

基于 CVaR 两步核估计量的投资组合管理^①

黄金波¹, 李仲飞^{2*}, 姚海祥³

(1. 广东财经大学金融学院, 广州 510320; 2. 中山大学管理学院, 广州 510275;
3. 广东外语外贸大学金融学院, 广州 510006)

摘要: 在不做任何分布假设的条件下, 利用非参数核估计方法对风险度量条件风险价值(conditional value-at-risk, CVaR) 进行估计, 得到 CVaR 的两步核估计公式. 然后用估计出来的 CVaR 代替理论上的 CVaR 建立均值-CVaR 模型, 实现对风险估计与投资组合优化同时进行, 并基于迭代思想设计求解该模型的简单算法. 蒙特卡洛模拟结果表明基于两步核估计方法的投资组合优化模型和算法比现有的方法更加有效, 估计出来的组合边界误差更小. 引入无风险资产后, 文中的模型和算法同样适用. 最后, 为说明其应用价值, 采用中国 A 股市场的日收益率数据进行了实例分析.

关键词: 均值-CVaR 模型; 两步核估计量; 组合边界; 中国 A 股市场

中图分类号: F830.9; O212.7 文献标识码: A 文章编号: 1007-9807(2016)05-0114-13

0 引言

投资组合选择的定量分析可追溯到 Markowitz^[1]建立的均值-方差模型, 此后, 均值-风险框架成为现代投资组合选择理论的基本分析框架之一. 用期望收益率度量投资收益已被广泛接受, 然而, 以收益率的方差作为风险度量指标, 则受到多方面的批评. 许多学者在批判方差的基础上发展了多种风险度量工具, 从而推动了现代投资组合选择理论的发展. 风险价值(value-at-risk, VaR) 和条件风险价值(conditional value-at-risk, CVaR) 风险度量正是在这个背景下被提出并很快被运用于投资组合选择和风险管理研究中. VaR 是指给定置信水平下某一个资产或资产组合在未来一定期限内的最大可能损失^[2]. 由于 VaR 不满足一致

性风险度量理论的次可加性公理^[3-4], 从而破坏投资组合理论中的风险分散化原理. 另外, VaR 不能对超过 VaR 水平的损失给出任何信息. 所以在 VaR 基础上, Rockafellar 和 Uryasev^[5-6]给出了 CVaR 的概念. CVaR 度量的是损失超过 VaR 水平的条件期望值. CVaR 满足一致性风险度量要求, 弥补了 VaR 不满足次可加性、未考虑尾部风险等缺陷.

VaR 和 CVaR 被提出之后, 对它们的研究沿着两个方向展开, 一个是 VaR 和 CVaR 估计问题的研究, 另一个是基于均值-VaR 和均值-CVaR 模型的投资组合优化问题研究. 一直以来, 这两个方向的研究相对独立发展. 风险估计问题的研究注重开发更加精确的计量模型和方法来捕捉金融市场的特征, 进而更加准确地估计金融市场风险, 且

① 收稿日期: 2013-08-17; 修订日期: 2014-03-24.

基金项目: 国家自然科学基金重点资助项目(71231008); 国家自然科学基金资助项目(71471045); 中国博士后科学基金特别资助项目(2015T80896); 中国博士后科学基金资助项目(2014M562246; 2014M560658); 全国统计科学研究计划资助项目(2013LY101); 广东省自然科学基金研究团队资助项目(2014A030312003); 广东省高等学校高层次人才资助项目; 广东省高等院校科技创新资助项目(2012KJ CX0050); 广东省普通高校特色创新资助项目(人文社科类); 广州市哲学社会科学规划资助项目(14G42).

通信作者: 李仲飞(1963—), 男, 内蒙古鄂尔多斯人, 博士, 教授, 博士生导师. Email: lnslzf@mail.sysu.edu.cn

前这方面的研究已相当完备,并始终处于不断发展中.相对而言,因 VaR 和 CVaR 的优化问题较难处理,基于 VaR 和 CVaR 的投资组合研究大多在特定分布下进行,这使得投资组合选择理论的应用受到局限.

关于均值-VaR 模型和均值-CVaR 模型较为成熟的研究大多在正态(或椭圆)分布假设下进行^[7-8].在正态(或椭圆)分布的假设下,VaR 和 CVaR 可以表达成均值和方差的线性函数,均值-VaR 模型和均值-CVaR 模型退化成均值-方差模型^[5],而均值-方差模型的研究已经相当成熟.所以,在正态(或椭圆)分布假设下,VaR 和 CVaR 与方差在风险度量方面没有本质区别,这将无法凸显 VaR 和 CVaR 在风险度量方面的优越性.另外,在实际的金融市场上,金融时间序列数据通常表现出尖峰厚尾、非对称等非正态(或椭圆)分布特征,简单的正态(或椭圆)分布假设将会导致风险估计的系统偏差,进而会误导投资者,使得人们无法进行有效的风险管理和组合优化.因此,在不做任何分布设定的条件下,如何利用计量方法估计出现实金融市场中的实际风险,进而把实际风险嵌入到投资组合优化模型中进行投资决策是非常有意义的课题.

针对 VaR 和 CVaR 的估计问题,理论界提出了很多方法. Engle 和 Manganelli^[9] 将它们分为 3 大类.第 1 类是参数法,主要包括 GARCH 族模型和 Copula 函数法.第 2 类是半参数法,主要包括极值理论 EVT (extreme value theory) 和条件自回归 VaR.第 3 类是非参数法,主要包括经验分布函数法和核估计方法.参数法与半参数法都假设收益率(在极值理论下是尾部收益率)服从某一事先设定的模型,然后估计出模型中的参数,进而得到风险度量 VaR 和 CVaR 的估计值^[10-11].相对于参数法与半参数法,非参数法不需要对收益率做任何形式的模型设定,避免人为的模型设定风险和参数估计偏差,能够给出较为准确的风险估计.更重要的是,非参数核估计方法允许金融时间序列之间相互依赖^[12-13],Bellini 和 Figa-Talamanca^[14] 证实收益率序列数据显示出非常强的尾部依赖,而参数法与半参数法对于这类相互依赖变

量问题的处理较为棘手^[12].

近年来,非参数核估计法因具备上述几方面的优势而备受广大学者关注.利用核估计法估计金融风险始于 Gouriéroux 等^[15] 的研究,他们首次考察了 VaR 的核估计.随后 Scaillet^[16-17] 把核估计法应用到对期望损失(expected shortfall,简称 ES)^②的估计. Scaillet^[16] 提出 ES 的两步核估计法,并用它来估计资产组合的期望损失和期望损失对组合头寸的敏感性. Scaillet^[17] 研究了条件 VaR 和条件 ES 的非参数核估计,并在平稳过程满足强混合条件下,导出了条件 ES 核估计量的渐进性质. Chen^[18] 同时用经验分布函数和核平滑分布函数估计 ES,得出二者在估计的方差和均方误差方面并无明显的差异.刘静和杨善朝^[19] 放松 Scaillet^[16] 的前提条件,在 α 混合序列具有幂衰减混合系数条件下,用两步核估计法估计 ES,得到了 ES 核估计量的 Bahadur 表示、均方误差和渐近正态性的收敛速度.刘晓倩和周勇^[20] 比较两步核光滑 ES 估计与 ES 完全经验估计及一步核光滑估计的优劣,得到两步光滑化并不能减小 ES 估计的方差,该发现与 Chen^[18] 的结论一致.由于 CVaR 的核估计量具有良好的连续性和光滑性,可以方便地处理投资组合优化问题,这一优点是 ES 完全经验估计不具备的.所以,许多学者倾向于利用核估计法来研究组合的 CVaR 及相关优化问题^[16-21].

虽然近年来,学者对 CVaR 的非参数核估计法做了诸多研究,但还鲜有学者把 CVaR 的核估计与风险优化、投资组合选择问题结合起来考虑.问题是不仅需要知道风险有多大,而且还要知道如何去对冲和管理风险,风险估计只是解决了前面的问题,而后面的问题往往更为重要. Yao 等^[21] 对此做了有益尝试,他们利用 Rockafellar 和 Uryasev^[5] 给出的 CVaR 特殊表达式并结合非参数核估计方法,得到核估计框架下的均值-CVaR 模型,并利用优化算法求解模型得到投资组合的组合边界.不同 Yao 等^[21] 的研究,本文直接利用两步核估计方法对 CVaR 进行估计,并将 CVaR 的两步核估计式嵌入均值-CVaR 模型,这样就不需

② 在分布函数满足连续性的条件下,ES 与 CVaR 是同一个风险度量指标的两个不同称呼.

