2011 International Conference on Economics and Finance Research
IPEDR vol.4 (2011) © (2011) IACSIT Press, Singapore

Fortified financial forecasting models: non-linear searching approaches

Mohammad R. Hamidizadeh, Ph.D.
Professor, Mgt. and Accounting Faculty,
Shahid Beheshti University, USB
Eveen, Tehran, Iran
E-mail: M-Hamidizadeh@sbu.ac.ir

Mohammad E. Fadaeinejad, Ph.D.
Associate Professor, Mgt. and Accounting Faculty,
Shahid Beheshti University, SBU
Eveen, Tehran, Iran
E-Mail: M-Fadaei@ sbu.ac.ir

***Abstract*- When researchers are going to forecast data with variations involving at times series data, they believe the best forecasting model is the one which realistically considers the underlying causal factors in a situational relationship and therefore has the best "track records" in generating data for forecasting. Naïve forecasting models may be adjusted for variations in related data. When a time series processes a great deal of randomness, smoothing techniques may improve the accuracy of the financial forecasts. But neither Naïve models nor smoothing techniques are capable to identify the major future changes in the direction of a situational data series. Hereby, Nonlinear techniques, like direct and sequential search approaches, overcome those shortcomings can be used. This paper is to manipulate how to prepare and integrate heuristic non-linear searching methods to serve calculating adjusted factors to produce the best forecast data.**

Keyword; Naïve forecasting models; Smoothing techniques; Fibonacci and Golden section search; Line search by curve fit.

I. INTRODUCTION

The article's aim is to construct, give a new method and to present a proper analysis of the factors which will effect the forecasting equations. The new calculating process is based on nonlinear intelligent methods which searching to find a new point in time series processes for residual minimization. We will try to present how to fortify forecasting financial models with nonlinear techniques characteristics based on line search methods for economic, finance and business forecasting techniques.

You may study many interesting nonlinear techniques on the one hand on Luenberger (1989) and Bazaraa (1993), but on the other hand, with analysis of forecasting techniques that is presented by McGuigan (1999), Gourieroux and Jasiak(2001), this idea got to present itself to improve significantly its predictive power by applying nonlinear methods advantages of generating stationary points in different ways to advocate the adjustment factor of the forecasting techniques to extent that minimize difference $e_t = (Y_t - \hat{Y}_t)$ or $(Y_t - Y_{t-1})$ [1]. As in the univariate case, the pure vector autoregressive process of order 1[VAR (1)] provides a simple framework for exploring the multiplier effects and forecasting. Moreover, this multivariate autoregressive process can accommodate quite complex dynamics of individual component series. In empirical research, the VAR (1) model often provides a satisfactory fit to multivariate return series. Therefore, the process (y_t; t<Z) is integrated 1, denoted I(1), if and only if it satisfies the recursive equation $y_t = y_{t-1} + e_t$; where (e_t) is a weak white noise [3]. This paper tries to minimize e_t [2, 9].

II. INITIAL PREREQUISITE

Researchers often deal with processes that vary as time passes. Time series are analyzed to better understand, describe, control, and predict the underlying process. The analysis usually involves a study of the components of the time series (TS) – such as secular trends, cyclical variations, seasonal effects, and stochastic variations- the particular components of interest tendency to vary from one problem to another [13].

The forecast is merely a prediction concerning the future and is required in virtually all areas of the operation [12]. Obviously, good forecasting is essential to reduce the uncertainty of the TS in which most decisions are made. The level of sophistication required in forecasting techniques varies directly with significance of the problem being examined [1, 8].

The forecasting techniques used in any particular situation depend on two major factors: aggregation factors such as international / regional economics; national economics; Industry economics; individual firm economics; the internal of the firm. Diagnostic factors are cost and potential gain; the complexity of the relationships; time period; accuracy required; lead time [4].

III. FORECASTING TECHNIQUES

I. Naïve forecasting models are based on TS observations of the values of the variable being forecast. Let ŷ denote the forecast data of the variable of interest, y denote an actual observed data of the variable and the subscript t identify the time period. The simplest model states that the forecast value of the variable for the next period will be the same as the value of that variable for the present period [7].

