

文章编号: 1002-1566(2013)01-0042-15

贝叶斯方法在信用风险度量中的应用研究综述

丁东洋¹ 周丽莉¹ 刘乐平²

(1. 南昌大学, 江西 南昌 330031; 2. 天津财经大学, 天津 300222)

摘要: 贝叶斯方法可以有效的处理信用风险度量中常见的数据缺失问题, 而且为科学使用专家意见等主观经验提供了有效途径, 已被广泛应用于信用风险度量领域。本文从模型构建、估计方法及模型比较三个方面对应用贝叶斯方法度量信用风险的重要文献进行综述, 重点关注信用风险的违约相关性和风险蔓延性等最新研究热点, 为深入研究信用风险度量问题提供参考, 并引起国内风险分析人员对贝叶斯方法的兴趣。

关键词: 贝叶斯方法; 信用风险度量; 违约依赖性

中图分类号: O212

文献标识码: A

Bayesian Methods in Credit Risk Measurement: A Survey

DING Dong-yang¹ ZHOU Li-li¹ LIU Le-ping²

(1. Nanchang University, Jiangxi Nanchang 330031, China;

2. Tianjin University of Finance and Economics, Tianjin 300222, China)

Abstract: Bayesian methods can effectively deal with the common problem of missing data, and provide a formal way for the scientific use of subjective experience, has been widely used in credit risk measurement. This paper presents a survey of important literature of credit risk measurement using Bayesian methods in three aspects of modeling, estimation methods and model comparison, and focus on the hot issues of default correlation and risk contagion of the latest credit risk research, in order to provide reference for further research in the credit risk measurement, and raise the domestic risk analyzer interest in Bayesian methods.

Key words: Bayesian methods, credit risk measurement, default dependence

0 引言

信用风险指金融市场上由交易方信用质量变动引发金融损失的可能性。在统计理论中一般采用信用损失分布 (Credit Loss Distribution) 描述损失的可能性, 而违约概率 (Probability of Default) 的估计是关键问题。债务人信用质量的快速波动, 不仅出现了信用违约, 而且市场风险急剧上升, 导致 2008 年的美国次贷危机最终引发全球金融海啸。随着巴塞尔协议的逐步

收稿日期: 2010 年 3 月 20 日;

收到最终修改稿日期: 2012 年 10 月 11 日

基金项目: 天津社科规划项目“宏观统计数据可靠性评估方法研究”(TJ TJ10-651) 资助; 全国统计科研计划项目“小域估计理论及其在我国统计调查中的应用”(2009LZ020) 资助。

实施, 国内金融机构如何提高信用风险管理水平, 在评级系统中应用有效的信用风险度量方法, 是当前值得关注的重要问题。

信用风险度量主要包括以下步骤: 一是明确影响信用风险的主要因素; 二是获取影响因素的数据信息及变量的动态特征; 三是构建和选择模型度量风险, 在此过程中, 概率论和统计方法的应用不可或缺。概率论从诞生至今已超过三个半世纪, 而关于“概率”理解的争论至今仍未停止。Neapolitan(1992)^[1]总结了概率的三种解释及应用方向: 古典概率论将概率解释为试验结果中部分事件相同的可能性, 能够处理“掷两个骰子得到点数之和最少为6的概率”之类的问题; 传统概率论来源于频率论学派的观点, 将概率应用于所有随机的、重复的事件, 从而能够处理“如果化肥没有效果, 农作物不会减产的可能性”等复杂问题; 贝叶斯学派进一步将概率应用于不确定性问题, 能够处理没有潜在随机过程的事件, 如“化肥没有效果的概率”。将概率论和统计方法应用于信用风险度量, 频率论和贝叶斯学派在模型设定和计算结果上都存在一定程度的差异, 这使得信用风险分析人员需要选择度量方法, 或是探索综合两种理论的风险度量技术。贝叶斯方法一般来说可分为标准贝叶斯方法和稳健贝叶斯方法, 不同方法内部存在一定程度的差异, 而本文主要描述能够应用于信用风险度量问题中的标准贝叶斯方法。

国外已有大量文献基于频率论学派的理论研究信用风险度量问题, 如 Lando and Skodeberg (2002)^[2]的综述涉及了较为全面的信用风险度量方法, 并给出了深刻的评价, Jafry and Schuermann(2004)^[3]不仅考察了违约概率的测度, 而且探讨了信用等级转移矩阵的选择和比较。从这些文献中, 可以发现在信用风险度量中的三个主要问题 (Scott, 2004)^[4]: 第一, 在变量选择和模型设定时都面临经验数据缺乏的问题; 第二, 无论是违约概率度量, 还是信用评级时都大量使用了专家意见或经验判断; 第三, 模型预测和比较都需要处理不确定性问题。贝叶斯方法能够科学的使用专家意见等主观信息, 在处理不确定性问题上具有一定的优势, 已较为广泛的应用于信用风险度量领域。本文以当前信用风险度量中的热点问题为主线, 从贝叶斯理论视角对信用风险度量中的模型构建、参数估计及模型比较等研究进展状况做重点的介绍, 以期引起国内风险分析人员对贝叶斯方法的兴趣, 并推动基于贝叶斯理论的信用风险度量的研究。值得说明的是本文并非对所有信用风险度量中的问题都提倡使用贝叶斯方法, 而是在需要及合适的情况下将贝叶斯方法与传统方法更好的结合, 以期更加有效准确的度量信用风险。

1 基于贝叶斯方法的信用风险度量模型

1909年, 著名评级公司穆迪 (Moody) 的创始人 John Moody 在《铁路投资分析年刊》中首次提出信用评级 (Credit Rating) 的概念, 至今信用风险度量已走过 100 年的历程。结构模型 (Structure Model) 和强度模型 (Intensity Model) 是传统信用风险度量模型的两种主要形式, 前者利用资产负债的变量关系描述违约, 后者利用特征参数强度描述违约事件。贝叶斯信用风险度量模型与传统模型并不相同, 在模型中既可以考虑资产价值的变动也可以涉及违约强度变量, 具有更强的适应性和灵活性 (Philosophov, 2011)^[5]。贝叶斯信用风险度量模型的分类目前仍无确定评述, 下面以一些重要文献中常见的几类模型为标准, 重点研究一般动态模型、潜因素模型和广义线性混合模型。

1.1 应用贝叶斯方法度量信用风险的早期思想

Qin(1996)^[6]在“Bayesian Econometrics: The First Twenty Years”的文中指出: “60 年代初伴随 Fisher、Hildreth、Zellner 以及 Tiao 等人研究的问世 (如 Fisher 的“Estimation in the Linear Decision Mode”, Zellner 的“Bayesian Analysis of the Regression Model with Autocorrelated

Errors”等), 贝叶斯方法处理经济管理类问题的优点逐渐显现”。哈罗德·比尔曼 (Harold Bierman, JR.) 和沃伦·豪斯曼 (Warren H. Hausman) 于 1970 年在《Management Science》上发表的“The Credit Granting Decision”一文采用贝叶斯方法构建模型估计授信企业的信用质量, 是将贝叶斯方法引入信用风险度量中的开创性文献^[7]。在此之前线性判别模型 (Linear Discriminant Model, LDM) 是度量信用风险的主要方法, 通过赋予贷款申请人不同的指标权重综合评分判断信用等级。性判别模型中的重要假定是不同等级的方差协方差矩阵相同, 且评分数据服从正态分布, 但这些假定过于严格 (Majone, 1968)^[8]。也有学者提出利用马尔科夫模型描述信用等级转移 (Sandor and Wedel 2001)^[9], 但前提是转移概率矩阵保持不变。哈罗德和沃伦教授提出借助贝叶斯方法放松以上假设, 包括使用债务人的盈利可能等先验信息, 通过更新技术修正后验概率预测信用风险。虽然他们只考虑可以授信和不能授信两种信用质量, 但模型已包含动态信息, 为后期贝叶斯多期动态模型的发展奠定了基础 (Montgomery and Rossi 1999)^[10]。

