Multi-linear Algebra
– Tensor Ranks & CP decomposition–
Lecture 15

F. M. Faulstich

19/03/2024

Matrix rank

Recall:

A matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ is of rank r if and only if:

- There are exactly r linearly independent columns in $\mathbf{A}$
- There are exactly r linearly independent row in $\mathbf{A}$
- The image of the linear map induced by $\mathbf{A}$ is of dimension r
- r is the smallest number such that exist $\mathbf{u}_i \in \mathbb{R}^m$ and $\mathbf{v}_i \in \mathbb{R}^n$ and real numbers $\sigma_i > 0$ such that

$$\mathbf{A} = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^\top = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \otimes \mathbf{v}_i$$

- r is the smallest number, such that there exist r-dimensional subspaces $V \subseteq \mathbb{R}^m$ and $U \subseteq \mathbb{R}^n$, such that A is an element of the induced tensor space $V \otimes U \subseteq \mathbb{R}^{m \times n}$

Canonical Polyadic (CP) Decomposition

Definition:
Let $\mathbf{X} \in \mathbb{R}^{n_1 \times \ldots \times n_d}$ be a tensor of order d. A representation of $\mathbf{X}$ as a sum of elementary tensors

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

for $\mathbf{v}_{i,p} \in \mathbb{R}^{n_i}$ is called a canonical polyadic (CP) representation of $\mathbf{X}$. The number of terms r is called the "rank of the representation". The minimal r, such that there exists a CP decomposition of X with rank r, is called the canonical rank or CP-rank of $\mathbf{X}$.

Example

$$\left[\begin{bmatrix}1.5 & -2.5\\ 2.5 & -2.5\end{bmatrix}, \begin{bmatrix}-2 & 2\\ -2 & 2\end{bmatrix}\right]$$

[1] J. Håstad, Journal of Algorithms, 1990

Example

$$\left[\begin{bmatrix}1.5 & -2.5\\ 2.5 & -2.5\end{bmatrix}, \begin{bmatrix}-2 & 2\\ -2 & 2\end{bmatrix}\right] = \left[\begin{bmatrix}2 & -2\\ 2 & -2\end{bmatrix}, \begin{bmatrix}-2 & 2\\ -2 & 2\end{bmatrix}\right] + \left[\begin{bmatrix}-.5 & -.5\\ .5 & .5\end{bmatrix}, \begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}\right]$$

$$= 2\begin{bmatrix}1\\ 1\end{bmatrix} \otimes \begin{bmatrix}1\\ -1\end{bmatrix} \otimes \begin{bmatrix}1\\ -1\end{bmatrix} + (-.5)\begin{bmatrix}1\\ -1\end{bmatrix} \otimes \begin{bmatrix}1\\ 1\end{bmatrix} \otimes \begin{bmatrix}1\\ 0\end{bmatrix}$$

[1] J. Håstad, Journal of Algorithms, 1990

Example

$$\left[\begin{bmatrix}1.5 & -2.5\\2.5 & -2.5\end{bmatrix},\begin{bmatrix}-2 & 2\\-2 & 2\end{bmatrix}\right] = \left[\begin{bmatrix}2 & -2\\2 & -2\end{bmatrix},\begin{bmatrix}-2 & 2\\-2 & 2\end{bmatrix}\right] + \left[\begin{bmatrix}-.5 & -.5\\.5 & .5\end{bmatrix},\begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}\right]$$
$$= 2\begin{bmatrix}1\\1\end{bmatrix}\otimes\begin{bmatrix}1\\-1\end{bmatrix}\otimes\begin{bmatrix}1\\-1\end{bmatrix} + (-.5)\begin{bmatrix}1\\-1\end{bmatrix}\otimes\begin{bmatrix}1\\1\end{bmatrix}\otimes\begin{bmatrix}1\\0\end{bmatrix}$$

The rank of this decomposition is 2

However, we also find

$$\left[\begin{bmatrix}1.5 & -2.5\\2.5 & -2.5\end{bmatrix},\begin{bmatrix}-2 & 2\\-2 & 2\end{bmatrix}\right]$$

