National Univaesity of Singapore

IE5202 - Forecasting Methods

Project 2

Currency Forecasting

Author:
Gabriel Azevedo Ferreira

Supervisor:
Chen Nan

November 19, 2017

1 Introduction

The goal of this project is to predict foreign exchange rates. We are given a dataset with exchange rates of eight countries relative to the United States Dollar, and our goal is to predict the Japanese yen currency.

This is a time series analysis problem and, in this project, we tackled it by applying the main tools taught during the IE5202 classes.

In order to predict the exchange rate for a given date, we are only allowed to use all past data relative to that date. This simulates a real forecasting situation, where the future data has not yet occured.

The evaluation of the model is done through the one step ahead error. For each method, we minimize the sum of squared errors for the one step ahead prediction (SSE). Because the mean squared erros (MSE) differs from the SSE by only a multiplied constant (the number of samples), we compared the different methods according to their MSE.

On the first part of this project we only used the time series to perform the prediction, while, on the second part, we used all the exchange rates (including the "JPY/USD" itself) to perform this task.

On the programming part, we chose to work in python, making use of the case studies used during the classes. A part of the code is done in R, where the package *forecasting* played an important role.

2 Visualization

The first step on the project was to look at the data and try to draw conclusions about it.

In order for the visualization to be more accurate, we scaled JPY/USD (only during plottings).

We also plot the first order differences (Figure 3).

We conclude from Figure 3 and 1 that, although data does not change smoothly, the noise seems to be relatively small compared to the JPY/USD overall value.

In working with the first order differences, we manage to have a data more closer to a stationary model (having less variation on properties like the mean and variance) we take the differences between the two successive currency exchanges. The variance and mean are shown in Figure 4, before (left hand side) and after (right hand side) the difference transformation.

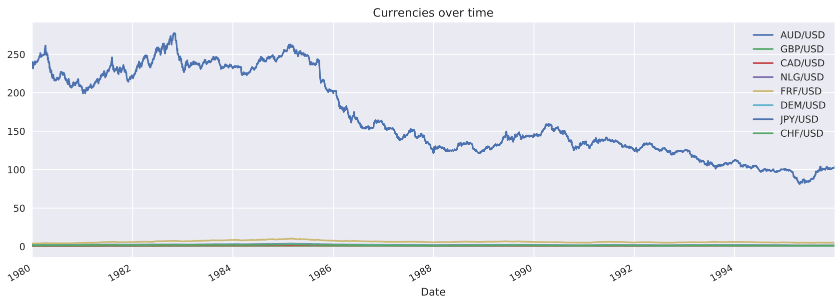

Figure 1: Currencies over time

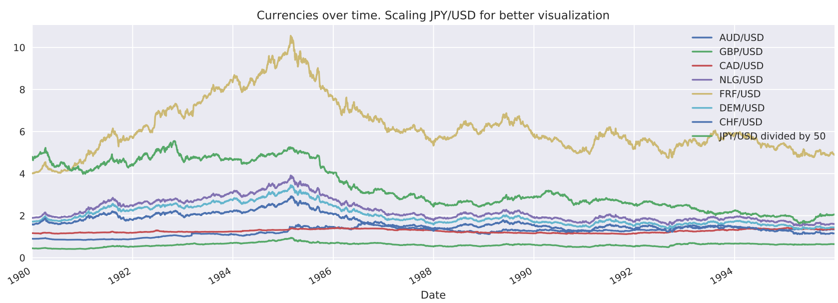

Figure 2: Currencies over time - JPY scaled for visualization

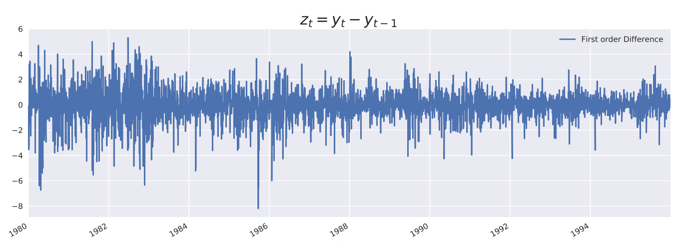

Figure 3: $z_t = y_t - y_{t-1}$. z_t Versus time

The plots in Figure 4 show that after the transformation ($z_t = y_t - y_{t-1}$) the mean is rather constant (slightly oscilating around zero) and eventhough the variance changes over time, the difference in absolute values is rather low. However, the variance on the data without transformation is less stable,

