灰色预测理论及其应用

姚天祥　巩在武　著

科学出版社

灰色预测理论及其应用

姚天祥　巩在武　著

国家自然科学基金(71171116)
教育部人文社会科学基金(09YJC630129)　联合资助
江苏高校优势学科建设工程
江苏高校哲学社会科学重点研究基地“中国制造业发展研究院”

科　学　出　版　社
北　京

内 容 简 介

本书介绍了灰色预测建模的基本理论、基本方法及实际应用，总结了作者长期从事灰色系统理论研究和教学工作过程中取得的新成果，反映了灰色预测建模理论与方法的前沿发展动态. 全书共 8 章，包括灰色预测理论发展概况、GM(1, 1) 模型及其特性、GM(1, 1) 模型的改进、内涵型 GM(1, 1) 模型、离散灰色预测模型、非等间距灰色预测模型、序列缓冲算子模型、灰色预测理论在能源领域中的应用. 书中大部分内容是作者的研究成果.

本书可作为高等院校和科研院所的经济、管理及相关专业大学生和研究生的参考教材，也可供政府机关、企业决策部门的有关人员参考.

图书在版编目(CIP)数据

灰色预测理论及其应用 / 姚天祥，巩在武著. —北京：科学出版社，2014.10

ISBN 978-7-03-042092-3

Ⅰ. ①灰… Ⅱ. ①姚… ②巩… Ⅲ. ①灰色预测模型–建立模型–研究 Ⅳ. ①N949

中国版本图书馆 CIP 数据核字(2014)第 230140 号

责任编辑：伍宏发　曾佳佳　陈会迎 / 责任校对：郭瑞芝
责任印制：徐晓晨 / 封面设计：许　瑞

科学出版社出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

北京京华虎彩印刷有限公司 印刷
科学出版社发行　各地新华书店经销

*

2014 年 10 月第 一 版　开本：720 × 1000 1/16
2014 年 10 月第一次印刷　印张：11 5/8
字数：234 000

定价：58.00 元

(如有印装质量问题，我社负责调换)

前　言

灰色系统理论是邓聚龙教授1982年创立的新学科，而灰色预测模型是灰色系统理论近年来发展最快、研究最活跃的领域．国内外应用灰色预测理论与方法撰写的硕士、博士学位论文日益增多，每年都有大量灰色预测领域的学术论文在国内外期刊发表．

本书是在作者的博士学位论文——《灰色预测建模理论与方法研究》的基础上，结合近年来最新研究成果，经修改、扩充而成．

本书由姚天祥提出写作方案、统稿．第1章～第7章由姚天祥执笔；第8章由巩在武、姚天祥执笔．

本书的出版得到了南京信息工程大学气象灾害预报预警与评估协同创新中心的支持，得到了国家自然科学基金（71171116）、教育部人文社会科学基金(09YJC630129)、江苏高校优势学科建设工程、江苏高校哲学社会科学重点研究基地“中国制造业发展研究院”的资助．本书的出版得到了很多老师和同学的支持与指导．衷心感谢灰色系统专业委员会理事长刘思峰教授，作者读博士期间得到了刘老师的多方面指导，刘老师严谨的治学态度与宽厚待人的长者之风是作者一生学习的榜样！感谢南京航空航天大学的党耀国教授、方志耕教授、谢乃明博士！感谢研究生顾红、程文荣及其他提供帮助的老师和同学！

由于作者水平有限，书中存在不足之处在所难免，恳请读者批评指正．

姚天祥

2014年5月

目　　录

第 1 章　灰色预测理论发展概况

1.1　灰色系统理论的发展

自 1982 年华中工学院邓聚龙教授提出灰色系统理论以来，该理论在经济、科教、工程、军事等众多领域得到广泛应用[1-3]．灰色系统理论的研究对象是“部分信息已知，部分信息未知”的“贫信息”不确定系统，它通过对“部分”已知信息的生成、开发实现对现实世界的确切描述和认识．由于系统的复杂性，人们得到的有关研究对象的信息总是不完全的，贫信息不确定性系统的存在决定了灰色系统理论具有十分广阔的发展前景．据中国科学引文数据库发布的信息(《中国科学时报》，1997 年 11 月 26 日)，华中理工大学邓聚龙教授的灰色系统理论被引用 533 次，居全国第一．中国国家科技部编撰出版的《中国科学技术蓝皮书(第 8 号)》把灰色系统理论作为中国学者创立的软科学新方法并加以肯定[4]．经过短短二十多年的发展，灰色系统理论已经初步构建起学科理论体系，灰色系统理论在以下几个方面取得了较大进展．