If a recognizable pattern exists, it may be incorporated by adjusting related equations. Two possible cases are a linear trends and nonlinear trends. A linear trend follows a constant rate of growth pattern [13]. Since the related equation is a nonlinear relationship, the parameters cannot be estimated directly with the least-squares method.

When seasonal variations are introduced into a naive forecasting model, it may be possible to improve significantly its short-run predictive power. The best known of these is the ratio-to-trend method. Trend projection equations are most useful for intermediate and long-term forecasting [1, 9].

Another model that is often very effective in generating forecasts when there is a significant seasonal component is the exponentially weighted moving-average (EWMA) forecasting model. At each point in time, the EWMA model estimates a smoothed average from past data, an average from past data, an average trend gain, and the seasonal factor. These three components are then combined to compute a forecast [3].

II. Smoothing techniques are higher form of naïve forecasting models which assume that an underlying pattern can be found in the TS of a variable that is being forecast. It is assumed that these historical observations represent not only the underlying pattern but also random variations. By taking some form of an average of post observations, smoothing techniques attempt to eliminate the distributions arising from random variations in the series and to those the forecast on a smoothed average of several past observations [8, 9].

A. Moving average.

If a data series possesses a large random factor, in an effort to eliminate the effects to this randomness, a series of recent observations can be average to arrive at a forecast. This is the moving average method. A number of observed values are chosen, their average is computed, and this average serves as a forecast for the next period.

1. Simple MA may be defined as:

$$\hat{Y}_{t+1} = \frac{1}{N}\left[\sum_{t=1}^{N} Y_t\right] = \hat{Y}_t + \frac{1}{N}\left[Y_t - Y_{t-N}\right] = \hat{Y}_t + \frac{1}{N}\Delta Y$$

(9)

$$\hat{Y}_{t+1} = \hat{Y}_t + \alpha$$

Where α is $\frac{1}{N}\Delta Y$ [7].

2. Weighted moving average may be viewed as a centered MA, where the observations in a sequence don't receive equal weights.

$\hat{Y}_{t+1}=\alpha[\Sigma w_t y_t]$, Where α is the reverse of the sum of the weights, and N is the number of observations in a moving average. The greater the number of observations N used in moving average, the greater the smoothing effect because each new observation receive less weight (1/N) as N increases. Hence, generally, the greater the randomness in the data series and the slower the change in the underlying pattern, the more preferable it is to use a relatively large number of past observations in developing the forecast.

The choice of an appropriate MA period, that is, the choice of N, should be based on a comparison of the results of the model in forecasting past observations. One criteria that is often used for such comparative purpose is the minimization of the average forecast error, or root mean square error.

$$RMSE = \left[\Sigma \bar{d}^2\right]^{1/2} = \sqrt{\frac{1}{m}\sum_{t=1}^{m}(\hat{y}_t - y_t)^2}$$

Where m is the number of time periods [8].

B. First-order exponential smoothing.

To forecast a stationary time series with the exponentials smoothing model, we first smooth the time series with a MA similar to those described in order to isolate the systematic or smooth component of the series. We then project this smooth component into the future. The MA employed by the exponential smoothing model for stationary TS is a special type of weighted MA.

The smooth component of a stationary time series may be considered as a succession of estimate of the underlying mean level of the stationary process. Update the smoothed estimate:

$$Y_t^s = \alpha y_t + (1-\alpha)Y_{t-1}^s$$

$$Y_t^s = Y_{t-1}^s + \alpha(Y_t - Y_{t-1}^s) = Y_{t-1}^s + \alpha d_t$$

$$Y_t^s = \alpha Y_t + \alpha(1-\alpha)Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + \alpha(1-\alpha)^3 Y_{t-3} + ...$$

$$\hat{Y}_{t+k} = Y_t^s, k = 1,2,3,....$$

Where Y_t^s is the smoothed estimate for the current period t, Y_{t-1}^s is the smoothed estimate for the proceeding period t-1, y_t is the observation for the current period t, α is the smoothing constant, and $\hat{Y}_{t+k}$ the forecast for k periods ahead. These models weight the most recent observation by $0<\alpha<1$, and the past forecasting by (1-α). A large α indicates that a heavy weight is being placed on the most recent observation [8].

