

基于加速梯度求法的置信规则库参数训练方法*

吴伟昆¹, 杨隆浩¹, 傅仰耿¹, 张立群¹, 巩晓婷²⁺

1. 福州大学 数学与计算机科学学院 福州 350116

2. 福州大学 经济与管理学院 福州 350116

Parameter Training Approach for Belief Rule Base Using the Accelerating of Gradient Algorithm*

WU Weikun¹, YANG Longhao¹, FU Yanggeng¹, ZHANG Liqueun¹, GONG Xiaoting²⁺

1. College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350116, China

2. College of Economics and Management, Fuzhou University, Fuzhou 350116, China

+ Corresponding author: E-mail: xtgong@126.com

WU Weikun, YANG Longhao, FU Yanggeng, et al. Parameter training approach for belief rule base using the accelerating of gradient algorithm. Journal of Frontiers of Computer Science and Technology, 2014, 8(8) :989-1001.

Abstract: The problem of training parameters for belief rule base (BRB) is essentially a nonlinear optimization problem with constraints, which is mainly solved by the FMINCON function or the swarm intelligence algorithms. However, these approaches have many shortages, such as poor portability, difficult to be implemented and requiring a large amount of calculation. To solve these problems, this paper proposes a new parameter training approach for BRB using the accelerating of gradient algorithm, which is improved from the existing parameter training methods, and is applied to the parameter training of multimodal function and pipeline leak detection. The proposed approach is compared with other traditional approaches in terms of convergence error, convergence precision and Pearson correlation coefficient in experiment analysis. The results show the better comprehensive benefits of the proposed approach, including convergence accuracy and convergence speed.

* The National Natural Science Foundation of China under Grant Nos. 71371053, 61300026, 61300104 (国家自然科学基金); the Science and Technology Project of Fujian Education Department under Grant No. JA13036 (福建省教育厅科技项目); the National Collegiate Innovation and Entrepreneurship Training Program of China under Grant No. 121038607 (国家级大学生创新创业训练计划项目); the Science and Technology Development Foundation of Fuzhou University under Grant No. 2014-XQ-26 (福州大学科技发展基金项目).

Received 2014-02, Accepted 2014-04.

CNKI网络优先出版 2014-04-22, <http://www.cnki.net/kcms/doi/10.3778/j.issn.1673-9418.1403061.html>

Key words: belief rule base (BRB); parameter training; nonlinear optimization problem; the accelerating of gradient algorithm

摘要 置信规则库(belief rule base ,BRB)的参数训练问题实质上是一个带有约束条件的非线性优化问题,目前在求解该问题上主要使用 FMINCON 函数及群智能算法,但在算法的应用中存在移植性差,难实现,计算时间长等局限性。通过对这些问题的研究,结合现有的参数训练方法提出了基于加速梯度求法的置信规则库参数训练方法,并将其应用在多峰函数、输油管道泄漏检测的置信规则库的参数训练上。以收敛误差、收敛时间和皮尔森相关系数作为衡量指标,对新方法与其他传统方法进行了对比,实验结果表明,新算法在收敛精度和收敛速度上具有更理想的综合效益。

关键词: 置信规则库(BRB), 参数训练, 非线性优化问题, 加速梯度求法

文献标志码: A **中图分类号**: TP18

1 引言

置信规则库(belief rule base ,BRB)由带有置信度分布表示类型的 IF-THEN 规则^[1]构成,其将输入的定量信息或定性知识在现有的知识库中进行推理、分析,并最终为决策者提供有效的决策依据,因此在解决现实中的决策问题时,BRB 系统能够准确地模拟专家的决策行为。BRB 系统中的核心算法是基于证据推理的置信规则库推理方法^[2](belief rule-base inference methodology using the evidential reasoning approach ,RIMER),RIMER 是在 D-S 证据理论^[3-4]、决策理论^[5]、模糊理论^[6]和传统 IF-THEN 规则库^[2]的基础上发展起来的,具有对带含糊或模糊不确定性、不完整性或概率不确定性,以及非线性特征的数据进行建模的能力。目前 BRB 系统已成功应用于武器装备体系能力的评估^[7]、输油管道检漏^[8]、石墨成分检测^[9]等领域。

BRB 系统中包含的参数有规则权重、前提属性候选值、前提属性权重、置信度、评级等级等,这些参数通常由专家根据经验和历史信息给定,致使准确性受限于专家对所研究对象的认知程度。当 BRB 系统变得庞大复杂时,专家很难确定这些参数的精确值,因此有必要建立一个参数优化模型来训练 BRB。目前可用于求解该参数优化模型的有 Matlab 中的 FMINCON 函数^[9]、梯度法^[10]、粒子群算法^[11]、差分进化算法^[12]、Wolf 简约梯度法^[13]等。但在实际应用中这些方法仍存在难实现,算法复杂度高等不足,诸

如粒子群算法的最优解对初始值的依赖性较大,Wolf 简约梯度法中的基变量变换十分繁琐和困难等。此外,常瑞等人^[10]基于梯度法提出的参数训练方法,虽然优于 FMINCON 函数,但是参数训练仅涉及部分参数,算法的收敛精度还有待提升。

鉴于此,本文结合梯度法及梯度快速求解方法,提出了基于加速梯度求法的置信规则库参数训练方法。对于新方法中涉及的参数,将对常瑞方法的局部模型进行扩展,提出更加全面的参数优化模型,进而提升经参数训练后 BRB 系统的推理性能。在实验分析中,本文首先引入了多峰函数的实例,重点对比分析了本文方法与常瑞方法,并以此说明了本文方法在收敛速度及收敛精度上具有更合理的综合效益。然后在输油管道检漏的实例中,通过对比现有的其他参数训练方法验证了本文方法的有效性。

本文组织结构如下:第 2 章介绍 BRB 系统的现有理论知识及本文拟解决的问题,第 3 章介绍快速求解梯度的方法及改进的参数训练模型,并在此基础上提出解决现有问题的新参数训练方法,第 4 章通过两个实例分析新参数训练方法的有效性,最后对全文进行总结。

2 BRB 系统与问题提出

为处理定量信息或定性知识中的不确定信息,Yang 等人^[2]构建了 BRB 系统,通过在传统规则库中加入分布式置信框架、权重参数和参数训练方法来

提高BRB系统决策的准确性。以下对BRB系统的基础知识进行简要介绍。

2.1 BRB 的表示

$$R_k: \text{If } A_1^k \wedge A_2^k \wedge \cdots \wedge A_{T_k}^k \quad (1)$$

$$\text{Then } \{(D_1, \bar{\beta}_{1,k}), (D_2, \bar{\beta}_{2,k}), \dots, (D_N, \bar{\beta}_{N,k})\}$$

式(1)中, $A_i^k (i=1, 2, \dots, T_k, k=1, 2, \dots, L)$ 表示第 k 条规则的第 i 个前提属性的候选值, T_k 表示第 k 条规则中前提属性的个数, L 表示规则的总条数; $D_j (j=1, 2, \dots, N)$ 表示规则结果的评价等级, N 为评价等级的个数; $\bar{\beta}_{j,k} (j=1, 2, \dots, N, k=1, 2, \dots, L)$ 表示在第 k 条规则中输出部分相对于第 j 个评价等级 D_j 的置信度, 若 $\sum_{j=1}^N \bar{\beta}_{j,k} = 1$, 则第 k 条规则包含的信息是完整的, 否则第 k 条规则包含的信息不是完整的。

此外第 k 条规则中还包含规则权重 $\theta_k (k=1, 2, \dots, L)$ 和前提属性权重 $\delta_{k,i} (k=1, 2, \dots, L, i=1, 2, \dots, T_k)$ 。前者反映第 k 条规则相对BRB中其他规则的重要程度; 后者反映第 i 个前提属性相对其他前提属性的重要程度。

2.2 RIMER 方法

RIMER方法是BRB系统的核心内容, 其在规则推理时主要包含3个步骤: 首先是激活权重的计算; 然后是置信度的修正; 最后再使用证据推理(evidential reasoning, ER)算法合成激活规则。

2.2.1 激活权重的计算

激活权重的计算取决于输入数据、前提属性权重和规则权重, 其中第 k 条规则的激活权重的计算公式可表示为:

$$\left\{ \begin{aligned} \omega_k &= \frac{\theta_k \prod_{i=1}^{T_k} (\alpha_i^k)^{\delta_{k,i}}}{\sum_{l=1}^L \theta_l \prod_{i=1}^{T_k} (\alpha_i^l)^{\delta_{l,i}}} \\ \bar{\delta}_{k,i} &= \frac{\delta_{k,i}}{\max_{i=1, 2, \dots, T_k} \{\delta_{k,i}\}} \end{aligned} \right. \quad (2)$$

