



AGH

Akademia Górniczo-Hutnicza
im. Stanisława Staszica
w Krakowie

Praca magisterska

Aproksymacja funkcji kawałkami regularnych przy
użyciu informacji dokładnej i niedokładnej

Tomasz Czyż

Kierunek: Matematyka
Specjalność: Matematyka obliczeniowa
i komputerowa

Nr albumu: 290565

Promotor
dr Maciej Goćwin



Wydział Matematyki Stosowanej

Kraków 2019

Oświadczenie studenta

Upředzony(-a) o odpowiedzialności karnej na podstawie art. 115 ust. 1 i 2 ustawy z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych (t.j. Dz.U. z 2018 r. poz. 1191 z późn. zm.): „Kto przywłaszcza sobie autorstwo albo wprowadza w błąd co do autorstwa całości lub części cudzego utworu albo artystycznego wykonania, podlega grzywnie, karze ograniczenia wolności albo pozbawienia wolności do lat 3. Tej samej karze podlega, kto rozpowszechnia bez podania nazwiska lub pseudonimu twórcy cudzy utwór w wersji oryginalnej albo w postaci opracowania, artystyczne wykonanie albo publicznie zniekształca taki utwór, artystyczne wykonanie, fonogram, wideogram lub nadanie.”, a także upředzony(-a) o odpowiedzialności dyscyplinarnej na podstawie art. 307 ust. 1 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2018 r. poz. 1668 z późn. zm.) „Student podlega odpowiedzialności dyscyplinarnej za naruszenie przepisów obowiązujących w uczelni oraz za czyn uchylający godności studenta.”, oświadczam, że niniejszą pracę dyplomową wykonałem(-am) osobiście i samodzielnie i nie korzystałem(-am) ze źródeł innych niż wymienione w pracy. Jednocześnie Uczelnia informuje, że zgodnie z art. 15a ww. ustawy o prawie autorskim i prawach pokrewnych Uczelni przysługuje pierwszeństwo w opublikowaniu pracy dyplomowej studenta. Jeżeli Uczelnia nie opublikowała pracy dyplomowej w terminie 6 miesięcy od dnia jej obrony, autor może ją opublikować, chyba że praca jest częścią utworu zbiorowego. Ponadto Uczelnia jako podmiot, o którym mowa w art. 7 ust. 1 pkt 1 ustawy z dnia 20 lipca 2018 r. — Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2018 r. poz. 1668 z późn. zm.), może korzystać bez wynagrodzenia i bez konieczności uzyskania zgody autora z utworu stworzonego przez studenta w wyniku wykonywania obowiązków związanych z odbywaniem studiów, udostępniać utwór ministrowi właściwemu do spraw szkolnictwa wyższego i nauki oraz korzystać z utworów znajdujących się w prowadzonych przez niego bazach danych, w celu sprawdzania z wykorzystaniem systemu antyplagiatowego. Minister właściwy do spraw szkolnictwa wyższego i nauki może korzystać z prac dyplomowych znajdujących się w prowadzonych przez niego bazach danych w zakresie niezbędnym do zapewnienia prawidłowego utrzymania i rozwoju tych baz oraz współpracujących z nimi systemów informatycznych.

.....
(Podpis czytelny studenta)

Oświadczenie promotora

Oświadczam, że praca spełnia wymogi stawiane pracom magisterskim.

.....
(Podpis promotora)

Spis treści

Streszczenie	2
Abstract	3
Wprowadzenie	4
Rozdział 1. Definicje	5
1.1. Informacja, algorytm, aproksymacja	5
1.2. Model obliczeniowy	7
1.3. Klasy funkcji	7
Rozdział 2. Ograniczenia na błąd	9
2.1. Ograniczenie z dołu	9
2.2. Ograniczenia z góry	10
Rozdział 3. Algorytmy	12
3.1. Algorytm oparty na wielomianach Lagrange’a	12
3.2. Algorytm oparty na różnicach dzielonych	13
Rozdział 4. Analiza algorytmów	16
4.1. Analiza algorytmu opartego o wielomiany Lagrange’a	16
4.2. Analiza algorytmu opartego o różnice dzielone	18
Rozdział 5. Testy numeryczne	24
Bibliografia	25

Streszczenie

Streszczenie

Słowa kluczowe

słowa kluczowe

Abstract

Abstract

Key words

keywords

Wprowadzenie

Aproksymacja funkcji oparta na dostępnej informacji jest problemem badanym od lat. Powstają coraz bardziej zaawansowane algorytmy, działające przy coraz słabszych założeniach o funkcji. Często jednak w rozważaniach teoretycznych pomijany jest czynnik zewnętrzny, który może powodować zaburzenia dostępnych informacji. W tej pracy pokazujemy, jak znaczący wpływ na wyniki numeryczne może mieć zaniedbanie tego faktu.

W tym celu przedstawimy dwa algorytmy z artykułów [5] i [4]. Pierwszy z nich bazuje na wielomianach Lagrange’a i jego analiza nie uwzględnia zaburzenia danych wejściowych. Algorytm z [5] dopuszcza informacje niedokładną, a kluczowy krok opiera się na różnicach dzielonych.

Omniane algorytmy aproksymują funkcje kawałkami regularne. Mówimy, że funkcja skalarna g jest (r, ϱ) -regularna na przedziale $[a, b]$, jeśli $g \in C^r([a, b])$ oraz g^r jest Hölderowsko ciągła z wykładnikiem $\varrho \in (0, 1]$. Rozważmy przestrzeń $F_{r, \varrho}$ T -okresowych funkcji f , które składają się z dwóch (r, ϱ) -regularnych części oddzielonych nieznany punktem osobliwym s_f .

Oba algorytmy osiągają ten sam minimalny błąd najgorszego przypadku proporcjonalny do $n^{-(r+\varrho)}$ dzięki zastosowaniu adaptacji. Adaptacyjny wybór punktów siatki jest niezbędny, aby otrzymać taki błąd. Ograniczenia związane z algorytmami nieadaptacyjnymi zostały omówione w [3], [2] oraz [?] (...).

Rozdział 1

Definicje

1.1. Informacja, algorytm, aproksymacja

W tym rozdziale wyjaśnimy co rozumiemy przez aproksymację i w jaki sposób ją otrzymujemy. W tym celu wprowadzimy fundamentalne pojęcia, takie jak operator rozwiązania, informacja zaburzona oraz algorytm. Szczególną uwagę poświęcimy informacji, która jest najważniejszym czynnikiem naszej analizy. Informacja, mówiąc w skrócie, jest tym co wiemy o problemie do rozwiązania. W niniejszej pracy kluczowym założeniem jest to, że informacja jest *zaburzona*, to znaczy, nie jest dokładna, czyli ma jakiś błąd.

Niech F będzie przestrznią liniową a G przestrzenią unormowaną, obie nad ciałem liczb rzeczywistych. Odwzorowanie

$$S : F \rightarrow G$$

nazywamy *operatorem rozwiązania*. Dla każdego elementu f z F chcemy obliczyć aproksymację $S(f)$. Niech $U(f)$ będzie obliczoną aproksymacją.

Niech $\varepsilon \geq 0$. Mówimy, że $S(f)$ jest ε -aprosymacją funkcji f wtw, gdy $\|S(f) - U(f)\| \leq \varepsilon$. Celem jest znalezienie takiej aproksymacji $U(f)$, że jest ona ε -aprosymacją dla wszystkich elementów f z F . Aby to zrobić potrzebujemy posiadać pewną wiedzę f .

Operatorem informacji (lub informacją) nazywamy odwzorowanie

$$N : F \rightarrow 2^Y,$$

gdzie Y jest zbiorem skończonych ciągów liczb rzeczywistych, $Y \subset \bigcup_{n=1}^{\infty} \mathbb{R}^n$, czyli $N(f)$ jest podzbiorem Y . Na ogół nie mamy dostępu do pełnej wiedzy o funkcji, dlatego musimy założyć, że możemy zbierać informacje o f poprzez formę (?tłum?) $L(f)$, gdzie $L : F \rightarrow H$ dla pewnego zbioru H .

Przez Λ oznaczmy klasę dopuszczalnych operacji L , czyli $L \in \Lambda$ wtw, gdy $L(f)$ może zostać obliczone dla każdego elementu f z F . Rozważmy teraz dwa sposoby doboru informacji. Jeśli informacja N jest liczona dla każdego f niezależnie, to taką informację nazywamy nieadaptacyjną. Innymi słowy, informacja N jest nieadaptacyjna wtw, gdy istnieje $L_1, \dots, L_n \in \Lambda$ takie, że

$$N(f) = [L_1(f), L_2(f), \dots, L_n(f)] \quad \forall f \in F.$$

Inna klasa informacji jest informacja adaptacyjna, w której możemy wybierać wartości bazując na poprzednich wynikach. Mówiąc dokładniej, informacja N jest adaptacyjna wtw, gdy

$$N(f) = [L_1(f), L_2(f; y_1), \dots, L_i(f; y_1, \dots, y_{n(f)-1})],$$

gdzie $y_1 = L(f_1)$ i $y_i = L_i(f; y_1, y_2, \dots, y_{i-1})$ dla $i = 2, 3, \dots, n(f)$. Musimy również założyć, że $L_i(\cdot; y_1, \dots, y_{i-1})$ należą do operacji dozwolonych. W przypadku informacji adaptacyjnej nie możemy z góry określić liczby operacji na problemie f , ponieważ jest to dynamicznie ustalone podczas procesu obliczenia kolejnych wartości y_i .

