

BORKOWSKA Dorota

ZASTOSOWANIE RADIALNYCH FUNKCJI BAZOWYCH W INTERPOLACJI DANYCH ROZPROSZONYCH CZ. II

Streszczenie

Praca poświęcona jest zagadnieniu interpolacji rozproszonych danych przez radialne funkcje bazowe (RBF). W tej części artykułu, dla wybranej RBF, zaprezentowano metody wyboru parametru kształtu oraz przebadano jego wpływ na jakość interpolacji. Przedstawiona symulacja numeryczna potwierdza zależność dokładności interpolacji od parametru kształtu.

WSTĘP

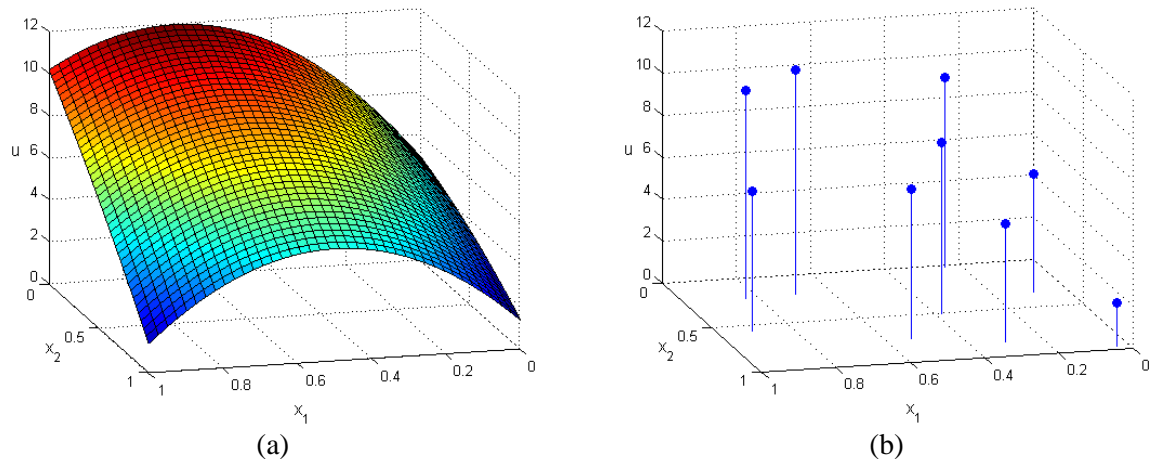
W zagadnieniu interpolacji funkcjami RBF problem sprowadza się do rozwiązania układu równań liniowych, w którym wskaźnik uwarunkowania macierzy interpolacji wzrasta, gdy wartość parametru kształtu dąży do zera. Jednocześnie najlepszą dokładność otrzymuje się dla małej wartości parametru kształtu, tzn. dla stosunkowo płaskich funkcji bazowych. W [3] wspomina się także o występowaniu zjawiska podobnego do efektu Rungego występującego w przypadku interpolacji wielomianowej. Na złagodzenie tego efektu wpływa wybór optymalnej wartości parametru kształtu c . Celem pracy jest omówienie metod wyboru wartości parametru c oraz analiza jego wpływu na dokładność interpolacji dla wybranej RBF.

