Multi-linear Algebra
– Hierarchical Tucker Decomposition –
Lecture 19

F. M. Faulstich

02/04/2024

Recall

Recall

- CP decomposition: Let $\mathbf{A} \in \mathbb{R}^{n_1 \times \ldots \times n_d}$. Then

$$\mathbf{A} = \sum_{p=1}^{r} \bigotimes_{i=1}^{d} \mathbf{v}_{i,p}$$

Storage of CP format:
CP rank:

Recall

- CP decomposition: Let $\mathbf{A} \in \mathbb{R}^{n_1 \times \ldots \times n_d}$. Then

$$\mathbf{A} = \sum_{p=1}^{r} \bigotimes_{i=1}^{d} \mathbf{v}_{i,p}$$

Storage of CP format: $\mathcal{O}(rnd)$
CP rank: minimal r s.t. we can express $\mathbf{A}$ in the above format

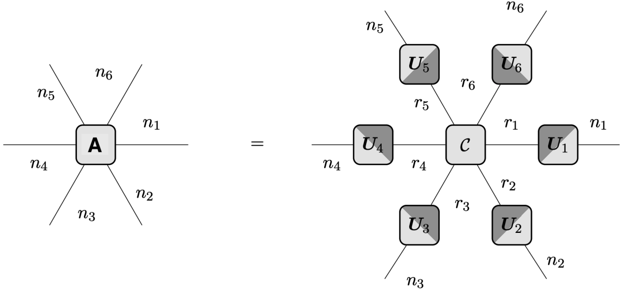
Recall

- Tucker decomposition: Let $\mathbf{A} \in \mathbb{R}^{n_1 \times \ldots \times n_d}$. Then

$$\begin{aligned}\mathbf{A} &= \sum_{i_1=1}^{r_1} \cdots \sum_{i_d=1}^{r_d} \mathbf{C}[i_1, ..., i_d] \cdot \mathbf{u}_{1,i_1} \otimes \mathbf{u}_{2,i_2} \otimes \cdots \otimes \mathbf{u}_{d,i_1} \\ &= \mathbf{C} *_1 \mathbf{U}_1 *_2 \mathbf{U}_2 ... *_d \mathbf{U}_d\end{aligned}$$

Recall

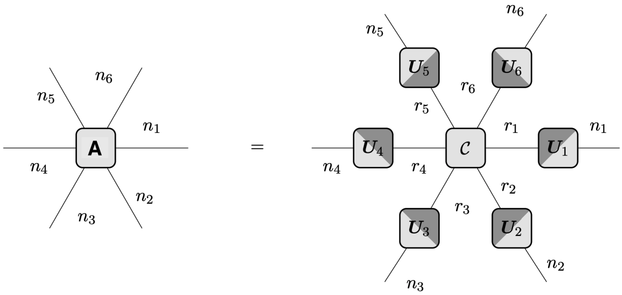
- Tucker decomposition: Let $\mathbf{A} \in \mathbb{R}^{n_1 \times \ldots \times n_d}$. Then



Storage of Tucker format:
T-rank:

Recall

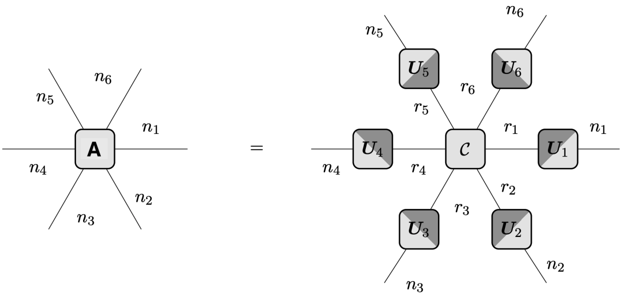
- Tucker decomposition: Let $\mathbf{A} \in \mathbb{R}^{n_1 \times \ldots \times n_d}$. Then



Storage of Tucker format: $\mathcal{O}(r^d + rnd)$
T-rank: $\mathbf{r} = (r_1, \ldots, r_d)$
Advantage:

Recall

- Tucker decomposition: Let $\mathbf{A} \in \mathbb{R}^{n_1 \times \dots \times n_d}$. Then