要借助于 Rockafellar 和 Uryasev^[5] 的 CVaR 特殊形式,而且模型的自变量维数比 Yao 等^[21] 的少. 本文基于迭代思想设计了简单的算法对该模型进行求解. 蒙特卡洛模拟结果显示,在偏差意义下,基于两步核估计方法的均值-CVaR 模型和算法准确有效,比现有方法的估计误差小^③. 最后将模型和算法拓展到存在无风险资产时的情形,并将它们应用到中国 A 股市场.

1 组合风险 CVaR 的两步核估计

1.1 非参数核估计方法基础

非参数核估计方法是近年发展起来的统计计量方法,它不需要事先设定任何的分布函数形式,而是在非常一般的条件下(如未知函数满足一定的光滑性等正则条件),利用收集的数据来拟合出分布函数,能够给出分布函数的稳健估计. 当对金融时间序列的分布函数进行估计时,大多数情况下,并不知道样本数据服从什么类型的分布,这时需要借助于非参数核估计方法. 给定一维时间序列的样本数据 $\{x_i\}_{i=1}^T$,假设其密度函数和分布函数为连续可导的未知函数 $f(x)$ 和 $F(x)$,则它们的非参数核估计量分别为^[22]

$$\begin{aligned} \hat{f}(x) &= \frac{1}{Th} \sum_{i=1}^T g\left(\frac{x_i - x}{h}\right), \\ \hat{F}(x) &= \int_{-\infty}^x \hat{f}(\tau) d\tau = \frac{1}{T} \sum_{i=1}^T G\left(\frac{x - x_i}{h}\right) \end{aligned} \quad (1)$$

其中函数 $G(y) = \int_{-\infty}^y g(\tau) d\tau$,而 $g(\tau)$ 称核函数,满足 $0 < g(\tau) < \infty$, $g(-\tau) = g(\tau)$, $\int_{-\infty}^{\infty} g(\tau) d\tau = 1$, $\int_{-\infty}^{\infty} \tau g(\tau) d\tau = 0$, $\int_{-\infty}^{\infty} \tau^2 g(\tau) d\tau = \sigma^2 < \infty$.

分布函数的核估计量考虑了局部加权平均,在每个点处的分布函数估计值都利用了所有的样本数据,比经验分布函数包含了更多的样本信息. Chen 和 Tang^[12] 得出分布函数的核估计量和经验分布函数都是真实分布函数的一致估计量,但前者的方差更小. 在核估计中,核函数起到平滑的作用,由于非参数核估计结果对不同的核函数并不太敏感,所以只要满足一定的正则性条件,核函数

可由研究者根据问题的不同自由选取. 如果是对单变量的密度函数或分布函数进行估计,二阶 Gauss 核函数能够给出更稳健的结果. 所谓二阶 Gauss 核函数就是 $g(\cdot)$ 取标准正态分布的密度函数. 可以验证二阶 Gauss 核函数满足前面对核函数的 5 点要求. 通常情况下,非参数核估计方法需要大量的观察数据才能够拟合得比较准确,证券市场上大量的高频数据可以满足这一要求.

在式(1)中, h 为平滑窗宽,在非参数核估计方法里,估计结果对不同的窗宽非常敏感,窗宽取得过大,会导致过度平滑,而当窗宽取得太小,会导致平滑不足. 在实践中,常用的窗宽选择方法是拇指法则(rule-of-thumb)^[22],即 $h = 1.06(T^{-0.2})\hat{\sigma}(x)$, $\hat{\sigma}(x)$ 为 $\{x_i\}_{i=1}^T$ 的样本标准差.

1.2 组合风险 CVaR 的两步核估计

假设存在 $n(n > 1)$ 种风险资产,第 i 种资产的收益率 r_i 为随机变量,则 $r = (r_1, r_2, \dots, r_n)'$ 为 n 种风险资产的收益率向量. 投资者在第 i 种资产上持有头寸为 x_i ,则 $x = (x_1, x_2, \dots, x_n)'$ 就构成了投资组合,组合的收益率 $R = x'r$. n 种资产在 T 期内的收益率记为 $\{r_t\}_{t=1}^T$,其中 $r_t = (r_{1t}, r_{2t}, \dots, r_{nt})'$,则样本均值向量 $\bar{r} = T^{-1} \sum_{t=1}^T r_t$,样本协方差阵 $S^2 = (T-1)^{-1} \sum_{t=1}^T (r_t - \bar{r})(r_t - \bar{r})'$. 投资组合的收益率样本为 $\{R_t\}_{t=1}^T$,其中 $R_t = x'r_t$,投资组合收益率的样本均值 $\bar{R} = x'\bar{r}$,样本方差为 $x'S^2x$. 记投资组合 x 在损失概率 α 下的 VaR 和 CVaR 分别为 $v(x, \alpha)$ 和 $u(x, \alpha)$.

根据 VaR 的定义,投资组合的 VaR 数学表达式为

$$\begin{aligned} v(x, \alpha) &:= -\inf\{z: F_R(z) \geq \alpha\} \Rightarrow \\ F_R(-v(x, \alpha)) &= \alpha \end{aligned} \quad (2)$$

式中 $F_R(\cdot)$ 为 R 的分布函数,设其连续可导.

根据 CVaR 的定义,投资组合的 CVaR 数学表达式为^[16]

$$\begin{aligned} u(x, \alpha) &:= E[-R | -R > v(x, \alpha)] \\ &= -\alpha^{-1} \int_{-\infty}^{-v(x, \alpha)} z f_R(z) dz \end{aligned} \quad (3)$$

③ 误差指标的定义可见后文的式(19).

式中 $f_R(\cdot)$ 为 R 的密度函数, $E[\cdot]$ 为期望算子, 边际 CVaR (marginal CVaR, MCVaR) 被定义为组合 CVaR 对头寸的导数^[16]

$$\Delta_x u(\mathbf{x}, \alpha) = E[-r \mid -R > v(\mathbf{x}, \alpha)] \quad (4)$$

由式(2)可知 VaR 实际上是组合收益率的下 α 分位数的相反数. 所以, 可以利用前节介绍的核估计方法先估计 $f_R(z)$ 和 $F_R(z)$, 然后基于估计出的分布函数来估计组合 VaR 和 CVaR. 记 $v(\mathbf{x}, \alpha)$ 和 $u(\mathbf{x}, \alpha)$ 的核估计量分别为 $\hat{v}(\mathbf{x}, \alpha)$ 和 $\hat{u}(\mathbf{x}, \alpha)$, 根据 Scaillet^[16] 及刘静和杨善朝^[19] 的研究, 以下给出 $u(\mathbf{x}, \alpha)$ 的两步核估计法.

第一步估计 $\hat{v}(\mathbf{x}, \alpha)$. 定义

$$[\zeta_{1t}, \zeta_{2t}, \zeta_{3t}] := (Th)^{-1} \sum_{i=1}^T \zeta_{it} g\left(\frac{\zeta_{3t} - \zeta_{2t}}{h}\right)$$

其中 $g(\cdot)$ 为核函数, h 为窗宽, 则 R 的密度函数 $f_R(z)$ 和分布函数 $F_R(z)$ 的核估计量分别为

$$\begin{aligned} \hat{f}_R(z) &= [1, R_t, z], \\ \hat{F}_R(z) &= \int_{-\infty}^z [1, R_t, \tau] d\tau \end{aligned} \quad (5)$$

上式经化简即为式(1)的特例. 从而, $\hat{v}(\mathbf{x}, \alpha)$ 可以通过下式得到

$$\hat{F}_R(-\hat{v}(\mathbf{x}, \alpha)) = \int_{-\infty}^{-\hat{v}(\mathbf{x}, \alpha)} [1, R_t, \tau] d\tau = \alpha \quad (6)$$

第二步估计 $\hat{u}(\mathbf{x}, \alpha)$. 令

$$\hat{I}(\xi) = \int_{-\infty}^{-\xi} [r_t, R_t, \tau] d\tau$$

条件期望 $E[r \mid -R > v(\mathbf{x}, \alpha)]$ 的估计量为 $\frac{\hat{I}(\hat{v}(\mathbf{x}, \alpha))}{\alpha}$ ^[16], 从而 CVaR 和 MCVaR 的核估计量为

$$\hat{u}(\mathbf{x}, \alpha) = \frac{-\mathbf{x} \hat{I}(\hat{v}(\mathbf{x}, \alpha))}{\alpha}, \quad (7)$$

$$\Delta_x \hat{u}(\mathbf{x}, \alpha) = \frac{-\hat{I}(\hat{v}(\mathbf{x}, \alpha))}{\alpha}$$

1.3 组合风险 CVaR 的凸性

引理 1^[23] 组合风险 CVaR 对组合头寸的二阶导数矩阵的解析式为

$$\begin{aligned} \frac{\partial^2 u(\mathbf{x}, \alpha)}{\partial \mathbf{x} \partial \mathbf{x}'} &= p(0) [1 - P(0)]^{-1} \times \\ &V[r \mid R = -v(\mathbf{x}, \alpha)] \end{aligned} \quad (8)$$

式中 $p(\cdot)$ 为 $-R - v(\mathbf{x}, \alpha)$ 的概率密度函数; $P(\cdot)$ 为它的分布函数; $V[\cdot]$ 为方差算子.

