XIANXING GUOCHENG DE RUOGAN JIXIAN LILUN JIQI YINGYONG

线性过程的若干极限理论及其应用

李云霞 著

ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

线性过程的若干极限理论及其应用

李云霞著

浙江大学出版社

图书在版编目（CIP）数据

线性过程的若干极限理论及其应用／李云霞著．—
杭州：浙江大学出版社，2010.11
ISBN 978-7-308-08121-4

Ⅰ．①线… Ⅱ．①李… Ⅲ．①线性系统理论—研究
Ⅳ．①O231

中国版本图书馆 CIP 数据核字（2010）第 219458 号

线性过程的若干极限理论及其应用

李云霞 著

责任编辑 徐 静
封面设计 刘依群
出版发行 浙江大学出版社
（杭州市天目山路 148 号 邮政编码 310007）
（网址：http://www.zjupress.com）
排 版 杭州中大图文设计有限公司
印 刷 杭州杭新印务有限公司
开 本 710mm×1000mm 1/16
印 张 14
字 数 290 千
版 印 次 2010 年 11 月第 1 版 2010 年 11 月第 1 次印刷
书 号 ISBN 978-7-308-08121-4
定 价 35.00 元

作者简介

李云霞，2005年毕业于浙江大学数学系，获概率论与数理统计专业博士学位。同年，进入浙江财经学院数学与统计学院任教，先后获财院“高质量论文”奖，“优秀教师”的称号以及“浙江财经学院科研成果奖”一等奖等。2006年，破格晋升为副教授，并被推选为“浙江财经学院优秀中青年骨干教师”和“中青年学科带头人”，入选浙江省“新世纪151人才工程”第三层次培养人员，入选2007年度浙江省高等学校优秀青年教师资助计划。近五年来，在国内、国际重要期刊上发表十几篇学术论文，多数被SCI收录，如《Journal of Mathematical Analysis and Applications》、《Metrika》、《Statistics and Probability Letters》、《Acta Mathematics Sinica-English Series》等。主持国家自然科学基金项目、教育厅科研项目等，并参与多项国家自然科学基金项目、国家社科基金项目、浙江省社科项目等建设。

序　　言

概率论是从数量上研究随机现象的规律性的学科. 它在自然科学、技术科学、管理科学中都有着广泛的应用, 因此从 20 世纪 30 年代以来, 发展甚为迅速, 而且不断有新的分支学科涌出. 概率极限理论就是其主要分支之一, 也是概率统计学科中极为重要的理论基础. 苏联著名概率论学者 Gnedenko 和 Kolmogrov 曾说过: “概率论的认识论的价值只有通过极限定理才能被揭示, 没有极限定理就不可能去理解概率论的基本概念的真正含义.” 经典极限理论是概率论发展上的重要成果, 而对时间序列中最具代表性的模型之一 —— 线性过程各类极限性质的研究是近代概率极限理论研究中的方向之一, 本书就是对线性过程的弱极限性质、强极限性质以及在变点中的应用进行了深入的研究.

线性过程在时间序列分析中具有非常重要的地位, 有大量文献都讨论了线性过程的各种性质, 它对于经济、工程及物理学科都有着极其广泛的应用. 因此很多学者致力于研究线性过程的误差项满足不同条件时线性过程的极限定理. 例如当误差项为鞅差随机变量序列 (Fakhre-Zakeri(1997)), 误差项为强混合随机变量序列 (Birkel (1993)) 以及误差项在线性坐标正相依 (LPQD) 条件限制下 (Tae-Sung (2001)), 已经得到了相应的关于线性过程的中心极限定理 (CLT) 和泛函中心极限定理 (FCLT). 在一些适当的条件下, 对于线性过程还有很多极限结果. 比如, Burton 和 Dehling (1990) 得到了线性过程的大偏差原理, Yang (1996) 建立了中心极限定理以及重对数律, Li *et al.* (1992) 和 Zhang (1996) 都得到了完全收敛性方面的结果.

