MH5200 Advanced Investigations in Linear Algebra I

Problem Sheet 7– Questions & Solutions

Academic Year 2025/2026, Semester 1

Quantitative Research Society @NTU

November 4, 2025

Overview of This Problem Sheet

- **Main themes.** This sheet consolidates core ideas around
  - similarity and invariants (Problems 1–3),
  - structure of eigenvectors (Problem 4),
  - spectral properties of real symmetric matrices (Problems 5, 10),
  - perturbation / "nearly symmetric" behaviour (Problem 6),
  - linear recurrences and matrix powers (Fibonacci, Problem 7),
  - rotations and trigonometric identities in matrix form (Problem 8),
  - characteristic polynomials, trace and determinant (Problem 9).
- **Skills targeted.**
  - Using similarity transforms and understanding what they do / do not preserve.
  - Working with characteristic polynomials and minimal polynomials.
  - Interpreting eigenvalues/eigenvectors for symmetric and positive (semi)definite matrices.
  - Connecting algebraic recurrences (Fibonacci) and geometric operations (rotations) to matrix powers.
  - Using Rayleigh quotients and spectral decompositions.

Preliminaries

Definition I (Eigenvalue and Eigenvector). Let A be an $n \times n$ matrix. A scalar λ is an **eigenvalue** of A if there exists a nonzero vector $\boldsymbol{v}$ (called a corresponding **eigenvector**) such that

$$A\boldsymbol{v} = \lambda \boldsymbol{v}.$$

Definition II (Characteristic Polynomial). Rewriting the eigenvalue equation as $(\lambda I - A)\boldsymbol{v} = \mathbf{0}$, $\boldsymbol{v} \neq \mathbf{0}$, we see that λ is an eigenvalue of A if and only if

$$\det(\lambda I - A) = 0.$$

The **characteristic polynomial** of A is the formal polynomial

$$p_A(t) := \det(tI - A),$$

and the equation $p_A(t) = 0$ is called the **characteristic equation** of A.

Problem 1: Similar Matrices and Eigenvalues

Problem

Matrices A and B are said to be **similar** if there exists an invertible matrix C such that

$$A = CBC^{-1}.$$

Prove that similar matrices have the same eigenvalues.

Solution

Method 1 (Characteristic polynomial)

We show that A and B have the same characteristic polynomial. If $A = CBC^{-1}$ with C invertible, then

$$\begin{aligned} p_A(\lambda) &= \det(\lambda I - A) = \det(\lambda I - CBC^{-1}) \\ &= \det(\lambda CC^{-1} - CBC^{-1}) = \det\big(C(\lambda I - B)C^{-1}\big) \\ &= \det(C)\ \det(\lambda I - B)\ \det(C^{-1}) \\ &= \det(\lambda I - B) = p_B(\lambda). \end{aligned}$$

Hence $p_A = p_B$; therefore A and B have exactly the same eigenvalues (including algebraic multiplicities).

Method 2 (Mapping eigenvectors)

Assume $B\boldsymbol{v} = \lambda\boldsymbol{v}$ with $\boldsymbol{v} \neq 0$. Then

$$A(C\boldsymbol{v}) = CBC^{-1}(C\boldsymbol{v}) = CB\boldsymbol{v} = C(\lambda\boldsymbol{v}) = \lambda(C\boldsymbol{v}).$$

Since C is invertible, $C\boldsymbol{v} \neq 0$, so $C\boldsymbol{v}$ is an eigenvector of A with eigenvalue λ. Thus every eigenvalue of B is an eigenvalue of A. Reversing the roles of A and B (because similarity is symmetric) shows the converse, so A and B have exactly the same eigenvalues.

Problem 2: Similarity and Powers / Transposes

Problem

Suppose that the $n \times n$ matrix B is similar to the $n \times n$ matrix A.

(a) Show that B^k is similar to A^k for positive integer k.

(b) Show that $B^\top$ is similar to $A^\top$.

Solution

Let $B = CAC^{-1}$ with C invertible.

