Tensor networks, TT (Matrix Product States) and Hierarchical Tucker decomposition

R. Schneider (TUB Matheon)

John von Neumann Lecture – TU Munich, 2012

Setting - Tensors

$$V_\nu := \mathbb{R}^n \ , \ \ \mathcal{H}_d = \mathcal{H} := \bigotimes_{\nu=1}^{d} V_\nu \ \ d\text{-fold tensor product Hilbert-s.},$$

$$\mathcal{H} \simeq \{(x_1, \ldots, x_d) \mapsto U(x_1, \ldots, x_d) \in \mathbb{R} : x_i = 1, \ldots, n_i\} \ .$$

The function $U \in \mathcal{H}$ will be called an order d-tensor.
For notational simplicity, we often consider $n_i = n$. Here $x_1, \ldots, x_d \in \{1, \ldots, n\}$ will be called variables or indices.

$$\mathbf{k} \mapsto U(k_1, \ldots, k_d) = (U_{k_1, \ldots, k_d}) \ \ , \ \ k_i = 1 \ldots, n_i \ .$$

Or in index (vectorial) notation

$$\mathbf{U} = \left(U_{k_1, \ldots, k_d}\right)_{k_i=0, 1 \leq i \leq d}^{n_i}$$

$$\dim \mathcal{H} = n^d \ \ \text{curse of dimensions!!!}$$

E.g. wave function $\Psi(\mathbf{r}_1, s_1, \ldots, \mathbf{r}_N, s_N)$

Contracted tensor representations

Espig & Hackbusch & Handschuh & S.
Let $V_\nu := \mathbb{R}^{r_\nu}$

$$\mathcal{K}_K = \mathcal{K} := \bigotimes_{\nu=1}^{K} V_\nu \quad d\text{-fold tensor product space,}$$

$$\mathcal{K} \simeq \{(k_1, \dots, k_K) \mapsto V(k_1, \dots, k_K) : k_i = 1, \dots, r_i\} .$$

Here $k_1, \dots, k_N$ will be called contraction variables,
$x_1, \dots, x_d$ will be called exterior variables.
We define an order $d + K$ tensor product space

$$\widetilde{\mathcal{H}} := \mathcal{H} \otimes \mathcal{K} \quad \simeq \quad \{(x_1, \dots x_d, k_1, \dots, k_K) \mapsto \widetilde{U}(x_1, \dots, x_d, k_1, \dots, k_K) : x_i = 1, \dots, n; 1 \leq i \leq d, k_j = 1, \dots, r_j, 1 \leq j \leq K\} .$$

Contracted tensor representations

Definition

Subset $\mathcal{V}^r \subseteq \mathcal{H}$: Let $d < d' \leq d + K$,

$$U \in \mathcal{V}^r \quad \Leftrightarrow U = \sum_{k_1}^{r_1} \dots \sum_{k_K}^{r_K} \bigotimes_{\nu=1}^{d'} u_\nu(x_\nu, k_{\nu(1)}, k_{\nu(2)}, k_{\nu(3)})$$

where $k_{\nu(i)} \in \{k_1, \dots, k_K\}$, $i = 1, 2, 3$, and

$$\begin{array}{rl} u_\nu(x_\nu, k_{\nu(1)}, k_{\nu(2)}) & \text{if} \quad \nu = 1, \dots, d \\ u_\nu(k_{\nu(1)}, k_{\nu(2)}, k_{\nu(3)}) & \text{if} \quad \nu = d+1, \dots, d' \ , \end{array}$$

I.e. the d' tensors $u_\nu(.,.,.)$ depends at most on 3 variables! We will write shortly $u_\nu(x_\nu, \mathbf{k}_\nu) \in \mathbb{R}$, keeping in mind that u_ν depends only on at most three variables $x_\nu, k_{\nu(1)}, k_{\nu(2)}, k_{\nu(3)}$.

Example

Canonical format: $K = 1$, $d' = d$, $u_\nu(x_\nu, k)$, $k = 1, \dots, r$.

