Tensor Decompositions, Matrix Completion and Singular Values

Harm Derksen

University of Michigan

ICERM, Computational Nonlinear Algebra
June 3, 2014

Tensor Product Spaces

$\mathbb{F}$ a field
$V^{(i)} \cong \mathbb{F}^{n_i}$ $\mathbb{F}$-vector space for $i = 1, 2, \ldots, d$
$V = V^{(1)} \otimes V^{(2)} \otimes \cdots \otimes V^{(d)} \cong \mathbb{F}^{n_1 \times \cdots \times n_d}$ tensor product space

Definition

A *pure* tensor is a tensor of the form $v^{(1)} \otimes v^{(2)} \otimes \cdots \otimes v^{(d)}$ $(v^{(i)} \in V^{(i)})$.

Tensor Product Spaces

$\mathbb{F}$ a field
$V^{(i)} \cong \mathbb{F}^{n_i}$ $\mathbb{F}$-vector space for $i = 1, 2, \dots, d$
$V = V^{(1)} \otimes V^{(2)} \otimes \cdots \otimes V^{(d)} \cong \mathbb{F}^{n_1 \times \cdots \times n_d}$ tensor product space

Definition

A *pure* tensor is a tensor of the form $v^{(1)} \otimes v^{(2)} \otimes \cdots \otimes v^{(d)}$ ($v^{(i)} \in V^{(i)}$).

Problem

Write a given tensor as a sum of the smallest number of pure tensors.

(for $\mathbb{F} = \mathbb{R}, \mathbb{C}$: Canonical Polyadic (CP) decompositions, PARAFAC, CANDECOMP)

Applications

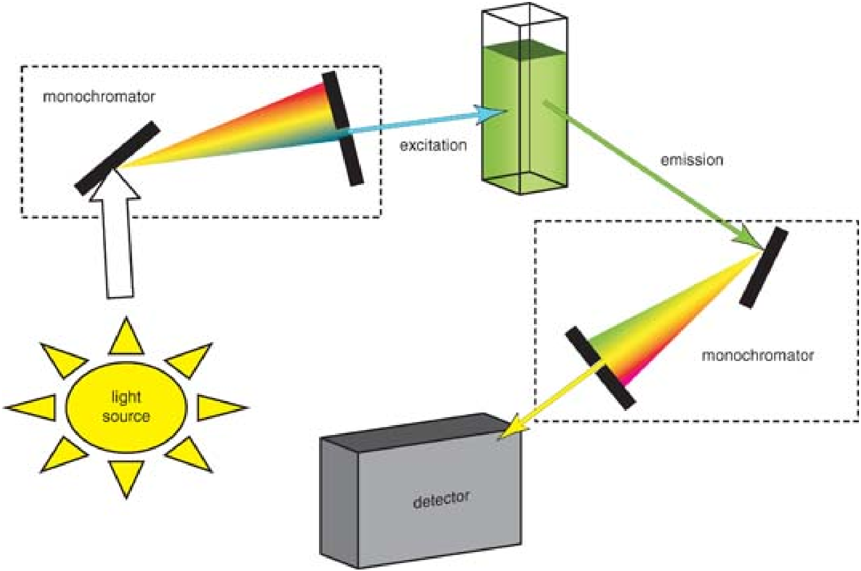
- psychometrics
- chemometrics
- algebraic complexity theory
- signal processing
- numerical linear algebra
- computer vision
- numerical analysis
- data mining
- graph analysis
- neuroscience
- economics/finance

Application: Fluorescence Spectroscopy

We have p samples of mixtures of unknown chemical compounds. Every mixture is excited with light of m different wavelengths. Light of n wavelengths is emitted from the mixture . Measuring the intensities, one obtains an $m \times n$ excitation-emisson matrix for every sample. This yields a $p \times m \times n$ matrix, which is a 3-way tensor T.
Every chemical compound corresponds to a rank 1 tensor. By writing T as the sum of $r = \text{rank}(T)$ pure tensors, we can distinguish r chemical compounds and find the excitation-emission matrix for each of them.

Application:Fluorescence Spectroscopy

(Image: Lei Li, Andrew Barron)



Tensor rank

Definition

The *rank* of a tensor $T \in V$ is the smallest nonnegative integer r such that we can write T as a sum of r pure tensors.

Tensor rank

Definition

The *rank* of a tensor $T \in V$ is the smallest nonnegative integer r such that we can write T as a sum of r pure tensors.

For example,

$$T = e_1 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_2 + e_2 \otimes e_1 \otimes e_2 + e_2 \otimes e_2 \otimes e_1$$

Tensor rank

Definition

The *rank* of a tensor $T \in V$ is the smallest nonnegative integer r such that we can write T as a sum of r pure tensors.