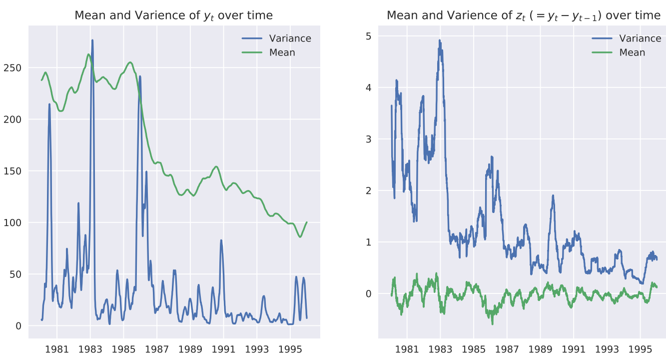

Figure 4: Variance and mean over time. The Left side represents $z_t = y_t - y_{t-1}$, that is data after the transformation and the right hand side represents data before the transformation. The variance and mean were computed by considering the last 100 samples (from $t - 100$ to $t - 1$)

varying in a much larger range of values (from 200 to 400). The mean of the transformed data also changes over time in a significant way. We conclude, therefore, that the assumption that the data holds the same characteristics is not reasonable to the non transformed data but might be for $z_t = y_t - y_{t-1}$. This assumption specially important for ARMA models.

We also looked at the autocorrelation on y_t and z_t. This is shown on the Figures 5, 6 and 7.

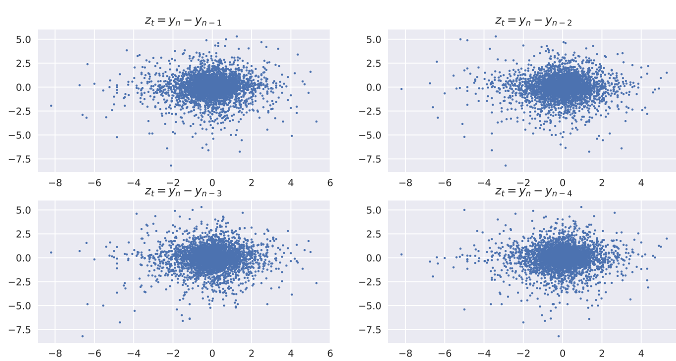

Figure 5: Four plots of z_t Vs z_{t-1}, where $z_t = y_t - y_{t-i}$

Figure 5 simulates the distribution of $P(z_t, z_{t-1})$ where $z_t = y_t - y_{t-i}$. We can see that the data seems to have a normal shape with mean close to $(0, 0)$ (meaning that it seems not to have a skew and has a bell-shape, although this distribution is not necessarily normal).

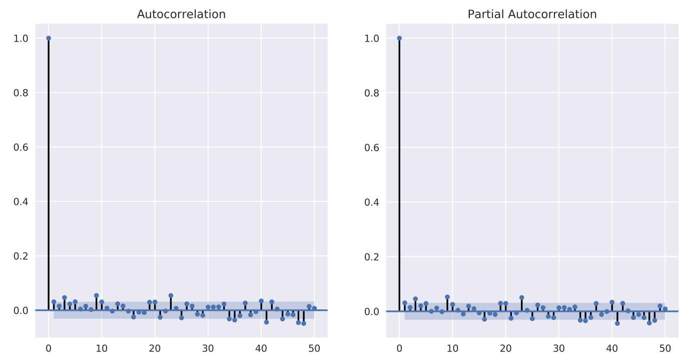

Figure 6: AACF and PACF for $z_t = y_t - y_{t-1}$

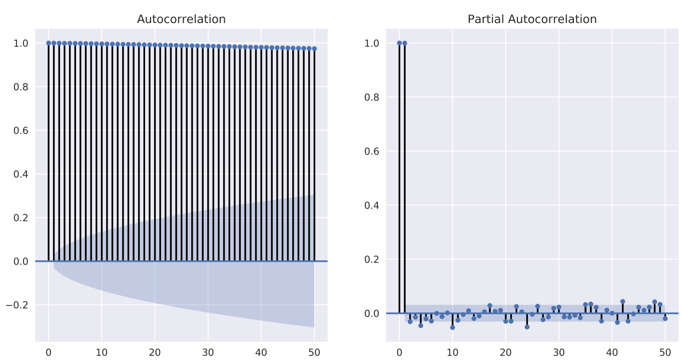

Figure 7: AACF and PACF for y_t

Figure 6 and Figure 7 show that data is strongly correlated to the past values and, given the influence of the last information, the past data might not be highly important.

