MATH 423 - Assignment 3

Emir Sevinc 260682995

November 23, 2018

```r
library(scatterplot3d)
```

Question 1)

From MSE for the given model is:

$\sum_{i=1}^{n}(Y_i - (X_{1i}\hat{\beta}_1 + X_{2i}\hat{\beta}_2))^2$

$= \sum_{i=1}^{n}(Y_i^2 + (X_{1i}\hat{\beta}_1 + X_{2i}\hat{\beta}_2)^2 - 2Y_i(X_{1i}\hat{\beta}_1 + X_{2i}\hat{\beta}_2))$

$= \sum_{i=1}^{n}(Y_i^2 + X_{1i}^2\hat{\beta}_1^2 + 2\hat{\beta}_1\hat{\beta}_2X_{1i}X_{2i} + X_{2i}^2\hat{\beta}_2^2 - 2Y(X_{1i}\hat{\beta}_1) - 2Y(X_{2i}\hat{\beta}_2))$

$= \sum_{i=1}^{n}Y_i^2 + \sum_{i=1}^{n}X_{1i}^2\hat{\beta}_1^2 + 2\hat{\beta}_1\hat{\beta}_2\sum_{i=1}^{n}X_{1i}X_{2i} + \sum_{i=1}^{n}X_{2i}^2\hat{\beta}_2^2 - \sum_{i=1}^{n}2Y(X_{1i}\hat{\beta}_1) - \sum_{i=1}^{n}2Y(X_{2i}\hat{\beta}_2))$

since $\sum_{i=1}^{n}X_{1i}X_{2i} = 0$ as given by the question, this is eqaul to:

$= \sum_{i=1}^{n}Y_i^2 + \sum_{i=1}^{n}X_{1i}^2\hat{\beta}_1^2 + \sum_{i=1}^{n}X_{2i}^2\hat{\beta}_2^2 - \sum_{i=1}^{n}2Y(X_{1i}\hat{\beta}_1) - \sum_{i=1}^{n}2Y(X_{2i}\hat{\beta}_2))$

$= \sum_{i=1}^{n}(Y_i^2 + X_{1i}^2\hat{\beta}_1^2 + X_{2i}^2\hat{\beta}_2^2 - 2Y(X_{1i}\hat{\beta}_1) - 2Y(X_{2i}\hat{\beta}_2))$

This is equal to: $\sum_{i=1}^{n}(Y_i - X_{1i}\hat{\beta}_1)^2 + X_{2i}^2\hat{\beta}_2^2 - 2Y(X_{2i}\hat{\beta}_2)$. If we differentiate this with respect to $\hat{\beta}_1$, the $\hat{\beta}_2$ terms will disappear and we will be left with $d/d\hat{\beta}_1\sum_{i=1}^{n}(Y_i - X_{1i}\beta_1)^2$; which is the normal equation needed to minimise the simple linear regression MSE for $\hat{\beta}_1$, that is $\sum_{i=1}^{n}(Y_i - X_{1i}\hat{\beta}_1)^2$
Similarly, the expression is also equal to $\sum_{i=1}^{n}(Y_i - X_{2i}\hat{\beta}_2)^2 + X_{1i}^2\hat{\beta}_1^2 - 2Y(X_{1i}\hat{\beta}_1)$
Once again, since differentiating with respect to $\hat{\beta}_2$ cancel all the $\hat{\beta}_1$ terms, this will be equivalent to optimising the simple linear regression MSE for $\hat{\beta}_2$, that is $\sum_{i=1}^{n}(Y_i - X_{2i}\hat{\beta}_2)^2$
This shows that given the condition; optimising the MSE for this model is equivalent to optimising the separate simple linear regression MSE's, so they will provide the same least squares estimators.