Contracted tensor representations

Definition

Subset $\mathcal{V}^r \subseteq \mathcal{H}$: Let $d < d' \leq d + K$,

$$U \in \mathcal{V}^r \quad \Leftrightarrow U = \sum_{k_1}^{r_1} \dots \sum_{k_K}^{r_K} \bigotimes_{\nu=1}^{d'} u_\nu(x_\nu, k_{\nu(1)}, k_{\nu(2)}, (k_{\nu(3)}))$$

the d' tensors $u_\nu(.,.,.)$ depends at most on 3 variables! (Not necessary!)

Complexity:

Let $r := \max\{r_i : i = 1, \dots, K\}$ then storage requirement is

$$\boxed{\sharp DOF's \ \leq \max\{d' \times n \times r^2 \ , \ d' \times r^3\} \ .}$$

Contracted tensor representations

- Due to the multi-linear ansatz, the present extremely general format allow differentiation and local optimization methods! (See e.g. previous talk!)
- Redundancy has not been removed, in general redundancy is enlarged.
- Closedness would be violated in general.

Tensor networks

Definition

A minimal tensor network is a representation in the above format, where each contraction variable k_i, $i = 1, \ldots, K$, appear exactly twice. In this case, they can be labeled by $i \sim (\nu, \mu)$, $1 \leq \nu \leq d'$, $1 \leq \mu < \nu$. It is often assumed, that each exterior variable x_i appear at most once!

Tensor networks

Diagram - (Graph) of a tensor network:

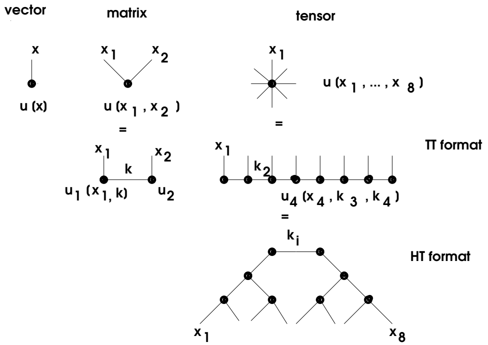


Diagramm:

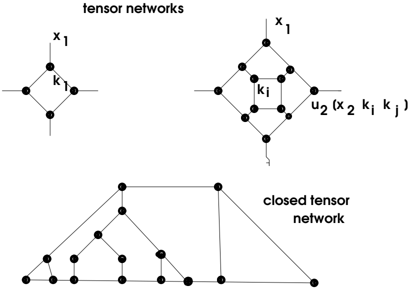
- Each node corresponds to a factor u_ν, $\nu = 1, \ldots, d'$,
- each line to a variable x_ν or a contraction variable k_μ.
- Since k_ν is an edge connecting 2 tensors u_{μ_1} and u_{μ_2}, one has to sum over connecting lines (edges).

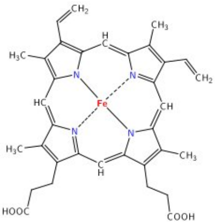
Tensor networks has been introduced in quantum information theory.

Example

1. TT (tensor trains, Oseledets & Tyrtyshnikov) and HT (Hackbusch & Kühn) formats are tensor networks.
2. the canonical format is not (directly) a tensor network.

tensor networks





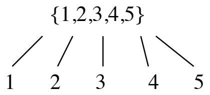
Tensor formats

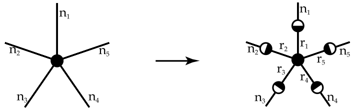
- Hierarchical Tucker format
  (HT; Hackbusch/Kühn, Grasedyck, Kressner : Tree-tensor networks)

Tensor formats

- ▷ Hierarchical Tucker format
  (HT; Hackbusch/Kühn, Grasedyck, Kressner : Tree-tensor networks)
- ▷ Tucker format (Q: MCTDH(F))

$$U(x_1,..,x_d) = \sum_{k_1=1}^{r_1} \dots \sum_{k_d=1}^{r_d} B(k_1,..,k_d) \bigotimes_{i=1}^{d} U_i(k_i,x_i)$$

