Math 123 **Problem Set 2** **Due: Thursday 2/25 at 11:59 p.m. EST**

Instruction: Read the homework policy. For problems 4 and 6, include printed copies of your code with your final homework submission. A simple demo MATLAB code to load images can be found in the HW2 folder on Canvas. You should submit a PDF copy of the homework and any associated codes on Canvas. Your PDF must be a single file, not multiple images.

1. Consider the following 2 by 3 matrix, $\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix}$.

(a) Find the right singular vectors i.e. eigenvectors of $\mathbf{A}^T\mathbf{A}$. Let $\mathbf{V}$ denote the matrix of the normalized eigenvectors.

(b) Find the singular values of $\mathbf{A}$. Let $\mathbf{\Sigma}$ be a diagonal matrix that consists the singular values in the main diagonal.

(c) Find the left singular vectors of $\mathbf{A}$. Let $\mathbf{U}$ denote the matrix of the normalized left singular vectors.

(d) Write the singular value decomposition (SVD) of $\mathbf{A}$ i.e. show that $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$.

(e) Write the reduced SVD of $\mathbf{A}$.

(f) Using the SVD of $\mathbf{A}$, compute $||\mathbf{A}||_2$ and $||\mathbf{A}||_F$.

[**Remark**: Show detailed work. You can certainly check your result numerically but no credit will be given to answers that rely on the results of numerical solvers].

2. Consider the matrix $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 20 & 0 & 0 \\ 1 & 0 & 0 & 7 \end{bmatrix}$.

(a) Find the singular decomposition of $\mathbf{A}$.

(b) Express the matrix $\mathbf{A}$ as a sum of rank 1 outer products.

(c) What is the rank of $\mathbf{A}$? Justify your answer using the SVD of $\mathbf{A}$.

(d) Let $\hat{\mathbf{A}}_1$ be the best rank-1 approximation of $\mathbf{A}$. Give an explicit form of $\hat{\mathbf{A}}_1$.

(e) Let $\hat{\mathbf{A}}_2$ be the best rank-2 approximation of $\mathbf{A}$. Give an explicit form of $\hat{\mathbf{A}}_2$.

(f) Compute $||\hat{\mathbf{A}}_1 - \mathbf{A}||_F$ and $||\hat{\mathbf{A}}_1 - \mathbf{A}||_2$.

(g) Compute $||\hat{\mathbf{A}}_2 - \mathbf{A}||_F$ and $||\hat{\mathbf{A}}_2 - \mathbf{A}||_2$.

[**Remark**: To compute the singular value decomposition in (a), you can use any solver in MATLAB or Python].

3. The squared Frobenius norm of the m by n matrix $\mathbf{A}$ is defined as $||\mathbf{A}||_F^2 = \sum_{i,j} a_{i,j}^2$, where $a_{i,j}$ is the (i,j)-th entry of $\mathbf{A}$. More compactly, $||\mathbf{A}||_F^2 = \text{trace}(\mathbf{A}^T\mathbf{A})$. trace denotes the trace of a matrix defined as the sum of the diagonal values of matrix. Let $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ denote the singular value decomposition of the m by n matrix $\mathbf{A}$ of rank r.

(a) For any m by n matrix $\mathbf{X}$, prove that $||\mathbf{UX}||_F^2 = ||\mathbf{X}||_F^2$. [**Hint**: Use the definition of the Frobenius norm in terms of the trace product.]

(b) For any m by n matrix $\mathbf{Z}$, prove that $||\mathbf{ZV}||_F^2 = ||\mathbf{Z}||_F^2$. [**Hint**: Use the definition of the Frobenius norm in terms of the trace product and the cyclic property of the trace i.e. $\text{trace}(\mathbf{ABC}) = \text{trace}(\mathbf{BCA}) = \text{trace}(\mathbf{CAB})$.]