1. WPŁYW PARAMETRU KSZTAŁTU NA DOKŁADNOŚĆ INTERPOLACJI RBF

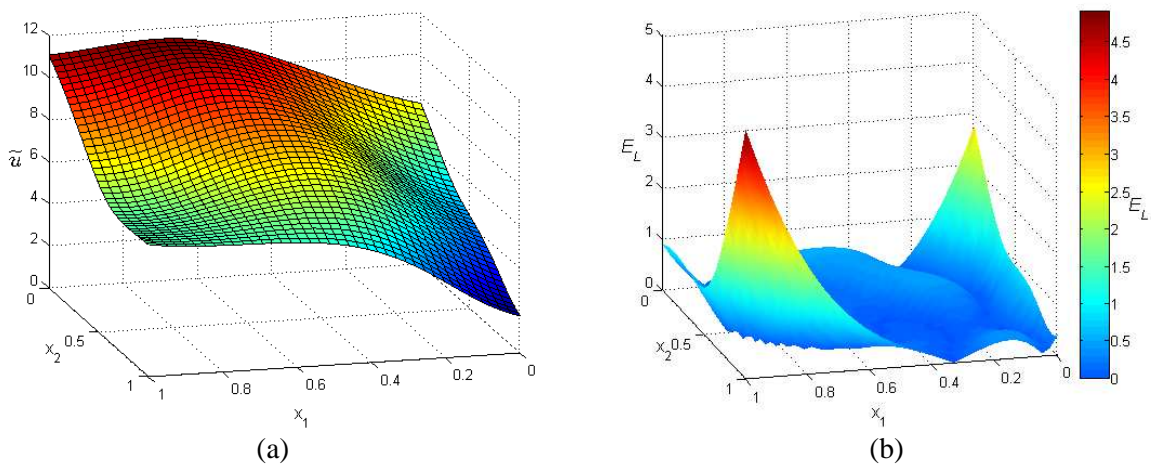
W celu przeanalizowania wpływu wartości parametru c na błąd interpolacji zostanie rozpatrzony przykład interpolacji funkcji Currina, danej wzorem [1]:

$$u(x_1, x_2) = 4.9 + 21.15x_1 - 2.17x_2 - 15.88x_1^2 - 1.38x_2^2 - 5.26x_1x_2 \quad (1)$$

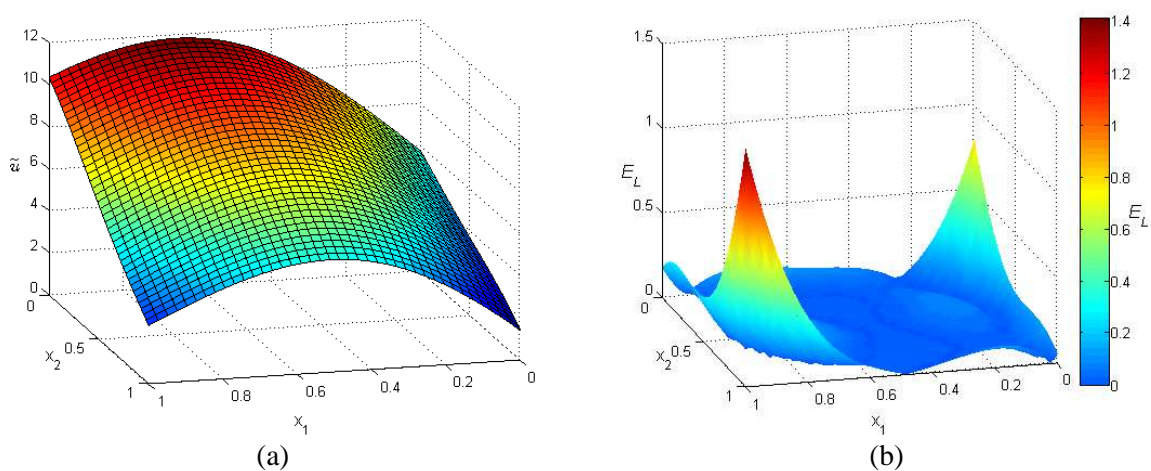
Wykres tej funkcji, w obszarze $[0,1] \times [0,1]$, przedstawia rys. 1(a). Dla 9 węzłów równomiernie rozproszonych na obszarze tak jak na rys. 1(b) interpolowano tę funkcję wielokwadratowymi RBF dla dwóch różnych wartościach parametru kształtu $c = 0.2$ oraz $c = 1$. Wykresy funkcji interpolujących oraz odpowiadające im wykresy błędów E_L przedstawiono na rys. 2 i 3. Wartości błędów E_R w obu przypadkach wynoszą odpowiednio 0.7645 dla $c = 0.2$ oraz 0.1798 dla $c = 1$.