本书主要是对由各种相依随机变量产生的线性过程的各类极限性质进行了讨论. 众所周知, 现实生活中发生的事情大多并不是互不相干的, 而是彼此之间

具有某种联系的．正确地用数学方法描述这种相关性，就可以用数学 —— 这一精确的工具来对事物进行精确地研究．由此可见，研究非独立的随机变量序列有着十分深刻的理论和实际意义．其实，关于相依随机变量的极限性质的研究可以追溯到 20 世纪二三十年代，当时就有 Bernstein (1927) 、 Hopf (1937) 和 Robbins (1948) 等学者相继对其进行研究．一直到现在，仍有新的相依变量类型及其结果层出不穷．

本书的第一章就线性过程弱收敛方面的结果进行了深入的讨论．其中第二节主要讨论了由渐近线性坐标负相依 (ALNQD) 随机变量序列产生的平稳线性过程，获得了一个泛函中心极限定理．第三节则是证明了只要满足其中一个关键的不等式，线性过程的误差项在很多种相依条件的假设下，都可使与第二节相同的泛函中心极限定理成立．并且第三节还叙述了一个简单应用，就是将此结果应用于计量经济中一种很常用过程 —— 单位根过程检验中统计量的极限分布．然而往往在许多实际问题中，误差项不是一个简单的实值随机变量，常常是一个过程．第四节讨论的就是一列由 $\rho-$ 混合的过程序列产生的线性过程，得到了部分和的弱收敛性、两参数随机过程的弱收敛性以及随机足标和的弱收敛．

在第二、三章中，我们讨论了关于线性过程的强极限性质．在第二章中，主要研究由两种较为常见的相依随机变量序列产生的线性过程的强极限性质．完全收敛性的概念是由 Hsu 和 Robbins (1947) 引入的， Erdös (1949, 1950) 和 Spitzer (1956) 也作了相应的研究．到了 20 世纪 60 年代， Katz (1963) 及 Baum 和 Katz (1965) 推广了他们的结果，得到如下结论：设 $X_1,\cdots,X_n,\cdots$ 为 i.i.d. 随机变量序列，记 $S_n=\sum\limits_{j=1}^{n}X_j$. 令 $p<2,\ r\geqslant p$. 那么

$$\sum_{n=1}^{\infty}n^{r/p-2}\mathsf{P}\{|S_n|>\varepsilon n^{1/p}\}<\infty,\quad \varepsilon>0$$

成立的充要条件为 $\mathsf{E}|X_1|^r<\infty$, 且当 $r\geqslant 1$ 时 $\mathsf{E}X=0$. 白志东、苏淳 (1985) 对这结果进行进一步的推广与改进．其后 Davis (1968) 证明了：对任意 $\varepsilon>0$,

$$\sum_{n=1}^{\infty}\frac{\log n}{n}\mathsf{P}\{|T_n|\geqslant\varepsilon\sqrt{n\log n}\}<\infty$$

成立的充要条件为 $\mathsf{E}X_1=0$ 且 $\mathsf{E}X^2<\infty$. 随后，又产生了各种各样的上述结果的推广形式．最近， Chen (1978) 及 Gut 和 Spătaru (2000a) 讨论了当 $\epsilon\searrow 0$ 时 i.i.d. 随机变量 Baum-Katz 及 Davis 大数律的精确渐近性 (定理 2.B). 国内，王

岳宝、张立新等在精确渐近性方面也取得了很多成果. 而对于线性过程, Zhang (1996) 得到了 Baum-Katz 大数律形式的完全收敛性的结果 (定理 2.A), 关于线性过程精确渐近性方面的结果非常之少. 因此我们在第二章的第二节中, 证明了定理 2.B 类型的完全收敛性的精确渐近结果对于在 φ- 混合或 NA 两种相依条件下的线性过程也同样成立. 另一方面, 重对数律是概率极限理论中极为深刻的结果, 它是强大数律的精确化, 对它的研究一直为众多的学者所关注和重视, 并已经得到了许多经典的结论. 最近, Gut 和 Spătaru (2000b) 还发表了一篇关于 i.i.d. 随机变量序列的重对数律的精确渐近性质的文章, 结果见定理 2.C. 而线性过程关于重对数律的精确渐近性方面的结果也非常之少. 因此第三节中, 我们将证明线性过程在 φ- 混合或 NA 两种相依条件下关于类似定理 2.C 的重对数律的精确渐近性结果也同样成立. 将 $\sqrt{n\log\log n}$ 替换成 $\sqrt{n\log n}$ 后, 本节同时证明了类似定理 2.D 的对数律的精确渐近结果也成立. 最后, 矩完全收敛性问题是由 Chow (1988) 提出的, 并讨论了 i.i.d. 随机变量序列的矩完全收敛性 (定理 2.E). 王定成和苏淳 (2002) 讨论了 B 值独立同分布随机变元序列的矩完全收敛性. 则第四节中我们讨论了由 NA 随机变量序列产生的线性过程关于矩的完全收敛性.

