

# 基于广义逆矩阵的多层径向基函数网络遗传算法

盛国敏, 庄健

(安徽工业大学商学院, 安徽马鞍山 243032)

**摘要:**为了解决以往正则最小二乘法求权重向量时遇到的矩阵接近奇异而无法求逆的问题,采用广义逆矩阵的方法求多层径向基函数网络中各层的权重向量,并将这种方法引至多层径向基函数网络的遗传算法中。采用实函数逼近,混沌时间序列建模与预测等仿真实验对算法进行验证。结果表明,采用广义逆矩阵的方法要比正则最小二乘法在逼近精度上高1至2个数量级。  
**关键词:**多层径向基函数网络;遗传算法;广义逆矩阵;实函数逼近;混沌时间序列

**中图分类号:**TP 389.1   **文献标志码:**A   **doi:**10.3969/j.issn.1671-7872.2016.04.015

## Genetic Algorithm for Multilayer Radial Basis Function Networks Based on Generalized Inverse Matrix

SHENG Guomin, ZHUANG Jian

(School of Business, Anhui University of Technology, Ma'anshan 243032, China)

**Abstract:** In order to solve the problem that maxtrice are nearly singular when using regular least squares method, a method of generalized inverse matrix was employed to obtain the weight vectors of each layer in the multilayer radial basis function network, which was introduced to the genetic algorithm for training multilayer radial basis function networks. By using real function approximation, chaotic time series modeling and forecasting simulation experiments, the algorithm was verified. The results show that the generalized inverse matrix method is much more better than regular least squares method on the approximation precision, which can be up to a 1 to 2 orders of magnitude.

**Key words:** multilayer radial basis function networks; genetic algorithm; generalized inverse matrix; real function approximation; chaos sequence

聚类算法经历了从单层径向基函数网络<sup>[1]</sup>向多层径向基函数(RBF)网络<sup>[2]</sup>,并由最初的采用多层径向基函数网络的聚类法<sup>[2-4]</sup>发展到了采用多层径向基函数网络的遗传算法<sup>[5]</sup>,进一步发展为复合多层RBF网络<sup>[6]</sup>的发展历程,仿真实验表明,多层RBF网络逼近实函数的能力很强,将其用于混沌时间序列<sup>[7]</sup>的建模与预测中,可取得良好的效果,将其用于偏微分方程的无网格法<sup>[8-9]</sup>和求解期权定价值<sup>[10-12]</sup>可获得高精度的数值解。然而,多层RBF网络的诸算法均存在数值计算问题,即逆矩阵的计算遇到矩阵接近奇异而难以求逆的问题。这使得网络的学习速度变慢,权向量误差大,逼近实数精度的提高受到限制。为解决问题,本文尝试使用广义逆矩阵<sup>[13]</sup>代替普通逆矩阵求最小二乘解,并将其引入多层RBF网络的遗传算法。

## 1 多层径向基函数网络原理

单层RBF网络学习速度快,具有一定的实函数逼近能力,但是其拟合函数的精度有限。多层RBF网络的提出进一步提高RBF网络的实函数逼近能力,基本思想:用第2层RBF网络去拟合第1层网络的拟合残差

收稿日期:2015-07-29

作者简介:盛国敏(1989-),男,安徽天长人,硕士,助教,研究方向为机器学习。

通信作者:庄健(1957-),男,上海市人,博士,研究员,研究方向为机器学习。

函数,然后再用第3层网络RBF网络去拟合第2层网络的拟合残差函数,如此进行下去,后一层网络拟合前一层的拟合误差,就得到一个高精度的多层径向函数网络。

设第1层RBF网络为

$$f^{(1)}(x) = w_0^{(1)} + \sum_{k=1}^{n_1} w_k^{(1)} \Phi_k^{(1)}(x) \tag{1}$$

这里

$$\Phi_k^{(1)}(x) = \exp\left(-\frac{\|x - \bar{x}_k^{(1)}\|^2}{c_k^{(1)}}\right) \tag{2}$$