[1] J. Håstad, Journal of Algorithms, 1990

Example

$$\left[\begin{bmatrix}1.5 & -2.5\\ 2.5 & -2.5\end{bmatrix}, \begin{bmatrix}-2 & 2\\ -2 & 2\end{bmatrix}\right] = \left[\begin{bmatrix}2 & -2\\ 2 & -2\end{bmatrix}, \begin{bmatrix}-2 & 2\\ -2 & 2\end{bmatrix}\right] + \left[\begin{bmatrix}-.5 & -.5\\ .5 & .5\end{bmatrix}, \begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}\right]$$
$$= 2\begin{bmatrix}1\\1\end{bmatrix} \otimes \begin{bmatrix}1\\-1\end{bmatrix} \otimes \begin{bmatrix}1\\-1\end{bmatrix} + (-.5)\begin{bmatrix}1\\-1\end{bmatrix} \otimes \begin{bmatrix}1\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix}$$

The rank of this decomposition is 2
However, we also find

$$\left[\begin{bmatrix}1.5 & -2.5\\ 2.5 & -2.5\end{bmatrix}, \begin{bmatrix}-2 & 2\\ -2 & 2\end{bmatrix}\right] = 1.5\begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} + (-2)\begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}$$
$$+ (-2.5)\begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} + 2\begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}$$
$$+ 2.5\begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} + (-2)\begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}$$
$$+ (-2.5)\begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} + 2\begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}$$

[1] J. Håstad, Journal of Algorithms, 1990

Example

$$\left[\begin{bmatrix}1.5 & -2.5\\2.5 & -2.5\end{bmatrix}, \begin{bmatrix}-2 & 2\\-2 & 2\end{bmatrix}\right] = \left[\begin{bmatrix}2 & -2\\2 & -2\end{bmatrix}, \begin{bmatrix}-2 & 2\\-2 & 2\end{bmatrix}\right] + \left[\begin{bmatrix}-.5 & -.5\\.5 & .5\end{bmatrix}, \begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}\right]$$
$$= 2\begin{bmatrix}1\\1\end{bmatrix} \otimes \begin{bmatrix}1\\-1\end{bmatrix} \otimes \begin{bmatrix}1\\-1\end{bmatrix} + (-.5)\begin{bmatrix}1\\-1\end{bmatrix} \otimes \begin{bmatrix}1\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix}$$

The rank of this decomposition is 2

However, we also find

$$\left[\begin{bmatrix}1.5 & -2.5\\2.5 & -2.5\end{bmatrix}, \begin{bmatrix}-2 & 2\\-2 & 2\end{bmatrix}\right] = 1.5\begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} + (-2)\begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}$$
$$+ (-2.5)\begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} + 2\begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}$$
$$+ 2.5\begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} + (-2)\begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}$$
$$+ (-2.5)\begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} + 2\begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}$$

The rank of this decomposition is 8.

[1] J. Håstad, Journal of Algorithms, 1990

Example

$$\left[\left[\begin{bmatrix}1.5 & -2.5\\ 2.5 & -2.5\end{bmatrix}, \begin{bmatrix}-2 & 2\\ -2 & 2\end{bmatrix}\right]\right] = \left[\left[\begin{bmatrix}2 & -2\\ 2 & -2\end{bmatrix}, \begin{bmatrix}-2 & 2\\ -2 & 2\end{bmatrix}\right]\right] + \left[\left[\begin{bmatrix}-.5 & -.5\\ .5 & .5\end{bmatrix}, \begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}\right]\right]$$
$$= 2\begin{bmatrix}1\\1\end{bmatrix} \otimes \begin{bmatrix}1\\-1\end{bmatrix} \otimes \begin{bmatrix}1\\-1\end{bmatrix} + (-.5)\begin{bmatrix}1\\-1\end{bmatrix} \otimes \begin{bmatrix}1\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix}$$

The rank of this decomposition is 2
However, we also find

$$\left[\left[\begin{bmatrix}1.5 & -2.5\\ 2.5 & -2.5\end{bmatrix}, \begin{bmatrix}-2 & 2\\ -2 & 2\end{bmatrix}\right]\right] = 1.5\begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} + (-2)\begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}$$
$$+ (-2.5)\begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} + 2\begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}$$
$$+ 2.5\begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} + (-2)\begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}$$
$$+ (-2.5)\begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} + 2\begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}$$

The rank of this decomposition is 8.
Deciding whether a rational tensor has CP-rank r is NP-hard [1]

[1] J. Håstad, Journal of Algorithms, 1990

CP deocposition

Given a tensor $\mathbf{X}$, we seek to find

$$\mathbf{X}_* = \underset{\text{CP}-\text{rank}(\mathbf{X}_r)\leq r}{\operatorname{argmin}} \|\mathbf{X} - \mathbf{X}_r\| \tag{1}$$

Matrices:

- Eckart–Young gives insight for unitarily invariant matrices

Tensors

- For many tensor ranks $r \geq 2$ and all orders $d \geq 3$, regardless of the choice of $\| \cdot \|$:

  Eq. (1) is ill-defined[2]!