其中, $\omega_k \in [0, 1]$, $k=1, 2, \dots, L$; α_i^k 表示在第 k 条规则中第 i 个输入相对于候选值 A_i^k 的个体匹配度。

2.2.2 置信度的修正

当输入数据包含模糊、不确定数据时, 需要对结果部分的各评价等级的置信度进行修正。第 k 条规则的第 i 个评价等级 D_i 的置信度 $\bar{\beta}_{i,k}$ 的修正公式为:

$$\beta_{i,k} = \bar{\beta}_{i,k} \frac{\sum_{t=1}^{T_k} \left(\tau(t, k) \sum_{j=1}^{|A_t|} \alpha_{t,j} \right)}{\sum_{t=1}^{T_k} \tau(t, k)} \quad (3)$$

$$\tau(t, k) = \begin{cases} 1, & A_t \in R_k, t=1, 2, \dots, T_k \\ 0, & \text{otherwise} \end{cases}$$

其中, $|A_t|$ 表示前提属性的个数。如果输入数据是完整的, 则 $\beta_{i,k} = \bar{\beta}_{i,k}$ 。

2.2.3 激活规则的合成

在ER算法中, 由每个基本属性的置信度和属性权重可求得基本属性的基本可信值^[14-15]:

$$m_{n,i} = \omega_i \beta_{n,i} \quad (4.1)$$

$$m_{H,i} = 1 - \omega_i \sum_{n=1}^N \beta_{n,i} = \bar{m}_{H,i} + \tilde{m}_{H,i} \quad (4.2)$$

$$\tilde{m}_{H,i} = \omega_i (1 - \sum_{n=1}^N \beta_{n,i}) \quad (4.3)$$

$$\bar{m}_{H,i} = 1 - \omega_i \quad (4.4)$$

式中, $n=1, 2, \dots, N$, $i=1, 2, \dots, L$ 。

在基本可信值的基础上利用ER的解析公式^[16]可将所有的基本属性一次合成, 合成公式为:

$$C_n = k \left[\prod_{j=1}^L (m_{n,j} + \bar{m}_{H,j} + \tilde{m}_{H,j}) - \prod_{j=1}^L (\bar{m}_{H,j} + \tilde{m}_{H,j}) \right] \quad (5.1)$$

$$\tilde{C}_H = k \left[\prod_{j=1}^L (\bar{m}_{H,j} + \tilde{m}_{H,j}) - \prod_{j=1}^L \bar{m}_{H,j} \right] \quad (5.2)$$

$$\bar{C}_H = k \prod_{j=1}^L \bar{m}_{H,j} \quad (5.3)$$

$$k^{-1} = \sum_{n=1}^N \prod_{j=1}^L (m_{n,j} + \bar{m}_{H,j} + \tilde{m}_{H,j}) - (N-1) \prod_{j=1}^L (\bar{m}_{H,j} + \tilde{m}_{H,j}) \quad (5.4)$$

新合成属性的置信度为:

$$\beta_n = \frac{C_n}{1 - \bar{C}_H}, n=1, 2, \dots, N \quad (6.1)$$

$$\beta_H = \frac{\tilde{C}_H}{1 - \bar{C}_H} \quad (6.2)$$

其中, β_n 表示评价等级 D_n 的置信度; β_H 表示不完备信息的置信度。根据上述的合成公式可得 BRB 系统分布式置信度输出:

$$S(y) = \{(D_n, \beta_n), n = 1, 2, \dots, N\} \quad (7.1)$$

假设 $\mu(D_n)$ 为第 n 个评价等级 D_n 的效用值, 则 BRB 系统的数值型输出表示为:

$$y = \sum_{n=1}^N \mu(D_n) \beta_n \quad (7.2)$$

2.3 BRB 的参数训练

针对利用初始的 BRB 系统进行决策时准确性不高的问题, Yang 等人^[9]最早提出使用 Matlab 中 FMINCON 函数对初始的 BRB 进行参数训练, 但由于只有规则权重、属性权重和置信度参与训练, 算法的收敛精度与收敛效率均不够理想。Chen 等人^[17]在 Yang 的基础上拓宽了参与训练的参数范围, 提高了 BRB 系统的推理性能。然而算法仍旧是建立在 Matlab 优化函数的基础上, 同样存在收敛速度慢、不易移植的问题。另一方面, 为了提高参数训练时的收敛精度和收敛效率, 常瑞等人^[10]引入梯度下降法和二分法, 在 Yang 的参数训练模型的基础上提出了更高效的算法, 其所采用的参数训练模型如图 1 所示。

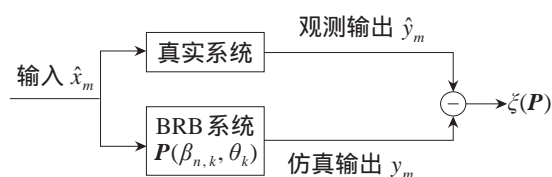


Fig.1 Chang Rui parameter training model of BRB

图1 常瑞的 BRB 参数训练模型

图 1 中 \hat{x}_m 是真实系统的输入; \hat{y}_m 是真实系统的观测输出; y_m 是 BRB 系统的仿真输出; $\zeta(P)$ 表示 \hat{y}_m 和 y_m 的差异值; P 是包含 $\beta_{n,k}$ 、 θ_k 的训练参数向量。显然, $\zeta(P)$ 越小, BRB 系统的推理性能越好。

由 Yang 的原始参数训练模型^[9]可知, 模型中各个参数还需满足以下约束条件:

(1) 置信规则中每个等级上的置信度均不能小于 0 或大于 1, 其中第 k 条规则的第 n 个评价等级上的置信度约束条件表示如下:

$$0 \leq \beta_{n,k} \leq 1, n = 1, 2, \dots, N, k = 1, 2, \dots, L \quad (8.1)$$

(2) 若第 k 条规则的输入不包含模糊、不确定信息, 则说明第 k 条规则是完整的, 即结果中所有置信度的和为 1。

$$\sum_{n=1}^N \beta_{n,k} = 1, k = 1, 2, \dots, L \quad (8.2)$$

(3) θ_k 表示第 k 条规则对整个规则库的贡献度, 经标准化后 θ_k 的值应落在 0 和 1 之间。

$$0 \leq \theta_k \leq 1, k = 1, 2, \dots, L \quad (8.3)$$

假设已知有 M 组输入-输出观测值 (\hat{x}_m, \hat{y}_m) ($m = 1, 2, \dots, M$), 依据 (7.1)、(7.2), 真实系统的观测输出可表示为:

$$\hat{y}_m = \sum_{j=1}^N \mu(D_j) \hat{\beta}_j(m) \quad (9.1)$$

同理, BRB 系统的仿真输出可表示为:

$$y_m = \sum_{j=1}^N \mu(D_j) \beta_j(m) \quad (9.2)$$

通过分析 BRB 的训练过程可知, 求解置信规则库训练模型的目标函数为:

$$\begin{cases} \min_P \zeta_j(P) \\ \text{s.t. (8.1)~(8.3)} \end{cases} \quad (10.1)$$

对于 BRB 系统与实际系统间的误差表示形式, 常用的有均方差 (mean square error, MSE), 因此目标函数可表示如下:

$$\zeta(P) = \frac{1}{M} \sum_{m=1}^M (y_m - \hat{y}_m)^2 \quad (10.2)$$

2.4 现有参数训练方法的瓶颈问题

在求解多参数下带有多个约束条件的最优值的参数训练中, 算法设计要解决的集中问题是如何提高收敛精度、收敛效率及简明算法流程。传统的参数训练方法以 Matlab 中的 FMINCON 函数为基础, 其收敛速度慢且不易移植。文献[11]中以粒子群算法 (particle swarm optimization, PSO) 作为参数优化算法, 但算法的初始值具有很大的随意性, 且最优解对初始值的依赖性较大。而差分进化算法存在搜索停滞和早熟收敛等缺陷^[18]。

常瑞等人使用图 1 的优化模型, 提出了基于梯度法和二分法^[10]的联合算法对 BRB 的参数进行训练, 并取得了比利用 FMINCON 函数更好的收敛精度和

收敛效率。但该模型只对部分参数进行训练,训练后BRB的推理性能存在很大的局限性,且联合算法在解决复杂问题时算法的设计上也存在不足之处。为解决现有参数训练的瓶颈问题,本文对图1优化模型进行了改进,提出了基于加速梯度求法的BRB参数训练方法。