C. Double exponential smoothing Models.

Data that are collected over time frequently exhibit a linear trend when a linear trend is not apparent in the data, single exponential smoothing is well suited as a forecasting technique. How if a trend is present in the data, DES is more appropriate for obtaining forecasts than is single ES. The DES is closely related to the trend analysis technique of simple linear regression, in which the forecast for Y_{t+1}, given TS up to and including time t. In DES, the coefficients a_t and b_t are considered to be functions of time and are updated as each successive observation becomes available. These updated coefficients are given by the following expressions:

$$\hat{y}_{t+1}=a_t+b_t(1+t)$$

$$b_t=[\alpha(1-\alpha)](S_t-S_{t-1}^{(2)})$$

$$a_t=2S_t-S_t^{(2)}-tb_t$$

$$S_t=\alpha y_t+(1-\alpha)S_{t-1}$$

$$S_t^{(2)}=\alpha S_t+(1-\alpha) S_{t-1}^{(2)}$$

Where α is a smoothing constant chosen to minimize the sum of the squared forecast errors, S_t is the single smoothed statistic computed recursively, $S_t^{(2)}$ is the double smoothed statistic also computed recursively, and the initial value S_0 and $S_0^{(2)}$ are obtained using the related equations as follows:

$$S_0=\alpha_0-[(1-\alpha)/\alpha]b_0$$

$$S_0^{(2)}=\alpha_0-2[(1-\alpha)/\alpha]b_0$$

The quantities α_0 and b_0 are initial values of the regression coefficients. If data are available, say, to time t, α_0

and b_0 are simply regression coefficients computed using the model $Y_t=\alpha_0+b_0t$. If no data are available, values for α_0 and b_0 are assigned subjectively, or if this is not possible S_0 and $S_0^{(2)}$ are both assigned the initial value of the series, Y_0 [8,13].

D. ES for TS with trend secular.

Step1. Update the smoothed estimate:

$$Y_t^s = \alpha Y_t + (1-\alpha)(Y_{t+1}^s + d_{t-1})$$

*Step2.*Update the trend estimate:

$$d_t = b(Y_t^s - Y_{t-1}^s) + (1-b)\, d_{t-1} \qquad ,0<b<1$$

Step3. Forecast for period t+k:

$$\hat{Y}_{t+k} = Y_t^s + kd_t$$

Where s_t is the trend estimate for the current period t, d_{t-1} is the trend estimate for the preceding period t-1, b is the trend adjustment constant, and Y_{t+k} is the forecast k periods ahead for future period t + k [12].

E. ES for TS with trend and seasonal components.

Step 1. Update the smoothed estimate:

$$Y_t^s = \alpha(\frac{Y_t}{s_{t-p}}) + (1-\alpha)(Y_{t-1}^s + d_{t-1}) \qquad ,$$

0<a<1

Step 2. Update the trend estimate:

$$d_t = b(Y_t^s - Y_{t-1}^s) + (1-b)d_{t-1}$$

Step 3. Update the smoothed index estimate:

$$S_t = s(\frac{Y_t}{Y_t^s}) + (1-s)S_{t-p}$$

Step 4. Forecast for period t+k:

$\hat{Y}_{t+k} = (Y_t^s + kd_t)S$, Where s_t is the seasonal index estimate for current period t, S_{t-p} is the seasonal index estimate for period t-p, one seasonal cycle earlier, and s is the seasonal index adjustment constant, 0<s<1 [8, 9].

