Time Series Analysis

Chenyin Gao

(Dept.：School of Math.,Statistics)

2.4 Kullback-Leibler Information

To begin with the p.d.f of $f(y;\theta_1), f(y:\theta_2), \theta = (\beta', \sigma^2)$

$$f(y;\theta_1) = \frac{1}{\sqrt{(2\pi)^n \det(\sigma_1^2)}} \exp(-\frac{1}{2}(\mathbf{Y}-\beta_1)'\sigma_1^- 2(\mathbf{Y}-\beta_1)) \tag{1a}$$

$$f(y;\theta_2) = \frac{1}{\sqrt{(2\pi)^n \det(\sigma_2^2)}} \exp(-\frac{1}{2}(\mathbf{Y}-\beta_2)'\sigma_2^{-2}(\mathbf{Y}-\beta_2)) \tag{1b}$$

Then substitute the (1a) and (1b) in (2)

$$I(\theta_1;\theta_2) = n^{-1}\mathbb{E}_1 \log \frac{f(y;\theta_1)}{f(y;\theta_2)} \tag{2}$$

convert to the integral expression (3)

$$\begin{aligned}
I(\theta_1;\theta_2) &= n^{-1}\int_{\mathbb{R}^n} \log(\frac{f(\mathbf{Y};\theta_1)}{f(\mathbf{Y};\theta_2)}) f(\mathbf{Y};\theta_1) d\mathbf{Y} \\
&= n^{-1}\int_{\mathbb{R}^n} \log(\frac{\sigma_1}{\sigma_2}\exp(-\frac{1}{2}(\mathbf{Y}-\beta_1)'\sigma_1^{-2}(\mathbf{Y}-\beta_1) + \frac{1}{2}(\mathbf{Y}-\beta_2)'\sigma_2^{-2}(\mathbf{Y}-\beta_2)')) f(\mathbf{Y};\theta_1) d\mathbf{Y} \\
&= \frac{1}{n}\int_{\mathbb{R}^n} [\log(\frac{\sigma_1}{\sigma_2}) - \frac{1}{2}(\mathbf{Y}-\beta_1)'\sigma_1^{-2}(\mathbf{Y}-\beta_1) + \frac{1}{2}(\mathbf{Y}-\beta_2)'\sigma_2^{-2}(\mathbf{Y}-\beta_2)] f(\mathbf{Y};\theta_1) d\mathbf{Y} \\
&= \frac{1}{n}\{\log(\frac{\sigma_1}{\sigma_2})n - \frac{n}{2} + \int_{\mathbb{R}^n} \frac{1}{2}(\mathbf{Y}-\beta_1+\beta_1-\beta_2)'\sigma_2^{-2}(\mathbf{Y}-\beta_1+\beta_1-\beta_2) f(\mathbf{Y};\theta_1) d\mathbf{Y}\} \\
&= \log(\frac{\sigma_1}{\sigma_2}) - \frac{1}{2} + \int_{\mathbb{R}^n} \frac{1}{2}\{(\mathbf{Y}-\beta_1)'\sigma_2^{-2}(\mathbf{Y}-\beta_1) + 2(\mathbf{Y}-\beta_1)'\sigma_2^{-2}(\beta_1-\beta_2) + (\beta_1-\beta_2)'\sigma_2^{-2}(\beta_1-\beta_2)\} f(\mathbf{Y};\theta_1) d\mathbf{Y} \\
&= \frac{1}{2}(\log\frac{\sigma_1^2}{\sigma_2^2} - 1 + \frac{\sigma_1^2}{\sigma_2^2}) + \frac{1}{2}\frac{(\beta_1-\beta_2)'(\beta_1-\beta_2)}{n\sigma_2^2}
\end{aligned} \tag{3}$$