由式(8)容易看出理论上 CVaR 的二阶导数矩阵为半正定矩阵, 即 CVaR 满足凸性. 下面给出

$\frac{\partial^2 u(\mathbf{x}, \alpha)}{\partial \mathbf{x} \partial \mathbf{x}'}$ 的核估计量, 由式(1), 可得 $p(0)$ 和 $P(0)$ 的核估计值分别为

$$\begin{aligned} \hat{p}(0) &= (Th)^{-1} \sum_{i=1}^T g\left(\frac{-R_i - \hat{v}(\mathbf{x}, \alpha)}{h}\right), \\ \hat{P}(0) &= T^{-1} \sum_{i=1}^T G\left(\frac{R_i + \hat{v}(\mathbf{x}, \alpha)}{h}\right) \end{aligned} \quad (9)$$

又

$$\begin{aligned} V[r \mid R = -v(\mathbf{x}, \alpha)] &= \\ E[rr' \mid R = -v(\mathbf{x}, \alpha)] &- \\ E[r \mid R = -v(\mathbf{x}, \alpha)] \times \\ E[r' \mid R = -v(\mathbf{x}, \alpha)] & \end{aligned}$$

则其核估计公式为^④

$$\begin{aligned} \hat{V}[r \mid R = -v(\mathbf{x}, \alpha)] &= \\ \frac{\sum_{i=1}^T r_i r_i' g\left(\frac{R_i + \hat{v}(\mathbf{x}, \alpha)}{h}\right)}{\sum_{i=1}^T g\left(\frac{R_i + \hat{v}(\mathbf{x}, \alpha)}{h}\right)} &- \\ \left[\frac{\sum_{i=1}^T r_i g\left(\frac{R_i + \hat{v}(\mathbf{x}, \alpha)}{h}\right)}{\sum_{i=1}^T g\left(\frac{R_i + \hat{v}(\mathbf{x}, \alpha)}{h}\right)} \right] \times \\ \left[\frac{\sum_{i=1}^T r_i' g\left(\frac{R_i + \hat{v}(\mathbf{x}, \alpha)}{h}\right)}{\sum_{i=1}^T g\left(\frac{R_i + \hat{v}(\mathbf{x}, \alpha)}{h}\right)} \right] & \end{aligned} \quad (10)$$

由此可得二阶导数矩阵的核估计量为

$$\begin{aligned} \frac{\partial^2 \hat{u}(\mathbf{x}, \alpha)}{\partial \mathbf{x} \partial \mathbf{x}'} &= \hat{p}(0) [1 - \hat{P}(0)]^{-1} \times \\ &\hat{V}[r \mid R = -v(\mathbf{x}, \alpha)] \end{aligned} \quad (11)$$

定理 1 组合风险 CVaR 的二阶导数矩阵的

④ 这里用到非参数核估计的回归技术, 具体可见参考文献[22]的第 60-66 页.

核估计量 $\frac{\partial^2 \hat{u}(x, \alpha)}{\partial x \partial x}$ 为半正定矩阵.

此定理的证明可参见文献 [15] 和文献 [23].

2 均值-CVaR 模型基础知识

2.1 均值-CVaR 建模

假设存在 $n (n > 1)$ 种风险资产, 资产交易无摩擦, 市场上不存在卖空限制. 投资者的财富标准化为 1, 其它条件同上. 记 e 为元素全为 1 的 n 维列向量, \mathfrak{R} 为投资者要求的最低期望收益率, u 为投资者愿意承受的用 CVaR 度量的最大风险, λ 为投资者的风险厌恶系数, 均值-CVaR 最优化模型可由以下 3 种方式构建

$$\begin{aligned}
 (\Psi_1) \quad & \begin{cases} \min_{x \in \mathbb{R}^n} & u(x, \alpha) \\ \text{s. t.} & E[R] \geq \mathfrak{R}, x'e = 1 \end{cases} \\
 (\Psi_2) \quad & \begin{cases} \max_{x \in \mathbb{R}^n} & E[R] \\ \text{s. t.} & u(x, \alpha) \leq u, x'e = 1 \end{cases} \\
 (\Psi_3) \quad & \begin{cases} \max_{x \in \mathbb{R}^n} & E[R] - \lambda u(x, \alpha), \lambda \geq 0 \\ \text{s. t.} & x'e = 1 \end{cases}
 \end{aligned}$$

令参数 $\mathfrak{R}, \mu, \lambda$ 变动, 便产生了各自意义下的均值-CVaR 组合边界. Krokmal 等 [24] 证明在一定条件下 3 种模型得到的组合边界是等价的, 所以, 下文只讨论模型 Ψ_1 .

假设市场引入一个无风险资产, r_f 为无风险资产的收益率. $x = (x_1, x_2, \dots, x_n)'$ 为投资者在风险资产上持有的头寸, 则 $(1 - x'e)$ 为投资者在无风险资产上持有的头寸. 投资组合的收益率 $R = (1 - x'e)r_f + x'r$. 根据 CVaR 的平移不变性 [4], 存在无风险资产时组合 CVaR 为

$$u(r_f, x, \alpha) = -(1 - x'e)r_f + u(x, \alpha) \quad (12)$$

由此, 可以建立含有无风险资产时的均值-CVaR 模型

$$\begin{aligned}
 (\Lambda_1) \quad & \begin{cases} \min_{x \in \mathbb{R}^n} & -(1 - x'e)r_f + u(x, \alpha) \\ \text{s. t.} & E[R] \geq \mathfrak{R} \end{cases} \\
 (\Lambda_2) \quad & \begin{cases} \max_{x \in \mathbb{R}^n} & E[R] \\ \text{s. t.} & -(1 - x'e)r_f + u(x, \alpha) \leq u \end{cases} \\
 (\Lambda_3) \quad & \max_{x \in \mathbb{R}^n} E[R] - \\
 & \lambda [-(1 - x'e)r_f + u(x, \alpha)], \lambda \geq 0
 \end{aligned}$$

2.2 正态分布假设下的均值-CVaR 模型与显示解

在 $n (n > 1)$ 种风险资产的收益率 r 服从联合正态分布 $N(\mu, \Sigma)$ 的假设下, 投资组合的 CVaR 可表达成期望和标准差的线性组合 [25]

$$u(x, \alpha) = -E[R] + \frac{\phi(z_\alpha) \sigma(R)}{\alpha}$$

式中 $\sigma(\cdot)$ 为标准差算子; $\phi(\cdot)$ 为标准正态分布的概率密度函数; z_α 为标准正态分布的下 α 分位数. 均值-CVaR 模型退化为均值-方差模型, 以问题 Ψ_1 为例, 退化为

$$(\Psi_1) \quad \begin{cases} \min_{x \in \mathbb{R}^n} & u(x, \alpha) = -E[R] + \frac{\phi(z_\alpha) \sigma(R)}{\alpha} \\ \text{s. t.} & E[R] = x'\mu \geq \mathfrak{R}, x'e = 1 \end{cases}$$

利用均值-方差的组合边界表达式, 可以得出均值-CVaR 的组合边界表达式为 [25]

$$\frac{C(u(x, \alpha) + \mathfrak{R})^2}{\left(\frac{\phi(z_\alpha)}{\alpha}\right)^2} - \frac{\left(\mathfrak{R} - \frac{A}{C}\right)^2}{\frac{D}{C^2}} = 1 \quad (13)$$

其中 $A = e'\Sigma^{-1}\mu$; $B = \mu'\Sigma^{-1}\mu$; $C = e'\Sigma^{-1}e$; $D = BC - A^2$.

