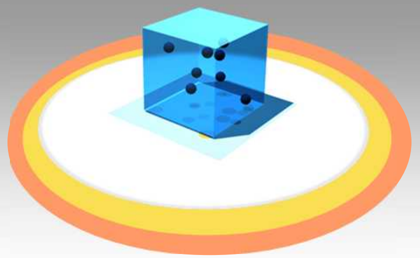
SAND2007-4534C

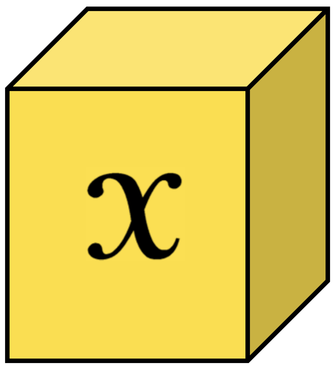
The MATLAB tensor toolbox for efficient computations with sparse and factored tensors

Brett W. Bader & Tamara G. Kolda
Sandia National Laboratories



Tensor Basics: Fibers and Unfolding

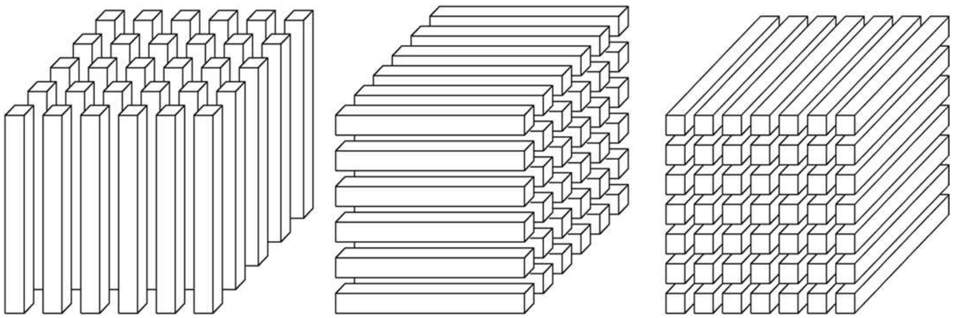
I x J x K



Mode-1 Fibers

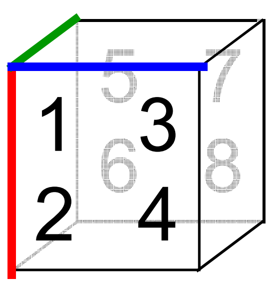
Mode-2 Fibers

Mode-3 Fibers



$\mathbf{X}_{(n)}$: The mode-**n** fibers are rearranged to be the columns of a matrix

$\mathcal{X} =$



$$\mathbf{X}_{(1)} = \begin{bmatrix} 1 & 3 & 5 & 7 \\ 2 & 4 & 6 & 8 \end{bmatrix}$$

$$\mathbf{X}_{(2)} = \begin{bmatrix} 1 & 2 & 5 & 6 \\ 3 & 4 & 7 & 8 \end{bmatrix}$$

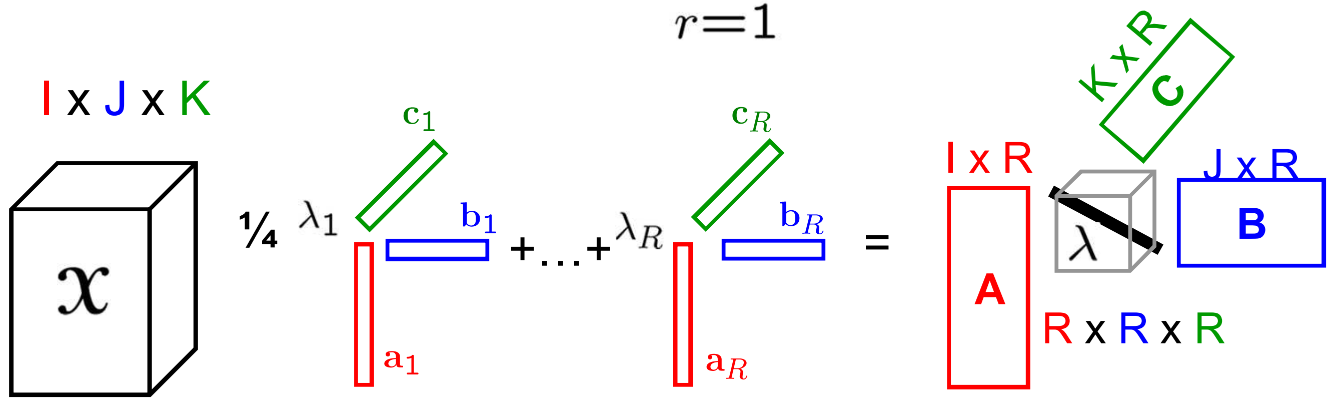
$$\mathbf{X}_{(3)} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}$$

$$\mathrm{vec}(\mathcal{X}) = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{bmatrix}$$

Sandia National Laboratories

CANDECOMP/PARAFAC

$$\mathcal{X} \approx [\![\lambda\,;\mathbf{A},\mathbf{B},\mathbf{C}]\!] = \sum_{r=1}^{R} \lambda_r \, \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r$$



$$\left.\begin{aligned} \mathbf{X}_{(1)} &= \mathbf{A}\Lambda(\mathbf{C} \odot \mathbf{B}) \\ \mathbf{X}_{(2)} &= \mathbf{B}\Lambda(\mathbf{C} \odot \mathbf{A}) \\ \mathbf{X}_{(3)} &= \mathbf{C}\Lambda(\mathbf{B} \odot \mathbf{A}) \end{aligned}\right\}$$

