

本人学号_____

批改人学号_____

成绩_____

特注:

- (1) 你可以与同学共同讨论，但最后必须亲自完成作业。如发现抄袭，将以零分计。
- (2) 如果作业中写了姓名，则该作业无效。
- (3) 只接受纸质版（回归结果须打印），不接受电子版。手写时请注意书写端正。
- (4) 为了公平起见，如果逾期未交作业，将不再接受补交作业。

Problem Set 1

1. 二维随机向量 $\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ 的期望为 $E(\mathbf{X}) = \boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$ 。 $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$ 为常数矩阵。证明以下等式。

(1) $E(\mathbf{AX}) = \mathbf{A}\boldsymbol{\mu}$

(2) $\text{Var}(\mathbf{AX}) = \mathbf{A} \text{Var}(\mathbf{X}) \mathbf{A}'$ （提示：使用定义及(1)的结果）

2. 假设数据矩阵为 $\mathbf{X} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \\ 1 & 2 \\ 1 & 2 \\ 1 & 2 \end{pmatrix}$ 。

- (1) 此数据矩阵是否满列秩？
- (2) 写出数据矩阵 \mathbf{X} 的转置。
- (3) 计算矩阵 $\mathbf{X}'\mathbf{X}$ ，其逆矩阵 $(\mathbf{X}'\mathbf{X})^{-1}$ 是否存在？

3. 考虑只有常数项的回归：

$$y_i = \alpha + \varepsilon_i$$

其中，常数项 α 是唯一的解释变量。推导 α 的 OLS 估计量，并证明此回归的 R^2 等于 0。

4. 数据集 `airq.dta` 包含 1972 年美国加州 30 个大城市的如下变量: `airq`(空气质量指数, 越低越好), `vala` (公司的增加值, 千美元), `rain` (降雨量, 英寸), `coast` (是否为海岸城市), `density` (人口密度, 每平方英里), `income` (人均收入, 美元)。

- (1) 把 `airq` 对其他变量进行 OLS 回归。
- (2) 检验原假设“平均收入对空气质量没有影响”。
- (3) 检验经济变量 `density` 与 `income` 的联合显著性。
- (4) 检验环境变量 `rain` 与 `coast` 的联合显著性。
- (5) 检验所有解释变量的联合显著性。

5. 穷国能否赶上富国? 由于穷国的资本较少, 故资本的边际产出较高。因此, 一种理论认为, 穷国的经济增长速度应比富国快, 并收敛于富国, 称为“绝对收敛”(absolute convergence)。另一种观点则认为, 只有在控制其他因素(比如人力资本)的情况下, 穷国的增长速度才快于富国, 称为“条件收敛”(conditional convergence)。使用 Gallup, Sachs and Mellinger (1999) 的部分跨国数据集 `geodata_short.dta`, 检验是否存在绝对收敛或条件收敛。该数据集的被解释变量为 `gdpg6590`(1965-1990 年人均 GDP 的增长率), 而解释变量包括 `lgdp65`(1965 年人均 GDP 的对数), 以及 `syr1965`(1965 年平均受中学教育年限的对数)。

- (1) 以 5% 的显著性水平检验是否存在绝对收敛。
- (2) 以 5% 的显著性水平检验是否存在条件收敛。