Method 1 (Direct algebra)

(a) Powers. We prove by induction on k that

$$B^k = CA^kC^{-1}.$$

For $k = 1$, this is precisely $B = CAC^{-1}$. Assume it holds for some $k \geq 1$:

$$B^k = CA^kC^{-1}.$$

Then

$$\begin{aligned} B^{k+1} = B^kB = (CA^kC^{-1})(CAC^{-1}) = CA^k(C^{-1}C)AC^{-1} \\ = CA^{k+1}C^{-1}. \end{aligned}$$

Thus, by induction, $B^k = CA^kC^{-1}$ for all $k \in \mathbb{N}$, so B^k is similar to A^k.

(b) Transpose. Take transpose in $B = CAC^{-1}$:

$$\begin{aligned} B^\top &= (CAC^{-1})^\top \\ &= (C^{-1})^\top A^\top C^\top \\ &= (C^\top)^{-1} A^\top C^\top. \end{aligned}$$

Hence $B^\top$ is similar to $A^\top$ via the similarity transform $C^\top$.

Method 2 (Polynomials in a matrix)

(a) Powers via polynomials. For any polynomial p, one can show

$$p(B) = C\,p(A)\,C^{-1}.$$

Indeed, for monomials $p(t) = t^k$ this follows from $B^k = CA^kC^{-1}$; linearity then extends it to arbitrary polynomials.

In particular, setting $p(t) = t^k$ gives $B^k = CA^kC^{-1}$.

(b) Transpose via similarity relation. Similarity is an equivalence relation preserved by transpose:

$$B \sim A \quad \Longrightarrow \quad B^\top \sim A^\top,$$

because any similarity $A = CBC^{-1}$ gives

$$A^\top = (CBC^{-1})^\top = (C^{-1})^\top B^\top C^\top,$$

so

$$B^\top = C^\top A^\top (C^\top)^{-1},$$

which is again a similarity relation.

Problem 3: Similarity Invariants and Counterexamples

Problem

(a) Let

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Show that B has the same rank, trace, and characteristic polynomial as A, but that B is not similar to A.

(b) Extend the result in part (a) to $n \times n$ matrices. That is, show that for an $n \times n$ matrix B to be similar to A, it is not sufficient for B to have the same rank, determinant, trace and characteristic polynomial as A.

Solution

Method 1 (Identity vs. Jordan block)

(a) The 2×2 example. First compute the invariants.

Rank. Both A and B are invertible (upper-triangular with nonzero diagonal), so

$$\operatorname{rank}(A) = \operatorname{rank}(B) = 2.$$

Trace.

$$\operatorname{tr}(A) = 1 + 1 = 2, \qquad \operatorname{tr}(B) = 1 + 1 = 2.$$

Characteristic polynomial.

$$p_A(t) = \det(tI - A) = \det \begin{bmatrix} t-1 & 0 \\ 0 & t-1 \end{bmatrix} = (t-1)^2.$$

$$p_B(t) = \det(tI - B) = \det \begin{bmatrix} t-1 & -1 \\ 0 & t-1 \end{bmatrix} = (t-1)^2.$$

So A and B share rank, trace and characteristic polynomial.

However, $A = I_2$ is the identity. If there were an invertible C such that $B = CAC^{-1}$, then

$$B = CI_2C^{-1} = CC^{-1} = I_2,$$

contradiction since $B \neq I_2$. Therefore B is not similar to A.

Equivalently: A has two linearly independent eigenvectors (indeed, every nonzero vector is an eigenvector), whereas B has only *one* independent eigenvector (it is a single Jordan block). Similar matrices must have the same geometric multiplicities for each eigenvalue, so $A \not\sim B$.

(b) The $n \times n$ version. Take

$$A = I_n, \qquad B = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Then:

– B is upper-triangular with all diagonal entries equal to 1, so $\det B = 1 = \det A$.

– $\operatorname{rank}(A) = \operatorname{rank}(B) = n$ (both invertible).

