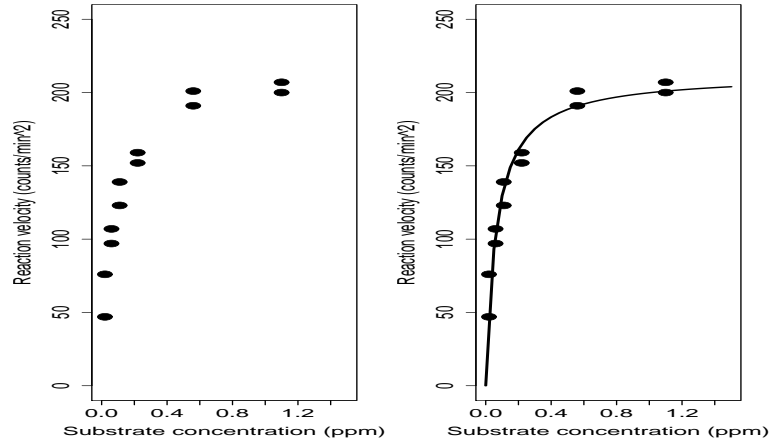


1. Use data set taken from table published in *Nonlinear Regression Analysis and its Applications* by Bates and Watts.

(a) My “eye-estimates” of the parameters based on the plot below are: $\theta_1^0 = 200$ and $\theta_2^0 = 0.1$.

```
> puromycin <- read.table("hw06.data1.txt",header=T) # It has three columns: y, x and z
> d <- puromycin[puromycin$z==0,]
> plot(d$x,d$y,xlab="Substrate concentration (ppm)",ylab="Reaction velocity (counts/min^2)",
      xlim=c(0,1.6),ylim=c(0,250),pch=16)
```



(b) The least squares estimate of the parameter vector $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2) = (212.684, 0.064)$. The “deviance” (error sum of squares) is 1195.449.

```
> library(nls)
> REACT.fm <- nls(y~theta1*x/(theta2+x),data=d,start=c(theta1=200,theta2=0.1))
> REACT.fm
Nonlinear regression model
  model: y ~ theta1 * x/(theta2 + x)
  data: d
      theta1      theta2
212.6836297  0.0641211
residual sum-of-squares: 1195.449
```

(c) After the last two commands in (a).

```
> conc <- seq(0,1.5,.05)
> theta <- coef(REACT.fm)
> velocity <- theta[1]*conc/(theta[2]+conc)
> lines(conc,velocity,lwd=3)
```

(d) i. > summary(REACT.fm)

```
Formula: y ~ theta1 * x/(theta2 + x)
Parameters:
      Estimate Std. Error t value Pr(>|t|)
theta1 2.127e+02  6.947e+00  30.615 3.24e-11 ***
theta2 6.412e-02  8.281e-03   7.743 1.57e-05 ***
```

Residual standard error: 10.93 on 10 degrees of freedom

Correlation of Parameter Estimates:

```
      theta1
theta2 0.7651
```

```
> round(vcov(REACT.fm),6)
      theta1  theta2
theta1 48.262879 0.044014
theta2  0.044014 0.000069
```

$$\text{ii. } D = \begin{bmatrix} \frac{x}{\theta_2+x} & -\frac{\theta_1 x}{(\theta_2+x)^2} \end{bmatrix}$$

```
> D <- cbind(d$x/(theta[2]+d$x), -theta[1]*d$x/(theta[2]+d$x)^2)
> MSE <- summary(REACT.fm)$sigma^2
> round(MSE*solve(t(D)%*%D),6) # Matrix obtained with vcov(REACT.fm)
[1,] 48.262879 0.044014
[2,] 0.044014 0.000069
> round(sqrt(diag(MSE*solve(t(D)%*%D))),6) # Std errors in summary(REACT.fm)
[1] 6.947149 0.008281
```

iii. An approximate 95% prediction interval for one additional reaction velocity, for substrate concentration .50 ppm is (162.24, 214.78) counts/min².

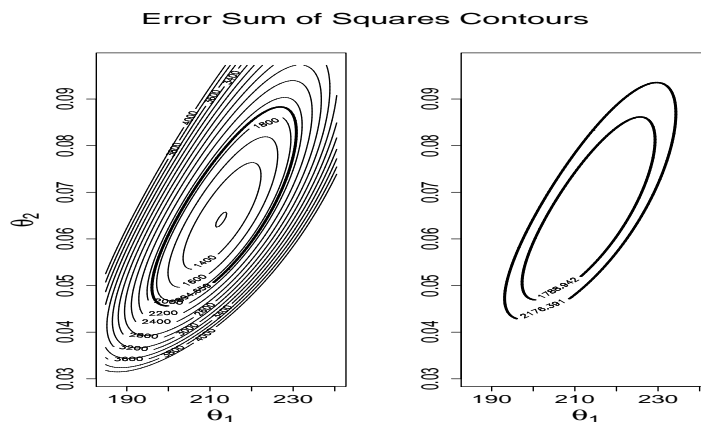
```
> x.new <- 0.5
> yhat <- theta[1]*x.new/(theta[2]+x.new)
> ll <- yhat - qt(0.975,12-2)*sqrt(MSE)*sqrt(1+Ghat%*%solve(t(D)%*%D)%*%Ghat)
> ul <- yhat + qt(0.975,12-2)*sqrt(MSE)*sqrt(1+Ghat%*%solve(t(D)%*%D)%*%Ghat)
```