在学科发展方面，灰色系统理论已经成为系统科学理论群体的一个重要的新成员，在经济管理、信息科学、机械工程、水利工程、航空航天、通信工程、农业工程、医药卫生等学科得到广泛应用，其中在图像处理、石油天然气勘测、有色金属探矿、机床故障诊断等领域取得了一定成果[3-5]．以灰色数学和灰色哲学为基础的灰色系统理论在学科发展上出现了一些新的分支，如灰色水文学、灰色地质学、灰色育种学、区域经济灰色系统分析等分支应运而生[4, 5]．灰色系统理论的研究范围也由最初的灰控制、灰关联、灰预测发展到灰决策、灰聚类、灰规划、灰博弈、灰色投入产出等领域．灰色系统理论与其他学科的交叉取得丰硕成果，灰色系统理论与博弈论、模糊数学、粗糙集、神经网络、遗传算法、可拓学等学科结合，出现了一系列杂合方法和模型．

在国际影响方面，英国、美国、德国、日本、澳大利亚、加拿大、奥地利、俄罗斯等国家，中国台湾、香港地区，以及联合国等国际组织有许多知名学者从事灰色系统的研究工作．邓聚龙教授创办的国际期刊 *Journal of Grey System* 已经被科学引文索引(Science Citation Index，SCI)检索．中国台湾的 *Journal of Grey System* 国际期刊近几年发展也很快．2011 年，刘思峰教授任主编的英文国际期刊 *Grey System：Theory and Application* 得到了 Emerald 的全额资助．*Technological*

Forecasting and Social Change[6-11]、*Decision Support Systems*[12]、*Expert Systems with Applications*[13-16]、*Fuzzy Sets and Systems*[17-19]、*Kybernetes: The International Journal of Systems & Cybernetics*[20-22]、*European Journal of Operational Research*[23-28]、*Journal of Computational and Applied Mathematics*[29]、*Information Sciences*[30]、*Applied Mathematics and Computation*[31-34]、*Applied Mathematics Letters*[35, 36]、*Applied Mathematical Modelling*[37-39]、*Computers & Operations Research*[40]、*Mathematics and Computers in Simulation*[41]等社会科学引文索引(Social Science Citation Index，SSCI)、SCI 源刊开始发表灰色系统理论方面的论文．截至 2008 年 5 月，以“grey system model”为主题词检索，在 SCI 中检索到 363 篇论文，在工程索引(Engineering Index，EI)中检索到 1421 篇论文．通过对检索论文的内容进行分析，检索到的 SCI 收录论文 80%以上应用灰色系统的理论与方法，而检索到的 EI 论文 95%以上应用灰色系统的理论与方法．国内外很多出版机构出版有灰色系统理论专著，其中世界著名出版商 Springer-Verlag 出版了 *Grey Information: Theory and Practical Applications*．很多国际会议将灰色系统理论列为讨论专题，2002 年，在美国匹兹堡召开的系统与控制世界组织(World Organisation of Systems and Cybernetics，WOSC)第 12 届年会和国际一般系统研究会(International Institute for General Systems Studies，IIGSS)第 4 届年会上，南京航空航天大学的刘思峰教授由于在灰色系统理论研究中取得的创造性成果，获得了系统与控制世界组织突出贡献奖．美国电子电气工程师协会(Institute of Electrical and Electronics Engineers，IEEE)的 IEEE International Conference on Systems，Man，and Cybernetics(IEEE SMC)在 2004～2008 年连续 5 年安排灰色系统专题会议，系统与控制世界组织第 12 届至第 14 届年会都连续安排灰色系统专题会议．2003 年 8 月在爱尔兰利默瑞克召开的第 32 届计算机与工业工程国际会议，为灰色系统理论安排了 4 场专题会议[3-5]．2007 年 11 月，首届 IEEE 灰色系统与智能服务国际会议(IEEE International Conference on Grey Systems and Intelligent Services)在南京航空航天大学召开．2008 年，IEEE 灰色系统委员会在南京航空航天大学成立．从 2009 年开始，IEEE 灰色系统与智能服务国际会议每两年举办一次，2013 年在中国澳门举行．

在学科支持方面，国家自然科学基金、各省市自然科学基金、社会科学基金、教育部博士学科点基金、国防基础研究重点项目等积极资助灰色系统研究，每年都有一大批灰色系统理论或应用研究项目获得各类基金资助．灰色系统全国会议连续两次得到中国高等科学技术中心全额资助．

在人才培养方面，世界上有许多大学和机构开设了灰色系统理论课程[3, 5]，如中国的华中科技大学、中国人民大学、清华大学、浙江大学、山东大学、南京航空航天大学、台湾“中央大学”、成功大学、台湾“大同大学”，美国的马里兰大学，日本的丰桥大学、神奈川大学，维也纳经济大学、法国宇航中心等．中国

的华中科技大学、南京航空航天大学、武汉理工大学、福州大学招收、培养灰色系统专业方向的博士研究生. 世界各国高等学校计有上千名博士、硕士研究生运用灰色系统的思想方法开展科学研究，撰写学位论文. 以“灰色系统”为主题词检索，1999 年至 2008 年 5 月，在中国优秀博硕士学位论文库中检索到博士学位论文 499 篇，硕士学位论文 1947 篇，对部分检索到的博硕士论文内容进行查阅发现，95%以上的检索论文应用了灰色系统的基本理论方法从事相关研究.

