Class 8

BUS 696

Prof. Jonathan Hersh

BUS 696: Class 8 Announcements

1. Pset 6 solutions
2. Problem Set 7 Posted
3. Final Project
   - Tell me about your projects
   - Qs?
4. Reminder: midterm exam in two weeks
   - Qs?
5. Any other Questions?

BUS 696: Class 8 Outline

1. AI in the News
2. Lasso Regression
3. Lasso Estimation
4. Comparing Lasso and Ridge
5. ElasticNet theory
6. Estimating ElasticNet Model

AI in the News

THE WALL STREET JOURNAL.

English Edition ▾ | October 14, 2019 | Print Edition | Video

Jonathan Hersh ▾

GET NEWS ALERTS

Home World U.S. Politics Economy **Business** Tech Markets Opinion Life & Arts Real Estate WSJ. Magazine Search

JOURNAL REPORTS: TECHNOLOGY Oil and Gas Companies Turn ...

JOURNAL REPORTS: TECHNOLOGY Should Consumers Be Able to Sell ...

JOURNAL REPORTS: LEADERSHIP Don't Ask Me to Do Office Housework!

JOURNAL REPORTS: LEADERSHIP The Military Offers Women Pay ...

JOURNA LEADER The Re Don't C

BUSINESS | JOURNAL REPORTS: TECHNOLOGY

Oil and Gas Companies Turn to AI to Cut Costs

Artificial intelligence helps predict equipment glitches, manage workers and increase output

By *Neanda Salvaterra*

Oct. 13, 2019 10:01 pm ET

FROM THE EXPERTS

To Combat Climate Change, We Can't Ignore Agriculture

Why Medical Schools Need to Focus More on Nutrition

What Happens When New Supermarkets Open in Food Deserts? Not What You May Think.

What Consumers Want from Online Grocery Shopping

How to Make Sure AI Devices Pass Ethical Muster

JOURNAL REPORTS

College Rankings

College Rankings Highlights

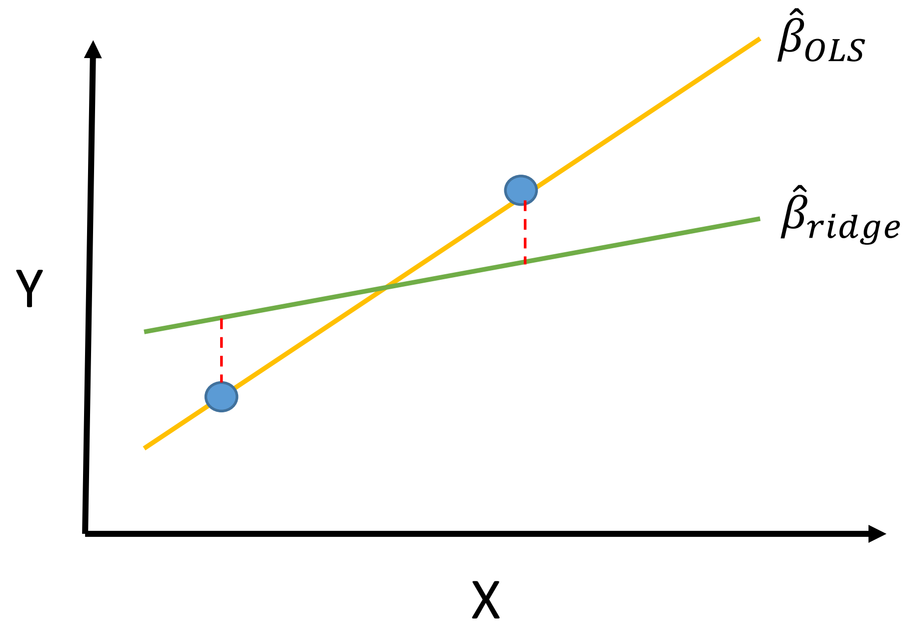
Energy

Funds/ETFs

Health Care

Ridge Regression Summary

$\hat{\beta}_{ridge}\ minimizes\colon\ residuals + \lambda \cdot (slope)^2$

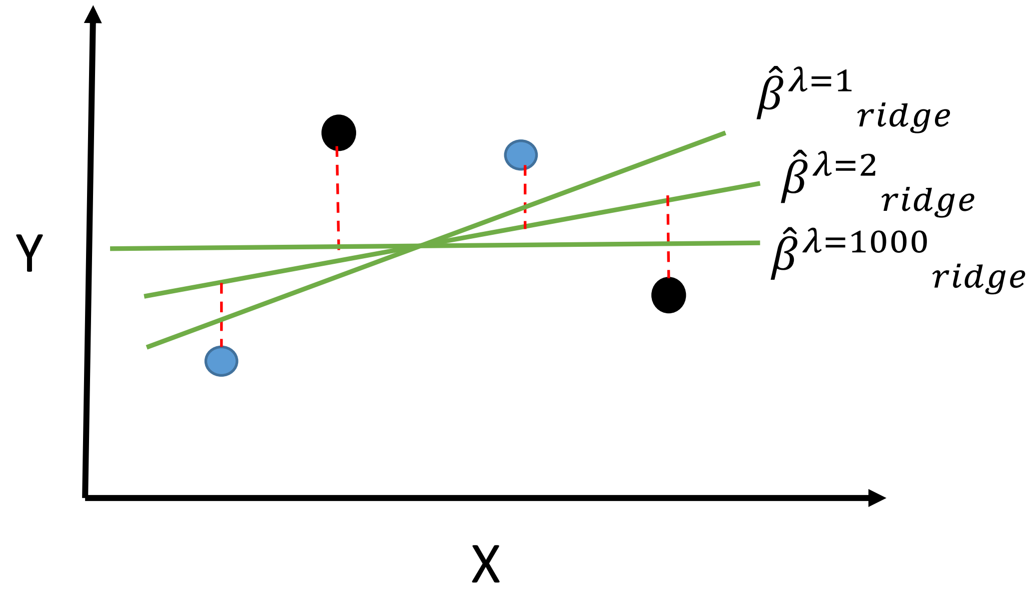


Ridge regression minimizes the OLS residuals plus the squared slope times λ

Idea: we accept a little bias (higher residuals) for less variance (better test performance)

Larger λ => More Penalization, Smaller Coefficients

$\hat{\beta}_{ridge}$ $minimizes$: $residuals + \lambda \cdot (slope)^2$



So how do we choose λ?

In practice we estimate a many models with many different values of λ

We pick a min and max lambda (say 0 and 1000), then choose some points in-between

Optimal λ^* minimizes cross-validated error

Even with very high λ ridge coefficients will never equal 0. Will always be slightly small

glmnet package

glmnet

From glmnet v2.0-18
by Trevor Hastie

99.99th
Percentile

Fit A GLM With Lasso Or Elasticnet Regularization

Fit a generalized linear model via penalized maximum likelihood. The regularization path is computed for the lasso or elasticnet penalty at a grid of values for the regularization parameter lambda. Can deal with all shapes of data, including very large sparse data matrices. Fits linear, logistic and multinomial, poisson, and Cox regression models.

Keywords models, regression

Usage

```
glmnet(x, y, family=c("gaussian","binomial","poisson","multinomial","cox","mgaussian"),
    weights, offset=NULL, alpha = 1, nlambda = 100,
    lambda.min.ratio = ifelse(nobs
```

