1. **Some basic properties of eigenvalues.** Show the following:
   
   a) The eigenvalues of $A$ and $A^T$ are the same.
   
   b) $A$ is invertible if and only if $A$ does not have a zero eigenvalue.
   
   c) If the eigenvalues of $A$ are $\lambda_1, \ldots, \lambda_n$ and $A$ is invertible, then the eigenvalues of $A^{-1}$ are $1/\lambda_1, \ldots, 1/\lambda_n$.
   
   d) The eigenvalues of $A$ and $T^{-1}AT$ are the same.

   *Hint:* you’ll need to use the facts that $\det A = \det(A^T)$, $\det(AB) = \det A \det B$, and, if $A$ is invertible, $\det A^{-1} = 1/\det A$.

2. **Detecting linear relations.** Suppose we have $N$ measurements $y_1, \ldots, y_N$ of a vector signal $x_1, \ldots, x_N \in \mathbb{R}^n$: 
   
   $$y_i = x_i + d_i, \ i = 1, \ldots, N.$$ 

   Here $d_i$ is some small measurement or sensor noise. We hypothesize that there is a linear relation among the components of the vector signal $x$, i.e., there is a nonzero vector $q$ such that $q^T x_i = 0, \ i = 1, \ldots, N$. The geometric interpretation is that all of the vectors $x_i$ lie in the hyperplane $q^T x = 0$. We will assume that $\|q\| = 1$, which does not affect the linear relation. Even if the $x_i$’s do lie in a hyperplane $q^T x = 0$, our measurements $y_i$ will not; we will have $q^T y_i = q^T d_i$. These numbers are small, assuming the measurement noise is small. So the problem of determining whether or not there is a linear relation among the components of the vectors $x_i$ comes down to finding out whether or not there is a unit-norm vector $q$ such that $q^T y_i, \ i = 1, \ldots, N$, are all small. We can view this problem geometrically as well. Assuming that the $x_i$’s all lie in the hyperplane $q^T x = 0$, and the $d_i$’s are small, the $y_i$’s will all lie close to the hyperplane. Thus a scatter plot of the $y_i$’s will reveal a sort of flat cloud, concentrated near the hyperplane $q^T x = 0$. Indeed, for any $z$ and $\|q\| = 1$, $|q^T z|$ is the distance from the vector $z$ to the hyperplane $q^T x = 0$. So we seek a vector $q$, $\|q\| = 1$, such that all the measurements $y_1, \ldots, y_N$ lie close to the hyperplane $q^T x = 0$ (that is, $q^T y_i$ are all small). How can we determine if there is such a vector, and what is its value? We define the following normalized measure:

   $$\rho = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (q^T y_i)^2} / \sqrt{\frac{1}{N} \sum_{i=1}^{N} \|y_i\|^2}.$$ 

   This measure is simply the ratio between the *root mean square distance* of the vectors to the hyperplane $q^T x = 0$ and the *root mean square length* of the vectors. If $\rho$ is small, it means that the measurements lie close to the hyperplane $q^T x = 0$. Obviously, $\rho$ depends on $q$. Here is the problem: explain how to find the minimum value of $\rho$ over all unit-norm vectors $q$, and the unit-norm vector $q$ that achieves this minimum, given the data set $y_1, \ldots, y_N$. 


3. Properties of symmetric matrices. In this problem $P$ and $Q$ are symmetric matrices. For each statement below, either give a proof or a specific counterexample.

a) If $P \geq 0$ then $P + Q \geq Q$.

b) If $P \geq Q$ then $-P \leq -Q$.

c) If $P > 0$ then $P^{-1} > 0$.

d) If $P \geq Q > 0$ then $P^{-1} \leq Q^{-1}$.

e) If $P \geq Q$ then $P^2 \geq Q^2$.

Hint: you might find it useful for part (d) to prove $Z \geq I$ implies $Z^{-1} \leq I$.

