MTL104 Linear Algebra and Its Applications
I Semester 2025-26
Practice Sheet IV

This practice sheet focuses on characteristic and minimal polynomials, invariant subspaces, diagonalization, the primary decomposition theorem, and basic concepts of inner product spaces.

1. Let T be a linear operator on a finite-dimensional vector space V over $\mathbb{C}$. Then T is diagonalizable if and only if there exists a polynomial $p(x) \in \mathbb{C}[x]$ with *distinct* roots that annihilates T, that is, $p(T) = 0$.

   Solution:

   ($\Rightarrow$) Assume T is diagonalizable. Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T, and write $E_{\lambda_i} = \ker(T - \lambda_i I)$. Since T is diagonalizable,

   $$V = \bigoplus_{i=1}^{k} E_{\lambda_i}.$$

   Define

   $$p(x) = \prod_{i=1}^{k} (x - \lambda_i).$$

   Then we have proved in our class that $p(T) = 0$.

   ($\Leftarrow$) Now suppose $p(x) = \prod_{i=1}^{k}(x - \lambda_i)$ has distinct roots and satisfies $p(T) = 0$. By the primary decomposition theorem,

   $$V = \bigoplus_{i=1}^{k} \ker((T - \lambda_i I)^{m_i}),$$

   where m_i is the multiplicity of $(x - \lambda_i)$ in the minimal polynomial of T. Since p has each root only to the first power, we must have $m_i = 1$ for all i, so

   $$V = \bigoplus_{i=1}^{k} \ker(T - \lambda_i I) = \bigoplus_{i=1}^{k} E_{\lambda_i}.$$

   Thus V has a basis consisting of eigenvectors of T, proving that T is diagonalizable.

2. Let T be the integral operator defined by

   $$(Tf)(x) = \int_0^x f(t)\, dt,$$

   on the space of continuous functions on the interval $[0, 1]$.

   Determine whether the following subspaces are invariant under T:

   (a) the space of polynomial functions,

   (b) the space of differentiable functions,

   (c) the space of functions which vanish at $x = \frac{1}{2}$.

Solution: Let $T : C([0,1]) \to C([0,1])$ be the operator

$$(Tf)(x) = \int_0^x f(t)\,dt.$$

A subspace $S \subset C([0,1])$ is invariant under T if $T(S) \subseteq S$. We examine each of the given subspaces.

(a) **Polynomials.**
If f is a polynomial, then $\int_0^x f(t)\,dt$ is again a polynomial.. Thus T maps polynomials to polynomials, so this subspace is invariant.

(b) **Differentiable functions.**
If f is differentiable on $[0,1]$, then by the Fundamental Theorem of Calculus the function

$$(Tf)(x) = \int_0^x f(t)\,dt$$

is differentiable and satisfies $(Tf)'(x) = f(x)$. Hence T maps differentiable functions to differentiable functions, and this subspace is invariant.

(c) **Functions vanishing at $x = \frac{1}{2}$.**
Let

$$S = \{f \in C([0,1]) : f(1/2) = 0\}.$$

For $f \in S$ we have

$$(Tf)\left(\tfrac{1}{2}\right) = \int_0^{1/2} f(t)\,dt,$$

which need not be zero even if $f(1/2) = 0$. For example, let $f(t) = t - \frac{1}{2}$. Then $f(1/2) = 0$, but

$$(Tf)\left(\tfrac{1}{2}\right) = \int_0^{1/2} (t - \tfrac{1}{2})\,dt = -\frac{1}{8} \neq 0.$$

Thus $T(f) \notin S$ for this f, and this subspace is *not* invariant under T.

3. Let V be a finite-dimensional vector space.

(a) What is the minimal polynomial for the identity operator on V?

(b) What is the minimal polynomial for the zero operator on V?

Solution.

Let V be a finite-dimensional vector space.

1. Minimal polynomial of the identity operator.

Let I be the identity operator on V and let

$$p(x) = a_0 + a_1 x + \cdots + a_k x^k$$

be a polynomial. Then

$$p(I) = a_0 I + a_1 I^2 + \cdots + a_k I^k.$$

Since $I^n = I$ for all $n \geq 1$, we obtain

$$p(I) = (a_0 + a_1 + \cdots + a_k) I = p(1)\, I.$$

Thus

$$p(I) = 0 \quad \Longleftrightarrow \quad p(1) = 0.$$

The monic polynomial of minimal degree with $p(1) = 0$ is

$$m_I(x) = x - 1.$$

2. Minimal polynomial of the zero operator.

Let 0 denote the zero operator. For a polynomial $p(x) = a_0 + a_1 x + \cdots + a_k x^k$ we have

$$p(0) = a_0.$$

Thus

$$p(0) = 0 \quad \Longleftrightarrow \quad a_0 = 0.$$

The monic polynomial of least degree with zero constant term is

$$m_0(x) = x.$$

Therefore,

$$m_I(x) = x - 1, \qquad m_0(x) = x.$$

4. Find a 3×3 matrix whose minimal polynomial is x^2.

solution

To find a 3×3 matrix whose minimal polynomial is x^2, we look for a nonzero matrix A satisfying $A^2 = 0$. This means we want

$$\mathrm{rg}(A) \subseteq \ker(A), \qquad A \neq 0.$$

Choose a basis $\{e_1, e_2, e_3\}$ and define

$$A(e_1) = e_2, \qquad A(e_2) = 0, \qquad A(e_3) = 0.$$

Then $A \neq 0$ and $A^2 = 0$, so the minimal polynomial is x^2.