Matricized versions of PARAFAC

Khatri-Rao Product

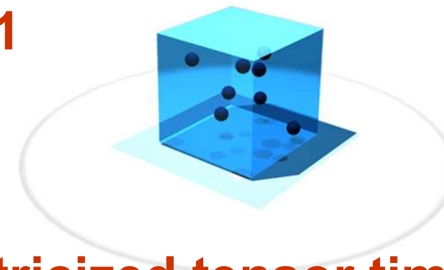
$$\underset{M \times R}{\mathbf{A}} \odot \underset{N \times R}{\mathbf{B}} = \underbrace{\begin{bmatrix} \mathbf{a}_1 \otimes \mathbf{b}_1 & \mathbf{a}_2 \otimes \mathbf{b}_2 & \cdots & \mathbf{a}_R \otimes \mathbf{b}_R \end{bmatrix}}_{MN \times R}$$

$$(\mathbf{A} \odot \mathbf{B})^{\mathsf{T}}(\mathbf{A} \odot \mathbf{B}) = (\mathbf{A}^{\mathsf{T}}\mathbf{A} * \mathbf{B}^{\mathsf{T}}\mathbf{B})$$

$$(\mathbf{A} \odot \mathbf{B})^{\dagger} \equiv (\mathbf{A}^{\mathsf{T}}\mathbf{A} * \mathbf{B}^{\mathsf{T}}\mathbf{B})^{\dagger}(\mathbf{A} \odot \mathbf{B})^{\mathsf{T}}$$

PARAFAC-ALS for Sparse Tensors

Specify R. $\boxed{\text{Initialize } \mathbf{A}, \mathbf{B}, \mathbf{C}.}$ **Tucker1**



Repeat

$$\mathbf{A} = \boxed{\mathbf{X}_{(1)}(\mathbf{C} \odot \mathbf{B})}(\mathbf{C}^{\mathsf{T}}\mathbf{C} * \mathbf{B}^{\mathsf{T}}\mathbf{B})^{\dagger}$$
$$\mathbf{B} = \boxed{\mathbf{X}_{(2)}(\mathbf{C} \odot \mathbf{A})}(\mathbf{C}^{\mathsf{T}}\mathbf{C} * \mathbf{A}^{\mathsf{T}}\mathbf{A})^{\dagger}$$
$$\mathbf{C} = \boxed{\mathbf{X}_{(3)}(\mathbf{B} \odot \mathbf{A})}(\mathbf{B}^{\mathsf{T}}\mathbf{B} * \mathbf{A}^{\mathsf{T}}\mathbf{A})^{\dagger}$$

Matricized tensor times Khatri-Rao product (mttkrp)

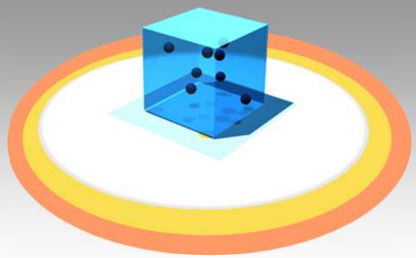
Normalize columns of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C} \Rightarrow \lambda$

$$\text{fit} = \|\, \mathcal{X} - [\![\lambda\,; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \,\|^2$$
$$= \boxed{\|\, \mathcal{X} \,\|^2} - 2\boxed{\langle \mathcal{X}, [\![\lambda\,; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \rangle} + \boxed{\|\, [\![\lambda\,; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \,\|^2}$$

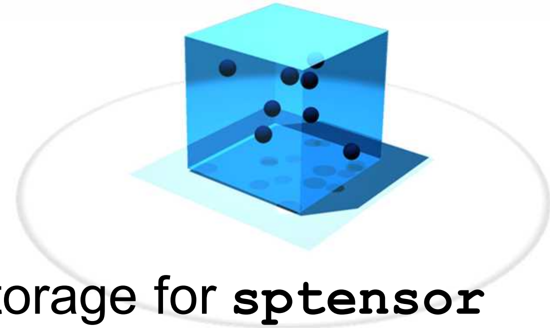
Norm **Inner Product** **Norm**

Until $\Delta_{\text{fit}} <$ tol

Sandia National Laboratories



Exploiting Sparsity: Sparse Tensors (sptensor)



- *Sparse* if majority of entries (x_{ijk}) are zero
- Some storage options
  - Each two-dimensional slice stored as sparse matrix
  - Unfold and store as sparse matrix
    - Lin, Liu, Chung, IEEE Trans. Computers, 2002 & 2003
  - Coordinate format

- Storage for **`sptensor`**
  - P = # nonzeros
  - **`vals`** = P x 1 vector of nonzero values
  - **`subs`** = P x 3 matrix of subscripts

Norm of sparse tensor

$$\| \mathcal{X} \|^2 = \sum_{p=1}^{P} [\texttt{vals}(p)]^2$$

Sandia National Laboratories

Roadmap

Specify R. Initialize $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$. **Tucker1**

