MA110
Lecture 20

Saurav Bhaumik
Department of Mathematics
IIT Bombay

Spring 2025

Cayley-Hamilton theorem

Theorem. Let $\mathbf{A} \in \mathbb{K}^{n\times n}$ and let $p(t)$ be its characteristic polynomial. Then $p(\mathbf{A}) = \mathbf{0}$.

Proof. Let $\mathbf{B}$ be the adjugate of the matrix $(t\mathbf{I}_n - \mathbf{A})$. Then we have the following formula for determinant, where $p(t)$ is the characteristic polynomial:

$$(t\mathbf{I}_n - \mathbf{A})\mathbf{B} = \det(t\mathbf{I}_n - \mathbf{A})\mathbf{I}_n = p(t)\mathbf{I}_n.$$

Now the (j,k)-th entry of the adjugate of $(t\mathbf{I}_n - \mathbf{A})$ is obtained by taking determinant of an $(n-1)\times(n-1)$-submatrix. Thus, the (j,k)-th entry of $\mathbf{B}$ is a polynomial in t of degree $\leq n-1$. Thus we can say, $\mathbf{B} = \mathbf{B}_0 + t\mathbf{B}_1 + \dots t^{n-1}\mathbf{B}_{n-1}$ for matrices $\mathbf{B}_j \in \mathbb{K}^{n\times n}$.

$$
\begin{aligned}
p(t)\mathbf{I}_n &= (t\mathbf{I}_n - \mathbf{A})\mathbf{B} && (1)\\
&= (tI_n - \mathbf{A})\sum_{i=0}^{n-1} t^i\mathbf{B}_i && (2)\\
&= \sum_{i=0}^{n-1} t\mathbf{I}_n \cdot t^i\mathbf{B}_i - \sum_{i=0}^{n-1}\mathbf{A}\cdot t^i\mathbf{B}_i && (3)\\
&= \sum_{i=0}^{n-1} t^{i+1}\mathbf{B}_i - \sum_{i=0}^{n-1} t^i\mathbf{A}\mathbf{B}_i && (4)\\
&= t^n\mathbf{B}_{n-1} + \sum_{i=1}^{n-1} t^i(\mathbf{B}_{i-1} - \mathbf{A}\mathbf{B}_i) - \mathbf{A}\mathbf{B}_0. && (5)
\end{aligned}
$$

It is clear that $p(A) = 0$. $\square$

Recall: We have defined the notion of (abstract) vector space over $\mathbb{K}$ and the notion of subspace of a vector space. We looked at some of the examples of vector spaces such as

- Subspaces of $\mathbb{K}^{n\times 1}$
- The space $\mathbb{K}^{m\times n}$ of all $m \times n$ matrices with entries in $\mathbb{K}$
- The spaces $\mathbb{K}[X]$ and $\mathcal{P}_n$ of polynomials in one variable
- The spaces $C[a, b]$ and $C^1[a, b]$ of functions $[a, b] \to \mathbb{K}$
- The space c of convergent sequences of real numbers

We have also seen that several notions and results discussed over $\mathbb{R}^n$ have a straightforward analogue in the context of a general vector space V over $\mathbb{K}$. These are as follows.

- Linear combination
- span
- Linear dependence
- Linear independence

Proposition (Crucial Result)

Let S be a subset of s elements and R be a set of r elements of V. If $S \subset \operatorname{span} R$ and $s > r$, then S is linearly dependent.

Examples

$\boxed{1}$. Let $m, n \in \mathbb{N}$, and let V be the vector space $\mathbb{K}^{m\times n}$. of all $m\times n$ matrices with entries in $\mathbb{K}$. For $j = 1, \ldots, m$ and $k = 1, \ldots, n$, let $\mathbf{E}_{jk}$ denote the $m\times n$ matrix whose (j, k)th entry is equal to 1 and all other entries are equal to zero. Then the set $S := \left\{\mathbf{E}_{jk} : 1 \leq j \leq m, 1 \leq k \leq, n\right\}$ is linearly independent. Moreover S spans V.

$\boxed{2}$. Let $V := c_0$ be the subpace of c consisting of all sequences in $\mathbb{R}$ which converge to 0. For $j \in \mathbb{N}$, let e_j denote the element of S whose jth term is equal to 1 and all other terms are equal to 0. Then the set $S := \left\{e_j : j \in \mathbb{N}\right\}$ is linearly independent. However, S does not span V. To see this, let $e = (1/n)$ be the sequence whose nth term is $1/n$ for $n \in \mathbb{N}$. Then $e \in c_0$, but e is not a (finite) linear combination of elements of S. Thus $S_1 := S \cup \{e\}$ is also linearly independent

.

