4.1 思考与讨论

**1.在对另一个产品进行线性回归建模时，选取了与本实验同样的自变量与因变量。但是在模型的结果中发现，促销幅度对产品销量的系数是个负值。为了分析原因，又绘制了相关系数矩阵图，发现促销幅度与销量的相关系数为正值。线性模型中出现这种状况的原因可能是？**

参考答案：模型出现了较严重的多重共线性。由于促销幅度与销售手段之间可能会存在较强的联系，比如，当该产品的销售手段中flyer与markdown的比例较大时，往往产品在此时的定价都会比较低，也就是促销幅度比较大。当促销幅度销售手段同时作为线性模型的解释变量时，可能会导致模型出现较为严重的多重共线性。如果销售手段对销量的影响比促销幅度要大，很可能会导致销售手段的系数偏大，促销幅度的系数偏小，甚至可能会变为负值。

**2.当上题中的线性回归模型中出现了多重共线性，应该如何来解决？**

参考答案：一般情况下，当模型出现了多重共线性时，可以考虑使用逐步回归法或主成分回归法来剔除存在多重共线性的变量。但是对于本实验中的模型来说，存在多重共线性的解释变量都是需要进行决策分析的。及我们需要知道价格每变动多少会对产品销量带来多大的影响，也需要知道哪种销售手段最能够增加产品的销量，所以我们不能剔除或合并这些解释变量。所以，当这种情况发生时，普通的线性模型可能就无法对我们的销售决策起到太大的作用，可以考虑对解释变量做对数变换或采用岭回归模型。

**3.在进行时间序列分解时需要计算移动平均。移动平均的作用主要是什么？**

参考答案：移动平均可以削弱短期的偶然因素引起的波动。当时间序列的数值由于受周期变动和随机波动的影响，起伏较大，不易显示出事件的发展趋势时，使用移动平均法可以消除这些因素的影响，显示出事件的发展方向与趋势（即趋势线），然后依趋势线分析预测序列的长期趋势。