"十二五"普通高等教育本科国家级规划教材

新编21世纪经济学系列教材

计量经济学

第五版

Econometrics

主 编 赵国庆

中国人民大学出版社

“十二五”普通高等

新编21世纪经济学系列教材

计量经济学

第五版

Econometrics

主　编　赵国庆

中国人民大学出版社
·北京·

图书在版编目（CIP）数据

计量经济学/赵国庆主编. —5版. —北京：中国人民大学出版社，2016.6
新编21世纪经济学系列教材
ISBN 978-7-300-22897-6

Ⅰ.①计… Ⅱ.①赵… Ⅲ.①计量经济学-高等学校-教材 Ⅳ.①F224.0

中国版本图书馆CIP数据核字（2016）第103894号

"十二五"普通高等教育本科国家级规划教材
新编21世纪经济学系列教材
计量经济学（第五版）
主　编　赵国庆
Jiliang Jingjixue

出版发行	中国人民大学出版社		
社　　址	北京中关村大街31号	**邮政编码**	100080
电　　话	010－62511242（总编室）		010－62511770（质管部）
	010－82501766（邮购部）		010－62514148（门市部）
	010－62515195（发行公司）		010－62515275（盗版举报）
网　　址	http://www.crup.com.cn		
	http://www.ttrnet.com（人大教研网）		
经　　销	新华书店		
印　　刷	北京鑫丰华彩印有限公司	**版　　次**	2001年2月第1版
规　　格	185mm×260mm　16开本		2016年6月第5版
印　　张	15.75	**印　　次**	2016年6月第1次印刷
字　　数	371 000	**定　　价**	32.00　元

主编简介

赵国庆，1996年获日本京都大学经济学博士学位，现任中国人民大学经济学院教授、博士生导师。兼任浙江大学教授、日本关西学院大学客座教授、中国数量经济学会常务理事、学术委员。

研究方向：计量经济学理论与应用

主要成果：

1. Uncovering the Relationship between FDI, Human Capital and Technological Progress in Chinese High-technology Industries, *China & World Economy*, 2010 (with Zhang Z.).
2. Heuristics for Replenishment with Linear Decreasing Demand, *International Journal of Production Economics*, 2001 (with Yang J. and Rand, G. K.).
3. Unit Root Analyses of the Causality between Japanese Money and Income, *Japanese Economic Review*, 1997 (with Morimune, K.).

引 言

经济理论的数量化研究是计量经济学（econometrics）的目的，包括经济模型的设计、建立、估计、检验及使用经济模型进行预测和政策评价的整个过程。随着经济的飞速发展，统计学、经济理论、数学已成为理解现代经济中的数量关系所不可缺少的条件，计算机的日益更新和计量软件的多样化又给现代经济的数量化研究提供了强有力的工具，这些条件的组合形成了计量经济分析的基础。

本书旨在使学生理解计量经济模型的思想，掌握常用的计量经济模型。同时本书研究计量经济模型在经济领域中分析问题以及辅助决策的作用和功能。本书通过各类经济问题训练学生的数量化基本技能，提高学生数学分析水平，尤其是培养学生对各类经济问题的研究能力及综合分析能力。本书不仅介绍计量经济学的基础理论，而且训练学生参与计量经济模型的整个过程，以便学生在今后的工作中更好地运用计量经济模型的方法和技巧解决实际问题。

本书作为经济类大学本科的计量经济学课程教材，与一般计量经济学的入门教科书相比，内容要广泛一些。具体表现在构造变化的 F 检验、分布滞后模型、离散选择模型、受限因变量模型、面板数据分析、经济变量的单位根检验和协整性分析等内容。这些内容我们认为是大学本科计量经济学课程必须包括的部分。如果不使用计算机，上述方法的运用很复杂，但随着近年计算机和各种计量经济学软件的发展，对于本书包含的上述内容，本科生完全可以接受和掌握。

本书把计量经济学的方法应用到国内外诸多经济实例中，并且对估计结果在统计上给出解释的同时，对其经济含义也进行了详细的讨论。主要包括：消费函数、菲利普斯曲线、生产函数、进口函数、Klein 模型和误差修正模型（ECM）等经济实例。