In this basis the matrix of A is

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

Since $A^2 = 0$ but $A \neq 0$, the minimal polynomial is

$$m_A(x) = x^2.$$

5. Consider the polynomial space $V = \mathcal{P}_2(\mathbb{R})$ and its standard basis $\mathcal{B} = 1, x, x^2$. Let $T : \mathcal{P}_2(\mathbb{R}) \to \mathcal{P}_2(\mathbb{R})$ be a linear tranformation defined by

$$T(1) = 2 - 3x^2, \quad T(x) = 4 - 4x - 7x^2, \quad T(x^2) = -2 + 3x + 4x^2.$$

Find

(a) the characteristic polynomial of T.

(b) the minimal polynomial of T.

(c) a basis $\mathcal{B}'$ for V such that $[T]_{\mathcal{B}'}$ is diagonal.

Solution:

The matrix of T in the standard basis $\mathcal{B} = \{1, x, x^2\}$ is

$$[A]_{\mathcal{B}} = \begin{pmatrix} 2 & 4 & -2 \\ 0 & -4 & 3 \\ -3 & -7 & 4 \end{pmatrix},$$

since

$$T(1) = 2-3x^2 \mapsto \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}, \quad T(x) = 4-4x-7x^2 \mapsto \begin{pmatrix} 4 \\ -4 \\ -7 \end{pmatrix}, \quad T(x^2) = -2+3x+4x^2 \mapsto \begin{pmatrix} -2 \\ 3 \\ 4 \end{pmatrix}.$$

The characteristic polynomial is

$$\chi_A(\lambda) = \det(A - \lambda I) = \lambda^3 - 2\lambda^2 - \lambda + 2 = (\lambda - 2)(\lambda - 1)(\lambda + 1),$$

so the eigenvalues are $2, 1, -1$, each simple. Hence the minimal polynomial is

$$m_A(x) = (x - 2)(x - 1)(x + 1).$$

Eigenvectors are, for instance

$$v_{-1} = (-2, 3, 3)^\top \quad (\text{polynomial } -2 + 3x + 3x^2),$$

$$v_1 = (-2, 3, 5)^\top \quad (\text{polynomial } -2 + 3x + 5x^2),$$

$$v_2 = (-1, 1, 2)^\top \quad (\text{polynomial } -1 + x + 2x^2).$$

These form a basis

$$\mathcal{B}' = \{-2 + 3x + 3x^2,\ -2 + 3x + 5x^2,\ -1 + x + 2x^2\},$$

and with respect to $\mathcal{B}'$ the matrix of T is diagonal, e.g.

$$[T]_{\mathcal{B}'} = \operatorname{diag}(-1, 1, 2)$$

if the basis vectors are ordered as (v_{-1}, v_1, v_2).

6. Let V be a finite dimensional vector space over an algebraically closed field F, let $T\colon V \to V$ be linear, and let $f \in F[x]$ be a polynomial. Show that a scalar $c \in F$ is a characteristic value of $f(T)$ if and only if there exists a characteristic value t of T such that $f(t) = c$.

Solution

($\Rightarrow$) Let t be a characteristic value of T and let $\beta \neq 0$ be a characteristic vector, so $T\beta = t\beta$.

Then by induction,

$$T^k\beta = t^k\beta \qquad (k \geq 1).$$

If $f(x) = a_0 + a_1x + \cdots + a_mx^m$ is a polynomial, then

$$f(T)\beta = a_0\beta + a_1T\beta + \cdots + a_mT^m\beta = (a_0 + a_1t + \cdots + a_mt^m)\beta = f(t)\,\beta.$$

Since $\beta \neq 0$, this shows that $f(t)$ is a characteristic value of $f(T)$. Thus

$$\{\, f(t) : t \text{ a characteristic value of } T \,\} \subseteq \sigma(f(T)).$$

($\Leftarrow$) Now assume that c is a characteristic value of $f(T)$, i.e. $f(T) - cI$ is not invertible. To show that $c = f(t)$ for some characteristic value t of T, we use the fact that T is triangularizable. For completeness, we include a proof.

Claim: If F is algebraically closed and $\dim V = n$, then T is triangularizable.

Proof of claim. Since F is algebraically closed, the characteristic polynomial $\chi_T(x)$ splits, so T has an eigenvalue $\lambda \in F$. Choose a nonzero eigenvector v_1 and extend it to a basis $v_1, v_2, \ldots, v_n$ of V. In this basis, the matrix of T has the form

$$\begin{pmatrix} \lambda & * & \cdots & * \\ 0 & & & \\ \vdots & & T_1 & \\ 0 & & & \end{pmatrix},$$

where T_1 is the matrix of T restricted to $V_1 = \operatorname{span}\{v_2, \ldots, v_n\}$, a subspace of dimension $n-1$. The characteristic polynomial of T_1 divides χ_T and hence also splits, so T_1 has an eigenvalue. Choose an eigenvector $v_2 \in V_1$ and extend $\{v_1, v_2\}$ to a basis of V. Repeating this procedure inductively yields a basis in which the matrix of T is upper triangular. This proves the claim. $\square$

Having such a basis, let $A = P^{-1}TP$ be the upper triangular matrix representing T, whose diagonal entries $t_1, \ldots, t_n$ are exactly the characteristic values of T, counted with multiplicity.

Since a polynomial of an upper triangular matrix is again upper triangular with diagonal entries obtained by applying the polynomial to the diagonal entries, the matrix

$$P^{-1}f(T)P = f(P^{-1}TP) = f(A)$$

is upper triangular with diagonal entries $f(t_1), \ldots, f(t_n)$. Therefore the matrix of $f(T) - cI$ in this basis,

$$P^{-1}(f(T) - cI)P = f(A) - cI,$$

is also triangular with diagonal entries $f(t_i) - c$.

An upper triangular matrix is invertible iff all its diagonal entries are nonzero. Hence $f(T) - cI$ is *not* invertible exactly when $f(t_i) - c = 0$ for some i, i.e. when $c = f(t_i)$ for some characteristic value t_i of T.

