AMATH 584 Assignment 2

Tyler Chen

Exercise 3.5

Example 3.6 shows that if E is an outer product $E = uv^*$, then $\|E\|_2 = \|u\|_2 \|v\|_2$. Is the same true for the Frobenius norm, i.e. $\|E\|_F = \|u\|_F \|v\|_F$? Prove it or give a counterexample.

Solution

Let $E = uv^*$ for some $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^n$. Denote the i-th component of v by v_i. We can then write $E = [\overline{v_1}u, ..., \overline{v_n}u]$.

Observe that the Frobenius norm of a column vector is the 2-norm. Moreover, recall that the sum of the squares of the two norm of the columns of a matrix is equal to the square of the Frobenius norm of that matrix.

Thus,

$$\|E\|_F^2 = \sum_{i=1}^{n} \|\overline{v_i}u\|_2^2 = \sum_{i=1}^{n} |\overline{v_i}|^2 \|u\|_2^2 = \|u\|_2^2 \sum_{i=1}^{n} |\overline{v_i}|^2 = \|u\|_2^2 \sum_{i=1}^{n} |v_i|^2 = \|u\|_2^2 \|v\|_2^2 = \|u\|_F^2 \|v\|_F^2$$

This proves that $\|E\|_F = \|u\|_F \|v\|_F$ for $E = uv^*$. $\square$

Exercise 4.1

Determine the SVDs of the following matrices:

$$\text{(a)} \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}, \quad \text{(b)} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \quad \text{(c)} \begin{bmatrix} 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{(d)} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad \text{(e)} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Solution

Note that if A can be written $A = U\Sigma V^*$ for U, V unitary, Σ real diagonal, then this is a SVD decomposition of A. That is, we can simply attempt to manipulate A into a form which looks like the SVD and we will have found the SVD.

(a) Here we simply have to switch the sign of the 2. We do this by right multiplying by a matrix which switches the sign of the second column and left multiplying by the identity.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

(b) Here we need to switch the 2 and 3 so that the singular values are decreasing along the main diagonal. We switch the first and second columns and the first and second rows.

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

(c) Here we simply switch the first and second columns.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

(d) We observe this matrix is a rank 1 outer product xy^* of $x = [1; 0]$, $y = [1, 1]$. Therefore it has 2-norm equal to $\|x\|_2 \|y\|_2 = 1\sqrt{2} = 2$. Therefore, the first signular value is $\sigma_1 = \sqrt{2}$. But as this matrix is rank 1, it has only 1 nonzero singular value.

We have $Av_1 = \sigma u_1$, for unit vectors u_1, v_1. Thus $[v_{11} + v_{12}; 0] = \sqrt{2}[u_{11}; u_{12}]$ so $u_{12} = 0$. Since $\|u_1\|_2 = 1$, WLOG let $u_{11} = 1$. We also have $v_{11} + v_{12} = \sqrt{2}$ and $v_{11}^2 + v_{12}^2 = 1$. Together these give $v_{11} = v_{12} = 1/\sqrt{2}$.

Since U is unitary, WLOG let $u_{21} = 0$ and $u_{22} = 1$. Similarly, we require $v_{11}^2 + v_{21}^2 = 1$ so, $|v_{21}| = 1/\sqrt{2}$. Likewise, $v_{12}^2 + v_{22}^2 = 1$ so, $|v_{22}| = \sqrt{2}$. We also have $Av_2 = [v_{21} + v_{22}, 0] = 0u_2 = 0$. So $v_{21} = -v_{22}$. Finally, we see that the sign has no impact. Therefore, we write,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

(e) We observe this matrix is a rank 1 outer product xy^* of $x = [1; 1]$, $y = [1, 1]$. Therefore it has 2-norm equal to $\|x\|_2 \|y\|_2 = \sqrt{2}\sqrt{2} = 2$. Therefore, the first singular value is 2. But as this matrix is rank 1, it has only 1 nonzero singular value.