在引入一个无风险资产的情况下, 均值-CVaR 的组合边界方程为 [25]

$$\frac{(u(x, \alpha) + \mathfrak{R})^2}{\left(\frac{\phi(z_\alpha)}{\alpha}\right)^2} = \frac{(\mathfrak{R} - r_f)^2}{H} \quad (14)$$

其中 $H = Cr_f^2 - 2Ar_f + B$.

在实践中, 通常并不知道真实均值向量和协方差矩阵, 所以还要估计这 n 种风险资产收益率的均值向量和协方差矩阵. 通常的做法是, 用 n 种风险资产的样本均值向量 \bar{r} 和样本协方差矩阵 S^2 代替理论均值向量 μ 和理论协方差阵 Σ , 这样可以得到正态分布假设下的均值-CVaR 曲线的近似估计. 本文把这种在正态分布假设下将样本均值向量和方差阵带入到均值-CVaR 曲线表达式的做法叫做传统方法.

3 基于 CVaR 两步核估计量的均值-CVaR 模型

3.1 模型与求解

在两步核估计的框架下, 模型 $\Psi_1 \sim \Psi_3$ 里的

理论 CVaR 都用核估计量 $\hat{u}(x, \alpha)$ 代替, 投资组合的期望收益率用其样本均值代替. 以问题 Ψ_1 为例, 在两步核估计框架下, 问题变为

$$(\Psi_1) \begin{cases} \min_{x \in \mathbb{R}^n} & \hat{u}(x, \alpha) = \frac{-x \hat{I}(\hat{v}(x, \alpha))}{\alpha} \\ \text{s. t.} & \bar{R} = x \bar{r} \geq \mathfrak{R}, \quad x' e = 1 \end{cases}$$

由于 $\hat{u}(x, \alpha)$ 的两步核估计中用到 VaR 的核估计量 $\hat{v}(x, \alpha)$ 和窗宽 h , 所以两步核估计框架下的均值-CVaR 模型必须同时满足如下约束条件.

条件 1 h 满足窗宽的估计公式

$$h = 1.06T^{-0.2} \sqrt{x' S^2 x}$$

条件 2 $\hat{v}(x, \alpha)$ 满足 VaR 的估计式 (6)

$$\int_{-\infty}^{-\hat{v}(x, \alpha)} [1 - R_t | \sigma] d\tau = \alpha$$

直接求解上面优化问题难度是很大的, 由于 $\hat{v}(x, \alpha)$ 没有显示表达式, 条件 1 和条件 2 无法以显示表达式的形式进入优化问题, 所以问题 Ψ_1 的目标函数也不是决策变量的显示表达式. 然而, 仔细观察可以发现, 条件 1 和条件 2 实际不是对 x 取值的约束. 给定 x 值, 可以通过条件 1 和条件 2 计算出 h , $\hat{v}(x, \alpha)$, 然后基于此可以估计出 $\hat{u}(x, \alpha)$, 这正是 CVaR 两步核估计的思路. 所以, 这一特点为利用迭代算法求解问题提供了方便.

设定 Lagrange 函数

$$L(x, \mu_1, \mu_2) = \hat{u}(x, \alpha) + \mu_1(x' \bar{r} - \mathfrak{R}) + \mu_2(x' e - 1) \quad (15)$$

记 $y = (x' \mu_1 \mu_2)'$, 则一阶条件为

$$f = \frac{\partial L}{\partial y} = (L_x, L_{\mu_1}, L_{\mu_2})' = \vec{0}$$

其中

$$\begin{aligned} L_x &= \Delta_x \hat{u}(x, \alpha) + \mu_1 \bar{r} + \mu_2 e, \\ L_{\mu_1} &= x \bar{r} - \mathfrak{R}, \\ L_{\mu_2} &= x' e - 1 \end{aligned} \quad (16)$$

由于 $\Delta_x \hat{u}(x, \alpha)$ 的复杂性, 从一阶条件中得到显示解是十分困难的. 定义二阶导数矩阵

$$H = \frac{\partial^2 L}{\partial y \partial y'} = \begin{pmatrix} L_{xx} & L_{x\mu_1} & L_{x\mu_2} \\ L_{\mu_1 x} & L_{\mu_1 \mu_1} & L_{\mu_1 \mu_2} \\ L_{\mu_2 x} & L_{\mu_2 \mu_1} & L_{\mu_2 \mu_2} \end{pmatrix} \quad (17)$$

其中

$$L_{xx} = \frac{\partial^2 \hat{u}(x, \alpha)}{\partial x \partial x'}, L_{x\mu_1} = L_{\mu_1 x} = \bar{r}, \quad (18)$$

$$L_{x\mu_2} = L_{\mu_2 x} = e, L_{\mu_i \mu_j} = 0, i, j = 1, 2$$

对问题 Ψ_1 的求解, 可通过以下迭代算法实现:

步骤 1 给定初始 $y_0 = (x_0' \mu_{1,0} \mu_{2,0})'$ 和很小的正数 ε , 取迭代计数变量 $k := 0$;

步骤 2 依照条件 1 和条件 2 的公式分别计算 $h_k, \hat{v}(x_k, \alpha)$

$$h_k = 1.06T^{-0.2} \sqrt{x_k' S^2 x_k},$$

$$\hat{v}(x_k, \alpha) =$$

$$\arg \min_v \left| \int_{-\infty}^{-v} [1 - x_k' r_t | \sigma] d\tau - \alpha \right|$$

利用式 (7) 可得 $\hat{u}(x_k, \alpha)$ 和 $\Delta_x \hat{u}(x_k, \alpha)$, 同时可以计算 f_k 和 H_k . 如果 $\|f_k\| < \varepsilon$, 停止迭代, 输出结果 y_k ; 否则计算 H_k^{-1} , 如果 H_k^{-1} 不存在, 则迭代失败, 如果 H_k^{-1} 存在, 则转入下一步 (其中 $\|\cdot\|$ 表示欧式模);

步骤 3 计算 $y_{k+1} = y_k - H_k^{-1} f_k$, 如果 $\|y_{k+1} - y_k\| < \varepsilon$, 则停止迭代, 输出 y_{k+1} , 否则 $k := k + 1$, 转入步骤 2.

为了保证上面迭代算法能够顺利进行, 本文得到以下两个定理.

定理 2 如果 $g(\cdot)$ 取二阶 Gauss 核函数, 则方程 $\int_{-\infty}^{-v} [1 - R_t | \sigma] d\tau = \alpha$ 有且仅有 1 个解.

证明 令 $\vartheta(v) = \int_{-\infty}^{-v} [1 - R_t | \sigma] d\tau$, 由式 (6) 可得

$$\begin{aligned} \vartheta(v) &= \int_{-\infty}^{-v} [1 - R_t | \sigma] d\tau \\ &= (Th)^{-1} \sum_{t=1}^T \int_{-\infty}^{-v} g\left(\frac{\tau - R_t}{h}\right) d\tau \\ &= T^{-1} \sum_{t=1}^T G\left(\frac{-R_t - v}{h}\right) \end{aligned}$$

式中 $g(\cdot)$ 和 $G(\cdot)$ 是分别取标准正态分布的密度函数和分布函数. 容易算得, $\lim_{v \rightarrow -\infty} \vartheta(v) = 1$, $\lim_{v \rightarrow \infty} \vartheta(v) = 0$, 又因为 $\vartheta(v)$ 为连续函数且 $\alpha \in (0, 1)$, 由介值定理, 方程 $\vartheta(v) = \alpha$ 至少存在 1 个解. 简单推导可得

$$\vartheta'(v) = -T^{-1} \sum_{i=1}^T g\left(\frac{-R_i - v}{h}\right) < 0$$

$\vartheta(v)$ 为严格单调减函数, 所以方程 $\vartheta(v) = \alpha$ 有且仅有 1 个解. 证毕.

定理 3 如果集合

$$B = \{x \mid x \in \mathbb{R}^n, x^T r \geq \alpha, x^T e = 1\}$$

非空, 则问题 Ψ_1^* 是个凸优化问题.

证明 容易看出问题 Ψ_1^* 的可行集 B 是凸集, 所以只需要证明问题 Ψ_1^* 的目标函数的海塞矩阵是个半正定矩阵即可. 而 Ψ_1^* 的目标函数的海塞矩阵即为 $\frac{\partial^2 \hat{u}(x, \alpha)}{\partial x \partial x^T}$. 由式 (11) 及定理 1 可知, 它为半正定矩阵. 所以, 问题 Ψ_1^* 是个凸优化问题.

