**110學年度第二學期科學計算軟體作業九**

姓名： 蕭合亭 學號： F64109527

使用datasets package中附加的mtcars資料集進行分析，透過迴歸分析探討因子對於汽車油耗量的影響。內容需包含下列項目：

\*共5題，每題20%；答題提醒：請使用紅框標註**p值(sig.)**、**β(Beta，估計值)及R2等主要判斷依據**，若未達到或錯誤皆會斟酌扣分。

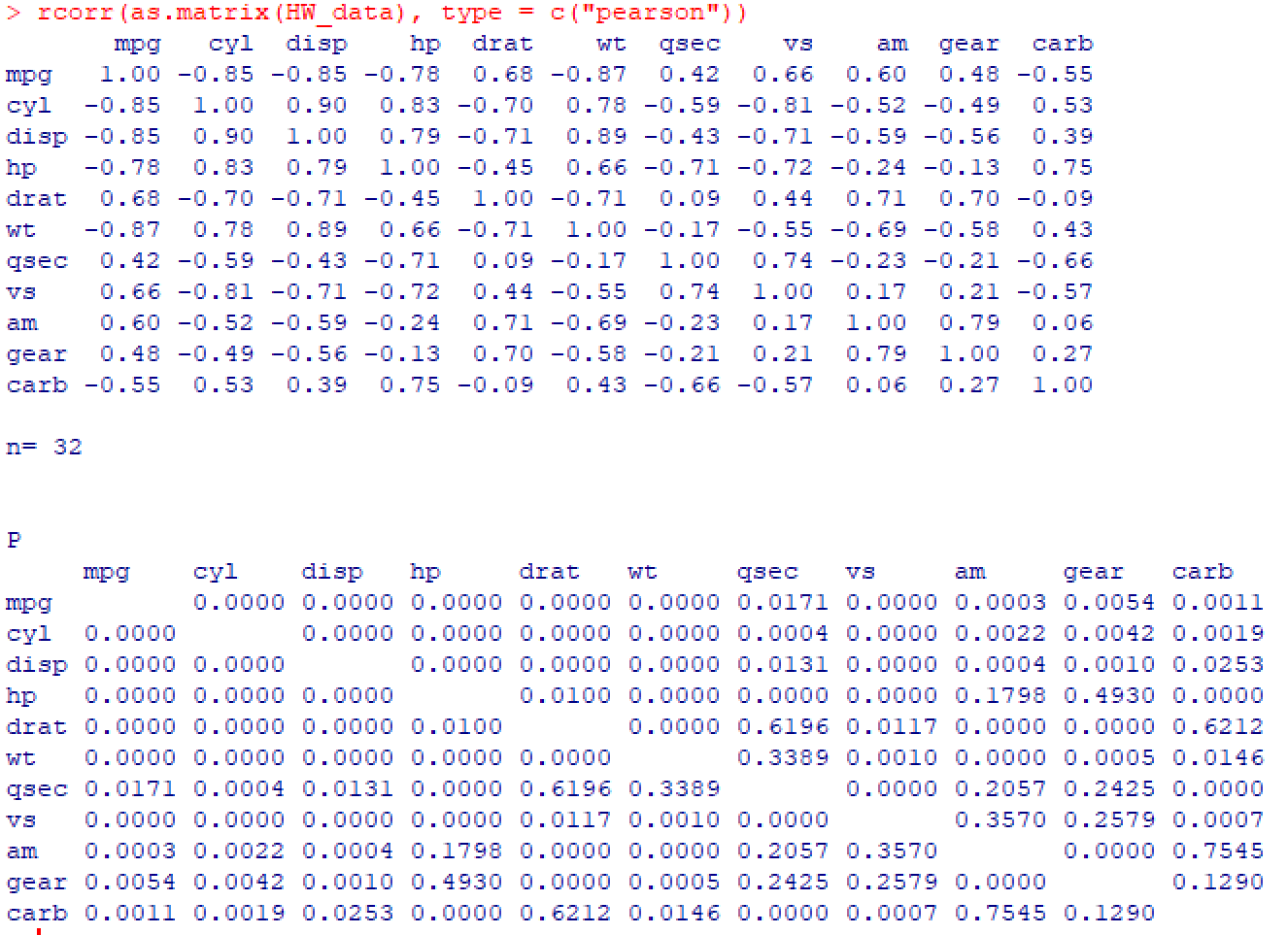
\*(2).~(5).需展示模型配適度(ANOVA)、模型摘要(Summary)、參數估計值(Estimate)等表。

