Journal of Southeast University (English Edition) Vol.20 No.1 Mar. 2004 ISSN 1003—7985

Parameter estimation methods in generalized weighted functional mean combining forecasting model

Wan Yucheng[1,2] Sheng Zhaohan[3]

([1] College of Economics and Management, Southeast University, Nanjing 210096, China)
([2] No.3 Department, Air Force Logistics College, Xuzhou 221002, China)
([3] School of Management Science and Engineering, Nanjing University, Nanjing 210093, China)

Abstract: A new kind of combining forecasting model based on the generalized weighted functional mean is proposed. Two kinds of parameter estimation methods with its weighting coefficients using the algorithm of quadratic programming are given. The efficiencies of this combining forecasting model and the comparison of the two kinds of parameter estimation methods are demonstrated with an example. A conclusion is obtained, which is useful for the correct application of the above methods.

Key words: combining forecasting; generalized weighted functional mean; parameter estimation; quadratic programming

The method of combining forecasting is an important method in forecasting. It has been paid more attention since Bates and Granger[1] originally proposed its theory and method in 1969. In solving the various actual forecasting problems, combining forecasting models may have different forms. But forecasting efficiencies may be different for different models. In order to get the best efficiencies of combining forecasting models, Wang and Fu[2] presented four kinds of aggregative methods of group forecasting. In this paper, we present a new kind of combining forecasting model based on the generalized weighted functional mean. This model is a new kind of aggregative method of group forecasting which has extensive representation. By seeking the appropriate parameters and determining the optimal combining form, we can effectively improve the forecasting precision. But, if an incorrect parameter estimation method is adopted, the effect will be weak. In order to make scientific forecasting, we present two kinds of parameter estimation methods with the weighting coefficients of the above model. Finally, this model is used to forecast air material consumption. Its efficiency is demonstrated with an example. Some useful conclusions are drawn, which can lay solid foundations for the correct application of the above methods.

1 Generalized Weighted Functional Mean Combining Forecasting Model

Suppose the real values of some forecasting problem in a period are $Y(t)$ $(t=1, 2, \cdots, n)$. There are m kinds of feasible individual forecasting models to forecast it, the forecasting values of which are $\hat{Y}_j(t)$ $(t=1, 2, \cdots, n; j=1, 2, \cdots, m)$. Furthermore, we assume that the weighting vector among the m kinds of models is $\boldsymbol{W}=\{w_1, w_2, \cdots, w_m\}^{\mathrm{T}}$, which satisfies the normalized constraint condition $\boldsymbol{E}^{\mathrm{T}}\boldsymbol{W}=1$ and the nonnegativity constraint condition $\boldsymbol{W}\geqslant 0$, where $\boldsymbol{E}=\{1, 1, \cdots, 1\}^{\mathrm{T}}$. We denote the values of the combining forecasting model by $\hat{Y}(t)$. Obviously, the less the approximation extents between $\hat{Y}(t)$ and $\hat{Y}_j(t)$, the better they are. Thus we choose the performance index as

$$\min J(t) = \sum_{j=1}^{m} w_j[(f(\hat{Y}(t)))^p - (g(\hat{Y}_j(t)))^p]^2 \qquad t=1, 2, \cdots, n; p\neq 0 \tag{1}$$

where f and g are both continuously differentiable functions, they should take the same functional forms customarily; and p is a nonzero variable parameter, which can take different values according to the different forecasting problems.

Suppose $(f(\hat{Y}(t))^{p-1}\neq 0$ and $\dfrac{\partial f(\hat{Y}(t))}{\partial \hat{Y}(t)}\neq 0$. Let $\dfrac{\partial J(t)}{\partial \hat{Y}(t)}=0$, we get

$$\hat{Y}(t) = f^{-1}\left[\left(\sum_{j=1}^{m} w_j(g(\hat{Y}_j(t)))^p\right)^{\frac{1}{p}}\right] \qquad t=1, 2, \cdots, n; p\neq 0 \tag{2}$$

Received 2003-09-18.