虽然采用贝叶斯方法度量信用风险的思想为人们所接受, 但直到 20 世纪 80 年代中期, 由于似然函数与后验分布的计算难题, 贝叶斯方法的应用进展缓慢。实际上如同“高收益高风险”的道理相同, 想要不花费计算成本就获得贝叶斯方法的优势也是不现实的。不过随着计算机技术和模拟方法的发展, 当前广泛使用的马尔科夫链蒙特卡罗 (MCMC) 算法在一定程度上缓解了贝叶斯的计算难题, 从而也推动了贝叶斯方法更广泛的应用于风险度量领域 (Wedel and Pieters, 2000)^[11]。至今 MCMC 方法已经可以应用于多种形式的模型, 使得在贝叶斯框架下采用更多类型的似然函数和先验分布成为可能 (Rossi and Allenby, 2002)^[12]。

1.2 应用贝叶斯动态模型解释违约相关性

早期的信用风险度量主要是考察特定公司的违约风险, 随着金融市场的发展, 大量衍生品和贷款组合出现, 信用风险度量更多面对的是组合资产的联合违约风险。资产组合的违约风险存在依赖性 (Default Dependence), 不能再单独的考察债务人自身违约概率。依赖性中表现出来的违约风险与时间变量相关, 称为违约相关性 (Correlation of Default), 这表明资产组合违约风险与整体经济环境密切相关。此外由于债务人之间的商业活动, 依赖性还表现出了风险蔓延 (Contagion of Risk) 特征, 即关联客户违约将导致债务人本身的信用风险增加。违约相关性和蔓延性会导致一定时期内出现大量不均衡的风险, 是当今金融市场上值得关注的重要问题。贝叶斯动态模型 (Bayesian Dynamic Model) 可以同时包含可观测变量和未知状态, 适于描述随时间不断变化的系统, 这与资产组合风险和信用等级随时间的变化等问题相一致。

对于信用风险的动态变化问题研究从 20 世纪 90 年代中期开始迅速增加, 人们首先认识到的是信用等级或违约概率受当期及滞后变量的共同影响, 即等级转移白相关问题。向量白回归模型 (Vector Auto-Regressive Model, VAR) 是解决白相关问题的常见模型。VAR 最早由 Sims(1972)^[13] 提出, 而后期的分析人员往往受限于样本数据, 同时要避免对滞后向量限定过多出现过度拟合 (Overfitting) 问题。Litterman(1980)^[14] 首次将贝叶斯方法引入 VAR 模型解决过度拟合问题, 避免了对模型中的某些相关系数默认为 0。在信用风险度量方面较有代表性的方法是采用信用等级转移矩阵评估信用风险并预测等级变化, 该建模方法最早由 Jarrow et al.(1997)^[15] 提出, Kijima(1998)^[16] 对其进行了拓展。不过前期使用的都是固定转移矩阵, Lando(1998)^[17] 和 Arvanitis et al.(1999)^[18] 等提出使用随机转移矩阵更加确切, Huge and Lando(1999)^[19] 将这种方法应用于掉期定价研究。Lando and Skodeberg(2002)^[2] 给出了应用连续时间变量估计转移概率的方法, Sanjiv et al.(2002)^[20] 则将贝叶斯方法应用于连续时间状态下信用等级转移的估计。

Wilson(1997)^[21] 针对违约相关性, 将宏观经济环境与违约概率的变化联系在一起, 运

用 Logit 模型给出回归结果。Nickell et al.(2000)^[22] 及 Bangia et al.(2002)^[23] 都认为违约概率和信用等级会随着 GDP 增长速度的变化而变化, 在经济衰退时期违约概率和等级下降强度都会上升。Kavvathas(2001)^[24] 及 Carling et al.(2002)^[25] 等基于观测的宏观变量采用久期方法建模, 结果发现违约时距 (Time to Default) 会在经济萧条时期缩短。Koopman and Lucas(2005)^[26] 利用时间序列数据证明了违约概率在长期内存在周期性。Fledelius et al.(2004)^[27] 则针对信用等级转移概率与宏观经济环境的关系进行分析, 明确指出资产组合信用风险存在时间波动性。以上这些文章的建模思路基本都与 Wilson(1997)^[21] 思想一致 (Rao and Wu, 2010)^[28]。

在考察经济环境对信用风险的影响时, 动态模型可以描述违约相关的两个重要特征: 波动性和非对称性, 前者指信用风险大小随经济环境出现周期性变化, 后者指的是信用质量时间序列的特征依赖于经济周期的状态。早期模型中的经济周期变量往往采用二值型 - 正增长和负增长。在随后的研究中模型形式和推断方法都有所拓展, 如 Kim(1994)^[29] 将马尔科夫开关 (Markov Switching) 引入动态线性模型; 接着 Kim and Nelson(1999)^[30] 基于贝叶斯方法和 MCMC 模拟技术处理广义马尔科夫开关模型; Kim and Murray(2002)^[31] 将经济周期分为三阶段: 经济萧条、正常增长及高速增长。另外一些学者考察不同经济周期阶段久期 (Duration) 对信用风险的影响, Sichel(1991)^[32] 和 Watson(1994)^[33] 等提出转移概率依赖于当前周期阶段的久期, Diebold and Rudebusch(1996)^[34] 在此基础上拓展了模型的多变量结构, 这就是信用风险度量中著名的信用等级转移久期模型, 也称为转移强度 (Transition Intensity) 模型 (Lando and Skodeberg, 2002)^[2]。

令解释变量 $x_t \in X$ 表示可观测的债务人特定因子和宏观经济变量, 被解释变量 $y_t \in Y$ 为风险度量指标, $\theta \in \Theta$ 为参向量。给定先验分布 $p(x_0|\theta)$, 观测密度 $p(y_t|x_t, y_{1:t-1}, \theta)$ 和转移密度 $p(x_t|x_{0:t-1}, y_{1:t-1}, \theta)$, $t = 1, 2, \dots, T$, 广义贝叶斯动态模型的一般形式为:

$$y_t \sim p(y_t|x_t, y_{1:t-1}, \theta), \quad (1)$$

$$x_t \sim p(x_t|x_{0:t-1}, y_{1:t-1}, \theta), \quad (2)$$

$$x_0 \sim p(x_0|\theta). \quad (3)$$

这种模型形式还可以处理变量的非高斯和非线性问题。 $x_{0:t} \triangleq (x_0, x_1, \dots, x_t)$ 表示到时间 t 为止, 向量 x 收集到的样本信息, 其中 $x_{-t} \triangleq (x_0, \dots, x_{t-1}, x_{t+1}, \dots, x_T)$ 表示剔除第 t 个元素信息。如果转移密度仅依赖于滞后一期的状态向量, 就变成常见的一阶马尔科夫模型:

$$(y_t|x_t) \sim p(y_t|x_t, y_{1:t-1}, \theta), \quad (4)$$

$$(x_t|x_{t-1}) \sim p(x_t|x_{t-1}, y_{1:t-1}, \theta), \quad (5)$$

$$x_0 \sim p(x_0|\theta). \quad (6)$$

这也是在大量文献中常用的假定, 即等级转移的马尔科夫性。在贝叶斯框架下, 即使遇到 p 阶马尔科夫模型, 也不必事先给定模型中某个变量的权重, 模型设定时仍然可以采用一阶形式, 参数分布的不确定性可以由最新获得的数据信息进行解释更新。Litterman(1986)^[35] 详细总结了含有宏观协变量时先验分布的设定要求: 一是体现宏观时间序列的主要趋势; 二是尽量采用近期的数据解释对未来的影响; 三是关键变量的滞后值应比其他变量的滞后值更能反映被解释变量的当前状态和特征。

1.3 应用潜因素模型描述宏观经济环境变化

在一般贝叶斯动态模型中, 有很多研究人员采用可观测的宏观经济变量解释违约概率相关性, 但是由于宏观环境对违约概率的影响存在滞后性, 选择宏观经济变量需要构建辅助预测模

