

```

In [4]: # 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
# 导入数据集划分工具
from sklearn.model_selection import train_test_split
# 导入线性回归模型
from sklearn.linear_model import LinearRegression
# 导入回归模型评价函数，均方误差和R方
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")
# 设置随机数种子
np.random.seed(42)
# 设置样本量和X的维度
n_samples, n_features = 200, 50
# 生成X
X = np.random.randn(n_samples, n_features)
# 设置真实系数
true_coef = 3 * np.random.randn(n_features)
# 对系数做非负稀疏限制
true_coef[true_coef < 0] = 0
# 生成y
y = np.dot(X, true_coef) + 5 * np.random.normal(
    size=(n_samples,)
)
# 划分数据集
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.5
)
# 构造非负回归模型
reg_nnls = LinearRegression(positive=True)
# 模型拟合
reg_nnls.fit(X_train, y_train)
# 预测
y_pred_nnls = reg_nnls.predict(X_test)
# R方
r2_score_nnls = r2_score(y_test, y_pred_nnls)
# MSE
mse_nnls = mean_squared_error(y_test, y_pred_nnls)
print("NNLS R2 score", r2_score_nnls)
print("NNLS MSE", mse_nnls)
# 拟合OLS模型
reg_ols = LinearRegression()
# 模型拟合
reg_ols.fit(X_train, y_train)
# 预测
y_pred_ols = reg_ols.predict(X_test)
# Rfang
r2_score_ols = r2_score(y_test, y_pred_ols)

```

```
# mse
mse_ols = mean_squared_error(y_test, y_pred_ols)
print("OLS R2 score", r2_score_ols)
print("OLS mse", mse_ols)
# 比较非负最小二乘的拟合系数和OLS的拟合系数
fig, ax = plt.subplots(figsize=(6,6))
# 两组系数的散点图
ax.scatter(reg_ols.coef_, reg_nnls.coef_, marker=".")
# 获取xy轴的范围
low_x, high_x = ax.get_xlim()
low_y, high_y = ax.get_ylim()
low = max(low_x, low_y)
high = min(high_x, high_y)
# 绘制系数的拟合回归线
ax.plot([low, high], [low, high], ls="--", c=".3", alpha=0.5)
# 绘制横纵轴标签
ax.set_xlabel(
    "OLS拟合的回归系数",
    fontweight="bold",
    fontproperties=font,
    fontsize=14
)
ax.set_ylabel(
    "NNLS拟合的回归系数",
    fontweight="bold",
    fontproperties=font,
    fontsize=14
)
plt.show()
fig.savefig("../codeimage/code2.pdf")
```

```
NNLS R2 score 0.8225220806196525
NNLS MSE 39.58532304418648
OLS R2 score 0.7436926291700342
OLS mse 57.16773167239297
```

