Jednostajna słaba podatność zadań wielowymiarowych

(Uniform Weak Tractability of Multivariate Problems)

Autoreferat rozprawy doktorskiej

Paweł Siedlecki

Wydział Matematyki, Informatyki i Mechaniki
Uniwersytet Warszawski

Wstęp

Modelowanie zjawisk świata rzeczywistego, takich jak procesy fizyczne, zachowanie rynków finansowych lub komórek organizmów żywych, jest nieodłącznie związane z numerycznym rozwiązywaniem zadań matematyki ciągłej. Przykładami takich zadań są wielowymiarowe całkowanie, zadanie aproksymacji funkcji czy rozwiązywanie równań różniczkowych. Mając do czynienia z problemem obliczeniowym dobrze jest wiedzieć, jak szybko może być on rozwiązany przy użyciu dostępnych środków, czyli jaka jest jego złożoność oraz który spośród rozwiązujących zagadnienie algorytmów jest optymalny. Komputer jest tworem dyskretnym, jest zatem jasne, że zadania obliczeniowe matematyki ciągłej nie są wiernie reprezentowane w maszynach cyfrowych i na ogół nie mogą być przy ich użyciu bezbłędnie rozwiązane. Każdy algorytm wykonywany na komputerze wykorzystuje jedynie częściową informację o zadaniu i na ogół jest w stanie dać jedynie rozwiązanie przybliżone. *Analityczna złożoność obliczeniowa* (ang. *information-based complexity* - *IBC*) [19] to teoria zajmująca się złożonością zadań matematyki ciągłej, mierzoną minimalną ilością środków potrzebnych do rozwiązania danego zadania z zadaną dokładnością ε.

Wśród zadań matematyki ciągłej szczególnie istotne z punktu widzenia zastosowań są zadania wielowymiarowe. Są to zadania zdefiniowane na przestrzeniach funkcji o "dużej" liczbie d argumentów. Zadania wielowymiarowe są często bardzo trudne do rozwiązania z powodu występowania *przekleństwa wymiaru* (ang. *curse of dimensionality*). Przykładem jest zadanie całkowania wielowymiarowego, gdzie w różnych problemach fizyki, chemii czy matematyki finansowej zachodzi konieczność całkowana funkcji o liczbie argumentów rzędu setek, tysięcy, a nawet milionów. Ze względu na powszechność zadań wielowymiarowych, niezwykle istotne jest zrozumienie, co wpływa na ich złożoność oraz jak konstruować skuteczne algorytmy dające przybliżone rozwiązania w rozsądnym czasie. *Podatność zadań wielowymiarowych* (ang. *tractability of multivariate problems*) [9, 10, 11] jest gałęzią IBC zajmującą sie badaniem relacji pomiędzy złożonością informacyjną a liczbą zmiennych d występujących w zadaniu wielowymiarowym. W szczególności, dziedzina ta bada metody przezwyciężania przekleństwa wymiaru.

Analityczna złożoność obliczeniowa i podatność zadań wielowymiarowych

Analityczna złożoność obliczeniowa (IBC) jest działem matematyki obliczeniowej którego przedmiotem zainteresowania są te zadania obliczeniowe, dla których dostępna informacja o problemie jest niekompletna, zaburzona i/lub kosztowna. IBC rozwijała się stopniowo w ciągu ostatnich dziesięcioleci dostarczając wielu istotnych wyników teoretycznych na temat złożoności zadań matematyki ciągłej, jak również praktycznych algorytmów które okazały się odpowiednie dla wielu zastosowań.

W IBC, zadania obliczeniowe interpretujemy jako odwzorowania pomiędzy odpowiednimi przestrzeniami. Zadanie jest reprezentowane przez *operator rozwiązania*

$$S : F \to G,$$

gdzie F jest przestrzenią unormowaną z normą $\|\cdot\|_F$ (być może wyposażoną w pewną dodatkową strukturę, jak iloczyn skalarny lub miara), a G jest przestrzenią unormowaną z normą $\|\cdot\|_G$. Zadanie polega na znalezieniu przybliżenia $S(f)$ dla każdego $f \in F$. Przybliżenie jest konstruowane w następujący sposób:

- uzyskujemy informację $N(f) \in \mathbb{R}^n$ o elemencie $f \in F$ (jest to skończony ciąg liczb rzeczywistych),

- obliczamy $\phi(N(f)) \in G$, gdzie ϕ jest odwzorowaniem przyporządkowującym skończonemu ciągowi liczb rzeczywistych $N(f)$ element przestrzeni G. Element $\phi(N(f))$ jest przybliżeniem $S(f)$.