3 PART 1

On the first part of the project we try to obtain the best predictor (acoording to the SSE loss and the one step ahead error) using only the "JPY/USD" time series (and no other features).

3.1 Moving average

The moving average is one of the simplest methods, but allowed us to reach very low errors. Because y_t is higly dependent on y_{t-1}, we obtained the best results using low number of past data in the average (that is, a low lag). In particular, The lowest MSE was obtained in the simplest model possible:

$$\hat{y}_t = y_{t-1}$$

The value of MSE obtained was 1.33712

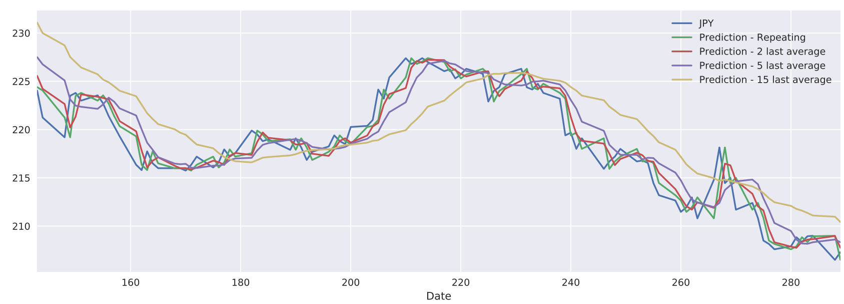

Figure 8: Mean and Variance over Time

3.2 Regression on time

Due to the considerable autocorrelation, we expect that a simple regression on time would not be sufficient for a good prediction. Linear Regression models assume that the samples are independent. Therefore, Ordinary Least square fails to take advantage of this information when the data is highly auto-correleted. On the other hand, the information that the data has high auto-correlation can be very useful if well explored.

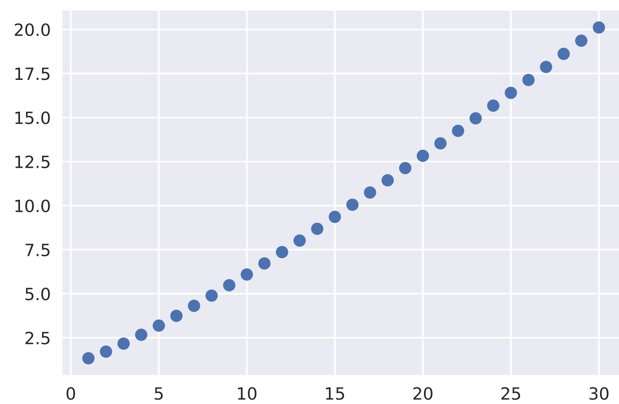

Figure 9: Mean and Variance over Time

However, in order to illustrate the non suitability of simple regression in ths situation, we created two linear regression models (Figure 10 and Figure 11). As expected, the results were far worse that the ones obtained with the moving average.

Figure 10: Mean and Variance over Time

Figure 11: Mean and Variance over Time

3.3 ETS

3.3.1 Exponential Smoothing Without Trend

The best parameter α (i.e. the one minimizing the SSE) was found to $\alpha = 1$, that is, the model equivalent to repeting the last day prediction.(Therefore, MSE $= 1.33712$) This is shown in Figure 12.

