

О сокращении нелинейной глубины многомерных сверточных нейронных схем

А. В. Хапкин¹

В работе рассматриваются многомерные сверточные схемы в базисе Маккалока-Питтса. Показано, что рассматриваемые схемы могут быть реализованы схемой из априорной и динамической части, в которой вычисления в априорной части не зависят от входных данных. При этом априорная и динамическая части имеют нелинейную глубину, равную 2.

Ключевые слова: сверточная нейронная сеть, нейронная схема, нелинейная сложность, модель Маккалока-Питтса.

1. Введение.

Сверточная нейронная сеть - специальная архитектура нейронных сетей, предложенная Яном Лекуном в 1988 году [1]. Основная идея таких сетей - последовательное использование сверточных слоев и субдискретизирующих слоёв, что от слоя к слою позволяет переходить ко все более и более высоким абстракциям, при этом на последних слоях выделяются признаки очень высокого уровня. Сеть в процессе обучения самонастраивается и учится сама выбирать такие детали путем фильтра маловажных деталей и выделения существенных. Основными преимуществами сверточных нейронных сетей являются не очень большое количество настраиваемых параметров по сравнению с обычными полносвязными сетями, удобное распараллеливание вычислений, что позволяет использовать для вычисления графические процессоры, относительная устойчивость к повороту и сдвигу исходных данных. Сверточные сети активно применяются в задачах компьютерного зрения - распознавании образов, детекции объектов, также такие сети используются для распознавания речи, обработки аудиосигналов и анализа временных рядов.

Поэтому нужно уметь оптимизировать работу сверточных нейронных сетей, особенно важно уметь обходиться сетями с не очень большой нелинейной глубиной, так как в таком случае вытекает ряд преимуществ

¹Хапкин Артем Владимирович — аспирант каф. математической теории интеллектуальных систем мех.-мат. ф-та МГУ, e-mail: 9880817@mail.com.

Khapkin Artem Vladimirovich — graduate student, Lomonosov Moscow State University, Faculty of Mechanics and Mathematics, Chair of Mathematical Theory of Intellectual Systems.

- это позволяет ускорять время работы сетей, делать сети более интерпретируемыми.

В данной работе рассматриваются сверточные одномерные нейронные схемы. ([2], [3]) в модели Маккалока-Питтса [4]. Выполняется разложение этих сетей на динамическую и априорную части, при этом вычисления априорной части не зависят от входных данных. Минимизируется нелинейная глубина каждой из этих частей.

2. Основные понятия, определения и текущие результаты.

Пусть $X = \{x_1, x_2, \dots\}$ - алфавит входных переменных. Введем следующие функции:

1) постоянная функция $g_c \equiv c, c \in \mathbb{R}$

2) сумматор $\Sigma_n(x_1, \dots, x_n) = x_1 + x_2 + \dots + x_n, x_i \in \mathbb{R}$

3) усилитель $f_\gamma(x) = \gamma x, \gamma, x \in \mathbb{R}$

4) функция Хевисайда, $\theta(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$

5) функция $F(x_1, x_2) = \begin{cases} x_1, & x_2 \geq 0 \\ 0, & x_2 < 0 \end{cases}$

Множество таких функций обозначим за Δ' .

По аналогии с автоматными схемами [5] определим схемы из функциональных элементов.

Каждой функции $f(x_1, \dots, x_n) : \mathbb{R}^n \rightarrow \mathbb{R}$ можно сопоставить графический объект S с n входными стрелками и одной выходной стрелкой (коротко входы и выход S). Входам объекта S приписаны слева направо переменные x_1, x_2, \dots, x_n .

Элементы из Δ' изображаются следующим образом:

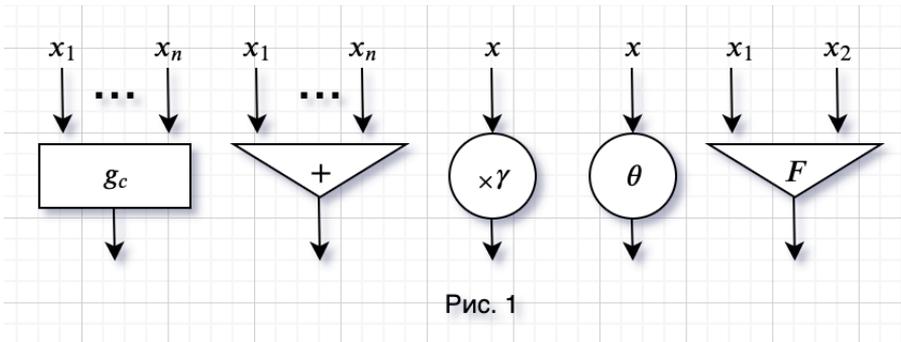


Рис. 1

Графический объект и соответствующая ему функция из Δ' называются функциональными элементами. Теперь индуктивно определим нейронные схемы без памяти. Функциональные элементы из 1-5 - нейронные схемы без памяти. Используя операции суперпозиции [6], а именно операции добавления фиктивного входа, изъятие фиктивного входа, склеивание ходов, переименование ходов без склеивания и последовательное соединение над функциональными элементами, индуктивно определяются нейронные схемы без памяти.

