姓名和学号：李胜志 2210180232

章节名称：套索回归

知识目标：掌握套索回归的基本原理与参数调节方法。

能力目标：能够理解并应用L1正则化技术来防止过拟合。

素质目标：学习基础知识，提高选择合适方法解决不同问题的能力；养成分析问题、事先做好准本的良好习惯

知识重点：L1正则化的原理及其在回归模型中的应用。

知识难点：正则化技术：L1正则化。

套索回归（Lasso Regression）原理

套索回归是一种线性回归技术，其原理主要通过在模型的损失函数中加入L1正则化项来实现特征选择和模型复杂度的控制。

套索回归是一种用于估计线性回归模型系数的方法，套索回归在损失函数中加入L1正则化减少部分特征来减小特征变量对预测结果的影响，从而实现特征选择。这种方法可以在保持预测准确性的同时，识别出对目标变量影响最显著的特征。  
 套索回归的目标函数可以表示为：

L1正则化项 的作用是将系数向量中小的权重变为0，从而实现特征选择和模型稀疏性。

套索回归的优缺点:

优点:

1. 特征选择能力：套索回归能够在保持预测准确性的同时，通过L1正则化项的引入，将某些不重要的特征系数压缩到零，从而实现特征选择。
2. 处理多重共线性：在存在多重共线性的情况下，套索回归能够选择出具有代表性的特征，并忽略其他冗余特征，从而减轻多重共线性对模型性能的影响。
3. 提高预测精度：通过特征选择，套索回归能够去除对目标变量影响较小的特征，使得模型更加专注于对目标变量有重要影响的特征，从而提高模型的预测精度。

缺点：

1. 对异常值敏感：由于套索回归的稀疏性质，它对数据中的异常值较为敏感。在存在大量异常值的情况下，套索回归可能会导致模型的不稳定，甚至产生错误的特征选择结果。
2. 正则化参数的选择：套索回归的性能受到正则化参数（λ）的显著影响。选择合适的正则化参数是一个挑战，需要通过交叉验证等方法进行调参。如果正则化参数选择不当，可能会导致模型欠拟合或过拟合。
3. 非线性数据拟合能力差：套索回归是一种线性回归技术，其对于非线性数据的拟合能力相对较弱。在数据存在明显非线性关系的情况下，套索回归可能无法捕捉到数据的真实规律。
4. 解释性受限：虽然套索回归能够通过特征选择提高模型的解释性，但在某些情况下（如特征之间高度相关时），套索回归可能会选择出难以解释的特征组合，从而限制模型的解释性。