·综合研究·

一种改进的LS-SVM 算法及其应用

董 瑶 *　潘国锋　夏克文　张志伟

（河北工业大学信息工程学院）

董瑶,潘国峰,夏克文,张志伟. 一种改进的 LS-SVM 算法及其应用. 石油地球物理勘探,2007,42(6):673～677

摘要　为了避免 LS-SVM 算法中存在的矩阵求逆问题,提出一种改进的 LS-SVM 算法,即利用改进 PSO 算法对 LS-SVM 算法中线性方程组进行迭代优化求解,这样既能加快算法训练速度和节省内存,又总能得到最小二乘解,提高计算精度。将此改进算法应用到长庆气田 C 井目的层井段进行气层识别,并与 BP 神经网络算法、经典的 SVM 算法和传统的 LS-SVM 算法比较,结果表明此算法识别精度高,收敛速度快,与试气结果吻合,效果显著。

关键词　最小二乘支持向量机　粒子群优化算法　迭代优化　气层识别

1　引言

随着计算智能技术的迅猛发展,神经计算、进化计算和粒计算等计算智能技术在许多领域得到成功应用,并已成为气层识别的一项主要技术。比如在小样本情形下,基于结构风险最小化(SRM)准则[1]的支持向量机(SVM)能够有效地避免经典学习方法中存在的过学习、维数灾难、局部极小等问题[2],且具有良好的泛化能力,在各种分类问题中已得到成功应用[3],但对于实际大规模问题则存在训练速度较慢的缺点。为此,Suyken 提出了最小二乘支持向量机(Least Squares Support Vector Machines,LS-SVM)[4],把二次优化问题转化为一个线性方程组的求解问题,克服了 SVM 的缺陷。但 LS-SVM 算法在求解过程中总会出现矩阵的求逆,对于实际工程的大规模问题在微机上是难以实现的。为此,基于迭代计算思想,本文提出了一种新的求解方法,即采用再现群智能的粒子群优化算法(Particle Swarm Optimization,PSO)来求解 LS-SVM 中的任意线性方程组,不仅加快了计算速度,而且避免了矩阵的求逆,节省了内存。

将本文提出的改进 LS-SVM 算法应用到气层识别领域,克服了以往常规线性、经验性测井解释技术[5]难以适应油井复杂的环境变化且识别准确率低的缺点,提高了识别准确率及计算速度。

* 天津市北辰区河北工业大学信息工程学院 390 信箱,300401
本文于 2007 年 1 月 29 日收到,修改稿于同年 2 月 8 日收到。
本研究为国家自然科学基金项目(60377020,60673087)。

2　基本 LS-SVM 及其改进算法

2.1　基本 LS-SVM 原理

设训练集 $\boldsymbol{S}=\{(x_k,y_k)\mid k=1,2,\cdots,N\}$,其中 $x_k\in\boldsymbol{R}^n$,$y_k\in\boldsymbol{R}$,分别为输入和输出数据。与经典 SVM 不同,LS-SVM 利用 SRM 准则构造如下最小化目标函数及其约束条件

$$\begin{cases}\min\limits_{w,b,e} J(\boldsymbol{w},e)=\dfrac{1}{2}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w}+\gamma\dfrac{1}{2}\sum\limits_{k=1}^{N}e_k^2\\ y_k=\boldsymbol{w}^{\mathrm{T}}\phi(x_k)+b+e_k\quad k=1,2,\cdots,N\end{cases}\tag{1}$$

其中:$\boldsymbol{w}$ 为权值;γ 为正规化参数;e_k 为误差变量;$b\in\boldsymbol{R}$ 为偏置参数。将求解式(1)的优化问题转化为求解如下线性方程组

$$\begin{bmatrix}0 & \boldsymbol{I}^{\mathrm{T}}\\ \boldsymbol{I} & K+\gamma^{-1}\boldsymbol{I}\end{bmatrix}\begin{bmatrix}b\\ \boldsymbol{a}\end{bmatrix}=\begin{bmatrix}0\\ \boldsymbol{Y}\end{bmatrix}\tag{2}$$

