Problem Set 10 due November 25, at 10 AM, on Gradescope (via Stellar)

Please list all of your sources: collaborators, written materials (other than our textbook and lecture notes) and online materials (other than Gilbert Strang's videos on OCW).

Give complete solutions, providing justifications for every step of the argument. Points will be deducted for insufficient explanation or answers that come out of the blue

Problem 1: Represent the US flag as a 13×25 matrix A, where each entry represents a color as follows: the entry 1 represents red, the entry 0 represents white, and the entry -1 represents blue. Then write this matrix A as a sum of rank 1 matrices (i.e. akin to formula (239) in the lecture notes).

Note on vexillology: you may ignore the stars, so just assume that the top left corner is a full-blue 7×10 submatrix of A. The height of all the stripes is one row. *(20 points)*

Proof. Let's write $+$ for 1 and $-$ for -1. Then A is the following matrix:

$$\begin{bmatrix}
- & - & - & - & - & - & - & - & - & - & + & + & + & + & + & + & + & + & + & + & + & + & + & + & + \\
- & - & - & - & - & - & - & - & - & - & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
- & - & - & - & - & - & - & - & - & - & + & + & + & + & + & + & + & + & + & + & + & + & + & + & + \\
- & - & - & - & - & - & - & - & - & - & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
- & - & - & - & - & - & - & - & - & - & + & + & + & + & + & + & + & + & + & + & + & + & + & + & + \\
- & - & - & - & - & - & - & - & - & - & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
- & - & - & - & - & - & - & - & - & - & + & + & + & + & + & + & + & + & + & + & + & + & + & + & + \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
+ & + \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
+ & + \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
+ & +
\end{bmatrix}$$

We have:

$$A = B + C$$

where:

$$B = \frac{\boldsymbol{e}_1 + \boldsymbol{e}_3 + \boldsymbol{e}_5 + \boldsymbol{e}_7 + \boldsymbol{e}_9 + \boldsymbol{e}_{11} + \boldsymbol{e}_{13}}{\sqrt{7}} \cdot 5\sqrt{7} \cdot \frac{\boldsymbol{e}_1^T + ... + \boldsymbol{e}_{25}^T}{5}$$

$$C = \frac{2\boldsymbol{e}_1 + \boldsymbol{e}_2 + 2\boldsymbol{e}_3 + \boldsymbol{e}_4 + 2\boldsymbol{e}_5 + \boldsymbol{e}_6 + 2\boldsymbol{e}_7}{\sqrt{19}} \cdot (-\sqrt{190}) \cdot \frac{\boldsymbol{e}_1^T + ... + \boldsymbol{e}_{10}^T}{\sqrt{10}}$$

You didn't necessarily need to renormalize the vectors above in order to have length 1 in this problem, since I never explicitly asked you to.

$\square$

Problem 2: If A is an $n \times n$ symmetric matrix with distinct eigenvalues $\lambda_1, ..., \lambda_n$ and orthonormal eigenvectors $\boldsymbol{q}_1, ..., \boldsymbol{q}_n$, what is the SVD of A? *(10 points)*

Proof. The diagonalization of a symmetric matrix is:

$$S = Q \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} Q^T \tag{1}$$

where the columns of Q are the orthonormal vectors $\boldsymbol{q}_1, ..., \boldsymbol{q}_n$. The SVD of S has $U = V = Q$ and Σ is the diagonal matrix with entries $\lambda_1, ..., \lambda_n$. *(8 points)*

However, the argument above only holds if $\lambda_1, ..., \lambda_n \geq 0$, since the singular values are positive. If some of the λ_i's are negative (let's assume $\lambda_1, ..., \lambda_i \geq 0 > \lambda_{i+1}, ..., \lambda_n$, since we can always rearrange the eigenvalues) then we can rewrite (1) as:

$$S = \begin{bmatrix} \boldsymbol{q}_1 & \dots & \boldsymbol{q}_i & -\boldsymbol{q}_{i+1} & \dots & -\boldsymbol{q}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_i & 0 & \dots & 0 \\ 0 & \dots & 0 & -\lambda_{i+1} & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & -\lambda_n \end{bmatrix} \begin{bmatrix} \boldsymbol{q}_1 & \dots & \boldsymbol{q}_i & \boldsymbol{q}_{i+1} & \dots & \boldsymbol{q}_n \end{bmatrix}^T$$

If one of the eigenvalues is 0, then we assume $\lambda_n = 0$ (the eigenvalues of a matrix have no preferred order, so they can be ordered in any way you want when you diagonalize the said matrix). *(2 points)*

□

Problem 3: All matrices in this problem are 2×2. A lower/upper triangular matrix with 1's on the diagonal has one degree of freedom (the bottom-left/top-right entry); a diagonal matrix has two degrees of freedom (the diagonal entries). Hence the LDU factorization has $1 + 2 + 1$ degrees of freedom, which is precisely the number of degrees of freedom in choosing a 2×2 matrix.