(e) A point estimate of x_{100} is 0.0569 ppm and a standard error of this estimate is 0.0052 ppm.

```
> xhat <- theta[2]*100/(theta[1]-100)
> Ghat <- c(-theta[2]*100/(theta[1]-100)^2, 100/(theta[1]-100))
> sqrt(Ghat%*%vcov(REACT.fm)%*%Ghat)
```

(f) Beale 90% confidence region for the parameter vector $\underline{\theta}$ includes all pairs (θ_1, θ_2) with sum of squares less than 1895.

```
> ss(theta)
[1] 1195.449 # Deviance obtained in (b)
> plot(th1,th2,type="n",main="Error Sum of Squares Contours")
> contour(th1,th2,SumofSquares,levels=c(seq(1000,4000,200)))
> dv*(1+(2/10)*qf(.90,2,10)) # For 90% confidence region for theta
[1] 1894.659
> contour(th1,th2,SumofSquares,levels=dv*(1+(2/10)*qf(.90,2,10)),add=T,lwd=3)
```



(g) A 95% confidence region for $\underline{\theta}$ includes all pairs (θ_1, θ_2) with sum of squares less than 2176. From the contour at 1789, “eye-estimates” of individual 95% confidence intervals for θ_1 and θ_2 are (197, 229) and (0.047, 0.085), respectively.

(h) Approximate 95% confidence intervals for θ_1 and θ_2 are (197.20, 228.16) and (0.046, 0.082), respectively. These estimates are close to those “guessed” from the graph.

```
> c(theta[1]+qt(0.025,10)*se[1], theta[1]+qt(0.975,10)*se[1]) # 95% c.i. for theta1
> c(theta[2]+qt(0.025,10)*se[2], theta[2]+qt(0.975,10)*se[2]) # 95% c.i. for theta2
```

```
(i) > ll <- sqrt(ss(theta)/qchisq(.975,10))
> ul <- sqrt(ss(theta)/qchisq(.025,10))
> c(ll,ul)
[1] 7.639533 19.187844 # 95% c.i. for sigma based on linear model result
```

```

> sigma2 <- seq(20,250,.01); n <- length(d$x)
> ll <- min(sigma2[(n/2)*log(sigma2)+dv/(2*sigma2)<=(n/2)*log(dv/n)+(n/2)+(1/2)*qchisq(.95,1)])
> ul <- max(sigma2[(n/2)*log(sigma2)+dv/(2*sigma2)<=(n/2)*log(dv/n)+(n/2)+(1/2)*qchisq(.95,1)])
> sqrt(c(ll,ul))
[1] 7.012132 15.811388 # 95% c.i. for sigma based on profile likelihood

```

(j) Intervals are similar to those obtained in (g).

```

> confint(REACT.fm,level=.95)
                2.5%          97.5%
theta1 197.30205011 229.29022954
theta2  0.04692625  0.08616203

```

(k) `> fit1 <- nls(y~(theta1+theta3*z)*x/(theta2+x),data=puromycin, start=c(theta1=213,theta2=.064,theta3=0))`

```

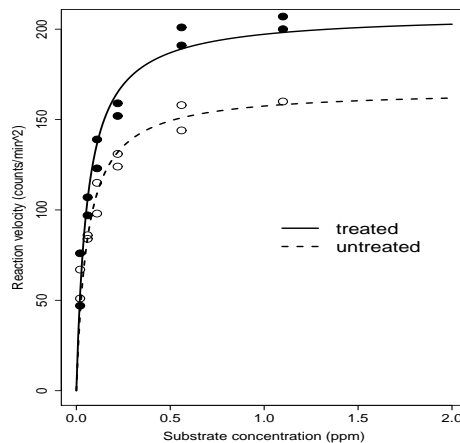
> fit1
Nonlinear regression model
model: y ~ (theta1 + theta3 * z) * x / (theta2 + x)
data: puromycin
      theta1      theta2      theta3
208.63012476  0.05797191 -42.02599166
residual sum-of-squares: 2240.891

```

```

> theta <- coef(fit1)
> se <- sqrt(diag(vcov(fit1)))
> c(theta[3]+qt(0.025,20)*se[3],theta[3]+qt(0.975,20)*se[3])
-55.10945 -28.94253 # 95% c.i. for theta3
> confint(fit1,level=.95)
                2.5%          97.5%
theta1 196.39422 221.50968754
theta2  0.04599  0.07234383
theta3 -55.19946 -28.95656324 # 95% c.i. for theta3

```



The negative sign says that the response y increases when the explanatory variable z decreases (from 1 to 0). We are 95% confident that, for the same concentration, the expected reaction velocity will increase between 29 and 55 counts/min² if enzymes are treated. The effect of the treatment is statistically significant.

(l) (Because I like to have “fun”)

```
th.cr <- th[SumofSquares<=dv*(1+(2/10)*qf(.90,2,10)),] # (theta1,theta2) inside c.r.
x <- seq(0,2,.1) # x-values covering (0,2)
xy <- NULL
for (i in 1:length(x)) {
  for (j in 1:dim(th.cr)[1]) {
    xy <- rbind(xy,c(x[i],th.cr[j,1]*x[i]/(th.cr[j,2]+x[i]))) # y's for thetas inside c.r.
  }
}
y.ll <- NULL; y.ul <- NULL
for (i in 1:length(x)) {
  tmp <- xy[xy[,1]==x[i],]
  y.ll[i] <- min(tmp[,2]) # lower limit of confidence band
  y.ul[i] <- max(tmp[,2]) # upper limit
}
plot(x,y.ul,type="l")
lines(x,y.ll)
```

