Problem Set 7

1. (50 points)

   (a) Use Matlab to find the singular value decomposition of the matrix A given below.

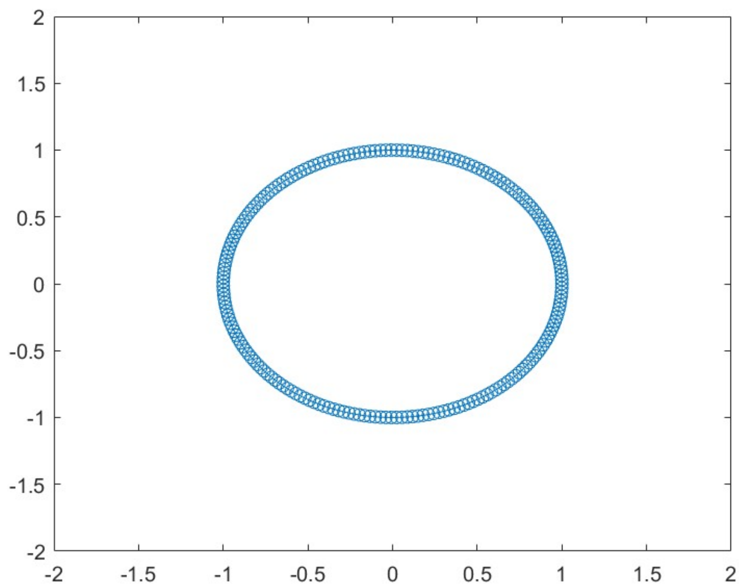
   $$A = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}$$

   The SVD is given by

   $$A = \begin{bmatrix} -0.3245 & -0.9459 \\ -0.9459 & 0.3245 \end{bmatrix} \begin{bmatrix} 6.67678 & 0 \\ 0 & 0.4433 \end{bmatrix} \begin{bmatrix} -0.6070 & -0.7947 \\ 0.7947 & -0.6070 \end{bmatrix} = U\Sigma V^{\intercal}$$

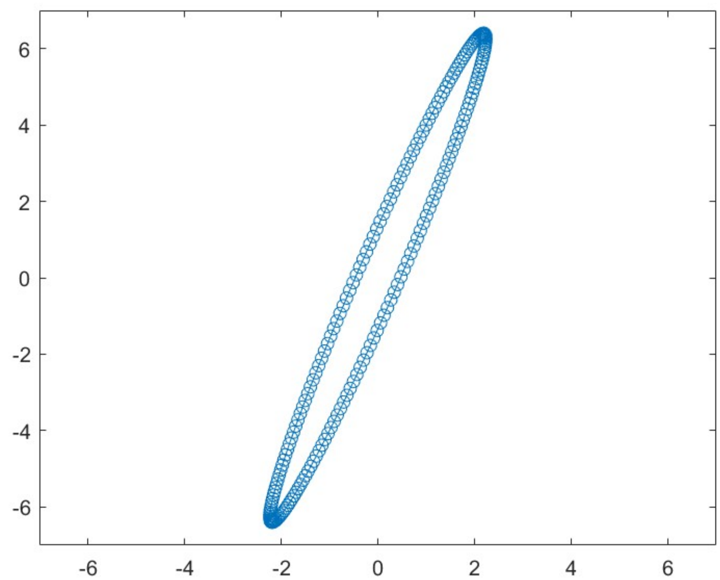
   (b) Write code to make a set of points giving a unit circle S (Hint: Use a common parametrization of a circle, given in terms of the angle θ, and then pick a list of θ values that are equally spaced. Plot S.

   See PS7N1.m for code. I used 200 points to make the following circle.



