Cayley-Hamilton Theorem

Jitendra Rathore

November 19, 2025

Throughout this note, we denote the base field by $\mathbb{F}$, which is $\mathbb{R}$ or $\mathbb{C}$. All the vector spaces are considered over $\mathbb{F}$.

We have learnt in the last lecture about generalized eigenspace. We also saw that how generalized eigenspaces helps us to decomposed a vector space into smaller subspaces. We begin by recalling this statement. In this lecture we prove Cayley-Hamilton Theorem, which says that characteristic polynomial of T annihilates T.

Theorem 0.1. *Let $T : V \to V$ be a linear operator on a finite–dimensional complex vector space V. For each eigenvalue λ of T, let*

$$G_\lambda = \ker(T - \lambda I)^n, \qquad n = \dim V,$$

denote the generalized eigenspace corresponding to λ. Then:

1. *The generalized eigenspace G_λ is a T–invariant subspace of V.*
2. *For each eigenvalue λ, the restriction operator*

   $$T_{G_\lambda} : G_\lambda \to G_\lambda, \qquad T_{G_\lambda}(v) = Tv,$$

   satisfies that $(T_{G_\lambda} - \lambda I)$ is nilpotent on G_λ.
3. *The vector space V is the direct sum of the generalized eigenspaces:*

   $$V = \bigoplus_{\lambda \in \sigma(T)} G_\lambda,$$

   where $\sigma(T)$ denotes the set of eigenvalues of T.

Remark 0.2. If $\lambda_1, ..., \lambda_m$ be the distinct m eigenvalues of T, then we have

$$\dim(E_{\lambda_1}) + \dim(E_{\lambda_2}) + ... + \dim(E_{\lambda_m}) \le \dim(V).$$

In general, this inequality might be strict. But the next result says that this becomes equality if we replace eigenspace by generalized eigenspaces.

Corollary 0.3. *Let $T : V \to V$ be a linear operator on a n-dimensional complex vector space V. Suppose $\lambda_1, ..., \lambda_m$ be m-distinct eigenvalues of T. Then*

$$\dim(G_{\lambda_1}) + \dim(G_{\lambda_2}) + ... + \dim(G_{\lambda_m}) = \dim(V).$$

Proof. This follows from Theorem 0.1 (3). $\square$

Note that an operator may not possess a basis consisting of eignevectors (exactly non-diagonalizable operators!). The next result shows that every operator on finite dimensional complex vector space has a basis consisting of generelized eigenspaces.

Corollary 0.4. *Let $T : V \to V$ be a linear operator on a finite–dimensional complex vector space V. Then V has a basis consisting of generalized eigenvectors of T.*

Proof. Use the part (3) of the Theorem 0.1. Choose a basis of G_λ for each λ, combining them yields a basis for V. This completes the proof. □

Example 0.5. Consider the linear operator $T : \mathbb{C}^2 \to \mathbb{C}^2$ represented by

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

The characteristic polynomial is

$$\det(T - \lambda I) = \det \begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix} = (1-\lambda)^2.$$

Thus the only eigenvalue is $\lambda = 1$. To find the eigenspace,

$$(T - I)v = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ 0 \end{pmatrix} = 0,$$

which gives $y = 0$. Hence $E_1 = \ker(T - I) = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$. To compute the generalized eigenspace. $G_1 = \ker(T - I)^2$. We compute

$$T - I = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad (T - I)^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Thus $(T - I)^2 = 0$, so $G_1 = \ker(T - I)^2 = \mathbb{C}^2$, and $V = \bigoplus_{\lambda \in \sigma(T)} G_\lambda$. In this example,

$$\sigma(T) = \{1\}, \qquad G_1 = \mathbb{C}^2.$$

.

1 Multiplicities of an Eigenvalue

Let $T : V \to V$ be a linear operator on a finite–dimensional complex vector space V, and let λ be an eigenvalue of T.