Warto zauważyć, że jeśli rozważany problem wymaga obliczenia bardzo dużej ilości informacji o funkcji w krótkim czasie, to zastosowanie informacji nieadaptacyjnej może przyspieszyć obliczenia, ze względu na możliwość zrównoleglenia obliczeń. W przypadku adaptacyjnym kolejność obliczeń ma znaczenie, więc informacje musimy pozyskiwać sekwencyjnie.

Możemy zapisać informacje nieadaptacyjną i adaptacyjną zakładając, że f jest funkcją oraz $N(f) = [f(t_1), f(t_2), \dots, f(t_{n(f)})]$. Jeżeli $n(f) = n$ i punkty t_i otrzymujemy *a priori*, wtedy N jest nieadaptacyjna. Natomiast, jeżeli $n(f)$ różni się lub wybór punktów t_i jest zależny od $f(t_1), f(t_2), \dots, f(t_i)$, to N jest adaptacyjna.

Kolejnym kluczowym założeniem jest fakt, że informacje o funkcji zazwyczaj uzyskujemy z pewnym błędem. Mówiąc dokładniej, informacje o funkcji przyjmują postać $y_i = f(x_i) + e_i$, dla $1 \leq i \leq n$, gdzie $e_i \leq \delta$ nazywamy *szumem*.

Dla przykładu, dla informacji nieadaptacyjnej składającej się z zaburzonych ewaluacji funkcji f w punktach x_1, \dots, x_n z precyzją δ , mamy:

$$N(f) = \{y \in \mathbb{R}^n : |y_i - f(x_i)| \leq \delta, \quad 1 \leq i \leq n\}$$

Każdy element $y \in N(f)$ będziemy nazywać *informacją o f* . Zauważmy, że dla $\delta = 0$, zbiór $N(f)$ ma dokładnie jeden element dla wszystkich $f \in F$, tzn. informacja N jest *dokładna*. Zakładamy, że $N(f)$ jest niepuste dla wszystkich $f \in F$. W przypadku, gdy istnieje f dla którego $N(f)$ ma przynajmniej dwa elementy, wtedy informacja jest *niedokładna (częściowa)*.

Naszym zadaniem jest aproksymacja elementów $S(f)$ dla f należącego do $E \subset F$, bazując wyłącznie na informacji zaburzonej o f .

Znając informację y o f możemy wprowadzić aproksymację, a dokładniej algorytm, który ją wyprodukuję. *Algorytmem* nazywamy odwzorowanie:

$$\varphi : Y \rightarrow G$$

Innymi słowy, aproksymacją $S(f)$ jest $\varphi(y)$, gdzie y jest informacją o f . Błąd aproksymacji zdefiniowany jest jako różnica $\|S(f) - \varphi(y)\|$, gdzie $\|\cdot\|$ jest normą w przestrzeni G .

1.2. Model obliczeniowy

W ogólności, optymalność algorytmu oraz jego złożoność zależą od przyjętego modelu obliczeniowego. Model jest określony poprzez sposób w jaki błąd i koszt algorytmu są zdefiniowane.

Jeżeli za błąd i koszt przyjmujemy wydajność na najtrudniejszym spośród wszystkich problemów w danej klasie, wtedy mówimy o *modelu najgorszego przypadku*. Innymi często rozważanymi modelami są: probabilistyczny, średni, mieszany, losowy czy asymptotyczny, jednak nimi nie będziemy zajmować się w tej pracy.

Niech $N : R \rightarrow 2^Y$ będzie operatorem informacji. Poprzez *błąd najgorszego przypadku* algorytmu $\varphi : Y \rightarrow G$ na zbiorze $E \subset F$ rozumiemy:

$$e_p^{wor}(\phi, N, E) = \sup_{f \in E} \sup_{y \in N(f)} \|S(f) - \varphi(y)\|$$

Oznaczmy przez $\mathcal{N}(n, \delta)$ klasę wszystkich (adaptacyjnych) informacji N , które używają co najwyżej n ewaluacji funkcji, z precyzją δ każda. Wtedy przez *minimalny błąd najgorszego przypadku* w klasie E , który może zostać osiągnięty przez algorytm używający informacji o co najwyżej n wartościach funkcji z precyzją δ rozumiemy:

$$r_p^{wor}(n, \delta, E) = \inf \{e_p^{wor}(\varphi, N, E) : \varphi \text{ używa } N \in \mathcal{N}(n, \delta)\}$$

W tej pracy porównujemy dwa algorytmy aproksymujące funkcje z osobliwością, które osiągają to samo, optymalne organiczenie na minimalny błąd najgorszego przypadku, jednak tylko w przypadku jednego z nich w analizie zostało uwzględnione zaburzenie.

1.3. Klasy funkcji

Dla liczby całkowitej $r \geq 0$, $\varrho \in (0, 1]$ oraz $a < b$, przez $H_{r, \varrho}(a, b)$ oznaczamy przestrzeń funkcji $g : [a, b] \rightarrow \mathbb{R}$ takich, że $g \in C^r([a, b])$ i $g^{(r)}$ jest Hölderowsko ciągła z wykładnikiem ϱ , tzn.

$$c(g) := \sup_{a \leq x \leq y \leq b} \frac{|g^{(r)}(x) - g^{(r)}(y)|}{|x - y|^\varrho} < \infty.$$

Dla danego $T > 0$ niech $F_{r, \varrho} = F_{r, \varrho}(T)$ będzie przestrzenią funkcji $f : \mathbb{R} \rightarrow \mathbb{R}$ spełniających następujące warunki: istnieje $s_f \in [0, T]$ i $g_f \in H_{r, \varrho}(0, T)$ takie, że

$$f(lT + s_f + x) = g_f(x) \quad \text{for all } l = 0, \pm 1, \pm 2, \dots \quad \text{i } x \in [0, T]$$

Można powiedzieć, że f jest 'kopią' g_f na każdym z przedziałów $(lT + s_f, (l+1)T + s_f]$ i f jest prawostronnie ciągła na $lT + s_f$. W związku z tym wszystkie punkty, które różnią się między sobą o wielokrotność T będą uważane za identyczne. Dla przykładu, jeżeli $0 < x_1 \leq T < x_2 \leq 2T$, to przedział $(x_1, x_2]$ będzie utożsamiany z $(x_1, T] \cup (0, x_2 - T] \subset (0, T]$.

Przez $\Delta_f^{(j)}$ oznaczmy *skoki nieciągłości* dla kolejnych pochodnych f w punkcie nieciągłości s_f ,

$$\Delta_f^{(j)} = f^{(j)}(s_f^+) - f^{(j)}(s_f^-) = g_f^{(j)}(0) - g_f^{(j)}(T) \quad 0 \leq j \leq r,$$

W tej pracy będziemy rozpatrywać różne aproksymacje $\varphi : Y \rightarrow L^p(0, T)$ funkcji $f \in F_{r,\varrho}$ względem normy L^p , gdzie $1 \leq p \leq \infty$. Czyli z definicji normy L_p , błąd aproksymacji funkcji f dla informacji y wynosi:

$$\|f - \varphi(y)\|_{L^p} = \left(\int_0^T |(f - \varphi(y))(x)|^p dx \right)^{1/p} \quad \text{dla } 1 \leq p < \infty$$

oraz

$$\|f - \varphi(y)\|_{L^\infty} = \operatorname{ess\,sup}_{0 < x \leq T} |(f - \varphi(y))(x)|$$

Rozróżniamy następujące klasy \mathcal{F} :

$$\begin{aligned} \mathcal{K} &= \{f = c\mathbb{1}_{\mathbb{R}}, c \in \mathbb{R}\}, \\ \mathcal{H}_{r,\varrho} &= \{f \in F_{r,\varrho} : c(g_f) \leq 1, \Delta_f^{(j)} = 0 \text{ for all } 0 \leq j \leq r\} \\ \mathcal{F}_{r,\varrho}^C &= \{f \in F_{r,\varrho} : c(g_f) \leq 1, \Delta_f^{(0)} = 0\} \\ \mathcal{F}_{r,\varrho}^D &= \{f \in F_{r,\varrho} : c(g_f) \leq 1, |\Delta_f^{(0)}| \leq 1\} \\ \mathcal{F}_{r,\varrho} &= \{f \in F_{r,\varrho} : c(g_f) \leq 1\} \end{aligned}$$

Oczywiście

$$\mathcal{K} \subset \mathcal{H}_{r,\varrho} \subset \mathcal{F}_{r,\varrho}^C \subset \mathcal{F}_{r,\varrho}^D \subset \mathcal{F}_{r,\varrho}$$

