1 On the power of threshold circuits with small weights

1.1 Что происходит

Рассматриваются схемы, состоящие из линейных пороговых элементов (linear threshold functions; LTF) вида $sgn(\sum w_i x_i + w_0),\ w_i \in Z, x_i \in \{-1, 1\}$; класс таких схем глубины d называем LT_d. Мотивация в изучении таких схем - нейронные сети, а именно многослойные перцептроны, которые фактически состоят из таких элементов. Ставится вопрос: можно ли в таких элементах ограничить веса полиномом от N (размера входного вектора), чтобы их было проще реализовывать. Схемы из таких элементов образуют классы $\widehat{LT_d}$.

1.2 Основные результаты

- Любой LTF может быть реализован схемой глубины 3, состоящей из элементов с ограниченным весом (т.е. $LT_1 \subset \widehat{LT_3}$).
- Обобщение: любую схему глубины d можно переделать в схему из элементов с ограниченным весом: $LT_d \subset \widehat{LT_{2d+1}}$
- Даны верхние оценки на реализацию некоторых арифметических функций:
  - $COMPARISON(X, Y) \in \widehat{LT_2}$
  - $SUM(X, Y) \in \widehat{LT_2}$
  - $MAXIMUM(X_1..X_m) \in \widehat{LT_4}$

1.3 Как доказывается

Доказываем в такой последовательности:

1. $COMPARISON \in \widehat{LT_2}$
   Пользуемся гармоническим анализом: представляем исходную функцию как $f(x) = \sum_\alpha a_\alpha X^\alpha, a_\alpha \in Q$, где α - всевозможные мономы из переменных $x_1..x_n$. Факт из гармонического анализа: если спектральная норма функции $||f(x)|| = \sum |a_\alpha|$ ограничена полиномом, то $f(x) \in \widehat{LT_2}$. Показываем, что, норма $COMPARISON$ растёт линейно с ростом n и, следовательно, $COMPARISON \in \widehat{LT_2}$.

2. $LT_1 \subset \widehat{LT_3}$
   Сводим вычисление произвольной взвешенной суммы под sgn к сумме двух чисел, которая, как показано выше, считается за 2 слоя. Можно показать, что сведение можно выполнить за 1 слой.

Дальше идут примеры:

1. $SUM(X, Y) \in \widehat{LT_2}$
   Замечаем, что сумма отличается от сравнения только знаком Y и, следовательно, лежит в том же классе.

2. $MAXIMUM(X_1..X_m) \in \widehat{LT_4}$
   Вводим вспомогательные переменные c_{ij}, которые обозначают результат сравнения чисел $i,\ j$. Все они считаются параллельно за 2 слоя. Затем для всех чисел требуем, чтобы оно было больше остальных. Взяв дизъюнкцию по всем числам, получаем итоговый максимум. Эти две операции реализуются конъюнкцией и дизъюнкцией, а они реализуются одним слоем каждая.

В конце доклада без доказательств приводилось обобщение: $LT_d \subset \widehat{LT_{2d+1}}$. Доказывается индукцией по глубине исходной схемы.