Linear Algebra

Linear algebra is the study of linear structures and functions that preserve this structure. Vector spaces are a mathematical sweet spot. They are completely characterized by their dimension. Linear operators also form a vector space. They can be factored into smaller pieces using invariant subspaces.

Vector Space

A *vector space* over a *field* $\boldsymbol{F}$ is set V with a binary operation $V \times V \to V$, $(v, w) \mapsto v + w$, and a scalar product $\boldsymbol{F} \times V \to V$, $(a, v) \mapsto av$, satisfying the distributive law. Typically $\boldsymbol{F}$ is either the real numbers $\boldsymbol{R}$ or complex numbers $\boldsymbol{C}$. The binary addition is *commutative* ($v + w = w + v$), *associative* ($(u+v)+w = u+(w+v)$), has an *identity element* $\mathbf{0}$ ($v+\mathbf{0} = v$), and each element has an inverse ($v + (-v) = \mathbf{0}$). The scalar product satisfies the *distributive laws* $a(v + w) = av + aw$, $(a + b)v = av + bv$, $(ab)v = a(bv)$, $a, b \in \boldsymbol{F}$, $v, w \in V$. We also require $1v = v$ and $av = va$ for $a \in \boldsymbol{F}$ and $v \in V$.

Exercise. *If* $v + z = v$ *for all* $v \in V$ *then* $z = \mathbf{0}$.

This shows the additive identity is unique.

Solution

Taking $v = \mathbf{0}$, $z = \mathbf{0} + z = \mathbf{0}$.

Exercise. *If* $v + v = v$ *then* $v = \mathbf{0}$.

Hint: $v + (-v) = \mathbf{0}$.

Exercise. *Show* $0v = \mathbf{0}$, $a\mathbf{0} = \mathbf{0}$, *and* $(-1)v = -v$, $a \in \boldsymbol{F}$, $v \in V$.

Solution

Note $0v + 0v = (0 + 0)v = 0v$ so $0v = \mathbf{0}$. If $a \neq 0$ then for any $v \in V$ we have $v + a\mathbf{0} = aa^{-1}v + a\mathbf{0} = a(a^{-1}v + \mathbf{0}) = aa^{-1}v = v$ so $a\mathbf{0} = \mathbf{0}$ since the identity is unique. Since $v + (-1)v = 1v + (-1)v = (1 + (-1))v) = 0v = \mathbf{0}$ we have $(-1)v = -v$.

If S is any set then the set of all functions from S to $\boldsymbol{F}$, $\boldsymbol{F}^S = \{v \colon S \to \boldsymbol{F}\}$, is a vector space. The addition is defined pointwise

$$(v + w)(s) = v(s) + w(s) \text{ for } v, w \in \boldsymbol{F}^S$$

as is the scalar product

$$(av)(s) = av(s) \text{ for } a \in \boldsymbol{F}, v \in \boldsymbol{F}^S.$$

The additive identity, $\mathbf{0}$, is the function $\mathbf{0}(s) = 0$ for all $s \in S$. Note there are two different brands of plus sign in $(v + w)(s) = v(s) + w(s)$. The one on the left is the vector space addition and the one on the right is addition in the underlying field.

Exercise. *Show $\boldsymbol{F}^S$ is a vector space.*

Solution

The vector space addition is commutative since $(v + w)(s) = v(s) + w(s) = w(s) + v(s) = (w + v)(s)$ and the field addition is commutative. Similarly, $(u + v) + w = u + (v + w)$ since the field addition is associative. The $\mathbf{0}$ vector satisfies $(v + \mathbf{0})(s) = v(s) + \mathbf{0}(s) = v(s) + 0 = v(s)$ so $v + \mathbf{0} = v$.

We will see later that every vector space has this form. The cardinality of S is the *dimension* of the vector space.

You are probably already familiar with the vector space $\boldsymbol{R}^n = \{(x_1, \ldots, x_n) \mid x_i \in \boldsymbol{R}, 1 \leq i \leq n\}$, where $\boldsymbol{R}$ is the real numbers. If $S = \{1, \ldots, n\}$ and $x \in \boldsymbol{R}^S$ then $x(i) = x_i$ provides a correspondance between $\boldsymbol{R}^{\{1,\ldots,n\}}$ and $\boldsymbol{R}^n$. The *Kronecker delta* is $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. The *standard basis* in $\boldsymbol{R}^n$ is $\{e_j\}_{1 \leq j \leq n} \subset \boldsymbol{R}^n$ where $(e_j)_i = \delta_{ij}$, $1 \leq i \leq n$.

