Linear Algebra Done Right

Notes – Linear Algebra

Zhijie Chen
January 19, 2025

Contents

Vector Spaces 3

Linear Maps 3

Kernal and Image of Linear Maps 3

Products and Quotients of Vector Spaces 3

Duality 4

Polynomials 6

Eigenvalues and Eigenvectors 7

Invariant Subspaces 7

Minimal Polynomials 8

Don’t just read it; fight it!

Ask your own questions.

Look for your own examples.

Discover your own proofs.

Is the hypothesis necessary?

Is the converse true?

What happens in the classical special case?

What about the degenerate cases?

Where does the proof use the hypothesis?

—Paul Holmos

Vector Spaces

Definition 1. The complexification of V, denoted by $V_{\mathbb{C}}$, equals $V \times V$ with normal addition and real scalar multiplication for product space. But we write an element (u, v) of $V_{\mathbb{C}}$ as $u + iv$. Complex scalar multiplication is defined by

$$(a + bi)(u + iv) = (au - bv) + i(av + bu)$$

for all $a, b \in \mathbb{R}$ and all $u, v \in V$.[1]

Lemma 2 (Linear dependence lemma)**.** *Suppose $v_1, \ldots, v_m$ is a linearly dependent set in V. Then there exists $k \in \{1, 2, \ldots, m\}$ such that*

$$v_k \in \operatorname{span}(v_1, \ldots, v_m).$$

Furthermore, removing the k^{th} term from the list does not change the span.[2]

Theorem 3. *Any two bases of a finite-dimensional vector space have the same length.*[3]

Proof. Suppose V is finite-dimensional. Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be two bases of V. Considering $\mathcal{B}_1$ as an independent set and $\mathcal{B}_2$ as a spanning set leads to $\#\mathcal{B}_1 \leq \#\mathcal{B}_2$. Interchanging the roles of $\mathcal{B}_1$ and $\mathcal{B}_2$ and we have $\#\mathcal{B}_2 \leq \#\mathcal{B}_1$. Thus $\#\mathcal{B}_1 = \#\mathcal{B}_2$. □

Linear Maps

Kernal and Image of Linear Maps

Exercise 4. Suppose U and V are finite-dimensional and $S \in \mathcal{L}(V, W)$ and $T \in \mathcal{L}(U, V)$. Prove that

$$\dim \ker ST \leq \dim \ker S + \dim \ker T.$$

Proof. Restrict to $Z = \ker ST$. By the fundamental theorem of linear maps,

$$\begin{aligned}
\dim Z &= \dim T(Z) + \dim \ker T|_Z \\
&\leq \dim T(Z) + \dim \ker T \\
&= \dim ST(Z) + \dim \ker S|_{T(Z)} + \dim \ker T \\
&\leq \dim \ker S + \dim \ker T.
\end{aligned}$$

□

Corollary 5 (Sylvester's rank inequality)**.** *Suppose $A \in \mathbb{F}^{m,n}$ and $B \in \mathbb{F}^{n,p}$ are two matrices. Then*[4]

$$\operatorname{rank} A + \operatorname{rank} B - n \leq \operatorname{rank}(AB)\,.$$

Products and Quotients of Vector Spaces

Notation 6. Suppose $T \in \mathcal{L}(V, W)$. Define $\widetilde{T} : V + \ker V \to V$ by

$$\widetilde{T}(v + \ker T) = Tv.$$

[1] Think of V as a subset of $V_{\mathbb{C}}$ by identifying $u \in V$ with $u + i0$. The construction of $V_{\mathbb{C}}$ from V can then be thought of as generalizing the construction of $\mathbb{C}^n$ from $\mathbb{R}^n$.

[2] This lemma lays the foundation for a series of basic results for vector spaces.

[3] This proposition ensures that the *dimension* of a vector space is well-defined.

[4] There is a slicker proof for this inequality using block matrices. But the proof here using linear maps is more informative.

Exercise 7. Suppose $V_1, \ldots, V_m$ are vector spaces. Prove that $\mathcal{L}(V_1 \times \cdots \times V_m, W)$ and $\mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$ are isomorphic vector spaces.

Proof. We construct an isomorphism T between the two vector spaces.

For every $\Gamma \in \mathcal{L}(V_1 \times \cdots \times V_m, W)$, define $\varphi_k : V_k \to W$ for each k by

$$\varphi_k(v_k) = \Gamma(0, \ldots, v_k, \ldots, 0)$$

with v_k in the k^{th} slot and 0 in all other slots. It can be verified that $\varphi_k \in \mathcal{L}(V_k, W)$.