Если функция F реализуется нейронной схемой S , то в этом случае будем использовать запись (S, F) . При этом элементы θ и F называются нелинейными, остальные элементы называются линейными. Множество функций $\mathbb{R}^n \rightarrow \mathbb{R}$, реализуемых нейронными схемами без памяти, будем обозначать \mathbb{L} . В работе [6] показано, что множество \mathbb{L} совпадает с множеством кусочно-линейных функций PL .

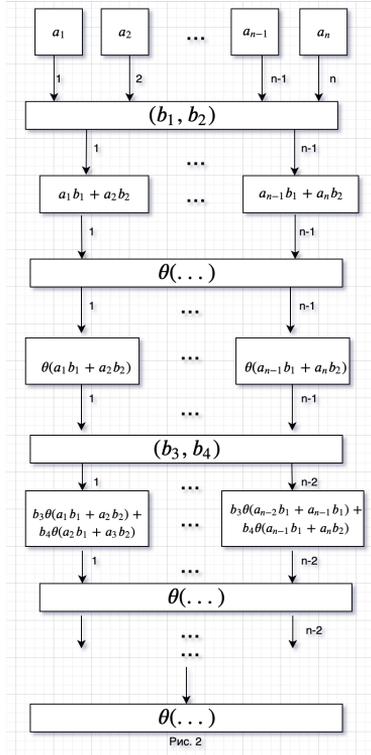
Число элементов в схеме называется сложностью нейронной схемы. Число нелинейных элементов в схеме называется нелинейной сложностью нейронной схемы.

Путем в нейронной схеме называется последовательность функциональных элементов

G_1, G_2, \dots, G_k , где вход G_1 является входом схемы, выход G_k является выходом схемы, и для любого $i, i = 1 \dots k - 1$ выход G_i является входом G_{i+1} . Нелинейной длиной пути называется число нелинейных элементов, содержащихся в нем. Нелинейной глубиной нейронной схемы называется длина самого длинного пути в ней.

Две нейронные схемы эквивалентны, если они реализуют одну и ту же функцию. В работе [6] показано, что для любой нейронной схемы без памяти существует эквивалентная ей схема с нелинейной глубиной 2.

Напомним определение одномерной сверточной схемы в базисе Маккалока-Питтса, приведенное в работе [7]. Имеется схема с n входами и одним выходом, в котором последовательно применяются операции свертки и Хевисайда. Приведем ее изображение:



Такую схему назовем одномерной сверточной схемой в базисе Маккалока-Питтса.

Нетрудно видеть, что приведенная схема на рис.2 имеет нелинейную глубину $n - 1$. Распишем подробнее, что происходит в процессе вычисления схемы на рис.2.

1. Выходы первого сверточного слоя нейронной схемы (которые передаются на вход второму слою):

$$(\theta(a_1b_1 + a_2b_2), \dots, \theta(a_{n-1}b_1 + a_nb_2)) \quad (1)$$

2. Выходы второго сверточного слоя нейронной схемы (которые передаются на вход третьему слою):

$$(\theta((b_3\theta(a_1b_1 + a_2b_2) + b_4\theta(a_2b_1 + a_3b_2)), \dots, \theta(b_3\theta(a_{n-2}b_1 + a_{n-1}b_2) + b_4\theta(a_{n-1}b_1 + a_nb_2))) \quad (2)$$

Далее идет третий сверточный слой с (b_5, b_6) , и так до конца до тех пор, пока у последнего сверточного слоя останется один выход.

Также напомним разложение, полученное в работе [7]. Если выразить результат вычисления схемы через $F_i^1, i = 1 \dots n - 1$, а конкретно то, что получается после первого сверточного слоя, то получится линейная комбинация слагаемых, при этом каждое слагаемое состоит из 2 множителей - первый множитель равен в F_i^1 или их всевозможных произведений, а второе слагаемое является комбинацией сверток, кроме первой. Следовательно, каждый такой множитель можно посчитать отдельно, в дальнейшем изобразим такие схемы.

