UPSC Civil Services Main 1981 - Mathematics
Linear Algebra

Brij Bhooshan

Asst. Professor

B.S.A. College of Engg & Technology

Mathura

Question 1(a) *State and prove the Cayley Hamilton theorem and verify it for the matrix* $\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$. *Use the result to determine* $\mathbf{A}^{-1}$.

Solution. See 1987 question 5(a) for the Cayley Hamilton theorem.

The characteristic equation of $\mathbf{A}$ is $\begin{vmatrix} x-2 & -3 \\ -3 & x-5 \end{vmatrix} = 0$, or $(x-2)(x-5) - 9 = 0 \Rightarrow x^2 - 7x + 1 = 0$. The Cayley Hamilton theorem implies that $\mathbf{A}^2 - 7\mathbf{A} + \mathbf{I} = \mathbf{0}$.

$\mathbf{A}^2 = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} = \begin{pmatrix} 13 & 21 \\ 21 & 34 \end{pmatrix}$.

Now $\mathbf{A}^2 - 7\mathbf{A} + \mathbf{I} = \begin{pmatrix} 13 & 21 \\ 21 & 34 \end{pmatrix} - 7\begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

So the theorem is verified. $\mathbf{A}^2 - 7\mathbf{A} + \mathbf{I} = \mathbf{0} \Rightarrow (\mathbf{A} - 7\mathbf{I})\mathbf{A} = -\mathbf{I} \Rightarrow \mathbf{A}^{-1} = 7\mathbf{I} - \mathbf{A}$. Thus $\mathbf{A}^{-1} = 7\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} = \begin{pmatrix} 5 & -3 \\ -3 & 2 \end{pmatrix}$. ∎

Question 1(b) *Let Q be the quadratic form*

$$Q = 5x_1^2 + 5x_2^2 + 2x_3^2 + 8x_1x_2 + 4x_1x_3 + 4x_2x_3$$

By using an orthogonal change of variables reduce Q to a form without the cross terms i.e. with terms of the form $a_{ij}x_ix_j, i \neq j$.

Solution. The matrix of the qiven quadratic form Q is $\mathbf{A} = \begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix}$.

The characteristic polynomial of $\mathbf{A}$ is

$$\begin{aligned}
&\begin{vmatrix} 5-\lambda & 4 & 2 \\ 4 & 5-\lambda & 2 \\ 2 & 2 & 2-\lambda \end{vmatrix} = 0 \\
\Rightarrow\quad & (5-\lambda)(5-\lambda)(2-\lambda) - 4(5-\lambda) - 4(8-4\lambda) + 16 + 16 - 4(5-\lambda) = 0 \\
\Rightarrow\quad & (\lambda^2 - 10\lambda + 25)(2-\lambda) - 20 + 4\lambda - 32 + 16\lambda + 12 + 4\lambda = 0 \\
\Rightarrow\quad & -\lambda^3 + 12\lambda^2 + \lambda(-25 + 4 + 16 + 4 - 20) + 50 - 20 - 32 + 12 = 0 \\
\Rightarrow\quad & \lambda^3 - 12\lambda^2 + 21\lambda - 10 = 0
\end{aligned}$$

Thus the eigenvalues are $\lambda = 1, 1, 10$. Let (x_1, x_2, x_3) be an eigenvector for $\lambda = 10$, then

$$\begin{pmatrix} -5 & 4 & 2 \\ 4 & -5 & 2 \\ 2 & 2 & -8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{0} \Rightarrow \begin{array}{rcll} -5x_1 + 4x_2 + 2x_3 & = & 0 & (i) \\ 4x_1 - 5x_2 + 2x_3 & = & 0 & (ii) \\ 2x_1 + 2x_2 - 8x_3 & = & 0 & (iii) \end{array}$$

Subtracting (ii) from (i), we get $-9x_1 + 9x_2 = 0 \Rightarrow x_1 = x_2 \Rightarrow x_1 = 2x_3$. Thus taking $x_3 = 1$, we get $(2, 2, 1)$ as an eigenvector for $\lambda = 10$.

