MATRIX COMPUTATIONS: HOMEWORK 2, 16 October 2023

In this homework, we study the *polar decomposition* of a matrix $A \in \mathbb{R}^{n\times n}$. As we will see, the polar decomposition has important applications in interpolation on manifolds, and in particular in computer animation. As an illustration, consider two interpolations of an image transformation; one with the "classical" interpolation, and one using the polar decomposition:

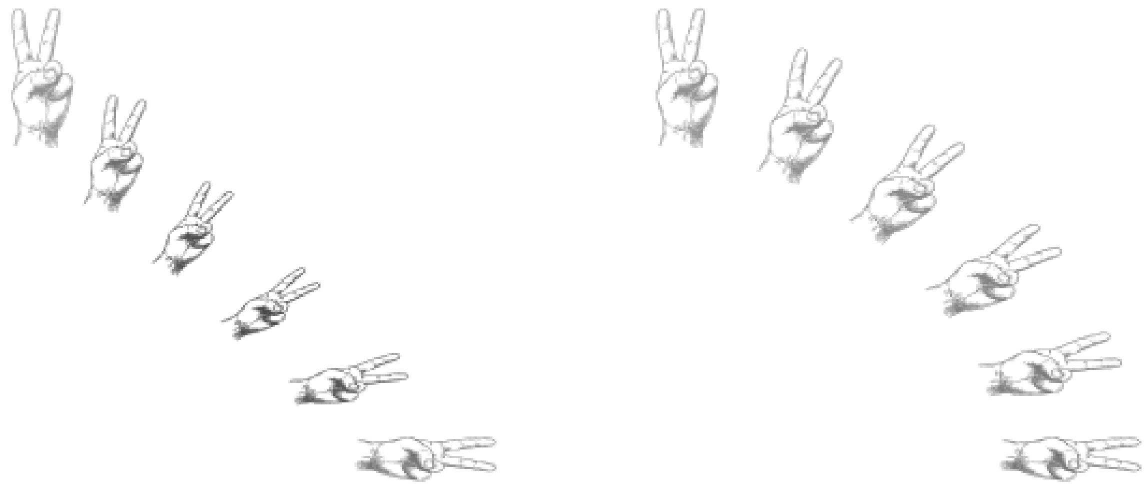

Figure 1: Interpolation of the transformation of a hand image. *Left.* With the classical interpolation technique. *Right.* With the polar decomposition technique.

We remind the definition of the polar decomposition and the existence result.[1].

Definition 1. Let $A \in \mathbb{R}^{n\times n}$. A *polar decomposition* of A is a factorization $A = HU$, wherein $H \in \mathbb{R}^{n\times n}$ is symmetric positive semidefinite and $U \in \mathbb{R}^{n\times n}$ is orthogonal. Formally, we will say that (H, U) is a polar decomposition of A if H and U satisfy the above conditions.

Theorem 1. Every $A \in \mathbb{R}^{n\times n}$ admits at least one polar decomposition.

Exercise A: Properties of the Polar Decomposition

(A1) Let $H \in \mathbb{R}^{n\times n}$ be symmetric positive definite and $U \in \mathbb{R}^{n\times n}$ be orthogonal. Show that $\mathrm{trace}(HU) \leq \mathrm{trace}(H)$. Also, show that $\mathrm{trace}(HU) < \mathrm{trace}(H)$ if $U \neq I$.

Hint. First, prove the result assuming that H is diagonal. Then, derive the general result using the eigenvalue decomposition of H. The following property may be useful as well.

Proposition 1. For all $M, N \in \mathbb{R}^{n\times n}$, it holds that $\mathrm{trace}(MN) = \mathrm{trace}(NM)$.

Theorem 1 guarantees the existence of *a* polar decomposition for every matrix. The following exercise shows that the polar decomposition is actually unique for invertible matrices.

(A2) Assume that A is invertible. Show that the polar decomposition is unique.

Hint. The result proved in A1 may be useful.

Remark 1. Note that the above is in contrast with the Singular Value Decomposition, which may be not unique, even for invertible matrices.