Exercise. *Show every vector $x = (x_1, \ldots, x_n) \in \boldsymbol{R}^n$ can be written using the standard basis as $x = \sum_j x_j e_j$.*

Independence

A fundamental concept in linear algebra is *independence*. A set of vectors $S \subset V$ is *independent* if every finite sum $\sum_{v \in S} a_v v = \mathbf{0}$, $a_v \in \boldsymbol{F}$, implies $a_v = 0$ for all $v \in S$. A *basis* of a vector space is a set of independent vectors that span V.

Exercise. *Show every singleton $\{v\}$ where $v \neq 0$ is independent.*

Exercise. *Show $\{v, w\}$, $v, w \neq \mathbf{0}$, is (not in)dependent if and only if $v = aw$ for some $a \in \boldsymbol{F}$.*

Exercise. *Show $\{v_i\}$ are dependent if and only if $v_j = \sum_{i \neq j} a_i v_i$, $a_i \in \boldsymbol{F}$, for some j.*

In this case we say v_j is in the *span* of $\{v_i\}_{i \neq j}$.

Subspace

A subset $U \subseteq V$ is a *subspace* of V if U is also a vector space. Clearly, $\{\mathbf{0}\}$ is the smallest subspace and V is the largest subspace of V.

Exercise. *If U and W are subspaces then so are $U \cap W$ and $U + W$.*

Hint: $U + W = \{u + w \mid u \in U, w \in W\}$.

Exercise. *If $v \in V$ then $\boldsymbol{F}v = \{av \mid a \in \boldsymbol{F}\}$ is a subspace.*

Given any set $S \subset V$ define $\operatorname{span} S$ to be the smallest subspace containing S.

Exercise. *If $S \subseteq V$ then* $\operatorname{span} S = \{\sum_{s \in S_0} a_s s \mid a_s \in \boldsymbol{F}, S_0 \subseteq S$ *finite* $\}$.

We need S_0 to be finite for the sum to be defined when S is infinite.

Solution

Taking $S_0 = \{s\}$ and $a_s = 1$ we see S is a subset of the right hand side. Every term of the form $\sum_{s \in S_0} a_s s$ must belong to $\operatorname{span} S$. Since the right hand side is a subspace (show this!) it must be equal to the span of S.

If $U \subseteq V$ is a subspace we can define the *quotient space* $V/U = \{v + U \mid v \in V\}$. Here $v + U = \{v + u \mid u \in U\}$. The quotient space is also a vector space with addition defined by $(v + U) + (w + U) = (v + w) + U$ and scalar multiplication by $a(v + U) = av + U$.

Exercise. *Show the addition and scalar multiplication are well-defined and V/U is a vector space.*

Hint: To show addition is well-defined show $v + U = v' + U$ and $w + U = w' + U$ imply $(v + w) + U = (v' + w') + U$.

Linear Operators

A *linear operator* from the vector space V to the vector space W, $T\colon V \to W$, is a function that preserves the vector space structure: $T(u + v) = Tu + Tv$ and $T(au) = aTu$ for $u, v \in V$, and $a \in \boldsymbol{F}$.

Exercise. *Show $T(au + v) = aTu + v$, $a \in \boldsymbol{F}$, $u, v \in V$ implies T is a linear operator.*

The set of all linear operators from a vector space V to a vector space W, $\mathcal{L}(V, W)$, is also a vector space. The addition is defined by $(S + T)v = Sv + Tv$, $S, T \in \mathcal{L}(V, W)$, $v \in V$ and scalar multiplication by $(aT)v = a(Tv)$, $a \in \boldsymbol{F}$.

Exercise. *Show $\mathcal{L}(V, W)$ is a vector space.*

The values of a linear transformation on a basis determine the linear transformation. If $\{v_i\}$ is a basis of V then for every $v \in V$ there exist unique $x_i \in \boldsymbol{F}$ with $v = \sum_i x_i v_i$ hence $Tv = \sum_i x_i T v_i$. Every vector space can be identified with $\boldsymbol{F}^S$ given a basis $S \subseteq V$.

