MA 402: Project 2

Instructions:

- Detailed instructions regarding submission are available on the class website[1].
- The zip file should contain five files hw2.pdf, hw2.tex, classnotes.sty, swift.mat, and deblur.mat.
- More instructions:
  - MATLAB users: use `loadmat` (type `who` to display what variables are in your workspace.
  - Python users: use `scipy.io.loadmat`. This will return a dictionary with all the necessary variables.
- For plotting, you may consider using `imshow`.

1 Pen-and-paper exercises

The problems from this section total 20 points.

1) (10 points) Consider the matrix A with the SVD

$$A = \begin{bmatrix} 4 & 0 \\ -5 & -3 \\ 2 & 6 \end{bmatrix} = U \begin{bmatrix} 6\sqrt{2} & 0 \\ 0 & 3\sqrt{2} \\ 0 & 0 \end{bmatrix} V^\top,$$

where

$$U = \frac{1}{3}\begin{bmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix} \qquad V = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

(a) (0 points) Verify for yourself that it is indeed the SVD of A, and that U, V are orthogonal.

(b) (1 point) What is the rank of this matrix?

(c) (2 points) From the full SVD of A, write down the thin SVD of A.

(d) (3 points) Compute the best rank-1 approximation of A.

(e) (2 points) Compute the 2-norm and the Frobenius norms of A.

(f) (2 points) Using the SVD of A, write down the SVD of $A^\top$ and $A^\top A$.

2) (10 points) Let $A \in \mathbb{R}^{m\times n}$. Recall: by definition, the Frobenius norm of A is $\|A\|_F = \left(\sum_i \sum_j |a_{ij}|^2\right)^{1/2}$. In this problem, we will derive the formula

$$\|A\|_F = \left(\sigma_1^2 + \cdots + \sigma_r^2\right)^{1/2}.$$

[1] https://github.ncsu.edu/asaibab/ma402/blob/master/project.md

(a) The trace of a square matrix is the sum of its diagonals entries. Show (an alternative representation for the Frobenius norm):

$$\|A\|_F = \left(\mathsf{trace}\,(A^\top A)\right)^{1/2}.$$

(b) Let C, D be $n \times n$ square matrices. Show: $\mathsf{trace}\,(CD) = \mathsf{trace}\,(DC)$.
Remark: This is known as the cyclic property of trace, which is true despite the fact that in general $CD \neq DC$. A consequence of the cyclic property is: if E has the same size as C, D, it implies

$$\mathsf{trace}\,(CDE) = \mathsf{trace}\,(DEC) = \mathsf{trace}\,(ECD).$$

(c) Using parts (a-c) complete the proof to show $\|A\|_F = \left(\sigma_1^2 + \cdots + \sigma_r^2\right)^{1/2}$.

(d) Show: $\|A\|_2 \leq \|A\|_F \leq \sqrt{r}\|A\|_2$.

2 Numerical exercises

The problems from this section total 30 points.

3) (15 points) *Compressing* and *Denoising* images.

(a) (0 points) Load the file 'swift.mat'. You will find the variables `A` and `An` which are both matrices of size 512×1024.

(b) (2 points) In a single figure with 2 subplots, plot the clean as well as the noisy matrices as images (use `imshow`). Denote the corresponding matrices as A and $\tilde{A} = A + E$, where E is the amount of noise added to the original image. Unfortunately, in real applications we do not know exactly how much noise is added.

(c) (2 points) Plot the first 100 singular values of A and $\tilde{A}$. (*Hint:* Use the `semilogy` plotting function).

(d) (3 points) In a single figure with 9 different subplots, plot A_k (the best rank-k approximation to A) as images for $k = 5, 10, \ldots, 45$ (use these same values of k for the rest of this problem).

(e) (2 points) As two subplots of the same figure, plot (left panel) the storage cost of the truncated SVD as a function of k, (right panel) relative error of A_k (in the Frobenius norm) as a function of k. Comment on these two subplots. (Assume that each floating point number requires 1 unit of storage.)

(f) (3 points) Our proposed algorithm to denoise the image is to use a truncated SVD of the matrix corresponding to the noisy image, i.e., computing $\tilde{A}_k$. In a single figure with 9 different subplots, plot $\tilde{A}_k$ (the best rank-k approximation to A) as images for $k = 5, 10, \ldots, 45$. Make sure to label each subplot.

(g) (2 points) Plot the relative error of the denoised image $\tilde{A}_k$ (in the Frobenius norm) as a function of the truncation index k. For (approximately) what value of k is the minimum attained?

(h) (2 bonus points) A result in perturbation analysis (due to Herman Weyl) says

$$\max_{1 \leq j \leq \min\{m,n\}} |\sigma_j(A + E) - \sigma_j(A)| \leq \|E\|_2.$$

In the context of the noisy images, give an interpretation of the above equation in your words. Based on this formula, can you obtain a lower bound for the amount of noise, measured as $\|E\|_2$?

Instructions: In total, you have to submit 6 separate plots. Make sure to label each plot/subplot, and label the axes of the singular value and the error plots.

4) (15 points) *Deblurring* an image.

(a) (0 points) Load the file 'deblur.mat'. You will find the variables `A` (blurring operator, size 4096×4096) and `bn` (blurred and noisy image, size 4096×1), `xtrue` (true image, size 4096×1).

(b) (2 points) In a single figure with 2 subplots, plot the true image, and the blurry image with noise. Note that you will have to reshape the vectors into 64×64 images.

(c) (2 points) Recall the naive solution $x_n = A^{-1} b_n$. Plot this solution as an image. (MATLAB users should look up backslash \, and Python users should look up `numpy.linalg.solve`. Do not compute the inverse of the matrix!)

(d) (3 points) Compute the condition number $\kappa_2(A) = \|A\|_2 \|A^{-1}\|_2$ of the matrix A. Using perturbation analysis explain why you expect the naive solution to perform poorly (you are given that $\|e\|_2/\|b\|_2 = 0.05$).

(e) (3 points) Implement the truncated SVD formula

$$x_k = \sum_{j=1}^{k} v_j \frac{u_j^\top b_n}{\sigma_j},$$

for $k = 400, 800, \ldots, 3600$. In a single figure with 9 subplots, plot the reconstructed vectors x_k as images.

(f) (3 points) Plot the relative error in the reconstructed solution as a function of k. For (approximately) what value of k is the minimum attained?

(g) (2 points) In your words, explain the behavior of the error as a function of k.

Instructions: In total, you have to submit 4 separate plots. Make sure to label each plot/subplot, and label the axes of the error plots.