1 Тропическая геометрия нейронных сетей

В разделе 5 нейронные сети определяются с помощью тропической алгебры, что позволяет нам изучать их с помощью тропической алгебраической геометрии. Мы покажем, что граница принятий решений нейронной сети — это подмножество тропической гиперповерхности, соответствующего тропического полинома(Раздел 6.1). Мы увидим, что в некотором смысле, зонотопы образуют геомтрические строительные блоки для нейронных сетей (Раздел 6.2). Затем мы докажем, что геометрия функции, представленной нейронной сетью, становится значиттельно более сложной с увеличением ее количества слоёв.

1.1 Границы решений нейронной сети

Мы будем использовать тропическую геометрию и идеи из Раздела 5 для изучения границ решений нейронных сетей, фокусируясь на случае классификации двух категорий для ясности. Как объяснено в Разделе 4, нейронная сеть $\nu : \mathbf{R}^d \to \mathbf{R}^p$ вместе с выбором функции оценки $s : \mathbf{R}^p \to \mathbf{R}$ дают нам классификатор. Если выходное значение $s(\nu(x))$ превышает некоторый порог принятия решений c, то нейронная сеть предсказывает, что x относится к одной категории (например, x — изображение кота), а в противном случае x относится к другой категории (например, x — изображение собаки). Таким образом входной пространство разделено на два непересекающихся подмножества *границей принятия решений* $B := \{x \in \mathbf{R}^d : \nu(x) = s^{-1}(c)\}$. Связанные области со значением выше порога и связные области со значением ниже порога будем называть *положительными* и *отрицательными областями* соответственно.

Предоставим оценки на количество положительных и отрицательных областей и покажем, что существует тропический многочлен, чья тропическая гиперповерхность содержит границу решений.

Утверждение 1 *(Тропическая геометрия границы решений). Пусть $\nu : \mathbf{R}^d \to \mathbf{R}$ — L-слойная нейронная сеть, удовлетворяющая предположению (a) − (c) с $t^{(L)} = -\infty$. Пусть функция счета $s : \mathbf{R} \to \mathbf{R}$ является инъективной с порогом принятия решений c в его диапазоне. Если $\nu = f \oslash g$, где f и g — тропические многочлены, тогда*

1. *Его граница решений $B = \{x \in \mathbf{R}^d : \nu(x) = s^{-1}(c)\}$ делит $\mathbf{R}^d$ на не более чем $N(f)$ связных положительных областей и не более, чем $N(g)$ связных отрицательных областей;*

2. *Его раница решений содержится в тропической гиперповерхности тропического многочлена $s^{-1} \odot g(x) \oplus f(x) = \max\{f(x), g(x) +$*

$s^{-1}(c)\}$, *то есть*

$$B \subset T(s^{-1}(c) \odot g \oplus f)$$

Функция $s^{-1}(c) \odot g \oplus f$ *не обязательно линейна на каждой положительной или отрицательной области и поэтому ее тропическая гиперповерхность* $T(s^{-1}(c) \odot g \oplus f)$ *может дальше делить положительные или отрицательные области, полученные из* B *на несколько линейных областей. В общем случае* $\subset$ *нельзя заменить на* $=$*.*

1.2 Зонотопы, как геометрические строительные блоки нейронной сети

Из раздела 3, мы знаем, что число областей тропической гиперповерхности $T(f)$ делит пространство на равное число вершин в двойственном разбиении многоугольника Ньютона, связанного с тропическим многочленом f. Это позволяет нам ограничить количество линейных областей нейронной сети, ограничивая число вершин в двойственном разбиении многоугольника Ньютона.

Мы начнём изучение того, как геометрия меняется от одного слоя к следующему в нейронной сети, более точнее:

Вопрос 1 *Как тропические гиперповерхности тропических многочленов в* $(l+1)$*-ом слое нейронной сети связаны с ими в* l*-ом слое? содержимое...*

Рекуррентное соотношение (2) описывает, как тропические многочлены, встречающиеся в $(l+1)$-ом слое получаются из многочленов в l-ом слое, а именно, через три операции: тропическая сумма, тропическую степень и тропическое умножение. Напомним, что тропическая гиперповерхность тропического многочлена — двойственно разбиение многогранника Ньютона тропического многочлена, который задается проекцией верхних граней на многогранники, определяемые формулой (1). Отсюда вопрос сводится к тому, как эти три операции преобразуют многогранники, а это рассматривается в утверждениях 3.1 и 3.2. Мы следуем обозначениям из Утверждения 5.1 для следующего результата.

