Traditional and novel tensor formats

R. Schneider (TUB Matheon)

John von Neumann Lecture – TU Munich, 2012

Setting - Tensors of order d

Goal: Generic perspective on methods for high-dimensional problems, i.e. problems posed on tensor spaces,

$$\mathcal{V} := \bigotimes_{i=1}^{d} V_i, \qquad \text{today:} \quad \mathcal{V} = \bigotimes_{i=1}^{d} \mathbb{R}^n = \mathbb{R}^{(n^d)}$$

Notation: $(x_1, \ldots, x_d) \mapsto U = U(x_1, \ldots, x_d) \in \mathcal{V}$

Main problem:

$$\dim \mathcal{V} = \mathcal{O}(n^d) \quad - \quad \text{Curse of dimensionality!}$$

e.g. $n = 100, d = 10 \rightsquigarrow 100^{10}$ basis functions,
$\rightsquigarrow$ coefficient vectors of 800×10^{18} Bytes $= 800$ Exabytes

Approach: Some higher order tensors can be constructed (data-) sparsely from lower order quantities.

As for matrices, incomplete SVD:

$$A(x_1, x_2) \approx \sum_{k=1}^{r} \sigma_k \big(\mathbf{u}_k(x_1) \otimes \mathbf{v}_k(x_2)\big)$$

Setting - Tensors of order d

Goal: Generic perspective on methods for high-dimensional problems, i.e. problems posed on tensor spaces,

$$\mathcal{V} := \bigotimes_{i=1}^{d} V_i, \qquad \text{today:} \quad \mathcal{V} = \bigotimes_{i=1}^{d} \mathbb{R}^n = \mathbb{R}^{(n^d)}$$

Notation: $(x_1, \ldots, x_d) \mapsto U = U(x_1, \ldots, x_d) \in \mathcal{V}$

Main problem:

$$\dim \mathcal{V} = \mathcal{O}(n^d) \quad - \quad \text{Curse of dimensionality!}$$

e.g. $n = 100, d = 10 \rightsquigarrow 100^{10}$ basis functions,
$\rightsquigarrow$ coefficient vectors of 800×10^{18} Bytes $=$ 800 Exabytes

Approach: Some higher order tensors can be constructed (data-) sparsely from lower order quantities.

As for matrices, incomplete SVD:

$$A(x_1, x_2) \approx \sum_{k=1}^{r} \sigma_k \big(\mathbf{u}_k(x_1) \otimes \mathbf{v}_k(x_2) \big)$$

Setting - Tensors of order d

Goal: Generic perspective on methods for high-dimensional problems, i.e. problems posed on tensor spaces,

$$\mathcal{V} := \bigotimes_{i=1}^{d} V_i, \qquad \text{today:} \quad \mathcal{V} = \bigotimes_{i=1}^{d} \mathbb{R}^n = \mathbb{R}^{(n^d)}$$

Notation: $\quad (x_1, \ldots, x_d) \mapsto U = U(x_1, \ldots, x_d) \in \mathcal{V}$

Main problem:

$$\dim \mathcal{V} = \mathcal{O}(n^d) \quad - \quad \text{Curse of dimensionality!}$$

e.g. $n = 100, d = 10 \rightsquigarrow 100^{10}$ basis functions,
$\rightsquigarrow$ coefficient vectors of 800×10^{18} Bytes $=$ 800 Exabytes

Approach: Some higher order tensors can be constructed (data-) sparsely from lower order quantities.

