

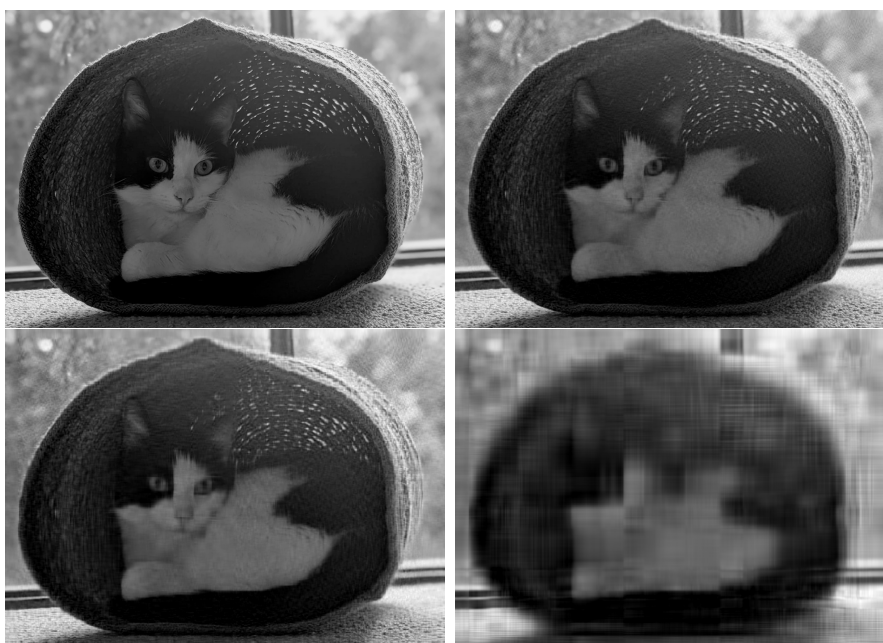
MAST10008 Written Assignment 2

Due 22 May 2026 on paper in class

Name:

Student ID:

Please hand in your answers for different questions on separate sheets of paper, if possible – they will be marked by different tutors.



You can see that the photos above all have different image quality – on the top left, you see the original, to its right a decent approximation, and on the bottom, two more versions that are further and further from the original. In this question, you will be guided through how these pictures were created, and why it works. A supporting lab-sheet from my old linear algebra collection can be found with the self-study problems. You are encouraged to play around with it and maybe compress some of your own pictures, but this does not form part of the assessment.

Our first task is to translate the problem into the setting of linear algebra: what is a digital photo, and what does it mean for your compressed picture to be any good?

Question 1 Your grayscale photo is encoded as an $m \times n$ -matrix with positive entries (really it is positive integer entries, but you may ignore that), where the magnitude of a given entry indicates the brightness of the corresponding pixel (so, an entry of 0 means that the pixel is black). So, let m and n be fixed throughout this assignment, and consider the space $M_{m \times n}(\mathbb{R})$ of $m \times n$ -matrices. Given $A, B \in M_{m \times n}(\mathbb{R})$, the Frobenius norm of A , is defined as

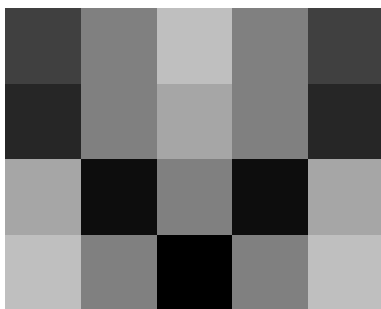
$$\|A\| = \sqrt{\text{tr}(A^T A)}.$$

The Frobenius distance between A and B is defined as $\|A - B\|$.

- Convince yourself that when $n = 1$, the Frobenius norm is the length of the vector as defined using the dot product.
- What does a small Frobenius norm tell you about the picture? How about the Frobenius distance between two pictures? Play around with this – work out some examples, make an educated guess, then prove it. Along the way, you will probably find a different but equivalent formula for $\|A\|$, which will be useful here.
- Prove $\|A^T\| = \|A\|$. What is the meaning of this for your digital photos?

Question 2

In this exercise, you will explore how the nature of matrix multiplication informs image compression. The grainy grayscale picture of your teddy bear on the left



is encoded in the 4×5 matrix

$$A = \begin{bmatrix} 5 & 10 & 15 & 10 & 5 \\ 3 & 10 & 13 & 10 & 3 \\ 13 & 1 & 10 & 1 & 13 \\ 15 & 10 & 0 & 10 & 15 \end{bmatrix}.$$

Think of each of the squares in the picture as a pixel. The numbers in the matrix A indicate brightness of the pixel in this position on a scale from 0 for black to 20 for white.