其中:$\boldsymbol{I}=[1,\cdots,1]^{\mathrm{T}}$;$K=\phi(x)^{\mathrm{T}}\phi(x_k)$是满足 Mercer 条件的核函数;$\boldsymbol{Y}=[y_1,\cdots,y_N]^{\mathrm{T}}$;$\boldsymbol{a}=[a_1,\cdots,a_N]^{\mathrm{T}}$;$\boldsymbol{I}$ 为单位矩阵。

用最小二乘法求出方程(2)中的 $\boldsymbol{a}$ 和 b,由此得到预测输出为

$$y_k=\sum_{k=1}^{l}a_kK(x,x_k)+b\tag{3}$$

返回目录

2.2 改进的 LS-SVM

我们将线性方程(2)改写为如下矩阵方程

$$\boldsymbol{AX}=\boldsymbol{z}\ (\boldsymbol{A}\in\boldsymbol{R}^{m\times n},\boldsymbol{z}\in\boldsymbol{R}^{m}) \tag{4}$$

在 LS-SVM 中,一般是采用最小二乘法求解。而对于实际工程的大规模问题,由于 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的维数较大,在微机上难以实现矩阵的求逆过程。为此,可以采用迭代计算的方法来求解矩阵方程。

用于解决优化问题的 PSO 算法是一种迭代寻优的计算方法。因此对于大型矩阵方程的求解,可以采用 PSO 算法来进行迭代求解,即将矩阵方程(4)转化为采用 PSO 算法来迭代求解最优粒子 $\boldsymbol{X}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 的问题,从而避免了矩阵求逆的过程。

基本 PSO 算法[6]是由 Kennedy 和 Eberhart 最早提出的,粒子通过跟踪两个"极值"来更新自己的速度 v_i 和位置 x_i,即局部极值 p_{best} 和全局极值 g_{best}。迭代更新公式为

$$v_{i+1}=v_i+\varphi_1\cdot r_1\cdot(p_{\mathrm{best}}-x_i)+\\ +\varphi_2\cdot r_2\cdot(g_{\mathrm{best}}-x_i) \tag{5}$$

$$x_{i+1}=x_i+v_{i+1} \tag{6}$$

式中:φ_1,φ_2 为加速常数;r_1,r_2 为随机参量。

改进 PSO 算法[7]即在基本 PSO 算法模型的基础上增加了收敛因子 χ、惯性权重 ω 和约束因子 α,保证了算法的收敛性,提高了收敛速度。粒子迭代公式为

$$v_{i+1}=\chi\cdot[\omega\cdot v_i+\varphi_1\cdot r_1\cdot(p_{\mathrm{best}}-x_i)+\\ +\varphi_2\cdot r_2\cdot(g_{\mathrm{best}}-x_i)] \tag{7}$$

$$x_{i+1}=x_i+\alpha\cdot v_{i+1} \tag{8}$$

为了验证上述改进 PSO 算法求解矩阵方程的有效性,下面用改进 PSO 算法和基本 PSO 算法分别求解在面波频散反演中常见的一个典型病态方程组[8]

$$\boldsymbol{AX}=\boldsymbol{z}\ (\boldsymbol{A}\in\boldsymbol{R}^{m\times n},\boldsymbol{z}\in\boldsymbol{R}^{m}) \tag{9}$$