3 基于加速梯度求法的BRB参数训练方法

3.1 梯度下降法及加速求法

梯度下降法(gradient descent algorithm, GDA)是一种求解问题最优解的算法,也称最速下降法,现广泛应用于求解无约束优化问题。其中相关的改进算法包括共轭梯度法、Wolf 简约梯度法、广义简约梯度法等。以下将结合梯度下降法在求解问题最优解时的算法步骤进行简要介绍,以BRB参数训练模型的目标函数为例,首先选取式(10.1)中目标函数的负梯度方向作为搜索方向:

$$d = -\nabla \zeta(P) \quad (11)$$

设 $\zeta(P)$ 在 P 附近连续可微,则搜索方向向量 d 为 $\zeta(P)$ 的各方向偏导值。为进一步分析搜索方向取负梯度方向的合理性,令 $g = \nabla \zeta(P)$,由泰勒展开式^[19]可得:

$$\zeta(P + \alpha d) = \zeta(P) + \alpha g^T d + o(\alpha), \alpha > 0 \quad (12)$$

其中, α 为由精确线搜索方法^[13]确定的步长因子。则目标函数 $\zeta(P)$ 在 P 处沿方向 d 下降的变化率可进行如下的推导:

$$\lim_{\alpha \rightarrow 0} \frac{\zeta(P + \alpha d) - \zeta(P)}{\alpha} = \lim_{\alpha \rightarrow 0} \frac{\alpha g^T d + o(\alpha)}{\alpha} = g^T d = \|g\| \|d\| \cos \theta \quad (13)$$

式(13)中, θ 表示 g 和 d 的夹角。显然,目标函数的变化率取决于 θ 的余弦值,因此为了让变化率达到最小,需要使 $\theta = \pi$,即式(11)中 d 取负梯度方向时,可求得让目标函数具有更优解的新迭代点 $P^{t+1} = P^t + \alpha d$ 。

对于负梯度方向的求解,目前主要是基于符号方程式求偏导的方法,其包括C++数学库函数、Matlab的符号求导函数 `diff()` 和符号赋值函数 `subs()`。但上述基于函数库的方法均表现出偏导符号公式难构造,算法效率低等不足^[20]。为此,本文根据偏导数的

定义提出了梯度的加速求解方法。基本思想是对待求偏导的变量加上一个极小量 h ,然后通过求极限的方式计算目标函数的偏导值。假设可偏导的目标函数为 $\zeta(x_1, x_2, \dots, x_n)$,则在点 (a_1, a_2, \dots, a_n) 处关于变量 $x_i (1 \leq i \leq n)$ 的偏导值求解公式如下:

$$\frac{\partial \zeta}{\partial x_i}(a_1, a_2, \dots, a_n) = \lim_{h \rightarrow 0} \frac{\zeta(a_1, a_2, \dots, a_i + h, \dots, a_n) - \zeta(a_1, a_2, \dots, a_i, \dots, a_n)}{h} \quad (14)$$

上述偏导无需构造符号方程式,实现简单且速度快。

3.2 新参数训练方法

根据梯度的加速求法,本文提出了一种基于加速梯度求法的BRB新参数训练方法。同时,相比常瑞的方法^[10],新参数训练方法中还将前提属性候选值、前提属性权重、结果的评价等级作为训练参数,改进后的参数训练模型如图2所示。

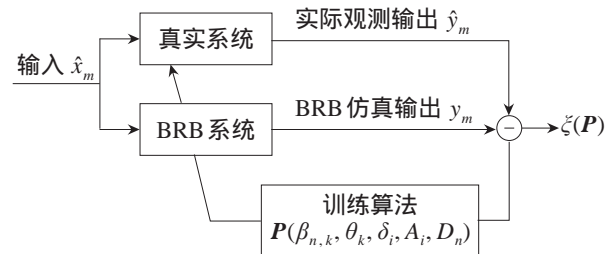


Fig.2 Modified parameter training model of BRB

图2 改进的BRB参数训练模型

图2中 P 是包含 $\beta_n, k, \theta_k, \delta_i, A_i, D_n$ 的训练参数向量。首先,对于新增训练参数需满足以下约束条件。

(1)假设 X 是一个有效输入,且落于区间 $[A_{i,k}, A_{i,k+1}]$,即前提属性候选值必须满足以下约束条件:

$$A_{i,k} - A_{i,k+1} \leq 0, i=1, 2, \dots, T, k=1, 2, \dots, L-1 \quad (15.1)$$

(2)前提属性权重 $\delta_{k,i}$ 经过标准化后应在0和1之间,当一条规则中只有一个前提属性时 $\delta_{k,i}$ 的值为1。

$$0 \leq \delta_{k,i} \leq 1, k=1, 2, \dots, N, i=1, 2, \dots, T \quad (15.2)$$

(3)评价等级 D_n 将输出结果表示为置信分布的形式, D_n 应满足:

$$D_n - D_{n+1} \leq 0, n=1, 2, \dots, N-1 \quad (15.3)$$

此时,待优化的目标函数(10.2)中包含的参数和约束条件分别为 $S_1 = N \times L + L + T + (L - 2) + (N - 2)$, $S_2 = N \times L + 2 \times L + T + (L - 1) + (N - 1)$, 本质上为求解一个含有 S_1 个参数,且带有 S_2 个约束条件的非线性优化问题。

新参数训练模型的目标是通过调整置信度 $\beta_{n,k}$ 、规则权重 θ_k 、前提属性权重 δ_k 、前提属性候选值 A_k 、评价等级 D_n 来极小化式(10.2)总的均方误差,使 BRB 系统更好地模拟真实系统。

首先,为了符合 BRB 训练模型中置信度的等式约束,需将式(8.2)转化为:

$$\beta_{N,k} = 1 - \sum_{n=1}^{N-1} \beta_{n,k}, k=1, 2, \dots, L \quad (16)$$

基于加速梯度求法的 BRB 参数训练方法利用式(16)来降低求解问题的维数,计算出使得目标函数更接近极小点的方向。使用梯度法确定搜索方向后,由式(8.1)、(8.3)、(15.1)~(15.3)的不等式约束可求得相应优化参数各自的最大步长,在此基础上利用二分搜索最优步长,直到目标函数达到最小迭代点。算法步骤如下:

步骤1 设置初始点 P^0 (即初始规则库矩阵),此时 $t=0$, 目标函数为式(10.2), $0 < \varepsilon < 1$ 。

步骤2 依据等式约束条件(8.2)对非线性优化问题进行降维处理,如式(16)所示,降维后减少了 L 个训练参数和 $2L$ 个约束条件。

步骤3 计算加速梯度求法的负梯度方向,利用不等式约束条件(8.1)、(8.3)、(15.1)~(15.3)求出各优化参数对应的最大步长,令 $\nabla \zeta(P^t)$ 为每一优化参数在初始点 P^0 的偏导值,计算 $g_k = \nabla \zeta(P)$ 。

$$\nabla \zeta(P) = \left(\frac{\partial \zeta}{\partial \beta_{11}}, \frac{\partial \zeta}{\partial \beta_{21}}, \dots, \frac{\partial \zeta}{\partial \beta_{(N-1)1}}, \frac{\partial \zeta}{\partial \beta_{12}}, \frac{\partial \zeta}{\partial \beta_{22}}, \dots, \frac{\partial \zeta}{\partial \beta_{(N-1)2}}, \dots, \frac{\partial \zeta}{\partial \beta_{1L}}, \frac{\partial \zeta}{\partial \beta_{2L}}, \dots, \frac{\partial \zeta}{\partial \beta_{(N-1)L}}, \frac{\partial \zeta}{\partial \theta_1}, \frac{\partial \zeta}{\partial \theta_2}, \dots, \frac{\partial \zeta}{\partial \theta_L}, \frac{\partial \zeta}{\partial A_2}, \frac{\partial \zeta}{\partial A_3}, \dots, \frac{\partial \zeta}{\partial A_{L-1}}, \frac{\partial \zeta}{\partial \delta_1}, \frac{\partial \zeta}{\partial \delta_2}, \dots, \frac{\partial \zeta}{\partial \delta_{|A_i|}}, \frac{\partial \zeta}{\partial D_2}, \frac{\partial \zeta}{\partial D_3}, \dots, \frac{\partial \zeta}{\partial D_{N-1}} \right)$$