– $\operatorname{tr}(A) = n = \operatorname{tr}(B)$.

– Since $B - I_n$ is nonzero nilpotent of rank 1, the characteristic polynomial of B is still $(t-1)^n$, matching that of A.

However, A is diagonalizable with n linearly independent eigenvectors, while B has a Jordan block of size at least 2 for eigenvalue 1, so it is not diagonalizable. Thus A and B cannot be similar.

Method 2 (Minimal polynomial viewpoint)

Similarity preserves not only rank, trace and characteristic polynomial, but also the *minimal polynomial.*

(a) For the 2×2 example.

$$A = I_2 \quad \Rightarrow \quad m_A(t) = t - 1.$$

$$B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad \Rightarrow \quad m_B(t) = (t-1)^2,$$

since $B - I_2 \neq 0$ but $(B - I_2)^2 = 0$.

Since $m_A \neq m_B$, A and B cannot be similar.

(b) For the $n \times n$ version. Similarly,

$$m_{I_n}(t) = t - 1,$$

while for the matrix B with a single superdiagonal 1,

$$m_B(t) = (t-1)^2.$$

Again the minimal polynomials differ, hence $B \not\sim A$, even though they share rank, determinant, trace and characteristic polynomial. This shows these invariants alone do not guarantee similarity.

Problem 4: When is a Linear Combination an Eigenvector?

Problem

Let $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ be eigenvectors of an $n \times n$ matrix A, and let c_1 and c_2 be two nonzero scalars. Under what circumstances is

$$\boldsymbol{y} := c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2$$

an eigenvector of A?

Solution

Method 1 (Direct comparison)

Compute

$$A\boldsymbol{y} = A(c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2) = c_1A\boldsymbol{x}_1 + c_2A\boldsymbol{x}_2 = c_1\lambda_1\boldsymbol{x}_1 + c_2\lambda_2\boldsymbol{x}_2,$$

where $A\boldsymbol{x}_1 = \lambda_1\boldsymbol{x}_1$ and $A\boldsymbol{x}_2 = \lambda_2\boldsymbol{x}_2$.

For $\boldsymbol{y}$ to be an eigenvector, there must exist some λ such that

$$A\boldsymbol{y} = \lambda\boldsymbol{y} = \lambda(c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2) = c_1\lambda\boldsymbol{x}_1 + c_2\lambda\boldsymbol{x}_2.$$

Hence

$$c_1(\lambda_1 - \lambda)\boldsymbol{x}_1 + c_2(\lambda_2 - \lambda)\boldsymbol{x}_2 = \boldsymbol{0}.$$

If $\boldsymbol{x}_1, \boldsymbol{x}_2$ are linearly independent, then we must have

$$\lambda_1 = \lambda = \lambda_2.$$

Thus, in the generic independent case, $\boldsymbol{y}$ is an eigenvector *if and only if* $\lambda_1 = \lambda_2$.

Method 2 (Subspace / eigenspace viewpoint)

There are two special situations:

- **Same eigenvalue.** If $\lambda_1 = \lambda_2 = \lambda$, then $\boldsymbol{x}_1, \boldsymbol{x}_2$ lie in the same eigenspace E_λ. Since eigenspaces are linear subspaces, any nontrivial combination

  $$\boldsymbol{y} = c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 \neq \boldsymbol{0}$$

  also satisfies $A\boldsymbol{y} = \lambda\boldsymbol{y}$, so $\boldsymbol{y}$ is an eigenvector.
- **Colinear eigenvectors.** Even if $\lambda_1 \neq \lambda_2$, it is possible that $\boldsymbol{x}_2$ is a scalar multiple of $\boldsymbol{x}_1$; but in that case they actually belong to the *same* eigenspace (since eigenvectors with different eigenvalues cannot be scalar multiples). So this reduces to the previous case.

If $\lambda_1 \neq \lambda_2$ and $\boldsymbol{x}_1, \boldsymbol{x}_2$ are linearly independent, then $\boldsymbol{y}$ is not an eigenvector for any nonzero c_1, c_2.