Arguments

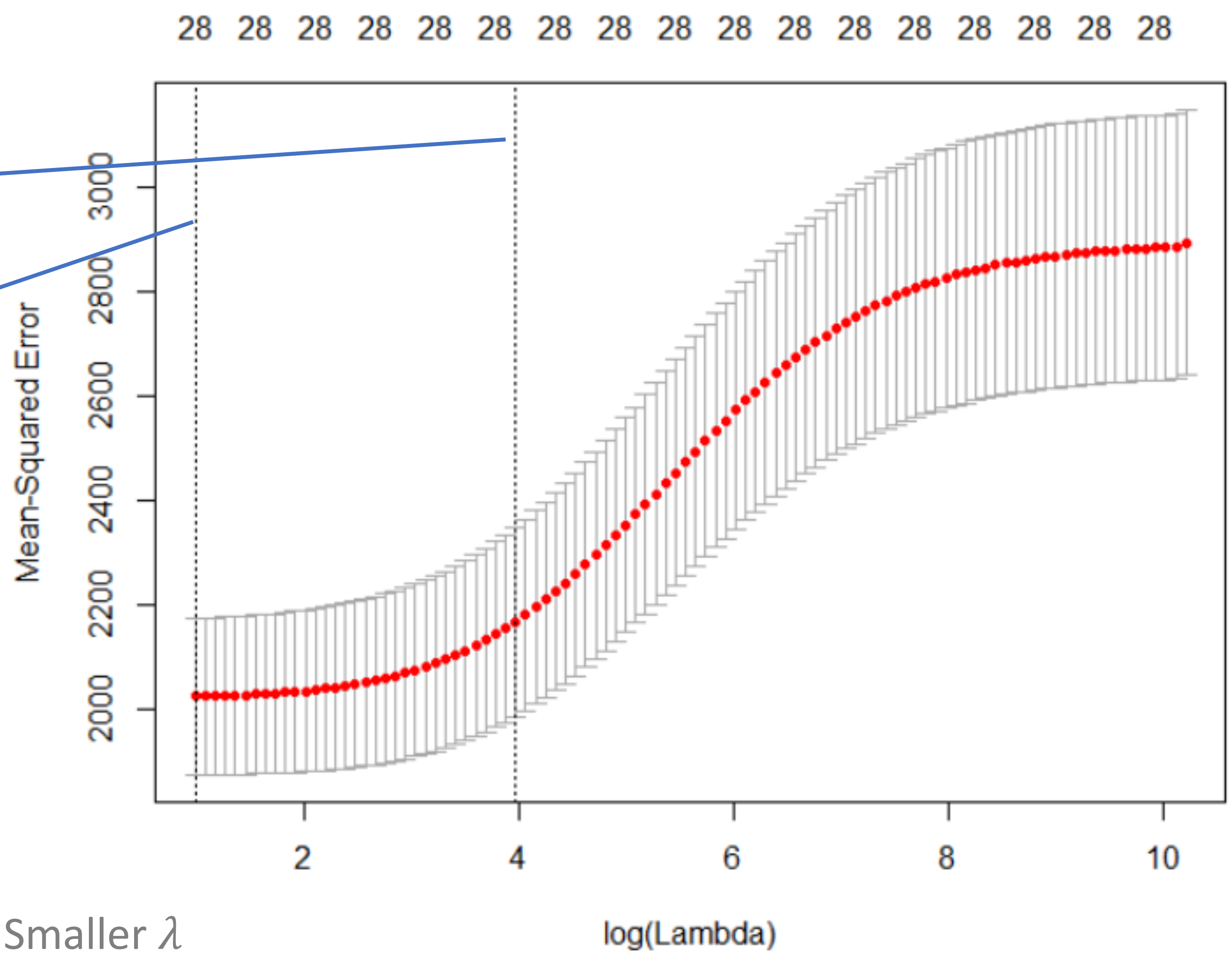
x input matrix, of dimension nobs x nvars; each row is an observation vector. Can be in sparse matrix format (inherit from class `"sparseMatrix"` as in package `Matrix` ; not yet available for `family="cox"`)

y response variable. Quantitative for `family="gaussian"` , or `family="poisson"` (non-negative counts). For `family="binomial"` should be either a factor with two levels, or a two-column matrix of counts or proportions (the second column is treated as the target class; for a factor, the last level in alphabetical order is the target class). For `family="multinomial"` , can be a `nc>=2` level factor, or a matrix with `nc` columns of counts or proportions. For either `"binomial"` or `"multinomial"` , if `y` is presented as a vector, it will be coerced into a factor. For `family="cox"` , `y` should be a two-column matrix with columns named 'time' and 'status'. The latter is a binary variable, with '1' indicating death, and '0' indicating right censored. The function `Surv()` in package survival produces such a matrix. For `family="mgaussian"` , `y` is a matrix of quantitative responses.

Cross-Validated MSE Plot As A Function of Lambda

Lambda.1se

Lambda.min

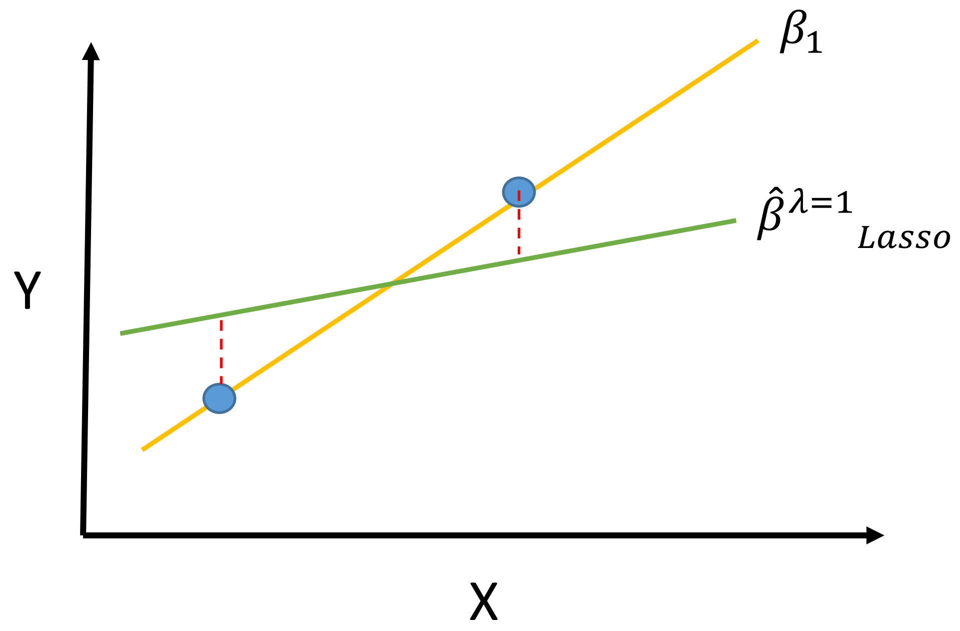


Smaller λ
Less penalization
Coefficients more
Similar to OLS

Larger λ
More penalization
Smaller coefficients

Lasso Regression Idea

$$\hat{\beta}_{Lasso}\, minimizes: \; residuals + \lambda \cdot (\, |\beta_1| + |\beta_2| + \cdots + |\beta_k|\,)$$



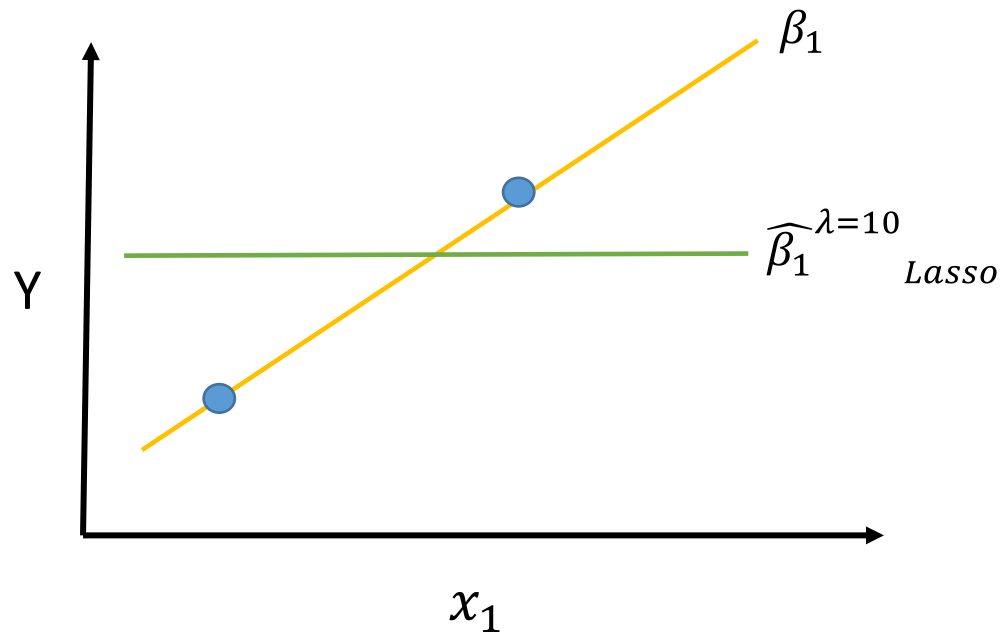
Lasso minimizes the residuals plus lambda times the absolute value of the slope coefficients

Lasso coefficients are still smaller than OLS coefficients

Lasso still accepts a little bias for (hopefully) less variance

Key Lasso Property: Variable Selection

$$\hat{\beta}_{Lasso}\, minimizes:\ residuals + \lambda \cdot (\,|\cancel{\beta_1}| + |\beta_2|)$$



For large values of λ, some slope coefficients will be chosen to be exactly zero

E.g. if we set $\lambda = 10$, maybe $\beta_1^{lasso} = 0$ but $\beta_2^{lasso}! = 0$

If that happens we effectively remove β_1 from the equation, and we have a variable selection mechanism

Estimating Lasso Model in R

```
# estimate Lasso mod
Lasso_mod <-

  cv.glmnet(profitM ~ .,
  data = movies_train %>%

    select(-c(director_name,actor_1_name,
  actor_2_name,actor_3_name,
  plot_keywords,movie_imdb_link,
  country,budgetM,grossM, genres,
  language, movie_title, budget, gross)),

  alpha = 1)
```

```
coef(Lasso_mod,
     s = Lasso_mod$lambda.min)

coef(Lasso_mod,
     s = Lasso_mod$lambda.1se)
```

The coef function outputs the coefficients for a specific lambda.