4. Drawing a graph. We consider the problem of drawing (in two dimensions) a graph with $n$ vertices (or nodes) and $m$ undirected edges (or links). This just means assigning an $x$- and a $y$-coordinate to each node. We let $x \in \mathbb{R}^n$ be the vector of $x$-coordinates of the nodes, and $y \in \mathbb{R}^n$ be the vector of $y$-coordinates of the nodes. When we draw the graph, we draw node $i$ at the location $(x_i, y_i) \in \mathbb{R}^2$. The problem, of course, is to make the drawn graph look good. One goal is that neighboring nodes on the graph (i.e., ones connected by an edge) should not be too far apart as drawn. To take this into account, we will choose the $x$- and $y$-coordinates so as to minimize the objective

$$J = \sum_{i<j, i \sim j} \left( (x_i - x_j)^2 + (y_i - y_j)^2 \right),$$

where $i \sim j$ means that nodes $i$ and $j$ are connected by an edge. The objective $J$ is precisely the sum of the squares of the lengths (in $\mathbb{R}^2$) of the drawn edges of the graph. We have to introduce some other constraints into our problem to get a sensible solution. First of all, the objective $J$ is not affected if we shift all the coordinates by some fixed amount (since $J$ only depends on differences of the $x$- and $y$-coordinates). So we can assume that

$$\sum_{i=1}^{n} x_i = 0, \quad \sum_{i=1}^{n} y_i = 0,$$

i.e., the sum (or mean value) of the $x$- and $y$-coordinates is zero. These two equations ‘center’ our drawn graph. Another problem is that we can minimize $J$ by putting all the nodes at $x_i = 0, y_i = 0$, which results in $J = 0$. To force the nodes to spread out, we impose the constraints

$$\sum_{i=1}^{n} x_i^2 = 1, \quad \sum_{i=1}^{n} y_i^2 = 1, \quad \sum_{i=1}^{n} x_i y_i = 0.$$

The first two say that the variance of the $x$- and $y$-coordinates is one; the last says that the $x$- and $y$-coordinates are uncorrelated. (You don’t have to know what variance or uncorrelated mean; these are just names for the equations given above.) The three equations above enforce ‘spreading’ of the drawn graph. Even with these constraints, the coordinates that minimize
J are not unique. For example, if x and y are any set of coordinates, and \( Q \in \mathbb{R}^{2 \times 2} \) is any orthogonal matrix, then the coordinates given by

\[
\begin{bmatrix}
\tilde{x}_i \\
\tilde{y}_i
\end{bmatrix} = Q \begin{bmatrix} x_i \\
y_i \end{bmatrix}, \quad i = 1, \ldots, n
\]

satisfy the centering and spreading constraints, and have the same value of J. This means that if you have a proposed set of coordinates for the nodes, then by rotating or reflecting them, you get another set of coordinates that is just as good, according to our objective. We’ll just live with this ambiguity. Here’s the question:

a) Explain how to solve this problem, i.e., how to find \( x \) and \( y \) that minimize \( J \) subject to the centering and spreading constraints, given the graph topology. You can use any method or ideas we’ve encountered in the course. Be clear as to whether your approach solves the problem exactly (i.e., finds a set of coordinates with \( J \) as small as it can possibly be), or whether it’s just a good heuristic (i.e., a choice of coordinates that achieves a reasonably small value of \( J \), but perhaps not the absolute best). In describing your method, you may not refer to any programming commands or operators; your description must be entirely in mathematical terms.

b) Implement your method, and carry it out for the graph given in \textbf{dg_data.json}. This JSON file contains the node adjacency matrix of the graph, denoted \( A \), and defined as \( A_{ij} = 1 \) if nodes \( i \) and \( j \) are connected by an edge, and \( A_{ij} = 0 \) otherwise. (The graph is undirected, so \( A \) is symmetric. Also, we do not have self-loops, so \( A_{ii} = 0 \), for \( i = 1, \ldots, n \).) Give the value of \( J \) achieved by your choice of \( x \) and \( y \), and verify that your \( x \) and \( y \) satisfy the centering and spreading conditions, at least approximately. If your method is iterative, plot the value of \( J \) versus iteration. Draw the corresponding graph by plotting nodes as small circles and edges as lines. For comparison, the JSON file also contains the vectors \( x_{\text{circ}} \) and \( y_{\text{circ}} \). These coordinates were obtained using a standard technique for drawing a graph, by placing the nodes in order on a circle. The radius of the circle has been chosen so that \( x_{\text{circ}} \) and \( y_{\text{circ}} \) satisfy the centering and spread constraints. Draw this graph on a separate plot.