Combining the two directions, we have shown:

$$c \in \sigma(f(T)) \quad \Longleftrightarrow \quad c = f(t) \text{ for some characteristic value } t \text{ of } T.$$

This completes the proof.

7. Let $T : V \to V$ be a linear operator such that every subspace of V is invariant under T. Prove that $T = \lambda I$ for some scalar λ.

   Solution If $\dim V = 0$, then $V = \{0\}$ and the desired result is plain.

   If $\dim V = 1$ choose a nonzero vector $v \in V$. Then $V = \operatorname{span}\{v\}$, so $T(v) = \lambda v$ for some $\lambda \in F$. For any $x \in V$, we have $x = av$,

$$T(x) = T(av) = aT(v) = a\lambda v = \lambda(av) = \lambda x = (\lambda I)(x).$$

   Thus $T = \lambda I$.

   Now suppose $\dim V > 1$. Choose linearly independent vectors $v, w \in V$. The one-dimensional subspace $\operatorname{span}\{v + w\}$ is also T-invariant, hence

$$T(v + w) = \lambda(v + w)(\, v + w\,)$$

   for some scalar $\lambda(v + w)$.

   On the other hand, using linearity and the formulas for $T(v)$ and $T(w)$,

$$T(v + w) = T(v) + T(w) = \lambda(v)v + \lambda(w)w.$$

   Equating the two expressions,

$$\lambda(v)v + \lambda(w)w = \lambda(v + w)v + \lambda(v + w)w.$$

   Since v, w are linearly independent, we conclude

$$\lambda(v) = \lambda(v + w) = \lambda(w).$$

   Thus $\lambda(u)$ is the same for every nonzero $u \in V$. Denote this common value by λ. Then for all nonzero $v \in V$ we have $T(v) = \lambda v$, and the equality trivially holds for $v = 0$ as well.

   Hence $T = \lambda I$, as desired.

8. Let $V = M_{n\times n}(F)$ be the vector space of all $n \times n$ matrices over a field F. Fix $A \in M_{n\times n}(F)$ and define a linear operator $T : V \to V$ by

$$T(B) = AB.$$

   Show that the minimal polynomial of the linear operator T equals the minimal polynomial of the matrix A.

   Solution

   Let $V = M_{n\times n}(F)$ and $T : V \to V$ be the linear map $T(B) = AB$ for a fixed $A \in M_{n\times n}(F)$. For a polynomial $p(x) = a_0 + a_1x + \cdots + a_kx^k \in F[x]$ we compute $p(T)$ acting on an arbitrary $B \in V$:

$$\begin{aligned} p(T)(B) &= (a_0I + a_1T + \cdots + a_kT^k)(B) \\ &= a_0B + a_1AB + \cdots + a_kA^kB \\ &= \big(a_0I + a_1A + \cdots + a_kA^k\big)B \end{aligned}$$

Thus for every $p \in F[x]$ we have the operator identity

$$p(T)(B) = p(A)B \qquad \text{for all } B \in V.$$

Now observe two implications:

- If $p(A) = 0$ then $p(T)(B) = p(A)B = 0$ for every B, so $p(T) = 0$ (the zero operator).
- Conversely, if $p(T) = 0$ as an operator, then in particular $0 = p(T)(I) = p(A)I = p(A)$, so $p(A) = 0$.

Therefore the set of polynomials that annihilate T equals the set of polynomials that annihilate A. The minimal polynomial of an operator/matrix is the unique monic polynomial of least degree in its annihilator ideal, hence the minimal polynomial of T equals the minimal polynomial of A.

9. If E is a projection and f is a polynomial, then show that

$$f(E) = aI + bE.$$

What are a and b in terms of the coefficients of f?

Solution Let $f(x) = c_0 + c_1x + c_2x^2 + \cdots + c_nx^n$. Since E is a projection, we have $E^2 = E$, and hence $E^k = E$ for all $k \geq 1$ (and $E^0 = I$). Therefore,

$$f(E) = c_0I + c_1E + c_2E^2 + \cdots + c_nE^n = c_0I + (c_1 + c_2 + \cdots + c_n)E.$$

Thus $f(E) = aI + bE$ where

$$a = c_0 = f(0), \qquad b = c_1 + c_2 + \cdots + c_n = f(1) - f(0).$$

10. Let V be an n-dimensional vector space over a field F and $T \in L(V)$. For $\lambda \in F$, let the *eigenspace* be $\ker(T - \lambda I)$. Its dimension is defined as the *geometric multiplicity* of λ. On the other hand, the *algebraic multiplicity* of λ is defined as its multiplicity as a root of the characteristic polynomial $\chi_T(x) := \det(T - xI) = \det([T]_{\mathcal{B}} - xI)$. Prove or disprove the following claims.

    (a) For each eigenvalue λ of T the geometric multiplicity of λ is at most its algebraic multiplicity.

    (b) T is diagonalizable if and only if for every eigenvalue λ the algebraic multiplicity equals the geometric multiplicity.

Solution

Notation and setup

Let V be an n-dimensional vector space over a field F and $T \in L(V)$. For $\lambda \in F$ the *eigenspace* is $E_\lambda := \ker(T - \lambda I)$; its dimension $g_\lambda := \dim E_\lambda$ is the *geometric multiplicity* of λ. The *algebraic multiplicity* m_λ is the multiplicity of λ as a root of the characteristic polynomial $\chi_T(x) = \det(T - xI)$.