Storage of Tucker format: $\mathcal{O}(r^d + rnd)$

Advantage:

1. Can be computed using HOSVD
2. Closed set of low-rank tensors
3. Manifold structure on the set of tensors with fixed rank
4. Can be sketched

Recall TT format

- TT decomposition: Let $\mathbf{A} \in \mathbb{R}^{n_1 \times \ldots \times n_d}$. Then

$$\begin{aligned}
\mathbf{A} &= \mathbf{U}_1 \circ \mathbf{U}_d \circ \ldots \circ \mathbf{U}_d \\
&= \mathbf{U}_1 *_{2,1} \mathbf{U}_d *_{3,1} \ldots *_{3,1} \mathbf{U}_d \\
&= \sum_{k_1=1}^{r_1} \ldots \sum_{k_{d-1}=1}^{r_{d-1}} \mathbf{U}_1[:, k_1]\mathbf{U}_2[k_1, :, k_2] \cdots \mathbf{U}_{d-1}[k_{d-2}, :, k_{d-1}]\mathbf{U}_d[k_{d-1}, :]
\end{aligned}$$

Recall TT format

- TT decomposition: Let $\mathbf{A} \in \mathbb{R}^{n_1 \times \ldots \times n_d}$. Then

$$\begin{aligned}\mathbf{A} &= \mathbf{U}_1 \circ \mathbf{U}_d \circ \ldots \circ \mathbf{U}_d \\ &= \mathbf{U}_1 *_{2,1} \mathbf{U}_d *_{3,1} \ldots *_{3,1} \mathbf{U}_d \\ &= \sum_{k_1=1}^{r_1} \ldots \sum_{k_{d-1}=1}^{r_{d-1}} \mathbf{U}_1[:, k_1]\mathbf{U}_2[k_1, :, k_2] \cdots \mathbf{U}_{d-1}[k_{d-2}, :, k_{d-1}]\mathbf{U}_d[k_{d-1}, :]\end{aligned}$$

or elementwise

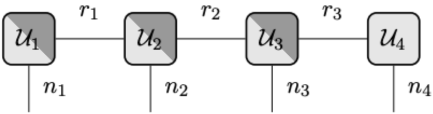
$$\begin{aligned}&\mathbf{A}[i_1, \ldots, i_d] \\ &= \sum_{k_1=1}^{r_1} \ldots \sum_{k_{d-1}=1}^{r_{d-1}} \mathbf{U}_1[i_1, k_1]\mathbf{U}_2[k_1, i_2, k_2] \cdots \mathbf{U}_{d-1}[k_{d-2}, i_{d-1}, k_{d-1}]\mathbf{U}_d[k_{d-1}, i_d]\end{aligned}$$

Recall TT format

- TT decomposition: Let $\mathbf{A} \in \mathbb{R}^{n_1 \times \ldots \times n_d}$. Then

$$
\begin{aligned}
\mathbf{A} &= \mathbf{U}_1 \circ \mathbf{U}_d \circ \ldots \circ \mathbf{U}_d \\
&= \mathbf{U}_1 *_{2,1} \mathbf{U}_d *_{3,1} \ldots *_{3,1} \mathbf{U}_d \\
&= \sum_{k_1=1}^{r_1} \ldots \sum_{k_{d-1}=1}^{r_{d-1}} \mathbf{U}_1[:, k_1] \mathbf{U}_2[k_1, :, k_2] \cdots \mathbf{U}_{d-1}[k_{d-2}, :, k_{d-1}] \mathbf{U}_d[k_{d-1}, :]
\end{aligned}
$$

or in diagrams



TT-SVD

$\mathbf{A}_{n_1,\ldots,n_d}$

TT-SVD

$$\mathbf{A}_{n_1,\ldots,n_d}$$

$$= \mathbf{A}^{n_1}_{n_2\cdots n_d} \qquad \text{reshape to } n_1 \times \prod_{j\neq i} n_j$$