Под схемой $\&$ подразумеваем получение всевозможных произведений (конъюнкций) из переданных в этот блок элементов, под схемой $+_{selected}$ подразумевается получение определенных линейных комбинаций. Теперь изобразим, как выглядит вид такой схемы с нелинейной глубиной 2, на основе приведенных выше преобразований:

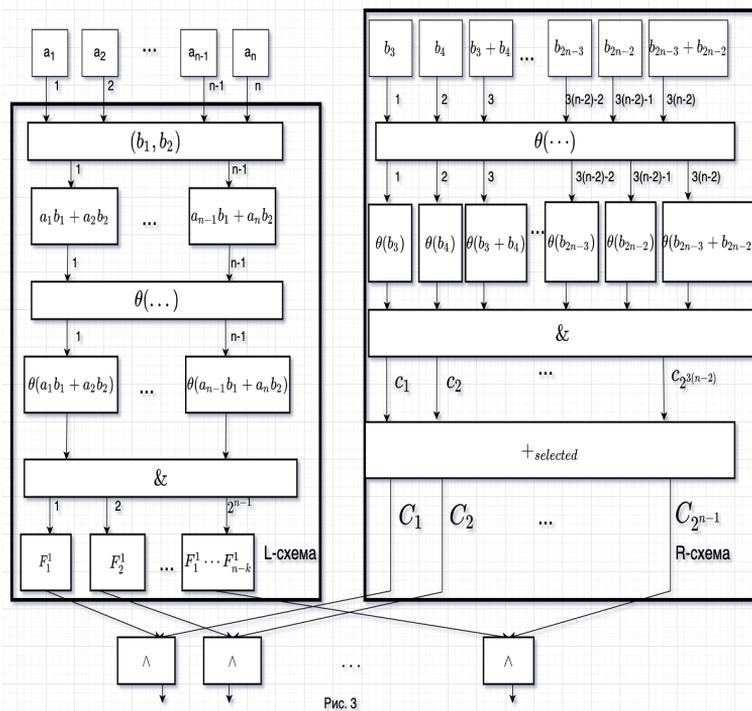


Рис. 3

Здесь c_i - различные произведения операций Хевисайда на множестве

$(b_3, b_4, b_3 + b_4, b_5, \dots, b_{2n-3}, b_{2n-2}, b_{2n-3} + b_{2n-2})$, а C_i - это определенные линейные комбинации таких операций Хевисайда, которые полу-

чаются при раскрытии нейронной схемы в явном виде. На выходах C_i схемы $+_{selected}$ реализуется конъюнкция с i подмножеством из функций $\{F_1, F_2, \dots, F_{n-1}\}$.

R -схемой назовем схему, реализующую комбинации сверток (априорная часть схемы, которая не зависит от исходных данных), а L -схемой назовем схему, реализующую преобразования с самими данными (динамическая часть схемы). LR -схемой назовем финальную комбинацию L и R -схемы, изображенную на рис.3.

По построению видно, нелинейная глубина R -схемы и L -схемы равна 2. Схема, изображенная на рис.3, имеет глубину 3, чтобы получить нелинейную глубину 2 после объединения R и L , нужно применить лемму 1.

Такой подход обладает интересным свойством - для новых данных нужно считать заново только преобразования из L -схемы, а преобразования R -схемы достаточно посчитать один раз, так как массивы сверток для разных массивов данных будет одинаков.

Также заметим, что полученную схему можно представить в другом виде - когда выходами L -схемы будут не F_i^1 и их всевозможные произведения, а F_i^k и их всевозможные произведения, где $1 < k < n$, а выходами R -схемы будут комбинации из меньшего числа сверток.

Имеет место следующая теорема, полученная в [7].

Теорема 1. *Любая одномерная сверточная сеть в базисе Маккалока-Питтса представляется в виде схемы нелинейной глубины 3, изображенной на рисунке 3.*

Также выпишем верхнюю оценку на нелинейную сложность, полученную в работе [7] для одномерных сверточных схем в базисе Маккалока-Питтса:

Теорема 2. *Порядок роста нелинейной сложности LR -схемы с нелинейной глубиной 2 не превосходит $n2^{3n}$. При этом нелинейная сложность L -схемы не превосходит $n2^n$.*

3. Рассмотрение двумерного случая.

Выше рассматривались только одномерные сверточные схемы, когда все операции изначально совершаются над одномерными массивами. Однако можно рассматривать и другие размерности, например размерность 2, когда изначально объектом выступают черно-белые изображения (матрицы). Более того, в практических задачах основную роль играют именно высокоразмерные объекты, а не одномерные.