Claim (1). For each eigenvalue λ of T we have

$$g_\lambda \leq m_\lambda.$$

Proof. Let $g := g_\lambda = \dim \ker(T - \lambda I)$. Choose a basis of the eigenspace

$$v_1, \ldots, v_g$$

and extend it to a basis of V by adding vectors

$$v_{g+1}, \ldots, v_n.$$

With respect to this basis the linear map T has a matrix of the block form

$$[T]_{\mathcal{B}} = \begin{pmatrix} \lambda I_g & * \\ 0 & A \end{pmatrix},$$

because $T(v_i) = \lambda v_i$ for $1 \leq i \leq g$ and T maps the complement into V (the lower-left block is zero). Hence the characteristic polynomial factors as

$$\chi_T(x) = \det([T]_{\mathcal{B}} - xI) = \det \begin{pmatrix} (\lambda - x)I_g & * \\ 0 & A - xI \end{pmatrix} = (\lambda - x)^g \det(A - xI).$$

Therefore $(x - \lambda)^g$ divides $\chi_T(x)$, so the algebraic multiplicity m_λ of λ satisfies $m_\lambda \geq g_\lambda$. $\square$

Claim (2). T is diagonalizable if and only if for every eigenvalue λ (in F) the algebraic multiplicity equals the geometric multiplicity.

Remark. One must assume the characteristic polynomial splits over F (i.e. all eigenvalues lie in F) to discuss algebraic multiplicities for a full set of eigenvalues; otherwise T cannot be diagonalized over F if there are no enough eigenvalues in F.

Proof. ($\Rightarrow$) If T is diagonalizable over F, there exists a basis of V consisting of eigenvectors of T. Grouping the basis vectors by eigenvalue shows that for each eigenvalue λ the eigenspace E_λ has a basis consisting of precisely the number of times λ appears on the diagonal of the diagonal matrix for T. Thus $g_\lambda = m_\lambda$ for every eigenvalue λ.

($\Leftarrow$) Conversely, suppose the characteristic polynomial of T splits over F and for every eigenvalue λ we have $g_\lambda = m_\lambda$. The algebraic multiplicities sum to n (the degree of χ_T), so

$$\sum_\lambda m_\lambda = n.$$

By hypothesis $m_\lambda = g_\lambda$, hence

$$\sum_\lambda g_\lambda = n.$$

For each λ choose a basis of the eigenspace E_λ; combining these bases for all distinct eigenvalues yields a collection of $\sum_\lambda g_\lambda = n$ linearly independent eigenvectors, hence a basis of V consisting of eigenvectors. Therefore T is diagonalizable. $\square$

Counterexample / caveat. If the characteristic polynomial does *not* split over F, there may be no eigenvalues in F (or not enough), and the statement "algebraic multiplicity equals geometric multiplicity for every eigenvalue" is vacuously true but T is not diagonalizable over F. For example, a 2×2 real rotation matrix has no real eigenvalues, so over $\mathbb{R}$ it is not diagonalizable even though there are no real eigenvalues to check. Thus the splitting hypothesis is important for the equivalence in (2).

11. Let T be a linear operator on a vector space V such that $\operatorname{rank}(T) = 1$. Prove that either T is diagonalizable or T is nilpotent, but not both.

Solution Let V be an n-dimensional vector space and $T : V \to V$ a linear map with $\operatorname{rank}(T) = 1$. Then $\dim \operatorname{Im}(T) = 1$, so $\dim \ker(T) = n - 1$.

Choose $0 \neq \beta \in \operatorname{Im}(T)$ and pick $\alpha_0 \in V$ with $T\alpha_0 = \beta$. Let $\{\alpha_1, \ldots, \alpha_{n-1}\}$ be a basis of $\ker(T)$. Then

$$\mathcal{B} = \{\alpha_0, \alpha_1, \ldots, \alpha_{n-1}\}$$

is a basis of V (indeed $\alpha_0 \notin \ker(T)$, so it is independent of the α_i's). For $i \geq 1$ we have $T\alpha_i = 0$, and $T\alpha_0 = \beta \in \operatorname{Im}(T) = \langle \beta \rangle$.

Write $T\alpha_0$ in the basis $\mathcal{B}$. There are two cases.

Case 1. $T\alpha_0 \in \ker(T)$.
Then $T\alpha_0 = c_1\alpha_1 + \cdots + c_{n-1}\alpha_{n-1}$ with at least one $c_i \neq 0$ (otherwise $\alpha_0 \in \ker(T)$, contradiction). The matrix of T in the basis $\mathcal{B}$ has the form

$$[T]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ c_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ c_{n-1} & 0 & \cdots & 0 \end{pmatrix}.$$

A direct computation shows $T^2 = 0$ (every column of $[T]_{\mathcal{B}}^2$ is zero), so T is nilpotent (of index at most 2).

Case 2. $T\alpha_0 \notin \ker(T)$.
Then $T\alpha_0$ is a nonzero scalar multiple of β (since the image is one-dimensional), say $T\beta = c_0\beta$ for some $c_0 \in F$. Put $\mathcal{B}' = \{\beta, \alpha_1, \ldots, \alpha_{n-1}\}$. In this basis

$$[T]_{\mathcal{B}'} = \begin{pmatrix} c_0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

which is diagonal, so T is diagonalizable.

Finally, show the two possibilities are mutually exclusive in the nonzero rank-one case: a nonzero nilpotent operator cannot be diagonalizable, because a nilpotent diagonalizable operator would have all eigenvalues equal to 0 and admit a basis of eigenvectors, hence be the zero operator. Since $\operatorname{rank}(T) = 1$, $T \neq 0$, so the nilpotent and diagonalizable cases cannot both occur.

Hence a linear operator of rank 1 is either nilpotent or diagonalizable, and not both.

12. Let T be a linear operator on a finite-dimensional vector space V over a field $\mathbb{F}$. Suppose the characteristic and minimal polynomials of T are given by

$$f(x) = \prod_{i=1}^{k} (x - \lambda_i)^{d_i} \quad \text{and} \quad p(x) = \prod_{i=1}^{k} (x - \lambda_i)^{r_i},$$

where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the distinct eigenvalues of T.