TT-SVD

$$
\begin{aligned}
&\mathbf{A}_{n_1,\ldots,n_d} \\
&= \mathbf{A}^{n_1}_{n_2\cdots n_d} && \text{reshape to } n_1 \times \prod_{j\neq i} n_j \\
&= (\mathbf{U}_1)^{n_1}_{r_1} \, (\boldsymbol{\Sigma}_1 \mathbf{V}_1^\top)^{r_1}_{n_2\cdots n_d} && \text{SVD}
\end{aligned}
$$

TT-SVD

$$
\begin{aligned}
&\mathbf{A}_{n_1,\ldots,n_d} \\
&= \mathbf{A}^{n_1}_{n_2\cdots n_d} && \text{reshape to } n_1 \times \prod_{j\neq i} n_j \\
&= (\mathbf{U}_1)^{n_1}_{r_1}\,(\mathbf{\Sigma}_1\mathbf{V}_1^\top)^{r_1}_{n_2\cdots n_d} && \text{SVD} \\
&= (\mathbf{U}_1)^{n_1}_{r_1}\,(\mathbf{\Sigma}_1\mathbf{V}_1^\top)^{r_1\cdot n_2}_{n_3\cdots n_d} && \text{reshape of } (\mathbf{\Sigma}_1\mathbf{V}_1^\top)
\end{aligned}
$$

TT-SVD

$$
\begin{aligned}
&\mathbf{A}_{n_1,\ldots,n_d} \\
&= \mathbf{A}^{n_1}_{n_2\cdots n_d} && \text{reshape to } n_1 \times \prod_{j\neq i} n_j \\
&= (\mathbf{U}_1)^{n_1}_{r_1}\,(\mathbf{\Sigma}_1\mathbf{V}_1^\top)^{r_1}_{n_2\cdots n_d} && \text{SVD} \\
&= (\mathbf{U}_1)^{n_1}_{r_1}\,(\mathbf{\Sigma}_1\mathbf{V}_1^\top)^{r_1\cdot n_2}_{n_3\cdots n_d} && \text{reshape of } (\mathbf{\Sigma}_1\mathbf{V}_1^\top) \\
&= (\mathbf{U}_1)^{n_1}_{r_1}\,(\mathbf{U}_2)^{r_1\cdot n_2}_{r_2}\,(\mathbf{\Sigma}_2\mathbf{V}_2^\top)^{r_2}_{n_3\cdots n_d} && \text{SVD of } (\mathbf{\Sigma}_1\mathbf{V}_1^\top)
\end{aligned}
$$

TT-SVD

$$
\begin{aligned}
&\mathbf{A}_{n_1,\dots,n_d} && \\
&= \mathbf{A}^{n_1}_{n_2\cdots n_d} && \text{reshape to } n_1 \times \prod_{j\neq i} n_j \\
&= (\mathbf{U}_1)^{n_1}_{r_1}\,(\mathbf{\Sigma}_1\mathbf{V}_1^\top)^{r_1}_{n_2\cdots n_d} && \text{SVD} \\
&= (\mathbf{U}_1)^{n_1}_{r_1}\,(\mathbf{\Sigma}_1\mathbf{V}_1^\top)^{r_1\cdot n_2}_{n_3\cdots n_d} && \text{reshape of } (\mathbf{\Sigma}_1\mathbf{V}_1^\top) \\
&= (\mathbf{U}_1)^{n_1}_{r_1}\,(\mathbf{U}_2)^{r_1\cdot n_2}_{r_2}\,(\mathbf{\Sigma}_2\mathbf{V}_2^\top)^{r_2}_{n_3\cdots n_d} && \text{SVD of } (\mathbf{\Sigma}_1\mathbf{V}_1^\top) \\
&= (\mathbf{U}_1)^{n_1}_{r_1}\,(\mathbf{U}_2)^{r_1\cdot n_2}_{r_2}\,(\mathbf{\Sigma}_2\mathbf{V}_2^\top)^{r_2\cdot n_3}_{n_4\cdots n_d} && \text{reshape of } (\mathbf{\Sigma}_2\mathbf{V}_2^\top)
\end{aligned}
$$

