累积法引论

曹定爱 张顺明 著

科学出版社

累积法引论

曹定爱　张顺明 著

科学出版社

1999

内 容 简 介

本书系统地介绍了计量经济学中的新方法：累积法的数学理论及其应用，其主要内容包括：最小二乘法及其应用；多阶段最小二乘法；累积法及其应用；多级估计法及其应用；间接累积法与多阶段累积法。

本书可供高校有关专业师生阅读，也可供计量经济工作者参阅。

图书在版编目(CIP)数据

累积法引论/曹定爱等著.-北京：科学出版社，1999.1
ISBN 7-03-007109-3

Ⅰ.累… Ⅱ.曹… Ⅲ.①累积误差 ②最小二乘法 Ⅳ.O 241.1

中国版本图书馆 CIP 数据核字(98)第 32962 号

科 学 出 版 社 出版
北京东黄城根北街 16 号
邮政编码：100717

新 蕾 印 刷 厂 印刷

新华书店北京发行所发行 各地新华书店经售

*

1999 年 1 月第 一 版 开本：850×1168 1/32
1999 年 1 月第一次印刷 印张：6
印数：1— 2000 字数：150 000

定价：15.00 元

（如有印装质量问题，我社负责调换〈新欣〉）

谨以本书来怀念我们尊敬的老师：

王毓云教授！

序（一）

Econometrics，有时译作经济计量学，有时译作计量经济学。研究经济学的人士多半采用前一种译名，他们认为经济计量学是一种用经验估算方法研究经济变量间相互关系的理论和方法，它可用来进行经济计量分析，而当这一计量分析方法用于历史学、科学学等其他学科的研究时，就产生了历史计量学、科学计量学等等。不是研究经济学的人士多半采用后一种计量经济学的译名。他们把计量经济学等同于数量经济学，即研究经济数量关系及其变化规律的经济学科，但这一译法的不足之处是分不清计量经济学与研究计量工作及其相关问题的经济学之区别。

经济计量学的分析步骤包括定式（Specification）、估算（Estimation）、检验（Validation）、应用（Application）。传统的经济计量学派把重点放在估算上，他们认为估算参数这一环节最能反映经济计量分析的特点，它依靠统计推断的理论和方法，利用样本资料，估测模型内结构参数的数值。革新的经济计量学派则把重点转移到定式上，强调经济理论的重要性，把合适的经济理论表述成为可作数理统计处理的数学模型。

对结构参数的估算方法有两类，一类是最小二乘法（LS），也叫最小平方法，在这类把函数最小化以求得参数估算值的方法，又有普通最小二乘法（OLS）、二阶段最小二乘法（2SLS）、三阶段最小二乘法（3SLS）、迭代的最小二乘法（TLS）等。另一类是最大似然法（ML），即把函数最大化以求得参数估算值的方法，在这类方法中包括有限信息最大似然法（LIML）、充分信息最大似然法（FLML）等。选用哪种估算方法，与系统的结构有关，需视模型和资料的情况以及经济计量估算人员的擅长而定。一般地说，二阶段最小二乘法估算的一致性较好，有限信息

最大似然法比较灵敏。经常有这样的情况，即对复杂的模型还得使用较简单的估算方法，如 OLS 法。

OLS 法的数学理论基础是由数学家 Guass 和 Markov 奠定的。但这种方法使结构参数有所偏离和不一致，且平均误差较大。曹定爱、张顺明两位同志根据意大利数学家 P. Marchsi 创立的累积法，探寻出累积算子的各阶通式，先是用来估计单一方程模型的参数，创造了累积算子表，提出可以取代普通二乘法的普通累积法，进而用来估计联立方程模型的参数，开发出间接累积法、二阶段累积法、三阶段累积法，以及广义累积法、多级估计法等一列估算参数的补充方法。

累积法原是一种曲线拟合技术，可用于数据平滑的时间数列分析。本书《累积法引论》的作者发展了这种方法，把它用于经济计量模型的结构参数的估算，以补最小二乘法之不足，如不够简便、估计误差较大等。这无疑是有益的探索和可贵的贡献。

《累积法引论》全书包括两个部分。第一部分分两章对最小二乘法作了说明和评价,这是本书的铺垫。第二部分分三章对累积法作了介绍和论述,这是本书的核心和精髓。作者在这部分就累积法的数学原理、基本概念、计算公式、估算分析、实际应用、误差处理、推广估计等各个问题展开研讨,并在相关部分的实证分析中与最小二乘法的结果进行比较考察,从而得出累积法是可以取代最小二乘法而成为估算经济计量模型结构参数新方法的结论。

