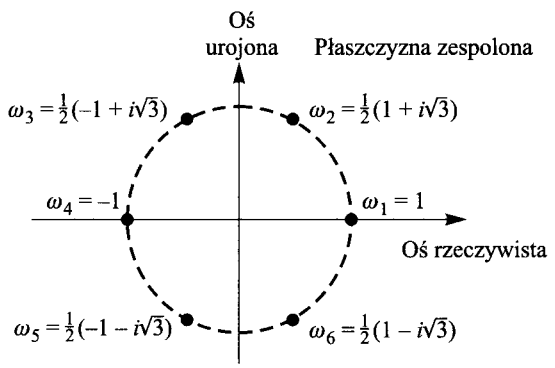
Zbieżność wielomianów interpolacyjnych

Dla funkcji f ciągłej w przedziale $[a, b]$ i dla n naturalnych utwórzmy wielomiany interpolacyjne p_n, każdy z równomiernie rozmieszczonymi węzłami. Można by przypuścić, że ciąg $\{p_n\}$ jest zbieżny *jednostajnie* do f, tzn. że ciąg norm $\|f - p_n\|_{[a,b]}$ dąży do 0, gdy $n \to \infty$. Z przykładu 6.1.4 wynika, że tak rzeczywiście jest dla $f(x) := \sin x$, ale to nie jest typowa funkcja ciągła. Istotnie, należy ona do klasy C^∞ na całej prostej rzeczywistej, a jako funkcja zmiennej zespolonej jest funkcją *całkowitą*, czyli nie mającą żadnych osobliwości na płaszczyźnie zespolonej.



RYS. 6.2. Sześć pierwiastków stopnia szóstego z 1

Zaskakujące jest to, że dla większości funkcji ciągłych powyższy ciąg norm nie dąży do 0. Pierwszy przykład tego typu podał Meray w 1884 r. Dotyczy on funkcji $f(z) = 1/z$ na okręgu jednostkowym $|z| = 1$. Niech, dla ustalonego n, liczby $\omega_1, \omega_2, \ldots, \omega_n$ będą n-tymi pierwiastkami z liczby 1. Są one równomiernie rozmieszczone na tym okręgu, co dla $n = 6$ pokazuje rys. 6.2. Dla tej funkcji f wielomian interpolacyjny p_{n-1} z węzłami ω_j jest równy po prostu z^{n-1}. Istotnie,

$$p_{n-1}(\omega_j) = \omega_j^{n-1} = \frac{1}{\omega_j}\omega_j^n = \frac{1}{\omega_j} = f(\omega_j) \quad (1 \leqslant j \leqslant n).$$

Norma różnicy $f - p_{n-1}$ na kole jednostkowym jest równa

$$\max_{|z|=1}|f(z) - p_{n-1}(z)| = \max_{|z|=1}|z^{-1} - z^{n-1}| = \max_{|z|=1}\frac{1}{|z|}|1 - z^n| = 2,$$

bo taką wartość otrzymuje się wtedy, gdy $z^n = -1$.

Inny przykład, odnoszący się do dziedziny rzeczywistej, podał Runge w 1901 r. Wykazał on, że dla funkcji $f(x) = (x^2 + 1)^{-1}$ rozpatrywanej

w przedziale $[-5,5]$ ciąg wielomianów interpolacyjnych z węzłami rozmieszczonymi tam w równych odstępach nie dąży do f. Nawet dla niewielkich wartości n (np. dla $n = 15$) wielomian p_n „szaleje". Jest to *zjawisko Rungego*, opisywane w wielu książkach i artykułach; zob. np. Epperson [1987]. Jego doświadczalne potwierdzenie jest tematem zad. 6.2.K1.

