MATRICES AND OPTIMIZATION

Eugene Tyrtyshnikov

Institute of Numerical Mathematics of Russian Academy of Sciences
eugene.tyrtyshnikov@gmail.com

Optimization at Work
MFTI, Dolgoprudny, October 27, 2017

Matrix approximation as a multivariate optimization problem

$$\min_{U,V \in \mathbb{R}^{n\times k}} ||A - UV^\top||$$

U and V are never unique!
The product $UV^\top$ is generically unique.

Solution heavily depends on the norm!

$2kn$ variables – quite a lot!

Singular Value Decomposition as a miracle of optimization

$$A = U\Sigma V^\top = \sum_{\alpha=1}^{r} \sigma_\alpha u_\alpha v_\alpha^\top \approx \sum_{\alpha=1}^{k} \sigma_\alpha u_\alpha v_\alpha^\top$$

Frobenius norm error equals $\sqrt{\sum\limits_{\alpha=k+1}^{r} \sigma_\alpha^2}$.

The same approximation is the best in any unitarily invariant norm (Mirsky theorem).

Cross approximation as a remedy for astronomically large matrices

$$A \approx CGR, \quad C \in \mathbb{R}^{n\times k}, \quad G \in \mathbb{R}^{k\times k}, \quad R \in \mathbb{R}^{r\times k}$$

S.Goreinov-E.Tyrtyshnikov (2000):

$$||A - C\hat{A}^{-1}R||_C \leqslant (k+1)E_2^{\mathrm{best}} = (k+1)\sigma_{k+1}$$

A.Osinsky- N.Zamarashkin (2017):

$$||A - C\hat{A}^{-1}R||_F \leqslant (k+1)E_F^{\mathrm{best}} = (k+1)\sqrt{\sum_{\alpha=k+1}^{r} \sigma_\alpha^2}$$

MAXIMIZATION OF VOLUME

$$A \approx A_r = Q \begin{bmatrix} A_{11} & A_{12} \end{bmatrix}$$

$$Q = \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} A_{11}^{-1}$$

THEOREM:

$$|Q_{ij}| \leqslant 1$$

PROOF

$$Q = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \\ q_{r+1,1} & \dots & q_{r+1,r} \\ \dots & \dots & \dots \\ q_{n1} & \dots & q_{nr} \end{bmatrix}$$

Necessary for the maximal volume:
$|q_{ij}| \leqslant 1, \quad r+1 \leqslant i \leqslant n, \quad 1 \leqslant j \leqslant r.$

Otherwise, swapping the rows increases the volume!

REPRESENTATION PROBLEM

Given vectors $a_1, \ldots, a_n \in \mathbb{C}^r$, select $k \geqslant r$ of them, $a_{i_1}, \ldots, a_{i_k}$, so that each expansion

$$a_i = q_{i,i_1} a_{i_i} + \ldots + q_{i,i_k} a_{i_k}$$

has *sufficiently small* coefficients.

Equivalently, if A has rows $a_1, \ldots, a_n$, then find B with rows $a_{i_1}, \ldots, a_{i_k}$ such that

$$A = QB,$$

where Q has *sufficiently small* entries.

SELECTING A FRAME USING VOLUMES OF RECTANGULAR MATRCES

A.Mikhalev and I.Oseledets: representations through a larger system but with smaller coefficients .

Given $A \in \mathbb{C}^{n \times r}$ with $r = \mathrm{rank}\, A$, find a submatrix $B \in \mathbb{C}^{k \times r}$ of maximal volume. Then

THEOREM. *There exists* $Q \in \mathbb{C}^{n \times k}$ *s.t.* $A = QB$ *and each row* q *of* Q *satisfies the inequality*

$$||q||_2 \leqslant \sqrt{\frac{r}{k - r + 1}}.$$

COROLLARY. *Each entry of* Q *can be made arbitrarily small by choosing* $k = cr$ *with* c *sufficiently large but independent of* n.