TT-SVD

$$
\begin{aligned}
&\mathbf{A}_{n_1,\ldots,n_d} \\
&= \mathbf{A}^{n_1}_{n_2\cdots n_d} && \text{reshape to } n_1 \times \prod_{j\neq i} n_j \\
&= (\mathbf{U}_1)^{n_1}_{r_1}\,(\boldsymbol{\Sigma}_1\mathbf{V}_1^\top)^{r_1}_{n_2\cdots n_d} && \text{SVD} \\
&= (\mathbf{U}_1)^{n_1}_{r_1}\,(\boldsymbol{\Sigma}_1\mathbf{V}_1^\top)^{r_1\cdot n_2}_{n_3\cdots n_d} && \text{reshape of } (\boldsymbol{\Sigma}_1\mathbf{V}_1^\top) \\
&= (\mathbf{U}_1)^{n_1}_{r_1}\,(\mathbf{U}_2)^{r_1\cdot n_2}_{r_2}\,(\boldsymbol{\Sigma}_2\mathbf{V}_2^\top)^{r_2}_{n_3\cdots n_d} && \text{SVD of } (\boldsymbol{\Sigma}_1\mathbf{V}_1^\top) \\
&= (\mathbf{U}_1)^{n_1}_{r_1}\,(\mathbf{U}_2)^{r_1\cdot n_2}_{r_2}\,(\boldsymbol{\Sigma}_2\mathbf{V}_2^\top)^{r_2\cdot n_3}_{n_4\cdots n_d} && \text{reshape of } (\boldsymbol{\Sigma}_2\mathbf{V}_2^\top) \\
&= (\mathbf{U}_1)^{n_1}_{r_1}\,(\mathbf{U}_2)^{r_1\cdot n_2}_{r_2}\,(\mathbf{U}_3)^{r_2\cdot n_3}_{r_3}\,(\boldsymbol{\Sigma}_3\mathbf{V}_3^\top)^{r_3}_{n_4\cdots n_d} && \text{SVD of } (\boldsymbol{\Sigma}_2\mathbf{V}_2^\top)
\end{aligned}
$$

TT-SVD

$$
\begin{aligned}
&\mathbf{A}_{n_1,\ldots,n_d} \\
&= \mathbf{A}^{n_1}_{n_2\cdots n_d} && \text{reshape to } n_1 \times \prod_{j\neq i} n_j \\
&= (\mathbf{U}_1)^{n_1}_{r_1}\,(\boldsymbol{\Sigma}_1\mathbf{V}_1^\top)^{r_1}_{n_2\cdots n_d} && \text{SVD} \\
&= (\mathbf{U}_1)^{n_1}_{r_1}\,(\boldsymbol{\Sigma}_1\mathbf{V}_1^\top)^{r_1\cdot n_2}_{n_3\cdots n_d} && \text{reshape of } (\boldsymbol{\Sigma}_1\mathbf{V}_1^\top) \\
&= (\mathbf{U}_1)^{n_1}_{r_1}\,(\mathbf{U}_2)^{r_1\cdot n_2}_{r_2}\,(\boldsymbol{\Sigma}_2\mathbf{V}_2^\top)^{r_2}_{n_3\cdots n_d} && \text{SVD of } (\boldsymbol{\Sigma}_1\mathbf{V}_1^\top) \\
&= (\mathbf{U}_1)^{n_1}_{r_1}\,(\mathbf{U}_2)^{r_1\cdot n_2}_{r_2}\,(\boldsymbol{\Sigma}_2\mathbf{V}_2^\top)^{r_2\cdot n_3}_{n_4\cdots n_d} && \text{reshape of } (\boldsymbol{\Sigma}_2\mathbf{V}_2^\top) \\
&= (\mathbf{U}_1)^{n_1}_{r_1}\,(\mathbf{U}_2)^{r_1\cdot n_2}_{r_2}\,(\mathbf{U}_3)^{r_2\cdot n_3}_{r_3}\,(\boldsymbol{\Sigma}_3\mathbf{V}_3^\top)^{r_3}_{n_4\cdots n_d} && \text{SVD of } (\boldsymbol{\Sigma}_2\mathbf{V}_2^\top) \\
&\ \vdots \\
&= \underbrace{(\mathbf{U}_1)^{n_1}_{r_1}}_{\mathbf{U}_1[n_d]}\cdots\underbrace{(\mathbf{U}_{d-1})^{r_{d-2}\cdot n_{d-1}}_{r_{d-1}}}_{=:\mathbf{U}_{d-1}[n_{d-1}]}\underbrace{(\boldsymbol{\Sigma}_{d-1}\mathbf{V}_{d-1}^\top)^{r_{d-1}}_{n_d}}_{=:\mathbf{U}_d[n_d]}
\end{aligned}
$$