IV. NONLINEAR TECHNIQUES

It is time to turn to a description of the nonlinear techniques used for the value of adjusted factor based on iteratively solving unconstrained variations minimization. These techniques are, of course, important for practical projection since they often offer the simplest, most direct alternatives for obtaining solutions; but perhaps their greatest importance is that they establish certain reference plateaus with respect to difficulty of implementation and speed of convergence. There is a fundamental underlying structure for almost all the descent algorithms we discuss [11]. One starts at an initial observation determines, according to a fixed rule, a direction of movement; and then moves in that direction to a (relative) minimum of the TS error on that line. At the new predicted value a new direction is determined and the process is repeated. Once the selection is made, all algorithms call for movement to the minimum forecast observes on the corresponding component [5].

The process of determining the minimum $\hat{Y}_t$ on a given variation is called variation search [9]. For general nonlinear functions that can not be minimized analytically, this process actually is accomplished by searching, in an intelligent manner, along the variation for the minimum forecast value error [10]. These line search techniques form the backbone of nonlinear programming algorithms, since higher dimensional problem are ultimately solved by executing a sequence of successive line searches [8].

I. FIBONACCI AND GOLDEN SECTION SEARCH

A. Fibonacci search.

The method determines the minimum value of a time series over a closed interval $[C_1,C_2]$. In projections, a TS may in fact be defined over a broader domain, but for this method a fixed interval of search must be specified [6].

To develop an appropriate search strategy, that is, a strategy for selecting observations based on the previously obtained values, we pose the following process: Find how to successively select N observations so that, with explicit knowledge of TS, we can determine the smallest possible region of uncertainty in which the minimum residual must lie. Thus, after values are known at N observations $y_1, y_2\ldots y_n$ with $C_1 \leq Y_1 < Y_2 < \ldots < Y_{n-1} < Y_n \leq C_2$

The region of uncertainly is the interval $[Y_{k-1}, Y_{k+1}]$ where $Y_k(=d_k)$ is the minimum error among the N, and we define $Y_0=C_1$, $Y_{n+1}=C_2$ for consistency. The minimum of d_k must lie some where in this interval.

The derivation of the optimal strategy for successively selecting observations to obtain the smallest region of uncertainty is fairly straight forward but somewhat tedious.

Let

$\alpha_1=C_2-C_1$, The initial width of uncertainly

α_k= width of uncertainly after k measurements.

Then, if a total of N observations are to be made, we have

$$\alpha_k = (\frac{F_{n-k+1}}{F_n})\alpha_1 = W\alpha_1$$

,Where the integers F_k are members of the Fibonacci sequence generated by the recurrence relation:

$F_N=F_{N-1}+F_{N-2}$,$F_0=F_1+1$, The resulting sequence is 1, 1,2,3,5,8,13,…

The procedure for reducing the width of uncertainty to α_N is this; The first two observations are made symmetrically at a distance of $(\frac{F_{N-1}}{F_N})\alpha_1 = W_1\alpha_1$ from the ends of the initial intervals; according to which of these is a lesser value, an uncertainty interval of width $\alpha_2=(W_1\alpha_1)$ is determined, the third observation is placed symmetrically in this new interval of uncertainty with respect to the measurement already in the interval. The results of this step gives an interval of uncertainty $\alpha_3=(F_{N-2}/F_N)\ \alpha_1=w_2\alpha_1$. In general, each successive observation is placed in the current interval of uncertainty symmetrically with the point already existing in that interval [6].

B. Search by Golden section.

If the number N of allowed observations in a Fibonacci search is made to approach infinity, we obtain the golden section method. It can be argued, based the optimal properly of the finite Fibonacci method, that the corresponding initial version yields a sequence of intervals of TS whose width tend to zero faster than that which would be obtained by other methods. The solution to the Fibonacci equation $F_N=F_{N-1}+F_{N-2}=\alpha f_1^N+b f_2^N$ is to finding, where f_1 and f_2 are roots of the characteristic equation $f^2=f+1$. Explicitly $f_1\approx1.618$ and $f_2\approx-0.618$. Therefore, 1.618 is known as the golden section ratio.