Para (N, ϕ) to, z formalnego punktu widzenia, algorytm dający przybliżenie $\phi(N(f))$ elementu $S(f)$ dla każdego $f \in F$. Operator informacji $N : F \to \mathbb{R}^n$ składa się z dozwolonych funkcjonałów informacji, to znaczy, dla każdego $f \in F$ mamy

$$N(f) = [L_1 f, \ldots, L_n f],$$

gdzie L_j jest funkcjonałem należącym do pewnego zbioru $\Lambda \subset F^*$. Funkcja $\phi : \mathbb{R}^n \to G$ może być wybrana dowolnie, chociaż dla wielu operatorów rozwiązania wystarczy rozważać ϕ liniowe.

Rozważmy na przykład zadanie całkowania wielowymiarowego. Wtedy F jest pewną przestrzenią unormowaną funkcji d-zmiennych $f : D \to \mathbb{R}$, gdzie $D \subset \mathbb{R}^d$, oraz

$$S(f) = \int_D f(t)\, dt.$$

Informacja o funkcji f może składać się z obliczeń jej wartości w punktach dziedziny:

$$N(f) = [f(t_1), f(t_2), \ldots, f(t_n)]$$

dla pewnych $t_i \in D$, $1 \leqslant i \leqslant n$. Typowym algorytmem w tym przypadku jest kwadratura, będąca pewną kombinacją liniową składowych wektora $N(f)$.

Istnieją różne sposoby definiowania *błędu* algorytmu (N, ϕ). Wybór konkretnego z nich zależy na ogół od naszych potrzeb, dodatkowych struktur w jakie wyposażone są zbiory F i G, oraz prowadzi do ustalenia *przypadku* (ang. *setting*) w jakim rozważamy nasze zadanie obliczeniowe. Najczęściej badanymi i używanymi w praktyce są *przypadek najgorszy* (ang. *worst case setting*) oraz *przypadek średni* (ang. *average case setting*), gdzie błędy są zdefiniowane następująco:

- *błąd najgorszy*:

$$e^{wor}(N, \phi) = \sup_{\|f\|_F \leqslant 1} \|S(f) - \phi(N(f))\|_G,$$

- *błąd średni*:

$$e^{avg}(N, \phi) = \left(\int_F \|S(f) - \phi(N(f))\|_G^2 \, \mu(df) \right)^{1/2},$$

przy założeniu, że F jest dodatkowo wyposażona w miarę probabilistyczną μ.

Powyżej zostały opisane przypadki deterministyczne. W *przypadku niedeterministycznym (zrandomizowanym)* (ang. *nondeterministic (randomized) setting*) algorytm generujący przybliżenie rozwiązania jest wybierany w sposób losowy. Typowym przykładem takiego algorytmu jest klasyczne *Monte Carlo* dla całkowania wielowymiarowego, gdzie przybliżona wartość całki jest liczona jako średnia z n obliczeń funkcji w losowo wybranych punktach jej dziedziny. Formalnie, *algorytm w przypadku zrandomizowanym* to rodzina $\{(N_\omega, \phi_\omega)\}_{\omega\in\Omega}$ wraz z miarą probabilistyczną ν na Ω. Uśredniając wszystkie możliwe wyniki tej losowej procedury definiujemy

- *błąd w przypadku zrandomizowanym*:

$$e^{ran}(\{(N_\omega, \phi_\omega)\}_{\omega\in\Omega}) = \sup_{f\in F} \Big(\mathbb{E}_\nu \|S(f) - \phi_\omega(N_\omega(f))\|_G^2\Big)^{1/2}.$$

Mając już pojęcie *błędu* algorytmu (N, ϕ), możemy zdefiniować *złożoność informacyjną* zadania S jako funkcję

$$(0,1) \to \mathbb{N} : \varepsilon \mapsto n(\varepsilon),$$

gdzie $n(\varepsilon)$ jest najmniejszą liczbą naturalną taką, że istnieją operator informacji N korzystający z co najwyżej $n(\varepsilon)$ funkcjonałów informacji z klasy Λ oraz funkcja ϕ takie, że błąd algorytmu (N, ϕ) wynosi, w zadanym *przypadku*, co najwyżej ε.