You can work through the linear algebra lab sheet to see (or take my promise) that $A = USV^T$, where (up to the fourth decimal)

$$U = \begin{bmatrix} -0.5113 & -0.5006 & -0.0450 & -0.6971 \\ -0.4382 & -0.5449 & 0.1195 & 0.7049 \\ -0.4604 & 0.3621 & -0.8001 & 0.1293 \\ -0.5785 & 0.5669 & 0.5861 & -0.0207 \end{bmatrix} \quad S = \begin{bmatrix} 38.7986 & 0 & 0 & 0 & 0 \\ 0 & 18.0044 & 0 & 0 & 0 \\ 0 & 0 & 11.0684 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$V = \begin{bmatrix} -0.4777 & 0.5040 & -0.1334 & 0.7071 & 0.0080 \\ -0.4057 & -0.2456 & 0.5245 & -0.0080 & 0.7071 \\ -0.4631 & -0.6093 & -0.6436 & 0 & 0 \\ -0.4057 & -0.2456 & 0.5245 & 0.0080 & -0.7071 \\ -0.4777 & 0.5040 & -0.1334 & -0.7071 & -0.0080 \end{bmatrix}.$$

For the discussion for this question, it really only matters which entries are zero and which are not, so you can write $*$ for each non-zero entry of U and V or, alternatively, use the notation

$\vec{u}_1, \dots, \vec{u}_4 \in \mathbb{R}^4$ for the column vectors of U and $\vec{v}_1, \dots, \vec{v}_5 \in \mathbb{R}^5$ for the column vectors of V and σ_1, σ_2 and σ_3 for the three non-zero entries of S . **No messy computations will be necessary.**

Delete the last column of U , the last two columns of V and the last row and last two columns from S . This give the new matrices

$$U_{new} = \begin{bmatrix} -0.5113 & -0.5006 & -0.0450 \\ -0.4382 & -0.5449 & 0.1195 \\ -0.4604 & 0.3621 & -0.8001 \\ -0.5785 & 0.5669 & 0.5861 \end{bmatrix} \qquad V_{new} = \begin{bmatrix} -0.4777 & 0.5040 & -0.1334 \\ -0.4057 & -0.2456 & 0.5245 \\ -0.4631 & -0.6093 & -0.6436 \\ -0.4057 & -0.2456 & 0.5245 \\ -0.4777 & 0.5040 & -0.1334 \end{bmatrix}.$$

and the diagonal matrix S_{new} with diagonal entries 38.7986, 18.0044 and 11.0684. We also write R for the 3×5 matrix obtained from S by deleting the last row.

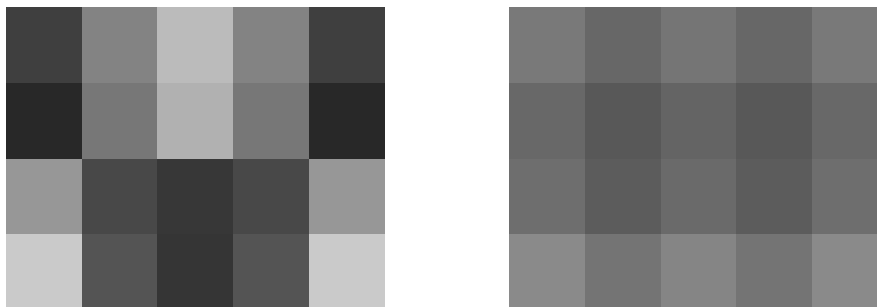
(a) Explain or prove each of the following matrix equations:

$$\begin{aligned} (1) \qquad \qquad \qquad US &= U_{new} R \\ (2) \qquad \qquad \qquad RV^T &= S_{new} V_{new}^T \\ (3) \qquad \qquad \qquad A &= U_{new} S_{new} V_{new}^T. \end{aligned}$$

Making the matrices U and V even smaller comes at the expense of changing our picture. Write

$$S_{small} = \begin{bmatrix} 38.7986 & 0 & 0 & 0 & 0 \\ 0 & 18.0044 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{and} \qquad S_{tiny} = \begin{bmatrix} 38.7986 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Then the matrices $B = US_{small}V^T$ and $C = US_{tiny}V^T$ encode the compressed teddy pictures:



- (b) Write C as the product of a 4×1 a 1×1 and a 1×5 matrix.
- (c) How many matrix entries to you need in total to produce C ?
- (d) You are given a greyscale photo of 1200×1800 pixels and apply the same method as for teddy, this time deleting all but the first 50 entries in S . How much space does this help you save?

