Journal of Computational Physics 281 (2015) 116–134

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

The functions $u_j^{(l)}(x_j)$ in (1.1) and the corresponding vectors $u_{i_j}^{(l)}$ in (1.2) are normalized to have unit Frobenius norm so that the size of the terms is carried by their positive s-values, σ_l.

Numerical computations using such representations require an algorithm to reduce the number of terms for a given accuracy, . Such reduction can be achieved via Alternating Least Squares (ALS) (see, e.g., [35,20,17,13,14,56,39]). Specifically, given a CTD of rank r, ALS allows us to find a representation of the same form but with fewer terms,

$$\tilde{\ }_{i_1\ldots i_d} = \mathop{\ }\limits_{l=1}^{k} \tilde{\sigma}_l \tilde{u}_{i_1}^{(l)} \tilde{u}_{i_2}^{(l)} \cdots \tilde{u}_{i_d}^{(l)}, \quad k < r, \tag{1.3}$$

so that $\| \ - \tilde{\ } \| \quad \| \ \|$, where is a user-selected accuracy. Standard operations on separated representations of rank k, such as multiplication, may result in a large number, e.g., $\mathcal{O}(k^2)$, of intermediate terms. If the intermediate separation rank $r = \mathcal{O}(k^2)$ is reducible to $\mathcal{O}(k)$, then the cost of ALS can be estimated as $\mathcal{O}(d \cdot k^4 \cdot M) \cdot (\textit{number of iterations})$, where we assumed that $M_j = M$, $j = 1, \ldots, d$. Noting that

$$= \sum_{l=1}^{r} \sigma_l \; ^{(l)}, \quad \text{where} \quad ^{(l)} = \prod_{j=1}^{d} \mathbf{u}_j^{(l)}. \tag{1.5}$$

The standard Frobenius inner product between any two tensors and is defined as

$$\langle \; , \; \rangle = \sum_{i_1=1}^{M_1} \cdots \sum_{i_d=1}^{M_d} \; _{i_1 \ldots i_d} \; _{i_1 \ldots i_d} \tag{1.6}$$

which for CTDs reduces to

$$\langle \; , \; \rangle = \sum_{l=1}^{r_u} \sum_{m=1}^{r_v} \sigma_l^u \sigma_m^v \; ^{(l)}, \; ^{(m)} = \sum_{l=1}^{r_u} \sum_{m=1}^{r_v} \sigma_l^u \sigma_m^v \prod_{j=1}^{d} \mathbf{u}_j^{(l)}, \mathbf{v}_j^{(m)} \;, \tag{1.7}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product between component vectors. The Frobenius norm is then defined as

$$\| \; \|_F = \sqrt{\langle \; , \; \rangle}. \tag{1.8}$$

Remark 1.1. The directional vectors $\mathbf{u}_j^{(l)}$ may represent objects of different types, including proper one dimensional vectors,
w matrices or even low dimensional tensors. The vectors $\mathbf{u}_j^{(l)}$ can have a complicated structure (e.g., sparse

We note that several alternative tensor formats, such as

2.2. Interpolative matrix decomposition

Our approach also relies on computing the interpolative decomposition of a matrix. The idea of matrix ID is to find, for a given accuracy , a near optimal set of columns (rows) of a matrix so that the rest of columns (rows) can be represented as a linear combination from the selected set. Algorithmically, this decomposition proceeds via pivoted QR factorization and so is

Lemma 2.4. *(See Observation 3.3 of [45].) The randomized matrix ID algorithm constructs matrices A_c and P such that*

1. *some subset of the columns of P makes up the $k \times k$ identity matrix*
2. *no entry of P has absolute values greater than* 2,
3. ⟨ than

3. Randomized tensor interpolative decomposition

3.1. On the interpolative decomposition of symmetric matrices

We briefly describe a matrix ID for symmetric matrices because it parallels the development for tensors (and, to our knowledge, its description is not available in the literature). Suppose A is an $m \times n$ matrix of rank k (for a given accuracy

Table 1
Estimates of computational complexity.

Reduction method	Computational complexity
ALS	$(d \cdot k^4 \cdot M) \cdot (\textit{number of iterations}) \quad (d \cdot k^6) \cdot (\textit{number of iterations})$
Tensor ID: random projection	$d \cdot k^3 \cdot M + k^4 \quad d \cdot k^5$

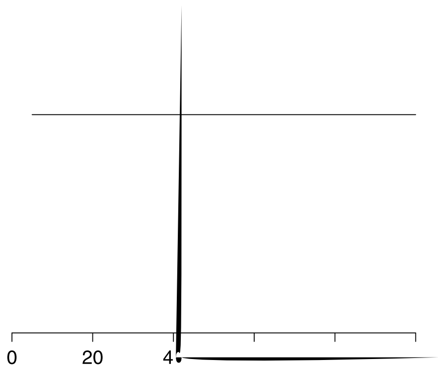
$$\mathsf{E}[\mathcal{R}_{i_1 \dots i_d}] = \mathsf{E}\left[\prod_{j=1}^{d} r_{i_j}\right] = \prod_{j=1}^{d} \mathsf{E}[r_{i_j}] = 0, \tag{3.12}$$