型,建模过程进一步复杂化。同时实证分析的经验结果也表明仅用可观测宏观变量解释信用风险变动的效果并不理想,因而有的学者认为在模型中引入宏观经济变量不如退一步从债务人自身数据直接判定违约风险,这样可以避免错误设定宏观变量导致出现风险评估偏差, Couderc and Renault(2004)^[36] 也认为直接从债务人数据度量信用风险有利于形成完整的风险度量框架。针对传统方法中的这些问题,贝叶斯学派的一些研究人员 (Gagliardini and Gouriou, 2004^[37]; McNeil and Wendin, 2004^[38]; Koopman et al., 2005^[39]) 提出采用潜因素 (Latent Factor) 与可观测变量一起描述宏观经济环境变化。潜因素实际上就是无法观测变量,出于各种目的构建的潜因素模型基本上都具有分层 (Hierarchy) 或错层 (Layering) 的特征 (Geweke, 2005)^[40], 也就是借助假定的参数决定潜因素的分布,进而综合潜因素以及潜因素与其他变量组合的协同影响描述被解释变量的分布。这种思想形成了贝叶斯框架下的分层模型 (Hierarchical Model), 日益广泛的应用于复杂数据问题的分析。分层模型的构建首先需要获得不同级别的共变 (Covariation) 信息, 分层能够显著简化后验分布,模型的推断可以采用吉布斯抽样等计算方法。

传统方法也将宏观经济变量和潜因素联合使用度量信用风险,近期的文献如 Dungey et al.(2000)^[41] 和 Diebold et al.(2006)^[42] 等在解释不同债券收益率的相关性时都运用了潜因素,而 Schuermann and Stroh(2006)^[43] 更是利用潜因素对股票收益率构建了动态模型。Bauwens and Hautsch(2003)^[44] 进一步拓展潜因素的应用,首次将潜因素强度 (Latent Factor Intensity, LFI) 引入股票交易点过程的风险度量模型。近些年,贝叶斯学派开始考察 LFI 模型的贝叶斯推断,如 Koopman et al.(2008)^[45] 将久期拓展为不仅包含等级转移发生所需时间而且涵盖即将发生事件 (等级的未来变动) 所需时间,模型中采用更加复杂的似然函数,运用多元变量扩展的蒙特卡罗技术 (Durbin and Koopman 1997, 2001)^[46,47] 进行计算推断,与传统的条件久期白回归模型 (Engle and Russell, 1998)^[48] 和条件强度白回归模型 (Russell, 1999)^[49] 具有本质区别,预测效果表现良好。

潜因素除了用于描述经济环境之外,还有很多研究人员建议采用潜因素解释债务人异质性 (Heterogeneity)。单个债务人的行为组合构成资产组合的信用风险变动特征,因而资产组合的风险度量需要整合个体的异质性 (Bradlow and Schmittlein, 1999)^[50]。在信用风险度量中,单个债务人的具体信息很多时候难以获得,传统的分布假设预测并不准确,而潜因素模型可以在一定程度上解决这个难题。潜因素通常采用连续变量,即使模型中的样本观测值是离散形式。描述债务人异质性的潜因素模型不仅可以用于风险度量,而且可以进一步拓展为信用衍生品定价模型 (Schönbucher, 2003^[51]; McNeil and Wendin, 2007^[52])。潜因素模型的一般形式以多元回归模型为基础:

$$Y = X\beta + \varepsilon, \quad \varepsilon \sim N(0, h^{-1}), \quad (7)$$

其中潜变量 Y 表示债务人信用质量, ε 为漂移项,表示可观测的宏观协变量的影响,解释变量 X 包含无法观测得到的协变量,用潜因素 \tilde{z} 替换上述模型中的 X 可得:

$$Y = \tilde{z}\beta + \varepsilon, \quad \tilde{z} \sim N(\underline{X}, \underline{H}^{-1}), \quad (8)$$

其中 h 表示精度,是方差 σ^2 的倒数, $\underline{X}, \underline{H}^{-1}$ 中的下划线表示均值和方差的先验设定。先验分布中对 X 的设定使得均值存在很大的不确定性,但相对于不确定性变量的均值将趋同,即 $E(X) = \underline{X}$, $\text{var}(X) = \underline{H}^{-1}$ 。潜因素模型的不同拓展会形成不同的模型。如有的分析人员 (Gagliardini and Gouriou, 2004)^[37] 将可观测的宏观协变量单独列入模型:

$$Y = \tilde{z}\beta + M\alpha + \varepsilon. \quad (9)$$

这是混合模型 (Mixture Model) 的一种形式,也称为潜因素混合模型,即模型中同时包含可观测变量和未知变量的影响,其中 $M\alpha = \sum_{i=1}^n \alpha_i m_i$, m_i 表示某宏观经济因素,如 GDP 增长率,

α_i 是对应的系数。有的分析人员 (McNeil and Wendin, 2004)^[38] 将信用质量的滞后值加入模型, 即:

$$Y_t = Y_{t-1} + \tilde{z}_t \beta + \varepsilon_t. \quad (10)$$

这是包含向量滞后值的潜因素模型, 其中 Y_{t-1} 表示前一次信用评估的结果, 实际上可以根据分析情况延长至第 i 期滞后值。

可以说, 潜因素模型的不断扩展和完善有效处理了信用风险相关性问题的, 但是信用风险除了受经济环境影响之外, 还受债务双方的商业活动影响, 也就是信用风险蔓延 (Egloff et al., 2004)^[53]。同时涵盖宏观经济环境的可观测变量和潜因素以及向量滞后值的模型就是广义线性混合模型, 可以用于描述包含风险蔓延在内更为广泛的违约依赖性。

1.4 应用广义线性混合模型处理风险蔓延性

广义线性混合模型 (Generalized Linear Mixed Model, GLMM) 用途极为广泛, 可以用于处理面板数据及空间数据的多重问题。信用风险分析中常用的 CreditRisk⁺ (Credit Suisse First Boston, 1997)^[54] 模型就属于 GLMM, 其基本原理假定违约变量服从泊松分布而计算违约概率。将债务人信用质量记作状态空间 $K = \{1, 2, \dots, k\}$, 如 $k = 8$ 即为常见的信用等级: AAA、AA、A、BBB、BB、B、CCC、D。令 $\gamma_{i,1} > \gamma_{i,2} > \dots > \gamma_{i,k-1} > \gamma_{i,k} = 0$ ($i \in \{1, 2, \dots, k-1\}$) 为标准阈值, 区间长度可以有所差异, 相同区间内债务人的违约概率相近。在时间 t 处于等级 i 的债务人 $N_i(t)$, 如果在时间 $t+1$ 代表债务人信用质量的潜变量变化至区间 $[\gamma_{i,l-1}, \gamma_{i,l}]$ 内, 则债务人信用等级从 i 转移到 l ; 如果潜变量小于 $\gamma_{i,k}$, 则债务人发生违约。因而 GLMM 中阈值的应用除了可以估计违约概率之外, 还可用于估计信用等级转移概率。

(1) 贝努利混合模型

在式 (9) 中设定合适的 \tilde{z} 和 M , 债务人违约的条件概率可表示为:

$$P(Y_j < \gamma | \tilde{z}, M) = f(\gamma - \varepsilon - \tilde{z}\beta - M\alpha), \quad (11)$$

其中 $f(\cdot)$ 为依赖于 \tilde{z} 分布的关联函数 (Stefanescu et al., 2009)^[55]。例如, \tilde{z} 具有极值分布时的 $f(\cdot)$ 为 logit 关联函数, 形式为 $f(x) = 1/(1 + e^{-x})$ 。 \tilde{z} 为正态分布时的 $f(\cdot)$ 为 probit 关联函数, 形式为 $f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$ 。从而债务人发生违约的联合概率分布为:

$$P(Y < \gamma | \tilde{z}, M) = \prod_{j=1}^{n_t} P(Y_j < \gamma | \tilde{z}, M)^{k_j} [1 - P(Y < \gamma | \tilde{z}, M)]^{1-k_j}, \quad (12)$$

其中 $k_j \in \{0, 1\}$, 1 表示发生违约, 0 表示未违约。式 (12) 具有标准的贝努利分布形式, 称为贝努利混合模型, 实际上著名的 CreditMetrics (KMV, 1997)^[56] 也是具有高斯随机效应的混合模型。向量 \tilde{z} 和 M 中既可以包含可观测的协变量, 也可以用虚拟变量指示不同的企业或行业。基于该模型可以计算不同债务人之间的风险蔓延性, 即 $cov(Y_i, Y_j)$, probit 关联函数和 logit 关联函数的计算有所差别 (参见 McCullagh and Nelder, 1989)^[57]。模型 (12) 在应用时的难点在于时间区间数量增多时的参数估计, 估计量的准确性和显著性难以满足 (Frey and McNeil, 2003)^[58]。