As for matrices, incomplete SVD:

$$A(x_1, x_2) \approx \sum_{k=1}^{r} \sigma_k \big(\mathbf{u}_k(x_1) \otimes \mathbf{v}_k(x_2)\big)$$

Setting - Tensors of order d

Goal: Generic perspective on methods for high-dimensional problems, i.e. problems posed on tensor spaces,

$$\mathcal{V} := \bigotimes_{i=1}^{d} V_i, \qquad \text{today:} \quad \mathcal{V} = \bigotimes_{i=1}^{d} \mathbb{R}^n = \mathbb{R}^{(n^d)}$$

Notation: $(x_1, \ldots, x_d) \mapsto U = U(x_1, \ldots, x_d) \in \mathcal{V}$

Main problem:

$$\dim \mathcal{V} = \mathcal{O}(n^d) \quad - \quad \text{Curse of dimensionality!}$$

e.g. $n = 100, d = 10 \rightsquigarrow 100^{10}$ basis functions,
$\rightsquigarrow$ coefficient vectors of 800×10^{18} Bytes $=$ 800 Exabytes

Approach: Some higher order tensors can be constructed (data-) sparsely from lower order quantities.

$\rightsquigarrow$ **Canonical decomposition** for order-d-tensors:

$$U(x_1, \ldots, x_d) \approx \sum_{k=1}^{r} \sigma_k \Big(\otimes_{i=1}^{d} \mathbf{u}_{i,k}(x_i) \Big).$$

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 $\Psi(\mathbf{r}_1, s_1, \ldots, \mathbf{r}_N, s_N)$

Tensors, basis vectors and linear operators

Tensor product basis from uni-variate basis $\{\mathbf{v}_i \in \mathcal{V}_i : i = 1, \ldots, n_i\}$:

$$\{\mathbf{V}_{k_1,\ldots,k_d} = \mathbf{v}_{k_1} \otimes \ldots \otimes \mathbf{v}_{k_d} = \bigotimes_{i=1}^{d} \mathbf{v}_{k_i} \in \mathcal{H} : 1 \le k_i \le n_i \, , \, 1 \le i \le d\}$$

Linear operators $\mathbf{A} : \mathcal{H} \to \mathcal{H}$ in index notation

$$\mathbf{A} =: \left(A_{\mathbf{x}}^{\mathbf{y}}\right) = \left(A_{x_1,\ldots,x_d}^{y_1,\ldots,y_l}\right)$$

and vectorized (tensorized)

$$(x_1, \ldots, x_d, y_1, \ldots, y_l) \to A(x_1, \ldots, x_d, y_1, \ldots, y_l) \in \mathcal{H}_{\mathbf{x}} \otimes \mathcal{H}_{\mathbf{y}} \; .$$

Matricisation or unfolding:

$$\mathbf{A} =: \left(A_{\mathbf{x}}^{\mathbf{k}}\right) \; , \; \mathbf{A} = \left(A_{x_1,\ldots,x_d}^{k_1,\ldots,k_l}\right) := \left(a_{x_1,\ldots,x_d;k_1,\ldots,k_l}\right) \; .$$

Canonical format

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

Single tensor product

$$(x_1, \ldots, x_d) \mapsto U(x_1, \ldots, x_d) = \Pi_{i=1}^{d} u_{i,k}(x_i) = \Pi_{i=1}^{d} u_i(x_i, k) \ ,$$

$$U = \bigotimes_{i=1}^{d} \mathbf{u}_{k,i} \ .$$

Canonical (CP) format, PARAFAC or r-term expansion,

$$\boxed{U(x_1, \ldots, x_d) = \sum_{k=1}^{r_C} U_k(\mathbf{x}) = \sum_{k=1}^{r_C} \Pi_{i=1}^{d} u_{i,k}(x_i).}$$

or

$$\boxed{U = \sum_{k=1}^{r_C} U_k = \sum_{k=1}^{r_C} \bigotimes_{i=1}^{d} \mathbf{u}_{i,k}}$$

Canonical format - pros

Definition (Canonical format)

$$U(x_1,\dots,x_d) = \sum_{k=1}^{r} \bigotimes_{i=1}^{d} u_{i,k}(x_i) = \sum_{i=1}^{r} \bigotimes_{\nu=1}^{d} u_i(x_i,k).$$

r - canonical rank (?)

Let $n := \max\{n_i : 1 \le i \le d\}$.

