MH5200 Advanced Investigations in Linear Algebra I

Problem Sheet 2– Problems & Solutions

Academic Year 2025/2026, Semester 1

Quantitative Research Society @NTU

September 14, 2025

Overview of This Problem Sheet

Where pedagogically helpful, we:

- expand arguments into explicit step-by-step derivations;
- add proof-structure hints (e.g. “consider the contrapositive”, “reduce to eigenbasis”, “apply spectral theorem”, “use Schur complement”);
- highlight advanced techniques frequently used in MH5200 (block matrix methods, invariant subspaces, polynomial functional calculus, spectral decompositions, singular value arguments, etc.).

Structure of the sheet.

- **Problem 1:** Magic squares; parametrisation of 3×3 magic matrices; all-ones vector as eigenvector with magic sum eigenvalue.
- **Problem 2:** Linear dependence of tridiagonal Toeplitz-type row vectors; determinant and spectrum of μI plus adjacency matrix.
- **Problem 3:** Norms in orthonormal coordinates; expressing $\|x\|$ via the coefficient vector β.
- **Problem 4:** Structure of skew-symmetric matrices; orthogonality $Ax \perp x$; characterisation via vanishing quadratic form.
- **Problem 5:** Matrix representation of the differentiation operator in the monomial basis; nilpotent upper-shift structure.
- **Problem 6:** Vandermonde matrices as evaluation operators; column independence via polynomial root bounds and Vandermonde determinant.

Problem 1

Problem

An $n \times n$ magic matrix is a matrix M_n with positive integer entries from 1 to n^2 whose rows, columns, and diagonal sum to the same number.

(a) Invent a 3×3 magic matrix whose first row is 8, 3, 4. Can you write down a general expression for a 3×3 magic matrix M_3? What is M_3 times $(1, 1, 1)^T$?

(b) What is M_4 times $(1, 1, 1, 1)^T$? How about M_n times $(1, \dots, 1)^T \in \mathbb{R}^n$?

Solution

Method 1: Direct Magic Sum Computation

(a) A standard 3×3 magic matrix with first row $(8, 3, 4)$ is

$$M_3 = \begin{bmatrix} 8 & 3 & 4 \\ 1 & 5 & 9 \\ 6 & 7 & 2 \end{bmatrix}.$$

Each row, column, and main diagonal sums to

$$8 + 3 + 4 = 1 + 5 + 9 = 6 + 7 + 2 = 15.$$

A general expression for a 3×3 magic matrix with centre entry 5 can be written as

$$M_3 = \begin{bmatrix} 5 + a & 5 - a + b & 5 - b \\ 5 - a - b & 5 & 5 + a + b \\ 5 + b & 5 + a - b & 5 - a \end{bmatrix},$$

for admissible integers a, b which ensure the entries are a permutation of $\{1, \dots, 9\}$. Multiplying by the all-ones vector,

$$M_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \text{row}_1 \text{ sum} \\ \text{row}_2 \text{ sum} \\ \text{row}_3 \text{ sum} \end{bmatrix} = \begin{bmatrix} 15 \\ 15 \\ 15 \end{bmatrix}.$$

(b) For any $n \times n$ magic matrix M_n, every row has the same sum S_n, called the magic constant. The total sum of all entries is

$$1 + 2 + \cdots + n^2 = \frac{n^2(n^2 + 1)}{2}.$$

Since there are n rows,

$$nS_n = \frac{n^2(n^2 + 1)}{2} \quad \Rightarrow \quad S_n = \frac{n(n^2 + 1)}{2}.$$

Thus

$$M_n \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \frac{1}{2} n(n^2+1) \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.$$

In particular, for $n = 4$,

$$S_4 = \frac{4(4^2+1)}{2} = \frac{4 \cdot 17}{2} = 34,$$

so

$$M_4 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 34 \\ 34 \\ 34 \\ 34 \end{bmatrix}.$$

Method 2: Eigenvector Viewpoint

Let $\mathbf{1}_n = (1, \ldots, 1)^T \in \mathbb{R}^n$. For a magic matrix M_n, each row sums to S_n, so

$$(M_n \mathbf{1}_n)_i = \sum_{j=1}^{n} (M_n)_{ij} = S_n \quad \text{for all } i.$$

Hence

$$M_n \mathbf{1}_n = S_n \mathbf{1}_n,$$

so $\mathbf{1}_n$ is an eigenvector with eigenvalue S_n. The trace of M_n is the sum of the diagonal entries, which is exactly S_n (the diagonal also sums to the magic constant), but we can also read off S_n as

$$S_n = \frac{1}{n} \mathbf{1}_n^T M_n \mathbf{1}_n = \frac{1}{n} \cdot (\text{sum of all entries}).$$

