Algebraic Methods for Tensor Data

Neriman Tokcan (Broad Institute, MIT/Harvard)
Harm Derksen (Dept. of Math., Northeastern University)
Jonathan Gryak (BCIL lab, Univ. of Michigan)
Kayvan Najarian (BCIL lab, Univ. of Michigan)

August 25, 2020

Tensors

$V_i = \mathbb{R}^{p_i}$
$V = V_1 \otimes V_2 \otimes \cdots \otimes V_d = \mathbb{R}^{p_1 \times p_2 \times \cdots \times p_d}$
$\mathcal{T} \in V$ tensor

Tensors

$V_i = \mathbb{R}^{p_i}$
$V = V_1 \otimes V_2 \otimes \cdots \otimes V_d = \mathbb{R}^{p_1 \times p_2 \times \cdots \times p_d}$
$\mathcal{T} \in V$ tensor

tensors appear in many data applications
e.g., chemometrics, psychometrics, imaging, algebraic complexity theory, signal processing, neural networks, . . .

Tensors

$V_i = \mathbb{R}^{p_i}$
$V = V_1 \otimes V_2 \otimes \cdots \otimes V_d = \mathbb{R}^{p_1 \times p_2 \times \cdots \times p_d}$
$\mathcal{T} \in V$ tensor

tensors appear in many data applications
e.g., chemometrics, psychometrics, imaging, algebraic complexity theory, signal processing, neural networks, ...

$\mathcal{T} \cdot \mathcal{S}$ inner product of tensors
$\|\mathcal{T}\| = |\mathcal{T}\|_F = \sqrt{\mathcal{T} \cdot \mathcal{T}}$ Euclidean (Frobenius) norm

Tensor Rank

Definition

rank($\mathcal{T}$) is the smallest integer r for which we have a decomposition

$$(\star) \qquad \mathcal{T} = \sum_{j=1}^{r} v_{1j} \otimes v_{2j} \otimes \cdots \otimes v_{dj}$$

Tensor Rank

Definition

rank($\mathcal{T}$) is the smallest integer r for which we have a decomposition

$$(\star) \qquad \mathcal{T} = \sum_{j=1}^{r} v_{1j} \otimes v_{2j} \otimes \cdots \otimes v_{dj}$$

if r is minimal then $(\star)$ is called the CP decomposition

Tensor Rank

Definition

rank($\mathcal{T}$) is the smallest integer r for which we have a decomposition

$$(\star) \qquad \mathcal{T} = \sum_{j=1}^{r} v_{1j} \otimes v_{2j} \otimes \cdots \otimes v_{dj}$$

if r is minimal then $(\star)$ is called the CP decomposition

tensor rank is not continuous or even semi-continuous, i.e., $\{\mathcal{T} \in V \mid \text{rank}(\mathcal{T}) \leq r\}$ not always a closed set

Nuclear Norm

use convex relaxation of tensor rank:

Definition

the nuclear norm $\|\mathcal{T}\|_\star$ of $\mathcal{T}$ is the minimal value of

$$(\#) \qquad \sum_{j=1}^{r} \|v_{i1} \otimes v_{i2} \otimes \cdots \otimes v_{id}\|$$

where r is arbitrary and

$$(\star) \qquad \mathcal{T} = \sum_{j=1}^{r} v_{1j} \otimes v_{2j} \otimes \cdots \otimes v_{dj}$$

(a minimum is attained)

Nuclear Norm

use convex relaxation of tensor rank:

Definition

the nuclear norm $\|\mathcal{T}\|_\star$ of $\mathcal{T}$ is the minimal value of

$$(\#) \qquad \sum_{j=1}^{r} \|v_{i1} \otimes v_{i2} \otimes \cdots \otimes v_{id}\|$$

where r is arbitrary and

$$(\star) \qquad \mathcal{T} = \sum_{j=1}^{r} v_{1j} \otimes v_{2j} \otimes \cdots \otimes v_{dj}$$

(a minimum is attained)

if $\|\mathcal{T}\|_\star$ is equal to $(\#)$, then $(\star)$ is called the nuclear decomposition or convex decomposition