当然，一种新的估算方法，从提出到被公认以至普及，还有许多工作要做。例如，对有关假设进一步作严格的理论论证；对新方法的应用范围和条件进行必要的相应说明；编新方法的计算机应用程序，推出规范的便于操作的应用软件，等等。我衷心希望在本书出版后的不久，作者以及有志于估算方法改进创新的其他同志，大家一起共同努力，来推进上述这些工作。

乌家培

一九九八年七月十二日

序（二）

计量经济学是一门新型的实证经济科学，自本世纪相交年代应运而生以来，已走过了 80 多年成长、演变和发展的路程。其间，由于几代学者坚持不懈的钻研和无私的奉献，计量经济学的基础理论和基本方法日趋成熟和完整，它在实证性的经济分析、决策研究、信息管理等方面，正发挥着愈益重要和积极的作用，并且为越来越多的经济理论工作者和广大经营管理人员所接受，这是无需多说的。

然而，就计量经济学的基本方法而言，至今沿用的仅是德国数学家高斯（K. Gauss，1777—1855）首创的最小二乘法及系列化体系，简言之，就是最小二乘法体系。当然，在估计各类计量经济学参数的过程中，最小二乘法体系依旧是独一无二的估计法。但随着社会经济体制、经济运行机制和决策机构的演变、发展和完善，各种经济变量之间的关系日趋复杂和多变，使人感到继续沿用最小二乘法体系，就会带来许多困难。尤其重要的是，理论上，最小二乘法有先天之不足，即误差处理要以系列无法验证的假设为必备前提，这是不合理和不科学的。有鉴于此，各国计量经济学家早已开始从事计量经济学新方法的研究工作，但除本书之外，至今无显著的突破。本书著者的贡献是：在创建计量经济学新方法体系的路径上，迈进了令人惊喜的一步，这将是本学科诞生以来的又一创新，必定引起各国计量经济学家的浓厚兴趣和广泛重视。

本书系统介绍的计量经济学新方法体系即累积法体系，最初是从意大利数学家马尔奇西（P. Marchsi）首创的累积法雏形中得到启发的。累积法曾试用于数据平滑的时间序列分析中，而在估计计量经济模型参数的过程中则没有得到运用。因累积法的计

算过程十分复杂，加上马尔奇西本人没有进一步归纳出一套通式来简化其运用的必不可少的累积算子（书中用语）的计算程序，所以正如书中介绍的那样，要用 30 组数据求知一个经济变量的二阶累积和，须写出该组数据不同期数资料共 4960 次，其三阶累积和为 40920 次，其四阶累积和为 278256 次，依此类推，其九阶累积和 163008020 次之多，即是说，阶数越多，计算就越复杂和烦琐。显而易见，这种烦杂的计算势必影响累积法在模型参数估计中的运用和推广，因而长期处于无人问津的地步。

然而问题的发现和寻找其解决方法的可能性是同时存在的。我还记得本书著者在上海财经大学经济学系读研究生期间，就已开始从事这项研究工作，并且十分敏锐地看出，普及累积法的关键在于，探寻累积算子的各阶通式，以概括客观规律性。于是就废寝忘食地开展研究工作，花费了大量的精力，经过无数次计算和验证，终于发现了累积算子的各阶通式，这是在我应邀赴日本讲学回国的第二天给我看的，至今记忆犹新。

从此以后，他就顺利完成了用累积法估计单一方程模型参数工作。其特点是，不直接处理误差项，而是运用有规律可循的累积和，径自估计模型参数。它既不同于数学中的方程组解法，又不同于最小二乘估计法。其效果不但与最小二乘法相同，而且比它简明得多，由于找到累积和的计算通式，并且据此创造了累积算子表，累积法的运用和推广就走上了一条平坦的轨道。例如，计算 30 个样本的八阶累积和，只需写出该样本各期资料各一次（共 30 次）就行，而不需要写出 163008020 次之多。这一改进和突破，从根本上为普及累积法，使之成为独立而完整的崭新的估计法体系开辟了一条捷径。本书把这种用于估计单一方程模型参数的累积法，仿照普通最小二乘法的称谓，简称普通累积法。