Conclusion. A nonzero combination $c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2$ is an eigenvector of A if and only if $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ lie in the same eigenspace, equivalently $\lambda_1 = \lambda_2$.

Problem 5: Real Symmetric Matrices

Problem

Let $A \in \mathbb{R}^{n\times n}$ be a real symmetric matrix.

(a) Prove that all of the eigenvalues of A are real.

(b) Prove that eigenvectors of A that correspond to distinct eigenvalues are always perpendicular.

Solution

Method 1 (Rayleigh quotient and symmetry)

(a) Eigenvalues are real. Let $\lambda \in \mathbb{C}$ be an eigenvalue with (possibly complex) eigenvector $\boldsymbol{x} \neq 0$. Then

$$A\boldsymbol{x} = \lambda\boldsymbol{x}.$$

Multiply on the left by $\boldsymbol{x}^*$ (conjugate transpose):

$$\boldsymbol{x}^* A\boldsymbol{x} = \lambda\boldsymbol{x}^*\boldsymbol{x}.$$

Since A is real symmetric, $A = A^\top = A^*$, and $\boldsymbol{x}^* A\boldsymbol{x}$ is real:

$$\boldsymbol{x}^* A\boldsymbol{x} = (\boldsymbol{x}^* A\boldsymbol{x})^* \in \mathbb{R}.$$

Also $\boldsymbol{x}^*\boldsymbol{x} > 0$. Hence

$$\lambda = \frac{\boldsymbol{x}^* A\boldsymbol{x}}{\boldsymbol{x}^*\boldsymbol{x}} \in \mathbb{R}.$$

Therefore every eigenvalue of A is real.

(b) Orthogonality of distinct eigenvectors. Let $\lambda_1 \neq \lambda_2$ be eigenvalues with (real) eigenvectors $\boldsymbol{x}_1, \boldsymbol{x}_2$:

$$A\boldsymbol{x}_1 = \lambda_1\boldsymbol{x}_1, \qquad A\boldsymbol{x}_2 = \lambda_2\boldsymbol{x}_2.$$

Compute

$$\boldsymbol{x}_2^\top A\boldsymbol{x}_1 = \lambda_1\boldsymbol{x}_2^\top\boldsymbol{x}_1.$$

On the other hand, by symmetry,

$$\boldsymbol{x}_2^\top A\boldsymbol{x}_1 = (A\boldsymbol{x}_2)^\top\boldsymbol{x}_1 = (\lambda_2\boldsymbol{x}_2)^\top\boldsymbol{x}_1 = \lambda_2\boldsymbol{x}_2^\top\boldsymbol{x}_1.$$

Thus

$$\lambda_1\boldsymbol{x}_2^\top\boldsymbol{x}_1 = \lambda_2\boldsymbol{x}_2^\top\boldsymbol{x}_1 \quad \Rightarrow \quad (\lambda_1 - \lambda_2)\boldsymbol{x}_2^\top\boldsymbol{x}_1 = 0.$$

Since $\lambda_1 \neq \lambda_2$, it follows that $\boldsymbol{x}_2^\top\boldsymbol{x}_1 = 0$, i.e. $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are orthogonal.

Method 2 (Spectral theorem viewpoint)

The spectral theorem states that a real symmetric matrix A can be orthogonally diagonalised:

$$A = Q\Lambda Q^\top,$$

where Q is orthogonal and Λ is real diagonal.

– From Λ being real diagonal, its diagonal entries (the eigenvalues) are real.
– The columns of Q are an orthonormal basis $\{\boldsymbol{q}_1, \dots, \boldsymbol{q}_n\}$ of eigenvectors of A. Distinct eigenvalues correspond to different columns; orthonormality gives perpendicularity.

Thus parts (a) and (b) follow immediately from the spectral decomposition.

