Exercises:

1. For each of the following matrices find a diagonalization or show one does not exist.

   (a) $\begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$.

   (b) $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$.

   (c) $\begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix}$.

2. Let $A \in M_{n\times n}$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ counting multiplicities. Prove the following formulas:

   (a) $\mathrm{Tr}(A) = \lambda_1 + \cdots + \lambda_n$.

   (b) $\det(A) = \lambda_1 \cdots \lambda_n$.

3. Suppose $A \in M_{n\times n}$ is diagonalizable and $\mathrm{char}_A(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots a_1 z + a_0$. Show that

$$a_n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I_n = O.$$

   [**Note:** this also holds for non-diagonalizable matrices and is called the Cayley–Hamilton theorem.]

4. Suppose $A \in M_{n\times n}$ has distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, and consider the following subspace of $M_{n\times n}$:

$$V := \mathrm{span}\{I, A, A^2, A^3, \ldots\}.$$

   Prove that $\dim(V) = n$.

5. Verify that $\langle A, B\rangle_2 := \mathrm{Tr}(B^* A)$ defines an inner product on $M_{m\times n}$.

6. Let V be an inner product space with a spanning system $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Prove that for $\mathbf{x} \in V$, $\langle \mathbf{x}, \mathbf{v}_k\rangle = 0$ for each $k = 1, \ldots, n$ if and only if $\mathbf{x} = \mathbf{0}$.

Solutions:

1. (a) By Exercise 5 on Homework 8, $\mathrm{char}_A(z) = (z-2)^2$ and $(1,1)^T$ is a basis for the eigenspace $\mathrm{Ker}(A-2I)$. Thus the geometric multipliicty of $\lambda = 2$ is strictly less than its algebraic multiplicity. Therefore A is not diagonalizable.

   (b) By Exercise 5 on Homework 8, $\sigma(A) = \frac{1\pm\sqrt{5}}{2}$. Denote these quantities by φ and τ, respectively. The same exercise showed $(\varphi, 1)^T$ is a basis for the eigenspace $\mathrm{Ker}(A - \varphi I)$, and $(\tau, 1)^T$ is a basis for the eigenspace $\mathrm{Ker}(A - \tau I)$. Thus a diagaonlization of A is given by

$$A = \begin{pmatrix} \varphi & \tau \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \varphi & 0 \\ 0 & \tau \end{pmatrix} \begin{pmatrix} \varphi & \tau \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \varphi & \tau \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \varphi & 0 \\ 0 & \tau \end{pmatrix} \frac{1}{\varphi - \tau} \begin{pmatrix} 1 & -\tau \\ -1 & \varphi \end{pmatrix}.$$

   (c) By Exercise 5 on Homework 8, $\sigma(A) = \{-2, 1\}$. Also $(-1,1,0)^T$ and $(-1,0,1)^T$ is a basis for the eigenspace $\mathrm{Ker}(A-(-2)I)$, and $(1,-1,1)^T$ is a basis for the eigenspace $\mathrm{Ker}(A - 1I)$. Thus a diagonalization of A is given by

$$A = \begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}^{-1}$$

To compute the inverse, we do row operations on an augmented matrix:

$$\left(\begin{array}{rrr|rrr} -1 & -1 & 1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \overset{R1\to R_3\to R_2\to R_1}{\longrightarrow} \left(\begin{array}{rrr|rrr} 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ -1 & -1 & 1 & 1 & 0 & 0 \end{array}\right)$$

$$\overset{R_3\mapsto R_3+R_1+R_2}{\longrightarrow} \left(\begin{array}{rrr|rrr} 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right)$$

$$\underset{R_2\mapsto R_2-R_3}{\overset{R_1\mapsto R_1+R_3}{\longrightarrow}} \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & 1 & 2 & 1 \\ 0 & 1 & 0 & -1 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right).$$

Thus A has diagonalization:

$$A = \begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ -1 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