Spectral Norm

Definition

the spectral norm of $\mathcal{T}$ is

$$\|\mathcal{T}\|_\sigma = \max\{\mathcal{T}\cdot(v_1 \otimes v_2 \otimes \cdots \otimes v_d) \mid \|v_1\| = \|v_2\| = \cdots = \|v_d\| = 1\}$$

(a maximum is attained)

Spectral Norm

Definition

the spectral norm of $\mathcal{T}$ is

$$\|\mathcal{T}\|_\sigma = \max\{\mathcal{T}\cdot(v_1\otimes v_2\otimes\cdots\otimes v_d) \mid \|v_1\| = \|v_2\| = \cdots = \|v_d\| = 1\}$$

(a maximum is attained)

$\|\cdot\|_\sigma$ and $\|\cdot\|_\star$ are dual norms

$$|\mathcal{T}\cdot\mathcal{S}| \leq \|\mathcal{T}\|_\star\|\mathcal{S}\|_\sigma$$

Matrices ($d = 2$)

$A \in \mathbb{R}^{p \times q}$ has a singular value decomposition
$A = UDV^t$ with $U \in \mathrm{O}_p$, $V \in \mathrm{O}_q$, D diagonal with the nonzero diagonal entries $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$

Matrices ($d = 2$)

$A \in \mathbb{R}^{p \times q}$ has a singular value decomposition
$A = UDV^t$ with $U \in \mathrm{O}_p$, $V \in \mathrm{O}_q$, D diagonal with the nonzero diagonal entries $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$

$r = \mathrm{rank}(A)$ is the matrix rank and the tensor rank of A

Matrices ($d = 2$)

$A \in \mathbb{R}^{p\times q}$ has a singular value decomposition
$A = UDV^t$ with $U \in \mathrm{O}_p$, $V \in \mathrm{O}_q$, D diagonal with the nonzero diagonal entries $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$

$r = \mathrm{rank}(A)$ is the matrix rank and the tensor rank of A

$\|A\|_\star = \mathrm{Tr}(\sqrt{AA^t}) = \lambda_1 + \lambda_2 + \cdots + \lambda_r$

Matrices ($d = 2$)

$A \in \mathbb{R}^{p \times q}$ has a singular value decomposition
$A = UDV^t$ with $U \in \mathrm{O}_p$, $V \in \mathrm{O}_q$, D diagonal with the nonzero diagonal entries $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$

$r = \mathrm{rank}(A)$ is the matrix rank and the tensor rank of A

$\|A\|_\star = \mathrm{Tr}(\sqrt{AA^t}) = \lambda_1 + \lambda_2 + \cdots + \lambda_r$

$\|A\|_\sigma = \lambda_1$ is operator norm

Matrices ($d = 2$)

$A \in \mathbb{R}^{p \times q}$ has a singular value decomposition
$A = UDV^t$ with $U \in \mathrm{O}_p$, $V \in \mathrm{O}_q$, D diagonal with the nonzero diagonal entries $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$

$r = \mathrm{rank}(A)$ is the matrix rank and the tensor rank of A

$\|A\|_\star = \mathrm{Tr}(\sqrt{AA^t}) = \lambda_1 + \lambda_2 + \cdots + \lambda_r$

$\|A\|_\sigma = \lambda_1$ is operator norm

the norms and the rank of A are easy to compute

negative results ($d = 3$)

Theorem (Håstad)

computing tensor rank is NP-complete

negative results ($d = 3$)

Theorem (Håstad)

computing tensor rank is NP-complete

Theorem (Friedland–Lim)

computing the nuclear norm is NP-complete

negative results ($d = 3$)

Theorem (Håstad)

computing tensor rank is NP-complete

Theorem (Friedland–Lim)

computing the nuclear norm is NP-complete

Theorem (Hillar–Lim)

computing the spectral norm is NP-complete

Approximating Spectral Norm

(case $d = 3$ for simplicity)

$$\|\mathcal{T}\|_{\overline{\sigma},d} = \left(\int_{S^{p-1} \times S^{q-1} \times S^{r-1}} |\mathcal{T} \cdot (x \otimes y \otimes z)|^d \right)^{\frac{1}{d}} .$$

normalize

$$\|\mathcal{T}\|_{\sigma,d} = \frac{\|\mathcal{T}\|_{\overline{\sigma},d}}{\|e_1 \otimes e_1 \otimes e_1\|_{\overline{\sigma},d}}$$