灰色系统理论中应用最为广泛的是 GM(1, 1)灰色预测模型，其特点在于建模机理与其他模型不同，对数据进行累加处理是其所独创，在建模的数据处理上，需要对数据进行累加生成和累减生成，通过灰色序列生成找寻数据演变的规律性. 灰色预测模型是基于最小二乘方法的指数拟合曲线，具有微分、差分、指数兼容等性质，在建模时不需要大量的时间序列数据就能够取得较好的预测效果，达到较高的精度.

尽管以 GM(1, 1)模型为基础的灰色建模理论得到广泛应用，但该模型存在着很多理论问题需要解决. GM(1, 1)模型是最常用的灰色预测模型，但是基于级比、光滑比判断的建模条件并不完善. 灰色预测建模的样本数目选择具有很大的盲目性，缺乏足够的理论依据. 有必要对样本数据选取规则和选取不同样本数据对模拟误差的影响进行研究. 由于 GM(1, 1)模型是有偏差的指数模型，而离散 GM(1, 1)模型是 GM(1, 1)模型的精确形式，有必要对离散 GM(1, 1)模型的特性进行研究. 对于离散 GM(1, 1)模型与 GM(1, 1)模型的背景值构造方面有待进一步研究. 灰色预测建模应用累加数据，将导致累加数据构成的矩阵具有病态性，数据的微小变化导致参数的巨大变化，对灰色预测模型矩阵的病态性有待进一步研究.

1.2　灰色预测模型的研究内容

灰色预测模型是灰色系统理论最重要的内容之一，也是预测理论体系中一个新的研究分支. 灰色预测模型主要针对现实世界中大量存在的灰色不确定性预测问题，利用少量有效数据和灰色不确定性数据，通过序列的累加生成，揭示系统的未来发展趋势. 鉴于现实世界大量存在的灰色不确定信息，人们对系统的未来趋势难以把握，同时也对经典预测理论和方法提出了挑战，灰色预测模型正是为解决这样一类问题而提出的. 1982 年，北荷兰出版公司出版的《系统与控制通讯》(*Systems & Control Letters*)杂志刊载了我国学者邓聚龙教授的第一篇灰色系统论文——《灰色系统的控制问题》(*Control problems of grey systems*)[42]；同年，《华中工学院学报》刊载了邓聚龙教授的第一篇中文灰色系统论文——《灰色控制系统》[43]. 这两篇开创性论文的公开发表，标志着灰色系统理论这一新兴横断学科经过其创始人邓聚龙教授多年卓有成效的努力，开始问世. 邓聚龙提出了灰色朦胧集，对灰色系统理论的基础进行了研究[44-46]. 王清印等对灰色数学基础和不确

定性信息的概念、类别及其数学表示进行了研究[47-50].

经过二十多年的发展，灰色预测理论已经在工业、农业、社会、经济、能源、交通、石油等众多领域得到应用，成功地解决了生产、生活和科学研究中的大量实际问题，灰色预测模型也由 GM(1, 1)模型扩展到数列预测、区间预测、灾变预测、季节灾变预测、波形预测和系统预测等多种类型，展现出了重要的学术意义和实际应用价值.

经典 GM(1, 1)模型是灰色预测理论的核心模型，邓聚龙教授最早研究了该模型的建模条件，提出了级比检验、光滑比检验和后验误差检验等检验方法，并给出了 GM(1, 1)模型的多种扩展形式及其参数包的求解方法[51-53]. 邓聚龙还从灰色数理资源的角度对 GM(1, 1)模型的含义进行了多角度分析[54-60]. 此后，众多学者参与到灰色预测模型的理论研究并积极对经典方法进行拓展，到目前为止，关于经典 GM(1, 1)模型的研究主要集中在以下六个方面.

