Техніка наближення нейтрального елемента

М. В. Лешко

У цій оглядові статті спробуємо елементарно описати техніку наближення нейтрального елементу, проте певні загальні слова будуть сказані. У 1.1 описано загальний принцип техніки, а у 1.4 її конкретизовано для випадку **R**. У §2 окреслено як реалізувати план з 1.1, 1.4 для наближення неперервних функцій многочленами (також у зауваженні 2.2 дано інше доведення за допомогою многочленів Бернштейна і пояснено зв'язок з цією технікою). Також у 2.3 дано приклади застосування техніки для інших (цікавих) випадків — реалізація залишена як вправа для читача. Зокрема цікавим для мене питанням є поставлене в 1.6. У кінці, §3 містить деякі прості приклади застосування теорем як основного інгредієнту.

§1. Наближення одиниці

1.1. Розглядаючи операцію *згортки* (див 1.3) для деяких топологічний векторних просторів зазвичай виявляється, що нейтрального елементу згортки не міститься в цьому просторі, проте може бути гарно наближено послідовністю елементів простору. Використовуючи евристичну ідею, що *згортка унаслідує гарні властивості обох функцій* ми можемо отримати гарні теореми про наближення елементів цього простору.

1.2. Будемо розглядати стандартний випадок **R**. У такому випадку, *нейтральний елемент* (одиниця) це (узагальнена) функція, що носить назву *дельта функція Дірака* і визначена наступним чином

$$\delta(x) = \begin{cases} +\infty, & x = 0 \\ 0, & x \neq 0. \end{cases}$$

1.3. Ознайомимо читача з поняттям згортки (у випадку **R**). Припустимо, що функція f неперервна та набуває нульових значень (має *компактний носій*) поза якимось відрізком $[a, b]$. Будемо писати $\int_{-\infty}^{+\infty}$ маючи на увазі $\int_a^b$. *Згортка* $f * g$ двох неперервних функцій з компактними носіями f, g як

$$(f * g)(x) := \int_{-\infty}^{+\infty} f(y)g(x - y)dy.$$

Ми очікуємо від згортки наступні гарні і природні властивості:

1. $f * g$ — неперервна функція з компактним носієм,
2. $f * g = g * f$ — коммутативність,
3. $f * (g_1 + g_2) = f * g_1 + f * g_2$ та $f * (cg) = c(f * g)$.
4. $(f * g) * h = f * (g * h)$ — асоціативність,
5. $(f * g)' = f' * g = f * g'$ — згортка коммутує з дифференціюванням.

1.4. Отже, якщо ми можемо наблизити δ якоюсь послідовністю функцій δ_a то ми зможемо гарно наблизити інші хороші (наприклад неперервні або L^p функції). Стратегія окреслена в 1.1 підсумована в діаграмі

$$\begin{tikzcd}
\{\delta_a(x)\}_a \arrow[r] \arrow[d, "f*\cdot"] & \delta(x) \arrow[d, "f*\cdot"] \\
\{(f * \delta_a)(x)\}_a \arrow[r, dashed] & f(x)
\end{tikzcd}$$

1.5. У часткових випадках (мені невідомо як узагальнювати) такий підхід може видозмінюватись. Наприклад, маючи *неприємну* неперервну функцію на якомусь повному метричному просторі (X, d) можемо визначити функції

$$f_t(x) = \sup_{y \in \mathbf{R}} \{f(y) - td(x, y)\}.$$

Вони будуть Ліпшецивими (константа функції f_t буде t) і $f_t \geq f$. Тоді $\{f_t\}_t$ наближують f при $t \to \infty$.