Repeat

$$\mathbf{A} = \mathbf{X}_{(1)}(\mathbf{C} \odot \mathbf{B})(\mathbf{C}^{\mathsf{T}}\mathbf{C} * \mathbf{B}^{\mathsf{T}}\mathbf{B})^{\dagger}$$

$$\mathbf{B} = \mathbf{X}_{(2)}(\mathbf{C} \odot \mathbf{A})(\mathbf{C}^{\mathsf{T}}\mathbf{C} * \mathbf{A}^{\mathsf{T}}\mathbf{A})^{\dagger}$$

$$\mathbf{C} = \mathbf{X}_{(3)}(\mathbf{B} \odot \mathbf{A})(\mathbf{B}^{\mathsf{T}}\mathbf{B} * \mathbf{A}^{\mathsf{T}}\mathbf{A})^{\dagger}$$

Matricized tensor times Khatri-Rao product (mttkrp)

Normalize columns of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ $\Rightarrow \lambda$

$$\text{fit} = \| \mathcal{X} - [\![\lambda \, ; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \|^2$$

$$= \| \mathcal{X} \|^2 - 2 \langle \mathcal{X}, [\![\lambda \, ; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \rangle + \| [\![\lambda \, ; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \|^2$$

Norm **Inner Product** **Norm**

Until $\Delta_{\text{fit}} <$ tol

Tucker1 (aka HO-SVD)

- Need to compute leading left singular vectors of $X_{(n)}$ where corresponding tensor is sparse
  - Equivalent to eigenvectors of $X_{(n)}^T X_{(n)}$
- $X_{(n)}$ is a narrow, wide matrix
  - Worst possible aspect ratio for MATLAB's CSC format
- Instead, let $U = X_{(n)}^T$, which is long and skinny
- Compute eigenvectors of $V = UU^T$

Roadmap

Specify R. $\boxed{\text{Initialize } \mathbf{A}, \mathbf{B}, \mathbf{C}.}$ Tucker1

Repeat

$$\mathbf{A} = \boxed{\mathbf{X}_{(1)}(\mathbf{C} \odot \mathbf{B})}(\mathbf{C}^{\mathsf{T}}\mathbf{C} * \mathbf{B}^{\mathsf{T}}\mathbf{B})^{\dagger}$$

$$\mathbf{B} = \boxed{\mathbf{X}_{(2)}(\mathbf{C} \odot \mathbf{A})}(\mathbf{C}^{\mathsf{T}}\mathbf{C} * \mathbf{A}^{\mathsf{T}}\mathbf{A})^{\dagger}$$

$$\mathbf{C} = \boxed{\mathbf{X}_{(3)}(\mathbf{B} \odot \mathbf{A})}(\mathbf{B}^{\mathsf{T}}\mathbf{B} * \mathbf{A}^{\mathsf{T}}\mathbf{A})^{\dagger}$$

Matricized tensor times Khatri-Rao product (mttkrp)

Normalize columns of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C} \Rightarrow \lambda$

$$\text{fit} = \| \mathcal{X} - [\![\lambda \, ; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \|^2$$

$$= \boxed{\| \mathcal{X} \|^2} - 2 \boxed{\langle \mathcal{X}, [\![\lambda \, ; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \rangle} + \boxed{\| [\![\lambda \, ; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \|^2}$$

Inner Product

Norm

Until $\Delta_{\text{fit}} < \text{tol}$

Matricized Tensor Times Khatri-Rao Product (`mttkrp`)

Specific example for 3-way tensor
(but idea is easily generalized):

$$\mathbf{Z} = \mathbf{X}_{(1)}(\mathbf{C} \odot \mathbf{B})$$

Compute Z column-wise (for r=1,…,R):

$$\mathbf{z}_r = \mathbf{X}_{(1)}(\mathbf{c}_r \otimes \mathbf{b}_r)$$

Rewrite the above equation element-wise:

$$z_r(i) = \sum_{j=1}^{J} \sum_{k=1}^{K} x(i,j,k) \cdot b_r(j) \cdot c_r(k)$$

Storage for **`sptensor`**

- P = # nonzeros
- **`vals`** = P x 1 vector of nonzero values
- **`subs`** = P x 3 matrix of subscripts

Loop through nonzeros only:

expand to array of length P
(with some repeated elements)

$$z_r(i) = \sum_{\substack{p=1 \\ \texttt{subs}(p,1)=i}}^{P} \texttt{vals}(p) \cdot b_r(\texttt{subs}(p,2)) \cdot c_r(\texttt{subs}(p,3))$$

Laboratories

Matlab Translation

$$z(i) = \sum_{j=1}^{J} \sum_{k=1}^{K} x(i,j,k) \cdot b(j) \cdot c(k)$$

$$= \sum_{\substack{p=1 \\ \texttt{subs}(p,1)=i}}^{P} \texttt{vals}(p) \cdot b(\texttt{subs}(p,2)) \cdot c(\texttt{subs}(p,3))$$

2x2x2 Tensor with 4 nonzeros	Vectors
$\texttt{subs} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 2 & 1 & 2 \\ 2 & 2 & 2 \end{bmatrix}$ $\texttt{vals} = \begin{bmatrix} 1.5 \\ 2.7 \\ 3.3 \\ 8.5 \end{bmatrix}$	$b = \begin{bmatrix} .1 \\ .2 \end{bmatrix}$ $c = \begin{bmatrix} 30 \\ 40 \end{bmatrix}$

```
z = accumarray(vals.*b(subs(:,2)).*c(subs(:.3)),subs(:,1))
```

$z(1)$	$\begin{bmatrix} 1.5 \\ 2.7 \end{bmatrix}$	$\begin{bmatrix} .1 \\ .2 \end{bmatrix}$	$\begin{bmatrix} 30 \\ 30 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$
$z(2)$	$\begin{bmatrix} 3.3 \\ 8.5 \end{bmatrix}$	$\begin{bmatrix} .1 \\ .2 \end{bmatrix}$	$\begin{bmatrix} 30 \\ 40 \end{bmatrix}$	$\begin{bmatrix} 2 \\ 2 \end{bmatrix}$

