

A TEXTBOOK SERIES IN
ECONOMICS AND MANAGEMENT



现代经济学管理学教科书系列



高等时间序列 经济计量学

*Advanced Time
Series Econometrics*

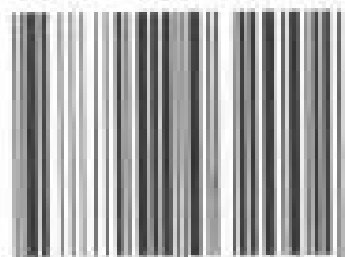


陆懋祖 著



上海人民出版社

ISBN 7-208-03200-9



9 787208 032002 >

定价 20.00 元

图书在版编目(CIP)数据

高等时间序列经济计量学/陆懋祖著. —上海:上海人民出版社,1999

(现代经济学管理学教科书系列/易刚,海闻主编)

ISBN 7-208-03200-9

I. 高… II. 陆… III. 时间序列分析-计量经济学 IV. F224.0

中国版本图书馆 CIP 数据核字(1999)第 17958 号

责任编辑 张 启

封面装帧 桑吉芳

·现代经济学管理学教科书系列·

高等时间序列经济计量学

陆懋祖 著

上海人民出版社出版、发行

(上海绍兴路 54 号 邮政编码 200020)

新华书店上海发行所经销

商务印书馆上海印刷股份有限公司印刷

开本 850×1168 1/32 印张 11.5 插页 4 字数 257,000

1999 年 8 月第 1 版 1999 年 8 月第 1 次印刷

印数 1—5,100

ISBN 7-208-03200-9/F·636

定价 20.00 元

总 前 言

20 世纪与 21 世纪之交,世界在注视着中国。经过近 20 年的经济改革,中国社会主义市场经济新体制的雏形在世界的东方渐渐清晰。

中国改革的目标是建立社会主义市场经济体系,这一历史使命的完成需要千千万万掌握市场经济运行规则的企业家、经济学家、管理专业人才和国家干部。中国需要人才,培养人才靠教育,高水平的教育迫切需要有一套与国际接轨的一流教材,尤其是大学本科生的教科书。应市场之求,受同仁委托,我们组织编写了这一套“现代经济学管理学教科书系列”。

本套教科书的适用对象是各大专院校经济管理专业的本科生和研究生,以及具有同等文化程度的自学者。

实事求是精神是本套教科书的指导思想。中国正在摸索一条有自己民族特色的市场经济之路,许多问题还有待于进一步的探讨。本套教科书对有争议问题的取舍标准是邓小平同志提出的“三个有利于”的原则:即主要看是否有利于发展社会主义社会的生产力,是否有利于增强社会主义国家的综合国力,是否有利于提高人民的生活水平。

本套教科书有如下的特点:第一,涵盖了现代经济学和管理学的主要领域和许多前沿专题,力图准确、全面、系统地阐述每个学科的基本内容,努力做到现代经济学、管理学教学研究的主流框架与我国经济改革具体实践的有机结合。第二,最大限度地

方便读者。各教科书均在前言或第1章中告诉读者该书的层次(本科低年级、本科高年级、研究生)。教科书的语言简明通俗,结构科学严谨,适合于教学与自学。各章附有小结与习题,全书后有按汉语拼音顺序排列的索引,并附有便于查找的参考书目。第三,尽可能地为教员提供方便。教科书的前言部分有一个教学大纲并附有两套教案,以适合本科或专科等各种教学和不同学时需要。有些教科书配有《教学参考书》并附有题库及部分答案。

本套教科书的作者们不仅受过现代经济学、管理学的系统训练,而且都从事教学科研工作多年,其中有不少已是国内国外的一流专家教授。他们既熟练掌握现代经济学、管理学理论,又对中国和发达国家经济的实际运行情况有深刻的了解,从而使这套教科书能结合中国国情来系统地、准确地介绍经济、管理理论与案例。每本教科书都经过两次评审(一次对写作计划,一次对书的初稿),作者们均按评审意见对书稿进行了认真的修改。

亲爱的读者们,亲爱的同事们,我们将此套教科书献给你们,请评头论足,不吝赐教。

易纲 海闻

前 言

在过去的十几年中,经济计量学在其各个领域内都取得了重要的进展。这不仅使经济计量学本身成为一门日趋重要的学科,也使得它在现代经济学和现代金融学中起到越来越重要的作用。如今很多研究经济学或经济计量学的人称自己为宏观经济计量学家、微观经济计量学家和金融经济计量学家等,可见经济计量学与现代经济学已不可分割。

本书旨在介绍过去十几年中经济计量学在时间序列领域里的发展,重点讨论非稳定的单位根过程(unit root process)、同积过程和同积系统(cointegrated system)的一些主要理论。这些理论在80年代初兴起后,在很大程度上改变了传统的时间序列经济计量学的理论和方法。稳定过程不再是经济计量学研究的唯一对象,非稳定的时间序列也不再是不可涉足的领域,特别是其中的 $I(1)$ 过程和 $I(2)$ 过程与同积过程成为了研究的主要对象,它们已在经济学和金融学中得到了广泛的应用。经济计量学在这一方面的发展也引起了数学方法上的突破,维纳过程(Wiener process)和基于其上的泛函中心极限定理(functional central limit theorem)成为了研究的主要数学工具。这使得我们能在更为广泛的条件下研究时间序列统计量的极限和极限分布问题。

当然,要在一本书中全面介绍这些内容是有很困难。尽管这些理论和方法只有十几年的发展历史,它们已包含了如此

丰富的内容,以至于在一本书中对这些理论作详细介绍似乎不大可能。不仅如此,这浩瀚的文献中并存着不同的体系和观点,它们对整个学科的发展起着相辅相成的作用,但却很难在一本书中全面反映这些观点。

本书以介绍单位根过程、同积过程和 ARCH 过程的基础的数学结构作为主要目的,旨在帮助经济系或统计系的本科高年级学生、研究生及研究人员掌握现代化的经济计量学工具,使他们在理论上受到一定的训练,以便今后在此领域中作出自己的贡献。这也是本书称为“高等”的原因,以强调理论推导的严格性。书中对介绍的主要定理都作了较为详细的数学证明,理解这些证明对进一步掌握理论是很有帮助的。当然,初读者可将定理证明留到以后必要时再读。

本书的第 1 章至第 3 章介绍了单位根过程的结构和主要特征,以及对单位根过程的参数估计和假设检验。第 4 章至第 6 章介绍了同积过程和同积系统的结构、主要特征和表示形式,以及参数估计和假设检验的方法。第 7 章介绍 ARCH 模型和它的几种重要的衍生模型,以及它们在现代经济学、金融学中的应用。

本书的很大一部分是基于作者在南安普顿大学经济系讲授研究生的经济计量学专题课时的讲义。在这些讲义和本书的写作过程中,作者参考了 Banajee et al. (1994)、Engle (1995)、Hendry (1995)、Hamilton (1994)、Johansen (1995) 和 Rozanov (1967) 等的著作。读者如对本书中介绍的内容有进一步的兴趣,可在这些著作中找到有价值的参考资料。

本书可供经济系、统计系或其他相应专业的高年级学生和研究生作同名课程的教材或参考书。作为教材,本书适合作一学期用。

在本书的写作过程中,作者得到了易纲博士和秦朵博士的

不断鼓励和支持,在此谨表示感谢。

陆懋祖

1998 年于英国南安普顿大学经济系

E-mail:ml@soton.ac.uk

主编：易 纲 海 闻

编委（按姓氏笔画）：

王一江	王 燕	邓胜梁	田国强
卢 锋	李庆云	宋国青	余明德
张维迎	张 帆	林毅夫	欧阳明
易 纲	周其仁	周惠中	俞 乔
胡大源	秦 朵	海 闻	梁 能

组织：北京大学中国经济研究中心

目 录

总前言	1
前言	1
第 1 章 单位根过程	1
1.1 简介	1
1.2 单位根过程的定义	7
1.3 维纳过程	10
1.4 泛函中心极限定理	13
1.5 连续映照定理	16
1.6 有关随机游动的极限分布	19
1.6.1 几个重要的极限	19
1.6.2 随机积分	25
1.6.3 最小二乘估计 $\hat{\rho}_T$ 的极限分布	28
1.6.4 t_T 统计量的极限分布	30
1.7 带常数项的随机游动	31
1.8 有关单位根过程的极限分布	37
1.9 时间序列的去势	47
1.10 近单位根过程	52
1.11 本章小结	55
习题	55

第 2 章 单位根过程的假设检验	57
2.1 简介	57
2.2 迪基—福勒(DF)检验法	59
2.2.1 情况一和情况二的 DF 检验法	60
2.2.2 情况三和情况四的 DF 检验法	69
2.2.3 迪基—福勒(DF)检验法小结	76
2.3 菲利普斯—配荣(PP)检验法	80
2.3.1 情况二的 PP 检验法	80
2.3.2 情况一和情况四的 PP 检验法	92
2.4 增广的迪基—福勒(ADF)检验法	94
2.4.1 P 阶自回归过程	94
2.4.2 情况二的 ADF 检验法	97
2.4.3 情况一和情况四的 ADF 检验法	110
2.5 本章小结	112
习题	113
第 3 章 多变量单位根过程	115
3.1 简介	115
3.2 多变量单位根过程的极限定理	115
3.3 含单位根的向量自回归过程	128
3.3.1 VAR(P) 的表示形式	129
3.3.2 不带常数项的 VAR(P)	130
3.3.3 带常数项的 VAR(P)	137
3.4 伪回归	143
3.5 伪回归的纠正方法	155
3.6 本章小结	156
习题	157
第 4 章 同积过程的性质和表示形式	159

4.1	简介	159
4.2	同积过程的主要特征	161
4.3	同积过程的表示形式	165
4.3.1	误差修正形式	165
4.3.2	三角表示形式	167
4.3.3	同趋势表示形式	169
4.4	本章小结	171
	习题	171

第5章 同积过程的参数估计和假设检验

	——最小二乘方法	173
5.1	同积向量的最小二乘估计	173
5.2	同积关系的规范化	179
5.3	多个同积向量	180
5.4	检验随机向量的同积性	185
5.5	同积向量的假设检验	209
5.6	对 u_{1t} 和 u_{2t} 的相关系数的纠正	218
5.7	充分改进的最小二乘估计	223
5.8	本章小结	230
	习题	231

第6章 同积系统的最大似然方法

6.1	简介	233
6.2	同积过程与误差修正过程	234
6.3	典型相关	240
6.4	ECM 和集中的对数似然函数	247
6.5	最大似然估计和典型相关分析	251
6.6	参数矩阵 π 的最大似然估计	254

6.7	同积关系的假设检验	260
6.8	对同积向量的假设检验	266
6.9	对矩阵 α 的假设检验	270
6.10	本章小结.....	275
	习题.....	276
第7章	ARCH 模型	278
7.1	简介	278
7.2	自回归条件异方差过程(ARCH)的定义	280
7.3	ARCH 过程参数的最大似然估计	287
7.4	非正态的 ARCH(q)过程的最大似然估计	297
7.5	ARCH 模型的假设检验	299
7.6	广义的 ARCH 模型——GARCH 模型	302
7.6.1	GARCH(1,1)过程	306
7.6.2	GARCH 回归模型的参数估计	308
7.6.3	GARCH 模型的假设检验	311
7.7	ARCH 模型的其他推广形式	312
7.7.1	指数的 GARCH 模型(EGARCH)	312
7.7.2	向量的 GARCH 模型	315
7.7.3	ARCH-M 模型	316
7.8	ARCH 模型和随机微分方程	318
7.9	ARCH 过程的综合	321
7.10	本章小结.....	331
	习题.....	331
	参考文献.....	333
	附表.....	336
	索引.....	343

Contents

Preface	1
Chapter 1 Unit Root Processes	1
1.1 Introduction	1
1.2 Definitions of unit root processes	7
1.3 Wiener process	10
1.4 Functional central limit theorem	13
1.5 Continuous mapping theorem	16
1.6 Limiting distributions for random walk processes	19
1.6.1 Some important limits	19
1.6.2 Stochastic integration	25
1.6.3 Limiting distribution of OLS estimator $\hat{\rho}_T$	28
1.6.4 Limiting distribution of t_T statistic	30
1.7 Random walk with constant	31
1.8 Limiting distributions for unit root processes	37
1.9 De-trending of time series	47
1.10 Near unit root processes	52
1.11 Summary	55
Exercises	55

Chapter 2 Hypothesis Testing for Unit Root

Processes	57
2.1 Introduction	57
2.2 Dickey-Fuller (DF) tests	59
2.2.1 DF tests for case I and case II	60
2.2.2 DF tests for case III and case IV	69
2.2.3 A summary of Dickey-Fuller tests	76
2.3 Phillips-Perron (PP) tests	80
2.3.1 PP test for case II	80
2.3.2 PP tests for case I and case IV	92
2.4 Augmented Dickey-Fuller (ADF) tests	94
2.4.1 Autoregressive process of order P	94
2.4.2 ADF test for case II	97
2.4.3 ADF tests for case I and case IV	110
2.5 Summary	112
Exercises	113

Chapter 3 Multivariate Unit Root Processes

3.1 Introduction	115
3.2 Limiting theorems for multivariate unit root processes	115
3.3 Vector autoregressive processes with unit roots	128
3.3.1 Representation of VAR (P) processes	129
3.3.2 VAR (P) processes with no constant	130
3.3.3 VAR (P) processes with constant	137
3.4 Spurious regression	143
3.5 Correction of spurious regression	155
3.6 Summary	156

Exercises	157
Chapter 4 Characteristics and Representation of	
Cointegrated Processes	159
4.1 Introduction	159
4.2 The main characteristics of cointegrated processes	161
4.3 Representation of cointegrated processes	165
4.3.1 Error correction representation	165
4.3.2 Triangular representation	167
4.3.3 Common trend representation	169
4.4 Summary	171
Exercises	171
Chapter 5 Parameter Estimation and Hypothesis Testing	
for Cointegrated Processes; OLS Method	173
5.1 OLS estimation of cointegrating vector	173
5.2 Normalization of cointegration relationship ...	179
5.3 More than one cointegrating vectors	180
5.4 Testing for cointegration	185
5.5 Hypothesis testing for cointegrating vectors	209
5.6 Correcting the correlation between u_{1t} and u_{2t}	218
5.7 Fully modified OLS estimation	223
5.8 Summary	230
Exercises	231

Chapter 6	Maximum Likelihood Estimation for Cointegrated System	233
6.1	Introduction	233
6.2	Cointegrated processes and error correction models	234
6.3	Canonical correlation analysis	240
6.4	ECM and concentrated log-likelihood function	247
6.5	Maximum likelihood estimation and canonical analysis	251
6.6	Maximum likelihood estimation of parameter matrix π	254
6.7	Testing for cointegration	260
6.8	Hypothesis testing for cointegrating vectors	266
6.9	Hypothesis testing for matrix α	270
6.10	Summary	275
	Exercises	276
Chapter 7	ARCH Models	278
7.1	Introduction	278
7.2	Definitions of autoregressive conditional heteroskedasticity (ARCH) processes	280
7.3	Maximum likelihood estimation of ARCH processes	287
7.4	Maximum likelihood estimation of non-Gaussian ARCH (q) processes	297
7.5	Hypothesis testing for ARCH models	299

7.6	Generalized ARCH models; GARCH	
	models	302
7.6.1	GARCH (1,1) process	306
7.6.2	Parameter estimation of GARCH regression	
	models	308
7.6.3	Hypothesis testing for GARCH models	311
7.7	Other generalized ARCH models	312
7.7.1	Exponential GARCH (EGARCH) model	312
7.7.2	Vector GARCH model	315
7.7.3	ARCH-M model	316
7.8	ARCH model and stochastic differential	
	equation	318
7.9	Aggregation of ARCH processes	321
7.10	Summary	331
	Exercises	331
	References	333
	Statistical tables	336
	Index of subjects	343

第 1 章 单位根过程

1.1 简介

现代时间序列经济计量学的一个重要的研究课题,是探索经济时间序列数的动态结构,研究它们的统计性质,理解产生这些经济数据的数据生成过程(data generating process)的特点和性质,从而能更有效地利用经济数据构造和建立经济计量模型,用以作经济预测,检验各种经济理论的可靠性和可行性,并为各级政府和企业的经济决策提供数量化的建议。

传统的时间序列经济计量学在进行这些研究时,通常假设经济数据和产生这些数据的随机过程是稳定的过程(stationary process),在此基础上对经济计量模型中的参数作估计和假设检验。稳定过程是数理统计和随机过程理论中最常研究的一种过程,已有很多成熟的研究结果。图 1.1 和图 1.2 给出了两个稳定过程的图像。这两个过程的一个共同的特点是都有一固定的均值,且在每一时刻对均值的偏离基本相同。以下给出稳定过程的定义。

定义 1.1: 单变量的稳定过程

随机过程 $\{x_t, t = 1, 2, \dots\}$, 其中 x_t 为一随机变量,是矩稳定过程(moment stationary process),简称稳定过程,若

(a) 在每一时刻 $t, t = 1, 2, \dots, x_t$ 的期望为一常数,

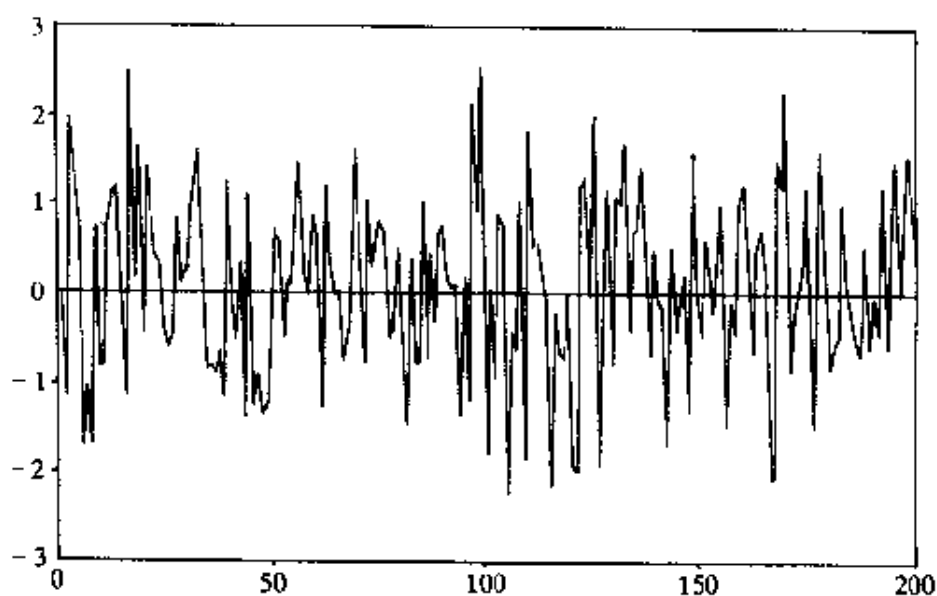


图 1.1 $y_t = \varepsilon_t; \varepsilon_t \sim \text{iid}(0, \sigma^2)$

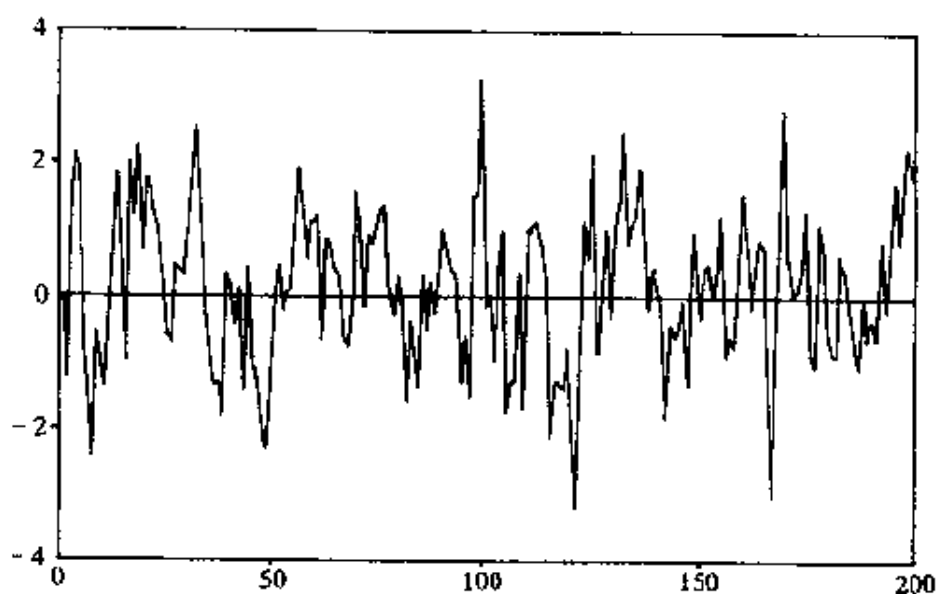


图 1.2 $y_t = 0.5y_{t-1} + \varepsilon_t; \varepsilon_t \sim \text{iid}(0, \sigma^2)$

$$E(x_t) = \mu < \infty;$$

(b) $x_t (t = 1, 2, \dots)$ 的协方差 $\text{cov}(x_t, x_j)$, 只与随机变量 x_t, x_j 在过程中的间隔 $t - j$ 有关, 而与它们的具体位置无关, 即

$$\text{cov}(x_t, x_{t-j}) = E\{(x_t - \mu)(x_{t-j} - \mu)\} = \mu_j < \infty \quad (1.1.1)$$

$\sqrt{T}(\hat{\rho} - \rho)$ 有正态的极限分布。

再如,在简单的两个变量的回归模型

$$y_t = \alpha + \beta x_t + \varepsilon_t; \quad t = 1, 2, \dots, T$$

中,若 $\{x_t\}$ 为稳定过程, $\{\varepsilon_t\}$ 为独立同分布, $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = \sigma^2$, 而且 x_t 与 ε_t 不相关,即 $\text{cov}(x_t, \varepsilon_t) = 0$, 那么 $\{y_t\}$ 也是稳定过程,未知参数 α 、 β 和 σ^2 可由最小二乘法估计,而且都是一致估计量。

但是,许多经济指标的时间序列数据并不具有稳定过程的特征。比如,图 1.3 中英国的 M1 货币的供应量和物价指数的年时间序列数据都不会由稳定的随机过程生成,因为它们显然不具有固定的期望值。

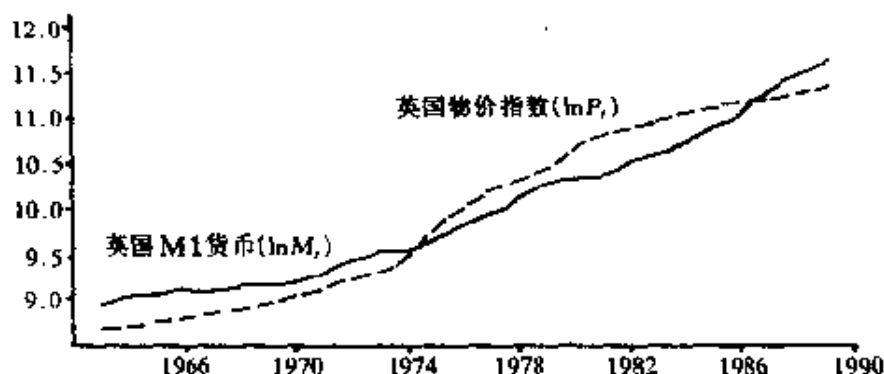


图 1.3 英国 M1 货币供应量和物价指数的年度数据

对于由非稳定过程 (non-stationary process) 生成的时间序列数据,传统的数理统计和经济计量学方法显得无能为力,特别是这时传统的中心极限定理不再适用。

值得注意的是,如对图 1.3 中的两个时间序列取一阶差分: $\Delta \ln M_t = \ln M_t - \ln M_{t-1}$, $\Delta \ln P_t = \ln P_t - \ln P_{t-1}$ 。其图像(图 1.4)表现出稳定过程的特征。

若随机过程 $\{x_t\}$ 的一阶差分过程 $\{\Delta x_t = x_t - x_{t-1}\}$ 为一稳定过程,则称 $\{x_t\}$ 服从单位根过程 (unit root process)。特别地,

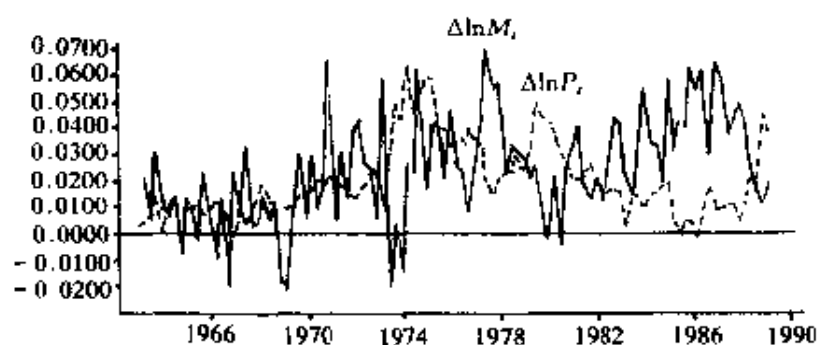


图 1.4 $\Delta \ln M_t$ 和 $\Delta \ln P_t$ 的图像

若 $\{\epsilon_t\}$ 为独立同分布, 且 $E(\epsilon_t) = 0$, $D(\epsilon_t) = \sigma^2$, 可以 ϵ_t 构造一随机游动过程 (random walk process) $\{y_t\}$:

$$y_t = y_{t-1} + \epsilon_t$$

随机游动过程是单位根过程的一个特例。

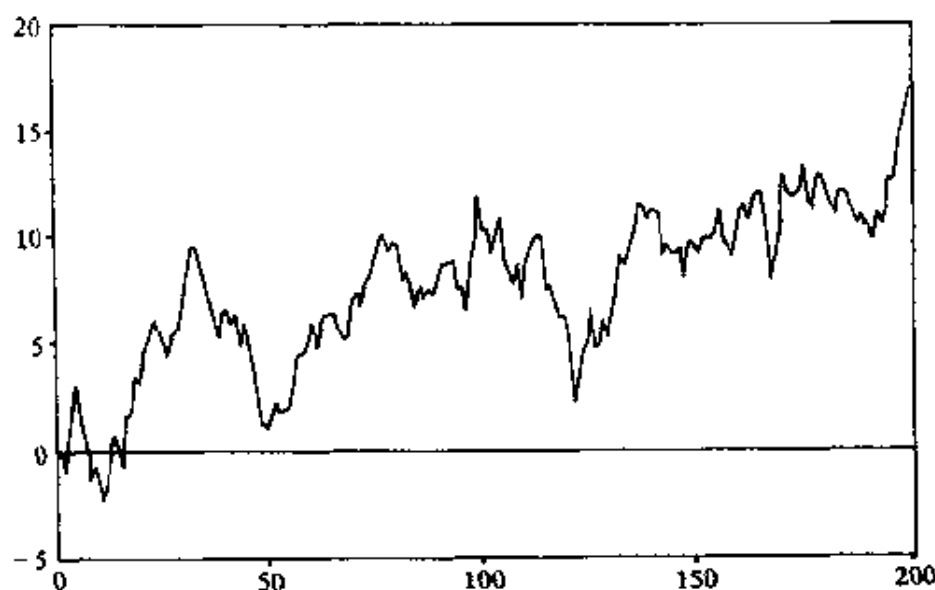


图 1.5 随机游动 $y_t = y_{t-1} + \epsilon_t$

单位根过程是最常见的非稳定过程之一, 由于它在现代金融学、宏观经济学的理论和实践中的广泛应用, 对单位根过程的研究成为当今经济计量学的主要课题之一, 特别自 80 年代以来, 出现了许多理论上和实践上的重大突破。这就使研究人员能

有效地处理以前不能处理的数据。以金融市场中的股票价格为例, 设 P_t 为某一股票在某一时刻 t 的价格, 根据金融学中有效市场 (efficient market) 的假设, 在时刻 $t+1$ 的股价 P_{t+1} 可由一单位根过程描述:

$$P_{t+1} = \rho P_t + u_{t+1} \quad (1.1.2)$$

这里, $\rho = 1$, $\{u_t\}$ 为独立同分布, 且 $E(u_t) = 0$, $D(u_t) = \sigma^2 < \infty$ 。可以将 (1.1.2) 中的过程看作是 AR(1) 过程的一个特例, 参数 ρ 和 σ^2 的最小二乘估计有非标准的分布。对 (1.1.2) 不断作迭代, 且令 $\rho = 1$, 则有

$$\begin{aligned} P_{t+1} &= P_t + u_{t+1} \\ &= P_{t-1} + u_t + u_{t+1} \\ &= u_1 + u_2 + \cdots + u_t + u_{t+1} \end{aligned}$$

所以 $D(P_{t+1}) = D(u_1 + \cdots + u_{t+1}) = (t+1)\sigma^2$ 。当 $t \rightarrow \infty$ 时, P_t 的方差趋于无穷大, 传统的中心极限定理在此是不适用的。

在两个变量的回归模型 $y_t = \alpha + \beta x_t + \varepsilon_t$ 中, 若 $\{x_t\}$ 是一单位根过程, 未知参数 α 和 β 的最小二乘估计量有非标准的分布, 因为这时传统的中心极限定理已不适用。一个折衷的方法是先对 x_t 和 y_t 取一阶差分:

$$\Delta y_t = y_t - y_{t-1}; \quad \Delta x_t = x_t - x_{t-1}$$

因为 $\{y_t\}$ 和 $\{x_t\}$ 都为单位根过程, 所以 $\{\Delta y_t\}$ 和 $\{\Delta x_t\}$ 都为稳定过程, 对它们可以应用中心极限定理。然后用 Δy_t 和 Δx_t 构造新的回归模型:

$$\Delta y_t = a + b \Delta x_t + v_t$$

其中的未知参数 a 和 b 的最小二乘估计是一致的, 并有正态的

极限分布。虽然这种方法克服了单位根过程的困难,在统计意义上有效,但由于水平数据 y_t 和 x_t 往往具有重要的经济意义,以 Δy_t 和 Δx_t 建立的模型不能对 y_t 和 x_t 之间的关系作充分的描述,从而达不到检验经济理论、进行经济预测的目的。

本章的以下几节将介绍有关单位根过程的一些基本概念和运算方法,着重讨论适合于单位根过程的中心极限定理。在此基础上建立最小二乘估计的极限分布。

1.2 单位根过程的定义

先定义随机游动过程。

定义 1.2: 随机游动过程

随机过程 $\{y_t, t = 1, 2, \dots\}$ 称为随机游动过程,若

$$y_t = y_{t-1} + \epsilon_t; \quad t = 1, 2, \dots$$

其中, $\{\epsilon_t\}$ 为独立同分布,且 $E(\epsilon_t) = 0$, $D(\epsilon_t) = E(\epsilon_t^2) = \sigma^2 < \infty$ 。

随机游动是一非稳定过程。虽然 y_t 的期望

$$E(y_t) = E(y_0 + \epsilon_1 + \epsilon_2 + \dots + \epsilon_t) = y_0$$

为一常数,但它的方差

$$D(y_t) = E(y_t - y_0)^2 = E(\epsilon_1 + \dots + \epsilon_t)^2 = t\sigma^2$$

是时间 t 的函数,而且随 t 发散到无穷大。

较随机游动更一般的,是单位根过程。

定义 1.3: 单位根过程

随机过程 $\{y_t, t = 1, 2, \dots\}$ 是一单位根过程,若

$$y_t = \rho y_{t-1} + u_t; \quad t = 1, 2, \dots \quad (1.2.1)$$

其中, $\rho = 1$, $\{u_t\}$ 为一稳定过程, 且 $E(u_t) = 0$, $\text{cov}(u_t, u_{t-s}) = \mu_s < \infty$, 这里 $s = 0, 1, 2, \dots$ 。

显然, 随机游动是单位根过程的一个特例。单位根过程中的随机干扰项 u_t 只需服从一般的稳定过程。 $\{\epsilon_t\}$ 和 $\{u_t\}$ 这种假设上的差异使它们在现代经济学和金融学上有不同的应用。当然, 从统计学的角度, 单位根过程在技术处理上更为困难。

将(1.2.1)式改写成以下形式

$$(1 - \rho L)y_t = u_t; \quad t = 1, 2, \dots$$

其中, L 为滞后算子, 使得 $Ly_t = y_{t-1}$ 。 $(1 - \rho L)$ 称为滞后多项式, 它的特征方程

$$1 - \rho z = 0 \quad (1.2.2)$$

有根 $\frac{1}{\rho}$ 。当 $\rho = 1$ 时, (1.2.2) 有一单位根。这就是称呼“单位根过程”的来历。分别以 $I(1)$ 和 $I(0)$ 表示单位根过程和稳定过程, 可将 y_t 和 Δy_t 记为:

$$y_t \sim I(1); \quad \Delta y_t \sim I(0)$$

虽然随机游动和单位根过程在定义上有所不同, 为叙述的简便起见, 在不引起混淆的情况下, 以下统称为单位根过程。

为进一步理解单位根过程和稳定过程的本质区别, 考虑一阶自回归过程

$$y_t = \rho y_{t-1} + \epsilon_t; \quad t = 1, 2, \dots \quad (1.2.3)$$

中参数 ρ 的最小二乘估计。若这时 $\{\epsilon_t\}$ 独立同分布, 并有 $E(\epsilon_t) = 0$, $D(\epsilon_t) = \sigma^2 < \infty$, 利用样本 y_1, \dots, y_T 构造 ρ 的最小二乘估计:

$$\hat{\rho}_T = \frac{\sum_{i=1}^T y_{i-1} y_i}{\sum_{i=1}^T y_{i-1}^2}$$

将(1.2.3)代入

$$\hat{\rho}_T = \rho + \frac{\sum_{i=1}^T y_{i-1} \epsilon_i}{\sum_{i=1}^T y_{i-1}^2} \quad (1.2.4)$$

当 $|\rho| < 1$ 时, $\{y_t\}$ 为一稳定过程。因为 $\{\epsilon_t\}$ 独立同分布, 所以滞后值 y_{i-1} 与 ϵ_i 不相关, 即 $\text{cov}(y_{i-1}, \epsilon_i) = 0$ 。根据大数定律, 当 $T \rightarrow \infty$ 时, $\hat{\rho}_T$ 以概率收敛于参数 ρ , 所以最小二乘估计是一致的估计。再根据中心极限定理, $\sqrt{T}(\hat{\rho}_T - \rho)$ 有正态的极限分布

$$\sqrt{T}(\hat{\rho}_T - \rho) \xrightarrow{d} N(0, \sigma^2(1 - \rho^2)) \quad (1.2.5)$$

这里, “ \xrightarrow{d} ” 表示以分布收敛。在上式中不难看出, 当 $\{y_t\}$ 为一稳定过程, $|\rho| < 1$ 时, 方差 $\sigma^2(1 - \rho^2)$ 为一大于零的正数, $\sqrt{T}(\hat{\rho}_T - \rho)$ 的极限分布是有定义的。但当 $\rho \rightarrow 1$ 时, $\{y_t\}$ 趋于一单位根过程, 方差 $\sigma^2(1 - \rho^2)$ 相应地趋于零, $\sqrt{T}(\hat{\rho}_T - \rho)$ 的极限分布趋于一退化的分布。这表明, 最小二乘估计 $\sqrt{T}(\hat{\rho}_T - \rho)$ 的极限分布在 $\rho = 1$ 这点上发生了质的变化, 传统的稳定过程理论和中心极限定理这时已无能为力, 需要有新的理论和工具。

研究单位根过程的一个有力的工具是维纳过程 (Wiener process) 理论, 和在此基础上建立起来的泛函中心极限定理 (functional central limit theorem)。正是在这些理论基础之上, 过去十几年中在单位根过程的研究中出现了许多重要的成果,

从而使我们能理解 $\sqrt{T}(\hat{\rho}_T - \rho)$ 在 $\rho = 1$ 时的极限分布。

1.3 维纳过程

维纳过程也称为布朗运动(Brownian motion),在现代随机过程理论中起重要作用。

定义 1.4: 标准维纳过程

标准维纳过程 $\{W(t), t \in [0, 1]\}$ 是定义在闭区间 $[0, 1]$ 上的连续变化的单变量的随机过程,满足以下条件:

(a) $W(0) = 0$;

(b) 对闭区间 $[0, 1]$ 上任何一组有限分割 $0 \leq t_1 < t_2 < \dots < t_k = 1$, 相应的 $W(t_j)$ ($j = 1, 2, \dots, k$) 的变化量

$$[W(t_2) - W(t_1)], [W(t_3) - W(t_2)], \dots, [W(t_k) - W(t_{k-1})]$$

为相互独立的随机变量;

(c) 对任何 $0 \leq s < t \leq 1$, 有

$$W(t) - W(s) \sim N(0, t - s) \quad (1.3.1)$$

标准维纳过程可看作是闭区间 $[0, 1]$ 上的连续变化的随机游动。事实上,若令 $s = t - \Delta t \geq 0$, 根据(1.3.1),对任何 $t \in [0, 1]$, 有

$$W(t) - W(t - \Delta t) = \eta_t \sim N(0, \Delta t)$$

也即

$$W(t) = W(t - \Delta t) + \eta_t; \quad \eta_t \sim N(0, \Delta t) \quad (1.3.2)$$

(1.3.2)是时间间隔为 Δt 的随机游动。

由标准维纳过程 $W(t)$,可定义一般的维纳过程。令

$$B(t) = \sigma W(t)$$

其中, $\sigma > 0$ 。 $B(t)$ 称为方差为 σ^2 的维纳过程。不难看出,对任何

$0 \leq s < t \leq 1$, 有

$$B(t) - B(s) \sim N(0, \sigma^2(t - s)) \quad (1.3.3)$$

特别地, 若令 $s = 0, t = 1$, 就有

$$B(1) \sim N(0, \sigma^2)$$

维纳过程 $B(t)$ 和标准维纳过程 $W(t)$ 是对正态分布 $N(0, \sigma^2)$ 和标准正态分布 $N(0, 1)$ 的推广, 它们具有连续函数和正态分布的良好性质, 许多有关单位根过程的极限分布可表示成维纳过程的泛函。比如, 可以定义:

$$V(t) = (B(t))^2$$

根据维纳过程 $B(t)$ 的性质, 在任何一时刻 t , $V(t)$ 有分布

$$V(t) \sim \sigma^2 t \chi^2(1)$$

给定 $\chi^2(1)$, $V(t)$ 是 t 的连续函数。这里, $\chi^2(1)$ 是自由度为 1 的 χ^2 -分布。

标准维纳过程 $W(t)$ 在任一给定时刻 t 为一随机变量, 同时它的轨线(path)(即给定 $W(t)$)是时间 t 的函数。维纳过程 $W(t)$ 的轨线对 t 的连续性是一个重要的特性。为了研究这一特性, 对于任何 $t_1, t_2 \in [0, 1]$ 和 $t_2 > t_1$, 将 $W(t_1)$ 和 $W(t_2)$ 之间的距离定义为

$$d(t_1, t_2) = \sqrt{E(W(t_2) - W(t_1))^2} \quad (1.3.4)$$

与一般的距离函数不同, $d(t_1, t_2)$ 基于 $(W(t_2) - W(t_1))^2$ 的期望值, 这是由于在任一时刻 t , $W(t)$ 为一随机变量。

定理 1.1: 标准维纳过程 $W(t)$ 的轨线在闭区间 $[0, 1]$ 上对于 t 处处连续。

证明: 对于任何 $t_0 \in [0, 1]$, 取 $\Delta t > 0$, 使得 $t_0, t_0 + \Delta t \in$

$[0, 1]$, 标准维纳过程 $W(t)$ 在 $t_0 + \Delta t$ 和 t_0 之间的距离为

$$d(t_0, t_0 + \Delta t) = \sqrt{E(W(t_0 + \Delta t) - W(t_0))^2}$$

若当 $\Delta t \rightarrow 0$ 时, $d(t_0, t_0 + \Delta t)$ 也趋于零, 则称 $W(t)$ 的轨线在 $t = t_0$ 处连续。由 $W(t)$ 的定义

$$W(t_0 + \Delta t) - W(t_0) \sim N(0, \Delta t)$$

因此

$$E(W(t_0 + \Delta t) - W(t_0))^2 = D(W(t_0 + \Delta t) - W(t_0)) = \Delta t$$

当 $\Delta t \rightarrow 0$ 时, 显然有

$$\begin{aligned} d(t_0, t_0 + \Delta t) &= \sqrt{E(W(t_0 + \Delta t) - W(t_0))^2} \\ &= \sqrt{\Delta t} \rightarrow 0 \end{aligned}$$

因此, 定理得证。

在闭区间 $[0, 1]$ 上的连续性, 是标准维纳过程 $W(t)$ 最重要的特性之一, 在以后对有关单位根过程的极限分布的讨论中起重要作用。尽管如此, $W(t)$ 的轨线在 $[0, 1]$ 上对于 t 不可微。

定理 1.2: 标准维纳过程 $W(t)$ 的轨线在闭区间 $[0, 1]$ 上对于 t 处处不可微。

证明: 对任一 $t_0 \in [0, 1]$, 取 $\Delta t > 0$, 使得 $t_0 + \Delta t \in [0, 1]$ 。标准维纳过程 $W(t)$ 的轨线在 $t = t_0$ 处可微, 若

$$\lim_{\Delta t \rightarrow 0} \frac{d(t_0, t_0 + \Delta t)}{\Delta t}$$

存在。由定理 1.1 可知, 此时 $d(t_0, t_0 + \Delta t) = \sqrt{\Delta t}$, 因此

$$\lim_{\Delta t \rightarrow 0} \frac{d(t_0, t_0 + \Delta t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{\sqrt{\Delta t}}{\Delta t} = \infty$$

标准维纳过程 $W(t)$ 的轨线在 $t = t_0$ 处不可微。定理证毕。

标准维纳过程 $W(t)$ 作为一个随机测度, 可以定义在闭区间 $[0, 1]$ 上的随机积分。由于标准维纳过程 $W(t)$ 在 $[0, 1]$ 上的不可微, 可以想象由它定义的随机积分与一般积分的不同。随机积分将在以后章节中作简要说明。

1.4 泛函中心极限定理

中心极限定理 (central limit theorem) 是概率论和数理统计中研究随机变量序列极限分布的一个重要工具, 但是它不适用于非稳定的时间序列过程。以下介绍的泛函中心极限定理 (functional central limit theorem), 是对一般中心极限定理的推广, 而且适用于单位根过程的极限分布。以下先介绍常见的林德贝格—勒维 (Lindeberg-Levy) 中心极限定理, 然后将其推广为泛函中心极限定理。

定理 1.3: 林德贝格—勒维中心极限定理

若 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_t, \dots$ 为一独立同分布的随机变量序列, 且有 $E(\varepsilon_t) = \mu, D(\varepsilon_t) = \sigma^2 < \infty, t = 1, 2, \dots$ 。那么序列的标准化的样本均值

$$\xi_T = \frac{1}{\sigma\sqrt{T}} \sum_{i=1}^T (\varepsilon_i - \mu)$$

有正态的极限分布, 即当 $T \rightarrow \infty$ 时,

$$\xi_T = \frac{1}{\sigma\sqrt{T}} \sum_{i=1}^T (\varepsilon_i - \mu) \xrightarrow{d} N(0, 1) \quad (1.4.1)$$

以上定理的证明在一般概率论和数理统计教科书中都能找到, 故在此从略。以下将此定理推广为泛函的中心极限定理。

对于给定的时间序列的样本 $\varepsilon_1, \dots, \varepsilon_T$, 采用以下构造部分

和的方法:对于给定的样本量 T , 仅用前一半的样本值构造部分和及样本均值:

$$\xi_{[T/2]} = \frac{1}{\sqrt{[T/2]}} \sum_{i=1}^{[T/2]} (\epsilon_i - \mu) \quad (1.4.2)$$

式中的 $[T/2]$ 表示 $T/2$ 的整数部分。显然, 若 T 为一偶数, 那么 $\left[\frac{T}{2}\right] = \frac{T}{2}$; 若 T 为一奇数, 则有 $\left[\frac{T}{2}\right] = \frac{T-1}{2}$ 。假设 $E(\epsilon_i) = 0, i = 1, 2, \dots, T$, 根据上述的林德贝格—勒维定理, (1.4.2) 中的样本均值有正态的极限分布:

$$\xi_{[T/2]} = \frac{1}{\sqrt{[T/2]}} \sum_{i=1}^{[T/2]} \epsilon_i \xrightarrow{d} N(0, \sigma^2) \quad (1.4.3)$$

将以上想法进一步推广。设 r 为闭区间 $[0, 1]$ 上的任一实数, 对于给定的时间序列样本 $\epsilon_1, \epsilon_2, \dots, \epsilon_T$, 取前 $[Tr]$ 部分的样本, 并构造统计量:

$$X_T(r) = \frac{1}{T} \sum_{i=1}^{[Tr]} \epsilon_i \quad (1.4.4)$$

这一统计量将成为泛函中心极限定理的基础。

当 r 在闭区间 $[0, 1]$ 上从 0 到 1 连续变化, 对于给定的样本 $\epsilon_1, \dots, \epsilon_T$, $X_T(r)$ 是闭区间 $[0, 1]$ 上的一阶梯函数, 取值为

$$X_T(r) = \begin{cases} 0 & 0 \leq r < \frac{1}{T} \\ \epsilon_1/T & \frac{1}{T} \leq r < \frac{2}{T} \\ \vdots & \\ (\epsilon_1 + \epsilon_2 + \dots + \epsilon_T)/T & r = 1 \end{cases} \quad (1.4.5)$$

将 $X_T(r)$ 乘上 \sqrt{T} , 再令 $T \rightarrow \infty$, 可求得 $\sqrt{T} X_T(r)$ 的极限分布, 它将是 r 的函数。将 $\sqrt{T} X_T(r)$ 改写成以下形式:

$$\sqrt{T} X_T(r) = \frac{1}{\sqrt{T}} \sum_{i=1}^{[Tr]} \epsilon_i = \frac{\sqrt{[Tr]}}{\sqrt{T}} \left\{ \frac{1}{\sqrt{[Tr]}} \sum_{i=1}^{[Tr]} \epsilon_i \right\} \quad (1.4.6)$$

那么由(1.4.3)中的讨论,立即可知

$$\frac{1}{\sqrt{[Tr]}} \sum_{i=1}^{[Tr]} \epsilon_i \xrightarrow{d} N(0, \sigma^2)$$

另一方面,对于任一 $r \in [0, 1]$

$$\lim_{T \rightarrow \infty} \frac{\sqrt{[Tr]}}{\sqrt{T}} = \sqrt{r}$$

综合起来, $\sqrt{T} X_T(r)$ 有极限分布

$$\sqrt{T} X_T(r) = \frac{1}{\sqrt{T}} \sum_{i=1}^{[Tr]} \epsilon_i \xrightarrow{d} \sqrt{r} N(0, \sigma^2) \equiv N(0, \sigma^2 r)$$

然后,再将正态分布 $N(0, \sigma^2 r)$ 用标准维纳过程表示。因为对任一 $r \in [0, 1]$, 有

$$W(r) \sim N(0, r)$$

这表明, $\sqrt{T} X_T(r)$ 的极限分布与维纳过程

$$B(r) = \sigma W(r)$$

的分布是一致的。从而有以下定理。

定理 1.4: 泛函中心极限定理

设 $\epsilon_1, \epsilon_2, \dots, \epsilon_r, \dots$ 为一列独立同分布的随机变量, 对所有 $t = 1, 2, \dots$, 有 $E(\epsilon_t) = 0, D(\epsilon_t) = E(\epsilon_t^2) = \sigma^2 < \infty$; r 为闭区间 $[0, 1]$ 中的任一正实数。给定样本 $\epsilon_1, \epsilon_2, \dots, \epsilon_r$, 取前 $[Tr]$ 部分样本作统计量

闭区间 $[0, 1]$ 上任何右连续并有左有限极限的函数有界可积, 所以 $[0, 1]$ 上的积分算子对度量空间 $D[0, 1]$ 中的所有元素有定义。对任意给定的 $\varepsilon > 0$, 取 $\delta = \varepsilon$, 当 $f_1(t), f_2(t) \in D[0, 1]$, 并有

$$\rho(f_1, f_2) = \sup_{0 \leq t \leq 1} |f_1(t) - f_2(t)| < \delta$$

那么

$$\begin{aligned} \left| \int_0^1 f_1(t) dt - \int_0^1 f_2(t) dt \right| &= \left| \int_0^1 (f_1(t) - f_2(t)) dt \right| \\ &\leq \int_0^1 |f_1(t) - f_2(t)| dt \leq \int_0^1 \sup_{0 \leq t \leq 1} |f_1(t) - f_2(t)| dt \\ &< \delta \int_0^1 dt = \delta = \varepsilon \end{aligned}$$

定理 1.6: 设 $\{S_t(r), t = 1, 2, \dots\}$ 为一列随机函数, $S_t(r) \in D[0, 1], r \in [0, 1], g(\cdot)$ 是定义在 $D[0, 1]$ 上到 R^1 内的连续映照。若序列 $S_t(r) (t = 1, 2, \dots)$ 弱收敛于随机函数 $S(r)$, 则有

$$g(S_t(r)) \Rightarrow g(S(r)); \quad r \in [0, 1]$$

上述定理的证明需要一些准备知识, 故在此不作介绍, 有兴趣的读者可参阅霍尔和海德(Hall and Heyde, 1980, p. 276)的定理 3 及其证明。

以下的例子利用了泛函中心极限定理和泛函的连续映照定理。

例 1.2 求 $T \rightarrow \infty$ 时, $\int_0^1 \sqrt{T} X_T(r) dr$ 的极限。

根据泛函中心极限定理, 部分和

$$\sqrt{T} X_T(r) = \frac{1}{\sqrt{T}} \sum_{i=1}^{[Tr]} \epsilon_i$$

弱收敛于维纳过程 $B(r)$:

$$\sqrt{T} X_T(r) \Rightarrow B(r) \equiv \sigma W(r)$$

对于给定的样本 $\epsilon_1, \epsilon_2, \dots, \epsilon_T$, $X_T(r) (r \in [0, 1])$ 是闭区间 $[0, 1]$ 上的阶梯函数, 所以, $X_T(r) \in D[0, 1]$ 。另一方面, $\int_0^1 \cdot dt$ 是 $D[0, 1]$ 上的连续映照, 根据连续映照定理

$$\int_0^1 \sqrt{T} X_T(r) dr \Rightarrow \sigma \int_0^1 W(r) dr$$

1.6 有关随机游动的极限分布

现在继续讨论在本章开始时提出的问题: 在 AR(1) 过程

$$y_t = \rho y_{t-1} + \epsilon_t \quad (1.6.1)$$

中, $\{\epsilon_t\}$ 为独立同分布, 且 $E(\epsilon_t) = 0$, $D(\epsilon_t) = E(\epsilon_t^2) = \sigma^2$, 当参数 $\rho = 1$ 时, 最小二乘估计

$$\hat{\rho}_T = \frac{\sum_{t=1}^T y_t y_{t-1}}{\sum_{t=1}^T y_{t-1}^2}$$

有怎样的极限分布。

1.6.1 几个重要的极限

若 AR(1) 过程 (1.6.1) 为一单位根过程, 那么

$$y_t = \sum_{j=1}^t \epsilon_j$$

这里设 $y_0 = 0$ 。对于任何 $r \in [0, 1]$ 和给定的样本 $\epsilon_1, \dots, \epsilon_T$, 部分和

$$X_T(r) = \frac{1}{\sqrt{T}} \sum_{i=1}^{[Tr]} \epsilon_i$$

为闭区间 $[0, 1]$ 上的阶梯函数

$$X_T(r) = \begin{cases} 0 & 0 \leq r < \frac{1}{T} \\ y_1/T & \frac{1}{T} \leq r < \frac{2}{T} \\ \vdots & \vdots \\ y_T/T & r = 1 \end{cases}$$

式中 $y_i = \sum_{t=1}^T \epsilon_t$ 。阶梯函数 $X_T(r)$ 在 $[0, 1]$ 上的积分可由下图定义。

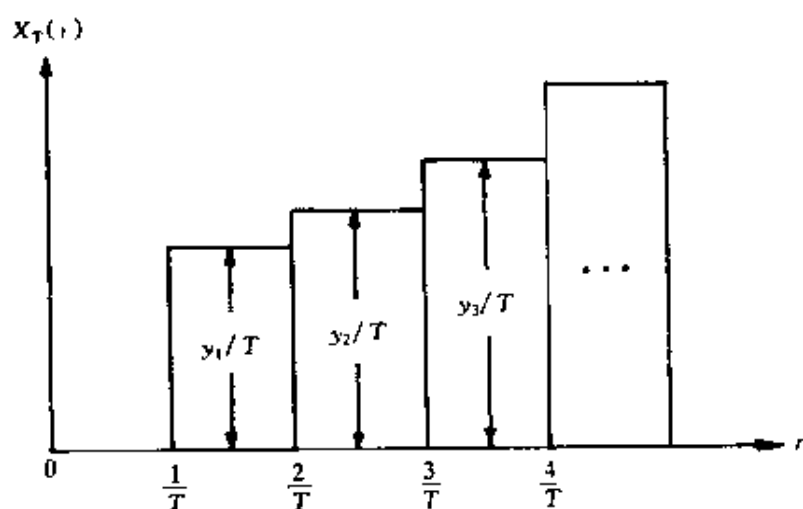


图 1.6 $X_T(r)$ 在 $[0, 1]$ 上的积分

图中每块小矩形的面积为 $(y_j/T) \cdot (1/T)$, $j = 0, 1, 2, \dots, T$ 。这些小矩形面积的和就定义为阶梯函数 $X_T(r)$ 在 $[0, 1]$ 上的积分:

$$\int_0^1 X_T(r) dr = \frac{y_1}{T^2} + \frac{y_2}{T^2} + \dots + \frac{y_{T-1}}{T^2} = T^{-2} \sum_{i=1}^T y_{i-1}$$

以 \sqrt{T} 乘上式, 得

$$\int_0^1 \sqrt{T} X_T(r) dr = T^{-3/2} \sum_{i=1}^T y_{i-1}$$

因为, $\sqrt{T} X_T(r)$ 弱收敛于 $\sigma W(r)$, 根据连续映照定理, 有

$$T^{-3/2} \sum_{i=1}^T y_{i-1} = \int_0^1 \sqrt{T} X_T(r) dr \Rightarrow \sigma \int_0^1 W(r) dr \quad (1.6.2)$$

从而证明了 $T^{-3/2} \sum_{i=1}^T y_{i-1}$ 的极限为维纳过程的泛函。还可显示,

$\sigma \int_0^1 W(r) dr$ 有正态的概率分布。以 $y_i = \sum_{j=1}^i \epsilon_j$ 代入 $T^{-3/2} \sum_{i=1}^T y_{i-1}$, 得

$$\begin{aligned} T^{-3/2} \sum_{i=1}^T y_{i-1} &= T^{-3/2} [\epsilon_1 + (\epsilon_1 + \epsilon_2) + (\epsilon_1 + \epsilon_2 + \epsilon_3) \\ &\quad + \cdots + (\epsilon_1 + \epsilon_2 + \cdots + \epsilon_{T-1})] \\ &= T^{-3/2} [(T-1)\epsilon_1 + (T-2)\epsilon_2 \\ &\quad + \cdots + 2\epsilon_{T-2} + \epsilon_{T-1}] \\ &= T^{-3/2} \sum_{i=1}^T (T-i)\epsilon_i \\ &= T^{-1/2} \sum_{i=1}^T \epsilon_i - T^{-3/2} \sum_{i=1}^T i\epsilon_i \quad (1.6.3) \\ &= (1, -1)h_i \end{aligned}$$

其中

$$h_i = \begin{bmatrix} T^{-1/2} \sum_{i=1}^T \epsilon_i \\ T^{-3/2} \sum_{i=1}^T i\epsilon_i \end{bmatrix}$$

不难计算(见习题),当 $T \rightarrow \infty$ 时,随机向量 \mathbf{h}_t 有联合极限分布

$$N\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma^2 \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix}\right] \quad (1.6.4)$$

所以 $(1, -1)\mathbf{h}_t$ 也有正态的极限分布,其方差为

$$\sigma^2(1, -1) \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{3}\sigma^2$$

因此

$$T^{-3/2} \sum_{i=1}^I y_{t-1} \Rightarrow \sigma \int_0^1 W(r) dr \sim N\left(0, \frac{1}{3}\sigma^2\right) \quad (1.6.5)$$

由(1.6.5)还可进一步推出单位根过程 $\{y_t\}$ 的样本均值 $\bar{y} = \frac{1}{T} \sum_{t=1}^T y_t$ 是发散的。

(1.6.3)式右边的第一项有极限

$$T^{-1/2} \sum_{t=1}^T \epsilon_t \Rightarrow \sigma W(1)$$

将(1.6.3)重新组织还可得

$$T^{-3/2} \sum_{t=1}^T t \epsilon_t \rightarrow \sigma W(1) - \sigma \int_0^1 W(r) dr$$

以上几个极限为计算最小二乘估计 $\hat{\rho}_T$ 的极限分布作了准备。下面的定理收集了几个重要的极限。

定理 1.7: 设随机游动

$$y_t = y_{t-1} + \epsilon_t$$

中, $\{\epsilon_t\}$ 为独立同分布, 且对任何 $t = 1, 2, \dots$, 有 $E(\epsilon_t) = 0$, $D(\epsilon_t) = E(\epsilon_t^2) = \sigma^2 < \infty$, 若 $y_0 = 0$, 以下极限成立:

$$(1) T^{-1/2} \sum_{i=1}^T \epsilon_i \Rightarrow \sigma W(1)$$

$$(2) T^{-1} \sum_{i=1}^T y_{i-1} \epsilon_i \Rightarrow \frac{1}{2} \sigma^2 (W^2(1) - 1)$$

$$(3) T^{-3/2} \sum_{i=1}^T t \epsilon_i \Rightarrow \sigma W(1) - \sigma \int_0^1 W(r) dr$$

$$(4) T^{-1/2} \sum_{i=1}^T y_{i-1} \Rightarrow \sigma \int_0^1 W(r) dr$$

$$(5) T^{-5/2} \sum_{i=1}^T t y_{i-1} \Rightarrow \sigma \int_0^1 r W(r) dr$$

$$(6) T^{-2} \sum_{i=1}^T y_{i-1}^2 \Rightarrow \sigma^2 \int_0^1 W^2(r) dr$$

证明：(1)、(3)和(4)中的结果前面已有推导，现证其他结论。

结论(2)的证明：首先，注意到可将 y_i^2 表示为

$$y_i^2 = (y_{i-1} + \epsilon_i)^2 = y_{i-1}^2 + \epsilon_i^2 + 2y_{i-1}\epsilon_i$$

重新组合后，有

$$y_{i-1}\epsilon_i = \frac{1}{2}(y_i^2 - y_{i-1}^2 - \epsilon_i^2)$$

将上式两边对 i 作和，并除以 T ，得

$$\frac{1}{T} \sum_{i=1}^T y_{i-1}\epsilon_i = \frac{1}{2} \left(\frac{1}{T} y_T^2 - \frac{1}{T} \sum_{i=1}^T \epsilon_i^2 \right)$$

注意到

$$X_T(1) = \frac{1}{T} \sum_{i=1}^T \epsilon_i = \frac{1}{T} y_T$$

并代入(1.6.5)

$$\frac{1}{T} \sum_{i=1}^T y_{i-1}\epsilon_i = \frac{1}{2} \left[\left(\sqrt{T} X_T(1) \right)^2 - \frac{1}{T} \sum_{i=1}^T \epsilon_i^2 \right]$$

由泛函中心极限定理,可得

$$\sqrt{T}X_T(1) \Rightarrow \sigma W(1)$$

而 $\frac{1}{T} \sum_{i=1}^T \epsilon_i^2$ 符合大数定律的条件,从而有

$$\frac{1}{T} \sum_{i=1}^T \epsilon_i^2 \xrightarrow{p} \sigma^2$$

综合以上结果,就有

$$\frac{1}{T} \sum_{i=1}^T y_{i-1} \epsilon_i \Rightarrow \frac{1}{2} \sigma^2 (W^2(1) - 1)$$

结论(5)的证明:将 $T^{-5/2} \sum_{i=1}^T t y_{i-1}$ 改写成以下形式:

$$T^{-5/2} \sum_{i=1}^T t y_{i-1} = T^{1/2} \sum_{i=1}^T \frac{t}{T} \left(\frac{y_{i-1}}{T} \right) \frac{1}{T} \quad (1.6.6)$$

由阶梯函数 $X_T(r)$ 的表示形式,可知当 $\frac{t-1}{T} \leq r < \frac{t}{T}$ 时,有

$$\frac{y_{i-1}}{T} = X_T(r); \quad [Tr] + 1 = t$$

所以

$$\begin{aligned} T^{-5/2} \sum_{i=1}^T t y_{i-1} &= T^{1/2} \sum_{i=1}^T \frac{[Tr] + 1}{T} (X_T(r)) \cdot \frac{1}{T} \\ &= \int_0^1 \frac{[Tr] + 1}{T} (\sqrt{T} X_T(r)) dr \end{aligned}$$

因为

$$\lim_{T \rightarrow \infty} \frac{[Tr]}{T} = r$$

根据泛函中心极限定理

$$W(t_2) - W(t_1) \sim N(0, t_2 - t_1)$$

以 dt 表示在时刻 t 的微小的时间增量, 定义

$$dW(t) = W(t + dt) - W(t)$$

显然

$$dW(t) \sim N(0, dt)$$

并有

$$E(dW(t)) = 0, D(dW(t)) = E(dW(t))^2 = dt$$

令 $\Phi(t)$ 为闭区间 $[0, 1]$ 上的函数(或随机函数), 考虑和式

$$S_n = \sum_{k=1}^n \Phi(t_{k-1}) [W(t_k) - W(t_{k-1})] \quad (1.6.7)$$

其中, $t_k (k = 1, \dots, n)$ 为闭区间上的一组分割点, 并有 $0 \leq t_1 < t_2 < \dots < t_n = 1$, 满足条件

$$\lim_{n \rightarrow \infty} \max_k (t_k - t_{k-1}) = 0$$

若 $n \rightarrow \infty$ 时, S_n 的极限存在, 则称函数 $\Phi(t)$ 在闭区间 $[0, 1]$ 对于 $dW(r)$ 可积, 并记为

$$\begin{aligned} \lim_{n \rightarrow \infty} S_n &= \lim_{n \rightarrow \infty} \sum_{k=1}^n \Phi(t_{k-1}) [W(t_k) - W(t_{k-1})] \\ &= \int_0^1 \Phi(t) dW \end{aligned}$$

尽管随机积分的定义和一般的积分很相似, 由于 dW 的随机性, 它们的计算规则有所不同。考虑随机积分

$$\int_0^1 W(t) dW$$

作和式

$$S_n = \sum_{k=1}^n W(t_{k-1})[W(t_k) - W(t_{k-1})]$$

因为

$$\begin{aligned} & W(t_{k-1})[W(t_k) - W(t_{k-1})] \\ &= \frac{1}{2}[W^2(t_k) - W^2(t_{k-1})] - \frac{1}{2}[W(t_k) - W(t_{k-1})]^2 \end{aligned}$$

可将和式 S_n 改写成

$$S_n = \frac{1}{2} \sum_{k=1}^n [W^2(t_k) - W^2(t_{k-1})] - \frac{1}{2} \sum_{k=1}^n [W(t_k) - W(t_{k-1})]^2$$

令 $t_k = \frac{k}{n}$, $k = 1, 2, \dots, n$, 注意到 $W(0) = 0$, 可将 S_n 简化为

$$S_n = \frac{1}{2} W^2(1) - \frac{1}{2} \sum_{k=1}^n \left[W\left(\frac{k}{n}\right) - W\left(\frac{k-1}{n}\right) \right]^2$$

令

$$V(k) = W\left(\frac{k}{n}\right) - W\left(\frac{k-1}{n}\right) = \sqrt{\frac{1}{n}} \cdot Z(k)$$

这里, $Z(k)$ ($k = 1, 2, \dots, n$) 相互独立, 并有标准正态分布。因此

$$V^2(k) = \frac{1}{n} Z^2(k) \sim \frac{1}{n} \chi^2(1)$$

将 $V(k)$ 代入 S_n , 根据大数定律, 有

$$\lim_{n \rightarrow \infty} S_n = \frac{1}{2} W^2(1) - \frac{1}{2} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n Z^2(k) = \frac{1}{2} (W^2(1) - 1)$$

因为

$$\frac{1}{n} \sum_{k=1}^n Z^2(k) \xrightarrow{P} E(Z^2(k)) = 1$$

所以有

$$\int_0^1 W(t) dW = \frac{1}{2} (W^2(1) - 1) \quad (1.6.8)$$

1.6.3 最小二乘估计 $\hat{\rho}_T$ 的极限分布

这节中讨论随机游动

$$y_t = \rho y_{t-1} + \varepsilon_t; \quad t = 1, 2, \dots$$

中最小二乘估计 $\hat{\rho}_T$ 的极限分布。其中, $\rho = 1$, $\{\varepsilon_t\}$ 独立同分布, 且 $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = \sigma^2 < \infty$ 。对于一般的单位根过程, 我们留在以后讨论。

参数 $\rho = 1$ 的最小二乘估计为

$$\hat{\rho}_T = \frac{\sum_{t=1}^T y_{t-1} y_t}{\sum_{t=1}^T y_{t-1}^2}$$

其中, T 为给定的样本量。若 $\{y_t\}$ 是随机游动, 则有 $\rho = 1$, 那么

$$\hat{\rho}_T = 1 + \frac{\sum_{t=1}^T y_{t-1} \varepsilon_t}{\sum_{t=1}^T y_{t-1}^2}$$

根据定理 1.7 的结论(2)和(6)

$$T^{-1} \sum_{t=1}^T y_{t-1} \varepsilon_t \Rightarrow \frac{1}{2} \sigma^2 (W^2(1) - 1)$$

$$T^{-2} \sum_{t=1}^T y_{t-1}^2 \Rightarrow \sigma^2 \int_0^1 W^2(r) dr$$

可得

会小于 1, 尽管这时 ρ 的真值为 1。

(c) 由于 $T(\hat{\rho} - 1)$ 的极限分布的非标准性, 传统的计算分布的临界值的方法不再适用。书后的附表只是由蒙的可罗 (Monte Carlo) 模拟方法计算所得。

1.6.4 t_T 统计量的极限分布

若一阶自回归过程

$$y_t = \rho y_{t-1} + \epsilon_t$$

是稳定的, 参数 ρ 的假设 $H_0: \rho = \rho_0$ 可由 t_T 统计量检验:

$$t_T = \frac{\hat{\rho}_T - \rho_0}{\hat{\eta}_T}$$

其中, $\hat{\rho}_T$ 为最小二乘估计, $\hat{\eta}_T$ 为 $\hat{\rho}_T$ 的标准差

$$\hat{\eta}_T = \left[\frac{\hat{\sigma}_T^2}{\sum_{t=1}^T y_{t-1}^2} \right]^{1/2}$$

$\hat{\sigma}_T^2$ 是 ϵ_t 的方差的最小二乘估计

$$\hat{\sigma}_T^2 = \frac{\sum_{t=1}^T (y_t - \hat{\rho}_T y_{t-1})^2}{T-1} \quad (1.6.10)$$

当原假设 $H_0: \rho = \rho_0$ 为真时, t_T 有自由度为 $T-1$ 的 t -分布。

现在的问题是, t_T 统计量是否还能用来检验假设 $H_0: \rho = 1$; 在原假设 H_0 为真时, t_T 有怎样的分布。以下就来回答这些问题。将 $\hat{\rho}_T$ 的表达式代入 t_T , 得

$$t_T = \frac{\sum_{t=1}^T y_{t-1} \epsilon_t}{\hat{\sigma}_T \sqrt{\sum_{t=1}^T y_{t-1}^2}}$$

当 $\rho = 1$ 为真时,最小二乘估计 $\hat{\rho}_T$ 是超一致的估计量,所以 ϵ_t 的方差的最小二乘估计 $\hat{\sigma}_T^2$ 是 σ^2 的一致估计,即 $\hat{\sigma}_T^2 \xrightarrow{P} \sigma^2$ 。另一方面,因为

$$T^{-1} \sum_{t=1}^T y_{t-1} \epsilon_t \Rightarrow \frac{1}{2} \sigma^2 (W^2(1) - 1)$$

$$T^{-2} \sum_{t=1}^T y_{t-1}^2 \Rightarrow \sigma^2 \int_0^1 W^2(r) dr$$

所以当原假设 $H_0: \rho = 1$ 为真时,有

$$t_T = \frac{T^{-1} \sum_{t=1}^T y_{t-1} \epsilon_t}{\hat{\sigma}_T \sqrt{T^{-2} \sum_{t=1}^T y_{t-1}^2}} \Rightarrow \frac{\frac{1}{2} (W^2(1) - 1)}{\left\{ \int_0^1 W^2(r) dr \right\}^{1/2}} \quad (1.6.11)$$

统计量 t_T 弱收敛于维纳过程的泛函,具有非对称的极限分布。与最小二乘估计 $\hat{\rho}_T$ 的分布值一样, t_T 的值由蒙的可罗方法模拟得到。

1.7 带常数项的随机游动

在随机过程

$$y_t = \alpha + \rho y_{t-1} + \epsilon_t \quad (1.7.1)$$

中,若 $\alpha \neq 0$, $\rho = 1$, $\{\epsilon_t\}$ 独立同分布,且对任何 $t = 1, 2, \dots$, $E(\epsilon_t) = 0$, $D(\epsilon_t) = \sigma^2 < \infty$, 则称 $\{y_t\}$ 服从带常数项的随机游动。

带常数项的随机游动 $\{y_t\}$ 的增长率是一非零的常数 α 和随机干扰的和,它适合于描述带有趋势的经济变量,如国民经济总值等。

本节考虑对参数 α 和 ρ 的最小二乘估计 $\hat{\alpha}_T$ 和 $\hat{\rho}_T$, 讨论它

可见常数项 α 改变了 $\{y_t\}$ 的统计性质, 使得规范化的部分和 $T^{-3/2} \sum_{t=1}^T y_{t-1}$ 不再有非退化的概率分布。另一方面, 如以 $T^{3/2}$ 除 $\sum_{t=1}^T y_{t-1}$, 并令 $T \rightarrow \infty$, 得

$$T^{-3/2} \sum_{t=1}^T y_{t-1} = T^{-1/2} y_0 + \frac{T(T-1)\alpha}{2T^{3/2}} + T^{-3/2} \sum_{t=1}^T \xi_{t-1}$$

尽管这时第一项和第三项存在极限, 起主导作用的 $\frac{T(T-1)\alpha}{2T^{3/2}}$ 发散到无穷大, 所以 $T^{-3/2} \sum_{t=1}^T y_{t-1}$ 亦发散。

为导出 $\hat{\alpha}_T$ 和 $\hat{\rho}_T$ 的极限分布, 还需计算 $T^{-3} \sum_{t=1}^T y_{t-1}^2$ 和 $T^{-3/2} \sum_{t=1}^T y_{t-1} \xi_t$ 的极限。

先将和式 $\sum_{t=1}^T y_{t-1}^2$ 作以下处理:

$$\begin{aligned} \sum_{t=1}^T y_{t-1}^2 &= \sum_{t=1}^T (y_0 + (t-1)\alpha + \xi_{t-1})^2 \\ &= Ty_0^2 + \sum_{t=1}^T \alpha^2 (t-1)^2 + \sum_{t=1}^T \xi_{t-1}^2 \\ &\quad + 2 \sum_{t=1}^T y_0 \alpha (t-1) + 2 \sum_{t=1}^T y_0 \xi_{t-1} + 2 \sum_{t=1}^T \alpha (t-1) \xi_{t-1} \end{aligned} \quad (1.7.5)$$

再除以 T^3 , 并令 $T \rightarrow \infty$ 。因为

$$T^{-2} \sum_{t=1}^T \xi_{t-1}^2 \Rightarrow \sigma^2 \int_0^1 W^2(r) dr$$

$$T^{-3} \sum_{t=1}^T \alpha^2 (t-1)^2 = \frac{\alpha^2}{6} \cdot \frac{T(T-1)(2T+1)}{T^3} \longrightarrow \frac{\alpha^2}{3}$$

$$T^{-3/2} \sum_{t=1}^T \xi_{t-1} \Rightarrow \sigma \int_0^1 W(r) dr$$

所以在和式(1.7.5)中,第二项 $\sum_{t=1}^T \alpha^2 (t-1)^2$ 在 $T \rightarrow \infty$ 时起主导作用,从而

$$\begin{aligned} T^{-3} \sum_{t=1}^T y_{t-1}^2 &= T^{-3} (T y_0^2) + T^{-3} \sum_{t=1}^T \alpha^2 (t-1)^2 + T^{-3} \sum_{t=1}^T \xi_{t-1}^2 \\ &\quad + 2y_0 \alpha T^{-3} \sum_{t=1}^T (t-1) + 2y_0 T^{-3} \sum_{t=1}^T \xi_{t-1} \\ &\quad + 2\alpha T^{-3} \sum_{t=1}^T (t-1) \xi_{t-1} \xrightarrow{P} \frac{\alpha^2}{3} \end{aligned} \quad (1.7.6)$$

最后,考虑和式 $\sum_{t=1}^T y_{t-1} \epsilon_t$ 。因为

$$\begin{aligned} \sum_{t=1}^T y_{t-1} \epsilon_t &= \sum_{t=1}^T (y_0 + \alpha(t-1) + \xi_{t-1}) \epsilon_t \\ &= y_0 \sum_{t=1}^T \epsilon_t + \alpha \sum_{t=1}^T (t-1) \epsilon_t + \sum_{t=1}^T \xi_{t-1} \epsilon_t \end{aligned} \quad (1.7.7)$$

根据前面的讨论

$$T^{-1} \sum_{t=1}^T \epsilon_t \xrightarrow{P} 0; \quad T^{-1} \sum_{t=1}^T \xi_{t-1} \epsilon_t \Rightarrow \frac{1}{2} \sigma^2 (W^2(1) - 1);$$

$$T^{-3/2} \sum_{t=1}^T (t-1) \epsilon_t \Rightarrow \sigma W(1) - \sigma \int_0^1 W(r) dr$$

这等价于

$$\begin{bmatrix} T^{1/2}(\hat{\alpha}_T - \alpha) \\ T^{3/2}(\hat{\rho}_T - 1) \end{bmatrix} = \begin{bmatrix} 1 & T^{-2} \sum_{t=1}^T y_{t-1} \\ T^{-2} \sum_{t=1}^T y_{t-1} & T^{-3} \sum_{t=1}^T y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} T^{-1/2} \sum_{t=1}^T \epsilon_t \\ T^{-3/2} \sum_{t=1}^T y_{t-1} \epsilon_t \end{bmatrix} \quad (1.7.9)$$

令 $T \rightarrow \infty$, 利用 (1.7.4) 和 (1.7.6) 中的结果, 可得

$$\begin{bmatrix} 1 & T^{-2} \sum_{t=1}^T y_{t-1} \\ T^{-2} \sum_{t=1}^T y_{t-1} & T^{-3} \sum_{t=1}^T y_{t-1}^2 \end{bmatrix} \xrightarrow{p} \begin{bmatrix} 1 & \frac{\alpha}{2} \\ \frac{\alpha}{2} & \frac{\alpha^2}{3} \end{bmatrix} \equiv Q \quad (1.7.10)$$

这里 Q 为一对称正定的实矩阵。再将 (1.7.9) 右侧的向量写成下列形式:

$$\begin{bmatrix} T^{-1/2} \sum_{t=1}^T \epsilon_t \\ T^{-3/2} \sum_{t=1}^T y_{t-1} \epsilon_t \end{bmatrix} = \begin{bmatrix} T^{-1/2} \sum_{t=1}^T \epsilon_t \\ T^{-3/2} \sum_{t=1}^T \alpha(t-1) \epsilon_t \end{bmatrix} \\ - \begin{bmatrix} 0 \\ T^{-3/2} \sum_{t=1}^T (y_0 + \hat{\xi}_{t-1}) \epsilon_t \end{bmatrix}$$

根据定理 1.7 的结论 (2)

$$T^{-1} \sum_{t=1}^T \hat{\xi}_{t-1} \epsilon_t \Rightarrow \frac{1}{2} \sigma^2 (W^2(1) - 1)$$

所以上式右侧的第二项当 $T \rightarrow \infty$ 时趋于零。参照 (1.6.4) 中的结果, 可知上式右侧的第一项有极限分布:

$$\begin{bmatrix} T^{-1/2} \sum_{t=1}^T \varepsilon_t \\ T^{-3/2} \sum_{t=1}^T y_{t-1} \varepsilon_t \end{bmatrix} \xrightarrow{d} N \left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma^2 \begin{bmatrix} 1 & \frac{\alpha}{2} \\ \frac{\alpha}{2} & \frac{\alpha^2}{3} \end{bmatrix} \right] \equiv N(\mathbf{0}, \sigma^2 \mathbf{Q})$$

综合以上讨论, 就有

$$\begin{bmatrix} T^{1/2}(\hat{a}_T - \alpha) \\ T^{3/2}(\hat{\rho}_T - 1) \end{bmatrix} \xrightarrow{d} N(\mathbf{0}, \sigma^2 \mathbf{Q}^{-1} \mathbf{Q} \mathbf{Q}^{-1}) \equiv N(\mathbf{0}, \sigma^2 \mathbf{Q}^{-1}) \quad (1.7.11)$$

与不带常数项的随机游动比较, $T^{1/2}(\hat{a}_T - \alpha)$ 和 $T^{3/2}(\hat{\rho}_T - 1)$ 的极限分布有以下特点:

(a) 它们有正态的极限分布, 但在不带常数项的情况下, $T(\hat{\rho}_T - 1)$ 的分布是非标准的和非对称的。

(b) \hat{a}_T 和 $\hat{\rho}_T$ 都为一致的估计量, 但有不同的收敛速度。

(c) $\hat{\rho}_T$ 的收敛速度为 $T^{-3/2}$, 而在不带常数项的情况下, $\hat{\rho}_T$ 的收敛速度为 T^{-1} 。

1.8 有关单位根过程的极限分布

与随机游动过程相比, 单位根过程

$$y_t = \rho y_{t-1} + u_t; \quad t = 1, 2, \dots$$

具有更一般的统计性质, 其中 $\{u_t\}$ 是一稳定过程, 有无穷阶的 MA 表示形式:

$$\begin{aligned}
 u_t &= \varepsilon_t + \varphi_1 \varepsilon_{t-1} + \varphi_2 \varepsilon_{t-2} + \cdots \\
 &= \varphi(L) \varepsilon_t = \sum_{j=0}^{\infty} \varphi_j \varepsilon_{t-j},
 \end{aligned} \tag{1.8.1}$$

其中, $\{\varepsilon_t\}$ 为独立同分布, 且 $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = \sigma^2$, L 为滞后算子, $\varphi(L)$ 为无穷阶的滞后多项式, 其系数满足条件:

$$\sum_{j=0}^{\infty} j |\varphi_j| < \infty \tag{1.8.2}$$

大多数常见的稳定过程都有(1.8.1)的表示形式, 白噪声过程是最简单的一种。

以下将前面讨论的参数 α 和 ρ 的最小二乘估计的极限分布推广到一般的单位根过程, 为此先介绍一个有用的引理。

引理 1.1: BN 分解

设 $\{u_t\}$ 为一稳定过程, 可表示为

$$u_t = \varphi(L) \varepsilon_t = \sum_{j=0}^{\infty} \varphi_j \varepsilon_{t-j},$$

这里, $\{\varepsilon_t\}$ 独立同分布, 且 $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = \sigma^2 < \infty$ 。若系数 $\varphi_j (j = 1, 2, \cdots)$ 满足条件(1.8.2), 那么 u_t 的部分和 $\sum_{t=1}^T u_t$ 有 BN 分解:

$$\sum_{t=1}^T u_t = \varphi(1) \sum_{t=1}^T \varepsilon_t + \eta_T - \eta_0 \tag{1.8.3}$$