Using the total sum $\frac{n^2(n^2+1)}{2}$ again yields

$$S_n = \frac{1}{n} \cdot \frac{n^2(n^2+1)}{2} = \frac{n(n^2+1)}{2},$$

and hence the same formula

$$M_n \mathbf{1}_n = \frac{1}{2} n(n^2+1)\, \mathbf{1}_n.$$

Problem 2

Problem

Find the values of the scalar μ for which the row vectors $(\mu, 1, 0)$, $(1, \mu, 1)$, and $(0, 1, \mu)$ are linearly dependent. For what values of μ are they linearly independent? Explain your reasoning.

Solution

Method 1: Determinant of the Row Matrix

Place the three row vectors into a 3×3 matrix:

$$A(\mu) = \begin{bmatrix} \mu & 1 & 0 \\ 1 & \mu & 1 \\ 0 & 1 & \mu \end{bmatrix}.$$

The rows are linearly dependent if and only if $\det A(\mu) = 0$.

Compute the determinant, for instance by expanding along the first row:

$$\begin{aligned} \det A(\mu) &= \mu \begin{vmatrix} \mu & 1 \\ 1 & \mu \end{vmatrix} - 1 \begin{vmatrix} 1 & 1 \\ 0 & \mu \end{vmatrix} + 0 \cdot (\dots) \\ &= \mu(\mu^2 - 1) - 1 \cdot (\mu - 0) \\ &= \mu(\mu^2 - 1) - \mu \\ &= \mu(\mu^2 - 1 - 1) \\ &= \mu(\mu^2 - 2). \end{aligned}$$

Thus

$$\det A(\mu) = 0 \iff \mu(\mu^2 - 2) = 0 \iff \mu \in \{0, \sqrt{2}, -\sqrt{2}\}.$$

Therefore the row vectors are linearly dependent if $\mu = 0$ or $\mu = \pm\sqrt{2}$, and linearly independent for all other values of μ.

Method 2: Spectral Interpretation

Write

$$A(\mu) = \mu I + B, \qquad B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

The matrix B is symmetric and represents a path graph on three nodes. Its eigenvalues are well-known (or can be computed directly) to be

$$\lambda_1 = \sqrt{2}, \quad \lambda_2 = 0, \quad \lambda_3 = -\sqrt{2}.$$

Thus the eigenvalues of $A(\mu) = \mu I + B$ are

$$\mu + \lambda_1,\ \mu + \lambda_2,\ \mu + \lambda_3 = \mu + \sqrt{2},\ \mu,\ \mu - \sqrt{2}.$$

The matrix $A(\mu)$ is singular exactly when one of these eigenvalues is zero, i.e.

$$\mu + \sqrt{2} = 0 \quad \text{or} \quad \mu = 0 \quad \text{or} \quad \mu - \sqrt{2} = 0$$

which again gives $\mu \in \{0, \pm\sqrt{2}\}$.

For all other μ, $A(\mu)$ is invertible, so its rows are linearly independent.

Problem 3

Problem

Suppose $a_1, \dots, a_k$ are orthonormal vectors in $\mathbb{R}^n$ and $x = \beta_1 a_1 + \dots + \beta_k a_k$ where $\beta_1, \dots, \beta_k$ are scalars. Express $\|x\|$ in terms of $\beta = (\beta_1, \dots, \beta_k)$.

Solution

Method 1: Direct Expansion Using Orthonormality

Compute the squared norm:

$$\begin{aligned}\|x\|^2 &= x^T x \\ &= (\beta_1 a_1 + \dots + \beta_k a_k)^T (\beta_1 a_1 + \dots + \beta_k a_k) \\ &= \sum_{i=1}^k \sum_{j=1}^k \beta_i \beta_j \, a_i^T a_j.\end{aligned}$$

Since the vectors are orthonormal,

$$a_i^T a_j = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

so all cross terms vanish and we obtain

$$\|x\|^2 = \beta_1^2 + \cdots + \beta_k^2.$$

Thus

$$\|x\| = \|\beta\|_2 = \sqrt{\beta_1^2 + \cdots + \beta_k^2}.$$

Method 2: Matrix Formulation

Form the $n \times k$ matrix

$$A = [\, a_1 \ \cdots \ a_k \,],$$

whose columns are orthonormal. Then $A^T A = I_k$. Writing $\beta = (\beta_1, \dots, \beta_k)^T$, we can express

$$x = \sum_{i=1}^k \beta_i a_i = A\beta.$$

Therefore

$$\|x\|^2 = \|A\beta\|^2 = (A\beta)^T (A\beta) = \beta^T A^T A \beta = \beta^T I_k \beta = \beta^T \beta,$$

so again $\|x\| = \|\beta\|_2$.

