SHANGHAI
JINRONG XUEYUAN GUOJI JINRONG YANJIUYUAN
XUESHU ZHUZUO CONGSHU

上海金融学院国际金融研究院学术著作丛书

线性模型中参数估计的可容许性理论

鹿长余 著

中国财政经济出版社

上海金融学院国际金融研究院学术著作丛书

线性模型中参数估计的可容许性理论

鹿长余 著

中国财政经济出版社

图书在版编目（CIP）数据

线性模型中参数估计的可容许性理论/鹿长余著．—北京：中国财政经济出版社，2015.3

（上海金融学院国际金融研究院学术著作丛书）

ISBN 978-7-5095-5990-1

Ⅰ.①线… Ⅱ.①鹿… Ⅲ.①线性模型-参数估计-可容许性-研究 Ⅳ.①O212

中国版本图书馆 CIP 数据核字（2015）第 014633 号

责任编辑：吕小军　　　　责任校对：王　英
封面设计：思梵星尚　　　　版式设计：董生平

中国财政经济出版社 出版

URL：http：//www.cfeph.cn

E-mail：cfeph@cfeph.cn

社址：北京市海淀区阜成路甲 28 号　邮政编码：100142

营销中心电话：88190406　北京财经书店电话：64033436　84041336

北京财经印刷厂印刷　各地新华书店经销

787×1092 毫米　16 开　9.5 印张　145 000 字

2015 年 3 月第 1 版　2015 年 3 月北京第 1 次印刷

定价：22.00 元

ISBN 978-7-5095-5990-1/F·4823

（图书出现印装问题，本社负责调换）

本社质量投诉电话：010-88190744

打击盗版举报热线：010-88190492，QQ：634579818

谨以此书献给：

我的父母、我的妻儿、我的姐姐们。

是他们给了我无微不至的爱，才有了本书的问世。

鹿长余

2014 年 10 月 2 日

前 言 | Preface

本书系统地介绍了线性模型参数估计的可容许性理论，集资料性和研究性于一体。在线性模型方差阵奇异的情形下，介绍了中国统计工作者关于回归系数函数估计的可容许性的工作，其方法有着简洁、漂亮的空间结构；提出并系统地研究了不完全椭球约束下线性模型中参数估计的可容许性理论，把历史上关于线性模型可容许估计分为参数无约束情况和参数受椭球约束的情况统一了起来；特别研究了方差分量模型的非负二次估计的可容许性问题，指出了非负二次估计和局部最优估计之间的关系；把回归系数的随机控制和泛控制的难题也向前推进了一大步。本书的最大特色在于证明方法的创新，对于进一步研究这方面的问题，具有较大的启发意义。

鉴于统计判决理论和线性模型理论等有关方面的书籍已出版很多，因此有关统计判决理论、线性模型理论的基本概念，本书没有涉及。建议初学者在知晓这些基本的理论后再阅读本书，推荐参阅陈希孺院士著的《高等数理统计》（中国科技大学出版社 2009 年 8 月版）、陈希孺院士等著的《线性模型参数的估计理论》（科学出版社 2010 年 9 月再版）。

为了给读者提供有关线性模型参数估计理论的全面研究方法，有必要把关于参数估计可容许性的关键文献内容提供给大家，为了不打破体系，这部分内容以扩展资料列在其后。这些文献都是可容许性研究中某个方面的重要文献，其结论与证明方法在可容

许性研究的文献中具有重要地位：扩展资料一解决了线性模型中的协方差阵可以是奇异阵时，回归系数的线性估计可容许性的充分必要条件，这是一篇研究线性模型中线性估计可容许性当方差阵非负定时的奠基性文献，证明方法独特。线性模型中参数估计的可容许性最初由世界著名的统计学家 C. R. RAO 在纪念统计判决理论的创始人 A. Wald 的文章中给出（Ann. Statist.，4：6（1976）. 1023－1037）。然而，遗憾的是，当方差阵非负定时，Rao 的结果是错的，陈希孺院士在他的《近代回归分析》（内部讲义）中指出了这个问题。扩展资料二给出了非齐次线性估计和齐次线性估计的可容许关系，这个关系适合几乎所有的线性模型的线性估计的可容许性。后来（目前还在继续）很多论文都是这篇论文结论的推理。