3.2 蒙特卡洛模拟

本节给出一些模拟算例, 验证 3.1 节模型和算法的准确性. 假设 $n (n > 1)$ 种资产收益率向量 r 服从 n 维正态分布 $N(\mu, \Sigma)$, 协方差阵 Σ 正定, 则由 Cholesky 分解得 $\Sigma = Q Q^T$, Q 为上三角矩阵. 令 $r = \mu + Q \varepsilon$, ε 为 n 维标准正态分布, 通过简单推导可知 $r \sim N(\mu, \Sigma)$. 这样通过 ε 就可以生成多维正态分布 $N(\mu, \Sigma)$ 的随机数. 本文以下模拟数据都通过 Cholesky 分解法生成.

令 $\mu = (1.0 \quad 1.5 \quad 2.0)^T$ ⑤,

$$\Sigma = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 4 & 3 \\ 0 & 3 & 9 \end{pmatrix}$$

模拟 1 设定风险资产数量 $n = 3$, 资产收益率服从正态分布 $N(\mu, \Sigma)$. 模拟过程如下.

第 1 步 采用 Cholesky 分解生成多元正态分布的样本, 取样本容量 $T = 500, 1\ 000, 1\ 500, 2\ 000, 4\ 000, 8\ 000$, 设定损失概率分别为 1%、5% 和 10%.

第 2 步 在每个样本容量和概率下, 取 30 个不同的收益率 $\mathfrak{R}_i, i = 1, 2, \dots, 30$, 用以下 4 种方法计算每个收益率对应的最小 CVaR: 1) 把 $\mathfrak{R}_i, \mu, \Sigma$ 代入式 (13), 得到真实的最小 CVaR, 记为 u_{true} ; 2) 基于两步核估计框架下的非参数模型和优化算法 (记为 NP 方法) 估计出最小 CVaR, 记为 $u_{\text{np}, i}$; 3) 利用 Yao 等^[21] 的模型 (记为 YL 方法) 估计出最小 CVaR, 记为 $u_{\text{yl}, i}$; 4) 基于 Rockafellar 和 Uryasev^[5-6] 的线性规划方法 (记为 LP 方法), 估计出最小 CVaR, 记为 $u_{\text{lp}, i}$. 为比较后 3 种估计方法的精度, 以 NP 方法为例, 定义如下误差指标: 绝对误差 (absolute error, Ae) 和相对误差 (relative error, Re), 有

$$Ae_{\text{np}} = \frac{1}{30} \sum_{i=1}^{30} |(u_{\text{np}, i} - u_{\text{true}, i})|, \tag{19}$$

$$Re_{\text{np}} = \frac{1}{30} \sum_{i=1}^{30} \frac{|(u_{\text{np}, i} - u_{\text{true}, i})|}{u_{\text{true}, i}}$$

第 3 步 重复前面两个步骤 40 次, 即进行 40 次蒙特卡洛模拟. 然后将 40 次得到的误差指标进行平均, 得到平均绝对误差和平均相对误差 (见表 1 和表 2).

表 1 平均绝对误差模拟结果

Table 1 Simulation results of average absolute error

T	方法								
	NP			YL			LP		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
500	0.556 6	0.547 7	0.364 4	0.629 0	0.407 0	0.387 5	0.692 2	0.433 2	0.421 4
1 000	0.462 5	0.377 6	0.294 0	0.485 6	0.322 2	0.266 7	0.427 3	0.283 6	0.219 2
1 500	0.419 2	0.305 5	0.245 2	0.394 0	0.285 2	0.209 3	∞	0.248 4	0.284 5
2 000	0.380 0	0.252 1	0.167 8	0.371 7	0.215 3	0.178 8	∞	0.220 8	0.196 8
4 000	0.262 7	0.221 2	0.143 3	0.234 5	0.171 6	0.134 9	∞	0.154 8	0.158 6
8 000	0.156 3	0.105 8	0.098 5	0.177 6	0.124 3	0.105 3	∞	∞	0.085 0

⑤ 需要说明的是, 这里的数据是虚拟数据, 数据的量纲 (或单位) 可以是任意的. 如果是金融资产的收益率数据, 单位为 1% 或 1‰.

从表 1 可以得出以下结论: 1) 3 种估计方法的平均绝对误差都随着样本容量的增加而减小, 反映基于 3 种估计方法得到的组合边界随着样本容量的增加而收敛于真实的组合边界. 基于两步核估计方法和 YL 方法得到的组合边界收敛是因为随机变量的分布函数的核估计量一致收敛于真实的分布函数; 2) 3 种估计方法的平均绝对误差都随着损失概率的增加而减小. 这是因为 CVaR 度量的是损失超过 VaR 的期望值, 各种 CVaR 估计方法准确性受到质疑的重要方面就是极端风险的样本数据太少, 以至不能全面反映尾部分布的特征. α 值越大, α 分位数以上的样本点越多, 从而能够有效利用的数据信息越多, 估计也更加准确. 另外 α 值越大, CVaR 的绝对数值越小, 从而估计的绝对误差也越小是符合预期的; 3) 从 3 种

估计方法的平均绝对误差的大小来看, 3 种方法的估计精度相当, 在不同的样本容量和损失概率下, 各有优劣. 但当损失概率较小, 比如取 1% 或者样本容量较大时, LP 方法失效. 这是因为 LP 模型里有 $n + 1 + T$ 个自变量, 2 个等式约束和 $2T$ 个不等式约束^[21], LP 模型里的自变量数和约束方程数随样本容量增加而增加, 当样本容量 T 太大时, LP 算法的收敛速度会降低甚至失效(表 1 和表 2 中, ∞ 表示优化算法失效). 而 NP 方法的模型里只有 n 个自变量, 2 个等式约束, 而且随着样本容量 T 增加, 两步核估计方法的估计精度越来越高. YL 方法的模型里有 $n + 1$ 个自变量, 3 个等式约束^[21]. 绝对误差指标没有考虑真值的大小, CVaR 的真实值越大, 它的估计值与真实值之间的绝对误差也会越大, 所以剔除了这种水平效应的相对误差指标更有意义.

表 2 平均相对误差模拟结果

Table 2 Simulation results of average relative error

T	方法								
	NP			YL			LP		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
500	0.118 9	0.169 8	0.148 8	0.141 8	0.143 4	0.166 1	0.145 7	0.131 3	0.168 3
1 000	0.102 9	0.114 8	0.117 5	0.109 7	0.116 6	0.123 8	0.090 3	0.084 1	0.090 7
1 500	0.089 0	0.093 8	0.099 6	0.090 9	0.097 1	0.102 3	∞	0.077 2	0.105 6
2 000	0.080 8	0.077 7	0.068 7	0.085 0	0.078 0	0.087 4	∞	0.065 1	0.077 4
4 000	0.056 9	0.068 9	0.060 9	0.057 0	0.062 2	0.070 4	∞	0.048 1	0.062 7
8 000	0.034 3	0.034 2	0.042 2	0.042 5	0.046 1	0.054 0	∞	∞	0.032 9

从表 2 可以得出以下结论: 1) 同绝对误差指标相同, 3 种估计方法的平均相对误差都随着样本容量的增加而减小, 这说明 3 种估计方法满足大样本性质; 2) 除了 $\alpha = 5\%$, $T = 500$ 和 $\alpha = 5\%$, $T = 4 000$ 外, NP 方法的平均相对误差都小于 YL 方法, 说明 NP 方法的估计精度在大多数时候高于 YL 方法, 这可能是因为 NP 方法里的自变量和约束条件更少; 3) 与 LP 方法相比, NP 方法在小概率和大样本情况下表现更好, 在实际的金融风险监管和金融机构的风险计算中, 通常要计算小概率(小于 1%) 下的 CVaR, 所以 NP 方法在实际应用中更加有效.

法和 LP 方法得到的组合边界偏差较大. 表 2 的数据也显示样本容量为 4 000, 损失概率为 10% 时, NP 方法的相对误差(0.060 9) 小于 YL 方法(0.070 4) 和 LP 方法(0.062 7).

