

HW 1

Shivalika Chavan

2026-02-15

a

```
set.seed(1234)

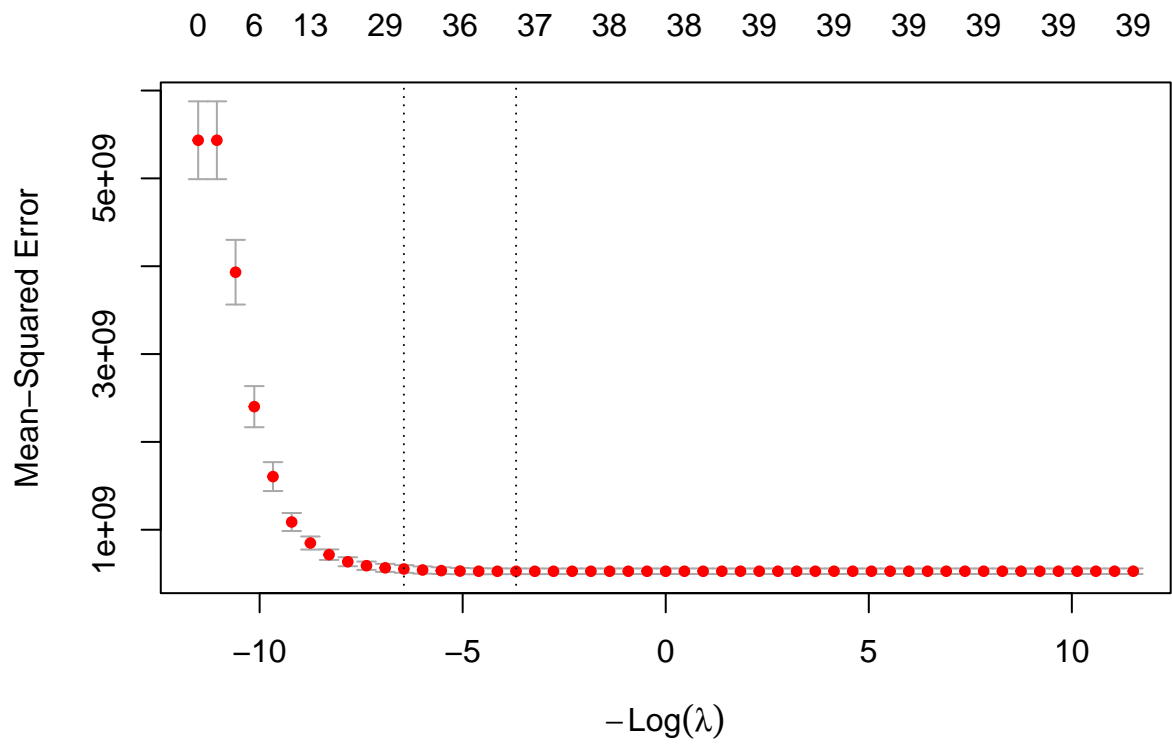
y = housing_train |> pull(sale_price)
x = model.matrix(sale_price ~., housing_train)[,-1]

lambda = 10^(seq(-5, 5, 0.2))
lasso_cv = cv.glmnet(x, y,
                     alpha = 1,
                     lambda = lambda)
lasso_fit = glmnet(x, y,
                  alpha = 1,
                  lambda = lambda)

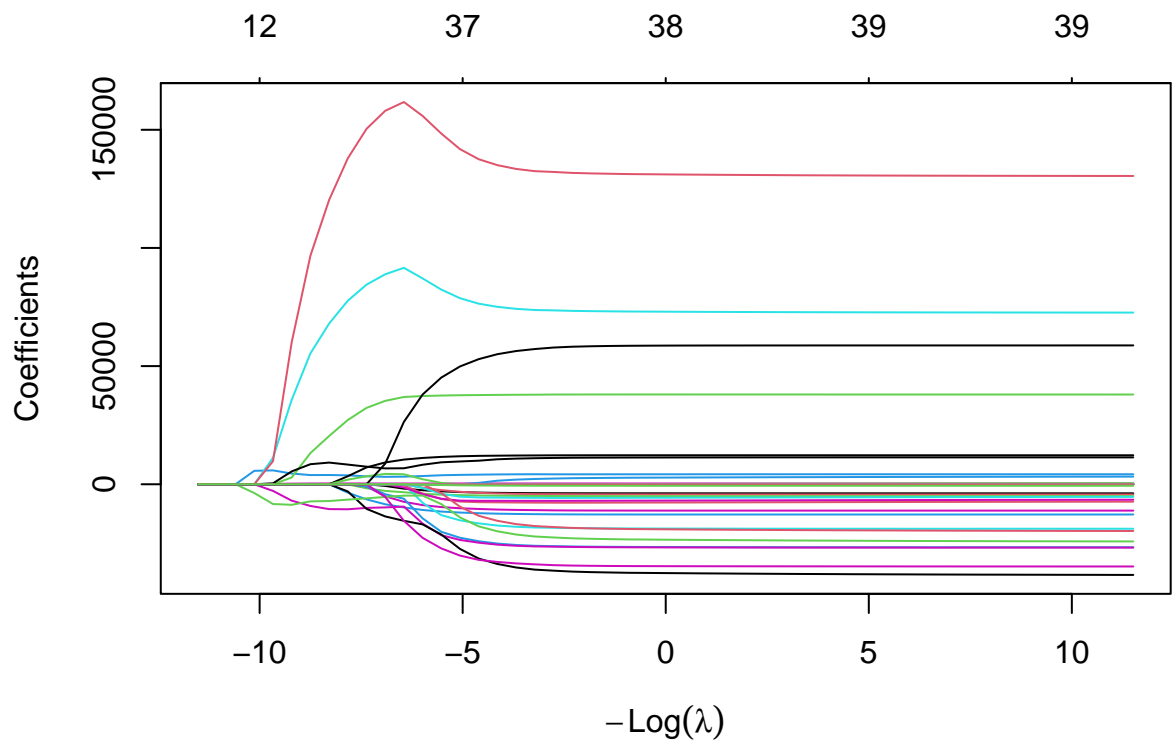
lambda_min = lasso_cv[["lambda.min"]] # minimum mean cross-validated error
lambda_1se = lasso_cv[["lambda.1se"]] # largest value of lambda such that error is within 1 standard error
CVM_min = lasso_cv |> broom::tidy() |> filter(lambda == lambda_min) |> pull(estimate)
CVM_1se = lasso_cv |> broom::tidy() |> filter(lambda == lambda_1se) |> pull(estimate)
```