Define

$$W_i = \ker\big((T - \lambda_i I)^{r_i}\big).$$

(a) Prove that W_i is the set of all vectors $a \in V$ such that

$$(T - \lambda_i I)^m a = 0 \quad \text{for some positive integer } m \text{ (which may depend upon } a).$$

(b) Prove that the dimension of W_i is d_i.

Solution

Let

$$f(x) = \prod_{i=1}^{k} (x - \lambda_i)^{d_i}, \qquad p(x) = \prod_{i=1}^{k} (x - \lambda_i)^{r_i},$$

be the characteristic and minimal polynomials of T, where $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T. Define

$$W_i = \ker\big((T - \lambda_i I)^{r_i}\big).$$

(1) Description of W_i as generalized eigenvectors

Let

$$S_i = \{\, v \in V : (T - \lambda_i I)^m v = 0 \text{ for some } m \geq 1 \,\}.$$

Clearly,

$$W_i \subset S_i.$$

Now take $v \in S_i$. Then $(T - \lambda_i I)^m v = 0$ for some $m \geq 1$. Using the primary decomposition

$$V = W_1 \oplus \cdots \oplus W_k,$$

write

$$v = v_1 + \cdots + v_k, \qquad v_j \in W_j.$$

Since each W_j is T-invariant,

$$0 = (T - \lambda_i I)^m v = \sum_{j=1}^{k} (T - \lambda_i I)^m v_j,$$

and each term on the right lies in W_j. Because the sum is direct, each term must be zero:

$$(T - \lambda_i I)^m v_j = 0 \quad \text{for all } j.$$

Fix $j \neq i$. The polynomials $(x-\lambda_i)^m$ and $(x-\lambda_j)^{r_j}$ are relatively prime. Thus, there exist polynomials $a(x), b(x)$ such that

$$a(x)(x-\lambda_i)^m + b(x)(x-\lambda_j)^{r_j} = 1.$$

Applying this identity with $x \mapsto T$ to $v_j \in W_j$ gives

$$v_j = a(T)(T-\lambda_i I)^m v_j + b(T)(T-\lambda_j I)^{r_j} v_j = 0 + 0 = 0.$$

Thus $v_j = 0$ for all $j \neq i$, so $v = v_i \in W_i$.

Hence $S_i \subset W_i$, and therefore

$$S_i = W_i.$$

(2) Dimension of W_i

From the primary decomposition

$$V = W_1 \oplus \cdots \oplus W_k$$

and T-invariance of each W_j, we may choose a basis adapted to this decomposition, making the matrix of T block diagonal with blocks $T|_{W_j}$. Thus the characteristic polynomial splits as

$$\chi_T(x) = \prod_{j=1}^{k} \chi_{T|_{W_j}}(x).$$

On each W_j, the operator $(T-\lambda_j I)^{r_j}$ is nilpotent, so $T|_{W_j}$ has the single eigenvalue λ_j. Therefore

$$\chi_{T|_{W_j}}(x) = (x-\lambda_j)^{\dim W_j}.$$

Hence

$$\chi_T(x) = \prod_{j=1}^{k} (x-\lambda_j)^{\dim W_j}.$$

Comparing this with the given characteristic polynomial

$$\chi_T(x) = \prod_{j=1}^{k} (x-\lambda_j)^{d_j},$$

we conclude that

$$\dim W_j = d_j \qquad \text{for all } j.$$

In particular,

$$\dim W_i = d_i.$$

$\square$

13. Let V be an inner product space. Prove the following identities:

(a) *Parallelogram identity*: For all $x, y \in V$,

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

Look at $\mathbb{R}^2$, and interpret this identity geometrically, that justifies its name.

(b) *Polarization identity*:

$$\langle u, v\rangle = \frac{1}{4}\Big(\|u+v\|^2 - \|u-v\|^2 + i\big(\|u+iv\|^2 - \|u-iv\|^2\big)\Big), \qquad \forall\, u, v \in V.$$

Solution

Parallelogram identity. Let V be an inner product space. For all $x, y \in V$,

$$\|x+y\|^2 + \|x-y\|^2 = 2\big(\|x\|^2 + \|y\|^2\big).$$

Proof. Expand using bilinearity:

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle,$$

$$\|x-y\|^2 = \langle x-y, x-y\rangle = \langle x, x\rangle - \langle x, y\rangle - \langle y, x\rangle + \langle y, y\rangle.$$

Adding the above gives

$$\|x+y\|^2 + \|x-y\|^2 = 2\langle x, x\rangle + 2\langle y, y\rangle = 2(\|x\|^2 + \|y\|^2).$$

Geometric Interpretation in $\mathbb{R}^2$. If x and y form two adjacent sides of a parallelogram, then $x+y$ and $x-y$ are its diagonals. The identity states:

$$\text{(sum of squares of diagonal lengths)} = 2 \times \text{(sum of squares of side lengths)},$$

which is the classical parallelogram law.

