Study on Geological Space Variables Function

Wang Na*
Jilin Engineering Normal University
Changchun Jilin 130052,China
E-mail: arnold0110@sina.com
* Corresponding Author

Yang Zhen Yu
Jilin Provincial Bureau of Statistics
Jilin 130000,China

Zhang Na
College of Art and Design
Shenyang Jianzhu University
ShenYang, China

Abstract-Geological variables are the basic units to carry information and the basic material of the construction of the geological spatial modeling . The quality of geological space model depends on the quality of the model of geological variables , so the geological variables are the precondition of stability and effect the geological space model . The role of geological variables focuses on the function of the geological variables . This paper first introduces the scientific nature of geological variables function research.Then discusses the mathematical models that fit the five functions , Finally, discuss the process and technique method for the extraction of geological variable. The purpose is to provide a series of general rule and method for geological modeling.

Keywords-Geological modeling , Geological variables function , Universal standards.

I. PROBLEM PRESENTATION

In the quantitative prediction of Geology and mineral resources in mathematics, the variables that general geological signs or geological phenomena with the formation age, location and property changes and take different values , called the geological variables . Research on geological variables is a key content . Geological variable refers to the geological time and space observations of various geological , geophysical , geochemical , geography and other related digital symbol set . We define the all the effectiveness in the model as the geological variables function that is variable capacity in the overall evaluation model . The strong and the weak of functions of geological variables , advantages and disadvantages determine the quality and effect of the quantitative model.As we known , quantitative prediction of mineral resources often involves model to the unknown area ,its core content is the quantitative prediction model of structural stability . The optimal function equation problem depends on the geological variables .

II. THE SCIENTIFIC NATURE OF GEOLOGICAL VARIABLES FUNCTION RESEARCH

The essence of geological variables function research mainly includes the following three parts:

The first. In mathematical geology field , the final result is achieved by processing the original data the concentrated geological information in order to extract useful information and clearly depict geological target . High concentration of geological information directly reflected in the variable set research , correlation , information and the target of geological variables of the information carried by the direct impact on the quality of geological information extraction and screening .

The second . In mineral resources evaluation and prediction, reliability and stability of prediction model depends on the degree of perfection of geological variables function and level . In general , the the construction of quantitative prediction model is based on the geological variable support completely by the optimal combination of variables . The optimal combination of variables at various variables function entirely compose of bearing , variable function is widely variable , the effect is more obvious ,the quality of the optimal combination of variables is higher , so the more robust quantitative prediction model .

The third . The level of detail geological variable role orientation clear or not depends on the geological variable function division , five division of function of geological variables cover all aspects of the geological space information symbol and reflects the overall relationship between the geological variables and objectives .Therefore the geological variable function detailed discussion is helpful to improve the geological target recognition ability .

III. THE ROLE OF GEOLOGICAL VARIABLES AND THE ROLE OF NATURE

The geological variables are not sure about the geological variables, which are not sure. In fact, a geological evolution history, and thegeological variables are related or related, and the function of the variables is characterized by the interaction between the two sides. The function of the variables is determined. The action direction of the variable is the choice tendency or the behavior in the process of variable correlation. The function of the variable is the interaction between the variables, which is reflected by some dynamic constraints, but not the nature of the traditional meaning, which is not divided into two aspects.

The function of variable is the variable group or combination of variables, and the relationship between the variables and the target is not by any one or some other variables. In other words, the variable is related to the interaction and coordination, and the interaction between them is different, and sometimes the difference between different scales and different combinations.

In fact, the variable isolation can not explain the characteristics of them. For example: Ar - Archean is a gold deposit is of great significance to the variable, but not only by its appearance that can indicate the mineralization, but also needs to have the structure, magmatic heat sources

only. Variables involved in the Archean not only for gold good for iron ore and other mineral deposits will be favorable. Ar and other variables are coordinated, can form

a complete set of variables. Variables in different form, the nature and function of the direction is different, sometimes very different. Such as deposit, in small scale, the result different size of the deposit, each of the best combination of different variables. Belong to the old stratum Ar (Archean) are the variables, when placed in different combinations, sometimes indicating large ore, sometimes indicating small mines exist. Therefore, a combination of variables Organic whole at a level.

The meaning of existence, the variables are independent, random; its relationship with the study objective is to connect with each other. The research object is divided into geological grade, such as ore bearing units with ore deposits and ore bodies of different levels, the corresponding variable combination is divided into levels. Some variables may also apply to the role of different levels in the local sense, but the combination of variables on their own level is strictly different.