其中

$$\boldsymbol{A}=\begin{bmatrix}
0.901\mathrm{D}+00 & -0.705\mathrm{D}-03 & 0.122\mathrm{D}-11 & 0.733\mathrm{D}-16 & 0.000\mathrm{D}+00 & 0.000\mathrm{D}+00\\
0.716\mathrm{D}+00 & -0.146\mathrm{D}-02 & -0.454\mathrm{D}-10 & 0.400\mathrm{D}-10 & 0.000\mathrm{D}+00 & 0.000\mathrm{D}+00\\
0.542\mathrm{D}+00 & -0.183\mathrm{D}-02 & -0.119\mathrm{D}-10 & -0.388\mathrm{D}-14 & 0.000\mathrm{D}+00 & 0.000\mathrm{D}+00\\
0.392\mathrm{D}+00 & 0.113\mathrm{D}-03 & -0.179\mathrm{D}-09 & -0.240\mathrm{D}-12 & 0.715\mathrm{D}-16 & -0.535\mathrm{D}-16\\
0.263\mathrm{D}+00 & 0.377\mathrm{D}-02 & -0.861\mathrm{D}-08 & -0.595\mathrm{D}-09 & -0.464\mathrm{D}-13 & -0.951\mathrm{D}-15\\
0.129\mathrm{D}+00 & 0.160\mathrm{D}-02 & 0.159\mathrm{D}-07 & -0.676\mathrm{D}-08 & -0.235\mathrm{D}-09 & -0.900\mathrm{D}-12\\
0.196\mathrm{D}-01 & 0.287\mathrm{D}-03 & 0.706\mathrm{D}-06 & 0.447\mathrm{D}-06 & 0.149\mathrm{D}-02 & -0.363\mathrm{D}-02
\end{bmatrix}$$

$$\boldsymbol{z}=(0.180130\mathrm{D}+01\quad -0.143054\mathrm{D}+01\quad -0.108217\mathrm{D}+01\quad -0.784113\mathrm{D}+00\quad -0.529770\mathrm{D}+00\\ -0.259600\mathrm{D}+00\quad -0.473966\mathrm{D}-01)^{\mathrm{T}}$$

式(9)的真解为

$$\boldsymbol{x}=(-2.0\quad -1.0\quad 0.0\quad 1.0\quad 2.0\quad 3.0)^{\mathrm{T}}$$

$\boldsymbol{A}$ 的条件数高达 10^{13},可见式(9)是严重病态的。

采用基本 PSO 算法和改进 PSO 算法求解上述矩阵方程,参数设定如下:粒子数目 $N=10$,随机常数 $r_1=r_2=\mathrm{rand}(0,1)$;加速常数 $\varphi_1=\varphi_2=2.05$,$\varphi=\varphi_1+\varphi_2=4.1$;则收敛因子 $\chi=\dfrac{2}{\left|2-\varphi-\sqrt{\varphi^2-4\varphi}\right|}=0.729$。本文中将 ω 初始为 0.9 并使其随迭代次数的增加,依据下式

$$\omega=\omega_{\max}-\frac{(\omega_{\max}-\omega_{\min})T_{\max}}{T} \tag{10}$$

线性递减至 0.1,从而调整算法的搜索能力,以达到优化目的;约束因子 $\alpha=0.8$。两种算法训练网络的迭代误差曲线如图 1 所示。采用基本 PSO 算法求解迭代 247 次误差达到 7.0228×10^{-7},而改进 PSO 算法仅迭代 39 次就已达到这一误差精度,在迭代 108 次误差就已达到 8.9591×10^{-11},收敛速度及精度上均优于基本 PSO 算法。输出 $\boldsymbol{x}$ 的结果如表 1 所示。

表 1 两种算法的输出结果

期望输出 $\boldsymbol{x}$	基本 PSO 算法的输出	改进 PSO 算法的输出
−2	−2.0003	−2.0000
−1	−1.0202	−1.0000
0	0.2720	0.1441
1	1.4414	1.0174
2	1.1465	1.9999
3	2.6464	3.0000
均方差	1.87×10^{-1}	3.5×10^{-3}