For example,

$$T = e_1 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_2 + e_2 \otimes e_1 \otimes e_2 + e_2 \otimes e_2 \otimes e_1$$

has rank 2 because

$$T = \tfrac{1}{2}(e_1+e_2)\otimes(e_1+e_2)\otimes(e_1+e_2)+\tfrac{1}{2}(e_1-e_2)\otimes(e_1-e_2)\otimes(e_1-e_2)$$

Tensor rank

Definition

The *rank* of a tensor $T \in V$ is the smallest nonnegative integer r such that we can write T as a sum of r pure tensors.

For example,

$$T = e_1 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_2 + e_2 \otimes e_1 \otimes e_2 + e_2 \otimes e_2 \otimes e_1$$

has rank 2 because

$$T = \tfrac{1}{2}(e_1+e_2)\otimes(e_1+e_2)\otimes(e_1+e_2)+\tfrac{1}{2}(e_1-e_2)\otimes(e_1-e_2)\otimes(e_1-e_2)$$

If $d = 2$ and $V = V^{(1)} \otimes V^{(2)} \cong \mathbb{R}^{d_1 \times d_2}$ then the tensor rank is just the rank of a matrix.

Application: Algebraic Complexity Theory

$V = \mathrm{Mat}_{n,n}(\mathbb{F}) \otimes \mathrm{Mat}_{n,n}(\mathbb{F}) \otimes \mathrm{Mat}_{n,n}(\mathbb{F})$

$$T_n = \sum_{i,j,k=1}^{n} e_{i,j} \otimes e_{j,k} \otimes e_{k,i}$$

$\mathrm{rank}(T_n)$ is the number of multiplications needed to multiply two $n \times n$ matrices. Clearly $\mathrm{rank}(T_n) \leq n^3$.

Application: Algebraic Complexity Theory

$V = \mathrm{Mat}_{n,n}(\mathbb{F}) \otimes \mathrm{Mat}_{n,n}(\mathbb{F}) \otimes \mathrm{Mat}_{n,n}(\mathbb{F})$

$$T_n = \sum_{i,j,k=1}^{n} e_{i,j} \otimes e_{j,k} \otimes e_{k,i}$$

$\mathrm{rank}(T_n)$ is the number of multiplications needed to multiply two $n \times n$ matrices. Clearly $\mathrm{rank}(T_n) \leq n^3$.

Theorem

If $\mathrm{rank}(T_m) \leq k$, *then* $\mathrm{rank}(T_n) = O(n^{\log_m(k)})$ *and two* $n \times n$ *matrices can be multiplied using* $O(n^{\log_m(k)})$ *arithmetic operations.*

Application: Algebraic Complexity Theory

Theorem (Strassen 1969)

$\text{rank}(T_2) \leq 7$, *so* $\text{rank}(T_n) = O(n^{\log_2(7)}) = O(n^{2.8073\dots})$.

Theorem (Williams 2012)

$\text{rank}(T_n) = O(n^{2.3727})$

Theorem (Masseranti, Raviolo 2013)

$$\text{rank}(T_n) \geq 3n^2 - 2\sqrt{2}n^{3/2} - 3n.$$

Low Rank Matrix Completion

Problem

Given a partially filled matrix, complete the matrix such that the resulting matrix has the smallest possible rank.

Problem

Given a partially filled matrix, complete the matrix such that the resulting matrix has the smallest possible rank.

For example

$$\begin{pmatrix} -2 & \cdot & \cdot \\ 2 & 3 & \cdot \\ \cdot & 6 & 2 \end{pmatrix}$$

Low Rank Matrix Completion

Problem

Given a partially filled matrix, complete the matrix such that the resulting matrix has the smallest possible rank.

For example

$$\begin{pmatrix} -2 & \cdot & \cdot \\ 2 & 3 & \cdot \\ \cdot & 6 & 2 \end{pmatrix}$$

can be completed to a rank 1 matrix

$$\begin{pmatrix} -2 & -3 & -1 \\ 2 & 3 & 1 \\ 4 & 6 & 2 \end{pmatrix}$$

Application: The Netflix Problem

The DVD rental company Netflix has 480,189 users, and 17,770 movies. The user ratings for every user can be put in a $480,189 \times 17,770$ matrix $A = (a_{i,j})$ for which only few entries are known (since most users have only seen a fraction of the 17,770 movies).

Presumably, the rank of the matrix A is low. Using Low rank matrix completion, Netflix can predict whether users like a movie they have seen, and will recommend movies to the users.