Lasso Coefficients Varying Lambda

```
> coef(Lasso_mod,
+      s = Lasso_mod$lambda.min)
29 x 1 sparse Matrix of class "dgCMatrix"
                                       1
(Intercept)                1819.78418481152
color                         6.41017608125
color Black and White       -16.35960620832
colorColor                    .
num_critic_for_reviews        0.02211767248
duration                     -0.23393200706
director_facebook_likes      -0.00116419372
actor_3_facebook_likes       -0.00904290755
actor_1_facebook_likes       -0.00755678586
num_voted_users               0.00018767446
cast_total_facebook_likes     0.00741363560
facenumber_in_poster          0.02115206284
num_user_for_reviews         -0.00146081667
content_ratingPG              6.78772369843
content_ratingPG-13           .
content_ratingR             -10.13450044626
content_ratingOther           0.31720906586
title_year                   -0.90056061740
actor_2_facebook_likes       -0.00736972247
imdb_score                    2.30207570806
aspect_ratio                 -5.95837474898
movie_facebook_likes         -0.00003254246
genre_mainAction            -13.06485079319
genre_mainAdventure          -5.64953269308
genre_mainComedy              5.42502166372
genre_mainCrime             -10.57734063797
genre_mainDrama               .
genre_mainOther               5.91146204507
cast_total_facebook_likes000s 0.01963379775
```

```
> coef(Lasso_mod,
+      s = Lasso_mod$lambda.1se)
29 x 1 sparse Matrix of class "dgCMatrix"
                                      1
(Intercept)                   658.8312414722
color                           .
color Black and White           .
colorColor                      .
num_critic_for_reviews          .
duration                        .
director_facebook_likes         .
actor_3_facebook_likes          .
actor_1_facebook_likes          .
num_voted_users                 0.0001572202
cast_total_facebook_likes       .
facenumber_in_poster            .
num_user_for_reviews            .
content_ratingPG                0.4370257357
content_ratingPG-13             .
content_ratingR                -5.3765403476
content_ratingOther             .
title_year                     -0.3270781069
actor_2_facebook_likes          .
imdb_score                      .
aspect_ratio                   -2.2827904623
movie_facebook_likes            .
genre_mainAction               -3.5449274967
genre_mainAdventure             .
genre_mainComedy                2.7792761015
genre_mainCrime                 .
genre_mainDrama                 .
genre_mainOther                 .
cast_total_facebook_likes000s   .
```

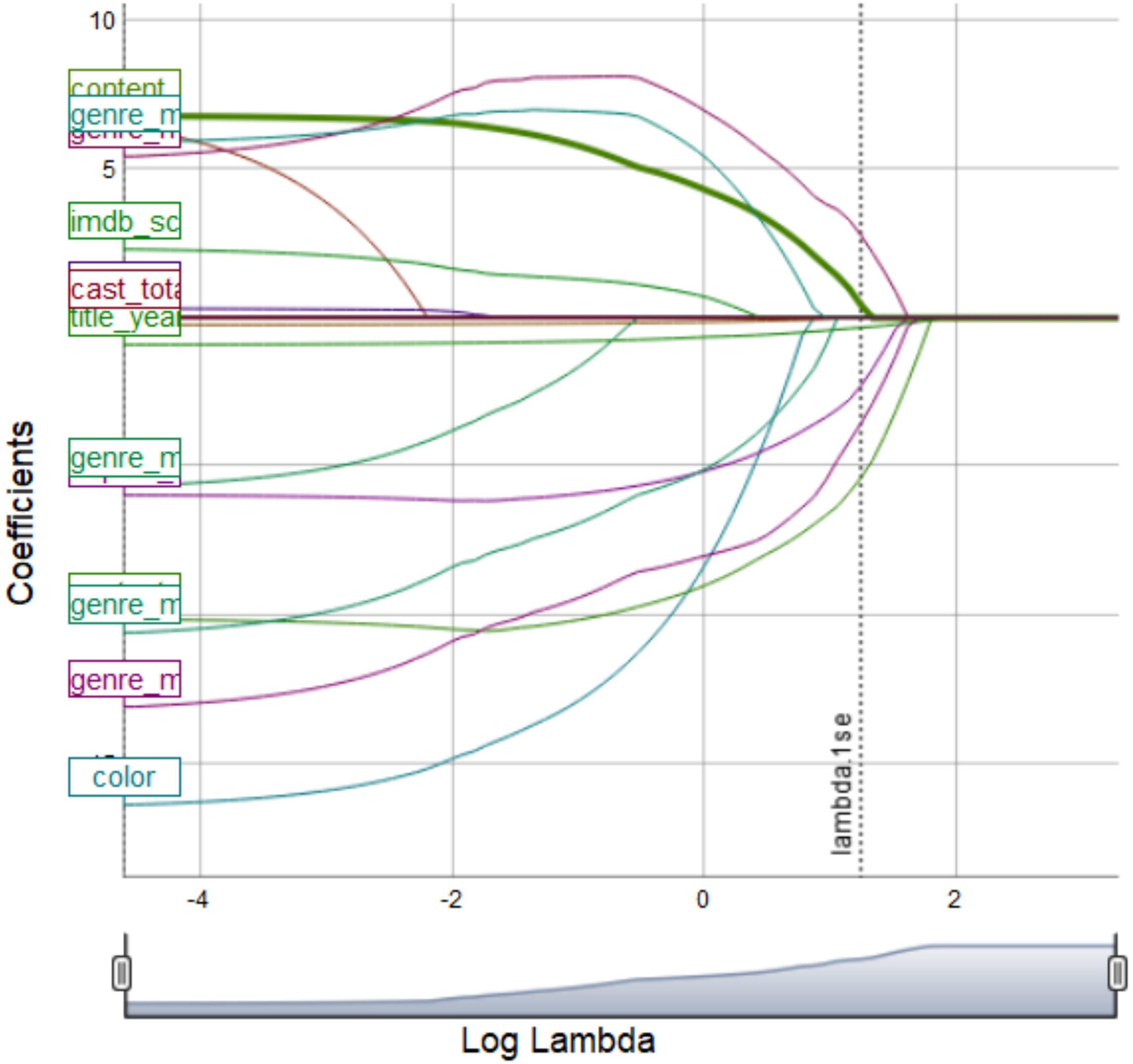
Make it a little prettier

```r
# put in a matrix
coef_mat <- data.frame(
  varname = rownames(coef(Lasso_mod)) %>%
            data.frame(),
  Lasso_min = as.matrix(coef(Lasso_mod,
                     s = Lasso_mod$lambda.min)) %>%
                round(3),
  Lasso_1se = as.matrix(coef(Lasso_mod,
                     s = Lasso_mod$lambda.1se)) %>%
                round(3)
) %>% rename(varname = 1,
             Lasso_min = 2,
             Lasso_1se = 3)  %>%
      remove_rownames()
```

```
> coef_mat
                       varname Lasso_min Lasso_1se
1                  (Intercept)  1819.784   658.831
2                        color     6.410     0.000
3          color Black and White   -16.360     0.000
4                   colorColor     0.000     0.000
5         num_critic_for_reviews     0.022     0.000
6                     duration    -0.234     0.000
7      director_facebook_likes    -0.001     0.000
8       actor_3_facebook_likes    -0.009     0.000
9       actor_1_facebook_likes    -0.008     0.000
10              num_voted_users     0.000     0.000
11    cast_total_facebook_likes     0.007     0.000
12        facenumber_in_poster     0.021     0.000
13         num_user_for_reviews    -0.001     0.000
14            content_ratingPG     6.788     0.437
15         content_ratingPG-13     0.000     0.000
16             content_ratingR   -10.135    -5.377
17         content_ratingOther     0.317     0.000
18                  title_year    -0.901    -0.327
19      actor_2_facebook_likes    -0.007     0.000
20                  imdb_score     2.302     0.000
21                aspect_ratio    -5.958    -2.283
22        movie_facebook_likes     0.000     0.000
23           genre_mainAction   -13.065    -3.545
24        genre_mainAdventure    -5.650     0.000
25           genre_mainComedy     5.425     2.779
26            genre_mainCrime   -10.577     0.000
27            genre_mainDrama     0.000     0.000
28            genre_mainOther     5.911     0.000
29 cast_total_facebook_likes000s     0.020     0.000
```