本书可作为高等院校概率、统计学专业的研究生教材和教学参考书，也可作为研究统计学的青年科研工作者的科学研究资料。

2014 年 10 月

目 录 | Contents

第 一 章

绪 论

本书涉及三方面内容：随机序意义下或泛类损失函数下的估计量的可容许性、线性模型中回归系数的线性估计的可容许性及方差参数的非负二次估计的可容许性。

随机控制与泛控制

在 Wald 的统计判决理论中，比较判决函数优良性的控制准则是关于特定的损失函数定义的。但是，实际上损失函数的精确形式难以确定，因此，从稳键性角度考虑，研究在一类损失函数下的控制问题是非常重要的。众所周知的 Rao - Blackwell 定理可以说是在一类损失下考虑估计问题的经典结果，这个定理指出对于任何一个不是基于充分统计量的估计 δ，可以找到一个基于充分统计量的估计 δ^*，δ^* 在所有凸损失函数组成的类中一致优于 δ。近十年来，许多知名的统计学家在如下的损失函数类 L 下研究了通常估计量的控制问题：

$$L = \{L(|\delta - \theta|_Q) : L(\cdot)\ \text{是任意的不减函数}\}$$

在 L 下的控制问题称为泛控制（Hwang，1985）。为了说明引入这个损失函数类的动机，我们引入另一个比较估计量的准则，此准则基于集中概率（concentration probability）：$P(\delta - \theta \in C)$，C 通常为包含原点的凸集。显然，集中概率越大，估计量越好。通常所说的随机序准则是这个准则的特例，在那里，C 通常取中心在原点的椭球。人们感兴趣的是找出集中概率最大的估计量，或在随机序下研究估计量的改进。许多作者证明了对于线性模型 $X = A\theta + \varepsilon$，其中 ε 服从椭球等高分布，则在所有线性无偏估计组成的类中，θ 的最小二乘估计具有集中概率最大的性质，且这个结论等价于 θ 的最小二乘估计在 L 下泛控制任一线性无偏估计（见Eaton，1988；Berk and Hwang，1989；Sinha and Drygas，1985；Hwang，1985；Ali and ponnalli，1990）。

Hwang（1985）证明了在 L 类下的泛控制等价于随机控制，而随机控制的思想近年来引起了统计学家的广泛关注，见 Hwang and Peddada（1994）、Rueda and Salvador（1995）、Rueda、Salvador and Fernandez（1997）、Kochar（1996a，b）、Dardanoni and Forcina（1998）、Dykstra（1997）、Shaked（1997）、Perlman（1996）、Belzunce，et.（1997）、Kelly（1989）、Hwang（1985，1986）、Brown and Hwang（1989）、Cohen and Sackrowitz（1970）、Rukhin（1978，1984）。所谓随机控制是指：若 $\forall c > 0$，

$$P(|\delta^*(X) - \theta|_Q \leqslant c) \geqslant P(|\delta(X) - \theta|_Q \leqslant c)$$

且上式不恒等，则称 $\delta^*(X)$ 随机控制 $\delta(X)$，因此研究在损失函数类 L 下估计量的控制问题有着十分强烈的概率背景。同时也为研究随机控制提供了另一条途径。

在前人研究的基础上，本书第一部分，我们将致力于研究在随机序意义下或泛控制意义下优于最小二乘估计的估计量存在性问题。将这方面的研究向前推进了一大步，这方面的部分结果

(Lu and Shi, 2000) 被该领域权威评论为"……this paper make a significant step about universal domination（该文在泛控制方面作出了重要的一步）"。

线性估计的可容许性的统一理论

我们首先考虑下面的线性模型：

$Y = X\beta + \varepsilon,\ E(\varepsilon) = 0, D(\varepsilon) = V$

其中 Y 是 $n\times 1$ 的随机变量；X 是已知的 $n\times p$ 阶矩阵；ε 是 $n\times 1$ 的误差变量，其中 $p\times 1$ 的向量 β 和 $n\times n$ 的非负定对称阵 V 是 T 中的未知参数变量；T 是笛卡尔乘积 $R^P\times V$ 的一个子集，其中 V 是所有非负定对称阵组成的集合。我们将这样的模型简记为 $(Y, X\beta, V \mid (\beta, V) \in T)$。

