**EMET8012 Business and Economic Forecasting**

**Final Project Report**

1. **Introduction**

In business and economic fields, forecasting has been a very prominent tool as number of outcomes from those two fields are observed in timely basis. In this paper, I am intended to fit several time series models to forecast the Quarterly Indonesian GDP Consumption Index. I am interested to forecast the Indonesian real GDP as GDP is one of the common indicators to examine the health of a nation’s economy. The Quarterly Indonesian GDP Consumption data used here is obtained from the CEIC official website and is originally released by Indonesian Bureau of Statistics (<https://insights-ceicdata-com.virtual.anu.edu.au/Untitled-insight/views>). This series is quarterly released somewhere between date 7 -11 for each quarter. Originally, the series being analysed in this paper was measured using different base price which are year 2000 price and year 2010 price. However, to make the data has the same base price, I recomputed the data set using the year 2000 base price. Afterwards, the further study will be analysed using the newest rescaling series. Furthermore, as the data is in billion unit, I rescaled the data into a unit by dividing the data with 105.

Given the previous information, the aim of this paper is to forecast the value of the quarterly Indonesian real GDP consumption for one year ahead. In regard to model fitting, I consider several time series models to be fitted into the series to answer the aims of this paper. However, prior to model fitting, it might be important to do a preliminary graphical data analysis on the series to see the behaviour of the series

1. **Methodology**

**II(a). Data**

The series used in this paper is quarterly Indonesian real GDP consumption starting from the first quarter of year 2000 to the second quarter of year 2019 (T=78). In addition, as we need to assess the performance of the fitted model in forecasting the series, both in-sample and out-sample goodness of fit measurement are proposed. In regard to the in-sample goodness-of-fit measurement, the MSE, AIC and BIC will be computed while the MSFE will be used as the out-sample goodness of fit which use the 40st until the 78th observation as the test data set. Afterwards, as several time series models are used to forecast the quarterly Indonesian real GDP Consumption, the model comparison will be undertaken to find the best fit for the series used in this paper.

**II(b). Graphical Data Analysis**

Before performing the time series model fitting, the graphical data analysis was firstly applied to the quarterly Indonesian real GDP consumption. The following is the time series plot of the given series;



Figure 1. Time series plot of the quarterly Indonesian GDP consumption

Based on Figure 1, we can see that that the series might not be stationary in variance. Hence, when fitting one of the Wold-Representation frameworks, it might be useful to check the stationary of the series. if the data is not stationary, the could do log transformation applied into the series or firstly differencing the series.

1. **Model Fitting**

**III(a). Benchmark**

In this paper, I used the random walk process as the benchmark of the model, assuming that the current real GDP consumption is affected only by the previous real GDP. The formula of this specified model is given below

To forecast the 1-step-ahead of the series *y*, the above random walk process gives the following equation;

Afterwards, by setting , I obtained the goodness-of-fit measurements of the random walk in forecasting the series as follow ;

Table 1. The goodness-of-fit of the random walk process

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | AIC | BIC | MSE | MSFE |
| Random Walk | 4.1885 | 6.5352 | 0.0284 | 0.0392 |

Here, the visualisation of the 1-step-ahead forecast of Indonesian real GDP consumption using random walk model is displayed at the Appendix A3(1). Based on the given plot, the random walk model is quite good in approaching the Indonesian real GDP consumption by which the MSFE and the MSE for the random walk is somewhat small at 0.0284 and 0.0392 respectively. However, it is of interest to fit the other time series models to forecast the Indonesian real GDP consumption rather than only fitting the random walk process.

**III(b). Seasonality Models**

Based on the preliminary graphical analysis, it can be seen that the Indonesian real GPD consumption experience a rise at the fourth quarter in each year. Hence, for the seasonality model fitting, it might be appropriate to fit a model where a dummy of the fourth quarter is constructed. However, instead of only assuming that the GDP consumption in the fourth quarter is significantly different compared to in the other quarter, it might be also useful to fit a model which can take account the different real GPD consumption value for each quarter. For a more clear explanation, the following are the model that I intended to fit. Let Dit be a dummy variable which denotes the quarter of the year, here to forecast the quarterly Indonesian real GDP consumption, I consider the following two specifications

After fitting the model, I obtained the measurement error of the specification given above in forecasting the Indonesian real GDP consumption given in Table 2. In addition, as the benchmark model is the random walk process, I used the 1-step-ahead forecast for model comparison based on the MSFE.

Table 2. The goodness-of-fit of the specification 1 and specification 2

|  |  |  |
| --- | --- | --- |
| Measurement Error | S1 | S2 |
| AIC | 21.9695 | 21.8510 |
| BIC | 33.6345 | 33.6345 |
| MSE | 0.1535 | 0.1519 |
| MSFE | 0.3276 | 0.3350 |

Based on Table 2, all the goodness of fit produced by the two seasonality process does not improve the benchmark model in forecasting the Indonesian real GDP consumption. Hence, for the further analysis such as model comparison, the seasonality specifications can be ignored.

