

Extra exercise 3.4

Simulation experiment with OLS, best subset regression, ridge regression, the lasso, least angle regression, principal components regression and partial least squares regression

Extend extra exercise 3.3 by adding principal components regression (PCR) and partial least squares regression (PLS). Perform an extra simulation experiment with negative correlations among the predictors.

a) Generate *one* training data sample with $N = 20$ observations, and estimate the the model by each of the regression methods. Select tuning parameters by 10-fold cross validation.

For PCR and PLS you can use the function `mvr` (or the separate functions `pcr` and `pls`) from the R package `pls`. Use the option `scale=TRUE` in the call to `mvr/pcr/pls` for standardizing the predictors to have the same standard deviation. You can use the function `coef` with option `intercept=TRUE` to extract the estimated coefficients, but you have to scale them back to the original x-scale to get the correct estimates. The scaling factors are found in the attribute `scale` of the object returned from `mvr/pcr/pls`.

b) Extend the simulation experiment by PCR and PLS.

c) Change the correlations between the predictors to -0.06 and repeat the simulation experiment.