返回目录

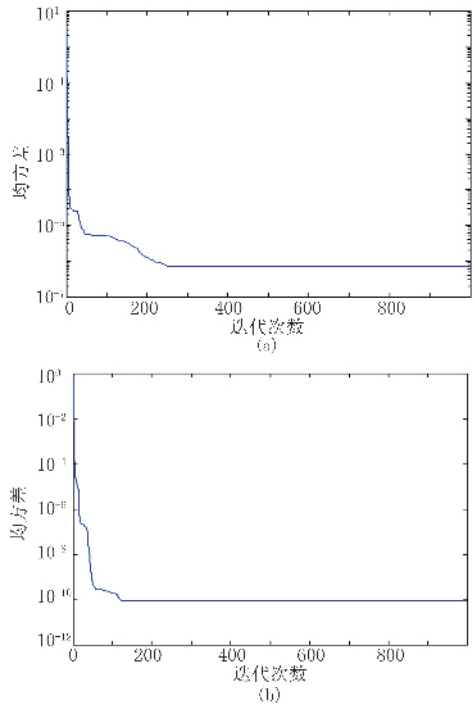

图 1　PSO 算法训练网络的误差演化曲线

(a)基本 PSO 算法;(b)改进 PSO 算法

由此可见,采用改进的 PSO 算法求解矩阵方程,与基本 PSO 算法相比收敛速度快,计算精度高,求解效果令人满意。为此提出基于 PSO 迭代优化的 LS-SVM 算法,其算法流程如下。

步骤 1　初始化粒子群。设定粒子群参数,在定义空间 $\boldsymbol{R}^n$ 中随机产生 n 个粒子 $x_1,x_2,\cdots,x_n$,组成初始种群 $\boldsymbol{X}(t)$;随机产生各粒子的初始速度 v_1,$v_2,\cdots,v_n$,组成速度矩阵 $\boldsymbol{V}(t)$;每个粒子的个体最优解 $p_{\text{best}i}$ 初始值为 x_i 的初始值。

步骤 2　评价各粒子适应度函数 $f(x)$。在具体矩阵方程求解中,可以按照残差($\boldsymbol{z}-\boldsymbol{Ax}$)的均方差来计算适应度函数,$f(x)$ 越小,适应能力越强。适应度函数可定义为

$$f(x)=\frac{1}{mn}\sum_{i=1}^{m}\sum_{i=1}^{n}\left(z_i-A_{ij}x_j\right)^2 \tag{11}$$

步骤 3　对每个粒子,比较当前适应度 $f(x_i)$ 和历史最好位置适应度 $f(p_{\text{best},i})$,如果 $f(x_i)<f(p_{\text{best},i})$,那么 $p_{\text{best},i}=x_i$,即可比较群体所有粒子当前适应度 $f(x_i)$ 和群体最好位置适应度 $f(g_{\text{best},i})$,如果 $f(x_i)<f(g_{\text{best},i})$,那么全局最优解 $g_{\text{best},i}=x_i$。

步骤 4　根据改进 PSO 算法式(7)和式(8)更新粒子的速度和位置,产生新种群 $\boldsymbol{X}(t+1)$,速度调整规则如下

$$\begin{cases} v_i=V_{\max} & v_i>V_{\max} \\ v_i=-V_{\max} & v_i\leqslant -V_{\max} \end{cases} \tag{12}$$

步骤 5　检查结束条件。若满足条件,则结束寻优,返回当前最优个体为结果,否则 $T=T+1$,转至步骤 2。设定结束条件为寻优达到最大迭代次数 $T_{\max}$ 或评价值小于给定精度。

步骤 6　输出结果。输出所求得的矩阵方程的最小二乘解,即对应式(2)中最优参数 b 和 $\{a_i\}_{i=1}^N$。

步骤 7　将 $\{a_i\}_{i=1}^N$ 和 b 代入式(3),得到识别函数式,再输入待识别样本进行识别。

本文选取径向基函数 $K(x,x_k)=\exp(-\|x-x_k\|^2/2\sigma^2)$ 为核函数,σ 为核函数的宽度。

3　应用举例

长庆气田为低产、低含气丰度、大面积分布的隐蔽岩性气藏,应用传统识别方法对气层测井资料进行定量评价存在很大困难。为此采用改进 LS-SVM 算法对 C 井目的井段(3300～3400m)进行定量评价。设计基于改进 LS-SVM 的气层训练识别模型如图 2。

