Time Series Analysis Project

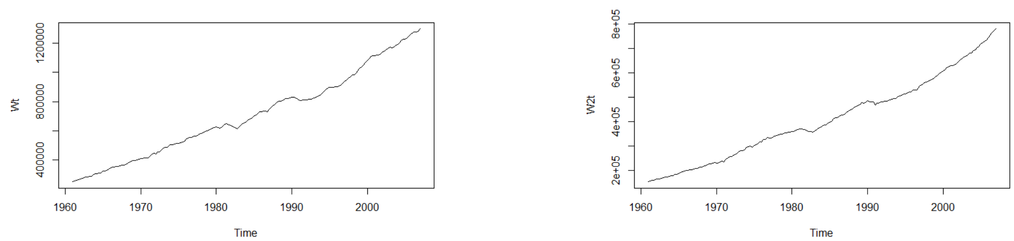
Alice Li

July 27, 2017

Contents

1 Introduction and Graphical Analysis

Consider the raw quarterly seasonally adjusted Canadian GDP series W_t, and the raw seasonally adjusted Canadian Personal Expenditure on Consumer Goods and Services time series W_{2t}. The data range covers the first quarter of the year 1961 till the first quarter of 2007, i.e. 185 observations. The figures below is a graphical representation of W_t and W_{2t}:



(a) W_t

(b) W_{2t}

Figure 1: Seasonally Adjusted Real GDP (a) and Consumption (b) Time Series

2 Extracting the Cycle using TS/DS Approaches

We now perform a log-linear transformation of the raw GDP time series W_t, and we consider two approaches to model the cycle Y_t: *Trend Stationary Approach (TS)* and *Difference Stationary Approach (DS)*.

Using the TS approach, the cycle Y_t is extracted as the residual from the regression:
$X_t = log(W_t) = \alpha + \mu t + Y_t$, where μ is the growth rate, and Y_t is the residual.
The regression results are reported as follows:

$$\underset{(t)}{X_t} = \underset{(1298.21)}{12.61} + \underset{(91.26)}{0.00826t} + Y_t \tag{1}$$

$$n = 185 \quad F-ratio = 8328 \quad RSS = 0.0658^2 \quad R^2 = 0.9785$$

Notice that both coefficients are highly significant, as shown from the t-statistics figures in brackets. The quarterly growth rate 0.83% per quarter, or 3.32% per year. Also, because $\alpha = 12.61$, we can say the economic series has a positive long-run trend.

Using the DS approach, the cycle Y_t is extracted as the residual from the regression:
$\Delta X_t = \mu + Y_t$ where μ is the growth rate and ΔX_t is the differenced time series of X_t.
The regression results are reported as follows:

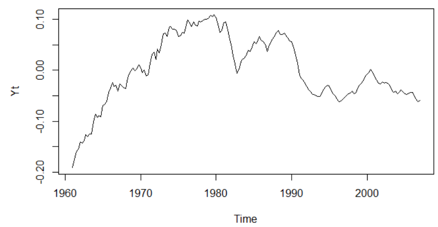
$$\underset{(t)}{\Delta X_t} = \underset{(13.84)}{0.00898} + Y_t \tag{2}$$

The estimated growth rate is 0.90% per quarter, or equivalently 3.60% per year.
The extracted cycle Y_t for both approaches are plotted in the figure below. Notice in plot(a), Y_t is above zero for around 1975-1980 and 1985-1990, indicating expansion in those periods.

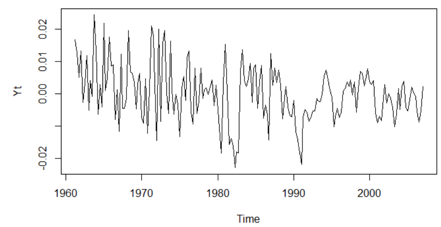
3 Fitting of an AR(p) process

3.1 Select Optimal Model

We now assume the cycle Y_t follows an $AR(p)$ process as defined as follows:



(a) TS



(b) DS

Figure 2: Extracted Cycle Y_t from the TS(a) and DS(b) Approach

Definition 1. *Y_t follows a p-th order autoregressive process or $Y_t \sim AR(p)$ if*

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + a_t, \quad where\ a_t \sim i.i.N(0, \sigma^2)$$

Now we try to fit both the TS and DS Y_t using the Bayesian Information Criteria (BIC) to decide on the approprate value of p for the $AR(p)$. The BIC(defined below) is constructed to choose the model which maximizes the trade-off between fit and parsimony. The following table summarizes the $BIC(k)$ results for both TS and DS.

Definition 2. *For an $AR(k)$ process*

$$BIC(k) = log(\hat{\sigma_k}^2) + \frac{log(N) \times k}{N}$$

where N is the number of effective observations and $\sigma_k{}^2$ is the standard deviance of the residual of the AR(k).The BIC estimator of p is $\hat{p_B}$, the value of k that minimizes BIC(k).

Lag k	$BIC(k)$ for TS	$BIC(k)$ for DS
0	-5.454	-9.469
1	-9.437	-9.554^*
2	-9.527^*	-9.530
3	-9.503	-9.517
4	-9.492	-9.489
5	-9.464	-9.461
6	-9.437	-9.438
7	-9.413	-9.415
8	-9.392	-9.396
9	-9.371	-9.416

Table 1: $BIC(k)$ values for TS and DS

From the above table, we find that for the TS model, $BIC(k)$ has a minimum at $k = 2$. For the DS model, $BIC(k)$ has a minimum at $k = 1$. As a result, for the purpose of this report, we will choose $AR(2)$ as the optimal AR model for both TS and DS.