Problem 6: A Nearly Symmetric Matrix

Problem

Consider

$$A = \begin{bmatrix} 1 & \epsilon \\ 0 & 1+\epsilon \end{bmatrix}, \qquad \epsilon > 0 \text{ small.}$$

For very small ϵ, the matrix A is "nearly symmetric". Find the eigenvectors of A and the angle between them. What does this example demonstrate?

Solution

Method 1 (Direct eigen-computation)

The characteristic polynomial is

$$\det(tI - A) = \det \begin{bmatrix} t-1 & -\epsilon \\ 0 & t-1-\epsilon \end{bmatrix} = (t-1)(t-1-\epsilon).$$

Hence

$$\lambda_1 = 1, \qquad \lambda_2 = 1+\epsilon.$$

Eigenvector for λ_1. Solve $(A-I)\boldsymbol{x} = \mathbf{0}$:

$$A - I = \begin{bmatrix} 0 & \epsilon \\ 0 & \epsilon \end{bmatrix}, \qquad \begin{bmatrix} 0 & \epsilon \\ 0 & \epsilon \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0} \;\Rightarrow\; x_2 = 0.$$

So we may take

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Eigenvector for λ_2. Solve $(A-(1+\epsilon)I)\boldsymbol{x} = \mathbf{0}$:

$$A - (1+\epsilon)I = \begin{bmatrix} -\epsilon & \epsilon \\ 0 & 0 \end{bmatrix}, \qquad -\epsilon x_1 + \epsilon x_2 = 0 \;\Rightarrow\; x_1 = x_2.$$

So we may take

$$\boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

The angle θ between $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ is

$$\cos\theta = \frac{\boldsymbol{v}_1 \cdot \boldsymbol{v}_2}{\|\boldsymbol{v}_1\|\,\|\boldsymbol{v}_2\|} = \frac{1}{1\cdot\sqrt{2}} = \frac{1}{\sqrt{2}} \quad \Rightarrow \quad \theta = \frac{\pi}{4} = 45^\circ.$$

Note that this angle is *independent* of ϵ; it does not become closer to 90° as $\epsilon \to 0$.

Conclusion. Even if a matrix is very close (in entries) to a symmetric matrix, its eigenvectors need not be close to orthogonal. Orthogonality of eigenvectors is a *qualitative* property of exact symmetry, not a continuous one.

Method 2 (Symmetric + skew-symmetric decomposition)

Write

$$A = S + N, \qquad S := \frac{A + A^\top}{2}, \quad N := \frac{A - A^\top}{2}.$$

Compute

$$A^\top = \begin{bmatrix} 1 & 0 \\ \epsilon & 1+\epsilon \end{bmatrix} \quad \Rightarrow \quad S = \begin{bmatrix} 1 & \frac{\epsilon}{2} \\ \frac{\epsilon}{2} & 1+\epsilon \end{bmatrix}, \quad N = \begin{bmatrix} 0 & \frac{\epsilon}{2} \\ -\frac{\epsilon}{2} & 0 \end{bmatrix}.$$

Here S is symmetric, N is skew-symmetric and small in norm ($\|N\| = O(\epsilon)$). However, even a small skew-symmetric perturbation can destroy orthogonality of eigenvectors: S has orthogonal eigenvectors, but $A = S + N$ does *not*.

This illustrates that eigenvectors of nonsymmetric matrices can be highly sensitive to perturbations, even when the matrix is close to symmetric.

Problem 7: Fibonacci Numbers and Matrix Powers

Problem

The Fibonacci numbers satisfy

$$F_{k+2} = F_{k+1} + F_k, \qquad F_0 = 0, \quad F_1 = 1,$$

giving the sequence $0, 1, 1, 2, 3, 5, 8, 13, \ldots$.

(a) Let $\boldsymbol{u}_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}$. Show that

$$\boldsymbol{u}_{k+1} = A\boldsymbol{u}_k$$

for a suitable 2×2 matrix A, and find A.

(b) Use this to express the n-th Fibonacci number F_n in closed form, and hence determine F_{100}.

