

第一作者简介:秦伟良,男,1965年9月生,研究生,讲师

在 $L^2(R^n)$ 也是稠密的。所以, 对于形如 (1) 的函数可用 Φ_i 中元素的线性组合逼近。

基于以上讨论, 我们构造如下形式的神经网络来逼近 (1) 中的函数

$$G(X) = \sum_{i=1}^N w_i \Psi[\mathbf{D}_i \mathbf{R}_i (X - \mathbf{t}_i)] + \bar{g} \tag{5}$$

其中参数 $\bar{g}$ 是为了逼近非零均值的函数而设置的, 伸缩矩阵 $\mathbf{D}_i$ 是对角阵, $\mathbf{R}_i$ 是旋转矩阵, $\mathbf{t}_i$ 是平移因子组成的向量。

$$\mathbf{D}_i = \text{diag}(\frac{1}{S_i^1}, \cdots, \frac{1}{S_i^n})$$

此网络可用一个具有输入层、隐含层和输出层的神经网络来描述如图 1 所示。

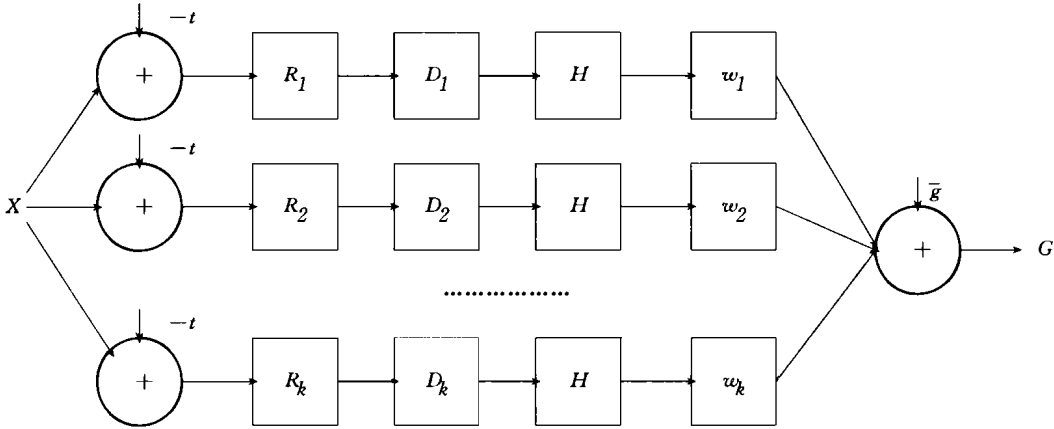


图 1 小波神经网络

Fig. 1 Sketch of wavelet neural network used

设 θ 表示所有参数, $G_{\theta}(X)$ 表示 (5) 给出的函数, 如果目标函数为

$$C(\theta) = \frac{1}{2} \sum_{k=1}^K [G_{\theta}(X_k) - y_k]^2 \tag{6}$$

对此神经网络采用类似于 BP 算法的梯度法对网络参数 $w_i, t_i, \mathbf{D}_i, \mathbf{R}_i, \bar{g}$ ($i=1, 2, \cdots, N$) 进行训练学习即可确定它们的值。为了方便起见, 引进记号

$$\begin{aligned} \Psi(X) &= \frac{\partial \Psi(X)}{\partial X} \\ e_k &= G_{\theta}(X_k) - y_k \\ z_i &= \mathbf{D}_i \mathbf{R}_i (X_k - \mathbf{t}_i) \end{aligned}$$

则相应的偏导数为

$$\begin{aligned} \frac{\partial C}{\partial \bar{g}} &= \sum_{k=1}^K e_k \\ \frac{\partial C}{\partial v_i} &= \sum_{k=1}^K e_k \Psi(z_i) \\ \frac{\partial C}{\partial a_i} &= - \sum_{k=1}^K e_k w_i \mathbf{R}_i^T \Psi'(z_i) \\ \frac{\partial C}{\partial s_i} &= - \sum_{k=1}^K e_k w_i \mathbf{D}_i^2 \text{diag}[\mathbf{R}_i (X_k - \mathbf{t}_i)] \Psi'(z_i) \end{aligned}$$

$$\frac{\partial C}{\partial R} = \sum_{k=1}^K e_{kWi} D_i \Psi(z_i) (X_k - t_i)^T$$

2 温度数据的拟合

我们用以上方法对 1958~1978 年共 21 年的月平均温度数据进行处理。
首先对原始数据作预处理。因为温度资料是季节性的,所以先作差分运算,以 $y(t)$ 表示原始数据,则

$$\Delta y(t) = y(t) - y(t - 12)$$

再把差分后的数据按以下方法初始化在[- 1, 1]内

$$Max = \max_i \{ \Delta y(t) \}$$

$$Min = \min_i \{ \Delta y(t) \}$$

$$dif = Max - Min$$

$$X_t = \frac{2 \times [\Delta y(t) - Min]}{dif - 1}$$

经过以上预处理后,再采用小波网络方法来拟合函数

$$x_t = F(x_{t-1}, \cdots, x_{t-d})$$

取基本波函数为

$$\Psi(x) = - x e^{-x^2/2}$$

小波数 N 取 5,我们用误差平方和 RE 来衡量拟合效果,对 d 取不同的数值时(即阶数取不同的值),结果为

$$\begin{aligned} N = 5, \quad d = 2, \quad RE &= 113.052\ 800 \\ N = 5, \quad d = 3, \quad RE &= 3.610\ 800 \\ N = 5, \quad d = 4, \quad RE &= 2.432\ 686 \\ N = 5, \quad d = 5, \quad RE &= 2.451\ 170 \end{aligned}$$

