Ridge and LASSO: Step by Step

Bookmark this page

Ridge regularization with validation only: step by step

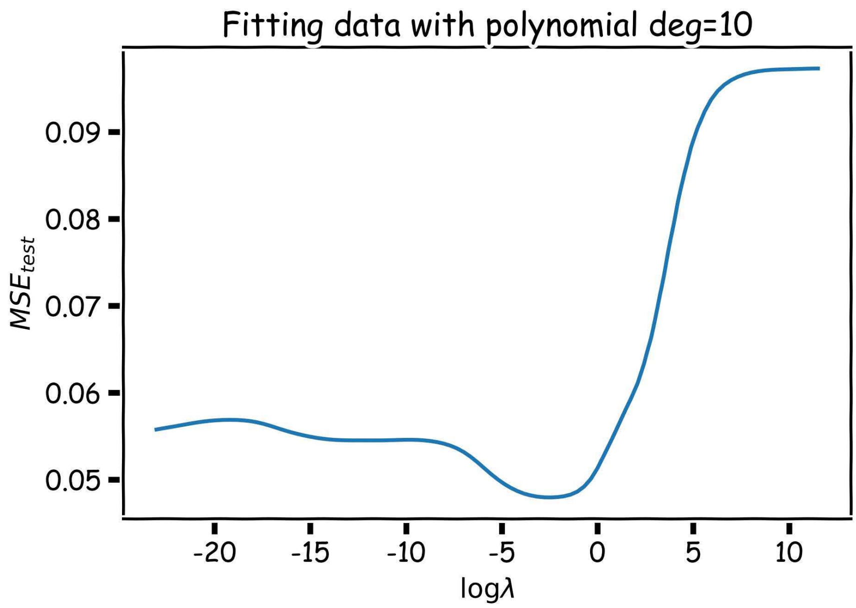
Here we will go through Ridge regularization using using a single validation set using MSE as our loss.

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 $\lambda_{min} \ldots \lambda_{max}$:
   - Determine the β that minimizes the L_{ridge} using the train data, $\beta_{ridge}(\lambda) = (X^T X + \lambda I)^{-1} X^T Y$
   - Record the MSE loss for this λ using the validation data, $L_{MSE}(\lambda)$.
3. Select the λ that minimizes the MSE loss on the validation data,

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

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



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 $\lambda_{min} \ldots \lambda_{max}$:
   - Determine the β that minimizes the L_{lasso}, $\beta_{lasso}(\lambda)$ using the train data. **This is done using a solver**.
   - Record the MSE loss for this λ using the validation data, $L_{MSE}(\lambda)$.
3. Select the λ that minimizes the MSE loss on the validation data,

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

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

Calculator

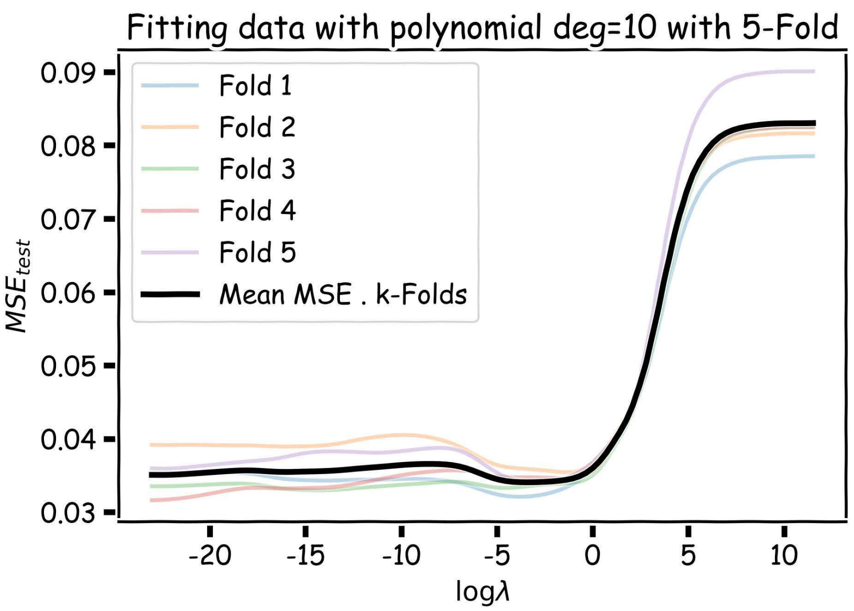
Ridge regularization with CV: step by step

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



Discussion Board (External resource)

Click OK to have your username and e-mail address sent to a 3rd party application.

OK

‹ Previous Next ›

Legal

Connect

Calculator