AMSC 660 Homework 3
Due 09/20/23
By Marvyn Bailly

Problem 1

Let A be an $n \times n$ matrix. The Rayleigh quotient $Q(x)$ is the following function defined on all $x \in \mathbb{R}^n$:

$$Q(x) := \frac{x^T A x}{x^T x}$$

(a) Let A be symmetric. Prove that $\nabla Q(x) = 0$ if and only if x is an eigenvector of A.

(b) Let A be asymmetric. What are the vectors x at which $\nabla Q(x) = 0$?

Solution

Proof. Let A be an $n \times n$ matrix. Recall that the Rayleigh quotient $Q(x)$ is the following function defined on all $x \in \mathbb{R}^n$

$$Q(x) := \frac{x^T A x}{x^T x}.$$

Let's first compute $\nabla Q(x)$. Let $N = x^T A x$ and $D = x^T x$. Then

$$\frac{\partial}{\partial x} D = \frac{\partial}{\partial x_k}\left(\sum_{k=1}^n x_k^2\right) = 2x.$$

To compute $\frac{\partial}{\partial x} N$, observe that $\frac{\partial}{\partial x_1} N$ is given by

$$\begin{aligned}
\frac{\partial}{\partial x_1}(N) &= \frac{\partial}{\partial x_1} \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_j x_i \\
&= \frac{\partial}{\partial x_1} \sum_{j=1}^n a_{1j} x_j x_1 + \sum_{i=1}^n a_{i1} x_1 x_i + \sum_{i=2}^n \sum_{j=2}^n a_{ij} x_j x_i \\
&= \sum_{j=1}^n a_{1j} x_j + \sum_{i=1}^n a_{i1} x_i.
\end{aligned}$$

Extending this yields

$$\frac{\partial}{\partial x}(N) = \begin{pmatrix} \sum_{j=1}^n a_{1j} x_j + \sum_{i=1}^n a_{i1} x_i \\ \vdots \\ \sum_{j=1}^n a_{nj} x_j + \sum_{i=1}^n a_{in} x_i \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j} x_j \\ \vdots \\ \sum_{j=1}^n a_{nj} x_j \end{pmatrix} + \begin{pmatrix} \sum_{j=1}^n a_{i1} x_j \\ \vdots \\ \sum_{j=1}^n a_{in} x_j \end{pmatrix} = \left(A + A^T\right)x.$$

Then we can compute

$$\nabla Q(x) = \frac{N'D - D'N}{D^2}.$$

(a) Assume that A is symmetric. Notice that

$$\frac{\partial}{\partial x}(N) = (A + A^T)x = 2Ax.$$

Now we can compute

$$\begin{aligned}\nabla Q(x) &= \frac{N'D - D'N}{D^2} \\ &= \frac{2Axx^T x - 2xx^T Ax}{(x^T x)^2} \\ &= \frac{2}{\|x\|_2^4}\big(Axx^T x - xx^T Ax\big).\end{aligned}$$

We want to show that $\nabla Q(x) = 0$ if and only if x is an eigenvector of A. It suffices to show that

$$Axx^T x = xx^T Ax \iff Ax = \lambda x,$$

where λ is the corresponding eigenvalue of x. Notice that $x^T x$ and $x^T Ax$ are scalars and by the Rayleigh Quotient, we are considering $x \in \mathbb{R}^n$ such that $x^T x, \neq 0$. Thus we can rearrange terms to get

$$Ax(x^T x) = (x^T Ax)x \iff Ax = \frac{x^T Ax}{x^T x}x \iff Ax = \lambda x,$$

where $\lambda = \frac{x^T Ax}{x^T x}$. Therefore $\nabla Q(x) = 0$ if and only if x is an eigenvector of A.

