Linear Algebra Problems

1 January 2002, Day 1, Question 8

Question

Let

$$\boldsymbol{A} = \begin{pmatrix} 7 & 2 & 4 & 6 & 3 \\ 2 & 5 & 1 & 9 & 8 \\ 4 & 1 & 6 & 4 & 7 \\ 6 & 9 & 4 & 2 & 9 \\ 3 & 8 & 7 & 9 & 6 \end{pmatrix},$$

and let $\boldsymbol{v}_{(0)} = (1,1,1,1,1)^T$ be a column vector. For the iteration scheme

$$\boldsymbol{v}_{(m+1)} = \boldsymbol{A}\boldsymbol{v}_{(m)}, \qquad m = 0,1,2,\ldots,$$

with initial guess $\boldsymbol{v}_{(0)}$, after m iterations $\boldsymbol{v}_{(m)} = \boldsymbol{A}^m\boldsymbol{v}_{(0)}$. Show that as $m \to \infty$,

$$\frac{\boldsymbol{v}_{(m)}^T\boldsymbol{v}_{(m+1)}}{\boldsymbol{v}_{(m)}^T\boldsymbol{v}_{(m)}} = \lambda_1 + \mathcal{O}\left(\left(\frac{|\lambda_2|}{\lambda_1}\right)^{2m}\right),$$

where $\lambda_1 > |\lambda_i|$, $i = 2,3,4,5$, are the eigenvalues of $\boldsymbol{A}$.

Answer

Note that $\boldsymbol{A}$ is real symmetric, so it is *orthogonally diagonalizable*, *i.e.* there is an orthogonal matrix $\boldsymbol{R}$ such that

$$\boldsymbol{A} = \boldsymbol{R}\begin{pmatrix} \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 & 0 \\ 0 & 0 & 0 & \lambda_4 & 0 \\ 0 & 0 & 0 & 0 & \lambda_5 \end{pmatrix}\boldsymbol{R}^{-1}.$$

Since $\boldsymbol{R}$ is orthogonal, $\boldsymbol{R}^{-1} = \boldsymbol{R}^T$. For simplicity of notation, let $\boldsymbol{\Lambda}$ be the diagonal matrix in the factorization above. Note that $\boldsymbol{A}^m = \boldsymbol{R}\boldsymbol{\Lambda}^m\boldsymbol{R}^T$.

$$\begin{aligned}
\boldsymbol{v}_{(m)} &= \boldsymbol{A}^m\boldsymbol{v}_{(0)} = \boldsymbol{R}\boldsymbol{\Lambda}^m\boldsymbol{R}^T\boldsymbol{v}_{(0)} \\
\boldsymbol{v}_{(m+1)} &= \boldsymbol{A}^{m+1}\boldsymbol{v}_{(0)} = \boldsymbol{R}\boldsymbol{\Lambda}^{m+1}\boldsymbol{R}^T\boldsymbol{v}_{(0)} \\
\boldsymbol{v}_{(m)}^T\boldsymbol{v}_{(m+1)} &= \left(\boldsymbol{R}\boldsymbol{\Lambda}^m\boldsymbol{R}^T\boldsymbol{v}_{(0)}\right)^T \boldsymbol{R}\boldsymbol{\Lambda}^{m+1}\boldsymbol{R}^T\boldsymbol{v}_{(0)} \\
&= \boldsymbol{v}_{(0)}^T\boldsymbol{R}\left(\boldsymbol{\Lambda}^m\right)^T\boldsymbol{R}^T\boldsymbol{R}\boldsymbol{\Lambda}^{m+1}\boldsymbol{R}^T\boldsymbol{v}_{(0)} \\
&= \boldsymbol{v}_{(0)}^T\boldsymbol{R}\boldsymbol{\Lambda}^{2m+1}\boldsymbol{R}^T\boldsymbol{v}_{(0)} \qquad (\boldsymbol{R}^T\boldsymbol{R} = \boldsymbol{I};\ \boldsymbol{\Lambda} \text{ is symmetric}) \\
\boldsymbol{v}_{(m)}^T\boldsymbol{v}_{(m)} &= \boldsymbol{v}_{(0)}^T\boldsymbol{R}\boldsymbol{\Lambda}^{2m}\boldsymbol{R}^T\boldsymbol{v}_{(0)}
\end{aligned}$$