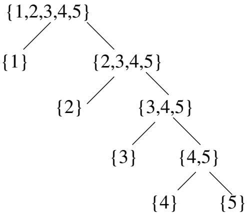


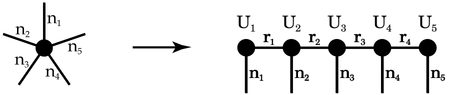


Tensor formats

- ▷ Hierarchical Tucker format
  (HT; Hackbusch/Kühn, Grasedyck, Kressner : Tree-tensor networks)
- ▷ Tucker format (Q: MCTDH(F))
- ▷ Tensor Train (TT-)format
  (Oseledets/Tyrtyshnikov, $\simeq$ MPS-format of quantum physics)

$$U(\underline{x}) = \sum_{k_1=1}^{r_1} \cdots \sum_{k_{d-1}=1}^{r_{d-1}} \prod_{i=1}^{d} B_i(k_{i-1}, x_i, k_i) = \mathbf{B}_1(x_1) \cdots \mathbf{B}_d(x_d)$$





Noteable special case of HT:

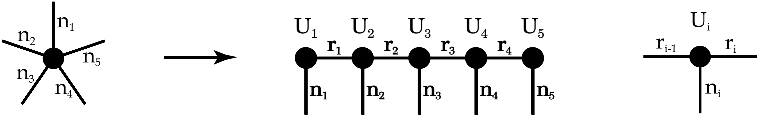
TT format (Oseledets & Tyrtyshnikov, 2009)
(matrix product states (MPS), Vidal 2003, Schöllwock et al.)

TT tensor U can be written as matrix product form

$$U(\mathbf{x}) = \mathbf{U}_1(x_1) \cdots \mathbf{U}_i(x_i) \cdots \mathbf{U}_d(x_d)$$

$$= \sum_{k_1=1}^{r_1} .. \sum_{k_{d-1}=1}^{r_{d-1}} U_1(x_1, k_1) U_2(k_1, x_2, k_2) ... U_{d-1}(k_{d-2} x_{d-1}, k_{d-1}) U_d(k_{d-1}, x_d, k_d)$$

with matrices $\mathbf{U}_i(x_i) = \left(u_{k_{i-1}}^{k_i}(x_i)\right) \in \mathbb{R}^{r_{i-1} \times r_i}$, $r_0 = r_d := 1$



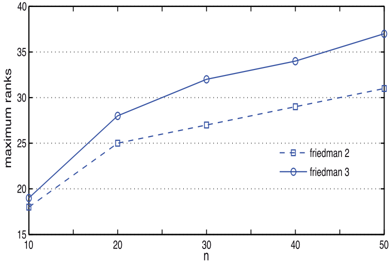
Redundancy: $U(\mathbf{x}) = \mathbf{U}_1(x_1) \mathbf{G}\mathbf{G}^{-1} \mathbf{U}_2(x_2) \cdots \mathbf{U}_i(x_i) \cdots \mathbf{U}_d(x_d)$.

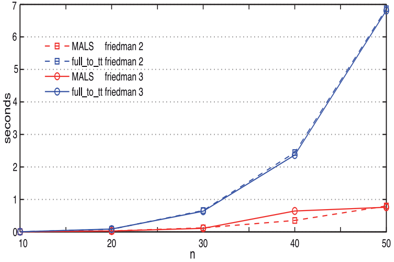
TT approximations of Friedman data sets

$$f_2(x_1, x_2, x_3, x_4) = \sqrt{(x_1^2 + (x_2 x_3 - \frac{1}{x_2 x_4})^2},$$

$$f_3(x_1, x_2, x_3, x_4) = \tan^{-1}\left(\frac{x_2 x_3 - (x_2 x_4)^{-1}}{x_1}\right)$$

on $4 - D$ grid, n points per dim. $\rightsquigarrow n^4$ tensor, $n \in \{3, \dots, 50\}$.



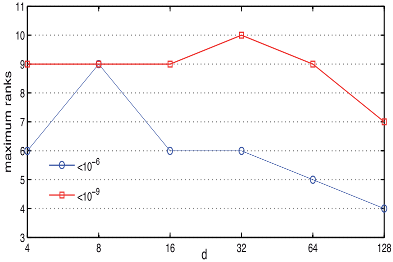


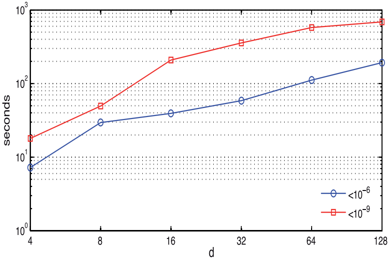
`full_to_tt` (Oseledets, successive SVDs)

and `MALS` (with A = I) (Holtz & Rohwedder & S.)

