Tensor Decompositions & Tensor Rank.

Goal: Thm for order-3 tensors important primitive for studying also higher order tensors.

Definition of Tensor Rank.

Rank-one tensors (elementary bricks)

- Take three vectors $\underline{a} \in \mathbb{R}^{I_1}$, $\underline{b} \in \mathbb{R}^{I_2}$, $\underline{c} \in \mathbb{R}^{I_3}$

  a^α, $\alpha = 1 \dots I_1$. b^β and c^γ.

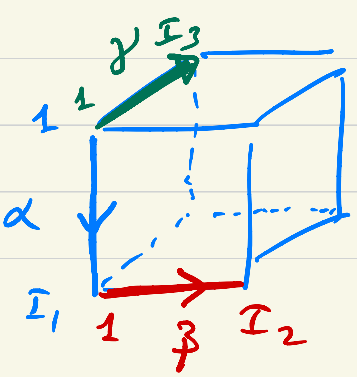
  $\beta = 1 \dots I_2$ $\quad \gamma = 1 \dots I_3$.

- Form the tensor product (or outer product)

$$\underline{a} \otimes \underline{b} \otimes \underline{c} = T$$

$$T^{\alpha\beta\gamma} = a^\alpha b^\beta c^\gamma$$

3-dim array of $I_1 I_2 I_3$ numbers.



Rank-R tensor. $R \geq 1$.

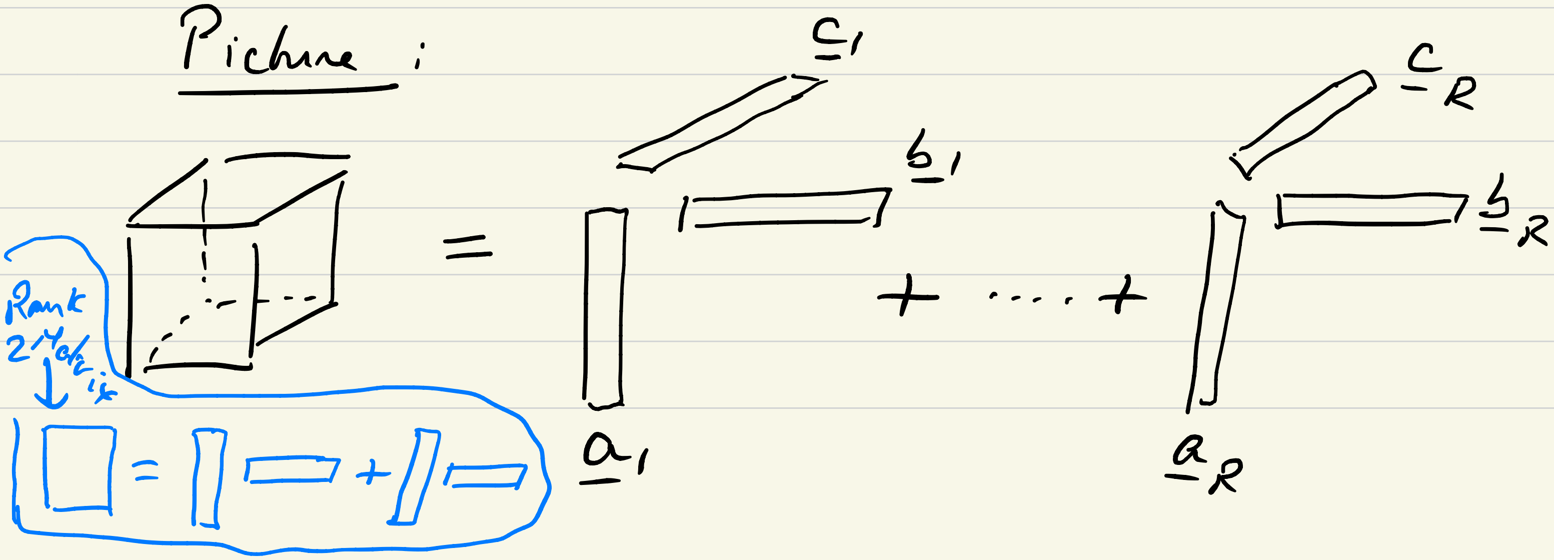
is a multi-array (3-arrays) that can be written as a sum of R rank-one tensors with R being the minimal possible number of terms.

$$T = \sum_{i=1}^{R} \underline{a}_i \otimes \underline{b}_i \otimes \underline{c}_i$$

R should be the smallest possible to qualify for Tensor-Rank.

Terminology: such decomp are also called "tensor factorizations" or "polyadic decompositions".