INTRODUCE A REDUCED VOLUME

A.Osinsky – N.Zamarashkin

$$V_r(A) := \prod_{i=1}^{r} \sigma_i(A)$$

ADVANTAGES OF REDUCED VOLUME WITH A LARGER CROSS

THEOREM (A.Osinsky – N.Zamarashkin)

Let

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \in \mathbb{C}^{M \times N},$$

where $A_{11} \in \mathbb{C}^{m \times n}$ is of maximal r-reduced volume among all $m \times n$ submatrices and $\operatorname{rank} A_{11} \geqslant r$. Then

$$||A - CA_{11}^{\dagger}R||_C \leqslant \sqrt{1 + \frac{r}{m-r+1}}\sqrt{1 + \frac{r}{n-r+1}}\ \sigma_{r+1},$$

$$C = \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix}, \quad R = \begin{bmatrix} A_{11} & A_{12} \end{bmatrix}.$$

COROLLARY OF THE REDUCED VOLUME THEOREM

Taking a cross with the intersection matrix of size $(2r-1) \times (2r-1)$ we can guarantee the estimate

$$||A - CA_{11}^{\dagger}R||_C \leqslant 2\sigma_{r+1}(A).$$

It holds if A_{11} is of maximal r-reduced volume.

A.Osinsky, N.Zamarashlin, *Pseudo-skeleton approximations with better accuracy estimates*, submitted to LAA, 2016.

Cross approximation is a chase for large elements

MAXVOL for rank-1 approximation delivers a close to maximal element with high probability depending on the choice of the initial column (A.Osinsky'2017).

CROSS INTERPOLATION HISTORY

1985 Knuth: Semi-optimal bases for linear dependencies

1995 Tyr., Goreinov, Zamarashkin: $A = CGR$ pseudoskeleton

2000 Tyr.: incomplete cross approximation with ALS maxvol

2000 Bebendorf: ACA = Gaussian elimination

2001 Tyr., Goreinov: maximum volume principle,
quasioptimality $\|\text{ cross }\|_C \le (r+1)\|\text{ best }\|_2$

2006 Mahoney et al: randomized CUR algorithm

2008 Oseledets, Savostyanov, Tyr.: Cross3D

2009 Oseledets, Tyr.: TT-Cross

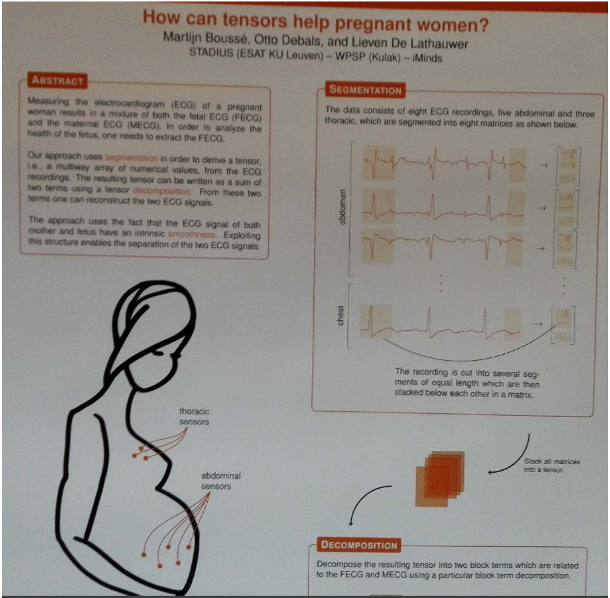
2010 J.Schneider: function-related quasioptimality
$\|\text{ cross }\|_C \le (r+1)^2\|\text{ best }\|_C$

2011 Tyr., Goreinov: quasioptimality
$\|\text{ cross }\|_C \le (r+1)^2\|\text{ best }\|_C$

2013 Ballani, Grasedyck, Kluge: HT-Cross

2013 Townsend, Trefethen -- Chebfun2

TENSORS ARE EVERYWHERE



INTRACTABLY BIG DATA MUST POSSESS STRUCTURE

Arrays with d indices of size $n \times \ldots \times n$:

$$a(i_1, \ldots, i_d), \quad 1 \leqslant i_1, \ldots, i_d \leqslant n$$

$n = 2,\ d = 300 \quad \Rightarrow \quad$ the entries $2^{300} \gg 10^{83}$
more than atoms in the universe.

Curse of dimensionality!

Multiplication of tensors

Given tensors $A^{(1)}, \dots, A^{(s)}$ with the elements

$$A^{(1)}_{i^1_1,\dots,i^1_{d_1}}, \quad \dots, \quad A^{(s)}_{i^s_1,\dots,i^s_{d_s}}$$

assume that in the whole list of indices $i_1, \dots, i_k$ occur only once and $j_1, \dots, j_l$ occur twice or more times. By *product of tensors* $A^{(1)}, \dots, A^{(s)}$ we mean a tensor B with the elements

$$B_{i_1,\dots,i_k} = \sum_{j_1,\dots,j_l} A^{(1)}_{i^1_1,\dots,i^1_{d_1}} \dots A^{(s)}_{i^s_1,\dots,i^s_{d_s}}.$$

We may write

$$B = A^{(1)} \dots A^{(s)}$$

but the detailed specification of indices is still needed.