Solution of $-\Delta U = b$

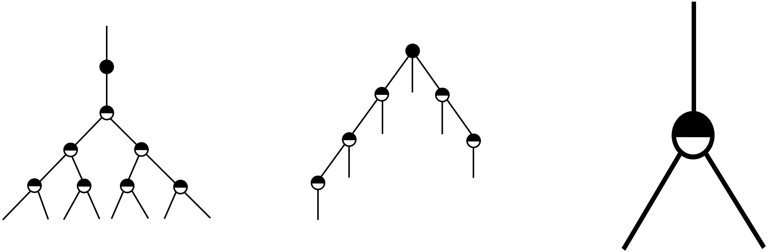
- Dimension $d = 4, \ldots, 128$ varying
- Gridsize $n = 10$
- Right-hand-side b of rank 1
- Solution U has rank 13





Hierarchical Tucker as subspace approximation

1) $D = \{1, \ldots, d\}$, tensor space $\mathbf{V} = \bigotimes_{j \in D} V_j$
2) T_D dimension partition tree, vertices $\alpha \in T_D$ are subsets $\alpha \subset T_D$, root: $\alpha = D$
3) $\mathbf{V}_\alpha = \bigotimes_{j \in \alpha} V_j$ for $\alpha \in T_D$



4) $\mathbf{U}_\alpha \subset \mathbf{V}_\alpha$ subspaces of dimension r_α with the characteristic nesting

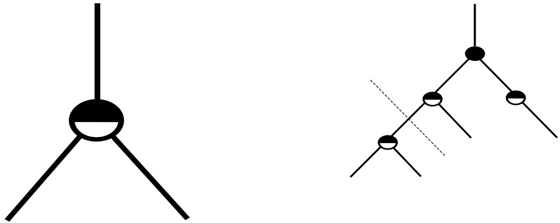
$$\mathbf{U}_\alpha \subset \mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2} \qquad (\alpha_1, \alpha_2 \text{ sons of } \alpha)$$

5) $\mathbf{v} \in \mathbf{U}_D$ (w.l.o.g. $\mathbf{U}_D = \text{ span } (\mathbf{v})$).

Subspace approximation: formulation with bases

$$\mathbf{U}_\alpha = \operatorname{span}\{\mathbf{b}_i^{(\alpha)} : 1 \le i \le r_\alpha\}$$

$$\mathbf{b}_\ell^{(\alpha)} = \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{ij}^{(\alpha,\ell)}\, \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)} \qquad (\alpha_1, \alpha_2 \text{ sons of } \alpha \in T_D).$$



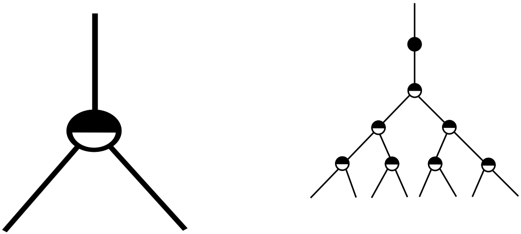
Coefficients $c_{ij}^{(\alpha,\ell)}$ form the matrices $C^{(\alpha,\ell)}$.

Final representation of $\mathbf{v}$ is

$$\mathbf{v} = \sum_{i=1}^{r_D} c_i^{(D)} \mathbf{b}_i^{(D)}\ , \quad \text{(usually with } r_D = 1\text{)}.$$

Orthonormal basis and recursive description

We can choose $\{\mathbf{b}_i^{(\alpha)} : 1 \leq i \leq r_\alpha\}$ orthonormal basis for $\alpha \in T_D$.