1.6. Зауваження до 1.5. Ідея 1.5, хоч і може бути доведена напряму з використанням теореми 2 (Стоуна-Веєрштраса) в цілому схожа з тим, що відбувається для 1.4. У цьому випадку $-td(x, y)$ веде себе в цілому як $t\delta(t(x - y))$, згортка заміняється на sup, добуток на суму. У цілому ми, напевно, заміняємо кільце дійсних чисел на його *тропічні* півкільце. Мені цікаво чи маючи діаграму виду 1.4 (для якихось інших просторів) будувати такого виду 'дуальні' теореми про наближення (замінюючи кільця). Вище описано про

$$(\mathbf{R}, +, \times) \xrightarrow{tropic} (\mathbf{R}, \min, +)$$

§2. Теорема Стоуна-Веєрштраса

2.1. У наступних параграфах ми освітлюємо як ця стратегія може бути застосована для доведення *теореми Веєрштраса.*

Теорема 1 (Веєрштрас). Нехай f — неперервна функція на відрізку $[a,b]$. Тоді для будь-якого $\varepsilon > 0$ існує многочлен P_ε такий, що

$$\sup_{\mathbf{x}\in[\mathbf{a},\mathbf{b}]} |f(x) - P_\varepsilon(x)| < \varepsilon.$$

Без обмеження загальності будемо вважати відрізок $[a,b]$ рівним $[-1,1]$.

2.2. Зауваження до 2.1. Існує підхід Бернштейна, який конструктивно будує наближення $f \in C([0,1])$ многочленами

$$B_n(f)(x) = \sum_{k=0}^{n} f(k/n) b_{k,n}(x),$$

де $b_{k,n}(x)$ — *базисні многочлени Бернштейна*

$$b_{k,n}(x) = \binom{n}{k} x^k (1-x)^{n-k}$$

Тоді $B_n(f) \xrightarrow{n\to\infty} f$ рівномірно на $[0,1]$.

Ідея доведення Бернштейна наступна.

Крок 1. Розглядається випадкова величина $\xi \sim B(n,p)$ (біноміальний розподіл), тоді ймовірність $\mathbf{P}(\xi = k)$ виражається саме *базисним многочленом Бернштейна* $b_{k,n}(p)$. Далі застосовується (слабкий) закон великих чисел,

$$\lim_{n\to\infty} \mathbf{P}\left(|\xi/n - p| > \varepsilon\right) = 0$$

рівномірно по p для всіх $\varepsilon > 0$.

Крок 2. Оскільки f неперервна на компакті, вона рівномірно неперервна, тому

$$\lim_{n\to\infty} \mathbf{P}\left(|f(\xi/n) - f(p)| > \varepsilon\right) = 0,$$

а тому (формально це треба акуратно довести, поділити на маленьку і велику долю сподівання)

$$\lim_{n\to\infty} \mathbf{E}|f(\xi/n) - f(p)| = 0,$$

Крок 3. Далі виносимо модуль і $\mathbf{E}(f(\xi/n)) = B_n(f)(p)$.

Для простоти сприйняття це доведення викладено як у Вікіпедії (було забрано лишні слова), але його можна підсумувати в такій красивій діаграмі, підозріло схожій на 1.4, що, насправді, не випадково (наприклад, якщо p — ірраціональне, то ξ/n цей елемент ‘недосяжний’; $E \circ f$

можна сприймати в певному сенсі як 'згортку' для випадку коли одна з функцій це p пам'ятаючи, що $\mathbf{E}$ та інтеграл мають однакову суть):

$$\begin{tikzcd}
 & \{\xi/n\} \arrow[rr, "LLN"] \arrow[d, "f(\cdot)"] \arrow[dd, bend right, "\mathbf{E}\circ f"'] & & p \arrow[d, "f(\cdot)"] \arrow[dd, bend left, "\mathbf{E}\circ f"] \\
 & \{f(\xi/n)\} \arrow[rr, "LLN+unif"] \arrow[d, "\mathbf{E}"] & & f(p) \arrow[d, "\mathbf{E}"] \\
\{B_n(f)\} \arrow[r, equal] & \{\mathbf{E}(\xi/n)\} \arrow[rr, dashed] & & f(p)
\end{tikzcd}$$