(c) Using the results in (a) and (b), prove that $||\mathbf{A}||_F^2 = ||\mathbf{\Sigma}||_F^2$.

(d) Using your result in (c), prove that $||\mathbf{A}||_F^2 = \sum_{i=1}^{r} \sigma_i^2$.

4. Load the matrix `A_test` from the HW2 folder on Canvas. Let $\mathbf{A}$ denote this matrix.

(a) Implement the power method and find the first right singular vector $\mathbf{v}_1$ and corresponcing largest singular value σ_1 of $\mathbf{A}$.

(b) Check if your result in (a) agrees with SVD of $\mathbf{A}$ obtained from a solver of your interest e.g. in MATLAB or Python.

(c) Let $\mathbf{A}$ be a generic m by n matrix with $m \ll n$. What is the computational complexity of $\mathbf{Bz}$ where $\mathbf{B} = \mathbf{A}^T\mathbf{A}$ and $\mathbf{z}$ is an n-dimensional vector? What is the computational complexity of $\mathbf{A}^T(\mathbf{Az})$ i.e. computing $\mathbf{Bz}$ in two steps by first applying $\mathbf{A}$ and then $\mathbf{A}^T$? Briefly discuss the implication of this result. [**Remark**: The cost of the elementary operations is assumed to be 1. For instance, the sum of two numbers a and b costs 1 i.e. the computational complexity is 1. The dot product of two n-dimensional vectors costs $2n - 1$ since we require n multiplications and $n - 1$ additions.]

[**Note on (c)**: You could ignore the cost of transposing and assume that $\mathbf{B} = \mathbf{A}^T\mathbf{A}$ is given i.e. the focus is on the cost of $\mathbf{Bz}$ as compared to the cost of $\mathbf{A}^T(\mathbf{Az})$.]

5.[**Bonus**: 5 pts] Let S be the set of matrices defined as follows

$$S = \left\{ \mathbf{A} | \mathbf{A} = \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & d \end{pmatrix}, a > 0, b > 0 \right\}.$$

Give an example of a matrix $\mathbf{X}$ in S with satisfying all the following conditions: (1) $\mathbf{X}$ has two non-zero singular values σ_1 and σ_2, (2) $\sigma_1^2 + \sigma_2^2 = \frac{37}{4}$ and (3) $||\mathbf{X}||_2 = 3$.

6. In this problem, we apply the singular value decomposition to the image compression problem. We consider the image of the panises flower shown in Figure 1. In mathematical terms, the image is a a 1528×1225 matrix where each entry has a value from 0 to 255. The value 0 indicates black and the value 255 indicates white. The values in between represent the different intensities of gray.

(a) In a programming language of your choice, load the image and store it as a matrix $\mathbf{A}$ of size 1528×1225. Compute the singular decomposition of $\boldsymbol{A}$. To compute the singular decomposition of $\boldsymbol{A}$, you can use any existing SVD algorithm. For example, in MATLAB, you can compute the SVD using `[U,S,V] = svd(A)`.

Figure 1: panises are not too happy about being compressed

(b) Using your result in (a), what is the rank of the matrix $\mathbf{A}$?

(c) Order the singular values in decreasing order and plot the values. What could you say about the decay of singular values?

(d) Compute the truncated singular value decomposition of $\mathbf{A}$ by considering only the first $k = 10$ largest singular values. Let $\hat{\mathbf{A}}$ denote the resulting matrix. Show the resulting matrix $\hat{\mathbf{A}}$ as an image. Describe the quality of the image.

(e) Repeat the experiment in (d) by considering the following cases: (a) $k = 25$, (b) $k = 50$, (c) $k = 100$. Describe its quality of the images for each case. The image refers to the visualization of the matrix $\hat{\mathbf{A}}$.

(f) In practice, it is important to identify a suitable truncation parameter k. Using the Eckart-Young theorem, propose a systematic way to choose k. What value do you obtain for this data?

