀疏投资组合模型, 通过变换原始问题, 降低了求解难度, 可以极大地提高大规模稀疏 Min-CVaR 投资组合模型的计算效率, 这对在大数据时代高频

交易场景下高效构建投资组合模型具有一定的指导借鉴意义。

2 稀疏 Min-CVaR 模型

2.1 基本概念

本文假设投资组合中有 N 个标的资产，样本的时间长度为 T ，投资组合的数据观测记为 $\mathbf{R}=[\mathbf{R}_1,\cdots,\mathbf{R}_N]\in\mathbb{R}^{T\times N}$ ，资产 i 在时间 t 的收益率记为 $r_{ti}(i=1,2,\cdots,N;t=1,2,\cdots,T)$ ，假设资产 i 的配资比例(权重)记为 $w_i(\mathbf{w}=(w_1,\cdots,w_N)^T\in\mathbb{R}^N)$ ，则投资组合在 t 时刻的收益率为

$$\mathbf{y}(\mathbf{w})=\mathbf{R}\mathbf{w}=\left[\sum_{i=1}^Nr_{ti}w_i\right]_{t=1,2,\cdots,T}$$

(1)

式中： $\mathbf{y}(\mathbf{w})\in\mathbb{R}^T$ ；记 $p_t\in[0,1]$ 为过去每个时刻可能发生的概率($\mathbf{p}\in\mathbb{R}^T,\mathbf{1}^T\mathbf{p}=1$)，并假设 $p_t=1/T$ ，即所有时刻重现的概率相等，则投资组合的期望收益率为

$$\mu(\mathbf{w})=\mathbf{p}^T\mathbf{R}\mathbf{w}=\sum_{t=1}^Ty_t(\mathbf{w})p_t$$

(2)

上述 $\mathbf{1}$ (后续的 $\mathbf{0}$)为维数与上下文匹配的全 $\mathbf{1}$ ($\mathbf{0}$)向量或矩阵。投资组合各标的资产间的协方差矩阵记为 $\mathbf{\Sigma}=\text{cov}(\mathbf{R})\in\mathbb{R}^{N\times N}$ ， $\mathbf{\Sigma}$ 具有对称性和半正定性。整个投资组合收益率的方差为 $\sigma^2(\mathbf{w})=\mathbf{w}^T\mathbf{\Sigma}\mathbf{w}$ 。

2.2 条件在险价值 CVaR

在险价值 VaR 有 3 个关键要素：置信水平或概率 α ，时间段 Δt ，损失值 VaR_α 。即基于当前时刻 t 的信息，预测在一定置信水平 α 下经过时间段 Δt 后的上限损失值，即未来 Δt 时间后损失超过 $VaR_\alpha(\mathbf{w})$ 的概率为 $1-\alpha$ 。离散情况下的数学表达为

$$VaR_\alpha(\mathbf{w})=-\inf\left\{y_K(\mathbf{w})\left|\sum_{k=1}^Kp_k\geq 1-\alpha\right.\right\}$$

(3)

式中： $y_k(\mathbf{w}),k=1,2,\cdots,K$ 满足 $y_1(\mathbf{w})\leq y_2(\mathbf{w})\leq\cdots\leq y_K(\mathbf{w})$ ； p_k 为 $y_K(\mathbf{w})$ 对应的概率。CVaR 定义为超过 VaR 损失的期望值，离散情况下 CVaR 的数学表达为

$$CVaR_\alpha(\mathbf{w})=\frac{1}{1-\alpha}\sum_{-y_t(\mathbf{w})>VaR_\alpha}-p_ty_t(\mathbf{w})$$

(4)

2.3 标准 Min-CVaR 模型

最小化式(4)等同于最小化式(5)，即

$$f(\mathbf{w},z)=(1-\alpha)z+\sum_{t=1}^Tp_t(-y_t(\mathbf{w})-z)^+$$

(5)

式中， $(\cdot)^+=\max(\cdot,0)$ ，并且由最小化式(5)得到的

z 值与 VaR 值相等^[13]，则标准的 Min-CVaR 模型可以表达为线性规划问题：

$$\min_{\mathbf{w}\in\mathbb{R}^N,\mathbf{s}\in\mathbb{R}^T,\zeta\in\mathbb{R}}\left\{\zeta+\frac{1}{1-\alpha}\sum_{t=1}^Tp_ts_t\right\}$$

s.t.
$$\begin{cases} s_t\geq -y_t(\mathbf{w})-\zeta, & t=1,\cdots,T \\ s_t\geq 0, & t=1,\cdots,T \\ \mathbf{1}^T\mathbf{w}=1 \end{cases}$$

(6)

因此，求解最小化模型(6)可以同时得到配资比例 $\mathbf{w}_{\text{opt}}$ 、在险价值 $VaR=\zeta_{\text{opt}}$ 和条件在险价值 $CVaR=f(\mathbf{w}_{\text{opt}},\zeta_{\text{opt}})/(1-\alpha)$ 。由于模型(6)引入辅助变量 $\mathbf{s}\in\mathbb{R}^T$ ，使得总的变量个数为 $T+N+1$ ，同时增加了 T 个结构复杂的约束，大规模情景下这无疑会使线性规划问题的求解难度增加。

