ORF 523	PROBLEM SET 1	Spring 2019, Princeton University
Instructor: A.A. Ahmadi		
TA: Cemil Dibek		Due at 1:30pm, in class, on February 21, 2019

Problem 1: Image compression on Valentine's day

A life without love is of no account. Don't ask yourself what kind of love you should seek, spiritual or material, divine or mundane, Eastern or Western. Divisions only lead to more divisions. Love has no labels, no definitions. It is what it is, pure and simple. Love is the water of life. And a lover is a soul of fire. The universe turns differently when fire loves water. - Shams Tabrizi

In honor of Valentine's day, we study the problem of compressing the image of Shams Tabrizi, a Persian spiritualist, famous for being the mentor of Rumi, and for the *"Forty Rules of Love"*, from which the quote above is extracted.

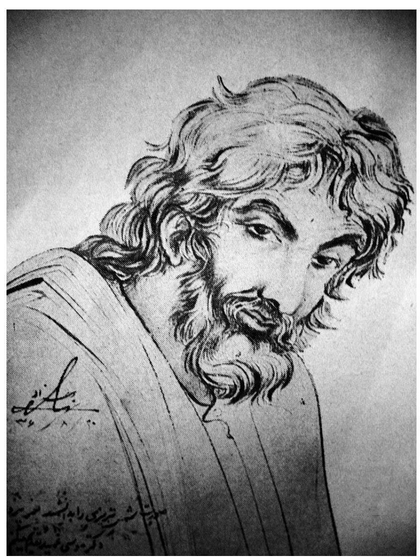

Figure 1: Shams Tabrizi (1185 –1248). The image to be compressed.

To do this compression, we introduce you to the concept of *singular value decomposition* (SVD) which is arguably one of the most important topics in computational linear algebra. This question covers one application of this concept in image processing, but the SVD has many other applications, notably in statistics (principal component analysis).

Let A be a real $m \times n$ matrix of rank r. (Recall that the rank of A is the number of linearly independent columns of A.) The singular value decomposition of A is a decomposition of the form

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

where U, Σ, and V are respectively $m \times m$, $m \times n$, and $n \times n$; U and V are orthogonal matrices (i.e., satisfy $U^T U = I$ and $V^T V = I$), and Σ is a matrix with r positive scalars $\sigma_1, \ldots, \sigma_r$ on the diagonal of its upper left $r \times r$ block and zeros everywhere else. The scalars $\sigma_1, \ldots, \sigma_r$ are called the *singular values* of A. They are given as

$$\sigma_i = \sqrt{i\text{-th eigenvalue of } A^T A},$$

and by convention they appear in descending order:

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r.$$

The columns of U and V are respectively called the left and right *singular vectors* of A and can be obtained by taking an orthonormal set of eigenvectors[1] for the matrices AA^T and $A^T A$. In Matlab, the command `svd` handles these eigenvector computations for you and outputs the three matrices U, Σ, V.

First, you need to answer some very basic questions about the SVD.

1. (a) Show that eigenvalues of $A^T A$ are always nonnegative. (Hence singular values are well-defined as real, nonnegative scalars.)

   (b) Show that if A is symmetric then the singular values of A are the same as the absolute value of the eigenvalues of A.

What does this have to do with optimization? Let A be a real $m \times n$ matrix with an SVD given by $A = U\Sigma V^T$ defined as above. For a positive integer $k \leq \min\{m, n\}$, we let $A_{(k)}$ denote an $m \times n$ matrix which is an "approximation" of the matrix A obtained from its top k singular values and singular vectors. Formally, we have

$$A_{(k)} := U_{(k)} \Sigma_{(k)} V_{(k)}^T,$$

where $U_{(k)}$ has the first k columns of U, $V_{(k)}$ has the first k columns of V, and $\Sigma_{(k)}$ is the upper left $k \times k$ block of Σ.

[1] By an orthonormal set of vectors we mean a collection of vectors that are pairwise orthogonal and each have 2-norm equal to one.

2. Show that

$$A_{(k)} = \min_{B \in \mathbb{R}^{m \times n}, rank(B) \leq k} ||A - B||_2.$$

Here, $||.||_2$ denotes the spectral norm of a matrix defined as $||C||_2 = \max_{||x||_2=1} ||Cx||_2$. (Hint: You may want to first prove that the spectral norm of a matrix is *unitarily invariant*, i.e., it does not change when the matrix is multiplied from left or right by an orthogonal matrix. You may also want to use the following fact from linear algebra: For any matrix $E \in \mathbb{R}^{m \times n}$, $\text{rank}(E) + \dim \text{null}(E) = n$.)