本书的编著者都是长期从事计量经济学、宏观经济学、微观经济学教学与研究的学者，他们当中大多有国外留学和讲学的经历。具体分工如下：中国人民大学沈民鸣（第一、二章），中国人民大学赵国庆（第三、四章），武汉大学何耀（第五章），上海交通大学朱保华（第六章），武汉大学童光荣（第七章），南开大学张晓峒（第八章），全书由主编赵国庆负责定稿。

本书在编著过程中始终得到中国人民大学和中国人民大学出版社的关心与支持，相关的编辑也为本书的出版付出了辛勤的劳动，在此表示由衷的感谢。

本书中的一些经济实例引自国家“211”工程项目“中国宏观经济运行模拟和分析系

统”，在此向国家信息中心预测部和“211”工程项目组表示感谢。

本书作为大学本科经济学类核心课程“计量经济学”教材的一个尝试，还存在着不少不足之处，恳请各位专家和读者指正。

赵国庆

第五版说明

经济理论的数量化研究是计量经济学的目的，构造计量经济模型是研究经济现象的重要方法，它使人们能够客观解释经济现象背后的形成机制。近年计算机的日益更新与计量软件的多样化发展也给计量经济方法的应用提供了强有力的支撑。过去计量经济学的学习，多是在掌握计量经济理论的基础上，对经济模型进行统计推断，今天则是在估算模型的同时学习必要的计量分析理论。规避难懂的计量理论的推导，掌握必要的计量分析方法与计量软件的时代已经到来，对于计量经济学本身而言这可能是一个发展趋势。《计量经济学》的第五版基于这一理念对教材进行了全面修订。第五版主要在以下几个方面做出修改。一是对例题中案例的数据进行更新，二是把复杂的证明推导过程调整到章后的附录，三是删去部分复杂的推导证明与结论。

第五版的具体分工如下：沈民鸣（第一、二章，中国人民大学副教授），赵国庆（第三、四章，中国人民大学教授），何耀（第五章，武汉大学教授），朱保华（第六章，上海交通大学教授），童光荣（第七章，武汉大学教授），张晓峒（第八章，南开大学教授），全书由主编赵国庆统一进行最后的修改定稿。

教材中案例的数据见中国人民大学出版社网页（www.crup.com.cn/jingji），或者给 zhaogq@163.com 发邮件。

在本书的定稿过程中，中国人民大学信息学院范红岗博士，山东师范大学尹慧博士，国家信息中心预测部邬琼博士，中国人民大学数量经济学专业研究生李本钊、惠炜、姚青松等参与了部分工作，中国人民大学出版社王美玲编辑也为本书的出版付出了辛勤的劳动，值本书付梓之际，向他们表示感谢。

《计量经济学》一书从第一版（2001 年）开始，一直得到各位专家学者的关心与支持。本书为教育部推荐教材、北京高等教育精品教材、普通高等教育“十一五”国家级规划教材与“十二五”普通高等教育本科国家级规划教材。主编赵国庆为北京高等学校精品课程《计量经济学》的负责人。在此向一直支持我们的各位专家与读者表示由衷的感谢。

赵国庆

2016 年 5 月

人大明德楼

目录

第一章

一元线性回归分析基础

回归分析是计量经济分析的基础，对于两个变量 X 和 Y 之间关系的度量，在 X 是原因、Y 是结果的因果关系已知的情况下，回归分析比相关分析更为有效。本章将讨论一元线性回归模型的主要估计方法——最小二乘法，包括估计结果的解释、假设检验与预测等内容。

第一节　模型的假定

社会经济活动可以用某些经济变量形式表示，例如，投资额、产量、销售量、价格、利润、利率、股票价格、国内生产总值、人均国民生产总值等。在生产、分配、交换和消费过程中，各种生产要素、产品等，无论是以实物形态出现，还是以货币形态表示，最终总要表现为一定的数量。对经济问题的研究，不仅要分析该问题的基本性质，也需要对经济变量之间的数量关系进行具体分析。常用的分析方法有回归分析、相关分析、方差分析等。这些分析方法各有特点，其中应用最广泛的是回归分析。相关分析和方差分析本身虽然可以独立进行某些方面的数量分析，但是在大多数情况下，则是与回归分析结合在一起，进行综合分析，并且作为回归分析方法的补充分析方法。

