Assignment 12, MECO 6315

Meisam Hejazinia

04/20/2013

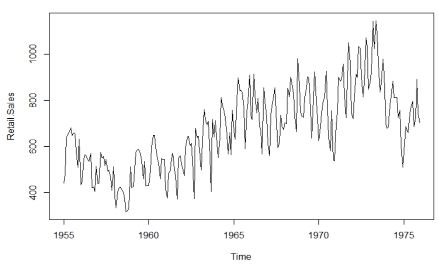

Figure 1: Monthly Retail Sales

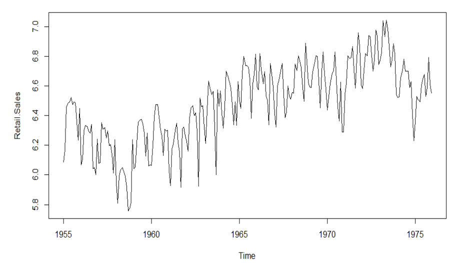

Figure 2: Log Transformed Monthly Retail Sales

1 General Analysis

From figure 1 we can infer that the model should be multiplicative, since the fluctuation in data is not constant over time. As a result, we used log transformation to transform the time series data. As figure 2 shows the transformed curve, the size of the seasonal fluctuation and the random fluctuation do not depend on the level of time series anymore. The transfomred data could be used to fit additive model. Next we tried to decompose the time series to trend and seasonal component. Figure 3 shows the result of smoothing procedure for order of 6. Still there are many random fluctuation in the data, so we smoothed by order 12, which is illustrated in figure 4. Next we used 'R' to decompose the trend and seasonal component, shown in figure 5.

