HW10

Alex Parra

5/5/2022

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.6     v dplyr   1.0.7
## v tidyr   1.2.0     v stringr 1.4.0
## v readr   2.1.2     v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

Problem 1

```
y <- c(1, 0, 0, 1, 2, 3, 3)
x1 <- c(-3, -2, -1, 0, 1, 2, 3)
x2 <- c(5, 0, -3, -4, -3, 0, 5)
x3 <- c(-1, 1, 1, 0, -1, -1, 1)
df <- data.frame(y, x1, x2, x3)
df
```

```
##   y x1 x2 x3
## 1 1 -3  5 -1
## 2 0 -2  0  1
## 3 0 -1 -3  1
## 4 1  0 -4  0
## 5 2  1 -3 -1
## 6 3  2  0 -1
## 7 3  3  5  1
```

Part (a)

```
m1 <- lm(y ~ x1 + x2 + x3, data = df)
summary(m1)
```

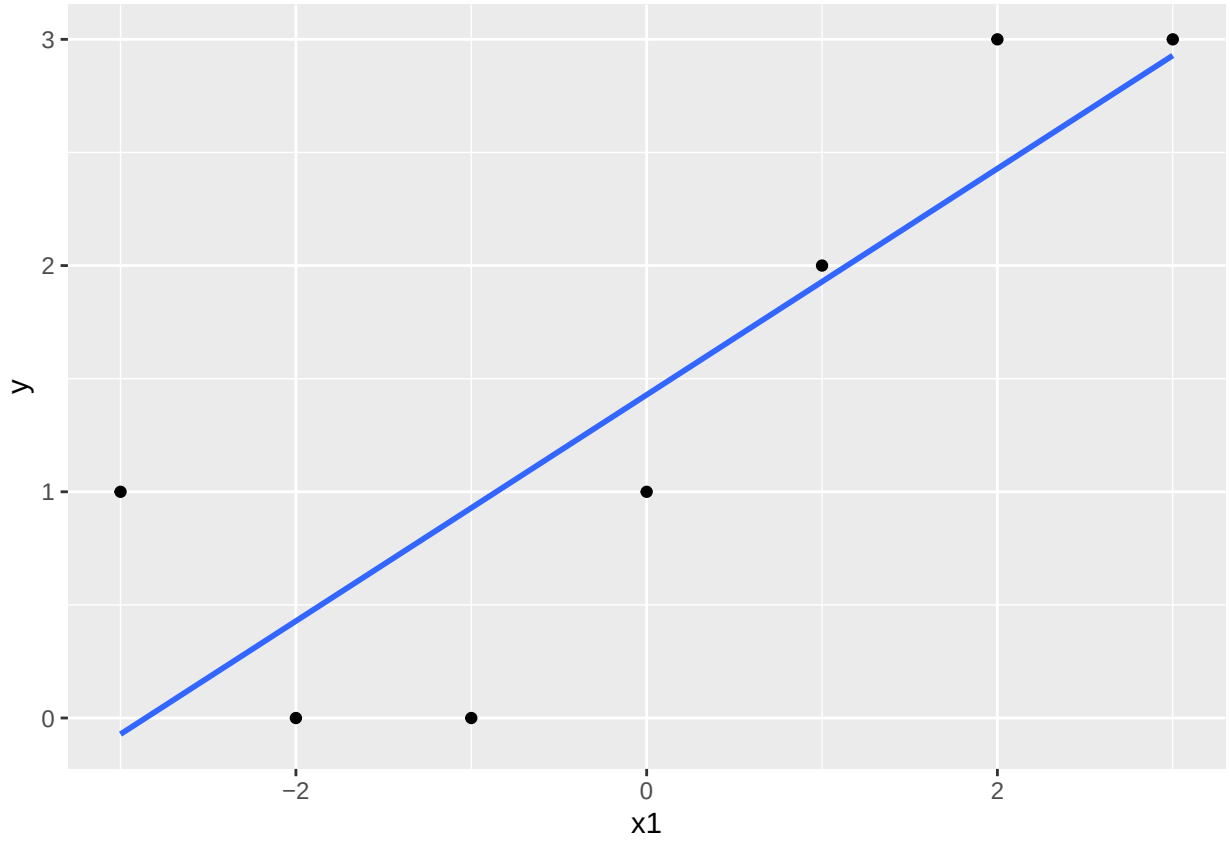
```
##
## Call:
## lm(formula = y ~ x1 + x2 + x3, data = df)
##
## Residuals:
##       1       2       3       4       5       6       7
## -0.02381  0.07143 -0.07143  0.04762 -0.07143  0.07143 -0.02381
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.42857    0.03367   42.43 2.88e-05 ***
## x1           0.50000    0.01684   29.70 8.38e-05 ***
## x2           0.11905    0.00972   12.25 0.001172 **
## x3          -0.50000    0.03637  -13.75 0.000833 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.08909 on 3 degrees of freedom
## Multiple R-squared:  0.9975, Adjusted R-squared:  0.9951
## F-statistic:   407 on 3 and 3 DF,  p-value: 0.0002058
```