Polarization identity (complex case). For all $u, v \in V$,

$$\langle u, v\rangle = \frac{1}{4}\left(\|u+v\|^2 - \|u-v\|^2 + i\big(\|u+iv\|^2 - \|u-iv\|^2\big)\right).$$

Proof. First,

$$\|u+v\|^2 - \|u-v\|^2 = 4\,\Re\langle u, v\rangle.$$

Next compute the imaginary part using $\langle u, iv\rangle = -i\langle u, v\rangle$:

$$\|u+iv\|^2 = \|u\|^2 + \|v\|^2 + 2\,\Im\langle u, v\rangle,$$

$$\|u-iv\|^2 = \|u\|^2 + \|v\|^2 - 2\,\Im\langle u, v\rangle.$$

Subtracting,

$$\|u+iv\|^2 - \|u-iv\|^2 = 4\,\Im\langle u, v\rangle.$$

Combining real and imaginary parts recovers the inner product:

$$\langle u, v\rangle = \frac{1}{4}\Big(\|u+v\|^2 - \|u-v\|^2 + i(\|u+iv\|^2 - \|u-iv\|^2)\Big).$$

14. (a) Suppose a, b, c, d are positive numbers. Prove that

$$(a+b+c+d)\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right) \geq 16.$$

(b) Prove that for all real numbers $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$,

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 \leq \left(\sum_{k=1}^{n} k a_k^2\right)\left(\sum_{k=1}^{n} \frac{b_k^2}{k}\right).$$

Solution

Suppose $a, b, c, d > 0$. Consider the inner product space $\mathbb{R}^4$ with the standard inner product

$$\langle x, y\rangle = \sum_{i=1}^{4} x_i y_i.$$

Apply the Cauchy–Schwarz inequality to the vectors

$$u = (\sqrt{a}, \sqrt{b}, \sqrt{c}, \sqrt{d}), \qquad v = \left(\frac{1}{\sqrt{a}}, \frac{1}{\sqrt{b}}, \frac{1}{\sqrt{c}}, \frac{1}{\sqrt{d}}\right).$$

Then

$$\langle u, v\rangle = \sum_{i=1}^{4} \sqrt{a_i}\,\frac{1}{\sqrt{a_i}} = 4.$$

Cauchy–Schwarz gives

$$4^2 = \langle u, v\rangle^2 \leq \langle u, u\rangle\,\langle v, v\rangle = (a+b+c+d)\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right),$$

which proves the required inequality.

Consider the inner product space $\mathbb{R}^n$ with the weighted inner product

$$\langle x, y\rangle = \sum_{k=1}^{n} w_k\, x_k y_k,$$

where the weights $w_k \geq 0$ is fixed. Consider the fixed weight $(1, 2, \ldots, n)$. Apply the Cauchy–Schwarz inequality to the vectors

$$u = (a_1, a_2, \ldots, a_n), \qquad v = \left(\frac{b_1}{1}, \frac{b_2}{2}, \ldots, \frac{b_n}{n}\right).$$

Then

$$\langle u, v\rangle = \sum_{k=1}^{n} k\, a_k \cdot \frac{b_k}{k} = \sum_{k=1}^{n} a_k b_k.$$

By Cauchy–Schwarz,

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 = \langle u, v\rangle^2 \leq \langle u, u\rangle\,\langle v, v\rangle = \left(\sum_{k=1}^{n} k a_k^2\right)\left(\sum_{k=1}^{n} \frac{b_k^2}{k}\right).$$

This proves the desired inequality.

15. Let $V = \mathcal{P}(\mathbb{R})$ and define

$$\langle p, q\rangle = \int_0^1 p(x)q(x)\,dx \text{ for all } p, q \in \mathcal{P}(\mathbb{R}).$$

(a) Verify that $\langle , \rangle$ is indeed an inner product on V.

(b) Apply the Gram–Schmidt procedure to the basis $\{1, x, x^2\}$ to produce an orthonormal basis of $\mathcal{P}_2(\mathbb{R})$.

Solution

Let $V = \mathcal{P}(\mathbb{R})$ be the vector space of real polynomials and define

$$\langle p, q\rangle := \int_0^1 p(x)q(x)\,dx, \qquad p, q \in V.$$

(a) Verification that $\langle \cdot, \cdot \rangle$ is an inner product

We check the three properties.

(i) Bilinearity. For any polynomials $p, q, r \in V$ and scalars $\alpha, \beta \in \mathbb{R}$,

$$\langle \alpha p+\beta q,\, r\rangle = \int_0^1 (\alpha p(x)+\beta q(x))r(x)\,dx = \alpha \int_0^1 p(x)r(x)\,dx+\beta \int_0^1 q(x)r(x)\,dx = \alpha\langle p, r\rangle+\beta\langle q, r\rangle.$$

Thus the map is linear in the first slot (and by symmetry also linear in the second slot for real scalars).

(ii) Symmetry. For all $p, q \in V$,

$$\langle p, q\rangle = \int_0^1 p(x)q(x)\,dx = \int_0^1 q(x)p(x)\,dx = \langle q, p\rangle.$$

(iii) Positive-definiteness. For any $p \in V$,

$$\langle p, p\rangle = \int_0^1 p(x)^2\,dx \geq 0,$$

and if $\langle p, p\rangle = 0$ then $\int_0^1 p(x)^2\,dx = 0$. Since $p(x)^2 \geq 0$ on $[0, 1]$ and is a continuous function, this integral vanishing forces $p(x)^2 = 0$ for all $x \in [0, 1]$, hence $p \equiv 0$. Thus $\langle p, p\rangle = 0$ iff $p = 0$.

Therefore $\langle \cdot, \cdot \rangle$ is an inner product on V.

Note that the integral of a nonnegative function f is zero does not imply the integrand f is a zero function, without the continuity assumption.