Let (x_1, x_2, x_3) be an eigenvector for $\lambda = 1$, then

$$\begin{pmatrix} 4 & 4 & 2 \\ 4 & 4 & 2 \\ 2 & 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{0} \Rightarrow \begin{array}{rcl} 4x_1 + 4x_2 + 2x_3 & = & 0 \\ 2x_1 + 2x_2 + x_3 & = & 0 \end{array}$$

Take $x_3 = 0, x_1 = 1 \Rightarrow x_2 = -1$ to get $(1, -1, 0)$ as an eigenvector for $\lambda = 1$. Take $x_1 = x_2 = 1 \Rightarrow x_3 = -4$ to get $(1, 1, -4)$ as another eigenvector for $\lambda = 1$, orthogonal to the first.

Thus

$$\mathbf{O} = \begin{pmatrix} \frac{2}{\sqrt{9}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{18}} \\ \frac{2}{\sqrt{9}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{18}} \\ \frac{1}{\sqrt{9}} & 0 & -\frac{4}{\sqrt{18}} \end{pmatrix}$$

is an orthogonal matrix such that $\mathbf{O}'\mathbf{A}\mathbf{O} = \mathbf{O}^{-1}\mathbf{A}\mathbf{O} = \begin{pmatrix} 10 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. If (X_1, X_2, X_3) are new variables, then $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{O} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$ takes $Q(x_1, x_2, x_3)$ to $10X_1^2 + X_2^2 + X_3^2$. ∎

Note: If the orthogonal transformation was not required for the diagonalization, we could do it easily by completing squares:

$$\begin{aligned}
& 5x_1^2 + 5x_2^2 + 2x_3^2 + 8x_1x_2 + 4x_1x_3 + 4x_2x_3 \\
= \; & 5[x_1^2 + \frac{8}{5}x_1x_2 + \frac{4}{5}x_1x_3] + 5x_2^2 + 2x_3^2 + 4x_2x_3 \\
= \; & 5[x_1 + \frac{4}{5}x_2 + \frac{2}{5}x_3]^2 + (5 - \frac{16}{25})x_2^2 + (2 - \frac{4}{5})x_3^2 + (4 - \frac{16}{5})x_2x_3 \\
= \; & 5[x_1 + \frac{4}{5}x_2 + \frac{2}{5}x_3]^2 + \frac{9}{5}(x_2^2 + \frac{4}{9}x_2x_3) + \frac{6}{5}x_3^2 \\
= \; & 5[x_1 + \frac{4}{5}x_2 + \frac{2}{5}x_3]^2 + \frac{9}{5}[x_2 + \frac{2}{9}x_3]^2 + \frac{10}{9}x_3^2 \\
= \; & 5X^2 + \frac{9}{5}Y^2 + \frac{10}{9}Z^2
\end{aligned}$$

where $X = x_1 + \frac{4}{5}x_2 + \frac{2}{5}x_3, Y = x_2 + \frac{2}{9}x_3, Z = x_3$, or $x_3 = Z, x_2 = Y - \frac{2}{9}Z, x_1 = X - \frac{4}{5}Y - \frac{2}{9}Z$.

Question 2(a) *Define a vector space. Show that the set $\mathcal{V}$ of all real-valued functions on $[0, 1]$ is a vector space over the set of real numbers with respect to the addition and scalar multiplication of functions.*

Solution. See 1984 question 4(a). ∎

Question 2(b) *If zero is a root of the characteristic equation of a matrix $\mathbf{A}$, show that the corresponding linear transformation cannot be one to one.*

Solution. If zero is a root of $|\mathbf{A} - \lambda\mathbf{I}| = 0$, the characteristic equation of $\mathbf{A}$, then 0 is an eigenvalue of $\mathbf{A}$, so there is a non-zero eigenvector $\mathbf{x}$ such that $\mathbf{Ax} = \mathbf{0}$, thus $\mathbf{A}$ is not 1-1. ∎

Question 2(c) *Show that a linear transformation $\mathbf{T}$ from a Euclidean space $\mathcal{V}$ to $\mathcal{V}$ is orthogonal if and only if the matrix corresponding to it with respect to any orthonormal basis is orthogonal.*

Solution. $\mathbf{T} : \mathcal{V} \longrightarrow \mathcal{V}$ is said to be orthogonal if $\langle \mathbf{T}(\mathbf{u}), \mathbf{T}(\mathbf{v}) \rangle = \langle \mathbf{u}, \mathbf{v} \rangle$ for any $\mathbf{u}, \mathbf{v} \in \mathcal{V}$.