(g) Assume that storing a single pixel costs 1 byte of memory. How many bytes do we need to store a general $m \times n$ matrix $\mathbf{A}$? Assume that $\mathbf{A}$ is rank r. How many bytes do we need to store the reduced SVD of $\mathbf{A}$? If we do truncated SVD with $k \ll r$, how many bytes do we need to store the truncated SVD? Apply your results to the panises image matrix and discuss the computational advantages of working with a truncated SVD.

1 $A = \begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix}$

(a) find the right singular vectors of A^TA

$$A^TA = \begin{bmatrix} 1 & -2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 5 & -2 & 1 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$A^TA - \lambda \mathbb{I} = \begin{bmatrix} 5-\lambda & -2 & 1 \\ -2 & 1-\lambda & 0 \\ 1 & 0 & 1-\lambda \end{bmatrix} \Rightarrow \det =$$

$$(5-\lambda)(1-\lambda)^2 -$$

$$-2(-2(1-\lambda))$$

$$+ -(1-\lambda)$$

$$= (5-\lambda)(1-\lambda)^2 - 5(1-\lambda)$$

$$= (1-\lambda)((5-\lambda)(1-\lambda) - 5)$$

$$= (1-\lambda)(5 - 6\lambda + \lambda^2 - 5)$$

$$= (1-\lambda)(\lambda^2 - 6\lambda)$$

$$= \lambda(1-\lambda)(\lambda - 6)$$

so $\lambda = 6$, $\lambda = 1$, $\lambda = 0$ are the eigenvalues

$$A^TA - 6\mathbb{I} = \begin{bmatrix} -1 & -2 & 1 \\ -2 & -5 & 0 \\ 1 & 0 & -5 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & -2 \\ 0 & -2 & -4 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{matrix} x_1 + 2x_2 - x_3 = 0 \\ x_2 + 2x_3 = 0 \end{matrix} \Rightarrow \begin{matrix} x_1 = 5x_3 \\ x_2 = -2x_3 \end{matrix} \qquad \sqrt{5^2+2^2+1}$$

$$\Rightarrow \text{Null}(A^TA - 6\mathbb{I}) = \left\{ x_3 \begin{pmatrix} 5 \\ -2 \\ 1 \end{pmatrix} \right\} \Rightarrow \text{normalize to } \begin{pmatrix} 5/\sqrt{30} \\ -2/\sqrt{30} \\ 1/\sqrt{30} \end{pmatrix}$$

$$A^TA - \mathbb{I} = \begin{bmatrix} 4 & -2 & 1 \\ -2 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow x_1 = 0 \quad x_2 = \tfrac{1}{2}x_3 \Rightarrow \text{Null}(A^TA - \mathbb{I}) = \left\{ x_3 \begin{pmatrix} 0 \\ 1/2 \\ 1 \end{pmatrix} \right\}$$

$\Downarrow$ normalize $\Downarrow$ $\begin{pmatrix} 0 \\ 1/2\sqrt{1.25} \\ 1/\sqrt{1.25} \end{pmatrix}$

$$A^TA = \begin{bmatrix} 5 & -2 & 1 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & -2 & -4 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow \text{Null}(A^TA) = \left\{ x_3 \begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix} \right\} \Rightarrow \text{normalize} \Rightarrow \begin{pmatrix} -1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix}$$

$$V = \begin{bmatrix} 5/\sqrt{30} & 0 & -1/\sqrt{6} \\ -2/\sqrt{30} & 1/2\sqrt{1.25} & -2/\sqrt{6} \\ 1/\sqrt{30} & 1/\sqrt{1.25} & 1/\sqrt{6} \end{bmatrix} \qquad \Lambda = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

↑ These are the right singular vectors!

