基于“序列相对贴近度”的组合预测权值分配

吕永乐 郎荣玲 谈展中

(北京航空航天大学 电子信息工程学院,北京 100191)

摘 要: 在时间序列的组合预测权值分配问题上,为克服传统的均方误差倒数加权、熵权和最优化方法之不足,从预测值序列与评价样本序列间的贴近性出发,提出新方法综合衡量单一参与模型的适用性,并据此分配权值.详细给出了序列相对贴近度(SRND, Sequence Relative Nearness Degree)及与之相关的“序列趋势关联度”和“尺度区间熵”的概念,并提出基于SRND的权值分配方法.将SRND权值分配方法应用于航空发动机排气温度裕度参数时间序列的联合自回归滑动平均模型、函数系数自回归模型和径向基函数网络预测,有效地提高了预测准确度,获得优于均方误差倒数加权和熵权方法的组合性能,且运算量远小于最优化方法.

关 键 词: 时间序列分析; 组合预测; 建模; 性能

中图分类号: O 23; TP 18

文献标识码: A **文 章 编 号**: 1001-5965(2009)12-1434-04

Weight allocation of combination prediction based on sequence relative nearness degree

Lü Yongle Lang Rongling Tan Zhanzhong

(School of Electronics and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China)

Abstract: Aiming at the weight allocation problems of combination prediction for a time series, a new method was proposed to evaluate the applicability of the employed models and allocate weights, based on the “nearness” between the test sequence and the corresponding prediction value sequence, which overcame the shortages of existing methods such as mean square error reciprocal weight (1/MSE), entropy weight and optimization weight. The definitions of sequence relative nearness degree (SRND), related sequence trend association and scale interval entropy were given and well discussed, as well as the weight allocation expressions based on SRND. By the example which combined the autoregressive moving average model, functional-coefficient autoregressive model and radial basis function prediction networks in the prediction analysis for the take-off exhaust gas temperature margin time series, the conclusion is drawn that the prediction accuracy can be effectively improved with the proposed method, compared to 1/MSE and entropy weight methods, while the calculation mount is far lower than optimization weight method.

Key words: time series analysis; combination prediction; model buildings; performance

自从Bates J. M.和Granger C. W. J.首次提出并证明[1]两种单项预测可以组合出更优的预测结果以来,组合预测逐渐成为该领域中的一个重要研究方向.对于线性加权组合预测而言,如何合理分析、评价参与组合的单一模型的适用性,并由此确定权值,使得组合模型能够更有效地提高预测准确度是关键.目前常用的权值分配方法主要有简单线性加权、均方误差倒数加权、熵权[2]和最优化方法[3]等.

1 时间序列组合预测

时间序列[4]是按照时间顺序、常以等时间间

收稿日期: 2008-11-12

基金项目: 国防十一五预研项目资助(102010201)

作者简介: 吕永乐(1981-),男,河南洛阳人,博士生, lv_yongle@ee.buaa.edu.cn.

隔取得的一系列观测值的集合. 对于单变量时间序列 $\{S_t,\ t\in\mathbf{N}\}$,设存在 K个预测模型; $O_t^i\ (i=1,2,\cdots,K)$表示各单一模型输出的 t时刻的预测值; 经线性加权后所得组合模型的预测输出为

$$Q_t^{\mathrm{comb}}=\sum_{i=1}^{K}\lambda_i O_t^i \tag{1}$$

式中,权值 $\lambda_i>0$且 $\sum_{i=1}^{K}\lambda_i=1$.

在已有的权值分配方法中,简单线性加权方法,如平均权法,无法客观评价各单一模型的适用度差异,组合效果不佳.

均方误差倒数加权方法是目前普遍采用的一种组合方法. 虽然理论上证明[1]了:对于几个相互独立的无偏预测(预测误差的期望等于 0),该方法能够有效提高预测准确度. 但实际应用中,"独立的无偏预测"条件较为苛刻,难以达到理想的模型适用性评价效果.

熵权法本质上是依据预测误差的概率分布来分配权值:分布越集中的模型,适用度越高,相应权值越大. 但此方法[2]单独采用所存在的主要不足是没有考虑不同模型预测误差均值水平的差异.