Lemma 1. $\mathbf{T}$ is orthogonal iff $\mathbf{T}$ takes an orthonormal basis to an orthonormal basis.

Proof: Let $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ be an orthonormal basis. Then

1. $\langle \mathbf{T}(\mathbf{v_i}), \mathbf{T}(\mathbf{v_j}) \rangle = \langle \mathbf{v_i}, \mathbf{v_j} \rangle = 0$ if $i \neq j$
2. $\langle \mathbf{T}(\mathbf{v_i}), \mathbf{T}(\mathbf{v_i}) \rangle = \langle \mathbf{v_i}, \mathbf{v_i} \rangle = 1$
3. If $\sum_{i=1}^n \alpha_i \mathbf{T}(\mathbf{v_i}) = 0$, then $\langle \sum_{i=1}^n \alpha_i \mathbf{T}(\mathbf{v_i}), \mathbf{v_j} \rangle = \alpha_j = 0$ for all j, so $\mathbf{T}(\mathbf{v_i})$ are linearly independent.

Thus $\mathbf{T}(\mathbf{v_1}), \dots, \mathbf{T}(\mathbf{v_n})$ form an orthonormal basis.

Conversely, let $\mathbf{T}(\mathbf{v_1}), \dots, \mathbf{T}(\mathbf{v_n})$ be an orthonormal basis of $\mathcal{V}$. Let $\mathbf{v} = \sum_{i=1}^n \alpha_i \mathbf{v_i}, \mathbf{w} = \sum_{i=1}^n \beta_i \mathbf{v_i}$, then $\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^n \alpha_i \beta_i$ and $\langle \mathbf{T}(\mathbf{v}), \mathbf{T}(\mathbf{w}) \rangle = \langle \sum_{i=1}^n \alpha_i \mathbf{T}(\mathbf{v_i}), \sum_{i=1}^n \beta_i \mathbf{T}(\mathbf{v_i}) \rangle = \sum_{i=1}^n \alpha_i \beta_i$. Thus $\langle \mathbf{T}(\mathbf{v}), \mathbf{T}(\mathbf{w}) \rangle = \langle \mathbf{v}, \mathbf{w} \rangle$, so $\mathbf{T}$ is orthogonal.

Lemma 2. Let $\mathbf{T}^*$ be defined by $\langle \mathbf{T}(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{T}^*(\mathbf{w}) \rangle$. Then $\mathbf{T}^*$ is a linear transformation, and $\mathbf{T}$ is orthogonal iff $\mathbf{T}^*\mathbf{T} = \mathbf{T}\mathbf{T}^* = \mathbf{I}$.

Proof: The fact that $\mathbf{T}^*$ is a linear transformation can be easily checked. If $\mathbf{T}$ is orthogonal, then $\langle \mathbf{v}, \mathbf{T}^*\mathbf{T}(\mathbf{w}) \rangle = \langle \mathbf{T}(\mathbf{v}), \mathbf{T}(\mathbf{w}) \rangle = \langle \mathbf{v}, \mathbf{w} \rangle$, so $\mathbf{T}^*\mathbf{T} = \mathbf{I}$. From this and the fact that $\mathbf{T}$ is 1-1, it follows that $\mathbf{T}\mathbf{T}^* = \mathbf{I}$.

Lemma 3. If the matrix of $\mathbf{T}$ w.r.t. the orthonormal basis $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ is $\mathbf{A} = (a_{ij})$, then the matrix of $\mathbf{T}^*$ is the transpose, i.e. (a_{ji}).