Approximating Spectral Norm

(case $d = 3$ for simplicity)

$$\|\mathcal{T}\|_{\overline{\sigma},d} = \left(\int_{S^{p-1}\times S^{q-1}\times S^{r-1}} |\mathcal{T}\cdot(x\otimes y\otimes z)|^d \right)^{\frac{1}{d}}.$$

normalize

$$\|\mathcal{T}\|_{\sigma,d} = \frac{\|\mathcal{T}\|_{\overline{\sigma},d}}{\|e_1\otimes e_1\otimes e_1\|_{\overline{\sigma},d}}$$

$$\|\mathcal{T}\|_{\sigma} = \lim_{d\to\infty} \|\mathcal{T}\|_{\sigma,d}.$$

Approximating Spectral Norm

(case $d = 3$ for simplicity)

$$\|\mathcal{T}\|_{\overline{\sigma},d} = \left(\int_{S^{p-1}\times S^{q-1}\times S^{r-1}} |\mathcal{T}\cdot(x\otimes y\otimes z)|^d\right)^{\frac{1}{d}}.$$

normalize

$$\|\mathcal{T}\|_{\sigma,d} = \frac{\|\mathcal{T}\|_{\overline{\sigma},d}}{\|e_1\otimes e_1\otimes e_1\|_{\overline{\sigma},d}}$$

$$\|\mathcal{T}\|_{\sigma} = \lim_{d\to\infty} \|\mathcal{T}\|_{\sigma,d}.$$

when d is even there is an algebraic method for computing $\|\mathcal{T}\|_{\sigma,d}$!

Invariant Tensors

$$V = \mathbb{R}^n, \ V^{\otimes d} = \underbrace{V \otimes V \otimes \cdots \otimes V}_{d}$$

Invariant Tensors

$$V = \mathbb{R}^n,\ V^{\otimes d} = \underbrace{V \otimes V \otimes \cdots \otimes V}_{d}$$

O_n acts on V and $V^{\otimes d}$

Invariant Tensors

$$V = \mathbb{R}^n,\ V^{\otimes d} = \underbrace{V \otimes V \otimes \cdots \otimes V}_{d}$$

O_n acts on V and $V^{\otimes d}$

there is an O_n-equivariant linear isomorphism between $V^{\otimes d}$ and the space of multilinear maps $V^d \to \mathbb{R}$

Invariant Tensors

$$V = \mathbb{R}^n,\ V^{\otimes d} = \underbrace{V \otimes V \otimes \cdots \otimes V}_{d}$$

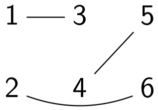
O_n acts on V and $V^{\otimes d}$

there is an O_n-equivariant linear isomorphism between $V^{\otimes d}$ and the space of multilinear maps $V^d \to \mathbb{R}$

so there is a linear isomorphism between the space $(V^{\otimes d})^{\mathrm{O}_n}$ of O_n-invariant tensors and the space of O_n-invariant multilinear maps $V^d \to \mathbb{R}$

Brauer Diagrams

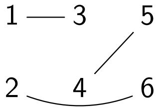
a Brauer diagram is a perfect matching, for example

D = 

Brauer Diagrams

a Brauer diagram is a perfect matching, for example

D =



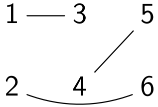
to a diagram E on d vertices we can associate an O_n-invariant multilinear map $\mathcal{M}_E : V^d \to \mathbb{R}$, for example

$$\mathcal{M}_D(v_1, v_2, \ldots, v_6) = (v_1 \cdot v_3)(v_2 \cdot v_6)(v_4 \cdot v_5)$$