3.2 TS model

By fitting the $AR(2)$ model with Y_t from the TS model, we get the following:

$$Y_t = \underset{(0.070)}{1.290} Y_{t-1} - \underset{(0.0685)}{0.313} Y_{t-2} + a_t, \quad where\ a_t \sim i.i.N(0, \sigma^2) \tag{3}$$

and $\hat{\sigma}^2 = 6.51 \times 10^{-5}$

Note that $\hat{\phi_1} = 1.290$ and $\hat{\phi_2} = -0.313$, and the values of standard errors of these estimates are included in the brackets underneath.
Now let's check if this AR(2) process is stationary, i.e. if it satisfies the following necesarry conditions for stationarity:
(1) $|\phi_p| < 1$ (2) $\phi_1 + \phi_2 + \ldots + \phi_p < 1$ (3) $-\phi_1 + \phi_2 + \ldots + (-1)^p \psi_p < 1$.
Because $|\hat{\phi_2}| = 0.313 < 1$, $\hat{\phi_1} + \hat{\phi_2} = 0.977 < 1$, -$\hat{\phi_1} + \hat{\phi_2} = -1.603 < 1$, we can conclude that this AR(2) process is stationary, which allows us to the next subsection.

3.2.1 Infinite Moving Average Weights

Due to stationarity, Wold Representation indicates that the estimated AR model equation can be written as the sum of infinite moving average weights ψ_k in the form of

$$Y_t = \psi(L)a_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \psi_3 a_{t-3} + \ldots$$

Additionally, for the $AR(p)$ model, given knowledge of the previous values $\psi_{k-1}, \psi_{k-2}, \ldots, \psi_{k-p}$, where $\psi_j = 0$ for any $j < 0$ and $\psi_0 = 1$, we can find the ψ_k values using the p-th order difference equation: $\psi_k = \phi_1\psi_{k-1} + \phi_2\psi_{k-2} + \ldots + \phi_p\psi_{k-p}$.
Hence for the estimated $AR(2)$ model, we have starting values $\psi_0 = 1$ and $\psi_{-1} = 0$ and second order difference equation

$$\psi_k = 1.290\psi_{k-1} - 0.313\psi_{k-2} \tag{4}$$

. The ψ_k values are summarized in the Table 2 below for $k = 0, 1, \ldots, 8$.

3.2.2 Autocorrelations

Similar to the infinite moving average weights, we can find $\rho(k)$ using the p-th difference equation: $\rho_k = \phi_1\rho(k-1) + \phi_2\rho(k-2) + \ldots + \phi_p\rho(k-p)$ where the starting values are $\rho(0) = 1$ and $\rho(-k) = \rho(k)$. For the estimated $AR(2)$, the difference equation is as follows

$$\rho(k) = 1.290\rho(k-1) - 0.313\rho(k-2) \tag{5}$$

with starting values $\rho(0) = 1$ and $\rho(-1) = \rho(1)$ The $\rho(k)$ values are summarized in the table below for $k = 0, 1, \ldots, 8$.

k	ψ_k	$\rho(k)$
0	1	1
1	1.290	0.982
2	1.351	0.954
3	1.338	0.923
4	1.303	0.892
5	1.262	0.861
6	1.219	0.831
7	1.178	0.803
8	1.137	0.775

Table 2: Partial Summary of ψ_k and ρ_k under the TS model

3.2.3 Standard Deviance

Suppose we want to find the standard deviation of Y_t, i.e. $\gamma(0)^{1/2}$. Note that for an AR(p) process, we calculate $\gamma(0)$ with the following the result:

$$\gamma(0) = \frac{\sigma^2}{1 - \phi_1\rho(1) - \phi_2\rho(2) - \ldots - \phi_p\rho(p)}$$

where for $i = 1, 2, 3 \ldots p$.

In the previous subsections, we found that for the estimated model, $\hat{\sigma}^2 = 6.51 \times 10^{-5}$, $\hat{\phi}_1 = 1.290$, $\hat{\phi}_2 = -0.313$, $\rho(1) = 0.982$, and $\rho(2) = 0.954$, therefore

$$\gamma(0) = \frac{\sigma^2}{1 - \phi_1\rho(1) - \phi_2\rho(2)} = 0.00205 \tag{6}$$

Hence $\gamma(0)^{1/2} = 0.0453$.

3.3 DS Model

Similar to Section 3.2, we find the following results for the DS model, assuming it's an AR(2) process.

$$Y_t = \underset{(0.0743)}{0.2970}Y_{t-1} + \underset{(0.0736)}{0.05743}Y_{t-2} + a_t, \quad where\ a_t \sim i.i.N(0, \sigma^2) \tag{7}$$

and $\hat{\sigma}^2 = 6.70 \times 10^{-5}$, $\hat{\phi}_1 = 0.2970$, $\hat{\phi}_2 = 0.05743$.

k	ψ_k	$\rho(k)$
0	1	1
1	0.2970	0.3150
2	0.1456	0.1510
3	0.0603	0.06230
4	0.0263	0.02736
5	0.0113	0.01174
6	0.00485	0.005057
7	0.002088	0.002176
8	0.000899	0.0009365

Table 3: Partial Summary of ψ_k and ρ_k under the DS model

4 Forecasting

4.1 TS Model Forecasting

First we look for the forecasts for the cycle lag k Y_{t+k}. Except for the starting values $E_t[Y_{t+k-j}] = Y_{t+k-j}$ for $j \geq k$, the forecasts $E_t[Y_{t+k}]$ follows the same p-th order difference equation as ψ_k and $\rho(k)$:

$$E_t[Y_{t+k}] = \phi_1 E_t[Y_{t+k-1}] + \phi_2 E_t[Y_{t+k-2}] + \ldots + \phi_p E_t[Y_{t+k-p}] \tag{8}$$

Recall from Section 3.2, $\hat{\phi}_1 = 1.290$ and $\hat{\phi}_2 = -0.313$. Hence, for the estimated $AR(2)$ model, we have

$$E_t[Y_{t+k}] = 1.290 E_t[Y_{t+k-1}] - 0.313 E_t[Y_{t+k-2}] \tag{9}$$

with starting values $E_t[Y_{185}] = Y_{185} = -0.0756$ and $E_t[Y_{184}] = Y_{184} = -0.0794$.

