

직접표본선택에 기초한 지질대상의 조건모의

김연호, 최광우

위대한 령도자 김정일동지께서는 다음과 같이 교시하시였다.

《과학자, 기술자들은 현실에 튼튼히 발을 붙이고 사회주의건설의 실천이 제기하는 문제들을 연구대상으로 삼고 과학연구사업을 진행하여야 하며 연구성과를 생산에 도입하는 데서 나서는 과학기술적문제들을 책임적으로 풀어야 합니다.》(《김정일선집》 증보판 제15권 492페이지)

현재 지질통계학에서는 복잡한 지질대상들을 모형화할수 있는 다점지질통계학에 대한 연구가 기본으로 되고있다.[1-3] 지금까지 다점지질통계학에 기초한 조건모의에서는 SNESIM알고리즘을 리용한 방법이 널리 리용되였다. 이 방법은 불연속자료만을 모의할수 있는 제한성을 가지고있다. 그러므로 우리는 연속자료도 모의할수 있는 직접표본선택에 기초한 지질대상의 조건모의방법을 연구하였다.

1. 지질조건모의를 위한 직접표본선택알고리즘

직접표본선택(Direct Sampling)에 기초한 지질조건모의에서는 우연함수 $Z(x)$ 를 모의하는데 이때 입력자료는 모의살창자료와 학습화상, 조건자료(추공자료와 같은 각종 탐사자료)이다. 모의살창자료의 마디점은 x , 학습화상의 마디점은 y , N 개의 조건자료를 $Z(x_i), i=1, N$ 로 각각 표시한다.

직접표본선택알고리즘은 다음과 같다.

① 매 조건자료를 모의살창자료에서 가장 가까운 살창마디점에 할당한다. 만일 몇개의 조건자료가 같은 살창마디점에 할당된다면 가장 가까운 점의 조건자료를 살창마디점에 할당한다.

② 모의살창자료에서 마디점들을 탐색하는 경로를 설정한다. 경로는 차례로 모의하려는 살창마디점들의 순서에 대한 벡토르자료로 표현된다.

③ 경로의 매 마디점 x 에 대하여 근방을 탐색한다. 근방은 모의살창점들가운데서 이미 조건자료가 할당되었거나 모의가 진행된 n 개의 가장 가까운 살창마디점 $\{x_1, x_2, \dots, x_n\}$ 으로 구성된다.

④ x 의 근방 $N(x, L) = \{x + h_1, \dots, x + h_n\}$ 을 정의할수 있는 다음의 벡토르를 계산한다.

$$L = \{h_1, \dots, h_n\} = \{x_1 - x, \dots, x_n - x\}$$

⑤ 자료사건 $d_n(x, L) = \{Z(x + h_1), \dots, Z(x + h_n)\}$ 을 구성한다. 자료사건은 근방의 모든 마디점들에서 우연함수값으로 구성되는 벡토르이다.

⑥ 학습화상에서 탐색창문을 설정하고 자료사건 $d_n(y, L)$ 을 검색한다.

⑦ 모의살창의 자료사건과 학습화상의 자료사건사이의 거리 $d\{d_n(x, L), d_n(y, L)\}$ 을 계산한다.

만약 계산값이 이 자료사건에 대하여 계산한 값들중에서 최소값이라면 $y, Z(y), d\{d_n(x, L), d_n(y, L)\}$ 를 보관한다.

⑧ $d\{d_n(x, L), d_n(y, L)\}$ 가 턱값 t 보다 작다면 $Z(y)$ 를 $Z(x)$ 에 할당한다.

⑨ ⑥-⑧순환의 반복회수가 설정한 한계값을 초과하는 경우 계산값이 가장 작은 마디점 y 를 선택하고 $Z(y)$ 를 $Z(x)$ 에 할당한다.

직접표본선택방법의 원리는 그림 1과 같다.

