stat430 a4 q4

Yiming Shen 20891774

26/11/2023

```
y = c(58,59,47,46,41,62,51,51,51,49,50,50,49,48)
x1 = c(1,1,-1,-1,-sqrt(3),sqrt(3),0,0,0,0,0,0,0,0)
x2 = c(1,-1,1,-1,0,0,-sqrt(3),sqrt(3),0,0,0,0,0,0)
x3 = c(1,-1,-1,1,0,0,0,0,-sqrt(3),sqrt(3),0,0,0,0)
datq4 = data.frame(y,x1,x2,x3)
datq4
```

```
##     y        x1        x2        x3
## 1  58  1.000000  1.000000  1.000000
## 2  59  1.000000 -1.000000 -1.000000
## 3  47 -1.000000  1.000000 -1.000000
## 4  46 -1.000000 -1.000000  1.000000
## 5  41 -1.732051  0.000000  0.000000
## 6  62  1.732051  0.000000  0.000000
## 7  51  0.000000 -1.732051  0.000000
## 8  51  0.000000  1.732051  0.000000
## 9  51  0.000000  0.000000 -1.732051
## 10 49  0.000000  0.000000  1.732051
## 11 50  0.000000  0.000000  0.000000
## 12 50  0.000000  0.000000  0.000000
## 13 49  0.000000  0.000000  0.000000
## 14 48  0.000000  0.000000  0.000000
```

(a)

```
# dot product of x1 and x2
sum(x1 * x2)==0
```

```
## [1] TRUE
```

```
# dot product of x1 and x3
sum(x1 * x3)==0
```

```
## [1] TRUE
```

```
# dot product of x2 and x3
sum(x2 * x3)==0
```

```
## [1] TRUE
```

Comments: column x1 and x2; column x2 and x3 are orthogonal. But column x1 and x3 are not not orthogonal.

(b)

```
plus_a <- sqrt(3)
minus_a <- -sqrt(3)

# temperature
UH <- 250
UL <- 200
plus_U <- plus_a * (UH - UL)/2 + (UH + UL)/2
minus_U <- minus_a * (UH - UL)/2 + (UH + UL)/2
center_U <- (UH + UL)/2
datq4$temperature[datq4$x1==-1] <- UL
datq4$temperature[datq4$x1==1] <- UH
datq4$temperature[datq4$x1==minus_a] <- minus_U
datq4$temperature[datq4$x1==plus_a] <- plus_U
datq4$temperature[datq4$x1==0] <- center_U

# pressure
UH <- 600
UL <- 400
plus_U <- plus_a * (UH - UL)/2 + (UH + UL)/2
minus_U <- minus_a * (UH - UL)/2 + (UH + UL)/2
center_U <- (UH + UL)/2
datq4$pressure[datq4$x2==-1] <- UL
datq4$pressure[datq4$x2==1] <- UH
datq4$pressure[datq4$x2==minus_a] <- minus_U
datq4$pressure[datq4$x2==plus_a] <- plus_U
datq4$pressure[datq4$x2==0] <- center_U

# concentration
UH <- 60
UL <- 50
plus_U <- plus_a * (UH - UL)/2 + (UH + UL)/2
minus_U <- minus_a * (UH - UL)/2 + (UH + UL)/2
center_U <- (UH + UL)/2
datq4$concentration[datq4$x3==-1] <- UL
datq4$concentration[datq4$x3==1] <- UH
datq4$concentration[datq4$x3==minus_a] <- minus_U
datq4$concentration[datq4$x3==plus_a] <- plus_U
datq4$concentration[datq4$x3==0] <- center_U

# final matrix
print(datq4)
```

```
##     y        x1        x2        x3 temperature pressure concentration
## 1  58  1.000000  1.000000  1.000000    250.0000 600.0000      60.00000
## 2  59  1.000000 -1.000000 -1.000000    250.0000 400.0000      50.00000
## 3  47 -1.000000  1.000000 -1.000000    200.0000 600.0000      50.00000
## 4  46 -1.000000 -1.000000  1.000000    200.0000 400.0000      60.00000
## 5  41 -1.732051  0.000000  0.000000    181.6987 500.0000      55.00000
## 6  62  1.732051  0.000000  0.000000    268.3013 500.0000      55.00000
## 7  51  0.000000 -1.732051  0.000000    225.0000 326.7949      55.00000
## 8  51  0.000000  1.732051  0.000000    225.0000 673.2051      55.00000
## 9  51  0.000000  0.000000 -1.732051    225.0000 500.0000      46.33975
```

```
## 10 49  0.000000  0.000000  1.732051    225.0000 500.0000      63.66025
## 11 50  0.000000  0.000000  0.000000    225.0000 500.0000      55.00000
## 12 50  0.000000  0.000000  0.000000    225.0000 500.0000      55.00000
## 13 49  0.000000  0.000000  0.000000    225.0000 500.0000      55.00000
## 14 48  0.000000  0.000000  0.000000    225.0000 500.0000      55.00000
```

(c)

```
# fit a second order linear model
mod <- lm(y ~ x1 * x2 * x3 + I(x1^2) + I(x2^2) + I(x3^2), data = datq4)
summary(mod)
```

```
##
## Call:
## lm(formula = y ~ x1 * x2 * x3 + I(x1^2) + I(x2^2) + I(x3^2),
##     data = datq4)
##
## Residuals:
##          1          2          3          4          5          6          7
##  5.551e-17  3.946e-17  7.259e-17  2.972e-19 -8.196e-18 -3.114e-17  9.045e-18
##          8          9         10         11         12         13         14
## -2.108e-17 -5.663e-17 -4.736e-17  7.500e-01  7.500e-01 -2.500e-01 -1.250e+00
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.925e+01  4.787e-01 102.880 2.02e-06 ***
## x1           6.062e+00  3.909e-01  15.510 0.000582 ***
## x2          -8.620e-17  3.909e-01   0.000 1.000000
## x3          -5.774e-01  3.909e-01  -1.477 0.236154
## I(x1^2)      7.500e-01  2.764e-01   2.714 0.072942 .
## I(x2^2)      5.833e-01  2.764e-01   2.111 0.125298
## I(x3^2)      2.500e-01  2.764e-01   0.905 0.432389
## x1:x2        7.735e-02  6.180e-01   0.125 0.908314
## x1:x3        3.761e-15  6.180e-01   0.000 1.000000
## x2:x3       -6.218e-02  6.180e-01  -0.101 0.926208
## x1:x2:x3     1.667e+00  6.180e-01   2.697 0.073986 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9574 on 3 degrees of freedom
## Multiple R-squared:  0.993,  Adjusted R-squared:  0.9697
## F-statistic: 42.65 on 10 and 3 DF,  p-value: 0.00518
```

```
# when temp change from U (natural unit) to U+1, the change of x (coded unit) is:
delta_x <- (1-225)/25 - (-225/25)
delta_x
```

```
## [1] 0.04
```

```
# which means from x (coded unit) to x+0.04

# the change of E[Y] will be delta_x * (estimate of beta1)
delta_x * summary(mod)$coefficients[2]
```

```
## [1] 0.2424871
```

Comments: when the temperature increases 1C, the expected yield will increase 0.2424871.