Weekly 5 Summary

Gaurang Kakade

Table of contents

Tuesday, Feb 7

> **!** TIL
>
> Today, I learnt the following concepts in class:
>
> 1. Integration of regression coefficients
> 2. Categorical covariates
> 3. Multiple regression
>   - Extension from single regression
>   - Other topics

Packages we will require this week

```
library(tidyverse)
```

```
-- Attaching packages --------------------------------------- tidyverse 1.3.2 --
v ggplot2 3.4.0     v purrr   1.0.1
v tibble  3.1.8     v dplyr   1.1.0
v tidyr   1.3.0     v stringr 1.5.0
v readr   2.1.3     v forcats 1.0.0
-- Conflicts ------------------------------------------ tidyverse_conflicts() --
```

```
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
```

```
library(ISLR2)
library(cowplot)
library(kableExtra)
```

```
Attaching package: 'kableExtra'

The following object is masked from 'package:dplyr':

    group_rows
```

Integration of regression coefficients

What is the interpretation of β_0 and β_1 ?

The regression model is given as follows:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

where:

1. y_i is the response
2. x_1 is the covariate
3. ϵ_i is the error (vertical black line in lecture 4 notes)
4. β_0 and β_1 are the regression coefficients
5. $i = 1, 2, \ldots, n$ are the indices for the observations

What is the interpretation for the regression coefficients ?

β_0 is the intercept and β_1 is the slope.

Let's consider the following example using `mtcars`

```
library(ggplot2)
mtcars %>% head() %>% kable()
```

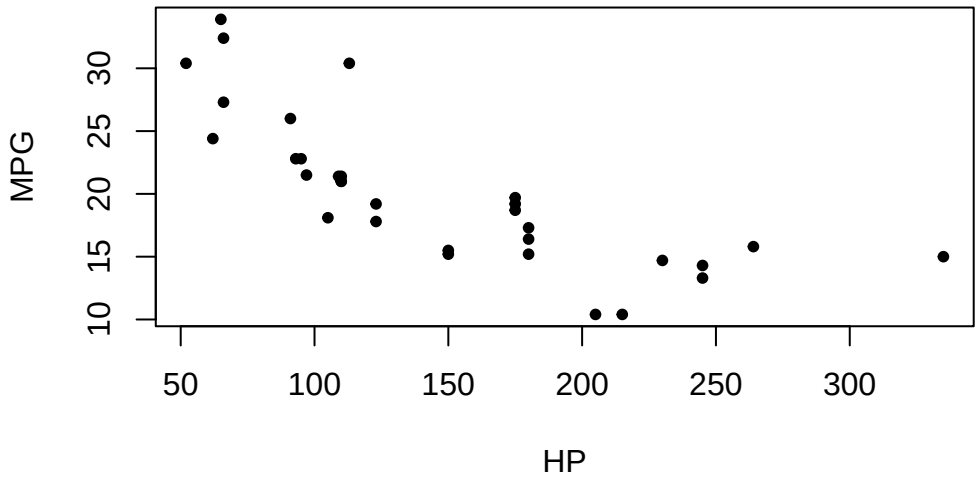
	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360	175	3.15	3.440	17.02	0	0	3	2
Valiant	18.1	6	225	105	2.76	3.460	20.22	1	0	3	1

The above code uses the mtcars dataset that comes pre-installed with R, the code is using the pipe operator %>% to pass the `mtcars` data to the function head(). This function returns the first 6 rows of the mtcars dataset. The result of head() is then passed to the kable() function from the knitr package. This function formats the data as a nice-looking table and outputs it in the R console.

Consider the following relationship

```
x <- mtcars$hp
y <- mtcars$mpg

plot(x, y, pch = 20, xlab = "HP", ylab = "MPG")
```




```
model <- lm(y~x) # This line of code creates a linear
# regression model object in R. The response variable "y"
# is modeled as a linear function of the predictor
# variable "x". The syntax of the lm() function stands
# for "linear model". After running this code, the object
# model will contain information about the fit of the
# regression model, such as the coefficients, residuals,
```