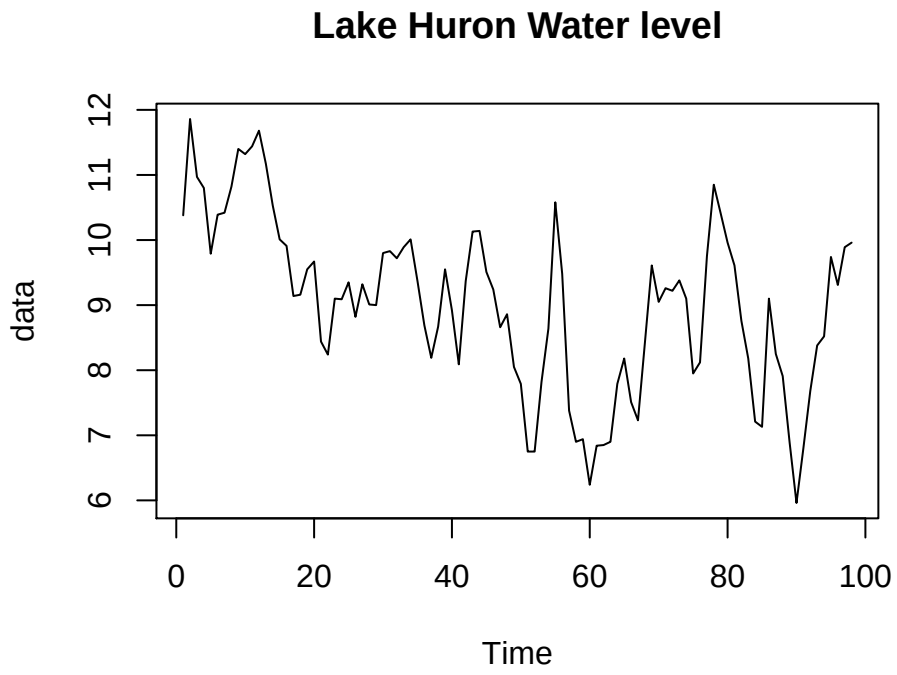
Lake Huron anlysis example

Here we will see whether Lake Huron data need to detrend, and the residuals are IID.

```
# set your own working directory, for example,
# setwd("/Users/accountname/program/CH1-program")
rm(list=ls(all=TRUE))
source("TS-library.R")
data = scan("huron.txt")
```

First step is to plot the data.

```
plot.ts(data);
title("Lake Huron Water level")
```



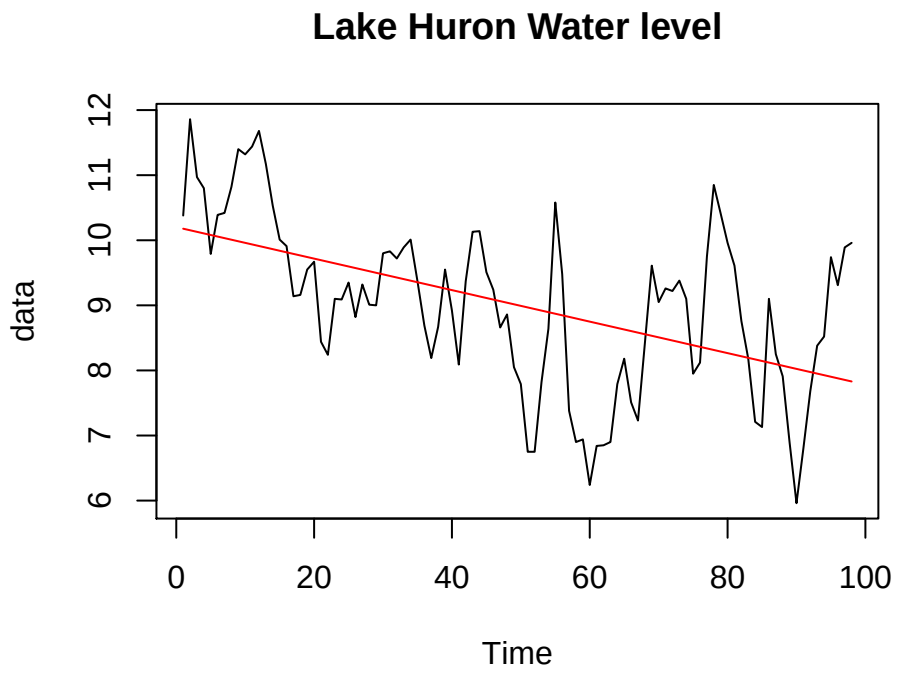
We can observe linear (or quadratic) trend. We will apply simple linear regression.

```
n = length(data);
x = seq(from=1, to = n, by=1);
out.lm = lm(data ~ 1 + x);
summary(out.lm)
```

```
##
## Call:
## lm(formula = data ~ 1 + x)
```