其中

$$\varphi(1) = \sum_{j=0}^{\infty} \varphi_j; \quad \eta_T = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{T-j}; \quad \alpha_j = - \sum_{i=1}^{\infty} \varphi_{j+i}; \quad \sum_{j=0}^{\infty} |\alpha_j| < \infty$$

此引理由贝弗里奇和纳尔逊(Beveridge and Nelson, 1981)提出, 故得名。引理的证明与本书内容无直接关系, 在此从略。

根据 BN 分解引理, 形如 (1.8.1) 的稳定过程 $\{u_t\}$ 的部分和可分解成两部分, 其中 $\varphi(1) \sum_{i=1}^T \varepsilon_i$ 以概率与 $T^{1/2}$ 同阶 ($O_P(T^{1/2})$); 另一部分 $\eta_T - \eta_0$ 以概率与常数同阶 ($O_P(1)$)。这样, 如以 $T^{-1/2}$ 乘 $\sum_{i=1}^T u_i$, 并令 $T \rightarrow \infty$, 只有第一部分有非退化的极限分布, 第二部分以概率趋向于零。

利用 BN 分解方法, 可将泛函中心极限定理 (定理 1.4) 推广到一般的单位根过程。

定理 1.8: 一般形式的泛函中心极限定理

给定稳定过程 $\{u_t\}$:

$$u_t = \varphi(L)\varepsilon_t = \sum_{i=1}^{\infty} \varphi_i \varepsilon_{t-i},$$

其中, 系数 $\{\varphi_i\}$ 满足条件 (1.8.2), $\{\varepsilon_t\}$ 独立同分布, 对任何 $t = 1, 2, \dots$, 有 $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = \sigma^2 < \infty$, 构造随机函数 $X_T(r)$:

$$X_T(r) = \frac{1}{T} \sum_{i=1}^{[Tr]} u_i; \quad r \in [0, 1]$$

那么, $\{\sqrt{T}X_T(r)\}$ 为弱收敛序列, 且有

$$\sqrt{T}X_T(r) \Rightarrow \sigma\varphi(1)W(r) \quad (1.8.4)$$

证明: 利用引理 1.1, 将 $\sqrt{T}X_T(r)$ 作 BN 分解

$$\begin{aligned} \sqrt{T}X_T(r) &= \frac{1}{\sqrt{T}} \sum_{i=1}^{[Tr]} u_i \\ &= \frac{1}{\sqrt{T}} \left(\varphi(1) \sum_{i=1}^{[Tr]} \varepsilon_i + \eta_{[Tr]} - \eta_0 \right) \\ &= \varphi(1) \frac{1}{\sqrt{T}} \sum_{i=1}^{[Tr]} \varepsilon_i + R_T(r) \end{aligned} \quad (1.8.5)$$

因为 $(\eta_{[T]r} - \eta_0) = O_p(1)$, 当 $T \rightarrow \infty$ 时, 有

$$R_T(r) = \frac{1}{\sqrt{T}}(\eta_{[Tr]} - \eta_0) \xrightarrow{p} 0$$

另由定理 1.4

$$\frac{1}{\sqrt{T}} \sum_{i=1}^{[Tr]} \varepsilon_i \Rightarrow \sigma W(r)$$

所以

$$\begin{aligned} \sqrt{T} X_T(r) &= \varphi(1) \frac{1}{\sqrt{T}} \sum_{i=1}^{[Tr]} \varepsilon_i + R_T(r) \\ &\Rightarrow \sigma \varphi(1) W(r) \end{aligned} \quad (1.8.6)$$

定理证毕。

不难看出, 对任一给定的 $r \in [0, 1]$, 有

$$\sigma \varphi(1) W(r) \sim N(0, \sigma^2 \varphi^2(1) r)$$

定理 1.9: 定理 1.7 的推广形式

给定稳定过程 $\{u_t\}$ 满足定理 1.8 中的条件, 令

$$r_j = E(u_t u_{t-j}) = \sigma^2 \sum_{s=0}^{\infty} \varphi_s \varphi_{s+j}; \quad j = 0, 1, 2, \dots$$

$$\xi_t = u_1 + u_2 + \dots + u_t; \quad t = 1, 2, \dots, T$$

若 $\xi_0 = 0$, 那么

$$(1) T^{-1/2} \sum_{t=1}^T u_t \Rightarrow \sigma \varphi(1) W(1)$$

$$(2) T^{-1/2} \sum_{t=1}^T u_{t-j} \varepsilon_t \xrightarrow{d} N(0, \sigma^2 r_0); \quad j = 1, 2, \dots$$

$$(3) T^{-1} \sum_{t=1}^T u_t u_{t-j} \xrightarrow{p} r_j; \quad j = 0, 1, 2, \dots$$

$$(4) T^{-1} \sum_{i=1}^T \xi_{i-1} \epsilon_i \Rightarrow \frac{1}{2} \sigma^2 \varphi(1) \{W^2(1) - 1\}$$

$$(5) T^{-1} \sum_{i=1}^T \xi_{i-1} u_{i-j} \Rightarrow \frac{1}{2} \{ \sigma^2 \varphi^2(1) W^2(1) - r_0 \}; \quad j = 0$$

$$\Rightarrow \frac{1}{2} \{ \sigma^2 \varphi^2(1) W^2(1) - r_0 \} + \sum_{i=0}^{j-1} r_i;$$

$$j = 1, 2, \dots$$

$$(6) T^{-3/2} \sum_{i=1}^T \xi_{i-1} \Rightarrow \sigma \varphi(1) \int_0^1 W(r) dr$$

$$(7) T^{-3/2} \sum_{i=1}^T t u_{i-j} \Rightarrow \sigma \varphi(1) \{W(1) - \int_0^1 W(r) dr\};$$

$$j = 0, 1, 2, \dots$$

$$(8) T^{-2} \sum_{i=1}^T \xi_{i-1}^2 \Rightarrow \sigma^2 \varphi^2(1) \int_0^1 W^2(r) dr$$

$$(9) T^{-5/2} \sum_{i=1}^T t \xi_{i-1} \Rightarrow \sigma \varphi(1) \int_0^1 r W(r) dr$$

$$(10) T^{-3} \sum_{i=1}^T t \xi_{i-1}^2 \Rightarrow \sigma^2 \varphi^2(1) \int_0^1 r W^2(r) dr$$

证明:

结论(1):在(1.8.6)中令 $r = 1$ 即得。

结论(2):对任何 $j = 1, 2, \dots$, u_{i-j} 与 ϵ_i 独立,因此 $E(u_{i-j}, \epsilon_i) = 0$, (u_{i-j}, ϵ_i) 的方差为

$$D(u_{i-j}, \epsilon_i) = E(\epsilon_{i-j} \epsilon_i + \varphi_1 \epsilon_{i-j-1} \epsilon_i + \dots)^2$$

$$= \sigma^2 (\varphi_0^2 + \varphi_1^2 + \dots) = \sigma^2 r_0; \quad j = 1, 2, \dots$$

序列 $\{u_{i-j}, \epsilon_i, i = 1, 2, \dots\}$ 独立同分布,根据中心极限定理,得

$$\frac{1}{\sqrt{T}} \sum_{i=1}^T u_{i-j}, \epsilon_i \xrightarrow{d} N(0, \sigma^2 r_0)$$

结论(3):直接利用大数定律即得。

结论(4);对 $\hat{\xi}_{t-1}$ 作 BN 分解:

$$\hat{\xi}_{t-1} = \sum_{j=1}^{t-1} u_j = \varphi(1) \sum_{j=1}^{t-1} \varepsilon_j + \eta_{t-1} - \eta_0$$

再乘以 ε_t , 并作和式

$$\begin{aligned} T^{-1} \sum_{t=1}^T \hat{\xi}_{t-1} \varepsilon_t &= T^{-1} \sum_{t=2}^T \left\{ \varphi(1) \sum_{j=1}^{t-1} \varepsilon_j + \eta_{t-1} - \eta_0 \right\} \varepsilon_t \\ &= \varphi(1) T^{-1} \sum_{t=2}^T (\varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_{t-1}) \varepsilon_t + T^{-1} \sum_{t=2}^T (\eta_{t-1} - \eta_0) \varepsilon_t \end{aligned}$$

由定理 1.7 结论(2)

$$T^{-1} \sum_{t=2}^T (\varepsilon_1 + \cdots + \varepsilon_{t-1}) \varepsilon_t \Rightarrow \frac{1}{2} \sigma^2 \{W^2(1) - 1\}$$

$(\eta_{t-1} - \eta_0)$ 与 ε_t 互相独立, 所以有

$$T^{-1} \sum_{t=2}^T (\eta_{t-1} - \eta_0) \varepsilon_t \xrightarrow{p} 0$$

综上所述, 得

$$T^{-1} \sum_{t=1}^T \hat{\xi}_{t-1} \varepsilon_t \Rightarrow \frac{1}{2} \sigma^2 \varphi(1) \{W^2(1) - 1\}$$

结论(5): 先考虑 $j = 0$ 。因为

$$\xi_t^2 = (\xi_{t-1} + u_t)^2 = \xi_{t-1}^2 + u_t^2 + 2\xi_{t-1}u_t$$

所以

$$\begin{aligned} T^{-1} \sum_{t=1}^T \xi_{t-1} u_t &= \frac{1}{2} T^{-1} \sum_{t=1}^T (\xi_t^2 - \xi_{t-1}^2 - u_t^2) \\ &= \frac{1}{2} T^{-1} \xi_T^2 - \frac{1}{2} T^{-1} \sum_{t=1}^T u_t^2 \end{aligned}$$

注意到

$$\frac{1}{2}T^{-1}\xi_T^2 = \frac{1}{2}\left[T^{-1/2}\sum_{i=1}^T u_i\right]^2$$

因为

$$T^{-1/2}\sum_{i=1}^T u_i \Rightarrow \sigma\varphi(1)W(1)$$

由连续映照定理,得

$$\frac{1}{2}T^{-1}\xi_T^2 = \frac{1}{2}\left[T^{-1/2}\sum_{i=1}^T u_i\right]^2 \Rightarrow \frac{1}{2}\sigma^2\varphi^2(1)W^2(1)$$

另一方面,由结论(3)

$$T^{-1}\sum_{i=1}^T u_i^2 \xrightarrow{p} r_0$$

所以

$$T^{-1}\sum_{i=1}^T \xi_{i-1}u_i \Rightarrow \frac{1}{2}\{\sigma^2\varphi^2(1)W^2(1) - r_0\} \quad (1.8.7)$$

其次,考虑 $j > 0$ 。因为

$$\xi_{i-1} = \xi_{i-2} + u_{i-1} = \xi_{i-j-1} + u_{i-j} + u_{i-j+1} + \cdots + u_{i-1}$$

以 u_{i-j} 乘上式,并从 $i = j+1$ 到 T 作和式

$$\begin{aligned} T^{-1}\sum_{i=j+1}^T \xi_{i-1}u_{i-j} &= T^{-1}\sum_{i=j+1}^T (\xi_{i-j+1} - u_{i-j} \\ &\quad + u_{i-j+1} + \cdots + u_{i-1})u_{i-j} \\ &= T^{-1}\sum_{i=j+1}^T \xi_{i-j+1}u_{i-j} + T^{-1}\sum_{i=j+1}^T \\ &\quad (u_{i-j} + u_{i-j+1} + \cdots + u_{i-1})u_{i-j} \quad (1.8.8) \end{aligned}$$

对上式中的第一项应用(1.8.7)中的结果,即有

$$T^{-1} \sum_{i=j+1}^T \xi_{t-j-1} u_{t-j} = \frac{T-j}{T} \left\{ \frac{1}{T-j} \sum_{i=j+1}^T \xi_{t-j-1} u_{t-j} \right\} \\ \Rightarrow \frac{1}{2} \{ \sigma^2 \varphi^2(1) W^2(1) - r_0 \}$$

对(1.8.8)中的第二个和式应用结论(3)

$$T^{-1} \sum_{i=j+1}^T (u_{t-j} + u_{t-j+1} + \cdots + u_{t-1}) u_{t-j} \xrightarrow{p} \sum_{i=0}^{j-1} r_i$$

综上所述,得

$$T^{-1} \sum_{i=j+1}^T \xi_{t-1} u_{t-j} \Rightarrow \frac{1}{2} \{ \sigma^2 \varphi^2(1) W^2(1) - r_0 \} + \sum_{i=0}^{j-1} r_i$$

最后,对于任何给定的 $j > 0$, $T^{-1} \sum_{i=1}^T \xi_{t-1} u_{t-j}$ 和

$T^{-1} \sum_{i=j+1}^T \xi_{t-1} u_{t-j}$ 有相同的极限,因为

$$T^{-1} \sum_{i=1}^T \xi_{t-1} u_{t-j} - T^{-1} \sum_{i=j+1}^T \xi_{t-1} u_{t-j} = T^{-1} \sum_{i=1}^j \xi_{t-1} u_{t-j} \xrightarrow{p} 0$$

结论(6):根据定义, $X_T(r) = \frac{1}{T} \sum_{i=1}^{[Tr]} u_i$ 在闭区间 $[0, 1]$ 上的

积分可表示为

$$\int_0^1 X_T(r) dr = T^{-2} \sum_{i=1}^T \xi_{i-1}$$

以 \sqrt{T} 乘上式,根据结论(1)和连续映照定理,当 $T \rightarrow \infty$,有

$$T^{-3/2} \sum_{i=1}^T \xi_{i-1} = \int_0^1 \sqrt{T} X_T(r) dr \Rightarrow \sigma \varphi(1) \int_0^1 W(r) dr$$

结论(7):将和式 $T^{-3/2} \sum_{i=1}^T t u_{t-j}$ 写成以下形式:

$$T^{-3/2} \sum_{i=1}^T t u_{t-j} = T^{-3/2} \sum_{i=1}^T (t - j + j) u_{t-j}$$

$$T^{-3} \sum_{t=1}^T t \xi_{t-1}^2 \Rightarrow \sigma^2 \varphi^2(1) \int_0^1 r W^2(r) dr$$

定理证毕。

1.9 时间序列的去势

许多时间序列数据,特别是宏观经济数据,常常显示出明显的时间趋势(trend),即序列的期望值是时间 t 的函数。这样的趋势往往在时间序列中起主导作用,左右序列的统计性质。习惯上常用去势的方法(de-trending),消除时间趋势在序列中的作用,然后再对其他参数作统计分析。

时间序列中的趋势有不同的表现形式。比如,在带趋势的稳定过程

$$y_t = c + rt + u_t \quad (1.9.1)$$

中, rt 表示了时间序列 $\{y_t\}$ 的确定的趋势(deterministic trend),这时 y_t 的期望是时间 t 的线性函数, y_t 的值在 $c + rt$ 周围波动。

有些时间序列中的趋势隐性地存在,并不在表达式中直接表示出来。如在带常数项的单位根过程

$$y_t = \alpha + y_{t-1} + u_t \quad (1.9.2)$$

中,对 y_t 不断向后迭代,可得

$$\begin{aligned} y_t &= \alpha + (\alpha + y_{t-2} + u_{t-1}) + u_t \\ &= \dots \\ &= \alpha t + \sum_{j=1}^t u_j \end{aligned} \quad (1.9.3)$$

经处理后的 y_t 包含确定的时间趋势 αt ,它由单位根中的常数项

积累而成。尽管(1.9.1)和(1.9.2)都含有确定的时间趋势,而且从图形上看它们也很相似,但它们有非常不同的统计性质,在经济上的应用也不一样。所以,严格区分这两种过程在理论上和实践上都有重要意义。

在实际工作中,过程(1.9.1)中是否存在确定的时间趋势,由对假设

$$H_0: r = 0$$

的检验而确定。对原假设 H_0 的拒绝,往往解释成接受过程(1.9.1),即认为数据是由过程(1.9.1)产生的。

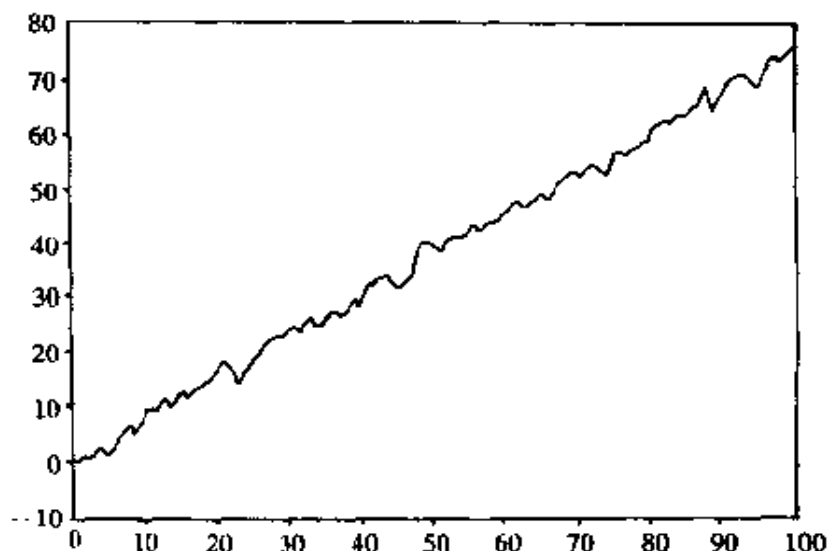


图 1.7 带趋势的稳定过程: $y_t = \alpha + u_t$

以下的讨论表明,以对 $H_0: r = 0$ 的检验,作为是否接受模型(1.9.1)的标准,在许多场合下是不恰当的。

假设数据 y_t 由带常数项的单位根过程(1.9.2)产生,但我们却选取了

$$y_t = c + rt + u_t$$

作为描述 y_t 的模型,并在其中检验参数 c 和 r 的显著性。以下分别考虑参数 c 和 r 的最小二乘估计 \hat{c}_T 和 \hat{r}_T ,以及它们的极限。

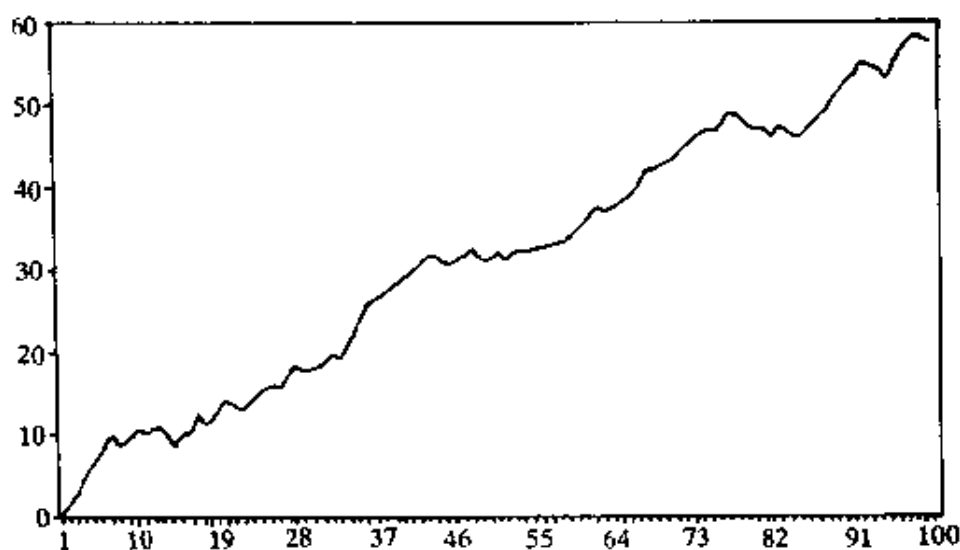


图 1.8 带常数项的单位根过程: $y_t = \alpha + y_{t-1} + u_t$

最小二乘估计 \hat{c}_T 可表示为

$$\hat{c}_T = d^{-1} \left\{ \left(\sum_{t=1}^T t^2 \right) \left(\sum_{t=1}^T y_t \right) - \left(\sum_{t=1}^T t y_t \right) \left(\sum_{t=1}^T t \right) \right\} \quad (1.9.4)$$

式中 $d = T \sum_{t=1}^T t^2 - \left(\sum_{t=1}^T t \right)^2$ 。

由于 y_t 由过程(1.9.2)产生,所以

$$y_t = \alpha + y_{t-1} + u_t = \alpha t + \hat{\epsilon}_t; \quad \hat{\epsilon}_t = u_1 + u_2 + \cdots + u_t$$

这里, $u_t = \sum_{j=1}^{\infty} \varphi_j \epsilon_{t-j}$, $\{\epsilon_t\}$ 为独立同分布, 且 $E(\epsilon_t) = 0$, $D(\epsilon_t) = \sigma^2$ 。代入(1.9.4), 得

$$\begin{aligned} T^{-1/2} \hat{c}_T &= (T^{-4}d)^{-1} \left\{ T^{-3} \left(\sum_{t=1}^T t^2 \right) \left(T^{-3/2} \sum_{t=1}^T \hat{\epsilon}_t \right) \right. \\ &\quad \left. - T^{-2} \left(\sum_{t=1}^T t \right) \left(T^{-5/2} \sum_{t=1}^T t \hat{\epsilon}_t \right) \right\} \quad (1.9.5) \end{aligned}$$

由定理 1.9 的结论(6)和结论(9),得

$$T^{-3/2} \sum_{i=1}^T \xi_i \Rightarrow \sigma\varphi(1) \int_0^1 W(r)dr$$

$$T^{-5/2} \sum_{i=1}^T t \xi_i \Rightarrow \sigma\varphi(1) \int_0^1 rW(r)dr$$

另一方面

$$T^{-3} \sum_{i=1}^T t^2 = T^{-3} \left\{ \frac{1}{6} T(T+1)(2T+1) \right\} \longrightarrow \frac{1}{3}$$

$$T^{-2} \sum_{i=1}^T t = T^{-2} \left\{ \frac{1}{2} T(T+1) \right\} \longrightarrow \frac{1}{2}$$

所以

$$T^{-4}d = T^{-4} \left\{ T \sum_{i=1}^T t^2 - \left(\sum_{i=1}^T t \right)^2 \right\}$$

$$= T^{-3} \sum_{i=1}^T t^2 - \left(T^{-2} \sum_{i=1}^T t \right)^2$$

$$\longrightarrow \frac{1}{12}$$

综上所述

$$T^{-1/2} \hat{c}_T \Rightarrow 12 \left\{ \frac{1}{3} \sigma\varphi(1) \int_0^1 W(r)dr - \frac{1}{2} \sigma\varphi(1) \int_0^1 rW(r)dr \right\}$$

$$= 4\sigma\varphi(1) \int_0^1 W(r)dr - 6\sigma\varphi(1) \int_0^1 rW(r)dr$$

$$= 6\sigma\varphi(1) \int_0^1 \left(\frac{2}{3} - r \right) W(r)dr \quad (1.9.6)$$

$T^{-1/2} \hat{c}_T$ 的极限有非退化的分布,这说明最小二乘估计 \hat{c}_T 的绝

对值以概率发散到无穷大。

最小二乘估计 \hat{r}_T 可表示为

$$\hat{r}_T = d^{-1} \left\{ T \sum_{t=1}^T t y_t - \left(\sum_{t=1}^T t \right) \left(\sum_{t=1}^T y_t \right) \right\}$$

同样以 $y_t = \alpha t + \xi_t$ 代入, 整理后可得

$$\begin{aligned} \sqrt{T}(\hat{r}_T - \alpha) &= (T^{-4}d)^{-1} \left\{ T^{-3/2} \sum_{t=1}^T t \xi_t \right. \\ &\quad \left. T^{-3/2} \left(\sum_{t=1}^T t \right) \left(T^{-3/2} \sum_{t=1}^T \xi_t \right) \right\} \end{aligned}$$

令 $T \rightarrow \infty$, 根据定理 1.7 和以上的讨论, 可得

$$\begin{aligned} \sqrt{T}(\hat{r}_T - \alpha) &\Rightarrow 12\sigma\varphi(1) \left\{ \int_0^1 r W(r) dr - \frac{1}{2} \int_0^1 W(r) dr \right\} \\ &= 12\sigma\varphi(1) \int_0^1 \left(r - \frac{1}{2} \right) W(r) dr \end{aligned}$$

这说明, $(\hat{r}_T - \alpha)$ 以概率收敛于零, 即

$$\hat{r}_T \xrightarrow{P} \alpha$$

这样, 只要数据生成过程 (1.9.2) 中 $\alpha \neq 0$, (1.9.1) 中的参数 r 的最小二乘估计 \hat{r}_T 就有非零的极限, 检验原假设 $H_0: r = 0$ 的 t_T 统计量也会呈显著性。用以上的方法还可进一步证明 $T^{-1/2}t_T$ 有非退化的极限分布, 从而 $|t_T|$ 以概率发散到无穷大。

以上分析在方法论上有重要的意义。对一个从图像上显示出有时间趋势的序列, 仅用 t_T 统计量对 (1.9.1) 中的参数 c 和 r 作假设检验是不够的, 参数 c 和 r 的显著性, 并不能区分 (1.9.1) 和 (1.9.2)。正确的方法是先对序列中的单位根假设作检验, 只有在单位根假设被拒绝后, 才对 (1.9.1) 中的 c 和 r 作通常的 t -检验。

1.10 近单位根过程

若在一阶自回归过程

$$y_t = \rho y_{t-1} + \varepsilon_t \quad (1.10.1)$$

中,参数 ρ 接近于 1,但 $|\rho| < 1$,则称 $\{y_t\}$ 服从一阶近单位根过程 (near unit root process),近单位根过程介于稳定过程和单位根过程之间,对它的研究可以揭示当 ρ 从 $|\rho| < 1$ 趋向于 1 时,过程 (1.10.1) 的统计性质发生突变的过程。

菲利普斯 (Phillips, 1988) 提出了一个模式,其中“单位根过程”和“近单位根过程”用统一的方法处理。这一模式对研究一阶自回归过程从稳定到非稳定的转变,提供了有效的方法。

在一阶自回归过程 (1.10.1) 中,令参数 ρ 为样本量 T 的函数,即

$$\rho = \exp\left(\frac{c}{T}\right) \quad (1.10.2)$$

其中, c 是小于零的实数。当 $T \gg -c$ 时, ρ 是一个小于 1,但很接近 1 的正数。对于给定的 c ,当 $T \rightarrow \infty$ 时, $\rho \rightarrow 1$ 。

菲利普斯关心的是这样一个问题:若参数 ρ 根据 (1.10.2) 随样本量变化,它的最小二乘估计 $\hat{\rho}_T$,在 ρ 从小于 1 而趋于 1 时,将如何变化? 为回答这一问题,先介绍随机过程 $K_c(r)$:

$$K_c(r) = W(r) + c \int_0^r \exp[c(r-s)W(s)] ds \quad (1.10.3)$$

其中, $c < 0$, $W(r)$ 和 $W(s)$ 为标准维纳过程。随机过程 $K_c(r)$ 是维纳过程 $W(r)$ 的一个泛函。当 $\rho \rightarrow 1$ 时, $c \rightarrow 0$,随机过程 $K_c(r)$ 简化成标准维纳过程 $W(r)$ 。 $K_c(r)$ 称为奥恩斯坦—乌伦贝克过程 (Ornstein-Uhlenbeck),是随机微分方程

根据(1.10.4)中的结果,当 $T \rightarrow \infty$ 时, $T(\hat{\rho}_T - 1)$ 有极限:

$$\begin{aligned} T(\hat{\rho}_T - 1) &\rightarrow c + \left\{ \int_0^1 K_c^2(r) dr \right\}^{-1} \left\{ \frac{1}{2} [K_c^2(1) - 1] - \frac{c}{\sigma^2} \int_0^1 K_c^2(r) dr \right\} \\ &= \frac{K_c^2(1) - 1}{2 \int_0^1 K_c^2(r) dr} + c \left(1 - \frac{1}{\sigma^2} \right) \end{aligned} \quad (1.10.7)$$

从(1.10.7)中的极限可引出以下的推论:

(a) 若数据生成过程是一单位根过程,即 $\rho = 1$, 由(1.10.2),这时有 $c = 0$ 和 $K_c(r) = W(r)$, (1.10.7)简化为

$$T(\hat{\rho}_T - 1) \Rightarrow \frac{W^2(1) - 1}{2 \int_0^1 W^2(r) dr} \quad (1.10.8)$$

这和(1.6.9)中的结果是一致的。

(b) 若 $c \neq 0$, 极限(1.10.7)有一非零的位移。特别当 $\sigma^2 \ll 1$ 时, $c \left(1 - \frac{1}{\sigma^2} \right)$ 为一很大的正数。由于 $K_c(r)$ 的方差为 $\frac{1}{2c} [e^{2rc} - 1]$ (Phillips, 1988), 将其在 $c = 0$ 处用泰勒级数展开至 c^2 项, 可得

$$\frac{1}{2c} [e^{2rc} - 1] \approx r(1 + c^2) > r$$

这表明,当 $c \neq 0$ 时, $T(\hat{\rho}_T - 1)$ 的极限不仅有非零的位移,且比 $W(r)$ 有更大的方差。

(c) (1.10.7)中的非零位移 $c \left(1 - \frac{1}{\sigma^2} \right)$ 并不影响 $\hat{\rho}_T$ 的一致收敛,因为 $(\hat{\rho}_T - 1)$ 以概率收敛于零,即

$$\hat{\rho}_T \xrightarrow{P} 1$$

1.11 本章小结

本章讨论了单位根过程的结构和一些主要性质,着重介绍了泛函中心极限定理和连续映照定理,以及这两个定理在有关单位根过程的极限中的应用。

维纳过程在本章的讨论中起关键的作用,许多统计量的极限是维纳过程的泛函。维纳过程在闭区间 $[0, 1]$ 上连续,但几乎处处不可微。

本章着重讨论了一阶自回归过程

$$y_t = \rho y_{t-1} + \epsilon_t$$

中参数 ρ 的最小二乘估计 $\hat{\rho}_T$ 的极限分布。当 $\rho = 1$ 时, $\hat{\rho}_T$ 是超一致的估计量, $(\hat{\rho}_T - 1)$ 以 T^{-1} 的速率收敛于零; $\sqrt{T}(\hat{\rho}_T - 1)$ 的极限是维纳过程的泛函,有非对称的极限分布。

带常数项和不带常数项的单位根过程有不同的统计性质。前者由于受到累积的常数项的作用,其中参数 α 和 ρ 的最小二乘估计 $\hat{\alpha}_T$ 和 $\hat{\rho}_T$ 都有更高的收敛速度,而且有正态的极限分布。

本章还讨论了近单位根过程的性质,介绍了菲利普斯方法,在其中可对单位根过程和近单位根过程作统一的分析,这对理解在稳定过程向单位根过程的转化中,最小二乘估计的极限的变化规律有重要意义。

习 题

1. 对任一给定的实数 $r \in [0, 1]$, 证明:

$$\lim_{T \rightarrow \infty} \frac{[Tr]}{T} = r$$

这里, $[Tr]$ 表示实数 Tr 的整数部分。

2. 设 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_t, \dots$ 为一列独立同分布的随机变量, 且 $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = \sigma^2 < \infty$.

构造随机向量 z_t :

$$z_t = \begin{bmatrix} T^{-1/2} \sum_{t=1}^T \varepsilon_t \\ T^{-3/2} \sum_{t=1}^T t \varepsilon_t \end{bmatrix}$$

证明: 当 $T \rightarrow \infty$ 时, 有

$$z_t \xrightarrow{d} N\left(0, \sigma^2 \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix}\right)$$

3. 设 $\{u_t\}$ 为稳定过程, 并有表示形式

$$u_t = \sum_{j=0}^{\infty} \varphi_j \varepsilon_{t-j},$$

其中, $\{\varepsilon_t\}$ 独立同分布, 且 $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = \sigma^2 < \infty$. 对 $\sum_{i=1}^t u_i$ 作 BN 分解:

$$\sum_{i=1}^t u_i = \varphi(1) \sum_{i=1}^t \varepsilon_i + \eta_t - \eta_0$$

式中 η_t 为

$$\eta_t = \sum_{p=0}^{\infty} \alpha_p \varepsilon_{t-p}; \quad \alpha_p = - \sum_{k=1}^{\infty} \varphi_{p+k}$$

证明:

$$T^{-1} \sum_{i=2}^T (\eta_{i-1} - \eta_0) \varepsilon_i \xrightarrow{p} 0$$

4. 利用随机积分的定义, 证明:

$$\int_0^1 r dW(r) = W\left(\frac{1}{3}\right)$$

这里, $W(r)$ 是定义在 $[0, 1]$ 上的标准维纳过程。

第2章 单位根过程的假设检验

2.1 简介

经济时间序列数据,特别是宏观经济数据,常常呈现明显的时间趋势,但仅从图像上是无法确定数据是由带趋势的稳定过程,还是由带常数项的单位根过程产生的。

在上章的分析中已指出,若对带趋势的稳定过程

$$y_t = c + rt + u_t \quad (2.1.1)$$

中的参数 c 和 r 作 t -检验,即使数据不是由(2.1.1)产生的,而是由带常数项的单位根过程产生的, t_T 统计量仍会呈显著性,但由此决定选取(2.1.1)作为描述数据的模型却是错误的。

正确的方法是在检验时间趋势之前,先确定在时间序列中是否存在单位根。只有在单位根假设被拒绝后,才采用模型(2.1.1),并对其中的参数作假设检验。以下考虑在一阶自回归过程

$$y_t = \rho y_{t-1} + \varepsilon_t \quad (2.1.2)$$

和

$$y_t = \alpha + \rho y_{t-1} + \varepsilon_t \quad (2.1.3)$$

中,对单位根假设 $H_0: \rho = 1$ 作检验。检验基于传统的 $T(\hat{\rho}_T)$ 统计量和 t_T 统计量,它们的极限分布是以下讨论的重点。在

2.2 迪基—福勒(DF)检验法

迪基—福勒的单位根检验法(以下称 DF 检验法)是由迪基(Dickey)和福勒(Fuller)在 70 年代和 80 年代的一系列文章中建立起来的。为理解 DF 方法,先回忆一下在稳定的一阶自回归过程

$$y_t = \rho y_{t-1} + \varepsilon_t$$

中,对参数 ρ 的检验方法。若这时 $|\rho| < 1$, $\{\varepsilon_t\}$ 独立同分布,且 $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = \sigma^2 < \infty$, 检验假设

$$H_0: \rho = \rho_0; \quad H_1: \rho \neq \rho_0$$

的 t_T 统计量由下式给出:

$$t_T = \frac{\hat{\rho}_T - \rho_0}{\hat{\eta}_T}$$

其中, $\hat{\rho}_T$ 为 ρ 的最小二乘估计, $\hat{\eta}_T$ 为 $\hat{\rho}_T$ 的标准差的估计值。当原假设 $H_0: \rho = \rho_0$ 为真时,即数据由 $y_t = \rho_0 y_{t-1} + \varepsilon_t$ 产生,统计量 t_T 有自由度为 $T-1$ 的 t -分布。选取一适当的显著水平 α (如 5%、1% 等),可从 t -分布表中查得相应的临界值 $t_{\alpha/2}$ 。若 $|t_T| < t_{\alpha/2}$, 接受原假设 $H_0: \rho = \rho_0$; 否则,拒绝原假设 H_0 ,接受备选假设 $H_1: \rho \neq \rho_0$ 。

这种方法不能直接用来检验假设 $H_0: \rho = 1$, 因为当假设 $H_0: \rho = 1$ 为真时,最小二乘估计 $\hat{\rho}_T$ 和 t_T 有非标准和非对称的极限分布,只能用模拟方法计算出它们的分布值。这两个统计量不存在小样本分布,只有当样本量 T 足够大时,它们的极限分布才有实际的应用价值。

为了克服这些困难,迪基和福勒用模拟的方法对统计量 t_T 和 $T(\hat{\rho}_T - 1)$ 的极限分布值作了计算,并采用了不同的样本量,如 $T = 25, 50, 100$ 等。这些模拟值收集在书后的表 1 和表 2 中,表 1 给出了 $T(\hat{\rho}_T - 1)$ 分布的临界值,表 2 给出了 t_T 统计量的临界值。此两表适用于迪基—福勒和菲利普斯—配荣检验,对于给定的样本量 T 和显著水平 α ,在表 1 和表 2 中找到相应于 $T(\hat{\rho}_T - 1)$ 和 t_T 的临界值。若统计量的实际计算值小于临界值,则拒绝单位根假设;若大于或等于临界值,则接受单位根假设 $H_0: \rho = 1$ 。注意到这里采用的是单边检验法,即只在统计量的实际值小于临界值时拒绝原假设,这是因为这里选择的备选假设 $H_1: \rho < 1$ 是单边的。

以下分别考虑四种情况下的 DF 检验。

2.2.1 情况一和情况二的 DF 检验法

在这两种情况下都假设数据是由一阶自回归过程

$$y_t = \rho y_{t-1} + \varepsilon_t \quad (2.2.1)$$

产生的,其中 $\{\varepsilon_t\}$ 独立同分布,且 $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = \sigma^2 < \infty$ 。对于随机干扰 ε_t 的更一般的假设在以后章节中讨论。我们所关心的是检验原假设 $H_0: \rho = 1$ 和备选假设 $H_1: \rho < 1$ 。在原假设成立时, $\{y_t\}$ 服从一随机游动过程;在备选假设成立时, $\{y_t\}$ 服从一稳定的一阶自回归过程。

在情况一中,所用的模型为(2.2.1),参数 ρ 的最小二乘估计 $\hat{\rho}_T$ 为

$$\hat{\rho}_T = \frac{\sum_{t=1}^T y_{t-1} y_t}{\sum_{t=1}^T y_{t-1}^2}$$

相应的临界值,则拒绝原假设 $H_0: \rho = 1$ 。

例 2.1 利用 1947 年第二季度到 1989 年第一季度的数据,可对美国财政部债券的利息率作不带常数项的一阶自回归模型

$$i_t = 0.99694 i_{t-1} \quad (2.2.5) \\ (0.01059)$$

括弧中的数字为 $\hat{\rho}_T$ 的标准差。以下分别用统计量 $T(\hat{\rho}_T - 1)$ 和 t_T 对 $H_0: \rho = 1$ 作检验。

因为 $T = 168$, $\hat{\rho}_T = 0.99694$, $\hat{\eta}_T = 0.01059$, 所以

$$T(\hat{\rho}_T - 1) = 168(0.99694 - 1) = -0.51$$

选取显著水平 $\alpha = 5\%$, 从表 1 的情况一查得相应的临界值为 -7.9 。由于 $-0.51 > -7.9$, 接受原假设 $H_0: \rho = 1$ 。

利用这些数据还可计算统计量 t_T 的实际值

$$t_T = \frac{\hat{\rho}_T - 1}{\hat{\eta}_T} = \frac{0.99694 - 1}{0.01059} = -0.29$$

从表 2 的情况一查得相应于 $\alpha = 5\%$ 的临界值为 -1.95 。由于 $-0.29 > -1.95$, 再次接受原假设 $H_0: \rho = 1$ 。

模型(2.2.5)没有包含常数项,这是因为任一非零的常数隐含着利率以一个固定的速率增加或减少,从而会使模型失去经济学上的意义。但是,若不存在来自经济学意义上或其他意义上的约束条件,一般采用更为普通的模型

$$y_t = \alpha + \rho y_{t-1} + \varepsilon_t \quad (2.2.6)$$

并在其中检验假设:

$$H_0: \alpha = 0, \rho = 1$$

$$H_1: \alpha \neq 0, \rho < 1$$

这是所需讨论的情况二。

模型(2.2.6)的最小二乘估计为

$$\begin{pmatrix} \hat{\alpha}_T \\ \hat{\rho}_T \end{pmatrix} = \begin{pmatrix} T & \sum y_{t-1} \\ \sum y_{t-1} & \sum y_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum y_t \\ \sum y_{t-1} y_t \end{pmatrix}$$

当原假设 $H_0: \alpha = 0, \rho = 1$ 成立时, 可将上式改写成

$$\begin{pmatrix} \hat{\alpha}_T \\ \hat{\rho}_T - 1 \end{pmatrix} = \begin{pmatrix} T & \sum y_{t-1} \\ \sum y_{t-1} & \sum y_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum \epsilon_t \\ \sum y_{t-1} \epsilon_t \end{pmatrix}$$

以 $T^{1/2}$ 和 T 分别乘 $\hat{\alpha}_T$ 和 $\hat{\rho}_T - 1$, 得

$$\begin{aligned} & \begin{pmatrix} T^{1/2} & 0 \\ 0 & T \end{pmatrix} \begin{pmatrix} \hat{\alpha}_T \\ \hat{\rho}_T - 1 \end{pmatrix} \\ &= \left\{ \begin{pmatrix} T^{-1/2} & 0 \\ 0 & T^{-1} \end{pmatrix} \begin{pmatrix} T & \sum y_{t-1} \\ \sum y_{t-1} & \sum y_{t-1}^2 \end{pmatrix} \begin{pmatrix} T^{-1/2} & 0 \\ 0 & T^{-1} \end{pmatrix} \right\}^{-1} \\ & \quad \times \left\{ \begin{pmatrix} T^{-1/2} & 0 \\ 0 & T^{-1} \end{pmatrix} \begin{pmatrix} \sum \epsilon_t \\ \sum y_{t-1} \epsilon_t \end{pmatrix} \right\} \end{aligned}$$

也即

$$\begin{pmatrix} T^{1/2} \hat{\alpha}_T \\ T(\hat{\rho}_T - 1) \end{pmatrix} = \begin{pmatrix} 1 & T^{-3/2} \sum y_{t-1} \\ T^{-3/2} \sum y_{t-1} & T^{-2} \sum y_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} T^{-1/2} \sum \epsilon_t \\ T^{-1} \sum y_{t-1} \epsilon_t \end{pmatrix} \quad (2.2.7)$$

在原假设成立时, $y_t = \sum_{t=1}^T \epsilon_t$, 根据定理 1.7

$$T^{-3/2}\Sigma y_{t-1} \Rightarrow \sigma \int_0^1 W(r) dr$$

$$T^{-1/2}\Sigma \varepsilon_t \Rightarrow \sigma W(1)$$

$$T^{-2}\Sigma y_{t-1}^2 \Rightarrow \sigma^2 \int_0^1 W^2(r) dr$$

$$T^{-1}\Sigma y_{t-1}\varepsilon_t \Rightarrow \frac{1}{2}\sigma^2(W^2(1) - 1)$$

从而有

$$\begin{aligned} \begin{pmatrix} 1 & T^{-3/2}\Sigma y_{t-1} \\ T^{-3/2}\Sigma y_{t-1} & T^{-2}\Sigma y_{t-1}^2 \end{pmatrix} &\Rightarrow \begin{pmatrix} 1 & \sigma \int_0^1 W(r) dr \\ \sigma \int_0^1 W(r) dr & \sigma^2 \int_0^1 W^2(r) dr \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & \sigma^{-1} \end{pmatrix} \begin{pmatrix} 1 & \int_0^1 W(r) dr \\ \int_0^1 W(r) dr & \int_0^1 W^2(r) dr \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 \\ 0 & \sigma^{-1} \end{pmatrix} \\ \begin{pmatrix} T^{-1/2}\Sigma \varepsilon_t \\ T^{-1}\Sigma y_{t-1}\varepsilon_t \end{pmatrix} &\Rightarrow \begin{pmatrix} \sigma W(1) \\ \frac{1}{2}\sigma^2(W^2(1) - 1) \end{pmatrix} \\ &= \begin{pmatrix} \sigma & 0 \\ 0 & \sigma^2 \end{pmatrix} \begin{pmatrix} W(1) \\ \frac{1}{2}(W^2(1) - 1) \end{pmatrix} \end{aligned}$$

综上所述,可得 $T^{1/2}\hat{a}_T$ 和 $T(\hat{\rho}_T - 1)$ 的极限

$$\begin{pmatrix} T^{1/2}\hat{a}_T \\ T(\hat{\rho}_T - 1) \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 \\ 0 & \sigma^{-1} \end{pmatrix} \begin{pmatrix} 1 & \int_0^1 W(r) dr \\ \int_0^1 W(r) dr & \int_0^1 W^2(r) dr \end{pmatrix}^{-1}$$

统计量 $T(\hat{\rho}_T - 1)$ 的极限 (2.2.9) 与情况一中的极限 (2.2.2) 有不同的分布, 它的临界值列在表 1 的情况二中。作为检验原假设 $H_0: \rho = 1$ 的统计量, $T(\hat{\rho}_T - 1)$ 的作用与情况一一样。

检验 $H_0: \rho = 1$ 的统计量 t_T 与情况一中的 t_T 有一样的形式:

$$t_T = \frac{\hat{\rho}_T - 1}{\hat{\eta}_T}$$

式中的标准差 $\hat{\eta}_T$ 为

$$\hat{\eta}_T^2 = \hat{\sigma}_T^2 [0, 1] \begin{bmatrix} T & \Sigma y_{t-1} \\ \Sigma y_{t-1} & \Sigma y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

其中的 $\hat{\sigma}_T^2$ 是 σ^2 的最小二乘估计:

$$\hat{\sigma}_T^2 = \frac{1}{T-2} \sum_{t=1}^T (y_t - \hat{a}_T - \hat{\rho}_T y_{t-1})^2$$

以 T^2 乘 $\hat{\eta}_T^2$, 整理后得

$$\begin{aligned} T^2 \hat{\eta}_T^2 &= \hat{\sigma}_T^2 [0, T] \begin{bmatrix} T & \Sigma y_{t-1} \\ \Sigma y_{t-1} & \Sigma y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ T \end{bmatrix} \\ &= \hat{\sigma}_T^2 [0, 1] \begin{bmatrix} 1 & T^{-3/2} \Sigma y_{t-1} \\ T^{-3/2} \Sigma y_{t-1} & T^{-2} \Sigma y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{aligned}$$

因为

$$\begin{aligned} &\begin{bmatrix} 1 & T^{-3/2} \Sigma y_{t-1} \\ T^{-3/2} \Sigma y_{t-1} & T^{-2} \Sigma y_{t-1}^2 \end{bmatrix} \\ \Rightarrow &\begin{bmatrix} 1 & 0 \\ 0 & \sigma^{-1} \end{bmatrix} \begin{bmatrix} 1 & \int_0^1 W(r) dr \\ \int_0^1 W(r) dr & \int_0^1 W^2(r) dr \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & \sigma^{-1} \end{bmatrix} \end{aligned}$$

所以, $T^2 \hat{\eta}_T^2$ 有以下极限

$$\begin{aligned}
 & T^2 \hat{\eta}_T^2 \\
 \Rightarrow & \sigma^2 [0, 1] \begin{bmatrix} 1 & 0 \\ 0 & \sigma^{-1} \end{bmatrix} \begin{bmatrix} 1 & \int_0^1 W(r) dr \\ \int_0^1 W(r) dr & \int_0^1 W^2(r) dr \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & \sigma^{-1} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
 = & [0, 1] \begin{bmatrix} 1 & \int_0^1 W(r) dr \\ \int_0^1 W(r) dr & \int_0^1 W^2(r) dr \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
 = & \frac{1}{\int_0^1 W^2(r) dr - \left(\int_0^1 W(r) dr \right)^2} \quad (2.2.10)
 \end{aligned}$$

$T(\hat{\rho}_T - 1)$ 和 $T^2 \hat{\eta}_T^2$ 的极限给出了统计量 t_T 的极限:

$$\begin{aligned}
 t_T = \frac{T(\hat{\rho}_T - 1)}{(T^2 \hat{\eta}_T^2)^{1/2}} & \Rightarrow \frac{\frac{1}{2}(W^2(1) - 1) - W(1) \int_0^1 W(r) dr}{\int_0^1 W^2(r) dr - \left(\int_0^1 W(r) dr \right)^2} \\
 & \times \left\{ \int_0^1 W^2(r) dr - \left(\int_0^1 W(r) dr \right)^2 \right\}^{1/2} \\
 = & \frac{\frac{1}{2}(W^2(1) - 1) - W(1) \int_0^1 W(r) dr}{\left\{ \int_0^1 W^2(r) dr - \left(\int_0^1 W(r) dr \right)^2 \right\}^{1/2}} \quad (2.2.11)
 \end{aligned}$$

与情况一相比, 情况二中的 t_T 统计量有较为复杂的极限, 它的分布的临界值列在表 2 的情况二中。 t_T 在检验 $\rho = 1$ 中的作用与情况一一样。

从表 1 和表 2 中可看出, 情况二中的值都小于相应的情况一中的值, 这表明对于带常数的单位根过程, $T(\hat{\rho}_T - 1)$ 和 t_T

的极限分布更偏向于左侧。以 $\alpha = 0.05$, $T = 100$ 为例, 表 1 的情况一、二中的临界值分别为 -7.9 和 -13.7 , 而表 2 中相应的临界值分别为 -1.95 和 -2.89 。

例 2.2 利用例 2.1 中的数据, 估计带常数项的一阶自回归模型

$$i_t = \begin{matrix} 0.211 & + & 0.96691 & i_{t-1} \\ (0.112) & & (0.01933) \end{matrix}$$

并在其中检验假设 $H_0: \rho = 1$ 。给定 $T = 168$, $T(\hat{\rho}_T - 1)$ 的实际值为

$$T(\hat{\rho}_T - 1) = 168(0.96691 - 1) = -5.56$$

由表 1 情况二可查得相应于 $\alpha = 0.05$ 的临界值为 -13.8 , 因为 $-5.56 > -13.8$, 所以接受原假设 $H_0: \rho = 1$ 。

统计量 t_T 的实际值可由下式计算:

$$t_T = \frac{\hat{\rho}_T - 1}{\hat{\eta}_T} = \frac{0.96691 - 1}{0.01933} = -1.71$$

由表 2 情况二查得相应于 $\alpha = 0.05$ 的临界值为 -2.89 , 因为 $-1.71 > -2.89$, 所以又一次接受单位根假设 $H_0: \rho = 1$ 。

严格地说, 在上例中还需对假设 $\alpha = 0$ 作检验, 因为我们所用的 $T(\hat{\rho}_T - 1)$ 与 t_T 的极限 (2.2.9) 和 (2.2.11) 在 $\rho = 1$ 和 $\alpha = 0$ 同为真时才成立。但 $T^{1/2}\hat{\alpha}_T$ 的极限 (2.2.8) 不能用来检验 $\alpha = 0$, 因为其中包含了未知参数 σ 。

原假设 $H_0: \rho = 1, \alpha = 0$ 可以用 F_T 统计量作检验。在模型

$$y_t = \alpha + \rho y_{t-1} + \varepsilon_t$$

中, 对假设 $H_0: \rho = \rho_0, \alpha = 0$ 的 F_T 统计量为

$$F_T = \frac{(\bar{R}^2 - \hat{R}^2)/2}{\hat{R}^2/(T-2)}$$

其中的 \bar{R}^2 和 \hat{R}^2 分别为有约束的和无约束的残差平方和：

$$\bar{R}^2 = \sum_{t=1}^T (y_t - \rho_0 y_{t-1})^2$$

$$\hat{R}^2 = \sum_{t=1}^T (y_t - \hat{\alpha}_T - \hat{\rho}_T y_{t-1})^2$$

当 $|\rho_0| < 1$ 时, $\{y_t\}$ 为稳定过程, 在原假设 H_0 下, F_T 统计量有自由度为 $(2, T-2)$ 的 F -分布。当 $\rho_0 = 1$ 时, F_T 的极限存在, 但有非标准的分布。迪基和福勒将 F_T 极限分布的模拟值列在表 3 中。

例 2.3 在例 2.2 的估计模型

$$i_t = \begin{matrix} 0.211 & + & 0.96691 & i_{t-1} \\ (0.112) & & (0.01933) \end{matrix}$$

中, 检验假设 $H_0: \rho = 1, \alpha = 0$ 。 F_T 统计量的实际值为 1.81, 从表 3 的情况二中查得相应于 $\alpha = 0.05$ 的临界值为 4.67, 因为 $1.81 < 4.67$, 我们接受联合假设 $H_0: \alpha = 0, \rho = 1$ 。

2.2.2 情况三和情况四的 DF 检验法

和情况一、二不同, 在情况三、四中, 我们假设数据由带常数项的单位根过程

$$y_t = \alpha + \rho y_{t-1} + \varepsilon_t \quad (2.2.12)$$

生成, 其中 $\{\varepsilon_t\}$ 独立同分布, $E(\varepsilon_t) = 0, D(\varepsilon_t) = \sigma^2 < \infty$ 。

情况三考虑在 (2.2.12) 中检验单位根假设。对于最小二乘估计量 $\hat{\alpha}_T$ 和 $\hat{\rho}_T$, 上章中的分析表明, $T^{1/2}(\hat{\alpha}_T - \alpha)$ 和 $T^{3/2}(\hat{\rho}_T -$

1) 都有正态的极限分布(参见 1.7.11 式及以后的讨论)。因此, 情况三的单位根假设 $H_0: \rho = 1$ 可从一般的正态分布和 t -分布表中得到临界值。

情况四考虑在估计模型

$$y_t = \alpha + \rho y_{t-1} + \delta t + \varepsilon_t \quad (2.2.13)$$

中, 对单位根假设作检验。以下先讨论对假设 $\rho = 1$ 的检验, 然后再讨论对假设 $\rho = 1, \delta = 0$ 的联合检验。

当假设 $\rho = 1$ 和 $\delta = 0$ 成立时, 可将 y_t 表示为

$$y_t = \alpha t + y_0 + \xi_t$$

这里

$$\xi_t = \varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_t = \sum_{j=1}^t \varepsilon_j$$

由于(2.2.13)中的 ρy_{t-1} 和 δt 都含有时间趋势, 用 y_t 对常数项、 y_{t-1} 和 t 作回归会产生共线性问题(collinearity)。为了对(2.2.13)中的参数作最小二乘估计, 先对模型作以下变换:

$$\begin{aligned} y_t &= \alpha + \rho y_{t-1} + \delta t + \varepsilon_t \\ &= (1 - \rho)\alpha + \rho[y_{t-1} - \alpha(t-1)] + (\delta + \rho\alpha)t + \varepsilon_t \\ &= \alpha^* + \rho^* \xi_{t-1} + \delta^* t + \varepsilon_t \end{aligned} \quad (2.2.14)$$

这里, $\alpha^* = (1 - \rho)\alpha$, $\rho^* = \rho$, $\delta^* = \delta + \rho\alpha$, $\xi_t = y_t - \alpha t$ 。因此, 原假设

$$H_0: \alpha = \alpha_0, \rho = 1, \delta = 0$$

等价于

$$H_0: \alpha^* = 0, \rho^* = 1, \delta^* = \alpha_0 \quad (2.2.15)$$

经整理后可简化为

$$\begin{aligned}
 & \begin{pmatrix} T^{1/2} \hat{\alpha}_T^* \\ T(\hat{\rho}_T^* - 1) \\ T^{3/2}(\hat{\delta}_T^* - \alpha_0) \end{pmatrix} \\
 &= \begin{pmatrix} 1 & T^{-3/2} \Sigma \hat{\xi}_{t-1} & T^{-2} \Sigma t \\ T^{-3/2} \Sigma \hat{\xi}_{t-1} & T^{-2} \Sigma \hat{\xi}_{t-1}^2 & T^{-5/2} \Sigma \hat{\xi}_{t-1} t \\ T^{-2} \Sigma t & T^{-5/2} \Sigma \hat{\xi}_{t-1} t & T^{-3} \Sigma t^2 \end{pmatrix}^{-1} \begin{pmatrix} T^{-1/2} \Sigma \epsilon_t \\ T^{-1} \Sigma \hat{\xi}_{t-1} \epsilon_t \\ T^{-3/2} \Sigma t \epsilon_t \end{pmatrix}
 \end{aligned}$$

由定理 1.7 中的结果, 可得以下极限:

$$\begin{aligned}
 & \begin{pmatrix} T^{1/2} \hat{\alpha}_T^* \\ T(\hat{\rho}_T^* - 1) \\ T^{3/2}(\hat{\delta}_T^* - \alpha_0) \end{pmatrix} \\
 &\Rightarrow \begin{pmatrix} 1 & \sigma \int_0^1 W(r) dr & \frac{1}{2} \\ \sigma \int_0^1 W(r) dr & \sigma^2 \int_0^1 W^2(r) dr & \sigma \int_0^1 r W(r) dr \\ \frac{1}{2} & \sigma \int_0^1 r W(r) dr & \frac{1}{3} \end{pmatrix}^{-1} \\
 &\quad \times \begin{pmatrix} \sigma W(1) \\ \frac{1}{2} \sigma^2 \{W^2(1) - 1\} \\ \sigma \{W(1) - \int_0^1 W(r) dr\} \end{pmatrix} \\
 &= \sigma \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & \int_0^1 W(r) dr & \frac{1}{2} \\ \int_0^1 W(r) dr & \int_0^1 W^2(r) dr & \int_0^1 r W(r) dr \\ \frac{1}{2} & \int_0^1 r W(r) dr & \frac{1}{3} \end{pmatrix}^{-1}
 \end{aligned}$$

$$\begin{aligned}
& \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} W(1) \\ \frac{1}{2} \{W^2(1) - 1\} \\ W(1) - \int_0^1 W(r) dr \end{bmatrix} \\
& = \begin{bmatrix} \sigma & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sigma \end{bmatrix} \begin{bmatrix} 1 & \int_0^1 W(r) dr & \frac{1}{2} \\ \int_0^1 W(r) dr & \int_0^1 W^2(r) dr & \int_0^1 r W(r) dr \\ \frac{1}{2} & \int_0^1 r W(r) dr & \frac{1}{3} \end{bmatrix}^{-1} \\
& \times \begin{bmatrix} W(1) \\ \frac{1}{2} \{W^2(1) - 1\} \\ W(1) - \int_0^1 W(r) dr \end{bmatrix} \quad (2.2.16)
\end{aligned}$$

值得注意的是,上式中只有 $T(\hat{\rho}_T - 1)$ (也即 $T(\hat{\rho}_T - 1)$) 的极限不含有未知参数 σ , 可以用来检验假设 $\rho = 1$ 。在推导统计量 $T(\hat{\rho}_T - 1)$ 和 t_T 的极限前, 先考虑 $\hat{\rho}_T$ 的标准差 $\hat{\eta}_T$ 的极限分布。 $\hat{\eta}_T$ 由下式给出:

$$\hat{\eta}_T^2 = \hat{\sigma}_T^2 [0, 1, 0] \begin{bmatrix} T & \Sigma \xi_{t-1} & \Sigma t \\ \Sigma \xi_{t-1} & \Sigma \xi_{t-1}^2 & \Sigma \xi_{t-1} t \\ \Sigma t & \Sigma \xi_{t-1} t & \Sigma t^2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

以 T^2 乘上式, 可得

$$T^2 \hat{\eta}_T^2 = \hat{\sigma}_T^2 [0, 1, 0] \begin{bmatrix} T^{1/2} & 0 & 0 \\ 0 & T & 0 \\ 0 & 0 & T^{3/2} \end{bmatrix} \begin{bmatrix} T & \Sigma \xi_{t-1} & \Sigma t \\ \Sigma \xi_{t-1} & \Sigma \xi_{t-1}^2 & \Sigma \xi_{t-1} t \\ \Sigma t & \Sigma \xi_{t-1} t & \Sigma t^2 \end{bmatrix}^{-1}$$

其中 $|A|$ 表示矩阵

$$A = \begin{pmatrix} 1 & \int_0^1 W(r)dr & \frac{1}{2} \\ \int_0^1 W(r)dr & \int_0^1 W^2(r)dr & \int_0^1 rW(r)dr \\ \frac{1}{2} & \int_0^1 rW(r)dr & \frac{1}{3} \end{pmatrix}$$

的行列式。

有了 $T^2 \hat{\eta}_T^2$ 的极限,就可进一步计算统计量 $T(\hat{\rho}_T - 1)$ 和 t_T 的极限。从 $T^{1/2} \hat{\alpha}_T^*$, $T(\hat{\rho}_T^* - 1)$ 和 $T^{3/2}(\hat{\delta}_T^* - \alpha_0)$ 的极限 (2.2.16) 中,可用代数余子式的方法抽取出 $T(\hat{\rho}_T^* - 1)$, 即 $T(\hat{\rho}_T - 1)$ 的极限:

$$\begin{aligned} T(\hat{\rho}_T - 1) &\Rightarrow \frac{1}{|A|} \left\{ \frac{1}{6} W(1) \int_0^1 W(r)dr - \frac{1}{2} W(1) \int_0^1 rW(r)dr \right. \\ &\quad \left. + \frac{1}{24} (W^2(1) - 1) + \int_0^1 W(r)dr \int_0^1 rW(r)dr \right. \\ &\quad \left. - \frac{1}{2} \left(\int_0^1 W(r)dr \right)^2 \right\} \end{aligned} \quad (2.2.19)$$

对于原假设 $H_0: \rho = 1$, 统计量 t_T 由下式给出:

$$t_T = \frac{\hat{\rho}_T - 1}{\hat{\eta}_T} = \frac{T(\hat{\rho}_T - 1)}{(T^2 \hat{\eta}_T^2)^{1/2}}$$

由 Q 的表示形式 (2.2.18), 统计量 t_T 的极限为

$$t_T = \frac{T(\hat{\rho}_T - 1)}{(T^2 \hat{\eta}_T^2)^{1/2}} \Rightarrow \frac{\sqrt{3}}{12} |A|^{-\frac{1}{2}} \left\{ 4W(1) \int_0^1 W(r)dr \right.$$

$$\begin{aligned}
& - 12W(1)\int_0^1 rW(r)dr + (W^2(1) - 1) \\
& + 24\int_0^1 W(r)dr\int_0^1 rW(r)dr - 12\left(\int_0^1 W(r)dr\right)^2 \Big\} \\
& \qquad \qquad \qquad (2.2.20)
\end{aligned}$$

$T(\hat{\rho}_T - 1)$ 和 t_T 的极限都有非标准的分布,但都不依赖于未知参数 α_0 和 σ^2 ,所以可以用来检验假设 $H_0: \rho = 1$, 而不受到其他参数影响。 $T(\hat{\rho}_T - 1)$ 和 t_T 的临界值也由模拟方法计算得到,分别列在表 1 和表 2 的情况四中。

联合假设 $H_0: \rho = 1, \delta = 0$ 可由 F_T 统计量检验,它由下式给出:

$$F_T = \frac{(\tilde{R}^2 - \hat{R}^2)/2}{\hat{R}^2/(T-3)} \quad (2.2.21)$$

分子中的 2 表示假设中受约束的个数,分母中的 3 表示式中所有参数的个数。当联合假设成立时, F_T 统计量的极限是维纳过程的泛函,它的分布是非标准的。表 3 的情况四中列出了 F_T 的临界值。

2.2.3 迪基—福勒(DF)检验法小结

迪基—福勒方法基于参数的最小二乘估计。在单位根过程(带常数项或不带常数项)中,由给定的样本构造统计量 $T(\hat{\rho}_T - 1)$ 、 t_T 或 F_T ,可计算出它们的实际值,并分别与表 1、表 2 或表 3 中的临界值相比较。若实际值明确地小于临界值,则拒绝单位根假设 $H_0: \rho = 1$ 。运用迪基—福勒检验方法的成败,还取决于如何选择正确的原假设和备用假设,因为统计量 $T(\hat{\rho}_T - 1)$ 、 t_T 和 F_T 不仅决定于数据生成过程的形式(带常数项或不带常数

项的单位根过程),还决定于估计模型的形式。在运用迪基—福勒方法前作以下考虑是有必要的。首先,对于现有的数据,尽量从经济学和统计学的角度考虑其数据生成过程可能有的形式。比如,在处理利息率数据时(见例 2.1 和 2.2),应仔细考虑包括常数项在经济学上是否合理。若经济学上、统计学上和其他方面的考虑不能得出明确的结论,则应尽量采用较为一般的模型,如情况二和四,作为数据生成过程和估计模型,然后在其中作单位根的假设检验。为了加深对迪基—福勒方法的理解,让我们再次考虑例 2.1 和 2.2 中所用的数据。图 2.1 给出了美国财政部季度债券从 1947 年一季度到 1989 年二季度的名义利率的数据。

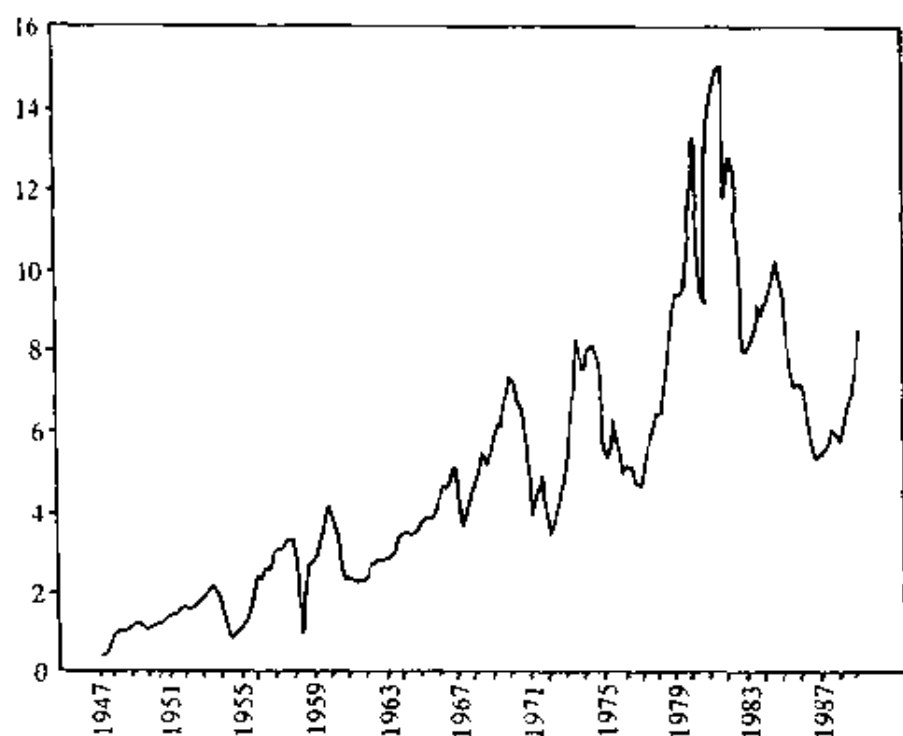


图 2.1 美国财政部债券名义利率:1947(1)—1989(2)

从图像上看,数据中似乎存在明显的上升趋势;但从经济学的观点,却没有任何理由认为名义利率含有确定性的趋势(见例 2.1 中的讨论)。基于这两方面的考虑,似乎应以不带常数项的单位根过程

$$i_t = \rho i_{t-1} + \epsilon_t \quad (2.2.22)$$

作为数据生成过程和估计模型,并在其中检验 $H_0: \rho = 1$ 。但这使得我们在选择备选假设 H_1 时遇到了困难,因为图 2.1 中的数据显然有大于零的平均值,所以,如果通常的备选假设 $H_1: \rho < 1$ 成立,还必须在(2.2.22)中包含一正常数项,解释图中的趋势,即在

$$i_t = \alpha + \rho i_{t-1} + \epsilon_t \quad (2.2.23)$$

中,假设 $\alpha > 0$ 和 $\rho < 1$ 同时成立。综合以上分析,合理的检验方法应是根据情况二,假设(2.2.22)为数据生成过程,但以(2.2.23)作为估计模型,并在其中检验 $H_0: \rho = 1$ 和 $H_1: \rho < 1$,或者 $H_0: \alpha = 0, \rho = 1$ 和 $H_1: \rho < 1, \alpha \neq 0$ 。

再考虑图 2.2 中的数据,其中给出的是 1947 年一季度至 1989 年二季度美国国民总产值的实际值,即已去除了价格指数的影响。

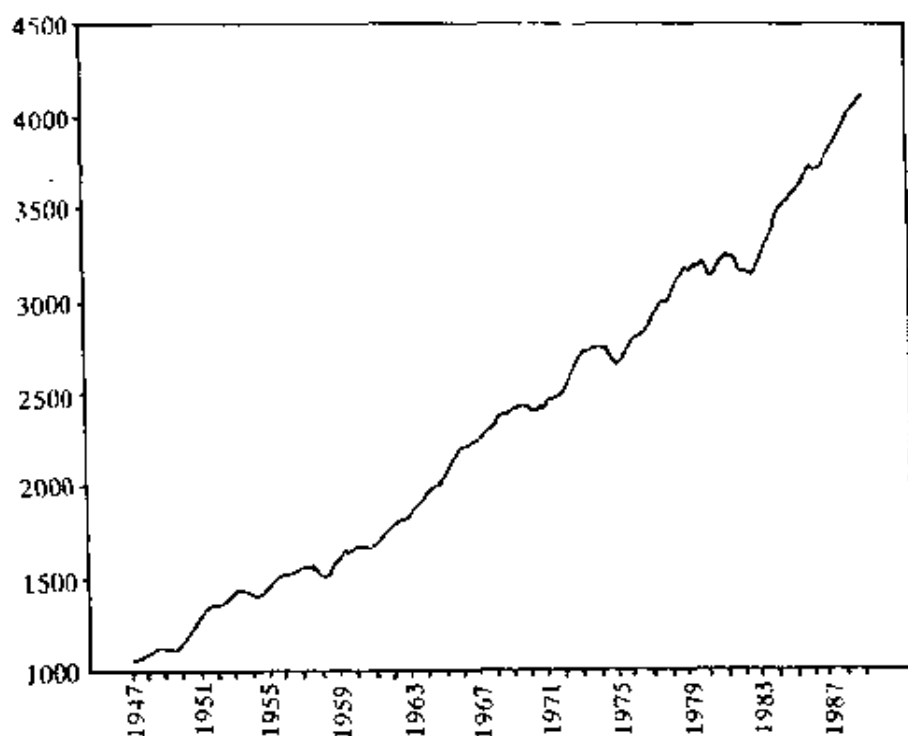


图 2.2 美国国民总产值:1947(1)—1989(2)

$$t_T = \frac{\hat{\rho}_T - 1}{\hat{\eta}_T} = \frac{0.96252 - 1}{0.0193} = -1.94$$

由表 2 的情况四中查得相应于 $T = 168$ 、 $\alpha = 5\%$ 的临界值为 -3.44 ，因为 $-1.94 > -3.44$ ，我们又一次在 5% 的水平上接受原假设 $H_0: \rho = 1$ 。

最后，对假设 $H_0: \rho = 1, \delta = 0$ 作联合检验。统计量 F_T 由公式(2.2.21)给出，在本例中它的实际值为 2.44。根据表 3 的情况四，相应于 5% 显著水平的临界值为 6.42。因为 $2.44 < 6.42$ ，接受联合假设 $H_0: \rho = 1, \delta = 0$ 。

在以上两例中，若选择不合适的模型作为数据生成过程或估计模型，不仅会导致错误的统计结论，还会引起不正确的经济解释。

2.3 菲利普斯—配荣(PP)检验法

菲利普斯—配荣方法(以下简称 PP 方法)是菲利普斯(1987)和菲利普斯、配荣(1988)提出的。PP 方法处理的是具有一般形式的单位根过程，其中的随机干扰 u_t 服从一稳定过程。以下先讨论情况二中的 PP 检验法，然后讨论情况一和情况四的 PP 方法。由于情况三中的最小二乘估计有正态的极限分布(见前节中的讨论)，在此不作专门讨论。

2.3.1 情况二的 PP 检验法

情况二考虑的是带常数项的单位根过程

$$y_t = \alpha + \rho y_{t-1} + u_t \quad (2.3.1)$$

其中， $\rho = 1$ ， $\{u_t\}$ 服从一稳定过程，可表示为

$$u_t = \varphi(L)\varepsilon_t = \sum_{j=0}^{\infty} \varphi_j \varepsilon_{t-j}$$

这里, $\{\varepsilon_t\}$ 独立同分布, $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = \sigma^2 < \infty$ 。 $\varphi(L)$ 为无穷阶的滞后多项式, 其系数满足条件 $\sum_{j=0}^{\infty} j|\varphi_j| < \infty$ 。

与迪基—福勒方法一样, PP 方法也基于参数 α 和 ρ 的最小二乘估计:

$$\begin{bmatrix} \hat{\alpha}_T \\ \hat{\rho}_T \end{bmatrix} = \begin{bmatrix} T & \Sigma y_{t-1} \\ \Sigma y_{t-1} & \Sigma y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \Sigma y_t \\ \Sigma y_{t-1} y_t \end{bmatrix} \quad (2.3.2)$$

当原假设 $H_0: \alpha = 0, \rho = 1$ 为真时, 因为

$$\Delta y_t = y_t - y_{t-1} = u_t = \varphi(L)\varepsilon_t = \sum_{j=0}^{\infty} \varphi_j \varepsilon_{t-j}$$

可将(2.3.2)改写成

$$\begin{bmatrix} T^{1/2} \hat{\alpha}_T \\ T(\hat{\rho}_T - 1) \end{bmatrix} = \begin{bmatrix} 1 & T^{-3/2} \Sigma y_{t-1} \\ T^{-3/2} \Sigma y_{t-1} & T^{-2} \Sigma y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} T^{-1/2} \Sigma u_t \\ T^{-1} \Sigma y_{t-1} u_t \end{bmatrix} \quad (2.3.3)$$

根据定理 1.7 的结论

$$T^{-3/2} \sum_{t=1}^T y_{t-1} \Rightarrow \sigma \varphi(1) \int_0^1 W(r) dr$$

$$T^{-2} \sum_{t=1}^T y_{t-1}^2 \Rightarrow \sigma^2 \varphi^2(1) \int_0^1 W^2(r) dr$$

从而有

$$\begin{bmatrix} 1 & T^{-3/2} \Sigma y_{t-1} \\ T^{-3/2} \Sigma y_{t-1} & T^{-2} \Sigma y_{t-1}^2 \end{bmatrix}^{-1}$$

$$\begin{aligned}
& \Rightarrow \begin{bmatrix} 1 & \sigma\varphi(1)\int_0^1 W(r)dr \\ \sigma\varphi(1)\int_0^1 W(r)dr & \sigma^2\varphi^2(1)\int_0^1 W^2(r)dr \end{bmatrix}^{-1} \\
& = \begin{bmatrix} 1 & 0 \\ 0 & \sigma\varphi(1) \end{bmatrix}^{-1} \begin{bmatrix} 1 & \int_0^1 W(r)dr \\ \int_0^1 W(r)dr & \int_0^1 W^2(r)dr \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & \sigma\varphi(1) \end{bmatrix}, \\
& \hspace{25em} (2.3.4)
\end{aligned}$$

和

$$\begin{aligned}
& \begin{bmatrix} T^{-1/2}\Sigma u_t \\ T^{-1}\Sigma y_{t-1}u_t \end{bmatrix} \Rightarrow \begin{bmatrix} \sigma\varphi(1)W(1) \\ \frac{1}{2}\{\sigma^2\varphi^2(1)W^2(1) - r_0\} \end{bmatrix} \\
& = \begin{bmatrix} \sigma\varphi(1)W(1) \\ \frac{1}{2}\sigma^2\varphi^2(1)\{W^2(1) - 1\} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{2}\{\sigma^2\varphi^2(1) - r_0\} \end{bmatrix} \\
& = \sigma\varphi(1) \begin{bmatrix} 1 & 0 \\ 0 & \sigma\varphi(1) \end{bmatrix} \begin{bmatrix} W(1) \\ \frac{1}{2}\{W^2(1) - 1\} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{2}\{\sigma^2\varphi^2(1) - r_0\} \end{bmatrix} \\
& \hspace{25em} (2.3.5)
\end{aligned}$$

其中, $r_0 = E(u_t^2)$ 。综合(2.3.4)和(2.3.5),可得

$$\begin{aligned}
& \begin{bmatrix} T^{1/2} \hat{a}_T \\ T(\hat{\rho}_T - 1) \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 \\ 0 & \sigma\varphi(1) \end{bmatrix}^{-1} \begin{bmatrix} 1 & \int_0^1 W(r)dr \\ \int_0^1 W(r)dr & \int_0^1 W^2(r)dr \end{bmatrix}^{-1} \\
& \quad \times \begin{bmatrix} 1 & 0 \\ 0 & \sigma\varphi(1) \end{bmatrix}^{-1} \left\{ \sigma\varphi(1) \begin{bmatrix} 1 & 0 \\ 0 & \sigma\varphi(1) \end{bmatrix} \right.
\end{aligned}$$

$$\begin{aligned}
& \times \left\{ \begin{array}{c} W(1) \\ \frac{1}{2}\{W^2(1) - 1\} \end{array} \right\} + \left\{ \begin{array}{c} 0 \\ \frac{1}{2}\{\sigma^2\varphi^2(1) - r_0\} \end{array} \right\} \Bigg\} \\
& = \begin{pmatrix} \sigma\varphi(1) & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \int_0^1 W(r)dr \\ \int_0^1 W(r)dr & \int_0^1 W^2(r)dr \end{pmatrix}^{-1} \\
& \times \left\{ \begin{array}{c} W(1) \\ \frac{1}{2}\{W^2(1) - 1\} \end{array} \right\} - \begin{pmatrix} 1 & 0 \\ 0 & \sigma^{-1}\varphi^{-1}(1) \end{pmatrix} \\
& \times \begin{pmatrix} 1 & \int_0^1 W(r)dr \\ \int_0^1 W(r)dr & \int_0^1 W^2(r)dr \end{pmatrix}^{-1} \\
& \times \left\{ \begin{array}{c} 0 \\ \frac{1}{2}\{\sigma^2\varphi^2(1) - r_0\}/\sigma\varphi(1) \end{array} \right\}
\end{aligned}$$

从中可将 $T(\hat{\rho}_T - 1)$ 的极限分离出来:

$$\begin{aligned}
T(\hat{\rho}_T - 1) & \Rightarrow [0, 1] \begin{pmatrix} 1 & \int_0^1 W(r)dr \\ \int_0^1 W(r)dr & \int_0^1 W^2(r)dr \end{pmatrix}^{-1} \begin{pmatrix} W(1) \\ \frac{1}{2}\{W^2(1) - 1\} \end{pmatrix} \\
& + \frac{1}{2} \frac{\sigma^2\varphi^2(1) - r_0}{\sigma^2\varphi^2(1)} [0, 1] \\
& \times \begin{pmatrix} 1 & \int_0^1 W(r)dr \\ \int_0^1 W(r)dr & \int_0^1 W^2(r)dr \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
&= s_T^2 [0, 1] \begin{bmatrix} T^{1/2} & 0 \\ 0 & T \end{bmatrix} \begin{bmatrix} T & \Sigma y_{t-1} \\ \Sigma y_{t-1} & \Sigma y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} T^{1/2} & 0 \\ 0 & T \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
&= s_T^2 [0, 1] \begin{bmatrix} 1 & T^{-3/2} \Sigma y_{t-1} \\ T^{-3/2} \Sigma y_{t-1} & T^{-2} \Sigma y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}
\end{aligned}$$

当 $T \rightarrow \infty$ 时, 有

$$\begin{aligned}
T^2 \hat{\eta}_T^2 &\Rightarrow r_0 [0, 1] \begin{bmatrix} 1 & 0 \\ 0 & \sigma^{-1} \varphi^{-1}(1) \end{bmatrix} \begin{bmatrix} 1 & \int_0^1 W(r) dr \\ \int_0^1 W(r) dr & \int_0^1 W^2(r) dr \end{bmatrix}^{-1} \\
&\quad \times \begin{bmatrix} 1 & 0 \\ 0 & \sigma^{-1} \varphi^{-1}(1) \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
&= \frac{r_0}{\sigma^2 \varphi^2(1)} \cdot \frac{1}{\int_0^1 W^2(r) dr - \left(\int_0^1 W(r) dr \right)^2} \quad (2.3.7)
\end{aligned}$$

与表达式(2.2.10)相比, 上式多了因子 $r_0/\sigma^2 \varphi^2(1)$, 它的大小反映了稳定过程 $\{u_t\}$ 的自相关程度, 当 $\{u_t\}$ 独立同分布时, $\varphi_0 = 1$, $\varphi_j = 0$, $j = 1, 2, \dots$, 就有

$$r_0/\sigma^2 \varphi^2(1) = 1$$

对统计量 $T(\hat{\rho}_T - 1)$ 作如下修正:

$$T(\hat{\rho}_T - 1) = \frac{1}{2} \{ \sigma^2 \varphi^2(1) - r_0 \} \{ T^2 \hat{\eta}_T^2 / s_T^2 \}$$