(2) 广义线性混合模型 (GLMM)

实际上贝努利混合模型、Logit 回归模型以及对数线性回归模型等都是 GLMM 的特殊形式或简化形式。在采用无信息先验分布时, GLMM 的后验分布可表示为:

$$P(p, \alpha, \beta | N, n) = \frac{\prod_{t=1}^T P(n_t | N_t, p, \alpha, \beta) p(p) p(\alpha) p(\beta)}{P(n | N)}, \quad (13)$$

其中信用等级 i 包含 N 个债务人, 发生违约的数量用 n 表示, p 为特定债务人的违约概率。实际上式 (12) 为如果不设定为贝努利分布, 即为 GLMM 的一般形式, 其中 M 称为固定效应 (Fixed Effect), \tilde{z} 称为随机效应 (Random Effect), 影响系数分别为 α 和 β , 二者共同构成外部风险 (Gelman, A. et.al., 2003)^[59]。随机效应 \tilde{z} 的表现形式可以是随机的截距或影响系数 (Skrondal and Rabe, 2004)^[60]。固定效应 M 可以包含所有债务人的共同影响因子, 也可含有影响特定债务人的因子。共同影响因子能够说明违约概率的非时齐性 (Koopman et al., 2005)^[39], 影响特定债务人的因子能够说明不同评分的债务人处于同一信用等级的异质性问题 (Kadam and Lenk, 2008)^[61]。此外, 影响特定债务人的因子可以是在建模初期就使用的解释变量, 也可以包括在得出结论前刚认识的解释变量, 对于新认知变量的描述需要构建辅助模型。这有别于传统固定分析历史数据的方法, 提升了预测能力。此外, 随机效应的形式既可以是一元也可以拓展为多元形式, 多元随机效应有助于建立分层模型, 充分的体现了模型的灵活性和适应性。

2 贝叶斯信用风险度量模型推断方法

2.1 过滤、预测和平滑

贝叶斯信用风险度量模型往往比较复杂, 模型的推断主要应用三个技术, 即过滤 (Filtering)、平滑 (Smoothing)、预测 (Prediction)。此外, Kalman(1960)^[62] 和 Kalman and Bucy(1961)^[63] 在高斯线性动态系统的分析中采用了卡尔曼滤波技术 (Kalman Filter), Koopman et al.(2008)^[45] 则将滤波用于贝叶斯信用风险度量模型的推断。

信用风险度量中的状态指债务人所处的信用等级或是否为违约, 密度函数 $p(x_t|y_{1:s}, \theta)$ 的估计值可以判断状态。在密度函数 $p(x_t|y_{1:s}, \theta)$ 中, 当 $t = s$ 时叫做过滤密度, 当 $t < s$ 时叫做平滑密度, 当 $t > s$ 时叫做预测密度, 特别的当 $t = 1$ 时, 密度 $p(x_0|y_0, \theta) = p(x_0|\theta)$ 就是动态模型的先验分布。假定在时间 t 已知密度 $p(x_{t-1}|y_{1:t-1}, \theta)$, 应用查普曼 - 科尔莫戈罗夫 (Chapman-Kolmogorov) 等式, 即

$$P\{X_u = z|X_s = x\} = \sum_{y \in E} P\{X_t = y|X_s = x\}P\{X_u = z|X_t = y\} = \sum_{y \in E} P_{s,t}(x, y)P_{t,u}(y, z) \\ \Leftrightarrow P_{s,u}(x, z) = P_{s,t}P_{t,u},$$

可得下一时点的预测密度为:

$$p(x_t|y_{1:t-1}, \theta) = \int_X p(x_t|x_{t-1}, y_{1:t-1}, \theta)p(x_{t-1}|y_{1:t-1}, \theta)dx_{t-1}. \quad (14)$$

当获得新的样本观测值 y_t , 可以更新预测密度, 从而过滤密度表示为:

$$p(x_t|y_{1:t}, \theta) = \frac{p(y_t, x_t|y_{1:t-1}, \theta)}{p(y_t|y_{1:t-1}, \theta)} = \frac{p(y_t|x_t, y_{1:t-1}, \theta)p(x_t|y_{1:t-1}, \theta)}{\int_X p(y_t|x_t, y_{1:t-1}, \theta)p(x_t|y_{1:t-1}, \theta)dx_t}, \quad (15)$$

其中 $p(x_t|y_{1:t-1}, \theta)$ 是从式 (14) 得到的预测密度, 在每一个时点 t 都可以推断状态向量位于 K 个时段后的预测密度, 采用迭代方法:

$$1 \text{ 个时段: } p(x_{t+1}|y_{1:t}, \theta) = \int_X p(x_{t+1}|x_t, y_{1:t}, \theta)p(x_t|y_{1:t}, \theta)dx_t, \quad (16)$$

$$K \text{ 个时段: } p(x_{t+k}|y_{1:t}, \theta) = \int_X p(x_{t+k}|x_{t+k-1}, y_{1:t}, \theta)p(x_{t+k-1}|y_{1:t}, \theta)dx_{t+k-1}. \quad (17)$$

进而

$$p(x_{t+k}|x_{t+k-1}, y_{1:t}, \theta) = \int_{Y^{k-1}} p(x_{t+k}|x_{t+k-1}, y_{1:t+k-1}, \theta)p(dy_{t+1:t+k-1}|y_{1:t}, \theta), \quad (18)$$

其中 $Y^k = \bigotimes_{i=1}^k Y_i$ 表示状态空间的 k 次笛卡尔乘积, $k = 1, 2, \dots, K$ 。在此基础上, 位于 k 个时段之后的条件预测密度:

$$p(y_{t+K}|y_{1:t}, \theta) = \int_Y p(y_{t+K}|x_{t+K}, y_{1:t+K-1}, \theta) p(dy_{t+1:t+K-1}|y_{1:t}, \theta) p(dx_{t+K}|y_{1:t}, \theta). \quad (19)$$

在广义动态模型中, 上述推断过程需要大量的复杂积分, 而 MCMC 算法或粒子滤波能够解决积分难题。实际风险分析人员还经常对变量的动态特征做出一些简化的假定, 如假定变量动态特征的变化独立于滞后值 $y_{1:t}$, 这样上述模型 (4)-(6) 可以进一步简化为:

$$(y_t|x_t) \sim p(y_t|x_t, \theta), \quad (20)$$

$$(x_t|x_{t-1}) \sim p(x_t|x_{t-1}, \theta), \quad (21)$$

$$x_0 \sim p(x_0|\theta). \quad (22)$$

相应的过滤和预测密度可简化为:

$$p(x_t|y_{1:t-1}, \theta) = \int_X p(x_t|x_{t-1}, \theta) p(x_{t-1}|y_{1:t-1}, \theta) dx_{t-1}, \quad (23)$$

$$p(x_t|y_{1:t}, \theta) = \frac{p(y_t|x_t, \theta) p(x_t|y_{1:t-1}, \theta)}{p(y_t|y_{0:t-1}, \theta)}, \quad (24)$$

$$p(x_{t+K}|y_{1:t}, \theta) = \int_X p(x_{t+K}|x_{t+K-1}, \theta) p(x_{t+K-1}|y_{1:t}, \theta) dx_{t+K-1}, \quad (25)$$

$$p(y_{t+K}|y_{1:t}, \theta) = \int_X p(y_{t+K}|x_{t+K}, \theta) p(x_{t+K}|y_{1:t}, \theta) dx_{t+K}. \quad (26)$$

上述推导过程中包含两个非常重要的递归关系, 第一个就是顺序滤波方程:

$$p(x_{0:T}|y_{1:T}, \theta) = p(x_{0:T-1}|y_{1:T-1}, \theta) = \frac{p(y_T|x_T, \theta) p(x_T|x_{T-1}, \theta)}{p(y_T|y_{1:T-1}, \theta)}. \quad (27)$$

该方程主要用于处理顺序数据, 同时也是粒子滤波 (Doucet et al., 2001)^[64] 方法的基础。第二个递归关系为状态向量平滑密度的因素分解形式:

$$p(x_{0:T}|y_{1:T}, \theta) = p(x_T|y_{1:T}, \theta) \prod_{t=0}^{T-1} p(x_t|x_{t+1}, y_{1:t}, \theta), \quad (28)$$