**III(c). Holt-Winter Smoothing Seasonality Approach**

Based on the previous section, the seasonality models are not being able to capture the trend of the Indonesian real GDP consumption series. Hence, in this part, I attempted to fit the Holt-Winters smoothing seasonality to forecast the *y* series. Based on Figure 1, the series used in this paper seems to have quarter seasonality. As we know that the series I analysed here is observed quarterly, then I fitted the quarterly seasonality Holt-Winters smoothing and the algorithm is the same as given in the lecture notes. In regard to the smoothing parameters, I put , , and . The Holt-Winters seasonality smoothing to forecast the *y* series is written in the following;

where is the level, is the slope and is a term that capture the seasonality term. In addition, those three terms are obtained by using the following formula

;

and

In this study, by using the three parameter smoothing given above, I obtained the corresponding Holt-Winters goodness-of-fit measurement as follow;

Table 4. The goodness-of-fit of the Holt-Winters quarter seasonality

|  |  |
| --- | --- |
| Forecast Horizon | MSFE |
| h=1 | 0.0036 |
| h=3 | 0.0042 |

As the benchmark I used in this paper is the random walk process, it is better to compare the benchmark with the 1-step-ahead forecast. Based on the MSFE value given in Table 4, the Holt-Winters quarterly smoothing significantly improve the forecast of the Indonesian real GDP consumption. Furthermore, based on the 1-step-ahead forecasting plot given in Appendix A3(4), it is obvious that the Holt-Winters quarterly smoothing well capture the behaviour of the series which result in really small MSFE. However, as I deal with quarterly series, I also attempted to constructed the 3-step-ahead forecast for the series. And the corresponding error measurement are also given in Table 4. Here, as the Holt-winters quarterly smoothing improve the forecast of the series, this fitted model can be included in the model comparison.

**III(d).** **AR(p) Model Fitting**

From the previous Holt-Winters seasonality smoothing, we know that the process significantly improve the forecast. However, the Holt-Winters smoothing is sensitice if there is a break in the series. Hence, I tried to explore another time series model which can take account the drawback posed by the Hold-Winters smoothing. Hence, I attempted to plot the ARMA(p,q) process to forecast the Indonesian real GDP consumption.

The series plot given in Figure 1 suggest there might be a relationship between the data. Hence, fitting the ARMA(p,q) into the given series might be reasonable. However, prior to the ARMA(p,q) model fitting, it is better to plot the covariance of the series to identify which Wold Representation process fit our series best. The autocovariance plot of the series used in this paper is given below;



Figure 2. The autocovariance plot of the Indonesian GDP consumption

The covariance plot given above suggests that the series is gradually decreasing which indicate that AR(p) or either ARMA(p,q) are the appropriate models to be fitted into the series. Before fitting the more complex model, I consider to fit the simpler model which are AR(1) and AR(4) process. Here, I considered fitting the AR(4) as the series being analysed is observed quarterly. The following is the model formula which I intended to fit;

AR(1) :

AR(2) :

After fitting the proposed models into the data set, I obtained the goodness-of-fit measurement of both AR(1) and AR(4) in forecasting 1-step-ahead forecast;

Table 5. The goodness-of-fit of AR(1)

|  |  |  |
| --- | --- | --- |
| Model | Forecast Horizon | MSFE |
| AR(1) | 1 | 0.0293 |
| AR(4) | 1 | 0.0032 |

Based on Table 3, it is obvious that the AR(4) process perform better in forecasting the Indonesian real GDP consumption. Furthermore, the AR(4) model improve the forecasting results compared to the benchmark and all the previous proposed model including the Holt-Winters quarterly smoothing. Hence, AR(4) is the best model to forecast the Indonesian real GDP by far. At this stage, we can consider AR(4) as the best candidate to forecast the *y* series based on the MSFE.

**III(e).** **IMA(1) and IMA(1,4) Model Fitting**

From the time series plot given in Figure 1, we can see that the series might not stationary in variance. The said problem can probably be resolved by fitting the MA(q) process. Since the series is seemingly not stationary, the original series was differenced prior to fit the MA(q) process. Therefore, the model I considered to fit is IMA(1,1):

Similarly, as the series I used in this paper is quarter observation, I also consider fitting the IMA(1,4). After fitting the two models and forecast the series using the out-sample method, I obtained the corresponding MSFE given in Table 4.

