

**Note:** The tensor product is denoted by  $\otimes$ . In other words, for vectors  $\underline{a}, \underline{b}, \underline{c}$  we have that  $\underline{a} \otimes \underline{b}$  is the square array  $a^\alpha b^\beta$  where the superscript denotes the components, and  $\underline{a} \otimes \underline{b} \otimes \underline{c}$  is the cubic array  $a^\alpha b^\beta c^\gamma$ . We often denote components by superscripts because we need the lower index to label vectors themselves.

**Problem 1: Tensors**

1. Consider the tensor  $M = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ .

- (a) Write down the matrix components of  $M$ .
- (b) For the matrix  $M$  of part (a) exhibit an uncountable number of decompositions of the form  $M = \vec{a} \otimes \vec{b} + \vec{c} \otimes \vec{d}$  using the rotation matrices

$$R = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad \theta \in \mathbb{R}.$$

2. Consider the following tensor decomposition

$$T = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Is this decomposition unique? Justify your answer. What is the rank of  $T$ ?

3. Let  $\vec{a}_1, \vec{a}_2 \in \mathbb{R}^2$  be linearly independent and  $\vec{b}_1, \vec{b}_2 \in \mathbb{R}^2$  be linearly independent as well. We define  $T = \vec{a}_1 \otimes \vec{b}_1 \otimes \vec{c} + \vec{a}_2 \otimes \vec{b}_2 \otimes \vec{c}$  where  $\vec{c} \in \mathbb{R}^2$  is not the zero vector.
- (a) Does Jennrich's theorem apply?
  - (b) Prove that the tensor rank of  $T$  is 2.

**Problem 2: Kronecker and Khatri-Rao products**

The *Kronecker product*  $\otimes_{\text{Kro}}$  of two vectors  $\underline{a} \in \mathbb{R}^{I_1}$  and  $\underline{b} \in \mathbb{R}^{I_2}$  is a vectorization of the tensor (or outer) product. This amounts to take the  $I_1 \times I_2$  array  $a^\alpha b^\beta = (\underline{a} \otimes \underline{b})^{\alpha\beta}$  and view it as a vector of size  $I_1 I_2$ . More precisely, we define the Kronecker product as the column vector:

$$\underline{a} \otimes_{\text{Kro}} \underline{b} = [a^1 \underline{b}^T \quad \dots \quad a^{I_1} \underline{b}^T]^T \in \mathbb{R}^{I_1 I_2}.$$

Let  $A = [\underline{a}_1 \quad \dots \quad \underline{a}_R]$  and  $B = [\underline{b}_1 \quad \dots \quad \underline{b}_R]$  be matrices of dimensions  $I_1 \times R$  and  $I_2 \times R$ . We define the *Khatri-Rao* product as the  $I_1 I_2 \times R$  matrix

$$A \odot_{\text{KhR}} B = [\underline{a}_1 \otimes_{\text{Kro}} \underline{b}_1 \quad \dots \quad \underline{a}_R \otimes_{\text{Kro}} \underline{b}_R].$$

- 1) Assume that both  $A$  and  $B$  are full column rank. Prove that the Khatri-Rao product  $A \odot_{\text{KhR}} B$  is also full column rank.
- 2) Explain in detail in which step of Jennrich's algorithm this fact is used (see Figure ??).

**Problem 3: Jennrich's type algorithm for order 4 tensors**

Consider an order four tensor

$$T = \sum_{r=1}^R \underline{a}_r \otimes \underline{b}_r \otimes \underline{c}_r \otimes \underline{d}_r$$

where  $A = [\underline{a}_1 \ \cdots \ \underline{a}_R] \in \mathbb{R}^{I_1 \times R}$ ,  $B = [\underline{b}_1 \ \cdots \ \underline{b}_R] \in \mathbb{R}^{I_2 \times R}$ ,  $C = [\underline{c}_1 \ \cdots \ \underline{c}_R] \in \mathbb{R}^{I_3 \times R}$  and  $D = [\underline{d}_1 \ \cdots \ \underline{d}_R] \in \mathbb{R}^{I_4 \times R}$  are full column rank.

