CHAPTER -5

MULTILINEAR ALGEBRA

This chapter aims at discussing the basic concepts of tensor algebra along with their notations and symbols. As our method 3D-ITSL is itself a tensor based, so understanding tensors is an indispensable requirement. Section 5.1 introduces the tensor by providing its mathematical representation. The notations are discussed in section 5.2 and the concept of rank is considered in section 5.3. The chapter is concluded with a review of other tensor decomposition techniques in section 4.

5.1 Definition

A tensor is a multidimensional array. More precisely, an N-way or Nth-order tensor is an element of the tensor product of N vector spaces, each of which has its own coordinate system.

The tensor concept that has been employed in our work is different from the conventional one which is used in physics and engineering and are generally referred to as tensor fields in mathematics. A third-order tensor has three indices as shown in Figure 5.1. A first-order tensor is a vector, a second-order tensor is a matrix and tensors of order three or higher are called higher-order tensors. A p^{th} - order tensor $\widehat{\boldsymbol{A}}$ can be defined as a multi-way array with p indices. A third-order tensor, for example, is written

$$\widehat{\boldsymbol{A}} = \left(\boldsymbol{a_{ijk}}\right) \in \boldsymbol{R^{I\,xJxK}} \tag{5.1}$$

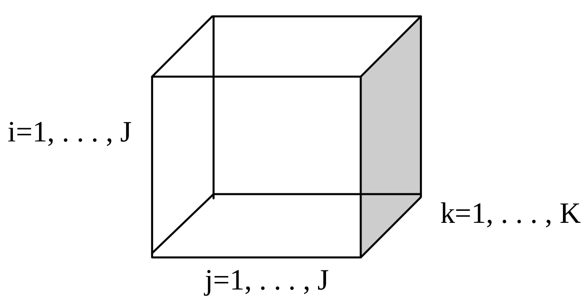


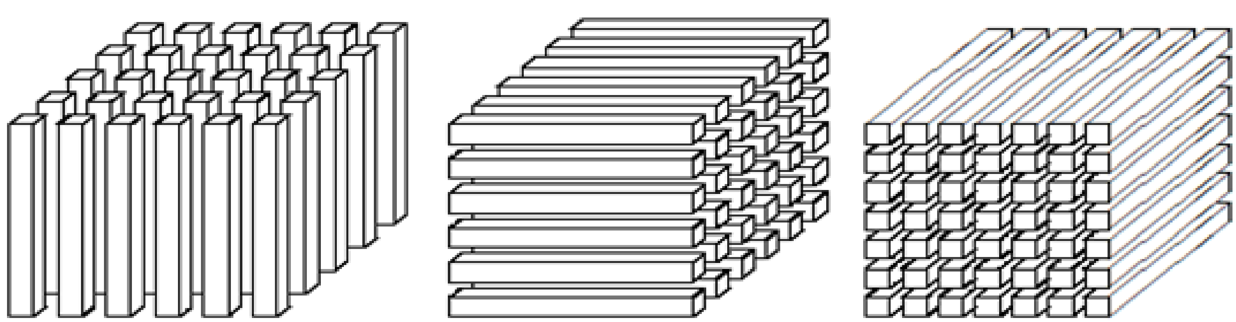
Figure 5.1 A third order order tensor $\widehat{\boldsymbol{A}} = (\boldsymbol{a_{ijk}}) \in \boldsymbol{R^{I\,xJxK}}$

Thus, a third-order tensor can be viewed as a “box,” and so forth. If x and y are real-valued vectors, it is well known that $\boldsymbol{xy^T}$ = $x\ o\ y$ is a rank-one matrix (“°” denotes the outer product).