If $\{w_j\}$ is a basis of W then $Tv_i = \sum_j t_{ij} w_j$ for some $t_{ij} \in \boldsymbol{F}$ so T is represented by a matrix (t_{ij}) given bases of V and W.

Exercise. *If $S\colon U \to V$ and $T\colon V \to W$ are linear transformations represented by (s_{ij}) and (t_{jk}) respectively, then $TS\colon U \to W$ is represented by $(\sum_j s_{ij} t_{jk})$.*

Matrix multiplication is composition of linear operators.

Solution

Let $\{u_i\}$ be a basis of U, $\{v_j\}$ be a basis of V, and $\{w_k\}$ be a basis of W. We have

$$(TS)u_i = T(\sum_j s_{ij}v_j) = \sum_j s_{ij}Tv_j = \sum_j s_{ij}\sum_k t_{jk}w_k = \sum_k(\sum_j s_{ij}t_{jk})w_k.$$

Note how working in terms of a basis can be tedious.

If a linear operator $T\colon V \to W$ is one-to-one and onto then T is an *isomorphism* and we write $V \cong W$. *One-to-one* means $Tu = Tv$ implies $u = v$ and *onto* means for every $w \in W$ there exists $v \in V$ with $Tv = w$. The inverse of such an operator is defined by $T^{-1}w = v$ if and only if $Tv = w$, just as for any function.

Exercise. *Show T^{-1} is linear.*

Isomorphism is an *equivalence relation* on vector spaces.

Exercise. *Show $V \cong V$, $V \cong W$ implies $W \cong V$, and $U \cong V$, $V \cong W$ imply $U \cong W$.*

Solution

The identity function $I\colon V \to V$ is an isomorphism. If $T\colon V \to W$ is an isomorphism then so is $T^{-1}\colon W \to V$. If $S\colon U \to V$ and $T\colon V \to W$ are isomorphisms then so is the composition $TS\colon U \to W$.

The fundamental theorem of linear algebra is that two vector spaces are isomorphic if and only if they have the same dimension. The non-trivial proof of this is omitted.

Vector spaces are classified up to isomorphism by their dimension. Contrast this with, e.g., the classification of finite simple groups.

Invariant Subspace

If $T\colon V \to V$ is a linear operator and U is a subspace of V then it is *invariant* under T if $TU \subseteq U$. Clearly, $\{\mathbf{0}\}$ and V are invariant subspaces.

Exercise. *If U and W are invariant subspaces then so are $U \cap W$ and $U + W$.*

The *kernel* of a linear transformation $T\colon V \to W$ is $\ker T = \{v \in V \mid Tv = 0\}$ and the *range* is $\operatorname{ran} T = \{Tv \mid v \in V\} \subseteq W$.

Exercise. *If $T\colon V \to W$ is a linear operator show the kernel is a subspace of V and the range is a subspace of W.*

Exercise. *Show $\ker T = \{\mathbf{0}\}$ implies T is one-to-one.*

Hint: Show $Tu = Tv$ imples $u = v$, $u, v \in V$.

If $T\colon V \to W$ is one-to-one we can define the inverse $T^{-1}\colon \operatorname{ran} T \to V$ by $T^{-1}w = v$ if and only if $w = Tv$.

Exercise. *If $T\colon V \to V$ is a linear operator then the kernel and range are invariant under T.*

Solution

We have $T(\ker T) = \{\mathbf{0}\} \subseteq \ker T$ and $T(\operatorname{ran} T) = T(TV) \subseteq TV = \operatorname{ran} T$.

If $v_1, \ldots, v_n$ is a basis for V we can define a linear operator $T\colon V \to \boldsymbol{F}^n$ by $Tv_i = e_i$ where $\{e_i\}$ is the standard basis of $\boldsymbol{F}^n$. By linearity $T(\sum_i a_i v_i) = \sum_i a_i e_i = (a_1, \ldots, a_n) \in \boldsymbol{F}^n$.

Exercise *Show T is an isomorphism.*

Eigenvectors/values

If $T\colon V \to V$ is a linear operator and $\boldsymbol{F}v$ is invariant under T for some $v \neq \mathbf{0}$ then v is an *eigenvector* of T. The number $\lambda \in \boldsymbol{F}$ with $Tv = \lambda v$ is the *eigenvalue* corresponding to v. If v is an eigenvector then av is an eigenvector for all nonzero $a \in \boldsymbol{F}$.

