250085 VU
TENSOR METHODS FOR DATA SCIENCE AND SCIENTIFIC COMPUTING
WINTER SEMESTER 2021

HOMEWORK ASSIGNMENT 3

VLADIMIR KAZEEV

Due by **13:00 on Monday, 15th November 2021**. To be submitted via Moodle.

The theoretical part may be (i) typeset, (ii) handwritten digitally or (iii) handwritten on a physical medium and then digitized by scanning or photographing. For the programming part, any programming language and environment may be used.

For each student submitting a solution, the submission should consist of (i) a single PDF file presenting the solution of the theoretical part and the results of the programming part in a self-contained fashion (so that running the student's code not be required for understanding the results) and (ii) the code (if any), ready to run and reproduce the results presented in the PDF file, organized in any reasonable number of files.

1. CPD approximation by ALS.

Implement the alternating-least-squares (ALS) algorithm for the low-rank approximation of a tensor of dimension $d \in \mathbb{N}$ at least two and mode sizes $n_1, \ldots, n_d \in \mathbb{N}$. For $R, r \in \mathbb{N}$ such that $r < R$, the implementation should take, as input parameters, a rank-R CPD of the tensor given in the form of matrices $U_k \in \mathbb{R}^{n_k \times R}$ with $k \in \{1, \ldots, d\}$ and a rank-r CPD of the initial guess given in the form of matrices $V_k \in \mathbb{R}^{n_k \times r}$ with $k \in \{1, \ldots, d\}$. The number of iterations to be performed should be an input parameter.

The best-accuracy rank-r least-squares approximation problem is based on the cost function given by

$$\phi(V_1, \ldots, V_d) = \|\Psi_r(V_1, \ldots, V_d) - \Psi_R(U_1, \ldots, U_d)\|_{\mathrm{F}}$$

for all $V_k \in \mathbb{R}^{n_k \times r}$ with $k \in \{1, \ldots, d\}$, where Ψ_r and Ψ_R denote the CPD multilinear representation maps transforming CP decompositions of ranks r and R into tensors $\mathbb{R}^{n_1 \times \cdots \times n_d}$.

In your implementation, a single ALS iteration starting at a CPD given by $V_k \in \mathbb{R}^{n_k \times r}$ with $k \in \{1, \ldots, d\}$ should consists in updating V_k for fixed

$$V_1, \ldots, V_{k-1}, V_{k+1}, \ldots, V_d$$

with a solution of the optimization problem

$$\phi(V_1, \ldots, V_d) \to \min_{V_k \in \mathbb{R}^{n_k \times r}}$$

sequentially for $k = 1, 2, \ldots, d-1, d, d-1, d-2, \ldots, 3, 2$.

An additional requirement is as follows: the cost of a single iteration should be linear with respect to d (certain elementwise products of Gram matrices accumulating from the left and from the right should be stored, and exactly one of these should be updated at each step of a single iteration).

Make sure that your implementation evaluate the above cost function ϕ and the 2-norm of $\nabla \frac{1}{2}\phi^2$ after each iteration and returns the history of computed values.

2. Testing the ALS algorithm on the matrix-multiplication tensor.

Consider the matrix-multiplication tensor for $n = 2$ and $r = 7$, $n = 3$ and $r = 23$, $n = 4$ and $r = 49$. For each $n \in \{2, 3, 4\}$, a CPD of rank $R = n^3$ is available (and was constructed in Assignment 2).

Test your implementation of the ALS algorithm on these three multiplication tensors with random initial guesses V_1, V_2, V_3, each matrix with all columns unitary (i.e., of norm one).

a) Start with seven runs (seven different initial guesses) for $n = 2$. In this case, the algorithm can be expected to achieve the accuracy of 10^{-8} within $10,000$ iterations in at least one of the seven runs. Plot the error ϕ and the 2-norm of $\nabla \frac{1}{2}\phi^2$ at the end of each iteration with respect to the iteration count for several runs.

b) Try repeating the experiment with $n = 3$ and $n = 4$. For each n, plot the error ϕ and the 2-norm of $\nabla \frac{1}{2}\phi^2$ at the end of each iteration with respect to the iteration count for several runs.

c) From a rank-7 CPD of the tensor corresponding to $n = 2$, derive (analytically) a rank-49 CPD of the tensor corresponding to $n = 4$. Implement the construction of this CPD.

Consider random perturbations of this CPD obtained by adding random perturbations to the entries of the CPD matrices drawn from the uniform distribution on $[-\varepsilon, \varepsilon]$. For $\varepsilon = 10^{-1}, \varepsilon = 10^{-2}, \dots, \varepsilon = 10^{-10}$, study the behavior of the ALS algorithm with the perturbed tensor used as an initial guess. For each ε, as above, plot the error ϕ and the 2-norm of $\nabla \frac{1}{2}\phi^2$ at the end of each iteration.