Sandia National Laboratories

Roadmap

Specify R. $\boxed{\text{Initialize } \mathbf{A}, \mathbf{B}, \mathbf{C}.}$ Tucker1

Repeat

$$\mathbf{A} = \boxed{\mathbf{X}_{(1)}(\mathbf{C} \odot \mathbf{B})}(\mathbf{C}^{\mathsf{T}}\mathbf{C} * \mathbf{B}^{\mathsf{T}}\mathbf{B})^{\dagger}$$

$$\mathbf{B} = \boxed{\mathbf{X}_{(2)}(\mathbf{C} \odot \mathbf{A})}(\mathbf{C}^{\mathsf{T}}\mathbf{C} * \mathbf{A}^{\mathsf{T}}\mathbf{A})^{\dagger}$$

$$\mathbf{C} = \boxed{\mathbf{X}_{(3)}(\mathbf{B} \odot \mathbf{A})}(\mathbf{B}^{\mathsf{T}}\mathbf{B} * \mathbf{A}^{\mathsf{T}}\mathbf{A})^{\dagger}$$

Matricized tensor times Khatri-Rao product (mttkrp)

Normalize columns of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ $\Rightarrow \lambda$

$$\text{fit} = \| \mathcal{X} - [\![\lambda \,; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \|^2$$

$$= \boxed{\| \mathcal{X} \|^2} - 2 \boxed{\langle \mathcal{X}, [\![\lambda \,; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \rangle} + \boxed{\| [\![\lambda \,; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \|^2}$$

Until $\Delta_{\text{fit}} <$ tol

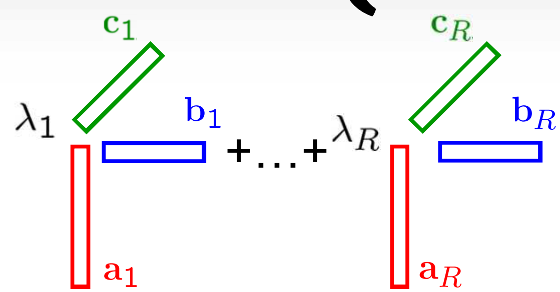
Norm

Inner Product

Norm

Exploiting Structure: Kruskal Tensors (`ktensor`)

$$[\![\lambda \,;\, \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \quad = $$



- Memory constraints prohibit calculating the “full” tensor
  - IJK memory
- Instead, store
  - λ – R x 1 vector
  - A – I x R matrix
  - B – J x R matrix
  - C – K x R matrix
- Memory: R(I+J+K+1)

Calculate the norm: O(R^2(I+J+K)) work

$$\| [\![\lambda \,;\, \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \|^2$$

vectorized version…

$$= \| \underbrace{(\mathbf{C} \odot \mathbf{B} \odot \mathbf{A})}_{\text{IJK x R}} \lambda \|^2$$

$$= \lambda^{\mathsf{T}} (\mathbf{C} \odot \mathbf{B} \odot \mathbf{A})^{\mathsf{T}} (\mathbf{C} \odot \mathbf{B} \odot \mathbf{A}) \lambda$$

property of the Khatri-Rao product…

$$= \lambda^{\mathsf{T}} (\underbrace{\mathbf{C}^{\mathsf{T}}\mathbf{C}}_{\text{R x R}} * \underbrace{\mathbf{B}^{\mathsf{T}}\mathbf{B}}_{\text{R x R}} * \underbrace{\mathbf{A}^{\mathsf{T}}\mathbf{A}}_{\text{R x R}}) \lambda$$

Roadmap

Specify R. $\boxed{\text{Initialize } \mathbf{A}, \mathbf{B}, \mathbf{C}.}$ Tucker1

Repeat

$$\mathbf{A} = \boxed{\mathbf{X}_{(1)}(\mathbf{C} \odot \mathbf{B})}(\mathbf{C}^{\mathsf{T}}\mathbf{C} * \mathbf{B}^{\mathsf{T}}\mathbf{B})^{\dagger}$$

$$\mathbf{B} = \boxed{\mathbf{X}_{(2)}(\mathbf{C} \odot \mathbf{A})}(\mathbf{C}^{\mathsf{T}}\mathbf{C} * \mathbf{A}^{\mathsf{T}}\mathbf{A})^{\dagger}$$

$$\mathbf{C} = \boxed{\mathbf{X}_{(3)}(\mathbf{B} \odot \mathbf{A})}(\mathbf{B}^{\mathsf{T}}\mathbf{B} * \mathbf{A}^{\mathsf{T}}\mathbf{A})^{\dagger}$$

Matricized tensor times Khatri-Rao product (mttkrp)

Normalize columns of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ $\Rightarrow \lambda$

$$\text{fit} = \| \mathcal{X} - [\![\lambda \,;\, \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \|^2$$

$$= \underbrace{\boxed{\| \mathcal{X} \|^2}}_{\text{Norm}} - 2 \underbrace{\boxed{\langle \mathcal{X}, [\![\lambda \,;\, \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \rangle}}_{\text{Inner Product}} + \underbrace{\boxed{\| [\![\lambda \,;\, \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \|^2}}_{\text{Norm}}$$