2.3. Зауваження до 2.1. Теорему 1 можна сформулювати по-іншому: алгебра многочленів на відрізку $[a, b]$ щільна в $C([a, b])$. Існують схожі теореми які доводяться стратегією з 1.1 та 1.4,

1. $C_0^\infty(\mathbf{R})$ (простір гладких функцій з компактним носієм) щільний в $L^p(\mathbf{R})$.
2. $\mathcal{S}(\mathbf{R})$ (простір Шварца) щільний в $L^p(\mathbf{R})$.
3. Простір многочленів Лорана щільний в $L^2(S^1)$.
4. Це все ще узагальнюється на компактні Хаусдорфові абелеві групи та міри Хаара відповідно[1]

2.4. Реалізуємо перший рядок діаграми в 1.4. Треба знайти послідовність *многочленів*, що наближаються до дельта функції Дірака. Для цього побудуємо спочатку функції $f_n(x) = (1 - x^2)^n$. Замітимо, що на відрізку $[-1, 1]$ вони ведуть себе як *bump functions*, тобто f_n

1. парні,
2. нулі в $x = \pm 1$,
3. зростають на $[-1, 0]$,
4. спадають на $[0, 1]$,
5. $f_n(0) = 1$.

Тобто $f_n(x) \longrightarrow \delta_{x,0}$ де δ_{ij} — символ Кронекера. Тепер, щоб зробити $f_n(x) \longrightarrow \delta(x)$, де $\delta(x)$ — дельта функція Дірака, потрібно щоб $f(0) \xrightarrow{n\to\infty} +\infty$. Для цього зробимо невеличкий фокус, а саме обчислимо замітимо,

[1] Ідея схожа на минулий пункт. Там $\{z^n\}$ утворювало ортонормований базис, лінійна оболонка якого утворювала щільний підпростір. Тут характери G утворюють схожу конструкцію.

що для інтегралів

$$I_n := \int_{-1}^{1} (1 - x^2)^n dx$$

виконується рекурентне співвідношення

$$I_n = \left(1 - \frac{1}{2n+1}\right) I_{n-1}$$

Далі можна довести (вправа), що $I_n \xrightarrow{n\to\infty} 0$, а тому $\delta_n(x) = f_n(x)/I_n$ будуть наближати дельта функцію Дірака.

2.5. Зауваження до 2.4. Існує багато інших гарних способів отримати дельта функцію Дірака як (слабку) границю, наприклад, якщо взяти

1. Гаусові функції
$$\delta(x) = \lim_{a\to 0} \frac{1}{|a|\sqrt{\pi}} e^{-(x/a)^2}.$$
2. Трикутні функції
$$\delta(x) = \lim_{a\to 0} \frac{1}{a} \max\{1 - |x/a|, 0\}$$
3. Sincs (особливо гарно з працює з спадщиною Фур'є),
$$\delta(x) = \lim_{n\to\infty} \frac{\sin \pi n x}{\pi x}.$$
4. Півколовий розподіл Вігнера (припускаємо, що функція 0 поза областю визначення)
$$\delta(x) = \lim_{a\to 0} \frac{2}{\pi a^2} \sqrt{a^2 - x^2}.$$
5. Теплове ядро, ядро Пуасонна, тощо

2.6. Наступний крок: довести стрілочки вниз, а саме потрібно довести, що згортка (неперервної) функції f та многочленів δ_n є многочленом. Це також просто і залишається як вправа (потрібно скористуватись $(f * g)' = f' * g = f * g'$ та диференціюванням під знаком інтеграла). Тоді залишиться лише перевірити нижній рядок, що теж не є складоною задачею.