由于在 Chow 等人的启发下, 蒋烨的博士论文 (2004) 讨论了 i.i.d. 随机变量序列关于矩的完全收敛及重对数律方面的精确渐近性质. 那么在蒋烨的博士论文 (2004) 的启发下, 第三章中我们主要考虑矩形式的精确渐近结果对线性过程是否也同样成立. 第二节中我们得到了由 i.i.d. 随机变量序列产生的线性过程关于矩形式完全收敛的精确渐近结果. 第三节则得到了关于矩的重对数律的精确渐近结果.

在第四章中, 我们主要讨论了线性过程在变点方面的一点应用. 一般地说, 变点就是 "模型中的某个或某些量起突然变化之点". 这种突然变化往往反映事物的某种质的变化, 在自然界、社会及各种领域中很常见且具有重要性, 虽然从统计学发展的角度看, 变点的统计分析这个课题还不能说已发展得很充分成熟了 (它迄今只有四十余年的历史), 但变点问题在许多应用中都非常重要, 因此针对一些常见的问题发展了若干行之有效的方法, 对应用家来说不失为一个有用的工具. 一旦变点被合适定位后, 原始模型就需要修改, 以便得到更好的数据解释以及更精确的预测. 因此变点估计在经济建模中起了很重要的作用. 许多统计与经济文献中都包含了大量关于变点问题的著作. 近期比较全面的文献可

以参见 Csőrgö 和 Horváth (1997). 单变点均值变动估计的问题是其中一个热门的研究方向，引起了学术界的长期关注. Sen 和 Srivastava (1975a, b), Hawkins (1977), Worsley (1979, 1986), James *et al.* (1987) 以及 Srivastava 和 Worsley (1986) 都考虑的是对一列正态序列提出单变点均值变动的检验. Hinkley (1970), Bhattacharya (1987), Yao (1987) 和其他很多学者考虑的是一列独立变量的单变点估计. 对于序列相关的数据，Picard (1985) 对阶已知的高斯自回归过程进行了估计. 以上这些作者考虑的都是极大似然估计 (MLE). 本章主要是由最小二乘估计 (LS) 方法来讨论线性过程未知变点估计的极限性质. 线性过程的 LS 方法是由 Bai (1994) 提出来的. 这个方法不同于 MLE, 无须对模型中的随机误差的分布有特定的假设, 而且计算相对简便. Bai (1994) 用 LS 方法考虑了由 i.i.d. 随机变量序列产生的线性过程的单变点估计. 然而对相依随机变量序列变点的研究无疑是学术界更加感兴趣的问题. 第二节就是考虑在相依假设下线性过程单变点估计的极限性质. 在两方面改进了 Bai (1994) 的结果: (i) 将条件 $\sum_{j=0}^{\infty} j|a_j| < \infty$ 减弱到 $\sum_{j=0}^{\infty} |a_j| < \infty$, (ii) 在更多的相依假设下用 LS 估计得到了类似的极限性质. 许多早期的努力都致力于单一变点的估计. 相对而言，涉及多变点的文献则较少. 当变点数目未知时多变点问题更加复杂, 因此更少的文章致力于此问题. 许多作者只考虑独立随机变量序列的特殊情况. 特别的， Yao (1988) 由 Schwarz 准则估计了独立正态序列均值的变点数目. 但是近期相依观测方面的研究也引起了学术界的广泛关注. 例如 Bai (1994); Davis et al. (1995); Horváth (1993,1997); Picard (1985); Epps (1988) 以及 Bai 和 Perron (1998) 等. 第三节就是讨论了各种相依假设下线性过程多变点的相合性以及相合速度. 当变点数目已知时，主要运用了 Bai (1994) 提出的 LS 方法来估计变点. 而变点数目未知的情况下，则是通过惩罚最小二乘方法来估计的. 此方法根据惩罚性可以视为模型 选择问题 (参见 Schwarz (1978)). 利用弱或强不变原理对观测序列进行逼近检验是变点分析中一种很重要的工具. 当弱不变原理成立时， Horváth (2000) 讨论了变点估计的逼近 CUSUM 检验. 第四节的主要目的就是在弱不变原理成立的条件下，对长程记忆过程进行均值和方差的基于最小二乘残差的逼近 CUSUM 检验.