紧接着，著者把累积法推广到联立方程模型参数的估计中，相继建立了间接累积法、二阶段累积法、三阶段累积法，以及广义累积法、多级估计法等补充方法，使之成为一整套体系，因而可以替代与之相对应的最小二乘法，而不影响模型参数的估计和

应用。

显然，这些成果都带有创造性，尽管其中包括借鉴前人成果的成分，但对计量经济学新方法体系的产生而言，毕竟创新为主，因而即使意大利数学家马尔奇西（P. Marchsi）本人看到后人的这一整套成果，也会感到十分欣慰。

总之，《累积法引论》是一本有很高学术价值的好书，其在计量经济学新方法体系及其实际运用方面作出的贡献是突出的，应当予以肯定。

当然，作为一本开拓新园地的专著，难免有许多不是之处，诸如论证不够严谨和周到，所用样本较少，模型设计中个别变量遗漏等等，但这些难免的因素并不影响本书的学术价值，因为它至少把计量经济学的一套新方法奉献给读者。因此，我们不宜像对待最小二乘法体系的计量经济学著作那样，要求其内容完善和严密，其不是之处应由学界同行的共同努力来加以弥补。

最后，我衷心希望本书著者继续发扬勇于探索创新的精神，不断完善计量经济学的新方法——累积法的体系，将其推向成熟(包括电子计算机应用程序在内)，也希望学界专家教授给予指导和帮助，使其成为学界的共同财富。我坚信，通过学界同行的共同努力，累积法体系必将与最小二乘法并驾齐驱，充当计量经济学领域的另一主角。

李柱锡

写于日本神户大学交换教授室

1998 年 3 月 9 日

目　　录

第一章　最小二乘法及其应用

第一节　最小二乘法的数学理论

普通最小二乘法(Ordinary Least Square,简写为 OLS)是估计单个方程多元线性计量经济模型的主要方法,也是其他特殊估计方法的基础。计量经济学和数理统计学中最小二乘法是在下面一般线性模型上得到的,最小二乘理论与线性模型中的参数估计有关。这种理论的基础是数学家 Gauss(1809)和 Markov(1900)所奠定的。

1. 线性模型

设因变量 Y 和自变量 $X_1,X_2,\cdots,X_n$ 服从线性关系

$$Y = \beta_1 X_1 + \cdots + \beta_n X_n + \varepsilon \tag{1.1}$$

$(Y_k, x_{k1},\cdots,x_{kn}),k=1,\cdots,m$ 是$(Y,X_1,\cdots,X_n)$的 m 个观察,它们满足关系

$$Y_k = \beta_1 x_{k1} + \cdots + \beta_n x_{kn} + \varepsilon_k,\quad k = 1,\cdots,m \tag{1.2}$$

记

$$Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_m \end{pmatrix},\quad X = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \cdots & \cdots & \cdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix}$$

$$\beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix},\quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_m \end{pmatrix}$$

则(1.2)式可表示为

$$Y = X\beta + \varepsilon \tag{1.3}$$

其中 X 是已知的 $m \times n$ 常数矩阵，β 是未知参数向量，ε 是均值为零的随机向量，满足条件

$$E\varepsilon = 0 \qquad \operatorname{cov}(\varepsilon, \varepsilon) = \sigma^2 I_m \tag{1.4}$$

I_m 是 $m \times m$ 单位阵，σ^2 是未知参数。

习惯上，我们称上述的 Y 服从线性模型，简单记为 $(Y, X\beta, \sigma^2 I_m)$。

由(1.3)和(1.4)式，容易计算得出

$$EY = X\beta, \qquad \operatorname{cov}(Y, Y) = \sigma^2 I_m \tag{1.5}$$

或者写为

$$EY_k = \sum_{t=1}^{n} x_{kt}\beta_t$$

$$\operatorname{cov}(Y_{k_1}, Y_{k_2}) = \begin{cases} \sigma^2, k_1 = k_2 \\ 0, k_1 \neq k_2 \end{cases} \quad k_1, k_2 = 1, 2, \cdots, m$$

对于由(1.3)式和(1.4)式所确定的线性模型 $(Y, X\beta, \sigma^2 I_m)$，通常所考虑的统计推断问题是：对未知参数 β 和 σ^2 进行估计；对关于 β 的某种假设进行检验；对于 Y 进行预报等等。

