Homework Week01: Classic Model
最小二乘法与模型形式

胡华平

2024-09-26

学生姓名：\_\_\_\_\_\_\_\_\_\_\_\_\_\_；学生学号：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_；专业班级：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**问：** 你将使用什么编程分析软件完成本次作业？

**答：** 我选择的分析软件是：\_\_\_\_\_\_\_\_。

**温馨提示**：建议大家优先选择R、python、stata。当然，也可以使用Matlab或EViews的编程功能。

# 作业提交

**作业发布时间**：2023-09-11（周一）24:00:00

**提交截止时间**：2023-09-24（周日）24:00:00

**作业提交材料**：

（1）根据作业要求，完成Office Word电子文档一份（注意不能是wps文档），提交前请将文件命名为下述格式：lab01\_word\_张三\_2019000001.docx。

（2）根据作业要求，完成相关编程分析操作，保存并提交1份**原始代码文件**（其中10位数字为学生的学号）：

* 如果使用R软件，请将编程代码文件保存并命名为：lab01\_code\_Rscript\_2019000001.r；或者lab01\_code\_Rmarkdown\_2019000001.Rmd。
* 如果使用Python软件，请将编程代码文件保存并命名为：lab01\_code\_Python\_2019000001.py。
* 如果使用Stata软件，请将编程代码文件保存并命名为：lab01\_code\_Stata\_2019000001.do。
* 如果使用Matlab软件，请将编程代码文件保存并命名为：lab01\_code\_Matlab\_2019000001.m。
* 如果使用EViews软件，请将编程代码文件保存并命名为：lab01\_code\_EViews\_2019000001.prg。

**作业提交方式**：

* 按上述要求命名各个文件，然后将全部作业提交材料压缩为zip格式，并命名为：lab01-专业年级-姓名-学号.zip（如：lab01-2023应经-张帅帅-2023120208.zip）
* 发送上述zip文件到电子邮箱：huhuaping01@qq.com。请填写“邮件主题”为：lab01-专业年级-姓名-学号（如：lab01-2023应经-张帅帅-2023120208）。
* 点击发送邮件！你将收到一份邮件已收到的“自动回复”！

# 1. 作业案例

## 1.1 案例数据集

下面的图 [图 1](#fig-data-show) 给出数据取自1980-1982年间英国家庭支出调查中1519个家庭的家庭食物支出（$foodexp$，作为被解释变量$Y$）和家庭总支出（$totexp$，作为解释变量$X$）。数据只包括住在伦敦市区和市郊有1~2个子女的家庭，样本不包括自我雇佣和退休家庭。

**温馨提示**： （1）作业配套数据请在作业发布界面中自行下载。 （2）表中只筛选了部分数据作为展示。

|  |
| --- |
| 图 1: 英国家庭食物支出数据集（n=1519） |

## 1.2 多种模型函数形式及计算参考

考虑到本次实验课内容，操作运算中可能需要计算不同模型形势下的弹性和斜率，下面表 [表 1](#tbl-calcu) 给出理论计算公式，供大家在操作中参照使用。

**温馨提示**： （1）$Y$对$X$的斜率的理论表达式为$b=\frac{dY}{dX}$；$Y$对$X$的弹性的理论表达式为$η=\frac{dY}{dX}⋅\frac{X\_{i}}{Y\_{i}}$。 （2）二者的关系为：$η=b⋅\frac{X\_{i}}{Y\_{i}}$。

Warning: Using `all\_of()` outside of a selecting function was
deprecated in tidyselect 1.2.0.
ℹ See details at
 <https://tidyselect.r-lib.org/reference/faq-selection-context.html>