Problem 4

Problem

An $n \times n$ matrix A is said to be skew-symmetric if $A^T = -A$.

(a) Find all 2×2 skew-symmetric matrices.

(b) Explain why the diagonal entries of a skew-symmetric matrix must be zero.

(c) Show that for a skew-symmetric matrix A, and any vector $x \in \mathbb{R}^n$, Ax is orthogonal to x: $(Ax) \perp x$.

(d) Now suppose that A is any matrix for which $(Ax) \perp x$ for any $x \in \mathbb{R}^n$. Show that A must be skew-symmetric.

Solution

Method 1: Coordinate Computations

(a) Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

The condition $A^T = -A$ gives

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix} = - \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix}.$$

Equating entries,

$$a = -a, \quad d = -d, \quad c = -b, \quad b = -c.$$

Thus $a = d = 0$ and $c = -b$. Let $\omega = c$; then $b = -\omega$, and

$$A = \omega \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \qquad \omega \in \mathbb{R}.$$

(b) For a general $n \times n$ skew-symmetric matrix A, we have $A_{ii} = (A^T)_{ii} = -A_{ii}$. Hence $2A_{ii} = 0$, so $A_{ii} = 0$ for all i.

(c) For skew-symmetric A, consider the scalar

$$x^T(Ax) = (Ax) \cdot x.$$

We have

$$x^T Ax = (x^T Ax)^T = x^T A^T x = x^T(-A)x = -x^T Ax.$$

Thus $x^T Ax = -x^T Ax$, so $2x^T Ax = 0$ and $x^T Ax = 0$ for all x. Therefore $(Ax) \cdot x = 0$, which means $Ax \perp x$.

Equivalently, working in indices,

$$x^T Ax = \sum_{i,j} A_{ij}x_ix_j = \sum_{i<j} A_{ij}x_ix_j + \sum_{i>j} A_{ij}x_ix_j = \sum_{i<j} A_{ij}x_ix_j + \sum_{i<j} A_{ji}x_jx_i$$

and since $A_{ji} = -A_{ij}$, each pair of terms cancels, giving zero.

(d) Assume now that A is an arbitrary matrix satisfying $(Ax) \perp x$ for all $x \in \mathbb{R}^n$. Then

$$x^T(Ax) = 0 \quad \text{for all } x.$$

In particular, take $x = e_i$ (the i-th standard basis vector). Then

$$0 = e_i^T Ae_i = A_{ii}$$

for all i, so the diagonal entries are zero.

Next, take $x = e_i + e_j$ with $i \neq j$. Then

$$\begin{aligned} 0 &= (e_i + e_j)^T A(e_i + e_j) \\ &= e_i^T Ae_i + e_i^T Ae_j + e_j^T Ae_i + e_j^T Ae_j \\ &= 0 + A_{ij} + A_{ji} + 0, \end{aligned}$$

since $e_i^T Ae_j = A_{ij}$ and we already know $A_{ii} = A_{jj} = 0$. Thus

$$A_{ij} + A_{ji} = 0 \quad \Rightarrow \quad A_{ji} = -A_{ij}$$

for all $i \neq j$. Together with $A_{ii} = 0$, this is exactly the condition $A^T = -A$. Hence A must be skew-symmetric.

Method 2: Bilinear Form and Polarisation

Define the bilinear form $B(x, y) = x^T Ay$. Then

$$x^T Ax = B(x, x).$$

If A is skew-symmetric, then

$$B(x, y) = x^T Ay = y^T A^T x = -y^T Ax = -B(y, x),$$

so $B(x, x) = -B(x, x)$, which forces $B(x, x) = 0$ and hence $x^T Ax = 0$, i.e. $Ax \perp x$.

Conversely, suppose $x^T Ax = 0$ for all x. Using the polarisation identity for bilinear forms,

$$B(x, y) + B(y, x) = \frac{1}{2}\left(B(x + y, x + y) - B(x, x) - B(y, y)\right),$$

and substituting $B(z, z) = z^T Az = 0$ for all z gives

$$B(x, y) + B(y, x) = 0.$$

Thus $x^T Ay = -y^T Ax = -x^T A^T y$ for all x, y, so $A = -A^T$, i.e. A is skew-symmetric.