注 1　一般情形，记 $\operatorname{cov}(\varepsilon, \varepsilon) = \sigma^2 \Sigma$，注意 Σ 作为 ε 的协方差矩阵是正定或者半正定的。如果 Σ 是正定矩阵（$|\Sigma| \neq 0$），那么 $\operatorname{cov}(\Sigma^{-\frac{1}{2}}\varepsilon, \Sigma^{-\frac{1}{2}}\varepsilon) = \sigma^2 I_m$，我们可以作线性变换 $Y' = \Sigma^{-\frac{1}{2}} Y$，$X' = \Sigma^{-\frac{1}{2}} X$ 和 $\varepsilon' = \Sigma^{-\frac{1}{2}}\varepsilon$，从而

$$Y' = X'\beta + \varepsilon' \tag{1.6}$$

满足条件

$$E\varepsilon' = 0, \quad \operatorname{cov}(\varepsilon', \varepsilon') = \sigma^2 I_m \tag{1.7}$$

这也就是说 Y' 服从线性模型 $(Y', X'\beta, \sigma^2 I_m)$。如果 Σ 是半正定矩阵（$|\Sigma| = 0$），我们在一般线性模型中详细讨论。

在下面的讨论中，我们假定 $m \geq n$。

2. $\boldsymbol{\beta}$ 的估计

对于线性模型 $(Y, X\beta, \sigma^2 I_m)$，常采用最小二乘法寻找未知参

数 β 的估计量$\hat{\beta}$,它要求 β 的估计$\hat{\beta}$满足下面的条件

$$\| Y - X\hat{\beta} \|^2 = \min_{\beta} \| Y - X\beta \|^2 = \min_{\beta}(Y - X\beta)^T(Y - X\beta) \tag{1.8}$$

满足(1.6)式的估计$\hat{\beta}_L = \begin{pmatrix} \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_n \end{pmatrix}$称为$\beta$ 的最小二乘法估计,这种求估计量的方法称为最小二乘法。

下面用微分法求满足(1.6)式的解$\hat{\beta}_L$。记

$$\begin{aligned} f(\beta) &= (Y - X\beta)^T(Y - X\beta) \\ &= Y^TY - 2\beta^TX^TY + \beta^TX^TX\beta \end{aligned}$$

令$\dfrac{\partial f}{\partial \beta}=0$ 得方程组 $-2X^TY+2X^TX\beta=0$,即

$$X^TY = (X^TX)\beta \tag{1.9}$$

(1.9)式称为正规方程。

首先考虑 rank $X=n$ 的情形,如果 rank $X<n$,我们也在一般线性模型中详细讨论。

由于 X 的秩为 n,所以 X^TX 是正定矩阵,存在逆矩阵$(X^TX)^{-1}$,解方程(1.9),得到 β 的最小二乘估计是

$$\hat{\beta}_L = (X^TX)^{-1}X^TY \tag{1.10}$$

将$\hat{\beta}_L$ 代入方程(1.1),由于 ε 一般表示模型误差不可预测,略去 ε 就得到因变量 Y 和自变量$X_1,\cdots,X_n$ 之间的经验关系式$\hat{Y}=X\hat{\beta}$,即

$$\hat{Y} = \hat{\beta}_1X_1 + \cdots + \hat{\beta}_nX_n \tag{1.11}$$

(1.11)式常称为经验模型或经验回归方程,当自变量取值为$(X_1,\cdots,X_n)$时。它可以用来预测因变量 Y 的平均值。

最小二乘估计$\hat{\beta}_L$ 具有如下基本性质:

性质 1.1 $\hat{\beta}_L$ 是 Y 的线性函数,即$\hat{\beta}_L$ 是一线性估计。

性质1.2 $\hat{\beta}_L$ 是β的无偏估计。

证明
$$
\begin{aligned}
E\hat{\beta}_L &= E[(X^TX)^{-1}X^TY] \\
&= (X^TX)^{-1}X^TEY \\
&= (X^TX)^{-1}X^TX\beta \\
&= \beta
\end{aligned}
$$

