ASM - Ridge Regression

Sergi Carol, Laura Cebollero, Alex Rodriguez

8th November, 2018

Choosing the penalization parameter

As asked, we have written two functions to choose an adequate parameter λ for the Ridge Regression.

The implementation of such functions can be found at `parameter_lambda.R` with its correspondent comments inline to understand how the code works.

We use 25 different λ that range from 0 to 100.000.

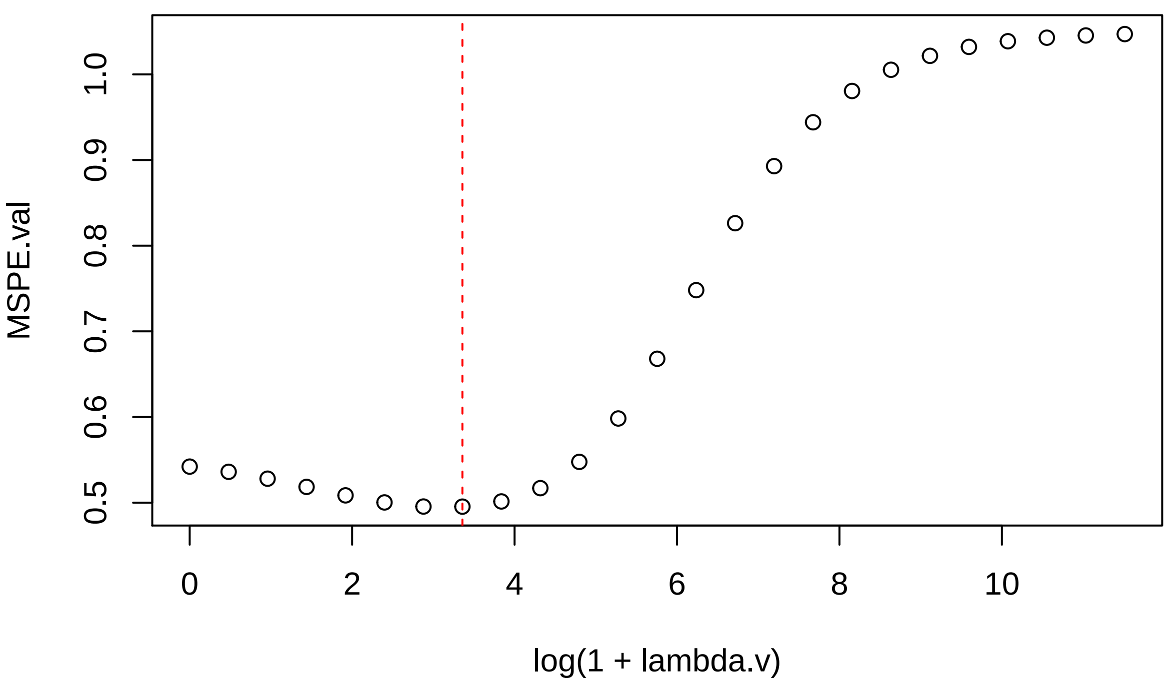
Choosing λ with MPSE and a validations set

The first function implemented is the one that allows us to choose λ using a validation set.

Such validation set has been selected by using the column `train`, which tells us which data can be used for training purposes.

This leaves us with a validation set of 30 observations, and 67 observations as the training sample.

Below is the plotting the resulting MSPE values for each lambda:




```
(minPos = which.min(MSPE.val))
```

```
## [1] 8
```

```
lambda.v[minPos]
```

