Logarytmiczne uwypuklenie wielomianów

Abdulljabar Naji Ahmed Abdullah

STRESZCZENIE ROZPRAWY DOKTORSKIEJ

Jednym z podstawowych problemów analizy, techniki, ekonomii i innych gałęzi nauki jest poszukiwanie minimów i punktów krytycznych funkcji. Jedną z metod prowadzących do tego celu jest deformacja danej funkcji do funkcji wypukłej, poszukiwanie punktów krytycznych tej deformacji i iterowanie tego procesu. Sprowadzanie danej funkcji do funkcji wypukłej, czy silnie wypukłej prowadzi do łatwego wyznaczania punktów krytycznych i minimów tej deformacji. Są to dokładnie te punkty, w których gradient się zeruje. Klasycznym podejściem do uwypuklania funkcji $f : \mathbb{R}^n \to \mathbb{R}$ na zbiorach ograniczonych i wypukłych jest dodanie do tej funkcji takiej funkcji silnie wypukłej $b : \mathbb{R}^n \to \mathbb{R}$, że $f + b$ jest funkcją silnie wypukłą na tym zbiorze (patrz na przykład prace W.B. Liu, C.A. Floudas, A.N. Tikhonov, V.Y. Arsenin, S. Zlobec, dla funkcji kwadratowej $b(x) = \gamma|x|^2$, $\gamma > 0$). Omówimy to dokładniej.

Niech $b : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^k$, $k \geq 2$, μ-silnie wypukłą, $\mu > 0$. Niech $X \subset \mathbb{R}^n$ będzie zbiorem zwartym i wypukłym, niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^k$ i niech $D \in \mathbb{R}$ będzie liczbą dodatnią taką, że

$$|\partial^2_\beta f(x)| \leq D \quad \textit{dla } x \in X \text{ i } \beta \in S^{n-1},$$

gdzie S^{n-1} jest sfera jenostkowa w $\mathbb{R}^n$. W klasycznym podejściu do uwypuklania funkcji na zbiorze zwartym, kluczową rolę odgrywa następujący łatwy do sprawdzenia (patrz uwaga 2.2.3 i fakt 3.1.1):

Dla dowolnego $\xi \in \mathbb{R}^n$ oraz $N > D/\mu$, funkcja $\phi_{N,\xi} : \mathbb{R}^n \to \mathbb{R}$ określona wzorem

$$\phi_{N,\xi}(x) = Nb(x - \xi) + f(x), \quad x \in \mathbb{R}^n,$$

jest silnie wypukła na zbiorze X (dokładniej $N\mu - D$-silnie wypukła).

W pracy porównamy powyższe podejście do uwypuklania funkcji, która przyjmuje tylko wartości dodatnie, z innym podejściem polegającym na pomnożeniu jej przez pewną potęgę funkcji silnie wypukłej (patrz rozdział 2 i rozdział 3). To drugie podejście zostało zaproponowane w 2015 w pracy K. Kurdyki i S. Spodziei i kontynuowane w pracy K. Kurdyki, K. Rudnickiej i S. Spodziei. Dokładniej, w pracy K. Kurdyki i S. Spodziei uzyskano uwypuklenie funkcji dodatniej f klasy $\mathscr{C}^2$ na zbiorze zwartym i wypukłym $X \subset \mathbb{R}^n$ przez pomnożenie funkcji f przez $(1 + |x|^2)^N$ dla pewnego N, a w drugiej – przez pomnożenie funkcji f przez $\exp(N|x|^2)$.

W rozdziale 2 uogólnimy te wyniki i pokażemy, że jeśli $X \subset \mathbb{R}^n$ jest zbiorem zwartym i wypukłym oraz $f : X \to \mathbb{R}$ jest funkcją klasy $\mathscr{C}^2$ przyjmującą tylko wartości dodatnie, to dla dowolnej funkcji silnie wypukłej $b : \mathbb{R}^n \to \mathbb{R}$ istnieje $N_0 > 0$ takie, że dla każdego $N \geq N_0$ oraz $\xi \in X$, funkcja

$$\varphi_{N,\xi}(x) = b^N(x - \xi)f(x), \quad x \in \mathbb{R}^n, \tag{1}$$

jest silnie wypukła na zbiorze X.