Biographies: Wan Yucheng (1970—), male, graduate; Sheng Zhaohan (corresponding author), male, professor, zhsheng@nju.edu.cn.

Formula (2) is the generalized weighted functional mean combining forecasting model to be given in this paper. It includes the following combining forecasting models.

1.1 Generalized weighted arithmetic mean combining forecasting model[2]

In formula (2), let $f(\hat{Y}(t)) = \hat{Y}(t)$ and $g(\hat{Y}_j(t)) = \hat{Y}_j(t)$, we have

$$\hat{Y}(t) = \left[\sum_{j=1}^{m} w_j(\hat{Y}_j(t))^p\right]^{\frac{1}{p}} \qquad t = 1, 2, \cdots, n;\ p \neq 0 \tag{3}$$

This is generalized weighted arithmetic mean combining forecasting model. And, when $p = 1$, it is a simple weighted arithmetic mean combining forecasting model. When $p = -1$, it is a simple weighted harmonic mean combining forecasting model. When $p = 2$, it is a simple weighted square sum mean combining forecasting model. When $p = 1/2$, it is a simple weighted square root sum mean combining forecasting model, etc.

1.2 Generalized weighted logarithmic mean combining forecasting model[2]

In formula (2), let $f(\hat{Y}(t)) = \ln\hat{Y}(t)$ and $g(\hat{Y}_j(t)) = \ln\hat{Y}_j(t)$, we have

$$\hat{Y}(t) = \exp\left[\left(\sum_{j=1}^{m} w_j(\ln\hat{Y}_j(t))^p\right)^{\frac{1}{p}}\right] \qquad t = 1, 2, \cdots, n;\ p \neq 0 \tag{4}$$

This is a generalized weighted logarithmic mean combining forecasting model. When $p = 1$, it is a simple weighted geometric mean combining forecasting model.

1.3 Generalized weighted exponential mean combining forecasting model

In formula (2), let $f(\hat{Y}(t)) = \exp\hat{Y}(t)$ and $g(\hat{Y}_j(t)) = \exp\hat{Y}_j(t)$, we have

$$\hat{Y}(t) = \ln\left[\left(\sum_{j=1}^{m} w_j(\exp\hat{Y}_j(t))^p\right)^{\frac{1}{p}}\right] \qquad t = 1, 2, \cdots, n;\ p \neq 0 \tag{5}$$

This is a generalized weighted exponential mean combining forecasting model.

Furthermore, if f and g are taken in any other forms of functions, we can get some additional new combining forecasting models.

2 Methods of Parameter Estimation in Generalized Weighted Functional Mean Combining Forecasting Model

In order to estimate the weighting coefficients of formula (2), we transfer formula (2) into:

$$(f(\hat{Y}(t)))^p = \sum_{j=1}^{m} w_j(g(\hat{Y}_j(t)))^p \qquad t = 1, 2, \cdots, n \tag{6}$$

If no errors are considered, there should ideally exist the following formula:

$$(f(Y(t)))^p = \sum_{j=1}^{m} w_j(g(\hat{Y}_j(t)))^p \qquad t = 1, 2, \cdots, n \tag{7}$$

In fact, forecasting errors always exist objectively and inevitably. So, Eq.(7) doesn't hold in general cases. Now, we can introduce the following error item:

$$\varepsilon_p(t) = (f(\hat{Y}(t)))^p - (f(Y(t)))^p = \sum_{j=1}^{m} w_j[(g(\hat{Y}_j(t)))^p - (f(Y(t)))^p] \tag{8}$$

Denote $\boldsymbol{E}_p = \{\varepsilon_p(1), \varepsilon_p(2), \cdots, \varepsilon_p(n)\}^{\mathrm{T}}$ and $\boldsymbol{Y}_p = ((g(\hat{Y}_j(t)))^p - (f(Y(t)))^p)_{n\times m}$, then, Eq.(8) can be written as $\boldsymbol{E}_p = \boldsymbol{Y}_p\boldsymbol{W}$. In order to make the combining forecasting model optimal, we should choose the appropriate weighting coefficients to make min $J_p = \sum_{t=1}^{n}\varepsilon_p^2(t) = \boldsymbol{E}_p^{\mathrm{T}}\boldsymbol{E}_p = \boldsymbol{W}^{\mathrm{T}}(\boldsymbol{Y}_p^{\mathrm{T}}\boldsymbol{Y}_p)\boldsymbol{W}$. Then, to solve the weighting vector is to solve the following optimization problem:

Model 1 $\min J_p = \boldsymbol{W}^{\mathrm{T}}(\boldsymbol{Y}_p^{\mathrm{T}}\boldsymbol{Y}_p)\boldsymbol{W}$

$$\text{s. t.}\begin{cases}\boldsymbol{E}^{\mathrm{T}}\boldsymbol{W} = 1\\ \boldsymbol{W} \geqslant 0\end{cases}$$

But on the other hand, we can transfer formula (2) into:

$$\sum_{j=1}^{m} w_j \left(\frac{g(\hat{Y}_j(t))}{f(\hat{Y}(t))}\right)^p = 1 \qquad t = 1, 2, \cdots, n \tag{9}$$

In the same way, if no errors are considered, there should ideally exist the following formula:

$$\sum_{j=1}^{m} w_j \left(\frac{g(\hat{Y}_j(t))}{f(Y(t))}\right)^p = 1 \qquad t = 1, 2, \cdots, n \tag{10}$$

But Eq.(10) still doesn't hold in general cases. We can introduce the following error item:

$$\eta_p(t) = \sum_{j=1}^{m} w_j \left(\frac{g(\hat{Y}_j(t))}{f(Y(t))}\right)^p - 1 \qquad t = 1, 2, \cdots, n \tag{11}$$

Denote $\boldsymbol{H}_p = \{\eta_p(1), \eta_p(2), \cdots, \eta_p(n)\}^{\mathrm{T}}$ and $\boldsymbol{G}_p = \left(\left(\frac{g(\hat{Y}_j(t))}{f(Y(t))}\right)^p - 1\right)_{n\times m}$, then, Eq.(11) can be written as $\boldsymbol{H}_p = \boldsymbol{G}_p\boldsymbol{W}$. In order to make the combining forecasting model optimal, we must choose the appropriate weighting coefficients to make $\min K_p = \sum_{t=1}^{n} \eta_p^2(t) = \boldsymbol{H}_p^{\mathrm{T}} \boldsymbol{H}_p = \boldsymbol{W}^{\mathrm{T}}(\boldsymbol{G}_p^{\mathrm{T}} \boldsymbol{G}_p)\boldsymbol{W}$. Then, to solve the weighting vector is to solve the following optimization problem:

Model 2 $\min K_p = \boldsymbol{W}^{\mathrm{T}}(\boldsymbol{G}_p^{\mathrm{T}} \boldsymbol{G}_p)\boldsymbol{W}$

$$\text{s.t.}\begin{cases}\boldsymbol{E}^{\mathrm{T}}\boldsymbol{W} = 1\\ \boldsymbol{W} \geqslant 0\end{cases}$$

Model 1 and model 2 are completely the same in solution means. So, we take model 2 as an example to study their solving procedures.

For model 2, it is clear that it is a quadratic programming problem. According to the theory of quadratic programming, when $\boldsymbol{G}_p^{\mathrm{T}} \boldsymbol{G}_p$ is a positive semidefinite matrix, there must exist an optimal solution. But it doesn't have a universal solution. For the above problem, if we don't consider the nonnegativity constraint condition, we can use the Lagrangian multiplier method to solve it as follows.

