

文章编号: 1672-9315(2008)04-0707-04

基于支持向量机回归的煤层含气量预测

连承波¹, 赵永军¹, 李汉林¹, 渠芳¹, 蔡福龙², 张军涛³

(1. 中国石油大学 地球资源与信息学院, 山东 257061; 2. 中国科学院 地质与地球物理研究所, 北京 100029; 3. 南京大学 地球科学系, 南京 210093)

摘要: 为了探讨煤层含气量的有效预测方法, 将支持向量机回归方法用于建立煤层含气量预测模型。利用所选的测井参数, 采用基于小样本理论的支持向量机回归方法建立测井参数与煤层含气量的关系模型, 对煤层含气量进行预测。实例分析表明, 选取适当的测井参数, 利用支持向量机回归方法建立的煤层气含量预测模型, 其预测结果与实测结果的误差小。

关键词: 支持向量机; 测井参数; 煤层含气量

中图分类号: TD 712 **文献标识码:** A

0 引言

煤层中的含气量不仅是煤矿生产的重要灾害因子之一^[1], 也是决定一个地区煤层气资源能否进行商业化勘探开发的先决条件^[2]。因此, 无论是为了煤矿生产安全, 还是为了煤层气资源而准确地评价和预测煤层气开发前景以及制定开发方案, 煤层气含量都是一个至关重要的参数。通过煤岩取样测试和试井分析, 可获分析点附近煤层的含气量。但由于取样和测试的费用高, 样品分析数量有限, 且煤层含气量分布不均衡, 导致难以掌握工区煤层含气量的分布特征。测井技术是煤层气勘探开发中的重要手段, 所需要的费用较低, 测井资料覆盖面较广。由于测井信息相对丰富且分辨率较高, 具有弥补取心、试井及煤心分析等方面的不足的优点, 可用来预测煤层的含气量^[3-5]。基于测井信息的煤层含气量预测方法主要有模型法和统计学方法^[6-11]。模型法的评价结果过分依赖于测井解释模型的选择, 常用统计方法只对样本数量要求比较高, 而在实际应用中样本数目通常都是有限的, 因此传统统计方法难以取得理想的效果。近年提出基于小样本理论的支持向量机方法, 它避免了人工神经网络等方法的网络结构难于确定、过学习和欠学习以及局部极小等问题, 被认为是目前针对小样本的分类、回归等问题的最佳理论^[12,13]。本文采用支持向量机回归的方法建立测井参数与煤层含气量关系模型, 利用建立的模型进行煤层含气量预测, 并以算例验证其有效性, 试图寻求一种便捷可行的煤层含气量预测方法。

1 支持向量回归的基本原理

支持向量机是由 Vapnik 等提出的一种新型的机器学习算法, 它的理论基础是 Vapnik 由创建的统计学习理论 (statistic learning theory, SLT)^[12]。与传统统计学相比, 统计学习理论着重研究有限样本情况下的统计规律和学习方法。支持向量机 (SVM) 的核心是支持向量, 其基本思想是升维和线性化。即基于 Mercer 展开定理, 通过非线性映射, 把样本空间映射到一个高维特征空间, 然后在这个特征空间中求取最优线性分类面, 寻求最优回归超平面的问题转化为求解一个二次凸规划问题, 并可以求得全局最优解。在支持向量机回归中, SVM 首先选择一个非线性变换把原始空间中的数据映射到一个高维特征空间中, 再在高维特征空间中进行线性估计^[12,13]。问题的数学描述为: 假定一个训练样本集 $\{(x_i, x_i) | i = 1, 2, \dots,$

* 收稿日期: 2008-01-20

责任编辑: 杨忠民

基金项目: 国家重点基础研究规划项目 (973 项目) (2002CB211702)

作者简介: 连承波 (1979-), 男, 山东荣成人, 博士生, 主要从事计算机技术在地质中的应用研究。

n }, 其中为 x_i 输入向量, y_i 为对应的输出值, n 为样本个数, 要求拟合的函数形式为 $f(x) = \omega\phi(x) + b$. 支

持向量回归实际上是在条件
$$\begin{cases} y_i - [\omega^T \phi(x_i) + b] \leq \varepsilon + \xi_i, \\ [\omega^T \phi(x_i) + b] - y_i \leq \varepsilon + \xi_i^*, \\ \xi_i, \xi_i^* \geq 0, i = 1, \Lambda, n \end{cases}$$
 的约束下, 求解一个优化问题

$$\min_{\omega, b, \xi, \xi^*} \frac{1}{2} \omega^T \omega + C \sum_{i=1}^n (\xi_i + \xi_i^*),$$

其中 ϕ 为非线性映射函数, ξ 和 ξ^* 分别为在误差 ε 约束下训练误差的上限, C 是一个常数, 控制对错分样本的惩罚程度. 通过拉格郎日优化方法可将其转化为对偶问题, 归结为在条件

$$\begin{cases} \sum_{i=1}^n (a_i - a_i^*) = 0, \\ 0 \leq a_i, a_i^* \leq C \end{cases}, i = 1, \Lambda, n$$

约束下求解

$$\max_{a_i, a_i^*} = -\frac{1}{2} \sum_{i,j=1}^n (a_i - a_i^*)(a_j - a_j^*)K(x_i, x_j) - \varepsilon \sum_{i=1}^n (a_i + a_i^*) + \sum_{i=1}^n y_i(a_i - a_i^*).$$

其中 $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$ 为满足 mercer 条件的核函数, a_i 和 a_i^* 是拉格郎日乘子. 求解上述问题后得到回归函数

$$f(x) = \sum_{i=1}^N (a_i - a_i^*)K(x_i, x) + b.$$

其中 N 为支持向量个数.

