ENCODING ORTHONORMAL TENSOR DECOMPOSITION AS SVD

Theorem - 1 : $u_1, u_2 \in \mathbb{R}^m$ $v_1, v_2 \in \mathbb{R}^n$ and $u_1 \perp u_2$, then

i) $\langle A, B \rangle = Tr(A^T B) = 0$ where $A = u_1 v_1^T$, $B = u_2 v_2^T$

ii) $\langle C, D \rangle = Tr(C^T D) = 0$ where $C = v_1 u_1^T$, $D = v_2 u_2^T$

Proof: $u_1 \perp u_2 \iff u_1^T u_2 = 0 = u_2^T u_1$

i) $\langle A, B \rangle = \langle u_1 v_1^T, u_2 v_2^T \rangle = Tr((u_1 v_1^T)^T u_2 v_2^T)$

$$= Tr(v_1 \underbrace{u_1^T u_2}_{0} v_2^T) = Tr(v_1 \cdot 0 \cdot v_2^T) = Tr(0) = 0$$

ii) $\langle C, D \rangle = \langle A^T, B^T \rangle = Tr(A^{TT} B^T) = Tr(A B^T)$

$$= Tr(u_1 v_1^T (u_2 v_2^T)^T) = Tr(u_1 v_1^T v_2 u_2^T)$$

$$\qquad\qquad\qquad\qquad\qquad\qquad \Vert$$

$$0 = Tr(0) = Tr(\underbrace{u_2^T u_1}_{0} v_1^T v_2)$$

Alternatively, notice $\langle A^T, B^T \rangle = \langle A, B \rangle$.

Orthonormal CP - decomposition

Let $T \in \mathbb{R}^{m \times n \times p}$, then $T = \sum_i^r \alpha_i (\vec{a}_i \circ \vec{b}_i \circ \vec{c}_i)$ is called the rank-r orthonormal CP-decomposition of T iff $a_i \perp a_j \ \forall i \neq j$, $c_i \perp c_j \ \forall i \neq j$, and $\|a_i\|_2 = 1$, $\|b_i\|_2 = 1$ and $\|c_i\|_2 = 1$.

This can be said more compactly as $c_i^T c_j = \delta_{ij} = a_i^T a_j$

where $\delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & \text{otherwise} \end{cases}$

Theorem - 2 :- Let $\vec{u}$ & $\vec{v}$ be in $\mathbb{R}^m$ & $\mathbb{R}^n$ respectively.

Let $u^T \cdot u = 1 = v^T \cdot v$ Then $\|A\|_F = 1 = \|vec(A)\|_2$

Proof :- $A = UV^T$, $\|A\|_F^2 = \langle A, A\rangle = Tr(A^TA) = Tr((UV^T)^T UV^T)$

$$1 = Tr(\underbrace{V^TV}_{1}) = Tr(VV^T) = Tr(V\underbrace{U^TU}_{1}V^T)$$

Rank r - orthonormal CP decomposition $\equiv$ rank r SVD

Let $T = \sum_{i=1}^{r} \alpha_i (\vec{a}_i \circ \vec{b}_i \circ \vec{c}_i)$ be the orthonormal CP decomposition of $T \in \mathbb{R}^{m\times n\times p}$. Let $A_i = a_i \circ b_i = a_i b_i^T$. Let $\vec{z}_i = vec(A_i)$

$$z_i^T z_j = \langle A_i, A_j\rangle = 0 \text{ as } a_i \perp a_j \text{ (theorem-1)}$$

and $\|A_i\|_F^2 = \|z_i\|_2^2 = \langle A_i, A_i\rangle = 1$

Hence $z_i^T z_j = \delta_{ij} = \begin{cases} 1, & \text{if } i=j \\ 0, & \text{otherwise} \end{cases}$

Let $\mathbb{R}^{mn} \ni dflate(T)_{:,i} := vec(T_{:,:,i})$.

Such that $M = dflate(T) \in \mathbb{R}^{mn\times p}$

Corollary

$dflate(a\circ b\circ c) = vec(a \circ b) \circ c$

Linearity of vec(·) & dflate(·):

It isn't hard to see $vec(\alpha A + \beta B) = \alpha\, vec(A) + \beta\, vec(B)$

& $dflate(\alpha_1 T_1 + \alpha_2 T_2) = \alpha_1 dflate(T_1) + \alpha_2 dflate(T_2)$

$\forall\, \alpha, \beta \in \mathbb{R}$

$$M = dflate(T) = dflate\left(\sum_{i=1}^{r} \alpha_i a_i \circ b_i \circ c_i\right) = \sum_{i=1}^{r} \alpha_i dflate(a_i \circ b_i \circ c_i)$$

$$= \sum_{i=1}^{r} \alpha_i\, vec(a_i \circ b_i) \circ c_i = \sum_{i=1}^{r} \alpha_i\, vec(A_i) \circ c_i = \sum_{i=1}^{r} \alpha_i z_i \circ c_i$$

$$\| \qquad \sum_{i=1}^{r} \alpha_i z_i c_i^T$$

where $Z = \begin{bmatrix} | \\ z_i \\ | \end{bmatrix}$

$\alpha = diag([\alpha_i])$

$C = \begin{bmatrix} | \\ c_i \\ | \end{bmatrix}$

$\Bigg\{ \; Z \alpha C^T = \sum_{i=1}^{r} \alpha_i z_i c_i^T$
(SVD)

