ZESZYTY NAUKOWE POLITECHNIKI ŚLĄSKIEJ 2004
Seria: BUDOWNICTWO z. 102 Nr kol. 1644

Joanna KOZŁOWSKA*
Akademia Rolnicza we Wrocławiu

METODA KRIGINGOWA APROKSYMACJI WYNIKÓW POMIARÓW ŚRODOWISKOWYCH

Streszczenie. Temat referatu obejmuje problemy aproksymacji wyników pomiarów środowiskowych. Opisano jedną z metod aproksymacji przestrzennego (jednak dwuwymiarowego) rozkładu cech pewnego zjawiska. W szczególności przedstawiono istotę metody krigingowej, podano algorytm obliczeniowy, opracowany przez autorkę i sformułowano wnioski o znaczeniu praktycznym.

THE KRIGING METHOD APPROXIMATION OF RESULTS OF ENVIRONMENTAL MEASUREMENT

Summary. The paper comprises approximations problems of results of environmental measurements, at the use of geostatistics methods. These methods are used to create topographical maps, contour lines. The literature of the subject distinguishes a considerable number of methods, credited to the geostatistics group. One of them is the kriging procedure, treated in a special manner. Kriging makes possible constructing, among other things, numerical models of the ground. The report presents the kriging method and introduces the computational algorithm.

1. Wstęp

Metody statystyki przestrzennej są stosowane w różnych dziedzinach, np. w górnictwie, topografii oraz celem monitoringu środowiska. W rezultacie tworzone są mapy topograficzne, warstwice, izolinie itp. W literaturze przedmiotu wyróżnia się znaczną liczbę metod zaliczanych do omawianej grupy. W referacie skupiono uwagę na procedurze krigingowej (zaczerpniętej z literatury) w zastosowaniu do monitoringu środowiska. Referat ma charakter studyjny (opisano istotę metody oraz algorytm obliczeniowy) i jego treść stanowi przygotowanie do badań eksperymentalnych. W procesie eksperymentów kriging

*Opiekun naukowy: Dr hab. inż. Janusz Łomotowski, prof. AR we Wrocławiu będzie weryfikowany przez porównanie z innymi metodami. Na uwagę zasługuje wkład własny autorki w przedmiotowy problem: algorytm sporządzono w formie umożliwiającej tworzenie programu komputerowego. Jest to pierwsze tego typu opracowanie dotyczące procedur krigingowych.

2. Istota metody krigingowej

Metoda krigingowa umożliwia szacowanie jakości i wielkości złóż oraz ilustrowanie rozmieszczenia wartości parametrów złożowych za pomocą map izorytm (np. warstwic). Metoda statystyki przestrzennej (krigingu) jest szczególnie wykorzystywana do oceny wskaźników zanieczyszczeń i zasobów złóż.

W ujęciu statystyki przestrzennej parametr złożowy traktowany jest jako tzw. zmienna zregionalizowana (zwana również przestrzenną lub zlokalizowaną), którą definiuje się jako ciągłą funkcję współrzędnych przestrzeni. Wartości tej zmiennej zregionalizowanej znane są tylko w obrębie próbek geologicznych lub pól pomiarowych o znikomo małych rozmiarach w porównaniu z przestrzenią złożową. Noszą one nazwę bazy geometrycznej pomiarów [1,2].

Aby względem zmiennej zregionalizowanej można było przeprowadzić interpretację statystyczną, nakłada się na nią pewne ograniczenia. W szczególności przyjmuje się hipotezę słabej stacjonarności, co oznacza, że wartość oczekiwana zmiennej nie zależy od miejsca pomiaru, a jej kowariancja jest jedynie funkcją odległości pomiędzy punktami pomiarów. W praktyce, ze względu na to, że hipoteza słabej stacjonarności w odniesieniu do realnych parametrów złożowych jest zbyt rygorystyczna, stosuje się znacznie łagodniejsze ograniczenie, które zakłada słabą stacjonarność nie tyle samej zmiennej zregionalizowanej, lecz jej przyrostów. Istotne znaczenie ma tu wariancja przyrostów, która definiuje podstawową funkcję charakterystyczną geostatystyki, zwaną wariogramem:

$$D^2[Z(x+h) - Z(x)] = E[Z(x+h) - Z(x)]^2 = 2\,\gamma\,(h) \qquad (1)$$

gdzie:

D, E - operatory wariancji i wartości oczekiwanej,

$2\gamma(h)$-wariogram,

$\gamma\,(h)$ - semiwariogram (półwariogram),

$Z(x+h)$, $Z(x)$ - wartości zmiennej zregionalizowanej w punkcie początkowym i odległym o h.

W zastosowaniach metod geostatystycznych wykorzystuje się najczęściej połowę wartości wariogramu, tzw. semiwariogram. Podstawę geostatystycznego opisu zmienności stanowi funkcja ujmująca zależność pomiędzy średnim zróżnicowaniem wartości parametrów złożowych a odległością pomiędzy miejscami ich pomiarów. W warunkach rutynowych pomiarów wartości parametrów złożowych mają postać semiwariogramu dla dyskretnej i regularnej sieci pomiarów; określa się je z klasycznej formuły G. Matherona [3,4,5]:

$$\gamma(h)=\frac{1}{2n_h}\sum_{i=1}^{n_h}(z_{h+i}-z_i)^2 \qquad (2)$$

gdzie z_i, z_{i+h} są wartościami parametru złożowego w punktach oddalonych o wektor h, natomiast n_h jest liczbą par punktów pomiarowych odległych o wektor h. Semiwariogram, którego postać została określona na podstawie wyników pomiarów w złożu, nosi nazwę semiwariogramu empirycznego i przedstawia on w syntetycznej formie strukturę zróżnicowania parametrów złożowych.

Wyznaczanie semiwariogramu w istocie polega na obliczaniu średniego kwadratu różnic dla wszystkich par utworzonych z wartości parametru złożowego, określonych w punkcie wyróżnionym (bazowym) i w każdym z punktów, który znalazł się w obszarze grupowania danych, a następnie na przypisaniu średniego kwadratu różnic średniej odległości pomiędzy punktem bazowym a punktami z rozpatrywanego sektora zliczania. Następnie całą procedurę powtarza się dla kolejnego przedziału odległości i dalej dla kolejnych punktów pomiarowych, przejmujących rolę punktów bazowych.

Semiwariogram empiryczny w formie wykresu punktowego nie może być wykorzystany jednak do rozwiązywania zadań oceny parametrów geologicznych. By można to było wykonać, przybliża się go różnymi funkcjami analitycznymi, które w dalszym postępowaniu traktowane są jako geostatystyczne modele zmienności. Do najczęściej używanych zalicza się model sferyczny Matherona opisany równaniem [6,7,8]:

$$\gamma(h)=C\left[\frac{3}{2}\frac{h}{a}-\frac{h^3}{2a^3}\right]+C_0 \quad dla \quad h\le a \qquad (3)$$

oraz

$$\gamma(h)=C_0+C=\delta^2 \quad dla \quad h\ge a \qquad (4)$$

Z definicji semiwariogramu powinno wynikać, że jego wartość dla $h = O$ wynosi również zero. Jednakże dla malejących do zera wartości h realne semiwariogramy empiryczne rzadko dążą do zera, na ogół dążą do pewnej wartości określanej w literaturze geostatystycznej jako stałej *efektu samorodków* C_0 *(nugget effect)*. Charakteryzuje ona zmienność lokalną badanego parametru złożowego i odpowiada składnikowi losowemu zmienności dla $h \to 0$.