对于线性模型，一个重要的问题是研究回归参数 β 的线性估计的可容许性，这一问题由 Cohen（1966）提出并得到很快发展。其研究主要沿着两个方向进行，假设 T 具有结构：$T = R^P \times \sigma^2 V$，一是参数不受约束，即 $(\beta, \sigma^2) \in R^P \times (0, +\infty)$，也就是说当参数 (β, σ^2) 没有限制时。见 Shinozaki（1975）、Rao（1976）、LaMotte（1982）、Zhu and Lu（1986、1987）、Stepniak（1984、1989）、Mathew, Rao and Sinha（1984）、Wu（1986）、Klonecki and Zontek（1988）、Baksalary and Markiewicz（1988）、Techene（1998）。二是参数受椭球约束，即 $\beta' N\beta \leqslant \sigma^2$ 时，其中 N 是已知正定阵，见 Hoffmann（1977）、Mathew（1985）、Zhu&Zhang（1989）。本书的第二部分试图将这两个方向的结果统一起来，考虑椭球约束的推广形式：

$(\beta - \beta_0)' N(\beta - \beta_0) \leqslant \sigma^2, N$ 为非负定阵，β_0 已知。

由于 N 非负定，故这种约束称为不完全椭球约束，因为可能只是回归参数的一部分分量受约束，另一部分不受约束。注意到不完全椭球约束包含了参数不受约束（$rkN=0$）和参数受椭球约束（$rkN=p$）的情形，因此我们的研究将历史上沿两个方向的研究统一了起来，且更广泛得多。从而，我们建立了线性估计可容许的统一理论。同时，在这一部分中，我们也研究了在不等式约束下线性估计的可容许性问题，这个问题以前还从未有人研究过。

方差分量的非负估计的可容许性

关于方差分量的二次估计的研究可以说始于许宝禄先生的关于线性模型误差估计的著名论文（Hsu，1938）。由于这个模型在生物育种、数量遗传、心理学研究、工业质量控制和计量经济等方面有着广泛的应用。因此有大量的文献涉及方差分量的估计问题。1988 年，Rao and Kleffe 出版了一本有关方差分量估计的非常出色的专著。这里我们主要研究方差分量的非负二次估计的可容许性问题。与此相关的是不变二次估计的可容许性问题，由于附加了不变性的要求，使方差分量的不变二次估计在不变二次估计类中可容许性等价于某个线性模型中线性估计的可容许性，从而使问题得以解决。文献中有关这方面结果已相当丰富。见 Klonecki and Zontek（1992、1989、1987），Olsen、Seely and Birkes（1976），LaMotte（1982）。

但是，不变性的要求对于方差分量的二次估计不太适当，因为不变二次估计不能保证所给的估计量具有非负性，这不能不认为是一个不容忽视的缺陷，因为我们欲估计的方差分量参数是非负的。Verdooren（198S）指出，在方差分量估计中非负性应该是

估计量必须具备的条件，因此，近年来，各种估计是否具有非负性以及怎样用非负估计去改进常用的估计成为方差分量估计研究的又一热点，如，Pukelskeim（1981）研究了方差分量线性函数的非负二次无偏估计及非负 MINQUE。Chaubey（1984，1991）研究了最接近 MINQUE 的非负二次估计。Gnot、Kleffe and Zmyslony（1985）研究了可容许的不变二次估计的非负性。Mathew and Sinha（1992）、Sinha（1999）、Kubokawa、Saleh and Makita（1993）研究了用非负估计改进通常的估计量问题，吴启光、成平、李国英（1981）研究了误差方差的非负二次型估计的可容许性。

在本书中，我们去掉了不合适的不变性，研究了方差分量的非负二次估计在非负二次估计类中可容许性，“这是一个非常困难的问题，目前尚无完整的结果”（叶兹南，1988；王松桂，1987）。叶兹南（1988）在一定的条件下给出了一个充分条件和一个必要条件，但必要条件是在非常苛该的限制下给出的，由于这些条件的限制，使得该文关于可容许估计的该划很不理想。我们去掉了所有的限制条件，对于最一般的方差分量模型，给出了与叶兹南（1988）一样的必要条件，这从本质上改进了叶兹南（1988）的结果。同时也研究了非负二次可容许估计与局部最优估计之间的关系，给出了当方差阵非负定时，如何研究可容许性的一个非常有效的方法，其中第六章是作者早期的工作成果，考虑到体系的完整性，将其放入本书中。