Rys. 1. (a) Funkcja Currina w obszarze $[0,1] \times [0,1]$ oraz (b) rozmieszczenie równomiernie rozproszonych próbek użytych w interpolacji ($n_i = 9$)



Rys. 2. Interpolacja funkcji Currina dla $c = 0.2$: (a) - \tilde{u} , (b) - E_L



Rys. 3. Interpolacja funkcji Currina dla $c = 1$: (a) - \tilde{u} , (b) - E_L

Jak można zauważyć rezultatem obliczeń numerycznych dla różnych wartości parametru kształtu jest otrzymanie różnych \tilde{u} . Zatem wybór parametru c ma wpływ na dokładność interpolacji.

2. METODY WYBORU WARTOŚCI PARAMETRU KSZTAŁTU

Możliwość poprawienia dokładności interpolacji poprzez odpowiedni wybór wartości parametru kształtu jest możliwa zarówno w przypadku użycia tradycyjnej metody RBF, jak również poprzez wykorzystanie metod stabilizacji interpolacji [3]. Poszukiwanie metody znalezienia optymalnej wartości parametru c , która umożliwiałaby uzyskiwanie najdokładniejszych wyników jest aktualnym tematem podejmowanym w wielu publikacjach [3], [4], [5], [8], [10]. Ustalenie optymalnej wartości c polega na:

- metodzie prób i błędów (ang. trial and error method) – polega na wyborze najlepszej wartości c w wyniku symulacji numerycznych, powtarzanych aż do momentu uzyskania zadowalających rezultatów,
- szacowaniu błędu przy użyciu tzw. funkcji potęgowej (ang. power function) [2], która pozwala na rozdzielenie analizy błędu na analizę dwóch niezależnych składowych tj. postaci interpolowanej funkcji u oraz RBF użytych do interpolacji,
- zastosowaniu odmiany metody walidacji krzyżowej – „leave-one-out cross validation” (LOOCV) [2], [3], w której błąd interpolacji szacuje się kolejno w każdym x_i -tym węźle poprzez porównanie danego u_i z wartością $\tilde{u}(x_i)$, przy czym \tilde{u} jest liczone na podstawie pozostałych węzłów. Procedura ta prowadzi do powstania wektora błędów. Porównanie norm takich wektorów, obliczonych dla różnych c pozwala na wybór optymalnej wartości. Jej przewaga nad dwiema powyższymi metodami polega na tym, że do oszacowania błędu nie jest wymagana znajomość interpolowanej funkcji.

Innym podejściem do zagadnienia jest wybór parametru kształtu bez uwzględniania wymogu optymalności dla poszukiwanej wartości. Tego typu metody określania wartości c a priori opierają się głównie na analizie rozmieszczenia węzłów interpolacji. Spotykane w literaturze formuły dotyczą głównie wielokwadratowych i wielokwadratowych odwrotnych RBF. Są to m.in.:

- $c = \frac{1}{0.815d}$, gdzie $d = \frac{1}{i} \sum_i d_i$, przy czym d_i jest odległością między węzłem interpolacji x_i a najbliższym jego sąsiadem [7],
- $c = \frac{0.8\sqrt{n_i}}{D}$, gdzie D oznacza średnicę najmniejszego okręgu zawierającego wszystkie węzły interpolacji [6],
- $c = \left(\frac{1}{10} \max \{ \max_{i,j} |x_i - x_j|, \max_{i,j} |y_i - y_j| \} \right)^{1/2}$ [11], [13].

