Спецкурс “Многомерная аппроксимация и анализ булевых функций”. Лекция 6: Меры сложности булевых функций

Определения. Пусть $f\colon \{-1,1\}^n \to \{-1,1\}$.

Мы уже рассмотрели одну из мер сложности — *степень* функции $\deg f$. Она равна степени многочлена, представляющего f. Эквивалентно, $\deg f = \max\{|S|\colon \hat{f}(S) \neq 0\}$.

Аппроксимативная степень функции $\widetilde{\deg} f$ равна минимальной степени многочлена $P(x_1,\dots,x_n)$, приближающего f:

$$\|f - P\|_\infty := \max_{x\in\{-1,1\}^n} |f(x) - P(x)| \leqslant 2/3.$$

Погрешность $2/3$ не принципиальна; главное, чтобы значения -1 и 1 после аппроксимации были гарантировано отделены положительной константой.

Перейдём теперь к комбинаторным мерам сложности.

Чувствительность функции f в точке x равна количеству бит, изменение которых меняет значение функции:

$$s(f,x) = \{i \in [n]\colon f(x^{\oplus i}) \neq f(x)\}.$$

Чувствительность функции $s(f) = \max_x s(f,x)$.

Вопрос: чему равна средняя чувствительность $\mathsf{E}s(f,x)$?

Разрешим теперь менять не один бит, а все биты в некотором блоке, т.е. для $x \in \{-1,1\}^n$ определяем $x^{\oplus B} \in \{-1,1\}^n$ так:

$$x_i^{\oplus B} = \begin{cases} -x_i, & i \in B, \\ x_i, & i \notin B. \end{cases}$$

Блок называется чувствительным, если $f(x^{\oplus B}) \neq f(x)$.

Блочная чувствительность функции f в точке x равна максимальному количеству непересекающихся чувствительных блоков: $B_1,\dots,B_b \subset [n]$, $B_i \cap B_j = \varnothing$, $f(x^{\oplus B_i}) \neq f(x)$ для всех i. Блочная чувствительность $\mathrm{bs}(f) = \max_x \mathrm{bs}(f,x)$.

Сертификатом функции f в точке x называется всякое множество координат $S \subset [n]$, которое гарантирует значение $f(x)$, т.е. такое что

$$x'\big|_S \equiv x|_S \quad \Rightarrow f(x') = f(x)$$

для любого x'. *Сертификационной сложностью* f в точке x называется минимальный размер сертификата в этой точке, $C(f,x)$. Сертификационная сложность функции равна $C(f) = \max_x C(f,x)$.

Наконец, рассмотрим представление функции в виде *дерева решений*. Это двоичное дерево, в каждом узле которого производится запрос значения некоторого бита x_i, возникают две ветви, соответствующие ответам $x_i = -1$ и $x_i = 1$. В листе такого дерева записано некоторое значение. Дерево вычисляет функцию f, если на любом входе x, если пройти по соответствующему пути до листа, в листе будет записано значение $f(x)$. *Глубина дерева* это максимальная длина пути.

Древесная сложность $D(f)$ равна минимальной глубине дерева, вычисляющего f.

Случай $\{0,1\}$-функций. Случай $f\colon \{-1,1\}^n \to \{-1,1\}$ вполне аналогичен случаю $g\colon \{0,1\}^n \to \{0,1\}$. Все комбинаторные меры сложности (s, bs, C, D) переносятся дословно. Что происходит с алгебраическими мерами сложности? Напомним, что $\{-1,1\}$ отождествляется с $\{0,1\}$ так: $-1 \mapsto 1$, $1 \mapsto 0$. Другими словами, $y_i = (1-x_i)/2$ и $x_i = 1-2y_i$. Значит, функции $f(x_1,\ldots,x_n)$ сопоставляется функция

$$g(y_1,\ldots,y_n) = \frac{1}{2}(1 - f(1-2y_1,\ldots,1-2y_n)).$$

Ясно, что при таких преобразованиях многочлены остаются многочленами той же степени, $\deg f = \deg g$. А вот для аппроксимативной степени погрешность становится $1/3$:

$$\|g - Q\|_\infty \leqslant 1/3.$$

Далее нам будет удобнее работать с $\{0,1\}$-функциями.

Элементарные неравенства. Докажем цепочку неравенств

$$s(f) \leqslant \mathrm{bs}(f) \leqslant C(f) \leqslant D(f).$$

Неравенство $\mathrm{bs} \geqslant s$ очевидно: в качестве чувствительных блоков для x возьмём однобитные блоки с чувствительными битами.