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 表 1: 模型函数形式及斜率和弹性的计算参考

| 序号 | 名称 | 表达式 | 斜率 | 点弹性 | 平均弹性 |
| --- | --- | --- | --- | --- | --- |
| $M\_{1}$ | 线性模型 | $Y\_{i}=β\_{1}+β\_{2}X\_{i}+u\_{i}$ | $β\_{2}$ | $β\_{2}\frac{X\_{i}}{Y\_{i}}$ | $β\_{2}\frac{‾}{‾}$ |
| $M\_{2}$ | 过原点模型 | $Y\_{i}=β\_{2}X\_{i}+u\_{i}$ | $β\_{2}$ | $β\_{2}\frac{X\_{i}}{Y\_{i}}$ | $β\_{2}\frac{‾}{‾}$ |
| $M\_{3}$ | 双对数模型 | $ln\left(Y\_{i}\right)=β\_{1}+β\_{2}ln\left(X\_{i}\right)+u\_{i}$ | $β\_{2}\frac{Y\_{i}}{X\_{i}}$ | $β\_{2}$ | $β\_{2}$ |
| $M\_{4}$ | 线性到对数模型 | $ln\left(Y\_{i}\right)=β\_{1}+β\_{2}X\_{i}+u\_{i}$ | $β\_{2}Y\_{i}$ | $β\_{2}X\_{i}$ | $β\_{2}‾$ |
| $M\_{5}$ | 对数到线性模型 | $Y\_{i}=β\_{1}+β\_{2}ln\left(X\_{i}\right)+u\_{i}$ | $\frac{β\_{2}}{X\_{i}}$ | $\frac{β\_{2}}{Y\_{i}}$ | $\frac{β\_{2}}{‾}$ |
| $M\_{6}$ | 倒数模型 | $Y\_{i}=β\_{1}+\frac{β\_{2}}{X\_{i}}+u\_{i}$ | $−\frac{β\_{2}}{X\_{i}^{2}}$ | $−\frac{β\_{2}}{\left(X\_{i}Y\_{i}\right)}$ | $−\frac{β\_{2}}{\left(‾‾\right)}$ |
| $M\_{7}$ | 对数倒数模型 | $ln\left(Y\_{i}\right)=β\_{1}+\frac{β\_{2}}{X\_{i}}+u\_{i}$ | $−β\_{2}\frac{Y\_{i}}{X\_{i}^{2}}$ | $−\frac{β\_{2}}{X\_{i}}$ | $−\frac{β\_{2}}{‾}$ |

 |

# 2. 作业任务

## 2.1 任务1：多种形式的OLS回归

根据研究目标，我们将尝试使用软件对如下7个模型进行OLS回归分析。

$$\begin{matrix}Y\_{i}&=β\_{1}+β\_{2}X\_{i}+u\_{i}&&(m1)\end{matrix}  \left(1\right)$$

$$\begin{matrix}Y\_{i}&=β\_{2}X\_{i}+u\_{i}&&(m2)\end{matrix}  \left(2\right)$$

$$\begin{matrix}ln\left(Y\_{i}\right)&=β\_{1}+β\_{2}ln\left(X\_{i}\right)+u\_{i} &&(m3)\end{matrix}  \left(3\right)$$

$$\begin{matrix}ln\left(Y\_{i}\right)&=β\_{1}+β\_{2}X\_{i}+u\_{i}&&(m4)\end{matrix}  \left(4\right)$$

$$\begin{matrix}Y\_{i}&=β\_{1}+β\_{2}ln\left(X\_{i}\right)+u\_{i}&&(m5)\end{matrix}  \left(5\right)$$

$$\begin{matrix}Y\_{i}&=β\_{1}+β\_{2}/X\_{i}+u\_{i}&&(m6)\end{matrix}  \left(6\right)$$

$$\begin{matrix}ln\left(Y\_{i}\right)&=β\_{1}+β\_{2}/X\_{i}+u\_{i}&&(m7)\end{matrix}  \left(7\right)$$