We have $Av_1 = \sigma u_1$, for unit vectors u_1, v_1. Thus $[v_{11} + v_{12}; v_{11} + v_{12}] = 2[u_{11}; u_{12}]$ so $u_{11} = u_{12}$. Since $\|u_1\|_2 = 1$, then WLOG let $u_{11} = u_{12} = 1/\sqrt{2}$. Therefore, $v_{11} = v_{12} = 1/\sqrt{2}$.

We require $v_{11}^2 + v_{21}^2 = 1$ so, $|v_{21}| = 1/\sqrt{2}$. Likewise, $v_{12}^2 + v_{22}^2 = 1$ so, $|v_{22}| = \sqrt{2}$. We also have $Av_2 = [v_{21} + v_{22}, v_{21} + v_{22}] = 0u_2 = 0$. So $v_{21} = -v_{22}$. The sign has no impact so WLOG pick $v_{21} = -v_{22} = 1/\sqrt{2}$. Finally, by the same argument, note $u_{21} = -u_{22}$ with $|u_{21}| = |u_{22}| = 1/\sqrt{2}$. However, in this case we require $u_{21} = -u_{22} = 1/\sqrt{2}$. Therefore, we write,

$$\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

Exercise 4.5

Theorem 4.1 asserts that every $A \in \mathbb{C}^{m\times n}$ has an SVD $A = U\Sigma V^*$. Show that if A is real, then it has a real SVD ($U \in \mathbb{R}^{m\times m}, V \in \mathbb{R}^{n\times n}$).

Solution

We first prove the following: *If $A \in \mathbb{R}^{m\times m}$ has real eigenvalue λ, then there exists a real unit eigenvector corresponding to λ.*

Indeed, suppose $v \in \mathbb{C}^m$ is an eigenvector corresponding to λ. That is, $Av = \lambda v$. We can decompose v into its real and imaginary parts, x and y so that $v = x + iy$. Then,

$$\lambda x + i\lambda y = \lambda(x + iy) = \lambda v Av = A(x_i y) = Ax + iAy$$

Since λ is real, then $\lambda x, \lambda y$ are real. Similarly, since A is real, then Ax, Ay are real. We can then equate real and imaginary parts to give,

$$Ax = \lambda x \qquad\qquad Ay = \lambda y$$

Since w is an eigenvector of A, w must be nonzero. This means at least one of x and y is nonzero. This vector is a real eigenvector of A. Clearly we can scale this vector to obtain a real unit eigenvector. $\square$

Next prove the following: *If $A \in \mathbb{R}^{m\times m}$ is symmetric then there is an eigendecompsition $AV = V\Lambda$, for Λ real and V unitary.*

Recall that for a Hermetian matrix all eigenvalues are real, and eigenvectors corresponding to distinct eigenvalues are orthogonal. Suppose λ is an eigenvalue with multiplicity k. Then the eigenvectors corresponding to λ form a k-dimensional subspace. But all vectors in this space are orthogonal to eigenvectors outside this space. Thus, by choosing an orthogonal basis for this set, we have k eigenvectors orthogonal to all other eigenvectors of A.

This proves we can construct a basis for $\mathbb{C}^m$ of orthogonal eigenvectors for A. Clearly these can be normalized. Let V be a matrix with the columns being the real, normal, orthogonal, eigenvectors of A. Then V is real and unitary. Let Λ be a diagonal matrix with the eigenvalues corresponding to the eigenvectors in V placed on the diagonal. Then $AV = V\Lambda$. $\square$

If we order the eigenvalues of A in decreasing order, then $AV = V\Lambda$ is unique up to scalar multiplication and rotation of any of the basis vectors of the subspaces corresponding to repeat eigenvalues.

Let $A \in \mathbb{R}^{m\times n}$.

Suppose v is a unit eigenvalue of A^*A. Then,

$$\lambda = \lambda v^* v = v^*\lambda v = v^*(A^*A)v = (v^*A^*)(Av) = (Av)^*(Av) = \|Av\| \geq 0$$

This proves the eigenvalues of A^*A are positive.

We have A^*A is real Hermetian, so by the above results we have decomposition $A^*AV = V\Lambda$ for some unitary $V \in \mathbb{R}^{n\times n}$. Moreover, we can reorder V and Λ such that the entries of Λ are decreasing in magnitude.

