УДК 621.391

Тензорно-матричные модели U-Net и генеративно-состязательных нейросетей

***В.И. Слюсар**, д.т.н., проф.*

Центральный научно-исследовательский институт
вооружения и военной техники Вооруженных Сил Украины

Рассмотренные в [1] варианты применения тензорно-матричной теории для формализации описания нейросетей касались преимущественно свёрточных нейронных сетей и их гиперансамблей [2], отличающихся последовательно сужающейся архитектурой слоёв. Между тем, нейросети не всегда являются системами, трансформирующими набор больших данных во множества меньшей размерности. Существуют архитектуры нейросетей, в которых используется расширяющийся сегмент, чья размерность входного слоя намного меньше размерности выхода. В качестве примера таких архитектур можно указать сети U-Net, применяемые для сегментации изображений, деконволюционные нейросети, а также генеративно-состязательные сети (GAN), синтезирующие новые изображения из набора исходных [3].

Рассмотрим варианты использования тензорно-матричного аппарата для формализации соответствующих нейронных структур. Традиционно используемый при этом подход предполагает использование кронекеровского произведения матриц. Однако его возможности сравнительно ограничены и характеризуются излишними вычислительными затратами. Более гибким в этом плане является семейство торцевых произведений матриц [4 - 7], когда размерность исходной матрицы может быть избирательно увеличена либо вдоль строк (торцевое произведение и его блочные модификации), либо вдоль столбцов (столбцевое произведение Хатри-Рао).

В качестве примера рассмотрим на входе расширяющегося сегмента нейросети матрицу признаков **A**. Ее торцевое произведение на матрицу коэффициентов **B** с тем же количеством строк позволяет получить матрицу большей размерности с увеличенным количеством столбцов. Для последующего приведения к требуемому виду получившейся в результате матрицы избыточного размера могут использоваться указанные в [1] операции умножения на вектор или вектор-строку единиц, матрицу меньшей размерности, а также процедуры фильтрации.

В случае гиперсетей данный подход может быть обобщён за счет использования блочных модификаций торцевых произведений [7], предложенных в [4 - 6]. В этом случае блоки блочных матриц позволяют описать функционирование различных сегментов гиперсети без взаимного влияния. Примечательно, что блочная версия торцевого произведения также имеет место при описании градиентов от торцевых произведений матриц [8], применение которых лежит в основе метода обратного распространения ошибок. Альтернативный вариант формализации моделей нейросетей состоит в применении проникающего прямого произведения [1]:

$$\mathbf{A}\left[\blacksquare\right]\mathbf{B}=\left[A_{ij}\blacksquare\,\mathbf{B}\right]=\left[A_{ij}\circ B_{mr}\right], \tag{1}$$

где $\blacksquare$ – символ проникающего торцевого произведения [5], $\circ$ – произведение Адамара.

Применительно к задаче генерации синтетических видеопотоков такой вариант матричного умножения позволяет получить поэлементное произведение каждого блока матрицы пикселей **A** на все блоки матрицы коэффициентов нейросети **B**:

$$\begin{bmatrix} A_{11} & \dots & A_{1T} \\ A_{21} & \dots & A_{2T} \\ \vdots & \vdots & \vdots \\ A_{P1} & \dots & A_{PT} \end{bmatrix}\left[\blacksquare\right]\begin{bmatrix} B_{11} & B_{12} & \dots & B_{1G} \\ B_{21} & B_{12} & \dots & B_{2G} \\ \vdots & \vdots & \vdots & \vdots \\ B_{P1} & B_{22} & \dots & B_{PG} \end{bmatrix}=\left[\begin{array}{c|c|c} A_{11}\blacksquare\begin{bmatrix} B_{11} & B_{12} & \dots & B_{1G} \\ B_{21} & B_{12} & \dots & B_{2G} \\ \vdots & \vdots & \vdots & \vdots \\ B_{P1} & B_{22} & \dots & B_{PG} \end{bmatrix} & \dots & A_{1T}\blacksquare\begin{bmatrix} B_{11} & B_{12} & \dots & B_{1G} \\ B_{21} & B_{12} & \dots & B_{2G} \\ \vdots & \vdots & \vdots & \vdots \\ B_{P1} & B_{22} & \dots & B_{PG} \end{bmatrix} \\ \vdots & \vdots & \vdots \\ A_{P1}\blacksquare\begin{bmatrix} B_{11} & B_{12} & \dots & B_{1G} \\ B_{21} & B_{12} & \dots & B_{2G} \\ \vdots & \vdots & \vdots & \vdots \\ B_{P1} & B_{22} & \dots & B_{PG} \end{bmatrix} & \dots & A_{PT}\blacksquare\begin{bmatrix} B_{11} & B_{12} & \dots & B_{1G} \\ B_{21} & B_{12} & \dots & B_{2G} \\ \vdots & \vdots & \vdots & \vdots \\ B_{P1} & B_{22} & \dots & B_{PG} \end{bmatrix} \end{array}\right].$$

В случае генерации последовательности кадров одного видеопотока **A** параллельно в нескольких расширяющихся сегментах нейросетей можно записать:

$$\left[A_{11}\dots A_{1T}\right]\left[\blacksquare\right]\begin{bmatrix} B_{11} & B_{12} & \dots & B_{1G} \\ B_{21} & B_{12} & \dots & B_{2G} \\ \vdots & \vdots & \vdots & \vdots \\ B_{P1} & B_{22} & \dots & B_{PG} \end{bmatrix} = \left[\begin{array}{c|c|c|c} A_{11}\blacksquare\begin{bmatrix} B_{11} & B_{12} & \dots & B_{1G} \\ B_{21} & B_{12} & \dots & B_{2G} \\ \vdots & \vdots & \vdots & \vdots \\ B_{P1} & B_{22} & \dots & B_{PG} \end{bmatrix} & \dots & A_{1T}\blacksquare\begin{bmatrix} B_{11} & B_{12} & \dots & B_{1G} \\ B_{21} & B_{12} & \dots & B_{2G} \\ \vdots & \vdots & \vdots & \vdots \\ B_{P1} & B_{22} & \dots & B_{PG} \end{bmatrix} \end{array}\right].$$

В данном выражении блок-строки матрицы **B** соответствуют коэффициентам одной нейросети в их ансамбле.

В более общем случае для построения модели генеративно-состязательной нейронной гиперсети следует воспользоваться блочной версией проникающего кронекеровского произведения [1]. Суть ее сводится к тому, что для двух матриц с одинаковым количеством блоков первого уровня, содержащих произвольное количество блоков второго уровня размерностью p×g каждый, результат соответствующего умножения имеет вид:

$$\mathbf{A}[[\blacksquare]]\mathbf{B} = \left[\mathbf{A}_{ij}\left[\blacksquare\right]\mathbf{B}_{ij}\right] = \left[\left[A_{bc}\circ B_{mr}\right]_{ij}\right], \qquad (2)$$

где *i, j* – индексы нумерации блоков второго уровня; *b, c* и *m, r* – индексы нумерации блоков первого уровня внутри *ij*-го блока второго уровня матрицы **A** и **B** соответственно.

Пример.

$$A = \left[\begin{array}{cc|cc} A_{111} & A_{121} & A_{112} & A_{122} \\ A_{211} & A_{221} & A_{212} & A_{222} \end{array}\right], \quad B = \left[\begin{array}{cc|cc} B_{111} & B_{121} & B_{112} & B_{122} \\ B_{211} & B_{221} & B_{212} & B_{222} \\ B_{311} & B_{321} & B_{312} & B_{322} \end{array}\right],$$