2 煤层含气量预测

测井信息是岩石物理特征的综合响应, 选择合理的测井参数是准确评价煤层特征的前提. 从煤岩的物理特征和测井相应原理角度分析, 煤层含气量主要与密度 (DEN), 声波时差 (AC), 自然伽马 (GR) 测井参数有关^[7-9]. 煤层内气体是在煤化作用过程中由成煤有机质通过生物化学和热解作用后生成的. 在后期地质作用下, 这些气体有的通过各种途径运移外逸, 有的还残留于煤层内. 其中大部分吸附于煤的裂隙表面和各种大小的微孔隙内, 以固溶状态存在. 含气量是指煤中实际储存的气体含量, 煤层吸附气体后, 基于测井体积模型的基本原理, 不同含气量所引起煤层的密度、传播速度等性质发生变化, 进而影响测井参数密度、声波时差、自然伽马响应特征. 测井信息是岩石物理特征的综合响应, 但从岩石地球物理学的角度来看, 煤层含气量与测井参数之间并无确定的线性关系. 为了利用测井信息对煤层含气量进行预测, 本次采用支持向量机回归的方法, 建立煤层含气量与测井参数之间的关系模型, 用其进行煤层含量预测.

支持向量机回归具有很强的自适应能力, 可以建立测井参数与煤层含气量的关系模型, 对煤层含气量进行预测. 本次以具有煤层测井响应和对应测试煤层含气量数据^[7]的样品为例进行分析, 随机选取 15 组样本用来训练建立预测模型, 模型结构如图 1 所示. 本次采用的核函数为双曲正割核函数 (*Hyperbolic Secant Function Kernel*).

$$K(x, y) = \prod_{i=1}^n \frac{2}{e^{u(x_i - y_i)} + e^{-u(x_i - y_i)}}$$

参数 u 和 c 的取值分别为 3 和 100. 为了检验模型的推广能力, 选取 10 组样品作为检验样品进行预测分析, 其预测结果如表 2 所示. 从表 2 可以看出, 采用支持向量机回归的方法, 建立的煤层含气量与测井参数之间的关系模型, 用其预测的煤层含气量在一定程度上能够反映出了煤层实际含气量的变化趋势, 煤层含气量的预测值与实测值之间的误差较小, 准确性较高.

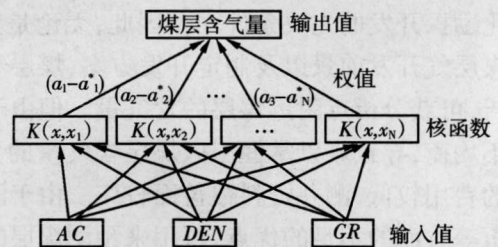


图 1 煤层含气量预测的支持向量回归模型结构
Fig.1 SVM-Regression model of coalbed gas content prediction

表2 煤层含气量预测结果
Tab.2 The result of prediction

标号	$AC/\mu s \cdot m^{-1}$	$DEN/g \cdot m^{-3}$	GR 脉冲/ min^{-1}	预测含气量/ $m^3 \cdot t^{-1}$	实测含气量/ $m^3 \cdot t^{-1}$	绝对误差	相对误差
1	402.00	1.43	2 077.23	10.36	10.58	0.22	0.02
2	424.65	1.44	3 410.95	12.09	13.24	1.15	0.09
3	431.16	1.42	3 293.65	12.39	13.62	1.23	0.09
4	426.51	1.40	2 040.39	11.87	10.09	1.78	0.18
5	400.25	1.55	5 626.08	7.95	7.63	0.32	0.04
6	433.52	1.48	7 948.32	2.07	3.32	1.25	0.38
7	387.31	1.57	77 38.42	5.73	4.23	1.50	0.35
8	370.48	1.78	12 257.17	9.54	7.18	2.36	0.33
9	424.73	1.54	9 844.30	7.63	7.73	0.10	0.01
10	410.47	1.48	3 772.09	11.10	11.83	0.73	0.06
平均误差						1.06	0.16

4 结 语

煤层含气量是煤矿安全生产和准确地评价和预测煤层气开发前景以及制定开发方案的重要评价参数。测井信息是岩石物理特征的综合响应,选择合理的测井参数是准确评价煤层特征的前提。基于小样本理论的支持向量机回归具有很强的自适应能力,可以建立测井参数与煤层含气量的关系模型,对煤层含气量进行预测。实例分析表明,选取适当的测井参数,利用支持向量机回归方法建立的煤层气含量预测模型,其预测结果误差小,可为煤层气资源的勘探开发提供一定的参考依据。

