Homework: 2024/11/13

1.

1.

$$C_p = \hat{e}'\hat{e} + 2ks^2 = \hat{e}'\hat{e}\left(1 + \frac{2k}{n-k}\right)$$

$$E[\hat{e}'\hat{e}] = E[(MY)'(MY)] = E[e'Me] = (n-k)\sigma^2$$

$$E[C_p] = E[\hat{e}'\hat{e}]\left(1 + \frac{2k}{n-k}\right) = \sigma^2(n+k)$$

and since $R = E[\|m - \hat{m}\|^2] = E[\|X\beta - X\hat{\beta}\|^2] = E[\|Pe\|^2] = E[e'Pe] = k\sigma^2$

$$\therefore E[C_p] = R + n\sigma^2. \quad QED.$$

2.

2.

$$E[\hat{e}_1'\hat{e}_1] = E[(M_1Y)'(M_1Y)] = E[(M_1e + M_1X_2'\beta_2)'(M_1e + M_1X_2'\beta_2)]$$

$$= E[e'M_1e] + E[(M_1X_2'\beta_2)'(M_1X_2'\beta_2)] + 2E[e'M_1(M_1X_2'\beta_2)]$$

we know that $E[e] = 0$ and e is independent of X

$$= E[e'M_1e] + E[(M_1X_2'\beta_2)'(M_1X_2'\beta_2)]$$

$$= (n-k_1)\sigma^2 + E[\beta_2'X_2'M_1X_2\beta_2]$$

$$\therefore E[C_{1,p}] = E[\hat{e}_1'\hat{e}_1] + \frac{2k_1}{n-k}E[\hat{e}'\hat{e}]$$
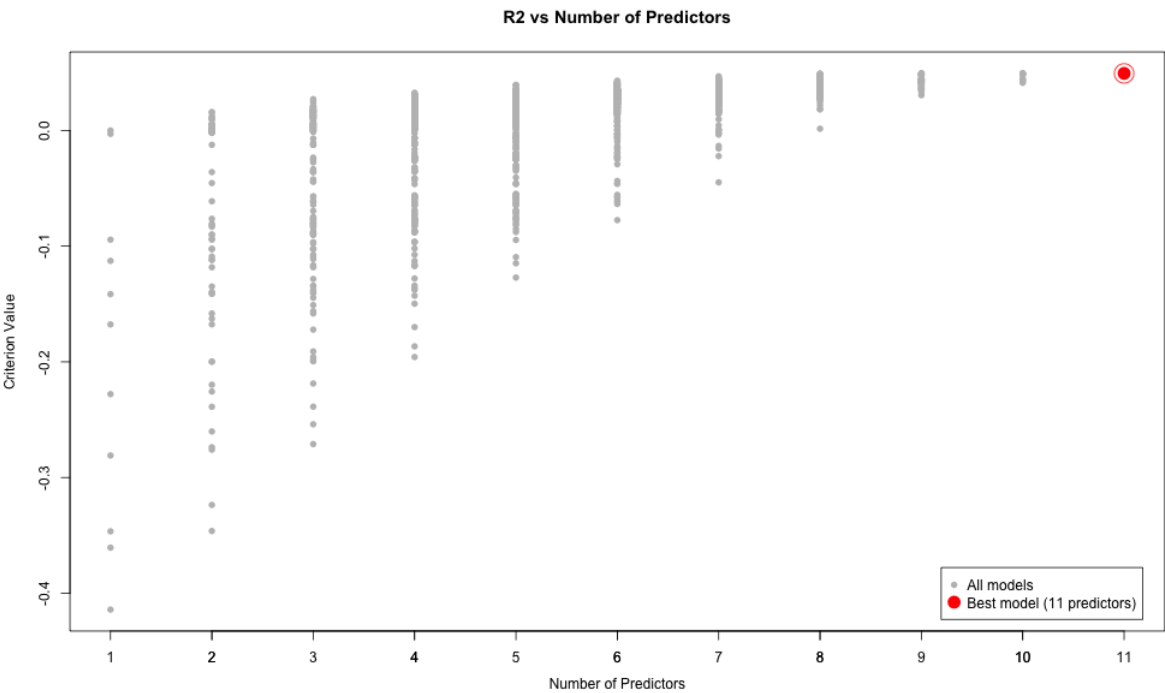
$$= (n-k_1)\sigma^2 + E[\beta_2'X_2'M_1X_2\beta_2] + 2k_1\sigma^2$$