请依次完成如下回归分析过程。

**温馨提示**： （1）$ln\left(\right)$变换操作要注意选择对数的底，大家可以直接使用$log\_{10}\left(\right)$，因为大部分统计软件默认都是以10为底的。

（1）构建过经典线性回归模型 [式 1](#eq-m1) ，利用数据软件中进行回归分析操作，保存回归结果并命名为eq\_m1。最后截图到下列空白处。

答：

（2）构建过原点回归模型 [式 2](#eq-m2) ，利用数据在软件中进行回归分析操作，保存回归结果，并命名为eq\_m2。最后截图到下列空白处。

答：

（3）构建双对数回归模型 [式 3](#eq-m3) ，利用数据在软件中进行回归分析操作，保存回归结果并命名为eq\_m3。最后截图到下列空白处。

答：

（4）构建线性到对数回归模型 [式 4](#eq-m4) ，利用数据软件中进行回归分析操作，保存回归结果并命名为eq\_m4。最后截图到下列空白处。

答：

（5）构建对数到线性回归模型 [式 5](#eq-m5) ，利用数据在软件中进行回归分析操作，保存回归结果并命名为eq\_m5。最后截图到下列空白处。

答：

（6）构建倒数回归模型 [式 6](#eq-m6) ，利用数据在软件中进行回归分析操作，保存回归结果并命名为eq\_m6。最后截图到下列空白处。

答：

（7）构建对数倒数回归模型 [式 7](#eq-m7) ，利用数据在软件中进行回归分析操作，保存回归结果并命名为eq\_m7。最后截图到下列空白处。

答：

## 2.2 任务2：提取7个方程的斜率系数

（1）给定数据点$\left(X\_{0},Y\_{0}\right)=\left(100,30\right)$，请分别在软件中创建标量对象，并分别命名为x0和y0。

答：此问不用作答，完成指定操作即可！

（2）根据样本数据点$\left(X\_{i},Y\_{i}\right)$，请分别在软件中计算样本均值$\overline{X},\overline{Y}$对象，并分别命名为x\_bar和y\_bar。

答：样本均值分别为$\overline{X}=$\_\_\_\_\_；$\overline{Y}=$\_\_\_\_

（3）提取所有方程的回归斜率系数。请利用软件分别提取前述7个回归方程的斜率系数$\hat{β}\_{2}$。你需要创建1个$\left(7×1\right)$的向量对象，保存并命名为beta2\_hat。最后请将beta2\_hat的结果值复制/截图到下面空白处。

答：7个回归方程的斜率系数$\hat{β}\_{2}$结果值截图如下

## 2.3 任务3：计算点斜率

根据前面的回归分析结果，给定数据点$\left(X\_{0},Y\_{0}\right)=\left(100,30\right)$，分别计算各模型$Y\_{i}$对$X\_{i}$的点斜率$b\_{m} \left(m=1,2,…,7\right)$。

**温馨提示**：我们可以利用矩阵计算，一次性算出所有的点斜率向量$b\_{\left(7×1\right)}$。关键在于准确得到两个向量：回归方程的斜率向量$\hat{β}\_{\left(2,m\right)}$，以及常数向量$c\_{m}$。根据点斜率计算 [表 1](#tbl-calcu) ，它可以转换为如下的矩阵计算问题。其中$⊙$表示“矩阵要素相乘”（也即对应矩阵元素直接相乘），注意不是“矩阵相乘”！

$$\begin{matrix}&&b\_{m}=&&\hat{β}\_{\left(2,m\right)}⊙&&c\_{m}\\&&​\_{\left(1×7\right)}=&&​\_{\left(1×7\right)}⊙&&​\_{\left(1×7\right)}\end{matrix}$$

