

University of Vienna
Faculty of Mathematics

TENSOR METHODS FOR DATA SCIENCE AND
SCIENTIFIC COMPUTING
1st Exercise

Milutin Popovic

24. October, 2021

Contents

1	Introduction	1
2	Matrix approximation by SVD	2
3	Matrix approximation by LU without pivoting	2
4	Matrix approximation by LU with pivoting	3
5	Understanding the basis rows and columns	4

1 Introduction

In the following we consider two functions f and g defined pointwise on $[-1, 1]^2$ as

$$f(x, y) := \frac{1}{1 + x^2 + y^2} \quad \text{and} \quad (1)$$

$$g(x, y) := \sqrt{x^2 + y^2} \cdot \left(1 + \frac{1}{2} \cos(15x + 22y)\right), \quad (2)$$

for $x, y \in [-1, 1]$, with a grid of points t of size $n = 901$

$$t_i = 2 \frac{i - 1}{n - 1} - 1 \quad i = 1, \dots, n. \quad (3)$$

On this $n \times n$ grid we define two matrices A and B as the following

$$(A)_{ij} = a_{ij} := f(t_i, t_j) \quad \text{and} \quad (4)$$

$$(B)_{ij} = b_{ij} := g(t_i, t_j), \quad (5)$$

for all $i, j \in \{1, \dots, n\}$. Throughout we denote r as the rank of the approximate factorization of A or B , which actually only makes sense for $r \leq \text{rank}(A)$ or $r \leq \text{rank}(B)$ accordingly. Since the matrices A and B have the following ranks

$$\text{rank}(A) = 8 \quad \text{rank}(B) = 57, \quad (6)$$

only low rank approximations are valid and any rank approximation above the matrix rank is not valid, e.g. $r = 80$ approximation does not make any sense.

2 Matrix approximation by SVD

In this section we go through the error analysis in Forbenius and max norms of the truncated SVD on the matrices A and B . Considering a $n \times n$ (real or complex matrix) M them the truncated SVD of rank r will give an approximation of M , \tilde{M}

$$\tilde{M} = U\Sigma V^*, \quad (7)$$

where $U \in \mathbb{F}^{n \times r}$ ($\mathbb{F} = \mathbb{R}$ or \mathbb{C}), $\Sigma \in \mathbb{F}^{r \times r}$ with r largest singular values are the diagonal (else zero) and $V^{r \times n}$. The computational errors of the TSVD approximation of A and B are shown in figure 3 below.

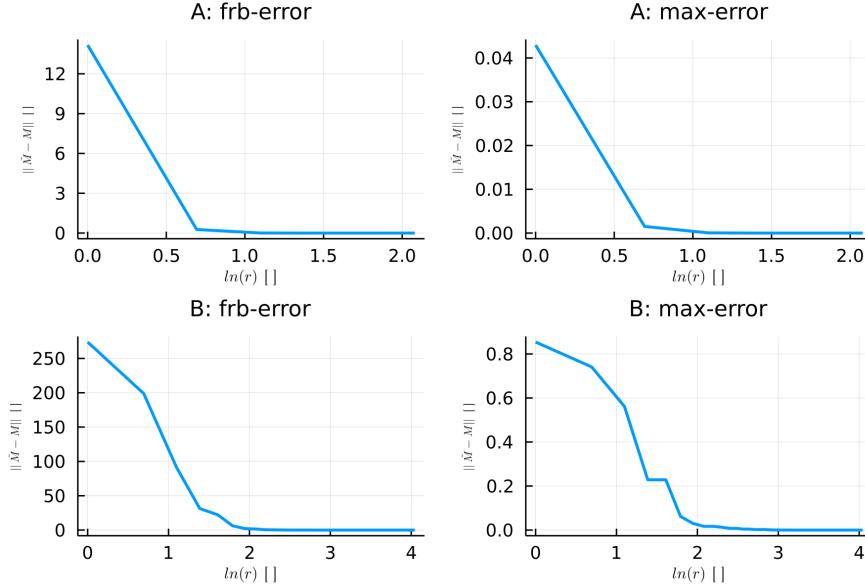


Figure 1: Error of rank r truncated SVD approximation of matrices A and B on a logarithmic scale

The error convergance can be explained by applying the Eckart-Young-Mirsky theorem for functions f , g and their singular value decomposition of rank r explicitly. In this sense both norms would be equivalent. Since we for a $p = \{0, \dots, r\}$ by the lecture script we have

$$A_p = \sum_{k=1}^p \sigma_k u_k v_k^*. \quad (8)$$

Thus for $r \rightarrow \text{rank}(A)$ the sum would converge to the original matrix A making the difference

$$\|A_p - A\|_F \text{ or } 2 \quad (9)$$

minimal, converging to zero.

