Time-series methods: Marketing Engineering Technical Note[1]

Table of Content

Introduction

The logic behind time series methods is that past data incorporate enduring patterns that will carry forward into the future and that can be uncovered through quantitative analysis. Thus the forecasting task becomes, in essence, a careful analysis of the past plus an assumption that the same patterns and relationships will hold in the future. There are a number of time-series analysis and forecasting methods, differing mainly in the way past observations are related to the forecasts. Many of these methods, such as moving averages, and exponential smoothing are available in Excel, as add-ins.

Time Series Techniques

Smoothing techniques: The notion underlying smoothing methods is that

[1] This technical note is a supplement to Chapter 5 of Principles of Marketing Engineering, by Gary L. Lilien, Arvind Rangaswamy, and Arnaud De Bruyn (2007). Visit www.decisionpro.biz for additional information.

there is some specific pattern in the values of the variables to be forecast, which is represented in past observations, along with random fluctuations or noise. Using smoothing methods, the analyst tries to distinguish the underlying pattern from the random fluctuations by eliminating the latter. For example, by averaging out short-term fluctuations in a sales data series could reveal the longer-term patterns or cycles in sales.

Formally, for simple moving averages let

S_t = forecast at time t,

X_t = actual value at time t, and

N = number of values included in average.

Then forecasting with ***moving averages*** can be represented as

$$S_{t+1} = \frac{1}{N}\sum_{i=t-N+1}^{t} X_i = \frac{X_t - X_{t-N}}{N} + S_t, \tag{1}$$

Thus, the moving average is simply the unweighted mean of the previous N observations. Eq. (1) makes it clear that the new forecast S_{t+1} is a function of the preceding moving-average forecast S_t. Furthermore, if X_t corresponds to a change (e.g., step change) in the basic pattern of variable X, it is difficult for the method to account for that change. Note also that the larger N is, the smaller $(X_t\text{-}X_{t-N})/N$ will be and the greater the smoothing effect will be.

In the *double moving average*, one starts by computing a set of single moving averages and then computes another moving average based on the values of the first. With a trend, a single or double moving average lags the actual series. Also, the double moving average is always below the simple moving average. Thus it is possible to forecast by taking the difference between the single moving average and the double moving average and adding it back to the single moving average. This forecasting technique is called the *double moving averages with trend adjustments*.

The *exponential-smoothing* approach is very similar to the moving-average method, differing in that the weights given to past observations are not constant—they decline exponentially so that more recent observations get more weight than earlier values. Choice of the smoothing factor is left to the analyst. Most often the analyst selects a value experimentally from a set of two or three different trial values.

With the foregoing notation, the procedure can be represented by

$$S_{t+1} = \alpha X_t + (1-\alpha) S_t, \tag{2}$$

where $0 \leq \alpha \leq 1$ is selected empirically by the analyst. A high value of α gives past forecasts and past data (included in S_t) little weight, whereas a low value of α weights the most recent period very lightly compared with all other past observations.

The method of *double exponential smoothing* is analogous to that of double moving averages, and easily adapts to changes in patterns, such as step changes.

Adaptive filtering (i.e., removing noise from signal) is another approach for determining the most appropriate set of weights, where the weights change to adjust to the changes in the time series being filtered. Notice that all the methods outlined so far are based on the idea that a forecast can be made as a weighted sum of past observations:

$$S_{t+1} = \sum_{i=t-N+1}^{t} W_i X_i, \tag{3}$$

where

S_{t+1} = *forecast for period t+1;*

W_i = *weight assigned to observation i;*

X_i = *observed value at i, as before; and*

N = *number of observations used in computing* S_{t+1} *(and so is equal to the number of weights required).*

The weights are determined by an iterative process that minimizes the average mean-squared forecasting error.

Box-Jenkins: This refers to a class of methods and a philosophy for approaching forecasting problems. Using it an analyst can develop an adequate model for almost any pattern of data. However, it is sufficiently complex that its users must have a certain amount of expertise. Box and Jenkins propose three general classes of models for describing any type of stationary process (processes that remain in equilibrium about a constant mean level): (1) autoregressive (AR), (2) moving average (MA), and (3) mixed autoregressive and moving average (ARMA).

