An introduction to tensors

Ion Nechita (CNRS, LPT Toulouse)

3rd online Sakura meeting, July 1st 2021

Talk outline

Definition and graphical notation

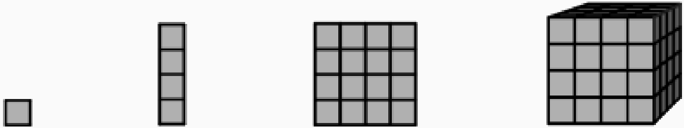
Tensor rank(s)

Tensor decompositions

Definition and graphical notation

What is a tensor?

- A scalar (0D object): $x \in \mathbb{C}$
- A vector (1D object): $v = (v_i)_{i\in[d]} \in \mathbb{C}^d$
- A matrix (2D object): $A = (A_{ij})_{i,j\in[d]} \in \mathcal{M}_d(\mathbb{C}) \cong \mathbb{C}^d \otimes \mathbb{C}^d$
- A tensor (3D object): $T = (T_{ijk})_{i,j,k\in[d]} \in \mathbb{C}^d \otimes \mathbb{C}^d \otimes \mathbb{C}^d$



by Anna Seigal

Definition

A *n*-mode tensor is an element $T \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \otimes \cdots \otimes \mathbb{C}^{d_n}$. It can be decomposed in the computational basis as

$$T = \sum_{i\in[d_1]\times\cdots\times[d_n]} T_{i_1 i_2 \cdots i_n} \underbrace{e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_n}}_{=:|i_1 i_2 \cdots i_n\rangle} .$$

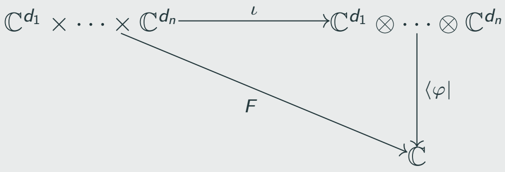
Abstract definition

Tensors are related to multilinear forms via the universal-property.

Proposition

Any $F : \mathbb{C}^{d_1} \times \mathbb{C}^{d_2} \times \cdots \times \mathbb{C}^{d_n} \to \mathbb{C}$ *factors through* $\mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_n}$ *uniquely. There exists a unique linear form* $\langle \varphi | \in [\mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_n}]^*$ *such that*

$$F(x_1, x_2, \ldots, x_n) = \langle \varphi, x_1 \otimes x_2 \otimes \cdots \otimes x_n \rangle, \qquad \forall x_i.$$



where $\iota(x_1, \ldots, x_n) = x_1 \otimes \cdots \otimes x_n$.

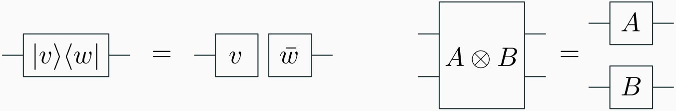
Graphical notation: tensor diagrams

A tensor diagram is a collection of boxes and wires.

❶ Boxes ⇝ tensors

- A box with n legs corresponds to a n-tensor. Each leg is associated to a vector space $\mathbb{C}^{d_i}$.
- Scalars are leg-less boxes, vectors have 1 leg, matrices have two legs, etc.

- The Kronecker product of tensors corresponds to juxtaposition.



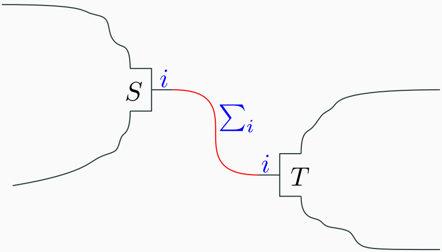
- Writing indices over a leg gives the corresponding coordinate.

Graphical notation: tensor diagrams

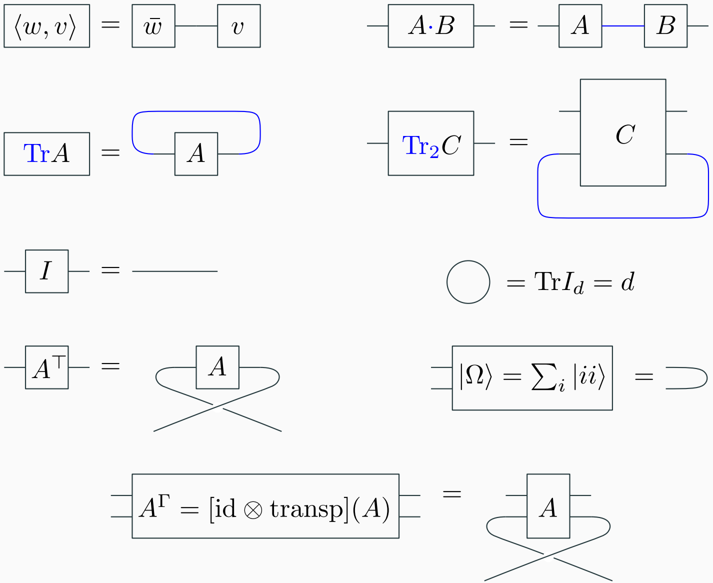
❷ Wires ⇝ tensor contractions

- A wire between two tensors corresponds to a sum over a common index. Only legs corresponding to the same vector space can be connected by a wire. Mathematically, a wire corresponds to the evaluation map:

$$\begin{aligned} \mathrm{ev} : V^* \times V &\to \mathbb{C} \\ (\alpha, v) &\mapsto \alpha(v). \end{aligned}$$



Operations on tensors



Tensor rank(s)

Tensor rank

A tensor of the form $x = x_1 \otimes x_2 \otimes \cdots \otimes x_n$ is called a simple (or product) tensor. In QIT, these are the separable pure states.