Solution

Method 1 (Matrix recurrence and diagonalisation)

(a) Constructing the recurrence. We have

$$\boldsymbol{u}_{k+1} = \begin{bmatrix} F_{k+2} \\ F_{k+1} \end{bmatrix} = \begin{bmatrix} F_{k+1} + F_k \\ F_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}.$$

Thus

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \boldsymbol{u}_{k+1} = A\boldsymbol{u}_k.$$

(b) Closed form via eigenvalues. We have

$$\boldsymbol{u}_k = A^k \boldsymbol{u}_0, \qquad \boldsymbol{u}_0 = \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

The characteristic polynomial of A is

$$\det(tI - A) = \det \begin{bmatrix} t-1 & -1 \\ -1 & t \end{bmatrix} = t^2 - t - 1.$$

Its roots are

$$\phi = \frac{1+\sqrt{5}}{2}, \qquad \psi = \frac{1-\sqrt{5}}{2}.$$

One finds eigenvectors

$$\boldsymbol{v}_1 = \begin{bmatrix} \phi \\ 1 \end{bmatrix}, \qquad \boldsymbol{v}_2 = \begin{bmatrix} \psi \\ 1 \end{bmatrix},$$

so with $P = [\boldsymbol{v}_1\ \boldsymbol{v}_2]$ and $D = \operatorname{diag}(\phi, \psi)$,

$$A = PDP^{-1}, \qquad A^n = PD^nP^{-1}.$$

From $\boldsymbol{u}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, solving $\boldsymbol{u}_0 = \alpha\boldsymbol{v}_1 + \beta\boldsymbol{v}_2$ gives

$$\alpha = \frac{1}{\sqrt{5}}, \qquad \beta = -\frac{1}{\sqrt{5}}.$$

Thus

$$\boldsymbol{u}_n = A^n\boldsymbol{u}_0 = \alpha\phi^n\boldsymbol{v}_1 + \beta\psi^n\boldsymbol{v}_2,$$

and the first component yields

$$F_n = \frac{1}{\sqrt{5}}\big(\phi^n - \psi^n\big),$$

the classical Binet formula.

For $n = 100$,

$$F_{100} = \frac{1}{\sqrt{5}}\left(\phi^{100} - \psi^{100}\right) = 354224848179261915075.$$

Method 2 (Solving the scalar recurrence directly)

Consider the scalar recurrence

$$F_{n+2} = F_{n+1} + F_n, \quad F_0 = 0,\ F_1 = 1.$$

Look for solutions of the form $F_n = r^n$, giving

$$r^{n+2} = r^{n+1} + r^n \quad \Rightarrow \quad r^2 = r + 1 \quad \Rightarrow \quad r = \phi, \psi.$$

Thus the general solution is

$$F_n = \alpha\phi^n + \beta\psi^n.$$

From $F_0 = 0$ and $F_1 = 1$:

$$0 = \alpha + \beta, \qquad 1 = \alpha\phi + \beta\psi,$$

so $\beta = -\alpha$ and

$$1 = \alpha(\phi - \psi) = \alpha\sqrt{5} \quad \Rightarrow \quad \alpha = \frac{1}{\sqrt{5}},\ \beta = -\frac{1}{\sqrt{5}}.$$

Hence Binet's formula as before, and the same value for F_{100}.

Problem 8: Powers of a Rotation Matrix

Problem

Let

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

be the matrix of a counterclockwise rotation by angle θ in $\mathbb{R}^2$. Show that

$$R^n = \begin{bmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{bmatrix}$$

for all integers $n \geq 1$.

Solution

Method 1 (Induction with angle-addition formulas)

For $n = 1$, the formula is trivially true. Assume it holds for some $n = k$:

$$R^k = \begin{bmatrix} \cos k\theta & -\sin k\theta \\ \sin k\theta & \cos k\theta \end{bmatrix}.$$

Then

$$\begin{aligned} R^{k+1} &= R^k R \\ &= \begin{bmatrix} \cos k\theta & -\sin k\theta \\ \sin k\theta & \cos k\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos k\theta \cos\theta - \sin k\theta \sin\theta & -\cos k\theta \sin\theta - \sin k\theta \cos\theta \\ \sin k\theta \cos\theta + \cos k\theta \sin\theta & -\sin k\theta \sin\theta + \cos k\theta \cos\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos((k+1)\theta) & -\sin((k+1)\theta) \\ \sin((k+1)\theta) & \cos((k+1)\theta) \end{bmatrix}, \end{aligned}$$

using the angle-addition formulas for sine and cosine. Hence the statement holds for all n by induction.