最优化方法是将权值 $\lambda_1,\lambda_2,\cdots,\lambda_K$作为 $\mathbf{R}^+$空间中的一组变量,通过求解最优化问题来获得最佳权值分配;但该方法计算量较大,尤其当 K较大时,易陷入"局部极小值"问题.

针对上述权值分配方法之不足,研究试图从预测值序列与评价样本序列间的"贴近性"出发,评价单一模型的适用性,为组合预测权值分配寻求一种更切实际的数值依据,且保持低运算量.

2　基于序列相对贴近度的权值分配

2.1　序列相对贴近度

定义"贴近性"为两时间序列在任一时刻变化趋势的相关性和取值状态的一致性. 为度量"贴近性",提出了序列相对贴近度 SRND(Sequence Relative Nearness Degree)及与之相关的序列趋势关联度和 Δ尺度区间熵的概念.

定义 1(序列趋势关联度)　预测值序列 $\{O_1,O_2,\cdots,O_n\}$与评价样本序列 $\{S_1,S_2,\cdots,S_n\}$间的趋势关联度定义为

$$\alpha_{OS}=\frac{\sum_{t=1}^{n-1}\left[(O_{t+1}-O_t)(S_{t+1}-S_t)\right]}{\sqrt{\sum_{t=1}^{n-1}(O_{t+1}-O_t)^2}\cdot\sqrt{\sum_{t=1}^{n-1}(S_{t+1}-S_t)^2}} \tag{2}$$

序列趋势关联度刻画了预测值序列与评价样本序列一步变化趋势的相关性.

定义 2(Δ尺度区间熵)　记预测误差 $\varepsilon_t=O_t-S_t,\ t=1,2,\cdots,n$;设误差序列 $\{\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\}$的一维概率密度函数为 $f(\varepsilon),\ \varepsilon\in(-\infty,+\infty)$; 在区间 $[a,b]$上定义 Δ尺度区间熵:

$$\left.\begin{aligned}&\beta_\varepsilon(a,b,\Delta)=-\sum_{i=1}^{N}p_i\ln p_i\\&0\cdot\ln 0\equiv 0\end{aligned}\right\} \tag{3}$$

其中

$$p_i=\frac{\int_{a+(i-1)\Delta}^{\min(b,\,a+i\Delta)}f(\varepsilon)\,\mathrm{d}\varepsilon}{\int_a^b f(\varepsilon)\,\mathrm{d}\varepsilon} \tag{4}$$

且 $\sum_{i=1}^{N}p_i=1$, $N=\mathrm{Int}\left[\frac{b-a}{\Delta}\right]+1$, Int表示取整数.

Δ尺度区间熵描述了给定区间内(图 1所示),预测误差 ε_t在 Δ尺度下取值状态的不确定性. 可作为 ε_t在区间 $[a,b]$内取值分布离散程度的一种度量.

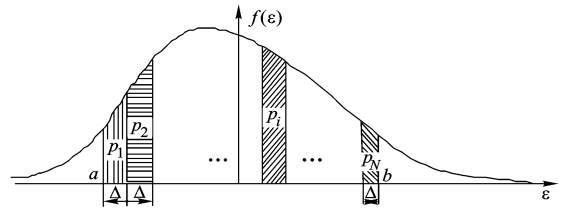


图 1　概率密度 $f(\varepsilon)$的 Δ尺度区间划分

定义 3(序列相对贴近度)　预测值序列 $\{O_1,O_2,\cdots,O_n\}$相对于评价样本序列 $\{S_1,S_2,\cdots,S_n\}$的相对贴近度定义为

$$\zeta_{O/S}=\frac{e^{\alpha_{OS}-1}}{\dfrac{\beta_\varepsilon(m_\varepsilon-3\sigma_\varepsilon,\,m_\varepsilon+3\sigma_\varepsilon,\,\sigma_\varepsilon)}{1.4426}\cdot\sigma_\varepsilon^2+m_\varepsilon^2}\cdot\sigma_S^2 \tag{5}$$