Define T by $T(\Gamma) = (\varphi_1, \ldots, \varphi_m)$. It can be verified that T is a linear map. We prove T is an isomorphism by constructing its inverse linear map S.

For every $(\varphi_1, \ldots, \varphi_m) \in \mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$, let

$$S(\varphi_1, \ldots, \varphi_m)\,(v_1, \ldots, v_m) = \varphi_1(v_1) + \cdots + \varphi_m(v_m)\,.$$

It can be shown that S is a linear map, and that $S \circ T = I$ and $T \circ S = I$. That proves T is indeed an isomorphism between the two vector spaces. □

Proposition 8. *A nonempty subset A of V is a translate of some subspace of V if and only if $\lambda v + (1 - \lambda)\, w \in A$ for all $v, w \in A$ and all $\lambda \in \mathbb{F}$.*

Exercise 9. Suppose $A_1 = v + U_1$ and $A_2 = w + U_2$ for some $v, w \in V$ and some subspaces U_1, U_2 of V. Prove that $A_1 \cap A_2$ is either the empty set or a translate of some subspace of V.[5]

Proposition 10. *Suppose U is a subspace of V and $v_1 + U, \ldots, v_m + U$ is a basis of V/U and $u_1, \ldots, u_n$ is a basis of U. Then $v_1, \ldots, v_m, u_1, \ldots, u_n$ is a basis of V. In other words, $V = \operatorname{span}(v_1, \ldots, v_m) \oplus U$.*[6]

Exercise 11. Suppose U is a subspace of V such that V/U is finite-dimensional.

(a) Prove that if W is a finite-dimensional subspace of V and $V = U + W$, then $\dim W \geq \dim V/U$.

(b) Prove that there exists a finite-dimensional subspace W of V such that $V = U \oplus W$ and $\dim W = \dim V/U$.

Proof. Let $\overline{w}_1 + U, \ldots, \overline{w}_m + U$ be a basis of V/U. Then by Proposition 10, we have $V = \operatorname{span}(\overline{w}_1, \ldots, \overline{w}_m) \oplus U$. Let $W_0 = \operatorname{span}(\overline{w}_1, \ldots, \overline{w}_m)$, then $V = U \oplus W_0$, as desired.

Now we prove that for each subspace W of V such that $V = U + W$, we have $\dim W \geq m = \dim V/U$.

For each $\overline{w}_i \in V$ above, by definition we have $\overline{w}_i = u_i + w_i$ for some $u_i \in U$ and $w_i \in W$. It can be shown from the linear independence of $\overline{w}_1 + U, \ldots, \overline{w}_m + U$ that $\overline{w}_1 - u_1, \ldots, \overline{w}_m - u_m$ are independent vectors in W. Hence $\dim W \geq m$. □

Duality

Theorem 12. *Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then*

$$T \textit{ is surjective} \iff T' \textit{ is injective} \quad \textit{and} \quad T \textit{ is injective} \iff T' \textit{ is surjective.}$$[7]

[5] Recall Proposition 8.

[6] $V = \operatorname{span}(v_1, \ldots, v_m) \oplus U$ still holds without the hypothesis that U is finite-dimensional.

[7] This result can be useful because sometimes it is easier to verify that T' is injective (surjective) than to show directly that T is surjective (injective).

Proposition 13. *Suppose V is finite-dimensional and U is a subspace of V. Then*

$$U = \left\{ v \in V : \varphi(v) = 0 \text{ for every } \varphi \in U^0 \right\}.$$

Exercise 14. Suppose V is finite-dimensional and U and W are subspaces of V.

(a) Prove that $W^0 \subseteq U^0$ if and only if $U \subseteq W$.

(b) Prove that $W^0 = U^0$ if and only if $U = W$.[8]

Exercise 15. Suppose V is finite-dimensional and U and W are subspaces of V.

(a) Prove that $(U + W)^0 = U^0 \cap W^0$.

(b) Prove that $(U \cap W)^0 = U^0 + W^0$.

Proposition 16. *Suppose V is finite-dimensional and $v_1, \ldots, v_n$ is a basis of V. Then $\varphi_1, \ldots, \varphi_n \in V'$ is the dual basis of $v_1, \ldots, v_n$ if and only if*

$$\begin{bmatrix} \mathcal{M}(\varphi_1, (v_1, \ldots, v_n)) \\ \vdots \\ \mathcal{M}(\varphi_n, (v_1, \ldots, v_n)) \end{bmatrix} = I.$$

Exercise 17. Suppose V is finite-dimensional and $\varphi_1, \ldots, \varphi_n$ is a basis of V'. Prove that there exists a basis of V whose dual basis is $\varphi_1, \ldots, \varphi_n$.