Лемма 1 *Пусть* $f_i^{(l)}, g_i^{(l)}, h_i^{(l)}$ *тропические многочлены, созданные* i*-ым узлом в* l*-ом слое нейронной сети, то есть они определяются как (2). Тогда* $P(f_i^{(l)}, P(g_i^{(l)}, P(h_i^{(l)}$*, являющиеся подмножествами* $\mathbf{R}^{d+1}$*, задаются следующим образом:*

1. $P(g_i^{(1)\text{ и } P(h_i^{(1)}}$ *являются точками.*

2. $P(f_i^{(1)}$ — *отрезок.*

3. $P(g_i^{(1)\text{ и } P(h_i^{(1)}}$ — *зонотопы.*

4. *Для* $l \geq 1$,

$$P(h_i^{(l)}) = Conv[P(g_i^{(l)} \odot t_i^{(l)}) \cup P(h_i^{(l)})]$$

Если $t_i^{(l)} \in \mathbf{R}$, *и* $P(f_i^{(l)}) = P(h_i^{(l)})$, *если* $t_i^{(l)} = -\infty$

5. *Для* $l \geq 1, P(g_i^{(l+1)})$ *и* $P(h_i^{(l+1)})$ *взвешены суммы Минковского,*

$$P(g_i^{(l+1)}) = \sum_{j=1}^{n_l} a_{ij}^{-} P(f_i^{(l)}) + \sum_{j=1}^{n_l} a_{ij}^{+} P(g_i^{(l)}),$$

$$P(h_i^{(l+1)}) = \sum_{j=1}^{n_l} a_{ij}^{+} P(f_i^{(l)}) + \sum_{j=1}^{n_l} a_{ij}^{-} P(g_i^{(l)})$$
$$+\{b_i e\},$$

Где a_{ij}, b_i *записаны в матрице весов* $A^(l+1) \in \mathbf{Z}^{n_{l+1} \times n_l}$ *и вектор смещения* $b^(l+1) \in \mathbf{R}^{n_l+1}$ *и* $e := (0, \ldots, 0, 1) \in \mathbf{R}^{d+1}$.

Завершение леммы 6.2 состоит в том, что зонотопы являются строительными блоками в тропической геометрии нейронных сетей. Зонотопы широко изучены в выпуклой геометрии и, среди прочего, они тесно связаны с расположением гиперплоскостей. Лемма 6.2 связывает нейронные сети с этим обширным объёмом работы, но полный смысл этого еще предстоит изучить. В разделе C.2, кроме того, мы покажем, как можно построить эти многогранники для двуслойных нейронных сетей.

1.3 Геометрическая сложность глубоких нейронных сетей

Мы обращаемся к инструментам из раздела 3 для изучения сложности нейронной сети, показывая, что глубокая сеть более выразительна, чем неглубокая. Наша мера сложности является геометрической: мы будему следовать (Montufar et al., 2014; Raghu et al., 2017) и использовать количество линейных областей кусочно-линейной функции $\nu : \mathbf{R}^d \to \mathbf{R}^p$ для измерения сложности ν.

Теорема 1 *Пусть $\nu : \mathbf{R}^d \to \mathbf{R}$ является L-слоем вещественной нейронной сетью с прямой связъю, удовлетворяющей (a)-(c). Пусть $t^{(L)} = -\infty$ и $n_l \geq d$ для всех $l = 1, \ldots, L-1$. Тогда $\nu = \nu^{(L)}$ имеет максимум*

$$\prod_{l=1}^{L-1} \sum_{i=0}^{d} \binom{n_l}{i}$$

линейных областей. В частности, если $d \leq n_1, \ldots, n_{L-1} \leq n$, то число линейных областей ν ограничено $O(n^{d(L-1)})$.

Доказательство. Если $L = 2$, то это следует непосредственно из Леммы 6.2 и Следствия 3.4. Случай, когда $L \geq 3$ находится в разделе 7, в дополнении. ■

Как отмечалось в (Raghu et al., 2017), эта верхняя граница близко соответствует нижней границе $\Omega((\frac{n}{d}^{(l-1)d})n^d)$ в (Montufar et al., 2014, Corollary 5), когда $n_1 = \cdots = n_{L-1} = n \geq d$. отсюда мы предполагаем, что число линейных областей нейронной сети растёт полиномиально с шириной n и экспоненциально с количеством слоёв L.

1.4 Заключение

Мы утверждаем, что прямые нейронные сети с выпрямленными узлами не что иное, как тропические рациональные карты. Чтобы понять их, нам зачастую нужно понимать соответствующую тропическую геометрию.

В этой статье мы сделали первый шаг, чтобы предоставить подтверждение концепции: вопросы, касающиеся границ решений, линейных областей, как глубина влияет на выразительность и т.д. можно перевести на вопросы, касающиеся тропических гиперповерхностей, разбиений многоугольника Ньютона, многогранников, построенных из зонотопов и др.

Как новая ветвь алгебраической геометрии, новшество тропической геометрии происходит из алгебры и геометрии и их взаимодействует друг с другом. Она связана множеством других областей математики. Среди прочих вещей, существует тропический аналог линейной алгебры и тропический аналог выпуклой геометрии. Мы затронули лишь небольшую часть этого богатого предмета. Мы надеемся, что дальнейшее исследование под тропическим углом поможет разгадать другие загадки глубоких нейронных сетей.