

Research on Statistical Methods of Mining Failure Depth of Floor

Xiangshuai Meng*, Haifeng Lu, Manman Zhang, Guifang Zhang, Chao Li

School of Earth and Environment, Anhui University of Science & Technology, Huainan Anhui
Email: *19966590832@189.cn

Received: Jun. 26th, 2020; accepted: Jul. 9th, 2020; published: Jul. 16th, 2020

Abstract

The research on mining failure depth of floor is of great significance to the risk assessment of water inrush in coal mines. With the rise of scientific calculation, it is more and more common to study the failure depth of floor by statistical methods. Based on the previous research results, this paper systematically combs the principle of regression analysis, support vector machine, grey prediction and BP neural network and their application in the determination of failure depth of floor, and looks forward to the research direction of statistical methods of failure depth of floor. It points out that with the development of computer technology, the determination of failure depth of floor will be more diversified and accurate. In the future, multi method fusion analysis based on machine learning is one of the main methods to study the mining failure depth of floor.

Keywords

Floor Failure Depth, Regression Analysis, Support Vector Machine, Grey Theory, BP Neural Network

底板采动破坏深度统计学方法研究

孟祥帅*, 鲁海峰, 张曼曼, 张桂芳, 李超

安徽理工大学地球与环境学院, 安徽 淮南
Email: *19966590832@189.cn

收稿日期: 2020年6月26日; 录用日期: 2020年7月9日; 发布日期: 2020年7月16日

摘要

底板采动破坏深度的研究对煤矿突水危险性评价具有重要意义, 随着科学计算的兴起, 以统计学方法研究底板采动破坏深度。*通讯作者。

究底板破坏深度越来越普遍。文章基于前人研究成果系统梳理了回归分析、支持向量机、灰色预测和BP神经网络四种算法的原理及其在底板破坏深度确定方面的应用,对底板破坏深度统计学方法研究的方向进行了展望,指出随着计算机技术的发展,底板破坏深度的确定会朝着更加多元化、精准化方向发展,基于机器学习的多方法融合分析是未来底板采动破坏深度研究的主要方法之一。

关键词

底板破坏深度, 回归分析, 支持向量机, 灰色理论, BP神经网络

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着我国工业的快速发展,对煤炭的需求量持续增大,煤矿开采深度不断增加,面临的地质环境也更加复杂[1][2]。尤其是在我国中东部地区,很多煤层直覆于承压含水层之上,因底板破裂引起的突水事故时常发生,给安全开采带来了极大的障碍。为解决这一问题,大量学者和现场工作者对此进行了研究,其中关于底板采动破坏深度的研究一直是热点课题之一[3][4]。目前,底板采动破坏深度确定的常用方法有四种:理论计算、现场实测、数值模拟和相似模拟。理论计算法不但需要现场实测数据作为基础,为简化计算还进行了诸多假设,导致结果往往与实际存在较大误差;现场实测虽然是目前获得底板破坏深度最准确、最可靠的一种方法,但该方法受限于地质环境,实施较为复杂;数值模拟虽计算简单,处理方便,但采用有限差分原理,计算结果往往也不符合实际;相似模拟是对现场环境的室内还原,可靠度较高,但成本较高。不宜大量使用。随着科学计算的兴起,基于统计学原理对底板破坏深度进行确定成为了当前研究的重要手段[5],该类方法较常规方法更为方便,结果可靠度更高。前人在这方面也做了大量工作。本文在前人研究成果的基础上,通过对基于统计学方法的底板采动破坏深度的相关研究进行了阐述,包括回归分析、支持向量机、灰色预测和BP神经网络,系统地梳理了各种研究方法的原理、计算过程以及前人所作的工作等,旨在能够为相关研究提供一定的参考。

2. 底板采动破坏深度确定的统计学方法

2.1. 回归分析

回归分析是以定量形式表征系统变量相关关系的一种统计分析方法,其中最常用的是多元线性回归,其假设多个自变量和因变量之间存在线性关系,通过求取参数建立数学模型[6],如公式(1)所示:

$$y_i = \alpha_i + \beta_{i1}x_1 + \beta_{i2}x_2 + \cdots + \beta_{ik}x_k + \varepsilon_i \quad (1)$$

上式可以简化表示为:

$$Y = a + X\beta + \varepsilon \quad (2)$$

式中: $Y = \begin{Bmatrix} y_1 \\ \vdots \\ y_n \end{Bmatrix}$ 为因变量, $\alpha = \begin{Bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{Bmatrix}$ 为截距项, $\beta = \begin{Bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{Bmatrix}$ 为待求参数, $X = \begin{Bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{Bmatrix}$ 为自变量,

$$\varepsilon = \begin{Bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{Bmatrix} \text{ 为误差项。}$$

