1 Spring 2002 – Linear Algebra

Problem 1.1. Let $\varphi : M_3(\mathbf{Q}) \to M_3(\mathbf{Q})$ be the map sending m to $\varphi(m) = m^2 + 3m + 3$. Show that $\varphi(m) \neq 0$ for all $m \in M_3(\mathbf{Q})$.

Solution: Suppose $\varphi(m) = 0$ for some $m \in M_3(\mathbf{Q})$. Let $f = x^2 + 3x + 3$. Then $m_{\mathbf{Q}}(m) \mid f$, where $m_{\mathbf{Q}}(m)$ is the minimal polynomial of m. Since f is irreducible by Eisenstein's Criterion, it must be the case that $m_{\mathbf{Q}}(m) = f$. Now let g be the characteristic polynomial of m, then $f \mid g$. Since f is irreducible over $\mathbf{Q}$ this implies that $g = f^k$. But $\deg(g) = 3 \neq 2k = \deg(f^k)$ for any (positive) integer k. This is a contradiction so we conclude $\varphi(k) \neq 0, \forall m \in M_3(\mathbf{Q})$. □

- Minimal Polynomial: Let T be a linear operator on a finite dimensional vector space V over a field F. The minimal polynomial p of T is the monic generator of the ideal of polynomials which annihilate T, i.e.
  - p is a monic polynomial over F.
  - $p(T) = 0$.
  - Any g such that $g(T) = 0$ is a multiple of p.

  Similarly define the minimal polynomial of a matrix A. Notice that
  - Similar matrices have the same minimal polynomial (they represent the same linear operator in different bases).
  - The minimal polynomial is invariant under field extensions, i.e. if $F \subset F_1$, then the minimal polynomial of A regarded as an element of $\mathbb{M}_n(F)$ is the same as that of A regarded as an element of $\mathbb{M}_n(F_1)$.

- Characteristic Polynomial: Let V be a vector space over some field F and let $A \in \mathbb{M}_n(F)$. The polynomial
$$f \equiv \det(xI - A)$$
is called the characteristic polynomial of A. The roots of f are the characteristic values (or eigenvalues) of A in F. c is a characteristic value if and only if there exists some $0 \neq \alpha \in V$ such that
$$A\alpha = c\alpha.$$
  - Similar matrices have the same characteristic polynomial, so the characteric polynomial of an operator T is well–defined (take the matrix representation in any basis).
  - The characteristic polynomial is a monic polynomial of degree n.

- Cayley–Hamilton: The minimal polynomial divides the characteristic polynomial. Moreover, the roots of the two polynomials are the same.
- Factorization of Polynomials: If F is an algebraically closed field, then all polynomials factor into linear terms. In particular, if T is a linear operator of a vector space over F, then F contains all the eigenvalues of T.
- Invariance Under Field Extensions: The following things are invariant under field extensions.
  - The minimal polynomial of a matrix A.
  - The quotient and remainder from the Division Algorithm.
  - The (monic) greatest common divisor of two polynomials (since they can be obtained from the Euclidean Algorithm).
- Eisenstein's Criterion: Let p be a prime in $\mathbb{Z}$ and let
  $$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{Z}[x], n \geq 1.$$
  Suppose
  $$p \mid a_i, 0 \leq i < n \text{ but } p^2 \nmid a_0.$$
  Then $f(x)$ is irreducible in both $\mathbb{Z}[x]$ and $\mathbb{Q}[x]$.

Problem 1.2. Let A be a real matrix with column vectors $A_1, A_2, \ldots, A_n$. If the A_j are mutually orthogonal, then
$$|\det\ A| = \prod_{j=1}^{n} |A_j|.$$
This follows because $|\det({}^tA \cdot A)| = |\det\ A|^2$ and ${}^tA \cdot A$ is a diagonal matrix with diagonal entries $|A_1|^2, |A_2|^2, \ldots, |A_n|^2$. Prove that a general matrix satisfies the inequality
$$|\det\ A| \leq \prod_{j=1}^{n} |A_j|.$$
Hint: apply the Gram–Schmidt orthogonalization process to the columns.