The λ value with the smallest CVM (5.2673109×10^8) is 39.8107171.

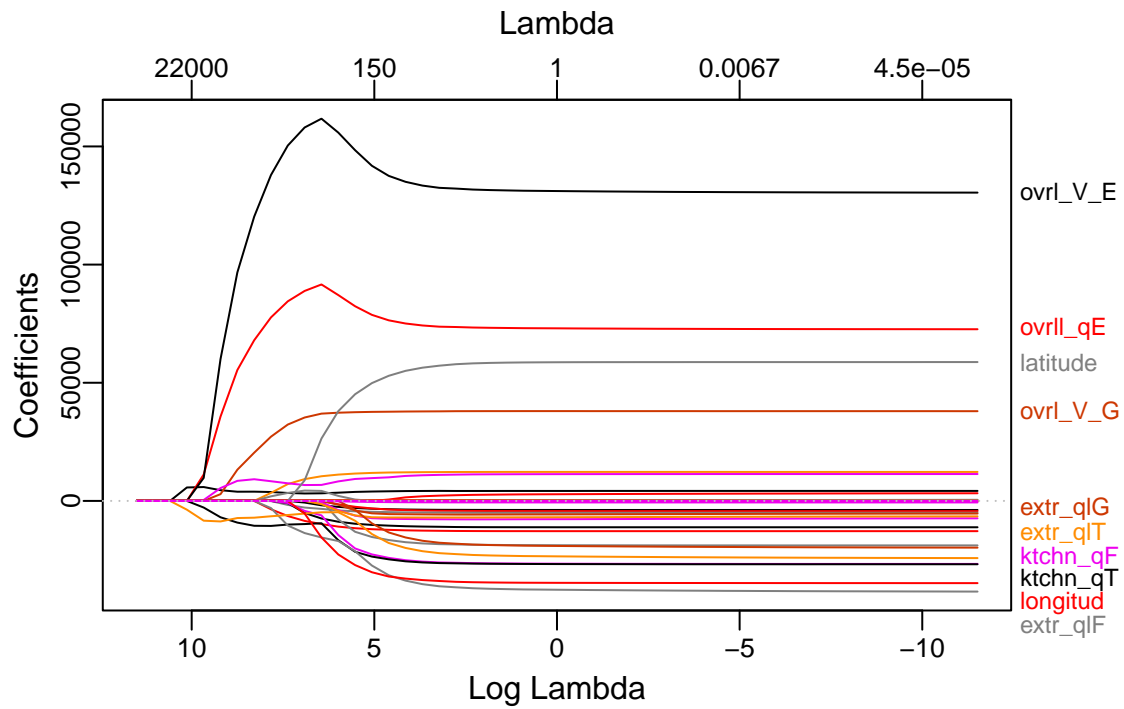
```
plot(lasso_cv)
```



```
plot(lasso_fit)
```



```
plot_glmnet(lasso_cv$glmnet.fit)
```



Test Error with $\lambda = 39.81$:

```
x_test = model.matrix(sale_price ~ ., housing_test)[-1]
y_test = housing_test$sale_price

y_pred = predict(lasso_cv, newx = x_test, s = lambda_min, type = "response")

test_error <- mean((y_test - y_pred)^2)
```

The test error is 4.4304636×10^8 .

```
coef_1se = predict(lasso_cv, type = "coefficients", s = lambda_1se)
num_predictors_1se = sum(coef_1se != 0) - 1
```

When using λ_{1SE} , there are 31 predictors.