



基于高斯过程回归的岩体结构面粗糙度系数预测模型

郑可馨¹, 吴益平^{1*}, 李江², 苗发盛¹, 柯超¹

1. 中国地质大学(武汉)工程学院, 武汉 430074;

2. 湖北省自然资源厅信息中心, 武汉 430071

第一作者, E-mail: zhx1998@cug.edu.cn

* 通信作者, E-mail: ypwu@cug.edu.cn

2023-03-06 收稿; 2023-04-18 修回; 2023-04-19 接受

基金项目: 湖北省自然科学基金项目(2023AFB580); 贵州省省级科技计划项目(黔科合支撑[2023]一般 127); 国家自然科学基金项目(42377161; 41977244)

中图分类号: P642

文章编号: 2096-8523(2024)04-0252-10

doi:10.19509/j.cnki.dzkq.tb20230113

摘要: 岩体结构面粗糙度系数(JRC)的估算是岩体力学性质评价的重要环节, 由于单一统计参数法难以全面表征岩体结构面的复杂粗糙形貌, 单一统计参数法建立的 JRC 计算模型精度较低。选取表征结构面粗糙形态的 8 种统计参数, 结合主成分分析法(PCA)和高斯过程回归(GPR)算法, 构建基于多参数融合的 JRC 预测模型。以公开的 112 条岩体结构面剖面线数据集(其中 95 条作为训练样本, 17 条为验证样本)为例进行分析研究, 最后将预测所得 JRC 与实测值对比并分析预测效果。结果表明: 由高斯过程回归构建的 JRC 预测模型决定系数(R^2)高达 0.972, 均方根误差(MSE)为 0.517, 反映出高斯过程回归方法在小样本条件下构建多统计参数与 JRC 值隐式关系的适用性, 为今后人工智能在 JRC 指标预测方面实现合理预测提供了思路。

关键词: 岩体结构面; 粗糙度; 高斯过程回归; 统计参数; 预测

岩体内部发育大量结构面, 结构面的力学性质对于岩体的稳定性评估有重要意义, 其中岩体结构面的剪切峰值强度对于岩体的稳定、变形起决定性的控制作用^[1-4]。结构面粗糙度系数(JRC)作为无充填岩体结构面峰值剪切强度评价的基础指标, 被广泛应用于岩体结构面峰值强度预测^[5-9]。 JRC 的定量评价是一以来来岩土领域关注的热点。

国内外学者提出许多定量获取 JRC 方法, 主要分为试验反算^[10-11]、统计参数法^[12-13]、分形数学法^[14]3 类, 其中试验方法可以获得精确的 JRC , 但需要耗费大量成本, 在实际工程中的应用存在局限性^[15-16]。在工程实践中, 统计参数法直观且易于实现, 是最常用的方法之一, 目前基于统计参数的 JRC 定量研究已有丰硕的成果。TSE 等^[17]和 YANG^[18]

深入探究了坡度均方根与 JRC 之间的定量关系。SAYLES 等^[19]和 ZHANG 等^[20]在考虑结构面取样长度的条件下探究起伏高度与 JRC 的密切关系。YU 等^[21]、MAERZ 等^[22]和陈世江等^[23]研究了用剖面线粗糙系数表示 JRC 的关系式。李化等^[24]建议用伸长率与相对起伏度共同评价 JRC 。蔡毅等^[25]研究了抗剪强度与其在垂直于剪切方向的面上的投影面积之间的关系, 提出了三维层面表征结构面粗糙形貌的 I_{PAP} 指标。然而结构面是复杂的微元起伏组合, 上述单一或者少量参数的 JRC 模型, 难以充分考虑到结构面的起伏角、起伏高、起伏分布等特征^[26-27]。为了得到更可靠的 JRC 值, 有必要对多参数的 JRC 预测模型深入研究。

随着智能化学习的不断发展^[28-29], 学者们^[30-33]

著录信息: 郑可馨, 吴益平, 李江, 等. 基于高斯过程回归的岩体结构面粗糙度系数预测模型[J]. 地质科技通报, 2024, 43(4): 252-261.

ZHENG Kexin, WU Yiping, LI Jiang, et al. A prediction model of the joint roughness coefficient based on Gaussian process regression[J]. Bulletin of Geological Science and Technology, 2024, 43(4): 252-261.

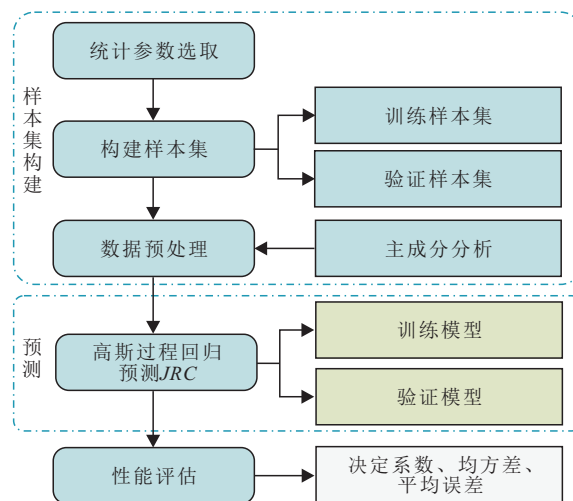
尝试用机器学习的方法建立多参数 JRC 预测模型,旨在用足够多的统计参数全面描述结构面的几何形貌,并利用机器学习的方法构建多种参数与 JRC 之间的映射关系。如 WANG 等^[31]提出用 SVR 结合因子分析的方法构建 JRC 预测模型的新思路。苗发盛等^[32]建立了基于 boosting 的决策树的 JRC 预测模型。尹宏等^[33]分析了因子分析结合神经网络预测 JRC 的适用性。虽然学者们为解决多统计参数与 JRC 的非线性回归问题提供了诸多思路,但是上述方法对数据量的要求较高,且模型优化过程较为复杂,参数众多,难以满足工程实践的要求。为此,引入参数较少和小样本条件下泛化能力较强的高斯过程回归(Gaussian process regression,简称 GPR)算法^[34],建立 JRC 预测模型,其基于贝叶斯理论框架为回归预测的 JRC 提供严格的概率意义,目前此方法在岩体结构面粗糙度预测方面的应用较少。

本研究选定能全面考虑结构面几何形态(起伏高、起伏角、起伏分布)的 8 个统计参数,构建 8 种统计参数的高斯过程回归 JRC 智能预测模型。首先将搜集到的 112 条结构面剖面线划分为 95 条训练样本与 17 条验证样本,使用主成分分析(principal component analysis,简称 PCA)的方法,降低统计参数的维度并构建复合指标系统,然后用高斯过程回归算法构建 8 种统计参数与 JRC 的非线性映射关系。该模型能够衡量输入指标对模型的贡献重要性,进而解译降维后的复合输入指标和 JRC 之间的相关关系,以此佐证复合参数构建的合理性。本研究得出的高斯过程回归预测模型具有较高解译性、准确性的优点,通过样本预测与实测 JRC 结果对比,证实了 JRC 预测模型的准确性。