根据(2.3.6)和(2.3.7)可知上式有极限

$$T(\hat{\rho}_T - 1) = \frac{1}{2} \{ \sigma^2 \varphi^2(1) - r_0 \} \{ T^2 \hat{\eta}_T^2 / s_T^2 \}$$

$$\begin{aligned}
& \Rightarrow \frac{\frac{1}{2}\{W^2(1) - 1\} - W(1)\int_0^1 W(r)dr}{\int_0^1 W^2(r)dr - \left(\int_0^1 W(r)dr\right)^2} \\
& \quad + \frac{\frac{1}{2}\{\sigma^2\varphi^2(1) - r_0\}}{\sigma^2\varphi^2(1)\left\{\int_0^1 W^2(r)dr - \left(\int_0^1 W(r)dr\right)^2\right\}} \\
& \quad - \frac{\frac{1}{2}\{\sigma^2\varphi^2(1) - r_0\}}{\sigma^2\varphi^2(1)\left\{\int_0^1 W^2(r)dr - \left(\int_0^1 W(r)dr\right)^2\right\}} \\
& = \frac{\frac{1}{2}\{W^2(1) - 1\} - W(1)\int_0^1 W(r)dr}{\int_0^1 W^2(r)dr - \left(\int_0^1 W(r)dr\right)^2} \quad (2.3.8)
\end{aligned}$$

修正后的极限与(2.2.9)一致,不再依赖于任何未知参数,它的临界值列在表1的情况二中。当然,修正后的统计量本身仍包含未知参数 $\sigma^2\varphi^2(1)$ 和 r_0 ,我们将在下面讨论对它们的估计方法,这是PP方法的一个重要组成部分。

用以上的方法,同样可以考虑统计量 t_T 的极限和修正方法。对于 t_T 统计量

$$t_T = \frac{\hat{\rho}_T - 1}{\hat{\eta}_T}$$

利用 $T^2\hat{\eta}_T^2$ 的极限分布(2.3.7),可得

$$\begin{aligned}
t_T &= \frac{\hat{\rho}_T - 1}{\hat{\eta}_T} = \frac{T(\hat{\rho}_T - 1)}{(T^2\hat{\eta}_T^2)^{1/2}} \\
&\Rightarrow \frac{\frac{1}{2}\{W^2(1) - 1\} - W(1)\int_0^1 W(r)dr}{\int_0^1 W^2(r)dr - \left(\int_0^1 W(r)dr\right)^2} \cdot \left\{\frac{\sigma^2\varphi^2(1)}{r_0}\right\}^{1/2}
\end{aligned}$$

$$\begin{aligned}
& \cdot \left\{ \int_0^1 W^2(r) dr - \left(\int_0^1 W(r) dr \right)^2 \right\}^{1/2} \\
& + \frac{1}{2} \cdot \frac{1}{\sigma^2 \varphi^2(1)} \cdot (\sigma^2 \varphi^2(1) - r_0) \cdot \left(\frac{\sigma^2 \varphi^2(1)}{r_0} \right)^{1/2} \\
& \quad / \left\{ \int_0^1 W^2(r) dr - \left(\int_0^1 W(r) dr \right)^2 \right\}^{1/2} \\
& = \frac{\frac{1}{2} \{W^2(1) - 1\} - W(1) \int_0^1 W(r) dr}{\left\{ \int_0^1 W^2(r) dr - \left(\int_0^1 W(r) dr \right)^2 \right\}^{1/2}} \cdot \frac{\sigma \varphi(1)}{r_0^{1/2}} \\
& \quad + \frac{\frac{1}{2} (\sigma^2 \varphi^2(1) - r_0)}{\sigma \varphi(1) r_0^{1/2} \cdot \left\{ \int_0^1 W^2(r) dr - \left(\int_0^1 W(r) dr \right)^2 \right\}^{1/2}} \\
& \hspace{15em} (2.3.9)
\end{aligned}$$

与 $T(\hat{\rho}_T - 1)$ 一样, t_T 的极限也由两项组成, 其中的第二项反映了 $\{u_t\}$ 的自相关程度。若 $\{u_t\}$ 独立同分布, 且 $E\{u_t\} = 0$, $D(u_t) = \sigma^2$, 那么 $\varphi(1) = 1$, $r_0 = \sigma^2$, (2.3.9) 式简化成 (2.2.11) 式, 后者是作 DF 检验的 t_T 的极限。

对统计量作如下的修正:

$$\frac{r_0^{1/2}}{\sigma \varphi(1)} \cdot t_T = \left\{ \frac{1}{2} (\sigma^2 \varphi^2(1) - r_0) / \sigma \varphi(1) \right\} \cdot \{T \hat{\eta}_T / s_T\}$$

由 (2.3.7) 中的分析, 立即可得

$$\begin{aligned}
& \frac{r_0^{1/2}}{\sigma \varphi(1)} \cdot t_T = \left\{ \frac{1}{2} (\sigma^2 \varphi^2(1) - r_0) / \sigma \varphi(1) \right\} \cdot \{T \hat{\eta}_T / s_T\} \\
& \Rightarrow \frac{\frac{1}{2} \{W^2(1) - 1\} - W(1) \int_0^1 W(r) dr}{\left\{ \int_0^1 W^2(r) dr - \left(\int_0^1 W(r) dr \right)^2 \right\}^{1/2}} \hspace{10em} (2.3.10)
\end{aligned}$$

$$\hat{r}_j = T^{-1} \sum_{i=j+1}^T \hat{u}_i \hat{u}_{i-j}$$

估计得到。参数 $\varphi_j (j = 0, 1, 2, \dots)$ 和自相关系数 $r_j (j = 0, 1, 2, \dots)$ 的关系由下式给出：

$$\sigma^2 \varphi^2(1) = r_0 + 2 \sum_{j=1}^{\infty} r_j \quad (2.3.11)$$

其中, $r_0 = E(u_i^2)$, $r_j = E(u_i u_{i-j})$, $j = 1, 2, \dots$ 。

由于样本量的限制,在实际中只存在有限阶的样本自相关系数 \hat{r}_j ,这就需要构造一个有高收敛速度的一致估计量。以下的估计量由纽威和韦斯特(Newey and West, 1987)提出:

$$\hat{\sigma}^2 \hat{\varphi}^2(1) = \hat{r}_0 + 2 \sum_{j=1}^q \left[1 - \frac{j}{q+1}\right] \hat{r}_j \quad (2.3.12)$$

这里, q 为一有限的正整数,当 $q \rightarrow \infty$ 时,上式与(2.3.11)一致;对于给定的 q ,估计量 $\hat{\sigma}^2 \hat{\varphi}^2(1)$ 中的前几项有更多的权,对于较大的 j , \hat{r}_j 在和式(2.3.12)中的作用很小。估计量 $\hat{\sigma}^2 \hat{\varphi}^2(1)$ 是一致估计量,它的收敛速度高于平均加权的估计量

$$\tilde{\sigma}^2 \tilde{\varphi}^2(1) = \hat{r}_0 + 2 \sum_{j=1}^q \hat{r}_j$$

将估计量 \hat{r}_0 和 $\hat{\sigma}^2 \hat{\varphi}^2(1)$ 代入修正的 $T(\hat{\rho}_T - 1)$ 统计量和 t_T 统计量,即得

$$T(\hat{\rho}_T - 1) = \frac{1}{2} (\hat{\sigma}^2 \hat{\varphi}^2(1) - \hat{r}_0) (T^2 \hat{\eta}_T^2 / s_T^2) \quad (2.3.13)$$

和

$$\frac{\hat{r}_0^{1/2}}{\hat{\sigma} \hat{\varphi}(1)} \cdot t_T = \left\{ \frac{1}{2} (\hat{\sigma}^2 \hat{\varphi}^2(1) - \hat{r}_0) / \hat{\sigma} \hat{\varphi}(1) \right\} \{T \hat{\eta}_T / s_T\} \quad (2.3.14)$$

它们分别称为 Z_ρ 检验法的 ρ -统计量和 Z_t 检验法的 t -统计量。它们的极限分别由(2.3.8)和(2.3.10)给出。

对情况二的单位根假设 $H_0: \rho = 1$, PP 检验的步骤如下:

(1) 以最小二乘法估计模型

$$y_t = \alpha + \rho y_{t-1} + u_t$$

得估计量 $\hat{\alpha}_T$ 、 $\hat{\rho}_T$ 和估计残差 $\hat{u}_t = y_t - \hat{\alpha}_T - \hat{\rho}_T y_{t-1}$;

(2) 计算

$$\hat{r}_0 = \frac{1}{T} \sum_{t=1}^T \hat{u}_t^2; \quad \hat{r}_j = \frac{1}{T} \sum_{t=j+1}^T \hat{u}_t \hat{u}_{t-j},$$

$$\hat{\sigma}^2 \hat{\varphi}^2(1) = \hat{r}_0 + 2 \sum_{j=1}^q \left[1 - \frac{j}{q+1} \right] \hat{r}_j$$

(3) 计算标准差 $\hat{\eta}_T$ 和方差估计量 s_T^2 ;

(4) 由(2.3.13)和(2.3.14)计算 ρ -统计量和 t -统计量的值, 并与表 1 和表 2 的情况二中的临界值比较。

以下例子说明 PP 检验的应用。

例 2.4 在例 2.2 中, 随机干扰项 u_t 假设为独立同分布, 以下假设 $\{u_t\}$ 服从一般的稳定过程:

$$u_t = \sum_{j=0}^{\infty} \varphi_j \varepsilon_{t-j},$$

估计残差 \hat{u}_t 为

$$\hat{u}_t = i_t - 0.211 - 0.96691 i_{t-1}; \quad t = 1, 2, \dots, 168$$

由此计算四阶样本自相关系数:

$$\hat{r}_0 = \frac{1}{T} \sum \hat{u}_t^2 = 0.630; \quad \hat{r}_1 = \frac{1}{T} \sum \hat{u}_t \hat{u}_{t-1} = 0.114;$$

$$\hat{r}_2 = \frac{1}{T} \sum \hat{u}_t \hat{u}_{t-2} = -0.162; \quad \hat{r}_3 = \frac{1}{T} \sum \hat{u}_t \hat{u}_{t-3} = 0.064;$$

$$\hat{r}_4 = \frac{1}{T} \sum \hat{u}_t \hat{u}_{t-4} = 0.047$$

并计算估计量 $\hat{\sigma}^2 \hat{\varphi}^2(1)$ 的值

$$\begin{aligned} \hat{\sigma}^2 \hat{\varphi}^2(1) &= 0.630 + 2 \left\{ \frac{4}{5} \right\} (0.114) + 2 \left\{ \frac{3}{5} \right\} (-0.162) \\ &\quad + 2 \left\{ \frac{2}{5} \right\} (0.064) + 2 \left\{ \frac{1}{5} \right\} (0.047) = 0.688 \end{aligned}$$

和 $\hat{\eta}_T$ 、 \hat{s}_T^2 的值

$$\hat{\eta}_T = 0.0191; \quad \hat{s}_T^2 = 0.6376$$

将以上计算结果代入 Z_ρ 检验的 ρ -统计量, 可得

$$\begin{aligned} T(\hat{\rho}_T - 1) &= \frac{1}{2} (T^2 \hat{\eta}_T^2 / \hat{s}_T^2) (\hat{\sigma}^2 \hat{\varphi}^2(1) - \hat{r}_0) \\ &= 168(0.966191 - 1) = \frac{1}{2} \{ [(168)(0.0191)]^2 / (0.6376) \} \\ &\quad \times \{ 0.688 - 0.630 \} \\ &= -6.03 \end{aligned}$$

以 5% 为显著水平, 由表 1 情况二查得相应于 $T = 168$ 的临界值为 -13.8 , 因为 $-6.03 > -13.8$, 所以接受原假设 $H_0: \rho = 1$ 。

同样可以计算 Z_t 检验的 t -统计量:

$$\begin{aligned} \frac{\hat{r}_0^{1/2}}{\hat{\sigma} \hat{\varphi}(1)} \cdot t_T &= \left\{ \frac{1}{2} (\hat{\sigma}^2 \hat{\varphi}^2(1) - \hat{r}_0) / \hat{\sigma} \hat{\varphi}(1) \right\} \{ T \hat{\eta}_T / \hat{s}_T \} \\ &= \{ (0.63) / (0.688) \}^{1/2} \{ (0.96691 - 1) / (0.0191) \} \end{aligned}$$

$$-\left\{\frac{1}{2}(0.688-0.630)[(168)(0.0191)/\sqrt{0.6376}]/\sqrt{0.688}\right\}$$

$$=-1.80$$

由表 2 情况二查得相应于 5% 显著水平的临界值为 -2.89, 因为 $-1.80 > -2.89$, 又一次接受单位根假设。

在计算中, 相关系数 \hat{r}_q 的阶数 q 的选取应根据实际情况而定。在上例中, 第四阶相关系数 \hat{r}_4 在 $\hat{\sigma}^2 \hat{\varphi}^2(1)$ 的值中只占 3% 左右, 更高阶的 \hat{r}_j 对 $\hat{\sigma}^2 \hat{\varphi}^2(1)$ 的贡献是微不足道的。

2.3.2 情况一和情况四的 PP 检验法

上节讨论的 PP 检验法可以平行地推广到情况一和情况四。在情况一中, 我们假设数据由

$$y_t = \rho y_{t-1} + u_t$$

产生, 并在其中检验 $\rho = 1$; 而在情况四中, 假设数据由

$$y_t = \alpha + \rho y_{t-1} + u_t$$

产生, 且在模型

$$y_t = \alpha + \rho y_{t-1} + \delta t + u_t$$

中检验, 假设 $\rho = 1$, 或联合假设 $H_0: \rho = 1, \delta = 0$ 。对于这两种情况, 我们都假设 $\{u_t\}$ 为一稳定过程, 并可表示为

$$u_t = \sum_{j=0}^{\infty} \varphi_j \varepsilon_{t-j} \quad (2.3.15)$$

对于假设 $H_0: \rho = 1$, 仍可用 ρ -统计量 (2.3.13) 和 t -统计量 (2.3.14) 作检验, 只是在表 1 和表 2 的情况一和情况四中选取适当的临界值。以下例子说明具体的检验步骤。

同样可计算 t -统计量的值:

$$\begin{aligned} & \frac{\hat{r}_0^{1/2}}{\hat{\sigma} \hat{\varphi}(1)} \cdot t_T = \left\{ \frac{1}{2} (\hat{\sigma}^2 \hat{\varphi}^2(1) - \hat{r}_0) / \hat{\sigma} \hat{\varphi}(1) \right\} \cdot \{T \hat{\eta}_T / s_T\} \\ & = \{(1.136)/(2.117)\}^{1/2} \{(0.96252-1)/(0.0193)\} - \left\{ \frac{1}{2} (2.117 \right. \\ & \quad \left. - 1.136) [(168)(0.0193) / \sqrt{1.15627}] / \sqrt{2.117} \right\} \\ & = -2.44 \end{aligned}$$

仍以 5% 为显著水平, 由表 2 的情况四查得相应于 $T = 168$ 的临界值为 -3.44 。因为 $-2.44 > -3.44$, 我们又一次接受单位根假设。

在此阶段还不能将 PP 检验法与上节中讨论的 DF 检验法作比较, 因为在讨论 DF 检验法时, 我们假设随机干扰 $\{\epsilon_t\}$ 为独立同分布的。以下将 DF 检验法推广到一般的稳定过程。

2.4 增广的迪基—福勒(ADF)检验法

增广的迪基—福勒检验法 (Augmented Dickey-Fuller Test), 以下简称 ADF 检验, 将检验单位根的 DF 方法推广到一般的单位根过程, 其中的随机干扰 $\{u_t\}$ 服从一稳定过程。ADF 方法由迪基和福勒在 1979 年提出 (Dickey and Fuller, 1979)。

2.4.1 P 阶自回归过程

ADF 方法的一个重要特点, 也是它区别于 PP 方法的一个主要特征, 在于它假设数据生成过程 $\{y_t\}$ 服从有单位根的 P 阶自回归过程, 而 PP 方法假设 y_t 中的随机干扰 u_t 服从无穷阶的

移动平均过程, 即 $u_t = \sum_{i=0}^{\infty} \varphi_i \epsilon_{t-i}$ 。这两种过程 (即 $MA(\infty)$ 和

AR(P))往往可以互相转换,它们之间的异同在以下讨论中会显得更明确。

随机过程 $\{y_t\}$ 服从 P 阶的自回归过程AR(P),如果 y_t 有表示形式

$$B(L)y_t = (1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p)y_t = \epsilon_t \quad (2.4.1)$$

其中, $\{\epsilon_t\}$ 为独立同分布过程,且 $E(\epsilon_t) = 0$, $D(\epsilon_t) = \sigma^2$, $E(\epsilon_t^2) < \infty$; $B(L)$ 为 P 阶的滞后多项式,它的特征方程

$$1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p = 0$$

有一个单位根 $z = 1$,其他根都在单位圆外。令

$$\rho = \phi_1 + \phi_2 + \cdots + \phi_p$$

$$\zeta_j = -(\phi_{j+1} + \cdots + \phi_p); j = 1, 2, \cdots, p-1$$

可将滞后多项式 $B(L)$ 改写成

$$\begin{aligned} B(L) &= (1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p) \\ &= (1 - \rho L) - (\zeta_1 L + \zeta_2 L^2 + \cdots + \zeta_{p-1} L^{p-1})(1 - L) \end{aligned} \quad (2.4.2)$$

相应地,可将 y_t 的AR(P)过程改写成

$$\begin{aligned} B(L)y_t &= \{(1 - \rho L) - (\zeta_1 L + \zeta_2 L^2 + \cdots \\ &\quad + \zeta_{p-1} L^{p-1})(1 - L)\}y_t = \epsilon_t \end{aligned}$$

重新组织后,可得

$$y_t = \rho y_{t-1} + \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots + \zeta_{p-1} \Delta y_{t-p+1} + \epsilon_t \quad (2.4.3)$$

根据假设,特征多项式 $B(z)$ 有且只有一个单位根,所以当 $z = 1$ 时,有

$$B(1) = (1 - \phi_1 - \phi_2 - \cdots - \phi_p) = 1 - \rho = 0$$

上式等价于 $\rho = 1$ 。

若 $\rho = 1$ 为真,可将 $B(L)$ 进一步改写成

$$\begin{aligned} B(L) &= (1 - L) - (\zeta_1 L + \zeta_2 L^2 + \cdots + \zeta_{p-1} L^{p-1})(1 - L) \\ &= (1 - \zeta_1 L - \zeta_2 L^2 - \cdots - \zeta_{p-1} L^{p-1})(1 - L) \end{aligned}$$

相应地,有

$$\begin{aligned} B(L)y_t &= (1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p)y_t \\ &= (1 - \zeta_1 L - \zeta_2 L^2 - \cdots - \zeta_{p-1} L^{p-1})(1 - L)y_t \\ &= C(L)\Delta y_t = \varepsilon_t \end{aligned} \quad (2.4.4)$$

由于特征多项式

$$C(z) = 1 - \zeta_1 z - \zeta_2 z^2 - \cdots - \zeta_{p-1} z^{p-1}$$

的根都在单位圆外,所以滞后多项式 $C(L)$ 的逆存在,从而有

$$\Delta y_t = C^{-1}(L)\varepsilon_t = \varphi(L)\varepsilon_t = u_t \quad (2.4.5)$$

$\Psi(L)$ 一般为一无穷阶的滞后多项式。容易看出, (2.4.5) 与 (2.3.1) 中的过程是等价的。

增广的迪基-福勒方法 (ADF 方法) 在模型 (2.4.3) 中检验单位根假设 $H_0: \rho = 1$ 。模型 (2.4.3) 只含有回归变量 y_{t-1} 和 y_{t-1} 的各阶差分, 它们中并不包含任何未知参数和未知变量。由于 $\{\varepsilon_t\}$ 的独立同分布, (2.4.3) 中的参数可由最小二乘法估计。当样本量 $T \rightarrow \infty$ 时, 这些估计都是一致的。

$$\sum_{t=1}^T \mathbf{x}_t \mathbf{x}_t' = \begin{pmatrix} \Sigma \Delta y_{t-1}^2 & \Sigma \Delta y_{t-1} \Delta y_{t-2} & \cdots & \Sigma \Delta y_{t-1} \Delta y_{t-p+1} & \Sigma \Delta y_{t-1} & \Sigma \Delta y_{t-1} y_{t-1} \\ \Sigma \Delta y_{t-2} \Delta y_{t-1} & \Sigma \Delta y_{t-2}^2 & \cdots & \Sigma \Delta y_{t-2} \Delta y_{t-p+1} & \Sigma \Delta y_{t-2} & \Sigma \Delta y_{t-2} y_{t-1} \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ \Sigma \Delta y_{t-p+1} \Delta y_{t-1} & \Sigma \Delta y_{t-p+1} \Delta y_{t-2} & \cdots & \Sigma \Delta y_{t-p+1}^2 & \Sigma \Delta y_{t-p+1} & \Sigma \Delta y_{t-p+1} y_{t-1} \\ \Sigma \Delta y_{t-1} & \Sigma \Delta y_{t-2} & \cdots & \Sigma \Delta y_{t-p+1} & T & \Sigma y_{t-1} \\ \Sigma y_{t-1} \Delta y_{t-1} & \Sigma y_{t-1} \Delta y_{t-2} & \cdots & \Sigma y_{t-1} \Delta y_{t-p+1} & \Sigma y_{t-1} & \Sigma y_{t-1}^2 \end{pmatrix}$$

和

$$\sum_{t=1}^T \mathbf{x}_t \boldsymbol{\varepsilon}_t = \begin{pmatrix} \Sigma \Delta y_{t-1} \boldsymbol{\varepsilon}_t \\ \Sigma \Delta y_{t-2} \boldsymbol{\varepsilon}_t \\ \vdots \\ \Sigma \Delta y_{t-p+1} \boldsymbol{\varepsilon}_t \\ \Sigma \boldsymbol{\varepsilon}_t \\ \Sigma y_{t-1} \boldsymbol{\varepsilon}_t \end{pmatrix}$$

定义对角矩阵

$$\mathbf{D}_T = \begin{pmatrix} \sqrt{T} & 0 & \cdots & 0 & 0 \\ 0 & \sqrt{T} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \sqrt{T} & 0 \\ 0 & 0 & \cdots & 0 & T \end{pmatrix}$$

并将其左乘(2.4.8)式,整理后可得

$$\mathbf{D}_T (\mathbf{b}_T - \boldsymbol{\beta}) = \{ \mathbf{D}_T^{-1} \left[\sum_{t=1}^T \mathbf{x}_t \mathbf{x}_t' \right] \mathbf{D}_T^{-1} \}^{-1} \{ \mathbf{D}_T^{-1} \left[\sum_{t=1}^T \mathbf{x}_t \boldsymbol{\varepsilon}_t \right] \}$$

矩阵 \mathbf{D}_T 对角线上的前 p 个元素均为 \sqrt{T} , 因此矩阵 $\mathbf{D}_T^{-1} \left[\sum_{t=1}^T \mathbf{x}_t \mathbf{x}_t' \right] \mathbf{D}_T^{-1}$ 左上角的 $(p \times p)$ 维子矩阵的每个元素都有因

子 $\frac{1}{T}$, 最后一行及最后一列的每个元素都有因子 $T^{-3/2}$, 右下角的最后一个元素有因子 T^{-2} 。

当原假设 $H_0: \alpha = 0, \rho = 1$ 成立时, 根据定理 1.9, 可得

$$T^{-1} \sum \Delta y_{t-i} \Delta y_{t-j} \xrightarrow{P} r_{|i-j|}; \quad i \neq j$$

$$T^{-1} \sum \Delta y_{t-j} \xrightarrow{P} E(\Delta y_{t-j}) = 0$$

$$T^{-3/2} \sum y_{t-1} \Delta y_t \xrightarrow{P} 0$$

$$T^{-3/2} \sum y_{t-1} \Rightarrow \sigma \Psi(1) \int_0^1 W(r) dr$$

$$T^{-2} \sum y_{t-1}^2 \Rightarrow \sigma^2 \Psi^2(1) \int_0^1 W^2(r) dr$$

其中, $r_j = E(\Delta y_t \cdot \Delta y_{t-j})$, $\Psi^2(1) = \frac{1}{1 - \zeta_1 - \zeta_2 - \dots - \zeta_{p-1}}$ 。

从而可得矩阵 $D_T^{-1} \left[\sum_{i=1}^T x_i x_i' \right] D_T^{-1}$ 的极限:

$$\begin{aligned} & D_T^{-1} \left[\sum_{i=1}^T x_i x_i' \right] D_T^{-1} \\ \Rightarrow & \begin{pmatrix} r_0 & r_1 & \cdots & r_{p-2} & 0 & 0 \\ r_1 & r_0 & \cdots & r_{p-3} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ r_{p-2} & r_{p-3} & \cdots & r_0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 1 & \sigma \Psi(1) \int_0^1 W(r) dr \\ 0 & 0 & \cdots & 0 & \sigma \Psi(1) \int_0^1 W(r) dr & \sigma^2 \Psi^2(1) \int_0^1 W^2(r) dr \end{pmatrix} \\ = & \begin{bmatrix} V & 0 \\ 0 & Q \end{bmatrix} \end{aligned} \quad (2.4.9)$$

其中的 V 和 Q 分别为 $(p-1) \times (p-1)$ 维和 2×2 维矩阵:

$$V = \begin{pmatrix} r_0 & r_1 & \cdots & r_{p-2} \\ r_1 & r_0 & \cdots & r_{p-3} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p-2} & r_{p-3} & \cdots & r_0 \end{pmatrix} \quad (2.4.10)$$

$$Q = \begin{pmatrix} 1 & \sigma\Psi(1)\int_0^1 W(r)dr \\ \sigma\Psi(1)\int_0^1 W(r)dr & \sigma^2\Psi^2(1)\int_0^1 W^2(r)dr \end{pmatrix} \quad (2.4.11)$$

其次,再考虑向量 $D_T^{-1}\left[\sum_{i=1}^T x_i \varepsilon_i\right]$

$$D_T^{-1}\left[\sum_{i=1}^T x_i \varepsilon_i\right] = \begin{pmatrix} T^{-1/2}\Sigma\Delta y_{t-1}\varepsilon_t \\ T^{-1/2}\Sigma\Delta y_{t-2}\varepsilon_t \\ \vdots \\ T^{-1/2}\Sigma\Delta y_{t-p+1}\varepsilon_t \\ T^{-1/2}\Sigma\varepsilon_t \\ T^{-1}\Sigma y_{t-1}\varepsilon_t \end{pmatrix} \quad (2.4.12)$$

(2.4.12)式中的前 $p-1$ 个元素具有 $T^{-1/2}\Sigma\Delta y_{t-i}\varepsilon_t$ ($i=1, 2, \dots, p-1$) 的形式,因为 $\{\Delta y_{t-i}\}$ 和 $\{\varepsilon_t\}$ 都为稳定过程,根据中心极限定理,前 $p-1$ 维子向量有正态的极限分布,并有协方差矩阵

$$\begin{aligned} & E\left\{\begin{pmatrix} \Delta y_{t-1}\varepsilon_t \\ \Delta y_{t-2}\varepsilon_t \\ \vdots \\ \Delta y_{t-p+1}\varepsilon_t \end{pmatrix} [\Delta y_{t-1}\varepsilon_t, \Delta y_{t-2}\varepsilon_t, \dots, \Delta y_{t-p+1}\varepsilon_t]\right\} \\ &= \sigma^2 \begin{pmatrix} r_0 & r_1 & \cdots & r_{p-2} \\ r_1 & r_0 & \cdots & r_{p-3} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p-2} & r_{p-3} & \cdots & r_0 \end{pmatrix} = \sigma^2 V \end{aligned}$$

即

$$\begin{bmatrix} T^{-1/2} \sum \Delta y_{t-1} \epsilon_t \\ T^{-1/2} \sum \Delta y_{t-2} \epsilon_t \\ \vdots \\ T^{-1/2} \sum \Delta y_{t-p+1} \epsilon_t \end{bmatrix} \Rightarrow h_1 \sim N(0, \sigma^2 V) \quad (2.4.13)$$

同样根据定理 1.9, 还可知

$$\begin{bmatrix} T^{-1/2} \sum \epsilon_t \\ T^{-1} \sum y_{t-1} \epsilon_t \end{bmatrix} \Rightarrow h_2 = \begin{bmatrix} \sigma W(1) \\ \frac{1}{2} \sigma^2 \Psi(1) \{W^2(1) - 1\} \end{bmatrix}$$

综上所述, $D_T(b_T - \beta)$ 有以下极限:

$$\begin{aligned} D_T(b_T - \beta) &= \{D_T^{-1} [\sum x_t x_t' D_T^{-1}]^{-1} \{D_T^{-1} [\sum x_t \epsilon_t]\} \\ &\Rightarrow \begin{bmatrix} V & 0 \\ 0 & Q \end{bmatrix}^{-1} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} V^{-1} h_1 \\ Q^{-1} h_2 \end{bmatrix} \end{aligned} \quad (2.4.14)$$

若将参数向量 β 写成分块形式:

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

$\beta'_1 = [\zeta_1, \zeta_2, \dots, \zeta_{p-1}]$, $\beta'_2 = [\alpha, \rho]$, 可分别对它们进行假设检验。以下先考虑对子向量 β_1 的检验。若对 β_1 的假设可表示成线性约束 $R_1 \beta_1 = \gamma_1$ 的形式, 其中的 R_1 为一给定的 $m \times (p-1)$ 维矩阵, 并有秩 m , γ_1 为一给定的 m 维向量, 由 (2.4.13) 和 (2.4.14) 中的讨论, β_1 的最小二乘估计 b_1^+ 有正态的极限分布:

$$\sqrt{T}(b_1^+ - \beta_1) \Rightarrow V^{-1} h_1 \sim N(0, \sigma^2 V^{-1})$$

当假设 $H_0: R_1 \beta_1 = \gamma_1$ 为真时

$$\sqrt{T} R_1 b_1^+ \xrightarrow{d} N(\gamma_1, \sigma^2 R_1 V^{-1} R_1') \quad (2.4.15)$$

从而可以构造二次型

$$T(\mathbf{R}_1 \mathbf{b}_T^1 - \mathbf{Y}_1)' [s_T^2 \mathbf{R}_1 \hat{\mathbf{V}}^{-1} \mathbf{R}_1']^{-1} (\mathbf{R}_1 \mathbf{b}_T^1 - \mathbf{Y}_1)$$

在 $H_0: \mathbf{R}_1 \boldsymbol{\beta}_1 = \mathbf{Y}_1$ 成立时, 它有极限分布 $\chi^2(m)$, 这里的 s_T^2 和 $\hat{\mathbf{V}}$ 分别为 σ^2 和 \mathbf{V} 的最小二乘估计, 由下式给出:

$$s_T^2 = \frac{1}{T - (p+1)} \sum_{i=1}^T (y_i - \hat{\xi}_1 \Delta y_{i-1} - \dots \\ - \hat{\xi}_{p-1} \Delta y_{i-p+1} - \hat{\alpha}_T - \hat{\rho}_T y_{i-1})^2$$

$$\hat{\mathbf{V}} = [\mathbf{I}_{p-1}, \mathbf{0}_{p-1,2}] \left\{ \mathbf{D}_T^{-1} \left[\sum_{i=1}^T \mathbf{x}_i \mathbf{x}_i' \right] \mathbf{D}_T^{-1} \right\} \begin{bmatrix} \mathbf{I}_{p-1} \\ \mathbf{0}_{p-1,2}' \end{bmatrix}$$

这里的 \mathbf{I}_{p-1} 为 $(p-1) \times (p-1)$ 维单位矩阵, $\mathbf{0}_{p-1,2}$ 为 $(p-1) \times 2$ 维零矩阵。 s_T^2 和 $\hat{\mathbf{V}}$ 都是一致的估计量, 当 $T \rightarrow \infty$, 分别有

$$s_T^2 \xrightarrow{P} \sigma^2, \quad \hat{\mathbf{V}} \xrightarrow{P} \mathbf{V}$$

以上分析表明, 由于 $\boldsymbol{\beta}_1 = [\xi_1, \xi_2, \dots, \xi_{p-1}]'$ 的最小二乘估计有正态的极限分布, 对假设 $\mathbf{R}_1 \boldsymbol{\beta}_1 = \mathbf{Y}_1$ 的检验可用 χ^2 -检验法进行, 检验所需的临界值可从 χ^2 -表中查到。

以下讨论对参数 $\boldsymbol{\beta}_2 = [\alpha, \rho]'$ 的假设检验。由以上讨论, 统计量 $\sqrt{T} \hat{\alpha}_T$ 和 $T(\hat{\rho}_T - 1)$ 在原假设 $H_c: \alpha = 0, \rho = 1$ 为真时, 有以下极限:

$$\begin{bmatrix} T^{1/2} \hat{\alpha}_T \\ T(\hat{\rho}_T - 1) \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & \sigma \Psi(1) \int_0^1 W(r) dr \\ \sigma \Psi(1) \int_0^1 W(r) dr & \sigma^2 \Psi^2(1) \int_0^1 W^2(r) dr \end{bmatrix}^{-1} \\ \times \begin{bmatrix} \sigma W(1) \\ \frac{1}{2} \sigma^2 \Psi(1) \{W^2(1) - 1\} \end{bmatrix}$$

$$\begin{aligned}
&= \sigma \begin{bmatrix} 1 & 0 \\ 0 & \sigma\Psi(1) \end{bmatrix}^{-1} \begin{bmatrix} 1 & \int_0^1 W(r)dr \\ \int_0^1 W(r)dr & \int_0^1 W^2(r)dr \end{bmatrix}^{-1} \\
&\quad \times \begin{bmatrix} 1 & 0 \\ 0 & \sigma\Psi(1) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sigma\Psi(1) \end{bmatrix} \\
&\quad \times \begin{bmatrix} W(1) \\ \frac{1}{2}\{W^2(1) - 1\} \end{bmatrix} \\
&= \begin{bmatrix} \sigma & 0 \\ 0 & 1/\Psi(1) \end{bmatrix} \begin{bmatrix} 1 & \int_0^1 W(r)dr \\ \int_0^1 W(r)dr & \int_0^1 W^2(r)dr \end{bmatrix}^{-1} \\
&\quad \times \begin{bmatrix} W(1) \\ \frac{1}{2}\{W^2(1) - 1\} \end{bmatrix}
\end{aligned}$$

从中可分离出相应于 $T(\hat{\rho}_T - 1)$ 的极限, 整理后得

$$T \cdot \Psi(1)(\hat{\rho}_T - 1) \Rightarrow \frac{\frac{1}{2}\{W^2(1) - 1\} - W(1)\int_0^1 W(r)dr}{\int_0^1 W^2(r)dr - \left(\int_0^1 W(r)dr\right)^2} \quad (2.4.16)$$

此极限与(2.2.9)一致。上式的左侧多了因子 $\Psi(1)$, 这是在 $\{u_i\}$ 服从稳定过程时对 $T(\hat{\rho}_T - 1)$ 的修正。当 $\{u_i\}$ 为独立同分布时, $\Psi(1) = 1$, (2.4.16) 就与(2.2.9)完全一致。

因为

$$\Psi(1)^{-1} = (1 - \zeta_1 - \zeta_2 - \cdots - \zeta_{p-1})$$

$\Psi(1)$ 的估计值可由 $\zeta_j (j = 1, 2, \dots, p-1)$ 的最小二乘估计

$$s_i^2 = \frac{1}{T - (p+1)} \sum_{t=1}^T (y_t - \hat{\xi}_1 \Delta y_{t-1} - \dots \\ - \hat{\xi}_{p-1} \Delta y_{t-p-1} - \hat{a}_T - \hat{\rho}_T y_{t-1})^2$$

由此可将 t_T 统计量写成以下形式:

$$t_T = \frac{(\hat{\rho}_T - 1)}{\{s_T^2 \cdot \mathbf{e}'_{p+1} (\Sigma \mathbf{x}, \mathbf{x}')^{-1} \mathbf{e}_{p+1}\}^{1/2}} \\ = \frac{T(\hat{\rho}_T - 1)}{\{s_T^2 \cdot \mathbf{e}'_{p+1} [T^2 (\Sigma \mathbf{x}, \mathbf{x}')^{-1}] \mathbf{e}_{p+1}\}^{1/2}} \quad (2.4.18)$$

$\mathbf{e}'_{p+1} (\Sigma \mathbf{x}, \mathbf{x}')^{-1} \mathbf{e}_{p+1}$ 是矩阵 $(\Sigma \mathbf{x}, \mathbf{x}')^{-1}$ 的右下角元素, 由 (2.4.10) 和 (2.4.11) 中矩阵 V 和 Q 的形式, 可知当 $T \rightarrow \infty$ 时, 有

$$s_T^2 \mathbf{e}'_{p+1} [T^2 (\Sigma \mathbf{x}, \mathbf{x}')^{-1}] \mathbf{e}_{p+1} = s_T^2 \mathbf{e}'_{p+1} \{D_T^{-1} (\Sigma \mathbf{x}, \mathbf{x}') D_T^{-1}\}^{-1} \mathbf{e}_{p+1} \\ \Rightarrow \sigma^2 \mathbf{e}'_{p+1} \begin{bmatrix} V^{-1} & \mathbf{0} \\ \mathbf{0} & Q^{-1} \end{bmatrix} \mathbf{e}_{p+1}$$

将 Q 的表示式 (2.4.11) 代入, 又得

$$s_T^2 \mathbf{e}'_{p+1} [T^2 (\Sigma \mathbf{x}, \mathbf{x}')^{-1}] \mathbf{e}_{p+1} \Rightarrow \frac{1}{\Psi^2(1) \left\{ \int_0^1 W^2(r) dr - \left(\int_0^1 W(r) dr \right)^2 \right\}}$$

这样, t_T 统计量有极限:

$$t_T = \frac{T(\hat{\rho}_T - 1)}{\{s_T^2 \mathbf{e}'_{p+1} [T^2 (\Sigma \mathbf{x}, \mathbf{x}')^{-1}] \mathbf{e}_{p+1}\}^{1/2}} \\ \Rightarrow \frac{1}{\Psi(1)} \cdot \left\{ \frac{\frac{1}{2} \{W^2(1) - 1\} - W(1) \int_0^1 W(r) dr}{\int_0^1 W^2(r) dr - \left(\int_0^1 W(r) dr \right)^2} \right\} \\ \times \left\{ \Psi^2(1) \left[\int_0^1 W^2(r) dr - \left(\int_0^1 W(r) dr \right)^2 \right] \right\}^{1/2}$$

$$= \frac{\frac{1}{2} \{W^2(1) - 1\} - W(1) \int_0^1 W(r) dr}{\left\{ \int_0^1 W^2(r) dr - \left(\int_0^1 W(r) dr \right)^2 \right\}^{1/2}} \quad (2.4.19)$$

上式中统计量 t_T 的极限与(2.2.11)式是一致的,后者是在 $\{\epsilon_t\}$ 独立同分布的假设下得到的。这表明,ADF 检验的 t -统计量不需对统计量 t_T 作修正,这与 PP 检验法的 t -统计量(2.3.14)形成明显的对照。这一区别是 ADF 和 PP 的不同估计方法引起的。在 PP 检验法中,最小二乘估计 $\hat{\rho}_T$ 是由 y_t 对 y_{t-1} 作回归取得的,稳定过程 $\{u_t\}$ 中的自相关并未在估计中得到处理。这并不影响估计量 $\hat{\rho}_T$ 的收敛性,因为若 $\{y_t\}$ 服从单位根过程, $\hat{\rho}_T$ 是超一致估计量。但在 ADF 方法中, $\hat{\rho}_T$ 是和其他估计量 $\hat{\xi}_1, \hat{\xi}_2, \dots, \hat{\xi}_{p-1}$ 一同取得的, $\{u_t\}$ 中的自相关已得到了处理,在构造 t -统计量时不再需要作修正。

例 2.6 利用例 2.3 中的数据,估计下列模型

$$\begin{aligned} i_t = & \frac{0.335}{(0.0788)} \Delta i_{t-1} - \frac{0.388}{(0.0808)} \Delta i_{t-2} + \frac{0.276}{(0.0800)} \Delta i_{t-3} \\ & - \frac{0.107}{(0.0794)} \Delta i_{t-4} + \frac{0.195}{(0.109)} + \frac{0.96904}{(0.01860)} i_{t-1} \end{aligned}$$

其中不仅包括了 i_{t-1} 项和常数项,还包括了 Δi_{t-i} 项 ($i = 1, 2, 3, 4$)。计算 ADF 的 ρ -统计量的值:

$$\begin{aligned} \frac{T(\hat{\rho}_T - 1)}{1 - \hat{\xi}_1 - \dots - \hat{\xi}_{p-1}} &= \frac{164}{1 - 0.335 + 0.388 - 0.276 + 0.107} \\ &\quad \times (0.96904 - 1) = -5.74 \end{aligned}$$

由表 1 的情况二查得相应于 $T = 164$, 显著水平为 5% 的临界值为 -13.8 , 因为 $-5.74 > -13.8$, 美国财政部债券利率有单位根的假设再次被接受。

容易计算 t_T 统计量的值:

$$t_T = \frac{(\hat{\rho}_T - 1)}{\hat{\eta}_T} = \frac{0.96904 - 1}{0.01860} = -1.66$$

由表 2 情况二查得相应的临界值为 -2.89 , 因为 $-1.66 > -2.89$, 接受单位根假设。

还可对关于滞后项 Δi_{i-1} ($i = 1, 2, 3, 4$) 的参数作检验。比如, 检验假设 $H_0: \zeta_i = 0$, 相应的 t_T 统计量的值为:

$$t_T = \frac{\hat{\zeta}_4}{\hat{\eta}_T} = \frac{-0.107}{0.0794} = -1.35$$

由 t -分布表查得此时相应于显著水平 5% 的临界值为 -1.95 , 接受假设 $\zeta_i = 0$ 。

对原假设 $H_0: \alpha = 0, \rho = 1$ 还可用 F_T 统计量作联合检验。将 H_0 写成矩阵形式

$$H_0: R\beta = \gamma$$

这里, $\beta = [\beta_1', \beta_2']$, $\beta_1 = [\zeta_1, \dots, \zeta_{p-1}]'$, $\beta_2 = [\alpha, \rho]'$; $R = [0_{2,p-1}, I_2]$, $0_{2,p-1}$ 为 $2 \times (p-1)$ 维零矩阵, I_2 为 2×2 维的单位矩阵; $\gamma = [0, 1]'$ 。以 b_T 表示参数 β 的最小二乘估计, 可将检验 H_0 的 F_T 统计量表示为:

$$F_T = (b_T - \beta)' R' \left\{ s_T^2 R \left(\sum_{t=1}^T x_t x_t' \right)^{-1} R' \right\}^{-1} R (b_T - \beta) / 2 \quad (2.4.20)$$

F_T 统计量也可用最小二乘估计的残差表示:

$$F_T = \frac{(\hat{R}^2 - \hat{R}^2)/2}{\hat{R}^2/(T - P - 1)}$$

其中, \tilde{R}^2 和 \hat{R}^2 分别为有约束的和无约束的残差平方和:

$$\hat{R}^2 = \sum_{t=1}^T (y_t - \hat{\xi}_1 \Delta y_{t-1} - \cdots - \hat{\xi}_{p-1} \Delta y_{t-p+1} - \hat{\alpha}_T - \hat{\rho}_T y_{t-1})^2$$

$$\tilde{R}^2 = \sum_{t=1}^T (\Delta y_t - \tilde{\xi}_1 \Delta y_{t-1} - \cdots - \tilde{\xi}_{p-1} \Delta y_{t-p+1})^2$$

其中, $\tilde{\xi}_j, \hat{\xi}_j$ ($j = 1, 2, \dots, p-1$) 分别为有约束的和无约束的最小二乘估计。

定义 2×2 维和 $(p+1) \times (p+1)$ 维对角矩阵 \bar{D}_T 和 D_T 如下:

$$\bar{D}_T = \begin{bmatrix} \sqrt{T} & 0 \\ 0 & T \end{bmatrix}; \quad D_T = \begin{bmatrix} \sqrt{T} & 0 & \cdots & \cdots & 0 & 0 \\ 0 & \sqrt{T} & 0 & \cdots & 0 & 0 \\ \vdots & 0 & \ddots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{T} & 0 \\ 0 & 0 & 0 & \cdots & 0 & T \end{bmatrix}$$

可将(2.4.20)式写成以下形式:

$$F_T = (\mathbf{b}_T - \boldsymbol{\beta})' \mathbf{R}' \bar{D}_T \left\{ s_T^2 \bar{D}_T \mathbf{R} \left(\sum_{t=1}^T \mathbf{x}_t \mathbf{x}_t' \right)^{-1} \mathbf{R}' \bar{D}_T \right\}^{-1} \bar{D}_T \mathbf{R} (\mathbf{b}_T - \boldsymbol{\beta}) / 2 \quad (2.4.21)$$

由于

$$\bar{D}_T \mathbf{R} = \mathbf{R} D_T$$

而且

$$\mathbf{R} D_T (\mathbf{b}_T - \boldsymbol{\beta}) \Rightarrow \mathbf{Q}^{-1} \mathbf{h}_2$$

这里的矩阵 \mathbf{Q} 和向量 \mathbf{h}_2 由(2.4.14)定义, 就可得统计量 F_T 的

极限:

$$\begin{aligned}
 F_T &= (\mathbf{b}_T - \boldsymbol{\beta})' (\mathbf{R} \mathbf{D}_T)' \left\{ s_T^2 \mathbf{R} \mathbf{D}_T \left(\sum_{i=1}^T \mathbf{x}_i \mathbf{x}_i' \right)^{-1} \mathbf{D}_T \mathbf{R}' \right\}^{-1} \mathbf{R} \mathbf{D}_T (\mathbf{b}_T - \boldsymbol{\beta}) / 2 \\
 &\Rightarrow (\mathbf{Q}^{-1} \mathbf{h}_2)' (\sigma^2 \mathbf{Q}^{-1})^{-1} (\mathbf{Q}^{-1} \mathbf{h}_2) / 2 = \mathbf{h}_2' \mathbf{Q}^{-1} \mathbf{h}_2 / 2 \sigma^2 \\
 &= \frac{1}{2 \sigma^2} \left[\sigma W(1), \frac{1}{2} \sigma^2 \Psi(1) \{W^2(1) - 1\} \right] \\
 &\quad \times \begin{bmatrix} 1 & \sigma \Psi(1) \int_0^1 W(r) dr \\ \sigma \Psi(1) \int_0^1 W(r) dr & \sigma^2 \Psi^2(1) \int_0^1 W^2(r) dr \end{bmatrix}^{-1} \\
 &\quad \times \begin{bmatrix} \sigma W(1) \\ \frac{1}{2} \sigma^2 \Psi(1) \{W^2(1) - 1\} \end{bmatrix} \\
 &= \frac{1}{2} \left[W(1), \frac{1}{2} \{W^2(1) - 1\} \right] \begin{bmatrix} 1 & 0 \\ 0 & \sigma \Psi(1) \end{bmatrix} \\
 &\quad \times \begin{bmatrix} 1 & 0 \\ 0 & \sigma \Psi(1) \end{bmatrix}^{-1} \begin{bmatrix} 1 & \int_0^1 W(r) dr \\ \int_0^1 W(r) dr & \int_0^1 W^2(r) dr \end{bmatrix}^{-1} \\
 &\quad \times \begin{bmatrix} 1 & 0 \\ 0 & \sigma \Psi(1) \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & \sigma \Psi(1) \end{bmatrix} \begin{bmatrix} \sigma W(1) \\ \frac{1}{2} \{W^2(1) - 1\} \end{bmatrix} \\
 &= \frac{1}{2} \left[W(1), \frac{1}{2} \{W^2(1) - 1\} \right] \begin{bmatrix} 1 & \int_0^1 W(r) dr \\ \int_0^1 W(r) dr & \int_0^1 W^2(r) dr \end{bmatrix}^{-1} \\
 &\quad \times \begin{bmatrix} W(1) \\ \frac{1}{2} \{W^2(1) - 1\} \end{bmatrix} \tag{2.4.22}
 \end{aligned}$$

情况四:数据生成过程为

$$y_t = \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots + \zeta_{p-1} \Delta y_{t-p+1} + \alpha + y_{t-1} + \epsilon_t$$

其中, $\{\epsilon_t\}$ 独立同分布, $E(\epsilon_t) = 0$, $D(\epsilon_t) = \sigma^2 < \infty$ 。估计模型为

$$y_t = \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots + \zeta_{p-1} \Delta y_{t-p+1} \\ + \alpha + \rho y_{t-1} + \delta t + \epsilon_t$$

对假设 $H_0: \rho = 1$, 可用 ADF 的 ρ -统计量

$$\frac{T(\hat{\rho}_T - 1)}{1 - \hat{\zeta}_1 - \hat{\zeta}_2 - \cdots - \hat{\zeta}_{p-1}}$$

和 t -统计量

$$t_T = \frac{\hat{\rho}_T - 1}{\hat{\eta}_T}$$

作检验, 它们的临界值分别列在表 1 和表 2 的情况四中。

对任何有关参数 $\zeta_1, \cdots, \zeta_{p-1}$ 的假设可由一般的 t 统计量和 F 统计量作检验, 它们的临界值可在一般的 t -分布和 F -分布表中查得。

对联合假设 $H_0: \rho = 1, \delta = 0$, 可用统计量

$$F_T = \frac{(\tilde{R}^2 - \hat{R}^2)/2}{\hat{R}^2/(T - P - 2)}$$

作检验, 其中的 \tilde{R}^2 和 \hat{R}^2 分别为有约束的和无约束的残差平方和:

$$\tilde{R}^2 = \sum_{t=1}^T (y_t - y_{t-1} - \tilde{\zeta}_1 \Delta y_{t-1} - \cdots - \tilde{\zeta}_{p-1} \Delta y_{t-p+1} - \tilde{\alpha}_T)$$

$$\hat{R}^2 = \sum_{t=1}^T (y_t - \hat{\xi}_1 \Delta y_{t-1} - \cdots - \hat{\xi}_{p-1} \Delta y_{t-p+1} - \hat{a}_T - \hat{\rho}_T y_{t-1} - \hat{\delta}_T t)$$

“ \sim ”和“ \wedge ”分别表示有约束的和无约束的最小二乘估计。 F_T 的临界值由表 3 的情况四给出。

2.5 本章小结

本章讨论了单位根过程的假设检验,着重介绍了检验单位根的迪基—福勒(DF)方法、增广的迪基—福勒(ADF)方法和菲利普斯—配荣(PP)方法。这三种方法都基于统计量 $T(\hat{\rho}_T - 1)$ 和 t_T 的极限分布。若单位根过程 $\{y_t\}$ 的随机干扰 $\{u_t\}$ 为独立同分布,则它们的极限是标准维纳过程的泛函,且不依赖于任何未知参数;若 $\{u_t\}$ 为一般的稳定过程,ADF 和 PP 方法对统计量 $T(\hat{\rho}_T - 1)$ 和 t_T 作适当的修正,使得修正后的极限仍与 $\{u_t\}$ 为独立同分布时一样。

在运用 ADF 和 PP 方法时,需确定数据生成过程的形式,建立实际的估计模型,以及检验原假设——单一假设 $H_0: \rho = 1$ 或联合假设 $H_0: \rho = 1, \delta = 0$, 然后在情况一、二和四中选择合适的临界值作为检验的依据。对临界值的不合理的选择,不仅会引起不正确的统计结论,也会导致不正确的经济解释。

如何在 ADF 和 PP 之间作选择,不存在一般的原则。在实际中,常根据模型的结构作选择。若在估计模型中不仅包含 y_{t-1} 项,还包含了 Δy_{t-1} 项,模型对数据的动态结构已作了描述,这时用 ADF 检验较合理;若模型只描述了 y_t 和 y_{t-1} 之间的关系,对随机干扰 $\{u_t\}$ 作了较一般的假设,那么可以选用 PP 检验方法。

这里,

$$u_t = \Delta y_t - \mu, \mu = \frac{\alpha}{1 - \zeta_1 - \zeta_2 - \dots - \zeta_{p-1}}$$

$$\mu^* = (1 - \rho)\mu, \delta^* = \delta + \rho\mu$$

$$\xi_{t-1} = y_{t-1} - \mu(t-1)$$

(b) 若 $\rho = 1, \delta = 0$, 证明:

$$u_t = (1 - \zeta_1 L - \zeta_2 L^2 - \dots - \zeta_{p-1} L^{p-1})^{-1} \varepsilon_t$$

$$\xi_{t-1} = y_0 + u_1 + \dots + u_{t-1}$$

(c) 令 $r_t = E(u_t u_{t-1})$, $\lambda = \sigma(1 - \zeta_1 - \zeta_2 - \dots - \zeta_{p-1})^{-1}$,

证明: 当 $\rho = 1, \delta = 0$ 为真时, 有

$$\begin{bmatrix} T^{1/2}(\hat{\xi}_T - \xi) \\ T^{1/2} \hat{\mu}_T \\ T(\hat{\rho}_T - 1) \\ T^{3/2}(\hat{\delta}_T^* - \delta^*) \end{bmatrix} \Rightarrow \begin{bmatrix} V & 0 & 0 & 0 \\ 0' & 1 & \lambda \int_0^1 W(r) dr & 1/2 \\ 0' & \lambda \int_0^1 W(r) dr & \lambda^2 \int_0^1 W^2(r) dr & \lambda \int_0^1 W(r) dr \\ 0' & 1/2 & \lambda \int_0^1 W(r) dr & 1/3 \end{bmatrix}^{-1} \\ \times \begin{bmatrix} h_1 \\ \sigma W(1) \\ \frac{1}{2} \sigma \lambda \{W^2(1) - 1\} \\ \sigma \left\{ W(1) - \int_0^1 W(r) dr \right\} \end{bmatrix}$$

这里 $h_1 \sim N(0, \sigma^2 V)$, V 由 (2.4.10) 给出。

$[W(t_2) - W(t_1)], [W(t_3) - W(t_2)], \dots, [W(t_k) - W(t_{k-1})]$ 为独立的正态向量, 且有 $[W(s) - W(r)] \sim N(0, (s - r)I_n)$, $\forall s > r \in [0, 1]$ 。

n 维向量随机过程 $\{y_t\}$ 为一单位根过程, 若 $\Delta y_t \sim I(0)$ 为 n 维的稳定过程, 这时可将 y_t 表示为

$$y_t = y_{t-1} + V_t$$

这里, $V_t \sim I(0)$ 。进一步假设 $\{V_t\}$ 独立同分布, $E(V_t) = 0$, $D(V_t) = E(V_t V_t') = I_n$, I_n 为 $n \times n$ 维单位矩阵。构造 V_t 的部分和均值 \bar{X}_T^* :

$$\bar{X}_T^* = T^{-1}(V_1 + V_2 + \dots + V_{[Tr]})$$

其中 $r \in [0, 1]$ 。由于 $\{V_t\}$ 的独立性, 它的任一分量 $\{V_{it}\}$ 也独立同分布, 且有 $E(V_{it}) = 0$, $D(V_{it}) = 1$ 。根据单变量的泛函中心极限定理, 对给定的 i , $1 \leq i \leq n$, 有

$$\sqrt{T} \bar{X}_{iT}^* = \frac{1}{\sqrt{T}}(V_{i1} + V_{i2} + \dots + V_{i[Tr]}) \Rightarrow W_i(r)$$

因此对向量 \bar{X}_T^* 就有

$$\sqrt{T} \bar{X}_T^* = T^{-1}(V_1 + \dots + V_{[Tr]}) \Rightarrow W(r) \quad (3.2.1)$$

$\{V_t\}$ 为最简单的 n 维独立同分布的稳定过程, 它的各分量互不相关, $E(V_{it} V_{jt}) = 0$, $i \neq j$ 。以下考虑 n 维独立同分布的稳定过程 $\{\varepsilon_t\}$, $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = E(\varepsilon_t \varepsilon_t') = \Omega$, Ω 为 $n \times n$ 维非异正定矩阵。若 $\Omega \neq I_n$, ε_t 的各分量相关, 可将正定矩阵作分解:

$$\Omega = PP'$$

其中 P 为 $n \times n$ 维非异矩阵, 从而可将 ε_t 表示为

$$\varepsilon_t = PV_t$$

对 $\{\epsilon_t\}$ 作部分和均值 X_T^*

$$\begin{aligned} X_T^* &= T^{-1}(\epsilon_1 + \epsilon_2 + \cdots + \epsilon_{[Tr]}) = T^{-1} \sum_{i=1}^{[Tr]} \epsilon_i \\ &= T^{-1} \sum_{i=1}^{[Tr]} P V_i = P \left\{ T^{-1} \sum_{i=1}^{[Tr]} V_i \right\} \\ &= P \cdot \tilde{X}_T^* \end{aligned}$$

根据 (3.2.1), 则有

$$\sqrt{T} X_T^* = P \{ \sqrt{T} \tilde{X}_T^* \} \Rightarrow P \cdot W(r) \quad (3.2.2)$$

$PW(r)$ 称为方差为 Ω 的 n 维维纳过程。对于给定的 $r \in [0, 1]$, 有

$$\begin{aligned} PW(r) &\sim P \cdot N(0, rI_n) \\ &= N(0, rPP') = N(0, r\Omega) \end{aligned}$$

(3.2.2) 的结果可进一步推广到一般的稳定过程。设 $\{u_t\}$ 为 n 维稳定过程, 并有表示形式:

$$u_t = \sum_{s=0}^{\infty} \Psi_s \epsilon_{t-s} \quad (3.2.3)$$

其中, $\{\epsilon_t\}$ 独立同分布, $E(\epsilon_t) = 0$, $D(\epsilon_t) = E(\epsilon_t \epsilon_t') = \Omega$; $\Psi_s (s = 1, 2, \cdots)$ 为 $n \times n$ 维矩阵。以 $\Psi_{ij}^{(s)}$ 表示第 s 个矩阵的第 i, j 个元素, 对给定的 i 和 j , 序列 $\Psi_{ij}^{(s)} (s = 1, 2, \cdots)$ 满足条件

$$\sum_{s=0}^{\infty} s \cdot |\Psi_{ij}^{(s)}| < \infty \quad (3.2.4)$$

以下引理给出了多变量稳定过程的 BN 分解形式。

引理 3.1: 多变量稳定过程的 BN 分解

设 $u_i = \Psi(L)\epsilon_i = \sum_{s=0}^{\infty} \Psi_s \epsilon_{i-s}$ 为 n 维稳定过程, 其中 $\{\epsilon_i\}$ 独立同分布, $E(\epsilon_i) = 0$, $D(\epsilon_i) = E(\epsilon_i \epsilon_i') = \Omega$ 。若 $\{\Psi_i\}$ 中每一元素组成的序列满足条件 (3.2.4), $\{u_i\}$ 的部分和 $\sum_{i=1}^T u_i$ 有以下 BN 分解:

$$\sum_{i=1}^T u_i = \Psi(1) \sum_{i=1}^T \epsilon_i + \eta_T - \eta_0 \quad (3.2.5)$$

其中, $\Psi(1) = (\Psi_0 + \Psi_1 + \cdots)$, $\eta_T = \sum_{i=1}^{\infty} \alpha_i \epsilon_{T-i+1}$, $\alpha_i = -(\Psi_{i+1} + \Psi_{i+2} + \cdots)$; $\{\eta_T, T = 1, 2, \cdots\}$ 为稳定过程, 矩阵列 $\{\alpha_i\}$ 的每一元素列 $\{\alpha_{ij}^{(s)}, s = 1, 2, \cdots\}$ 绝对可和:

$$\sum_{s=1}^{\infty} |\alpha_{ij}^{(s)}| < \infty$$

以下将第 1 章的定理 1.9 推广为向量形式。

对于 n 维稳定过程 $\{u_i\}$ 构造部分和均值:

$$X_T(r) = T^{-1} \sum_{i=1}^{[Tr]} u_i$$

并可得 BN 分解形式

$$X_T(r) = T^{-1} \left\{ \Psi(1) \sum_{i=1}^{[Tr]} \epsilon_i + \eta_{[Tr]} - \eta_0 \right\} \quad (3.2.6)$$

$(\eta_{[Tr]} - \eta_0)$ 的每一分量都为—稳定过程, 因此有

$$\sup_{r \in [0, 1]} T^{-1/2} |\eta_{i, [Tr]} - \eta_{i0}| \xrightarrow{p} 0; \quad i = 1, 2, \cdots, n$$

以 \sqrt{T} 乘 $X_T(r)$, 并令 $T \rightarrow \infty$, 得

$$\begin{aligned}\sqrt{T}X_T(r) &= \frac{1}{\sqrt{T}}(\Psi(1) \sum_{i=1}^{[Tr]} \varepsilon_i) + \frac{1}{\sqrt{T}}(\eta_{[Tr]} - \eta_0) \\ &\Rightarrow \Psi(1)PW(r)\end{aligned}\quad (3.2.7)$$

其中, P 为 $n \times n$ 维非异矩阵, 且 $PP' = \Omega$ 。对于任意给定的 $r \in [0, 1]$, $W(r) \sim N(0, rI_n)$, 这样 $\Psi(1)PW(r)$ 就有分布:

$$\begin{aligned}\Psi(1)PW(r) &\sim N(0, r\Psi(1)PP'\Psi(1)') \\ &= N(0, r\Psi(1)\Omega\Psi(1)')\end{aligned}$$

(3.2.7)称为多变量的泛函中心极限定理。

作为多变量泛函中心极限定理的应用, 考虑以下例子。若 $\{\xi_t\}$ 为 n 维单位根过程, $\xi_t = \xi_{t-1} + u_t$, ξ_t 是 t 个随机干扰的和:

$$\xi_t = \sum_{j=1}^t u_j$$

作 ξ_t 的部分和 $\sum_{i=1}^T \xi_i$, 并进行 BN 分解, 可得

$$\begin{aligned}\sum_{i=1}^T \xi_i &= \sum_{i=1}^T \left(\sum_{j=1}^i u_j \right) = \sum_{i=1}^T (\Psi(1) \sum_{j=1}^i \varepsilon_j + \eta_i - \eta_0) \\ &= \sum_{i=1}^T \left(\Psi(1) \sum_{j=1}^i PV_j \right) + \sum_{i=1}^T (\eta_i - \eta_0) \\ &= \Psi(1) \sum_{i=1}^T \delta_i + \sum_{i=1}^T (\eta_i - \eta_0)\end{aligned}$$

这里, $\delta_i = \sum_{j=1}^i V_j$, 所以 $\delta_i = \delta_{i-1} + V_i$, $\{V_i\}$ 为独立同分布, $E(V_i) = 0$, $D(V_i) = E(V_i V_i') = I_n$ 。由 $(\eta_i - \eta_0)$ 的稳定性, 可知:

$$\sup_{r \in [0, 1]} T^{-1/2} |\eta_{i, [Tr]} - \eta_{i, 0}| \xrightarrow{P} 0; i = 1, 2, \dots$$

从而有

$$T^{-3/2} \sum_{i=1}^T (\eta_i - \eta_0) \xrightarrow{P} 0$$

所以以下结论成立:

$$\begin{aligned} T^{-3/2} \sum_{i=1}^T \xi_i &= \Psi(1) \cdot P \left\{ T^{-3/2} \sum_{i=1}^T \delta_i \right\} + T^{-1/2} \sum_{i=1}^T (\eta_i - \eta_0) \\ &\Rightarrow \Psi(1) P \int_0^1 W(r) dr \end{aligned} \quad (3.2.8)$$

这是定理 1.9 的结论(6)的推广。

以下定义向量形式的随机积分。

若 $\delta_i = V_1 + V_2 + \cdots + V_i$, 其中每一分量构成序列 $\{\delta_u$, $u = 1, 2, \dots\}$, 根据定理 1.9 的结论(5), 有

$$T^{-1} \sum_{i=1}^T \delta_{i, i-1} V_u \Rightarrow \frac{1}{2} \{W_i^2(1) - 1\}; \quad i = 1, 2, \dots, n$$

根据 1.6 节关于随机积分的讨论, 下列关系式成立:

$$\int_0^1 W_i(r) dW_i(r) = \frac{1}{2} \{W_i^2(1) - 1\}$$

从而有

$$T^{-1} \sum_{i=1}^T \delta_{i, i-1} V_u \Rightarrow \int_0^1 W_i(r) dW_i(r); \quad i = 1, 2, \dots, n \quad (3.2.9)$$

对于 $i \neq j$, 可推得相应的极限

$$T^{-1} \sum_{i=1}^T \delta_{i, i-1} V_u \Rightarrow \int_0^1 W_i(r) dW_j(r) \quad (3.2.10)$$

综合(3.2.9)和(3.2.10)可获得以下结果

$$T^{-1} \sum_{i=1}^T \delta_{t-1} V'_i \Rightarrow \int_0^1 W(r) dW'(r) \\ = \frac{1}{2} \{W(1)W'(1) - I_n\} \quad (3.2.11)$$

这一随机积分的结果以后常会用到。

定理 3.1: 向量形式的定理 1.9

设 u_t 为 n 维随机向量, 有表示式

$$u_t = \Psi(L)\epsilon_t = \sum_{s=0}^{\infty} \Psi_s \epsilon_{t-s}$$

其中矩阵列 $\{\Psi_s\}$ 满足条件:

$$\sum_{s=0}^{\infty} s |\Psi_{ij}^{(s)}| < \infty; \quad i, j = 1, 2, \dots, n$$

$\{\epsilon_t\}$ 独立同分布, $E(\epsilon_t) = 0$, $D(\epsilon_t) = E(\epsilon_t \epsilon'_t) = \Omega$, $\Omega = PP'$ 为非异正定。 $\sigma_{ij} = E(\epsilon_{it} \epsilon_{jt})$ 为矩阵 Ω 的第 i, j 个元素, 令

$$\Gamma_s = E(u_t u'_{t-s}) = \sum_{k=0}^{\infty} \Psi_{s+k} \Omega \Psi'_k; \quad s = 0, 1, 2, \dots$$

$$z_t = [u'_{t-1}, u'_{t-2}, \dots, u'_{t-k}]'; \quad k \geq 1$$

$$V = E(z_t z'_t) = \begin{bmatrix} \Gamma_0 & \Gamma_1 & \dots & \Gamma_{k-1} \\ \Gamma_{-1} & \Gamma_0 & \dots & \Gamma_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_{-k+1} & \Gamma_{-k+2} & \dots & \Gamma_0 \end{bmatrix}$$

$$\Lambda = \Psi(1) \cdot P = (\Psi_0 + \Psi_1 + \dots) \cdot P; \quad \xi_t = u_1 + \dots + u_t$$

若 $\xi_0 = 0$, 以下结论成立:

$$(1) T^{-1/2} \sum_{i=1}^T u_i \Rightarrow \Lambda W(1)$$

$$(2) T^{-1/2} \sum_{i=1}^T z_i \varepsilon_{it} \xrightarrow{d} N(0, \sigma_{\varepsilon} V); i = 1, 2, \dots, n$$

$$(3) T^{-1} \sum_{i=1}^T u_i u'_{i-s} \xrightarrow{p} \Gamma_s; s = 0, 1, 2, \dots$$

$$(4) T^{-1} \sum_{i=1}^T (\xi_{i-1} u'_{i-s} + u_{i-s} \xi'_{i-1})$$

$$\Rightarrow \begin{cases} \Lambda W(1)W(1)' \Lambda' - \Gamma_0 & s = 0 \\ \Lambda W(1)W(1)' \Lambda' + \sum_{k=-s+1}^{s-1} \Gamma_k & s = 1, 2, \dots \end{cases}$$

$$(5) T^{-1} \sum_{i=1}^T \xi_{i-1} u'_i \Rightarrow \Lambda \left\{ \int_0^1 W(r) [dW(r)]' \Lambda' + \sum_{k=1}^{\infty} \Gamma'_k \right\}$$

$$(6) T^{-1} \sum_{i=1}^T \xi_{i-1} \varepsilon'_i \Rightarrow \Lambda \left\{ \int_0^1 W(r) [dW(r)]' \right\} P'$$

$$(7) T^{-3/2} \sum_{i=1}^T \xi_{i-1} \Rightarrow \Lambda \int_0^1 W(r) dr$$

$$(8) T^{-3/2} \sum_{i=1}^T t u_{i-s} \Rightarrow \Lambda \left\{ W(1) - \int_0^1 W(r) dr \right\};$$

$$s = 0, 1, 2, \dots$$

$$(9) T^{-2} \sum_{i=1}^T \xi_{i-1} \xi'_{i-1} \Rightarrow \Lambda \left\{ \int_0^1 W(r) W(r)' dr \right\} \Lambda'$$

$$(10) T^{-5/2} \sum_{i=1}^T t \xi_{i-1} \Rightarrow \Lambda \int_0^1 r W(r) dr$$

$$(11) T^{-3} \sum_{i=1}^T t \xi_{i-1} \xi'_{i-1} \Rightarrow \Lambda \left\{ \int_0^1 r W(r) W(r)' dr \right\} \Lambda'$$

证明:

(1) 在(3.2.7)中令 $r = 1$, 即得结论。

(2) 利用一般的中心极限定理, 即得结论。

(3) 利用大数定律, 即得结论。

(4) 由于

$$\begin{aligned}\sum_{i=1}^T \xi_i \xi'_i &= \sum_{i=1}^T (\xi_{i-1} + u_i)(\xi'_{i-1} + u'_i) \\ &= \sum_{i=1}^T (\xi_{i-1} \xi'_{i-1} + \xi_{i-1} u'_i + u_i \xi'_{i-1} + u_i u'_i)\end{aligned}$$

对 $s = 0$, 可得

$$\begin{aligned}\sum_{i=1}^T (\xi_{i-1} u'_i + u_i \xi'_{i-1}) &= \sum_{i=1}^T \xi_i \xi'_i - \sum_{i=1}^T \xi_{i-1} \xi'_{i-1} - \sum_{i=1}^T u_i u'_i \\ &= \xi_T \xi'_T - \sum_{i=1}^T u_i u'_i\end{aligned}$$

由 (3.2.6) $X_T(r)$ 的 BN 分解形式:

$$\xi_T = T X_T(1)$$

可得

$$\begin{aligned}T^{-1} \sum_{i=1}^T (\xi_{i-1} u'_i + u_i \xi'_{i-1}) &= T^{-1} \xi_T \xi'_T - T^{-1} \sum_{i=1}^T u_i u'_i \\ &= [\sqrt{T} X_T(1)][\sqrt{T} X_T(1)]' - T^{-1} \sum_{i=1}^T u_i u'_i\end{aligned}$$

因为

$$\sqrt{T} X_T(1) \Rightarrow \Psi(1) P W(1) = \Lambda W(1)$$

$$T^{-1} \sum_{i=1}^T u_i u'_i \xrightarrow{p} \Gamma_0$$

所以

$$T^{-1} \sum_{i=1}^T (\xi_{i-1} u'_i + u_i \xi'_{i-1}) \Rightarrow \Lambda W(1) W(1)' \Lambda' - \Gamma_0$$

另一方面, 对 $s > 0$, 有

$$\begin{aligned}
& T^{-1} \sum_{t=s+1}^T (\xi_{t-1} u'_{t-s} + u_{t-s} \xi'_{t-1}) \\
&= T^{-1} \sum_{t=s+1}^T \{ (\xi_{t-s-1} + u_{t-s} + u_{t-s-1} + \cdots + u_{t-1}) u'_{t-s} \\
&\quad + u_{t-s} (\xi'_{t-s-1} + u'_{t-s} + u'_{t-s+1} + \cdots + u'_{t-1}) \} \\
&= T^{-1} \sum_{t=s+1}^T (\xi_{t-s-1} u'_{t-s} + u_{t-s} \xi'_{t-s-1}) \\
&\quad + T^{-1} \sum_{t=s+1}^T \{ (u_{t-s} u'_{t-s}) + \cdots + (u_{t-1} u'_{t-s}) \\
&\quad + (u_{t-s} u'_{t-1}) + (u_{t-s} u'_{t-s+1}) + \cdots + (u_{t-1} u'_{t-1}) \}
\end{aligned}$$

从而有

$$\begin{aligned}
& T^{-1} \sum_{t=1}^T (\xi_{t-1} u'_{t-s} + u_{t-s} \xi'_{t-1}) \\
&= T^{-1} \left(\sum_{t=1}^s + \sum_{t=s+1}^T \right) (\xi_{t-1} u'_{t-s} + u_{t-s} \xi'_{t-1}) \\
&= \sum_{t=s+1}^T (\xi_{t-1} u'_{t-s} + u_{t-s} \xi'_{t-1}) + o_p(T^{-1}) \\
&\Rightarrow \Lambda W(1)W(1)' \Lambda' - \Gamma_0 + [\Gamma_0 + \Gamma_1 + \cdots \\
&\quad + \Gamma_{s-1} + \Gamma_0 - \Gamma_{-1} + \cdots + \Gamma_{-s+1}] \\
&= \Lambda W(1)W(1)' \Lambda' + \sum_{k=-s+1}^{s-1} \Gamma_k
\end{aligned}$$

结论得证。

(5) 当 $s = 0$, 可将结论(4)改写成:

$$T^{-1} \sum_{t=1}^T (\xi_{t-1} u'_t + u_t \xi'_{t-1}) \Rightarrow \Lambda \{W(1)W(1)' - I_n\} \Lambda' + (\Lambda \Lambda' - \Gamma_0)$$

因为

$$\begin{aligned}\Lambda\Lambda' &= \Psi(1)PP' \Psi(1)' = \Psi(1)\Omega\Psi'(1) \\ &= \Gamma_0 + \sum_{k=1}^{\infty} (\Gamma_k + \Gamma_k')\end{aligned}$$

根据(3.2.11),就可得

$$\begin{aligned}& T^{-1} \sum_{i=1}^T (\xi_{i-1}u_i' + u_i\xi_{i-1}') \\ & \Rightarrow 2\Lambda \left\{ \int_0^1 W(r)[dW(r)]' \right\} \Lambda' + \sum_{k=1}^{\infty} (\Gamma_k + \Gamma_k')\end{aligned}$$

结论(5)得证。

(6) 因为 $\xi_{i-1} = \sum_{j=1}^{i-1} u_j$, $u_i = \Psi(L)\varepsilon_i$, 对 u_i 作 BN 分解, 就可得

$$\sum_{i=1}^T \xi_{i-1}\varepsilon_i' = \Psi(1) \sum_{i=1}^T \delta_{i-1}^*\varepsilon_i' + \sum_{i=1}^T (\eta_{i-1} - \eta_0)\varepsilon_i' \quad (3.2.12)$$

其中, $\delta_{i-1}^* = \sum_{j=1}^{i-1} \varepsilon_j$, $(\eta_{i-1} - \eta_0) \sim I(0)$ 。将 δ_{i-1}^* 、 ε_i 改写成

$$\delta_{i-1}^* = P\delta_{i-1}; \quad \varepsilon_i = PV_i$$

$\{V_i\}$ 独立同分布, $E(V_i) = 0$, $D(V_i) = E(V_iV_i') = I_n$, $\delta_{i-1} = \sum_{j=1}^{i-1} V_j$, 代入(3.2.12), 可得

$$\begin{aligned}T^{-1} \sum_{i=1}^T \xi_{i-1}\varepsilon_i' &= \Psi(1)P \left\{ T^{-1} \sum_{i=1}^T \delta_{i-1}V_i' \right\} P' + T^{-1} \sum_{i=1}^T (\eta_{i-1} - \eta_0)\varepsilon_i' \\ &\Rightarrow \Lambda \left\{ \int_0^1 W(r)[dW(r)]' \right\} P'\end{aligned}$$

(9) 由(3.2.8)中的推导, 将 $T^{-2} \sum_{i=1}^T \xi_{i-1} \xi'_{i-1}$ 写成以下形式:

$$T^{-2} \sum_{i=1}^T \xi_{i-1} \xi'_{i-1} = \int_0^1 [\sqrt{T} X_T(r)] [\sqrt{T} X_T(r)]' dr$$

令 $T \rightarrow \infty$, 则

$$T^{-2} \sum_{i=1}^T \xi_{i-1} \xi'_{i-1} \Rightarrow \Lambda \left\{ \int_0^1 W(r) W(r)' dr \right\} \Lambda'$$

结论得证。

(10) 先将 $T^{-5/2} \sum_{i=1}^T t \xi_{i-1}$ 改写成下式:

$$T^{-5/2} \sum_{i=1}^T t \xi_{i-1} = T^{1/2} \sum_{i=1}^T \left(\frac{t}{T} \right) (\xi_{i-1}/T^2)$$

根据随机向量积分的定义, 参照定理 1.9 结论(9)的证明方法, 可得

$$\begin{aligned} T^{-5/2} \sum_{i=1}^T t \xi_{i-1} &= T^{1/2} \int_0^1 \left\{ \frac{[Tr] + 1}{T} \right\} \{ (u_1 + \cdots + u_{[Tr]})/T \} dr \\ &= T^{1/2} \int_0^1 \frac{[Tr] + 1}{T} X_T(r) dr \end{aligned}$$

因为

$$\frac{[Tr] + 1}{T} \longrightarrow r; \quad \sqrt{T} X_T(r) \Rightarrow \Psi(1) P \int_0^1 W(r) dr$$

所以

$$T^{-5/2} \sum_{i=1}^T t \xi_{i-1} \Rightarrow \Lambda \int_0^1 r W(r) dr$$

结论(10)得证。

(11) 与结论(10)的证明相似,将 $T^{-3} \sum_{i=1}^T t \xi_{i-1} \xi'_{i-1}$ 改写成下式:

$$T^{-3} \sum_{i=1}^T t \xi_{i-1} \xi'_{i-1} = T \sum_{i=1}^T \left(\frac{t}{T} \right) (\xi_{i-1}/T) (\xi'_{i-1}/T) \left(\frac{1}{T} \right)$$

由随机积分的定义,可得

$$\begin{aligned} T^{-3} \sum_{i=1}^T t \xi_{i-1} \xi'_{i-1} &= T \sum_{i=1}^T \left(\frac{t}{T} \right) (\xi_{i-1}/T) (\xi'_{i-1}/T) \left(\frac{1}{T} \right) \\ &= T \int_0^1 \left\{ \frac{[Tr] + 1}{T} \right\} \{ (u_1 + \cdots + u_{[Tr]})/T \} \\ &\quad \times \{ (u_1 + \cdots + u_{[Tr]})'/T \} dr \\ &= \int_0^1 \left\{ \frac{[Tr] + 1}{T} \right\} \{ \sqrt{T} X_T(r) \} \{ \sqrt{T} X_T(r) \}' dr \end{aligned}$$

由于

$$\frac{[Tr] + 1}{T} \longrightarrow r; \quad \sqrt{T} X_T(r) \Rightarrow \Psi(1)P \int_0^1 W(r) dr$$

从而有

$$T^{-3} \sum_{i=1}^T t \xi_{i-1} \xi'_{i-1} \Rightarrow \Lambda \left\{ \int_0^1 r W(r) W(r)' dr \right\} \Lambda'$$

结论(11)得证。定理证毕。

3.3 含单位根的向量自回归过程

向量自回归过程(vector autoregressive process),简称 VAR 过程,是分析多变量时间序列的有力工具。以下主要介绍含有单位根的 VAR 过程,将以前的讨论进一步推广。VAR 过

$$= [I_n - (\Psi_p) + (\Psi_{p-1} + \Psi_p)]L^{p-1} - (\Psi_p)L^p \\ = I_n - \Psi_1 L - \Psi_2 L^2 - \cdots - \Psi_{p-1} L^{p-1} - \Psi_p L^p$$

