

《系统工程导论》黑箱建模作业 1

利用 Matlab 或者 Python, 编程实现一元线性回归

要求:

- 1) 实现函数 `linear_regression1(data, alpha)`;
- 2) 输入为 $N \times 2$ 的矩阵 `data`, 第一列为 `Y`, 第二列为 `X`, 显著性水平 `alpha`;
- 3) 打印出回归直线方程 (有必要的可以输出重要的中间数据);
- 4) 用 F 检验法进行统计检验, 显著性水平为输入 `alpha`; 提示: Matlab 中得到 F 分布对于给定显著性水平和自由度的分位数函数是 `finv`, 请大家直查其用法; 使用 Python 的同学请自行搜索相应的函数。输出检验结果, 如果输入数据满足线性关系, 那么继续 5) 和 6), 否则结束。
- 5) 打印出置信区间, 提示 Matlab 中得到标准正态分布相应分位数的函数是 `norminv`, 请大家直查其用法; 使用 Python 的同学请自行搜索相应的函数。
- 6) 在一个 figure 中, 给出: a. 所有的数据点, b. 回归直线, c. 置信区间相应的两条边界直线。
- 7) 使用给定的数据使用上面实现的函数进行处理, 并完成作业报告, 显著性水平取 0.05。

数据说明:

- 1) 数据文件: `us_covid19_data2021.txt`, 为 2021 年 1 月份美国的新型冠状病毒疫情数据, 包含总的确诊病例、疑似病例、治愈病例和死亡比例。数据来源:
<https://github.com/BlankerL/DXY-COVID-19-Data>
- 2) 对确诊病例和治愈病例进行一元线性回归分析, 以确诊病例为 `X`, 治愈病例为 `Y`

作业要求:

- 1) 独立完成
- 2) 必须提交电子版报告+代码, 使用 Python 的同学请使用 Python3.5 或以上版本, 只可以提交 .py 文件, 不允许提交 jupyter.notebook 文件
- 3) 提交的代码文件必须有一定的注释, 尤其是程序的关键步骤或者一些比较迷幻的切片 or 数据操作