Lista 2 [TI8419 - Multilinear Algebra]

Lucas Abdalah

Professors: André Lima e Henrique Goulart

Table of Contents

Problem 1

By using the properties of the outer product, show that

$$\mathcal{X} = \mathbf{a}_1 \circ \mathbf{b}_1 \circ \mathbf{c}_1 + \mathbf{a}_2 \circ \mathbf{b}_2 \circ \mathbf{c}_2$$

is a rank-one tensor whenever $\mathbf{b}_1 = \mathbf{b}_2$ and $\mathbf{c}_1 = \mathbf{c}_2$. Is this also true in general when $\mathbf{c}_1 = \mathbf{c}_2$ but $\mathbf{b}_1 \neq \mathbf{b}_2$?

1.1) O tensor $\mathcal{X}$ é construido a partir do seguinte modelo:

$$\mathcal{X} = \mathbf{a}_1 \circ \mathbf{b}_1 \circ \mathbf{c}_1 + \mathbf{a}_2 \circ \mathbf{b}_2 \circ \mathbf{c}_2$$

Assumindo $\mathbf{b}_1 = \mathbf{b}_2$ e $\mathbf{c}_1 = \mathbf{c}_2$

$$\begin{aligned} \mathcal{X} &= \mathbf{a}_1 \circ \mathbf{b}_1 \circ \mathbf{c}_1 + \mathbf{a}_2 \circ \mathbf{b}_1 \circ \mathbf{c}_1 \\ &= (\mathbf{a}_1 + \mathbf{a}_2) \circ \mathbf{b}_1 \circ \mathbf{c}_1 \end{aligned}$$

1.2) É conveniente observar a soma como um novo vetor: $\mathbf{v}_1 = \mathbf{a}_1 + \mathbf{a}_2$

1.3) Finalmente, esta representação torna mais evidente que apenas 1 tensor de posto-1 é suficiente para representar $\mathcal{X}$, i.e, tem posto $R = 1$.

$$\boxed{\mathcal{X} = \mathbf{v}_1 \circ \mathbf{b}_1 \circ \mathbf{c}_1}$$

1.4) Assumindo $\mathbf{b}_1 \neq \mathbf{b}_2$ e $\mathbf{c}_1 = \mathbf{c}_2$

$$\begin{aligned} \mathcal{X} &= \mathbf{a}_1 \circ \mathbf{b}_1 \circ \mathbf{c}_1 + \mathbf{a}_2 \circ \mathbf{b}_2 \circ \mathbf{c}_1 \\ &= (\mathbf{a}_1 \circ \mathbf{b}_1 + \mathbf{a}_2 \circ \mathbf{b}_2) \circ \mathbf{c}_1 \end{aligned}$$

1.5) Com estas premissas, apenas a reorganização não permite afirmar que $\mathcal{X}$ é representado por um tensor de posto 1.

1.6) Entretanto, ao assumir que existe α tal que $\mathbf{b}_2 = \alpha \mathbf{b}_1$, a equação fica mais simples.

$$\begin{aligned} \mathcal{X} &= \mathbf{a}_1 \circ \mathbf{b}_1 \circ \mathbf{c}_1 + \mathbf{a}_2 \circ \alpha \mathbf{b}_1 \circ \mathbf{c}_1 \\ &= \mathbf{a}_1 \circ \mathbf{b}_1 \circ \mathbf{c}_1 + \alpha \mathbf{a}_2 \circ \mathbf{b}_1 \circ \mathbf{c}_1 \\ &= (\mathbf{a}_1 + \alpha \mathbf{a}_2) \circ \mathbf{b}_1 \circ \mathbf{c}_1 \end{aligned}$$

É conveniente observar a soma como um novo vetor: $\mathbf{v}_2 = \mathbf{a}_1 + \alpha \mathbf{a}_2$

1.7) Finalmente, esta representação torna mais evidente que apenas 1 tensor de posto-1 é suficiente para representar $\mathcal{X}$, i.e, tem posto $R = 1$.

$$\mathcal{X} = \mathbf{v}_2 \circ \mathbf{b}_1 \circ \mathbf{c}_1$$

1.8) Entretanto, se não existe α, ou seja $\mathbf{b}_1$ e $\mathbf{b}_2$ não são colineares, consequentemente $\mathbf{v}_2$ não existe e apenas 1 tensor de posto-1 é insuficiente para representar $\mathcal{X}$. Sendo utilizados 2 tensores de posto-1 para representação, implica em $posto(\mathcal{X}) = 2$.

$$\boxed{\mathcal{X} = (\mathbf{a}_1 \circ \mathbf{b}_1 + \mathbf{a}_2 \circ \mathbf{b}_2) \circ \mathbf{c}_1}$$

