

一、程序总体功能

本程序的主要目标是**判断二维或多维点集是否线性可分**，利用**感知机算法**验证线性可分性，出于可视化的目的同时在一张图中绘制三个示例数据集的分布情况与分类结果标注。最终输出：

1. 控制台打印每个示例数据集的线性可分性结果（1 表示可分，-1 表示不可分）；
2. 生成一张包含三个子图的可视化图片 linear_separability.png。

二、功能模块划分与核心算法

(1) 函数 whetherLinearSeparable(X)

功能：判断输入数据集是否线性可分。

输入 X 为 $N \times (d + 1)$ 的数组，每行前 d 列为样本特征，最后一列为类别标签（取值 ± 1 ）。

主要算法：感知机线性分类算法

- 通过引入偏置项（在特征前添加常数 1 列），形成扩展特征矩阵 $X' = [1, x_1, x_2, \dots]$ 。
- 初始化权重向量 $w = 0$ 。
- 迭代更新：若样本 x_i 被误分类（即 $\text{sign}(w^T x_i) \neq y_i$ ），则执行更新：

$$w \leftarrow w + y_i x_i$$

- 若在有限次迭代（设定上限 1000）内所有样本都正确分类，则认为数据**线性可分**；否则不可分。

返回值：

- 1 \rightarrow 线性可分
- -1 \rightarrow 线性不可分

该方法基于感知机收敛定理：若样本线性可分，感知机算法必在有限步内收敛。

(2) 函数 plot_all_datasets(X1, X2, X3)

功能：将三个样本数据集的分布情况同时绘制到一张图的三个子图中。

实现思路：

- 使用 matplotlib 的 `subplots(1,3)` 创建三个子图。
- 对每个数据集：
 - 根据标签分离正类（红色圆点）与负类（蓝色叉号）；
 - 调用 `whetherLinearSeparable(X)` 判断线性可分性；
 - 在子图标题中动态显示 “Linear Separable: 1/-1”；
 - 添加坐标轴标签、网格、图例。
- 最后调用 `plt.tight_layout()` 使排版紧凑，并保存图片为 `linear_separability.png`。

此模块用于直观展示数据分布与算法判断结果。

(3) 主程序模块 `if __name__ == "__main__":`

功能：创建三个测试数据集，并调用上述函数执行判定与绘图。

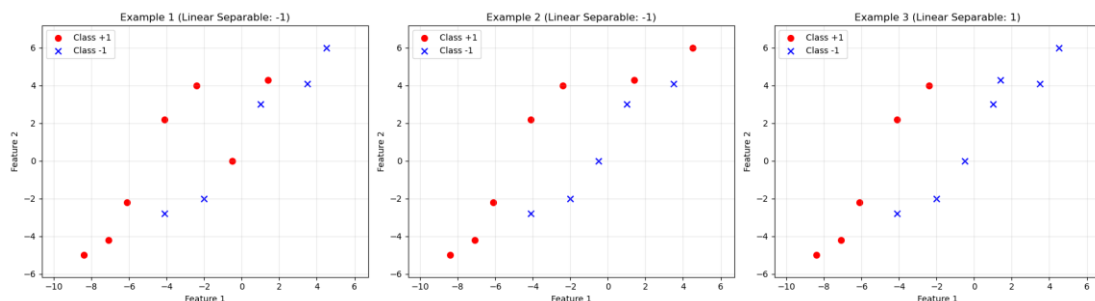
三个示例数据分别构造为：

- **示例 1 (X1)：**混合样本标签随机，无法用直线划分 → **线性不可分**。
- **示例 2 (X2)：**标签略有调整，仍不可分。
- **示例 3 (X3)：**正负类分布明显分开 → **线性可分**。

三、实验结果与性能说明

```
D:\Programs\miniconda3\python.exe E:\课程作业\机器学习\线性可分判断\whetherLinearSeparable.py
示例1结果： -1
示例2结果： -1
示例3结果： 1

进程已结束，退出代码为 0
```



该算法对于肉眼不易分辨的复杂二分类问题也可以给出正确的划分结果（示例二），并且可视化简化了初步的验证工作。