TENSORES E A COMPLEXIDADE MULTIPLICATIVA

Erick Luz Aquino[1] (IC), Rick Antonio Rischter (PQ)[1]
[1]*Universidade Federal de Itajubá.*

Palavras-chave: Matrizes. multiplicação. posto.

Introdução

O estudo sobre a complexidade computacional (ou custo computacional) busca quantificar a complexidade de se efetuar certas operações computacionalmente. Por exemplo, sobre operações aritméticas, a complexidade multiplicativa quantifica a complexidade de se efetuar uma certa multiplicação computacionalmente. Em geral, a complexidade multiplicativa é muito maior do que a complexidade somativa, e por isso, muitos estudos vêm sendo feitos sobre a complexidade multiplicativa. Nesta pesquisa, buscamos estudar a complexidade multiplicativa de matrizes com tensores, tendo como base o livro texto [1] e os resultados dos artigos [2] e [3].

No livro texto [1], vemos que tensores são transformações multilineares vistas de uma forma geométrica, cuja definição busca generalizar a noção de escalares, vetores e matrizes.

No artigo [2], é apresentado um algoritmo para obter o resultado da multiplicação entre matrizes 2×2 mais eficaz computacionalmente, por calcular tal resultado com apenas sete multiplicações numéricas, enquanto o algoritmo usual utiliza oito.

No artigo [3], é mostrado que a cota $R(m, n, p)$ do posto máximo de tensores $m \times n \times p$ é igual à cota da complexidade multiplicativa máxima de uma família de p matrizes $m \times n$. O número dessa complexidade multiplicativa é a menor quantidade de matrizes de posto 1 necessárias para gerar qualquer família de p matrizes $m \times n$.

Ao fim desta pesquisa, obtivemos a cota das multiplicações matriciais com o algoritmo do artigo [2] (encontrando assim, uma cota para o posto do tensor da multiplicação matricial) e a comparamos com a cota das multiplicações do algoritmo usual. Também obtivemos uma cota para $R(m, n, p)$, num caso particular.

Metodologia

Nesta Iniciação Científica buscamos estudar os resultados do artigo [2] do matemático alemão Volker Strassen, publicado em 1969, e o artigo [3], dos matemáticos Atkinson e Stephens, publicado em 1978. Como ambos tratam da complexidade multiplicativa de matrizes, após estudarmos o livro texto [1] "Tensors: Geometry and Applications", de Landsberg, tentamos analisar ambos os artigos de um ponto de vista tensorial, obtendo assim alguns pequenos lemas e proposições.

Dados os espaços vetoriais $A_1, \cdots, A_n$, o posto de um tensor $T \in A_1 \otimes \cdots \otimes A_n$ é definido como sendo o menor número $r =: R(T)$ tal que

$$T = \sum_{u=1}^{r} Z_u,$$

onde os tensores $Z_u \in A_1 \otimes \cdots \otimes A_n$ são dados por $Z_u = a_{1u} \otimes \cdots \otimes a_{nu}$, tal que $a_{iu} \in A_i$, para todo $i = 1, \cdots, n$. Tensores como Z_u são ditos de posto 1.

O algoritmo de Strassen mostra que dadas A e B, matrizes 2×2 cujas entradas são a_{ij} e b_{ij}, respectivamente, as entradas c_{ij} da matriz $A \cdot B$ são dadas pelas constantes

$$I = (a_{11} + a_{22})(b_{11} + b_{22}),$$
$$II = (a_{21} + a_{22})b_{11},$$
$$III = a_{11}(b_{12} - b_{22}),$$
$$IV = a_{22}(-b_{11} + b_{21}),$$
$$V = (a_{11} + a_{12})b_{22},$$
$$VI = (-a_{11} + a_{21})(b_{11} + b_{12}),$$
$$VII = (a_{12} - a_{22})(b_{21} + b_{22}),$$

através das combinações lineares

$$c_{11} = I + IV - V + VII,$$
$$c_{12} = III + V,$$
$$c_{21} = II + IV,$$
$$c_{22} = I + III - II + VI.$$

Ao todo, utilizam-se 7 multiplicações numéricas e 18 adições numéricas (o algoritmo usual utiliza 8 multiplicações e 4 adições). Strassen também mostra um algoritmo para a inversão de matrizes. Ele também prova que

Teorema 1: O produto de duas matrizes de ordem n pode ser computado com no máximo $4,7n^{\log_2 7}$ operações aritméticas.