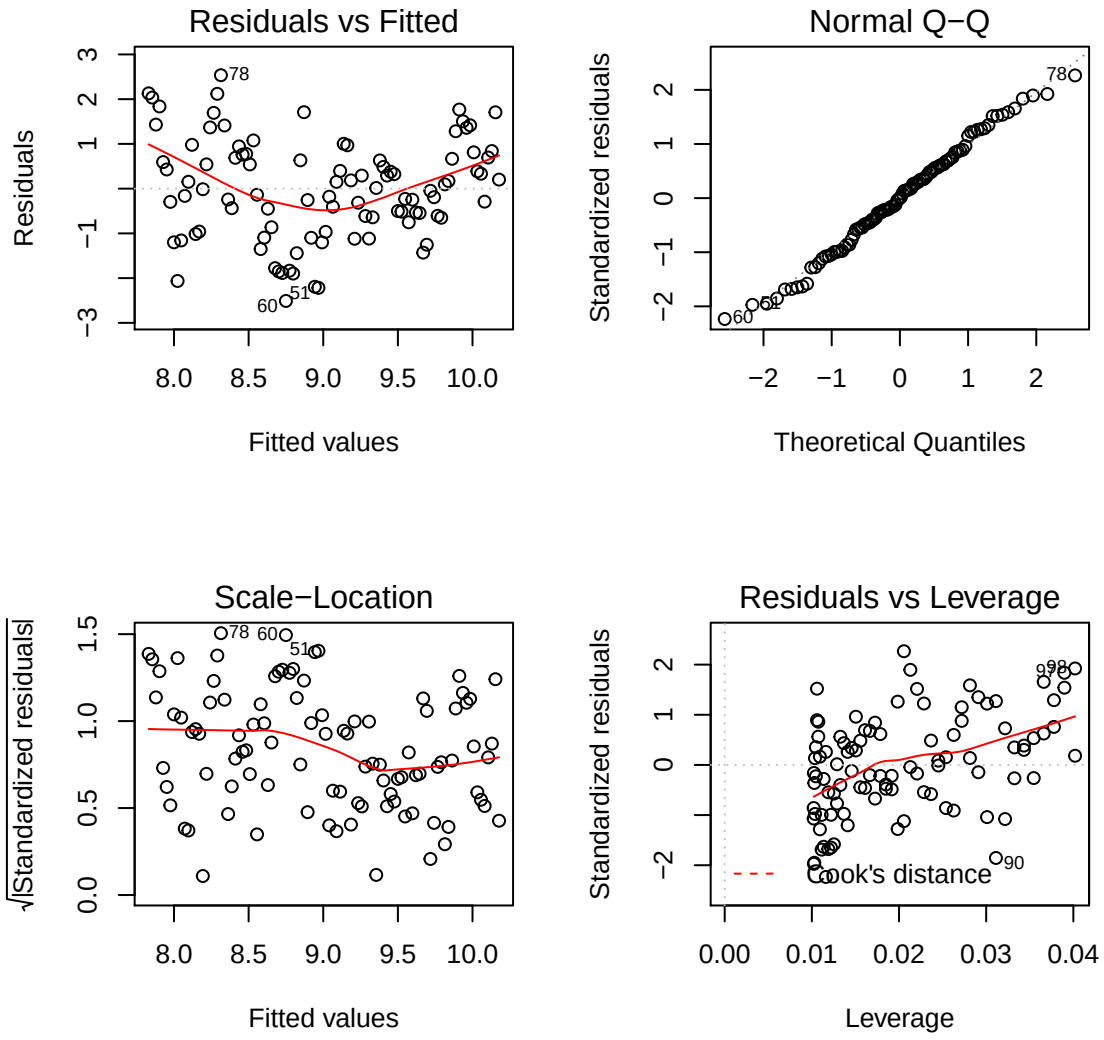
```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.50997 -0.72726  0.00083  0.74402  2.53565
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 10.202037   0.230111  44.335  < 2e-16 ***
## x           -0.024201   0.004036  -5.996 3.55e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.13 on 96 degrees of freedom
## Multiple R-squared:  0.2725, Adjusted R-squared:  0.2649
## F-statistic: 35.95 on 1 and 96 DF,  p-value: 3.545e-08
```

```
plot.ts(data);
title("Lake Huron Water level")
lines(out.lm$fitted.values, col="red")  Linear Regression Line
```