2.4 稀疏 Min-CVaR 模型

在上述 Min-CVaR 模型中，配资比例是在所有标的资产之间计算的。首先模型求解是分散在所有标的资产上的，无法实现挑选标的资产的目的，其次，在假设允许卖空时，为了实现最优，模型会产生在某些标的资产上的极端头寸(比如看跌资产 i 并借10倍杠杆，看涨资产 j 并配资11倍)。通过增加 L_1 范数约束可以避免以上问题， $\|\mathbf{w}\|_1=\sum_{i=1}^N|w_i|\leq n$ ，会对数值较大的权重加以约束，可同时实现只挑选少数标的资产以及均衡被挑选标的资产间的配资比例，达到权重的稀疏性(较多零元)。稀疏 Min-CVaR 模型可以表达为

$$\min_{\mathbf{w}\in\mathbb{R}^N,\mathbf{s}\in\mathbb{R}^T,\zeta\in\mathbb{R}}\left\{\zeta+\frac{1}{1-\alpha}\sum_{t=1}^Tp_ts_t\right\}$$

s.t.
$$\begin{cases} s_t\geq -y_t(\mathbf{w})-\zeta, & t=1,\cdots,T \\ s_t\geq 0, & t=1,\cdots,T \\ \mathbf{1}^T\mathbf{w}=1 \\ \|\mathbf{w}\|_1\leq n \end{cases}$$

(7)

模型(7)中可以通过调整参数 n 控制解的稀疏度， n 越大，最优解中的非零元个数越多且分布范围越大。当 $n\rightarrow+\infty$ ，带 L_1 范数约束的 Min-CVaR 模型将退化为标准的 Min-CVaR 模型。在一定条件下，模型(7)的等价形式为

$$\min_{\mathbf{w}\in\mathbb{R}^N,\mathbf{s}\in\mathbb{R}^T,\zeta\in\mathbb{R}}P(\mathbf{w},\mathbf{s},\zeta)=\zeta+\frac{1}{1-\alpha}\sum_{t=1}^Tp_ts_t+\lambda\|\mathbf{w}\|_1$$

s.t.
$$\begin{cases} s_t\geq -y_t(\mathbf{w})-\zeta, & t=1,\cdots,T \\ s_t\geq 0, & t=1,\cdots,T \\ \mathbf{1}^T\mathbf{w}=1 \end{cases}$$

(8)

其中， $\lambda>0$ 为调和参数。对于模型(8)，其拉格朗

日函数为

$$\mathcal{L}(\zeta, s, \mathbf{w}; \mathbf{v}_1, \mathbf{v}_2, \eta) = \zeta + \frac{1}{1-\alpha} \sum_{t=1}^T p_t s_t + \lambda \|\mathbf{w}\|_1 - \mathbf{v}_1^T (\mathbf{s} + \mathbf{R}\mathbf{w} + \zeta \mathbf{1}) - \mathbf{v}_2^T \mathbf{s} - \eta (\mathbf{1}^T \mathbf{w} - 1) \quad (9)$$

其中, $\mathbf{v}_1 \in \mathbb{R}_+^T$, $\mathbf{v}_2 \in \mathbb{R}_+^T$ 和 $\eta \in \mathbb{R}$ 分别为模型(8)中约束条件对应的拉格朗日乘子向量。对拉格朗日函数求偏导并令偏导数为零可得最优性条件:

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial \zeta} = 1 - \mathbf{v}_1^T \mathbf{1} = 0 \\ \frac{\partial \mathcal{L}}{\partial \mathbf{s}} = \frac{1}{1-\alpha} \mathbf{p} - \mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0} \\ \mathbf{0} \in \lambda \partial \|\mathbf{w}\|_1 - \mathbf{R}^T \mathbf{v}_1 - \eta \mathbf{1} \\ \mathbf{v}_1 \geq \mathbf{0}, \mathbf{v}_1^T (\mathbf{s} + \mathbf{R}\mathbf{w} + \zeta \mathbf{1}) = \mathbf{0}, \mathbf{s} + \mathbf{R}\mathbf{w} + \zeta \mathbf{1} \geq \mathbf{0} \\ \mathbf{v}_2 \geq \mathbf{0}, \mathbf{v}_2^T \mathbf{s} = \mathbf{0}, \mathbf{s} \geq \mathbf{0} \\ \mathbf{1}^T \mathbf{w} = 1 \end{cases} \quad (10)$$

其中, $\partial \|\mathbf{w}\|_1$ 表示 L_1 范数的次微分。

$$\partial \|\mathbf{w}\|_1 = \left\{ \mathbf{h} \in \mathbb{R}^N \mid \mathbf{h}_i \begin{cases} = 1, & w_i > 0 \\ \in [-1, 1], & w_i = 0, i = 1, \dots, N \\ = -1, & w_i < 0 \end{cases} \right\} \quad (11)$$

则模型(8)的拉格朗日对偶函数为

$$\begin{aligned} D(\eta, \mathbf{v}) &= \inf_{\zeta, s, \mathbf{w}} \mathcal{L}(\zeta, s, \mathbf{w}; \mathbf{v}_1, \mathbf{v}_2, \eta) = \\ &\begin{cases} \eta + \mathbf{R}_i^T \mathbf{v} \in [-\lambda, \lambda], i = 1, \dots, N & \text{(I)} \\ 0 \leq v_t \leq p_t \frac{1}{(1-\alpha)}, t = 1, \dots, T & \text{(II)} \\ \mathbf{1}^T \mathbf{v} = 1 & \text{(III)} \end{cases} \quad (12) \\ &-\infty, \text{其他} \end{aligned}$$

得对偶模型:

$$\max_{\eta \in \mathbb{R}, \mathbf{v} \in \mathbb{R}^T} \eta, \quad \text{s.t. } \mathbf{v} \in \mathbf{Dom} \quad (13)$$