HT decomposition

- Similar idea as TT format:
  $\rightarrow$ instead of using the nested subspace use a hierarchy of subspaces
- Define a partition tree for the set of mode indices:
  $\rightarrow$ the root of the tree contains the complete set
  $\rightarrow$ each leaf contains a single mode index
  $\rightarrow$ each inner node of the tree contains the union of its children

Tree tensor network

What is a partition tree for the set of mode indices?

Tree tensor network

What is a partition tree for the set of mode indices?

- It is a hierarchical tree
- TT format:

$\mathbb{R}^{n_1 \times n_2 \times n_3 \times n_4}$	$\{1,2,3,4\}$
$U_1 \otimes \mathbb{R}^{n_2 \times n_3 \times n_4}$	$\{1\}$ $\{2,3,4\}$
$U_2 \otimes \mathbb{R}^{n_3 \times n_4}$	$\{2\}$ $\{3,4\}$
$U_3 \otimes \mathbb{R}^{n_4}$	$\{3\}$ $\{4\}$

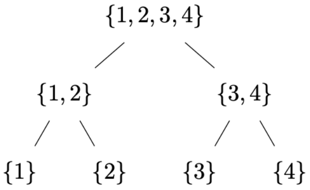
Balanced tree tensor network

- Consider $\mathbf{U} \in \mathbb{R}^{n_1 \times n_2 \times n_3 \times n_4}$
- Balanced tree of mode indices:

$$\mathbb{R}^{n_1 \times n_2 \times n_3 \times n_4}$$

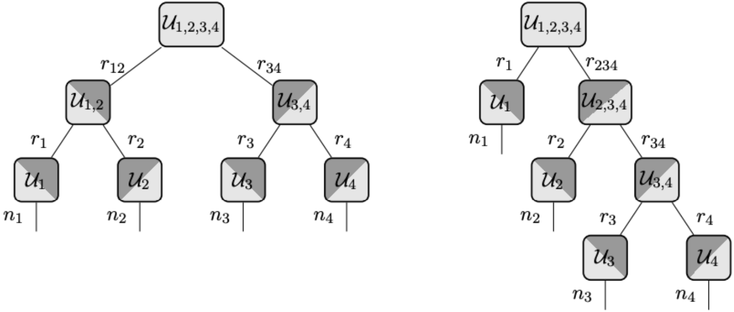
$$\tilde{U}_1 \otimes \tilde{U}_2, \quad \tilde{U}_1 \subseteq \mathbb{R}^{n_1 \times n_2},\ \tilde{U}_2 \subseteq \mathbb{R}^{n_3 \times n_4}$$

$$U_1 \otimes U_2 \otimes U_3 \otimes U_4$$



Tree tensor networks in diagrams

- Compare TT vs HT:



Recall matricization

- For a tensor $\mathbf{A} \in \mathbb{R}^{n_1 \times \ldots \times n_d}$, a collection of dimension indices $t \subset \{1, ..., d\}$, and its complement $s = \{1, ..., d\} \setminus t$, we define

$$\mathbf{A}^{(t)} \in \mathbb{R}^{n_t \times n_s}, \quad n_t = \bigtimes_{k \in t} n_k, \; n_s = \bigtimes_{k \in s} n_k$$

  elementwise defined as

$$[\mathbf{A}^{(t)}]_{(i_k)_{k \in t}, (i_\ell)_{\ell \in s}} = \mathbf{A}_{i_1, \ldots, i_d}$$

Matricization example

Consider

$$\mathbf{A}_{i_1,i_2,i_3,i_4} = i_1 + 2(i_1 - 1) + 4(i_3 - 1) + 8(i_4 - 1), \quad i_1, i_2, i_3, i_4 \in \{1, 2\}$$

