

```

In [1]: # 导入操作系统库
import os
# 更改工作目录
os.chdir(r"D:\softwares\applied statistics\pythoncodelearning\chap1\sourcecode")
# 导入基础计算库
import numpy as np
# 导入绘图库
import matplotlib.pyplot as plt
# 导入数据集生成工具
from sklearn.datasets import make_regression
# 导入线性回归模型
from sklearn.linear_model import RidgeCV
# 导入数据集划分工具
from sklearn.model_selection import train_test_split
# 导入模型评估的工具
from sklearn.metrics import mean_squared_error, r2_score
# 导入绘图库中的字体管理包
from matplotlib import font_manager
# 实现中文字符正常显示
font = font_manager.FontProperties(fname=r"C:\Windows\Fonts\SimKai.ttf")
# 使用seaborn风格绘图
plt.style.use("seaborn-v0_8")
# 生成Hilbert矩阵作为X
X, y = make_regression(
    n_features=20, # 维度
    n_samples=300, # 样本量
    n_informative=10 # 有效显著变量的个数
)
# 划分数据集
x_train, x_test, y_train, y_test = train_test_split(
    X, y, test_size=0.3, random_state=10
)
# 设置不同的惩罚系数alpha的个数
n_alphas = 200
# alpha系数向量
alphas = np.linspace(0.01, 2, n_alphas)
# 构造岭回归模型
ridge = RidgeCV(
    alphas=alphas, # 惩罚系数
    fit_intercept=False, # 不拟合
    cv=5 # 五折交叉验证
)
# 模型拟合
ridge.fit(x_train, y_train)
# 预测
y_pred = ridge.predict(x_test)
# 测试集上的mse
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
print("在测试集上的MSE为: ", mse)
print("在测试集上的R方为: ", r2)

```

在测试集上的MSE为: 0.00012988709824135518

在测试集上的R方为: 0.9999999974707636