学校编码：10384　　　　分类号 ______ 密级 ______

学号：200423037　　　　UDC ______________

厦　门　大　学

硕　士　学　位　论　文

广义 G-M 模型参数估计的一种新的相对效率

A New Relative Efficiency of the Parameter Estimate in Generalized Gauss-Markov Models

王　娜　娜

指导教师姓名： 林建华　教授

专 业 名 称： 概 率 统 计

论文提交日期： 2007 年 5 月

论文答辩日期： 2007 年　月

学位授予日期： 2007 年　月

答辩委员会主席： ______________

评　阅　人： ______________

2007 **年** 5 **月**

A New Relative Efficiency of the Parameter Estimate in Generalized Gauss-Markov Models

By

Nana Wang

Supervisor: Professor Jianhua Lin

Speciality: Probability Statistics

Institution: College of Mathematics Science

Xiamen University

Xiamen, P. R. China

May , 2007

学　位　论　文

广义 G-M 模型参数估计的一种新的相对效率

王　娜　娜

厦　门　大　学

二　0　0　七　年　五　月

厦门大学学位论文原创性声明

兹呈交的学位论文，是本人在导师指导下独立完成的研究成果。本人在论文写作中参考的其它个人或集体的研究成果，均在文中以明确方式标明。本人依法享有和承担由此论文而产生的权利和责任。

声明人（签名）：

年　　月　　日

厦门大学学位论文著作权使用声明

本人完全了解厦门大学有关保留、使用学位论文的规定。厦门大学有权保留并向国家主管部门或其指定机构送交论文的纸质版和电子版，有权将学位论文用于非赢利目的的少量复制并允许论文进入学校图书馆被查阅，有权将学位论文的内容编入有关数据库进行检索，有权将学位论文的标题和摘要汇编出版。保密的学位论文在解密后适用本规定。

本学位论文属于

1、保密（　　），在　　　　年解密后适用本授权书。

2、不保密（　　）。

（请在以上相应括号内打"√"）

作者签名：　　　　　　　　　日期：　　　年　　月　　日

导师签名：　　　　　　　　　日期：　　　年　　月　　日

目　录

Contents

摘　要

线性模型是非常重要的一种统计模型, 它包括一系列统计模型, 例如线性回归模型、方差分析模型、协方差分析模型等等。许多医学、经济、管理、气象、农业、工业等领域的现象都可以用线性模型来近似描述。因此线性模型成为现代统计学中应用最为广泛的模型之一。本文着重研究线性回归模型。

对于广义 G-M 模型, 如果最小二乘估计 (LS 估计) 与最佳线性无偏估计 (BLU 估计) 相等, 就可以用 LS 估计代替 BLU 估计; 反之, 用 LS 估计代替 BLU 估计就要蒙受一些损失。有时, 这种损失可能是很大的, 因而研究这种损失的大小就显得颇为重要。本文提出了一种新的相对效率, 并给出了该相对效率的上下界, 讨论了该相对效率与广义相关系数的关系。

关键词: 线性模型, 广义 G-M 模型, 相对效率, 广义相关系数, LS 估计, BLU 估计.

Abstract

Linear Models are very important statistic models. Linear models include many statistic models, for example linear regression models, variance analysis models, covariance analysis models and so on. A lot of phenomenon of research in medical science, economy, manegement, meteorology, agriculture, industry can be described approximately by linear models. So linear models become one of the most popular models in modern statistics. This paper mainly researches linear regression models.

For generalized Gauss-Markov models, the best linear unbiased estimate can be substituted by the least square estimate if they are equal to each other ; on the contrary, this method will bring some loss. Sometimes the loss is very great , so it is very important to study the loss. This paper proposes a relative efficiency . After giving its upper bound and lower bound, its relation with generalized coefficient is discussed at last.

Key words: linear regression models, generalized G-M models, relative efficiency, generalized coefficient, LSE, BLUE.