2.5 Model Selection

Finding an unbiased estimator for $\mathbb{E}_1[I(\beta_1,\sigma_1^2;\hat{\beta},\hat{\sigma^2})]$ where

$$I(\beta_1,\sigma_1^2;\hat{\beta},\hat{\sigma^2}) = \frac{1}{2}(\log\frac{\sigma_1^2}{\sigma_2^2} - 1 + \frac{\sigma_1^2}{\sigma_2^2}) + \frac{1}{2}\frac{(\beta_1-\beta_2)'(\beta_1-\beta_2)}{n\sigma_2^2} \tag{4}$$

Since

$$\begin{aligned}
\xi &= \frac{n\hat{\sigma^2}}{\sigma_1^2} \sim \chi^2_{n-k} \\
\eta &= \frac{(\hat{\beta}-\beta)'\mathbf{Z}'\mathbf{Z}(\hat{\beta}-\beta_1)}{\sigma_1^2} \sim \chi^2_k \\
\zeta &= \frac{n\hat{\sigma^2}/\sigma_1^2}{(\hat{\beta}-\beta)'\mathbf{Z}'\mathbf{Z}(\hat{\beta}-\beta_1)/\sigma_1^2}\frac{n-k}{k} \sim F(k,n-k)
\end{aligned} \tag{5a}$$

and

$$\begin{aligned}
&\mathbb{E}_1(\frac{\sigma_1^2}{\hat{\sigma^2}}) = \frac{n}{n-k-2} \\
&\mathbb{E}_1(\log\frac{\sigma_1^2}{\hat{\sigma^2}}) = -\log\sigma_1^2 + \mathbb{E}_1\log\hat{\sigma^2} \\
&\mathbb{E}_1(\frac{(\beta_1-\beta_2)'(\beta_1-\beta_2)}{n\sigma_2^2}) = \frac{n-k}{n-k-2}\frac{k}{n-k} = \frac{k}{n-k-2}
\end{aligned} \tag{5b}$$

Therefore,the expectation of (4) could be

$$\begin{aligned}
\mathbb{E}_1[I(\beta_1,\sigma_1^2;\hat{\beta},\hat{\sigma^2})] &= \frac{1}{2}(\frac{n}{n-k-2} - \log\sigma_1^2 + \mathbb{E}_1\log\hat{\sigma^2} + \frac{k}{n-k-2} - 1) \\
&= \frac{1}{2}(-\log\sigma_1^2 + \mathbb{E}_1\log\hat{\sigma^2} + \frac{n+k}{n-k-2} - 1)
\end{aligned} \tag{6}$$

2.8 glacial varve

(a)

(1) shows the original and log-transformed varves with its respective first and second half variation

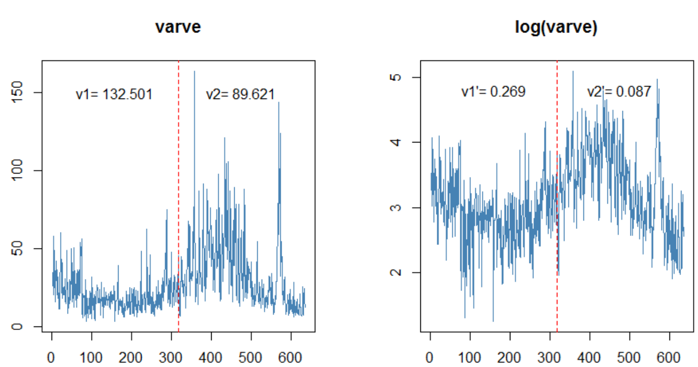

Figure 1: *varves* (top) compared with log transformed *varves* (bottom)

