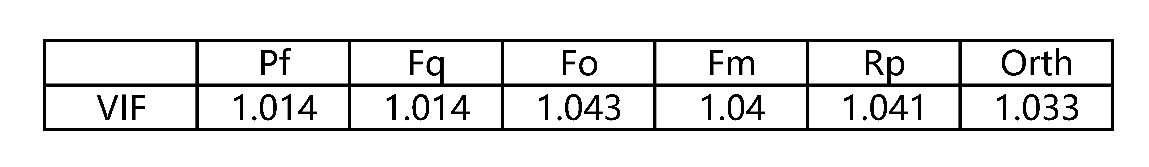
1. 由第一问对单词属性的挖掘，利用各属性作为自变量建立模型
2. 对各属性内部做统计分析

由相关系数图，发现某些属性之间存在比较强的相关性，例如Rp与Fo之间的相关系数为-0.16，Rp与Fm之间的相关系数为-0.13，但相关性并不是特别显著。

于是利用SPSS进行多重共线性检验:

1. 经检验，六个属性的VIF值均在1左右，小于3



再参考共线性诊断中的特征值与条件指标。当特征值约等于0、条件指标的值大于10、方差比例接近1时（其中一项符合即可），均可说明存在比较严重的共线性。



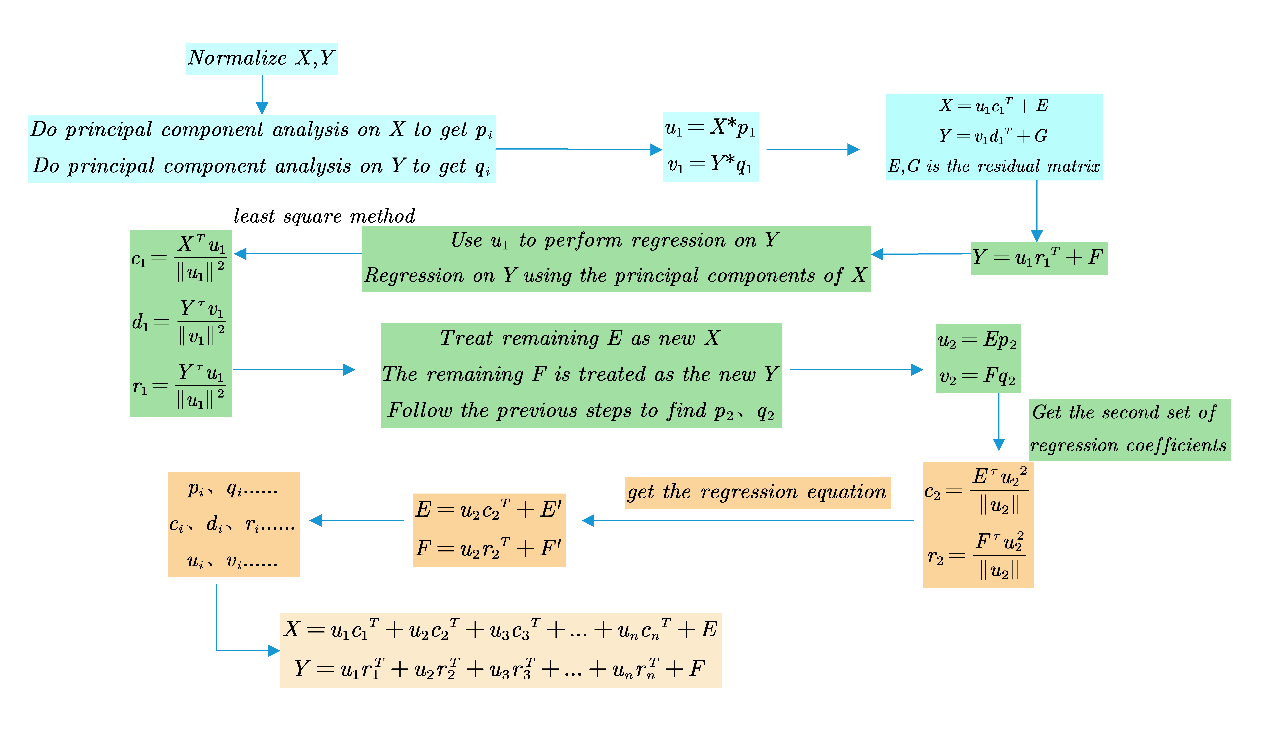
由多重共线性检验结果

1. 维度5的特征值约等于0

2. Fq维度2的方差比例、Orth维度3的方差比例、Mi、Rp维度4的方差比例均接近1

说明存在较严重的共线性。

1. 由于本题数据量较少，且自变量间存在较严重的共线性，所以采用偏最小二重回归建立模型



1. 使用SPSSPRO进行偏最小二乘法回归
2. 得到预测结果



1. 不确定因素
2. 使用MSE、RMSE、MAE、MAPE评价预测结果