(b) Assume that A is asymmetric. Observe that

$$\nabla Q(x) = \frac{N'D - D'N}{D^2} = \frac{(A + A^T)xx^T x - 2xx^T Ax}{(x^T x)^2}.$$

We want to find the vectors x such that $\nabla Q(x) = 0$. Thus let's find the vectors that satisfy

$$(A + A^T)xx^T x = 2xx^T Ax \iff \frac{(A + A^T)}{2}x = \frac{x^T Ax}{x^T x}x,$$

notice that $\frac{A+A^T}{2} = B$ is the symmetric decomposition of A. Now if we let $\lambda = \frac{x^T Ax}{x^T x}$, we have

$$Bx = \lambda x,$$

and thus $\nabla Q(x) = 0$ is achieved if there exists an eigenvector x of $\frac{A+A^T}{2}$ that has a corresponding eigenvalue such that $\lambda = \frac{x^T Ax}{x^T x}$.

$\square$

Problem 2

The goal of this exercise is to understand how one can compute a QR decomposition using *Householder reflections*.

(a) Let u be a unit vector in $\mathbb{R}^n$, i.e. $\|u\|_2 = 1$. Let $P = I - 2uu^T$. This matrix performs reflections with respect to the hyperplane orthogonal to the vector u. Show that $P = P^T$ and $P^2 = I$.

(b) Let $x \in \mathbb{R}^n$ be any vector, $x = [x_1, \dots, x_n]^T$. Let u be defined as follows:

$$\tilde{u} := \begin{pmatrix} x_1 - \mathrm{sign}(x_1)\|x\|_2 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \equiv x - \mathrm{sign}(x_1)\|x\|_2 e_1, \quad u = \frac{\tilde{u}}{\|\tilde{u}\|_2},$$

where $e_1 = [1, 0, \dots, 0]^T$. The matrix with the vector u construct according to (1) will be denoted `House(x)`:

$$P = I - 2uu^T \equiv I - 2\frac{\tilde{u}\tilde{u}^T}{\tilde{u}^T\tilde{u}} \equiv \mathrm{House(x)}.$$

Calculate Px.

(c) Let A be an $m \times n$ matrix, $m \geq n$, with columns a_j, $j = 1, \dots, n$. Let $A_0 = A$. Let $P_1 = \mathrm{House}(a_1)$. Then $A_1 := P_1 A_0$ has the first column with the first entry nonzero and the other entries being zero. Next, we define P_2 as

$$P_2 = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{P}_2 \end{pmatrix},$$

where the matrix $\tilde{P}_2 = \mathrm{House}((A_1)(2:n, 2))$. The notation $A_1(2:n, 2)$ is Matlab's syntax indicating this is the vector formed by entries 2 through n of the 2nd column on $A_!$. Then we set $A_2 = P_2 A_1$. And so on. This algorithm can be described as follow. Let $A_0 = A$. Then for $j = 1, 2, \dots, n$ we set

$$P_j = \begin{pmatrix} I_{(j-1)\times(j-1)} & 0 \\ 0 & \tilde{P}_j \end{pmatrix}; \quad \tilde{P}_j = \mathrm{House}(A_{j-1}(j:n, j)), \quad A_j = P_j A_{j-1}.$$

Check that the resulting matrix A_n is upper triangular, its entries $(A)_{ij}$ are all zeros for $i > j$. Propose an `if`-statement in this algorithm that will guarantee that A_n has positive entries $(A_n)_{jj}, 1 \leq j \leq n$.

(d) Extract the QR decomposition of A given the matrices $P_j, 1 \leq j \leq n$, and A_n.

Solution

Proof. (a) Let u be a unit vector in $\mathbb{R}^n$. Let $P = I - 2uu^T$. Notice that

$$P^T = (I - 2uu^T)^T = I - 2uu^T = P,$$

and

$$P^2 = (I - 2uu^T)^2 = I^2 - 2uu^T - 2uu^T + 4uu^Tuu^T = I - 4uu^T + 4uu^T = I,$$

where $u^Tu = I$ since u is a unit vector and thus u^T and u are orthogonal.