Problem 2

Show that the tensor rank is indeed a tensor property: in other words, it is invariant with respect to a multilinear transformation by nonsingular matrices, that is, if

$$\mathcal{X} = \mathcal{S} \times_1 \mathbf{A}^{(1)} \cdots \times_N \mathbf{A}^{(N)}$$

where $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times I_n}$ is nonsingular for every n, then

$$rank(\mathcal{X}) = rank(\mathcal{S}).$$

(Hint: write $\mathcal{S}$ as a PD with a minimal number of terms, and then use the properties of the multilinear transformation to bound the rank of $\mathcal{X}$; similarly, use the invertibility of the multilinear transformation to bound the rank of $\mathcal{S}$. More generally, conclude that the same property holds for matrices $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R_n}$ having linearly independent columns (and thus $R_n \leq I_n$).

2.1) O tensor *core* $\mathcal{S}$ reescrito em função de fatores da decomposição CP resulta em:

$$\mathcal{S} = \sum_{r=1}^{R} \mathbf{s}_r^{(1)} \circ \cdots \circ \mathbf{s}_r^{(N)}$$

2.2) Também é possível reorganizar a equação em função $\mathcal{S}$ de acordo com as equações (11) e (17) das notas de aula, utilizando as propriedades do operador transposto $\left({}^{\top}\right)$, caso $\mathbf{A}^{(n)} \in \mathbb{R}$, e operador hermitiano/autoadjunto $\left({}^{H}\right)$

$$\mathcal{S} = \mathcal{X} \times_1 \mathbf{A}^{(1)^H} \cdots \times_N \mathbf{A}^{(N)^H}$$

2.3) Desenvolvendo $\mathcal{X}$ em função de 2.1, obtém-se a representação em função do somatório ponderado pelas matrizes de mudança de base $\mathbf{A}^{(n)}$.

$$\mathcal{X} = \left(\sum_{r=1}^{R} \mathbf{s}_r^{(1)} \circ \cdots \circ \mathbf{s}_r^{(N)} \right) \times_1 \mathbf{A}^{(1)} \cdots \times_N \mathbf{A}^{(N)}$$

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{A}^{(1)} \mathbf{s}_r^{(1)} \circ \cdots \circ \mathbf{A}^{(N)} \mathbf{s}_r^{(N)}$$

2.4) Dado as considerações anteriores para $\mathcal{S}$, o tensor *core* $\mathcal{S}$ também pode ser reescrito como:

$$\mathcal{S} = \left(\sum_{r=1}^{R} \mathbf{A}^{(1)} \mathbf{s}_r^{(1)} \circ \cdots \circ \mathbf{A}^{(N)} \mathbf{s}_r^{(N)} \right) \times_1 \mathbf{A}^{(1)^H} \cdots \times_N \mathbf{A}^{(N)^H}$$

2.5) Dado as propriedas que relacionam os produtos de modo-N, o tensor $\mathcal{S}$, convenientemente é reorganizado de modo que as matrizes hermitianas multiplicam a matriz de transformação original de mesmo modo.

$$\mathcal{S} = \sum_{r=1}^{R} \mathbf{A}^{(1)^H} \mathbf{A}^{(1)} \mathbf{s}_r^{(1)} \circ \cdots \circ \mathbf{A}^{(N)^H} \mathbf{A}^{(N)} \mathbf{s}_r^{(N)}$$

2.6) Sendo $\mathbf{A}^{(n)^H} \mathbf{A}^{(n)} = \mathbf{I}$, obtemos:

$$\mathcal{S} = \sum_{r=1}^{R} \mathbf{s}_r^{(1)} \circ \cdots \circ \mathbf{s}_r^{(N)}$$

2.7) Consequentemente, $\mathcal{S}$ preserva o posto de $\mathcal{X}$, i.e, $\mathrm{rank}(\mathcal{X}) = \mathrm{rank}(\mathcal{S})$.

2.8) Para o primiro caso onde a matriz era quadrada, os operadores de inversa e hermitiano foram utilizdos. Já para o caso $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R_n}$, é necessário estender o conceito para pseudo-inversa de uma matriz. Assumindo a sua existência, obtém-se

$$\mathbf{A}^{(n)^\dagger} \mathbf{A}^{(n)} = \mathbf{I}$$

E isso permite a extensão da demonstração para matrizes retangulares.

Problem 3

Let $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ be given by

$$\mathcal{X} = \mathbf{a}_1 \circ \mathbf{b}_1 \circ \mathbf{c}_1 + \mathbf{a}_2 \circ \mathbf{b}_2 \circ \mathbf{c}_1 + \mathbf{a}_1 \circ \mathbf{b}_2 \circ \mathbf{c}_2$$

where the vectors are assumed to satisfy the following:

- a1 is not collinear with a2;
- b1 is not collinear with b2;
- c1 is not collinear with c2.

