

生命周期动态模拟技术及其在税收领域应用

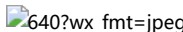
2021/4/22 8:54:02 新闻来源: 中国仿真学会生命系统建模仿真专业委员会

生命周期动态模拟技术脱胎于两维时间分解技术,以后者为基础的建模和预测在欧美已有15年以上的历史,是银行进行长期战略规划和压力测试的重要手段。本文首先回顾了生命周期动态模拟技术的历史沿革,其次介绍了该技术的理论基础及核心概念,再次以美联储用其对金融业进行压力测试这一典型案例来展示其工作原理,最后对该技术如何在税收领域应用进行了分析和探讨。

一、生命周期动态模拟技术的历史沿革

生命周期动态模拟技术的前身是两维时间分解技术。以后者为基础的建模和预测在欧美已有15年以上的历史,是银行进行长期战略规划和压力测试的重要手段。该技术中使用的基本概念和分析方法已深入到银行和监管机构的业务实践与合规要求中。两维时间分解技术由Joseph Breeden博士创立,是应用建模领域中的一个里程碑。传统计量经济学中把面板数据(Panel Data)结构定义在两维空间轴上,而Breeden博士把这一定义扩展到了两维时间轴上,从而为两维时间分解提供了思路。随着时间的推移,两维时间分解中使用的主要概念,特别是同龄组分析(Vintage Analysis)和因素分解(Factor Decomposition)等,在个贷风控领域已是众人皆知。

触发生命周期动态模拟技术在两维时间分解技术基础上发展起来的原因主要有三个:一是两维时间分解中繁杂的数学演绎超出了压力测试过程所能承受的极限;二是非大数据思维下建立的体系极大地限制了两维时间分解技术的应用空间,从而排除了其在动态评分和状态周期模型领域的应用;三是生命周期动态模拟技术重新建立了因素分解的数学体系以及在软件实施平台编码中的基本算法,从而在根本上拓展了其应用空间。这种拓展可以归纳为图所示的三个方面。



二、生命周期动态模拟技术的理论基础及核心概念

(一) 技术的理论基础

生命周期动态模拟技术假设生命周期过程是事物演进的主导因素,而与日历时间有关的因素则是对生命周期自然过程的外生干扰。同时,个体的生命周期表现和总体生命周期行为是有差异的,个体对时间因素的响应行为与总体的响应行为也是有差异的;这些差异可以通过个性化参数来量化。

生命周期动态模拟技术的核心是对历史成因在两维时间上的因素进行分解。这一分解在技术流程中被称为一级分解。一级分解是解析过程,其产生的结果是结构模型。从原理上讲,从一级分解中得出的任何因素,无论是生命周期行为、时间响应行为或是个体拟合参数等,都是可以进一步分解建模,即进入二级建模过程的。二级建模都是在单一时间维度上进行的。最为常见的二级分解集中在时间响应层面,而对个体拟合参数的二级分解也有重要的应用。

生命周期模拟技术在一级分解时使用的是分步解析与迭代滤波技术;在二级分解时使用的是线性回归和时间序列模型。

[业界动态](#)[业界动态](#)[图片中心](#)[点击排行](#)[自主知识产权的仿真软件需求](#)[航天系统仿真重点实验室2020年](#)[中国仿真学会:“智汇”仿真科技人](#)[湖南省系统仿真学会成立](#)[北京仿真中心航天系统仿真重点实](#)[航天科工集团二院北京仿真中心](#)[纯粹数学的雪崩效应:庞加莱猜想](#)[航天系统仿真重点实验室 召开20](#)[青年工作委员会“智能人机交互领](#)[“大数据分析与应用技术国家工程](#)

(二) 技术的核心概念

生命周期。生命周期动态模拟过程首先需要生命周期行为进行刻画，因为这是一切生物和生态过程、社会与商业过程最为本质的特征。生命周期过程可以是关于生命特质或指标的，也可以是关于现象的，亦或是关于过程和结局的。

时间响应。每一个具体的生命周期过程都是在特定的历史时段下走完的。如果不把环境因素剔除，就得不到纯粹的生命周期行为曲线。对环境因素影响的量化分析除了能对历史提供解释，还能为模拟未来提供依据。

