ISYE 8803: High-Dimensional Data Analytics

Homework 3

Problem 1. Multilinear algebra

Suppose $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$. Please Calculate

a) $A \otimes B$
b) $A \odot B$
c) $(A \otimes B)^{-1}$
d) $(A^{-1} \otimes B^{-1})$
e) $A * B$

Problem 2. CP and Tucker decomposition

a. To minimize the least square problem

$$\{C, \boldsymbol{V}_1, \dots, \boldsymbol{V}_d\} = \mathrm{argmin}\left\| Y - \sum_{\mathrm{i}=1}^{\mathrm{p}} \mathrm{X_i B} \right\|$$

$$B_j = C_j \times_1 U_{j1} \times_2 U_{j2} \times_3 \dots \times_{lj} U_{jl_j} \times_{l_j+1} V_1 \times_{l_j+2} \dots \times_{l_j+d} V_d$$

$$V_i^T V_i = I_{\tilde{Q}_i}$$

Where $Y \in R^{Q_1 \times \dots \times Q_d}$ and $\mathrm{X_i} \in \mathrm{R}^{\mathrm{P_{i1}} \times \dots \times P_{il_i}}$ $C_j \in R^{\tilde{P}_1 \times \dots \times \tilde{P}_{lj} \times_{l_j+1} \tilde{Q}_1 \times_{l_j+2} \dots \times_{l_j+d} \tilde{Q}_d}$ is a core tensor with $\tilde{P}_{ji} \ll P_{ji}$ and $\tilde{Q}_i \ll Q_i$; $\{U_{ji}: j = 1, \dots, p; i = 1, \dots, l_j\}$ is a set of bases that spans the j^{th} input space; and $\{V_i: i = 1, \dots, d\}$ is a set of bases that spans the output space.

Prove the following theorem:

When U_{ji}, V_i and R_j are known, a reshape form of the core tensor C_j can be estimated as

$$\tilde{C}_j = R_j \times_1 \left(\boldsymbol{Z}_j^T \boldsymbol{Z}_j\right)^{-1} \boldsymbol{Z}_j^T \times_2 \boldsymbol{V}_1^T \times_3 \boldsymbol{V}_2^T \dots \times_{d+1} \boldsymbol{V}_d^T$$

Where $Z_j = X_{j(1)}\left(\boldsymbol{U}_{jl} \otimes \boldsymbol{U}_{jl-1} \otimes \dots \otimes \boldsymbol{U}_{j1}\right)$ and $R_j = Y - \sum_{i \neq j}^{p} B_j * X_j$. Note that $\tilde{C}_j$ has fewer modes $(d+1)$ than the original core tensor in (4), but it can be transformed into C by a simple reshape operation.

This can be done by the following steps:

1) Prove $\mathrm{argmin_C}\left\| R_{j(1)} - X_{j(1)} B_j \right\|_F^2 = \mathrm{argmin_C}\left\| vec\left(R_{j(1)}\right) - \left(V_d \otimes V_{d-1} \dots \otimes V_1 \otimes Z_j\right) vec\left(C_j\right) \right\|_F^2$ where $vec(X)$ stacks the columns of matrix X on top of each other. $Hint: \left(vec(ABC^T) = (C \otimes A) vec(B)\right)$
2) Prove $\mathrm{argmin_C}\left\| vec\left(R_{j(1)}\right) - \left(V_d \otimes V_{d-1} \dots \otimes V_1 \otimes Z_j\right) vec\left(C_j\right) \right\|_F^2$ has optimal solution:

$$vec\left(C_j\right) = \left(V_d^T \otimes V_{d-1}^T \dots . \otimes V_1^T \otimes \left(Z_j^T Z_j\right)^{-1} Z_j^T\right) vec\left(R_{j(1)}\right)$$

Hint $(A \otimes B)^{-1} = (A^{-1} \otimes B^{-1})$

b. Show that CP decomposition is the special case of tucker decomposition

Problem 3. Heat transfer process

Consider a heat transfer process that follows the following equations:

$$\frac{\partial S(x,y,t)}{\partial t} = \alpha\left(\frac{\partial^2 S}{\partial x^2} + \frac{\partial^2 S}{\partial y^2}\right)$$

Where $0 \leq x, y \leq 0.05$ represents the location of each image pixel, α is the thermal diffusivity coefficient, and t is the time frame. The initial and boundary conditions are set such that $S|_{t=1} = 0$ and $S|_{x=0} = S|_{x=0.05} = S|_{y=0} = S|_{y=0.05} = 1$. At each time t, the image is recorded at locations $x = \frac{j}{n+1}, y = \frac{k}{n+1}, j, k = 1, \ldots, n,$ resulting in an $n * n$ matrix. Here we set $n = 21$ and $t = 1, \ldots, 10$, which leads to 10 images of size 21*21, that can be represented as a $21 * 21 * 10$ tensor.

The thermal diffusivity coefficient depends on the material being heated. In the dataset heatT.mat, we have tensor 1 corresponding to a heat transfer process in material 1, and tensor 2 corresponding to a heat transfer process in material 2. Additionally, we have a third tensor. The goal of this problem is to determine to which material does tensor 3 correspond to.

In order to reach the goal, use CP and Tucker decomposition. Select the optimal rank for CP and tucker decomposition using AIC. Present your conclusion and the steps that lead to it. In particular, which decomposition method gives better result?