Пусть имеются $n \times n$ матрица A ,

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

и последовательность из $n - 1$ матриц (назовем эти матрицы ядрами):

$$B_1 = \begin{pmatrix} b_{11} & b_{12} \\ b_{13} & b_{14} \end{pmatrix}, B_2 = \begin{pmatrix} b_{21} & b_{22} \\ b_{23} & b_{24} \end{pmatrix}, \dots, B_{n-1} = \begin{pmatrix} b_{n-1,1} & b_{n-1,2} \\ b_{n-1,3} & b_{n-1,4} \end{pmatrix}$$

По аналогии с одномерными сверточными схемами последовательно применяются операции свертки и Хевисайда, обобщим эти операции для двумерного случая. В этой схеме последовательно применяются операции свертки к каждому 2×2 матрице окну A (применяются 4 операции усилителя и затем одна операция сумматора между ними) и последующем применении функции Хевисайда к получившемуся результату. К полученному выходу повторяем такую же последовательность операций. Последовательное применение операций свертки и функции Хевисайда будем называть одним сверточным слоем нейронной схемы. Приведенная схема содержит $n - 1$ сверточных слоев.

Дадим более формальное определение. Для начала определим двумерную свертку. Пусть есть $n \times n$ матрица A , 2×2 матрица, и $n - 1 \times n - 1$ матрица C , элементы все матриц изображены на рисунке. Тогда операцией $conv B$ будем называть нейронную схему, на вход которой подается матрица A , а на выходе получается матрица C .

$$C = \begin{pmatrix} \theta(b_{11}a_{11} + b_{12}a_{12} + b_{13}a_{21} + b_{14}a_{22}) & \cdots & \theta(b_{11}a_{1,n-1} + b_{12}a_{1n} + b_{13}a_{2,n-1} + b_{14}a_{2n}) \\ \cdots & \cdots & \cdots \\ \theta(b_{11}a_{n-1,1} + b_{12}a_{n-1,2} + b_{13}a_{n,1} + b_{14}a_{n2}) & \cdots & \theta(b_{11}a_{n-1,n-1} + b_{12}a_{n-1,n} + b_{13}a_{n,n-1} + b_{14}a_{nn}) \end{pmatrix}$$

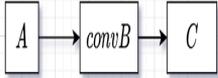
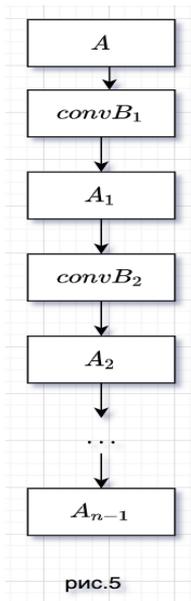
$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{13} & b_{14} \end{pmatrix} \quad A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{2,2} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$


Рис. 4

Теперь дадим определение двумерной сверточной схемы в базисе Маккалока-Питтса. Будем обозначать матрицами A_1, A_2, \dots, A_{n-1} матрицы, которые получаются на соответствующем слое. На последнем слое, на месте A_{n-1} , будет матрица размерности 1.



Двумерной сверточной сетью в базисе Маккалока-Питтса назовем сеть, изображенную на рисунке 5.

В результате таких операций размер соответствующей матрицы на каждом последующем слое будет уменьшаться на 1, и в финале у схемы получится только один выход. Полученная нелинейная глубина такой

схемы будет равна $n - 1$, так как количество блоков с функцией Хевисайда здесь (длина пути) равно $n - 1$.

Пример для $n=3$. Рассмотрим пример с небольшой размерностью при $n = 3$, получится ровно 2 сверточных слоя. Даны матрицы:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, B_1 = \begin{pmatrix} b_{11} & b_{12} \\ b_{13} & b_{14} \end{pmatrix}, B_2 = \begin{pmatrix} b_{21} & b_{22} \\ b_{23} & b_{24} \end{pmatrix}$$

Выходом первого сверточного слоя будет:

$$\begin{pmatrix} \theta(a_{11}b_{11} + \dots + a_{14}b_{14}) & \theta(a_{12}b_{11} + \dots + a_{23}b_{14}) \\ \theta(a_{21}b_{11} + \dots + a_{32}b_{14}) & \theta(a_{22}b_{11} + \dots + a_{33}b_{14}) \end{pmatrix}$$