其中:  $\bar{x}_k^{(1)}$  ( $1 \leq k \leq n_1$ ) 为第1层的核函数中心;  $c_k^{(1)}$  则是宽度系数,  $w_k^{(1)}$  为权重。在文献[3]提出的算法中,核函数中心和宽度系数的优化是使广义交叉率  $G$  最小,权重采用正则最小二乘法得到,即如式(3)拟合残差平方和

$$R_1 = \sum_{i=1}^N (y_i - f^{(1)}(x_i))^2 + \lambda w_1^T K w_1 \tag{3}$$

取最小值,式中  $w_1$  为权重  $w_k^{(1)}$  ( $0 \leq k \leq n_1$ ) 的向量表达形式。记

$$e_i^{(1)} = y_i - f^{(1)}(x_i) \quad 1 \leq i \leq N \tag{4}$$

这是第1层网络在  $x_i$  处的输出值与期望值  $y_i$  的拟合残差。第2层RBF网络为

$$f^{(2)}(x) = w_0^{(2)} + \sum_{k=1}^{n_2} w_k^{(2)} \Phi_k^{(2)}(x) = w_0^{(2)} + \sum_{k=1}^{n_2} w_k^{(2)} \exp\left(-\frac{\|x - \bar{x}_k^{(2)}\|^2}{c_k^{(2)}}\right) \tag{5}$$

式中各参数的含义和第1层类同。核函数中心和宽度系数的取值也使第2层的广义交叉率  $G$  最小,权重则使

$$R_2 = \sum_{i=1}^N (e_i^{(1)} - f^{(2)}(x_i))^2 + \lambda w_2^T K w_2 \tag{6}$$

最小,其中  $w_2$  为权重  $w_k^{(2)}$  ( $0 \leq k \leq n_2$ ) 的向量表达形式。于是可得一个更精确的模型

$$F_2(x) = f^{(1)}(x) + f^{(2)}(x) \tag{7}$$

如此进行下去,用后一层径向基函数网络去拟合前一层网络的拟合残差,便得到一个多层RBF网络

$$F_p(x) = f^{(1)}(x) + f^{(2)}(x) + f^{(3)}(x) + \dots + f^{(p)}(x) \tag{8}$$

根据适当的判别准则,可以决定增加到几层最为适合。

## 2 多层径向基函数网络的遗传算法

采用自适应遗传算法训练多层RBF网络,具体如下。

1) 对RBF网络的进行遗传编码 采用实数编码的方法,设第  $k$  层网络核函数的个数为  $n_k$ ,每个核函数的输入维数为  $m$ 。此时,每个个体用  $(m+1) \times n_k$  维的向量来表示,向量的前  $m \times n_k$  个分量表示RBF网络的中心,后  $n_k$  个分量表示中心所对应的宽度系数。

2) 选取适应度函数 为了既能得到最大限度的高逼近精度,又不出现过拟合现象,每个个体的适应度函数可以选取为RBF网络的广义交叉率的倒数,即对第  $i$  个个体,其适应度为

$$f_{\text{fitness}}(i) = 1/G(i) \tag{9}$$

这里  $G(i)$  为第  $i$  个个体的广义交叉率<sup>[3]</sup>。

3) 选择操作 首先根据每一代种群的个体的适应度的大小,得到最优个体和最劣个体,用最优个体替换最劣个体,直接复制给下一代,不参加交叉和变异的过程。剩下的种群个体通过赌轮盘法进行筛选,然后进行交叉和变异操作。

4) 交叉和变异 记交叉概率为  $p_c$ ,变异概率  $p_m$ ,其自适应计算方法如下:

$$p_c = \begin{cases} a - (a - c)(f' - f_{\text{avg}}) / (f_{\text{max}} - f_{\text{avg}}), & f' \geq f_{\text{avg}} \\ a & f' < f_{\text{avg}} \end{cases} \quad (10)$$

$$p_m = \begin{cases} b - (b - d)(f - f_{\text{avg}}) / (f_{\text{max}} - f_{\text{avg}}), & f \geq f_{\text{avg}} \\ b, & f < f_{\text{avg}} \end{cases} \quad (11)$$

其中:  $a=0.9; b=0.1, c=0.6, d=0.001$ ;  $f_{\text{max}}$  表示群体中最大适应度值;  $f_{\text{avg}}$  表示每一代群体的平均适应度值;  $f'$  表示要交叉的两个个体中较大的适应度值;  $f$  表示要变异个体的适应度值。

5) 权重的计算 采用正则最小二乘法,

$$\mathbf{w} = (\mathbf{B}^T \mathbf{B} + \lambda \mathbf{I})^{-1} \mathbf{Bz} \quad (12)$$

