New decomposed rank for tensors

Tucker decomposition of a tensor

Higher order singular value decomposition.

Applications: data compression for multi-arrays.

1) Recap SVD for matrices.

2) Concept of Multilinear Rank of a Tensor

3) HOSVD, and a Measure of Tucker.
1) Decompose SVD:

\[ A \in \mathbb{R}^{M \times N}, \text{ always exist } U \in \mathbb{R}^{M \times M} \]

and \( V \in \mathbb{R}^{N \times N} \) that form orthogonal

\[
\begin{pmatrix}
U & U^T = U^T U = I \\
V & V^T = V^T V = I
\end{pmatrix}
\]

such that

\[ A = U \Sigma V^T \]

\( M \times N \) matrix of singular values

\[
\Sigma = \begin{bmatrix}
\sigma_1 & \cdots & 0 \\
0 & \sigma_2 & \cdots \\
\vdots & \ddots & \vdots \\
0 & \cdots & 0
\end{bmatrix}
\]

\( M \leq N \)

Minimum \( \sigma_i \in \mathbb{C} \in (M, N) \)

Singular values: \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min(M, N)} \).

Remark: Scaled like this, \( M \) \( N \) SVD is not unique.
We will use in fact a restatement of SVD:

Suppose \( \text{rank}(A) = R \leq \min(r, n) \)

in fact: only \( R \) singular values \( \sigma_i \) are non-vanishing.

\[
\sigma_1 > \sigma_2 > \ldots > \sigma_R > 0
\]

\[
A = U \Sigma V^T
\]

\[
\Sigma_{R \times R} = \begin{bmatrix}
\sigma_1 & 0 & \cdots & 0 \\
0 & \sigma_2 & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & \cdots & \sigma_R
\end{bmatrix}
\]

\[
U_{M \times R} \text{ given by first } R \text{ of } A \\
V_{N \times R} \text{ given by first } R \\
U_{M \times N}^T U_{N \times R} = I_{R \times R} \text{ & idem for } V.
\]

Remark: Now if all \( \sigma_1, \ldots, \sigma_R \) are distinct then the SVD is unique.
One can also write:

\[ A = \sum_{r=1}^{R} \sigma_r u_r v_r^T \]  

\[
\begin{bmatrix}
\mathbf{u}_1 & \cdots & \mathbf{u}_r & \cdots & \mathbf{u}_R
\end{bmatrix} = U_{R \times M}^T
\]

\[
\begin{bmatrix}
\mathbf{v}_1 & \cdots & \mathbf{v}_r & \cdots & \mathbf{v}_R
\end{bmatrix} = V_{R \times N}^T
\]

Theorem: Eckart–Young Thm. (best of dim. Red)

\[
\arg\min_{\tilde{A}} \| \mathbf{A} - \tilde{A} \|_F \text{ and } \tilde{A} = \sum_{r=1}^{K} \sigma_r \mathbf{u}_r \mathbf{v}_r^T
\]

\[
\text{rank}(\tilde{A}) \leq K \quad \text{rank}({A}) \leq K \quad \text{where } K \leq R.
\]

\[
\| \mathbf{M} \|_F^2 = \sum_{i,j} \mathbf{M}_{ij}^2.
\]
2) **Concept of Multilinear Rank of a Tensord**

- To fix ideas: order three term $\mathbf{T} = (\mathbf{T})$

  (but the whole discourse is general for any order $p$).

- Take the three matricizations $\mathbf{T}_{(1)}$, $\mathbf{T}_{(2)}$, $\mathbf{T}_{(3)}$

  recall: $\mathbf{T} = \{\text{fibers or vectors (column)}\} \text{ with } n_1 \text{ components}$

  align the fibers $\rightarrow$ Matrix $\mathbf{T}_{(1)}: I_1 \times I_2 I_3$

  $\mathbf{T} \times \mathbf{T}$ align the fibers $\rightarrow$ $\mathbf{T}_{(2)}: I_2 \times I_1 I_3$

  $\mathbf{T} \times \mathbf{T}$ fibers $\rightarrow$ $\mathbf{T}_{(3)}: I_3 \times I_1 I_2$

Say that usual matrix ranks are

\[ R_1 = \text{rank}(\mathbf{T}_{(1)}) \quad R_2 = \text{rank}(\mathbf{T}_{(2)}) \quad R_3 = \text{rank}(\mathbf{T}_{(3)}) \]

By definition, multilinear rank of $\mathbf{T}$ is

\[ \text{rank}_M(\mathbf{T}) = \{ R_1, R_2, R_3 \} \]
What is related with the "tensor rank" 

\[ \text{rank}_r(T) = \min \{ \text{# of terms in a decom of } T \text{ into rank-}r \text{ elementary tensors} \} \]

\[ T = \sum_{r=1}^{R} a_r \otimes b_r \otimes c_r. \]

\[ \text{rank}_r(T) = R. \]

\[ \begin{align*}
\text{rank}_r(T) & = \left\{ R_1, R_2, R_3 \right\} \\
\text{inequality in the exercise space} & \end{align*} \]

Remark: \( \text{rank}_r(T) \) is difficult to compute.

(Sometimes applies Schur's Theorem, but in general not always.)

\[ \text{rank}_r(T) \text{ is easy to compute by } \]

\[ \text{and lin alg methods.} \]

Finally: For matrices \( \text{rank}_p(M) = \text{rank}_p(M^\top) \) \( \text{order } p = 2 \text{ tensor} \) \( R_1 = R_2 = R_3 \).
3) Statement of the Tucker decomposability of a tensor / Higher Order SVD for tensors.

Theorem. Let \( T = (T_{\alpha \beta \gamma}) \in \mathbb{R}^{I_1 \times I_2 \times I_3} \) a multi-array s.t.

\[
\text{rank}_{\mathcal{T}}(T) = \{ R_1, R_2, R_3 \}.
\]

It is always possible to decompose \( T \) as

\[
T = \sum_{p=1}^{R_1} \sum_{q=1}^{R_2} \sum_{r=1}^{R_3} G_{pq r} \mu_p \otimes \nu_q \otimes \omega_r,
\]

where

\[
\begin{align*}
\begin{bmatrix} \mu_1 & \cdots & \mu_{R_1} \end{bmatrix} & \text{ are orthogonal vectors } I_1 \times R_1, \\
\begin{bmatrix} \nu_1 & \cdots & \nu_{R_2} \end{bmatrix} & \text{ idem } I_2 \times R_2, \\
\begin{bmatrix} \omega_1 & \cdots & \omega_{R_3} \end{bmatrix} & \text{ idem } I_3 \times R_3
\end{align*}
\]

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

\[
H = \sum_{r=1}^{R} \sigma_r \mathbf{u}_r \otimes \mathbf{v}_r \otimes \mathbf{w}_r \quad \text{and} \quad T = \begin{bmatrix} \sigma_1 \mathbf{u}_1 \otimes \mathbf{v}_1 \otimes \mathbf{w}_1 & \cdots \end{bmatrix}.\]
1) The core bank $G$ is not direct for $p > 3$.

2) This claim is NOT unique & hence an infinite number of them.

3) For another harmonic of SPD, the best low rank approx (Eckart-Young Theorem).

But for $p = 2$, its harmonic is not unique.