Aproksymacyjne własności projekcji

Niech $f \in C([a,b])$ oraz $\|\cdot\| = \|\cdot\|_{C([a,b])}$. Niech dalej

$$V_0 \subset V_1 \subset V_2 \subset \cdots$$

będzie ustalonym, nieskończonym ciągiem podprzestrzeni liniowych, $\dim(V_n) = n$. Dalej będziemy zakładać, że $V_n = \Pi_{n-1}$ jest podprzestrzenią wielomianów algebraicznych stopnia co najwyżej $n-1$. Jak zwykle,

$$\mathrm{dist}(f, V_n) \;=\; \min_{v \in V_n} \|f - v\|$$

i przez v_f^n oznaczymy element optymalny dla f w V_n. Z punktu widzenia praktycznej aproksymacji funkcji f chcielibyśmy mieć algorytm $A_n : C([a,b]) \to V_n$, który z jednej strony produkuje przybliżenie $A_n f$ z małym błędem $\|f - A_n f\|$, a z drugiej jest nie tylko realizowalny, ale też stosunkowo tani obliczeniowo. *Realizowalność* oznacza tu, że aproksymacja konstruowana jest na podstawie informacji o wartościach funkcji w skończonej liczbie punktów, tzn. A_n jest postaci

$$A_n f \;=\; \Phi\Big(f(x_1), f(x_2), \ldots, f(x_n)\Big),$$

gdzie $\Phi : \mathbb{R}^n \to C([a,b])$ jest pewnym przekształceniem. (To implikuje, w szczególności, że funkcjom o tych samych wartościach w x_i dla $0 \leqslant i \leqslant n$ odpowiada ta sama aproksymacja.) Natomiast *taniość* oznacza koszt proporcjonalny do n. Ten postulat spełniają algorytmy liniowe,

$$A_n f \;=\; \sum_{i=1}^{n} f(x_i)\phi_i,$$

gdzie $\phi_i \in C([a,b])$, $0 \leqslant i \leqslant n$, są pewnymi ustalonymi funkcjami. Przykładami są aproksymacja wielomianami Bernsteina,

$$(B_n f)(x) \;=\; (b-a)^{1-n} \sum_{i=1}^{n} f\left(a + \frac{i-1}{n-1}(b-a)\right) \binom{n-1}{i} (x-a)^{i-1}(b-x)^{n-i},$$

albo po prostu interpolacja Lagrange'a z węzłami $a \leqslant x_1 < x_2 < \ldots < x_n \leqslant b$,

$$(P_n f)(x) \;=\; \sum_{i=1}^{n} f(x_i)\, l_i(x), \qquad l_i(x) = \prod_{j=1, j\neq i}^{n} \frac{x - x_j}{x_i - x_j}.$$

Natomiast optymalna aproksymacja $f \mapsto v_f^n$ nie wchodzi w rachubę, bo v_f^n zależy nieliniowo od f, a poza tym znalezienie v_f^n jest trudne obliczeniowo i wymaga w szczególności znalezienia *alternansu.*

Dodatkowo chcielibyśmy, aby algorytm A_n "odtwarzał" funkcje z V_n, tzn.

$$A_n f = f \qquad \text{o ile} \quad f \in V_n.$$

Ten warunek spełnia interpolacja Lagrange'a, ale nie wielomiany Bernsteina. Można bowiem zauważyć, że już dla wielomianu $w_2(x) = x^2$ i (dla uproszczenia) $[a,b] = [0,1]$ mamy

$$(B_n w_2)(x) = \frac{n-1}{n} x^2 + \frac{1}{n} x,$$

skąd $w_2(x) - (B_n w_2)(x) = x(1-x)/n$. Już ten fakt niejako dyskredytuje ten typ aproksymacji. (Porównaj również z twierdzeniem 1 poniżej.)

Teraz warto zadać sobie pytanie, czy istnieją algorytmy A_n spełniające powyższe warunki, dla których błąd $\|f - A_n f\|$ zbiega do zera jak błąd optymalny $\mathrm{dist}(f, V_n)$, gdy $n \to \infty$. W szczególności, chcielibyśmy "odtworzyć" następujące twierdzenie, który jest jednym z wariantów słynnego w teorii aproksymacji *twierdzenia Jacksona* (bez dowodu).