Brauer Diagrams

a Brauer diagram is a perfect matching, for example

D =



to a diagram E on d vertices we can associate an O_n-invariant multilinear map $\mathcal{M}_{\mathrm{E}} : V^d \to \mathbb{R}$, for example

$$\mathcal{M}_{\mathrm{D}}(v_1, v_2, \ldots, v_6) = (v_1 \cdot v_3)(v_2 \cdot v_6)(v_4 \cdot v_5)$$

the invariant multi-linear map $\mathcal{M}_{\mathrm{E}}$ corresponds to an invariant tensor $\mathcal{T}_{\mathrm{E}}$ using the linear isomorphism of before, for example

$$\mathcal{T}_{\mathrm{D}} = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} e_i \otimes e_j \otimes e_i \otimes e_k \otimes e_k \otimes e_j$$

Brauer Diagrams

Theorem (FFT of Invariant Theory for O_n)

the space $(V^{\otimes d})^{\mathrm{O}_n}$ is spanned by all $\mathcal{T}_{\mathrm{D}}$ where D runs over all Brauer diagrams on d vertices

Brauer Diagrams

Theorem (FFT of Invariant Theory for O_n)

the space $(V^{\otimes d})^{\mathrm{O}_n}$ is spanned by all $\mathcal{T}_{\mathrm{D}}$ where D runs over all Brauer diagrams on d vertices

the tensors $\mathcal{T}_{\mathrm{D}}$ are independent if $n \geq d$

Brauer Diagrams

Theorem (FFT of Invariant Theory for O_n)

the space $(V^{\otimes d})^{\mathrm{O}_n}$ is spanned by all $\mathcal{T}_{\mathrm{D}}$ where D runs over all Brauer diagrams on d vertices

the tensors $\mathcal{T}_{\mathrm{D}}$ are independent if $n \geq d$

the number of Brauer diagrams on d vertices is $1 \cdot 3 \cdot 5 \cdots d - 1$ (when d even)

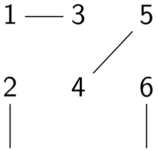
Partial Brauer Diagrams

to a partial Brauer diagram with d vertices and e edges we associate an $\mathrm{O}(V)$-equivariant multi-linear map $\mathcal{M}_\mathrm{D} : V^d \to V^{\otimes(d-2e)}$ and a linear map $\mathcal{L}_\mathrm{D} : V^{\otimes d} \to V^{\otimes d-2e}$

Partial Brauer Diagrams

to a partial Brauer diagram with d vertices and e edges we associate an $\mathrm{O}(V)$-equivariant multi-linear map $\mathcal{M}_{\mathrm{D}} : V^d \to V^{\otimes(d-2e)}$ and a linear map $\mathcal{L}_{\mathrm{D}} : V^{\otimes d} \to V^{\otimes d-2e}$

for example, if

D = 

then

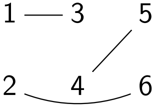
$$\mathcal{M}_{\mathrm{D}}(v_1, v_2, \ldots, v_6) = (v_1 \cdot v_3)(v_4 \cdot v_5) v_2 \otimes v_6 \in V^{\otimes 2}$$

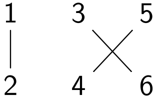
the inner product of two tensors $\mathcal{T}_{\mathsf{D}}$ and $\mathcal{T}_{\mathsf{E}}$ is n^c, were c is the number of cycles if we overlay the diagrams D and E

Brauer Diagram Calculus

the inner product of two tensors $\mathcal{T}_{\mathsf{D}}$ and $\mathcal{T}_{\mathsf{E}}$ is n^c, were c is the number of cycles if we overlay the diagrams D and E