(b) find the singular values of A

by the work in (a), $\sigma_i = \sqrt{\lambda_i}$ for each λ_i eigenvalue of $A^T A$

$\Rightarrow \sigma_1 = \sqrt{6} \quad \sigma_2 = 1 \quad \sigma_3 = 0$

$$\Rightarrow \Sigma = \begin{bmatrix} \sqrt{6} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

(c) find the left singular vectors of A

$$U = AV\Sigma^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} .9129 & 0 & -.4082 \\ -.3651 & .4472 & -.8165 \\ .1826 & .8944 & .4082 \end{bmatrix} \begin{bmatrix} \sqrt{6} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 6/\sqrt{30} & 1/\sqrt{5/4} & 0 \\ -12/\sqrt{30} & 1/2\sqrt{5/4} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 6/\sqrt{30}\sqrt{6} & 1/\sqrt{5/4} & 0 \\ -12/\sqrt{30}\sqrt{6} & 1/2\sqrt{5/4} & 0 \end{bmatrix}$$ = left singular vectors of A!

(d) Compute the SVD of A

based on (a), (b), (c) we have

$$A = U \Sigma V^T$$

and to verify:

$$= \begin{bmatrix} 6/\sqrt{6}\sqrt{30} & 1/\sqrt{1.25} & 0 \\ -12/\sqrt{6}\sqrt{30} & 1/2\sqrt{1.25} & 0 \end{bmatrix} \begin{bmatrix} \sqrt{6} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 5/\sqrt{30} & -2/\sqrt{30} & 1/\sqrt{30} \\ 0 & 1/2\sqrt{1.25} & 1/\sqrt{1.25} \\ -1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \end{bmatrix}$$

$$= \begin{bmatrix} 6/\sqrt{30} & 1/\sqrt{1.25} & 0 \\ -12/\sqrt{30} & 1/2\sqrt{1.25} & 0 \end{bmatrix} \begin{bmatrix} 5/\sqrt{30} & -2/\sqrt{30} & 1/\sqrt{30} \\ 0 & 1/2\sqrt{1.25} & 1/\sqrt{1.25} \\ -1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \end{bmatrix}$$

$$= \left[\begin{array}{c|c|c} 30/\sqrt{30}^2 & -12/\sqrt{30}^2 + 1/2\sqrt{1.25}^2 & 6/\sqrt{30}^2 + 1/\sqrt{1.25}^2 \\ -60/\sqrt{30}^2 & 24/\sqrt{30}^2 + \frac{1}{(2\sqrt{1.25})^2} & -12/\sqrt{30}^2 + 1/2\sqrt{1.25}^2 \end{array}\right]$$

$$= \left[\begin{array}{c|c|c} 1 & -12/30 + 4/10 & 6/30 + 4/5 \\ -2 & 24/30 + 1/5 & -12/30 + 4/10 \end{array}\right]$$

$$= \begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} = A$$

(e) the reduced SVD of A is

$$\underbrace{\begin{bmatrix} 6/\sqrt{6}\sqrt{30} & 1/\sqrt{1.25} \\ -12/\sqrt{6}\sqrt{30} & 1/2\sqrt{1.25} \end{bmatrix}}_{U_r} \underbrace{\begin{bmatrix} \sqrt{6} & 0 \\ 0 & 1 \end{bmatrix}}_{\Sigma_r} \underbrace{\begin{bmatrix} 5/\sqrt{30} & -2/\sqrt{30} & 1/\sqrt{30} \\ 0 & 1/2\sqrt{1.25} & 1/\sqrt{1.25} \end{bmatrix}}_{V_r^T}$$

(f) Compute $\|A\|_2$ and $\|A\|_F$

$$\|A\|_2 = \left(\sum \sigma_i^2\right)^{1/2} = \left(\sqrt{6}^2 + 1^2\right)^{1/2} = \sqrt{7}$$

$$\|A\|_F = \sqrt{\text{trace}(A^T A)} = \sqrt{7}$$

2

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 20 & 0 & 0 \\ 1 & 0 & 0 & 7 \end{bmatrix}$$