We chose the lexicographical ordering for compound indices

$$(i_p, ..., i_q) \mapsto \ell = i_p + n_p(i_{p+1} - 1) + n_p n_{p+1}(i_{p+2} - 1) + ... + i_q n_p \cdots n_{q-1}$$

Compute

$$\mathbf{A}^{(\{1\})} =$$

Matricization example

Consider

$$\mathbf{A}_{i_1,i_2,i_3,i_4} = i_1 + 2(i_1 - 1) + 4(i_3 - 1) + 8(i_4 - 1), \quad i_1, i_2, i_3, i_4 \in \{1, 2\}$$

We chose the lexicographical ordering for compound indices

$$(i_p, ..., i_q) \mapsto \ell = i_p + n_p(i_{p+1} - 1) + n_p n_{p+1}(i_{p+2} - 1) + ... + i_q n_p \cdots n_{q-1}$$

Compute

$$\mathbf{A}^{(\{1\})} = \begin{bmatrix} 1 & 3 & 5 & 7 & 9 & 11 & 13 & 15 \\ 2 & 4 & 6 & 8 & 10 & 12 & 14 & 16 \end{bmatrix}$$

Matricization example

Consider

$$\mathbf{A}_{i_1,i_2,i_3,i_4} = i_1 + 2(i_1 - 1) + 4(i_3 - 1) + 8(i_4 - 1), \quad i_1, i_2, i_3, i_4 \in \{1, 2\}$$

We chose the lexicographical ordering for compound indices

$$(i_p, ..., i_q) \mapsto \ell = i_p + n_p(i_{p+1} - 1) + n_p n_{p+1}(i_{p+2} - 1) + ... + i_q n_p \cdots n_{q-1}$$

Compute

$$\mathbf{A}^{(\{3\})} =$$

Matricization example

Consider

$$\mathbf{A}_{i_1,i_2,i_3,i_4} = i_1 + 2(i_1 - 1) + 4(i_3 - 1) + 8(i_4 - 1), \quad i_1, i_2, i_3, i_4 \in \{1, 2\}$$

We chose the lexicographical ordering for compound indices

$$(i_p, ..., i_q) \mapsto \ell = i_p + n_p(i_{p+1} - 1) + n_p n_{p+1}(i_{p+2} - 1) + ... + i_q n_p \cdots n_{q-1}$$

Compute

$$\mathbf{A}^{(\{3\})} = \begin{bmatrix} 1 & 2 & 3 & 4 & 9 & 10 & 11 & 12 \\ 5 & 6 & 7 & 8 & 13 & 14 & 15 & 16 \end{bmatrix}$$

Matricization example

Consider

$$\mathbf{A}_{i_1,i_2,i_3,i_4} = i_1 + 2(i_1 - 1) + 4(i_3 - 1) + 8(i_4 - 1), \quad i_1, i_2, i_3, i_4 \in \{1, 2\}$$

We chose the lexicographical ordering for compound indices

$$(i_p, ..., i_q) \mapsto \ell = i_p + n_p(i_{p+1} - 1) + n_p n_{p+1}(i_{p+2} - 1) + ... + i_q n_p \cdots n_{q-1}$$

Compute

$$\mathbf{A}^{(\{2,3,4\})} =$$

Matricization example

Consider

$$\mathbf{A}_{i_1,i_2,i_3,i_4} = i_1 + 2(i_1 - 1) + 4(i_3 - 1) + 8(i_4 - 1), \quad i_1, i_2, i_3, i_4 \in \{1, 2\}$$

We chose the lexicographical ordering for compound indices

$$(i_p, ..., i_q) \mapsto \ell = i_p + n_p(i_{p+1} - 1) + n_p n_{p+1}(i_{p+2} - 1) + ... + i_q n_p \cdots n_{q-1}$$

Compute

$$\mathbf{A}^{(\{2,3,4\})} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \\ 9 & 10 \\ 11 & 12 \\ 13 & 14 \\ 15 & 16 \end{bmatrix} = (\mathbf{A}^{(\{1\})})^\top$$

Matricization Tucker format

Consider the tensor

$$\mathbf{A} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \otimes \begin{bmatrix} 3 \\ 4 \end{bmatrix} \otimes \begin{bmatrix} 5 \\ 6 \end{bmatrix}$$

