**Best Subset, Forward Stepwise, or Lasso?**

* **Best Subset Regression**
  + Given a subset size *k* between 0 and min{*n*, *p*}, finds the subset of *k* predictors that best fits data in terms of squared error, i.e. minimize subject to
  + It ideally finds a global optimum (given subset size)
  + Most computationally challenging
* **Forward Stepwise Selection**
  + Initializing at an empty model, iteratively adds variables that best improve fit
  + Greedy algorithm
* **Lasso**
  + Similar to best subset, but with the penalty for tuning parameter *t*
  + Equivalent to minimizing , where is a tuning parameter
* Generally, in terms of “aggressiveness” in selecting and estimating coefficients, lasso is less aggressive than best subset, and forward selection is in the middle. This is more explicit by saying:
  + Lasso applies shrinkage to its nonzero coefficients, but the other two do not
  + Forward selection is the locally optimal version of best subset
* Comparisons:
  + Lasso more accurate than best subset in low signal-to-noise ratio (SNR) settings, while best subset more accurate with in high SNR settings
  + Best subset can take significantly more time (because it is np-hard)
  + Forward selection performs quite similarly to best subset (though forward selection is much faster)
  + There is little difference in proportion of variance explained (PVE) between methods, so one can focus on computational accuracy/ease in selecting which to use
* Authors also discuss a “relaxed Lasso” developed in another paper, which generally performs equal to or better than the other listed methods