利用关系式(3.3.3)可将 $\text{VAR}(P)$ 表示成下列形式:

$$[I_n - \Psi_1 L - \Psi_2 L^2 - \cdots - \Psi_p L^p]y_t \\ = [(I_n - \rho L) - (\zeta_1 L + \zeta_2 L^2 + \cdots + \zeta_{p-1} L^{p-1})(1 - L)]y_t \\ = y_t - \rho y_{t-1} - \zeta_1 \Delta y_{t-1} - \zeta_2 \Delta y_{t-2} - \cdots - \zeta_{p-1} \Delta y_{t-p+1} \\ = \alpha + \varepsilon_t$$

从而有

$$y_t = \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots + \zeta_{p-1} \Delta y_{t-p+1} + \alpha + \rho y_{t-1} + \varepsilon_t \quad (3.3.4)$$

若在以上的 $\text{VAR}(P)$ 中, 随机向量 y_t 的每一个分量都含有单位根, 即

$$\Delta y_t = y_t - y_{t-1} \sim I(0)$$

那么, $\rho = I_n$, 且有

$$\Psi_1 + \Psi_2 + \cdots + \Psi_p = I_n$$

3.3.2 不带常数项的 $\text{VAR}(P)$

以下假设数据的生成过程是 n 维不带常数项的 $\text{VAR}(P)$ 过程:

$$y_t = \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots + \zeta_{p-1} \Delta y_{t-p+1} + \rho y_{t-1} + \varepsilon_t \quad (3.3.5)$$

其中, ρ 和 $\zeta_1, \cdots, \zeta_{p-1}$ 为 $n \times n$ 维矩阵, $\{\varepsilon_t\}$ 为独立同分布, $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = E(\varepsilon_t \varepsilon_t') = \Omega = PP'$, P 为 $n \times n$ 维满秩矩阵。

由矩阵 $\zeta_1, \cdots, \zeta_{p-1}$ 构成的特征方程

$$|I_n - \xi_1 z - \xi_2 z^2 - \cdots - \xi_{p-1} z^{p-1}| = 0$$

的根都在单位圆以外。这样,当 $\rho = I_n$ 时,可将数据生成过程 (3.3.5) 转换成无穷阶的移动平均过程:

$$\begin{aligned} \Delta y_t = u_t &= (I_n - \xi_1 L - \xi_2 L^2 - \cdots - \xi_{p-1} L^{p-1})^{-1} \varepsilon_t \\ &= \sum_{s=0}^{\infty} \Psi_s \varepsilon_{t-s} = \Psi(L) \varepsilon_t \end{aligned} \quad (3.3.6)$$

和以前讨论单变量单位根过程一样,给定不带常数项的 VAR(P) 作为数据生成过程,将常数项 α 包含在估计模型中:

$$y_t = \xi_1 \Delta y_{t-1} + \xi_2 \Delta y_{t-2} + \cdots + \xi_{p-1} \Delta y_{t-p+1} + \alpha + \rho y_{t-1} + \varepsilon_t \quad (3.3.7)$$

然后利用最小二乘估计 $\hat{\alpha}_T$ 和 $\hat{\rho}_T$ 对假设 $\alpha = 0$ 和 $\rho = I_n$ 作检验。

以 y_{it} 表示向量 y_t 的第 i 个元素, $i = 1, 2, \dots, n$; ξ'_i 表示矩阵 ξ_i 的第 i 行元素,这里, $s = 1, 2, \dots, p-1, i = 1, 2, \dots, n$; ρ'_i 表示矩阵 ρ 的第 i 行。若对随机向量 y_t 的单位根假设成立, $\rho = I_n$, 那么 $\rho'_i = e'_i$, 这里 e'_i 为 $n \times n$ 维单位矩阵 I_n 的第 i 行。以下先考虑对 (3.3.7) 中的第 i 个方程的参数作估计。第 i 个方程 ($i = 1, 2, \dots, n$) 可表示为:

$$\begin{aligned} y_{it} &= \xi'_{i1} \Delta y_{t-1} + \xi'_{i2} \Delta y_{t-2} + \cdots \\ &\quad + \xi'_{i,p-1} \Delta y_{t-p+1} + \alpha_i + \rho'_i y_{t-1} + \varepsilon_{it} \end{aligned} \quad (3.3.8)$$

若单位根假设 $H_0: \rho'_i = e'_i$ 成立,上式中的 Δy_{t-i} 可由 u_t , ($i = 1, 2, \dots, p-1$) 代替:

$$\begin{aligned} y_{it} &= \xi'_{i1} u_{t-1} + \xi'_{i2} u_{t-2} + \cdots \\ &\quad + \xi'_{i,p-1} u_{t-p+1} + \alpha_i + \rho'_i y_{t-1} + \varepsilon_{it} \end{aligned}$$

将上式中的参数 $\xi_{i1}, \dots, \xi_{i,p-1}, \alpha_i, \rho_i$ 列成一行向量:

$$\beta'_i = [\xi'_{i1}, \xi'_{i2}, \dots, \xi'_{i,p-1}, \alpha_i, \rho'_i]$$

并令

$$x'_t = [u'_{t-1}, u'_{t-2}, \dots, 1, y'_{t-1}]$$

可将(3.3.8)写成向量形式:

$$y_{it} = \beta' x_t + \epsilon_{it}; \quad i = 1, 2, \dots, n; t = 1, 2, \dots, T$$

参数向量 β 可由最小二乘法估计得

$$b_T = \left[\sum_{t=1}^T x_t x'_t \right]^{-1} \left[\sum_{t=1}^T x_t y_{it} \right] = \beta + \left[\sum_{t=1}^T x_t x'_t \right]^{-1} [x_t \epsilon_{it}] \quad (3.3.9)$$

其中的矩阵 $\left(\sum_{t=1}^T x_t x'_t \right)$ 可被展开为

$$\begin{pmatrix} \sum u_{t-1} u'_{t-1} & \sum u_{t-1} u'_{t-2} & \cdots & \sum u_{t-1} u'_{t-p+1} & \sum u_{t-1} & \sum u_{t-1} y'_{t-1} \\ \sum u_{t-2} u'_{t-1} & \sum u_{t-2} u'_{t-2} & \cdots & \sum u_{t-2} u'_{t-p+1} & \sum u_{t-2} & \sum u_{t-2} y'_{t-1} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ \sum u_{t-p+1} u'_{t-1} & \sum u_{t-p+1} u'_{t-2} & \cdots & \sum u_{t-p+1} u'_{t-p+1} & \sum u_{t-p+1} & \sum u_{t-p+1} y'_{t-1} \\ \sum u'_{t-1} & \sum u'_{t-2} & \cdots & \sum u'_{t-p+1} & T & \sum y'_{t-1} \\ \sum y_{t-1} u'_{t-1} & \sum y_{t-1} u'_{t-2} & \cdots & \sum y_{t-1} u'_{t-p+1} & \sum y_{t-1} & \sum y_{t-1} y'_{t-1} \end{pmatrix}$$

还可将向量 $\left(\sum_{t=1}^T x_t \epsilon_{it} \right)$ 表示为:

$$\begin{pmatrix} \sum u_{t-1} \epsilon_{it} \\ \sum u_{t-2} \epsilon_{it} \\ \vdots \\ \sum u_{t-p+1} \epsilon_{it} \\ \sum \epsilon_{it} \\ \sum y_{t-1} \epsilon_{it} \end{pmatrix}$$

当单位根假设成立时, b_T 是 β 的一致估计。 $(b_T - \beta)$ 有表

示式:

$$(\mathbf{b}'_r - \beta_r) = \begin{bmatrix} \hat{\xi}_{r,1} - \xi_{r,1} \\ \hat{\xi}_{r,2} - \xi_{r,2} \\ \vdots \\ \hat{\xi}_{r,p-1} - \xi_{r,p-1} \\ \hat{\alpha}_r \\ \hat{\rho}_r - e_r \end{bmatrix}$$

定义对角矩阵 \mathbf{D}_T :

$$\mathbf{D}_T = \begin{bmatrix} T^{1/2} \mathbf{I}_{n(p-1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0}' & T^{1/2} & \mathbf{0}' \\ \mathbf{0} & \mathbf{0} & T \mathbf{I}_n \end{bmatrix} \quad (3.3.10)$$

其中, $\mathbf{I}_{n(p-1)}$ 和 \mathbf{I}_n 分别为 $[n(p-1)] \times [n(p-1)]$ 维和 $n \times n$ 维单位矩阵, \mathbf{D}_T 为 $(np+1) \times (np+1)$ 维满秩的对角矩阵。

将矩阵 \mathbf{D}_T 左乘方程(3.3.9), 可得

$$\mathbf{D}_T [\mathbf{b}'_r - \beta_r] = [\mathbf{D}_T^{-1} \left(\sum_{t=1}^T \mathbf{x}_t \mathbf{x}'_t \right) \mathbf{D}_T^{-1}] \cdot \left[\mathbf{D}_T^{-1} \sum_{t=1}^T \mathbf{x}_t \varepsilon_t \right] \quad (3.3.11)$$

根据定理 3.1 的结论(1)、(3)、(4)、(7)和(9), 当 $T \rightarrow \infty$, 有

$$\begin{aligned} & \mathbf{D}_T^{-1} \left(\sum_{t=1}^T \mathbf{x}_t \mathbf{x}'_t \right) \mathbf{D}_T^{-1} \\ &= \begin{bmatrix} T^{-1} \sum_{t=1}^T \mathbf{u}_{t-1} \mathbf{u}'_{t-1} & \cdots & T^{-1} \sum_{t=1}^T \mathbf{u}_{t-1} \mathbf{u}'_{t-p+1} & T^{-1} \sum_{t=1}^T \mathbf{u}_{t-1} & T^{-3/2} \sum_{t=1}^T \mathbf{u}_{t-1} \mathbf{y}'_{t-1} \\ T^{-1} \sum_{t=1}^T \mathbf{u}_{t-2} \mathbf{u}'_{t-1} & \cdots & T^{-1} \sum_{t=1}^T \mathbf{u}_{t-2} \mathbf{u}'_{t-p+1} & T^{-1} \sum_{t=1}^T \mathbf{u}_{t-2} & T^{-3/2} \sum_{t=1}^T \mathbf{u}_{t-2} \mathbf{y}'_{t-1} \\ \vdots & & \vdots & \vdots & \vdots \\ T^{-1} \sum_{t=1}^T \mathbf{u}_{t-p+1} \mathbf{u}'_{t-1} & \cdots & T^{-1} \sum_{t=1}^T \mathbf{u}_{t-p+1} \mathbf{u}'_{t-p+1} & T^{-1} \sum_{t=1}^T \mathbf{u}_{t-p+1} & T^{-3/2} \sum_{t=1}^T \mathbf{u}_{t-p+1} \mathbf{y}'_{t-1} \\ T^{-1} \sum_{t=1}^T \mathbf{u}'_{t-1} & \cdots & T^{-1} \sum_{t=1}^T \mathbf{u}'_{t-p+1} & 1 & T^{-3/2} \sum_{t=1}^T \mathbf{y}'_{t-1} \\ T^{-3/2} \sum_{t=1}^T \mathbf{y}_{t-1} \mathbf{u}'_{t-1} & \cdots & T^{-3/2} \sum_{t=1}^T \mathbf{y}_{t-1} \mathbf{u}'_{t-p+1} & T^{-3/2} \sum_{t=1}^T \mathbf{y}_{t-1} & T^{-2} \sum_{t=1}^T \mathbf{y}_{t-1} \mathbf{y}'_{t-1} \end{bmatrix} \end{aligned}$$

$$\Rightarrow \begin{pmatrix} V & 0 \\ 0 & Q \end{pmatrix} \quad (3.3.12)$$

这里

$$V = \begin{pmatrix} \Gamma_0 & \Gamma_1 & \cdots & \Gamma_{p-2} \\ \Gamma_{-1} & \Gamma_0 & \cdots & \Gamma_{p-3} \\ \vdots & \vdots & \cdots & \vdots \\ \Gamma_{-p-2} & \Gamma_{-p+3} & \cdots & \Gamma_0 \end{pmatrix};$$

$$\Gamma_s = E(u, u'_{t-s}); \quad s = 0, 1, \dots, p-2$$

$$Q = \begin{pmatrix} 1 & \left(\int_0^1 W(r) dr \right)' \Lambda' \\ \Lambda \int_0^1 W(r) dr & \Lambda \left\{ \int_0^1 W(r) W(r)' dr \right\} \Lambda' \end{pmatrix} \quad (3.3.13)$$

这里

$$\Lambda = (I_n - \zeta_1 - \zeta_2 - \cdots - \zeta_{p-1})^{-1} P; \quad E(\epsilon_t \epsilon_t') = \Omega = P P'$$

$$\left(D_T^{-1} \sum_{t=1}^T x_t \epsilon_{it} \right) = \begin{pmatrix} T^{-1/2} \sum u_{t-1} \epsilon_{it} \\ T^{-1/2} \sum u_{t-2} \epsilon_{it} \\ \vdots \\ T^{-1/2} \sum u_{t-p+1} \epsilon_{it} \\ T^{-1/2} \sum \epsilon_{it} \\ T^{-1} \sum y_{t-1} \epsilon_{it} \end{pmatrix} \Rightarrow \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \quad (3.3.14)$$

其中

$$h_1 \sim N(0, \sigma_u^2 V); \quad \sigma_u^2 = E(\epsilon_u^2)$$

$$h_2 = \begin{pmatrix} e' P W(1) \\ \Lambda \left\{ \int_0^1 W(r) [dW(r)]' \right\} P' e \end{pmatrix}$$

e_i 为 $n \times n$ 维单位矩阵的第 i 列。

将(3.3.12)至(3.3.14)中的结果代入(3.3.11),可得

$$\begin{aligned} D_T[b_T^i - \beta_i] &= [D_T^{-1} \left(\sum_{t=1}^T x_t x_t' \right) D_T^{-1}]^{-1} [D_T^{-1} \sum_{t=1}^T x_t \epsilon_{it}] \\ &\Rightarrow \begin{bmatrix} V^{-1} h_1 \\ Q^{-1} h_2 \end{bmatrix} \end{aligned} \quad (3.3.15)$$

由于向量 h_1 有多变量的正态分布 $N(0, \sigma_u^2 V)$, 所以 $D_T[b^i - \beta_i]$ 的前 $n(p-1)$ 维的子向量也有正态的极限分布:

$$\begin{aligned} \sqrt{T}[b_{iT} - \beta_{it}] &= \sqrt{T} \begin{bmatrix} \hat{\xi}_{i1} - \xi_{i1} \\ \vdots \\ \hat{\xi}_{i, p-1} - \xi_{i, p-1} \end{bmatrix} \\ &\Rightarrow V^{-1} h_1 \sim N(0, \sigma_u^2 V^{-1}) \end{aligned} \quad (3.3.16)$$

子向量 $\sqrt{T}[b_{iT} - \beta_{it}]$ 的极限分布与另一子向量 $[T^{1/2} \hat{\alpha}_i, T(\hat{\rho}_i - e_i)']'$ 的极限 $Q^{-1} h_2$ 的分布不相关, 这是因为这两个子向量有不同的收敛速度。这表明, 若在下列模型

$$\Delta y_{it} = \alpha_i + \xi'_{i1} \Delta y_{t-1} + \cdots + \xi'_{i, p-1} \Delta y_{t-p+1} + \epsilon_{it} \quad (3.3.17)$$

中, 对参数 $\xi_{i1}, \dots, \xi_{i, p-1}$ 作最小二乘估计, 那么统计量的极限与(3.3.16)一致。模型(3.3.16)和(3.3.17)都可用来对参数 $\xi_{i1}, \dots, \xi_{i, p-1}$ 作估计和假设检验, 所用的统计量有相同的极限。只有当样本量 T 不很大时, 才会出现明显的差异。

由(3.3.14)和(3.3.15), 子向量 $[T^{1/2} \hat{\alpha}_i, T(\hat{\rho}_i - e_i)']'$ 的极限为

中, $\sqrt{T}(\hat{\rho} - \rho)$ 以概率趋于零。这样, 在 $\text{VAR}(P)$ 的 (3.3.1) 表示式

$$y_t = \alpha + \Psi_1 y_{t-1} + \cdots + \Psi_p y_{t-p} + \varepsilon_t$$

中, 参数 Ψ_j ($j = 1, 2, \dots, p$) 的最小二乘估计都有正态的极限分布。虽然单位根系数 ρ 的估计值出现在 $\hat{\Psi}_1$ 中, 由于收敛速度的不同, $\hat{\Psi}_1$ 的极限分布由 $\hat{\xi}_1$ 起主导作用。这样就很难在 (3.3.1) 的 $\text{VAR}(P)$ 中分析单位根的作用。对参数 Ψ_s ($s = 1, 2, \dots, p$) 的假设检验可用一般的 F -统计量, 尽管这时 $\hat{\Psi}_s$ 是非稳定变量 y_{t-1}, \dots, y_{t-p} 的函数。

3.3.3 带常数项的 $\text{VAR}(P)$

以下假设数据的生成过程为带常数项的 $\text{VAR}(P)$:

$$y_t = \xi_1 \Delta y_{t-1} + \xi_2 \Delta y_{t-2} + \cdots + \xi_{p-1} \Delta y_{t-p+1} + \alpha + \rho y_{t-1} + \varepsilon_t \quad (3.3.19)$$

其中参数 $\alpha \neq 0$, 即 α 中至少有一非零的分量。当单位根假设成立时, $\rho = I_n$ 为 $n \times n$ 维单位矩阵, 上式可简化为

$$\Delta y_t = \delta + u_t$$

其中

$$\delta = (I_n - \xi_1 - \cdots - \xi_{p-1})^{-1} \alpha$$

$$u_t = \Psi(L)\varepsilon_t; \Psi(L) = (I_n - \xi_1 L - \xi_2 L^2 - \cdots - \xi_{p-1} L^{p-1})^{-1}$$

特征方程

$$|I_n - \xi_1 z - \xi_2 z^2 - \cdots - \xi_{p-1} z^{p-1}| = 0$$

的根都在单位圆外,所以矩阵 $(I_n - \zeta_1 L - \zeta_2 L^2 - \cdots - \zeta_{p-1} L^{p-1})$ 和矩阵多项式 $(I_n - \zeta_1 L - \zeta_2 L^2 - \cdots - \zeta_{p-1} L^{p-1})$ 的逆存在。上节中讨论的情况相应于 $\delta = 0$ 。

第1章曾指出,在带常数项的单位根过程 $y_t = \alpha + \rho y_{t-1} + \epsilon_t$ 中,若 $\rho = 1$, $\alpha \neq 0$, 最小二乘估计 $\hat{\alpha}_T$ 和 $\hat{\rho}_T$ 都有正态的极限分布,尽管这时 $\{y_t\}$ 是一 $I(1)$ 时间序列。以下讨论向量参数 α 在模型(3.3.19)的参数估计和假设检验中的作用。

将 $\Delta y_t = \delta + u_t$ 等价地写成以下形式:

$$y_t = y_0 + \delta t + u_1 + \cdots + u_t$$

并不失一般性地假设 δ 的第 n 个分量不为零,即 $\delta_n \neq 0$, 若 $\delta_n = 0$, 总能找到一不为零的 δ_i , 将 y_t 重新排列后使得 δ_i 成为新的向量中的第 n 个分量。 δ 的其他分量是否为零并不影响以下讨论的结果。

对向量 y_t 作如下变换:

$$y_{1t}^* = y_{1t} - (\delta_1/\delta_n)y_{nt}$$

$$y_{2t}^* = y_{2t} - (\delta_2/\delta_n)y_{nt}$$

$$\vdots$$

$$y_{n-1,t}^* = y_{n-1,t} - (\delta_{n-1}/\delta_n)y_{nt}$$

$$y_{nt}^* = y_{nt}$$

对任何 $i = 1, 2, \cdots, n-1$, 可得

$$\begin{aligned} y_{it}^* &= y_{it} - (\delta_i/\delta_n)y_{nt} \\ &= [y_{i0} + \delta_{it} + u_{i1} + u_{i2} + \cdots + u_{it}] \\ &\quad - (\delta_i/\delta_n)[y_{n0} + \delta_{nt} + u_{n1} + \cdots + u_{nt}] \end{aligned}$$

足同样可和性条件。 $(n-1)$ 维随机向量 y_t^* 满足定理 3.1 的条件,与定理中的随机向量 ξ_t 有相同的极限(只是维数不同)。给定随机向量的观察值 y_t^* 和 u_t ,由(3.3.19)所得的估计值,与下式中的参数估计值等价:

$$y_t = \xi_1 u_{t-1} + \xi_2 u_{t-2} + \cdots + \xi_{p-1} u_{t-p+1} + \alpha^* + \rho^* y_{t-1}^* + \gamma y_{n,t-1} + \varepsilon_t \quad (3.3.20)$$

其中, ρ^* 为 $n \times (n-1)$ 维系数矩阵; γ 为 n 维系数向量。回归模型(3.3.20)的自变量可分为三类:均值为零的稳定变量 $u_{t-s} = \Delta y_{t-s} - \delta$, $s = 1, 2, \cdots, p-1$; 常数项 α^* ; 不含常数项的 $I(1)$ 变量 y_{t-1}^* 和带常数项的 $I(1)$ 变量 $y_{n,t-1}$ 。因此,以(3.3.20)式分析估计量 $\hat{\xi}_1, \cdots, \hat{\xi}_{p-1}, \alpha^*$ 和 $\hat{\rho}^*$ 的分布,更有利于发现非零常数项的作用。根据(3.3.20),随机向量 y_t 的第 i 个分量 y_{it} 有表示式:

$$y_{it} = \hat{\xi}'_{i1} u_{t-1} + \hat{\xi}'_{i2} u_{t-2} + \cdots + \hat{\xi}'_{i,p-1} u_{t-p+1} + \alpha_i^* + \rho_i^{*'} y_{t-1}^* + \gamma_i y_{n,t-1} + \varepsilon_{it} \quad (3.3.21)$$

其中 $\hat{\xi}'_{i\cdot}$ 和 $\rho_i^{*'}$ 分别为矩阵 $\hat{\xi}_\cdot$ 和 ρ^* 的第 i 行。令

$$x_t' = [u_{t-1}', u_{t-2}', \cdots, u_{t-p+1}', 1, y_{t-1}^{*'}, y_{n,t-1}']'$$

$$D_T = \begin{bmatrix} TI_{(p-1) \times (p-1)} & \mathbf{0} & 0 & \mathbf{0} \\ \mathbf{0}' & T^{-1} & \mathbf{0}' & 0 \\ 0 & \mathbf{0} & TI_{n \times 1} & \mathbf{0} \\ \mathbf{0}' & 0 & \mathbf{0}' & T^{-1/2} \end{bmatrix} \quad (3.3.22)$$

$$\Lambda^* = \Psi^*(1) \cdot P$$

和

$$\beta^* = [\hat{\xi}'_{i1}, \hat{\xi}'_{i2}, \cdots, \hat{\xi}'_{i,p-1}, \alpha_i^*, \rho_i^{*'}]'$$

利用给定的样本 $\mathbf{x}_1^*, \dots, \mathbf{x}_T^*$, 参数 β^* 的最小二乘估计为

$$\mathbf{b}_T^{*'} = \beta^* + \left[\sum_{i=1}^T \mathbf{x}_i^* \mathbf{x}_i^{*'} \right]^{-1} \left[\sum_{i=1}^T \mathbf{x}_i^* \epsilon_{it} \right]$$

以矩阵 D_T 左乘, 可得

$$D_T [\mathbf{b}_T^{*'} - \beta^*] = \left[D_T^{-1} \left(\sum_{i=1}^T \mathbf{x}_i^* \mathbf{x}_i^{*'} \right) D_T^{-1} \right]^{-1} \left[D_T^{-1} \sum_{i=1}^T \mathbf{x}_i^* \epsilon_{it} \right]$$

由于 $D(\epsilon_t) = E(\epsilon_t \epsilon_t') = \Omega = PP'$, 根据定理 3.1, 可得

$$\begin{aligned} & D_T^{-1} \left(\sum_{i=1}^T \mathbf{x}_i^* \mathbf{x}_i^{*'} \right) D_T^{-1} \\ \Rightarrow & \begin{pmatrix} V & \mathbf{0} & 0 & \mathbf{0} \\ \mathbf{0}' & 1 & \left[\int_0^1 \mathbf{W}(r) dr \right]' \Lambda^{*'} & \delta_n/2 \\ 0 & \Lambda^* \int_0^1 \mathbf{W}(r) dr & \Lambda^* \left\{ \int_0^1 \mathbf{W}(r) \mathbf{W}(r)' dr \right\} \Lambda^{*'} & \partial_n \Lambda^* \int_0^1 r \mathbf{W}(r) dr \\ \mathbf{0}' & \delta_n/2 & \delta_n \left[\int_0^1 r \mathbf{W}(r) dr \right]' \Lambda^{*'} & \delta_n^2/3 \end{pmatrix} \end{aligned}$$

其中的子矩阵 V 为

$$V = \begin{pmatrix} \Gamma_0 & \Gamma_{-1} & \cdots & \Gamma_{p-2} \\ \Gamma_{-1} & \Gamma_0 & \cdots & \Gamma_{p-3} \\ \vdots & \vdots & \cdots & \vdots \\ \Gamma_{p-2} & \Gamma_{p-3} & \cdots & \Gamma_0 \end{pmatrix} \quad (3.3.23)$$

同理可得

$$D_T^{-1} \sum_{i=1}^T \mathbf{x}_i^* \epsilon_{it} \Rightarrow \begin{pmatrix} h_1 \\ h_2 \\ h_3 \\ \vdots \\ h_{p+1} \end{pmatrix}$$

$$Q = \begin{pmatrix} 1 & \left[\int_0^1 W(r) dr \right]' \Lambda^{*'} & \delta_n/2 \\ \Lambda^* \int_0^1 W(r) dr & \Lambda^* \left\{ \int_0^1 W(r) W'(r) dr \right\} \Lambda^{*'} & \delta_n \Lambda^* \int_0^1 r W(r) dr \\ \delta_n/2 & \delta_n \left[\int_0^1 r W(r) dr \right]' \Lambda^{*'} & \delta_n^2/3 \end{pmatrix}$$

统计量 $T^{1/2}(\hat{\omega}_i - \omega_i)$ 有正态的极限分布:

$$T^{1/2}(\hat{\omega}_i - \omega_i) \xrightarrow{d} N(0, \sigma_{ii}^2 V^{-1}) \quad (3.3.24)$$

若在带常数项的 $\text{VAR}(P)$ (3.3.19) 中令 $\rho = I_n$, 以 Δy_i 对其他自变量作回归, 所得到的估计量 $\hat{\xi}_1, \dots, \hat{\xi}_{p-1}$ 的极限与 (3.3.24) 有相同的形式。这结果与上节相应的不带常数项的 $\text{VAR}(P)$ 的结果是一致的。

同样也可直接对模型 (3.3.1) 的参数作最小二乘估计, 估计量 $\hat{\Psi}_s$ ($s = 1, \dots, p$) 是参数 $\hat{\omega}_i, \hat{\alpha}_i, \hat{\rho}_i$ 和 $\hat{\gamma}_i$ 的线性组合。由于 $\hat{\rho}_i$ 和 $\hat{\gamma}_i$ 的收敛速度都高于 $T^{-1/2}$, $\hat{\Psi}_s$ 的极限只是 $\hat{\omega}_i$ 和 $\hat{\alpha}_i$ 极限的线性组合, 所以也有正态的极限分布。这结果与上节的讨论也是一致的。因此, 不管产生数据的生成过程是否带有常数项, 模型 (3.3.1) 参数的假设检验都可用一般的 F -检验法, 统计量 F_T 在原假设为真时有 χ^2 极限分布。

3.4 伪回归

当随机变量服从单位根过程, 在用这些变量构造经济计量模型时, 需注意伪回归 (spurious regression) 现象的出现。以下讨论伪回归现象产生的原因和纠正方法。

在线性回归模型

$$y_t = x_t\beta + u_t; \quad t = 1, 2, \dots, T$$

中,若最小二乘估计 $\hat{\beta}$ 的 t -统计量是显著的,从统计学角度我们不拒绝 y_t 和 x_t 之间存在线性关系的假设。但也存在这样的情况,尽管 $\hat{\beta}$ 的 t -统计量呈显著性, y_t 和 x_t 之间根本不存在任何线性关系,这种现象称为伪回归现象。经济学家们早已认识到在经济变量之间存在伪回归现象,但在怎样的条件下会产生伪回归现象,长期以来无统一认识。直到70年代中,哥兰格和纽博尔德(Granger and Newbold, 1974)用蒙的可罗模拟方法表明,当 y_t 和 x_t 都服从单位根过程时,即使它们之间不存在任何线性相关关系,以 y_t 对 x_t 作回归得到的 β 的最小二乘估计 $\hat{\beta}$ 有显著的 t 值。这对认识伪回归现象产生的原因作出了重要的贡献。在他们工作的启发下,菲利普斯(Phillips, 1986)第一次在理论上证明了不相关的单位根变量之间会产生伪回归现象。

考虑回归模型

$$y_t = \alpha + \beta'x_t + u_t$$

其中, x_t 和 β 都为 g 维向量, y_t 和 x_t 都服从单位根过程。随机变量 y_t 和 x_t 之间不存在任何相关关系。比如,设 y_t 是英国累计的失业人数, x_t 是在时间 t 时累计的太阳黑子数。以上模型纯属人为构造。

参数 α 和 β 的最小二乘估计由下式给出:

$$\begin{bmatrix} \alpha \\ \hat{\beta} \end{bmatrix} = \begin{bmatrix} T & \sum x_t' \\ \sum x_t & \sum x_t x_t' \end{bmatrix}^{-1} \begin{bmatrix} \sum y_t \\ \sum x_t y_t \end{bmatrix} \quad (3.4.1)$$

对参数向量 β 的假设 $H_0: R\beta = \gamma$ 作检验,其中 R 为给定的 $m \times g$ 维已知矩阵, γ 为 m 维已知向量。若 $R = I_g$, $\gamma = 0$, H_0 成为 $\beta = 0$ 。由最小二乘估计 $\hat{\alpha}$ 和 $\hat{\beta}$ 构造检验假设 H_0 的 F -统

计量:

$$F_T = \{\mathbf{R}'\hat{\boldsymbol{\beta}} - \gamma\}' \left\{ s_T^2 [\mathbf{0}, \mathbf{R}] \begin{bmatrix} T & \sum \mathbf{x}_t \\ \sum \mathbf{x}_t' & \sum \mathbf{x}_t \mathbf{x}_t' \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{0}' \\ \mathbf{R}' \end{bmatrix} \right\}^{-1} \{\mathbf{R}'\hat{\boldsymbol{\beta}} - \gamma\} / m \quad (3.4.2)$$

其中的 s_T^2 为方差的最小二乘估计:

$$s_T^2 = \frac{1}{T - g - 1} \sum_{t=1}^T \hat{u}_t^2; \quad \hat{u}_t^2 = (y_t - \hat{\alpha} - \hat{\boldsymbol{\beta}}' \mathbf{x}_t)^2$$

以下证明,若 y_t 和 \mathbf{x}_t 都服从单位根过程,尽管它们之间不存在任何相关性, F_T 统计量将随着样本量 T 的增加,以概率发散到无穷大。所以只要样本量足够大, F_T 值一定呈显著性。

定理 3.2: 设 \mathbf{z}_t 为 n 维的 $I(1)$ 随机向量,一阶差分 $\Delta \mathbf{z}_t$ 为 n 维的稳定过程,可表示为

$$\Delta \mathbf{z}_t = \boldsymbol{\Psi}(L) \boldsymbol{\varepsilon}_t = \sum_{i=0}^{\infty} \boldsymbol{\Psi}_i \boldsymbol{\varepsilon}_{t-i}$$

其中, $\{\boldsymbol{\varepsilon}_t\}$ 独立同分布, $E(\boldsymbol{\varepsilon}_t) = \mathbf{0}$, $D(\boldsymbol{\varepsilon}_t) = E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') = \boldsymbol{\Omega} = \mathbf{P} \mathbf{P}'$ 。矩阵列 $\{s \cdot \boldsymbol{\Psi}_i\}_{i=0}^{\infty}$ 的每一元素列绝对可和。令 $g = n - 1$, $\boldsymbol{\Lambda} = \boldsymbol{\Psi}(1) \mathbf{P}$, 将向量 \mathbf{z}_t 分解为 $\mathbf{z}_t = [y_t, \mathbf{x}_t']'$, 并相应地将矩阵 $\boldsymbol{\Lambda} \boldsymbol{\Lambda}'$ 分块:

$$\boldsymbol{\Lambda} \boldsymbol{\Lambda}' = \begin{bmatrix} \Sigma_{11} & \Sigma_{21}' \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

这里, $\boldsymbol{\Lambda} \boldsymbol{\Lambda}'$ 为 $n \times n$ 维对称矩阵, Σ_{11} 为一标量, Σ_{21} 为 g 维向量, Σ_{22} 为 $g \times g$ 维矩阵。 Σ_{11} 和 Σ_{22} 分别为 Δy_t 和 $\Delta \mathbf{x}_t$ 的方差。对矩阵 Σ_{22}^{-1} 作三角分解:

$$\Sigma_{22}^{-1} = \mathbf{L}_{22} \mathbf{L}_{22}'$$

其中 \mathbf{L}_{22} 为下三角矩阵。令 $(\sigma_1^*)^2 = \Sigma_{11} - \Sigma_{21}' \Sigma_{22}^{-1} \Sigma_{21}$, 以 y_t 对 \mathbf{x}_t

作回归,得到最小二乘估计 $\hat{\alpha}$ 和 $\hat{\beta}$ 。以下结果成立:

(a) 最小二乘估计 $\hat{\alpha}$ 和 $\hat{\beta}$ 有极限

$$\begin{bmatrix} T^{-1/2} \hat{\alpha} \\ \hat{\beta} - \Sigma_{22}^{-1} \Sigma_{21} \end{bmatrix} \Rightarrow \begin{bmatrix} \sigma_1^* h_1 \\ \sigma_1^* L_{22} \end{bmatrix}$$

其中

$$\begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} 1 & \int_0^1 W_2^*(r)' dr \\ \int_0^1 W_2^*(r) dr & \int_0^1 W_2^*(r) W_2^*(r)' dr \end{bmatrix}^{-1} \begin{bmatrix} \int_0^1 W_1^*(r) dr \\ \int_0^1 W_2^*(r) W_1^*(r) dr \end{bmatrix} \quad (3.4.3)$$

$W_1^*(r)$ 为单变量的标准维纳过程, $W_2^*(r)$ 为 g 维的标准维纳过程,
 $W_1^*(r)$ 与 $W_2^*(r)$ 互相独立。

(b) y_t 对 x_t 作回归的残差平方和 RSS_T 有极限

$$T^{-2} RSS_T \Rightarrow (\sigma_1^*)^2 \cdot H$$

式中

$$\begin{aligned} H = & \int_0^1 W_1^*(r)^2 dr - \left\{ \left[\int_0^1 W_1^*(r) dr, \int_0^1 W_1^*(r) W_2^*(r)' dr \right] \right. \\ & \times \left. \begin{bmatrix} 1 & \int_0^1 W_2^*(r)' dr \\ \int_0^1 W_2^*(r) dr & \int_0^1 W_2^*(r) W_2^*(r)' dr \end{bmatrix}^{-1} \begin{bmatrix} \int_0^1 W_1^*(r) dr \\ \int_0^1 W_2^*(r) W_1^*(r) dr \end{bmatrix} \right\} \end{aligned} \quad (3.4.4)$$

(c) 对于(3.4.2)中的 F_T ,以下极限成立:

$$T^{-1} F_T \Rightarrow \{ \sigma_1^* R^* h_2 - Y^* \}' \{ (\sigma_1^*)^2 H [0, R^*] \}$$

$$\begin{aligned} & \times \begin{bmatrix} 1 & \int_0^1 \mathbf{W}_2^*(r)' dr \\ \int_0^1 \mathbf{W}_2^*(r) dr & \int_0^1 \mathbf{W}_2^*(r) \mathbf{W}_2^*(r)' dr \end{bmatrix}^{-1} \\ & \times \begin{bmatrix} \mathbf{0}' \\ \mathbf{R}^{*'} \end{bmatrix} \}^{-1} \times \{ \sigma_1^* \mathbf{R}^* \mathbf{h}_2 - \mathbf{Y}^* \} \end{aligned} \quad (3.4.5)$$

这里, $\mathbf{R}^* = \mathbf{R} \cdot \mathbf{L}_{22}$, $\mathbf{Y}^* = \mathbf{Y} - \mathbf{R} \Sigma_{22}^{-1} \Sigma_{21}$ 。

证明: 为便于推导, 先对 $\mathbf{z}_i = [y_i, \mathbf{x}_i']'$ 作线性变换:

$$y_i^* = y_i - \Sigma_{21}' \Sigma_{22}^{-1} \Sigma_{21} \quad (3.4.6)$$

$$\mathbf{x}_i^* = \mathbf{L}_{22}' \mathbf{x}_i$$

不难验证, 经变换后的 y_i^* 和 \mathbf{x}_i^* 是不相关的, 而且 $D(\Delta \mathbf{x}_i^*) = \mathbf{I}_g$ 。构造回归模型

$$y_i^* = \alpha^* + \beta^{*'} \mathbf{x}_i^* + u_i^*$$

可得参数 α^* 和 β^* 的最小二乘估计

$$\begin{bmatrix} \hat{\alpha}^* \\ \hat{\beta}^* \end{bmatrix} = \begin{bmatrix} T & \Sigma \mathbf{x}_i^{*'} \\ \Sigma \mathbf{x}_i^* & \Sigma \mathbf{x}_i^* \mathbf{x}_i^{*'} \end{bmatrix}^{-1} \begin{bmatrix} \Sigma y_i^* \\ \Sigma \mathbf{x}_i^* y_i \end{bmatrix}$$

和估计残差

$$\hat{u}_i^* = y_i^* - \hat{\alpha}^* - \hat{\beta}^{*'} \mathbf{x}_i^*$$

若以 $\hat{\alpha}$ 和 $\hat{\beta}$ 表示 y_i 对 \mathbf{x}_i 作回归的估计值, $\hat{u}_i = y_i - \hat{\alpha} - \hat{\beta}' \mathbf{x}_i$ 为估计残差, 容易验证 (见本章后习题), 残差 \hat{u}_i^* 和 \hat{u}_i 是一致的, 所以

$$\begin{aligned} y_i - \hat{\alpha} - \hat{\beta}' \mathbf{x}_i &= y_i^* - \hat{\alpha}^* - \hat{\beta}^{*'} \mathbf{x}_i^* \\ &= (y_i - \Sigma_{21}' \Sigma_{22}^{-1} \Sigma_{21}) - \hat{\alpha}^* - \hat{\beta}^{*'} \mathbf{L}_{22}' \mathbf{x}_i \end{aligned}$$

$$= \begin{pmatrix} \frac{1}{(\sigma_1^*)^2}(\Sigma_{11} - \Sigma'_{21}\Sigma_{22}^{-1}\Sigma_{21}) & \mathbf{0}' \\ \mathbf{0} & L'_{22}\Sigma_{22}L_{22} \end{pmatrix} \quad (3.4.8)$$

矩阵(3.4.8)是分块对角的, 其对角线上的子矩阵都为零。又因为

$$(\sigma_1^*)^2 = (\Sigma_{11} - \Sigma'_{21}\Sigma_{22}^{-1}\Sigma_{21}); \quad \Sigma_{22}^{-1} = L_{22}L'_{22}$$

所以

$$L' \Lambda \Lambda' L = I_n \quad (3.4.9)$$

为 n 维的单位矩阵。利用(3.4.9)可将一标准维纳过程变换成另一标准维纳过程, 即下式

$$W^*(r) = L' \Lambda W(r)$$

也是 n 维标准维纳过程。

对随机向量作部分和 $\sum_{i=1}^T z_i$, 由定理 3.1 的结论(7)可得

$$T^{-3/2} \sum_{i=1}^T z_i \Rightarrow \Lambda \int_0^1 W(r) dr$$

以 $T^{-3/2}$ 乘向量(3.4.7), 并对 t 作和, 就可得

$$\begin{aligned} T^{-3/2} \sum_{i=1}^T \begin{bmatrix} y_i^* / \sigma_1^* \\ x_i^* \end{bmatrix} &= T^{-3/2} \sum_{i=1}^T L' z_i = L' \left\{ T^{-3/2} \sum_{i=1}^T z_i \right\} \\ &\Rightarrow L' \Lambda \int_0^1 W(r) dr = \int_0^1 W^*(r) dr \end{aligned} \quad (3.4.10)$$

利用定理 3.1 的结论(9), 还可知

$$\begin{aligned}
& \begin{bmatrix} T^{-2}\Sigma(y_i^*)^2/(\sigma_1^*)^2 & T^{-2}\Sigma y_i^* x_i^{*'} / \sigma_1^* \\ T^{-2}\Sigma x_i^* y_i^* / \sigma_1^* & T^{-2}\Sigma x_i^* x_i^{*'} \end{bmatrix} \\
&= L' \left\{ T^{-2} \sum_{i=1}^T z_i z_i' \right\} L \Rightarrow L' \Lambda \left\{ \int_0^1 W(r) W(r)' dr \right\} \Lambda' L \\
&= \int_0^1 W^*(r) W^*(r)' dr \quad (3.4.11)
\end{aligned}$$

有了以上这些结果就可证明定理的结论(a)、(b)和(c)。

以 σ_1^* 除向量 $[\hat{\alpha}^*, \hat{\beta}^{*'}]'$, 并以矩阵

$$\begin{bmatrix} T^{-1/2} & \mathbf{0}' \\ \mathbf{0} & I_s \end{bmatrix}$$

左乘, 可得

$$\begin{aligned}
& \begin{bmatrix} T^{-1/2} & \mathbf{0}' \\ \mathbf{0} & I_s \end{bmatrix} \begin{bmatrix} \hat{\alpha}^* / \sigma_1^* \\ \hat{\beta}^* / \sigma_1^* \end{bmatrix} \\
&= \begin{bmatrix} T^{-1/2} & \mathbf{0}' \\ \mathbf{0} & I_s \end{bmatrix} \begin{bmatrix} T & \Sigma x_i^{*'} \\ \Sigma x_i^* & \Sigma x_i^* x_i^{*'} \end{bmatrix}^{-1} \begin{bmatrix} T^{-3/2} & \mathbf{0}' \\ \mathbf{0} & T^{-2} I_s \end{bmatrix}^{-1} \\
&\quad \times \begin{bmatrix} T^{-3/2} & \mathbf{0}' \\ \mathbf{0} & T^{-2} I_s \end{bmatrix} \begin{bmatrix} \Sigma y_i^* / \sigma_1^* \\ \Sigma x_i^* y_i^* / \sigma_1^* \end{bmatrix}
\end{aligned}$$

也即

$$\begin{bmatrix} T^{-1/2} \hat{\alpha}^* / \sigma_1^* \\ \hat{\beta}^* / \sigma_1^* \end{bmatrix} = \begin{bmatrix} 1 & T^{-3/2} \Sigma x_i^{*'} \\ T^{-3/2} \Sigma x_i^* & T^{-2} \Sigma x_i^* x_i^{*'} \end{bmatrix}^{-1} \begin{bmatrix} T^{-3/2} \Sigma y_i^* / \sigma_1^* \\ T^{-2} \Sigma x_i^* y_i^* / \sigma_1^* \end{bmatrix}$$

由(3.4.10)和(3.4.11)的结果, 可得上式的极限

$$\begin{bmatrix} T^{-1/2} \hat{\alpha}^* / \sigma_1^* \\ \hat{\beta}^* / \sigma_1^* \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & \int_0^1 W_2^*(r)' dr \\ \int_0^1 W_2^*(r) dr & \int_0^1 W^*(r) W^*(r)' dr \end{bmatrix}^{-1}$$

$$\times \begin{bmatrix} \int_0^1 \mathbf{W}_1^*(r) dr \\ \int_0^1 \mathbf{W}_2^*(r) \mathbf{W}_1(r) dr \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \quad (3.4.12)$$

以 $\hat{\alpha}^* = \hat{\alpha}$ 和 $\hat{\beta}^* = L_{22}^{-1} \hat{\beta} - L_{22} \Sigma_{21}$ 代入, 并以矩阵

$$\begin{bmatrix} \sigma_1^* & \mathbf{0}' \\ \mathbf{0} & \sigma_1^* L_{22} \end{bmatrix}$$

左乘, 即可得结论(a)。

为证明结论(b), 先计算 $RSS_T = \sum_{i=1}^T \hat{u}_i^2$ 。

$$\begin{aligned} RSS_T &= \Sigma (y_i - \hat{\alpha} - \hat{\beta}' x_i)^2 = \Sigma (y_i^*)^2 \\ &\quad - \left\{ [\Sigma y_i^*, \Sigma y_i^* x_i^{*'}] \begin{bmatrix} T & \Sigma x_i^{*'} \\ \Sigma x_i^* & \Sigma x_i^* x_i^{*'} \end{bmatrix}^{-1} \begin{bmatrix} \Sigma y_i^* \\ \Sigma y_i^* x_i^* \end{bmatrix} \right\} \\ &= \Sigma (y_i^*)^2 - \left\{ [\Sigma y_i^*, \Sigma y_i^* x_i^{*'}] \begin{bmatrix} T^{1/2} & \mathbf{0}' \\ \mathbf{0} & I_g \end{bmatrix} \right. \\ &\quad \left. \left[\begin{bmatrix} T^{-3/2} & \mathbf{0}' \\ \mathbf{0} & T^{-2} I_g \end{bmatrix} \begin{bmatrix} T & \Sigma x_i^{*'} \\ \Sigma x_i^* & \Sigma x_i^* x_i^{*'} \end{bmatrix} \begin{bmatrix} T^{1/2} & \mathbf{0}' \\ \mathbf{0} & I_g \end{bmatrix} \right]^{-1} \right. \\ &\quad \left. \times \begin{bmatrix} T^{-3/2} & \mathbf{0}' \\ \mathbf{0} & T^{-2} I_g \end{bmatrix} \begin{bmatrix} \Sigma y_i^* \\ \Sigma y_i^* x_i^* \end{bmatrix} \right\} \quad (3.4.13) \end{aligned}$$

以 $(T\sigma_1^*)^2$ 除上式, 并令 $T \rightarrow \infty$, 可得

$$\begin{aligned} &\frac{1}{(\sigma_1^*)^2 T^2} RSS_T \\ &= T^{-2} \Sigma (y_i^* / \sigma_1^*)^2 - \{ [T^{-3/2} \Sigma (y_i^* / \sigma_1^*), T^{-2} \Sigma (y_i^* / \sigma_1^*) x_i^{*'}] \\ &\quad \times \begin{bmatrix} 1 & T^{-3/2} \Sigma x_i^{*'} \\ T^{-1/2} \Sigma x_i^* & T^{-2} \Sigma x_i^* x_i^{*'} \end{bmatrix}^{-1} \begin{bmatrix} T^{-3/2} \Sigma (y_i^* / \sigma_1^*) \\ T^{-2} \Sigma x_i^* y_i^* / \sigma_1^* \end{bmatrix} \} \end{aligned}$$

$$\Rightarrow \int_0^1 W_1^*(r)^2 dr = \left\{ \left[\int_0^1 W_1^*(r) dr, \int_0^1 W_1^*(r) W_2^*(r)' dr \right] \right. \\ \left. \times \begin{bmatrix} 1 & \int_0^1 W_2^*(r)' dr \\ \int_0^1 W_2^*(r) dr & \int_0^1 W_2^*(r) W_2^*(r)' dr \end{bmatrix}^{-1} \begin{bmatrix} \int_0^1 W_1^*(r) dr \\ \int_0^1 W_2^*(r) W_1^*(r) dr \end{bmatrix} \right\}$$

结论(b)得证。

为证明结论(c),对 $y_t = \alpha + \beta' x_t + u_t$ 的假设

$$H_0: R\beta = \gamma$$

作检验。由参数向量 β 和 β^* 的关系,可将假设 H_0 以 β^* 表示,因

$$\begin{aligned} R\beta = \gamma &= R(L_{22}\beta^* + \Sigma_{22}^{-1}\Sigma_{21}) = \gamma \\ &= (RL_{22})\beta^* = (\gamma - R\Sigma_{22}^{-1}\Sigma_{21}) \\ &= R^*\beta^* = \gamma^* \end{aligned}$$

其中的 $R^* = RL_{22}$, $\gamma^* = \gamma - R\Sigma_{22}^{-1}\Sigma_{21}$ 。

另一方面,对 $y_t^* = \alpha^* + \beta^{*'} x_t^* + u_t^*$ 的假设 $H_0: R^*\beta^* = \gamma^*$ 的检验统计量 F_T 与(3.4.2)有一样的形式(见习题)。这样,为证明结论(c),可用检验 $R^*\beta^* = \gamma^*$ 的 F_T 统计量:

$$\begin{aligned} F_T = \{R^* \hat{\beta}^* - \gamma^*\}' &\left\{ s_T^{*-2} [\mathbf{0}, R^*] \begin{bmatrix} T & \Sigma x_t^{*'} \\ \Sigma x_t^* & \Sigma x_t^* x_t^{*'} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{0}' \\ R^{*'} \end{bmatrix} \right\}^{-1} \\ &\times \{R^* \hat{\beta}^* - \gamma^*\} / m \end{aligned}$$

以 T 除上式,可得表示式:

$$T \cdot F_T = \{R^* \hat{\beta}^* - \gamma^*\}' \left\{ T^{-1} s_T^{*-2} [\mathbf{0}, R^*] \right.$$

$$\begin{aligned} & \times \begin{pmatrix} 1 & T^{-3/2} \sum \mathbf{x}_i^* \\ T^{-3/2} \sum \mathbf{x}_i^* & T^{-2} \sum \mathbf{x}_i^* \mathbf{x}_i^{*'} \end{pmatrix}^{-1} \\ & \times \begin{pmatrix} \mathbf{0}' \\ \mathbf{R}^{*'} \end{pmatrix} \Bigg\}^{-1} \{ \mathbf{R}^* \hat{\boldsymbol{\beta}}^* - \boldsymbol{\gamma}^* \} / m \end{aligned}$$

其中, $s_i^{*2} = (T-n)^{-1} \sum_{i=1}^T (\hat{u}_i^*)^2 = (T-n)^{-1} \sum_{i=1}^T \hat{u}_i^2$ 。由 RSS_T 的表示式(3.4.13)和以上(b)的结论,可得

$$\begin{aligned} T^{-1}(s_T^{*2}) &= T^{-1}(s_T^2) = \left(\frac{T}{T-n} \right) T^{-2} RSS_T \\ &\Rightarrow (\sigma_1^*)^2 \mathbf{H} \end{aligned} \quad (3.4.14)$$

矩阵 \mathbf{H} 由(3.4.4)式给出。另一方面,由(3.4.3)的推导,可得

$$\begin{aligned} \begin{pmatrix} 1 & T^{-3/2} \sum \mathbf{x}_i^{*'} \\ T^{-3/2} \sum \mathbf{x}_i^* & T^{-2} \sum \mathbf{x}_i^* \mathbf{x}_i^{*'} \end{pmatrix}^{-1} &\Rightarrow \begin{pmatrix} 1 & \int_0^1 \mathbf{W}_2^*(r)' dr \\ \int_0^1 \mathbf{W}_2^*(r) dr & \int_0^1 \mathbf{W}_2^*(r) \mathbf{W}_2^*(r)' dr \end{pmatrix}^{-1} \\ \hat{\boldsymbol{\beta}}^* &\Rightarrow \sigma_1^* \mathbf{h}_2 \end{aligned}$$

因此可得 $T^{-1}F_T$ 的极限

$$\begin{aligned} T^{-1}F_T &\Rightarrow \{ \sigma_1^* \mathbf{R}^* \mathbf{h}_2 - \boldsymbol{\gamma}^* \}' \left\{ (\sigma_1^*)^2 \mathbf{H}[\mathbf{0}, \mathbf{R}^*] \right. \\ &\quad \times \begin{pmatrix} 1 & \int_0^1 \mathbf{W}_2^*(r)' dr \\ \int_0^1 \mathbf{W}_2^*(r) dr & \int_0^1 \mathbf{W}_2^*(r) \mathbf{W}_2^*(r)' dr \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{0}' \\ \mathbf{R}^{*'} \end{pmatrix} \Bigg\}^{-1} \\ &\quad \{ \sigma_1^* \mathbf{R}^* \mathbf{h}_2 - \boldsymbol{\gamma}^* \} / m \end{aligned}$$

结论(c)得证。定理证毕。

这样,最小二乘估计 $\hat{\alpha}$ 和 $\hat{\beta}$ 都不收敛于它们的真值,而且 $\hat{\alpha}$ 还随着样本量 T 的增加而发散,这也说明为什么 F_T 值是显著的。

3.5 伪回归的纠正方法

以下介绍三种纠正伪回归现象的方法。

第一种方法,是在回归模型中包括自变量和应变量的一阶滞后变量,比如在上节末的例子中包括 y_{t-1} 和 x_{t-1} ,使估计模型成为

$$y_t = \alpha + \beta x_t + \phi y_{t-1} + \delta x_{t-1} + u_t \quad (3.5.1)$$

若这时 $\{y_t\}$ 和 $\{x_t\}$ 是独立的随机游动,那么不难验证参数 β 和 δ 的最小二乘估计 $\hat{\beta}$ 和 $\hat{\delta}$ 将以概率收敛于零,而且以 $T^{1/2}$ 的速度收敛于正态分布。这样,在(3.5.1)中包括滞后变量后,就解决了伪回归问题。

纠正伪回归的第二种方法,是在建立和估计模型之前先作 y_t 和 x_t 的一阶差分。由于 $\{y_t\}$ 和 $\{x_t\}$ 为随机游动过程,一阶差分 $\{\Delta y_t\}$ 和 $\{\Delta x_t\}$ 都为稳定过程。然后建立和估计模型

$$\Delta y_t = \alpha + \beta \Delta x_t + u_t$$

若这时 $\{y_t\}$ 和 $\{x_t\}$ 不相关,参数 α 和 β 的最小二乘估计 $\hat{\alpha}$ 和 $\hat{\beta}$ 以概率趋于零,并以 $T^{1/2}$ 的速度收敛于正态分布,它们的 t_T 和联合检验的 F_T 统计量分别有正态的和 χ^2 的极限分布。

第三种方法,是用所谓的科克伦—奥克特(Cochrane-Orcutt)方法估计模型 $y_t = \alpha + \beta x_t + u_t$ 。这个方法最早由科克伦和奥克特提出的,目的在于纠正由一阶自回归随机干扰 $u_t = \rho u_{t-1} + \epsilon_t$ 引起的估计偏差。它的具体做法如下:若随机变量适

合模型

$$y_t = \alpha + \beta x_t + u_t; \quad u_t = \rho u_{t-1} + \varepsilon_t \quad (3.5.2)$$

以 ρ 乘 y_{t-1} , 可得

$$\rho y_{t-1} = \rho\alpha + \rho\beta x_{t-1} + \rho u_{t-1}$$

再将其从 y_t 中减去, 可得

$$y_t - \rho y_{t-1} = \alpha(1 - \rho) + \beta(x_t - \rho x_{t-1}) + \varepsilon_t \quad (3.5.3)$$

科克伦—奥克特方法利用方程(3.5.2)和(3.5.3)对估计值 $\hat{\alpha}$ 、 $\hat{\beta}$ 和 $\hat{\rho}$ 不断更新。首先, 从(3.5.2)得到最小二乘估计 $\hat{\alpha}$ 和 $\hat{\beta}$, 及估计残差 $\hat{u}_t = y_t - \hat{\alpha} - \hat{\beta}x_t$; 然后, 以残差 \hat{u}_t 对 \hat{u}_{t-1} 作回归, 得到参数 ρ 的最小二乘估计 $\hat{\rho}$; 最后, 以估计值 $\hat{\rho}$ 代入(3.5.3), 从而得到第二阶段的 α 和 β 的估计值。以上过程可循环多次, 以提高估计的精确度。可以证明, 若以科克伦—奥克特方法估计

$$y_t = \alpha + \beta x_t + u_t$$

其中 $\{y_t\}$ 和 $\{x_t\}$ 为不相关的随机游动, 估计量 $\hat{\alpha}$ 和 $\hat{\beta}$ 以概率趋于它们的真值 $\alpha = 0$ 、 $\beta = 0$ 。

3.6 本章小结

本章讨论了多变量的单位根过程, 将许多重要的单变量单位根过程的极限分布推广为向量形式。

向量的单位根过程常可表示成 VAR(P)形式

$$y_t = \alpha + \Psi_1 y_{t-1} + \cdots + \Psi_p y_{t-p} + \varepsilon_t \quad (3.6.1)$$

或其等价形式

$$y_t = \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots + \zeta_{p-1} \Delta y_{t-p+1} + \alpha + \rho y_{t-1} + \varepsilon_t \quad (3.6.2)$$

后者更便于对参数 $\zeta_1, \dots, \zeta_{p-1}$ 和 α, ρ 作估计和假设检验。

由 (3.6.1) 所得到的估计量 $\hat{\Psi}_1, \dots, \hat{\Psi}_p$ 是参数 Ψ_1, \dots, Ψ_p 的一致估计, 它们都以 \sqrt{T} 的速度收敛于正态分布。因此, 对参数 Ψ_1, \dots, Ψ_p 的假设检验可用一般的 F_T 统计量, 在原假设 $H_0: R\beta = \gamma$ 成立时, F_T 有 χ^2 极限分布。

本章还讨论了伪回归现象, 从理论上证明了以两个或多个相互独立的时间序列构造回归模型, 它们 (y_t 和 x_t) 之间会产生显著的 F_T 值, 尽管它们真正的相关系数等于零。最后, 还介绍了三种纠正伪回归的方法, 这些方法使最小二乘估计成为一致的估计。

习 题

1. 设 $u_t = \sum_{k=0}^{\infty} \Psi_k \varepsilon_{t-k}$, 为 n 维随机向量, 其中 $\{\varepsilon_t\}$ 独立同分布, $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = E(\varepsilon_t \varepsilon_t') = \Omega = PP'$, 令

$$\Gamma_s = E(u_t u_{t-s}') = \sum_{k=0}^{\infty} \Psi_{s+k} \Omega \Psi_k', \quad s = 0, 1, 2, \dots$$

$$\Lambda = \Psi(1)P = (\Psi_0 + \Psi_1 + \dots)P$$

证明: $\Lambda \Lambda' = \Psi(1)PP' \Psi'(1) = \Psi(1)\Omega \Psi'(1)$

$$= \Gamma_0 + \sum_{k=1}^{\infty} (\Gamma_k + \Gamma_k')$$

2. 对上题中的 u_t , 证明:

$$\frac{1}{T} \sum_{t=1}^T u_t u_{t-s}' \xrightarrow{P} \Gamma_s; \quad s = 0, 1, 2, \dots$$

3. 令 $\xi_{t-1} = u_1 + u_2 + \cdots + u_{t-1}$, $X_T(r) = \frac{1}{T} \sum_{i=1}^{[Tr]} u_i$, 根据随机积分的定义验证:

$$(a) T^{-2} \sum_{i=1}^T \xi_{i-1} \xi'_{i-1} = \int_0^1 [\sqrt{T} X_T(r)] [\sqrt{T} X_T(r)]' dr$$

$$(b) T^{-2} \sum_{i=1}^T \xi_{i-1} \xi'_{i-1} \Rightarrow \int_0^1 W(r) W(r)' dr$$

4. 考虑回归模型

$$y_t = \alpha + \beta' x_t + u_t \quad (*)$$

其中 $\{y_t\}$ 和 $\{x_t\}$ 服从单位根过程, 并有

$$\begin{pmatrix} \Delta y_t \\ \Delta x_t \end{pmatrix} \sim N \left(\mathbf{0}, \begin{pmatrix} \Sigma_{11} & \Sigma'_{21} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \right)$$

对 y_t 和 x_t 作线性变换:

$$y_t^* = y_t - \Sigma'_{21} \Sigma_{22}^{-1} x_t$$

$$x_t^* = L'_{22} x_t$$

其中, L_{22} 为下三角矩阵, $\Sigma_{22}^{-1} = L_{22} L'_{22}$ 。证明: 在 $(*)$ 中检验假设 $H_0: R\beta = \gamma$ 的 F_T 统计量, 与在模型

$$y_t^* = \alpha^* + \beta^{*'} x_t^* + u_t^*$$

中检验假设 $H_0: R^* \beta^* = \gamma^*$ 的 F_T 统计量是等同的。这里, $R^* = RL_{22}$, $\gamma^* = \gamma - R\Sigma_{22}^{-1}\Sigma_{21}$ 。

第4章 同积过程的性质和表示形式

4.1 简介

从本章开始,我们将讨论一种特殊的向量单位根过程——同积过程^①(co-integrated process)。对这一随机过程的研究不仅在经济计量学中开创了新的理论分支,也为实际研究人员在处理复杂的宏观经济问题时提供了有力的工具。本章主要讨论同积过程的一些重要的性质和几种常用的表示形式,对同积过程的参数估计和假设检验将在以后几章中讨论。

设 $\{y_t, t = 1, 2, \dots\}$ 为一 n 维的向量单位根过程,它的每一个分量序列 $\{y_{it}\} (i = 1, 2, \dots, n)$ 为一单变量单位根过程, $y_{it} \sim I(1)$ 。如果这时存在一非零的 n 维向量 α ,使得 y_t 的线性组合 $\alpha'y_t$ 成为一稳定过程,即 $\alpha'y_t \sim I(0)$,则称随机向量 y_t 是同积的, α 为其同积向量。以下用一个二维的随机向量说明同积的意义。考虑二维系统:

$$y_{1t} = \gamma y_{2t} + u_{1t}$$

① co-integration 似无统一的中文译名,秦朵首先建议以“协整过程”译之。本书采用译名“同积过程”,因其似与译名“积分过程”(integrated process)较为一致。

$$y_{2t} = y_{2,t-1} + u_{2t} \quad (4.1.1)$$

其中的 $\{u_{1t}\}$ 和 $\{u_{2t}\}$ 为独立的白噪声过程。系统(4.1.1)的第二个方程表明 y_{2t} 为一随机游动过程,第一个方程 y_{1t} 也为一个 $I(1)$ 过程。第一个方程还表明 $\{y_{1t}\}$ 和 $\{y_{2t}\}$ 是同积的,因为 $\alpha'y_t \sim I(0)$,这里 $\alpha = [1, -\gamma]'$, $y_t = [y_{1t}, y_{2t}]'$, α 为随机向量 y_t 的同积向量。

若向量单位根过程 y_t 的维数 $n > 2$,则可能存在多个线性独立的 n 维常数向量 α_s , $s = 1, 2, \dots, k$,使得 $\alpha_s'y_t \sim I(0)$ 。将 k 个同积向量列在 $n \times k$ 维矩阵 A 中:

$$A = [\alpha_1, \dots, \alpha_k]$$

由于向量 α_s ($s = 1, 2, \dots, k$) 线性独立,矩阵 A 的秩为 k ,记为 $k = rk(A)$, k 也称为 y_t 的同积秩(co-integrating rank)。

同积理论为研究宏观经济规律提供了有力的理论工具和数学模型。长期以来人们发现许多经济时间序列,特别是宏观经济时间序列,如国民收入(y_t)和总体消费(c_t)等都呈现明显的非稳定的单位根过程的特征。为了研究这些时间序列之间的关系,如 y_t 和 c_t 之间的关系,需先对它们取一阶差分,然后构造差分后的变量 Δy_t 和 Δx_t 之间的关系。因为若直接对变量 y_t 和 c_t 之间的关系作分析,则可能产生伪回归现象,从而导出不正确的结论。但是只对一阶差分 Δy_t 和 Δc_t 之间的关系作分析,往往又不是经济学家的目标,因为这最多只能揭示收入增长和消费增长之间的关系,而不是收入和消费之间的关系。同积理论使我们能直接对 $I(1)$ 变量 y_t 和 x_t 在水平上的关系作分析。

在过去的十几年中,同积理论在经济学和经济计量学的各个主要的分支中都有应用,其中的一个主要原因是“同积”概念与宏观经济学的“均衡”(equilibrium)概念有本质上的联系。两个 $I(1)$ 时间序列之间的同积关系往往作为它们之间存在长期

的均衡关系的证据。

4.2 同积过程的主要特征

仍以二维的同积系统(4.1.1)为例,将其中的 y_{1t} 作一阶差分,得

$$\Delta y_{1t} = \gamma \Delta y_{2t} + \Delta u_{1t} = (\gamma u_{2t} + u_{1t} - u_{1,t-1})$$

令 $\epsilon_{1t} = \gamma u_{2t} + u_{1t}$, 可将 Δy_{1t} 和 Δy_{2t} 写成移动平均形式:

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} 1-L & \gamma L \\ 0 & L \end{bmatrix} \begin{bmatrix} \epsilon_{1t} \\ u_{2t} \end{bmatrix} = \Psi(L) \epsilon_t \quad (4.2.1)$$

滞后多项式 $\Psi(L)$ 的特征方程 $|\Psi(z)| = 0$ 有一单位根, 因为

$$|\Psi(1)| = \begin{vmatrix} 1-1 & \gamma \\ 0 & 1 \end{vmatrix} = 0$$

这表明, 移动平均系统 $\Delta y_t = [\Delta y_{1t}, \Delta y_{2t}]'$ 是不可逆的, 即不存在滞后多项式 $\Psi^{-1}(L)$, 使得 $\Psi^{-1}(L) \Delta y_t$ 有有限阶自回归表示形式。一般地说, 若向量单位根过程 y_t 同积, 尽管 y_t 的一阶差分 Δy_t 是稳定的, 用有限阶的 VAR(p) 过程

$$\Delta y_t + \alpha_1 \Delta y_{t-1} + \cdots + \alpha_p \Delta y_{t-p} = \epsilon_t$$

(这里 p 为任一有限正整数) 来描述 Δy_t 是不合理的。

进一步对 $\Delta y_t = \Psi(L) \epsilon_t$ 作 BN 分解, 并对 t 作部分和, 可得

$$\sum_{i=1}^T \Delta y_i = \Psi(1) \sum_{i=1}^T \epsilon_i + \eta_T - \eta_0$$

计算统计量 $\left(\frac{1}{T} \sum_{i=1}^T \Delta y_i \right)$ 的方差, 并令 $T \rightarrow \infty$, 得

$$Q = \lim_{T \rightarrow \infty} D \left(\frac{1}{T} \sum_{i=1}^T \Delta y_i \right) = \lim_{T \rightarrow \infty} \Psi(1) \left(\frac{1}{T} \sum_{i=1}^T \epsilon_i \epsilon_i' \right) \Psi'(1)$$

$$= \Psi(1)\Omega\Psi'(1)$$

其中, $\Omega = D(\epsilon_t)$ 。矩阵 Q 称为随机向量 Δy_t 的长期方差矩阵, 它等于 $2\pi f(0)$, 这里 $f(0)$ 为 Δy_t 的谱密度函数在零频率的值。由于 $|\Psi(1)| = 0$, 矩阵 Q 也是退化的:

$$|Q| = |\Psi(1)||\Omega||\Psi'(1)| = 0$$

一阶差分 Δy_t 有退化的长期方差矩阵, 这是同积过程 y_t 的最本质的特征之一。以下对此作更一般的讨论。

设 $y_t \sim I(1)$ 为一 n 维随机向量, 则 Δy_t 为一稳定过程。令 $E(\Delta y_t) = \delta$, $u_t = \Delta y_t - \delta$, 并将 u_t 写成移动平均形式:

$$u_t = \Delta y_t - \delta = \epsilon_t + \Psi_1 \epsilon_{t-1} + \cdots = \Psi(L)\epsilon_t$$

式中的 $\Psi_s (s = 1, 2, \cdots)$ 为 $n \times n$ 维矩阵, $\{\epsilon_t\}$ 独立同分布, $E(\epsilon_t) = 0$, $D(\epsilon_t) = E(\epsilon_t \epsilon_t') = \Omega$ 。

若 $\{y_t\}$ 为同积过程, 有同积秩 k , 则存在 k 个线性独立的 n 维向量 $\alpha_1, \cdots, \alpha_k$, 使得 $\alpha' y_t \sim I(0)$ 。令

$$A = [\alpha_1, \cdots, \alpha_k] \quad (4.2.2)$$

并将 y_t 表示成

$$y_t = y_{t-1} + \delta + u_t = y_0 + \delta t + \sum_{j=1}^t u_j$$

对 $\sum_{j=1}^t u_j$ 作 BN 分解, 得

$$y_t = y_0 + \delta t + \Psi(1) \sum_{i=1}^t \epsilon_i + \eta_t - \eta_0 \quad (4.2.3)$$

以同积向量矩阵 A' 左乘上式。因为 A 包含所有 y_t 的同积向量, 因此 $A' y_t \sim I(0)$, 从而有

$$A' y_t = A' (y_0 - \eta_0) + A' \delta t + A' \Psi(1) \sum_{j=1}^t \varepsilon_j \sim I(0)$$

上式中的第二项和第三项都是非稳定的, 上式成为 $I(0)$ 过程的充分和必要条件为

$$A' \delta = 0; \quad A' \Psi(1) = 0 \quad (4.2.4)$$

这说明矩阵 $\Psi(1)$ 是退化的, 由此 Δy_t 的长期方差矩阵也是退化的, 因为

$$\begin{aligned} Q &= \lim_{T \rightarrow \infty} \left| D \left[\frac{1}{T} \sum_{i=1}^T \Delta y_i \right] \right| = | \Psi(1) \Omega \Psi'(1) | \\ &= | \Psi(1) | | \Omega | | \Psi'(1) | = 0 \end{aligned}$$

$\Psi(1)$ 的奇异性说明移动平均过程

$$\Delta y_t = \delta + \Psi(L) \varepsilon_t$$

是不可逆的, Δy_t 不存在有限阶的 $\text{VAR}(P)$ 的表示形式。

向量单位根过程 y_t 本身可由 $\text{VAR}(P)$ 形式表示:

$$y_t = \alpha + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \cdots + \Phi_p y_{t-p} + \varepsilon_t$$

或者

$$\Phi(L) y_t = \alpha + \varepsilon_t$$

其中, $\{\varepsilon_t\}$ 独立同分布, $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = E(\varepsilon_t \varepsilon_t') = \Omega = PP'$, 滞后多项式 $\Phi(L)$ 为

$$\Phi(L) = I_n - \Phi_1 L - \Phi_2 L^2 - \cdots - \Phi_p L^p$$

另一方面, 一阶差分过程 $\{\Delta y_t\}$ 有移动平均表示形式:

$$\Delta y_t = (1 - L) y_t = \delta + \Psi(L) \varepsilon_t \quad (4.2.5)$$

其中

$$\Psi(L) = I_n + \Psi_1 L + \Psi_2 L^2 + \dots$$

以滞后多项式 $\Phi(L)$ 左乘(4.2.5),得

$$(1 - L) \Phi(L) y_t = \Phi(1) \delta + \Phi(L) \Psi(L) \varepsilon_t \quad (4.2.6)$$

并以 $\Phi(L) y_t = \alpha + \varepsilon_t$ 代入,可得

$$(1 - L)(\alpha + \varepsilon_t) = (1 - L)\varepsilon_t = \Phi(1)\delta + \Phi(L)\Psi(L)\varepsilon_t \quad (4.2.7)$$

等式(4.2.7)成立的必要条件为

$$\Phi(1)\delta = 0 \quad (4.2.8)$$

而且滞后多项式 $(1 - L)I_n$ 与 $\Phi(L)\Psi(L)$ 相等。这意味着对任何实数 z ,下式

$$(1 - z)I_n = \Phi(z)\Psi(z)$$

成立。特别地,取 $z = 1$,则有

$$\Phi(1)\Psi(1) = 0 \quad (4.2.9)$$

以 π'_i 表示矩阵 $\Phi(1)$ 的第 i 行,对任何 $i = 1, 2, \dots, n$,由(4.2.8)和(4.2.9),有

$$\pi'_i \delta = 0; \quad \pi'_i \Psi(1) = 0$$

与(4.2.4)比较,可知 π_i 也为 $\{y_t\}$ 的同积向量,可由 $\alpha_1, \dots, \alpha_k$ 的线性组合表示成:

$$\pi_i = A b_i = [\alpha_1, \dots, \alpha_k] b_i$$

或者

$$\pi'_i = b'_i A'; \quad i = 1, 2, \dots, n$$

这意味着矩阵 $\Phi(1)$ 可以分解为

设 y_t 为 n 维 $I(1)$ 随机向量, 它的一阶差分 Δy_t 有移动平均表示形式:

$$\Delta y_t = \delta + \Psi(L)\epsilon_t$$

其中, $\{\epsilon_t\}$ 为独立同分布, $E(\epsilon_t) = 0$, $D(\epsilon_t) = \Omega$, $\{\delta + \Psi(L)\epsilon_t\}_{t=0}^{\infty}$ 中的所有元素列绝对可和。若 y_t 中有 k 个线性独立的同积关系, 则存在 $n \times k$ 维矩阵 A , 其秩为 k , 使得

$$z_t = A'y_t$$

为一 k 维的 $I(0)$ 过程, 这里矩阵 A 满足条件

$$A'\Psi(1) = 0$$

若将 y_t 表示成 (4.3.3) 中的 $\text{VAR}(P)$ 形式:

$$y_t = \alpha + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \cdots + \Phi_p y_{t-p} + \epsilon_t$$

则存在 $n \times k$ 维矩阵 B , 使得

$$\Phi(1) = BA'$$

从而可将同积过程 y_t 表示成误差修正形式

$$\Delta y_t = \zeta_1 \Delta y_{t-1} + \cdots + \zeta_{p-1} \Delta y_{t-p+1} + (\alpha - BA'y_{t-1}) + \epsilon_t$$

4.3.2 三角表示形式

以 $n \times k$ 维矩阵 A 表示 y_t 的 k 个线性独立的同积向量, A 的 k 个列向量 $\alpha_1, \dots, \alpha_k$ 可作为一组基, 构成 n 维线性空间的一个 k 维子空间, 称为 y_t 的同积空间。若列向量 α_1 的第一个元素不为零, 将其规范为 1; 若它为零, 则总可将 y_t 重新排列, 使得列向量 α_1 的第一个分量不为零, 这种重新排列并不影响以下讨论的普遍性。因此矩阵 A' 有形式

$$A' = \begin{bmatrix} \alpha'_1 \\ \vdots \\ \alpha'_k \end{bmatrix} = \begin{bmatrix} 1 & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \alpha_{k1} & \alpha_{k2} & \alpha_{k3} & \cdots & \alpha_{kn} \end{bmatrix}$$

将 A' 的第一行中的每一元素乘 α_{21} , 并将乘得的结果从相应的第二行的元素中减去。对 A' 中的其他各行作相同的处理, 可得

$$A'_1 = \begin{bmatrix} 1 & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1n} \\ 0 & \alpha_{22}^* & \alpha_{23}^* & \cdots & \alpha_{2n}^* \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \alpha_{k2}^* & \alpha_{k3}^* & \cdots & \alpha_{kn}^* \end{bmatrix}$$

同理可设 α_{22}^* 不为零, 并将其除遍第二行, 再以上述方法将第二列中除了 α_{22}^* 以外的其他元素都变为零。对其他各行、列都作同样的处理, 最后可得矩阵 \bar{A}' 为

$$\begin{aligned} \bar{A}' &= \begin{bmatrix} 1 & 0 & \cdots & 0 & -\gamma_{1,k-1} & \cdots & -\gamma_{1n} \\ 0 & 1 & \cdots & 0 & -\gamma_{2,k+1} & \cdots & -\gamma_{2n} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -\gamma_{k,k+1} & \cdots & -\gamma_{kn} \end{bmatrix} \\ &= [I_k, -\Gamma'] \end{aligned} \quad (4.3.6)$$

这里, Γ' 为 $k \times (n-k)$ 维矩阵。从矩阵 A' 到 \bar{A}' 的线性变换是非异的, 因此 \bar{A}' 中 k 个行向量仍为线性独立的, $\{\bar{A}' y_i\}$ 为 k 维的稳定过程。以 μ_i 表示 $(\bar{A}' y_i)$ 的期望值 $E(\bar{A}' y_i)$, 可将 $\bar{A}' y_i$ 表示成

$$\bar{A}' y_i = \mu_i + v_i \quad (4.3.7)$$

其中 v_i 为均值为零的稳定过程。将 (4.3.6) 代入 (4.3.7), 可得

$$\begin{bmatrix} I_k & -\Gamma' \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \mu_1 + v_t$$

其中 y_{1t} 和 y_{2t} 分别为 k 维和 $(n-k)$ 维子向量, 从而有

$$y_{1t} = \mu_1 + \Gamma' y_{2t} + v_t \quad (4.3.8)$$

其中 y_{2t} 有表示式

$$y_{2t} = y_{2,t-1} + \delta_2 + u_{2t} \quad (4.3.9)$$

(4.3.8) 和 (4.3.9) 可由联立方程形式表示:

$$\begin{bmatrix} I_k & -\Gamma' \\ \mathbf{0} & I_{(n-k)} \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \delta_2 \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I_{(n-k)} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} v_t \\ u_{2t} \end{bmatrix} \quad (4.3.10)$$

上式左侧的矩阵为一 $n \times n$ 维上三角矩阵, 因此称 (4.3.10) 为同积过程 $\{y_t\}$ 的三角表示形式 (triangular representation)。

4.3.3 同趋势表示形式

最后介绍同积过程的同趋势表示形式 (common trend representation)。若 y_t 为一 n 维的 $I(1)$ 随机向量, 其中有 k 个独立的同积向量, 那么 y_t 就有 (4.3.10) 的三角表示形式。若再进一步假设 (4.3.10) 中的随机向量 $[v_t', u_{2t}']'$ 有以下的移动平均形式:

$$\begin{bmatrix} v_t \\ u_{2t} \end{bmatrix} = \sum_{s=0}^{\infty} \begin{bmatrix} H_s \\ J_s \end{bmatrix} \varepsilon_{t-s} \quad (4.3.11)$$

其中的 H_s 和 J_s 分别为 $k \times n$ 维和 $(n-k) \times n$ 维矩阵; $\{\varepsilon_t\}$ 独立同分布, $E(\varepsilon_t) = \mathbf{0}$, $D(\varepsilon_t) = E(\varepsilon_t \varepsilon_t') = \Omega$ 。矩阵 $\{s \cdot H_s\}_{s=0}^{\infty}$ 和

$\{s \cdot J_s\}_{s=0}^{\infty}$ 中的所有元素列都绝对可和。对 $u_{2t} = \sum_{s=0}^{\infty} J_s \varepsilon_{t-s}$ 作 BN 分解, 则可将 y_{2t} 写成下列形式:

$$\begin{aligned} y_{2t} &= y_{2,t-1} + \delta_2 + u_{2t} = y_{20} + \delta_2 t + \sum_{s=1}^t u_{2s} \\ &= y_{20} + \delta_2 t + J(1) \sum_{s=1}^t \varepsilon_s + \eta_{2t} - \eta_{20} \end{aligned}$$

其中, $J(1) = J_0 + J_1 + \dots$ 。令 $\mu_2 = y_{20} - \eta_{20}$, $\xi_{2t} = J(1) \sum_{s=1}^t \varepsilon_s$, 可将上式写成

$$y_{2t} = \mu_2 + \delta_2 t + \xi_{2t} + \eta_{2t} \quad (4.3.12)$$

以此代入(4.3.8)式, 得

$$\begin{aligned} y_{1t} &= \mu_1 + \Gamma' y_{2t} + v_t = \mu_1 + \Gamma' (\mu_2 + \delta_2 t + \xi_{2t} + \eta_{2t}) + v_t \\ &= \bar{\mu}_1 + \Gamma' (\delta_2 t + \xi_{2t}) + \bar{\eta}_{1t} \end{aligned} \quad (4.3.13)$$

这里, $\bar{\mu}_1 = \mu_1 + \Gamma' \mu_2$, $\bar{\eta}_{1t} = v_t + \Gamma' \eta_{2t}$ 。将关系式(4.3.12)和(4.3.13)写成联立方程的形式:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \left\{ \begin{bmatrix} \bar{\mu}_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \bar{\eta}_{1t} \\ \eta_{2t} \end{bmatrix} \right\} + \left\{ \begin{bmatrix} \Gamma' \\ I_{(n-k)} \end{bmatrix} \begin{bmatrix} \delta_2 t + \xi_{2t} \end{bmatrix} \right\} \quad (4.3.14)$$

(4.3.14)称为同积过程 y_t 的同趋势表示形式, 其中右侧第一个括弧中的两个向量都是 $I(0)$ 变量, 第二个括弧中的 $(\delta_2 t + \xi_{2t})$ 是 $(n-k)$ 维向量, 是非随机的时间趋势 $\delta_2 t$ 与向量随机游动 ξ_{2t} 的和。系统中的 y_{1t} 和 y_{2t} 都由 $(\delta_2 t + \xi_{2t})$ 生成, 只是有不同的系数矩阵 Γ' 和 $I_{(n-k)}$ 。同趋势表示形式着眼于系统 y_t 中的独立的单位根变量, 若 y_t 中有 k 个同积关系, 那么真正驱使系统的是

的 k 维子向量。证明:若 x_{1t} 和 y_{1t} 同积,即存在实向量 α_1 和 α_2 ,使得 $\alpha_1'x_{1t} + \alpha_2'y_{1t} \sim I(0)$, 那么 x_t 和 y_t 也同积。

2. 若上题中的 x_t 和 y_t 同积,那么子向量 x_{1t} 和 y_{1t} 是否也同积?