This is equivalent to
1) At the leaves ($\alpha = \{j\}, j \in D$): $\{\mathbf{b}_i^{(j)} : 1 \leq i \leq r_j\}$ is chosen orthonormal
2) The matrices $\left\{C^{(\alpha,\ell)} : 1 \leq \ell \leq r_\alpha\right\}$ are orthonormal (w.r.t. the Frobenius scalar product).

Hierarchical Tucker tensors

- ▷ Canonical decomposition
- ▷ Subspace approach (Hackbusch/Kühn, 2009)

(Example: $d = 5, \mathbf{U}_i \in \mathbb{R}^{n \times k_i}, \mathbf{B}_t \in \mathbb{R}^{k_t \times k_{t_1} \times k_{t_2}}$)

Hierarchical Tucker tensors

- Canonical decomposition — not closed, no embedded manifold!
- Subspace approach (Hackbusch/Kühn, 2009)

(Example: $d = 5, \mathbf{U}_i \in \mathbb{R}^{n \times k_i}, \mathbf{B}_t \in \mathbb{R}^{k_t \times k_{t_1} \times k_{t_2}}$)

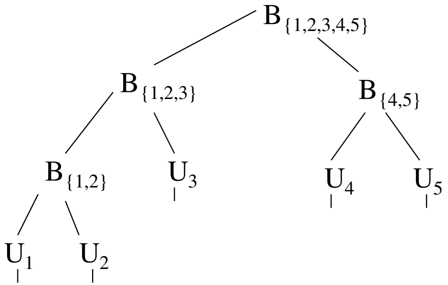
Hierarchical Tucker tensors

- Canonical decomposition   not closed, no embedded manifold!
- Subspace approach (Hackbusch/Kühn, 2009)

(Example: $d = 5, \mathbf{U}_i \in \mathbb{R}^{n \times k_i}, \mathbf{B}_t \in \mathbb{R}^{k_t \times k_{t_1} \times k_{t_2}}$)

Hierarchical Tucker tensors

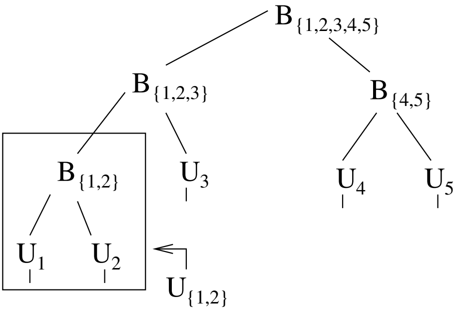
- Canonical decomposition   not closed, no embedded manifold!
- Subspace approach (Hackbusch/Kühn, 2009)



(Example: $d = 5, \mathbf{U}_i \in \mathbb{R}^{n \times k_i}, \mathbf{B}_t \in \mathbb{R}^{k_t \times k_{t_1} \times k_{t_2}}$)

Hierarchical Tucker tensors

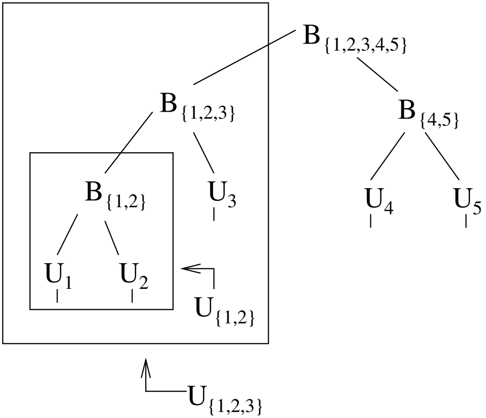
- Canonical decomposition — not closed, no embedded manifold!
- Subspace approach (Hackbusch/Kühn, 2009)



(Example: $d = 5, \mathbf{U}_i \in \mathbb{R}^{n \times k_i}, \mathbf{B}_t \in \mathbb{R}^{k_t \times k_{t_1} \times k_{t_2}}$)

Hierarchical Tucker tensors

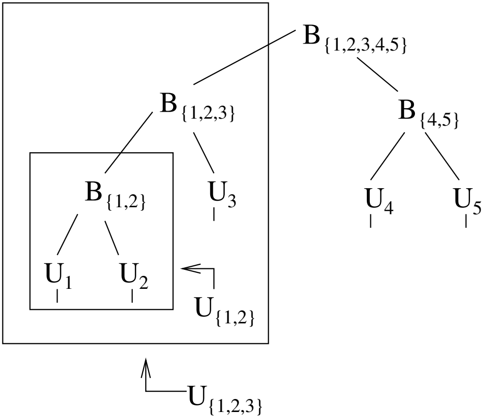
- Canonical decomposition   not closed, no embedded manifold!
- Subspace approach (Hackbusch/Kühn, 2009)