Table 5. Goodness-of-fit of IMA(1,1) and IMA(1,4)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | IMA(1,1) | IMA(1,4) | AR(4) | Random Walk |
| MSFE | 0.0356 | 0.0193 | 0.0032 | 0.0392 |

Based on the Table 5, it can be said that the AR(4) still the best fitted model to forecast the Indonesian real GDP consumption as it produce the smallest MSFE compared to other proposed models. Both newest proposed model IMA(1,1) and IMA(1,4) does not improve the forecast results of the Indonesian real GDP consumption based on the MSFE value. In this case, even after differencing the series, the differenced process does not able to surpass the performance of AR(4) to forecast the *y* series. Furthermore, it seems that fitting both AR(p) and MA(q) using degree 4 in this series is a better choice.

**III(f). IAR (1, 1)**

As I already fitted the IMA(1,1) and IMA(1,4) in the previous section, I consider to fit the IAR(*d, p*) assuming that after fitting the AR(p) model given in section III(d), the errors have not achieved stationary. As suggested in the previous section, using degree 4 for p and q in this series might be better. However, as it is better to perform the model from the lower degree i.e 1, therefore, to forecast the Indonesian GDP consumption, I fitted the IAR(1,1). The MSFE for the fitted model is given in Table 6.

Table 6. Goodness of fit of the IAR(1,1)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | IAR(1,1) | IMA(1,1) | IMA(1,4) | AR(4) | Random Walk |
| MSFE | 0.0325 | 0.0356 | 0.0193 | 0.0032 | 0.0392 |

The MSFE of IAR(1,1) suggest that the model does not improve the forecast significantly compared to the best identified model by far which is AR(4). At this state, it can be said that applying the first difference into the series does not significantly fixed the error in forecasting the Indonesian real GDP consumption.

**III(e). ARMA(p,q)**

The forecasting outputs in section III(a) until III(d) suggest that differencing the series in this case does not improve the forecasting results. But, it is also important to check whether fitting the ARMA(*p,q*) model can eliminate the correlation between the data and the error in the Indonesian real GDP consumption series. here, I determined the *p* and *q* order for the ARMA model by examining the ACF and PACF plots. The ACF and PACF plots for the series given below

|  |  |
| --- | --- |
| Figure 3. ACF plots of the Indonesian GDP consumption | Figure 4. PACF plots of the Indonesian GDP consumption |

Based on the ACF and PACF plots given above, it is obvious that the ACF of the series are exponentially decay to zero and the PACF cuts off at fourth lag. Such condition suggest that to eliminate the autocorrelation in the series, we can apply the AR(p) model to the series. In this case, I already fitted the AR(4) process to the series by which I obtained a precise forecast.

However, in order to make sure that fitting AR(p) model are already sufficient to explain the behaviour of the given series, I considered to for the ARMA(1,1), ARMA(1,4) and ARMA(2,2). After fitted the proposed models into the series, the MSFE of the three models are given below

Table 7. MSFE of the ARMA(*p, q*) model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | ARMA(1,1) | ARMA(2,2) | ARMA(1,4) | AR(4) |
| MSFE | 0.0166 | 0.0145 | 0.0074 | 0.0032 |

Table 7 given above indicate that all the proposed ARMA(*p,q*) models could not surpass the performance of AR(4) model to forecast the quarterly Indonesian real GDP consumption. Furthermore, the outputs further confirm that AR(4) is the best model to fitted into the series used in this study.

1. **Conclusion**

Based on the outputs of the time series modelling given in the results section, it can be concluded that

1. Before fitting more sophisticated time series models into the real Indonesian GDP consumption, the random walk process could capture the behaviour of the given series. Hence, it has a fairly small MSFE to forecast the 1-step-ahead the real Indonesian GDP consumption.
2. In this study, I only consider to assessed the 1-step-ahead forecast as I set random walk as the benchmark model.
3. After fitting several time series models, the AR(4) produced the best 1-step-ahead forecasting outputs for the real Indonesian GDP consumption based on the MSFE error measurement. As AR(4) process is the best fitted model that can capture and forecast the real Indonesian GDP consumption, this specification will be used as the forecast model for the given series.
4. The fitted model and the one year forecasts using the 1-step-ahead by AR(4) model is given below

* The 95% confidence interval for the third quarter (the 1-ahead-forecast) is given below

|  |  |
| --- | --- |
|  | AR(4) |
|  | 0.0023 |
| Confidence Interval | (15.47087, 15.65886) |

* Point forecast for the following one year given

, ,

Table 10. One year quarterly forecast for

the real Indonesian GDP consumption

|  |  |
| --- | --- |
| Month | AR(4) |
| December 2019 | 15.56487 |
| March 2020 | 15.63657 |
| June 2020 | 15.63884 |
| September 2020 | 15.89088 |

1. In this study, I did not consider to fit the VAR model since the AR(4) model already gives a significantly small MSFE.