(a) How many degrees of freedom does an orthogonal 2×2 matrix Q have? Explain. *(5 points)*

(b) What is the total number of degrees of freedom of the QR factorization? What about the total number of degrees of freedom of the SVD $U\Sigma V^T$? Explain. *(5 points)*

(c) What is the total number of degrees of freedom of $Q\Lambda Q^T$, where Q is orthogonal and Λ is diagonal? Still in the 2×2 case. *(5 points)*

(d) Why didn't you get 4 in part (c)?
Hint: it's because matrices $Q\Lambda Q^T$ are special, i.e. they are ______________ *(5 points)*

Proof. (a) There are 2 degrees of freedom in choosing a vector in the plane, but only one degree of freedom in choosing a vector of length 1 (it has to lie on a circle of radius 1 centered at the origin). Therefore, you have 1 degree of freedom in choosing the first column of an orthogonal matrix Q. But then, you have no more degrees of freedom in choosing the second column, because it has to be a length 1 vector perpendicular to the already chosen first column. Therefore, the answer is 1.

(b) There are 3 degrees of freedom in choosing the upper triangular matrix R with no restriction on the diagonal entries (namely the $(1,1)$, $(1,2)$ and $(2,2)$ entries). Therefore, the total number of degrees of freedom in the QR factorization is $1+3=4$, which is the same as the total number of degrees of freedom in choosing a 2×2 matrix.

Meanwhile, there are $1+2+1=4$ degrees of freedom in choosing the SVD $U\Sigma V^T$, since U and V must be orthogonal matrices and Σ must be diagonal.

(c) The total number of degrees of freedom is $1+2$: one for Q and two for Λ.

(d) We got $1+2=3$ in part (c) because matrices of the form $Q\Lambda Q^T$ are not general 2×2 matrices, but just symmetric ones. You only have 3 degrees of freedom in choosing a symmetric 2×2 matrix, because the $(1,2)$ entry must be equal to the $(2,1)$ entry.

□

Problem 4: Consider the matrix $A = \begin{bmatrix} 1 & -1 & 0 \\ -2 & 3 & 1 \\ 0 & 2 & 2 \end{bmatrix}$.

(a) Compute the SVD of A, and the pseudo-inverse A^+. *(15 points)*

(b) Compute the vector $\boldsymbol{v}^+$ defined by formula (261) in the lecture notes, which will have the property that $A\boldsymbol{v}^+ = \boldsymbol{p}$ is as close to $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ as possible. *(5 points)*

(c) Compute all solutions to $A\boldsymbol{v} = \boldsymbol{p}$ and prove that $\boldsymbol{v}^+$ is the shortest one. *(10 points)*

Proof. (a) First compute the matrix:

$$A^TA = \begin{bmatrix} 5 & -7 & -2 \\ -7 & 14 & 7 \\ -2 & 7 & 5 \end{bmatrix}$$

and its characteristic polynomial $-\lambda^3 + 24\lambda^2 - 63\lambda = -\lambda(\lambda-21)(\lambda-3)$. Therefore, the eigenvalues of A^TA are $\lambda_1 = 21$, $\lambda_2 = 3$ and $\lambda_3 = 0$. Its eigenvectors (scaled so that they have unit length) are:

$$\boldsymbol{v}_1 = \frac{1}{\sqrt{6}}\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \qquad \boldsymbol{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \qquad \boldsymbol{v}_3 = \frac{1}{\sqrt{3}}\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \tag{2}$$

As for A, it has rank 2 and its singular values are $\sigma_1 = \sqrt{21}$ and $\sigma_2 = \sqrt{3}$. Its left singular vectors, and to compute the right singular vectors, we must apply the matrix A to the vectors (2):

$$\boldsymbol{u}_1 = \frac{1}{\sigma_1} A\boldsymbol{v}_1 = \frac{1}{\sqrt{14}} \begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{u}_2 = \frac{1}{\sigma_2} A\boldsymbol{v}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

Because $A\boldsymbol{v}_3 = 0$, we may just pick $\boldsymbol{u}_3$ so that it is orthonormal to $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ above. To find such a $\boldsymbol{u}_3$, just start from an arbitrary vector in $\mathbb{R}^3$ and perform Gram-Schmidt with respect to $\boldsymbol{u}_1$, $\boldsymbol{u}_2$:

$$\boldsymbol{u}_3 = \frac{1}{\sqrt{21}} \begin{bmatrix} -4 \\ -2 \\ 1 \end{bmatrix}$$

So we conclude that:

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

where:

$$U = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3 \end{bmatrix} = \frac{1}{\sqrt{42}} \begin{bmatrix} -\sqrt{3} & \sqrt{7} & -4\sqrt{2} \\ 3\sqrt{3} & -\sqrt{7} & -2\sqrt{2} \\ 2\sqrt{3} & 2\sqrt{7} & \sqrt{2} \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sqrt{21} & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$V = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 & \sqrt{3} & -\sqrt{2} \\ 2 & 0 & -\sqrt{2} \\ 1 & \sqrt{3} & \sqrt{2} \end{bmatrix}$$

Then the pseudo-inverse is:

$$A^+ = V\Sigma^+U^T = \frac{1}{21} \begin{bmatrix} 4 & -5 & 6 \\ -1 & 3 & 2 \\ 3 & -2 & 8 \end{bmatrix}$$

(b) The answer is $\boldsymbol{v}^+ = A^+\boldsymbol{b} = \frac{1}{21} \begin{bmatrix} 5 \\ 4 \\ 9 \end{bmatrix}$.

