LECTURE 2
MOVING AVERAGES AND EXPONENTIAL SMOOTHING

OVERVIEW

This lecture introduces time-series smoothing forecasting methods. Various models are discussed, including methods applicable to nonstationary and seasonal time-series data. These models are viewed as classical time-series model; all of them are univariate.

LEARNING OBJECTIVES

- Moving averages
- Forecasting using exponential smoothing
- Accounting for data trend using Holt's smoothing
- Accounting for data seasonality using Winter's smoothing
- Adaptive-response-rate single exponential smoothing

1. Forecasting with Moving Averages

The naive method discussed in Lecture 1 uses the most recent observations to forecast future values. That is, $\hat{Y}_{t+1} = Y_t$. Since the outcomes of Y_t are subject to variations, using the mean value is considered an alternative method of forecasting. In order to keep forecasts updated, a simple moving-average method has been widely used.

1.1. The Model

Moving averages are developed based on an average of weighted observations, which tends to smooth out short-term irregularity in the data series. They are useful if the data series remains fairly steady over time.

Notations

$M_t \equiv \hat{Y}_{t+1}$ - Moving average at time t, which is the forecast value at time $t+1$,

Y_t - Observation at time t,

$e_t = Y_t - \hat{Y}_t$ - Forecast error.

A moving average is obtained by calculating the mean for a specified set of values and then using it to forecast the next period. That is,

$$M_t = (Y_t + Y_{t-1} + \cdots + Y_{t-n+1})/n \qquad (1.1.1)$$

$$M_{t-1} = (Y_{t-1} + Y_{t-2} + \cdots + Y_{t-n})/n \qquad (1.1.2)$$

Subtracting Equation (1.1.2) from Equation (1.1.1), we obtain:

$$M_t = M_{t-1} + (Y_t - Y_{t-n})/n \qquad\qquad\qquad (1.1.3)$$

Equation (1.1.3) allows us to update the data, making the forecasting process much easier. This equation states that the moving average can be updated by using a previous moving average plus the average changes in actual value from time *t* to *t-n*. Using either Equation (1.1.1) or (1.1.3) should yield the same result.

1.2. A Numerical Example

To illustrate how a moving average is used, consider Table 3-1, which contains the exchange rate between the Japanese yen and the US dollar from 1983Q1 through 1998Q4. To calculate the three-quarter moving average requires first that we sum the first three observations (239.3, 239.8, and 236.1). This three-quarter total is then divided by 3 to obtained 238.40, as shown in the third cell of column 4 in Table 1. This smoothed number, 238.40, becomes the forecast for 1983Q4, displayed in the fourth cell of column 5 of 3-Q MAF.

By the same token, we can obtain the forecast for 1984Q1 by moving one quarter ahead and dropping the most distant quarter. That is,

$$Y_{t+1} = (Y_t + Y_{t-1} + Y_{t-2})/3 \quad \text{(assume } n = 3),$$

$$\frac{232 + 236.1 + 239.8}{3} = 235.97.$$

The last value of the moving average is 130.29, which is the forecast for 1999.Q1.

$$\frac{115.2 + 135.72 + 139.95}{3} = 130.29.$$

It is of interest to calculate the squared errors (SE) and the sum of squared errors (SSE). The squared errors of using moving average are presented in column 8, labeled by SE_MA. The resulted mean-squared error (MSE) is 244.21 (The last row of Table 1). This figure (244.21) appears to be larger than the MSE of 218.94 obtained by a naive model, the random-walk process.

　　Not surprisingly, if you are familiar with the research in international finance, this result is consistent with most empirical findings. It has been shown that not many models can beat the random-walk process since the current exchange rate contains all the historical information pertinent to predict exchange rate movements, as stated by the efficient market hypothesis.