偏导值利用式(14)求解。

$$\alpha = (\alpha_{11}^\beta, \alpha_{21}^\beta, \dots, \alpha_{(N-1)1}^\beta, \alpha_{12}^\beta, \alpha_{22}^\beta, \dots, \alpha_{(N-1)2}^\beta, \dots, \alpha_{1L}^\beta, \alpha_{2L}^\beta, \dots, \alpha_{(N-1)L}^\beta, \alpha_1^\theta, \alpha_2^\theta, \dots, \alpha_L^\theta, \alpha_2^A, \alpha_3^A, \dots, \alpha_{L-1}^A, \alpha_1^\delta, \alpha_2^\delta, \dots, \alpha_{|A_i|}^\delta, \alpha_2^D, \alpha_3^D, \dots, \alpha_{N-1}^D)$$

α 为步长向量。 A_1 、 A_L 和 D_1 、 D_N 为端点,不参与训练。

(1)若 $\nabla \zeta(P_i^t) > 0$, 根据梯度下降法有 $P_i^{t+1} = P_i^t - \alpha_i \times \nabla \zeta(P_i^t)$, 当 $\nabla \zeta(P_i^t)$ 为 A_2, A_3, \dots, A_{L-1} 或 D_2, D_3, \dots, D_{N-1} 的偏导结果时, $left = A_i - A_\varphi (A_\varphi < A_i - A_{i-1})$ 或 $left = D_i - D_\varphi (D_\varphi < D_i - D_{i-1})$, 否则 $left = 0$ 。根据优化模型要求 $P_i^{t+1} = P_i^t - \alpha_i \times \nabla \zeta(P_i^t)$ $left$ 则 $\alpha_i^{\max} = \frac{P_i^t - left}{\nabla \zeta(P_i^t)}$ 。

(2)若 $\nabla \zeta(P_i^t) = 0$, 此时对于该点的优化训练会陷入停滞。为了避免发生该情况,为偏导数加一个小小的摄动,即设 $\nabla \zeta(P_i^t) = \varepsilon$, 当 $\nabla \zeta(P_i^t)$ 为 A_2, A_3, \dots, A_{L-1} 或 D_2, D_3, \dots, D_{N-1} 的偏导结果时, $left = A_i - A_\varphi (A_\varphi < A_i - A_{i-1})$ 或 $left = D_i - D_\varphi (D_\varphi < D_i - D_{i-1})$, 否则 $left = 0$ 。而 $P_i^{t+1} = P_i^t - \alpha_i \times \nabla \zeta(P_i^t)$ $left$ 则 $\alpha_i^{\max} = \frac{P_i^t - left}{\nabla \zeta(P_i^t)}$ 。

(3)若 $\nabla \zeta(P_i^t) < 0$, 根据梯度下降法 $P_i^{t+1} = P_i^t - \alpha_i \times \nabla \zeta(P_i^t)$, 当 $\nabla \zeta(P_i^t)$ 为 A_2, A_3, \dots, A_{L-1} 或 D_2, D_3, \dots, D_{N-1} 的偏导结果时, $right = A_i + A_\varphi (A_\varphi < A_{i+1} - A_i)$ 或 $right = D_i + D_\varphi (D_\varphi < D_{i+1} - D_i)$, 否则 $right = 1$ 。根据优化模型要求 $P_i^{t+1} = P_i^t - \alpha_i \times \nabla \zeta(P_i^t)$ $right$ 则 $\alpha_i^{\max} = \frac{P_i^t - right}{\nabla \zeta(P_i^t)}$ 。

步骤4 通过对步骤3中求出的训练参数 $\beta_{n,k} (n=1, 2, \dots, N-1)$ 对应的最大步长不断进行二分,使其满足式(8.1)的不等式约束条件。而 $\beta_{N,k} (k=1, 2, \dots, L)$ 可通过式(16)求得。由式(8.1)可知对于新迭代点的 $\beta_{N,k}^{t+1}$ 应满足 $0 < \beta_{N,k}^{t+1} < 1$, 则要求 $0 < \sum_{n=1}^{N-1} \beta_{n,k}^{t+1} < 1$ 成立。假设对 $\beta_{n,k} (n=1, 2, \dots, N-1)$ 的最大步长进行 n 次二分后产生新的迭代点使得 $0 < \sum_{n=1}^{N-1} \beta_{n,k}^{t+1} < 1$ 成立, 即有 $0 < \sum_{n=1}^{N-1} \beta_{n,k}^t - \frac{1}{2^n} \sum_{n=1}^{N-1} \nabla \zeta(\beta_{n,k}^t) \times \alpha_{n,k}^{\max} < 1$ 。

(1) 若 $\sum_{n=1}^{N-1} \nabla \zeta(\beta_{n,k}^t) \times \alpha_{n,k}^{\max} > 0$ 且 $\sum_{n=1}^{N-1} \beta_{n,k}^t \neq 0$, 则 $n = \left\lfloor \lg \left(\frac{\sum_{n=1}^{N-1} \nabla \zeta(\beta_{n,k}^t) \times \alpha_{n,k}^{\max}}{\sum_{n=1}^{N-1} \beta_{n,k}^t} \right) \right\rfloor + 1$, 即至少要经过 n 次二分才能实现 $0 < \beta_{N,k}^{t+1} < 1$ 。

(2) 若 $\sum_{n=1}^{N-1} \nabla \zeta(\beta_{n,k}^t) \times \alpha_{n,k}^{\max} > 0$ 且 $\sum_{n=1}^{N-1} \beta_{n,k}^t = 0$ 令 $\alpha_{n,k}^{\max} = 0$ ($n = 1, 2, \dots, N-1, k = 1, 2, \dots, L$)。

(3) 若 $\sum_{n=1}^{N-1} \nabla \zeta(\beta_{n,k}^t) \times \alpha_{n,k}^{\max} = 0$, 以下两式等于正无穷:

$$\left\lfloor \lg \left(\frac{\sum_{n=1}^{N-1} \nabla \zeta(\beta_{n,k}^t) \times \alpha_{n,k}^{\max}}{\sum_{n=1}^{N-1} \beta_{n,k}^t} \right) \right\rfloor + 1$$

或 $\left\lfloor \lg \left(\frac{\sum_{n=1}^{N-1} \nabla \zeta(\beta_{n,k}^t) \times \alpha_{n,k}^{\max}}{\left(\sum_{n=1}^{N-1} \beta_{n,k}^t - 1 \right)} \right) \right\rfloor + 1$

令 $\alpha_{n,k}^{\max} = 0$ ($n = 1, 2, \dots, N-1, k = 1, 2, \dots, L$)。

(4) 若 $\sum_{n=1}^{N-1} \nabla \zeta(\beta_{n,k}^t) \times \alpha_{n,k}^{\max} < 0$ 且 $\sum_{n=1}^{N-1} \beta_{n,k}^t \neq 1$, 则 $n = \left\lfloor \lg \left(\frac{\sum_{n=1}^{N-1} \nabla \zeta(\beta_{n,k}^t) \times \alpha_{n,k}^{\max}}{\left(\sum_{n=1}^{N-1} \beta_{n,k}^t - 1 \right)} \right) \right\rfloor + 1$, 即至少要经过 n 次二分才能实现 $0 < \beta_{N,k}^{t+1} < 1$ 。

(5) 若 $\sum_{n=1}^{N-1} \nabla \zeta(\beta_{n,k}^t) \times \alpha_{n,k}^{\max} < 0$ 且 $\sum_{n=1}^{N-1} \beta_{n,k}^t = 1$ 令 $\alpha_{n,k}^{\max} = 0$ ($n = 1, 2, \dots, N-1, k = 1, 2, \dots, L$)。

取方向 $d_k = -g_k$, 经过上述处理求得最新的步长向量 α , 则 $P^{t+1} = P^t + \alpha d_k$, 设新的迭代点为 P_4^t 。

步骤5 判断步骤4中新迭代点 P_4^t 是否使目标函数(10.2)下降。若 $\zeta(P_4^t) < \zeta(P^t)$, 则转到步骤6, 否则在原来的基础上对步长不断进行二分, 直到求出新的迭代点 P_5^t , 使得 $\zeta(P_5^t) < \zeta(P^t)$ 成立。

步骤6 对步长再进行一次二分, 设产生的新的迭代点为 P_6^t , 如果 $\zeta(P_6^t) < \zeta(P_5^t)$, 则 $P^{t+1} = P_6^t$, 否则 $P^{t+1} = P_5^t$ 。

步骤7 判断终止条件。如果前后迭代误差 $|\zeta(P^{t+1}) - \zeta(P^t)| < \varepsilon$, 且 $\left| \zeta\left(\frac{P^{t+1} + P^t}{2}\right) - \zeta(P^t) \right| < \varepsilon$, 同时为使训练结果达到使得拟合效果较优的最优解, 避免欠拟合或过拟合, 还应满足 $\zeta(P^{t+1}) < \gamma$ ($0 < \gamma < 1$,

