

# 银行信用组合风险多成分重要性抽样算法研究<sup>①</sup>

龚 朴, 邓 洋, 胡祖辉

(华中科技大学管理学院, 武汉 430074)

摘要: 银行信用组合违约风险的度量和计算对银行监管有着重要的意义. 使蒙特卡洛研究信用组合违约概率时, 为提高模拟效率, 越来越多的学者采用了重要性抽样技术来实现. 它主要通过条件独立性和“均值移动”两个步骤实现. 本文基于前人研究结果的基础之上, 提出了一种基于违约相关性矩阵的多因子变方差的重要性抽样算法. 该算法通过主成分分析选择违约结构中的占优成分并扩大其方差来实现. 数值算例证明了该方法在信用组合遭遇极值事件时, 能够提高模拟效率及计算精度, 具有一定的计算优势.

关键词: 信用组合风险; 蒙特卡洛模拟; 重要性抽样; 主成分分析

中图分类号: F832.33 文献标识码: A 文章编号: 1007-9807(2012)11-0003-08

## 0 引 言

由债务人违约而导致的损失, 也就是人们常说的信用风险, 是包括银行在内的所有金融机构重点关心的问题之一. 因此, 准确度量包含各类金融资产的投资组合的信用风险获得了学术界的广泛关注. 近年来, 随着信用衍生品市场的发展和成熟, 信用风险的流动性逐渐增加. 这种新的金融市场环境为对冲或投机信用风险提供了更多可能的机会, 迫使商业银行精确地计算由贷款和债券等资产构成的投资组合的违约风险, 更加积极地管理其信用组合风险敞口.

商业银行的风险敞口一般来自于不同的公司、地区和行业. 例如, 商业银行普遍偏好于贷款给某个地区的公司, 或者某个行业的公司, 以减小信息不对称所带来的额外风险, 这种情况将导致当其中一个贷款违约时, 极有可能引起其它贷款同时违约的骨牌效应<sup>[1, 2]</sup>. 新古典金融框架下, 对信用资产违约事件的建模大致可分为结构化模型和简约化模型两类<sup>[3]</sup>. 其中, 结构化模型为构建

其它更简单的违约模型提供了一种重要的概念性框架. 许多较为成功的商业信用风险管理软件成功地使用了这一框架, 如 KMV 商业信用风险管理模型等. 在此框架下计算信用组合的违约概率时, 困难之处在于如何从投资组合管理的角度处理信用间的违约相关性<sup>[4]</sup>, 同时提高计算速度, 降低计算成本.

蒙特卡洛模拟可以用来处理较为复杂的组合违约相关性, 然而, 计算时间较长. 对于一个具有较高信用等级银行所持有的贷款组合而言, 估计小概率大损失事件的发生概率会需要数天的计算时间. 因此, 直接使用蒙特卡洛方法效率不高, 降低了信用组合压力测试的可行性. 所以, 各种加速方法, 如重要性抽样算法, 得到了越来越多学者的关注. Glasserman 和 Li<sup>[5]</sup>率先提出了一种两阶段重要性抽样(importance sampling, IS)算法: 首先, 在经典 Credit Metrics 模型<sup>[6]</sup>下得到单因素条件独立的违约事件; 然后, 对条件独立违约事件使用指数抽样算法, 结果表明重要性抽样算法提高了计算速度且计算效率在很大程度上受信用事件

① 收稿日期: 2012-07-11; 修订日期: 2012-09-05.

基金项目: 国家自然科学基金资助项目(71071067); 教育部高等学校博士学科点专项科研基金资助项目(20110142110068).

作者简介: 龚 朴(1954—), 男, 湖北武汉人, 教授, 博士生导师. Email: gongpu@mail.hust.edu.cn

之间的违约相关性的影响. Glasserman 等<sup>[7,8]</sup>将上述工作扩展至多因素模型,用于处理不同信用等级度的银行贷款组合,并将异质性信用组合违约损失问题归结至一类“最小子集的问题”.在此框架下,学者们从不同的角度研究了蒙特卡洛的计算效率. Glasserman<sup>[9]</sup>使用蒙特卡洛模拟和重要性抽样算法研究了信用组合中边际风险的贡献率问题. Dunkel 等<sup>[10]</sup>讨论了如何有效地使用蒙特卡洛模拟估计凸风险度量. Bassamboo 等<sup>[11]</sup>将极值相依关系引入模型,并给出了渐进有效的重要性抽样算法.与上述工作不同的是, Huang 和 Oosterlee<sup>[12]</sup>采用自适应蒙特卡洛积分替代了蒙特卡洛模拟和重要性抽样算法.

值得注意的是,这些模型在计算信用组合违约概率时都需要条件独立性,以便使样本尽可能集中在对计算结果有影响的重要性区域.然而,在参数空间中确定重要性区域的方位和大小涉及到一系列较为困难的优化问题<sup>[13,14,15]</sup>,导致算法求解复杂,计算效率较差. Morokoff<sup>[16]</sup>指出在计算信用组合的违约概率时,条件独立性并不是必要的.在此基础上,他构建了单因素常方差重要性抽样模型.但是,该模型在组合遭遇极值损失事件(如次债危机)时,计算效率将会下降.本文运用主成分分析的基本原理,将其工作推广至多个占优因子的情形,并给出了多因子情形下的最优方差扩大倍数.数值算例说明多因子变方差扩大模型能够较好地处理极端损失事件,改善计算效率,实现降低计算成本的目的.

