Homework Assignment #2

Due Date: April 19 (Campus) and April 18 (Evening), before class

Notes:

- Use 5% level in all tests.
- The notation $\rho_i$ is the lag-$i$ autocorrelation coefficient.
- In some of the problems, I provide guidances to specify a time series model. This is to help you gain experience in empirical data analysis. You can try your own models to gain further experience. The assignments show that multiple models can fit a given data set well and seasonally adjusted data might still have some residual seasonality.

Assignment:

1. (Commodity price). Consider the daily gold fixing price 10:30 am (London time) in London Bullion Market in U.S. dollars per Troy ounce from January 3, 1995 to March 30, 2017. The data can be obtained from FRED using the `quantmod` package. Since there are some missing values, we need to remove them before analysis. Let $x_t = \log(\text{gold price})$. See instructions below.

   ```r
   require(quantmod)
   getSymbols('GOLDAMGBD228NLBM', src='FRED')
   GOLD <- GOLDAMGBD228NLBM[6982:12784]
   idx <- c(1:nrow(GOLD))[is.na(GOLD)]
   GOLD <- GOLD[-idx]
   xt <- log(as.numeric(GOLD))
   rt <- 100*diff(xt)
   ```

   (a) Obtain the time plots of $x_t$ and $r_t$ (in one page, using the command `par(mfcol=c(2,1))`).

   (b) Compute the first 12 lags of ACF of $x_t$. Based on the ACF, is there a unit root in $x_t$? Why?

   (c) Let $r_t = 100 \times (x_t - x_{t-1})$ be the return series of the gold prices, in percentages. Consider the $r_t$ series. Test $H_0 : \rho_1 = \cdots = \rho_{12} = 0$ versus $H_a : \rho_i \neq 0$ for some $1 \leq i \leq 12$. Draw your conclusion.

   (d) Use the command `ar(rt,method='mle',order.max=20)` to specify the order of an AR model for $r_t$.

   (e) Build an AR model for $r_t$, including model checking. Refine the model by excluding all estimates with $t$-ratio less than 1.645. Write down the fitted model.
(f) Use the fitted AR model to compute 1-step to 4-step ahead forecasts of \( r_t \) at the forecast origin March 30, 2017. Also, compute the corresponding 95% interval forecasts.

2. (Model comparison). Consider, again, the log return series of gold price of Problem 1.

(a) Build an MA(7) model for \( r_t \). Refine the model by removing coefficient estimates with \( t \)-ratio less than 1.645. Write down the fitted model.

(b) Compute the Ljung-Box statistic \( Q(10) \) of the residuals of the fitted MA(7) model. Is there serial correlation in the residuals? Why?

(c) Consider in-sample fits of the AR model of Problem 1 and the MA(7) model. Which model is preferred? Why?

(d) Use \texttt{backtest} at the forecast origin \( t = 5585 \) with horizon \( h = 1 \) to compare the two models. Which model is preferred? Why?

3. (Volatility modeling). Consider the CBOE daily volatility index (VIX) from January 2, 1990 to March 31, 2017. The data are available from FRED via \texttt{quantmod}. Again, there are some missing values, which need to be removed. Use the commands below:

```r
getSymbols('VIXCLS', src='FRED')
VIXCLS <- VIXCLS[1:7109]
idx <- c(1:7109)[is.na(VIXCLS)]
VIXCLS <- VIXCLS[-idx]
vix <- as.numeric(VIXCLS)
```

(a) Find an AR model for the VIX series. Remove insignificant coefficient estimates (based on \( t \)-ratio 1.645). Provide model checking to confirm that the model is adequate. Write down the model.

(b) Use the fitted model to obtain 1-step to 10-step ahead predictions at the forecast origin March 31, 2017.

4. (GDP revisited). Consider, again, the U.S. quarterly real GDP growth rate from 1947.II to 2016.IV. The data are available from FRED. See the command below.

```r
getSymbols('A191RL1Q225SBEA', src='FRED')
gdp <- as.numeric(A191RL1Q225SBEA)
```

(a) Obtain time-series plot of the real GDP growth rates.

(b) Find an AR model for the real GDP growth rate, including model checking. Write down the fitted model.

(c) Does the model imply existence of business cycle? Why?

(d) If business cycles are present, compute the average length of the cycles. Otherwise, the length is infinity.

(e) Obtain 95% interval forecasts of 1-step to 4-step ahead GDP growth rates at the forecast origin 2016.IV.
5. (Seasonal model). Consider the monthly U.S. unemployment rates used in the lecture. Data are available from FRED and is entitled UNRATE. Let $r_t$ be the unemployment at time $t$. Ignore any possible outliers in the data.

(a) Fit the seasonal model as below

```r
m1 <- arima(unrate, order=c(3,0,1), seasonal=list(order=c(1,0,1), period=12))
```

Perform model checking and write down the fitted model.

(b) Fit an AR(11) model as shown in the Lecture Note 2.

(c) Use `backtest` at the forecast origin $t=770$ and horizon $h=1$ to compare the two models. Which model is preferred? Why?

**Reading assignments**: Chapter 2 of the textbook.