Until $\Delta_{\text{fit}} <$ tol

Sandia National Laboratories

Inner Product: `sptensor` & `ktensor`

$$\langle\, \mathcal{X}, [\![\lambda\,;\mathbf{A},\mathbf{B},\mathbf{C}]\!]\,\rangle$$

$$= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} x_{ijk}\left(\sum_{r=1}^{R}\lambda_r a_{ir} b_{jr} c_{kr}\right)$$

write out product

$$= \sum_{r=1}^{R}\lambda_r \underbrace{\left(\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} x_{ijk} a_{ir} b_{jr} c_{kr}\right)}_{\boxed{\sum_{p=1}^{P} \mathtt{vals}(p)\cdot a_r(\mathtt{subs}(p,1))\cdot b_r(\mathtt{subs}(p,2))\cdot c_r(\mathtt{subs}(p,3))}}$$

rearrange terms

Roadmap

Specify R. $\boxed{\text{Initialize } \mathbf{A}, \mathbf{B}, \mathbf{C}.}$ **Tucker1**

Repeat

$$\mathbf{A} = \boxed{\mathbf{X}_{(1)}(\mathbf{C} \odot \mathbf{B})}(\mathbf{C}^{\mathsf{T}}\mathbf{C} * \mathbf{B}^{\mathsf{T}}\mathbf{B})^{\dagger}$$

$$\mathbf{B} = \boxed{\mathbf{X}_{(2)}(\mathbf{C} \odot \mathbf{A})}(\mathbf{C}^{\mathsf{T}}\mathbf{C} * \mathbf{A}^{\mathsf{T}}\mathbf{A})^{\dagger}$$

$$\mathbf{C} = \boxed{\mathbf{X}_{(3)}(\mathbf{B} \odot \mathbf{A})}(\mathbf{B}^{\mathsf{T}}\mathbf{B} * \mathbf{A}^{\mathsf{T}}\mathbf{A})^{\dagger}$$

Matricized tensor times Khatri-Rao product (mttkrp)

Normalize columns of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ $\Rightarrow \lambda$

$$\text{fit} = \| \mathcal{X} - [\![\lambda \,; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \|^2$$

$$= \boxed{\| \mathcal{X} \|^2} - 2 \boxed{\langle \mathcal{X}, [\![\lambda \,; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \rangle} + \boxed{\| [\![\lambda \,; \mathbf{A}, \mathbf{B}, \mathbf{C}]\!] \|^2}$$

Norm **Inner Product** **Norm**

Until $\Delta_{\text{fit}} <$ tol

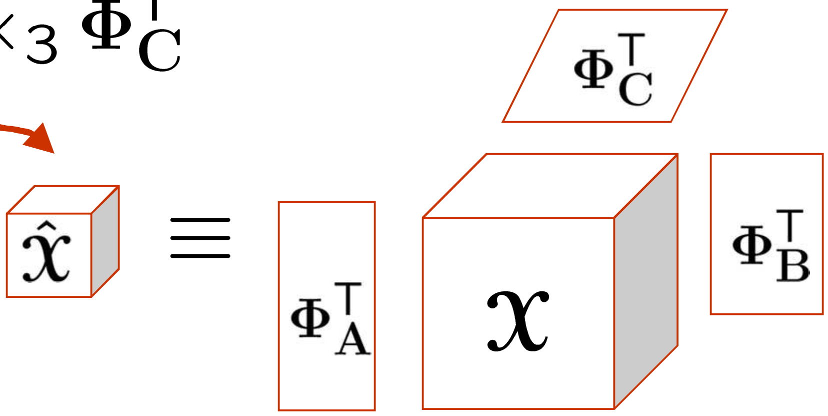
Compressed PARAFAC (CANDELINC)

Choose orthonormal matrices Φ_A, Φ_B, Φ_C for compression

Mode-n Multiplication

$$\mathcal{Y} = \mathcal{X} \times_n \mathbf{U} \Leftrightarrow \mathbf{Y}_{(n)} = \mathbf{U}\mathbf{X}_{(n)}$$

$$\hat{\mathcal{X}} \equiv \mathcal{X} \times_1 \Phi_A^T \times_2 \Phi_B^T \times_3 \Phi_C^T$$



Calculate PARAFAC on compressed matrix

$$\hat{\mathcal{X}} \approx [\![\hat{A}, \hat{B}, \hat{C}]\!]$$

Reassemble into PARAFAC for full matrix

$$\mathcal{X} \approx [\![\underbrace{\Phi_A \hat{A}}_{A}, \underbrace{\Phi_B \hat{B}}_{B}, \underbrace{\Phi_C \hat{C}}_{C}]\!]$$

compression matrices = mode-n principal components, i.e., eigenvectors of $X_{(n)}^T X_{(n)}$

Sandia National Laboratories

Roadmap

Specify R. $\boxed{\text{Initialize } \hat{A},\ \hat{B},\ \hat{C}.}$ Random

Repeat

$$\hat{A} = \boxed{\hat{X}_{(1)}(\hat{C} \odot \hat{B})}(\hat{C}^{\mathsf{T}}\hat{C} * \hat{B}^{\mathsf{T}}\hat{B})^{\dagger}$$

$$\hat{B} = \boxed{\hat{X}_{(2)}(\hat{C} \odot \hat{A})}(\hat{C}^{\mathsf{T}}\hat{C} * \hat{A}^{\mathsf{T}}\hat{A})^{\dagger}$$

$$\hat{C} = \boxed{\hat{X}_{(3)}(\hat{B} \odot \hat{A})}(\hat{B}^{\mathsf{T}}\hat{B} * \hat{A}^{\mathsf{T}}\hat{A})^{\dagger}$$

