Financial Statistics: Time Series, Forecasting, Mean Reversion, and High Frequency Data

FINM 33170 and STAT 33910 Winter 2021

HW 2, due Friday 29 January, 2021

Homework should be scanned and e-mailed to both Ahmed Bou-Rabee, ahmedb@uchicago.edu, and Yi Wang, samwang8899@uchicago.edu

1. **Some more OEF data**

   (a) (Creating a time series of volatilities.) We return to the file "f657ab9f9ecd5e33.csv" from last time. Also, load the file TSCV.R (on the class home page) into R, by putting it in your directory, and then give the command

   ```
   source("TSCV.R")
   TSCV #if you would like to look at the function
   ```

   The TSCV function provides better estimates for volatilities than the sum of squares. (We shall revisit this method later on.) To more easily operate on the file, you can now again give the command

   ```
   attach(series)
   ```

   To operate on all dates of the file, set

   ```
   dates<-unique(DATE)
   dtno<-length(dates)
   ```

   To calculate the time series of volatilities for the year 2017, use the commands

   ```
   vol<-c(1:dtno)*0
   for (i in 1:dtno) {
   dt <- dates[i]
   x<-log(PRICE[DATE==dt])
   y<-x
   vol[i]<-TSCV(x,y,4,20)
   }
   ```

   (This assumes a choice of smoothing parameters, and we shall return to this issue later.) Plot this series, along with the series

```
lvol<-log(vol)
```

(b) (A time series analysis.) You may now fit an ARMA(1,1) model to the log volatility series by the command

```
lvol.mod<-arima(lvol,order=c(1,0,1),method="ML")
```

Try the commands

```
lvol.mod$coef
lvol.mod$var.coef
lvol.mod$sigma2,
```

to determine what time series model has been estimated, and whether you would like to eliminate any of the parameters.

Use plots, and the commands acf and pacf on the original lvol series and on the residuals vol.mod$residuals to see if the fitted model makes sense.

2. **AR and estimation** (Optional. We would like you to give a plausible but not necessarily rigorous explanation.)

Suppose that as in Section 2.1.4 of the high frequency notes, on each day t, the log stock price follows a Brownian motion $X_{t,u} = X_{t_0} + \mu_t u + \sigma_t W_{t,u}$, where $u \to W_{t,u}$ is a Brownian motion (in time u) for each day t. (Hence, u is the time from the start of the day, while t is the day number.) Assume, for each day t, that σ_t^2 is estimated by its maximum likelihood estimator $\hat{\sigma}_t^2$, as given in the notes. Note that following the development on p. 113,

$$\log(\hat{\sigma}_t^2) = \log(\sigma_t^2) + U_t \tag{1}$$

where U_t has the distribution of $\log(\chi^2_{n-1}/n)$.

Assume that $\log(\sigma_t^2)$ follows an AR(1) model:

$$\log(\sigma_t^2) = \varphi_1 \log(\sigma_{t-1}^2) + \epsilon_t. \tag{2}$$

Also assume that the U_t and the ϵ_t are independent sequences. Set $Y_t = \log(\hat{\sigma}_t^2)$. Show that Y_t is an ARMA(1,1) in a second order sense (in other words, based on autocovariances). (You do not need to find the innovations Z_t in this process.)