性质1.3 $\hat{\beta}_L$ 的协方差矩阵是$\sigma^2(X^TX)^{-1}$。

证明
$$
\begin{aligned}
\sum\nolimits_{\hat{\beta}_L} &= \mathrm{cov}(\hat{\beta}_L, \hat{\beta}_L) \\
&= (X^TX)^{-1}X^T\mathrm{cov}(Y,Y)X(X^TX)^{-1} \\
&= (X^TX)^{-1}X^T\sigma^2 I_m X(X^TX)^{-1} \\
&= \sigma^2(X^TX)^{-1}
\end{aligned}
$$

性质1.4(Gauss-Markov 定理) $\hat{\beta}_L$ 是β的最小方差线性无偏估计。

在这里,若 T 是β的另一线性无偏估计,有

$$\sum\nolimits_{\hat{\beta}_L} \leq \sum\nolimits_T = \mathrm{cov}(T,T)$$

其中$\leq$表示$\sum_T - \sum_{\hat{\beta}_L}$是非负定矩阵,则称$\hat{\beta}_L$ 的方差不大于T 的方差。

证明 设 T 是β的任一线性无偏估计,则 T 可以表示为

$$T = CY$$

其中 C 是$n \times m$ 矩阵,由无偏性要求得

$$ET = E[CY] = CEY = CX\beta = \beta \tag{1.12}$$

由于(1.12)式要求对一切 β 成立,所以

$$CX = I_m$$

而

$$
\begin{aligned}
\sum\nolimits_T &= \mathrm{cov}(T,T) \\
&= \mathrm{cov}(CY, CY) \\
&= C\mathrm{cov}(Y,Y)C^T \\
&= C\sigma^2 I_m C^T
\end{aligned}
$$

$$= \sigma^2 CC^T \tag{1.13}$$

因为

$$\begin{aligned} 0 &\leq [C - (X^TX)^{-1}X^T][C - (X^TX)^{-1}X^T]^T \\ &= CC^T + (X^TX)^{-1} - (X^TX)^{-1}X^TC^T - CX(X^TX)^{-1} \\ &= CC^T + (X^TX)^{-1} - (X^TX)^{-1} - (X^TX)^{-1} \\ &= CC^T - (X^TX)^{-1} \end{aligned}$$

所以

$$CC^T \geq (X^TX)^{-1} \tag{1.14}$$

从而

$$\sum\nolimits_{\hat{\beta}_L} = \sigma^2(X^TX)^{-1} \leq \sigma^2 CC^T = \sum\nolimits_T$$

由于 T 是任意选取的一个线性无偏估计,所以最小二乘估计 $\hat{\beta}_L$ 是 β 的最小方差线性无偏估计。

下面给出最小二乘估计的几何解释。记 m 维向量 $x_t = \begin{pmatrix} x_{1t} \\ \vdots \\ x_{mt} \end{pmatrix}$, $t=1,\cdots,n$,它是矩阵 X 的第 t 列元素,用 $\mu(X)$ 表示由向量 x_t, $t=1,\cdots,n$ 的全部线性组合所构成的一个线性空间,则(1.8)式表示要求在 $\mu(X)$ 中寻找一个向量 $\xi = \sum_{t=1}^{n} \beta_t x_t$,使得 ξ 与 Y 之间的距离 $\| Y - \xi \|$ 达到最小。而只有当 ξ 是 Y 在 $\mu(X)$ 中的投影时,才使 $\| Y - \xi \|$ 达到最小。

注2 注 1 中的线性模型 $(Y, X\beta, \sigma^2 \Sigma)$,也就是线性模型 $(Y', X'\beta, \sigma^2 I_m)$,其中 $Y' = \Sigma^{-\frac{1}{2}} Y$, $X' = \Sigma^{-\frac{1}{2}} X$ 和 $\varepsilon' = \Sigma^{-\frac{1}{2}} \varepsilon$, β 的估计 $\hat{\beta}$

$$\begin{aligned} \| Y' - X'\hat{\beta} \|^2 &= \min_{\beta} \| Y' - X'\beta \|^2 \\ &= \min_{\beta} (Y' - X'\beta)^T (Y' - X'\beta) \\ &= \min_{\beta} (Y - X\beta)^T \sum\nolimits^{-1} (Y - X\beta) \end{aligned} \tag{1.15}$$

用微分法求满足(1.15)式的解$\hat{\beta}_L$。记

$$f(\beta) = (Y - X\beta)^T \sum^{-1} (Y - X\beta)$$
$$= Y^T \sum^{-1} Y - 2\beta^T X^T \sum^{-1} Y + \beta^T X^T \sum^{-1} X\beta$$

