

# Week11

## LMR - Problem 8.2

Using the divusa data, fit a regression model with divorce as the response and unemployed, femlab, marriage, birth and military as predictors.

a). Make two graphical checks for correlated errors. What do you conclude?

Here, we see that successive errors are correlated as noted in the correlation (0.85) between the vector of residuals with the first and last term omitted. From plots below, we see that the residual plot shows some sort of cyclical behavior.

b). Allow for serial correlation with an AR(1) model for the errors. (Hint: Use maximum likelihood to estimate the parameters in the GLS fit by `gls(..., method = "ML", ...)`). What is the estimated correlation and is it significant? Does the GLS model change which variables are found to be significant?

The estimated correlation is 0.85 and the Durbin-Watson test indicates evidence of serious positive autocorrelation with the actual value of the DW statistic  $0.37429 < \text{critical value range of } 1.36 - 1.62$ .

c). Speculate why there might be correlation in the errors.

The data is collected over time and hence there is a probability that the errors will be correlated.

```
library(faraway)
library(nlme)
library(lmtest)
```

```
## Loading required package: zoo
```

```
##
```

```
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      as.Date, as.Date.numeric
```

```
lmmod <- lm(divorce ~ unemployed + femlab + marriage + birth + military, divusa)
summary(lmmod)
```

```
##
```

```
## Call:
```

```
## lm(formula = divorce ~ unemployed + femlab + marriage + birth +
##      military, data = divusa)
```

```
##
```

```
## Residuals:
```

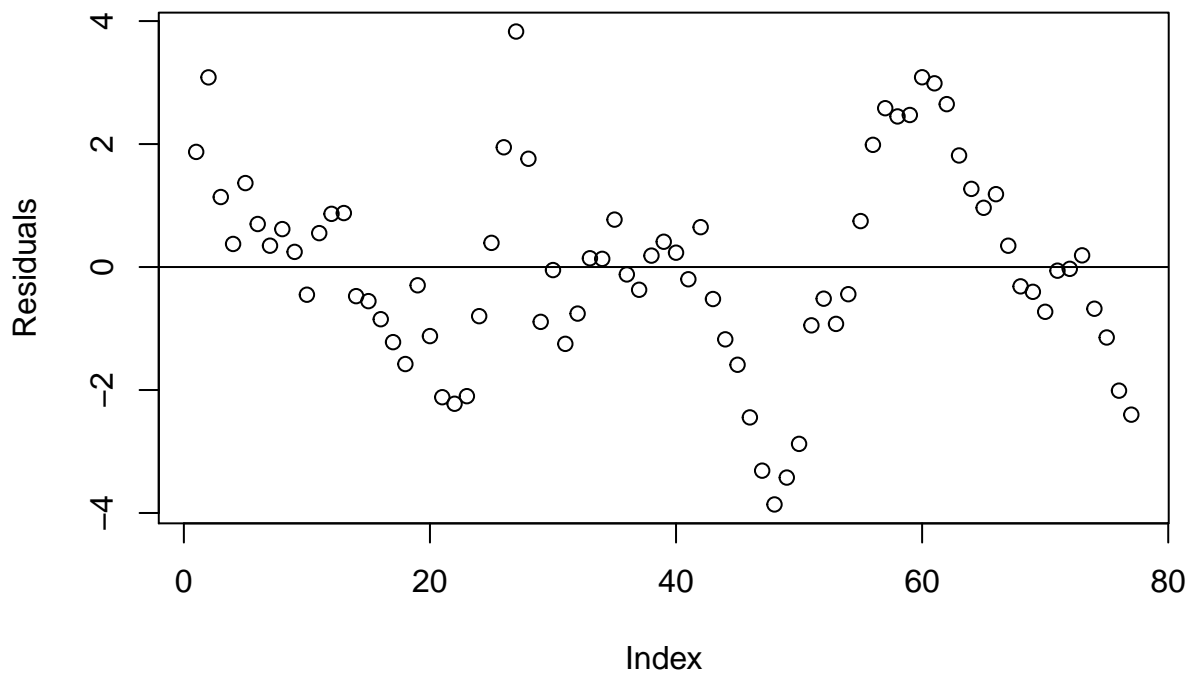
```
##      Min       1Q   Median       3Q      Max
```

```
## -3.8611 -0.8916 -0.0496  0.8650  3.8300
```

```
##
```