根据问题精度设置) 且 $\zeta\left(\frac{P^{t+1} + P^t}{2}\right) < \gamma$, 结束循环得到最优解 P^{t+1} ; 否则 $t = t + 1$, 返回步骤3。

3.3 收敛性分析

取 $0 < \rho < 1$, 参数取值集合为 B , 由 $d_k = -g_k$ ($k = 1$) 可知:

引理1 $-g_k^T d_k = (1 - \rho) \|g_k\|^2$ 。

引理2 $\|d_k\| = \frac{1}{1 - \rho} \|g_k\|$ 。

为下面证明需要, 本文作如下假设:

H1 目标函数 $\zeta(P)$ 在集合 B 上存在最优解, 及 $\zeta(P)$ 有下界。

H2 梯度函数 $g(P) = \nabla \zeta(P)$ 在集合 B 内满足 Lipschitz 连续, 即存在常数 $l > 0$, 使

$$\|g(P) - g(P')\| \leq l \|P - P'\|, \forall P, P' \in B$$

引理3^[21] 若 H1 和 H2 成立, 则对任意 $\alpha > 0$ 及 $P_k + \alpha d_k \in B$, 有:

$$\zeta(P_k + \alpha d_k) - \zeta(P_k) \leq -\alpha g_k^T d_k + \frac{1}{2} \alpha^2 l \|d_k\|^2$$

定理1 若 H1 和 H2 成立, $\{P_k\}$ 是由本文算法产生的迭代点列, 则 $\lim_{k \rightarrow +\infty} \|g_k\| = 0$ 。

证明 令 $\hat{\alpha}_k = \frac{-g_k^T d_k}{l \|d_k\|^2}$, 则由引理3及精确线搜索

规则^[13]知:

$$\zeta_k - \zeta_{k+1} = \zeta(P_k) - \zeta(P_k + \hat{\alpha}_k d_k)$$

$$= -\hat{\alpha}_k g_k^T d_k - \frac{1}{2} \hat{\alpha}_k^2 l \|d_k\|^2 = \frac{1}{2l} \left(\frac{g_k^T d_k}{\|d_k\|} \right)^2 \quad (17)$$

由引理1、引理2和上式可知:

$$\zeta_k - \zeta_{k+1} = \frac{1}{2l} \frac{(1 - \rho)^2 \|g_k\|^4}{\|d_k\|^2} = \frac{(1 - \rho)^2 \|g_k\|^2}{2l(1 + \rho)^2} \quad (18)$$

因为 $\{P_k\}$ 单调不减且有下界, 故知 P_k 有极限。从而由式(17)知 $\lim_{k \rightarrow +\infty} \|g_k\| = 0$ 。□

4 实验结果与分析

为了验证基于加速梯度求法的 BRB 参数训练方法的有效性, 本文从收敛精度和收敛时间两个方面进行实验分析, 并分别在多峰函数和输油管道检漏

两个实例中将本文方法与传统的参数训练方法进行比较。实验环境为: Intel® Core i3-3210@3.20 GHz; 4 GB 内存; Windows 7 操作系统; 算法由 Visual C++ 6.0 编写。

4.1 多峰函数中的参数训练

为验证本文方法的寻优能力和收敛速度, 首先引入一个多峰函数^[22], 函数表达式如下所示:

$$g(x) = e^{-(x-2)^2} + 0.5e^{-(x+2)^2}, -5 \leq x \leq 5 \quad (19)$$

函数的曲线如图3所示。

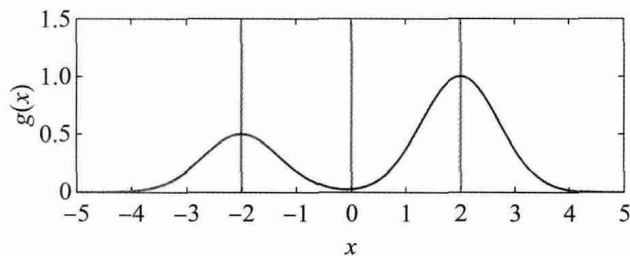


Fig.3 Graph of multimodal function

图3 多峰函数的曲线图

由图3可知, 当变量 x 取值为 -2 、 0 、 2 时为多极值函数的极值点, 而由各个极值点的函数值可设定规则结果集的评级等级和等级效用值:

$$\{D_1, D_2, D_3, D_4, D_5\} = \{-0.5, 0, 0.5, 1.0, 1.5\} \quad (20)$$

在 BRB 的构建中, 首先提取多峰函数的变量 x 作为前提属性。图3中定义域内有4个单调区间, 因此假定极值点对应的 x 取值作为前提属性候选值, 其分别为 $\{-5, -2, 0, 2, 5\}$ 。然后将前提属性候选值代入多极值函数求得函数值, 再依据基于规则的信息转换技术^[2]初始化规则结果集的分布式置信度。其中由5个候选值构建的 BRB 如表1所示。

使用本文所提的 BRB 参数训练方法对表1中的 BRB 进行参数训练。在参数训练过程中, 通过在定义域区间 $[-5, 5]$ 上均匀选值的方式选取1000个值作为训练数据, 同时还规定在求偏导时 $h = 1 \times 10^{-13}$ 。步骤3对前提属性 A_i 和评价等级 D_i 进行约束时 $A_\varphi = 1.65$, $D_\varphi = 0.45$ 。步骤7判断终止条件时 $\varepsilon = 1 \times 10^{-6}$, $\gamma = 5.16 \times 10^{-5}$ 。经参数训练后所得的 BRB 如表2所示。

为直观反映经本文方法训练后的 BRB 系统具有更高的推理准确性, 以下选取 FMINCON 函数的训练结果和常瑞方法^[10]的训练结果作为比较对象。各个 BRB 系统拟合多峰函数曲线的效果如图4所示。

由图4可以看出, 因常瑞方法只对部分参数进行训练, 其在 $[0, 1]$ 和 $[1, 2]$ 区间上的拟合效果不够理想。

Table 1 Initial BRB of multimodal function

表1 多峰函数初始 BRB

编号	规则权重	候选值	$g(x)$	结果集 $\{D_1, D_2, D_3, D_4, D_5\} = \{-0.5, 0, 0.5, 1.0, 1.5\}$
1	1	-5	0.000 1	$\{(D_1, 0), (D_2, 0.999\ 9), (D_3, 0.000\ 1), (D_4, 0), (D_5, 0)\}$
2	1	-2	0.500 0	$\{(D_1, 0), (D_2, 0), (D_3, 1.000\ 0), (D_4, 0), (D_5, 0)\}$
3	1	0	0.027 5	$\{(D_1, 0), (D_2, 0.950\ 0), (D_3, 0.050\ 0), (D_4, 0), (D_5, 0)\}$
4	1	2	1.000 0	$\{(D_1, 0), (D_2, 0), (D_3, 0), (D_4, 1.000\ 0), (D_5, 0)\}$
5	1	5	0.000 1	$\{(D_1, 0), (D_2, 0.999\ 8), (D_3, 0.000\ 2), (D_4, 0), (D_5, 0)\}$

Table 2 Trained BRB of multimodal function

表2 多峰函数训练后的 BRB

编号	规则权重	候选值	结果集 $\{D_1, D_2, D_3, D_4, D_5\} = \{-0.500\ 0, 0.001\ 5, 0.478\ 5, 1.059\ 3, 1.500\ 0\}$
1	0.981 2	-5.000 0	$\{(D_1, 0), (D_2, 0.999\ 9), (D_3, 0.000\ 1), (D_4, 0), (D_5, 0)\}$
2	0.350 1	-2.063 7	$\{(D_1, 0.008\ 2), (D_2, 0.050\ 6), (D_3, 0.867\ 3), (D_4, 0.059\ 0), (D_5, 0.014\ 9)\}$
3	0.425 4	-0.113 7	$\{(D_1, 0), (D_2, 0.950\ 0), (D_3, 0.050\ 0), (D_4, 0), (D_5, 0)\}$
4	0.311 9	2.063 5	$\{(D_1, 0.011\ 1), (D_2, 0.120\ 2), (D_3, 0.005\ 7), (D_4, 0.661\ 7), (D_5, 0.201\ 3)\}$
5	0.843 9	5.000 0	$\{(D_1, 0.066\ 5), (D_2, 0.919\ 2), (D_3, 0.000\ 2), (D_4, 0), (D_5, 0.014\ 1)\}$