It is clear that the first half sample exhibits more *heterosecdasticity* implying by the sample variance. Plot the histograms of x_t and $y_t = log(x_t)$ in (2) to see whether the normality is improved

```r
len1 <- length(varve)
par(mfrow=c(1,2))
#1
plot(varve,main="varve", ylab="",xlab='',col='steelblue')
abline(v=len1/2,col='red',lty=2)
v1 <- var(varve[1:len1/2])
v2 <- var(varve[len1/2:len1])
text(len1/4,150,paste('v1=',round(v1,3)))
text(3*len1/4,150,paste('v2=',round(v2,3)))
#2
plot(log(varve),main="log(varve)", ylab="",xlab='',col='steelblue')
abline(v=len1/2,col='red',lty=2)
v1_ln <- var(log(varve[1:len1/2]))
v2_ln <- var(log(varve[len1/2:len1]))
text(len1/4,4.8,paste('v1\'=',round(v1_ln,3)))
text(3*len1/4,4.8,paste('v2\'=',round(v2_ln,3)))
```

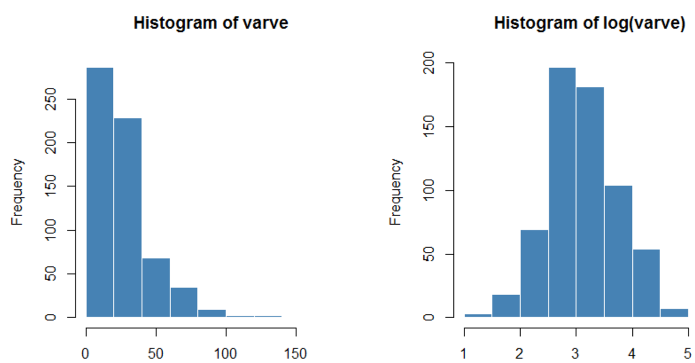

Figure 2: the histogram of *varves*(top) and *varves*(bottom)

From (2), we could conclude that log-transformed could improved the normality of the data. Furthermore, by quantile testing, original correlation value is 0.896. However, the log-transformed correlation value is 0.997, which is improved significantly.

```r
par(mfrow=c(2,1))
hist(varve,col='steelblue',border='white',xlab='')
hist(log(varve),col='steelblue',border='white',xlab='')
# correlation testing
#1
q1 <- qqnorm(varve)
#2
q2 <- qqnorm(log(varve))
cor(q1$x,q1$y)
cor(q2$x,q2$y)
```

(b)

Kernel smoother $m_t = \sum_{i=-50}^{50} w_{t+i}(t) y_t$. (3) uses weights $a_{-50} = a_{-49} = \cdots = a_{50} = 0.01; k = 100$. This particular method removes the obvious annual temperature cycle and helps emphasize the overall trend.

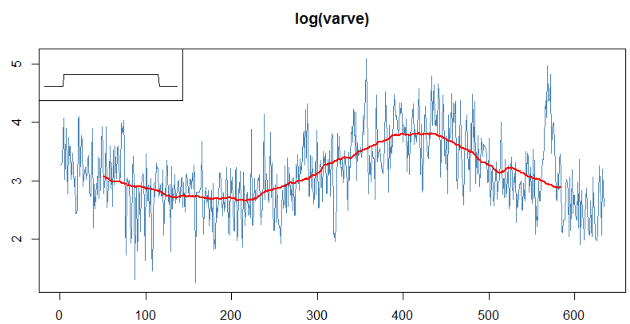

Figure 3: The moving average of y_t.The insert shows the shapes of the kernel

From the (3), we observed the comparable tendency of *varves* to that observed in the global temperature records

```
y <- log(varve)
# log-transformed
plot(y,col='steelblue',main='log(varve)',ylab="",xlab='')
w1 <- rep(1,100)/100
y_w1 <- filter(y,sides = 2,filter = w1)
# moving average
lines(y_w1,col='red',lwd=2)
par(fig = c(0, 0.35, .5, 1), new = TRUE) # the insert
nwgts = c(rep(0,20), w1, rep(0,20))
# kenerl function
plot(nwgts, type="l", xaxt='n', ylim=c(-0.01,0.03),yaxt='n', main='',xlab = '',ylab='')
```

(c)

The definition of **ACF** in (7)

$$\rho(s,t) = \frac{\gamma(s,t)}{\sqrt{\gamma(s,s)\gamma(t,t)}} \tag{7}$$

Examine the sample *ACF* and plot in (4) to comment,which (4) reflects that ACF of y_t peaking at *lag=1* and remaining the same level regardless the *lag*.To comment on, y_t is not a stationary process.

```
acf(y,main='log(varve)')
```

(d)

Compute the difference $u_t = y_t - y_{t-1}$ and examine its time plot and sample ACF in (5)

From the sample ACF of u_t in (5), we conclude that because the value is near zeros when *lag* is larger than 2, u_t is stationary process.