2. By Theorem 4.12, there exists an invertible matrix $Q \in M_{n\times n}$ so that

$$A = Q \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ & & \lambda_n \end{pmatrix} Q^{-1}.$$

Denote the upper triangular matrix by T. Then

$$\mathrm{Tr}(A) = \mathrm{Tr}(QTQ^{-1}) = \mathrm{Tr}(Q^{-1}QT) = \mathrm{Tr}(T) = \lambda_1 + \cdots + \lambda_n.$$

Also,

$$\det(A) = \det(QTQ^{-1}) = \det(Q)\det(T)\det(Q^{-1}) = \det(T) = \lambda_1 \cdots \lambda_n.$$

□

3. Let $A = QDQ^{-1}$ be a diagonalization of A, where $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A counting multiplicities. By the functional calculus, we have

$$a_n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I_n = \mathrm{char}_A(A) = Q\mathrm{char}_A(D)Q^{-1}.$$

Now, since $\lambda_1, \ldots, \lambda_n$ are the roots of $\mathrm{char}_A(z)$, we have

$$\mathrm{char}_A(D) = \begin{pmatrix} \mathrm{char}_A(\lambda_1) & & \\ & \ddots & \\ & & \mathrm{char}_A(\lambda_n) \end{pmatrix} = \begin{pmatrix} 0 & & \\ & \ddots & \\ & & 0 \end{pmatrix} = O.$$

Hence $a_n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I_n = QOQ^{-1} = O$. □

4. We claim that $\mathcal{B} := \{I, A, A^2, \ldots, A^{n-1}\}$ is a basis for V, in which case we will have $\dim(V) = n$. We first show that the system is spanning. Define

$$p(z) := \mathrm{char}_A(z) = (z - \lambda_1)\cdots(z - \lambda_n).$$

Note that by expanding we have $p(z) = z^n + q(z)$ where $q(z)$ is degree at most $n-1$. Also note that by Corollary 4.10, A is diagonalizable.

We claim that for any $k \geq n$ we have $A^k \in \mathrm{span}\,\mathcal{B}$. We will proceed by induction on k. For the base case, $k = n$, we have by Exercise 3 that $p(A) = O$. Thus

$$O = p(A) = A^n + q(A),$$

so that $A^n = -q(A)$. Since q is degree at most $n-1$, $A^n = -q(A) \in \operatorname{span}\mathcal{B}$. Now, for the induction step assume the claim holds for k; that is,

$$A^k = \alpha_0 I + \alpha_1 A + \cdots + \alpha_{n-2}A^{n-2} + \alpha_{n-1}A^{n-1}$$

for some scalars $\alpha_0, \ldots, \alpha_{n-1}$. Then using the base case again we have

$$A^{k+1} = \alpha_0 A + \alpha_1 A^2 + \cdots + \alpha_{n-2}A^{n-1} + \alpha_{n-1}A^n = \alpha_0 A + \alpha_1 A^2 + \cdots + \alpha_{n-2}A^{n-1} + \alpha_{n-1}(-q(A)),$$

which is in $\operatorname{span}\mathcal{B}$. So by induction $A^k \in \operatorname{span}\mathcal{B}$ for all $k \geq 0$ and hence $\mathcal{B}$ is spanning.

Towards showing $\mathcal{B}$ is linearly independent (and hence a basis) suppose

$$\alpha_0 I + \alpha_1 A + \cdots + \alpha_{n-1}A^{n-1} = O$$

for some scalars $\alpha_0, \ldots, \alpha_{n-1}$. Assume, towards a contradiction, that not all of these scalars are zero. Then we can define a non-trivial polynomial

$$r(z) := \alpha_0 + \alpha_1 z + \cdots + \alpha_{n-1}z^{n-1}$$

and $r(A) = O$. If $A = QDQ^{-1}$ is a diagonalization of A, then we have

$$r(D) = Q^{-1}Qr(D)Q^{-1}Q = Q^{-1}r(A)Q = Q^{-1}OQ = O.$$

However, $r(D)$ is the diagonal matrix with entries $r(\lambda_1), \ldots, r(\lambda_n)$. So if it equals the zero matrix, then $\lambda_1, \ldots, \lambda_n$ must all be roots of r. Since $\lambda_1, \ldots, \lambda_n$ are all distinct, this would imply $r(z)$ has n distinct roots which contradicts $r(z)$ being of degree at most $n-1$. Thus it must be that $\alpha_0 = \cdots = \alpha_{n-1} = 0$, and so $\mathcal{B}$ is a linearly independent system. $\square$