其中, $\mathbf{Dom}$ 为满足式(12)中约束(I)~(III)的集合。模型(13)中变量总数为 $T+1$, 约束总数为 $2(T+N)+1$ 。实际上, 由于约束(II)为简单的盒子约束(box-constraints), 不会影响求解问题的复杂程度, 所以

$$\begin{aligned} \mu_F &= \begin{pmatrix} 0.7828 \\ 0.5180 \\ 0.4100 \end{pmatrix}, \quad \Sigma_F = \begin{pmatrix} 0.02914 & 0.02387 & 0.01018 \\ 0.02387 & 0.05395 & -0.00696 \\ 0.01018 & -0.00696 & 0.08685 \end{pmatrix} \\ \mu_B &= \begin{pmatrix} 0.02355 \\ 0.01298 \\ 0.02071 \end{pmatrix}, \quad \Sigma_B = \begin{pmatrix} 1.2507 & -0.0350 & -0.2042 \\ -0.0350 & 0.3156 & -0.0023 \\ -0.2042 & -0.0023 & 0.1930 \end{pmatrix} \\ a_\varepsilon &= 3.3286, \quad b_\varepsilon = 0.1876 \end{aligned}$$

文献[30]进行了大规模情景数据模拟实验($T=50\,000$), 本文也采用相似规模设置, 不同的是, 本文收益矩阵由式(14)描述的三因子模型生成。

通过表1可以发现在计算大规模情景数据

只有结构稍微复杂的约束(I)和(III)会影响问题求解, 因而影响对偶模型(13)求解难度的约束个数只受股票池规模(矩阵 $\mathbf{R}$ 的列数 N)的影响。当计算大规模情景数据时($T \gg N$), 求解原始模型的难度骤升, 而通过对偶方式则会极大地提升求解效率。

3 数值实验

在数值实验中, 本文使用模拟数据和真实数据测试求解对偶模型策略的时间优越性。首先介绍了利用三因子模型模拟生成大规模数据, 以及选择调和参数的具体步骤, 结果显示, 在模拟数据上, 随着情景规模的增加, 对偶模型提升求解效率的表现更显著。另外, 给出了在真实高频数据(标普500)上的试验结果, 同时和标准 Min-CVaR 模型进行了对比, 结果显示, 稀疏模型可以明显地增加解的零元个数, 达到挑选标的资产的目的, 对偶模型能够有效提升求解效率。本文试验环境为 MATLAB2017b, Inter(R) Core(TM) i5-3230M, CPU@2.6 GHz, RAM 8 GB, GURIOBI 求解器通过 MATLAB2017b 调用。

3.1 模拟数据

假设收益矩阵记为 $\mathbf{R} \in \mathbb{R}^{T \times N}$, 本文根据经典的三因子模型^[29]生成随机的收益矩阵为

$$\mathbf{R}_{T \times N} = \mathbf{F} \mathbf{B}_{T \times 3}^T + \boldsymbol{\varepsilon}_{T \times N} \quad (14)$$

式中: $\mathbf{F}$ 为因子矩阵, 服从多元正态分布 $\mathcal{N}(\mu_F, \Sigma_F)$; $\mathbf{B}$ 为因子载荷矩阵, 服从多元正态分布 $\mathcal{N}(\mu_B, \Sigma_B)$; 假设噪声矩阵 $\boldsymbol{\varepsilon}$ 与因子矩阵独立并且服从伽马分布 $\Gamma(a_\varepsilon, b_\varepsilon)$ 。上述的分布参数选择依据文献[29]中的表1, 这里重新给出:

时, 对偶方法在计算效率上具有明显的优越性。表1统计了每个问题计算20个 λ 对应的问题的总耗时, $\lambda \in \lambda_{\max} [10^{-3}, 1]$, 其中 $\lambda_{\max}$ 是可确定的合适的调和参数, 使得 $\forall \lambda \geq \lambda_{\max}$ 时, 原问题(8)的最优解不再变化。实际上, 当 λ 非常小的时候, 对偶模型

表 1 不同概率水平下计算整个λ路径所需时间 (T=50 000)

Tab.1 Time consuming for solving a path of λ with different probability levels

N	模型	α = 0.99	α = 0.95	α = 0.90	α = 0.80
20	原始	227.1	237.5	203.3	209.8
	对偶	43.4	42.3	43.4	45.3
50	原始	300.4	277.7	258.3	265.7
	对偶	103.8	104.0	106.2	107.7
100	原始	556.0	517.3	484.2	461.5
	对偶	237.9	255.4	259.5	261.9
200	原始	1 143.3	1 153.6	1 063.1	977.4
	对偶	589.1	607.1	609.1	601.8

的求解效率提升不太明显，稍微比原始问题快 2~3 倍；当λ较大的时候，效率提升至 10 倍左右（依赖于收益矩阵的规模，情景维数越大，提升效果越显著）。具体地，随着λ的增大，求解原始模型的耗时呈上升趋势，求解对偶模型的耗时呈下降趋势。图 1 展示了 T = 50 000, N = 20, α = 0.9 时，λ大小对求解时间的影响（其中 i 对应于 λ 序列的位次， $i/20 = -3\lg(\lambda/\lambda_{\max})$, $i = 1, \dots, 20$, i 越小，对应的 λ 越大）。从数学优化的角度看，当 $\lambda \rightarrow 0$ 时， L_1 范数的影响降低，原始模型逐渐趋向于标准 Min-CVaR 模型，即模型变得简单，而对偶模型中约束 (I) 的上下界间隙越来越小，即模型变得复杂。因而当 λ 非常小时，求解对偶模型的效率提升相对较小。一般情况下，模型倾向于简洁时，会选取稍大的 λ 值，此时，求解对偶模型的效率提升相对较大。