Выходом второго сверточного слоя будет 1 элемент (матрица размерности 1×1):

$$\theta(b_{21}\theta(a_{11}b_{11} + \dots + a_{14}b_{14}) + b_{22}\theta(a_{12}b_{11} + \dots + a_{23}b_{14}) + b_{23}\theta(a_{21}b_{11} + \dots + a_{32}b_{14}) + b_{24}\theta(a_{22}b_{11} + \dots + a_{33}b_{14}))$$

Запишем результат такой схемы в другой форме. Пусть: $F_{11}^1 = \theta(a_{11}b_{11} + a_{12}b_{12} + a_{13}b_{13} + a_{14}b_{14})$
 $F_{12}^1 = \theta(a_{12}b_{11} + a_{13}b_{12} + a_{22}b_{13} + a_{23}b_{14})$
 $F_{21}^1 = \theta(a_{21}b_{11} + a_{22}b_{12} + a_{31}b_{13} + a_{32}b_{14})$
 $F_{22}^1 = \theta(a_{22}b_{11} + a_{23}b_{12} + a_{32}b_{13} + a_{33}b_{14})$

Обозначим этот финальный элемент как F_1^2 . Заметим, что F_1^2 можно воспринимать как функциональный элемент, входами которого являются $F_{11}^1, F_{12}^1, F_{21}^1, F_{22}^1$. И поэтому его можно записать как $F_1^2(F_{11}^1, F_{12}^1, F_{21}^1, F_{22}^1)$.

Действительно, в явном виде выпишем всевозможные варианты:

$$\begin{aligned}
0000 : 0 &= K_{0000} \\
0001 : \theta(b_{24}) &= K_{0001} \\
0010 : \theta(b_{23}) &= K_{0010} \\
0011 : \theta(b_{24} + b_{23}) &= K_{0011} \\
0100 : \theta(b_{22}) &= K_{0100} \\
0101 : \theta(b_{22} + b_{24}) &= K_{0101} \\
0110 : \theta(b_{22} + b_{23}) &= K_{0110} \\
0111 : \theta(b_{22} + b_{23} + b_{24}) &= K_{0111} \\
1000 : \theta(b_{21}) &= K_{1000} \\
1001 : \theta(b_{21} + b_{24}) &= K_{1001} \\
1010 : \theta(b_{21} + b_{23}) &= K_{1010} \\
1011 : \theta(b_{21} + b_{23} + b_{24}) &= K_{1011} \\
1100 : \theta(b_{21} + b_{22}) &= K_{1100} \\
1101 : \theta(b_{21} + b_{22} + b_{24}) &= K_{1101} \\
1110 : \theta(b_{21} + b_{22} + b_{23}) &= K_{1110} \\
1111 : \theta(b_{21} + b_{22} + b_{23} + b_{24}) &= K_{1111}
\end{aligned}$$

Таким образом, финальный выход возможно представить в виде следующей суммы:

$$\begin{aligned}
F_1^2 &= \theta(b_{24})(1 - F_{11}^1)(1 - F_{12}^1)(1 - F_{21}^1)F_{22}^1 + \\
&\theta(b_{23})(1 - F_{11}^1)(1 - F_{12}^1)F_{21}^1(1 - F_{22}^1) + \\
&\theta(b_{23} + b_{24})(1 - F_{11}^1)(1 - F_{12}^1)F_{21}^1F_{22}^1 + \dots
\end{aligned}$$

Случай произвольного n . Пользуясь аналогичными рассуждениями, разберем случай двумерной сверточной сети в базисе Маккалока-Питтса для произвольного n . В явном виде выпишем рекуррентные преобразования схемы, которые сократят нелинейную глубину до 2.

Входной матрицей размерности $(n \times n)$ будет:

$$\begin{pmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{31} & a_{32} & \cdots & a_{nn}
\end{pmatrix}$$

Сверточными матрицам будут:

$$B_1 = \begin{pmatrix} b_{11} & b_{12} \\ b_{13} & b_{14} \end{pmatrix}, B_2 = \begin{pmatrix} b_{21} & b_{22} \\ b_{23} & b_{24} \end{pmatrix}, \dots, B_{n-1} = \begin{pmatrix} b_{n-1,1} & b_{n-2,2} \\ b_{n-2,3} & b_{n-2,4} \end{pmatrix}$$

Выходом первого сверточного слоя будет матрица размерности $(n - 1 \times n - 1)$:

$$\begin{pmatrix} \theta(a_{11}b_{11} + \dots + a_{14}b_{14}) & \dots & \theta(a_{1,n-1}b_{11} \dots + a_{1n}b_{14}) \\ \vdots & \ddots & \vdots \\ \theta(a_{n-1,1}b_{11} + \dots + a_{n2}b_{14}) & \dots & \theta(a_{n-1,n-1}b_{11} \dots + a_{nn}b_{14}) \end{pmatrix}$$

