Linear Algebra: A complete Guide

Linear Algebra

August 12, 2021

Elijah Thompson,
Physics and Math Honors

Solo Pursuit of Learning

Contents

Part I

Vector Space Theory

Chapter 1

Vector Spaces

1.A.0 Constructions

Chapter 2

Linear Maps

Chapter 3

Matrix Operations

Chapter 4

Determinants

Chapter 5

Spectral Theory

Chapter 6

Canonical Forms

6.1.0 Triangular Form

In this section we aim to construct the triangular form of linear maps, and demonstrate some properties this canonical form has. Throughout this section we presume V to be a finite-dimensional vector space over a field $\mathbb{F}$.

One important result we aim to show is that the matrix of a linear operator $T : V \to V$ with respect to a basis $\beta = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$, $[T]_\beta^\beta$, is in upper triangular form if and only if $T(W_i) \subseteq W_i$ for all $W_i = \text{span}(\mathbf{x}_1, ..., \mathbf{x}_i)$.

Definition 6.1.1 (Invariant Subspaces). *Let $T : V \to V$ be a linear mapping. A subspace $W \subseteq V$ is said to be* invariant *under T if $T(W) \subseteq W$.*

Example 6.1.1. a $\{\mathbf{0}\}$ and V are invariant under all linear mappings $T : V \to V$

b $\ker(T)$ and $\text{Im}(T)$ are invariant subspaces for all T as well

c If λ is an eigenvector of T then the eigenspace E_λ is invariant under T

Proposition 6.1.1. *Let V be a vector space, let $T : V \to V$ be a linear mapping, and let $\beta = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$ be a basis for V. Then $[T]_\beta^\beta$ is upper-triangular if and only if each of the subspaces $W_i = \text{span}(\{\mathbf{x}_1, ..., \mathbf{x}_i\})$ is invariant under T. Note that*

$$\{\mathbf{0}\} \subset W_1 \subset W_1 \subset ... \subset W_{n-1} \subset W_n = V$$

Definition 6.1.2. *We say that a linear mapping $T : V \to V$ on a finite-dimensional vector space V is triangularizable if there exists a basis β such that $[T]_\beta^\beta$ is upper triangular.*

Remark 6.1.1. To show that a mapping is triangularizable we must produce the increasing sequence of invariant subspaces

$$\{\mathbf{0}\} \subset W_1 \subset W_1 \subset ... \subset W_{n-1} \subset W_n = V$$

spanned by the subsets of the basis β and show that each W_i is invariant under T.

Proposition 6.1.2. *Let $T : V \to V$ and let $W \subset V$ be an invariant subspace. Then the characteristic polynomial of the restriction $T|_W$ divides the characteristic polynomial of T.*

Proof. Let $\alpha = \{\mathbf{x}_1, ..., \mathbf{x}_k\}$ be a basis for W, and extend α to a basis $\beta = \{\mathbf{x}_1, ..., \mathbf{x}_k, ..., \mathbf{x}_n\}$ of V. Since W is invariant under T, for each $i \leqslant k$ we have

$$T(\mathbf{x}_i) = a_{1i}\mathbf{x}_1 + ... + a_{ki}\mathbf{x}_k + 0\mathbf{x}_{k+1} + ... + 0\mathbf{x}_n$$

Hence, the matrix $[T]_\beta^\beta$ has a block form composition

$$[T]_\beta^\beta = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

where $A = [T|_W]_\alpha^\alpha$ is a $k \times k$ matrix. When we compute the characteristic polynomial of T using this matrix, we have $\det(T - \lambda I) = \det(A - \lambda I)\det(C - \lambda I)$ where $\det(A - \lambda I)$ is the characteristic polynomial of $T|_W$, as desired. ∎

Theorem 6.1.3. *Let V be a finite-dimensional vector space over a field $\mathbb{F}$, and let $T : V \to V$ be a linear mapping. Then T is triangularizable if and only if the characteristic polynomial equation of T has $dim(V)$ roots (counted with multiplicities) in the field F.*

