MGMT 237E: Empirical Methods in Finance

Homework 3

Yi-Chi Chan Zhaofang Shi
Ahswin Kumar Ashok Kumar George Bonebright

January 26, 2016

1.1. $y_t = \phi_0 + \phi_2 y_{t-2} + \epsilon_t$
$\mu = \phi_0 + \phi_2\mu$
$y_t - \mu = \phi_2(y_{t-2} - \mu) + \epsilon_t$
$(y_t - \mu)(y_{t-1} - \mu) = \phi_2(y_{t-2} - \mu)(y_{t-1} - \mu) + \epsilon_t(y_{t-1} - \mu)$
$\gamma_1 = \phi_2\gamma_1, |\phi_2| < 1$
First order auto-covariance: $\gamma_1 = 0$

$\mu = \frac{\phi_0}{1-\phi_2}$
$\gamma_0 = \phi_2\gamma_2 + \sigma^2$
$\gamma_2 = \phi_2\gamma_0$
$\gamma_0 = \frac{\sigma^2}{1-\phi_2^2}$
Second order auto-covariance: $\gamma_2 = \phi_2 \frac{\sigma^2}{1-\phi_2^2}$

1.2. $\phi_{pp} = corr(y_t, y_{t-p}|y_{t-p}, y_{t-p+1}, y_{t-p+2}, ..., y_{t-1})$
$\phi_{11} = corr(y_t, y_{t-1}|y_{t-1}) = corr(\phi_0 + \phi_2 y_{t-2} + \epsilon_t, y_{t-1}|y_{t-1})$
$= corr(\phi_0, y_{t-1}|y_{t-1}) + corr(\phi_2 y_{t-2}, y_{t-1}|y_{t-1}) + corr(\epsilon_t, y_{t-1}|y_{t-1})$
$= 0$

$\phi_{22} = corr(y_t, y_{t-2}|y_{t-2}, y_{t-1}) = corr(\phi_0 + \phi_2 y_{t-2} + \epsilon_t, y_{t-2}|y_{t-2}, y_{t-1})$
$= corr(\phi_0, y_{t-2}|y_{t-2}) + corr(\phi_2 y_{t-2}, y_{t-2}|y_{t-2}) + corr(\epsilon_t, y_{t-2}|y_{t-2})$
$= \phi_2$

1.3. $y_t = \phi_0 + \phi_2 y_{t-2} + \epsilon_t$
$= \phi_0 + \phi_2(\phi_0 + \phi_2 y_{t-4} + \epsilon_{t-2}) + \epsilon_t$
$= \phi_0(1 + \phi_2) + \phi_2^2 y_{t-4} + \phi_2\epsilon_{t-2} + \epsilon_t$
$= \phi_0(1 + \phi_2) + \phi_2^2(\phi_0 + \phi_2 y_{t-6} + \epsilon_{t-4}) + \phi_2\epsilon_{t-2} + \epsilon_t$
$= \phi_0(1 + \phi_2 + \phi_2^2) + \phi_2^3 y_{t-6} + \phi_2^2\epsilon_{t-4} + \phi_2\epsilon_{t-2} + \epsilon_t$
$= \phi_0(1 + \phi_2 + \phi_2^2 + ... + \phi_2^{k-1} + \phi_2^k) + \phi_2^{k+1} y_{t-2k-2} + \phi_2^k\epsilon_{t-2k} + ... + \phi_2^2\epsilon_{t-4} +$
$\phi_2\epsilon_{t-2} + \epsilon_t$
$\psi_{2k} = \frac{\delta y_t}{\epsilon_{t-2k}} = \phi_2^k$
$\psi_{2k-1} = \frac{\delta y_t}{\epsilon_{t-2k+1}} = 0$
$\psi_k = \frac{\phi_2^{\frac{k}{2}}}{2} + (-1)^k \frac{\phi_2^{\frac{k}{2}}}{2}$

