S1−14

Comparative Study of Improved MGGP and SVR Models for Spatial Grade Estimation

Ri Kwang Su, Han Chang Ik*

Faculty of Resources Exploration, Kim Chaek University of Technology, Pyongyang, DPRK
*Corresponding author: Email: rgs67514@star−co. net. kp

Accurate estimation of spatial grade is one of the most important aspects of assessing mineral resources and reserves. The aim of this study is to compare support vector regression (SVR) and improved multi−gene genetic programming (MGGP) methods in spatial grade estimation. We noticed that MGGP ignored spatial structural features reflecting geological process, which is very difficult to be found by the MGGP itself, for grade prediction. To solve this problem, we introduced and embed the spatial variation evaluation term in the paradigm of MGGP. And then the performance of each model participating in the evolutionary process is evaluated using the following criterion given by:

$$criterion = \sqrt{\frac{\sum_{i=1}^{N}\left|G_i - Z_i\right|^2}{N}} \cdot \sqrt{\frac{\sum_{j=1}^{M}\left|\frac{\gamma_j^G - \gamma_j^Z}{\gamma_j^Z}\right|^2}{M}}$$

Where the first term is the root mean square error of MGGP model on the training data and the second term is the spatial variation evaluation term, in which G and Z are the variograms computed by the predicted and actual observed values in the j−th pair of samples, respectively, and M is the number of sample pairs used in variogram computation. The spatial variation evaluation term, which is the basis of the proposed method, is derived through spatial structure analysis corresponding to the first step of kriging. Above criterion multiply the evaluation term about spatial variation to the MGGP model error and thereby penalizes the fitness of models ignored spatial structure features and prevents over−fitting. In this way, the proposed method improves prediction accuracy and generalization ability of models.

The steps generally followed in improved MGGP are:

ⅰ) Construct spatial variation model through the spatial structure analysis.
ⅱ) Create an initial population of individuals (i.e. programs or equations).
ⅲ) Define criterion using spatial variation evaluation term.
ⅳ) Evaluate individuals based on above criterion.
ⅴ) Select the fittest individuals as parents.
ⅵ) Create new individuals (also called the offspring) through the genetic operations of crossover, mutation, and reproduction.
ⅶ) Replace the weaker parents in the population by the stronger ones.
ⅷ) Repeat steps 2 through 5 until the user defined termination criterion is satisfied. The termination criterion can be completion of a specified number of generations or fitness criterion such as step 2 reached.

We demonstrate the performance of improved MGGP method compared with SVR in spatial grade estimation. For this illustration, we have selected data collected from an iron ore deposit located in the western part of DPRK. We randomly selected 152 boreholes as training data for model estimation and used the rest 44 boreholes as validating data for cross−validation. The data of orebody thickness are known to have a spherical variograms of nested structure with a range of 64.04m, a sill of 2.01 and nugget effect of 55.73. The best MGGP model selected based on minimum fitness on training data is used to cross validation, which is as follows:

(x, y)=plog(atan(psqrt(−sin(x*0.2062+0.3375))))*−3.1649−sin(y*0.1318−atan(y*6.4079)*0.1318+psqrt(7.6425)*sin(x)*0.1318−sin(x*y)*sin(y−3.8033)*0.1318+0.6244)*3.7449−sin(y*−6.1779+psqrt(x)+mypower(x, 1.2523))*2.889−cos(plog(sin(psqrt(y+1.6363)))/tanh(atan(tanh(y*6.1779))−cos(x)*1.3136))*5.0940+cos(square(mypower(cos(psqrt(y)), cos(y)))−psqrt(−x*(y−4.8986))*0.4116)*3.165+psqrt(cos(y+12.1835)+0.9815)*2.8890+mypower(psqrt(sin(x*6.1918))*exp(atan(psqrt(plog(x)))), square(−sin(psqrt(x))*sin(y+6.1779)+atan(tanh(y))*mypower(sin(y), x*2.0))+square(mypower(square(mypower(cos(y), tanh(y))), psqrt(sin(y*5.9141)))))*2.1689+mypower(psqrt(sin(psqrt(sin(x*y))))*psqrt(4.9842), square(square(square(sin(psqrt(x)))))+square(y*mypower(sin(y), x+1.6363)*6.4079−sin(sin(y))^2))*6.6120+10.9099

Correlation coefficient (R), mean absolute prediction error (MAPE) and root mean square prediction error (RMSPE) are used to evaluate the validation result of two models for grade prediction.

We can also know that MAPE and RMSPE of improved MGGP model are smaller than SVR and that R higher, indicating that proposed model is reliable in predicting the spatial grade. In addition, MGGP method overcomes the defect of SVR that does not provide explicit formulation for grade prediction model.