

Simulated Annealing 기법을 이용한 실험적 베리오그램의 모델링

정대인(jdi@geofluid.snu.ac.kr), 최종근(johnchoe@snu.ac.kr), 기세일(ksi@geofluid.snu.ac.kr)

서울대학교 지구환경시스템공학부

요약문

실험적 베리오그램의 모델링에 SA(Simulated Annealing)기법을 이용하였다. 최소 자승법의 해를 구하기 위하여 기존의 상용 프로그램에서 많이 이용되고 있는 반복법에 근거한 방법에 비해서 SA 기법은 초기 가정값에 크게 영향을 받지 않고 일정한 모델 인자의 값을 제시하였다. 임의의 초기 가정값을 입력하여도 충분한 반복 계산을 통하여 목적함수의 값이 광역적 최소값으로 수렴하는 것을 확인할 수 있었다. 베리오그램 모델이 일반적으로 비선형 모델이기 때문에 목적함수의 지역적 최소값으로의 수렴이 문제가 되고 이로 인하여 구해지는 인자의 값이 정확하지 않을 수 있지만 SA 기법을 이용하여 최소 자승법의 해를 구하게 되면 정확한 인자의 값을 구할 수 있음을 확인하였다.

key word: 베리오그램, 모델링, SA(Simulated Annealing), 최소 자승법

1. 서론

우리는 여러 형태의 지구물리학적 자료를 효율적, 통합적으로 이용하기 위하여 공간자료의 분석을 하게 된다. 대표적인 분석 방법으로 크리깅, 다각형 방법, 회귀 분석, 역거리 추정법 등이 있다. 이 중 크리깅의 방법이 가장 널리 이용되는데 크리깅을 하기 위해서는 실제 데이터로부터 구한 실험적 베리오그램을 이론적 베리오그램으로 모델링 해야 한다. 실험적 베리오그램의 모델링은 수동적인 시각적 분석이나 몇 가지 인자 값들의 단순 반복 평가를 통해서 이루어진다. 그러나 수동적인 시각적 분석이나 단순 반복 평가법에 의해서 베리오그램 모델을 결정하고 모델의 인자 값들을 추정하는 것은 분석자의 주관이나 초기 가정값에 의해서 크게 영향을 받게 된다.

분석적인 방법으로 베리오그램의 모델링을 할 때 가장 많이 쓰이는 방법이 최소 자승법이다. 최소 자승법은 실제 데이터와 모델로부터 구한 데이터의 차이의 제곱이 최소가 되도록 모델의 인자를 결정하는 것이다. 일반적으로 베리오그램 모델은 비선형 모델이기 때문에 목적함수의 형태도 비선형이다. 이런 경우에 기울기 기반의 최적화 기법이나 단순 반복법은 목적 함수를 지역적 최소값에 수렴하게 하는 문제점을 가지게 된다. 이 문제의 해결을 위한 여러 방법들이 제시되었고 가장 대표적인 방법 중의 하나가 SA(Simulated Annealing) 기법이다. SA 기법은 주로 광역 최적화 기법에 사용되며 원하는 목적함수의 값을 최소화하기

위하여 관심 있는 변수의 일부를 임의로 변화시켜 가면서 목적함수가 최소값에 도달하도록 하는 방법이다. 이 방법이 다른 최적화 기법과 다른 점은 목적함수의 값이 이전 단계에 비해서 증가하더라도 그러한 변화를 수용할 확률을 가지게 함으로써 지역적 최소값이 아닌 광역 최소값에 수렴하게 한다는 것이다. 이 연구에서는 실험적 베리오그램의 모델링 방법으로 최소 자승법을 이용하였고 최소 자승법의 해를 구하는데 SA 기법을 이용하였다.

2. 방법

1) 최소 자승법

최소 자승법은 식 1과 같이 실제 데이터로부터 구한 베리오그램과 이론적 베리오그램 모델로부터 구한 데이터의 차이의 제곱이 최소가 되도록 모델의 인자들을 결정하는 방법이다.

$$\min E = \sum_{i=1}^N (\gamma(h_i) - \bar{\gamma}(h_i))^2 \quad (1)$$

여기서, h_i : 분리거리, $\gamma(h_i)$: 실험적 베리오그램, $\bar{\gamma}(h_i)$: 모델로부터 구한 베리오그램

본 연구에서 이용한 베리오그램 모델은 구형 모델, 지수 모델, 파워 모델, 가우스 모델, 큐빅 모델 등이며 그 각각의 식은 다음과 같다.