- Number of terms $r = r_c$ (canonical rank (?)) is invariant w.r.t. basis transformations
- canonical rank $r \le \sharp$ DOF for a given tensor product basis - best N-term approximation *(super adaptivity)!*
- there is an additional cost storing the components $\mathbf{u}_{i,k_i}$
- degrees of freedom (DOF) or better strorage complexity: $\boxed{\mathcal{O}(drn)}$
- complexity scaling is $\mathcal{O}(drn)$ instead of $\mathcal{O}(n^d)$ for the full tensor!

Counter example of Silva and Lim

Let $A = \bigotimes_{i=1}^d \mathbf{a}_i$, $B = \bigotimes_{i=1}^d \mathbf{b}_i \in \mathcal{H}$, (possibly $\langle \mathbf{a}_i, \mathbf{b}_i \rangle = 0 \ \forall i$).

$$
\begin{aligned}
\text{Let} \quad & U(x_1, \ldots, x_d) \\
:= \quad & U_1(x_1, \ldots, x_d) + \ldots + U_d(x_1, \ldots, x_d) \ , \ (\ U_i \perp U_j \ ) \\
:= \quad & b_1(x_1) a_2(x_2) \ldots a_d(x_d) + \cdots + a_1(x_1) \ldots a_{d-1}(x_{d-1}) b_d(x_d) \\
= \quad & \frac{1}{\epsilon} (a_1(x_1) + \epsilon b_1(x_1)) \cdots (a_d(x_d) + \epsilon b_d(x_d)) \\
& - \frac{1}{\epsilon} a_1(x_1) \cdots a_d(x_d) + \mathcal{O}(\epsilon) \\
=: \quad & V_\epsilon(x_1, \ldots, x_d) + \mathcal{O}(\epsilon) \ , \ \text{product -rule for } A' \ .
\end{aligned}
$$

$V_\epsilon \to U$ as $\epsilon \to 0$, but rank $r_c(U) = d \neq r_c(V_\epsilon) = 2$ if $d \geq 3$!!!

$$\Rightarrow$$

- $\mathcal{K}^{\leq r} := \{ U \in \mathcal{H} : U = \sum_{k=1}^r U_k \}$ is not closed.
  (nor weakly closed)
- The notion of canonical rank is not well defined! Border rank problem.

Remark: The above example shows the product rule for the directional derivative

Optimization Problems

Problem (Generic optimization problem (OP))

Given a cost functional $F : \mathcal{H} \to \mathbb{R}$ *and an admissible set* $\mathcal{M} \subset \mathcal{H}$ *finding*

$$\mathit{argmin} \left\{ F(W) : W \in \mathcal{M} \right\} .$$

Problem (Tensor product optimization problem (TOP))

$$U := \mathit{argmin} \left\{ F(W) : W \in \mathcal{M} \cap \mathcal{K}^{\leq r} \right\} \tag{1}$$

Here we consider a modified optimization problem where the original admissible set is confined to tensors of rank at most r. Most problems can cast into this form.

Example

1. Approximation: for given $U \in \mathcal{H}$ minimize

$$\boxed{F(W) = \|U - W\|^2_{\mathcal{H}} == \|U - W\|^2 \ , \ W \in \mathcal{K}^r}$$

2. solving equations: where $A, g : \mathcal{V} \to \mathcal{H}$,

$$AU = B \ \ or \ \ g(U) = 0$$

   here

$$\boxed{F(W) := \|AW - B\|^2 \ \text{ resp. } \ F(W) := \|g(W)\|^2 \ .}$$

3. or, if $A : \mathcal{V} \to \mathcal{V}'$ is symmetric and $B \in \mathcal{V}'$, $\mathcal{V} \subset \mathcal{H} \subset \mathcal{V}'$,

$$\boxed{F(W) := \frac{1}{2}\langle AW, W\rangle - \langle B, W\rangle}$$

4. computing the lowest eigenvalue of a symmetric operator $A : \mathcal{V} \to \mathcal{V}'$,

$$\boxed{U = \text{argmin}\,\{F(W) = \langle AW, W\rangle : \langle W, W\rangle = 1\} \ .}$$

In many cases $\mathcal{M} \cap \mathcal{K}^{\leq r} = \mathcal{K}^{\leq r}$. and most F are quadratic.