值得一提的是，本书所涉及的关于重对数律的精确渐近性质以及长程记忆过程的极限性质的课题是近几年极限理论中的热门课题. 且文中有些结果所需的条件已经达到与独立同分布序列相关已知结果对等的程度，如第二章中的关于由 φ- 混合、NA 序列产生的线性过程的结论大多达到了最一般的独立情形完

全对等的程度. 而所加条件在独立情形时是充分必要的, 因此在我们讨论的情形中, 这些条件也是不可减弱的, 关于线性过程的条件也是一般性的. 然而, 限于个人的学识能力, 文中仍有一些结果还没达到最佳的程度.

Preface

Theory of Probability is a science of quantitatively studying regularity of random phenomena, which is extensively applied in natural science, technological science, social science and managerial science etc. Hence, it has been developing rapidly since 1930's and many new branches have emerged from time to time. Limit Theory is one of the branches and also an important theoretical basis of science of Probability and Statistics. As stated in the classical book "Limit distributions for sums of independent random variables"(1949) by B.V.Gendenko and A.N.Kolmogrov, "The epistemological value of the theory of probability is revealed only by limit theorems. Without limit theorems it is impossible to understand the real content of the primary concept of all our sciences — the concept of probability." Classical limit theory is the signify achievement in the progress of Probability. The linear processes are the most representative model in time series. Studying various limiting properties of linear processes is one of orientations of the current study of Limit Theory. Some significant results of the linear processes about weak limit properties, strong limit properties and application in change-points problem have been reached through deep research in this dissertation.

The linear processes are of special important in time series analysis and they arise in a wide variety of contexts. Applications to economics, engineering and physical sciences are extremely broad and a vast amount of literature is devoted to the study of the limiting theorems for the linear processes under various condition on errors. For example, under the martingle difference assumption on error, under the strong mixing condition on error and under LPQD condition on error, the central limit theorem (CLT) and the functional central limit theorem (FCLT) of the linear processes are proved. Under some suitable conditions, other limiting results have been obtained for the linear processes. For example, Burton and Dehling (1990) have obtained a large deviation principle for the linear processes, Yang (1996) has established CLT and the law of the iterated logarithm (LIL), Li *et al.* (1992) and Zhang (1996) have obtained the results on the complete convergence *etc.*

Some kinds of limiting properties of the linear processes under various dependence assumptions are discussed in this paper. As is known to all, everything has correlations between one another in the world. If we can properly describe these correlations by mathematics, we can analyze subjects accurately by the precise tool– mathematics. Hence one can see that, the study on dependent random variables has momentous significance. In fact, the study on the limit properties of dependent random variables may be dated back to 1920's and 1930's. At that time, scholars such as Bernstein (1927), Hopf (1937) and Robbins (1948) had carried on studies on this topic. Till now, new kinds of dependent random variables and their corresponding conclusions have emerged in a endless stream.

The first chapter presents a insightful discussion over the results about weak convergence of the linear processes. In the second section of this chapter, the stationary linear processes generated by asymptotically linear negative quadrant dependent (ALNQD) are considered, and the FCLT is obtained. The third section considers a more general linear process with dependent errors. It is shown that if the dependent errors satisfy a key inequality, the FCLT is also true. As a simple application, the limit distribution of the statistics in testing the unit-root process is obtained. The unit-root process is a important process in theory of econometrics. However, the error involved in many practical problems is usually a process rather than a simple real random variable. The forth section copes with such a linear process in question, which generated from a sequence of ρ-mixing processes, and obtains the weak convergence about the partial sums of this process, two parameters stochastic process and the random sum.

