

- Greedy algorithms: fast, but only return a **local optimal** solution (which might be good enough in practice).
 - **Backward**: start with the full model and sequentially delete predictors until the score does not improve.
 - **Forward**: start with the null model and sequentially add predictors until the score does not improve.
 - **Stepwise**: consider both deleting and adding one predictor at each stage.

What if $p > n$?

Variable Screening

- When $p \gg n$, stepwise (starting with the full model) cannot be used.
Then we can apply the following screening procedure to pick a model as the starting point for stepwise.
- A simple screening procedure: rank the p predictors by the absolute value of their (**marginal**) correlation with Y ; keep the top K predictors (e.g., $K = n/3$).
- Such a simple screening procedure is likely to miss some important variables, which hopefully could be added back by the stepwise procedure.