同龄组拟合参数。生命周期曲线和时间响应曲线都是共性行为曲线。生命周期动态模拟技术对每一个最小的“细胞”单元，即每一个同龄组和特征组的交集，都会计算出一组特征参数，以保证对个体模拟的精度。

同龄组划分和同龄组分析。同龄组是生命周期技术的基石概念，生命周期曲线要刻画的就是同龄组从生成到终结的平均行为过程。同龄组分析是分析当前趋势的重要手段。

特征组划分和特征组分析。特征组分析直接涉及分析的深度和对决策的实用价值。例如，同一年出生的人口，不同的出生和成长背景对于任何生命周期行为（如升学、就业、消费等等）和时间响应行为都是不同的。揭示这些差别不但是对历史进行研究的目的，也是使预测结果细致化和准确化的前提。

三、生命周期动态模拟技术的价值

十余年来银行数据的一体化和以面板数据为基础的结构化为动态模拟技术的实现创造了必要条件，而动态模拟技术可以应用于整体风控和业务管理的各个功能与环节，并适用于对个贷行业已知的合规应用，这主要包括：

(一) 把业务运行纳入一个完整的系统

以个贷业务为例的业务流程含市场营销、授信、存量账户管理、催收和财务管理等。在内部数据整体化和结构化的基础上，对整个业务的全流程实现可视化监测、一体化建模与决策分析就变得顺理成章了，这样对多个变量、多种用途和不同应用场景的表象监测、建模和预测就可以使用同一个过程来完成。

(二) 建立动态的控制机制

风控是一个典型的控制论课题，即是一个无休止的反馈调节过程。生命周期动态模拟技术的贡献是对操作程式的建立和对操作平台的提供，其涵盖的范围包括试错分析、边缘人群分析、人群精细划分、对决策选项的事先模拟、可能性分析和路径优化等。

动态评分对于任何针对个人的评分系统都是重要的。以个贷行业为例，它可以对历史短暂而变化迅速的新金融贷款产品进行客户评分，同时对未来进行任意时段的评分，使评分同时反映真实概率和风险排序。信用分数在不同环境下的变化，对产品、银行和行业诸个层面风险指数的建立具有重要价值。

(三) 对控制资源的集约化使用

集中的数据组织和平台化的技术手段可以在有限的投资下以少量的人员更好更快地实现以往劳动密集型技术所无法实现的产出。具体到个贷领域，动态模拟系统的实施可以同时为下述三大种类的决策提供方案：

一是围绕客户个体的决策——即在精准营销、授信、贷后管理和催收中的动态分析与评分；二是围绕业务运行决策的应用——包括长期规划、年度计划以及季（月）度滚动短期计划；备选策略的事前模拟及对实施偏差的敏感度分析；对损失与利润指标的表现监测、影响因素分析及未来预测；行业走向分析，企业表现与行业表现的基准比较。三是围绕着合规要求的应用——包括数据的一体化和规范化、损失压力测试、行业风险的量化（指数化）监控等。

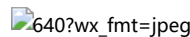
四、生命周期动态模拟技术典型应用案例

生命周期模拟技术一个典型应用是美联储用其对金融业进行压力测试。从美国始发的金融危机源于过度房贷导致的大规模个人信用违约。随之而来的连锁效应波及了整个世界金融界和实体经济，使各国监管当局对金融业系统性震荡的严重性有了切实的体验。而更为深层的忧虑是这次金融危机的发生始料未及。有鉴于此，美国和欧洲相继出台了有关法案，要求各大银行在对未来损失的预测中加入压力测试，并由此决定银行企业层面的储备金总量，以防范个体或局部危机触发系统性危机。

就模型的流程而言，美联储在其独立于个体贷款机构之外所进行的全行业压力测试中，使用的是巴塞尔协议中的PD-LGD-EAD流程体系。在该流程体系下，银行根据标准参数或内部估计确定其风险要素，并计算得出银行所面临的风险。这些风险要素主要包括：

1. **违约概率 (PD)** ——指债务人违反贷款相关规定，没有按时偿还本金和利息的概率；
2. **违约损失率 (LGD)** ——指债务人没有按时偿还本金和利息给银行带来的损失的状况，它表现为单位债务的损失均值；
3. **违约风险敞口 (EAD)** ——指交易对象违约时，对银行所面临的风险的估计。该体系的主要目标就是使得资本的配置更加精确，与银行内部的信用风险更加匹配。

