New decomp method for tensors

$$\boxed{\begin{array}{l}\text{Tucker decomposition of a tensor} \\ \text{Higher order singular value decomposition.}\end{array}}$$

Applications: data compression for multiarrays.

#.

1) Recap SVD for matrices.

2) Concept of Multilinear Rank of a Tensor

3) HOSVD. and a theorem of Tucker.

#.

1) Recap of SVD:

$A \in \mathbb{R}^{M \times N}$. Always exist $U \in \mathbb{R}^{M \times M}$ and $V \in \mathbb{R}^{N \times N}$ that are orthogonal

$$\begin{pmatrix} U U^T = U^T U = I \\ V V^T = V^T V = I \end{pmatrix} \quad (*)$$

such that

$$A = U \underbrace{\Sigma}_{} V^T$$

$M \times N$ matrix of singular values

$$\Sigma = \left[\begin{array}{cccc|c} \sigma_1 & & & 0 & \\ & \sigma_2 & & & \\ & & \ddots & & 0 \\ 0 & & & \sigma_M & \end{array}\right]$$

R R

$M \leq N$

$\min(M, N)$

singular values: $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min(M,N)}$.

Remark: stated like this the SVD is not unique.

We will use in fact a restatement of SVD:

suppose $\operatorname{rank}(A) = R \leq \min(M, N)$

a fact: only R singular values that are non vanishing.

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_R > 0$$

$$\underset{\uparrow\atop M\times N}{A} = U_{M\times R}\, \Sigma_{R\times R}\, V^T_{R\times N}$$

$$\Sigma_{R\times R} = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_R \end{bmatrix}$$

$\left\{\begin{array}{l} U_{M\times R} \text{ given by first } R \text{ columns of } \bar{U}_{M\times M} \text{ (} \bar{U} \text{ before)} \\ V_{N\times R} \text{ given by first } R \text{ " " } \bar{V}_{N\times N} \end{array}\right.$

$U^T_{M\times R}\, U_{M\times R} = I_{R\times R}$ & idem for V.

Remark: Now if all $\sigma_1 \dots \sigma_R$ are distinct then the SVD is unique.

One can also write:

$$A = \sum_{r=1}^{R} \sigma_r \vec{u}_r \vec{v}_r^{T} \qquad (*)$$

$$\underset{\in \mathbb{R}^M}{[\vec{u}_1} \dots \vec{u}_R] = U_{M \times R}$$

$$\underset{\in \mathbb{R}^N}{[\vec{v}_1} \dots \vec{v}_R] = V_{N \times R}$$

singular values $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_R$.

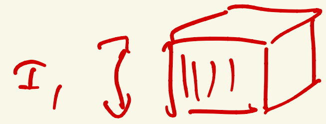
(Again if sing values are all distinct then $(*)$ is unique).

Theorem: Eckart-Young thm. (basis of dim Red)

$$\underset{\substack{\tilde{A} \\ \mathrm{rank}(\tilde{A}) \leq K}}{\mathrm{argmin}} \| A - \underset{\mathrm{rank}(\tilde{A}) \leq K}{\tilde{A}} \|_F \quad \tilde{A} = \sum_{r=1}^{K} \sigma_r \vec{u}_r \vec{v}_r^{T}.$$

where $K \leq R$.

$$\|M\|_F^2 = \sum_{i,j} |M_{ij}|^2.$$



2) Concept of Multilinear Rank of a Tensor.

- To fix ideas: order three tensor $T = \left(T^{\alpha\beta\gamma}\right)$ (but the whole discussion is general for any order p).