3. 在系统

$$\begin{aligned} y_t &= \beta x_t + u_t \\ x_t &= x_{t-1} + v_t \end{aligned} \quad (*)$$

中, $[u_t, v_t]'$ ($t = 1, \dots, T$) 独立同分布,且

$$\begin{bmatrix} u_t \\ v_t \end{bmatrix} \sim N\left(0, \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix}\right)$$

(a) 令 $z_t = [y_t, x_t]'$, $\epsilon_t = [u_t, v_t]'$, 试将以上模型表示成 VAR(1) 形式

$$z_t = \Phi z_{t-1} + \epsilon_t \quad (**)$$

并解释为何直接估计矩阵 Φ 是不恰当的。

(b) 由 (**) 进一步说明:将 Δz_t 表示成 VAR(P) 形式是不恰当的。

(c) 将模型 (**) 表示成误差纠正形式

$$\begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} = \pi \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \epsilon_t$$

并讨论矩阵 π 的结构。

第 5 章 同积过程的参数估计和假设检验

——最小二乘方法

同积过程 $\{y_t\}$ 的同积向量往往是未知的,需由观察到的样本值估计得到。本章讨论对同积向量的最小二乘估计、估计量的极限分布和对同积向量的假设检验。

5.1 同积向量的最小二乘估计

若 n 维随机向量同积,并有同积向量 a ,那么 a 可由最小二乘法一致地估计得到。为说明最小二乘方法的合理性,作以下考虑。 a 为 y_t 的同积向量,因此 $z_t = a' y_t$ 为一单变量的 $I(0)$ 过程。根据大数定律,就有

$$T^{-1} \sum_{t=1}^T z_t^2 = T^{-1} \sum_{t=1}^T (a' y_t)^2 \xrightarrow{P} E(z_t^2) < \infty$$

但是,若 a 不是 y_t 的同积向量,那么 $z_t = a' y_t$ 仍为 $I(1)$ 变量,根据定理 3.1 的结论(9),可得

$$\begin{aligned} T^{-2} \sum_{t=1}^T z_t^2 &= T^{-2} \sum_{t=1}^T (a' y_t)^2 = a' \left\{ T^{-2} \sum_{t=1}^T y_t y_t' \right\} a \\ &\Rightarrow a' \Lambda \left\{ \int_0^1 W(r) W(r)' dr \right\} \Lambda' a \end{aligned} \quad (5.1.1)$$

其中 Λ 在定理 3.1 中定义, 它决定于 Δy_t 的自相关系数矩阵。

(5.1.1) 式是一正定的二次型, 从而有

$$T^{-1} \sum_{t=1}^T z_t^2 = T \left(T^{-2} \sum_{t=1}^T z_t^2 \right) \xrightarrow{P} +\infty$$

因此, 只要 y_t 是同积的, 并有同积向量 a , 以下最优问题

$$\min_{\hat{a} \in R^n} \left\{ \frac{1}{T} \sum_{t=1}^T (a' y_t)^2 \right\} \quad (5.1.2)$$

的解 \hat{a} , 即 a 的最小二乘估计, 是同积向量的一致估计。

不失一般性, 设同积向量 a 有以下形式:

$$a = \begin{bmatrix} 1 \\ -\gamma_2 \\ -\gamma_3 \\ \vdots \\ -\gamma_n \end{bmatrix}$$

目标函数(5.1.2)成为:

$$T^{-1} \sum_{t=1}^T (a' y_t)^2 = T^{-1} \sum_{t=1}^T (y_{1t} - \gamma_2 y_{2t} - \cdots - \gamma_n y_{nt})^2$$

它的最小值由 y_{1t} 对 $y_{it} (i = 2, \cdots, n)$ 作回归取得。

同积向量的最小估计通常在同积过程的三角表示形式中进行。同积过程 $\{y_t\}$ 的三角表示形式为

$$\begin{aligned} y_{1t} &= \alpha + \gamma' y_{2t} + u_{1t} \\ \Delta y_{2t} &= u_{2t} \end{aligned} \quad (5.1.3)$$

这里, y_{1t} 为单变量随机变量, y_{2t} 为 $(n-1)$ 维随机向量。参数 α 和 γ 的最小二乘估计及它们的极限由定理 5.1 给出。

定理 5.1: 设在同积过程的三角表示形式中, y_{1t} 为一单变量随机变量, y_{2t} 为 $(n-1)$ 维随机向量, n 维随机向量 $[u_{1t}, u'_{2t}]'$ 有移动平均表示形式

$$\begin{bmatrix} u_{1t} \\ u'_{2t} \end{bmatrix} = \Psi(L)\varepsilon_t$$

其中, $\{\varepsilon_t\}$ 为独立同分布, $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = E(\varepsilon_t \varepsilon'_t) = \Omega = PP'$;

$\Psi(L)$ 为无穷阶滞后多项式, 可表示为 $\Psi(L) = \sum_{s=0}^{\infty} \Psi_s L^s$, Ψ_s 为

$n \times n$ 维矩阵, 矩阵列 $\{\Psi_s\}_{s=0}^{\infty}$ 中的每一元素列绝对可和, 矩

阵 $\Psi(1) = \sum_{s=0}^{\infty} \Psi_s$ 为满秩的 $n \times n$ 维矩阵。

若同积过程 $\{y_t\}$ 只有一个独立的同积向量 $[1, -\gamma']'$, 参数 α 和 γ 的最小二乘估计由下式给出:

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} T & \sum y'_{2t} \\ \sum y_{2t} & \sum y_{2t} y'_{2t} \end{bmatrix}^{-1} \begin{bmatrix} \sum y_{1t} \\ \sum y_{2t} y_{1t} \end{bmatrix}$$

将矩阵 $\Psi(1) \cdot P$ 分解为

$$\Psi(1) \cdot P = \begin{bmatrix} \lambda'_1 \\ \Lambda_2 \end{bmatrix}$$

这里, λ_1 为 n 维向量, Λ_2 为 $(n-1) \times n$ 维矩阵。分别以 $T^{1/2}$ 和 T 乘 $(\hat{\alpha} - \alpha)$ 和 $(\hat{\gamma} - \gamma)$, 可得

$$\begin{bmatrix} T^{1/2}(\hat{\alpha} - \alpha) \\ T(\hat{\gamma} - \gamma) \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & \left\{ \int_0^1 W(r)' dr \right\} \Lambda'_2 \\ \Lambda_2 \int_0^1 W(r) dr & \Lambda_2 \left\{ \int_0^1 W(r) W(r)' dr \right\} \Lambda'_2 \end{bmatrix}^{-1} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \quad (5.1.4)$$

这里, $W(r)$ 为 n 维标准维纳过程, h_1 和 h_2 分别为

$$h_1 = \lambda_1' W(1)$$

$$h_2 = \Lambda_2 \left\{ \int_0^1 W(r) [dW(r)]' \right\} \lambda_1 + \sum_{s=0}^{\infty} E(u_{2t} \cdot u_{1, t+s})$$

证明: 令 $\bar{y}_{1t} = \sum_{s=1}^t u_{1s}$, $t = 1, 2, \dots, T$; $\bar{y}_{10} = 0$, 由 \bar{y}_{1t} 和 y_{2t} 组成的向量可表示为

$$\begin{bmatrix} \bar{y}_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 0 \\ y_{20} \end{bmatrix} + \xi_t$$

这里

$$\xi_t = \sum_{s=1}^t \begin{bmatrix} u_{1s} \\ u_{2s} \end{bmatrix}$$

由定理 1.9 的结论(5), 可得以下结果:

$$\begin{aligned} & T^{-1} \sum_{t=1}^T \begin{bmatrix} \bar{y}_{1, t-1} \\ y_{2, t-1} \end{bmatrix} [u_{1t}, u'_{2t}] \\ & \Rightarrow \Psi(1) \cdot P \left\{ \int_0^1 W(r) [dW(r)]' \right\} P' \Psi(1)' + \sum_{s=1}^{\infty} \Gamma'_s \end{aligned} \quad (5.1.5)$$

其中的矩阵 Γ'_s 为

$$\Gamma'_s = E \left\{ \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} [u_{1, t+s}, u'_{2, t+s}] \right\}$$

再由定理 1.9 的结论(1)、(7)和(9), 又可得

$$T^{-1/2} \sum_{t=1}^T \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \Rightarrow \Psi(1) P W(1) \quad (5.1.6)$$

$$T^{-3/2} \sum_{t=1}^T \begin{bmatrix} \bar{y}_{1t} \\ y_{2t} \end{bmatrix} \Rightarrow \Psi(1)P \int_0^1 W(r)dr \quad (5.1.7)$$

$$T^{-2} \sum_{t=1}^T \begin{bmatrix} \bar{y}_{1t} \\ y_{2t} \end{bmatrix} [\bar{y}_{1t}, y'_{2t}] \Rightarrow \Psi(1)P \left\{ \int_0^1 W(r)W(r)'dr \right\} P' \Psi(1)' \quad (5.1.8)$$

将 α 和 γ 分别从 $\hat{\alpha}$ 和 $\hat{\gamma}$ 中减去, 并以 $T^{1/2}$ 和 T 分别相乘, 可得

$$\begin{aligned} \begin{bmatrix} T^{1/2}(\hat{\alpha} - \alpha) \\ T(\hat{\gamma} - \gamma) \end{bmatrix} &= \left\{ \begin{bmatrix} T^{-1/2} & \mathbf{0}' \\ \mathbf{0} & T^{-1}I_{n-1} \end{bmatrix} \begin{bmatrix} T & \Sigma y'_{2t} \\ \Sigma y_{2t} & \Sigma y_{2t} y'_{2t} \end{bmatrix} \right. \\ &\quad \times \left. \begin{bmatrix} T^{-1/2} & \mathbf{0}' \\ \mathbf{0} & T^{-1}I_{n-1} \end{bmatrix} \right\}^{-1} \\ &\quad \times \left\{ \begin{bmatrix} T^{-1/2} & \mathbf{0}' \\ \mathbf{0} & T^{-1}I_{n-1} \end{bmatrix} \begin{bmatrix} \Sigma u_{1t} \\ \Sigma y_{2t} u_{1t} \end{bmatrix} \right\} \\ &= \begin{bmatrix} 1 & T^{-3/2} \Sigma y'_{2t} \\ T^{-3/2} \Sigma y_{2t} & T^{-2} \Sigma y_{2t} y'_{2t} \end{bmatrix}^{-1} \begin{bmatrix} T^{-1/2} \Sigma u_{1t} \\ T^{-1} \Sigma y_{2t} u_{1t} \end{bmatrix} \end{aligned} \quad (5.1.9)$$

令 $T \rightarrow \infty$, 对于 $T^{-1} \Sigma y_{2t} u_{1t}$, 则有

$$\begin{aligned} T^{-1} \sum_{t=1}^T y_{2t} u_{1t} &= [\mathbf{0}, I_{n-1}] \left\{ T^{-1} \sum_{t=1}^T \begin{bmatrix} \bar{y}_{1t} \\ y_{2t} \end{bmatrix} [u_{1t}, u'_{2t}] \right\} \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix} \\ &\Rightarrow [\mathbf{0}, I_{n-1}] \Psi(1) \cdot P \left\{ \int_0^1 W(r) [dW(r)]' \right\} \\ &\quad \times P' \Psi(1)' \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix} + [\mathbf{0}, I_{n-1}] \sum_{v=0}^{\infty} \Gamma_v' \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix} \end{aligned}$$

$$T^{-1} \sum_{i=1}^T (y_{1i} - \bar{y}_1)^2 \xrightarrow{p} +\infty$$

这就意味着,当样本量 T 足够大时,有

$$R^2 = 1 - \frac{T^{-1} \sum_{i=1}^T \hat{u}_{1i}^2}{T^{-1} \sum_{i=1}^T (y_{1i} - \bar{y}_1)^2} \xrightarrow{p} 1$$

因此,若 y_{1i} 和 y_{2i} 同积,在评价模型 $y_{1i} = \alpha + \gamma' y_{2i}$ 的拟合程度时, R^2 失去了通常的作用。

5.2 同积关系的规范化

在以最小二乘法估计同积关系时,常将某一个变量(如 y_{1i})的系数规范成 1,并以此变量对其他变量(如 y_{2i}, y_{3i}, \dots)作回归。但同积过程 $y_i = [y_{1i}, y_{2i}, \dots, y_{ni}]'$ 中各分量的排序并非唯一确定,问题是若对不同的变量的系数作规范,是否会得到不同的同积向量的估计值? 比如,在一个二维的同积系统 $[y_{1i}, y_{2i}]'$ 中,可以对 y_{1i} 的系数作规范,并由下式

$$y_{1i} = \alpha + \gamma y_{2i} + u_i$$

得到最小二乘估计 $\hat{\alpha}$ 和 $\hat{\gamma}$;也可对 y_{2i} 的系数作规范,然后由

$$y_{2i} = \theta + \eta y_{1i} + v_i$$

得到最小二乘估计 $\hat{\theta}$ 和 $\hat{\eta}$ 。以上两个回归分别给出了同积向量 $[1, -\gamma]'$ 和 $[-\eta, 1]'$ 的估计值,它们之间存在线性关系:

$$\begin{bmatrix} 1 \\ -\gamma \end{bmatrix} = \gamma \begin{bmatrix} \eta \\ -1 \end{bmatrix}$$

也即 $\gamma\gamma = 1$ 。在样本量有限的情况下, $\hat{\gamma}\hat{\gamma} = 1$ 不一定能成立, 因此以哪个变量作为回归的应变变量, 对分析的结果有直接的影响。解决这一问题的方法是采用下一章介绍的全信息似然估计方法, 这种方法不需要先对同积向量规范化, 各分量是联合估计得到的。

5.3 多个同积向量

在以上的讨论和定理 5.1 的证明中, 我们假设同积过程 $y_i = [y_{1i}, y'_{2i}]'$ 只有一个独立的同积向量, 而且这一向量相应于 y_{1i} 的分量不为零。但在许多情况下, n 维同积过程 y_i 可有多同积向量, 若以 y_{1i} 对其余变量作回归, 所得到的到底是哪个同积向量的估计值? 以下就来回答这一问题。

设 n 维同积过程 $y_i = [y_{1i}, y_{2i}]'$ 有 $k > 1$ 个独立的同积向量, 将 y_i 重新分块如下:

$$y_i = \begin{bmatrix} y_{1i} \\ y_{2i}^* \end{bmatrix}$$

其中, y_{1i} 和 y_{2i}^* 分别为 k 维和 $(n - k)$ 维子向量, 并设子向量 y_{2i}^* 各分量之间不存在同积关系。以三角形表示系统 $[y'_{1i}, y_{2i}^{*'}]'$ 为

$$\begin{aligned} y_{1i} &= \mu_1 + \Gamma' y_{2i}^* + u_i^* \\ \Delta y_{2i}^* &= \delta_2 + u_{2i}^* \end{aligned} \quad (5.3.1)$$

其中 $u_i^* = [u_{1i}^*, \dots, u_{ki}^*]'$ 是均值为零的稳定过程。对于给定的 $u_{1i}^*, \dots, u_{ki}^*$, 考虑由 $u_{2i}^*, \dots, u_{ni}^*$ 构成的 $(k - 1)$ 维子空间, 将 u_{1i}^* 对子空间投影, 可将 u_{1i}^* 表示为

证明：不失一般性，假设三角系统(5.3.1)中的 $\delta_2 = \mathbf{0}$ ，在方程(5.3.2)中添加 θ 和 $\eta' y_{2t}$ 两项，考虑回归模型：

$$\begin{aligned} u_{1t}^* &= \beta_2 u_{2t}^* + \beta_3 u_{3t}^* + \cdots + \beta_k u_{kt}^* + \theta + \eta' y_{2t} + u_t \\ &= \beta' u_{2t}^* + \theta + \eta' y_{2t} + u_t \end{aligned} \quad (5.3.3)$$

显然，参数 θ 和 η 的真值都为零。如果有随机变量 u_t^* 的观察值，参数 β 、 θ 和 η 的估计值可由最小二乘法得到：

$$\begin{bmatrix} \hat{\beta} \\ \hat{\theta} \\ \hat{\eta} \end{bmatrix} = \begin{bmatrix} \sum u_{2t}^* u_{2t}^{*'} & \sum u_{2t}^* & \sum u_{2t}^* y_{2t}' \\ \sum u_{2t}^{*'} & T & \sum y_{2t}' \\ \sum y_{2t} u_{2t}^{*'} & \sum y_{2t} & \sum y_{2t} y_{2t}' \end{bmatrix}^{-1} \begin{bmatrix} \sum u_{2t}^* u_{1t}^* \\ \sum u_{1t}^* \\ \sum y_{2t} u_{1t}^* \end{bmatrix} \quad (5.3.4)$$

从上式中减去各估计量的真值，以 \sqrt{T} 乘 $\hat{\eta}$ ，可得

$$\begin{aligned} \begin{bmatrix} \hat{\beta} - \beta \\ \hat{\theta} \\ \sqrt{T} \hat{\eta} \end{bmatrix} &= \begin{bmatrix} I_{k-1} & \mathbf{0} & 0 \\ \mathbf{0}' & 1 & \mathbf{0}' \\ 0 & 0 & \sqrt{T} I_{n-1} \end{bmatrix} \\ &\times \begin{bmatrix} \sum u_{2t}^* u_{2t}^{*'} & \sum u_{2t}^* & \sum u_{2t}^* y_{2t}' \\ \sum u_{2t}^{*'} & T & \sum y_{2t}' \\ \sum y_{2t} u_{2t}^{*'} & \sum y_{2t} & \sum y_{2t} y_{2t}' \end{bmatrix}^{-1} \begin{bmatrix} \sum u_{2t}^* u_{1t}^* \\ \sum u_{1t}^* \\ \sum y_{2t} u_{1t}^* \end{bmatrix} \\ &= \begin{bmatrix} T^{-1} \sum u_{2t}^* u_{2t}^{*'} & T^{-1} \sum u_{2t}^* & T^{-3/2} \sum u_{2t}^* y_{2t}' \\ T^{-1} \sum u_{2t}^{*'} & 1 & T^{-3/2} \sum y_{2t}' \\ T^{-3/2} \sum y_{2t} u_{2t}^{*'} & T^{-3/2} \sum y_{2t} & T^{-2} \sum y_{2t} y_{2t}' \end{bmatrix}^{-1} \begin{bmatrix} T^{-1} \sum u_{2t}^* u_{1t}^* \\ T^{-1} \sum u_{1t}^* \\ T^{-3/2} \sum y_{2t} u_{1t}^* \end{bmatrix} \end{aligned}$$

由(5.3.3)式， u_{1t}^* 与 u_t 不相关， $E(u_{1t}^* u_t) = 0$ ，根据大数定律，得

$$T^{-1} \sum u_{2t}^* u_t \xrightarrow{p} \mathbf{0}; \quad T^{-1} \sum u_t \xrightarrow{p} 0$$

另一方面，由(5.1.10)式，得

$$T^{-1} \sum_{t=1}^T y_{2t} u_t \Rightarrow \Lambda_2 \left\{ \int_0^1 W(r) [dW(r)]' \right\} \lambda_1 + \sum_{s=0}^{\infty} E(u_{2t} u_{t+s})$$

因此

$$T^{-3/2} \sum_{t=1}^T y_{2t} u_t \xrightarrow{P} 0$$

同时

$$\begin{aligned} & \begin{pmatrix} T^{-1} \Sigma u_{2t}^* u_{2t}' & T^{-1} \Sigma u_{2t}^* & T^{-3/2} \Sigma u_{2t}^* y_{2t}' \\ T^{-1} \Sigma u_{2t}^{*'} & 1 & T^{-3/2} \Sigma y_{2t}' \\ T^{-3/2} \Sigma y_{2t} u_{2t}' & T^{-3/2} \Sigma y_{2t} & T^{-2} \Sigma y_{2t} y_{2t}' \end{pmatrix} \\ \Rightarrow & \begin{pmatrix} E(u_{2t}^* u_{2t}') & 0 & 0 \\ 0' & 1 & \left\{ \int_0^1 W(r)' dr \right\} \Lambda_2' \\ 0 & \Lambda_2 \int_0^1 W(r) dr & \Lambda_2 \left\{ \int_0^1 W(r) W(r)' dr \right\} \Lambda_2 \end{pmatrix} \end{aligned}$$

其中, Λ_2 为非异矩阵 $\Psi(1) \cdot P$ 的最后 $(n-1)$ 行, 因此是 $(n-1) \times n$ 维行满秩矩阵。因此以上矩阵为非异的 $n \times n$ 维矩阵, 其逆存在。综上所述, 可得

$$\begin{pmatrix} \hat{\beta} - \beta \\ \hat{\theta} \\ \sqrt{T} \hat{\eta} \end{pmatrix} \xrightarrow{P} 0$$

在同积过程 $\{y_t\}$ 中, 以 y_{1t} 对其他分量作回归, 得

$$\begin{aligned} y_{1t} &= \alpha + \gamma_2 y_{2t} + \gamma_3 y_{3t} + \cdots + \gamma_n y_{nt} + v_t \\ &= \alpha + \gamma' y_{2t} + v_t \end{aligned} \quad (5.3.5)$$

其中各参数的最小二乘估计, 可由 (5.3.4) 中的估计值 $\hat{\beta}$ 、 $\hat{\theta}$ 和

$\hat{\eta}$ 变换得到。因为根据(5.3.3), 有

$$[1, -\beta']u_i^* = \theta + \eta'y_{2i} + u_i$$

以 $u_i^* = y_{1i} - \mu_1 - \Gamma'y_{2i}^*$ 代入, 得

$$[1, -\beta'](y_{1i} - \mu_1 - \Gamma'y_{2i}^*) = \theta + \eta'y_{2i} + u_i$$

整理后可得

$$y_{1i} = \beta_2 y_{2i} + \beta_3 y_{3i} + \cdots + \beta_k y_{ki} + \alpha + \eta^{*'} y_{2i} + u_i \quad (5.3.6)$$

其中的参数为

$$\begin{aligned} \alpha &= \theta + [1, -\beta']\mu_1 = [1, -\beta']\mu_1 \\ \eta^{*'} &= \eta' + [1, -\beta']\Gamma' = [1, -\beta']\Gamma' \end{aligned} \quad (5.3.7)$$

(因为参数 θ 和 η 的真值都为零。)因为模型(5.3.3)的参数估计 $\hat{\beta}$ 、 $\hat{\theta}$ 和 $\hat{\eta}$ 都是一致的, 相对应的模型(5.3.4)的参数估计 $\hat{\alpha}$ 和 $\hat{\gamma} = [\hat{\beta}', \hat{\eta}^*]'$ 也是一致的, 分别有极限

$$\begin{aligned} \hat{\alpha} &\xrightarrow{p} [1, -\beta']'\mu_1 \\ \hat{\gamma} &= \begin{bmatrix} \hat{\gamma}_1 \\ \vdots \\ \hat{\gamma}_n \end{bmatrix} = \begin{bmatrix} \hat{\beta} \\ \hat{\eta}^* \end{bmatrix} \xrightarrow{p} \begin{bmatrix} \beta \\ m \end{bmatrix} \end{aligned}$$

其中

$$m = \Gamma \begin{bmatrix} 1 \\ -\beta \end{bmatrix}$$

定理证毕。

定理 5.2 表明,当同积过程 $\{y_t\}$ 有 $k > 1$ 个同积向量时,以 y_{1t} 对 y_t 的其他分量作回归得到的参数估计 $\hat{\alpha}$ 和 $\hat{\gamma}$ 并不收敛于某个特定的同积向量,而是趋于 k 个同积向量的线性组合。 $\hat{\alpha}$ 和 $\hat{\gamma}$ 的极限还受到 u_{1t}^* 和 $u_{2t}^*, u_{3t}^*, \dots, u_{kt}^*$ 之间的相关系数的影响。

5.4 检验随机向量的同积性

若 $I(1)$ 随机变量 y_{1t} 和 y_{2t} 是同积的,它们之间的同积关系可表示为

$$y_{1t} = \alpha + \beta y_{2t} + u_t \quad (5.4.1)$$

由此得到的最小二乘估计 $\hat{\alpha}$ 和 $\hat{\beta}$ 是一致的估计量,估计残差 $\hat{u}_t = y_{1t} - \hat{\alpha} - \hat{\beta} y_{2t}$ 构成一 $I(0)$ 过程。但若 y_{1t} 和 y_{2t} 不同积, (5.4.1) 给出的是 y_{1t} 和 y_{2t} 之间的伪回归关系,由此得到的估计残差 \hat{u}_t 仍为 $I(1)$ 过程。所以 $\{\hat{u}_t\}$ 是否为单位根过程反映了 y_{1t} 和 y_{2t} 是否同积。对残差 \hat{u}_t 的单位根假设可用迪基—福勒 (DF) 或增广的迪基—福勒 (ADF) 方法,以及菲利普斯—配荣 (PP) 方法检验,但由于 \hat{u}_t 是最小二乘估计的残差,对它作单位根检验与第 2 章中的情况有所不同。

设 y_t 为一 $I(1)$ 的 n 维随机向量,并可分解为

$$y_t = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix}$$

其中, y_{1t} 为一标量, y_{2t} 为 $(n-1)$ 维随机向量。若随机向量 y_t 同积,而且同积向量中对应于 y_{1t} 的分量不为零,可用 y_{1t} 对 y_{2t} 中的各分量作回归:

最后,以 \hat{c}_j 表示 \hat{e}_t 的 j 阶样本自相关系数:

$$\hat{c}_j = (T-1)^{-1} \sum_{t=j+2}^T \hat{e}_t \hat{e}_{t-j}; \quad j = 0, 1, 2, \dots, T-2$$

构造检验假设 $\rho = 1$ 的 PP 方法的 ρ -统计量(参见 2.3.13 式)

$$\begin{aligned} Z_\rho = (T-1)(\hat{\rho}_T - 1) - \frac{1}{2} \{ (T-1)^2 \cdot \hat{\sigma}_\epsilon^2 / s_T^2 \} \\ \times (\hat{\sigma}^2 \hat{\varphi}^2(1) - \hat{c}_0) \end{aligned} \quad (5.4.5)$$

上式中的 $\hat{\sigma}^2 \hat{\varphi}^2(1)$ 可表示为

$$\hat{\sigma}^2 \hat{\varphi}^2(1) = \hat{c}_0 + 2 \sum_{j=1}^q \left[1 - \frac{j}{q+1} \right] \hat{c}_j$$

用 ρ -统计量检验单位根假设的方法已有详细讨论。若 Z_ρ 是显著的,则拒绝假设 $\rho = 1$,从而接受假设 \hat{u}_t 是稳定过程, y_{1t} 和 y_{2t} 是同积的。反之,若 ρ -统计量是不显著的,即它的计算值大于适当的临界值,则接受假设 $\rho = 1$,从而拒绝 y_{1t} 和 y_{2t} 同积的假设。

对假设 $\rho = 1$ 也可由 t -统计量作检验。PP 检验方法的 t -统计量为

$$\begin{aligned} Z_t = \frac{\hat{c}_0^{1/2}}{\hat{\sigma} \hat{\varphi}(1)} \cdot t_T - \frac{1}{2} \{ (T-1) \hat{\sigma}_\epsilon^2 / s_T \} \\ \times \{ \hat{\sigma}^2 \hat{\varphi}^2(1) - \hat{c}_0 \} / \hat{\sigma} \hat{\varphi}(1) \end{aligned}$$

其中, $t_T = (\hat{\rho}_T - 1) / \hat{\sigma}_\rho$ 。 t -统计量的用法与 ρ -统计量是相似的,只有当检验统计量的值是显著的,才拒绝 $\{\hat{u}_t\}$ 是单位根过程的假设,从而接受 y_{1t} 和 y_{2t} 同积的假设。

除了 PP 检验法以外,也可用 ADF 方法对 $\rho = 1$ 的假设作检验。将残差 \hat{u}_t 表示成

$$\hat{u}_t = \zeta_1 \Delta \hat{u}_{t-1} + \zeta_2 \Delta \hat{u}_{t-2} + \cdots + \zeta_{p-1} \Delta \hat{u}_{t-p+1} + \rho \hat{u}_{t-1} + \varepsilon_t \quad (5.4.6)$$

并假设 $\{\varepsilon_t\}$ 为独立同分布,对假设 $H_0: \rho = 1$ 的 ADF 检验在第 2 章中已有详细的讨论。

以下继续讨论统计量 Z_ρ 和 Z_t 在 $H_0: \rho = 1$ 为真时的极限和临界值。我们将指出,检验同积关系的统计量 Z_ρ 和 Z_t 的临界值不同于表 1 和表 2 中的值。以下定理 5.3 给出了统计量 Z_ρ 和 Z_t 的极限。

定理 5.3: 设 y_t 为 n 维的 $I(1)$ 随机向量,它的一阶差分可表示为

$$\Delta y_t = \sum_{s=0}^{\infty} \Psi_s \varepsilon_{t-s}$$

其中, $\{\varepsilon_t\}$ 为独立同分布的 n 维随机向量, $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = E(\varepsilon_t \varepsilon_t') = \Omega = PP'$, 矩阵列 $\{s \cdot \Psi_s\}_{s=0}^{\infty}$ 中的每一元素列都绝对可和。令 $\Lambda = \Psi(1)P$, 而且 $\Lambda\Lambda'$ 为非异的 $n \times n$ 维矩阵, 对 $(\Lambda\Lambda')^{-1}$ 作三角分解:

$$(\Lambda\Lambda')^{-1} = LL'$$

L 为满秩的 $n \times n$ 维下三角矩阵。若 y_t 不同积, 即在 (5.4.2) 中 $\rho = 1$, 以下结论成立。

(a) 统计量 $(T-1)(\hat{\rho}_T - 1)$ 有极限

$$(T-1)(\hat{\rho}_T - 1) \Rightarrow \left\{ \frac{1}{2} [1, -h_2'] W(1) W(1)' \begin{bmatrix} 1 \\ -h_2 \end{bmatrix} \right\}$$

$$\begin{aligned}
& -h_1 \mathbf{W}(1)' \begin{bmatrix} 1 \\ -\mathbf{h}_2 \end{bmatrix} \\
& - \frac{1}{2} [1, -\mathbf{h}_2'] \mathbf{L}' \{E(\Delta \mathbf{y}_t)(\Delta \mathbf{y}_t')\} \mathbf{L} \begin{bmatrix} 1 \\ -\mathbf{h}_2 \end{bmatrix} / H_n
\end{aligned}
\tag{5.4.7}$$

其中, $\mathbf{W}(r)$ 为 n 维标准维纳过程:

$$\mathbf{W}(r) = \begin{bmatrix} W_1(r) \\ \mathbf{W}_2(r) \end{bmatrix}$$

$W_1(r)$ 和 $\mathbf{W}_2(r)$ 分别为 1 维和 $(n-1)$ 维标准维纳过程; h_1 和 \mathbf{h}_2 分别为 1 维和 $(n-1)$ 维随机泛函:

$$\begin{bmatrix} h_1 \\ \mathbf{h}_2 \end{bmatrix} = \begin{bmatrix} 1 & \int_0^1 \mathbf{W}_2(r)' dr \\ \int_0^1 \mathbf{W}_2(r) dr & \int_0^1 \mathbf{W}_2(r) \mathbf{W}_2(r)' dr \end{bmatrix}^{-1} \begin{bmatrix} \int_0^1 W_1(r) dr \\ \int_0^1 \mathbf{W}_2(r) W_1(r) dr \end{bmatrix},$$

H_n 为

$$H_n = \int_0^1 W_1(r)^2 dr - \left[\int_0^1 W_1(r) dr, \int_0^1 \mathbf{W}_2(r) \mathbf{W}_2(r)' dr \right] \begin{bmatrix} h_1 \\ \mathbf{h}_2 \end{bmatrix}$$

(b) 若当 $T \rightarrow \infty$, 估计量

$$\hat{\sigma}^2 \hat{\varphi}^2(1) = \hat{c}_0 + 2 \sum_{j=1}^q \left[1 - \frac{j}{q+1} \right] \hat{c}_j,$$

中的正整数 q 也趋于无穷大, 但 $q/T \rightarrow 0$, 那么 PP 检验法的 ρ -统计量 Z_ρ 有极限

$$Z_\rho \Rightarrow Z_n \tag{5.4.8}$$

式中的 Z_n 为

$$Z_n = \left\{ \frac{1}{2} [1, -h_2'] W(1) W(1)' \begin{bmatrix} 1 \\ -h_2 \end{bmatrix} - h_1 W(1)' \begin{bmatrix} 1 \\ -h_2 \end{bmatrix} - \frac{1}{2} (1 + h_2' h_2) \right\} / H_n$$

(c) 仍假设 $T \rightarrow \infty$, $q \rightarrow \infty$, 但 $q/T \rightarrow 0$, 那么 PP 检验法的 t -统计量有极限

$$Z_t \Rightarrow Z_n \cdot \sqrt{H_n} / (1 - h_2' h_2)^{1/2}$$

(d) 若进一步假设 Δy_t 服从均值为零的自回归移动平均过程, 并当 $T \rightarrow \infty$ 时, Δy_t 的自回归阶数 P 也趋于无穷大, 但 $P/T^{1/3} \rightarrow 0$, 那么 ADF 检验法的 t -统计量和 PP 检验法的 ρ -统计量 Z_ρ 有一样的极限 Z_n 。

证明: 将 $n \times n$ 维矩阵 $\Lambda \Lambda'$ 写成

$$\Lambda \Lambda' = \begin{pmatrix} \Sigma_{11} & \Sigma'_{21} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

并定义上三角矩阵 L' 为

$$L' = \begin{pmatrix} \frac{1}{\sigma_1^*} & -\frac{1}{\sigma_1^*} \Sigma'_{21} \Sigma_{22}^{-1} \\ \mathbf{0} & L'_{22} \end{pmatrix} \quad (5.4.9)$$

其中, Σ_{11} 、 Σ_{21} 和 Σ_{22} 分别为 1 维、 $(n-1)$ 维和 $(n-1) \times (n-1)$ 维矩阵; L_{22} 为子矩阵 Σ_{22}^{-1} 的三角分解因子:

$$\Sigma_{22}^{-1} = L_{22} L'_{22}$$

标量 σ_1^* 可表示为

$$\sigma_1^* = \sqrt{\Sigma_{11} - \Sigma'_{21} \Sigma_{22}^{-1} \Sigma_{21}}$$

显然, $L' \Lambda \Lambda' L = I_n$ 。

对 y_{1t} 和 y_{2t} 作线性变换:

$$y_{1t}^* = y_{1t} - \Sigma'_{21} \Sigma_{22}^{-1} y_{2t}$$

$$y_{2t}^* = L'_{22} y_{2t}$$

不难验证, 以下的两个回归模型

$$y_{1t} = \alpha + \gamma' y_{2t} + u_t \quad (5.4.10)$$

$$y_{1t}^* = \alpha^* + \gamma^* y_{2t}^* + u_t^* \quad (5.4.11)$$

有同样的估计残差

$$\hat{u}_t = \hat{u}_t^*$$

根据定理 3.2 的结论(a), 回归模型(5.4.11)的参数的最小二乘估计 $\hat{\alpha}^*$ 和 $\hat{\gamma}^*$ 在 y_{1t} 和 y_{2t} 不同积时, 即 y_{1t} 和 y_{2t} 之间存在伪回归关系时, 有以下极限:

$$\begin{bmatrix} T^{1/2} \hat{\alpha}^* / \sigma_1^* \\ \hat{\gamma}^* / \sigma_1^* \end{bmatrix} \Rightarrow \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \quad (5.4.12)$$

以下证明定理的主要结论。

(a) 因为 $\hat{u}_t = \hat{u}_t^*$, 所以

$$\begin{aligned} (T-1)(\hat{\rho}_T - 1) &= (T-1) \left\{ \frac{\sum \hat{u}_{t-1}^* \hat{u}_t^*}{\sum \hat{u}_{t-1}^{*2}} - 1 \right\} \\ &= (T-1) \left\{ \frac{\sum \hat{u}_{t-1}^* \hat{u}_t^*}{\sum \hat{u}_{t-1}^{*2}} - 1 \right\} \end{aligned}$$

$$= \frac{(T-1)^{-1} \sum_{i=2}^T \hat{u}_{i-1}^* (\hat{u}_i^* - \hat{u}_{i-1}^*)}{(T-1)^{-2} \sum_{i=2}^T \hat{u}_i^{*2}} \quad (5.4.13)$$

其中的估计残差 \hat{u}_i^* 可表示为

$$\begin{aligned} \hat{u}_i^* &= y_{1i}^* - \hat{\alpha}^* - \hat{\gamma}^{*'} y_{2i}^* \\ &= \sigma_1^* \{ [1, -\hat{\gamma}^{*'} / \sigma_1^*] \xi_i^* - (\hat{\alpha}^* / \sigma_1^*) \} \end{aligned} \quad (5.4.14)$$

向量 ξ_i^* 为

$$\xi_i^* = \begin{bmatrix} y_{1i}^* / \sigma_1^* \\ y_{2i}^* \end{bmatrix} = L' y_i$$

对估计残差 \hat{u}_i^* 取一阶差分:

$$\Delta \hat{u}_i^* = \hat{u}_i^* - \hat{u}_{i-1}^* = \sigma^* [1, -\hat{\gamma}^{*'} / \sigma_1^*] \Delta \xi_i^*$$

由此, 可将(5.4.13)的分子写成

$$\begin{aligned} & (T-1)^{-1} \sum_{i=2}^T \hat{u}_{i-1}^* (\hat{u}_i^* - \hat{u}_{i-1}^*) \\ &= \sigma_1^{*2} (T-1)^{-1} \sum_{i=2}^T \{ [1, -\hat{\gamma}^{*'} / \sigma_1^*] \xi_{i-1}^* - (\hat{\alpha}^* / \sigma_1^*) \} \\ & \quad \times \left\{ (\Delta \xi_i^*)' \begin{bmatrix} 1 \\ -\hat{\gamma}^{*'} / \sigma_1^* \end{bmatrix} \right\} \\ &= \sigma_1^{*2} [1, -\hat{\gamma}^{*'} / \sigma_1^*] \left\{ (T-1)^{-1} \sum_{i=2}^T \xi_{i-1}^* (\Delta \xi_i^*)' \begin{bmatrix} 1 \\ -\hat{\gamma}^{*'} / \sigma_1^* \end{bmatrix} \right\} \\ & \quad - \sigma_1^{*2} (T-1)^{-1/2} (\hat{\alpha}^* / \sigma_1^*) \left\{ (T-1)^{-1/2} \sum_{i=2}^T (\Delta \xi_i^*)' \begin{bmatrix} 1 \\ -\hat{\gamma}^{*'} / \sigma_1^* \end{bmatrix} \right\} \end{aligned} \quad (5.4.15)$$

因为

$$[1, -\hat{Y}^{*'} / \sigma_1^*] \left\{ (T-1)^{-1} \sum_{i=2}^T \xi_{i-1}^* (\Delta \xi_i^*)' \right\} \begin{bmatrix} 1 \\ -\hat{Y}^* / \sigma_1^* \end{bmatrix}$$

为一标量, 所以等于它的转置, 从而有

$$\begin{aligned} & [1, -\hat{Y}^{*'} / \sigma_1^*] \left\{ (T-1)^{-1} \sum_{i=2}^T \xi_{i-1}^* (\Delta \xi_i^*)' \right\} \begin{bmatrix} 1 \\ -\hat{Y}^* / \sigma_1^* \end{bmatrix} \\ &= \frac{1}{2} \left\{ [1, -\hat{Y}^{*'} / \sigma_1^*] \left\{ (T-1)^{-1} \sum_{i=2}^T [\xi_{i-1}^* (\Delta \xi_i^*)' \right. \right. \right. \\ & \quad \left. \left. \left. + (\Delta \xi_i^*) \xi_{i-1}^{*'}] \right\} \begin{bmatrix} 1 \\ -\hat{Y}^* / \sigma_1^* \end{bmatrix} \right\} \end{aligned} \quad (5.4.16)$$

由定理 3.1 的结论(4), 可得

$$\begin{aligned} & (T-1)^{-1} \sum_{i=2}^T [\xi_{i-1}^* (\Delta \xi_i^*)' + (\Delta \xi_i^*) \xi_{i-1}^{*'}] \\ &= L' \left\{ (T-1)^{-1} \sum_{i=2}^T [y_{i-1} \Delta y_i' + \Delta y_i y_{i-1}'] \right\} L \\ &\Rightarrow L' \{ \Lambda W^*(1) W^{*'}(1)' \Lambda' - E(\Delta y_i \Delta y_i') \} L \\ &= W(1) W(1)' - E(\Delta \xi_i^* \Delta \xi_i^{*'}) \end{aligned} \quad (5.4.17)$$

其中的 $W^*(1)$ 和 $W(1)$ 都为 n 维标准维纳过程, 因 $L' \Lambda \Lambda L = I_n$, 故 $W(1) = L' \Lambda W^*(1)$ 。由(5.4.12)和(5.4.17)结果, 可得

$$[1, -\hat{Y}^{*'} / \sigma_1^*] \left\{ (T-1)^{-1} \sum_{i=2}^T \xi_{i-1}^* (\Delta \xi_i^*)' \right\} \begin{bmatrix} 1 \\ -\hat{Y}^* / \sigma_1^* \end{bmatrix}$$

$$\begin{aligned}
&= (T-1)^{-1} \sum_{i=j+2}^T \{ \Delta \hat{u}_{i-1}^* - (\hat{\rho}_T - 1) \hat{u}_{i-1}^* \} \\
&\quad \times \{ \Delta \hat{u}_{i-1}^* - (\hat{\rho}_T - 1) \hat{u}_{i-1}^* \} \quad (5.4.19)
\end{aligned}$$

还可将其中的交叉项进一步展开为

$$\begin{aligned}
&(T-1)^{-1} \sum_{i=j+2}^T (\hat{\rho}_T - 1) \hat{u}_{i-1}^* \Delta \hat{u}_{i-1}^* \\
&= \sigma_1^{*2} (\hat{\rho}_T - 1) (T-1)^{-1} \sum_{i=j+2}^T \{ [1, -\hat{\gamma}^{*'} / \sigma_1^*] \xi_{i-1}^* \\
&\quad - (\hat{\alpha}^{*'} / \sigma_1^*) \Delta \xi_{i-1}^{*'} \left[\begin{array}{c} 1 \\ -\hat{\gamma}^{*'} / \sigma_1^* \end{array} \right] \} \\
&= \sigma_1^{*2} [(T-1)^{1/2} (\hat{\rho}_T - 1)] \\
&\quad \times \left\{ [1, -\hat{\gamma}^{*'} / \sigma_1^*] (T-1)^{-3/2} \sum_{i=j+2}^T \xi_{i-1}^* \Delta \xi_{i-1}^{*'} \left[\begin{array}{c} 1 \\ -\hat{\gamma}^{*'} / \sigma_1^* \end{array} \right] \right\} \\
&= \sigma_1^{*2} [(T-1)^{1/2} (\hat{\rho}_T - 1)] \\
&\quad \times \left\{ [(T-1)^{-1/2} (\hat{\alpha}_1^{*'} / \sigma_1^*)] (T-1)^{-1} \sum_{i=j+2}^T \Delta \xi_{i-1}^{*'} \left[\begin{array}{c} 1 \\ -\hat{\gamma}^{*'} / \sigma_1^* \end{array} \right] \right\} \quad (5.4.20)
\end{aligned}$$

首先,由本定理的结论(a),有 $(T-1)^{1/2} (\hat{\rho}_T - 1) \xrightarrow{P} 0$; 再由定理 3.1 的结论(1)和(5)以及(5.4.12)式中的结果,可知上式的花括弧中的和式都有有限极限,因此

$$(T-1)^{-1} \sum_{i=j+2}^T (\hat{\rho}_T - 1) \hat{u}_{i-1}^* \Delta \hat{u}_{i-1}^* \xrightarrow{P} 0 \quad (5.4.21)$$

同理还可得

$$\begin{aligned}
& (T-1)^{-1} \sum_{t=j+2}^T (\hat{\rho}_T - 1)^2 \hat{u}_{t-1}^* \hat{u}_{t-j-1}^* \\
&= \sigma_1^{*2} (T-1)^{-1} \sum_{t=j+2}^T (\hat{\rho}_T - 1)^2 \{ [1, -\hat{\gamma}^{*'} / \sigma_1^*] \xi_{t-1}^* - (\hat{\alpha}^* / \sigma_1^*) \} \\
&\quad \times \{ [1, -\hat{\gamma}^{*'} / \sigma_1^*] \xi_{t-j-1}^* - (\hat{\alpha}^* / \sigma_1^*) \} \\
&= \sigma_1^{*2} (T-1)^{-1} \sum_{t=j+2}^T (\hat{\rho}_T - 1)^2 [1, -\hat{\gamma}^{*'} / \sigma_1^{*'}] \\
&\quad - (T-1)^{-1/2} \hat{\alpha}^* / \sigma_1^*] \\
&\quad \times \begin{bmatrix} \xi_{t-1}^* \\ (T-1)^{1/2} \end{bmatrix} [\xi_{t-j-1}^{*'}, (T-1)^{1/2}] \begin{bmatrix} 1 \\ -\hat{\gamma}^{*'} / \sigma_1^* \\ -(T-1)^{-1/2} \hat{\alpha}^* / \sigma_1^* \end{bmatrix} \\
&= \sigma_1^{*2} [(T-1)^{1/2} (\hat{\rho}_T - 1)]^2 [1, -\hat{\gamma}^{*'} / \sigma_1^*, \\
&\quad - (T-1)^{-1/2} \hat{\alpha}^* / \sigma_1^*] \\
&\quad \times \left\{ (T-1)^{-2} \sum_{t=j+2}^T \begin{bmatrix} \xi_{t-1}^* \xi_{t-j-1}^{*'} & (T-1)^{1/2} \xi_{t-1}^* \\ (T-1)^{1/2} \xi_{t-j-1}^{*'} & T-1 \end{bmatrix} \right\} \\
&\quad \times \begin{bmatrix} 1 \\ -\hat{\gamma}^{*'} / \sigma_1^* \\ -(T-1)^{-1/2} \hat{\alpha}^* / \sigma_1^* \end{bmatrix} \xrightarrow{p} 0 \quad (5.4.22)
\end{aligned}$$

另一方面, 根据定理 3.1 的结论 (7) 和 (9), $(T-1)^{-2} \times \sum_{t=j+2}^T \xi_{t-1}^* \xi_{t-j-1}^{*'}$ 和 $(T-1)^{-3/2} \sum_{t=j+2}^T \xi_{t-1}^*$, 都有有限极限。(5.4.21) 式和 (5.4.22) 式说明 \hat{c}_j 和 $(T-1)^{-1} \sum_{t=j+2}^T \Delta \hat{u}_t^* \Delta \hat{u}_{t-j}^*$ 有相同的极限, 而后者的极限为

$$\begin{aligned}
& (T-1)^{-1} \sum_{i=j+2}^T \Delta \hat{u}_i^* \Delta \hat{u}_{i-}^*, \\
& = \sigma_1^{*2} [1, -\hat{\gamma}^{*'} / \sigma_1^*] \left\{ (T-1)^{-1} \sum_{i=j+2}^T \Delta \xi_i^* \Delta \xi_{i-}^* \right\} \begin{bmatrix} 1 \\ -\hat{\gamma}^{*'} / \sigma_1^* \end{bmatrix} \\
& \Rightarrow \sigma_1^{*2} [1, -h_2'] [E(\Delta \xi_i^* \Delta \xi_{i-}^*)] \begin{bmatrix} 1 \\ -h_2 \end{bmatrix} \\
& = \sigma_1^{*2} [1, -h_2'] L' E(\Delta y_i \Delta y_{i-}) L \begin{bmatrix} 1 \\ -h_2 \end{bmatrix}
\end{aligned}$$

对任意给定的正整数 q , 可得

$$\begin{aligned}
\hat{\lambda}^2 &= \hat{\sigma}^2 \hat{\varphi}^2(1) = \hat{c}_0 + 2 \sum_{j=1}^q \left[1 - \frac{j}{q+1} \right] \hat{c}_j, \\
&\Rightarrow \sigma_1^{*2} [1, -h_2'] L' \left\{ \sum_{j=-q}^q \left[1 - \frac{|j|}{q+1} \right] E(\Delta y_i \Delta y_{i-j}') \right\} L \begin{bmatrix} 1 \\ -h_2 \end{bmatrix}
\end{aligned}$$

若再令 $T \rightarrow \infty$, $q \rightarrow \infty$ 和 $q/T \rightarrow 0$, 则有

$$\begin{aligned}
\hat{\lambda}^2 &\Rightarrow \sigma_1^{*2} [1, -h_2'] L' \left\{ \sum_{j=-\infty}^{\infty} E(\Delta y_i \Delta y_{i-j}') \right\} L \begin{bmatrix} 1 \\ -h_2 \end{bmatrix} \\
&= \sigma_1^{*2} [1, -h_2'] L' \Psi(1) P P' \Psi(1)' L \begin{bmatrix} 1 \\ -h_2 \end{bmatrix} \\
&= \sigma_1^{*2} [1, -h_2'] I_n \begin{bmatrix} 1 \\ -h_2 \end{bmatrix} \\
&= \sigma_1^{*2} (1 + h_2' h_2)
\end{aligned} \tag{5.4.23}$$

同时, $\hat{\sigma}_\varepsilon^2 = s_T^2 / \sum_{i=2}^T \hat{u}_i^2$, 根据结论(a), 可得

$$(T-1)^2 \hat{\sigma}_\rho^2 / s_T^2 = \frac{1}{(T-1)^2 \sum_{t=2}^T \hat{u}_{t-1}^2} \Rightarrow (\sigma_1^*)^{-2} \cdot H_n^{-1} \quad (5.4.24)$$

综合以上讨论,就有

$$\begin{aligned} & \{(T-1)^2 \hat{\sigma}_\rho^2 / s_T^2\} \{ \hat{\sigma}^2 \hat{\varphi}^2(1) - \hat{c}_0 \} \\ & \Rightarrow [1, -h_2'] [I_n - L'E(\Delta y_t \Delta y_{t-1}') L] \begin{bmatrix} 1 \\ -h_2 \end{bmatrix} / H_n \end{aligned}$$

以 $\frac{1}{2}$ 乘上式,并将乘积从(5.4.7)的 $(T-1)(\hat{\rho}_T-1)$ 的极限中减去,就可得

$$\begin{aligned} Z_p &= (T-1)(\hat{\rho}_T-1) - \frac{1}{2} \{ (T-1)^2 \hat{\sigma}_\rho^2 / s_T^2 \} (\hat{\sigma}^2 \hat{\varphi}^2(1) - \hat{c}_0) \\ &\Rightarrow \left\{ \frac{1}{2} [1, -h_2'] W(1) W(1)' \begin{bmatrix} 1 \\ -h_2 \end{bmatrix} \right. \\ &\quad \left. - h_1 W(1)' \begin{bmatrix} 1 \\ -h_2 \end{bmatrix} - \frac{1}{2} (1 + h_2' h_2) \right\} / H_n \end{aligned}$$

结论(b)得证。

(c) 将统计量 Z_t 改写成

$$\begin{aligned} Z_t &= \left[\frac{\hat{c}_0}{\hat{\sigma}^2 \hat{\varphi}^2(1)} \right]^{1/2} \cdot t_T - \frac{1}{2} \{ (T-1) \hat{\sigma}_\rho^2 / s_T \} \\ &\quad \times \{ \hat{\sigma}^2 \hat{\varphi}^2(1) - \hat{c}_0 \} / \hat{\sigma}^2 \hat{\varphi}^2(1) \end{aligned}$$

$$T(\hat{\rho}_T - 1) \Rightarrow \frac{\frac{1}{2}(W^2(1) - 1)}{\int_0^1 W^2(r)dr}$$

因此,对估计残差 \hat{u}_t 作单位根检验,由此对 y_t 同积作检验时,不能用表 1 和表 2 中的临界值。尽管统计量 Z_ρ 和 Z_t 的极限有较为复杂的形式,其中的 h_1 、 h_2 和 H_n 并不依赖于未知参数,只决定于回归中所包含的回归变量的个数,因此可对 Z_ρ 和 Z_t 的极限的分布值作模拟计算,它们的极限分布的临界值分别列于表 4 与表 5 中。

表 4 的情况一适用于不带常数项的同积模型:

$$y_{1t} = \gamma_2 y_{2t} + \gamma_3 y_{3t} + \cdots + \gamma_n y_{nt} + u_t$$

表中的数值是统计量 Z_ρ 在不同显著水平上的临界值,这些值由蒙的可罗方法模拟计算所得。比如,要检验四个变量 y_{1t} 、 y_{2t} 、 y_{3t} 和 y_{4t} 之间的同积关系,先以 y_{1t} 对其他变量作回归得估计残差 \hat{u}_t ,再以 \hat{u}_t 对 \hat{u}_{t-1} 作回归得估计量 $\hat{\rho}_T$,从而构造 Z_ρ 统计量:

$$Z_\rho = (T-1)(\hat{\rho}_T - 1) - \frac{1}{2} \{ (T-1)^2 \cdot \hat{\sigma}_\rho^2 / s_T^2 \} (\hat{\sigma}^2 \hat{\varphi}(1) - \hat{c}_0)$$

若这时样本量 $T = 500$,在原假设 $H_0: \rho = 1$ 为真时,统计量 Z_ρ 有 95% 的概率大于临界值 -27.9。若 Z_ρ 的实际计算值小于 -27.9,则拒绝 \hat{u}_t 中存在单位根的假设,从而接受变量 y_{1t} 、 y_{2t} 、 y_{3t} 和 y_{4t} 同积的假设。

表 4 的情况二适用于带常数项的同积模型:

$$y_{1t} = \alpha + \gamma_2 y_{2t} + \gamma_3 y_{3t} + \cdots + \gamma_n y_{nt} + u_t$$

仍以四个变量 y_{1t} 、 y_{2t} 、 y_{3t} 和 y_{4t} 为例。以 y_{1t} 对常数项、 y_{2t} 、 y_{3t} 和 y_{4t} 作回归,得估计残差 \hat{u}_t ,再以 \hat{u}_t 对 \hat{u}_{t-1} 作回归得估计量 $\hat{\rho}_T$,

并对估计残差 $\hat{u}_t = y_{1t} - \hat{a} - \hat{b}t$ 用 ADF 方法作单位根检验,即在

$$\begin{aligned}\hat{u}_t = & \zeta_1 \Delta \hat{u}_{t-1} + \zeta_2 \Delta \hat{u}_{t-2} + \cdots + \zeta_{p-1} \Delta \hat{u}_{t-p+1} \\ & + \rho \hat{u}_{t-1} + e_t\end{aligned}$$

中检验 $\rho = 1$ 的统计量,与在下式

$$\begin{aligned}y_{1t} = & \zeta_1 \Delta y_{1,t-1} + \zeta_2 \Delta y_{1,t-2} + \cdots + \zeta_{p-1} \Delta y_{1,t-p+1} \\ & + \alpha + \rho y_{1,t-1} + \delta t + u_t\end{aligned}$$

中,对 $\rho = 1$ 作检验的统计量是渐近等价的。利用这一等价关系可以进一步证明:当 y_{2t} 中含有不为零的常数项时,在下式

$$\hat{u}_t = \zeta_1 \Delta \hat{u}_{t-1} + \zeta_2 \Delta \hat{u}_{t-2} + \cdots + \zeta_{p-1} \Delta \hat{u}_{t-p+1} + \rho \hat{u}_{t-1} + e_t \quad (5.4.27)$$

中检验 $\rho = 1$ 的 t -统计量,与在下式

$$\begin{aligned}y_{1t}^* = & \zeta_1 \Delta y_{1,t-1}^* + \zeta_2 \Delta y_{1,t-2}^* + \cdots + \zeta_{p-1} \Delta y_{1,t-p+1}^* \\ & + \alpha + \rho y_{1,t-1}^* + \delta t + u_t\end{aligned} \quad (5.4.28)$$

中检验 $\rho = 1$ 的 t -统计量是等价的。这里,(5.4.27)式中的 \hat{u}_t 是同积回归

$$y_{1t} = \alpha + \gamma y_{2t} + u_t$$

的估计残差;(5.4.28)式中的 $y_{1t}^* = y_{1t} - (\delta_1/\delta_2)y_{2t}$ 。根据以前的讨论可知,对(5.4.28)作 $\rho = 1$ 检验的 t -统计量的临界值列在表 2 的情况四中。

这表明,若 y_{2t} 含有非零的常数项,对同积关系

$$y_{1t} = \alpha + \gamma y_{2t} + u_t \quad (5.4.29)$$

的检验,可从(5.4.27)得估计残差 \hat{u}_t , 然后对 \hat{u}_t 用 ADF 方法作单位根检验,即在

$$\hat{u}_t = \zeta_1 \Delta \hat{u}_{t-1} + \zeta_2 \Delta \hat{u}_{t-2} + \cdots + \zeta_{p-1} \Delta \hat{u}_{t-p+1} + \rho \hat{u}_{t-1} + e_t$$

中检验 $\rho = 1$; 或用 PP 方法作单位根检验。计算所得的 t -统计量 Z_t 的临界值列在表 2 的情况四中,而不是表 5 的情况二中。为方便起见,表 2 的情况四中相应于 $T = 500$ 的临界值,重新列于表 5 的情况三的第一行。

当然,也可用 PP 方法的 Z_ρ 统计量,或 ADF 方法的 ρ -统计量对 \hat{u}_t 作单位根检验,只是计算所得的 Z_ρ 统计量和 ADF 的 ρ -统计量的值应与表 1 的情况四中的临界值作比较,这些临界值也重新列于表 4 的情况三中。

以下进一步考虑在 n 维同积向量 y_t 中,非零常数项 δ 在同积假设的检验中所起的作用。在同积模型

$$y_{1t} = \alpha + \gamma_2 y_{2t} + \gamma_3 y_{3t} + \cdots + \gamma_n y_{nt} + u_t \quad (5.4.30)$$

中,令 $\delta_i (i = 1, 2, \cdots, n)$ 为 y_{it} 的一阶差分的期望值:

$$E(\Delta y_{it}) = \delta_i$$

并假设 $\delta_n \neq 0$ 。对(5.4.30)作同积假设检验,可先取得估计残差 \hat{u}_t , 然后以 \hat{u}_t 构造 ADF 方法的 ρ -统计量或 t -统计量,或者构造 PP 方法的 Z_ρ 和 Z_t 统计量,并与表 4 和表 5 的情况三中的临界值相比较。

例 5.1 购买力平价(purchasing power parity)是经济学中的一个著名原理,它常用来解释不同货币兑换率的关系。在有效市场的假设下,两个国家的物价水平和两种货币的交换率应有下列关系式:

$$p_t = p_t^* + s_t \quad (5.4.31)$$

其中的 p_t 和 p_t^* 是取对数后的物价水平, s_t 为取对数后的兑换率。关系式(5.4.31)是否成立,可作为判断市场是否开放和是否充分竞争的证据。物价水平 p_t 、 p_t^* 和兑换率 s_t 一般为非稳定的 I(1)变量,因此只有当它们同积时,(5.4.31)才可能成立。以下用美国和意大利 1973 年 1 月到 1989 年 10 月的月数据检验购买力平价的假设。构造回归模型:

$$p_t = \alpha + \beta s_t + \gamma p_t^* + u_t$$

其中 p_t 和 p_t^* 分别为美国和意大利的物价水平的对数值, s_t 为美元对里拉的交换率的对数值。由数据可得

$$p_t = \begin{matrix} 2.71 \\ (0.37) \end{matrix} + \begin{matrix} 0.051 \\ (0.012) \end{matrix} s_t + \begin{matrix} 0.5300 \\ (0.0067) \end{matrix} p_t^* + \hat{u}_t \quad (5.4.32)$$

为检验 p_t 、 p_t^* 和 s_t 之间的同积关系,以 \hat{u}_t 对 \hat{u}_{t-1} 作一阶自回归估计,可得

$$\hat{u}_t = \begin{matrix} 0.98331 \\ (0.01172) \end{matrix} \hat{u}_{t-1} + \hat{e}_t$$

$$s_T^2 = (T-2)^{-1} \sum_{t=2}^T \hat{e}_t^2 = (0.40374)^2$$

$$\hat{c}_0 = 0.1622$$

$$\hat{c}_j = (T-1)^{-1} \sum_{t=j+2}^T \hat{e}_t \hat{e}_{t-j}; \quad j = 1, 2, \dots, 12$$

$$\hat{\sigma}^2 \hat{\varphi}^2(1) = \hat{c}_0 + 2 \sum_{j=1}^{12} \left[1 - \frac{j}{13}\right] \hat{c}_j = 0.4082 \quad (5.4.33)$$

这里 j 取值 1, 2, ..., 12, 是为了反映数据在一年内的相关性。

采用 PP 检验法, 计算 Z_ρ 统计量的值:

$$\begin{aligned} Z_\rho &= (T-1)(\hat{\rho}_T - 1) - \frac{1}{2} \{ (T-1)^2 \hat{\sigma}_\rho^2 / s_T^2 \} \{ \hat{\sigma}^2 \hat{\varphi}^2(1) - \hat{c}_0 \} \\ &= (201)(0.98331 - 1) - \frac{1}{2} \left\{ (201)^2 \frac{(0.01172)^2}{(0.40374)^2} \right\} \\ &\quad \times (0.4082 - 0.1622) \\ &= -7.54 \end{aligned}$$

由于在这一时期中, 美国和意大利(特别是意大利)都经历了持续的通货膨胀, 因此可认为 p_t 和 p_t^* 都为含有非零常数项的单位根过程。这样, 同积模型(5.4.32)有两个含非零常数的回归变量(s_t 和 p_t^*), 从表 4 情况三中可查得相应于 $(n-1) = 2$, 显著水平为 5% 的临界值为 -27.1 。因为 $-7.54 > -27.1$, 所以接受 \hat{u}_t 有单位根的假设, 从而拒绝 p_t 、 p_t^* 和 s_t 同积。

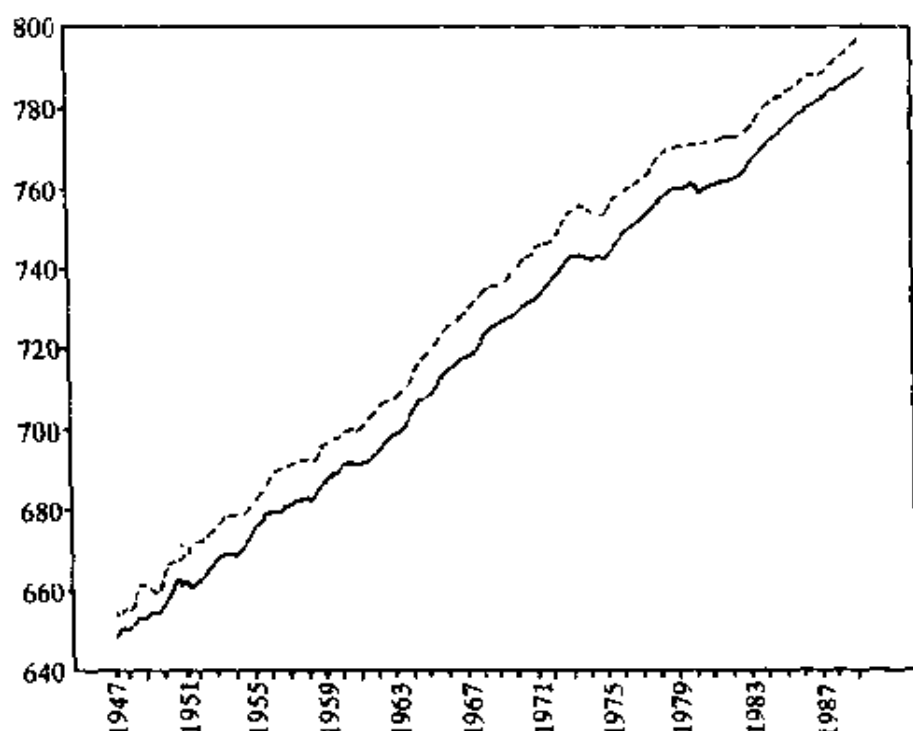
同样, 也可用 Z_t 统计量作检验。因为

$$\begin{aligned} Z_t &= (\hat{c}_0 / \hat{\sigma}^2 \hat{\varphi}^2(1))^{1/2} (\hat{\rho}_T - 1) / \hat{\sigma}_\rho \\ &\quad - \frac{1}{2} \{ (T-1) \hat{\sigma}_\rho^2 / s_T^2 \} (\hat{\sigma}^2 \hat{\varphi}^2(1) - \hat{c}_0) / \hat{\sigma}^2 \hat{\varphi}^2(1) \\ &= \{ (0.1622) / (0.4082) \}^{1/2} (0.98331 - 1) / (0.01172) \\ &\quad - \frac{1}{2} \{ (201)(0.01172) / (0.40374) \} \\ &\quad \times (0.4082 - 0.1622) / (0.4082)^{1/2} \\ &= -2.02 \end{aligned}$$

由表 5 的情况三中查得相应于 $n-1 = 2$ 和显著水平为 5% 的临界值为 -3.80 。因为 $-2.02 > -3.80$, 又一次拒绝 p_t 、 p_t^* 和 s_t 同积的假设。估计模型(5.4.32)是一个很好的伪回归关系

的例子,式中的每一个估计值都有很小的标准差(括弧中的值),因此相应的每一个 t_T 值都很显著,但 p_t 、 p_t^* 和 s_t 之间实际上不存在同积关系。

例 5.2 下图给出了从 1947 年到 1989 年美国的个人可支配收入总数 y_t 和同期的国民消费总数 c_t 的季度数据。其中 y_t 和 c_t 都为对数值,分别以虚线和实线表示。



以下检验 y_t 和 c_t 之间的同积关系。从图上可看出, y_t 和 c_t 都带有明显的时间趋势。对 y_t 用 ADF 方法作单位根检验:

$$y_t = \alpha + \sum_{j=1}^6 \beta_j \Delta y_{t-j} + \delta t + \rho y_{t-1} + u_t$$

得 ADF 的 t 值为 -1.28 。再对 c_t 作单位根检验:

$$c_t = \alpha + \sum_{j=1}^6 \beta_j \Delta c_{t-j} + \delta t + \rho c_{t-1} + u_t$$

得 ADF 的 t 值为 -1.88 。由于这两个 t 值都大于相应的临界

$$\begin{aligned}
Z_\rho &= (T-1)(\hat{\rho}_T - 1) - \frac{1}{2} \left\{ (T-1)^2 \hat{\sigma}_\rho^2 / s_T^2 \right\} \\
&\quad \times (\hat{\sigma}^2 \hat{\varphi}^2(1) - \hat{c}_0) \\
Z_t &= (\hat{c}_0 / \hat{\sigma}^2 \hat{\varphi}^2(1))^{1/2} (\hat{\rho}_T - 1) / \hat{\sigma}_\rho \\
&\quad - \frac{1}{2} \left\{ (T-1) \hat{\sigma}_\rho / s_T \right\} (\hat{\sigma}^2 \hat{\varphi}^2(1) - \hat{c}_0)
\end{aligned}$$

它们的临界值分别由表 4 和表 5 的情况一给出。在上式中

$$\begin{aligned}
s_T^2 &= (T-2)^{-1} \sum_{i=2}^T \hat{e}_i^2, & \hat{e}_i &= \hat{u}_i - \hat{\rho}_T \hat{u}_{i-1} \\
\hat{\sigma}_\rho^2 &= s_T^2 / \sum_{i=2}^T \hat{u}_{i-1}^2, & \hat{c}_j &= (T-1)^{-1} \sum_{i=2}^T \hat{e}_i \hat{e}_{i-1} \\
\hat{\sigma}^2 \hat{\varphi}^2(1) &= \hat{c}_0 + 2 \sum_{j=1}^q \left[1 - \frac{j}{q+1} \right] \hat{c}_j
\end{aligned}$$

也可用估计残差 \hat{u}_i 构造 ADF 检验法的 t -统计量,即在回归模型

$$\hat{u}_i = \zeta_1 \Delta \hat{u}_{i-1} + \cdots + \zeta_{p-1} \Delta \hat{u}_{i-p+1} + \rho \hat{u}_{i-1} + e_i$$

中,计算检验 $\rho = 1$ 的 t -统计量,它与 Z_t 有相同的临界值。

类型二: 检验同积关系

$$y_{1t} = \alpha + \gamma_2 y_{2t} + \gamma_3 y_{3t} + \cdots + \gamma_n y_{nt} + u_t$$

用估计残差 \hat{u}_i 作一阶自回归估计,从而计算 ADF 方法的 t_T 值,或 PP 方法的 Z_ρ 和 Z_t 的值,并与表 4 和表 5 的情况二中的临界值作比较。

类型三: 检验同积关系

$$y_{1t} = \alpha + \gamma_2 y_{2t} + \gamma_3 y_{3t} + \cdots + \gamma_n y_{nt} + u_t$$

但此时 $y_t = [y_{1t}, y_{2t}, \dots, y_{nt}]'$ 是带非零常数项的向量单位根过程,有表示式

$$\Delta y_t = \delta + \sum_{s=0}^{\infty} \Psi_s \varepsilon_{t-s}$$

与前面一样,以估计残差 \hat{u}_t 构造一阶自回归模型

$$\hat{u}_t = \hat{\rho}_T \hat{u}_{t-1} + \hat{e}_t$$

从而计算 Z_ρ 、 Z_t 或 ADF 的 t_T 值,并与表 4 和表 5 的情况三中的临界值相比较。

5.5 同积向量的假设检验

如果 $I(1)$ 变量 $y_{1t}, y_{2t}, \dots, y_{nt}$ 之间存在同积关系,则存在一分量不全为零的实数向量 $\alpha' = [\alpha_1, \alpha_2, \dots, \alpha_n]$, 使得

$$\alpha_1 y_{1t} + \alpha_2 y_{2t} + \dots + \alpha_n y_{nt} \sim I(0)$$

以下讨论对同积向量 α 各分量之间的约束条件作假设检验。若已知同积向量 α 存在,识别这些约束条件不仅能帮助识别 α 中的各分量,还往往赋予 α 重要的经济意义。本节讨论对同积向量 α 的线性约束的检验方法。考虑以三角形式表示的同积系统:

$$y_{1t} = \alpha + Y' y_{2t} + u_{1t}$$

$$y_{2t} = y_{2,t-1} + u_{2t} \quad (5.5.1)$$

其中的 y_{2t} 为 $(n-1)$ 维随机向量, u_{1t} 和 u_{2t} 分别独立同分布,且互不相关:

当原假设 H_0 为真时, 条件分布 (5.5.3) 是期望为零的正态分布。以统计量 s_T^2 作为未知参数 σ_1^2 的估计量, 即

$$s_T^2 = (T - n)^{-1} \sum_{t=1}^T (y_{1t} - \hat{\alpha}_T - \hat{\gamma}'_T y_{2t})^2$$

并构造二次型

$$Q = (R_1 \hat{\alpha}_T + R_2 \hat{\gamma}_T - d)' \left\{ s_T^2 [R_1, R_2] \begin{bmatrix} T & \Sigma y'_{2t} \\ \Sigma y_{2t} & \Sigma y_{2t} y'_{2t} \end{bmatrix}^{-1} \begin{bmatrix} R'_1 \\ R'_2 \end{bmatrix} \right\}^{-1} \\ \times (R_1 \hat{\alpha}_T + R_2 \hat{\gamma}_T - d)/m$$

这里的正数 m 为约束条件的个数。在原假设 $H_0: R_1 \alpha + R_2 \gamma = d$ 为真时, 二次型 Q 的条件分布是自由度为 $(m, T - n)$ 的 F 分布, 记为

$$H_0 \\ Q \sim F(m, T - n)$$

以下进一步考虑当 $T \rightarrow \infty$ 时, 二次型 Q 的极限分布。分别以 $T^{1/2}$ 和 T 乘 $(\hat{\alpha}_T - \alpha)$ 和 $(\hat{\gamma}_T - \gamma)$, 可得

$$\begin{bmatrix} T^{1/2}(\hat{\alpha}_T - \alpha) \\ T(\hat{\gamma}_T - \gamma) \end{bmatrix} \sim N \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma_1^2 \begin{bmatrix} T^{1/2} & 0' \\ 0 & T \cdot I_{n-1} \end{bmatrix} \begin{bmatrix} T & \Sigma y'_{2t} \\ \Sigma y_{2t} & \Sigma y_{2t} y'_{2t} \end{bmatrix}^{-1} \right. \\ \left. \times \begin{bmatrix} T^{1/2} & 0' \\ 0 & T \cdot I_{n-1} \end{bmatrix} \right\} \\ = N \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma_1^2 \begin{bmatrix} 1 & T^{-3/2} \Sigma y'_{2t} \\ T^{-3/2} \Sigma y_{2t} & T^{-2} \Sigma y_{2t} y'_{2t} \end{bmatrix}^{-1} \right\}$$

利用定理 5.1 的结果, 令 $\Psi(L) = I_n$, 以及

$$P = \begin{pmatrix} \sigma_1 & \mathbf{0}' \\ \mathbf{0} & P_{22} \end{pmatrix}$$

P_{22} 为非异的下三角矩阵, 满足条件

$$\Omega_{22} = P_{22} P_{22}'$$

定理 5.1 中的矩阵 $\Psi(1)P$ 可表示成:

$$\Psi(1)P = P = \begin{pmatrix} \sigma_1 & \mathbf{0}' \\ \mathbf{0} & P_{22} \end{pmatrix}$$

根据定理 5.1, 可得

$$\begin{aligned} \begin{pmatrix} T^{1/2}(\hat{a}_T - \alpha) \\ T(\hat{\gamma}_T - \gamma) \end{pmatrix} &= \begin{pmatrix} 1 & T^{-3/2} \Sigma y'_{2t} \\ T^{-3/2} \Sigma y_{2t} & T^{-2} \Sigma y_{2t} y'_{2t} \end{pmatrix}^{-1} \begin{pmatrix} T^{-1/2} \Sigma u_{1t} \\ T^{-1} \Sigma y_{2t} u_{1t} \end{pmatrix} \\ \Rightarrow &\begin{pmatrix} 1 & \int_0^1 W(r)' dr \begin{pmatrix} \mathbf{0}' \\ P'_{22} \end{pmatrix} \\ [\mathbf{0}, P_{22}] \int_0^1 W(r) dr & [\mathbf{0}, P_{22}] \int_0^1 W(r) W(r)' dr \begin{pmatrix} \mathbf{0}' \\ P'_{22} \end{pmatrix} \end{pmatrix}^{-1} \\ &\times \begin{pmatrix} [\sigma_1, \mathbf{0}'] W(1) \\ [\mathbf{0}, P_{22}] \int_0^1 W(r) [dW(r)]' \begin{pmatrix} \mathbf{0}' \\ P'_{22} \end{pmatrix} \end{pmatrix} \end{aligned}$$

上式中的 $W(r)$ 是由 $[u_{1t}, u'_{2t}]'$ 生成的 n 维标准维纳过程, 将其分解成

$$W(r) = \begin{pmatrix} W_1(r) \\ W_2(r) \end{pmatrix}$$

$W_2(r)$ 为 $(n-1)$ 维标准维纳过程, 可将以上的条件分布写成

$$\begin{aligned} \begin{bmatrix} T^{1/2}(\hat{\alpha}_T - \alpha) \\ T(\hat{\gamma}_T - \gamma) \end{bmatrix} &\Rightarrow \begin{bmatrix} 1 & \left(\int_0^1 \mathbf{W}_2(r)' dr \right) \mathbf{P}'_{22} \\ \mathbf{P}_{22} \int_0^1 \mathbf{W}_2(r) dr & \mathbf{P}_{22} \left(\int_0^1 \mathbf{W}_2(r) \mathbf{W}_2(r)' dr \right) \mathbf{P}'_{22} \end{bmatrix}^{-1} \\ &\times \begin{bmatrix} \sigma_1 W_1(1) \\ \sigma_1 \mathbf{P}_{22} \int_0^1 \mathbf{W}_2(r) (dW_1(r)) \end{bmatrix} = \sigma_1 \begin{bmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \end{bmatrix} \end{aligned}$$

这里的 $[\mathbf{V}_1, \mathbf{V}_2]'$ 为

$$\begin{aligned} \begin{bmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \end{bmatrix} &= \begin{bmatrix} 1 & \left(\int_0^1 \mathbf{W}_2(r)' dr \right) \mathbf{P}'_{22} \\ \mathbf{P}_{22} \int_0^1 \mathbf{W}_2(r) dr & \mathbf{P}_{22} \left(\int_0^1 \mathbf{W}_2(r) \mathbf{W}_2(r)' dr \right) \mathbf{P}'_{22} \end{bmatrix}^{-1} \\ &\times \begin{bmatrix} W_1(r) \\ \mathbf{P}_{22} \int_0^1 \mathbf{W}_2(r) (dW_1(r)) \end{bmatrix} \quad (5.5.4) \end{aligned}$$

由于 u_{1t} 和 u_{2t} 的相互独立性, 维纳过程 $W_1(r)$ 和 $W_2(r)$ 也是互相独立的。根据 $(dW_1(r))$ 的定义, 可以证明, 在给定维纳过程 $W_2(r)$ 时, $\int_0^1 \mathbf{W}_2(r) dW_1(r)$ 有条件分布 (见本章习题):

$$\int_0^1 \mathbf{W}_2(r) dW_1(r) \sim N\left(\mathbf{0}, \int_0^1 \mathbf{W}_2(r) \mathbf{W}_2(r)' dr\right)$$

(5.5.4) 式右侧的第一个矩阵只含有常数矩阵 \mathbf{P}_{22} 和维纳过程 $W_2(r)$ 的泛函, 它们都与 $W_1(1)$ 不相关。因此在给定 $W_2(r)$ 时, $[T^{1/2}(\hat{\alpha}_T - \alpha), T(\hat{\gamma}_T - \gamma)']'$ 有条件极限:

$$\begin{bmatrix} T^{1/2}(\hat{\alpha}_T - \alpha) \\ T(\hat{\gamma}_T - \gamma) \end{bmatrix} \bigg|_{y_{2t}, t=1, 2, \dots, T} \Rightarrow \sigma_1 \begin{bmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \end{bmatrix} \bigg|_{\mathbf{W}_2(r)}$$

$$\begin{aligned}
&= \begin{bmatrix} 1 & \left(\int_0^1 \mathbf{W}_2(r)' dr \right) \mathbf{P}'_{22} \\ \mathbf{P}_{22} \int_0^1 \mathbf{W}_2(r) dr & \mathbf{P}_{22} \left(\int_0^1 \mathbf{W}_2(r) \mathbf{W}_2(r)' dr \right) \mathbf{P}'_{22} \end{bmatrix}^{-1} \\
&\quad \times \begin{bmatrix} \sigma_1 \mathbf{W}_1(1) \\ \sigma_1 \mathbf{P}_{22} \left(\int_0^1 \mathbf{W}_2(r) d\mathbf{W}_1(r) \right) \end{bmatrix} \\
&\sim N \left\{ \begin{bmatrix} 0 \\ \mathbf{0} \end{bmatrix}, \sigma_1^2 \begin{bmatrix} 1 & \left(\int_0^1 \mathbf{W}_2(r)' dr \right) \mathbf{P}'_{22} \\ \mathbf{P}_{22} \int_0^1 \mathbf{W}_2(r) dr & \mathbf{P}_{22} \left(\int_0^1 \mathbf{W}_2(r) \mathbf{W}_2(r)' dr \right) \mathbf{P}'_{22} \end{bmatrix}^{-1} \right\} \\
&\hspace{25em} (5.5.5)
\end{aligned}$$