### 1 问题提出

考虑某商业银行持有由贷款、债券等  $N$  个资产构成的投资组合.假定第  $i (i \in [1, N])$  个资产的边际违约概率为  $p_i \in (0, 1)$  (在实际中  $p_i$  一般通过与该标的资产相关的历史数据和计量经济学模型所得).假设存在表示相应的标的资产的经济价值的代理变量  $X = (X_1, \dots, X_N)$ , 设定当  $X_i$  小于某个给定的阈值  $x_i$  时,第  $i$  个资产违约,记其违约损失为  $l_i$ <sup>[5]</sup>.

令  $f_{X_i}(x)$  表示资产  $i$  的边际密度函数,则资产的违约阈值和违约概率存在如下关系

$$P\{X_i < x_i\} = \int_{-\infty}^{x_i} f_{X_i}(u) du = p_i \quad (1)$$

商业银行的资产组合的违约损失可表示为

$$L = \sum_{i=1}^N l_i 1_{\{X_i < x_i\}} \quad (2)$$

其中  $1_{\{\cdot\}}$  为示性函数.

那么组合发生大损失的概率  $l(x)$  为

$$l(x) = \Pr(L > x) \quad (3)$$

由式(2)可看到,标的资产违约事件  $1_{\{X_i < x_i\}}$  之间的相关关系可以由代理变量  $X_i$  之间的相关关系决定.更具体的,相关性模型通常经由一个线性因子模型来给出<sup>[15]</sup>.

$$X_i = a_{i1}Z_1 + \dots + a_{id}Z_d + b_i\varepsilon_i \quad (4)$$

其中  $Z_1, \dots, Z_d$  为相互独立且同分布的标准正态变量,其经济意义为所有资产都面临的系统风险;  $\varepsilon_i$  是均值为 0 的正态随机变量,且与  $Z_i$  是相互独立,表示第  $i$  个资产的特有风险.同时因子载荷满足  $a_{i1}^2 + \dots + a_{id}^2 + b_i^2 = 1$ .

Glasserman 和 Li<sup>[5]</sup> 以及 Glasserman 和 Wanmo<sup>[7,8]</sup> 的重要性区域是通过“移动”代理变量  $X_i$  的分布的均值而得到.其存在的问题是,如果模型(4)中影响因子较多,如何在多维空间中确定分布均值的移动方向和大小是一个较难处理的优化问题. Morokoff<sup>[16]</sup>指出通过扩大模型中占优成分的方差也能够得到重要性区域,达到使计算样本将更多地来自联合分布的尾部的目的,且这种处理较为直观并避免了复杂的优化过程.但是,直接扩大某个因子或某几个因子的方差并不能带来令人满意的结果,其原因在于其它因子的取值可能会消除方差扩大的影响.

主成分分析法 (principal components analysis, PCA) 可以将所有因子按照各自的相对重要性程度组成新的成分,再按照某一特殊比例关系同时扩大所有因子的方差,以满足算法的需要.根据 Chen 和 Glasserman<sup>[17]</sup>,代理变量  $X$  的相关性矩阵可以表示为

$$\Sigma = AA^T + BB^T \quad (5)$$

其中矩阵  $A$  表示模型中的载荷矩阵,矩阵  $B$  是载荷  $A$  的补充矩阵.显然,相关性矩阵  $\Sigma$  的主对角线为 1.一般而言,相关性矩阵  $\Sigma$  应至少具有  $m (m \leq M)$  个正的特征值  $\lambda_1 \geq \dots \geq \lambda_m > 0$ .令其对应的正交标准特征向量表示为  $\{v_1, \dots, v_m\}$ .

假设需要扩大  $m$  个占优因子的方差. 由 PCA 的基本原理可知, 相应成分的相对重要性可以由特征值  $\lambda_i$  决定. 特征值  $\lambda_i$  越大, 其对总方差的解释力度越强. 因此, 本文将使用特征值的大小关系来确定成分之间的占优关系: 在识别出  $m$  个占优成分之后, 扩大相应的  $\lambda_i$ , 处理后相当于扩大了分布的方差. 在模拟过程中将会有更多的样本取自那些对计算结果有影响的重要性区域, 进而提高模拟效率.

## 2 多成分变方差重要性抽样算法

重要性抽样技术是从初始的概率度量到一个新的、有偏的、能够更好的满足模拟需要的概率度量的变化<sup>[18]</sup>. 其基本原理为: 通过改变随机变量的分布函数, 使更多的样本来自于对计算结果有影响的区域. 假设需要估计的函数为  $E_f[h(W)]$ , 其中  $f$  表示随机变量  $W$  的密度函数,  $h(\cdot)$  为相应概率密度下可测函数. 假设  $g(\cdot)$  是改变后的密度函数, 当  $f(W) > 0$  时, 有  $g(W) > 0$ . 则  $E_f[h(W)]$  的重要性抽样估计量为

$$E_f[h(W)] = \int h(w)f(w) dw = \int h(w) \frac{f(w)}{g(w)} g(w) dw$$

$$= E_g \left[ h(w) \frac{f(w)}{g(w)} \right] \quad (6)$$

其中  $f(w)/g(w)$  为似然比率. 式(6)说明可通过模拟服从  $g(\cdot)$  的随机变量  $W$ , 通过  $h(w) \frac{f(w)}{g(w)}$  来无偏地估计原估计量.