Tal resultado é relevante, levando em conta que o algoritmo usual utiliza n^3 multiplicações numéricas e $2n^2$ adições numéricas. Para provar o Teorema 1, Strassen não considera o uso exclusivo de seu algoritmo, mas também considera o uso do algoritmo usual, quando este é oportuno.
Denota-se por $R(m, n, p)$ a cota do posto máximo de tensores $m \times n \times p$ (chamados de paralelepípedos). $R(m, n, p)$ é igual à cota da complexidade multiplicativa máxima de uma família de p matrizes $m \times n$ (que é igual a menor quantidade de matrizes de posto 1 necessárias para gerar qualquer família de p matrizes $m \times n$). Tal informação é muito relevante, pois os dois principais resultados do artigo mostram que

Teorema 2: Se $m \leq n$, então $R(m, n, p) \leq m + [p/2]n$.

Teorema 3: Se $k \leq \min\{m, n\}$, então $r(m, n, mn - k) = mn - k^2 + R(k, k, k^2 - k)$.

O livro texto [1] apresenta a seguinte representação do algoritmo de Strassen em tensores:

$$\begin{aligned} M_{222} = {} & (A_{11} + A_{22}) \otimes (A_{11} + A_{22}) \otimes (C_{11} + C_{22}) \\ & + (A_{21} + A_{22}) \otimes A_{11} \otimes (C_{21} - C_{22}) \\ & + A_{11} \otimes (A_{12} - A_{22}) \otimes (C_{12} + C_{22}) \\ & + A_{22} \otimes (-A_{11} + B_{21}) \otimes (C_{21} + C_{11}) \\ & + (A_{11} + A_{12}) \otimes A_{22} \otimes (-C_{11} + C_{12}) \\ & + (-A_{11} + A_{21}) \otimes (A_{11} + A_{12}) \otimes C_{22} \\ & + (A_{12} - A_{22}) \otimes (A_{21} + A_{22}) \otimes C_{11}, \end{aligned}$$

onde M_{222} é a transformação bilinear que corresponde à multiplicação matricial 2×2, tal que $M_{222}(A, B) = A \cdot B$ (os elementos A_{ij},B_{ij} e C_{ij} são matrizes 2×2 cuja única entrada não-nula é a ij e seu valor numérico é 1). Note que $R(M_{222}) = 7$. Tal decomposição, e os resultados dos artigos aqui citados, abrem os horizontes para analisarmos os resultados sobre a complexidade multiplicativa de matrizes (de mesma ordem) através do posto de tensores.

Resultados e discussão

O algoritmo de multiplicação de Strassen pode ser aplicado para matrizes de ordem $n = m2^k$ (m é ímpar), através do completamento de zeros. Primeiramente, foi provado que

Proposição 1: O produto de duas matrizes de ordem n pode ser computado com no máximo $2,15n^{\log_2 7}$ multiplicações numéricas.

Observamos que para $n \geq 87$, com Strassen utiliza-se menos operações aritméticas (Figura 1) e para $n \geq 54$, com Strassen utiliza-se menos multiplicações (Figura 2), ambos em comparação com a utilização exclusiva do algoritmo usual.
Os tensores $m \times n \times p$ são chamados de paralelepípedos devido a sua representação geométrica. Concluímos que eles podem ser vistos como uma família de: p matrizes $m \times n$, ou de m matrizes $n \times p$ e ou n matrizes $m \times p$. Essas famílias de matrizes são chamadas de fatias desse paralelepípedo. Veja o exemplo da Figura 3.
Como $R(m, n, p)$ é igual a complexidade multiplicativa máxima de uma família de p matrizes $m \times n$, com a Proposição 1 podemos concluir, dada o formato tensorial de M_{222} na Seção Metodologia, que o posto de M_{nnn} (transformação bilinear que representa a multiplicação matricial $n \times n$) é menor ou igual a $2,15n^{\log_2 7}$, ou melhor, $R(M_{nnn}) \leq 2,15n^{\log_2 7}$.
O último resultado que registramos foi o

Proposição 2: Se $p = 3q$ ou $p = 3q + 1$, então $R(m, n, p) \leq m + [p/3](m + n)$ se $[n/2] < m \leq n$ e $R(m, n, p) \leq m + [p/3]3m$ se $m \leq [n/2]$.

A cota para $R(m, n, p)$ oferecida pela Proposição 2 não é tão relevante se comparada ao do Teorema 2 de [3], pois observamos que ela é um pouco menor em pouquíssimos casos. Nós provamos a Proposição 2 seguindo a ideia da prova do Teorema 2.