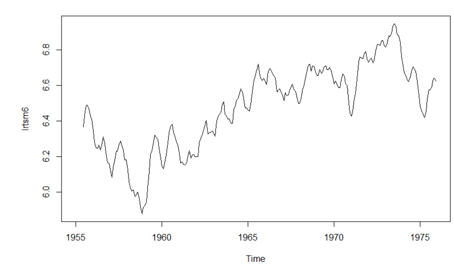

Figure 3: Result of Smoothing with Order 6

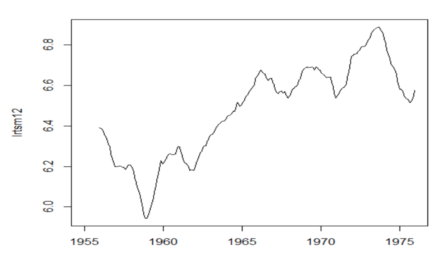

Figure 4: Result of Smoothing with Order 12

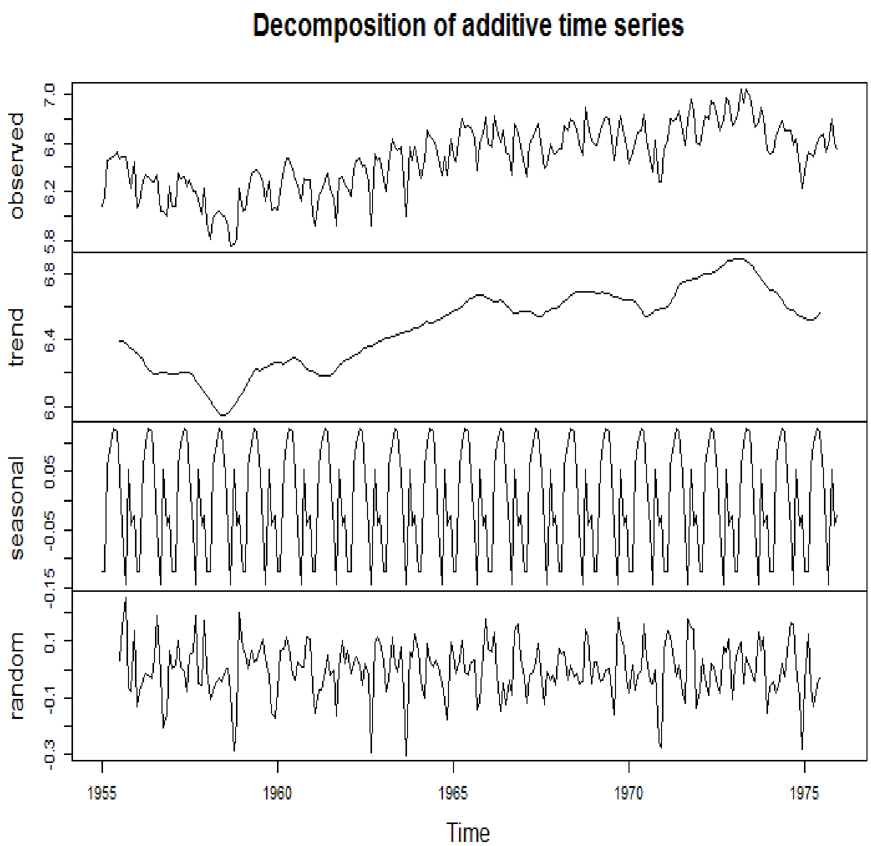

Figure 5: Result of Decomposition into Seasonal and Trend Components

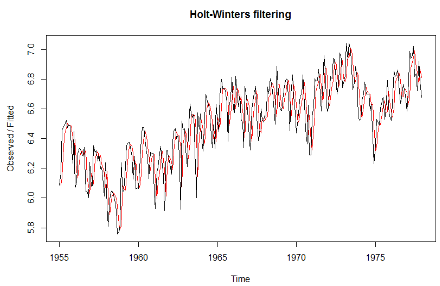

Figure 6: Holts Winter Prediction Versus Actual. Red curve illustrates predicted time series.

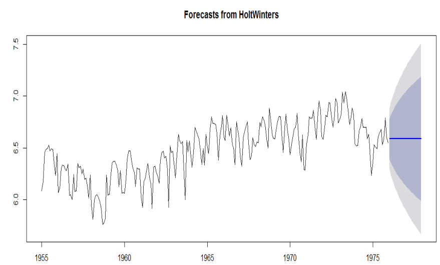

Figure 7: Holts Winter Forecast

2 Holt Winter

Holt Winter is used to forcast using exponential smoothing. Result of running Holt Winter method tells us that $\alpha = 0.6259951$. This is not close to zero, so recent observations have more effect on determining the current value. Coefficient is calculated 6.588841. Sum of square is alculate $SSE = 5.522659$ for Holt Winter multiplicative procedure.

By default, HoltWinters() just makes forecasts for the same time iginalperiod covered by our or time series. As a result we had no choice other than run the model on all years, rather than holding two last years to compare. Figure 6 shows how forecast, in red, is compared with time series. To be able to forecast another package called forecast is used. We excluded last two years, and created the model, and forecast procedure created forecast values that are shown in table 1. Figure 7 shows these forecasts. As I evaluated whether the real value is inside the confidence interval, I could not reject that they are inside of it.

Ljung-Box test is used for autocorrelation between the residuals. The result was showing that we have autocorrelation, since we had $p - value < 2.2e - 16$. This tells us that we need to use another method to improve the forecast. Also I plotted corologram, on lag 2 and 3, and 7 and 8 the confidence level is touched. Holt winter residuals is illustrated in figure 8. In sample forecast error seems to have constant variance over time. Figure 9 shows normal probability plot of residuals of the model, illustrating quite fit.

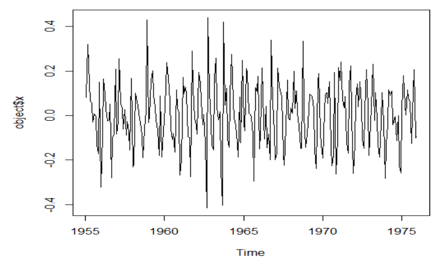

Figure 8: Holts Winter Residual Plot

Although exponential smoothing methods such as Holt Winter are useful for making forecast, they make no assumption about the correlation between time series elements. To have prediction using exponential smoothing method, it is required that forecast error be uncorrelated and normally distributed with mean zero, and constant variance. Autoregressive Integrated Moving Average (ARIMA) take into concern non-zero autocorrelation in irregular component of time series. In the next section we will run ARIMA procedure on the data.