前　言

在线性模型参数估计理论与方法中,最小二乘估计方法占有中心基础的地位 [22] 。在通常情况下,最佳线性无偏估计很难计算出来,当用最小二乘估计来代替最佳线性无偏估计时，估计的精度就要蒙受一些损失 [21] 。为了度量这种损失，一些作者研究了最小二乘估计相对于最佳线性无偏估计的相对效率 [17], [19], [20] ，常用的几种方式有

$$e_1(\hat{\beta}) = \frac{|\mathrm{cov}(\beta^*)|}{|\mathrm{cov}(\hat{\beta})|},$$

$$e_2(\hat{\beta}) = \frac{\mathrm{trcov}(\beta^*)}{\mathrm{trcov}(\hat{\beta})},$$

$$e_3(\hat{\beta}) = \frac{\|\ \mathrm{cov}(\beta^*)\ \|}{\|\ \mathrm{cov}(\hat{\beta})\ \|},$$

此处 $|\ A\ |$ 表示矩阵 A 的行列式， $\mathrm{tr}(A)$ 表示矩阵 A 的迹， $\|\ A\ \|$ 表示矩阵 A 的欧氏模。

本文定义了一种新的相对效率

$$e_4(\hat{\beta}) = \Big[\frac{\mathrm{tr}(\mathrm{cov}(\beta^*))^q}{\mathrm{tr}(\mathrm{cov}(\hat{\beta}))^q}\Big]^{\frac{1}{q}}, \qquad q \geq 1.$$

并给出了相对效率的上下界以及与广义相关系数的联系。

本文共四章，第一章是基础知识，由三部分组成。第一节介绍线性模型的概念以及一般线性模型的形式；第二节介绍线性模型中的最小二乘估计与最佳线性无偏估计；第三节介绍广义 G-M 模型的概念。第二章介绍线性模型相对效率的概念，以及相对效率几种常用的表示方法。第三章是全文的主体，给出了广义 G-M 模型相对效率的概念，相对效率常用的几种表示方法，这一部分定义了一种新的相对效率，并给出了相对效率的上下界以及与广义相关系数的联系。最后第四章是结论部分。

第一章 基础知识

§ 1.1 线性模型

一 引言

研究一个 (或一组) 变量与另一个 (或另一组) 变量之间的关系，在几乎所有的科学分支中都是非常重要的。而变量之间的关系大体上可分为两类，一类是确定性关系，即数学上的函数关系；另一类是非确定性关系，又叫统计相关关系。例如某种农作物在单位面积上的产量 Y 与施肥量 X_1 、浇水量 X_2 、种籽 X_3(定性变量) 就有着非常密切的关系，但对给定的 $X_1 = x_1, X_2 = x_2, X_3 = x_3$ 来说， Y 还是不确定的。事实上， Y 是一随机变量，其概率分布与自变量 $(X_1, X_2, X_3)^T := X$ 所取的值 $(x_1, x_2, x_3)^T := x$ 密切相关。作为自变量的 X 可以是一维的，也可以是多维的；可以是定量变量，也可以是定性变量；可以是非随机变量，也可以是随机变量 (甚至一部分分量随机，另一部分分量不随机) 。因变量 (又叫响应变量) Y 与自变量 (又叫预测变量或预报变量) $X = (X_1, X_2, \cdots, X_{p-1})^T$(在上面的例子中 $p = 4$) 之间的这种统计相关关系可表示为

$$Y = \mu(x_1, x_2, \cdots, x_{p-1}) + \epsilon,$$

其中 $\mu(x_1, x_2, \cdots, x_{p-1})$ 是在给定 $X := (X_1, X_2, \cdots, X_{p-1})^T = (x_1, x_2, \cdots, x_{p-1})$ 的条件下，因变量 Y 的平均值 (条件数学期望) ，即

$$\mu(x_1, x_2, \cdots, x_{p-1}) = E(Y|X_1 = x_1, X_2 = x_2, \cdots, X_{p-1} = x_{p-1}),$$

且 $\mu(x_1, x_2, \cdots, x_{p-1})$ 中含有若干个 (比如说 p 个) 未知参数 $\beta_0, \beta_1, \cdots, \beta_{p-1}$ ，而 ϵ 为随机误差，满足 $E\epsilon = 0$.

特别,当 $\mu(x_1, x_2, \cdots, x_{p-1})$ 为未知参数 $\beta_0, \beta_1, \cdots, \beta_{p-1}$ 的线性关系,即 $\mu(x_1, x_2, \cdots, x_{p-1}) = \sum_{j=0}^{p-1} \beta_j f_j(x_1, x_2, \cdots, x_{p-1})$ 时,称因变量 Y 与自变量 $X = (X_1, X_2, \cdots, X_{p-1})^T$ 之间的统计相关关系为线性统计模型,简称为线性模型。这样,一个线性模型是因变量 Y 与自变量 $X = (X_1, X_2, \cdots, X_{p-1})^T$ 之间如下类型的一种统计相关关系

$$Y = \sum_{j=0}^{p-1} \beta_j f_j(x_1, x_2, \cdots, x_{p-1}) + \epsilon,$$

其中 β_j 为未知参数,$f_j(x_1, x_2, \cdots, x_{p-1})(j = 0, 1, \cdots, p-1)$ 为 $(x_1, x_2, \cdots, x_{p-1})^T := x$ 的某种已知函数 (不含未知参数),ϵ 为随机误差,满足 $E\epsilon = 0$,通常还假定 $D\epsilon = \sigma^2 > 0$.