Denote $L = \boldsymbol{W}^{\mathrm{T}}(\boldsymbol{G}_p^{\mathrm{T}}\boldsymbol{G}_p)\boldsymbol{W} + \lambda(\boldsymbol{E}^{\mathrm{T}}\boldsymbol{W} - 1)$, where λ is a Lagrangian multiplier corresponding to the constraint condition $\boldsymbol{E}^{\mathrm{T}}\boldsymbol{W} = 1$. Suppose $\boldsymbol{G}_p^{\mathrm{T}} \boldsymbol{G}_p$ is a positive definite matrix and let $\frac{\partial L}{\partial \boldsymbol{W}} = 0$, we have $2\boldsymbol{G}_p^{\mathrm{T}}\boldsymbol{G}_p\boldsymbol{W} + \lambda\boldsymbol{E} = 0$. And with $\boldsymbol{E}^{\mathrm{T}}\boldsymbol{W} = 1$, we have

$$\boldsymbol{W} = \frac{(\boldsymbol{G}_p^{\mathrm{T}}\boldsymbol{G}_p)^{-1}\boldsymbol{E}}{\boldsymbol{E}^{\mathrm{T}}(\boldsymbol{G}_p^{\mathrm{T}}\boldsymbol{G}_p)^{-1}\boldsymbol{E}} \tag{12}$$

But this solution can yield the negativity weighting coefficients, which have no practical meanings. So, the nonnegativity constraint condition should be considered. The known solution methods of model 2 are the linear programming method[3], nonlinear programming method[4], dynamic programming method[5] and neural network method[6], etc. But they are all very complicated. In this paper, we use the algorithm of quadratic programming to solve it. Its Kuhn-Tucker conditions can be expressed as

$$\left.\begin{array}{l}(\mathrm{G}_p^{\mathrm{T}} \boldsymbol{G}_p)\boldsymbol{W} - \lambda\boldsymbol{E} - \boldsymbol{U} = 0\\ \boldsymbol{E}^{\mathrm{T}}\boldsymbol{W} = 1\\ u_i w_i = 0 \qquad i = 1, 2, \cdots, m\\ \boldsymbol{W},\ \boldsymbol{U} \geqslant 0\end{array}\right\} \tag{13}$$

where λ is a Lagrangian multiplier corresponding to the constraint condition $\boldsymbol{E}^{\mathrm{T}}\boldsymbol{W} = 1$, $\boldsymbol{U} = \{u_1, u_2, \cdots, u_m\}^{\mathrm{T}}$ is the Kuhn-Tucker multiplier corresponding to the weighting vector $\boldsymbol{W} = \{w_1, w_2, \cdots, w_m\}^{\mathrm{T}}$, and $\boldsymbol{E} = \{1, 1, \cdots, 1\}^{\mathrm{T}}$. In order to solve (13), we construct the aid linear programming. And since λ without the nonnegativity restriction, let $\lambda = \lambda' - \lambda''$ which satisfies $\lambda', \lambda'' \geqslant 0$ and $\lambda'\lambda'' = 0$. Now we can construct the following aid linear programming model:

$$\min J = v$$

$$\text{s. t.}\begin{cases}(\boldsymbol{G}_p^{\mathrm{T}} \boldsymbol{G}_p)\boldsymbol{W} - \lambda'\boldsymbol{E} + \lambda''\boldsymbol{E} - \boldsymbol{U} = 0\\ \boldsymbol{E}^{\mathrm{T}}\boldsymbol{W} + v = 1\\ \boldsymbol{W},\ \boldsymbol{U} \geqslant 0;\ \lambda',\ \lambda'',\ v \geqslant 0\end{cases} \tag{14}$$

It must be noticed that λ' and λ'' as well as u_i and w_i can't be basic variables simultaneously. By solving model (14), we can obtain the optimal combining weighting vector $\boldsymbol{W}^*$ easily. Generally speaking, the optimal combining weighting vector W^* can't be identical for different parameters p. But there must exist a group of weighting vectors which can produce the best combining forecasting effects.

3 Evaluation of Forecasting Effects

According to the evaluation conventions of forecasting effects, we choose the following indices as our evaluation criteria to evaluate and compare the two kinds of parameter estimation methods in the generalized weighted functional mean combining forecasting model:

① Sum of squares error SSE $= \sum_{t=1}^{n}(Y(t) - \hat{Y}(t))^2$, where $Y(t)$ and $\hat{Y}(t)$ represent the real values and the forecasting values, respectively; ② Mean absolute error MAE $= \frac{1}{n}\sum_{t=1}^{n}|Y(t) - \hat{Y}(t)|$; ③ Mean square error MSE $= \frac{1}{n}\sqrt{\sum_{t=1}^{n}(Y(t) - \hat{Y}(t))^2}$; ④ Mean absolute percent error MAPE $= \frac{1}{n}\sum_{t=1}^{n}\left|\frac{Y(t) - \hat{Y}(t)}{Y(t)}\right|$; ⑤ Mean square percent error MSPE $= \frac{1}{n}\sqrt{\sum_{t=1}^{n}\left(\frac{Y(t) - \hat{Y}(t)}{Y(t)}\right)^2}$.

