Tensor Algebra

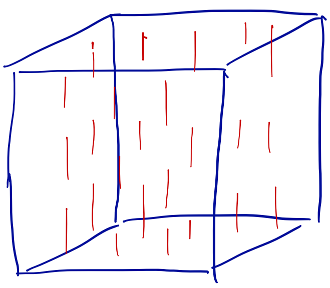
Indexing

Tensor $\underline{T} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ has elements

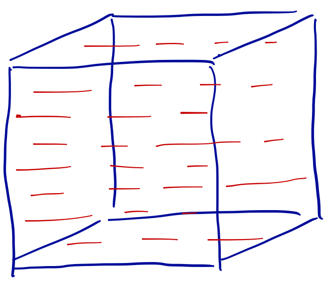
$$\underline{T} = (t_{i_1 i_2 \cdots i_N})$$

We can write $\vec{i} = i_1 i_2 \cdots i_N$ to denote an *index vector* $\vec{i} \in [I_1] \times [I_2] \times \cdots \times [I_N]$, where $[n] = \{1, 2, 3, \dots, n\}$.

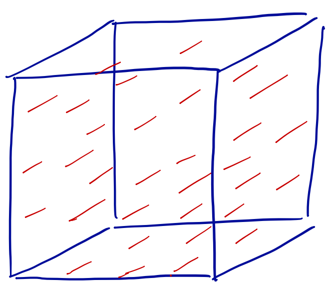
The *fibres* are denoted $\vec{t}_{i_1 i_2 \cdots i_{n-1} : i_{n+1} \cdots i_N}$



$t_{:i_2 i_3}$



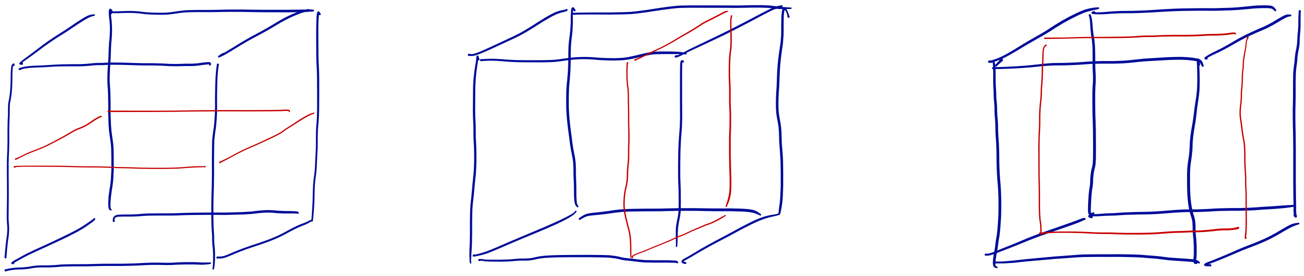
$t_{i_1 : i_3}$



$t_{i_1 i_2 :}$

The <u>slices</u> are denoted

$$T_{i_1 \cdots i_{j-1} : i_{j+1} \cdots i_{k-1} : i_{k+1} \cdots i_N}$$



$T_{i_1 ::}$ $\qquad$ $T_{: i_2 :}$ $\qquad$ $T_{:: i_3}$

For 3-way tensors, the frontal slice $T_{:: i_3}$ has a special short-hand notation

$$T_{:: i_3} = T_{i_3}$$

<u>Matricization</u>

A tensor $\underline{T}$ can be <u>matricized</u> or <u>unfolded</u> or <u>flattened</u> in a matrix by reordering its elements. <u>Mode-n matricization</u> takes the mode-n fibres and stacks them as the columns of the new matrix. If $\underline{T} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, its mode-1 matricization is a matrix

$$T_{(1)} \in \mathbb{R}^{I_1 \times (I_2 I_3 \cdots I_N)},$$

and in general, the mode-n matricization is a matrix

$$T_{(n)} \in \mathbb{R}^{I_n \times (I_1 I_2 \cdots I_{n-1} I_{n+1} \cdots I_N)}.$$

In mode-n matricization, element at $(i_1, i_2, \dots i_N)$ maps ~~to~~ (i_n, j), where

$$j = 1 + \sum_{\substack{k=1 \\ k \neq n}}^{N} (i_k - 1) J_k, \quad \text{where}$$

$$J_k = \prod_{\substack{m=1 \\ m \neq n}}^{k-1} I_m .$$

An example (next page) will clarify the process.