$\boxed{3}$. Let $V := \mathbb{K}[x]$ be the vector space of all polynomials in the indeterminate x with coefficients in $\mathbb{K}$. Then the set $S := \{x^j : j = 0, 1, 2, \ldots\}$ is linearly independent. Moreover S spans V. For a fixed $n \in \mathbb{N}$, the set $S_n := \{x^j : 0 \le j \le n\}$ is linearly independent and it spans the subspace $\mathcal{P}_n$ of $\mathbb{K}[X]$.

$\boxed{4}$. Let $V := C[-\pi, \pi]$. For $n \in \mathbb{N}$, define $u_n, v_n \in V$ by

$$u_n(t) := \cos nt \quad \text{and} \quad v_n(t) := \sin nt \quad \text{for } t \in [-\pi, \pi].$$

Then the set $S := \{u_1, u_2, \ldots\} \cup \{v_1, v_2, \ldots\}$ is linearly independent. (Note that the zero element of this vector space is the function having all its values on $[-\pi, \pi]$ equal to 0.)

But S doesn't span V. To see this, consider $w(t) := t$ for $t \in [-\pi, \pi]$. Then the set $S_1 := S \cup \{w\}$ is also linearly independent, since $w(\pi) \neq w(-\pi)$, and so $w \notin \operatorname{span} S$.

Definition

A vector space V is said to be **finite dimensional** *if there is a finite subset S of V such that $V = \operatorname{span} S$; otherwise the vector space V is said to be* **infinite dimensional**.

If a vector space V is infinite dimensional, then V is larger than the span of any finite subset of V, and so V must contain an infinite linearly independent subset. Conversely, if V contains an infinite linearly independent subset, then V must be infinite dimensional.

Examples: Let $n, m \in \mathbb{N}$. The vector spaces $\mathbb{K}^{n\times 1}$, $\mathbb{K}^{1\times n}$ and $\mathbb{K}^{m\times n}$ are finite dimensional, and so is the vector space $\mathcal{P}_n$ of all polynomials in the indeterminate x having degree less than or equal to n. But the vector spaces $\mathbb{K}[x]$, $C[-\pi, \pi]$, c, and c_0 are infinite dimensional.

Definition

Any linearly independent subset of a finite dimensional vector space V which spans V is called a **basis** *for V.*

Here is the most important result about finite dimensional vector spaces. The proof is similar to that in the case of subspaces of $\mathbb{K}^n$.

Proposition

Let V be a finite dimensional vector space over $\mathbb{K}$. Then:

- *V has a basis.*
- *Every set that spans V has a subset which is a basis of V.*
- *Every linearly independent subset of V can be extended to a basis of V.*
- *Any two bases of V have the same cardinality, called the* **dimension** *of V and denoted by* $\dim V$.

Linear Transformations

Definition

Let V and W be vector spaces over $\mathbb{K}$. A **linear transformation** *or a* **linear map** *from V to W is a function $T : V \to W$ which 'preserves' the operations of addition and scalar multiplication, that is, for all $u, v \in V$ and $\alpha \in \mathbb{K}$,*

$$T(u+v) = T(u) + T(v) \quad \textit{and} \quad T(\alpha\, v) = \alpha\, T(v).$$

It is clear that if $T : V \to W$ is linear, then $T(0) = 0$. Also, T 'preserves' linear combinations of elements of V:

$$T(\alpha_1 v_1 + \cdots + \alpha_k v_k) = \alpha_1 T(v_1) + \cdots + \alpha_k T(v_k)$$

for all $k \in \mathbb{N}$, $v_1, \ldots, v_k \in V$ and $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$.

Remark: A linear transformation from a vector space V to itself is often called a linear operator on V.

Examples

1. Let $\mathbf{A}$ be an $m \times n$ matrix with entries in $\mathbb{K}$. Then the map $T : \mathbb{K}^{n\times 1} \to \mathbb{K}^{m\times 1}$ defined by $T(\mathbf{x}) := \mathbf{A}\,\mathbf{x}$ is linear. Similarly, the map $T' : \mathbb{K}^{1\times m} \to \mathbb{K}^{1\times n}$ defined by $T'(\mathbf{y}) := \mathbf{yA}$ is linear. More generally, the map

$$T : \mathbb{K}^{n\times p} \to \mathbb{K}^{m\times p} \quad \text{defined by} \quad T(\mathbf{X}) := \mathbf{A}\,\mathbf{X}$$

is linear, and the map

$$T' : \mathbb{K}^{p\times m} \to \mathbb{K}^{p\times n} \quad \text{defined by} \quad T'(\mathbf{Y}) := \mathbf{YA}$$

is linear.