W przypadku, gdy funkcja f jest wielomianem, wykładnik N można oszacować efektywnie w terminach promienia zbioru X (to znaczy $\sup\{|x| : x \in X\}$) wielkości modułów współczynników wielomianu oraz $m = \inf\{f(x) : x \in X\}$ (wniosek 2.2.2). Wbec tego, w przypadku funkcji dodatnich na zbiorach zwartych i wypukłych, zarówno dodanie do funkcji pewnej wielokrotności funkcji silnie wypukłej jak i pomnożenie jej przez potęgę takiej funkcji dają podobny efekt, jednak pierwsza metoda prowadzi do użycia mniejszego współczynnika N. Jeśli dodatkowo założymy, że funkcja b jest logarytmicznie silnie wypukła (to znaczy $\ln b$ jest funkcją silnie wypukłą), to funkcja $\varphi_{N,\xi}$ również jest logarytmicznie silnie wypukła (patrz wniosek 2.2.9). W przypadku, gdy zbiór X jest semialgebraiczny, zwarty i wypukły, współczynniki wielomianów opisujących zbiór X oraz współczynniki wielomianu f są liczbami całkowitymi (lub wymiernymi), to wykładnik N można wyznaczyć w pełni efektywnie (patrz twierdzenia 2.2.10 i 2.2.12). Twierdzenia te uzyskujemy stosując wynik G. Jeronimo, D. Perrucci, E. Tsigaridas z 2013 roku.

W rozdziale 3 porównamy klasyczne podejście do problemu uwypuklania funkcji z powyższym dla dowolnej funkcji silnie wypukłej $b : \mathbb{R}^n \to \mathbb{R}$ i funkcji f dodatniej na zbiorze domkniętym i wypukłym (niekoniecznie ograniczonym). W tym rozdziale będziemy uwypuklać funkcje f przez pomnożenie jej przez $b(N(x-\xi))$, zamiast przez $b^N(x-\xi)$. Podejście takie upraszcza niektóre obliczenia.

Trudno jest zastosować uwaga 2.2.3 i fakt 3.1.1 w przypadku zbiorów nieograniczonych. Mianowicie mamy

Fakt 3.1.2. Niech $b : \mathbb{R}^n \to (0, +\infty)$ będzie funkcją wypukłą klasy $\mathscr{C}^2$, niech $X \subset \mathbb{R}^n$ będzie zbiorem wypukłym i domkniętym, niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^2$ i niech $N > 0$. Jeśli dla każdego $\xi \in X$, funkcja $\phi_{N,\xi}$ określona wzorem (1) jest wypukła na zbiorze X, to

$$\partial^2_\beta \phi_{N,\xi}(\xi) = N\partial^2_\beta b(0) + \partial^2_\beta f(\xi) \geq 0 \quad \text{dla każdego } \beta \in S^{n-1}.$$

W szczególności pochodne $\partial^2_\beta f$, $\beta \in S^{n-1}$, są wspólnie ograniczone od dołu na zbiorze X. .

W związku z powyższym, uwagę 2.2.3 i fakt 3.1.1 można przenieść do przypadku zbiorów nieograniczonych tylko przy założeniu, że pochodne kierunkowe drugiego rzędu $\partial^2_\beta f$, $\beta \in S^{n-1}$, są ograniczone od dołu na zbiorze X. W ogólnym przypadku zamiast stałej N musimy wybrać funkcję zależną od $|\xi|$. Mianowicie, załóżmy, że funkcja f posiada wzrost wielomianowy drugiego rzędu, to znaczy

$$|\partial^2_\beta f(x)| \leq D(1 + |x|)^\alpha \quad \text{dla } x \in X \text{ i } \beta \in S^{n-1},$$

dla pewnych $D > 0$ i $\alpha \in \mathbb{N}$ oraz $b : \mathbb{R}^n \to \mathbb{R}$ jest funkcją klasy $\mathscr{C}^k$, $k \geq 2$, i logarytmicznie μ-silnie wypukłą, $\mu > 0$, taką że $0 = \operatorname{argmin}_{\mathbb{R}^n} b$ i $b(0) = 1$, gdzie $\operatorname{argmin}_X b$ jest jedynym punktem zbioru X, w którym funkcja b przyjmuje najmniejszą wartość w zbiorze X.