Method 2 (Complex-number representation)

Identify $\mathbb{R}^2$ with $\mathbb{C}$ via $(x, y) \leftrightarrow z = x + iy$. Then rotation by θ corresponds to multiplication by $e^{i\theta}$.

Applying the rotation n times corresponds to multiplication by $e^{in\theta}$. The associated real 2×2 matrix is

$$\begin{bmatrix} \Re e^{in\theta} & -\Im e^{in\theta} \\ \Im e^{in\theta} & \Re e^{in\theta} \end{bmatrix} = \begin{bmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{bmatrix},$$

which is exactly R^n.

Problem 9: Coefficients of the Characteristic Polynomial

Problem

Let $A \in \mathbb{R}^{n\times n}$ and let $p_A(t) = \det(tI - A)$ be its characteristic polynomial. Show that

$$p_A(t) = t^n - (\operatorname{tr} A)t^{n-1} + \cdots + (-1)^n \det A.$$

Write the trace $\operatorname{tr} A$ and determinant $\det A$ in terms of the eigenvalues of A.

Solution

Method 1 (General expansion)

Regard $tI - A$ as an $n \times n$ matrix whose diagonal entries are $t - a_{ii}$ and off-diagonal entries are $-a_{ij}$.

In expanding $\det(tI - A)$ as a polynomial in t:

– The highest-degree term t^n arises by choosing t from each of the n diagonal entries; hence its coefficient is 1.

– The coefficient of t^{n-1} arises from choosing t from $n-1$ diagonal positions and $-a_{ii}$ from the remaining diagonal position. Summing these contributions gives

$$-\sum_{i=1}^{n} a_{ii} = -\operatorname{tr} A.$$

– The constant term (coefficient of t^0) is $\det(-A) = (-1)^n \det A$.

Thus

$$p_A(t) = t^n - (\operatorname{tr} A)t^{n-1} + \cdots + (-1)^n \det A.$$

Method 2 (Eigenvalue factorisation)

Let $\lambda_1, \ldots, \lambda_n$ be the (complex) eigenvalues of A, counted with algebraic multiplicity. Then

$$p_A(t) = \det(tI - A) = \prod_{i=1}^{n}(t - \lambda_i).$$

Expanding this product:

$$p_A(t) = t^n - \Big(\sum_{i=1}^{n} \lambda_i\Big)t^{n-1} + \cdots + (-1)^n \prod_{i=1}^{n} \lambda_i.$$

Comparing this with the previous expression, we read off

$$\operatorname{tr} A = \sum_{i=1}^{n} \lambda_i, \qquad \det A = \prod_{i=1}^{n} \lambda_i.$$

Thus the trace is the sum of eigenvalues and the determinant is their product.

Problem 10: Positive Semidefinite Matrix and Rayleigh Quotients

Problem

An $n \times n$ symmetric matrix S is said to be **positive semidefinite** if all its eigenvalues are nonnegative. An equivalent condition is

$$\boldsymbol{x}^\top S\boldsymbol{x} \geq 0 \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^n.$$

Suppose S is symmetric positive semidefinite with eigenvalues

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0.$$

(a) What are the eigenvalues of $\lambda_1 I - S$?

(b) Show that $\lambda_1 \boldsymbol{x}^\top \boldsymbol{x} \geq \boldsymbol{x}^\top S\boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^n$.