Similarly, if $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(p)}$ are real-valued vectors, then $\widehat{\boldsymbol{A}} = \boldsymbol{x}^{(1)} \;\boldsymbol{o}\; \boldsymbol{x}^{(2)} \;\boldsymbol{o}\; \ldots\ldots \;\boldsymbol{o}\; \boldsymbol{x}^{(p)}$ is a rank-one tensor with $\widehat{\boldsymbol{A}}\,(\boldsymbol{i_1}, \boldsymbol{i_2}, \ldots, \boldsymbol{i_p}) = \boldsymbol{x}_{\boldsymbol{i_1}}^{(1)} \boldsymbol{x}_{\boldsymbol{i_2}}^{(2)} \ldots \boldsymbol{x}_{\boldsymbol{i_p}}^{(p)}$

5.2 Glimpse of the notations

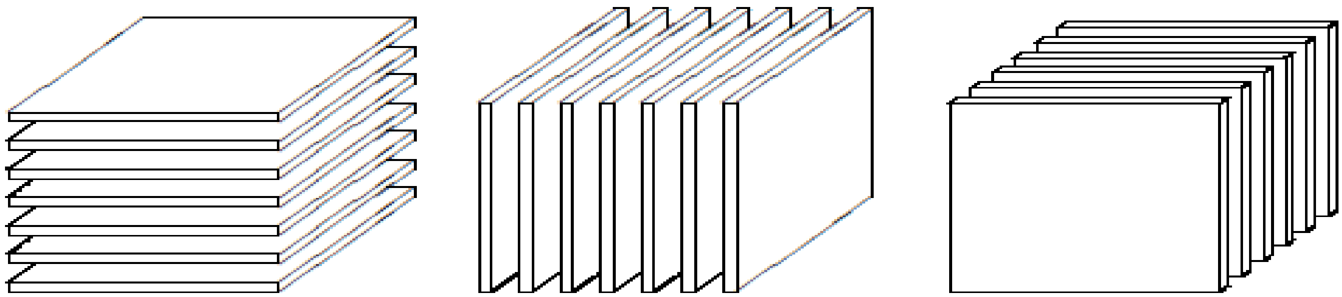
The order of a tensor is the number of dimensions, also known as ways or modes. Scalars are denoted by lowercase letters, e.g., a. Vectors are denoted by boldface lowercase letters, e.g., a. Matrices (tensors of order two) are denoted by boldface capital letters, e.g., A. Higher-order tensors (order three or higher) are denoted by a bold accented capital letters, e.g., $\widehat{\boldsymbol{X}}$. The i^{th} entry of a vector a is denoted by a_i, element (i ; j) of a matrix A is denoted by a_{ij} , and element (i; j; k) of a third-order tensor $\widehat{\boldsymbol{X}}$ is denoted by x_{ijk}. Indices typically range from 1 to their capital version, e.g., i = 1; : : : ; I. The nth element in a sequence is denoted by a superscript in parentheses, e.g., $A_{(n)}$ denotes the nth matrix in a sequence. Subarrays are formed when a subset of the indices is fixed. For matrices, these are the rows and columns. A colon is used to indicate all elements of a mode. Thus, the j^{th} column of A is denoted by $a_{:j}$, and the i^{th} row of a matrix A is denoted by $a_{i:}$. Alternatively, the j^{th} column of a matrix, $a_{:j}$, may be denoted more compactly as a_j . Fibers are the higher order analogue of matrix rows and columns. A fiber is defined by fixing every index but one. A matrix column is a mode-1 fiber and a matrix row is a mode-2 fiber. Third-order tensors have column, row, and tube fibers, denoted by $x_{:jk}$, $x_{i:k}$, and $x_{ij:}$, respectively; see Figure 5.2. When extracted from the tensor, fibers are always assumed to be oriented as column vectors



a) Model-1 (column) fibers: $x_{:jk}$ b) Mode-2 (row) fibers: $x_{i:k}$ c) Mode-3 (tube) fibers: $x_{ij:}$

Figure 5.2 Fibers of a 3rd order tensor

Slices are two-dimensional sections of a tensor, defined by fixing all but two indices. Figure 5.3 shows the horizontal, lateral, and frontal slides of a third-order tensor $\widehat{\boldsymbol{X}}$, denoted by $X_{i::}$, $X_{:j:}$, and $X_{::k}$, respectively. Alternatively, the kth frontal slice of a third-order tensor, $X_{::k}$, may be denoted more compactly as X_k.



a) Horizontal slices: $X_{i::}$ b) Lateral slices: $X_{:j:}$ c) Frontal slices: $X_{::k}$ (or X_k)

