TIP8419 - Tensor Algebra — PPGETI/UFC
Exercise list n° 1:
Multilinear transformations and multilinear rank

Semester: 2022-1

1) The multilinear transformation

$$\mathcal{Y} = \mathcal{X} \times_1 \mathbf{A}^{(1)} \cdots \times_N \mathbf{A}^{(N)}$$

corresponds to a change of basis in the tensor space of $\mathcal{X}$ whenever the matrices $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times I_n}$ are nonsingular. It is legitimate to say that a property P of $\mathcal{X}$ is a property of the *tensor* $\mathcal{X}$, and not merely of the *coordinate array* $\mathcal{X}$, whenever P is invariant under a change of basis. Tensors are therefore objects with a well-defined geometric meaning which is independent of the choice of coordinates, and thus are characterized by their invariants with respect to multilinear transformations by nonsingular matrices.

Show that the multilinear rank of $\mathcal{X}$ is indeed a tensor property, that is, it is invariant with respect to a multilinear transformation by nonsingular matrices.

2) Recall that a matrix $\mathbf{P} \in \mathbb{C}^{I \times I}$ corresponds to a projection onto a subspace of $\mathbb{C}^I$ if and only if it is idempotent and self-adjoint:

$$\mathbf{P}^2 = \mathbf{P}, \qquad \mathbf{P}^{\mathsf{H}} = \mathbf{P}.$$

The projection of a vetor $\mathbf{x}$ onto that subspace is given by the matrix-vector product $\mathbf{y} = \mathbf{P}\mathbf{x}$. This leads to the properties

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{C}^I, \qquad \mathbf{P}(\mathbf{P}\mathbf{x}) = \mathbf{P}\mathbf{x}, \qquad \langle \mathbf{P}\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{P}\mathbf{y} \rangle. \tag{1}$$

When multiplied by a projection matrix $\mathbf{P}^{(n)} \in \mathbb{C}^{I_n \times I_n}$ in the nth mode, a tensor can be also be projected onto a subspace of the original tensor space:

$$\mathcal{Y} = \mathcal{X} \times_n \mathbf{P}^{(n)}.$$

Since

$$[\mathcal{Y}]_{(n)} = \mathbf{P}^{(n)} [\mathcal{X}]_{(n)},$$

it can be seen that the colums of $[\mathcal{Y}]_{(n)}$, that is, the nth-mode fibers of $\mathcal{Y}$, are given by the projection of the respective colums (fibers) of $[\mathcal{X}]_{(n)}$ (of $\mathcal{X}$) onto the subspace associated with $\mathbf{P}^{(n)}$. This operation can be generalized to comprise projections in all modes:

$$\mathcal{Y} = \mathcal{X} \times_1 \mathbf{P}^{(1)} \cdots \times_N \mathbf{P}^{(N)}.$$

In fact, performing a best approximation of multilinear rank $(R_1, \ldots, R_n)$ of a tensor $\mathcal{X}$ corresponds to projecting $\mathcal{X}$ onto a tensor subspace, according to N per-mode projections, in such a way that the nth-mode projection matrix $\mathbf{P}^{(n)}$ spans an R_n-dimensional subspace of $\mathbb{C}^{I_n}$.

(a) Show that $\mathcal{Y} = \mathcal{X} \times_n \mathbf{P}^{(n)}$ is indeed a projection, in the sense that, similarly to (1),

$$(\mathcal{X} \times_n \mathbf{P}^{(n)}) \times_n \mathbf{P}^{(n)} = \mathcal{X} \times_n \mathbf{P}^{(n)}$$

and

$$\left\langle \mathcal{X} \times_n \mathbf{P}^{(n)}, \mathcal{Y} \right\rangle = \left\langle \mathcal{X}, \mathcal{Y} \times_n \mathbf{P}^{(n)} \right\rangle.$$

(b) Show that, for any projection matrix $\mathbf{P}^{(n)}$,

$$\left\| \mathcal{X} \times_n \mathbf{P}^{(n)} \right\|_{\mathsf{F}} \leqslant \|\mathcal{X}\|_{\mathsf{F}} ,$$

that is, the nth-mode projection is a *non-expansive* operation. (Hint: decompose $\mathcal{X}$ as a sum of projections onto the subspace associated with $\mathbf{P}^{(n)}$ and onto the orthogonal complement of that subspace, and then develop the squared norm of the obtained expression).

3) Let $(R_1, \ldots, R_N)$ be the multilinear rank of $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$. Show that

$$\forall n \in \{1, \ldots, N\}, \quad R_n \leqslant \min \left\{ I_n, \prod_{m \neq n} I_m \right\}.$$

4)

(a) Show that the multilinear transformation is linear in the transformed tensor, and also in each involved matrix separately:

$$\begin{aligned}(\alpha_1 \mathcal{X}_1 + \alpha_2 \mathcal{X}_2) \times_1 \mathbf{A}^{(1)} \cdots \times_N \mathbf{A}^{(N)} = {} & \alpha_1 \left(\mathcal{X}_1 \times_1 \mathbf{A}^{(1)} \cdots \times_N \mathbf{A}^{(N)} \right) \\ & + \alpha_2 \left(\mathcal{X}_2 \times_1 \mathbf{A}^{(1)} \cdots \times_N \mathbf{A}^{(N)} \right)\end{aligned}$$

and

$$\begin{aligned}\mathcal{X} \times_1 \mathbf{A}^{(1)} \cdots \times_n (\alpha \mathbf{A}^{(n)} + \beta \mathbf{B}^{(n)}) \cdots \times_N \mathbf{A}^{(N)} = {} & \\ \alpha \left(\mathcal{X} \times_1 \mathbf{A}^{(1)} \cdots \times_n \mathbf{A}^{(n)} \cdots \times_N \mathbf{A}^{(N)} \right) & \\ + \beta \left(\mathcal{X} \times_1 \mathbf{A}^{(1)} \cdots \times_n \mathbf{B}^{(n)} \cdots \times_N \mathbf{A}^{(N)} \right). &\end{aligned}$$