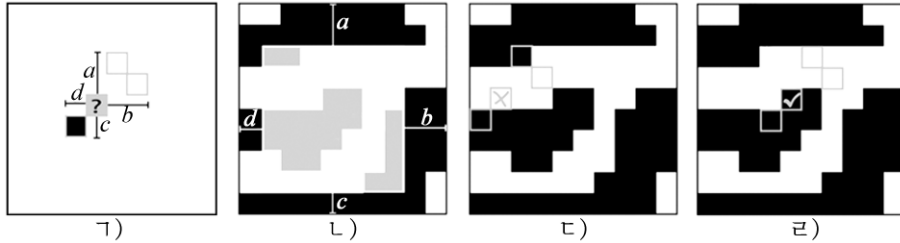


그림 1. 직접표본선택방법의 원리

가) 자료사건, 나) 탐색창문, 다) 탐색창문에서의 검색과정, 라) 검색된 자료사건; a, b, c, d-자료사건의 크기

두 자료사건들사이의 거리 $d\{d_n(x, L), d_n(y, L)\}$ 는 연속자료와 불연속자료의 모의에 모두 적용할수 있다.

불연속자료인 경우 거리를 다음과 같이 정의한다.

$$d\{d_n(x, L), d_n(y, L)\} = \frac{1}{n} \sum_{i=1}^n a_i \quad (1)$$

여기서 $a_i = \begin{cases} 0, & Z(x_i) = Z(y_i) \\ 1, & Z(x_i) \neq Z(y_i) \end{cases}$ 이다.

연속자료인 경우 다음과 같은 무계화된 유클리드거리를 이용한다.

$$d\{d_n(x), d_n(y)\} = \sqrt{\sum_{i=1}^n \alpha_i [Z(x_i) - Z(y_i)]^2} \quad (2)$$

여기서 $\alpha_i = \frac{\|h_i\|^{-\delta}}{d_{\max}^2 \sum_{j=1}^n \|h_j\|^{-\delta}}$, $d_{\max} = \max_{y \in \text{TI}} Z(y) - \min_{y \in \text{TI}} Z(y)$, TI는 학습화상을 의미한다.

실천에서 특히 연속자료인 경우 $d_n(x)$ 와 완전히 일치하는 자료사건 $d_n(y)$ 를 학습화상에서 찾을수 없다. 그러므로 모의계산에서는 턱값 t 를 도입하여 $d\{d_n(x), d_n(y)\}$ 가 t 보다 작으면 자료사건 $d_n(y)$ 를 보관한다.

2. 직접표본선택에 기초한 조건모의결과

우리는 직접표본선택방법을 연속자료와 불연속자료로 표시되는 지질대상의 조건모의에 적용하였다.

먼저 맥모양의 반암형동광상자료[2]를 모의하였다.(그림 2)

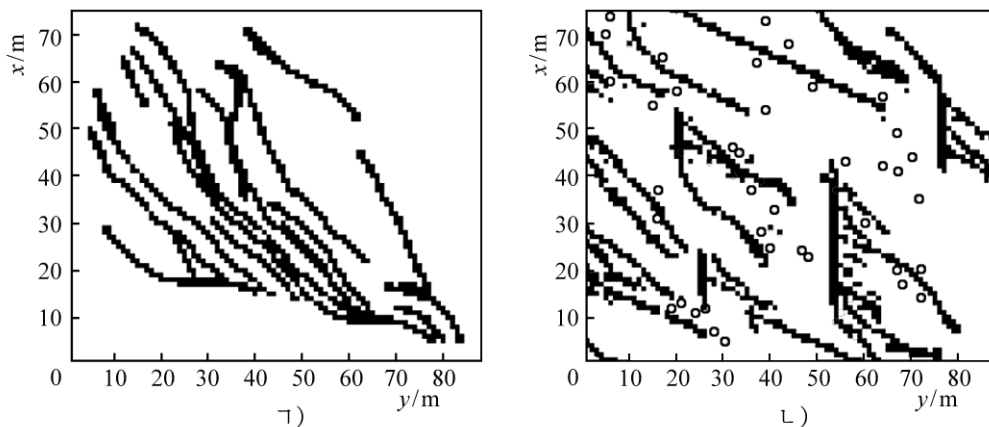


그림 2. 불연속자료로 표시되는 지질대상의 모의

ㄱ) 학습화상, ㄴ) 모의결과; 검은색 점들은 광체,
흰색바탕은 배태암, 동그라미점들은 추공자료

모의에서는 자료사건들사이의 거리를 식 (1)을 리용하여 계산하였다. 그리고 턱값을 0~0.5에서 0.05간격으로 변화시키면서 모의특성을 평가하였다. 결과 턱값을 0.2이하로 설정하는 경우에 가장 좋은 모의결과가 얻어졌다.