Algorytm \mathcal{A}^{KP} przedstawiony w pracy [4], oryginalnie bazuje na lekko zmodyfikowanej klasie funkcji, co wynika z innej natury problemu rozważanego w pracy. Dla spójności i lepszego przedstawiania problemu, wprowadzimy oryginalnie rozważaną klasę funkcji i sprecyzujemy różnice między klasami przedstawionymi wcześniej.

$$\begin{aligned} G_{reg}^{r,\varrho}([a, b]) &= \{g : [a, b] \rightarrow \mathbb{R}^d \mid g \in C^{(r)}([a, b]), \|g^{(j)}(t)\| \leq D_j, j = 0, 1, \dots, r \\ &\quad \|g^{(r)}(t) - g^{(r)}(s)\| \leq L_r |t - s|^\varrho, \|g(s) - g(t)\| \leq L_0 |s - t|, t, s \in [a, b]\} \end{aligned}$$

Rozważmy następującą klasę $G_{reg}^{r,\varrho}([a, b])$ funkcji $g : [a, b] \rightarrow \mathbb{R}^d$ z co najwyżej jednym (nieznanym) punktem osobliwym \hat{t}_g . Funkcja $g : [a, b] \rightarrow \mathbb{R}^d, g = [g^1, g^2, \dots, g^d]^T$, należy do $G_{reg}^{r,\varrho}([a, b])$, jeśli istnieje punkt $\hat{t}_g \in (a, b)$ taki, że

$$g \in G_{reg}^{r,\varrho}([a, \hat{t}_g]) \cap G_{reg}^{r,\varrho}([\hat{t}_g, b])$$

oraz lewostronna granica każdej składowej pochodnej $(g^k)^{(j)}$ istnieje w \hat{t}_g . W punkcie osobliwym \hat{t}_g , pochodne są rozumiane jako prawostronne.

Dla $g \in G_{reg}^{r,\varrho}([a, b])$, definiujemy wektor skoków

$$\Delta_g^j = g^{(j)}(\hat{t}_g+) - g^{(j)}(\hat{t}_g-), \quad j = 0, 1, \dots, r$$

oraz wielomian

$$s_g(t) = \sum_{j=0}^r \frac{1}{j!} \Delta_g^j (t - \hat{t}_g)^j, \quad t \in [\alpha, \beta]$$

Jeżeli $\Delta_g^j = 0$ dla wszystkich $j = 0, 1, \dots, r$, wtedy g jest regularna, czyli $g \in G_{reg}^{r,\varrho}([a, b])$. Jeżeli $g \in G_{reg}^{r,\varrho}([a, b])$ and $\Delta_g^0 = 0$, wtedy g jest Lipschitzowsko ciągła w $[a, b]$.

Ograniczenia na błąd

2.1. Ograniczenie z dołu

Do określania błędu będziemy używać notacji Ω , Θ , \mathcal{O} , o (wersji Knutha). Gdy dwie funkcje f i g zdefiniowane na \mathbb{R}_+ i przyjmują wartości nieujemne, to piszemy $f(x) = \Omega(g(x))$ wtedy, gdy istnieją dodatnie stałe c_1 i c_2 takie, że $f(x) \geq c_1 g(x)$ dla $x \in [0, c_2]$. Przez $f(x) = \Theta(g(x))$ rozumiemy $f(x) = \Omega(g(x))$ and $g(x) = \Omega(f(x))$, czyli istnieją takie stałe c_1 , c_2 i c_3 , że $c_1 g(x) \leq f(x) \leq c_2 g(x)$ dla $x \in [0, c_3]$. Natomiast przez $f(x) = \mathcal{O}(g(x))$ rozumiemy, że $g(x) = \Omega(f(x))$, a przez $f(x) = o(g(x))$ rozumiemy, że $\lim_{x \rightarrow 0} \frac{f(x)}{g(x)} = 0$.

Na początku przytoczymy znane wyniki dotyczące ograniczeń z dołu dla informacji dokładnej, które uzasadniają użycie algorytmów adaptacyjnych. Wiadomo, że w klasie funkcji r -regularnych, najlepszym tempem zbieżności dla błędu jest n^{-r} . W pracy [2] udowodniono, że wprowadzenie osobliwości, powoduje pogorszenie się tej własności dla algorytmów nieadaptacyjnych do $n^{1/p}$. Pokazuje to następujące twierdzenie.

Twierdzenie 2.1. *Niech \mathcal{A}_n będzie dowolnym algorytmem nieadaptacyjnym oraz niech $\Delta > 0$. Istnieje kawałkami stała funkcja $f \in F_{r,1}^1$ taka, że $|\Delta_1^{(0)}| \leq \Delta$ oraz*

$$\|f - \mathcal{A}_n f\|_{L^p} \geq \frac{1}{2} \Delta \left(\frac{T}{n+1} \right)^{1/p}$$

Dowód. Załóżmy, że \mathcal{A}_n oblicza f w punktach $x_0 < \dots < x_n$. Niech $x_0 = 0$ i $x_n = T$. Istnieją $0 < a < b < T$ takie, że $b - a \geq T/(n+1)$ i $[a, b] \subset [x_k, x_{k+1}]$ dla pewnego $0 \leq k \leq n$. Weźmy teraz funkcje $g_1 = \Delta \mathbb{1}_{(a,T]}$ oraz $g_2 = \Delta \mathbb{1}_{(b,T]}$. Ponieważ g_1 i g_2 używają tej samej informacji, tj. $g_1(x_i) = g_2(x_i)$ dla wszystkich $1 \leq i \leq n$ i $\|g_1 - g_2\|_{L^p} \geq \Delta(T/(n+1))^{1/p}$, to błąd algorytmu nie może być mniejszy niż $\Delta(T/(n+1))^{1/p}/2$ dla przynajmniej jednej z funkcji g_i . Co dowodzi tezę. \square

Wiemy już jakie minimalne błędy mogą zostać osiągnięte przez algorytmy aproksymujące bazujące na informacji dokładnej. Poniższe stwierdzenie wprowadza kilka własności wynikających z znanych rezultatów przy obecności zaburzenia danych.

Stwierdzenie 2.2. *Dla każdego n i $\delta \geq 0$ mamy:*

- (i) $r_p^{wor}(n, \delta, \mathcal{K}) \geq \delta T^{1/p}$
- (ii) $r_p^{wor}(n, \delta, \mathcal{H}_{r,\varrho}) \geq a_{r,\varrho} n^{-(r+\varrho)}$ dla pewnego $a_{r,\varrho} > 0$
- (iii) $r_p^{wor}(n, \delta, \mathcal{F}_{r,\varrho}) = \infty$, $r \geq 1$

Dowód. W celu udowodnienia (i) wystarczy zauważyć, że $y = (0, \dots, 0)$ jest informacją o funkcji stałej postaci $f_{\pm} = \pm\delta$ dla każdego N z precyzją δ . Wynika z tego, że error dowolnego algorytmu używającego N jest równy conajmniej $\|f_{+\delta} - f_{-\delta}\|_{L^p}/2 = \delta T^{1/p}$.

Nierówność (ii) wynika z znanych rezultatów dotyczących minimalnego błędu aproksymacji dla informacji dokładnej, zobacz [?].

Aby pokazać (iii), użyjemy rozumowania podobnego do [2] [sekcja, 5.2], gdzie przeprowadzono dowód dla $\varrho = 1$ i $\delta = 0$. Niech $S(M) \subset \mathcal{F}_{r,\varrho}$ będzie rodziną funkcji f_s dla $s \in [0, T)$

$$f_s(x) = \frac{M}{T} (x \mathbb{1}_{[0,s)}(x) + (x - T) \mathbb{1}_{[s,T)}(x)), \quad 0 \leq x \leq T$$

Niech N będzie dowolną (adaptacyjną) informacją używającą nie więcej niż n ewaluacji funkcji. Ponieważ dla każdego ustalonego x , funkcja $f_s(x)$ może przyjmować tylko dwie wartości w zależności czy $s \leq x$ lub $s > x$, to całkowita liczba punktów użytych przez N dla klasy $S(M)$ wynosi co najwyżej $2^n - 1$. Dlatego istnienie przedział $[s_1, s_2] \subset (0, T)$ o długości $T2^{-(n-1)}$, który nie zawiera żadnego z tych punktów. To oznacza, że $N(f_{s_1}) = N(f_{s_2})$, a więc błąd dowolnego algorytmu używającego informacji N wynosi przynajmniej $\|f_{s_1} - f_{s_2}\|_{L^p}/2 = \delta M(T2^{-(n+p+1)})^{1/p}$. Z uwagi na to, że M jest dowolnie duże, błąd również może być dowolnie duży. \square

Stwierdzenie (iii) mówi, że nie możemy uogólnić wyników na klasę kawałkami Hölderowskich funkcji $\mathcal{F}_{r,\varrho}$ z $r \geq 1$. Z tego powodu rozważania będziemy prowadzić głównie na klasie $\mathcal{F}_{r,\varrho}^D$ funkcji kawałkami Hölderowskich z jednostajnie ograniczonymi skokami nieciągłości $\Delta_f^{(0)}$ oraz na klasie $\mathcal{F}_{r,\varrho}^C$ funkcji kawałkami Hölderowskich ciągłych.

