Homework #1

Problem 1

$$A = \begin{bmatrix} 1 & 0.5 \\ 0 & 1+\epsilon \end{bmatrix}$$

a) Find the eigenvalues/eigenvectors of A.

Let $det(\lambda I - A) = 0$
$(\lambda - 1)(\lambda - (1+\epsilon)) = 0$
$\lambda_1 = 1, \lambda_2 = 1+\epsilon$

For λ_1:
$AX_1 = X_1$
$X_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$
$Y_1^T A = Y_1^T$
$Y_1 = \frac{1}{\sqrt{\epsilon^2+0.25}} \begin{bmatrix} \epsilon \\ -0.5 \end{bmatrix}$

For λ_2:
$AX_2 = (1+\epsilon)X_2$
$X_2 = \frac{1}{\sqrt{\epsilon^2+0.25}} \begin{bmatrix} 0.5 \\ \epsilon \end{bmatrix}$
$Y_2^T A = Y_2^T$
$Y_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

b) Diagonalize A.

$T = \begin{bmatrix} 1 & \frac{0.5}{\sqrt{\epsilon^2+0.25}} \\ 0 & \frac{\epsilon}{\sqrt{\epsilon^2+0.25}} \end{bmatrix}$
$T^{-1} = \begin{bmatrix} \frac{\epsilon}{\sqrt{\epsilon^2+0.25}} & \frac{-0.5}{\sqrt{\epsilon^2+0.25}} \\ 0 & 1 \end{bmatrix}$

$$T^{-1}AT = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1+\epsilon \end{bmatrix}$$

$\epsilon \to 0 : \frac{0.5}{\sqrt{\epsilon^2+0.25}} \to 1, \frac{\epsilon}{\sqrt{\epsilon^2+0.25}} \to 0, \frac{\epsilon}{\sqrt{\epsilon^2+0.25}} \to 0, \frac{-0.5}{\sqrt{\epsilon^2+0.25}} \to -1$
$\Rightarrow T = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, T^{-1} = \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}$

Therefore, when $\epsilon = 0$, A is not diagonalizable

Problem 2

a) trace(ABT) = trace(BTA) = trace(BAT) = trace(ATB)

$trace(AB^T) = \sum_{i=1}^{k}\sum_{j=1}^{m} a_{ij}b_{ij}$
$trace(B^TA) = \sum_{i=1}^{m}\sum_{j=1}^{k} b_{ji}a_{ji}$
$trace(BA^T) = \sum_{i=1}^{k}\sum_{j=1}^{m} b_{ij}a_{ij}$
$trace(A^TB) = \sum_{i=1}^{m}\sum_{j=1}^{k} a_{ji}b_{ji}$
Therefore, $trace(AB^T) = trace(B^TA) = trace(BA^T) = trace(A^TB)$

b) E[xTAx]

$E(x^TAx) = E(trace(x^TAx)) = E(trace(Axx^T))$
Consider $trace(AB^T) = \sum_{i=1}^{k}\sum_{j=1}^{m} a_{ij}b_{ij}$:
$E(trace(AB^T)) = E(\sum_{i=1}^{k}\sum_{j=1}^{m} a_{ij}b_{ij}) = \sum_{i=1}^{k}\sum_{j=1}^{m} E(a_{ij}b_{ij})$, because $E(X+Y) = E(X)+E(Y)$ when random variable X and Y are independent.
Hence, $E(x^TAx) = trace(E(Axx^T)) = trace(A \times E(xx^T))$
$E(xx^T) = Q \Rightarrow E(x^TAx) = trace(AQ)$

c) trace(A) = trace(UAU−1)

$$trace(UAU^{-1}) = trace(AU^{-1}U) = trace(AI) = \sum_{i=1}^{k}\sum_{j=1}^{m} a_{ij} = trace(A)$$

d) if matrix A of dimensions k×k is diagonalizable then its trace is equal to the sum of its eigenvalues

$$trace(A) = trace(TAT^{-1}) = trace(\Lambda) = \sum_{i=1}^{k} \lambda_i$$

e) how do you explain this equality given that when A is real the trace is also real whereas the eigenvalues can be complex?

For each $\lambda_k = x + yi$, where $y \neq 0$, there must be another $\lambda_{k'} = x - yi$, i.e. $\lambda_k, \lambda_{k'}$ are conjugated.
Actually, $\lambda_k, \lambda_{k'}$ are a pair of conjugate roots of the characteristic polynomial.

f) what can you say about the coefficient ck−1 of the characteristic polynomial

$$c_{k-1} = -\sum_{i=1}^{k} \lambda_i = -trace(\Lambda) = -trace(A)$$

Problem 3

Ar = 0 for some integer r > 1

a) all eigenvalues of A must be equal to 0.

$$det(\lambda I - A^r) = 0$$
$$\lambda_i' = 0, i = 1, 2, ...$$
$$\lambda_i' = \lambda_i^r \Rightarrow \lambda_i^r = \lambda_i = 0$$

b) Is such a matrix diagonalizable?

If $A = 0$, A itself has been diagonalized.
If $A \neq 0$, A cannot be diagonalizable. Prove:
Assume $A \neq 0$ is diagonalizable, i.e. $\exists T, T^-1, A = T\Lambda T^{-1}$
$\lambda_i = 0$, where $i = 1, 2, ... \Rightarrow \Lambda = 0 \Rightarrow T\Lambda T^{-1} = A = 0$
Contradiction!
Hence, A is diagonalizable if and only if $A = 0$

c) Give an example of a 2×2 matrix

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$
where $r = 2$:
$$A^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

d) If we apply a similarity transformation to a nilpotent is the resulting matrix nilpotent?