- There are methods calculating approximate CP decompositions of higher-order tensors
  → Challenging and expensive task
  → In practice approached using optimization algorithms

[2] De Silva & Lim, SIAM Journal on Matrix Analysis and Applications, 2008

Set of Tensors with Fixed Canonical Rank

Ill-Definedness of Eq. (1) can be connected to the following problem:

Let's consider

$$\mathcal{M}_{\leq r} = \left\{ \mathbf{X} \in \mathbb{R}^{n_1 \times \ldots \times n_d} \mid \mathrm{CP} - \mathrm{rank}(\mathbf{X}) \leq r \right\}$$

the sequence

$$\mathbf{X}_n = n \left(\mathbf{u} + \frac{1}{n} \mathbf{v} \right) \otimes \left(\mathbf{u} + \frac{1}{n} \mathbf{v} \right) \otimes \left(\mathbf{u} + \frac{1}{n} \mathbf{v} \right) - n \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{u}$$

with $\mathbf{u}$, $\mathbf{v} \in \mathbb{R}^m$, $\|\mathbf{v}\| = \|\mathbf{u}\| = 1$ and $\langle \mathbf{v}, \mathbf{u} \rangle \neq 1$.

Set of Tensors with Fixed Canonical Rank

Ill-Definedness of Eq. (1) can be connected to the following problem:

Let's consider

$$\mathcal{M}_{\leq r} = \left\{ \mathbf{X} \in \mathbb{R}^{n_1 \times \ldots \times n_d} \mid \mathrm{CP} - \mathrm{rank}(\mathbf{X}) \leq r \right\}$$

the sequence

$$\mathbf{X}_n = n \left(\mathbf{u} + \frac{1}{n}\mathbf{v} \right) \otimes \left(\mathbf{u} + \frac{1}{n}\mathbf{v} \right) \otimes \left(\mathbf{u} + \frac{1}{n}\mathbf{v} \right) - n\mathbf{u} \otimes \mathbf{u} \otimes \mathbf{u}$$

with $\mathbf{u}$, $\mathbf{v} \in \mathbb{R}^m$, $\|\mathbf{v}\| = \|\mathbf{u}\| = 1$ and $\langle \mathbf{v}, \mathbf{u} \rangle \neq 1$.
Note that $\mathbf{X}_n \in \mathcal{M}_{\leq r}$ for all $n \in \mathbb{N}$, however

$$\lim_{n \to \infty} \mathbf{X}_n = \mathbf{v} \otimes \mathbf{u} \otimes \mathbf{u} + \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{u} + \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{v} \notin \mathcal{M}_{\leq r}$$

Set of Tensors with Fixed Canonical Rank

Ill-Definedness of Eq. (1) can be connected to the following problem:

Let's consider

$$\mathcal{M}_{\leq r} = \left\{ \mathbf{X} \in \mathbb{R}^{n_1 \times \ldots \times n_d} \mid \mathrm{CP} - \mathrm{rank}(\mathbf{X}) \leq r \right\}$$

the sequence

$$\mathbf{X}_n = n \left(\mathbf{u} + \frac{1}{n} \mathbf{v} \right) \otimes \left(\mathbf{u} + \frac{1}{n} \mathbf{v} \right) \otimes \left(\mathbf{u} + \frac{1}{n} \mathbf{v} \right) - n \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{u}$$

with $\mathbf{u}$, $\mathbf{v} \in \mathbb{R}^m$, $\|\mathbf{v}\| = \|\mathbf{u}\| = 1$ and $\langle \mathbf{v}, \mathbf{u} \rangle \neq 1$.
Note that $\mathbf{X}_n \in \mathcal{M}_{\leq r}$ for all $n \in \mathbb{N}$, however

$$\lim_{n \to \infty} \mathbf{X}_n = \mathbf{v} \otimes \mathbf{u} \otimes \mathbf{u} + \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{u} + \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{v} \notin \mathcal{M}_{\leq r}$$

Similarly

$$\mathcal{M}_{r} = \left\{ \mathbf{X} \in \mathbb{R}^{n_1 \times \ldots \times n_d} \mid \mathrm{CP} - \mathrm{rank}(\mathbf{X}) = r \right\}$$

is not closed.