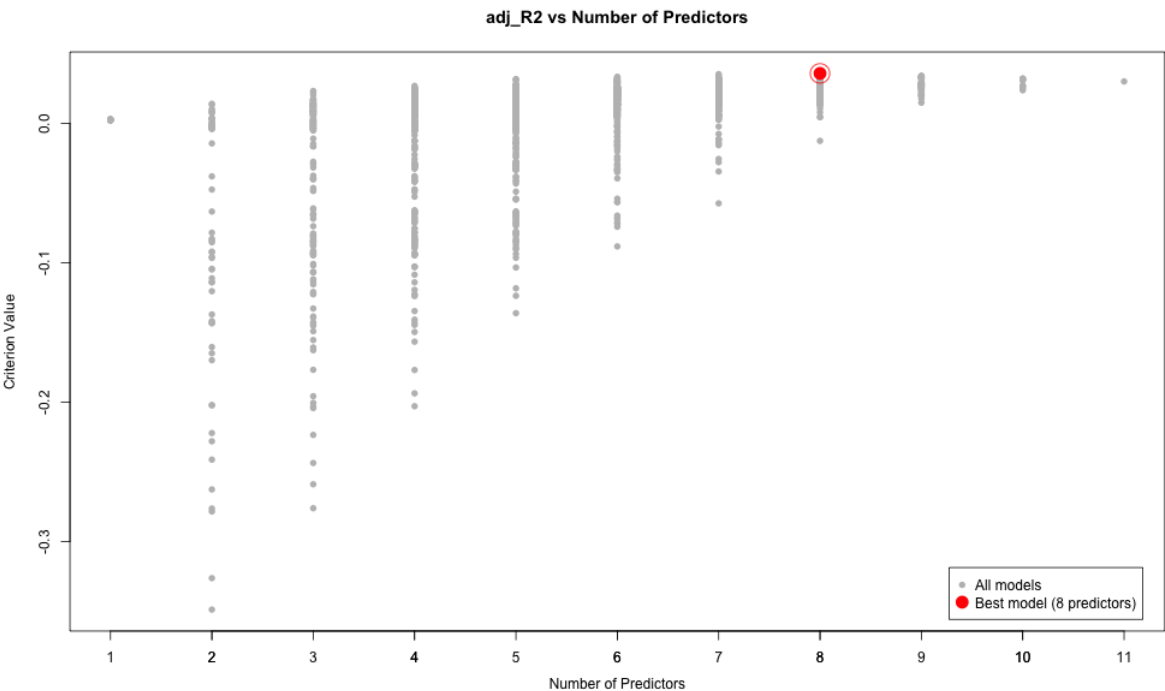$$= n\sigma^2 + k_1\sigma^2 + E[\beta_2'X_2'M_1X_2\beta_2]$$