Niech

$$N(|\xi|) = \frac{D}{\mu}\left(|\xi| + 1 + \sqrt{\frac{\alpha}{\mu}}\right)^\alpha + 1 \tag{3.2}$$

Wówczas dla dowolnego $\xi \in \mathbb{R}^n$ funkcja $\phi_\xi : \mathbb{R}^n \to \mathbb{R}$ określona wzorem

$$\phi_\xi(x) = N(|\xi|)b(x-\xi) + f(x), \quad x \in \mathbb{R}^n,$$

jest silnie wypukła na zbiorze X (dokładniej μ-silnie wypukła), patrz lemat 3.2.1.

W szczególności można uzyskać tezę powyższego wniosku dla funkcji $\psi_\xi(x) = Nb(\xi)b(x-\xi) + f(x)$, dla dostatecznie dużej stałej N (patrz lemat 3.2.5).

W przypadku, gdy uwypuklenie funkcji otrzymujemy przez pomnożenie jej przez $x \mapsto b(N(x-\xi))$, gdzie b jest funkcją silnie wypukłą lub logarytmicznie silnie wypukłą, musimy oczywiście założyć, że funkcja przyjmuje tylko wartości dodatnie na zbiorze X. Wówczas

Fakt 3.1.5. Niech $b : \mathbb{R}^n \to (0, +\infty)$ będzie funkcją μ-silnie wypukłą klasy $\mathscr{C}^2$ taką, że $0 = \operatorname{argmin}_{\mathbb{R}^n} b$, niech $X \subset \mathbb{R}^n$ będzie zbiorem wypukłym i domkniętym oraz niech $f : \mathbb{R}^n \to (0, +\infty)$ będzie funkcją klasy $\mathscr{C}^2$. Niech

$$\varphi_{N,\xi}(x) = b(N(x-\xi))f(x), \ N > 0$$

(i) Jeśli dla dowolnego $\xi \in X$ funkcja $\varphi_{N,\xi}$ jest ściśle wypukła na zbiorze X, to $C\partial_\beta^2 f(x) \geq -f(x)$ dla dowolnych $x \in X$ i $\beta \in S^{n-1}$ i pewnej stałej $C > 0$.

(ii) Jeśli fankcja b jest logarytmicznie μ-silnie wypukła i $C\partial_\beta^2 f(x) \geq -f(x)$ oraz $Cf(x) \geq |\partial_\beta f(x)|$ dla dowolnych $x \in X$, $\beta \in S^{n-1}$ i pewnej stałej $C > 0$, to dla $N > \frac{2C}{\sqrt{\mu}} + \frac{1}{C\mu}$, funkcja $\varphi_{N,\xi}$ jest ściśle wypukła na zbiorze X.

Główną trudnością w stosowaniu powyższego faktu jest oszacowanie stałej C. Trudność tę można pokonać w przypadku, gdy uwypuklamy wielomiany. Dokładniej, niech $f \in \mathbb{R}[x]$, gdzie $x = (x_1, \ldots, x_n)$ jest układem zmiennym, będzie wielomienem stopnia d i niech $f = f_0 + \cdots + f_d$, gdzie f_j jest wielomianem jednorodnym stopnia j lub zerem. Niech $f_{d*} = \min_{|x|=1} f_d(x)$. Oczywiście $f_{d*} > 0$ wtedy i tylko wtedy, gdy forma wiodąca f_d wielomianu f przyjmuje tylko wartości dodatnie w $\mathbb{R}^n \setminus \{0\}$. Wówczas możemy uzyskać uwypuklenie wielomianu f przez pomnożenie go przez funkcję $b(N(x-\xi))$, $\xi \in \mathbb{R}^n$. Mianowicie, mamy