Remark 2. If A is singular, then the polar decomposition may be not unique, as shown by the trivial example $A = [0]$ that admits two polar decompositions: $A = [0] \cdot [-1]$ and $A = [0] \cdot [1]$.

[1] Lecture notes, Theorem 3.21.

(A3) Let $A \in \mathbb{R}^{n\times n}$ and let (H, U) be a polar decomposition of A. Show that A is normal[2] if and only if H and U commute, i.e., if and only if $HU = UH$.

Hint. For the "only if" direction, you may find the following result useful.

Proposition 2. Let $H \in \mathbb{R}^{n\times n}$ be symmetric positive semidefinite and $A \in \mathbb{R}^{n\times n}$. It holds that A commutes with H if and only if it commutes with H^2.

(A4) Let $A \in \mathbb{R}^{n\times n}$ and (H, U) be a polar decomposition of A. Show that U minimizes the Frobenius distance to A among all orthogonal matrices, i.e.,

$$\|A - U\|_F = \min\{\|A - V\|_F : V \in \mathbb{R}^{n\times n},\ V^\top V = I\}$$

wherein $\|M\|_F \doteq \sqrt{\mathrm{trace}(M^\top M)}$ is the Frobenius norm.

Hint. You may find A1 useful.

Exercise B: Algorithm to Compute the Polar Decomposition

We saw in the course that the polar decomposition can be computed from the SVD (Theorem 3.21). However, this might not be the most efficient way to compute it in practice, as the computation of the SVD can be quite expensive. In this section, we will see another method to compute the polar decomposition of an invertible matrix. This method stems from the following iteration that converges to 1 for any initial value $s_0 > 0$:

$$s_{k+1} = \frac{s_k + \frac{1}{s_k}}{2}. \tag{1}$$

Theorem 2. For any value $s_0 > 0$, the iteration (1) satisfies that for all $k \in \mathbb{N}$, $s_k > 0$, and $\lim_{k\to\infty} s_k = 1$.

(B1) Let $S_0 \in \mathbb{R}^{n\times n}$ be a *diagonal and positive definite matrix.* Consider the iteration:

$$S_{k+1} = \frac{S_k + S_k^{-1}}{2}. \tag{2}$$

Show that for all $k \in \mathbb{N}$, S_k is diagonal positive definite, and $\lim_{k\to\infty} \|S_k - I\|_2 = 0$.

Hint. For a rigorous proof of the first part (namely, showing that S_k is diagonal and positive definite), use mathematical induction.

(B2) Let $S_0 \in \mathbb{R}^{n\times n}$ be *symmetric* positive definite. Consider the iteration:

$$S_{k+1} = \frac{S_k + S_k^{-1}}{2}. \tag{3}$$

Show that for all $k \in \mathbb{N}$, S_k is symmetric positive definite, and $\lim_{k\to\infty} \|S_k - I\|_2 = 0$.

Hint. Use the eigenvalue decomposition of symmetric positive definite matrices.

(B3) Let $A \in \mathbb{R}^{n\times n}$ be invertible and let (H, U) be the polar decomposition of A. Let $Q_0 = A$ and consider the iteration:

$$Q_{k+1} = \frac{Q_k + Q_k^{-\top}}{2}. \tag{4}$$

Show that for all $k \in \mathbb{N}$, Q_k is invertible, and $\lim_{k\to\infty} \|Q_k - U\|_2 = 0$.

Hint. Start from the polar decomposition of A, and use B2.

[2] We remind that A is normal if and only if $A^\top A = AA^\top$.

(B4) On your favorite computer program, apply the iteration (4) on the matrix

$$A = \begin{bmatrix} 1 & -1 \\ 1 & -a \end{bmatrix}$$

for $a \in \{2, 1.1, 1.01, 1.001\}$. Stop when $\|Q_k^\top Q_k - I\|_2 \leq 10^{-9}$, indicating that Q_k is close enough to being orthogonal. Plot the value of $\|Q_k^\top Q_k - I\|_2$ for the different values of k, for each a.

Hint. For good readability, use a logarithmic scale for the y-axis of the plot.