Reduction LRMC to Tensor Rank

A is $n \times m$ matrix
are allowed to change entries in positions $(i_1, j_1), \ldots, (i_s, j_s)$.
$\text{mrank}(A)$ minimal possible rank

Reduction LRMC to Tensor Rank

A is $n \times m$ matrix
are allowed to change entries in positions $(i_1, j_1), \ldots, (i_s, j_s)$.
$\text{mrank}(A)$ minimal possible rank

Define

$$T = \sum_{k=1}^{s} e_{i_k} \otimes e_{j_k} \otimes e_k + A \otimes e_{s+1} \in \mathbb{C}^n \otimes \mathbb{C}^m \otimes \mathbb{C}^{s+1}$$

Theorem (D.)

$\text{rank}(T) = \text{mrank}(A) + s$

Example

$$A = \begin{pmatrix} 1 & t \\ \cdot & 1 \end{pmatrix}$$

For $T = e_2 \otimes e_1 \otimes e_1 + e_1 \otimes (e_1 + te_2) \otimes e_2 + e_2 \otimes e_2 \otimes e_2$ we have $\text{rank}(T) = 1 + \text{mrank}(A)$.

Example

$$A = \begin{pmatrix} 1 & t \\ \cdot & 1 \end{pmatrix}$$

For $T = e_2 \otimes e_1 \otimes e_1 + e_1 \otimes (e_1 + te_2) \otimes e_2 + e_2 \otimes e_2 \otimes e_2$ we have $\text{rank}(T) = 1 + \text{mrank}(A)$.

$\text{mrank}(A) = 1$ if $t \neq 0$ and $\text{mrank}(A) = 2$ if $t = 0$.

Example

$$A = \begin{pmatrix} 1 & t \\ \cdot & 1 \end{pmatrix}$$

For $T = e_2 \otimes e_1 \otimes e_1 + e_1 \otimes (e_1 + te_2) \otimes e_2 + e_2 \otimes e_2 \otimes e_2$ we have $\text{rank}(T) = 1 + \text{mrank}(A)$.

$\text{mrank}(A) = 1$ if $t \neq 0$ and $\text{mrank}(A) = 2$ if $t = 0$.

If $t = 0$, then $T = e_2 \otimes e_1 \otimes e_1 + e_1 \otimes e_1 \otimes e_2 + e_2 \otimes e_2 \otimes e_2$ has rank 3.

Example

$$A = \begin{pmatrix} 1 & t \\ \cdot & 1 \end{pmatrix}$$

For $T = e_2 \otimes e_1 \otimes e_1 + e_1 \otimes (e_1 + te_2) \otimes e_2 + e_2 \otimes e_2 \otimes e_2$ we have $\text{rank}(T) = 1 + \text{mrank}(A)$.

$\text{mrank}(A) = 1$ if $t \neq 0$ and $\text{mrank}(A) = 2$ if $t = 0$.

If $t = 0$, then $T = e_2 \otimes e_1 \otimes e_1 + e_1 \otimes e_1 \otimes e_2 + e_2 \otimes e_2 \otimes e_2$ has rank 3.

If $t \neq 0$ then $\text{rank}(T) = 2$ and

$$T = e_2 \otimes e_1 \otimes (e_1 - t^{-1}e_2) + (e_1 + t^{-1}e_2) \otimes (e_1 + te_2) \otimes e_2$$

Rank Minimization Problem (RM)

Problem (Rank Minimization)

Given $A, B_1, \dots, B_s \in \mathrm{Mat}_{p,q}(\mathbb{F})$, *minimize* $\mathrm{rank}(A + x_1 B_1 + \cdots + x_s B_s)$.

LRMC is a special case of RM where $B_k = e_{i_k, j_k}$.

Rank Minimization Problem (RM)

Problem (Rank Minimization)

Given $A, B_1, \ldots, B_s \in \mathrm{Mat}_{p,q}(\mathbb{F})$, *minimize* $\mathrm{rank}(A + x_1 B_1 + \cdots + x_s B_s)$.

LRMC is a special case of RM where $B_k = e_{i_k, j_k}$.

Theorem (D.)

The Tensor Rank problem can be reduced to RM and RM can be reduced to LRMC.