Figure 5.3 Slices of a 3rd- order tensor

The norm of a tensor $\widehat{\boldsymbol{X}} \in \boldsymbol{R}^{I_1 \times I_2 \times I_3 \times \cdots \times I_N}$ is the square root of the sum of the squares of all its elements, i.e

$$\left\|\widehat{\boldsymbol{X}}\right\| = \sqrt{\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \dots \sum_{I_N=1}^{I_N} x_{i_1 i_2 .. i_N}^2} \tag{5.2}$$

This is analogous to the matrix Frobenius norm, which is denoted ||A|| for a matrix A. The inner product of two same-sized tensors $\widehat{\boldsymbol{X}}, \widehat{\boldsymbol{Y}} \in \boldsymbol{R}^{I_1 \times I_2 \times I_3 \times \cdots \times I_N}$ is the sum of the products of their entries, i.e.,

$$< \widehat{\boldsymbol{X}}, \widehat{\boldsymbol{Y}} > = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \dots \sum_{I_N=1}^{I_N} x_{i_1 i_2 \dots i_N} \, y_{i_1 i_2 \dots i_N} \tag{5.3}$$

5.2.1 Rank-1 Tensors

An N-way tensor $\widehat{\boldsymbol{X}} \in \boldsymbol{R}^{I_1 \times I_2 \times I_3 \times \cdots \times I_N}$ is rank one if it can be written as the outer product of N vectors, i.e.,= $\boldsymbol{a}^{(1)} \circ \boldsymbol{a}^{(2)} \circ \dots . \circ \boldsymbol{a}^{(N)}$. The symbol “o” represents the vector outer product. This means that each element of the tensor is the product of the corresponding vector elements

$$x_{t_1 t_2 \dots t_N} = a_{t_1}^{(1)} a_{t_2}^{(2)} a_{t_N}^{(N)} \quad \text{For all } 1 \le \boldsymbol{i_n} \le \boldsymbol{I_n} \tag{5.4}$$

5.2.2 Symmetry

A tensor is called cubical if every mode is the same size, i.e.,$\hat{\boldsymbol{Y}} \in \boldsymbol{R}^{\boldsymbol{I}\times\boldsymbol{I}\times\cdots\times\boldsymbol{I}}$, X ∈ $\boldsymbol{R}^{\boldsymbol{I}\times\boldsymbol{I}\times\ldots\times\boldsymbol{I}}$. A cubical tensor is called supersymmetric if its elements remain constant under any permutation of the indices. For instance, a three-way tensor X ∈ $\boldsymbol{R}^{\boldsymbol{I}\times\boldsymbol{I}\times\boldsymbol{I}}$ is super symmetric if

$$\boldsymbol{x_{ijk}} = \boldsymbol{x_{ikj}} = \boldsymbol{x_{jik}} = \boldsymbol{x_{jki}} = \boldsymbol{x_{kij}} = \boldsymbol{x_{kli}} \quad \text{For all i, j, k = 1,…,I} \tag{5.5}$$

Tensors can be symmetric in two or more modes as well. For example, a three-way tensor $\boldsymbol{x} \in \boldsymbol{R}^{\boldsymbol{I}\times\boldsymbol{I}\times\boldsymbol{k}}$ is symmetric in modes one and two if all its frontal slices are symmetric, i.e.,

$$\boldsymbol{X_k} = \boldsymbol{X_k^T} \quad \text{for all k= 1,. . . , K} \tag{5.6}$$

5.2.3 Diagonal Tensor

A tensor $\hat{\boldsymbol{X}} \in \boldsymbol{R}^{\boldsymbol{I_1}\times\boldsymbol{I_2}\times\boldsymbol{I_3}\times\cdots\times\boldsymbol{I_N}}$ is diagonal if $\boldsymbol{x_{t_1 t_2 \ldots t_N}} \neq \boldsymbol{0}$ only if$\boldsymbol{i_1} = \boldsymbol{i_2} = \cdots = \boldsymbol{i_N}$. Figure 5.4 illustrates a cubical tensor with ones along the super diagonal.