Finally result of accuracy analysis of Holt winter method is shown in the following:

3 ARIMA

In order to run ARIMA we need to have stationary time series. To achieve stationary time series, it was

	Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jan	1976	6.588841	6.398409	6.779273	6.297601	6.880081
Feb	1976	6.588841	6.364174	6.813508	6.245243	6.932439
Mar	1976	6.588841	6.334507	6.843175	6.19987	6.977812
Apr	1976	6.588841	6.307955	6.869727	6.159263	7.018419
May	1976	6.588841	6.283705	6.893977	6.122176	7.055506
Jun	1976	6.588841	6.261246	6.916436	6.087827	7.089855
Jul	1976	6.588841	6.24023	6.937452	6.055687	7.121995
Aug	1976	6.588841	6.220411	6.957271	6.025377	7.152305
Sep	1976	6.588841	6.201606	6.976076	5.996616	7.181066
Oct	1976	6.588841	6.183672	6.99401	5.969188	7.208494
Nov	1976	6.588841	6.166499	7.011183	5.942924	7.234758
Dec	1976	6.588841	6.149997	7.027685	5.917687	7.259995
Jan	1977	6.588841	6.134094	7.043588	5.893366	7.284316
Feb	1977	6.588841	6.118729	7.058953	5.869866	7.307816
Mar	1977	6.588841	6.10385	7.073832	5.847111	7.330571
Apr	1977	6.588841	6.089414	7.088268	5.825033	7.352649
May	1977	6.588841	6.075384	7.102298	5.803576	7.374106
Jun	1977	6.588841	6.061727	7.115955	5.78269	7.394992
Jul	1977	6.588841	6.048415	7.129267	5.762331	7.415351
Aug	1977	6.588841	6.035424	7.142258	5.742462	7.43522
Sep	1977	6.588841	6.02273	7.154952	5.723049	7.454633
Oct	1977	6.588841	6.010315	7.167367	5.704061	7.47362
Nov	1977	6.588841	5.99816	7.179521	5.685473	7.492209
Dec	1977	6.588841	5.986251	7.191431	5.66726	7.510422

Table 1: Holt winter forecasts

ME	RMSE	MAE	MPE	MAPE	MASE
0.003195335	0.148332824	0.118331919	0.013616809	1.836795686	1.027507875

Table 2: Holt winter model percision

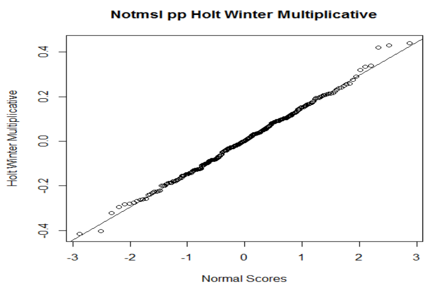

Figure 9: Holts Winter Residual Normal Probability Plot

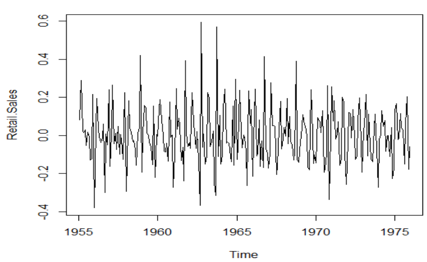

Figure 10: ARIMA First Difference Curve

need to take differences. As first difference was taken the variance becomes stationary, since according to figure 10, mean and variance become constant. Next it is needed to determine the degree of model (p,q). To determine the degree, auto correlogram is ploted in figure 11. Moreover, partial autocorrelogram is illustrated in figure 12. The two figures are indicative of either $(p, q) = (19, 0)$ or $(p, q) = (1, 11)$. We have also option with greater 'p' and 'q'. Model with less parameter is prefered to more parameter. As a result $(p, q) = (1, 11)$ is selected. The result of running the model is shown in the following.

Series: lrtsm
ARIMA(0,1,11)

Coefficients:
ma1 ma2 ma3 ma4 ma5 ma6 ma7 ma8
-0.3977 -0.2035 0.1365 -0.3439 0.1385 -0.0073 -0.1453 0.3126