1 研究方法

1.1 模型构建思路

高斯过程回归在小样本条件下对于高维度、非线性数据的回归问题有较强的优势,其原理基于贝叶斯概率的思想,且参数有较强的解译能力,近年来在岩土工程领域得到了广泛的应用^[35-36]。本研究拟构建高斯过程回归的 JRC 预测模型,构建的过程如图 1 所示。数据集被划分为 95 组训练样本与 17 组验证样本 2 部分:训练样本集被用于确定模型超参数与训练预测模型;验证样本用于验证模型的预测准确度以及泛化能力。训练模型的性能由均方误差(MSE)、决定系数(R^2)、平均误差(E_a) 3 个指标



JRC. 岩体结构面粗糙度系数

图 1 高斯过程回归模型预测 JRC 流程

Fig. 1 Flowchart of the JRC prediction based on GPR

评估。

1.2 数据集来源

本研究用于模型训练以及预测的结构面样本收集自文献^[31]中公开发表的 112 条岩体结构面剖面线数据(其中 102 条为真实岩体的实测剖面线,10 条为 Barton 标准剖面线^[9]),剖面线长度区间在 72 ~ 119.6 mm 之间,采样间距为 0.4 mm。图 2 展示了样本中岩体结构面剖面线的 JRC 值分布情况,对应剖面的 JRC 值是通过试验反算获得的实测值。102 条实测结构面剖面线样本是来自世界各国多个地区的岩样^[37-39],其涵盖的岩体类型广泛,包括不同风化程度(未风化、弱风化、微风化)的砂岩、石灰石、大理石、花岗岩、板岩、蛇纹岩、玄武岩和粉砂岩,因此数据集有较好的代表性。

1.3 岩体结构面粗糙度预测指标选取

本节旨在选取多个反映结构面几何形貌特性的

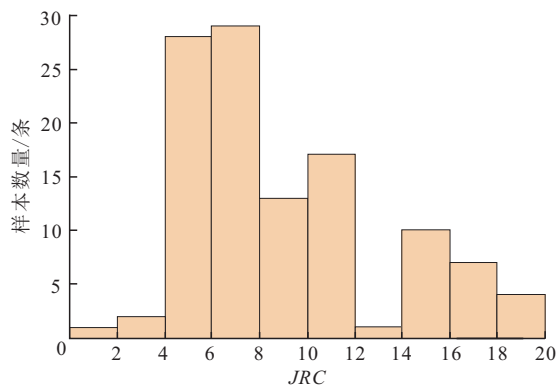


图 2 数据集中样本的 JRC 分布

Fig. 2 Distribution of the JRC values in the dataset

统计参数并构建能较全面表征岩体结构面形貌特征的指标体系。 JRC 是表征结构面粗糙度对峰值强度影响的参数,且最终要服务于岩体的力学评价,因此选取指标时应关注岩体结构面的剪切行为与结构面粗糙度之间的密切联系。由于结构面受剪切时,表面微元起伏高度越高、角度越大,突起磨损越严重^[40-41],指标体系应能够综合反映出结构面粗糙表面的起伏角度与起伏高度、起伏分布的形貌。本研究所选的岩体结构面形貌统计参数^[12,17-22]为:平均相对起伏高度(R_{ave})、起伏高度标准差(SD_h)、平均起伏角(i_{ave})、起伏角标准差(SD_i)、最大相对起伏高度(R_{max})、剖面线粗糙系数(R_p)、结构函数(SF)、坡度均方根参数(Z_2)。其中, R_{ave} 、 R_{max} 、 SD_h 定量表征岩体结构面起伏高度的形貌特征, i_{ave} 、 SD_i 、 Z_2 定量表征结构面起伏角度的几何形貌特征, R_p 、 SF 表征岩体结构面的总体起伏形貌特征,因此 8 种常用的统计参数构成的本研究模型的输入指标体系能够较全面地表征岩体结构面的几何形貌。8 种参数的计算公式如下:

(1) 平均相对起伏高度 R_{ave}

$$R_{ave} = \frac{h_{ave}}{L}, \quad (1)$$

$$h_{ave} = \frac{1}{N} \sum_{n=1}^{n=N} |y_n|, \quad (2)$$

式中: L 为剖面线在 x 轴方向的投影长度; N 为每条剖面线上的采样点数量; y_n 为剖面线上第 n 个点的 y 坐标; h_{ave} 为结构面剖面线的平均起伏高度。

(2) 起伏高度标准差 SD_h

$$SD_h = \left[\frac{1}{N-1} \sum_{n=1}^{n=N-1} (y_n - h_{ave})^2 \right]^{1/2}, \quad (3)$$

(3) 最大相对起伏高度 R_{max}

$$R_{max} = \frac{h_p - h_v}{L}, \quad (4)$$

式中: h_p 为剖面线上最高点的 y 坐标; h_v 为剖面线最低点的 y 坐标。

(4) 平均起伏角 i_{ave}

$$i_{ave} = \arctan \left[\frac{1}{N-1} \sum_{n=1}^{n=N-1} \left| \frac{y_{n+1} - y_n}{x_{n+1} - x_n} \right| \right], \quad (5)$$

式中: x_n 为剖面线上第 n 个点的 x 的坐标。

(5) 起伏角标准差 SD_i

$$SD_i = \arctan \left[\frac{1}{N-1} \sum_{n=1}^{n=N-1} \left(\frac{y_{n+1} - y_n}{x_{n+1} - x_n} - \tan i_{ave} \right)^2 \right]^{1/2}, \quad (6)$$

(6) 剖面线粗糙系数 R_p

$$R_p = \frac{L^*}{L} = \frac{\sum_{n=1}^{n=N-1} [(x_{n+1} - x_n)^2 + (y_{n+1} - y_n)^2]^{1/2}}{\sum_{n=1}^{n=N-1} (x_{n+1} - x_n)}, \quad (7)$$

式中: L^* 为结构面剖面线的真实迹线长度。

(7) 结构函数 SF

$$SF = \frac{1}{N-1} \sum_{n=1}^{n=N-1} (y_{n+1}^2 - y_n^2), \quad (8)$$

(8) 坡度均方根参数 Z_2

$$Z_2 = \left[\frac{1}{N-1} \sum_{n=1}^{n=N-1} \left(\frac{y_{n+1} - y_n}{x_{n+1} - x_n} \right)^2 \right]^{1/2}, \quad (9)$$