(Example: $d = 5, \mathbf{U}_i \in \mathbb{R}^{n \times k_i}, \mathbf{B}_t \in \mathbb{R}^{k_t \times k_{t_1} \times k_{t_2}}$)

Hierarchical Tucker tensors

- ▷ Canonical decomposition   not closed, no embedded manifold!
- ▷ Subspace approach (Hackbusch/Kühn, 2009)



(Example: $d = 5, \mathbf{U}_i \in \mathbb{R}^{n \times k_i}, \mathbf{B}_t \in \mathbb{R}^{k_t \times k_{t_1} \times k_{t_2}}$)

HOSVD bases (Hackbusch)

1) Assume $r_D = 1$. Then

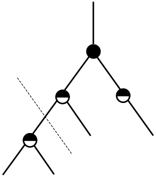
$$C^{(D,1)} = \Sigma_\alpha := diag\{\sigma_1^{(D)}, \ldots\},$$

where $\sigma_i^{(D)}$ are the singular values of $\mathcal{M}_{\alpha_1}(\mathbf{v})$
2) For non-leaf vertices $\alpha \in T_D$, $\alpha \neq D$, we have

$$\sum_{\ell=1}^{r_\alpha} (\sigma_\ell^{(\alpha)})^2 \, C^{(\alpha,\ell)} C^{(\alpha,\ell)\mathsf{H}} = \Sigma_{\alpha_1}^2,$$

$$\sum_{\ell=1}^{r_\alpha} (\sigma_\ell^{(\alpha)})^2 \, C^{(\alpha,\ell)\mathsf{T}} \overline{C^{(\alpha,\ell)}} = \Sigma_{\alpha_2}^2,$$

where α_1, α_2 are the first and second son of $\alpha \in T_D$ and Σ_{α_i} the diagonal of the singular values of $\mathcal{M}_{\alpha_i}(\mathbf{v})$.



Historical remarks – hierarchical Tucker approx.

Physics: (`I do not have suff. complete knowledge`)
The ideas are well established in many body quantum physics, quantum optics and quantum information theory

1. DMRG: S. White (91)
2. MPS: Roemmer & Ostlund (94), Vidal (03), Verstraete et al.
3. (tree) tensor networks : Verstrate, Cirac, Schollwöck, Wolf, Eisert .... (?)
4. MERA, PEPS, ...
5. Chemistry: Multilevel MCTDH $\simeq$ HT : Meyer et al. (2000)

Mathematics: Completely new in tensor product approximation

1. Khoromskij et al. (2006), remark about multi level Tucker
2. Lubich (book) (2008): review about Meyers paper
3. HT - Hackbusch (2009) et al.: Introduction of the subspace approximation!!!!
4. Grasedyck (2009): HT -sequential SVD
5. Oseledets & Tyrtishnikov (2009): TT and sequential SVD

Successive SVD for e.g. TT tensors

- Vidal (2003), Oseledets (2009), Grasedyck (2009)

1. Matricisation - unfolding

$$F(x_1, \ldots, x_d) \approx \mathbf{F}_{x_1}^{x_2, \ldots, x_d} \ .$$

   low rank approximation up to an error ϵ_1, e.g. by SVD.

$$\mathbf{F}_{x_1}^{x_2, \ldots, x_d} = \sum_{k_1=0}^{r_1} \mathbf{u}_{x_1}^{k_1} \mathbf{v}_{k_1}^{x_2, \ldots, x_d} \ ,$$

$$U_1(x_1, k_1) := \mathbf{u}_{x_1}^{k_1} \ , \ \ k_1 = 1, \ldots, r_1 \ .$$

2. Decompose $V(k_1, x_2, \ldots, x_d)$ via matricisation up to an accuracy ϵ_2,

$$\mathbf{V}_{k_1, x_2}^{x_3, \ldots, x_d} = \sum_{k_2}^{r_2} \mathbf{u}_{k_1, x_2}^{k_2} \mathbf{v}_{k_2}^{x_3, \ldots, x_d}$$

$$U_2(k_1, x_2, k_2) := \mathbf{u}_{k_1, x_2}^{k_2} \ .$$

3. repeat with $V(k_2, x_3, \ldots x_d)$ until one ends with $\left[\mathbf{v}_{k_{d-1}, x_d}\right] \mapsto U_d(k_{d-1}, x_d)$.