For example, suppose that $T = (t_{i,j,k})_{i,j,k=1}^{2}$ is a $2 \times 2 \times 2$ tensor.

$$
\left(\begin{array}{cccccccc|cccccccc|c|c|c}
t_{1,1,1} & t_{1,2,1} & a_{1,1} & a_{1,2} & a_{1,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
t_{2,1,1} & t_{2,2,1} & a_{2,1} & a_{2,2} & a_{2,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & \lambda_{1,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & \lambda_{1,2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & \lambda_{1,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
b_{1,1} & b_{2,1} & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
b_{1,2} & b_{2,2} & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
b_{1,3} & b_{2,3} & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & t_{1,1,2} & t_{1,2,2} & a_{1,1} & a_{1,2} & a_{1,3} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & t_{2,1,2} & t_{2,2,2} & a_{2,1} & a_{2,2} & a_{2,3} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \lambda_{2,1} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \lambda_{2,2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \lambda_{2,3} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & b_{1,1} & b_{2,1} & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & b_{1,2} & b_{2,2} & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & b_{1,3} & b_{2,3} & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{1,1} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{2,1} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & b_{1,1} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & b_{2,1} & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{1,2} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{2,2} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & b_{1,2} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & b_{2,2} & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{1,3} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{2,3} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & b_{1,3} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & b_{2,3}
\end{array}\right)
$$

The smallest rank over all $a_{i,j}, b_{i,j}, \lambda_{i,j}$ $(1 \leq i \leq 2, 1 \leq j \leq 3)$ is $12 + \text{rank}(T)$.

Motivation: compressed sensing and convex relaxation

For $x \in \mathbb{R}^n$, its sparsity is measured by $\|x\|_0 = \#\{i \mid x_i \neq 0\}$.

Problem

Given $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$, $b \in \mathbb{R}^m$, find a solution $x \in \mathbb{R}^n$ for $Ax = b$ with $\|x\|_0$ minimal (a sparsest solution).

But, $\|\cdot\|_0$ is not convex and this optimization problem is difficult,

Motivation: compressed sensing and convex relaxation

For $x \in \mathbb{R}^n$, its sparsity is measured by $\|x\|_0 = \#\{i \mid x_i \neq 0\}$.

Problem

Given $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$, $b \in \mathbb{R}^m$, *find a solution* $x \in \mathbb{R}^n$ *for* $Ax = b$ *with* $\|x\|_0$ *minimal (a sparsest solution).*

But, $\|\cdot\|_0$ is not convex and this optimization problem is difficult, so instead we consider:

Problem (Basis Pursuit)

Given $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$, $b \in \mathbb{R}^m$, *find a solution* $x \in \mathbb{R}^n$ *for* $Ax = b$ *with* $\|x\|_1$ *minimal.*

Basis Pursuit can be solved by linear programming and is generally fast. Under reasonable assumptions, Basis Pursuit also gives the sparsest solutions (Candès-Tao, Donoho).

Nuclear Norm

$V^{(i)}$ Hilbert space. Instead of the CP model, we consider:

Problem (Convex Decomposition)

Given a tensor T, write $T = \sum_{i=1}^{r} v_i$ for some r and some pure tensors $v_1, \dots, v_r$ such that $\sum_{i=1}^{r} \|v_i\|_2$ is minimal.

Finding a convex decomposition seems to be easier than finding a CP decomposition. Heuristically, we expect convex decompositions to give a CP decomposition or at least a low rank decomposition.

Nuclear Norm

$V^{(i)}$ Hilbert space. Instead of the CP model, we consider:

Problem (Convex Decomposition)

Given a tensor T, write $T = \sum_{i=1}^{r} v_i$ for some r and some pure tensors $v_1, \ldots, v_r$ such that $\sum_{i=1}^{r} \|v_i\|_2$ is minimal.

Finding a convex decomposition seems to be easier than finding a CP decomposition. Heuristically, we expect convex decompositions to give a CP decomposition or at least a low rank decomposition.

Definition (Lim-Comon)

The nuclear norm $\|T\|_\star$ is the smallest value of $\sum_{i=1}^{r} \|v_i\|_2$ where $T = \sum_{i=1}^{r} v_i$ and $v_1, \ldots, v_r$ are pure tensors.

For some tensors we know the nuclear norm but not the rank.

Spectral Norm

Definition

The spectral norm is defined by

$$\|T\|_\sigma = \max\{|\langle T, v\rangle| \mid v \text{ pure tensor with } \|v\|_2 = 1\}.$$

The spectral norm is *dual* to the nuclear norm, in particular

$$|\langle T, S\rangle| \leq \|T\|_\star \|S\|_\sigma$$

for all tensors S, T.

Example: Determinant Tensor

Consider the tensor

$$D_n = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) e_{\sigma(1)} \otimes e_{\sigma(2)} \otimes \cdots \otimes e_{\sigma(n)} \in \mathbb{C}^n \otimes \cdots \otimes \mathbb{C}^n.$$

Clearly $\operatorname{rank}(D_n) \leq n!$, actually $\binom{n}{\lfloor n/2 \rfloor} \leq \operatorname{rank}(D_n) \leq (\frac{5}{6})^{\lfloor n/3 \rfloor} n!$.