$p(x_t|x_{t+1}, y_{1:t}, \theta) = p(x_{t+1}|x_t, y_{1:t}, \theta) p(x_t|y_{1:t}, \theta) / p(x_{t+1}|y_{1:t}, \theta)$, 即分解为过滤密度和预测密度的运算。

在处理更加复杂的信用风险分析问题时, 一些模型的推断应用简单的过滤、平滑和预测技术无法解决, 如常见的条件高斯线性模型:

$$y_t = F(s_t)x_t + V(s_t)\varepsilon_t, \quad \varepsilon_t \sim N(0, I); \quad x_{t+1} = G(s_t)x_t + W(s_t)\eta_t, \quad \eta_t \sim N(0, I). \quad (29)$$

Harrison and West(1997)^[65] 将上述模型称作多进程模型, 式 (29) 中 s_t 是顺序随机变量, ε_t 与 η_t 相互独立, 包括 $s_t = s_{t-1} = s$ 的普通形式和 s_t 为随机过程的特殊形式。特别当 s_t 为离散时间变量, 在给定转移密度有限状态马尔科夫链的前提下, 该模型就是风险分析中常见的跳跃马尔科夫线性模型, 或称为马尔科夫开关线性模型。当采用马尔科夫开关隐藏进程描述向量状态时, 即马尔科夫开关潜因素模型:

$$y_t = \alpha x_t + \sigma_\varepsilon \varepsilon_t, \quad \varepsilon_t \sim N(0, 1), \quad (30)$$

$$x_{t+1} = \mu(s_{t+1}) + \rho x_t + \sigma_\eta \eta_{t+1}, \quad \eta_{t+1} \sim N(0, 1), \quad (31)$$

$$s_t \sim \text{Markov}(P), \quad s_t \in \{0, 1\}, \quad (32)$$

其中 P 是转移矩阵, $\mu(s_t) = \mu + \nu s_t$ 。这种形式的模型缺少可解析的过滤密度, 采用上述简化方法无法推断, 此时 MCMC 等模拟技术是常用的计算方法。

2.2 后验分布模拟

贝叶斯信用风险度量模型的计算难题之一就是后验分布的多种积分。如果后验均值服从独立同分布, 那么可简化计算如下:

$$I = E_{\theta|y}[b(\theta)] = \int b(\theta)p(\theta|y)d\theta; \quad \hat{I} = \frac{1}{R} \sum_{r=1}^R b(\theta_r). \quad (33)$$

简化计算仍可以得到任何精度下参数 θ 的估计值。除利用独立同分布简化推断外, 还可以利用马尔科夫链计算转移密度生成含有 θ 的序列, 再进一步计算推断 (Gelfand and Smith 1990)^[66]。如 $p(\theta|y)$ 是马尔科夫链的静态分布, 可以通过递归运算消除先验分布的影响而推断 \hat{I} 。当模拟数据出现自相关等问题时, 将后验积分表示为参数的期望分布函数, 利用模拟技术可以估算不同精度的 I , 这也是 MCMC 方法的基本思想。

构建马尔科夫链的常见方法是 Metropolis-Hastings 算法, 在一定条件下, 该算法可使后验分布成几何速率收敛 (Tierney, 1994)^[67]。Metropolis-Hastings 算法一种常见的形式是吉布斯抽样, 为了推出 $\theta_r|\theta_{r-1}$ 循环使用条件分布

$$p_1(\theta_{r,1}|\theta_{r-1,2}, \theta_{r-1,3}, \dots, \theta_{r-1,K}, y), \dots, p_K(\theta_{r,K}|\theta_{r,1}, \theta_{r,2}, \dots, \theta_{r,K-1}, y),$$

使其收敛于模型参数的联合后验分布 $p(\theta_1, \dots, \theta_K|y)$, 这样就可以估计模型中的特定参数:

$$p(\theta_i|y) = \int p(\theta_1, \dots, \theta_K|y)d\theta_{-i}. \quad (34)$$

式 (34) 这种闭合的条件分布形式在信用风险度量中十分常见, 且分布多为熟知的形式, 如泊松分布、贝塔分布、伽马分布等。不仅线性回归模型能够使用吉布斯抽样方法, 即使在数据增广算法 (Data Augmentation) 中, 也能够利用吉布斯抽样建立 probit 模型。随着模拟技术的不断发展, Gilks(1992)^[68] 等提出了舍选抽样法 (Adaptive Rejection Sampling), 对参变量的每个元素顺序抽样, 在理论上可以用于任何模型。当前的模拟技术已经可以处理非标准的似然函数和先验分布族, 一旦模拟结果存在高度自相关, 可以将参向量的元素分组, 通过自定义 Metropolis 算法继续推断。

3 贝叶斯信用风险度量模型的比较和选择

当信用风险度量模型的可能形式不确定时, 需要进行模型比较和选择。除了进行模型变量的选择外, 还涉及混合模型中参数成份的设定以及参数族的选择, 相应结论不仅可用来选择最优模型, 而且可以用来改善推断结论, 如模型平均法 (Congdon, 2001)^[69]。

在贝叶斯框架下, 模型比较方法主要两种 (Alston et al., 2005)^[70], 即隔离估计 (Separate Estimation) 和联合估计 (Simultaneous Estimation)。隔离估计方法包括单独计算每个模型的后验预测分布 (Posterior Predictive Distributions)、贝叶斯因子 (Bayes Factors)、贝叶斯信息准则 (BIC) 以及离差信息准则 (DIC) 等。联合估计方法主要涉及可逆跳跃 (Reversible Jump) MCMC 模拟和生灭进程 (Birth and Death Processes) 等方法。

3.1 隔离估计

信用风险度量建模的目的是预测未来的信用风险, 在贝叶斯框架下很自然的想到对模型的后验预测分布进行比较。令 y 的边际分布为:

$$p(y) = \int p(y, \theta) d\theta = \int p(\theta) p(y|\theta) d\theta. \quad (35)$$

式 (35) 在也称为先验预测分布 (Prior Predictive Distribution), 不以前期观测到的数据为条件。当观测得到数据 y 后, 可进一步预测后期数据 \tilde{y} 。 $y = (y_1, y_2, \dots, y_n)$ 为从时间 1 到 n 的观测向量值, $\theta = (\mu, \sigma^2)$ 是未知参数的期望和方差, \tilde{y} 为未观测到的数据, 其分布为后验预测分布 (Posterior Predictive Distribution):

$$p(\tilde{y}|y) = \int p(\tilde{y}, \theta|y) d\theta = \int p(\tilde{y}|\theta, y) p(\theta|y) d\theta = \int p(\tilde{y}|\theta) p(\theta|y) d\theta. \quad (36)$$

可见, 后验预测分布是以 θ 后验分布为条件的预测分布期望值, 不同模型的后验预测分布能够用来比较推断的准确程度 (Gelman et al., 1995)^[71], 类似的方法还包括后验分布的预测 P 值和条件 P 值的比较 (Perez and Berger, 2002)^[72]。

另外一种可行的方法是利用贝叶斯因子比较模型。以模型 M_1 和 M_2 为例, 按照贝叶斯法则, 先验分布 $p(M)$ 和似然函数 $p(y|M)$ 的乘积与后验密度成比例, 则模型 M_1 和 M_2 的后验比率为:

$$p(M_2|y)/p(M_1|y) = \{p(M_2)/p(M_1)\} \times \{p(y|M_2)/p(y|M_1)\}. \quad (37)$$

式 (37) 中边际似然函数的比率就是贝叶斯因子 (Kass and Raftery, 1995)^[73], 比值偏大表明建议使用模型 M_2 。贝叶斯因子也有一些常见的变形, 但离差和似然率等统计量并不发生改变, 如对数形式, 其他更多的形式转换可参见 Berger et al.(2003)^[74], Robert and Casella(2004)^[75] 等。