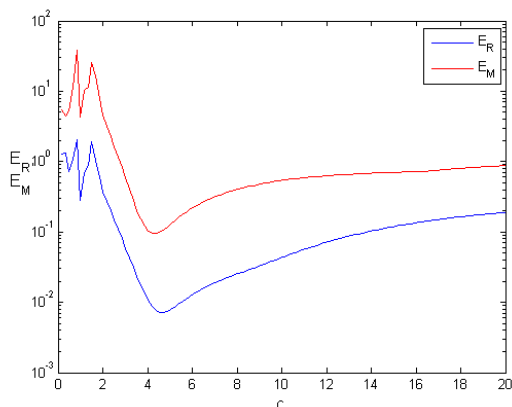
Są to metody stacjonarne (ang. stationary interpolation) [2], co oznacza, że parametr kształtu zmienia się w zależności od liczby węzłów użytych do interpolacji i co się bezpośrednio z tym wiąże – odległości między nimi (promień pokrycia – ang. covering radius). Skalowanie parametru c w takich przypadkach prowadzi do stosowania „ostrzych” RBF przy gęsto rozmieszczonych danych oraz „płaskich” RBF – dla danych rozmieszczonych rzadziej. W literaturze zamiennie z promieniem pokrycia używane jest pojęcie wypełnienie przestrzeni (fill distance) [2].

Przeciwnieństwem tego typu podejścia są metody niestacjonarne (ang. non-stationary interpolation), w których ustalone c stosuje się niezależnie od gęstości siatki węzłów interpolacji. Przykład użycia tego typu techniki został zaprezentowany w I części artykułu w rozdziale 2.1.

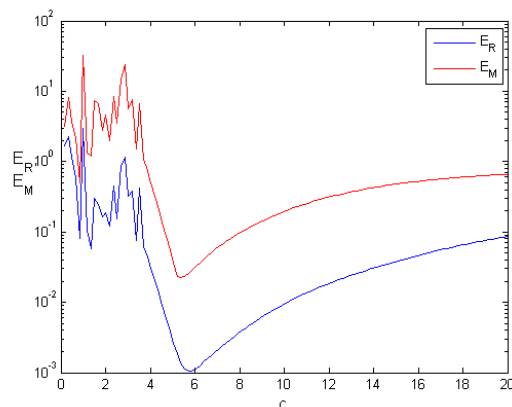
3. WYBÓR PARAMETRU KSZTAŁTU METODĄ PRÓB I BŁĘDÓW

Jedną z wymienionych metod wyboru optymalnej wartości c jest metoda *prób i błędów*. Polega ona na wykonaniu serii symulacji numerycznych dla różnych wartości parametru kształtu. Analiza wykresów błędów pozwala na wybór najlepszej.

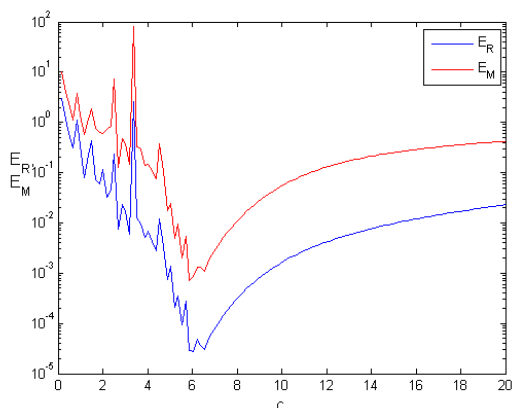
Niech będzie dany problem interpolacji funkcji Franke'go przy pomocy RBF Gaussa. Rys. 4 przedstawia wykresy błędów E_R i E_M w zależności od wartości parametru c , dla ustalonych $\{n_i\}$ elementowych zbiorów równomiernie rozproszonych danych. Na ich podstawie wybiera się quasi-optymalną wartość c .