Example: Determinant Tensor

Consider the tensor

$$D_n = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) e_{\sigma(1)} \otimes e_{\sigma(2)} \otimes \cdots \otimes e_{\sigma(n)} \in \mathbb{C}^n \otimes \cdots \otimes \mathbb{C}^n.$$

Clearly $\operatorname{rank}(D_n) \leq n!$, actually $\binom{n}{\lfloor n/2 \rfloor} \leq \operatorname{rank}(D_n) \leq (\frac{5}{6})^{\lfloor n/3 \rfloor} n!$.

$$\|D_n\|_\sigma = \max\{|\det(v_1 v_2 \cdots v_n)| \mid \|v_1\|_2 = \cdots = \|v_n\|_2 = 1\} = 1$$

by Hadamard's inequality.

$$\|D_n\|_\star = \|D_n\|_\star \|D_n\|_\sigma \geq \langle D_n, D_n \rangle = n!, \text{ so}$$

Theorem (D.)

$$\|D_n\|_\star = n!$$

Example: Permanent Tensor

$$P_n = \sum_{\sigma \in S_n} e_{\sigma(1)} \otimes e_{\sigma(2)} \otimes \cdots \otimes e_{\sigma(n)} \in \mathbb{C}^n \otimes \cdots \otimes \mathbb{C}^n.$$

$$\|P_n\|_\sigma = \max\{|\operatorname{perm}(v_1 v_2 \cdots v_n)| \mid \|v_1\|_2 = \cdots = \|v_n\|_2 = 1\}$$

Example: Permanent Tensor

$$P_n = \sum_{\sigma \in S_n} e_{\sigma(1)} \otimes e_{\sigma(2)} \otimes \cdots \otimes e_{\sigma(n)} \in \mathbb{C}^n \otimes \cdots \otimes \mathbb{C}^n.$$

$$\|P_n\|_\sigma = \max\{|\operatorname{perm}(v_1 v_2 \cdots v_n)| \mid \|v_1\|_2 = \cdots = \|v_n\|_2 = 1\}$$

Theorem (Carlen, Lieb and Moss, 2006)

$$\max\{\operatorname{perm}(v_1 v_2 \cdots v_n) \mid \|v_1\|_2 = \cdots = \|v_n\|_2 = 1\} = n!/n^{n/2}$$

$$\|P_n\|_\star = \frac{n^{n/2}}{n!} \|P_n\|_\star \|P_n\|_\sigma \geq \frac{n^{n/2}}{n!} \langle P_n, P_n \rangle = n^{n/2}.$$

Example: Permanent Tensor

Theorem (Glynn 2010)

$$P_n = \frac{1}{2^{n-1}} \sum_{\delta} \Big(\prod_{i=1}^n \delta_i \Big) (\sum_{i=1}^n \delta_i e_i) \otimes \cdots \otimes (\sum_{i=1}^n \delta_i e_i)$$

where δ *runs over* $\{1\} \times \{-1, 1\}^{n-1}$.

Example: Permanent Tensor

Theorem (Glynn 2010)

$$P_n = \frac{1}{2^{n-1}} \sum_{\delta} \Big(\prod_{i=1}^n \delta_i \Big) (\sum_{i=1}^n \delta_i e_i) \otimes \cdots \otimes (\sum_{i=1}^n \delta_i e_i)$$

where δ runs over $\{1\} \times \{-1,1\}^{n-1}$.

In particular, $\binom{n}{\lfloor n/2 \rfloor} \leq \operatorname{rank}(P_n) \leq 2^{n-1}$ and $\|P_n\|_\star \leq n^{n/2}$, so

Theorem (D.)

$$\|P_n\|_\star = n^{n/2}$$

t-Orthogonality

Definition

Pure tensors $v_1, v_2, \ldots, v_r$ are t-orthogonal if

$$\sum_{i=1}^{r} |\langle v_i, w\rangle|^{2/t} \leq 1$$

for every pure tensor w with $\|w\|_2 = 1$.

t-Orthogonality

Definition

Pure tensors $v_1, v_2, \ldots, v_r$ are t-orthogonal if

$$\sum_{i=1}^{r} |\langle v_i, w\rangle|^{2/t} \leq 1$$

for every pure tensor w with $\|w\|_2 = 1$.

