

Jiyang Liu 4731134

文本

描述已自动生成

**The code is in HW2\_1.ipynb.**

**Synthetic 1**

图表, 散点图

描述已自动生成

文本

描述已自动生成

**Compared with LDA:**

文本

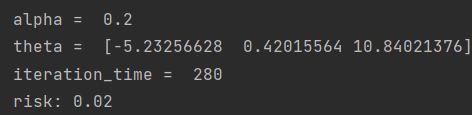
描述已自动生成

**Risks is all 0, and they have completed the classification very well.**

**Synthetic 2**

图表, 散点图

描述已自动生成



**Compared with LDA:**

文本

描述已自动生成

**The Risks are very similar, both are low, and they both complete the classification well.**

**Synthetic 3**

**图表, 散点图

描述已自动生成**

图形用户界面, 应用程序

描述已自动生成

**Compared with LDA:**

文本

描述已自动生成

**If using the original step size in data with a high degree of mixing, the risk will be very large. We reduced the step size and the risks of the final result were very close to LDA.**

**Synthetic 4**

**图表, 散点图

描述已自动生成**

图形用户界面, 文本

描述已自动生成

**Compared with LDA:**

文本

描述已自动生成

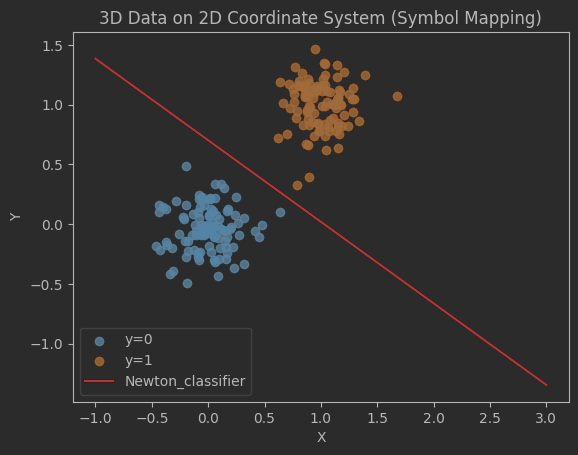
**The mixing degree of the data in Synthetic 4 is similar to that in Synthetic 3, but the data in Synthetic 4 is more concentrated. The risk of gradient descent is much smaller than the risk of LDA, indicating that the gradient descent method is more suitable in this situation.**

文本

描述已自动生成

**The code is in HW2\_2.ipynb.**

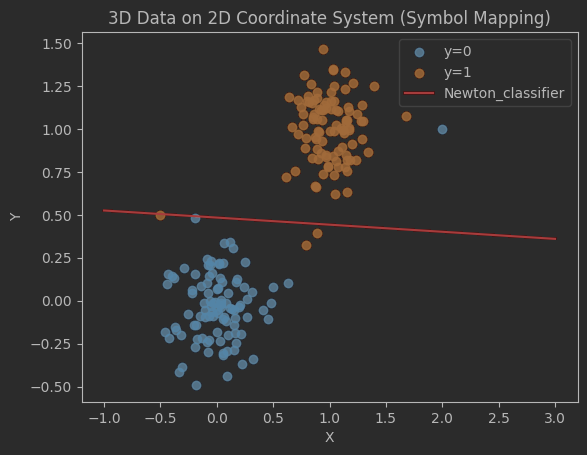
**Synthetic 1**

****

文本

描述已自动生成

**Synthetic 2**

****

图形用户界面, 文本

描述已自动生成

**Discussion:**

**From Synthetic 1 and Synthetic 2, we can see that the number of iterations is much smaller than gradient decent. This shows that the Newton method can make the gradient drop very fast, and does not require multiple experiments to determine the step size, which can save a lot of time.**

**For Synthetic 3 and Synthetic 4, I didn’t get valid result. It shows that my theta didn’t converge. I think the Newton method may not be stable for messy data, or there may be something wrong with my code.**

文本

描述已自动生成

**The code is in HW2\_3.ipynb.**

**Synthetic 1**

**图表, 散点图

描述已自动生成**

图形用户界面, 文本

中度可信度描述已自动生成

**Synthetic 2**

**图表, 散点图

描述已自动生成**

文本

中度可信度描述已自动生成

**Synthetic 3**

**图表, 散点图

描述已自动生成**

图形用户界面

描述已自动生成

**Synthetic 4**

**图表, 散点图

描述已自动生成**

图形用户界面, 文本

中度可信度描述已自动生成

**Discussion:**

**For Synthetic 1 and Synthetic 2, the stochastic gradient descent method has the largest number of iterations among the three methods. This shows that stochastic gradient descent is not a good method for data that is easy to classify.**

**For Synthetic 3 and Synthetic 4, the risk and number of iterations of stochastic gradient descent are the least. This shows that this method works well for complex and large amounts of data.**