$$= n\sigma^2 + R, \quad QED.$$
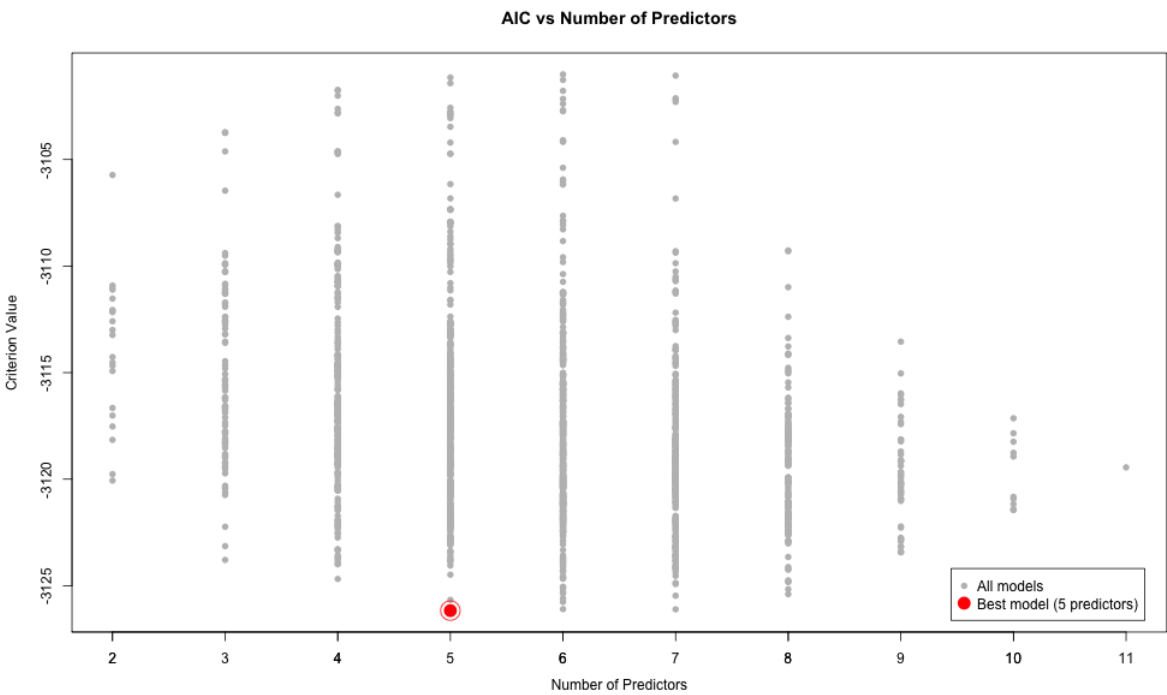
3. Model Selection

Plot of R2

Plot of Adjusted_R2

Observing the plots of R2 and adjusted R2, we can see the trend of R2 is the same as adj_R2, as the model complexity increases, criterion value of R2 and adj_R2 tend to increase. The difference is that the criterion value of adjusted R2 is affected by the model complexity, while the criterion value of R2 is not. The best model selected by R2 and adj_R2 are the model that has the largest criterion value.

Plot of AIC

Plot of BIC
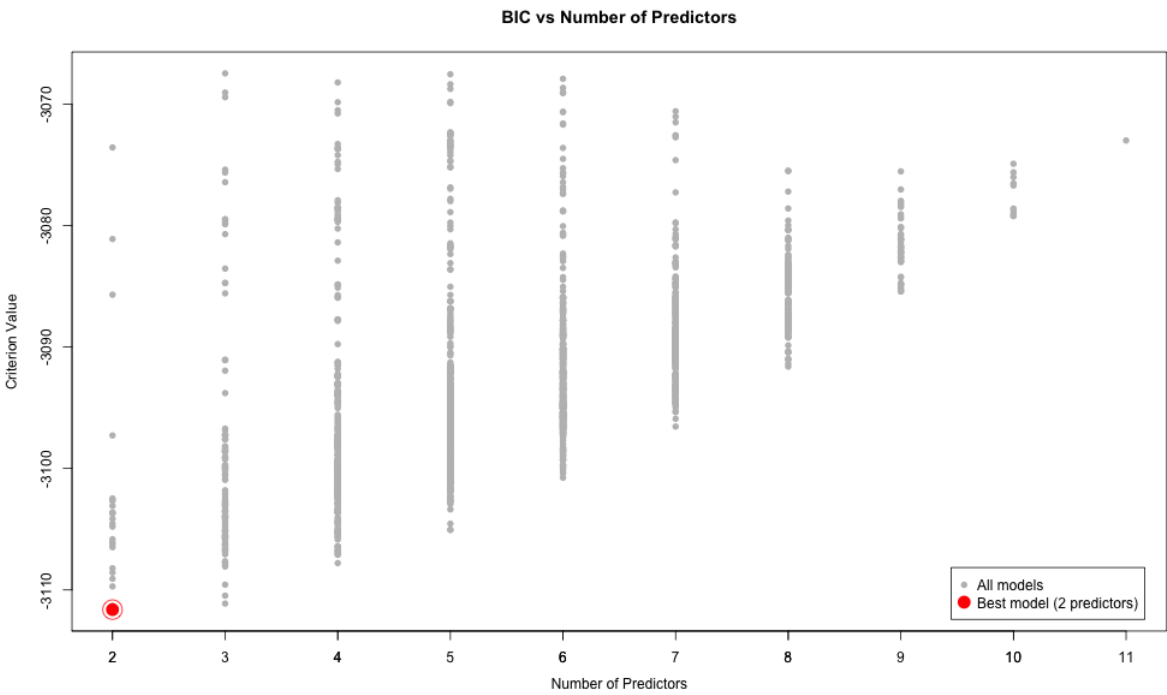
Observing the plots of AIC and BIC, we can see that the model selected are different from R2 and adjustet R2. Since AIC and BIC adds a penalty for the model complexity, the model selected by AIC and BIC are usually have smaller model complexity than that selected by R2 and adjusted R2. The best model selected by AIC and BIC are the model that has the smallest criterion value.