Residuals look like

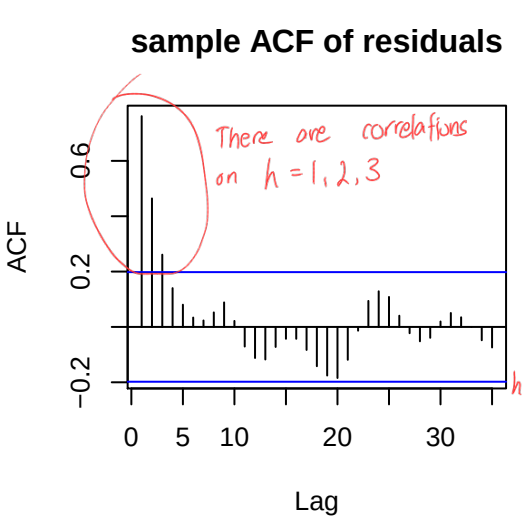
```
# residual diagnostics
par(mfrow=c(2,2))    4개의 그래프를 2행 2열로 정렬
plot(out.lm)
```

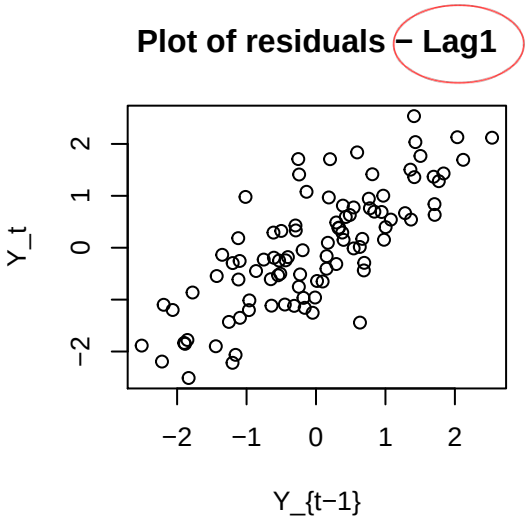


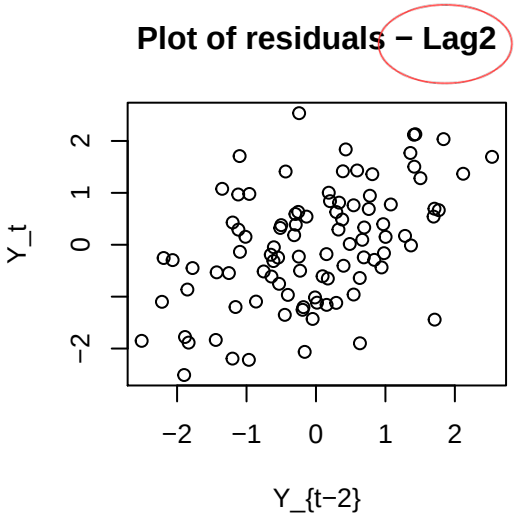
$\gamma(h) = Cov(X_t, X_{t+h})$

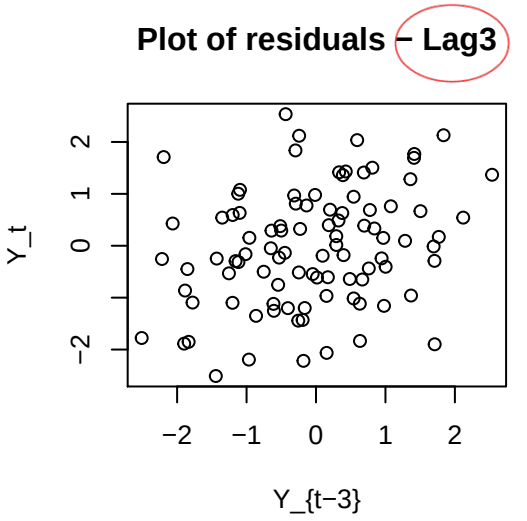
If you see the ACF of residuals, then there still remains correlations on the residuals.

```
par(mfrow=c(2,2))
acf2(out.lm$residuals)
title("sample ACF of residuals");

plot(out.lm$residuals[1:(n-1)], out.lm$residuals[2:n], xlab="Y_{t-1}", ylab="Y_t")
title("Plot of residuals - Lag1")

plot(out.lm$residuals[1:(n-2)], out.lm$residuals[3:n], xlab="Y_{t-2}", ylab="Y_t")
title("Plot of residuals - Lag2")

plot(out.lm$residuals[1:(n-3)], out.lm$residuals[4:n], xlab="Y_{t-3}", ylab="Y_t")
title("Plot of residuals - Lag3")
```


sample ACF of residuals
There are correlations on h = 1, 2, 3
ACF
0.6
0.2
−0.2
0
5
10
20
30
h
Lag