**Appendices**

**Appendix A1. Matlab codes for all the analysis given in the paper**

|  |
| --- |
| %% Graphical Data Analysis  clear  GDP = readtable ('Consumption.csv');  y = GDP (29:end,3) ;  y = table2array (y);  plot(y);  xlabel('t');  ylabel('GDP Consumption');  %% Random Walk Process (Benchmark)  randwalk=(T0:T)';  for t=T0:T  randwalk(t)=y(t-1);  end  randwalk;    %% Random walk full  T0=2;  randwalk.full=(T0:T)';  for t=T0:T  randwalk.full(t)=y(t-1);  end  randwalk.full;    err= (y(2:end) - randwalk.full(2:end)).^2;  MSE.rand = mean(err);  k=1;  AICrand = (T-1)\*MSE.rand + 1\*2;  BICrand = (T-1)\*MSE.rand + 1\*log(T-1);  goodness3 =[MSE3 AIC3 BIC3];  plot(2:T, randwalk.full(2:T))  randwalk.full(t)  %% MSFE Benchmark Model  MSFE\_hist= mean((ytph-yhat\_mean(50:106)).^2);  MSFE\_1 = mean(MSFE\_hist);  MSFE\_rand = mean((ytph-randwalk(30:end)).^2);    %% Plotting Model  plot(y);  hold on  plot ((T0:T), yhat\_mean(T0:T));  hold on  plot ((T0:T), randwalk(T0:T));  xlabel('t');  ylabel('GDP Consumption');  legend('actual GDP Consumption', 'Random Walk Forecast');  %% Seasonality Model with dummy variable for the fourth quarter  t = (1:T)';  X2 = [ ones(T,1) t D4];  betahat2 = (X2'\*X2)\(X2'\*y);  yhat2 = X2\*betahat2;  MSE2 = mean((y-yhat2).^2);  AIC2 = T\*MSE2 + 5\*2;  BIC2 = T\*MSE2 + 5\*log(T);    T0=40;  syhat2 = zeros(T-h-T0+1,1); %66 observations  ytph2 = y(T0+h:end); %66 observations    h=1;  %h=3;    for t= T0:T-h  yt = y(1:t);  D1t =D1(1:t); D2t = D2(1:t);  D3t =D3(1:t); D4t = D4(1:t);    Xt2 = [ones(t,1) (1:t)' D4t];  beta2 = (Xt2'\*Xt2)\(Xt2'\*yt);  yhat2 = [1 t+h D4(t+h)]\*beta2;  syhat2(t-T0+1) = yhat2;    end    MSFE2 = mean((ytph2-syhat2).^2);  plot(y, 'k');  title('GDP Index Consumption');  xlabel('t');  ylabel('Consumption Index');  hold on  plot((T0+h:T), syhat2);  legend('GDP Consumption', '3-step ahead forecast ');  set(legend,'location','best')  %% Seasonality Models Dummy for all the quarter  clear  GDP = readtable ('Consumption.csv');  y = GDP (29:end,3) ;  y = table2array (y);  x = (1:length(y))';  plot(y);  hold on  xlabel('t');  ylabel('GDP Consumption');  title('Quarterly Indonesian GDP Consumption');    %% Quartely Dummy Variable  Q = GDP(29:end,4); Q= table2array(Q);  T = length(y); t = (1:T)';  %% construct 4 dummy variables  D1 = (Q == 1); D2 = (Q == 2);  D3 = (Q == 3); D4 = (Q == 4);    %% First Model specification  X = [t D1 D2 D3 D4];  betahat = (X'\*X)\(X'\*y);  yhat = X\*betahat;  MSE = mean((y-yhat).^2);  AIC = T\*MSE + 5\*2;  BIC = T\*MSE + 5\*log(T);    %% MSFE 1-step-ahead Dummy for the fourth quarter (Jelek)  T0= 40;  h = 1;  %h = 3;  syhat1 = zeros(T-h-T0+1,1); %66 observations  ytph1 = y(T0+h:end); %66 observations    for t= T0:T-h  yt = y(1:t);  D1t =D1(1:t); D2t = D2(1:t);  D3t =D3(1:t); D4t = D4(1:t);    Xt1 = [(1:t)' D1t D2t D3t D4t];  beta1 = (Xt1'\*Xt1)\(Xt1'\*yt);  yhat1 = [t+h D1(t+h) D2(t+h) D3(t+h) D4(t+h)]\*beta1;  syhat1(t-T0+1) = yhat1;    end    MSFE1 = mean((ytph1-syhat1).