Proof. We start from an arbitrary basis $u_1, \ldots, u_n$ of V. Let $\psi_1, \ldots, \psi_n$ be its dual basis. In this proof, we take standard basis $e_1, \ldots, e_n$ as the basis of $\mathbb{F}^n$.

Define $S, T \in \mathcal{L}(V, \mathbb{F}^n)$ by

$$T(v) = (\varphi_1(v), \ldots, \varphi_n(v)), \quad S(v) = (\psi_1(v), \ldots, \psi_n(v)).$$

Then by Proposition 16, $\mathcal{M}(S, (u_1, \ldots, u_n)) = I$.

Let A be the change of basis matrix from ψ's to φ's, i.e.,

$$\begin{bmatrix} \varphi_1 & \cdots & \varphi_n \end{bmatrix} = \begin{bmatrix} \psi_1 & \cdots & \psi_n \end{bmatrix} A.$$

Then by the definition of change of basis matrix, we have

$$\mathcal{M}(T, (u_1, \ldots, u_n)) = \begin{bmatrix} \mathcal{M}(\varphi_1, (u_1, \ldots, u_n)) \\ \vdots \\ \mathcal{M}(\varphi_n, (u_1, \ldots, u_n)) \end{bmatrix} = A^t \begin{bmatrix} \mathcal{M}(\psi_1, (u_1, \ldots, u_n)) \\ \vdots \\ \mathcal{M}(\psi_n, (u_1, \ldots, u_n)) \end{bmatrix}$$

$$= A^t \cdot \mathcal{M}(S, (u_1, \ldots, u_n)) = A^t.$$

Consider basis $v_1, \ldots, v_n$ of V such that the change of basis matrix from u's to v's is $\left(A^t\right)^{-1}$.[9] Thus

$$\mathcal{M}(T, (v_1, \ldots, v_n)) = \mathcal{M}(T, (u_1, \ldots, u_n)) \cdot \mathcal{M}(I, (v_1, \ldots, v_n), (u_1, \ldots, u_n)) = I.$$

Then by Proposition 16, the dual basis of $v_1, \ldots, v_n$ is precisely $\varphi_1, \ldots, \varphi_n$, as desired. □

[8] Recall Proposition 13.

[9] The change of basis for $V' \to V'$ corresponds to the transpose of $V \leftarrow V$, where transpose and inverse both come from duality. That gives the idea of considering $\left(A^t\right)^{-1}$.

Exercise 18 (A natural isomorphism from primal space onto double dual space)**.** Define $\Lambda : V \to V''$ by

$$(\Lambda v)(\varphi) = \varphi(v)$$

for each $v \in V$ and $\varphi \in V'$.

(a) Prove that if $T \in \mathcal{L}(V)$, then $T'' \circ \Lambda = \Lambda \circ T$.

(b) Prove that if V is finite-dimensional, then Λ is an isomorphism from V onto V''.[10,11]

Polynomials

Theorem 19. *Suppose $p \in \mathcal{P}(\mathbb{F})$ is a nonconstant polynomial of degree m. Then $\lambda \in \mathbb{F}$ is a zero of p if and only if there exists a polynomial $q \in \mathcal{P}(\mathbb{F})$ of degree $m-1$ such that $p(z) = (z-\lambda)q(z)$ for every $z \in \mathbb{F}$.*

Theorem 20. *Suppose $p \in \mathcal{P}(\mathbb{F})$ is a nonconstant polynomial of degree m. Then p has at most m zeros in $\mathbb{F}$.*[12,13]

Theorem 21 (Division algorithm for polynomials)**.** *Suppose that $p, s \in \mathcal{P}(\mathbb{F})$, with $s \neq 0$. Then there exist unique polynomials $q, r \in \mathcal{P}(\mathbb{F})$ such that $p = sq + r$.*

Proof. Let $n = \deg p$ and $m = \deg s$. The case where $n < m$ is trivial. Thus we now assume that $n \geq m$.

The list

$$1, z, \dots, z^{m-1}, s, zs, \dots, z^{n-m}s$$

is linearly independent in $\mathcal{P}_n(\mathbb{F})$. And it also has length $n+1$. Hence the list is a basis of $\mathcal{P}_n(\mathbb{F})$.