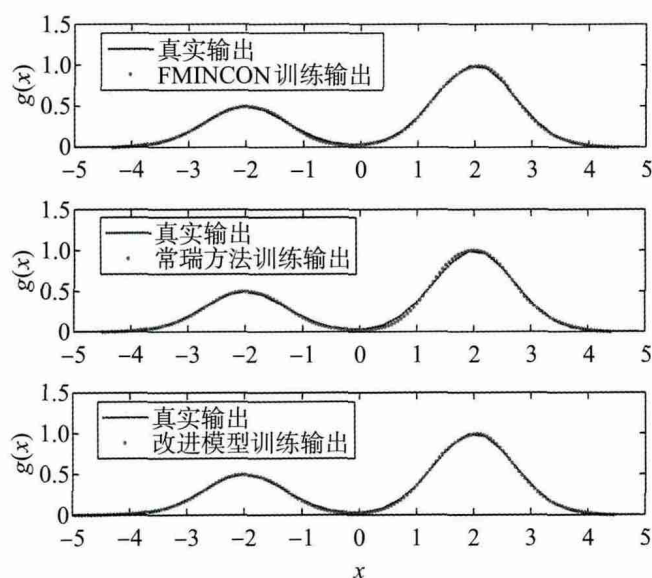


Fig.4 Comparison on BRB output with different training methods

图4 不同训练方法下的BRB输出比较

为进一步说明改进优化模型下本文方法在收敛精度和收敛速度上的优势,绘制传统FMINCON函数方法、常瑞方法和本文方法的收敛精度随时间变化的曲线,如图5所示。

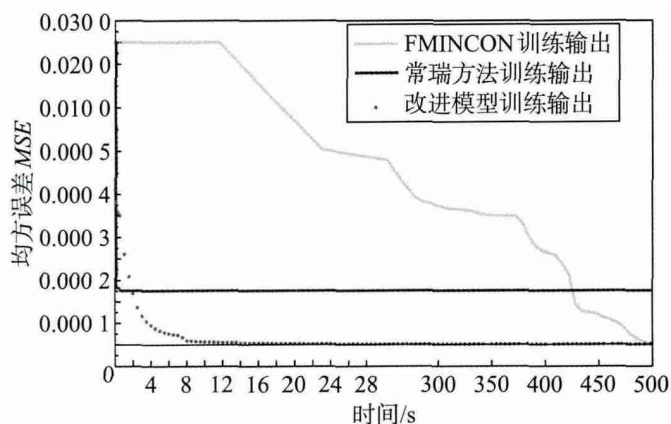


Fig.5 Graph with time on convergence accuracy of different methods

图5 不同方法的收敛精度随时间变化的曲线

由图5可知,本文方法和传统FMINCON函数方法的收敛精度最高,常瑞方法由于训练参数较少,减弱了BRB系统的推理能力,收敛精度只能达到 1.7×10^{-4} 。再比较收敛所需时间,相对传统使用FMINCON函数进行参数训练^[9],在收敛精度同时为 $MSE=5.16 \times 10^{-5}$ 时,FMINCON函数所需时间约为

500 s,而本文方法只需约28 s,算法在速度上有很大的提升,且算法用C语言,实现简单,程序有很强的移植性。综上可知,相比现有参数训练方法,本文方法在收敛精度和收敛速度上具有更理想的综合效益。

4.2 输油管道检漏中的参数训练

为验证本文方法同样适用于具有动态特性的BRB系统,引入输油管道检漏的BRB系统。输油管道检漏问题中,当输油管道某处发生泄漏时,管道的泄漏大小(LS)由油液流量(FD)和压力(PD)变化决定,因此将 FD 和 PD 作为BRB系统的输入, LS 作为输出。为构建BRB系统,给出如下前提条件:

以 FD 和 PD 为前提属性, FD 选取8个候选值,即 $FD \in \{-10, -5, -3, -1, 0, 1, 2, 3\}$; PD 选取7个候选值,即 $PD \in \{-0.042, -0.025, -0.010, 0, 0.010, 0.025, 0.042\}$; 对于输出 LS ,选取5个评价等级,分别为零、很小、中、高、很高,转换为定量等级后为 $D=\{0, 2, 4, 6, 8\}$ 。

由 FD (8个候选值)和 PD (7个候选值)可知,在BRB中包含56条规则。在泄漏为25%时,从2 008组数据中随机选取500组训练数据,然后利用本文提出的BRB参数训练方法对BRB系统进行训练。

