Swedish Summer School

Part I: Tensor Decompositions

Charles Spearman (1904): There are two types of intelligence

(1) eductive : make sense out of complexity

(2) reproductive : store and reproduce info

To test his theory, invented factor analysis

$$\begin{array}{cc} & \text{students (1000)} \\ \begin{array}{c}\text{tests} \\ (10)\end{array} & \begin{bmatrix} \quad & M & \quad \end{bmatrix} \end{array} \approx \begin{array}{c} \overbrace{\quad}^{\text{inner-dimension (2)}} \\ \begin{bmatrix} A \end{bmatrix} \end{array} \begin{bmatrix} \quad B \quad \end{bmatrix}$$

Setup: Given $M = \sum_{i=1}^{r} a_i b_i^T$

$$= AB^T$$

"correct" factors

However for any rotation R, we have

$$M = (AR)(R^{-1}B^{-1})$$

alternative factorization

Q: when can we find the factors $\{a_i\}$ and $\{b_i\}$ uniquely?

e.g. up to the trivial rescaling

$$a_i \leftarrow \alpha a_i$$
$$b_i \leftarrow \frac{1}{\alpha} b_i$$

and permutation

Claim: the factors $\{a_i\}$ and $\{b_i\}$ are not uniquely determined, unless we impose additional conditions on them

e.g. if $\{a_i\}$ are orthogonal, same for $\{b_i\}$

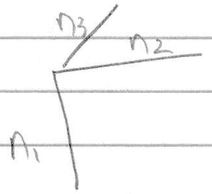
or if $\mathrm{rank}(M)=1$

This is called the rotation problem and is a major issue in factor analysis

It motivates the study of tensor methods

Tensor Decompositions: Definitions

A third-order tensor T has three dimensions, sometimes called rows, columns, tubes



The size of T is $n_1 \times n_2 \times n_3$

definition: A rank one third order tensor T is the tensor product of three vectors u, v and w and its entries are

$$T_{i,j,k} = u_i v_j w_k$$

Also written as $T = u \otimes v \otimes w$

The rank of T is the smallest r s.t.

$$T = \sum_{i=1}^{r} u^{(i)} \otimes v^{(i)} \otimes w^{(i)}$$

We can view a tensor T as a stacked collection of matrices

$$T_1 = T_{(\cdot,\cdot,1)}, \; T_2 = T_{(\cdot,\cdot,2)}, \text{etc}$$

Claim: If $\text{rank}(T) \leq r$ then for all a, $\text{rank}(T_a) \leq r$ too

This follows from the definition of rank, since

$$T = \sum_{i=1}^{r} u^{(i)} \otimes v^{(i)} \otimes w^{(i)} \implies$$

$$T_a = \sum_{i=1}^{r} w_a^{(i)} \left(u^{(i)} \otimes v^{(i)} \right)$$

However a low-rank tensor is not just a collection of arbitrary low rank matrices

Lemma: Consider a rank $\leq r$ tensor T with

$$T = \sum_{i=1}^{r} u^{(i)} \otimes v^{(i)} \otimes w^{(i)}$$

Then for all a we have

$$\text{colspan}(T_a) \subseteq \text{span}(\{u^{(i)}\})$$

$$\text{rowspan}(T_a) \subseteq \text{span}(\{v^{(i)}\})$$

Proof Left as exercise

Intuitively we have

matrix $\triangleq$ one "view" of vectors $\{u^{(i)}\}$ and $\{v^{(i)}\}$

tensor $\triangleq$ multiple "views"

and this is how we will solve the rotation problem

The Trouble with Tensors

Many of the properties we know and love about matrices will break for tensors, e.g.

For any matrix M, we have:

$$\text{rank}(M) = \dim(\text{colspan}(M)) = \dim(\text{rowspan}(M))$$

Does this type of relation hold for tensors? No!

Claim: The rank of a tensor depends on the field you are working over.

e.g. Consider

$$T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

It can be shown that

$$\text{rank}_{\mathbb{R}}(T) \geq 3$$

only real values are allowed for the factors

However $\text{rank}_{\mathbb{C}}(T) = 2$; in particular

$$T = \frac{1}{2}\left(\begin{bmatrix} 1 \\ -i \end{bmatrix} \otimes \begin{bmatrix} 1 \\ i \end{bmatrix} \otimes \begin{bmatrix} 1 \\ -i \end{bmatrix} + \begin{bmatrix} 1 \\ i \end{bmatrix} \otimes \begin{bmatrix} 1 \\ -i \end{bmatrix} \otimes \begin{bmatrix} 1 \\ i \end{bmatrix} \right)$$

Even though T is real-valued, can use fewer rank one tensors if we use complex numbers

Claim: There are tensors of rank 3, but which are arbitrarily close to tensors of rank 2

definition: The border rank of T is the minimum r s.t. $\forall \varepsilon > 0$ $\exists$ a tensor S of rank at most r s.t. T and S are ε-close entry-wise

e.g. consider $T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

It can be shown that $\text{rank}_{\mathbb{R}}(T) = 3$.

Yet it admits an arbitrarily good rank 2 approximation. Let

$$S_n = \begin{bmatrix} n & 1 \\ 1 & \frac{1}{n} \end{bmatrix}; \begin{bmatrix} 1 & \frac{1}{n} \\ \frac{1}{n} & \frac{1}{n^2} \end{bmatrix}$$

$$R_n = \begin{bmatrix} n & 0 \\ 0 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$\text{rank}_{\mathbb{R}}(S_n - R_n) = 2$ and yet

$S_n - R_n$ and T are $\frac{1}{n}$-close entry-wise

Exercise: Use Eckhart-Young to show that this cannot happen for matrices, i.e. rank and border rank are the same

For matrices, consider the best rank k approximation to M in Frobenius norm

$$M_k = \text{bestrank}_k(M)$$

Then we have

$$\text{bestrank}_{k-1}(M) = \text{bestrank}_{k-1}(M_k)$$

<u>Claim</u>: For tensors, the best rank $k-1$ approximation may not share any common factors with the best rank k approximation

For me, the root of those problems is computational

<u>Theorem</u> [Hastad] It is NP-hard to compute the rank of a tensor

So of course it cannot be equal to the dimension of the span of its rows, etc

[Hillar, Lim] showed a laundry list of tensor problems are NP-hard

"Most Tensor Problems are Hard"

We'll explore an important (previously forgotten) positive result and its myriad applications