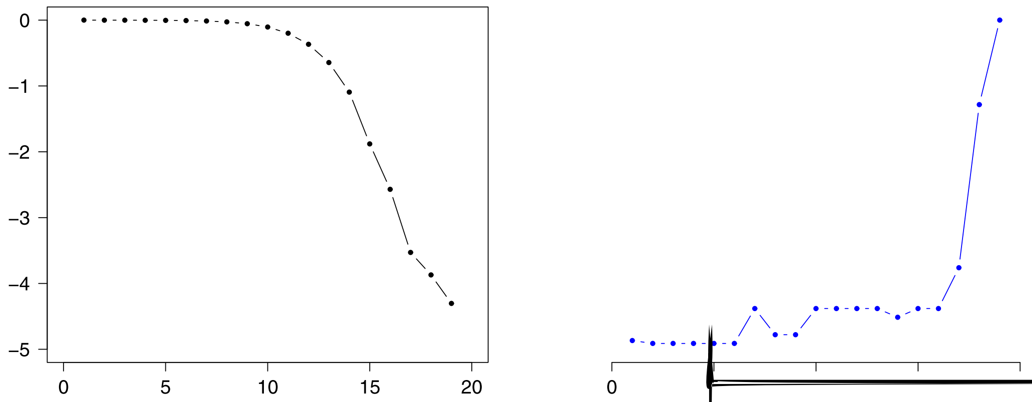
and

$$\mathsf{E}\left(\mathcal{R}_{i_1 \dots i_d} - \mathsf{E}[\mathcal{R}_{i_1 \dots i_d}]\right)^2 = \mathsf{E}\,\mathcal{R}^2_{i_1 \dots i_d} = \prod_{j}^{d}$$

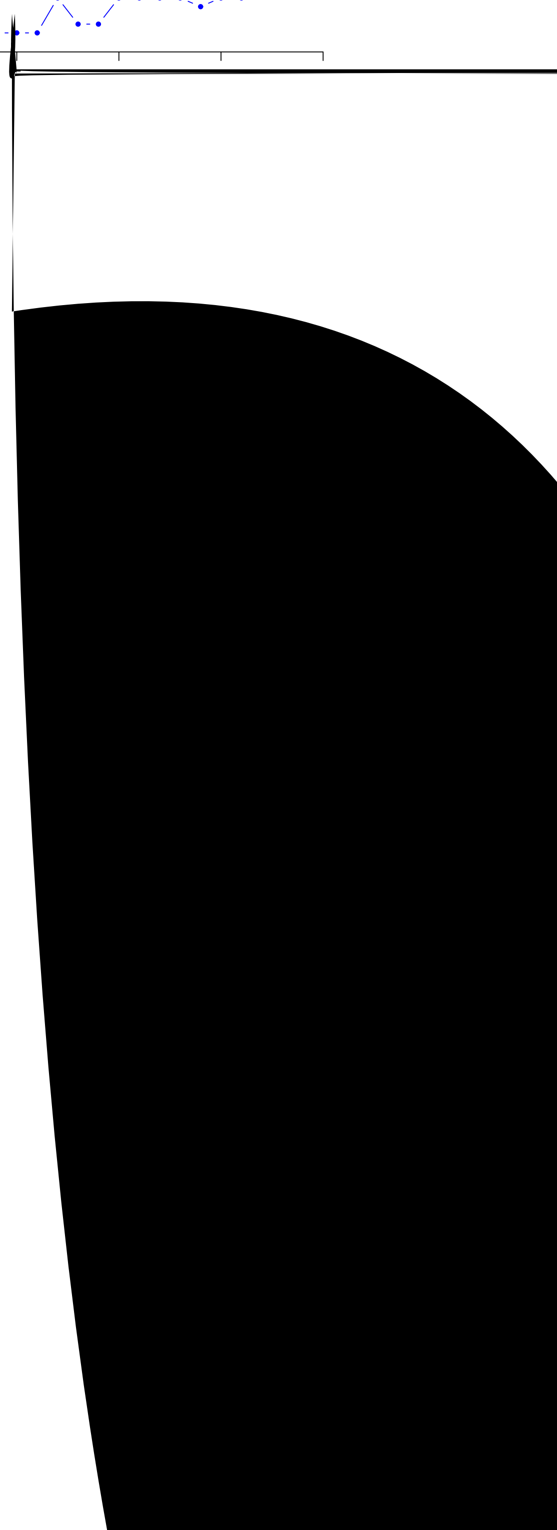
Remark 4.2. Although the definition of the s-*norm* parallels the matrix 2-norm and can be useful as a way of estimating errors, computing the s-norm exactly for arbitrary dense tensors is claimed to be an NP-hard problem in [36]. We note that for symmetric tensors, the global convergence of the iteration described above has been claimed in [40].

We discuss our approach to initialization below but first consider the cost of estimating the


0
20
4



(a)



- - - - - - - -

[18] E.J. Candès, T. Tao, Near-optimal signal recovery from random projections: universal encoding strategies?, IEEE Trans. Inf. Theory 52 (12) (2006) 5406–5425.
[19] E.J. Candès, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, IEEE Trans. Inf. Theory 52 (2) (Feb. 2006) 489–509.
[20] J.D. Carroll, J.J. Chang, Analysis of individual differences in multidimensional trika/T1_2 1 Tf 0 Tc 9.9626 0 0 9.9626 159.4407 606.533 6660(dif /T1_0 1 Tf -