Classic tensor decompositions

CANONICAL POLYADIC:

$$a = g_1 \dots g_d$$

$$a(i_1, \dots, i_d) = \sum_{\alpha} g_1(i_1, \alpha) \ \dots \ g_d(i_d, \alpha)$$

TUCKER:

$$a = t\, u_1 \dots u_d$$

$$a(i_1, \dots, i_d) = \sum_{\alpha_1, \dots, \alpha_d} t(\alpha_1, \dots, \alpha_d)\, u_1(i_1, \alpha_1) \ \dots \ u_d(i_d, \alpha_d)$$

SOME TENSOR DECOMPOSITIONS REDUCE TENSORS TO MATRICES

Two new names are now widely used in numerical analysis:

- TT (Tensor Train) – Moscow, INM (2009)
- HT (Hierarchical Tucker) – Leipzig, MPI (2009)

Both use *low-rank matrices*.

Both use the same *dimensionality reduction tree*.

TENSOR TRAIN IN d DIMENSIONS

$$a(i_1 \dots i_d) = A^{(1)}_{i_1} A^{(2)}_{i_2} \dots A^{(d)}_{i_d} =$$

$$\sum g_1(i_1 \alpha_1) g_2(\alpha_1 i_2 \alpha_2) \dots$$

$$\dots g_{d-1}(\alpha_{d-2} i_{d-1} \alpha_{d-1}) g_d(\alpha_{d-1} i_d)$$

d-tensor reduces to 3-tensors $g_k(\alpha_{k-1} i_k \alpha_k)$.

If the maximal size is $r \times n \times r$ then
the number of tensor-train elements does not exceed

$$dnr^2 \ll n^d.$$

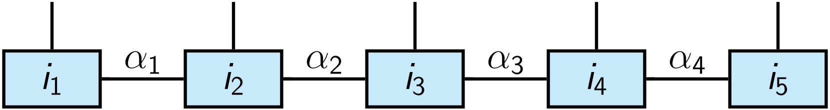
DIFFERENT FACES OF ONE THING

Tensor Train = Matrix Product State = Linear Tensor Network

$$a(i_1, i_2, i_3, i_4, i_5) =$$

$$\sum g_1(i_1, \alpha_1) g_2(\alpha_1, i_2, \alpha_2) g_3(\alpha_2, i_3, \alpha_3) g_4(\alpha_3, i_4, \alpha_4) g_5(\alpha_4, i_3)$$

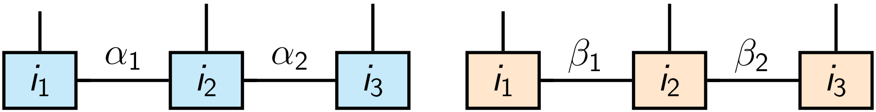
$$= \underbrace{A_1^{(i_1)}}_{1 \times r_1} \underbrace{A_2^{(i_2)}}_{r_1 \times r_2} \underbrace{A_3^{(i_3)}}_{r_2 \times r_3} \underbrace{A_4^{(i_4)}}_{r_3 \times r_4} \underbrace{A_5^{(i_5)}}_{r_4 \times 1}$$



EASY OPERATIONS ON TENSORS

e.g. summation

$$a(i_1, i_2, i_3) = \underbrace{A_1^{(i_1)}}_{1\times r_1} \underbrace{A_2^{(i_2)}}_{r_1\times r_2} \underbrace{A_3^{(i_3)}}_{r_2\times 1}, \qquad b(i_1, i_2, i_3) = \underbrace{B_1^{(i_1)}}_{1\times s_1} \underbrace{B_2^{(i_2)}}_{s_1\times s_2} \underbrace{B_3^{(i_3)}}_{s_2\times 1}$$



$$(a+b)(i_1, i_2, i_3) = \begin{bmatrix} A_1^{(i_1)} & B_1^{(i_1)} \end{bmatrix} \begin{bmatrix} A_2^{(i_2)} & \\ & B_2^{(i_2)} \end{bmatrix} \begin{bmatrix} A_3^{(i_3)} \\ B_3^{(i_3)} \end{bmatrix}$$

A NEW PARADIGM OF COMPUTATIONS

only through low-parametric formats

- $A = A(p), \quad B = B(q), \quad C = C(s)$
- To implement $C = A * B$ we should devise *fast algorithms* for getting s from p and q.
- *General algebraic method* for a wide class of applications!
- We can use classical methods of numerical analysis together with TT-approximation.