其中:  $\mathbf{z} = (z_1, z_2, \dots, z_N)$  为期望输出向量;  $\mathbf{B}$  的计算如下

$$\mathbf{B} = \begin{bmatrix} 1 & \Phi_1(x_1) & \Phi_2(x_1) & \cdots & \Phi_n(x_1) \\ 1 & \Phi_1(x_2) & \Phi_2(x_2) & \cdots & \Phi_n(x_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \Phi_1(x_n) & \Phi_2(x_n) & \cdots & \Phi_n(x_n) \end{bmatrix} \quad (13)$$

$$\mathbf{K} = \mathbf{D}^T \mathbf{D} \quad (14)$$

$$\mathbf{D} = \begin{bmatrix} 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & -2 & 1 \end{bmatrix} \quad (15)$$

这里,  $\lambda$  为求权重时的正则最小二乘法的参数。当  $\lambda=0$  时,  $\mathbf{w}$  即为普通最小二乘解。对第 1 层网络,  $\mathbf{z}_\alpha = \mathbf{y}_\alpha$ ,  $f(\cdot) = f_1(\cdot)$ 。对第  $i$  层网络 ( $i \geq 2$ ),  $\mathbf{z}_\alpha = e_\alpha^{i-1}$ ,  $f(\cdot) = f_i(\cdot)$ 。对第  $k$  层网络 ( $k \geq 1$ ),  $\Phi_i(x_j) = \Phi_i^k(x_j)$ 。

6) RBF 网络层数的确定 设第  $m$  层的广义交叉率为  $G(m)$ , 第  $m+1$  层则为  $G(m+1)$ , 如果

$$(G(m) - G(m+1)) / G(m) < \varepsilon \quad (16)$$

则停止加层, 网络层数为  $m$ 。

### 3 广义逆矩阵法在多层径向基函数网络遗传算法中应用

在大多数情况下, 由于式(12)中  $\mathbf{B}^T \mathbf{B}$  接近奇异, 使得计算逆矩阵  $(\mathbf{B}^T \mathbf{B} + \lambda \mathbf{I})^{-1}$  很困难, 勉强得到的数值解精度很差, 从而使整个网络的精度大大下降, 为解决这一问题, 采用广义逆矩阵求最小二乘法<sup>[8]</sup>。设  $\mathbf{A}$  为  $m \times n$  类数矩阵,  $\mathbf{A}$  的广义逆矩阵<sup>[9]</sup>  $\mathbf{X}$  满足下面 4 个条件, 即,

$$\mathbf{AXA} = \mathbf{A} \quad (17)$$

$$\mathbf{XAX} = \mathbf{X} \quad (18)$$

$$(\mathbf{AX})^T = \mathbf{AX} \quad (19)$$

$$(\mathbf{XA})^T = \mathbf{XA} \quad (20)$$

$\mathbf{X}$  常记为  $\mathbf{A}^+$ 。于是, 线性模型

$$\mathbf{y} = \lambda \boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (21)$$

的最小二乘估计  $\hat{\boldsymbol{\beta}}$  为

$$\hat{\boldsymbol{\beta}} = (\lambda^T \lambda)^+ \lambda^T \mathbf{y} \quad (22)$$

将式(20)用于 RBF 网络, 并采用最小二乘法, 得到

$$\mathbf{W} = (\mathbf{B}^T \mathbf{B})^+ \mathbf{Bz} \quad (23)$$

这里  $\mathbf{B}$  由式(13)给出, 再给出本文使用的广义逆矩阵的一个迭代算法, 取初始值

$$\lambda_0 = \alpha \mathbf{A}^T \quad (24)$$

这里,  $0 < \alpha < \frac{\alpha}{\lambda_1(A^T A)}$

其中  $\lambda_1(A^T A)$  为  $A^T A$  的最大特征值, 迭代公式

$$A_{k+1} = A_k + \lambda_0(I - AA_k), k=0, 1, 2, \dots \tag{25}$$

则  $\lim_{x \rightarrow \infty} A_k = A^+$ 。

### 4 仿真实验

遗传算法的参数为:进化的代数为 100, 种群的规模为 200, 学习参数为  $a=0.9, b=0.1, c=0.6, d=0.001$ 。根据  $G$  决定层数确定层数, 控制参数为 0.05。

#### 4.1 实验一

考虑函数

$$y = 4x(1-x), x \in (0, 1)$$

随机生成 400 个在 (0, 1) 上服从均匀分布的样本。200 个样本用于训练, 其余 200 个用于检验。网络构造为: 第 1 层 RBF 网络取 10 个径向基函数, 第 2 层取 20 个, 第 3 层以后取 30 个。

表 1 给出了预测结果以及采用基于正则最小二乘法的遗传算法训练得到的网络的学习误差和预测误差。由表 1 可见基于广义逆矩阵的遗传算法的精度要比基于正则最小二乘法的遗传算法高约 1 个数量级。