Coefpath to examine the shrinkage path

```
# explore how coefficients
# change as we change lambda
library(coefplot)
coefpath(Lasso_mod)
```



Another way to write Lasso

Lasso

$$\min_{\beta} \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 \quad \text{subject to} \quad \sum_{j=1}^{p} |\beta_j| \le s$$

Lasso with two variables

$$\min_{\beta} \sum_{i=1}^{N} (y_i - \beta_0 - \beta_1 x_{i1} - \beta_1 x_{i2})^2 \quad \text{subject to} \quad |\beta_1| + |\beta_2| \le s$$

In other words: I give you s as a budget (like setting some lambda)

You can increase your coefficients but the sum of the absolute value of them must be less than s

Another way to write Ridge

Ridge

$$\min_{\beta} \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 \quad \text{subject to} \quad \sum_{j=1}^{p} \left(\beta_j\right)^2 \leq s$$

Ridge with two variables

$$\min_{\beta} \sum_{i=1}^{N} (y_i - \beta_0 - \beta_1 x_{i1} - \beta_1 x_{i2})^2 \quad \text{subject to} \quad (\beta_1)^2 + (\beta_2)^2 \leq s$$

In other words: I give you s as a budget (like setting some lambda)

You can increase your coefficients but the sum of the absolute value of them must be less than s

Ridge Versus Lasso Penalty

Ridge penalty

$$(\beta_1)^2 + (\beta_2)^2 \leq s$$

Lasso penalty

$$|\beta_1| + |\beta_2| \leq s$$

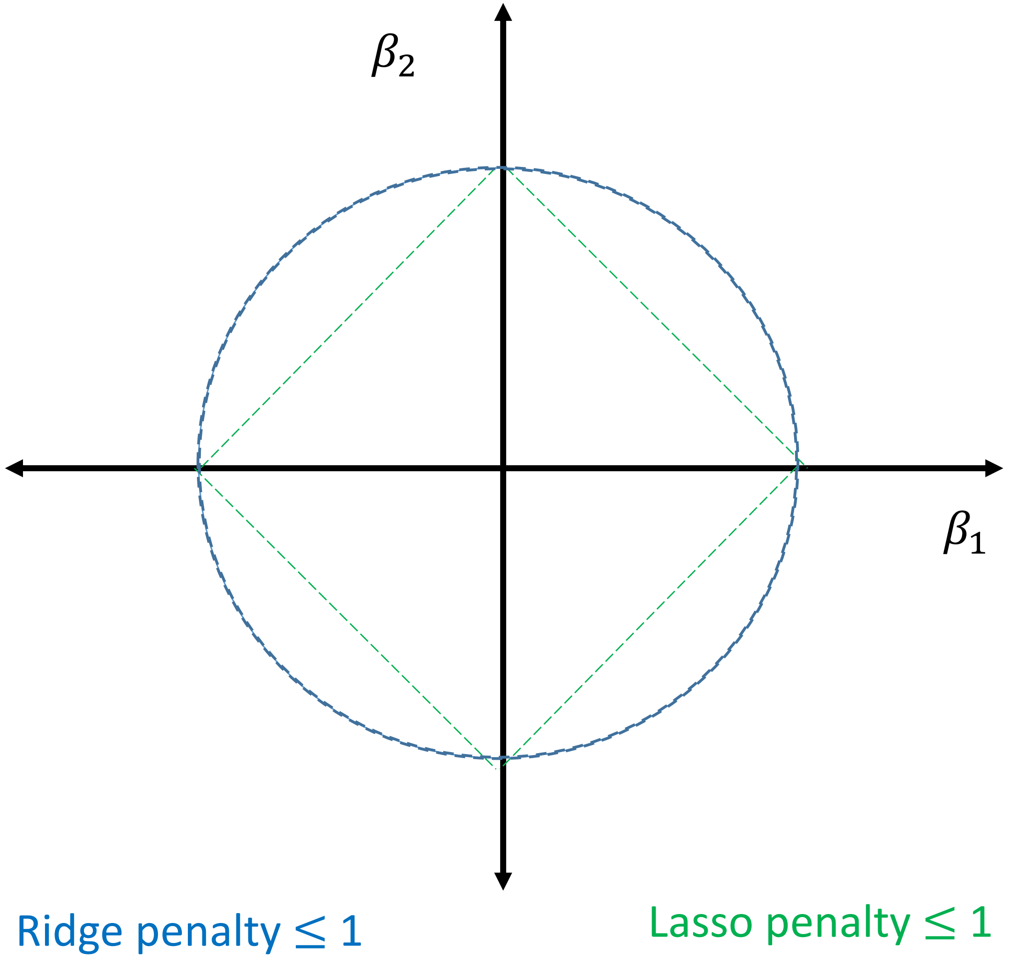
Ridge Versus Lasso Penalty

Ridge penalty $(\beta_1)^2 + (\beta_2)^2 \leq 1$

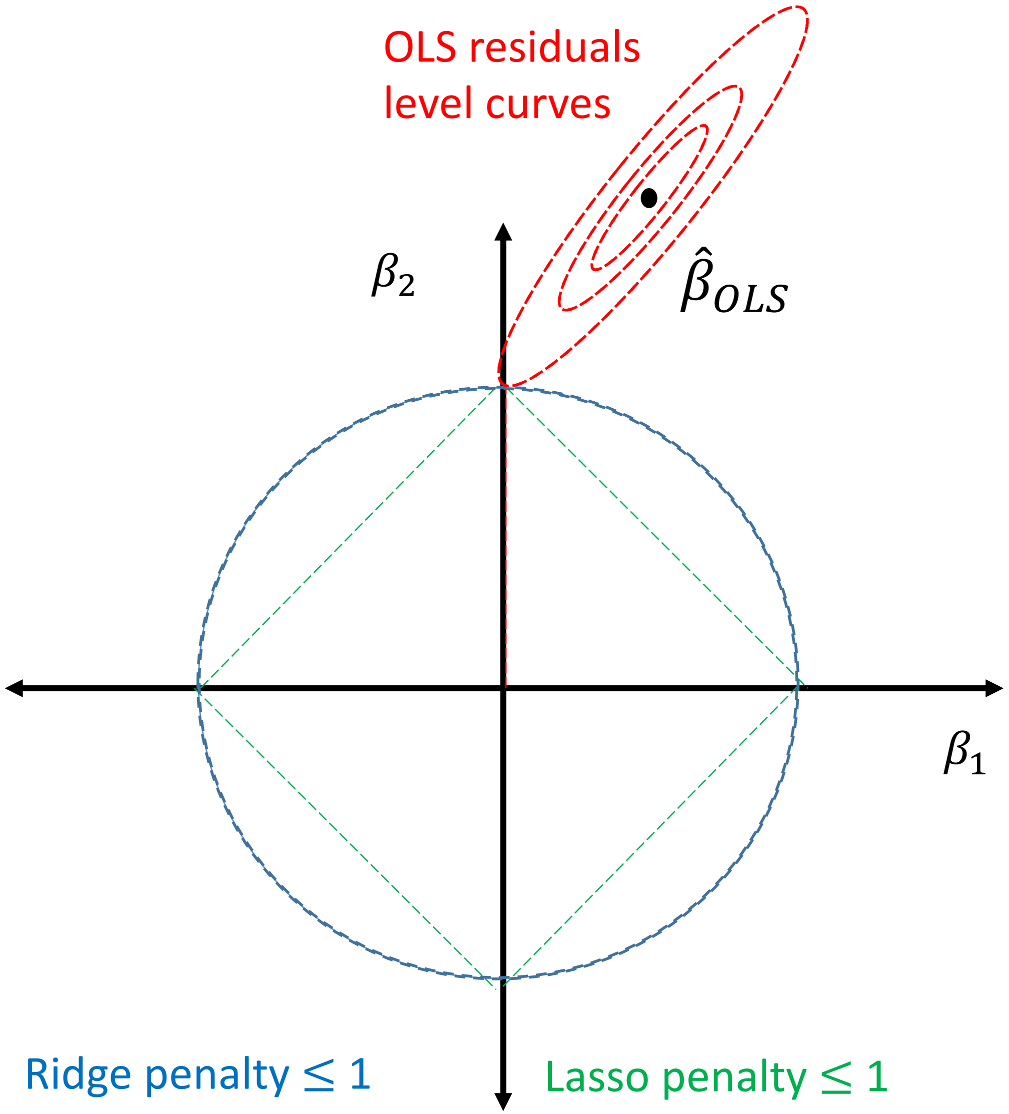
Lasso penalty $|\beta_1| + |\beta_2| \leq 1$

Let's pick an arbitrary value of s = 1

What do these look like graphically?



