lasso

2022-03-30

```
install.packages("ISLR")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.1'
## (as 'lib' is unspecified)
```

```
install.packages("dplyr")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.1'
## (as 'lib' is unspecified)
```

```
install.packages("tidyr")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.1'
## (as 'lib' is unspecified)
```

```
library(ISLR)
library(glmnet)
```

```
## Loading required package: Matrix
```

```
## Loaded glmnet 4.1-4
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(tidyr)
```

```
##
## Attaching package: 'tidyr'
```

```
## The following objects are masked from 'package:Matrix':
##
##     expand, pack, unpack
```

```
rent.bed.kde.df <- read.csv("Costar R processed, All column, no size, by bed, no pop, no crime.csv")
names(rent.bed.kde.df)
```

```
##  [1] "X"                          "OID"
##  [3] "Field1"                     "Property_Address"
##  [5] "Property_Name"              "Effective_Rent_Bed"
##  [7] "Latitude"                   "Longitude"
##  [9] "City"                       "Style"
```

```
## [11] "Zoning"                       "Number_Of_Units"
## [13] "Transit_Time"                 "Drive_Time"
## [15] "Parking_Ratio"                "Bed_Count"
## [17] "Amenities_AC"                 "Amenities_Safty"
## [19] "Amenities_Pool"               "Amenities_Indoor_Gather"
## [21] "Amenities_Entertainment"      "Amenities_Outdoor_Gather"
## [23] "Amenities_Gym"                "Amenities_EV"
## [25] "Amenities_Service"            "NEAR_FID"
## [27] "NEAR_DIST"                    "Distance_Bike"
## [29] "Distance_Beach"               "overlay_flood_hazard__1__Pol"
## [31] "noise_look_up"                "kde_park"
## [33] "kde_resturant"                "kde_bus"
## [35] "kde_bike"                     "kde_gym"
## [37] "kde_student"                  "kde_store"
## [39] "Noise"                        "Flood"
```

```
library(dplyr)
#no size
rent.bed.kde.short.stand.pdf = rent.bed.kde.df %>% select (Effective_Rent_Bed, City, Style, Number_Of_U
names(rent.bed.kde.short.stand.pdf)
```

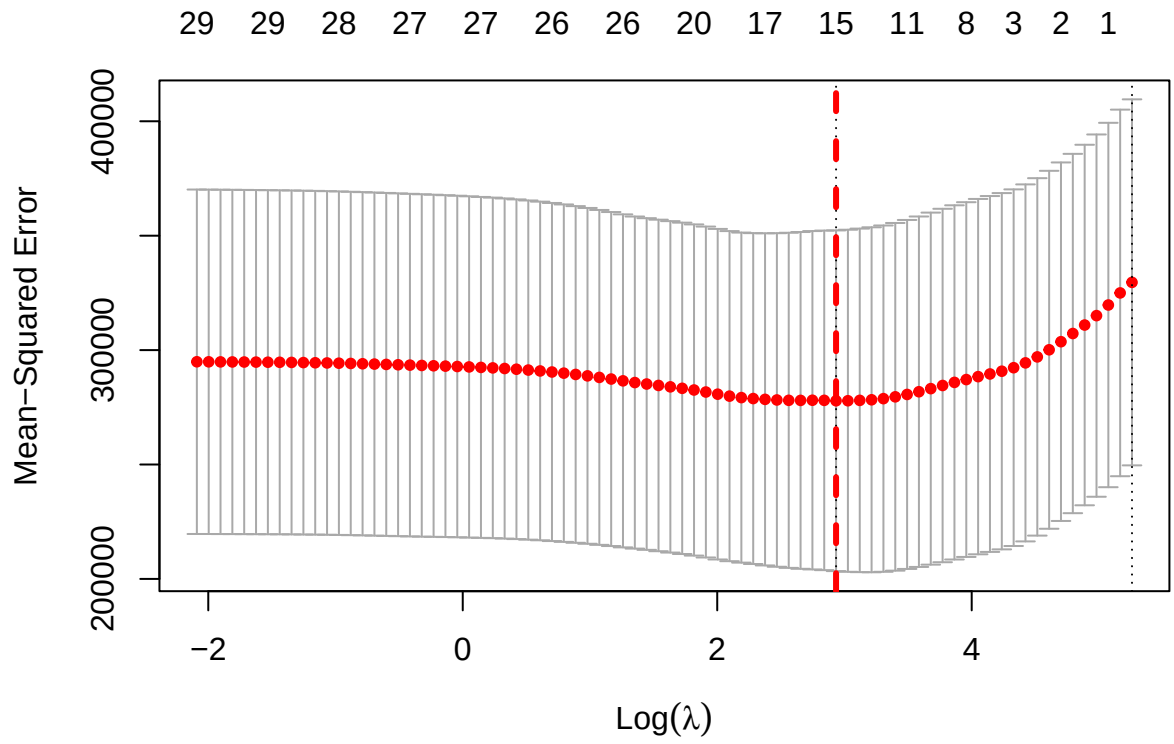
```
##  [1] "Effective_Rent_Bed"       "City"
##  [3] "Style"                    "Number_Of_Units"
##  [5] "Transit_Time"             "Drive_Time"
##  [7] "Parking_Ratio"            "Bed_Count"
##  [9] "Amenities_AC"             "Amenities_Safty"
## [11] "Amenities_Pool"           "Amenities_Indoor_Gather"
## [13] "Amenities_Entertainment"  "Amenities_Outdoor_Gather"
## [15] "Amenities_Gym"            "Amenities_EV"
## [17] "Amenities_Service"        "Distance_Beach"
## [19] "Distance_Bike"            "kde_bike"
## [21] "kde_resturant"            "kde_bus"
## [23] "kde_gym"                  "kde_store"
## [25] "Noise"                    "Flood"
```

```
x = model.matrix(rent.bed.kde.short.stand.pdf$Effective_Rent_Bed~., data = rent.bed.kde.short.stand.pdf
#y is not standardized
y = rent.bed.kde.short.stand.pdf$Effective_Rent_Bed
```