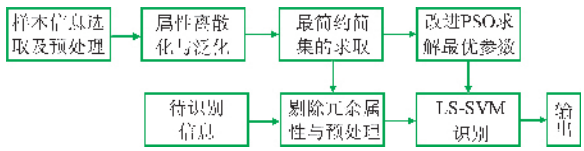

图 2　基于改进 PSO 的 LS-SVM 气层识别模型

3.1　样本信息选取与预处理

样本集资料的选取要完备、全面,应与气层评价密切相关,尽量保证选取的资料不重叠。在长庆气田 C 井目的井段(3300～3400m)处提取 190 个样本点测井数据作为训练样本,其中气层样本 85 个,非气层样本 105 个。为避免出现计算饱和现象,要对样本数据进行归一化处理,使输入的样本数据在区间[0,1],归一化公式为

$$x'=\frac{x-x_{\min}}{x_{\max}-x_{\min}} \tag{13}$$

式中:$x\in[x_{\min},x_{\max}]$;$x_{\min}$、$x_{\max}$ 分别为样本数据的最小值和最大值。

3.2　属性离散化和泛化

样本信息的决策属性为{非气层,气层},将决策属性 D 泛化为{−1,1},分别代表非气层和气层,样

返回目录

本信息的条件属性有 13 个，即：GR、DT、SP、WQ、LLD、LLS、DEN、NPHI、PE、U、TH、K、CALI，采用黄金分割优选法实现连续属性离散化，各属性离散化区间如表 2 所示。

表 2　13 个条件属性离散化区间

条件属性	离　散　化　区　间
GR	[26.2090,80.1413]，(80.1413,167.3930]
DT	[204.2380,239.9491]，(239.9491,262.0230]
SP	[-132.3970,-108.8555]，(-108.8555,-94.3040]
WQ	[5.9760,123.5468]，(123.5468,196.2200]
LLD	[16.9160,89.4111]，(89.4111,134.2220]
LLS	[16.5520,84.5147]，(84.5147,126.5240]
DEN	[2.3180,2.5806]，(2.5806,2.7430]
NPHI	[4.6180,18.5690]，(18.5690,41.1390]
PE	[1.7510,2.9833]，(2.9833,3.7450]
U	[0.2950,3.7774]，(3.7774,5.9300]
TH	[0.7480,9.2307]，(9.2307,22.9540]
K	[0.4060,2.3249]，(2.3249,3.5110]
CALI	[22.6490,23.8918]，(23.8918,24.6600]

3.3　样本信息属性约简

将 13 个条件属性通过预处理和离散化处理后，采用基于相似度的属性约简算法[9]剔除样本集中的冗余属性，从而得到反映气层特性的最简条件属性集{GR,DT,SP,LLD,LLS,DEN,K}，如表 3 所示，为约简后的 7 个样本属性的数值范围。

表 3　约简后的样本属性数值范围

属性	GR	DT	SP	LLD	LLS	DEN	K
最小值	20	190	−135	10	10	2	0
最大值	180	280	−90	170	150	3	4

将这 7 个属性按式(13)在井段 3300～3400m 之间进行归一化处理，如图 3 所示。

3.4　迭代寻优

设粒子群规模为 30，解空间为 191 维，即寻优参数的个数，加速常数 φ_1 和 φ_2 为 2.05，则收敛因子 $\chi=0.729$，初始惯性权重 $\omega=0.9$，粒子的最大速度 $V_{max}=10$。通过交叉验证的方式选取改进 LS-SVM 中正规化参数 $\gamma=1000$，径向基核函数的宽度参数 $\sigma^2=0.125$，利用基本 PSO 和改进 PSO 求解 LS-SVM 算法中的式(2)，所得到的误差曲线如图 4 所示。基本 PSO 迭代 298 次误差精度达到 1.4667×10^{-8}，而改进 PSO 达到同一精度仅需迭代 43 次，在迭代 96 次误差精度就已达到 2.0115×10^{-12}，可见改进 PSO 在收敛速度和精度上均远远优于基本 PSO。