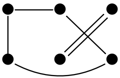
for example, if

D = 

E = 

then

$$\mathcal{T}_{\mathsf{D}} \cdot \mathcal{T}_{\mathsf{E}} = n^2 \qquad \text{because}$$



has 2 cycles

Integrating

let

$$\mathcal{S}_d = \sum_{\mathsf{D}} \mathcal{T}_{\mathsf{D}}$$

where D runs over all Brauer diagrams on d vertices

Integrating

let

$$\mathcal{S}_d = \sum_{\mathrm{D}} \mathcal{T}_{\mathrm{D}}$$

where D runs over all Brauer diagrams on d vertices

if we integrate $v^{\otimes d} \in V^{\otimes d}$ over the unit sphere $S^{n-1} \subset V$ (where S^{n-1} has measure 1) then we get a tensor that is invariant under O_n and the symmetric group Σ_d, so

$$\int_{S^{n-1}} v^{\otimes d} = C\mathcal{S}_d$$

for some constant C

Integrating

let

$$\mathcal{S}_d = \sum_{\mathrm{D}} \mathcal{T}_{\mathrm{D}}$$

where D runs over all Brauer diagrams on d vertices

if we integrate $v^{\otimes d} \in V^{\otimes d}$ over the unit sphere $S^{n-1} \subset V$ (where S^{n-1} has measure 1) then we get a tensor that is invariant under O_n and the symmetric group Σ_d, so

$$\int_{S^{n-1}} v^{\otimes d} = C\mathcal{S}_d$$

for some constant C

from

$$1 = \int_{S^{n-1}} 1 = \int_{S^{n-1}} v^{\otimes d} \cdot \mathcal{T}_{\mathrm{E}} = C\mathcal{S}_d \cdot \mathcal{T}_{\mathrm{E}} = Cn(n+2)\cdots(n+d-2)$$

follows $C = \frac{1}{n(n+2)\cdots(n+d-2)}$

Colored Brauer Diagrams

$R = \mathbb{R}^p$, $G = \mathbb{R}^q$, $B = \mathbb{R}^r$
$V = R \otimes G \otimes B$ space of order 3 tensors
$\psi : R^{\otimes d} \otimes G^{\otimes d} \otimes B^{\otimes d} \to V^{\otimes d}$ linear isomorphism
$H = O_p \times O_q \times O_r$ acts on V, $V^{\otimes d}$

Colored Brauer Diagrams

$R = \mathbb{R}^p$, $G = \mathbb{R}^q$, $B = \mathbb{R}^r$
$V = R \otimes G \otimes B$ space of order 3 tensors
$\psi : R^{\otimes d} \otimes G^{\otimes d} \otimes B^{\otimes d} \to V^{\otimes d}$ linear isomorphism
$H = \mathrm{O}_p \times \mathrm{O}_q \times \mathrm{O}_r$ acts on V, $V^{\otimes d}$

isomorphism $\psi : (R^{\otimes d})^{\mathrm{O}_p} \otimes (G^{\otimes d})^{\mathrm{O}_q} \otimes (B^{\otimes d})^{\mathrm{O}_q} \to (V^{\otimes d})^H$

Colored Brauer Diagrams

$R = \mathbb{R}^p$, $G = \mathbb{R}^q$, $B = \mathbb{R}^r$
$V = R \otimes G \otimes B$ space of order 3 tensors
$\psi : R^{\otimes d} \otimes G^{\otimes d} \otimes B^{\otimes d} \to V^{\otimes d}$ linear isomorphism
$H = \mathrm{O}_p \times \mathrm{O}_q \times \mathrm{O}_r$ acts on V, $V^{\otimes d}$

isomorphism $\psi : (R^{\otimes d})^{\mathrm{O}_p} \otimes (G^{\otimes d})^{\mathrm{O}_q} \otimes (B^{\otimes d})^{\mathrm{O}_q} \to (V^{\otimes d})^H$

$(V^{\otimes d})^H$ spanned by all $\psi(\mathcal{T}_{\mathrm{D}_R} \otimes \mathcal{T}_{\mathrm{D}_G} \otimes \mathcal{T}_{\mathrm{D}_B})$ where $\mathrm{D}_R, \mathrm{D}_G, \mathrm{D}_B$ are Brauer diagrams on d vertices