Matricization example

Let $\underline{T} \in \mathbb{R}^{4\times3\times2}$ with frontal slices

$$T_1 = \begin{pmatrix} 1 & 5 & 9 \\ 2 & 6 & 10 \\ 3 & 7 & 11 \\ 4 & 8 & 12 \end{pmatrix} \text{ and } T_2 = \begin{pmatrix} 13 & 17 & 21 \\ 14 & 18 & 22 \\ 15 & 19 & 23 \\ 16 & 20 & 24 \end{pmatrix}.$$

Now we have:

$$T_{(1)} = \begin{pmatrix} 1 & 5 & 9 & 13 & 17 & 21 \\ 2 & 6 & 10 & 14 & 18 & 22 \\ 3 & 7 & 11 & 15 & 19 & 23 \\ 4 & 8 & 12 & 16 & 20 & 24 \end{pmatrix}$$

$$T_{(2)} = \begin{pmatrix} 1 & 2 & 3 & 4 & 13 & 14 & 15 & 16 \\ 5 & 6 & 7 & 8 & 17 & 18 & 19 & 20 \\ 9 & 10 & 11 & 12 & 21 & 22 & 23 & 24 \end{pmatrix}$$

$$T_{(3)} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & \cdots & 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 & 17 & \cdots & 21 & 22 & 23 & 24 \end{pmatrix}$$

Vectorizations

A vectorization of a tensor stacks the columns of its mode-1 matricization in a column vector. $\underline{T}$ as above, we have $\mathrm{vec}(\underline{T}) = (1, 2, 3, \ldots, 24)^T$.

Tensor times {scalar, vector, matrix, tensor}

Tensor times a scalar, $\alpha \underline{T}$, scales every element of T by $\alpha \in \mathbb{R}$,

$$\alpha \underline{T} = (\alpha t_{\vec{i}}).$$

The <u>n-mode vector-tensor product</u> of a tensor $\underline{T} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ with vector $\vec{v} \in \mathbb{R}^{I_n}$ is

$$\underline{T} \,\bar{\times}_n\, \vec{v} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_{n-1} \times I_{n+1} \times \dots \times I_N},$$

$$\left(\underline{T} \,\bar{\times}_n\, \vec{v}\right)_{i_1 \dots i_{n-1} i_{n+1} \dots i_N} = \sum_{i_n = 1}^{I_n} t_{i_1 i_2 \dots i_N} v_{i_n}$$

$$= \left[\left\langle \vec{t}_{i_1 \dots i_{n-1} : i_{n+1} \dots i_N}, \vec{v} \right\rangle\right]_{i_1 \dots i_{n-1} i_{n+1} \dots i_N}$$

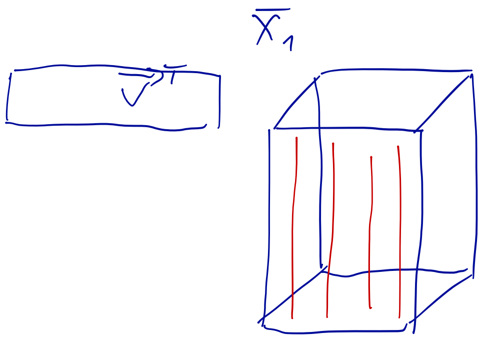
where $\langle \cdot, \cdot \rangle$ is the <u>inner product</u> of two vectors. The idea is to take the inner products between all mode-n fibres and $\vec{v}$.

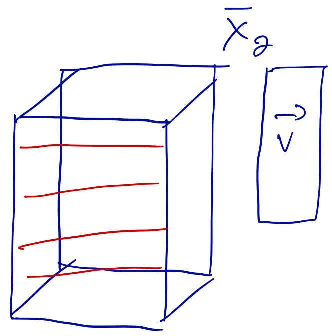
In tensor-vector products, precedence matters:

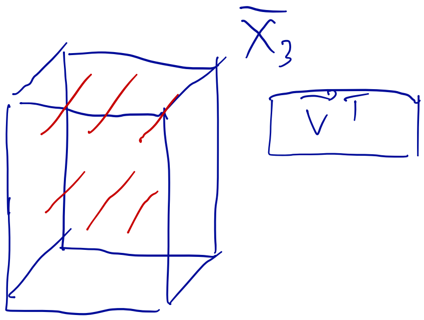
$$\underline{T} \,\bar{\times}_m\, \vec{a} \,\bar{\times}_n\, \vec{b} = \left(\underline{T} \,\bar{\times}_m\, \vec{a}\right) \bar{\times}_{n-1}\, \vec{b} = \left(\underline{T} \,\bar{\times}_n\, \vec{b}\right) \bar{\times}_m\, \vec{a}$$

If $\underline{T}$ is 2-way (a matrix), we have

$$\underline{T} \bar{\times}_1 \vec{v} = \vec{v}^T \underline{T} \quad \text{and} \quad \underline{T} \bar{\times}_2 \vec{v} = \underline{T} \vec{v}.$$







Let $\underline{T} \in \mathbb{R}^{4\times 3\times 2}$ with frontal slices

$$T_1 = \begin{pmatrix} 1 & 5 & 9 \\ 2 & 6 & 10 \\ 3 & 7 & 11 \\ 4 & 8 & 12 \end{pmatrix} \quad \text{and} \quad T_2 = \begin{pmatrix} 13 & 17 & 21 \\ 14 & 18 & 22 \\ 15 & 19 & 23 \\ 16 & 20 & 24 \end{pmatrix},$$

and let $\vec{v} = (5, 10)^T$. Then

$$\underline{T} \bar{\times}_3 \vec{v} = \begin{pmatrix} 5+130 & 25+170 & 45+210 \\ 10+140 & 30+180 & 50+220 \\ 15+150 & 35+190 & 55+230 \\ 20+160 & 40+200 & 60+240 \end{pmatrix} = \begin{pmatrix} 135 & 195 & 255 \\ 150 & 210 & 270 \\ 165 & 225 & 285 \\ 180 & 240 & 300 \end{pmatrix}.$$