训练后的BRB如表3所示,其中 FD 权重为0.680, PD 权重为0.333。用2 008组数据来测试BRB训练的推理性能,结果如图6所示。

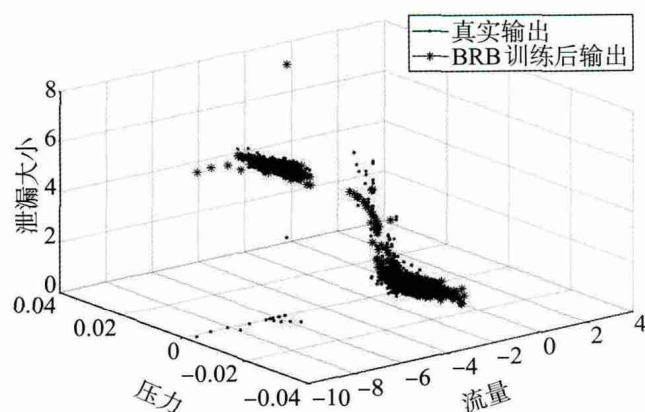


Fig.6 Actual system output and BRB output after training

图6 真实系统输出和BRB训练后输出

从图6中可以看出,训练后BRB系统的输出和真实系统的输出匹配度较高,能够真实反映真实系统的动态特性。

Table 3 Trained BRB of oil pipeline

表3 输油管道训练后的BRB

编号	规则权重	流量(FD)	压力(PD)	结果集 $D=\{0, 2.851, 5.180, 6.817, 8.000\}$
1	0	-10.000	-0.042	$\{(D_1, 0), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 1.000)\}$
2	0.013	-10.000	-0.031	$\{(D_1, 0), (D_2, 0), (D_3, 0), (D_4, 0.010), (D_5, 0.990)\}$
3	0.737	-10.000	-0.015	$\{(D_1, 0.001), (D_2, 0.004), (D_3, 0), (D_4, 0.984), (D_5, 0.011)\}$
4	0.053	-10.000	-0.002	$\{(D_1, 0), (D_2, 0), (D_3, 0.535), (D_4, 0.462), (D_5, 0.003)\}$
5	0.914	-10.000	0.009	$\{(D_1, 0.047), (D_2, 0.028), (D_3, 0.004), (D_4, 0.749), (D_5, 0.173)\}$
6	0.261	-10.000	0.023	$\{(D_1, 0.187), (D_2, 0.149), (D_3, 0.172), (D_4, 0.031), (D_5, 0.461)\}$
7	0	-10.000	0.042	$\{(D_1, 0.030), (D_2, 0.002), (D_3, 0), (D_4, 0), (D_5, 0.968)\}$
8	0	-5.108	-0.042	$\{(D_1, 0), (D_2, 0), (D_3, 0.003), (D_4, 0.029), (D_5, 0.968)\}$
9	0.013	-5.108	-0.031	$\{(D_1, 0.017), (D_2, 0.017), (D_3, 0.056), (D_4, 0.017), (D_5, 0.894)\}$
10	0.606	-5.108	-0.015	$\{(D_1, 0.055), (D_2, 0.190), (D_3, 0.113), (D_4, 0.612), (D_5, 0.030)\}$
11	0.146	-5.108	-0.002	$\{(D_1, 0.052), (D_2, 0.125), (D_3, 0.040), (D_4, 0.687), (D_5, 0.096)\}$
12	0.224	-5.108	0.009	$\{(D_1, 0.171), (D_2, 0.117), (D_3, 0.073), (D_4, 0.385), (D_5, 0.254)\}$
13	0.290	-5.108	0.023	$\{(D_1, 0.190), (D_2, 0.146), (D_3, 0.172), (D_4, 0.031), (D_5, 0.461)\}$
14	0	-5.108	0.042	$\{(D_1, 0.032), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0.968)\}$
15	0	-2.719	-0.042	$\{(D_1, 0), (D_2, 0), (D_3, 0.013), (D_4, 0.019), (D_5, 0.968)\}$
16	0	-2.719	-0.031	$\{(D_1, 0), (D_2, 0), (D_3, 0.026), (D_4, 0.006), (D_5, 0.968)\}$
17	0.585	-2.719	-0.015	$\{(D_1, 0.059), (D_2, 0.156), (D_3, 0.427), (D_4, 0.266), (D_5, 0.092)\}$
18	0.113	-2.719	-0.002	$\{(D_1, 0.100), (D_2, 0.700), (D_3, 0.200), (D_4, 0), (D_5, 0)\}$
19	1.000	-2.719	0.009	$\{(D_1, 0.241), (D_2, 0.157), (D_3, 0.290), (D_4, 0.290), (D_5, 0.022)\}$
20	0	-2.719	0.023	$\{(D_1, 0.029), (D_2, 0.003), (D_3, 0), (D_4, 0), (D_5, 0.968)\}$
21	0	-2.719	0.042	$\{(D_1, 0.032), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0.968)\}$
22	0	-1.151	-0.042	$\{(D_1, 0.001), (D_2, 0.004), (D_3, 0.012), (D_4, 0.015), (D_5, 0.968)\}$
23	0	-1.151	-0.031	$\{(D_1, 0.006), (D_2, 0.050), (D_3, 0.008), (D_4, 0), (D_5, 0.936)\}$
24	0.013	-1.151	-0.015	$\{(D_1, 0.524), (D_2, 0.041), (D_3, 0), (D_4, 0), (D_5, 0.435)\}$
25	0.158	-1.151	-0.002	$\{(D_1, 0.840), (D_2, 0.003), (D_3, 0.004), (D_4, 0.004), (D_5, 0.150)\}$
26	0.163	-1.151	0.009	$\{(D_1, 0.452), (D_2, 0.097), (D_3, 0.093), (D_4, 0.132), (D_5, 0.227)\}$
27	0	-1.151	0.023	$\{(D_1, 0.055), (D_2, 0.015), (D_3, 0.015), (D_4, 0.015), (D_5, 0.901)\}$
28	0.025	-1.151	0.042	$\{(D_1, 0.055), (D_2, 0.015), (D_3, 0.015), (D_4, 0.015), (D_5, 0.901)\}$
29	0.025	-0.281	-0.042	$\{(D_1, 0.055), (D_2, 0.015), (D_3, 0.015), (D_4, 0.015), (D_5, 0.901)\}$
30	0.173	-0.281	-0.031	$\{(D_1, 0.577), (D_2, 0.022), (D_3, 0.009), (D_4, 0.009), (D_5, 0.384)\}$
31	0.013	-0.281	-0.015	$\{(D_1, 0.954), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0.046)\}$
32	0.700	-0.281	-0.002	$\{(D_1, 0.998), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0.002)\}$
33	0	-0.281	0.009	$\{(D_1, 0.078), (D_2, 0.031), (D_3, 0.031), (D_4, 0.031), (D_5, 0.830)\}$
34	0	-0.281	0.023	$\{(D_1, 0.078), (D_2, 0.031), (D_3, 0.031), (D_4, 0.031), (D_5, 0.830)\}$
35	0.056	-0.281	0.042	$\{(D_1, 0.078), (D_2, 0.031), (D_3, 0.031), (D_4, 0.031), (D_5, 0.830)\}$
36	0.056	1.117	-0.042	$\{(D_1, 0.049), (D_2, 0.059), (D_3, 0.031), (D_4, 0.031), (D_5, 0.830)\}$
37	0.038	1.117	-0.031	$\{(D_1, 0.490), (D_2, 0.044), (D_3, 0.016), (D_4, 0.016), (D_5, 0.434)\}$
38	0.013	1.117	-0.015	$\{(D_1, 0.920), (D_2, 0.022), (D_3, 0), (D_4, 0), (D_5, 0.058)\}$
39	1.000	1.117	-0.002	$\{(D_1, 0.998), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0.002)\}$
40	0	1.117	0.009	$\{(D_1, 0.078), (D_2, 0.031), (D_3, 0.031), (D_4, 0.031), (D_5, 0.830)\}$
41	0	1.117	0.023	$\{(D_1, 0.078), (D_2, 0.031), (D_3, 0.031), (D_4, 0.031), (D_5, 0.830)\}$
42	0.056	1.117	0.042	$\{(D_1, 0.078), (D_2, 0.031), (D_3, 0.031), (D_4, 0.031), (D_5, 0.830)\}$
43	0.056	2.128	-0.042	$\{(D_1, 0.035), (D_2, 0.073), (D_3, 0.031), (D_4, 0.031), (D_5, 0.830)\}$
44	0.056	2.128	-0.031	$\{(D_1, 0.045), (D_2, 0.064), (D_3, 0.031), (D_4, 0.031), (D_5, 0.830)\}$
45	0.056	2.128	-0.015	$\{(D_1, 0.071), (D_2, 0.038), (D_3, 0.031), (D_4, 0.031), (D_5, 0.830)\}$
46	0.221	2.128	-0.002	$\{(D_1, 0.745), (D_2, 0.026), (D_3, 0.024), (D_4, 0.024), (D_5, 0.180)\}$
47	0.124	2.128	0.009	$\{(D_1, 0.704), (D_2, 0.031), (D_3, 0.031), (D_4, 0.031), (D_5, 0.204)\}$
48	0.056	2.128	0.023	$\{(D_1, 0.078), (D_2, 0.031), (D_3, 0.031), (D_4, 0.031), (D_5, 0.830)\}$
49	0.056	2.128	0.042	$\{(D_1, 0.078), (D_2, 0.031), (D_3, 0.031), (D_4, 0.031), (D_5, 0.830)\}$
50	0.056	3.000	-0.042	$\{(D_1, 0.073), (D_2, 0.035), (D_3, 0.031), (D_4, 0.031), (D_5, 0.830)\}$
51	0.056	3.000	-0.031	$\{(D_1, 0.077), (D_2, 0.031), (D_3, 0.031), (D_4, 0.031), (D_5, 0.830)\}$
52	0.056	3.000	-0.015	$\{(D_1, 0.078), (D_2, 0.031), (D_3, 0.031), (D_4, 0.031), (D_5, 0.830)\}$
53	0.056	3.000	-0.002	$\{(D_1, 0.078), (D_2, 0.031), (D_3, 0.031), (D_4, 0.031), (D_5, 0.830)\}$
54	0.056	3.000	0.009	$\{(D_1, 0.078), (D_2, 0.031), (D_3, 0.031), (D_4, 0.031), (D_5, 0.830)\}$
55	0.056	3.000	0.023	$\{(D_1, 0.078), (D_2, 0.031), (D_3, 0.031), (D_4, 0.031), (D_5, 0.830)\}$
56	0.056	3.000	0.042	$\{(D_1, 0.078), (D_2, 0.031), (D_3, 0.031), (D_4, 0.031), (D_5, 0.830)\}$

图7为训练后的BRB对真实系统的模拟效果。

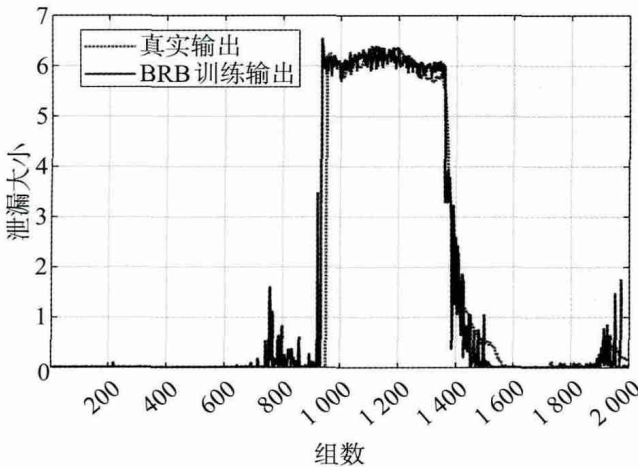


Fig.7 Comparison on actual system output and BRB output after training

图7 真实系统输出和BRB训练后输出的比较

由图7可知,采用本文方法对BRB训练后能够很好地拟合真实系统的输出,图中存在偏差较大的点是因为原始的泄漏数据中含有噪声,主要是由测量仪器和外界影响所造成的泄漏误报。

为比较本文方法与现有方法对BRB系统推理性能的影响,以平均绝对误差(mean absolute error, MAE)、均方根误差(root mean square error, RMSE)、皮尔森相关系数^[23](Pearson correlation coefficient, PCC)作为衡量指标,结果如表4所示。

Table 4 Comparison on reasoning performance of BRB with different training methods

表4 不同训练方法下BRB推理性能的比较

训练方法	MAE	RMSE	PCC	时间/s
Matlab中FMINCON函数	0.171 68	0.638 16	0.966 28	7 180
专家干预下差分进化算法	0.165 31	0.614 24	0.968 85	147
常瑞传统模型梯度下降法	0.223 26	0.745 88	0.954 76	57
改进模型的加速梯度求法	0.182 78	0.667 47	0.963 41	50