$$\begin{matrix}\left(\begin{matrix}b\_{1}\\b\_{2}\\b\_{3}\\b\_{4}\\b\_{5}\\b\_{6}\\b\_{7}\end{matrix}\right)=\left(\begin{matrix}\hat{β}\_{21}\\\hat{β}\_{12}\\\hat{β}\_{23}\\\hat{β}\_{24}\\\hat{β}\_{25}\\\hat{β}\_{26}\\\hat{β}\_{27}\end{matrix}\right)⊙\left(\begin{matrix}c\_{1}\\c\_{2}\\c\_{3}\\c\_{4}\\c\_{5}\\c\_{6}\\c\_{7}\end{matrix}\right)=\left(\begin{matrix}\hat{β}\_{21}\\\hat{β}\_{12}\\\hat{β}\_{23}\\\hat{β}\_{24}\\\hat{β}\_{25}\\\hat{β}\_{26}\\\hat{β}\_{27}\end{matrix}\right)⊙\left(\begin{matrix}1\\1\\\frac{Y\_{0}}{X\_{0}}\\Y\_{0}\\\frac{1}{X\_{0}}\\−\frac{1}{X\_{0}^{2}}\\−\frac{Y\_{0}}{X\_{0}^{2}}\end{matrix}\right)=\left(\begin{matrix}\hat{β}\_{21}\\\hat{β}\_{12}\\\hat{β}\_{23}\frac{Y\_{0}}{X\_{0}}\\\hat{β}\_{24}Y\_{0}\\\hat{β}\_{25}\frac{1}{X\_{0}}\\−\hat{β}\_{26}\frac{1}{X\_{0}^{2}}\\−\hat{β}\_{27}\frac{Y\_{0}}{X\_{0}^{2}}\end{matrix}\right)\end{matrix}$$

请在软件中计算得到7个模型中$Y\_{i}$对$X\_{i}$的斜率$b\_{m} \left(m=1,2,…,7\right)$。请参考计算表 [表 1](#tbl-calcu) 。分步骤回答如下问题：

（1）（步骤1）首先，你需要根据给定的数据点$\left(X\_{0},Y\_{0}\right)=\left(100,30\right)$，构建1个$\left(7×1\right)$的向量对象，保存并命名为c\_m，用于存放计算常数量值（参看计算表 [表 1](#tbl-calcu) 和上述提示）。 请将常数向量对象c\_m的结果复制或截图到如下空白处。

答：

（2）（步骤2）：然后，利用矩阵要素相乘的方法（参看计算表@tbl-calcu 和上述提示），构建并计算1个$\left(7×1\right)$的向量对象，保存并命名为b\_point，用于存放点斜率的计算量值（参看计算表@tbl-calcu）。 请将向量对象b\_point的结果复制或截图到如下空白处。

答：

## 2.4 任务4：计算点弹性

根据前面的软件回归分析结果，给定数据点$\left(X\_{0},Y\_{0}\right)=\left(100,30\right)$，分别计算各模型$Y\_{i}$对$X\_{i}$的点弹性$η\_{m} \left(m=1,2,…,7\right)$。

**温馨提示**：我们可以利用矩阵计算，一次性算出所有的点斜率向量$η\_{\left(7×1\right)}$。关键在于准确得到两个向量：回归方程的斜率向量$\hat{β}\_{\left(2,m\right)}$，以及常数向量$g\_{m}$。根据点斜率计算表 [表 1](#tbl-calcu) ，它可以转换为如下的矩阵计算问题。其中$⊙$表示“矩阵要素相乘”（也即对应矩阵元素直接相乘），注意不是“矩阵相乘”！

$$\begin{matrix}&&η\_{m}=&&\hat{β}\_{\left(2,m\right)}⊙&&g\_{m}\\&&​\_{\left(1×7\right)}=&&​\_{\left(1×7\right)}⊙&&​\_{\left(1×7\right)}\end{matrix}$$

