

文章编号:1673-5005(2010)02-0067-04

基于相关向量机的地层可钻性级值预测

马海¹, 王延江¹, 胡睿², 魏茂安²

(1. 中国石油大学 信息与控制工程学院, 山东 东营 257061; 2. 胜利油田 钻井工艺研究院, 山东 东营 257017)

摘要:对录井资料及测井资料与地层可钻性级值的关系进行分析,提出一种基于相关向量机算法的地层可钻性级值预测的新方法。通过标准化钻速、测井声波时差、地层密度、泥质含量和地层深度进行学习训练相关向量机,建立地层可钻性级值预测的相关向量机模型。对准噶尔盆地部分井的地层可钻性级值进行预测的结果表明,该方法优于BP神经网络方法,具有预测精度高、收敛速度快、推广能力强等优点。

关键词:地层可钻性; 相关向量机; 录井资料; 测井资料; 预测模型

中图分类号:TE 249 **文献标志码:**A **doi:**10.3969/j.issn.1673-5005.2010.02.013

Formation drillability prediction based on relevance vector machine

MA Hai¹, WANG Yan-jiang¹, HU Rui², WEI Mao-an²

(1. College of Information and Control Engineering in China University of Petroleum, Dongying 257061, China;
2. Drilling Technology Research Institute of Shengli Oilfield, Dongying 257017, China)

Abstract: The relation between log data, well-log data and formation drillability was analyzed, and a novel method for predicting formation drillability based on relevance vector machine (RVM) was proposed. A prediction model for formation drillability was established using the data of normalized drilling rate, well-log acoustic velocity, formation density, shaliness and formation depth by training the RVM. The proposed method was applied to predict the formation drillability of some wells in Junggar Basin. The results show that the RVM method has higher prediction precision, faster convergence speed and better generalization effect than BP neural network approach.

Key words: formation drillability; relevance vector machine; log data; well-log data; prediction model

地层可钻性是描述在一定的条件下钻进岩石难易程度的一个指标。地层可钻性指标是钻井工程中必需的基本数据,它在钻井工程中有重要的应用价值,可以指导钻头选型、用于钻头参数选优、预测钻速等^[1]。目前,国内外学者普遍采用的地层可钻性评价方法主要有室内岩心微钻头实验法、声波测井资料计算法和录井资料反算法^[1-5]。相关向量机(relevance vector machine, RVM)是 Michael E. Tipping 等提出的一种稀疏概率模型^[6-9],它的训练是在贝叶斯框架下进行的,可以用它进行回归估计预测,从而获得预测值的分布。目前相关向量机方法在非线性函数估计及模式识别方面已取得了较好的效果^[10-12],运算简单,收敛速度快,精度高。基于此,

笔者提出一种基于相关向量机算法的地层可钻性级值预测方法,将录井资料和测井资料相结合,建立地层可钻性级值预测模型。

1 相关向量机算法

RVM 采用一个核函数线性组合加权的函数进行分类和回归,它的理论在于设定权重满足均值为零、方差不同的高斯概率分布,然后通过贝叶斯推断获得权重概率分布中的参数优化值^[6],给定训练样本集 $\{x_i, t_i\}_{i=1}^N, x_i \in \mathbf{R}^d, t_i \in \mathbf{R}$, RVM 的模型输出为

$$y(\mathbf{x}; \mathbf{w}) = \sum_{i=1}^N w_i K(\mathbf{x}, x_i) + w_0. \quad (1)$$

式中, $K(\mathbf{x}, x_i)$ 为非线性基函数; w_i 为模型的权值; N 为样本数。选用以每个训练样本为中心的高斯函数

收稿日期:2009-04-19

基金项目:中国石化科技攻关项目(JP04014);山东省自然科学基金项目(ER2009FL029)

作者简介:马海(1981-),男(汉族),山东蓬莱人,博士研究生,从事信号与信息处理及信息融合方法研究。

作为基函数。在定义了模型(1)的基函数之后,可以在贝叶斯框架下用最大似然方法来训练模型权值 w_i , 这样可回避过学习问题,提高模型的泛化能力。因此,RVM 为每个权值定义了先验概率分布

$$p(\mathbf{w}|\boldsymbol{\alpha}) = \prod_{i=0}^N N(w_i|0, \alpha_i^{-1}) \quad (2)$$

