Exercise 7 – Low-rank approximation techniques

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Prof. Dr. D. Kressner
Dr. L. Periša

Problem 1. Let $\mathcal{X}$ and $\mathcal{Y}$ be $n_1 \times \cdots \times n_d$ tensors of tensor rank 1:

$$\mathcal{X} = x_1 \circ \cdots \circ x_d, \quad \mathcal{Y} = y_1 \circ \cdots \circ y_d.$$

Show that $\mathcal{X} + \mathcal{Y}$ is of rank 1 if and only if all but at most one of the components x_i and y_i are equal (within a scalar multiple). For the "only if" statement prove only the case $d = 2$ and $n_1 = n_2 = 2$.

Problem 2. What is the complexity of computing the inner product of two tensors of tensor rank at most R given in CP decompositions (with R terms)?

Problem 3. The aim of this exercise is to prove the upper bound from Slide 41 of Lecture 6:

$$\operatorname{rank}(\mathcal{X}) \leq \min\{R_2R_3, R_1R_3, R_1R_2\},$$

where (R_1, R_2, R_3) is the multilinear rank of $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$. Without loss of generality, the proof can be restricted to showing $\operatorname{rank}(\mathcal{X}) \leq R_2R_3$.

1. Show that $\operatorname{rank}(\mathcal{X}) \leq n_2n_3$. (Hint: One way to see this is to consider $X^{(1)}$ column by column.)
2. Establish a relation between the tensor rank of $\mathcal{X}$ and the tensor rank of the core tensor in a Tucker decomposition of $\mathcal{X}$.
3. Combine Points 1 and 2 to conclude the result.

Problem 4. Show that the following relations hold:

- $(A \odot B) \odot C = A \odot (B \odot C)$,
- $(A \odot B)^T (A \odot B) = A^T A * B^T B$,

where $*$ denotes the elementwise product, and one assumes that all involved matrix sizes are suitably chosen.

Problem 5. In Matlab, implement the Alternating least squares procedure from slide 55. Compare its convergence for a random tensor (of full generic rank) and a random tensor of exact tensor rank R (generated by $[\![A, B, C]\!]$ for random matrices A, B, C).

Problem 6. An $n \times \cdots \times n$ tensor $\mathcal{X}$ of order d is called symmetric if

$$\mathcal{X}_{i_1,\ldots,i_d} = \mathcal{X}_{\sigma(i_1),\ldots,\sigma(i_d)}$$

holds for every permutation $\sigma : \{1, \ldots, d\} \to \{1, \ldots, d\}$. What can you say about the multilinear rank of such a tensor?

Problem 7. Show that for a tensor $\mathcal{X}$ of order 3, U_1, U_2, U_3 columnwise orthogonal matrices and $\mathcal{C} = U_1^T \circ_1 U_2^T \circ_2 U_3^T \circ_3 \mathcal{X}$, the following equality holds

$$\|\mathcal{X} - U_1 \circ_1 U_2 \circ_2 U_3 \circ_3 \mathcal{C}\|^2 = \|\mathcal{X}\|^2 - \|U_1^T \circ_1 U_2^T \circ_2 U_3^T \circ_3 \mathcal{X}\|^2.$$

Problem 8.

1. In Matlab, implement the HOSVD algorithm from slide 60.
2. There is a more efficient variant, called the Sequentially Truncated HOSVD (STHOSVD), which proceeds as follows:
   (a) Calculate SVD $X^{(1)} = \widetilde{U}_1 \widetilde{\Sigma}_1 \widetilde{V}_1^T$. Truncate $U_1 := \widetilde{U}_1(:, 1 : r_1)$ and update $\mathcal{X} \leftarrow U_1^T \circ_1 \mathcal{X}$.
   (b) Calculate SVD $X^{(2)} = \widetilde{U}_2 \widetilde{\Sigma}_2 \widetilde{V}_2^T$. Truncate $U_2 := \widetilde{U}_2(:, 1 : r_2)$ and update $\mathcal{X} \leftarrow U_2^T \circ_2 \mathcal{X}$.
   (c) Calculate SVD $X^{(3)} = \widetilde{U}_3 \widetilde{\Sigma}_3 \widetilde{V}_3^T$. Truncate $U_3 := \widetilde{U}_3(:, 1 : r_3)$ and update $\mathcal{X} \leftarrow U_3^T \circ_3 \mathcal{X}$.
   (d) Set $\mathcal{C} = \mathcal{X}$.

Any order of truncation is possible. It is usually most efficient to proceed from the largest to the smallest mode size.

Implement STHOSVD in Matlab and compare the performance of both algorithms for random tensors of different sizes.

3. Adjust both algorithms to work with predescribed accuracy ε (discard involved singular values lower than ε). Test on function-related tensors, e.g. from Problem 7 of last week's exercises.

4. Bonus: Extend the results from Slides 61 and 62 on the error of HOSVD to STHOSVD.