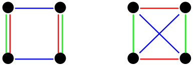
Colored Brauer Diagrams

$R = \mathbb{R}^p,\ G = \mathbb{R}^q,\ B = \mathbb{R}^r$
$V = R \otimes G \otimes B$ space of order 3 tensors
$\psi : R^{\otimes d} \otimes G^{\otimes d} \otimes B^{\otimes d} \to V^{\otimes d}$ linear isomorphism
$H = \mathrm{O}_p \times \mathrm{O}_q \times \mathrm{O}_r$ acts on $V,\ V^{\otimes d}$

isomorphism $\psi : (R^{\otimes d})^{\mathrm{O}_p} \otimes (G^{\otimes d})^{\mathrm{O}_q} \otimes (B^{\otimes d})^{\mathrm{O}_q} \to (V^{\otimes d})^H$

$(V^{\otimes d})^H$ spanned by all $\psi(\mathcal{T}_{\mathrm{D}_R} \otimes \mathcal{T}_{\mathrm{D}_G} \otimes \mathcal{T}_{\mathrm{D}_B})$ where $\mathrm{D}_R, \mathrm{D}_G, \mathrm{D}_B$ are Brauer diagrams on d vertices

we draw the Brauer diagrams D_R in red, D_G in green, and D_B in blue using the same d vertices and we get a *colored Brauer diagram* D, for example



Tensor Invariants

$V = R \otimes G \otimes B$, space of $p \times q \times r$ tensors

$(V^{\otimes d})^H$ is spanned by all $\mathcal{T}_{\mathrm{D}} := \psi(\mathcal{T}_{\mathrm{D}_R} \otimes \mathcal{T}_{\mathrm{D}_G} \otimes \mathcal{T}_{\mathrm{D}_B})$ where D runs over all colored Brauer diagrams

Tensor Invariants

$V = R \otimes G \otimes B$, space of $p \times q \times r$ tensors

$(V^{\otimes d})^H$ is spanned by all $\mathcal{T}_{\mathrm{D}} := \psi(\mathcal{T}_{\mathrm{D}_R} \otimes \mathcal{T}_{\mathrm{D}_G} \otimes \mathcal{T}_{\mathrm{D}_B})$ where D runs over all colored Brauer diagrams

for a colored Brauer diagram, define a polynomial function $\mathcal{I}_{\mathcal{D}} : V \to \mathbb{R}$ by $\mathcal{I}_{\mathcal{D}}(\mathcal{U}) = \mathcal{U}^{\otimes d} \cdot \mathcal{T}_{\mathrm{D}}$

Theorem

the ring of H-invariant polynomials on the space of $p \times q \times r$ tensors is spanned by all $\mathcal{I}_{\mathrm{D}}$, where D is a colored Brauer diagram

Computing the Integral

Up to a constant

$$\int_{S^{p-1}\times S^{q-1}\times S^{r-1}} (a\otimes b\otimes c)^{\otimes d}$$

is equal to $\psi(\mathcal{S}_d \otimes \mathcal{S}_d \otimes \mathcal{S}_d)$, the sum of all $\mathcal{T}_{\mathrm{D}}$ where D is a colored Brauer diagram on d vertices

Computing the Norm

$$\|\mathcal{U}\|_{\sigma,4}^4 = \int (\mathcal{U} \cdot a \otimes b \otimes c)^d = \int \mathcal{U}^{\otimes d} \cdot (a \otimes b \otimes c)^{\otimes d} =$$
$$= \mathcal{U}^{\otimes d} \cdot \left(\int (a \otimes b \otimes c)^{\otimes d} \right)$$

is up to scalar equal to the sum of all $\mathcal{I}_{\mathrm{D}}$ where D is a colored diagram on d vertices

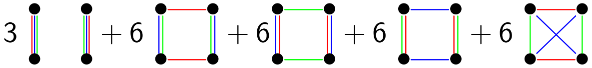
Computing the Norm

$$\|\mathcal{U}\|_{\sigma,4}^4 = \int (\mathcal{U} \cdot a \otimes b \otimes c)^d = \int \mathcal{U}^{\otimes d} \cdot (a \otimes b \otimes c)^{\otimes d} =$$

$$= \mathcal{U}^{\otimes d} \cdot \left(\int (a \otimes b \otimes c)^{\otimes d} \right)$$