Dowód tego, że zjawisko Rungego występuje, można znaleźć w książce Steffensena [1950] (I wyd. opublikowane w 1927 r.). Książka Davisa [1982] daje ujęcie tego tematu związane z dziedziną zespoloną. Funkcja Rungego, pozornie „niewinna", ma jednak osobliwość na osi urojonej w pobliżu przedziału $[-5,5]$ i to właśnie jest przyczyną kłopotów. Uzasadnienie tego wykracza jednak poza zakres niniejszej książki. Niżej podano dwa twierdzenia – tw. 6.1.9 pokazuje, że pod pewnymi warunkami ciąg $\{p_n\}$ jest zbieżny do f, a tw. 6.1.8 świadczy o tym, że nie zawsze tak jest. Różnica kryje się w kwantyfikatorach!

W 1914 roku Faber udowodnił następujące bardzo ogólne twierdzenie:

Twierdzenie 6.1.8 (Faber). *Dla dowolnego ciągu układów węzłów*

$$a \leqslant x_0^{(n)} < x_1^{(n)} < \ldots < x_n^{(n)} \leqslant b \quad (n \geqslant 0) \tag{6.1.12}$$

istnieje taka funkcja ciągła w $[a,b]$, że ciąg wielomianów interpolacyjnych zbudowanych dla tych węzłów nie jest do niej zbieżny.

Jest to twierdzenie bardziej subtelne niż mogłoby się wydawać na pierwszy rzut oka, skoro zachodzi też takie twierdzenie o zbieżności:

Twierdzenie 6.1.9. *Jeśli f jest funkcją ciągłą w $[a,b]$, to istnieje taki ciąg układów węzłów* (6.1.12), *że zbudowane dla nich wielomiany interpolacyjne tworzą ciąg zbieżny do f.*

To twierdzenie wynika z zestawienia dwóch bardzo ważnych twierdzeń, odpowiednio – autorstwa Weierstrassa (o aproksymacji) i Czebyszewa (o alternansie). To drugie znajduje się w podrozdz. 6.9, dowód pierwszego (tw. 6.1.11) opiera się na pomocniczym tw. 6.1.10 i własnościach wielomianów, które w 1912 r. zdefiniował i zbadał S. Bernstein.

Poniższe twierdzenie dotyczy operatorów *liniowych* i *dodatnich* przekształcających przestrzeń funkcji ciągłych $C[a,b]$ w nią samą. Operator L ma te własności, jeśli odpowiednio

$$L(\varphi f + \psi g) = \varphi L f + \psi L g \quad (\varphi, \psi \in \mathbb{R},\ f, g \in C[a,b]),$$
$$Lf \geqslant 0, \quad \text{gdy} \quad f \geqslant 0.$$

Okazuje się, że są to własności niezmiernie istotne dla teorii aproksymacji.

Twierdzenie 6.1.10 (Bohman-Korowkin). *Niech $\{L_n\}$ $(n \geqslant 1)$ będzie ciągiem operatorów liniowych dodatnich, przekształcających przestrzeń $C[a,b]$ w nią samą. Jeśli dla $f(x) := 1,\ x,\ x^2$ jest $\{\|L_n f - f\|_{[a,b]}\} \to 0$, to jest tak również dla każdej innej funkcji $f \in C[a,b]$.*

Dowód. Jeśli L jest operatorem liniowym i dodatnim, to zachodzą następujące implikacje:

$$f \geqslant g \Rightarrow f - g \geqslant 0 \Rightarrow L(f-g) \geqslant 0 \Rightarrow Lf - Lg \geqslant 0 \Rightarrow Lf \geqslant Lg.$$

Prócz tego jest $|f| \geqslant f$ i $|f| \geqslant -f$, a zatem $L(|f|) \geqslant Lf$ i $L(|f|) \geqslant -Lf$, czyli $L(|f|) \geqslant |Lf|$.

Niech teraz będzie $h_k(x) = x^k$ dla $k = 0, 1, 2$. Zdefiniujmy też funkcje:

$$\alpha_n = L_n h_0 - h_0, \quad \beta_n = L_n h_1 - h_1, \quad \gamma_n = L_n h_2 - h_2.$$

Z założenia wynika, że

$$\{\|\alpha_n\|\} \to 0, \quad \{\|\beta_n\|\} \to 0, \quad \{\|\gamma_n\|\} \to 0$$

(dla uproszczenia tu i dalej pomijamy w symbolu normy przedział $[a,b]$).