^2);  plot(y, 'k');  title('GDP Index Consumption');  xlabel('t');  ylabel('Consumption Index');  hold on  plot((T0+h:T), syhat1);  legend('GDP Consumption', '3-step ahead forecast ');  set(legend,'location','best')  %% Holt-Winter Forecasting  clear  GDP = readtable ('Consumption.csv');  y = GDP (29:end,3) ; y = table2array (y);  x = (1:length(y))';  T = length(y); t = (1:T)';    T0 = 40;  %h=1;  h=3;  sq=4;  syhatq= zeros(T-h-T0+1,1);  ytphq = y(T0+h:end);  %alpha = 0.0285; beta=0.0465; gamma= 0.4144;  alpha=0.3; beta=0.3; gamma=0.3;  St = zeros(T-h,1);  %Initialize  Lt=mean(y(1:sq)); bt=0; St(1:sq)= y(1:sq)-Lt;    for t = sq+1:T-h  newLt = alpha\*(y(t) - St(t-sq)) + (1-alpha)\*(Lt+bt);  newbt = beta\*(newLt-Lt) + (1-beta)\*bt;  St(t)= gamma\*(y(t)-newLt)+(1-gamma)\*St(t-sq);  yhatq=newLt + h\*newbt + St(t+h-sq);  Lt = newLt; bt=newbt;  if t>= T0  syhatq(t-T0+1,:)= yhatq;  end  end  MSFEq=mean((ytphq-syhatq).^2);    %% Plotting the 1-step-ahead forecast  plot(y, 'k');  title('GDP Index Consumption');  xlabel('t');  ylabel('Consumption Index');  hold on  plot((T0+h:T), syhatq);  legend('GDP Consumption', '1-step ahead forecast ');  set(legend,'location','best')  %% AR(1)  clear  GDP = readtable ('Consumption.csv');  y = GDP (29:end,3); y = table2array (y);  m=4;  h=1;  T0 = 40;  y0 = y(1:m); y = y(m+1:end); T = length(y);  ytph = y(T0+h: end);  yhatAR = zeros(T-h-T0+1,1);  h=1;  for t = T0:T-h  yt = y(h:t);  zt = [ones(t-h+1,1) [y0(m); y(1:t-h)]];  betahat = (zt'\*zt)\(zt'\*yt);  yhatAR(t-T0+1,:)= [1 y(t) ]\*betahat;  end  MSFE\_AR1 = mean((ytph - yhatAR).^2); %332    %% MSE, AIC, BIC for AR(1)    GDP = readtable ('Consumption.csv');  y = GDP (29:end,3) ; y = table2array (y);  T = length(y);  Z = [ones(T-1,1) y(1:T-1)];  betahat = (Z'\*Z)\(Z'\*y(2:end));  yhatAR1= Z\*betahat;  MSE\_AR1 = mean((yhatAR1-y(2:end)).^2);  k=2;  AIC\_AR1 = MSE\_AR1\*(T) + k\*2;  BIC\_AR1 = MSE\_AR1\*(T-1)+ k\*log(T-1);  %% AR(4)  clear  GDP = readtable ('Consumption.csv');  y = GDP (29:end,3) ; y = table2array (y);    %% %Covariance Plot data  [cov\_y, lags\_y] = xcov(y, 10, 'coeff');  stem(lags\_y, cov\_y);    x = (1:length(y))';  T = length(y); t = (1:T)';  m=4;  T0=40-4;  h=1;  %h=3;  y0 = y(1:m); y = y(m+1:end); T = length(y);  ytph = y(T0+h: end);  yhatAR = zeros(T-h-T0+1,1);  T0 = 40;    for t = T0:T-h  yt = y(h:t);  zt = [ones(t-h+1,1) [y0(m); y(1:t-h)] [y0(m-1:end); y(1:t-h-1)]...  [y0(m-2:end); y(1:t-h-2)] [y0(m-3:end); y(1:t-h-3)]];  betahat = (zt'\*zt)\(zt'\*yt);  yhatAR(t-T0+1,:)= [1 y(t) y(t-1) y(t-2) y(t-3)]\*betahat;  end  MSFE\_AR4 = mean((ytph - yhatAR).^2); %332    %% MSE, AIC, BIC for AR(4)  MSFE\_AR4 = mean((ytph-yhatAR).^2); %332    GDP = readtable ('Consumption.csv');  y = GDP (29:end,3) ; y = table2array (y);  T = length(y);  Z = [ones(T-4,1) y(4:T-1) y(3:T-2) y(2:T-3) y(1:T-4)];  betahat = (Z'\*Z)\(Z'\*y(5:end));  yhatAR4= Z\*betahat;  MSE\_AR4 = mean((yhatAR4-y(5:end)).^2);  k=5;  AIC\_AR4 = MSE\_AR4\*(T-3) + k\*2;  BIC\_AR4 = MSE\_AR4\*(T-4)+ k\*log(T-4);    %% Sigma square  Sigsquare = ((yhatAR4-y(5:end)).