Let $\boldsymbol{u} = \boldsymbol{R}^T \boldsymbol{v}_{(0)}$. Then the ratio we seek can be written as

$$\begin{aligned}
\frac{\boldsymbol{v}_{(m)}^T \boldsymbol{v}_{(m+1)}}{\boldsymbol{v}_{(m)}^T \boldsymbol{v}_{(m)}} &= \frac{\lambda_1^{2m+1} u_1^2 + \lambda_2^{2m+1} u_2^2 + \cdots}{\lambda_1^{2m} u_1^2 + \lambda_2^{2m} u_2^2 + \cdots} \\
&= \frac{\lambda_1^{2m+1} \left[u_1^2 + \left(\frac{\lambda_2}{\lambda_1}\right)^{2m+1} u_2^2 + \cdots \right]}{\lambda_1^{2m} \left[u_1^2 + \left(\frac{\lambda_2}{\lambda_1}\right)^{2m} u_2^2 + \cdots \right]} \\
&= \lambda_1 \left(\frac{u_1^2 + \left(\frac{\lambda_2}{\lambda_1}\right)^{2m+1} u_2^2 + \cdots}{u_1^2 + \left(\frac{\lambda_1}{\lambda_2}\right)^{2m} u_2^2 + \cdots} \right).
\end{aligned}$$

If $u_1 \neq 0$, then we can re-write this as

$$\frac{\boldsymbol{v}_{(m)}^T \boldsymbol{v}_{(m+1)}}{\boldsymbol{v}_{(m)}^T \boldsymbol{v}_{(m)}} = \lambda_1 \left(\frac{1 + \left(\frac{\lambda_1}{\lambda_2}\right)^{2m+1} \left(\frac{u_2}{u_1}\right)^2 + \cdots}{1 + \left(\frac{\lambda_1}{\lambda_2}\right)^{2m} \left(\frac{u_2}{u_1}\right)^2 + \cdots} \right).$$

Since $|\lambda_2/\lambda_1| < 1$, for large-enough m,

$$\epsilon = \left(\frac{\lambda_2}{\lambda_1}\right)^{2m} \left(\frac{u_2}{u_1}\right)^2 + \cdots < 1.$$

In this case, we can use the series

$$\frac{1}{1+\epsilon} = \sum_{k=0}^{\infty} (-1)^k \epsilon^k.$$

which converges for $|\epsilon| < 1$. We get

$$\frac{\boldsymbol{v}_{(m)}^T \boldsymbol{v}_{(m+1)}}{\boldsymbol{v}_{(m)}^T \boldsymbol{v}_{(m)}} = \lambda_1 \left(1 + \left(\frac{\lambda_1}{\lambda_2}\right)^{2m+1} \left(\frac{u_2}{u_1}\right)^2 + \cdots \right) \left(1 - \epsilon + \epsilon^2 - \cdots\right).$$

The largest of the small terms on the right-hand side is ϵ, which has terms with exponent $2m$. Hence

$$\frac{\boldsymbol{v}_{(m)}^T \boldsymbol{v}_{(m+1)}}{\boldsymbol{v}_{(m)}^T \boldsymbol{v}_{(m)}} = \lambda_1 \left(1 - \epsilon + \cdots\right) = \lambda_1 \left[1 + \mathcal{O}\left(\left(\frac{|\lambda_2|}{\lambda_1} \right)^{2m} \right) \right].$$

We haven't referred even once to $\boldsymbol{v}_{(0)}$. Does it play a rôle here at all? It does, but it's a minor rôle. The forms of $\boldsymbol{v}_{(0)}$ and $\boldsymbol{A}$ ensure that each entry in $\boldsymbol{v}_{(1)}$ is a positive number. Further, each entry in $\boldsymbol{v}_{(m)}$ will be positive for each m. Thus we are guaranteed to have nonzero u_1 in the manipulations above, so the argument works.