3.2 参数选择

在原始模型和对偶模型中，参数λ扮演着不同

的角色，但都会导致解的稀疏性(0 除外)。λ在原始模型中被称为调和参数，具有平衡稀疏性和最小化目标函数的作用。调和参数越大，获得的解稀疏性越高(零元越多)。在对偶模型中λ是复杂结构约束(I)的上下界，影响着约束的严苛程度，根据对偶性质，原始问题也是其对偶问题的对偶问题，原始变量与对偶问题的对偶变量间存在一一对应关系。较大的λ意味着对偶问题中约束的上下界距离较远，则积极约束的个数会相应减少，积极约束集的大小与对偶问题的对偶变量非零元个数一致，即原始问题的解非零元少(稀疏性大)。模型参数的选择有不同的准则，如交叉检验(Cross-Validation)^[31]、赤池弘次(Akaike)信息准则(AIC)^[32]。本文选取贝叶斯信息准则(BIC)^[33-34]作为判断不同参数下模型好坏的依据，BIC 越小则模型越优越。记 w^λ 为问题(8)的最优解，则

$$\hat{\lambda} = \arg \min_{\lambda} BIC(\lambda) = \sum_{t=1}^T s_t + \|w^\lambda\|_0 \frac{\ln T \ln N}{2T} \quad (15)$$

其中，零范数 $\|w^\lambda\|_0$ 衡量 w^λ 中非零元的个数。从统计模型和实际操作的角度，大家都倾向于选择更简单的模型。结合 Tibshirani^[26] 的一个标准差最简原则，即可确定既满足 BIC 准则下模型优越性又兼具模型简洁性的参数 $\hat{\lambda}$ 。图 2 中的两条竖线分别对应 BIC 准则下的最优参数(图 2(a)右侧竖线)和兼具模型简洁性的参数(图 2(a)左侧竖线)。根据图 2(b)可以看出，λ越大，解中非零元的个数 $\|w\|_0$ 越少，依据一个标准差最简原则挑选了较大的 λ，因而会得到更稀疏的解。根据图 3(a)可知，较大的λ能够得到兼具模型简洁性解，即挑选更少的标的资产。

3.3 结果分析

现给出其他维数规模的模拟数据实验结果，仍然用式(15)的三因子模型生成模拟数据，情景维数T分别为 1 000，5 000 和 10 000，资产维数N分别为 50 和 100，给出了 3 个置信水平(α= 0.99, 0.95, 0.90)下的计算结果。其中，总时间表示为求解 3 个置信水平下的最优模型所用的总时间。在实际操作中，资产池和数据的情景维度一般是保持不变的，因此，最优调和参数λ可事先设定(表 2)，亦无须变化，这缓解了必须求解多个逐渐变化的λ才可以确定模型优越性的压力。根据前文的分析，实验中利用的对偶模型，主要受资产池规模的影响，情景规模对对偶模型的求解效率影响不