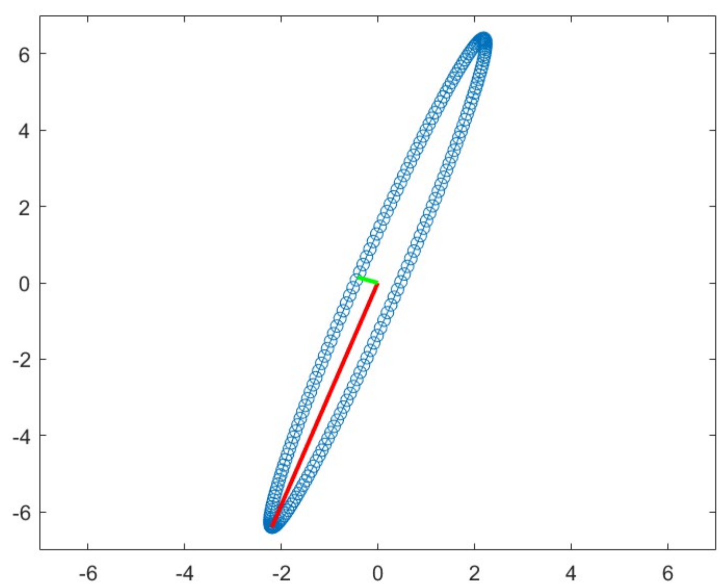
   (c) Apply the matrix A to S to get the resulting ellipse, and plot it.

   By letting S be the list of x values (from the circle) in row 1 and the list of y values in row 2, and then multiplying AS, we can obtain the new x and y values, plotted below.

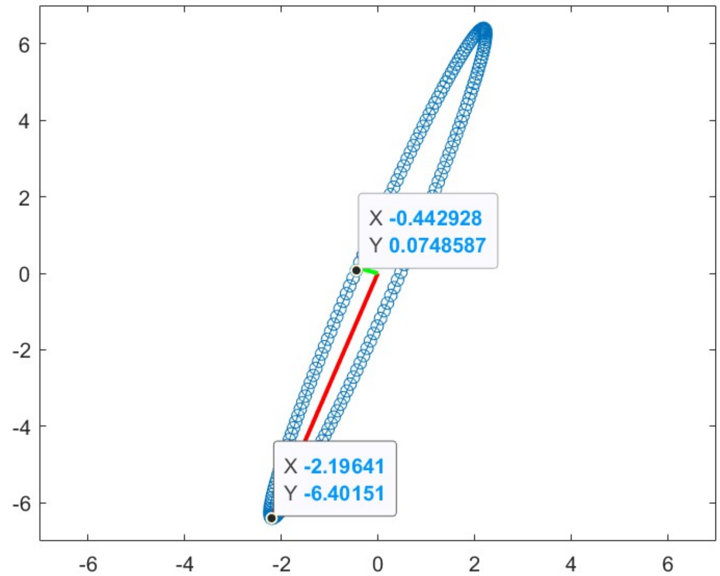


(d) Does your image match what you saw in the SVD? Explain.

We could investigate this a few different ways. Firstly, we can plot $\sigma_i \mathbf{u}_i$ for $i = 1, 2$ to see if these give the principal semiaxes of the ellipse. That plot is shown below, confirming that these axes can be obtained from the SVD of A.



It would also be acceptable to sort of estimate where two end-points of the semiaxes are, shown below and then find the lengths of the semiaxes in order to see that they match up with σ_1 an σ_2. For the estimated points below, we would get that the largest principal semiaxes has length of about $\sqrt{6.4^2 + 2.2^2} \approx 6.77$ and the length of smallest axis is about $\sqrt{0.44^2 + 0.075^2} \approx 0.446$. These both approximately match the singular values we saw in part (a).



2. (75 points) Take the almost singular linear system $A\mathbf{x} = \mathbf{b}$ given below.

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 6 & 8 & 12 & 2 \\ 1 & 2 & 3 + 10^{-15} & 4 \\ 2 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 30 \\ 66 \\ 30 \\ 7 \end{bmatrix}$$

(a) Use Matlab's backslash to estimate a solution to this system. What do you expect the size of the relative error to be? Based on this relative error, what part of your solution can you trust?