We can also find the variance of the forecasts for Y_{t+k}, by using the formula:

$$Var_t[Y_{t+k}] = \sigma^2 \sum_{j=1}^{k-1} (\psi_j^2) \tag{10}$$

Second, we look for forecasts for the growth rate ΔX_{t+k} using Result 3.31 by Sampson (2013), which shows that if X_t is Trend Stationary, then

$$E_t[\Delta X_{t+k}] = \mu + E_t[Y_{t+k}] + E_t[Y_{t+k-1}] \tag{11}$$

Because we have obtained $\mu = 0.00826$ from equation (1), and can find $E_t[Y_{t+k}]$ and $E_t[Y_{t+k-1}]$ using equation (9), we will be able to find $E_t[\Delta X_{t+k}]$ recursively.
Third, we look for the forecasts for the variance of ΔX_{t+k}. Result 3.31 from Sampson(2013) shows that

$$Var_t[\Delta X_{t+k}] = \sigma^2 [1 + \sum_{j=1}^{k-1} (\psi_j - \psi_{j-1})^2] \tag{12}$$

Because we obtained $\sigma^2 = 6.51 \times 10^{-5}$ from equation (3), and can get ψ_j for $j = 0, 1, \ldots, k$ recursively using equation (4), we will be able to find $Var_t[\Delta X_{t+k}]$ for all k's.

Table 4 displays the forecasts, conditional variances and 95% Confidence Interval(CI) for $k = 0, 1, 2, \ldots, 8$. Figure 3 plots the forecasts and CI.

k	$E_t[Y_{t+k}]$	$E_t[\Delta X_{t+k}]$	$Var_t[Y_{t+k}]$	$Var_t[\Delta X_{t+k}]$	Confidence Interval
0	-0.05864	0.01119	0	0	[0.01119]
1	-0.05636	0.01055	6.51×10^{-5}	6.51×10^{-5}	[-0.00527 , 0.02636]
2	-0.05433	0.01029	1.73×10^{-4}	7.06×10^{-5}	[-0.00617 , 0.02676]
3	-0.05243	0.01016	2.92×10^{-4}	7.08×10^{-5}	[-0.00633 , 0.02666]
4	-0.05062	0.01008	4.09×10^{-4}	7.08×10^{-5}	[-0.00642 , 0.02657]
5	-0.04887	0.01001	5.19×10^{-4}	7.09×10^{-5}	[-0.00649 , 0.02651]
6	-0.04719	0.00995	6.23×10^{-4}	7.10×10^{-5}	[-0.00657 , 0.02646]
7	-0.04556	0.00989	7.20×10^{-4}	7.11×10^{-5}	[-0.00664 , 0.02642]
8	-0.04399	0.00983	8.10×10^{-4}	7.12×10^{-5}	[-0.00671 , 0.02638]

Table 4: Partial Summary of Forecasted Growth Rates, Variance and 95% Confidence Interval under the TS model

4.2 DS Model Forecasting

First we look for the forecasts for the cycle Y_{t+k}. Same as the TS Model, we can find $E_t[Y_{t+k}]$ using the Equation (8). Specifically for the DS model, as we have $\hat{\phi_1} = 0.2970$, $\hat{\phi_2} = 0.05743$, Equation (8) can be written as:

$$E_t[Y_{t+k}] = 0.2970 E_t[Y_{t+k-1}] + 0.05743 E_t[Y_{t+k-2}] \tag{13}$$

with starting values $E_t[Y_{184}] = Y_{184} = 0.002207$ and $E_t[Y_{183}] = Y_{183} = -0.005422$.

$$E_t[Y_{t+k}] = 0.2970 E_t[Y_{t+k-1}] + 0.05743 E_t[Y_{t+k-2}] \tag{14}$$

$Var_t[Y_{t+k}]$ can calculated using equation (10) as defined in the TS Model. Because we obtained $\sigma^2 = 6.70 \times 10^{-5}$ from equation (3), and can get ψ_j for $j = 0, 1, \ldots, k$ recursively, we will be able to find $Var_t[Y_{t+k}]$ for all k's.

Second, we look for forecasts for the growth rate ΔX_{t+k}. Recall regression under DS Model gives $\Delta X_t = \mu + Y_t$. As a result, $E_t[\Delta X_{t+k}]$ written in the following form:

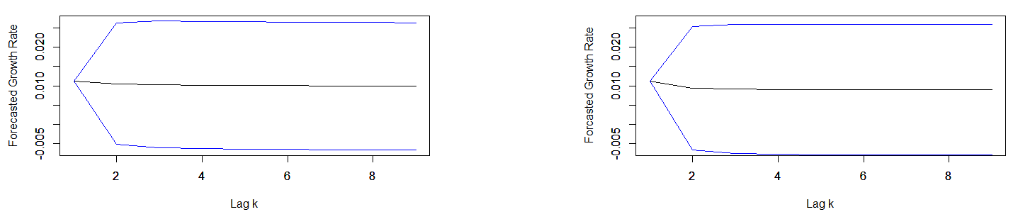
$$E_t[\Delta X_{t+k}] = \mu + E_t[Y_{t+k}] \tag{15}$$

As we have found $\mu = 0.00898$ from Equation (2), and can derive $E_t[Y_{t+k}]$ with Equation (14), we can calculate $E_t[\Delta X_{t+k}]$ recursively.
Third, we look for the forecasts for the variance of ΔX_{t+k}. Since the constant μ in equation $\Delta X_t = \mu + Y_t$ does not affect the variance, we have $Var_t[\Delta X_{t+k}] = Var_t[Y_{t+k}]$.