Matricized tensor times Khatri-Rao product (mttkrp)

Normalize columns of $\hat{A},\ \hat{B},\ \hat{C} \Rightarrow \hat{\lambda}$

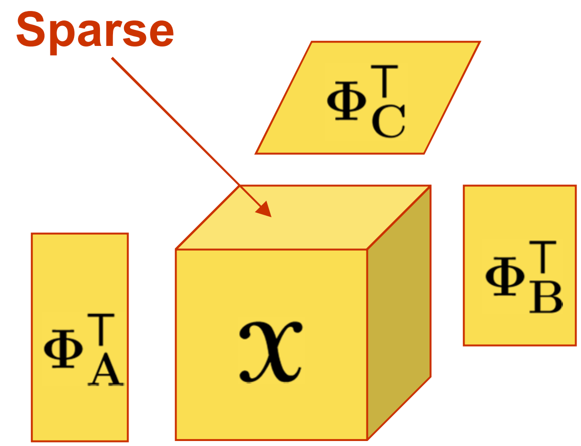
$$\text{fit} = \left\| \hat{\mathfrak{X}} - [\![\hat{\lambda}\,;\hat{A},\hat{B},\hat{C}]\!] \right\|^2$$

$$= \underbrace{\boxed{\left\| \hat{\mathfrak{X}} \right\|^2}}_{\text{Norm}} - 2\underbrace{\boxed{\langle \hat{\mathfrak{X}}, [\![\hat{\lambda}\,;\hat{A},\hat{B},\hat{C}]\!] \rangle}}_{\text{Inner Product}} + \underbrace{\boxed{\left\| [\![\hat{\lambda}\,;\hat{A},\hat{B},\hat{C}]\!] \right\|^2}}_{\text{Norm}}$$

Until $\Delta_{\text{fit}} <$ tol

Tucker Tensor (`ttensor`)

- Store "core" tensor X
  - Can be sparse
- Store multiplier matrices
  - Φ_A, Φ_B, Φ_C
- If core is sparse, cannot compute norm
  - Dense intermediate products



$$\hat{x}_{mnp} = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} x_{ijk} \, (\Phi_A)_{mi} \, (\Phi_B)_{nj} \, (\Phi_C)_{pk}$$

Sandia National Laboratories

Roadmap

Specify R. $\boxed{\text{Initialize } \hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}.}$ **Random**

Repeat

$$\hat{\mathbf{A}} = \boxed{\hat{\mathbf{X}}_{(1)}(\hat{\mathbf{C}} \odot \hat{\mathbf{B}})}(\hat{\mathbf{C}}^{\mathsf{T}}\hat{\mathbf{C}} * \hat{\mathbf{B}}^{\mathsf{T}}\hat{\mathbf{B}})^{\dagger}$$

$$\hat{\mathbf{B}} = \boxed{\hat{\mathbf{X}}_{(2)}(\hat{\mathbf{C}} \odot \hat{\mathbf{A}})}(\hat{\mathbf{C}}^{\mathsf{T}}\hat{\mathbf{C}} * \hat{\mathbf{A}}^{\mathsf{T}}\hat{\mathbf{A}})^{\dagger}$$

$$\hat{\mathbf{C}} = \boxed{\hat{\mathbf{X}}_{(3)}(\hat{\mathbf{B}} \odot \hat{\mathbf{A}})}(\hat{\mathbf{B}}^{\mathsf{T}}\hat{\mathbf{B}} * \hat{\mathbf{A}}^{\mathsf{T}}\hat{\mathbf{A}})^{\dagger}$$

Matricized tensor times Khatri-Rao product (mttkrp)

Normalize columns of $\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}} \Rightarrow \hat{\lambda}$

$$\text{fit} = \left\| \hat{\mathfrak{X}} - [\![\hat{\lambda} \,; \hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}]\!] \right\|^2$$

$$= \boxed{\cancel{\| \hat{\mathfrak{X}} \|^2}} - 2 \boxed{\langle \hat{\mathfrak{X}}, [\![\hat{\lambda} \,; \hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}]\!] \rangle} + \boxed{\left\| [\![\hat{\lambda} \,; \hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}]\!] \right\|^2}$$

Norm **Inner Product** **Norm**

Until $\Delta_{\text{fit}} <$ tol

Matricized Tensor Times Khatri-Rao Product (`mttkrp`)

Specific example for 3-way tensor (but idea is easily generalized):

$$
\begin{aligned}
\mathbf{Z} &= \hat{\mathbf{X}}_{(1)}(\hat{\mathbf{C}} \odot \hat{\mathbf{B}}) \\
&= [\mathbf{\Phi}_{\mathbf{A}}^{\mathsf{T}} \mathbf{X}_{(1)} (\mathbf{\Phi}_{\mathbf{C}} \otimes \mathbf{\Phi}_{\mathbf{B}})](\hat{\mathbf{C}} \odot \hat{\mathbf{B}}) \\
&= \mathbf{\Phi}_{\mathbf{A}}^{\mathsf{T}} [\mathbf{X}_{(1)} (\mathbf{\Phi}_{\mathbf{C}} \hat{\mathbf{C}} \odot \mathbf{\Phi}_{\mathbf{B}} \hat{\mathbf{B}})] \\
&= \mathbf{\Phi}_{\mathbf{A}}^{\mathsf{T}} [\underbrace{\mathbf{X}_{(1)} (\mathbf{C} \odot \mathbf{B})}_{\text{core ``sparse'' mttkrp}}]
\end{aligned}
$$