Using backslash, we get $\mathbf{x} = \begin{bmatrix} -0.2353 \\ 7.4706 \\ (10)^{-15} \\ 3.8235 \end{bmatrix}$

The condition numbers, in 2-norm and max-norm, respectively, are $\mathcal{O}(10^{16})$ and $\mathcal{O}(10^{17})$. The size of the relative error is therefore likely to be $\mathcal{O}(1)$ or $\mathcal{O}(10)$. This means that I cannot trust *any* digits of our solutions. There is therefore no reason to believe the solution is the above vector.

(b) Use Matlab to find the SVD of the matrix A. At what singular value do you think you should truncate? Explain. What will be the resulting rank r of your matrix A_r? What is the condition number of the resulting A_r (Hint: You do not need to form A_r to answer this.)

The SVD is given by

$A = U\Sigma V^\intercal \approx$

$$\begin{bmatrix} -0.267 & 0.651 & -0.067 & 0.707 \\ -0.919 & -0.359 & 0.163 & 0 \\ -0.267 & 0.651 & -0.067 & -0.707 \\ -0.117 & -.149 & -0.981 & -(10)^{-15} \end{bmatrix} \begin{bmatrix} 17.041 & 0 & 0 & 0 \\ 0 & 4.689 & 0 & 0 \\ 0 & 0 & 1.276 & 0 \\ 0 & 0 & 0 & (10)^{-15} \end{bmatrix}$$

$$. \qquad \begin{bmatrix} -0.368 & -0.501 & -0.748 & -0.233 \\ -0.245 & -0.089 & -0.118 & 0.958 \\ -0.875 & 0.045 & 0.452 & -0.164 \\ -0.194 & 0.860 & -0.0471 & -0.028 \end{bmatrix}.$$

Since the first three singular values are all approximately the same size, but the last one is *much* smaller, that is the value before which we should truncate. We would then get a matrix A_3 of rank $r = 3$. The resulting condition number would be $17.041/1.276 \approx 13.35$, which is *significantly* smaller than the original condition number.

(c) Write Matlab code to find the solution to the now over-determined problem resulting from the rank-r approximation of A. Use `format long` to see more digits of your solution.

See PS7N2.m for the code. The resulting solution is $\mathbf{x} = \begin{bmatrix} 1.000000000000004 \\ 2.000000000000001 \\ 2.999999999999999 \\ 3.999999999999999 \end{bmatrix}.$

(d) Does your solution satisfy $A\mathbf{x} = \mathbf{b}$? Explain. If it does not satisfy the solution, what is the residual? Why would we choose to "solve" this system this way?

The solution does not satisfy $A\mathbf{x} = \mathbf{b}$ because by eliminating one of the singular values, we *changed* the problem we were solving. However, the solution should be close to satisfying the linear system since the singular value we eliminated was very small, and therefore the contribution to the solution along the associated vector should be very small.

The residual is $\mathbf{r} = \mathbf{b} - A\mathbf{x} = \begin{bmatrix} 0 \\ -1.42(10)^{-14} \\ 0 \\ -8.88(10)^{-15} \end{bmatrix}.$