式中, α_i 是决定权值 w_i 先验分布的超参数。

假定目标值 t_i 是独立的,并且数据的噪声服从方差为 σ^2 的高斯分布,则相应的训练样本集的似然函数为

$$p(\mathbf{t}|\mathbf{w}, \sigma^2) = (2\pi\sigma^2)^{-N/2} \exp\left\{-\frac{1}{2\sigma^2}\|\mathbf{t} - \Phi\mathbf{w}\|^2\right\} \quad (3)$$

其中

$$\begin{aligned} \mathbf{t} &= (t_1, \dots, t_N)^T, \quad \mathbf{w} = (w_0, \dots, w_N)^T, \\ \Phi &= [\varphi(x_1), \varphi(x_2), \dots, \varphi(x_N)]^T, \\ \varphi(x_n) &= K[1, K(x_n, x_1), K(x_n, x_2), \dots, K(x_n, x_N)]^T. \end{aligned}$$

根据先验概率分布和似然分布,再用贝叶斯公式计算权值的后验概率分布,即

$$p(\mathbf{w}|\mathbf{t}, \boldsymbol{\alpha}, \sigma^2) = \frac{p(\mathbf{t}|\mathbf{w}, \sigma^2)p(\mathbf{w}|\boldsymbol{\alpha})}{p(\mathbf{t}|\boldsymbol{\alpha}, \sigma^2)} = (2\pi)^{-(N+1)/2} \times |\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{w} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{w} - \boldsymbol{\mu})\right\} \quad (4)$$

其中

$$\begin{aligned} \Sigma &= (\sigma^{-2}\Phi^T\Phi + \mathbf{A})^{-1}, \quad \mathbf{A} = \text{diag}(\alpha_0, \alpha_1, \dots, \alpha_N), \\ \boldsymbol{\mu} &= \sigma^{-2}\Sigma\Phi^T\mathbf{t}. \end{aligned}$$

式中, Σ 为协方差; $\boldsymbol{\mu}$ 为均值。

训练目标值的似然分布通过对权值变量进行积分实现边缘化^[6],从而求得超参数的边缘似然分布,即

$$p(\mathbf{t}|\boldsymbol{\alpha}, \sigma^2) = \int p(\mathbf{t}|\mathbf{w}, \sigma^2)p(\mathbf{w}|\boldsymbol{\alpha})d\mathbf{w} = N(0, \mathbf{C}) \quad (5)$$

其中

$$\mathbf{C} = \sigma^2\mathbf{I} + \Phi\mathbf{A}^{-1}\Phi^T.$$

最后,RVM 方法中的模型权值的估计值由后验分布的均值给出,同时它也是权值的最大后验(maximum a posterior, MAP)估计。权值的 MAP 估计取决于超参数 $\boldsymbol{\alpha}$ 和噪声方差 σ^2 ,它通过最大化超参数边缘似然分布^[6-7]找到其可能值 $\boldsymbol{\alpha}_{MP}$ 和 σ_{MP}^2 。后验分布反映出权值最优值的不确定性,可以表示模型预测的不确定性。根据权值的后验及最优超参数 $\boldsymbol{\alpha}_{MP}, \sigma_{MP}^2$,可以预测新的观测数据 x_* ,计算其预测分布

$$p(t_*|\mathbf{t}, \boldsymbol{\alpha}_{MP}, \sigma_{MP}^2) = \int p(t_*|\mathbf{w}, \sigma_{MP}^2)p(\mathbf{w}|\mathbf{t}, \boldsymbol{\alpha}_{MP}, \sigma_{MP}^2)d\mathbf{w} \quad (6)$$

由于两积分均为高斯分布,故

$$p(t_*|\mathbf{t}, \boldsymbol{\alpha}_{MP}, \sigma_{MP}^2) = N(t_*|y_*, \sigma_*^2).$$

其中预测均值 $y_* = \boldsymbol{\mu}^T \boldsymbol{\varphi}(x_*)$, 方差 $\sigma_*^2 = \sigma_{MP}^2 + \boldsymbol{\varphi}(x_*)^T \Sigma \boldsymbol{\varphi}(x_*)$ 。

