Statistical Learning

Computer Homework

109064509 楊暐之

1.

(a)

	Coefficient	Std. Error	t-statistic	p-value
β_0 (intercept)	-0.1785	0.1824	-0.9783	0.1647
X_1	0.0853	0.0560	1.5242	0.0647
X_2	0.0045	0.0113	0.4015	0.3443
X_3	-0.0907	0.4654	-0.1949	0.4229
X_4	-0.0578	0.1118	-0.5173	0.3028
X_5	0.1411	0.2354	0.5993	0.2749
X_6	0.0330	0.0060	5.4952	< 0.0001

(b)

	Coefficient	Std. Error	t-statistic	p-value
X_1	0.0843	0.0542	1.5553	0.0609
X_6	0.0344	0.0053	6.4921	< 0.0001

To compare with the result in (c), more precisely, the coefficient of X_1, X_6 are 0.08428384 0.03444787, resp.
Note that the two predictors that yields the smallest RSS are also the two predictors with smalles p-value in (a)

(c)

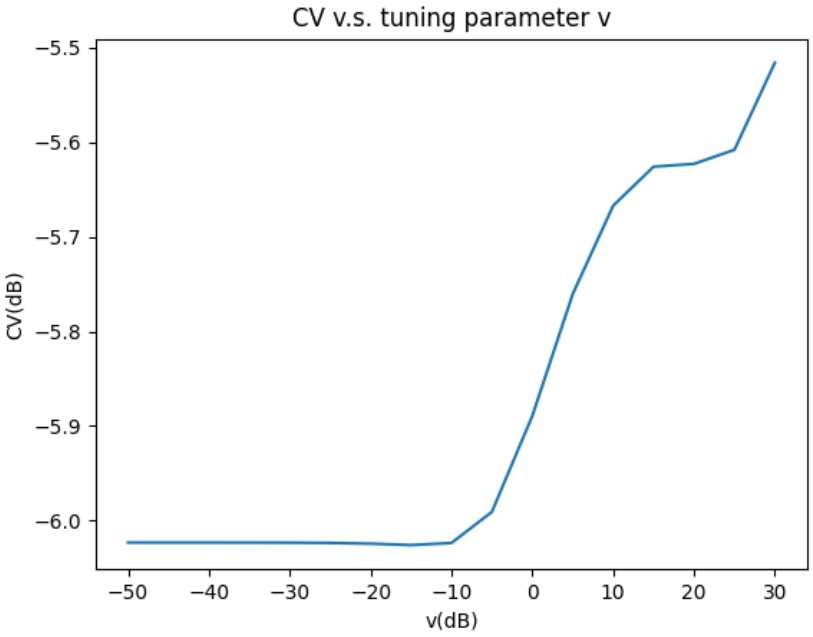
$\lambda = 10^{-1.5}$	Coefficient
X_1	0.0843
X_6	0.0344

To compare with the result in (b), more precisely, the coefficient of X_1, X_6 are 0.08428384 0.03444787, resp.

I use all predictors to caculate $\mathrm{CV}_{(5)}$ for each λ and then choose λ^* as

$$\lambda^* = \arg\min_{\lambda} \mathrm{CV}_{(5)}$$

Then, find the predictors that yield the smallest RSS with $\lambda = \lambda^*$



2.

(a)

Initial conditions :

step size : 0.1

β : $\vec{0}$

Termination condition :

$\nabla J(\beta) \leq 10000$

		True label		
		1	2	Total
Predicted label	1	618	23	641
	2	406	1001	1407
	Total	1024	1024	2048

error rate of label 1 is 39.6%

error rate of label 2 is 2.2%

average error rate is 20.9%

Change termination condition to

$\nabla J(\beta) \leq 100$

		True label		
		1	2	Total
Predicted label	1	639	35	674
	2	385	989	1374
	Total	1024	1024	2048

error rate of label 1 is 37.6%

error rate of label 2 is 3.4%

average error rate is 20.5%

The performance is only slightly improved by changing the termination condition, and it recognizes label 2 much more accurately than label 1. It is possible that label 1 is similar to label 2 , but the converse is not true for the training results in some degree.

(b)