For large N the first term on the right side of f_N equation dominates the second and hence $\lim_{N\to\infty}\frac{F_{N-1}}{F_N}=\frac{1}{F_1}\cong 0.618$

It follows from (16) that the interval of TS at any observations in the TS has width $\alpha_t=\alpha_1.(\frac{1}{f_1})^{t-1}$

And from this it follows that $\frac{\alpha_{t+1}}{\alpha_t}=\frac{1}{f_1}=0.618$

Therefore, we conclude that, with respect to the width of the TS interval, the search by golden section covers linearly [4, 6].

II. LINE SEARCH BY CURVE FITTING.

A. Newton's method.

Suppose that the TS error of a single variable Y is to minimized, and suppose that at a value y_k where an observation is made it is possible to evaluate the forecast, first derivative ($\triangle$) and second derivative ($\triangle^2$) of the predicted values on unit of time t: $\hat{y}_t$, $\triangle\hat{y}_t$, $\triangle^2\hat{y}_t$. Hereby, it is possible to construct a quadratic time series q which at y_t agrees with TS up to second derivatives ($\triangle^2$), that is

$$q(y)=\hat{Y}_1+\triangle\hat{Y}_1(y_0-y_1)+\frac{1}{2}\triangle^2\hat{Y}_t(Y_0-Y_{-1})^2$$

Now, we can calculate an estimate Y_{t+k} of minimum residual of TS by finding the residual where the derivative of q vanishes. Thus, setting

$$0=\triangle\hat{Y}_t+\triangle^2\hat{Y}_t(Y_{t+1}-Y_t)$$

Or $\triangle\hat{Y}_t+\triangle^2\hat{Y}_t.\ \triangle Y=0$

We then find forecast for period t+k: $\hat{Y}_{t+k}=Y_t+\alpha_t$, k=1,2,… where $\alpha_t=\triangle Y_t/\triangle\hat{Y}_t$.

B. Method of false position.

Newton's method for residual minimization is based on fitting a quadratic on the basis of information at a single observation; by using more points, less information is required at each of them. . Thus, using

$\hat{Y}_t,\triangle\hat{Y}_t$, $\triangle\hat{Y}_{t-1}$ it is possible to fit the quadratic

$$q(Y)=\hat{Y}_t+\triangle\hat{Y}_t(Y_0-Y_t)+\frac{\Delta\hat{Y}_{t+1}-\Delta\hat{Y}_t}{Y_{t-1}-Y_t}.\frac{(Y_0-Y_t)}{2}$$

which has the same corresponding values. An estimate Y_{t+k} can then be determined by finding the observation where the derivative of q vanishes; Thus, we find forecast for period t+k:

$\hat{Y}_{t+k}=Y_t-\alpha.\triangle\hat{Y}_t$ Where $\alpha=[Y_{t-1}-Y_t]/(\triangle\hat{Y}_{t-1}-\triangle\hat{Y}_t)$.

Thus, using $\hat{Y}_t,\triangle\hat{Y}_t$, $\triangle\hat{Y}_{t-1}$ it is possible to fit the quadratic which has the same corresponding values. An estimate Y_{t+k} can then be determined by finding the observation where the derivative of q vanishes; Thus, we find forecast for period t+k. Comparing this equation with Newton's method, we see again that the value $\hat{Y}_t$ does not enter; hence, our fit could have been passed through either $\hat{Y}_t$ or $\hat{Y}_{t-1}$. Also the model can be regarded as an approximation to Newton's model where the second derivative is replaced by the difference of two first derivatives[2].

C. Cubic Fit.

Given the observation, Y_{t-1} and Y_t with corresponding forecast and derivative values $\hat{Y}_{t-1}$, $\triangle\hat{Y}_{t-1}$, $\hat{Y}_t$, $\triangle\hat{Y}_t$, It is possible to fit a cubic model to the observations having corresponding values. The next observe Y_{t+1} can then be determined as the relative minimum point of this cubic. This leads to find forecast for period t+k; $\hat{Y}_{t+k}=Y_t-\alpha(Y_t-Y_{t-1})=Y_t-\alpha.\triangle Y_t$ with

$\alpha=[\triangle\hat{Y}_t+u_2-u_1]/[\triangle\hat{Y}_t-\triangle\hat{Y}_{t-1}+2u_2]$

Where $u_1=\triangle\hat{Y}_{t-1}+\triangle\hat{Y}_t-3([\hat{Y}_{t-1}-\hat{Y}_t])/(Y_{t-1}-Y_t)$

$u_2=[u^2_1-\triangle\hat{Y}_{t-1}\triangle\hat{Y}_t]^{1/2}$

It can be shown that the order of convergence of the cubic fit method is 2. Thus, although the method is exact for cubic models indicating that its order might be three; its order is actually only two[4, 5].