4 Example Analysis

We take the example of air material consumption[7]. The consumption of the air material from 1992 to 1997 is listed in Tab.1 in detail. Its dominating forecasting models are[7] as follows.

Causality forecasting model: $\hat{Y}_1(t) = 17.73462 + 4.457063 \times 10^{-2}x_3$, where x_3 is times of flight landing which is one of the flight parameters[7].

Time series forecasting model: $\hat{Y}_2(t) = 156.3 - 0.247Y(t-1)$.

Tab.1 The actual consumption and the forecasting values of each model

Year	$Y(t)$	$\hat{Y}_1(t)$	$\hat{Y}_2(t)$	$\hat{Y}_{31}(t)$	$\hat{Y}_{32}(t)$	$\hat{Y}_{41}(t)$	$\hat{Y}_{42}(t)$	$\hat{Y}_{51}(t)$	$\hat{Y}_{52}(t)$
1992	119	118							
1993	131	123	127	125.34	124.67	125.33	124.68	125.33	124.61
1994	150	125	124	124.40	124.57	124.41	124.57	124.40	124.59
1995	101	124	120	121.58	122.26	121.59	122.24	121.58	122.35
1996	117	119	132	126.32	124.16	126.29	124.16	126.29	124.11
1997	130	137	128	131.46	133.00	131.50	132.95	131.45	133.23

In order to improve the forecasting precision, the effective consumption forecasting model should be the combination of these two kinds of forecasting models. By using the method of model 2 in this paper we can get the following optimal combining forecasting models.

Optimal generalized weighted arithmetic mean combining forecasting model ($p^* = -1.3$):

$$\hat{Y}_{32}(t) = [0.5744\hat{Y}_1^{-1.3}(t) + 0.4256\hat{Y}_2^{-1.3}(t)]^{-\frac{1}{1.3}}$$

Optimal generalized weighted logarithmic mean combining forecasting model ($p^* = -6.4$):

$$\hat{Y}_{42}(t) = \exp[0.5709(\ln\hat{Y}_1(t))^{-6.4} + 0.4291(\ln\hat{Y}_2(t))^{-6.4}]^{-\frac{1}{6.4}}$$

Optimal generalized weighted exponential mean combining forecasting model ($p^* = -0.00952$):

$$\hat{Y}_{52}(t) = \ln[0.5919(e^{\hat{Y}_1})^{-0.00952} + 0.4081(e^{\hat{Y}_2(t)})^{-0.00952}]^{-\frac{1}{0.00952}}$$

And, by using model 1, we can obtain the other optimal combining forecasting models as follows.

Optimal generalized weighted arithmetic mean combining forecasting model ($p^* = -1.5$):

$$\hat{Y}_{31}(t) = [0.4050\hat{Y}_1^{-1.5}(t) + 0.5950\hat{Y}_2^{-1.5}(t)]^{-\frac{1}{1.5}}$$

Optimal generalized weighted logarithmic mean combining forecasting model ($p^* = -5.9$):

$$\hat{Y}_{41}(t) = \exp[0.4081(\ln\hat{Y}_1(t))^{-5.9} + 0.5919(\ln\hat{Y}_2(t))^{-5.9}]^{-\frac{1}{5.9}}$$

Optimal generalized weighted exponential mean combining forecasting model ($p^* = -0.021$):

$$\hat{Y}_{51}(t) = \ln[0.4062(e^{\hat{Y}_1})^{-0.021} + 0.5938(e^{\hat{Y}_2(t)})^{-0.021}]^{-\frac{1}{0.021}}$$

We make a comparison among the above models. According to the formulae of these models, we get the forecasting values of these models listed in Tab.1 in detail. And the evaluations of the forecasting effects are described in Tab.2 in detail.