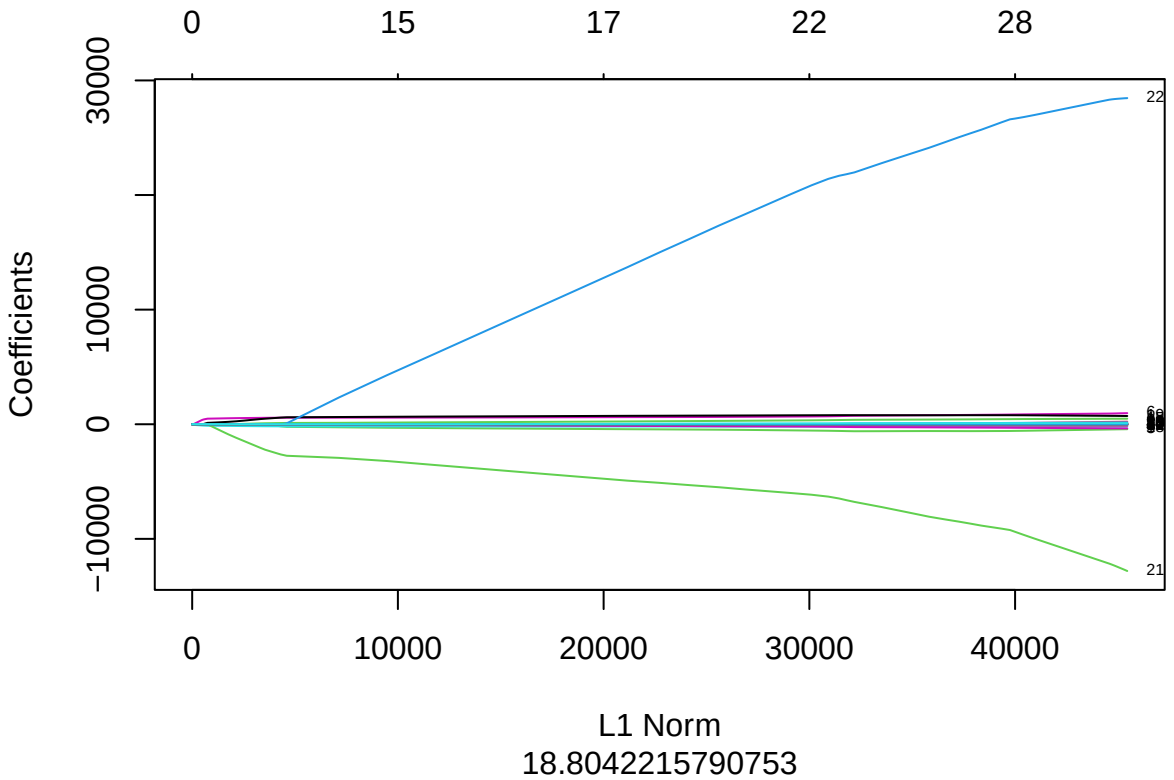
```
set.seed(1843)
cv.out.lasso=cv.glmnet(x,y,nfolds=10,alpha=1)
best.lam.lasso=cv.out.lasso$lambda.min
best.lam.lasso
```

```
## [1] 18.80422
```

```
plot(cv.out.lasso)
abline(v = log(cv.out.lasso$lambda.min), col="red", lwd=3, lty=2)
```




```
lasso.mod <- glmnet(x, y, alpha=1)
plot(lasso.mod,s = best.lam.lasso, cex = 0.2, label = TRUE)
```




```
names(rent.bed.kde.short.stand.pdf)[21]
```

```
## [1] "kde_resturant"
```

```
names(rent.bed.kde.short.stand.pdf)[22]
```

```
## [1] "kde_bus"
```

```
lasso.best <- glmnet(x, y, alpha=1,lambda=best.lam.lasso)
predict(lasso.best,type="coefficients",s=best.lam.lasso)
```

```
## 30 x 1 sparse Matrix of class "dgCMatrix"
##                                       s1
## (Intercept)                1.950698e+03
## CityIsla Vista            -6.136628e+00
## CitySanta Barbara          .
## CitySummerland            -3.019758e+02
## StyleGarden                .
## StyleLow-Rise              .
## StyleMid-Rise              5.783535e+02
## Number_Of_Units            6.223055e-01
## Transit_Time               .
## Drive_Time                 .
## Parking_Ratio              .
## Bed_Count                 -1.904652e+02
## Amenities_AC              -7.342785e+01
## Amenities_Safty            .
## Amenities_Pool             .
## Amenities_Indoor_Gather    1.452412e+02
## Amenities_Entertainment    .
## Amenities_Outdoor_Gather   .
## Amenities_Gym             -3.001012e+01
## Amenities_EV               6.508172e+02
## Amenities_Service          .
## Distance_Beach            -3.246874e+03
## Distance_Bike              4.280727e+03
## kde_bike                   .
## kde_resturant              .
## kde_bus                   -4.122995e-05
## kde_gym                   -4.110897e-06
## kde_store                 -1.319878e-01
## Noise                     -2.212608e+00
## Flood                      .
```