Do aproksymacji semiwariogramów empirycznych modelami teoretycznymi stosowane są najczęściej metody automatycznego dopasowania metodą *najmniejszych kwadratów*. Jest ona szczególnie przydatna przy dużej skali obliczeń. Wykonana poprawnie procedura obliczeniowa wymaga zastosowania aproksymacji ważonej ze względu na różną dokładność oceny poszczególnych wartości semiwariogramu empirycznego, co wynika z różnej liczebności par danych, na podstawie których są wyznaczane. Dlatego zazwyczaj pomija się wartości semiwariogramu określone dla zbyt ubogiej liczby danych (np. kilku par danych).

Metoda krigingu jest geostatystyczną metodą szacowania średnich wartości parametrów złożowych i ich wartości w punktach złoża, opartą na zmienności wyrażonej semiwariogramem. Estymator wartości średniej parametru ma postać średniej ważonej i określony jest wzorem [1,2,9]:

$$z_0 = \sum_{i=1}^{n} \lambda_i \cdot z_i \tag{5}$$

gdzie λ_i jest współczynnikiem wagowym krigingu, a z_i jest wartością parametru złożowego w i-tym punkcie pomiarowym.

Specyfika procedury krigingu polega na ustalaniu wartości współczynników wagowych. Powinny być one tak dobrane, by spełniły dwa postulaty: nieobciążoności i maksymalnej efektywności.

Pierwszy postulat jest spełniony, gdy

$$E(z_i - m) = 0 \tag{6}$$

skąd wynika wymóg, aby

$$\sum_{i=1}^{n} \lambda_i = 1 \tag{7}$$

Drugi postulat oznacza, iż wariancja różnicy pomiędzy rzeczywistą, średnią wartością parametru m i jej oceną z_i winna być możliwie najmniejsza, co można zapisać następująco:

$$\sigma^2 = E[(z_i - m)^2] = \min \tag{8}$$

Warunkiem koniecznym do spełnienia postulatu minimalizacji błędu jest zerowanie się pierwszych pochodnych wariancji z uwagi na wszystkie współczynniki wagowe, tzn.:

$$\frac{\partial\left[\sigma^2 - 2\nu\sum_{i=1}^{n}\lambda_i\right]}{\partial z_i} = 0 \text{ , dla i = 1,2, 3, ... n} \tag{9}$$

gdzie ν jest mnożnikiem Lagrange'a.

W wyniku różniczkowania otrzymuje się układ n równań z n niewiadomymi współczynnikami wagowymi λ_i , który w uproszczonej postaci można zapisać następująco:

$$\gamma(x_i, x_0) = \sum_{i=1}^{n}\lambda_i\gamma(x_i, x_0) + \nu \quad \text{dla i = 1,2, 3, ... n} \tag{10}$$

lub w postaci macierzowej [2,3,4]:

$$\begin{bmatrix} \gamma(x_i, x_j) & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_i \\ \nu \end{bmatrix} = \begin{bmatrix} \gamma(x_i, x_0) \\ 1 \end{bmatrix} \tag{11}$$

gdzie $\gamma(x_i, x_j)$ jest średnią wartością semiwariogramu dla odcinków łączących punkty pomiarów (opróbowań) $z(x_i)$ oraz $z(x_j)$, x_i, x_j – odległość między punktami x_i i x_j, natomiast x_i i x_j są określonymi punktami pomiaru przy wykorzystaniu przyjętego modelu zmienności badanego parametru. Rozwiązanie przedstawionego układu równań prowadzi do ustalenia liczbowych wartości współczynników wagowych, które są niezbędne do szacowania średniej wartości parametru z minimalnym błędem. Są one również konieczne do wyznaczenia przeciętnej wielkości błędu, zwanego *błędem krigingu*, którego wariancję określa wzór:

$$\sigma^2 = \sum_{i=1}^{n}\lambda_i \cdot \gamma(x_i, x_j) + \nu - \gamma(x_i, x_0) \tag{12}$$