Difficulties CP format

- CP decomposition sets have very little structure
- Low-rank matrices for manifolds
  $\rightarrow$ we can use optimization techniques on Manifolds
- CP rank tensor do not form any kind of manifold $\rightarrow$ optimization on such sets is extremely difficult
- The approximation is ambiguous
  $\rightarrow$ Many parameters $\mathbf{v}_{p,i}$ approximate the same tensor equally well

$\Rightarrow$ More in Mitchell & Burdick, Journal of Chemometrics, 1994

The CP format allows an unparalleled complexity reduction for tensors with small canonical rank!

Computational Aspects of CP decomposition

Recall:

Storing a tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times \ldots \times n_d}$ requires $\mathcal{O}(n^d)$, where $n = \max_i n_i$.

Computational Aspects of CP decomposition

Recall:

Storing a tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times \ldots \times n_d}$ requires $\mathcal{O}(n^d)$, where $n = \max_i n_i$.

In the CP format, we store the vector entries $\mathbf{v}_{i,p}$.
$\rightarrow$ requires $\mathcal{O}(ndr)$
$\rightarrow$ linearly in the dimension

What about operations?

Addition in CP format

Consider
Then the addition of $\mathbf{X}$ and $\bar{\mathbf{X}}$ i.e.,

$$\mathbf{X} + \bar{\mathbf{X}} = \sum_{p=1}^{r} \bigotimes_{i=1}^{d} \mathbf{v}_{i,p} + \sum_{p=1}^{\bar{r}} \bigotimes_{i=1}^{d} \bar{\mathbf{v}}_{i,p} = \sum_{p=1}^{\bar{r}+r} \bigotimes_{i=1}^{d} \mathbf{W}_{i,p}$$

with

$$\mathbf{w}_{i,p} = \{ \mathbf{v}_{i,p} \quad k \leq r \bar{\mathbf{v}}_{i,p} \qquad k > r \tag{2}$$

In order to access the element, we have to perform the following operation

$$(\mathbf{X} + \bar{\mathbf{X}})[i_1, ..., i_d] = \left(\sum_{p=1}^{\bar{r}+r} \bigotimes_{k=1}^{d} \mathbf{W}_{i,p} \right) [i_1, ..., i_d] = \sum_{p=1}^{\bar{r}+r} \prod_{k=1}^{d} \mathbf{W}_{i,p}[i_k]$$

Which scales as $\mathcal{O}(nd(\bar{r} + r))$, compared to adding two dense tensors $\mathcal{O}(n^d)$

kth-mode contraction

Given a matrix $\mathbf{A} \in \mathbb{R}^{n_k \times m}$. Then

$$\begin{aligned}
\mathbf{X} *_k \mathbf{A} &= \left(\sum_{p=1}^{\bar{r}+r} \bigotimes_{k=1}^{d} \mathbf{W}_{i,p} \right) *_k \mathbf{A} \\
&= \sum_{p=1}^{\bar{r}+r} \left(\bigotimes_{k=1}^{d} \mathbf{W}_{i,p} \right) *_k \mathbf{A} \\
&= \sum_{p=1}^{\bar{r}+r} \mathbf{v}_{1,p} \otimes \ldots \otimes \left(\mathbf{A}^\top \mathbf{v}_{k,p} \right) \otimes \ldots \otimes \mathbf{v}_{d,p}
\end{aligned}$$

Other tensor operations in CP format

Operation	CP-Format	dense tensor
Hadamard Product	$\mathcal{O}(ndr\bar{r})$	$\mathcal{O}(n^d)$
Frobenius Inner Product	$\mathcal{O}(ndr\bar{r})$	$\mathcal{O}(n^d)$
Frobenius Norm	$\mathcal{O}(ndr^2)$	$\mathcal{O}(n^d)$
k-mode product	$\mathcal{O}((d+m)nr)$	$\mathcal{O}(n^d m)$