选择一个合适的重要性抽样分布与待解决的问题有着紧密的关系. 如果新分布函数使得样本过于集中在某些区域, 可能会导致模拟精度的下降. 针对这个问题, 本文通过扩大占优因子的方差, 实现使分布的尾部变厚的效果, 可以避免出现样本过于集中的情形.

从第 1 部分的分析可知, 代理变量  $X_i$  的分布在处理前后发生了变化, 符合重要性抽样算法的工作原理. 处理后, 代理变量的联合分布仍然是多元正态分布, 唯一的差别在于它们的协方差矩阵特征值不尽相同.

假设新概率分布的协方差矩阵的特征值为  $\tilde{\lambda}_i = \alpha_i^2 \lambda_i, i = 1, \dots, m; \tilde{\lambda}_i = \lambda_i, i = m + 1, \dots, N$ , 那么新的协方差矩阵可以表示为

$$\tilde{\Sigma} = V\tilde{\Lambda}V^T \quad (7)$$

其中  $\tilde{\Lambda}$  为主对角元素为  $\tilde{\lambda}_1, \dots, \tilde{\lambda}_N$  的对角矩阵. 矩阵  $V$  的第  $i$  列元素为  $v_i$ . 那么似然比率可以表示为

$$\omega(x_{\tilde{\Sigma}}) = \sqrt{\frac{|\tilde{\Sigma}|}{|\Sigma|}} \exp\left(-\frac{1}{2} x_{\tilde{\Sigma}}^T (\Sigma^{-1} - \tilde{\Sigma}^{-1}) x_{\tilde{\Sigma}}\right) \quad (8)$$

其中  $x_{\tilde{\Sigma}}$  表示协方差矩阵为  $\tilde{\Sigma}$  的多元正态随机变量. 运用  $\Sigma = V\Lambda V^T, \tilde{\Sigma} = V\tilde{\Lambda}V^T$  和相应的标准正交条件, 得到

$$\omega(x_{\tilde{\Sigma}}) = \prod_{i=1}^m \alpha_i \exp\left(-\frac{1}{2} \sum_{i=1}^m \left(1 - \frac{1}{\alpha_i^2}\right) \frac{(x_{\tilde{\Sigma}}^T v_i)^2}{\lambda_i}\right) \quad (9)$$

为了能够直接比较多个重要性抽样算法的计算结果, 需要将其转化到未做概率变换的分布上, 也将似然比率表达为关于初始向量  $x_{\Sigma}$  的函数.

记对角矩阵为  $D$ , 其主对角元素分别为  $d_{ii} = \alpha_i, i = 1, \dots, m; d_{ii} = 1, i = m + 1, \dots, N$ , 则

$$x_{\tilde{\Sigma}} = (VDV^T)x_{\Sigma} \quad (10)$$

所以, 有

$$x_{\tilde{\Sigma}} = \sum_{i=1}^m (\alpha_i - 1) (v_i^T x_{\Sigma}) v_i + x_{\Sigma} \quad (11)$$

注意到

$$v_i v_j^T = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad (12)$$

有

$$\omega(x_{\Sigma}) = \prod_{i=1}^m \alpha_i \exp\left(-\frac{1}{2} \sum_{i=1}^m (\alpha_i^2 - 1) \frac{(x_{\Sigma}^T v_i)^2}{\lambda_i}\right) \quad (13)$$

假设  $\alpha_i$  已知, 通过利用标准正交性条件将 Morokoff<sup>[16]</sup> 推广至一般情形. 对于随机变量  $x_{\Sigma}$ , 如果  $\Sigma = CC^T$ , 则可以使用标准正态随机变量模拟  $x_{\Sigma} = Cz$ . 通常, 从新分布中抽取随机变量需要额外的计算成本. 然而, 式(11)说明了本节提出的重要性抽样算法所需的额外计算成本可以忽略不计. 在此新框架中, 若令  $\alpha = 1$ , 即为直接蒙特卡洛模拟算法; 若令  $m = 1$  且  $\alpha = \sqrt{2}$ , 即可得到 Morokoff<sup>[16]</sup> 的单因子常方差重要性抽样算法. 也就是说, 本文构建的多成分变方差重要性抽样算法能够同时包含直接蒙特卡洛模拟和 Morokoff 重要性抽样算法作为其特例. 如果对三个算法在模拟时使用同样的随机数, 即可以排除随机数的影

响,进而直接比较其计算效率.

Morokoff<sup>[16]</sup> 认为随机变量  $v_1 x_\Sigma / \sqrt{\lambda_1}$  是一个服从标准正态分布的随机变量. 那么,似然比率也可以看作服从如下分布的随机变量

$$F_\omega(w) = P(\omega < w) = 2\left(1 - \Phi\left(\sqrt{\frac{-2\log(w/\alpha)}{\alpha^2 - 1}}\right)\right) \quad (14)$$

并假定随着违约损失的增加似然比率值将会减小. 然而,该假设并不总是成立,信用组合的违约相关性与违约损失之间有着较强的相互关系. 数值算例说明了这种方法在当信用组合违约损失超过自身价值一半时,计算效率将下降.