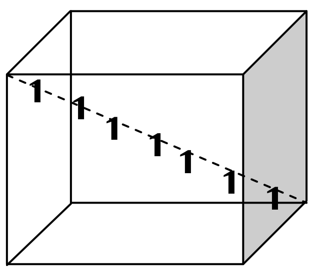


Figure5.4 Three way tensor of size IxIxI with ones along the superdiagonal

5.2.4 Matricization

Matricization is also known as unfolding or flattening and is the process of reordering the elements of an N- way array into a matrix. For instance, a 2 x 3 x 4 tensor can be arranged as a 6 x 4 matrix or a 3 x 8 matrix, and so on. Each order of a tensor is associated with a “mode”. By unfolding a tensor along a mode, a tensor’s unfolding matrix corresponding to this mode is obtained. For example, the mode-n unfolding matrix

$$\boldsymbol{A_{(n)}} \in \boldsymbol{R}^{\boldsymbol{I_n}\, x\, (\prod_{i\neq n} \boldsymbol{I_i})} \tag{5.7}$$

of $\widehat{\boldsymbol{A}}$ consists of I_n-dimensional mode-n column vectors which are obtained by varying the nth-mode index in and keeping indices of the other modes fixed, i.e. the column vectors of A(n) are just the mode-n vectors. Figure 5.5 shows the process of unfolding a 3-order tensor $\widehat{\boldsymbol{A}}$ into three matrices: the mode-1 matrix $A_{(1)}$ consisting of I_1-dimensional column vectors, the mode-2 matrix $A_{(2)}$ consisting of I_2-dimensional column vectors, and the mode-3 matrix $A_{(3)}$ consisting of I_3- dimensional column vectors. The inverse operation of the mode-n unfolding is the mode-n folding which restores the original tensor $\widehat{\boldsymbol{A}}$ from the mode-n unfolding matrix $A_{(n)}$, represented as $\hat{A} = \boldsymbol{fold}\ (\boldsymbol{A}_{(\boldsymbol{n})}, \boldsymbol{n})$. The mode-n rank R_n of $\widehat{\boldsymbol{A}}$ is defined as the dimension of the space generated by the mode-n vectors: $\boldsymbol{R_n} = \boldsymbol{rank}(\boldsymbol{A}_{(\boldsymbol{n})})$.

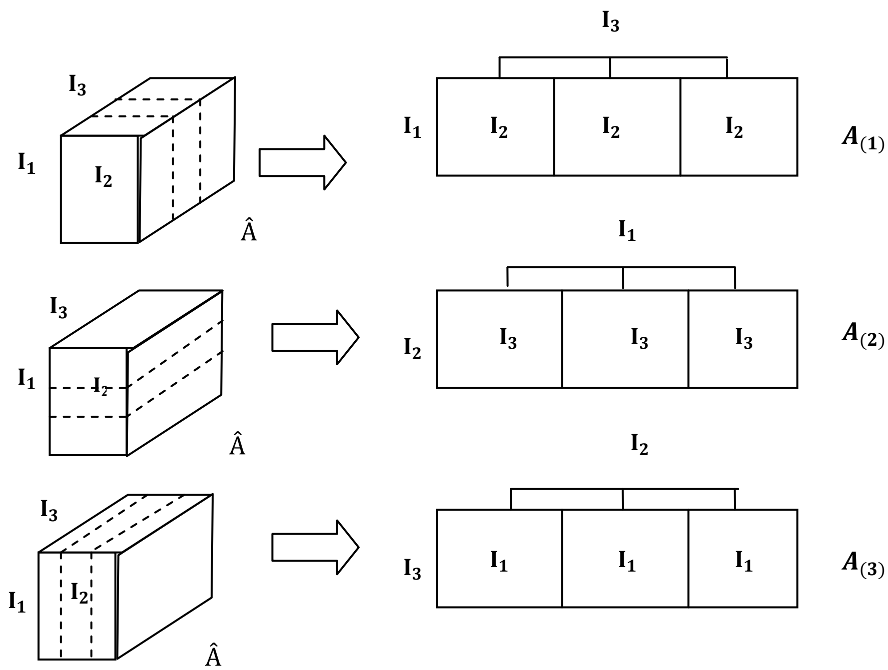


Figure 5.5 Illustration of unfolding a 3-order tensor [6]

