调和平均组合预测方法的进一步研究

王应明

厦门大学自动化系, 361005

摘 要 对调和平均组合预测两种不同的组合权系数估计方法,通过大量的仿真预测实例进行了对比分析,得到了明确的结论,为调和平均组合预测方法的正确应用奠定了基础。

主题词 预测,参数估计,+ 调和平均数。

Evaluation and Analysis of Harmonic-Means Combining Forecasting Effects

Wang Yingming

Dept. of Automation, Xiamen University, 361005

Abstract: Two kinds of weight coefficients estimation methods (I) and (II) and their combining forecasting effects based on harmonic-means are compared and analysed through a lots of simulation forecasting examples. A definite conclusion is obtained, which can lay solid foundations for correct application on above combining forecasting methods.

Keywords: Harmonic-means, Combining forecasts, Parameter estimation.

1 引 言

调和平均组合预测是一种比较常用的组合预测方法,在某些情况下,可以比加权算术平均和加权几何平均组合预测取得更好的组合预测效果。但是,如果使用不恰当的参数估计方法,则不但不能取得比较好的组合预测效果,而且还会影响调和平均组合预测方法的有效性。鉴于此,本文对调和平均组合预测权系数的估计方法进行实证分析,以便进行科学选择。

2 模型及参数估计

设某一预测问题在某一时段的实际值为 y_t ($t=1,2,\cdots,n$),对此预测问题共有 m 种可行的预测方法,其预测值或模型拟合值分别为 f_{tj} ($t=1,2,\cdots,n$; $j=1,2,\cdots,m$)。又设 m 种预测方法的加权向量为 $W=(W_1,W_2,\cdots,W_m)^T$,且满足归一化约束条件

$$e^T W = 1 \tag{1}$$

其中 $e^T=(1,1,\cdots,1)$。为了使组合预测模型能从经济意义上作出科学而合理的解释,加权向量 W 还必须满足非负性约束条件

$$W \geqslant 0 \tag{2}$$

根据加权调和平均数计算公式[1],令

$$\hat{y}_t = \frac{\sum_{j=1}^{m} W_j}{\sum_{j=1}^{m} \frac{W_j}{f_{tj}}} = \frac{1}{\sum_{j=1}^{m} \frac{W_j}{f_{tj}}} \quad t=1,2,\cdots,n \tag{3}$$

其中 $\hat{y}_t$ 为 m 种可行预测方法在 t 时刻的加权调和平均值,很显然,$\hat{y}_t$ 是 y_t 的一个很好的逼近,因此,(3)式即为本文给出的加权调和平均组合预测模型公式。为便于对模型(3)进行组合权系数的估计,可将(3)式转变为

$$\sum_{j=1}^{m} \frac{W_j}{f_{tj}} = \frac{1}{\hat{y}_t} \quad t=1,2,\cdots,n \tag{4}$$

或

$$\sum_{j=1}^{m} \frac{\hat{y}_t}{f_{tj}} W_j = 1 \quad t=1,2,\cdots,n \tag{5}$$

假设不考虑预测误差的存在,在理想的情况下,应有下式成立

收稿日期: 1997年 7月 30日

$$\sum_{j=1}^{m}\frac{W_j}{f_{tj}}=\frac{1}{y_t}\quad t=1,2,\cdots,n \tag{6}$$

$$\sum_{j=1}^{m}\frac{y_t}{f_{tj}}W_j=1\quad t=1,2,\cdots,n \tag{7}$$

但实际上,由于预测误差的客观存在性和不可避免性,(6)和(7)式在通常情况下是不成立的,为此引入误差项

$$X_t=\sum_{j=1}^{m}\frac{W_j}{f_{tj}}-\frac{1}{y_t}\quad t=1,2,\cdots,n \tag{8}$$