The table below displays the forecasts for the DS Model, conditional variances and 95% Confidence Interval(CI) for $k = 0, 1, \ldots, 8$. Forecasts and CI are also plotted in Figure 3.

k	$E_t[Y_{t+k}]$	$E_t[\Delta X_{t+k}]$	$Var_t[Y_{t+k}]$	$Var_t[\Delta X_{t+k}]$	Confidence Interval
0	2.2×10^{-3}	0.011193	0	0	[0.011193]
1	3.4×10^{-4}	0.00933	6.70×10^{-5}	6.70×10^{-5}	[-0.006712 , 0.025372]
2	2.3×10^{-4}	0.009215	7.29×10^{-5}	7.29×10^{-5}	[-0.007519 , 0.02595]
3	8.8×10^{-5}	0.009074	7.43×10^{-5}	7.43×10^{-5}	[-0.007823 , 0.025971]
4	3.9×10^{-5}	0.009025	7.46×10^{-5}	7.46×10^{-5}	[-0.007899 , 0.02595]
5	1.7×10^{-5}	0.009003	7.46×10^{-5}	7.46×10^{-5}	[-0.007927 , 0.025933]
6	7.2×10^{-6}	0.008993	7.46×10^{-5}	7.46×10^{-5}	[-0.007937 , 0.025924]
7	3.1×10^{-6}	0.008989	7.46×10^{-5}	7.46×10^{-5}	[-0.007941 , 0.02592]
8	1.3×10^{-6}	0.008988	7.46×10^{-5}	7.46×10^{-5}	[-0.007943 , 0.025918]

Table 5: Partial Summary of Forecasted Growth Rates, Variance and 95% Confidence Interval under the DS model



(a) *TS* (b) *DS*

Figure 3: Extracted Cycle Y_t from the TS(a) and DS(b) Approach

5 Dickey Fuller test

Previously we used the TS and DS approaches to model the cycle Y_t. To determine which approach we should use, we consider the Augmented Dickey Fuller (ADF)test in which null hypothesis H_0 is that the log transformation series X_t is Difference Stationary(DS).
We regress the differenced series ΔX_t on a constant, time trend t, X_{t-1}, and five lagged differences of the series. The regression result output by R using the adf.test function in the tseries package gives $p - value = 0.2865 > 0.05$, indicating we do not have sufficient evidence to reject the null hypothesis. Thus we conclude X_t is DS.

6 Box Jenkins Identification

Another method to determine whether a given series is better described by an AR(p) or an MA(q) model is the Box-Jenkins identification method.
Note that if Y_t is an AR(p), then $\rho(k)$ will be a damped exponential while ϕ_{kk} will have a cutoff at $k = p$. And if Y_t is an MA(q), then $\rho(k)$ will have a cut off at $k = q$ and ϕ_{kk} is a damped exponential. In practice, we estimate $\rho(k)$ and ϕ_{kk} with $\hat{\rho(k)}$ and $\hat{\phi_{kk}}$, which are non-zero, hence we need to verify null hypothesis H_0: $\rho(k) = 0$, $\phi_{kk} = 0$.
Result 8.2 and a simplification of Result 8.1 from Sampson(2013) shows that, under H_0,

$$\hat{\rho(k)} \overset{approx}{\sim} N(0, \frac{1}{n}) \quad and \quad \hat{\phi_{kk}} \overset{approx}{\sim} N(0, \frac{1}{n})$$

Hence by Two-Sigma rule, we take $\hat{\rho(k)}$ and $\hat{\phi_{kk}}$ to be significantly different than zero if

$$\hat{\rho(k)} > \frac{2}{\sqrt{T}}, \quad \hat{\phi_{kk}} > \frac{2}{\sqrt{T}} \tag{16}$$

where T is the number of observations.

6.1 TS model

By using the acf function in R to obtain the $\hat{\rho}(k)$ estimates and fitting Y_t to AR(k) models to get the $\hat{\phi}_{kk}$ estimates, we summarize the results for $k = 1, \ldots, 9$ in the following table:

k	1	2	3	4	5	6	7	8	9
$\rho(k)$	0.9656	0.9293	0.8938	0.857	0.8224	0.787	0.7511	0.7161	0.6805
ϕ_{kk}	0.9656	-0.3131	-0.0552	-0.1223	0.0064	0.0317	-0.0655	-0.0746	0.0891

Table 6: $\hat{\rho}(k)$ and $\hat{\phi}_{kk}$ values under the TS model for $k = 1, \ldots, 10$

Since there are 185 observations, we take $\hat{\rho}(k)$ and $\hat{\phi}_{kk}$ to be significantly different than 0 if they exceed in absolute value

$$2 \times \frac{1}{\sqrt{T}} = 2 \times \frac{1}{\sqrt{185}} = 0.1470.$$

This means that all of the $\hat{\rho}(k)$ are significant, and $\rho(k)$ appears to a damped exponential, so Y_t is an AR(p). As ϕ_{kk} has the cut-off property at k=2, Y_t appears to be an AR(2), aligning with our conclusion in Section 3.1 using BIC. Again, estimating an AR(2), we have:

$$Y_t = \underset{(0.070)}{1.290} Y_{t-1} - \underset{(0.0685)}{0.313} Y_{t-2} + a_t, \quad a_t \sim i.i.N(0, 6.51 \times 10^{-5}) \tag{17}$$

6.2 DS model

Using the same method as the TS model, we summarize $\hat{\rho}(k)$ and $\hat{\phi}_{kk}$ below:

k	1	2	3	4	5	6	7	8	9
$\rho(k)$	0.3246	0.1539	0.1738	0.0819	0.0156	0.0725	0.1025	-0.0253	0.1646
ϕ_{kk}	0.3246	0.0574	0.1239	-0.008	-0.0335	0.0647	0.0723	-0.0938	-0.1635