Lemma 6.1.4. *Let $T : V \to V$ be as in the theorem, and assume that the characteristic polynomial of T has $n = dim(V)$ roots in $\mathbb{F}$. If $W \subsetneq V$ is an invariant subspace under T, then there exists a vector $\mathbf{x} \neq \mathbf{0}$ in V such that $\mathbf{x} \notin W$ and $W + \text{span}(\{\mathbf{x}\})$ is also invariant under T.*

Proof. Let $\alpha = \{\mathbf{x}_1, ..., \mathbf{x}_k\}$ be a basis for W and adjoin $\alpha' = \{\mathbf{x}_{k+1}, ..., \mathbf{x}_n\}$ to form a basis $\beta = \alpha \cup \alpha'$ for V. Let $W' = \text{span}(\alpha')$. Now, consider the linear mapping $P : V \to V$ defined by

$$P(a_1\mathbf{x}_1 + ... + a_n\mathbf{x}_n) = a_1\mathbf{x}_1 + ... + a_k\mathbf{x}_k$$

Note P is the projection on W with kernel W', while $I - P$ is the projection on W' with kernel W. Let $S = (I - P)T$. Since $\text{Im}(I - P) = W'$, we see that $\text{Im}(S) \subset \text{Im}(I - P) = W'$. Thus, W' is an invariant subspace of S.

We claim the set of eigenvalues of $S|_{W'}$ is a subset of the set of eigenvalues of T. Since W is an invariant subspace of T, the matrix $[T]_\beta^\beta$ is in block upper triangular form:

$$[T]_\beta^\beta = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

From our constructions and definitions we find that $A = [T|_W]_\alpha^\alpha$ and $C = [S|_{W'}]_{\alpha'}^{\alpha'}$. Hence, $\det(T - \lambda I) = \det(T|_W - \lambda I)\det(S|_{W'} - \lambda I)$, so the characteristic polynomial of $S|_{W'}$ divides the characteristic polynomial of T.

Since all eigenvalues of $S|_{W'}$ will lie in $\mathbb{F}$ by assumption on T, there exists an eigenvector $\mathbf{x} \in W'$ for some $\lambda \in \mathbb{F}$ such that $S(\mathbf{x}) = \lambda\mathbf{x}$. This says $(I - P)T(\mathbf{x}) = \lambda\mathbf{x}$. Hence, $T(\mathbf{x}) = \lambda\mathbf{x} + PT(\mathbf{x})$. Since $PT(\mathbf{x}) \in W$ we find that $W + \text{span}(\{\mathbf{x}\})$ is also invariant under T. ∎

Corollary 6.1.5. *If $T : V \to V$ is triangularizable, with eigenvalues λ_i with respective multiplicities m_i, then there exists a basis β for V such that $[T]_\beta^\beta$ is upper triangular, and the diagonal entries of $[T]_\beta^\beta$ are m_1 λ_1's, followed by m_2 λ_2's, and so on.*

Theorem 6.1.6 (Cayley-Hamilton)**.** *Let $T : V \to V$ be a linear mapping on a finite-dimensional vector space V, and let $p(t) = \det(T - tI)$ be its characteristic polynomial. Assume that $p(t)$ has $dim(V)$ roots in the field $\mathbb{F}$ over which V is defined. Then $p(T) = 0$ is the zero operator.*

Remark 6.1.2. The conclusion of the Cayley-Hamilton theorem is valid even if some of the eigenvalues of T are contained in an extension of the field $\mathbb{F}$.

Remark 6.1.3. The result of the Cayley-Hamilton theorem can be used for inverse computations (although it is rather insufficient).