MATRIX FACTORIZATION AS TENSOR DECOMPOSITION

Notations: Let $\vec{v} \in \mathbb{R}^{qp}$.

define $mat(v) \in \mathbb{R}^{q \times p}$ the matrix such that

$\vec{v} = vec(mat(\vec{v}))$ and $A = mat(vec(A))$ for $A \in \mathbb{R}^{q \times p}$

i.e. $mat = vec^{-1}$

$\Rightarrow$ Let $M \in \mathbb{R}^{mn \times p}$ be any matrix.

$\Rightarrow$ Define $T = nflate(M) \in \mathbb{R}^{m \times n \times p}$ such that

$$nflate(M)_{:,:,i} = mat(M_{:,i})$$

$\Rightarrow$ Clearly nflate is linear, i.e. $nflate(\gamma_1 M_1 + \gamma_2 M_2) = \gamma_1\, nflate(M_1) + \gamma_2\, nflate(M_2)$

or $nflate\left(\sum_i \gamma_i M_i\right) = \sum_i \gamma_i\, nflate(M_i)$ for $\gamma_i \in \mathbb{R}$.

Remark: $dflate(nflate(M))_{:,i} = vec(nflate(M)_{:,:,i})$

$= vec(mat(M_{:,i})) = M_{:,i}$

Hence, $dflate(nflate(M)) = M$

Similarly $nflate(dflate(T))_{:,:,i} = mat(dflate(T)_{:,i})$

$= mat(vec(T_{:,:,i}))$

$= T_{:,:,i}$

Hence, $nflate(dflate(T)) = T$

So, we have $nflate = dflate^{-1}$

Corollary: If $\vec{w} = vec(\vec{x} \circ \vec{y})$, then $nflate(\vec{w} \circ \vec{z}) = \vec{x} \circ \vec{y} \circ \vec{z}$

$$nflate(vec(\vec{x} \circ \vec{y}) \circ \vec{z}) = \vec{x} \circ \vec{y} \circ \vec{z}$$

Matrix Factorization :- Let $M \in \mathbb{R}^{mn \times p}$ be a matrix such that

$M = \sum_{i=1}^{r} d_i e_i^T = \sum_{i=1}^{r} \vec{d}_i \circ \vec{e}_i$ any rank-r factorization of M such that $\vec{d}_i \in \mathbb{R}^{mn}$ & $\vec{e}_i \in \mathbb{R}^p$

Let $D_i = \text{mat}(\vec{d}_i) \in \mathbb{R}^{m \times n} \Rightarrow \text{vec}(D_i) = \vec{d}_i$

Let $D_i = \sum_{j=1}^{k} a_j b_j^T = \sum_{j=1}^{k} \vec{a}_j \circ \vec{b}_j$ be any rank-k factorization of D_i

Hence
$$M = \sum_{i=1}^{r} \vec{d}_i \circ \vec{e}_i = \sum_{i=1}^{r} \text{vec}(D_i) \circ \vec{e}_i$$
$$= \sum_{i=1}^{r} \text{vec}\left(\sum_{j=1}^{k} \vec{a}_j \circ \vec{b}_j\right) \circ \vec{e}_i$$
$$= \sum_{i=1}^{r} \sum_{j=1}^{k} \text{vec}(a_j \circ b_j) \circ \vec{e}_i$$

$$\Rightarrow \underbrace{\text{nflate}(M)}_{T} = \text{nflate}\left(\sum_{i=1}^{r} \sum_{j=1}^{k} \text{vec}(a_j \circ b_j) \circ e_i\right)$$
$$T = \sum_{i=1}^{r} \sum_{j=1}^{k} \text{nflate}[\text{vec}(a_j \circ b_j) \circ e_i)]$$
$$T = \sum_{i=1}^{r} \sum_{j=1}^{k} \vec{a}_j \circ \vec{b}_j \circ \vec{e}_i \quad \because \text{corollary on prev. page.}$$

Let $s = rk$, then $T = \sum_{s=1}^{rk} \vec{a}_{\lceil \frac{s}{r} \rceil} \circ \vec{b}_{\lceil \frac{s}{r} \rceil} \circ \vec{e}_{\lceil \frac{s}{k} \rceil}$

where $\lceil \cdot \rceil$ is the ceil function or the "smallest integer value" function.

$\text{nflate}(M) = T = \sum_{s=1}^{rk} \vec{a}_{\lceil \frac{s}{r} \rceil} \circ \vec{b}_{\lceil \frac{s}{r} \rceil} \circ \vec{e}_{\lceil \frac{s}{k} \rceil}$ is a rank 'rk' tensor decomposition/factorization of T where all of the tensor atoms are unique but the factors $\vec{a}_i$, $\vec{b}_i$ & $\vec{c}_i$ are sometimes repeated and chosen out of k of $\vec{a}$s / $\vec{b}$s and r of $\vec{e}$s.