式中,$m_\varepsilon,\sigma_\varepsilon^2$分别表示预测误差序列的均值和方差. $\beta_\varepsilon(\cdot,\cdot,\cdot)$项是 ε_t在区间 $[m_\varepsilon-3\sigma_\varepsilon,\,m_\varepsilon+3\sigma_\varepsilon]$上的 $\Delta=\sigma_\varepsilon$尺度区间熵;常数 1.4426是假定 ε_t服从正态分布时 $\beta_\varepsilon(\cdot,\cdot,\cdot)$的取值; $\beta_\varepsilon(\cdot,\cdot,\cdot)/1.4426$表示对正态分布进行归一化,可看作 σ_ε^2项的奖惩因子. σ_S^2是 $\{S_1,S_2,\cdots,S_n\}$的方差,放在定义式中既体现了"相对"含义,又保证了 SRND的无量纲性.

2.2　预测模型适用性的 SRND评价

实际应用中,由于对象序列的非平稳性、单一建模方法自身的局限性、模型参数估计偏差、建模

时样本数据含有观测噪声等因素存在,针对评价样本的预测值往往不能满足"无偏性"要求,且预测误差序列通常也不符合正态白噪声假定.因此,在区分、评价单一预测模型的适用性时,需要综合考虑误差序列的一维概率密度和自相关性.

SRND采用均值、方差、尺度区间熵3个指标来刻画误差序列的一维概率分布;采用预测值序列与评价样本序列间的趋势关联度来衡量误差序列的自相关性.用SRND评价预测模型的适用性,反映在数值上就是:SRND取值越大,模型越适用.

2.3 SRND权值分配

在基于SRND的权值分配中

$$\lambda_i = f(\zeta_{O_i/S}) \Big/ \sum_{i=1}^{K} f(\zeta_{O_i/S}) \tag{6}$$

函数 $f: \mathbf{R}^+ \to \mathbf{R}^+$ 单调递增,根据应用需要,可采用指数形式、对数形式和多项式形式等.文中采用最简单的函数形式

$$f(x) = x \tag{7}$$

3 应用实例

3.1 EGTM参数联合预测

图2显示的是某航空公司A300飞机装配的CF6-80C2型航空发动机从2004年5月29日至2006年5月4日连续600次飞行的起飞排气温度裕度(EGTM, Exhaust Gas Temperature Margin)参数时间序列($\{S_t,\ t=1, 2, \cdots, 600\}$,已经过预处理,单位:℃).将前400个数据作为训练样本,用于ARMA(Autoregressive Moving Average model),FAR(Functional-coefficient Auto Regressive model)模型参数估计和RBF预测网络(RBFPN, Radial Basis Function Prediction Networks)训练[5-7].后200个数据作为评价样本,用于单一模型适用性评价及组合权值确定.

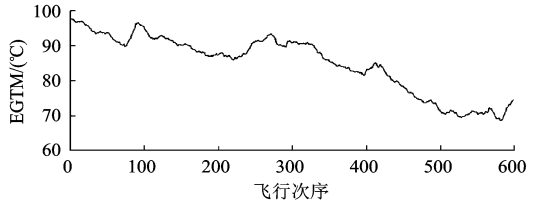


图2 某航空发动机EGTM参数监测数据序列

所建立ARMA(2,2)和FAR(3,1)预测模型的表达式分别为

$$O_{t+1}^{\text{ama}} = 1.869(S_t - 84.89) - 0.873(S_{t-1} - 84.89) - 0.828a_t + 0.025a_{t-1} + 84.89 \tag{8}$$

$$O_{t+1}^{\text{far}} = \tilde{f}_1(u)(S_t - 84.89) + \tilde{f}_1(u)(S_{t-1} - 84.89) + \tilde{f}_3(u)(S_{t-2} - 84.89) + 84.89 \tag{9}$$

其中,84.89是$\{S_t,\ t=1, 2, \cdots, 600\}$的均值;$O_{t+1}^{\text{ama}}$,$O_{t+1}^{\text{far}}$分别表示$t$时刻基于ARMA,FAR模型的1步预测值,对应于$S_{t+1}$.$\{a_t\}$是均值为0,方差为0.0749的独立同正态分布随机序列.$u = S_t - 84.89$,$\tilde{f}_i(u)$ $(i=1, 2, 3)$是对FAR(3,1)模型系数函数局部线性回归估计[6]的低阶多项式拟合,如图3所示.

