ELEC 501 | **Serkan Gugercin**
Approximation Theory | September 11, 2002

Homework 1

Due Date: **September 27, 2002**

Problem [1]. Find the singular value decompositions and compute the rank one approximations, which are optimal in the 2-norm, of the matrices listed below. Use paper and pencil.

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 6 & 3 \\ -1 & 2 \end{pmatrix}.$$

Problem [2]. **(a)** Let $A \in \mathbb{R}^{n\times n}$ and $\det A \neq 0$. What is the relationship between the singular values of A and A^{-1}?

(b) Let $\lambda \in \mathbb{R}$ be an eigenvalue of A. Show that $\sigma_n(A) \leq |\lambda| \leq \sigma_1(A)$.

(c) Find the singular values of $A = \begin{pmatrix} p & -q \\ q & p \end{pmatrix}$. Explain your answer geometrically. Find $u \in \mathbb{R}^2$, $||u||_2 = 1$, such that $||Au||_2 = \sigma_1$. Explain.

Problem [3]. **(a)** Show by direct computation that the **Least Squares** solution of $Ax = b$ is given by

$$x_{LS} = (A^TA)^{-1}A^Tb, \qquad A \in \mathbb{R}^{n\times m},\ n \geq m,\ \text{rank}A = m$$

(b) Using the **SVD** of A show that

$$x_{LS} = \sum_{i=1}^{m} \frac{u_i^* b}{\sigma_i} v_i, \qquad \text{and} \qquad \|Ax_{LS} - b\|^2 = \sum_{i=m+1}^{n} (u_i^* b)^2.$$

What is the geometrical interpretation of this problem? Do Ax_{LS} and $b - Ax_{LS}$ lie in the $\text{Im}(A)$, $\text{Ker}(A)$ or in a space perpendicular to $\text{Im}(A)$? Justify your results.

Problem [4]. **(a)** Prove that if $A \in \mathbb{R}^{m\times n}$, then there exists a unit 2-norm vector $z \in \mathbb{R}^n$ such that $A^TAz = \sigma_1^2 z$ where $\sigma_1 = \|A\|_2$.

(b) Using the above result show that if $A \in \mathbb{R}^{m\times n}$, then $\|A\|_2 \leq \sqrt{\|A\|_1\|A\|_\infty}$.

(c) Let $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^n$. Show that that if $E = uv^T$ and $v^Tu = 1$, then $\|E\|_F = \|E\|_2 = \|u\|_2\|v\|_2$.

Problem [5]. **(a)** For any arbitrary matrix A, let E be a matrix such that $\|E\|_2 < \sigma_{min}(A))$. Then prove that $rank(A+E) \geq rank(A)$.

(b) Let $\|\cdot\|$ be a vector norm on $\mathbb{R}^m$ and assume $A \in \mathbb{R}^{m\times n}$. Show that if rank$(A)$=$n$, then $\|x\|_A$=$\|Ax\|$ is a vector norm on $\mathbb{R}^n$.

Problem [6]. (a) Show that the Frobenius norm, i.e. the Schatten 2-norm,

$$\|A\|_F = \left(\sum_{i=1}^{n} \sigma_i^2(A) \right)^{\frac{1}{2}} \quad \text{where} \quad A \in \mathbb{R}^{n\times m}, \quad n \leq m$$

is unitary invariant, that is $\|A\|_F = \|UAV\|_F$ where $U \in \mathbb{R}^{n\times n}$ and $V \in \mathbb{R}^{m\times m}$ with $UU^T = I_n$ and $VV^T = I_m$. Also, prove that

$$\|A\|_F^2 = \text{trace}\,(A^T A) = \sum_{i=1}^{n} \sum_{j=1}^{m} \mid A(i,j) \mid^2$$

(b) Show that the Shatten 1-norm, also known as the trace norm,

$$|||A|||_1 = ||A||_{\text{trace}} := \sum_{i=1}^{n} \sigma_i(A), \quad \text{where} A \in \mathbb{R}^{n\times m}, \quad n \leq m$$

satisfies the triangle inequality.
Hint: Use the following Lemma:
Lemma: Let $A \in \mathbb{R}^{n\times n}$ have singular values $\sigma_i(1)$, $i = 1, \ldots, n$ in decreasing order and $C \in \mathbb{R}^{n\times n}$ be a rank k partial isometry. Then for each $k = 1, \ldots, n$, we have

$$\sum_{i=1}^{k} \sigma_i(A) = \max\{\mid \mathit{trace}(AC) \mid\}$$

Problem [7]. Approximation of the **clown.mat** image using **Matlab**: after starting **Matlab**, type

```
load clown;
Z = ind2gray(X, map);
[mz,nz] = size(Z);
imshow(Z,64);
mag = 2;
truesize(1, [mz*mag, nz*mag]);
```