2 August 2002, Day 1, Question 4

Question

Let $\boldsymbol{u}$ be a unit vector in $\mathbb{R}^n$ with $\boldsymbol{u}^T\boldsymbol{u} = 1$. Let

$$\boldsymbol{A} = \boldsymbol{I} - 2\boldsymbol{u}\boldsymbol{u}^T$$

be an $n \times n$ matrix.

(a) Show that $\boldsymbol{A}$ is an orthogonal matrix, that is, $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}$.
(b) Find the eigenvalues of $\boldsymbol{A}$. Find the determinant of $\boldsymbol{A}$.

Let $\boldsymbol{v}$ be a unit vector in $\mathbb{R}^n$ with $\boldsymbol{v}^T\boldsymbol{v} = 1$. Let

$$\boldsymbol{B} = \boldsymbol{I} - 2\boldsymbol{v}\boldsymbol{v}^T$$

be an $n \times n$ matrix. Let

$$\boldsymbol{C} = \boldsymbol{A}\boldsymbol{B}.$$

(c) Show that $\boldsymbol{C}$ is an orthogonal matrix, that is, $\boldsymbol{C}^T\boldsymbol{C} = \boldsymbol{I}$.
(d) Find the determinant of $\boldsymbol{C}$.

Answer

(a) First note that $\boldsymbol{u}\boldsymbol{u}^T$ is a symmetric matrix: $\left(\boldsymbol{u}\boldsymbol{u}^T\right)^T = \boldsymbol{u}\boldsymbol{u}^T$. Since the identity matrix $\boldsymbol{I}$ is also symmetric, $\boldsymbol{A}$ is symmetric.

$$\begin{aligned}
\boldsymbol{A}^T\boldsymbol{A} &= \boldsymbol{A}\boldsymbol{A} \\
&= \left(\boldsymbol{I} - 2\boldsymbol{u}\boldsymbol{u}^T\right)\left(\boldsymbol{I} - 2\boldsymbol{u}\boldsymbol{u}^T\right) \\
&= \boldsymbol{I} - 2\boldsymbol{u}\boldsymbol{u}^T - 2\boldsymbol{u}\boldsymbol{u}^T + 4\boldsymbol{u}\boldsymbol{u}^T\boldsymbol{u}\boldsymbol{u}^T \\
&= \boldsymbol{I} - 4\boldsymbol{u}\boldsymbol{u}^T + 4\boldsymbol{u}\left(\boldsymbol{u}^T\boldsymbol{u}\right)\boldsymbol{u}^T \\
&= \boldsymbol{I} - 4\boldsymbol{u}\boldsymbol{u}^T + 4\boldsymbol{u}\boldsymbol{u}^T \qquad \left(\boldsymbol{u}^T\boldsymbol{u} = 1\right) \\
&= \boldsymbol{I}
\end{aligned}$$

(b) Note immediately that $\boldsymbol{u}\boldsymbol{u}^T$ maps $\boldsymbol{u}$ to itself: $\boldsymbol{u}\boldsymbol{u}^T\boldsymbol{u} = \boldsymbol{u}\left(\boldsymbol{u}^T\boldsymbol{u}\right) = \boldsymbol{u}$. $\boldsymbol{I}$ also maps $\boldsymbol{u}$ to itself, so $\boldsymbol{A}$ maps $\boldsymbol{u}$ to some scalar multiple of itself. Hence $\boldsymbol{u}$ is an eigenvector of $\boldsymbol{A}$. In fact,

$$\left(\boldsymbol{I} - 2\boldsymbol{u}\boldsymbol{u}^T\right)\boldsymbol{u} = \boldsymbol{u} - 2\boldsymbol{u} = -\boldsymbol{u},$$

so -1 is the eigenvalue associated with the eigenvector $\boldsymbol{u}$.