Table 7: $\hat{\rho}(k)$ and $\hat{\phi}_{kk}$ values under the TS model for $k = 1, \ldots, 10$

Since there are 184 observations, we take $\hat{\rho}(k)$ and $\hat{\phi}_{kk}$ to be significantly different than 0 if they exceed in absolute value

$$2 \times \frac{1}{\sqrt{T}} = 2 \times \frac{1}{\sqrt{184}} = 0.1474.$$

$\hat{\rho}(k)$ appears to cut off at $k = 3$, and ϕ_{kk} has the cut-off property at $k = 2$. We thereby can interpret the results two different ways: Y_t is an AR(1) or Y_t is an MA(3).
Estimating the AR(1) Model, we get

$$Y_t = \underset{(0.069)}{0.32470} Y_{t-1} + a_t, \quad a_t \sim i.i.N(0, 6.75 \times 10^{-5}) \tag{18}$$

Estimating the MA(3) Model, we get

$$Y_t = a_t + \underset{(0.0734)}{0.2966} a_{t-1} + \underset{(0.0725)}{0.1114} a_{t-2} + \underset{(0.0741)}{0.1523} a_{t-2}, \quad a_t \sim i.i.N(0, 6.80 \times 10^{-5}) \tag{19}$$

7 Diagnostic Tests

Using Box Jenkins Identification, we identified the $AR(2)$ under the TS approach, and $AR(1)$ and $MA(3)$ under the DS approach. Now we perform the following 4 diagnostic tests to determine if these are good models:
1) Box-Pierce Test 2) Overfitting for r=4 3) Jarque-Bera 4) a test for ARCH(6).

7.1 AR(2) Under TS Approach

7.1.1 Box-Pierce Test

Definition 3. *The Box-Pierce Test: If* $M \approx \sqrt{T}$ *then under the null hypothesis* $H_0 : \rho_a(k) = 0 \ for \ k = 1, \ldots, M,$ *we have*

$$Q = T \times \hat{\rho_a}(1)^2 + \hat{\rho_a}(2)^2 + \cdots + \hat{\rho_a}(M)^2) \overset{a}{\sim} \chi^2_M$$

We use the Box-Pierce Test to test for serial autocorrelation of the error terms. In our case under the TS approach, with T = 185 observations, we have $M \approx \sqrt{T} = \sqrt{185} \approx 10$.
From R function Box.test, we get

$$Q = T \times \hat{\rho_a}(1)^2 + \hat{\rho_a}(2)^2 + \cdots + \hat{\rho_a}(10)^2) = 17.517 \overset{a}{\sim} \chi^2_{10} \tag{20}$$

and $p - value = P(\chi^2_{10} > 17.517) = 0.06369 > 0.05$
Therefore we accept H_0 that there is no serial correlation using the estimated AR(2) Model. The model passes this diagnostic test.

7.1.2 Overfitting with r=4

Another diagnostic is based on overfitting. If we correctly selected an $AR(p)$ as

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + a_t$$

Then under null hypothesis H_0: $\phi_{p+1} = \phi_{p+2} = \ldots = \phi_{p+r} = 0$, Y_t is also an $AR(p + r)$ as

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \phi_{p+1} Y_{t-(p+1)} + \ldots + \phi_{p+r} Y_{t-(p+r)} + a_t$$

We test this using a likelihood ratio test with statistic

$$\Lambda = T \times log(\frac{\hat{\sigma_p}^2}{\hat{\sigma_{p+r}}^2}) \sim \chi^2_r$$

To test our estimated $AR(2)$ model, we need to fit $AR(2 + 4) = AR(6)$ first, which gives us $\hat{\sigma_6}^2 = 6.28 \times 10^{-5}$. Hence

$$\Lambda = T \times log(\frac{\hat{\sigma_4}^2}{\hat{\sigma_6}^2}) = 179 \times log(\frac{6.51 \times 10^{-5}}{6.28 \times 10^{-5}}) = 6.403 \sim \chi^2_4$$

$$p - value = P(\chi^2_4 > 6.6174) = 0.171 > 0.05$$

Hence we do not reject H_0, meaning there is no evidence the original $AR(2)$ model is incorrect, passing the diagnostic test.

7.1.3 Normality and Jarque-Bera Test

If our model is correct and $a_t \sim N[0, \sigma^2]$, then we would expect the standardized residuals

$$z_t = \frac{1}{\sigma} a_t \sim N[0, 1].$$

Hence we can test normality informally using the standardized residuals test. We take the estimated residuals $\hat{a}_t$, normalize them as $\hat{z}_t = \dfrac{\hat{a}_t}{\hat{\sigma}}$ and plot $\hat{z}_t$.

In our $T = 183$ observations of residuals, we expect to see $|\hat{z}_t| > 2$ for about $183 \times 0.05 \approx 9$ observations, $|\hat{z}_t| > 3$ for about $183 \times 0.003 \approx 1$ observation, and none for $|\hat{z}_t| > 4$. The standardized residuals plot below for the estimated AR(2) model shows consistency with our expectations, therefore normal distribution to appears to be appropriate.

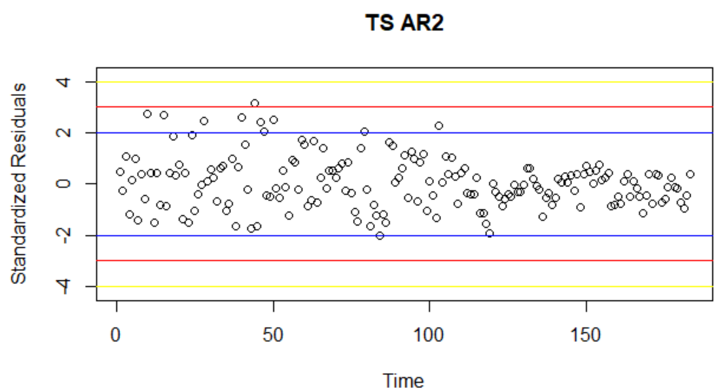


Figure 4: Standardized Residuals Plot for the estimated $AR(2)$ model

Now let's test normality with the Jarque-Bera(JB) Test as defined below.