Rozważmy teraz dowolną funkcję f z $C[a,b]$. Jest to więc funkcja jednostajnie ciągła, tzn. dla każdego $\varepsilon > 0$ istnieje takie $\delta > 0$, że jeśli $x, y \in [a,b]$ i $|x-y| < \delta$, to $|f(x) - f(y)| < \varepsilon$. Stąd wynika, że dla $c = 2\|f\|/\delta^2$ jest

$$|x-y| \geqslant \delta \Rightarrow |f(x) - f(y)| \leqslant 2\|f\| \leqslant 2\|f\| \frac{(x-y)^2}{\delta^2} = c(x-y)^2.$$

Dlatego dla wszystkich x i y z $[a,b]$ jest $|f(x) - f(y)| \leqslant \varepsilon + c(x-y)^2$. Używając funkcji $h_k(x)$, wyrażamy to tak:

$$|f - f(y)h_0| \leqslant \varepsilon h_0 + c(h_2 - 2yh_1 + y^2 h_0).$$

Uwzględniając uwagi z początku dowodu, wnioskujemy stąd, że

$$|L_n f - f(y) L_n h_0| \leqslant \varepsilon L_n h_0 + c(L_n h_2 - 2y L_n h_1 + y^2 L_n h_0).$$

Jest to nierówność wiążąca funkcje zmiennej x. Podstawmy w jej miejsce y:

$$\begin{aligned}
&|(L_n f)(y) - f(y)(L_n h_0)(y)| \leqslant \\
&\quad \leqslant \varepsilon (L_n h_0)(y) + c[(L_n h_2)(y) - 2y(L_n h_1)(y) + y^2 (L_n h_0)(y)] = \\
&\quad = \varepsilon[1 + \alpha_n(y)] + c[y^2 + \gamma_n(y) - 2y(y + \beta_n(y)) + y^2(1 + \alpha_n(y))] = \\
&\quad = \varepsilon + \varepsilon\alpha_n(y) + c\gamma_n(y) - 2cy\beta_n(y) + cy^2\alpha_n(y) \leqslant \\
&\quad \leqslant \varepsilon + \varepsilon\|\alpha_n\| + c\|\gamma_n\| + 2c\|h_1\|\,\|\beta_n\| + c\|h_2\|\,\|\alpha_n\|.
\end{aligned}$$

Zgodnie z założeniem istnieje takie m, że dla wszystkich $n \geqslant m$ prawa strona tej nierówności nie przewyższa 2ε. Wtedy

$$\|L_n - fL_nh_0\| \leqslant 2\varepsilon.$$

Kończąc dowód, zauważmy, że

$$\|L_nf - f\| \leqslant \|L_nf - fL_nh_0\| + \|fL_nh_0 - fh_0\| \leqslant 2\varepsilon + \|f\|\,\|\alpha_n\|.$$

W razie potrzeby zwiększamy m tak, żeby dla $n \geqslant m$ było $\|f\|\,\|\alpha_n\| < \varepsilon$. Dla tych n zachodzi więc nierówność

$$\|L_nf - f\| < 3\varepsilon.$$

ε było dowolne, a zatem twierdzenie zostało udowodnione. ■

Zapowiedziane już *wielomiany Bernsteina* są zdefiniowane, dla dowolnej funkcji $f \in C[0,1]$, wzorem

$$(B_nf)(x) := \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}x^k(1-x)^{n-k}.$$

Użyta symbolika podkreśla to, że wielomian Bernsteina jest w istocie operatorem liniowym; przekształca on funkcję ciągłą w $[0,1]$ na wielomian n-tego stopnia. Ponieważ w tym przedziale jest $x^k(1-x)^{n-k} \geqslant 0$, więc jest to operator dodatni. Stosujemy te wielomiany w elementarnym dowodzie[2] twierdzenia Weierstrassa, ale używa się ich także w projektowaniu wspomaganym komputerowo, choć tu wypierają je tzw. funkcje B-sklejane (zob. podrozdz. 6.5 i 6.6).