Solution: We apply the Gram–Schmidt orthogonalization process to the columns (without normalizing) to obtain $\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_N$ such that $\tilde{A}_i$'s are mutually orthogonal and have the same span as the A_i's. We have
$$\tilde{A}_1 = A_1, \quad \tilde{A}_k = A_k - \sum_{1 \leq j < k} \frac{\langle A_k, A_j \rangle}{\langle A_j, A_j \rangle} A_j, 1 < k \leq n,$$

where $\langle v, w\rangle$ denotes the inner product of v and w. If $(v_1, v_2, \ldots, v_n) = v \in \mathbb{R}^n$, then

$$\|v\| = \sqrt{v_1 + v_2 + \cdots + v_n} = \sqrt{\langle v, v\rangle}.$$

Let's compute $\|\tilde{A}_k\|$:

$$\begin{aligned}
\|\tilde{A}_k\|^2 = \langle \tilde{A}_k, \tilde{A}_k\rangle &= \|A_k\|^2 - 2\sum_{1\le j<k} \frac{|\langle A_k, A_j\rangle|^2}{\|A_j\|^2} + \sum_{1\le j<k} \frac{|\langle A_k, A_j\rangle|^2}{\|A_j\|^2} \\
&= \|A_k\|^2 - \sum_{1\le j<k} \frac{|\langle A_k, A_j\rangle|^2}{\|A_j\|^2} \\
&\le \|A_k\|^2,
\end{aligned}$$

since the sum in the penultimate line has all positive terms. The $\tilde{A}_i$'s are orthogonal, so if $\tilde{A}$ is the matrix with columns $\tilde{A}_i$, then

$$\det(\tilde{A}) = \prod_{j=1}^{n} |\tilde{A}_j|.$$

Next observe that to transform A into $\tilde{A}$ involves multiplying A by elementary matrices with determinant 1, hence since the determinant function is multiplicative, we have

$$|\det(\tilde{A}) = |\det(AE_1E_2\ldots E_n)| = |\det(A)\det(E_1)\ldots\det(E_n)| = \det(A),$$

where $E_1 \ldots E_n$ are the elementary matrices we multiply by to transform A into $\tilde{A}$ (e.g. to turn A_2 into $\tilde{A}_2$, we would multiply A on the right by the matrix which is the identity matrix with $\frac{\langle A_2, A_1\rangle}{\langle A_1, A_1\rangle}$ times the second column subtracted from the first column; the resulting matrix still has determinant 1). Finally, by the inequality in norms we derived ($|\tilde{A}_k| \le |A_k|$) we get

$$|\det(\tilde{A})| = \prod_{j=1}^{n} |\tilde{A}_j| \le \prod_{j=1}^{n} |A_j|.$$

□

- Determinant: Let K be a commutative ring with identity. Suppose

  $$D : \mathbb{M}_n(K) \longrightarrow K.$$

  Then D is called a determinant function if

  - D is n–linear (in rows).

- D is alternating (i.e. $D(A) = 0$ if two rows are the same and $D(A) = -D(A')$ if A' is obtained from A by interchanging two rows.
- $D(I) = 1$, where I is the identity matrix.

In addition, the following are true.

- If $A, B \in \mathbb{M}_n(K)$, then

$$\det(AB) = (\det A)(\det = B).$$

- If A^t is the transpose of A (i.e. rows of A^t are columns of A and vice versa), then

$$\det(A^t) = \det(A),$$

  so in particular the determinant is also linear in the columns.
- If B is obtained from A by adding a multiple of one row (column) of A to another, then

$$\det(B) = \det(A).$$

- Suppose we have a matrix in block form, then

$$\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = (\det\ A)(\det\ B).$$

- The following is of theoretical importance:

$$\det(A) \equiv D(A) = \sum_{\sigma} (\text{sgn}\ \sigma) A(1, \sigma 1) \ldots A(n, \sigma n),$$

  where $A(i, j)$ is the ij^{th} entry of A.
- The determinant is usually calculated by (cofactor expansion):

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij}\ \det[A(i|j)],$$

  where j is the index of some fixed row or column and $A(i|j)$ is the $(n-1) \times (n-1)$ matrix obtained from A by deleting the i^{th} row and j^{th} column.

• Inner Product: Let $F = \mathbb{R}$ or $\mathbb{C}$. Let V be a vector space over F. An inner product on V is a function

$$\langle , \rangle : V \times V \longrightarrow F : (u, v) \mapsto \langle u, v \rangle$$

such that for all $u, v, w \in V$ and $c \in F$, we have

- $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$
- $\langle cu, v\rangle = c\langle u, v\rangle$
- $\langle u, v\rangle = \overline{\langle v, u\rangle}$, where bar denotes complex conjugation
- $\langle u, u\rangle > 0$ if $u \neq 0$.