上述综合指标体系涵盖岩体结构面微小突起信息与宏观起伏信息,能较全面描述结构面的凹凸几何形貌特征。然而从统计参数的定义而言,指标体系内的指标间各自并不相互独立,即多指标的简单组合实质上存在结构面同一特征的重复表达。在统计学方面,各指标之间的信息重叠亦有所体现,表 1 总结了预测指标体系内各指标的统计学特征(均值、最大值、最小值、标准差),图 3 在对角线绘制了 JRC 以及各预测指标的分布直方图,并用散点图表征任意 2 个指标间的相关关系,二者共同反映出数据的分布特征:①各指标的分布范围差异较大;②各指标之间有显著的线性相关关系。指标体系中过多的重复信息将增大估算结果的偏差,因此,有必要对选定的形貌特征指标进行预处理,提取出其中有效信息。

表 1 形貌指标的统计学特征

Table 1 Statistical features of these morphological indicators

参数	均值	最大值	最小值	标准差
R_{ave}	0.006 9	0.036 0	0.001 1	0.005 8
SD_h/mm	0.453 0	2.585 5	0.076 4	0.405 5
$i_{ave}/(^{\circ})$	10.306 4	27.861 0	3.104 2	4.223 0
$SD_i/(^{\circ})$	16.867 3	40.323 1	5.159 1	6.166 0
R_{max}	0.033 9	0.165 3	0.006 6	0.027 0
R_p	1.032 6	1.181 3	1.002 8	0.029 6
SF/mm^2	0.011 4	0.072 6	0.000 9	0.011 7
Z_2	0.248 0	0.675 0	0.074 6	0.101 0

R_{ave} 、 R_{max} 分别为平均和最大相对起伏高度; SD_h 起伏高度标准差; i_{ave} 平均起伏角; SD_i 起伏角标准差; R_p 剖面线粗糙系数; SF 结构函数; Z_2 坡度均方根参数;下同

1.4 主成分分析

8 种统计参数表征的含义有诸多重叠之处,直

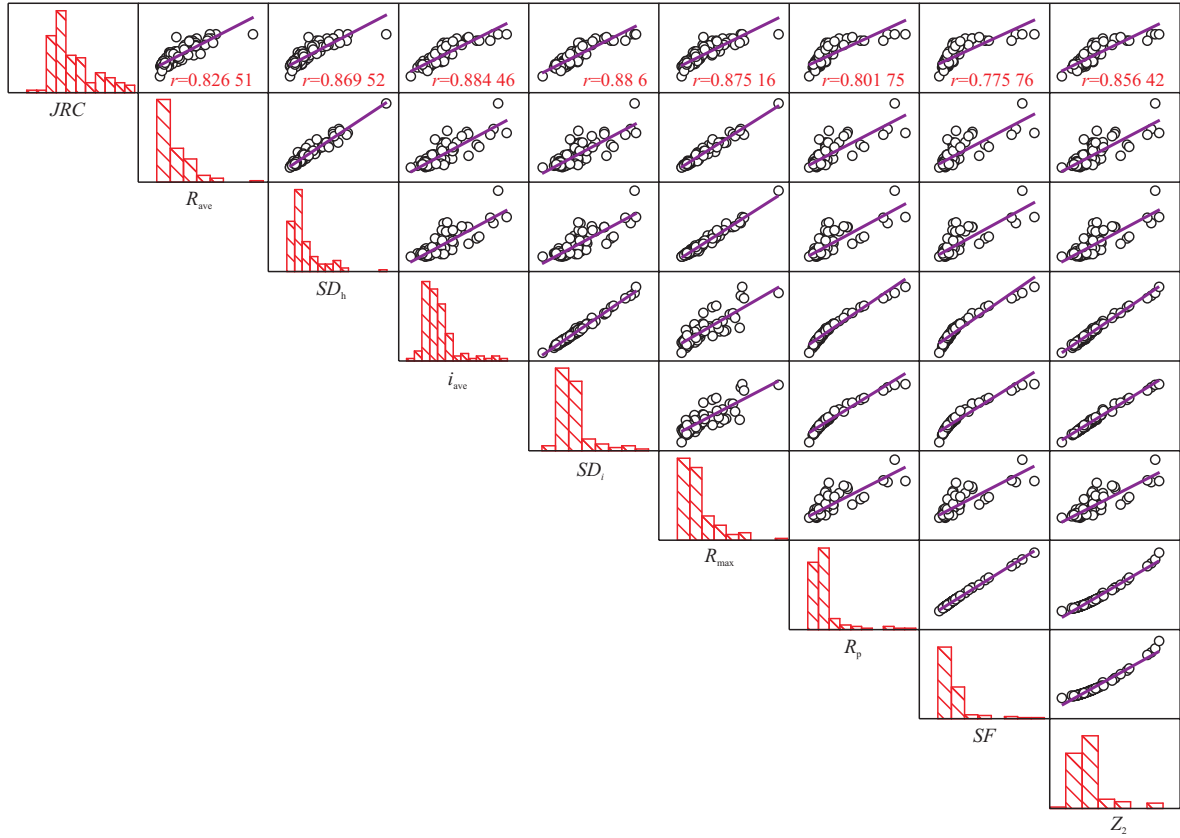


图3 岩体结构面粗糙度系数与各统计参数的直方图与散点图矩阵

Fig. 3 Scatterplot matrix of the JRC values and statistical parameters with the histogram on the diagonal

接组合有可能因为信息冗杂而削弱岩体结构面的几何特征信息,且在样本量相对较少的情况下输入维度过高会降低模型的计算效率与精度。主成分分析(PCA)^[42]作为解决复杂非线性问题的经典降维方法,能保留大部分原始结构面剖面线形貌参数信息,经此方法处理后的结构面统计参数是互不相关的复合形貌特征参数,具体步骤为:

(1) 8种统计参数组成样本经标准化处理,并用矩阵表示为:

$$\mathbf{P} = [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4, \mathbf{a}_5, \mathbf{a}_6, \mathbf{a}_7, \mathbf{a}_8]。$$

(2) 计算出样本对应协方差矩阵 $\mathbf{R}_{8 \times 8}$ 以及协方差矩阵的 q 个特征值。将特征值按从大到小的顺序表示为 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_q \geq 0$, 对应特征向量为 $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q$, 由此特征向量组成的矩阵表示为 $\mathbf{V}_{q \times 8}$ 。