The goal of this exercise is to show that any such tensor has rank three, that is, it cannot be expressed as a sum of fewer terms. We will proceed by steps.

(i) First, show that

$$\mathcal{X} = \mathcal{S} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C}$$

where

$$\mathbf{A} = [\mathbf{a}_1 \quad \mathbf{a}_2], \quad \mathbf{B} = [\mathbf{b}_1 \quad \mathbf{b}_2], \quad \mathbf{C} = [\mathbf{c}_1 \quad \mathbf{c}_2],$$

and

$$\mathbf{S}_{..1} = \mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{S}_{..2} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Then, using the result of Exercise 2), conclude that $\mathcal{X}$ and $\mathcal{S}$ have the same rank.

(ii) Hence, it suffices to show that $\mathrm{rank}(\mathcal{S}) = 3$. Suppose, for a contradiction, that $\mathrm{rank}(\mathcal{S}) = 2$. Using the properties of the PARAFAC decomposition, show that this imples the existence of matrices $\mathbf{U}, \mathbf{V}, \mathbf{D}_1, \mathbf{D}_2 \in \mathbb{C}^{2\times 2}$ such that $\mathbf{D}_1, \mathbf{D}_2$ are diagonal and

$$\mathbf{S}_{..1} = \mathbf{U}\mathbf{D}_1\mathbf{V}^{\top}, \quad \mathbf{S}_{..2} = \mathbf{U}\mathbf{D}_2\mathbf{V}^{\top}$$

(iii) Now, use the fact that $\mathbf{X}_{..1} = \mathbf{I}$ to show that (2) implies that $\mathbf{S}_{..2}$ can be diagonalized by $\mathbf{U}$, that is, there exists a diagonal matrix $\mathbf{D} such that

$$\mathbf{S} = \mathbf{U}\mathbf{D}\mathbf{U}^{-1}.$$

(iv) Conclude that this leads to a contradiction, by taking into account the Jordan form of $\mathbf{S}_{..2}$.

3.1) *Unfoldings* do core do tensor $\mathcal{S}$

$$[\mathbf{S}]_{(1)} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$[\mathbf{S}]_{(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

$$[\mathbf{S}]_{(2)} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$[\mathbf{X}]_{(1)} = \mathbf{A}[\mathbf{S}]_{(1)}(\mathbf{C} \otimes \mathbf{B})^{\top}$$

$$[\mathbf{X}]_{(1)} = [\mathbf{a}_1 \quad \mathbf{a}_2] \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \left[(\mathbf{c}_1 \otimes \mathbf{b}_1)^{\top} \quad (\mathbf{c}_1 \otimes \mathbf{b}_2)^{\top} \quad (\mathbf{c}_2 \otimes \mathbf{b}_1)^{\top} \quad (\mathbf{c}_2 \otimes \mathbf{b}_2)^{\top} \right]$$

$$[\mathbf{X}]_{(1)} = [\mathbf{a}_1 \quad \mathbf{a}_2] \begin{bmatrix} 1(\mathbf{c}_1 \otimes \mathbf{b}_1)^{\top} + 0(\mathbf{c}_1 \otimes \mathbf{b}_2)^{\top} + 0(\mathbf{c}_2 \otimes \mathbf{b}_1)^{\top} + 1(\mathbf{c}_2 \otimes \mathbf{b}_2)^{\top} \\ 0(\mathbf{c}_1 \otimes \mathbf{b}_1)^{\top} + 1(\mathbf{c}_1 \otimes \mathbf{b}_2)^{\top} + 0(\mathbf{c}_2 \otimes \mathbf{b}_1)^{\top} + 0(\mathbf{c}_2 \otimes \mathbf{b}_2)^{\top} \end{bmatrix}$$

$$[\mathbf{X}]_{(1)} = [\mathbf{a}_1 \quad \mathbf{a}_2] \begin{bmatrix} (\mathbf{c}_1 \otimes \mathbf{b}_1)^{\top} + (\mathbf{c}_2 \otimes \mathbf{b}_2)^{\top} \\ (\mathbf{c}_1 \otimes \mathbf{b}_2)^{\top} \end{bmatrix}$$

$$[\mathbf{X}]_{(1)} = \mathbf{a}_1 \left[(\mathbf{c}_1 \otimes \mathbf{b}_1)^{\top} + (\mathbf{c}_2 \otimes \mathbf{b}_2)^{\top} \right] + \mathbf{a}_2 \left[(\mathbf{c}_1 \otimes \mathbf{b}_2)^{\top} \right]$$

$$[\mathbf{X}]_{(1)} = \mathbf{a}_1 (\mathbf{c}_1 \otimes \mathbf{b}_1)^{\top} + \mathbf{a}_1 (\mathbf{c}_2 \otimes \mathbf{b}_2)^{\top} + \mathbf{a}_2 (\mathbf{c}_1 \otimes \mathbf{b}_2)^{\top}$$

```
3.)
```

