

## 系统工程导论 第四次作业

来昆 自 72 2017011607

1. 试说明：病态线性回归问题中，显著性检验是否需要？如果需要，是在自变量降维去线性之前，还是之后，还是前后都检验？给出理由。

答：需要。我认为是在自变量降维之前，即在最终得到的与原始自变量维数相同的空间中，因为本来的目的是对原空间的自变量进行线性回归，要对原始的自变量进行显著性检验，降维后显著性检验不能反映原空间的线性情况，可能有降维后线性关系存在显著性但原自变量空间中不存在线性关系显著性的情况。

2. 使用 python 实现多元线性回归

1) 结果：

设定阈值为 0.0005

```
Linear Regression Outcome: beta =  
    -9.9875  
    -0.082247  
     0.58624  
    -0.0076245  
     0.33289  
F=136.7117, F_a=5.0503, F>F_a, X and Y have Linear relation.  
Half width of Confidence Interval:1.0719  
Confidence Interval: [y_hat-1.0719,y_hat+1.0719]
```

设定阈值为 0.05

```
Linear Regression Outcome: beta =  
    -9.9875  
    -0.082247  
     0.58624  
    -0.0076245  
     0.33289  
F=136.7117, F_a=5.0503, F>F_a, X and Y have Linear relation.  
Half width of Confidence Interval:1.0719  
Confidence Interval: [y_hat-1.0719,y_hat+1.0719]
```

2) 代码介绍：

linear\_regression.m 为实现多元线性回归的函数，输入 Y, X, alpha, 输出回归参数, F 检验结果, 置信区间。

可根据设置的阈值判断是否为病态线性回归 (calc\_m.m)，若为病态，返回 bool=1 和降维后维数 m，并根据 m 进行病态线性回归 (ill\_linear\_regression.m)，若不为病态，返回 bool=0, m=n, 进行多元线性回归。

得到回归参数后，进行 F 检验(F\_test.m)，根据 F 和 F\_a 的关系来判断是否为线性关系。并计算打印置信区间。