(b) Let $x \in \mathbb{R}^n$ be any vector, $x = [x_1, \dots, x_n]^T$. Let u be defined as follows:

$$\tilde{u} := \begin{pmatrix} x_1 - \text{sign}(x_1)\|x\|_2 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \equiv x - \text{sign}(x_1)\|x\|_2 e_1, \quad u = \frac{\tilde{u}}{\|\tilde{u}\|_2},$$

where $e_1 = [1, 0, \dots, 0]^T$. The matrix with the vector u construct according to (1) will be denoted `House(x)`:

$$P = I - 2uu^T \equiv I - 2\frac{\tilde{u}\tilde{u}^T}{\tilde{u}^T\tilde{u}} \equiv \text{House(x)}.$$

Notice that

$$\begin{aligned}
Px &= (I - 2uu^T) \\
&= Ix - 2uu^Tx \\
&= x - \frac{2\tilde{u}\tilde{u}^Tx}{\tilde{u}^T\tilde{u}} \\
&= x - \frac{2(x - \text{sign}(x_1)\|x\|_2 e_1)(x^T - \text{sign}(x_1)\|x\|_2 e_1^T)x}{(x^T - \text{sign}(x_1)\|x\|_2 e_1^T)(x - \text{sign}(x_1)\|x\|_2 e_1)} \\
&= x - \frac{2(x - \text{sign}(x_1)\|x\|_2 e_1)(x^Tx - \text{sign}(x_1)\|x\|_2 e_1^Tx)}{x^Tx - \text{sign}(x_1)\|x\|_2 x^Te_1 - \text{sign}(x_1)\|x\|_2 e_1^Tx + \text{sign}(x)\text{sign}(x)\|x\|^2 e_1^Te_1} \\
&= x - \frac{2(x - \text{sign}(x_1)\|x\|_2 e_1)(x^Tx - \text{sign}(x_1)\|x\|_2 e_1^Tx)}{2(x^Tx - \text{sign}(x_1)\|x\|_2 x^Te_1)} \\
&= x - (x - \text{sign}(x_1)\|x\|_2 e_1) \\
&= \text{sign}(x_1)\|x\|_2 e_1,
\end{aligned}$$

where we used $\text{sign}(x)\text{sign}(x)\|x\|^2 e_1^Te_1 = \|x\|^2 = x^Tx$ and $x^Te_1 = e_1^Tx$. Notice that Px is a column vector with a nonzero entree in the first position and zeros everywhere else.

(c) Let A be an $m \times n$ matrix, $m \geq n$, with columns a_j, $j = 1, \ldots, n$. Let $A_0 = A$. Let $P_1 = \text{House}(a_1)$. Then $A_1 := P_1 A_0$ which gives

$$A_1 = P_1 A_0 = \left(\text{sign}(a_{11})\|a_1\|_2 e_1 \middle| \tilde{a_2} \middle| \cdots \middle| \tilde{a_n}\right),$$

which has the first column with the first entry nonzero and the other entries being zero. This directly follows from our observation in (b). Let $P_1 a_j = \tilde{a_j}$ for $j > 1$. Next, we define P_2 as

$$P_2 = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{P_2} \end{pmatrix},$$

where the matrix $\tilde{P_2} = \text{House}(A_1(2:n, 2))$. The notation $A_1(2:n, 2)$ is Matlab's syntax indicating this is the vector formed by entries 2 through n of the 2nd column on A_1. Then we set $A_2 = P_2 A_1$. Notice that $P_2 a_1 = P_2(\text{sign}(a_{11})\|a_1\|_2 e_1) = \tilde{a}_1 e_1$ and still has zeros below the first entree. Furthermore, $P_2 \tilde{a}_2 = \tilde{a}_{21} e_1 + \text{sign}(\tilde{a}_{22})\|\tilde{a}_2\| e_2$ and therefore the second column of A_2 has nonzero entries on and above the diagonal with zeros everywhere else. And so on each step, the updated $A_j = P_j A_{j-1}$ will zero the entrees below the diagonal of the jth column. Therefore the resulting A_n matrix will be upper triangular. To guarantee that the elements along the main diagonal are positive, we can impose the following `if`-statement: $\text{if}(\text{sign}(a_{ii}) < 0), \text{then set} \tilde{P_j} = -\tilde{P_j}$ which will force the elements along the main diagonal to be positive.