Exercise. *If v and w are eigenvectors having the same eigenvalue then $v + w$ is an eigenvector with the same eigenvalue.*

Hint: $\ker(T - \lambda I)$ is a subspace.

Exercise. *If v and w are eigenvectors with different eigenvalues then v and w are independent.*

Solution

Suppose $Tv = \lambda v$ and $Tw = \mu w$ with $\lambda \neq \mu$. If $av + bw = \mathbf{0}$ then $\mathbf{0} = (T - \lambda I)(av + bw) = b(\mu - \lambda)w$ so $b = 0$. Applying $T - \mu I$ shows $a = 0$.

Note if $Tv = \lambda v$ then $v \in \ker T - \lambda I \neq \{0\}$ and so $T - \lambda I$ is not invertible. The *spectrum* of an operator is the set $\sigma(T) = \{\lambda \in \boldsymbol{F} \mid T - \lambda I \text{ is not invertable}\}$. In finite dimensions it is equal to the set of eigenvalus.

It is not the case every linear operator on a vector space over the real numbers has an eigenvector, e.g., a rotation about the origin. It is the case every linear operator on a finite dimensional vector space over the complex numbers does, but proving that requires more machinery.

Exercise. *Let $T\colon \boldsymbol{C}^2 \to \boldsymbol{C}^2$ be defined by $Te_1 = e_2$ and $Te_2 = -e_1$. Show $\sigma(T) = \{i, -i\}$ and find the corresponding eigenvectors.*

Solution

If $Tv = \lambda T$ where $v = ae_1 + be_2$ then $Tv = ae_2 - be_1$ so $\lambda a = -b$ and $\lambda b = a$. Not both a and b are zero. If $a \neq 0$ we have $a = \lambda b = -\lambda^2 a$ so $\lambda = \pm i$. If $b \neq 0$ then $b = -\lambda a = -\lambda^2 a$ fo $\lambda = \pm i$. If $T(ae_1 + be_2) = i(ae_1 + be_2)$ then $-a = ib$

and $b = ia$ so $T(e_1 - ie_2) = e_2 + ie_1 = i(e_1 - ie_2)$ hence $e_1 - ie_2$ is an eigenvector with eigenvalue i. Since $T(e_1 + ie_2) = e_2 - ie_1 = -i(e_1 + ie_2)$ we have $e_1 + ie_2$ is an eigenvector with eigenvalue $-i$.

If $V = \boldsymbol{F}^{\boldsymbol{N}}$ where $\boldsymbol{N} = \{0, 1, 2, \ldots\}$ are the natural numbers define the *forward shift operator* $S\colon V \to V$ by $S(v_0, v_1, v_2, \ldots) = (0, v_0, v_1, \ldots)$.

Exercise. *If* $Sv = \lambda v$ *then* $v = \mathbf{0}$.

Solution

Note S is one-to-one so $Sv = 0$ implies $v = 0$ and we can assume $\lambda \neq 0$. If $Sv = \lambda v$ then $(0, v_0, v_1, \dots) = (\lambda v_0, \lambda v_1, \lambda v_2, \ldots)$. This implies $0 = \lambda v_0$ so $v_0 = 0$. Likewise $0 = v_0 = \lambda v_1$ so $v_1 = 0$. By induction $v_j = 0$ for all j so $v = \mathbf{0}$.

This shows S has no eigenvectors.

Exercise. *Show* $\sigma(S) = \{\lambda \in \boldsymbol{F} \mid |\lambda| \leq 1\}$.

Hint: If $|\lambda| > 1$ then $(\lambda I - S)^{-1} = I/\lambda + S/\lambda^2 + S^2/\lambda^3 + \cdots$.

The forward shift operator has lots of invariant subspaces.

Exercise. *Show* $\mathcal{M}_n = \{v \in V \mid v_j = 0, 0 \leq j \leq n\}$ *is an invariant subspace for* $n \geq 0$.

The structure of linear operators on finite dimensional spaces begins with finding invariant subspaces associated with each eigenvalue. If $\sigma(T) = \{\lambda_1, \ldots, \lambda_m\}$ and $V_i \subseteq V$ are invariant subspaces with $\sigma(T|_{V_i}) = \{\lambda_i\}$, $1 \leq i \leq m$ then $V_i \cap V_j = \{\mathbf{0}\}$ if $i \neq j$.