Matricization Tucker format

Consider the tensor

$$\mathbf{A} = \begin{bmatrix}1\\2\end{bmatrix} \otimes \begin{bmatrix}3\\4\end{bmatrix} \otimes \begin{bmatrix}5\\6\end{bmatrix} = \left[\begin{bmatrix}15 & 20\\30 & 40\end{bmatrix}, \begin{bmatrix}18 & 24\\36 & 48\end{bmatrix}\right]$$

Then its $(\{1,2\})$-matrizication is

$$\mathbf{A}^{(\{1,2\})}$$

Matricization Tucker format

Consider the tensor

$$\mathbf{A} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \otimes \begin{bmatrix} 3 \\ 4 \end{bmatrix} \otimes \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \left[\begin{bmatrix} 15 & 20 \\ 30 & 40 \end{bmatrix}, \begin{bmatrix} 18 & 24 \\ 36 & 48 \end{bmatrix} \right]$$

Then its $(\{1, 2\})$-matrizication is

$$\mathbf{A}^{(\{1,2\})} = \begin{bmatrix} 15 & 18 \\ 30 & 36 \\ 20 & 24 \\ 40 & 48 \end{bmatrix}$$

Matricization Tucker format

Consider the tensor

$$\mathbf{A} = \begin{bmatrix}1\\2\end{bmatrix} \otimes \begin{bmatrix}3\\4\end{bmatrix} \otimes \begin{bmatrix}5\\6\end{bmatrix} = \left[\begin{bmatrix}15 & 20\\30 & 40\end{bmatrix}, \begin{bmatrix}18 & 24\\36 & 48\end{bmatrix}\right]$$

Then its $(\{1,2\})$-matrizication is

$$\mathbf{A}^{(\{1,2\})} = \begin{bmatrix}15 & 18\\30 & 36\\20 & 24\\40 & 48\end{bmatrix} = \left(\begin{bmatrix}1\\2\end{bmatrix} \otimes \begin{bmatrix}3\\4\end{bmatrix}\right)^{(\{1,2\})} \cdot \left(\left(\begin{bmatrix}5\\6\end{bmatrix}\right)^{(\{1\})}\right)^{\top}$$

Matricization Tucker format

Consider the tensor $\mathbf{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$

$$\mathbf{A} = \sum_{j_1=1}^{k_1} \sum_{j_2=1}^{k_2} \sum_{j_3=1}^{k_3} c_{j_1,j_2,j_3} \mathbf{u}_{j_1,1} \otimes \mathbf{u}_{j_2,2} \otimes \mathbf{u}_{j_3,3}$$

Then

$$\mathbf{A}^{(\{1,2\})} = \sum_{j_1=1}^{k_1} \sum_{j_2=1}^{k_2} (\mathbf{u}_{j_1,1} \otimes \mathbf{u}_{j_2,2})^{(\{1,2\})} \cdot \left(\sum_{j_3=1}^{k_3} c_{j_1,j_2,j_3} \mathbf{u}_{j_3,3} \right)^\top$$

Matricization Tucker format

Consider the tensor $\mathbf{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$

$$\mathbf{A} = \sum_{j_1=1}^{k_1} \sum_{j_2=1}^{k_2} \sum_{j_3=1}^{k_3} c_{j_1,j_2,j_3} \mathbf{u}_{j_1,1} \otimes \mathbf{u}_{j_2,2} \otimes \mathbf{u}_{j_3,3}$$

Then

$$\mathbf{A}^{(\{1,2\})} = \sum_{j_1=1}^{k_1} \sum_{j_2=1}^{k_2} (\mathbf{u}_{j_1,1} \otimes \mathbf{u}_{j_2,2})^{(\{1,2\})} \cdot \left(\sum_{j_3=1}^{k_3} c_{j_1,j_2,j_3} \mathbf{u}_{j_3,3} \right)^{\top}$$