RVM 使用贝叶斯框架下的显著度解决了模型的参数选取问题,具有较好的适用性。

2 地层可钻性级值预测

大量研究表明,地层可钻性与地层岩性、物性、电性、流体特性密切相关,即与声波时差等测井资料关系密切。声波时差和岩石可钻性均间接反映了岩石的各种机械力学性质,是同一事物的不同体现,密度和泥质含量是影响岩石声波时差最主要、最直接的因素,而且声波时差和密度也是影响岩石弹性参数和岩石强度的重要参数。另外在实际钻井过程中,钻速与岩性具有一定的对应关系,由于影响钻速的因素很多,如钻压、转速、比水功率、钻头直径、井底压差等,本文中采用“标准化钻速法”^[13-14]将不同条件下取得的钻速转化为同一条件下的标准化钻速,使机械钻速更好地表征岩石可钻性。从理论上讲,声波时差、地层密度、泥质含量、标准化钻速和地层深度必然与地层可钻性级值有某种内在的联系^[15-16]。

利用相关向量机算法来建立地层可钻性级值预测模型,用与地层可钻性密切相关的录井资料(标准化钻速)、测井资料(声波时差、地层密度、泥质含量)和地层深度作为网络输入,以地层可钻性级值作为网络输出值,来构成地层可钻性级值预测模型结构(图1)。

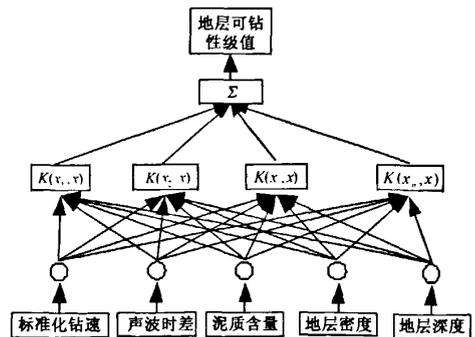


图1 地层可钻性级值预测模型结构图

Fig.1 Structural diagram of formation drillability prediction model

3 实例应用

试验中研究的区块为准噶尔盆地中部1区块莫西庄构造,其西面为盆1井西凹陷,东面为东道海子北凹陷,南面为昌吉凹陷,北接莫北突起。莫西庄背斜为一宽缓低幅度背斜。庄字号井主要包括庄2、庄102、庄103、庄104和庄105等重点井。首先在新区通过取心和室内可钻性测试,取得样本集,用相关向量机方法建立地层可钻性级值预测模型,对单井通过上部井段建立的模型对下部井段进行预测,然后通过多井资料建立的模型对其他井进行预测。

利用可钻性级值模型对庄2井进行地层可钻性级值预测。庄2井的部分数据如表1所示,其中可钻性级值数据为通过微钻头可钻性测定仪进行测量获取的。

表1 庄2井部分测试数据
Table 1 Test data in well Zhuang 2

样本	地层深度 h/m	标准化	声波	地层	泥质含量 V _{sh} /%	可钻性级值 K _d
		钻速 v _R / (m·h ⁻¹)	时差 Δt/ (μs·m ⁻¹)	密度 ρ/ (g·cm ⁻³)		
1	4253	8.077	230.39	2.400	18.2	3.34
2	4334	1.092	232.24	2.515	10.0	5.17
3	4404	1.509	219.27	2.508	6.6	4.82
4	4483	1.487	227.57	2.506	5.8	4.83
5	4548	1.250	219.52	2.554	6.5	5.64
6	4579	1.719	219.05	2.666	13.9	5.84
7	4627	8.154	211.66	2.457	4.2	4.32
8	4697	3.405	223.21	2.547	7.7	5.11
9	4731	2.082	231.27	2.617	8.3	5.56
10	4806	5.615	254.59	2.431	3.4	4.56

首先选择庄2井4.2~5.4 km井段的地层可钻性数据作为相关向量机算法的训练样本数据,尺度参数为35,超参数初始值为0.067。为了说明模型的先进性,采用BP神经网络方法和所建立的相关向量机预测方法分别对庄2井5.4~6.1 km井段的地层可钻性级值进行了预测。预测结果如图2所示,预测结果分析见表2。