Question 3

It is now time to understand the role of the matrices U and V . First, we need to understand which features of U and V guarantee that small changes in S do not blow up into big changes in A . Remember that all matrices in this assignment have real entries. An orthogonal matrix is a square matrix U satisfying $U^T U = I_n$. In particular, every orthogonal matrix is invertible.

- (a) Give three examples of orthogonal matrices and discuss their geometry.
- (b) Prove that, given an orthogonal $n \times n$ -matrix U and an $m \times n$ matrix A , one has $\|AU\| = \|A\|$.
- (c) Using what you just proved, compute the Frobenius distances $\|A - B\|$ and $\|A - C\|$ from the Teddy bear question above to quantify the quality of your compressed images. Justify your answer carefully. You may pretend that U and V orthogonal, although this is only almost true. For an optional extra challenge (unassessed) try to think about why we can get away with this.

Question 4

Finally, how to find U and V ? You may use without proof the *Spectral Theorem* (orthogonal diagonalisation) stating that every symmetric matrix with real entries possesses an orthonormal basis of eigenvectors.

Let λ be a non-zero real number.

- (a) Prove: if $\vec{v} \in \mathbf{R}^n$ is an eigenvector for $A^T A$ with eigenvalue λ , then $A\vec{v}$ is an eigenvector for AA^T with eigenvalue λ .
- (b) Assume that λ is a non-zero eigenvalue for $A^T A$. Prove that $\lambda > 0$. **Hint:** given an eigenvector \vec{v} with eigenvalue λ how do the lengths of \vec{v} and $A\vec{v}$ compare?
- (c) Prove that the null-spaces of A and of $A^T A$ agree, $\text{Null}(A) = \text{Null}(A^T A)$. **Hint:** use that only the zero vector has length zero.
- (d) How are the ranks of A , $A^T A$ and AA^T related?
- (e) Prove that, given two composable linear transformations $T: U \rightarrow V$ and $S: V \rightarrow W$, where U , V , and W are finite dimensional \mathbb{F} -vector spaces, then

$$\text{rank}(ST) \leq \min(\text{rank}(S), \text{rank}(T)).$$

- (f) Sort your eigenvalues of $A^T A$ by decreasing magnitude,

$$\lambda_1 > \lambda_2 > \dots > \lambda_k$$

(so zero is last if it is an eigenvalue). Given an orthonormal basis \mathcal{B} of eigenvectors $\vec{v}_1, \dots, \vec{v}_n$ of $A^T A$, sorted to match this ordering of the eigenvalues, construct an orthonormal eigenbasis \mathcal{C} of AA^T as follows: take $\vec{u}_1, \dots, \vec{u}_{\text{rank}(A)}$ to be the normalisations of $A\vec{v}_1, \dots, A\vec{v}_{\text{rank}(A)}$. Then find an orthonormal basis for the null-space of AA^T , and add these vectors to the end of the list. Let V be the change of basis matrix from \mathcal{B} to the standard basis of \mathbb{R}^n , and let U be the change of basis matrix from the basis \mathcal{C} to the standard basis of \mathbb{R}^m . Let $S = [A]_{\mathcal{C}-\mathcal{B}}$. Find S and explain the equation $A = USV^T$ in terms of base change.

Question 5

It is time to do an example. Work out the formalism of the last two question parts in the concrete example $A = \begin{bmatrix} 1 & -2 \\ 3 & 4 \\ 5 & 0 \end{bmatrix}$; you may use

$$\begin{bmatrix} 1 & -2 \\ 3 & 4 \\ 5 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 & 5 \\ -2 & 4 & 0 \end{bmatrix} = \begin{bmatrix} 5 & -5 & 5 \\ -5 & 25 & 15 \\ 5 & 15 & 25 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & 5 \\ -2 & 4 & 0 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 3 & 4 \\ 5 & 0 \end{bmatrix} = \begin{bmatrix} 35 & 10 \\ 10 & 20 \end{bmatrix}.$$

To make your task more precise: find an orthonormal basis of eigenvectors for either AA^T or $A^T A$, *whichever is easier*, transport as much of it as you can over to begin building an eigenbasis for the other. Take care to treat the eigenvalue 0 separately, if it exists. Scale the eigenvectors to unit length and convince yourself that the resulting base change matrices are orthogonal matrices. Find U , V and S and multiply them back together to fact-check that you indeed get A .

Finally, how would you proceed if you were now asked to go through the same exercise for the matrix $\begin{bmatrix} 1 & 3 & 5 \\ -2 & 4 & 0 \end{bmatrix}$?