Because $p \in \mathcal{P}_n(\mathbb{F})$, there exist unique constants $a_0, \dots, a_{m-1}, b_0, \dots, b_{n-m} \in \mathbb{F}$ such that

$$\begin{aligned} p &= a_0 + a_1 z + \cdots + a_{m-1}z^{m-1} + b_0 s + b_1 zs + \cdots + b_{n-m}z^{n-m}s \\ &= \left(a_0 + a_1 z + \cdots + a_{m-1}z^{m-1}\right) + s\left(b_0 + b_1 z + \cdots + b_{n-m}z^{n-m}\right). \end{aligned} \qquad \square$$

Theorem 22 (Fundamental theorem of algebra, first version)**.** *Every nonconstant polynomial with complex coefficients has a zero in $\mathbb{C}$.*

Proof. Suppose $p \in \mathcal{P}(\mathbb{C})$ is a nonconstant polynomial with highest-order nonzero term $c_m z^m$. Then $|p(z)| \to \infty$ as $|z| \to \infty$. Thus the continuous function $z \mapsto |p(z)|$ has a global minimum at some $\zeta \in \mathbb{C}$. Assume that $p(\zeta) \neq 0$.

Consider polynomial $q(z) = p(z+\zeta)/p(\zeta)$. The function $z \mapsto |q(z)|$ has a global minimum at $z = 0$. Write

$$q(z) = 1 + a_k z_k + \cdots + a_m z^m$$

where k is the smallest positive integer such that the coefficient of z_k is nonzero.

[10] Suppose V is finite-dimensional. Then V and V' are isomorphic, but finding an isomorphism from V onto V' generally requires choosing a basis of V. In contrast, the isomorphism Λ from V onto V'' does not require a choice of basis and thus is more natural.

[11] Another natural isomorphism is $\pi' \in \mathcal{L}((V/U)', V')$ where π is the normal quotient map.

[12] This theorem implies that when a polynomial p has too many zeros, $p = 0$.

[13] This theorem implies that the coefficients of a polynomial are uniquely determined. In particular, the *degree* of a polynomial is well-defined.

Let β be a k^{th} root of $-1/a_k$. There is a constant $c > 1$ such that if $t \in (0,1)$, then

$$|q(t\beta)| \leq \left|1 + a_k t^k \beta^k\right| + ct^{k+1} = 1 - t^k(1 - tc).$$

Thus taking t to be $1/(2c)$ leads to $|q(t\beta)| < 1$. The contradiction implies that $p(\zeta) = 0$, as desired. □

Theorem 23 (Fundamental theorem of algebra, second version)**.** *If $p \in \mathcal{P}(\mathbb{C})$ is a nonconstant polynomial, then p has a unique factorization of the form*

$$p(z) = c(z - \lambda_1) \cdots (z - \lambda_m),$$

where $c, \lambda_1, \ldots, \lambda_m \in \mathbb{C}$.

Theorem 24 (Factorization of a polynomial over $\mathbb{R}$)**.** *If $p \in \mathcal{P}(\mathbb{R})$ is a nonconstant polynomial, then p has a unique factorization of the form*

$$p(x) = c(x - \lambda_1) \cdots (x - \lambda_m)(x^2 + b_1 x + c_1) \cdots (x^2 + b_M x + c_M),$$

where $m, M \in \mathbb{N}$ and $c, \lambda_1, \ldots, \lambda_m, b_1, \ldots, b_M, c_1, \ldots, c_M \in \mathbb{R}$, with $b_k^2 < 4c_k$ for each k.

Exercise 25. Suppose $p, q \in \mathcal{P}(\mathbb{C})$ are nonconstant polynomials with no zeros in common. Let $m = \deg p$ and $n = \deg q$. Prove that there exist $r \in \mathcal{P}_{n-1}(\mathbb{C})$ and $s \in \mathcal{P}_{m-1}(\mathbb{C})$ such that $rp + sq = 1$.

Proof. Define $T : \mathcal{P}_{n-1}(\mathbb{C}) \times \mathcal{P}_{m-1}(\mathbb{C}) \to \mathcal{P}_{m+n-1}(\mathbb{C})$ by $T(r,s) = rp + sq$. It can be shown that T is an injective linear map. Because the domain space and target space have the same dimension, T is surjective, completing the proof. □

Eigenvalues and Eigenvectors

Invariant Subspaces

Theorem 26. *Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbb{F}$. Then the following are equivalent.*[14]

(a) λ is an eigenvalue of T.

(b) $T - \lambda I$ is not injective.

(c) $T - \lambda I$ is not surjective.

(d) $T - \lambda I$ is not invertible.

