Ridge and LASSO: Step by Step

Bookmark this page

Ridge regularization with validation only: step by step

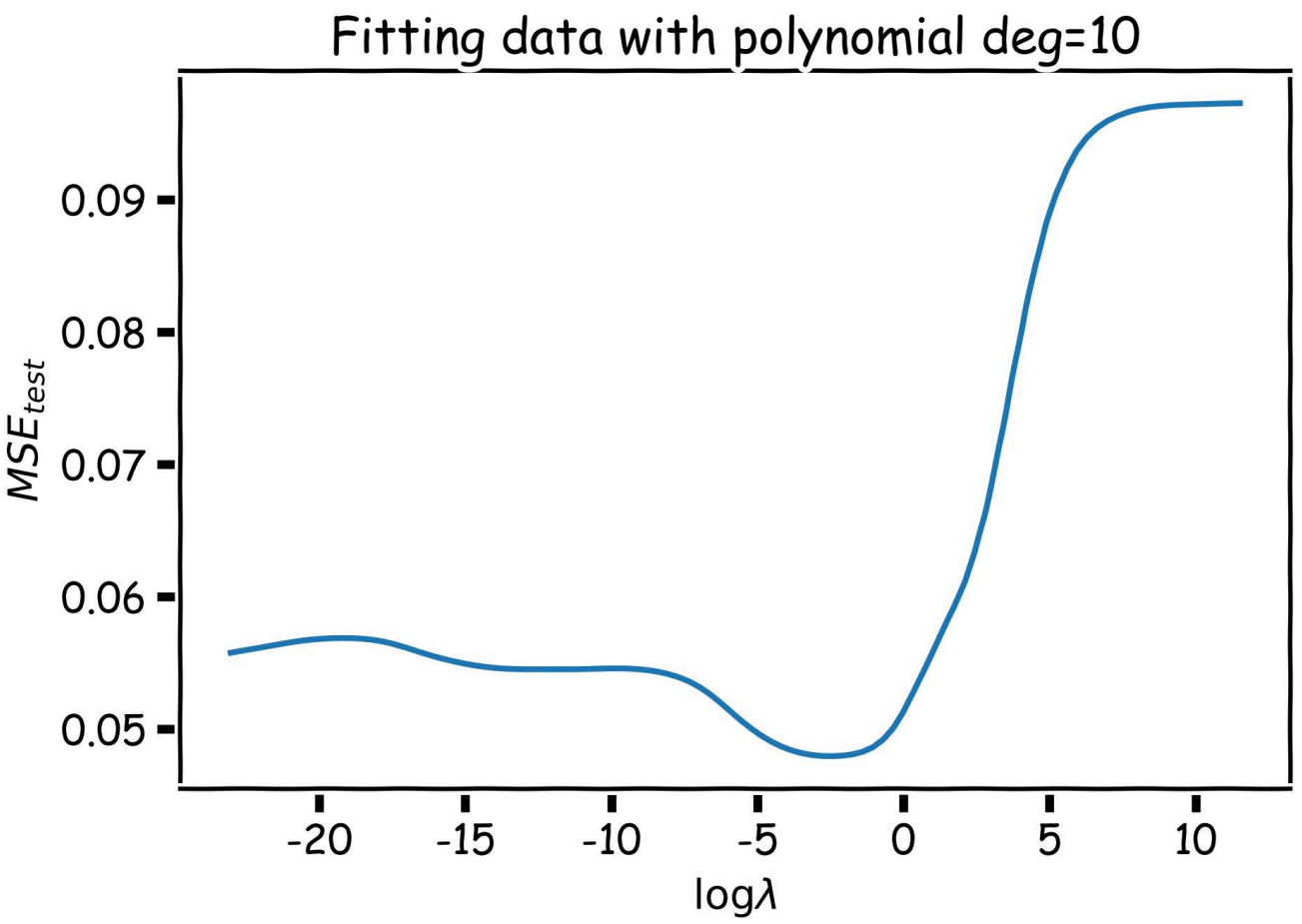
Here we will go through Ridge regularization using using a single validation set using MSE as our loss.