1-orthogonal $\Leftrightarrow$ orthogonal

If $t > s$ then t-orthogonal $\Rightarrow$ s-orthogonal.

t-Orthogonality

Definition

Pure tensors $v_1, v_2, \dots, v_r$ are t-orthogonal if

$$\sum_{i=1}^{r} |\langle v_i, w\rangle|^{2/t} \leq 1$$

for every pure tensor w with $\|w\|_2 = 1$.

1-orthogonal $\Leftrightarrow$ orthogonal
If $t > s$ then t-orthogonal $\Rightarrow$ s-orthogonal.

Theorem (D.)

If $v_1, \dots, v_r \in V$ are t-orthogonal, then $r \leq \dim(V)^{1/t}$.

Horizontal and Vertical Tensor Product

Theorem ("horizontal tensor product", D.)

If $v_1, \dots, v_r$ *are t-orthogonal, and* $w_1, \dots, w_r$ *are s-orthogonal, then* $v_1 \otimes w_1, \dots, v_r \otimes w_r$ *are* $(s+t)$*-orthogonal.*

Horizontal and Vertical Tensor Product

Theorem ("horizontal tensor product", D.)

If $v_1, \dots, v_r$ are t-orthogonal, and $w_1, \dots, w_r$ are s-orthogonal, then $v_1 \otimes w_1, \dots, v_r \otimes w_r$ are $(s+t)$-orthogonal.

If $V = V^{(1)} \otimes \cdots \otimes V^{(d)}$ and $W = W^{(1)} \otimes \cdots \otimes W^{(d)}$, then

$$V \boxtimes W := (V^{(1)} \otimes W^{(1)}) \otimes \cdots \otimes (V^{(d)} \otimes W^{(d)}).$$

Theorem ("vertical tensor product", D.)

If $v_1, v_2, \dots, v_r \in V$ and $w_1, \dots, w_s \in W$ are t-orthogonal, then $\{v_i \boxtimes w_j \mid 1 \leq i \leq r, 1 \leq j \leq s\}$ are t-orthogonal.

The Diagonal Singular Value Decomposition

Definition

Suppose that $(\star) : T = \sum_{i=1}^{r} \lambda_i v_i$ such that $\lambda_1 \geq \cdots \geq \lambda_r > 0$ and $v_1, \ldots, v_r$ are 2-orthogonal pure tensors of unit length, then $(\star)$ is called a *diagonal singular value decomposition* of T (DSVD).

If $d = 2$ (tensor product of 2 spaces) then the DSVD is the usual singular value decomposition. For $d > 2$, the DSVD is different from the *Higher Order Singular Value Decomposition* defined by De Lathauer, De Moor, and Vandewalle. Not every tensor has a DSVD.

The Diagonal Singular Value Decomposition

Theorem (D.)

If T has a DSVD then

$$\|T\|_\star = \sum_i \lambda_i, \quad \|T\|_2 = \sqrt{\sum_i \lambda_i^2}, \quad \|T\|_\sigma = \lambda_1$$

The Diagonal Singular Value Decomposition

Theorem (D.)

If T has a DSVD then

$$\|T\|_\star = \sum_i \lambda_i, \quad \|T\|_2 = \sqrt{\sum_i \lambda_i^2}, \quad \|T\|_\sigma = \lambda_1$$

Theorem (D.)

If $\lambda_1 > \lambda_2 > \cdots > \lambda_r$ then the DSVD is unique.

Theorem (D.)

If $v_1, \ldots, v_r$ are t-orthogonal with $t > 2$, then the DSVD is unique.

Example: Matrix Multiplication Tensor

$e_1, \ldots, e_n \in \mathbb{C}^n$ are orthogonal
$e_1 \otimes e_1, \ldots, e_n \otimes e_n \in \mathbb{C}^n \otimes \mathbb{C}^n$ are 2-orthogonal

Example: Matrix Multiplication Tensor

$e_1, \ldots, e_n \in \mathbb{C}^n$ are orthogonal
$e_1 \otimes e_1, \ldots, e_n \otimes e_n \in \mathbb{C}^n \otimes \mathbb{C}^n$ are 2-orthogonal

$e_1 \otimes e_1 \otimes 1, \ldots, e_n \otimes e_n \otimes e_1 \in \mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}$ are 2-orthogonal
$e_1 \otimes 1 \otimes e_1, \ldots, e_n \otimes 1 \otimes e_n \in \mathbb{C}^n \otimes \mathbb{C} \otimes \mathbb{C}^n$ are 2-orthogonal
$1 \otimes e_1 \otimes e_1, \ldots, 1 \otimes e_n \otimes e_n \in \mathbb{C} \otimes \mathbb{C}^n \otimes \mathbb{C}^n$ are 2-orthogonal

