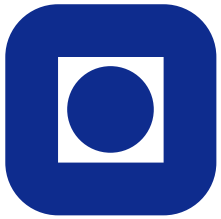


NTNU
Norwegian University of
Science and Technology

TMA4267 Linear Statistical Models V2014

Part 7: Model selection, regularization and dimension reduction
Quiz hosted by Kahoot!

Mette Langaas

To be lectured: April 1, 2014
wiki.math.ntnu.no/emner/tma4267/2014v/start/

1. Model selection

Which of the following is correct for the Best subset selection method? (SSE=sums-of-squares of error on the data set). p=number of predictors.

A The SSE can be used to select the optimal number of predictors.

B The SSE increase as a function of the number of predictors.

C The Mallows Cp, AIC, BIC or R^2_{adj} can be used to select the optimal number of predictors.

D Best subset selection can be used for all values of p.

2. Model selection

Which of the following is correct for the Forward stepwise selection method? (SSE=sums-of-squares of error on the data set). p=number of predictors.

A The SSE can be used to select the optimal number of predictors.

B The SSE increase as a function of the number of predictors.

C The Forward subsect selection method wil always have an SSE that is lower than the SSE for the Best subset selection.

D Forward subset selection can be used for all values of p.

3. Test SSE

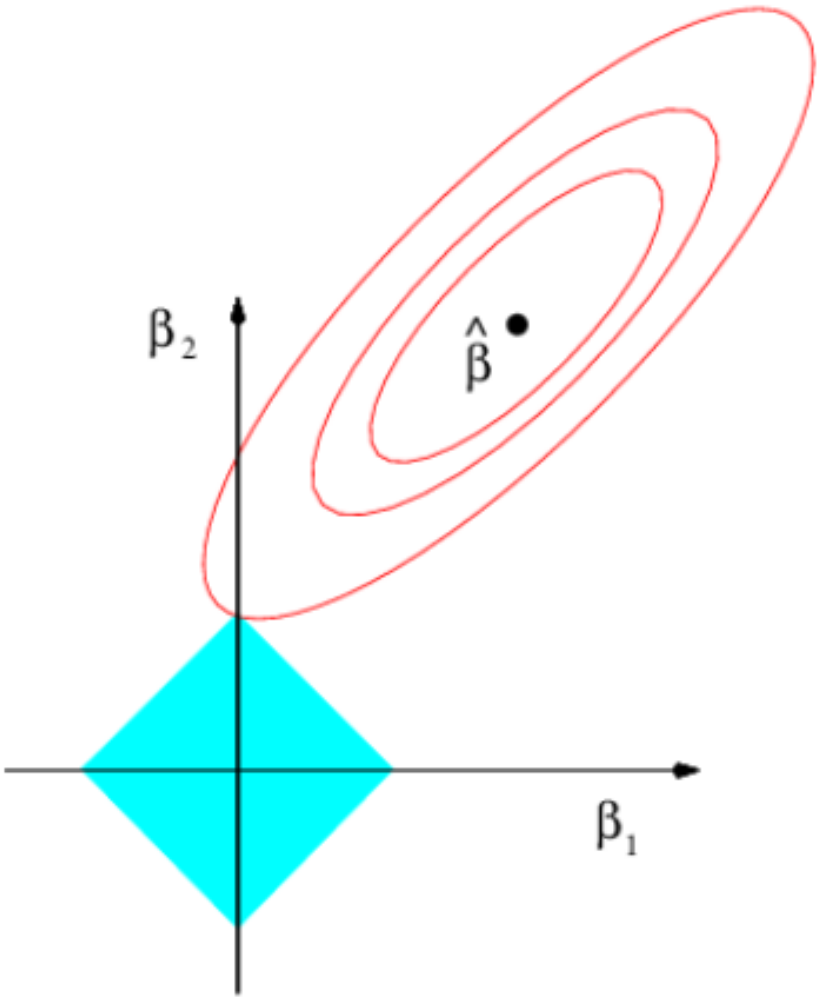
Assume we have a training set to fit the regression model, and a test set to assess the model fit. We estimate the regression coefficients by minimizing

$$(\boldsymbol{Y} - \boldsymbol{X}\beta)^T(\boldsymbol{Y} - \boldsymbol{X}\beta) + \lambda \sum_{j=1}^{p} \beta_j^2$$

for a particular value of λ.
As we increase λ from 0 the test SSE (sums-of-squares of error) will:

A Steadily increase.
B Steadily decrease.
C Increase initially, and then eventually start decrease (inverted U shape).
D Decrease initially, and then eventually start increasing (U shape).

4. Graphical display



Which method is depictured here?

A Best subset selection.

B Ridge regression.

C Lasso regression.

D Principal component regression

5. PCR (quiz problem from ISLR MOOC)

You are working an a regression problem with many variables, so you decide to do Principal component analysis and then fit the regression to the 2 first principal components. Which of the following would you expect to happen?

A A subset of the original predictors will be selected.

B Model bias will decrease relative to fitting least squares to the original predictors.

C Variance of the fitted values will decrease relative to fitting the least squares to the original predictors.

D Model interpretability will improve relative to fitting the least squares to the original predictors.

6. PCR interpretation (quiz problem from ISLR MOOC)

We have a data set where each data point represent a single student's score on a math test, a physics test, a reading comprehension test, and a vocabulary test.
We find the first two principal components, which capture 90% of the variability of the data, and interpret their loadings. We conclude that the first PC represents overall academic ability, the second represents a contrast between quantitative ability and verbal ability.
What loadings would be consistent with that interpretation?

A (0.5,0.5,0.5,0.5) and (0.71,0.71,0,0)

B (0.5,0.5,0.5,0.5) and (0.5,0.5,-0.5,-0.5)

C (0.71,0.71,0,0) and (0,0,0.71,-0.71)

D (0.71,0,-0.71,0) and (0 ,0.71,0,-0.71)

Correct?

Are you sure you want to read the correct answers? Maybe try first?

Answers

CDDCCB