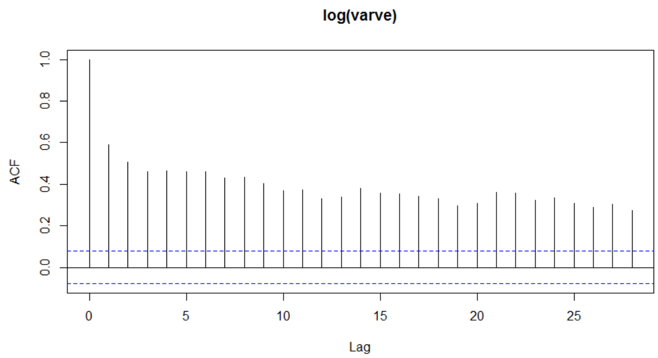

Figure 4: ACF of y_t

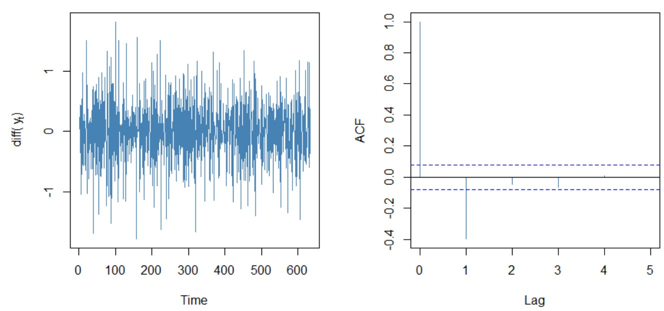

Figure 5: Time plot of u_t(top) and sample ACF(bottom)

Therefore the practical interpretation of u_t, recall *footnote* log(1+p)=$p - \frac{p^2}{2} + \frac{p^3}{3} - \ldots$ for $-1 < p \leq 1$.. If p is near zero, the higher-order terms in the expansion are negligible

$$u_t = y_t - y_{t-1} = \log(x_t) - \log(x_{t-1}) = \log(\frac{x_t}{x_{t-1}} - 1 + 1) \approx \frac{x_t}{x_{t-1}} - 1 \tag{8}$$

So,$x_{t-1}(u_t + 1) \approx x_t$, the practical interpretation of u_t is the increase ratio.

(e)

The generalization of the model is $m_t = \alpha + \beta_1 u_t + \beta_2 u_{t-1}$, the ACF of the stationary process could be expressed as (9)

$$\begin{aligned} \rho(h) &= \frac{\gamma(t+h,t)}{\sqrt{\gamma(t+h,t+h)\gamma(t,t)}} = \frac{\gamma(h)}{\gamma(0)} \\ &= \frac{\gamma(\hat{h})}{\gamma(\hat{0})} = \begin{cases} \frac{\beta_1\beta_2}{\beta_1^2+\beta_2^2} & h = 1 \\ 0 & h > 1 \end{cases} \end{aligned} \tag{9}$$

The (9) is according with the ACF of the difference of u_t, which is near zero after $lag \geq 2$. Therefore, the model is reasonable. Assume $u_t = \mu + w_t + \theta w_{t-1}$, where w_t are assumed independent with mean 0 and variance σ_w^2.So,

$$\begin{aligned} &\gamma_u(0) = var(\mu + w_t + \theta w_{t-1}) \overset{w_t \perp w_{t-1}}{=} \sigma_w^2(1+\theta^2) \\ &\gamma_u(\pm 1) = cor(u_{t\pm 1}, u_t) = \theta cor(w_t, w_t) = \theta\sigma_2^2 \\ &\gamma_u(h) = cor(\mu + w_{t+h} + \theta w_{t+h-1}, \mu + w_t + \theta w_{t-1}) = 0 \;\; |h| > 1 \end{aligned} \tag{10}$$