Tensor optimization problem

$U = \mathrm{argmin} F(W)$ means finding $u_{i,k} \in V$ s. t.

$U = \sum_{k=1}^{r} \bigotimes_{i=1}^{d} u_{i,k}$ The unknow components are

$$\mathbf{u} := (u_{i,k})_{k=1,\ldots,r\ ,i=1,\ldots,d} \in V^{r\times d} = \times_{i=1}^{d} \times_{k=1}^{r} V\ .$$

Definition

We introduce the (multi-linear) mapping

$$T : V^{r\times d} \to \mathcal{H} = \bigotimes_{\nu=1}^{d} V\ , \mathbf{u} \mapsto U = T(\mathbf{u}) := \sum_{k=1}^{r} \bigotimes_{k=1}^{d} u_{i,k}\ .$$

Then our original optimization problem takes the form TOP1: Finding

$$\boxed{\mathbf{u} = \mathrm{argmin}\{J(\mathbf{w}) := F(T(\mathbf{w})) : T(\mathbf{w}) \in \mathcal{M}, \mathbf{w} \in V^{r\times d}\}\ .}$$

Even, if OP is well posed, TOP, or equivalently TP1, will be no longer be well posed!

Canonical format - cons

In order to get rid of the non-closedness, further assumptions are required

- assuming an upper bound $\|U_k\| \leq C$, by penalization, see Espig et al.
- assuming positivity $U_k(\mathbf{x}) \geq 0$, by inequality constraints
- assuming orthogonality $\langle U_k, U_l \rangle = 0$, $k \neq l$, (???? not yet clear how to do this).

Redundancy could only partially be removed. Obivious redundancy could be avoided, see Espig etc.

Approximation or optimization suffer from all kind of troubles:

- no uniqueness
- often trapped by local minima
- Canonical rank r is not low, in real applications.
- Usually, it becomes worse with increasing ranks.
- there is no dynamical formulation

Canonical format cannot be recommended without serious warnings!

Algorithmic treatment of (TP3)

Goal: First/second order methods for (TP3).

Directional derivatives of $T(\mathbf{u}) := \sum_{i=1}^{r} \bigotimes_{\nu=1}^{d} u_{i,\nu}$ *w.r.t.* $u_{k,\alpha}, u_{\ell,\beta}$:

$$\left[T'_{(k,\alpha)}(\mathbf{u})\right](v) = \Big(\bigotimes_{\nu=1}^{\alpha-1} u_{k,\nu}\Big) \otimes v \otimes \Big(\bigotimes_{\nu=\alpha+1}^{d} u_{k,\nu}\Big),$$

$$\left[T''_{(k,\alpha),(\ell,\beta)}(\mathbf{u})\right](v,w) = \delta_{k,\ell}(1-\delta_{\alpha,\beta})\Big(\bigotimes_{\nu=1}^{\alpha-1} u_{k,\nu}\Big)\otimes v\otimes\Big(\bigotimes_{\nu=\alpha+1}^{\beta-1} u_{k,\nu}\Big)\otimes w\otimes\Big(\bigotimes_{\nu=\beta+1}^{d} u_{k,\nu}\Big).$$

Directional derivatives of $J(\mathbf{u}) := F(T(\mathbf{u}))$*:*

$$\left[J'_{(k,\alpha)}(\mathbf{u})\right](v) = \Big\langle F'(T(\mathbf{u})),\ \left[T'_{(k,\alpha)}(\mathbf{u})\right](v)\Big\rangle,$$

$$\begin{aligned}\left[J''_{(k,\alpha),(\ell,\beta)}(\mathbf{u})\right](v,w) \quad &= \quad \Big\langle F''(T(\mathbf{u}))\left[T'_{(\ell,\beta)}(\mathbf{u})\right](w),\ \left[T'_{(k,\alpha)}(\mathbf{u})\right](v)\Big\rangle \\ &\quad + \Big\langle F'(T(\mathbf{u})), \left[T''_{(k,\alpha),(\ell,\beta)}(\mathbf{u})\right](v,w)\Big\rangle.\end{aligned}$$