在信用风险度量中, 一些模型由于包含随机效应或未知先验信息, 会导致贝叶斯因子难以计算, 从而出现了贝叶斯因子的近似值, 如 $BIC = \log p(y|\hat{\theta}, M) - (p/2) \log n$ 。以及用于处理更高维度模型的 $DIC = \overline{dev} + pD$, $dev = -2 \log(\text{likelihood}) = -2 \log p(y|\theta)$, \overline{dev} 表示离差的后验均值 (Spiegelhalter et al., 2002)^[76]。Celeux et al.(2003)^[77] 提出了用来计算含有潜因素的模型 DIC 完善方法。

3.2 联合估计

扩大参数空间是模型选择方法的基本思想 (Carlin and Chib, 1995^[78]), 联合估计中较为完善的一种方法称为可逆跳跃 MCMC (Reversible Jump MCMC, RJMCMC), 在有的文献中也叫做转移强度 MCMC (Green, 1995^[79])。可逆跳跃 MCMC 的基本做法是将模型看做变量, 应用 MCMC 模拟模型间的跳跃, 对不同可能的模型不需要单独隔离估计。对于具有混合效应的信用风险度量模型, 建模时存在混合成份未知的情形, Richardson and Green(1997)^[80] 提出需要进行额外的 Metropolis-Hastings 计算, 在模拟过程中决定增减、分离或是合并成份。对复杂模型比较, 还可能需要对不同模型中的参数建桥, 如基于已有的两个均值生成新的均值, 最后通过比较模型在 MCMC 中接受的时间比例选择模型, 具体操作方法参见 Robert and Casella(2004)^[81]。

Stephens(2000)^[82] 对生灭 MCMC (Birth and Death, BDMCMC) 方法进行了拓展, 将高维度模型的跳跃时间间隔看做随机变量, 模拟推断变量步入较低维度模型所需的时间。不同于 RJMCMC, BDMCMC 方法在选择模型时, 主要依据 MCMC 在该模型滞留时间的长短 (Cappe et al., 2003)^[83]。

除上述方法外,贝叶斯模型比较方法还包括比估计 (Comparative Estimation),主要用于生物、工程、物理等领域的模型选择,涉及的方法包括熵距离和 Kullback-Leibler 散度等。这种距离测度方法的基本思想就是两个模型后验分布间的距离如果足够小,则倾向于选择模型中的相对简单的形式。距离测度方法的拓展有很多形式, Mengersen and Robert(1996)^[84] 使用 Kullback-Liebler 测度给出了模型被接受的“无差异区域”, Sahu and Cheng(2003)^[85] 对熵距离的应用给出了更为全面的说明。隔离估计仍是当前贝叶斯信用风险度量模型比较的常用方法。

4 国内研究

从 20 世纪 90 年代末期开始,国内研究信用风险度量方法的文献逐渐增加,但已有研究以对比违约概率度量方法为主,涉及变量挖掘的文献较少。在一些具有代表性的文献中,王春峰等 (1998)^[86] 对多元判别分析方法做了较为全面的比较,而且提出了信用风险度量的小样本方法;方洪全等 (2004)^[87] 借助多元统计分析方法筛选信用评级指标,对多标准等级判别模型进行了拓展;于立勇等 (2004)^[88] 利用贝叶斯判别方法构建信用评级指标体系;姚潇和余乐安 (2012)^[89] 在信用风险度量中应用了模糊近似支持向量机方法。

采用贝叶斯方法度量信用风险的文献近年来逐渐显现,但在已有的文献中模型多保留着传统形式,将贝叶斯方法作为辅助性计算工具。在贝叶斯模型的应用中,程建等 (2009)^[90] 针对对小样本违约数据,提出了改进模型预测力的先验加权方法;钟波和江青松 (2007)^[91] 对金融风险值测算的贝叶斯模型进行完善;周丽莉和丁东洋 (2011)^[92] 应用贝叶斯分层模型对相关数据进行实证分析。总而言之,国内研究多是进行实际数据的经验分析或是方法间的对比研究,缺少对模型比较和选择方法的探索。

国内更多的早期文献评述还可参见管七海和冯宗宪 (2004)^[93] 以及龚朴和何旭彪 (2005)^[94],他们对传统的信用风险度量模型 (主要是结构模型和强度模型) 给出了比较全面的介绍。

5 结语

哈罗德和沃伦教授 (1970)^[7] 将贝叶斯方法引入信用风险度量的文章发表于“管理科学 (Management Science)”,而王春峰教授等人 (1998)^[86] 研究信用风险度量方法的文章发表于“管理科学学报 (Journal of Management Sciences in China)”,虽然这仅仅是个巧合,但是两篇文章差近 30 年可以发现该领域的差距,而到中国学者将贝叶斯方法用于度量信用风险,时间相差 35 年以上。在统计学内部争论激烈的年代,波依雷尔 (Poirier, 1988)^[95] 将那些非贝叶斯学派视为异教徒。如今,频率论学派和贝叶斯学派的争论趋于宽容,但风险度量领域是一个新兴而热门的领域,方法问题的争论才刚刚开始,但这种争论有助于该领域的完善与发展。从研究进展来看,需要探讨的问题仍比较多,值得关注的主要有三个方面:

(1). 贝叶斯方法在资产定价中的应用。信用风险分析不仅要度量违约风险,考察债务人信用质量,而且还涉及重要的信用衍生品定价问题。由于资产定价相关的数据信息难以获得,至今对资产定价的贝叶斯研究仅有少量文献。未来研究需解决的问题仍然较多,既包括古老的难题,如股权溢价之谜 (Mehra and Prescott, 1985)^[96], 无风险利率的数理论证 (Weil, 1989)^[97], 以及超额波动性缘由等,又涉及新出现的争议,如市场风险与信用风险综合度量方法的选择 (Campbell, 2003)^[98]。

(2). 贝叶斯方法处理半参数模型的稳健性。半参数 (Semiparametrics) 模型是二十世纪八十年代发展起来的一种重要的统计模型,模型中既含有参数分量又含有非参数分量,非参数分量用于表示模型误差。半参数模型的贝叶斯处理方法仍存在不足,特别是在信用风险分析

中涉及违约概率动态特征, 如随机效应的分布族选择以及宏观经济影响的滞后性说明等问题。此外在信用资产组合中, Mongin(1995)^[99] 指出贝叶斯模型不能同时考虑所有决策者的个人偏好, 从而难以实现帕累托最优。

(3). 信用风险度量问题的国内研究重点。国外相关文献对理论研究具有具有重要参考价值, 可是从我国信用风险度量的应用情况来看, 仍面临不少问题, 如风险数据库等基础设施还有待进一步完善。未来的研究应侧重以下几点: 第一, 注重模型关键变量探寻的研究。当前金融机构在评价信用质量时, 往往通过主观判别选取指标变量, 今后应该加深理论推导, 结合国内现实问题和信用风险的特点探寻关键变量; 第二, 提高模型比较和选择的技术层次。信用风险领域相关文献的样本数据来源差异非常明显, 一些研究人员仅根据自身的数据特征选择变量和设定模型, 未来的研究需提高比较和选择模型的标准, 关注信用风险度量模型在样本范围外的统计检验结果; 第三, 推动贝叶斯理论的普及和应用。信用风险度量中的常见难题, 如数据缺乏, 样本容量小等问题十分突出, 风险分析人员需要利用专家观点和经验, 这些特征十分适于采用贝叶斯数据分析方法, 将贝叶斯方法融于信用风险度量技术有助于提高我国信用风险管理水平。

总之, 贝叶斯方法在信用风险度量中发挥日益重要的技术支撑作用, 尽管本文展示的贝叶斯信用风险度量模型有限, 但可以看出信用风险度量问题涉及的基本知识结构和归纳逻辑确实属于贝叶斯方法的范畴。引用哈佛大学 Feldstein(2004)^[100] 教授的一句话作为本文的结尾: “风险分析的本质是处理不确定性, 而关键是贝叶斯统计决策方法的运用”。

[参考文献]