Podatność zadań wielowymiarowych jest działem IBC zajmującym się złożonością informacyjną zadań wielowymiarowych. Jest to stosunkowo nowa dziedzina, została zapoczątkowana pracą H. Woźniakowskiego [21] z roku 1994.

Niech S będzie ciągiem operatorów rozwiązania:

$$S = \{S_d : F_d \to G_d\}_{d\in\mathbb{N}}.$$

W praktyce, F_d jest na ogół podzbiorem pewnej przestrzeni funkcji d zmiennych, a operatory S_d dla $d > 1$ są, w pewnym sensie, powiązane z S_1 i "trudność" ich aproksymacji rośnie wraz z d.

Definiujemy *złożoność informacyjną* zadania wielowymiarowego S jako funkcję

$$(0,1) \times \mathbb{N} \to \mathbb{N} : (\varepsilon, d) \mapsto n(\varepsilon, d),$$

gdzie $n(\varepsilon, d)$ jest równe *złożoności informacyjnej* $n(\varepsilon)$ operatora rozwiązania S_d.

Podatność zadań wielowymiarowych bada zachowanie złożoności informacyjnej nie tylko względem wymaganej dokładności rozwiązania ε, lecz również względem liczby zmiennych d. Z reguły im więcej zmiennych, tym trudniejszy problem, a w wielu praktycznych zadaniach liczba występujących zmiennych jest duża lub bardzo duża (wiele takich zadań pojawiło się w ostatnich 20 latach, na przykład w matematyce finansowej), zatem zależność złożoności informacyjnej problemu nie tylko od zadanej dokładności rozwiązania ε, lecz również od liczby występujących zmiennych d jest kluczowym czynnikiem w zrozumieniu wewnętrznej trudności zadania wielowymiarowego i poszukiwaniu efektywnych algorytmów jego rozwiązania.

Mówimy, że zadanie S jest *niepodatne*, jeśli jego złożoność informacyjna $n(\varepsilon, d)$ jest funkcją wykładniczą ε^{-1} i/lub d. Jeśli złożoność informacyjna zadania S jest funkcją wykładniczą liczby d występujących zmiennych, to mówimy, że zadanie S jest dotknięte *przekleństwem wymiaru* (ang. *curse of dimensionality*); terminologia ta pochodzi od R.E. Bellmana z roku 1957 [1]. Takie zadania nie mają praktycznych rozwiązań, przynajmniej nie dla "dużych" d. Podatne zadania wielowymiarowe są klasyfikowane na podstawie zachowania ich złożoności informacyjnej. Do najbardziej znanych rodzajów podatności należą silnie wielomianowa podatność, wielomianowa podatność, quasi-wielomianowa podatność, i słaba podatność. Dokładniej, zadanie wielowymiarowe jest:

- *silnie wielomianowo podatne* wtw

$$n(\varepsilon, d) \leqslant C\varepsilon^{-p} \qquad \text{dla pewnych } C, p > 0,$$

- *wielomianowo podatne* wtw

$$n(\varepsilon, d) \leqslant C\varepsilon^{-p} d^q \qquad \text{dla pewnych } C, p, q > 0,$$

- *quasi-wielomianowo podatne* wtw

$$n(\varepsilon, d) \leqslant C \exp\Big(t(1 + \ln \varepsilon^{-1})(1 + \ln d)\Big) \qquad \text{dla pewnych } C, t, p, q > 0,$$

- *słabo podatne* wtw
$$\lim_{\varepsilon^{-1}+d\to\infty} \frac{\ln n(\varepsilon,d)}{\varepsilon^{-1}+d} = 0.$$