Podsumowując, z stwierdzenia 2.2 (i)-(ii) otrzymujemy ograniczenia z dołu

$$r_p^{wor}(n, \delta, \mathcal{F}_{r,\varrho}^D) \geq r_p^{wor}(n, \delta, \mathcal{F}_{r,\varrho}^C) \geq \max(\delta T^{1/p}, a_{r,\varrho} n^{-(r+\varrho)})$$

W dalszej części pracy udowodnimy, że te nierówności są ostre, z wyjątkiem pierwszej dla $p = \infty$. To jest główny wynik artykułu [5].

2.2. Ograniczenia z góry

Dla uproszczenia, w dalszej części pracy przyjmujemy oznaczenia pochodzące od pierwszych liter nazwisk autorów poszczególnych artykułów. Oznaczenia \mathcal{A}^{KP} oraz φ^{KP} odnoszą się odpowiednio do algorytmu z pracy [4] opartego na wielomianach Lagrange'a oraz do aproksymacji zwróconej przez ten algorytm. Analogicznie, oznaczenia \mathcal{A}^{MP} oraz φ^{MP} odnoszą się odpowiednio do algorytmu z pracy [5] opartego na różnicach dzielonych oraz do aproksymacji zwróconej przez ten algorytm

Górne ograniczenia na błąd otrzymujemy poprzez analizę skonstruowanych algorytmów, szczegóły w rozdziale 4.

Omawiane algorytmy osiągają te same ograniczenia z góry, jednak jak wspomnieliśmy, algorytm przedstawiony w pracy [4] bazuje na informacji dokładnej, w przeciwieństwie do algorytmu z pracy [5], który uwzględnia zaburzenie danych. Ta różnica wpłynęła na to, że do uzyskania tych samych wyników autorzy doszli w odmienny sposób. Aby lepiej przedstawić przebieg rozumowania, nie uogólniamy wyników, które są bardziej szczegółowe niż założenia tej pracy wymagają. Tyczy się to przede wszystkim algorytmu \mathcal{A}^{KP} , ponieważ jest on tylko częścią rozwiązania innego problemu, który jest tematem pracy [4]. Analiza błędów algorytmu \mathcal{A}^{MP} opiera się na badaniu właściwości jego poszczególnych kroków. Oszacowanie błędów dla \mathcal{A}^{KP} wynika z dokładnej analizy własności wielomianów Lagrange’a i testu na nich opartego.

Poniższe twierdzenia prezentują główne właściwości algorytmów.

Twierdzenie 2.3 (Błąd algorytmu \mathcal{A}^{MP}).

1. $r_p^{wor}(n, \delta, \mathcal{F}_{r,\varrho}^D) = \Theta(\max(\delta, n^{-(r+\varrho)}))$ dla $1 \leq p \leq \infty$
2. $r_\infty^{wor}(n, \delta, \mathcal{F}_{r,\varrho}^C) = \Theta(\max(\delta, n^{-(r+\varrho)}))$

Twierdzenie 2.4 (Błąd algorytmu \mathcal{A}^{KP}). *Niech $r + \varrho \geq 1$. Istnieją stałe C i m_0 takie, że dla $m > m_0$, przedziału $[a, b]$ oraz wszystkich $g \in G^{r,\varrho}([a, b])$ z $\Delta_g^0 = 0$ zachodzi:*

$$\sup_{t \in [a, b]} \|g(t) - q(t)\| \leq C m^{-(r+\varrho)}$$

Uwaga 2.5. *Oba algorytmy używają $\mathcal{O}(m)$ ewaluacji wartości aproksymowanej funkcji. Dodatkowo w przypadku \mathcal{A}^{MP} otrzymujemy również informacje o progu dla m , którego przekroczenie nie zwiększy dokładności otrzymanego przybliżenia, $m_{max} = \Theta(\delta^{-(r+\varrho)})$. Ta informacja jest istotna ze względu na precyzję reprezentacji liczb na komputerze. Wspomniane własności są wyprowadzone w rozdziale 4.*

Rozdział 3

Algorytmy

3.1. Algorytm oparty na wielomianach Lagrange’a

Algorytm przedstawiony w pracy [4] lokalizuje osobliwość przy pomocy wielomianów Lagrange’a w_g^r . Na wejściu algorytm otrzymuje $g \in G^{r,e}([a, b])$, przedział $[a, b]$, regularność r oraz współczynnik Höldera ϱ . Kluczowym elementem algorytmu jest zdefiniowana poniżej wielkość (*test*), która jest użyta do wykrycie punktu osobliwego.

$$A_g(a, \bar{a}, \bar{b}, b) = \max_{0 \leq j \leq r} \frac{\|w_g^r([\bar{b}, b])(z_j) - w_g^r([a, \bar{a}]) (z_j)\|}{\bar{h}^{r+e}}, \quad (3.1)$$

gdzie $a < \bar{a} < \bar{b} < b$, $z_j = \bar{a} + (\bar{b} - \bar{a})j/r$, dla $j = 0, 1, \dots, r$ oraz $\bar{h} = b - a$ jest długością przedziału, na którym *test* jest zdefiniowany.

Algorytm oparty na wielomianach Lagrange'a

- K1:** Niech $\omega := h^{r+e}$, $B := \{a, b\}$
jeżeli $\max_{0 \leq i \leq p-1} (c_{i+1} - c_i) \leq 4\omega$ **wtedy**
 idź do *Krok 3*
- w p.p.**
 Niech, dla $j = 0, 1, \dots, p-1$,
 $A = \max \{A_g(c_j, c_j + \omega, c_{j+1} - \omega, c_{j+1}) \mid c_{j+1} - c_j > 4\omega\}$
 Niech $A_g^i = A_g(c_i, c_i + \omega, c_{i+1} - \omega, c_{i+1})$
jeżeli istnieją różne k i l takie, że $A = A^k \wedge A = A^l$ **wtedy**
 idź do *K3*
- K2:** Niech $[c_k, c_{k+1}]$ - przedział otrzymany w *K1* oraz niech $[a_1, b_1] = [a, b]$
dopóki $b_1 - a_1 > \omega$ **wykonuj:**
 Oblicz $v = (a_1 + b_1)/2$ oraz $B = B \cup \{v\}$
jeżeli $A_g(a_1, a_1 + \omega, v - \omega, v) = A_g(v, v + \omega, b_1 - \omega, b_1)$ **wtedy**
 idź do *K3*
- w p.p.**
 za następny przedział $[a_1, b_1]$ wybierz podprzedział $[a_1, v]$ lub $[v, b_1]$,
 dla którego wartość testu była większa
- K3:** Niech $\bar{M} = \{c_0, \dots, c_p\} \cap B$
- $$q(t) = \begin{cases} g(c_i) & \text{gdy } t \in [c_i, c_{i+1}) \wedge c_{i+1} - c_i \leq 4\omega \\ g(c_i) & \text{gdy } t \in [c_i, c_i + \omega) \wedge c_{i+1} - c_i > 4\omega, \\ w_g^r([c_i + \omega, c_{i+1} - \omega])(t) & \text{gdy } t \in [c_i + \omega, c_{i+1} - \omega) \wedge c_{i+1} - c_i > 4\omega \\ g(c_{i+1} - \omega) & \text{gdy } t \in [c_{i+1} - \omega, c_{i+1}) \wedge c_{i+1} - c_i > 4\omega \end{cases}$$
- dla $i = 0, 1, \dots, k-1$ z $q(b) =$ zdefiniowanym przez ciągłość na ostatnim przedziale
 $B := B \cup \{c_i + \omega, c_{i+1} - \omega \mid c_{i+1} - c_i > 4\omega, i = 0, \dots, k-1\}$

3.2. Algorytm oparty na różnicach dzielonych

W tym rozdziale opiszemy algorytm bazujący na informacji zaburzonej przedstawiony w artykule [5]. Analizowany algorytm używa co najwyżej n wartości funkcji z precyzją δ oraz w najgorszym przypadku ma błąd proporcjonalny do $\max(\delta, n^{-1/r+e})$ w klasie funkcji $\mathcal{F}_{r,e}^D$ dla $p < \infty$ oraz w klasie $\mathcal{F}_{r,e}^C$ dla $p \leq \infty$. Kluczowym parametrem algorytmu jest

$$h = T/m \quad \text{with} \quad m \geq 2r + 1,$$

gdzie m jest początkową gęstością siatki. Dodatkowo, niech $\omega = \omega(h)$ spełnia $0 < \omega < (r+1)h$.