上式中的分布称为混合正态分布 (mixed normal distribution), 它的协方差矩阵包含维纳过程 $\mathbf{W}_2(r)$ 的泛函。给定 $\mathbf{W}_2(r)$ 和 \mathbf{P}_{22} 的值, (5.5.5) 式有条件正态分布。在以上推导中, u_{1t} 和 u_{2t} 的相互独立是 (5.5.5) 式成立的关键。

利用 (5.5.5) 中的极限, 可进一步考虑二次型 Q 在 $T \rightarrow \infty$ 时的极限分布。以正整数 m 乘 Q , 得

$$\begin{aligned}
mQ &= (\mathbf{R}_1 \hat{\alpha}_T + \mathbf{R}_2 \hat{\mathbf{y}}_T - \mathbf{d})' \left\{ s_T^2 [\mathbf{R}_1, \mathbf{R}_2] \begin{bmatrix} T & \Sigma \mathbf{y}'_{2t} \\ \Sigma \mathbf{y}_{2t} & \Sigma \mathbf{y}_{2t} \mathbf{y}'_{2t} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{R}'_1 \\ \mathbf{R}'_2 \end{bmatrix} \right\}^{-1} \\
&\quad \times (\mathbf{R}_1 \hat{\alpha}_T + \mathbf{R}_2 \hat{\mathbf{y}}_T - \mathbf{d})
\end{aligned}$$

为了叙述方便, 假定原假设 H_0 只涉及同积向量 \mathbf{y} , 而与 α 无关。这样, 在以上表达式中 $\mathbf{R}_1 = \mathbf{0}$, 所以

$$\begin{aligned}
mQ &= [\mathbf{R}_2 \hat{\mathbf{y}}_T - \mathbf{d}]' \left\{ s_T^2 [\mathbf{0}, \mathbf{R}_2] \begin{bmatrix} T & \Sigma \mathbf{y}'_{2t} \\ \Sigma \mathbf{y}_{2t} & \Sigma \mathbf{y}_{2t} \mathbf{y}'_{2t} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{0}' \\ \mathbf{R}'_2 \end{bmatrix} \right\}^{-1} \\
&\quad \times [\mathbf{R}_2 \hat{\mathbf{y}}_T - \mathbf{d}]
\end{aligned}$$

$$= [\mathbf{R}_2 T(\hat{\mathbf{Y}}_T - \mathbf{Y})]' s_T^{-2} \left\{ [\mathbf{0}, \mathbf{R}_2] \begin{pmatrix} 1 & T^{-3/2} \sum \mathbf{y}'_{2t} \\ T^{-3/2} \sum \mathbf{y}_{2t} & T^{-2} \sum \mathbf{y}_{2t} \mathbf{y}'_{2t} \end{pmatrix}^{-1} \right. \\ \left. \times \begin{pmatrix} \mathbf{0}' \\ \mathbf{R}'_2 \end{pmatrix} \right\}^{-1} [\mathbf{R}_2 T(\hat{\mathbf{Y}}_T - \mathbf{Y})]$$

当 $T \rightarrow \infty$ 时, $s_T^2 \xrightarrow{p} \sigma_1^2$, mQ 有条件极限:

$$mQ \Rightarrow [\mathbf{R}_2 \sigma_1 \mathbf{V}_2]' \sigma_1^{-2} \left\{ [\mathbf{0}, \mathbf{R}_2] \right. \\ \times \left. \begin{pmatrix} 1 & \left\{ \int_0^1 \mathbf{W}_2(r)' dr \right\} \mathbf{P}'_{22} \\ \mathbf{P}_{22} \int_0^1 \mathbf{W}_2(r) dr & \mathbf{P}_{22} \left\{ \int_0^1 \mathbf{W}_2(r) \mathbf{W}_2(r)' dr \right\} \mathbf{P}'_{22} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{0}' \\ \mathbf{R}'_2 \end{pmatrix} \right\}^{-1} \\ \times [\mathbf{R}_2 \sigma_1 \mathbf{V}_2]$$

根据(5.5.4)式,随机向量 $[\mathbf{R}_2 \sigma_1 \mathbf{V}_2]$ 有期望为零的条件正态分布,并有协方差矩阵

$$D[\mathbf{R}_2 \sigma_1 \mathbf{V}_2] = \sigma_1^2 [\mathbf{0}, \mathbf{R}_2]$$

$$\times \left\{ \begin{pmatrix} 1 & \left\{ \int_0^1 \mathbf{W}_2(r)' dr \right\} \mathbf{P}'_{22} \\ \mathbf{P}_{22} \int_0^1 \mathbf{W}_2(r) dr & \mathbf{P}_{22} \left\{ \int_0^1 \mathbf{W}_2(r) \mathbf{W}_2(r)' dr \right\} \mathbf{P}'_{22} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{0}' \\ \mathbf{R}'_2 \end{pmatrix} \right\}$$

因此二次型 mQ 有极限分布:

$$mQ \Rightarrow [\mathbf{R}_2 \sigma_1 \mathbf{V}_2]' \{D[\mathbf{R}_2 \sigma_1 \mathbf{V}_2]\}^{-1} [\mathbf{R}_2 \sigma_1 \mathbf{V}_2] \sim \chi^2(m) \quad (5.5.6)$$

其中 m 为 χ^2 分布的自由度,是矩阵 \mathbf{R}_2 的秩。根据以上讨论,若 \mathbf{y}_{1t} 和 \mathbf{y}_{2t} 同积,对同积向量 \mathbf{Y} 的约束条件可用通常的 χ^2 检验法检验,不过这时的 $\chi^2(m)$ 分布只是渐近地成立。(5.5.6)的结果

在 $R_1 \neq 0$ 时也成立。

以上的讨论假定 u_{1t} 和 u_{2t} 不相关, 而且 u_{1t} 和 u_{2t} 分别独立同分布。以下进一步指出, 只要 u_{1t} 和 u_{2t} 互不相关, (5.5.6) 中的结果就能成立, u_{1t} 和 u_{2t} 的独立同分布是非必要的。

假设 $[u_{1t}, u'_{2t}]'$ 可表示为

$$\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \Psi(L)\varepsilon_t = \sum_{s=0}^{\infty} \Psi_s \varepsilon_{t-s}$$

其中矩阵列 $\{\Psi_s\}_{s=0}^{\infty}$ 的每一元素列都绝对可和, 随机向量 $\{\varepsilon_t\}$ 独立同分布, $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = E(\varepsilon_t \varepsilon'_t) = \Omega = PP'$ 。为保证 u_{1t} 和 u_{2t} 不相关, 还假设矩阵滞后多项式 $\Psi(L)$ 和矩阵 P 有分块对角表示形式:

$$\Psi(L) = \begin{bmatrix} \Psi_{11}(L) & 0' \\ 0 & \Psi_{22}(L) \end{bmatrix}$$

$$P = \begin{bmatrix} \sigma_1 & 0' \\ 0 & P_{22} \end{bmatrix}$$

所以, 矩阵 $\Psi(1) \cdot P$ 也是分块对角的, 即

$$\begin{aligned} \Psi(1) \cdot P &= \begin{bmatrix} \Psi_{11}(1) & 0' \\ 0 & \Psi_{22}(1) \end{bmatrix} \begin{bmatrix} \sigma_1 & 0' \\ 0 & P_{22} \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1 \Psi_{11}(1) & 0' \\ 0 & \Psi_{22}(1) \cdot P_{22} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0' \\ 0 & \Lambda_{22} \end{bmatrix} \end{aligned}$$

因此, 只要将 (5.5.5) 式中的 σ_1 和 P_{22} 分别换成 λ_1 和 Λ_{22} 就可得 $T^{1/2}(\hat{\alpha}_T - \alpha)$ 和 $T(\hat{\gamma}_T - \gamma)$ 在更一般情况下的极限:

$$\begin{bmatrix} T^{1/2}(\hat{\alpha}_T - \alpha) \\ T(\hat{\gamma}_T - \gamma) \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & \left\{ \int_0^1 W_2(r)' dr \right\} \Lambda_{22}' \\ \Lambda_{22} \int_0^1 W_2(r) dr & \Lambda_{22} \left\{ \int_0^1 W_2(r) W_2(r)' dr \right\} \Lambda_{22}' \end{bmatrix}^{-1}$$

$$\times (\mathbf{R}_1 \hat{\alpha}_T + \mathbf{R}_2 \hat{\gamma}_T - \mathbf{d}) \xrightarrow{d} \chi^2(m)$$

以上讨论表明, 只要 u_{1t} 和 u_{2t} 不相关, $[T^{1/2}(\hat{\alpha}_T - \alpha), T(\hat{\gamma}_T - \gamma)']'$ 就有正态的条件极限分布。 u_{1t} 和 u_{2t} 的自相关只影响正态分布(5.5.7)的方差矩阵, 但不改变分布的正态性。因此, 在 y_{1t} 和 y_{2t} 同积时, 可以 $\chi^2(m)$ 分布对同积向量的假设作检验。

以下讨论在 u_{1t} 和 u_{2t} 相关时, 对同积向量的假设检验。

5.6 对 u_{1t} 和 u_{2t} 的相关系数的纠正

以前的讨论指出, 在同积系统

$$\begin{aligned} y_{1t} &= \alpha + \gamma' y_{2t} + u_{1t} \\ y_{2t} &= y_{2, t-1} + u_{2t} \end{aligned} \quad (5.6.1)$$

中, 无论 u_{1t} 和 u_{2t} 是否相关, 或 u_{1t} 和 u_{2t} 本身是否自相关, 最小二乘估计 $\hat{\alpha}_T$ 和 $\hat{\gamma}_T$ 是未知参数 α 和 γ 的超一致估计量。尽管如此, 如果 u_{1t} 和 u_{2t} 之间的相关系数不为零, 统计量 $T(\hat{\gamma}_T - \gamma)$ 将不再有正态的条件极限分布。因为这时由 u_{1t} 和 u_{2t} 生成的维纳过程 $W_1(r)$ 和 $W_2(r)$ 不再相互独立, 泛函 $\int_0^1 W_2(r) dW_1(r)$ 不再有正态的条件分布, 因此二次型 mQ 不再有 $\chi^2(m)$ 极限分布。为了仍能用 $\chi^2(m)$ 分布对同积关系作假设检验, 需对二次型 mQ 作修正, 从中去除 u_{1t} 和 u_{2t} 的相关系数。

将 u_{1t} 对由 $u_{2, t-p}, u_{2, t-p+1}, \dots, u_{2t}, u_{2, t-1}, \dots, u_{2, t-p}$ 构成的线性空间作投影, 得表达式

$$u_{1t} = \sum_{s=p}^t l'_s u_{2, t-s} + \tilde{u}_{1t} \quad (5.6.2)$$

其中 p 为一正整数。由于 u_{1t} 和 u_{2t} 之间的相关系数不为零， $I_t(s = -p, \dots, p)$ 中至少有一不全为零的向量。投影残差 $\tilde{u}_{1t}(t = 1, 2, \dots)$ 与向量 $u_{2, t-s}(s = -p, \dots, p)$ 不相关。以 (5.6.2) 式代入 (5.6.1) 中的第一个方程，得

$$\begin{aligned} y_{1t} &= \alpha + \gamma' y_{2t} + \sum_{s=-p}^p I'_s u_{2, t-s} + \tilde{u}_{1t} \\ y_{2t} &= y_{2, t-1} + u_{2t} \end{aligned} \quad (5.6.3)$$

这时 \tilde{u}_{1t} 和 u_{2t} 不再相关。假定 $[\tilde{u}_{1t}, u'_{2t}]'$ 有无穷阶移动平均表示形式：

$$\begin{bmatrix} \tilde{u}_{1t} \\ u_{2t} \end{bmatrix} = \sum_{s=0}^{\infty} \tilde{\Psi}_s \varepsilon_{t-s}$$

$\{s \cdot \tilde{\Psi}_s\}_{s=0}^{\infty}$ 为 $n \times n$ 维矩阵列，其中每一元系列都绝对可和， $\{\varepsilon_t\}$ 独立同分布， $E(\varepsilon_t) = 0$ ， $D(\varepsilon_t) = \Omega = PP'$ 。由于 \tilde{u}_{1t} 与 u_{2t} 不相关，所以矩阵 P 和 $\tilde{\Psi}(L)$ 有分块对角形式：

$$P = \begin{bmatrix} \sigma_1 & 0' \\ 0 & P_{22} \end{bmatrix}$$

和

$$\tilde{\Psi}(L) = \begin{bmatrix} \tilde{\Psi}_{11}(L) & 0' \\ 0 & \tilde{\Psi}_{22}(L) \end{bmatrix} \quad (5.6.4)$$

其中， P_{22} 和 $\tilde{\Psi}_{22}(L)$ 均为 $(m-1) \times (n-1)$ 维矩阵。

定义随机向量

$$z_t = [u'_{2, t-p}, \dots, u'_{2t}, \dots, u'_{2, t+p}]'$$

和参数向量

$$l = [l'_p, l'_{p-1}, \dots, l'_{-p}]'$$

可将系统(5.6.3)的第一个方程改写成以下形式:

$$\begin{aligned} y_{1t} &= \alpha + \gamma' y_{2t} + \sum_{s=-p}^p l'_s u_{2, t-s} + \tilde{u}_{1t} \\ &= l' z_t + \alpha + \gamma' y_{2t} + \tilde{u}_{1t} \end{aligned} \quad (5.6.5)$$

根据以前的讨论, 参数 l 、 α 和 γ 的最小二乘估计有以下极限:

$$\begin{bmatrix} T^{1/2}(\hat{l}_T - l) \\ T^{1/2}(\hat{\alpha}_T - \alpha) \\ T(\hat{\gamma}_T - \gamma) \end{bmatrix} \Rightarrow \begin{bmatrix} G^{-1}h \\ \tilde{\lambda}_{11}V_1 \\ \tilde{\lambda}_{11}V_2 \end{bmatrix}$$

其中 $G = E(z_t z_t')$, $T^{-1/2} \sum_{t=1}^T z_t \tilde{u}_{1t} \Rightarrow h$, $\tilde{\lambda}_{11} = \sigma_1 \tilde{\Psi}_{11}(1)$, 以及

$$\begin{aligned} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} &= \begin{bmatrix} 1 & \left\{ \int_0^1 W_2(r)' dr \right\} \tilde{\Lambda}'_{22} \\ \tilde{\Lambda}_{22} \int_0^1 W_2(r) dr & \tilde{\Lambda}_{22} \left\{ \int_0^1 W_2(r) W_2(r)' dr \right\} \tilde{\Lambda}'_{22} \end{bmatrix}^{-1} \\ &\quad \times \begin{bmatrix} W_1(1) \\ \tilde{\Lambda}_{22} \left\{ \int_0^1 W_2(r) dW_1(r) \right\} \end{bmatrix} \end{aligned}$$

其中 $\tilde{\Lambda}_{22} = \tilde{\Psi}_{22}(1)P_{22}$, $W_1(r)$ 为单变量的标准维纳过程, $W_2(r)$ 为 $(n-1)$ 维标准维纳过程, $W_1(r)$ 和 $W_2(r)$ 相互独立. 这样, 在给定 $W_2(r)$ 时, $[V_1, V_2]'$ 有条件正态分布:

$$\begin{pmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \end{pmatrix} \middle| \mathbf{W}_2(r) \sim N \left\{ \begin{pmatrix} 0 \\ \mathbf{0} \end{pmatrix}, \begin{bmatrix} 1 & \left\{ \int_0^1 \mathbf{W}_2(r)' dr \right\} \bar{\Lambda}'_{22} \\ \bar{\Lambda}_{22} \int_0^1 \mathbf{W}_2(r) dr & \bar{\Lambda}_{22} \left\{ \int_0^1 \mathbf{W}_2(r) \mathbf{W}_2(r)' dr \right\} \bar{\Lambda}'_{22} \end{bmatrix} \right\}^{-1}$$

利用上节的方法,构造二次型 mQ , 当假设 $\mathbf{R}_2 \mathbf{Y} = \mathbf{d}$ 成立时, mQ 有条件极限分布:

$$\begin{aligned} mQ &= (\mathbf{R}_2 \hat{\mathbf{Y}}_T - \mathbf{d})' \left\{ s_T^2 [\mathbf{0}, \mathbf{0}, \mathbf{R}_2] \right. \\ &\quad \times \left. \begin{bmatrix} \Sigma \mathbf{z}_t \mathbf{z}_t' & \Sigma \mathbf{z}_t & \Sigma \mathbf{z}_t \mathbf{y}_{2t}' \\ \Sigma \mathbf{z}_t' & T & \Sigma \mathbf{y}_{2t}' \\ \Sigma \mathbf{y}_{2t} \mathbf{z}_t' & \Sigma \mathbf{y}_{2t} & \Sigma \mathbf{y}_{2t} \mathbf{y}_{2t}' \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{0}' \\ \mathbf{0}' \\ \mathbf{R}_2' \end{bmatrix} \right\}^{-1} (\mathbf{R}_2 \hat{\mathbf{Y}}_T - \mathbf{d}) \\ &\Rightarrow (\tilde{\lambda}_{11}^2 / \text{plims } s_T^2) [\mathbf{R}_2 \mathbf{V}_2]' \left\{ [\mathbf{0}, \mathbf{R}_2] \right. \\ &\quad \times \left. \begin{bmatrix} 1 & \left\{ \int_0^1 \mathbf{W}_2(r)' dr \right\} \bar{\Lambda}'_{22} \\ \bar{\Lambda}_{22} \int_0^1 \mathbf{W}_2(r) dr & \bar{\Lambda}_{22} \left\{ \int_0^1 \mathbf{W}_2(r) \mathbf{W}_2(r)' dr \right\} \bar{\Lambda}'_{22} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{0}' \\ \mathbf{R}_2' \end{bmatrix} \right\}^{-1} \\ &\quad \times [\mathbf{R}_2 \mathbf{V}_2] \end{aligned}$$

与上节的(5.5.6)式比较,立即可得

$$mQ \xrightarrow{d} (\tilde{\lambda}_{11}^2 / \text{plims } s_T^2) \chi^2(m) \quad (5.6.6)$$

由于 \tilde{u}_{1t} 自相关, $\tilde{\lambda}_{11}^2$ 一般不等于 $\text{plims } s_T^2$ 。以下,我们用两种不同的方法构造 $\tilde{\lambda}_{11}^2$ 的一致估计量。

首先,对回归模型

$$y_{1t} = \alpha + \gamma' y_{2t} + \sum_{i=-p}^p \beta_i \Delta y_{2, t-i} + \bar{u}_{1t}$$

中的参数作最小二乘估计,并以 \hat{e}_t 表示回归残差,再以 $\hat{e}_t (t = 1, 2, \dots, T)$ 构造 P 阶自回归模型

$$\hat{e}_t = \phi_1 \hat{e}_{t-1} + \phi_2 \hat{e}_{t-2} + \dots + \phi_p \hat{e}_{t-p} + \eta_t \quad (5.6.7)$$

并对参数作最小二乘估计 $\hat{\phi}_1, \dots, \hat{\phi}_p$ 。因为

$$\bar{\lambda}_{11} = \sigma_1 \bar{\Psi}_{11}(1)$$

所以可用最小二乘估计 $\hat{\phi}_1, \dots, \hat{\phi}_p$ 构造参数 $\bar{\lambda}_{11}$ 的一致估计量

$$\hat{\bar{\lambda}}_{11} = \frac{\hat{\sigma}_1^2}{1 - \hat{\phi}_1 - \dots - \hat{\phi}_p}$$

其中, $\hat{\sigma}_1^2 = (T-p)^{-1} \sum_{t=p+1}^T \hat{\eta}_t^2$, $\hat{\eta}_t$ 为 (5.6.7) 的回归残差。当样

本量 $T \rightarrow \infty$ 时, $\hat{\bar{\lambda}}_{11} \xrightarrow{p} \bar{\lambda}_{11}$ 。以 $\hat{\bar{\lambda}}_{11}$ 代替 mQ 中的 s_T , 可得统计量 \bar{mQ} , 它有极限分布

$$\begin{aligned} \bar{mQ} &= (R_2 \hat{Y}_T - d)' \left\{ \hat{\bar{\lambda}}_{11}^2 [0, \mathbf{0}, R_2] \right. \\ &\quad \times \left[\begin{array}{ccc} \Sigma z_t z_t' & \Sigma z_t & \Sigma z_t y_{2t}' \\ \Sigma z_t' & T & \Sigma y_{2t}' \\ \Sigma y_{2t} z_t' & \Sigma y_{2t} & \Sigma y_{2t} y_{2t}' \end{array} \right]^{-1} \left[\begin{array}{c} 0' \\ \mathbf{0}' \\ R_2' \end{array} \right] \left. \right\}^{-1} (R_2 \hat{Y}_T - d) \\ &\xrightarrow{d} \left[\frac{\bar{\lambda}_{11}^2}{\text{plim } \hat{\bar{\lambda}}_{11}^2} \right] \chi^2(m) = \chi^2(m) \end{aligned} \quad (5.6.8)$$

形式

$$\begin{aligned} y_{1t} &= \alpha + \gamma' y_{2t} + u_{1t} \\ \Delta y_{2t} &= u_{2t} \end{aligned} \quad (5.7.1)$$

$(u_{1t}, u_{2t})'$ 有表示形式

$$\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \Psi(L) \varepsilon_t$$

其中, $\{\varepsilon_t\}$ 独立同分布, $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = E(\varepsilon_t \varepsilon_t') = \Omega = PP'$ 。
令

$$\Lambda = \Psi(1)P$$

$$\Sigma = \Lambda \Lambda' = \Psi(1)PP' \Psi(1) = \begin{bmatrix} \Sigma_{11} & \Sigma'_{21} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \quad (5.7.2)$$

矩阵 Σ 为非异正定, 可表示为

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma'_{21} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} = \sum_{v=-\infty}^{\infty} \begin{bmatrix} E(u_{1t} u_{1, t-v}) & E(u_{1t} u_{2, t-v}) \\ E(u_{2t} u_{1, t-v}) & E(u_{2t} u_{2, t-v}) \end{bmatrix}$$

系统 (5.7.1) 中的参数 α 和 γ 的最小二乘估计的极限由定理 5.1 给出。当子矩阵 Σ_{21} 为零时, 由 u_{1t} 和 u_{2t} 生成的维纳过程 $W_1(r)$ 和 $W_2(r)$ 互相独立, 统计量 $T(\hat{\gamma}_T - \gamma)$ 有正态的条件极限分布。一般来说, 若 Σ_{21} 不为零, $T(\hat{\gamma}_T - \gamma)$ 的极限分布是非标准的。以下介绍的改进的 OLS 方法, 旨在克服这一困难。

若 Σ_{21} 不为零, 构造 $\Sigma'_{21} \Sigma_{22}^{-1} \Delta y_{2t}$, 将其从 y_{1t} 中减去, 可得

$$y_{1t}^+ = y_{1t} - \Sigma'_{21} \Sigma_{22}^{-1} \Delta y_{2t}$$

修改后的同积系统为

$$y_{1t}^+ = \alpha + \gamma' y_{2t} + u_{1t}^+$$

$$\Delta y_{2t} = u_{2t}$$

其中的 u_{1t} 为

$$u_{1t}^- = u_{1t} - \Sigma'_{21} \Sigma_{22}^{-1} \Delta y_{2t}$$

随机向量 $[u_{1t}^+, u_{2t}']'$ 满足

$$\begin{bmatrix} u_{1t}^+ \\ u_{2t}' \end{bmatrix} = L' \begin{bmatrix} u_{1t} \\ u_{2t}' \end{bmatrix}$$

其中的矩阵 L' 为

$$L' = \begin{bmatrix} 1 & -\Sigma'_{21} \Sigma_{22}^{-1} \\ \mathbf{0} & I_{n-1} \end{bmatrix} = \begin{bmatrix} L'_1 \\ L'_2 \end{bmatrix}$$

对 y_{1t}^+ 中的参数作估计, 得最小二乘估计 $\hat{\alpha}_T^+$ 和 $\hat{\gamma}_T^+$

$$\begin{bmatrix} \hat{\alpha}_T^+ \\ \hat{\gamma}_T^+ \end{bmatrix} = \begin{bmatrix} T & \Sigma y'_{2t} \\ \Sigma y_{2t} & \Sigma y_{2t} y'_{2t} \end{bmatrix}^{-1} \begin{bmatrix} \Sigma y_{1t}^+ \\ \Sigma y_{2t} y_{1t}^+ \end{bmatrix}$$

由定理 5.1, 立即可得

$$\begin{aligned} & \begin{bmatrix} T^{1/2}(\hat{\alpha}^+ - \alpha) \\ T(\hat{\gamma}^+ - \gamma) \end{bmatrix} = \begin{bmatrix} 1 & T^{-3/2} \Sigma y'_{2t} \\ T^{-3/2} \Sigma y_{2t} & T^{-2} \Sigma y_{2t} y'_{2t} \end{bmatrix}^{-1} \begin{bmatrix} T^{-1/2} \Sigma u_{1t}^+ \\ T^{-1} \Sigma y_{2t} u_{1t}^+ \end{bmatrix} \\ & \Rightarrow \begin{bmatrix} 1 & \left\{ \int_0^1 W(r)' dr \right\} \Lambda' L_2 \\ L_2' \Lambda \int_0^1 W(r) dr & L_2' \Lambda \left\{ \int_0^1 W(r) W(r)' dr \right\} \Lambda' L_2 \end{bmatrix}^{-1} \\ & \quad \times \begin{bmatrix} L_1' \Lambda W(1) \\ L_2' \Lambda \left\{ \int_0^1 W(r) [dW(r)]' \right\} \Lambda' L_1 + \delta^- \end{bmatrix} \end{aligned}$$

其中, $W(r)$ 为 n 维标准维纳过程, 参数 δ^+ 有表示式

$$\begin{aligned}\delta^+ &= \sum_{v=0}^{\infty} E(u_{2t} u_{1, t+v}) = \sum_{v=0}^{\infty} E\{u_{2t} [u_{1, t+v} - \Sigma'_{21} \Sigma_{22}^{-1} u_{2, t+v}]\} \\ &= \sum_{v=0}^{\infty} E\{u_{2t} [u_{1, t+v}, u'_{2, t+v}]\} \begin{bmatrix} 1 \\ -\Sigma_{22}^{-1} \Sigma_{21} \end{bmatrix}\end{aligned}$$

定义维纳过程 $B(r)$ 为

$$B(r) = \begin{bmatrix} L_1' \\ L_2' \end{bmatrix} \Lambda W(r) = \begin{bmatrix} L_1' \Lambda W(r) \\ L_2' \Lambda W(r) \end{bmatrix} = \begin{bmatrix} B_1(r) \\ B_2(r) \end{bmatrix}$$

$B_1(r)$ 和 $B_2(r)$ 分别为 1 维和 $n-1$ 维的维纳过程。因为

$$\begin{aligned}E\{B(1)B(1)'\} &= \begin{bmatrix} L_1' \\ L_2' \end{bmatrix} \Lambda \Lambda' \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -\Sigma'_{21} \Sigma_{22}^{-1} \\ \mathbf{0} & I_{n-1} \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \Sigma'_{21} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}' \\ -\Sigma_{22}^{-1} \Sigma_{21} & I_{n-1} \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1^{+2} & \mathbf{0}' \\ \mathbf{0} & \Sigma_{22} \end{bmatrix}\end{aligned}$$

所以维纳过程 $B_1(r)$ 和 $B_2(r)$ 是互相独立的。上式中的 σ_1^{+2} 为

$$\sigma_1^{+2} = \Sigma_{11} - \Sigma'_{21} \Sigma_{22}^{-1} \Sigma_{21}$$

令 $\Sigma_{22} = P_{22}^+ P_{22}^{+'}$, 可将维纳过程 $B(r)$ 写成以下形式:

$$B(r) = \begin{bmatrix} B_1(r) \\ B_2(r) \end{bmatrix} = \begin{bmatrix} \sigma_1^+ & \mathbf{0}' \\ \mathbf{0} & P_{22}^+ \end{bmatrix} \begin{bmatrix} W_1^+(r) \\ W_2^+(r) \end{bmatrix}$$

其中的 $W_1^+(r)$ 和 $W_2^+(r)$ 分别为 1 维和 $n-1$ 维互相独立的标准维纳过程。由此, 可进一步将 $[T^{1/2}(\hat{\alpha}_T^+ - \alpha), T(\hat{\gamma}_T^+ - \gamma)']'$ 的极限写成下列形式:

给出。式中的 \hat{u}_{1t} 是 y_{1t} 对常数项和 y_{2t} 的回归残差, $\hat{u}_{2t} = \Delta y_{2t}$ 。由于参数 δ^+ 有表示式

$$\begin{aligned}\delta^+ &= \sum_{v=0}^{\infty} E \left[\begin{pmatrix} u_{1t} & u'_{2, t-v} \\ u_{2t} & u'_{2, t-v} \end{pmatrix} \right]' \begin{pmatrix} 1 \\ -\Sigma_{22}^{-1} \Sigma_{21} \end{pmatrix} \\ &= \sum_{v=0}^{\infty} \begin{pmatrix} r_v^{12} \\ r_v^{22} \end{pmatrix}' \begin{pmatrix} 1 \\ -\Sigma_{22}^{-1} \Sigma_{21} \end{pmatrix}\end{aligned}$$

可以构造参数 δ^+ 的一致估计量 $\hat{\delta}^+$ 为

$$\hat{\delta}^+ = \sum_{v=0}^T \left\{ 1 - \frac{v}{q+1} \right\} (\hat{r}_v^{12'}, \hat{r}_v^{22'}) \begin{pmatrix} 1 \\ -\hat{\Sigma}_{22}^{-1} \hat{\Sigma}_{21} \end{pmatrix}$$

然后再将 $T\hat{\delta}^+$ 从估计量 \hat{y}_T 中减去, 就可得充分改进的最小二乘估计 $\hat{\alpha}_T^{++}$ 和 \hat{y}_T^{++} :

$$\begin{pmatrix} \hat{\alpha}_T^{++} \\ \hat{y}_T^{++} \end{pmatrix} = \begin{pmatrix} T & \Sigma y'_{2t} \\ \Sigma y_{2t} & \Sigma y_{2t} y'_{2t} \end{pmatrix}^{-1} \begin{pmatrix} \Sigma \hat{y}_{1t} \\ \Sigma y_{2t} \hat{y}_{1t}^+ - T\delta^+ \end{pmatrix}$$

其中的 \hat{y}_{1t}^+ 为

$$\hat{y}_{1t}^+ = y_{1t} - \hat{\Sigma}'_{21} \hat{\Sigma}_{22}^{-1} \Delta y_{2t}$$

总结起来, 构造充分改进的最小二乘估计 $\hat{\alpha}_T^{++}$ 和 \hat{y}_T^{++} 需要采取以下步骤: 先以 y_{1t} 对常数和 y_{2t} 作回归, 取得残差 \hat{u}_{1t} ; 然后用 \hat{u}_{1t} 和 \hat{u}_{2t} 构造估计量 \hat{r}_v 和 $\hat{\Sigma}$; 对 y_{1t} 作调整, $\hat{y}_{1t}^+ = y_{1t} - \hat{\Sigma}'_{21} \hat{\Sigma}_{22}^{-1} \Delta y_{2t}$; 计算 $\hat{\delta}^+$; 最后用 \hat{y}_{1t}^+ 对常数和 y_{2t} 作回归, 并将 $T\hat{\delta}^+$ 从 $(\Sigma y_{2t} \hat{y}_{1t}^+)$ 中减去。

根据上述讨论和定理 5.1 的结果, $\hat{a}_{T^{++}}$ 和 $\hat{\gamma}_{T^{++}}$ 均为一致估计量, $T^{1/2}(\hat{a}_{T^{++}} - a)$ 和 $T(\hat{\gamma}_{T^{++}} - \gamma)$ 有极限:

$$\begin{aligned} \begin{bmatrix} T^{1/2}(\hat{a}_{T^{++}} - a) \\ T(\hat{\gamma}_{T^{++}} - \gamma) \end{bmatrix} &= \begin{bmatrix} 1 & T^{-3/2} \Sigma y'_{2t} \\ T^{-3/2} \Sigma y_{2t} & T^{-2} \Sigma y_{2t} y'_{2t} \end{bmatrix}^{-1} \begin{bmatrix} T^{-1/2} \Sigma \hat{u}^+_{it} \\ T^{-1} \Sigma y_{2t} \hat{u}^+_{it} - \hat{\delta} \end{bmatrix} \\ &\Rightarrow \sigma_1^+ \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \end{aligned} \quad (5.7.4)$$

其中的 V_1 和 V_2 为

$$\begin{aligned} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} &= \begin{bmatrix} 1 & \left\{ \int_0^1 W_2^+(r)' dr \right\} P_{22}^{+'} \\ P_{22}^+ \int_0^1 W_2^+(r) dr & P_{22}^+ \left\{ \int_0^1 W_2^+(r) W_2^+(r)' dr \right\} P_{22}^{+'} \end{bmatrix}^{-1} \\ &\quad \times \begin{bmatrix} W_1^+(1) \\ P_{22}^+ \left\{ \int_0^1 W_2^+(r) [dW_1^+(r)] \right\} \end{bmatrix} \end{aligned}$$

$W_1^+(r)$ 和 $W_2^+(r)$ 为互相独立的维纳过程, 在 $W_2^+(r)$ 给定时, $[V_1, V_2]'$ 有期望为零的正态条件分布, 记为

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \bigg| W_2^+(r) \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, H^{-1} \right)$$

矩阵 H 为

$$H = \begin{bmatrix} 1 & \left\{ \int_0^1 W_2^+(r)' dr \right\} P_{22}^{+'} \\ P_{22}^+ \int_0^1 W_2^+(r) dr & P_{22}^+ \left\{ \int_0^1 W_2^+(r) W_2^+(r)' dr \right\} P_{22}^{+'} \end{bmatrix} \quad (5.7.5)$$

$[T^{1/2}(\hat{a}_{T^{++}} - a), T(\hat{\gamma}_{T^{++}} - \gamma)']'$ 的正态极限分布, 是进一

步对同积向量作假设检验的基础,由此构造的二次型 mQ 就有 $\chi^2(m)$ 的极限分布。(5.7.4)式中的未知参数 σ_1^+ 可由下式

$$\hat{\sigma}_1^+ = \sqrt{\hat{\Sigma}_{11} - \hat{\Sigma}_{21}' \hat{\Sigma}_{22}^{-1} \hat{\Sigma}_{21}}$$

一致地估计。其中的估计量 $\hat{\Sigma}_{11}$ 、 $\hat{\Sigma}_{21}$ 和 $\hat{\Sigma}_{22}$ 由 (5.7.3) 式给出。基于 $[T^{1/2}(\hat{\alpha}_T^{++} - \alpha), T(\hat{\gamma}_T^{+-} - \gamma)']'$ 的极限分布,可以构造检验假设 $H_0: R\gamma = d$ 的二次型 mQ 为

$$mQ = (R\hat{\gamma}_T^{+-} - d)' \left\{ (\hat{\sigma}_1^+)^2 [0, R] \begin{bmatrix} T & \Sigma y'_{2t} \\ \Sigma y_{2t} & \Sigma y_{2t} y'_{2t} \end{bmatrix}^{-1} \begin{bmatrix} 0' \\ R' \end{bmatrix} \right\}^{-1} \\ \times (R\hat{\gamma}_T^{++} - d)$$

当假设 $H_0: R\gamma = d$ 为真时, mQ 有 $\chi^2(m)$ 的条件极限分布:

$$mQ \Rightarrow (\sigma_1^+)^2 (RV_2)' \left\{ (\sigma_1^+)^2 [0, R] H^{-1} \begin{bmatrix} 0' \\ R' \end{bmatrix} \right\}^{-1} (RV_2) \sim \chi^2(m)$$

值得注意的是,以上的分析并不依赖于 u_{1t} 和 u_{2t} 的具体的参数结构,所以从这意义上说,改进的 OLS 方法是一种非参数的方法。

5.8 本章小结

本章讨论了同积过程的参数估计和对同积向量的假设检验,讨论基于同积向量的最小二乘估计。

本章指出,若随机过程 y_{1t} 和 y_{2t} 同积,它们之间的同积向量的最小二乘估计是超一致收敛的。定理 5.1 给出了最小二乘估计的极限和极限分布。定理 5.1 的一个重要结论是,最小二乘估计 $\hat{\alpha}_T$ 和 $\hat{\gamma}_T$ 的超一致性不受到 u_{1t} 和 u_{2t} 之间的相关性的影响。

这种相关性只影响 $T^{1/2}(\hat{\alpha}_T - \alpha)$ 和 $T(\hat{\gamma}_T - \gamma)$ 的极限分布. 如果 u_{1t} 和 u_{2t} 之间的相关系数不为零, 它们的条件极限分布将不是正态的。

本章介绍了检验同积关系的 ADF 和 PP 方法。这两种方法的原假设都是 y_{1t} 和 y_{2t} 不同积。如果残差的单位根假设被拒绝, 则拒绝 y_{1t} 和 y_{2t} 不同积的假设, 也即接受 y_{1t} 和 y_{2t} 同积。

在给定 y_{1t} 和 y_{2t} 同积的情况下, 可以进一步对同积向量的约束条件进行假设检验。本章的讨论表明, 这些检验可以用通常的 t_T 和 F_T 统计量进行, 这和在稳定过程时的情况一样。只要 u_{1t} 和 u_{2t} 之间不相关, 即使它们本身分别有自相关, t_T 和 F_T 检验法也是一致的检验。

当 u_{1t} 和 u_{2t} 相关时, 本章介绍了两种改进的方法, 使得改进后的统计量 $T(\hat{\gamma}_T - \gamma)$ 仍有正态的条件极限分布, 从而二次型 mQ 有 $\chi^2(m)$ 的条件极限分布。

本章的讨论基于 y_{1t} 和 y_{2t} 之间只存在一个同积向量的情况, 若它们之间存在一个以上的同积关系, 本章的方法有其局限性。这些弱点将在下章讨论的系统方法中加以改善。

习 题

1. 设 $W_1(r)$ 为一维的标准维纳过程, 随机微分 $dW_1(r)$ 定义为

$$dW_1(r) = W_1(r + dt) - W_1(r)$$

因此

$$dW_1(r) \sim N(0, dt)$$

若 n 维标准维纳过程 $W_2(r)$ 与 $W_1(r)$ 互相独立, 根据 1.6.2 节中随机积分的定义, 证明:

$$\int_0^1 W_2(r) dW_1(r) \sim N\left(0, \int_0^1 W_2(r) W_2(r)' dr\right)$$

2. 若随机向量 $[u_{1t}, u_{2t}]'$ 有表示形式

$$\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \sum_{s=0}^{\infty} \Psi_s \epsilon_{t-s}$$

其中 $\{\epsilon_t\}$ 为独立同分布, $E(\epsilon_t) = \mathbf{0}$, $D(\epsilon_t) = \Omega = PP'$ 。验证:若

$$P = \begin{pmatrix} \sigma_1 & \mathbf{0}' \\ \mathbf{0} & P_{22} \end{pmatrix}, \quad \Psi(L) = \begin{pmatrix} \Psi_{11}(L) & \mathbf{0}' \\ \mathbf{0} & \Psi_{22}(L) \end{pmatrix}$$

则有 $E(u_{1t} u_{2t}) = \mathbf{0}$ 。

3. 比较 5.6 节中介绍的纠正 u_{1t} 和 u_{2t} 之间相关系数的两种方法, 讨论它们各自的优缺点, 并举例说明这两种方法所适用的情况。

第 6 章 同积系统的最大似然方法

6.1 简介

前两章介绍的处理同积过程的方法,主要基于对同积向量的最小二乘估计,在此基础上对各种假设作检验。这种方法利用了同积系统的三角表示形式:

$$\begin{aligned}y_{1t} &= \alpha + Y' y_{2t} + u_{1t} \\ y_{2t} &= y_{2, t-1} + u_{2t}\end{aligned}\tag{6.1.1}$$

这种表示形式的主要特点是:(1)系统中只存在一个同积关系,它由系统的第一个方程表示;(2)同积关系中不包含动态结构, y_{1t} 和 y_{2t} 的滞后项都被压缩在随机干扰 u_{1t} 和 u_{2t} 中。这种分析方法简单明确,应用方便,在许多场合下很有效,特别是当系统中只含有唯一的同积关系,并且事先能确定对哪一变量进行规范时。但是,在现实中,以上两个条件往往难以满足。在一个复杂的经济系统中,许多相关的变量交织在一起,形成多个同积关系系统,而且常常难以从中分离出一个明确的同积关系,更难以事先确定对哪一变量进行规范。

本章介绍对同积系统作全面分析的最大似然方法。这一方法对系统中所有独立的同积关系作总体分析,不事先假定系统

中同积关系的个数,也无需确定对哪一个变量作规范,这与前面介绍的单方程的方法形成明显的对照。本章的讨论基于约翰逊(Johansen, 1988)的工作,并采用他的符号和表达形式,这与前几章的约定有些不同,但不会引起误解。

6.2 同积过程与误差修正过程

令 x_t 为一 n 维的 $I(1)$ 随机过程,它的一阶差分 Δx_t 为 $I(0)$ 过程,并有表示形式

$$\Delta x_t = C(L)(m + \varepsilon_t) \quad (6.2.1)$$

这里的 m 为一 n 维的常数向量, $\{\varepsilon_t\}$ 为独立同分布的正态向量, $\varepsilon_t \sim N(0, \Omega)$, Ω 为 $n \times n$ 维协方差矩阵。常数向量 m 表示随机过程中的时间趋势,当 $m = 0$ 时, x_t 为不含时间趋势的单位根过程。随机向量 x_t 的各个分量之间可以存在一个以上独立的同积关系,它们的个数在分析过程中确定。 $C(L)$ 为矩阵的滞后多项式,有表示式

$$C(L) = \sum_{i=0}^{\infty} C_i L^i$$

其中, $C_i (i = 0, 1, 2, \dots)$ 为 $n \times n$ 维矩阵,并有 $C_0 = I_n$ 。矩阵 $C(1)$ 称为 $C(L)$ 的总积累效应,且满足

$$|C(1)| = \left| I_n + \sum_{i=1}^{\infty} C_i \right| < \infty$$

所以矩阵 $C(1)$ 中的每一个元素都有限。利用 $C(1)$ 将滞后多项式 $C(L)$ 分解为

$$C(L) = C(1) + (1 - L)C^*(L) \quad (6.2.2)$$

这里, $C^*(L) = \sum_{i=0}^{\infty} C_i^* L^i$, $C_i^* = - \sum_{j=i+1}^{\infty} C_j$, 显然, $C_0^* = I_n - C(1)$ 。利用分解式(6.2.2)可将 Δx_i 改写成

$$\begin{aligned}\Delta x_i &= C(L)(m + \varepsilon_i) = C(1)m + C(L)\varepsilon_i \\ &= C(1)m + C(1)\varepsilon_i + C^*(L)\Delta \varepsilon_i\end{aligned}$$

或者

$$\Delta x_i - \mu = C(1)\varepsilon_i + C^*(L)\Delta \varepsilon_i \quad (6.2.3)$$

式中 $\mu = C(1)m$ 。以(6.2.3)式表示同积系统 x_i 有许多方便之处,为进一步研究它的结构,先作定义。

定义 6.1: 令 P 为 $n \times q$ 维列满秩矩阵, $rk(P) = q < n$, $n \times (n - q)$ 维矩阵 P_{\perp} 称为矩阵 P 的正交补矩阵,若

$$(i) P'_{\perp} P = 0; \quad (ii) rk(P_{\perp}) = n - q$$

下面对滞后多项式 $C(z) = \sum_{i=0}^{\infty} C_i z^i$ 作以下假设:

假设一:无穷阶多项式 $C(z) = \sum_{i=0}^{\infty} C_i z^i$ 的所有的根都不在单位圆内,即若 $|C(z)| = 0$, 则有 $|z| \geq 1$ 。

假设二:若 $n \times n$ 维矩阵 $C(1)$ 的秩为 $n - r$, 其中 r 为 x_i 中独立的同积关系的个数,则存在 $n \times (n - r)$ 维矩阵 A 和 B , 并有 $rk(A) = rk(B) = n - r$, 使得 $C(1) = AB'$ 。

假设三:以 A_{\perp} 和 B_{\perp} 分别表示矩阵 A 和 B 的正交补阵,那么 $A'_{\perp} C^*(1) B_{\perp}$ 为满秩的 $n \times n$ 维矩阵。

以 A' 同时左乘(6.2.3)式的两侧,得

$$\begin{aligned}A'(\Delta x_i - \mu) &= A'C(1)\varepsilon_i + A'C^*(L)\Delta \varepsilon_i \\ &= A'AB'\varepsilon_i + A'C^*(L)\Delta \varepsilon_i\end{aligned} \quad (6.2.4)$$

$$\begin{aligned}
A'(\Delta x_t - \mu) &= A'AB'\varepsilon_t + A'C^*(L)\Delta\varepsilon_t \\
&= A'AB'(B\bar{B}' + B_\perp\bar{B}'_\perp)\varepsilon_t + A'C^*(L)(B_\perp y_t + B\Delta z_t) \\
&= (A'A)(B'B)z_t + A'C^*(L)B\Delta z_t + A'C^*(L)B_\perp y_t \quad (6.2.7)
\end{aligned}$$

$$\begin{aligned}
A'_\perp(\Delta x_t - \mu) &= A'_\perp C^*(L)\Delta\varepsilon_t \\
&= A'_\perp C^*(L)B\Delta z_t + A'_\perp C^*(L)B_\perp y_t \quad (6.2.8)
\end{aligned}$$

定义矩阵 $H(z)$ 为

$$H(z) = \begin{bmatrix} (A'A)(B'B) + A'C^*(z)B(1-z) & A'C^*(z)B_\perp \\ A'_\perp C^*(z)B(1-z) & A'_\perp C^*(z)B_\perp \end{bmatrix}$$

就可将(6.2.7)式和(6.2.8)式写成系统形式

$$[A, A_\perp]'(\Delta x_t - \mu) = H(L)[z'_t, y'_t]' \quad (6.2.9)$$

根据假设三, $n \times n$ 维矩阵 $A'_\perp C^*(1)B_\perp$ 满秩, 因此当 $z = 1$ 时, 矩阵 $H(1)$ 的行列式为

$$\begin{aligned}
|H(1)| &= |(A'A)(B'B)(A'_\perp C^*(1)B_\perp)| \\
&= |A'A| |B'B| |A'_\perp C^*(1)B_\perp| \neq 0
\end{aligned}$$

因此, 多项式 $|H(z)|$ 不具有单位根。当 $|z| > 1$ 时, 根据 $C(z)$ 和 $C^*(z)$ 的关系, 可得

$$C^*(z) = \frac{1}{1-z} (C(z) - C(1))$$

以此代入矩阵 $H(z)$, 并注意 $C(1) = AB'$, $A'_\perp A = B'_\perp B_\perp = 0$, 则有

$$H(z) = [A, A_\perp]'C(z)[B, B_\perp(1-z)^{-1}]$$

当 $|z| > 1$ 时, 上式中的矩阵 $[A, A_\perp]'$ 和 $[B, B_\perp(1-z)^{-1}]$ 都非异, 因此行列式 $|H(z)|$ 等于零, 当且仅当 $|C(z)| = 0$ 时。由

$$F(1) = B_{\perp} (A'_{\perp} C^*(1) B_{\perp})^{-1} A'_{\perp}$$

另一方面, 因为 $\mu = C(1)m = (AB')m$, 所以

$$F(1)\mu = B_{\perp} (A'_{\perp} C^*(1) B_{\perp})^{-1} A'_{\perp} AB'm = 0$$

所以, 最终可得

$$F(L)x_t = x_0 + \varepsilon_t \quad (6.2.10)$$

这里 $F(L)$ 为矩阵的滞后多项式。这表明, 任何同积的随机向量都有自回归表示形式。这一结果对研究同积过程的一般特性有重要意义。从以上讨论中不难看出, $n \times n$ 维矩阵 $F(1)$ 是不满秩的。因为在 $F(1) = B_{\perp} (A'_{\perp} C^*(1) B_{\perp})^{-1} A'_{\perp}$ 中, B_{\perp} 和 A_{\perp} 的秩都为 $r < n$, 所以 $rk(F(1)) = r$ 。这样, 可将 $n \times n$ 维矩阵 $F(1)$ 分解为

$$F(1) = \alpha\beta'$$

这里, 矩阵 α 和 β 均为 $n \times r$ 维, 且 $rk(\alpha) = rk(\beta) = r$ 。矩阵 $F(1)$ 的奇异性, 使得系统 (6.2.10) 有误差修正形式 (ECM) (参见第 4 章的讨论):

$$\begin{aligned} \Delta x_t &= D(L)\Delta x_{t-1} + x_0 + F(1)x_{t-k} + \varepsilon_t \\ &= D(L)\Delta x_{t-1} + x_0 + \alpha\beta'x_{t-k} + \varepsilon_t \end{aligned} \quad (6.2.11)$$

根据第 4 章的讨论, 矩阵 β 给出了 r 个独立的同积向量, 矩阵 α 给出了 r 个同积关系 ($\beta'x_{t-k}$) 对 Δx_t 各分量的贡献。应该指出, 分解式 $F(1) = \alpha\beta'$ 并非唯一的, 因此 β 中的 r 个独立的同积向量也非唯一决定的, 因为利用任何非异的 $r \times r$ 维矩阵 P , 可将 $\alpha\beta'$ 改写成

$$\alpha\beta' = (\alpha P)(P^{-1}\beta') = \alpha^*\beta^{*'}.$$

从而矩阵 β^* 同样给出了 x_t 中的 r 个独立的同积向量。尽管 β 和

β^* 一般不同, 它们之间存在一对一的非异映照, 所以它们生成的线性空间是一致的。由矩阵 β 和 β^* 的 r 个列向量生成的线性空间, 称为随机过程 x_t 的同积空间。

约翰逊方法(Johansen method)对同积系统作整体的分析, 对误差修正形式中的参数作联合的最大似然估计。作为理论准备, 以下先介绍典型相关的概念和方法。

6.3 典型相关

本节对典型相关(canonical correlation)方法作简明的介绍, 这对理解同积系统的结构和约翰逊系统分析方法有重要意义。分别以 y_t 和 x_t 表示 n_1 维和 n_2 维稳定的随机向量, 并假设它们的期望值都为零。以 Σ 表示随机向量 $[y_t', x_t']'$ 的协方差矩阵:

$$\Sigma = \begin{bmatrix} E(y_t y_t') & E(y_t x_t') \\ E(x_t y_t') & E(x_t x_t') \end{bmatrix} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

其中, $\Sigma_{12} = \Sigma'_{21}$ 是 y_t 和 x_t 的协方差。令 $n = \min(n_1, n_2)$, 构造线性变换:

$$\eta_t = G_1' y_t$$

$$\xi_t = G_2' x_t$$

G_1 和 G_2 分别为 $n_1 \times n$ 维和 $n_2 \times n$ 维矩阵。经变换后的随机向量 η_t 和 ξ_t 满足条件:

(a) 随机向量 η_t 和 ξ_t 的所有的分量都互不相关, η_t 和 ξ_t 都有单位协方差矩阵, 即

$$D(\eta_t) = E(\eta_t \eta_t') = G_1' E(y_t y_t') G_1 = G_1' \Sigma_{11} G_1 = I_n$$

$$D(\xi_t) = E(\xi_t \xi_t') = G_2' E(x_t x_t') G_2 = G_2' \Sigma_{22} G_2 = I_n$$

(b) 若 $i \neq j$, η_i 的第 i 个分量与 ξ_i 的第 j 个分量不相关; 若 $i = j$, 则有

$$E(\eta_i \xi_i) = r_i$$

从而有

$$E(\eta_i \xi_i') = G_1' E(y_i x_i') G_2 = G_1' \Sigma_{12} G_2 = R$$

R 为 $n \times n$ 维对角矩阵

$$R = \begin{pmatrix} r_1 & 0 & \cdots & 0 \\ 0 & r_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_n \end{pmatrix} \quad (6.3.1)$$

(c) 调整矩阵 G_1 和 G_2 中各行的前后次序, 使得对角矩阵 R 的对角元素按以下次序排列:

$$1 \geq r_1 \geq r_2 \geq \cdots \geq r_n \geq 0$$

对角元素 $r_i (i = 1, 2, \cdots, n)$ 称为随机向量 y_i 和 x_i 的典型相关系数, 它与 y_i 和 x_i 的方差矩阵和协方差矩阵有密切的联系。

构造 $n_1 \times n_1$ 维矩阵 $\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$, 并设 $\lambda_1, \cdots, \lambda_{n_1}$ 为它的 n_1 个特征值, $\tilde{k}_1, \cdots, \tilde{k}_{n_1}$ 为相应于 $\lambda_i (i = 1, 2, \cdots, n_1)$ 的 n_1 维特征向量, 即

$$\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \tilde{k}_i = \lambda_i \tilde{k}_i; \quad i = 1, 2, \cdots, n_1$$

给定特征值 λ , 特征向量 \tilde{k}_i 并非唯一确定, 因为 \tilde{k}_i 的倍数仍为 λ 的特征向量, 所以可将 \tilde{k}_i 规范化, 使得 $\tilde{k}_i' \Sigma_{11} \tilde{k}_i = 1$ 。再令

$$k_i = \tilde{k}_i / \sqrt{\tilde{k}_i' \Sigma_{11} \tilde{k}_i} \quad (6.3.2)$$

从而有

$$(c) G_1' \Sigma_{11} G_1 = I_n, G_2' \Sigma_{22} G_2 = I_n;$$

$$(d) G_2' \Sigma_{21} G_1 = R, \text{ 这里}$$

$$R^2 = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

证明: 以下对定理的四个结论分别作证明。

(a) 矩阵 $\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$ 的特征值 $\lambda_1, \dots, \lambda_{n_1}$ 是方程

$$|\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} - \lambda I_{n_1}| = 0$$

的解。由于矩阵 Σ_{11} 正定非异, 所以上式等价于

$$|\Sigma_{11}^{-1}| |\Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} - \lambda \Sigma_{11}| = 0$$

因为 $|\Sigma_{11}^{-1}| \neq 0$, 因此

$$|\lambda \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}| = 0 \quad (6.3.3)$$

矩阵 $(\Sigma_{11} \quad \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21})$ 是随机向量 y_i 在给定 x_i 时的条件协方差矩阵, 所以是正定非异的, 因此 $|\Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}| \neq 0$ 。这说明 $\lambda = 1$ 不是 (6.3.3) 的解。另一方面, 因为

$$\lambda \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} = (\lambda - 1) \Sigma_{11} + (\Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}),$$

所以当 $\lambda > 1$ 时, 上式是两个正定矩阵的和, 因此也是正定的, 从而有

$$|\lambda \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}| = |(\lambda - 1) \Sigma_{11} + (\Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21})| \neq 0$$

这表明矩阵 $\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$ 所有的特征根 $\lambda_i (i = 1, 2, \dots, n_1)$ 都满足 $\lambda < 1$ 。

以下再证 $\lambda \geq 0 (i = 1, 2, \dots, n_1)$ 。若不然, 则至少存在某

一个 $\lambda_i < 0$, 使得 $\lambda_i \Sigma_{11}$ 为负定矩阵, 从而 $(\lambda_i \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21})$ 也为负定, 行列式 $|\lambda_i \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}|$ 不可能为零。综上所述

$$0 \leq \lambda_i < 1; \quad i = 1, 2, \dots, n_1$$

同理可证

$$0 \leq \mu_j < 1; \quad j = 1, 2, \dots, n_2$$

(b) k_i 为相应于特征值 λ_i 的特征向量, 有

$$\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} k_i = \lambda_i k_i \quad (6.3.4)$$

以 Σ_{21} 左乘上式, 可得

$$(\Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}) (\Sigma_{22}^{-1}) (\Sigma_{21} k_i) = \lambda_i (\Sigma_{21} k_i)$$

向量 $\Sigma_{21} k_i$ 为非零向量, 不然 (6.3.4) 式左侧为零, 所以 λ_i 也是矩阵 $(\Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1})$ 的特征值, 其相应的特征向量为 $(\Sigma_{21} k_i)$ 。由于矩阵 $(\Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1})$ 和它的转置 $(\Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12})$ 有同样的特征值, λ_i 也为 $(\Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12})$ 的特征值, 所以有

$$\lambda_i = \mu_i; \quad i = 1, 2, \dots, n$$

(c) 由于

$$\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} k_i = \lambda_i k_i$$

以 $k_i' \Sigma_{11}$ 左乘上式, 可得

$$k_i' \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} k_i = \lambda_i k_i' \Sigma_{11} k_i$$

同理也有

$$k_j' \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} k_j = \lambda_j k_j' \Sigma_{11} k_j$$

将以上两式相减, 就可得

$$(\lambda_i - \lambda_j) k_i' \Sigma_{11} k_j = 0$$

上式对任何 i, j 成立。因此当 $i \neq j$ 时, 必有 $k'_i \Sigma_{11} k_j = 0$; 而当 $i = j$ 时, $k'_i \Sigma_{11} k_i = 1$ 。综合起来, 即有 $G'_1 \Sigma_{11} G_1 = I_n$ 。同理可得, $G'_2 \Sigma_{22} G_2 = I_n$ 。

(d) 因为

$$\Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} h_i = \mu_i h_i$$

将其转置, 并以 $\Sigma_{21} k_j$ 右乘, 即得

$$h'_i \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} k_j = \mu_i h'_i \Sigma_{21} k_j$$

其中 k_j 是矩阵 $(\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21})$ 相应于特征值 λ_j 的特征向量, 因此上式等价于

$$h'_i \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} k_j = \lambda_j h'_i \Sigma_{21} k_j$$

根据(b)中的结论 $\lambda_i = \mu_i$, 从以上两式的差可得

$$(\lambda_i - \lambda_j) h'_i \Sigma_{21} k_j = 0$$

从而当 $i \neq j$ 时, $h'_i \Sigma_{21} k_j = 0$ 。

当 $i = j$ 时, 首先注意到

$$h'_i \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} h_i = \mu_i h'_i \Sigma_{22} h_i = \mu_i = \lambda_i \quad (6.3.5)$$

其次, 不失一般性, 可设 $n = n_1$, 由以上结论(c)可知 $G'_1 \Sigma_{11} G_1 = I_n$ 。并且由于这时 G_1 非异, 其逆阵存在, 所以

$$\Sigma_{11} = (G'_1)^{-1} G_1^{-1}$$

或等价地表示为

$$\Sigma_{11}^{-1} = G_1 G'_1$$

以上式代入(6.3.5)式, 可得

$$h'_i \Sigma_{21} G_1 G'_1 \Sigma_{12} h_i = \lambda_i$$

$$\lambda_1 = r_1^2 = \frac{\Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}}{\Sigma_{11}}$$

若以 y_t 和 x_t 的样本 ($t = 1, 2, \dots, T$) 对 Σ_{12} 、 Σ_{22} 和 Σ_{11} 作估计, 可得估计值 $\hat{\lambda}_1$ 为

$$\hat{\lambda}_1 = \hat{r}_1^2 = \frac{\hat{\Sigma}_{12} \hat{\Sigma}_{22}^{-1} \hat{\Sigma}_{21}}{\hat{\Sigma}_{11}}$$

其中, $\hat{\Sigma}_{11} = \frac{1}{T} \sum_{t=1}^T y_t^2$, $\hat{\Sigma}_{12} = \frac{1}{T} \sum_{t=1}^T y_t x_t$, $\hat{\Sigma}_{22} = \frac{1}{T} \sum_{t=1}^T x_t^2$ 。 \hat{r}_1^2 称为样本典型相关系数 (sample canonical correlation), 可写成以下形式:

$$\hat{r}_1^2 = \frac{(\sum y_t x_t)^2}{(\sum x_t^2)(\sum y_t^2)}$$

不难看出, 这正是以 y_t 对 x_t 作回归所得的拟合优度 (goodness of fit) R^2 。

6.4 ECM 和集中的对数似然函数

以 \mathbf{x}_t 表示 n 维 I(1) 随机向量, $\Delta \mathbf{x}_t$ 为其一阶差分。 \mathbf{x}_t 的误差修正形式为

$$\Delta \mathbf{x}_t = \sum_{i=1}^{k-1} \mathbf{D}_i \Delta \mathbf{x}_{t-i} + \pi \mathbf{x}_{t-k} + \boldsymbol{\varepsilon}_t; \quad t = 1, 2, \dots, T \quad (6.4.1)$$

式中的 \mathbf{D}_i ($i = 1, 2, \dots, k-1$) 和 π 为 $n \times n$ 维参数矩阵; $\{\boldsymbol{\varepsilon}_t\}$ 为独立同分布的正态随机向量, 且有 $E(\boldsymbol{\varepsilon}_t) = \mathbf{0}$, $D(\boldsymbol{\varepsilon}_t) = E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') = \Omega$ 。

同积过程的误差修正形式适用于对 \mathbf{x}_t 的系统分析, 它是以

下介绍的约翰逊方法的基础。矩阵 π 的秩 r 等于系统中独立的同积关系的个数。若 $r = n$, x_t 中有 n 个独立的同积关系, 从而 x_t 为一 $1(0)$ 随机向量; 若 $r = 0$, π 为一零矩阵, 因此 x_t 中不存在任何同积关系。除了这两种极端情况外, 有实际研究和应用价值的是 $0 < r < n$ 的情况, 此时 $n \times n$ 维矩阵 π 有分解形式

$$\pi = \alpha\beta'$$

其中 α 和 β 为列满秩的 $n \times r$ 维矩阵。约翰逊方法的一个重要特征, 是同积向量的个数 r 不是事先决定的, 而是假设检验的结果。在 r 确定以后, 再进一步估计矩阵 α 和 β 的值。

根据 (6.4.1) 式, Δx_t 有对数似然函数:

$$L(D_1, \dots, D_{k-1}, \pi, \Omega) = -\frac{Tn}{2} \ln(2\pi) - \frac{T}{2} \ln |\Omega| - \frac{1}{2} \sum_{t=1}^T \epsilon_t' \Omega^{-1} \epsilon_t \quad (6.4.2)$$

其中的 $D_1, \dots, D_{k-1}, \pi, \Omega$ 均为未知参数矩阵。当然, 最重要的是分析和估计矩阵 π , 它包含了同积系统的主要特征。为了更有效地估计和分析矩阵 π , 先将对数似然函数 (6.4.2) 集中在 π 上, 即除了矩阵 π 以外, 其他参数矩阵均用它们的最大似然估计代替, 然后再对 π 作最大似然估计。固定 (6.4.2) 中的 π , 对对数似然函数 $L(D_1, \dots, D_{k-1}, \pi, \Omega)$ 求关于 D_1, \dots, D_{k-1} 和 Ω 的一阶偏导, 并令其等于零, 可解出最大似然估计 D_1^*, \dots, D_{k-1}^* 和 Ω^* :

$$\Omega^* = T^{-1} \sum_{t=1}^T \hat{\epsilon}_t \hat{\epsilon}_t'$$

其中 $\hat{\epsilon}_t$ 为估计残差

$$\hat{\epsilon}_t = (\Delta x_t - \pi x_{t-k}) - \left\{ \sum_{i=1}^{k-1} D_i^* \Delta x_{t-i} \right\} \quad (6.4.3)$$

与(6.4.3)式相比,只要 $D_i^* = \hat{D}_i - \pi \bar{D}_i$, $R_{0i} - \pi R_{ki}$ 和 $\hat{\varepsilon}_i$ 就有同样的表示形式。事实上,因为残差 R_{0i} 和 R_{ki} 都与 $\Delta x_{i-1}, \dots, \Delta x_{i-k+1}$ 正交,因此 $R_{0i} - \pi R_{ki}$ 也与 Δx_i , ($i = 1, 2, \dots, k-1$) 正交。由正交投影的唯一性可知 $D_i^* = \hat{D}_i - \pi \bar{D}_i$, $i = 1, \dots, k-1$ 。这样,可以 $R_{0i} - \pi R_{ki}$ 代替 $L(D_1^*, \dots, D_{k-1}^*, \pi, \Omega^*)$ 中的 $\hat{\varepsilon}_i$:

$$\begin{aligned} & L(D_1^*, \dots, D_{k-1}^*, \pi, \Omega^*) \\ &= K - \frac{T}{2} \ln |\Omega^*| - \frac{1}{2} \sum_{i=1}^T (R_{0i} - \pi R_{ki})' \Omega^{*-1} (R_{0i} - \pi R_{ki}) \end{aligned}$$

其中的 Ω^* 为

$$\Omega^* = \frac{1}{T} \sum_{i=1}^T (R_{0i} - \pi R_{ki})(R_{0i} - \pi R_{ki})'$$

以下引理可使 $L(D_1^*, \dots, D_{k-1}^*, \pi, \Omega^*)$ 进一步简化。

引理 6.1: 设 A 、 B 分别为 $n \times k$ 维和 $k \times n$ 维矩阵,那么

$$\text{tr}(AB) = \text{tr}(BA)$$

其中 tr 为矩阵的迹数(trace),定义为:若 $n \times n$ 维矩阵 $A = \{a_{ij}; i, j = 1, 2, \dots, n\}$, A 的迹数为

$$\text{tr}(A) = \sum_{j=1}^n a_{jj}$$

证明见习题。

以 R^* 表示 $L(D_1^*, \dots, D_{k-1}^*, \pi, \Omega^*)$ 中的最后一项:

$$R^* = \frac{1}{2} \sum_{i=1}^T (R_{0i} - \pi R_{ki})' \Omega^{*-1} (R_{0i} - \pi R_{ki})$$

由于 R^* 为一标量,它的迹数即为自身,从而有

$$\begin{aligned}\hat{\Sigma}_{00} &= \frac{1}{T} \sum_{t=1}^T \mathbf{R}_{0t} \mathbf{R}'_{0t}, \quad \hat{\Sigma}_{kk} = \frac{1}{T} \sum_{t=1}^T \mathbf{R}_{kt} \mathbf{R}'_{kt} \\ \hat{\Sigma}_{0k} &= \frac{1}{T} \sum_{t=1}^T \mathbf{R}_{0t} \mathbf{R}'_{kt} = \hat{\Sigma}'_{k0}\end{aligned}\quad (6.5.1)$$

为讨论方便起见, 假设残差 \mathbf{R}_{0t} 和 \mathbf{R}_{kt} 已经规范化, 呈典型形式 (canonical form), 即有

$$\begin{aligned}\hat{\Sigma}_{00} &= \mathbf{I}_n, \quad \hat{\Sigma}_{kk} = \mathbf{I}_n, \quad \hat{\Sigma}_{0k} = \hat{\Sigma}_{k0} = \hat{\mathbf{R}} \\ \hat{\mathbf{R}} &= \begin{bmatrix} \hat{r}_1 & & & \\ & \hat{r}_2 & & \\ & & \ddots & \\ & & & \hat{r}_n \end{bmatrix}\end{aligned}\quad (6.5.2)$$

$\hat{\mathbf{R}}$ 的对角线元素 $\hat{r}_i (i = 1, 2, \dots, n)$ 为样本典型相关系数。

在 (6.4.7) 式中求未知参数矩阵 π 的最大似然估计 $\hat{\pi}$, 使得 $L^*(\hat{\pi})$ 取极大值, 等价于使下式

$$\left| T^{-1} \sum_{t=1}^T (\mathbf{R}_{0t} - \pi \mathbf{R}_{kt})(\mathbf{R}_{0t} - \pi \mathbf{R}_{kt})' \right| \quad (6.5.3)$$

取极小值, 并满足约束条件: 矩阵 π 的秩为 r , r 为系统中独立同积向量的个数。这一约束条件在以下的讨论中起重要作用。若不对矩阵 π 的秩施加约束条件, π 的第 i 行 π_i 的最大似然估计, 可由 R_{0t} 对 \mathbf{R}_{kt} 作回归得到, 这里 R_{0t} 为向量 \mathbf{R}_{0t} 的第 i 个元素;

$$\hat{\pi}'_i = \left\{ T^{-1} \sum_{t=1}^T \mathbf{R}_{kt} \mathbf{R}'_{kt} \right\}^{-1} \left\{ T^{-1} \sum_{t=1}^T \mathbf{R}_{kt} R_{0t} \right\}$$

根据 (6.5.2) 中的条件, 这时显然有

$$\hat{\pi}_i = \hat{r}_i e_i$$

这里 e_i 为 $n \times n$ 维单位矩阵的第 i 列。由此可计算回归估计残差的平方的均值:

$$\begin{aligned} & \frac{1}{T} \sum_{t=1}^T (R_{i0t} - \hat{\pi}_i' R_{kt})^2 \\ &= \left\{ \frac{1}{T} \sum_{t=1}^T (R_{i0t})^2 \right\} - \left\{ \frac{1}{T} \sum_{t=1}^T R_{i0t} R_{kt}' \right\} \left\{ \frac{1}{T} \sum_{t=1}^T (R_{kt} R_{kt}') \right\}^{-1} \left\{ \frac{1}{T} \sum_{t=1}^T R_{kt} R_{i0t} \right\} \\ &= 1 - \hat{r}_i' e_i' I_n e_i \hat{r}_i = 1 - \hat{r}_i^2 \end{aligned}$$

另一方面, 根据 (6.5.2) 中的条件, 第 i 个回归残差 $R_{i0t} - \hat{\pi}_i' R_{kt}$ 与第 j 个回归残差 $R_{j0t} - \hat{\pi}_j' R_{kt}$, 在 $i \neq j$ 时正交, 即

$$\sum_{t=1}^T (R_{i0t} - \hat{\pi}_i' R_{kt})(R_{j0t} - \hat{\pi}_j' R_{kt}) = 0$$

这样, 在参数矩阵不受约束的条件下, (6.5.3) 式在 $\hat{\pi}$ 处取极小值, 并有形式

$$\begin{aligned} & \min_{\pi} \left| T^{-1} \sum_{t=1}^T (R_{0t} - \pi R_{kt})(R_{0t} - \pi R_{kt})' \right| \\ &= \left| \begin{array}{cccc} 1 - \hat{r}_1^2 & & & \\ & 1 - \hat{r}_2^2 & & \\ & & \ddots & \\ & & & 1 - \hat{r}_n^2 \end{array} \right| = \prod_{i=1}^n (1 - \hat{r}_i^2) \end{aligned}$$

如果系统中有 r 个独立的同积向量, 且 $r < n$, 矩阵 π 的秩也为 r 。这时, 在估计中施加约束条件 $rk(\pi) = r$, 可以提高估计和假设检验的效率。另一方面, 由于

$$\hat{\pi}_i = \hat{r}_i e_i; \quad i = 1, 2, \dots, n$$

其中的 \hat{r}_i 满足 $\hat{r}_1 > \hat{r}_2 > \cdots > \hat{r}_n$, 因此条件 $rk(\pi) = r$ 表明

$$\hat{r}_{r+1} = \hat{r}_{r+2} = \cdots = \hat{r}_n = 0$$

从而, (6.5.3) 的条件极小值为

$$\begin{aligned} & \min_{\pi} \left| T^{-1} \sum_{t=1}^T (R_{0t} - \pi R_{kt}) (R_{0t} - \pi R_{kt})' \right| \\ &= \begin{vmatrix} 1 - \hat{r}_1^2 & & & & \\ & 1 - \hat{r}_2^2 & & & \\ & & \ddots & & \\ & & & 1 - \hat{r}_r^2 & \\ & & & & 1 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{vmatrix} = \prod_{i=1}^r (1 - \hat{r}_i^2) \end{aligned} \quad (6.5.4)$$

$$s. t. \quad rk(\pi) = r$$

(6.5.4) 式揭示了同积系统的最大似然估计方法与典型相关分析之间的密切联系。

6.6 参数矩阵 π 的最大似然估计

在集中的对数似然函数

$$L^*(\pi) = K_0 - \frac{T}{2} \ln \left| T^{-1} \sum_{t=1}^T (R_{0t} - \pi R_{kt}) (R_{0t} - \pi R_{kt})' \right|$$

中, 未知参数矩阵 π 的秩为 $r < n$, 且有分解式

$$\pi = \alpha \beta' \quad (6.6.1)$$

$$L^{**}(\beta) = K_0 - \frac{T}{2} \ln |S_{00} - S_{0k} \beta (\beta' S_{kk} \beta)^{-1} \beta' S_{k0}|$$

但是,对 $L^{**}(\beta)$ 求关于 β 的微分远非易事。以下引理,给出了简化 $L^{**}(\beta)$ 的方法。

引理 6.2: 若 $n \times n$ 维矩阵 A 有分块形式

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

其中, A_{11} 和 A_{22} 分别为 $n_1 \times n_1$ 维和 $n_2 \times n_2$ 维非异子矩阵, $n_1 + n_2 = n$ 。那么 A 的行列式可表示为

$$\begin{aligned} |A| &= |A_{11}| |A_{22} - A_{21} A_{11}^{-1} A_{12}| \\ &= |A_{22}| |A_{11} - A_{12} A_{22}^{-1} A_{21}| \end{aligned}$$

证明见习题。

考虑矩阵 S :

$$S = \begin{bmatrix} S_{00} & S_{0k} \beta \\ \beta' S_{k0} & \beta' S_{kk} \beta \end{bmatrix}$$

其中的子矩阵 S_{ij} ($i, j = 0, k$) 由 (6.6.2) 式给出。由 S_{00} 和 $\beta' S_{kk} \beta$ 的构造, 可知它们都为非异子矩阵, 因此 $|S_{00}| \neq 0$, $|\beta' S_{kk} \beta| \neq 0$ 。根据引理 6.2, 矩阵 S 的行列式为

$$\begin{aligned} |S| &= |S_{00}| |\beta' S_{kk} \beta - \beta' S_{k0} S_{00}^{-1} S_{0k} \beta| \\ &= |\beta' S_{kk} \beta| |S_{00} - S_{0k} \beta (\beta' S_{kk} \beta)^{-1} \beta' S_{k0}| \quad (6.6.5) \end{aligned}$$

因此有

$$|S_{00} - S_{0k} \beta (\beta' S_{kk} \beta)^{-1} \beta' S_{k0}| = |\beta' S_{kk} \beta|^{-1} |S_{00}| |\beta' (S_{kk} - S_{k0} S_{00}^{-1} S_{0k}) \beta|$$

将上式与 $L^{**}(\beta)$ 比较, 可知求 $L^{**}(\beta)$ 关于 β 的极大值, 等价于求下式

$$\frac{|\beta'(S_{kk} - S_{k0}S_{00}^{-1}S_{0k})\beta|}{|\beta'S_{kk}\beta|} \quad (6.6.6)$$

关于 β 的极小值。

由于在分解式 $\pi = \alpha\beta'$ 中, 矩阵 α 和 β 的值并非唯一确定, 因此总可以选择 β , 使得 $\beta'S_{kk}\beta = I_r$, 并在其中寻找(6.6.6)式的最小值。(6.6.6)式的极值问题由此简化成以下的条件极值:

$$\begin{aligned} \min_{\beta} & |\beta'(S_{kk} - S_{k0}S_{00}^{-1}S_{0k})\beta| \\ \text{s. t.} & \quad \beta'S_{kk}\beta = I_r \end{aligned} \quad (6.6.7)$$

其相应的拉格朗日函数为

$$|\beta'(S_{kk} - S_{k0}S_{00}^{-1}S_{0k})\beta| - \phi[\text{tr}(\beta'S_{kk}\beta - I_r)] \quad (6.6.8)$$

这里的 ϕ 为拉格朗日乘数, tr 表示矩阵的迹数。

求(6.6.7)式关于矩阵 β 的极值, 在文献中已有成熟的方法。以下介绍 T. W. 安德森(T. W. Anderson, 1958)的方法, 详细推导过程请参阅他书中的第十二章。

安德森指出, 求(6.6.8)式关于矩阵 β 的极值, 等价于在下列线性方程组

$$(\lambda S_{kk} - S_{k0}S_{00}^{-1}S_{0k})\beta = 0 \quad (6.6.9)$$

中求关于 β 的解。

其中的 λ 为特征方程

$$|\lambda S_{kk} - S_{k0}S_{00}^{-1}S_{0k}| = 0 \quad (6.6.10)$$

的解。根据 6.3 节中关于典型相关的讨论, (6.6.10)式的 n 个特征值满足条件

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$$

将 V_1, \dots, V_n 记作相应于 $\lambda_1, \dots, \lambda_n$ 的 n 个特征向量, 并满足条件

$$V_i' S_{kk} V_i = 1; \quad i = 1, 2, \dots, n$$

从而有

$$(\lambda_i S_{kk} - S_{k0} S_{00}^{-1} S_{0k}) V_i = 0 \quad (6.6.11)$$

选取 r 个特征向量 V_1, V_2, \dots, V_r , 使其分别相应于最大的 r 个特征值 $\lambda_1, \lambda_2, \dots, \lambda_r$, 令

$$\hat{\beta} = [V_1, V_2, \dots, V_r]$$

显然, $\hat{\beta}$ 满足方程 (6.6.9) 和 (6.6.10), 从而是条件极值 (6.6.7) 的一组解。令 P 为 $n \times r$ 维选择矩阵:

$$P' = [I_r, 0']$$

其中, 0 为 $(n-r) \times r$ 维零矩阵, $\hat{\beta}$ 可表示为

$$\hat{\beta} = VP$$

这里 $V = [V_1, V_2, \dots, V_n] = [\hat{\beta}, \hat{\delta}]$ 包含了所有 n 个特征向量, 子矩阵 $\hat{\delta}$ 包含的是相应于 $(n-r)$ 个最小的特征值的特征向量。

有了估计值 $\hat{\beta} = VP$, $L'(\pi)$ 中的其他参数的估计值就可相应地得到。将 $\hat{\beta}$ 代回 (6.6.4) 式, 可得 α 的估计值为

$$\hat{\alpha} = S_{0k} \hat{\beta} (\hat{\beta}' S_{kk} \hat{\beta})^{-1} = S_{0k} \hat{\beta}$$

最后, 矩阵 π 的最大似然估计为

$$\hat{\pi} = \hat{\alpha} \hat{\beta}' = S_{0k} \hat{\beta} \hat{\beta}' \quad (6.6.12)$$

$$(\mathbf{S}_{k0}\mathbf{S}_{00}^{-1}\mathbf{S}_{0k})\mathbf{V}_i = \lambda_i\mathbf{S}_{kk}; \quad i = 1, 2, \dots, n$$

因此

$$(\mathbf{S}_{k0}\mathbf{S}_{00}^{-1}\mathbf{S}_{0k})\mathbf{V} = \mathbf{S}_{kk}\mathbf{V}\mathbf{\Lambda}$$

这里

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

以 \mathbf{V}' 左乘上式, 因为 $\mathbf{V}'\mathbf{S}_{kk}\mathbf{V} = \mathbf{I}_n$, 所以

$$\mathbf{V}'\mathbf{S}_{k0}\mathbf{S}_{00}^{-1}\mathbf{S}_{0k}\mathbf{V} = (\mathbf{V}'\mathbf{S}_{kk}\mathbf{V})\mathbf{\Lambda} = \mathbf{\Lambda} \quad (6.6.14)$$