$$Z_t=\sum_{j=1}^{m}\frac{y_t}{f_{tj}}W_j-1\quad t=1,2,\cdots,n \tag{9}$$

人们自然总是希望误差愈小愈好,为此,定义误差性能指标为

$$\min J=\sum_{t=1}^{n}X_t^2 \tag{10}$$

或

$$\min\tilde{J}=\sum_{t=1}^{n}Z_t^2 \tag{11}$$

由此可以得到两种不同的组合权系数估计方法,即

$$\min J=\sum_{t=1}^{n}X_t^2 \tag{12}$$

方法(I):

$$\text{s. t.}\begin{cases}\sum_{j=1}^{m}\frac{W_j}{f_{tj}}-\frac{1}{y_t}=X_t\quad t=1,2,\cdots,n & (13)\\ e^TW=1 & (14)\\ W\geqslant 0 & (15)\end{cases}$$

$$\min\tilde{J}=\sum_{t=1}^{n}Z_t^2 \tag{16}$$

方法(II):

$$\text{s. t.}\begin{cases}\sum_{j=1}^{m}\frac{y_t}{f_{tj}}W_j-1=Z_t\quad t=1,2,\cdots,n & (17)\\ e^TW=1 & (18)\\ W\geqslant 0 & (19)\end{cases}$$

方法(I)和方法(II)在解法上完全一样,现以方法(II)为例研究其统一解法。令

$$\tilde{f}_{tj}=y_t/f_{tj}\quad t=1,2,\cdots,n \tag{20}$$

则(17)式可写成向量形式

$$\tilde{F}W-e=H \tag{21}$$

其中 $\tilde{F}=(\tilde{f}_{tj})_{n\times m}$,$e=(1,1,\cdots,1)^T$,$H=(Z_1,Z_2,\cdots,Z_n)^T$,于是,方法(II)可用向量形式表示为

$$\min\tilde{J}=\sum_{t=1}^{n}Z_t^2=H^TH$$

$$=(\tilde{F}W-e)^T(\tilde{F}W-e) \tag{22}$$

$$\text{s. t.}\begin{cases}e^TW=1 & (23)\\ W\geqslant 0 & (24)\end{cases}$$

显然,这是一个二次规划问题,在 $W>0$的情况下,可用最小二乘法求解,但这种解法不具有一般性,根据二次规划理论可知,上述二次规划问题的最优解一定存在,且其 Kuhn-Tucker条件可表示为

$$\begin{cases}\tilde{F}^T(\tilde{F}W-e)-\lambda e-\tilde{U}=0 & (25)\\ e^TW=1 & (26)\\ \tilde{u}_iW_i=0\quad i=1,2,\cdots,m & (27)\\ W,\tilde{U}\geqslant 0 & (28)\end{cases}$$

式中 λ 为与约束条件 $e^TW=1$ 相对应的 Lagrange乘子,$\tilde{U}=(\tilde{u}_1,\tilde{u}_2,\cdots,\tilde{u}_m)^T$ 为与组合权向量 $W=(W_1,W_2,\cdots,W_m)^T$ 相对应的 Kuhn-Tucker乘子。由于 λ 无非负约束,故可令 $\lambda=\lambda'-\lambda''$且满足 $\lambda',\lambda''\geqslant 0$,$\lambda'\cdot\lambda''=0$,同时构造辅助线性规划模型(ALP)为

$$\min\tilde{J}'=v \tag{29}$$

$$\text{s. t.}\begin{cases}(\tilde{F}^T\tilde{F})W-\lambda'e+\lambda''e-\tilde{U}=\tilde{F}^Te & (30)\\ e^TW+v=1 & (31)\\ W,\tilde{U}\geqslant 0;\lambda',\lambda'',v\geqslant 0 & (32)\\ \lambda'\text{与}\lambda''\text{及}\tilde{u}_i\text{与}W_i\text{不能同时为基变量}\ (i=1,2,\cdots,m) & (33)\end{cases}$$