^2);  sum = sum(Sigsquare);  sig = sum/(T-1);  %% Wold Modelling IMA(1,1) Process  clear  GDP = readtable ('Consumption.csv');  y = GDP (29:end,3) ; y = table2array (y);  x = (1:length(y))';  m = 4; %% the first m points as lags  h = 1;  T0 = 30;    dely = y(m+h:end) - y(m:end-h);  y0 = y(1:m);  y = y(m+1:end);  T = length(y);    plot (y)  hold on    % h step ahead forecast  yhatMA = zeros(T-h-T0+1,1); %% IMA(1,1) forecasts  ytph = y(T0+h:end); % observed y {t+h}  f = @(theta) loglike\_MA1(theta,dely(1:T0));  thetaMLE = fminsearch(f,[0.5, 0.5]);    % forecast  for t = T0:T-h  delyt = dely(1:t);    % find the MLE  f = @(theta) loglike\_MA1(theta,delyt);  thetaMLE = fminsearch(f, thetaMLE);    % prediction  H = speye(t) + sparse(2:t,1:t-1,ones(1,t-1),t,t)\*thetaMLE(2);  uhat = H\delyt;  % store the forecasts  yhatMA(t-T0+1,:) = y(t) + thetaMLE(2)\*uhat(end);  end  MSFE\_MA3 = mean((ytph-yhatMA).^2);  plot(y, 'k');  title('GDP Index Consumption');  xlabel('t');  ylabel('Consumption Index');  hold on  plot((T0+h:T), yhatMA);  legend('GDP Consumption', '1-step ahead forecast ');  set(legend,'location','best')  %% Wold Modelling IMA(1,4) Process  clear  GDP = readtable ('Consumption.csv');  y = GDP (29:end,3) ; y = table2array (y);  x = (1:length(y))';  m = 4; %% the first m points as lags  h = 1;  T0 = 40;    dely = y(m+h:end) - y(m:end-h);  y0 = y(1:m);  y = y(m+1:end);  T = length(y);    plot (y)  hold on    % h step ahead forecast  yhatMA = zeros(T-h-T0+1,1); %% IMA(1,4) forecasts  ytph = y(T0+h:end); % observed y {t+h}  f = @(theta) loglike\_MA4(theta,dely(1:T0));  thetaMLE = fminsearch(f,[0.5, 0.5, 0.5, 0.5, 0.5]);    % forecast  for t = T0:T-h  delyt = dely(1:t);    % find the MLE  f = @(theta) loglike\_MA4(theta,delyt);  thetaMLE = fminsearch(f, thetaMLE);    % prediction  H = speye(t) + sparse(2:t,1:t-1,ones(1,t-1),t,t)\*thetaMLE(2) + sparse(3:t,1:t-2,ones(1,t-2),t,t)\*thetaMLE(3) + sparse(4:t,1:t-3,ones(1,t-3),t,t)\*thetaMLE(4) + sparse(5:t,1:t-4,ones(1,t-4),t,t)\*thetaMLE(5);  uhat = H\delyt;    % store the forecasts  yhatMA(t-T0+1,:) = y(t) + thetaMLE(2)\*uhat(end) + thetaMLE(3) \* uhat(end-1)+ thetaMLE(4) \* uhat(end-2)+ thetaMLE(5) \* uhat(end-3);  end    MSFE\_MA4 = mean((ytph-yhatMA).^2);  plot(y, 'k');  title('GDP Index Consumption');  xlabel('t');  ylabel('Consumption Index');  hold on  plot((T0+h:T), yhatMA);  legend('GDP Consumption', '1-step ahead forecast ');  set(legend,'location','best')  %% IAR (1,1) Process  clear  GDP = readtable ('Consumption.csv');  y = GDP (29:end,3); y = table2array (y);    dely = y(2:end) - y(1:end-1);  m=3; T0 = 40; h=1;  % y0 = y(1:m); y = y(m+1:end);  N = length(y);  dely0 = dely(1:m); dely = dely(m+1:end); T = length(dely);  ydelhat = zeros(T-h-T0+1,1);  yhat\_IAR = zeros(T-h-T0+1,1);  ytph = y(m+T0+1: end);    for t = T0:T  delyt = dely(h:t);  zt = [ones(t-h+1,1) [dely0(m); dely(1:t-h)]];  betahat = (zt'\*zt)\(zt'\*delyt);  ydelhat(t-T0+1,:) = [1 dely(t) ]\*betahat;  yhat\_IAR(t-T0+1,:) = ydelhat(t-T0+1,:) + y(m+t);  end    MSFE\_IAR1 = mean((ytph - yhat\_IAR).