Wraz z rozwojem dyscypliny stało się jasne, że bardzo ważnym czynnikiem mającym wpływ na podatność zadania wielowymiarowego jest poziom istotności kolejnych zmiennych zadania. Jeśli wszystkie zmienne są jednakowo istotne, to wtedy najczęściej problem okazuje się niepodatny. Z drugiej strony, jeśli rola poszczególnych zmiennych nie jest jednorodna, to podatność problemu może zależeć od tego, jak istotną rolę odgrywają poszczególne zmienne. Inne czynniki zaburzające jednorodność problemu również mogą mieć wpływ na jego podatność. Na przykład, wpływ regularności problemu ze względu na kolejne zmienne na jego podatność był już badany, lecz uzyskano jedynie warunki na (silną) wielomianową podatność, quasi-wielomianową podatność i słabą podatność. W przypadku problemów z wagami przypisanymi zmiennym sytuacja jest analogiczna; znane są jedynie warunki wiążące wagi z (silną) wielomianową podatnością, quasi-wielomianową podatnością i słabą podatnością.

Wyniki rozprawy

W prezentowanej rozprawie doktorskiej wprowadzamy nowy rodzaj podatności zadań wielowymiarowych:

jednostajną słabą podatność (ang. uniform weak tractability).

- Zadanie wielowymiarowe jest *jednostajnie słabo podatne* wtw $n(\varepsilon,d)$ nie jest wykładnicza w żadnej dodatniej potędze ε^{-1} i d, to znaczy dla każdych $\alpha,\beta>0$ zachodzi
$$\lim_{\varepsilon^{-1}+d\to\infty} \frac{\ln n(\varepsilon,d)}{\varepsilon^{-\alpha}+d^{\beta}} = 0.$$

W **Rozdziale 2** rozprawy badamy jednorodne liniowe zadania tensorowe dla klasy Λ^{all} złożonej ze wszystkich ciągłych funkcjonałów liniowych.

W przypadku najgorszym (worst case setting) uzyskujemy twierdzenia charakteryzujące jednostajną słabą podatność zadania $S=\{S_d\}_{d\in\mathbb{N}}$ dla bezwzględnego i znormalizowanego kryterium błędu. Warunek konieczny i dostateczny na jednostają słabą podatność jest podany przy użyciu ciągu $\{\lambda_k\}_{k\in\mathbb{N}}$ wartości własnych operatora $W_1=S_1^*S_1$. Zakładamy przy tym, że:

- jeśli rozważamy kryterium błedu bezwzględnego, to $\lambda_1 = 1 > \lambda_2$ lub $1 > \lambda_1 \geqslant \lambda_1$,
- jeśli rozważamy kryterium błędu znormalizowanego, to $\lambda_1 > \lambda_2 > 0$.

Wtedy dla obu kryteriów błędu zachodzi następujące twierdzenie.

Twierdzenie.

Zadanie S jest jednostajnie słabo podatne $\Leftrightarrow \forall p > 0 \ \lim_{k\to\infty} \frac{\lambda_k}{[\ln k]^{-p}} = 0.$

W przypadku średnim (average case setting) uzyskujemy twierdzenie charakteryzujące jednostajną słabą podatność zadania $S = \{S_d\}_{d\in\mathbb{N}}$ dla bezwzględnego kryterium błędu. Warunek konieczny i dostateczny na jednostajną słabą podatność podany jest przy użyciu ciągu $\{t_k\}_{k\in\mathbb{N}}$, gdzie $t_k = \sum_{j=k+1}^{\infty} \lambda_k$, a $\{\lambda_j\}_{j\in\mathbb{N}}$ jest ciągiem wartości własnych operatora kowariancji C_{ν_1} miary Gaussowskiej ν_1 indukowanej na przestrzeni G_1 przez operator rozwiązania $S_1 : F_1 \to G_1$ oraz miarę gaussowską μ_1 na przestrzeni F_1. Zakładamy, że $\sum_{j=1}^{\infty} \lambda_j < 1$. Zachodzi następujące twierdzenie.

Twierdzenie.

Zadanie S jest jednostajnie słabo podatne $\Leftrightarrow \forall p > 0 \ \lim_{k\to\infty} \frac{t_k}{[\ln k]^{-p}} = 0.$

Ostatnia część **Rozdziału 2** poświęcona jest problemom aproksymacyjnym oraz relacji pomiędzy ich jednostajną słabą podatnością dla klas informacji Λ^{all} oraz Λ^{std}, gdzie Λ^{std} jest klasą złożoną ze wszystkich ciągłych ewaluacji. Pokazujemy, że przy odpowiednich założeniach jednostajna słaba podatność problemu aproksymacyjnego dla klasy Λ^{all} jest równoważna jego jednostajnej słabej podatności dla klasy Λ^{std}. Wyniki tego typu uzyskujemy dla przypadku najgorszego, średniego i zrandomizowanego.