由此可见,当变量个数为 4 时,误差平方和已经非常小,此时的拟合效果已相当好。此时网络参数值为

$$\bar{g} = -0.139\ 407$$

$$W[i] (i = 1, 2, 3, 4, 5)$$

i				
1	2	3	4	5
- 2.204 959	- 5.699 178	4.291 621	2.663 922	4.635 011

$$D[i][j] (i = 1, 2, 3, 4, 5; j = 1, 2, 3, 4)$$

j	i				
	1	2	3	4	5
1	0.631 732	1.218 065	1.018 380	1.233 198	0.998 829
2	0.997 238	1.441 960	1.037 655	0.000 500	1.162 965
3	1.551 825	1.563 475	0.787 674	0.678 567	0.869 472
4	1.013 967	0.989 462	1.000 911	1.363 571	2.703 972

t[i][j](i= 1, 2, 3, 4, 5; j= 1, 2, 3, 4)

j	i				
	1	2	3	4	5
1	- 0. 648 533	- 0. 836 817	- 0. 505 240	- 0. 936 022	- 0. 515 148
2	- 1. 422 195	- 1. 286 090	- 1. 159 272	0. 985 999	1. 459 674
3	1. 500 000	1. 435 902	- 0. 911 617	- 0. 811 838	0. 022 623
4	0. 635 895	- 1. 421 685	- 0. 679 645	- 0. 247 362	- 0. 681 385

R[i][j][k](i= 1, 2, 3, 4, 5; j, k= 1, 2, 3, 4)

k	j	i				
		1	2	3	4	5
1	1	0. 782 380	0. 084 281	- 0. 511 546	- 0. 345 107	- 0. 068 594
	2	0. 994 353	0. 004 363	0. 080 862	0. 500 716	0. 040 822
	3	0. 855 797	- 0. 123 407	0. 363 950	- 0. 049 932	- 0. 076 902
	4	0. 926 894	0. 609 637	- 0. 707 540	0. 238 085	- 0. 266 542
2	1	0. 089 497	- 0. 315 140	- 0. 115 948	0. 937 674	- 0. 551 877
	2	- 0. 583 067	- 0. 557 173	- 0. 212 184	0. 561 929	0. 245 165
	3	- 0. 787 038	- 0. 068 558	0. 313 641	- 0. 186 481	- 0. 547 654
	4	- 0. 752 946	- 0. 002 999	0. 979 050	- 0. 157 190	- 0. 129 399
3	1	0. 481 936	0. 075 925	0. 782 377	- 0. 387 115	0. 818 143
	2	0. 030 353	- 0. 251 498	0. 516 206	0. 926 826	0. 166 136
	3	- 0. 251 164	0. 224 295	- 0. 199 821	0. 921 848	- 0. 282 891
	4	- 0. 173 898	0. 316 604	0. 125 156	0. 619 185	- 0. 707 607
4	1	- 0. 028 743	0. 327 003	0. 688 113	0. 647 103	0. 334 876
	2	0. 593 239	0. 512 403	- 0. 522 846	0. 248 189	0. 668 291
	3	- 0. 637 285	0. 292 672	0. 561 958	- 0. 089 477	0. 443 141
	4	0. 692 693	- 0. 714 468	0. 439 827	0. 367 338	0. 401 436

ERROR= 2. 432 686

同时也可看出, 用非线性模型处理此动态数据时, 只要四阶数即可, 而对同样的问题采用线性模型来处理, 用有关的定阶准则定阶, 那么阶数将达到 12 阶以上。因此, 用小波网络处理非线性时间序列分析问题是一个有效的方法。当然, 用小波网络时, 小波数和变量数都不宜太大, 否则, 相应的参数个数会迅速增加, 在训练学习网络参数时计算量和收敛性都会受影响。

参 考 文 献

1 Zhang Qinghua, Albert Benveniste. Wavelet Networks. IEEE Ttans. On Neural Networks, 1992, 3: 889~898

2 Cheng Bing, Titterington D M. Neural networks: a review from astatistical perspective. statistical science, 1994, 9(1): 2~54

3 曹良月, 高辉清. 用确定性方法预测工业总产值. 预测, 1995, (2): 53~56

4 陈兆国. 时间序列及其谱分析. 北京: 科学出版社, 1988

5 焦李成. 神经网络的应用与实现. 西安: 西安电子科技大学出版社, 1995

USE OF WAVELET NETWORK IN NONLINEAR MODELS OF TIMESERIES ANALYSIS

Qin Weiliang

(*Department of Basic Courses, NIM, Nanjing 210044*)

Jin Long

(*Jiangsu Meteorological Administration, Nanjing 210008*)

Abstract The paper is devoted to application of the wavelet network technique to nonlinear models for timeseries analysis and the nonlinear approximation by the method is able to construct a nonlinear function that is used to process temperature data, resulting in good fitting accuracy.

Keywords timeseries, wavelet network, nonlinear model