Assignment 2

2022-10-22

Question 1

a)

The test statistic for the hypothesis is

$$T = \frac{\hat{\theta}_3 - 2}{\sqrt{\hat{\Sigma}_{3,3}}} = \frac{4-2}{\sqrt{1.1}} = 1.9069252.$$

Since $1.9069252 > t_{97,0.95} = 1.6607146$ so we reject H_0.

b)

Let $\alpha = 0.05$, the 95% confidence interval is

$$\hat{\theta}_3 \pm t_{97,0.975}\sqrt{\hat{\Sigma}_{3,3}} = 4 \pm 1.98 \times 1.04 = [1.94, 6.05].$$

c)

We have the statistic

$$T = \frac{f(0,\hat{\theta})}{\sqrt{\hat{v}_x^T \hat{\Sigma} \hat{v}}} = \frac{0.54}{0.85} = 0.63$$

which is lower than $t_{97,0.975}$ so we can not reject H_0.

d)

The 95% confidence interval is

$$f(0,\hat{\theta}) \pm t_{97,0.975} \times \sqrt{\hat{v}_x^T \hat{\Sigma} \hat{v}} = [-1.14, 2.22].$$

e)

Since the dataset with the funtion $f(x,\theta)$, we can compute the matrix $\hat{V}$ where

$$\hat{V}_{ij} = \partial f(x_i, \hat{\theta}) / \partial \theta_j.$$

Thus, we can compute $(\hat{V}^T \hat{V})^{-1}$ and

$$\hat{\sigma}^2 = \hat{\Sigma}(\hat{V}^T \hat{V})$$

where $\hat{\Sigma}$ is the given estimated covariance matrix.

Question 2

a)

The estimates of θ are $\hat{\theta} = (0.81, -0.44, 1.98, 1.27)$ and $\hat{\sigma}^2 = \frac{S(\hat{\theta})}{n-p}$ where $S(\hat{\theta}) = RSE^2(n-p)$ where RSE is the residual standard error. Thus, we have $\hat{\sigma}^2 = RSE^2 = 0.275$. The residual sum of squares can be recovered from the residual standard error and the degree of freedoms because

$$RSE = \sqrt{\frac{RSS}{n-p}}$$

where p is the number of parameters θ, which in this case, is 4.

b)

We first obtain the estimated covariance matrix or just the values on the diagonal of $\hat{\Sigma}$ using $x_i = \frac{3(i-1)}{n-1}$ with $n = 100$ and the partial derivatives $\partial f(x_i, \hat{\theta})/\partial \theta_j$. Note that the partial derivatives calculations are quite straightforward so they will be omitted. We have $\hat{\Sigma}_{1,1} = 0.114$, $\hat{\Sigma}_{2,2} = 0.028$, $\hat{\Sigma}_{3,3} = 0.0076$, and $\hat{\Sigma}_{4,4} = 0.1104$. Like before, we have the test statistic

$$T = \frac{0.8}{\sqrt{0.114}} = 2.37 > t_{96,1-0.05/2} = 1.984$$

so we reject H_0. The 95% confidence interval is

$$\hat{\theta_1} \pm t_{96,1-0.05/2}\sqrt{\hat{\Sigma}_{1,1}} = 0.81 \pm 1.984 \times 0.337 = [0.141, 1.478].$$

c)

Likewise, we have the statistic

$$T = \frac{\hat{\theta}_4 - 1}{\sqrt{\hat{\Sigma}_{4,4}}} = 0.81 < t_{96,1-0.05/2}$$

so we can not reject H_0. Let $\alpha = 0.02$ then the 98% confidence interval is

$$\hat{\theta_2} \pm t_{96,1-0.02/2}\sqrt{\hat{\Sigma}_{2,2}} = -0.44 \pm 2.36 \times 0.028 = [-0.5, -0.37].$$

d)

e)