Table 1. The Japanese Yen / US Dollar Rate: 1983Q1 - 1998Q4

Period	Actual	1-Q RW	3-Q MA	3-Q MAF	XS(α=0.8)	SE_RW	SE_MA	SE_XS
Mar-83	239.3	Missing	Missing	Missing	Missing			
Jun-83	239.8	239.3	Missing	Missing	239.30			
Sep-83	236.1	239.8	238.40	Missing	239.70			
Dec-83	232	236.1	235.97	238.40	236.82	16.81	40.96	23.23
Mar-84	224.75	232	230.95	235.97	232.96	52.56	125.81	67.47
Jun-84	237.45	224.75	231.40	230.95	226.39	161.29	42.25	122.26
Sep-84	245.4	237.45	235.87	231.40	235.24	63.20	196.00	103.25
Dec-84	251.58	245.4	244.81	235.87	243.37	38.19	246.91	67.44
Mar-85	250.7	251.58	249.23	244.81	249.94	0.77	34.69	0.58
Jun-85	248.95	250.7	250.41	249.23	250.55	3.06	0.08	2.55
Sep-85	216	248.95	238.55	250.41	249.27	1085.70	1184.05	1106.86
Dec-85	200.6	216	221.85	238.55	222.65	237.16	1440.20	486.37
Mar-86	179.65	200.6	198.75	221.85	205.01	438.90	1780.84	643.17
Jun-86	163.95	179.65	181.40	198.75	184.72	246.49	1211.04	431.48
Sep-86	153.63	163.95	165.74	181.40	168.10	106.50	771.17	209.51
Dec-86	160.1	153.63	159.23	165.74	156.52	41.86	31.85	12.78
Mar-87	145.65	160.1	153.13	159.23	159.38	208.80	184.33	188.65
Jun-87	146.75	145.65	150.83	153.13	148.40	1.21	40.66	2.71
Sep-87	146.35	146.75	146.25	150.83	147.08	0.16	20.10	0.53
Dec-87	122	146.35	138.37	146.25	146.50	592.92	588.06	600.05
Mar-88	124.5	122	130.95	138.37	126.90	6.25	192.28	5.76
Jun-88	132.2	124.5	126.23	130.95	124.98	59.29	1.56	52.13
Sep-88	134.3	132.2	130.33	126.23	130.76	4.41	65.07	12.56
Dec-88	125.9	134.3	130.80	130.33	133.59	70.56	19.65	59.15
Mar-89	132.55	125.9	130.92	130.80	127.44	44.22	3.06	26.13
Jun-89	143.95	132.55	134.13	130.92	131.53	129.96	169.87	154.31
Sep-89	139.35	143.95	138.62	134.13	141.47	21.16	27.21	4.48
Dec-89	143.4	139.35	142.23	138.62	139.77	16.40	22.88	13.15
Mar-90	157.65	143.4	146.80	142.23	142.67	203.06	237.67	224.26
Jun-90	152.85	157.65	151.30	146.80	154.65	23.04	36.60	3.26
Sep-90	137.95	152.85	149.48	151.30	153.21	222.01	178.22	232.90
Dec-90	135.4	137.95	142.07	149.48	141.00	6.50	198.34	31.38
Mar-91	140.55	135.4	137.97	142.07	136.52	26.52	2.30	16.24
Jun-91	138.15	140.55	138.03	137.97	139.74	5.76	0.03	2.54
Sep-91	132.95	138.15	137.22	138.03	138.47	27.04	25.84	30.46
Dec-91	125.25	132.95	132.12	137.22	134.05	59.29	143.20	77.51
Mar-92	133.05	125.25	130.42	132.12	127.01	60.84	0.87	36.47
Jun-92	125.55	133.05	127.95	130.42	131.84	56.25	23.68	39.59
Sep-92	119.25	125.55	125.95	127.95	126.81	39.69	75.69	57.13
Dec-92	124.65	119.25	123.15	125.95	120.76	29.16	1.69	15.12
Mar-93	115.35	124.65	119.75	123.15	123.87	86.49	60.84	72.63
Jun-93	106.51	115.35	115.50	119.75	117.05	78.15	175.30	111.19
Sep-93	105.1	106.51	108.99	115.50	108.62	1.99	108.23	12.38
Dec-93	111.89	105.1	107.83	108.99	105.80	46.10	8.43	37.04
Mar-94	102.8	111.89	106.60	107.83	110.67	82.63	25.33	61.98
Jun-94	98.95	102.8	104.55	106.60	104.37	14.82	58.47	29.43
Sep-94	98.59	98.95	100.11	104.55	100.03	0.13	35.48	2.09
Dec-94	99.83	98.59	99.12	100.11	98.88	1.54	0.08	0.90
Mar-95	88.38	99.83	95.60	99.12	99.64	131.10	115.42	126.78
Jun-95	84.77	88.38	90.99	95.60	90.63	13.03	117.29	34.36
Sep-95	98.18	84.77	90.44	90.99	85.94	179.83	51.65	149.76
Dec-95	102.91	98.18	95.29	90.44	95.73	22.37	155.42	51.52

(continued)

Table 3-1 (continued)