(a) find the SVD of A (using MatLab)

[u, s, v] = SVD(A)

$$\Longrightarrow \underbrace{\begin{bmatrix} 0 & .0204 & -.9998 \\ -1 & 0 & 0 \\ 0 & .9998 & .0204 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 20 & 0 & 0 & 0 \\ 0 & 7.0725 & 0 & 0 \\ 0 & 0 & .9897 & 0 \end{bmatrix}}_{S} \underbrace{\begin{bmatrix} 0 & -1 & 0 & 0 \\ .1442 & 0 & 0 & .9895 \\ -.9895 & 0 & 0 & .1442 \\ 0 & 0 & -1 & 0 \end{bmatrix}}_{V^T}$$

(b) Express A as a sum of rank 1 outer products.

$$A = \sum_{i=1} \sigma_i u_i v_i^T$$

$$= 20 \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} [0\ {-1}\ 0\ 0] + 7.0725 \begin{pmatrix} .0204 \\ 0 \\ .9998 \end{pmatrix} [.1442\ 0\ 0\ .9895]$$

$$+ .9897 \begin{pmatrix} -.9998 \\ 0 \\ .0204 \end{pmatrix} [-.9895\ 0\ 0\ .1442]$$

(c) What is the rank of A?

rank(A) = 3 because A has 3 non-zero singular values, corresponding to A being a sum of 3 rank 1 outer products.

(d) Let $\hat{A}_1$ be the best rank 1 approximation of A

then from our proof in class we know $\hat{A}_1$ is the truncated SVD at $k=1$

hence $\hat{A}_1 = U_1 \sigma_1 V_1^T = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} [20] [0 \ -1 \ 0 \ 0]$

$$= \begin{pmatrix} 0 \\ -20 \\ 0 \end{pmatrix} [0 \ -1 \ 0 \ 0]$$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 20 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

(e) Similarly,

$$\hat{A}_2 = \begin{bmatrix} 0 & .0204 \\ -1 & 0 \\ 0 & .9998 \end{bmatrix} \begin{bmatrix} 20 & 0 \\ 0 & 7.0725 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 & 0 \\ .1442 & 0 & 0 & .9895 \end{bmatrix}$$

$$= \begin{bmatrix} .0208 & 0 & 0 & .1428 \\ 0 & 20 & 0 & 0 \\ 1.0197 & 0 & 0 & 6.9968 \end{bmatrix}$$

$$\hat{A}_2 = 20 \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} [0 \ -1 \ 0 \ 0] + 7.0725 \begin{pmatrix} .0204 \\ 0 \\ .9998 \end{pmatrix} [.1442 \ 0 \ 0 \ .9895]$$

(f) Compute $\|\hat{A}_1 - A\|_F$ and $\|\hat{A}_1 - A\|_2$

$$\|A - \hat{A}_1\|_F = \|A - \hat{A}_1\|_2 = \sqrt{(7.0725)^2 + (.9897)^2}$$

$$= 7.1414$$

(g) Compute $\|\hat{A}_2 - A\|_F$ and $\|\hat{A}_2 - A\|_2$

$$\|A - \hat{A}_2\|_F = \|A - \hat{A}_2\|_2 = \sqrt{(.9897)^2}$$

$$= .9897$$

3 $\|A\|_F^2 = \text{trace}(A^T A)$ $\quad A = U \Sigma V^T$

SVD of A

(a) for any $m \times n$ matrix X prove $\|UX\|_F^2 = \|X\|_F^2$

$$\|UX\|_F^2 = \text{trace}\left((UX)^T (UX)\right)$$

$$= \text{trace}(X^T U^T U X)$$ since U is orthonormal

$$= \text{trace}(X^T X)$$

$$= \|X\|_F^2$$

(b) for any $m \times n$ matrix Z prove $\|ZV\|_F^2 = \|Z\|_F^2$

$$\|ZV\|_F^2 = \text{trace}\left((ZV)^T ZV\right)$$

$$= \text{trace}(\underbrace{V^T}_{a} \underbrace{Z^T}_{b} \underbrace{ZV}_{c})$$ since trace(abc) = trace(bca)