本文通过最小化待估计值的方差,或者二阶矩来选择  $\alpha_i$ . 信用违约损失概率的二阶矩可以表示为

$$M_2(x) = M_2(x, \alpha_i) = E_\Sigma[1\{L > x\} \omega^2(x)] \quad (15)$$

所以

$$E_\Sigma[1\{L > x\} \omega^2(x)] = E_\Sigma[1\{L > x\} \omega(x)] \leq E_\Sigma[\omega(x)] \quad (16)$$

有

$$E_\Sigma[\omega(x)] = \int_{R^N} \frac{\prod_{i=1}^m \alpha_i}{\sqrt{(2\pi)^N |\Sigma|}} \times \exp\left(-\frac{1}{2}\left(\sum_{i=1}^m (\alpha_i^2 - 1) \frac{(x_\Sigma^T v_i)^2}{\lambda_i} + x^T \Sigma^{-1} x\right)\right) dx \quad (17)$$

相关性矩阵  $\Sigma$  是半正定的,可以表示为  $\Sigma = V \Lambda V^T$ , 它的逆可以表示成为  $\Sigma^{-1} = V^T \Lambda^{-1} V$ , 其中  $\Lambda^{-1}$  是一个对角矩阵,其主对角元素为  $1/\lambda_i$ , 或者零. 注意到  $\lambda_i > 0$ , 因此对于  $x$  的标准二次型, 当对角矩阵的某些主对角元素设为零时, 有  $x^T \Lambda^{-1} x \geq x^T (\Lambda^T)^{-1} x$ . 所以

$$-\frac{1}{2}\left(\sum_{i=1}^m (\alpha_i^2 - 1) \frac{(x_\Sigma^T v_i)^2}{\lambda_i} + x^T \Sigma^{-1} x\right) \leq -\frac{1}{2}\left(\sum_{i=1}^m \frac{\alpha_i^2}{\lambda_i} (x_\Sigma^T v_i)^2\right) \quad (18)$$

此时

$$\sum_{i=1}^m \frac{\alpha_i^2}{\lambda_i} (x_\Sigma^T v_i)^2 = \sum_{j=1}^N \sum_{i=1}^m \frac{\alpha_i^2}{\lambda_i} (v_{i,j} x_{jN} + v_{i,j+1} x_{j+1N})^2 - \sum_{j=1}^N \sum_{i=1}^m \frac{\alpha_i^2}{\lambda_i} (v_{i,j} x_j)^2 \quad (19)$$

( $\cdot \lfloor N$ ) 为取余操作.  $v_{i,j}$  表示向量  $v_i$  的第  $j$  个元

素. 因此,有

$$M_2(x) \leq \int_{R^N} \frac{\prod_{i=1}^m \alpha_i}{\sqrt{(2\pi)^N |\Sigma|}} \times \exp\left(-\frac{1}{2} \sum_{j=1}^N \sum_{i=1}^m \frac{\alpha_i^2}{\lambda_i} (v_{i,j} x_j)^2\right) dx \quad (20)$$

当  $\alpha_i = \sqrt{\lambda_i}$  时,  $M_2(x)$  表示为  $N$  个正态概率累积函数的乘积,所以

$$M_2(x) \leq \frac{\prod_{i=1}^m \sqrt{\lambda_i}}{\sqrt{|\Sigma|}} \prod_j \sqrt{\sum_{i=1}^m v_{i,j}^2} \quad (21)$$

同时,由于相关性矩阵的半正定性,可以选择

$m = N$ , 有  $|\Sigma| = \prod_{i=1}^N \lambda_i$ , 则

$$M_2(x) \leq \sqrt{\sum_{i=1}^N v_{i,j}^2} \quad (22)$$

近似地,  $\sum_{i=1}^N v_{i,j}^2 \leq 1$ , 因此  $M_2(x) \leq 1$ .

特别的,如果令  $m = 1$ , 那么

$$M_2(x) \leq \sqrt{\frac{\lambda_1}{|\Sigma|}} \quad (23)$$

通过以上分析,可以得到  $M_2(x)$  的一个上限. 尽管上限仍然取决于相关性矩阵  $\Sigma$ , 却是在有限范围内的. 如果标的信用间的相关性很高,相关性矩阵行列式的值较小,特征值  $\lambda_1$  也较小,因此,上限值较小. 如果标的信用间的相关性较低,相关性矩阵行列式的值较大,此时特征值  $\lambda_1$  也相对较大.

### 3 数值检验

本节将通过数值算例讨论和比较直接蒙特卡洛模拟方法、Morokoff 重要性抽样算法和多成分方差扩大重要性抽样的适用性并对其计算效率进行对比. 第一个数值算例取自 Glasserman 和 Jingyi<sup>[5]</sup>: 假设信用组合由  $N = 1\,000$  个信用组成, 含有 10 个因子. 信用的边际违约概率和边际违约损失由下式给出

$$p_i = 0.01 \times \left(1 + \sin \frac{16i\pi}{N}\right) \quad i = 1, \dots, N \quad (24)$$

$$l_i = \left[\frac{5i}{N}\right]^2, i = 1, \dots, N \quad (25)$$

因子载荷系数  $a_{ij}$  由相互独立的均匀分布  $(0, 1/$

$\sqrt{d}$ )  $d = 10$  构成. 它代表了一个典型的异质信用组合.

图 1 给出了直接蒙特卡洛模拟算法 (plain MC), Morokoff 重要性抽样算法 (Morokoff MC), 单一成分重要性抽样 (PCA 1 MC) 和两成分重要性抽样算法 (PCA 2 MC) 的结果对比. 更多成分的重要性抽样算法在计算过程中会产生数据向下溢出现象, 因此不再对比其计算结果. 所有的蒙特卡洛模拟均使用了相同的随机数. 为了说明计算结果的稳定性, 直接蒙特卡洛模拟使用的路径数为 10 000 次, 重要性抽样算法使用的路径数为 1 000 次.