3 Matrix approximation by LU without pivoting

In this section we look at the errors formed by the LU approximation of the leading principal submatrix without pivoting, when considering rank r cross approximation of the given matrix. Again we atone, that this procedure does not make sense for a submatrix of a size above the rank of the original matrix, since for the approximation the inverse of the submatrix is needed. Also one of the criteria is that the leading principal submatrix is a matrix with maximal volume and maximal rank. Obtaining the maximum volume submatrix is very difficult [1], thus we will use an algorithm called “Maxvol” [2] to obtain a quasioptimal maximum volume submatrix. In this let us say we are considering appropriate matrix M to approximate, to compute the quasioptimal submatrix we first need to select random r columns of M , then the algorithm “Maxvol” will do the rest of the work. Further out we apply the LU decomposition on the leading principal submatrix without pivoting and construct an approximation of M by the cross approximation

(as in the lecture notes). The computational errors of the LU-cross approximation without pivoting of A and B are shown in figure ?? below.

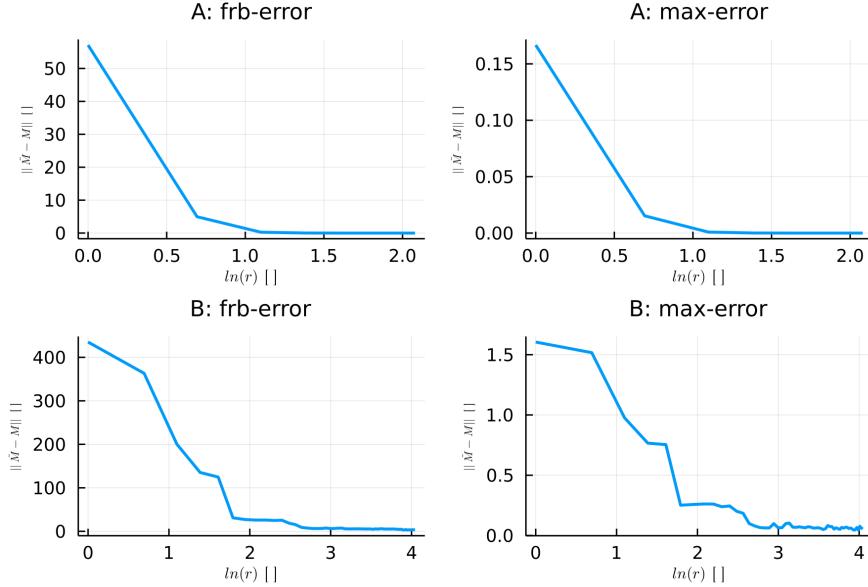


Figure 2: Error of rank r LU-cross approximation of matrices A and B on a logarithmic scale without pivoting

4 Matrix approximation by LU with pivoting

Here we follow the same procedure, where we cross approximate the matrix. Only that now we consider a pivoted LU decomposition of the leading principal submatrix. The computational errors of the LU-cross approximation with pivoting of A and B are shown in figure ?? below.

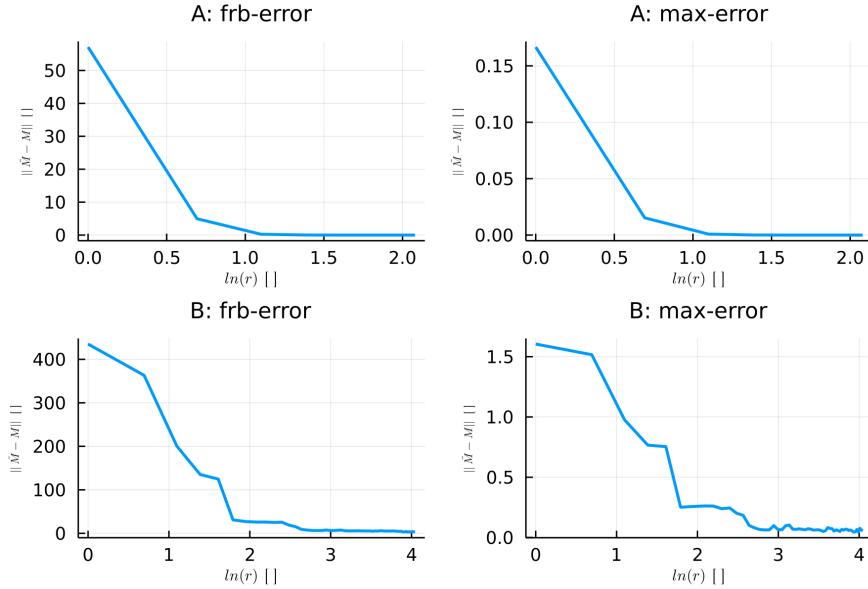


Figure 3: Error of rank r LU-cross approximation of matrices A and B on a logarithmic scale with pivoting

The convergence observed is that the bigger r , the less the error is. By [3], for a matrix A and its according leading principal submatrix A_{11} of a rank exceeding r we have

$$\|\hat{A} - A\|_{F/2} \leq (r+1)^2 \min_{\text{rank}(B) \leq r} \|A - B\|_{F/2}. \quad (10)$$

Every increasing of r in the LU-cross approximation we are “nearer” the original matrix A .