$$= \text{trace}(Z V V^T Z^T)$$

$$= \text{trace}(Z Z^T)$$ since V is orthonormal

$$= \text{trace}(Z^T Z)$$

$$= \|Z\|_F^2$$

(c) prove $\|A\|_F^2 = \|\Sigma\|_F^2$

$$\|A\|_F^2 = \|U\Sigma V^T\|_F^2 = \|\Sigma V^T\|_F^2 = \|\Sigma V^T V\|_F^2$$

$$= \|\Sigma\|_F^2$$

(d) prove $\|A\|_F^2 = \sum_i^r \sigma_i^2$

$$\|A\|_F^2 = \|\Sigma\|_F^2 = \mathrm{trace}(\Sigma^T\Sigma) = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_r^2$$

$$= \sum_i^r \sigma_i^2$$

4

(a) see Matlab implementation

(b) 〃

(c) let A be an $m \times n$ matrix w/ $m << n$

What is the computational complexity of Bz where $B = A^T A$ and $z \in \mathbb{R}^n$

$$A = \begin{array}{cc} & n \\ m & \begin{bmatrix} - & v_1 & - \\ & \vdots & \\ - & v_m & - \end{bmatrix} \end{array} \qquad A^T = \begin{array}{cc} & m \\ n & \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} \end{array}$$

$$z = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$$

$$A^T A = \begin{array}{cc} & n \\ n & \begin{bmatrix} \bullet & & \\ & & \\ & & \end{bmatrix} \end{array}$$

$2m-1$ calculations

$n^2(2m-1)$ total for the matrix transpose

then $(2n-1)n$ more operations to multiply by the vector z

but we ignore the cost of transposing

if we did it in two steps,

$Az \rightarrow m(2n-1)$ steps is a $\mathbb{R}^m$ vector

$A^T(Az) \rightarrow n(2m-1)$ steps

$\Longrightarrow$

to compare:

$$\cancel{n^2(2m-1)} + n(2n-1) \quad \text{vs} \quad m(2n-1) + n(2m-1)$$

ignore this

$$= 2n^2 - n$$

$$= 2mn - m + 2mn - n$$

$$= 4mn - m - n \text{ steps}$$

since n is assumed to be much larger than m in this case, it seems like the second method, doind Az then $A^T(Az)$ is much faster, because n^2 will dominate mn for much larger n.

5 (Bonus)

let $S = \left\{ A = \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & d \end{pmatrix},\ a>0,\ b>0 \right\}$

Find $X \in S$ such that

1) X has two singular values $\sigma_1, \sigma_2 > 0$

2) $\sigma_1^2 + \sigma_2^2 = \frac{37}{4}$

3) $\|X\|_2 = 3$

I think this is impossible since

$$\|X\|_2 = \sum_{i=1}^{2} \sigma_i^2 = \frac{37}{4} \neq 3$$

6

(a) see Matlab implementation

(b) the rank of A = rank of S = 1225

(c) looking at the plot in Matlab, its clear the singular values decay exponentially

(d) after truncating the SVD of A at $k=10$, the resulting image is barely anything at all, no resemblance.

(e) the image increases in quality with higher k values.

(f) Systematic way to choose k:

- decide how much loss is acceptable, set it to ε
- start with $k=1$ and compute $\|A-[A]_k\|_F$
- check that $\|A-[A]_k\|_F < \varepsilon$, if not $k += 1$

for this data, setting tolerance to 1.2×10^3 gives $k \approx 100$

(g) 1 pixel = 1 byte

$m \times n$ matrix = $m \times n$ bytes memory

reduced SVD is $m \times r + r + r \times n$

truncated SVD is $m \times K + K + r \times K$