

Quiz 6

Started: Nov 1 at 3:53pm

Quiz Instructions

This quiz is open book, open notes, "open R". The expected duration is 60 minutes. Two attempts are allowed. If both attempts are taken, the score for the second attempt will "overwrite" that from the first attempt (regardless if it is higher or lower). Even though the quiz has 24 points, it will be graded out of 20 points (i.e., 4 points bonus).

You are allowed to use any of the class materials from our SML class, but no other materials (no internet browsing or communication with other parties online/offline).

Even if a question is asking for a numerical value or True/False answer, in order to receive full credit (if your "guess" is correct) or partial credit (if appropriate, if your "guess" is incorrect), please provide your rationale as comments in the uploaded file requested at the end of the quiz.

If you solved some of the questions analytically, provide a clear scan of your work (possibly, with a smartphone app, not a raw photo), in pdf format. In this case, create a folder, place all files in it, make an archive (zip or rar) and submit.

Only if, for whatever reason, the form at the end of the quiz is not working, please submit the supplementary files via Dropbox file request link,

tinyurl.com/nbliznyuk-submit-files

Question 1

4 pts

Consider the lasso and ridge regression solutions (i.e., vectors of estimated coefficients $\hat{\beta}_\lambda$) in multiple linear regression. Briefly discuss the properties of shrinkage and sparsity of these solutions, relative to the (unpenalized) multiple linear regression solution $\hat{\beta} = (X^T X)^{-1} X^T Y$, as the penalty parameter λ increases.

Edit View Insert Format Tools Table

12pt ▾ Paragraph ▾ | **B** *I* U A ▾  ▾ T² ▾ | ⋮

p



0 words



Question 2

4 pts

Briefly (in, at most, three sentences) discuss practical applicability (in what scenarios?), strengths and weaknesses (one of each) for best subsets selection (i.e., exhaustive enumeration) in variable selection.

Edit View Insert Format Tools Table

12pt ▾

Paragraph ▾

B

I

U

A ▾

 ▾

T^2 ▾

⋮

p



0 words

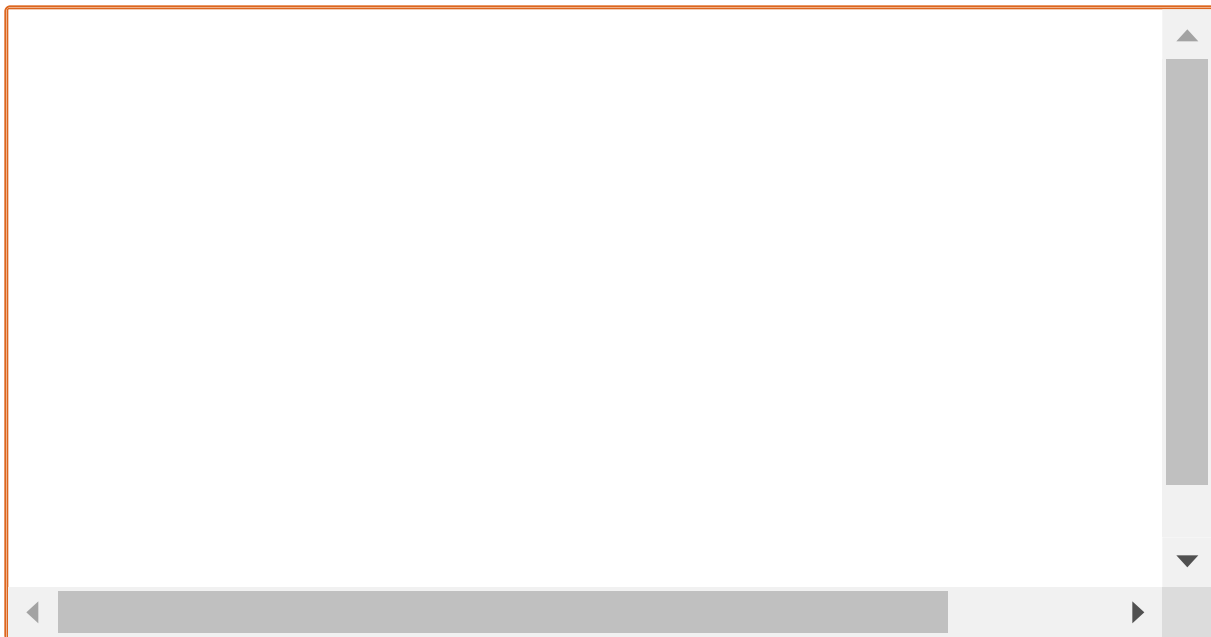


Question 3

4 pts

Recall the setup for the best subset selection in multiple linear regression, where all p predictors are numerical and $n > (p + 1)$ (e.g., page 5 of Chapter 6 slides). There are a total of 2^p models. The set of all models is partitioned into $(p+1)$ subsets, where the k -th subset consists of all models that have exactly k predictors, $k=0, 1, \dots, p$. Briefly explain why, within the k -th class, the best model selected using the unadjusted RSS (the lowest RSS) will coincide with the best model with respect to the adjusted R^2 (in both cases, based on the training data only). Assume that RSS is different for every model (so that there are no "ties" when the models are ranked).