(d) From the algorithm described above, we have

$$A_n = P_n A_{n-1} = P_n P_{n-1} \cdots P_2 P_1 A = PA,$$

where A_n is upper triangular. Notice that by the construction of P_j, it is orthogonal and symmetric and thus so is P. Now letting $Q^T = P_n \cdots P_1$ then $Q = P_1^T \cdots P_n^T = P_1 \cdots P_n$. And letting $A_n = R$ we have

$$A = QR,$$

where Q is orthogonal and R is upper triangular.

$\square$

Problem 3

Prove items (1)-(6) of the following Theorem:

Let $A = U\Sigma V^T$ be the SVD of the $m \times n$ matrix $A, m \geq n$.

1. Suppose A is symmetric and $A = U\Lambda U^T$ be an eigendecomposition of A. Then the SVD of A is $U\Sigma V^T$ where $\sigma_i = |\lambda_i|$ and $v_i = u_i \text{sign}(\lambda_i)$, where $\text{sign}(0) = 1$.

2. The eigenvalues of the symmetric matrix A^TA are σ^2. The right singular vectors v_i are the corresponding orthonormal eigenvectors.

3. The eigenvalues of the symmetric matrix AA^T are σ_i^2 and $m-n$ zeros. The left singular vectors u_i are the corresponding orthonormal eigenvectors for the eigenvalues σ_i^2. One can take any $m-n$ orthogonal vectors as eigenvectors for the eigenvalue 0.

4. If A has full rank, the solution of
$$\min_x \|Ax - b\| \text{ is } x = V\Sigma^{-1}U^Tb.$$

5. If A is square and nonsingular, then
$$\|A^{-1}\|_2 = \frac{1}{\sigma_n}.$$

6. Suppose
$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0.$$
Then
$$\text{rank}(A) = r,$$
and
$$\text{null}(A) = \{x \in \mathbb{R}^n : Ax = 0 \in \mathbb{R}^m\} = \text{span}(v_{r+1}, \ldots, v_n), \ \text{range}(A) = \text{span}(u_1, \ldots, u_r).$$

Solution

Proof. Let $A = U\Sigma V^T$ be the SVD of the $m \times n$ matrix A, $m \geq n$.

1. Suppose A is symmetric and $A = U\Lambda U^T$ is the eigendecomposition of A where the columns of U are u_j. Let $\sigma_i = |\lambda_i|$, $\Sigma = \text{diag}(\sigma_i)$, and V be a matrix with columns $v_i = u_i \text{sign}(\lambda_i)$. Observe that
$$U\Sigma V^T = U \begin{pmatrix} |\lambda_1| & & \\ & \ddots & \\ & & |\lambda_n| \end{pmatrix} \begin{pmatrix} u_1 \text{sign}(\lambda_1) \\ \vdots \\ u_n \text{sign}(\lambda_1) \end{pmatrix}$$

$$
\begin{aligned}
&= U \begin{pmatrix} |\lambda_1|\text{sign}(\lambda_1) & & \\ & \ddots & \\ & & |\lambda_n|\text{sign}(\lambda_1) \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \\
&= U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^T \\
&= U \Lambda U^T.
\end{aligned}
$$

Since U is constructed from the eigenvectors of A, U and V are orthogonal. By construction Σ is diagonal. Thus the SVD of A is $U\Sigma V^T$ where $\sigma_i = |\lambda_i|$ and $v_i = u_i \text{sign}(\lambda_i)$.

2. Consider that $A^T A$ is symmetric and thus has an eigendecomposition of the form $A^T A = Q\Lambda Q^T$. By (1), the SVD of $A^T A$ is $Q\Sigma V^T$ where $\sigma_i = |\lambda_i|$ and $v_i = q_i \text{sign}(\lambda_i)$. Notice that the SVD of $A^T A$ is also given by

$$
\begin{aligned}
A^T A &= \left(U\Sigma V^T\right)^T \left(U\Sigma V^T\right) \\
&= V\Sigma U^T U \Sigma V^T \\
&= V\Sigma^2 V^T \\
&= V \begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{pmatrix} V^T,
\end{aligned}
$$

where $U^T U = I$ as U is orthogonal. So $A = Q\Lambda Q = V\Sigma V^T$. Thus the eigenvalues of $A^T A$ are σ_i^2 as $\Sigma = \Lambda$ and the right singular vectors v_i are the corresponding orthonormal eigenvectors as $Q = V$. Note, that since $A^T A$ is symmetric, the eigenvectors of $A^T A$ form an orthogonal basis which can be scaled to be orthonormal.

