Приближение линейными пространствами

Пусть H — гильбертово пространство, $L \subset H$. Мы знаем, что $H = L \oplus L^{\perp}$, т.е. любой $x \in H$ однозначно представляется в виде $x = l + l'$, $l \in L$, $l' \perp L$ (напомним: l — ближайший к x в L, разность $x - l$ ортогональна, т.к. иначе образуется острый угол и расстояние можно уменьшить). Отображение $P_L \colon x \mapsto l$ называется ортопроектором на подпространство L. Напомним его свойства:

- P_L — линейный оператор на H;
- P_L — проектор на L, то есть $P_L x \in L$ и $P_L y = y$ для $y \in L$;
- $P_L x$ — единственный ближайший к x элемент L:

$$|x - y|^2 = |x - P_L x|^2 + |P_L x - y|^2 \quad \text{для } y \in L.$$

  Тем самым, метрическая проекция x на L есть $P_L x$ (точнее, одноточечное множество $\{P_L x\}$), что позволяет нам использовать для них одно и то же обозначение P.

Следствие: евклидовы пространства устроены проще в смысле аппроксимации, наилучшее приближение линейным подпространством задаётся линейным оператором.

Определение: система $u_1, \dots, u_n \in H$ ортонормирована, если $|u_i| = 1$ и $\langle u_i, u_j \rangle = 0$ при $i \neq j$.

Утверждение: если $u_1, \dots, u_n$ — ортонормированный базис L, то ортопроектор имеет вид

$$P_L x = \sum_{k=1}^{n} \langle x, u_k \rangle u_k.$$

Если H сепарабельно, то в нём существует ортонормированный базис (полная ортонормированная система) $u_1, \dots, u_n, \dots$. Пусть $\dim H = \infty$. Для любого $x \in H$ имеем

$$x = \sum_{k=1}^{\infty} \langle x, u_k \rangle u_k \quad \text{— ряд Фурье по системе } \{u_k\},$$

$$|x|^2 = \sum_{k=1}^{\infty} \langle x, u_k \rangle^2 \quad \text{— равенство Парсеваля.}$$

При этом наилучшее приближение первыми n функциями u_k даётся частичной суммой ряда Фурье.

Ортогонализация: систему линейно независимых векторов $x_1, \ldots, x_n \in H$ можно превратить в ортогональную — из x_k вычитаем ортопроекцию на подпространство, порождённое $x_1, \ldots, x_{k-1}$.

Матрица Грама: $G_{i,j} = \langle x_i, x_j \rangle$. Матрица невырождена, т.к. её определитель не меняется при ортогонализации системы.

Пример: пространство $L_2[0, 2\pi]$ со скалярным произведением

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(x) g(x)\, dx.$$

Тригонометрическая система: $\{1,\ \sqrt{2} \cos kx,\ \sqrt{2} \sin kx\}_{k=1}^{\infty}$ — ортонормирована, ряд Фурье:

$$f = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos kx + b_k \sin kx,$$

где $a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx\, dx$, $b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx\, dx$.

Замечание: теория евклидовых пространств переносится на случай пространств над полем $\mathbb{C}$. В этом случае от скалярного произведения требуются свойства: $\langle y, x \rangle = \overline{\langle x, y \rangle}$, линейность по первому и и антилинейность по второму аргументу: $\langle x, \lambda y \rangle = \overline{\lambda} \langle x, y \rangle$, и $\langle x, x \rangle > 0$ при $x \neq 0$.