限于篇幅, 图 1 仅给出了样本容量为 4 000, 损失概率为 10% 时某一次的模拟结果(共模拟 40 次), 图上直观地显示: 基于 NP 方法得到的组合边界几乎与真实的组合边界重合, 而基于 YL 方

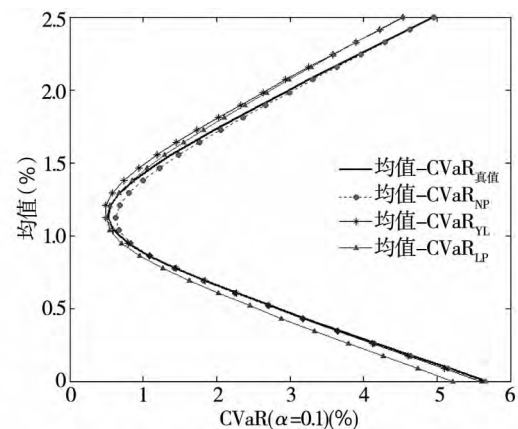


图 1 3 种估计方法的精度比较

Fig. 1 Comparison of estimation accuracy based on NP, LP and YL methods

模拟2 假设资产数量 $n = 3$,资产收益率向量 r 服从3维 t 分布 $t_m(\mu, \Sigma)$,其中均值 μ 和散度 Σ 取值同上, m 为自由度,本例取5. 设定样本容量为8 000,损失概率为1% , t 分布随机数的生成可采用Cholesky分解法^⑥. 本文分别基于NP方法、YL方法和传统方法(记为MN方法)估计多维 t 分布下的均值-CVaR曲线,并与真实的均值-CVaR曲线比较. 在多维 t 分布的情形下 CVaR 可表达成均值和方差的线性函数^[23]

$$u(x, \alpha) = -E[R] + k_\alpha \sigma(R) \quad (20)$$

其中

$$k_\alpha = \frac{1}{\alpha} \frac{m}{m-1} \frac{\Gamma((m+1)/2)}{\sqrt{m\pi} \Gamma(m/2)} \left(1 + \frac{t_{1-\alpha}^2}{m}\right)^{\frac{1-m}{2}}$$

式中 $t_{1-\alpha}$ 为自由度为 m 的经典一维 t 分布的下 $1 - \alpha$ 分位数; $\Gamma(\cdot)$ 为Gamma函数. 将式(13)中的 $\varphi(z_\alpha) / \alpha$ 替换成 k_α 可得到多维 t 分布下真实的均值-CVaR曲线.

结果见图2. 从图中可以得出以下结论: 首先, MN方法, 即传统方法, 是在正态分布假设下将样本均值和样本方差阵代入式(13)得到均值-CVaR曲线的估计, 如果真实分布不是正态分布(本例中是多维 t 分布) 这种方法将是有偏的, 图2直观地显示 MN方法系统地低估投资组合的风险; 其次, 对比图1和图2可以发现, 在相同的均值向量、方差阵(或散度)和损失概率下, 多维 t 分布下的极端(即小概率下)风险值要大于正态分布, 即在相同收益率下, 多维 t 分布下的有效前沿

的 CVaR 更大. 第三, NP方法和YL方法估计出来的均值-CVaR曲线与真实的均值-CVaR曲线比较接近, 而NP方法的估计偏差更小一些.

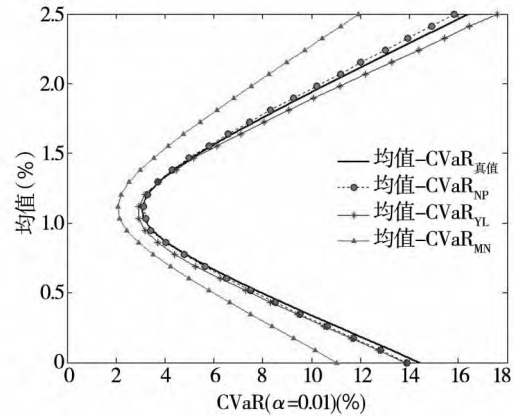


图2 多维 t 分布下的‘均值-CVaR’曲线估计

Fig. 2 Mean-CVaR curves estimation under multivariate t distribution

3.3 实例分析

本节选取我国A股市场10只股票的日收益率数据进行实例分析, 它们是: 万科A、深物业A、深深宝A、云南白药、铜陵有色、格力电器、罗牛山、承德露露、新希望和青岛啤酒. 数据期间为2007-01-01~2012-12-31, 由于在某些交易日, 有些股票由于各种原因会停牌, 所以必须选出那些每只股票都有交易的交易日收盘价数据, 经过删减和匹配处理后, 每只股票的可用日对数收益率数据为1 278个. 由于日对数收益率数据都很小, 为了方便, 把所有的日对数收益率数据都乘以100. 容易算得10只股票日对数收益率的样本均值和样本协方差阵分别为

$$\bar{r} = (-0.012 \ 1 \ 0.058 \ 1 \ 0.103 \ 9 \ 0.090 \ 9 \ 0.038 \ 1 \ 0.091 \ 3 \ 0.043 \ 6 \ 0.027 \ 9 \ 0.062 \ 3 \ 0.030 \ 6)$$

$$S^2 = \begin{pmatrix} 7.633 \ 5 & 0.089 \ 3 & -0.478 \ 5 & -0.084 \ 9 & 0.058 \ 5 & 0.396 \ 9 & -0.280 \ 5 & 0.104 \ 3 & 0.015 \ 4 & -0.173 \ 7 \\ 0.089 \ 3 & 11.684 \ 8 & -0.071 \ 8 & 0.190 \ 9 & 0.870 \ 7 & 0.515 \ 8 & -0.048 \ 2 & 0.885 \ 5 & -0.185 \ 0 & -0.016 \ 2 \\ -0.478 \ 5 & -0.071 \ 8 & 14.046 \ 0 & 0.118 \ 9 & -0.169 \ 3 & 0.562 \ 9 & 0.480 \ 7 & 0.265 \ 7 & -0.033 \ 3 & -0.268 \ 8 \\ -0.084 \ 9 & 0.190 \ 9 & 0.118 \ 9 & 4.901 \ 1 & -0.099 \ 2 & 0.145 \ 1 & 0.381 \ 6 & 0.682 \ 0 & 0.057 \ 8 & -0.055 \ 6 \\ 0.058 \ 5 & 0.870 \ 7 & -0.169 \ 3 & -0.099 \ 2 & 12.612 \ 0 & 0.139 \ 0 & 0.615 \ 9 & 0.608 \ 0 & 0.431 \ 1 & 0.266 \ 8 \\ 0.396 \ 9 & 0.515 \ 8 & 0.562 \ 9 & 0.145 \ 1 & 0.139 \ 0 & 6.994 \ 0 & 0.383 \ 6 & 0.554 \ 0 & -0.112 \ 1 & 0.381 \ 5 \\ -0.280 \ 5 & -0.048 \ 2 & 0.480 \ 7 & 0.381 \ 6 & 0.615 \ 9 & 0.383 \ 6 & 10.264 \ 6 & 0.316 \ 2 & 0.296 \ 8 & 0.648 \ 9 \\ 0.104 \ 3 & 0.885 \ 5 & 0.265 \ 7 & 0.682 \ 0 & 0.608 \ 0 & 0.554 \ 0 & 0.316 \ 2 & 6.986 \ 6 & -0.083 \ 0 & 0.178 \ 1 \\ 0.015 \ 4 & -0.185 \ 0 & -0.033 \ 3 & 0.057 \ 8 & 0.431 \ 1 & -0.112 \ 1 & 0.296 \ 8 & -0.083 \ 0 & 10.911 \ 3 & -0.022 \ 0 \\ -0.173 \ 7 & -0.016 \ 2 & -0.268 \ 8 & -0.055 \ 6 & 0.266 \ 8 & 0.381 \ 5 & 0.648 \ 9 & 0.178 \ 1 & -0.022 \ 0 & 5.282 \ 0 \end{pmatrix}$$

本文分别基于传统方法(即MN方法)、NP方法、YL方法和LP方法4种估计方法得到 $\alpha =$

5%时的组合边界(见图3). 传统方法假设这10支股票的联合分布为正态分布, 而另外3种估计

⑥ 多维 t 分布随机数的生成过程可见参考文献[26].

方法都不需要对分布进行事前设定. 在实践中, 事先并不知道这 10 只股票的收益率服从何种分布, 从估计的结果来看, 如果这 10 只股票的联合分布确实服从正态分布, 那么传统方法与另外 3 种方法得到的组合边界趋向一致, 图 3 并没有显示出这种趋势. 特别是在组合边界的下半部分, 传统方法估计出来的组合边界明显不同于其它 3 种估计方法得到的结果, 而其它 3 种估计方法得到的组合边界趋向于一致, 这也就说明这 10 只股票的联合分布不服从正态分布. 在不服从正态分布情况下, 传统方法将是有偏的, 它给出的组合边界不是最有效的, 也就是说它基于正态分布假设给出的 CVaR 估计式本身是不恰当的, 所以在优化模型里得到的最小化 CVaR 并不是真正最小 CVaR.