In words, the result you are proving states that among all $m \times n$ matrices of rank at most k, the matrix $A_{(k)}$ obtained from truncating the SVD is the one that best approximates A in the spectral norm. The benefit in approximating a matrix with low-rank matrices is that low-rank matrices admit a much more succinct representation. It turns out that the same result holds for the Frobenius norm; i.e.,

$$A_{(k)} = \min_{B \in \mathbb{R}^{m \times n}, \text{rank}(B) \leq k} ||A - B||_F.$$

(Recall that the Frobenius norm of a matrix is defined as $||C||_F = \sqrt{\sum_{i,j} C_{i,j}^2}$.) Let's see how this applies to our image compression problem.

Download the file `Shams.jpg` into your Matlab path. You can read this file in by typing:

```
A=imread('Shams.jpg');
A=im2double(A);
A=rgb2gray(A);
```

The result is a 1024×768 matrix A, with each entry representing a single pixel in the picture with a number between 0 and 1. To upload this picture on Instagram, you would need to upload $1024 \times 768 = 786432$ numbers (pixels).

3. For $k = 25, 50, 100, 150$, use Matlab to compute $A_{(k)}$ as defined above. Report the value of $||A - A_{(k)}||_F$ in each case. (Include your code for this part and the next.)

4. Use the commands `subplot` and `imshow` to produce on the same figure the original image, as well as your compressed images $A_{(k)}$ for $k = 25, 50, 100, 150$. Label your subplots. In addition, produce two separate plots demonstraing (i) $||A - A_{(k)}||_F$ versus k, and (ii) "total savings" versus k. Total savings is to be interpreted as the answer to the question: How many fewer numbers do you need in order to store $A_{(k)}$ than you did to store A? Explain why this number is equal to $mn - (n + m + 1)k$. How much are you saving for $k = 150$?

5. Use the Matlab function `imwrite` to create two images from `imshow(A)` and `imshow(`$A_{(200)}$`)`. Can you tell them apart? Does Shams Tabrizi look any less mystical?

Problem 2: Tests for local optimality of quadratics

Without using arguments based on convexity, prove that for any quadratic function $f(x) = x^T Q x + b^T x + c$, the following statements hold:

1. $\bar{x}$ is a local min $\Leftrightarrow \nabla f(\bar{x}) = 0$ and $\nabla^2 f(\bar{x}) \succeq 0$.

2. $\bar{x}$ is a strict local min $\Leftrightarrow \nabla f(\bar{x}) = 0$ and $\nabla^2 f(\bar{x}) \succ 0$.

Problem 3: Norms, dual norms, and induced norms (see notes)

1. Let $Q \in S^{n \times n}$ and assume $Q \succ 0$. Show that

$$f(x) = \sqrt{x^T Q x}$$

is a norm.

2. Show that Q^{-1} exists and is positive definite. Show that the dual norm of f is given by

$$g(x) = \sqrt{x^T Q^{-1} x}.$$

(Hint: You may want to bring in $\sqrt{Q}$, i.e., a matrix whose square is Q. If you do, you have to first prove that this matrix exists.)

3. Let $A \in \mathbb{R}^{m \times n}$. Prove the following expression for its induced 2-norm:

$$||A||_2 = \sqrt{\lambda_{\max}(A^T A)}.$$

Problem 4: Properties of positive semidefinite matrices

Prove or disprove the following statements. All matrices are $n \times n$ and with real entries.

(a) Suppose $A \succeq 0$. Then the largest entry in absolute value of A must be on the diagonal.[2]

(b) If $A \succeq 0$ and trace$(A) = 0$, then $A = 0$.

(c) If $A \succeq 0, B \succeq 0$, and $A + B = 0$, then $A = B = 0$.

(d) If $A \succeq 0, B \succeq 0$, and $AB = 0$, then $A = 0$ or $B = 0$.

[2] In other words, the largest entry of $|A|$ must be on the diagonal, where $|A|$ is the matrix whose entries are the absolute values of those of A.