5. We must check each of the four parts of the definition of an inner product. We first note that

$$\operatorname{Tr}(B^*A) = \sum_{j=1}^{n}(B^*A)_{j,j} = \sum_{j=1}^{n}\sum_{i=1}^{m}(B^*)_{j,i}(A)_{i,j} = \sum_{i=1}^{m}\sum_{j=1}^{n}\overline{(B)_{i,j}}(A)_{i,j}.$$

(1) **Conjugate Symmetry:** using the above formula, we have

$$\langle B, A\rangle_2 = \operatorname{Tr}(A^*B) = \sum_{i=1}^{m}\sum_{j=1}^{n}\overline{(A)_{i,j}}(B)_{i,j} = \overline{\sum_{i=1}^{m}\sum_{j=1}^{n}(A)_{i,j}\overline{(B)_{i,j}}} = \overline{\operatorname{Tr}(B^*A)} = \overline{\langle A, B\rangle_2}.$$

(2) **Linearity:** using the linearity of the trace we have

$$\begin{aligned}\langle \alpha A + \beta B, C\rangle_2 &= \operatorname{Tr}(C^*(\alpha A + \beta B)) = \operatorname{Tr}(\alpha C^*A + \beta C^*B)\\ &= \alpha\operatorname{Tr}(C^*A) + \beta\operatorname{Tr}(C^*B) = \alpha\langle A, C\rangle_2 + \beta\langle B, C\rangle_2.\end{aligned}$$

(3) **Non-negativity:** using the above formula we have

$$\langle A, A\rangle_2 = \sum_{i=1}^{m}\sum_{j=1}^{m}\overline{(A)_{i,j}}(A)_{i,j} = \sum_{i=1}^{m}\sum_{j=1}^{m}|(A)_{i,j}|^2,$$

which is greater than or equal to zero as a sum of non-negative numbers.

(4) **Non-degeneracy:** using the computation from the previous part, we see that $\langle A, A\rangle_2 = 0$ if and only if

$$\sum_{i=1}^{m}\sum_{j=1}^{m}|(A)_{i,j}|^2 = 0,$$

which is possible if and only if $(A)_{i,j} = 0$ for each $i = 1, \ldots, m$ and $j = 1, \ldots, n$. That is, if and only if $A = O$. $\square$

6. If $\mathbf{x} = \mathbf{0}$ then we immediately have $\langle \mathbf{x}, \mathbf{v}_k\rangle = 0$ for each $k = 1, \ldots, n$. Conversely, suppose $\langle \mathbf{x}, \mathbf{v}_k\rangle = 0$ for each $k = 1, \ldots, n$. Since $\mathbf{v}_1, \ldots, \mathbf{v}_n$ span V, there exists scalars $\alpha_1, \ldots, \alpha_n$ such that $\mathbf{x} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n$. Consequently,

$$\langle \mathbf{x}, \mathbf{x}\rangle = \langle \mathbf{x}, \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n\rangle = \bar{\alpha}_1\langle \mathbf{x}, \mathbf{v}_1\rangle + \cdots + \bar{\alpha}_n\langle \mathbf{x}, \mathbf{v}_n\rangle = \bar{\alpha}_1 0 + \cdots + \bar{\alpha}_n 0 = 0$$

Thus $\langle \mathbf{x}, \mathbf{x}\rangle = 0$ and so by non-degeneracy $\mathbf{x} = \mathbf{0}$. $\square$