Ridge and Lasso Equations Redux



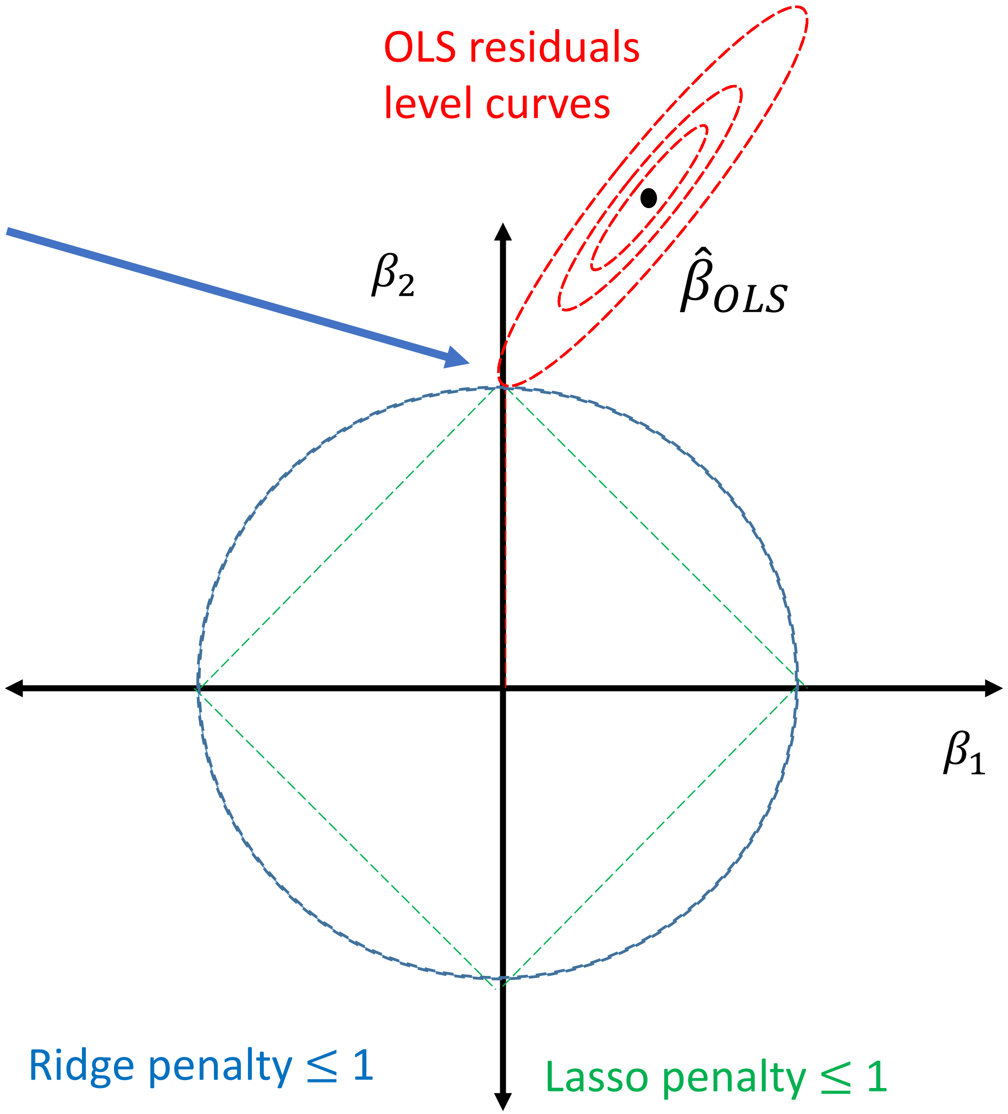
Suppose the optimal OLS beta is this point in black

Meaning, without constraints this point achieves a minimum of the residuals

We can represent that graphically as a series of contour lines where the black dot (OLS beta) is the minimum

Level curves farther from the OLS point are higher residuals

Ridge and Lasso Equations Redux



Graphically what the ridge equation is asking is: "find the lowest residual level curve while staying within the blue circle"

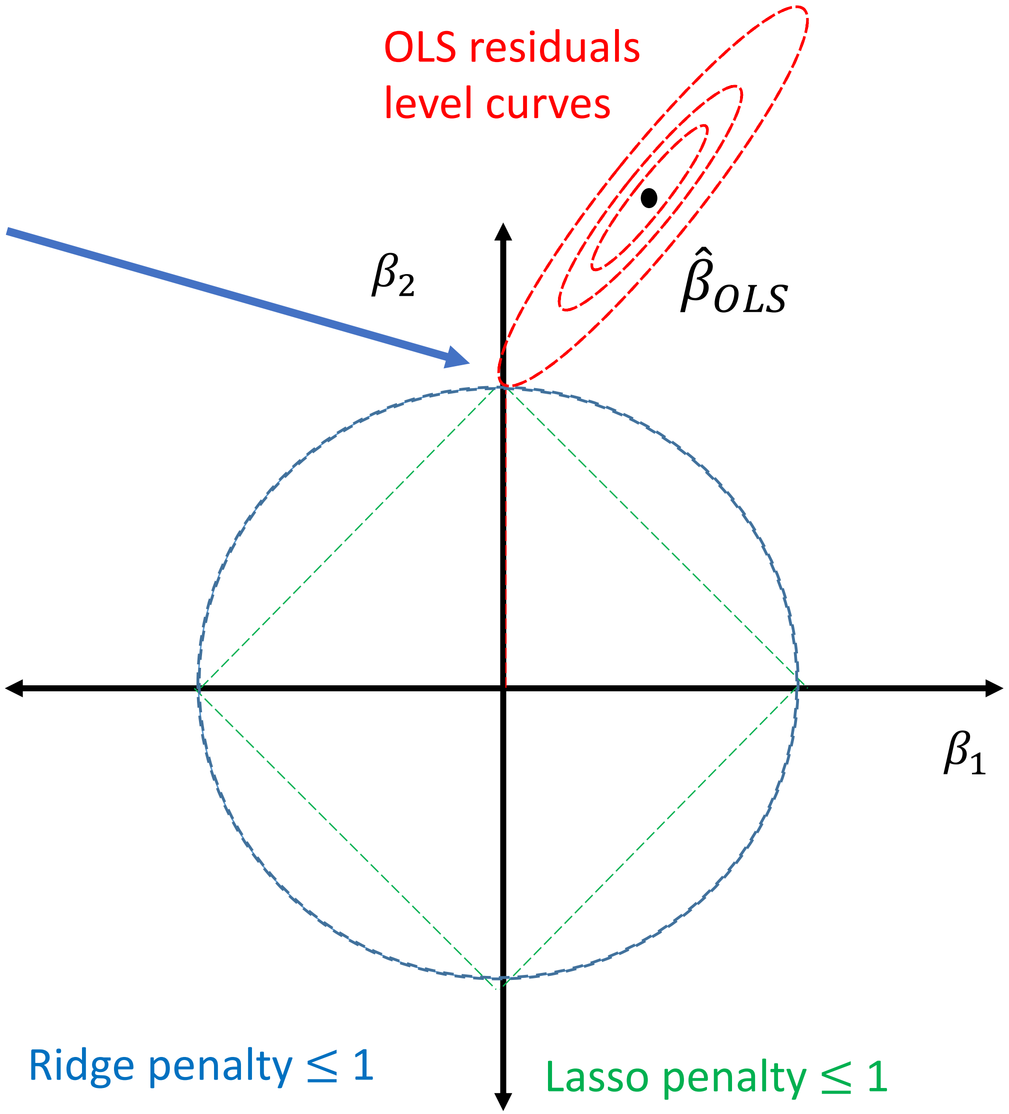
That is the level curve tangent to the blue line