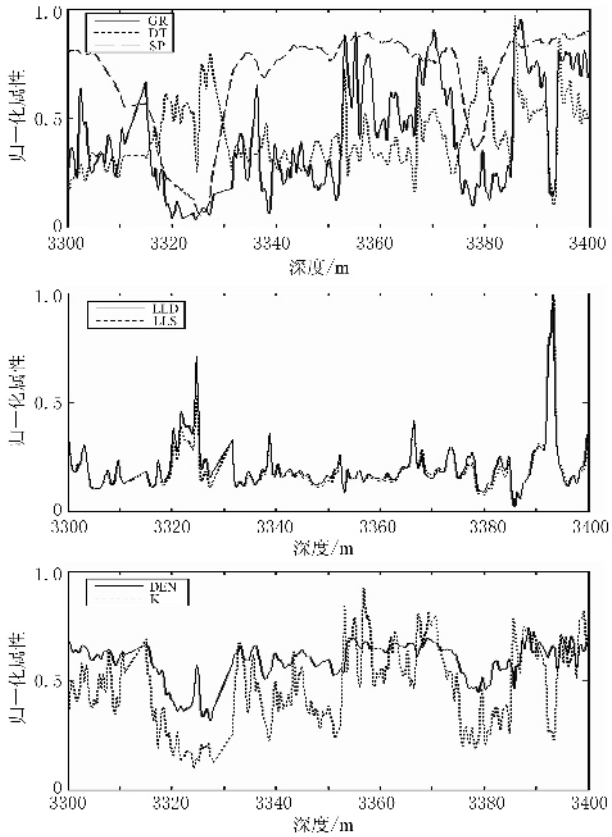

图 3　经过归一化后的 7 个属性的曲线图

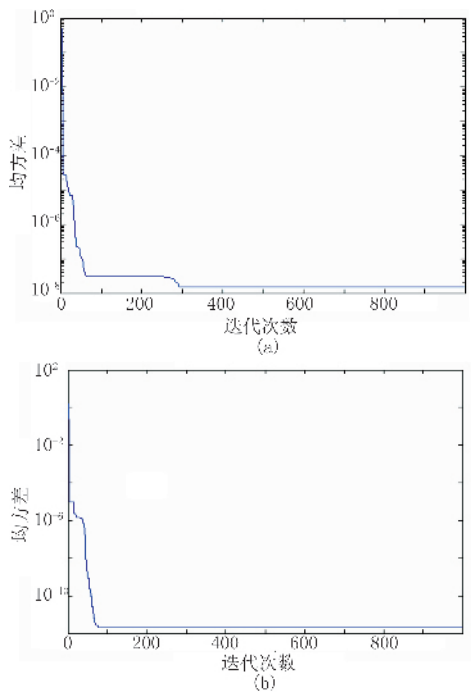

图 4　PSO 算法训练网络的误差演化曲线

(a)基本 PSO 算法；(b)改进 PSO 算法

返回目录

4　结果分析比较

将迭代寻优后得到的预测模型对全部井段样本进行气层识别,并与采用 BP 算法、经典 SVM 算法和传统 LS-SVM 算法的识别结果相比较。为了衡量识别结果的性能,定义以下几个性能指标

$$\text{均方根误差:RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}e_i^2} \tag{14}$$

$$\text{最大正误差:MAXPE}=\max\{e_i,0\} \tag{15}$$

$$\text{最大负误差:MAXNE}=\min\{e_i,0\} \tag{16}$$

式中:$e_i=y_i-\hat{y}_i$,$\hat{y}$ 和 y_i 分别为识别输出值和期望输出值。

其中,BP 算法采用"7-8-1"的网络结构,学习因子取为 0.05,隐含层和输出层的传输函数均选取 Tansig 函数($f(x)=2/(1+e^{-2x})-1$),经典 SVM 模型亦采用径向基核函数,通过交叉验证的方式选取传统 LS-SVM 中正规化参数 $\gamma=256$,径向基核函数的宽度参数 $\sigma^2=0.125$,性能指标如表 4 所示。