We needn't find the other eigenvectors to determine the other eigenvalues. Note that the span of $\{\boldsymbol{u}\}$ is a one-dimensional subspace of $\mathbb{R}^n$, so there is an $(n-1)$-dimensional subspace of vectors orthogonal to $\boldsymbol{u}$. That is, there are linearly independent vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}$ such that $\boldsymbol{u}^T \boldsymbol{x}_k = 0$ for $k = 1, \ldots, n-1$. For each of these $\boldsymbol{x}_K$, then,

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x}_k &= \left(\boldsymbol{I} - 2\boldsymbol{u}\boldsymbol{u}^T\right) \boldsymbol{x}_k \\ &= \boldsymbol{x}_k - 2\boldsymbol{u}\left(\boldsymbol{u}^T \boldsymbol{x}_k\right) \\ &= \boldsymbol{x}_k \qquad\qquad (\boldsymbol{u}^T \boldsymbol{x}_k = 0), \end{aligned}$$

so each $\boldsymbol{x}_k$ is an eigenvector with eigenvalue 1.

$\boldsymbol{A}$ has one eigenvalue equal to -1 and $n-1$ eigenvalues equal to 1. Since the determinant is equal to the product of all n eigenvalues, $\det \boldsymbol{A} = (-1)(1)^{n-1} = 1$.

(c) First note that since that $\boldsymbol{B}$ has the same type of structure that $\boldsymbol{A}$ does. Hence $\boldsymbol{B}$ is symmetric, orthogonal, and has determinant -1. Then note that $\boldsymbol{C}^T = (\boldsymbol{A}\boldsymbol{B})^T = \boldsymbol{B}^T \boldsymbol{A}^T = \boldsymbol{B}\boldsymbol{A}$.

$$\begin{aligned} \boldsymbol{C}^T \boldsymbol{C} &= \boldsymbol{B}\boldsymbol{A}\boldsymbol{A}\boldsymbol{B} \\ &= \boldsymbol{B}\boldsymbol{I}\boldsymbol{B} \qquad\qquad (\boldsymbol{A}\boldsymbol{A} = \boldsymbol{I}) \\ &= \boldsymbol{B}\boldsymbol{B} \\ &= \boldsymbol{I} \end{aligned}$$

(d) The determinant of a product of matrices is equal to the product of the matrices' determinants, so

$$\det \boldsymbol{C} = (\det \boldsymbol{A})(\det \boldsymbol{B}) = (-1)^2 = 1.$$

Comment

It is worth noting here that there are two types of orthogonal matrices: those with determinant 1 and those with determinant -1. Suppose that $\boldsymbol{A}$ is an orthogonal (not necessarily symmetric) $n \times n$ real matrix. Then $\boldsymbol{A}^T \boldsymbol{A} = \boldsymbol{I}$, and

$$1 = \det \boldsymbol{I} = \det\left(\boldsymbol{A}^T \boldsymbol{A}\right) = \left(\det \boldsymbol{A}^T\right)(\det \boldsymbol{A}) = (\det \boldsymbol{A})^2 .$$

Here we've used the fact that $\det \boldsymbol{A}^T = \det \boldsymbol{A}$. The conclusion is that $\det \boldsymbol{A}$ is 1 or -1. Those $n \times n$ orthogonal matrices with determinant 1 form the *special orthogonal group (of order n)*, often written $SO(n)$. Special orthogonal matrices correspond to rotations of $\mathbb{R}^n$, while those with determinant -1 correspond to rotations combined with reflections.

Since $\boldsymbol{A} = \boldsymbol{I} - 2\boldsymbol{u}\boldsymbol{u}^T$ has determinant -1, it corresponds to reflection in $\mathbb{R}^n$. Since $\boldsymbol{A}\boldsymbol{u} = -\boldsymbol{u}$ and since all of $\boldsymbol{A}$'s other eigenvalues are equal to 1, the mapping $\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}$ performs a reflection in the hyperplane orthogonal to $\boldsymbol{u}$.

A matrix of the form $\boldsymbol{A} = \boldsymbol{I} - 2\boldsymbol{u}\boldsymbol{u}^T$, where $\boldsymbol{u}$ is a unit vector, is called a *Householder matrix*, and the transformation it performs is called a *Householder reflection*. It is an essential tool for the computation of the *QR decomposition* of a real square matrix. Each real square matrix can be expressed as a product $\boldsymbol{Q}\boldsymbol{R}$, where $\boldsymbol{Q}$ is orthogonal and $\boldsymbol{R}$ is upper triangular.