· 구형 모델(Spherical Model)

$$\gamma(h) = \begin{cases} C_0 + C_1 \left[1.5 \times \left(\frac{h}{a} \right) - 0.5 \times \left(\frac{h}{a} \right)^3 \right] & 0 \leq h < a \\ C_0 + C_1 & a \leq h \end{cases} \quad (2)$$

· 지수 모델(Exponential Model)

$$\gamma(h) = C_0 + C_1 [1 - \exp(-\frac{3h}{a})] \quad 0 \leq h \quad (3)$$

· 파워 모델(Power Model)

$$\gamma(h) = C_0 + ah^\beta \quad 0 < \beta < 2, \quad 0 \leq h \quad (4)$$

· 가우스 모델(Gaussian Model)

$$\gamma(h) = C_0 + C_1 [1 - \exp(-3(\frac{h}{a})^2)] \quad 0 \leq h \quad (5)$$

· 큐빅 모델(Cubic Model)

$$\gamma(h) = \begin{cases} C_0 + C_1 [7 \times (\frac{h}{a})^2 - 8.75 \times (\frac{h}{a})^3 + 3.5 \times (\frac{h}{a})^5 - 0.75 \times (\frac{h}{a})^7] & 0 \leq h < a \\ C_0 + C_1 & a \leq h \end{cases} \quad (6)$$

2) SA(Simulated Annealing) 기법

SA 기법은 지역적 최소값으로의 수렴 문제를 해결하기 위하여 목적함수의 값이 지역적 최소값을 벗어날 수 있는 확률을 가지게 함으로써 광역적 최소값으로 수렴하게 하는 방법으로 그 구체적인 과정은 다음과 같다.

- ① 모델의 인자값을 가정하고 초기 자료를 이용하여 목적함수의 초기값을 계산한다.
- ② 목적함수의 초기값이 허용 오차를 만족하지 않는다면 모델의 인자의 값을 변화시켜서 목적함수의 값을 다시 구한다.
- ③ 새로 구한 목적 함수의 값과 이전 단계의 목적함수의 값을 이용하여 인자들의 변화를 받아들일 것인지를 결정한다.
- ④ 인자들의 변화를 허용한다면 초기 인자들의 값을 변화된 값으로 업데이트한다.
- ⑤ 목적함수의 값이 stopping criteria에 도달할 때까지 위의 과정을 반복 수행한다.

위의 과정 중 인자들의 변화를 허용할 것인지를 결정하는 알고리즘으로 메트로폴리스(Metropolis) 알고리즘이 많이 이용된다. 메트로폴리스 알고리즘의 구체적인 과정은 다음과 같다.

- ① 초기 모델의 인자 m_i 를 가정하여 초기 목적함수 값 E_i 를 구한다.
- ② 이 모델 인자에 변화를 준 후 m_j , 목적함수의 값 E_j 를 구한다.

$$m_j = m_i + \Delta m_i, \quad \Delta E_{ij} = E_j - E_i$$

- ③ $\Delta E_{ij} \leq 0$ 이면, 변화를 채택한다.

$\Delta E_{ij} \geq 0$ 이면, metropolis criterion을 이용해서 채택 여부를 결정한다.

· Metropolis criterion

$P = \exp(-\frac{\Delta E_{ij}}{T})$ 를 계산하고 0과 1사이에서 난수를 발생하여 이 값을 r이라고 한다.

만약 $r > P$ 이면 새 값을 기각하고 $r < P$ 이면 새 값을 채택한다.

SA 기법은 초기 온도, 일정한 온도에서의 시행 횟수, 온도의 감소 비율 등을 적절히 정해 준다면 목적함수를 항상 광역적 최소값에 수렴하게 한다는 장점이 있다. 일반적인 모델의 인자를 추정해 주는 상용 프로그램의 경우 대체로 기울기에 근거한 방법들을 이용하기 때문에 인자들의 초기 값을 적절히 가정하지 않을 경우에는 모델 인자들을 정확하게 추정하지 못하지만 SA 기법을 적용한 경우에는 인자들의 초기 가정 값에 관계없이 정확한 인자 추정이 이루어진다. SA 기법은 하나의 온도에서 몇 번의 시뮬레이션을 수행하고 온도를 낮추어 준 다음 다시 일정한 시뮬레이션을 수행하기 때문에 시뮬레이션 시간이 상당히 많이 걸리는 단점이 있다. 이러한 문제점을 해결하기 위해 Fielding(2000)은 최적의 고정된 온도에서 SA를 수행하는 방법을 제시하기도 하였다.

3. 결과 및 고찰

그림 1-5는 SA를 이용하여 베리오그램 모델링을 하는 프로그램 결과이다. 주어진 자료는 임의의 카드뮴의 오염 농도 자료이다. 개발된 프로그램을 통하여 그림 1-5와 같이 각각

의 모델에 대하여 최소의 오차를 가진 모델을 제시할 수 있으며, 그림 1에 제시된 구형모델의 경우가 최소의 오차를 나타내었다.