Plot of Cp
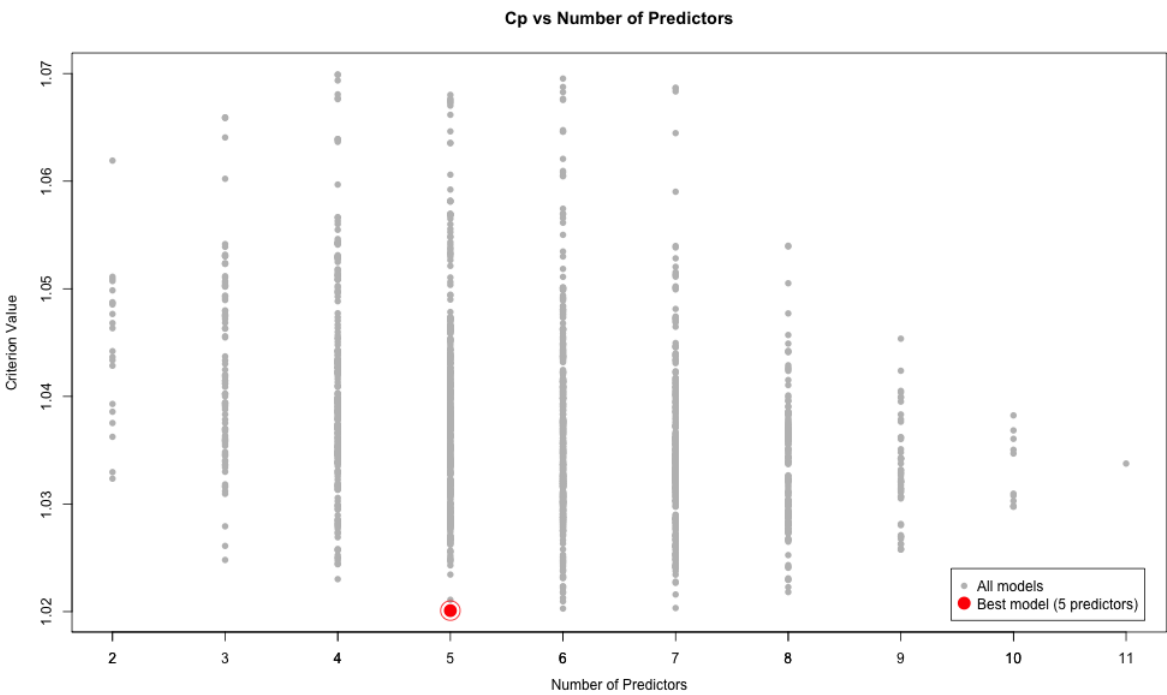
Plot of LOOCV
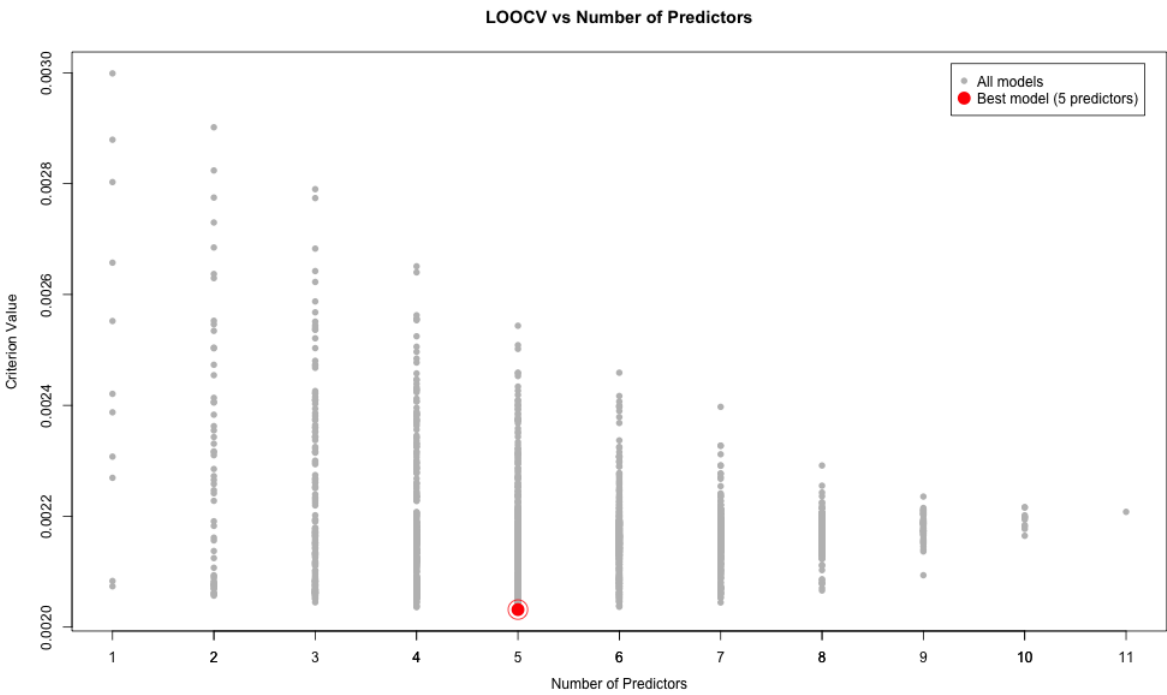
The best model selected by Cp and LOOCV are the model that has the smallest criterion value.

Result of Model Selection

```
[lin1214@Lins-MacBook-Pro hw8 % Rscript homework8.r
Best models:

 R2 :
 [1] "ones"          "dfy"           "infl"          "svar"          "tms"
 [6] "tbl"           "dfy_squared"   "infl_squared"  "svar_squared"  "tms_squared"
[11] "tbl_squared"

 adj_R2 :
[1] "ones"          "dfy"           "svar"          "tms"           "infl_squared"
[6] "svar_squared"  "tms_squared"   "tbl_squared"

 AIC :
[1] "dfy"           "tms"           "tbl"           "infl_squared"  "tms_squared"

 BIC :
[1] "dfy"           "dfy_squared"

 Cp :
[1] "dfy"           "tms"           "tbl"           "infl_squared"  "tms_squared"

 LOOCV :
[1] "dfy"           "tms"           "tbl"           "infl_squared"  "tms_squared"
```

Number of predictors that the best model selected

- R2: 11 (predictors)
- Adjusted_R2: 8
- AIC: 5
- BIC: 2
- Cp: 5
- LOOCV: 5

Also, we can see that AIC, Cp, and LOOCV method chooses the same model respectively.

4. Source Code

Source Code