(c) All other solutions $\boldsymbol{v}$ to $A\boldsymbol{v} = \boldsymbol{p} = A\boldsymbol{v}^+$ have the property that:

$$\boldsymbol{v} = \boldsymbol{v}^+ + \boldsymbol{w}$$

for some $\boldsymbol{w}$ in the nullspace of A. The nullspace of A is one-dimensional and spanned by the eigenvector $\boldsymbol{v}_3$. Therefore, the answer is:

$$\boldsymbol{v} = \frac{1}{21} \begin{bmatrix} 5 \\ 4 \\ 9 \end{bmatrix} + \alpha \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

for all numbers α. Since $\boldsymbol{v}^+ \perp \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$, we have:

$$||\boldsymbol{v}||^2 = ||\boldsymbol{v}^+||^2 + \alpha^2 \cdot 3$$

which implies that $\boldsymbol{v}^+$ is shorter than $\boldsymbol{v}$.

$\square$

Problem 5: (a) Compute the sixth roots of unity (i.e. the complex number z such that $z^6 = 1$) in both Cartesian (i.e. $a + bi$) and polar (i.e. $re^{i\theta}$) form. Draw them all on a picture of the plane. *(10 points)*

(b) Prove the double angle and triple angle formulas:

$$\cos(2\theta) = 2(\cos\theta)^2 - 1 \qquad \text{and} \qquad \cos(3\theta) = 4(\cos\theta)^3 - 3\cos\theta \tag{3}$$

by the following logic:

- think of $\cos\theta$ as the real part of the complex number $z = e^{i\theta} = a+bi$ where $a = \cos\theta$, $b = \sin\theta$

- then compute z^2 (respectively z^3) first in polar form, and

- finally compute z^2 (respectively z^3) in Cartesian form

By equating the results in the last two bullets, you should obtain (3). *(10 points)*

Proof. (a) In polar form, if $z = re^{i\theta}$ satisfies $z^6 = 1$, then:

$$1 = r^6 e^{i6\theta}$$

By taking absolute values in the equality above, we get $1 = r^6$, and because r is a positive real number this implies that $r = 1$. Meanwhile, by taking angular parts in the equality above, we must have:

$$6\theta = \text{integer multiple of } 2\pi$$

because only the integer multiples of 2π have the same angular coordinate as the number 1. Hence:

$$\theta \in \left\{0, \frac{2\pi}{6}, \frac{4\pi}{6}, \frac{6\pi}{6}, \frac{8\pi}{6}, \frac{10\pi}{6}\right\}$$

The reason why we don't take other multiples of $\frac{2\pi}{6}$ is that from $\frac{12\pi}{6}$, they start repeating themselves with period 2π, hence we would just be getting the same angles over and over again. Therefore, the answer is:

$$z \in \left\{1, e^{\frac{\pi}{3}}, e^{\frac{2\pi}{3}}, e^{\frac{3\pi}{3}}, e^{\frac{4\pi}{3}}, e^{\frac{5\pi}{3}}\right\}$$

in polar form, or equivalently:

$$z \in \left\{1, \frac{1+i\sqrt{3}}{2}, \frac{-1+i\sqrt{3}}{2}, -1, \frac{-1-i\sqrt{3}}{2}, \frac{1-i\sqrt{3}}{2}\right\}$$

Geometrically, the 6 points corresponding to the 6 complex numbers above are the vertices of a regular hexagon on the unit circle (radius 1 centered at the origin) which has 1 as one of its vertices.

(b) Let $z = e^{i\theta} = a + bi$ where $a = \cos\theta$ and $b = \sin\theta$. On one hand, we have:

$$z^2 = e^{i2\theta} = \cos(2\theta) + i\sin(2\theta) \qquad \text{and} \qquad z^3 = e^{i3\theta} = \cos(3\theta) + i\sin(3\theta)$$

But on the other hand, in Cartesian coordinates, we have:

$$z^2 = (a+bi)^2 = a^2 - b^2 + 2abi \qquad \text{and} \qquad z^3 = (a+bi)^3 = a^3 - 3ab^2 + (3a^2b - b^3)i$$

By equating the real parts in the two equations above, we get:

$$\cos(2\theta) = a^2 - b^2 = (\cos\theta)^2 - (\sin\theta)^2 \qquad \text{and} \qquad \cos(3\theta) = (\cos\theta)^3 - 3(\cos\theta)(\sin\theta)^2$$

If you substitute $(\sin\theta)^2 = 1 - (\cos\theta)^2$, you get precisely the double/triple angle formulas.

$\square$