第一，对 GM(1, 1)模型的性质进行研究. 邓聚龙建立了 GM(1, 1)模型参数包的求解方法，对只需要 4 个数据就能建立 GM(1, 1)模型进行了理论证明[61]. 李希灿的研究表明原始序列乘以常数 ρ 后，发展系数及误差均不变，但灰色输出 b 及预测值均扩大 ρ 倍[62]. 文献[4]、[63]则分别讨论了 GM(1, N)模型和 GM(0, N)模型的数乘变换影响. 李炳乾研究了原始序列特征与发展系数的关系，结果表明：若原始序列是单调递增序列，则 $a<0$；若原始序列为单调递减序列，则 $a>0$[64]. 刘思峰和邓聚龙对 GM(1, 1)模型的适用范围和模拟精度进行了研究[65]. 邓聚龙对 GM(1, 1)模型的改进型及其参数性质进行了研究[66-69]. 冯利华得到预测值增大的三个充分条件，即计算零点升高、序列的第二项减少、累加次数增多[70]. 但是该文献是采用列表方式得出这些结论，并没有进行严格的数学证明. 针对累加矩阵可能产生的病态矩阵，郑照宁等[71]、党耀国等[72]对灰色预测模型是否产生病态矩阵进行了研究. 发展 GM(1, 1)模型建模新技术，提高灰色预测精度，其关键在于研究 GM(1, 1)模型的性质，找出 GM(1, 1)模型建模的条件[4]，因此以上研究对于灰色预测建模发展具有重要意义.

第二，重点讨论经典 GM(1, 1)模型的初始值选取问题. 由于经典 GM(1, 1)模型的初始值选择为数据序列的第一个数据，而通过理论证明第一个数据与模型的发展系数无关，这不符合灰色系统理论充分利用已有信息的原理，因此，在原有初始值选取的基础上，出现了多种改进方法的讨论，主要有选取序列终点数据为初始值(符合灰色系统理论的不动点原理)[73]，选取最小偏差中间数据为初始值(符合样本综合最优化条件)，以及根据原始序列与模拟序列偏差平方和最小、原始累加序列与模拟累加序列偏差平方和最小与平均相对偏差最小等方法[74-79].

第三，重点研究经典 GM(1, 1)模型的背景值的选取. 通过优化模型的背景值，进而提高模型的模拟和预测精度. 谭冠军[80-82]、张彬和西桂权[83]主要讨论了背景值直

接构造方法，通过对背景值中系数的选取与优化，构造了若干种背景值，在一定程度上提高了模型的预测精度. 以李俊峰为代表的学者主要针对背景值的插值方法展开研究，通过引进经典数学的各种不同的插值方法，从而起到优化背景值的效果[84].

第四，运用缓冲算子提高模型模拟精度. 值得一提的是，刘思峰教授根据灰色系统理论所特有的序列数据生成建模的基本原理，提出缓冲算子概念并构建缓冲算子公理体系，通过对不符合建模的数据序列进行缓冲处理，以取得更好的模拟、预测效果[85]. 宋中民和张曙红提出了平移算子的概念，并将平移算子与 GM(1, 1) 模型结合进行分离建模，提高灰色模型的预测精度[86]. 党耀国和刘思峰[87]、王正新等[88]、谢乃明和刘思峰[89]在原有灰色序列生成的基础上，对缓冲算子进行了扩展研究，构造了大量新型实用的弱化缓冲算子和强化缓冲算子. 许秀莉和罗健运用权系数和影响因子改进现有序列算子，将序列算子应用到人口增长预测数据处理，加入影响因子后用于招生人数预测和农业总产值预测数据处理[90]. 尹春华和顾培亮利用缓冲算子的预测原理，用二阶弱化算子对我国能源消费进行短期预测[91]. 邓聚龙[92-95]、宋中民[96]对累加生成在灰色预测模型方面的应用进行了理论探讨，并构建了累加生成空间.

第五，通过提高原始序列光滑度的方法提高经典 GM(1, 1) 模型的模拟、预测精度. 文献[97]～[103]通过运用对数变换法、幂函数变换法、对数幂函数变换法以及各种复合变换对光滑度的影响，通过提高序列的光滑度以达到提高模拟精度的目的. 其中，李学全和李松仁从理论上分析了拓广灰色系统建模的方法，既分析了改变原始数据列光滑性的变换函数的构造条件[101]，也给出了一类既改进原始数据列的光滑度，同时又缩小逆变换误差的变换函数，提出了选择适当变换的方法. 王子亮对灰色建模理论进行了比较全面的研究，指出提高光滑度并不是提高精度的充分条件[104]. 另外有众多学者采用灰色序列生成的方法对建模的基础数据进行处理，以获得更贴近实际的预测效果[105-108].

第六，研究与传统 GM(1, 1) 模型对应的新模型. 王义闹提出了一种适用于任意时距灰指数单序列的逐步优化直接建模方法，对灰色系统直接建模法进行了深入研究，并对其升、降、凹、凸特性进行了严格证明[109-115]. 在 GM(1, 1) 直接建模方法基础上进一步提出了一种逐步优化方法，弥补了原方法不具有白指数律吻合性的缺陷. 刘孝贤研究了含有负数序列的 GM(1, 1) 模型建模问题，通过数据提升再还原的方法解决了负数序列的建模问题[116]. 陈绵云提出了系统云灰色模型(system cloud grey model，SCGM)，对于模型的特性进行了研究. 系统云模型主要用于研究贫信息、多因子、不确定的错综复杂事物. SCGM$(1, h)$ 模型是以系统云为背景，按基于积分生成变换和趋势关联分析的灰色动态建模原理构造而成的. 部分学者对 SCGM 的应用和扩展进行了研究[117-122]. 宋中民等建立了中心逼近式 GM(1, 1) 模型和反向累加生成模型[123-125]. 谢乃明和刘思峰[126]提出了离散 GM(1, 1) 模型，并对其与原 GM(1, 1) 模型的关系做了深入研究，找出了原模型预测不稳定的原因，利用麦克劳林公式对