Example: Second order methods, linear equation

$$\text{Solve } AU = B \text{ for } A = \sum_{j=1}^{R} \bigotimes_{\nu=1}^{d} A_{j,\nu}, \quad B = \sum_{j=1}^{s} \bigotimes_{\nu=1}^{d} b_{j,\nu}.$$

$$\big[F'(U)\big](V) = \langle AU, V\rangle - \langle B, V\rangle, \quad \big[F''(U)\big](V, W) = \langle AU, V\rangle.$$

Gradient of $J(\mathbf{u}) := F(T(\mathbf{u}))$:

$$\big[J'_{(k,\alpha)}(\mathbf{u})\big](.) = \sum_{i=1}^{r}\sum_{j=1}^{R}\Big(\prod_{\nu\neq\alpha}\langle A_{j,\nu}u_{i,\nu}, u_{k,\nu}\rangle\Big)A_{j,\alpha}u_{i,\alpha} - \sum_{j=1}^{s}\Big(\prod_{\nu\neq\alpha}\langle b_{j,\nu}, u_{k,\nu}\rangle\Big)b_{j,\alpha}$$

Approximate Hessian of $J(\mathbf{u})$:

$$\big[J''_{(k,\alpha),(\ell,\beta)}(\mathbf{u})\big](v, w) = \delta_{\alpha,\beta}\cdot\sum_{j=1}^{R}\Big(\prod_{\nu\neq\alpha}\langle A_{j,\nu}u_{k,\nu}, u_{\ell,\nu}\rangle\Big)\langle A_{j,\alpha}v, w\rangle + (1-\delta_{\alpha,\beta})[\ldots]$$

Example: Second order methods, linear equation

Solve $AU = B$ for $A = \sum_{j=1}^{R} \bigotimes_{\nu=1}^{d} A_{j,\nu}, \quad B = \sum_{j=1}^{s} \bigotimes_{\nu=1}^{d} b_{j,\nu}.$

$$\big[F'(U)\big](V) = \langle AU, V\rangle - \langle B, V\rangle, \quad \big[F''(U)\big](V, W) = \langle AU, V\rangle.$$

Gradient of $J(\mathbf{u}) := F(T(\mathbf{u}))$:

$$\big[J'_{(k,\alpha)}(\mathbf{u})\big](.) \quad = \quad \sum_{i=1}^{r}\sum_{j=1}^{R}\Big(\prod_{\nu\neq\alpha}\langle A_{j,\nu}u_{i,\nu}, u_{k,\nu}\rangle\Big)A_{j,\alpha}u_{i,\alpha} \; - \; \sum_{j=1}^{s}\Big(\prod_{\nu\neq\alpha}\langle b_{j,\nu}, u_{k,\nu}\rangle\Big)b_{j,\alpha}$$

Approximate Hessian of $J(\mathbf{u})$:

$$\big[J''_{(k,\alpha),(\ell,\beta)}(\mathbf{u})\big](v, w) \quad = \quad \delta_{\alpha,\beta}\cdot\sum_{j=1}^{R}\Big(\prod_{\nu\neq\alpha}\langle A_{j,\nu}u_{k,\nu}, u_{\ell,\nu}\rangle\Big)\langle A_{j,\alpha}v, w\rangle$$
$$+(1-\delta_{\alpha,\beta})[\ldots]$$

Example: Second order methods, linear equation

$$\text{Solve } AU = B \text{ for } A = \sum_{j=1}^{R} \bigotimes_{\nu=1}^{d} A_{j,\nu}, \quad B = \sum_{j=1}^{s} \bigotimes_{\nu=1}^{d} b_{j,\nu}.$$

$$\left[F'(U)\right](V) = \langle AU, V\rangle - \langle B, V\rangle, \quad \left[F''(U)\right](V, W) = \langle AU, V\rangle.$$