^2);  plot(y, 'k');  title('GDP Index Consumption');  xlabel('t');  ylabel('Consumption Index');  hold on  plot((T0:T), yhat\_IAR);  legend('GDP Consumption', '1-step ahead forecast ');  set(legend,'location','best')  %% ARMA (1,1) Process  clear  GDP = readtable ('Consumption.csv');  y = GDP (:,3) ; y = table2array (y);  x = (1:length(y))';  m = 4; %% the first m points as lags  h = 1;  T0 = 30;    hold on  plot(y);    y0 = y(1:m); y = y(m+1:end); T = length(y);  yhatARMA11 = zeros(T-h-T0+1,1);  % theta = [phi1 phi2 mu psi];  f = @(theta) loglike\_ARMA11(theta,y0,y(1:T0));  thetahat = fminsearch(f, [.5;0;.5]);    for t = T0:T-h  yt = y(1:t);  % MLE  f = @(theta) loglike\_ARMA11(theta, y0, yt);  thetahat = fminsearch(f, thetahat);  % make uhat  H = speye(t) + spdiags(ones(t-1,1),[-1],t,t)\*thetahat(3);  X = [[y0(m); y(1:t-1)] ones(t,1)];  uhat = H\(yt - X\*[thetahat(1) thetahat(2)]');  % store  yhatARMA11(t-T0+1,:) = thetahat(2) + thetahat(1)\*y(t) + thetahat(3)\*uhat(end);  end    ytph = y(T0+h:end); % observed y {t+h}  MSFE\_ARMA11 = mean((ytph-yhatARMA11).^2);  plot(y)  hold on  plot((T0+h:T), yhatARMA11);  m = 4;  GDP = readtable ('Consumption.csv');  y = GDP (:,3) ; y = table2array (y);  plot(y);  hold on  y0 = y(1:m); y = y(m+1:end); T = length(y);  T0 = 40; h = 1;  yhatARMA22 = zeros(T-h-T0+1,1);  % theta = [phi1 phi2 mu psi];  f = @(theta) loglike\_ARMA22\_new(theta,y0,y(1:T0));  thetahat = fminsearch(f, [.5;.5;0;.5;.5]);    for t = T0:T-h  yt = y(1:t);  % MLE  f = @(theta) loglike\_ARMA22\_new(theta, y0, yt);  thetahat = fminsearch(f, thetahat);  % make uhat  H = speye(t) + spdiags(ones(t-1,1),[-1],t,t)\*thetahat(4)...  +spdiags(ones(t-2,1),[-2],t,t)\*thetahat(5);  X = [[y0(m); y(1:t-1)] [y0(m-1:end); y(1:t-2)] ones(t,1)];  uhat = H\(yt - X\*[thetahat(1) thetahat(2) thetahat(3)]');  % store  yhatARMA22(t-T0+1,:) = thetahat(3) + thetahat(1)\*y(t) + ...  thetahat(2)\*y(t-1) + thetahat(4)\*uhat(end)+ thetahat(5)\*uhat(end-1);  end  plot((T0+h:T), yhatARMA22);  ytph = y(T0+h:end);  MSFA\_ARMA22 = mean((ytph-yhatARMA22).^2);  %% ARMA(1,4) Process  clear    GDP = readtable ('Consumption.csv');  y = GDP (29:end,3) ;  y = table2array (y);  m = 4;  y0 = y(1:m);  y = y(m+1:end);    plot (y)  hold on    T = length(y);  T0 = 40;  h = 1; % h?step?ahead forecast  yhatMA14 = zeros(T-h-T0+1,1); %% ARMA(2,1) forecasts  %theta = [phi\_1, phi\_2, mu, psi, sigma2]  f = @(theta) loglike\_ARMA14(theta,y0,y(1:T0));  thetahat = fminsearch(f,[.5 ; .5 ; 0.5 ; 0.5 ; 0.5; 0.5; 0.5]);  for t = T0:T-h  yt = y(1:t);  % MLE  f = @(theta) loglike\_ARMA14(theta, y0, yt);  thetahat = fminsearch(f, thetahat);  % make uhat  H = speye(t) + spdiags(ones(t-1,1),[-1],t,t)\*thetahat(3)...  +spdiags(ones(t-2,1),[-2],t,t)\*thetahat(4) + spdiags(ones(t-3,1),[-3],t,t)\*thetahat(5)...  +spdiags(ones(t-4,1),[-4],t,t)\*thetahat(6);  X = [[y0(m); y(1:t-1)] ones(t,1)];  uhat = H\(yt - X\*[thetahat(1) thetahat(2)]');  % store  yhatARMA14(t-T0+1,:) = thetahat(2) + thetahat(1)\*y(t)...  + thetahat(3)\*uhat(end)+ thetahat(4)\*uhat(end-1)+ thetahat(5)\*uhat(end-2) ...  + thetahat(6)\*uhat(end-3) ;  end    ytph = y(T0+h:end); % observed y {t+h}  MSFE\_ARMA14 = mean((ytph-yhatARMA14).^2);  plot(y, 'k');  title('GDP Index Consumption');  xlabel('t');  ylabel('Consumption Index');  hold on  plot((T0+h:T), yhatARMA14);  legend('GDP Consumption', '1-step ahead forecast ');  set(legend,'location','best') |

