

REVIEW OF SUBSET SELECTION

Best Subset

$M_0: \{ \}$ null model
 $M_1: \{X_1\}, \dots, \{X_p\}$
 $M_2: \{X_1, X_2\}, \{X_1, X_3\}, \dots, \{X_1, X_p\},$
 $\{X_2, X_3\}, \dots, \{X_{p-1}, X_p\}$
 \vdots
 $M_p: \{X_1, \dots, X_p\}$ full model

• Pick M_0, \dots, M_p where M_j is the best model in layer j using in-sample R^2 or RSS.

• Select the best out of M_0, \dots, M_p using one of these criterion: CV, validation set approach, AIC, BIC, Cp.

Forward Stepwise

$M_0: \{ \}$ null model
 $M_1: \{X_1\}, \{X_2\}, \dots, \{X_p\}$
 First select M_1 using in-sample R^2 or RSS. Then create the models in the next layer

$M_2: \{X_2, X_1\}, \{X_2, X_3\}, \dots, \{X_2, X_p\}$

→ Do this again to get M_2 and create the next one.

$M_3: \{X_2, X_3, X_1\}, \{X_2, X_3, X_4\}, \dots, \{X_2, X_3, X_p\}$

$M_p: \{X_1, \dots, X_p\}$ full model

• Select the best out of M_0, \dots, M_p using one of these criterion: CV, validation set approach, AIC, BIC, Cp.

Backward Stepwise

$M_p: \{X_1, \dots, X_p\}$ full model
 $M_{p-1}: \{X_2, \dots, X_p\}, \{X_1, X_3, \dots, X_p\}, \dots, \{X_1, \dots, X_{p-1}\}$

First select M_{p-1} using in-sample R^2 or RSS. Then create the next layer

$M_{p-2}: \{X_3, \dots, X_p\}, \{X_1, X_4, \dots, X_p\}, \dots, \{X_1, X_3, \dots, X_{p-1}\}$

$M_0: \{ \}$ null model

• Select the best out of M_0, \dots, M_p using one of these criterion: CV, validation set approach, AIC, BIC, Cp.

Can stop the search early at some $k^* < p$.

Q:

① How many models are scanned for each of the following?

Best Subset Selection:

2^p

Forward Stepwise Selection:

Backward stepwise Selection:

$\left\{ \begin{array}{l} 2^p \\ 1 + \frac{p(p+1)}{2} \end{array} \right.$

② Does the in-sample R^2 increase (in-sample RSS decrease) from M_0 to M_p ?