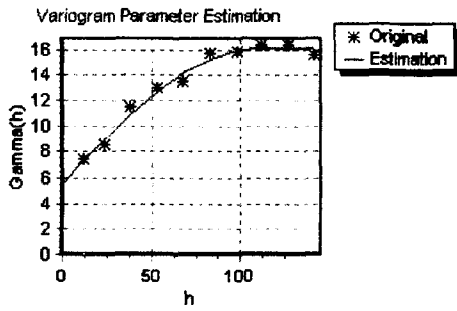


그림 2 - 구형 모델을 이용한 모델링

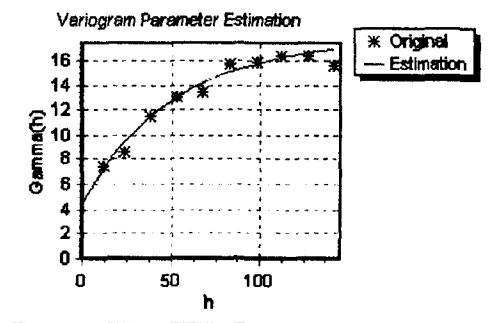


그림 1 - 지수 모델을 이용한 모델링

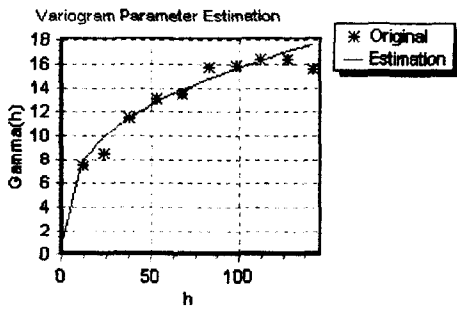


그림 3 - 파워 모델을 이용한 모델링

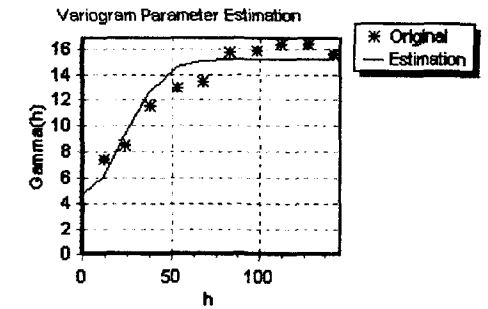


그림 4 - 가우스 모델을 이용한 모델링

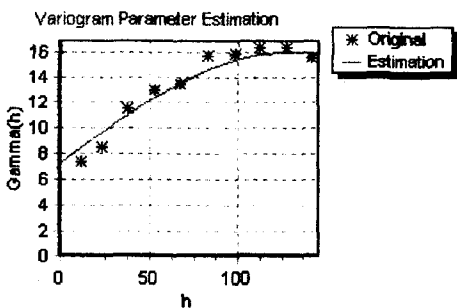


그림 5 - 큐빅 모델을 이용한 모델링

	구간	문턱값	너깃
초기 가정	30	0	0
추정 결과	108	10.4	5.68
초기 가정	30	10	10
추정 결과	108	10.1	5.54
초기 가정	100	10	10
추정 결과	108	10.2	5.74

표 1 초기 가정 인자값에 따른 모델 인자값의 변화

표 1은 초기 가정값에 따라서 얻어지는 구형 모델의 문턱값, 너깃, 구간의 변화를 보여준다. 표의 데이터로부터 SA 기법으로 베리오그램 모델의 인자를 결정하게 되면 초기 가정값에 관계없이 거의 일정한 문턱값, 너깃, 구간을 얻을 수 있음을 알 수 있다. 이는 SA 기법을 이용하게 되면 기존의 베리오그램 모델링 프로그램이 가지는 단점인 초기 가정값에 대해

서 추정된 모델의 인자값이 민감하게 변화하는 현상을 제거될 수 있음을 보여준다. 적절한 모델을 선택해서 SA 기법으로 베리오그램 모델링을 하면 초기 가정값을 임의로 주어도 SA 기법은 반복 계산을 충분히 하게 되면 목적함수의 값이 광역적 최소값에 수렴하게 된다. 따라서 얻어지는 모델의 인자값도 거의 일정한 값을 가지게 되며 그 인자값을 이용하여 얻은 이론적 베리오그램 값은 크리깅 기법을 이용한 미지값의 예측에 효율적으로 사용될 수 있다.

4 결론

1. SA 기법을 실험적 베리오그램의 모델링에 적용하여도 정확한 모델 인자의 추정이 가능하였다.
2. SA 기법을 이용하게 되면 모델 인자의 초기 가정값에 관계없이 거의 일정한 모델 인자 값을 구할 수 있었다.
3. 5가지 모델에 대해서 실험적 베리오그램의 모델링이 가능한 사용자 중심의 윈도우용 프로그램을 개발하였다.

5. 사사

이 연구는 한국과학재단의 특정기초연구비(R01-2000-00058)에 의해 지원되었습니다.

6. 참고문헌

1. Jian, X., Olea, R.A., Yu, Y. "Semivariogram modeling by weighted least squares", *Computer & Science.*, 22, pp.387-397(1996).
2. Fielding, M. "Simulated annealing with an optimal fixed temperature", *SIAM J. OPTIM.*, 11, pp.289-307(2000).
3. 최중근, 공간정보 모델링: 크리깅과 최적화, 구미서관, 서울, 2002.