Definition

The rank of a tensor x is the minimum number of terms in a decomposition of x as a sum of simple tensors

$$R(x) := \min\{r : x = \sum_{i=1}^{r} x_1^{(i)} \otimes x_2^{(i)} \otimes \cdots \otimes x_n^{(i)}\}.$$

For matrices, this is the usual notion of rank:

$$R(A) = \min\{r : A = \sum_{i=1}^{r} |x_i\rangle\langle y_i|\}.$$

Examples for 3-qubit tensors:

- The GHZ state $|\mathrm{GHZ}\rangle = |000\rangle + |111\rangle$ has rank 2.
- The W state $|\mathrm{W}\rangle = |001\rangle + |010\rangle + |100\rangle$ has rank 3.

Border rank

In the case of matrices, the set of tensors of rank $\leq r$ is an algebraic variety (i.e. 0-set of some polynomials), hence closed.

For 3-tensors, this is no longer true: one can write a tensor of rank 3 as a limit of rank-2 tensors:

$$|\mathrm{W}\rangle = |001\rangle + |010\rangle + |100\rangle = \lim_{\varepsilon \to 0} \frac{(|0\rangle + \varepsilon |1\rangle)^{\otimes 3} - |000\rangle}{\varepsilon}.$$

Definition

Border rank of a tensor

$$\underline{R}(x) := \min\{r : x = \lim_{k \to \infty} x_k, \text{ with } R(x_k) = r \quad \forall k\}.$$

So $\underline{R}(|\mathrm{W}\rangle) = 2 < 3 = R(|\mathrm{W}\rangle)$.

Symmetric tensors

An element of $\vee^n(\mathbb{C}^d)$ is called a symmetric tensor. The tensors $|\mathrm{GHZ}\rangle$ and $|\mathrm{W}\rangle$ are symmetric.

Since $\mathrm{span}\{x^{\otimes n}\}_{x\in\mathbb{C}^d} = \vee^n(\mathbb{C}^d)$, one can define the symmetric rank

$$R_s = \min\{r \,:\, x = \sum_{i=1}^{r} x_i^{\otimes n}\}.$$

Clearly, $R_s \geq R$ for symmetric tensors. The inequality can be strict (Shitov's counterexample to Comon's conjecture from 2018, tensor of size 800^3, with $R \leq 903$ and $R_s \geq 904$).

Remarkably, symmetric tensors are in bijection with homogeneous polynomials in d variables of degree n

$$t \in \vee^n(\mathbb{C}^d) \leftrightarrow p(x_1, \dots x_d) = \langle t, (x_1, \dots, x_d)^{\otimes n}\rangle.$$

This correspondence has many interesting properties (GHZ-type tensors correspond to norms, partial traces correspond to derivatives, etc).

Tensor decompositions

The singular value decomposition

Any $d \times d$ matrix A can be decomposed as

$$A = \sum_{i=1}^{d} s_i |x_i\rangle\langle y_i|,$$

where $s_i \geq 0$ are the singular values and $\{x_i\}$, resp $\{y_i\}$ are orthonormal bases of $\mathbb{C}^d$. The number of non-zero singular values is the rank of A.

In QIT, this is know as the Schmidt decomposition

$$\mathbb{C}^d \otimes \mathbb{C}^d \ni |\varphi\rangle = \sum_{i=1}^{d} \sqrt{\lambda_i}\, |x_i\rangle \otimes |y_i\rangle \, .$$

If the vector $|\varphi\rangle$ is normalized, the Schmidt coefficients $\{\lambda_i\}_{i \in [d]}$ form a probability vector. The Schmidt coefficients contain most of the physically relevant information in the bipartite pure state $|\varphi\rangle$, such as the entropy of entanglement

$$E(|\varphi\rangle) = H(\lambda) = -\sum_{i=1}^{d} \lambda_i \log \lambda_i.$$

Tensor decompositions

Not all tensors ($n \geq 3$) can be decomposed as

$$x = \sum_{i=1}^{d} s_i x_i^{(1)} \otimes x_i^{(2)} \otimes \cdots \otimes x_i^{(n)},$$

with orthonormal $\{x_i^{(k)}\}_{i \in [d]}$ for all k. Naive dimension count:

$$\dim_{\mathbb{R}} (\mathbb{C}^d)^{\otimes n} = 2d^n \gg d + nd^2 - d.$$

Tensors admitting such a decomposition are called udeco. Their best rank-1 approximation is given by the term with the largest s_i.

Tucker and HOSVD decompositions: isometries + a core tensor

$$\mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_n} \ni T = (V_1 \otimes V_2 \otimes \cdots \otimes V_n) C,$$

with $C \in \mathbb{C}^{p_1} \otimes \cdots \otimes \mathbb{C}^{p_n}$ and $V_i : \mathbb{C}^{p_i} \to \mathbb{C}^{d_i}$ are isometries. For some applications, we want C to be as "diagonal" as possible.