Na początku algorytm aproksymuje punkt osobliwy s_f . Jest to realizowane w trzech krokach. W kroku 1. przy pomocy siatki rozmiarze o długości h i różnic dzielonych lokalizowany jest punkt s_f na przedziale $[u_1, v_1]$ o długości $(r+1)h$. W kroku 2. używamy wielomianów interpolujących \tilde{p}_+ i \tilde{p}_- do zwężenia tego przedziału do $[u_2, v_2]$. Krok 3.

produkuje przedział $[u_3, v_3] \subseteq [u_2, v_2]$, w którym różnica $|\tilde{p}_+ - \tilde{p}_-|$ jest nierosnąca na $[u_3, \xi]$ i niemalejąca na $[\xi, v_3]$, gdzie ξ jest finalną aproksymacją s_f .

Powyższe kroki mogą być zapisane następująco; dla $t_i = ih \ \forall i$.

Algorytm oparty na różnicach dzielonych

Krok 1 Oblicz różnice dzielone $\tilde{d}_i = \tilde{f}[t_i, \dots, t_{i+r+1}]$ for $1 \leq i \leq m$ oraz znajdź

$$i^* = \arg \max_{1 \leq i \leq m} |\tilde{d}_i|$$

Niech $u_1 = t_{i^*}$ i $v_1 = t_{i^*+r+1}$.

Krok 2 Oznaczmy przez \tilde{p}_+ i \tilde{p}_- wielomiany stopnia $\leq r$, które interpolują węzły $(t_j, \tilde{f}(t_j))$ odpowiednio dla $i^* - r \leq j \leq i^*$ oraz dla $i^* + r + 1 \leq j \leq i^* + 2r + 1$. Następnie wykonaj iterację:

$u := u_1, v := v_1$

dopóki $v - u > \omega$ **wykonuj**:

$$z_j := u + j(v - u)/(r + 2), \quad j = 1, 2, \dots, r + 1$$

$$j^* := \arg \max_{1 \leq j \leq r+1} |\tilde{p}_+(z_j) - \tilde{p}_-(z_j)|$$

jeżeli $|\tilde{f}(z_{j^*}) - \tilde{p}_-(z_{j^*})| \leq |\tilde{f}(z_{j^*}) - \tilde{p}_+(z_{j^*})|$ **wtedy**

$$u := z_{j^*}$$

w p.p.

$$v := z_{j^*}$$

koniec

Niech $u_2 = u$ i $v_2 = v$.

Krok 3 Wykonaj iterację:

$u := u_2, v := v_2$

dopóki istnieje maksimum lokalne $|\tilde{p}_+ - \tilde{p}_-|$ na (u, v) **wykonuj**

$z :=$ największe maksimum lokalne $|\tilde{p}_+ - \tilde{p}_-|$ na (u, v)

jeżeli $|\tilde{f}(z) - \tilde{p}_-(z)| \leq |\tilde{f}(z) - \tilde{p}_+(z)|$ **wtedy**

$$u := z$$

w p.p.

$$v := z$$

koniec

Niech $u_3 = u$ i $v_3 = v$.

Finalną aproksymacją s_f jest

$$\xi := \arg \max_{u_3 \leq x \leq v_3} |\tilde{p}_+ - \tilde{p}_-|$$

Przedstawiony algorytm używa m wartości funkcji w kroku 1 oraz jedną wartość funkcji w każdej iteracji w krokach 2 i 3. Czyli w kroku 2 używamy co najwyżej

$$\left\lceil \frac{\ln \left(\frac{(r+1)h}{\omega(h)} \right)}{\ln \left(\frac{r+2}{r+1} \right)} \right\rceil$$

wartości funkcji i $(r - 1)$ w kroku 3. Stąd otrzymujemy, że jeżeli $\omega = \omega(h) \geq kh^\alpha$ dla pewnego ustalonego k i α , wtedy w najgorszym przypadku liczba użytych wartości funkcji równa się asymptotycznie $m = \frac{T}{h}$ dla $h \rightarrow 0^+$.

Analiza algorytmów

4.1. Analiza algorytmu opartego o wielomiany Lagrange’a

Zacznijmy od wyjaśnienia własności testu 3.1 służącego do wykrywania osobliwości. Rozważmy błąd interpolacji Lagrange’a dla nieciągłej funkcji $g \in G^{r,e}([a, b])$. Błąd jest ograniczony za względu na wielomian s^g (TODO: reference).

Lemat 4.1. *Istnieje stała C taka, że dla wszystkich $[a, b]$, wszystkich $g \in G^{r,e}([a, b])$ oraz $s = 0, 1, \dots, r$, mamy*

$$\sup_{t \in [a, b]} \|g(t) - w_g^s([a, b])(t)\| \leq C \left(\min \left\{ \sup_{t \in [a, \hat{t}_g)} \|s_g(t)\|, \sup_{t \in [\hat{t}_g, b]} \|s_g(t)\| \right\} + \bar{h}^{\min\{s+1, r+e\}} \right)$$

Dowód. (...) □

Wniosek 4.2. *Istnieje stała C taka, że dla wszystkich $[a, b]$, wszystkich $g \in G^{r,e}([a, b])$ z $\delta_g^0 = 0$, $0 \leq \delta \leq \min 1, \bar{h}$ oraz $s = 0, 1, \dots, r$, mamy*

$$\hat{t}_g \in (a, a + \delta] \cup [b - \delta, b) \implies \sup_{t \in [a, b]} \|g(t) - w_g^s([a, b])(t)\| \leq C (\delta + \bar{h}^{\min\{s+1, r+e\}})$$

Dowód. (...) □

Lemat 4.3. *Istnieje stała C zależna od r i L_r taka, że dla wszystkich $[a, b]$, $\bar{a} \in (a, b)$, wszystkich $g \in G^{r,e}([a, b])$, mamy*

$$\hat{t}_g \in (\bar{a}, b) \implies g(t) - w_g^r([a, \bar{a}]) (t) = s_g(t) \mathbb{1}_{[\hat{t}_g, b]}(t) + R_g(t), \quad t \in [\bar{a}, b],$$

gdzie $\|R_g(t)\| \leq C \bar{h}^{r+e}$, dla $t \in [\bar{a}, b]$

Dowód. (...) □

Wniosek 4.4. Istnieje stała \bar{C} zależna od r i L_r taka, że dla wszystkich $[a, b]$, $\bar{a} \in (a, b)$, wszystkich $g \in G^{r,e}([a, b])$, mamy

$$\hat{t}_g \in (a, \bar{a}) \implies g(t) - w_g^r([\bar{a}, a])(t) = s_g(t) \mathbb{1}_{[a, \hat{t}_g]}(t) + R_g(t), \quad t \in [a, \bar{a}],$$

gdzie $\|R_g(t)\| \leq \bar{C} \bar{h}^{r+e}$, dla $t \in [a, \bar{a}]$

Dowód. (...)

□

Stwierdzenie 4.5. Istnieje stała C^* zależna od r i L_r taka, że dla wszystkich $a < \bar{a} < \bar{b} < b$ i $[a, b]$ oraz wszystkich $g \in G^{r,e}([a, b])$, mamy

$$\hat{t}_g \text{ z niezerowym wielomianem } s_g \text{ nie jest w } (a, b) \implies A_g(a, \bar{a}, \bar{b}, b) \leq C^*$$

Dowód. (...)

□

Uwaga 4.6. Stwierdzenie 4.5 pokazuje, że procedura LOCATE-APPROXIMATE wraz z procedurą BISECTION, sukcesywnie wybiera przedziały bazując a wartościach testu. Zauważmy, że jeżeli \hat{t}_g jest unikalna, to wtedy dla jakiegokolwiek przedziału $[a, b]$, który nie został wybrany, mamy $A_g(a, \bar{a}, \bar{b}, b) \leq C^*$

Stwierdzenie 4.7. Niech $D > 0$. Istnieją stałe C i \bar{N} , zależne tylko od parametrów klast $G^{r,e}([a, b])$ i D , takie, że dla wszystkich $[a, b]$, $[\bar{a}, \bar{b}] \subset (a, b)$, $g \in G^{r,e}([a, b])$ oraz $s = 0, 1, \dots, r$, mamy

$$\hat{t}_g \in (\bar{a}, \bar{b}] \wedge b - a \leq D(\bar{b} - \bar{a}) \implies \text{dla } [\gamma, \omega] = [a, b] \vee [\gamma, \omega] = [\bar{a}, \bar{b}] \text{ zachodzi}$$

$$\sup_{t \in [\gamma, \omega]} \|g(t) - w_g^s([\gamma, \omega])(t)\| \leq C (1 + A_g(a, \bar{a}, \bar{b}, b)) \bar{h}^{\min\{s+1, r+e\}}$$

oraz ponadto

$$\sup_{t \in [\gamma, \omega]} \left\| (w_g^s([\gamma, \omega]))^{(j)}(t) \right\| \leq \bar{N} (1 + \bar{h}^{\min\{s+1-j, r+e-j\}} + (1 + A_g(a, \bar{a}, \bar{b}, b)) \bar{h}^{r+e-j})$$

dla $j = 0, 1, \dots, s$.

Dowód. (...)