6.2.0 Canonical Form for Nilpotent Mappings

Definition 6.2.1. *A linear map $N : V \to V$ is nilpotent if there exists some $k \geqslant 1$ such that $N^k = 0$ the zero mapping.*

Lemma 6.2.1. *A mapping is nilpotent if and only if the mapping has one eigenvalue $\lambda = 0$ with multiplicity $n = dim(V)$.*

Remark 6.2.1. Let $dim(V) = n$ and let $N : V \to V$ be a nilpotent mapping. Then, given any vector $\mathbf{x} \in V$, either $\mathbf{x} = \mathbf{0}$ or there exists a unique integer k, $1 \leqslant k \leqslant n$, such that $N^k(\mathbf{x}) = \mathbf{0}$ but $N^{k-1}(\mathbf{x}) \neq \mathbf{0}$. It follows that if $\mathbf{x} \neq \mathbf{0}$ then the set $\{N^{k-1}(\mathbf{x}), ..., N(\mathbf{x}), \mathbf{x}\}$ consists of distinct nonzero vectors.

Definition 6.2.2. *Let N, $\mathbf{x} \neq 0$ and k be as before:*

1. *The set $\{N^{k-1}(\mathbf{x}), ..., N(\mathbf{x}), \mathbf{x}\}$ is called the* cycle *generated by $\mathbf{x}$. $\mathbf{x}$ is called the* initial vector *of the cycle.*

2. *The subspace* $\text{span}\{N^{k-1}(\mathbf{x}), ..., N(\mathbf{x}), \mathbf{x}\}$ *is called the* cyclic subspace *generated by $\mathbf{x}$, and denoted $C(\mathbf{x})$*

3. *The integer k is called the* length *of the cycle*

Proposition 6.2.2. *With all notations as before:*

1. *$N^{k-1}(\mathbf{x})$ is an eigenvector of N with eigenvalue $\lambda = 0$*

2. *$C(\mathbf{x})$ is an invariant subspace of V under N*

3. *The cycle generated by $\mathbf{x} \neq \mathbf{0}$ is a linearly independent set. Hence, $dim(C(\mathbf{x})) = k$, the length of the cycle.*

Observation 6.2.2. If N is a nilpotent mapping with cyclic subspace $C(\mathbf{x})$ generated by some

$\mathbf{x} \in V$ with α being the cycle generated by $\mathbf{x}$, which we view as a basis, it follows that

$$[N|_{C(\mathbf{x})}]_\alpha^\alpha = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & 0 \end{bmatrix}$$

Hence, the matrix of N restricted to a cyclic subspace is a special type of upper triangular matrix.

Proposition 6.2.3. *Let $\alpha_i = \{N^{k_i-1}(\mathbf{x}_i), ..., \mathbf{x}_i\}$, ($1 \leqslant i \leqslant r$) be cycles of lengths k_i, respectively. If the set of eigenvectors $\{N^{k_1-1}(\mathbf{x}_1), ..., N^{k_r-1}(\mathbf{x}_r)\}$ is linearly independent then $\alpha_1 \cup ... \cup \alpha_r$ is linearly independent.*

Definition 6.2.3. *We say that the cycles $\alpha_i = \{N^{k_i-1}(\mathbf{x}_i), ..., \mathbf{x}_i\}$ are* non-overlapping cycles *if $\alpha_1 \cup ... \cup \alpha_r$ is linearly independent.*

Definition 6.2.4. *Let $N : V \to V$ be a nilpotent mapping on a finite-dimensional vector space V. We call a basis β for V a* canonical basis *with respect to N if β is the union of a collection of non-overlapping cycles for N.*

Theorem 6.2.4 (Canonical Form for Nilpotent Maps)**.** *Let $N : V \to V$ be a nilpotent mapping on a finite-dimensional vector space. There exists a canonical basis β of V with respect to N.*

Lemma 6.2.5. *Consider the cycle tableau corresponding to a canonical basis for a nilpotent mapping $N : V \to V$. As before, let r be the number of rows and let k_i be the number of boxes in the ith row ($k_1 \geqslant ... \geqslant k_r$). For each $1 \leqslant j \leqslant k_1$, the number of boxes in the jth column of the tableau is $dim(\ker(N^j)) - dim(\ker(N^{j-1}))$.*

Corollary 6.2.6. *The canonical form of a nilpotent mapping is unique (provided the cycles in the canonical basis are arranged so the lengths satisfy $k_1 \geqslant k_2 \geqslant ... \geqslant k_r$)*