**Appendix A2. Log Likelihood**

|  |
| --- |
| %% negative of the log-lilkelihood for IMA(1,1)  %% input: x = [sig, theta(1)]; y = data  function ell = loglike\_MA1(x,y)  sig = x(1); theta1 = x(2);  T = length(y);  A = speye(T);  B = sparse(2:T,1:T-1,ones(1,T-1),T,T);  Gam = A + B\*theta1;  Gam2 = Gam\*Gam';  ell = -T/2\*log(2\*pi\*sig)-.5\*log(det(Gam2))-.5/sig\*y'\*(Gam2\y);  ell = - ell;  %% IMA(1,4) Process  %% negative of the log-lilkelihood for MA(4)  %% input: x = [sig, theta(1) theta(2) theta(3) theta(4)]; y = data  function ell = loglike\_MA4(x,y)  sig = x(1); theta1 = x(2); theta2 = x(3); theta3 = x(4); theta4 = x(5);  T = length(y);  A = speye(T);  B = sparse(2:T,1:T-1,ones(1,T-1),T,T);  C = sparse(3:T,1:T-2,ones(1,T-2),T,T);  D = sparse(4:T,1:T-3,ones(1,T-3),T,T);  E = sparse(5:T,1:T-4,ones(1,T-4),T,T);  Gam = A + B\*theta1 + C\*theta2 + D\*theta3 + E\*theta4;  Gam2 = Gam\*Gam';  ell = -T/2\*log(2\*pi\*sig)-.5\*log(det(Gam2))-.5/sig\*y'\*(Gam2\y);  ell = - ell;  % negative of the log likeliood for ARMA(1,1)  % input: x = [phi1 mu psi]  % input: y = data; y0 = lags;  function ell = loglike\_ARMA11(x,y0,y)  phi = zeros(2,1);  phi(1) = x(1);  phi(2) = x(2); psi = x(3);  N = length(y); m = length(y0);  A = speye(N);  B = sparse(2:N, 1:N-1, ones(1,N-1), N,N);  gam = A + B\*psi;  gam2 = gam\*gam';  X = [[y0(m);y(1:N-1)] ones(N,1)];  ell = -(y-X\*phi)'\*(gam2\(y-X\*phi));  ell = -ell;  % negative of the log likeliood for ARMA(2,2)  % input: x = [phi1 phi2 mu psi1 psi2]  % input: y = data; y0 = lags;  function ell = loglike\_ARMA22\_new(x,y0,y)  phi = zeros(3,1);  phi(1) = x(1); phi(2) = x(2);  phi(3) = x(3); % phi(3) = mu  psi1 = x(4); psi2 = x(5);  N = length(y); m = length(y0);  A = speye(N);  B = sparse(2:N, 1:N-1, ones(1,N-1), N,N);  C = sparse(3:N, 1:N-2, ones(1,N-2), N,N);  gam = A + B\*psi1 + C\*psi2;  gam2 = gam\*gam';  X = [[y0(m);y(1:N-1)] [y0(m-1:end);y(1:N-2)] ones(N,1)];  ell = -(y-X\*phi)'\*(gam2\(y-X\*phi));  ell = -ell;  %% ARMA(1,4) Process  % negative of the log likeliood for ARMA(1,4)  % input: x = [phi1 mu psi(1) psi(2) psi(3) psi(4) sig]  % input: y = data; y0 = lags;  function ell = loglike\_ARMA14(x,y0,y)  phi = zeros(2,1);  phi(1) = x(1);  phi(2) = x(2); %mu  psi (1) = x(3);  psi (2) = x(4);  psi (3) = x(5);  psi (4) = x(6);  sig2 = x(7);  N = length(y); m = length(y0);  A = speye(N);  B = sparse(2:N, 1:N-1, ones(1,N-1), N,N);  C = sparse(3:N, 1:N-2, ones(1,N-2), N,N);  D = sparse(4:N, 1:N-3, ones(1,N-3), N,N);  E = sparse(5:N, 1:N-4, ones(1,N-4), N,N);  gam = A + B\*psi(1)+C\*psi(2) + D\*psi(3) + E\*psi(4);  gam2 = gam\*gam';  X = [[y0(m);y(1:N-1)] ones(N,1)];  ell = -(y-X\*phi)'\*(gam2\(y-X\*phi));  ell = -ell; |

**Appendix A3. Out-of-Sample Forecasting Plots**

|  |  |
| --- | --- |
| **Appendix A3(1). 1-step-ahead forecast using Random Walk Process**    **Appendix A3(3). 3-step-ahead forecast using specification given in section III(b)**    **Appendix A4. 3-step-ahaed forecast using the second specification given in section III(B)**    **Appendix A5. 1-step-ahead IMA(1,1)** | **Appendix A3(2). 1-step-ahead forecast using the first specification given in section III(b)**    **Appendix A3(4). 1-step-ahead forecast using the second specification given in section III(b)**    **Appendix A4. 1-step-ahead Holt-Winters quarterly seasonality smoothing**    **Appendix A6. 1-step-ahead IMA(1,4)** |