Twierdzenie 3.3.1. *Załóżmy, że $f_{d*} > 0$ oraz istnieje $m > 0$ takie, że*

$$f(x) \geq m \quad \textit{dla } x \in \mathbb{R}^m.$$

Wówczas istnieje, efektywnie wyliczalne, N_0 takie, że dla dowolnego $N > N_0$ oraz dla dowolnego $\xi \in \mathbb{R}^n$, funkcja $\varphi_{N,\xi} : \mathbb{R}^n \to \mathbb{R}$ określona wzorem

$$\varphi_{N,\xi}(x) = b(N(x-\xi))f(x)$$

jest μ-silnie wypukła w $\mathbb{R}^n$.

W rozdziałach 4, 5 i 6 zajmujemy się iteracjami odwzorowania, które każdemu punktowi przypisuje jedyny punkt krytyczny uwypuklenia funkcji f.

W rozdziale 4, zakładamy, że $X \subset \mathbb{R}^n$ jest zbiorem wypukłym i zwartym oraz, że funkcja $b : \mathbb{R}^n \to \mathbb{R}$ jest silnie wypukła klasy $\mathscr{C}^k$, $k \geq 2$, taka że $0 = \operatorname{argmin}_{\mathbb{R}^n} b$. Niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^k$. Wówczas istnieje liczba $N \geq 1$, taka, że dla dowolnego $\xi \in \mathbb{R}^n$, funkcja

$$\phi_{N,\xi}(x) = Nb(x - \xi) + f(x)$$

jest silnie wypukła na zbiorze X. Określamy odwzorowanie

$$\kappa_N : X \ni \xi \mapsto \operatorname{argmin}_X \phi_{N,\xi} \in \mathbb{R}^n,$$

Jeśli

$$X_{f \leq r} := \{x \in \mathbb{R}^n : f(x) \leq r\} \subset X,$$

to pokazujemy, że (patrz lematy 4.1.3 i 4.1.4)

(i) Odwzorowanie κ_N jest dyfeomorfizmem klasy $\mathscr{C}^{k-1}$ ze zbioru $X_{f \leq r}$ na zbiór $Y = \kappa_N(X_{f \leq r}) \subset X_{f \leq r}$.

(ii) Zbiorem punktów stałych odwzorowania $\kappa_N|_{X_{f \leq r}}$ jest $\Sigma_f \cap X_{f \leq r}$, gdzie Σ_f jest zbiorem punktów krytycznych funkcji f.

W twierdzeniu 4.2.1 (c) pokazujemy, że : Jeśli funkcja f jest semialgebraiczna, to dla dowolnego $\xi \in X_{f \leq r}$, punkt graniczny $\lim_{\nu \to \infty} \kappa_N^\nu(\xi)$ istnieje i należy do zbioru Σ_f.

Dowód tego twierdzenia opiera sie na pokazaniu monotoniczności ciągu $f(\xi_\nu)$ (patrz wniosek 4.1.5) i zastosowaniu zasady porównania z pracy K. Kurdyki i S. Spodziei do pokazania, że szereg $\sum_{\nu=0}^{\infty} \operatorname{dist}(\kappa_N^\nu(\xi), f^{-1}(f(\kappa_N^{\nu+1}(\xi))))$ jest zbieżny. Idea tego dowodu jest wzorowana na dowodzie (Theorem 7.5) z pracy K. Kurdyki i S. Spodziei. Dowód tego twierdzenia nie jest bezpośrednim przeniesieniem dowodu (Theorem 7.5) z pracy K. Kurdyki i S. Spodziei, gdyż w tej pracy rozważano uwypuklanie funkcji f przez pomnożenie jej przez $(1 + |x|^2)^N$, a my rozważamy uwypuklanie tej funkcji przez dodanie do niej Nb.

Przy założeniu, że funkcja f jest semialgebraiczna, powyższe twierdzenie pozwala określić odwzorowanie

$$\kappa_* : X_{f \leq r} \to \Sigma_f \cap X_{f \leq r}\,,$$

wzorem $\kappa_{N,*}(\xi) = \lim_{\nu \to \infty} \kappa_N^\nu(\xi)$.