一、一元线性回归模型

各种经济变量之间的关系，可以划分为两种类型。一类是变量之间有唯一确定的关系，即函数关系。例如，X_1，X_2，…，X_n 与 Y 之间的函数关系可以用隐函数形式表示为

$$F(X_1, X_2, \cdots, X_n, Y)=0 \tag{1—1}$$

也可以用显函数形式表示为

$$Y=f(X_1, X_2, \cdots, X_n) \tag{1—2}$$

其中，最简单的形式为一元线性函数关系。例如，当某种商品单价 P 固定不变时，这种商品的销售收入 Y 与销售的商品数量 X 之间的关系为一元线性关系，即

$$Y=PX \tag{1—3}$$

如果用 X 和 Y 构成的直角坐标图来表示，式（1—3）所表示的函数关系为一条经过坐标

原点的直线。所有可能的点都在这条直线上。

经济变量之间的另一类关系为不完全确定的相关关系。例如，家庭消费支出 Y 与家庭收入 X 之间的关系，就不是完全确定的。虽然每个家庭的收入 X 必然会影响并且制约着这个家庭的消费支出 Y，但是消费支出 Y 还要受到其他多种因素的影响。例如，家庭人口、消费习惯、银行存款利率、商品价格水平变化趋势等。即使对于同一个家庭在每月收入相同的条件下，每月的消费支出也不会完全相同。这类变量之间不完全确定的关系可以表示为

$$F(X_1, X_2, \cdots, X_n, Y, u) = 0 \tag{1—4}$$

或者表示为

$$Y = f(X_1, X_2, \cdots, X_n, u) \tag{1—5}$$

其中最简单的形式为一元线性回归模型

$$Y = \beta_1 + \beta_2 X + u \tag{1—6}$$

其中 u 包含了除家庭收入 X 之外的各种对家庭消费支出 Y 有影响的因素。这种不完全确定的关系，如果用 X 和 Y 构成的直角坐标图表示，可以表示为平面上的一系列散点，如图 1—1 所示。

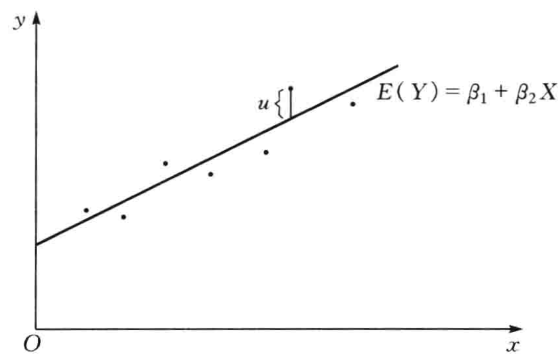

图 1—1　不完全确定的相关关系

这些点均匀分布在直线

$$Y = \beta_1 + \beta_2 X \tag{1—7}$$

附近，但是不完全落在该直线上。散点到该直线的垂直坐标距离为 u。

因为绝大多数经济变量都受到多种其他经济变量的影响，所以变量之间有完全确定的函数关系的情况在经济问题中很少见。通常变量之间完全确定的函数关系，如式（1—1），不属于计量经济学讨论的范围。计量经济学只讨论变量之间不完全确定的关系，如式（1—4）或式（1—5）所表示的关系。

式（1—6）所表示的关系式，称为一元线性回归模型。

“一元”是指只有一个自变量 X，这个自变量 X 可以解释引起因变量 Y 变化的部分原因。因此，X 称为解释变量，Y 称为被解释变量，β_1 和 β_2 为参数。其中 β_1 决定了直线 $Y=\beta_1+\beta_2 X$ 的截距，β_2 决定了该直线的斜率。参数 β_2 确定了解释变量 X 影响被解释变量 Y 的基本关系，不确定的部分由变量 u 表示。u 称为误差项。

“线性”一词在这里有两重含义。它一方面指被解释变量 Y 与解释变量 X 之间为线性

关系，另一方面也指 Y 与参数 β_1，β_2 之间为线性关系。

在数学分析中，“线性”一般指 Y 与 X 为线性关系。在计量经济学中，更重视被解释变量 Y 与参数 β_1，β_2 之间的线性关系。只要 Y 与 β_1，β_2 之间满足线性关系，即使 Y 与 X 不为线性关系，也可以通过线性变换，使变换后的被解释变量与解释变量之间的关系实现线性化。