**Hint.** You are welcome to use the results described below, without proving them. Let \( A \in \mathbb{R}^{n \times n} \) be symmetric, with eigenvalue decomposition \( A = \sum_{i=1}^{n} \lambda_i q_i q_i^\top \), with \( \lambda_1 \geq \cdots \geq \lambda_n \), and \( \{q_1, \ldots, q_n\} \) orthonormal. You know that a solution of the problem

\[
\begin{align*}
\text{minimize} & \quad x^\top Ax \\
\text{subject to} & \quad x^\top x = 1,
\end{align*}
\]

where the variable is \( x \in \mathbb{R}^n \), is \( x = q_n \). The related maximization problem is

\[
\begin{align*}
\text{maximize} & \quad x^\top Ax \\
\text{subject to} & \quad x^\top x = 1
\end{align*}
\]

with variable \( x \in \mathbb{R}^n \). A solution to this problem is \( x = q_1 \). Now consider the following generalization of the first problem:

\[
\begin{align*}
\text{minimize} & \quad \text{trace}(X^\top AX) \\
\text{subject to} & \quad X^\top X = I_k
\end{align*}
\]
where the variable is \( X \in \mathbb{R}^{n \times k} \), and \( I_k \) denotes the \( k \times k \) identity matrix, and we assume \( k \leq n \). The constraint means that the columns of \( X \), say, \( x_1, \ldots, x_k \), are orthonormal; the objective can be written in terms of the columns of \( X \) as \( \text{trace}(X^TAX) = \sum_{i=1}^{k} x_i^T Ax_i \). A solution of this problem is \( X = [q_{n-k+1} \cdots q_n] \). Note that when \( k = 1 \), this reduces to the first problem above. The related maximization problem is

\[
\begin{align*}
\text{maximize} & \quad \text{trace}(X^TAX) \\
\text{subject to} & \quad X^TX = I_k
\end{align*}
\]

with variable \( X \in \mathbb{R}^{n \times k} \). A solution of this problem is \( X = [q_1 \cdots q_k] \).

5. Set descriptions. Let \( L \) be the line through the origin in the direction \( z \) where

\[
z = \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix}
\]

That is, \( L = \{ \lambda z \mid \lambda \in \mathbb{R} \} \). Let \( S \subset \mathbb{R}^3 \) be the set of all points a distance 2 or less from \( L \).

a) find a matrix \( Q = Q^T \) for which

\[
S = \{ x \in \mathbb{R}^3 \mid x^TQx \leq 1 \}
\]

b) Is \( Q \) unique? Give a proof or counterexample.

c) The sum of two sets \( A, B \subset \mathbb{R}^n \) is defined to be

\[
A + B = \{ a + b \mid a \in A, b \in B \}
\]

that is, the set of all possible sums \( a + b \) where \( a \in A \) and \( b \in B \). Let \( E \) be the ellipsoid

\[
E = \{ x \in \mathbb{R}^3 \mid x^TWx \leq 1 \}
\]

where

\[
W = \begin{bmatrix} 5 & 4 & 1 \\ 4 & 5 & 2 \\ 1 & 2 & 3 \end{bmatrix}
\]

Find the matrix \( P \) for which

\[
E + L = \{ x \in \mathbb{R}^3 \mid x^TPx \leq 1 \}
\]

d) Suppose \( M \) is the line \( M = \{ \lambda w \mid \lambda \in \mathbb{R} \} \). where

\[
w = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}
\]

Find a matrix \( R \) such that the ellipsoid

\[
F = \{ x \in \mathbb{R}^3 \mid x^TRx \leq 1 \}
\]

satisfies

\[
F = \{ x \in \mathbb{R}^3 \mid \text{dist}(x, L)^2 + \text{dist}(x, M)^2 \leq 1 \}
\]

Here dist\((x, L)\) is the distance between \( x \) and the closest point in \( L \), that is

\[
\text{dist}(x, L) = \min_{y \in L}\|x - y\|
\]
6. **Square matrices and the SVD.** Let $A$ be an $n \times n$ real matrix. State whether each of the following statements is true or false. Do not give any explanation or show any work.

a) If $x$ is an eigenvector of $A$, then $x$ is either a left or right singular vector of $A$

b) If $\lambda$ is an eigenvalue of $A$, then $|\lambda|$ is a singular value

c) If $A$ is symmetric, then every singular value of $A$ is also an eigenvalue of $A$

d) If $A$ is symmetric, then every singular vector of $A$ is also an eigenvector of $A$

e) If $A$ is symmetric with the following singular value decomposition
\[
A = U\Sigma V^T
\]
then $U = V$

f) If $A$ is invertible, then
\[
\sigma_i \neq 0 \quad \text{for all } i = 1, \ldots, n
\]