Proof: $\mathbf{T}(\mathbf{v_i}) = \sum_{j=1}^n a_{ij} \mathbf{v_j}$. Let $\mathbf{T}^*(\mathbf{v_i}) = \sum_{j=1}^n b_{ij} \mathbf{v_j}$. Now $b_{ij} = \langle \mathbf{T}^*(\mathbf{v_i}), \mathbf{v_j} \rangle = \langle \mathbf{v_i}, \mathbf{T}(\mathbf{v_j}) \rangle = \langle \mathbf{v_i}, \sum_{k=1}^n a_{jk} \mathbf{v_k} \rangle = a_{ji}$. Since $\mathbf{T}\mathbf{T}^* = \mathbf{I}$, $\mathbf{A}'\mathbf{A} = \mathbf{A}\mathbf{A}' = \mathbf{I}$, so $\mathbf{A}$ is orthogonal.

The converse is also obvious now. ∎

Question 3(a) *Investigate for what values of λ and μ does the following system of equations*

$$\begin{aligned} x + y + z &= 6 \\ x + 2y + 3z &= 10 \\ x + 2y + \lambda z &= \mu \end{aligned}$$

have (1) a unique solution (2) no solution (3) an infinite number of solutions?

Solution.

1. A unique solution exists when $\begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 2 & \lambda \end{vmatrix} \neq 0$, whatever μ may be. Thus $2\lambda - 6 - (\lambda - 3) \neq 0 \Rightarrow \lambda \neq 3$. Thus for all $\lambda \neq 3$ and for all μ we have a unique solution given by $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 2 & \lambda \end{pmatrix}^{-1} \begin{pmatrix} 6 \\ 10 \\ \mu \end{pmatrix}$

2. A unique solution does not exist if $\lambda = 3$. If $\mu \neq 10$, then the second and third equations are inconsistent. Thus if $\lambda = 3, \mu \neq 10$, the system has no solution.

3. If $\lambda = 3, \mu = 10$, then the system is $x+y+z = 6, x+2y+3z = 10$. The coefficient matrix is of rank 2, so the space of solutions is one dimensional. $y + 2z = 4 \Rightarrow y = 4 - 2z$, and thus $x = 2 + z$. The space of solutions is $(2 + z, 4 - 2z, z)$ for $z \in \mathbb{R}$.

∎

Question 3(b) *Let $(x_i, y_i), i = 1, \ldots, n$ be n points in the plane, no two of them having the same abscissa. Find a polynomial $f(x)$ of degree $n-1$ which takes the value $f(x_i) = y_i, 1 \leq i \leq n$.*

Solution. Let $f(x) = a_0 + a_1 x + \ldots + a_{n-1}x^{n-1}$. We want to determine $a_0, \ldots, a_{n-1}$ such that

$$\mathbf{A}\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & x_1 & \ldots & x_1^{n-1} \\ 1 & x_2 & \ldots & x_2^{n-1} \\ \vdots & & & \\ 1 & x_n & \ldots & x_n^{n-1} \end{pmatrix}\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

This is possible as $|\mathbf{A}| \neq 0$, as $x_1, \ldots, x_n$ are distinct. $\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix} = \mathbf{A}^{-1}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$.

Note: We can also use Lagrange's interpolation formula from numerical analysis, giving

$$f(x) = \sum_{i-1}^{n} y_i \frac{\prod_{j\neq i}(x - x_j)}{\prod_{j\neq i}(x_i - x_j)}$$

The two methods give the same polynomial, which is unique. ∎

Paper II

Question 4(a) *Find a set of three orthonormal eigenvectors for the matrix* $\mathbf{A} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 4 & \sqrt{3} \\ 0 & \sqrt{3} & 6 \end{pmatrix}$

Solution. The characteristic equation of $\mathbf{A}$ is

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 3-\lambda & 0 & 0 \\ 0 & 4-\lambda & \sqrt{3} \\ 0 & \sqrt{3} & 6-\lambda \end{vmatrix} = 0$$

Thus $(3-\lambda)(4-\lambda)(6-\lambda) - 3(3-\lambda) = 0 \Rightarrow \lambda = 3, \lambda^2 - 10\lambda + 21 = 0$. Thus the eigenvalues of $\mathbf{A}$ are $3, 3, 7$.