Tab.2 Evaluation results of forecasting effects

Indices of forecasting effects	$\hat{Y}_1(t)$	$\hat{Y}_2(t)$	$\hat{Y}_{31}(t)$	$\hat{Y}_{32}(t)$	$\hat{Y}_{41}(t)$	$\hat{Y}_{42}(t)$	$\hat{Y}_{51}(t)$	$\hat{Y}_{52}(t)$
SSE	1 271	1 282	1 199.895 2	1 198.919 2	1 199.963 4	1 198.031 9	1 199.321 0	1 203.239 9
MAE	13.0	13.2	12.524 7	12.635 3	12.530 0	12.622 2	12.517 0	12.698 1
MSE	7.130 2	7.161 0	6.927 9	6.925 1	6.928 1	6.922 5	6.926 2	6.937 5
MAPE	0.105 3	0.107 1	0.101 7	0.102 5	0.101 8	0.102 4	0.101 6	0.103 0
MSPE	0.058 84	0.057 63	0.056 21	0.056 45	0.056 22	0.056 42	0.056 20	0.056 59

It can be seen that the best effects of the generalized weighted arithmetic mean combining forecasting model and the generalized weighted logarithmic mean combining forecasting model can be obtained if and only if the parameter estimation model 2 is adopted. But for the generalized weighted exponential mean combining forecasting model, its best effect can be obtained if and only if the parameter estimation model 1 is adopted. Many other examples also support the above conclusion.

5 Conclusion

This paper presents a new kind of combining forecasting model and its parameter estimation methods of weighting coefficients. From the example of air material consumption forecasting, we can see that the generalized weighted functional mean combining forecasting model has extensive representation. It is a new kind of aggregative method of group forecasting. By taking the suitable combining form of the forecasting models and adopting a correct parameter estimation method, we can get the optimal combining form. Thereby, we can improve the forecasting precision and gain preferable forecasting results.

References

[1] Bates J M, Granger C W J. The combination of forecasts [J]. *Operations Research Quarterly,* **1969, 20**(4): 451－468.

[2] Wang Yingming, Fu Guowei. Research on the aggregative methods of group forecasting [J]. *Forecasting,* **1993, 12**(3): 42－45. (in Chinese)

[3] Zeng Yong, Tang Xiaowo. A study on the computing method of optimum combined prediction of non-negative weight [J]. *Statistical Research,* 1994(3): 70－74. (in Chinese)

[4] Li Xuequan. Research on nonnegative weighting coefficient calculating method of optimal combining predictor [J]. *Forecasting,* **1995, 14**(4): 49－50, 60. (in Chinese)

[5] Yang Guiyuan, Tang Xiaowo. Integration forecasting model with non-negative weights [J]. *Quantitative & Technical Economics,* 1998(3): 56－60. (in Chinese)

[6] Zhang Qing. Application research on an optimal mix forecasting method based on ANN [J]. *Systems Engineering — Theory & Practice,* **2001, 21**(9): 90－93. (in Chinese)

[7] Wan Yucheng. Generalized weighted mean combining forecasting and its application in the forecasting of air materials consumption [J]. *Journal of Systems Science and Systems Engineering,* **1999, 8**(4): 445－450.

广义加权函数平均组合预测模型参数估计方法

万玉成[1,2]　　盛昭瀚[3]

([1] 东南大学经济管理学院, 南京 210096)
([2] 空军后勤学院三系, 徐州 221002)
([3] 南京大学工程管理学院, 南京 210093)

摘要: 提出一种新的组合预测模型——广义加权函数平均组合预测模型, 并利用二次规划算法给出其加权系数的 2 种不同的参数估计方法.最后, 以预测实例说明了模型的有效性并对 2 种参数估计方法进行了比较, 得到一个对参数估计方法进行正确应用的有用结论.

关键词: 组合预测; 广义加权函数平均; 参数估计; 二次规划

中图分类号: F224