Let the global model Ω be the full model, i.e, $S(\hat{\theta_q}) = \min_{\theta_q} ||Y - f_\Omega(x, \theta_q)||^2$, where $\theta_q \in \mathbb{R}^q$, and the linear submodel to be $\omega : S(\hat{\theta_p}) = \min_{\theta_p} ||Y - f_\omega(x, \theta_p)||^2$ where $\theta_p \in \mathbb{R}^p$ with $p < q$ such that $f_\omega(x, \theta_p)$ is linear. We can test for H_0 that the linear submodel fits well at significance level α by using the test statistic

$$V = \frac{[S(\hat{\theta_p}) - S(\hat{\theta_q})]/(q-p)}{S(\hat{\theta_q})/(n-q)}.$$

If $V > F_{q-p,n-q,1-\alpha}$ then we reject H_0 that the linear submodel fits well.

Question 3

a)

Question 4

a)

```r
stormer <- data.frame(stormer);
smod_lin <- lm(Wt * Time ~ Viscosity + Time, data = stormer);
summary(smod_lin)
```

```
##
## Call:
## lm(formula = Wt * Time ~ Viscosity + Time, data = stormer)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -330.7 -153.8    4.7  170.7  368.3
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 220.1381    82.0080   2.684  0.01426 *
## Viscosity    28.0987     0.5663  49.620  < 2e-16 ***
## Time          2.0818     0.7302   2.851  0.00987 **
## [...]
```

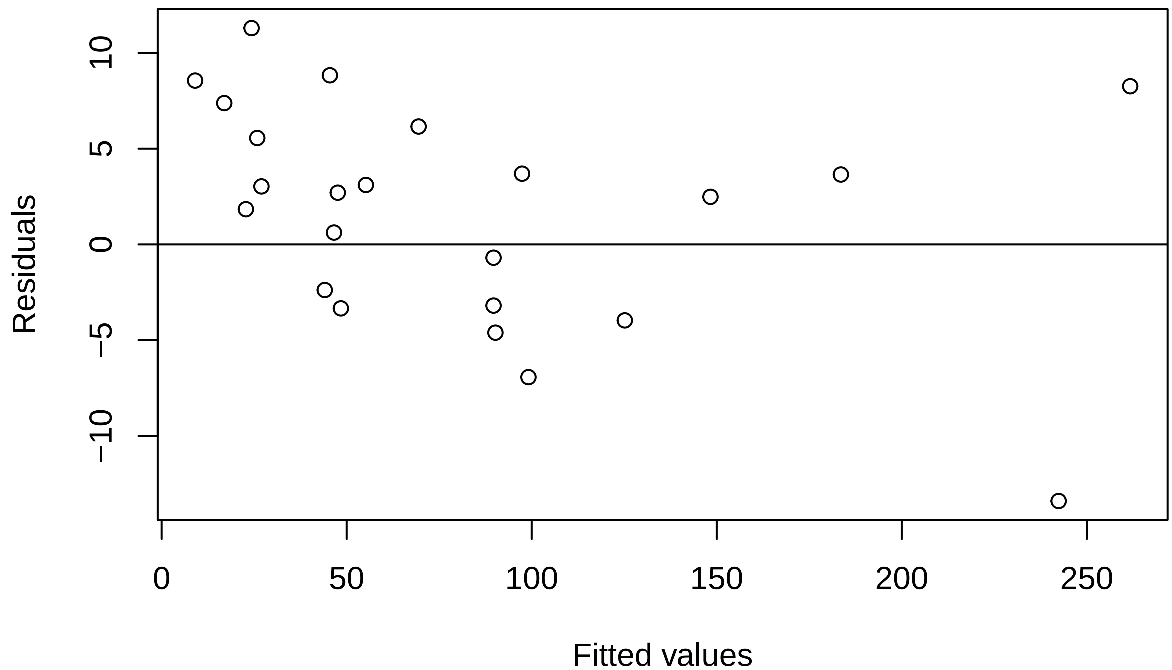
Fitting the linear regression $wT = \theta_1 v + \theta_2 T + (w - \theta_2)\varepsilon$ returns the estimates $\hat{\theta}_1 = 28$ and $\hat{\theta}_2 = 2$ which we will use for the initial values of the non-linear regression.

```r
n = nrow(stormer);
p = 2;
smod.nls <- nls(Time ~ (the1 * Viscosity)/(Wt - the2),
               data = stormer,
               start=c(the1=28, the2=2)
               );
RSS=deviance(smod.nls);
smod.evar <- round(RSS/(n - p), 2);
summary(smod.nls)
```