Sandia National Laboratories

Roadmap

Specify R. $\boxed{\text{Initialize } \hat{\mathbf{A}},\ \hat{\mathbf{B}},\ \hat{\mathbf{C}}.}$ **Random**

Repeat

$$\hat{\mathbf{A}} = \boxed{\hat{\mathbf{X}}_{(1)}(\hat{\mathbf{C}} \odot \hat{\mathbf{B}})}(\hat{\mathbf{C}}^{\mathsf{T}}\hat{\mathbf{C}} * \hat{\mathbf{B}}^{\mathsf{T}}\hat{\mathbf{B}})^{\dagger}$$

$$\hat{\mathbf{B}} = \boxed{\hat{\mathbf{X}}_{(2)}(\hat{\mathbf{C}} \odot \hat{\mathbf{A}})}(\hat{\mathbf{C}}^{\mathsf{T}}\hat{\mathbf{C}} * \hat{\mathbf{A}}^{\mathsf{T}}\hat{\mathbf{A}})^{\dagger}$$

$$\hat{\mathbf{C}} = \boxed{\hat{\mathbf{X}}_{(3)}(\hat{\mathbf{B}} \odot \hat{\mathbf{A}})}(\hat{\mathbf{B}}^{\mathsf{T}}\hat{\mathbf{B}} * \hat{\mathbf{A}}^{\mathsf{T}}\hat{\mathbf{A}})^{\dagger}$$

Matricized tensor times Khatri-Rao product (mttkrp)

Normalize columns of $\hat{\mathbf{A}},\ \hat{\mathbf{B}},\ \hat{\mathbf{C}} \Rightarrow \hat{\lambda}$

$$\text{fit} = \left\| \hat{\mathfrak{X}} - [\![\hat{\lambda}\,; \hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}]\!] \right\|^2$$

$$= \cancel{\boxed{\|\hat{\mathfrak{X}}\|^2}} - 2\boxed{\langle \hat{\mathfrak{X}}, [\![\hat{\lambda}\,; \hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}]\!] \rangle} + \boxed{\left\| [\![\hat{\lambda}\,; \hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}]\!] \right\|^2}$$

Norm **Inner Product** **Norm**

Until $\Delta_{\text{fit}} \leq$ tol

Inner Product: `ttensor` & `ktensor`

$$
\begin{aligned}
&\langle \hat{\mathcal{X}}, [\![\lambda \,;\, \hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}]\!] \rangle \\
&= \langle \mathcal{X} \times_1 \mathbf{\Phi}_{\mathbf{A}}^{\mathsf{T}} \times_2 \mathbf{\Phi}_{\mathbf{B}}^{\mathsf{T}} \times_3 \mathbf{\Phi}_{\mathbf{C}}^{\mathsf{T}}, [\![\lambda \,;\, \hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}]\!] \rangle \\
&= \langle \mathcal{X}, [\![\lambda \,;\, \mathbf{\Phi}_{\mathbf{A}}\hat{\mathbf{A}}, \mathbf{\Phi}_{\mathbf{B}}\hat{\mathbf{B}}, \mathbf{\Phi}_{\mathbf{C}}\hat{\mathbf{C}}]\!] \rangle \\
&= \langle \underbrace{\mathcal{X}, [\![\lambda \,;\, \mathbf{A}, \mathbf{B}, \mathbf{C}]\!]}_{\text{Inner product of sparse \& ktensor}} \rangle
\end{aligned}
$$

Roadmap

Specify R. $\boxed{\text{Initialize } \hat{\mathrm{A}},\ \hat{\mathrm{B}},\ \hat{\mathrm{C}}.}$ **Random**

Repeat

$$\hat{\mathrm{A}} = \boxed{\hat{\mathrm{X}}_{(1)}(\hat{\mathrm{C}} \odot \hat{\mathrm{B}})}(\hat{\mathrm{C}}^{\mathsf{T}}\hat{\mathrm{C}} * \hat{\mathrm{B}}^{\mathsf{T}}\hat{\mathrm{B}})^{\dagger}$$

$$\hat{\mathrm{B}} = \boxed{\hat{\mathrm{X}}_{(2)}(\hat{\mathrm{C}} \odot \hat{\mathrm{A}})}(\hat{\mathrm{C}}^{\mathsf{T}}\hat{\mathrm{C}} * \hat{\mathrm{A}}^{\mathsf{T}}\hat{\mathrm{A}})^{\dagger}$$

$$\hat{\mathrm{C}} = \boxed{\hat{\mathrm{X}}_{(3)}(\hat{\mathrm{B}} \odot \hat{\mathrm{A}})}(\hat{\mathrm{B}}^{\mathsf{T}}\hat{\mathrm{B}} * \hat{\mathrm{A}}^{\mathsf{T}}\hat{\mathrm{A}})^{\dagger}$$

Matricized tensor times Khatri-Rao product (mttkrp)

Normalize columns of $\hat{\mathrm{A}},\ \hat{\mathrm{B}},\ \hat{\mathrm{C}} \Rightarrow \hat{\lambda}$

$$\text{fit} = \left\| \hat{\mathfrak{X}} - [\![\hat{\lambda}\,;\hat{\mathrm{A}},\hat{\mathrm{B}},\hat{\mathrm{C}}]\!] \right\|^2$$