We get this equation:

$$\hat{y} = 1.42857 + 0.50000 \times x_1 + 0.11905 \times x_2 - 0.50000 \times x_3$$

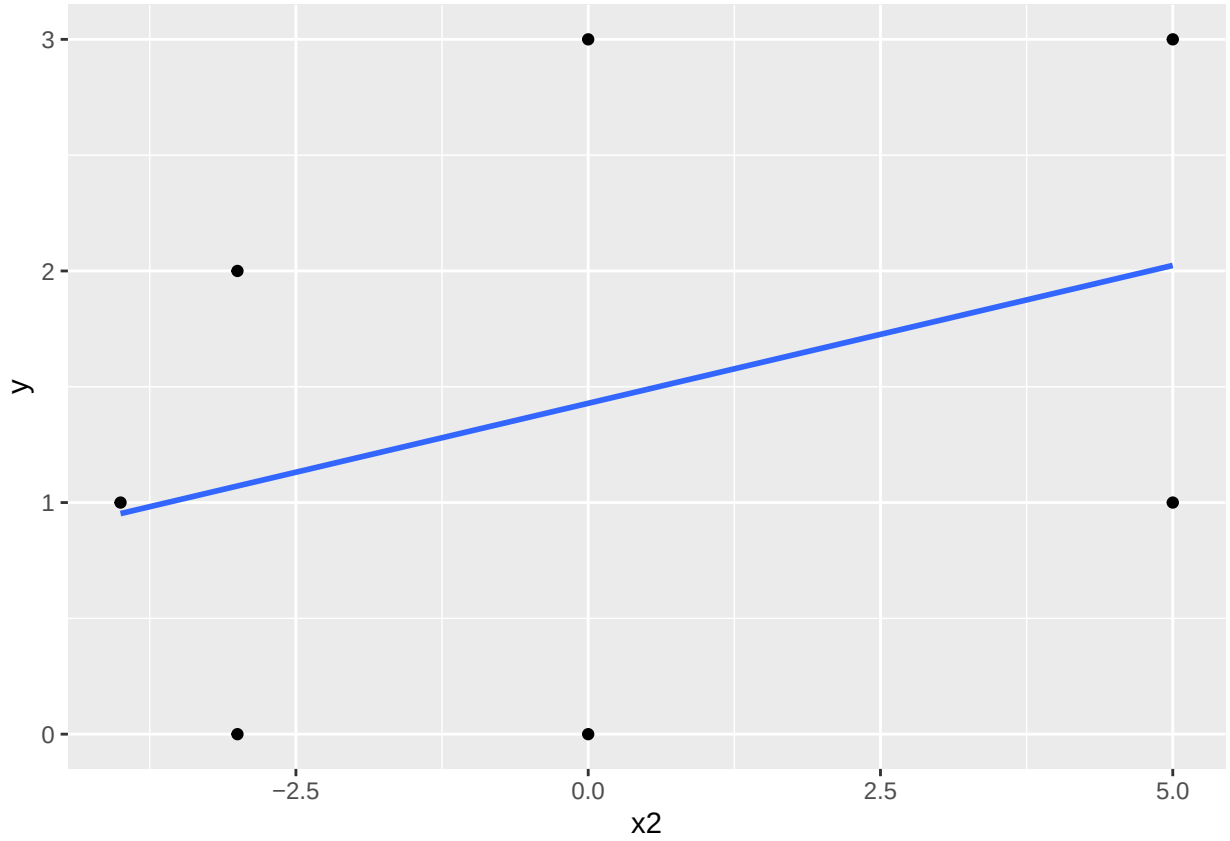
```
ggplot(data = df, aes(x = x1, y = y)) +
  geom_point() +
  stat_smooth(method = "lm", se = FALSE)
```

```
## `geom_smooth()` using formula 'y ~ x'
```