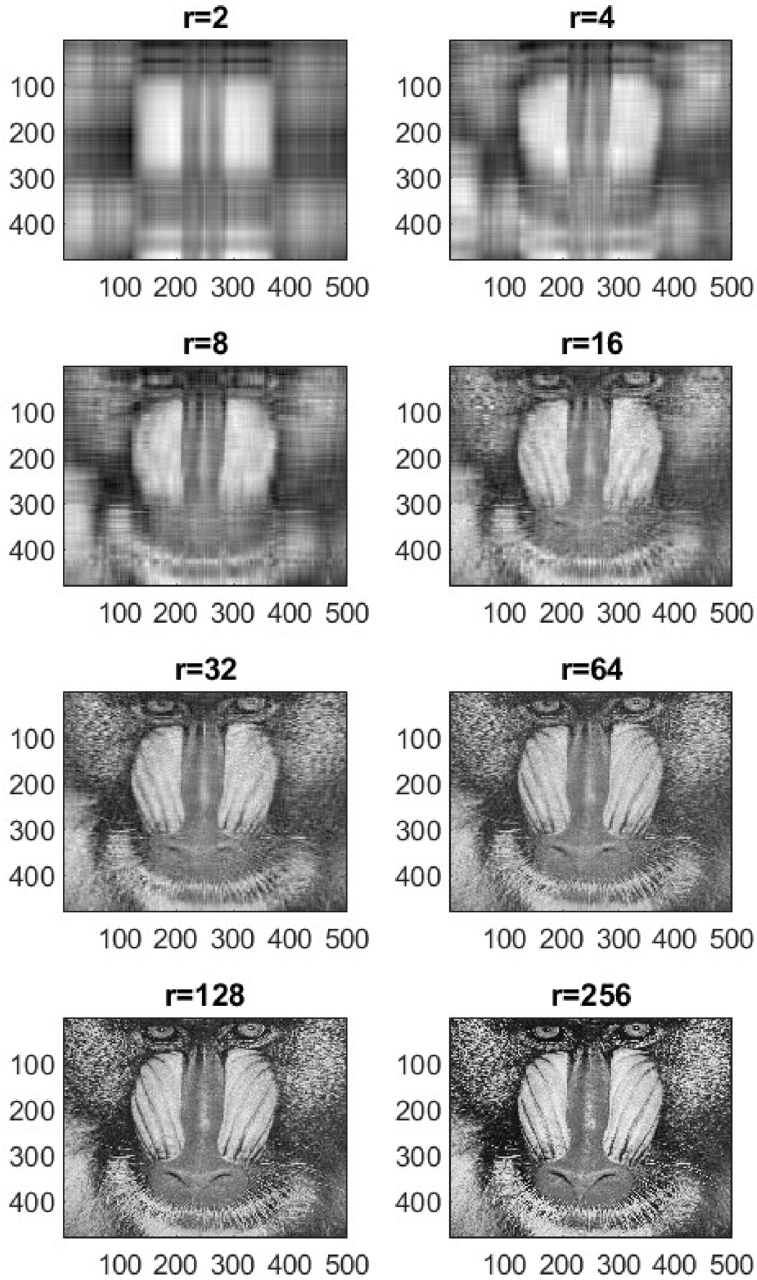
By getting rid of the small singular value of the matrix, the resulting condition number is much smaller than it was originally. Therefore, even though we are introducing a small amount of error by solving a slightly different problem, we are avoiding the gigantic errors we would introduce by working on the original matrix.

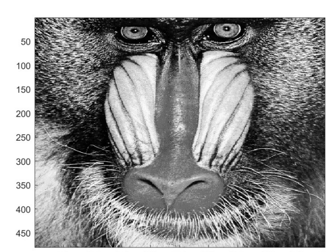
3. (75 points)