第 二 章
随机控制与泛控制

第一节 引 言

在 Wald 的统计判决理论中，控制准则（或估计量的改进）是关于特定的损失函数定义的：假设总体的观测量 X 具有分布函数 $F(x,\theta)$，$\theta\in\Theta$ 为未知参数，$L(\theta,a)$ 表示用判决 a 去估计 θ 所引起的损失，$R(\theta,a)=E_\theta L(\theta,a)$ 表示风险函数，称 $\delta_1(x)$ 至少和 $\delta_2(x)$ 一样好，若

$$R(\theta,\delta_1)\leqslant R(\theta,\delta_2),\forall\theta\in\Theta \tag{2-1}$$

如果 δ_1 至少和 δ_2 一样好，且对某个 $\theta_0\in\Theta$，(2-1) 式成立不等号，则称 δ_1 控制 δ_2，也称 δ_1 一致优于 δ_2。

但是，实际上，损失函数的精确形式难于确定，因此，研究在一类损失函数下通常估计量的控制问题是非常重要的。众所周知的 Rao - Blackwell 定理可以说是在一类损失下考虑估计问题的经典结果，这个定理指出对于任何一个不是基于充分统计量的估计 δ，可以找到一个基于充分统计量的估计 δ^*，δ^* 在所有凸损失函数组成的类中一致优于 δ。近十年来，许多知名的统计学家在如下

的损失函数类 L 下，研究了通常估计量的控制问题：

$$L = \{L(|\delta - \theta|_Q) : L(\cdot) \text{ 是任意的不减函数}\} \tag{2-2}$$

其中 $(|\delta - \theta|_Q) = [(\delta - \theta)^t Q(\delta - \theta)]^{\frac{1}{2}}$ 称为广义欧氏误差，Q 为给定的非负定矩阵。

显然，损失函数类 L 包含的损失函数非常宽泛，包含了所有看似合理的损失函数，在 L 下的控制问题称为泛控制（见 Hwang，1985）。

为了说明引入这个损失函数类的动机，我们引入另一个比较估计量优良性的准则，此准则基于集中概率（concentration probability）：$P(\delta - \theta \in C)$，C 通常为包含原点的凸集。显然，集中概率越大，估计量越好。通常所说的随机序准则是这个准则的特例，在那里，C 通常取中心在原点的椭球。人们感兴趣的是找出集中概率最大的估计量，或在随机序下，研究估计量的改进。许多作者证明了对于线性模型 $X = A\theta + \varepsilon$，其中 ε 服从椭球等高分布，则在所有线性无偏估计组成的类中，θ 的最小二乘估计具有集中概率最大的性质，且这个结论等价于 θ 的最小二乘估计在 L 下泛控制任一线性无偏估计。见 Eaton（1988）、Berk and Hwang（1989）、Sinha and Drygas（1985）、Hwang（1985）、Ali and ponnalli（1990）、张宝学与鹿长余（1994）等。

1985 年，Hwang 证明了在 L 类下的泛控制等价于随机控制，而随机控制的思想近年来引起了统计学家的广泛关注，见 Hwang and Peddada（1994），Ruedaand SMvador（1995），Kuada、Salvador and Fernandez（1997）、Dardanoni and Forcina（1998），Dykstra（1997），Perlman（1996），Belzunce，et.（1997），Kelly（1989），Hwang（1985、1986），Brown and Hwaag（1989），Cohen and Sackrowitz（1970），Rukhin（1978、1984）。

所谓随机控制是指：若 $\forall c > 0, P(|\delta^*(X) - \theta|_Q \leqslant c) \geqslant P(|\delta(X) - \theta|_Q \leqslant c)$，且该式不恒等，则称 $\delta^*(X)$ 随机控制