Definition 4. *The Jarque-Bera Test: Under the joint hypothesis* $H_0 : \kappa_3 = 0,\ \kappa_4 = 0,$

$$JB_{stat} = \hat{t}_3^2 + \hat{t}_4^2 = T(\frac{\hat{\kappa_3}^2}{6} + \frac{(\hat{\kappa_4} - 3)^2}{24}) \overset{a}{\sim} \chi_2^2$$

where $\hat{\kappa_3} = \dfrac{1}{T}\sum_{t=1}^{T} \hat{z_t}^3$ *and* $\hat{\kappa_4} = \dfrac{1}{T}\sum_{t=1}^{T} \hat{z_t}^4$ *are the sample skewness and kurtosis.*

Note that at $\alpha = 5\%$, $JB_{stat} \approx 6$, so we reject H_0 if $JB_{stat} > 6$. Therefore for the estimated $AR(2)$ model, we get $\hat{t}_3 = 4.169$ and $\hat{t}_4 = 1.437$, so

$$JB_{stat} = \hat{t}_3^2 + \hat{t}_4^2 = 19.454 > 6$$

and we reject the null hypothesis of normality, hence fail this diagnostic test.

7.1.4 ARCH(6)

Let $Y_t = \beta_1 + \beta_2 X_{2t} + \ldots + \beta_k X_{kt} + a_t$, where $a_t \sim (0, \sigma^2)$.
Earlier our Box-Pierce pierce test determined linear independence(zero correlation)the error terms a_t. To model the non-linear dependence in a_t, we introduce the q-th order Autoregressive Conditional Heteroskedasticity Model, or $ARCH(q)$ specified as

$$a_t = z_t \times (\sigma^2 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-2}^2 + \ldots + +\alpha_q a_{t-q}^2)^{1/2} \quad where\ z_t \sim i.i.N[0, 1]$$

To test null hypothesis $H_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_q \;\Rightarrow\; a_t = z_t \times \sigma \sim i.i.N[0, \sigma^2]$
Engle (1982) has constructed a test of H_0 (non-linear independence) which consists of running the auxiliary regression

$$\hat{a}_t^2 = \psi_0 + \psi_1 \hat{a}_{t-1}^2 + \ldots + \psi_q \hat{a}_{t-q}^2 + error_t \tag{21}$$

and calculating the R^2. Under H_0, $LM_{stat} = T \times R^2 \overset{a}{\sim} \chi_q^2$.
Setting $q = 6$, we will test

$$H_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_6$$

in

$$a_t = z_t \times (\sigma^2 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-2}^2 + \ldots + +\alpha_6 a_{t-6}^2)^{1/2}.$$

We regress $\hat{a}_t$ on a constant and $\hat{a}_{t-1}, \ldots, \hat{a}_{t-6}$ as

$$\hat{a}_t^2 = \psi_0 + \psi_1 \hat{a}_{t-1}^2 + \ldots + \psi_6 \hat{a}_{t-6}^2 + error_t \tag{22}$$

With $R^2 = 0.1217$ and $T = 183$, we get

$$LM_{stat} = T \times R^2 = 22.262 \sim \chi_6^2$$

and $p - value = P(\chi_6^2 > 22.262) = 0.0010 < 0.05$, so we reject H_0 and conclude there exists non-linear evidence in a_t, hence failing the diagnostic test.

7.2 AR(1) Under DS Approach

Similar to testing the estimated $AR(1)$ model in 7.1, we conduct the 4 diagnostic tests.

7.2.1 Box-Pierce Test

With T = 184 observations, we have $M \approx \sqrt{T} = \sqrt{183} \approx 10$, so we test
$H_0 : \rho_a(k) = 0 \; for \; k = 1, \ldots, 10$, which gives $Q = 17.502 \overset{a}{\sim} \chi_{10}^2$
and $p - value = P(\chi_{10}^2 > 17.502) = 0.06369 > 0.05$.
Hence we accept H_0 that there is no serial correlation using the estimated model. The model passes this diagnostic test.

7.2.2 Overfitting with r=4

We overfit model $AR(1 + 4) = AR(5)$, which gives us $\hat{\sigma_5}^2 = 6.28 \times 10^{-5}$. Hence
$\Lambda = T \times log(\frac{\hat{\sigma_1}^2}{\hat{\sigma_5}^2}) = 180 \times log(\frac{6.75 \times 10^{-5}}{6.50 \times 10^{-5}}) = 6.690 \sim \chi_4^2$
As $p - value = P(\chi_4^2 > 6.690) = 0.1489 > 0.05$, we do not reject H_0, hence passing the diagnostic test.

7.2.3 Normality and Jarque-Bera Test

The standardized residuals plot for the estimated $AR(1)$ model shows consistency with our expectations specified in 7.1.3. Thus normal distribution to appears to be appropriate.
Now let's test normality formally with the Jarque-Bera(JB) Test for the estimated $AR(1)$ model, we get $\hat{t}_3 = 2.760$ and $\hat{t}_4 = 1.233$, so

$$JB_{stat} = \hat{t}_3^2 + \hat{t}_4^2 = 9.1394 > 6$$

and we reject the null hypothesis of normality, hence fail this diagnostic test.