From (10), we could show that

$$\gamma_u(h) = \begin{cases} \sigma_w^2(1+\theta^2) & h = 0 \\ \theta\sigma_2^2 & h = \pm 1 \\ 0 & |h| > 1 \end{cases} \tag{11}$$

(f)

Derive the parameters by equating sample moments to theoretical moments.

$$\begin{aligned} \hat{\rho}_u(1) &= \frac{\gamma_u(1)}{\gamma_u(0)} = \frac{\theta}{1+\theta^2} \\ \hat{\gamma}_u(0) &= \sigma_w^2(1+\theta^2) \end{aligned} \tag{12}$$

By solving (12), we obtain the estimator of θ, σ_w^2 in (13)

$$\begin{aligned} \hat{\theta} &= \frac{1 \pm \sqrt{1 - 4\hat{\rho}_u^2(1)}}{2} \\ \hat{\sigma_w^2} &= \frac{\hat{\gamma}_u(0)}{1+\hat{\theta}^2} \end{aligned} \tag{13}$$

2.10

(a)

Plot the *oil* and *gas* on the same graph, which is most resemble to the *random walk with drift*, $x_t = \delta + x_{t-1} + w_t$

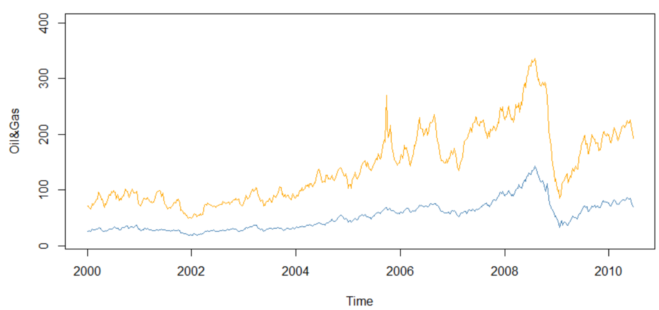

Figure 6: Time plot of oil and gas

From the (6), it is clear that these two data are not stationary because the **mean** of the data is keeping increasing, which contradicts with the same **mean** definition of stationary process.

```
plot(oil, type='l', ylim=c(10,400), ylab='Oil&Gas', col='steelblue')
lines(gas, col='orange')
```

(b)

The transformation $y_t = \nabla \log x_t$ could be expressed in (14)

$$y_t = \nabla \log x_t = \log(\frac{x_t}{x_{t-1}} - 1 + 1) \approx \frac{x_t}{x_{t-1}} - 1 \tag{14}$$

So, $\nabla \log x_t$ is the percentage change in price, which is important in economics.

(c)

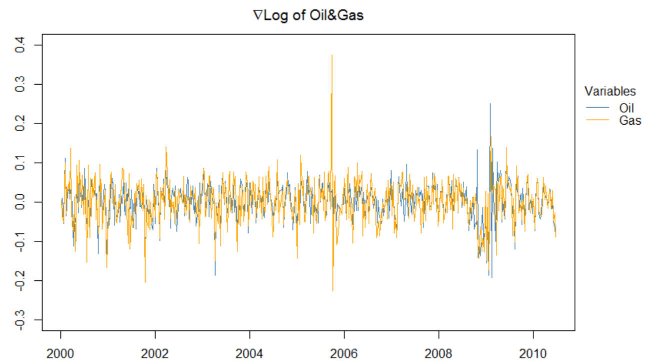

Figure 7: $\nabla \log *$ transformation plot of oil and gas

In (7), we could conclude that, after the $\nabla \log *$ transformation, the time series are intuitively stationary(stay the same mean and variance). Then we plot the sample ACFs for more details

