Problem set 1

In this section we'll explore why the SVD is very useful.

1. Let $M \in \mathbb{R}^{n\times n}$ and such that $M^\top = M$. Remember that the two norm of such matrix is given by:

$$\|M\|_2 = \max_{x\neq 0} \frac{\|Mx\|_2}{\|x\|_2} = \max_{\|x\|=1} \|Ax\|_2.$$

Let $\lambda_1, \lambda_2, ..., \lambda_n$ be the Eigenvalues of M and such that $|\lambda_1| \geq |\lambda_2| \geq ... \geq |\lambda_n|$. Prove the following:

$$\|M\|_2 = |\lambda_1|.$$

2. Using a similar argument prove that for a general rectangular matrix $A \in \mathbb{R}^{m\times n}$ we have that:

$$\|A\|_2 = \max_i \sqrt{\lambda_i},$$

where λ_i are the Eigenvalues of $A^\top A$. Don't use the SVD here.

3. Consider the SVD of the matrix $A \in \mathbb{R}^{m\times n}$:

$$A = USV^\top.$$

We know that we can't just write A^{-1} because A is a rectangular matrix. But our intuition says that if A is a square matrix and if it's invertible, then we could be able to write its inverse as:

$$A^{-1} = VS^{-1}U^\top.$$

A way to "define an inverse" is with the Moore-Penrose inverse (or pseudo inverse). This pseudo inverse is given by:

$$A^\dagger = (A^\top A)^{-1}A^\top,$$

This is the left pseudo inverse of A, i.e. $A^\dagger A = I$. We can define the right pseudo inverse in a similar matter. Using the SVD prove that indeed, $A^\dagger = VS^{-1}U^\top$.

4. Prove the following:

- If A is a square matrix and if A is invertible then $A^{-1} = A^{\dagger}$.
- Write the least square solution of

$$Ax = b$$

  in terms of $A^{\dagger}$.

5. Now using the SVD, prove using $A = USV^{\top}$ that the condition number of any matrix defined as:

$$\kappa(A) = \|A\|_2 \|A^{\dagger}\|_2$$

is given by:

$$\kappa(A) = \frac{\sigma_1}{\sigma_k},$$

where σ_1 is the largest singular value of A and σ_k is the smallest singular value of A.

6. Suppose that we want to solve the following system:

$$Ax = b. \tag{1}$$

By now we already know that our computer rounds up the values we're working with. Hence let $\hat{b}$ be the rounded version of b, then although our intention is to solve 1 we might end up solving:

$$A\hat{x} = \hat{b}.$$

Suppose that this round up is to precision ϵ:

$$\frac{\|b - \hat{b}\|}{\|b\|} = \epsilon.$$

Derive a bound of the round of of x:

$$\frac{\|x - \hat{x}\|}{\|x\|}.$$

(Hint: SVD)

7. Geometrically, the non reduced version of the SVD, $A = USV^{\top}$, where $U \in \mathbb{R}^{m \times m}$, $S \in \mathbb{R}^{m \times n}$, $V \in \mathbb{R}^{n \times n}$ we have that:

- The columns of U are the Eigenvectors of $AA^\top$.
- The columns of V are the Eigenvectors of $A^\top A$.
- The k singular values on the diagonal of S are the square roots of the nonzero Eigenvalues of both $AA^\top$ and $A^\top A$.

Prove this.

8. The nice part of the SDV is that U and V give orthonormal bases for all four fundamental subspaces:

   - The first k columns of U form an orthonormal base for the **column space** of A, $C(A)$
   - The last $m-k$ columns of U form an orthonormal base for the **left nullspace** of A, $N(A^\top)$
   - The first k columns of V form an orthonormal base for the **row space** of A, $C(A^\top)$
   - The last $n-k$ columns of V form an orthonormal base for the **nullspace** of A, $N(A)$

   Calculate the SVD of:

$$A = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix},$$

   and give bases for $C(A)$, $N(A)$, $C(A^\top)$, $N(A^\top)$.

Problem set 2

A trigonometric polynomial of degree N is a polynomial that has the following form:

$$t(x) = a_0 + \sum_{k=1}^{N} a_k \cos(kx) + b_k \sin(kx). \tag{2}$$

The inner product in the space is:

$$\langle t_1, t_2 \rangle = \int_{-\pi}^{\pi} t_1(x) t_2(x) dx.$$

And thus the norm we are going to define here is:

$$\|t\| = \int_{-\pi}^{\pi} (t(x))^2 dx.$$

1. Prove that $\mathcal{B}$ is a basis for $\mathcal{T}$, the space of trigonometric polynomials of degree at most N defined in $[-\pi, \pi]$:

$$\mathcal{B} = \{1, \cos(x), \sin(x), \cos(2x), \sin(2x), ..., \cos(Nx), \sin(Nx)\},$$

conclude that T is a space of dimension $2N + 1$.

2. Find an orthonormal basis of $\mathcal{T}$ using G-S in $\mathcal{B}$.

3. For any trigonometric polynomial of the form 2 find the coefficients a_0, a_k and b_k given the orthonormal basis you previously found. These coefficients are commonly known as the Fourier coefficients.

Coding part

Since this is just quick examples on why the SVD is important, feel free to use Python libraries to compute the necessary matrix decomposition.

Naive image compression

A black-and-white image can be identified with a rectangular matrix A (if the image is rectangular of course). Each of its entries a_{ij} belong to $[0, 1]$. Suppose that the size of the image is very large so we can't store it reasonably. The SVD can be used to minimize the storage size of the image by replacing it by an approximation that is "visually equivalent". This is because the SVD can be seen as:

$$A = USV^\top = \sum_{i=1}^{k} \sigma_i u_i v_i^\top,$$

where u_i, v_i are the columns of U and V respectively, σ_i are the singular values of A and k is the range of A.

then we can use the r range approximation of A:

$$A \approx \sum_{i=1}^{r} \sigma_i u_i v_i^\top.$$

Download the image "imageComp.jpg" and output the compression for $r = 5, 10, 15, 20, 25, 50, 100, 200$.