回归方程分线性部分和误差项两部分，其共同解释因变量的变化规律。对于回归模型，其待求参数一般通过最小二乘法求得，确定 α 和 β 即确定了回归方程，即可对结果进行预测。回归分析可广泛应用于各领域，基于多元线性回归对底板破坏深度进行预测是在理论计算的基础上应用最早、也是最为简单的一类统计学研究方法，很多学者对此进行了研究。李小孟[7]等结合工程实例采用多元线性回归对底板破坏深度进行了预测，对比实测数据，二者结果基本相同。段宏飞[8]、张风达[9]等均在分析影响底板破坏深度的因素的基础上，采用回归分析方法建立了底板破坏深度预测模型，实现了底板破坏深度的预测。

2.2. 支持向量机

支持向量机是一种基于深度学习的广义分类算法，常用以进行识别分类及回归分析。支持向量机与其它算法相比优势在于可以对复杂非线性方程进行深度学习。支持向量机用以进行回归分析称为支持向量回归[10]。

在支持向量回归模型中，存在初始样本 $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ ， $y_i \in R$ ，期望通过反复学习得到回归模型 $f(x) = w^T + b$ ，使得 $|f(x) - y| \rightarrow 0$ ， $f(x)$ 即为所求结果，其计算过程如下：

假设 $f(x)$ 与 y 之间存在最大偏差 ε ，当且仅当 $|f(x) - y| > \varepsilon$ 时才计算损失，如图 1 所示。

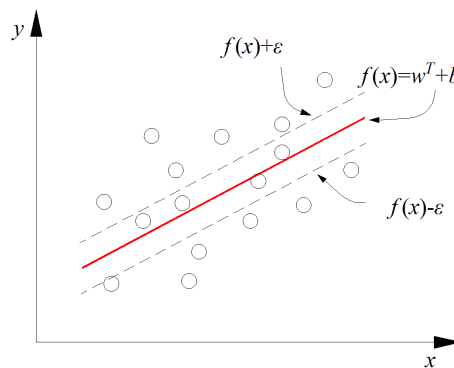


Figure 1. Support vector regression model
图 1. 支持向量回归模型

则支持向量回归问题可表示为：

$$\min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m l_\varepsilon(f(x_i), y_i) \quad (3)$$

式中： C ——正则化常数， w, b ——待求参数， l_ε ——不敏感损失函数， l_ε 函数示意图如图 2 所示。

l_ε 函数表达式为：

$$l_\varepsilon(z) = \begin{cases} 0 & |z| \leq \varepsilon \\ |z| - \varepsilon & |z| > \varepsilon \end{cases} \quad (4)$$

引入松弛变量 ξ_i 和 $\hat{\xi}_i$ ，则式(3)可重写为：

$$\min_{w,b,\xi_i,\hat{\xi}_i} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m (\xi_i, \hat{\xi}_i) \quad (5)$$

引入拉格朗日乘子 μ_i ，

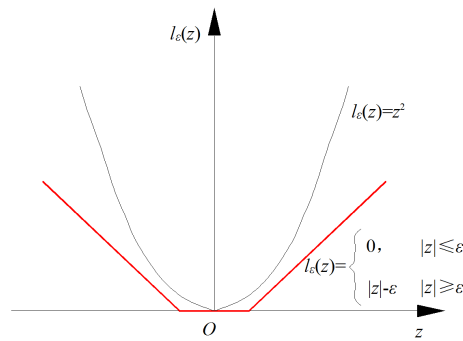


Figure 2. Insensitive loss function l_ϵ
图 2. 不敏感损失函数 l_ϵ

$$L(w, b, \alpha, \hat{\alpha}, \zeta, \hat{\zeta}, \mu, \hat{\mu}) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m (\xi_i + \hat{\xi}_i) - \sum_{i=1}^m \xi_i \mu_i - \sum_{i=1}^m \hat{\xi}_i \hat{\mu}_i + \sum_{i=1}^m \alpha_i (f(x_i) - y_i - \epsilon - \xi_i) + \sum_{i=1}^m \hat{\alpha}_i (y_i - f(x_i) - \epsilon - \hat{\xi}_i) \quad (6)$$