WHAT IS OUR CLASS OF TENSORS?

$$A_k = [a(i_1 \dots i_k \, ; \, i_{k+1} \dots i_d)] =$$

$$\left[\sum u_k(i_1 \dots i_k \, ; \, \alpha_k) \; v_k(\alpha_k \, ; \, i_{k+1} \dots i_d) \right] = U_k V_k^\top$$

$$u_k(i_1 \dots i_k \alpha_k) = \sum g_1(i_1 \alpha_1) \dots g_k(\alpha_{k-1} i_k \alpha_k)$$

$$v_k(\alpha_k i_{k+1} \dots i_d) = \sum g_{k+1}(\alpha_k i_{k+1} \alpha_{k+1}) \dots g_d(\alpha_{k-1} i_d)$$

THE MAIN PROPERTY OF THE CLASS:
all matrices A_k must be (close to) low-rank matrices
(I.Oseledets-E.Tyrtyshnikov)

EVERYTHING REDUCES TO MATRICES

Tensor train can be viewed as a *rank-structured representation* for matrices $A_1, \dots, A_{d-1}$.

- **Structured SVD can be computed for them simultaneously just in $O(dnr^3)$ operations!**
- **Tensor train can be constructed if we know low-rank decompositons for matrices $A_1, \dots, A_{d-1}$.**
- **Moreover, it can be construced from cleverly chosen *crosses* in some small submatrices of those matrices.**

Density Matrix Renormalization Group (DMRG)

$$g_1 g_2 g_3 \cdots g_{d-1} g_d \to (g_1 g_2) g_3 \cdots g_{d-1} g_d \to$$

$$g_1 (g_2 g_3) \cdots g_{d-1} g_d \to \cdots \to g_1 \cdots g_{d-1} g_d$$

A recent alternative is Alternating Minimization of Energy (AMEN) by S.Dolgov and D.Savostyanov.

TENSOR TRAIN COMES

FROM SMALL CROSSES IN THE UNFOLDING MATRICES

$$\widetilde{A}(i_1 \dots i_d) = \sum \prod_{k=1}^{d} A(J_{\leqslant k-1},\ i_k,\ J_{>k})\,[A(J_{\leqslant k},\ J_{>k})]^{-1}$$

A QUASI-OPTIMALY RESULT

$$\widetilde{A}(i_1 \dots i_d) = \sum \prod_{k=1}^{d} A(J_{\leqslant k-1},\, i_k,\, J_{>k})\, [A(J_{\leqslant k},\, J_{>k})]^{\dagger}_{\tau_k}$$

THEOREM.

$$||A - \widetilde{A}||_C \leqslant c(r) \cdot \underbrace{||F||_C}_{\texttt{BEST APPROXIMATION ERROR}}$$

$$c(r) = \frac{(4r)^{\lceil \log_2 d \rceil} - 1}{4r - 1}(r+1)^2$$

Tensorisation of vectors and matrices

Any vector of size $N = n_1 \dots n_d$ can be viewed as a d-tensor and any $N \times N$ matrix

$$a(i,j) = a(i_1 \dots i_d,\ j_1 \dots j_d)$$

can be viewed as a $2d$-tensor, and as a d-tensor, e.g.

$$a(i_1 j_1, \dots, i_d j_d)$$

of size $n_1^2 \times \dots \times n_d^2$.

Tensorization with TT may crucially decrease the number of representation parameters!

FAST SUMMATION OF ELEMENTS OF AN ASTRONOMICALLY HUGE VECTOR

$$i = \overline{i_1 i_2 \dots i_d} \qquad d = 83$$

$$a(i) = a(i_1, \dots, i_d) = \sum_{\alpha_1, \dots, \alpha_{d-1}} g_1(i_1, \alpha_1) g_2(\alpha_1, i_2, \alpha_2) \dots g_d(\alpha_{d-1}, i_d)$$

$$\sum_{i_1, \dots, i_d} a(i_1, \dots, i_d) = \sum_{\alpha_1, \dots, \alpha_{d-1}} \hat{g}_1(\alpha_1) \hat{g}_2(\alpha_1, \alpha_2) \dots \hat{g}_d(\alpha_{d-1})$$

$$\hat{g}_k = \sum_{i_k} g_k$$

TENSOR-TRAIN INTEGRATOR

Compute a d-dimensional integral

$$I(d) = \int \sin(x_1 + x_2 + \ldots + x_d)\ dx_1 dx_2 \ldots dx_d =$$

$$\mathrm{Im} \int_{[0,1]^d} e^{i(x_1+x_2+\ldots+x_d)}\ dx_1 dx_2 \ldots dx_d = \mathrm{Im}((\frac{e^i - 1}{i})^d).$$