Обозначим эти выходы за

$$\begin{pmatrix} F_{11}^1 & F_{12}^1 & \dots & F_{1,n-1}^1 \\ F_{21}^1 & F_{22}^1 & \dots & F_{2,n-1}^1 \\ \vdots & \vdots & \ddots & \vdots \\ F_{n-1,1}^1 & a_{n-1,2}^1 & \dots & F_{n-1,n-1}^1 \end{pmatrix}$$

Далее, по аналогии с одномерными сверточными схемами применяем операцию свертки и операцию Хевисайда, получаем

$$\begin{pmatrix} F_{11}^2 & F_{12}^2 & \dots & F_{1,n-2}^2 \\ F_{21}^2 & F_{22}^2 & \dots & F_{2,n-2}^2 \\ \vdots & \vdots & \ddots & \vdots \\ F_{n-2,1}^2 & F_{n-2,2}^2 & \dots & F_{n-2,n-2}^2 \end{pmatrix}$$

Далее последовательно применяются операции свертки и Хевисайда до тех пор, пока не достигнем последнего элемента. Матрицей 2×2 предпоследнего выхода слоя будет

$$\begin{pmatrix} F_{11}^{n-2} & F_{12}^{n-2} \\ F_{21}^{n-2} & F_{22}^{n-2} \end{pmatrix}$$

А сам финальный выход обозначим за F_1^{n-1} .

По аналогии с одномерным случаем заметим, что F_1^{n-1} явно определяется через предпоследний слой, по той же формуле, которая была использована в примере 3:

$$\begin{aligned} F_1^{n-1} &= \theta(b_{n-2,4})(1 - F_{11}^{n-2})(1 - F_{12}^{n-2})(1 - F_{21}^{n-2})F_{22}^{n-2} + \\ &\quad \dots \\ &+ \theta(b_{n-2,1} + b_{n-2,2} + b_{n-2,3} + b_{n-2,4})F_{11}^{n-2}F_{12}^{n-2}F_{21}^{n-2}F_{22}^{n-2} \end{aligned}$$

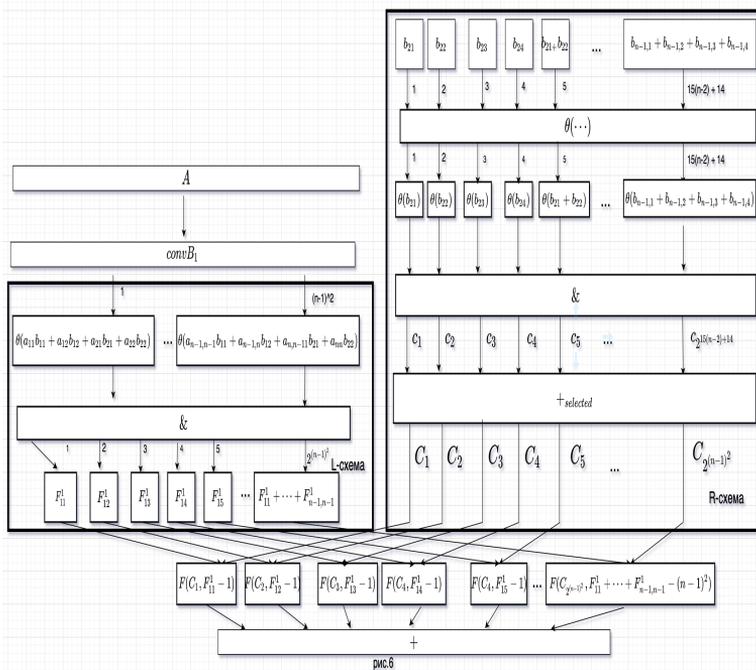
Также заметим, что по аналогии с одномерным случаем каждый из элементов $F_{11}^{n-2} F_{12}^{n-2}, F_{21}^{n-2}, F_{22}^{n-2}$ представляется в виде такой же формулы как комбинации элементов из слоя $n-3$. а те в свою очередь как комбинации элементов из слоя $n-4$, и так до тех пор, пока не будут достигнуты элементы из первого слоя, F_{ij}^k .