$$\mathcal{X} = \mathbf{a}_1 \circ \mathbf{b}_1 \circ \mathbf{c}_1 + \mathbf{a}_1 \circ \mathbf{b}_2 \circ \mathbf{c}_2 + \mathbf{a}_2 \circ \mathbf{b}_2 \circ \mathbf{c}_1$$

```
3.) Dado a demonstração do problema 2, o posto do tensor é uma propriedade.
```

$$\mathcal{S} = \sum_{r=1}^{2} \mathbf{s}_r^{(1)} \circ \mathbf{s}_r^{(2)} \circ \mathbf{s}_r^{(3)}$$

$$\mathcal{S}_{..1} = \sum_{r=1}^{2} \mathbf{s}_{1,r}^{(3)} \circ \mathbf{s}_r^{(1)} \circ \mathbf{s}_r^{(2)^{\top}} = \mathbf{S}^{(1)} \mathbf{D}_1 \left(\mathbf{S}^{(3)} \right) \mathbf{S}^{(2)^{\top}}$$

$$\mathcal{S}_{..2} = \sum_{r=1}^{2} \mathbf{s}_{2,r}^{(3)} \circ \mathbf{s}_r^{(1)} \circ \mathbf{s}_r^{(2)^{\top}} = \mathbf{S}^{(1)} \mathbf{D}_2 \left(\mathbf{S}^{(3)} \right) \mathbf{S}^{(2)^{\top}}$$

```
3.) Identidade e SVD
```

$$\mathcal{S}_{..1} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\top} = \mathbf{S}^{(1)} \mathbf{D}_1 \left(\mathbf{S}^{(3)} \right) \mathbf{S}^{(2)^{\top}} = \mathbf{I}$$

$$\mathbf{S}^{(1)} = \mathbf{D}_1 \left(\mathbf{S}^{(3)} \right) = \mathbf{S}^{(2)} = \mathbf{U} = \mathbf{\Sigma} = \mathbf{V}^{\top} = \mathbf{I}$$

```
3.) Rescrevendo $\mathcal{S}_{..2}$ a partir dos resultados obtidos em função de $\mathcal{S}_{..1}$.
```

$$\mathcal{S}_{..2} = \mathbf{S}^{(1)} \mathbf{D}_2 \left(\mathbf{S}^{(3)} \right) \mathbf{S}^{(2)^{\top}}$$

```
3.) Multiplicando
```

$$\mathcal{S}_{..2} = \mathbf{S}^{(1)} \left[\mathbf{S}^{(1)} \mathbf{D}_2 \left(\mathbf{S}^{(3)} \right) \mathbf{S}^{(2)^{\top}} \right] \mathbf{S}^{(1)^{-1}}$$

$$\mathcal{S}_{..2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{D}_2 \left(\mathbf{S}^{(3)} \right) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

```
3.) Dado que a identidade é o elemento neutro na multiplicação de matrizes, podemos reescrever a equação omitindo-as.
```

$$\mathcal{S}_{..2} = \mathbf{D}_2 \left(\mathbf{S}^{(3)} \right)$$

$$\mathcal{S}_{..2} = \begin{bmatrix} \mathbf{s}_{21} & 0 \\ 0 & \mathbf{s}_{22} \end{bmatrix}$$

3.) and thus it is possible to diagonalized the second frontal slice S...2 with matrix S(1). However, this similarity transformation leads to an invalid Jordan matrix as defined by S...2. Thus, it is not possible to the core tensor S be rank two and it must be in fact rank three

Problem 4

In this last exercise, we will show that, although the tensors of the form considered in the last exercise have rank 3, they are limits of sequences of rank-2 tensors. Thus, unlike happens for matrices, a sequence of rank-R tensors can converge to a rank-S tensor with $S > R$.

(i) First, show that the rank-1 tensor

$$\mathcal{Y}_m = m(\mathbf{a}_1 + m^{-1}\mathbf{b}_2) \circ (\mathbf{b}_2 + m^{-1}\mathbf{b}_1) \circ (\mathbf{c}_1 \circ m^{-1}\mathbf{c}_2)$$

is equal to $\mathcal{X}$ (as given by (1)) plus an $O(m)$ term $\mathcal{Z}_m$ and an $O(1/m)$ term.

(ii) Subtract the $O(m)$ term to get:

$$\mathcal{X}_m = \mathcal{Y}_m - \mathcal{Z}_m.$$

What is the rank of $\mathcal{X}_m$?

(iii) Use the expression obtained for $\mathcal{X}_m$ to conclude that $\lim_{m\to\infty} \mathcal{X}_m = \mathcal{X}$.

⋮

⋮