$n = 11$ nodes along each dimension $\Rightarrow$ in total n^d values! Only a very small part of them in needed for the constrcution of TT.

d	$I(d)$	Relative error	Time
1000	-2.637513e-19	3.482065e-11	11.60
2000	2.628834e-37	8.905594e-12	33.05
4000	9.400335e-74	2.284085e-10	105.49

TENSORISATION FOR 1-D INTEGRALS

For the integral

$$\int_0^\infty \frac{\sin x}{x} dx = \frac{\pi}{2}$$

the first step is to reduce integration to an appropriate bounded interval, the latter is computed using the rule of rectangles.

For machine precision we need about 2^{77} nodes. The vector of the function values at those nodes is considered as a tensor of size $2 \times 2 \times \ldots \times 2$.

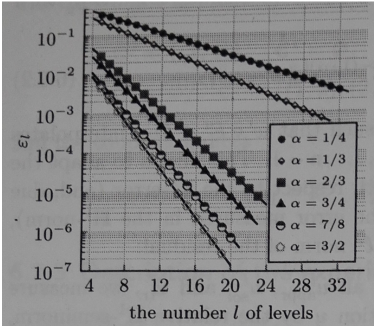
The TT-ranks were $\leqslant 12$. Time is less than 1 second on a labtop.

TT+FE+AMEN

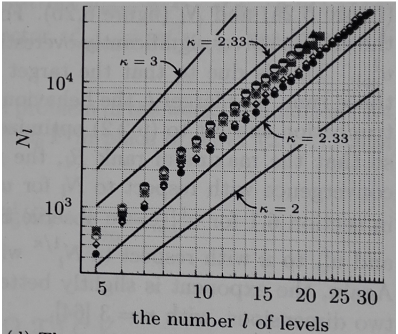
$$u_\Gamma(x) = r^\alpha \sin \alpha\phi(x), \quad x \in \Omega = (0,1)^2$$

$$\varepsilon_l \leqslant \exp\{-cN_l^{\frac{1}{\kappa}}\}, \quad N_l \text{ – the number of TT-elements}$$

THEOREM (V.Kazeev & C.Schwab). $\kappa \leqslant 5.$

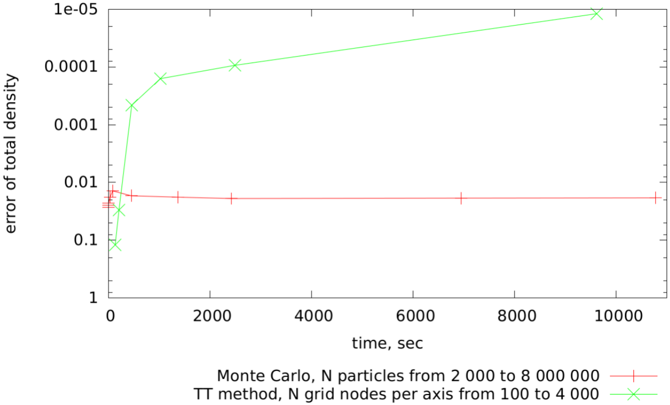


(a) Convergence with respect to l. The reference lines correspond to the exponential convergence $\varepsilon_l = C_\alpha 2^{-\tilde{\alpha} l}$ with C_α independent of l and with $\tilde{\alpha} = \min\{\alpha, 1\}$.



(d) The number N_l (3.3.2) of QTT parameters vs. l. The reference lines correspond to the algebraic growth $N_l = C_\alpha l^\kappa$ with κ and C_α independent of l.

TENSOR TRAIN VS MONTE CARLO FOR 2D SMOLUCHOWSKI EQUATIONS



Ballistic kernel: $K(u,v) = (u^{\frac{1}{3}} + v^{\frac{1}{3}})^2 \sqrt{\frac{1}{u} + \frac{1}{v}}$.

WELCOME THE BLESSING OF DINENSIONALITY

- Fokker-Planck, Smoluchovski equations
- Differential equations with parameters
- Green functions in integral equations
- Spin dynamics
- Global optimization algorithms
- Many others

RECENT BOOKS:

G. Golub and Ch. Van Loan, Matrix Computations, 4th edition, 2013.

W. Hackbusch, Tensor Spaces and Numerical Tensor Calculus, Springer, 2012.