Ridge

$$\min_{\beta} \sum_{i=1}^{N} (y_i - \beta_0 - \beta_1 x_{i1} - \beta_1 x_{i2})^2$$

subject to $(\beta_1)^2 + (\beta_2)^2 \leq s$

Ridge and Lasso Equations Redux



Graphically what the Lasso equation is asking is: "find the lowest residual level curve while staying within the green diamond"

That is the level curve tangent to the green line

Lasso

$$\min_{\beta} \sum_{i=1}^{N} (y_i - \beta_0 - \beta_1 x_{i1} - \beta_1 x_{i2})^2$$

subject to $$|\beta_1| + |\beta_2| \leq s$$

Ridge and Lasso Equations Redux

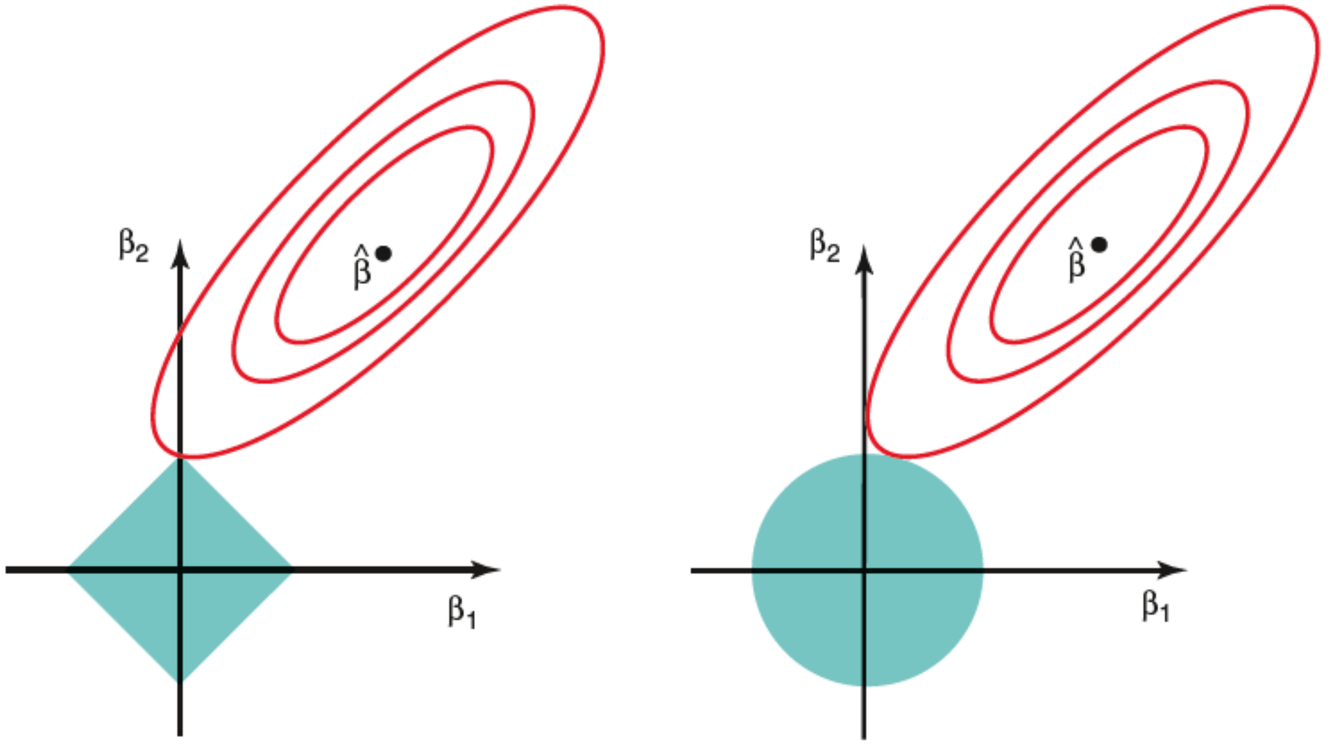


FIGURE 6.7. *Contours of the error and constraint functions for the lasso* (left) *and ridge regression* (right)*. The solid blue areas are the constraint regions, $|\beta_1| + |\beta_2| \le s$ and $\beta_1^2 + \beta_2^2 \le s$, while the red ellipses are the contours of the RSS.*

Lasso acts as a variable selector because the point of tangency for Lasso is often such that one of the variables (here beta_1) is zero

Ridge does not have this property, and we see there's still some small value for beta_1 in the right plot

Ridge versus Lasso

- Use Lasso when the "data generating process" (DGP, how the data is really formed) is **sparse**
- What is a sparse DGP?
  - Only a few variables really matter!
- Ridge should be used when many variables matter a little

BUS 696: Class 8 Outline

1. AI in the News
2. Lasso Regression
3. Lasso Estimation
4. Comparing Lasso and Ridge
5. ElasticNet theory
6. Estimating ElasticNet Model

Ridge and Lasso?

- Why do we have to choose between ridge and Lasso?
- Conceptually we could have both a squared and an absolute value penalty on the coefficients
- That model is called the ElasticNet model

ElasticNet Equation

$$\beta_{ENet} = \underset{\beta}{\operatorname{argmin}} \left\{ \underbrace{\sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{K} x_{ij}\,\beta_j \right)^2}_{OLS\ Sum\ of\ squared\ residuals} \quad \lambda \left[\alpha \left(\sum_{j=1}^{p} \left|\beta_j\right| \right) + (1-\alpha) \sum_{j=1}^{p} \left(\beta_j\right)^2 \right] \right\}$$

- Let’s set p=2 (the number of variables) to make this easier to read

ElasticNet Equation

$$\beta_{ENet} = \min_{\beta} \sum_{i=1}^{N} (y_i - \beta_0 - \beta_1 x_{i1} - \beta_1 x_{i2})^2$$

$$+ \lambda[\underbrace{\alpha(|\beta_1| + |\beta_2|)}_{\text{Lasso penalty}} + \underbrace{(1-\alpha)(\beta_1^2 + \beta_2^2)}_{\text{Ridge penalty}}]$$

- $\alpha \in [0,1]$ controls the amount of ridge versus lasso penalty
- λ functions as before -> controlling total amount of shrinkage penalty

How to choose λ and α? Grid Search

$$\beta_{ENet} = \min_{\beta} \sum_{i=1}^{N} (y_i - \beta_0 - \beta_1 x_{i1} - \beta_1 x_{i2})^2$$

$$+ \lambda[\alpha(|\beta_1| + |\beta_2|) + (1-\alpha)(\beta_1^2 + \beta_2^2)]$$

λ	$\boxed{\alpha = 0}$	$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 0.75$	$\alpha = 1$
$\boxed{0.5}$	0.1	1.1	1.1	1.1	1.1
1.0	1.1	1.1	1.1	1.1	1.1
1.5	1.1	1.1	1.1	1.1	1.1
2	1.1	1.1	1.1	1.1	1.1

- We try out a number of different combinations of hyper-parameters
- For each hyper-parameter combination we calculate cross-validate MSE
- Optimal combination has lowest cross-validated MSE

ElasticNet in R

cva.glmnet

From glmnetUtils v1.1.2
by Hong Ooi

99.99th
Percentile

Do Elastic Net Cross-Validation For Alpha And Lambda Simultaneously

Do elastic net cross-validation for alpha and lambda simultaneously

Usage

```
cva.glmnet(x, ...)

# S3 method for default
cva.glmnet(x, y, alpha = seq(0, 1, len = 11)^3,
  nfolds = 10, foldid = sample(rep(seq_len(nfolds), length = nrow(x))),
  ..., outerParallel = NULL, checkInnerParallel = TRUE)

# S3 method for formula
cva.glmnet(formula, data, ..., weights = NULL,
  offset = NULL, subset = NULL, na.action = getOption("na.action"),
  drop.unused.levels = FALSE, xlev = NULL, sparse = FALSE,
  use.model.frame = FALSE)

# S3 method for cva.glmnet
predict(object, newx, alpha, which = match(TRUE,
  abs(object$alpha - alpha) < 1e-08), ...)

# S3 method for cva.glmnet.formula
predict(object, newdata, alpha,
  which = match(TRUE, abs(object$alpha - alpha) < 1e-08),
  na.action = na.pass, ...)

# S3 method for cva.glmnet
coef(object, alpha, which = match(TRUE,
  abs(object$alpha - alpha) < 1e-08), ...)

# S3 method for cva.glmnet.formula
print(x, ...)

# S3 method for cva.glmnet
plot(x, ...)

minlossplot(x, ...)

# S3 method for cva.glmnet
minlossplot(x, ..., cv.type = c("1se", "min"))
```