□

Uwaga 4.8. Stwierdzenie 4.7 pokazuje, że w przypadku z osobliwością, ograniczenie górne na błąd interpolacji możemy wyrazić za pomocą $A_g(a, \bar{a}, \bar{b}, b)$

Poniższy lemat dotyczy przypadku, gdy osobliwość znajduje się na brzegu przedziału $[a, b]$

(lemat... dowod...)

(Dowód tw1 z 2014)

Uwaga 4.9. Twierdzenie 2.4 zachodzi również dla funkcji g , która ma skok w punkcie c_i początkowego podziału M oraz ma niezerowy wielomian s_g dla co najwyżej jednego nieznanego punktu t_g , $t_g \neq c_i \forall i$.

Dodatkowo, obliczenie q wymaga $O(p + \log m)$ ewaluacji fun g , gdzie p jest liczbą przedziałów w początkowym podziale M przedziału $[a, b]$. Czyli, aby otrzymać optymalną aproksymację $g \in G^{r,e}([a, b])$ na przedziale $[a, b]$, wystarczy wziąć podział M z $m + 1$ równoodległymi punktami $x_i = a + (b - a)i/m$, dla $i = 0, 1, \dots, m$. wtedy obliczenie q wymaga $O(m)$ ewaluacji funkcji g .

4.2. Analiza algorytmu opartego o różnice dzielone

Zanim przejdziemy do analizy samego algorytmu, potrzebujemy przedstawić kilka właściwości związanych z informacją niedokładną i rozważaną klasą funkcji.

Niech $m \geq 2r + 1$, $h + \frac{T}{m}$ oraz $t_i = ih$ dla każdego i . Przez d_i oznaczmy różnicę dzieloną stopnia $r + 1$ bazującą na wartościach $f(t_i)$:

$$d_i = f[t_i, \dots, t_{i+r+1}] = \sum_{j=1}^{i+r+1} \gamma_j \prod_{k=1 \wedge k \neq j}^{i+r+1} (t_k - t_j)^{-1}$$

Następnie oznaczmy przez \tilde{d}_i (niedokładną) różnicę dzieloną stopnia $r + 1$ bazującą na wartościach $y_j = F(t_j) + e_j$, gdzie $|e_j| \leq \delta$

$$\tilde{d}_i = \tilde{f}[t_i, \dots, t_{i+r+1}] = \sum_{j=1}^{i+r+1} \gamma_j \prod_{\substack{k=1 \\ k \neq j}}^{i+r+1} (t_k - t_j)^{-1}$$

Lemat 4.10. *Jeżeli $f \in H_{r,\varrho}(t_i, t_{i+r+1})$, wtedy*

$$|\tilde{d}_i| \leq \frac{c(g_f)(r+1)^\varrho}{(r+1)!} h^{\varrho-1} + \delta \frac{2^{r+1}}{(r+1)!} h^{-(r+1)}$$

Dowód. Korzystając z nierówności trójkąta $|\tilde{d}_i| \leq |d_i| + |\tilde{d}_i - d_i|$ możemy oszacować pierwszy człon:

$$\begin{aligned} |d_i| &= \frac{|f[x_{i+1}, \dots, x_{i+r+1}] - f[x_i, \dots, x_{i+r}]|}{x_{i+r+1} - x_i} = \frac{1}{r!} \frac{|f^{(r)}(\xi_1) - f^{(r)}(\xi_2)|}{x_{i+r+1} - x_i} \\ &\leq \frac{c(g_f)}{r!} \frac{|\xi_1 - \xi_2|^\varrho}{x_{i+r+1} - x_i} \leq \frac{c(g_f)}{r!} (x_{i+r+1} - x_i)^{\varrho-1} \leq \frac{c(g_f)(r+1)^\varrho}{(r+1)!} h^{\varrho-1} \end{aligned} \quad (4.1)$$

oraz drugi człon:

$$\begin{aligned} |\tilde{d}_i - d_i| &= h^{-(r+1)} \left| \sum_{i=0}^{r+1} e_i \prod_{\substack{\ell=0 \\ \ell \neq i}}^{r+1} (\ell - j)^{-1} \right| \\ &\leq \delta h^{-(r+1)} \sum_{i=0}^{r+1} \prod_{\substack{\ell=0 \\ \ell \neq i}}^{r+1} |\ell - j|^{-1} = \delta \frac{2^{r+1}}{(r+1)!} h^{-(r+1)} \end{aligned} \quad (4.2)$$

co dowodzi lemta. □

Teraz oszacujemy błąd interpolacji i ekstrapolacji w obecności zaburzenia wartości funkcji. Niech p_i i \tilde{p}_i odpowiadają wielomianom stopnia co najwyżej r interpolujących f opartych na dokładnych i niedokładnych wartościach funkcji f w punktach $t_i, t_{i+1}, \dots, t_{i+r}$. Dla $r \geq 1$, wprowadźmy oznaczenia:

$$\beta_r = \max_{0 \leq t \leq r} \left| \prod_{k=0}^r (t - k) \right|, \quad \Lambda_r = \max_{0 \leq t \leq r} \sum_{k=0}^r \prod_{\substack{l=0 \\ l \neq k}}^r \left| \frac{t - l}{k - l} \right|, \quad \tilde{\Lambda}_r = \sum_{k=0}^r \prod_{\substack{l=0 \\ l \neq k}}^r \left| \frac{2r+1-l}{k-l} \right|$$

Lemat 4.11. Niech $f \in H_{0,\varrho}$, wtedy:

dla $x \in [t_{i-\frac{1}{2}}, t_{i+\frac{1}{2}}]$:

$$|f(x) - \tilde{p}_1(x)| \leq C_{0,\varrho}(f)h^\varrho + \delta, \quad C_{0,\varrho}(f) = c(g_f)2^{-\varrho}$$

dla $x \in [t_{i-1}, t_{i-\frac{1}{2}}) \cup (t_{i+\frac{1}{2}}, t_{i+1}]$:

$$|f(x) - \tilde{p}_i(x)| \leq \overline{C}_{0,\varrho}(f)h^\varrho + \delta, \quad \overline{C}_{0,\varrho}(f) = c(g_f)$$

Niech $f \in H_{r,\varrho}$ i $r \geq 1$, wtedy:

dla $x \in [t_i, t_{i+r}]$:

$$|f(x) - \tilde{p}_i(x)| \leq C_{r,\varrho}(f)h^{r+\varrho} + \delta\Lambda_r, \quad C_{r,\varrho}(f) = c(g_f)2^{-\varrho}$$

dla $x \in [t_{i-r-1}, t_i) \cup (t_{i+r}, t_{i+2r+1}]$:

$$|f(x) - \tilde{p}_i(x)| \leq \overline{C}_{r,\varrho}(f)h^{r+\varrho} + \delta\overline{\Lambda}_r, \quad \overline{C}_{r,\varrho}(f) = c(g_f)\frac{(2r+1)!(2r+1)^\varrho}{r(r!)^2}$$

Dowód. Przypadek dla $r = 0$ jest oczywisty. Niech $r \geq 1$, korzystając z nierówności trójkąta:

$$|f(x) - \tilde{p}_i(x)| \leq |f(x) - p_i(x)| + |\tilde{p}_i(x) - p_i(x)|$$

Jeżeli $x \in [t_i, t_{i+r}]$, wtedy dla pierwszego członu powyższej sumy mamy:

$$\begin{aligned} |f(x) - p_i(x)| &= |(x - t_i) \cdots (x - t_{i+r}) f[t_i, \dots, t_{i+r}, x]| \\ &\leq \beta_r h^{r+1} \frac{|f[t_{i+1}, \dots, t_{i+r}, x] - f[t_i, \dots, t_{i+r-1}, x]|}{t_{i+r} - t_i} \\ &\leq \beta_r h^{r+1} \frac{c(g_f)}{r^{1-\varrho} r!} h^{\varrho-1} = C_{r,\varrho}(f)h^{r+\varrho} \end{aligned}$$

a dla drugiego członu mamy:

$$|\tilde{p}_i(x) - p_i(x)| = \left| \sum_{k=i}^{i+r} e_k \prod_{\substack{\ell=i \\ \ell \neq k}}^{i+r} \frac{x - t_\ell}{t_k - t_\ell} \right| \leq \delta\Lambda_r \quad (4.3)$$

Przypadek dla $x \in [t_{i-r-1}, t_i) \cup (t_{i+r}, t_{i+2r+1}]$ jest analogiczny. \square

Lemat 4.12. Niech $f \in F_{r,\varrho}$ oraz

$$s_f \in \begin{cases} (t_{i-\frac{1}{2}}, t_{i+\frac{1}{2}}], & \text{gdy } r = 0 \\ (t_i, t_{i-r}], & \text{gdy } r \geq 0 \end{cases}$$