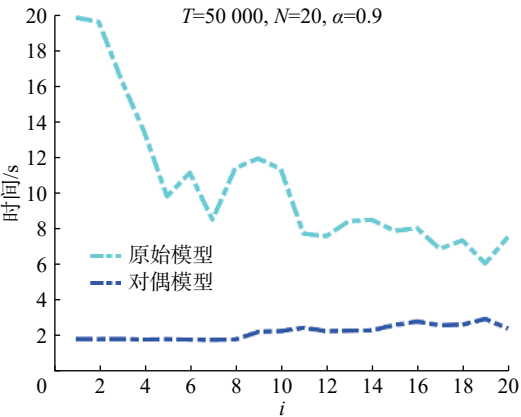


图 1 求解对偶模型与原始模型的耗时变化

Fig. 1 Time paths for solving the primal and dual model

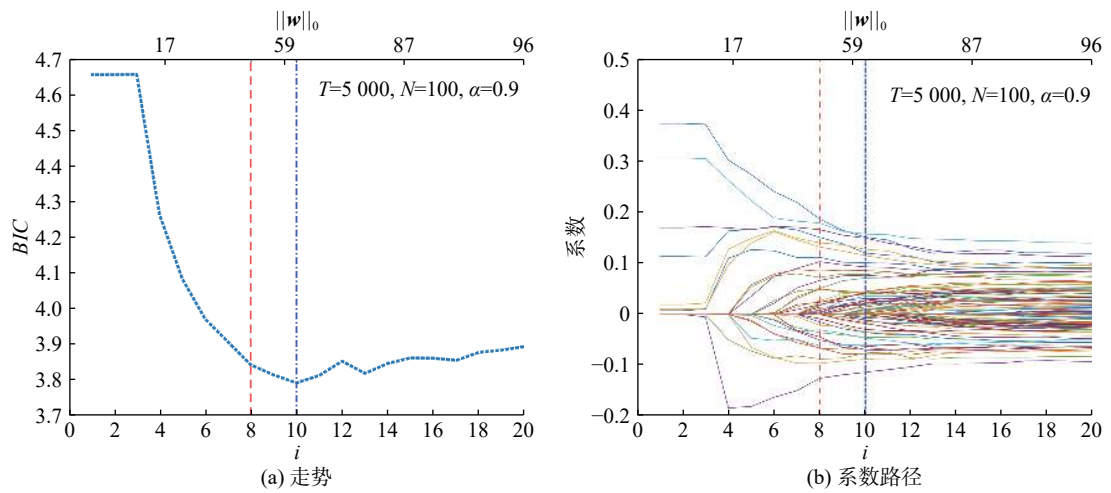


图 2 BIC 走势图与投资组合系数路径图

Fig.2 BIC path and portfolio weights solution paths

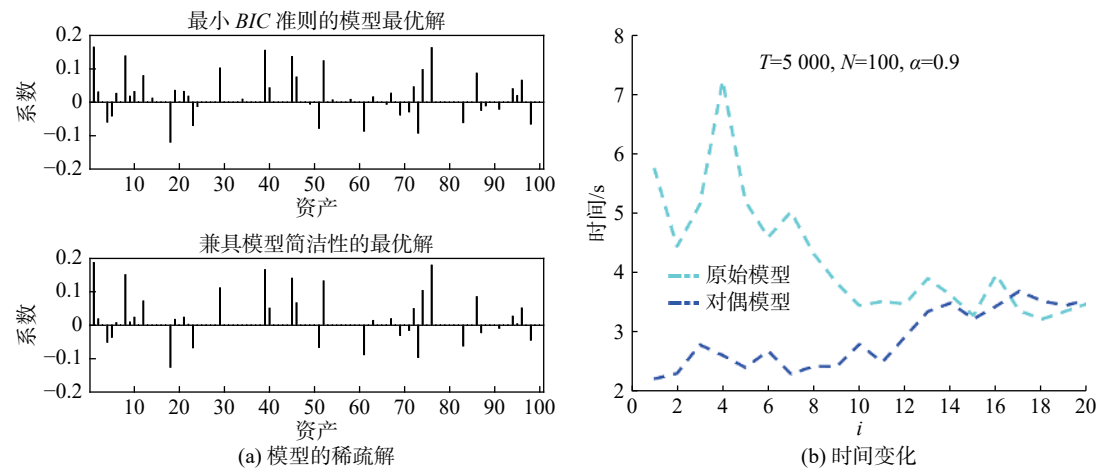


图 3 模型的稀疏解和时间变化图

Fig.3 Sparse weights solutions; Time paths for solving primal and dual model

大。实验结果表明(表 3 第 3 列),随着情景规模的增加,对偶模型在求解效率上的优越性表现更明显。表 3 中,EWP(equal weight portfolio)指等权重组合策略。

3.4 真实数据

本文采用标普 500 高频交易数据作为真实数据进行了实验。资产池共有 483 个资产,情景维数为 25 805,(391 条/d,共 66 d)。由于高频交易

数据本身的性质,价格变动幅度小,收益率量级普遍较小,在结果分析中,均将收益率转换为日收益率格式。

3.5 结果分析

表 4 给出了在真实高频数据——大规模情景数据上的实验结果。表 4 所用的参数 λ 同样遵循 BIC 准则下兼具模型简洁性的原则。相比于标准 Min-CVaR 模型,本文提出的对偶策略提高了求解效率,同时兼顾了稀疏性以利于同步筛选投资标的。稍微遗憾的是得到的 CVaR 比 Min-CVaR 大,尽管差别不大。这也可以从数学优化的角度解释,由表 4 可知 Min-CVaR 的最优解稀疏度非常小,因为 Min-CVaR 利用资产池里几乎所有标的资产构建投资组合,势必会得到更优的目标函数。因此,Min-CVaR 不具有同步挑选标的资产的特性。本文原始模型的解具有稀疏性,兼顾了自动挑选标的的特性,

表 2 最优模型对应的调和参数

概率 α	例子规模($T N$)					
	1000 50	1000 100	5000 50	5000 100	10000 50	10000 100
0.99	0.0397	0.2506	0.0792	0.0079	0.0251	0.0158
0.95	0.2506	0.0315	0.0500	0.2506	0.0251	0.0251
0.90	0.1991	0.1581	0.0629	0.1991	0.0500	0.0629

表 3 在模拟算例上的表现

Tab.3 Portfolio performance on a variety of simulation data cases

例子规模($T N$)	方法	总时间/s	$\alpha = 0.99$		$\alpha = 0.95$		$\alpha = 0.90$	
			CVaR	VaR	CVaR	VaR	CVaR	VaR
1 000 50	EWP	0.001 6	1.569 5	1.264 9	1.086 0	0.715 5	0.819 0	0.450 4
	Min-CVaR	1.689 4	-0.174 1	-0.174 1	-0.214 3	-0.286 1	-0.271 3	-0.370 2
	原始	1.684 2	-0.167 9	-0.167 9	-0.083 0	-0.196 3	-0.173 2	-0.281 3
	对偶	0.316 0	-0.167 9	-0.167 9	-0.083 0	-0.196 3	-0.173 2	-0.281 3
1 000 100	EWP	0.001 8	1.620 0	1.290 4	1.064 2	0.724 4	0.810 2	0.373 2
	Min-CVaR	2.584 7	-0.417 2	-0.417 2	-0.423 5	-0.433 7	-0.435 8	-0.469 4
	原始	2.899 3	-0.274 8	-0.274 8	-0.421 4	-0.431 6	-0.347 0	-0.412 6
	对偶	0.702 2	-0.274 8	-0.274 8	-0.421 4	-0.431 6	-0.347 0	-0.412 6
5 000 50	EWP	0.006 0	1.678 4	1.333 5	1.104 4	0.763 5	0.848 4	0.430 5
	Min-CVaR	5.525 6	-0.081 7	-0.124 1	-0.180 2	-0.257 0	-0.241 4	-0.338 4
	原始	5.688 4	-0.054 4	-0.096 0	-0.162 6	-0.244 6	-0.217 9	-0.321 7
	对偶	1.506 6	-0.054 4	-0.096 0	-0.162 6	-0.244 6	-0.217 9	-0.321 7
5 000 100	EWP	0.010 2	1.631 2	1.361 7	1.138 9	0.775 5	0.868 9	0.443 4
	Min-CVaR	9.505 2	-0.284 8	-0.297 3	-0.339 4	-0.390 7	-0.377 7	-0.440 9
	原始	12.025 7	-0.054 4	-0.296 9	-0.162 6	-0.259 4	-0.217 9	-0.336 7
	对偶	2.585 2	-0.054 4	-0.296 9	-0.162 6	-0.259 4	-0.217 9	-0.336 7
10 000 50	EWP	0.008 4	1.608 2	1.279 8	1.083 6	0.724 6	0.827 4	0.431 9
	Min-CVaR	10.604 6	-0.085 3	-0.142 5	-0.198 0	-0.284 0	-0.262 0	-0.362 1
	原始	15.366 6	-0.081 1	-0.135 6	-0.192 5	-0.278 3	-0.247 6	-0.348 1
	对偶	3.042 9	-0.081 1	-0.135 6	-0.192 5	-0.278 3	-0.247 6	-0.348 1
10 000 100	EWP	0.018 7	1.623 7	1.342 7	1.112 4	0.762 8	0.853 8	0.452 5
	Min-CVaR	20.478 4	-0.247 0	-0.278 3	-0.319 8	-0.378 9	-0.365 0	-0.434 7
	原始	25.151 1	-0.244 5	-0.278 2	-0.312 6	-0.374 2	-0.340 4	-0.417 6
	对偶	6.296 2	-0.244 5	-0.278 2	-0.312 6	-0.374 2	-0.340 4	-0.417 6