Question 2)

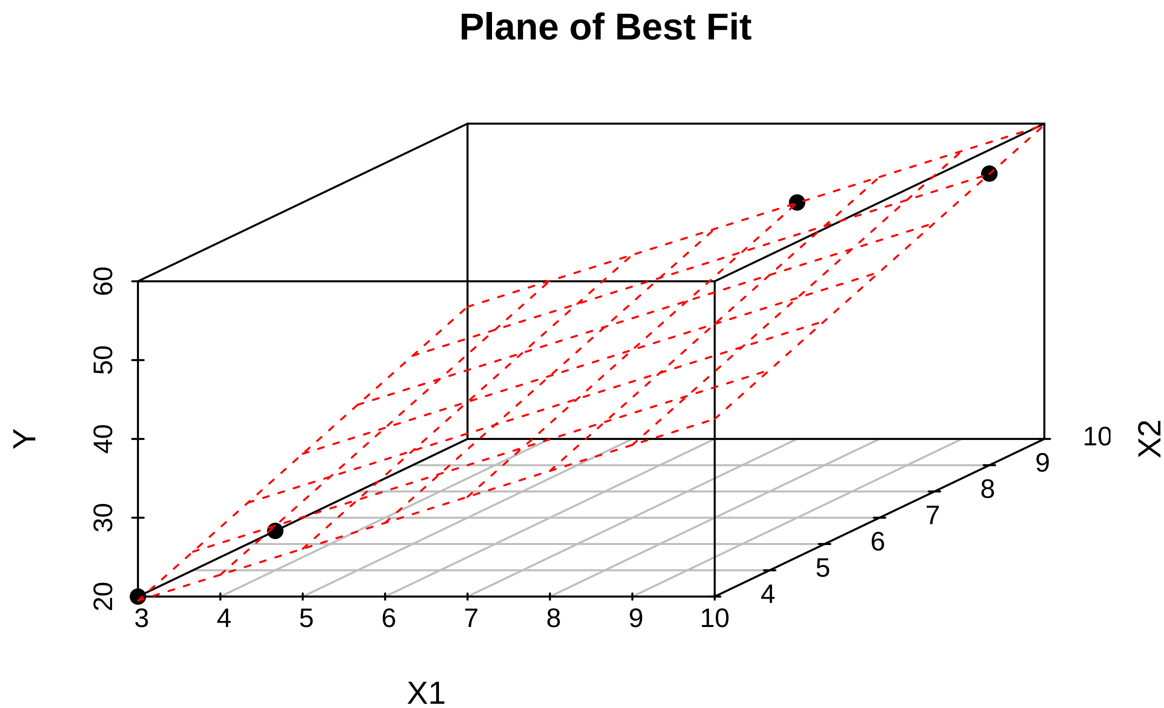
1)

```r
X1 <- c(4,3,10,7)
X2 <- c(5,4,9,10)
Y <- c(25,20,57,50)
fit.Y<-lm(Y~X1+X2)
summary(fit.Y)
```

```
##
## Call:
## lm(formula = Y ~ X1 + X2)
##
## Residuals:
##        1        2        3        4
## -0.64706  0.52941  0.05882  0.05882
```

```
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -1.9412     1.2353  -1.571   0.3608
## X1            3.2941     0.2999  10.983   0.0578 .
## X2            2.8824     0.3222   8.946   0.0709 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8402 on 1 degrees of freedom
## Multiple R-squared:  0.9993, Adjusted R-squared:  0.9979
## F-statistic: 706.4 on 2 and 1 DF,  p-value: 0.0266
```

```
s3d <-scatterplot3d(X1,X2,Y, pch=19, grid=TRUE, main="Plane of Best Fit",angle=30)
s3d$plane3d(fit.Y, col="red")
```



The computed results are:

$\hat{\beta}_0 = -1.9412$, $\hat{\beta}_1 = 3.2941$, and $\hat{\beta}_2 = 2.8824$ and thus the fitted model is $Y = -1.9412 + 3.2941 * X1 + 2.8824 * X2$