Twierdzenie 1 *Jeśli* $f \in C^r([a,b])$ *to* $\mathrm{dist}(f, V_n) \leqslant C n^{-r} \|f^{(r)}\|$ *dla pewnej* $C > 0$.

Aby odpowiedzieć na to pytanie, rozpatrzmy sytuację nieco ogólniejszą. Niech

$$L_n : C([a,b]) \to V_n, \qquad n = 0, 1, 2, \ldots,$$

będzie ciągiem *projekcji*, tzn. L_n jest liniowy i ciągły oraz $L_n f = f$ gdy $f \in V_n$. Oczywiście, P_n jest projekcją, ale B_n już nie, jak pokazaliśmy powyżej na przykładzie funkcji $x \mapsto x^2$. Przypomnijmy jeszcze, że norma projekcji (ogólniej, przekształcenia liniowego) zdefiniowana jest jako

$$\|L_n\| := \sup_{\|f\| \leqslant 1} \|L_n f\|.$$

Łatwo widać, że dla algorytmu A_n mamy

$$\|A_n\| = \sup_{a \leqslant x \leqslant b} \Big| \sum_{i=1}^{n} \phi_i(x) \Big|,$$

skąd, w szczególności, $\|B_n\| = 1$ oraz $\|P_n\| = \max_{a \leqslant x \leqslant b} \sum_{i=1}^{n} |l_i(x)|$.

Dla dowolnej funkcji $f \in C([a,b])$ mamy

$$\begin{aligned} \|f - L_n f\| &\leqslant \|f - v_f^n\| + \|v_f^n - L_n f\| \;=\; \mathrm{dist}(f, V_n) + \|L_n(v_f^n - f)\| \\ &\leqslant (1 + \|L_n\|)\, \mathrm{dist}(f, V_n), \end{aligned}$$

co oznacza, że jeśli normy projekcji są wspólnie ograniczone, $\|L_n\| \leqslant M \;\forall n$, to błąd zbiega jak błąd optymalny. Ale też odwrotnie, jeśli istnieje K taka, że dla wszystkich $f \in C([a,b])$ jest $\|f - L_n f\| \leqslant K \,\mathrm{dist}(f, V_n) \;\forall n$, to

$$\|L_n f\| \leqslant \|f - L_n f\| + \|f\| \leqslant (K+1)\, \|f\|$$

(gdzie wykorzystaliśmy fakt, że $\mathrm{dist}(f, V_n) \leqslant \|f\|$), czyli normy $\|L_n\|$ są wspólnie ograniczone.

Odpowiedź na nasze pytanie sprowadza się teraz do odpowiedzi na inne pytanie: czy istnieją projekcje na $V_n = \Pi_{n-1}$ o wspólnie ograniczonych normach? Niestety, odpowiedź brzmi NIE. Mówi o tym ważne twierdzenie *Łozińskiego-Charsziladze*, które podajemy bez dowodu i w uproszczonej formie.

Twierdzenie 2 *Dla dowolnych projekcji* L_n *mamy* $\lim_{n \to \infty} \|L_n\| = \infty$.

Z naszych rozważań można w szczególności wysnuć wniosek, że nie istnieją projekcje, a więc również algorytmy A_n (w sensie zdefiniowanym powyżej) takie, że dla funkcji $f \in C^r([a,b])$ błąd $\|f - A_n f\|$ zbiega do zera jak n^{-r}. Okazuje się, że problemem jest wybór podprzestrzeni $V_n = \Pi_{n-1}$, czyli aproksymacja wielomianami coraz wyższego stopnia. Sytuacja zmienia się radykalnie gdy zamiast aproksymacji wielomianami rozpatrzymy aproksyamcję kawałkami wielomianami ustalonego stopnia $r-1$. Wtedy dla $f \in C^r([a,b])$ mamy zbieżność n^{-r} - PATRZ ROZDZIAŁ 7.3 na str. 83-85 SKRYPTU. Poza tym, aproksymacja kawałkami wielomianowa okazuje się optymalna w klasie funkcji z $\|f^{(r)}\| \leqslant M$, w sensie, błędu najgorszego - PATRZ U. 7.2 na str. 85 SKRYPTU.