Неравенство $\mathrm{bs} \leqslant C$ тоже докажем поточечно: $\mathrm{bs}(f,x) \leqslant C(f,x)$. Для этого возьмём минимальный сертификат S для f в x и рассмотрим непересекающиеся чувствительные блоки $B_1,\ldots,B_b$; нужно показать, что $b \leqslant |S|$. Если блок чувствительный, то он должен пересекаться с сертификатом, поскольку изменение любых битов вне сертификата не меняет значение функции. Значит, $B_j \cap S \neq \varnothing$. Но блоки не пересекаются между собой, поэтому их не больше $|S|$.

Неравенство $C \leqslant D$. Дерево решений даёт следующий сертификат в точке x: возьмём все биты, запрошенные в дереве решений на пути к листу, при входе x. (Нарисуйте пример!) Они однозначно определяют значение и поэтому дают сертификат.

Докажем также, что

$$\widetilde{\deg} f \leqslant \deg f \leqslant D(f).$$

Первое неравенство очевидно. Второе: дереву соответствует многочлен. Допустим, дерево в корневой вершине запрашивает значение бита x_i; левая ветвь, $x_i = 0$, вычисляет функцию $L(x)$, а правая, соответствующая $x_i = 1$, вычисляет функцию $R(x)$. Тогда всё дерево вычисляет $(1-x_i)L(x) + x_iR(x)$. Продолжая это разложения для L и R, мы получим многочлен, степень которого равна глубине, и который совпадает с нашей функцией.

Аппроксимативная степень симметричных функций. Напомним, что функция f называется симметричной, если

$$f(x_1,\ldots,x_n) = f(\sigma x) := f(x_{\sigma(1)},\ldots,x_{\sigma(n)})$$

для любой перестановки σ и любого x. Такие функции определяются своими значениями вида

$$f_k = f(1,1,\ldots,1,0,\ldots,0) \quad (k \text{ единиц}).$$

Тогда f_k это значение функции на множестве векторов с $x_1 + \ldots + x_n = k$.

Покажем, что для приближения симметричных функций достаточно использовать многочлены вида $q(x_1 + \ldots + x_k)$. Пусть $P(x_1,\ldots,x_n)$ — многочлен, приближающий симметричную булеву функцию $g\colon \{0,1\}^n \to \mathbb{R}$. Мы можем его симметризовать:

$$P^{\mathrm{sym}}(x_1,\ldots,x_n) = \frac{1}{n!}\sum_\sigma P(\sigma x).$$

Тогда, используя симметричность g,

$$\|g - P^{\mathrm{sym}}\|_\infty = \|g^{\mathrm{sym}} - P^{\mathrm{sym}}\|_\infty \leqslant \|g - P\|_\infty,$$

т.е. P^{sym} приближает g не хуже. (Это хорошо известный эффект в теории приближений: если нужно приблизить чётную функцию на отрезке $[-1, 1]$, то можно считать многочлен чётным, заменив $P(x)$ на $(P(x) + P(-x))/2$.)

Что произойдёт с мономами вида $x_{i_1} \cdots x_{i_m}$? Для них симметризация равна (с точностью до множителя) элементарным симметрическим многочленам

$$\sum_\sigma x_{\sigma(1)} \cdots x_{\sigma(m)}. \tag{1}$$

Проанализируем их, пользуясь тем, что x на булевом кубе $\{0,1\}^n$. Предположим, $x_1 + \ldots + x_n = k$, чему равно выражение (1)? Если $m > k$, то выражение равно нулю (среди m чисел всегда найдётся ноль); если $m \leqslant k$, то оно равно $\binom{k}{m}$. Биномиальный коэффициент можно считать многочленом по k:

$$\binom{k}{m} = \frac{1}{m!} k(k-1)\cdots(k-m+1),$$

причём такая формула работает и при $k < m$. Значит, (1) это многочлен от $k = x_1 + \ldots + x_m$ той же степени m. Поэтому на булевом кубе

$$P^{\mathrm{sym}}(x_1, \ldots, x_n) = q(x_1 + \ldots + x_n), \quad \deg q = \deg P^{\mathrm{sym}}.$$

Итак, задача аппроксимации g сводится к построению q, $\deg q \leqslant d$, такого что

$$\max_{0 \leqslant k \leqslant n} |g_k - q(k)| \leqslant 1/3. \tag{2}$$

Теорема 1 (Paturi)**.** *Пусть функция $g\colon \{0,1\}^n \to \{0,1\}$ симметрична. Рассмотрим величину*

$$r(g) := \max_{k\colon\, g_k \neq g_{k+1}} \min\{k+1, n-k\}.$$

Тогда

$$c\sqrt{n\, r(g)} \leqslant \widetilde{\deg}(g) \leqslant C\sqrt{n\, r(g)},$$

где c, C абсолютные постоянные.