Given an inner product, a norm can be defined as

$$\|v\| = \sqrt{\langle v, v\rangle}.$$

The norm satisfies the following properties:

- $\|cv\| = |c|\|v\|$
- $\|v\| > 0$ for $v \neq 0$
- $|\langle u, v\rangle| \leq \|u\|\|v\|$ (or $|\langle u, v\rangle|^2 \leq \langle u, u\rangle\langle v, v\rangle$)
- $\|u + v\| \leq \|u\| + \|v\|$

The third item is called the Cauchy–Schwarz Inequality and the fourth item is called the Triangle Inequality.

- Gram–Schmidt: Let V be an inner product space and $v_1, \dots, v_n$ be any independent vectors in V. Then the set of vectors $\{\tilde{v_1}, \dots, \tilde{v_n}\}$ given by

$$\tilde{v_1} = v_1, \quad \tilde{v_k} = v_k - \sum_{j=1}^{k-1} \left\langle v_k, \frac{\tilde{v_j}}{\|\tilde{v_j}\|} \right\rangle \frac{\tilde{v_j}}{\|\tilde{v_j}\|} = v_k - \sum_{j=1}^{k-1} \frac{\langle v_k, \tilde{v_j}\rangle}{\langle \tilde{v_j}, \tilde{v_j}\rangle} \tilde{v_j}$$

form an orthogonal set with the same span as $v_1, \dots, v_n$.

Problem 1.3. Let $T \in M_3(\mathbf{C})$ and let $\mathcal{A}_T$ be the centralizer of T in $M_3(\mathbf{C})$. Show that $\dim(\mathcal{A}_T) \geq 3$ and describe (up to similarity) the linear transformations T such that $\dim(\mathcal{A}_T) = 3$.

- Linear Operators $\rightsquigarrow$ $F[x]$–Module: Let V be a finite–dimensional vector space over some field F. Let $T : V \to V$ be a linear transformation. If $v \in V$, define

$$x \cdot v = T(v)$$

So if $f(x) \in F[x]$,

$$f(x) \cdot v = [f(T)](v).$$

This gives V a $F[x]$–module structure.

- $F[x]/(f(x))$: If
$$f(x) = x^k + b_{k-1}x^{k-1} + \cdots + b_1 x + b_0 \in F[x],$$
then
$$\{1, \overline{x}, \overline{x}^2, \ldots, \overline{x}^{k-1}\}$$
is a basis for $F[x]/(f(x))$ viewed as an F–vector space. Moreover, in this basis, T (multiplication by x) acts like:
$$\begin{aligned} 1 &\mapsto \overline{x} \\ \overline{x} &\mapsto \overline{x}^2 \\ &\vdots \\ \overline{x}^{k-1} &\mapsto \overline{x}^k = -b_0 - b_1\overline{x} - \cdots - b_{k-1}\overline{x}^{k-1}. \end{aligned}$$
The corresponding matrix, a $k \times k$ matrix called the companion matrix of $f(x)$, and denoted $\mathcal{C}_{f(x)}$, looks like
$$\mathcal{C}_{f(x)} \equiv \begin{pmatrix} 0 & 0 & \ldots & \ldots & -b_0 \\ 1 & 0 & \ldots & \ldots & -b_1 \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 & -b_{k-1} \end{pmatrix}.$$

- Rational Canonical Form: V is finite dimensional over F, but $F[x]$ is infinite dimensional over F, hence V must be a torsion $F[x]$–module. By the Fundamental Theorem of Finitely Generated Modules over a PID, we must then have
$$V \cong F[x]/(a_1(x)) \oplus \cdots \oplus F[x]/(a_m(x))$$
as $F[x]$–modules and such that
$$a_1(x) \mid a_2(x) \mid \cdots \mid a_m(x).$$
By the previous item (and since we have a direct sum), we then see that there is a basis for T with corresponding matrix
$$\begin{pmatrix} \mathcal{C}_{a_1(x)} & & & \\ & \mathcal{C}_{a_2(x)} & & \\ & & \ddots & \\ & & & \mathcal{C}_{a_m(x)} \end{pmatrix}.$$
This is the Rational Canonical Form of T.

- Observations and Consequences: By the uniqueness statement of the fundamental theorem, given a linear transformation T, the rational canonical form of T is unique. Given a matrix A, we can define $Tv = Av$, hence we immediately obtain:
  - Every matrix is similar to a matrix in rational canonical form.
  - 2 matrices are similar if and only if they have the same rational canonical form.
  - The rational canonical form is invariant under field extensions.
  - Similarity of matrices is invariant under field extensions.