3. Consider that the AA^T is a symmetric $m \times m$ matrix and has an eigendecomposition of the form $AA^T = Q\Lambda Q^T$ where Q and Λ are $m \times m$ matrices. By (1), the SVD of AA^T is $Q\Sigma V^T$ where $\sigma_i = |\lambda_i|$ and $v_i = q_i \text{sign}(\lambda_i)$. Notice that the SVD of AA^T is also given by

$$
\begin{aligned}
AA^T &= \left(U\Sigma V^T\right) \left(U\Sigma V^T\right)^T \\
&= U\Sigma V^T V \Sigma U^T \\
&= U\Sigma^2 U^T \\
&= U \begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{pmatrix} U^T,
\end{aligned}
$$

where $V^TV = I$ as V is orthogonal, U is $m \times n$ and Σ is $n \times n$. To make the dimensions of the two forms of the SVD match, we can extend Σ by $m-n$ rows and columns of zeros and add $m-n$ columns of orthogonal vectors $\{\tilde{u_{m-n}}, \ldots, \tilde{u_m}\}$ (orthogonal to each other and to the columns of U) to get

$$AA^T = [u_1 \big| \cdots \big| u_n \big| \tilde{u}_{m-n} \big| \cdots \big| \tilde{u}_m] \begin{pmatrix} \sigma_1^2 & & & & & \\ & \ddots & & & & \\ & & \sigma_n^2 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} [u_1 \big| \cdots \big| u_n \big| \tilde{u}_{m-n} \big| \cdots \big| \tilde{u}_m]^T = \tilde{U}\tilde{\Sigma}\tilde{U}^T.$$

Now that the dimensions match, we have that $AA^T = Q\Lambda Q^T = \tilde{U}\tilde{\Sigma}\tilde{U}^T$. Thus the eigenvalues of the symmetric matrix AA^T are σ_i^2 and $m-n$ zeros as $\Lambda = \tilde{\Sigma}$. The left singular vectors u_i are the corresponding orthonormal eigenvectors for the eigenvalues σ_i^2 and one can take any $m-n$ orthogonal vectors as eigenvectors for the eigenvalue 0 as $Q = \tilde{U}$. Note, that since AA^T is symmetric, the eigenvectors of A^TA form an orthogonal basis which can be scaled to be orthonormal.

4. Assume that A has full rank. Recall that the minimizer of the least squares problem

$$\min_x \|Ax - b\|,$$

is given by the normal equation $A^TAx^* = A^Tb$ as A is full rank. Plugging in the SVD form of A yields

$$\begin{aligned} A^TAx^* &= A^Tb \\ V\Sigma^2V^Tx^* &= V\Sigma U^Tb \\ \Sigma^2V^Tx^* &= \Sigma U^Tb \\ V^Tx^* &= \Sigma^{-1}U^Tb \\ x^* &= V\Sigma^{-1}U^Tb. \end{aligned}$$

Therefore, the solution of

$$\min_x \|Ax - b\| \text{ is } x = V\Sigma^{-1}U^Tb.$$

5. Assume A is square and nonsingular. Recall that

$$\|A\|_2 = \max_i \sigma_i = \sigma_1.$$

Observe that

$$A^{-1} = (U\Sigma V^T)^{-1} = U^{-1}\Sigma^{-1}V^{-T} = U^{-1} \begin{pmatrix} \frac{1}{\sigma_1} & & \\ & \ddots & \\ & & \frac{1}{\sigma_n} \end{pmatrix} V^{-T},$$