再令 $L(w, b, \alpha, \hat{\alpha}, \zeta, \hat{\zeta}, \mu, \hat{\mu})$ 对 $w, b, \xi_i, \hat{\xi}_i$ 的偏导数为零可得:

$$w = \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) x_i \quad (7)$$

上述过程中需满足下列条件:

$$\begin{cases} \alpha_i (f(x_i) - y_i - \epsilon - \xi_i) = 0 \\ \hat{\alpha}_i (y_i - f(x_i) - \epsilon - \hat{\xi}_i) = 0 \\ \alpha_i \hat{\alpha}_i = 0, \xi_i \hat{\xi}_i = 0 \\ \xi_i (C - \alpha_i) = 0, \hat{\xi}_i (C - \hat{\alpha}_i) = 0 \end{cases} \quad (8)$$

则模型的解为:

$$f(x) = \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) k(x_i^T x) + b \quad (9)$$

若 $0 < \alpha_i < C$, 则必有 $\xi_i = 0$ 。

$$b = y_i + \epsilon - \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) x_i^T x \quad (10)$$

基于以上原理, 确定 w 和 b 的支持向量回归模型即可实现参数预测。支持向量回归算法较普通回归算法精度更高、且适用性强, 很多学者采用支持向量回归算法对底板破坏深度进行了研究。柳聪亮[11]等采用支持向量机创建了预测底板采动导水破坏带深度的回归计算模型。通过国内一些矿区实测成果对模型进行反复训练, 模型预测结果和实测结果基本相同, 可以较好地满足工程设计需要。韩进[12]等运用遗传算法和支持向量机融合分析, 以深部煤层底板破坏数据为样本, 构建了底板破坏深度与其影响因素间的非线性模型, 得出了模型预测结果。

2.3. 灰色预测

灰色预测法是一种基于灰色理论对具有波动因素的系统进行预测的统计学方法。其原理主要是通过

鉴别单系统或多系统因素之间发展趋势的内在相异性和变化程度，确定系统影响因素的关联度，生成初始数据序列，确定系统变化的基本规律，从而实现对事物发展趋势的预测。

GM(0, N)是常用的灰色预测数学模型[13]，其可以表示为：

$$x_1^{(1)} = a + b_1x_2^{(1)} + b_2x_3^{(1)} + \dots + b_{n-1}x_n^{(1)} \tag{11}$$

式中： a 、 b_i 为待定系数， i 为样本数据序列个数， $i = 1, 2, \dots, n-1$ ；

记样本数据序列为： $\{x_i^{(0)}(t)\}, i = 1, 2, \dots, n; t = 1, 2, \dots, m$ ， n 表示样本数据序列个数， m 表示单个样本数据序列中的元素个数。一次累加后，数据序列为： $\{x_i^{(1)}(t)\}, i = 1, 2, \dots, n; t = 1, 2, \dots, m$ 。

$$\text{记： } X(B) = \begin{bmatrix} 1 & x_2^{(1)}(1) & x_3^{(1)}(1) & \dots & x_n^{(1)}(1) \\ 1 & x_2^{(1)}(2) & x_3^{(1)}(2) & \dots & x_n^{(1)}(2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_2^{(1)}(m) & x_3^{(1)}(m) & \dots & x_n^{(1)}(m) \end{bmatrix} \tag{12}$$

$$Y_m = [x_1^{(1)}(1), x_1^{(1)}(2), \dots, x_1^{(1)}(m)]^T \tag{13}$$

$$\tilde{a} = [a, b_1, b_1, \dots, b_{n-1}] \tag{14}$$

由最小二乘法可求得：

$$\tilde{a} = [X(B)^T X(B)]^{-1} X(B)^T Y_m \tag{15}$$

将式(12)、式(13)、式(14)、计算结果代入式(15)，得到 GM(0, N)灰色模型。采用灰色预测进行分析可以充分考虑样本因素的权重，结果准确性更高，一些学者对此进行了研究。张文泉[14]等从力学角度研究了底板采动破坏的主要影响因素，利用灰色理论获取了影响底板破坏深度各因素的关联度，在此基础上以线性与非线性方式对破坏深度进行拟合，得到了新的计算公式。路畅[15]等为确定底板破坏深度主要影响因素权重，采用灰色关联分析法对其进行了分析，结合具体工作面实测数据，建立了底板破坏深度的灰色预测模型。并基于工程实例对模型进行了验证。

2.4. BP 神经网络

BP 神经网络是由多层神经元组成的人工神经网络，其主要由三部分构成，分别为：输入层、输出层与隐含层。输入层用以接收初始数据；隐含层用以表示或存贮知识；输出层反馈出决策或预测结果[16]。多层 BP 神经网络结构图如图 3 所示。