Metoda krigingu, ze względu na powyższe własności, stanowi dogodną metodę interpolacyjną, wysoce przydatną przy sporządzaniu map izarytm (izolinii) parametrów złożowych. Z uwagi na minimalizację błędu oceny parametru procedura krigingu zapewnia większą dokładność w porównaniu z innymi procedurami interpolacyjnymi, np. opartymi na ważeniu odwrotności odległości (lub ich kwadratu) węzła interpolacyjnego od punktów pomiaru. Zasady zastosowania tej metody polegają na nałożeniu na mapę rozmieszczenia punktów pomiarowych (otworów, próbek) na ogół kwadratowej sieci punktów (interpolacyjnych), w których będzie wykonana interpolacja wartości parametru. Przy ocenie wartości parametru złożowego w tych punktach bierze się pod uwagę wszystkie obserwacje w złożu, znajdujące się w obrębie koła o założonej z góry wielkości promienia, którego środek stanowi punkt interpolacyjny. Promień tego koła dobiera się uwzględniając zasięg semiwariogramu oraz charakter rozmieszczenia obserwacji w złożu. Wagi poszczególnych obserwacji określa się, rozwiązując wspomniany wcześniej układ równań krigingu.

Ze względu na ogromną liczbę punktów interpolacji, realizacja zadania możliwa jest wyłącznie przy zastosowaniu komputerów. Mając określone wagi obserwacji dla poszczególnych punktów interpolacji, komputer wyznacza dla każdego z nich wartości parametru (jako średnią ważoną w myśl procedury krigingu) oraz błąd krigingu.

Metoda krigingu pozwala zatem na oszacowanie (predykcję) wartości zmiennej dla określonych punktów interpolacyjnych z danych otaczających punktów pomiarowych, a z nieregularnej siatki pomiarowej uzyskuje się regularny rozkład przestrzenny zmiennej w układzie geograficznym.

3. Algorytm obliczeniowy metody krigingowej

Algorytm obejmuje 12 kroków, których kolejność w treści podano poniżej.

Krok 1. Wyznaczamy położenie, np. współrzędne geograficzne punktu x_0, którego wartość z_0 zamierzamy oszacować, np. wysokość nad poziomem morza .

Krok 2. Wyznaczamy zbiór punktów bazowych, których znamy położenie oraz wartości z.

Krok 3. Tworzymy tabelę wartości Z dla poszczególnych punktów bazowych:

	x_1	x_2	x_3
z_i	z_1	z_2	z_3

Krok 4. Tworzymy symetryczną macierz Q odległości między punktami bazowymi.

	x_1	x_2	x_3
x_1	$h_{11} = 0$	h_{12}	h_{13}
x_2	h_{21}	$h_{22} = 0$	h_{23}
x_3	h_{31}	h_{32}	$h_{33} = 0$

Krok 5. Wybieramy model wariogramu oraz wyznaczamy stałe. Przeważnie wybierany jest model sferyczny. Wartość zmiennej h nie będzie większa niż maksymalna odległość pomiędzy punktami bazowymi. Stała C i stała a powinny być mniejsze od maksymalnej odległości pomiędzy punktami bazowymi h.

$$\gamma(h) = C\left[\frac{3}{2}\frac{h}{a} - \frac{h^3}{2a^3}\right] + C_0 \quad \text{dla} \quad h \le a$$

$$\gamma(h) = C_0 + C = \delta^2 \quad \text{dla} \quad h \ge a$$

Krok 6. Macierz symetryczną Q transformujemy na Q_t, czyli wyliczamy wartość $\gamma(h_{ij})$ oraz dopisujemy wiersz i kolumnę *jedynek*.

	x_1	x_2	x_3	
x_1	$\gamma(h_{11})$	$\gamma(h_{12})$	$\gamma(h_{13})$	1
x_2	$\gamma(h_{31})$	$\gamma(h_{32})$	$\gamma(h_{33})$	1
x_3	$\gamma(h_{31})$	$\gamma(h_{32})$	$\gamma(h_{33})$	1
	1	1	1	1