为了准确地区分违约损失概率及其置信区间, 将违约损失概率的样本方差置于图 2 至图 5

中, 并对直接蒙特卡洛模拟和重要性抽样算法计算了相同违约损失发生的概率. 本文发现, 当违约损失较小时, 重要性抽样算法 (无论是 Morokoff MC 还是 PCA MC) 比直接蒙特卡洛模拟算法效率要低. 这一点可以从图中观察样本的方差得出. 然而, 随着违约损失的增大, 重要性抽样算法变得越来越具有效率. 当违约损失超过某一阈值时, PCA 重要性抽样算法变得比 Morokoff 重要性抽样算法有效. 随着违约损失的进一步扩大, PCA 2 IS 变成了最具效率的重要性抽样算法. 这也意味着, 当违约损失越来越大时, 单方差常数扩大算法效率下降. 随着违约损失的扩大, 应改变方差扩大倍数.

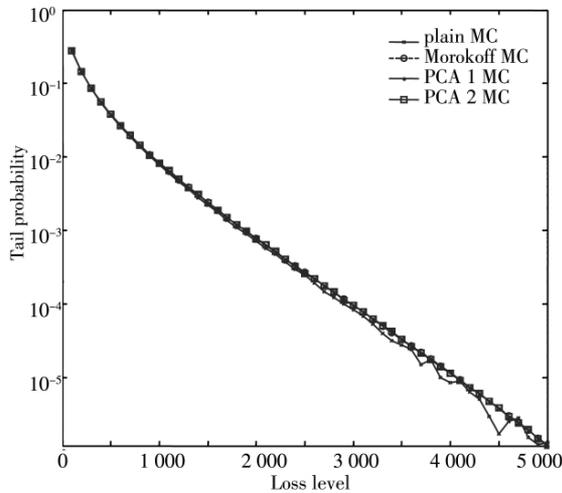


图 1 10 因子信用组合违约概率

Fig. 1 The default probability of credit portfolio with 10 factors

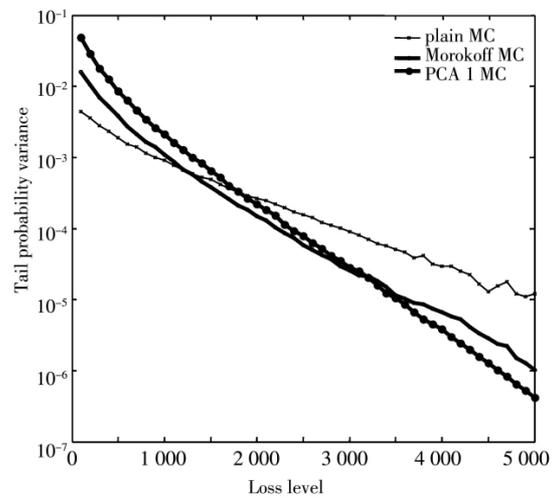


图 2 10 因子信用组合违约概率方差对比 1

Fig. 2 The variance of the estimated default probability of credit portfolio with 10 factors (a)

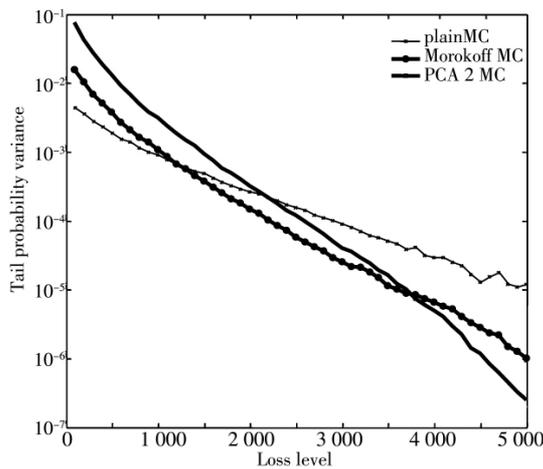


图 3 10 因子信用组合违约概率方差对比 2

Fig. 3 The variance of the estimated default probability of credit portfolio with 10 factors (b)

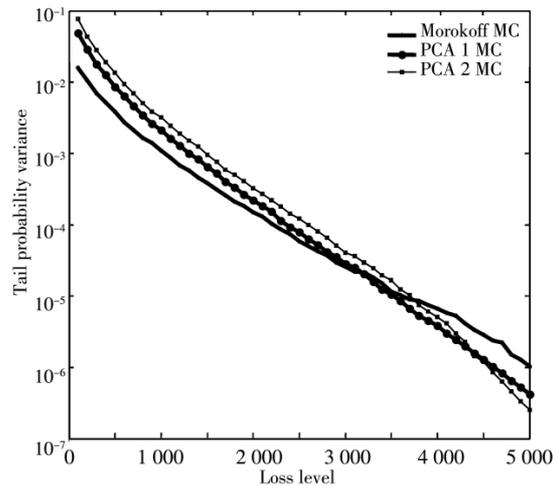


图 4 10 因子信用组合违约概率方差对比 3

Fig. 4 The variance of the estimated default probability of credit portfolio with 10 factors (c)