这些原因进行了全面解释．文献[127]～[129]对离散灰色拓展模型及基本模型的特性进行了研究，对预测效果进行了实证检验，提出了灰色拓展模型的求解算法．

除对经典 GM(1, 1)模型进行研究外，其他灰色预测模型的研究也得到了一定发展，主要表现在以下三个方面．

第一，对非等间距灰色预测模型的研究．沈建国采用内插法计算出正整数序列中缺失点的数值，然后利用最小二乘法求解灰色微分方程[130]．邓聚龙给出了定义型非等间距模型的求解方法[131, 132]．罗佑新和周继容把非等间距灰色预测模型广泛地应用于机械工程领域[133]．各种插值方法通过在非等间距序列中插入一定数值，从而将非等间距序列变为等间距序列，但是插值过程具有较多人为因素．胡斌直接利用原始数据的有限信息，不生成人为的原始数据，通过得出理想的逼近曲线解决非等间距灰色预测建模问题．但是该方法采用了近似求解方法，参数求解有较大误差，没有考虑到初始值的优化问题．谭冠军通过重构背景值，提出了非等间距灰色建模的新方法[81]．戴文战和李俊峰提出了一种基于积分定义重构 GM(1, 1)模型背景值的方法[134]．王丰效基于灰色模型的指数特性和积分定义，提出了构造多变量非等间距序列的 GM(1, m)模型背景值的方法[135]．构造背景值的方法尽管假定原始序列类似指数增长，但是在求解过程中仍然假定在任意两个间距之间，原始序列是均匀增长．

第二，对含跳跃点的灰色预测模型的研究．对于一般系统而言，如果没有瞬间的外部因素作用，其行为输出量是平稳的，但当系统遇到瞬间外力作用时，其后的输出量可能产生跳跃变化的现象．文献[136]～[138]对跳变序列进行了研究．金义富等研究了序列缺失数据的灰插值问题[139]．

第三，对灰色组合模型的研究．由于 GM(1, 1)模型和离散 GM(1, 1)模型的预测结果是一个等比序列，在预测方面具有较大局限性，很多学者对灰色预测组合模型进行了研究．唐万梅[140]分析了灰色预测方法和支持向量机各自的优缺点，提出了将两者相结合灰色支持向量机预测模型．新模型发挥了灰色预测方法中累加生成的优点，弱化了原始序列中随机扰动因素的影响，增强了数据的规律性，同时避免了灰色预测方法及模型存在的理论缺陷．谢开贵等[141]应用差分格式，研究了基于遗传算法的 GM(1, 1, λ)模型．夏军和赵红英[142]结合“人工神经网络模型”与“信息论”中的编码、译码，提出一种灰色人工神经网络模型，并就其在径流短期预报中的应用进行了初步的探讨．张超等[143]结合灰色预测和马尔可夫链理论的优点，提出了一种灰色马尔可夫 SCGM(1, 1)模型．用单因子系统云灰色 SCGM(1, 1)模型拟合系统的发展变化趋势，并以此为基础进行了马尔可夫预测．其他一些学者也对灰色组合预测模型的应用进行了研究[144-148]．尽管灰色组合模型对于振荡序列的建模进行了有益的探索，但是总体来说，通常采用两种或多种模型的简单加权或依次应用，缺乏创新，理论探讨方面的内容也很少，研究进展缓慢．

1.3　灰色预测模型存在的主要问题

尽管许多学者对经典 GM(1, 1) 模型的完善付出了努力，但是其理论体系仍不够完善，以上各类模型都在一定程度上提高了经典 GM(1, 1) 模型的模拟和预测精度，但是对经典 GM(1, 1) 模型所满足的机理，或者说建模数据所满足的条件缺乏必要理论研究，另外，对于模型模拟、预测效果的检验方法的研究也亟待完善. 目前研究主要存在以下问题.

(1) GM(1, 1) 模型的建模条件不完善，对发展系数的存在区间证明不严格，导致对级比判定区间得出结论不够严密；很多文献对如何提高光滑比进行了研究，认为提高光滑比能够提高建模精度，而实际上提高光滑比并不是提高建模精度的充分条件. 灰色预测建模条件有待进一步研究. 既有的级比、光滑比、后验误差检验等方法均缺乏严密的理论推导；对于序列数据建模的样本数目选取具有随意性，缺乏足够的数理基础.