巴塞尔PD-LGD-EAD体系基本流程如图所示。在该流程图中，其表现变量是违约账户的数量和金额，其驱动变量是违约概率、平均损失金额和净损失率。

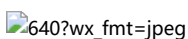


(一) 压力测试的基本要求

对压力测试的产出要求是提供上述诸变量在不同经济景气假设条件下的预测值。例如，图3.2显示的是在美联储三种不同的经济景气假设条件下某假想银行对违约概率PD的预测结果，即违约损失余额在未来18个月中的数值和走向。由此我们看到，美联储等监管机构对压力测试在操作上的最基本要求是：

建立违约损失和经济景气变量之间在历史上的数量关系；

将未来经济景气的假设带入上述关系，得出未来损失在不同经济景气假设下的预测。

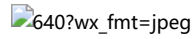


(二) 经济环境影响的分离

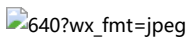
为实现上述基本要求，我们需要对违约概率进行两级分解。一级分解首先将生命周期因素（即月龄增长因素）和与日历时间相关联的因素分离开来，二级分解要把经济影响从与日历时间相关的总体影响中分离出来。实现这个目的的总体思路是把时间响应曲线分解为：事件（如果存在）因素的影响、季节波动因素的影响和环境因素的影响。

1. 一级分解

一级因素分解中产生的生命周期曲线刻画贷款增龄对违约概率的决定作用。如图所示，所有特征组（高度风险-分行渠道、中度风险-分行渠道、低度风险-分行渠道）的违约概率都明显地是月龄的函数；而对于不同的特征组，这一函数是十分不同的。通过这种比较，我们可以对违约表现做出特征组分析。由于生命周期行为中滤除了环境因素的影响，因而它可以成为监测贷款业务的风险特征，并进行银行间风险偏好的比较依据。

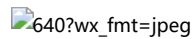


接下来分解的是与日历时间相关的因素对违约概率的影响。在时间响应曲线里面，通常将一定时期内违约概率的正常水平（比如3.5%）定义为100%（基准值）。每个时点违约概率采用相对于基准值的比率来刻画。比如，如果某个时点违约概率升至7%，则该时点在时间响应曲线里对应的纵轴坐标是200%；如果某个时点违约概率降至1.75%，则该时点在时间响应曲线里对应的纵轴坐标是50%。如图所示，在金融危机逐步加深的过程中，违约概率从50%左右上升到了150%左右（2008年-2009年间的平均水平）。在其后的复苏和发展时期，环境因素走向了利好的一面，从而使违约概率逐步下降。

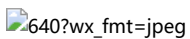


2. 二级分解

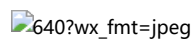
二级分解是对图3.4中的时间响应曲线做出的进一步分解。分解中首先是将事件和季节性波动滤出，季节性波动曲线如图所示。



将上述两个因素剥离以后，就得出纯粹的环境影响曲线，如图所示。

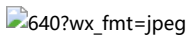


图则显示了环境影响曲线和失业率之间的相关关系，而对这种关系的刻画是建立经济响应方程的课题。

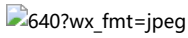


3. 经济响应方程的建立

建立经济响应方程在方法上是使用回归方程，即用经济景气变量（例如本例中的失业率）来解释上述环境影响曲线。显示的是回归分析的结果，深色曲线代表被失业率解释了的部分。



在环境曲线与经济景气变量之间的历史关系以回归方程的形式确定以后，这一方程就可用来进行压力测试。图3.9中重合的曲线表示的是违约概率的历史数据，三条不同颜色的分叉曲线是在三种经济景气假设下对违约概率的预测。



测

五、生命周期动态模拟技术在税收领域的应用场景

生命周期动态模拟技术在发展中形成了其特有的分析方式，在金融领域取得了广泛的应用。那么，生命周期动态模拟技术是否可以应用在税收领域呢？答案是肯定的。该技术的核心功能之一是为管理者提供各种决策依据。税收领域的诸多现象，比如税收贡献、风险特征、纳税人的经济行为等都会随着生命周期而发生变化，并也都将毫无例外地受到环境因素的影响。因此，无论是主动的税收政策措施还是被动的税收应变手段，都可以依靠动态模拟技术使决策过程有可视化和系统化的依据。该技术可以应用在税收领域的以下诸多方面：