Example

The function

$$U(x_1,\ldots,x_d) = x_1 + \ldots + x_d \in \mathcal{H}$$

in canonical representation

$$U(x_1,\ldots,x_d) \;=\; x_1\cdot 1\cdot ..\cdot 1 \;+\; 1\cdot x_2\cdot 1\cdot ..\cdot 1 \;+\; \ldots \;+\; 1\cdot ..\cdot 1\cdot x_d$$

In Tucker format, let

$\mathbf{U}_{i,1} = 1, \mathbf{U}_{i,2} = x_i$ be the basis of $\mathbf{U}_i$ (non-orthogonal), $r_i = 2$.

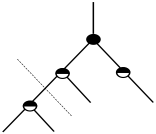
U inTT or MPS representation

$$U(\underline{x}) \;=\; x_1 + \ldots + x_d \;=\; \begin{pmatrix} x_1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ x_2 & 1 \end{pmatrix} \cdot \ldots \cdot \begin{pmatrix} 1 & 0 \\ x_{d-1} & 1 \end{pmatrix} \begin{pmatrix} 1 \\ x_d \end{pmatrix}.$$

(non-orthogonal), $r_1 = \ldots = r_{d-1} = 2$

Successive SVD for e.g. TT tensors

The above algorithm can be extended to HT tensors!



Error analysis

- Quasi best approximation

$$\|F(\mathbf{x}) - \mathbf{U}(x_1)\dots\mathbf{U}_d(x_d)\|_2 \leq \Big(\sum_{i=1}^{d-1}\epsilon_i^2\Big)^{1/2}$$
$$\leq \sqrt{d-1}\,\inf_{U\in\mathcal{T}^{\leq r}}\|F(\mathbf{x}) - U(\mathbf{x})\|_2 \,.$$

- Exact recovery: if $U \in \mathcal{T}^{\leq r}$, then is will be recovered exactly! (up ot rounding errors)

Ranks of TT and HT tensors

The minimal numbers r_i for a TT or HT respresentation of a tensor U is well defined.

1. The optimal ranks r_i, $\underline{r} = (r_1, \ldots, r_{d-1})$, of the TT decomposition are equal to the ranks of the following matrices, $i = 1, \ldots, d-1$

$$r_i = \text{rank of } \mathbf{A}_i = U_{x_1,\ldots,x_i}^{x_{i+1},\ldots,x_d} .$$

2. If $\nu \subset D$ then $\mu := D \backslash \mu$ the the *mode* ν-rank r_ν is given by

$$r_\nu = \text{rank of } \mathbf{A}_i = U_{\underline{x}_\nu}^{\underline{x}_\mu}.$$

(If $\nu = \{1, 2, 7, 8\}$ and $\mu = \{3, 4, 5, 6\}$ then $\underline{x}_\nu = (x_1, x_2, x_7, x_8)$.)

Closedness

A tree T_D is characterized by the property, if one remove one egde yields two seperate trees.

Observation: Let $\mathbf{A}_i$ with ranks rank$\mathbf{A}_i \leq r$. If $\lim_{i\to\infty} \|\mathbf{A}_i - \mathbf{A}\|_2 = 0$ then rank$\mathbf{A} \leq r$: $\Rightarrow$ closedness of Tucker and HT tensor in $\mathbb{T}^{\leq \underline{r}}$ (Falco & Hackbsuch).

$$\boxed{\mathbb{T}^{\leq \underline{r}} = \bigcup_{s \leq r} \mathbb{T}_{\underline{s}} \subset \mathcal{H} \text{ is closed!}}$$

due to Hackbusch & Falco

Landsberg & Ye: If a tensor network has not a tree structure, the set of all tensor of this form need not to be closed!