Remark 6.2.3. In order to find a canonical basis for a nilpotent mapping N we first determine the cycle tableau using the previous Lemma. First, the final vector of a cycle of length k must be in $\ker(N) \cap \mathrm{Im}(N^{k-1})$, corresponding to a row of length k in the tableau. We find an eigenvector of N that is also an element of $\mathrm{Im}(N^{k-1})$. If $\mathbf{y}$ is such a final vector then we solve the system $N^{k-1}(\mathbf{x}) = \mathbf{y}$ to determine an initial vector, from which we can then construct the cycle (row). Care must be taken to ensure the final vectors of the cycles are linearly independent.

6.3.0 Jordan Canonical Form

Proposition 6.3.1. *Let $T : V \to V$ be a linear mapping whose characteristic polynomial had $dim(V)$ roots (λ_i with respective multiplicities m_i) in the field $\mathbb{F}$ over which V is defined.*

1. *There exists subspaces $V_i' \subset V$ such that*

(a) Each V_i' is invariant under T

(b) $T|_{V_i'}$ has exactly one distinct eigenvalue λ_i, and

(c) $V = V_1' \oplus V_2' \oplus ... \oplus V_k'$

2. *There exists a basis β for V such that $[T]_\beta^\beta$ has a direct sum decomposition into upper-triangular blocks of the form, with the entries above the diagonal arbitrary and those in the diagonal equal to λ_i.*

Definition 6.3.1. *Let $T : V \to V$ be a linear mapping on a finite-dimensional vector space V. Let λ be an eigenvalue of T with multiplicity m.*

1. *The λ-*generalized eigenspace*, denoted K_λ, is the kernel of the mapping $(T - \lambda I)^m$ on V.*

2. *The nonzero elements of K_λ are called* generalized eigenvectors *of T*

Proposition 6.3.2.

1. *For each eigenvalue λ of T, K_λ is an invariant subspace of V*

2. *If λ_i $(1 \leqslant i \leqslant k)$ are the distinct eigenvalues of T then $V = K_{\lambda_1} \oplus ... \oplus K_{\lambda_k}$*

3. *If λ is an eigenvalue of multiplicity m, then $dim(K_\lambda) = m$*

Definition 6.3.2.

1. *A matrix of the form*

$$\begin{bmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & 1 \end{bmatrix}$$

   is called a Jordan block matrix

2. *A matrix $A \in M_{n\times n}(\mathbb{F})$ is said to be in* Jordan canonical form *if A is a direct sum of Jordan block matrices.*

Theorem 6.3.3 (Jordan Canonical Form)**.** *Let $T : V \to V$ be a linear mapping on a finite-dimensional vector space V whose characteristic polynomial has $dim(V)$ roots in the field $\mathbb{F}$ over which V is defined.*

1. *There exists a basis γ (called the* canonical basis*) of V such that $[T]_\gamma^\gamma$ has a direct sum decomposition into Jordan block matrices*

2. *In this decomposition the number of Jordan blocks and their sizes are uniquely determined by T. (though the order of the blocks may differ)*

Remark 6.3.1. Process for computing the Jordan Canonical Form:

1. Find all eigenvalues of T and their multiplicities.

2. For each distinct eigenvalue λ_i in turn, construct the cycle tableau for a canonical basis of K_{λ_i} with respect to the mapping $N_i = (T - \lambda_i I)|_{K_{\lambda_i}}$. For each j, the number of boxes in the jth column of the tableau for λ_i will be

$$dim(\ker(T - \lambda_i I)^j) - dim(\ker(T - \lambda_i I)^{j-1})$$

(computed on V).

3. Form the corresponding Jordan blocks and assemble the matrix of T.

Chapter 7

Inner Product Spaces

Part II

Module Theory

Chapter 8

Free Modules

Chapter 9

Linear Transformations

Chapter 10

Matrix Theory for Free Modules

Chapter 11

Modules over PIDs

Part III

Multilinear Algebra

Chapter 12

Tensor Products

Appendices

.1.0 Important Algebraic Structures

.2.0 Euclidean Space and Geometry