5.2.5 Tensor Multiplication

Tensors can be multiplied together, though obviously the notation and symbols for this are much more complex than for matrices. Here we consider only the tensor n-mode product, i.e., multiplying a tensor by a matrix (or a vector) in mode n. N-mode multiplication The n-mode (matrix) product of a tensor $\boldsymbol{x} \in \boldsymbol{R}^{\boldsymbol{I_1} \times \boldsymbol{I_2} \times \boldsymbol{I_3} \times \ldots \times \boldsymbol{I_N}}$ with a matrix $\boldsymbol{U} \in \boldsymbol{R}^{\boldsymbol{J} \times \boldsymbol{I_N}}$ In is denoted by $\boldsymbol{x} \times_{\boldsymbol{n}} \boldsymbol{U}$ and is of size $\boldsymbol{I_1} \times \boldsymbol{I_2} \times \boldsymbol{I_3} \times \ldots \boldsymbol{I_{n-1}} \times \boldsymbol{J} \times \boldsymbol{I_{n+1}} \ldots \boldsymbol{I_N}$. Element-wise, we have

$$(\boldsymbol{x} \times_{\boldsymbol{n}} \boldsymbol{U})_{t_1 \ldots t_{n-1} j_{n+1} \ldots t_N} = \sum\nolimits_{t_n=1}^{I_n} \mathbf{x}_{\mathbf{t_1 t_2 \ldots t_N}} \mathbf{u}_{\mathbf{j t_n}} \tag{5.8}$$

Each mode-n fiber is multiplied by the matrix U. The idea can also be expressed in terms of unfolded tensors:

$$\boldsymbol{y} = \boldsymbol{x} \times_{\boldsymbol{n}} \boldsymbol{U} \;\Leftrightarrow\; \boldsymbol{Y}_{(\boldsymbol{n})} = \boldsymbol{U}\boldsymbol{X}_{(\boldsymbol{n})} \tag{5.9}$$

The n-mode product of a tensor with a matrix is related to a change of basis in the case when a tensor defines a multi-linear operator. The operation of mode-*n* product of a tensor and a matrix forms a new tensor. The mode-*n* product of tensor Â and matrix *U* is denoted as $\hat{A} \times_{\boldsymbol{n}} \boldsymbol{U}$. Let matrix U $\boldsymbol{\epsilon}\, \boldsymbol{R}^{J_n \times I_n}$ Then, $\hat{A} \times_{\boldsymbol{n}} \boldsymbol{U} \,\boldsymbol{\epsilon}\, \boldsymbol{R}^{I_1 \times \ldots \times I_{n-1} \times J_n \times I_{n+1} \times \ldots \times I_N}$ and its elements are calculated by:

$$(\hat{A} \times_{\boldsymbol{n}} \boldsymbol{U})_{i_1 \ldots i_{n-1} j_n i_{n+1} \ldots i_N} = \sum\nolimits_{\mathbf{i_n}} \mathbf{a}_{\mathbf{i_1 \ldots i_N}} \mathbf{u}_{\mathbf{j_n i_n}}. \tag{5.10}$$

Of course, , $\hat{A} \times_{\boldsymbol{n}} \boldsymbol{U}$ can be obtained by calculating *U · A(n)* first where the operation "." represents matrix multiplication, and then operating mode-*n* folding on *U · A(n)*.
Given a tensor $\hat{A} \boldsymbol{\epsilon}\, \boldsymbol{R}^{I_1 \times I_2 \times I_3 \times \cdots \times I_N}$ and three matrices C $\boldsymbol{R}^{J_n \times I_n}$, D $\boldsymbol{\epsilon}\, \boldsymbol{R}^{K_n \times J_n}$ and E $\boldsymbol{\epsilon}\, \boldsymbol{R}^{J_m \times I_m}$ $(\boldsymbol{n} \neq \boldsymbol{m})$, tensor's mode-*n* product has the following properties:

1. $(\hat{A} \times_{\boldsymbol{n}} \boldsymbol{C}) \times_{\boldsymbol{m}} \boldsymbol{E} = \left(\hat{A} \times_{\boldsymbol{m}} \boldsymbol{E}\right) \times_{\boldsymbol{n}} \boldsymbol{C} = \hat{A} \times_{\boldsymbol{n}} \boldsymbol{C} \times_{\boldsymbol{m}} \boldsymbol{E}$ (5.11)
2. $(\hat{A} \times_{\boldsymbol{n}} \boldsymbol{C}) \times_{\boldsymbol{n}} \boldsymbol{D} = \hat{A} \times_{\boldsymbol{n}} (\boldsymbol{D}.\boldsymbol{C})$ (5.12)

