Tensor Decompositions Explainer

Ibrohim Nosirov, Julian Bellavita

Basics and Notation

This section covers the subset of basic tensor notation that is necessary to understand the rest of this document. Throughout the section, let $\mathcal{X} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ be an order d tensor with mode j of size n_j. $\mathcal{X}$ is indexed using a d-tuple of indices $(i_1 \dots i_d)$. The Frobenius norm generalizes in a straightforward manner to tensors, meanining

$$\|\mathcal{X}\|_F = \sum_{i_1 \dots i_d}^{n_1 \dots n_d} \sqrt{\mathcal{X}^2_{i_1,\dots,i_d}}$$

Slices and Fibers

A slice of a tensor is a subset of tensor entries with one fixed index. For example, an i_j slice of $\mathcal{X}$ is given by

$$\mathcal{X}_{:,\dots,:,i_j,:,\dots,:} \in \mathbb{R}^{n_1 \times \cdots \times n_{j-1} \times n_{j+1} \times \dots n_d}$$

A tensor fiber is a subset of tensor entries with all but one fixed index. If we consider the so-called "mode i_j fibers" of $\mathcal{X}$, there are $\prod_{k \neq j} n_k$ mode-i_j fibers, and each is a vector in $\mathbb{R}^{n_j}$ uniquely identified by a $(d-1)$-tuple of indices. Formally, each mode-i_j fiber of $\mathcal{X}$ is given by

$$\mathcal{X}_{i_1 \dots i_{j-1} : i_{j+1} \dots i_d} \in \mathbb{R}^{n_j}$$

Matricizations/Unfoldings

Oftentimes it is necessary to store and perform some operation on all mode-k fibers of a tensor. In such scenarios, the so-called *matricization* operation is useful. Informally, the mode-k matricization of $\mathcal{X}$ 'unfolds' the entries of $\mathcal{X}$ into a matrix with one column per mode-k fiber of $\mathcal{X}$. This essentially transforms the tensor into a matrix which can be analyzed using classical numerical linear algebra methods. Formally, the mode-k matricization of $\mathcal{X}$ is given by $\mathbf{X}_{(k)}$ and is a matrix with n_k rows and $\prod_{j \neq k} n_j$ columns.

Special Matrix Products

Kronecker Products

The Kronecker product of two matrices $A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times n}$ is defined as

$$A \otimes B = \begin{bmatrix} A_{11}B & A_{12}B & \dots & A_{1n}B \\ A_{21}B & A_{22}B & \dots & A_{2n}B \\ \vdots & \vdots & \vdots & \vdots \\ A_{n1}B & A_{n2}B & \dots & A_{nn}B \end{bmatrix}$$

Essentially, the Kronecker product multiplies each entry of A with the entirety of B and stores the resulting matrix in a single block of the output.

Hadamard Products

The Hadamard product of $A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times n}$ is the elementwise product of the two matrices. Formally, we have

$$A * B = \begin{bmatrix} A_{11}B_{11} & A_{12}B_{12} & \dots & A_{1n}B_{1n} \\ A_{21}B_{21} & A_{22}B_{22} & \dots & A_{2n}B_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ A_{n1}B_{n1} & A_{n2}B_{n1} & \dots & A_{nn}B_{nn} \end{bmatrix}$$

Khatri-Rao Products

The Khatri-Rao product is a column-wise Kronecker product of two matrices $A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times n}$. The output is therefore a tall-skinny matrix, with the number of rows being extremely large. Formally, the Khatri-Rao product is given by

$$A \odot B = \begin{bmatrix} A_{:1}B_{:1} & A_{:2}B_{2:} & \dots & A_{:n}B_{:n} \end{bmatrix}$$

Tensor Times Matrix Product

Tensors can be multiplied with matrices using an operation with similar logic to classical matrix multiplication. The so-called *Tensor Times Matrix Product* of a tensor $\mathcal{X}$ and a matrix $U \in \mathbb{R}^{m \times n_k}$ computes an inner product between each mode-k fiber of $\mathcal{X}$ and all rows of U to produce an output tensor $\mathcal{Y} \in \mathbb{R}^{n_1 \times \cdots \times n_{k-1} \times m \times n_{k+1} \times \ldots n_k}$. Formally, this is given by

$$\mathcal{Y}_{i_1 \ldots i_{k-1}, j, i_{k+1} \ldots i_n} = \sum_{i_k}^{n_k} \mathcal{X}_{i_1 \ldots i_n} U_{j, i_k}$$

This operation can be written in two ways: in terms of normal tensors and in terms of matricized tensors

$$\mathcal{Y} = \mathcal{X} \times_k U \iff \mathbf{Y}_k = U \mathbf{X}_k$$