ElasticNet in R

cva.glmnet {glmnetUtils} R Documentation

Do elastic net cross-validation for alpha and lambda simultaneously

Description

Do elastic net cross-validation for alpha and lambda simultaneously

Usage

```
cva.glmnet(x, ...)

## Default S3 method:
cva.glmnet(x, y, alpha = seq(0, 1, len = 11)^3,
  nfolds = 10, foldid = sample(rep(seq_len(nfolds), length = nrow(x))),
  ..., outerParallel = NULL, checkInnerParallel = TRUE)
```

ElasticNet in R

```
#-----------------------------------------------
### ElasticNet
#-----------------------------------------------
alpha_list <- seq(0,1,len = 5)
alpha_list
```

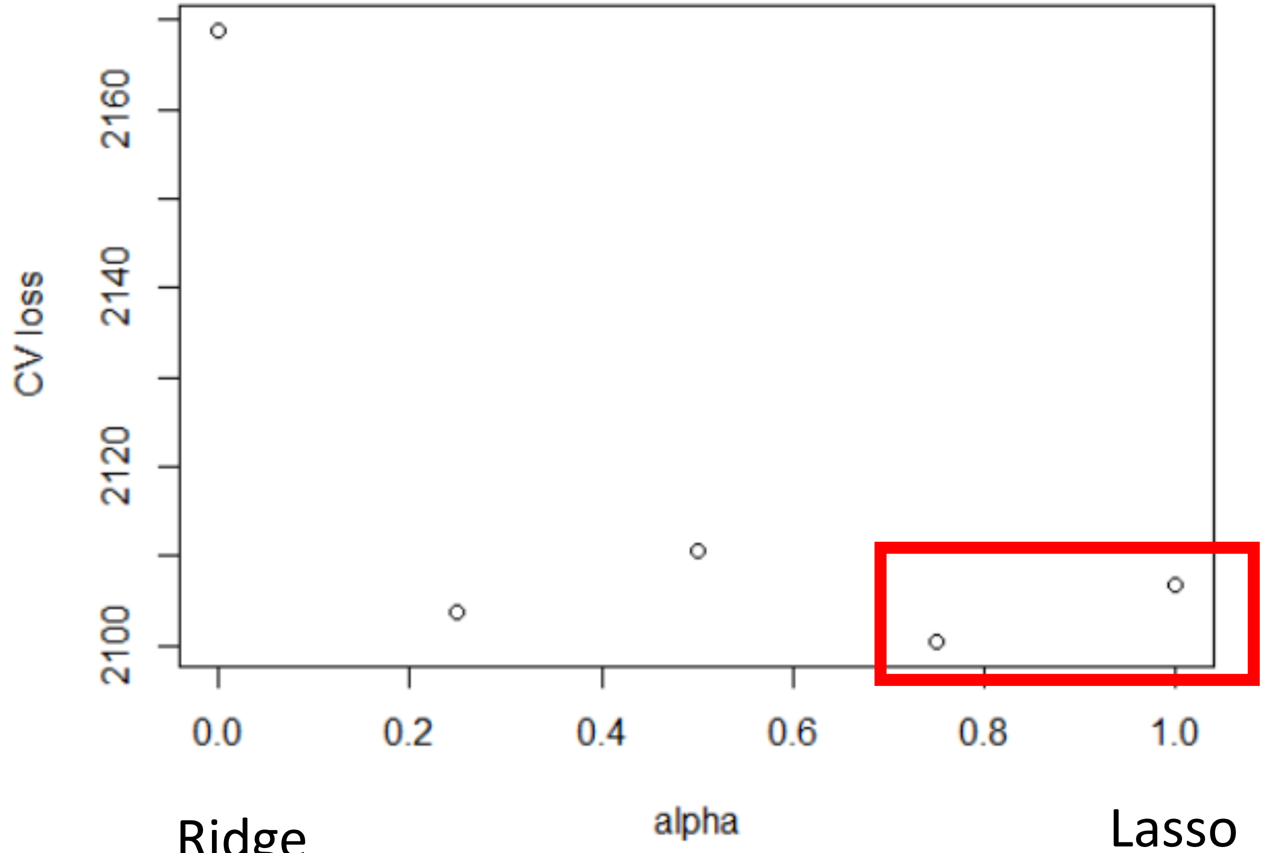
```
> alpha_list
[1] 0.00 0.25 0.50 0.75 1.00
```

```
enet_fit <- cva.glmnet(profitM ~ . ,
                       data = movies_train,
                       alpha = alpha_list)
```

```
> enet_fit
Call:
cva.glmnet.formula(formula = profitM ~ ., data = movies_train,
    alpha = alpha_list)

Model fitting options:
    Sparse model matrix: FALSE
    Use model.frame: FALSE
    Alpha values: 0 0.25 0.5 0.75 1
    Number of crossvalidation folds for lambda: 10
```

ninlossplot(enet_fit) to view CV error against alpha



Ridge

Lasso

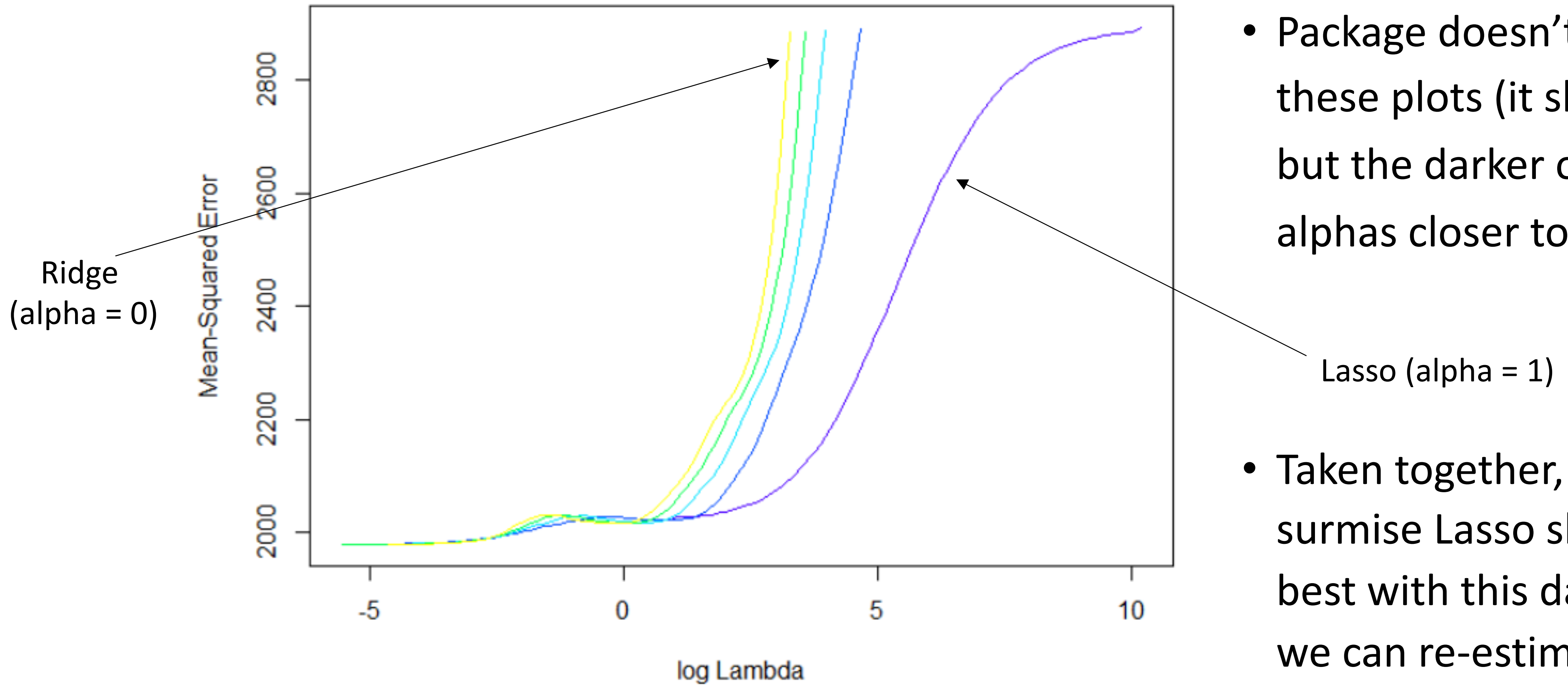
More Ridge penalty

More Lasso penalty

- The package doesn't automatically choose alpha for us, because this is a noisy statistical process -> hard to optimize