Gradient of $J(\mathbf{u}) := F(T(\mathbf{u}))$:

$$\left[J'_{(k,\alpha)}(\mathbf{u})\right](.) \quad = \quad \sum_{i=1}^{r} \sum_{j=1}^{R} \Big(\prod_{\nu\neq\alpha} \langle A_{j,\nu} u_{i,\nu}, u_{k,\nu}\rangle\Big) A_{j,\alpha} u_{i,\alpha} \; - \; \sum_{j=1}^{s} \Big(\prod_{\nu\neq\alpha} \langle b_{j,\nu}, u_{k,\nu}\rangle\Big) b_{j,\alpha}$$

Approximate Hessian of $J(\mathbf{u})$:

$$\begin{aligned}\left[J''_{(k,\alpha),(\ell,\beta)}(\mathbf{u})\right](v, w) \quad = \quad & \delta_{\alpha,\beta} \cdot \sum_{j=1}^{R} \Big(\prod_{\nu\neq\alpha} \langle A_{j,\nu} u_{k,\nu}, u_{\ell,\nu}\rangle\Big) \langle A_{j,\alpha} v, w\rangle \\ & + (1 - \delta_{\alpha,\beta})[\ldots]\end{aligned}$$

Greedy algorithm

Best $\mathcal{H}$ approximation: Given tensor $U \in \mathcal{H}$ find $T(\mathbf{u}) \in \mathcal{K}^{\leq r}$, which approximates

$$\|U - T(\mathbf{u})\| \to 0 \ , \ r \to \infty \ .$$

Start $V_1 := U$ For $k = 1, \ldots, r$

- Find best rank one approximation $\operatorname{argmin}\{\|U_k - V_k\| : U_k = \bigotimes_{i=1}^{d} \mathbf{u}_{k,i}\}$
- $V_{k+1} := V_k - U_k$
- repeat

Ehrlacher et al. have shown convergence of the greedy method for strongly convex F.
Result from Lim:

Theorem

For $d \geq 3$, computing a minimizer $\operatorname{argmin}\{\|U_k - V_k\| : U_k = \bigotimes_{i=1}^{d} \mathbf{u}_{k,i}\}$ is NP hard.

Remarks and conclusions

Further remarks:

1. Sum of exponentials (e.g. discretized Laplace-Transform)

$$\frac{1}{x_1 + \ldots + x_d} = \int_0^\infty e^{-s(x_1+\ldots+x_d)} ds \approx \sum_k \omega_k e^{-s_k(x_1+\ldots+x_d)} \; , \; x_i > 0 \, , 1 \leq i \leq d \, .$$

2. Sufficient conditions for uniqueness of the CP format are stated by Kruskal.
3. For symmetric tensors a recovery scheme has been provided by P. Comon.
4. The canonical format is often considered as the `true` tensor product format.
5. Its complexity scaling is ideal $\mathcal{O}(ndr)$!
6. .... there is much more to tell you ...

Conclusion:

The canonical format is like a *Queen of Night*, you and everybody like to dance with. But you must be an extremely good dancer, not to get dissapointed.

Tucker format - subspace approximation

Tucker format

$$U(x_1, \ldots, x_d) = \sum_{k_1=1}^{r_1} \ldots \sum_{k_d=1}^{r_d} V(k_1, \ldots, k_d) \bigotimes_{i=1}^{d} u_{k_i}(x_i)$$

$\underline{r} = (r_1, \ldots, r_d)$ rank (array) $r := \max\{r_i : 1 \leq i \leq d\}$ - max.
Tucker rank.
$(k_1, \ldots, k_d) \to V(k_1, \ldots, k_d)$ - core tensor in $\mathbb{R}^{r_1 \times \cdots \times r_d}$.