Twierdzenie 6.1.11 (Weierstrass). *Jeśli funkcja f jest ciągła w przedziale $[a,b]$, to dla każdego $\varepsilon > 0$ istnieje taki wielomian p, że $\|f-p\|_{[a,b]} \leqslant \varepsilon$.*

Dowód. Możemy ograniczyć się do przedziału $[0,1]$, ponieważ przekształcenie liniowe zmiennej $x = a + t(b-a)$ zmienia $[0,1]$ na $[a,b]$, a wielomian p na inny wielomian tego samego stopnia.

Zgodnie z tw. 6.1.10 pozostaje wykazać, że dla $h_k(x) := x^k$ i $k = 0,1,2$ jest $\{B_nh_k\} \to h_k$. Dla $k = 0$ mamy równość

$$(B_nh_0)(x) = \sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k} = [x + (1-x)]^n = 1.$$

[2] Dowód odwołuje się jednak do raczej wyrafinowanego tw. 6.1.10. Pierwotny dowód Bernsteina jest chyba bardziej naturalny; zob. Achiezer [*1957]. Istnieje zresztą wiele innych jeszcze dowodów (*przyp. tłum.*).

Dla $k = 1$ korzystamy z zad. 22:

$$(B_n h_1)(x) = \sum_{k=0}^{n} \frac{k}{n}\binom{n}{k} x^k (1-x)^{n-k} = \sum_{k=1}^{n} \binom{n-1}{k-1} x^k (1-x)^{n-k} =$$
$$= x \sum_{k=0}^{n-1} \binom{n-1}{k} x^k (1-x)^{n-1-k} = x.$$

Dla $k = 2$ stosujemy to zadanie dwukrotnie, zakładając, że $n > 1$:

$$(B_n h_2)(x) = \sum_{k=0}^{n} \left(\frac{k}{n}\right)^2 \binom{n}{k} x^k (1-x)^{n-k} = \sum_{k=1}^{n} \frac{k}{n}\binom{n-1}{k-1} x^k (1-x)^{n-k} =$$
$$= \sum_{k=1}^{n} \left(\frac{n-1}{n}\,\frac{k-1}{n-1} + \frac{1}{n}\right)\binom{n-1}{k-1} x^k (1-x)^{n-k} =$$
$$= \frac{n-1}{n} x^2 \sum_{k=2}^{n} \binom{n-2}{k-2} x^{k-2} (1-x)^{n-k} + \frac{x}{n} =$$
$$= \frac{n-1}{n} x^2 + \frac{x}{n} \to x^2. \qquad \blacksquare$$

ZADANIA 6.1

1. Udowodnić, że jeśli g interpoluje funkcję f w węzłach $x_0, x_1, \ldots, x_{n-1}$, a h interpoluje f w węzłach $x_1, x_2, \ldots, x_n$, to funkcja

$$g(x) + \frac{x_0 - x}{x_n - x_0}[g(x) - h(x)]$$

interpoluje f we wszystkich węzłach $x_0, x_1, \ldots, x_n$ (g i h nie muszą być wielomianami).

2. Udowodnić, że jeśli funkcja g (wielomian lub nie) interpoluje funkcję f w węzłach $x_0, x_1, \ldots, x_{n-1}$, a h jest funkcją taką, że $h(x_i) = \delta_{in}$ $(0 \leqslant i \leqslant n)$, to dla pewnego c funkcja $g + ch$ interpoluje f w $x_0, x_1, \ldots, x_n$.

3. Niech E będzie $(n+1)$-wymiarową przestrzenią wektorową funkcji określonych w obszarze D. Niech $x_0, x_1, \ldots, x_n$ będą różnymi punktami z D. Udowodnić, że dla $f \in E$ zadanie interpolacyjne

$$f(x_i) = y_i \quad (0 \leqslant i \leqslant n)$$

ma dla dowolnych rzędnych y_i jednoznaczne rozwiązanie wtedy i tylko wtedy, gdy jedyną funkcją z E znikającą we wszystkich x_i jest funkcja tożsamościowo równa 0.