$$\begin{matrix}\left(\begin{matrix}η\_{1}\\η\_{2}\\η\_{3}\\η\_{4}\\η\_{5}\\η\_{6}\\η\_{7}\end{matrix}\right)=\left(\begin{matrix}\hat{β}\_{21}\\\hat{β}\_{12}\\\hat{β}\_{23}\\\hat{β}\_{24}\\\hat{β}\_{25}\\\hat{β}\_{26}\\\hat{β}\_{27}\end{matrix}\right)⊙\left(\begin{matrix}g\_{1}\\g\_{2}\\g\_{3}\\g\_{4}\\g\_{5}\\g\_{6}\\g\_{7}\end{matrix}\right)=\left(\begin{matrix}\hat{β}\_{21}\\\hat{β}\_{12}\\\hat{β}\_{23}\\\hat{β}\_{24}\\\hat{β}\_{25}\\\hat{β}\_{26}\\\hat{β}\_{27}\end{matrix}\right)⊙\left(\begin{matrix}\frac{X\_{0}}{Y\_{0}}\\\frac{X\_{0}}{Y\_{0}}\\1\\X\_{0}\\\frac{1}{Y\_{0}}\\−\frac{1}{X\_{0}Y\_{0}}\\−\frac{1}{X\_{0}}\end{matrix}\right)=\left(\begin{matrix}\hat{β}\_{21}\frac{X\_{0}}{Y\_{0}}\\\hat{β}\_{12}\frac{X\_{0}}{Y\_{0}}\\\hat{β}\_{23}\\\hat{β}\_{24}X\_{0}\\\hat{β}\_{25}\frac{1}{Y\_{0}}\\−\hat{β}\_{26}\frac{1}{X\_{0}Y\_{0}}\\−\hat{β}\_{27}\frac{1}{X\_{0}}\end{matrix}\right)\end{matrix}$$

请在软件中计算得到7个模型中$Y\_{i}$对$X\_{i}$的点弹性$η\_{m} \left(m=1,2,…,7\right)$。请参考计算表 [表 1](#tbl-calcu) 。分步骤回答如下问题：

（1）（步骤1）首先，你需要根据给定的数据点$\left(X\_{0},Y\_{0}\right)=\left(100,30\right)$，构建1个$\left(7×1\right)$的向量对象，保存并命名为g\_m，用于存放计算常数量值（参看计算表 [表 1](#tbl-calcu) 和上述提示）。 请将常数向量对象g\_m的结果复制或截图到如下空白处。

答：

（2）（步骤2）：然后，利用矩阵要素相乘的方法（参看计算表 [表 1](#tbl-calcu) 和上述提示），构建并计算1个$\left(7×1\right)$的向量对象，保存并命名为eta\_point，用于存放点弹性的计算量值（参看计算表 [表 1](#tbl-calcu) ）。 请将向量对象eta\_point的结果复制或截图到如下空白处。

答：

## 2.5 任务5：计算平均弹性

根据前面的回归分析结果，根据样本数据$\left(X\_{i},Y\_{i}\right)$，分别计算各模型$Y\_{i}$对$X\_{i}$的平均弹性$\overline{η}\_{m} \left(m=1,2,…,7\right)$。

**温馨提示**：我们可以利用矩阵计算，一次性算出所有的点斜率向量$\overline{η}\_{\left(7×1\right)}$。关键在于准确得到两个向量：回归方程的斜率向量$\hat{β}\_{\left(2,m\right)}$，以及常数向量$‾\_{m}$。根据点斜率计算表 [表 1](#tbl-calcu) ，它可以转换为如下的矩阵计算问题。其中$⊙$表示“矩阵要素相乘”（也即对应矩阵元素直接相乘），注意不是“矩阵相乘”！

$$\begin{matrix}&&\overline{η}\_{m}=&&\hat{β}\_{\left(2,m\right)}⊙&&‾\_{m}\\&&​\_{\left(1×7\right)}=&&​\_{\left(1×7\right)}⊙&&​\_{\left(1×7\right)}\end{matrix}$$