(3) 由式(10)计算主成分:

$$\mathbf{S} = [\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_q]^T。$$

$$\mathbf{S} = \mathbf{V}\mathbf{P}^T, \quad (10)$$

(4) 计算成分各成分贡献率 ω_a :

$$\omega_a = \frac{\lambda_a}{\sum_{m=1}^q \lambda_m}, \quad a = 1, 2, \dots, q \quad (11)$$

式中: λ_a 为第 a 个特征向量 $\mathbf{V}_a$ 对应的特征值。

(5) 计算累计贡献率 β_t :

$$\beta_t = \sum_{a=1}^t \omega_a \geq 95\%, \quad (12)$$

累计贡献率 $\beta_t \geq 95\%$ 时则认为大部分数据信息被保留,因此选择前 t 个主成分作为模型的复合输入指标。

1.5 高斯过程回归

高斯过程回归(GPR)是一种基于贝叶斯概率框架的非参数回归分析方法。本研究的目的是通过GPR方法建立8种统计参数行向量指标 $\mathbf{X} = (R_{ave}, SD_h, i_{ave}, SD_i, R_{max}, R_p, SF, Z_2)$ 与 JRC 的非线性对应关系^[43], 利用GPR方法从训练样本中获取 $JRC = f(R_{ave}, SD_h, i_{ave}, SD_i, R_{max}, R_p, SF, Z_2)$ 的先验分布,并假设后验分布依旧服从先验的多维正态分布,比如在GPR中 JRC 的未知函数服从 $f(\mathbf{X}) \sim GP(\mu(\mathbf{X}), K(\mathbf{X}, \mathbf{X}'))$ 形式的高斯分布,其中 $f(\mathbf{X})$ 表示不考虑测量误差的 JRC 值,受均值函数 $\mu(\mathbf{X})$ 和协方差(核函数) $K(\mathbf{X}, \mathbf{X}')$ 共同控制。

由于取样偏差或者测量精度差异等问题, JRC 样本通常涉及一定程度的测量误差,可通过在训练

过程中加入噪声以表现样本的不确定误差。通常设置均值函数为0,即 $m(X^*) = 0$ 。因此为方便推导用列向量 y^* 表示 k 个训练样本中的 JRC 数据,其对应 k 组已知剖面线形貌参数指标数据 X^* 。先验分布满足均值为0,协方差为 $K(X^*, X^*) + \sigma_{\text{noise}}^2 \mathbf{I}$ 的多维高斯分布,表示为:

$$y^* \sim GP(0, K(X^*, X^*) + \sigma_{\text{noise}}^2 \mathbf{I}), \quad (13)$$

式中: σ_{noise} 为服从高斯分布的白噪声; $\mathbf{I}$ 为单位矩阵; $K(X^*, X^{*'})$ 为训练集中 X^* 自身的协方差。预测集样本中统计参数数据表示为 $X^{*'}$, f_* 表示与 $X^{*'}$ 相对应的待预测的 JRC 。 y^* 和 f_* 的联合先验分布可以表示为:

$$\begin{bmatrix} y^* \\ f_* \end{bmatrix} \sim F\left(0, \begin{bmatrix} K(X^*, X^*) + \sigma_{\text{noise}}^2 \mathbf{I} & K(X^*, X^{*'}) \\ K(X^{*'}, X^*) & K(X^{*'}, X^{*'}) \end{bmatrix}\right), \quad (14)$$

根据高斯过程的解析特性和贝叶斯理论框架,在给定 y^* 的条件下,预测值 f_* 的后验概率分布依旧满足多维高斯分布,其表达式为:

$$f_* | X^*, y^*, X^{*'} \sim N(\bar{f}_*, \text{cov}(f_*)), \quad (15)$$

式中: $\bar{f}_*$ 为待预测 JRC 的均值; $\text{cov}(f_*)$ 为待预测 JRC 的方差,具体表达式为:

$$\begin{cases} \bar{f}_* = K(X^{*'}, X^*) [K(X^*, X^*) + \sigma_{\text{noise}}^2 \mathbf{I}]^{-1} y^* \\ \text{cov}(f_*) = K(X^{*'}, X^{*'}) - K(X^{*'}, X^*) \cdot [K(X^*, X^*) + \sigma_{\text{noise}}^2 \mathbf{I}]^{-1} K(X^*, X^{*'}), \end{cases} \quad (16)$$

由式(16)可知,待预测的 JRC 准确性取决于训练样本 X^* 以及待预测集合 $X^{*'}$ 之间的相似程度,相当于通过衡量输入与训练集的相似度,确定一个相似权重并与已知 JRC 数据叠加;相似度的衡量标准是关于 $K(X^*, X^{*'})$ 的函数,表示为:

$$K(X^*, X^{*'}) = \sigma_f^2 \exp\left(-\sum_{i=1}^d \frac{(X_i^* - X_i^{*'})^2}{2\psi_i^2}\right), \quad (17)$$

式中: ψ 为方差尺度; σ_f^2 为信号方差; d 为输入的参数特征总数。输入的参数值与已有试验样本之间相似程度越高,则相似权重越大,待预测 JRC 与原有样本中实测 JRC 值越接近。

本研究中的超参数设定为 $\theta = \{\psi_i, \sigma_f^2, \sigma_{\text{noise}}^2\}$ 。超参数的选取决定了式(17)中协方差函数的值,对预测结果有很大的影响。

1.6 高斯过程回归模型超参数的极大似然估计

本研究介绍的是超参数 θ 优化时极大似然方法估计的步骤。极大似然法优化超参数的原理是找到使得边缘似然函数 $P(y^* | X^*, \theta)$ 最大的 θ 组合,即已知核函数表达式,在给定结构面形貌参数数据集

X^* 和相对应的 JRC 数据 y^* 的条件下,在训练过程中选取最优超参数集合 θ 使所有给定的训练样本 y^* 值出现的概率最大。其中,为了降低超参数初值对预测结果的影响, θ 的初值是用多次试算的方法确定的。

根据贝叶斯理论,在给定超参数 θ 的条件下,获得预测数据 D 的可能性为 $P(y^* | \theta)$,表示为:

$$P(y^* | \theta) = \frac{1}{\sqrt{(2\pi)^d \det(K(X_y^*, X_y^*) + \sigma_{\text{noise}}^2 \mathbf{I})}} \cdot \exp\left(-\frac{y^*(K(X_y^*, X_y^*) + \sigma_{\text{noise}}^2 \mathbf{I})^{-1} y^*}{2}\right), \quad (18)$$

由于 θ 是非负超参数,通常对训练样本的负对数似然函数 $L(\theta)$ 求偏导,通过寻找偏导的最小值来获取超参数 θ 的最优值,即:

$$L(\theta) = -\frac{1}{2} y^T K(X^*, X^*)^{-1} y - \frac{1}{2} \ln |K(X^*, X^*)| - \frac{n}{2} \ln 2\pi, \quad (19)$$