1.4. $\hat{r_t}(h) = \mu + \psi_h\epsilon_t + \psi_{h+1}\epsilon_{t-1} + ...$
$= \frac{\phi_0}{1-\phi_2} + (\frac{\phi_2^{\frac{h}{2}}}{2} + (-1)^h \frac{\phi_2^{\frac{h}{2}}}{2})\epsilon_t + (\frac{\phi_2^{\frac{h+1}{2}}}{2} + (-1)^{(h+1)} \frac{\phi_2^{\frac{h+1}{2}}}{2})\epsilon_{t-1} + ...$

$v_t(h) = \epsilon_{t+h} + \psi_1\epsilon_{t+h-1} + \psi_2\epsilon_{t+h-2} + ... + \psi_{h-1}\epsilon_{t+1}$
$= \epsilon_{t+h} + \phi_2\epsilon_{t+h-2} + \phi_2^2\epsilon_{t+h-4} + ... + (\frac{\phi_2^{\frac{h-1}{2}}}{2} + (-1)^{(h-1)} \frac{\phi_2^{\frac{h-1}{2}}}{2})\epsilon_{t+1}$

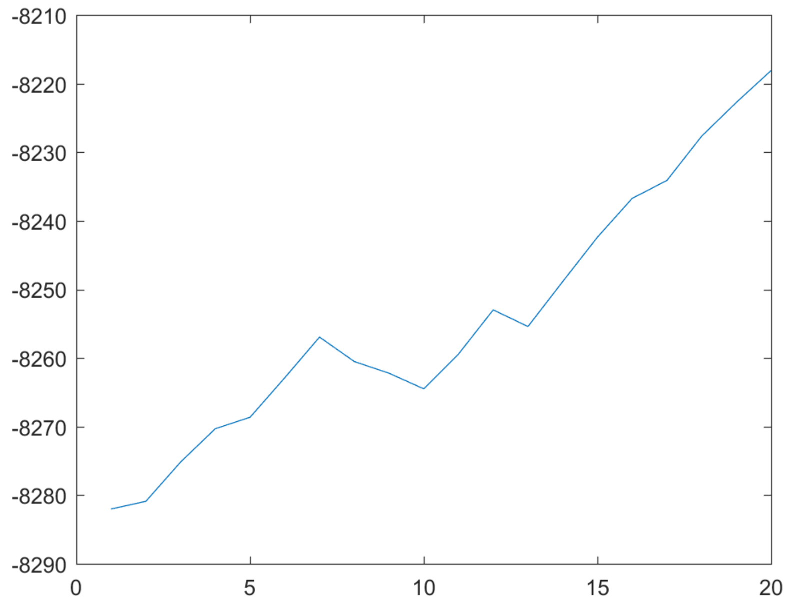
$$Var[v_t(h)] = (1 + \psi_1^2 + \psi_2^2 + ... + \psi_{h-1}^2)\sigma_\epsilon^2$$
$$= (1 + \phi_2^2 + \phi_4^2 + (\frac{\phi_2^{\frac{h-1}{2}}}{2} + (-1)^{(h-1)}\frac{\phi_2^{\frac{h-1}{2}}}{2})^2)\sigma_\epsilon^2$$

3.1. The best model for the data is the AR(1) model. Although higher order lags do have some auto-correlation, it is most likely due to carry over from the first lag.