Пример: пространство $L_2^{\mathbb{C}}[0, 2\pi]$ со скалярным произведением

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(x) \overline{g(x)}\, dx.$$

Тригонометрическая система в комплексной форме: $\{e^{ikx}\}_{k \in \mathbb{Z}}$, ряд Фурье

$$f = \sum_{k \in \mathbb{Z}} c_k e^{ikx}, \quad c_k = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx}\, dx.$$

Пространство тригонометрических полиномов порядка не выше n:

$$\mathcal{T}_n = \operatorname{span}\{\cos(kx), \sin(kx)\}_{k=0}^{n},$$

$$\mathcal{T}_n^{\mathbb{C}} = \operatorname{span}\{e^{ikx}\}_{k=-n}^{n}.$$

Ясно, что

$$\mathcal{T}_n = \{T \in \mathcal{T}_n^{\mathbb{C}} \colon T(x) \in \mathbb{R} \;\; \forall x\} = \{\sum_{k=-n}^{n} c_k e^{ikx} \colon c_k \equiv \overline{c_{-k}}\}.$$

Наилучшее приближение в L_2 пространством $\mathcal{T}_n$ даётся начальным отрезком ряда Фурье:

$$f \approx S_n(f) = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos kx + b_k \sin kx = \sum_{|k| \leqslant n} c_k e^{ikx}.$$

В других L_p нахождение наилучшего полинома — сложная задача!

Пример: система Хаара в $L_2[0,1]$: $h_0 = 1$, далее разбиваются на "пачки"

$$h_{k,j}(x) = \begin{cases} 2^{k/2}, & \frac{j-1}{2^k} < x < \frac{j-1/2}{2^k}, \\ -2^{k/2}, & \frac{j-1/2}{2^k} < x < \frac{j}{2^k}. \end{cases}$$

Задача: для непрерывной функции $f \not\equiv \mathrm{const}$ невозможно, чтобы $\max_j h_{k,j} = o(2^{-3k/2})$ при $k \to \infty$.

Пример: система Радемахера $r_k(x) = (-1)^{x_k}$, где $x = \sum_{k=1}^{\infty} x_k 2^{-k}$. Не полна (почему?).

Пример: триг. система $\{e^{ikx}\}$ обладает тем свойством, что $|e^{ikx}| \equiv 1$. Существует ли вещественная система с тем же свойством? Да, например, система Уолша: $\{r_{k_1}(x) \cdot r_{k_2}(x) \cdot \ldots r_{k_s}(x)\}$. Она уже полна.

Пример: при ортогонализации мономов $\{1, x, x^2, \ldots\}$ в $L_2[-1,1]$ получаем *многочлены Лежандра* $P_n(x)$. Они обладают рядом свойств, например, $P_n(x) = 0$ имеет n решений на $[-1,1]$ (почему?).

Конечномерный случай.

Изучим подробнее случай $H = \mathbb{R}^n$. Подпространство $L \subset \mathbb{R}^n$ размерности k задаётся матрицей A размера $n \times k$, в столбцы которой записаны координаты какого-либо базиса $\{v_1, \ldots, v_k\}$ пространства L. Тогда $L = \{Az \colon z \in \mathbb{R}^k\}$. Найдём матрицу оператора $P_L \colon \mathbb{R}^n \to \mathbb{R}^n$. Имеем $P_L x = \sum_{i=1}^{k} z_i v_i$ с неизвестным $z \in \mathbb{R}^k$. Условие $x - P_L x \perp L$ равносильно тому, что $x - P_L x \perp v_j$, $j = 1, \ldots, k$, то есть

$$\langle x, v_j \rangle = \langle P_L x, v_j \rangle = \sum_{i=1}^{k} z_i \langle v_i, v_j \rangle, \quad j = 1, \ldots, k,$$