Landsbergs result - tensor networks

Let μ are the vertices and $s_\mu \in S_\mu$ are incident edges of a vertex μ, and V_μ the vector space attached to it. (it can be empty!)
A vertex μ is called *super critical* if

$$\dim V_\mu \geq \Pi_{s(\mu)\in S(\mu)} r_{s(\mu)} \ ,$$

and *critical* if $\dim V_\mu = \Pi_{s(\mu)\in S(\mu)} r_{s(\mu)}$

Lemma

Any supercritical vertex can be reduced to a critical one by Tucker decomposition.

Theorem (Landsberg-Ye)

If a (proper) loop a tensor network contains only (super)-critical vertices, then the network is not (Zariski)-closed.

I.e. if the ranks are too small! For $\bigotimes_{i=1}^{d} \mathbb{C}^2$, i.e. $n = 2$, Landsbergs result has no consequences.

Summary

For Tucker and HT redundancy can be removed (see next talk)

Table: Comparison

	canonical	Tucker	HT
complexity	$\mathcal{O}(ndr)$ ++	$\mathcal{O}(r^d + ndr)$ $-$	$\mathcal{O}(ndr + dr^3)$ TT- $\mathcal{O}(ndr^2)$ +
rank	no $r_c \geq$	defined r_T	defined r_{HT} , $r_T \leq r_{HT}$
closedness	no	yes	yes
essential redundancy	yes	no	no
recovery quasi best approx. best approx.	?? no no	yes yes exist but NP hard	yes yes exist but NP hard

Thank you for your attention.

References:

UNDER CONSTRUCTION,

G. Beylkin, M.J. Mohlenkamp *Numerical operator calculus in higher dimensions*, Proc. nat. Acad. Sci. USA, 99 (2002), 10246 – 10251.

G. Beylkin, M.J. Mohlenkamp *Algorithms for numerical analysis high dimensions*, SIAM J. Sci. Comput. 26, (2005), 2133 –2159.

W. Hackbusch, *Tensor spaces and numerical tensor calculus,SSCM, vol. 42, to appear, Springer, 2012.*

W. Hackbusch, S. Kühn, A new scheme for the tensor representation, Preprint 2, MPI MIS, 2009.

A. Falco, W. Hackbusch, On minimal subspaces in tensor representations Preprint 70, MPI MIS, 2010.

M. Espig, W. Hackbusch, S. Handschuh and R. Schneider *Optimization Problems in Contracted Tensor Networks*, MPI, MIS Preprint 66/2011, submitted to *Numer. Math.*

L. Grasedyck, *Hierarchical singular value decomposition of tensors*, SIAM. J. Matrix Anal. & Appl. 31, p. 2029, 2010.

C.J. Hillar and L.-H. Lim, *Most tensor problems are NP hard* preprint, (2012) arXiv:0911.1393v3

S. Holtz, Th. Rohwedder, R. Schneider, *On manifolds of tensors with fixed TT rank*, Num. Math. DOI: 10.1007/s00211-011-0419-7, 2011.

O. Koch, C. Lubich, *Dynamical tensor approximation*, SIAM J. Matrix Anal. Appl. 31, p. 2360, 2010.

I. Oseledets, E. E. Tyrtyshnikov, *Breaking the curse of dimensionality , or how to use SVD in many dimensions*, SIAM J. Sci. Comput., Vol.31 (5), 2009, 3744–3759.

J. M. Landsberg, Yang Qi, Ke Ye , On the geometry of tensor network states arXiv:1105.4449v2 [math.AG] to appear Quantum Information Computation

D. Kressner and C. Tobler `htucker` *A Matlab toolbox for tensor in hierarchical Tucker format* Tech Report, SAM ETH Zürich (2011)

G. Vidal Phys. Rev. Lett **101**, 110501 (2003)

Uschmajew, *Well-posedness of convex maximization problems on Stiefel manifolds and orthogonal tensor product approximations*, Numer. Math." 115-2, 309–331", 2010

`I do not have suff.  complete knowledge about physics literatur`