从表4中MAE和RMSE的值可知,本文方法的训练结果相对于常瑞方法具有更高的推理准确性。传统FMINCON函数方法^[9]的收敛精度虽然更高,但训练时间长,可移植性差。差分进化算法^[12]训练结果虽也较理想,但算法存在搜索停滞和早熟收敛的不足^[18]。由表4分析还可知,本文方法求得的PCC值比常瑞

方法所得的PCC值更大,说明其对观测输出的动态变化进行快速拟合时性能更好。同时本文方法收敛速度快,训练结果的MAE和RMSE更小。

5 结束语

本文针对传统参数训练方法中存在的算法复杂度高、收敛精度差及收敛速度慢等瓶颈问题,在现有梯度法的参数训练方法的基础上,改进梯度法中求解负梯度方向的过程,并新增前提属性候选值、前提属性权重、结果评价等级作为BRB训练参数,从而提出了具有高收敛精度和高收敛效率的基于加速梯度求法的置信规则库参数训练方法。在实验分析中,通过在多峰函数和输油管道检漏的实例中进行参数训练,验证了本文方法的有效性。同时还对比了传统FMINCON函数方法和常瑞方法,实验结果表明本文方法在收敛精度和收敛速度上具有更理想的综合效益。在将来的研究工作中,将进一步针对多参数情形下的参数训练进行实验研究,以期设计出更合理更有效的参数训练方法。

References:

- [1] Sun R. Robust reasoning: integration rule-based and similarity-based reasoning[J]. Artificial Intelligence, 1995, 75(2): 241-295.
- [2] Yang Jianbo, Liu Jun, Wang Jin, et al. Belief rule-base inference methodology using the evidential reasoning approach RIMER[J]. IEEE Transactions on Systems, Man, and Cybernetics: Part A Systems and Humans, 2006, 36(2): 266-285.
- [3] Dempster A P. A generalization of Bayesian inference[J]. Journal of the Royal Statistical Society: Series B Methodological, 1968, 30(2): 205-247.
- [4] Shafer G. A mathematical theory of evidence[M]. Princeton, USA: Princeton University Press, 1976.
- [5] Hwang C, Yoon K. Methods for multiple attribute decision making[J]. Multiple Attribute Decision Making Lecture Notes in Economics and Mathematical Systems, 1981, 186: 58-191.
- [6] Zadeh L Z. Fuzzy sets[J]. Information and Control, 1965, 8(3): 338-353.
- [7] Jiang Jiang, Li Xuan, Zhou Zhijie, et al. Weapon system capability assessment under uncertainty based on the evidential reasoning approach[J]. Expert Systems with Applications,

- 2011, 38(11): 13773-13784.
- [8] Zhou Zhijie, Hu Changhua, Yang Jianbo, et al. Online updating belief-rule-based system for pipeline leak detection under expert intervention[J]. Expert Systems with Applications, 2009, 36(4): 7700-7709.
- [9] Yang Jianbo, Liu Jun, Xu Dongling, et al. Optimization model for training belief-rule-based systems[J]. IEEE Transactions on Systems, Man, and Cybernetics: Part A Systems and Humans, 2007, 37(4): 569-585.
- [10] Chang Rui, Zhang Su. An algorithm for training parameters in belief rule-bases based on gradient methods with optimization step size[J]. North China Institute of Water Conservancy and Hydroelectric Power, 2011, 32(1): 154-157.
- [11] Lee B H, Kim S, Seok J W, et al. Nonlinear system identification based on support vector machine using particle swarm optimization[C]//Proceedings of the International Joint Conference on SICE-ICASE. Piscataway, NJ, USA: IEEE, 2006: 5614-5618.
- [12] Chang Leilei, Li Mengjun, Jiang Jiang, et al. Parameter learning for the belief rule base system in the residual life probability prediction of metalized film capacitor[J]. European Journal of Operational Research, submitted.
- [13] Ma Changfeng. Optimization method and Matlab programming[M]. Beijing: Science Press, 2010.
- [14] Wang Yingming, Yang Jianbo, Xu Dongling. Environmental impact assessment using the evidential reasoning approach[J]. European Journal of Operational Research, 2006, 174(3): 1885-1913.
- [15] Duan Xinsheng. Evidence theory, decision & artificial intelligence[M]. Beijing: Renmin University of China Press, 1993.
- [16] Wang Yingming, Yang Jianbo, Xu Dongling, et al. The evidential reasoning approach for multiple attribute decision analysis using interval belief degrees[J]. European Journal of Operational Research, 2006, 175(1): 35-66.
- [17] Chen Yuwang, Yang Jianbo, Xu Dongling, et al. Inference analysis and adaptive training for belief rule based systems[J]. Expert Systems with Applications, 2011, 38(10): 12845-12860.
- [18] Jia Dongli. Modified differential evolution algorithm and its application in communications signal processing[D]. Shanghai: Shanghai University, 2011.
- [19] Department of Mathematics, Tongji University. Advanced mathematics/Volume 1[M]. Beijing: Higher Education Press, 2007.
- [20] Wang Moran. MATLAB and scientific computing[M]. Beijing: Electronic Industry Press, 2003.
- [21] Shi Zhenjun. A new super-memory gradient method for unconstrained optimization[J]. Advances in Mathematics, 2006, 3(35): 265-274.
- [22] Chen Yuwang, Yang Jianbo, Xu Dongling, et al. On the inference and approximation properties of belief rule based systems[J]. Information Sciences, 2013, 234: 121-135.
- [23] Hao Liren. SPSS practical statistics[M]. Beijing: China Water & Power Press, 2003.

附中文参考文献：

- [10] 常瑞, 张速. 基于优化步长和梯度法的置信规则库参数学习方法[J]. 华北水利水电学院学报, 2011, 32(1): 154-157.
- [13] 马昌凤. 最优化方法及其 Matlab 程序设计[M]. 北京: 科学出版社, 2010.
- [15] 段新生. 证据理论与决策、人工智能[M]. 北京: 中国人民大学出版社, 1993.
- [18] 贾东立. 改进的差分进化算法及其在通信信号处理中的应用研究[D]. 上海: 上海大学, 2011.
- [19] 同济大学数学系. 高等数学/上册[M]. 北京: 高等教育出版社, 2007.
- [20] 王沫然. MATLAB 与科学计算[M]. 北京: 电子工业出版社, 2003.
- [23] 郝黎仁. SPSS 实用统计分析[M]. 北京: 中国水利水电出版社, 2003.



WU Weikun was born in 1991. He is a student at College of Mathematics and Computer Science, Fuzhou University. His research interests include intelligent decision technology and data mining, etc.

吴伟昆(1991)男, 福建泉州人, 福州大学数学与计算机科学学院学生, 主要研究领域为智能决策技术, 数据挖掘等。



YANG Longhao was born in 1990. He is a master candidate at College of Mathematics and Computer Science, Fuzhou University. His research interests include intelligent decision-making technology and belief rule base inference, etc.

杨隆浩(1990) ,男 ,福建南平人 ,福州大学数学与计算机科学学院硕士研究生 ,主要研究领域为智能决策技术 ,置信规则库推理等。



FU Yanggeng was born in 1981. He received the Ph.D. degree from Fuzhou University in 2013. Now he is a lecturer at College of Mathematics and Computer Science, Fuzhou University, and the member of CCF. His research interests include multi-criteria decision making under uncertainty, belief rule base inference and mobile Internet applications, etc.

傅仰耿(1981) ,男 ,福建泉州人 ,2013年于福州大学获得博士学位 ,现为福州大学数学与计算机科学学院讲师 ,CCF会员 ,主要研究领域为不确定多准则决策 ,置信规则库推理 ,移动互联网应用等。



ZHANG Liqun was born in 1991. He is a student at College of Mathematics and Computer Science, Fuzhou University. His research interests include intelligent decision technology and data mining, etc.

张立群(1991) ,男 ,福建泉州人 ,福州大学数学与计算机科学学院学生 ,主要研究领域为智能决策技术 ,数据挖掘等。



GONG Xiaoting was born in 1982. She received the M.S degree from Fuzhou University in 2006. Now she is a lecturer at College of Economics and Management, Fuzhou University. Her research interests include multi-criteria decision making under uncertainty and information hiding technology, etc.

巩晓婷(1982) ,女 ,河南漯河人 ,2006年于福州大学获得硕士学位 ,现为福州大学经济与管理学院讲师 ,主要研究领域为不确定多准则决策 ,信息隐藏技术等。