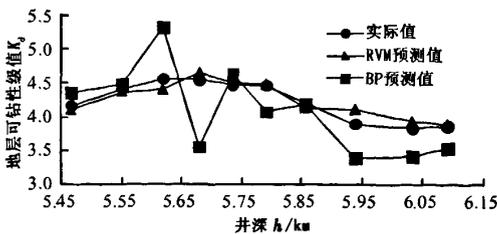


图2 单井地层可钻性级值预测结果
Fig. 2 Formation drillability prediction results of single well

由表2发现,利用RVM方法进行地层可钻性

级值预测比BP神经网络法预测的误差小,精度高。RVM方法的计算方差为0.0104,相关系数为0.9345,而BP方法的方差为0.2331,相关系数为0.6799。RVM方法预测结果相对BP方法比较稳定,与真实结果的相似程度高。综上所述,RVM方法可以提高地层可钻性级值的预测精度。

表2 庄2井5.4~6.1 km井段地层可钻性级值预测结果分析

Table 2 Formation drillability prediction results in 5.4~6.1 km depth of well Zhuang 2

样本	井深 h/m	可钻性级值 K _d		绝对误差 e _a		相对误差 e _r /%		
		实际值	RVM方法	BP方法	RVM方法	BP方法	RVM方法	BP方法
1	5466	4.17	4.11	4.36	0.06	0.19	-1.44	4.56
2	5551	4.42	4.37	4.49	0.05	0.07	-1.13	1.58
3	5619	4.57	4.42	5.33	0.15	0.76	-3.28	16.63
4	5679	4.55	4.66	3.55	0.11	1.00	2.42	-21.98
5	5736	4.49	4.50	4.64	0.01	0.15	0.22	3.34
6	5791	4.47	4.48	4.08	0.01	0.39	0.22	-8.72
7	5857	4.20	4.13	4.19	0.07	0.01	-1.67	-0.24
8	5940	3.90	4.11	3.39	0.21	0.51	5.38	-13.08
9	6033	3.84	3.95	3.42	0.11	0.42	2.86	-10.94
10	6091	3.86	3.89	3.53	0.03	0.33	0.78	-8.55

为了进一步验证所提方法的可行性和普遍性,选取庄2井、庄102井、庄103井和庄104井的录井资料、测井资料和可钻性级值数据,分别利用相关向量机方法和BP神经网络方法建立地层可钻性级值预测模型,对庄105井进行地层可钻性级值预测。预测结果如图3所示。

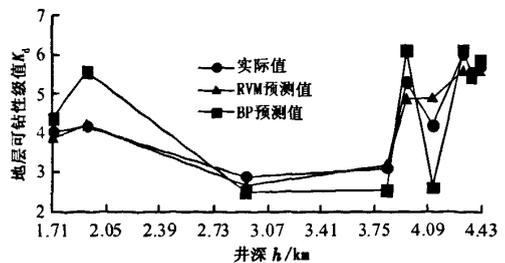


图3 多井联合地层可钻性级值预测结果
Fig. 3 Formation drillability prediction results of multi-well

由图3可以看出,相比BP神经网络方法,RVM方法能够更好地对地层可钻性级值进行预测。另外,随着区块已钻井数目的增多和收集到的样本集的丰富,待钻井的地层可钻性级值的预测模型将不断修正和完善。

4 结论

(1)将录井资料和测井资料结合进行地层可钻性级值预测可以在一定程度上克服室内岩心微钻头

实验法及单独使用声波测井资料和录井资料进行计算的不足。

(2) 基于相关向量机的可钻性级值预测模型比传统的 BP 神经网络算法收敛速度快, 预测精度高, 泛化效果好。

(3) 该方法已在中石化科技攻关项目“基于钻井工程地质数据库的钻井模拟”中得到应用, 通过它可以快速建立起某地区的地层可钻性剖面, 为钻井模拟提供依据。

致谢 胜利油田钻井工艺研究院钻井信息中心的孙正义副总工程师、曹锡玲高工、高兴坤高工为本研究提供了数据并给予了許多宝贵的意见, 在此一并表示感谢。

参考文献:

- [1] 邓理, 李黔, 高自力. 岩石可钻性评价方法研究新进展[J]. 钻采工艺, 2007, 30(6): 27-29.
DENG Li, LI Qian, GAO Zi-li. New evaluation method of rock drillability[J]. Drilling & Production Technology, 2007, 30(6): 27-29.
- [2] 张辉, 高德利. 钻头下部未钻开地层的可钻性预测新方法[J]. 石油学报, 2006, 27(1): 97-100.
ZHANG Hui, GAO De-li. A new method for predicting drillability of un-drilled formation[J]. Acta Petrolei Sinica, 2006, 27(1): 97-100.
- [3] 邹德水, 程远方, 查永进, 等. 利用岩屑波速随钻检测地层可钻性及优选钻头类型[J]. 石油大学学报: 自然科学版, 2005, 29(1): 37-40.
ZOU De-yong, CHENG Yuan-fang, ZHA Yong-jin, et al. Rock-drillability evaluation and bit selection through ultrasonic velocity measurement on cuttings[J]. Journal of the University of Petroleum, China (Edition of Natural Science), 2005, 29(1): 37-40.
- [4] SOMERTON W H. Further studies of the relation of physical properties to rock drillability [R]. SPE 2390, 1969: 57-67.
- [5] 潘起峰, 高德利. 岩石可钻性各向异性评价模型研究[J]. 天然气工业, 2005, 25(10): 64-66.
PAN Qi-feng, GAO De-li. Model to evaluate anisotropy of rock drill-ability [J]. Natural Gas Industry, 2005, 25(10): 64-66.
- [6] TIPPING M E. Sparse bayesian learning and the relevance vector machine [J]. Journal of Machine Learning Research, 2001, 1(3): 211-244.
- [7] TIPPING M E. The relevance vector machine [C]//SOLLA S A, LEEN T K, MÜLLER K R. Proceedings of Advances in Neural Information Processing Systems. Cambridge, Mass: MIT Press, c2000: 652-658.
- [8] MACKAY D J C. Bayesian interpolation [J]. Neural Computation, 1992, 4(3): 415-447.
- [9] BISHOP C M, TIPPING M E. Variational relevance vector machines [C]//BOUTILIER C, GOLDSZMIDT M. Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence. San Francisco, CA: Morgan Kaufmann, c2000: 46-53.
- [10] CHEN S, GUNN S R, HARRIS C J. The relevance vector machine technique for channel equalization application [J]. IEEE Trans on Neural Networks, 2001, 12(6): 1529-1532.
- [11] WEI Li-yang, YANG Yong-yi, NISHIKAWA R M, et al. Relevance vector machine for automatic detection of clustered microcalcifications[J]. IEEE Trans on Medical Imaging, 2005, 24(10): 1278-1285.
- [12] 刘遵雄, 张德运, 孙钦东, 等. 基于相关向量机的电力负荷中期预测[J]. 西安交通大学学报, 2004, 38(10): 1005-1008.
LIU Zun-xiong, ZHANG De-yun, SUN Qin-dong, et al. Mid-term electric load prediction based on the relevant vector machine[J]. Journal of Xi'an Jiaotong University, 2004, 38(10): 1005-1008.
- [13] 王彦棋. 标准化钻速法检测地层压力技术的改进与应用[J]. 钻采工艺, 1998, 21(6): 16-18.
WANG Yan-qi. Improvement and application of testing formation pressure using standardized drilling velocity method [J]. Drilling & Production Technology, 1998, 21(6): 16-18.
- [14] 陈中普, 黄国芳. 应用 NDR 计算地层孔隙压力[J]. 录井技术, 2000, 11(1): 21-27.
CHEN Zhong-pu, HUANG Guo-fang. Calculating formation pore pressure using NDR [J]. Mud Logging Technology, 2000, 11(1): 21-27.
- [15] 夏宏泉, 刘之的, 陈平, 等. 基于 BP 神经网络的岩石可钻性测井计算研究[J]. 测井技术, 2004, 28(2): 148-157.
XIA Hong-quan, LIU Zhi-di, CHEN Ping, et al. On calculating rock drillability from log data based on BP neural network technology [J]. Well Logging Technology, 2004, 28(2): 148-157.
- [16] 李荣, 何世明, 罗勇, 等. 钻井工程中岩石可钻性求取综合研究[J]. 钻采工艺, 2004, 27(5): 1-3.
LI Rong, HE Shi-ming, LUO Yong, et al. Comprehensive research on drillability in drilling project [J]. Drilling & Production Technology, 2004, 27(5): 1-3.

(编辑 修荣荣)