W **Rozdziale 3** rozprawy badamy wielowymiarowe zadania liniowe określone na przestrzeniach funkcji o rosnącej regularności względem kolejnych zmiennych. Dla każdego z tych zadań niemalejący ciąg liczb rzeczywistych

$$r_1 \leqslant r_2 \leqslant r_3 \leqslant \dots$$

zadaje regularność problemu.

W przypadku najgorszym badamy trzy zadania aproksymacyjne dla klasy Λ^{all}. Pierwsze zadanie polega na aproksymacji funkcji z przestrzeni Korobova. Otrzymujemy następujące twierdzenie.

Twierdzenie. *Rozważmy zadanie aproksymacji* APP *określone na przestrzeniach Korobova. Wtedy quasi-wielomianowa podatność jest równoważna jednostajnej słabej podatności oraz słabej podatności i zachodzi wtw* $r_1 > 0$. *Ponadto wykładnikiem quasi-wielomianowej podatności jest* $t^* = r_1^{-1}$.

Kolejne dwa zadania aproksymacyjne określone są na przestrzeniach Soboleva. Rozważamy dwie klasy przestrzeni Soboleva. Odpowiednie przestrzenie obu klas określone są na tych samych zbiorach, a różnią się jedynie wyborem iloczynu skalarnego: dla zadanego parametru regularności $r \in \mathbb{N}$ iloczyn skalarny przestrzeni należących do pierwszej klasy wykorzystuje jedynie zerowe i r-te pochodne, natomiast iloczyn skalarny przestrzeni należących do drugiej klasy wykorzystuje wszystkie pochodne od zerowej do r-tej włącznie.

Dla pierwszej klasy przestrzeni Soboleva otrzymujemy następujące twierdzenie.

Twierdzenie. *Rozważmy zadanie aproksymacji* APP^1 *określone na przestrzeniach Soboleva pierwszego rodzaju. Wtedy quasi-wielomianowa podatność jest równoważna jednostajnej słabej podatności oraz słabej podatności i zachodzi wtw* $r_k = 1$ *dla każdej liczby* $k \in \mathbb{N}$. *Ponadto wykładnikiem quasi-wielomianowej podatności jest* $t_1^* = 1$.

Dla drugiej klasy przestrzeni Soboleva otrzymujemy następujące twierdzenie.

Twierdzenie. *Rozważmy zadanie aproksymacji* APP^2 *określone na przestrzeniach Soboleva drugiego rodzaju. Wtedy quasi-wielomianowa podatność jest równoważna jednostajnej słabej podatności oraz słabej podatności i zachodzi wtw* $r_1 \geqslant 1$. *Ponadto wykładnik quasi-wielomianowej podatności* t_2^* *spełnia* $t_2^* \in \left[\frac{2}{\ln 13}, 1\right]$.

W ostatniej części **Rozdziału 3** badamy dwa zadania aproksymacyjne w przypadku średnim dla klasy Λ^{all}. Oba zadania polegają na aproksymacji funkcji z przestrzeni $C([0,1]^d)$ wyposażonej w odpowiednią miarę gaussowską skupioną na podprzestrzeni tych funkcji ciągłych, które są r_k-krotnie różniczkowalne względem k-tej zmiennej. Dla pierwszego zadania przestrzeń funkcji ciągłych wyposażona jest w miarę gaussowską o średniej zero, której operator kowariancji zadany jest przez jądro kowariancji będące produktem funkcji kowariancji scałkowanego procesu Eulera. Dla drugiego zadania przestrzeń funkcji ciągłych wyposażona jest w miarę gaussowską o średniej zero, której operator kowariancji zadany jest przez jądro kowariancji będące produktem funkcji kowariancji scałkowanego procesu Wienera.

Dla zadania aproksymacji APP zadanego przez scałkowany proces Eulera otrzymujemy następujące twierdzenie.

