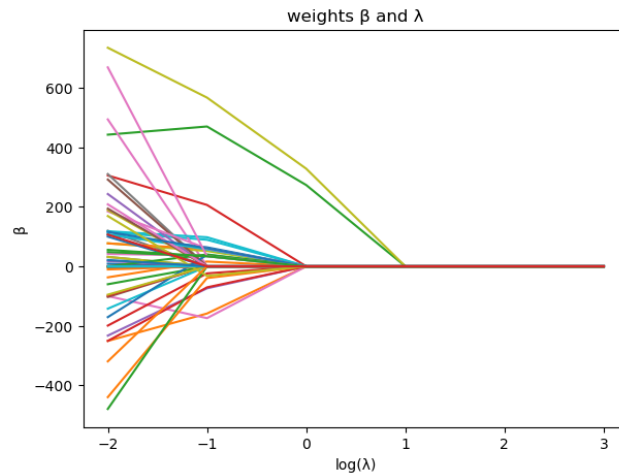


Problem 1 (LASSO and Ridge regression)

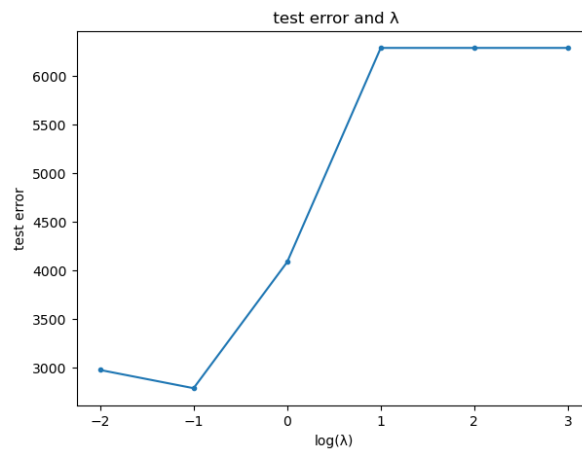
- (a) Plot the weights β at different λ . Discuss how the weights are changed by varying λ .

λ 增加時 β 收縮，當 $\lambda = \infty$ 時 $\beta = 0$ 。



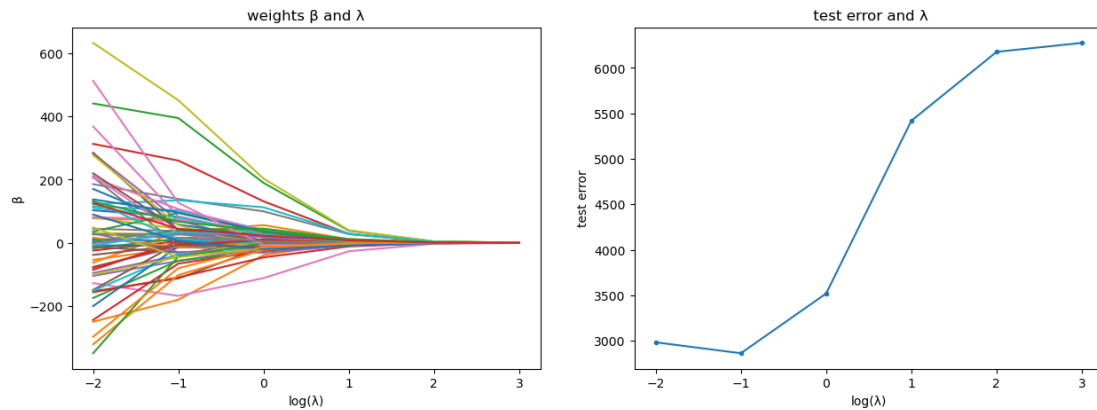
- (b) Plot test error at different λ . Discuss how the test error is changed by varying λ .

$\lambda = 0.1$ 時有最小的 test error。



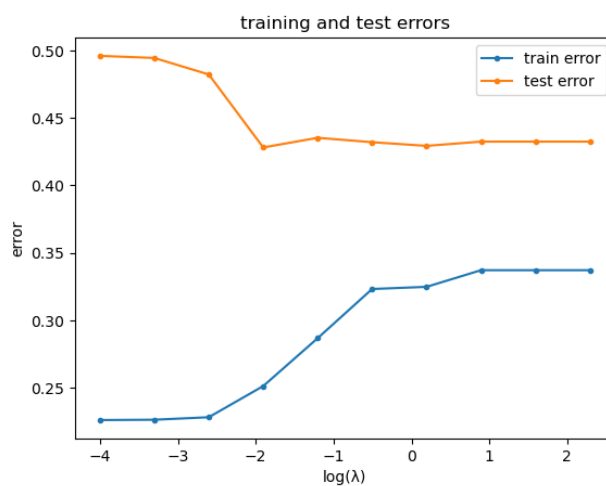
- (c) Repeat the same procedure by using the ridge regression (built-in function provided by MATLAB) with the same set of λ . Comparing the solution β obtained by the ridge regression to that by LASSO.

當 $\lambda = 1$ 時 LASSO 已幾乎收縮，而 Ridge 則在 $\lambda = 100$ 時才幾乎收縮，而 $\lambda = 0.1$ 時 LASSO 和 ridge regression 均有最小的 test error。



Problem 2 (LASSO regression)

(a) Plot the (square) training and test errors versus different λ .



(b) Briefly explain the fitting behavior (i.e., over-fitting and under-fitting) of the models with different λ , where $\lambda = \{0.0001, 0.0005, 0.0025, 0.0125, 0.0625, 0.3125, 1.5625, 7.815, 39.0625, 195.3125\}$.

1. 當 $\lambda < 0.0125$ 時，test error 隨著 λ 下降而上升，應為 **overfitting** 之情形
2. 當 $\lambda = 0.0125$ 時，test error 最小且 training error 也不會過大，為最理想之情況
3. $\lambda > 0.0125$ 時，test error 和 training error 均隨著 λ 上升而上升，應為 **underfitting** 之情形

(c) Which λ would you choose to train your final model? Why?

我會選擇 $\lambda = 0.0125$ ，因此時 test error 最小