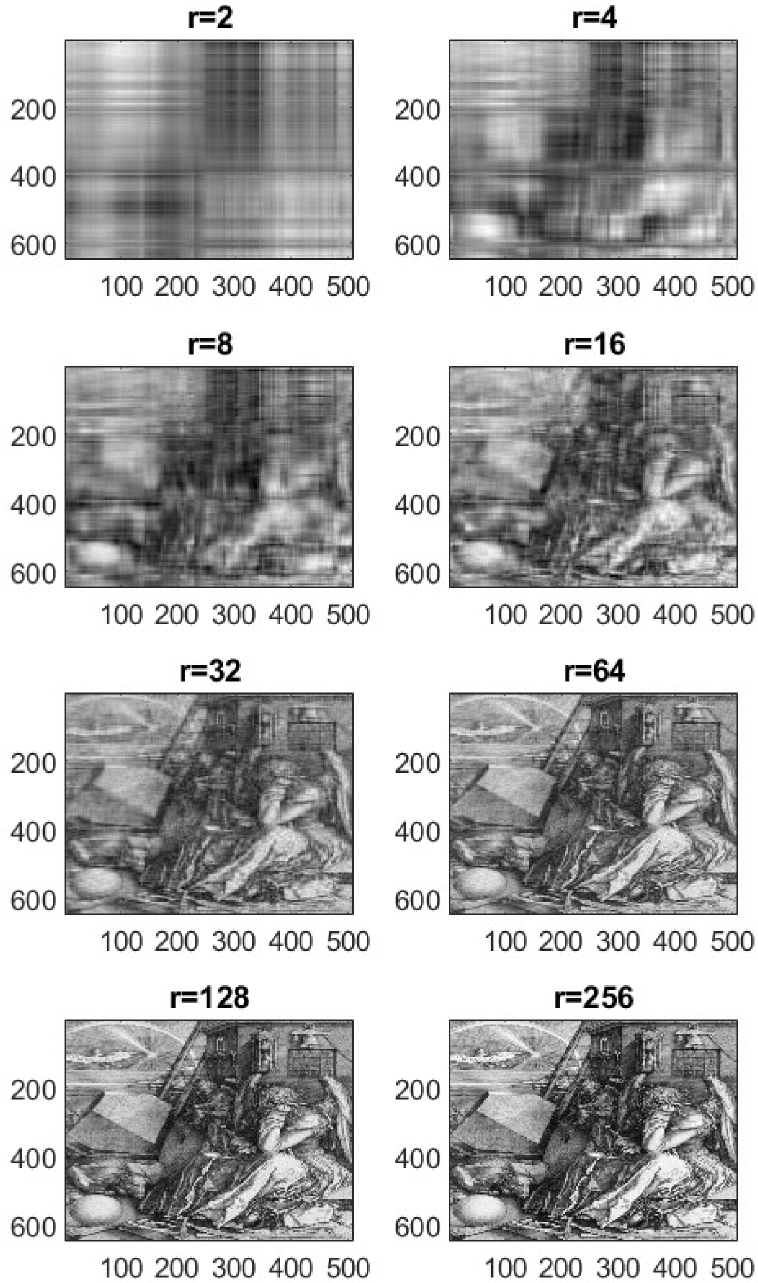
(a) Load the two pictures from Matlab's repository of images, titled "mandrill.mat" and "durer.mat". Use `colormap(gray)` and `imagesc` for all processes in this problem. What are the sizes and ranks of the associated matrices (X)? Write a Matlab script to computing the truncated SVD. For both pictures, examine a range of rank values at which to truncate, r, starting with 2 and going up by powers of 2 to 256. Provide the output. (You may want to use `subplot` to help with the presentation.)

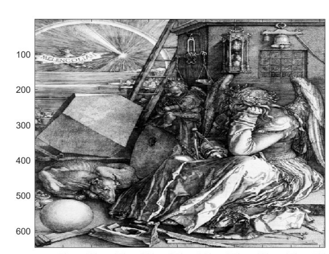
See PS7N3.m for code. For the mandrill, X is a 480×500 matrix with rank 480. For durer, the X is 648×509 with rank 509. Therefore, both matrices are full-rank.

Below are the resulting compressed images, along with the original images.


r=2
r=4
r=8
r=16
r=32
r=64
r=128
r=256







(b) Comment on the performance of the truncated SVD for each image. Explain the difference in the effectiveness of the technique for the two images for small r.

For the mandrill, I would say that we can see what the image is with only a rank of 16, and I would say the image is quite clear by rank 64. For the Durer drawing, there are a lot more sharp details in the image, and for this reason, we need to use more modes to capture the details of the image. I would say that we can make out the large person in the foreground for rank 32, but to get other details, we need rank 64 or 128. Therefore, if an image has many small, sharply contrasting details, a smaller rank may not be enough to fully capture the image, but we are still able to capture the image with a rank that is half of the rank of the original image matrix.

(c) State how much storage is required as a function of r and how much storage is required for the original image.

For the mandrill, the storage required is $981r$, and the original storage was 240000. For the Durer, the storage required is $1158r$ and the original storage was 329832. For $r = 128$, the mandrill requires 125568 and the Durer requires 148224, so by compressing the image to rank 128 we need about half the storage for the image.