因此可以引入拟牛顿法(Quasi-Newton methods)迭代求取 θ 的最优组合,并代入式(16)以实现 JRC 的预测。

2 岩体结构面粗糙度预测

2.1 数据预处理

数据集中样本共有112组,采用随机抽取的95组样本对模型进行训练,划分样本集时采用多次随机试验调整训练样本分布以保证训练集具有典型性,训练集中部分典型岩体结构面样本的二维剖面线形态如图4所示。选取的统计参数量纲不同,为避免数量级差异而对结果产生误差,有必要在构建

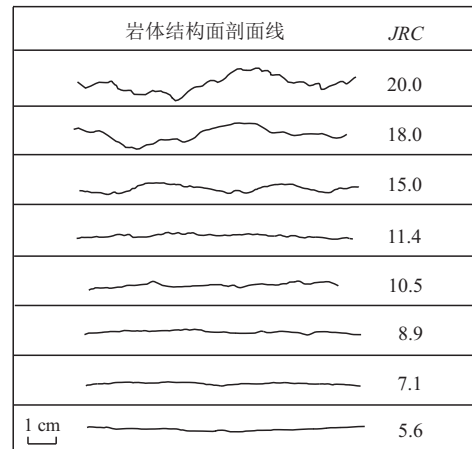


图4 训练集中典型岩体结构面样本的剖面线

Fig.4 Profiles of typical rock joint samples in the training set

模型前对统计参数的数据进行标准化预处理,其公式如下^[44]:

$$x_{ij}^* = \frac{x_{ij}^* - x_j^{*\min}}{x_j^{*\max} - x_j^{*\min}}, \quad (20)$$

式中: x_{ij}^* 为标准化后第 m 个样本的第 j 项统计参数; $x_j^{*\max}$ 为第 j 项统计参数的最大值; $x_j^{*\min}$ 为第 j 项统计参数的最小值,经标准化处理后的统计参数范围落在区间 $[0, 1]$ 内;其中, $j=1, 2, 3, 4, 5, 6, 7, 8$, $m=1, 2, \dots, 112$ 。

数据标准化后,通过 PCA 方法对 8 个特征参数降维处理,为防止数据污染,此步骤中验证样本集和训练样本集的降维处理是单独进行的,其中验证样本集的降维由训练样本集降维过程中获取的特征向量矩阵实现。

如图 5 所示,训练集中各因子对方差的贡献率是对原始信息表达程度的衡量,前 2 个因子的方差贡献率分别是 90.312%, 6.392%, 累计超过 95% 的输入信息。因此认为选取前 2 个复合特征作为输入指标已经可以表达大部分几何参数信息,且大幅降低了输入维度。

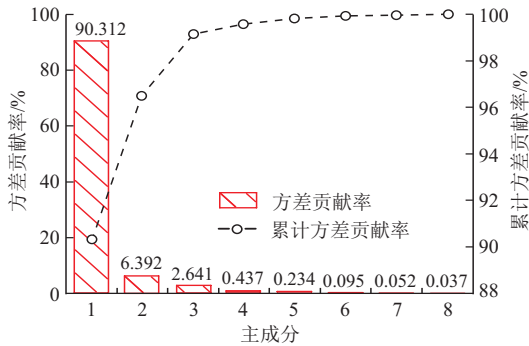


图 5 主成分方差贡献率分布图

Fig. 5 Distribution of the variance contribution of the principal components

由此确定的 2 个复合指标 $PC1$ 、 $PC2$ 在几何含义上依旧表征结构面起伏高、起伏角的复合特点,降维后的指标如下所示:

$$\begin{cases} PC1 = 0.32R_{ave} + 0.31SD_h + 0.37i_{ave} + 0.40SD_i + \\ \quad 0.34R_{max} + 0.35R_p + 0.34SF + 0.38Z_2 \\ PC2 = 0.49R_{ave} + 0.46SD_h - 0.21i_{ave} - 0.23SD_i + \\ \quad 0.47R_{max} - 0.28R_p - 0.29SF - 0.25Z_2 \end{cases}, \quad (21)$$

2.2 高斯过程回归 JRC 模型构建

基于上述 2 个复合输入因子,利用训练样本集构建 JRC 模型,通过拟牛顿法优化得到最优 $\{\psi_i\}$ 为

$\psi_1 = 1.976, \psi_2 = 5.397$ 。 ψ 的含义为方差尺度,反映高斯过程中变量之间的相关程度。优化后 ψ 的大小可解译因子分别在高斯过程建模中的重要性,从式(17)上看较小的 ψ 对协方差的值影响更为显著,即实质上模型中的点与点相互关系由 $\{\psi_i\}$ 中较小的值控制。由训练结果可知, ψ_1 (对应输入因子 $PC1$) 比 ψ_2 (对应输入因子 $PC2$) 相对更小,说明因子 $PC1$ 对于模型预测的贡献更大,与图 5 方差贡献率分布图中表示因子 $PC1$ 涵盖的起伏高度与起伏角度信息较因子 $PC2$ 更多一致,从模型角度侧面佐证了本研究“降维处理后因子仍然保留原本表征结构面粗糙形态的几何含义”的观点。

2.3 高斯过程回归 JRC 模型验证

为了验证模型的预测效果,用验证样本集中的 17 组岩体结构面剖面线样本数据对模型的泛化能力进行验证,预测获得的 JRC 与实际的 JRC 对比如图 6a 所示。为了更好地说明预测结果,图 6b 中绘制了预测点的对角线误差图,预测样本点越接近对角线说明误差越小,所有的预测点位于上下边界区间内说明误差在 20% 以内。由图 6 可知, GPR 模型预测的 JRC 值在绝大部分情况下与实验获取的 JRC 非常接近,高斯过程方法适用于小样本条件下的多指标 JRC 值预测。验证集中 JRC 的预测值与实验值的平均误差为 0.435, 决定系数为 0.972, 均方根误差为 0.517, 这说明模型具有较高的精确度。