(b) Orthonormal Basis

$$u_1 = 1, \qquad u_2 = x, \qquad u_3 = x^2.$$

Step I: Orthogonalization

$$v_1 = u_1 = 1.$$

$$v_2 = u_2 - \frac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1 = x - \frac{1}{2}.$$

Compute:

$$\langle u_3, v_1\rangle = \int_0^1 x^2\,dx = \frac{1}{3}, \qquad \langle v_2, v_2\rangle = \int_0^1 (x - \tfrac{1}{2})^2\,dx = \frac{1}{12}.$$

$$\langle u_3, v_2\rangle = \int_0^1 x^2(x - \tfrac{1}{2})\,dx = \frac{1}{12}.$$

Thus

$$v_3 = u_3 - \tfrac{1}{3} - v_2 = x^2 - x + \tfrac{1}{6}.$$

Hence the orthogonal set is

$$v_1 = 1, \qquad v_2 = x - \tfrac{1}{2}, \qquad v_3 = x^2 - x + \tfrac{1}{6}.$$

Step II: Normalization

$$\|v_1\| = 1, \qquad \|v_2\| = \frac{1}{2\sqrt{3}}, \qquad \|v_3\| = \frac{1}{6\sqrt{5}}.$$

Thus

$$\begin{aligned} \varphi_1 &= v_1 = 1, \\ \varphi_2 &= \frac{v_2}{\|v_2\|} = 2\sqrt{3}\,(x - \tfrac{1}{2}) = \sqrt{3}(2x - 1), \\ \varphi_3 &= \frac{v_3}{\|v_3\|} = 6\sqrt{5}\,(x^2 - x + \tfrac{1}{6}) = \sqrt{5}(6x^2 - 6x + 1). \end{aligned}$$

Verification of orthonormality

$$\langle \varphi_1, \varphi_2\rangle = \int_0^1 \sqrt{3}(2x - 1)\,dx = 0,$$

$$\langle \varphi_1, \varphi_3\rangle = \int_0^1 \sqrt{5}(6x^2 - 6x + 1)\,dx = 0,$$

$$\langle \varphi_2, \varphi_3\rangle = \sqrt{15}\int_0^1 (12x^3 - 18x^2 + 8x - 1)\,dx = 0.$$

$$\|\varphi_1\| = \|\varphi_2\| = \|\varphi_3\| = 1.$$

Hence $\{\varphi_1, \varphi_2, \varphi_3\}$ is an orthonormal basis of $\mathcal{P}_2(\mathbb{R})$.

16. If we apply the Gram–Schmidt orthogonalization process to a set of vectors that is not known a priori to be linearly independent, what will happen?

Solution Let V be an inner-product space and let $u_1, \ldots, u_n \in V$. Apply the Gram–Schmidt orthogonalization producing vectors $v_1, \ldots, v_n$ by

$$v_1 := u_1, \qquad v_k := u_k - \sum_{j=1}^{k-1} \frac{\langle u_k, v_j\rangle}{\langle v_j, v_j\rangle} v_j \quad (2 \le k \le n),$$

when the denominators are nonzero. Then, for each $k \ge 2$,

Proof. We prove the two implications separately.

(If $v_k = 0$ then $u_k \in \operatorname{span}\{u_1, \dots, u_{k-1}\}$.) If $v_k = 0$, with k as the least such index, then by the defining relation for v_k,

$$0 = v_k = u_k - \sum_{j=1}^{k-1} \frac{\langle u_k, v_j \rangle}{\langle v_j, v_j \rangle} v_j,$$

so

$$u_k = \sum_{j=1}^{k-1} \frac{\langle u_k, v_j \rangle}{\langle v_j, v_j \rangle} v_j.$$

Each v_j lies in $\operatorname{span}\{u_1, \dots, u_j\} \subseteq \operatorname{span}\{u_1, \dots, u_{k-1}\}$, hence $u_k \in \operatorname{span}\{u_1, \dots, u_{k-1}\}$.

(If $u_k \in \operatorname{span}\{u_1, \dots, u_{k-1}\}$ then $v_k = 0$.) Assume u_k is a linear combination of $u_1, \dots, u_{k-1}$. Since Gram–Schmidt preserves spans at each step, there exist scalars $d_1, \dots, d_{k-1}$ with $u_k = \sum_{j=1}^{k-1} d_j v_j$. Using orthogonality of the v_j's,

$$\langle u_k, v_j \rangle = \Big\langle \sum_{i=1}^{k-1} d_i v_i, v_j \Big\rangle = d_j \langle v_j, v_j \rangle.$$

Thus $\dfrac{\langle u_k, v_j \rangle}{\langle v_j, v_j \rangle} = d_j$ for each j, and substituting into the formula for v_k yields

$$v_k = u_k - \sum_{j=1}^{k-1} d_j v_j = 0.$$

Combining the two implications gives the stated equivalence. The remaining assertions follow: discard zero vectors to obtain an orthogonal basis of the span, and normalize to obtain an orthonormal basis. □

17. Suppose V is finite-dimensional, $T \in L(V)$, and U is a subspace of V. Prove that

$$U \text{ and } U^\perp \text{ are both invariant under } T \iff P_U T = T P_U,$$

where P_U denotes the orthogonal projection of V onto U.

Solution Recall the following properties of the orthogonal projection P_U: P_U is linear and idempotent ($P_U^2 = P_U$), $\operatorname{range}(P_U) = U$, $\ker(P_U) = U^\perp$, and every $v \in V$ has a unique decomposition $v = P_U v + (v - P_U v)$ with $P_U v \in U$ and $v - P_U v \in U^\perp$.

$\Rightarrow$ Assume U and $U^\perp$ are T-invariant. Let $v \in V$ and write $v = P_U v + (v - P_U v)$. Since T preserves U and $U^\perp$, we have $T(P_U v) \in U$ and $T(v - P_U v) \in U^\perp$. Applying P_U to Tv yields

$$P_U T v = P_U\big(T(P_U v) + T(v - P_U v)\big) = T(P_U v) + 0 = T(P_U v).$$

Hence $P_U T v = T P_U v$ for every v, i.e. $P_U T = T P_U$.

$\Leftarrow$ Conversely assume $P_U T = T P_U$. If $u \in U$ then $P_U u = u$, and thus

$$Tu = T(P_U u) = P_U T u \in \operatorname{range}(P_U) = U,$$

so U is T-invariant. If $w \in U^\perp$ then $P_U w = 0$, and

$$P_U(Tw) = T(P_U w) = T0 = 0,$$

so $Tw \in \ker(P_U) = U^\perp$, proving $U^\perp$ is T-invariant.