“回归”一词来自生物学。在遗传现象中，后代的遗传特征与前代的特征有关系。在前代的特征偏离该物种特征平均值的情况下，其后代的遗传特征有可能比前代偏离特征平均值更远，也有可能不如前代偏离该物种平均值那么远。后一种情况出现的概率高于前一种情况出现的概率，即后代的遗传特征有返回该物种特征平均值的倾向，从而使该物种的基本特征得以延续。人们把这种现象称为回归。在数理统计学中，“回归”通常指散点分布在一条直线（或曲线）附近，并且越靠近该直线（或曲线），点的分布越密集的情况。

“模型”一词通常指满足某些假设条件的方程或方程组。以前面提到的家庭收入 X 与家庭消费支出 Y 之间的关系为例，每个家庭的消费支出 Y 主要取决于该家庭的收入 X，但是也要受到其他多种因素影响。一般来说，收入较高的家庭消费支出比较高，这类高收入家庭在消费支出方面的选择余地比较大。其中，一部分家庭选择高消费，也有一部分家庭选择比较低的消费和比较高的储蓄或投资。从统计角度来看，这类家庭的消费支出的离散性比较大，即方差比较大。对于低收入家庭来说，消费支出的选择余地就小得多。尽管不同家庭在收入相同的情况下消费支出有差别，但是支出中的大部分用于维持必不可少的基本生活消费。这类家庭消费支出的离散性比较小，即方差比较小。

一般来说，消费支出 Y 的分布函数是多种多样的，不一定是正态分布，也不一定是相同的分布。分布函数的方差、均值都不相同，分布函数的形式也不同，如图 1—2 所示。

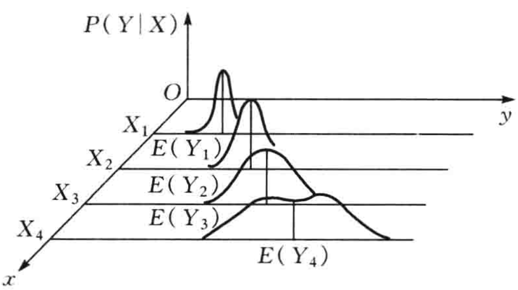

图 1—2　不同 X 对应的 Y 的一般概率分布

从图 1—2 中可以看出，家庭消费支出 Y 是家庭收入 X 的条件概率函数 $P(Y|X)$。这个概率函数 $P(Y|X)$ 有三个明显特征：

(1) 对于不同的 X，条件概率函数 $P(Y|X)$ 的分布函数形式不同。

(2) 对于不同的 X，条件概率函数 $P(Y|X)$ 的方差不同。

(3) 对于不同的 X，条件概率函数 $P(Y|X)$ 的均值 $E(Y)$ 一般不在同一条直线上。

对于这样的概率函数进行数学分析是非常困难的，目前还没有比较好的解决办法。为了简化数学分析，通常对实际情况进行抽象，作一些假设：

(1) 假设概率函数 $P(Y|X)$ 的分布函数形式相同。例如，服从正态分布。

（2）假设概率函数 $P(Y|X)$ 的分布函数的方差相同，均为常数 σ_u^2。即

$$\mathrm{var}(Y_t)=\mathrm{var}(u_t)=\sigma_u^2 \qquad (t=1,\ 2,\ \cdots,\ n) \tag{1—8}$$

（3）对于不同的 X，Y 的均值 $E(Y)$ 在同一条直线上。即

$$E(Y_t)=\beta_1+\beta_2 X_t \qquad (t=1,2,\cdots,\ n) \tag{1—9}$$

这个假设是满足一元线性回归要求的。

满足这些假设条件的 Y 的概率分布函数可以用图 1—3 表示。

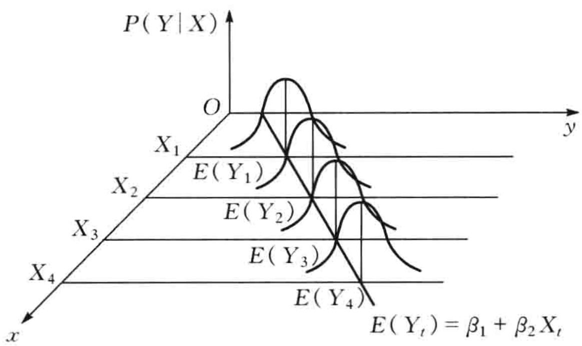

图 1—3 满足经典假设条件的 Y 的概率分布

二、误差项的性质

与精密数学中的函数关系相比，回归模型式（1—4）、式（1—5）、式（1—6）中的显著特点是多了误差项 u。误差项 u 包含了丰富的内容。产生误差项的原因主要有以下几方面：