为了进一步说明模型的准确性,基于验证样本集,在图 7 中绘制由 R_p 、 SF 、 Z_2 公式计算获得的 JRC 值折线图, R_p 、 SF 、 Z_2 的 JRC 计算公式^[18, 22]如下:

$$JRC = 441(R_p - 1), \quad (22)$$

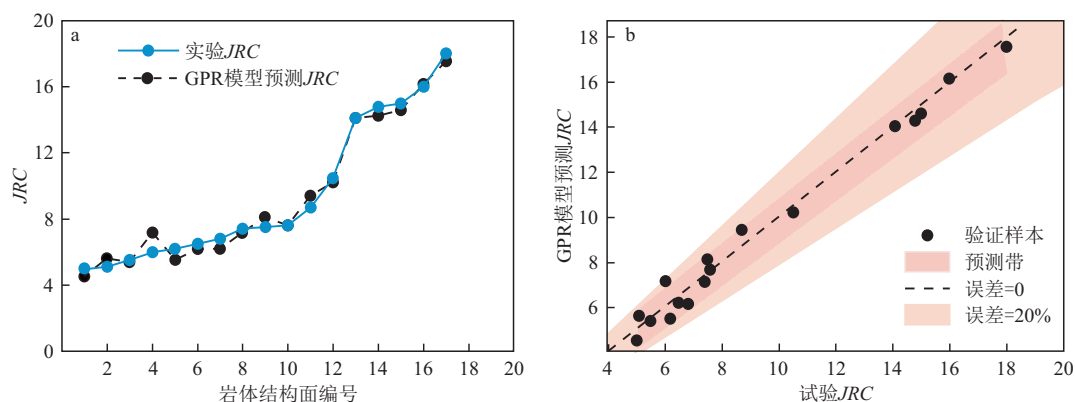
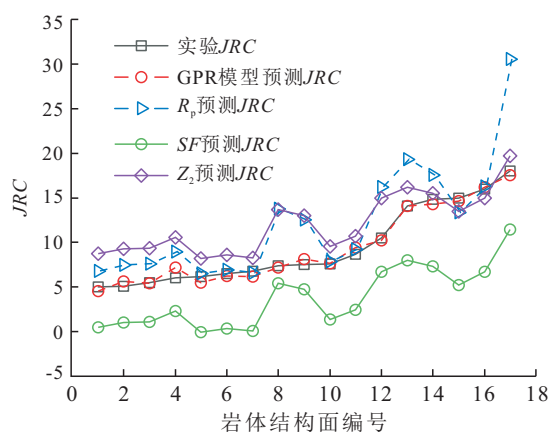
$$JRC = 37.63 + 16.51 \lg SF, \quad (23)$$

$$JRC = 32.69 + 32.98 \lg Z_2, \quad (24)$$

显而易见,由单一统计参数的经验公式(22)~(24)获得的 JRC 值与本研究 JRC 计算的方法相比,本研究提出的模型与实测值更相近。这实质上证实了“单一统计参数表达的信息很难全面反映出真实岩体结构面的粗糙形态”这一观点,表明在实际情况下利用单一统计参数与 JRC 之间构建的回归公式预测 JRC 有难度,而本研究提出的模型对于真实 JRC 值预测更加有优势。

3 结 论

(1) 利用主成分分析法在大程度保留原始起伏高度、角度形态信息的同时,将 8 种复杂相关的特征

图 6 验证样本集预测 JRC 与实际 JRC 值对比Fig. 6 Comparison between the predicted and measured JRC values based on the validation database图 7 基于 GPR 模型的 JRC 预测与单参数预测对比Fig. 7 Comparison of JRC values prediction based on the GPR method with various single-parameter models

参数转化为 2 种复合输入参数,实现数据间有效信息的提取,提升了计算效率。

(2) 本研究由高斯过程回归构建的 JRC 模型预测的决定系数高达 0.972,均方根误差为 0.517,反映出高斯过程回归方法对于小样本的复杂多参数 JRC 值预测的适用性。同时,高斯过程回归不需人工确定各参数的权重,实现智能化预测 JRC ,展现了高斯过程模型的数据驱动优势。

(3) 通过计算岩体结构面的形态参数与 JRC 之间的相关性,刻画了数据之间的联系。本研究模型解译了 JRC 与降维后变量之间的依赖关系,是复合指标有效表征结构面起伏高度、起伏角度的重要佐证。

(4) 本研究构建的预测模型在划分样本集时,采用多次随机试验的方法调整训练样本的分布,使训练集尽可能符合真实岩体结构面的数据分布,未探讨训练数据集中不同种类岩体结构面样本组成对

模型预测性能的影响,应在之后的研究中进行补充。

所有作者声明不存在利益冲突。

The authors declare that no competing interests exist.

参考文献:

- [1] 唐辉明. 工程地质学基础[M]. 北京:化学工业出版社,2008.
TANG H M. Fundamentals of engineering geology [M]. Beijing: Chemical Industry Press,2008. (in Chinese with English abstract)
- [2] ZHANG X B,JIANG Q H,CHEN N,et al. Laboratory investigation on shear behavior of rock joints and a new peak shear strength criterion[J]. Rock Mechanics and Rock Engineering,2016,49(9): 3495-3512.
- [3] 唐雅婷,谭杰,李长冬,等. 基于模型试验的动水驱动型顺层岩质滑坡启滑机制初探[J]. 地质科技通报,2022,41(6): 137-148.
TANG Y T,TAN J,LI C D,et al. Preliminary study on the initiation mechanism of hydrodynamic-driven bedding rock landslides based on physical model tests[J]. Bulletin of Geological Science and Technology,2022,41(6): 137-148. (in Chinese with English abstract)
- [4] 谢妮,王丁浩,吕阳,等. 酸腐蚀作用下川渝红层砂岩蠕变特性试验研究[J]. 地质科技通报,2022,41(5): 141-149.
XIE N,WANG D H,LÜ Y,et al. Experimental study on creep behaviors of red sandstone in Sichuan and Chongqing under acid corrosion[J]. Bulletin of Geological Science and Technology,2022,41(5): 141-149. (in Chinese with English abstract)
- [5] LIU Q S,TIAN Y C,LIU D F,et al. Updates to JRC-JCS model for estimating the peak shear strength of rock joints based on quantified surface description[J]. Engineering Geology,2017,228: 282-300.
- [6] SINGH H K,BASU A. Evaluation of existing criteria in estimating shear strength of natural rock discontinuities[J]. Engineering Geology,2018,232: 171-181.
- [7] 宋磊博,亢倩倩,杜时贵,等. 基于表面微凸体磨损特征演化规律的节理剪切强度各向异性机制研究[J]. 岩土力学,2021,42