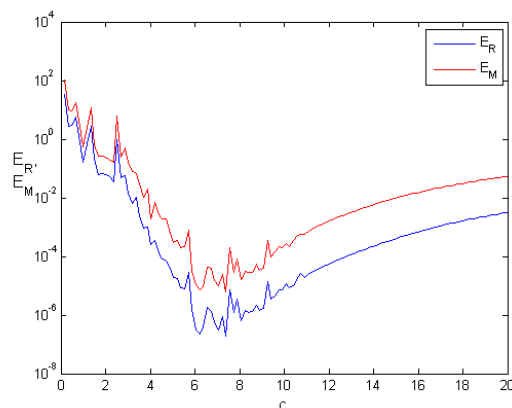
(a) - $n_i = 100$



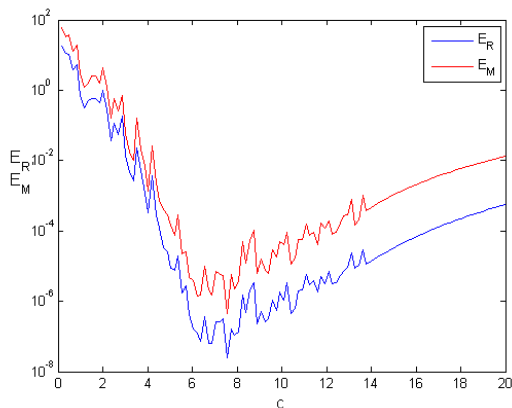
(b) - $n_i = 225$



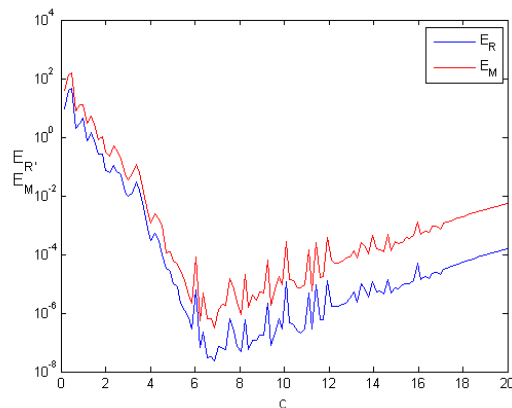
(c) - $n_i = 625$



(d) - $n_i = 1600$



(e) - $n_i = 2500$



(f) - $n_i = 3600$

Rys. 4. Wykresy zależności błędów (E_R , E_M) od wartości parametru kształtu

Przedstawione wyniki uwiadcniają problemy interpolacji RBF spowodowane dużą liczbą węzłów interpolacji. Jak widać na rysunkach, wykresy błędów nie są monotoniczne. Dla małych wartości parametru kształtu macierz interpolacji jest źle uwarunkowana, co jest zauważalne w postaci nieregularnych wykresów obydwu błędów. Po przekroczeniu pewnej wartości granicznej n_i wykresy te wygładzają się, co oznacza polepszenie uwarunkowania zadania.

Drugą ważną kwestią jest występowanie minimum błędu, które zarówno dla E_R jak i E_M występuje prawie w tym samym miejscu. Jak można zauważyć minimum błędów dla małej liczby węzłów interpolacji otrzymuje się dla „stabilnego” c , natomiast wraz ze wzrostem n_i zakres niestabilności c rośnie. Zakres ten nazwany nazywany jest w literaturze „szarą strefą” (ang. grey zone) [2].

Zagadnienie to jest dość istotne, ponieważ dla pewnych przypadków minima błędów otrzymuje się dla małych wartości parametru kształtu, leżących właśnie w „szarej strefie”. Ten konflikt między dokładnością i stabilnością rozwiązywany jest często przy pomocy tzw. reguły kompromisu (ang. trade-off principle) [2], [8], [12]. W takich przypadkach decyzję odnośnie tego czy „optymalną” wartością c jest minimum leżące w szarej strefie, czy też najmniejsza wartość leżąca już poza nią należy podjąć indywidualnie. Należy także dodać, że istnieją również techniki stabilizacji interpolacji dla małych wartości parametru kształtu (metoda Contour-Padé, RBF-QR) [5], [9], jednak podejścia te mają pewne ograniczenia związane głównie z ograniczoną liczbą węzłów interpolacji [4].