다음으로 연속자료로 주어지는 지질대상을 모의하였다.(그림 3)

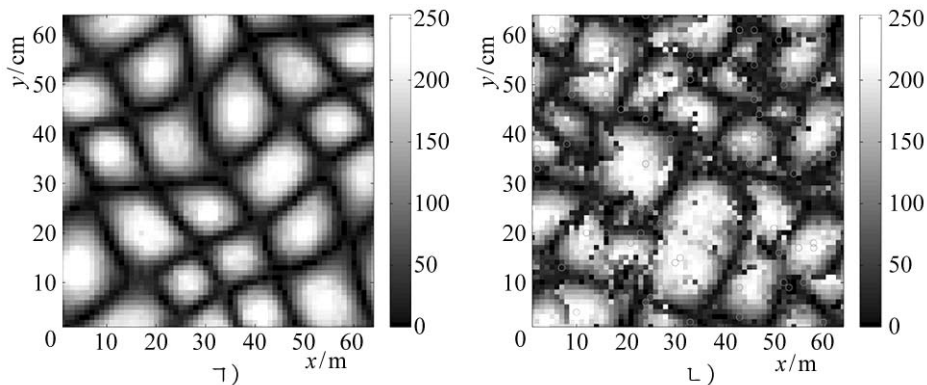


그림 3. 연속자료로 표시되는 지질대상의 모의

ㄱ) 학습화상, ㄴ) 모의결과

그림 3에서 보는바와 같이 모의대상은 암석식별을 위한 사진자료로서 0~250사이의 밝음도값을 가진다. 이와 같은 자료는 조건모의특성을 평가하는데 널리 리용된다.

연속자료의 조건모의에서는 자료사건들사이의 거리로서 무게화된 유클리드거리와 정규화된 거리[3]를 리용할수 있다. 우리는 무게화된 유클리드거리를 리용하여 모의를 진행하였다. 이때 턱값을 0~0.5사이에서 0.05간격으로 변화시키면서 모의특성을 평가하였다. 턱값을 0.1이하로 설정하는 경우 가장 좋은 모의결과를 얻었다.(그림 3의 ㄴ)) 이로부터 직접 표본선택에 기초한 방법은 연속자료로 표시되는 복잡한 지질대상을 효과적으로 모의할수 있다는것을 알수 있다.

맺 는 말

1) 지질대상의 조건모의를 위한 직접표본선택방법은 불연속자료뿐만아니라 연속자료도 모의할수 있으며 이때 자료사건들사이의 거리를 모의하려는 지질대상의 특성을 잘 고려하여 선택하는것이 중요하다.

2) 직접표본선택에 기초한 조건모의에서는 턱값을 불연속자료인 경우 0.2이하로, 연속자료인 경우 0.1이하로 설정하는것이 가장 합리적이다.

참 고 문 헌

- [1] 김일성종합대학학보(자연과학), 58, 7, 165, 주체101(2012).
- [2] P. Renard et al.; Mathematical Geosciences, 43, 8, 879, 2011.
- [3] G. Mariethoz et al.; Water Resources Research, 48, 307, 2012.

주체104(2015)년 2월 5일 원고접수

Conditional Simulation of Geological Objects based on Direct Sampling

Kim Yon Ho, Choe Kwang U

In direct sampling method for conditional simulation of geological objects, it is important to choose distance between data events by properly considering the characteristics of geological objects to be simulated.

In the conditional simulation based on direct sampling, it is most reasonable to set threshold smaller than 0.2 for discrete data and threshold smaller than 0.1 for continuous data.

Key words: direct sampling, conditional simulation, geostatistics