For ridge regression there exist an analytical solution for the coefficients:

1. Split the data into train, validation, and test sets, $X, Y_{train}, X, Y_{validation}, X, Y_{test}$
2. Iterate over a range of λ values for λ in $\lambda_{min} \dots \lambda_{max}$:
   - Determine the β that minimizes the L_{ridge} using the train data, $\beta_{ridge}\left(\lambda\right) = \left(X^T X + \lambda I\right)^{-1} X^T Y$
   - Record the MSE loss for this λ using the validation data, $L_{MSE}\left(\lambda\right)$.
3. Select the λ that minimizes the MSE loss on the validation data,

$$\lambda_{ridge} = argmin_{\lambda} L_{MSE}\left(\lambda\right)$$

4. Refit the model using **both train and validation data combined** using the selected λ, $X, Y_{train}, X, Y_{validation}$, now using λ_{ridge}, resulting to $\hat{\beta} ridge\left(\lambda_{ridge}\right)$
5. Report the MSE on the test set, X, Y_{test} given the $\hat{\beta}_{ridge}\left(\lambda_{ridge}\right)$.



LASSO regularization with validation only: step by step

Here we will go through Lasso regularization using using a single validation set using MSE as our loss .

The steps are largely the same as with Ridge regression except that there is **no** analytical solution for the coefficients in Lasso regression, so we use a **solver**.

1. Split the data into train, validation, and test sets, $X, Y_{train}, X, Y_{validation}, X, Y_{test}$
2. Iterate over a range of λ values for λ in $\lambda_{min} \dots \lambda_{max}$:
   - Determine the β that minimizes the L_{lasso}, $\beta_{lasso}\left(\lambda\right)$ using the train data. **This is done using a solver**.
   - Record the MSE loss for this λ using the validation data, $L_{MSE}\left(\lambda\right)$.
3. Select the λ that minimizes the MSE loss on the validation data,

$$\lambda_{lasso} = argmin_{\lambda} L_{MSE}\left(\lambda\right)$$

4. Refit the model using **both train and validation data combined** using the selected λ,

$X, Y_{train}, X, Y_{validation}$, now using λ_{lasso}, resulting to $\hat{\beta}_{ridge}\left(\lambda_{lasso}\right)$

5. Report the MSE on the test set, X, Y_{test}, given the $\hat{\beta}_{lasso}\left(\lambda_{lasso}\right)$.

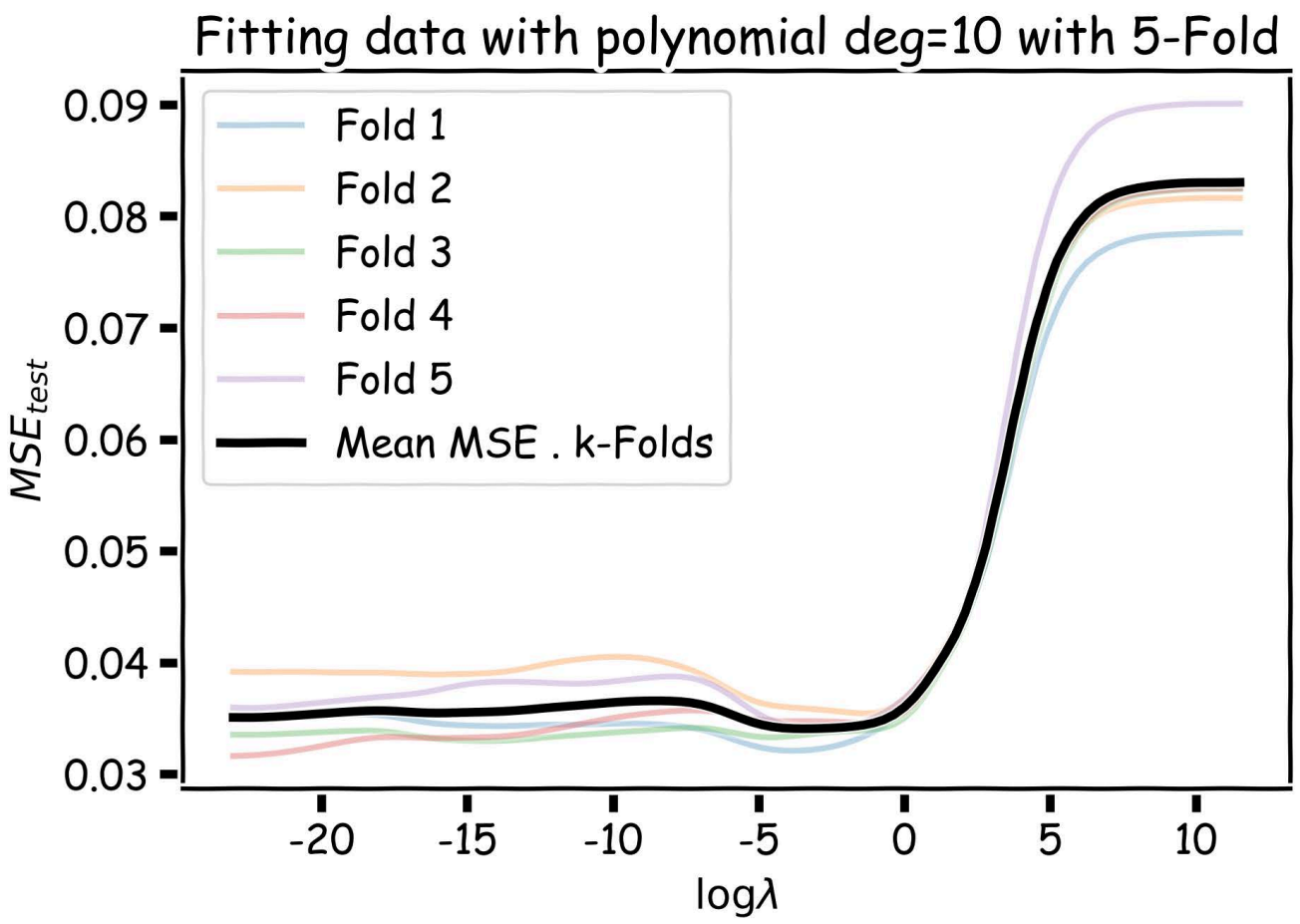
Ridge regularization with CV: step by step