1. 被忽略的影响因素造成的误差

在一般情况下，每一个经济变量通常要受到多种因素的影响。但是为了简化分析，突出主要矛盾，在构造回归模型时，通常只选取最重要的解释变量与被解释变量构成回归模型，将次要的影响因素忽略。这些被忽略的影响因素对被解释变量 Y 的影响就归入了误差项 u。以家庭消费支出问题为例。家庭消费支出 Y 要受家庭收入、家庭人口、消费习惯、存款利率、商品价格水平变化趋势等多种因素影响。如果家庭收入 X 能够解释家庭消费支出 Y 的变化原因的大部分，就可以忽略其他影响因素。这些次要影响因素对 Y 的影响总和都进入了误差项 u。

2. 模型关系不准确造成的误差

在一般情况下，解释变量与被解释变量之间的关系是比较复杂的非线性关系。在构造模型时，为了简化模型，用线性模型代替了非线性关系，或者用简单的非线性模型代替了复杂的非线性关系，造成了模型关系不准确的误差。在很多情况下，由于人们对经济规律的认识与客观经济规律本身不完全一致，也会造成模型关系不准确的误差。

3. 变量观察值的计量误差

测量工具的精确度和测量方法不正确使得观察值与真实值不完全一致，从而造成误差。

4. 随机误差

对以上三种误差，总可以通过改变模型形式，改进测量设备和技术来减小相应的误差。但是经济变量本身受很多随机因素影响，不具有确定性和重复性。同时，社会经济问

题涉及人的思维和行动，也涉及各阶级、各阶层的物质利益。人的行为具有很多不确定因素。由此造成的误差是随机的，随机误差无法减小，这些随机误差也归入误差项 u 中。

总之，误差项的存在是计量经济学模型的特点，是计量经济学模型与精密数学中完全确定的函数关系的主要区别。计量经济学中遇到的各种困难问题几乎都是由误差项 u 的存在造成的。计量经济学中的多种估计、检验、预测等分析方法，也是针对不同性质的误差项 u 引入的。

三、经典假设条件

经典的一元线性回归模型

$$Y_t=\beta_1+\beta_2 X_t+u_t \qquad (t=1,\ 2,\ \cdots,\ n) \tag{1—10}$$

通常要满足五个假设条件：

假设 1　误差项 u_t 的数学期望（均值）为零，即

$$E(u_t)=0 \qquad (t=1,\ 2,\ \cdots,\ n) \tag{1—11}$$

假设 2　误差项 u_t 的方差与 t 无关，为一个常数，即

$$\begin{aligned}\mathrm{var}(u_t)&=E((u_t-E(u_t))^2)\\&=E(u_t^2)\\&=\sigma_u^2 \qquad (t=1,\ 2,\ \cdots,\ n)\end{aligned} \tag{1—12}$$

假设 3　不同的误差项 u_t 和 u_s 之间不相关，即

$$\mathrm{cov}(u_t,u_s)=E((u_t-E(u_t))(u_s-E(u_s)))=0 \tag{1—13}$$

$$(t\neq s;\ t=1,\ 2,\ \cdots,\ n;\ s=1,\ 2,\ \cdots,\ n)$$

或

$$E(u_t u_s)=0 \tag{1—14}$$

假设 4　解释变量 X_t 与误差项 u_t 不相关，即

$$\begin{aligned}\mathrm{cov}(X_t,\ u_t)&=E((X_t-E(X_t))(u_t-E(u_t)))\\&=E((X_t-E(X_t))u_t)\\&=0 \qquad (t=1,\ 2,\ \cdots,\ n)\end{aligned} \tag{1—15}$$

如果 X_t 为非随机变量，这个假设自动满足，即

$$E(X_t)=X_t$$

并且

$$E(X_t u_t)=X_t E(u_t)=0 \tag{1—16}$$

也成立。

假设 5　u_t 为服从正态分布的随机变量，即

$$u_t \sim N(0,\ \sigma_u^2)$$

以上五个假设条件称为经典假设条件。

假设 1 表示模型式（1—10）为线性模型，即有 $E(Y_t|X_t)=\beta_1+\beta_2 X_t$。由于一元线性回归模型中有参数 β_1 一项，这个参数本身是待估计的，总可以改变 β_1 的数值，使