Yes.
If $P^{-1}AP = B$:
$(P^{-1}AP)^r = P^{-1}A^rP = 0 = B^r$

Say $P = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$,

$$P^{-1} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{5} \end{bmatrix}$$

$$B = P^{-1}AP = \begin{bmatrix} 0 & \frac{5}{2} \\ 0 & 0 \end{bmatrix}$$

Problem 4

Denote with Qi the matrix we obtain from Q by eliminating its ith column and ith row.

a) Show that Qi is also symmetric and positive definite.

$$Q = \begin{bmatrix} A_{(i-1)\times(i-1)} & E_{(i-1)\times 1} & B_{(i-1)\times(k-i)} \\ F_{1\times(i-1)} & q_{ii} & G_{1\times(k-i)} \\ C_{(k-i)\times(i-1)} & H_{(k-i)\times 1} & D_{(k-i)\times(k-i)} \end{bmatrix}$$

$Q^T = Q \Rightarrow A^T = A, D^T = D, B^T = C$
$$Q_i = \begin{bmatrix} A_{(i-1)\times(i-1)} & B_{(i-1)\times(k-i)} \\ C_{(k-i)\times(i-1)} & D_{(k-i)\times(k-i)} \end{bmatrix}$$
Therefore, $Q_i^T = Q_i$

Consider $x^T = \begin{bmatrix} x_1 & x_2 & \ldots & x_i & x_{i+1} & \ldots & x_k \end{bmatrix}$
For $\forall x \neq 0, x^T Q x > 0$
Denot with x' the vector we obtain form x by eliminating x_i
$x'^T Q_i x' = x^T Q x > 0$, when $x_i = 0$
Therefore, Q_i is also symmetric and positive definite.

Since Q_i is symmetric and positive definite, Assign $Q' = Q_i$. Now $Q'_j = Q_{ij}$ is also symmetric and positive definite.
By Mathematical Induction, we can eliminate more than one columns and rows till the remaining is a one by one matrix, or say scalar.

b) Show that the largest eigenvalue of Q is larger than the largest eigenvalue of Qi and, that the smallest eigenvalue of Q is smaller than the smallest eigenvalue of Qi

Assume there is a eigenvalue of Q_i, say λ' is larger than all eigenvalue of Q, and $Q_i x' = \lambda' x'$.
Hence, $x^T Q x = x'^T Q_i x' = \lambda' x'^T x' = \lambda' x^T x$, where $x_i = 0, i.e. \frac{x^T Q x}{x^T x} = \lambda' > \max\limits_i(\lambda_i)$
Contradiction!
Hence, the largest eigenvalue of Q is larger than or equal to the largest eigenvalue of Q_i

Similarly, assume λ'' is smaller than all eigenvalue of Q. There is another contradiction that $\frac{x^T Q x}{x^T x} = \lambda'' > \min\limits_i(\lambda_i)$.
Hence, the smallest eigenvalue of Q is smaller than or equal to the smallest eigenvalue of Q_i

c) min iσi ≤ min i |λi|≤ max i |λi|≤ max iσi.

To show $\min\limits_i \sigma_i \leq \min\limits_i |\lambda_i| \leq \max\limits_i |\lambda_i| \leq \max\limits_i \sigma_i$ is to show $\min\limits_i \lambda(A^T A)_i \leq \min\limits_i |\lambda(A)_i|^2 \leq \max\limits_i |\lambda(A)_i|^2 \leq \max\limits_i \lambda(A^T A)_i$, because $0 \leq \min\limits_i \sigma_i \leq \min\limits_i |\lambda_i| \leq \max\limits_i |\lambda_i| \leq \max\limits_i \sigma_i$

Consider $\min\limits_i |\lambda(A)|_i^2 \leq \max\limits_i |\lambda(A)|_i^2$ by defination. What we need show is $\min\limits_i \lambda(A^T A)_i \leq \min\limits_i |\lambda(A)|_i^2$ and $\max\limits_i |\lambda(A)|_i^2 \leq \max\limits_i \lambda(A^T A)_i$

Consider $f(x) = \frac{(x^*)^T A^T A x}{(x^*)^T x}$:
Let $U^T A^T A U = \Lambda$, where $U^T U = I$:
Let $y = Ux$:

$f(x) = f(y) = \frac{(y^*)^T \Lambda y}{(y^*)^T y} = \frac{\sum_{i=1}^{k} \lambda(A^T A)_i |y_i|^2}{\sum_{i=1}^{k} |y_i|^2}$

Therefore, $\min\limits_x f(x) = \min\limits_i \lambda(A^T A)_i$ and $\max\limits_x f(x) = \max\limits_i \lambda(A^T A)_i$, i.e.

$\min\limits_i \lambda(A^T A)_i \le f(x) \le \max\limits_i \lambda(A^T A)_i$

Let x_i be the eigenvector of $\lambda(A)_i$

Therefore, $f(x_i) = \frac{(x_i^*)^T A^T A x_i}{(x_i^*)^T x_i} = \frac{((Ax_i)^*)^T (Ax_i)}{|x_i|^2} = \frac{|\lambda(A)_i|^2 |x_i|^2}{|x_i|^2} = |\lambda(A)_i|^2$, i.e.

$\min\limits_i \lambda(A^T A)_i \le |\lambda(A)_i|^2 \le \max\limits_i \lambda(A^T A)_i$, for $\forall i = 1, 2, ..., k$

Therefore, $\min\limits_i \lambda(A^T A)_i \le \min\limits_i |\lambda(A)|_i^2 \le \max\limits_i |\lambda(A)|_i^2 \le \max\limits_i \lambda(A^T A)_i$