IDS 702: Module 1.12

Bringing the MLR pieces together II (ILLUSTRATION)

DR. OLANREWAJU MICHAEL AKANDE



Let's try model selection for our diamonds example. We will do this on the log scale - recall our analysis in the previous module.

First, forward selection using AIC

```
diamonds <- read.csv("data/diamonds.csv", header= T,</pre>
                      colClasses = c("numeric","factor","factor","factor","numeric"))
diamonds$CaratsCent <- diamonds$Carats - mean(diamonds$Carats)</pre>
diamonds$CaratsCent2 <- diamonds$CaratsCent^2</pre>
NullModel <- lm(log(Price)~1,data=diamonds)</pre>
FullModel <- lm(log(Price)~CaratsCent+CaratsCent2+
                   Color*Clarity+Color*Certification+
                   Clarity*Certification,
                 data=diamonds)
Model forward <- step(NullModel, scope = formula(FullModel), direction="forward", trace=0)
# Remove the trace=0 option if you want to function to print the entire process
# Let's see the variables the model selected
Model forward$call
## lm(formula = log(Price) ~ CaratsCent + CaratsCent2 + Color +
      Clarity + Certification + Color:Clarity + Color:Certification,
##
      data = diamonds)
##
#run summary(Model forward) to see the results of the final model
```



Let's do the same using BIC

Backward selection using AIC

Same result as forward selection using AIC

Backward selection using BIC

```
Model_backward <- step(FullModel,direction="backward",trace=0,k = log(n))
# Let's see the variables the model selected
Model_backward$call

## lm(formula = log(Price) ~ CaratsCent + CaratsCent2 + Color +
## Clarity, data = diamonds)

#run summary(Model_backward) to see the results of the final model</pre>
```

Same result as forward selection using BIC

Stepwise selection using AIC

```
Model_stepwise <- step(NullModel, scope = formula(FullModel),direction="both",trace=0)
# Let's see the variables the model selected
Model_stepwise$call

## lm(formula = log(Price) ~ CaratsCent + CaratsCent2 + Color +
## Clarity + Certification + Color:Clarity + Color:Certification,
## data = diamonds)

#run summary(Model_backward) to see the results of the final model</pre>
```

Same result as previous results using AIC

Stepwise selection using BIC

Same result as previous results using BIC

Let's use the regsubsets function.

```
library(leaps)
Model_forward <- regsubsets(log(Price)~CaratsCent+CaratsCent2+Color*Clarity+</pre>
                              Color*Certification+Clarity*Certification,data=diamonds,
                            method="forward")
Select results <- summary(Model forward)</pre>
coef(Model forward, which.max(Select results$adjr2)) # Adj R-sq
                                           ColorG
## (Intercept) CaratsCent CaratsCent2
                                                       ColorH
                                                                   ColorT
    8.6185951 3.0050895 -2.0109553
                                       -0.1275071 -0.2147009 -0.3185926
## ClarityVS1 ClarityVS2 ClarityVVS2
## -0.1688242 -0.2525954 -0.1116575
coef(Model_forward, which.min(Select_results$bic)) #BIC
                                           ColorG
                                                       ColorH
## (Intercept) CaratsCent CaratsCent2
                                                                   ColorT
  8.6185951 3.0050895 -2.0109553
                                       -0.1275071 -0.2147009 -0.3185926
## ClarityVS1 ClarityVS2 ClarityVVS2
## -0.1688242 -0.2525954 -0.1116575
```



```
Model_backward <- regsubsets(log(Price)~CaratsCent+CaratsCent2+Color*Clarity+</pre>
                             Color*Certification+Clarity*Certification,data=diamonds,
                           method="backward")
Select results <- summary(Model backward)</pre>
coef(Model_backward, which.max(Select_results$adjr2)) # Adj R-sq
                                      ColorG
## (Intercept) CaratsCent CaratsCent2
                                                      ColorH
                                                                 ColorI
## 8.6185951 3.0050895 -2.0109553
                                      -0.1275071 -0.2147009 -0.3185926
## ClarityVS1 ClarityVS2 ClarityVVS2
## -0.1688242 -0.2525954 -0.1116575
coef(Model_backward, which.min(Select_results$bic)) #BIC
## (Intercept) CaratsCent CaratsCent2
                                          ColorG
                                                      ColorH
                                                                 ColorI
## 8.6185951 3.0050895 -2.0109553
                                      -0.1275071 -0.2147009 -0.3185926
## ClarityVS1 ClarityVS2 ClarityVVS2
## -0.1688242 -0.2525954 -0.1116575
```



WHAT'S NEXT?

MOVE ON TO THE READINGS FOR THE NEXT MODULE!