Это значит, что элемент на слое k на месте в матрице i, j однозначно представляется как комбинация элементов предыдущего слоя по следующей формуле:

$$F_{ij}^k = \theta(b_{k-1,4})(1 - F_{ij}^{k-1})(1 - F_{i,j+1}^{k-1})(1 - F_{i+1,j}^{k-1})F_{i+1,j+1}^{k-1} + \dots + \theta(b_{k-1,1}) + \theta(b_{k-1,2}) + \theta(b_{k-1,3}) + \theta(b_{k-1,4})F_{ij}^{k-1}F_{i,j+1}^{k-1}F_{i+1,j}^{k-1}F_{i+1,j+1}^{k-1} +$$

После того, как будет достигнут первый слой, применяем лемму 1 и получаем нелинейную глубину 2.

Таким образом, исходная схема представляется в виде схемы, изображенной на рисунке 6 (все обозначения на этом рисунке такие же, как и в одномерном случае):



Имеет место следующая теорема.

Теорема 3. Любая двумерная сверточная сеть в базисе Маккалока-Питтса представляется в виде схемы нелинейной глубины 3, изображенной на рисунке 6.

Заметим, что в конце, в отличие от одномерного случая, используются функции F и вместо всевозможных конъюнкций считаются всевозможные суммы. Нетрудно проверить, что такая схема эквивалентна схеме, когда считаются конъюнкции и дальше идет умножение число.

Действительно, рассмотрим к примеру самый последний элемент: $F_{11}^1 \cdots F_{n-1, n-1}^1$ принимает значение один, а следовательно, после домножения на часть из правой схемы, значение $C_{2^{(n-1)^2}}$ когда все элемента равны 1, и 0, если хотя бы один из них равен 0, значит достаточно, чтобы сумма $F_{11}^1 + F_{n-1, n-1}^1 - (n-1)^2$. Следовательно, $F(C_{2^{(n-1)^2}}, F_{11}^1 + \cdots + F_{n-1, n-1}^1 - (n-1)^2)$ реализует именно это.

Теперь оценим нелинейную сложность построения такой схемы. Напомним, что нелинейной сложностью схемы называется количество нелинейных элементов в такой схеме. Для начала оценим нелинейную сложность L-схемы. Заметим, что с использованием функции F из первоначального базиса получилось сильно соркатить нелинейную сложность - теперь достаточно посчитать значение каждого F_{ij}^1 на первом слое, а их всего $(n-1)^2$, так как сумматор не дает прироста к нелинейной сложности. Итого, нелинейная сложность L - не будет превосходить n^2 .

Теперь посмотрим на нелинейную сложность R-схемы. В ней у нас имеется не более 2^{15n} комбинаций с элементами, каждый из которых имеет длину, не более чем $k*n$, где k - это константа, которую в асимптотике можно отбросить. И так как линейные комбинации в такой схеме не будут давать прироста к нелинейной сложности, то сложность R-схемы не будет превышать $n2^{15n}$.

Видно, в чем основное отличие от одномерного случая - наибольший вклад в рост нелинейной сложности вносит L-схема из-за того, что количество различных элементов растет по порядку степени, а R-схема растет по линейному закону.

Теорема 4. Порядок роста нелинейной сложности LR-схемы с нелинейной глубиной 2 не превосходит $n^2 2^{15n}$ при $n \rightarrow \infty$.

4. Случай с произвольной размерностью k , где $k > 2$.

В современных задачах машинного обучения часто возникают ситуации, когда нужно работать с объектами большой размерностью - к примеру, цветные изображения (размерность 3) или видео (размерность 4). В общем случае такие объекты называются тензорами, их мы и будем рассматривать. Общий вид такой схемы изображен на рис. 5, при этом понятие свертки обобщается для случая с более высокой размерностью.

Пусть имеет тензор размерности k : $A = \{a_{i_1, i_2, \dots, i_k} : i_1 \in \{1, \dots, n\}, \dots, i_k \in \{1, \dots, n\}\}$, и точно также есть $n - 1$ тензоров:

$$B_1 = \{b_{i_1, \dots, i_k}^1 : i_1 \in \{1, 2, \dots\}, \dots, i_k \in \{1, 2, \dots\}\}$$

$$B_2 = \{b_{i_1, \dots, i_k}^2 : i_1 \in \{1, 2, \dots\}, \dots, i_k \in \{1, 2, \dots\}\}$$

...

$$B_{n-1} = \{b_{i_1, \dots, i_k}^{n-1} : i_1 \in \{1, 2, \dots\}, \dots, i_k \in \{1, 2, \dots\}\}$$

Далее к тензору A последовательно применяются сверточные операции с помощью тензоров B_i и операции Хевисайда до тех пор, пока не останется один выход финальной сверточной схемы.

Такую схему назовем k -мерной сверточной схемой в базисе Маккалока-Питтса.

4.1. Разложение такой схемы.