$$\left.\begin{aligned}\tilde{f}_1(u) &= -0.0004u^3 + 0.0080u^2 - 0.0453u + 1.2243\\ \tilde{f}_2(u) &= 0.0002u^3 - 0.0100u^2 + 0.1235u - 0.3319\\ \tilde{f}_3(u) &= 0.0002u^3 + 0.0020u^2 - 0.0785u + 0.1075\end{aligned}\right\} \tag{10}$$

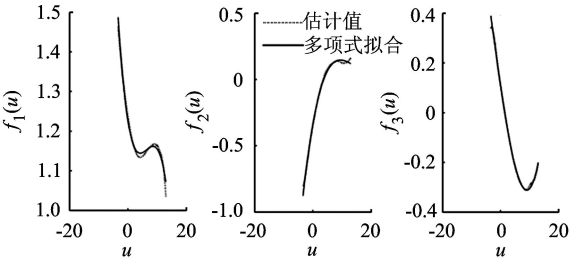


图3 FAR(3,1)模型的系数函数

RBFPN结构如图4所示,隐节点作用函数为

$$\psi_j(S_t) = \exp\left(-\frac{1}{2b_j^2}\|S_t - c_j\|^2\right) \quad j = 1, 2, \cdots, 23 \tag{11}$$

c_j,b_j和连接权ω_j由二阶段学习算法得到(略).

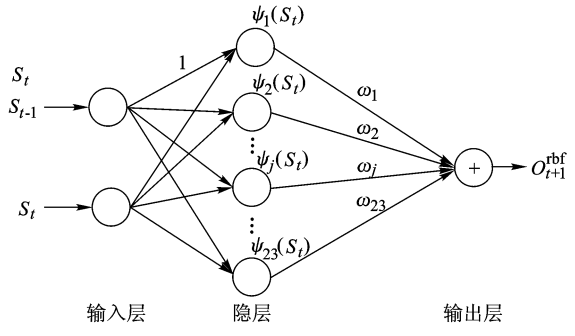


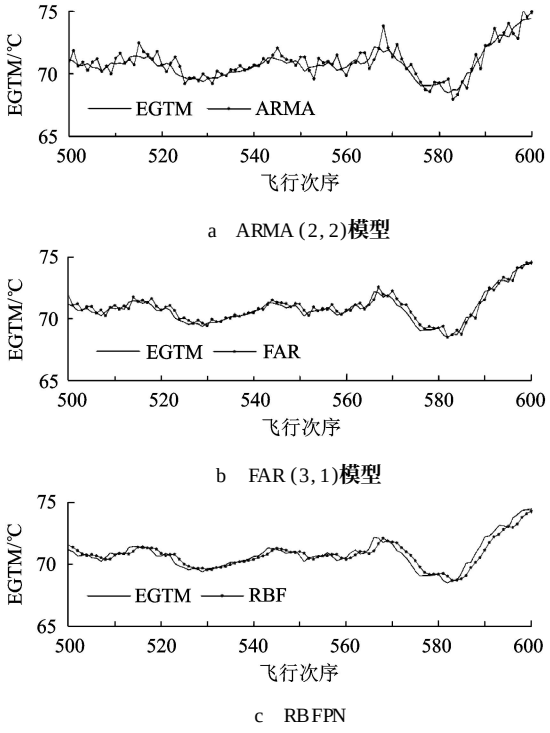
图4 RBFPN结构

ARMA(2,2),FAR(3,1)和RBFPN针对评价样本序列的一步预测效果如图5所示.预测值序列相对于评价样本序列的SRND分别为$\zeta_{O_{\text{ama}}/S} = 27.45$,$\zeta_{O_{\text{far}}/S} = 102.33$,$\zeta_{O_{\text{rbf}}/S} = 90.86$.于是由式

(6)和式(7)得到

$$\lambda_{arma} = 0.1244 \quad \lambda_{far} = 0.4638$$

$$\lambda_{rbf} = 0.4118$$



a　ARMA(2,2)模型

b　FAR(3,1)模型

c　RBFPN

图 5　评价样本序列(部分)的一步预测值

3.2　SRND与其它权值分配方法的性能比较

采用预测准确度指标(MSE,Mean Square Error)来评价组合性能.表 1列出 SRND权、均方误差倒数权、熵权、最优化等方法的权值分配结果,以及一步组合预测 MSE和计算时间.计算时间是在具有 Inter(R) Core(TM)2处理器(2.40 GHz)和 1.99 GB内存的计算机上(OS:WinXp),在 Matlab7.0集成环境中编写程序测得的.最优化方法的权值求解采用 QP(Quadratic Programming)算法实现.