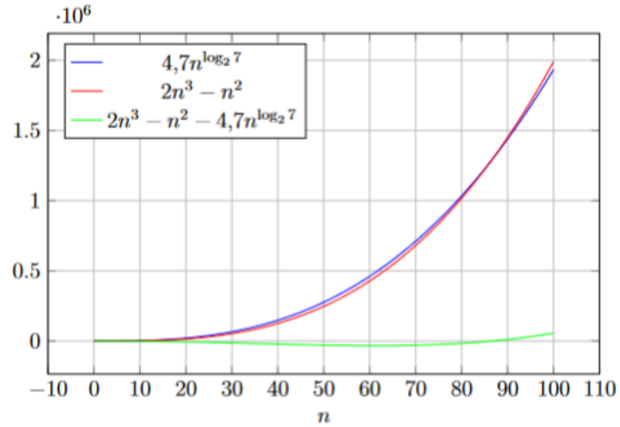

Figura 1 – Gráficos da cota das operações aritméticas com o algoritmo de Strassen (azul), da cota das operações aritméticas do algoritmo usual (vermelho) e da diferença entre estas cotas (verde).

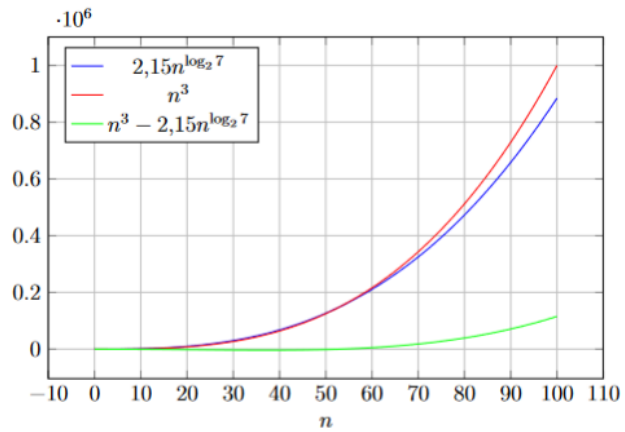

Figura 2 – Gráficos da cota das multiplicações numéricas com o algoritmo de Strassen (azul), da cota das multiplicações numéricas do algoritmo usual (vermelho) e da diferença entre estas cotas (verde).

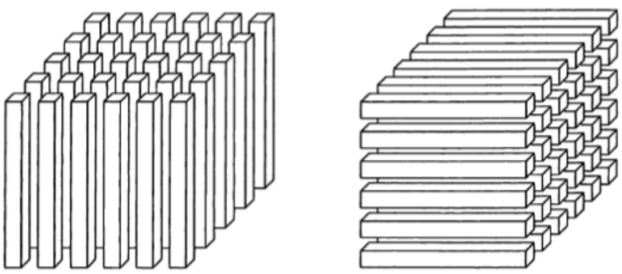

Figura 3 – Representação geométrica de um tensor $m \times n \times p$. São as colunas das fatias horizontais e as linhas das fatias verticais do paralelepípedo, respectivamente (da esquerda para a direita).

Conclusões

No livro texto [1] vimos que um tensor é a representação geométrica de transformações multilineares. Com os resultados do artigo [3], concluímos que a complexidade multiplicativa de matrizes pode ser estudada a partir do cálculo do posto de tensores $m \times n \times p$, e a partir disso, após termos determinado na Proposição 1 que o produto entre matrizes $n \times n$ pode ser calculado com menos de $2,15n^{log_2 7}$ multiplicações numéricas, concluímos que o posto do tensor M_{nnn} da multiplicação matricial $n \times n$, é menor ou igual a $2,15n^{log_2 7}$. Seguindo a ideia da prova do Teorema 2, concluímos na Proposição 2 que $R(m, n, p)$ é menor ou igual ao mínimo de $\{m + [p/3](m + n), m + [p/3]3m\}$, em certos casos.

Agradecimentos

À UNIFEI, por todo o suporte, à PIBIC UNIFEI, pela bolsa de Iniciação Científica, e ao professor Rick Antonio Rischter, por me ensinar.

Referências

[1] Joseph M. Landsberg. **Tensors: Geometry and Applications**. American Mathematical Society. Graduate Studies in Mathematics, vol. 128, ISBN 978-0-8218-6907-9, 2012.

[2] Strassen, Volker. **Gaussian Elimination is not Optimal**. University of Zurich. 1969, Numerische Mathematik, vol. 13, pages 354-356.M.

[3] D. Atkinson, N. M. Stephens. **On the Maximal Multiplicative Complexity of a Family of Bilinear Forms**. Department of Computing Mathematics, University College, Cardiff. March, 1978.