The three kinds of geological variables are related to each other, which can be attributed to the following three properties:

The equivalence number range, examine the relationship between the two two variables, when a relationship between variable X and variable Y satisfies reflexivity, symmetry, transitivity axiom three, said the relationship between the variables is the equivalence relation is reflexive. Here refers to the variable X and Y itself have associated symmetry; refers to when the variable X connecting to Y Y in the X case, also have connection function. The equivalence relation is widespread, is the most thorough discussion. The correlation coefficient of the traditional definition of the similarity coefficient of the statistics is the equivalence of the index.

A relationship between the non equivalence of two two is sometimes not satisfies three axioms of equivalence relations, which only satisfy one or sometimes two axioms, which at this time, this relationship is called non equivalence. Here is divided into several situations: (1) does not meet the reflexivity. This means that the variable itself does not produce association effect. Such as: deposit value is not associated with the formation of itself. In this case, the main diagonal elements of variable incidence matrix is zero. (2) the relationship between the variable does not satisfy the symmetry, not by X of the Y Association Y also launched related to X. This means that a variable in direction is one-way thus, the relationship is not reversible. (3) does not meet the transfer. The relationship between variables that can not transfer the information transfer is not unconditional. When the relationship between variables with reflexivity and symmetry and has not passed, if we consider the classification problem, should first according to the transitive closure The theory carries on the synthesis processing, which makes it into an equivalent relationship, and can be classified.

Variables were associated with other variables, which is the number of variables associated with other variables, such as: interactive. In this case, the relevant variables are defined: the variables are defined by the correlation variables. In this case, that is, if the variables are similar. As a result of the variables, variables such as: correlation. In this case, the measure variables. The relationship between the value of the ore deposit and the geological variable is a partial correlation. The complex relationship, which requires that all aspects of the variables should be considered in the process of modeling.

IV. THE EXTRACTION PROCESS OF GEOLOGICAL VARIABLES

Geological variables contain geographic information in the different scale , different time and space , different in content and form , including both variable attribute information and graphics information and including the related information between the variables . The latter is often concerned . Once variable decides , it becomes the carrier of information but it sometimes can not be directly used . Geological information need to improve the utilization value by conversion . The transformation of variables or information extraction is a very important link .

Variables are obtained from raw data, but generally are not equivalent to the original data. The structure variable is a creative process, which requires the whole geological material to be processed, transformed, compressed and constructed to form a new statistical method which satisfies the above four properties: Comprehensive approach: a number of signs are integrated into a new variable, so that information can be enhanced and concentrated. Practice shows that the integrated variables are more inclusive and stable than a single variable.

Correlation method. By using the calculation method, the variable quantity of supply as much information as possible on the research target, and achieve the optimization variables. This method can make the important variables can be preserved, removing unimportant variables. Correlation method is based on the similarity variable correlation and sample measurement.

The variables are taken from the original data , but generally not equivalent to the original data . Structure variable is a creative process requirements analysis that needs all the geological material processing , transform , compression and structure to the formation of new that has meet the above four kinds of statistics .

In general , the typical method of variable structure are as follows :

Regression and step-wise regression method : The method uses m dimension from n sample to process the linear regression calculation and obtain the regression coefficients by solving the normal equation group of the least squares method to weight the measure of geological variables and thus participate in approximate calculation model to measure function ; On the other hand , through the gradual introduction of interactive operation , selecting the optimal equation under the single variable regression weight conditions . This method is common through linear regression unbiased estimation assumes that the construction of a set of variables .

Multiple regression mathematical model , its general form :

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_m x_m + \varepsilon$$

For the order of $n \cdot (m+1)$ observation data matrix

$$X=\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1m} & y_1 \\ x_{21} & x_{22} & \cdots & x_{2m} & y_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} & y_n \end{pmatrix}$$

Where $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$,are independent and $\varepsilon_i \sim N(0,\sigma^2), i=1,\cdots n$, here $Y=X\beta+\varepsilon$, where $\beta=(\beta_1,\beta_2,\cdots\beta_m)'$, $Y=(y_1,y_2,\cdots,y_n)'$, we can get the estimate of using the least square method , therefor , let $\frac{\partial Q}{\partial \beta}=0$.

$$Q=\sum_{\alpha=1}^{n}(y_\alpha-\hat{y}_\alpha)^2$$

Where , is the sum of square error . Here $\hat{y}_\alpha=\beta_0+\beta_1 x_{\alpha 1}+\beta_2 x_{\alpha 2}+\cdots+\beta_p x_{\alpha p}$, then $\hat{\beta}=(X'X)^{-1}X'Y$.

Step-wise discriminant analysis method : this method is calculating the discriminant function of obtaining variable measure to be as the variable selection criteria under the assumption of known samples belonging to .

The specific is : calculate the increment of the residual sum of square Q or reduction of the regression sum of square U , that is $V_j=Q'-Q$ or $V_j=U-U'$, where V_j is the changes when the j -th variable X_j is introduced in the group of given the number of variables . It is called Partial regression sum of square . The value of V_j indicates the importance of X_j so as to the standard of measure whether to introduce or eliminate the variable .