$$\begin{matrix}\left(\begin{matrix}\overline{η}\_{1}\\\overline{η}\_{2}\\\overline{η}\_{3}\\\overline{η}\_{4}\\\overline{η}\_{5}\\\overline{η}\_{6}\\\overline{η}\_{7}\end{matrix}\right)=\left(\begin{matrix}\hat{β}\_{21}\\\hat{β}\_{12}\\\hat{β}\_{23}\\\hat{β}\_{24}\\\hat{β}\_{25}\\\hat{β}\_{26}\\\hat{β}\_{27}\end{matrix}\right)⊙\left(\begin{matrix}‾\_{1}\\‾\_{2}\\‾\_{3}\\‾\_{4}\\‾\_{5}\\‾\_{6}\\‾\_{7}\end{matrix}\right)=\left(\begin{matrix}\hat{β}\_{21}\\\hat{β}\_{12}\\\hat{β}\_{23}\\\hat{β}\_{24}\\\hat{β}\_{25}\\\hat{β}\_{26}\\\hat{β}\_{27}\end{matrix}\right)⊙\left(\begin{matrix}\frac{‾}{‾}\\\frac{‾}{‾}\\1\\‾\\\frac{1}{‾}\\−\frac{1}{‾‾}\\−\frac{1}{‾}\end{matrix}\right)=\left(\begin{matrix}\hat{β}\_{21}\frac{‾}{‾}\\\hat{β}\_{12}\frac{‾}{‾}\\\hat{β}\_{23}\\\hat{β}\_{24}‾\\\hat{β}\_{25}\frac{1}{‾}\\−\hat{β}\_{26}\frac{1}{‾‾}\\−\hat{β}\_{27}\frac{1}{‾}\end{matrix}\right)\end{matrix}$$

请在软件中计算得到7个模型中$Y\_{i}$对$X\_{i}$的平均弹性$\overline{η}\_{m} \left(m=1,2,…,7\right)$。请参考计算表 [表 1](#tbl-calcu) 。分步骤回答如下问题：

（1）（步骤1）首先，你需要根据给定的数据点$\left(X\_{i},Y\_{i}\right)$算出均值点$\left(‾,‾\right)$，构建1个$\left(7×1\right)$的向量对象，保存并命名为g\_m\_bar，用于存放计算常数量值（参看计算表 [表 1](#tbl-calcu) 和上述提示）。 请将常数向量对象g\_m\_bar的结果复制或截图到如下空白处。

答：

（2）（步骤2）：然后，利用矩阵要素相乘的方法（参看计算表 [表 1](#tbl-calcu) 和上述提示），构建并计算1个$\left(7×1\right)$的向量对象，保存并命名为eta\_mean，用于存放点弹性的计算量值（参看计算表 [表 1](#tbl-calcu) ）。 请将向量对象eta\_mean的结果复制或截图到如下空白处。

答：

## 2.6 任务6：逆变模型OLS回归比较

对于一元线性回归

$$\begin{matrix}Y=α+βX+u\end{matrix}  \left(8\right)$$

我们利用OLS方法可以估计得到$\hat{α}$和$\hat{β}$。

根据上述模型(见式 [式 8](#eq-simple-normal) )，A同学直接进行如下等价变换（逆变模型）：

$$\begin{matrix}X=−\frac{α}{β}+\frac{1}{β}Y−\frac{1}{β}u\end{matrix}  \left(9\right)$$

A同学通过上述等价变换，认为经典模型 [式 8](#eq-simple-normal) 就是$Y$对$X$的线性回归，而逆变模型 [式 9](#eq-simple-reverse) 则是$X$对$Y$的线性回归。因此A同学认为：如果对二者分别采取OLS回归估计，那么逆变模型的截距参数估计一定等于$−\frac{\hat{α}}{\hat{β}}$，逆变模型的斜率参数估计一定等于$\frac{1}{\hat{β}}$。

请你回答如下问题：

（1）A同学关于逆变模型中参数估计的说法是否正确？并使用老师提供的配套数据集进行OLS对照检查。（要求：分别给出两个模型的回归结果，请分别复制截图如下）。

（2）如果A同学的参数估计结论是不正确的，那么两个模型之间正确的参数关系又是如何？请你做出简要分析证明。