Przypuśćmy, że

$$|\tilde{d}_k| \leq B h^{\varrho-1} \forall k. \quad (4.4)$$

Wtedy dla każdego $x \in [t_{i-1}, t_{i+1}]$, gdy $r = 0$ lub dla każdego $x \in [t_{i-r-1}, t_{i+2r+1}]$, gdy $r \geq 1$, mamy:

$$|f(x) - \tilde{p}_i(x)| \leq D_r(B, f)h^{r+\varrho} + \delta\Lambda_r,$$

gdzie $D_0(B, f) = c(g_f) + B$ i

$$D_r(B, f) = c(g_f) \frac{\beta_r(r+1)^\varrho}{r r!} + B(2^{r+1} - 1) \frac{(2r)!}{(r-1)!} \quad \text{for } r \geq 1.$$

Dowód. Przypadki, gdy $s_f \leq x$ i $s_f > x$ są analogiczne. Weźmy $s_f \leq x$. Jeżeli $r = 0$, wtedy

$$\begin{aligned} |f(x) - \tilde{p}_i| &\leq |f(x) - p_{i+1}| + |p_{i+1} - \tilde{p}_{i+1}| + |\tilde{p}_{i+1} - \tilde{p}_i| \\ &\leq c(g_f) h^e + \delta + B h^e = (c(g_f) + B) h^e + \delta \end{aligned}$$

Pokazaliśmy pierwszą część lematu. Załóżmy, że $r \geq 1$ i $s_f \leq x < t_{i+r}$. Wybierzmy najmniejszy indeks j taki, że $s_f \leq t_j$. Oczywiście $i+1 \leq j \leq i+r$ oraz $x \in [t_{j-1}, t_{j+r}]$. Otrzymujemy

$$|f(x) - \tilde{p}_i(x)| \leq |f(x) - p_j(x)| + |p_j(x) - \tilde{p}_j(x)| + |\tilde{p}_j(x) - \tilde{p}_i(x)| \quad (4.5)$$

A ponieważ $s_f \notin (t_j, t_{j+r}]$, to

$$|f(x) - p_j(x)| \leq c(g_f) \beta_r h^{r+1} \frac{1}{r!} \frac{(t_{j+r} - t_{j-1})^e}{t_{j+r} - t_j} = C_{r,e}(f) \left(1 + \frac{1}{r}\right)^e h^{r+e}$$

Tak jak w równaniu 4.3, mamy

$$|p_j(x) - \tilde{p}_j(x)| \leq \delta \Lambda_r$$

Możemy teraz oszacować pozostały człon $|\tilde{p}_j(x) - \tilde{p}_i(x)|$. Dla $i+r+1 \leq k \leq j+r$, mamy

$$\left(\tilde{f} - \tilde{p}_i\right)[t_i, \dots, t_{i+r}, t_k] = \frac{y_k - \tilde{p}_i(t_k)}{(k-i)(k-i-1) \dots (k-i-r) h^{r+1}} \quad (4.6)$$

oraz

$$\left|\left(\tilde{f} - \tilde{p}_i\right)[t_i, \dots, t_{i+r}, t_k]\right| = \left|\tilde{f}[t_i, \dots, t_{i+r}, t_k]\right| \leq \max_{i \leq \ell \leq k-r-1} |\tilde{d}_\ell| \leq B h^{Q-1} \quad (4.7)$$

gdzie pierwsza nierówność wynika z [3](Lemat 1), natomiast druga z 4.4. Biorąc 4.6 oraz 4.7, otrzymujemy:

$$|y_k - \tilde{p}_i(t_k)| \leq \frac{(2r)!}{(r-1)!} B h^{r+e} \quad (4.8)$$

Także ostatni człon nierówności 4.5 możemy oszacować następująco

$$\begin{aligned} |\tilde{p}_j(x) - \tilde{p}_i(x)| &= \left| \sum_{k=j}^{j+r} (\tilde{p}_j(t_k) - \tilde{p}_i(t_k)) \prod_{\substack{\ell=j \\ \ell \neq k}}^{j+r} \frac{x - t_\ell}{t_k - t_\ell} \right| \\ &\leq \left(\max_{j \leq k \leq j+r} |y_k - \tilde{p}_i(t_k)| \right) \left(\max_{0 \leq t \leq r+1} \sum_{k=0}^r \prod_{\substack{\ell=0 \\ \ell \neq k}}^r \left| \frac{t - \ell}{k - \ell} \right| \right) \end{aligned}$$

Pierwsze maksimum z powyższego równania jest oszacowane poprzez 4.8. Natomiast drugie maksimum jest osiągnięte dla $t = r+1$ i jest równe

$$\sum_{k=0}^r \prod_{\substack{\ell=0 \\ \ell \neq k}}^r \left| \frac{r+1-\ell}{k-\ell} \right| = \sum_{k=0}^r \binom{r+1}{k} = 2^{r+1} - 1$$

Stąd

$$|\tilde{p}_j(x) - \tilde{p}_i(x)| \leq \frac{(2r)!}{(r-1)!} (2^{r+1} - 1) B h^{r+e}$$

□

Przeprowadzimy teraz punktową analizę błędu, tzn. analizę wartości $|f(x) - \varphi_h^*(x)|$ dla każdego x . Rozważmy kilka przypadków w zależności od lokalizacji punktu osobliwego s_f . Przypomnijmy, że

$$(u_3, v_3] \subseteq (u_2, v_2] \subseteq (u_1, v_1],$$

gdzie u_i, v_i dla $i = 1, 2, 3$ są punktami zlokalizowanymi w krokach 1-3 algorytmu. Dla przypadku I przyjmijmy, że

$$\delta \leq b h^{r+e}, \quad \text{dla pewnej stałej } b > 0 \quad (4.9)$$

Uzasadnienie tego wyjaśnimy później [ref].

Przypadek I: $s_f \notin (u_1, v_1]$. Przy takich założeniach, z lematu 4.10 oraz z 4.9 mamy, że

$$|\tilde{d}_i| \leq B_r(b, f) h^{e-1} \quad z B_r(b, f) = \frac{c(g_f)(r+1)^e + b 2^{r+1}}{(r+1)!}$$

Przypuśćmy, że $\varphi_h^*(y)(x) = \tilde{p}_i(x)$ dla pewnego i . Jeżeli $x \notin [u_1, v_1]$ (albo $x \notin [t_{i-1/2}, t_{i+1/2})$ dla $r = 0$) oraz $[t_i, t_{i+r}] \cap (u_1, v_1) = \emptyset$. Wtedy z lematu 4.11 i 4.12 wiemy, że błąd aproksymacji w punkcie x jest ograniczony z góry przez $C_{r,\varrho}(f)h^{r+e} + \delta\Lambda_r$ dla pewnego $s_d \notin (t_i, t_{i+r}]$ (albo $s_f \notin [x_{i-1/2}, x_{i+1/2})$ dla $r = 0$) lub przez $E_r(f)h^{r+e} + \delta\Lambda_r$ w przeciwnym przypadku, gdzie $E_r(f) = D_r(B_r(b, f), f)$. Z drugiej strony, jeżeli $x \in [u_1, v_1]$, wtedy znów z 4.12 wiem, że błąd jest ograniczony przez $\bar{C}_{r,\varrho}h^{r+e} + \delta\bar{\Lambda}_r$ dla $s_f \notin (t_i, t_{i+r}]$ (albo $s_f \notin [t_{i-1/2}, t_{i+1/2}]$ dla $r = 0$), lub przez $E_r(f)h^{r+e} + \delta\Lambda_r$ w przeciwnym przypadku.

Przypadek II: $s_f \in (u_1, v_1]$. Bez straty ogólności założmy, że

$$u_1 < s_f \leq \xi \leq v_1$$

Z 4.11 dla $x \notin [u_1, v_1]$ wiemy, że error jest ograniczony z góry przez $C_{r,\varrho}(f)h^{r+e} + \delta\Lambda_r$ a dla $x \in [u_1, s_f) \cap [\xi, v_1)$ przez $\bar{C}_{r,\varrho}h^{r+e} + \delta\bar{\Lambda}_r$. Z tego powodu możemy założyć, że $x \in [s_f, \xi)$, co jest najbardziej interesujący przypadek. Mamy trzy możliwości:

Przypadek IIa: $s_f \in (u_1, v_1] \setminus (u_2, v_2]$. Z przyjętych założeń wynika, że w kroku 2. algorytmu, w pewnej iteracji musi zachodzić

$$\left| \tilde{f}(z_{j*}) - \tilde{p}_-(z_{j*}) \right| \leq \left| \tilde{f}(z_{j*}) - \tilde{p}_+(z_{j*}) \right| \quad \text{and} \quad s_f \in (u, z_{j*}]$$

co powoduje, że dla wszystkich $1 \leq j \leq r+1$ zachodzi

$$\begin{aligned} |\tilde{p}_+(z_j) - \tilde{p}_-(z_j)| &\leq |\tilde{p}_+(z_{j*}) - \tilde{p}_-(z_{j*})| \leq \left| \tilde{f}(z_{j*}) - \tilde{p}_-(z_{j*}) \right| + \left| \tilde{f}(z_{j*}) - \tilde{p}_+(z_{j*}) \right| \\ &\leq 2 \left| \tilde{f}(z_{j*}) - \tilde{p}_+(z_{j*}) \right| \leq 2 (C_{r,\varrho}(f)h^{r+e} + \delta\Lambda_r) \end{aligned}$$