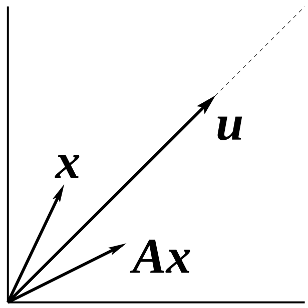

Figure 1: An example of a reflection in $\mathbb{R}^2$ with $\boldsymbol{A} = \boldsymbol{I} - 2\boldsymbol{u}\boldsymbol{u}^T$

3 Math 583-A, Fall 2002, Final Exam, Question 1

Question

Let $\mathbb{V}$ be the space of 2×2 real matrices and $F : \mathbb{V} \to \mathbb{V}$ be the linear map defined by

$$F(\boldsymbol{A}) = \boldsymbol{AM} - \boldsymbol{MA},$$

where $\boldsymbol{A} \in \mathbb{V}$ and

$$\boldsymbol{M} = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}.$$

Find a basis and the dimension of the kernel (null-space) of F.

Answer

Let

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Then the matrix equation $\boldsymbol{AM} - \boldsymbol{MA} = \boldsymbol{O}$, where the left-hand side is

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} -2a_{21} & 2a_{11} + 2a_{12} - 2a_{22} \\ -2a_{21} & 2a_{21} \end{pmatrix},$$

is equivalent to a system of four scalar linear equations:

$$\begin{aligned} -2a_{21} &= 0, \\ 2a_{11} + 2a_{12} - 2a_{22} &= 0, \\ -2a_{21} &= 0, \\ 2a_{21} &= 0. \end{aligned}$$

Note that we do not have four independent equations here, so the solution is not unique. In fact the solutions are the form

$$\begin{pmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{12} \\ 0 \\ -a_{11} - a_{12} \end{pmatrix}.$$

where a_{11} and a_{12} can be any real numbers. The corresponding 2×2 matrix is

$$\begin{pmatrix} a_{11} & a_{12} \\ 0 & -a_{11} - a_{12} \end{pmatrix} = a_{11} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + a_{12} \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}.$$

Each 2×2 real matrix $\boldsymbol{A}$ of this form satisfies $F(\boldsymbol{A}) = \boldsymbol{O}$, so the kernel of F is 2-dimensional, and one example of a basis for this kernel is

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} \right\}.$$

4 Math 583-A, Fall 2002 Final Exam, Question 2

Question

(a) For a general 2×2 matrix $\boldsymbol{A}$, write down expressions for the eigenvalues of $\boldsymbol{A}$ in terms of the determinant and trace of $\boldsymbol{A}$, *i.e.* $\det \boldsymbol{A}$ and $\mathrm{Tr}\boldsymbol{A}$.

(b) Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be 2×2 matrices such that

$$\boldsymbol{AB} - 2\boldsymbol{BA} = \boldsymbol{O}.$$

Show that all the eigenvalues of the matrix $\boldsymbol{C} = \boldsymbol{AB}$ are equal to zero.

Answer

(a) First we compute the determinant of $\boldsymbol{A} - \lambda \boldsymbol{I}$:

$$\begin{aligned} \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} &= (a_{11} - \lambda)(a_{22} - \lambda) - a_{21}a_{12} \\ &= \lambda^2 - (\underbrace{a_{11} + a_{22}}_{\mathrm{Tr}\boldsymbol{A}})\lambda + \underbrace{a_{11}a_{22} - a_{21}a_{12}}_{\det \boldsymbol{A}}. \end{aligned}$$

The solutions of the quadratic equation are

$$\lambda_{\pm} = \frac{1}{2}\left(\mathrm{Tr}\boldsymbol{A} \pm \sqrt{(\mathrm{Tr}\boldsymbol{A})^2 - 4\det \boldsymbol{A}} \right).$$

