

Lasso Regression

Jan-Philipp Kolb

25 September 2018

Lasso Regression

Lasso - least absolute shrinkage and selection operator

- ▶ lasso is a regression analysis method that performs variable selection and regularization (reduce overfitting)
- ▶ in order to enhance the prediction accuracy and interpretability of the statistical model it produces.
- ▶ Originally introduced in geophysics literature in 1986
- ▶ Independently rediscovered and popularized in 1996 by Robert Tibshirani, who coined the term and provided further insights into the observed performance.

Lasso was originally formulated for least squares models and this simple case reveals a substantial amount about the behavior of the estimator, including its relationship to ridge regression and best subset selection and the connections between lasso coefficient estimates and so-called soft thresholding. It also reveals that (like standard linear regression) the coefficient estimates need not be unique if covariates are collinear.

Lasso for other models than least squares

Though originally defined for least squares, lasso regularization is easily extended to a wide variety of statistical models including generalized linear models, generalized estimating equations, proportional hazards models, and M-estimators, in a straightforward fashion.

- ▶ Lasso's ability to perform subset selection relies on the form of the constraint and has a variety of interpretations including in terms of geometry, Bayesian statistics, and convex analysis.

The LASSO is closely related to basis pursuit denoising.

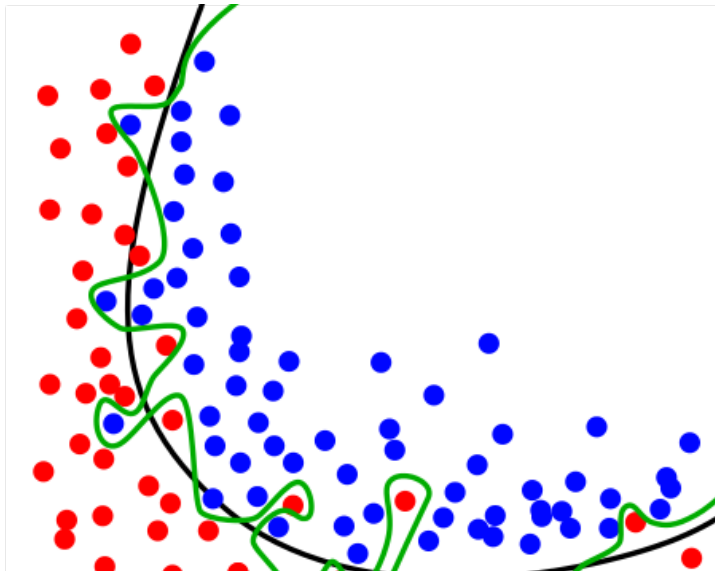
What is lasso regression

- ▶ Lasso regression uses shrinkage
- ▶ data values are shrunk towards a central point
- ▶ Ridge and lasso regularization work by adding a penalty term to the log likelihood function.
- ▶ A tuning parameter, λ controls the strength of the L1 penalty.

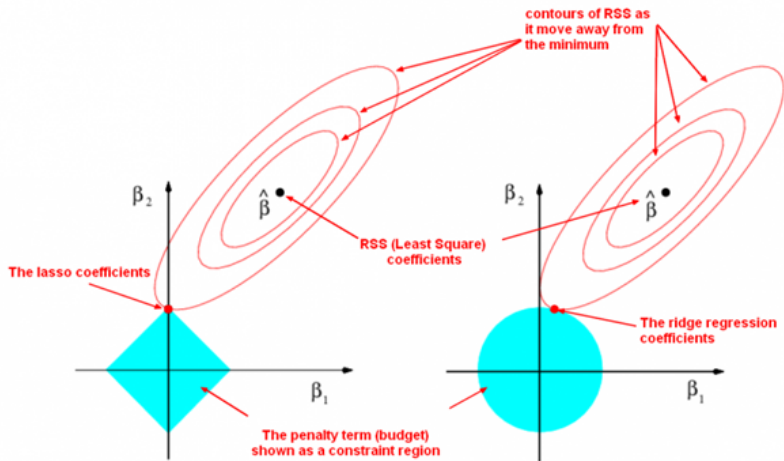
$$\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 + \lambda \sum_{j=1}^p |\beta_j| = RSS + \lambda \sum_{j=1}^p |\beta_j|.$$

Regularization

regularization is the process of adding information in order to solve an ill-posed problem or to prevent overfitting.



The L1 norm explained



LASSO

RIDGE REGRESSION

Ridge Regression and the Lasso

```
swiss <- datasets::swiss  
x <- model.matrix(Fertility~., swiss)[,-1]  
y <- swiss$Fertility  
lambda <- 10^seq(10, -2, length = 100)
```

Test and train dataset

```
library(glmnet)
set.seed(489)
train = sample(1:nrow(x), nrow(x)/2)
test = (-train)
ytest = y[test]
```


A first ols model

#OLS

```
swisslm <- lm(Fertility~., data = swiss)  
coef(swisslm)
```

##	(Intercept)	Agriculture	Examination
##	66.9151817	-0.1721140	-0.2580082
##	Catholic Infant.Mortality		
##	0.1041153	1.0770481	

A ridge model

#ridge

```
ridge.mod <- glmnet(x, y, alpha = 0, lambda = lambda)
predict(ridge.mod, s = 0, type = 'coefficients')[1:6,]
```

##	(Intercept)	Agriculture	Examination
##	66.8911177	-0.1714307	-0.2603091
##	Catholic Infant.Mortality		
##	0.1037196	1.0776950	

Lasso regression with package glmnet

```
install.packages("glmnet")
```

```
library(glmnet)
```

```
x=matrix(rnorm(100*20),100,20)  
g2=sample(1:2,100,replace=TRUE)  
fit2=glmnet(x,g2,family="binomial")
```

```
caret::varImp(fit2,lambda=0.0007567)
```

##		Overall
## V1	0.150397416	
## V2	0.208434920	
## V3	0.633739581	
## V4	0.546691236	
## V5	0.040870939	
## V6	0.131795416	
## V7	0.039071452	
## V8	0.131155125	

- ▶ LASSO is a feature selection method.
- ▶ LASSO regression has inbuilt penalization functions to reduce overfitting.

- ▶ The logarithmic function is used for the link between probability and logits
- ▶ The Logit function is used to linearize sigmoid curves.

The package caret

- ▶ Classification and Regression Training

```
install.packages("caret")
```

```
library("caret")
```

- ▶ **Vignette caret package** -

```
?caret::train
```

```
logit<-train(data = gp.train.c,  
             method = 'glm',  
             family = 'binomial',  
             trControl = ctrl0)
```

Links

A comprehensive beginners guide for Linear, Ridge and Lasso Regression

- ▶ Course for statistical learning - Youtube - Videos
- ▶ pcLasso: a new method for sparse regression
- ▶ Youtube - lasso regression - clearly explained
- ▶ Glmnet Vignette
- ▶ Regularization Methods in R
- ▶ A gentle introduction to logistic regression and lasso regularisation using R
- ▶ Penalized Regression in R