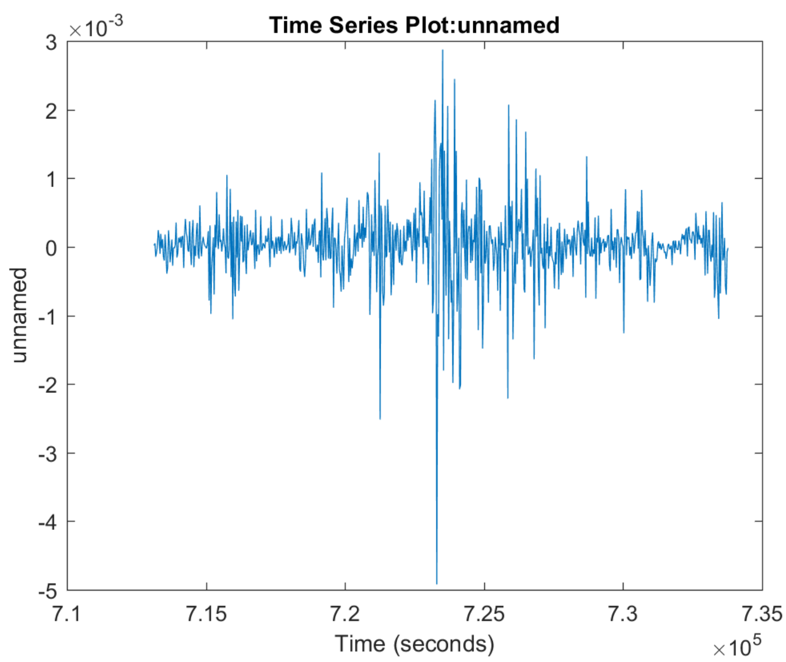
```
ARIMA(20,0,0) Model:
--------------------
Conditional Probability Distribution: Gaussian

                                  Standard          t
 Parameter       Value             Error        Statistic
-----------   -----------      ------------   -----------
 Constant     0.000106277      5.88213e-05        1.80678
    AR{1}        0.879565         0.025217        34.8798
    AR{2}       0.0743894        0.0344552        2.15901
    AR{3}      -0.0597999         0.041664       -1.43529
    AR{4}      -0.0148535        0.0394274      -0.376731
    AR{5}       0.0889085        0.0324679        2.73835
    AR{6}       0.0399696        0.0371973        1.07453
    AR{7}       -0.144663        0.0422749       -3.42195
    AR{8}        0.231795        0.0379251        6.11192
    AR{9}      0.00882295         0.044628         0.1977
   AR{10}       -0.180959        0.0352317       -5.13626
   AR{11}       0.0730484        0.0445958        1.63801
   AR{12}       0.0988266        0.0452402        2.18449
   AR{13}       -0.116297        0.0401625       -2.89565
   AR{14}      -0.0083523        0.0357916      -0.233359
   AR{15}      -0.0149868        0.0425696      -0.352055
   AR{16}       0.0993057        0.0532711        1.86416
   AR{17}      -0.0829113        0.0563352       -1.47175
   AR{18}       0.0560752        0.0557955        1.00501
   AR{19}     -0.00384941        0.0437847     -0.0879167
   AR{20}      -0.0516555        0.0287052       -1.79952
 Variance     2.65011e-07      9.73801e-08        2.72141
```

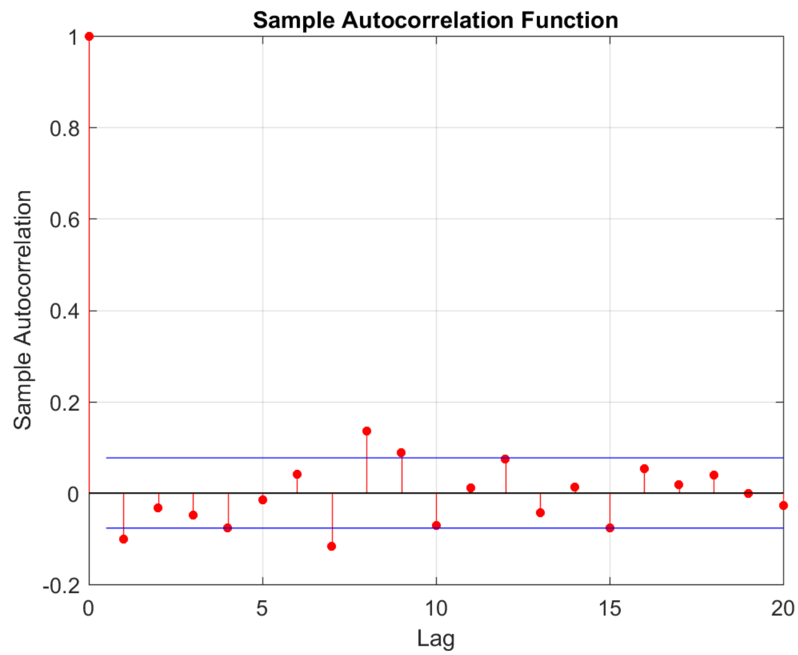
Graph of BIC values:



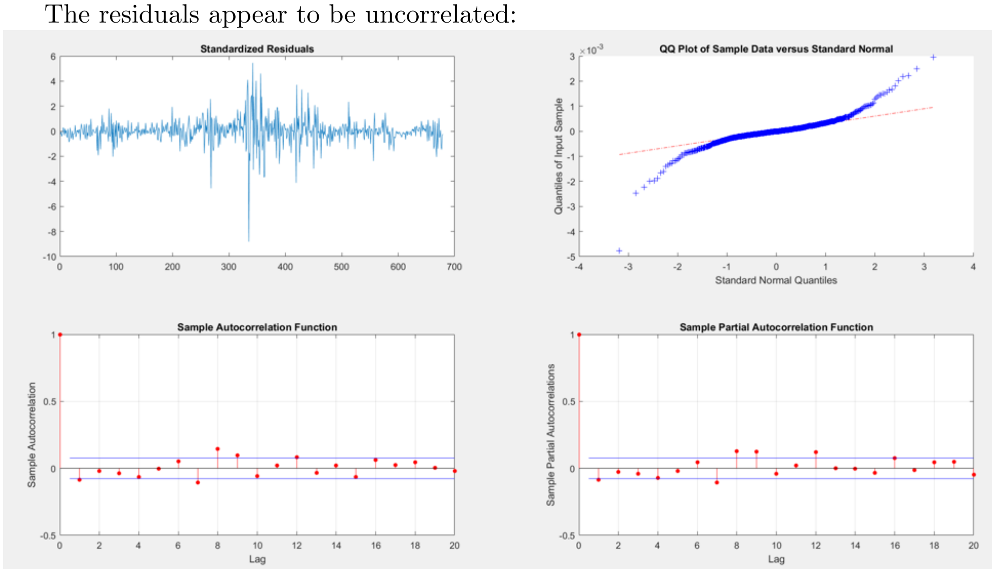
Graph of Time Series:



Graph of Autocorrelation Function:

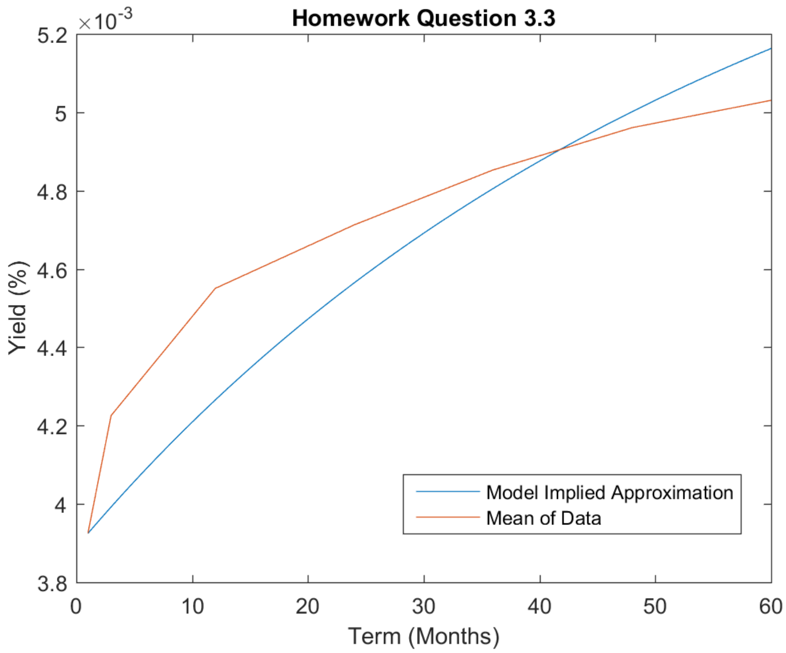


3.2. For AR(1) Estimate: $\mu = .0039; \phi = .9725; \sigma = 5.46 * 10^{-4}$
The residuals appear to be uncorrelated:



The model does not capture new information such as policy changes which are independent of past monthly yields. Furthermore, the AR(1) model assumes stationarity so variation over past periods is not captured.