表 1 多层 RBF 网络的训练与预测结果

Tab. 1 Training and prediction results of multilayer RBF network

算法类型	网络层数	误差类型	RSS	RMSE	MAE
广义逆矩阵法	8	学习误差	2.79e-16	1.18e-10	9.24e-11
		预测误差	5.29e-16	1.63e-10	1.12e-10
正则最小二乘法	10	学习误差	-	4.21e-09	2.73e-09
		预测误差	-	5.50e-08	9.21e-09

注:RSS 为最后一层的残差平方和;RMSE 为均方根误差;MAE 为平均绝对误差。下表同。

#### 4.2 实验二

考虑函数

$$z = \exp(-(x-0.5)^2/(2 \cdot 0.2^2) - (y-0.5)^2/(2 \cdot 0.2^2)) \quad x, y \in [0, 1]$$

生成 3 200 个服从在 (0, 1) 上均匀分布的样本, 其中 1 600 个作为训练用样本, 其余 1 600 个用于检验。网络构造为: 第 1 层 RBF 网络取 40 个径向基函数, 第 2 层以后取 80 个。

表 2 给出了预测结果以及采用基于正则最小二乘法遗传算法训练得到的网络学习误差和预测误差。由表 2 可见基于广义逆矩阵的遗传算法精度要比基于正则最小二乘法的遗传算法高约 2 个数量级。

表 2 多层 RBF 网络的训练与预测结果

Tab. 2 Training and prediction results of multilayer RBF network

算法类型	网络层数	误差类型	RSS	RMSE	MAE
广义逆矩阵法	15	学习误差	1.62e-09	3.19e-08	2.42e-08
		预测误差	6.54e-09	6.40e-08	3.01e-08
正则最小二乘法	5	学习误差	-	6.12e-06	4.80e-06
		预测误差	-	7.17e-06	5.12e-06

#### 4.2 实验三

考虑 Logistic 模型

$$x_{n+1} = Rx_n(1-x_n)$$

其中  $R$  是参数, 取  $R=3.7$ 。取初始值  $x_0=0.27$ 。共产生 700 个时间序列样本, 前 600 个用于学习, 后 100 个用于检验预测效果。前 700 个中, 前 650 个用于学习, 后 50 个用于计算可预测步长。网络构造为: 第 1 层 RBF 网络取 80 个径向基函数, 第 2 层以后取 120 个, 嵌入维数取  $p=7$ 。

表3给出了一步预测结果以及采用基于正则最小二乘法遗传算法训练而得到的网络学习误差和一步预测误差。由表3可见基于广义逆矩阵的遗传算法精度要比基于正则最小二乘法的遗传算法高约1个数量级。

表3 多层RBF网络的训练和一步预测结果

Tab. 3 Training and one step prediction results of multilayer RBF network

算法类型	网络层数	误差类型	RSS	RMSE	MAE
广义逆矩阵法	6	学习误差	2.13e-11	5.76e-08	4.240 9e-08
		预测误差	3.01e-12	7.84e-08	5.890 6e-08
正则最小二乘法	7	学习误差	2.17e-10	5.81e-07	4.384 6e-07
		预测误差	8.09e-11	1.29e-06	8.385 6e-07

表4给出了多步预测的可预测步长和预测精度以及基于正则最小二乘法遗传算法训练而得到的网络可预测步长和预测精度。可见在同样的误差容许范围内,基于广义逆矩阵的遗传算法的可预测步长要比基于正则最小二乘法的遗传算法长许多。

表4 多层RBF网络的多步预测结果

Tab. 4 Multi-step prediction results of multilayer RBF network

算法类型	网络层数	可预测步长	RSS	RMSE	MAE
广义逆矩阵法	6	32	0.183 2	0.076 9	0.027 2
正则最小二乘法	7	21	0.244 2	0.110 5	0.048 3

#### 4.4 实验四

考虑Mackey\_Glass混沌时间序列,它由如下Mackey Glass方程所产生

$$x_t = x_{t-1} + \left[ \frac{ax_{t-\tau}}{1 + (x_{t-\tau})^c} - bx_{t-1} \right]$$

$\tau$ 为时滞参数。这里取 $a=0.2$ , $b=0.1$ , $c=10$ , $\tau=30$ , $x_0=1.2$ 。产生800个时间序列的样本,前700个作为学习用样本,后100个作为预测用样本。