Plot of residuals − Lag1
Y_t
2
1
0
−2
−2
−1
0
1
2
Y_{t−1}


Plot of residuals − Lag2
Y_t
2
1
0
−2
−2
−1
0
1
2
Y_{t−2}


Plot of residuals − Lag3
Y_t
2
1
0
−2
−2
−1
0
1
2
Y_{t−3}

Now, we will try moving avarge to detrend. Detrend by MA(25) and see whether MA successfully produce IID errors.

Bandwidth selection by CV is implemented as in the below

```
ma.cv
```

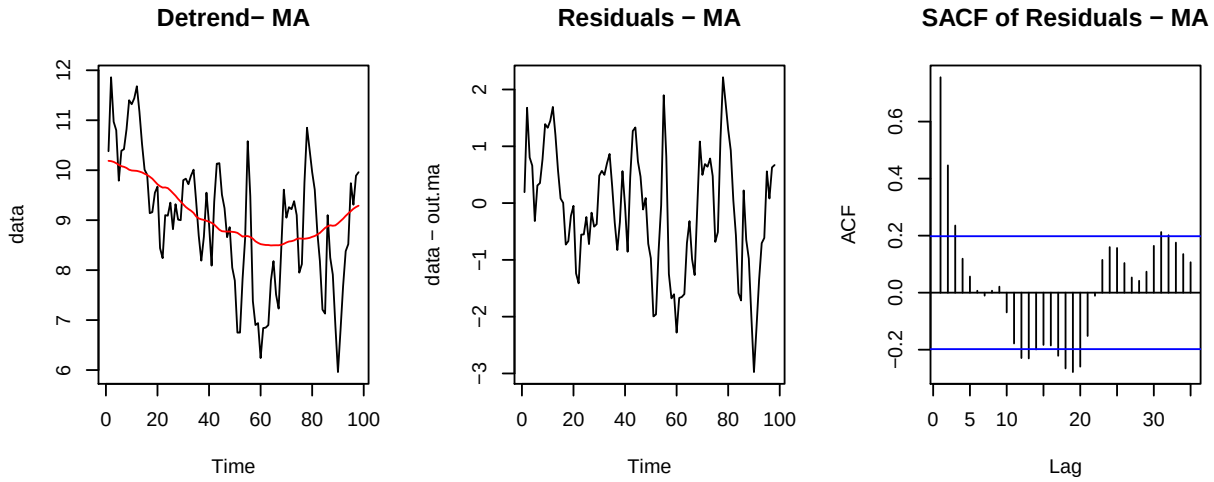
bandwidth dataset — number of test observation

```
## function (h, Y, l)
## {
##     Y <- as.vector(Y)
##     n <- length(Y)
##     cv <- 0
##     ind = 1:n
##     eps <- 1e-16
##     for (i in 1:n) {
##         del = seq(max(i - l, 1), min(i + l, n), by = 1)
##         id = ind[-del]
##         Z = Y[-del]
##         tmp <- (id - i)/h
##         s0 <- (abs(tmp) <= 1)
##         s1 <- Z * s0
##         m <- sum(s1)/max(eps, sum(s0))
##         cv <- cv + (Y[i] - m)^2
##     }
##     return(cv/n)
## }
```

trying to find the best model

```
h.ma = optimize(f=ma.cv, interval=c(5, length(data)/2), Y=data, l=1, tol = .Machine$double.eps^0.25)
out.ma = smooth.ma(data, h.ma$minimum)
```

```
out.ma = smooth.ma(data, 32)
par(mfrow=c(1,3))
plot.ts(data);
lines(out.ma, col="red")
title("Detrend- MA")
plot.ts(data-out.ma);
title("Residuals - MA")
acf2(data-out.ma)
title("SACF of Residuals - MA")
```



Save the results in R

```
save.image("huron.Rdata")
```