18. In $\mathbb{R}^4$, let

$$U = \operatorname{span}\{(1,1,0,0),(1,1,1,2)\}.$$

Find $u \in U$ such that $\|u-(1,2,3,4)\|$ is as small as possible.

Solution

$$a_1 = (1,1,0,0), \qquad a_2 = (1,1,1,2), \qquad b = (1,2,3,4).$$

Apply Gram–Schmidt to $\{a_1, a_2\}$ to get an ONB for U.

$$\|a_1\|^2 = 2, \qquad e_1 = \frac{a_1}{\|a_1\|} = \frac{1}{\sqrt{2}}(1,1,0,0).$$

$$\langle a_2, e_1\rangle = \frac{1+1}{\sqrt{2}} = \sqrt{2}, \qquad v_2 = a_2 - \langle a_2, e_1\rangle e_1 = a_2 - a_1 = (0,0,1,2).$$

$$\|v_2\|^2 = 5, \qquad e_2 = \frac{1}{\sqrt{5}}(0,0,1,2).$$

Thus $\{e_1, e_2\}$ is an orthonormal basis of U, and the orthogonal projection of b onto U is

$$u = P_U b = \langle b, e_1\rangle e_1 + \langle b, e_2\rangle e_2.$$

$$\langle b, e_1\rangle = \frac{1+2}{\sqrt{2}} = \frac{3}{\sqrt{2}}, \qquad \langle b, e_2\rangle = \frac{3+8}{\sqrt{5}} = \frac{11}{\sqrt{5}}.$$

Therefore

$$u = \frac{3}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}}(1,1,0,0) + \frac{11}{\sqrt{5}} \cdot \frac{1}{\sqrt{5}}(0,0,1,2) = \left(\tfrac{3}{2}, \tfrac{3}{2}, \tfrac{11}{5}, \tfrac{22}{5}\right).$$

Hence the vector in U closest to b is $u = (3/2, 3/2, 11/5, 22/5)$.

19. Let $\{a_1, \ldots, a_n\}$ be an orthogonal set of nonzero vectors in an inner product space $(V, \langle\cdot,\cdot\rangle)$. For any $x \in V$ we have

$$\sum_{k=1}^{n} \frac{|\langle x, a_k\rangle|^2}{\|a_k\|^2} \le \|x\|^2,$$

called Bessel's inequality.

Solution

Consider the orthogonal projection of x onto the subspace $\operatorname{span}\{a_1, \ldots, a_n\}$:

$$p = \sum_{k=1}^{n} \frac{\langle x, a_k\rangle}{\|a_k\|^2}\, a_k.$$

Since the set is orthogonal, p is the unique vector in the span of $\{a_k\}$ closest to x. Because $x - p$ is orthogonal to every a_k, we have

$$\langle x - p, p\rangle = 0.$$

Thus,

$$\|x\|^2 = \|x-p\|^2 + \|p\|^2,$$

Next compute $\|p\|^2$. Using orthogonality again,

$$\|p\|^2 = \left\langle \sum_{k=1}^{n} \frac{\langle x, a_k\rangle}{\|a_k\|^2} a_k, \sum_{j=1}^{n} \frac{\langle x, a_j\rangle}{\|a_j\|^2} a_j \right\rangle = \sum_{k=1}^{n} \frac{|\langle x, a_k\rangle|^2}{\|a_k\|^2}.$$

Substituting this into the identity $\|x\|^2 = \|x-p\|^2 + \|p\|^2$ gives

$$\sum_{k=1}^{n} \frac{|\langle x, a_k\rangle|^2}{\|a_k\|^2} = \|p\|^2 \le \|x\|^2,$$

because $\|x-p\|^2 \ge 0$. This proves Bessel's inequality.

20. Let $V = C([-1,1])$ with the inner product

$$\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)\,dt.$$

Let $W = \{f \in V : \ f(-t) = -f(t) \text{ for all } t\}$ be the subspace of odd functions. Show that

$$W^{\perp} = \{g \in V : \ g(-t) = g(t) \text{ for all } t\},$$

the subspace of even functions.

Solution

Define also

$$E = \{g \in V : g(-t) = g(t) \text{ for all } t\},$$

the subspace of even functions. We show that $W^{\perp} = E$.

Step 1: $E \subset W^{\perp}$. If $g \in E$ is even and $f \in W$ is odd, then the product $t \mapsto f(t)g(t)$ is an odd function. Hence

$$\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)\,dt = 0,$$

since the integral of any odd function over the symmetric interval $[-1,1]$ is 0. Thus every even function is orthogonal to every odd function, and therefore $E \subset W^{\perp}$.

Step 2: $W^{\perp} \subset E$. Take any $g \in V$ and decompose it into its even and odd parts:

$$g_e(t) = \frac{g(t)+g(-t)}{2}, \qquad g_o(t) = \frac{g(t)-g(-t)}{2}.$$

Then $g = g_e + g_o$, where g_e is even and g_o is odd.

Suppose now that $g \in W^{\perp}$. Then for every odd function $f \in W$,

$$0 = \langle f, g\rangle = \langle f, g_e\rangle + \langle f, g_o\rangle.$$

Since g_e is even and f is odd, the product fg_e is odd, and therefore

$$\langle f, g_e\rangle = 0 \quad \text{for all odd } f.$$

Thus $\langle f, g_o\rangle = 0$ for every odd function f. Taking $f = g_o$ (which is itself odd) gives

$$0 = \langle g_o, g_o\rangle = \int_{-1}^{1} g_o(t)^2\,dt.$$

Since the integrand is nonnegative and continuous, this implies $g_o \equiv 0$. Hence $g = g_e$ is even, so $g \in E$. Combining the two steps, we conclude that $W^{\perp} = E$, which completes the proof.