解此辅助线性规划模型(ALP),即可得到加权调和平均组合预测模型的最优组合加权向量 W^*。对于参数估计方法(I),同样也可得到一组最优组合加权向量 W^*,一般而言,这两组加权向量是不一样的,但其中必有一组能使调和平均组合预测效果最佳。

3　预测效果评价

为了评价和比较两种不同参数估计方法的需要,必须制定一套可行的评价指标体系,对调和平均组合预测效果进行全方位的综合性评价。按照预测效果评价原则和惯例,本文选择下列指标作为评判准则。

(1)平方和误差

$$\mathrm{SSE}=\sum_{t=1}^{n}(y_t-\hat{y}_t)^2 \tag{34}$$

式中 y_t为预测事物实际观测值,$\hat{y}_t$为预测值。

(2)平均绝对误差

$$\mathrm{MAE}=\frac{1}{n}\sum_{t=1}^{n}|y_t-\hat{y}_t| \tag{35}$$

(3)均方误差

$$\mathrm{MSE}=\frac{1}{n}\sqrt{\sum_{t=1}^{n}(y_t-\hat{y}_t)^2} \tag{36}$$

(4)平均绝对百分比误差

$$\mathrm{MAPE}=\frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t-\hat{y}_t}{y_t}\right| \tag{37}$$

(5)均方百分比误差

$$\mathrm{MSPE}=\frac{1}{n}\sqrt{\sum_{t=1}^{n}\left(\frac{y_t-\hat{y}_t}{y_t}\right)^2} \tag{38}$$

4 应用举例

本文以文献[2~4]所给预测实例为例进行调和平均组合预测效果的评价分析。已知例1例2和例3的原始数据分别取自于文献[2~4],如表1所示。调和平均组合预测方法的效果评价如表2所示。为了便于比较分析,表2中还同时给出了各单个预测方法及其他常用组合预测方法的效果评价。

表 1 预测实例原始数据简表

t		1	2	3	4	5	6	7	8	9	10	11	12
例 1	y_t	14.9	18.6	22.2	17.6	19.6	24	31.6	43.7	37	47.2		
	f_{t1}	10	14.9	23.3	26.1	17.5	20.2	26.4	36.8	52.5	38.5		
	f_{t2}	12	15.48	18.95	22.43	25.9	29.38	32.85	36.33	39.8	43.28		
例 2	y_t	57.0	65.4	75.4	82.5	92.8	102.7	119.5	143.8	169.7	201.0	251.2	
	f_{t1}	54.52	62.89	72.54	83.67	96.51	111.32	128.41	148.11	170.84	197.06	227.31	
	f_{t2}	64.68	66.74	68.72	76.61	88.42	104.15	123.79	147.35	174.82	206.21	241.51	
例 3	y_t	11.49	13.06	15.34	20.58	23.28	26.46	27.33	34.22	40.19	53.37	77.79	100.63
	f_{t1}	18.47	14.54	12.84	13.28	16.15	21.16	28.40	37.87	49.58	63.53	79.00	98.12
	f_{t2}	10.03	11.23	15.24	18.67	27.78	26.36	29.67	27.40	42.73	47.36	71.00	109.32

表 2 预测效果评价简表

预测效果评价指标					SSE	MAE	MSE	MAPE	MSPE
例 1	个体预测			方法(I)	520.60	6.04	2.28	0.2251	0.0825
				方法(II)	199.76	4.11	1.41	0.1696	0.0599
	组合预测	加权算术平均		$W_1=0.1158$ $W_2=0.8842$	194.16	4.05	1.39	0.1649	0.0579
		加权几何平均		$W_1=0.2159$ $W_2=0.7841$	191.35	3.97	1.38	0.1590	0.0561
		加权调和平均	方法(I)	$W_1=0.0393$ $W_2=0.9607$	192.71	4.05	1.39	0.1663	0.0584
			方法(II)	$W_1=0.1634$ $W_2=0.8366$	183.33	3.88	1.35	0.1570	0.0555
例 2	个体预测			方法(I)	795.59	5.78	2.56	0.0440	0.0156
				方法(II)	338.25	4.96	1.67	0.0474	0.0179
	组合预测	加权算术平均		$W_1=0.1259$ $W_2=0.8741$	328.56	4.80	1.65	0.0443	0.0159
		加权几何平均		$W_1=0.5652$ $W_2=0.4348$	449.04	4.28	1.92	0.0334	0.0122
		加权调和平均	方法(I)	$W_1=0.7017$ $W_2=0.2983$	536.54	4.53	2.11	0.0332	0.0126
			方法(II)	$W_1=0.5596$ $W_2=0.4404$	448.64	4.26	1.93	0.0332	0.0122