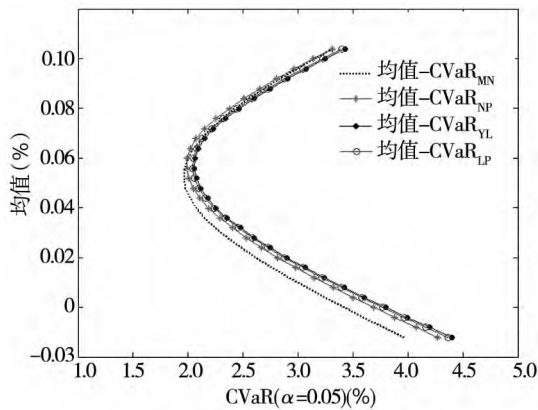


图 3 中国 A 股市场的‘均值-CVaR’曲线估计
Fig. 3 Mean-CVaR curves estimation of Chinese A-share market

4 引入无风险资产

4.1 模型与求解

在两步核估计的框架下, 模型 $\Lambda_1 \sim$ 模型 Λ_3 里的理论 CVaR 都用其两步核估计量 $\hat{u}(x, \alpha)$ 代替, 投资组合收益率的期望用其样本均值代替. 以模型 Λ_1 为例, 在核估计框架下模型变为

$$(\Lambda_1) \begin{cases} \min_{x \in \mathbb{R}^n} & -(1 - x'e) r_f + \hat{u}(x, \alpha) \\ \text{s. t.} & \bar{R} = (1 - x'e) r_f + x' \bar{r} \geq \beta \end{cases}$$

由于 $\hat{u}(x, \alpha)$ 的两步核估计需用 VaR 的核估计量 $\hat{v}(x, \alpha)$ 和窗宽 h , 所以模型 $\Lambda_1 \sim \Lambda_3$ 需满足

前述的条件 1 和条件 2. 同问题 Ψ_1^n 一样, 很难得到问题 Λ_1 的解析解, 但能求出其数值解, 求解算法与 3.1 节相同, 只是引入无风险资产后拉格朗日函数里的变量少了 1 个, 在迭代算法步骤 2, 组合 CVaR 的计算公式变成

$$\hat{u}(r_f, x_k, \alpha) = -(1 - x_k'e) r_f + \hat{u}(x_k, \alpha)$$

4.2 蒙特卡洛模拟

模拟 3 仿照基于均值-方差模型的资本市场线定义, 本文把存在无风险资产时均值-CVaR 模型的有效边界定义为资本市场线 (capital market line, CML). 假设存在 3 种风险资产和 1 种无风险资产, 风险资产的收益率服从联合正态分布 $N(\mu, \Sigma)$, μ 和 Σ 取值同上. 无风险资产的收益率为 0.5, 取样本容量为 4 000, 损失概率 5%. 将真实参数代入式 (13) 和式 (14) 可以得到真实的组合边界曲线和 CML, 基于两步核估计框架可以得到估计出来的组合边界曲线和 CML. 从模拟结果来看 (见图 4), 基于两步核估计方法得到的 CML 和真实 CML 非常接近, 说明利用两步核估计方法对 CVaR 进行估计是合适的, 引入无风险资产后, 本文的模型和算法也是可行且准确的.

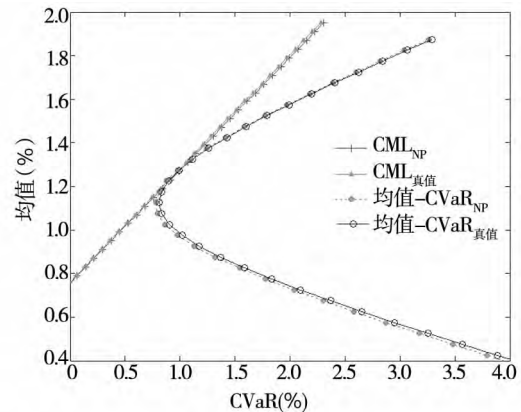


图 4 均值-CVaR' 曲线和 CML (α=5%)
Fig. 4 Mean-CVaR curves and capital market lines (α=5%)

4.3 实例分析

继续使用 3.3 节的 10 只股票作为风险资产, 取无风险资产收益率为 1%^⑦, 分别估计损失概率为 3% 和 5% 时的组合边界曲线与 CML. 值得注意的是, 与基于均值-方差模型的资本市场线不同, 基于均值-CVaR 模型的资本市场线不过点 (0,

⑦ 由于前面 10 只股票的日收益率都乘以 100, 此处无风险资产的日利率取为 1% 是合适的.

r_f) 而是过点 $(-r_f, r_f)$ ⑧, 模拟结果也证实了这一点, 即 CML 与纵轴交点的纵坐标不是 1% (见图 5). 另外, 虽然没有严格的证明, 但模拟的结果显示在各自的损失概率下, 组合边界曲线和 CML 相切, 这一点说明在非正态分布情况下, CAPM 依然成立. 最后值得一提的是, 这里的估计结果完全是利用样本数据基于本文的两步核估计方法建模并通过简单迭代算法计算得到, 不需要任何的分布假设和参数设定, 只需要一组样本数据, 前面的蒙特卡洛模拟说明这种方法是准确有效的.

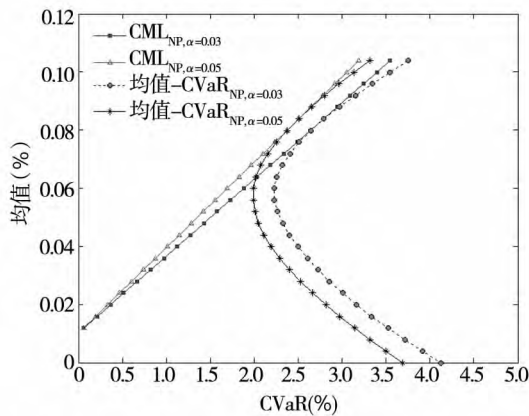


图 5 中国 A 股市场的‘均值-CVaR’曲线和 CML 估计
Fig. 5 Mean-CVaR curves and capital market lines of Chinese A-share market

5 结束语

一直以来, 金融风险的估计和基于均值 -

风险模型的投资组合选择理论相对独立发展, 本文尝试将二者结合起来, 建立一个完整框架实现风险估计与投资组合优化同步进行. 近年来核估计方法因具有模型设定灵活、能够抓住金融序列的尾部风险特征、允许数据序列相依等优点, 备受国内外学者的关注, 是对金融市场风险进行估计的理想方法. 本文不需对分布函数做任何假设, 也不需要知道 CVaR 的函数形式, 直接利用核估计方法得到 CVaR 的两步核估计式, 进而将 CVaR 的两步核估计式嵌入均值-CVaR 模型, 并设计个简单算法对该模型进行求解. 蒙特卡洛模拟结果表明基于两步核估计方法的模型和算法得到的结果同已知真实分布情况下的结果高度一致, 这要归功于非参数核估计方法能够很好地拟合分布函数. 进一步, 基于我国 A 股市场数据的实例分析说明我国金融市场的收益率数据不服从正态分布, 基于正态分布假设的投资组合选择模型会误导投资者.

当然, 如何结合风险估计和投资组合优化是目前学者思考的前沿课题, 本文的研究是初步的. 下一步的研究可以沿着两个方向进行: 1) 将本文核估计框架扩展到其它风险度量指标, 如 VaR、半方差、下偏矩、绝对偏差、安全首要、EaR、CaR、WCT 等; 2) 放松本文的假设, 加入卖空限制、不同借贷利益、存在交易费用等约束, 在更加现实的市场环境下拓宽本文的研究.