二 线性回归模型

线性模型是一类统计模型的总称,它包括了线性回归模型、方差分析模型、协方差分析模型和线性混合效应模型 (或称方差分量模型) 等。

定义 若

$$\begin{aligned} \mu(x_1, x_2, \cdots, x_{p-1}) &= E(Y|X_1 = x_1, X_2 = x_2, \cdots, X_{p-1} = x_{p-1}) \\ &= \beta_0 + \beta_1 x_1 + \cdots + \beta_{p-1} x_{p-1}, \end{aligned}$$

其中 $\beta_0, \beta_1, \cdots, \beta_{p-1}$ 为未知参数,则称模型

$$Y = \mu(x_1, x_2, \cdots, x_{p-1}) + \epsilon, \quad E\epsilon = 0, \quad D\epsilon = \sigma^2$$

为线性回归模型。

易见,线性回归模型是一种特殊的线性模型,即 Y 关于 X 的理论回归函数是关于未知参数 $\beta_0, \beta_1, \cdots, \beta_{p-1}$ 的线性函数,且 β_0 的系数为 1 。

若在第 i 次实验 (观测) 中，自变量 $X_1 = x_{i1}, X_2 = x_{i2}, \cdots, X_{p-1} = x_{i,p-1}$, 因变量 Y 的观测值为 $Y_i, i = 1, 2, \cdots, n$, 独立，则应有

$$Y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_{p-1} x_{i,p-1} + \epsilon_i, \qquad i = 1, 2, \cdots, n$$

独立。

用矩阵的形式记这些方程，有

$$Y = X\beta + \epsilon,$$

其中

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, X = \begin{bmatrix} 1 & x_{11} \ldots & x_{1,p-1} \\ 1 & x_{21} \ldots & x_{2,p-1} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} \ldots & x_{n,p-1} \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{bmatrix}, \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix},$$

且 ϵ 满足 $E\epsilon = 0, D\epsilon = \sigma^2 I_n, \sigma^2 > 0$, 我们称此矩阵形式的回归模型为线性回归模型。

§ 1.2 LS 估计与 BLU 估计

一 LS 估计与 BLU 估计

考虑一般线性模型

$$y = X\beta + \epsilon, \quad \epsilon \sim N(0, \sigma^2 I_n), \tag{1.1}$$

其中 y 为 $n \times 1$ 观测向量， X 为 $n \times p$ 的设计阵， β 为 $p \times 1$ 的未知参数向量， ϵ 为 $n \times 1$ 的随机误差向量， $\sigma^2 > 0$ 为未知参数。如果矩阵 X 的秩 $\mathrm{rk}(X) = r \le p$,

称 (1.1) 为降秩线性模型，否则，称为满秩线性模型。我们先讨论参数 β 的估计问题。

获得参数向量估计的基本方法是最小二乘法，其思想是， β 的真值应该使误差向量 $\epsilon = y - X\beta$ 达到最小，也就是它的长度平方

$$Q(\beta) = \| \epsilon \|^2 = \| y - X\beta \|^2 = (y - X\beta)^T(y - X\beta)$$

达到最小。这一最小问题的正则方程为

$$X^T X\beta = X^T y.$$

正则方程的解为

$$\hat{\beta} = (X^T X)^- X^T y,$$

这里 $(X^T X)^-$ 是 $X^T X$ 的任意一个广义逆。

若矩阵 X 的秩 $\mathrm{rk}(X) = p$, 则 $X^T X$ 可逆，这时 $\hat{\beta} = (X^T X)^{-1} X^T y$ ，且有 $E(\hat{\beta}) = \beta$, 即 $\hat{\beta}$ 是 β 的无偏估计。我们称 $\hat{\beta} = (X^T X)^{-1} X^T y$ 为 β 的最小二乘估计 (least square estimate, 简记为 LS 估计).