- [1] Neapolitan R E. A survey of uncertain and approximate inference [A]. Fuzzy Logic for the Management of Uncertainty [C], New York: John Wiley & Sons, 1992: 55-82.
- [2] Lando D, Skodeberg T. Analyzing Rating Transitions and Rating Drift with Continuous Observations [J]. Journal of Banking and Finance, 2002, 26(2): 423-444.
- [3] Jafry Y, Schuermann T. Measurement, Estimation and Comparison of Credit Migration Matrices [J]. Journal of Banking and Finance, 2004, 28(2): 2603-2639.
- [4] Scott F. Bayesian Methods in Risk Assessment [R]. Working Paper. Head of Waste & Storage Unit, 2004.
- [5] Philosophov L V. Valuation of Risky Debt: A Multi-Period Bayesian Model [R]. Working Paper, Moscow Committee of Bankruptcy Affairs, 2011.
- [6] Qin D. Bayesian Econometrics: The First Twenty Years [J]. Econometric Theory, 1996, 12(1): 500-516.
- [7] Harold Bierman J R, Warren H. The Credit Granting Decision [J]. Management Science, 1970, 16(8): 519-532.
- [8] Majone G. Classification and Discrimination in the Analysis of Credit Risks: II [A]. Management Sciences Research Report [C], Carnegie-Mellon University, 1968: 128.
- [9] Sandor Zsolt, Michel Wedel. Designing Conjoint Choice Experiments Using Managers' Prior Beliefs [J]. Journal of Marketing Research, 2001, 28(1): 430-444.
- [10] Montgomery Alan L, Peter E Rossi. Estimating Price Elasticities with Theory Based Priors [J]. Journal of Marketing Research, 1999, 36(1): 413-423.
- [11] Wedel M, Rik Pieters. Eye Fixations on Advertisements and Memory for Brands: A Model and Findings [J]. Marketing Science, 2000, 19(2): 297-312.
- [12] Rossi P E, Allenby G M. Bayesian Statistics and Marketing [R]. Working Paper, University of Chicago, 2002.
- [13] Sims C. Money, Income and Causality [J]. American Economic Review, 1972, 62(1): 540-553.
- [14] Litterman R. Techniques for Forecasting with Vector Autoregressions [D]. University of Minnesota, Ph. D. Dissertation, 1980.

- [15] Jarrow R A, Lando S M, Turnbull A. A Markov Model for the Term Structure of Credit Risk Spreads [J]. *The Review of Financial Studies*, 1997, 10(2): 481-523.
- [16] Kijima M. Monotonicities in a Markov Chain Model for Valuing Corporate Bonds Subject to Credit Risk [J]. *Mathematical Finance*, 1998, 8(3): 229-247.
- [17] Lando D. On Rating Transition Analysis and Correlation [A]. *Credit Derivatives: Applications for Risk Management, Investment and Portfolio Optimization* [C], New York: Risk Publications, 1998: 147-155.
- [18] Arvanitis A, Gregory J, and Laurent J P. Building Models for Credit Spread [J]. *The Journal of Derivatives*, 1999(6): 27-43.
- [19] Huge B, and Lando D. Swap Pricing with Two Sides Default Risk in a Rating Based Model [J]. *European Finance Review*, 1999(3): 239-268.
- [20] Sanjiv R, Rong F, and Gary G. Bayesian Migration in Credit Ratings Based on Probabilities of Default [J]. *The Journal of Fixed Income*, 2002, 12(1): 17-23.
- [21] Wilson T C. Portfolio Credit Risk (I and II)[J]. *Risk*, 1997, 10(3): 56-61; 111-117.
- [22] Nickell P, Perraudin W, Varotto S. Stability of Rating Transitions [J]. *Journal of Banking and Finance*, 2000, 24(1): 203-227.
- [23] Bangia A et al. Ratings Migration and the Business Cycle, Applications to Credit Portfolio Stress Testing [J]. *Journal of Banking and Finance*, 2002, 26(2): 445-474.
- [24] Kavvathas D. Estimating Credit Rating Transition Probabilities for Corporate Bonds [R]. Working Paper. University of Chicago, 2001.
- [25] Carling K, Jacobson T, Linde J, Roszbach K. Capital Charges under Basel II: Corporate Credit Risk Modeling and the Macro Economy [R]. Working Paper. Sveriges Riksbank, 2002.
- [26] Koopman S J, Lucas A. Business and Default Cycles for Credit Risk [J]. *Journal of Applied Econometrics*, 2005, 20(2): 311-323.
- [27] Fledelius P, Lando D, Nielsen J P. Non-parametric Analysis of Rating Transition and Default Data [J]. *Journal of Investment Management*, 2004, 2(2): 71-85.
- [28] Rao N, Wu C. Bayesian Pseudo-empirical-likelihood Intervals for Complex Surveys [J]. *Journal of Royal Statistical Society, Ser. B*, 2010, 72(1): 533-544.
- [29] Kim C J. Dynamic Linear Models with Markov Switching [J]. *Journal of Econometrics*, 1994, 60(1): 1-22.
- [30] Kim C J, and Nelson C R. *State-Space Models with Regime Switching* [M]. Cambridge: MIT Press, 1999.
- [31] Kim C J, Murray C J. Permanent and Transitory Components of Recessions [J]. *Empirical Economics*, 2002(2): 163-183.
- [32] Sichel D E. Business Cycle Duration Dependence: A Parametric Approach [J]. *Review of Economics and Statistics*, 1991, 73(3): 254-256.
- [33] Watson J. Business Cycle Durations and Postwar Stabilization of the U.S. Economy [J]. *American Economic Review*, 1994(84): 24-46.
- [34] Diebold F X, Rudebusch G D. Measuring Business Cycles: A Modern Perspective [J]. *The Review of Economics and Statistics*, 1996, 78(3): 67-77.
- [35] Litterman R. Forecasting with Bayesian Vector Autoregressions: Five Years of Experience [J]. *Journal of Business and Economic Statistics*, 1986(4): 25-38.
- [36] Couderc F, Renault O. Time-to-default: Life Cycle, Global and Industry Cycle Impacts [R]. Working Paper. University of Geneva, 2004.
- [37] Gagliardini P, and Gouriéroux C. Stochastic Migration Models with Application to Corporate Risk [R]. Working Paper. CREST, 2004.
- [38] McNeil, A J, and Wendin J. Dependent Credit Migrations [R]. Working Paper. NCCR FinRisk, 2004.
- [39] Koopman S J, Lucas A, Daniels R J O. A non-Gaussian Panel Time Series Model for Estimating and Decomposing Default Risk [R]. Working Paper. Vrije Universiteit Amsterdam, 2005.
- [40] Geweke J. *Contemporary Bayesian Econometrics and Statistics* [M]. New York: Wiley, 2005.
- [41] Dungey M, Martin V L, Pagan A R. A Multivariate Latent Factor Decomposition of International Bond Yield Spreads [J]. *Journal of Applied Econometrics*, 2000, 15(1): 697-715.