表 4 在真实算例上的表现

Tab.4 Portfolio performance on real-world data cases

方法	α	耗时/s	CVaR	VaR	稀疏度/%
EWP	0.99	0.1	0.985 4	0.573 6	0
	0.95	0.1	0.527 0	0.323 4	0
	0.90	0.1	0.397 3	0.223 2	0
Min-CVaR	0.99	233.6	0.256 2	0.255 5	0.4
	0.95	268.0	0.191 6	0.136 3	0.4
	0.90	246.3	0.151 9	0.093 4	0.4
原始模型	0.99	384.3	0.273 1	0.241 4	13.5
	0.95	357.2	0.207 6	0.136 9	48.7
	0.90	336.9	0.180 7	0.101 3	82.7
对偶模型	0.99	159.2	0.273 1	0.241 4	13.5
	0.95	101.7	0.207 6	0.136 9	48.7
	0.90	96.2	0.180 7	0.101 3	82.7

同时利用对偶模型求解, 数倍地节省了原有模型的求解时间。

表 5 给出了在真实数据上, 不同概率下稀疏 Min-CVaR 模型所确定的两个调和参数, 其中最小 BIC 准则的模型对应的调和参数为 $\hat{\lambda}_1$, 最小 BIC 的基础上, 为了使模型更加精简, 遵循一个标准差

内最简洁原则确定的模型调和参数为 $\hat{\lambda}_2$ 。观察可知 $\hat{\lambda}_2$ 比 $\hat{\lambda}_1$ 更大, 根据前述分析, 越大的调和参数将导致越稀疏的解。观察表 5 和图 4 均可发现, 较大的调和参数 $\hat{\lambda}_2$ 对应的模型的解具有更高的稀疏度, 而相比于 $\hat{\lambda}_1$, 它们对应的风险度量(分别参考表 5 中 VaR 与 CVaR 值)相差并不明显。

表 5 最小 BIC 准则对应的调和参数与兼具模型简洁性的调和参数
Tab.5 Tuning parameters corresponding to minimum BIC and the ones with model simplicity

α	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\zeta_1(\text{VaR}_1)$	$\zeta_2(\text{VaR}_2)$	CVaR ₁	CVaR ₂	稀疏度1/%	稀疏度2/%
0.99	1.6×10^{-3}	4.5×10^{-3}	0.2358	0.2414	0.2599	0.2731	11.6	13.5
0.95	1.2×10^{-3}	4.7×10^{-3}	0.1345	0.1369	0.1937	0.2076	19.2	48.7
0.90	7.0×10^{-3}	9.8×10^{-3}	0.0983	0.1013	0.1731	0.1807	67.1	82.7