2.7. Маючи теорему 1 можемо провести певні топологічні аргументи і отримати таке гарне узагальнення (у термінах 2.3)

Теорема 2 (Стоун, Веєрштрас). Підалгебра $A \subset C(X)$ (X — компактний Хаусдорфів простір) щільна в $C(X)$ тоді і тільки тоді, коли вона *розділяє точку*, тобто для будь-яких різних точок $x, y \in X$ існує функція $f \in A$, така, що $f(x) \neq f(y)$.

Наприклад, наближення Ліпшецивими функціями з 1.6 майже напряму слідує з цієї теореми.

§3. Деякі приклади застосувань

3.1. Простий, але ефективний приклад, що пояснює, наприклад, як застосовувати *ε-наближення.*

Теорема 3 (Ріман, Лебег)**.** Довести, що якщо $f \in L^1(\mathbf{R})$ то перетворення Фур'є

$$\hat{f}(\xi) = \int_{\mathbf{R}} f(x)e(x\xi)dx$$

прямує до нуля коли $\xi \longrightarrow \infty$.

Крок 1. За 2.3 маємо, що $C_0^\infty(\mathbf{R})$ щільний в $L^1(\mathbf{R})$, тому для будь-якого $\varepsilon > 0$ існує функція f_ε така, що

$$\|f - f_\varepsilon\|_1 = \int_{\mathbf{R}} |f(x) - f_\varepsilon(x)|dx \leq \varepsilon.$$

Твердження теореми очевидне (чому?) для гладких функцій з компактним носієм.

Крок 2. Тепер треба акуратно перейти до $\varepsilon \to 0$. Для цього потрібна рівномірна збіжність $\hat{f}_\varepsilon$ до $\hat{f}$, що дійсно правда, оскільки

$$\sup_\xi |\hat{f}(\xi) - \hat{f}_\varepsilon(\xi)| \leq \|f - f_\varepsilon\|_1 \leq \varepsilon.$$

3.2. Просте, гарне і дуже корисне використання варіації теореми 1 для тригонометричних многочленів подано в наступній теоремі.

Теорема 4 (Вейль). Послідовність (a_n) називається[a] *рівномірно розподіленою* на $\mathbf{R}/\mathbf{Z}$ (по модулю 1, на колі), якщо для будь-якої інтегрованої по Ріману функції $f : \mathbf{R}/\mathbf{Z} \longrightarrow \mathbf{C}$ має місце

$$\frac{1}{N}\sum_{n=1}^{N} f(a_n) \xrightarrow{N\to\infty} \int_{\mathbf{R}/\mathbf{Z}} f(x)dx. \tag{1}$$

Тоді має місце такий критерій. Послідовність (a_n) є рівномірно розподіленою на $\mathbf{R}/\mathbf{Z}$ тоді і тільки тоді, коли для всіх $k \in \mathbf{Z} \setminus \{0\}$ має місце

$$\frac{1}{N}\sum_{n=1}^{N} e^{2\pi i k a_n} \longrightarrow 0. \tag{2}$$

[a] інколи вводять трішки інше визначення в стилі кількісної міри, наприклад на Вікіпедії.

Одразу замітимо, що якщо послідовність рівномірно розподілена, то (2) очевидно виконується.

Якщо (2) виконується то для таких функцій $f(x) = e^{2\pi i k x}$ виконується (1), а тому за лінійністю маємо (1) і для всіх *тригонометричних многочленів*. Далі за теоремою 1 для тригонометричних многочленів (а краще 2) маємо (1) для всіх неперервних функцій. Далі неперервними функціями наближаємо інтегровані по Ріману.

Метод експоненційних сум (див (2)) як і критерій Вейля гарно використовується в *аналітичній теорії чисел*.

3.3. Зауваження до 3.2. Спробуйте довести дещо схоже (але простіше) твердження:

Теорема 5. Якщо $f : [0,1] \longrightarrow \mathbf{R}$ — неперервна функція, така, що для всіх цілих невід'ємних n має місце

$$\int_0^1 f(x)e^{nx}dx = 0,$$

то $f \equiv 0$ всюди.