Problem 5

Problem

Let p be a polynomial of degree $n-1$ or less, given by $p(t) = c_1 + c_2 t + \ldots + c_n t^{n-1}$. Its derivative with respect to t is a polynomial of degree $n-2$ or less, given by $p'(t) = d_1 + d_2 t + \ldots + d_{n-1} t^{n-2}$. Find a matrix D for which $d = Dc$.

Solution

Method 1: Matching Coefficients

We have

$$p(t) = \sum_{k=0}^{n-1} c_{k+1} t^k, \quad \Rightarrow \quad p'(t) = \sum_{k=1}^{n-1} k\, c_{k+1} t^{k-1}.$$

Thus

$$d_1 = 1 \cdot c_2, \quad d_2 = 2 \cdot c_3, \quad \ldots, \quad d_{n-1} = (n-1) \cdot c_n.$$

In vector form,

$$d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_{n-1} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & (n-1) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} =: Dc.$$

Equivalently,

$$D = \begin{bmatrix} \mathbf{0}_{(n-1)\times 1} & \operatorname{diag}(1, 2, \ldots, n-1) \end{bmatrix}.$$

Method 2: Operator View via Basis Transformation

Let $\{1, t, \ldots, t^{n-1}\}$ be the standard basis of the space of polynomials of degree at most $n-1$. The derivative operator $\frac{d}{dt}$ acts on this basis by

$$\frac{d}{dt}(t^k) = k t^{k-1}, \quad k = 1, \ldots, n-1, \qquad \frac{d}{dt}(1) = 0.$$

Therefore, relative to the ordered basis $(1, t, \ldots, t^{n-1})$, the matrix of the derivative operator has entries

$$D_{ij} = \begin{cases} j-1, & i = j-1,\ j \geq 2, \\ 0, & \text{otherwise}, \end{cases}$$

which gives exactly the same upper-shift diagonal matrix as above:

$$D = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & n-1 \end{bmatrix}.$$

Thus $d = Dc$ represents $p' = \frac{d}{dt}p$ in coefficient form.

Problem 6

Problem

A Vandermonde matrix is an $m \times n$ matrix of the form

$$V = \begin{bmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^{n-1} \\ 1 & t_2 & t_2^2 & \cdots & t_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_m & t_m^2 & \cdots & t_m^{n-1} \end{bmatrix}$$

where $t_1, \dots, t_m$ are real numbers.

(a) Provide an interpretation of the result of multiplying a vector $c = (c_1, \dots, c_n)^T \in \mathbb{R}^n$ by the Vandermonde matrix V.

(b) Assume that $m \geq n$. Show that the columns of a Vandermonde matrix are linearly independent if the numbers $t_1, \dots, t_m$ are distinct.

Solution

Method 1: Polynomial Evaluation Argument

(a) Let

$$c = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \quad \text{and} \quad p(t) = c_1 + c_2 t + \cdots + c_n t^{n-1}.$$

Then the i-th entry of Vc is

$$(Vc)_i = c_1 + c_2 t_i + \cdots + c_n t_i^{n-1} = p(t_i).$$

Thus Vc is the vector of polynomial evaluations

$$Vc = \begin{bmatrix} p(t_1) \\ p(t_2) \\ \vdots \\ p(t_m) \end{bmatrix}.$$

(b) Suppose $Vc = 0$. Then

$$p(t_i) = 0 \quad \text{for } i = 1, \dots, m.$$

If the t_i are distinct, this says that the polynomial p of degree at most $n-1$ has at least m distinct roots. Since $m \geq n$, we have at least n distinct roots. A non-zero polynomial of degree at most $n-1$ can have at most $n-1$ distinct roots, so the only possibility is that $p \equiv 0$, i.e. all coefficients $c_1, \dots, c_n$ vanish. Hence $Vc = 0$ implies $c = 0$, so the columns of V are linearly independent.

Method 2: Vandermonde Determinant (Square Case) and Reduction

Assume first $m = n$, so V is a square $n \times n$ Vandermonde matrix. Its determinant is explicitly

$$\det V = \prod_{1 \le i < j \le n} (t_j - t_i).$$

If the t_i are distinct, each factor $t_j - t_i \neq 0$, so $\det V \neq 0$, and hence V is invertible. Therefore its columns are linearly independent.

Now if $m > n$, consider the $n \times n$ submatrix V' formed by any n distinct rows, corresponding to n distinct points among $\{t_1, \ldots, t_m\}$. By the same determinant formula, $\det V' \neq 0$, so V' has linearly independent columns. If the original columns of V were linearly dependent, then restricting to these n rows would also give a linear dependence among the columns of V', contradicting invertibility. Hence the columns of V are linearly independent for all $m \ge n$ when the t_i are distinct.