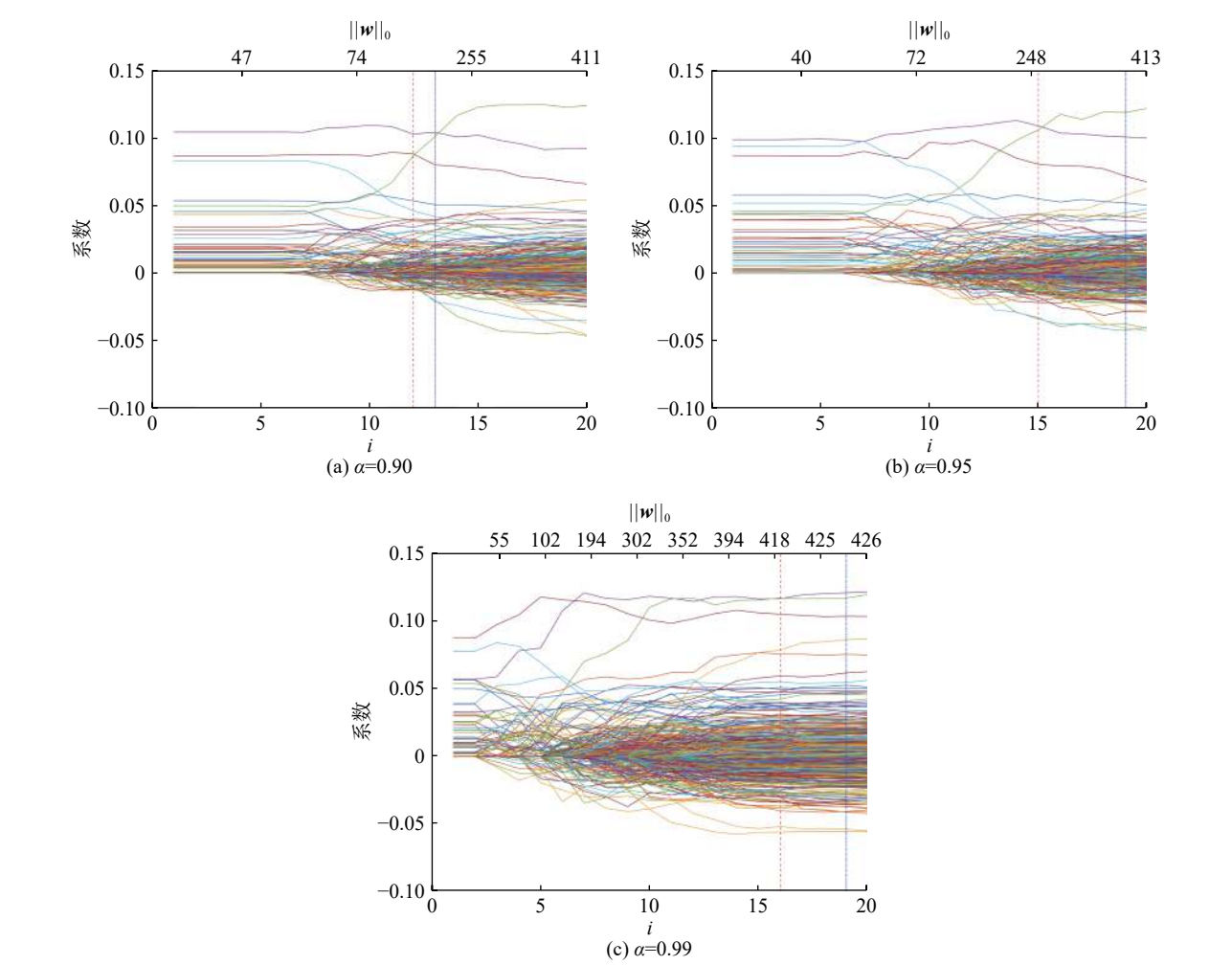


图 4 权重路径图
Fig.4 Portfolio weights solution paths

4 结 论

本文研究了大规模情景数据背景下的 L_1 范数稀

疏 Min-CVaR 投资组合模型。 L_1 范数稀疏 Min-CVaR 模型可以同步完成从资产池中挑选标的资产和确定各标的资产配置比例的工作, 且避免极端配资

的情况, 但此模型求解复杂性较高。本文通过对原始模型进行变换, 得到对偶模型, 利用单纯形法求解对偶模型得到最优解继而求得原始模型的最优解。本文对对偶模型和原始模型求解效率差异的理论原因进行了分析和解释, 同时通过数值试验证实了求解对偶模型的策略在时间上的优越性。在计算大规模数据时, 情景维数越大, 效率提升的效果越显著。

参考文献:

- [1] MARKOWITZ H. Portfolio selection[J]. *The Journal of Finance*, 1952, 7(1): 77–91.
- [2] MAYANJA F, MATARAMVURA S, CHARLES W M. A mathematical approach to a stocks portfolio selection: the case of uganda securities exchange (use)[J]. *Journal of Mathematical Finance*, 2013, 3(4): 487–501.
- [3] LI W C, WU Y, OJIAKO U. Using portfolio optimisation models to enhance decision making and prediction[J]. *Journal of Modelling in Management*, 2014, 9(1): 36–57.
- [4] MOHR E, DOCHOW R. Risk management strategies for finding universal portfolios[J]. *Annals of Operations Research*, 2017, 256(1): 129–147.
- [5] POLAK G G, ROGERS D F, SWEENEY D J. Risk management strategies via minimax portfolio optimization[J]. *European Journal of Operational Research*, 2010, 207(1): 409–419.
- [6] XU Q F, ZHOU Y Y, JIANG C X, et al. A large CVaR-based portfolio selection model with weight constraints[J]. *Economic Modelling*, 2016, 59: 436–447.
- [7] LWIN K T, QU R, MACCARTHY B L. Mean-VaR portfolio optimization: a nonparametric approach[J]. *European Journal of Operational Research*, 2017, 260(2): 751–766.
- [8] BIGLOVA A, ORTOBELLI S, RACHEV S T, et al. Different approaches to risk estimation in portfolio theory[J]. *The Journal of Portfolio Management*, 2004, 31(1): 103–112.
- [9] MARKOWITZ H M. Portfolio selection, efficient diversification of investments[J]. *Journal of the Institute of Actuaries*, 1992, 119(1): 165–166.
- [10] BASAK S, SHAPIRO A. Value-at-risk-based risk management: optimal policies and asset prices[J]. *The Review of Financial Studies*, 2001, 14(2): 371–405.
- [11] ROCKAFELLAR R T, URYASEV S. Optimization of conditional value-at-risk[J]. *Journal of Risk*, 2000, 2(3): 21–42.
- [12] WANG D H, SONG J B, LIN Y Z. Does the VaR measurement using monte-carlo simulation work in China?—evidence from Chinese listed banks[J]. *Journal of Financial Risk Management*, 2017, 6(1): 66–78.
- [13] ROCKAFELLAR R T, URYASEV S. Conditional value-at-risk for general loss distributions[J]. *Journal of Banking & Finance*, 2002, 26(7): 1443–1471.
- [14] PFLUG G C. Some remarks on the value-at-risk and the conditional value-at-risk[M]//URYASEV S. *Probabilistic Constrained Optimization: Methodology and Applications*. Dordrecht: Kluwer Academic Publishers, 2000.
- [15] SAWIK B T. Conditional value-at-risk vs. value-at-risk to multi-objective portfolio optimization[J]. *Applications of Management Science*, 2012, 15: 277–305.
- [16] URYASEV S, SARYKALIN S, SERRAINO G, K K. VaR vs CVaR in risk management and optimization[C]//CARISMA Conference, 2010. La Grande Motte, 2010.
- [17] KONNO H, WAKI H, YUUKI A. Portfolio optimization under lower partial risk measures[J]. *Asia-Pacific Financial Markets*, 2002, 9(2): 127–140.
- [18] RYSZ M, VINEL A, KROKHMAL P, et al. A scenario decomposition algorithm for stochastic programming problems with a class of downside risk measures[J]. *Inform Journal on Computing*, 2015, 27(2): 416–430.
- [19] FÁBIÁN C I. Handling CVaR objectives and constraints in two-stage stochastic models[J]. *European Journal of Operational Research*, 2008, 191(3): 888–911.
- [20] DENTCHEVA D, RUSZCZYŃSKI A. Inverse cutting plane methods for optimization problems with second-order stochastic dominance constraints[J]. *Optimization*, 2010, 59(3): 323–338.
- [21] BRODIE J, DAUBECHIES I, DE MOL C, et al. Sparse and stable markowitz portfolios[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2009, 106(30): 12267–12272.
- [22] 吕智翔, 孟志青. 稀疏化 Mean-CVaR 模型的应用研究 [J]. *经营与管理*, 2016(12): 111–113.
- [23] 陶一鑫. 基于正则化的投资组合分析 [D]. 杭州: 浙江工商大学, 2014.
- [24] CHENG R, GAO J J. On cardinality constrained mean-CVaR portfolio optimization[C]//Proceedings of the 27th Chinese Control and Decision Conference. Qingdao, China: IEEE, 2015.
- [25] OGRYCZAK W, ŚLIWIŃSKI T. On solving the dual for portfolio selection by optimizing Conditional Value at Risk[J]. *Computational Optimization and Applications*, 2011, 50(3): 591–595.
- [26] TIBSHIRANI R. Regression shrinkage and selection via the lasso[J]. *Journal of the Royal Statistical Society: Series B (Methodological)*, 1996, 58(1): 267–288.
- [27] BECK A, TEOULLE M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems[J]. *SIAM Journal on Imaging Sciences*, 2009, 2(1): 183–202.