and thus

$$\|A^1\| = \max_i \sigma_i = \frac{1}{\sigma_n}.$$

6. Suppose

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0.$$

Recall the geometric definition of an SVD. That is $Av_i = \sigma_i u_i$, i.e. A transforms the basis $\{v_i\}$ to the basis $\{u_i\}$ with scaling of each v_i given by σ_i. Since $\sigma_{r+1} = \cdots = \sigma_n = 0$, $Av_i = 0$ for $i = r+1, \ldots, n$. Thus $\text{null}(A) = \text{span}(v_{r+1}, \ldots, v_n)$. Furthermore, for $i = 1, \ldots, r$, $Av_i = \sigma_i u_i$ which gives $\text{range}(A) = \text{span}(u_1, \ldots, u_r)$. By the rank-nullity theorem, we have that $\text{rank}(A) = \dim(A) - \text{nullity}(A)$ which tells $\text{rank}(A) = n - (n - r) = r$.

$\square$

Problem 4

Let A be an $m \times n$ matrix where $m < n$ and rows of A are linearly independent. Then the system of linear equations $Ax = b$ is underdetermined, i.e., infinitely many solutions. Among them, we want to find the one that has the minimum 2-norm. Check that the minimum 2-norm solution is given by

$$x^* = A^T(AA^T)^{-1}b.$$

Hint. One way to solve this problem is the following. Check that x^* is a solution to $Ax = b$. Show that is $x^* + y$ is also a solution of $Ax = b$ then $Ay = 0$. Then check that the 2-norm of $x^* + y$ is minimal if $y = 0$.

Solution

Proof. Let A be an $m \times n$ matrix where $m < n$ and rows of A are linearly independent. Then the system of linear equations $Ax = b$ is underdetermined. Among them, we want to find the one that has the minimum 2-norm. Consider

$$x^* = A^T * (AA^T)^{-1}b.$$

First, notice that x^* is a solution of $Ax = b$ as

$$Ax^* = AA^T(AA^T)^{-1}b = b,$$

since A is orthogonal by construction. Next observe that $x^* + y$ is also a solution if $Ay = 0$ as

$$A(x^* + y) = AA^T(AA^T)^{-1}b + Ay = b + Ay.$$

Finally, notice that $x^* + y$ is minimal if $y = 0$ as

$$\begin{aligned}\min_y \|x^* + y\|_2^2 &=< x^* + y, x^* + y > \\ &=< x^*, x > +2 < x^*, y > + < y, y > \\ &= \|x^*\|_2^2 + \|y\|_2^2 + 2 < A^T(AA^T)^{-1}b, y > \\ &= \|x^*\|_2^2 + \|y\|_2^2 + 2b^T(AA^T)^{-1}Ay \\ &= \|x^*\|_2^2 + \|y\|_2^2,\end{aligned}$$

as $Ay = 0$. Therefore the minimum $2 - norm$ solution is given by

$$x^* = A^T(AA^T)^{-1}b.$$

$\square$

Problem 5

Let A be a 3×3 matrix, and let T be its Schur form, i.e., there is a Hermitian matrix Q (i.e. $Q^*Q = QQ^* = I$ where Q^* denotes the transpose and complex conjugate of Q) such that

$$A = QTQ^*, \quad \text{where } T = \begin{pmatrix} \lambda_1 & t_{12} & t_{13} \\ 0 & \lambda_2 & t_{23} \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

Assume that λ_j, $j = 1, 2, 3$ are all distinct.

(a) Show that if v is an eigenvector of T then Qv is the eigenvector of A corresponding to the same eigenvalue.

(b) Find eigenvectors of T. *Hint: Check that* $v_1 = [1, 0, 0]^T$. *Look for* v_2 *of the form* $v_2 = [a, 1, 0]^T$*, and then for* v_3 *of the form* $v_3 = [b, c, 1]^T$*, where* a, b, c *are to be expressed via entries of matrix* T.

(c) Write out eigenvectors of A in terms of the found eigenvectors of T and the columns of Q: $Q = [q_1, q_2, q_3]$.

Solution

Proof. Let A be a 3×3 matrix, and let T be its Schur form, i.e., there is a Hermitian matrix Q (i.e. $Q^*Q = QQ^* = I$ where Q^* denotes the transpose and complex conjugate of Q) such that

$$A = QTQ^*, \quad \text{where } T = \begin{pmatrix} \lambda_1 & t_{12} & t_{13} \\ 0 & \lambda_2 & t_{23} \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

Assume that λ_j, $j = 1, 2, 3$ are all distinct.