(2) 建立 GM(1, 1) 模型时通常遇到的问题是选择样本数据越多，模拟误差越大，尽管部分文献发现了这个问题，但只是限于举例说明，没有进行严格的理论证明和分析. 有必要对样本数据选取规则和选取不同样本数据对模拟误差的影响进行研究.

(3) 由于 GM(1, 1) 模型对纯指数序列进行拟合时具有偏差，而离散 GM(1, 1) 模型是 GM(1, 1) 模型的精确形式，有必要对离散 GM(1, 1) 模型的特性进行研究. 尽管离散 GM(1, 1) 模型可以完全拟合指数序列，但是离散 GM(1, 1) 模型与 GM(1, 1) 模型的模拟序列都是等比序列，并且并非离散 GM(1, 1) 模型在任何情况下都比 GM(1, 1) 模型模拟精度高. 两种模型在初始迭代点选择、背景值构造方面都有待进一步研究. 由于实际预测序列多数并非单调序列，需要对离散 GM(1, 1) 模型的各种拓展模型进行研究以适应不同序列的建模需要.

(4) 在对灰色预测模型进行改进时，通常采用最优化理论，应用最小二乘法对初始迭代点进行优化. 在选择最优化目标函数时通常采用模拟序列与原始序列偏差平方和最小或模拟值累加序列与原始序列累加序列偏差平方和最小，而在模型精度检验中通常采用平均相对偏差最小，因此存在优化目标与模型精度检验标准不一致的情况. 关于最优化目标函数和模型精度检验标准的一致性问题有待研究.

(5) 以往文献多关注等间距序列的灰色建模问题，对于非等间距的灰色预测模型研究较少，现有的非等间距灰色模型存在不同程度的问题，因此有待建立新的非等间距灰色预测模型.

(6) 经典 GM(1, 1) 模型存在多种表现形式，其中内涵型 GM(1, 1) 模型是直接由定义型推导而来的，但是以往文献缺乏相关研究. 有关内涵型 GM(1, 1) 模型的特性有待进一步研究.

第 2 章　GM(1, 1) 模型及其特性

GM(1, 1) 模型是灰色预测的核心模型，体现了灰色系统的小样本特性．首先，本章对灰色系统的贫信息和小样本特性给予新的解释，从而扩展了灰色系统理论的适用范围．其次，本章以等比序列为例，研究了 GM(1, 1) 模型的样本数目选择与模拟精度的关系．再次，本章研究了 GM(1, 1) 模型的建模条件．最后，本章研究了新信息 GM(1, 1) 模型的参数性质．

2.1　GM(1, 1) 模型的小样本特性与样本数目选择

建立 GM(1, 1) 模型的前提是假定原始序列近似服从指数分布，然而即使是等比序列，也存在着样本数据越多、模拟精度越差的现象．而建模所采用的级比、光滑比、后验差等检验条件仅仅是必要条件，等比数列完全符合这些条件，但样本数据过多时仍然存在较大偏差．原 GM(1, 1) 模型得到的模拟数据实际上是一个等比数列，很多文献在讨论建模精度时采用纯指数序列作为模型精度的比较实例[149-152]，因此对于纯指数序列的建模特性进行研究就极为重要．

尽管部分文献发现了 GM(1, 1) 模型对于小样本的模拟具有较高精度，样本数目较多时模拟精度较差，但是以往的文献对灰色系统小样本概念的解释存在不足之处．传统文献认为 GM(1, 1) 模型只能处理小样本，但是很多文献将 GM(1, 1) 模型应用于大样本，理论与应用之间存在着矛盾．有必要对小样本概念给予新的解释以扩展灰色系统的应用范围．第一，由于统计学与灰色系统有着本质区别，灰色系统的小样本不同于统计学上的小样本．第二，如何理解灰色系统理论中 GM(1, 1) 模型的小样本特性，当样本量较多时如何应用 GM(1, 1) 模型．第三，对于小样本具有较高精度缺少数学证明，仅仅停留于定性分析．由于 GM(1, 1) 模型是灰色系统的核心模型之一，而“贫信息和小样本”特性又是灰色系统的本质特点，对于小样本数据建模进行严格的理论证明极其重要．无论原始序列是否为等比序列，都可以建立 GM(1, 1) 模型，但是当发展系数一定时，等比序列的模拟精度高于非等比序列的模拟精度．如果原始序列为非等比序列，无法对 GM(1, 1) 模型的小样本特性进行严格的数学证明．由于 GM(1, 1) 模型的模拟值序列是等比序列，本书采用等比序列为例，对 GM(1, 1) 模型的小样本特性进行严格的数学证明．