$$E(u_t)=0$$

成立，也就是说，如果模型式（1—10）成立，这个假设自动成立。

假设 2 表示无论 X_t 随 t 如何变化，误差项的方差不发生变化。对于不同的解释变量 X_t，如果误差项的方差不同，那么与其相对应的观察值 Y_t 的可靠程度也不相同。通常误差项 u_t 的方差小表示它所对应的观察值 Y_t 的可靠程度高，应给予较高程度的重视。相反，误差项 u_t 的方差大表示它所对应的观察值 Y_t 的可靠程度低，应给予较低程度的重视。即对不同的误差项 u_t 所对应的不同观察值 Y_t，应该加上不同的权数。这会使参数的检验和利用模型进行预测复杂化。而满足同方差假设，将使检验和预测简化。

假设 3 表示不同的误差项之间不相关。同时不同的被解释变量在统计上也是不相关的。即

$$\begin{aligned}\operatorname{cov}(Y_t, Y_s)&=E((Y_t-E(Y_t))(Y_s-E(Y_s)))\\&=E(u_t u_s)\\&=0\end{aligned}$$

其中，$t\neq s$。如果这个假设成立，参数的检验和利用模型进行预测将被简化。

假设 4 表示解释变量 X_t 与误差项 u_t 不相关。在一般情况下，X_t 为非随机变量。在这种情况下，这个假设条件将自动满足。

假设 5 表示误差项 u_t 服从正态分布。如果只利用最小二乘法进行参数估计，就不需要这个假设条件。但是，如果要进行假设检验和预测，就必须知道总体 Y_t 的分布情况。如果 X_t 为非随机变量，总体 Y_t 与误差项 u_t 服从相同的分布，Y_t 与 u_t 之间仅有均值 $E(Y_t)$ 的差别。

如果样本容量足够大，由于 u_t 代表了所有被忽略的对 Y_t 有影响的因素的总和，根据中心极限定理，只要这些影响因素是随机的和互相独立的，并且具有有限的数学期望和方差（这些条件通常都可以成立），那么，作为这些影响因素的总和，误差项 u_t 近似服从正态分布。因此，假设 5 对于检验和预测是必要的。

对于大样本问题，无论 u_t 中包含的每一种影响因素服从什么分布，u_t 都近似服从正态分布。即在大样本条件下这个假设自动成立。对于小样本问题，如果这个假设不成立，就无法进行检验和预测。

综上所述，一元线性回归模型可以归结为

$$\begin{aligned}&Y_t=\beta_1+\beta_2 X_t+u_t \qquad (t=1, 2, \cdots, n) \qquad (1—17)\\&E(u_t)=0\\&\operatorname{var}(u_t)=\sigma_u^2 \qquad (\text{常数})\\&\operatorname{cov}(u_t, u_s)=0 \qquad (t\neq s;\ t, s=1, 2, \cdots, n)\\&\operatorname{cov}(X_t, u_t)=0\\&u_t \sim N(0, \sigma_u^2)\end{aligned}$$

第二节　参数的最小二乘估计

对于家庭收入 X 影响家庭消费支出 Y 的问题，如果通过调查得到一组数据，可列表

显示，见表1—1。

表1—1　**家庭收入与家庭消费支出**　单位：元

	家庭收入 X	家庭消费支出 Y
1	800	770
2	1 200	1 100
3	2 000	1 300
4	3 000	2 200
5	4 000	2 100
6	5 000	2 700
7	7 000	3 800
8	9 000	3 900
9	10 000	5 500
10	12 000	6 600