Z faktu, że jeżeli wielomian p stopnia co najwyżej r jest $|p(z_j)| \leq a, 1 \leq j \leq r+1$, wtedy dla wszystkich $u_1 \leq x < v_1$ is $|p(x)| \leq (2^{r+1} - 1)a$ [zob]. Otrzymaliśmy

$$\begin{aligned} |f(x) - \varphi_h^*(\mathbf{y})(x)| &= |f(x) - \tilde{p}_-(x)| \leq |f(x) - \tilde{p}_+(x)| + |\tilde{p}_+(x) - \tilde{p}_-(x)| \\ &= 2^{r+1} (\bar{C}_{r,\varrho}(f)h^{r+\varrho} + \delta\bar{\Lambda}_r) \end{aligned}$$

Przypadek IIb: $s_f \in (u_2, v_2] \setminus (u_3, v_3]$ Z przyjętych założeń wynika, że w kroku 3. algorytmu, w pewnej iteracji musi zachodzić

$$\left| \tilde{f}(z) - \tilde{p}_-(z) \right| \leq \left| \tilde{f}(z) - \tilde{p}_+(z) \right| \quad \text{and} \quad s_f \in (u, z]$$

a to wraz z [ref] wskazują na to, że

$$|\tilde{p}_+(x) - \tilde{p}_-(x)| \leq \max(|\tilde{p}_+(z) - \tilde{p}_-(z)|, |\tilde{p}_+(s_f) - \tilde{p}_-(s_f)|)$$

a ponieważ w przeciwnym przypadku z nie byłoby największym lokalnym maksimum wyrażenia $|\tilde{p}_+ - \tilde{p}_-|$ na przedziale interval (u, v) . Następnie otrzymujemy

$$\begin{aligned} |\tilde{p}_+(s_f) - \tilde{p}_-(s_f)| &\leq |f(s_f^+) - \tilde{p}_+(s_f)| + |f(s_f^-) - \tilde{p}_-(s_f)| + |f(s_f^+) - f(s_f^-)| \\ &\leq 2(\bar{C}_{r,\varrho}(f)h^{r+\varrho} + \delta\bar{\Lambda}_r) + \left| \Delta_f^{(0)} \right| \end{aligned}$$

a z [ref] mamy

$$|\tilde{p}_+(z) - \tilde{p}_-(z)| \leq 2 \left| \tilde{f}(z) - \tilde{p}_+(z) \right| \leq 2(\bar{C}_{r,\varrho}(f)h^{r+\varrho} + \delta\bar{\Lambda}_r)$$

Końcowo otrzymujemy

$$\begin{aligned} |f(x) - \varphi_h^*(\mathbf{y})(x)| &= |f(x) - \tilde{p}_-(x)| \leq |f(x) - \tilde{p}_+(x)| + |\tilde{p}_+(x) - \tilde{p}_-(x)| \\ &\leq 3(\bar{C}_{r,\varrho}(f)h^{r+\varrho} + \delta\bar{\Lambda}_r) + \left| \Delta_f^{(0)} \right| \end{aligned}$$

Przypadek IIc: $s_f \in (u_3, v_3]$ Ponieważ funkcja $|\tilde{p}_+ - \tilde{p}_-|$ nie ma jakiegokolwiek lokalnego maksimum w przedziale (u_3, v_3) , to wiemy, że jest nierosnąca w przedziale $[s_f, \xi)$. Dlatego znów otrzymujemy

$$|\tilde{p}_+(x) - \tilde{p}_-(x)| \leq |\tilde{p}_+(s_f) - \tilde{p}_-(s_f)| \leq 2(\bar{C}_{r,\varrho}(f)h^{r+\varrho} + \delta\bar{\Lambda}_r) + \left| \Delta_f^{(0)} \right|$$

oraz

$$\begin{aligned} |f(x) - \varphi_h^*(\mathbf{y})(x)| &= |f(x) - \tilde{p}_-(x)| \leq |f(x) - \tilde{p}_+(x)| + |\tilde{p}_+(x) - \tilde{p}_-(x)| \\ &\leq 3(\bar{C}_{r,\varrho}(f)h^{r+\varrho} + \delta\bar{\Lambda}_r) + \left| \Delta_f^{(0)} \right| \end{aligned}$$

Podsumowując analizę błędu dla każdego punktu otrzymujemy, że gdy $\delta \leq bh^{r+\varrho}$ wtedy:

$$\begin{cases} |f(x) - \phi_h^*(y_h)(x)| \propto \max(1, c(g_f))h^{r+\varrho} & \text{dla } x \notin (u_2, v_2] \\ |f(x) - \phi_h^*(y_h)(x)| \propto \max(1, c(g_f))h^{r+\varrho} + \left| \Delta_f^{(0)} \right| & \text{dla } x \in (u_2, v_2] \end{cases}$$

gdzie $v_2 - u_2 \leq \omega$

Mamy również, że liczba ewaluacji funkcji n jest proporcjonalna do h^{-1} , tak więc $h^{r+\varrho}$ jest proporcjonalne do $n^{-(r+\varrho)}$. Z te jobserwacji wynika poniższe stwierdzenie:

Stwierdzenie 4.13. *Niech $1 \leq p \leq \infty$. Jeżeli $\delta \leq bh^{r+\varrho}$ oraz $\omega(h) = h^{(r+\varrho)p+1}$, wtedy*

$$e_p^{wor}(\varphi_h^*, N_h^*; \mathcal{F}_{r,\varrho}^D) = \mathcal{O}(n^{-(r+\varrho)})$$

Przypomnijmy, że powyższe ograniczenia górne nie może zostać spełnione przez algorytmy nieadaptacyjne, co zostało pokazane w [2]. Pokazano tam również, że dla $p = \infty$ nie istnieje algorytm z błędem zbiegającym do zera, dlatego założenia $p < \infty$ jest niezbędne. Dodatkowo, gdy rozważymy klasę $\mathcal{F}_{r,\varrho}^C \subset \mathcal{F}_{r,\varrho}^D$, to możemy uprościć algorytm biorąc $\omega(h) = (r+1)h$ i unikając iteracji w kroku 2. Otrzymujemy w ten sposób algorytm, który dla $r = 0, 1$ jest nieadaptacyjny, a dla $r \geq 2$ używa co najwyżej $r-1$ dodatkowych punktów, niezależnie od tego jak małe jest h . Co więcej, ograniczenie górne zachodzi dla $p = \infty$. Stosując powyższą modyfikację możemy sformułować następujące stwierdzenie.

Stwierdzenie 4.14. *Jeżeli $\delta \leq bh^{r+\varrho}$ i $\omega(h) = (r+1)h$, wtedy:*

$$e_\infty^{wor}(\varphi_h^*, N_h^*; \mathcal{F}_{r,\varrho}^C) = \mathcal{O}(n^{-(r+\varrho)}). \quad (4.10)$$

Ponownie, dla $r \geq 2$ użycie informacji adaptacyjnej jest konieczne. Łącząc wyniki 4.13, 4.14 i 2.2(i) otrzymujemy 2.3. Faktycznie, dla ustalonego δ i n możemy wybrać $h = \frac{T}{m}$ takie, że

$$m = m(n, \delta) = \left\lfloor \min \left(\beta n, \frac{1}{T} \left(\frac{b}{\delta} \right)^{\frac{1}{r+\varrho}} \right) \right\rfloor = \Theta(\min(n, \delta^{-1/(r+\varrho)})), \quad (4.11)$$

Uwaga 4.15. *Zauważmy, że dla ustalonej precyzji δ nie ma sensu brać m większego niż $m_{max} = \Theta(\delta^{-1/(r+\varrho)})$ wartości funkcji, ponieważ dla $m = m_{max}$ osiągamy maksymalną dokładność dla danego δ .*

Rozdział 5

Testy numeryczne

porównanie algorytmów

Bibliografia

- [1] F. Arandiga, A. Cohen, R. Donat, N. Dyn, *Interpolation and approximation of piecewise smooth functions*, SIAM J. Numer. Anal. 43 (2005) 41–57
- [2] L. Plaskota, G. W. Wasilkowski, Y. Zhao, *The power of adaption for approximating functions with singularities*, Mathematics Of Computation 77 2008, p. 2309–2338
- [3] L. Plaskota, G. W. Wasilkowski, *Uniform approximation of piecewise r -smooth and globally continuous functions*, SIAM Journal on Numerical Analysis, Vol. 47, No. 1 (2008/2009)
- [4] B. Kacewicz, P. Przybyłowicz, *Complexity of the derivative-free solution of systems of IVPs with unknown singularity hypersurface*, Journal of Complexity
- [5] P. M. Morkisz, L. Plaskota, *Approximation of piecewise Hölder functions from inexact information*, Journal of Complexity
- [6] J. F. Traub, H. Woźniakowski, G. W. Wasilkowski *Information-Based Complexity*, Academic Press, New York, 1988