参考文献:

[1] Markowitz H. Portfolio selection [J]. Journal of Finance, 1952, 7(1): 77-91.
 [2] Jorion P. Value at Risk [M]. 2nd Edition. New York: McGraw-Hill, 2001.
 [3] Artzner P, Delbaen F, Eber J M, et al. Thinking coherently [J]. Risk, 1997, (10): 68-71.
 [4] Artzner P, Delbaen F, Eber J M, et al. Coherent measures of risk [J]. Mathematical Finance, 1999, 9(3): 203-228.
 [5] Rockafellar R T, Uryasev S. Optimization of conditional value-at-risk [J]. The Journal of Risk, 2000, 2(3): 21-41.
 [6] Rockafellar R T, Uryasev S. Conditional value-at-risk for general loss distributions [J]. Journal of Banking and Finance, 2002, 26(7): 1443-1471.
 [7] Alexander G, Baptista A. Economic implications of using a mean-var model for portfolio selection: A comparison with mean-variance analysis [J]. Journal of Economic Dynamics and Control, 2002, 26(7/8): 1159-1193.
 [8] 姚京, 李仲飞. 基于 VaR 的金融资产配置模型 [J]. 中国管理科学, 2004, 12(1): 8-14.

⑧ 此时所有资产投在无风险资产上, 组合的收益率为 r_f , 组合的 CVaR 为 $-r_f$; 实际上, 当风险资产的收益率服从联合正态分布时, 可以从式 (14) 直接推出这个结论.

- Yao Jing , Li Zhongfei. The asset allocation model based on VaR [J]. Chinese Journal of Management Science , 2004 , 12 (1) : 8 - 14. (in Chinese)
- [9] Engle R F , Manganelli S. CAViaR: Conditional autoregressive value at risk by regression quantiles [J]. Journal of Business & Economics Statistics , 2004 , 22 (4) : 367 - 381.
- [10] 林宇, 黄登仕, 魏宇. 胖尾分布及长记忆下的动态 EVT-VaR 测度研究 [J]. 管理科学学报, 2011, 14(7): 71 - 82.
- Lin Yu , Huang Dengshi , Wei Yu. Study on financial markets dynamic EVT-VaR measuring based on fated-tail distribution and long memory volatility [J]. Journal of Management Sciences in China , 2011 , 14(7) : 71 - 82. (in Chinese)
- [11] 叶五一, 缪柏其. 已实现波动与日内价差条件下的 CVaR 估计 [J]. 管理科学学报, 2012, 15(8): 60 - 71.
- Ye Wuyi , Miao Baiqi. Estimating of CVaR with consideration of realized volatility and price range [J]. Journal of Management Sciences in China , 2012 , 15(8) : 60 - 71. (in Chinese)
- [12] Chen S X , Tang C Y. Nonparametric inference of value-at risk for dependent financial returns [J]. Journal of Financial Econometrics , 2005 , 3(2) : 227 - 255.
- [13] 赵晓玲, 陈雪蓉, 周勇. 金融风暴中基于非参估计 VaR 和 ES 方法的风险度量 [J]. 数理统计与管理, 2012, 31(3): 381 - 383.
- Zhao Xiaoling , Chen Xuerong , Zhou Yong. Risk measure based on nonparametric estimate of VaR and ES in financial crisis [J]. Journal of Applied Statistics and Management , 2012 , 31(3) : 381 - 383. (in Chinese)
- [14] Bellini F , Figa-Talamanca G. Detecting and modeling tail dependence [J]. International Journal of Theoretical and Applied Finance , 2004 , 7(3) : 269 - 287.
- [15] Gourieroux C , Laurent J P , Scaillet O. Sensitivity analysis of values at risk [J]. Journal of Empirical Finance , 2000 , (7) : 225 - 245.
- [16] Scaillet O. Nonparametric estimation and sensitivity analysis of expected shortfall [J]. Mathematical Finance , 2004 , 14 (1) : 115 - 129.
- [17] Scaillet O. Nonparametric estimation of conditional expected shortfall [J]. Insurance and Risk Management Journal , 2005 , (74) : 639 - 660.
- [18] Chen S X. Nonparametric estimation of expected shortfall [J]. Journal of Financial Econometrics , 2008 , 6(1) : 87 - 107.
- [19] 刘静, 杨善朝. 风险度量 ES 的非参数估计 [J]. 工程数学学报, 2009, 26(4): 577 - 585.
- Liu Jing , Yang Shanchao. Nonparametric estimation of expected shortfall [J]. Chinese Journal of Engineering Mathematics , 2009 , 26(4) : 577 - 585. (in Chinese)
- [20] 刘晓倩, 周勇. 金融风险度量 ES 度量的非参数方法的比较及其应用 [J]. 系统工程理论与实践, 2011, 31(4): 631 - 642.
- Liu Xiaoqian , Zhou Yong. Nonparametric estimation and comparative analysis of ES in risk measure with applications [J]. Systems Engineering-Theory & Practice , 2011 , 31(4) : 631 - 642. (in Chinese)
- [21] Yao H X , Li Z F , Lai Y Z. Mean-CVaR portfolio selection: A nonparametric estimation framework [J]. Computers & Operations Research , 2013 , 40(4) : 1014 - 1022.
- [22] Li Q , Racine J S. Nonparametric Econometrics: Theory and Practice [M]. Princeton: Princeton University Press , 2007.
- [23] 刘小茂, 李楚霖. 资产组合的 CVaR 风险的敏感度分析 [J]. 数学物理学报, 2004, 24A(4): 442 - 448.
- Liu Xiaomao , Li Chulin. Sensitivity analysis of conditional value at risk [J]. Acta Mathematica Scientia , 2004 , 24A(4) : 442 - 448. (in Chinese)
- [24] Krokmal P , Palmquist J , Uryasev S. Portfolio optimization with conditional value-at-risk objective and constraints [J]. Journal of Risk , 2002 , 4(4) : 43 - 68.
- [25] 刘小茂, 李楚霖, 王建华. 风险资产组合的均值-CVaR 有效前沿 (II) [J]. 管理工程学报, 2005, 19(1): 1 - 5.
- Liu Xiaomao , Li Chulin , Wang Jianhua. Mean-CVaR efficient frontier and its economic implications (II) [J]. Journal of Industrial Engineering/Engineering Management , 2005 , 19(1) : 1 - 5. (in Chinese)
- [26] Cui X T , Zhu S S , Sun X L , et al. Nonlinear portfolio selection using approximate parametric Value-at-Risk [J]. Journal of Banking & Finance , 2013 , 37(6) : 2124 - 2139.

Investment portfolio management based on the two-step kernel estimator of CVaR

HUANG Jin-bo¹, LI Zhong-fei^{2*}, YAO Hai-xiang³

1. School of Finance, Guangdong University of Finance & Economics, Guangzhou 510320, China;
2. Sun Yat-Sen Business School, Sun Yat-Sen University, Guangzhou 510275, China;
3. School of Finance, Guangdong University of Foreign Studies, Guangzhou 510006, China

Abstract: The paper first applies nonparametric kernel estimation method to estimating CVaR which is currently a popular risk measurement tool, then derives a two-step kernel estimator of CVaR with distribution-free specification. Next, a two-step kernel estimator of CVaR is embed into the mean-CVaR portfolio optimization models to derive financial risk estimation and portfolio optimization at the same time. A simple iterative algorithm is designed to solve these models. Monte Carlo simulation result shows that the portfolio optimization models and the algorithm based on the two-step kernel estimator of CVaR is feasible and effective, and that the estimated error of portfolio frontier is very small. The models and algorithm above apply to a risk-free security. Finally, an empirical analysis of daily return data from Chinese A-stock market is presented to illustrate the application of this research.

Key words: mean-CVaR model; two-step kernel estimator; portfolio frontier; Chinese A-stock market

(上接第 113 页)

Cashflow news, cashflow risk and pricing of stock returns

SU Cheng-jian^{1,2}

1. School of Finance of Guizhou University of Finance and Economics, Guiyang 550025, China
2. Department of Finance, Jincheng College of Sichuan University, Chengdu 611731, China;

Abstract: The paper derives and tests the relationship between current-period unexpected returns and unexpected earnings that incorporates cashflow news and cashflow risk. A new multi-variable model based on a three-variable model is established to explain the unexpected individual stock returns. The paper estimates the model using data for individual stocks in China's market from 2002 to 2011. The main findings are: (i) The result of the cashflow news derived by revisions in forecasts of future earnings is an important determinant of cross-sectional unexpected stock returns and realized stock returns and excess stock returns; (ii) The cashflow risk thus derived to reflect systematic risk can explain expected stock returns; (iii) The new multi-variable model based on the three-variable of model and considering cashflow news and cashflow risk dominates the multi-variable model based based on the three-factor model Fama-French and considering cashflow news and cashflow risk in explaining excess stock returns.

Key words: expected stock returns; unexpected stock returns; cashflow news; cashflow risk