2)

```
c1 <- rep(1,4)
X<-cbind(c1,X1,X2)
X
```

```
##      c1 X1 X2
```

```
## [1,]  1  4  5
## [2,]  1  3  4
## [3,]  1 10  9
## [4,]  1  7 10
```

```r
XtX<-t(X) %*% X
XtX
```

```
##    c1  X1  X2
## c1  4  24  28
## X1 24 174 192
## X2 28 192 222
```

```r
XtXinv<-solve(XtX)

XtXinv
```

```
##              c1          X1         X2
## c1  2.16176471  0.05882353 -0.3235294
## X1  0.05882353  0.12745098 -0.1176471
## X2 -0.32352941 -0.11764706  0.1470588
```

3)

```r
B <- XtXinv %*% t(X) %*% Y

B
```

```
##         [,1]
## c1 -1.941176
## X1  3.294118
## X2  2.882353
```

So we computed $\hat{\beta}_0 = -1.941176$, $\hat{\beta}_1 = 3.294118$, and $\hat{\beta}_2 = 2.882353$. Note that these are almost exaclty the same as what we got from the regression we fit earier in part 1, and the minute difference can be attributed to roundoff error.

4)

```r
H <- X %*% B
H
```

```
##          [,1]
## [1,] 25.64706
## [2,] 19.47059
## [3,] 56.94118
## [4,] 49.94118
```

This corresponds to the fitted values of our regression model. Note that we could've gotten the same values with:

```r
fit = fitted(fit.Y)
fit
```

```
##        1        2        3        4
## 25.64706 19.47059 56.94118 49.94118
```

5)

We had:

```
XtXinv
```

```
##             c1          X1         X2
## c1  2.16176471  0.05882353 -0.3235294
## X1  0.05882353  0.12745098 -0.1176471
## X2 -0.32352941 -0.11764706  0.1470588
```

Thus $Var(\hat{\beta}_0) = \sigma^2 * 2.16176471$, $Var(\hat{\beta}_1) = \sigma^2 * 0.12745098$ and $Var(\hat{\beta}_2) = \sigma^2 * 0.1470588$ where σ^2 is the presumed variance of the error term ϵ_i

Question 3)

We have $e^T\hat{Y} = (Y - \hat{Y})^T HY = (Y^T - \hat{Y}^T)HY$
$= (Y^T - (HY)^T)HY$ as $\hat{Y} = HY$
$= (Y^T - Y^T H^T)HY$
$= (Y^T - Y^T H)HY$ as $H^T = H$
$= Y^T HY - Y^T HHY$
$= Y^T HY - Y^T HY$ as $HH = H$
$= 0$ which shows that they are orthogonal.

Question 4)

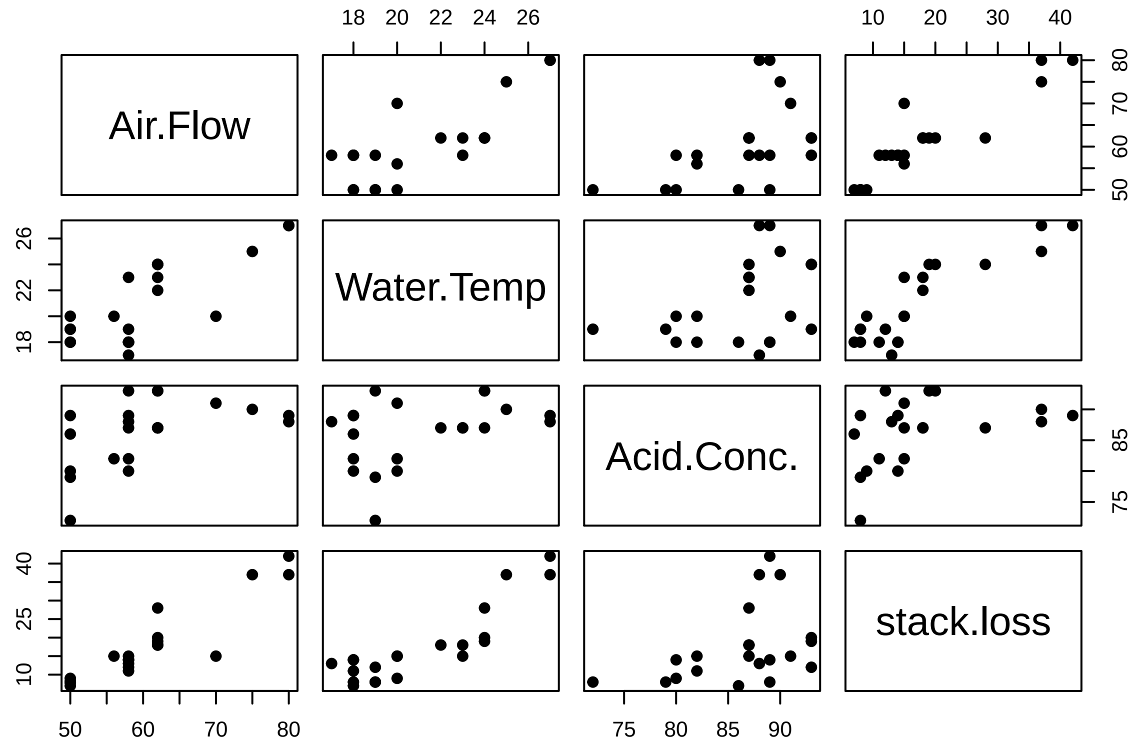
```
data(stackloss)
names(stackloss)
```

```
## [1] "Air.Flow"   "Water.Temp" "Acid.Conc." "stack.loss"
```

```
## [1] "Air.Flow" "Water.Temp" "Acid.Conc." "stack.loss"
```