(b) Multiplying an $n \times n$ matrix by a scalar α multiplies each column by that same scalar, so $\det(\alpha \boldsymbol{A}) = \alpha^n \det(\boldsymbol{A})$. In the case at hand, we're told that $\boldsymbol{AB} = 2\boldsymbol{BA}$, so

$$\det \boldsymbol{C} = \det(\boldsymbol{AB}) = \det(2\boldsymbol{BA}) = 2^2 \det(\boldsymbol{BA}).$$

But $\det(\boldsymbol{BA}) = \det(\boldsymbol{AB})$, so

$$\det \boldsymbol{C} = 4\det(\boldsymbol{BA}) = 4\det(\boldsymbol{AB}) = 4\det \boldsymbol{C}.$$

The conclusion is that $\det \boldsymbol{C} = 0$.

Comment

Each $n \times n$ matrix $\boldsymbol{A}$ has n eigenvalues, when the eigenvalues are counted according to (algebraic) multiplicity.

$n \times n$ matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are called *similar* (to each other) if there is some invertible matrix $\boldsymbol{C}$ such that $\boldsymbol{B} = \boldsymbol{CAC}^{-1}$. Recall that the eigenvalues of $\boldsymbol{B}$ are the zeroes of its characteristic polynomial $(\det(\boldsymbol{B} - \lambda \boldsymbol{I}))$; how are the characteristic polynomials of similar matrices related?

$$\begin{aligned}
\det(\boldsymbol{B} - \lambda\boldsymbol{I}) &= \det\left(\boldsymbol{CAC}^{-1} - \lambda\boldsymbol{I}\right) \\
&= \det\left(\boldsymbol{CAC}^{-1} - \lambda\boldsymbol{CIC}^{-1}\right) \\
&= \det\left(\boldsymbol{C}(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{C}^{-1}\right) \\
&= (\det\boldsymbol{C})(\det(\boldsymbol{A} - \lambda\boldsymbol{I}))\left(\det\boldsymbol{C}^{-1}\right) \\
&= (\det\boldsymbol{C})(\det\boldsymbol{C})^{-1}(\det(\boldsymbol{A} - \lambda\boldsymbol{I})) \\
&= \det(\boldsymbol{A} - \lambda\boldsymbol{I})
\end{aligned}$$

Similar matrices have identical characteristic polynomials; thus they also have identical eigenvalues and identical multiplicities of eigenvalues.

Each $n \times n$ matrix $\boldsymbol{A}$ (real or complex) has a *Schur factorization*:

$$\boldsymbol{A} = \boldsymbol{QTQ}^*,$$

where $\boldsymbol{Q}$ is unitary ($\boldsymbol{Q}^*\boldsymbol{Q} = \boldsymbol{I}$, *i.e.* $\boldsymbol{Q}^* = \boldsymbol{Q}^{-1}$) and $\boldsymbol{T}$ is upper triangular. You should be able to convince yourself that the characteristic polynomial of an upper triangular matrix has the form

$$\det(\boldsymbol{T} - \lambda\boldsymbol{I}) = \prod_{k=1}^{n}(t_{kk} - \lambda),$$

where t_{kk} is the kth element on the diagonal of $\boldsymbol{T}$. This indicates that the eigenvalues of an upper triangular matrix are found on the diagonal of such a matrix.

Let $\boldsymbol{A}$ have Schur factorization $\boldsymbol{QTQ}^*$, so that

$$\begin{aligned}
\det\boldsymbol{A} &= \det(\boldsymbol{QTQ}^*) \\
&= (\det\boldsymbol{Q})(\det\boldsymbol{T})(\det\boldsymbol{Q}^*) \\
&= (\det\boldsymbol{Q})(\det\boldsymbol{T})\left(\det\boldsymbol{Q}^{-1}\right) \\
&= (\det\boldsymbol{Q})(\det\boldsymbol{T})(\det\boldsymbol{Q})^{-1} \\
&= (\det\boldsymbol{Q})(\det\boldsymbol{Q})^{-1}(\det\boldsymbol{T}) \\
&= \det\boldsymbol{T} \;=\; \prod_{k=1}^{n} t_{kk} \;=\; \prod_{k=1}^{n}\lambda_k,
\end{aligned}$$

where $\lambda_1, \ldots, \lambda_n$ are the eiganvalues of $\boldsymbol{T}$. Since $\boldsymbol{A}$ and $\boldsymbol{T}$ are similar, these are also the eigenvalues of $\boldsymbol{A}$. Hence for general $n \times n$ matrix $\boldsymbol{A}$,

$$\det \boldsymbol{A} = \prod_{k=1}^{n} \lambda_k.$$

Now we prove a related fact concerning the trace. Let $\boldsymbol{A}$ again have Schur factorization $\boldsymbol{QTQ}^*$.