这个结果在计算中很有用。

将对角矩阵 $\mathbf{\Lambda}$ 中的元素 $\lambda_i (i = 1, 2, \dots, n)$ 按从大到小的次序排列。若以 $\mathbf{\Lambda}_r$ 和 $\mathbf{\Lambda}_{n-r}$ 分别表示 $\mathbf{\Lambda}$ 的子矩阵, 使得

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_r & \\ & \mathbf{\Lambda}_{n-r} \end{bmatrix}$$

根据 (6.6.14) 式, 则有

$$\mathbf{\Lambda}_r = \hat{\boldsymbol{\beta}}'\mathbf{S}_{k0}\mathbf{S}_{00}^{-1}\mathbf{S}_{0k}\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\alpha}}'\mathbf{S}_{00}^{-1}\hat{\boldsymbol{\alpha}}$$

$$\mathbf{\Lambda}_{n-r} = \hat{\boldsymbol{\delta}}'\mathbf{S}_{k0}\mathbf{S}_{00}^{-1}\mathbf{S}_{0k}\hat{\boldsymbol{\delta}} = \hat{\boldsymbol{\rho}}'\mathbf{S}_{00}^{-1}\hat{\boldsymbol{\rho}}$$

这里, $\hat{\boldsymbol{\rho}} = \mathbf{S}_{0k}\hat{\boldsymbol{\delta}}$ 为 $(n-r) \times n$ 维矩阵。

6.7 同积关系的假设检验

以 H_0 表示原假设 \mathbf{x}_i 中有 r 个独立的同积关系, $r = 0$,

1, ..., n-1。本节讨论对 H_r 的检验方法。

在集中的对数似然函数

$$L^{**}(\beta) = K_1 - \frac{T}{2} \ln |S_{00} - S_{0k} \beta (\beta' S_{kk} \beta)^{-1} \beta' S_{k0}|$$

中, 若以估计量 $\hat{\beta}$ 代入, 并注意到关系式 (6.6.5) 和 (6.6.6), 可得

$$L^{**}(\hat{\beta}) = K_2 - \frac{T}{2} \ln |\hat{\beta}' (S_{kk} - S_{k0} S_{00}^{-1} S_{0k}) \hat{\beta}|$$

又因为

$$\hat{\beta}' S_{kk} \hat{\beta} = I_r; \quad \hat{\beta}' S_{k0} S_{00}^{-1} S_{0k} \hat{\beta} = \Lambda_r$$

所以

$$\begin{aligned} L^{**}(\hat{\beta}) &= K_2 - \frac{T}{2} \ln |I_r - \Lambda_r| \\ &= K_2 - \frac{T}{2} \sum_{i=1}^r \ln(1 - \lambda_i) \end{aligned} \quad (6.7.1)$$

系统中有 r 个独立同积关系的假设, 等价于对角矩阵 Λ 中的后 $(n-r)$ 个元素为零的假设。构造迹数统计量 η_r 为

$$\eta_r = -T \sum_{i=r+1}^n \ln(1 - \lambda_i); \quad r = 0, 1, 2, \dots, n-1 \quad (6.7.2)$$

当原假设 H_r 为真时, $\eta_r = 0$, 反之亦然。

如果在估计参数矩阵 π 时, 不对其秩作 $rk(\pi) = r$ 的限制, 也不作分解 $\pi = \alpha\beta'$, 那么 π 的无限制的最大似然估计 $\hat{\pi}$ 使集中的对数似然函数

$$\eta_r \approx h\chi^2(2m^2)$$

这里, $h = 0.85 - \frac{0.58}{2m^2}$, $m = n - r$, $\chi^2(2m^2)$ 是自由度为 $2m^2$ 的 χ^2 -分布。

迹数统计量 η_r 检验的是联合假设 $\lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_n = 0$ 。但由于特征值 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 都为非负实数, 而且按从大到小的次序排列, 因此从逻辑上讲, 若已知 $\lambda_{r+1} = 0$, 即可推导出 $\lambda_{r+2} = \lambda_{r+3} = \cdots = \lambda_n = 0$ 。因此, 有检验原假设的第二种方法。构造统计量

$$\zeta_r = -T \ln(1 - \lambda_{r+1}); \quad r = 0, 1, 2, \cdots, n-1$$

它是检验第 $(r+1)$ 个特征值 λ_{r+1} 为零的似然比统计量, 称为 λ -max 统计量。 λ -max 统计量 ζ_r 的极限也有非标准的分布, 它的模拟值列在书后的表 7 中。当系统中有 $n-1$ 个独立的同积关系 ($n-r=1$) 时, 统计量 η_r 和 ζ_r 是一致的, 所以表 6 和表 7 的第一行是一样的。

需要指出的是, 在实际检验中, 迹数统计量 η_r 和 λ -max 统计量 ζ_r 会出现差异, 特别是在样本量有限的情况下, 这种差异可能会导致不同的结论。这方面的研究近来在文献中出现得很多, 但这已超出本书的范围, 在此不作详细介绍。

下面用一个实例说明以上两种检验方法的应用。

上一章的例 5.1 讨论了美国和意大利的物价水平 p_t 、 p_t^* 和两者之间的兑换率 s_t 之间的同积关系。根据菲利普斯—配荣的 Z_t 和 Z_p 检验结果, 这种同积关系的假设被拒绝了。以下用统计量 η_r 和 ζ_r 对这一同积关系再作检验。将 p_t 、 s_t 和 p_t^* ($t = 1, 2, \cdots, T$) 列入随机向量 \mathbf{x}_t , 有

$$\mathbf{x}_t' = [p_t, s_t, p_t^*]$$

并分别以 $\Delta \mathbf{x}_t$ 和 \mathbf{x}_{t-k} 对滞后变量 $\Delta \mathbf{x}_{t-1}, \cdots, \Delta \mathbf{x}_{t-k+1}$ 作回归, 可

得估计残差 R_{0t} 和 R_{kt} , 以此构造矩阵

$$\begin{aligned} S_{00} &= \frac{1}{T} \sum_{t=1}^T R_{0t} R_{0t}', & S_{0k} &= \frac{1}{T} \sum_{t=1}^T R_{0t} R_{kt}' \\ S_{kk} &= \frac{1}{T} \sum_{t=1}^T R_{kt} R_{kt}', & S_{k0} &= S_{0k}' \end{aligned}$$

利用例 5.1 的数据(从 1974 年 2 月至 1989 年 10 月, $T = 189$), 计算可得

$$\begin{aligned} S_{00} &= \begin{bmatrix} 0.0435 & -0.0316 & 0.0154 \\ -0.0316 & 4.687 & 0.0320 \\ 0.0154 & 0.0320 & 0.1799 \end{bmatrix} \\ S_{kk} &= \begin{bmatrix} 427.37 & -370.70 & 805.81 \\ -370.70 & 424.08 & -709.04 \\ 805.81 & -709.04 & 1525.45 \end{bmatrix} \\ S_{0k} &= \begin{bmatrix} -0.4849 & 0.4988 & -0.8377 \\ -1.8140 & -2.9593 & -2.4690 \\ -1.8084 & 1.4690 & -3.5899 \end{bmatrix} \end{aligned}$$

矩阵 $(S_{kk}^{-1} S_{k0} S_{00}^{-1} S_{0k})$ 的特征值为

$$\lambda_1 = 0.1105, \quad \lambda_2 = 0.05603, \quad \lambda_3 = 0.03039$$

它们都为正数, 并按从大到小的次序排列。

首先检验原假设 H_0 : 系统中无同积关系, 对应的备选假设 H_1 为系统中有一个同积关系。计算迹数统计量 η_0 为

$$\begin{aligned} \eta_0 &= -T \sum_{i=1}^3 \ln(1 - \lambda_i) \\ &= -189 [\ln(1 - 0.1105) + \ln(1 - 0.05603) \\ &\quad + \ln(1 - 0.03039)] = 38.85 \end{aligned}$$

由书后表 6 的情况三查得,相应于 5% 的临界值为 29.51,因为 $38.85 > 29.51$,所以拒绝原假设 H_0 ,接受系统中存在同积关系的假设。

再以 λ -max 统计量检验原假设 H_0 。因为

$$\zeta_0 = -T \ln(1 - \lambda_1) = 22.12$$

表 7 情况三给出的临界值为 20.78,所以又一次拒绝原假设,接受系统中有同积关系的假设。

以下进一步讨论系统中是否存在一个以上的同积关系,也即检验原假设 $H_1: r = 1$,对应的备选假设 $H_2: r = 2$ 。计算迹数统计量 η_1 为

$$\begin{aligned} \eta_1 &= -T \sum_{i=2}^3 \ln(1 - \lambda_i) \\ &= 10.90 + 5.83 = 16.73 \end{aligned}$$

和 λ -max 统计量 ζ_1 为

$$\zeta_1 = -T \ln(1 - \lambda_2) = 10.90$$

相应于 5% 显著水平的临界值分别为 15.2 和 14.0。由于 $16.73 > 15.2$,所以根据迹数统计量,应拒绝系统中只有一个同积关系的原假设,而接受备选假设 $H_2: r = 2$,即系统中有两个同积关系。但由于 $10.90 < 14.0$,所以根据 λ -max 统计量 ζ_1 ,不应拒绝系统中只有一个同积关系的原假设。显然, η_1 和 ζ_1 统计量给出了不同的统计结论。需注意的是, η_0 和 ζ_0 都接受系统中存在同积关系的假设,这与例 5.1 的结论是不一致的。在解释和评价这些检验结果时需十分谨慎,进一步的研究往往是必要的,如设法增加样本量等等。随意选择检验结果,对正确评价模型的可靠性是不利的。

若我们从以上的分析中接受系统中存在一个同积关系的假

设,就可进一步计算这一同积向量的值,它应是相应于最大的特征值 $\lambda_1 = 0.1105$ 的特征向量。在规范条件 $\beta_1' S_{xx} \beta_1 = 1$ 下,可得

$$\beta_1' = [-0.7579, 0.0280, 0.4220]$$

若将其中的第一个元素规范为 1,又得

$$\beta_1' = [1, -0.04, -0.56] \quad (6.7.5)$$

同积向量 β_1 与(5.4.31)式中的 OLS 方法得到的同积关系很接近(除了常数项以外)。

上例中检验同积关系个数的步骤是有普遍意义的。一般来说,检验从原假设 H_0 开始,若其被拒绝,则进一步检验 $H_1: r = 1$,一直到原假设 H_r 被接受为止。

6.8 对同积向量的假设检验

在确定了系统中同积关系的个数,估计出了同积向量的值以后,往往需要对同积向量的值作假设检验。这些假设往往有明显的经济和统计意义,可以提高模型的准确度和可靠性。比如,在同积向量(6.7.5)

$$\beta_1' = [1, -0.04, -0.56]$$

中,第二个分量的绝对值很小,所以它的真值可能为零。若这一假设成立,即

$$\beta_1' = [1, 0, -0.56]$$

那么 p_i 和 p_i^* 同积,但两者都不与 s_i 同积。

若在一个 n 维的系统 x_i 中存在 r 个同积关系,我们关心的是这些同积关系是否由 x_i 中的 q 个分量组成。显然,正数 q 满足 $r \leq q \leq n$ 。

以 \mathbf{R} 表示 $(n \times q)$ 维的选择矩阵,使得 $\mathbf{R}'\mathbf{x}_t$ 表示向量 \mathbf{x}_t 中组成同积关系的 q 个分量。比如在以上的例子中,矩阵 \mathbf{R} 为

$$\mathbf{R} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

检验同积关系是否由 $\mathbf{R}'\mathbf{x}_t$ 组成,仍可用以上讨论的方法,只是在求残差时,用 $\mathbf{R}'\mathbf{x}_{t-k}$ 对 $\Delta\mathbf{x}_{t-1}, \Delta\mathbf{x}_{t-2}, \dots, \Delta\mathbf{x}_{t-k-1}$ 作回归。这样求得的估计残差 $\tilde{\mathbf{R}}_{kt}$ 与 \mathbf{R}_{kt} 有如下关系:

$$\tilde{\mathbf{R}}_{kt} = \mathbf{R}'\mathbf{R}_{kt}$$

因此相应的各样本矩阵为

$$\tilde{\mathbf{S}}_{kk} = \frac{1}{T} \sum_{t=1}^T \tilde{\mathbf{R}}_{kt} \tilde{\mathbf{R}}'_{kt} = \frac{1}{T} \sum_{t=1}^T \mathbf{R}'\mathbf{R}_{kt} \mathbf{R}'_{kt} \mathbf{R} = \mathbf{R}'\mathbf{S}_{kk}\mathbf{R}$$

$$\tilde{\mathbf{S}}_{0k} = \frac{1}{T} \sum_{t=1}^T \mathbf{R}_{0t} \tilde{\mathbf{R}}'_{kt} = \frac{1}{T} \sum_{t=1}^T \mathbf{R}_{0t} \mathbf{R}'_{kt} \mathbf{R} = \mathbf{S}_{0k}\mathbf{R}$$

构造矩阵

$$\tilde{\mathbf{S}}_{kk}^{-1} \tilde{\mathbf{S}}_{k0} \mathbf{S}_{00}^{-1} \tilde{\mathbf{S}}_{0k} \quad (6.8.1)$$

其中, $\tilde{\mathbf{S}}_{0k} = \tilde{\mathbf{S}}'_{k0}$, $\mathbf{S}_{00} = \frac{1}{T} \sum_{t=1}^T \mathbf{R}_{0t} \mathbf{R}'_{0t}$, 以 $\tilde{\lambda}_i$ 表示矩阵(6.8.1)的第 i 个特征值。显然有

$$\tilde{\mathbf{S}}_{kk}^{-1} \tilde{\mathbf{S}}_{k0} \mathbf{S}_{00}^{-1} \tilde{\mathbf{S}}_{0k} = (\mathbf{R}'\mathbf{S}_{kk}\mathbf{R})^{-1} (\mathbf{R}'\mathbf{S}_{k0}\mathbf{S}_{00}^{-1}\mathbf{S}_{0k}\mathbf{R})$$

因为已知系统中有 r 个同积关系, $\Delta\mathbf{x}_t$ 的集中的对数似然函数 $L^{**}(\hat{\beta})$ 可表示为

$$L^{**}(\hat{\beta}) = K_2 - \frac{T}{2} \sum_{i=1}^r \ln(1 - \tilde{\lambda}_i)$$

$$\hat{\lambda}_1 = 0.1059; \quad \hat{\lambda}_2 = 0.0468$$

在上节的讨论中已经确定,系统 $\mathbf{x}_i' = [p_i, s_i, p_i^*]$ 中只有一个同积关系,因此 $r = 1$ 。由于 $\lambda_1 = 0.1105$, 所以检验统计量 $\tilde{\eta}_0$ 的值为

$$\begin{aligned}\tilde{\eta}_0 &= -T \sum_{i=1}^r \ln(1 - \lambda_i) + T \sum_{i=1}^r \ln(1 - \hat{\lambda}_i) \\ &= -189 \times \ln(1 - 0.1105) + 189 \times \ln(1 - 0.1059) \\ &= 22.12 - 21.15 = 0.97\end{aligned}$$

$\tilde{\eta}_0$ 在原假设成立时有 χ^2 极限分布,其自由度为

$$r(n - q) = 1 \times (3 - 2) = 1$$

$\chi^2(1)$ 的 5% 的临界值为 3.84, 因为 $0.97 < 3.84$, 所以不拒绝原假设。从而可以确认美国的物价水平 p_i 和意大利的物价水平 p_i^* 同积,但两者都不与兑换率 s_i 同积。

在经济学中,若 p_i 、 p_i^* 和 s_i 之间存在以下关系:

$$p_i = s_i + p_i^* + \varepsilon_i$$

其中, ε_i 为独立同分布, $E(\varepsilon_i) = 0$, $D(\varepsilon_i) = \sigma^2$, 则称美元和里拉之间存在购买力平价关系。这等价于在 $\mathbf{x}_i = [p_i, s_i, p_i^*]$ 中存在同积向量

$$\tilde{\beta}_1 = [1, -1, -1]$$

这一同积关系是否成立,也可由以上讨论的似然比检验法作检验。相应于这一假设的矩阵 \mathbf{R} 为

$$\mathbf{R}' = [1, -1, -1]$$

因此

$$\bar{S}_{kk} = R' S_{kk} R = 88.60$$

$$\bar{S}_{0k} = S_{0k} R = [-0.145, 3.614, 0.313]'$$

矩阵 $\bar{S}_{kk}^{-1} \bar{S}_{k0} S_{00}^{-1} \bar{S}_{0k}$ 为一标量, 因此它的特征值即为其本身, 所以

$$\bar{\lambda}_1 = \bar{S}_{kk}^{-1} \bar{S}_{k0} S_{00}^{-1} \bar{S}_{0k} = 0.042$$

从而有

$$T \ln(1 - \bar{\lambda}_1) = -8.20$$

检验假设 $\bar{\beta}_1 = [1, -1, -1]'$ 的统计量 $\tilde{\eta}_0$ 为

$$\begin{aligned} \tilde{\eta}_0 &= -T \sum_{i=1}^r \ln(1 - \lambda_i) + T \sum_{i=1}^r \ln(1 - \bar{\lambda}_i) \\ &= -189 \times \ln(1 - 0.1105) + 189 \times \ln(1 - 0.042) \\ &= 13.92 \end{aligned}$$

在原假设成立时, $\tilde{\eta}_0$ 有 χ^2 极限分布, 其自由度为

$$r(n - q) = 1 \times (3 - 1) = 2$$

$\chi^2(2)$ 分布的 5% 的临界值为 5.99, 因为 $13.92 > 5.99$, 所以应拒绝美元和里拉之间存在购买力平价的假设。

6.9 对矩阵 α 的假设检验

以上介绍的估计同积系统的最大似然方法的一个重要特点, 是对系统

$$\Delta x_t = \sum_{i=1}^{k-1} D_i \Delta x_{t-i} + \alpha \beta' x_{t-k} + \varepsilon_t$$

中的参数作联合的估计。由于利用了系统中所有 n 个方程包含

的信息,这一方法是最有效的。但在实际中,如果系统较大,其中包括的方程和参数的个数都较大,而样本量相对较小,那么对所有参数作联合估计会使估计方差增大,使假设检验的统计量的自由度过小。特别是当我们所感兴趣的只是系统的一部分,我们希望只对这一部分作估计,但又不减低估计的有效性,这样就要减少所需估计的参数的个数,在样本量有限的情况下提高估计的精确度,增加检验统计量的自由度。以下讨论这种做法的可行条件和对这些条件的检验方法。

考虑同积系统 $x_t = [x'_{1t}, x'_{2t}]'$ 的误差修正形式:

$$\Delta x_{1t} = \sum_{i=1}^{k-1} D_{1i} \Delta x_{t-i} + \alpha_1 \beta' x_{t-k} + \varepsilon_{1t} \quad (6.9.1)$$

$$\Delta x_{2t} = \sum_{i=1}^{k-1} D_{2i} \Delta x_{t-i} + \alpha_2 \beta' x_{t-k} + \varepsilon_{2t} \quad (6.9.2)$$

假设我们所感兴趣的只是系统中有关 x_{1t} 的方程和其中的参数,问题是能否仅用(6.9.1)中的方程对参数 α_1 、 β 和 D_{1i} ($i = 1, 2, \dots, k-1$) 作估计。如果这时在(6.9.2)式中 $\alpha_2 \neq 0$, 那么 r 个同积关系 $(\beta' x_{t-k})$ 也出现在 Δx_{2t} 的方程中,因此只利用 Δx_{1t} 中的方程估计 β , 就会失去包含在 Δx_{2t} 中的信息。这样得到的估计值 $\hat{\beta}$ 不是有效的。只有当 $\alpha_2 = 0$ 时,单独估计子系统 Δx_{1t} 才是可行的。这时称 x_{2t} 对于子系统 Δx_{1t} 中的参数 α_1 、 β 和 D_{1i} ($i = 1, 2, \dots, k-1$) 是弱外生的。只有当 x_{2t} 对于我们所感兴趣的方程中的参数是弱外生的, α_1 和 β 的最大似然估计才可由子系统 Δx_{1t} 中的方程估计所得,而不降低估计的有效性。

以下定理由约翰逊得到,证明见约翰逊(Johansen, 1995, pp. 123—124)。

定理 6.2: 在系统(6.9.1)和(6.9.2)中,若 $\alpha_2 = 0$, 那么 x_{2t} 对于子系统 Δx_{1t} 中的参数 α_1 和 β 是弱外生的, α_1 和 β 的最大似

然估计可由子系统 Δx_{1t} 中的方程得到。

因此,在系统(6.9.1)和(6.9.2)中对假设 $\alpha_2 = 0$ 作检验,对提高参数估计的准确性有重要意义。当然,检验 $\alpha_2 = 0$ 只是对矩阵 α 作假设检验的一个特例。以下讨论对矩阵 α 的线性约束条件作假设检验。这些约束条件采取

$$\alpha = A\Psi \quad (6.9.3)$$

的形式。其中, A 为 $(n \times s)$ 维的已知实矩阵, Ψ 为 $(s \times r)$ 维参数矩阵,它的元素需由估计确定, r 为系统中同积关系的个数,且 $r \leq s \leq n$ 。显然, $\alpha_2 = 0$ 是(6.9.3)式的一个特例。因为,若令

$$A = \begin{bmatrix} I \\ 0 \end{bmatrix}; \quad \Psi = \alpha_1$$

那么根据(6.9.3),就有

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = A\Psi = \begin{bmatrix} \alpha_1 \\ 0 \end{bmatrix}$$

也即等价于

$$\alpha_2 = 0$$

在参数不受约束的情况下,矩阵 $\pi = \alpha\beta'$ 的最大似然估计可从集中的对数似然函数 $L^*(\alpha, \beta)$ 得到:

$$L^*(\alpha, \beta) = K_0 - \frac{T}{2} \ln \left| \frac{1}{T} \sum_{t=1}^T \hat{\varepsilon}_t \hat{\varepsilon}_t' \right|$$

其中, $\hat{\varepsilon}_t$ 满足关系式

$$R_{0t} = \alpha\beta' R_{kt} + \hat{\varepsilon}_t \quad (6.9.4)$$

R_{0t} 和 R_{kt} ($t = 1, 2, \dots, T$) 为估计残差。以 A_1 表示矩阵 A 的正

交补阵,使得 $A'A_{\perp} = 0$, $rk(A, A_{\perp}) = n$ 。分别以 A' 和 A'_{\perp} 左乘 (6.9.4) 式,得

$$A'R_{0t} = A'\alpha\beta'R_{kt} + A'\hat{\epsilon}_t = (A'A)\Psi\beta'R_{kt} + A'\hat{\epsilon}_t$$

$$A'_{\perp}R_{0t} = A'_{\perp}\alpha\beta'R_{kt} + A'_{\perp}\hat{\epsilon}_t = (A'_{\perp}A)\Psi\beta'R_{kt} + A'_{\perp}\hat{\epsilon}_t = A'_{\perp}\hat{\epsilon}_t$$

并令

$$\bar{\epsilon}_t = \begin{bmatrix} A'\hat{\epsilon}_t \\ A'_{\perp}\hat{\epsilon}_t \end{bmatrix}$$

就可计算和式

$$\sum_{t=1}^T \bar{\epsilon}_t \bar{\epsilon}'_t = \begin{bmatrix} S_{aa} - S_{a1}\beta\Psi'A'A - A'A\Psi\beta'S_{1a} - A'A\Psi\beta'S_{kk}\beta\Psi'A'A & S_{ab} & A'A\Psi\beta'S_{1b} \\ S_{ba} - S_{b1}\beta\Psi'A'A & S_{bb} & \end{bmatrix}$$

$$\text{其中, } S_{aa} = A'S_{00}A, S_{bb} = A'_{\perp}S_{00}A_{\perp}, S_{ab} = A'S_{00}A_{\perp}, S_{a1} = S'_{1a} \\ = A'S_{0k}, S_{00} = \frac{1}{T} \sum_{t=1}^T R_{0t}R'_{0t}, S_{0k} = \frac{1}{T} \sum_{t=1}^T R_{0t}R'_{kt}。$$

根据引理 6.2, 矩阵 $\sum_{t=1}^T \bar{\epsilon}_t \bar{\epsilon}'_t$ 的行列式由下式

$$\begin{aligned} \left| \sum_{t=1}^T \bar{\epsilon}_t \bar{\epsilon}'_t \right| &= |S_{bb}| |(S_{aa} - S_{ab}S_{bb}^{-1}S_{ba}) \cdots (S_{a1} - S_{ab}S_{bb}^{-1}S_{b1})\beta\Psi'A'A \\ &\quad - A'A\Psi\beta'(S_{1a} - S_{1b}S_{bb}^{-1}S_{ba}) \\ &\quad + A'A\Psi\beta'(S_{kk} - S_{1b}S_{bb}^{-1}S_{b1})\beta\Psi'A'A| \end{aligned}$$

给出。将上式的四个括弧中的表达式分别按次序定义为 $S_{aa \cdot b}$, $S_{a1 \cdot b}$, $S_{1a \cdot b}$ 和 $S_{11 \cdot b}$, 可将上式改写为

$$\left| \sum_{t=1}^T \bar{\epsilon}_t \bar{\epsilon}'_t \right| = |S_{bb}| |S_{aa \cdot b} - S_{a1 \cdot b}\beta\Psi'A'A$$

$$-A'A\Psi\beta'S_{1\alpha,b} + A'A\Psi\beta'S_{11,b}\beta\Psi'A'A|$$

由此可以构造受约束的对数似然函数 $\bar{L}(\alpha, \beta)$ 为

$$\bar{L}(\alpha, \beta) = K_0 - \frac{T}{2} \ln \left| \frac{1}{T} \sum_{i=1}^T \bar{\epsilon}_i \bar{\epsilon}_i' \right|$$

根据 6.6 节中的讨论,若进一步将 $\bar{L}(\alpha, \beta)$ 集中在 β 上,求 β 的最大似然估计,需先解特征方程

$$|\bar{\lambda}S_{11,b} - S_{1\alpha,b}S_{\alpha\alpha,b}^{-1}S_{\alpha1,b}| = 0$$

求得矩阵 $S_{11,b}^{-1}, S_{1\alpha,b}, S_{\alpha\alpha,b}^{-1}, S_{\alpha1,b}$ 的特征值 $\bar{\lambda}, i = 1, 2, \dots, n$, 相应于最大的 r 个特征值的特征向量构成矩阵 β 的最大似然估计 $\bar{\beta}$ 。类似于非约束的最大似然估计 $\hat{\alpha}$ 和 $\hat{\beta}$ 的关系 (6.6.12), 受约束的最大似然估计 $\bar{\beta}$ 和 $\bar{\Psi}$ 之间存在关系:

$$A'A\Psi = S_{\alpha1,b}\bar{\beta}(\bar{\beta}'S_{11,b}\bar{\beta})^{-1} = S_{\alpha1,b}\bar{\beta}$$

从而有

$$\bar{\alpha} = A(A'A)^{-1}S_{\alpha1,b}\bar{\beta}$$

将以上关系式代入集成的对数似然函数 $\bar{L}(\alpha, \beta)$, 可得

$$\bar{L}(\Psi, \beta) = K_0 - \frac{T}{2} \sum_{i=1}^r \ln(1 - \bar{\lambda}_i)$$

若对于系统中的 r 个同积关系, 约束条件 $\alpha = A\Psi$ 为真, 集成的对数似然函数 $\bar{L}(\Psi, \beta)$ 在 Ψ 和 $\bar{\beta}$ 取极大值。这样, 检验假设 $\alpha = A\Psi$ 的似然比统计量为

$$\begin{aligned} \bar{\eta}_0 &= 2(L^{**}(\hat{\beta}) - \bar{L}(\Psi, \bar{\beta})) \\ &= -T \left\{ \sum_{i=1}^r \ln(1 - \lambda_i) + \sum_{i=1}^r \ln(1 - \bar{\lambda}_i) \right\} \end{aligned}$$

4. 在回归模型

$$y_t = x_t' \beta + \varepsilon_t$$

中, β 为未知参数, ε_t 独立同分布, $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = \sigma^2$, 回归变量 x_t 对于参数是非弱外生的。若 $E(x_t \varepsilon_t) = C \neq 0$, 试证: 在这种情况下参数 β 的最小二乘估计 $\hat{\beta}_T$ 不是一致的估计量, 即 $\text{plim } \hat{\beta}_T \neq \beta$ 。

5. 在回积系统

$$\Delta x_{1t} = \alpha_1 \beta' x_{t-1} + \varepsilon_{1t}$$

$$\Delta x_{2t} = \alpha_2 \beta' x_{t-1} + \varepsilon_{2t}$$

中, $x_t = [x_{1t}, x_{2t}]'$, α_1 , α_2 和 β 为未知参数, $E(\varepsilon_{1t}, \varepsilon_{2t}) \neq 0$ 。试将 Δx_{1t} 表示成 Δx_{2t} 的函数:

$$\Delta x_{1t} = \rho \Delta x_{2t} + v_t$$

这里, $\rho = \alpha_1 / \alpha_2$, $v_t = \varepsilon_{1t} - \rho \varepsilon_{2t}$ 并证明此时 Δx_{2t} 对于参数 ρ 不是弱外生的。

第 7 章 ARCH 模型

7.1 简介

前几章介绍了处理非稳定时间序列的一些最新的发展,着重讨论了对单位根过程和同积过程的参数估计和假设检验方法。本章将介绍另一种重要的时间序列过程——自回归条件异方差过程(autoregressive conditional heteroskedastic process),简称 ARCH 过程。它在过去十几年中得到广泛的重视,如今已成为研究实证金融学和金融经济计量学的主要工具之一。ARCH 模型由恩格尔(Engle, 1982)首次提出,立即受到广大研究人员的重视,出现了许多重要的研究成果。ARCH 模型的理论 and 数学结构已经建立,它的主要性质和特点已为研究人员和实际工作者接受。ARCH 模型和它的各种推广形式,以及它们在金融领域里的应用,是本章讨论的中心。

在 ARCH 模型提出之前,描述股票价格等金融现象最常用的模型是随机游动模型。如以

$$P_t = P_{t-1} + \varepsilon_t$$

描述某股价在时刻 t 的值, ε_t 为白噪声干扰。根据这一模型,股价 P_t 在时刻 t 的增值 $\Delta P_t = P_t - P_{t-1}$ 是期望为零的白噪声过程,在样本量较大时可用正态分布作近似。这一模型符合金融学

中有效市场(efficient market)的理论,运用简便,常用来预测和估算股票价格。但随机游动模型在应用中有明显的局限性。金融市场中的一个常见的现象,是股价或其他类似的金融产品的变易率(volatility)不仅随时间 t 变化,而且常在某一时段中连续出现偏高或偏低的情况。图 7.1 和图 7.2 给出了两个常用的金融时间序列的数据,它们上下波动的范围相应于变易率的大小。显然,这两个时间序列都呈现变易率聚类(volatility clustering)

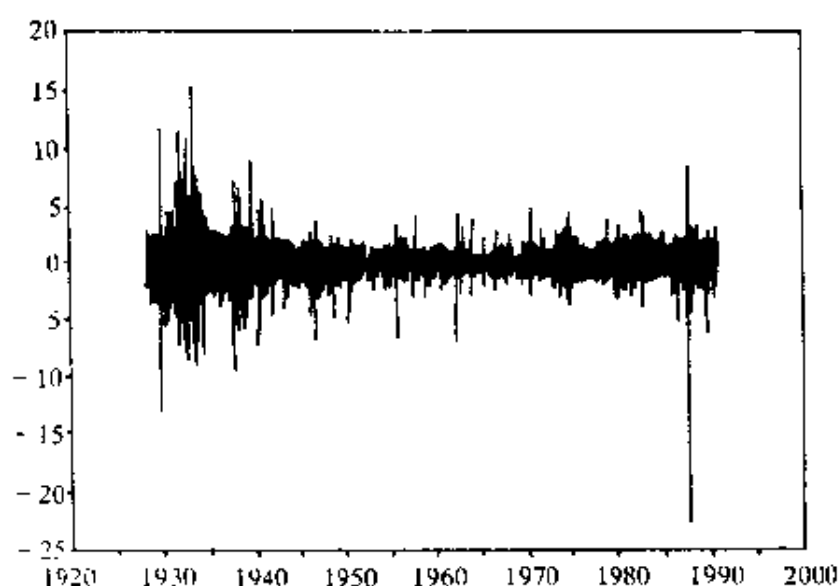


图 7.1 S & P 500 指数的日回报率(%):1953-1990

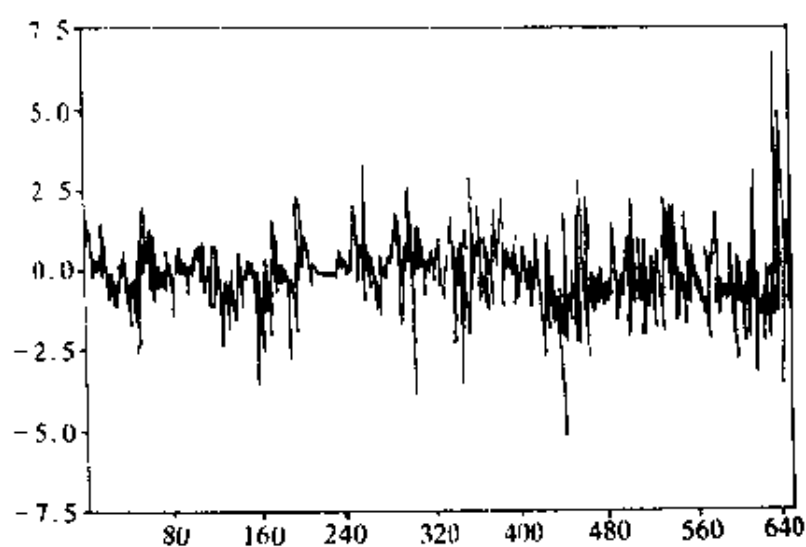


图 7.2 美元/英镑兑换率的周回报率(%)

的特点。

股价和其他金融变量(如汇率等)的这一特征很难由随机游动模型解释,但却可用 ARCH 模型及其各种推广模型来描述。正由于 ARCH 类模型的这一特点,它在过去的十几年里得到了普遍的重视,并由此出现了许多重要的成果。由于篇幅有限,本章只能对 ARCH 模型作简要的介绍。

7.2 自回归条件异方差过程(ARCH)的定义

自回归条件异方差过程(ARCH 过程)在文献中有多种不同的定义方法。以下介绍的是基于恩格尔在 1982 年提出的定义。

一个随机变量 x_t 有 P 阶的自回归表示形式 $AR(P)$, 如果

$$x_t = \beta_0 + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \beta_p x_{t-p} + \varepsilon_t \quad (7.2.1)$$

其中, $\{\varepsilon_t\}$ 为独立同分布的白噪声过程, 且有 $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = \sigma^2$ 。上式可等价地表示成滞后多项式

$$(1 - \beta_1 L - \beta_2 L^2 - \cdots - \beta_p L^p) x_t = \beta_0 + \varepsilon_t$$

$AR(P)$ 过程(7.2.1)是一稳定过程, 它的特征多项式

$$1 - \beta_1 z - \beta_2 z^2 - \cdots - \beta_p z^p = 0$$

所有的根都在单位圆外。

在(7.2.1)式中, 若固定变量 $x_{t-1}, x_{t-2}, \cdots, x_{t-p}$ 的值, 随机变量的条件期望为

$$E(x_t | x_{t-1}, \cdots, x_{t-p}) = \beta_0 + \beta_1 x_{t-1} + \cdots + \beta_p x_{t-p}$$

它是在给定 x_{t-1}, \cdots, x_{t-p} 时对 x_t 的最佳线性预测。另一方面, 不难计算 x_t 的无条件期望为

$$E(x_t) = \frac{\beta_0}{1 - \beta_1 - \dots - \beta_p} \quad (7.2.2)$$

它为一与时间 t 无关的常数, 而 x_t 的条件期望 $E(x_t | x_{t-1}, \dots)$ 却是时间 t 的函数。

若有一随机过程 $\{\epsilon_t\}$, 它的平方 ϵ_t^2 服从 $AR(q)$ 过程

$$\epsilon_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \dots + \alpha_q \epsilon_{t-q}^2 + \eta_t \quad (7.2.3)$$

其中 $\{\eta_t\}$ 独立同分布, 且有 $E(\eta_t) = 0$, $D(\eta_t) = \lambda^2$, $t = 1, 2, \dots$, 则称 $\{\epsilon_t\}$ 服从 q 阶的 ARCH 过程, 记作 $\epsilon_t \sim \text{ARCH}(q)$ 。在 (7.2.3) 式中一般还假设 $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, 2, \dots, q$ 。由于随机变量 ϵ_t^2 的非负性, 给定变量 $\epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2$ 的值, 白噪声过程 $\{\eta_t\}$ 的分布是受约束的, 因为它显然应满足

$$\eta_t \geq -\alpha_0; \quad t = 1, 2, \dots$$

为确保 $\{\epsilon_t^2\}$ 为一稳定过程, 假设 (7.2.3) 式的特征方程

$$1 - \alpha_1 z - \alpha_2 z^2 - \dots - \alpha_q z^q = 0$$

的所有的根都在单位圆外。若 $\alpha_0 > 0$, $\alpha_i \geq 0$ ($i = 1, 2, \dots, q$) 成立, 以上条件等价于

$$\alpha_1 + \alpha_2 + \dots + \alpha_q < 1 \quad (7.2.4)$$

这样, 若 $\epsilon_t \sim \text{ARCH}(q)$, 那么 ϵ_t 的无条件方差

$$\sigma^2 = E(\epsilon_t^2) = \frac{\alpha_0}{1 - \alpha_1 - \dots - \alpha_q} \quad (7.2.5)$$

为一常数。

ARCH 模型的一个重要特点是给出了计算时间序列的条件方差的方法。在每一时刻 t , ARCH 过程的条件方差是过去的随机干扰的函数, 可由递推公式计算。以 σ_t^2 表示 $\text{ARCH}(q)$ 过程 ϵ_t 在时刻 t 的条件方差, 给定随机变量 $\epsilon_{t-1}^2, \epsilon_{t-2}^2, \dots, \epsilon_{t-q}^2$ 的

值,则

$$\sigma_t^2 = E(\epsilon_t^2 | \epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2) = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \dots + \alpha_q \epsilon_{t-q}^2$$

所以,只要知道参数 $\alpha_0, \alpha_1, \dots, \alpha_q$ 的值,就可在时刻 $(t-1)$, 利用给定的数据 $\epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2$, 准确地预测 ϵ_t 在时刻 t 的条件方差 σ_t^2 。以下讨论在时刻 t , 对 ϵ_t 在时刻 $t+s (s \geq 0)$ 的条件方差的预测。

以 $\hat{\epsilon}_{t+1,t}^2$ 表示在时刻 t , 对 ϵ_t 在时刻 $t+s$ 的条件方差 σ_{t+s}^2 的估计值。显然,当 $s=1$ 时,有

$$\hat{\epsilon}_{t+1,t}^2 = \sigma_{t+1}^2 = \alpha_0 + \alpha_1 \epsilon_t^2 + \dots + \alpha_q \epsilon_{t-q+1}^2 \quad (7.2.6)$$

根据(7.2.5)式,可将未知参数 α_0 表示为

$$\alpha_0 = (1 - \alpha_1 - \alpha_2 - \dots - \alpha_q) \sigma^2$$

以此代入(7.2.6)式,可得

$$\hat{\epsilon}_{t+1,t}^2 = (1 - \alpha_1 - \dots - \alpha_q) \sigma^2 + \alpha_1 \epsilon_t^2 + \dots + \alpha_q \epsilon_{t-q+1}^2$$

稍加整理,又得

$$\begin{aligned} \hat{\epsilon}_{t+1,t}^2 - \sigma^2 &= \alpha_1 (\epsilon_t^2 - \sigma^2) + \alpha_2 (\epsilon_{t-1}^2 - \sigma^2) \\ &\quad + \dots + \alpha_q (\epsilon_{t-q+1}^2 - \sigma^2) \end{aligned}$$

同理可得在时刻 $t+1$, 对 ϵ_t 在时刻 $t+2$ 的条件方差的估计为

$$\begin{aligned} \hat{\epsilon}_{t+2,t+1}^2 - \sigma^2 &= \alpha_1 (\epsilon_{t+1}^2 - \sigma^2) + \alpha_2 (\epsilon_t^2 - \sigma^2) \\ &\quad + \dots + \alpha_q (\epsilon_{t-q+2}^2 - \sigma^2) \end{aligned}$$

$\hat{\epsilon}_{t+1,t}^2$ 也是对随机变量 ϵ_{t+1}^2 的最佳线性预测。以 $\hat{\epsilon}_{t+1,t}^2$ 代替

上式中的 ϵ_{t-1}^2 , 就可得在时刻 t , 对 ϵ_t 在 $t+2$ 时的条件方差为

$$\begin{aligned}\hat{\epsilon}_{t+2/t}^2 - \sigma^2 &= \alpha_1(\hat{\epsilon}_{t+1/t}^2 - \sigma^2) + \alpha_2(\epsilon_t^2 - \sigma^2) \\ &\quad + \cdots + \alpha_q(\epsilon_{t-q+2}^2 - \sigma^2)\end{aligned}\quad (7.2.7)$$

一般地, 可以定义在时刻 t , 对 ϵ_t 在 $t+s$ 时的条件方差的估计为

$$\hat{\epsilon}_{t+s/t}^2 = E(\epsilon_{t+s}^2 | \epsilon_t^2, \cdots)$$

它也是在时刻 t , 对 $\epsilon_{t+s}^2 (s=1, 2, \cdots)$ 的 s 步预测。

类似(7.2.7)式, 经连续往后迭代, 可得计算 $\hat{\epsilon}_{t+s/t}^2$ 的递推公式

$$\begin{aligned}\hat{\epsilon}_{t+s/t}^2 - \sigma^2 &= \alpha_1(\hat{\epsilon}_{t+s-1/t}^2 - \sigma^2) + \alpha_2(\hat{\epsilon}_{t+s-2/t}^2 - \sigma^2) + \cdots \\ &\quad + \alpha_q(\hat{\epsilon}_{t+s-q/t}^2 - \sigma^2)\end{aligned}$$

其中的 $\hat{\epsilon}_{t+s-i/t}^2 (i=1, 2, \cdots, q)$ 是在时刻 t , 对 ϵ_{t+s-i}^2 的最佳线性预测。当 $s \leq q$ 时, 显然有

$$\hat{\epsilon}_{t+s-q/t}^2 = \epsilon_{t+s-q}^2$$

为进一步研究 ARCH(q) 过程的性质, 以下将 $\epsilon_t \sim \text{ARCH}(q)$ 表示为

$$\epsilon_t = \sqrt{h_t} \cdot v_t \quad (7.2.8)$$

并假设 $\{v_t\}$ 独立同分布, $E(v_t) = 0$, $D(v_t) = 1$ 。 h_t 有表达式

$$h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-2}^2 + \cdots + \alpha_q \epsilon_{t-q}^2$$

显然, 在任何时刻 t , ϵ_t 的条件期望为

$$\lambda^2 - \left\{ \frac{\alpha_1^2 \lambda^2}{1 - \alpha_1^2} + \frac{\alpha_0^2}{(1 - \alpha_1)^2} \right\} \cdot E(v_i^2 - 1) \quad (7.2.10)$$

作为 λ 的二次方程, 上式对 λ 一般不存在实数解。比如, 可设 $v_i \sim N(0, 1)$, 则有 $E(v_i^2 - 1) = 2$, 上式简化为

$$\frac{(1 - 3\alpha_1^2)\lambda^2}{1 - \alpha_1^2} = \frac{2\alpha_0^2}{(1 - \alpha_1)^2}$$

显然, 当 $\alpha_1^2 \geq \frac{1}{3}$ 时, 上式不存在 λ 的实数解。

给定 $h_i = \alpha_0 + \alpha_1 \varepsilon_{i-1}^2 + \dots + \alpha_q \varepsilon_{i-q}^2$ 中参数 $\alpha_0, \alpha_1, \dots, \alpha_q$ 的值, ARCH(q) 过程 $\varepsilon_i = \sqrt{h_i} \cdot v_i$ 的值可由蒙的可罗模拟方法计算所得。以下的三个图由蒙的可罗方法模拟产生, 从中可明显地看出 ARCH 过程变易率聚类的特点。

图 7.3 给出的是正态白噪声过程的图像, 它相应于在 $\varepsilon_i = \sqrt{h_i} \cdot v_i$ 中 $\alpha_0 = 1, \alpha_1 = \alpha_2 = \dots = \alpha_q = 0$ 的情况。 v_i 由标准正态分布产生。

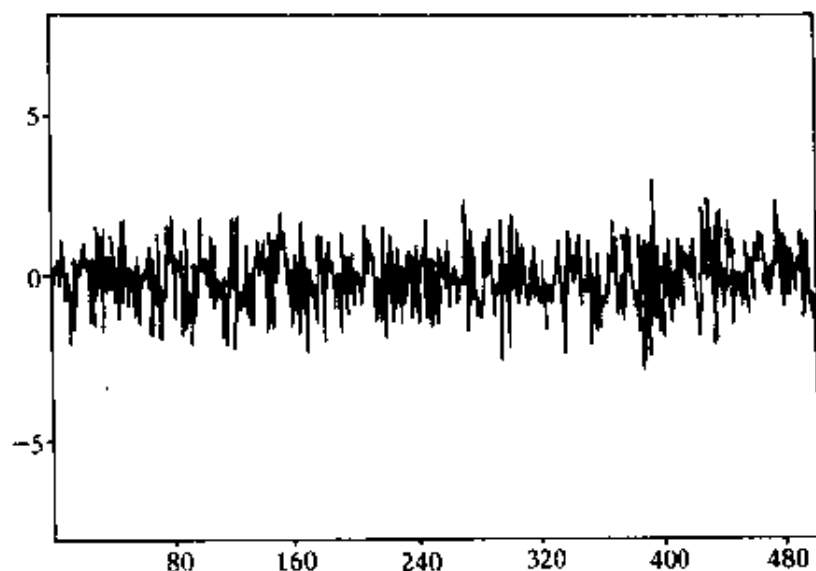


图 7.3 模拟的正态白噪声过程

利用上图中模拟产生的数据 $v_i (i = 1, 2, \dots, 500)$, 再令

$\alpha_0 = 0.1, \alpha_1 = 0.9, \alpha_2 = \alpha_3 = \cdots = \alpha_q = 0$, 则有

$$h_t = 0.1 + 0.9\epsilon_{t-1}^2$$

从而可以构造 ARCH(1)过程 $\epsilon_t = \sqrt{h_t} \cdot v_t$, 它的图像由图 7.4 给出。

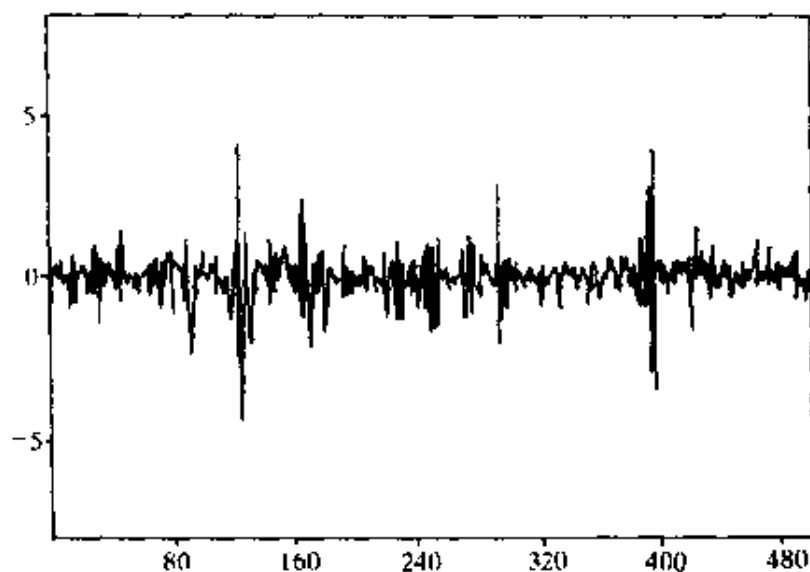


图 7.4 模拟的 ARCH(1)过程

再利用图 7.3 中的模拟数据 $v_t (t = 1, 2, \cdots, 500)$, 令

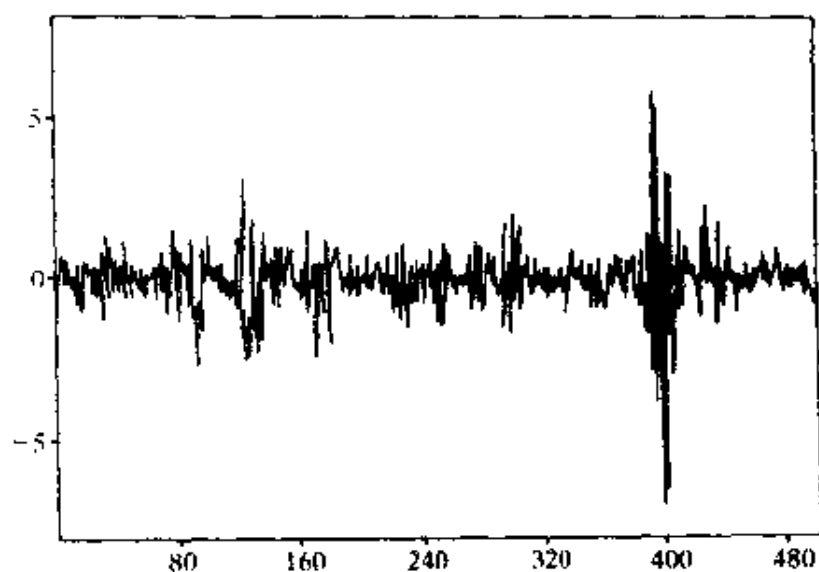


图 7.5 模拟的 ARCH(4)过程

$$h_t = 0.1 + 0.36\epsilon_{t-1}^2 + 0.27\epsilon_{t-2}^2 + 0.18\epsilon_{t-3}^2 + 0.09\epsilon_{t-4}^2$$

可构造 ARCH(4)过程,它的图像由图 7.5 给出。

在以上的模拟计算中,三个过程的无条件方差都设定为 1,这就使得图中条件方差的作用表现得更为明显。

7.3 ARCH 过程参数的最大似然估计

ARCH 过程最通常的应用,是在回归模型

$$y_t = x_t' \beta + \epsilon_t \quad (7.3.1)$$

中,假设 $\epsilon_t \sim \text{ARCH}(q)$ 。为计算方便起见,一般假设 y_t 的前 q 个观察值已知,并记为 $y_{-q+1}, y_{-q+2}, \dots, y_0$, 估计所用数据为 y_1, y_2, \dots, y_T 。模型(7.3.1)中的 x_t 为已知的回归变量,其中可以包括滞后的 y_t 值。 ϵ_t 服从 ARCH(q)过程,可表示为

$$\epsilon_t = \sqrt{h_t} \cdot v_t$$

其中, $h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \dots + \alpha_q \epsilon_{t-q}^2$ 。

以下首先讨论 v_t 服从标准正态分布时,ARCH(q)过程的参数的最大似然估计, v_t 服从非正态分布的情况在稍后讨论。

将 y_t 和 x_t ,以及它们的滞后值列在向量 Y_t 中:

$$Y_t = [y_t, y_{t-1}, \dots, y_1, y_0, \dots, y_{-q+1}, x_t', \dots, x_t', x_{t-1}', \dots, x_{-q+1}']'$$

根据模型(7.3.1)的构造,随机干扰 v_t 与 x_t, Y_{t-1} 相互独立。给定 x_t 和 Y_{t-1} 的值,随机变量 y_t 服从正态分布,并有条件密度函数:

$$\frac{\partial L(\theta)}{\partial \theta} = \sum_{i=1}^T l_i(\theta)$$

$l_i(\theta)$ 可表示为

$$l_i(\theta) = -\frac{1}{2} \frac{\partial \ln(h_i)}{\partial \theta} - \frac{1}{2} \left\{ \frac{1}{h_i} \frac{\partial (y_i - \mathbf{x}_i' \boldsymbol{\beta})^2}{\partial \theta} - \frac{(y_i - \mathbf{x}_i' \boldsymbol{\beta})^2}{h_i^2} \frac{\partial h_i}{\partial \theta} \right\}$$

由于

$$\frac{\partial (y_i - \mathbf{x}_i' \boldsymbol{\beta})^2}{\partial \theta} = \begin{bmatrix} \frac{\partial (y_i - \mathbf{x}_i' \boldsymbol{\beta})^2}{\partial \boldsymbol{\beta}} \\ \frac{\partial (y_i - \mathbf{x}_i' \boldsymbol{\beta})^2}{\partial \delta} \end{bmatrix} = \begin{bmatrix} -2\mathbf{x}_i \epsilon_i \\ 0 \end{bmatrix} \quad (7.3.4)$$

$$\begin{aligned} \frac{\partial h_i}{\partial \theta} &= \frac{\partial \left(\alpha_0 + \sum_{j=1}^q \alpha_j \epsilon_{i-j}^2 \right)}{\partial \theta} \\ &= \frac{\partial \alpha_0}{\partial \theta} + \sum_{j=1}^q \left(\frac{\partial \alpha_j}{\partial \theta} \right) \cdot \epsilon_{i-j}^2 + \sum_{j=1}^q \alpha_j \left(\frac{\partial \epsilon_{i-j}^2}{\partial \theta} \right) \\ &= \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \epsilon_{i-1}^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \epsilon_{i-q}^2 \end{bmatrix} + \sum_{j=1}^q \begin{bmatrix} -2\epsilon_{i-j} \mathbf{x}_{i-j} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} -2 \sum_{j=1}^q \alpha_j \epsilon_{i-j} \mathbf{x}_{i-j} \\ \mathbf{z}_i(\boldsymbol{\beta}) \end{bmatrix} \end{aligned} \quad (7.3.5)$$

所以有

$$l_i(\theta) = \frac{\partial \ln f(y_i | \mathbf{x}_i, \mathbf{Y}_{i-1}; \theta)}{\partial \theta}$$

$$= - \left\{ \frac{1}{2h_t} - \frac{\epsilon_t^2}{2h_t^2} \right\} \begin{bmatrix} 2 \sum_{j=1}^q \alpha_{j\epsilon_{t-j}} \mathbf{x}_{t-j} \\ \mathbf{z}_t(\boldsymbol{\beta}) \end{bmatrix} - \begin{bmatrix} \mathbf{x}_t \epsilon_t / h_t \\ \mathbf{0} \end{bmatrix} \quad (7.3.6)$$

因此,可将对数似然函数 $L(\boldsymbol{\theta})$ 对参数 $\boldsymbol{\theta}$ 的微分表示成

$$\begin{aligned} \frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{s}(\boldsymbol{\theta}) &= \sum_{t=1}^T \mathbf{l}_t(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} \\ \frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\delta}} \end{bmatrix} = \begin{bmatrix} \mathbf{s}_1(\boldsymbol{\theta}) \\ \mathbf{s}_2(\boldsymbol{\theta}) \end{bmatrix} \\ &= \frac{1}{2} \sum_{t=1}^T \left\{ h_t^{-2} (\epsilon_t^2 - h_t) \begin{bmatrix} 2 \sum_{j=1}^q \alpha_{j\epsilon_{t-j}} \mathbf{x}_{t-j} \\ \mathbf{z}_t(\boldsymbol{\beta}) \end{bmatrix} + \begin{bmatrix} \mathbf{x}_t \epsilon_t / h_t \\ \mathbf{0} \end{bmatrix} \right\} \end{aligned} \quad (7.3.7)$$

参数向量 $\boldsymbol{\theta}$ 的最大似然估计 $\hat{\boldsymbol{\theta}}$ 为方程

$$\frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{0}$$

的解,使得 $L(\boldsymbol{\theta})$ 在 $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ 时取极大值。这一最优化问题可由数值计算方法解决。以下介绍的是恩格尔在 1982 年提出的算法。

由 (7.3.7) 式,子向量 $\mathbf{s}_1(\boldsymbol{\theta})$ 可表示为

$$\begin{aligned} \mathbf{s}_1(\boldsymbol{\theta}) &= \frac{1}{2} \sum_{t=1}^T \left\{ \frac{\epsilon_t^2}{h_t} - 1 \right\} \frac{\partial \ln(h_t)}{\partial \boldsymbol{\beta}} - \frac{1}{2} \sum_{t=1}^T \frac{1}{h_t} \cdot \frac{\partial \epsilon_t^2}{\partial \boldsymbol{\beta}} \\ &= \sum_{t=1}^T \left\{ \frac{\mathbf{x}_t \epsilon_t}{h_t} - \frac{1}{h_t} \left(\frac{\epsilon_t^2}{h_t} - 1 \right) \sum_{j=1}^q \alpha_{j\epsilon_{t-j}} \mathbf{x}_{t-j} \right\} \end{aligned} \quad (7.3.8)$$

整理后,又得

以此可定义相应于 $s_1(\theta)$ 的信息矩阵

$$I_\beta = -\frac{1}{T} E \left\{ \frac{\partial^2 L(\theta)}{\partial \beta \partial \beta'} \middle| x_i, Y_{i-1} \right\}$$

根据(7.3.5)式中的结果,可得

$$\frac{\partial \ln(h_i)}{\partial \beta} = -2 \sum_{j=1}^q \alpha_j \frac{\epsilon_i}{h_i} x_{i-j}$$

由于 ϵ_{i-j} 和 ϵ_{i-i} 在 $i \neq j$ 时不相关,所以信息矩阵 I_β 可由下式一致地估计:

$$\hat{I}_\beta = \frac{1}{T} \sum_{i=1}^T \left\{ \frac{x_i x_i'}{h_i} + 2 \sum_{j=1}^q \alpha_j^2 \frac{\epsilon_{i-j}^2}{h_i^2} x_{i-j} x_{i-j}' \right\} \quad (7.3.10)$$

将 $x_i x_i'$ 提到括弧外,作适当整理后,上式可等价地表示为

$$\begin{aligned} \hat{I}_\beta &= \frac{1}{T} \sum_{i=1}^T x_i x_i' \left\{ h_i^{-1} + 2 \epsilon_i^2 \sum_{j=1}^q \alpha_j^2 h_{i+j}^{-2} \right\} \\ &= \frac{1}{T} \sum_{i=1}^T x_i x_i' n_i \end{aligned} \quad (7.3.11)$$

上式与(7.3.9)式一样,在以下的计算中经常用到。

接下来对子向量 $s_2(\theta)$ 作类似的处理。根据(7.3.7)式

$$s_2(\theta) = \frac{\partial L(\theta)}{\partial \delta} = \frac{1}{2} \sum_{i=1}^T \left\{ \frac{\epsilon_i^2}{h_i} - 1 \right\} \frac{\partial \ln(h_i)}{\partial \delta} = \sum_{i=1}^T \frac{1}{2h_i} \frac{\partial \epsilon_i^2}{\partial \delta}$$

由(7.3.4)式,可知

$$\frac{\partial \epsilon_i^2}{\partial \delta} = 0$$

因此

$$s_2(\theta) = \frac{1}{2} \sum_{i=1}^T \left\{ \frac{\epsilon_i^2}{h_i} - 1 \right\} \cdot \frac{\partial \ln(h_i)}{\partial \delta}$$

对 $s_2(\theta)$ 求关于 δ' 的微分, 得

$$\begin{aligned}\frac{\partial s_2(\theta)}{\partial \delta'} &= \frac{\partial^2 L(\theta)}{\partial \delta \partial \delta'} = \frac{1}{2} \sum_{i=1}^T \frac{\partial \ln(h_i)}{\partial \delta} \cdot \frac{\partial}{\partial \delta'} \left\{ \frac{\epsilon_i^2}{h_i} - 1 \right\} \\ &\quad + \frac{1}{2} \sum_{i=1}^T \left\{ \frac{\epsilon_i^2}{h_i} - 1 \right\} \frac{\partial^2 \ln(h_i)}{\partial \delta \partial \delta'}\end{aligned}$$

由于

$$\frac{\partial \epsilon_i^2}{\partial \delta} = 0; \quad \frac{\partial z_i(\beta)}{\partial \delta'} = 0$$

因此

$$\frac{\partial^2 \ln(h_i)}{\partial \delta \partial \delta'} = 0$$

从而可将 $(\partial^2 L(\theta) / \partial \delta \partial \delta')$ 简化为

$$\frac{\partial^2 L(\theta)}{\partial \delta \partial \delta'} = -\frac{1}{2} \sum_{i=1}^T \frac{\epsilon_i^2}{h_i} \cdot \frac{\partial \ln(h_i)}{\partial \delta} \cdot \frac{\partial \ln(h_i)}{\partial \delta'}$$

将相应于 $s_2(\theta)$ 的信息矩阵记作 I_δ , 则

$$I_\delta = -\frac{1}{T} E \left\{ \frac{\partial^2 L(\theta)}{\partial \delta \partial \delta'} \mid x_i, Y_{i-1} \right\}$$

当给定 x_i 和 Y_{i-1} 时, 有

$$E(\epsilon_i^2 \mid x_i, Y_{i-1}) = h_i$$

又根据(7.3.5)式, 有

$$\frac{\partial \ln(h_i)}{\partial \delta} = \frac{1}{h_i} \cdot z_i(\beta)$$

因此, 信息矩阵 I_δ 有一致估计

$$\hat{\mathbf{I}}_{\delta} = \frac{1}{2T} \sum_{i=1}^T \mathbf{z}_i(\boldsymbol{\beta}) \mathbf{z}_i(\boldsymbol{\beta})' / h_i^2 \quad (7.3.12)$$

综合以上讨论,就可以构造计算参数 $\boldsymbol{\beta}$ 和 δ 的最大似然估计的迭代公式。以上标 i 表示相应的统计量在第 i 步迭代时的值, $\hat{\boldsymbol{\beta}}^{i+1}$ 和 $\hat{\delta}^{i+1}$ 有递推关系:

$$\begin{aligned} \hat{\boldsymbol{\beta}}^{i+1} &= \hat{\boldsymbol{\beta}}^i + [\hat{\mathbf{I}}_{\beta}^i]^{-1} \left(\frac{1}{T} \frac{\partial L(\boldsymbol{\theta})'}{\partial \boldsymbol{\beta}} \right) \\ \hat{\delta}^{i+1} &= \hat{\delta}^i + [\hat{\mathbf{I}}_{\delta}^i]^{-1} \left(\frac{1}{T} \frac{\partial L(\boldsymbol{\theta})'}{\partial \delta} \right) \end{aligned}$$

将(7.3.9)式至(7.3.12)式中的结果代入上式,可得 $\hat{\boldsymbol{\beta}}^{i+1}$ 和 $\hat{\delta}^{i+1}$ 的递推计算公式为

$$\begin{aligned} \hat{\boldsymbol{\beta}}^{i+1} &= \hat{\boldsymbol{\beta}}^i + \left(\sum_{i=1}^T \mathbf{x}_i \mathbf{x}_i' \hat{n}_i \right)^{-1} \left(\sum_{i=1}^T \mathbf{x}_i \hat{\varepsilon}_i \hat{m}_i \right) \\ \hat{\delta}^{i+1} &= \hat{\delta}^i + \left(\sum_{i=1}^T \mathbf{z}_i(\hat{\boldsymbol{\beta}}^i) \mathbf{z}_i(\hat{\boldsymbol{\beta}}^i)' / h_i^i \right)^{-1} \left(\sum_{i=1}^T \mathbf{z}_i(\hat{\boldsymbol{\beta}}^i) (\hat{\varepsilon}_i^2 - h_i^i) / h_i^i \right) \end{aligned} \quad (7.3.13)$$

这里, m_i 和 n_i 的表示式由(7.3.9)和(7.3.11)给出。

用以上两公式对 $\hat{\boldsymbol{\beta}}$ 和 $\hat{\delta}$ 作递推计算。首先用最小二乘法从回归模型 $y_i = \mathbf{x}_i' \boldsymbol{\beta} + \varepsilon_i$ 中得到 $\boldsymbol{\beta}$ 的估计值和估计残差,并将它们对应为迭代计算的初始值 $\hat{\boldsymbol{\beta}}^0$ 和 $\hat{\varepsilon}_i^0$ ($i = 1, 2, \dots, T$); 其次计算出估计值 $\hat{\delta}^0$ 和 h_i^0 ; 最后将初始值代入(7.3.13)中的两个迭代公式,即可得估计值 $\hat{\boldsymbol{\beta}}^1$ 和 $\hat{\delta}^1$ 。不断重复上述过程,直到达到预先给定的收敛准则为止。

以迭代方法得到的估计值 $\hat{\boldsymbol{\beta}}$ 和 $\hat{\delta}$ 是一致的估计,在一定的准则条件下,有正态的极限分布:

$$\begin{aligned}\sqrt{T}(\hat{\beta} - \beta) &\xrightarrow{d} N(0, I_{\beta}^{-1}) \\ \sqrt{T}(\hat{\delta} - \delta) &\xrightarrow{d} N(0, I_{\delta}^{-1})\end{aligned}\quad (7.3.14)$$

以上迭代方法的一个重要特点,是不需对未知参数 β 和 δ 作联合估计,最大似然估计 $\hat{\beta}$ 和 $\hat{\delta}$ 分别由(7.3.13)中的方程迭代计算得到,这样就简化了计算过程,但又不降低估计的效率和精确度。这种对 $\hat{\beta}$ 和 $\hat{\delta}$ 作分别计算的方法并不是在任何情况下都可行的。一般来说,由联合方法得到的 $\hat{\beta}$ 和 $\hat{\delta}$ 是最优的。只有在以下条件满足时,由(7.3.13)式分别迭代取得的估计值 $\hat{\beta}$ 和 $\hat{\delta}$ 才是等价最优的:

$$I_{\beta\delta} = -\frac{1}{T}E\left\{\frac{\partial^2 L(\theta)}{\partial \delta \partial \beta'} \mid x_t, Y_{t-1}\right\} = 0 \quad (7.3.15)$$

不难证明,上述条件对 ARCH(q)模型成立。

以上讨论基于假设 $\epsilon_t = \sqrt{h_t} \cdot v_t$ 有独立的正态分布。事实上,当样本量 T 足够大时,根据中心极限定理,以上的计算方法在作适当修改后仍能适合非正态分布的情况,只要这时

$$E(v_t \mid x_t, Y_{t-1}) = 0; \quad E(v_t^2 \mid x_t, Y_{t-1}) = 1$$

成立。换言之,非正态分布的 v_t 并不影响 $\hat{\theta} = [\hat{\beta}', \hat{\delta}']'$ 极限分布的正态性,只是此时的正态分布有不同的协方差矩阵

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} N(0, D^{-1} S D^{-1}) \quad (7.3.16)$$

其中

$$S = \text{plim}_{T \rightarrow \infty} T^{-1} \sum_{t=1}^T l_t(\theta) \cdot l_t'(\theta)$$

$l_t(\theta)$ 由 (7.3.4) 式给出。矩阵 D 为

$$\begin{aligned} D &= \text{plim}_{T \rightarrow \infty} T^{-1} \sum_{t=1}^T \left\{ -E \left[\frac{\partial l_t(\theta)}{\partial \theta'} \middle| x_t, Y_{t-1} \right] \right\} \\ &= \text{plim}_{T \rightarrow \infty} T^{-1} \sum_{t=1}^T \left\{ \frac{1}{2h_t^2} \begin{bmatrix} \sum_{j=1}^q -2\alpha_j \varepsilon_{t-j} x_{t-j} \\ z_t(\beta) \end{bmatrix} \right. \\ &\quad \left. \times \left[\sum_{j=1}^q -2\alpha_j \varepsilon_{t-j} x'_{t-j}, z_t(\beta)' \right] - \frac{1}{h_t} \begin{bmatrix} x_t x'_t & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix} \right\} \end{aligned}$$

若随机干扰 v_t 有独立的正态分布, 矩阵 D 和 S 相等, $\hat{\theta}$ 的极限分布为

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} N(\mathbf{0}, S^{-1})$$

这里的协方差矩阵为

$$\begin{aligned} S &= \text{plim}_{T \rightarrow \infty} T^{-1} \sum_{t=1}^T l_t(\theta) \cdot l_t(\theta)' = \text{plim}_{T \rightarrow \infty} T^{-1} E \left\{ \frac{\partial^2 L(\theta)}{\partial \theta \partial \theta'} \middle| x_t, Y_{t-1} \right\} \\ &= \begin{bmatrix} \text{plim}_{T \rightarrow \infty} I_\beta & \mathbf{0} \\ \mathbf{0} & \text{plim}_{T \rightarrow \infty} I_\delta \end{bmatrix} \end{aligned}$$

这与 (7.3.14) 中的结果是一致的。

当随机干扰 v_t 服从非正态分布时, 矩阵 D 和 S 之间出现差异, 这差异可用来修正估计值 $\hat{\beta}$ 和 $\hat{\delta}$ 的标准差, 使得在样本有限时提高估计的精确度。矩阵 S 和 D 分别有一致估计:

$$\begin{aligned} \hat{S} &= T^{-1} \sum_{t=1}^T l_t(\hat{\theta}) \cdot l_t(\hat{\theta})' \\ \hat{D} &= T^{-1} \sum_{t=1}^T \left\{ \frac{1}{2\hat{h}_t^2} \begin{bmatrix} \sum_{j=1}^q -2\hat{\alpha}_j \hat{\varepsilon}_{t-j} x_{t-j} \\ z_t(\hat{\beta}) \end{bmatrix} \right. \end{aligned}$$

$$\times \left\{ - \sum_{i=1}^q - 2 \hat{\alpha}_1 \hat{\varepsilon}_{i-1} x'_{i-1} z_i (\hat{\beta})' - \frac{1}{\hat{h}_i} \begin{bmatrix} x_i x'_i & 0 \\ 0 & 0 \end{bmatrix} \right\}$$

这里, $\hat{h}_i = \hat{\alpha}_0 + \hat{\alpha}_1 \hat{\varepsilon}_{i-1}^2 + \dots + \hat{\alpha}_q \hat{\varepsilon}_{i-q}^2$ 。这样, 根据最大似然估计 $\hat{\theta}$ 的极限分布 (7.3.15), $\hat{\theta}$ 的标准差等于矩阵 $(T^{-1} \hat{D}^{-1} \hat{S} \hat{D}^{-1})$ 对角线上元素的平方根。

7.4 非正态的 ARCH(q) 过程的最大似然估计

在实践中人们注意到, 许多时间序列, 特别是金融时间序列的无条件分布往往具有比正态分布更厚的尾部。而且在 ARCH(q) 过程

$$\varepsilon_t = \sqrt{h_t} \cdot v_t \quad (7.4.1)$$

中, 即使随机干扰 v_t 有正态的条件分布, ε_t 也可以有非正态的无条件分布, 有较厚的尾部分布。为了更精确地描述这些时间序列分布的尾部特征, 本节假设随机变量 ε_t 服从 t 分布。与正态分布一样, t 分布也是对称的。它的分布曲线与正态分布很相似, 只是有较厚的尾部分布。当 t 分布的自由度 k 足够大时, 它很接近于标准正态分布。不难证明, 当 $k \rightarrow \infty$ 时, t 分布的极限分布为标准正态分布。由于 t 分布的特殊性, ARCH(q) 过程 ε_t 不再表示成 (7.4.1) 的形式。表示 ARCH 作用的条件异方差 h_t 将被直接引入 ε_t 的方差中。 t 分布的密度函数为

$$f(\varepsilon_t) = \frac{\Gamma[(k+1)/2]}{\sqrt{\pi k} \cdot \Gamma(k/2)} c_t^{-1/2} \left[1 + \frac{\varepsilon_t^2}{c_t k} \right]^{-\frac{k+1}{2}} \quad (7.4.2)$$

其中, $\Gamma(\cdot)$ 为 Γ -函数, c_t 为一比例参数, k 是 t 分布的自由度, 为一正整数。当 $k > 2$ 时, ε_t 的期望值为零, 方差为

$$D(\epsilon_t) = c_t \frac{k}{k-2}$$

为在 ϵ_t 中引入 ARCH 效应, 令比例参数 c_t 为

$$c_t = h_t \cdot \frac{k-2}{k}$$

从而, ϵ_t 有条件方差 h_t , 密度函数(7.4.2)可改写为

$$f(\epsilon_t) = \frac{\Gamma[(k+1)/2]}{\sqrt{\pi} \cdot \Gamma(k/2)} (k-2)^{-1/2} h_t^{-1/2} \left[1 + \frac{\epsilon_t^2}{h_t(k-2)} \right]^{-\frac{k+1}{2}} \quad (7.4.3)$$

其中, h_t 可表示为

$$h_t = \alpha_0 + \alpha_1 \epsilon_1^2 + \alpha_2 \epsilon_2^2 + \cdots + \alpha_q \epsilon_{t-q}^2$$

与上节一样, 我们的目的是估计回归模型

$$y_t = \mathbf{x}_t' \boldsymbol{\beta} + \epsilon_t; \quad t = 1, 2, \dots, T$$

中的参数, 这里 ϵ_t 服从有 t 分布的 ARCH(q) 过程。相应于样本 y_1, y_2, \dots, y_T 的对数似然函数为

$$\begin{aligned} L(\theta) &= \sum_{t=1}^T \ln f(y_t | \mathbf{x}_t, Y_{t-1}; \theta) \\ &= T \ln \left\{ \frac{\Gamma[(k+1)/2]}{\sqrt{\pi} \cdot \Gamma(k/2)} (k-2)^{-\frac{1}{2}} \right\} - \frac{1}{2} \sum_{t=1}^T \ln(h_t) \\ &\quad - \frac{k+1}{2} \sum_{t=1}^T \ln \left[1 + \frac{(y_t - \mathbf{x}_t' \boldsymbol{\beta})^2}{h_t(k-2)} \right] \end{aligned}$$

其中需要估计的参数为 k 、 $\boldsymbol{\beta}$ 和 $\boldsymbol{\delta}$, 它们的最大似然估计可由与上节相似的数值方法计算, 只是 $L(\theta)$ 的一阶和二阶偏微分有更为复杂的形式。

上述方法也适用于有其他分布形式的 ARCH(q) 模型。比

$$I_{\delta} = - \frac{1}{T} E \left\{ \frac{\partial^2 L(\theta)}{\partial \delta \partial \delta'} \middle| x_i, Y_{i-1} \right\}$$

有一致估计

$$\hat{I}_{\delta} = \frac{1}{2T} \sum_{i=1}^T z_i(\theta) z_i(\theta)' / h_i^2$$

以 θ 的估计值 $\hat{\theta}$ 代入 $L(\theta)$ 的一阶和二阶偏导, 并记 h_0 、 ϵ_i^0 和 $z_i^0(\hat{\beta})$ 为 h_i 、 ϵ_i 和 $z_i(\theta)$ 在原假设 H_0 成立时的值, 则有

$$h_0 = \hat{a}_0 = \frac{1}{T} \sum_{i=1}^T e_i^2, \quad \epsilon_i^0 = e_i, \quad z_i^0(\hat{\beta}) = [1, e_{i-1}^2, \dots, e_{i-q}^2]'$$

这里 e_i 为 y_i 对 x_i 作回归的估计残差。因此在原假设成立时, 有

$$\frac{\partial L(\hat{\theta})}{\partial \delta} = \frac{1}{2h_0} \sum_{i=1}^T \left\{ \frac{e_i^2}{h_0} - 1 \right\} \cdot z_i^0(\hat{\beta})$$

$$\hat{I}_{\delta} = \frac{1}{2h_0^2 T} \sum_{i=1}^T z_i^0(\hat{\beta}) z_i^0(\hat{\beta})'$$

这样, 就可以构造检验 $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_q = 0$ 的 LM 检验统计量

$$\begin{aligned} \xi &= T \cdot \left(\frac{\partial L(\hat{\theta})}{\partial \delta} \right)' \hat{I}_{\delta}^{-1} \left(\frac{\partial L(\hat{\theta})}{\partial \delta} \right) \\ &= \frac{1}{2} \left\{ \sum_{i=1}^T \left[\frac{e_i^2}{h_0} - 1 \right] z_i^0(\hat{\beta}) \right\}' \left\{ \sum_{i=1}^T z_i^0(\hat{\beta}) z_i^0(\hat{\beta})' \right\}^{-1} \\ &\quad \times \left\{ \sum_{i=1}^T \left[\frac{e_i^2}{h_0} - 1 \right] z_i^0(\hat{\beta}) \right\} \end{aligned}$$

在 H_0 成立时, 统计量 ξ 有 $\chi^2(q)$ 极限分布。

分别定义矩阵 Z^v 和向量 f^v 为

(2) 构造 $f_t^0 = \frac{e_t^2}{h_0} - 1$, $z_t^0(\hat{\beta}) = [1, e_{t-1}^2, \dots, e_{t-q}^2]'$ 。以 f_t^0 对 $z_t^0(\hat{\beta})$ 作回归, 得回归的拟合优度 R^2 。

7.6 广义的 ARCH 模型——GARCH 模型

ARCH(q)模型 $\varepsilon_t = \sqrt{h_t} \cdot v_t$ 决定于条件异方差 h_t 的结构:

$$h_0 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_q \varepsilon_{t-q}^2$$

若 ARCH(q)模型的阶数 q 过大, 在样本有限时, 上节介绍的迭代估计方法的效率会降低。为弥补这一弱点, 以下介绍广义的 ARCH 模型 — GARCH 模型。

假设在 ARCH(q)过程

$$\varepsilon_t = \sqrt{h_t} \cdot v_t$$

中, $\{v_t\}$ 独立同分布, 且有 $v_t \sim N(0, 1)$, $t = 1, 2, \dots, T$ 。令 ARCH 过程的阶数 $q \rightarrow \infty$, 条件异方差 h_t 可表示为

$$h_t = \alpha_0 + \pi(L)\varepsilon_t^2$$

这里, $\pi(L)$ 为无穷阶滞后多项式:

$$\pi(L) = \sum_{j=1}^{\infty} \pi_j L^j \quad (7.6.1)$$

它可表示成两个有限阶滞后多项式的商:

$$\pi(L) = \frac{\alpha(L)}{1 - \rho(L)} = \frac{\alpha_1 L + \alpha_2 L^2 + \dots + \alpha_q L^q}{1 - \rho_1 L - \rho_2 L^2 - \dots - \rho_r L^r} \quad (7.6.2)$$

其中, 滞后多项式 $(1 - \rho(L))$ 的特征方程

$$1 - \rho_1 z - \rho_2 z^2 - \cdots - \rho_r z^r = 0$$

的根都在单位圆之外。利用(7.6.2)式可将 h_t 改写成

$$\begin{aligned} h_t = & k_0 + \rho_1 h_{t-1} + \rho_2 h_{t-2} + \cdots + \rho_r h_{t-r} \\ & + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-2}^2 + \cdots + \alpha_q \epsilon_{t-q}^2 \end{aligned} \quad (7.6.3)$$

其中的常数项 k_0 为

$$k_0 = (1 - \rho_1 - \rho_2 - \cdots - \rho_r) \alpha_0$$

由(7.6.3)式中 h_t 定义的 ARCH 过程 $\epsilon_t = \sqrt{h_t} \cdot v_t$ 称为广义的 ARCH 过程, 简称为 GARCH 过程, 记为 $\epsilon_t \sim \text{GARCH}(r, q)$ 。显然, 当 $r = 0$ 时, $\epsilon_t \sim \text{ARCH}(q)$; 当 $r = q = 0$ 时, ϵ_t 为一白噪声过程。将(7.6.1)式与(7.6.3)式比较, 可以得到系数 π_i 和 ρ_i 、 α_i 之间的关系为

$$\pi_i = \begin{cases} \alpha_i + \sum_{j=1}^n \rho_j \pi_{i-j}, & i = 1, \cdots, q \\ \sum_{j=1}^n \rho_j \pi_{i-j}, & i = q+1, \cdots \end{cases}$$