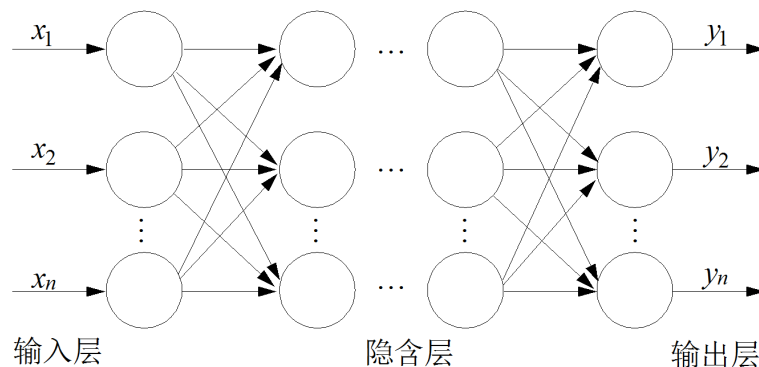


Figure 3. Structure of multilayer BP neural network
图 3. 多层 BP 神经网络结构图

BP 神经网络采用单向信号传递方式, 即神经元信号仅向下层传递, 同层神经元之间不相连, 相邻层神经元之间有连接, 以权重值表示其连接强度。每个神经元的状态均与一个作用函数 f 和阈值 θ 相对应, 在输入层中, $f(x) = x$, $\theta = 0$; 在隐含层和输出层中, $f(x) = 1/(1 + e^{-x})$ 。

则隐含层和输出层的输出可以分别由式(16)和式(17)计算得出:

$$b_j^k = f\left(\sum_{i=1}^n w_{ji} a_i^k - \theta_j\right) \quad (16)$$

$$c_i^k = f\left(\sum_{j=1}^p v_{ji} b_j^k - r_i\right) \quad (17)$$

式中: b_j^k ——隐含层第 j 个神经元对第 k 个样本的输出; n ——输入层神经元个数; w_{ji} ——输入层第 i 个神经元至隐含层第 j 个神经元的权值; a_j^k ——第 j 个输入神经元对应于第 k 个样本的输入; θ_j ——隐含层第 j 个神经元阈值; c_i^k ——输出层第 i 个神经元对第 k 个样本的输出; p ——隐含层神经元个数; v_{ji} ——隐含层第 j 个神经元至输出层第 i 个神经元的权值; r_i ——输出层第 i 个神经元的阈值。

由式(18)可以计算输出误差平方和 E 来作为判断训练过程是否结束的标准。

$$E = \sum_{k=1}^m \sum_{j=1}^q (y_j^k - c_j^k)^2 / 2 \quad (18)$$

式中: m ——训练样本个数; q ——输出神经元个数; y_j^k ——第 j 个输出神经元对第 k 个样本的理想输出。

在误差逆向传播的过程中, 校正误差由式(19)和式(20)求得。

$$d_i^k = (y_i^k - c_i^k) f'(L_i^k) \quad (19)$$

$$e_j^k = \left[\sum_{i=1}^q v_{ji} \cdot d_i^k \right] \cdot f'(s_j^k) \quad (20)$$

式中: q ——输出神经元个数; d_i^k , L_i^k ——输出层第 i 个神经元相对于第 k 个输入样本的校正误差和输入值; e_j^k , s_j^k ——隐含层第 j 个神经元相对于第 k 个输入样本的校正误差和输入值。

获得输出层和隐含层校正误差 d_i^k 和 e_j^k 之后, 为获取最优结果, 需要对相邻层之间的连接权值和阈值进行调整。其调整量可依据式(21)和式(22)计算:

$$\Delta v_{ji} = \alpha \cdot d_i^k \cdot b_j^k; \quad \Delta r_i = \alpha \cdot d_i^k \quad (21)$$

$$\Delta w_{ji} = \beta \cdot e_j^k \cdot a_i^k; \quad \Delta \theta_j = \beta \cdot e_j^k \quad (22)$$

式中: α , β ——学习系数, 且 $\alpha > 0$, $\beta > 1$, 将修正后的权值和阈值重新带入模型即可得到最优化输出层结果。

BP 神经网络是机器学习的经典算法之一, 经实践证明具有较高的可行性与准确性。很多学者在使用 BP 神经网络进行底板破坏深度预测方面进行了研究, 取得了大量的研究成果。于小鸽[17]依据典型煤矿突水案例, 基于 BP 神经网络构建了底板破坏深度预测模型, 并应用于工程实例, 通过对工程实测结果证实了模型的可靠性。郭文兵[18]对底板采动破坏深度影响因素进行了综合分析, 并应用神经网络方法建立了底板破坏深度计算模型, 对比分析了模型预测值、理论计算值和现场实测值。施龙青[19]等利用灰狼算法优化 BP 神经网络参数, 建立 PCA-GWO-BP 神经网络模型预测煤层底板破坏深度, 并与其他预测方法进行对比, 结果证明该模型误差小、准确度高。