- (9):2331-2343.
- SONG L B, KANG Q Q, DU S G, et al. Anisotropy mechanism of shear strength based on wear and shear failure evolution of asperities of joint surface[J]. Rock and Soil Mechanics, 2021, 42(9): 2331-2343. (in Chinese with English abstract)
- [8] 杨超, 蒋洪南, 岳六一, 等. 法向卸荷条件下非贯通节理岩体剪切力学特性试验研究[J]. 岩土工程学报, 2022, 44(9): 1741-1750.
- YANG C, JIANG H N, YUE L Y, et al. Shear mechanical properties of rock mass with discontinuous joints under unloading normal stress[J]. Chinese Journal of Geotechnical Engineering, 2022, 44(9): 1741-1750. (in Chinese with English abstract)
- [9] BARTON N, CHOUBEY V. The shear strength of rock joints in theory and practice[J]. Rock Mechanics, 1977, 10(1): 1-54.
- [10] BARTON N. Review of a new shear-strength criterion for rock joints[J]. Engineering Geology, 1973, 7(4): 287-332.
- [11] 肖维民, 刘伟超, 朱占元. 基于正反向直剪试验的岩石节理 *JRC* 值计算研究[J]. 岩石力学与工程学报, 2021, 40(增刊2): 3098-3106.
- XIAO W M, LIU W C, ZHU Z Y. Study on calculation of rock joint *JRC* value based on forward and reverse direct shear tests[J]. Chinese Journal of Rock Mechanics and Engineering, 2021, 40(S2): 3098-3106. (in Chinese with English abstract)
- [12] LEE S D, LEE C I, PARK Y. Characterization of joint profiles and their roughness parameters[J]. International Journal of Rock Mechanics and Mining Sciences, 1997, 34(3/4): 174. e1-174. e5.
- [13] TANG Z C, JIAO Y Y, WONG L N Y, et al. Choosing appropriate parameters for developing empirical shear strength criterion of rock joint: Review and new insights[J]. Rock Mechanics and Rock Engineering, 2016, 49(11): 4479-4490.
- [14] GE Y F, KULATILAKE P H S W, TANG H M, et al. Investigation of natural rock joint roughness[J]. Computers and Geotechnics, 2014, 55: 290-305.
- [15] 杜时贵, 黄曼, 罗战友, 等. 岩石结构面力学原型试验相似材料研究[J]. 岩石力学与工程学报, 2010, 29(11): 2263-2270.
- DU S G, HUANG M, LUO Z Y, et al. Similar material study of mechanical prototype test of rock structural plane[J]. Chinese Journal of Rock Mechanics and Engineering, 2010, 29(11): 2263-2270. (in Chinese with English abstract)
- [16] 范斌强. 基于精细雕刻的岩体结构面峰值抗剪强度模型及工程应用研究[D]. 武汉: 中国地质大学(武汉), 2021.
- FAN B Q. Research on the peak shear strength model of rock joint based on fine carving and its engineering application[D]. Wuhan: China University of Geosciences(Wuhan), 2021. (in Chinese with English abstract)
- [17] TSE R, CRUDEN D M. Estimating joint roughness coefficients[J]. International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts, 1979, 16(5): 303-307.
- [18] YANG Z Y, LO S C, DI C C. Reassessing the joint roughness coefficient (*JRC*) estimation using Z_2 [J]. Rock Mechanics and Rock Engineering, 2001, 34(3): 243-251.
- [19] SAYLES R S, THOMAS T R. The spatial representation of surface roughness by means of the structure function: A practical alternative to correlation[J]. Wear, 1977, 42(2): 263-276.
- [20] ZHANG G C, KARAKUS M, TANG H M, et al. A new method estimating the 2D joint roughness coefficient for discontinuity surfaces in rock masses[J]. International Journal of Rock Mechanics and Mining Sciences, 2014, 72: 191-198.
- [21] YU X B, VAYSSADE B. Joint profiles and their roughness parameters[J]. International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts, 1991, 28(4): 333-336.
- [22] MAERZ N H, FRANKLIN J A, BENNETT C P. Joint roughness measurement using shadow profilometry[J]. International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts, 1990, 27(5): 329-343.
- [23] 陈世江, 赵自豪, 王超. 基于修正线粗糙度法的岩石节理粗糙度估值[J]. 金属矿山, 2012(6): 22-25.
- CHEN S J, ZHAO Z H, WANG C. Estimation of rock joint roughness based on modified line-roughness[J]. Metal Mine, 2012(6): 22-25. (in Chinese with English abstract)
- [24] 李化, 黄润秋. 岩石结构面粗糙度系数 *JRC* 定量确定方法研究[J]. 岩石力学与工程学报, 2014, 33(增刊2): 3489-3497.
- LI H, HUANG R Q. Method of quantitative determination of joint roughness coefficient[J]. Chinese Journal of Rock Mechanics and Engineering, 2014, 33(S2): 3489-3497. (in Chinese with English abstract)
- [25] 蔡毅, 唐辉明, 葛云峰, 等. 岩体结构面三维粗糙度评价的新方法[J]. 岩石力学与工程学报, 2017, 36(5): 1101-1110.
- CAI Y, TANG H M, GE Y F, et al. A new method for evaluating the roughness of three-dimensional discontinuity surface of rock[J]. Chinese Journal of Rock Mechanics and Engineering, 2017, 36(5): 1101-1110. (in Chinese with English abstract)
- [26] HONG E S, LEE J S, LEE I M. Underestimation of roughness in rough rock joints[J]. International Journal for Numerical and Analytical Methods in Geomechanics, 2008, 32(11): 1385-1403.
- [27] 陈世江, 朱万成, 王创业, 等. 岩体结构面粗糙度系数定量表征研究进展[J]. 力学学报, 2017, 49(2): 239-256.
- CHEN S J, ZHU W C, WANG C Y, et al. Review of research progresses of the quantifying joint roughness coefficient[J]. Chinese Journal of Theoretical and Applied Mechanics, 2017, 49(2): 239-256. (in Chinese with English abstract)
- [28] 马泽栋, 马雷, 李科, 等. 基于岩石图像深度学习的多尺度岩性识别[J]. 地质科技通报, 2022, 41(6): 316-322.
- MA Z D, MA L, LI K, et al. Multi-scale lithology recognition based on deep learning of rock images[J]. Bulletin of Geological Science and Technology, 2022, 41(6): 316-322. (in Chinese with English abstract)
- [29] 宋琨, 孙驰, 安冬, 等. 数字钻孔全景影像中结构面特征智能识别方法[J]. 地质科技通报, 2020, 39(5): 17-22.
- SONG K, SUN C, AN D, et al. Intelligent identification method for rock discontinuities properties by digital borehole panoramic images[J]. Bulletin of Geological Science and Technology, 2020, 39(5): 17-22. (in Chinese with English abstract)
- [30] 宋康明, 姜阳厚, 谭志祥, 等. 基于随机森林方法的岩石节理粗糙度系数研究[J]. 地质科技情报, 2018, 37(3): 263-267.
- SONG K M, JIANG Y H, TAN Z X, et al. Method to calculate the