The <u>n-mode matrix-tensor product</u> of tensor $\underline{T} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ and matrix $M \in \mathbb{R}^{J \times I_n}$ is

$$\underline{T} \times_n M \in \mathbb{R}^{I_1 \times \dots \times I_{n-1} \times J \times I_{n+1} \times \dots \times I_N}$$

$$(\underline{T} \times_n M)_{i_1 \dots i_{n-1} j i_{n+1} \dots i_N} = \sum_{i_n = 1}^{I_n} t_{i_1 i_2 \dots i_N} m_{j i_n}$$

$$= \left[M \vec{t}_{i_1 \dots i_{n-1} : i_{n+1} \dots i_N} \right]_{i_1 \dots i_N}$$

That is, we multiply each mode-n fibre with M. Equivalently, we can use unfolding:

$$\underline{S} = \underline{T} \times_n M \iff S_{(n)} = M T_{(n)}.$$

If $\underline{T}$ is 2-way (i.e. a matrix), $T_{(1)} = \underline{T}$ and $T_{(2)} = \underline{T}^T$. So

$$\underline{T} \times_1 M = M \underline{T} \quad \text{and} \quad \underline{T} \times_2 M = M \underline{T}^T$$

The order of tensor-matrix multiplications over different modes doesn't matter: if $m \neq n$, then

$$\underline{T} \times_m A \times_n B = \underline{T} \times_n B \times_m A .$$

If the modes are the same and $A \in \mathbb{R}^{J \times I_n}$ and $B \in \mathbb{R}^{K \times J}$, then

$$\underline{T} \times_n A \times_n B = \underline{T} \times_n (BA) .$$

Let $\underline{T} \in \mathbb{R}^{4 \times 3 \times 2}$ with frontal slices

$$T_1 = \begin{pmatrix} 1 & 5 & 9 \\ 2 & 6 & 10 \\ 3 & 7 & 11 \\ 4 & 8 & 12 \end{pmatrix} \text{ and } T_2 = \begin{pmatrix} 13 & 17 & 21 \\ 14 & 18 & 22 \\ 15 & 19 & 23 \\ 16 & 20 & 24 \end{pmatrix},$$

and let

$$M = \begin{pmatrix} 1 & 2 & 4 & 6 \\ 10 & 20 & 40 & 60 \end{pmatrix}$$

Then, if $\underline{S} = \underline{T} \times_1 M$, we have

$s_{111} = 1 \cdot 1 + 2 \cdot 2 + 4 \cdot 3 + 6 \cdot 4 = 41$ and

$$S_1 = \begin{pmatrix} 41 & 93 & 145 \\ 410 & 930 & 1450 \end{pmatrix} \text{ and } S_2 = \begin{pmatrix} 197 & 249 & 301 \\ 1970 & 2490 & 3010 \end{pmatrix}$$

The tensor inner product of two tensors $\underline{S}, \underline{T} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is the sum of their element-wise products,

$$\langle \underline{S}, \underline{T} \rangle = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} s_{i_1 i_2 \cdots i_N} t_{i_1 i_2 \cdots i_N} .$$

Tensor norm

The norm of a tensor $\underline{T} \in \mathbb{R}^{I_1 \times I_2 \times \cdots I_N}$ is the square-root of the sum of the squares of its elements:

$$\|\underline{T}\| = \left(\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} t_{i_1 i_2 \cdots i_N}^2 \right)^{1/2}$$

This can be alternatively be defined as $\sqrt{\langle \underline{T}, \underline{T} \rangle}$, $\|T_{(1)}\|_F$, $\|T_{(i)}\|_F$, $\|\mathrm{vec}(\underline{T})\|_2$, or $\sqrt{\mathrm{trace}(T_{(1)} T_{(1)}^T)}$, or using any other way to define the Euclidean/Frobenius norm of a vector/matrix.

Symmetry

Tensor $\underline{T}$ is cubical if all of its modes have the same dimensionality:

$$\underline{T} \in \mathbb{R}^{I \times I \times I \times \cdots \times I}$$

Cubical tensor is (super-) symmetric, if its elements remain constant under any permutation of the indices. If $\underline{T} \in \mathbb{R}^{I \times I \times I}$, it is symmetric if and only if

$t_{ijk} = t_{ikj} = t_{jik} = t_{jki} = t_{kij} = t_{kji}$ for all $i, j, k \in [I]$

Tensor $\underline{T} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is (hyper-) diagonal if $t_{i_1 i_2 \cdots i_N} \neq 0$ only if $i_1 = i_2 = \cdots = i_N$. If $I_1 = I_2 = \cdots I_N$, $\underline{T}$ is also symmetric.

If the diagonal entries are all 1s, tensor behaves similarly to the identity matrix.