本节的创新点主要体现在三个方面：首先，提出了灰色系统理论小样本概念

的新解释，从而解决了灰色系统小样本概念在理论与应用中的矛盾问题．其次，首次采用定量分析方法，对于 GM(1, 1) 应用于小样本时存在较高精度给予理论证明．最后，提出了分段修正新信息模型，解决了多周期大样本数据的灰色建模精度问题．

本节对“贫信息、小样本”给予了新的解释．以指数序列为例，研究了指数序列建模的相关特性．从增长率的角度对 GM(1, 1) 模型的小样本性进行了深入的探讨，研究了灰色预测模型模拟数据的增长率特点．探讨了原始序列增长率与模拟序列增长率及发展系数之间的关系，研究了单调递增非负等比序列和单调递减非负等比序列的样本数目选择与模拟平均相对误差的关系．研究了新信息 GM(1, 1) 模型原始数据初始值的特性．基于 GM(1, 1) 模型的小样本特性，给出了分段修正新信息 GM(1, 1) 模型，实例研究表明了该模型的有效性．

2.1.1　GM(1, 1) 模型的小样本特性

定义 2.1.1[1]　设 $X^{(0)}=\{x^{(0)}(1),x^{(0)}(2),\cdots,x^{(0)}(n)\}$，$X^{(1)}=\{x^{(1)}(1),x^{(1)}(2),\cdots,x^{(1)}(n)\}$，$Z^{(1)}=\{z^{(1)}(1),z^{(1)}(2),\cdots,z^{(1)}(n)\}$，其中 $z^{(1)}(k)=0.5(x^{(1)}(k)+x^{(1)}(k-1))$，则称

$$x^{(0)}(k)+az^{(1)}(k)=b$$

为 GM(1, 1) 模型的基本形式．其中，$-a$ 为发展系数；b 为灰色作用量．

GM(1, 1) 模型的时间响应式[1]为

$$\hat{x}^{(1)}(k+1)=(x^{(0)}(1)-\frac{b}{a})\mathrm{e}^{-ak}+\frac{b}{a},\quad k=1,2,\cdots,n \tag{2.1.1}$$

还原值

$$\hat{x}^{(0)}(k+1)=\hat{x}^{(1)}(k+1)-\hat{x}^{(1)}(k)=(1-\mathrm{e}^{a})(x^{(0)}(1)-\frac{b}{a})\mathrm{e}^{-ak},\quad k=1,2,\cdots,n$$

称 $\hat{X}^{(0)}=\{\hat{x}^{(0)}(1),\hat{x}^{(0)}(2),\cdots,\hat{x}^{(0)}(n)\}$ 为 $X^{(0)}$ 的模拟序列，$\hat{x}^{(1)}(k+1)$ 为 $x^{(1)}(k+1)$ 的模拟值．

令 $\hat{u}(k)$ 为模拟序列的增长率，容易证明模拟序列是一个等比数列，模拟序列的增长率是一个定值，即

$$\hat{u}(k)=\frac{\hat{x}^{(0)}(k+1)-\hat{x}^{(0)}(k)}{\hat{x}^{(0)}(k)}=\frac{\hat{x}^{(0)}(k+1)}{\hat{x}^{(0)}(k)}-1=\mathrm{e}^{-a}-1 \tag{2.1.2}$$

由式(2.1.2)可以看出，GM(1, 1) 模型的模拟序列增长率是一个定值．增长率序列为单调递增序列或单调递减序列．随着时间的推移，系统会受到随机扰动因素的影响，原始序列很难具备增长率为常数这个条件．小样本更容易具备增长率相对稳定的条件，这也反映了灰色系统的“贫信息、小样本”特点．

一种新的理论的提出要求该理论能够解决其他理论不能解决的问题，至少该

理论在解决某类问题时能够比其他学科解决得更好．“灰性”的基本含义是“不完全信息”．文献[1]、[2]指出系统信息不完全的情况分为以下四种：①元素(参数)信息不完全；②结构信息不完全；③边界信息不完全；④运行行为信息不完全．

新学科的建立要求该学科能够解决其他学科不能解决的问题，至少也要达到在解决同一个问题时能够比其他学科解决得更好．灰色系统最显著的特色就是以“部分信息已知，部分信息未知”的“小样本、贫信息”不确定性系统为研究对象[1, 2]．在统计学中，“小样本”具有明确的含义．令 n 为样本数目，若 $n \geqslant 30$，则样本为大样本；若 $n < 30$，则样本为小样本[153]．在灰色系统中，不同的模型对小样本有不同的要求，以 GM(1, 1) 模型为例，通常仅需要 4～10 个样本数据建模，样本数据较多时反而偏差过大，因此需要对“小样本、贫信息”给予新的解释．“小样本、贫信息”包含以下五种含义．