- Consider multiple alphas if they are close together due to noisy error estimation

plot(enet_fit) to view CV-MSE across alphas



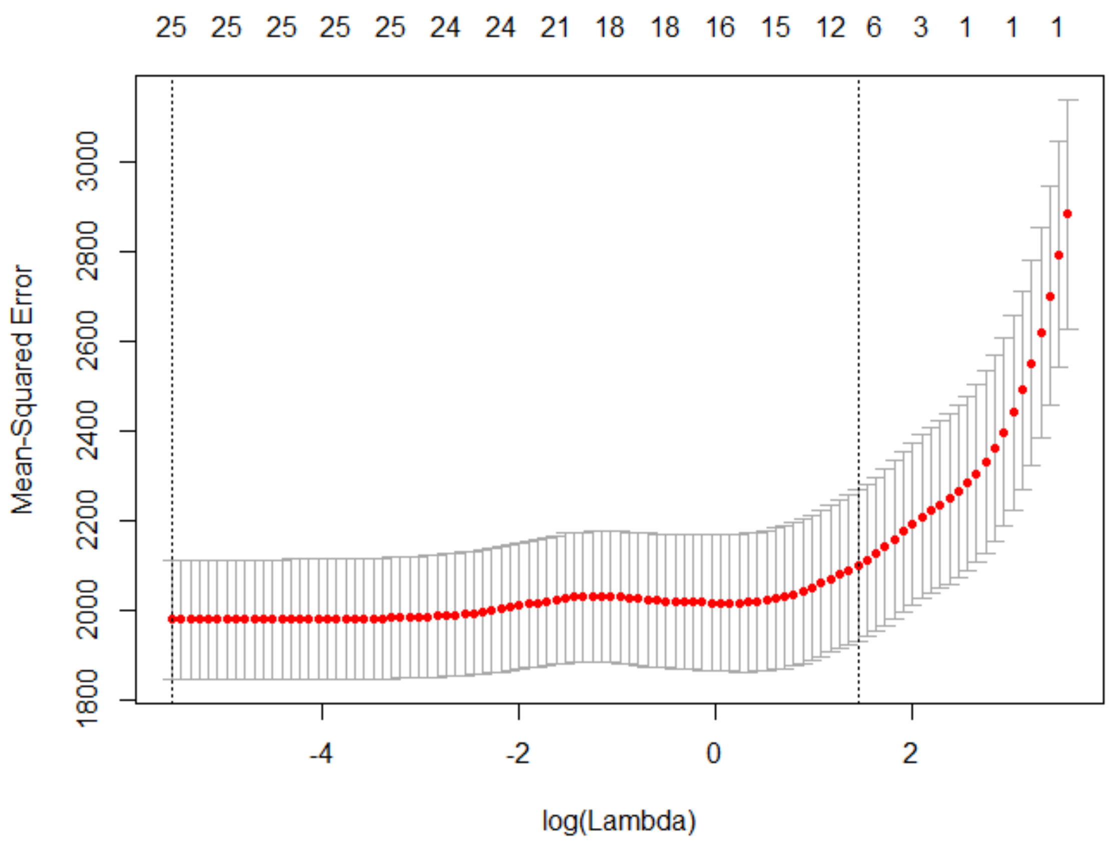
- Package doesn't label these plots (it should) but the darker ones have alphas closer to 1

- Taken together, we can surmise Lasso should do best with this data, but we can re-estimate to ensure alpha =0.75 isn't best

plot(enet_fit$modlist[[#]]) stores individual models

```
# look at each individual model
plot(enet_fit$modlist[[4]])
```

- Alpha = 0.75 model



coef(mod) must specify alpha

```r
# view coefficients
coef(enet_fit, alpha = 0.75) %>%
  round(3)
```

```
> coef(enet_fit, alpha = 0.75) %>%
+   round(3)
29 x 1 sparse Matrix of class "dgCMatrix"
                                   1
(Intercept)                  700.270
color                              .
color Black and White              .
colorColor                         .
num_critic_for_reviews             .
duration                           .
director_facebook_likes            .
actor_3_facebook_likes             .
actor_1_facebook_likes             .
num_voted_users                0.000
cast_total_facebook_likes          .
facenumber_in_poster               .
num_user_for_reviews               .
content_ratingPG               0.983
content_ratingPG-13                .
content_ratingR               -5.767
content_ratingOther                .
title_year                    -0.347
actor_2_facebook_likes             .
imdb_score                         .
aspect_ratio                  -2.762
movie_facebook_likes               .
genre_mainAction              -3.954
genre_mainAdventure                .
genre_mainComedy               3.174
genre_mainCrime                    .
genre_mainDrama                    .
genre_mainOther                    .
cast_total_facebook_likes000s      .
```

coef(mod) must specify alpha

```r
coef(enet_fit, alpha = 1) %>%
  round(3)
```

```
> coef(enet_fit, alpha = 1) %>%
+   round(3)
29 x 1 sparse Matrix of class "dgCMatrix"
                                    1
(Intercept)                   658.831
color                               .
color Black and White               .
colorColor                          .
num_critic_for_reviews              .
duration                            .
director_facebook_likes             .
actor_3_facebook_likes              .
actor_1_facebook_likes              .
num_voted_users                 0.000
cast_total_facebook_likes           .
facenumber_in_poster                .
num_user_for_reviews                .
content_ratingPG                0.437
content_ratingPG-13                 .
content_ratingR                -5.377
content_ratingOther                 .
title_year                     -0.327
actor_2_facebook_likes              .
imdb_score                          .
aspect_ratio                   -2.283
movie_facebook_likes                .
genre_mainAction               -3.545
genre_mainAdventure                 .
genre_mainComedy                2.779
genre_mainCrime                     .
genre_mainDrama                     .
genre_mainOther                     .
cast_total_facebook_likes000s       .
```

Compare coefficients across alphas

```
> enet_coefs
                    varname   ridge alpha025 alpha05 alpha075   lasso
1               (Intercept) 771.379  759.236 675.118  700.270 658.831
2                     color   3.081    0.000   0.000    0.000   0.000
3     color Black and White  -4.819    0.000   0.000    0.000   0.000
4                colorColor   4.733    0.000   0.000    0.000   0.000
5      num_critic_for_reviews   0.017    0.000   0.000    0.000   0.000
6                  duration  -0.042   -0.017   0.000    0.000   0.000
7    director_facebook_likes   0.000    0.000   0.000    0.000   0.000
8     actor_3_facebook_likes   0.001    0.001   0.000    0.000   0.000
9     actor_1_facebook_likes   0.000    0.000   0.000    0.000   0.000
10          num_voted_users   0.000    0.000   0.000    0.000   0.000
11   cast_total_facebook_likes   0.000    0.000   0.000    0.000   0.000
12     facenumber_in_poster  -0.008    0.000   0.000    0.000   0.000
13     num_user_for_reviews   0.013    0.004   0.000    0.000   0.000
14         content_ratingPG   5.331    2.325   0.930    0.983   0.437
15      content_ratingPG-13   1.511    0.000   0.000    0.000   0.000
16          content_ratingR  -5.087   -6.335  -5.542   -5.767  -5.377
17      content_ratingOther   6.017    0.000   0.000    0.000   0.000
18               title_year  -0.391   -0.377  -0.334   -0.347  -0.327
19    actor_2_facebook_likes   0.000    0.000   0.000    0.000   0.000
20               imdb_score   3.201    0.814   0.000    0.000   0.000
21             aspect_ratio  -5.467   -3.766  -2.694   -2.762  -2.283
22     movie_facebook_likes   0.000    0.000   0.000    0.000   0.000
23         genre_mainAction  -4.083   -4.213  -3.595   -3.954  -3.545
24      genre_mainAdventure   0.622    0.000   0.000    0.000   0.000
25         genre_mainComedy   4.908    3.978   2.879    3.174   2.779
26          genre_mainCrime  -4.073   -0.716   0.000    0.000   0.000
27          genre_mainDrama  -2.292    0.000   0.000    0.000   0.000
28          genre_mainOther   2.851    0.305   0.000    0.000   0.000
29 cast_total_facebook_likes000s   0.015    0.000   0.000    0.000   0.000
```