若矩阵 X 的秩 $\mathrm{rk}(X) < p$ ，则 $E(\hat{\beta}) \neq \beta$, 即 $\hat{\beta}$ 不是 β 的无偏估计。更进一步，在此条件下不存在 β 的线性无偏估计，此时称 β 是不可估的。

定义 1.1　若存在 $n \times 1$ 向量 a, 使得 $E(a^T y) = c^T \beta$ 对一切 β 成立，则称 $c^T\beta$ 是可估函数 (estimate function).

定理 1.1[22]　$c^T\beta$ 是可估函数 $\Longleftrightarrow c \in \mathcal{M}(X^T)$ ，其中 $\mathcal{M}(X^T)$ 表示 X^T 的列向量所张成的空间。

定义 1.2　对可估函数 $c^T\beta$ ，称 $c^T\hat{\beta}$ 为 $c^T\beta$ 的 LS 估计.

定理 1.2(Gauss-Markov **定理**)[22]　对任意的可估函数 $c^T\beta$ ， LS 估计 $c^T\beta$ 为

其惟一的 BLU 估计。

二 广义最小二乘估计

考虑线性模型

$$y = X\beta + \epsilon, \qquad E(\epsilon) = 0, \qquad \mathrm{Cov}(\epsilon) = \sigma^2\Sigma, \tag{1.2}$$

其中 $\Sigma > 0$.

因为 $\Sigma > 0$ ，故存在惟一的正定对称阵 $\Sigma^{\frac{1}{2}}$. 用 $\Sigma^{-\frac{1}{2}}$ 左乘 (1.2) 式，并记 $\tilde{y} = \Sigma^{-\frac{1}{2}}y, \tilde{X} = \Sigma^{-\frac{1}{2}}X, u = \Sigma^{-\frac{1}{2}}\epsilon$, 则得到

$$\tilde{y} = \tilde{X}\beta + u, \qquad E(u) = 0, \qquad \mathrm{Cov}(u) = \sigma^2 I. \tag{1.3}$$

对模型 (1.3) 用最小二乘法求 β 的 LS 解，即解 $Q(\beta) = \parallel \tilde{y} - \tilde{X}\beta \parallel^2$ 的最小值问题。等价地，解

$$\min Q(\beta) = \min (y - X\beta)^T \Sigma^{-1} (y - X\beta), \tag{1.4}$$

其正则方程为

$$X^T\Sigma^{-1}X\beta = X^T\Sigma^{-1}y,$$

于是， β 的 LS 解为

$$\beta^* = (X^T\Sigma^{-1}X)^- X^T\Sigma^{-1}y,$$

称为广义最小二乘解。特别当 $\Sigma = \mathrm{diag}(\sigma_1{}^2, \ldots, \sigma_n{}^2), \sigma_i{}^2 (i = 1, \ldots, n)$ 已知时，称 β^* 为加权最小二乘解。我们称 $c^T\beta^*$ 为可估函数 $c^T\beta$ 的广义最小二乘估计 (generalized least squares estimate), 简记为 GLS 估计。

定理 1.3[22] 对任一可估函数 $c^T\beta, c^T\beta^*$ 为其惟一的 BLU 估计，其方差为

$$\mathrm{Var}(c^T\beta^*) = \sigma^2 c^T (X^T\Sigma^{-1}X)^- c.$$

§ 1.3　广义 G-M 模型

考虑线性模型

$$y = X\beta + \epsilon, \qquad E(\epsilon) = 0, \qquad \mathrm{Cov}(\epsilon) = \sigma^2\Sigma, \tag{1.5}$$

如果 $\Sigma \geq 0$ ，则称该模型为广义 G-M 模型，此时 Σ^{-1} 不存在，所以不能通过最小化 (1.4) 式所定义的 $Q(\beta)$ 来求得 β 的最小二乘估计。

著名统计学家 Rao 应用推广的最小二乘法 [22] 所导出的估计以其形式简单便于理论研究而得到普遍采用，本文也采用这种方法。定义

$$T = \Sigma + XUX^T,$$

其中 $U \geq 0$ ，矩阵 T 的秩满足 $\mathrm{rk}(T) = \mathrm{rk}(\Sigma \vdots X)$, 令

$$Q(\beta) = (y - X\beta)^T T^-(y - X\beta),$$

最小化 $Q(\beta)$ ，求出其最小值点

$$\beta^* = (X^T T^- X)^- X^T T^- y.$$

定理 1.4[22]　对于线性模型 (1.5) 和任一可估函数 $c^T\beta$,

$$c^T\beta^* = c^T(X^T T^- X)^- X^T T^- y$$

为 $c^T\beta$ 的 BLU 估计。