- Take the three matricizations $\underline{T_{(1)}}$; $\underline{T_{(2)}}$; $\underline{T_{(3)}}$

<u>recall</u>: $T^{\hat{\bullet}\,\widehat{\beta\gamma}}$ = fibers or vectors (column) with I_1 components.

align the fibers $\rightarrow$ Matrix $T_{(1)}$: $\underbrace{I_1}\times\underbrace{I_2 I_3}$

$T^{\alpha\bullet\gamma}$ align the fibers $\rightarrow$ $T_{(2)} : I_2 \times I_1 I_3$

$T^{\widehat{\alpha\beta}\hat{\bullet}}$ fibers $\rightarrow$ $T_{(3)} : \underbrace{I_3} \times \underbrace{I_1 I_2}$

Say that usual matrix Ranks are

$$R_1 = \text{rank}\left(T_{(1)}\right) \quad R_2 = \text{rank}\left(T_{(2)}\right) \quad R_3 = \text{rank}\left(T_{(3)}\right)$$

<u>By Definition</u>: Multilinear Rank of T is

$$\text{rank}_{\boxplus}(T) = \{R_1, R_2, R_3\}$$

- What is relation with the "tensor rank"

is the <u>min</u> # of terms in a decomp of T into rank-one elementary tensors.

$$T = \sum_{r=1}^{R} \vec{a}_r \otimes \vec{b}_r \otimes \vec{c}_r .$$

$$\mathrm{rank}_{\otimes}(T) = R .$$

$$\begin{cases} \mathrm{rank}_{\boxplus}(T) = \{ \underline{R_1}, \underline{R_2}, R_3 \} \overset{?}{\longleftrightarrow} \mathrm{ran}_{\otimes}(T) = R \\ \text{inequality in the exercise session} \end{cases}$$

- <u>Remark</u>: $\mathrm{ran}_{\otimes}(T)$ is <u>difficult to compute</u>.

(Sometimes apply Sevrich's thm, but in general not always).

$\mathrm{rank}_{\boxplus}(T)$ is easy to compute by usual lin alg methods.

- <u>Finally</u>: <u>For matrices</u> $\mathrm{ran}_{\otimes}(\text{Matrix}) = \mathrm{ran}_{\boxplus}(\text{Matrix})$

order $p=2$ tensor $\quad R_1 = R_2 = R$

3) Statement of the Tucker decomposition of a tensor / Higher Order SVD for tensors.

Theorem. Let $T = (T^{\alpha\beta\gamma}) \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ multiarray s.t

$\mathrm{rank}_{\boxplus}(T) = \{R_1, R_2, R_3\}$. It is always possible to decompose T as follows:

$$T = \sum_{p=1}^{R_1} \sum_{q=1}^{R_2} \sum_{r=1}^{R_3} \underbrace{G_{pqr}} \, \vec{u}_p \otimes \vec{v}_q \otimes \vec{w}_r$$

where $[\vec{u}_1 \ldots \vec{u}_{R_1}]$ are orthogonal vectors $I_1 \times R_1$

$[\vec{v}_1 \ldots \vec{v}_{R_2}]$ idem $I_2 \times R_2$

$[\vec{w}_1 \ldots \vec{w}_{R_3}]$ idem $I_3 \times R_3$

and G is an order 3 tensor in $\mathbb{R}^{R_1 \times R_2 \times R_3}$.

* G is not diagonal.  called the core tensor

$R_1 \leq I_1$, $R_2 \leq I_2$ & $R_3 \leq I_3$.

* Kind of analogous to SVD for matrices. but here

$$M = \sum_{r=1}^{R} \sigma_r \vec{u}_r \otimes \vec{v}_r \equiv \vec{u}_r \vec{v}_r^T \qquad \Sigma = [\sigma_1 \ldots \sigma_R] \text{ is diag.}$$

Remarks.

1) The core tensor G is not diag for $p \geq 3$.

2) This decomp is NOT unique & there are an infinite number of them

↑ exercise session.

3) For matrices truncates of SVD give best low rank approx (Eckart-Young Thm).

But for Tensors truncates of HOSVD do not gives best low rank approx.

What happens here that is bad is that you can have sequence of rank K