Заметим, что по построению k -мерная сверточная схема практически не отличается от построения одномерной и двумерной - добавляются лишь новые размерности.

Будем обозначать через F -нотацию выходы сверточных слоев. Например, за $F_{i_1, i_2, \dots, i_k}^j$ будет обозначаться элемент, который будет стоять на i_1, i_2, \dots, i_k месте на слое j .

Для примера, распишем, что будет после первого слоя (операция свертки и операция Хевисайда) на произвольном месте i_1, i_2, \dots, i_k :

$$F_{i_1, i_2, \dots, i_k}^1 = \theta \left(\sum_{i_1=1}^2 \sum_{i_2=1}^2 \dots \sum_{i_k=1}^2 a_{i_1 i_2, \dots, i_k} * b_{i_1, i_2, \dots, i_k}^1 \right) \quad (3)$$

По аналогии, получаем все остальные элементы выхода первого слоя, и далее к ним применяются новые операции свертки и операция Хевисайда, до тех пор, пока не останется только один выход у последнего слоя, его, как и ранее, обозначим за F_1^{n-1} .

Далее, по аналогии с двумерным случаем, точно также расписываются рекуррентные соотношения, которые позволяют через $F_{i_1, i_2, \dots, i_k}^1$ (выходы первого слоя) и последующие свертки, расписать, что получится у финального выхода, и схема получается такой же, как и на рисунке 6, с поправкой на размерность и на большее количество слагаемых в свертках.

4.2. Оценка сложности схем с размерностью $k > 2$

Также, по аналогии с двумерным случаем, нелинейная сложность полученной схемой будет определяться в основном R-схемой, и верна следующая теорема, которая является обобщением аналогичной теоремы для двумерного случая.

Теорема 5. *Порядок роста нелинейной сложности LR-схемы с нелинейной глубиной 2 не превосходит $n2^{(2^k-1)n}$ при $n \rightarrow \infty$. При этом основной вклад вносит R-схема, порядок роста сложности L-схемы не превосходит n^k .*

Автор выражает благодарность своему научному руководителю, доценту кафедры МаТИС механико-математического факультета МГУ, к.ф.-м.н., Часовских Анатолию Александровичу, за помощь в проведении исследования и постановке задачи.

Список литературы

- [1] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jacke, *Backpropagation applied to Handwritten Zip Code Recognition, Neural Computation*, 1989.
- [2] *Автоматы. Сборник статей. Под ред. К. Э. Шеннона и Дж. Маккарти. Пер. с англ. под ред. А. А. Ляпунова*, Изд. иностр. лит., 1956.
- [3] С. Хайкин, *Нейронные сети: полный курс, 2-е издание, Вильямс*, 2006.

- [4] W.S. McCulloch, W. Pitts, *A logical calculus of the ideas immanent in nervous activity*, **5**, 1943.
- [5] Кудрявцев В.Б., Алешин С.В., Подколзин А.С., *Введение в теорию автоматов*, «Наука», Москва, 1985, 320 pp.
- [6] Половников В.С., *Об оптимизации структурной реализации нейронных сетей*, МГУ, 2007
- [7] Хапкин А.В., *Об оптимизации структуры сверточных нейронных сетей*, МГУ, 2020

On reducing the nonlinear depth of multidimensional convolutional neural schemes

Khapkin A.V.

The paper considers multidimensional convolutional schemes in the McCulloch-Pitts basis. It is shown that the considered schemes can be implemented by a scheme from the a priori and dynamic parts, in which the calculations in the a priori part are independent of the input data. In this case, the a priori and dynamic parts have a nonlinear depth equal to 2.

Keywords: convolutional neural network, neural scheme, nonlinear complexity, McCulloch-Pitts model.

References

- [1] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jacke, *Backpropagation applied to Handwritten Zip Code Recognition*, *Neural Computation*, 1989
- [2] *Automatic machines. Digest of articles. Ed. C.E. Shannon and J. McCarthy. Tr. from English ed. A. A. Lyapunova*, Publishing foreign literature, 1956
- [3] S. Hykin, *Neural Networks: The Complete Course, 2nd Edition*, Williams, 2006
- [4] W.S. McCulloch, W. Pitts, *A logical calculus of the ideas immanent in nervous activity*, **5**, 1943
- [5] Kudryavtsev V.B., Aleshin S.V., Podkolzin A.S., *Introduction to automata theory*, «Наука», Moscow, 1985, 320 pp.

- [6] Polovnikov V.S., *On optimization of structured implementation of neural networks*, MSU, 2007
- [7] Khapkin A.V., *On optimizing the structure of convolutional neural networks*, MSU, 2020