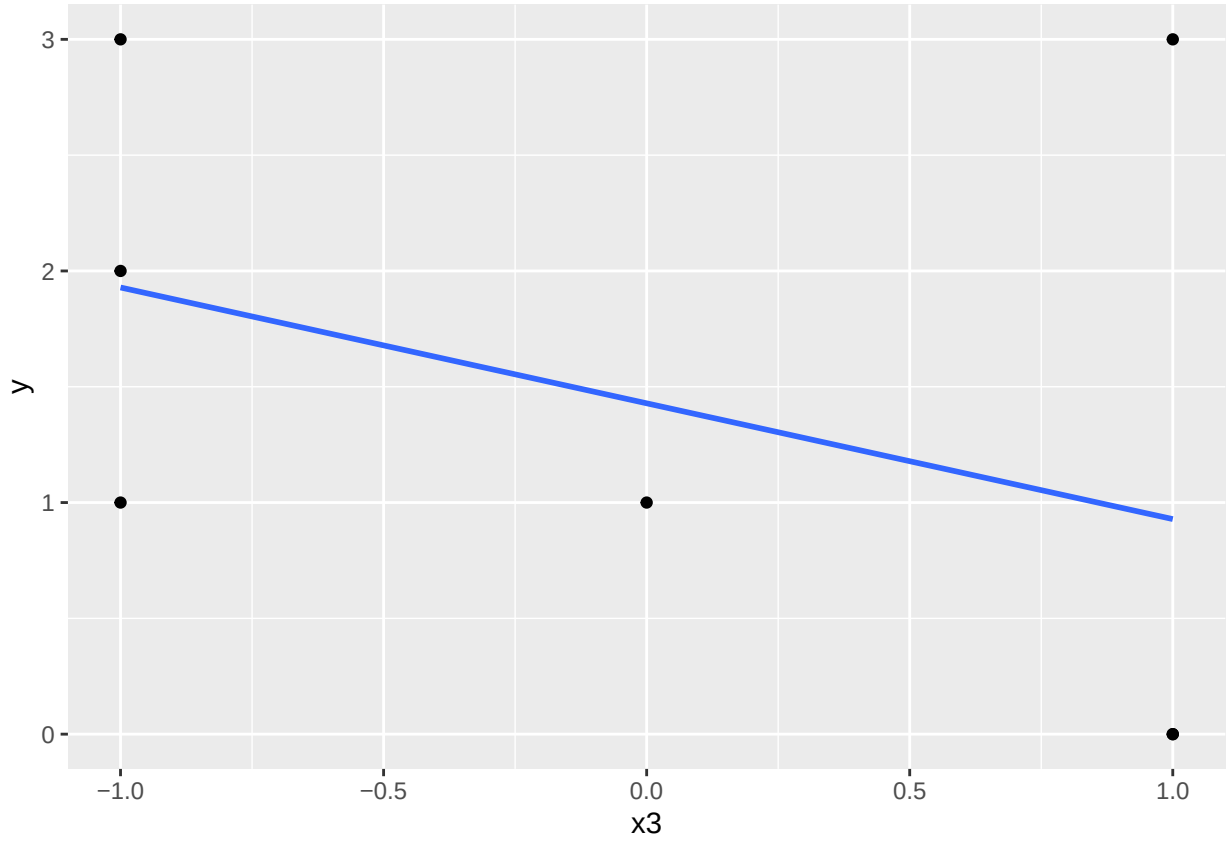

```
ggplot(data = df, aes(x = x2, y = y)) +
  geom_point() +
  stat_smooth(method = "lm", se = FALSE)
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
ggplot(data = df, aes(x = x3, y = y)) +
  geom_point() +
  stat_smooth(method = "lm", se = FALSE)
```

```
## `geom_smooth()` using formula 'y ~ x'
```



Part (b)

```
x_1=1
x_2=-3
x_3=-1
y_hat = 1.42857 + 0.50000 * x_1 + 0.11905 * x_2 -0.50000 * x_3
y_hat
```

```
## [1] 2.07142
```

Part (c)

```
X = matrix(c(c(1,1,1,1,1,1,1),x1,x2,x3), ncol=4, nrow=7)
Y = matrix(y, ncol=1, nrow=7)
B = matrix(c(1.42857, 0.50000, 0.11905, -0.50000), ncol=1, nrow=4)

SSE=(t(Y)%*%Y)-t(X%*%B)%*%Y
print("SSE: ")
```

```
## [1] "SSE: "
```

```
SSE
```

```
##        [,1]
## [1,] 0.0238
```

```
sigma_sq=SSE/(7-3-1)
print("sigma_sq: ")
```

```
## [1] "sigma_sq: "
```

```
sigma_sq
```

```
##              [,1]
## [1,] 0.007933333
```

```
sigma=sqrt(sigma_sq)
print("sigma: ")
```

```
## [1] "sigma: "
```

```
sigma
```

```
##            [,1]
## [1,] 0.08906926
```

```
a=matrix(c(1,1,-3,-1), ncol=1, nrow=4)
x=matrix(c(1,-3,-1), ncol=1, nrow=3)
t_alpha=3.182

fix=(t(a)%*%B)
range=t_alpha*sigma*sqrt(1+t(a)%*%solve(t(X)%*%X)%*%a)
fix
```

```
##         [,1]
## [1,] 2.07142
```

```
range
```

```
##          [,1]
## [1,] 0.341561
```

```
PI=c(fix-range, fix+range)
print("PI: ")
```

```
## [1] "PI: "
```

```
PI
```

```
## [1] 1.729859 2.412981
```