Definition 1.1. The **generalized multiplicity** (or simply the *multiplicity*) of the eigenvalue λ is defined as the $\dim G_\lambda$, where $G_\lambda = \ker(T - \lambda I)^{\dim V}$ is the generalized eigenspace corresponding to λ.

Definition 1.2. The **algebraic multiplicity** of λ, denoted $m_{\mathrm{alg}}(\lambda)$, is the multiplicity of λ as a root of the characteristic polynomial of T.

The algebraic multiplicity counts how many times λ appears as a root of the characteristic polynomial.

Definition 1.3. The **geometric multiplicity** of λ is defined as the $\dim\ker(T-\lambda I)$, the dimension of the eigenspace of T corresponding to λ. We denote it by $m_{\text{geo}}(\lambda)$

The geometric multiplicity counts the number of linearly independent eigenvectors for λ.

Remark 1.4. For every eigenvalue λ of T, the following inequalities hold:

$$1 \leq m_{\text{geo}}(\lambda) \leq \dim G_\lambda \tag{1}$$

$$1 \leq m_{\text{geo}}(\lambda) \leq m_{\text{alg}}(\lambda) \tag{2}$$

Here, the first inequality holds since every eigenvector is also generalized eigenvector. The second inequality is proved in the class during lectures when we were studying the diagonalizability.

Remark 1.5. Note that, $m_{\text{geo}}(\lambda) = m_{\text{alg}}(\lambda)$ if and only if T is diagonalizable at λ (Did you remember this theorem?).

Example 1.6. $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Since the characteristic polynomial is $(\lambda-1)^2$, so only eigenvalue is $\lambda = 1$ and $m_{\text{alg}}(1) = 2$. Also, the eigenspace $\ker(T-I) = \ker\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \operatorname{span}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}$. Hence $m_{\text{geo}}(1) = 1$. Now since $(T-I)^2 = 0$. We have

$$G_1 = \ker(T-I)^2 = \mathbb{C}^2,$$

so, $\dim G_1 = 2$, and hence multilplicity of λ is 2.

Definition 1.7. A matrix is called **block diagonal** if it has the form

$$\begin{pmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_k \end{pmatrix},$$

where each B_i is a square matrix (called a *block*). All off-diagonal blocks are zero matrices.

Example 1.8. Consider a block diagonal matrix of 6×6, where we have two blocks B_1 and B_2, which are 4×4 and 2 respectively..

$$\begin{pmatrix} \boxed{\begin{matrix} 2 & 1 & 0 & 3 \\ 0 & 2 & 4 & 0 \\ 5 & 0 & 1 & 1 \\ 0 & 0 & 2 & 1 \end{matrix}} & \begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix} & \boxed{\begin{matrix} 7 & -1 \\ 2 & 5 \end{matrix}} \end{pmatrix}$$

Proposition 1.9. *Let $T : V \to V$ be a linear operator on a finite–dimensional complex vector space V with dinstinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$ having multiplicities $d_1, d_2, \ldots, d_m$ respectively. Then there exists a basis of V with respect to which T has a* block diagonal *matrix of the form*

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_m \end{pmatrix},$$

where each block A_j is a $d_j \times d_j$ upper triangular matrix of the form

$$A_j = \begin{pmatrix} \lambda_j & * & \cdots & * \\ 0 & \lambda_j & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_j \end{pmatrix},$$

.