Przy założeniu, że funkcja f ma tylko jedną wartość krytyczną na zbiorze $X_{f \leq r}$, pokażemy, że odwzorowanie $\kappa_{N,*}$ jest ciągłe. Mianowicie, mamy

Twierdzenie 4.3.1. *Niech $0 \in \operatorname{Int} X_{f \leq r}$ oraz niech $f(0)$ będzie najmniejszą wartością funkcji f. Wtedy istnieje $f(0) < \delta < r$ taki, że ciąg κ_N^ν jest zbieżny jednostajnie do $\kappa_{N,*}$ w zbiorze $U = X_{f \leq \delta}$. W szczególności odwzorowanie*

$$\kappa_{N,*}|_U : U \to U \cap \Sigma_f$$

jest ciągłe oraz $\kappa_{N,*}(\xi) = \xi$ wtedy i tylko wtedy, gdy $\xi \in U \cap \Sigma_f$. W konsekwencji $\kappa_{N,*}|_U$ jest retrakcją, a zbiór $U \cap \Sigma_f$ jest retraktem zbioru U.

W rozdziale 5 przenosimy pewne własności odwzorowania κ_N na przypadek zbiorów nieograniczonych. Między innymi, w przypadku, gdy uwypuklenie funkcji f jest postaci

$$\psi_\xi(x) = N(\xi)b(x-\xi) + f(x), \qquad (\xi, x) \in \mathbb{R}^n \times \mathbb{R}^n$$

oraz dla zbioru wypukłego i domkniętego X,

$$\kappa(\xi) = \operatorname{argmin}_X \psi_\xi \in X,$$

pokazujemy: Dla dowolnego $\xi \in X$ punkt $\kappa(\xi)$ jest jedynym dolnym punktem krytycznym funkcji ψ_ξ na zbiorze X. Podobny fakt zachodzi dla funkcji

$$\kappa_N : X \ni \xi \mapsto \operatorname{argmin}_X b(N(x-\xi))f(x)$$

.

W rozdziale 6 przenosimy wyniki z rozdziału czwartego dla odwzorowania $X_{f\le r} \ni \xi \mapsto \operatorname{argmin}_{|x|\le R} b^N(x-\xi)f(x) \in X$, przy założeniu, że zbiór $X_{f\le r}$ jest zwarty i wypukły. W tym przypadku zachodzą wszystkie powyżej przedstawionme włsaności odwzorowania κ_N. Jeśli dodatkowo założymy, że $b(x) = \exp(|x|^2)$ oraz f jest wielomianem, to odwzorowanie κ_N ma pewne dodatkowe włsności, między innymi jest to odwzorowanie analityczne i semialgebraiczne, czyli jest odwzorowaniem Nasha. Mianowicie, odwzorowania $\kappa_N : X_{f,r} \to \kappa_N(X_{f,r})$ jest odwzorowaniem odwrotnym do odwzorowania

$$\kappa_N(X_{f\le r}) \ni x \mapsto x + \frac{1}{2Nf(x)}\nabla f(x) \in X_{f\le r}\,,$$

jest więc analityczne i semialgebraiczne, t.j., jest odwzorowaniem Nasha.

Na końcu rozdziału szóstego zajmujemy się problemem zbieżności ciągu $\frac{\xi_\nu}{|\xi_\nu|}$, gdzie $\xi_\nu = \kappa_N^\nu(\xi)$, $\nu \in \mathbb{N}$, oraz $\xi \in \mathbb{R}^n$, czyli problem zbieżności ciągu części sferycznych ciągu ξ_ν. Jest to przeniesienie problemu René Thoma dla trajektorii pola gradientowego (rozwiązanego w pracy K. Kurdyka, T. Mostowski, A, Parusiński) do przypadku dyskretnego. Rozważamy ten problem przy założeniu, że $\xi_\nu \to 0$ gdy $\nu \to \infty$ oraz kilku dodatkowych dość restrykcyjnych założeniach.

Część wyników tej pracy zostało już opublikowane w pracy A.N. Abdullah, K. Rosiak, S. Spodzieja. Dotyczy to punktu 2.2 i rozdziału 6.