Lastly, let us go through Ridge regularization using using a **cross-validation** using MSE as our loss.

1. Split the data into train, validation, and test sets, $X, Y_{train}, X, Y_{validation}, X, Y_{test}$
2. Split the train data into K folds, $X, Y_{train}^{-k}, X, Y_{validation}^{k}$
3. Iterate over these K folds for k in $1, \ldots, K$
4. Iterate over a range of λ values for λ in $\lambda_0 \ldots \lambda_n$:
   - Determine the β that minimizes the $L_{ridge}, \beta_{ridge}\left(\lambda, k\right) = \left(X^T X + \lambda I\right)^{-1} X^T Y$ using the train data of the fold, X, Y_{train}^{-k}
   - Record $L_{MSE}\left(\lambda, k\right)$ using the validation data of the fold $X, Y_{validation}^{k}$

At this point we have a 2-D matrix, rows are for different k, and columns are for different λ values.

1. Average the $L_{MSE}\left(\lambda, k\right)$ for each λ, $\bar{L}_{MSE}\left(\lambda\right)$
2. Find the λ that minimizes the $\bar{L}_{MSE}\left(\lambda\right)$, resulting to λ_{ridge}.
3. Refit the model using the full training data, $X, Y_{train}, X, Y_{validation}$, resulting to $\hat{\beta}_{ridge}\left(\lambda_{ridge}\right)$
4. Report the MSE on the test set, X, Y_{test} given the $\hat{\beta}_{ridge}\left(\lambda_{ridge}\right)$



Discussion Board (External resource)

Haga clic en Aceptar para que su nombre de usuario y dirección de correo electrónico se envíen a una aplicación de terceros.

Aceptar

What did you learn?

Help future students by selecting the skills or subjects you learned from this unit.

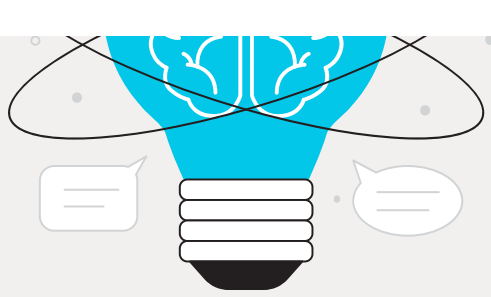
Problem Solving

HyperText Markup Language (HTML)

Programming Environments

Lasso (Programming Language)

Ridge/LASSO Regressions



☐ None of these subjects were covered

Enviar

‹ Previous

Next ›

edX

About
Affiliates
edX for Business
Open edX
Careers
News

Legal

Terms of Service & Honor Code
Privacy Policy
Accessibility Policy
Trademark Policy
Sitemap
Cookie Policy
Your Privacy Choices

Connect

Idea Hub
Contact Us
Help Center
Security
Media Kit

GET IT ON Google play

Download on the App Store

Calculator