Picture :



Adopt the following Notation:

$$T = \sum_{i=1}^{R} \underline{a}_i \otimes \underline{b}_i \otimes \underline{c}_i$$

$$A = [\,\underline{a}_1 \;\ldots\; \underline{a}_R\,] \leftarrow I_1 \times R.$$

column vectors $\underline{a}_1 = \left[a_1^{\alpha} \right]_{\alpha = I_1}^{\alpha = 1}$

$$B = [\,\underline{b}_1 \;\ldots\; \underline{b}_R\,] \leftarrow I_2 \times R$$

$$C = [\,\underline{c}_1 \;\ldots\; \underline{c}_R\,] \leftarrow I_3 \times R.$$

Remarks about the notion of tensor-rank :

order-2 tensors

1) For matrices "tensor-rank" $\Longleftrightarrow$ usual lin-algebra notion of rank.

dim (row span) = dim (column span).

For tensors (of order $p \geq 3$) "Tensor-Rank" is NOT a linear-alg notion.

Moreover there are other Notions of Rank (see later lects).

2) Non trivial to decompose a tensor in a minimal # of rank one terms.

- unicity ?
- algorithms ?

Rule of thumb : R is "small" w.r.t. to I_1, I_2, I_3
$\rightarrow$ decomp tends to be unique
$\rightarrow$ efficient algs are known.

R is "moderate"
$\rightarrow$ decomp is unique ...
$\rightarrow$ eff algs are not always known.

R is "large" $\rightarrow$ dont even have unicity.

Basic Theorem. [about order 3-tensors. 1970's Jennrich, Carroll, Harshman ...]

- Let $A = [\underline{a}_1, \dots, \underline{a}_R]$ $I_1 \times R$

  $B = [\underline{b}_1, \dots, \underline{b}_R]$ $I_2 \times R$.

  $C = [\underline{c}_1, \dots, \underline{c}_R]$ $I_3 \times R$.
- A & B have full column rank R. (in particular $R \le \min(I_1, I_2)$)
- C has pairwise indep columns: $\forall i,j$ $\underline{c}_i \ne \lambda_{ij} \underline{c}_j$
- Take a 3-tensor

$$T = \sum_{i=1}^{R} \underline{a}_i \otimes \underline{b}_i \otimes \underline{c}_i \rightsquigarrow T^{\alpha\beta\gamma}$$

- Then the tensor (or multiarray of numbers $T^{\alpha\beta\gamma}$) can be decomposed in a unique way in a sum of at most R rank-one terms. And the tensor rank of T is R.
- Qualification for unicity: up to trivial rescaling of $\underline{a}, \underline{b}, \underline{c}$ $[\lambda \underline{a}_i \otimes \frac{1}{\sqrt{\lambda}} \underline{b}_i \otimes \frac{1}{\sqrt{\lambda}} \underline{c}_i]$ & up to permutation of terms in sum.

Remarks:

1) The thm will be proved in a constructive way and the proof will thus give also an algorithm poly(I_1, I_2, I_3).

2) For this thm we have $R \leq \min(I_1, I_2)$ – in this sense R is "small" –

– unicity & algorithmic result –

3) For $R > \min(I_1, I_2)$. less is known.

As an aside state the following result:

- suppose that k_a, k_b, k_c are Kruskal ranks of A, B, C.

(Ka is the KR if it is the max integer such that all subsets of Ka vects in $[\underline{a}_1 \cdots \underline{a}_R]$ are lin-indep; $\exists$ Ka+1 subset that is lin-dep)

if $2R+2 \leq k_a + k_b + k_c$ in decomp of T Then this decomp is unique.

- For example $I_1 = I_2 = I_3 = N$ & $R > N$ imagine $k_a = k_b = k_c = N$, $R < \frac{3N-2}{2}$.

here you have unicity result for "Moderate rank" $N < R < \frac{3N}{2} - 2$

BUT NO EFFICIENT ALGO'S ARE KNOWN!

4) Then is "surprising" in the sense that for 2-tensors (matrices) it does not hold!

<u>Recall the "rotation problem":</u>

$A = [\underline{a}_1, \underline{a}_2]$ $I_1 \times 2$ $R=2$.

$B = [\underline{b}_1, \underline{b}_2]$ $I_2 \times 2$ $R=2$.

A & B have rank $R=2$.

$$T = \underline{a}_1 \otimes \underline{b}_1 + \underline{a}_2 \otimes \underline{b}_2 \quad \checkmark$$

This decomp is not unique! <u>Why</u>?

$$T = \underline{a}_1 \underline{b}_1^T + \underline{a}_2 \underline{b}_2^T \qquad \underline{a} \otimes \underline{b} = \underbrace{\underline{a}\,\underline{b}^T}$$

outer product of vectors.

$$= [\underline{a}_1, \underline{a}_2] \begin{bmatrix} \underline{b}_1^T \\ \underline{b}_2^T \end{bmatrix}$$

$$= A B^T$$

$$= A \underbrace{R R^T}_{I} B^T = \underbrace{(AR)}_{A'} \underbrace{(BR)^T}_{B'^T}$$

$$= \underline{a}_1' \otimes \underline{b}_1' + \underline{a}_2' \otimes \underline{b}_2' \quad \checkmark \#$$

BREAK NEXT VIDEO PROVE THE THM