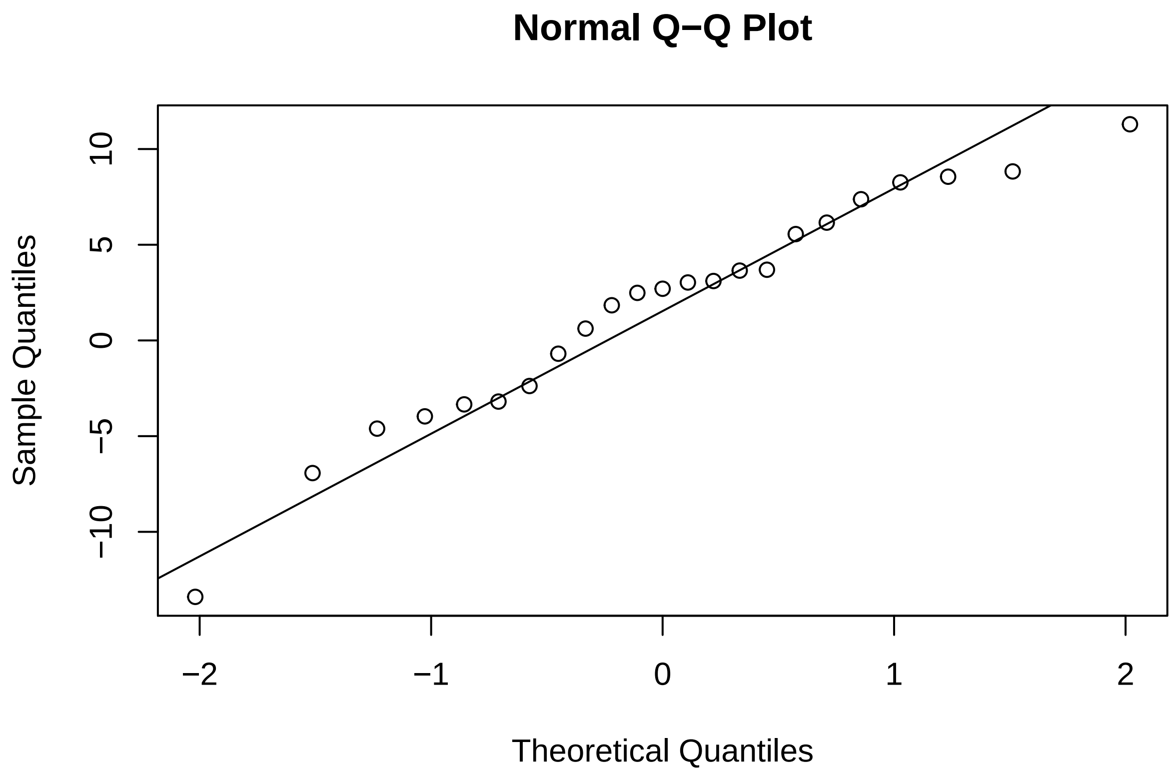
```
##
## Formula: Time ~ (the1 * Viscosity)/(Wt - the2)
##
## Parameters:
##      Estimate Std. Error t value Pr(>|t|)
## the1  29.4013     0.9155  32.114  < 2e-16 ***
## the2   2.2183     0.6655   3.333  0.00316 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## [...]
```