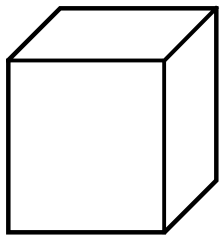
$$= \boxed{\cancel{\|\hat{\mathfrak{X}}\|^2}} - 2\boxed{\langle \hat{\mathfrak{X}}, [\![\hat{\lambda}\,;\hat{\mathrm{A}},\hat{\mathrm{B}},\hat{\mathrm{C}}]\!] \rangle} + \boxed{\left\| [\![\hat{\lambda}\,;\hat{\mathrm{A}},\hat{\mathrm{B}},\hat{\mathrm{C}}]\!] \right\|^2}$$

Norm **Inner Product** **Norm**

Until $\Delta_{\text{fit}} < \text{tol}$

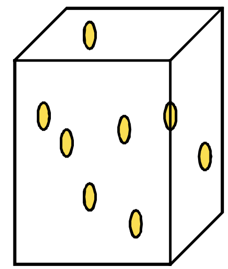
Sparse & Structured Tensors supported in Tensor Toolbox V2.0

Dense Tensors `tensor`



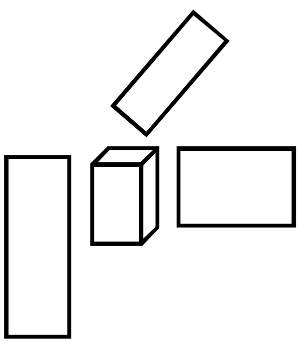
- Extends MATLAB's native MDA capabilities
- Can be converted to a matrix and vice versa

Sparse Tensors `sptensor`



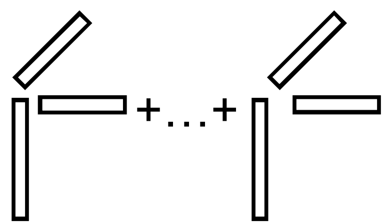
- Unique to Tensor Toolbox
- Can be converted to a (sparse) matrix and vice versa
- Effort to choose suitable representation
- Efficient functions for computation

Tucker Tensors `ttensor`



- Stores a tensor in decomposed form
- A different way to store a large-scale dense tensor
- Can do many operations in factored form

Kruskal Tensors `ktensor`



- Stores a tensor as sum of rank-1 tensors
- A different way to store a large-scale dense tensor
- Can do many operations in factored form

Sandia National Laboratories

Other Roles for Structure

- Missing value imputations
  - With Morten Morup
  - Data imputation

$$\mathcal{P} = [\![\mathbf{A}, \mathbf{B}, \mathbf{C}]\!]$$

Current Estimate

Binary weight matrix

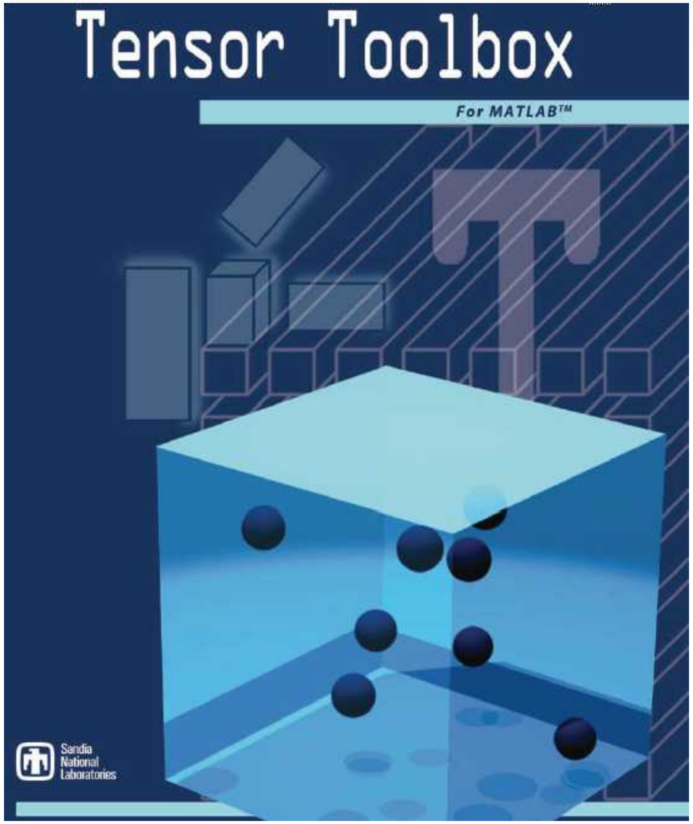
$$\hat{\mathcal{X}} = \mathcal{W} * \mathcal{X} + (1 - \mathcal{W}) * \mathcal{P}$$
$$= \underbrace{\mathcal{X} - \mathcal{W} * \mathcal{P}}_{\text{sparse}} + \underbrace{\mathcal{P}}_{\text{ktensor}}$$

- Structured slices in a 3-way array
  - With Teresa Selee
  - N x N x K tensors
  - Each slice is a similarity matrix
  - $X_k = F_k F_k^T$ where F_k is an N x P_k feature matrix
  - I-PARAFAC = Implicit PARAFAC

Sandia National Laboratories

Conclusions & Future Work

- Discussed exploiting sparsity & structure
  - sptensor – sparse tensors
  - Ktensor – sum of rank-1 tensors
  - Ttensor – core times matrices
- Motivation: Data mining
- Many other applications
  - 700+ Tensor Toolbox downloads since Sept 2007



More information:
Tammy Kolda
tgkolda@sandia.gov
http://csmr.ca.sandia.gov/~tgkolda/

Sandia National Laboratories