  Let A be an $n \times n$ matrix over F. By considering the rational canonical form of A, we also learn something about the minimal and characteristic polynomials of A.
  - From the decomposition $F[x] \cong F[x]/(a_1(x)) \oplus \cdots \oplus F[x]/(a_m(x))$, we see that
  $$m_A(x) = a_m(x).$$
  - A quick calculation shows the characteristic polynomial of a $\mathcal{C}_{a(x)}$ is exactly $a(x)$. Hence we have (using the fact that similar matrices have the same characteristic polynomial)
  $$c_A(x) = a_1(x)a_2(x)\ldots a_m(x),$$
  where $c_A(x)$ is the characteristic polynomial of A.
  - From the previous item the Cayley–Hamilton Theorem is immediate.
  - More precisely,
  $$c_A(x) \mid (m_A(x))^k,$$
  for some positive integer k. In particular, $c_A(x)$ and $m_A(x)$ have the same roots.

- Computational Aspects: Let A be a matrix. To find the rational canonical form, apply elementary row and column operations to $xI - A$ to put it into Smith Normal Form:
$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & a_1(x) & & & \\ & & & & a_2(x) & & \\ & & & & & \ddots & \\ & & & & & & a_m(x) \end{pmatrix}.$$
The $a_i(x)'s$ are then the invariant factors.

- The Algebraically Closed Case: Now suppose the field F is algebraically closed. Then the invariant factors $a_1(x), \dots, a_m(x)$ factor completely into linear terms (equivalently, F contains all eigenvalues of T). In $F[x]/(x-\lambda)^k$, the elements

$$\{\overline{1}, \overline{x}-\lambda, (\overline{x}-\lambda)^2, \dots, (\overline{x}-\lambda)^{k-1}\}$$

form a basis. In this basis, T acts like (write $x = \lambda + (x-\lambda)$):

$$\begin{aligned} 1 &\mapsto \lambda \cdot 1 + (\overline{x}-\lambda) \\ (\overline{x}-\lambda) &\mapsto \lambda(\overline{x}-\lambda) + (\overline{x}-\lambda)^2 \\ &\vdots \\ (\overline{x}-\lambda)^{k-1} &\mapsto \lambda(\overline{x}-\lambda)^{k-1} + (\overline{x}-\lambda)^k = \lambda(\overline{x}-\lambda)^{k-1}. \end{aligned}$$

The corresponding matrix, a $k \times k$ matrix called a Jordan Block, looks like

$$\begin{pmatrix} \lambda & & & & \\ 1 & \lambda & & & \\ & 1 & \ddots & & \\ & & \ddots & \ddots & \\ & & & 1 & \lambda \end{pmatrix}.$$

- Jordan Canonical Form: By the Fundamental Theorem of Finitely Generated Modules over PID II, we then have that

$$V \cong F[x]/(x-\lambda_1)^{k_1} \oplus \cdots \oplus F[x]/(x-\lambda_t)^{k_t},$$

where λ_i's are the eigenvalues of T. By the previous item, we then see that there is a basis for T with corresponding matrix

$$\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_t \end{pmatrix},$$

where each J_i is a $k_i \times k_i$ Jordan block. This is the Jordan Canonical Form of T.

- Observations and Consequences: Up to permutation of the Jordan blocks, the Jordan Canonical Form is unique.
  - Every matrix is similar to a matrix in Jordan Canonical Form.

- 2 matrices over a field F are similar if and only if they have the same Jordan Canonical Form over the algebraic closure of F.
- If a matrix A is similar to a diagonal matrix D, then D is the Jordan Canonical Form of A.
- The Jordan Canonica Form is NOT invariant under field extensions.

By considering the Jordan Canonical Form, we also have a criterion for diagonalizability: A matrix A is diagonalizable if and only if $m_A(x)$ has no repeated roots.

- A quick calculation shows that the minimal polynomial of a diagonal matrix has as roots exactly the distinct elements along the diagonal (no repeats).
- Conversely, the minimal polynomial of a Jordan block of size k with eigenvalue λ has minimal polynomial $(x-\lambda)^k$ (think $F[x]/(x-\lambda)^k$). The minimal polynomial of a Jordan Canonical Form is the least common multiple of the minimal polynomials of the Jordan blocks (use Smith Normal Form). The result follows.