For convenience denote r as the number of nonzero entries of Λ. That is $r = \operatorname{rank}(A^*A)$.

We have $r \leq \min(m, n)$ by rank arguments.

Define $\Sigma \in \mathbb{R}^{m\times n}$ by taking the square roots of the first r entries of Λ along the main diagonal. Leave all other entries zero.

For $j \leq r$, $\sigma_j \neq 0$, so define $u_j := Av_j/\sigma_j$. This gives a set $\{u_1, ..., u_r\}$ of real orthonormal vectors. Complete this set to a real orthonormal basis $\{\{u_1, ..., u_r, u_{r+1}, ..., u_m$ of $\mathbb{R}^m$.

Then observe that for all j, $Av_j = \sigma_j u_j$. That is, $AV = U\Sigma$ for $V \in \mathbb{R}^{n\times n}, U \in \mathbb{R}^{m\times m}$ unitary, and $\Sigma \in \mathbb{R}^{m\times n}$ diagonal with positive decreasing entries.

That is, $A = U\Sigma V^*$ is a real SVD for A. $\square$

Exercise 5.4

Suppose $A \in \mathbb{C}^{m\times m}$ has an SVD $A = U\Sigma V^*$. Find an eigenvalue decomposition of the $2m \times 2m$ Hermetian matrix

$$\begin{bmatrix} 0 & A^* \\ A & 0 \end{bmatrix}$$

Solution

Write the SVD of A as $A = U\Sigma V^*$ so $A^* = V\Sigma^* U^* = V\Sigma U^*$. Recall, for all $1 \leq j \leq m, Av_j = \sigma_j u_j$ and $A^* u_j = \sigma_j v_j$

Then,

$$\begin{bmatrix} 0 & A^* \\ A & 0 \end{bmatrix}\begin{bmatrix} v_j \\ u_j \end{bmatrix} = \begin{bmatrix} A^* u_j \\ Av_j \end{bmatrix} = \begin{bmatrix} \sigma_j v_j \\ \sigma_j u_j \end{bmatrix} = \sigma_j \begin{bmatrix} v_j \\ u_j \end{bmatrix}$$

and similarly,

$$\begin{bmatrix} 0 & A^* \\ A & 0 \end{bmatrix}\begin{bmatrix} v_j \\ -u_j \end{bmatrix} = \begin{bmatrix} -A^* u_j \\ Av_j \end{bmatrix} = \begin{bmatrix} -\sigma_j v_j \\ \sigma_j u_j \end{bmatrix} = -\sigma_j \begin{bmatrix} v_j \\ -u_j \end{bmatrix}$$

That is, $\begin{bmatrix} v_j \\ u_j \end{bmatrix}$ and $\begin{bmatrix} v_j \\ -u_j \end{bmatrix}$ are eignevalues of $\begin{bmatrix} 0 & A^* \\ A & 0 \end{bmatrix}$ with corresponding eigenvalues σ_j and $-\sigma_j$.

We can therefore write,

$$\begin{bmatrix} 0 & A^* \\ A & 0 \end{bmatrix}\begin{bmatrix} V & V \\ U & -U \end{bmatrix} = \begin{bmatrix} V & V \\ U & -U \end{bmatrix}\begin{bmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{bmatrix}$$

Therefore the above decomposition is close to the and eigen decomposition. However, observe,

$$\begin{aligned}
\begin{bmatrix} V & V \\ U & -U \end{bmatrix}\begin{bmatrix} V & V \\ U & -U \end{bmatrix}^* &= \begin{bmatrix} V & V \\ U & -U \end{bmatrix}\begin{bmatrix} V^* & U^* \\ V^* & -U^* \end{bmatrix} \\
&= \begin{bmatrix} VV^* + VV^* & VU^* - VU^* \\ UV^* - UV^* & UU^* + UU^* \end{bmatrix} \\
&= \begin{bmatrix} 2I_m & 0 \\ 0 & 2I_m \end{bmatrix} \\
&= 2I_{2m}
\end{aligned}$$