The scalar product of two tensors Â and ˆ*B* with the same set of indices is defined as:

$$< \widehat{\boldsymbol{A}}, \widehat{\boldsymbol{B}} >= \sum\nolimits_{i_1=1}^{I_1} \sum\nolimits_{i_2=1}^{I_2} \ldots \sum\nolimits_{I_N=1}^{I_N} \boldsymbol{a}_{i_1 i_2 \ldots i_N} \boldsymbol{b}_{i_1 i_2 \ldots i_N} \tag{5.13}$$

There are other tensor multiplication techniques eg: Matrix Kronecker, Khtri-Rao, Hadamard products

5.3 Tensor Decomposition

In context of matrices the singular value decomposition of a matrix A is a well-known, rank-revealing factorization. If the SVD of a matrix A is given by $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\boldsymbol{T}}$, then we write

$$A = \sum_{k=0}^{n} \sigma_i (u^{(i)} \circ v^{(i)}) \tag{5.14}$$

Where u(i) and v(i) are the i^{th} columns of U and V, respectively, the numbers σ_i on the diagonal of the diagonal matrix Σ are the singular values of A, and R is the rank of A. A matrix is a tensor of order 2. Analogous to above the tensor decomposition is higher-order SVD which is a generalization of the conventional matrix SVD. The SVD of a matrix $X \in R_{m\times n}$ can be represented as $\boldsymbol{X = U\, \Sigma\, V^T}$, where matrix U $\boldsymbol{R^{m\,x\,m}}$ matrix Σ $R^{m\,x\,n}$ and matrix V Σ $R^{n\,x\,n}$. The column vectors in U are the eigenvectors of $\boldsymbol{XX^T}$ and Σ is a diagonal matrix containing the singular values of X. The tensor decomposition of a N-order tensor Â which lies in N vector spaces involves N orthonormal matrices $\boldsymbol{U_{(1)}, U_{(2)}, \dots, U_{(N)}}$ to generate these N spaces respectively: the orthonormal column vectors of $U_{(N)}$ span the column space of the mode-n unfolding matrix $A_{(N)}$ $(1 \le n \le N)$. Then, the tensor Â is decomposed in the following way:

$$\hat{A} = \hat{B} \times_1 U^{(1)} \times_2 U^{(2)} \cdots \times_N U^{(N)} \tag{5.15}$$

Where $\boldsymbol{\hat{B}}$ is the core tensor controlling the interaction between the N mode matrices $U_{(1)}, \dots, U_{(N)}$.

In this way, each mode matrix $U_{(n)}$ $(1 \le n \le N)$ is computed by finding the SVD for themode-n unfolding matrix: $A_{(n)} = \tilde{U}_n \tilde{\Sigma}_n \tilde{V}_n^T$ and setting the mode matrix $U_{(n)}$ as the orthonormal $\breve{U}_{\boldsymbol{n}} (\breve{U}^{(\boldsymbol{n})} = \breve{U}_{\boldsymbol{n}})$.

$$\hat{B} = \hat{A} \times_1 U^{(1)^T} \cdots \times_n U^{(n)^T} \cdots \times_N U^{(N)^T}. \tag{5.16}$$

Such a decomposition can only be achieved offline, i.e. it cannot be used for incremental tensor subspace learning. In real applications, dimension reduction is necessary for a compact representation of a tensor. Lathauwer et al. [57] proposed a rank-$(R_1, R_2, \dots, R_N)$ approximation algorithm for the dimension reduction. The algorithm applies the technique of alternate least squares to find the dominant projection subspaces of a tensor. Given an N-order tensor, $\hat{A} \boldsymbol{\epsilon}\, \boldsymbol{R}^{I_1 \times I_2 \times I_3 \times \dots \times I_N}$, a rank – $(R_1, R_2, \dots, R_N)$ tensor is found to minimize the square of the Frobenius norm of the error tensor.