

Sample of Subset Selection Procedure

Lester Cajegas

March 15, 2018

Subset Selection

Subset selection for the least squares regression is possible to do in R using the `regsubsets` function in `leaps` package. The said function is akin to the `lm` procedure for the OLS method.

In the summary below, we can see the best model per number of included variables. Variables with asterisk(*) means they are included in the proposed model.

```
## (Intercept) AtBat Hits HmRun Runs RBI Walks Years CAtBat CHits
## 1 TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## 2 TRUE FALSE TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## 3 TRUE FALSE TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## 4 TRUE FALSE TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## 5 TRUE TRUE TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## 6 TRUE TRUE TRUE FALSE FALSE FALSE TRUE FALSE FALSE FALSE
## 7 TRUE FALSE TRUE FALSE FALSE FALSE TRUE FALSE TRUE TRUE
## 8 TRUE TRUE TRUE FALSE FALSE FALSE TRUE FALSE FALSE FALSE
## CHmRun CRuns CRBI CWalks LeagueN DivisionW PutOuts Assists Errors
## 1 FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## 2 FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## 3 FALSE FALSE TRUE FALSE FALSE FALSE FALSE TRUE FALSE FALSE
## 4 FALSE FALSE TRUE FALSE FALSE FALSE TRUE TRUE FALSE FALSE
## 5 FALSE FALSE TRUE FALSE FALSE FALSE TRUE TRUE FALSE FALSE
## 6 FALSE FALSE TRUE FALSE FALSE FALSE TRUE TRUE FALSE FALSE
## 7 TRUE FALSE FALSE FALSE FALSE FALSE TRUE TRUE FALSE FALSE
## 8 TRUE TRUE FALSE TRUE FALSE FALSE TRUE TRUE FALSE FALSE
## NewLeagueN
## 1 FALSE
## 2 FALSE
## 3 FALSE
## 4 FALSE
## 5 FALSE
## 6 FALSE
## 7 FALSE
## 8 FALSE
```