Therefore, define,

$$X = \frac{1}{\sqrt{2}}\begin{bmatrix} V & V \\ U & -U \end{bmatrix} \qquad \Lambda = \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma \end{bmatrix}$$

Then $XX^* = I$, so the columns of X are orthonormal (and therefore lienarly independent) and Λ is diagonal. Therefore, we have eigen decomposition,

$$\begin{bmatrix} 0 & A^* \\ A & 0 \end{bmatrix} = \left(\frac{1}{\sqrt{2}}\begin{bmatrix} V & V \\ U & -U \end{bmatrix}\right)\begin{bmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{bmatrix}\left(\frac{1}{\sqrt{2}}\begin{bmatrix} V & V \\ U & -U \end{bmatrix}\right)^* = X\Lambda X^* \qquad \square$$

Exercise (Image Compression)

In Matlab, type `imagedemo`. You will see a picture of an Albrecht Durer print. Type `who` to see what variables it has used and type `type imagedem` to see the actual Matlab code that you have run. You will see at the end that it executes the commands:

```
imagesc(X);
colormap(map);
axis off;
```

The 648 by 509 matrix `X` contains a grayscale number (from 1 to 128) for each pixel in a grid. This number determines how dark or light that pixel will be shaded when the command `imagesc(X)` is executed. This is fine if one can store a 648 by 509 matrix, but if there are many such images and they are, say, being sent from outer space, using this large a matrix to represent each one could be prohibitive!

Compute the SVD of `X`. Try executing the above commands with `X` replaced by some low rank approximations formed from the largest singular values and corresponding singular vectors, and decide about how many singular values/vectors are needed to make the picture recognizable. Turn in a few plots showing how the picture improves as you increase the rank of the approximation used. Label each plot with the rank of the approximation used. [You can put several plots on one page using the `subplot` command. Type `help subplot` to see exactly how it works. You can save your plots to a file by typing `print -depsc hw2plots.eps` where the filename `hw2plots` can be replaced by any name you like.]

Solution

We first export the image matrix `X` from MATLAB as a file `img.mat`.

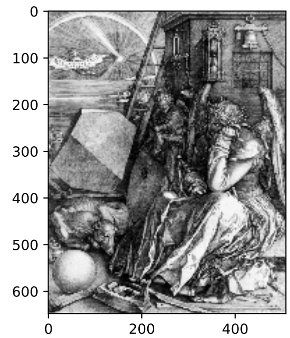
We then import with SciPy. We plot the original image. We compute the SVD of the matrix. We plot the rank-k approximation of the matrix for the listed k. Note that rather than computing the rank-k approximation from `X` we simply multiply the appropriate submatrices of $U, V, S := \Sigma$.

The outputs are saved and appended.

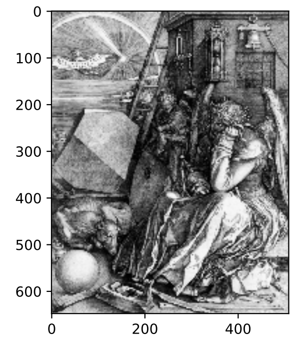
What it means for an image to be "recognizable" is vague, however the rank-50 approximation is pretty close to the original image, and the rank-150 image is almost indistinguishable from the original.

```
import scipy as sp
from matplotlib import pyplot as plt

def exercise_2_4():
    #import from matlab export
    M=sp.io.loadmat('img.mat')
    X=M['X']

    m,n = X.shape

    # original matrix
    fig=plt.figure()
    plt.imshow(X,cmap='gray')
    fig.savefig('img/original.pdf',bbox_inches='tight')


    # get SVD
    [U,s,V] = sp.linalg.svd(X) # full_matrices=True
```