1)

```
plot(stackloss, pch=19)
```



2)

```r
fit.Stackloss<-lm(stack.loss~Air.Flow+Water.Temp+Acid.Conc.,data=stackloss)
summary(fit.Stackloss)
```

```
##
## Call:
## lm(formula = stack.loss ~ Air.Flow + Water.Temp + Acid.Conc.,
##     data = stackloss)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.2377 -1.7117 -0.4551  2.3614  5.6978
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -39.9197    11.8960  -3.356  0.00375 **
## Air.Flow      0.7156     0.1349   5.307  5.8e-05 ***
## Water.Temp    1.2953     0.3680   3.520  0.00263 **
## Acid.Conc.   -0.1521     0.1563  -0.973  0.34405
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.243 on 17 degrees of freedom
## Multiple R-squared:  0.9136, Adjusted R-squared:  0.8983
## F-statistic:  59.9 on 3 and 17 DF,  p-value: 3.016e-09
```

The computed results are:

$\hat{\beta}_0 = -39.9197$, $\hat{\beta}_1 = 0.7156$, $\hat{\beta}_2 = 1.2953$ and $\hat{\beta}_3 = -0.1521$ thus the fitted model is $Y = -39.9197 + 0.7156 * X1 + 1.2953 * X2 - 0.1521 * X3$

3)

```
confint(fit.Stackloss,level=0.90)
```

```
##                      5 %        95 %
## (Intercept) -60.6140306 -19.2253183
## Air.Flow      0.4810400   0.9502404
## Water.Temp    0.6550686   1.9355036
## Acid.Conc.   -0.4240127   0.1197676
```

Thus the intevrals are: $[-60.6140306, -19.2253183]$ for β_0
$[0.4810400, 0.9502404]$ for β_1
$[0.6550686, 1.9355036]$ for β_2
and $[-0.4240127, 0.1197676]$ for β_3

4)

```
predict(fit.Stackloss, data.frame(Air.Flow=58, Water.Temp=20, Acid.Conc.=86),
        interval = "prediction", conf.level = 0.99)
```

```
##        fit      lwr      upr
## 1 14.41064 7.385265 21.43602
```

Thus the prediction is 14.41064, and the interval is [7.385265,21.43602]

5)

0 is includeed in the confidence interval for β_3 so we can not reject the hypothesis that $\beta_3 = 0$
From the output the test statistic is given as -0.973 for β_3 So:

```
p_val = pt(-0.973, 17, lower.tail=T)+pt(0.973, 17, lower.tail=F)
p_val
```

```
## [1] 0.3441956
```

Thus we find the p value as 0.3441956. At $\alpha = 0.1$, since $0.1 < 0.3441956$, we fail to reject.