PODSUMOWANIE

Wybór parametru kształtu ma duży wpływ na dokładność oraz stabilność numeryczną interpolacji funkcjami RBF. Przeprowadzone symulacje numeryczne potwierdzają, że quasi-optymalna wartość parametru kształtu zależy od liczby węzłów. Dodatkową trudnością jest fakt, że minima błędów w wielu przypadkach mogą występować w strefie niestabilnego c (źle uwarunkowanie macierzy interpolacji). W artykule wymieniono niektóre kryteria wyboru c . Niestety do tej pory nie została jeszcze sformułowana jedna uniwersalna metoda, którą można by było zastosować dla każdej wartości n_i i każdego rodzaju RBF; badania nad tymi zagadnieniami będą kontynuowane.

Program MATLAB wykorzystany do przeprowadzenia badań został zakupiony w wyniku realizacji Projektu nr UDA-RPPK.01.03.00-18-003/10-00 „Budowa, rozbudowa i modernizacja bazy naukowo-badawczej Politechniki Rzeszowskiej” współfinansowanego ze środków Unii Europejskiej w ramach Regionalnego Programu Operacyjnego Województwa Podkarpackiego na lata 2007-2013, Priorytet I. Konkurencyjna i Innowacyjna Gospodarka, Działanie 1.3 Regionalny system innowacji.

BIBLIOGRAFIA

1. Currin, C., Mitchell, T., Morris, M., & Ylvisaker, D., *Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments*. Journal of the American Statistical Association, 1991, 86:953-963,.
2. Fasshauer G. E., *Meshfree Approximation Methods with MATLAB*. World Scientific Publishing Co., Singapore 2007.
3. Fasshauer G. E., Zhang J. G. *On Choosing “Optimal” Shape Parameters for RBF Approximation*. Numer. Algor., 2007, 45:345-368.
4. Fornberg B., Larsson E., Flyer N., *Stable Computations with Gaussian Radial Basis Functions*. SIAM Journal on Scientific Computing, 2011, 33:869-892.
5. Fornberg B., Wright G., *Stable computation of multiquadratic interpolants for all values of shape parameter*. Computers and Mathematics with applications, 2004, 48:853-867.

6. Franke R. *Scattered data interpolation: tests of some methods*. Math. Comp., 1982, 48: 181-200.
7. Hardy R. L., *Multiquadratic equations of topography and other irregular surfaces*. J. Geophys. Res., 1971, 76:1905-1915.
8. Mongillo M., *Choosing Basis Functions and Shape Parameters for Radial Basis Function Methods*. SIAM Undergraduate Research Online (SIURO), 2011.
9. Piret C., *Analytical and Numerical Advances in Radial Basis Functions*. PhD thesis, University of Colorado at Boulder 2007.
10. Rippa S. *An algorithm for selecting a good value for the parameter c in radial basis function interpolation*. Adv. Comput. Mat., 1999, 11:193-210.
11. Stead S.E., *Estimation of gradients from scattered data*. Rocky Mountain J. Math., 1984, 14:265-279.
12. Wendland, H. *Scattered Data Approximation*. Cambridge University Press, Cambridge 2005.
13. Bozzini M., Rossini M., *Testing Methods for 3D Scattered Data Interpolation*. Department of Mathematics and Applications, Univ. of Milano Bicocca, I-20126, Milano, Italy; Monografías de la Academia de Ciencias de Zaragoza 01/2002; 20:111-135.

SCATTERED DATA INTERPOLATION USING RADIAL BASIS FUNCTIONS, PART 2

Abstract

The paper is devoted to the problem of radial basis function (RBF) interpolation from scattered data. In this part of the article the methods of choosing the shape parameter are presented and the impact which the shape parameter has on the error of interpolation is analyzed. A numerical experiment confirms that the shape parameter has great impact on the accuracy of the RBF interpolation.

Autorzy:

mgr inż. **Dorota Borkowska** – Politechnika Rzeszowska