Krok 7. Transformowaną macierz Q_t odwracamy i otrzymujemy $Q_t^{(-1)}$

Krok 8. Tworzymy macierz S odległości punktu x_0 od punktów bazowych.

	x_0
x_1	h_{10}
x_2	h_{20}
x_3	h_{30}

Krok 9. Transformujemy macierz S na S_t, czyli wyliczamy wartości $\gamma(h_{i0})$ oraz dopisujemy kolumnę *jedynek*.

	x_0
x_1	$\gamma(h_{10})$
x_2	$\gamma(h_{20})$
x_3	$\gamma(h_{30})$
	1

Krok 10. Obliczamy wagi krigingu λ_i oraz współczynik Lagranga z równania macierzowego:

$$\left[\frac{\gamma(x_i,x_j)\ 1}{1\quad 0}\right]\cdot\left[\frac{\lambda_i}{\nu}\right]=\left[\frac{\gamma(x_i,x_0)}{1}\right]$$

lub

$$\left[Q_t^{(-1)}\right]\cdot\left[S_t\right]=\left[\frac{\lambda_i}{\nu}\right]$$

Krok 11. Wyliczenie szukanej wartości z_0 punktu x_0

$$z_0=\sum_{i=1}^{n}\lambda_i\cdot z_i$$

wykorzystując dane z tabeli Z oraz z kroku 9.

Krok 12. Wykonanie wariogramu i sprawdzenie poprawności wybranego modelu metodą najmniejszych kwadratów. Jeżeli wyniki są niezadowalające, należy wybrać inny model lub inne stałe modelu i powtórzyć operację od początku.

4. Wnioski końcowe

Metoda krigingowa, należąca do metod statystyki przestrzennej, jest jedną najpełniej odpowiadającą założeniom stawianym przy tworzeniu map numerycznych terenu.

Kriging, ze względu na specyficzne właściwości stanowi dogodną metodę interpolacyjną, przydatną przy sporządzaniu map parametrów złożowych

Z uwagi na występującą znaczną liczbę punktów interpolacji realizacja zadania krigingowego możliwa jest wyłącznie na drodze numerycznej. Mając sprecyzowane wagi obserwacji dla poszczególnych punktów interpolacji, komputer określa dla każdego z nich wartość parametru (jako średnią ważoną) oraz błąd krigingu.

Metoda krigingowa umożliwia określenie wartości zmiennej dla przyjętych punktów interpolacyjnych na podstawie danych otaczających punkty pomiarowe.

Referat sporządzono na podstawie realizowanego projektu badawczego Nr 4 T12C 054 26, finansowanego przez MNiI.

LITERATURA

1. Cressie N. A. C.: Statistics for Spatial Data. John Wiley and Sons, Inc., New York 1991.
2. Deutsch C. V., Journel A. G.: GSLIB – Geostatistical Software Library and User's Guide. Oxford Uniwersity Press, New York 1998.
3. Furst J.: Anwendung von Geopraphischen Informationssystemen in Hydrologie und Wasserwirtschaft – Studienblatter und Materialien zur Vorlesung im SS 1998. WUM – Universitatsverlag der Hochschulerschaft an der Universitat Wien GmbH, Wien, 1998.
4. Isaaks E.H., Srivastava R.M.: Applied Geostatistics. Oxford University Press. New York 1989.
5. Kitanidis P.K.: Introduction to Geostatistics. Cambridge University Press, Melbourne, Australia 1997.
6. Maidment D. R.: Handbook of Hydrology. McGraw – Hill, Inc. USA 1993.
7. Pannatier Y.: VARIOWIN Software for Spatial Data Analysis in 2D. Spinger – Verlag, Berlin, Heidelberg 1996.
8. Wackernagel H.: Muitivariate geostatistics – An Introduction with Applications. Spinger – Verlag, Berlin, Heidelberg 1998.

Recenzent: Prof. dr hab. inż. Andrzej Majde