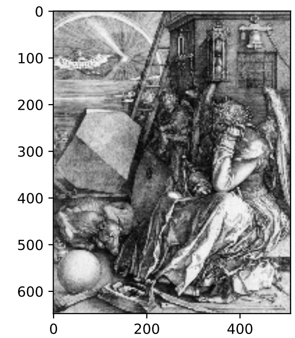
```
    S = sp.zeros((m,n))
    S[:n,:n] = sp.diag(s)

    for k in [509,300,150,100,50,30,20,10,5,1,0]:
        #plot rank k approximation
        fig=plt.figure()
        plt.imshow(sp.dot(U[:,:k],sp.dot(S[:k,:k],V[:k])),cmap='gray')
        fig.savefig('img/'+str(k)+'.pdf',bbox_inches='tight')

exercise_2_4()
```



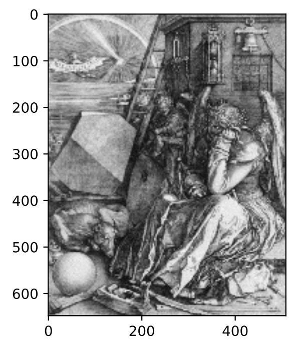
(a) original



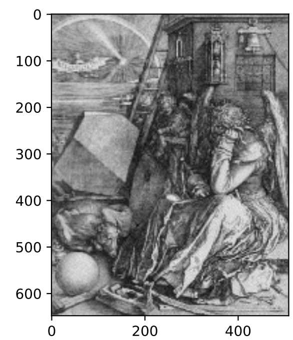
(b) $k = 509$



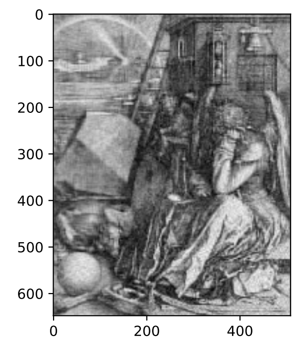
(c) $k = 300$



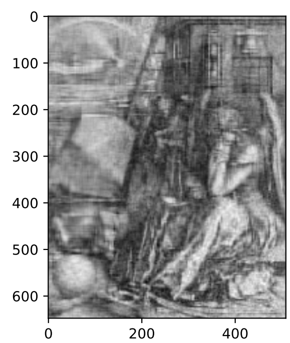
(d) $k = 150$



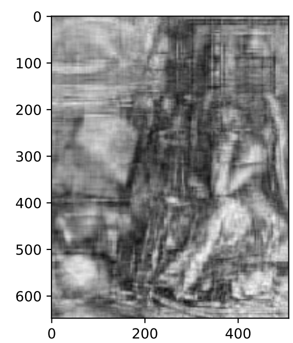
(e) $k = 100$



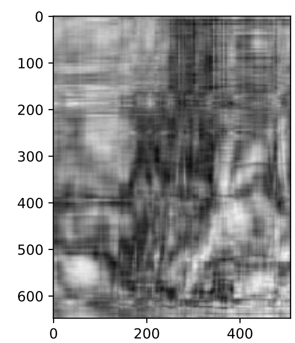
(f) $k = 50$



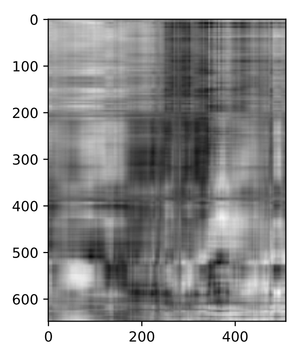
(g) $k = 30$



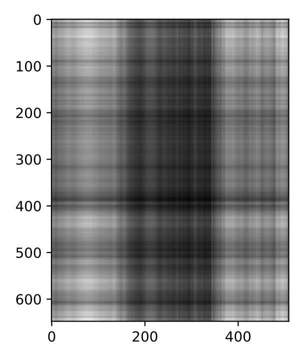
(h) $k = 20$



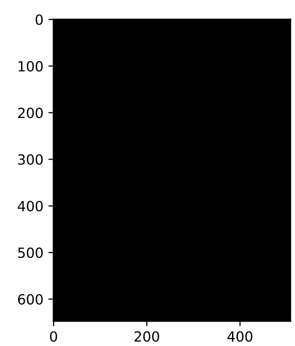
(i) $k = 10$



(j) $k = 5$



(k) $k = 1$



(l) $k = 0$