表 1　不同权值分配方法组合性能比较

权值分配方法	λ_{arma}	λ_{far}	λ_{rbf}	MSE/℃²	计算时间/ms
SRND	0.1244	0.4638	0.4118	0.1029	1.593
1/MSE	0.1591	0.4457	0.3952	0.1052	0.063
Entr	0.3253	0.3521	0.3226	0.1251	1.296
最优化	0.0719	0.4805	0.4476	0.1018	171.921

本实例中,"SRND"一步组合预测的 MSE最接近最优化方法,计算时间却不足其 1%.虽然 SRND权值分配的计算时间比"1/MSE"、"Entr"方法长,但为了获得更高的组合性能,一般应用中还是可以接受的.

表 2给出各单一、组合模型的 1,3,10步预测 MSE;多步预测采用迭代方法实现.对比发现,在"SRND","1/MSE","Entr"三者之中,研究提出的 SRND权值分配方法的组合性能最佳,"1/MSE"次之,"Entr"最差(失效).

表 2　各单一、组合模型的预测 MSE

预测步数	MSE/10⁻²℃²						
	ARMA	FAR	RBFPN	SRND	1/MSE	Entr	最优化
1	33.40	11.93	13.45	10.29	10.52	12.15	10.18
3	64.89	50.83	32.10	29.66	30.35	32.28	28.81
10	457.46	297.02	50.28	38.81	39.25	129.79	38.76

4　结束语

传统的均方误差倒数加权、熵权、最优化等权值分配方法存在着应用条件苛刻、评价指标单一或运算量大等不足.从预测值序列与评价样本序列间的贴近性出发,提出了"序列相对贴近度"概念,评价参与组合各单一模型的适用性,并为组合预测权值分配提供一种更切实际和全面的数值依据.函数 $f(\cdot)$对权重分配有直接影响;为了获得更好的组合性能,需进一步探索其他有价值的函数形式,将作为后续研究的重点.

参考文献 (References)

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

[2] 陈华友. 熵值法及其在确定组合预测权系数中的应用[J]. 安徽大学学报:自然科学版, 2003, 27(4): 1 - 6
Chen Huayou. Entropy method and application to determine weights of combination forecasting[J]. Journal of Anhui Univ Natural Science Edition, 2003, 27(4): 1 - 6(in Chinese)

[3] 马永开, 唐小我. 线性组合预测模型优化问题研究[J]. 系统工程理论与实践, 1998(9): 110 - 114
Ma Yongkai, Tang Xiaowo. Research on the problem of optimizing linear combination forecasting model[J]. Systems Engineering-Theory & Practice, 1998(9): 110 - 114(in Chinese)

[4] Box G P E, Jenkins G M, Reinsel G C. Time series analysis: forecasting and control[M]. 3rd ed. New Jersey: Prentice Hall/Pearson Education Asia Ltd, 2005

[5] 李晓白, 崔秀伶, 郎荣玲. 航空发动机性能参数预测方法[J]. 北京航空航天大学学报, 2008, 34(3): 253 - 256
Li Xiaobai, Cui Xiuling, Lang Rongling. Forecasting method for aeroengine performance parameters[J]. Journal of Beijing University of Aeronautics and Astronautics, 2008, 34(3): 253 - 256(in Chinese)

[6] Fan J Q, Yao Q W. Nonlinear time series: nonparametric and parametric methods[M]. Berlin: Springer, 2005

[7] Feng Liu, Chai Quek, Geok S N. Neural network model for time series prediction by reinforcement learning[C]// IJCNN2005. Hilton Montreal Bonaventure, Montreal, Canada: IEEE, 2005: 809 - 814

(编　辑:娄　嘉)