Exercise 27. Suppose $T \in \mathcal{L}(V)$ has no eigenvalues and $T^4 = I$. Prove that $T^2 = -I$.

Proof. $(T^2 + I)(T + I)(T - I) = 0$. Because T has no eigenvalues, $T + I$ and $T - I$ are both invertible. Hence $T^2 + I = 0$. □

Exercise 28. Suppose that $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ are pairwise distinct. Prove that the list $e^{\lambda_1 x}, \ldots, e^{\lambda_n x}$ is linearly independent in $\mathbb{R}^{\mathbb{R}}$.

[14] The equivalences are useful in that they allow identifying an eigenvalue without explicitly constructing an eigenvector.

Proof. Let $V = \operatorname{span}(e^{\lambda_1 x}, \dots, e^{\lambda_n x})$.[15] Define $D \in \mathcal{L}(V)$ by $Df = f'$. Then $e^{\lambda x}$ is an eigenvector of D corresponding to λ. A list of eigenvectors corresponding to distinct eigenvalues is linearly independent. □

Definition 29. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V invariant under T. The *quotient operator* $T/U \in \mathcal{L}(V/U)$ is defined by

$$(T/U)(v + U) = Tv + U$$

for each $v \in V$.

Exercise 30. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V invariant under T. Prove that each eigenvalue of the quotient operator T/U is an eigenvalue of T.

Proof. It suffices to show that $T/U - \lambda I = (T - \lambda I)/U$ is not injective $\implies$ $T - \lambda I$ is not injective. We prove that $T - \lambda I$ is invertible $\implies$ $(T - \lambda I)/U$ is injective.

Suppose $T - \lambda I$ is invertible. U being invariant under T implies that U is invariant under $T - \lambda I$. Thus $(T - \lambda I)v \in U \iff v \in U$. Suppose $((T - \lambda I)/U)(v + U) = 0$. Then $(T - \lambda I)v \in U$, which implies that $v \in U$, i.e., $v + U = 0$. That proves the injectivity of $(T - \lambda I)/U$. □

Exercise 31. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T has an eigenvalue if and only if there exists a subspace of V of dimension $\dim V - 1$ that is invariant under T.[16]

Proof. We first suppose that T has an eigenvalue λ. Then there exists $\varphi \in V'$ such that $\varphi \circ T = T'\varphi = \lambda\varphi$. Extend φ to a basis $\varphi, \varphi_2, \dots, \varphi_n$ of V' and let $v, v_2, \dots, v_n$ be a basis of V whose dual basis is $\varphi, \varphi_2, \dots, \varphi_n$. Then $(\varphi \circ T)v_k = 0$ for every k. Because $\varphi(Tv_k) = 0$ for every k, $Tv_k \in \operatorname{span}(v_2, \dots, v_n)$. That proves that $\operatorname{span}(v_2, \dots, v_n)$ is invariant under T.

To prove the other direction, reverse the steps to obtain an eigenvector of T'. □

Minimal Polynomials

Exercise 32 (Companion matrix of a polynomial)**.** Suppose $a_0, \dots, a_{n-1} \in \mathbb{F}$. Let $T \in \mathcal{L}(\mathbb{F}^n)$ be such that $\mathcal{M}(T)$ (with respect to the standard basis) is

$$\begin{bmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & 1 & \ddots & & -a_2 \\ & & \ddots & & \vdots \\ & & & 0 & -a_{n-2} \\ & & & 1 & -a_{n-1} \end{bmatrix}$$

Prove that the minimal polynomial of T is the polynomial[17]

$$a_0 + a_1 z + \dots + a_{n-1} z^{n-1} + z^n.$$

[15] Alternatively we can let V be the vector space of differentiable functions on $\mathbb{R}$.

[16] This proof is inspired by $\mathcal{M}(T)$ and its transpose, where an invariant subspace $\operatorname{span}(v_1, \dots, v_m)$ is characterized by zero entries at (i, j), where $m + 1 \le i \le \dim V$ and $1 \le j \le m$.

[17] This exercise implies that every monic polynomial is the minimal polynomial of some operator. Hence an algorithm that could produce exact eigenvalues for each operators on each $\mathbb{F}^n$ does not exist.

Proposition 33. *Suppose $T \in \mathcal{L}(V)$ and $p \in \mathcal{P}(\mathbb{F})$. Then there exists a unique $r \in \mathcal{P}(\mathbb{F})$ such that $p(T) = r(T)$ and $\deg r$ is less than the degree of the minimal polynomial of T.*[18]

[18] This proposition implies that every polynomial applied to an operator can be simplified to a polynomial of smaller degree.