Functional Calculus

If $T\colon V \to V$ is a linear operator then so is $p(T)\colon V \to V$ for any polynomial p. If q is a polynomial having no roots in the spectrum of T then $q(T)^{-1}$ is also a well-defined linear operator. This provides a *functional calculus* from rational function with no poles in the spectrum of T to linear operators on V

$$\Phi\colon \mathcal{R}(\sigma(T)) \to \mathcal{L}(V), p/q \mapsto p(T)q(T)^{-1}.$$

This is not only a linear map, it also preserves products.

Exercise. *Show* $(rs)(T) = r(T)s(T)$ *if* r *and* s *are rational functions with no poles in* $\sigma(T)$.

Theorem. (Spectral mapping theorem) *If* $r \in \mathcal{R}(\sigma(T))$ *then* $r(\sigma(T)) = \sigma(r(T))$.

Proof: For any $\lambda \in \boldsymbol{C}$ $p(z) - p(\lambda) = (z - \lambda)q(z)$ for some polynomial q. If $\lambda \in \sigma(T)$ then $P(T) - p(\lambda)I = (T - \lambda I)q(T)$.

We can also define the shift operator $J = J^n$ on $\boldsymbol{F}^n$ by $J(x_1, \ldots, x_n) = (0, x_1, \ldots, x_{n-1})$.

Exercise. *Show* e_n *is the only eigenvector and it has eigenvalue* 0.

Note $J^2(x_1, \ldots, x_n) = (0, 0, x_1, \ldots, x_{n-2})$ has eigenvectors e_{n-1} and e_n. Likewise, J^k has eigenvectors $e_{n-k+1}, \ldots, e_n$, $1 \leq k \leq n$. Clearly $J^n = \mathbf{0}$, the zero operator.

It is not hard to show $\sigma(J) = 0$ but we can use the *spectal mapping theorem* to give a simple proof. If p is a polynomial and $T\colon V \to V$ is a linear operator then $p(T)\colon V \to V$ can be defined.

Exercise. *If $T\colon V \to V$ and $T^m = 0$ for some m then $\sigma(T) = \{0\}$.*

Solution

Using the spectral mapping theorem we have $\{0\} = \sigma(T^m) = \sigma(T)^m$. If $0 = \lambda^m$ then $\lambda = 0$.

Jordan Canonical Form

Suppose $T\colon V \to V$ is a linear operator on an n-dimensional space V. For $v \in V$ define its *order*, $o(v)$, to be the minimum m such that $v, Tv, \ldots T^m v$ are linearly dependent. If $o(v)$ equals the dimension of V then v is a *cyclic vector* for T and $T^n v = \sum_{0 \leq j < n} a_j T^j v$ for some $a_j \in \boldsymbol{C}$. Using the basis v, Tv, $\ldots$, $T^{n-1}v$ gives a representation for T as a matrix.

$T(\sum_j a_j T_j) = \sum_j a_j T^{j+1}$

and $\sigma(T) = \{0\}$. For any $v \in V$ the vectors v, Tv, $\ldots$, T^n are linearly dependent so $p(T)v = 0$ from some polynomial p.

Given $v_1, \ldots, v_n \in V$ define the *shift operator* $J\colon V \to V$ by $Jv_i = v_{i+1}$, $1 \leq j < n$ and $Jv_n = \mathbf{0}$.

Dual

The *dual* of a vector space V is the set of all linear operators from the vector space to the underlying field $V^* = \mathcal{L}(V, \boldsymbol{F})$. The *dual pairing* $\langle ., . \rangle\colon V \times V^* \to \boldsymbol{F}$ is defined by $\langle v, v^* \rangle = v^*(v)$, $v \in V$, $v^* \in V*$.

Given any linear operator $T\colon V \to W$ define its *adjoint* $T^*\colon W^* \to V^*$ by $\langle v, T^* w^* \rangle = \langle Tv, w^* \rangle$, $w^* \in W$, $v \in V$. This defines $T^* w^* \in V^*$ at each point $v \in V$. We also have $T^{**}\colon V^{**} \to W^{**}$.

Define $\iota V \to V^{**}$ by $\langle v^*, \iota v \rangle = \langle v, v^* \rangle$.

Exercise. *If $I\colon V \to V$ is the identity, show $I^{**}\colon V^{**} \to V^{**}$ is one-to-one.*