The second and third chapters are about the strong limit properties of the linear processes, and the second one mainly considers the strong limit properties of the linear processes generated by two common dependent random variables. The results on this two chapters are related to the well-known complete convergence. Erdös (1949, 1950) and Spitzer (1956) have carried out the relevant research on the concept of the complete convergence, introduced first by Hsu and Robbins (1947), and later in the 1960's, Katz (1963) and Baum and Katz (1965) popularized the outcome and came to the following conclusion, Let $1 \leqslant p < 2$, $r \geqslant p$, then

$$\sum_{n=1}^{\infty} n^{r/p-2}\mathsf{P}\{|\sum_{k=1}^{n} X_k| > \varepsilon n^{1/p}\} < \infty, \quad \varepsilon > 0$$

holds, if and only if $\mathsf{E}|X_1|^r < \infty$, and, when $r \geqslant 1$, $\mathsf{E}X = 0$. Davis (1968) proved: for any $\varepsilon > 0$,

$$\sum_{n=1}^{\infty} \frac{\log n}{n}\mathsf{P}\{|\sum_{k=1}^{n} X_k| \geqslant \varepsilon\sqrt{n\log n}\} < \infty.$$

holds, if and only if $\mathsf{E}X_1 = 0$ and $\mathsf{E}X^2 < \infty$. Bai *et al.*(1985) established a more general result on this topic. Subsequently followed by a variety of ways to popularize. Furthermore, Chen (1978) and Gut and Spătaru (2000a) have studied the precise asymptotics of i.i.d. random variables in the Baum-Katz and Davis laws of large numbers as $\epsilon \searrow 0$ (Theorem 2.B). Some Chinese scholars have also studied the precise asymptotics. As to the linear processes, Zhang (1996) have obtained the result of the complete convergence in the form of Baum-Katz and Davis laws

(Theorem 2.A). However, the precise asymptotics results about the linear processes are very few. In section 2, chapter 2, the purpose is to show that the kind of precise asymptotics result such as Theorem 2.B also holds for the linear process under φ-mixing and NA dependence assumptions. On the other hand, it is well-known that LIL is a very propounding result in Probability Limit Theory. It is the precise phenomenon of the strong law of large numbers. Many scholars have always paid close attention to the study of LIL and a lot of classical conclusions have been drawn. Recently, Gut and Spătaru (2000b) published an article regarding the precise asymptotics in LIL of i.i.d. random variables (see Theorem 2.C). But few results for the precise asmyptotics in LIL about the linear processes are known. In section 3, we shall prove that the kind of precise asymptotics result on LIL such as Theorem 2.C also holds for the linear process both under φ-mixing and NA dependence assumptions. By replacing $\sqrt{n \log \log n}$ by $\sqrt{n \log n}$, this section also gives the result for the precise asymptotics on the law of the logarithm such as Theorem 2.D.

The moment type convergence of the complete convergence was introduced first by Chow (1988), and this convergence of i.i.d. random variables was discussed (Theorem 2.E). Then Wang and Su (2002) established the moment convergence of i.i.d. random elements on Banach space. In section 4, the moment convergence of the linear process under NA dependence assumption will be studied. Enlightened by Chow's result, Jiang (2004) discussed the precise asymptotics properties on moment convergence and LIL about i.i.d. random variables. Then, with the help of Jiang's research, we mainly consider whether the precise asymptotics on moment about the linear processes also hold in Chapter 3. In section 2, we obtain the result of the precise asymptotics on moment convergence for the linear processes of i.i.d. random variables. In section 3, the precise asymptotics on moment LIL is also studied.