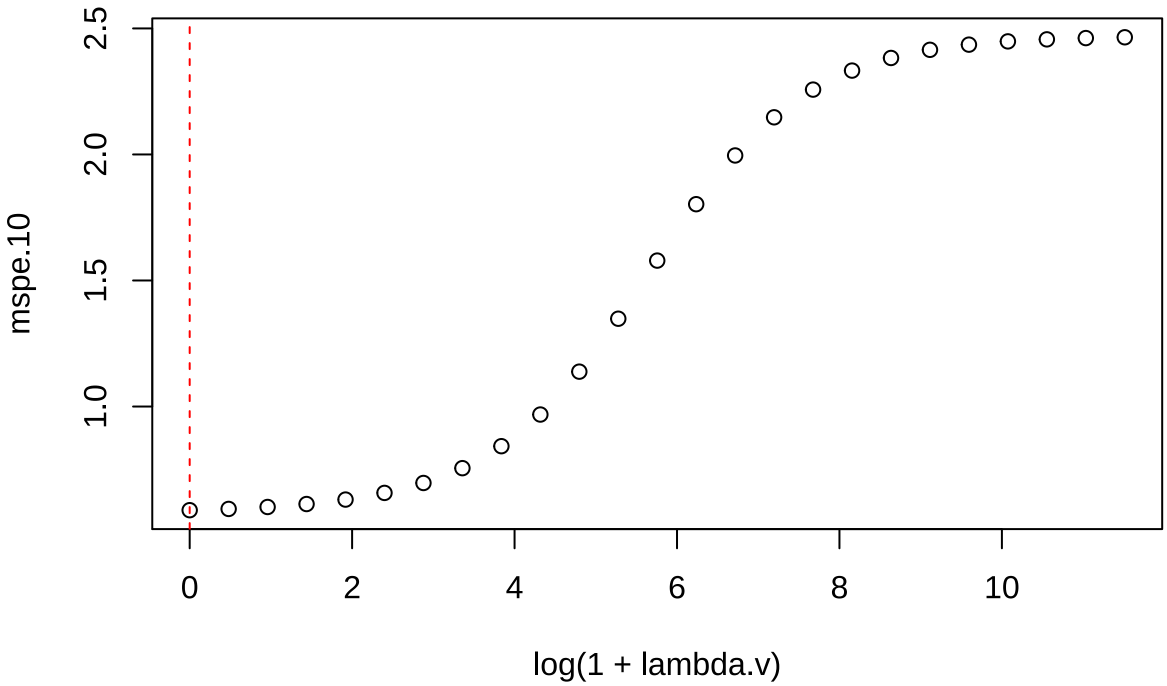
```
## [1] 27.72993
```

We can see that using this method, the λ that gives us the least error is the 8th one, with a value of 27.7.

Choosing λ with MPSE and K-fold cross-validation

We have also done the implementation of k-fold cross-validation.

Using a 10-fold CV, we obtain the following plot:




```
(minPos = which.min(mspe.10))
```

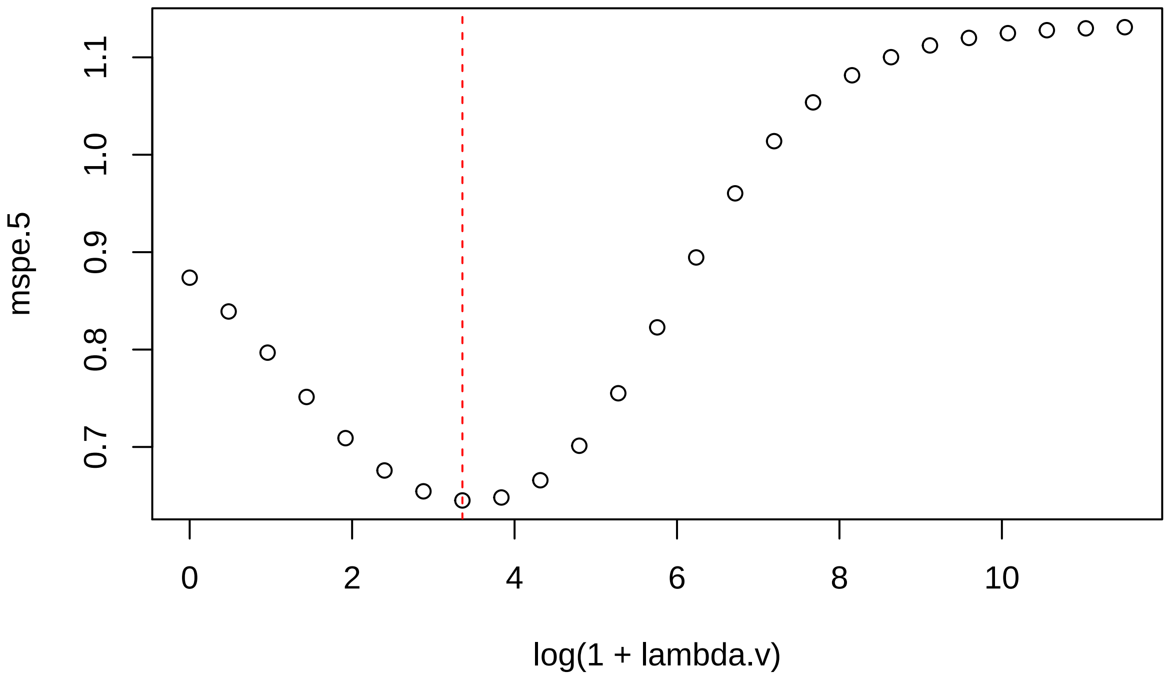
```
## [1] 1
```

```
lambda.v[minPos]
```

```
## [1] 0
```

It seems that in this case, we are selecting the one with $\lambda = 0$, which means we are not penalizing.

If we use a 5-fold CV:




```
(minPos = which.min(mspe.5))
```

```
## [1] 8
```

```
lambda.v[minPos]
```

```
## [1] 27.72993
```

Now the selected λ with the least error is the one on the 9th position with $\lambda = 45.41$.

TODO: Compare results with those obtained with leave 1 out and CV.