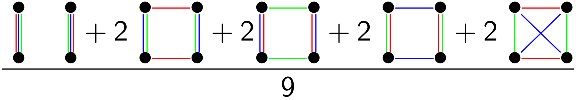
is up to scalar equal to the sum of all $\mathcal{I}_{\mathrm{D}}$ where D is a colored diagram on d vertices

note that $\mathcal{I}_{\mathrm{D}}$ does not depend on the labeling of the vertices, so up to a scalar, $\|\mathcal{U}\|_{\sigma,4}^4$ is equal to



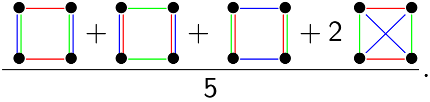
(by abuse of notation, the diagram D represents $\mathcal{I}_{\mathrm{D}}$)

$$\|\mathcal{U}\|_{\sigma,4}^4 =$$



a similar norm that in some ways is a better approximation of the spectral norm is given by

$$\|\mathcal{U}\|_{\#}^4 =$$



Tensor Amplification

if matrix A has singular values $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r$
then AA^tA has singular values $\lambda_1^3, \lambda_2^3, \ldots$
so the map $A \mapsto AA^tA$ enhances the low rank structure (largest singular values) and surpresses noise (small singular values)

Tensor Amplification

if matrix A has singular values $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r$
then AA^tA has singular values $\lambda_1^3, \lambda_2^3, \ldots$
so the map $A \mapsto AA^tA$ enhances the low rank structure (largest singular values) and surpresses noise (small singular values)

AA^tA is the gradient of $\frac{1}{4}\operatorname{Tr}((AA^t)^2)$

Tensor Amplification

if matrix A has singular values $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r$
then AA^tA has singular values $\lambda_1^3, \lambda_2^3, \ldots$
so the map $A \mapsto AA^tA$ enhances the low rank structure (largest singular values) and surpresses noise (small singular values)

AA^tA is the gradient of $\frac{1}{4}\operatorname{Tr}((AA^t)^2)$

we can do something similar with tensors and define

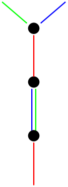
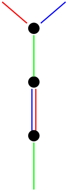
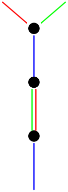
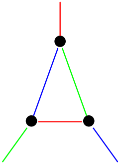
$$\Phi_{\sigma,4} = \nabla \| T \|_{\sigma,4}^4$$

$$\Phi_{\#} = \nabla \| T \|_{\#}^4$$

these maps are $\mathrm{O}_p \times \mathrm{O}_q \times \mathrm{O}_r$-equivariant and amplify the low rank structure of a tensor

Partial Colored Brauer Diagrams

in partial colored Brauer diagrams, we have

$\Phi_{\#} = \frac{4}{5}$  $+ \frac{4}{5}$  $+ \frac{4}{5}$  $+ \frac{8}{5}$ 

Numerical Experiments

choose a random tensor $30 \times 30 \times 30$ tensor $\mathcal{T}_n = \mathcal{T} + \mathcal{E}$
$\mathcal{T}$ is random rank 1 with $\|\mathcal{T}\| = 1$
$\mathcal{E}$ is random tensor with $\|\mathcal{E}\| = 10$
signal to noise ratio $-20dB$

experiment: use Alternating Least Squares (ALS) algorithm for best rank 1 approximation of $\mathcal{T}_n$ (to recover $\mathcal{T}$)

use ALS with random initial guess
compare with using ALS with amplified initial guess

Results

Random (10 runs)	Max Fit	Total # Iterations	Total Time
Average	0.7136	77.5080	0.0943
Standard Deviation	0.2715	12.0254	0.0159
Quick Rank 1	Fit	# Iterations	Time
Average	0.7848	2.94	0.0177
Standard Deviation	0.1618	1.2345	0.0025
$\Phi_{\sigma,4}$ **and Quick Rank 1**	Fit	# Iterations	Time
Average	0.8010	2	0.0210
Standard Deviation	0.1256	0	0.0027
$\Phi_{\#}$ **and Quick Rank 1**	Fit	# Iterations	Time
Average	0.8178	2	0.0205
Standard Deviation	0.0515	0	0.0025