(a) Assume that v is an eigenvector of T corresponding to the eigenvalue λ, i.e. $Tv = \lambda v$. Observe

$$AQv = QTQ^*Qv = QTv = \lambda QV.$$

Thus Qv is an eigenvector of A corresponding to the same eigenvalue.

(b) We wish to find the eigenvectors of T. Let's look for v_1 of the form $v_1 = [1, 0, 0]^T$ with the corresponding eigenvalue of λ. Observe that

$$\begin{pmatrix} \lambda_1 & t_{12} & t_{13} \\ 0 & \lambda_2 & t_{23} \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \iff \begin{pmatrix} \lambda_1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda \\ 0 \\ 0 \end{pmatrix}.$$

So $Tv_1 = \lambda_1 v_1$ is indeed an eigenvector of T. Next, let's look for v_2 of the form $v_2 = [a, 1, 0]^T$ with the corresponding eigenvalue of λ. Observe that

$$\begin{pmatrix} \lambda_1 & t_{12} & t_{13} \\ 0 & \lambda_2 & t_{23} \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} a \\ 1 \\ 0 \end{pmatrix} = \lambda \begin{pmatrix} a \\ 1 \\ 0 \end{pmatrix} \iff \begin{pmatrix} \lambda_1 a + t_{12} \\ \lambda_2 \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda a \\ \lambda \\ 0 \end{pmatrix}.$$

Solving the system of equations gives $\lambda = \lambda_2$ and $a = \frac{t_{12}}{\lambda_2 - \lambda_1}$. Thus $Tv_2 = \lambda_2 v_2$ where

$$v_2 = \left[\frac{t_{12}}{\lambda_2 - \lambda_1}, 1, 0\right]^T.$$

Next, let's look for v_3 of the form $v_3 = [b, c, 1]^T$ with the corresponding eigenvalue of λ. Observe that

$$\begin{pmatrix} \lambda_1 & t_{12} & t_{13} \\ 0 & \lambda_2 & t_{23} \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} b \\ c \\ 1 \end{pmatrix} = \lambda \begin{pmatrix} b \\ c \\ 1 \end{pmatrix} \iff \begin{pmatrix} \lambda_1 b + t_{12} c + t_{13} \\ \lambda_2 c + t_{23} \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} \lambda b \\ \lambda c \\ \lambda \end{pmatrix}.$$

Solving the system of equations gives $\lambda = \lambda_3, c = \frac{t_{23}}{\lambda_3 - \lambda_2}$, and $b = \frac{t_{12}c + t_{13}}{\lambda_3 - \lambda_1} = \frac{\frac{t_{12}t_{23}}{\lambda_3 - \lambda_2} + t_{13}}{\lambda_3 - \lambda_1}$. Thus $Tv_3 = \lambda_3 v_3$ where

$$v_3 = \left[\frac{\frac{t_{12}t_{23}}{\lambda_3 - \lambda_2} + t_{13}}{\lambda_3 - \lambda_1}, \frac{t_{23}}{\lambda_3 - \lambda_2}, 1\right]^T.$$

(c) Next, we wish to find the eigenvectors of A in terms of the eigenvectors of T and the columns of $Q = [q_1, q_2, q_3]$. From (a) we know that if v is an eigenvector of T then Qv is the eigenvector of A corresponding to the same eigenvalue. Thus let's compute

$$\begin{aligned} Qv_1 &= q_1, \\ Qv_2 &= aq_1 + q_2 = \frac{t_{12}}{\lambda_2 - \lambda_1} q_1 + q_2, \\ Qv_3 &= bq_1 + cq_2 + q_3 = \frac{\frac{t_{12}t_{23}}{\lambda_3 - \lambda_2} + t_{13}}{\lambda_3 - \lambda_1} q_1 + \frac{t_{23}}{\lambda_3 - \lambda_2} q_2 + q_3. \end{aligned}$$

where Qv_i correspond to λ_i.

$\square$