3.3. Comparison of Model and Actual Data:



Although the model does not appear to be a good fit for the data it isn't an unusable estimate. The model seems to capture an average of the slopes of the unconditional means of interest rates. Additionally, the Vasicek model is based on a single factor which prevents it from capturing the curvature of a multi-factor function.

```
clear all
%% Problem 3: AR(1) model of the yield curve
%% (1). estimate different AR(p) models for the 1-month yield
data = xlsread('fama_bliss_data.xlsx');
for i =1:length(data(:,1))
    data(i, 1) = datenum(num2str(data(i,1)), 'yyyymmdd');
end
dropped_data = data(1:(751-72), :);
y = dropped_data(:,2);

p=20;
aic=zeros(p,1);
bic=zeros(p,1);
for i =1:p
    Mdl=arima(i,0,0);
    [EstMdl,EstParamCov,logL,info] = estimate(Mdl,y);
    [aic(i,1), bic(i,1)]=aicbic(logL,i,679);
end
plot(bic)
% I choose AR(1) model because the BIC value is the lowest.
ToEstMdl = arima(1, 0, 0);
ToEstMdl.Constant = 0;
Y = dropped_data(:, 2);
[EstMdl, EstParamCov, logL, info] = estimate(ToEstMdl, Y );
OurModel = arima('AR',{0.992729},'Variance',0.2,'Constant',0);
[E,V] = infer(OurModel,Y);

ts1 = timeseries(E, dropped_data(:, 1));
plot(ts1)
datetick('x','mmm yyyy')
autocorr(E)
%%
%% (2). Estimate an AR(1) model on monthly data for the 1-month yield
Y = dropped_data(:,2);
X = zeros(length(Y), 3);
X(:, 1) = 1;
X(:, 2) = lagmatrix(Y,1);
[parameters, bint, residuals,rint,stats] = regress(Y,X);
mu = parameters(1)/(1-parameters(2));
phi = parameters(2);
sigma = sqrt(var(residuals, 'omitnan'));
%%
%% (3). calibrate the Vasicek model to get the best ?t for the average yield curve in the sample
largest = 60;
a_bar = zeros;
b_bar = zeros;
a_bar(1) = 0;
b_bar(1) = -1;
delta0 = 0;
delta1 = 1;
% initial guess for lamda_0 and lamda_1 from class notes
lamda_0 = -0.1144;
lamda_1 = -10.741;
step = 0.1;
step_size = 100;
```

```
l0 = lamda_0-step*step_size:step:lamda_0+ step*step_size;
l1 = lamda_1-step*step_size:step:lamda_1+step*step_size;
minsum = 1000;
l0min = 0;
l1min = 0;
Vasicek_Y = zeros;
Vasicek_Y(1) = mean(Y);


for lambda0 = l0
    for lambda1 = l1
        for i =2:largest
            b_bar(i) = b_bar(i-1)*(phi - sigma*lambda1) - delta1;
            a_bar(i) = a_bar(i-1) - delta0 + b_bar(i-1)*((1-phi)*mu - sigma*lambda0) + 1/2*sigma^2*(b_bar(i-1))^2;
        end
        g = Y;
        z = zeros(length(Y),6);
        k = 1;
        for i = [3,12,24,36,48,60]
            z(:,k) = -(a_bar(i)+b_bar(i)*g)/i;
            k = k+1;
        end
        if sum((mean(dropped_data(:,2:8))-mean([g,z]))'.^2) < minsum
            minsum = sum((mean(dropped_data(:,2:8))-mean([g,z]))'.^2);
            l0min = lambda0;
            l1min = lambda1;
        end
    end
end

%Using the l0min and l1min
for i =2:largest
     b_bar(i) = b_bar(i-1)*(phi - sigma*l1min) - delta1;
     a_bar(i) = a_bar(i-1) - delta0 + b_bar(i-1)*((1-phi)*mu - sigma*l0min) + 1/2*sigma^2*(b_bar(i-1))^2;
end

for i =2:largest
    Vasicek_Y(i) = -1/i*a_bar(i) -1/i*b_bar(i)*Vasicek_Y(1);
end
periods = [1 3 12 24 36 48 60];
TermStructure = zeros;
for i=1:length(periods)
    TermStructure(i) = mean(dropped_data(:, i+1));
end
x = 1:60;
plot(x, Vasicek_Y, periods, TermStructure);
%%
```