Period	Actual	1-Q RW	3-Q MA	3-Q MAF	XS(α=0.8)	SE_RW	SE_MA	SE_XS
Mar-96	106.49	102.91	102.53	95.29	101.47	12.82	125.51	25.16
Jun-96	109.88	106.49	106.43	102.53	105.49	11.49	54.07	19.30
Sep-96	111.45	109.88	109.27	106.43	109.00	2.46	25.23	6.00
Dec-96	115.98	111.45	112.44	109.27	110.96	20.52	44.98	25.20
Mar-97	123.97	115.98	117.13	112.44	114.98	63.84	133.02	80.89
Jun-97	114.3	123.97	118.08	117.13	122.17	93.51	8.03	61.96
Sep-97	121.44	114.3	119.90	118.08	115.87	50.98	11.27	30.98
Dec-97	129.92	121.44	121.89	119.90	120.33	71.91	100.33	92.03
Mar-98	133.39	129.92	128.25	121.89	128.00	12.04	132.33	29.04
Jun-98	139.95	133.39	134.42	128.25	132.31	43.03	136.89	58.33
Sep-98	135.72	139.95	136.35	134.42	138.42	17.89	1.69	7.30
Dec-98	115.2	135.72	130.29	136.35	136.26	421.07	447.46	443.54
Mar-99		115.2	Missing	130.29	119.41			
MSE						218.94	244.21	233.39

Notes:
1-Q RW: 1-quarter random walk process
3-Q MA : 3-quarter moving average
3-Q MAF: 3-quarter moving average forecast
XS(α=0.8): 1-quarter exponential smoothing with α =0.8
SE_RW: Squared errors by using random walk forecast
SE_MA: Squared errors by 3-quarter moving-average forecast
SE_XS: Squared errors by using exponential-smoothing forecast
MSE: Mean squared errors

1.3. Remarks on Moving-Average Method

The moving-average method provides an efficient mechanism for obtaining a value for forecasting stationary time series.   The technique is simply an arithmetic average as time passes, with some lag-length determined optimally by an underlying cycle present in the data. Thus, moving-averages and moving-average lines are frequently derived by technicians on Wall Street to generate market expectations, one of the most important input variables used by fund managers to allocate portfolios.

The difficulty in using moving averages is their inability to capture the peaks and troughs of the series. When the market (actual) data are moving down persistently, the moving average forecast tends to produce over-predicted valued; while when the market is moving up continually, the moving-average forecast will under-predict the market. Obviously, this method fails to deal with non-stationary data. Moreover, since all the data points in the moving-average process are given equal weight, this approach fails to reflect the importance of time ordering with respect to observations.   For this reason, a weighted moving-average method has been suggested. It merely imposes different weights on the observations being used for forecasting. The double moving-average method, taking the form of moving average on the first moving-averages, gives more weight on the middle point. Exponential smoothing methods are the techniques that place more weights on the recent observations.

2. Forecasting with Exponential Smoothing

2.1. The Model

Simple exponential smoothing takes the form of:

$$F_{t+1} = \alpha Y_t + (1-\alpha)F_t \tag{2.1.1}$$

$$\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha)\hat{Y}_t$$

Notations:

F_{t+1} = forecast value for period $t+1$ made at time t, which can be defined as $\hat{Y}_{t+1}$

Y_t = actual value in period t

(Wilson and Keating use X_t, others use Z_t, we use Y_t to maintain notation consistency)

F_t = forecast value for period t made by t-1.

α = smoothing constant (0< α <1)

By continuing to substitute previous forecasting values back to the starting point of the data:

$$\hat{Y}_{t+1} = \alpha Y_t + \alpha(1-\alpha)\hat{Y}_t + \alpha(1-\alpha)^2\hat{Y}_{t-1} + \ldots + \alpha(1-\alpha)^{t-1}Y_1 + (1-\alpha)^t Y_0 \tag{2.1.2}$$

Writing this equation in compact form:

$$\hat{Y}_{t+1} = \alpha \sum_{k=0}^{t-1}(1-\alpha)^k Y_{t-k} + (1-\alpha)^t Y_0 \tag{2.1.3}$$

It is clear that the weights, $\alpha, \alpha(1-\alpha), \alpha(1-\alpha)^2$, on Y_t, Y_{t-1}, Y_{t-2} ...as implied in Equation (2.1.3) are exponentially declining.

Two points deserve our attention before we proceed to make our forecasts. First, we need to decide the initial value, Y_0. A convenient way to accomplish this is to utilize the value of the initial data point or the average value of the first few observations of the data series. Second, we must determine the value of α. Usually, this selection can be achieved by minimizing the MSE or RMSE based on in-sample experiments.