1. **各變數間之相關性檢定**(本表結果表可不用標註相關資訊，但須說明各項因子與油耗量之間的關係正負及是否顯著)

程式碼

rcorr(as.matrix(HW\_data), type = c("pearson"))

結果



分析

mpg-cyl : 相關性=-0.85，為高度負相關；p=0.000<0.05，達顯著性。

mpg-disp : 相關性=-0.85，為高度負相關；p=0.000<0.05，達顯著性。

mpg-hp : 相關性=-0.78，為高度負相關；p=0.000<0.05，達顯著性。

mpg-drat : 相關性=0.68，為高度正相關；p=0.000<0.05，達顯著性。

mpg-wt : 相關性=-0.87，為高度負相關；p=0.000<0.05，達顯著性。

mpg-qsec: 相關性=0.42，為中度正相關；p=0.0171<0.05，達顯著性。

mpg-vs : 相關性=0.66，為高度負相關；p=0.000<0.05，達顯著性。

mpg-am : 相關性=0.60，為高度負相關；p=0.0003<0.05，達顯著性。

mpg-gear: 相關性=0.48，為高度負相關；p=0.0054<0.05，達顯著性。

mpg-carb : 相關性=-0.55，為高度負相關；p=0.0011<0.05，達顯著性。

1. **強制輸入法**(+共線性檢定(回答根據數值是否具有嚴重共線性)+列出模型完整公式)

程式碼

ols\_coll\_diag(data\_model)

data\_model<-lm(HW\_data$mpg~HW\_data$cyl+HW\_data$disp+HW\_data$hp

+HW\_data$drat+HW\_data$wt+HW\_data$qsec+HW\_data$vs

+HW\_data$am+HW\_data$gear+HW\_data$carb)

ols\_regress(data\_model)

結果

一張含有 文字 的圖片

自動產生的描述

特徵值接近0

CI>30

容忍度接近0

VIF>5

一張含有 桌 的圖片

自動產生的描述

分析

sig=0.000，mpg對於其他資料達到統計上的顯著性

R-squared=0.869，表示可以解釋約87%。

Adj R-squared=0.807，表示可以解釋約81%。

ANOVA的sum of squares

由Beta值得到共線式 :