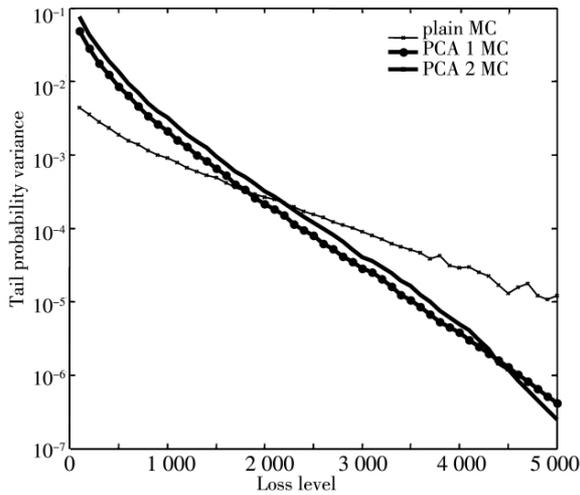


图5 10因子信用组合违约概率方差对比4

Fig. 5 The variance of the estimated default probability of credit portfolio with 10 factors (d)

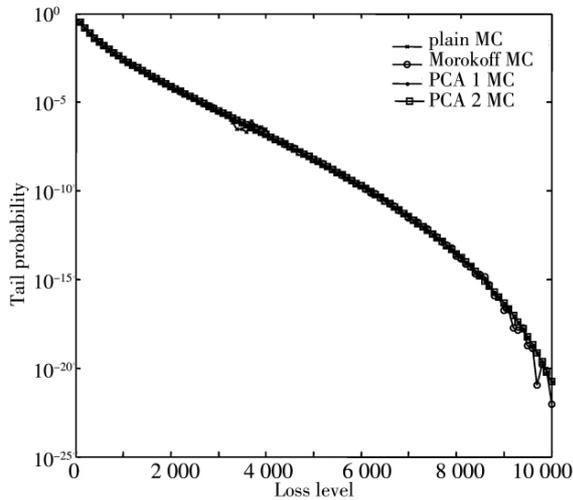


图6 50因子信用组合违约概率

Fig. 6 The default probability of credit portfolio with 50 factors

第二个算例是50因子信用组合模型,来自于Morokoff<sup>[16]</sup>的算例:组合资产个数为1000,它们的边际违约概率和违约损失分别服从式(24)和式(25).第1个载荷系数从均匀分布[0.21, 0.31]中随机取得.第2个到第5个载荷系数相互独立的从均匀分布[0.11, 0.21]中随机取得.最后两个非零载荷系数从剩余的系数中独立地挑选.如果选择的两个载荷系数是同一个,那么选择该载荷系数和相邻载荷系数,且服从[0, 0.1]的均匀分布.最后需要将这些载荷系数单位化,使其平方和为1.这个系数结构有如下解释:第1个因子代表了市场范围的公共风险因子,它作用于所有的信用;其它4个影响因子代表了产业和

地理因素的影响,剩余的两个影响因子表示了与公司相关的信用影响因素.

算例2的模拟结果显示在图6—图10中.根据设定,该信用组合最大可能的违约损失是11000,设定损失阈值为 $c = 10000$ 表示了非常极端的违约损失情况.表1给出的方差缩小因子是直接蒙特卡洛的样本方差与重要性抽样算法的样本方差的比值.就像本文所预期的那样,方差缩小因子随着违约损失的扩大而扩大.而本文重要性抽样算法所需要的计算时间与直接蒙特卡洛模拟基本相同,因此,这些方差缩小因子也是重要性抽样算法的效率扩大倍数.

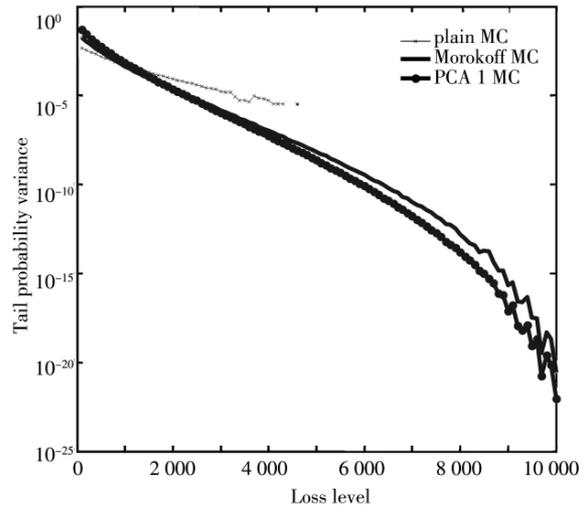


图7 50因子信用组合违约概率方差对比1

Fig. 7 The variance of the estimated default probability of credit portfolio with 50 factors (a)

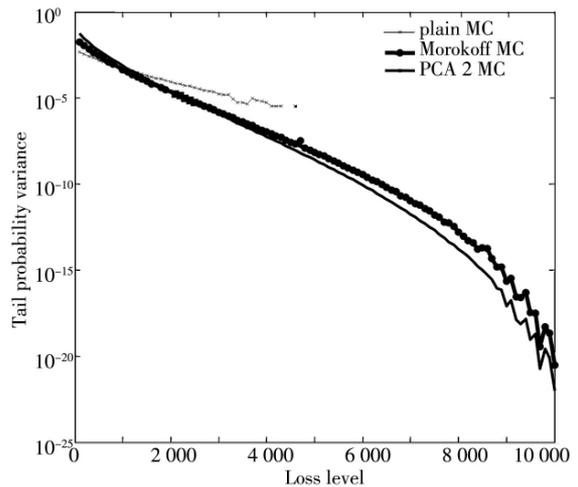


图8 50因子信用组合违约概率方差对比2

Fig. 8 The variance of the estimated default probability of credit portfolio with 50 factors (b)