D. Quadratic Fit.

This method is often most useful in line searching by fitting a quadratic through three given actual and forecast observations. This has the advantage of not requiring first and second derivatives information. If we have points Y_1, Y_2, Y_3 and corresponding forecast values $\hat{Y}_1$, $\hat{Y}2$, $\hat{Y}_3$. We construct the quadratic passing through these points:

$$q(y)=\sum_{i=1}^{3}\alpha\hat{Y}_i$$

Where $\alpha=[\prod_{j\pm i}(Y_0-Y_j)]/[\prod_{j\pm i}(Y_i-Y_j)]$ and determine a new observation Y_4 as the point where the derivative of q vanishes. Thus

$$Y_4=\frac{1}{2}.\frac{b_{23}\hat{Y}+b_{31}\hat{Y}+b_{12}\hat{Y}_3}{a_{23}\hat{Y}+a_{31}\hat{Y}_2+a_{12}\hat{Y}_3}$$

Where $a_{ij}=Y_i-Y_j$, $b_{ij}=Y_i^2-Y_j^s$ [6].

III. CONCLUSION

This article could make connections between two isolation areas: forecasting techniques and nonlinear programming techniques. The aim of this interaction was to be used nonlinear optimization methods for finding forecasting values of interest variables based on line searching methods.

The initial difference of forecasting methods is mainly on adjusting the data mechanism for optimal prediction of variables; hereby, the article first analyzed different kinds of forecasting techniques especially on adjusting variation components power, then it followed some basic descent methods to minimize the time series residual in relation to actual values of variable.

Because of making stationary time series advantage of these methods, it has tried to show how to integrate this characteristic for generating fortified forecasting data.

REFERENCES

[1] J. Scott, Armstrong, "Research of forecasting: a quarter century review, 1969-1984." *Interfaces*, January-February 1986, Pp.89-109.

[2] M. S., Bazaraa, H.D. Sherali and C.M. Shetty, *Nonlinear programming: theory and algorithms*, N.Y.: John Wiley & Sons, 1993.

[3] C. Gourieroux, & J. Jasiak, *Financial econometrics*, New Jersey: Princeton University Press, 2001.

[4] M.R. Hamidizadeh, *Nonlinear programming*, Tehran: SAMT Publishing Co., 2002.

[5] J., Kowalik, and M.R. Osborne, *Methods for unconstrained optimization problems*, N.Y.: Elsevier, 1968.

[6] D. G. Luenberger, *Linear and nonlinear programming*, Massachusetts: Addison-Wesley Publishing Co., Reading,1989.

[7] J.R. McGuigan, & R.C. Moyer, *Managerial economics*, St. Paul: West Publishing Co., 1989.

[8] J.R. McGuigan, R.C. Moyer, and F.H. Harris, *Managerial economics: applications, strategy, and tactics,* Cincinati: South-Western College Publishing, 1999.

[9] J. Neter, W. Wasserman & G.A. Whitmore, *Applied statistics*, Boston: Allyn and Bacon, Inc., 1988.

[10] J.F. Traub, *Iterative methods for the solution of equation*, N.J.: Prentice-Hall Englewood Cliffs, 1964.

[11] D. j. Wilde, and C.S. Beightler, *Foundation of optimization*, N.J.: Prentice-Hall Englewood Cliffs, 1967.

[12] W.I. Zangwill, "Nonlinear programming via penalty functions," *Management Science* vol. 13, no 5, 1967, pp.344-358.

[13] V. Zarnowitz, "Recent work on business cycles in historical perspective", *Journal of Economic Literature*, June 1985, pp.523-580.