在Mackey Glass方程中有一个线性项,由于高斯函数为峰状函数,因此RBF网络在拟合线性项的时候精度不高,如果仅仅使用高斯函数,模型的逼近精度不高,预测效果不好。为此在第1层网络中加上了线性项,其具体形式为

$$f(x) = a_0 + a_1x_1 + a_2x_2 + \cdots + a_nx_n + \sum_{i=1}^m w_i^{(0)} \Phi_i \left( \left\| x - \bar{x}_i^{(0)} \right\|^2 \right)$$

网络构造为第1层RBF网络取80个径向基函数,第2层以后取120个,嵌入维数取 $p=6$ 。学习用样本的前650个用于训练,后50个用于计算可预测步长。

表5给出了一步预测结果及采用基于正则最小二乘法遗传算法训练而得到的网络学习精度和一步预测精度。由表5可见基于广义逆矩阵遗传算法的精度要比基于正则最小二乘法遗传算法高约1到2个数量级。

表5 多层RBF网络的训练和一步预测结果

Tab. 5 Training and one step prediction results of multilayer RBF network

算法类型	网络层数	误差类型	RSS	RMSE	MAE
广义逆矩阵法	2	学习误差	1.16e-07	1.34e-06	9.39e-07
		预测误差	5.13e-09	1.02e-06	7.87e-07
正则最小二乘法	4	学习误差	4.58e-06	8.44e-05	5.93e-05
		预测误差	2.93e-07	7.74e-05	5.70e-05

表6给出了多步预测的可预测步长和预测精度,以及由基于正则最小二乘法遗传算法训练而得到的网络可预测步长和预测精度。可见在同样的误差容许范围内,基于广义逆矩阵的遗传算法的可预测步长要比基于正则最小二乘法的遗传算法长许多。

表6 多层RBF网络的多步预测结果

Tab. 6 Multi-step prediction results of multilayer RBF network

算法类型	网络层数	可预测步长	RSS	RMSE	MAE
广义逆矩阵法	2	38	0.100 6	0.052 9	0.043 1
正则最小二乘法	4	32	0.009 4	0.070 5	0.050 2

## 5 结 论

用广义逆矩阵计算多层RBF网络的每层权连接,仿真实验表明,采用广义逆矩阵遗传算法得到的权重其精度大大提高,网络的后一层能更好地拟合前一层的拟合误差,整个网络的实函数逼近能力就这样大大提高。将这个方法用于混沌时间序列的建模与预测得到了良好的效果。

### 参考文献:

- [1] MOODY J, DARKEN C. Fast learning in networks of locally-tuned processing units[J]. Neural Computation, 1989, 1(2):281-294.
- [2] ZHANG J, JIANG H F, WU R L. Multi layer RBF network for real functions approximation and nonlinear regression[C/CD]// Scientific Research Publishing. Proceedings of 2011 International Symposium on Statistics & Management Science. Chongqing: Scientific Research Publishing, 2011.
- [3] 吴瑞柳. 多层径向基函数网络的聚类学习方法及应用[D]. 马鞍山:安徽工业大学, 2011.
- [4] 李建. 多层径向基函数网络的算法改进及其应用[D]. 马鞍山:安徽工业大学, 2013.
- [5] 盛飞. 多层RBF网络的自适应遗传算法及其在实函数逼近中的应用[J]. 安徽工业大学学报(自然科学版), 2013, 30(2):192-196.
- [6] 徐光鲁. 复合多层径向基函数网络及其在偏微分方程中的应用[J]. 安徽工业大学学报(自然科学版), 2015, 32(1):76-80.
- [7] 余健, 郭平. 基于RBF网络的金融时间序列预测[J]. 湖南工程学院学报(自然科学版), 2007, 17(4):44-48.
- [8] LIU G R, GU Y T. 无网格法理论及程序设计[M]. 王建明, 周学军, 译. 济南:山东大学出版社, 2007:259-264.
- [9] 张雄, 刘岩. 无网格法[M]. 北京:清华大学出版社, 2004:95-175.
- [10] 侯木舟, 周耀琼. 二叉树期权定价方法的一种新推广[J]. 数学理论与应用, 2006, 26(3):111-115.
- [11] 牟旷凝. 蒙特卡洛方法和拟蒙特卡洛方法在期权定价中应用的比较研究[J]. 科学技术与工程, 2010, 10(8):1925-1929.
- [12] 蹇明, 宜娜, 张春晓. 期权定价模型的五点式混合差分方法[J]. 经济数学, 2011, 28(4):66-70.
- [13] 王宋桂, 杨振海. 广义逆矩阵及其应用[M]. 北京:北京工业大学出版社, 1996.

责任编辑:丁吉海