что в матричном виде записывается как $A^t x = (A^t A)z$, откуда $z = (A^t A)^{-1} A^t x$. *(Почему матрица $A^t A$ обратима?)* Окончательно, $P_L x = Az$,

$$P_L x = A(A^t A)^{-1} A^t x.$$

Пример: линейная регрессия (метод наименьших квадратов). Известна "выборка" $(X_1, Y_1), \ldots, (X_n, Y_n)$ пар, состоящих из векторов $X_i \in \mathbb{R}^k$ и соответствующих им чисел $Y_i \in \mathbb{R}$. Требуется аппроксимировать неизвестную нам зависимость $Y \approx F(X)$ линейной функцией от X:

$$\sum_{i=1}^{n} |Y_i - \langle v, X_i \rangle|^2 \to \min_v .$$

Мы должны приблизить вектор $Y = (Y_1, \ldots, Y_n)^t$ векторами вида $\mathbf{X}v$, где в матрице $\mathbf{X}$ размера $n \times k$ записаны по *строкам* вектора X_i, $i = 1, \ldots, n$.. Отсюда, оптимальная аппроксимация имеет вид:

$$Y \approx \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t X.$$

Прямые и обратные теоремы

Пусть $f = \sum c_k e^{ikx}$ (можно считать, что функция вещественная, но удобней всё равно работать с комплексной записью). Определим наилучшее приближение

$$E_n(f)_{L_2} = \inf_{T \in \mathcal{T}_n} \|f - T\|_2 = \|f - \sum_{|k| \leqslant n} c_k e^{ikx}\|_2 = (\sum_{|k|>n} |c_k|^2)^{1/2}$$

и модуль непрерывности

$$\omega(f, \delta)_{L_2} := \sup_{0 \leqslant h \leqslant \delta} \|f(x+h) - f(x)\|_{L_2}.$$

Теорема (Джексон — прямая теорема): Для любой функции $f \in L_2[0, 2\pi]$ имеем $E_n(f)_{L_2} \leqslant C\omega(f, \pi/n)_{L_2}$.

Доказательство.

$$\|f(x+h) - f(x)\|_2^2 = \| \sum c_k (e^{ikh} - 1) e^{ikx} \|_2^2 = \sum |c_k|^2 |e^{ikh} - 1|^2 = \sum |c_k|^2 4 \sin^2 \frac{kh}{2}.$$

Положим $\delta_n := \pi/n$ и усредним по отрезку $[0, \delta_n]$:

$$\omega^2(f, \delta_n) \geqslant \frac{1}{\delta_n} \int_0^{\delta_n} \|f(x+h) - f(x)\|_2^2 \, dh = 4 \sum |c_k|^2 \frac{1}{\delta_n} \int_0^{\delta_n} \sin^2 \frac{kh}{2} \, dh.$$

При $k > n$ последний интеграл

$$\frac{1}{\delta_n}\int_0^{\delta_n}\sin^2\frac{kh}{2}\,dh = \int_0^1 \sin^2\frac{\pi kx}{2n}\,dx \geqslant \text{const}.$$

Следовательно,

$$\omega^2(f,\delta_n) \geqslant C\sum_{|k|>n}|c_k|^2.$$

Теорема (Бернштейн — обратная теорема): Для любой функции $f \in L_2[0,2\pi]$ имеем

$$\omega^2(f,\pi/n)_{L_2} \leqslant \frac{C}{n^2}\sum_{k=1}^{n} kE_{k-1}^2(f)_{L_2}.$$

Доказательство.

$$\|f(x+h)-f\|_2^2 = \sum_{|k|<n}|c_k|^2 4\sin^2(kh/2) + \sum_{|k|\geqslant n}|c_k|^2 4\sin^2(kh/2) \leqslant$$
$$\leqslant \sum_{|k|<n}|c_k|^2k^2h^2 + 4\sum_{|k|\geqslant n}|c_k|^2.$$

Откуда

$$\omega^2(f,\pi/n) \leqslant \frac{\pi^2}{n^2}\sum_{k=1}^{n-1}(|c_k|^2+|c_{-k}|^2)k^2 + 4E_{n-1}^2(f) =$$
$$= \frac{\pi^2}{n^2}\sum_{k=1}^{n-1}k^2(E_{k-1}^2(f)-E_k^2(f)) + 4E_{n-1}^2(f).$$

Последнее слагаемое легко оценивается.