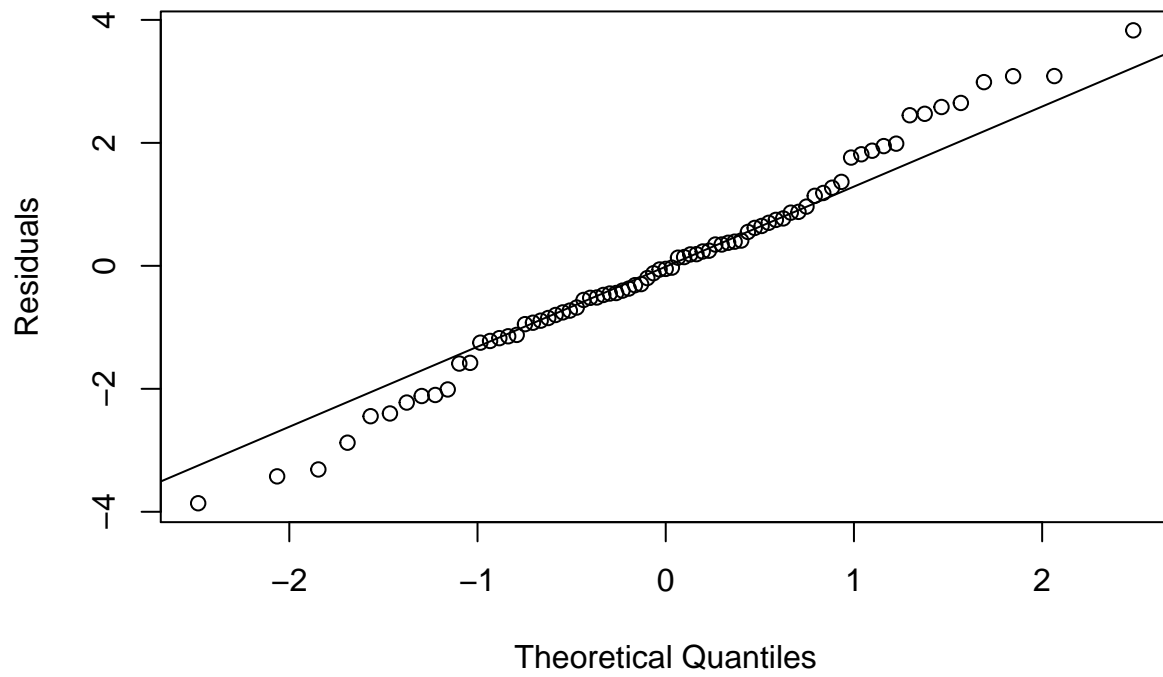
```
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.48784    3.39378   0.733  0.4659
## unemployed  -0.11125    0.05592  -1.989  0.0505 .
## femlab       0.38365    0.03059  12.543 < 2e-16 ***
## marriage     0.11867    0.02441   4.861 6.77e-06 ***
## birth       -0.12996    0.01560  -8.333 4.03e-12 ***
## military    -0.02673    0.01425  -1.876  0.0647 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.65 on 71 degrees of freedom
## Multiple R-squared:  0.9208, Adjusted R-squared:  0.9152
## F-statistic: 165.1 on 5 and 71 DF,  p-value: < 2.2e-16
```

```
plot(residuals(lmod),ylab = "Residuals")
abline(h=0)
```



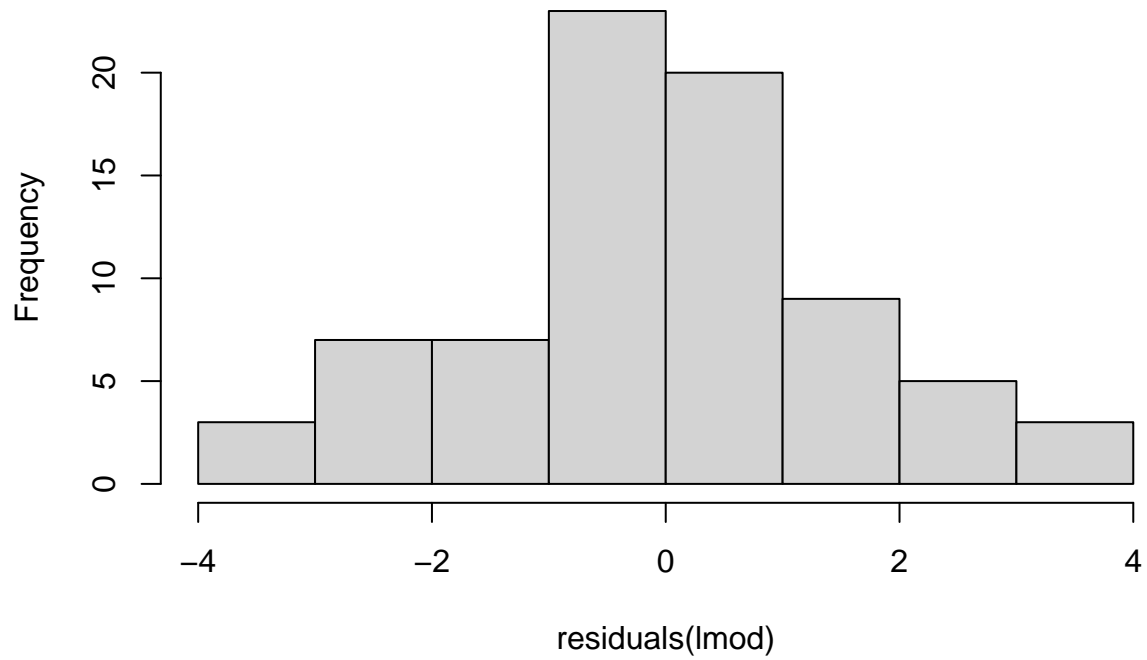
```
qqnorm(residuals(lmod),ylab = "Residuals")
qqline(residuals(lmod))
```