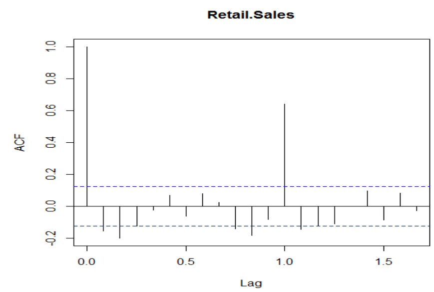

Figure 11: ARIMA Correlogram of Auto Correlation

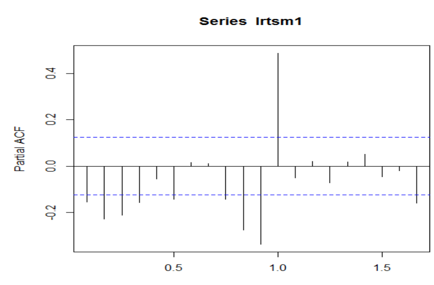

Figure 12: ARIMA Correlogram of Partial Auto Correlation

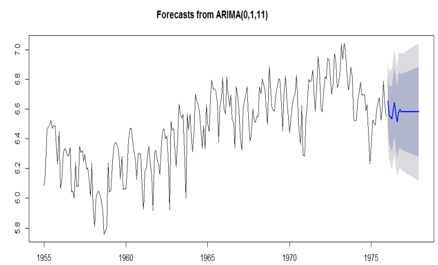

Figure 13: Manual ARIMA Model Prediction

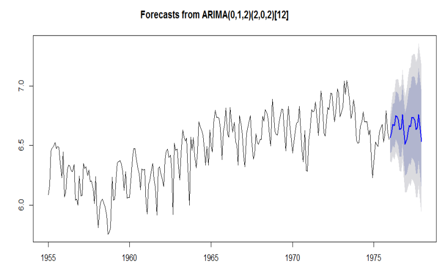

Figure 14: Automatic ARIMA Model Prediction

s.e. 0.0608 0.0633 0.0624 0.0612 0.0635 0.0599 0.0705 0.0642
ma9 ma10 ma11
−0.4709 0.0076 0.2473
s.e. 0.0613 0.0593 0.0551

$sigma^2$ estimated as 0.01685: log likelihood=154.58
$AIC = -285.15$ $AICc = -283.84$ $BIC = -242.85$

Predictive accuracy of the model is given in the following:

Table 2 shows confidence interval for my prediction based on ARIMA. Due to unknow reason the model had very poor predictability. Only 7 out of 24 real values were correctly inside the intervals. This made me to test automatic ARIMA, which finds the fit automatically, rather than using my procedure. Figure 13 shows prediction of ARIMA model ran manually.

The result of automatic ARIMA is shown in the following. Compared with my model its RMSE (Root Eean Square Error)and MASE (Mean Absolute Percentage Error/Deviation) are lower, indicating lower residuals. Figure 14 shows automatic ARIMA prediction.

$alpha = 0.6145$
$gamma = 0.2317$

Initial states:
$l = 6.276$
$s = 0.0944$ -0.103 -0.0056 -0.0464 0.008 0.0073 0.0643 0.1101 0.1177 0.0868 -0.1754 -0.1582

AIC	AICc	BIC
275.1742	276.9463	324.5862

Table 5: Automatic Exponential Fit Result

sigma: 0.1029

Table 3 shows the predicted confidence interval based on automatic ARIMA model. Interestingly, all the real parameters values are inside confidence interval of 95%.

After these I used automatic exponential model of 'R' to be able to compare result based on same measure of accuracy with ARIMA. The result is shown in the following:

ETS(A,N,A)

Call:
$ets(y = lrtsm)$

Smoothing parameters:
$alpha = 0.6145$
$gamma = 0.2317$

Initial states:
$l = 6.276$
$s = 0.0944$ -0.103 -0.0056 -0.0464 0.008 0.0073 0.0643 0.1101 0.1177 0.0868 -0.1754 -0.1582