2.2. Numerical Example

Simple exponential smoothing can be illustrated by using quarterly data of the yen/dollar exchange rate in Table 3-1. Assuming that $\alpha = 0.8$, calculations of the exponential smoothing of the exchange rate are as follows:

$$F_{t+1} = \alpha Y_t + (1-\alpha)F_t = 0.8(239.8) + (1-0.8)\,239.30 = 239.70 \text{ - Forecast for 1983.Q3,}$$

$$= 0.8(236.1) + (1-0.8)\,239.70 = 236.82 \qquad \text{Forecast for 1983.Q4.}$$

By calculating the smoothing values in the same manner, we obtain the figures presented in column 6 under "XS($\alpha = 0.8$)". Again, the squared errors are shown in column 9; the resulting MSE is still higher than the random-walk process although it is slightly better than the moving-average smoothing calculation. From our exercise, the naive model in the form of a random walk is not so naive; it is quite appropriate to describe an asset-price behavior.

As a guide for selecting the smoothing constant, it is suggested that α values close to 0 are selected if the series has small variations and values close to 1 are selected if the forecast values appear to depend on recent changes in actual values. Usually the MSE or RMSR can be used as the criterion for selecting an appropriate smoothing constant. For instance, by assigning α values from 0.1 to 0.99, we select the value that produces the smallest MSE.

2.3. Remarks on Simple Exponential Smoothing

- The model $\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha)\hat{Y}_t$ can be rewritten as:
  $\hat{Y}_{t+1} - \hat{Y}_t = \alpha(Y_t - \hat{Y}_t)$, change in forecasting value is proportionate to the forecast error. That is, $\hat{Y}_{t+1} - \hat{Y}_t = \alpha(\varepsilon_t)$.

- Exponential smoothing provides an effective mechanism for forecasting, especially when we have only a few observations in hand for conducting the forecast process. This method is appropriate for series that move randomly above and below a constant mean. However, if the series presents a trend or seasonal pattern, some modification is required.

3. Exponential Smoothing with Trend - Holt's Model

Holt's two-parameter exponential smoothing model extends simple exponential smoothing to include a linear-trend component. Accordingly, Holt's model is appropriate for non-stationary data. We shall briefly present the model below:

3.1. The Holt's Model

$$F_{t+1} = \alpha Y_t + (1-\alpha)(F_t + T_t) \tag{3.1.1}$$

$$T_{t+1} = \beta(F_{t+1} - F_t) + (1-\beta)T_t \tag{3.1.2}$$

$$H_{t+m} = F_{t+1} + mT_{t+1} \tag{3.1.3}$$

Notations:

$F_{t+1} =$ forecast value for period $t+1$ made at time t, which can be denoted as $\hat{Y}_{t+1}$

$Y_t =$ actual value in period t

$F_t =$ forecast value for period t made by $t-1$.

$T_t =$ trend

$\alpha =$ smoothing constant for the data ($0 < \alpha < 1$)

$\beta =$ smoothing constant for the trend estimate ($0 < \beta < 1$)

$m =$ number of periods ahead to be forecast

$H_{t+m} =$ Holt's forecast value of period $t+m$.

The model proposed by Holt contains two smoothing constants, one for the level of the series, and one for the trend. In equation (3.1.1), the smoothing value, F_{t+1} , is predicted based on the current observation and the previous smoothed value. However, the latter is adjusted by adding a trend factor. The trend in equation (3.1.2) evolves by weighting the average of the recent change of the smoothed value and the previous trend.

Equation (3.1.3) is a forecast equation which is used to forecast m periods into the future by using both updated smoothing value, F_{t+1}, and trend estimate, T_t, derived from Equations (3.1.1) and (3.1.2). To conduct forecasts in this model m periods ahead, we can follow the same procedure as that of simple exponential-smoothing model. What we have to add here is a trend variable. The initial trend value is usually set at 0; the increment is advanced by 1 over time. Here, we need to search for two smoothing constants, α and β by using MSE or RMSE criterion.

4. Winters' Exponential Smoothing

Due to the fact that previous models ignored the seasonal component, Winters' three-parameter, exponential-smoothing model extends Holt's model by adding a seasonality factor, which is itself smoothed. Accordingly, we have three smoothing parameters, one for

actual data, one for trend, and one for seasonal factors. Since a new variable is added to the system, there are four equations in Winters' model.