$$\begin{aligned}
\mathrm{Tr}\boldsymbol{A} &= \sum_{k=1}^{n} a_{kk} \\
&= \sum_{k=1}^{n} \left(\boldsymbol{ATQ}^*\right)_{kk} \\
&= \sum_{k=1}^{n}\sum_{j=1}^{n} \left(\boldsymbol{QT}\right)_{kj} \left(\boldsymbol{Q}^*\right)_{jk} \\
&= \sum_{k=1}^{n}\sum_{j=1}^{n} \left(\sum_{i=1}^{n} q_{ki}t_{ij}\right) \left(\boldsymbol{Q}^*\right)_{jk} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} t_{ij} \left(\sum_{k=1}^{n} \left(\boldsymbol{Q}^*\right)_{jk} q_{ki}\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} t_{ij} \left(\boldsymbol{Q}^*\boldsymbol{Q}\right)_{ji} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} t_{ij}\delta_{ji} \qquad (\boldsymbol{Q} \text{ is unitary}) \\
&= \sum_{j=1}^{n} t_{jj} \;=\; \sum_{j=1}^{n} \lambda_j, \qquad (\boldsymbol{T} \text{ is upper triangular})
\end{aligned}$$

where $\lambda_1, \ldots, \lambda_2$ are the eigenvalues of $\boldsymbol{T}$. Since $\boldsymbol{A}$ is similar to $\boldsymbol{T}$, these are the eigenvalues of $\boldsymbol{A}$ as well. We see, then, that the trace of $\boldsymbol{A}$ is equal to the sum of $\boldsymbol{A}$'s eigenvalues, counted according to multiplicity.

5 January 2003, Day 1, Question 4

Question

Let $\boldsymbol{A}$ be a complex rectangular $m \times n$ matrix, and let $\boldsymbol{A}^*$ be its conjugate transpose. Show that $\boldsymbol{A}^*\boldsymbol{A}$ and $\boldsymbol{AA}^*$ have the same nonzero eigenvalues.

Answer

Each matrix, regardless of whether it is square, has unique *singular value decomposition*:

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^*,$$

where $\boldsymbol{U}$ is an $m \times m$ complex unitary matrix, $\boldsymbol{V}$ is an $n \times n$ complex unitary matrix, and $\boldsymbol{\Sigma}$ is an $m \times n$ real "diagonal" matrix. $\boldsymbol{\Sigma}$, which is non-square if $m \neq n$, is diagonal in the sense that largest upper-left square sub-matrix of $\boldsymbol{\Sigma}$ is diagonal. The non-zero entries in $\boldsymbol{\Sigma}$ are called the *singular values* of $\boldsymbol{A}$. The rank of $\boldsymbol{A}$ is equal to the number of singular values $\boldsymbol{A}$ has, *i.e.* the number of non-zero entries in $\boldsymbol{\Sigma}$. This number is bounded by $\min\{m, n\}$.

The following is an equivalent expression of the singular value decomposition and will be easier to use in this problem:

$$\boldsymbol{A} = \sum_{j=1}^{r} \sigma_j \boldsymbol{u}_j \boldsymbol{v}_j^*,$$

where r is the rank of $\boldsymbol{A}$, σ_j is the jth element on the diagonal of $\boldsymbol{\Sigma}$, $\boldsymbol{u}_j$ is the jth column of $\boldsymbol{U}$, and $\boldsymbol{v}_j$ is the jth column of $\boldsymbol{V}$.