这组数据可以用图1—4表示。

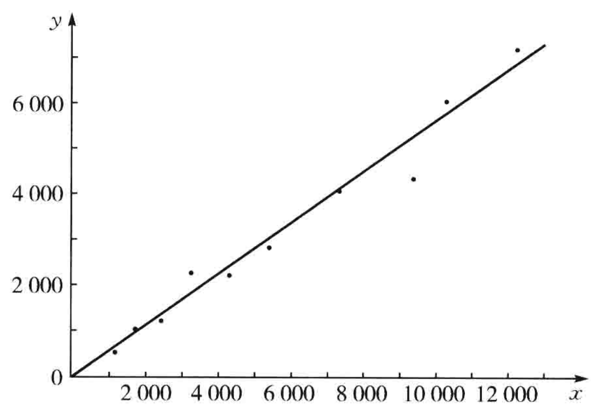

图1—4　家庭收入与家庭消费支出

假设这组数据满足一元线性回归模型式（1—17），如何找到一条直线

$$\hat{Y}_t=\hat{\beta}_1+\hat{\beta}_2X_t$$

使这条直线尽可能靠近所有的点？这个问题在几何学中等价于寻找一条拟合散点的直线。

首先，必须讨论如何判断一条直线是否靠近所有样本点。也就是说，拟合一条直线的准则是什么。

一、拟合准则与最小二乘估计

直线外一个点到直线上的点的距离有三种特殊情况：

（1）点到直线的垂直距离。

（2）点到直线的垂直坐标距离。

（3）点到直线的水平坐标距离。

其中，点到直线的水平坐标距离在回归分析中不使用。点到直线的垂直距离在几何中是指

该点与垂足之间的距离，但是计算不太方便，在回归分析中一般也不使用。回归分析中使用的是点到直线的垂直坐标距离，即（$Y_t-\hat{Y}_t$）。当点在直线上方时，垂直坐标距离为正值。当点在直线下方时，垂直坐标距离为负值。

拟合一条直线的准则有以下几种情况：

1. 使 $\left|\sum_{t=1}^{n}(Y_t-\hat{Y}_t)\right|$ 达到最小值

这种准则在计算上是最简单的。但是有一个严重的缺点：由于正负偏差在求和过程中互相抵消，即使大多数点远离该直线，也有可能使 $\left|\sum_{t=1}^{n}(Y_t-\hat{Y}_t)\right|$ 达到最小值的要求得到满足。

2. 使 $\sum_{t=1}^{n}|Y_t-\hat{Y}_t|$ 达到最小值

由于求和的每一项均为绝对值，不存在正负抵消的问题。这种准则的缺点是不能保证找到的直线具有无偏性。

3. 使 $\max|Y_t-\hat{Y}_t|$ 达到最小值

这种准则是使直线垂直坐标距离最远的点到直线对应的点的距离达到最小值。这种准则的缺点是容易受到远离大多数散点的个别点的干扰。也就是说，这种准则只考虑个别点的影响，没有考虑多数点的影响。

4. 使 $\sum_{t=1}^{n}(Y_t-\hat{Y}_t)^2$ 达到最小值

由于逐项平方，这种准则不存在正负抵消的问题。它不仅考虑了所有点的影响，而且具有无偏性，是一个很好的准则。这个准则称为最小二乘准则。用最小二乘准则寻找拟合直线的方法称为最小二乘法。

假设 $\hat{\beta}_1$ 和 $\hat{\beta}_2$ 分别为 β_1 和 β_2 的估计值。拟合直线为

$$\hat{Y}_t=\hat{\beta}_1+\hat{\beta}_2X_t \qquad (t=1, 2, \cdots, n) \tag{1—18}$$

观察值 Y_t 与 $\hat{Y}_t$ 不完全相等。

$$Y_t=\hat{Y}_t+e_t=\hat{\beta}_1+\hat{\beta}_2X_t+e_t \tag{1—19}$$

其中，e_t 称为残差。最小二乘准则是选择不同的 $\hat{\beta}_1$（即直线的截距）和 $\hat{\beta}_2$（即直线的斜率），使 $\sum_{t=1}^{n}(Y_t-\hat{Y}_t)^2$ 达到最小值，即使 $\sum_{t=1}^{n}e_t^2$ 达到最小值。

残差平方和 $\sum_{t=1}^{n}e_t^2$ 为估计量 $\hat{\beta}_1$ 与 $\hat{\beta}_2$ 的函数，由残差平方和最小化的一阶条件，有方程组

$$\begin{cases}\dfrac{\partial\left(\sum_{t=1}^{n}e_t^2\right)}{\partial\hat{\beta}_1}=\dfrac{\partial\left(\sum_{t=1}^{n}(Y_t-\hat{\beta}_1-\hat{\beta}_2X_t)^2\right)}{\partial\hat{\beta}_1}=0 & (1—20)\\[2ex] \dfrac{\partial\left(\sum_{t=1}^{n}e_t^2\right)}{\partial\hat{\beta}_2}=\dfrac{\partial\left(\sum_{t=1}^{n}(Y_t-\hat{\beta}_1-\hat{\beta}_2X_t)^2\right)}{\partial\hat{\beta}_2}=0 & (1—21)\end{cases}$$