Example: Matrix Multiplication Tensor

$e_1, \ldots, e_n \in \mathbb{C}^n$ are orthogonal
$e_1 \otimes e_1, \ldots, e_n \otimes e_n \in \mathbb{C}^n \otimes \mathbb{C}^n$ are 2-orthogonal

$e_1 \otimes e_1 \otimes 1, \ldots, e_n \otimes e_n \otimes e_1 \in \mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}$ are 2-orthogonal
$e_1 \otimes 1 \otimes e_1, \ldots, e_n \otimes 1 \otimes e_n \in \mathbb{C}^n \otimes \mathbb{C} \otimes \mathbb{C}^n$ are 2-orthogonal
$1 \otimes e_1 \otimes e_1, \ldots, 1 \otimes e_n \otimes e_n \in \mathbb{C} \otimes \mathbb{C}^n \otimes \mathbb{C}^n$ are 2-orthogonal

Using vertical tensor product, we get

$$\{(e_i \otimes e_j) \otimes (e_j \otimes e_k) \otimes (e_k \otimes e_i) \mid 1 \leq i, j, k \leq n\}$$

are 2-orthogonal.

Example: Matrix Multiplication Tensor

Theorem (D.)

The matrix multiplication tensor

$$T_n = \sum_{i,j,k=1}^{n} e_{i,j} \otimes e_{j,k} \otimes e_{k,i}$$

is a DSVD.

The singular values of T_n are

$$\underbrace{1, 1, \ldots, 1}_{n^3}$$

In particular,

$$\| T_n \|_\star = \sum_{i=1}^{n^3} 1 = n^3.$$

Example: Discrete Fourier Transform

Define

$$F_n = \sum_{\substack{1 \leq i,j,k \leq n \\ i+j+k \equiv 0 \bmod n}} e_i \otimes e_j \otimes e_k$$

This tensor is related to the multiplication of univariate polynomials. Clearly $\text{rank}(F_n) \leq n^2$ and $\|F_n\|_\star \leq n^2$.

Example: Discrete Fourier Transform

Define

$$F_n = \sum_{\substack{1 \leq i,j,k \leq n \\ i+j+k \equiv 0 \bmod n}} e_i \otimes e_j \otimes e_k$$

This tensor is related to the multiplication of univariate polynomials. Clearly $\mathrm{rank}(F_n) \leq n^2$ and $\|F_n\|_\star \leq n^2$.

Discrete Fourier Transform (DFT):

$$F_n = \sum_{j=1}^{n} \sqrt{n}\left(\tfrac{1}{\sqrt{n}} \textstyle\sum_{i=1}^n \zeta^{ij} e_i\right) \otimes \left(\tfrac{1}{\sqrt{n}} \textstyle\sum_{i=1}^n \zeta^{ij} e_i\right) \otimes \left(\tfrac{1}{\sqrt{n}} \textstyle\sum_{i=1}^n \zeta^{ij} e_i\right).$$

where $\zeta = e^{\pi i/n}$. This is the *unique* DSVD of F_n. So the singular values are $\sqrt{n}, \dots, \sqrt{n}$ (n times), $\mathrm{rank}(F_n) = n$ and $\|F_n\|_\star = n\sqrt{n}$.

Generalization: Group Algebra Multiplication Tensor

G is a group with n elements and $\mathbb{C}G \cong \mathbb{C}^n$ is the group algebra

$$T_G = \sum_{g,h \in G} g \otimes h \otimes h^{-1} g^{-1}.$$

DFT case corresponds to $G = \mathbb{Z}/n\mathbb{Z}$.

Generalization: Group Algebra Multiplication Tensor

G is a group with n elements and $\mathbb{C}G \cong \mathbb{C}^n$ is the group algebra

$$T_G = \sum_{g,h \in G} g \otimes h \otimes h^{-1}g^{-1}.$$

DFT case corresponds to $G = \mathbb{Z}/n\mathbb{Z}$.

Theorem (D.)

T_G has a DSVD and its singular values are

$$\underbrace{\sqrt{\frac{n}{d_1}}, \dots, \sqrt{\frac{n}{d_1}}}_{d_1^3}, \dots, \underbrace{\sqrt{\frac{n}{d_s}}, \dots, \sqrt{\frac{n}{d_s}}}_{d_s^3}$$

where $d_1, d_2, \dots, d_s$ are the dimension of the irreducible representations of G.