(b) Let

$$\mathcal{X} = \mathbf{u}_1 \circ \mathbf{v}_1 \circ \mathbf{w}_1 + \mathbf{u}_2 \circ \mathbf{v}_2 \circ \mathbf{w}_2.$$

Use the result of (i) to obtain an expression for

$$\mathcal{X} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C}$$

in terms of the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and of the vectors $\mathbf{u}_i$, $\mathbf{v}_i$ and $\mathbf{w}_i$. (Hint: write the scalar components of that tensor and apply the definition of the nth-mode product.)

5) We have seen along the course that an approximation of multilinear rank $(R_1, \ldots, R_N)$ of an Nth-order tensor $\mathcal{X}$ can be obtained by truncating the HOSVD of $\mathcal{X}$, that is, by computing

$$\hat{\mathcal{X}} = \mathcal{S} \bigtimes_{n=1}^{N} \mathbf{U}^{(n)},$$

where $\mathbf{U}^{(n)} \in \mathbb{C}^{I_n \times R_n}$ contains the first R_n left singular vectors of $[\mathcal{X}]_{(n)}$ and

$$\mathcal{S} = \mathcal{X} \bigtimes_{n=1}^{N} (\mathbf{U}^{(n)})^{\mathsf{H}}.$$

The goal of this exercise is to show that such an approximation is *quasi-optimal*, in the sense that

$$\|\mathcal{X} - \hat{\mathcal{X}}\|_{\mathsf{F}}^2 \leqslant N \|\mathcal{X} - \mathcal{X}^\star\|_{\mathsf{F}}^2, \tag{2}$$

where $\mathcal{X}^\star$ is a solution of the best multilinear-rank-$(R_1, \ldots, R_N)$ approximation of $\mathcal{X}$:

$$\mathcal{X}^\star \in \underset{\mathcal{U} \in M_{(R_1,\ldots,R_N)}}{\arg\min} \ \|\mathcal{X} - \mathcal{U}\|_{\mathsf{F}}^2,$$

where $M_{(R_1,\ldots,R_N)}$ is the set of tensors having multilinear rank bounded (entry-wise) by $(R_1, \ldots, R_N)$. For simplicity, we will take $N = 3$.

(a) First show that, for any orthogonal projector $\mathbf{P}$ and any mode n, we have

$$\langle \mathcal{X} \times_n \mathbf{P}, \mathcal{X} \times_n (\mathbf{I} - \mathbf{P}) \rangle = 0.$$

Conclude that

$$\|\mathcal{X} \times_n \mathbf{P} - \mathcal{X} \times_n (\mathbf{I} - \mathbf{P})\|_{\mathsf{F}}^2 = \|\mathcal{X} \times_n \mathbf{P}\|_{\mathsf{F}}^2 + \|\mathcal{X} \times_n (\mathbf{I} - \mathbf{P})\|_{\mathsf{F}}^2.$$

(b) Now, let $\mathbf{P}^{(n)} := \mathbf{U}^{(n)}(\mathbf{U}^{(n)})^{\mathsf{H}}$ be an orthogonal projector onto the subspace spanned by the dominating (i.e., the first) R_n singular vectors of $[\mathcal{X}]_{(n)}$. Write $\hat{\mathcal{X}}$ in terms of $\mathcal{X}$ and the projectors $\mathbf{P}^{(n)}$ to show that

$$\begin{aligned}\|\mathcal{X} - \hat{\mathcal{X}}\|_{\mathsf{F}}^2 = \|\mathcal{X} \times_1 (\mathbf{I} - \mathbf{P}^{(1)})\|_{\mathsf{F}}^2 + \|\mathcal{X} \times_1 \mathbf{P}^{(1)} \times_2 (\mathbf{I} - \mathbf{P}^{(2)})\|_{\mathsf{F}}^2 \\ + \|\mathcal{X} \times_1 \mathbf{P}^{(1)} \times_2 \mathbf{P}^{(2)} \times_3 (\mathbf{I} - \mathbf{P}^{(3)})\|_{\mathsf{F}}^2.\end{aligned}$$

(Hint: introduce the terms $\mathcal{X} \times_1 \mathbf{P}^{(1)} - \mathcal{X} \times_1 \mathbf{P}^{(1)} + \mathcal{X} \times_1 \mathbf{P}^{(1)} \times_2 \mathbf{P}^{(2)} - \mathcal{X} \times_1 \mathbf{P}^{(1)} \times_2 \mathbf{P}^{(2)}$ inside the norm and use the result of the previous item.)

(c) By using the non-expansiveness of nth-mode orthogonal projection, show then that

$$\|\mathcal{X} - \hat{\mathcal{X}}\|_{\mathsf{F}}^2 \leqslant \sum_{n=1}^{3} \|\mathcal{X} \times_n (\mathbf{I} - \mathbf{P}^{(n)})\|_{\mathsf{F}}^2 = \sum_{n=1}^{3} \|[\mathcal{X}]_{(n)} - \mathbf{P}^{(n)} [\mathcal{X}]_{(n)}\|_{\mathsf{F}}^2.$$

(d) Finally, use the Eckart–Young theorem (optimality of the SVD truncation) to obtain the result (2) from the previous item. (Hint: recall that $[\mathcal{X}^\star]_{(n)}$ has rank at most R_n.)