参考文献 References

- [1] Wang Y. Coal resource evaluation and management information system in underground mines[J]. Computer Application in the Mineral Industries, Netherlands, 2001, 65(1):155-158.
- [2] 桑树勋,刘焕杰,范炳恒,等.含煤盆地煤层气资源量计算方法[J].中国煤层气,1996,(2):106-109.
SANG Shu-yun, LIU Hhuan-jie, FAN Bing-heng, et al. Calculating method on coalbed methane resource in coal-bearing basin[J]. China Coalbed Methane, 1996, (2):106-109.
- [3] 王敦则,蔚远江,覃世银.煤层气地球物理测井技术发展综述[J].石油物探,2003,24(4):385-390.
WANG Dun-ze, YU Yuan-jiang, TAN Shi-yin. A review on the development of geophysical logging techniques for coalbed methane reservoir[J]. Geophysical Prospecting For Petroleum, 2003, 24(4):385-390.
- [4] 潘和平.煤层气储层测井评价[J].天然气工业,2005, 25(3):48-51.
PAN He-ping. Evaluating coalbed methane reservoir by log data[J]. Natural Gas Industry, 2005, 25(3):48-51.
- [5] 周尚忠.煤层气地球物理测井系列选择[J].中国煤层气,2006,3(2):25-27.
ZHOU Shang-zhong. Series option of coalbed methane geophysical logging[J]. China Coalbed Methane, 2006, 3(2):25-27.
- [6] 葛祥,李涛.煤层物性测井评价方法初探[J].测井技术,2003,27(2):129-131.
GE Xiang, LI Tao. On well logging evaluating method for physical property of coalbed[J]. Well Logging Technology, 2003, 27(2):129-131.
- [7] 潘和平,黄智辉.煤层含气量测井解释方法探讨[J].煤田地质与勘探,1998,26(2):58-60.
PAN He-ping, HUANG Zhi-hui. Discussion on the log interpretation method of coalbed gas content[J]. Coal Geology and Exploration, 1998, 26(2):58-60.
- [8] 孙耀庭,李涛.测井评价煤层天然气的方法[J].石油天然气学报,2005,27(4):464-467.
SUN Yao-ting, LI Tao. Well logging evaluating method for coalbed gas[J]. Journal of Oil and Gas Technology, 2005, 27(4):464-467.
- [9] 高绪晨,张春才,段铁梁.煤层气测井资料解释初探[J].中国煤田地质,2003,15(4):54-57.

- GAO Xu-chen, ZHANG Chun-cai, DUAN Tie-liang. Probe into interpretation on logging informations of coalbed gas[J]. Coal Geology of China, 2003, 15(4): 54-57.
- [10] 渠芳, 陈清华, 连承波, 等. 河流相储层细分与对比中存在的问题[J]. 西安科技大学学报, 2007, 27(1): 49-53.
QU Fang, CHEN Qing-hua, LIAN Cheng-bo, et al. Problems in the subdivision and correlation of fluvial reservoir[J]. Journal of Xi'an University of Science and Technology, 2007, 27(1): 49-53.
- [11] 马东民, 史纪敦, 张守刚. 沁南地区煤气的解吸特征实验研究[J]. 西安科技大学学报, 2007, 27(4): 581-583.
MA Dong-min, SHI Ji-dun, ZHANG Shou-gang. Experiment on Qinnan field CBM desorption[J]. Journal of Xi'an University of Science and Technology, 2007, 27(4): 581-583.
- [12] Vapnik V. The nature of statistical learning theory[M]. New York: Springer, 1995: 10-45.
- [13] 张学工. 关于统计学习理论与支持向量机[J]. 自动化学报, 2000, 26(1): 32-42.
ZHANG Xue-gong. Introduction to statistical learning theory and support vector machines[J]. Acta Automatica Sinica, 2000, 26(1): 32-42.

Prediction of coalbed gas content based on support vector machine regression

LIAN Cheng-bo¹, ZHAO Yong-jun¹, LI Han-lin¹, QU Fang¹, CAI Fu-long², ZHANG Jun-tao³

(1. School of Earth Resources and Information, Petroleum University, Shandong 257061, China;

2. Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China;

3. Dept. of Earth Science, Nanjing University, Nanjing 210093, China)

Abstract: Support vector machine regression was proposed to predict coalbed gas content based on well logging data. The model of correlation between coalbed gas content and well logging parameters was established based on support vector machine regression. The model was used to predict coalbed gas content. The practical application shows that the error between predicted results and those analyzed by sample test is small.

Key words: support vector machine; well logging parameters; coalbed gas content

* **Biography:** LIAN Cheng-bo, Candidate for Ph. D., Shandong 257061, P. R. China, Tel: 0086-13793957069, E-mail: lianchengbo1999@hotmail.com