Exercise C: Image Animation

Consider the image given in the file `thumb.txt`. Each line represents the position of a dot; see the black dots in Figure 2. On this image, we apply a linear transformation given by the matrix

$$A = \begin{bmatrix} -0.5 & 0.87 \\ -0.61 & -0.35 \end{bmatrix};$$

see the red dots in Figure 2.

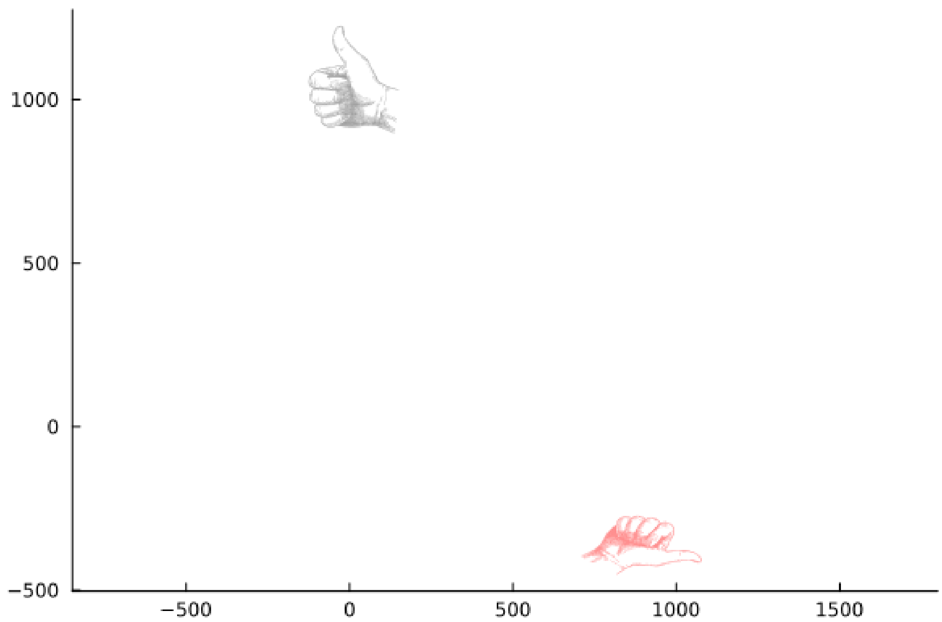

Figure 2: *Black dots.* Original image. *Red dots.* Image after linear transformation.

We want to interpolate the image, i.e., the position of the dots, between the original image and the transformed image. In other words, assume that the identity I is the transformation applied to the dots at time $t = 0$ and A is the transformation applied to the dots at time $t = 1$. We want to find a smooth time-dependent matrix A_t such that $A_0 = I$ and $A_1 = A$ that provides a "meaningful" transformation of the dots between $t = 0$ and $t = 1$. Therefore, we consider two approaches to define A_t: the classical "linear" interpolation, and the interpolation based on the polar decomposition.

(C1) The linear interpolation is defined by $A_t = (1-t)I + tA$. For $t \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$, draw the image after applying the transformation A_t.

(C2) Let (H, U) be the polar decomposition of A. The polar interpolation is defined by $A_t = H_t U_t$, wherein $H_t = (1-t)I + tH$ and $U_t = \exp(t \log(U))$. Note that most programming languages (e.g., Matlab, Python, Julia) include packages to compute the matrix exponential and matrix logarithm. For $t \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$, draw the image after applying the transformation A_t.

Practical information

The homework solution should be written in English.
Please submit a `zip` file on Moodle with your solution in a `pdf` file `Group_XX.pdf` and your code for B4, C1 and C2 in files `Group_XX_QQ.*`.[3]
As you are in master, we strongly recommend you to write your report in latex.
Deadline for turning in the homework: Monday 6 November 2023 (11:59 pm).
It is expected that each group makes the homework individually.
If your group has problems or questions, you are welcome to contact the teaching assistants: julien.calbert@uclouvain.be, guillaume.berger@uclouvain.be.

[3] E.g., `Group_01.pdf`, `Group_02.pdf`, `Group_12.pdf`, `Group_15_B4.py`, `Group_23_C2.jl`. Failure to comply with these requirements may result in point penalties.