2. $T : \mathbb{K}^{m\times n} \to \mathbb{K}^{n\times m}$ defined by $T(\mathbf{A}) := \mathbf{A}^{\mathsf{T}}$ is linear.

3. The map $T : \mathbb{K}^{n\times n} \to \mathbb{K}$ defined by $T(\mathbf{A}) := \operatorname{trace}\mathbf{A}$ is linear. But $\mathbf{A} \longmapsto \det\mathbf{A}$ does not define a linear map.

4. The map $T : \mathbb{K}[X] \to \mathbb{K}$ defined by $T(p(X)) = p(0)$ is linear.

5. Let $V := c_0$, the set of all sequences in $\mathbb{K}$ which converge to 0. Then the map $T : V \to V$ defined by

$$T(x_1, x_2, \ldots) := (0, x_1, x_2, \ldots)$$

is linear, and so is the map $T' : V \to V$ defined by

$$T'(x_1, x_2, \ldots) := (x_2, x_3, \ldots).$$

Note that $T' \circ T$ is the identity map on V, but $T \circ T'$ is not the identity map on V. The map T is called the **right shift operator** and T' is called the **left shift operator** on V.

6. Let $V := C^1([a, b])$, the set of all real-valued continuously differentiable functions, and let $W := C([a, b])$, the set of all real-valued continuous functions on $[a, b]$. Then the map $T' : V \to W$ defined by $T'(f) = f'$ is linear. Also, the map

$$T : W \to V \text{ defined by } \ T(f)(x) := \int_a^x f(t)dt \ \text{ for } x \in [a, b],$$

is linear. [Question. What are $T' \circ T$ and $T' \circ T$?]

Let V and W be vector spaces over $\mathbb{K}$, and let $T : V \to W$ be a linear map. Two important subspaces associated with T are

(i) $\mathcal{N}(T) := \{v \in V : T(v) = 0\}$, the **null space** of T, which is a subspace of V,

(ii) $\mathcal{I}(T) := \{T(v) : v \in V\}$, the **image space** of T, which is a subspace of W.

Suppose V is finite dimensional, and let $\dim V = n$. Since $\mathcal{N}(T)$ is a subspace of V, it is finite dimensional and $\dim \mathcal{N}(T) \leq n$

Let $v_1, \ldots, v_n$ be a basis for V. If $v \in V$, then there are $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ such that $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$, so that $T(v) = \alpha_1 T(v_1) + \cdots + \alpha_n T(v_n)$. This shows that $\mathcal{I}(T) = \text{span}\{T(v_1), \ldots, T(v_n)\}$. Hence $\mathcal{I}(T)$ is also finite dimensional and $\dim \mathcal{I}(T) \leq n$.

Definition

The dimension of $\mathcal{N}(T)$ is called the **nullity** *of the linear map T, and the dimension of $\mathcal{I}(T)$ is called the* **rank** *of T.*

The Rank-Nullity Theorem for a matrix $\mathbf{A}$ that we proved earlier is a special case of the following result.

Proposition (Rank-Nullity Theorem for Linear Maps)

Let V and W be vector spaces over $\mathbb{K}$, and let $T : V \to W$ be a linear map. Suppose $\dim V = n \in \mathbb{N}$. Then

$$\mathrm{rank}(T) + \mathrm{nullity}(T) = n.$$

Proof (Sketch): Let $s := \mathrm{nullity}(T)$ and let $\{u_1, \ldots, u_s\}$ be a basis of $\mathcal{N}(T)$. Extend the linearly independent set $\{u_1, \ldots, u_s\}$ to a basis $\{u_1, \ldots, u_s, u_{s+1}, \ldots, u_n\}$ of V. Check that the set $\{T(u_{s+1}), \ldots, T(u_n)\}$ is a basis of $\mathcal{I}(T)$. $\square$

Corollary

Let V, W be finite dimensional vector spaces with $\dim V = n$ and $\dim W = m$. Also, let $T : V \to W$ be a linear map. Then

$$T \textit{ is one-one} \iff \operatorname{rank}(T) = n.$$

In particular, if T is one-one, then $n \leq m$. Further,

$$\textit{if } m = n, \textit{ then } \quad T \textit{ is one-one} \iff T \textit{ is onto.}$$

Proof. The first assertion follows from the Rank-Nullity Theorem since

$$T \text{ is one-one} \iff \mathcal{N}(T) = \{0\} \iff \operatorname{nullity}(T) = 0.$$

If T is one-one, then $n = \operatorname{rank}(T) = \dim \mathcal{I}(T) \leq \dim W = m$. Further, if $m = n$, then $\operatorname{rank}(T) = n \iff T$ is onto. $\square$