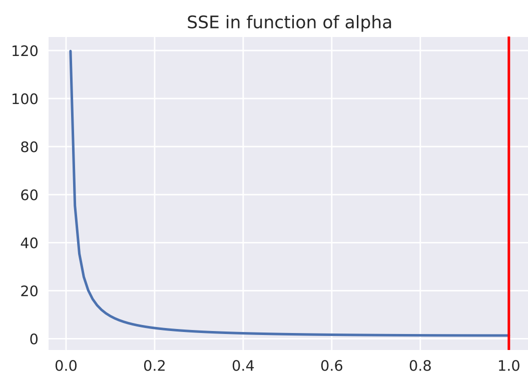

Figure 12: Mean and Variance over Time

3.3.2 Holt's Trend Corrected Smoothing

We now considered a model with a linear trend, that is, $\hat{y}_t = Bt + L$ (where we update B and L according to the exponential smoothing equations). The best model found by the *"forecasting"* package on R, by minimizing the MSE, was still practically the day-before prediction ($\hat{y}_t = y_{t-1}$).

The parameters for this model are $\alpha = 0.9999$ and $\beta = 0.0001$, and the MSE is 1.336714. We recall that the updating equations are:

$$L_n = \alpha y_n + (1\alpha)(L_{n1} + B_{n1})$$
$$Bn = \beta(L_n L_{n1}) + (1\beta)B_{n1}$$

The difference from the day before prediction ($\alpha = 1$ and $\beta = 0$) can be explained by the numerical optimization methods used. Therefore, the model found by the package can be considered equivalent to $\hat{y}_t = y_{t-1}$.

We consider a multiplicative trend, and a damped model, and the results were slightly better. The model used contains the following structure.

$$\hat{y}|_{t+h} tamp; = \ell_t b_t^{(\phi+\phi^2+\cdots+\phi^h)}$$
$$\ell_t amp; = \alpha y_t + (1-\alpha)\ell_{t-1} b_{t-1}^{\phi}$$
$$b_t amp; = \beta^* \frac{\ell_t}{\ell_{t-1}} + (1-\beta^*) b_{t-1}^{\phi}.$$

The coefficients found where the followig:

$$\alpha = 0.9999\beta = 0.0284\phi = 0.9347$$

And the MSE was: 1.331441

3.3.3 Holt-Winters Filtering

In this part we took the seasoonality factor into account. We tested a variaty of seasonalities periods, from 2 to 20. However, this feature didn't play an important role on the prediction, and the model presenting the least MSE was also roughly equivalent to the day before prediction. For period equals 11 days, we found

$$\alpha = 0.9999, \beta = 0.0076 \text{ and } \gamma = 1e - 04$$

The MSE was 1.337366.

Again, the parameters were virtually zero except for α, which is practically one. This makes the model equivalent to day-before prediction.

3.4 ARIMA

ARIMA is a generalization of the ARMA model. It has the follows the following structure:

$$\underset{\mathrm{AR}(p)}{\underset{\uparrow}{(1-\phi_1 B-\cdots-\phi_p B^p)}} \quad \underset{d\text{ differences}}{\underset{\uparrow}{(1-B)^d y_t}} \quad = \quad c+\underset{\mathrm{MA}(q)}{\underset{\uparrow}{(1+\theta_1 B+\cdots+\theta_q B^q)e_t}}$$

where B is the differentiating operator. The optimal values for the ϕ's and θ's are:

ϕ_1	0.134481483	θ_1	-0.10518325
ϕ_2	0.265028864	θ_2	-0.24936813
ϕ_3	0.038526979	θ_{20}	0.13254135
ϕ_{17}	-0.101683572	θ_{21}	-0.01394113
ϕ_{18}	0.013674558	θ_{22}	-0.03305159
ϕ_{19}	0.026949082	θ_{40}	0.11987226
ϕ_{20}	0.003917561	θ_{41}	0.01260855
ϕ_{40}	-0.073276886	θ_{42}	-0.02989232
ϕ_{41}	0.009854384	p	43
ϕ_{42}	0.019420490	d	1
ϕ_{43}	0.002823137	q	42

The MSE obtained was: 1.328867

4 Part 2

In Part 2 we tested many number of lags and chose the model with least MSE on the fraction of the test data for which the target is available.

4.1 Day before prediction

$$\hat{y}_t = y_{t-1}$$

This model gave MSE = 1.121705.