The fourth chapter is on the applications of the linear processes in change-point problems. Generally speaking, change-point is " the point which some quantities suddenly change in the models ", such a sudden change is usually a qualitative change of things, which is common but important in nature, society and various fields. Although, from the statistical viewpoint, the project of statistical analysis of change-points can not be considered the most developed, it is vital in many applications. As a result, several effective ways to common problems are emerged, which is useful for the applications. Once a change point is properly located, the original model should be modified accordingly to provide better interpretation of data and more accurate forecasts. Therefore change point estimation plays a extremely active role in econometric modeling, and there are many articles about change-point problems in statistical and economic literature. As for all-around literature, we can turn to Csőrgö and Horváth (1997). Among various issues, estimation of the single mean shift is no doubt a popular research topic, which arouses long-range attention in academic field. Sen and Srivastava (1975a, b), Hawkins (1977), Worsley (1979, 1986), James *et al.* (1987) and Srivastava and Worsley (1986) proposed tests for testing a shift in a sequence of normal means. Hinkley (1970), Bhattacharya (1987), Yao (1987) and many others considered the estimation of the shift point in a sequence of independent variables. For serially correlated data, Picard (1985) estimated a shift in Gaussian autoregressive process with a known order. These authors considered maximum likelihood estimation (MLE). This chapter discusses the least-square (LS) estimator of the unknown change-point in the linear process, which has been proposed by Bai (1994). Unlike the MLE, the LS method does not need to specify the underlying error distribution function and is computationally simple. The LS procedure also allows a broader specification of correlation structure in the data than MLE can typically permit. Bai (1994) has considered a linear process of i.i.d. variables by the LS method. However, it is undoubtedly more interesting to study change-points about dependent random variables. In section 2, we consider the limit properties of the change-point estimation for the linear processes under dependence assumptions, and at the same time, Bai's (1994) outcomes are improved from two aspects: (i) to weaken the condition $\sum_{j=0}^{\infty} j|a_j| < \infty$ to $\sum_{j=0}^{\infty} |a_j| < \infty$, (ii) similar limit properties are obtained under more dependence assumptions. Most early efforts have been devoted to the detection of a unique change-point. In comparison, less studies have been carried out on the issue of multiple structural changes in multiple changes. The problem is much more intricate when the number of changes is unknown, and only a few papers are published on this problem. Many people only consider the particular case of changes in a sequence of independent random variables. In particular, Yao (1988) estimated the number of jumps in an independent normal sequence via 'Schwarz' criterion. Some others also considered the problem of dependent data, for example, Bai (1994); Davis *et al.* (1995); Horváth (1993, 1997); Picard (1985); Epps (1988) and Bai and Perron (1998) and so on. In section 3, we discuss the consistency and the rate convergence of the multiple change-points estimation of the linear processes under various dependence assumptions. When the number of change-points is known, the configuration of change-points is estimated by LS method, which has been proposed by Bai (1994). When the number of changes is unknown, it is estimated by using penalized least-squares approach. This method of change-points detection can be seen as a problem of model selection via penalization (see Schwarz (1978)). One of the key tools in change-point analysis is to make use of a weak (or strong) invariance principle for the observed sequence and to develop an asymptotic test. Horváth (2000) derived asymptotic CUSUM tests for detecting changes of weak dependence processes for which a weak invariance principle is available. The main aim of the section 4 is to derive the asymptotic CUSUM tests (based on least squares residuals) for detecting changes in the mean or variance of a strong dependence process such as a linear process with long memory.

It should be pointed out that some subjects mentioned in this article, such as the precise asymptotics in LIL and the limit properties of the long memory processes, are hot topics in field of limit theorems. And we try our best to make each of our results as perfect as possible. For instance, in Chapter 2, the results on linear process are established under minimal conditions. These conditions are sufficient and necessary for partial sums of i.i.d. random variables. However, due to the limitation of academic ability, some results in the paper may not come to the optimality.

文中部分缩写及符号说明

$r.v.$	随机变量
$a.s.$	几乎必然
i.i.d.	互相独立且同分布
$\mathsf{E}X$	随机变量X的数学期望
$\mathsf{Var}X$	随机变量X的方差
$\mathsf{Cov}(X, Y)$	随机变量X与Y的协方差
$X_n \to X$ a.s.	随机变量序列$\{X_n\}$几乎必然收敛于随机变量X
$X_n \xrightarrow{P} X$	随机变量序列$\{X_n\}$依概率收敛于随机变量X
$X_n \xrightarrow{\mathcal{D}} X$	随机变量序列$\{X_n\}$依分布收敛于随机变量X
$\mu_n \Rightarrow \mu$	测度序列$\{\mu_n\}$弱收敛于测度μ
$U \stackrel{\mathcal{D}}{=} V$	U与V等价，即U与V有相同的有限维分布
$I\{A\}$	集合A的示性函数
$\sharp A$	集合A中元素的个数