$\delta(X)$。因此，研究在损失函数类 L 下估计量的控制问题有着十分强烈的概率背景。同时也为研究随机控制提供了另一条途径。

有关泛控制的比较完整的结果由 Hwang（1985）给出，他指出当 $X-\theta$ 具有自由度为 N 的 t 分布时，存在无穷多个 James - stein 正部估计 $\delta_\alpha = (1-\alpha/|X|^2)^+ X$ 泛控制 θ 的最小二乘估计；但对于重要的正态情形 $X-\theta \sim N(0,I)$，则任一 James - stein 正部估计均不能泛控制 θ 的最小二乘估计，那么是否存在泛控制 θ 的最小二乘估计的估计量呢？1989 年，Brown and Hwang 部分地解决了这个问题，他们指出泛控制严重地依赖泛类损失函数 $L(|\delta-\theta|_Q)$ 中的非负定阵 Q；当 $Q=I$ 时，不存在泛控制 θ 的最小二乘估计的估计量；当 $Q \neq I$ 时，只要 X 的维数 p 充分大，就存在泛控制最小二乘估计的估计量。特别当 $Q = diag(q_1, q_2, \cdots, q_p)$，$q_1 > q_2 = q_3 =, \cdots, = q_p$ 时，存在泛控制最小二乘估计的估计量的充要条件是 $p \geqslant 4$，Lu and Shi（2000）继续深入研究了这方面问题，对于给定的 Q，他们致力于研究泛控制最小二乘估计的估计量的存在性问题。给出了存在泛控制最小二乘估计的一个较好的充分条件，特别当 $q_1 > q_2 = q_3 \geqslant, \cdots, \geqslant q_p$ 时给出了一个充要条件，将 Brown and Hwang（1989）的结果向前推进了一大步。其中给出的解决问题的方法具有较大创新性，下面介绍有关这方面的结果。

第二节　概念和背景知识

假设 X 的概率分布函数为 $F(x,\theta)$，$\theta \in \Theta$ 为末知参数，假定 $\delta(X)$ 是 θ 的估计量，损失函数类取为（2 - 2）式中给出的 L，其中 Q 为给定的非负定矩阵，称 $|\delta(X)-\theta|_Q = [(\delta(X)-\theta)'Q(\delta(X)-\theta)]^{\frac{1}{2}}$ 为广义欧氏误差；当 $Q=I$ 时，简写为 $|\delta$

(X) $-\theta|$。在本章中，对于给定的矩阵或者向量 X, X'、X^t 均表示 X 的转置。

定义 2.1：(1) 若对 L 中的每一个损失函数 $L(\cdot)$, $\delta_1(X)$ 都至少和 $\delta_2(X)$ 一样好，且对某一损失而言，$\delta_1(X)$ 控制 $\delta_2(X)$，则称 $\delta_1(X)$ 关于 Q 泛控制 $\delta_2(X)$。

(2) 若不存在泛控制 $\delta(X)$ 的估计量，则称 $\delta(X)$ 是泛可容许的。

定义 2.2：若 $p_\theta(|\delta_1(X)-\theta|_Q<c)\geqslant p_\theta(|\delta_2(X)-\theta|_Q<c)$，$\forall c>0, \forall\theta\in\Theta$，且不等号对某个 $\theta_0\in\Theta, c_0>0$ 成立，则 $\delta_1(X)$ 关于 Q 随机控制 $\delta_2(X)$。

为了给出较全面的背景知识，我们引述有关泛控制的一些已有结果。

定理 2－1：等价性定理（Hwang，1985）。在广义欧氏误差 $|\delta(X)-\theta|_Q$ 下，δ_1 泛控制 δ_2，充分必要 δ_1 随机控制 δ_2。

定理 2－2：假设 $X=A\theta+\varepsilon$。ε 服从均值为 0，协方差为 $\sigma^2\Sigma$ 的椭球等高分布，则 θ 的最小二乘估计 $(A^t\sum^{-1}A)^{-1}A^tX$ 关于任意给定的非负定阵 Q 泛控制 θ 的任一线性无偏估计（Hwang，1985）。

定理 2－3：假设 p 维随机变量 $X\sim N(\theta,I_p)$，I_p 表示 p－阶单位阵。Q 为给定的非负定阵，$q_1,q_2,\cdots,q_p$ 为其特征根（Brown and Hwang，1989）。则：

(1) 当 $P\leqslant 3$ 时，θ 的最小二乘估计 X 关于 Q 是泛—可容许的。

(2) 当 $Q=I$ 时，对任意的 p, X 是泛—可容许的。

(3) 设 $q_1>q_2\geqslant,\cdots,\geqslant q_p>0$，若 $p>2\max\limits_c\left\{\dfrac{\max\limits_i(D_c(i,i))}{\min\limits_i(D_c(i,i))}\right\}+1$，则 X 是泛—可容许的。

(4) 设 $q_1>q_2=,\cdots,=q_p>0$，则 X 是泛—不可容许的充分必要条件是 $p\geqslant 4$。这里记号 $D_c(i,i)$ 的意义见下一节的（2－9）