Rank bound

Lemma:
Let $\mathbf{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ with

$$\mathbf{A} = \sum_{j_1=1}^{k_1} \sum_{j_2=1}^{k_2} \sum_{j_3=1}^{k_3} c_{j_1,j_2,j_3} \mathbf{u}_{j_1,1} \otimes \mathbf{u}_{j_2,2} \otimes \mathbf{u}_{j_3,3}.$$

Then

$$\operatorname{rank}(\mathbf{A}^{(\{1,2\})}) \leq \min(\{k_1 \cdot k_2, k_3\}).$$

Root-to-leaves truncated hierarchical SVD[1]

Set $\mathcal{I} = n_1 \times \ldots \times n_d$, and let $T_{\mathcal{I}}$ be a dimension tree of depth $p \in \mathbb{N}$. We call $\mathcal{L}(T_{\mathcal{I}})$ the leafs of the $T_{\mathcal{I}}$, and $\mathcal{I}(T_{\mathcal{I}})$ are the internal nodes of the $T_{\mathcal{I}}$.

Algorithm:

Input: $\mathbf{A} \in \mathbb{R}^{\mathcal{I}}$, $T_{\mathcal{I}}$, target rank $(r_t)_{t \in T_{\mathcal{I}}}$

Output: $(\mathbf{U}_t)_{t \in \mathcal{L}(T_{\mathcal{I}})}$, $(\mathbf{B}_t)_{t \in \mathcal{I}(T_{\mathcal{I}})}$

for $t \in \mathcal{L}(T_{\mathcal{I}})$

$\mathbf{U}_t, \mathbf{\Sigma}_t, \mathbf{V}_t = \mathrm{SVD}(\mathbf{A}^{(t)}, r_t)$

for $\ell = p - 1 : 0$

for $t \in \mathcal{I}(T_{\mathcal{I}})$ on level ℓ

$\mathbf{U}_t, \mathbf{\Sigma}_t, \mathbf{V}_t = \mathrm{SVD}(\mathbf{A}^{(t)}, r_t)$

$\mathbf{U}_{t_1}$ and $\mathbf{U}_{t_2}$ be successors of t.

compute $(\mathbf{B}_t)_{i,j,k} = \langle (\mathbf{U}_{t_1})_i, (\mathbf{U}_{t_2})_j \otimes (\mathbf{U}_{t_1})_k \rangle$

Compute $(\mathbf{B}_{\{1,\ldots,d\}})_{1,j,k} = \langle (\mathbf{A}, (\mathbf{U}_{t_1})_j \otimes (\mathbf{U}_{t_2})_k \rangle$

[1] Grasedyck, SIAM Journal on Matrix Analysis and Applications, 2010

Hierarcical Tucker format

Tree tensor network:

$$(\mathbf{U}_t)_{t\in\mathcal{L}(T_{\mathcal{I}})} \quad \text{and} \quad (\mathbf{B}_t)_{t\in\mathcal{I}(T_{\mathcal{I}})}$$

Storage:

[2] Grasedyck & Hackbusch, Computational Methods in Applied Mathematics, 2011

Hierarcical Tucker format

Tree tensor network:

$$(\mathbf{U}_t)_{t\in\mathcal{L}(T_\mathcal{I})} \quad \text{and} \quad (\mathbf{B}_t)_{t\in\mathcal{I}(T_\mathcal{I})}$$

Storage: $\mathcal{O}(rnd + dr^3)$

Compare to TT: $\mathcal{O}(dnr^2)$

Why HT?

[2]Grasedyck & Hackbusch, Computational Methods in Applied Mathematics, 2011

Hierarcical Tucker format

Tree tensor network:

$$(\mathbf{U}_t)_{t\in\mathcal{L}(T_\mathcal{I})} \quad \text{and} \quad (\mathbf{B}_t)_{t\in\mathcal{I}(T_\mathcal{I})}$$

Storage: $\mathcal{O}(rnd + dr^3)$

Compare to TT: $\mathcal{O}(dnr^2)$

Why HT?

$\Rightarrow$ Remember, the tensor rank is not uniquely defined!

A tensor $\mathbf{A}$ may be *exactly* represented in HT and TT format, however, the ranks appearing in HT can be lower!
(Consequence of general rank bounds[2])

[2] Grasedyck & Hackbusch, Computational Methods in Applied Mathematics, 2011