续表 2 预测效果评价简表

<table>
<tr><th colspan="5">预测效果评价指标</th><th>SSE</th><th>MAE</th><th>MSE</th><th>MAPE</th><th>MSPE</th></tr>
<tr><td rowspan="6">例 3</td><td colspan="3" rowspan="2">个体预测</td><td>方法(I)</td><td>401.56</td><td>4.88</td><td>1.67</td><td>0.1959</td><td>0.0731</td></tr>
<tr><td>方法(II)</td><td>245.58</td><td>3.59</td><td>1.30</td><td>0.0998</td><td>0.0334</td></tr>
<tr><td rowspan="4">组合预测</td><td colspan="2">加权算术平均</td><td>W_1= 0.4121
W_2= 0.5879</td><td>94.89</td><td>2.34</td><td>0.81</td><td>0.0791</td><td>0.0283</td></tr>
<tr><td colspan="2">加权几何平均</td><td>W_1= 0.2617
W_2= 0.7383</td><td>117.86</td><td>2.60</td><td>0.90</td><td>0.0737</td><td>0.0242</td></tr>
<tr><td rowspan="2">加权调和平均</td><td>方法(I)</td><td>W_1= 0.2473
W_2= 0.7527</td><td>126.06</td><td>2.67</td><td>0.94</td><td>0.0742</td><td>0.0250</td></tr>
<tr><td>方法(II)</td><td>W_1= 0.2481
W_2= 0.7519</td><td>125.83</td><td>2.66</td><td>0.93</td><td>0.0742</td><td>0.0250</td></tr>
</table>

例 1的组合预测效果评价表明,加权调和平均组合预测方法是最佳的组合预测方法,但是,如果使用参数估计方法(I),则不能取得此最佳效果。因此,最佳的调和平均组合预测效果必须以参数估计方法(II)为前提。例 2和例 3的组合预测效果评价表明,加权调和平均组合预测方法虽然不能取得最佳的组合预测效果,但是,使用参数估计方法(II)比使用参数估计方法(I)优越,参数估计方法(II)能使各项组合预测效果评价指标得到不同程度的改善。此外,笔者还对其他大量的组合预测实例进行了验证,得到了与此相同的结论。

综合以上评价分析可以看出,加权算术平均组合预测、加权几何平均组合预测和加权调和平均组合预测这三种常用的组合预测方法各有千秋,很难比较谁优谁劣,每种方法都有其存在的价值和优越性,但是,作为调和平均组合预测方法,由于存在不同的组合权系数估计方法,采用不同的组合权系数估计方法往往会取得不同的组合预测效果,大量的仿真预测实例表明,参数估计方法(II)比参数估计方法(I)优越,因此,在采用调和平均数进行组合预测的时候,应广泛采用参数估计方法(II)进行组合权系数的估计,这样才能取得比较好的组合预测效果。

参 考 文 献

1 于涛,栗方忠.社会经济统计学原理.武汉大学出版社,1992

2 周传世,罗国民.加权几何平均组合预测模型及其应用.数理统计与管理,1995(2):17~19

3 杨桂元,唐小我,马永开.最优加权几何平均组合预测方法研究.统计研究,1996(2):55~58

4 王应明,傅国伟.基于不同误差准则和范数的组合预测方法研究.控制与决策,1994(1):20~28