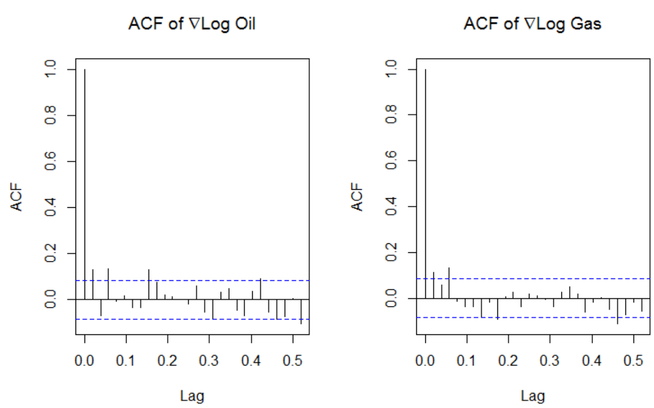

Figure 8: sample ACFs plot of transformed oil and gas

The (8)shows the ACFs plot. We could discover that the autocorrelation decreases after lag ≥ 2, which imply that the transformed time series is somehow stationary and is similar to the $MA(1)$ model.

```
par(mfrow=c(1,2))
acf(diff(log(oil)),main=expression(paste('ACF of ',nabla,'Log Oil')))
acf(diff(log(gas)),main=expression(paste('ACF of ',nabla,'Log Gas')))
```

(d)

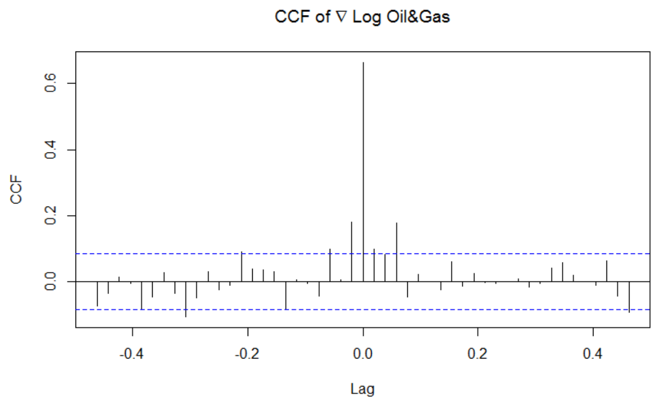

Figure 9: CCF plot of transformed oil and gas

In the Fig(9), the CCF of transformed oil and gas decline when lag or lead is getting larger. However, when gas leads oil about 3 steps, the plot shows relative small but significant peak which could be interpreted as the feedback of gas to oil. So, it is reasonable to propose that the turbulence of the gas price could affect the oil price in the same direct about 3 steps later

```
ccf(diff(log(oil)),diff(log(gas)),
    main=expression(paste('CCF of ',nabla,' Log Oil&Gas')),
    ylab='CCF')
```

(e)

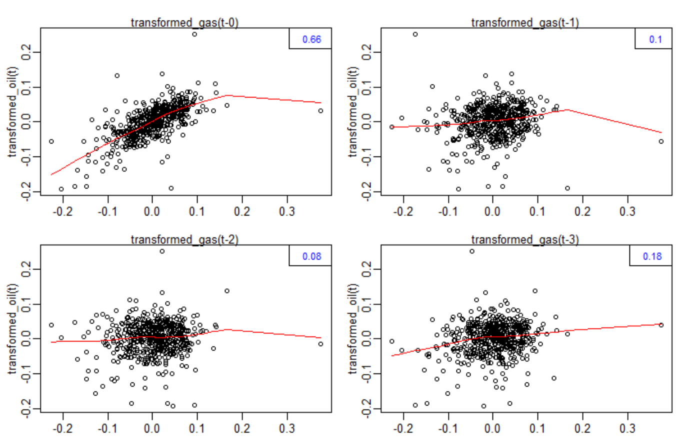

Figure 10: Scatterplots of the oil and gas growth rate series

We plot catterplots of the oil and gas growth rate series for up to three weeks of lead time of oil prices($gas_{(t-3,t-2,t-1,t-0)}, oil_t$) and adds a lowess[1] smoother in each scatterplot