令$\frac{\partial f}{\partial \beta}=0$得方程组$-2X^T\sum^{-1}Y+2X^T\sum^{-1}X\beta=0$,即得正规方程

$$X^T \sum^{-1} Y = (X^T \sum^{-1} X)\beta \tag{1.16}$$

由于X的秩为n,所以$X^T\sum^{-1}X$是正定矩阵,存在逆矩阵$(X^T\sum^{-1}X)^{-1}$,解方程(1.16),得到β的最小二乘估计是

$$\hat{\beta}_L = (X^T \sum^{-1} X)^{-1} X^T \sum^{-1} Y \tag{1.17}$$

最小二乘估计$\hat{\beta}_L$仍然具有如下基本性质:

性质1.5 $\hat{\beta}_L$是Y的线性函数,即$\hat{\beta}_L$是一线性估计。

性质1.6 $\hat{\beta}_L$是β的无偏估计。

性质1.7 $\hat{\beta}_L$的协方差矩阵是$\sigma^2(X^TX)^{-1}$。

性质1.8(Gauss-Markov 定理) $\hat{\beta}_L$是β的最小方差线性无偏估计。

注3 如果rank $X<n$,那么某些参数就没有无偏估计,而在我们的结构下,对这样的参数,不能作出任何推断。假如$(X^TX)^-$是X^TX的一个广义逆,那么正规方程(1.9)的通解是$(X^TX)^- X^TY+[I-(X^TX)^- X^TX]Z$,其中$Z$是任意$n\times 1$向量;而$(X^TX)^- X^TY$就是正规方程(1.9)的一个特解。

3. σ^2 的估计

记

$$\widetilde{Y} = Y - X\hat{\beta}_L = [I_m - X(X^TX)^{-1}X^T]Y \tag{1.18}$$

为剩余向量,表示Y被估计后所剩下的残差,故又称为残差向量。记R_0^2是剩余平方和

$$R_0^2 = \widetilde{Y}^T \widetilde{Y}$$

$$
\begin{aligned}
&= (Y - X\hat{\beta}_L)^T(Y - X\hat{\beta}_L) \\
&= Y^T[I_m - X(X^TX)^{-1}X^T]Y \\
&= Y^TY - \hat{\beta}_L^T(X^TY) \qquad (1.19)
\end{aligned}
$$

性质1.9 $\widetilde{Y}$与$\hat{\beta}_L$互不相关。

证明

$$
\begin{aligned}
&\text{cov}(\widetilde{Y}, \hat{\beta}_L) \\
&= [I_m - X(X^TX)^{-1}X^T]\text{cov}(Y, Y)[(X^TX)^{-1}X^T]^T \\
&= \sigma^2[I_m - X(X^TX)^{-1}X^T][X(X^TX)^{-1}] = 0 \qquad (1.20)
\end{aligned}
$$

从几何解释来看这一性质是显然的,那里$\widetilde{Y}$表示 Y 的垂线,(1.20)式意味着$\hat{Y}$和$\widetilde{Y}$互相垂直。

性质1.10 残差$\widetilde{Y}$的均值向量和协方差矩阵分别是

$$
E[\widetilde{Y}] = E[Y - X\hat{\beta}_L] = 0 \qquad (1.21)
$$

$$
\begin{aligned}
\text{cov}(\widetilde{Y}, \widetilde{Y}) &= D([I_m - X(X^TX)^{-1}X^T]Y) \\
&= \sigma^2[I_m - X(X^TX)^{-1}X^T] \qquad (1.22)
\end{aligned}
$$

记

$$
\hat{\sigma}^2 = \frac{1}{m-n}R_0^2
$$

称为剩余方差。

性质1.11 $\hat{\sigma}^2$是 σ^2 的无偏估计,即 $E[\hat{\sigma}^2] = \sigma^2$。

证明

$$
\begin{aligned}
E[R_0^2] &= E[\widetilde{Y}^T\widetilde{Y}] \\
&= tr\{\text{cov}(\widetilde{Y}, \widetilde{Y})\} \\
&= \sigma^2 tr\{I_m - X(X^TX)^{-1}X^T\} \\
&= \sigma^2(m - tr\{X(X^TX)^{-1}X^T\}) \\
&= \sigma^2(m - tr\{(X^TX)^{-1}X^TX\})
\end{aligned}
$$