- [42] Diebold F X, Rudebusch D, Boragan Aruoba S. The Macroeconomy and the Yield Curve: a Dynamic Latent Factor Approach [J]. *Journal of Econometrics*, 2006, 131(1): 309–338.
- [43] Schuermann T, Stiroh K J. Visible and Hidden Risk Factors for Banks [R]. Working Paper. Financial Institutions Center, Wharton School of Business, 2006.
- [44] Bauwens L, Hautsch H. Dynamic Latent Factor Models for Intensity Processes [R]. Working Paper. CORE, 2003.
- [45] Koopman S, Lucas A, Monteiro. The Multi-state Latent Factor Intensity Model for Credit Rating Transitions [J]. *Journal of Econometrics*, 2008, 142(5): 399–424.
- [46] Durbin J, Koopman S J. Monte Carlo Maximum Likelihood Estimation for Non-Gaussian State Space Models [J]. *Biometrika*, 1997, 84(3): 669–684.
- [47] Durbin J, Koopman S J. Time Series Analysis by State Space Methods [M]. Oxford: Oxford University Press, 2001.
- [48] Engle R F, Russell J R. Autoregressive Conditional Duration: a New Model for Irregularly Spaced Transaction Data [J]. *Econometrica*, 1998, 66(5): 1127–1162.
- [49] Russell J R. Econometric Modelling of Multivariate Irregularly Spaced High-frequency Data [R]. Working Paper. University of Chicago, 1999.
- [50] Bradlow Eric T, David Schmittlein. The Little Engines That Could Modeling the Performance of World Wide Web Search Engines [J]. *Marketing Science*, 1999, 19(1): 43–62.
- [51] Schönbucher P J. Credit Derivatives Pricing Models [M]. New Jersey: John Wiley & Sons, 2003.
- [52] McNeil A, Wendin J. Bayesian Inference for Generalized Linear Mixed Models of Portfolio Credit Risk [J]. *Journal of Empirical Finance*, 2007, 14(1): 131–149.
- [53] Egloff D, Leippold M, Vanini P. A Simple Model of Credit Contagion [R]. Working Paper. Swiss Banking Institute, University of Zurich, 2004.
- [54] Wilde T. CreditRisk+ A Credit Risk Management Framework [Z]. Technical Report, Credit Suisse First Boston, 1997.
- [55] Stefanescu C, Tunaru R, Turnbull S. The Credit Rating Process and Estimation of Transition Probabilities: A Bayesian Approach [J]. *Journal of Empirical Finance*, 2009, 16(1): 216–234.
- [56] KMV Corporation KMV and Credit Metrics [M]. San Francisco: KMV Corporation, 1997.
- [57] McCulloch C. Maximum Likelihood Algorithms for Generalized Linear Mixed Models [J]. *Journal of the American Statistical Association*, 1997, 92(3): 162–170.
- [58] Frey R, McNeil A. Dependent Defaults in Models of Portfolio Credit Risk [J]. *Journal of Risk*, 2003, 6(1): 59–92.
- [59] Gelman A et al. Bayesian Data Analysis, 2nd ed.[M]. London: Chapman and Hall, 2003.
- [60] Skrondal A, Rabe-Hesketh S. Generalized Latent Variable Modeling [M]. London: Chapman and Hall, 2004.
- [61] Kadam A, Lenk P. Bayesian Inference for Issuer Heterogeneity in Credit Rating Migration [J]. *Journal of Banking and Finance*, 2008, 32(1): 2267–2274.
- [62] Kalman R. A new Approach to Linear Filtering and Prediction Problems [J]. *Journal of Basic Engineering*, Transaction ASME, Series D, 1960, 82(2): 35–45.
- [63] Kalman R E, Bucy R S. New Results in Linear Filtering and Prediction Problems [J]. *Journal of Basic Engineering*, Transaction ASME, Series D, 1961, 83(3): 95–108.
- [64] Doucet A, Freitas J G. Gordon. Sequential Monte Carlo Methods in Practice [M]. New York: Springer Verlag, 2001.
- [65] Harrison J, West M. Bayesian Forecasting and Dynamic Models, 2nd Ed. [M]. New York: Springer Verlag, 1997.
- [66] Gelfand Alan E, Adrian F M Smith. Sampling-Based Approaches to Calculating Marginal Densities [J]. *Journal of the American Statistical Association*, 1990, 87(3): 523–532.
- [67] Tierney Luke. Markov Chains for Exploring Posterior Distributions [J]. *Annals of Statistics*, 1994, 23(4): 1701–1728.
- [68] Gilks W R, and Wild P. Adaptive Rejection sampling for Gibbs Sampling [J]. *Applied Statistics*, 1992, 41(1): 337–348.
- [69] Congdon P. Bayesian Statistical Modelling [M]. England: Wiley, 2001.
- [70] Alston C et al. Bayesian Model Comparison: Review and Discussion [R]. Working Paper, Queensland University of Technology, 2005.

- [71] Gelman A, Carlin J B, Stern H S, and Rubin D B. Bayesian Data Analysis [M]. London: Chapman and Hall, 1995.
- [72] Perez J M, Berger J. Expected Posterior Prior Distributions for Model Selection [J]. Biometrika, 2002, 89(2): 491-512.
- [73] Kass R, Raftery A. Bayes Factors [J]. Journal of the American Statistical Association, 1995, 90(2): 773-795.
- [74] Berger J, Ghosh J K, Mukhopadhyay N. Approximations to the Bayes Factor in Model Selection Problems and Consistency Issues [J]. Journal of Statistical Planning and Inference, 2003, 112(1): 241-258.
- [75] Robert C, Casella G. Monte Carlo Statistical Methods, 2nd ed.[M]. New York: Springer-Verlag, 2004.
- [76] Spiegelhalter D J, Best N G, Carlin B P, van der Linde A. Bayesian Measures of Model Complexity and Fit [J]. Journal of the Royal Statistical Society Series B, 2002, 64(3):583-639.
- [77] Celeux G, Forbes F, Robert C P, Titterton D M. Deviance Information Criteria for Missing Data Models [R]. Working Paper, National Institute of Statistics and Economic Studies, 2003.
- [78] Carlin B, Chib S. Bayesian Model Choice via Markov Chain Monte Carlo Methods [J]. Journal of the Royal Statistical Society Series B, 1995, 57(3): 473-484.
- [79] Green P. Reversible Jump MCMC Computation and Bayesian Model Determination [J]. Biometrika, 1995, 82(4): 711-732.
- [80] Richardson S, Green P. On Bayesian Analysis of Mixtures with an Unknown Number of Components [J]. Journal of the Royal Statistical Society Series B, 1997, 59(1): 731-792.
- [81] Robert C, Casella G. Monte Carlo Statistical Methods, 2nd ed.[M]. New York: Springer-Verlag, 2004.
- [82] Stephens M. Bayesian Analysis of Mixture Models with an Unknown Number of Components an Alternative to Reversible Jump Methods [J]. Annals of Statistics, 2000, 28(2): 40-74.
- [83] Cappe O, Robert C, Ryden T. Reversible Jump MCMC Converging to Birth-and-Death MCMC and More General Continuous Time Samplers [J]. Journal of the Royal Statistical Society Series B, 2002, 65(3): 679-700.
- [84] Mengersen K, and Robert C P. Testing for Mixtures: a Bayesian Entropic Approach [A]. Bayesian Statistics [C], New York: Oxford University Press, 1996: 255-276.
- [85] Sahu S, Cheng R. A fast Distance based Approach for Determining the Number of Components in Mixtures [J]. Canadian Journal of Statistics, 2003, 31(1): 3-22.
- [86] 王春峰, 万海晖, 张维. 商业银行信用风险评估及其实证研究 [J]. 管理科学学报, 1998, 1(1): 68-72.
- [87] 方洪全, 曾勇. 多标准等级判别模型在银行信用风险评估中的应用研究 [J]. 金融研究, 2004(9): 92-100.
- [88] 于立勇, 詹捷辉, 金建国. 内部评级法中违约概率与违约损失率的测算研究 [J]. 统计研究, 2004(12): 22-26.
- [89] 姚潇, 余乐安. 模糊近似支持向量机模型及其在信用风险评估中的应用 [J]. 系统工程理论与实践, 2012, 32(3): 549-554.
- [90] 程建, 连玉君, 刘奋军. 信用风险模型的贝叶斯改进研究 [J]. 国际金融研究, 2009(1): 63-68.
- [91] 钟波, 汪青松. 基于 Bayes 估计的金融风险值 -VaR 计算 [J]. 数理统计与管理, 2007, 26(5): 881-886.
- [92] 周丽莉, 丁东洋. 基于 MCMC 模拟的贝叶斯分层信用风险评估模型 [J]. 统计与信息论坛, 2011(12): 26-31.
- [93] 管七海, 冯宗宪. 信用违约概率测度研究: 文献综述与比较 [J]. 世界经济, 2004(11): 40-54.
- [94] 龚朴, 何旭彪. 信用风险评估模型与方法最新研究进展 [J]. 管理评论, 2005(5): 8-16.
- [95] Poirier. Frequentist and Subjectivist Perspectives on the Problems of Model Building in Economics [J]. Journal of Economic Perspective, 1988, 2(1): 121-144.
- [96] Mehra R, Prescott E C. The Equity Premium: A Puzzle [J]. Journal of Monetary Economics, 1985, 15(3): 145-161.
- [97] Weil P. The Equity Premium Puzzle and the Risk-Free Rate Puzzle [J]. Journal of Monetary Economics, 1989, 24(2): 401-421.
- [98] Campbell J Y. Consumption-Based Asset Pricing [A]. Handbook of the Economics of Finance [C], North-Holland Elsevier, 2003(1): 803-887.
- [99] Mongin P. Consistent Bayesian Aggregation [J]. Journal of Economic Theory, 1995, 6(1): 313-351.
- [100] Feldstein M. Innovations and Issues in Monetary Policy: Panel Discussion [J]. American Economic Review, 2004, 2(1): 41-48.