

$$1.) \min_w \|Xw - y\|_2^2 + a^T w + b$$

$X \in \mathbb{R}^{n \times d}$ $w \in \mathbb{R}^d$ $y \in \mathbb{R}^n$ $a \in \mathbb{R}^d$ $b \in \mathbb{R}$
 col of X are lin ind $\therefore X^T X$ is invertible

$$\text{show } \|Xw - y\|_2^2 + a^T w + b = \|Xw - y + f\|_2^2 + g$$

where $f \in \mathbb{R}^n$ $g \in \mathbb{R}$

expand left side:

$$\|Xw - y\|_2^2 + a^T w + b$$

$$= (Xw)^T (Xw) - 2y^T Xw + y^T y + a^T w + b$$

$$= (Xw)^T (Xw) + (a^T - 2y^T X)w + y^T y + b$$

expand right:

$$\|Xw - y + f\|_2^2 + g$$

$$= (Xw)^T (Xw) - 2y^T Xw + 2f^T Xw - 2y^T f + y^T y + f^T f + g$$

$$= (Xw)^T (Xw) + (2f^T X - 2y^T X)w - 2y^T f + y^T y + f^T f + g$$

set two sides equal, find f and g

$$\begin{aligned} & \underbrace{(Xw)^T(Xw)} + \underbrace{(2f^T X - 2y^T X)w} + \underbrace{(-2y^T f + y^T y + f^T f + g)} \\ = & \underbrace{(Xw)^T(Xw)} + \underbrace{(a^T - 2y^T X)w} + \underbrace{(y^T y + b)} \end{aligned}$$

$$2f^T X - 2y^T X = a^T - 2y^T X$$

$$\Rightarrow 2f^T X = a^T$$

$$\Rightarrow f^T X = \frac{a^T}{2} \quad \Rightarrow X^T f = \frac{a}{2}$$

$$\text{let } \boxed{f = X(X^T X)^{-1} \frac{a}{2}}$$

$$X^T \left(X(X^T X)^{-1} \frac{a}{2} \right) = \frac{a}{2} \quad \checkmark$$

$$-2y^T f + y^T y + f^T f + g = y^T y + b$$

$$\boxed{g = b - f^T f + 2y^T f}$$

Solve $\arg\min_w \|Xw - y + f\|_2^2 + g$

$$\frac{\partial}{\partial w} \left[\|Xw - y + f\|_2^2 + g \right] = 0$$

$$2X^T(Xw - y + f) = 0$$

$$2X^T X w - 2X^T y + 2X^T f = 0$$

$$X^T X w = X^T y - X^T f$$

$$\boxed{w = (X^T X)^{-1} X^T (y - f)}$$

where $f =$

CSCI 5525 HW1

Daniel Chang

January 2023

3: Ridge and Lasso Regression

MSE for Ridge Regression												
λ	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Mean	SD
0.01	0.501	0.533	0.502	0.532	0.570	0.493	0.556	0.607	0.497	0.513	0.530	0.035
0.1	0.499	0.533	0.501	0.533	0.571	0.494	0.560	0.606	0.501	0.513	0.531	0.035
1	0.510	0.550	0.518	0.555	0.597	0.516	0.598	0.620	0.533	0.535	0.553	0.037
10	0.546	0.591	0.561	0.600	0.646	0.561	0.662	0.656	0.587	0.580	0.599	0.040
100	0.558	0.604	0.575	0.612	0.658	0.575	0.672	0.646	0.600	0.591	0.609	0.036

MSE for Lasso Regression												
λ	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Mean	SD
0.01	0.510	0.543	0.515	0.536	0.574	0.503	0.555	0.559	0.504	0.516	0.531	0.024
0.1	0.574	0.614	0.593	0.616	0.648	0.581	0.629	0.618	0.587	0.587	0.605	0.023
1	0.898	0.969	0.952	0.951	0.936	0.927	0.987	0.991	0.912	0.975	0.950	0.030
10	1.260	1.349	1.332	1.321	1.276	1.315	1.400	1.392	1.295	1.387	1.333	0.046
100	1.261	1.349	1.333	1.321	1.276	1.317	1.401	1.391	1.295	1.389	1.333	0.047

Best ridge λ : 0.01

Best lasso λ : 0.01

Ridge test mse: 0.5272739878977297

Lasso test mse: 0.5345910890138215

Ridge performs the best!

4: Linear Discriminant Analysis

LDA k-fold Error Rate												
λ	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Mean	SD
-2	0.138	0.100	0.125	0.125	0.131	0.150	0.062	0.163	0.100	0.156	0.125	0.029
-1.8	0.100	0.062	0.113	0.088	0.088	0.081	0.069	0.125	0.088	0.125	0.094	0.021
-1.6	0.069	0.044	0.069	0.056	0.069	0.056	0.069	0.094	0.062	0.094	0.068	0.015
-1.4	0.038	0.031	0.038	0.050	0.056	0.025	0.056	0.056	0.044	0.062	0.046	0.012
-1.2	0.019	0.019	0.019	0.038	0.031	0.013	0.025	0.038	0.031	0.044	0.028	0.010
-1	0.006	0.013	0.006	0.031	0.019	0.006	0.013	0.025	0.025	0.031	0.018	0.010
-0.8	0.006	0.013	0.006	0.038	0.013	0.006	0.013	0.019	0.025	0.019	0.016	0.009
-0.6	0.006	0.006	0.000	0.038	0.019	0.006	0.031	0.006	0.013	0.013	0.014	0.011
-0.4	0.006	0.006	0.006	0.025	0.013	0.000	0.031	0.013	0.013	0.006	0.012	0.009
-0.2	0.013	0.006	0.006	0.038	0.013	0.006	0.038	0.019	0.019	0.013	0.017	0.011
0	0.025	0.019	0.019	0.038	0.019	0.006	0.056	0.025	0.013	0.013	0.023	0.014
0.2	0.025	0.019	0.025	0.044	0.025	0.019	0.062	0.025	0.019	0.013	0.028	0.014
0.4	0.025	0.038	0.031	0.081	0.031	0.056	0.081	0.038	0.025	0.025	0.043	0.021
0.6	0.044	0.088	0.038	0.094	0.075	0.062	0.113	0.062	0.056	0.069	0.070	0.022
0.8	0.050	0.119	0.062	0.119	0.100	0.088	0.131	0.094	0.075	0.075	0.091	0.025
1	0.081	0.175	0.106	0.163	0.119	0.125	0.150	0.144	0.106	0.106	0.128	0.028
1.2	0.125	0.200	0.150	0.200	0.150	0.138	0.188	0.156	0.125	0.119	0.155	0.029
1.4	0.144	0.231	0.188	0.238	0.169	0.175	0.238	0.213	0.163	0.131	0.189	0.037
1.6	0.181	0.275	0.213	0.269	0.225	0.194	0.288	0.256	0.213	0.150	0.226	0.043
1.8	0.256	0.300	0.269	0.288	0.263	0.225	0.356	0.306	0.256	0.200	0.272	0.042
2	0.281	0.306	0.325	0.331	0.294	0.275	0.400	0.388	0.294	0.238	0.313	0.047

Best LDA λ : -0.4

LDA test error rate: 0.005