Applying non-linear regression, we have the estimates for (θ_1, θ_2) as $(\hat{\theta}_1, \hat{\theta}_2) = (29.4, 2.2)$ which are very close to the inital values.The estimated variance of the error $\hat{\sigma}^2 = 39.29$. To test for the validity of the model's assumptions, we use residual plot and the qq plot, Shapiro-Wilk test to check for the normality of the residual.

```r
plot(fitted(smod.nls), resid(smod.nls), xlab="Fitted values", ylab="Residuals");
abline(0, 0)
```




```
qqnorm(resid(smod.nls))
qqline(resid(smod.nls))
```




```
shapiro.test(resid(smod.nls))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  resid(smod.nls)
## W = 0.96402, p-value = 0.5491
```

It can be seen that the pattern of the points are somewhat random so the non-linear model is somewhat good. Since the p-value of the Shapiro-Wilk test is 0.55, which is higher than 0.05, we reject the null hypothesis that the data is not normally distributed. The QQ-plot also confirms the normality assumption.

b)

Calculate the test statistic T using the covariance matrix, we have that $T = 0.3005162 < t_{21,1-0.05/2} = 2.08$ so we can not reject H_0.

c)

```
lb=numeric(2); ub=numeric(2);
for(i in 1:2) {lb[i]=coef(smod.nls)[i]-qt(0.975,n-length(coef(smod.nls)))*sqrt(smod.cov[i,i])
      ub[i]=coef(smod.nls)[i]+qt(0.975,n-length(coef(smod.nls)))*sqrt(smod.cov[i,i])}
ci=cbind(lb,ub); rownames(ci)=names(coef(smod.nls)); ci
```

```
##             lb        ub
## the1 27.497301 31.305213
## the2  0.834246  3.602302
```

The 95% confidence interval for $\hat{\theta}_1$ is $[27.49, 31.3]$ and $\hat{\theta}_2$ is $[0.83, 3.6]$.

d)

```
grad<-function(v,w,the){rbind(v/(w - the[2]), the[1]*v/(w - the[2])^2)};
gradvec <- grad(100, 60, coef(smod.nls));

se=sqrt(t(gradvec)%*%vcov(smod.nls)%*%gradvec);
f <- function(v, w, the) { the[1]*v/(w - the[2]) };
f4 <- f(100, 60, coef(smod.nls))

lb=f4-qt(0.05/2,n-length(coef(smod.nls)), lower.tail=FALSE)*se
ub=f4+qt(0.05/2,n-length(coef(smod.nls)), lower.tail=FALSE)*se

c(lb, ub)
```

```
## [1] 48.65760 53.10902
```

The 95% confidence interval for the expected value of T is therefore $[48.65, 53.1]$.

e)

```
form2 <- as.formula(Time ~ (the1 * Viscosity)/(Wt));
smod.nls2 <- nls(form2, data=stormer, start=c(the1=28));
anova(smod.nls, smod.nls2)
```

```
## Analysis of Variance Table
##
## Model 1: Time ~ (the1 * Viscosity)/(Wt - the2)
## Model 2: Time ~ (the1 * Viscosity)/(Wt)
##   Res.Df Res.Sum Sq Df  Sum Sq F value  Pr(>F)
## 1     21     825.05
## 2     22    1210.38 -1 -385.33  9.8078 0.00504 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Using ANOVA, the reduced model gives a worse fit than the full one since its residual sum of squares is $1210 > 825.05$. We also calculate the V statistic to compare against the F-distribution.

```
SSq <- deviance(smod.nls);
SSp <- deviance(smod.nls2);

n <- length(resid(smod.nls));
q<- length(coef(smod.nls));
p <- length(coef(smod.nls2))

fstat <- ((SSp-SSq)/(q-p))/(SSq/(n-q));
pval <- 1 - pf(fstat, q-p, n-p);
```

The obtained F-statistic is $9.8078075 > F_{1,21,0.95} = 4.32$
and its p-value is $0.004851 < 0.05$ so we reject the null hypothesis H_0 that the smaller model ω is appropriate.

```
AIC(smod.nls)
```

```
## [1] 153.6101
```

```
AIC(smod.nls2)
```

```
## [1] 160.4247
```

The AIC of the full model is 153.6 while the AIC for the smaller one is 160.4. Clearly, $153 < 160$ so the full model fits better.