Factor analysis and canonical correlation method : This method is a simplified conditions in multidimensional variable space , variables and sample scores obtained in low dimensional factor space , function and structure characterization of the relationship between variables ; In two groups of variables correlation measurement process , through the construction of typical variables related to reproduce the model variables , measure the degree of information transformation and structural optimization . The specific application of the formula is as follows :

Suppose the standardization of $n\cdot m$ the order of observation data matrix $X=(x_{ij})_{n\times m}=(x_1,x_2,\cdots,x_m)$, the general model of factor analysis is

$$x_j=\sum_{k=1}^{p}a_{kj}\cdot f_k+u_j\varepsilon_j, j=1,2,\cdots,m \quad (1)$$

Where , are common factors and are single factors and they are not the observed random variables . The common factors occur in the every expression of the primitive variables and we can take them as the common element of the primitive variables ; every common factor has effect on at least two primitive variables or it belongs to the single factor . Every single factor occurs in the -th corresponding expression of the primitive variable , it only has effect on the original variable . (1)can be expressed by the matrix , where

$$F=(f_1,f_2,\cdots,f_p), A=(a_1,a_2,\cdots,a_m), E=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_m), U=\begin{pmatrix} u_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & u_m \end{pmatrix}$$

,

Then we can obtain A by the first p characteristic value and the corresponding characteristic vector of correlation matrix $R=\frac{1}{n}X'X$; And $F=\lambda_p{}^{-1}X(a_1,a_2,\cdots,a_p)$,

$$\lambda_p^{-1}=\begin{pmatrix} \frac{1}{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\lambda_p} \end{pmatrix}.$$

Minimum dimension analysis method : this method is to calculate the dimension of the original sample matching coefficient and minimum dimension scale under the condition of the minimum distance difference based on reconstructed sample space , the implementation of the five functions of metric variables .

Suppose the qualitative variable matrix : $X=(x_{ij})_{n\times m}$.

First of all according to a defined between the two two variables of non proximity measure matrix :

$$S=(s_{ij})_{n\times m} \qquad (s_{ij}\geq 0, s_{ij}=s_{ji}, i\neq j) .$$

Then try to find a few samples in a certain minimum Euclidean space in a scale :

$$(x_{i1},x_{i2},\cdots,x_{ir})' \qquad i=1,2,\cdots,n, \qquad r<m$$

The squared Euclidean distance by this scale decision :

V. CONCLUSION

We have discussed above , function research of geological variables is an important part to achieve high quality of geological modeling , spatial prediction accuracy and stability and the prediction accuracy of the model is not the complexity of the model itself , but lies in the amount of fine degree of geological variables , the following conclusions are obtained :the essence of comprehensive calculation five functions is to simplify the process of structure of multidimensional space to low dimensional space . The simulation of the maximum correlation calculation and its essence is the main problem in linear algebra . The stability and accuracy of geological spatial model depends on the predicted concentration information level of geological variable and the advanced degree of expression and function , its essence is the choice of space conversion tool , so the orthogonality of the space expression and generalized correlation of spatial geological space are the five function research measurement content that must be solve .

REFERENCE

[1]LU Laijun . Several issues on the study of geological variables in the function[J].Journal of Changchun College of geology in Dr.1992,223-229.

[2]WANG Shicheng , CHEN Yongliang , XIA Lixian . The theory and method of comprehensive information mineral resources prediction[M].Science Press . 2000.

[3]ZHANG Jiatong , LU Laijun . Mixed type of geological variables close relationship model[J].Journal of Jilin University: Earth Science Edition2012,42(Suppl. 1): 224-227.

[4]Davidson D.A., Theocharopoulos S.P., Bloksma R.J.. A land evaluation project inGreece using GIS and based on Boolean and fuzzy set methodologies. InternationalJournal of Geographical Information System,1994,8(4):369-384.

[5]CHEN Shupeng , LU Xuejun , ZHOU Chenghu . Introduction to geographic information system[M].Beijing: Science press,1999.

[6]Thoen Bill. Internet resources for the geosciences, with an emphasis on GIS andmapping.ComPuters&Geosciences,1995,21(6):779-786.

[7]QIA Jinhai . Comprehensive information mineral resources prediction and information integration system[M].Changchun: Jilin University,2007.

[8]CAI Hongchun , ZHANG Chunming JIANG Shaofei et.al.Mineral resources information system based on GIS Technology[J].Geology and resources,2003,12(2):111-114.

[9] Zhoujixiang. Regression analysis [M]. Shanghai: east China normal university press, 1993.

[10] Hu Peng,Geographic information system course [M]. Wuhan: wuhan university press, 2003