Таким образом, для аппроксимации симметричных функций нужна всегда степень не меньше $n^{1/2}$ по порядку. Наиболее плохой случай скачка около середины, $r \approx n/2$. даёт степень $\approx n$.

Доказательство (только оценка сверху). Построим кусочно-линейную функцию $h(t)$, равную $g(k)$ в целых точках $t = k$. Отметим, что $|h'(t)| \leqslant 1$ и $h \equiv \mathrm{const}$ на отрезке $[r, n-r]$, где $r = r(g)$. Сделаем тригонометрическую подстановку:

$$t = \frac{n}{2}(1 - \cos\theta), \quad H(\theta) := h(t(\theta)), \quad 0 \leqslant \theta \leqslant \pi.$$

Тогда H — 2π-периодическая чётная функция. Вычислим её производную:

$$H'(\theta) = t'(\theta) g'(t(\theta)),$$

$$t'(\theta) = \frac{n}{2}\sin\theta = \frac{n}{2}\sqrt{1 - \cos^2\theta} = \frac{n}{2}\sqrt{1 - (1 - 2t/n)^2} =$$

$$= \frac{n}{2}\sqrt{2t/n \cdot (2 - 2t/n)} = \frac{n}{2}\frac{2}{n}\sqrt{t(n-t)} = \sqrt{t(n-t)}.$$

Поэтому $|t'(\theta)| \leqslant \sqrt{t(n-t)}$. Значит, в области $t \notin [r, n-r]$, где H не константа, справедлива оценка $|H'| \leqslant \sqrt{rn}$.

Теперь мы воспользуемся классическим результатом из теории аппроксимаций. Если F — 2π-периодическая функция, её естественно приближать тригонометрическими полиномами, т.е. функциями вида

$$T(\theta) = a_0 + \sum_{k=1}^{m} a_k \cos k\theta + b_k \sin k\theta.$$

Обозначим через $E_m(F)$ наилучшее приближение F такими функциями:

$$E_m(F) := \min_{a_k, b_k} \|F - T\|_{C[0,2\pi]} = \min_{a_k, b_k} \max_{\theta} \left| F(\theta) - (a_0 + \sum_{k=1}^{m} a_k \cos k\theta + b_k \sin k\theta) \right|.$$

(В пространстве L_2 наилучшее приближение достигается на тригонометрическом полиноме, равном начальному отрезку ряда Фурье функции F. В пространстве C это уже не обязательно так.)

Неравенство Джексона. Для любой непрерывной 2π-периодической функции F и любого $m \in \mathbb{N}$ справедливо неравенство

$$E_m(F) \leqslant C\omega(F, 1/m), \qquad \text{где } \omega(F, \delta) := \sup_{|\theta - \theta'| \leqslant \delta} |F(\theta) - F(\theta')|.$$

Наиболее точная и законченная форма этого неравенства — теорема Корнейчука:

$$E_{m-1}(F) \leqslant \omega(F, \pi/m).$$

Из неравенства Джексона вытекает, что $E_m(F) \leqslant (C/m)\|F'\|_\infty$.

Продолжаем доказательство теоремы. Подставляя оценку для нашей функции H, получаем $E_m(H) \leqslant (C/m)\sqrt{nr}$. При $m \asymp \sqrt{nr}$ получим оценку $E_m(H) \leqslant 1/3$, т.е. построим тригонометрический полином $T(\theta) = a_0 + \sum_{k=1}^{m}(a_k \cos k\theta + b_k \sin k\theta)$, $\|H - T\|_\infty \leqslant 1/3$.