Compute the **SVD** of Z, using **svd**: `[U,S,V] = svd(Z);`

1. Plot the singular values of Z on a logarithmic scale.
2. Compute approximant having error less than 10%, 5%, 3%, 2% of the largest singular value of Z. What is the rank of the corresponding approximants? What is the norm of corresponding errors? Also for each case compute the compression ratio (compression ratio is defined as the number of bytes required to store the approximant divided by the original image size in bytes.)
3. Attach with your homework the original image, the approximant and the error image, where the error is less than 2%.
4. Partition Z into four blocks. Then approximate each block such that error is less than 5% for each of them. Form Z_{app} by putting approxmiated blocks together. Compute the norm of the error. Attach the approximant image. Compare the 5% approximants in Part-2 and Part-4.

Problem [8]. Denoising and sorting of images using the **SVD**:

Given are the images of of the three numbers **1**, **7** and **0** represented in Figure-1 (12-pixel images):

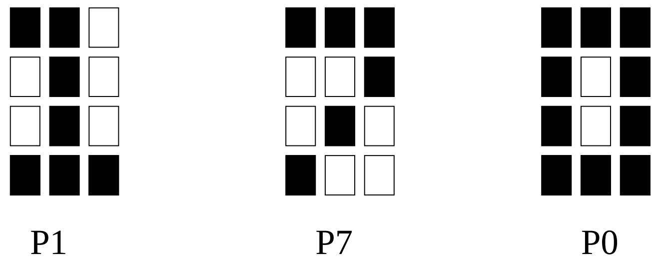

Figure 1: Numbers '1', '7' and '0'

1. Convert the three images in Figure-1 to three 4×3 matrices by replacing black pixels by 1 and white pixels by 0. Let the matrices be called $P1$, $P7$ and $P0$.
2. **Obtaining the noisy data:** Construct 500 noisy versions of each image as follows:

   ```
   P1+0.05*rand(4,3); P7+0.05*rand(4,3); P0+0.05*rand(4,3);
   ```

   This yields 1500 noisy data matrices $N_i \in \mathbb{R}^{4\times 3}$, $i = 1, \cdots, 1500$.
3. Vectorize each image N_i as follows:

   $$x_i := [\ N_i(:,1);\quad N_i(:,2);\quad N_i(:,3)\ ] \in \mathbb{R}^{12\times 1}, \qquad i = 1, \cdots, 1500.$$

   Then, form the following matrix: $A = [\ x_1\ \ x_2\ \ x_3\ \ \cdots\ \ x_{1500}] \in \mathbb{R}^{12\times 1500}$.
4. Take the SVD of A and examine the singular values. How many significant singular values do you see? Find the best rank 3 approximation of A in the 2-norm. Let A_3 be the optimal approximant with SVD $A_3 = U_3 \Sigma_3 V_3^T$ where $U_3 \in \mathbb{R}^{12\times 3}$, $\Sigma_3 \in \mathbb{R}^{3\times 3}$ and $V_3 \in \mathbb{R}^{1500\times 3}$. Each column of V_3^T corresponds to a point in 3-dimensional space. Plot all 1500 points each of which corresponds to a column of V_k^T. You should obtain 3 clusters of points. Find the center of gravity of these 3 clusters. Now, using the coordinates of these points form the 3×3 matrix W_3^T. And finally compute the matrix $\hat{A} = U_3 \Sigma_3 W_3^T \in \mathbb{R}^{12\times 3}$.
5. Reverse the procedure in Step-2 and obtain the matrices $\hat{P}1$, $\hat{P}7$, and $\hat{P}0 \in \mathbb{R}^{4\times 3}$. Compare these matrices with the original noise-free data matrices $P1$, $P7$ and $P0$. Comment on your results.

Grader: Andrew Mayo **Email:** ajmayo@rice.edu
Office: Duncan 2030