（一）纳税人群体分析

纳税人群体分析的目的是为纳税人分类分级管理提供依据。分类分级管理要求我们能够充分地细分纳税人群体，及时地观测细分群体的表现，并对表现上的差异进行因素分析，从而做出以数据分析为基础的、理智的、有内在反馈机制的动态决策。在实施上，特定纳税人群体分析同时依赖于同龄组分析、特征组分析和因素分解。

特征组分析涉及的是如何对纳税人群体进行精细分类，并对纳税人群体之间的表现差异进行比较。对于同一个同龄组，地域条件、自然环境、所属行业等特征的不同，其税收表现都是相异的。生命周期动态模拟技术可以在上述各个层面上刻画这些因素的影响。

因素分解的目的是为细分纳税人群体之间的表现差异提供解释。有两个层次的差异比较：一是直接的数据比较，从而得到观察性的结论；另一个是更深层次上的比较，是具有分析意义的比较，即对差异产生来源进行比较。

（二）税收收入预测

传统的税收收入预测方法是在单时间维度上展开的，从而在模型的大类划分上就有了生命周期模型(年龄时间)以及时间序列模型(日历时间)，非此即彼。生命周期动态模拟技术的特有机制是将生命周期因素和环境因素融入同一个模型体系，并通过机器学习过程将两个时间维度上的影响因素进行剥离。这对于税收收入预测研究来说尤其关键，因为税收收入从本质上说是由生命周期因素主导的；但各种环境因素（比如宏观经济变化、税制调整等）对税收收入都会有显著的影响。这些因素如果不能被有效地剔除，就得不到可靠的生命周期曲线。

（三）对纳税信用动态评分和建立纳税人动态画像

纳税信用动态评分提供了传统的逻辑回归评分无法或难于实现的功能，并且能够实现对未来进行任意时段而不是固定时段（如未来12个月）的纳税信用进行评分。信用分数在实现排序功能的同时也产生出真实的概率数值，并且信用分数能体现环境变化对风险的影响，进而为税收风险指数的建立提供计算基础。在此基础上，可以构建面向风险管控和决策管理的纳税人动态画像，涵盖宏观分析、行业趋势、外部环境等多个业务场景。

（四）税收政策效应分析

经济业态的发展日新月异，因此建模中须使用最全面最新近的数据，否则就会产生未来预测与历史之间的脱节。动态模拟的主旨是使决策者有一个评估税收政策影响的实验室。不同的税收政策选项，当它们被单一或叠加实施的时候会如何影响未来的税收收入？这种模拟对于决策过程本身无疑是重要的，而其衍生的意义也同样是重要的，即为未来实施结果的监测、为税收业务部门指标的设定和业绩评估提供参照。

六、结语

生命周期动态模拟就是将动态模拟技术引入对生物和生态行为、对社会和经济过程的研究和决策过程。它和众多的机器学习和图形识别技术在目的上是一脉相承的，但在实现方法上则是针对社会和经济过程的特性而建立的。

生命周期动态模拟可以演绎在不同假设历史条件下的虚拟结果，从而使我们更好地理解现实的(或实现了的)历史。动态模拟技术对未来的预测都是场景预测，因而可以演示出未来结果的多重可能性。例如，对一项或者多项税收优惠政策的调整，对税收收入的影响将是怎样的？

生命周期动态模拟技术对于未来业务决策方式的影响力是可以预期的，因为该技术同时顺应了大数据和动态模拟这两个发展趋势。其带来的变革将主要体现在：监管和业务运行中信息的一体化，从而避免了信息的扭曲和不对称性；历史表现监测和未来预测的一体化以及相应结果的可视化；对整个过程而不是仅在局部环节上的表现监测和历史成因分析；以最为集约的方式进行大规模建模，并为政策的制订与动态调整提供实证和实施手段。

相关链接：

地址：北京市海淀区学院路37号工程训练中心637室 电话：010-82317098 传真：010-82317098

中国仿真学会 版权所有 电子邮箱：cassimul@vip.sina.com

京ICP备17016611号-1; 技术支持：北京中捷京工科技发展有限公司(010-88516981)