P_n no DSVD

Suppose that P_n has singular values $\lambda_1, \ldots, \lambda_r$.

$$\|P_n\|_\star = n^{n/2} = \textstyle\sum_{i=1}^r \lambda_i$$

$$\|P_n\|_\sigma = \tfrac{n!}{n^{n/2}} = \lambda_1$$

$$\|P_n\|_2^2 = n! = \textstyle\sum_{i=1}^r \lambda_i^2$$

$$\lambda_1 \sum_{i=1}^r \lambda_i = n! = \sum_{i=1}^r \lambda_i^2$$

so $\lambda_1 = \cdots = \lambda_r$, and $r = \frac{r\lambda_1}{\lambda_1} = \frac{\|D_n\|_\star}{\|D_n\|_\sigma} = \frac{n^n}{n!}$.
If $n > 2$ then r is not an integer!

D_n no DSVD

Suppose that D_n has singular values $\lambda_1, \ldots, \lambda_r$. Then $\lambda_1 = \cdots = \lambda_r = 1$ and $r = n!$ (similar calculation).

$$v_1, v_2, \ldots, v_r$$

are 2-orthogonal, so $r \leq (\dim V)^{1/2} = n^{n/2}$. So

$$n! = r \leq n^{n/2}$$

Not possible for $n > 2$.

Slope decomposition

Definition

$T = T_1 + T_2 + \cdots + T_s$ is called a slope decomposition if $\langle T_i, T_j \rangle = \|T_i\|_\star \|T_j\|_\sigma$ for all $i \leq j$ and

$$\frac{\|T_1\|_\star}{\|T_1\|_\sigma} < \frac{\|T_2\|_\star}{\|T_2\|_\sigma} < \cdots < \frac{\|T_r\|_\star}{\|T_r\|_\sigma}.$$

Slope decomposition is unique if it exists. If T has a DSVD, then it also has a slope decomposition. But not every tensor has a slope decomposition.

Singular values for tensors with slope decomposition

Definition

Suppose that $T = T_1 + \cdots + T_s$ is a slope decomposition, $\mu_i = \|T_i\|_\sigma$, and $\lambda_i = \mu_i + \mu_{i+1} + \cdots + \mu_s$ for all i. Then T has singular value λ_i with multiplicity $\frac{\|T_i\|_\star}{\|T_i\|_\sigma} - \frac{\|T_{i-1}\|_\star}{\|T_{i-1}\|_\sigma}$.

Singular values for tensors with slope decomposition

Definition

Suppose that $T = T_1 + \cdots + T_s$ is a slope decomposition, $\mu_i = \|T_i\|_\sigma$, and $\lambda_i = \mu_i + \mu_{i+1} + \cdots + \mu_s$ for all i. Then T has singular value λ_i with multiplicity $\frac{\|T_i\|_\star}{\|T_i\|_\sigma} - \frac{\|T_{i-1}\|_\star}{\|T_{i-1}\|_\sigma}$.

D_n has singular value 1 with multiplicity $n!$
P_n has singular value $\frac{n!}{n^{n/2}}$ with multiplicity $\frac{n^n}{n!}$.

Singular values for tensors with slope decomposition

Definition

Suppose that $T = T_1 + \cdots + T_s$ is a slope decomposition, $\mu_i = \|T_i\|_\sigma$, and $\lambda_i = \mu_i + \mu_{i+1} + \cdots + \mu_s$ for all i. Then T has singular value λ_i with multiplicity $\frac{\|T_i\|_\star}{\|T_i\|_\sigma} - \frac{\|T_{i-1}\|_\star}{\|T_{i-1}\|_\sigma}$.

D_n has singular value 1 with multiplicity $n!$
P_n has singular value $\frac{n!}{n^{n/2}}$ with multiplicity $\frac{n^n}{n!}$.

Lemma

If T has slope decomposition and singular values $\lambda_1, \ldots, \lambda_r$ with multiplicities $m_1, \ldots, m_r$, then $\|T\|_\sigma = \lambda_1$, $\|T\|_\star = \sum_i m_i\lambda_i$ and $\|T\|_2^2 = \sum_i m_i\lambda_i^2$.

Generalizations

One can a singular spectrum for arbitrary tensors compatible with the singular spectrum for tensors with slope decomposition, but: one may get continuous spectra or negative multiplicities.

Generalizations

One can a singular spectrum for arbitrary tensors compatible with the singular spectrum for tensors with slope decomposition, but: one may get continuous spectra or negative multiplicities.

One can define a slope decomposition and singular values whenever we have a vector space with two norms dual to each other.

Generalizations

One can a singular spectrum for arbitrary tensors compatible with the singular spectrum for tensors with slope decomposition, but: one may get continuous spectra or negative multiplicities.

One can define a slope decomposition and singular values whenever we have a vector space with two norms dual to each other.

Thank You!