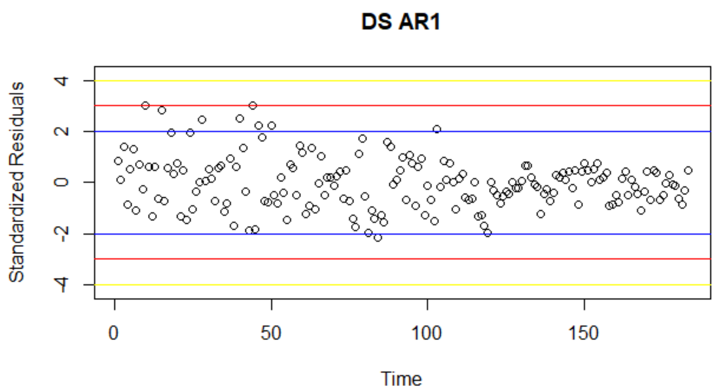


Figure 5: Standardized Residuals Plot for the estimated $AR(2)$ model

7.2.4 ARCH(6)

Performing the same steps in 7.1.4, we get

$$LM_{stat} = T \times R^2 = 183 \times 0.1347 = 24.653 \sim \chi_6^2$$

and $p - value = P(\chi_6^2 > 24.653) = 0.000396 < 0.05$, so we reject H_0 and conclude there exists non-linear correlation in a_t, hence failing the diagnostic test.

7.3 MA(3) Under DS Approach

7.3.1 Box-Pierce Test

With T = 184 observations, we have $M \approx 10$, giving $Q = 12.601 \overset{a}{\sim} \chi_{10}^2$
and $p - value = P(\chi_{10}^2 > 12.061) = 0.281 > 0.05$.
Hence we accept H_0 that there is no serial correlation using the estimated model. The model passes this diagnostic test.

7.3.2 Overfitting with r=4

First we overfit model with $MA(3+4) = MA(7)$, which gives us $\hat{\sigma_5}^2 = 6.28 \times 10^{-5}$.
Testing null hypothesis $H_0 : \theta_4 = \theta_5 = \theta_6 = \theta_7 = 0$,
we get $\Lambda = T \times log(\frac{\hat{\sigma_3}^2}{\hat{\sigma_7}^2}) = 180 \times log(\frac{6.80 \times 10^{-5}}{6.62 \times 10^{-5}}) = 4.842 \sim \chi_4^2$
and $p - value = P(\chi_4^2 > 4.842) = 0.304 > 0.05$.
Hence we do not reject H_0, meaning there is no evidence the original $MA(3)$ model is incorrect, passing the diagnostic test.

7.3.3 Normality and Jarque-Bera Test

The standardized residuals plot below for the estimated $MA(3)$ model shows consistency with our expectations in 7.1.3. Hence normal distribution appears to be appropriate.
JB Test for the estimated $MA(3)$ model gives $\hat{t}_3 = 3.209$ and $\hat{t}_4 = 1.720$, so
$JB_{stat} = \hat{t}_3^2 + \hat{t}_4^2 = 13.258 > 6$. Hence we reject the null hypothesis of normality, hence fail this diagnostic test.

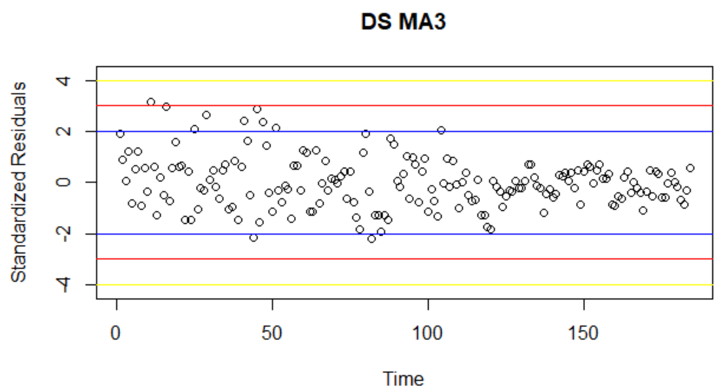


Figure 6: Standardized Residuals Plot for the estimated $AR(2)$ model

7.3.4 ARCH(6)

Performing the same steps in 7.1.4, we get

$$LM_{stat} = T \times R^2 = 183 \times 0.1284 = 23.627 \sim \chi_6^2$$

and $p - value = P(\chi_6^2 > 23.627) = 0.000612 < 0.05$, so we reject H_0 and conclude there exists non-linear correlation in a_t, hence failing the diagnostic test.

7.4 Conclusion

By conducting the diagnostics tests on the for the 3 estimated models, we found all 3 passing the Box-Pierce Test, Overfitting test and Standard Residuals test, but failing the Jarque-Bera Test and $ARCH(6)$ test. Hence we conclude that all three might be appropriate, but further analysis is to be conducted.

8 Financial Times Series

In this section, we use the monthly financial time series P_t for the S&P, and the goal is to determine if this series follows a random walk as

$$log(P_t) = \delta + log(P_{t-1}) + a_t \ with \ a_t \sim i.i.N[0, \sigma^2]$$

First, we test if $\phi = 1$ by running regression on

$$log(P_t) = \delta + \phi log(P_{t-1}) + a_t$$

Results are summarized below:

$$\underset{(t)}{log(P_t)} = \underset{(0.784)}{0.00546} + \underset{(610.622)}{1.000014} log(P_t) + a_t \tag{23}$$

$$n = 648 \quad F - ratio = 3.729 \times 10^5 \quad RSS = 0.04363^2 \quad R^2 = 0.9983$$

Under null hypothesis H_0: $\phi = 1$, we have

$$p - value = 2P(Z > |\frac{\hat{\phi} - 1}{se(\hat{\phi})}|) = P(Z > \frac{1.000014 - 1}{0.001638}) = 0.99318 > 0.05$$

Therefore we conclude that $\phi = 1$.