Edit View Insert Format Tools Table

12pt ▾ Paragraph ▾ | **B** *I* U A ▾  ▾ T^2 ▾ | ⋮


p



0 words



Question 4

4 pts

Consider multiple linear regression with p numerical predictors. Suppose p is moderate (e.g., 20) and n is very large (e.g., 1000). Consider the lasso solution where the optimal λ is calibrated by K -fold cross-validation.

True or False: the vector of regression coefficients corresponding to the optimal λ is guaranteed to be sparse, i.e., some of the coefficients are going to be exactly zero.

To receive full credit, please provide justification as the upload in the last item of this quiz.

☐ True

☐ False

Question 5

4 pts

Consider multiple linear regression with p numerical predictors, where $n > (p + 1)$; assume the design matrix is of full rank. Consider two regression problems (both unpenalized):

P1: Regression with the intercept and the p predictors on the original scale.

P2: Regression with the intercept and the p predictors that have been both centered (by subtracting the mean of each predictor column) and rescaled (by dividing each column by its standard deviation).

True or False: the training data coefficient of determination R^2 for the two model fits will be exactly the same.

To receive full credit, please provide justification as the upload in the last item of this quiz.

☐ True

☐ False

Question 6

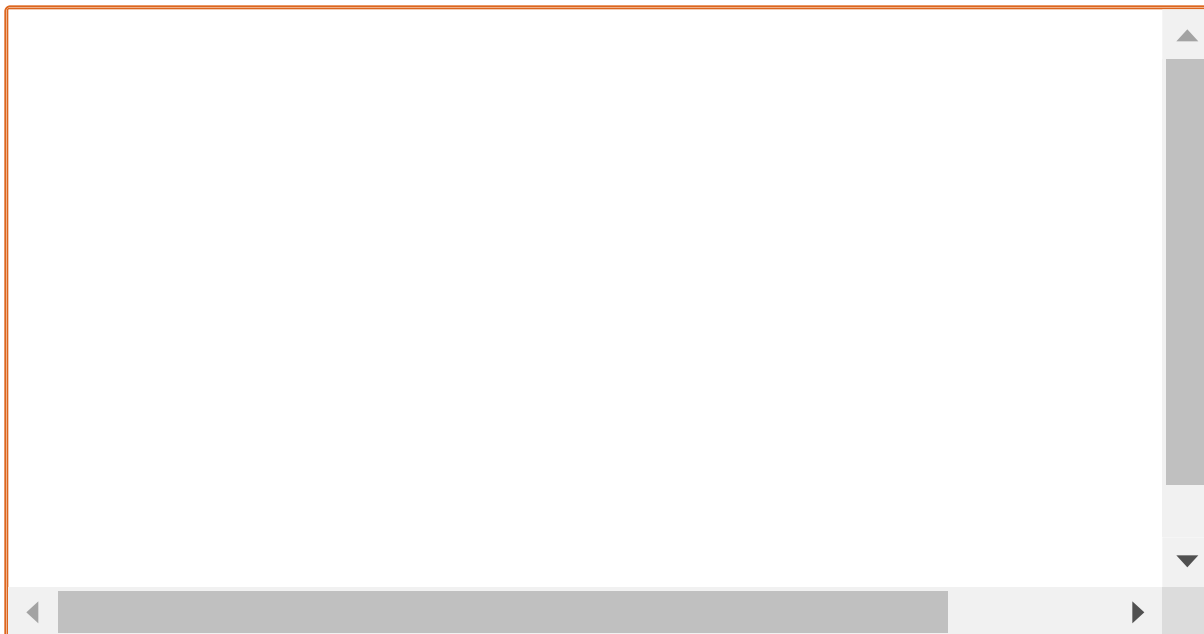
4 pts

Briefly (in at most three sentences) contrast how dimension reduction is achieved in principal components (regression) versus variable selection approaches (such as

best subsets selection or forward/backward stepwise selection) for multiple linear regression with p numerical predictors.

Edit View Insert Format Tools Table

12pt ▾ Paragraph ▾ | **B** *I* U A ▾  ▾ T^2 ▾ | ⋮



p



0 words



Question 7

0 pts

Enter your explanations for the True/False questions here.

Edit View Insert Format Tools Table

12pt ▾ Paragraph ▾ | **B** *I* U A ▾  ▾ T^2 ▾ | ⋮



A screenshot of a presentation editor interface. The main area is a large white rectangle with a thick orange border. At the bottom, there is a status bar. On the left of the status bar is the letter 'p'. In the center is a blue icon of a person at a podium. To the right of the icon is the text '0 words'. Further right are two blue icons: a code symbol '</>' and a zoom-in symbol with an arrow. On the far right is a vertical ellipsis menu icon (three dots).

0 pts

If you did not provide justification to the true/false questions in the field for the previous question, please upload it here.

Upload

Choose a File

Quiz saved at 3:54pm

Submit Quiz