Y= 12.303 -0.111X1+0.013X2-0.021X3+0.787X4-3.715X5+0.821X6

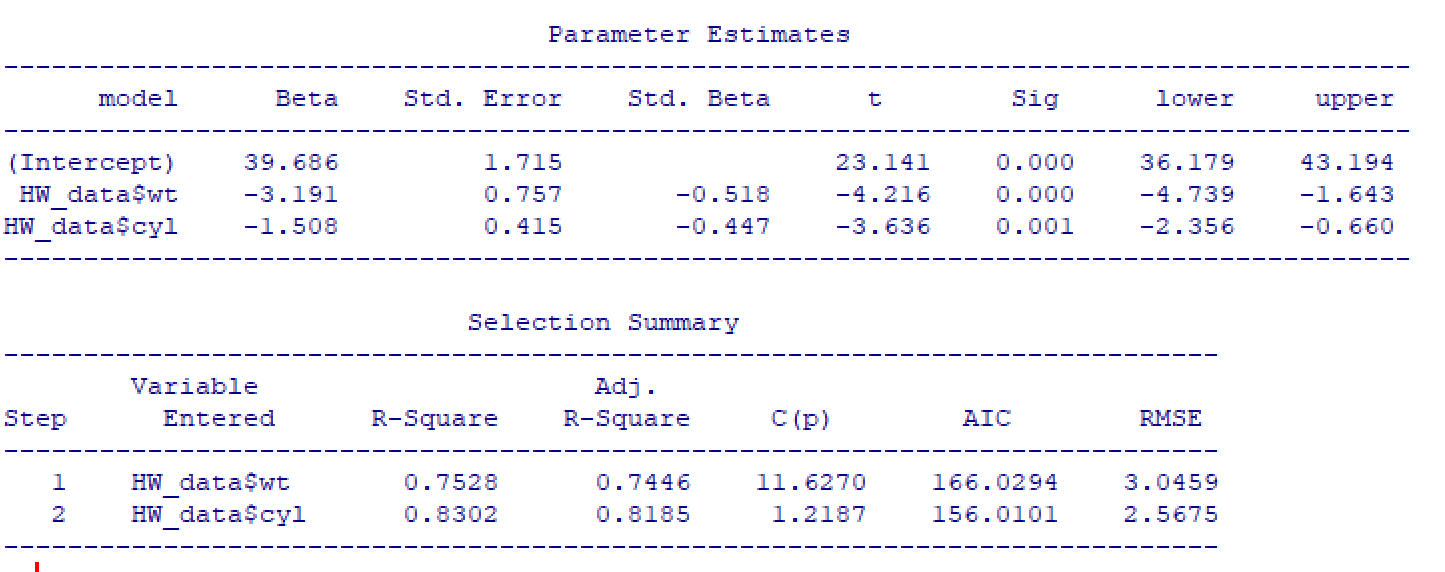
+0.318X7+2.520X8+0.655X9-0.199X10

1. **順向進入法**(+變數選入摘要表(Selection summary))

程式碼

ols\_step\_forward\_p(data\_model,penter=0.05,details=TRUE)

結果



分析

使用順向進入法，當p值<0.05時才會進入，選取了wt、cyl欄。

wt : sig= 0.000，達統計上顯著標準。

cyl : sig= 0.001，達統計上顯著標準。

選取摘要表 :

R-square :

wt 的R-square = 0.7528，表示可以解釋約75%。

cyl 的R-square = 0.8302，表示可以解釋約83%。

Adj R-square :

wt 修正過後的R-square = 0.7446，表示可以解釋約74%。

cyl 修正過後的R-square = 0.8185，表示可以解釋約82%。

由Beta值得到共線式 :

X1為wt項、X2為cyl

Y= 39.686 -3.191X1-1.508X2

1. **反向淘汰法**(+變數淘汰摘要表(Elimination summary))

程式碼

ols\_step\_backward\_p(data\_model,prem=0.1,details=TRUE)

結果

一張含有 桌 的圖片

自動產生的描述

分析

使用反向淘汰法，逐步將p值>0.1的項目刪除，最後淘汰了cyl、vs、carb、gear、drat、disp、hp項。

wt : sig= 0.000，達統計上顯著標準。

qsec : sig= 0.000，達統計上顯著標準。

am : sig=0.047，達統計上顯著標準。

淘汰摘要表 :

R-square :

cyl 的R-square = 0.8689，表示可以解釋約87%。

vs 的R-square = 0.8687，表示可以解釋約87%。

carb 的R-square = 0.8681，表示可以解釋約87%。

gear 的R-square = 0.8667，表示可以解釋約87%。

drat 的R-square = 0.8637，表示可以解釋約86%。

disp 的R-square = 0.8579，表示可以解釋約86%。

hp 的R-square = 0.8497，表示可以解釋約85%。

Adj R-square :

cyl 修正過後的R-square = 0.8153，表示可以解釋約82%。

vs 修正過後的R-square = 0.823，表示可以解釋約82%。

carb 修正過後的R-square = 0.8296，表示可以解釋約83%。

gear 修正過後的R-square = 0.8347，表示可以解釋約83%。

drat 修正過後的R-square = 0.8375，表示可以解釋約84%。

disp 修正過後的R-square = 0.8368，表示可以解釋約84%。

hp 修正過後的R-square = 0.8336，表示可以解釋約83%。

由Beta值得到共線式 :

X1為wt項、X2為qsec項、X3為am項

Y= 9.618 -3.917X1-1.226X2+2.936X3

1. **逐步分析法**(+變數選入/淘汰摘要表(Stepwise selection summary))

**\***選入(Enter)、移除(Remove)條件均0.05及0.1

程式碼

ols\_step\_both\_p(data\_model,penter=0.05,prem=0.1,details=TRUE)

結果

一張含有 桌 的圖片

自動產生的描述

一張含有 桌 的圖片

自動產生的描述

分析

使用逐步分析法(同時衡量選入與淘汰)

將wt、cyl加入參數：

wt : sig= 0.000，達統計上顯著標準。

cyl : sig= 0.001，達統計上顯著標準。

摘要表 :

加入項

R-square :

wt 的R-square = 0.753，表示可以解釋約75%。

cyl 的R-square = 0.830，表示可以解釋約83%。

Adj R-square :

wt 修正後的R-square = 0.745，表示可以解釋約75%。

cyl 修正後的R-square = 0.819，表示可以解釋約82%。

由Beta值得到共線式 :

X1為wt項、X2為cyl

Y= 39.686 -3.191X1-1.508X2

*>> data(mtcars, package = "datasets")*

*>> HW\_data <- mtcars*

*>> View(HW\_data)*