[1]First,a certain proportion of nearest neighbors to x_t are included in a weighting scheme; values closer to x_t in time get more weight. Then, a robust weighted regression is used to predict x_t and obtain the smoothed values m_t. The larger the fraction of nearest neighbors included, the smoother the fit will be. In Fig(10),one smoother uses 5% of the data to obtain the estimate of the data.

The Fig(10) indicate the nonlinear relationship between the gas and oil in three weeks lead. However, besides the high correlation when $t = 0$ which is *0.66*, we notice some outliers in the plots which could play a significant role in the lowess regression due to its reweighted procedures. As a result, we implicate that the gas has somehow small correlation with oil in terms of lead 3, indicated in plot is *0.18*.

```r
transformed_gas <- diff(log(gas))
transformed_oil <- diff(log(oil))
par(mfrow=c(1,3))
lag2.plot(transformed_gas, transformed_oil, 3)
```

(f)

Many researchers questioned that whether gasoline prices respond more quickly when oil prices are rising than when oil prices are falling ('asymmetry')

(i)

Fit the regression

$$G_t = \alpha_1 + \alpha_2 I_t + \beta_1 O_t + \beta_2 O_{t-1} + w_t \tag{15}$$

where $I_t = \begin{cases} 1, if\ O_t \geq 0 \\ 0,\ \ else \end{cases}$ (I_t is the indicator of no growth or positive growth in oil price)

```r
indi <- ifelse(transformed_oil < 0, 0, 1)
mess <- ts.intersect(transformed_gas, transformed_oil,
                     poilL = lag(transformed_oil,-1), indi)
summary(fit <- lm(transformed_gas~ transformed_oil + poilL + indi, data=mess))
```

Table 1: The summary of the regression

	Estimate	Std.Error	t value	$Pr(>\|t\|)$
α_1	-0.00645	0.003464	-1.86	0.06338
α_2	0.012368	0.005516	2.242	0.02534
β_1	0.683127	0.058369	11.704	0
β_2	0.111927	0.038554	2.903	0.00385

(ii)

We seperate the regression result into different scenario, the negativity of O_t and the positivity of O_t and obtain the regression equation respectively.

When O_t is negative

$$G_t = -0.006445 + 0.683127 O_t + 0.111927 O_{t-1} \tag{16a}$$

and when O_t is positive

$$G_t = 0.005923 + 0.683127 O_t + 0.111927 O_{t-1} \tag{16b}$$

However, these equations do not indicate the rapidity of the gas prices response, neither does the asymmetry hypothesis.

(iii)

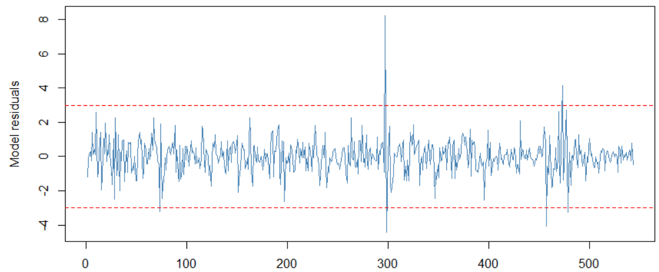

Figure 11: The residual of the model fit

In the Fig(11), we plot the standardized residuals from the model fit in order to further understand the meanings of the residuals. We calibrate the residuals by divide it by *MSE* to discover the outliers. Besides, the residuals show stationary property(same mean and stable variance)

```
ts.plot(fit$residuals/sd(fit$residuals),
        ylab='Model residuals',
        xlab='')
abline(h=3,col='red',lty=2)
abline(h=-3,col='red',lty=2)
anova(fit)
```