Solutions to Homework 9: 10 May 2022
CS-526 Learning Theory

Problem 1

B is true. If $w_i(\epsilon) \neq 0$ for all i then the three arrays have rank K and there are K terms in the tensor decomposition. Therefore, by Jennrich's theorem, the decomposition is unique and the rank of the tensor $T(\epsilon)$ is K.

A is not true if there exists some (i, ϵ) such that $w_i(\epsilon)$ is zero.

C is not true because all the functions w_i are continuous. Therefore, $\lim_{\epsilon \to 0} T(\epsilon) = T(0)$ and, by Jennrich's theorem, the rank is at most K (the rank is K if $\forall i : w_i(0) \neq 0$).

D is not true because if $w_i(\epsilon) \neq 0$ for all i and $\epsilon \in [0, 1]$ then the rank is K.

Problem 2

1. Working with components, we have on one hand

$$T_2^{\alpha\beta} = \sum_{i=1}^{k} w_i a_i^\alpha a_i^\beta \ ,$$

and on the other hand

$$\begin{aligned}(V\mathrm{Diag}(w_1, \cdots, w_k)V^T)^{\alpha\beta} &= \sum_{i,j=1}^{n} V^{\alpha i} w_i \delta_{ij} (V^T)^{j\beta} = \sum_{i,j=1}^{n} V^{\alpha i} w_i \delta_{ij} V^{\beta j} \\ &= \sum_{i=1}^{n} V^{\alpha i} w_i V^{\beta i} = \sum_{i=1}^{n} \mathbf{a}_i^\alpha w_i \mathbf{a}_i^\beta \ .\end{aligned}$$

Exactly the same calculation applies to

$$T_3(I, I, \mathbf{u}) = \sum_{i=1}^{k} w_i (\mathbf{u}^T \mathbf{a}_i)(\mathbf{a}_i \otimes \mathbf{a}_i)$$

with w_i replaced by $w_i(\mathbf{u}^T \mathbf{a}_i)$. It remains to notice that

$$\mathrm{Diag}(w_1(\mathbf{u}^T\mathbf{a}_1), \cdots, w_k(\mathbf{u}^T\mathbf{a}_k)) = \mathrm{Diag}(w_1, \cdots, w_k)\mathrm{Diag}(\mathbf{u}^T\mathbf{a}_1, \cdots, \mathbf{u}^T\mathbf{a}_k) \ .$$

2. When $n = k$, since the vectors $\mathbf{a}_i$ are linearly independent, the matrix V is square and full rank so invertible. Since the w_i's are non-zero, it directly follows that T_2 is also invertible and

$$T_2^{-1} = (V^{-1})^T \,\mathrm{Diag}\left(\frac{1}{w_1}, \dots, \frac{1}{w_k}\right) V^{-1} \ .$$

3. a) First, note that:

$$\begin{aligned}
M &= T_3(I, I, \mathbf{u})T_2^{-1} \\
&= V\text{Diag}(w_1, \dots, w_k)\text{Diag}(\mathbf{u}^T\mathbf{a}_1, \dots, \mathbf{u}^T\mathbf{a}_k)V^T(V^{-1})^T\text{Diag}\left(\frac{1}{w_1}, \cdots, \frac{1}{w_k}\right)V^{-1} \\
&= V\text{Diag}(\mathbf{u}^T\mathbf{a}_1, \dots, \mathbf{u}^T\mathbf{a}_k)V^{-1}
\end{aligned}$$

Thus,

$$MV = V\text{Diag}(\mathbf{u}^T\mathbf{a}_1, \cdots, \mathbf{u}^T\mathbf{a}_k)$$

which is equivalent to

$$\forall i \in \{1, \dots, K\} : M\mathbf{a}_i = \lambda_i\mathbf{a}_i \quad \text{with} \quad \lambda_i = \mathbf{u}^T\mathbf{a}_i \ .$$

When $\mathbf{u}$ is taken at random from a continuous distribution, the inner products $\mathbf{a}_i^T\mathbf{u}$ are all distinct and nonzero with probability one. Indeed, the set of $\mathbf{u}$'s satisfying either $\mathbf{a}_i^T\mathbf{u} = \mathbf{a}_j^T\mathbf{u}$ for some $i \neq j$ or $\mathbf{a}_i^T\mathbf{u}$ for some i has measure zero. Therefore, we can uniquely determine the (normalized) eigenvectors $\mathbf{a}_i$'s by diagonalizing M.

b) Once we have recovered V, we can determine the w_i's from the diagonal matrix $V^{-1}T_2(V^{-1})^T$.

Problem 3

The first identity simply follows from the definitions:

$$(\mathbf{c} \otimes_{\text{Kro}} \mathbf{b})^T = \begin{bmatrix} c_1\mathbf{b}^T & c_2\mathbf{b}^T & \cdots & c_J\mathbf{b}^T \end{bmatrix} = \mathbf{c}^T \otimes_{\text{Kro}} \mathbf{b}^T \ .$$

For the second identity on the inner product between the two column vectors $\mathbf{e} \otimes_{\text{Kro}} \mathbf{d}$ and $\mathbf{c} \otimes_{\text{Kro}} \mathbf{b}$, we simply have:

$$(\mathbf{e} \otimes_{\text{Kro}} \mathbf{d})^T(\mathbf{c} \otimes_{\text{Kro}} \mathbf{b}) = \begin{bmatrix} e_1\mathbf{d}^T & e_2\mathbf{d}^T & \cdots & e_J\mathbf{d}^T \end{bmatrix} \begin{bmatrix} c_1\mathbf{b} \\ c_2\mathbf{b} \\ \vdots \\ c_J\mathbf{b} \end{bmatrix} = \sum_{j=1}^{J} e_jc_j\mathbf{d}^T\mathbf{b} = (\mathbf{e}^T\mathbf{c})(\mathbf{d}^T\mathbf{b}) \ .$$

Finally, the product of the $R \times IJ$ matrix $(E \otimes_{\text{Khr}} D)^T$ and the $IJ \times R$ matrix $(C \otimes_{\text{Khr}} B)$ is the $R \times R$ matrix whose entries are $\forall(i, j) \in \{1, \dots, R\}^2$:

$$\begin{aligned}
\left[(E \otimes_{\text{Khr}} D)^T(C \otimes_{\text{Khr}} B)\right]_{ij} &= \sum_{k=1}^{IJ} \left[E \otimes_{\text{Khr}} D\right]_{ki}\left[C \otimes_{\text{Khr}} B\right]_{kj} \\
&= (\mathbf{e}_i \otimes_{\text{Kro}} \mathbf{d}_i)(\mathbf{c}_j \otimes_{\text{Kro}} \mathbf{b}_j) \\
&= (\mathbf{e}_i^T\mathbf{c}_j)(\mathbf{d}_i^T\mathbf{b}_j) \\
&= [E^TC]_{ij}[D^TB]_{ij} \\
&= [(E^TC) \circ (D^TB)]_{ij} \ .
\end{aligned}$$

The third equality follows from the identity on the inner product of two Kronecker products. Hence $(E \otimes_{\text{Khr}} D)^T(C \otimes_{\text{Khr}} B) = (E^TC) \circ (D^TB)$.