If a series is increasing or decreasing with time, we can remove this (trend) by taking differences,

$$\Delta Y_t = Y_t - Y_{t-1}. \tag{4}$$

and then developing an ARMA model for ΔY_t. The original series Y_t can be recovered by successively adding in the ΔY_t, starting at Y_0. If the trend is nonlinear, several successive differences (d) may be required to produce a stationary ARMA series. (Recall that if you differentiate $Y=X^2$ twice—d^2Y/dX^2—you get a constant, 2. The differencing operation here is analogous and produces the same result.) Again, the original series can be recovered by summing d times. Such a series is called an integrated ARMA series, denoted as ARIMA (p, d, q), where p is the order (number of periods used) of the AR part, q is the order of the MA part, and d is the level of difference used to produce stationarity. There are also multivariate extensions of the ARMA models, known as multivariate ARMA, or MARMA. They combine powerful time-series forecasting techniques with explanatory variables and causal models (Hanssens, Parsons, and Schultz 1990). Applying the ARMA and MARMA methods requires more technical expertise and experience than many of the other methods we describe.

EXAMPLE

Exhibit 1 shows how some of these forecasting methods perform on data drawn from the National Bureau of Economic Research. Using the mean-absolute-percent error (MAPE) as the measure of forecasting ability, Box-Jenkins does best in this case. However, the naive method is the third best out of the six methods, suggesting that more sophisticated methods do not always perform better than simple ones.

Year	Q1*	Q2	Q3	Q4
1969	11,445	11,573	11,516	11,990
1970	11,704	11,050	11,069	10,705
1971	10,729	10,931	11,832	12,172
1972	12,472	12,840	12,865	13,491
1973	14,324	14,684	14,689	15,473
1974	16,483	16,634	17,245	17,177
1975	16,230	16,562	17,614	18,318
1976	19,148	19,730	19,184	19,424
1977	20,774	21,184	21,052	22,121

* Q1 = quarter 1, and so on.

(a)

1978	(1) Actual	(2) Naive	(3) Averaged on Four Previous Quarters, Moving Average	(4) Moving Average with Trend Adjustment
Q1	22,433	22,121	21,283	22,666
Q2	23,792	22,433	21,698	23,219
Q3	23,980	23,792	22,350	23,772
Q4	25,840	23,980	23,082	24,325
MAPE*	3.78	3.78	7.85	2.53

Exponential Smoothing

1978	(5) $\rho = 0.90$	(6) $\rho = 0.50$	(7) Optimal Box-Jenkins
Q1	22,014	21,397	23,168
Q2	22,391	21,915	23,509
Q3	23,652	22,853	24,133
Q4	23,947	23,416	25,141
MAPE*	4.10	6.65	1.95

*MAPE = mean-absolute-percent error $= \frac{1}{n}\Sigma\,(|\text{actual} - \text{forecast}| / \text{actual}) \times 100.$

(b)

EXHIBIT 1

A comparison of the forecasting accuracy of six forecasting methods; (a) gives actual data for fabricated metal products while (b), columns (2) through (7), gives the forecasting accuracy of six methods. *Source*: National Bureau of Economic Research Series MDCSMS.

Decompositional methods: The forecasting methods described thus far are based on the idea that we can distinguish an underlying pattern in a data series from noise by smoothing (averaging) past values. The smoothing eliminates noise so that we can project the pattern into the future and use it as a forecast. These methods make no attempt to identify individual components of the basic underlying pattern. However, in many cases we can break the pattern down (decompose it) into sub-patterns that identify each component of the series separately. With such a breakdown we can frequently improve accuracy in forecasting and better understand the series.