Let (x_1, x_2, x_3) be an eigenvector for $\lambda = 7$. Then

$$\begin{pmatrix} -4 & 0 & 0 \\ 0 & -3 & \sqrt{3} \\ 0 & \sqrt{3} & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{0}$$

Thus $-4x_1 = 0, -3x_2 + \sqrt{3}x_3 = 0, \sqrt{3}x_2 - x_3 = 0$. Thus $x_1 = 0, x_3 = \sqrt{3}x_2$ with $x_2 \neq 0$ gives any eigenvector for $\lambda = 7$. Take $x_2 = 1$ to get $(0, 1, \sqrt{3})$, and normalize it to get $(0, \frac{1}{2}, \frac{\sqrt{3}}{2})$.

Let (x_1, x_2, x_3) be an eigenvector for $\lambda = 3$. Then

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & \sqrt{3} \\ 0 & \sqrt{3} & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{0}$$

Thus $x_2 + \sqrt{3}x_3 = 0$. Thus $(x_1, -\sqrt{3}x_3, x_3)$ with $x_1, x_3 \in \mathbb{R}$ gives any eigenvector for $\lambda = 3$. We can take $x_1 = 1, x_3 = 0$, and $x_1 = 0, x_3 = 1$ to get $(1, 0, 0), (0, -\sqrt{3}, 1)$ as eigenvectors for $\lambda = 3$ — these are orthogonal and therefore span the the eigenspace of $\lambda = 3$. Orthonormal vectors are $(1, 0, 0), (0, -\frac{\sqrt{3}}{2}, \frac{1}{2})$.

Thus the required orthonormal vectors are $(0, \frac{1}{2}, \frac{\sqrt{3}}{2}), (1, 0, 0), (0, -\frac{\sqrt{3}}{2}, \frac{1}{2})$.

In fact

$$\begin{pmatrix} 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2} \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 4 & \sqrt{3} \\ 0 & \sqrt{3} & 6 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \end{pmatrix} = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

■

Question 4(b) *Show that if $A = \mathbf{X'AX}$ and $B = \mathbf{X'BX}$ are two quadratic forms one of which is positive definite and $\mathbf{A}, \mathbf{B}$ are symmetric matrices, then they can be expressed as linear combinations of squares by an appropriate linear transformation.*

Solution. Let $\mathbf{B}$ be positive definite. Then there exists an orthogonal real non-singular matrix $\mathbf{H}$ such that $\mathbf{H'BH} = \mathbf{I_n}$, the unit matrix of order n. $\mathbf{A}$ is real-symmetric $\Rightarrow \mathbf{H'AH}$ is real symmetric. There exists $\mathbf{K}$ a real orthogonal matrix such that $\mathbf{K'H'AHK}$ is a diagonal matrix i.e. $\mathbf{K'H'AHK} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$ where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of $\mathbf{H'AH}$.

Now $\mathbf{K'H'BHK} = \mathbf{K'I_nK} = \mathbf{I_n}$. Then $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \mathbf{HK} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ diagonalizes $\mathbf{A}, \mathbf{B}$ simultaneously.

$$\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} \mathbf{A} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \lambda_1 X_1^2 + \ldots + \lambda_n X_n^2 \qquad \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} \mathbf{B} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = X_1^2 + \ldots + X_n^2$$

Note that $\lambda_1, \dots, \lambda_n$ are the roots of $|\mathbf{A} - \lambda\mathbf{B}| = 0$ because $|\mathbf{A} - \lambda\mathbf{B}| = |\mathbf{H'}||\mathbf{A} - \lambda\mathbf{B}||\mathbf{H}| = |\mathbf{H'AH} - \lambda\mathbf{H'BH}| = |\mathbf{H'AH} - \lambda\mathbf{I_n}| = \prod_{i=1}^n (\lambda - \lambda_i)$.

■