(1) 原始序列实际数据少．如改革开放初期的农业经济数据，一般只有 10 个以下．

(2) 原始序列数据量很多，但是如果对原始数据按照经济现象的周期进行分类，则每一周期的数据少，而反映最新周期的数据通常更少．如中国的外汇储备数据，尽管有很多数据，但是如果应用该数据进行预测建模，根本不可能选用全部数据，需要对原始数据按照周期进行分段，在每一个周期内的样本数据就成了“贫信息、小样本”数据．

(3) 客观上经济现象发展的不确定造成“贫信息、小样本”，反映新趋势的数据少．如中国外汇储备 1994 年以来增长较快，2004 年增长率达到 51.25%，2005～2006 年尽管年增长率有所下降，但每年的增长量仍然超过 2000 亿美元．很明显这种增长不可能长久持续下去，但将来如何发展、拐点在哪年出现不清楚，由于客观上反映外汇年增长率剧增的数据量少，这也是一种“贫信息、小样本”．

(4) 专家主观上能够比较准确预期的数据少，信息不充分．依据专家的经验，可以对未来经济现象的发展做出比较合理的估计，从而得到部分数据，但能够估计的数据量通常较少，属于“贫信息、小样本”．专家根据国内外经济发展状况，对中国未来几年的外汇储备可能做出相对比较准确的预期，但是这种预期数据量通常较少，属于一种“贫信息、小样本”．

(5) 尽管实际样本量较多，但相对于建模需要数据较少，数据量属于相对缺乏．此处建模泛指灰色系统的各种建模．

通过以上分析可以看出，将灰色系统主要定位于“贫信息、小样本”的研究，可以解决统计学等其他要求大样本的学科不能解决的问题．下面以 GM(1, 1) 模型为例，从理论上探讨该模型在大样本下通常产生较大偏差的原因．由于灰色系统与统计学具有较大差别，统计学中的精度检验方法也不同于灰色系统的精度检验方法，本书只采用灰色系统理论中被广泛接受的检验方法，采用平均相对误差来

检验模型的精度．由于本书的目的是证明 GM(1, 1)模型的小样本特性，应用平均相对误差来检验模型的精度．

2.1.2　样本选择与模拟精度

定理 2.1.1[154]　令 $X^{(0)}(k)=\{1,\lambda,\cdots,\lambda^{n-1}\}$，其中 $\lambda>0$ 且 $\lambda\neq1$，则建立 GM(1, 1) 模型得到的发展系数与 n 无关．

证明　令 $Y=\begin{bmatrix}\lambda\\ \lambda^2\\ \vdots\\ \lambda^{n-1}\end{bmatrix}$，$B=\begin{bmatrix}-\dfrac{2-\lambda^2-\lambda^3}{2(1-\lambda)} & 1\\ -\dfrac{2-\lambda^3-\lambda^4}{2(1-\lambda)} & 1\\ \vdots & \vdots\\ -\dfrac{2-\lambda^{(n-1)}-\lambda^n}{2(1-\lambda)} & 1\end{bmatrix}$，$B^{\mathrm T}B=\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}$，其中

$$a_{11}=\frac{4(n-1)(1-\lambda)+(1+\lambda)(\lambda-\lambda^n)(\lambda+\lambda^n-4)}{4(1-\lambda)^3}$$

$$a_{12}=a_{21}=-\frac{2(n-1)(1-\lambda)-(1+\lambda)(\lambda-\lambda^n)}{2(1-\lambda)^2},\qquad a_{22}=n-1$$

$$(B^{\mathrm T}B)^{-1}=\frac{1}{\left|B^{\mathrm T}B\right|}\begin{bmatrix}a_{22} & -a_{21}\\ -a_{12} & a_{11}\end{bmatrix}$$

$$\left|B^{\mathrm T}B\right|=\frac{(1+\lambda)(\lambda-\lambda^n)[(n-2)\lambda-n\lambda^2+n\lambda^n-(n-2)\lambda^{n+1}]}{4(1-\lambda)^4}$$

则

$$\begin{bmatrix}\hat a\\ \hat b\end{bmatrix}=(B^{\mathrm T}B)^{-1}B^{\mathrm T}Y=\frac{1}{\left|B^{\mathrm T}B\right|}\begin{bmatrix}b_{11} & b_{12} & \dots & b_{1(n-1)}\\ b_{21} & b_{22} & \dots & b_{2(n-1)}\end{bmatrix}Y$$

其中，

$$b_{1k}=\frac{(1+\lambda)[\lambda^n-\lambda+\lambda^k(n-1)(1-\lambda)]}{2(1-\lambda)^2},\quad k=1,2,\cdots,n-1$$

$$b_{2k}=\frac{2(1+\lambda)(\lambda^n-\lambda)-2(n-1)(1+\lambda)(1-\lambda)+(\lambda-\lambda^n)(\lambda+\lambda^n-\lambda^k-\lambda^{k+1})}{4(1-\lambda)^3},\quad k=1,2,\cdots,n-1$$