这里, $n = \min\{r, i-1\}$ 。以下的定理给出了 $\text{GARCH}(r, q)$ 过程是稳定过程的充要条件。

定理 7.1: 由(7.6.3)中的 h_t 定义的 $\text{GARCH}(r, q)$ 过程是稳定的随机过程, 并有 $E(\epsilon_t) = 0$ 、 $D(\epsilon_t) = k_0(1 - \alpha(1) - \rho(1))^{-1}$ 和 $\text{cov}(\epsilon_t, \epsilon_s) = 0 (t \neq s)$ 的充分必要条件为

$$\alpha(1) + \rho(1) < 1$$

这里 $\alpha(1) = \sum_{i=1}^q \alpha_i$, $\rho(1) = \sum_{i=1}^r \rho_i$ 。

证明: 在(7.6.3)式中以 $h_{t-r} \cdot v_t^2$ 代替 ϵ_{t-r}^2 , 可得

以此可推导出 $M(t, k)$ 的递推公式为

$$M(t, k+1) = \sum_{i=1}^r \rho_i M(t-i, k) + \sum_{i=1}^q \alpha_i v_{t-i}^2 M(t-i, k) \quad (7.6.4)$$

由于 $\{v_i^2\}$ 独立同分布, 因此 $M(t, k)$ 的各阶矩 (无条件矩) 与时刻 t 无关, 特别地, 有

$$E\{M(t, k)\} = E\{M(s, k)\} \quad (7.6.5)$$

根据以上两式, 可得

$$\begin{aligned} E\{M(t, k+1)\} &= \left\{ \sum_{i=1}^q \alpha_i + \sum_{i=1}^r \rho_i \right\} E\{M(t, k)\} \\ &= \dots\dots\dots \\ &= \left\{ \sum_{i=1}^q \alpha_i + \sum_{i=1}^r \rho_i \right\}^{k+1} E\{M(t, 0)\} \\ &= \left\{ \sum_{i=1}^q \alpha_i + \sum_{i=1}^r \rho_i \right\}^{k+1} \quad (7.6.6) \end{aligned}$$

最后, 计算 ϵ_t 的二阶矩为

$$\begin{aligned} E(\epsilon_t^2) &= k_0 E\left\{ \sum_{k=0}^{\infty} M(t, k) \right\} \\ &= k_0 \sum_{k=0}^{\infty} E\{M(t, k)\} = k_0 \sum_{k=0}^{\infty} (1 - \alpha(1) - \rho(1))^k \\ &= k_0 (1 - \alpha(1) - \rho(1))^{-1} \end{aligned}$$

上式成立的充要条件是

$$\alpha(1) + \rho(1) < 1$$

同理可证在此条件下, $E(\epsilon_t) = 0$, $\text{cov}(\epsilon_t, \epsilon_s) = 0 (t \neq s)$ 。因此

定理证毕。

7.6.1 GARCH(1, 1)过程

GARCH(1, 1)是最简单的一种 GARCH 过程。尽管形式简单,但它在经济学的许多领域,特别是在金融学中有很多应用。GARCH(1, 1)过程可表示为

$$\epsilon_t = \sqrt{h_t} \cdot v_t; h_t = k_0 + \rho_1 h_{t-1} + \alpha_1 \epsilon_{t-1}^2 \quad (7.6.7)$$

其中, $\{v_t\}$ 独立同分布, 且有 $v_t \sim N(0, 1)$, 参数满足条件 $k_0 > 0$, $\rho_1 \geq 0$, $\alpha_1 \geq 0$ 。根据定理 7.1, $\epsilon_t \sim \text{GARCH}(1, 1)$ 是稳定过程的充要条件为 $\rho_1 + \alpha_1 < 1$ 。

在有些应用中,不仅要求 $\epsilon_t \sim \text{GARCH}(1, 1)$ 是稳定过程,还要求它有高阶矩。以下定理给出 ϵ_t 存在 $2m$ 阶矩的充要条件。

定理 7.2: GARCH(1, 1) 过程 (7.6.7) 具有 $2m$ 阶矩的充分必要条件为

$$\mu(\rho_1, \alpha_1, m) = \sum_{j=0}^m \binom{m}{j} d_j \rho_1^{m-j} \alpha_1^j < 1 \quad (7.6.8)$$

这里, m 为一正整数; $d_0 = 1$, $d_j = \prod_{i=1}^j (2i-1)$, $j = 1, 2, \dots$ 。

而且, ϵ_t 的 $2m$ 阶矩满足递推公式

$$\begin{aligned} E(\epsilon_t^{2m}) &= d_m \left[\sum_{n=0}^m d_n^{-1} E(\epsilon_t^{2n}) k_0^{m-n} \binom{m}{m-n} \mu(\rho_1, \alpha_1, n) \right] \\ &\quad \times [1 - \mu(\rho_1, \alpha_1, m)]^{-1} \end{aligned}$$

证明: 由于 $\epsilon_t^{2m} = h_t^m v_t^{2m}$, $v_t \sim N(0, 1)$, 所以下列关系成立:

$$E(\epsilon_t^{2m}) = \prod_{j=1}^m (2j-1) E(h^m) = d_m E(h^m)$$

h^m 可由二项式公式展开为

$$\begin{aligned} h^m &= (k_0 + \rho_1 h_{t-1} + \varepsilon_{t-1}^2)^m \\ &= \sum_{n=0}^m \binom{m}{n} k_0^{m-n} \sum_{j=0}^n \binom{n}{j} \rho_1^{n-j} \alpha_j^1 h_{t-1}^{n-j} \varepsilon_{t-1}^{2j} \end{aligned}$$

上式中的 $h_{t-1}^{n-j} \varepsilon_{t-1}^{2j}$ 项有条件期望

$$E(h_{t-1}^{n-j} \varepsilon_{t-1}^{2j} | x_{t-1}, Y_{t-2}) = d_j h_{t-1}^n$$

因此

$$E(h_t^m | x_{t-1}, Y_{t-2}) = \sum_{n=0}^m h_{t-1}^n \binom{m}{n} k_0^{m-n} \sum_{j=0}^n d_j \rho_1^{n-j} \alpha_j^1 \quad (7.6.9)$$

令 $u_t = [h_t^m, h_t^{m-1}, \dots, h_t]^t$, 由上式可得

$$E(u_t | Y_{t-2}) = b + C u_{t-1}$$

其中, b 为 m 维常数向量, C 为 $m \times m$ 阶上三角矩阵, 其对角元素为

$$\mu(\rho_1, \alpha_1, i) = \sum_{j=0}^i \binom{i}{j} d_j \rho_1^{i-j} \alpha_j^1; \quad i = 1, 2, \dots, m. \quad (7.6.10)$$

由此可计算在给定 Y_{t-k-1} 时, 向量 u_t 的条件期望为

$$E(u_t | Y_{t-k-1}) = (I + C + C^2 + \dots + C^{k-1})b + C^k u_{t-k},$$

其中, C, \dots, C^k 为矩阵 C 的乘幂。令 $k \rightarrow \infty$, 上式存在极限的充要条件为 $C^k \rightarrow 0$, 这等价于矩阵 C 的所有特征值都在单位圆内。由于 C 为上三角阵, 其特征值等于它的对角元素, 所以对 $i = 1, 2, \dots, m$, 有

$$\mu(\rho_1, \alpha_1, i) < 1$$

不难验证, 上式意味着 $\mu(\rho_1, \alpha_1, i-1) < 1$, 因此 $\mu(\rho_1, \alpha_1, m) < 1$ 足以保证 ϵ_t 的 $2m$ 阶矩存在。定理证毕。

7.6.2 GARCH 回归模型的参数估计

与 ARCH 过程一样, GARCH 过程最常见的应用也是在回归分析中。以下考虑 GARCH 回归模型

$$\begin{aligned} y_t &= x_t' \beta + \epsilon_t; & \epsilon_t &= \sqrt{h_t} \cdot v_t \\ h_t &= k_0 + \sum_{i=1}^r \rho_i h_{t-i} + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2, & v_t &\sim N(0, 1) \end{aligned} \quad (7.6.11)$$

的参数最大似然估计。令

$$\begin{aligned} z_t &= [1, \epsilon_{t-1}^2, \epsilon_{t-2}^2, \dots, \epsilon_{t-q}^2, h_{t-1}, \dots, h_{t-r}]' \\ \delta &= [k_0, \alpha_1, \dots, \alpha_q, \rho_1, \dots, \rho_r]' \\ \theta &= [\beta', \delta']' \end{aligned}$$

可将 GARCH 模型(7.6.11)的对数似然函数表示为

$$L(\theta) = \sum_{t=1}^T l_t(\theta) = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T \ln(h_t) - \frac{1}{2} \sum_{t=1}^T \epsilon_t^2 h_t^{-1}$$

先对 $L(\theta)$ 求关于 δ 的一阶和二阶微分, 得

$$\frac{\partial L(\theta)}{\partial \delta} = \frac{1}{2} \sum_{t=1}^T h_t^{-1} \frac{\partial h_t}{\partial \delta} \left(\frac{\epsilon_t^2}{h_t} - 1 \right) \quad (7.6.12)$$

$$\frac{\partial^2 L(\theta)}{\partial \delta \partial \delta'} = \sum_{t=1}^T \left(\frac{\epsilon_t^2}{h_t} - 1 \right) \frac{\partial}{\partial \delta'} \left[\frac{1}{2} h_t^{-1} \frac{\partial h_t}{\partial \delta} \right] - \frac{1}{2} \sum_{t=1}^T h_t^{-2} \frac{\partial h_t}{\partial \delta} \frac{\partial h_t}{\partial \delta'} \frac{\epsilon_t^2}{h_t}$$

其中

$$\frac{\partial h_t}{\partial \delta} = z_t(\beta) + \sum_{i=1}^r \rho_i \frac{\partial h_{t-i}}{\partial \delta} \quad (7.6.13)$$

与 ARCH(q)模型的(7.3.5)式相比,上式中的 $(\partial h_t / \partial \delta)$ 多了第二项 $\left(\sum_{i=1}^r \rho_i \frac{\partial h_{t-i}}{\partial \delta} \right)$ 。

同时注意到

$$E \left\{ \sum_{i=1}^T \left(\frac{\epsilon_i^2}{h_i} - 1 \right) \frac{\partial}{\partial \delta'} \left[\frac{1}{2} h_i^{-1} \frac{\partial h_i}{\partial \delta} \right] \middle| x_i, Y_{i-1} \right\} = 0$$

因此, GARCH 回归模型(7.6.11)相应于参数 δ 的信息矩阵

$$I_\delta = - \frac{1}{T} E \left\{ \frac{\partial^2 L(\theta)}{\partial \delta \partial \delta'} \middle| x_i, Y_{i-1} \right\}$$

只需用 $\left(\frac{\partial h_t}{\partial \delta} \right)$ 就可得到一致的估计。

再对 $L(\theta)$ 求关于 β 的一阶和二阶微分,得

$$\begin{aligned} \frac{\partial L(\theta)}{\partial \beta} &= \sum_{i=1}^T \epsilon_i x_i h_i^{-1} + \frac{1}{2} \sum_{i=1}^T h_i \frac{\partial h_i}{\partial \beta} \left(\frac{\epsilon_i^2}{h_i} - 1 \right) \\ \frac{\partial^2 L(\theta)}{\partial \beta \partial \beta'} &= - \sum_{i=1}^T h_i^{-1} x_i x_i' - \frac{1}{2} \sum_{i=1}^T h_i^{-2} \frac{\partial h_i}{\partial \beta} \frac{\partial h_i}{\partial \beta'} \left(\frac{\epsilon_i^2}{h_i} \right) \\ &\quad - 2 \sum_{i=1}^T h_i^{-2} \epsilon_i x_i \frac{\partial h_i}{\partial \beta} + \sum_{i=1}^T \left(\frac{\epsilon_i^2}{h_i} - 1 \right) \frac{\partial}{\partial \beta'} \left[\frac{1}{2} h_i^{-1} \frac{\partial h_i}{\partial \beta} \right] \end{aligned} \quad (7.6.14)$$

其中

$$\frac{\partial h_t}{\partial \beta} = -2 \sum_{j=1}^q \alpha_j x_{t-j} \epsilon_{t-j} + \sum_{j=1}^r \rho_j \frac{\partial h_{t-j}}{\partial \beta}$$

与 ARCH(q) 模型的 (7.3.5) 式相比, $(\partial h_t / \partial \beta)$ 中多了 $\left(\sum_{j=1}^q \rho_j \partial h_{t-j} / \partial \beta \right)$ 这一项。

由于 h_t 对 β 和 δ 的一阶偏导中包含了自回归项 $\left(\sum_{j=1}^q \rho_j \partial h_{t-j} / \partial \beta \right)$ 和 $\left(\sum_{j=1}^q \rho_j \partial h_{t-j} / \partial \delta \right)$, 估计 ARCH(q) 模型的迭代公式不再适用。以下介绍的算法基于 BHHH 算法, 细节请参阅 Berndt, B. Hall, R. E. Hall and Hausman (1974)。

以 $\hat{\theta}^{i+1}$ 表示参数 $\theta = [\beta', \delta']'$ 在第 $i+1$ 步迭代时的估计值, $L(\theta) = \sum_{t=1}^T \ln(f_t(\theta))$, $\hat{\theta}^{i+1}$ 可由下式

$$\hat{\theta}^{i+1} = \hat{\theta}^i + \lambda \left(\sum_{t=1}^T \frac{\partial \ln'(f_t)}{\partial \theta} \cdot \frac{\partial \ln'(f_t)}{\partial \theta} \right)^{-1} \sum_{t=1}^T \frac{\partial \ln'(f_t)}{\partial \theta}$$

计算。这里, $(\partial \ln'(f_t) / \partial \theta)$ 表示一阶微分在 $\hat{\theta}^i$ 的值; λ 为一步长变量, 在给定的方向向量下, 它使似然函数 $L(\theta)$ 取极大值。方向向量可由一 T 维的单位向量 $[1, \dots, 1]'$ 对 $\partial \ln'(f_t) / \partial \theta$ ($t = 1, 2, \dots, T$) 作回归而得。

由于参数 β 和 δ 的分离性 (见 (7.3.14) 式及其讨论), 估计值 $\hat{\beta}^{i+1}$ 和 $\hat{\delta}^{i+1}$ 可由下列公式分别计算:

$$\begin{aligned} \hat{\beta}^{i+1} &= \hat{\beta}^i + \lambda \left(\sum_{t=1}^T \frac{\partial \ln'(f_t)}{\partial \beta} \cdot \frac{\partial \ln'(f_t)}{\partial \beta} \right)^{-1} \sum_{t=1}^T \frac{\partial \ln'(f_t)}{\partial \beta} \\ \hat{\delta}^{i+1} &= \hat{\delta}^i + \lambda \left(\sum_{t=1}^T \frac{\partial \ln'(f_t)}{\partial \delta} \cdot \frac{\partial \ln'(f_t)}{\partial \delta} \right)^{-1} \sum_{t=1}^T \frac{\partial \ln'(f_t)}{\partial \delta} \end{aligned} \quad (7.6.15)$$

由以上算法得到的迭代估计 $\hat{\theta} = [\hat{\beta}', \hat{\delta}']'$ 是一致的估计。当 $T \rightarrow \infty$, $\hat{\theta}$ 有正态的极限分布

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} N(\mathbf{0}, \mathcal{J}^{-1})$$

其中, $\mathcal{J} = -E(\partial^2 \ln(f_t) / \partial \theta \partial \theta')$ 。

当 GARCH 过程 $\varepsilon_t = \sqrt{h_t} \cdot v_t$ 中的随机干扰 v_t 有非正态的分布时, 迭代公式 (7.6.15) 也能适用, 只要这时有

$$E(\varepsilon_t | Y_{t-1}) = 0, \quad E(\varepsilon_t^2 h_t^{-1} | Y_{t-1}) = 1$$

$$E(\varepsilon_t^4 h_t^{-2} | Y_{t-1}) \leq M < \infty$$

迭代估计 $\hat{\theta} = [\hat{\beta}', \hat{\delta}']'$ 仍有正态的极限分布

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} N(\mathbf{0}, \mathcal{J}^{-1} F \mathcal{J}^{-1})$$

其中, $F = E(\partial \ln(f_t) / \partial \theta \cdot \partial \ln(f_t) / \partial \theta')$ 。当 $v_t \sim N(0, 1)$ 时, $F = \mathcal{J}$ 。

7.6.3 GARCH 模型的假设检验

对于 GARCH(r, q) 模型

$$y_t = x_t' \beta + \varepsilon_t; \quad \varepsilon_t = \sqrt{h_t} \cdot v_t, \quad v_t \sim N(0, 1)$$

$$h_t = k_0 + \sum_{i=1}^r \rho_i h_{t-i} + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 = z_t' \delta$$

可用 7.5 节讨论的检验方法, 对参数 δ 作假设检验。一般地, 可将参数 δ 分解为 $\delta = [\delta_1', \delta_2']'$, 将 h_t 表示成

$$h_t = z_t' \delta = z_{1t}' \delta_1 + z_{2t}' \delta_2$$

用拉格朗日检验法对原假设 $H_0: \delta_2 = \mathbf{0}$ 作检验。

根据 7.5 节中的讨论, 检验 $H_0: \delta_2 = \mathbf{0}$ 的 LM 统计量有如下形式:

$$\xi = \frac{1}{2} f^0{}' Z^0 (Z^0{}' Z^0)^{-1} Z^0{}' f^0$$

其中的向量 f^0 和矩阵 Z^0 分别为

$$f^0 = [e_1^2 h_1^{-1} - 1, \dots, e_T^2 h_T^{-1} - 1]'$$

$$Z^0 = \left[h_1 \frac{\partial h_1}{\partial \delta}, \dots, h_T \frac{\partial h_T}{\partial \delta} \right]'$$

并且都取值于 $\delta_2 = 0$ 。当原假设成立时,统计量 ξ 的极限分布为 $\chi^2(P)$, P 为子向量 δ_2 的维数。同样根据 7.5 节的讨论,统计量 ξ 的值可由辅助回归得到。计算回归

$$f^0 = Z^0 \pi + V$$

的拟合优度 R^2 ,那么 ξ 等价于

$$\xi' = TR^2$$

7.7 ARCH 模型的其他推广形式

除了广义的 ARCH 模型(GARCH)外,在过去几年中还出现了一些其他的 ARCH 模型的推广形式。以下对其中的几种作简要的介绍。

7.7.1 指数的 GARCH 模型(EGARCH)

奈尔逊(Nelson)在 1991 年提出了指数的 GARCH 模型(EGARCH),对 ARCH 作了进一步的推广。在 ARCH 过程

$$\epsilon_t = \sqrt{h_t} \cdot v_t$$

中,仍设 $\{v_t\}$ 为独立同分布, $E(v_t) = 0$, $D(v_t) = 1$, 并设条件方差 h_t 有如下形式:

$$\ln h_t = \alpha_0 + \sum_{j=1}^{\infty} \pi_j \{ |v_{t-j}| - E|v_{t-j}| + g \cdot v_{t-j} \} \quad (7.7.1)$$

则称 ϵ_t 服从 EGARCH 过程。

EGARCH 模型的一个重要特征是在条件方差 h_t 中引入了参数 g , 这使得 h_t 在随机干扰 v_t 取正、负值时有不同程度的变化, 从而能更准确地描述金融产品价格波动的情况。比如, 在股票市场中, 若将利多消息看作是对股价的正干扰, 将利空消息看作是负干扰, 人们注意到, 股价往往对同样程度的负干扰 (指绝对值相同) 的反应更强烈。这种对正负干扰反应的不对称性, 可由 EGARCH 模型描述, 如参数 g 为一负数, 且 $-1 < g < 0$, 那么一个负干扰 ($v_{t-j} < 0$) 所引起的 h_t 的变化, 比相同程度的正干扰 ($v_{t-j} > 0$) 所引起的要大; 反之, 若 $g > 0$, 同样程度的正干扰引起 h_t 的变化则更大; 若 $g = 0$, 则 h_t 对于正负干扰的变化是对称的。

由于 EGARCH 模型的条件方差 h_t 由指数形式表示, 所以无论 (7.7.1) 式中的参数 π_j ($j = 1, 2, \dots$) 取何实数, 条件方差 h_t 总是大于零的。这样, 在对 EGARCH 模型作参数估计时, 不需对 π_j 的值作约束, 从而可以减少很多计算量。可以验证, 当条件

$$\sum_{j=1}^{\infty} \pi_j^2 < \infty$$

成立时, EGARCH 过程 $\{\epsilon_t\}$ 为一稳定的随机过程。

在一般情况下, 可将无穷阶的滞后多项式 $\pi(L)$ 表示成两个有限阶的滞后多项式 $\alpha(L)$ 和 $1 - \rho(L)$ 的比, 即

$$\pi(L) = \frac{\alpha(L)}{1 - \rho(L)} = \frac{\alpha_1 L + \alpha_2 L^2 + \dots + \alpha_q L^q}{1 - \rho_1 L - \rho_2 L^2 - \dots - \rho_r L^r}$$

从而, $\ln h_t$ 有有限阶表示形式

$$\begin{aligned}\ln h_t = & k_0 + \rho_1 \ln h_{t-1} + \rho_2 \ln h_{t-2} + \cdots + \rho_q \ln h_{t-q} \\ & + \alpha_1 \{ |v_{t-1}| - E|v_{t-1}| + g \cdot v_{t-1} \} \\ & + \alpha_2 \{ |v_{t-2}| - E|v_{t-2}| + g \cdot v_{t-2} \} \\ & + \cdots + \alpha_q \{ |v_{t-q}| - E|v_{t-q}| + g \cdot v_{t-q} \} \quad (7.7.2)\end{aligned}$$

EGARCH 模型的参数可由最大似然方法估计。为了使 EGARCH 有最广泛的应用, 奈尔逊建议随机干扰 $\{v_t\}$ 服从广义误差分布 (generalized error distribution), 其密度函数为

$$f(v_t) = \frac{c \exp \left[-\frac{1}{2} |v_t/\lambda|^c \right]}{\lambda \cdot 2^{[(c+1)/c]} \Gamma(1/c)}; \quad 0 < c \leq \infty \quad (7.7.3)$$

其中, $\Gamma(\cdot)$ 为 Γ -函数, c 和 λ 均为常数, λ 的值为

$$\lambda = \left\{ \frac{2^{(-2/c)} \Gamma(1/c)}{\Gamma(3/c)} \right\}^{1/2}$$

它被称为尾部厚度参数 (tail-thickness parameter)。当参数 $c = 2$ 时, 随机干扰项 v_t 有正态分布; 当 $c < 2$ 时, v_t 的分布函数有较正态分布更厚的尾部; 当 $c > 2$ 时, v_t 的分布函数有较正态分布更薄的尾部; 特别当 $c = \infty$ 时, v_t 服从 $[-3^{1/2}, 3^{1/2}]$ 上的均匀分布。不难验证, 若 v_t 服从广义误差分布, 则有

$$E|v_t| = \frac{\lambda \cdot 2^{1/c} \Gamma(2/c)}{\Gamma(1/c)}$$

特别地, 当 $c = 2$ 时, 有

$$E|v_t| = \sqrt{\frac{2}{\pi}}$$

7.7.2 向量的 GARCH 模型

在 n 维的回归系统

$$y_t = \pi' x_t + \varepsilon_t \quad (7.7.4)$$

中, y_t 和 ε_t 都为 n 维随机向量, x_t 为 k 维回归变量, ε_t 为 n 维的独立同分布的随机向量。给定 $Y_{t-1} = [y_{t-1}, y_{t-2}, \dots, x_t, x_{t-1}, \dots]$, ε_t 的条件期望和方差分别为

$$E(\varepsilon_t | Y_{t-1}) = 0$$

$$E(\varepsilon_t \varepsilon_t' | Y_{t-1}) = H_t$$

随机向量 $\{\varepsilon_t\}$ 服从向量的 GARCH(r, q) 过程, 若条件方差 H_t 可表示为

$$\begin{aligned} H_t = & K + \Delta_1 H_{t-1} \Delta_1' + \Delta_2 H_{t-2} \Delta_2' + \dots + \Delta_r H_{t-r} \Delta_r' \\ & + A_1 \varepsilon_{t-1} \varepsilon_{t-1}' A_1' + A_2 \varepsilon_{t-2} \varepsilon_{t-2}' A_2' + \dots + A_q \varepsilon_{t-q} \varepsilon_{t-q}' A_q' \end{aligned}$$

其中 K 、 Δ_j 和 A_j ($j = 1, 2, \dots$) 均为 $n \times n$ 维系数矩阵。显然, 若矩阵 K 是对称正定的, 那么 H_t 也是正定的。

若向量 y_t 和 ε_t 的维数较大, 矩阵 Δ_j 和 A_j ($j = 1, 2, \dots$) 中的未知参数的个数就很大。因此, 有必要对这些未知参数作适当的约束, 使它们在样本量 T 有限时能被识别 (identified) 和估计。为了计算方便, 在许多情况下可假设参数矩阵 Δ_j 和 A_j ($j = 1, 2, \dots$) 为对角矩阵。在这样的约束条件下, 随机变量 ε_{it} 和 ε_{jt} 之间的条件协方差 $\text{cov}(\varepsilon_{it}, \varepsilon_{jt} | Y_{t-1})$ 只是 $(\varepsilon_{is, t-s}, \varepsilon_{js, t-s})$ 的函数 ($s = 1, 2, \dots$), 而不依赖于其他随机干扰 $\varepsilon_{ks, t-s}$ ($k \neq i, j$), 这使得参数估计大为简化。

另一个常用的简化参数矩阵 Δ_j 和 A_j ($j = 1, 2, \dots$) 的方法是假设随机向量 ε_t 中的各元素之间的相关系数不依赖于时

间 t , 尽管这时一般仍假设 ε_t 中每个元素的条件方差是时间的函数。以 $h_{ii}^{(t)}$ 表示矩阵 H_t 的第 i 行和第 i 列上的元素, 它是元素 ε_{it} 的条件方差, 即

$$h_{ii}^{(t)} = E(\varepsilon_{it}^2 | Y_{t-1})$$

若假设 ε_{it} 服从 GARCH(1, 1) 过程, 那么可将 $h_{ii}^{(t)}$ 表示为

$$h_{ii}^{(t)} = k_i + \delta_i h_{ii}^{(t-1)} + \alpha_i u_{it,t-1}^2$$

再对向量 ε_t 中的其他元素作同样的处理, 可得 n 个如上式的 GARCH(1, 1) 模型。以 $h_{ij}^{(t)}$ 表示矩阵 H_t 的第 i 行和第 j 列上的元素, 也即向量 ε_t 中的元素 ε_{it} 和 ε_{jt} 的条件协方差。若这时 ε_{it} 和 ε_{jt} 之间的相关系数

$$\rho_{ij} = \frac{\text{cov}(\varepsilon_{it}, \varepsilon_{jt} | Y_{t-1})}{\sqrt{D(\varepsilon_{it}) \cdot D(\varepsilon_{jt})}}$$

不依赖于时间 t , 则可将 $h_{ij}^{(t)}$ 表示为

$$\begin{aligned} h_{ij}^{(t)} &= \text{cov}(\varepsilon_{it}, \varepsilon_{jt} | Y_{t-1}) \\ &= \rho_{ij} \cdot \sqrt{h_{ii}^{(t)}} \cdot \sqrt{h_{jj}^{(t)}} \end{aligned}$$

这样, 矩阵 H_t 中的元素由其对角线上的元素 $h_{ii}^{(t)}$ 和 ρ_{ij} 所决定。

若进一步假设随机向量 ε_t 有多元正态分布, 向量 GARCH(r, q) 模型的参数一般可由最大似然方法作估计。在以上两种特殊情况下, 计算过程一般较为简单。

7.7.3 ARCH-M 模型

在前面的讨论中, ARCH、GARCH 和 EGARCH 过程主要描述回归模型的干扰项的条件方差, 一般与 y_t 的条件期望无关。但在实际中人们注意到, 条件方差的变化往往直接影响条件

$$\rho_1 = \frac{2\alpha_1^3\delta^2\alpha_0}{2\alpha_1^2\delta^2\alpha_0 + (1 - \alpha_1)(1 - 3\alpha_1^2)}$$

$$\rho_k = \alpha_1^{k-1}\rho_1; \quad k = 2, 3, \dots$$

显然,参数 α_0 、 α_1 和 δ 必须满足一定的约束条件,才能使以上三个表达式都具有意义。由于 $D(y_t)$ 、 ρ_1 和 ρ_k 都是参数 α_0 、 α_1 和 δ 的非线性函数,相应的约束条件也往往呈非线性,这就增加了参数估计的难度。为了简便起见,在实际应用中可取 $g(h_t) = \sqrt{h_t}$, 或 $g(h_t) = \ln h_t$ 。但对后者应多加小心,因为当 $h_t < 1$ 时, $\ln h_t$ 为负值,特别当 $h_t \rightarrow 0$ 时, $\ln h_t \rightarrow -\infty$ 。

7.8 ARCH 模型和随机微分方程

在现代金融学中,一般可假设某一金融资产的价格 Y_t 和它的收益(ΔY_t)的波动程度 σ_t 服从连续的随机过程,并可由连续的随机微分方程来描述。由于价格 Y_t 为正值,一般先对 Y_t 取对数,得 $y_t = \ln Y_t$ 。给定初始值 (y_0, σ_0) , y_t 和 σ_t 的联合运动过程可由以下微分方程表示:

$$dy_t = \left(\mu - \frac{\sigma_t^2}{2} \right) dt + \sigma_t dW_t \quad (7.8.1)$$

$$d(\ln \sigma_t^2) = -\beta(\ln \sigma_t^2 - \alpha)dt + \psi dV_t \quad (7.8.2)$$

其中, μ 、 ψ 、 β 和 α 为未知参数; W_t 和 V_t 为定义在 $[0, 1]$ 上的连续的维纳过程,并满足

$$\begin{bmatrix} dW_t \\ dV_t \end{bmatrix} [dW_t, dV_t] = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} dt$$

当 $\rho \neq 0$ 时,维纳过程 W_t 和 V_t 相关。

以下的讨论将表明,经过适当的离散化之后,(7.8.1)式和(7.8.2)式中连续的随机微分方程可由一个 EGARCH 模型近似地表示;反之,在适当的连续化后,一个 ARCH 模型也可由一个连续的随机微分方程近似地表示。

连续的随机微分方程有多种不同的离散化方法,以下介绍的是较为常用的欧拉近似方法。以 $h > 0$ 表示在时刻 t 的时间增量,当 h 足够小时, y_t 的微分 dy_t 可由 y_t 的一阶差分近似地表示为 $dy_t \approx y_{t+h} - y_t$ 。从而可将微分方程(7.8.1)和(7.8.2)近似地表示为

$$\begin{aligned} y_{t+h} &= y_t + \left(\mu - \frac{\sigma_t^2}{2} \right) h + h^{1/2} \sigma_t z_{1,t+h} \\ \ln \sigma_{t+h}^2 &= \ln \sigma_t^2 - h\beta(\ln \sigma_t^2 - \alpha) + h^{1/2} \psi z_{2,t+h} \end{aligned} \quad (7.8.3)$$

其中, t 取离散值 $h, 2h, 3h, \dots$; $(Z_{1,t}, Z_{2,t})'$ 服从期望值为零的二元正态分布,其协方差矩阵为

$$\begin{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} (z_{1,t}, z_{2,t}) \end{bmatrix} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \quad (7.8.4)$$

可以证明,在一定的准则条件下,当 $h \rightarrow 0$ 时,(7.8.3)中的随机差分方程组一致地收敛于随机微分方程组(7.8.1)和(7.8.3)。简单地说,由于(7.8.3)中的随机差分方程组的期望和协方差分别为

$$\begin{aligned} h^{-1} E_t \begin{bmatrix} y_{t+h} - y_t \\ \ln \sigma_{t+h}^2 - \ln \sigma_t^2 \end{bmatrix} &= \begin{bmatrix} \mu & -\frac{\sigma_t^2}{2} \\ -\beta(\ln \sigma_t^2 - \alpha) \end{bmatrix} \\ h^{-1} D \begin{bmatrix} y_{t+h} - y_t \\ \ln \sigma_{t+h}^2 - \ln \sigma_t^2 \end{bmatrix} &= \begin{bmatrix} \sigma_t^2 & \sigma_t \rho \psi \\ \sigma_t^2 \rho \psi & \psi^2 \end{bmatrix} \end{aligned} \quad (7.8.5)$$

这与 dy_t 和 $d\sigma_t$ 的期望值和协方差矩阵是一致的。以上两式对任

何 $h > 0$ 都成立。令 $h \rightarrow 0$, 就可得 (7.8.1) 和 (7.8.2) 中的随机微分方程组。

但 (7.8.3) 中的差分方程组还不具有 ARCH 模型的形式。我们的目的是要用第一个差分方程中的 $z_{1,t+h}$ 表示第二个方程中的随机项 $\psi z_{2,t+h}$ 。但只要 $z_{1,t+h}$ 和 $z_{2,t+h}$ 之间的相关系数 $\rho \neq 1$, σ_t^2 就不能完全由 $\{y_t, t=1, 2, \dots\}$ 来表示。另一方面, 只要 $\rho \neq 0$, 总可找到一个可测函数 $g(\cdot)$, 使得 $g(z_{1,t+h})$ 可以近似地表示 $\psi z_{2,t+h}$, 而且存在某个 $\delta > 0$, 使得

$$E(|g(z_{1,t+h})|^{2+\delta}) < \infty$$

$$D \begin{bmatrix} z_{1,t} \\ -g(z_{1t}) \end{bmatrix} = \begin{bmatrix} 1 & \rho\psi \\ \rho\psi & \psi^2 \end{bmatrix} \quad (7.8.6)$$

以 $g(z_{1,t+h})$ 代替 (7.8.3) 中第二个差分方程的 $\psi z_{2,t+h}$ 项, 就可得

$$\ln \sigma_{t+h}^2 = \ln \sigma_t^2 - h\beta(\ln \sigma_t^2 - \alpha) + h^{1/2}g(z_{1,t+h}) \quad (7.8.7)$$

其中的函数 $g(\cdot)$ 满足条件 (7.8.6)。这样的 $g(\cdot)$ 不难构造。比如, 因为 $E(|z_{1t}|) = (2/\pi)^{1/2}$, $D(|z_{1t}|) = 1 - \frac{2}{\pi}$, 从而可以构造如下函数:

$$g(z_{1t}) = \rho\psi z_{1t} + \psi \left(\frac{1 - \rho^2}{1 - 2/\pi} \right)^{1/2} \left[|z_{1t}| - \left(\frac{2}{\pi} \right)^{1/2} \right] \quad (7.8.8)$$

若将上式代入 (7.8.7) 式, 并与 (7.7.2) 式中的 $\ln h_t$ 相比较, 不难看出 (7.8.7) 有 EGARCH 表示形式。

以下再考虑用随机微分方程表示 ARCH 模型。为便于说明, 假设随机过程 $\{y_t\}$ 为一个单位根过程:

$$y_{t+h} = y_t + \epsilon_{t+h}$$

其中,随机干扰 ε_{t+h} 服从 GARCH(1, 1)过程,并有条件方差

$$\sigma_{t+h}^2 = D(\varepsilon_{t+h} | Y_t) = k_0 h + (1 - \theta h - \alpha h^{1/2}) \sigma_t^2 + h^{1/2} \alpha \varepsilon_t^2$$

给定 y_t 和 σ_t^2 的值,显然有

$$h^{-1} E \begin{bmatrix} y_{t+h} - y_t \\ \sigma_{t+h}^2 - \sigma_t^2 \end{bmatrix} = \begin{bmatrix} 0 \\ k_0 - \theta \sigma_t^2 \end{bmatrix} \quad (7.8.9)$$

和

$$h^{-1} D \begin{bmatrix} y_{t+h} - y_t \\ \sigma_{t+h}^2 - \sigma_t^2 \end{bmatrix} = \begin{bmatrix} \sigma_t^2 & 0 \\ 0 & 2\alpha^2 \sigma_t^4 \end{bmatrix} \quad (7.8.10)$$

(7.8.10)式成立是因为

$$h^{-1} D(\sigma_{t+h}^2 - \sigma_t^2) = h^{-1} E[(\sigma_{t+h}^2 - \sigma_t^2) - (k_0 - \theta \sigma_t^2)]^2$$

再令 $h \rightarrow 0$, $h^{-1}(y_{t+h} - y_t)$ 和 $h^{-1}(\sigma_{t+h}^2 - \sigma_t^2)$ 两式分别收敛于 dy_t/dt 和 $d\sigma_t^2/dt$,从而有

$$dy_t = \sigma_t dW_t$$

$$d\sigma_t^2 = (k_0 - \theta \sigma_t^2)dt + \sqrt{2} \alpha \sigma_t^2 dV_t$$

其中的 W_t 和 V_t 分别为定义在 $t \in [0, 1]$ 上的标准维纳过程。

ARCH 模型和连续的随机微分方程之间的互为近似的性质,使我们在实际研究工作中有了更大的灵活性。前者便于处理实际的金融数据,而后者便于理论上的研究。

7.9 ARCH 过程的综合

由于现代通讯技术和计算机技术、管理技术的发展,许多金融数据,如股价、兑换率等都以小时,甚至以分秒作为记录单位。这些高频率数据,提供了丰富的信息资源,使建立在大样本基础

上的统计推断和预测更为精确可靠。但另一方面,经济学家们有时更注重经济变量之间长期的均衡关系。用来描述这种长期关系的经济计量模型往往需要用低频率数据(如月、季度或年度数据)作参数估计和其他统计推断。时间序列的综合问题(aggregation)讨论这两种数据之间的关系,以及它们对经济计量模型的影响。

对于高频率时间序列 $y_t, t = 1, 2, \dots$, 一般可由以下两种方法构造相应的低频率时间序列:一是以一定的时间间隔在高频率时间序列 $y_t (t = 1, 2, \dots)$ 中抽取一个子序列 $y_t^{(m)} (t = m, 2m, 3m, \dots, m \text{ 为某一正整数})$;二是给定正整数 m , 作部分和 $\bar{y}_t^{(m)} = \sum_{i=0}^{m-1} y_{t-i} (t = m, 2m, 3m, \dots)$ 。由此产生的随机变量 $y_t^{(m)}$ 和 $\bar{y}_t^{(m)}$ 分别称为由 y_t 衍生的积变量(stock variable)和流变量(flow variable)。以下讨论的问题是,如果一个高频率的时间序列 $y_t (t = 1, 2, \dots)$ 服从 ARCH(或 GARCH、EGARCH)过程,由它所衍生的积变量 $y_t^{(m)}$ 和流变量 $\bar{y}_t^{(m)}$ 是否也服从同样的过程。若 $y_t^{(m)}$ 和 $\bar{y}_t^{(m)}$ 服从同样的过程,只是有不同的参数,则称 y_t 的分布是封闭的。下面以 GARCH(1, 1)为例讨论 GARCH 过程,对其他 ARCH 类模型可作类似的分析。

首先引入几个有关 GARCH 过程的定义。

若随机过程 $\{\epsilon_t\}$ 可表示为

$$\epsilon_t = \sqrt{h_t} \cdot v_t \quad (7.9.1)$$

其中, $\{v_t\}$ 独立同分布, 且有 $E(v_t) = 0, D(v_t) = 1, h_t = k_0 +$

$\sum_{i=1}^r \rho_i h_{t-i} + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2$, 则称 ϵ_t 服从强 GARCH(r, q)过程。

若(7.9.1)中的 ϵ_t 只满足条件

$$E(\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, \dots) = 0$$

$$E(\epsilon_t^2 | \epsilon_{t-1}, \epsilon_{t-2}, \dots) = h_t \quad (7.9.2)$$

则称 ϵ_t 服从半强 GARCH(r, q) 过程。

若以 $p[\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, \dots]$ 表示由元素 $1, \epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-1}^2, \epsilon_{t-2}^2, \dots$ 构成的对 ϵ_t 的最佳线性预测, 且 ϵ_t 满足条件

$$\begin{aligned} p[\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, \dots] &= 0 \\ p[\epsilon_t^2 | \epsilon_{t-1}, \epsilon_{t-2}, \dots] &= h_t \end{aligned} \quad (7.9.3)$$

则称 ϵ_t 服从弱 GARCH(r, q) 过程。不难验证, (7.9.3) 中的矩条件, 等价于

$$E\{(\epsilon_t - p[\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, \dots])\epsilon_{t-i}^r\} = 0 \quad (7.9.4)$$

其中 $i \geq 1, r = 0, 1, 2, \dots$ 。

显然, 强 GARCH(r, q) 过程是半强 GARCH(r, q) 过程的特例, 而它们都满足弱 GARCH(r, q) 的条件 (7.9.3)。

以下假设随机过程 ϵ_t 服从弱 GARCH(1, 1) 过程。我们分别考虑积变量 $\epsilon_t^{(m)}$ 和流变量 $\bar{\epsilon}_t^{(m)}$ 的分布情况。

定理 7.3: 若 ϵ_t 服从弱 GARCH(1, 1) 过程:

$$\epsilon_t = \sqrt{h_t} \cdot v_t; \quad h_t = k_0 + \rho_1 h_{t-1} + \alpha_1 \epsilon_{t-1}^2$$

其中, ϵ_t 满足矩条件 (7.9.3), 那么积变量 $\epsilon_t^{(m)}$ 在综合后是封闭的, 即 $\epsilon_t^{(m)}$ 也服从弱 GARCH(1, 1) 过程, 且有条件方差 $h_{tm}^{(m)}$ 为

$$h_{tm}^{(m)} = k_0^{(m)} + \rho_1^{(m)} h_{t-m}^{(m)} + \alpha_1^{(m)} (\epsilon_{t-m}^{(m)})^2$$

其中, m 为某一正整数,

$$k_0^{(m)} = k_0 \frac{1 - (\rho_1 + \alpha_1)^m}{1 - (\rho_1 + \alpha_1)}, \quad \alpha_1^{(m)} = (\rho_1 + \alpha_1)^m - \rho_1^{(m)} \quad (7.9.5)$$

$\rho_1^{(m)} \in (0, 1)$, 是以下方程

$$\frac{\rho_1^{(m)}}{1 + (\rho_1^{(m)})^2} = \frac{\rho_1(\rho_1 + \alpha_1)^{m-1}}{1 + \alpha_1^2 \frac{1 - (\rho_1 + \alpha_1)^{2m-2}}{1 - (\rho_1 + \alpha_1)^2} + \rho_1^2(\rho_1 + \alpha_1)^{2m-2}} \quad (7.9.6)$$

的解。

证明:下面只证明 $m = 2$ 的情况,对更一般的情况,证明方法类似。根据积变量 $\epsilon_t^{(m)}$ 的构造, $\{\epsilon_t^{(2)}\}$ 由元素 $\epsilon_2, \epsilon_4, \epsilon_6, \dots$ 组成。由于 ϵ_t 服从弱 GARCH(1, 1) 过程, 所以 $\epsilon_t (t = 2, 4, 6, \dots)$ 满足(7.9.3)中的第一个条件。根据(7.6.11), 可将 ϵ_t^2 等价地表示成

$$\epsilon_t^2 = k_0 + (\rho_1 + \alpha_1)\epsilon_{t-1}^2 + \eta_t - \rho_1\eta_{t-1} \quad (7.9.7)$$

其中 $\eta_t = \epsilon_t^2 - h_t, t = 1, 2, \dots$, 并且前后不相关。在上式中将 ϵ_{t-1}^2 往后迭代, 又得

$$\epsilon_t^2 = k_0(1 + \rho_1 + \alpha_1) + (\rho_1 + \alpha_1)^2\epsilon_{t-2}^2 + \delta_t \quad (7.9.8)$$

这里, $\delta_t = \eta_t + \alpha_1\eta_{t-1} - \rho_1(\rho_1 + \alpha_1)\eta_{t-2}$ 。只要能证明 ϵ_t^2 满足(7.9.3)中的第二个条件, 命题即可得证, 容易验证, 对于任何 $k > 1, E(\delta_t \cdot \delta_{t-2k}) = 0$ 。令

$$\omega_t = (1 - \lambda L^2)^{-1} \delta_t$$

其中 L 为滞后算子; λ 为常数, 使得偶数项的 ω_t 互不相关, 从而有

$$-\frac{\lambda}{1 + \lambda^2} = E(\delta_t \cdot \delta_{t-2})/E(\delta_t^2)$$

由于 ω_t 是 $\epsilon_t^2, \epsilon_{t-2}^2, \epsilon_{t-4}^2, \dots$ 的线性组合, 所以有

$$p[\epsilon_t^2 | \epsilon_{t-2}, \epsilon_{t-4}, \dots]$$

$$= \lambda p[\epsilon_{t-2}^2 | \epsilon_{t-4}, \epsilon_{t-6}, \dots] + k_0(1 + \rho_1 + \alpha_1) \\ + \{(\rho_1 + \alpha_1)^2 - \lambda\} \epsilon_{t-2}^2 \quad (7.9.9)$$

从而证明了 $\{\epsilon_{2t}, t=1, 2, \dots\}$ 服从弱 GARCH(1, 1) 过程, 且有参数

$$k_0^{(2)} = k_0(1 + \rho_1 + \alpha_1), \rho_1^{(2)} = \lambda, \alpha_1^{(2)} = (\rho_1 + \alpha_1)^2 - \lambda$$

以上的讨论表明, ARCH(1) 过程在综合后也是封闭的, 因为只需设 $\rho_1^{(m)} = 0$ 。注意到 (7.9.6) 式中的方程为一实数的二次方程, 且存在实数解, 这就保证了 $\rho_1^{(m)}$ 的存在。由 (7.9.5) 中的第二个关系式可知

$$\alpha_1^{(m)} + \rho_1^{(m)} = (\rho_1 + \alpha_1)^m$$

再由定理 7.1 可知

$$\rho_1 + \alpha_1 < 1$$

从而有

$$\lim_{m \rightarrow \infty} (\alpha_1^{(m)} + \rho_1^{(m)}) = \lim_{m \rightarrow \infty} (\rho_1 + \alpha_1)^m = 0$$

这说明, 随着综合度的提高 ($m \rightarrow \infty$), 积变量序列 $\epsilon_t^{(m)}$ 的条件方差的 ARCH 特征逐渐减弱。

定理 7.4: 在与定理 7.3 相同的条件下, 流变量序列 $\bar{\epsilon}_t^{(m)}$ 在综合后是封闭的, $\bar{\epsilon}_t^{(m)}$ 也服从弱 GARCH(1, 1) 过程, 且有条件方差

$$\bar{h}_{tm}^{(m)} = \bar{k}_0^{(m)} + \bar{\rho}_1^{(m)} \bar{h}_{tm-m}^{(m)} + \bar{\alpha}_1^{(m)} (\bar{\epsilon}_{tm-m}^{(m)})^2$$

其中

$$\bar{k}_0^{(m)} = mk_0 \frac{1 - (\rho_1 + \alpha_1)^m}{1 - (\rho_1 + \alpha_1)}, \quad \bar{\alpha}_1^{(m)} = (\rho_1 + \alpha_1)^m - \bar{\rho}_1^{(m)}$$

$\bar{\rho}_1^{(m)}$ 满足条件 $|\bar{\rho}_1^{(m)}| < 1$, 是以下二次方程

$$\begin{aligned} & \frac{\bar{\rho}_1^{(m)}}{1 + (\bar{\rho}_1^{(m)})^2} \\ &= \frac{A}{A(\rho_1, \alpha_1, C_\varepsilon, m)\{1 + (\rho_1 + \alpha_1)^{2m}\} - 2B(\rho_1, \alpha_1, m)} \end{aligned} \quad (7.9.10)$$

的解。参数 C_ε 为随机变量 ε_t 的峭度(kurtosis), 定义为

$$C_\varepsilon = \frac{E(\varepsilon_t^4)}{(E\varepsilon_t^2)^2}$$

函数 $A(\cdot)$ 和 $B(\cdot)$ 分别有表示式

$$\begin{aligned} & A(\rho_1, \alpha_1, C_\varepsilon, m) \\ &= m(1 - \rho_1)^2 + 2m(m-1) \frac{(1 - \rho_1 - \alpha_1)^2(1 - \rho_1^2 - 2\rho_1\alpha_1)}{(C_\varepsilon - 1)\{1 - (\rho_1 + \alpha_1)^2\}} \\ & \quad + 4 \frac{\{m-1-(\rho_1+\alpha_1)+(\rho_1+\alpha_1)^m\}\{\alpha_1 - \rho_1\alpha_1(\rho_1 + \alpha_1)\}}{1 - (\rho_1 + \alpha_1)^2} \\ & B(\rho_1, \alpha_1, m) = \{\alpha_1 - \rho_1\alpha_1(\rho_1 + \alpha_1)\} \frac{1 - (\rho_1 + \alpha_1)^{2m}}{1 - (\rho_1 + \alpha_1)^2} \end{aligned}$$

证明: 仍考虑 $m = 2$ 的特殊情况, 一般情况的证明可用同样方法得到。根据构造, 流变量 $\bar{\varepsilon}_t^{(2)}$ 显然满足(7.9.2)中的第一个条件

$$E(\bar{\varepsilon}_t^{(2)} | \bar{\varepsilon}_{t-2}^{(2)}, \bar{\varepsilon}_{t-4}^{(2)}, \dots) = 0$$

为了计算 $\bar{\varepsilon}_t^{(2)}$ 的条件方差, 令 $\eta_t = \varepsilon_t^2 - h_t$, 并以此代入(7.9.8)式, 可得

$$(\bar{\varepsilon}_t^{(2)})^2 = 2k_0(1 + \rho_1 + \alpha_1) + (\rho_1 + \alpha_1)^2(\bar{\varepsilon}_{t-2}^{(2)})^2 + \bar{\delta}_t$$

这里, $\bar{\varepsilon}_t^{(2)} = \varepsilon_t + \varepsilon_{t-1}$, $\bar{\delta}_t = \delta_t + \delta_{t-1}$ 。因为

$$\delta_t = \eta_t + \alpha_1 \eta_{t-1} - \rho_1(\rho_1 + \alpha_1) \eta_{t-2}$$

所以

$$\begin{aligned} \bar{\delta}_t &= \eta_t + (1 + \alpha_1) \eta_{t-1} + \{\alpha_1 - \rho_1(\rho_1 + \alpha_1)\} \eta_{t-2} \\ &\quad - \rho_1(\rho_1 + \alpha_1) \eta_{t-3} + 2\epsilon_t \epsilon_{t-1} - 2(\rho_1 + \alpha_1)^2 \epsilon_{t-2} \epsilon_{t-3} \end{aligned}$$

显然, $\bar{\delta}_t$ 中的各项前后不相关, 因此对任何大于 1 的正整数 k , 有

$$E(\bar{\delta}_t \cdot \bar{\delta}_{t-2k}) = 0$$

令 $\bar{\omega}_t = (1 - \bar{\lambda}L^2)^{-1} \bar{\delta}_t$, 其中 L 为滞后算子, $\bar{\lambda}$ 为选定的常数, 使得 $\bar{\omega}_t$ 互不相关。从而有

$$\{\bar{\epsilon}_t^{(2)}\}^2 = 2k_0(1 + \rho_1 + \alpha_1) + (\rho_1 + \alpha_1)^2 \{\bar{\epsilon}_{t-2}^{(2)}\}^2 + \bar{\omega}_t - \bar{\lambda} \bar{\omega}_{t-2} \quad (7.9.11)$$

运用推导关系式(7.9.9)的方法, 对上式作投影, 即可证明 $\{\bar{\epsilon}_t^{(2)}\}^2$ 服从 GARCH(1, 1) 过程, 其中的参数为

$$\bar{k}_0^{(2)} = 2k_0(1 + \rho_1 + \alpha_1), \quad \bar{\rho}_1^{(2)} = \bar{\lambda}, \quad \bar{\alpha}_1^{(2)} = (\rho_1 + \alpha_1)^2 - \bar{\lambda}$$

与积变量相比, 关于流变量 $\bar{\epsilon}_t^{(m)}$ 的结果和证明都要复杂得多, 这是因为 δ_t 中含有交叉项 $(\epsilon_t \epsilon_{t-1})$ 和 $(\epsilon_{t-2} \epsilon_{t-3})$, 从而使得 $\bar{\lambda}$ 值的计算较为繁琐, 在此不作详细介绍。

值得指出的是, 若在以上的结果(7.9.11)中令 $\rho_1 = 0$, 即假设 ϵ_t 服从 ARCH(1) 过程, 但(7.9.11)式并不由此简化为 ARCH(1) 过程。这说明, ARCH(1) 过程对于流变量在综合后是不封闭的, 这与积变量形成明显的对照。

定理 7.5: 强 GARCH(r, q) 过程

$$\epsilon_t = \sqrt{h_t} \cdot v_t$$

其中, $\{v_t\}$ 为独立同分布, 且 $v_t \sim N(0, 1)$, $h_t = k_0 + \sum_{i=1}^r \rho_i h_{t-i} + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2$, 对于综合是不封闭的。

证明: 考虑 $m = 2, r = 0, q = 1$ 的积变量的情况, 并假设 ϵ_t 服从 ARCH(1) 过程:

$$\epsilon_t = \sqrt{h_t} \cdot v_t = \sqrt{k_0 + \alpha_1 \epsilon_{t-1}^2} \cdot v_t$$

将

$$\epsilon_{t-1}^2 = (k_0 + \alpha_1 \epsilon_{t-2}^2) \cdot v_{t-1}^2$$

代入 h_t , 可将 ϵ_t^2 改写成

$$\begin{aligned} \epsilon_t^2 &= h_t \cdot v_t^2 = (k_0 + \alpha_1 \epsilon_{t-1}^2) \cdot v_t^2 \\ &= [k_0 + \alpha_1 (k_0 + \alpha_1 \epsilon_{t-2}^2) \cdot v_{t-1}^2] \cdot v_t^2 \end{aligned}$$

给定 ϵ_{t-2}, \dots 的值, 求 ϵ_t^2 的条件期望, 从而可得积变量 $\epsilon_t^{(2)}$ 的方差

$$h_t^{(2)} = E(\epsilon_t^2 | \epsilon_{t-2}, \dots) = k_0(1 + \alpha_1) + \alpha_1^2 \epsilon_{t-2}^2$$

令 $v_t^{(2)} = \epsilon_t / \sqrt{h_t^{(2)}}$, 则有

$$v_t^{(2)} = \epsilon_t / \sqrt{h_t^{(2)}} = \sqrt{\frac{k_0}{h_t^{(2)}} + \frac{\alpha_1}{h_t^{(2)}} \cdot \epsilon_{t-1}^2} \cdot v_t$$

再令 $\lambda_t = k_0/h_t^{(2)}$, 并注意到 $\epsilon_{t-1}^2 = (k_0 + \alpha_1 \epsilon_{t-2}^2) \cdot v_{t-1}^2$, 因此有

$$\begin{aligned} v_t^{(2)} &= \sqrt{\frac{k_0}{h_t^{(2)}} + \frac{\alpha_1 (k_0 + \alpha_1 \epsilon_{t-2}^2)}{h_t^{(2)}} v_{t-1}^2} \cdot v_t \\ &= \sqrt{\lambda_t + (1 - \lambda_t) v_{t-1}^2} \cdot v_t \end{aligned} \quad (7.9.12)$$

7.10 本章小结

本章介绍了 ARCH、GARCH 和 EGARCH 的一些主要的统计性质以及参数估计和假设检验的一些基本方法。ARCH(以及 GARCH、EGARCH)过程在满足一定的约束条件时是稳定过程,它的无条件期望和方差均为常数。但是它的条件方差却随时间变化,是过去的随机干扰的函数。这使得 ARCH 模型适合于描述某些随机变量(如股价、汇率等),这些变量常在某一时段内持续呈现较大或较小的条件方差。

ARCH 过程的参数结构较为复杂,一般需用数值方法计算参数的估计值。本章介绍的恩格尔方法是较常用的一种数值计算方法,已被许多经济计量软件,如 PcGive 和 Microfit 等采用。

ARCH 过程和连续的随机微分方程之间有着本质上的联系。在一定条件下,它们互为近似,这给研究工作带来了很大的灵活性。前者适用于处理离散的实际数据,后者则便于理论上的推导。

在实际应用 ARCH 模型时,必须注意数据的特点,如数据的间隔是天、星期、月或季度等,以及 ARCH 过程的综合性质。本章的讨论表明,当高频率的随机变量服从 ARCH 过程时,经过综合后的低频率变量(积变量或流变量)不一定再服从 ARCH 过程。

习 题

1. 随机变量 X 服从广义误差分布,若它的密度函数 $f(x)$ 可表示为

$$f(x) = \frac{c \cdot \exp\left[-\frac{1}{2} |x/\lambda|^c\right]}{\lambda \cdot 2^{[(c+1)/2]} \Gamma(1/c)}$$

其中, $0 < c \leq \infty$, $\Gamma(\cdot)$ 为 Γ -函数, λ 为常数, 其值为

$$\lambda = \left\{ \frac{2^{(-2/c)} \cdot \Gamma(1/c)}{\Gamma(3/c)} \right\}^{1/2}$$

验证:

- (a) 当 $c = 2$ 时, 随机变量 X 有标准正态分布。
 - (b) 当 $c < 2$ 时, X 的密度函数有较正态分布更厚的尾部。
 - (c) 当 $c > 2$ 时, X 的密度函数有较正态分布更薄的尾部。
 - (d) 当 $c \rightarrow \infty$ 时, X 的极限分布是 $[-3^{1/2}, 3^{1/2}]$ 上的均匀分布。
2. 若随机变量 X 服从广义误差分布, 证明:

$$E|X| = \frac{\lambda 2^{1/c} \Gamma(2/c)}{\Gamma(1/c)}$$

然后利用 Γ -函数的性质

$$\Gamma(2/c) = (2\pi)^{-1/2} \cdot 2^{2/c-1/2} \Gamma(1/c) \Gamma\left(1/c + \frac{1}{2}\right)$$

$$\Gamma(3/c) = (2\pi)^{-1} 3^{3/c-1/2} \Gamma(1/c) \Gamma\left(1/c + \frac{1}{3}\right) \Gamma\left(1/c + \frac{2}{3}\right)$$

证明:

$$\lim_{c \rightarrow \infty} \frac{\lambda \cdot 2^{1/c} \Gamma(2/c)}{\Gamma(1/c)} = \frac{\sqrt{3}}{2}$$

并与上题的结论(d)作比较。

3. 以 $p[\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, \dots]$ 表示由元素 $1, \epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-1}^2, \epsilon_{t-2}^2, \dots$ 构成的对 ϵ_t 的最佳线性预测, 证明:

$$E\{(\epsilon_t - p[\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, \dots]) \epsilon_{t-r}^2\} = 0$$

这里, $i \geq 1, r = 0, 1, 2$ 。

参考文献

1. Anderson, T. W. (1958), *An Introduction to Multivariate Statistical Analysis*, New York: John Wiley & Sons.
2. Banajee, A. , Dolado, J. , Galbraith, J. W. and Hendry, D. F. (1994), *Co-integration, Error Correction, and the Econometric Analysis of Non-Stationary Data*, Oxford University Press.
3. Beveridge, S. and Nelson, C. R. (1981), "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of 'Business Cycle'", *Journal of Monetary Economics*, 7, pp. 151 -174.
4. Berndt, E. K. , Hall, B. , Hall, R. E. and Hausman, J. A. (1974), "Estimation and Inference in Non-linear Structural Models", *Annals of Economic and Social Measurement*, 3, pp. 653—665.
5. Billingsley, P. (1968), *Convergence of Probability Measures*, New York: John Wiley & Sons.
6. Dickey, D. A. and Fuller, W. A. (1979), "Distribution of the Estimators of Autoregressive Time Series with a Unit Root", *Journal of the American Statistical Association*, 74, pp. 427—431.
7. Dickey, D. A. and Fuller, W. A. (1981), "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root", *Econometrica*, 49, pp. 355—367.
8. Engle, R. F. (1982), "Autoregressive Conditional Heteroskedas-

- ticity, with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, 50, pp. 987—1007.
9. Engle, R. F. (1995), *ARCH: Selected Readings*, Oxford University Press.
 10. Engle, R. F. and Granger, C. W. J. (1987), "Cointegration and Error Correction: Representation, Estimation and Testing", *Econometrica*, 55, pp. 119—139.
 11. Granger, C. W. J. and Newbold, P. (1974), "Spurious Regression in Econometrics", *Journal of Econometrics*, 2, pp. 111—120.
 12. Hall, P. and Heyde, C. C. (1980), *Martingale Limit Theory and its Applications*, London: Academic Press.
 13. Hamilton, J. D. (1994), *Time Series Analysis*, Princeton University Press.
 14. Hendry, D. F. (1995), *Dynamic Econometrics*, Oxford University Press.
 15. Hsieh, D. A. (1989), "Modelling Heteroskedasticity in Daily Foreign-Exchange Rates", *Journal of Business and Economic Statistics*, 7, pp. 307—317.
 16. Jorion, P. (1988), "On Jump Processes in the Foreign Exchange and Stock Markets", *Review of Financial Studies*, 1, pp. 427—445.
 17. Johansen, S. (1988), "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, 12, 231—254.
 18. Johansen, S. (1992a), "Cointegration in Partial Systems and the Efficiency of Single-equation Analysis", *Journal of Econometrics*, 52, pp. 389—402.
 19. Johansen, S. (1992b), "A Representation of Vector Autoregressive Processes Integrated of Order 2", *Econometric Theory*, 8, pp. 188—202.
 20. Johansen, S. (1995), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University press.

21. Nelson, D. B. (1990), "Stationarity and Persistence in the GARCH (1, 1) Model", *Econometric Theory*, 6, pp. 318—334.
22. Nelson, D. B. (1991), "Conditional Heteroskedasticity in Asset Returns: A New Approach", *Econometrica*, 59, pp. 347—370.
23. Newey, W. K. and West, K. D. (1987), "A Simple Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix", *Econometrica*, 55, pp. 703—708.
24. Phillips, P. C. B. (1986), "Understanding Spurious Regressions in Econometrics", *Journal of Econometrics*, 33, pp. 311—340.
25. Phillips, P. C. B. (1987), "Time Series Regression with a Unit Root", *Econometrica*, 55, pp. 277—301.
26. Phillips, P. C. B. (1988), "Regression Theory for Near-integrated Time Series", *Econometrica*, 56, pp. 1021—1043.
27. Phillips, P. C. B. (1991), "Optimal Inference in Cointegrated Systems", *Econometrica*, 59, pp. 283—306.
28. Phillips, P. C. B. and Ouliaris, S. (1990), "Asymptotic Properties of Residual Based Tests for Cointegration", *Econometrica*, 58, pp. 165—193.
29. Phillips, P. C. B. and Perron, P. (1988), "Testing for a Unit Root in Time Series Regression", *Biometrika*, 75, pp. 335—346.
30. Rozanov, Y. A. (1967), *Stationary Random Process*, San Francisco: Holden-Day.

表 1 菲利普斯—配荣的 Z_ρ 检验和迪基—福勒检验
(基于 OLS 估计值)的临界值

样本量 T	统计量 $T(\hat{\rho}_T - 1)$ 小于表中数值的概率							
	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
情况一								
25	-11.9	-9.3	-7.3	-5.3	1.01	1.40	1.79	2.28
50	-12.9	-9.9	-7.7	-5.5	0.97	1.35	1.70	2.16
100	-13.3	-10.2	-7.9	-5.6	0.95	1.31	1.65	2.09
250	-13.6	-10.3	-8.0	-5.7	0.93	1.28	1.62	2.04
500	-13.7	-10.4	-8.0	-5.7	0.93	1.28	1.61	2.04
∞	-13.8	-10.5	-8.1	-5.7	0.93	1.28	1.60	2.03
情况二								
25	-17.2	-14.6	-12.5	-10.2	-0.76	0.01	0.65	1.40
50	-18.9	-15.7	-13.3	-10.7	-0.81	-0.07	0.53	1.22
100	-19.8	-16.3	-13.7	-11.0	-0.83	-0.10	0.47	1.14
250	-20.3	-16.6	-14.0	-11.2	-0.84	-0.12	0.43	1.09
500	-20.5	-16.8	-14.0	-11.2	-0.84	-0.13	0.42	1.06
∞	-20.7	-16.9	-14.1	-11.3	-0.85	-0.13	0.41	1.04
情况四								
25	-22.5	-19.9	-17.9	-15.6	-3.66	-2.51	-1.53	-0.43
50	-25.7	-22.4	-19.8	-16.8	-3.71	-2.60	-1.66	-0.65
100	-27.4	-23.6	-20.7	-17.5	-3.74	-2.62	-1.73	-0.75
250	-28.4	-24.4	-21.3	-18.0	-3.75	-2.64	-1.78	-0.82
500	-28.9	-24.8	-21.5	-18.1	-3.76	-2.65	-1.78	-0.84
∞	-29.5	-25.1	-21.8	-18.3	-3.77	-2.66	-1.79	-0.87

资料来源: W. A. Fuller, *Introduction to Statistical Time Series*, New York: Wiley, 1976, p. 371.

表2 菲利普斯—配荣的 Z_t 检验和
迪基—福勒的 t 检验的临界值

样本量	统计量 $T(\hat{\rho}_T - 1)/\hat{\sigma}_\varepsilon$ 小于表中数值的概率							
T	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
情况一								
25	-2.66	-2.26	-1.95	-1.60	0.92	1.33	1.70	2.16
50	-2.62	-2.25	-1.95	-1.61	0.91	1.31	1.66	2.08
100	-2.60	-2.24	-1.95	-1.61	0.90	1.29	1.64	2.03
250	-2.58	-2.23	-1.95	-1.62	0.89	1.29	1.63	2.01
500	-2.58	-2.23	-1.95	-1.62	0.89	1.28	1.62	2.00
∞	-2.58	-2.23	-1.95	-1.62	0.89	1.28	1.62	2.00
情况二								
25	-3.75	-3.33	-3.00	-2.63	-0.37	0.00	0.34	0.72
50	-3.58	-3.22	-2.93	-2.60	-0.40	-0.03	0.29	0.66
100	-3.51	-3.17	-2.89	-2.58	-0.42	-0.05	0.26	0.63
250	-3.46	-3.14	-2.88	-2.57	-0.42	-0.06	0.24	0.62
500	-3.44	-3.13	-2.87	-2.57	-0.43	-0.07	0.24	0.61
∞	-3.43	-3.12	-2.86	-2.57	-0.44	-0.07	0.23	0.60
情况四								
25	-4.38	-3.95	-3.60	-3.24	-1.14	-0.80	-0.50	-0.15
50	-4.15	-3.80	-3.50	-3.18	-1.19	-0.87	-0.58	-0.24
100	-4.04	-3.73	-3.45	-3.15	-1.22	-0.90	-0.62	-0.28
250	-3.99	-3.69	-3.43	-3.13	-1.23	-0.92	-0.64	-0.31
500	-3.98	-3.68	-3.42	-3.13	-1.24	-0.93	-0.65	-0.32
∞	-3.96	-3.66	-3.41	-3.12	-1.25	-0.94	-0.66	-0.33

资料来源:W. A. Fuller, *Introduction to Statistical Time Series*, New York, Wiley, 1976, p. 373.

表 3 迪基—福勒 F 检验的临界值

样本量 T	F 统计量大于表中数值的概率							
	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01
情况二								
(F test of $\alpha=0, \rho=1$ in regression $y_t = \alpha + \rho y_{t-1} + u_t$)								
25	0.29	0.38	0.49	0.65	4.12	5.18	6.30	7.88
50	0.29	0.39	0.50	0.66	3.94	4.86	5.80	7.06
100	0.29	0.39	0.50	0.67	3.86	4.71	5.57	6.70
250	0.30	0.39	0.51	0.67	3.81	4.63	5.45	6.52
500	0.30	0.39	0.51	0.67	3.79	4.61	5.41	6.47
∞	0.30	0.40	0.51	0.67	3.78	4.59	5.38	6.43
情况四								
(F test of $\delta = 0, \rho = 1$ in regression $y_t = \alpha + \delta t + \rho y_{t-1} + u_t$)								
25	0.74	0.90	1.08	1.33	5.91	7.24	8.55	10.61
50	0.76	0.93	1.11	1.37	5.61	6.73	7.81	9.31
100	0.76	0.94	1.12	1.38	5.47	6.49	7.44	8.73
250	0.76	0.94	1.13	1.39	5.39	6.34	7.25	8.43
500	0.76	0.94	1.13	1.39	5.36	6.30	7.20	8.34
∞	0.77	0.94	1.13	1.39	5.34	6.25	7.16	8.27

资料来源: D. A. Dickey and W. A. Fuller, "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root", *Econometrica*, 49, p. 1063.

表 5 检验同积关系的菲利普斯—配荣的 Z 检验
和迪基—福勒 t 检验的临界值

不包括常 数项或时 间趋势的 回归变量 的个数减 1 ($n-1$)		样本量 (T)	统计量 $(\hat{\rho}_T - 1)/\hat{\sigma}_\rho$ 小于表中数值的概率						
			0.010	0.025	0.050	0.075	0.100	0.125	0.150
情况一									
1	500	-3.39	-3.05	-2.76	-2.58	-2.45	-2.35	-2.26	
2	500	-3.84	-3.55	-3.27	-3.11	-2.99	-2.88	-2.79	
3	500	-4.30	-3.99	-3.74	-3.57	-3.44	-3.35	-3.26	
4	500	-4.67	-4.38	-4.13	-3.95	-3.81	-3.71	-3.61	
5	500	-4.99	-4.67	-4.40	-4.25	-4.14	-4.04	-3.94	
情况二									
1	500	-3.96	-3.64	-3.37	-3.20	-3.07	-2.96	-2.86	
2	500	-4.31	-4.02	-3.77	-3.58	-3.45	-3.35	-3.26	
3	500	-4.73	-4.37	-4.11	-3.96	-3.83	-3.73	-3.65	
4	500	-5.07	-4.71	-4.45	-4.29	-4.16	-4.05	-3.96	
5	500	-5.28	-4.98	-4.71	-4.56	-4.43	-4.33	-4.24	
情况三									
1	500	-3.98	-3.68	-3.42	—	-3.13	—	—	
2	500	-4.36	-4.07	-3.80	-3.65	-3.52	-3.42	-3.33	
3	500	-4.65	-4.39	-4.16	-3.98	-3.84	-3.74	-3.66	
4	500	-5.04	-4.77	-4.49	-4.32	-4.20	-4.08	-4.00	
5	500	-5.36	-5.02	-4.74	-4.58	-4.46	-4.36	-4.28	

资料来源: P. C. B. Phillips and S. Ouliaris, "Asymptotic Properties of Residual Based Tests for Cointegration", *Econometrica*, 59, p. 190.

表 7 检验 r 个同积关系的约翰逊 λ -max 检验的临界值

随机游动 个数 ($g=n-r$) (g)	样本量 (T)	λ -max 统计量入于表中数值的概率					
		0.500	0.200	0.100	0.050	0.025	0.001
情况一							
1	400	0.58	1.82	2.86	3.84	4.93	6.51
2	400	4.83	7.58	9.52	11.44	13.27	15.69
3	400	9.71	13.31	15.59	17.89	20.02	22.99
4	400	14.94	18.97	21.58	23.80	26.14	28.82
5	400	20.16	24.83	27.62	30.04	32.51	35.17
情况二							
1	400	2.415	4.905	6.691	8.083	9.658	11.576
2	400	7.474	10.666	12.783	14.595	16.403	18.782
3	400	12.707	16.521	18.959	21.279	23.362	26.154
4	400	17.875	22.341	24.917	27.341	29.599	32.616
5	400	23.132	27.953	30.818	33.262	35.700	38.858
情况三							
1	400	0.447	1.699	2.816	3.962	5.332	6.936
2	400	6.852	10.125	12.099	14.036	15.810	17.936
3	400	12.381	16.324	18.697	20.778	23.002	25.521
4	400	17.719	22.113	24.712	27.169	29.335	31.943
5	400	23.211	27.899	30.774	33.178	35.546	38.341

资料来源:同表 6。

143-115、154、156、158-161、
163、172、185-187、201、205、
207、209、234、278、321)
带常数项的~ (48、79、80、138)
低频率数据 (323)
迪基-福勒检验 (DF 检验) (59、
76、112、185)
增广的~ (ADF 检验) (94、96、
112、185)
递推公式 (281、283、305、306)
典型相关 (240、246、251、254、257)
~系数 (241、242、246)
样本~ (247、252)
典型形式 (252)
独立同分布 (iid) (3-9、13、15、
19、22、28、31、38、39、41、49、56、
58-60、69、81、84、85、87、90、
93-97、103、106、110-113、
116-119、121、125、129、130、
145、154、157、162、163、167、
169、172、175、181、188、199、
209、216、219、224、231、234、
247、269、276、277、280、281、
283、302、305、306、312、316、
318、323、329、330)
对角矩阵 (71、98、133、241、260、
261、316)
分块~ (149、216、219)
对数似然函数 (248、254、255、
259、267、268、274、288、290、

298、299、308)
集中的~ (247、251、255、261、
268、272、274、275)

条件的~

多斯科定理 (16)

E

EGARCH 模型 (指数 GARCH 模
型) (312、314、315、320)

二次型 (102、154、174、211、214、
215、217、218、221-223、229、
231)

二项式 (307)

F

F-检验法 (143)

F-统计量 (68、69、76、107、110、
113、145、152、155、157、158)

泛函中心极限定理 (9、13-16、
18、24、25、39、45、55、116)
多变量的~ (115、119)

非标准分布 (142)

非参数 (230)

非退化 (29、33、39、50、51)

非稳定过程 (4、5、7)

菲利普斯方法 (55)

菲利普斯-配荣检验 (PP 检验)
(58、60、80、112、185)

G

Γ -函数 (297、315、333)
GARCH 模型(广义 ARCH 模型)
 (302、308、311、312、318)
 半强~ (324、330)
 强~ (323、324、328、330)
 弱~ (324)
 向量~ (316、317)
GARCH (1,1) (306、317、322—
 326、328、331)
高频率数据 (322)
哥兰格表示定理 (166、171)
共线性 (70)
购买力平价 (203、204、269、270)
估计残差 (88、90、147、156、185、
 186、191、192、200—203、207—
 209、248、249、253、263、267、
 272、294、300)
广义误差分布 (315、332、333)
规范化 (33、179、180、241、252)
国民经济 (31)

H

混合正态分布 (214)

J

积变量 (323—326、328 330、332)
几乎处处 (55)
 ~连续 (11)
 ~不可微 (12)
迹数 (250、257、276)
 ~统计量 (261—265、275)
假设检验 (1、29、51、57、77、101、
 102、112、115、135、137、138、
 143、157、159、171、173、186、
 203、209、218、229、230、248、
 253、260、266、270—272、275、
 278、299、311、332)
阶梯函数 (14、19、20、24、25)
近单位根过程 (52、55)
矩阵多项式 (138、139)
矩稳定过程(见稳定过程)
绝对可和 (118、139、145、167、170、
 175、181、188、216、219)
均衡关系 (160、166、171、323)

K

科克伦 奥克特方法 (155、156)
可测函数 (321)

L

拉格朗日

~乘数 (257、299)

~函数 (257)

~检验法 (311)

λ -max 检验 (263、265、275)

连续映照定理 (16、17、19、21、25、
43—45、55)

泛函的~ (18)

林德贝格—勒维中心极限定理
(13、16)

临界值 (30、59—62、66—70、76、
79、80、86、88、90—94、102、106、
107、110—112、187、188、200—
203、205—209、262、265、269、
270)

流变量 (323、324、326—328、330—
332)

M

Microfit (332)

蒙的可罗 (30、31、144、200、285)

N

拟合优度 (178、247、301、302、312)

P

PcGive (332)

谱密度函数 (162)

Q

强GARCH(r,q)过程

(见 GARCH 过程)

峭度 (327)

去势 (47)

R

ρ -统计量 (90、93、104、106、110、
111、187、189、190、203)

弱收敛 (16、18、19、21、31、39)

弱外生 (271、276、277)

S

三角

~表示形式 (167、169、171、
174、175、223、233)

~分解 (145、188、190)

时间趋势 (47、48、51、57、70、79、
170、171、201、206、207、234)

确定的~ (47、48)

识别 (209、255、276、316)

收敛速度 (29、37、55、89、135—137、143、178)

数据生成过程 (1、51、54、76—80、94、110—112、131)

随机 (1、4、10、11、13、16、17、52、56、93—95、112、113、116、119、121、127、129—131、140、143—145、149、153—155、157、159、160、162、165、167、169—171、173—175、182、185、188、201、209、215、216、219、224、230、231、233、234、239—241、243、246—248、259、263、280—282、284、287、297、299、303、311、314—319、321—324、327、332)

~差分方程 (320)

~函数 (16—18、26、39)

~积分 (13、25、26、56、120、121、128、157、231)

~微分方程 (52、319—322、332)

随机游动 (5、7、8、10、19、22、28、31、32、37、60、154—156、160、170、171、278—280)

带常数的~ (31、32、35)

T

t -统计量 (90—92、94、104、106、110、111、144、187、190、201—

203、208、210)

特征

~方程 (8、95、130、137、161、257、259、274、281、302)

~向量 (241、242、244、245、258、259、266、274、275)

~值 (241—246、257—259、263、264、266—268、270、274、275、307)

条件

~方差 (281—284、287、298、299、312、314、316—318、322、324、326、327、330—332)

~期望 (210、280、281、283、291、307、316—318、329)

同积

~过程 (159、161、162、165—167、169—171、173—175、179、180、183、185、230、233、234、239、247、278)

~空间 (167、240)

~理论 (115、129、160)

~向量 (159、160、162、164、166、167、169、171、173—175、179、180、185、203、209、214、215、218、223、229—231、233、239、248、252、253、266、268、269、275、276)

~秩 (160、162)

同趋势表示形式 (169-171)
投影 (180、218、219、250、328)

V

VAR (向量自回归) (128-130、
136、137、156、161、163、165、
167、172)
带常数的~ (137、143)
不带常数的~ (130、131、143)

W

维纳过程 (9、10-13、15、19、21、
25、31、52、55、56、76、112、115、
117、146、149、176、189、193、
212-214、217、218、220、224—
227、229、231、262、319、322)
标准的~ (29、115)
多维的~ (115)
尾部厚度参数 (315)
伪回归 (115、143、144、154、155、
157、160、185、191、205)
稳定过程 (1、3、4、6、8、9、29、37—
40、52、56、58、69、79、80、85、90、
92-94、100、103、106、110、112、
116-118、145、155、159、162、
165、166、168、180、187、231、
280、281、303、306、332)
带趋势的~ (47、48、57)

无条件
~方差 (281、284、287)
~矩 (305)
~期望 (280、284、318、332)

误差修正(ECM) (165、171)
~过程 (165、166、171、234)
~形式 (165-167、239、240、
247、248、259、271、275、276)

X

显著水平 (59-62、79、80、91—
94、106、107、110、200、205、265)
线性 (140、144、160、162、166—
168、179、209、218、240、280、
283、333)
~变换 (147、148、158、168、191、
240)
~组合 (143、159、164、166、185、
325)
向量自回归(VAR)
(见 VAR)
信息矩阵 (292、293、299、309)

Y

一致
~估计 (3、4、29、31、84、88、89、
104、132、157、174、217、221、
222、227、228、259、293、296、

300)

~收敛 (54)

移动平均(MA) (94、131、161、
163、167、169、171、175、181、
190、219)

有效市场 (6、203、279)

约翰逊方法 (240、246、248、275)

Z

Z_p -检验法 (90)

Z_t -检验法 (90)

正定矩阵 (116、210、243)

正交补矩阵 (235)

正态分布 (11、15、27、70、135、136、
142、155、157、210、211、214、
215、218、220、278、285、287、
295—297、315、317、320、333)

指数 GARCH 模型
(见 EGARCH)

滞后多项式 (8、38、81、95、96、129、
161、163—165、175、216、234、
235、239、280、302、314)

自回归条件异方差(ARCH)
(见 ARCH)

自相关系数 (88、89、174)

样本~ (88—90、187)

综合 (15、24、37、78、82、120、198、
245、294、322—324、326、328
330、332)

总积累效应 (234)

最大似然

~方法 (3、233、270、315、317)

~估计 (171、240、248、251、252、
254、255、258、259、261、271、
272、274、275、287、288、290、
294、295、297、298、308)

最佳线性预测 (283、324)

最小二乘方法(OLS) (273、279)
充分改进的~ (223、228)