sigma: 0.1029

Result of ETS (automatic exponential fit) of 'R' is

ME	RMSE	MAE	MPE	MAPE	MASE
0.005551083	0.129558696	0.106177122	0.053450937	1.648646594	0.921964505

Table 3: ARIMA model of $(p, q) = (1, 11)$

	Point Forecast	Low 80	Low 95	High 80	High 95
Jan-76	6.489844	6.401775	6.656209	6.822575	6.910644
Feb-76	6.366749	6.26394	6.560961	6.755172	6.857982
Mar-76	6.343575	6.234931	6.548808	6.75404	6.862684
Apr-76	6.309818	6.191387	6.533541	6.757263	6.875694
May-76	6.368988	6.249364	6.594965	6.820941	6.940566
Jun-76	6.413343	6.290241	6.645889	6.878435	7.001537
Jul-76	6.322524	6.196186	6.561182	6.79984	6.926178
Aug-76	6.274571	6.147272	6.515044	6.755517	6.882816
Sep-76	6.326108	6.191697	6.580017	6.833926	6.968337
Oct-76	6.340911	6.206489	6.594839	6.848767	6.983189
Nov-76	6.326868	6.192426	6.580835	6.834802	6.969244
Dec-76	6.322813	6.186225	6.580835	6.838857	6.975446
Jan-77	6.318821	6.180119	6.580835	6.842849	6.981551
Feb-77	6.314889	6.174106	6.580835	6.846782	6.987565
Mar-77	6.311014	6.168179	6.580835	6.850656	6.993491
Apr-77	6.307194	6.162337	6.580835	6.854477	6.999333
May-77	6.303427	6.156575	6.580835	6.858244	7.005095
Jun-77	6.29971	6.150891	6.580835	6.861961	7.01078
Jul-77	6.296041	6.14528	6.580835	6.865629	7.01639
Aug-77	6.292419	6.139741	6.580835	6.869251	7.021929
Sep-77	6.288843	6.134271	6.580835	6.872828	7.0274
Oct-77	6.285309	6.128867	6.580835	6.876362	7.032804
Nov-77	6.281817	6.123526	6.580835	6.879853	7.038144
Dec-77	6.278366	6.118248	6.580835	6.883305	7.043423

Table 4: ARIMA confidence interval for my Model

ME	RMSE	MAE	MPE	MAPE	MASE
0.001867907	0.100425412	0.077591936	0.013782450	1.206095820	0.673751645

Table 6: Automatic ARIMA Fit Result

	Point Forecast	Low 80	Low 95	High 80	High 95
Jan-76	6.565937	6.436982	6.694893	6.368717	6.763158
Feb-76	6.607679	6.446344	6.769014	6.360938	6.85442
Mar-76	6.674616	6.500495	6.848738	6.40832	6.940913
Apr-76	6.66898	6.482948	6.855012	6.384469	6.953491
May-76	6.749278	6.552054	6.946502	6.44765	7.050906
Jun-76	6.739237	6.531423	6.947051	6.421413	7.057061
Jul-76	6.717767	6.499877	6.935657	6.384533	7.051001
Aug-76	6.636195	6.408675	6.863716	6.288233	6.984158
Sep-76	6.644269	6.40751	6.881028	6.282177	7.006361
Oct-76	6.761017	6.515367	7.006668	6.385327	7.136708
Nov-76	6.624975	6.370743	6.879206	6.236161	7.013788
Dec-76	6.509787	6.247255	6.772319	6.108279	6.911296
Jan-77	6.547228	6.263095	6.831361	6.112684	6.981771
Feb-77	6.61214	6.311597	6.912683	6.152499	7.07178
Mar-77	6.666267	6.353268	6.979265	6.187576	7.144957
Apr-77	6.659518	6.33454	6.984496	6.162507	7.156528
May-77	6.7356	6.39907	7.07213	6.220921	7.250279
Jun-77	6.730885	6.383185	7.078585	6.199124	7.262646
Jul-77	6.714704	6.356183	7.073225	6.166394	7.263015
Aug-77	6.633988	6.264963	7.003013	6.069613	7.198363
Sep-77	6.64678	6.267542	7.026018	6.066785	7.226775
Oct-77	6.760761	6.371577	7.149944	6.165555	7.355966
Nov-77	6.634583	6.235701	7.033464	6.024546	7.244619
Dec-77	6.533912	6.125563	6.942261	5.909397	7.158427

Table 7: Automatic ARIMA Confidence Interval

ME	RMSE	MAE	MPE	MAPE	MASE
0.00278239	0.10286996	0.07862786	0.02780022	1.22239676	0.68274683

Table 9: Automatic Exponential Fit Result

AIC	AICc	BIC
275.1742	276.9463	324.5862

Table 8: Automatic Exponential Fit Result

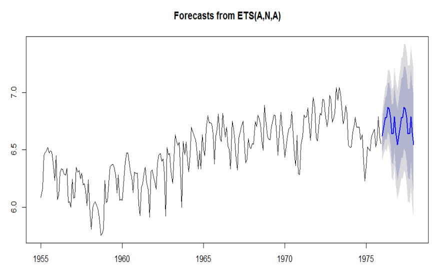

Figure 15: Automatic Exponential Model (ETS) Prediction

shown in figure 15.

The result shows that ARIMA fits the data much better than exponential model, according to its lower MAPE, and MASE. After all the models I fit the data worked worse than the models that 'R' automatically fited the time series data.