5 Understanding the basis rows and columns

Here we show contour and surface plots of f and g on $[-1, 1]^2$ and additionally color red the axis-parallel lines selected by the pivoted LU algorithm above used for the cross approximation at $r = 1, 2, 3, 4, 5, 10, 15, 20, 30, 40$. Where for the matrix A only lines for up to $\text{rank}(A) = 8$ are drawn in.

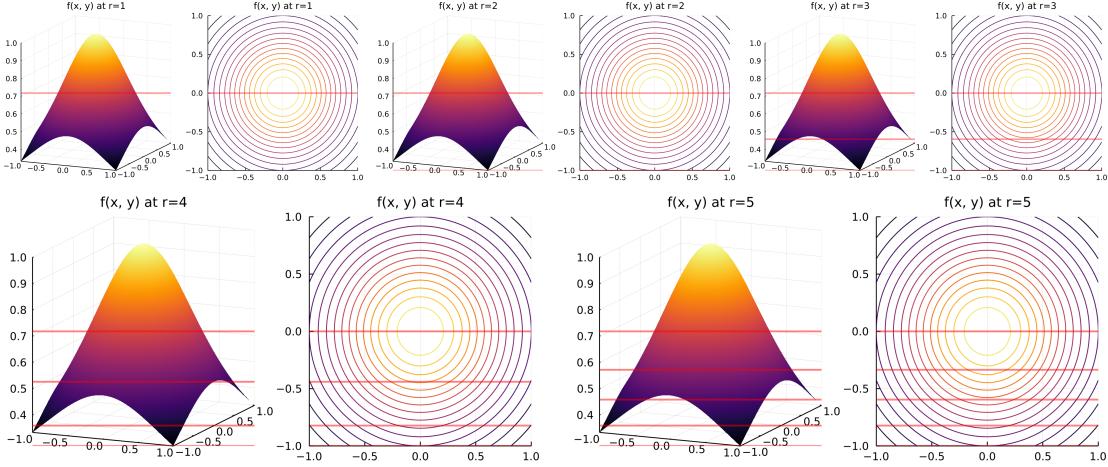


Figure 4: Axis-parallel lines selected by the pivoted LU, of $f(x, y)$

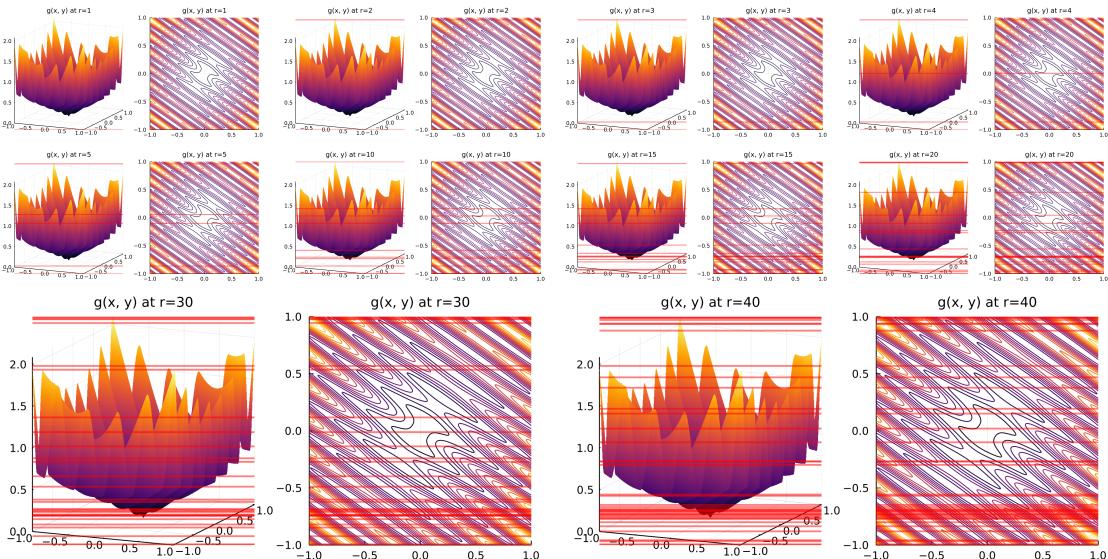


Figure 5: Axis-parallel lines selected by the pivoted LU, of $g(x, y)$

References

- [1] N. Kishore Kumar and Jan Shneider. *Literature survey on low rank approximation of matrices*. 2016. arXiv: [1606.06511 \[math.NA\]](https://arxiv.org/abs/1606.06511).
- [2] JuliaHub. *Maxvol*. URL: <https://juliahub.com/ui/Packages/Maxvol/B0c7T/1.0.0>.
- [3] Sergei Goreinov and E. Tyrtyshnikov. “Quasioptimality of Skeleton Approximation of a Matrix in the Chebyshev Norm”. In: *Doklady Mathematics - DOKL MATH* 83 (June 2011), pp. 374–375. DOI: [10.1134/S1064562411030355](https://doi.org/10.1134/S1064562411030355).
- [4] Popovic Milutin. *Git Instance, Tensor Methods for Data Science an Scientific Computing*. URL: git://popovic.xyz/tensor_methods.git.