4.2 Multiple regression

We analyzed the correlations between variations on JPY/USD and variations in other currencies exchanges. Table 13 show the results. The correlations were measured between the JPY/USD and the othe currencies shifted of one day backwards. For example, the element at row one and column two represents the correlation between JPY and AUD shifted one day backwards. The element at row two and column two represents the correlation between the first order difference of JPY and the first order difference of AUD. Since the coeficients of both first and second order dfference are close to zero, we can not conclude that changes on JPY are related to one day past changes of other currencies.

We also plot the correlations with different time-differences (when the correlation data is shifted backwards of step values greater than one) and the coefficients were close to zero in all cases tested.

	JPY	AUD	GBP	CAD	NLG	FRF	DEM	CHF
y_t	0.9997	-0.6091	0.0906	-0.1319	0.837	0.5435	0.8556	0.8823
$z_t = y_t - y_{t-1}$	0.0310	0.00212	-0.0051	0.00622	0.0149	0.0174	0.00898	0.0213
$w_t = z_t - z_{t-1}$	0.0069	0.00117	0.0035	0.00111	0.0044	0.0041	0.00460	0.0045

Figure 13: Each element of the table represents the correlation coefficient of JPY.USD against one of the eight currencies shifted of one day period and transformed. The first row represents no transformation, while the second row represents the first order difference and the third shows the second order difference.

We performed linear regression having as features the information about all currencies in the past k days (dat is, with lag equals k). We also included the days of the week, as dummy variables, but this information did not appear to improve the predictions.

For simplicity, we trained our model on the training set and validated it on the testing set.

We obtained reasonably good results for k = 2.

The coefficients are shown in the attachements section. as expected, the coefficient of JPY shifted of one time unit (JPY) is close to one.

The MSE for the validation set was: 1.3275

5 Conclusion

In this work we could use many important algorithms in the context of time series data analysis and compare it's results. In this especific problem, most of the algorithms pointed to the fact that the day before prediction might be a very good prediction in currency exchange prediction. The fact that it perfomed better than most of the other methods might be explained by the great correlation between y_t and y_{t-1}, and specially the low partial correlation with y_{t-i} , $i > 1$ (Figure 7).

Seasonality did not appear to play a significant role in this case.

6 Attachements

OLS Regression Results

Dep. Variable:	Target	R-squared:	1.000
Model:	OLS	Adj. R-squared:	1.000
Method:	Least Squares	F-statistic:	5.708e+05
Date:	Sun, 19 Nov 2017	Prob (F-statistic):	0.00
Time:	11:22:51	Log-Likelihood:	-6260.4
No. Observations:	4017	AIC:	1.255e+04
Df Residuals:	4000	BIC:	1.266e+04
Df Model:	16		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	1.3032	0.390	3.342	0.001	0.539	2.068
JPY1	1.0368	0.021	48.498	0.000	0.995	1.079
AUD1	-0.1557	2.333	-0.067	0.947	-4.730	4.419
GBP1	-12.2132	6.603	-1.850	0.064	-25.160	0.733
CAD1	1.4643	5.494	0.267	0.790	-9.306	12.235
NLG1	19.4638	7.478	2.603	0.009	4.803	34.125
FRF1	1.9163	1.205	1.590	0.112	-0.447	4.279
DEM1	-30.5945	8.932	-3.425	0.001	-48.107	-13.082
CHF1	6.0530	3.568	1.696	0.090	-0.943	13.049
JPY2	-0.0398	0.021	-1.864	0.062	-0.082	0.002
AUD2	-0.3845	2.334	-0.165	0.869	-4.960	4.191
GBP2	12.8218	6.608	1.940	0.052	-0.133	25.776
CAD2	-1.9707	5.490	-0.359	0.720	-12.735	8.793
NLG2	-17.4642	7.471	-2.338	0.019	-32.111	-2.817
FRF2	-2.0386	1.204	-1.693	0.091	-4.399	0.322
DEM2	29.1209	8.920	3.265	0.001	11.633	46.609
CHF2	-6.4390	3.571	-1.803	0.071	-13.441	0.563

Omnibus:	559.862	Durbin-Watson:	2.001
Prob(Omnibus):	0.000	Jarque-Bera (JB):	2998.918
Skew:	-0.552	Prob(JB):	0.00
Kurtosis:	7.086	Cond. No.	2.19e+05