Sunghyun Kim
5362318
kimx5261@umn.edu

Homework 1
CSci 5525 Spring 23

2023-02-02

1 Solve $\min_{\mathbf{w}} ||\mathbf{Xw} - \mathbf{y} + \mathbf{f}||_2^2 + \mathbf{a}^T\mathbf{w} + b$ to $\arg\min_{\mathbf{w}} ||\mathbf{Xw} - \mathbf{y} + \mathbf{f}||_2^2 + g$

$$
\begin{aligned}
&\frac{\partial}{\partial w} ||\mathbf{Xw} - \mathbf{y} + \mathbf{f}||_2^2 + \mathbf{a}^T\mathbf{w} + b \\
&\rightarrow \mathbf{a}^T = -2\mathbf{X}^T(\mathbf{Xw} - \mathbf{y}) \\
&\rightarrow ||\mathbf{Xw} - \mathbf{y}||^2 - 2\mathbf{X}^T(\mathbf{Xw} - \mathbf{y}) + b \\
&\rightarrow ||\mathbf{Xw} - \mathbf{y} - \mathbf{X}^T||^2 \\
&\rightarrow \frac{\partial}{\partial w} ||\mathbf{Xw} - \mathbf{y} - \mathbf{X}^T||^2 = 0 \\
&\rightarrow 2\mathbf{X}^T(\mathbf{Xw} - \mathbf{X}^T) = 0 \\
&\rightarrow \mathbf{Xw} - \mathbf{y}\mathbf{X}^T = 0 \\
&\rightarrow \mathbf{Xw} = \mathbf{y} + \mathbf{X}^T \\
&\rightarrow \mathbf{w} = \mathbf{X}^{-1}(\mathbf{y} + \mathbf{X}^T)
\end{aligned}
$$

2 Cross Validation

Please consult the my_cross_val.py code submitted.

3 Ridge Regression

For both Ridge and Lasso Regression, the most optimal λ value was 0.01.

Ridge Regression on Test Data, MSE by Fold

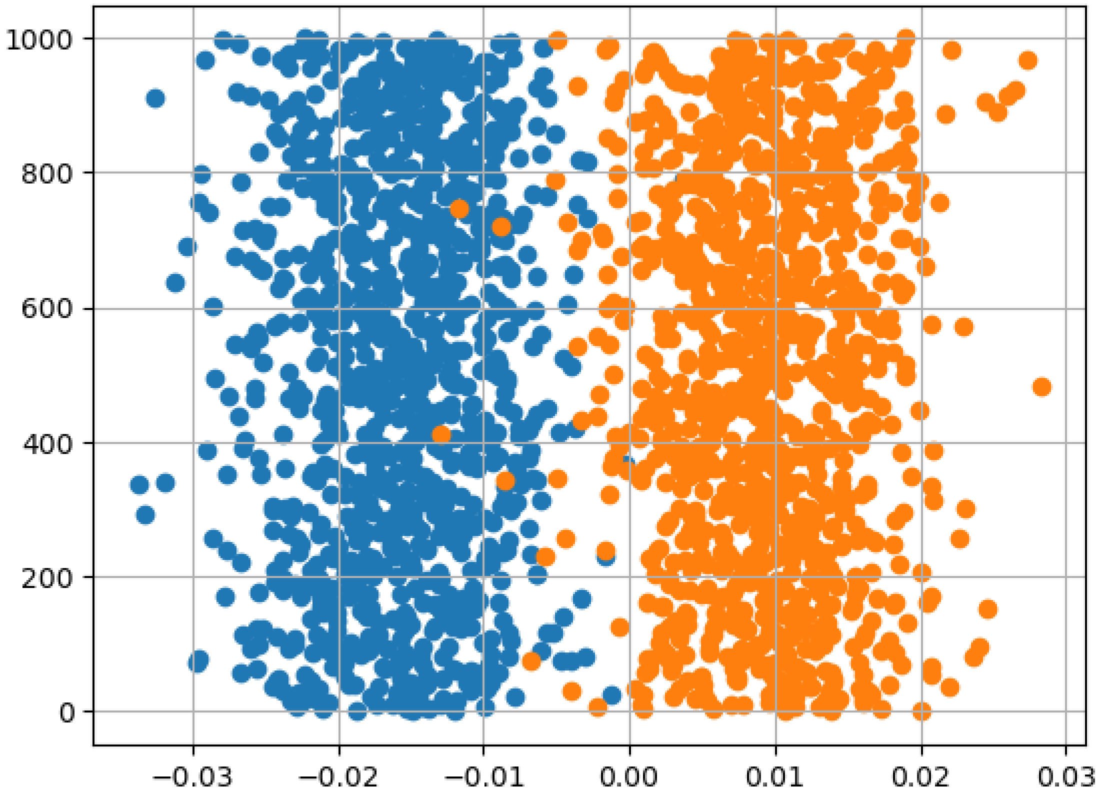
1	2	3	4	5	6	7	8	9	10	Mean	SD
0.4705	0.4338	0.4539	0.4746	0.5926	0.4915	0.4784	0.5087	0.5087	0.4923	0.4865	0.0405

Lasso Regression on Test Data, MSE by Fold

1	2	3	4	5	6	7	8	9	10	Mean	SD
0.4579	0.5012	0.5389	0.4036	0.4667	0.4788	0.6109	0.5724	0.4305	0.4422	0.4903	0.0624

Please consult the code's output for more detailed result.

4 Fischer's LDA



According to the projection, the most optimal value seems to be between [-0.01, 0].
Through cross validation, the most optimal λ value for this specific LDA model appeared to be -0.003.

Linear Discriminant Analysis on Test Data, Error Rate by Fold											
1	2	3	4	5	6	7	8	9	10	Mean	SD
0.0	0.0	0.025	0.0	0.0	0.0	0.0	0.025	0.025	0.0	0.0075	0.0115

Please consult the code's output for more detailed result.