Теперь заметим, что можно считать T чётным, поскольку функция H чётна (почему?). Т.е. в T входят только косинусы (почему?). Но чётный тригонометрический полином является алгебраическим многочленом от $\cos\theta$:

$$T(\theta) = p(\cos\theta), \quad \deg p \leqslant m.$$

Это следует, например, из того, что $\cos k\theta = T_k(\cos\theta)$, где T_k это многочлены Чебышёва (что можно доказать индукцией по k). Значит, $|g(t) - q(t)| \leqslant 1/3$, $0 \leqslant t \leqslant n$, для многочлена q, полученного из p линейной заменой переменной. Многочлен q и даёт искомую аппроксимацию (2). □

Примеры. Первый: функция $\mathrm{OR}(x_1, \ldots, x_n) = x_1 \vee x_2 \vee \ldots \vee x_n$, т.е. $\mathrm{OR}(x) = 1$ всюду, кроме $x = (0, \ldots, 0)$.

Найдём степень OR. Возьмём характер $\chi_{[n]}$; в 0,1-версии он равен $\chi_{[n]}(x_1, \ldots, x_n) = (-1)^{x_1 + \ldots + x_n}$. Вычисляем:

$$\widehat{\mathrm{OR}}(\{1, \ldots, n\}) = \langle \mathrm{OR}, \chi_{[n]} \rangle = \langle 1 - \mathbf{1}\{x_1 = \cdots = x_n = 0\}, \chi_{[n]} \rangle = \\ = \langle 1, \chi_{[n]} \rangle - 2^{-n}\chi_{[n]}(0, \ldots, 0) = -2^{-n}\chi_{[n]}(0, \ldots, 0).$$

Значит, последний коэффициент Фурье ненулевой и функция имеет степень ровно n.

Аппроксимативная степень. По теореме Paturi, $\widetilde{\deg}(\mathrm{OR}) \asymp n^{1/2}$ (скачок происходит только между $k = 0$ и $k = 1$). Но мы доказали только оценку сверху, докажем оценку снизу

непосредственно для OR (на самом деле, мы докажем оценку снизу $c\sqrt{n}$ для любой симметричной функции; просто не всегда эта оценка будет точна по порядку).

В случае OR задача (2) превращается в следующую: $|q(0)| \leqslant 1/3$, $|q(k) - 1| \leqslant 1/3$, $k = 1, \ldots, n$, $\deg q \leqslant d$. Перенесём эту картину на отрезок $[-1, 1]$. Получается многочлен $\widetilde{q}$ степени d на отрезке $[-1, 1]$, который на равномерной сетке с шагом $2/n$ принимает значения близкие к 0 или к 1. Воспользуемся ещё одним классическим результатом.

Неравенство Маркова. Для любого многочлена p степени d справедливо неравенство $\|p'\|_{C[-1,1]} \leqslant d^2 \|p\|_{C[-1,1]}$.

Выведем из неравенства Маркова, что степень нашего многочлена $\widetilde{q}$ не меньше $c\sqrt{n}$.

Предположим, $d < \sqrt{n}/10$. Пусть M это максимум модуля $\widetilde{q}$ на $[-1, 1]$. В силу неравенства Маркова $|\widetilde{q}'| \leqslant M \cdot n/100$ на этом отрезке. Значит, от точки максимума до ближайшей точки сетки модуль многочлена может уменьшиться до $0.99M$ и не более. Поэтому $M \leqslant 2$ и $|\widetilde{q}'| \leqslant n/50$. Но тогда многочлен не успеет за расстояние $2/n$ пройти от значения близкого к 0 до значения близкого к 1. Противоречие.

Итак, $\widetilde{\deg}(\mathrm{OR}) \asymp n^{1/2}$.

Нетрудно проверить, что $s(\mathrm{OR}) = n$ (взяв $x = 0$), откуда $\mathrm{bs}(\mathrm{OR}) = C(\mathrm{OR}) = D(\mathrm{OR}) = n$.

Второй пример: функция Рабиновича. Рассмотрим функцию от $n = m^2$ переменных $x_1, \ldots, x_n$, разбитых на m групп по m штук: $(x_1, \ldots, x_m)$, $(x_{m+1}, \ldots, x_{2m})$ и т.д. Функция $f(x) = 1$, если есть хотя бы одна "правильная" группа. Группа является правильной, если все x_i в ней равны нулю, кроме двух единиц, стоящих подряд.

Нетрудно проверить, что $s(f) \asymp m$ (сделайте это!). Докажем, что $\mathrm{bs}(f) \asymp m^2$. Действительно, для $x = 0$ чувствительные блоки это пары (x_{2i-1}, x_{2i}) соседних координат в группах.