Proof. Let λ_j be an eigenvalue of T with generalized eigenspace

$$G_{\lambda_j} = \ker(T - \lambda_j I)^{d_j},$$

where d_j is the dimension of G_{λ_j}. Note that the operator $N_j := T - \lambda_j I$ is nilpotent on G_{λ_j}, *i.e.*, $N_j^{d_j} = 0$ As this is nilpotent operators, then with respect to a suitable basis of G_{λ_j}, N_j has a strictly upper-triangular matrix (see the lecture on nilpotent operator for this fact). Since $T|_{G_{\lambda_j}} = N_j + \lambda_j I$, the matrix of T on G_{λ_j} in this basis has the form

$$A_j = \begin{pmatrix} \lambda_j & * & \cdots & * \\ 0 & \lambda_j & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_j \end{pmatrix},$$

where the $*$ entries come from the strictly upper-triangular form of N_j. The size of A_j is $d_j \times d_j$. We now use the fact that $V = \bigoplus_{j=1}^{m} G_{\lambda_j}$, taking the union of the bases of each G_{λ_j} gives a basis of V. In this basis, the operator T has the desired block diagonal form

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_m \end{pmatrix},$$

where each block A_j corresponds to the generalized eigenspace G_{λ_j}. This completes the proof. □

Remark 1.10. Over a complex vector space of dimension n, the characteristic polynomial of a linear operator T can be factored completely as

$$p(t) = (t - \lambda_1)^{d_1}(t - \lambda_2)^{d_2} \cdots (t - \lambda_m)^{d_m},$$

where $\lambda_1, \lambda_2, \ldots, \lambda_m$ are the eigenvalues of T and $d_1, d_2, \ldots, d_m$ are their respective algebraic multiplicities. Note that $d_1 + d_2 + ... + d_m = n$.

Theorem 1.11 (Cayley-Hamilton Theorem). *Let $T : V \to V$ be a linear operator on a finite-dimensional complex vector space V, and let $p(t)$ denote the characteristic polynomial of T. Then*

$$p(T) = 0,$$

where $p(T)$ is obtained by substituting T for t in $p(t)$.

Proof. We factor

$$p(t) = \det(tI - T) = (t - \lambda_1)^{d_1}(t - \lambda_2)^{d_2} \cdots (t - \lambda_m)^{d_m}$$

where $\lambda_1, \ldots, \lambda_m$ are the distinct eigenvalues of T and $d_1, \ldots, d_m$ are their algebraic multiplicities. Let $G_{\lambda_j} = \ker(T - \lambda_j I)^{d_j}$ denote the generalized eigenspace corresponding to λ_j. Since $\dim G_{\lambda_j} = d_j$, and the operator $N_j := T - \lambda_j I$ is nilpotent on G_{λ_j}, we have $N_j^{d_j} = (T - \lambda_j I)^{d_j} = 0$ on G_{λ_j}.

Now consider

$$p(T) = (T - \lambda_1 I)^{d_1}(T - \lambda_2 I)^{d_2} \cdots (T - \lambda_m I)^{d_m}.$$

For each j, the factor $(T - \lambda_j I)^{d_j}$ acts as zero on G_{λ_j} because $N_j^{d_j} = 0$. Since

$$V = \bigoplus_{j=1}^{m} G_{\lambda_j},$$

any vector $v \in V$ can be written as $v = v_1 + v_2 + \cdots + v_m$ with $v_j \in G_{\lambda_j}$. Applying $p(T)$ to v gives

$$p(T)v = p(T)(v_1 + \cdots + v_m) = 0 + \cdots + 0 = 0.$$

Hence $p(T) = 0$, which proves the Cayley-Hamilton theorem. □

Definition 1.12. Let $T : V \to V$ be a linear operator on a finite-dimensional vector space V. A monic polynomial $m_T(t)$ of least degree such that

$$m_T(T) = 0$$

is called the **minimal polynomial** of T. Equivalently, $m_T(t)$ is the unique monic polynomial of smallest degree such that $m_T(t) = 0$. Recall a ploynomial is said to be monic if coefficient of higher degree term is 1.

Example 1.13. Consider the polynomials

$$p_1(t) = t^3 - 2t + 1, \qquad p_2(t) = 3t^2 + t - 5.$$

- $p_1(t)$ is **monic** as the coefficient of t^3 is 1.
- $p_2(t)$ is **not monic** because the leading coefficient of t^2 is $3 \neq 1$.