Report of HW3

王焕宇 522030910212

Accuracy

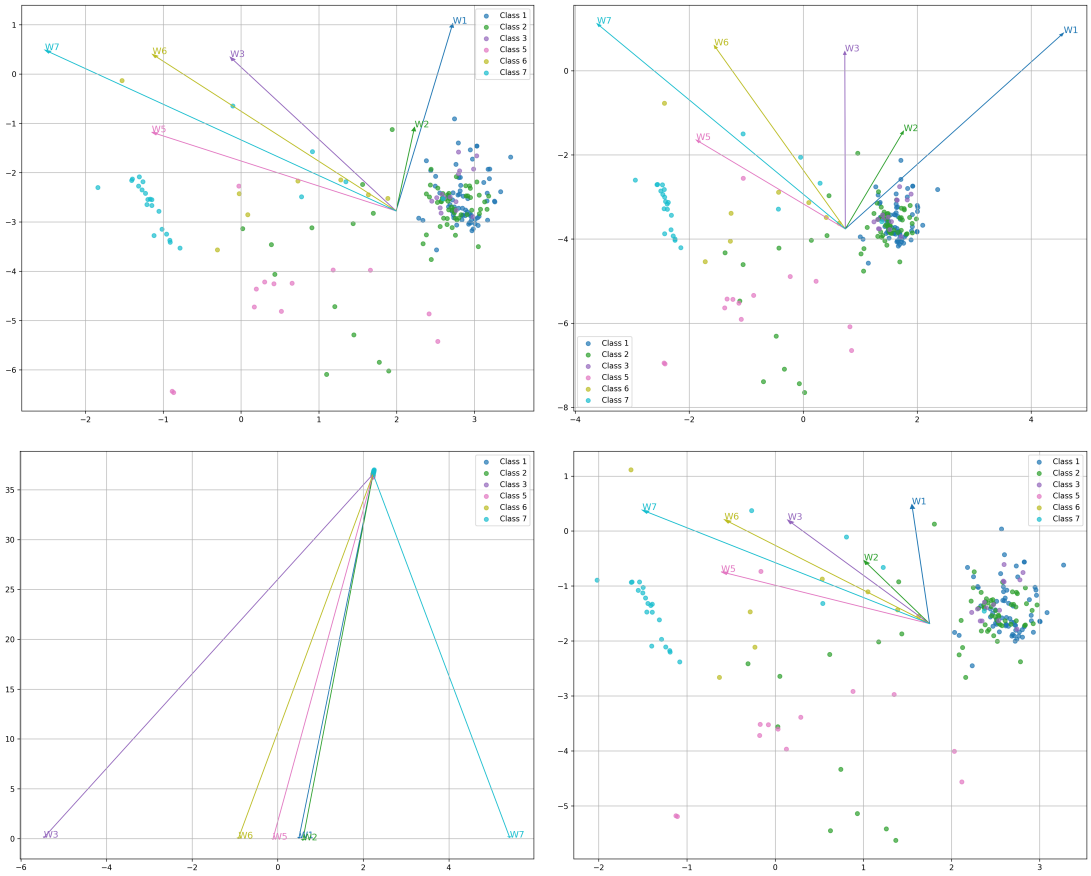
以下是最终四个方法在测试集上的预测准确率:

Method	Accuracy
Logistic Regression	54.55%
Linear Model	59.09%
LDA	50.00%
Logistic Loss	59.09%

可以看出，Linear Model 和 Logistic Regiression trained by logistic loss 取得了最好的结果，准确率为 59.09%；LDA 准确率最低，但仍有 50%。

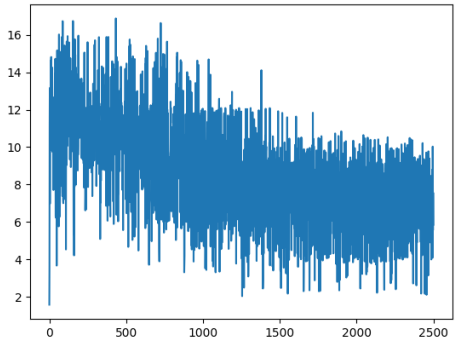
Principal Component Analysis

以下是四种方法的 PCA 结果 (从左往右，从上往下，依次为 Logistic Regression, Linear Model, LDA, Logistic Loss)



可以看出，LDA 方法得到的 PCA 效果不好，这可能是因为 LDA 生成了非常强的决策边界，导致参数矩阵相比于训练样本显得非常大，二者放在同一张图中进行可视化会很不平衡。从其余三种方法的 PCA 可视化结果来看，参数矩阵能够较好拟合训练样本的分布，但是仍然存在一些不平衡 (如参数矩阵方向较为集中)，我认为这是由于数据集本身不平衡，或训练过拟合导致的。

此外，以下是 Linear Model 的 Cross Entropy Loss 可视化，能够观察到 loss的下降:



Discussion

对比 Logistic Regression, Linear Model, LDA, Logistic Loss 四种方法，有以下方面：

1. Linear Model 直接使用一个参数矩阵 W 与样本进行矩阵向量乘法，通过 softmax 函数得到概率分布，取最高的概率作为分类结果；而其余三种方法都是通过将多分类问题分解为多个二分类问题来逐一解决。二者在问题建模上有本质区别 (我认为相比之下，多分类问题上 Linear Model 的建模更加优雅自然)
2. 损失函数角度，LDA 不涉及优化，Logistic Regression 和 Linear Model 使用交叉熵损失，而 Logistic Loss 使用对数损失，略有区别。其中 Logistic Regression 应用了两种损失函数，分别应对 0/1 标签和 +1/-1 标签的情况，但核心思想是相似的。
3. 除了 LDA，其余三种方法都使用了梯度下降方法来优化参数，这是机器学习的本质。