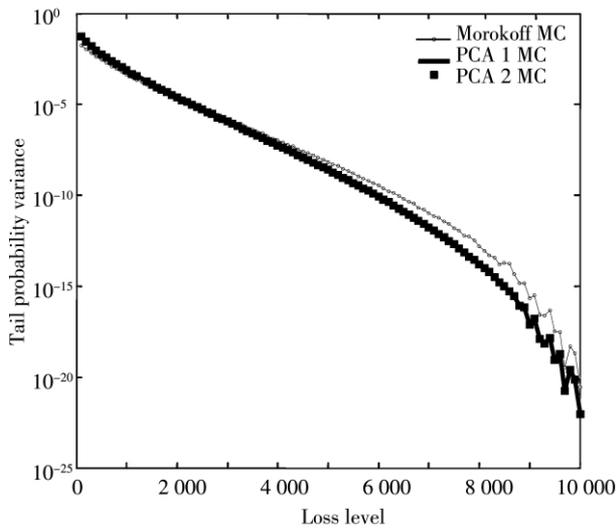


图 9 50 因子信用组合违约概率方差对比 3

Fig. 9 The variance of the estimated default probability of credit portfolio with 50 factors (c)

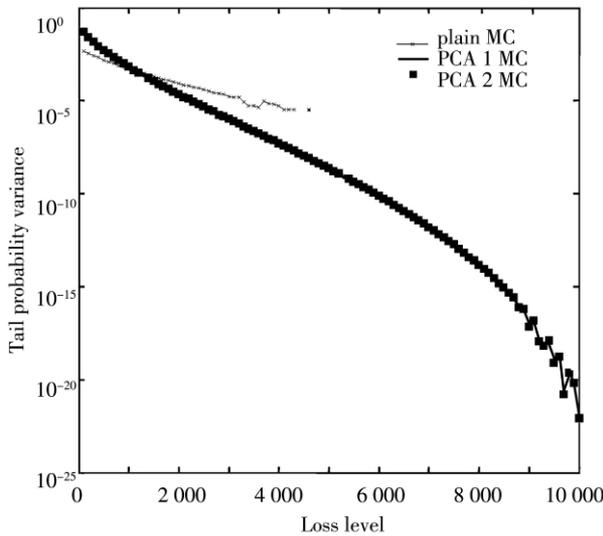


图 10 50 因子信用组合违约概率方差对比 4

Fig. 10 The variance of the estimated default probability of credit portfolio with 50 factors (d)

表 1 50 因子模型方差缩小值

Table 1 Variance reduction of the portfolio with 50 factors

损失水平	5 000	6 000	7 000	8 000	9 000	10 000
PCA 1 MC	2.80	4.52	5.97	9.92	27.08	30.96
PCA 2 MC	2.48	4.57	6.22	12.09	29.45	32.73

在第二个算例中,当违约损失  $c > 5\,000$  时,直接蒙特卡洛模拟方法给出的估计值将不再可靠.因此,本文将方差缩小因子的原始定义做出

了微小的调整:当违约损失超过 5 000 时,方差缩小因子使用 Morokoff 重要性抽样算法的样本方差作为分子,也就是以 Morokoff 重要性抽样算法作为比较基准.结果如表 1 所示.

从表 1 可以看出,本文提出的多成分变方差重要性抽样算法,在当组合违约损失大于 5 000 时,无论是 PCA 1 MC,还是 PCA 2 MC,都要优于 Morokoff 重要性抽样算法.而当违约损失越来越大时,PCA 2 MC 算法获得的方差因子比 PCA 1 MC 的要大.这进一步验证了本文所提出的多成分变方差重要性抽样算法在当大规模投资组合发生极值损失的情况时,相比较 Morokoff 重要性抽样算法和普通蒙特卡洛算法而言,在保证计算效率的前提下,达到了提高信用组合违约概率的度量精度的目的.

#### 4 结束语

本文针对商业银行信用组合违约概率度量效率较低的问题,采用了近年来学者们广泛使用的重要性抽样技术来构建模型.由于在多维空间中确定重要性区域设计到一系列难易处理的优化问题而导致计算效率不高,以及 Morokoff 重要性抽样算法在处理极端损失情况失效的缺点,本文基于 Morokoff 的算法,运用主成分分析和重要性抽样技术的基本原理,提出了多成分变方差重要性抽样算法.通过数值算例的结果可以发现,该算法通过适当地扩大占优成分的方差,在资产数较多的信用组合发生极端损失的情形下,能够相对精确有效的度量违约概率,提高违约概率的计算精度.

近年来 Copula 函数在金融领域的应用越来越广泛.针对金融数据所呈现的尖峰厚尾的统计特征,可考虑采用能够较好描述尾部相关性结构的学生 t Copula 函数来刻画投资组合内标的金融资产间的相关性结构,再融合本文提出的多成分重要性抽样算法,更进一步探讨商业银行信用风险的度量问题.有关这方面的研究将在以后工作中作进一步的深入展开.