- joint roughness coefficient based on random forest[J]. Geological Science and Technology Information, 2018, 37(3): 263-267. (in Chinese with English abstract)
- [31] WANG L Q, WANG C S, KHOSHNEVISAN S, et al. Determination of two-dimensional joint roughness coefficient using support vector regression and factor analysis[J]. Engineering Geology, 2017, 231: 238-251.
- [32] 苗发盛, 吴益平, 李麟玮, 等. 基于 Boosting-决策树 C5.0 的岩体结构面粗糙度预测[J]. 浙江大学学报(工学版), 2021, 55(3): 483-490.
- MIAO F S, WU Y P, LI L W, et al. Prediction of joint roughness coefficient of rock mass based on boosting-decision tree C5.0[J]. Journal of Zhejiang University (Engineering Science Edition), 2021, 55(3): 483-490 (in Chinese with English abstract)
- [33] 尹宏, 王述红, 董卓然, 等. 引入因子分析的结构面粗糙度 RBF 复合参数模型[J]. 岩土工程学报, 2022, 44(4): 721-730.
- YIN H, WANG S H, DONG Z R, et al. RBF composite parameter model for structural surface roughness with factor analysis[J]. Chinese Journal of Geotechnical Engineering, 2022, 44(4): 721-730. (in Chinese with English abstract)
- [34] RASMUSSEN C E, WILLIAMS C K I. Gaussian processes for machine learning[M]. Cambridge, Mass.: MIT Press, 2006.
- [35] 赵腾远, 宋超, 何欢. 小样本条件下江苏软土路基回弹模量的贝叶斯估计: 基于静力触探数据与高斯过程回归的建模分析[J]. 岩土工程学报, 2021, 43(增刊2): 137-141.
- ZHAO T Y, SONG C, HE H. Bayesian estimation of resilient modulus of Jiangsu soft soils from sparse data: Gaussian process regression and cone penetration test data-based modelling and analysis[J]. Chinese Journal of Geotechnical Engineering, 2021, 43(S2): 137-141. (in Chinese with English abstract)
- [36] 宋超, 赵腾远, 许领. 基于贝叶斯高斯过程回归与模型选择的岩石单轴抗压强度估计方法[J]. 岩土工程学报, 2023, 45(8): 1664-1673.
- SONG C, ZHAO T Y, XU L. Estimation of uniaxial compressive strength based on fully Bayesian Gaussian process regression and model class selection[J]. Chinese Journal of Geotechnical Engineering, 2023, 45(8): 1664-1673. (in Chinese with English abstract)
- [37] GRASSELLI G. Shear strength of rock joints based on quantified surface description[C]//OHNISHI Y, AOKI K. ISRM International Symposium/3rd Asian Rock Mechanics Symposium (ARMS). Rotterdam, Netherlands: Millpress Science Publisherspo Box, 2001.
- [38] BANDIS S C, LUMSDEN A C, BARTON N R. Fundamentals of rock joint deformation[J]. International Journal of Rock Mechanics and Mining Sciences, 1983, 20(6): 249-268.
- [39] STAVROS B. Experimental studies of scale effects on shear strength, and deformation of rock joints[D]. Leeds, Yorkshire and the Humber, UK: University of Leeds, 1980.
- [40] YANG Z Y, DI C C, YEN K C. The effect of asperity order on the roughness of rock joints[J]. International Journal of Rock Mechanics and Mining Sciences, 2001, 38(5): 745-752.
- [41] GENTIER S, RISS J, ARCHAMBAULT G, et al. Influence of fracture geometry on shear behavior[J]. International Journal of Rock Mechanics and Mining Sciences, 2000, 37(12): 161-174.
- [42] 董福瑞, 王述红, 侯钦宽. 基于主成分分析的多参数岩体结构面优势分组方法研究[J]. 岩土力学, 2022, 43(9): 2457-2470.
- DONG F R, WANG S H, HOU Q K. Multi-parameter dominant grouping method of rock mass discontinuity based on principal component analysis[J]. Rock and Soil Mechanics, 2022, 43(9): 2457-2470. (in Chinese with English abstract)
- [43] HAN L, WANG L, DING X M, et al. Similarity quantification of soil parametric data and sites using confidence ellipses[J]. Geoscience Frontiers, 2022, 13(1): 101280.
- [44] 陈朝亮, 彭树宏, 钱静, 等. 基于 AHP-Logistic 熵权模型的西南浅丘区地质灾害分布特征研究: 以内江市为例[J]. 长江科学院院报, 2020, 37(2): 55-61.
- CHEN C L, PENG S H, QIAN J, et al. Distribution characteristics of geological hazards in southwestern shallow hill based on AHP-Logistic entropy combined weight model: A case study of Neijiang city[J]. Journal of Yangtze River Scientific Research Institute, 2020, 37(2): 55-61. (in Chinese with English abstract)

Summary for “基于高斯过程回归的岩体结构面粗糙度系数预测模型”

A prediction model of the joint roughness coefficient based on Gaussian process regression

ZHENG Kexin¹, WU Yiping^{1*}, LI Jiang², MIAO Fasheng¹, KE Chao¹

1. Faculty of Engineering, China University of Geosciences (Wuhan), Wuhan 430074, China;

2. Information Center, Department of Natural Resources of Hubei Province, Wuhan 430071, China

First author, E-mail: zhx1998@cug.edu.cn

* Corresponding author, E-mail: ypwu@cug.edu.cn

Abstract: [**Objective**] Estimating the joint roughness coefficient (JRC) is essential for evaluating the mechanical properties of a rock mass. Due to the limitation of a single statistical parameter for characterizing morphology, JRC values estimation by a single statistical parameter may produce a sufficiently unreliable result. [**Methods**] To address the existing challenges in determining JRC values, a model based on Gaussian process regression (GPR) combined with principal component analysis (PCA) was proposed for the quantitative evaluation of JRC. Notably, eight parameters were selected as indicators for the comprehensive expression of the rock joint roughness. To analyse the model's performance, a publicly available dataset of 112 rock joint profiles was used as an example, of which 95 were chosen as training samples and 17 were chosen as validation samples. The reliability of the model was verified by comparing the predicted results with the measured JRC values. [**Results**] The results show that the derived GPR model demonstrates promising performance ($R^2=0.972$, $MSE=0.517$) for estimation of JRC values, indicating the high applicability of the model in constructing implicit relationships between multiple statistical parameters and JRC values even under small sample conditions. [**Conclusion**] In general, the GPR model may provide a new way of estimating JRC values with artificial intelligence.

Key words: rock joints; roughness; Gaussian process regression; statistical parameter; prediction

Received: 2023-03-06; Revised: 2023-04-18; Accepted: 2023-04-19

doi:10.19509/j.cnki.dzkg.tb20230113