- [7] SACHDEV S, BHATT R N. Bond-operator representation of quantum spins: mean-field theory of frustrated quantum Heisenberg antiferromagnets[J]. *Physical Review B*, 1990, 41(13): 9323–9329.
- [8] WANG H T, LIN H Q, SHEN J L. Bond-operator analysis of the magnetization of spin chains[J]. *Physical Review B*, 2000, 61(6): 4019–4025.
- [9] KUMAR M, SOOS Z G. Bond-order wave phase of the extended Hubbard model: electronic solitons, paramagnetism, and coupling to Peierls and Holstein phonons[J]. *Physical Review B*, 2010, 82(15): 155144–155156.
- [10] LI Y X, CHEN B. Bond operator theory of Heisenberg ferrimagnetic spin chain[J]. *Physics Letters A*, 2015, 379(5): 407–411.
- [11] CHEN Y G, LI Y X, TIAN L J, et al. Bond operator method of dimerized ferrimagnetic spin chain[J]. *Physica B: Condensed Matter*, 2017, 520: 65–68.
- [12] HASE M, TERASAKI I, UCHINOKURA K. Observation of the spin-peierls transition in linear Cu^{2+} (Spin-1/2) chains in an inorganic compound CuGeO_3 [J]. *Physical Review Letters*, 1993, 70(23): 3651–3654.
- [13] WANG H T, SHEN J L, SU Z B. Bond operator theory of isotropic spin-1 antiferromagnetic chains[J]. *Physical Review B*, 1997, 56(22): 14435–14440.
- [14] NIGHTINGALE M P, BLÖTE H W J. Gap of the linear spin-1 Heisenberg antiferromagnet: a Monte Carlo calculation[J]. *Physical Review B*, 1986, 33(1): 659–661.
- [15] KATO Y, TANAKA A. Numerical study of the $S=1$ antiferromagnetic spin chain with bond alternation[J]. *Journal of the Physical Society of Japan*, 1994, 63(4): 1277–1280.
- [16] YAMAMOTO S. Elementary excitations of $S=1$ antiferromagnetic Heisenberg chains with bond alternation[J]. *Physical Review B*, 1995, 51(22): 16128–16134.
- [17] HAGIWARA M, NARUMI Y, KINDO K, et al. Experimental verification of the gapless point in the $S=1$ antiferromagnetic bond alternating chain[J]. *Physical Review Letters*, 1998, 80(6): 1312–1315.
- [18] SAKAI T, TAKAHASHI M. Effect of the Haldane gap on quasi-one-dimensional systems[J]. *Physical Review B*, 1990, 42(7): 4537–4543.
- [19] CHEN W, HIDA K, SANCTUARY B C. Ground-state phase diagram of $S=1$ XXZ chains with uniaxial single-ion-type anisotropy[J]. *Physical Review B*, 2003, 67(10): 104401–104409.

(编辑: 石 瑛)

(上接第 532 页)

- [28] DAUBECHIES I, DEFRISE M, DE MOL C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint[J]. *Communications on Pure and Applied Mathematics*, 2004, 57(11): 1413–1457.
- [29] FAN J Q, ZHANG J J, YU K. Vast portfolio selection with gross-exposure constraints[J]. *Journal of the American Statistical Association*, 2012, 107(498): 592–606.
- [30] LIM C, SHERALI H D, URYASEV S. Portfolio optimization by minimizing conditional value-at-risk via nondifferentiable optimization[J]. *Computational Optimization and Applications*, 2010, 46(3): 391–415.
- [31] KOHAVI R. A study of cross-validation and bootstrap for accuracy estimation and model selection[C]//Proceedings of International Joint Conference on Artificial Intelligence. San Francisco, CA: Morgan Kaufmann Publishers, 1995.
- [32] AKAIKE H. Factor analysis and AIC[J]. *Psychometrika*, 1987, 52(3): 317–332.
- [33] KUHA J. AIC and BIC: comparisons of assumptions and performance[J]. *Sociological Methods & Research*, 2004, 33(2): 188–229.
- [34] LV Y Z, ZHANG R Q, ZHAO W H, et al. Quantile regression and variable selection of partial linear single-index model[J]. *Annals of the Institute of Statistical Mathematics*, 2015, 67(2): 375–409.

(编辑: 丁红艺)