3. 结论

煤层底板破坏深度的准确确定目前仍是一个非常重要的研究课题, 笔者在查阅研究大量资料的基础

上, 对使用统计学原理确定底板破坏深度的常用方法进行了阐述, 得出如下结论:

1) 随着机器学习与大数据的兴起, 使用统计学方法实现煤层底板破坏深度的预测已越来越成为一种重要手段。

2) 回归分析、支持向量机、灰色预测、BP 神经网络是目前在底板破坏深度研究中最常用的几种统计学方法。目前, 支持向量机和 BP 神经网络的应用较多, 结果也较为准确, 但计算步骤略为繁琐, 回归分析、灰色预测应用较少, 但计算简单, 结果可靠度也较高。

3) 底板采动破坏深度的确定目前正朝着多元化, 精细化, 智能化方向发展, 基于机器学习和大数据的多方法融合分析是未来底板破坏深度确定的必然趋势。

基金项目

安徽省教育厅 2019 年高校自然科学研究重大项目(KJ2019ZD11); 国家自然科学基金面上项目(No. 41977253)。

参考文献

- [1] 宋振骥, 文志杰. 煤炭资源的开采现状及发展方向[J]. 科技导报, 2011, 29(35): 3.
- [2] 袁亮. 煤及共伴生资源精准开采科学问题与对策[J]. 煤炭学报, 2019, 44(1): 1-9.
- [3] 徐智敏, 孙亚军, 巩思园, 等. 高承压水上采煤底板突水通道形成的监测与数值模拟[J]. 岩石力学与工程学报, 2012, 31(8): 1698-1704.
- [4] 鲁海峰, 姚多喜, 胡友彪, 等. 水压影响下煤层底板采动破坏深度弹性力学解[J]. 采矿与安全工程学报, 2017, 34(3): 452-458.
- [5] 白丽扬, 赵金海, 刘占新, 等. 基于数据挖掘算法的底板破坏深度预测[J]. 煤炭工程, 2017, 49(6): 92-95.
- [6] 施龙青, 徐东晶, 邱梅, 等. 采场底板破坏深度计算公式的改进[J]. 煤炭学报, 2013, 38(S2): 299-303.
- [7] 李小孟, 刘磊, 冯阳. 基于多元线性回归的底板破坏深度研究[J]. 山东煤炭科技, 2015(8): 148-150.
- [8] 段宏飞. 底板破坏深度六因素线性预测模型[J]. 岩土力学, 2014, 35(11): 3323-3330.
- [9] 张风达. 深部煤层底板破坏深度多元非线性预测模型[J]. 煤矿开采, 2018, 23(1): 92-95.
- [10] 赵云平, 邱梅, 刘绪峰, 等. 煤层底板破坏深度预测的 GRA-FOA-SVR 模型[J]. 中国科技论文, 2018, 13(3): 247-252.
- [11] 柳聪亮, 谭志祥, 李培现, 等. 底板采动导水破坏带深度求取方法研究[J]. 煤矿开采, 2010, 15(5): 38-41.
- [12] 韩进, 王颖, 施龙青, 等. 基于 GA-SVM 大采深条件下底板破坏深度回归分析[J]. 煤炭技术, 2017, 36(6): 7-9.
- [13] 高雅. 基于 GM(0,n)灰色预测模型的构造预测及定量评价[J]. 中国煤炭, 2019, 45(6): 38-42+101.
- [14] 张文泉, 赵凯, 张贵彬, 等. 基于灰色关联度分析理论的底板破坏深度预测[J]. 煤炭学报, 2015, 40(S1): 53-59.
- [15] 路畅, 尹立明, 李杨杨, 等. 基于灰色系统理论的煤层底板破坏深度预测模型[J]. 矿业研究与开发, 2015, 35(6): 63-67.
- [16] 于小鸽. 基于信息融合的底板突水预测[D]: [硕士学位论文]. 青岛: 山东科技大学, 2008.
- [17] 于小鸽, 韩进, 施龙青, 等. 基于 BP 神经网络的底板破坏深度预测[J]. 煤炭学报, 2009, 34(6): 731-736.
- [18] 郭文兵, 邹友峰, 邓喀中. 煤层底板采动导水破坏深度计算的神经网络方法[J]. 中国安全科学学报, 2003, 13(3): 37-40+84.
- [19] 施龙青, 张荣遨, 韩进, 等. 基于 GWO 改进的 PCA-BP 神经网络煤层底板破坏深度预测模型[J]. 矿业研究与开发, 2020, 40(2): 88-93.