$$\begin{aligned}
\boldsymbol{A}^*\boldsymbol{A} &= \left(\sum_{j=1}^{r} \sigma_j \boldsymbol{u}_j \boldsymbol{v}_j^*\right)^* \left(\sum_{k=1}^{r} \sigma_k \boldsymbol{u}_k \boldsymbol{v}_k^*\right) \\
&= \left(\sum_{j=1}^{r} \sigma_j \boldsymbol{v}_j \boldsymbol{u}_j^*\right) \left(\sum_{k=1}^{r} \sigma_k \boldsymbol{u}_k \boldsymbol{v}_k^*\right) \\
&= \sum_{j=1}^{r} \sum_{k=1}^{r} \sigma_j \sigma_k \boldsymbol{v}_j \left(\boldsymbol{u}_j^* \boldsymbol{u}_k\right) \boldsymbol{v}_k^* \\
&= \sum_{j=1}^{r} \sum_{k=1}^{r} \sigma_j \sigma_k \boldsymbol{v}_j \delta_{jk} \boldsymbol{v}_k^* \qquad \text{(columns of a unitary matrix form an orthonormal set)} \\
&= \sum_{j=1}^{r} \sigma_j^2 \boldsymbol{v}_j \boldsymbol{v}_j^*, \\
\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{v}_\ell &= \sum_{j=1}^{r} \sigma_j^2 \boldsymbol{v}_j \boldsymbol{v}_j^* \boldsymbol{v}_\ell \;=\; \sum_{j=1}^{r} \sigma_j^2 \boldsymbol{v}_j \delta_{k\ell} \;=\; \sigma_\ell^2 \boldsymbol{v}_\ell
\end{aligned}$$

The eigenvalues of $\boldsymbol{A}^*\boldsymbol{A}$, which is an $n \times n$ matrix, are σ_ℓ^2, $\ell = 1, \ldots, n$. On the other hand,

$$
\begin{aligned}
\boldsymbol{A}\boldsymbol{A}^* &= \left(\sum_{j=1}^{r} \sigma_j \boldsymbol{u}_j \boldsymbol{v}_j^*\right)\left(\sum_{k=1}^{r} \sigma_k \boldsymbol{u}_k \boldsymbol{v}_k^*\right)^* \\
&= \left(\sum_{j=1}^{r} \sigma_j \boldsymbol{u}_j \boldsymbol{v}_j^*\right)\left(\sum_{k=1}^{r} \sigma_k \boldsymbol{v}_k \boldsymbol{u}_k^*\right) \\
&= \sum_{j=1}^{r}\sum_{k=1}^{r} \sigma_j \sigma_k \boldsymbol{u}_j \left(\boldsymbol{v}_j^* \boldsymbol{v}_k\right) \boldsymbol{u}_k^* \\
&= \sum_{j=1}^{r}\sum_{k=1}^{r} \sigma_j \sigma_k \boldsymbol{u}_j \delta_{jk} \boldsymbol{u}_k^* \\
&= \sum_{j=1}^{r} \sigma_j^2 \boldsymbol{u}_j \boldsymbol{u}_j^*, \\
\boldsymbol{A}\boldsymbol{A}^*\boldsymbol{u}_\ell &= \sum_{j=1}^{r} \sigma_j^2 \boldsymbol{u}_j \boldsymbol{u}_j^* \boldsymbol{u}_\ell \;=\; \sum_{j=1}^{r} \sigma_j^2 \boldsymbol{u}_j \delta_{j\ell} \;=\; \sigma_\ell^2 \boldsymbol{u}_\ell.
\end{aligned}
$$

The eigenvalues of $\boldsymbol{A}\boldsymbol{A}^*$, which is an $m \times m$ matrix, are σ_ℓ^2, $\ell = 1, \ldots, m$. $\boldsymbol{A}^*\boldsymbol{A}$ and $\boldsymbol{A}\boldsymbol{A}^*$ have the same non-zero eigenvalues $\sigma_1^2, \ldots, \sigma_r^2$, where r is the rank of $\boldsymbol{A}$; if $m \neq n$, then $\boldsymbol{A}^*\boldsymbol{A}$ and $\boldsymbol{A}\boldsymbol{A}^*$ have different numbers of zero eigenvalues.