表 4　性能指标比较

算　法	RMSE	MAXPE	MAXNE	识别率 %	运行时间 s
BP	0.3020	1.2710	−1.0919	81.4	4.4913
经典 SVM	0.2613	1	−1	90.6	8.0938
LS-SVM	0.2321	1	−1	98.4	2.6875
改进 LS-SVM	0.1228	1	−1	99.1	1.0190

所得结果表明采用 LS-SVM 算法在识别率等性能指标上明显优于 BP 算法和经典 SVM 算法,LS-SVM 与 SVM 算法在识别精度、泛化能力上要优于 BP 算法,本文提出的改进 LS-SVM 不但有效地避免了 BP 算法中存在的过学习、维数灾难、局部极小等问题,而且解决了 SVM 算法对于大规模样本所出现的计算复杂度高、计算速度慢等问题,并且较传统 LS-SVM 算法在微机运行速度和识别率上均有所提高。

利用改进 LS-SVM 算法训练好的网络对 3300～3400m 井段进行识别,结果如图 5 所示,即在井段 3315～3326.5m、3336～3338m、3358～3360.5m、3374～3383m 这 4 个区间段均为气层,其他区间段均为非气层,识别结果与试气结论吻合,表明本文提出的改进 LS-SVM 算法适用于解决气层识别等大规模样本问题,且识别效果显著。

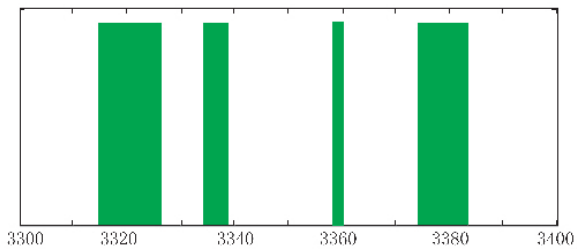

图 5　改进 LS-SVM 气层识别结果图

横坐标表示井深,单位为 m;纵坐标无量纲。
阴影区间表示气层,其它区间表示非气层

5　结论

基于结构风险最小化(SRM)准则的 LS-SVM 是一种很好的分类方法,本文提出的改进 LS-SVM 算法不仅避免了矩阵求逆计算,而且总能求得最优解,且其训练速度和求解精度均得到提高。实际应用结果表明,本文提出的改进的 LS-SVM 算法在气层识别中效果显著,且可以应用到其他大规模工程识别中,具有很好的应用前景。

参 考 文 献

[1] Vladimir N Vapnik 著,张学工译. 统计学习理论的本质. 北京:清华大学出版社,2000

[2] Cortes C, Vapnik V. Support vector machine. *Machine Learning*, 1995, 20:273～297

[3] Nello Cristianini, John Shawe-Taylor 著,李国正等译. 支持向量机导论. 北京:电子工业出版社,2005

[4] Suykens J A K, Vandewalle J. Least squares support vector machine classifiers. *Neural Processing Letters*, 1999, 9 (3):293～300

[5] 《测井学》编写组. 测井学. 北京:石油工业出版社,1998

[6] Kennedy J, Eberhart R C. Particle swarm optimization. *Proc IEEE Int Conf Neural Networks*. Piscataway, NJ: IEEE Press, 1995, 1942～1948

[7] Shi Y, Eberhart R. A modified particle swarm optimizer. In: *IEEE World Congress on Computational Intelligence*, 1998:69～73

[8] 李平,王椿镛. 地球物理反演中奇异值分解应用的若干问题研讨. 自然科学进展,2001,11(8):891～896

[9] 夏克文,刘明霄,张志伟,董瑶. 基于属性相似度的属性约简算法. 河北工业大学学报,2005,34(4):20～23

(本文编辑:冯小球)

返回目录