**Normal Q-Q Plot**

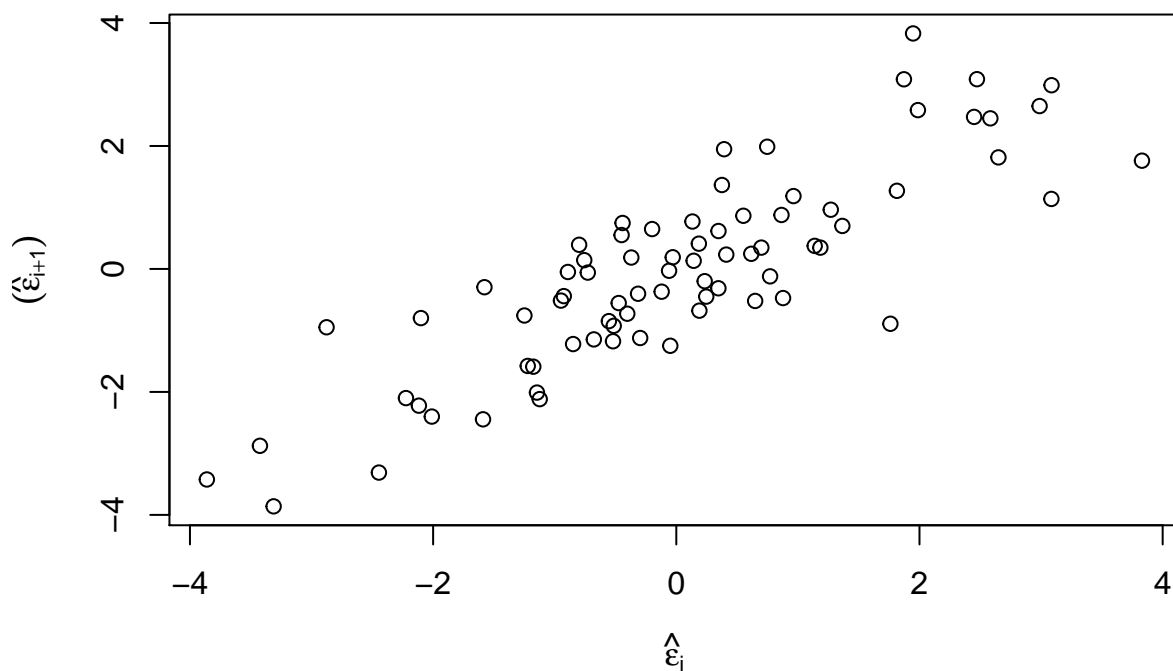


```
hist(residuals(lmod))
```

**Histogram of residuals(lmod)**



```
plot(residuals(lmod)[-77],residuals(lmod)[-1],xlab=expression(hat(epsilon)[i]),ylab = expression((hat(epsilon)[i+1]))
```



```
cor(residuals(lmod)[-1], residuals(lmod)[-length(residuals(lmod))])
```

```
## [1] 0.8469792
```

```
# Regression of successive residuals
```

```
lmod1 <- summary(lm(residuals(lmod)[-1] ~ 0-1+residuals(lmod)[-77]))
summary(lmod1)
```

```
##           Length Class  Mode
## call           2    -none- call
## terms          3     terms call
## residuals      76    -none- numeric
## coefficients    4    -none- numeric
## aliased         1    -none- logical
## sigma          1    -none- numeric
## df              3    -none- numeric
## r.squared       1    -none- numeric
## adj.r.squared   1    -none- numeric
## fstatistic      3    -none- numeric
## cov.unscaled    1    -none- numeric
```

```
dwtest(divorce ~., data = divusa)
```

```
##
## Durbin-Watson test
##
```

```
## data:  divorce ~ .  
## DW = 0.37429, p-value < 2.2e-16  
## alternative hypothesis: true autocorrelation is greater than 0
```