Decompositional methods assume that all series are made up of patterns plus error. The objective is to decompose the pattern of the series into trend, cycle, and seasonality:

$$X_t = f(I_t, T_t, C_t, E_t). \tag{5}$$

where

X_t = *time series at time t;*

I_t = *seasonal component (or index) at t;*

T_t = *trend component at t;*

C_t = *cyclical component at t; and*

E_t = *error or random component at t.*

The exact functional form of Eq. (5) depends on the decompositional method used. The most common form is a multiplicative model:

$$X_t = I_t \times T_t \times C_t \times E_t. \tag{6}$$

An additive form is used often, as well. An example of additive decomposition is given in Exhibit 2.

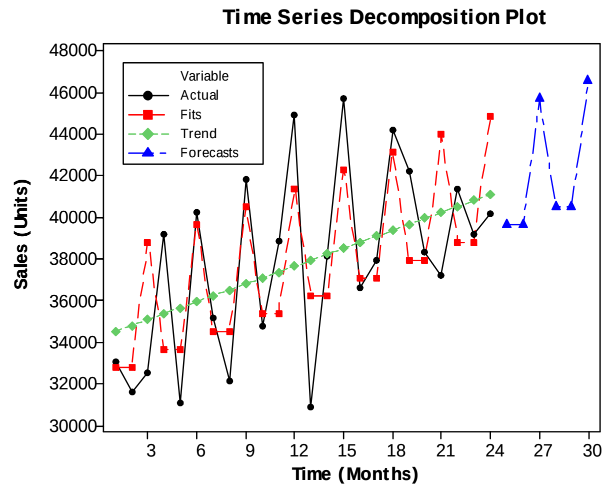

EXHIBIT 2:
This chart how the actual data for25 periods are decomposed into two additive components, namely trend and seasonality to create the fitted data. We also show how the forecasts would be generated from this series for periods 25 to 30.

Although there are a number of decompositional methods, they all seem to follow the same basic process:

1. For the series X_t compute a moving average of length N, where N is the length of the seasonality (e.g., N=12 with monthly data). This averaging will eliminate seasonality by averaging seasonally high periods with seasonally low periods; and because random errors have no systematic pattern, it reduces randomness as well.
2. Separate the outcome of the N-period moving average from the original data period to obtain trend and cyclicality. If the model is multiplicative, you do this by dividing the original series by the smoothed series, leaving seasonality and error:

$$\frac{X_1}{T_t + C_t}(= \text{moving average}) = \mathrm{I}_t \times E_t. \tag{7}$$

3. Isolate the seasonal factors by averaging them for each data point in a season over the complete length of the series.
4. Specify the appropriate form of the trend (linear, quadratic, exponential)

and calculate its value at each period T_t. You can do this by using regression analysis or moving averages with trend adjustments.

5. Use the results to separate out the cycle from the trend+cycle (i.e., the moving average).
6. When you have separated the seasonality, trend, and cyclicality from the original data series, you can identify the remaining randomness, E_t.

Decompositional methods are widely used and have been developed empirically and tested on thousands of series. Although they do not have a sound statistical base, the methods are intuitive and geared to the practitioner and, therefore, the opposite of such procedures as the Box-Jenkins approach, which is derived from theory. Decompositional methods appear to be most appropriate for short- or medium-term forecasting and are mainly suited to macroeconomic series.

Summary

Time series methods help marketers to generate forecasts as a function of time, i.e., what will happen to a particular data series (e.g., sales, trial rate) at time t in the future? Unlike naïve methods that simply project the past onto the future, the methods described in this note attempt to isolate the enduring patterns hidden in a data series by removing noise from the signal. Thus, time series methods are most useful in situations when enduring patterns repeat themselves in the future. They are also most useful when our interest centers mainly on forecasting, and not on explaining or diagnosing a data pattern. We have described some simple techniques that can be implemented in spreadsheets. More complex methods such as Kalman filters, Bayesian filters, and Fuzzy filters, can also be used in more sophisticated systems for analysis of time series data.

References

Hanssens, Dominique M.; Parsons, Leonard J.; and Schultz, Randall L., 1990, *Market Response Models: Econometric and Time Series Analysis*, Kluwer Academic Publishers, Boston.