(c) Determine

$$\max_{\boldsymbol{x}\in\mathbb{R}^n\setminus\{\boldsymbol{0}\}} \frac{\boldsymbol{x}^\top S\boldsymbol{x}}{\boldsymbol{x}^\top \boldsymbol{x}}.$$

Solution

Method 1 (Direct eigen-analysis)

(a) Eigenvalues of $\lambda_1 I - S$. If $S\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i$, then

$$(\lambda_1 I - S)\boldsymbol{v}_i = \lambda_1 \boldsymbol{v}_i - S\boldsymbol{v}_i = (\lambda_1 - \lambda_i)\boldsymbol{v}_i.$$

Hence the eigenvalues of $\lambda_1 I - S$ are

$$\mu_i = \lambda_1 - \lambda_i, \qquad i = 1, \ldots, n.$$

Since $\lambda_1 \geq \lambda_i$, each $\mu_i \geq 0$, so $\lambda_1 I - S$ is also positive semidefinite.

(b) The inequality $\lambda_1 \boldsymbol{x}^\top \boldsymbol{x} \geq \boldsymbol{x}^\top S\boldsymbol{x}$. Because $\lambda_1 I - S$ is positive semidefinite, we have

$$\boldsymbol{x}^\top(\lambda_1 I - S)\boldsymbol{x} \geq 0 \quad \text{for all } \boldsymbol{x}.$$

Expanding:

$$\lambda_1 \boldsymbol{x}^\top \boldsymbol{x} - \boldsymbol{x}^\top S\boldsymbol{x} \geq 0 \quad \Rightarrow \quad \boldsymbol{x}^\top S\boldsymbol{x} \leq \lambda_1 \boldsymbol{x}^\top \boldsymbol{x}.$$

(c) Maximising the Rayleigh quotient. The Rayleigh quotient is

$$R(\boldsymbol{x}) = \frac{\boldsymbol{x}^\top S\boldsymbol{x}}{\boldsymbol{x}^\top \boldsymbol{x}}, \quad \boldsymbol{x} \neq \mathbf{0}.$$

From (b), $R(\boldsymbol{x}) \leq \lambda_1$. On the other hand, if we choose $\boldsymbol{x} = \boldsymbol{v}_1$, an eigenvector corresponding to λ_1, then

$$R(\boldsymbol{v}_1) = \frac{\boldsymbol{v}_1^\top S\boldsymbol{v}_1}{\boldsymbol{v}_1^\top \boldsymbol{v}_1} = \frac{\boldsymbol{v}_1^\top(\lambda_1\boldsymbol{v}_1)}{\boldsymbol{v}_1^\top \boldsymbol{v}_1} = \lambda_1.$$

Thus the maximum value is

$$\max_{\boldsymbol{x}\neq 0} R(\boldsymbol{x}) = \lambda_1.$$

Method 2 (Spectral decomposition)

Write the spectral decomposition

$$S = Q\Lambda Q^\top,$$

where Q is orthogonal and

$$\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n).$$

Let $\boldsymbol{y} = Q^\top \boldsymbol{x}$ (a change of orthonormal basis), so $\|\boldsymbol{y}\| = \|\boldsymbol{x}\|$. Then

$$\boldsymbol{x}^\top S\boldsymbol{x} = \boldsymbol{x}^\top Q\Lambda Q^\top \boldsymbol{x} = \boldsymbol{y}^\top \Lambda \boldsymbol{y} = \sum_{i=1}^{n} \lambda_i y_i^2.$$

Thus

$$\frac{\boldsymbol{x}^\top S\boldsymbol{x}}{\boldsymbol{x}^\top \boldsymbol{x}} = \frac{\sum_{i=1}^n \lambda_i y_i^2}{\sum_{i=1}^n y_i^2} \leq \frac{\lambda_1 \sum_{i=1}^n y_i^2}{\sum_{i=1}^n y_i^2} = \lambda_1,$$

and equality is obtained by taking $\boldsymbol{y} = e_1$, i.e. $\boldsymbol{x} = \boldsymbol{v}_1$, an eigenvector corresponding to λ_1. This again shows that the maximum Rayleigh quotient equals λ_1.