4.1. The Winters' Model

$$F_t = \alpha(Y_t / S_{t-p}) + (1-\alpha)(F_{t-1} + T_{t-1}) \qquad (4.1.1)$$

$$S_t = \gamma(Y_t / F_t) + (1-\gamma)S_{t-p} \qquad (4.1.2)$$

$$T_t = \beta(F_t - F_{t-1}) + (1-\beta)T_{t-1} \qquad (4.1.3)$$

$$W_{t+m} = (F_t + mT_t)S_{t+m-p} \qquad (4.1.4)$$

Notations:

F_t = smoothed value of the level of series for period t

F_{t-1} = smoothed value for period t-1

Y_t = actual value in period t

T_t = trend estimate

S_t = seasonality estimate

α = smoothing constant for the data ($0 < \alpha < 1$)

β = smoothing constant for the trend estimate ($0 < \beta < 1$)

γ = smoothing constant for seasonality estimate ($0 < \gamma < 1$)

p = number of periods in seasonal cycle

m = number of periods ahead to be forecast

W_{t+m} = Winters' forecast for m periods into the future

A special feature of this model is that the element of a seasonal factor is added to the model. Equation (4.1.1) is similar to Holt's equation for smoothing the trend. A minor difference is that seasonal fluctuations in Y_t have been removed. As can be seen in the first term, Y_t is divided by S_{t-p} to adjust for seasonality. The seasonality estimate and trend are updated in a fashion similar to the simple exponential process as described Equations (4.1.2) and (4.1.3). Finally, Equation (4.1.4) is employed to compute the forecast for m periods into the future.

5. Adaptive-Response Approach

An alternative to simple exponential smoothing for stationary and non-seasonal time series is the adaptive-response approach to single exponential smoothing model in which an adaptive algorithm is used to determine a time-varying smoothing parameter. Accordingly, adaptive smoothing has the ability to adapt to a changing mean of an otherwise stationary and nonseasonal time series.

5.1. The Adaptive-Response model

$$F_{t+1} = \alpha_t Y_t + (1 - \alpha_t) F_t \qquad (5.1.1)$$

$$\alpha_t = \frac{S_t}{|A_t|} \qquad (5.1.2)$$

$$S_t = \beta e_t + (1 - \beta) S_{t-1} \qquad (5.1.3)$$

$$A_t = \beta |e_t| + (1 - \beta) A_{t-1} \qquad (5.1.4)$$

$$e_t = Y_t - F_t \qquad (5.1.5)$$

The basic equation (5.1.1) for forecasting is generally the same as the simple exponential smoothing model represented by equation (2.1.1). The only difference is that the smoothed term α in Equation (2.1.1) is replaced by α_t, and the later is adaptive over time, governed by the value of the smoothed error divided by the absolute smoothed error as expressed by Equation (5.1.2). Like most exponential smoothing models, both S_t (not a notation for seasonality) and A_t are smoothed out by using Equations (5.1.3) and (5.1.4).

The forecasting procedure for this model can be proceeded recursively. Given the values of Y_t and F_t, we can estimate e_t from (5.1.5). The e_t then is plugged into (5.1.3) and (5.1.4). Given the estimated value of β, we can obtain α_t by using (5.1.2). Finally, we use α_t, Y_t and F_t to predict F_{t+1} by employing equation (5.1.1).

5.2 Remarks

The advantage of this model is that it allows the smoothing value to change over time. However, the underlying rationale for time varying is less clear. If we want to have a time-varying coefficient model to reflect the changing pattern, the behavior of α_t can be specified as a random coefficient or autoregressive form, depending on the nature of the process. Another drawback for this model is that there is no explicit way to handle seasonality. Thus, in facing seasonal data, the data must first be deseasonalized and then reseasonalized to

generate forecasts.   Adaptive smoothing is an alternative to Winters' smoothing when handling seasonal time series.

A major flaw with smoothing models is their inability to predict cyclical reversals in the data, since forecasts depend solely on the past. Perhaps even more pernicious is the possibility of spurious cycles, since all smoothing models produce serially correlated forecasts.

Threaded question

Assume you were to use α values of 0.1, 0.5, and 0.95 in a simple exponential smoothing model. How would these different α values weight past observations of the variable to be forecasted? How would you know which of these α values provides the best forecasting model?   If the α = 0.99 value provides the best forecast for your data at the current calculation, would this imply that you should make forecasts in the future based on availability of new data? Does exponential smoothing place more or less weight on the most recent data when compared with the moving-average method? What weight is applied to each observation in a moving-average model? Why is smoothing (simple, Holt's, and Winters') also termed *exponential* smoothing?

Assignment

Consider the monthly stock index data for the US (USsp) market from 1995.01-2000.06 available to the class. Use Excel to calculate both the 12-month moving-average and simple exponential smoothing (α = 0.6) for these data and compare the forecasts of these two methods by calculating the root-mean-squared errors.   Make brief comments on your findings.

(Note the data file is: Monthly_US_JP_MacroData and variable name is USsp).