Twierdzenie. *Rozważmy zadanie aproksymacji* APP *określone przez scałkowany proces Eulera. Zadanie* APP *jest jednostajnie słabo podatne wtw* $\liminf_{k\to\infty} \frac{r_k}{\ln k} \geqslant \frac{1}{2\ln 3}$.

Odpowiedni wynik dla zadania aproksymacji APP zadanego przez scałkowany proces Wienera wygląda następująco.

Twierdzenie. *Rozważmy zadanie aproksymacji* APP *określone przez scałkowany proces Wienera. Zadanie* APP *jest jednostajnie słabo podatne wtw* $\liminf_{k\to\infty} \frac{\ln r_k}{\ln k} \geqslant \frac{1}{2}$.

Literatura

[1] R. E. Bellman, Dynamic programming, *Princeton University Press*, Princeton NJ, 1957

[2] F. Gao, J. Hanning, F. Torcaso, Integrated Brownian motions and exact L_2-small balls, *Ann. Probab.* 31, pp. 1320-1337, 2003.

[3] M. Gnewuch, H. Woźniakowski, Quasi-polynomial tractability, *J. Complexity* 27, pp. 312-330, 2011.

[4] F. J. Hickernell, G. W. Wasilkowski, H. Woźniakowski, Tractability of linear multivariate problems in the average case setting, in: *Monte Carlo and Quasi-Monte Carlo Methods 2006*, A. Keller, S. Heinrich, H. Niederreiter (Eds.), Springer, Berlin 2008, pp. 461-494.

[5] F. J. Hickernell, H. Woźniakowski, Tractability of multivariate integration for periodic functions, *J. Complexity* 17, pp. 660-682, 2001.

[6] M.A. Lifshits, A. Papageorgiou, H. Woźniakowski, Average case tractability of non-homogeneous tensor product problems, *J. Complexity* 28, pp. 539-561, 2012.

[7] M.A. Lifshits, A. Papageorgiou, H. Woźniakowski, Tractability of multi-parametric Euler and Wiener integrated processes, *Prob. and Math. Stat.* 32, pp. 131-165, 2012.

[8] E. Novak and H. Woźniakowski, Intractability results for integration and discrepancy, *J. Complexity* 17, pp. 388-441, 2001.

[9] E. Novak and H. Woźniakowski, Tractability of Multivariate Problems, Volume I: Linear Information, *European Mathematical Society*, Zürich, 2008.

[10] E. Novak and H. Woźniakowski, Tractability of Multivariate Problems, Volume II: Standard Information for Functionals, *European Mathematical Society*, Zürich, 2010.

[11] E. Novak and H. Woźniakowski, Tractability of Multivariate Problems, Volume III: Standard Information for Operators, *European Mathematical Society*, Zürich, 2012.

[12] A. Papageorgiou and I. Petras, On the tractability of linear tensor product problems in the worst case, *J. Complexity* 25, pp. 415-419, 2009.

[13] A. Papageorgiou and I. Petras, Tractability of tensor product problems in the average case setting, *J. Complexity* 27, pp. 273-280, 2011.

[14] A. Papageorgiou, H. Woźniakowski, Tractability through increasing smoothness, *J. Complexity* 26, pp. 409-421, 2010.

[15] S. H. Paskov, J. F. Traub, Faster evaluation of financial derivatives, *J. Portfolio Management*, 22(1), 113-120, 1995.

[16] L. Plaskota, Noisy Information and Computational Complexity, *Cambridge University Press*, Cambridge, 1996.

[17] P. Siedlecki, Uniform Weak Tractability, *J. Complexity* (2013), to appear.

[18] P. Siedlecki, Uniform Weak Tractability of Multivariate Problems with Increasing Smoothness, submitted for publication.

[19] J. F. Traub, G. W. Wasilkowski and H. Woźniakowski, Information-Based Complexity, *Academic Press*, New York, 1988.

[20] A.G. Wershulz, H. Woźniakowski, Tractability of Multivariate Approximation over a Weighted Unanchored Sobolev Spaces, *Constr. Approx.* 30, pp. 395-421, 2009.

[21] H. Woźniakowski, Tractability and strong tractability of multivariate tensor product problems, *J. of Computing and Information* 4, pp. 1-19, 1994.