Tucker format is a subspace approximation:
$V_i = \text{span}\,\{\mathbf{u}_{k_i} : k_i = 1, \ldots, r_i\}$ univariate basis $x_i \mapsto \mathbf{u}_{k_i}(x_i)$ can be choosen to ONB $\langle u_{i,k_i}, u_{i,l_i} \rangle = \delta_{l,k}$.

$$U = \sum_{k_1=1}^{r_1} \ldots \sum_{k_d=1}^{r_d} V(k_1, \ldots, k_d) \bigotimes_{i=1}^{d} \mathbf{u}_{k_i}$$

Remark: The two formats are identical if $d = 2$ or $r_c = 1$.

Tucker format

Degr. of freedom (DOF): $\boxed{DOF : \mathcal{O}(r_{max}^d + dr_{max}n)}$, n := $\max\{n_i : 1 \leq i \leq d\}$.
If $d = 2$ or $r = 1$ it is $\approx$ canonical format. Curse of dimensionality!

But: $\mathcal{T}^{\leq \underline{r}} = \bigcup_{s_i=1}^{r_i} \mathcal{T}^{\underline{s}} = \operatorname{clos} \mathcal{T}^{\underline{r}}$ is closed (Falco & Hackbusch)- Segre variety

$\mathcal{T}^{\underline{r}}$ is an embedded manifold (Lubich et al.)

Tucker format ios well established in phyiscs: in particular for anti-(symmetric) functions. It allows the treamtment of dynamical problems.
Examples: MCSCF (Multi configuration self-consistent field), Analysis: infinite-dim.: Friesecke, M. Lewin , TDMCH (time dep. multi-configurational Hartree approximation)

Stable (submanifold), But $\boxed{DOF : \mathcal{O}(r_{max}^d + dr_{max}n)}$, Curse of dimensionality!

- Is there a format breaking the curse of dimensionality, and inheriting the properties of the Tucker format?

Hierarchical Tucker (HT) format

Hackbusch & Kühn (2009), Grasedyck (2010)

Hierarchical MCTDH *Mayer et al.* (2000) Matrix product states (1992) – (Tree) Tensor networks: e.g. *Vidal, Schollwöck etc. (2003)*

Noteable special case of HT: TT format , Oseledets & Tyrtyshnikov

TT-representation of U

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

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

with component tensors $U_i(k_{i-1},x_i,k_i) \in \mathbb{R}^{r_{i-1}\times n_i\times r_i}$.

DOFs: $\mathcal{O}(ndr^2)$, TT and HT inherits nice properties from the Tucker format!

Thank you
for your attention.

References:

J. M. Landsberg, Tensors: Geometry and Applications, Graduate Studies in Mathematics, vol. 128, AMS

T. G. Kolda, B. W. Bader, Tensor decompositions and applications, SIAM Review Vol. 51, **3**, 455-500,

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 and W. Hackbusch *A Regularized Newton method for the Efficient Approximation of Tensors Represented in the Canonical Tensor Format* MIS Preprint 2010, submitted to: Numerische Mathematik

M. Espig, W. Hackbusch, Th. Rohwedder and R. Schneider, Variational calculus with sums of elementary tensors of fixed rank, to appear in Numer. Math. SPP 1324 Preprint ,

E. Cancs, T. Lelivre, V. Ehrlacher *Convergence of a greedy algorithm for high-dimensional convex nonlinear problems* Preprint 33 pages, 2010

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

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

P. Comon, G. Golub, L.-H. Lim, and B. Mourrain, *Symmetric tensors and symmetric tensor rank*, SIAM Journal on Matrix Analysis and Applications, 30 (2008), no. 3, pp. 12541279.

U. Schollwoeck, The density-matrix renormalization group in the age of matrix product states, arXiv:1008.3477v2 [cond-mat.str-el]

I.V. Oseledets, *Tensor-Train Decomposition*, SIAM J. Sci. Comput. 33, pp. 2295-2317, 2011.

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.