8.1 Autocorrelation for a_t

Second, we determine whether a_t are serial correlated with the estimated autocorrelation function $\rho_a(k)$. The table below summarizes $\rho_a(k)$ from $k = 1, \dots, 9$:

k	1	2	3	4	5	6	7	8	9
$\rho_a(k)$	-0.01155	-0.01353	0.00423	0.02082	0.08246	-0.05599	-0.00967	-0.05858	0.01524

Table 8: $\hat{\rho}_a(k)$ for $k = 1, \dots, 10$

Since there are 648 observations, we take $\hat{\rho}_a(k)$ to be significantly different than 0 if they exceed in absolute value $2/\sqrt{648} = 0.0785$. This means most $\rho_a(k)$'s are insignificant, with the exception of $k = 5$, which can be ignored being slightly significant. Hence a_t appears to have no correlation.
We verify this using the Box-Pierce test. we have $M \approx \sqrt{648} \approx 25$, giving $Q = 21.702 \overset{a}{\sim} \chi^2_{25}$ and $p - value = P(\chi^2_{25} > 21.702) = 0.6529 > 0.05$.
Hence the model passes this diagnostic test, and we accept the null hypothesis that there is no serial correlation. The time series appears to follow the random walk model.

8.2 Normality for a_t

Third, we plot the standardized residuals and perform the Jarque-Bera test to test normality in a_t. Notice that the plot below indicates that there are 4 data points outside of the 3 σ (red line) range, which is more than our expectations of $648 \times 0.003 \approx 1$ or 2 observations. However, as there appears be consistent variance in the plot, so normality might still hold.
Conducting the Jarque-Bera Test, we get $\hat{t}_3 = -4.478$ and $\hat{t}_4 = 11.735$, so

$$JB_{stat} = \hat{t}_3^2 + \hat{t}_4^2 = 555.59 > 6$$

and we reject the null hypothesis of normality.

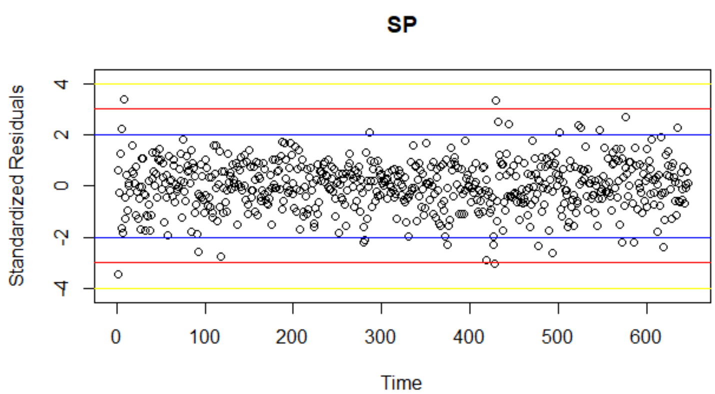


Figure 7: Standardized Residuals of a_t

8.3 Autocorrelation for a_t^2 and GARCH(1,1)

We now determine whether a_t^2 are serially correlated. the estimated autocorrelation function $\rho_{a^2}(k)$ are summarized the following table for $k = 1, \dots, 9$:

k	1	2	3	4	5	6	7	8	9
$\rho_{a^2}(k)$	0.07451	0.02345	0.01443	0.01547	-0.00913	0.04605	-0.04554	0.11019	0.07291

Table 9: $\hat{\rho}(k)$ for $k = 1, \ldots, 10$

Because the dataset is monthly, we consider it to be high frequency data, thus modeling with the Generalized Autogregressive Conditional Heteroskdasticity or $GARCH$ Model would be appropriate. $GARCH(q, p)$ is defined as follows:

Definition 5.

$$\sigma_t^2 = \sigma^2 + \sum_{i=1}^{q} \alpha_i a_{t-i}^2 + \sum_{j=1}^{p} \beta_j a_{t-j}^2$$

For stationarity, $\sigma^2 > 0$, $\alpha > 0$, $\beta > 0$ and $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$.

Using the garchFit in the R package fGarch, we fit $GARCH(1,1)$ as seen below:

$$\sigma_t^2 = 9.487 \times 10^{-5} + 0.06072 a_{t-1}^2 + 0.8884 a_{t-1}^2$$

Note that the estimates satisfy the stationary conditions $\hat{\sigma}^2 > 0$, $\hat{\alpha}_1 > 0$, $\hat{\beta}_1 > 0$ and $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j = \hat{\alpha}_1 + \hat{\beta}_1 < 1$, and the p-values indicate all coefficients are significant, therefore we conclude stationarity of this model.

9 Conclusion

This report demonstrates the statistical methodologies of time series analysis in the economic sector using the data from the government(Canada GDP) and financial markets (S&P Index). The report extracted the cycle component Y_t from the log-transformed time series data using the Trend Stationary Approach and Difference Stationary Approach. It also shows the process of fitting Y_t with an $AR(p)$ process and forecasting the growth rates and its 95% Confidence Intervals. An ADF test was performed, and suggested that for the GDP data, the DS approach is more appropriate than the TS approach. We then performed Box Jenkins Identification to determine the appropriate $ARMA(p, q)$ model that describes the GDP time series, for which we identified $AR(2)$ to be appropriate under the TS approach, and $AR(1)$ and $MA(3)$ under the DS approach. To test if these are good models, we applied 4 diagnostic tests: Box-Pierce Test, Overfitting, Normality (standardized residuals and Jarque-Bera), and ARCH(6). As all 3 models passed the first 2 tests and rejected the other 2, we could not determine which model is the best fitting, and concluded that further analysis is required. We later applied regression and tested for autocorrelation to determine if our monthly financial time series for S&P is a random walk,for which we concluded that there's no autocorrelation. The test of normality suggested that the time series errors for the S&P data may not follow a normal distribution. We estimated the autocorrelation of a_t's and fitted the time series data with the $GARCH(1,1)$, for which we concluded stationarity.