- 1) Check that you can apply Jennrich's algorithm to a "flattened" version of  $T$ , namely the order three tensor

$$\tilde{T} = \sum_{r=1}^R \underline{a}_r \otimes \underline{b}_r \otimes (\underline{c}_r \otimes_{\text{Kro}} \underline{d}_r).$$

where  $\otimes_{\text{Kro}}$  is the Kronecker product defined in the previous question.

- 2) Deduce that the rank  $R$  as well as the matrices  $A$ ,  $B$ ,  $C$ ,  $D$  can be uniquely determined from the four-dimensional array of numbers  $T^{\alpha\beta\gamma\delta}$  (up to trivial rank permutation and feature scaling).

4.1.1 *Jennrich's Algorithm.* If  $\mathbf{A}$ ,  $\mathbf{B}$ , and  $\mathbf{C}$  are all linearly independent (i.e. have full rank), then  $\mathcal{X} = \sum_{r=1}^R \lambda_r \mathbf{a}_r \odot \mathbf{b}_r \odot \mathbf{c}_r$  is unique up to trivial rank permutation and feature scaling and we can use Jennrich's algorithm to recover the factor matrices [23, 24]. The algorithm works as follows:

- (1) Choose random vectors  $\mathbf{x}$  and  $\mathbf{y}$ .
- (2) Take a slice through the tensor by hitting the tensor with the random vector  $\mathbf{x}$ :  

$$\mathcal{X}(\mathbf{I}, \mathbf{I}, \mathbf{x}) = \sum_{r=1}^R \langle \mathbf{c}_r, \mathbf{x} \rangle \mathbf{a}_r \odot \mathbf{b}_r = \mathbf{A} \text{Diag}(\langle \mathbf{c}_r, \mathbf{x} \rangle) \mathbf{B}^T.$$
- (3) Take a second slice through the tensor by hitting the tensor with the random vector  $\mathbf{y}$ :  

$$\mathcal{X}(\mathbf{I}, \mathbf{I}, \mathbf{y}) = \sum_{r=1}^R \langle \mathbf{c}_r, \mathbf{y} \rangle \mathbf{a}_r \odot \mathbf{b}_r = \mathbf{A} \text{Diag}(\langle \mathbf{c}_r, \mathbf{y} \rangle) \mathbf{B}^T.$$
- (4) Compute eigendecomposition to find  $\mathbf{A}$ :  

$$\mathcal{X}(\mathbf{I}, \mathbf{I}, \mathbf{x}) \mathcal{X}(\mathbf{I}, \mathbf{I}, \mathbf{y})^\dagger = \mathbf{A} \text{Diag}(\langle \mathbf{c}_r, \mathbf{x} \rangle) \text{Diag}(\langle \mathbf{c}_r, \mathbf{y} \rangle)^\dagger \mathbf{A}^\dagger$$
- (5) Compute eigendecomposition to find  $\mathbf{B}$ :  

$$\mathcal{X}(\mathbf{I}, \mathbf{I}, \mathbf{x})^\dagger \mathcal{X}(\mathbf{I}, \mathbf{I}, \mathbf{y}) = (\mathbf{B}^T)^\dagger \text{Diag}(\langle \mathbf{c}_r, \mathbf{x} \rangle)^\dagger \text{Diag}(\langle \mathbf{c}_r, \mathbf{y} \rangle) \mathbf{B}^T$$
- (6) Pair up the factors and solve a linear system to find  $\mathbf{C}$ .

Figure 1: Jennrich's algorithm (from *Introduction to Tensor Decompositions and their Applications in Machine Learning Review*, Rabanser, Shchur, Gunnemann)