$$\mathbf{A}[[\blacksquare]]\mathbf{B} = \left[\begin{array}{cc|cc} A_{111}\blacksquare\begin{bmatrix} B_{111} & B_{121} \\ B_{211} & B_{221} \\ B_{311} & B_{321} \end{bmatrix} & A_{121}\blacksquare\begin{bmatrix} B_{111} & B_{121} \\ B_{211} & B_{221} \\ B_{311} & B_{321} \end{bmatrix} & A_{112}\blacksquare\begin{bmatrix} B_{112} & B_{122} \\ B_{212} & B_{222} \\ B_{312} & B_{322} \end{bmatrix} & A_{122}\blacksquare\begin{bmatrix} B_{112} & B_{122} \\ B_{212} & B_{222} \\ B_{312} & B_{322} \end{bmatrix} \\ A_{211}\blacksquare\begin{bmatrix} B_{111} & B_{121} \\ B_{211} & B_{221} \\ B_{311} & B_{321} \end{bmatrix} & A_{221}\blacksquare\begin{bmatrix} B_{111} & B_{121} \\ B_{211} & B_{221} \\ B_{311} & B_{321} \end{bmatrix} & A_{212}\blacksquare\begin{bmatrix} B_{112} & B_{122} \\ B_{212} & B_{222} \\ B_{312} & B_{322} \end{bmatrix} & A_{222}\blacksquare\begin{bmatrix} B_{112} & B_{122} \\ B_{212} & B_{222} \\ B_{312} & B_{322} \end{bmatrix} \end{array}\right]$$

В целом, произведение (2) позволяет формализовать модель совокупности входных слоёв нескольких расширяющихся сегментов нейронной гиперсети, синтезирующих, например, множество видеопотоков в различных спектральных диапазонах. Предложенный подход обеспечивает унификацию аналитического описания моделей нейросетей расширяющейся структуры с тензорно-матричными моделями [1], а также сокращает вычислительные затраты при их формализации по сравнению с применением кронекеровского произведения матриц.

Литература

1. *Слюсар В.И.* Модели нейросетей на основе тензорно-матричной теории. //«Проблемы разработки перспективных микро- и наноэлектронных систем» (МЭС-2021). – 2021. – № 2. – С. 23 – 28. DOI: 10.31114/2078-7707-2021-2-23-28.

2. *Tomer Galanti, Lior Wolf.* On the Modularity of Hypernetworks. // 34th Conference on Neural Information Processing Systems (NeurIPS 2020), Vancouver, Canada. – 2020. – 11 p.

3. Nahian Siddique, Sidike Paheding, Colin P. Elkin, and Vijay Devabhaktuni. U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications. // IEEE Access, Vol. 9, 2021. – Pp. 82031 – 82057. – DOI: 10.1109/ACCESS.2021.3086020.

4. *Слюсар В.И.* Торцевые произведения матриц в радиолокационных приложениях. // Известия высших учебных заведений. Радиоэлектроника. – 1998. – Том 41, № 3. – С. 71 – 75.

5. *Слюсар В.И.* Семейство торцевых произведений матриц и его свойства. // Кибернетика и системный анализ. – 1999. – Том 35; № 3. – С. 379 – 384. – DOI: 10.1007/BF02733426.

6. *Слюсар В.И.* Обобщенные торцевые произведения матриц в моделях цифровых антенных решеток с неидентичными каналами. // Известия высших учебных заведений. Радиоэлектроника. – 2003. – Том 46, № 10. – С. 15 – 26.

7. *Слюсар В.И.* Тензорно-матричная теория искусственного интеллекта. // Труды 63-й Всероссийской научной конференции МФТИ. 23 – 24 ноября 2020. Радиотехника и компьютерные технологии. – Москва: МФТИ. – 2020. – С. 104 – 106.

8. *Слюсар В.И.* Информационная матрица Фишера для моделей систем, базирующихся на торцевых произведениях матриц. // Кибернетика и системный анализ. – 1999. – Том 35; № 4. – С. 636 – 643. DOI: 10.1007/BF02835859.