## 参 考 文 献:

- [1] 谢尚宇, 汪寿阳, 周 勇. 金融危机下带传染效应的违约预报[J]. 管理科学学报, 2011, 14(1): 1-12.  
Xie Shangyu, Wang Shouyang, Zhou Yong. Default prediction with credit contagion under financial crisis[J]. Journal of Management Sciences in China, 2011, 14(1): 1-12. (in Chinese)
- [2] 张 维, 邱 勇. 多银行贷款池的组合违约风险研究[J]. 管理科学学报, 2008, 11(4): 134-141.  
Zhang Wei, Qiu Yong. Portfolio default risks of multi-bank loan pools[J]. Journal of Management Sciences in China, 2008, 11(4): 134-141. (in Chinese)
- [3] 龚 朴, 高 原. 非理性预期对信用衍生产品定价的影响——美国次贷危机的启示[J]. 管理科学学报, 2010, 13(9): 55-67.  
Gong Pu, Gao Yuan. Impact of irrational belief on credit derivatives pricing: Implication of subprime crisis[J]. Journal of Management Sciences in China, 2010, 13(9): 55-67. (in Chinese)
- [4] 龚 朴, 胡祖辉. 信用衍生产品隐含相关性结构研究[J]. 金融研究, 2011, 1: 182-194.  
Gong Pu, Hu Zuhui. The implied correlation structure of credit derivative[J]. Journal of Financial Research, 2011, 1: 182-194. (in Chinese)
- [5] Glasserman P, Li J Y. Importance sampling for portfolio credit risk[J]. Management Science, 2005, 51(11): 1643-1656.
- [6] Gupta G M, Finger C C. Credit Metrics Technical Document[M]. New York: Morgan J P & Co, 1997.
- [7] Glasserman P, Wanmo K. Large deviations in multifactor portfolio credit risk[J]. Mathematical Finance, 2007, 17: 345-379.
- [8] Glasserman P, Wanmo K. Fast simulation of multifactor portfolio credit risk[J]. Operations Research, 2008, 56: 1200-1217.
- [9] Glasserman P. Measuring marginal risk contributions in credit portfolios[J]. Journal of Computational Finance, 2006, 2(9): 1-41.
- [10] Dunkel J, Weber S. Efficient Monte Carlo Methods for Convex Risk Measures in Portfolio Credit Risk Models[M]. Washington DC: Winter Simulation Conference, 2007.
- [11] Bassamboo A, Juneja S. Portfolio credit risk with extremal dependence: Asymptotic analysis and efficient simulation[J]. Operations Research, 2008, 56: 593-606.
- [12] Huang X, Oosterlee C W. Adaptive integration for multi-factor portfolio credit loss models[J]. Journal of Computational and Applied Mathematics, 2009, 231(2): 506-516.
- [13] Grundke P. Importance sampling for integrated market and credit portfolio models[J]. European Journal of Operational Research, 2009, 194(1): 206-226.
- [14] Kang W, Lee K. Optimization problems in the simulation of multifactor portfolio credit risk[J]. Computational Science and Its Applications, 2006, 3: 777-784.
- [15] Sak H, Hörmann W, Leydole J. Efficient risk simulations for linear asset portfolios in the t-Copula model[J]. European Journal of Operational Research, 2010, 1: 802-815.
- [16] Morokoff W J. An Importance Sampling Method for Portfolios of Credit Risky Assets[M]. Washington DC: Winter Simulation Conference, 2004.
- [17] Chen Z Y, Glasserman P. Fast pricing of basket default swaps[J]. Operations Research, 2008, 56(2): 286-303.
- [18] Rubinstein R Y, Kroese D P. Simulation and the Monte Carlo Method, Seconded[M]. New York: John Wiley & Sons, 2007.

(下转第 22 页)

being that the successful probability of the farmer's project affects only the expected return of the bank and the other being that the successful probability of the farmer's project affects on both the expected return of the bank and the expected return of the farmer respectively, it proposes two farmer credit loan decision models and gives two different optimal loan interest rate mechanisms on the bank by taking both individual rationality of the farmer and loan risk losing ratio the bank can tolerate the most as constraint conditions. It also gives an example. Aiming at the farmer's different credit grades and he got the bank credit in the course of the five levels of classification, it designs 4 groups of different combinatorial data. It discusses the changes of both the optimal project successful probability of the farmer and the optimal loan interest rate of the bank based on the changes of the loan fund, the farmer's self-wealth, the loan risk losing ratio the bank can tolerate the most and the expected yield of the farmer. It also discusses the changes of both the net expected yield and the reasonable interval of the expected yield for the farmer project.

**Key words:** loan risk losing ratio of bank; farmer credit rating; individual rationality; successful probability of the project; credit loan decision model

---

(上接第 10 页)

## Principal component importance sampling for bank credit portfolio risk management

GONG Pu, DENG Yang, HU Zu-hui

School of Management, Huazhong University of Science & Technology, Wuhan 430074, China

**Abstract:** The bank credit portfolio risk measurement has great significance to bank supervision. One of the most popular methods to estimate the default probability of credit asset is Monte Carlo simulation. In order to improve the simulation efficiency, more and more studies have adopted the important sampling technique to deal with it. In this paper, we propose an importance sampling procedure which does not need the conditional independence which previous studies had to base on. The procedure we provide uses principal component analysis to choose dominant factors. Numerical experiments are provided and the results show that our approach when a credit portfolio confronts extreme events, offers substantial variance reduction and outperforms plain Monte Carlo algorithm and Morokoff IS algorithm.

**Key words:** portfolio credit risk; Monte Carlo simulation; important sampling; principle component analysis