

Linear maps

Matthew Macauley

Department of Mathematical Sciences
Clemson University
<http://www.math.clemson.edu/~macaule/>

Math 8530, Spring 2017

Goal

Abstract the concept of a matrix as a linear mapping between vector spaces.

Advantages:

- simple, transparent proofs;
- better handles infinite dimensional spaces.

Definition

A **linear map** (or *mapping*, *transformation*, or *operator*) between vector spaces X and U over K is a function $T: X \rightarrow U$ that is:

- (i) additive: $T(x + y) = T(x) + T(y)$, for all $x, y \in X$;
- (ii) homogeneous: $T(ax) = aT(x)$, for all $x \in X$, $a \in K$.

The **domain space** is X and the **target space** is U .

Usually we'll write Tx for $T(x)$, and so the additive property is just the distributive law:

$$T(x + y) = Tx + Ty.$$

Examples of linear maps

Examples

- (i) Any isomorphism;
- (ii) $X = U = \{\text{polynomials of degree } < n \text{ in } s\}$, $T = \frac{d}{ds}$.
- (iii) $X = U = \mathbb{R}^2$, $T = \text{rotation about the origin}$.
- (iv) X any vector space, $U = K$ (1-dimensional), T any $\ell \in X'$.
- (v) $X = U = C_0(\mathbb{R})$, $(Tf)(x) = \int_{-1}^1 f(y)(x-y)^2 dy$.
- (vi) $X = \mathbb{R}^n$, $U = \mathbb{R}^m$, $u = Tx$, where $u_i = \sum_{j=1}^n t_{ij}x_j$, $i = 1, \dots, m$.
- (vii) $X = U = \{\text{piecewise cont. } [0, \infty) \rightarrow \mathbb{R} \text{ of "exponential order"}\}$,
 $(Tf)(s) = \int_0^{\infty} f(t)e^{-st} dt$. "Laplace transform"
- (viii) $X = U = \{\text{functions with } \int_{-\infty}^{\infty} |f(x)| dx < \infty\}$,
 $(Tf)(\xi) = \int_{-\infty}^{\infty} f(x)e^{i\xi x} dx$. "Fourier transform"

Basic properties

Theorem 3.1

Let $T: X \rightarrow U$ be a linear map.

- (a) The **image** of a subspace of X is a subspace of U .
- (b) The **preimage** of a subspace U is a subspace of X .

(Proof is a HW exercise.)

□

Definition

The **range** of T is the image $R_T := T(X)$. The **rank** of T is $\dim R_T$.

The **nullspace** of T is the preimage of 0:

$$N_T := T^{-1}(0) = \{x \in X : Tx = 0\}.$$

The **nullity** of T is $\dim N_T$.

Rank-nullity theorem

Theorem 3.2

Let $T: X \rightarrow U$ be a linear map. Then $\dim N_T + \dim R_T = \dim X$.

Proof

Since T maps N_T to 0, then $Tx_1 = Tx_2$ if $x_1 \equiv x_2 \pmod{N_T}$.

Thus, T extends to a well-defined map on the quotient space X/N_T :

$$T: X/N_T \longrightarrow U, \quad T\{x\} = Tx.$$

Note that this map is 1-1, and so $\dim(X/N_T) = \dim R_T$.

Therefore, $\dim X = \dim N_T + \dim X/N_T = \dim N_T + \dim R_T$. □

Consequences of the rank-nullity theorem

Corollary A

Suppose $\dim U < \dim X$. Then $Tx = 0$ for some $x \neq 0$.

Proof

We have $\dim R_T \leq \dim U < \dim X$, so by the R-N Theorem, $\dim N_T > 0$.

Thus, there is some nonzero $x \in N_T$. □

Example A

Take $X = \mathbb{R}^n$, $U = \mathbb{R}^m$, with $m < n$. Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ be any linear map (see Example (vi)).

Since $m = \dim U < \dim X < n$, Corollary A implies that the system of m equations

$$\sum_{j=1}^n t_{ij}x_j = 0 \quad i = 1, \dots, m$$

has a non-trivial solution, i.e., not all $x_j = 0$.

Consequences of the rank-nullity theorem

Corollary B

Suppose $\dim X = \dim U$ and the only vector satisfying $Tx = 0$ is $x = 0$. Then $R_T = U$.

Proof

We have $N_T = \{0\}$, which means that $\dim N_T = 0$.

Clearly, $R_T \leq U$ ["is a subspace of"]. We just need to show they have the same dimension.

By the R-N Theorem, $\dim U = \dim X = \dim R_N + \dim N_T = \dim R_N$. □

Example B

Take $X = U = \mathbb{R}^n$, and $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ given by $\sum_{j=1}^n t_{ij}x_j = u_i$, for $i = 1, \dots, n$.

If the related **homogeneous system** of equations $\sum_{j=1}^n t_{ij}x_j = 0$, for $i = 1, \dots, n$, has only the trivial solution $x_1 = \dots = x_n = 0$, then the **inhomogeneous system** T has a **unique** solution for all u_1, \dots, u_n .

[Reason: $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is an isomorphism.]

Applications of the rank-nullity theorem

Application 1: Polynomial interpolation

Take $X = \{p \in \mathbb{C}[x] : \deg p < n\}$, $U = \mathbb{C}^n$, and let $s_1, \dots, s_n \in \mathbb{C}$ all be distinct. Define

$$T: X \rightarrow U, \quad Tp = (p(s_1), \dots, p(s_n)).$$

Suppose $Tp = 0$ for some $p \in X$. Then $p(s_1) = \dots = p(s_n) = 0$, which is impossible because p has at most $n - 1$ distinct roots.

Therefore $N_T = \{0\}$, and so Corollary B implies that $R_T = U$.

Application 2: Average values of polynomials

Let $X = \{p \in \mathbb{R}[x] : \deg p < n\}$, $U = \mathbb{R}^n$, and I_1, \dots, I_n be pairwise disjoint intervals on \mathbb{R} .

The **average value** of p over I_j is the integral

$$\bar{p}_j := \frac{1}{|I_j|} \int_{I_j} p(s) ds.$$

Define $T: X \rightarrow U$ by $Tp = (\bar{p}_1, \dots, \bar{p}_n)$.

Suppose $Tp = 0$. Then $\bar{p}_j = 0$ for all j , and so p (if nonzero) must change sign in I_j .

But this would imply that p has n distinct roots, which is impossible.

Thus, $N_T = \{0\}$, and so $R_T = U$.

Application 3: Numerical solutions to Laplace's equation

Laplace's equation is $\Delta u = u_{xx} + u_{yy} = 0$, where $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is a linear operator.

Solutions to Laplace's PDE (“**harmonic functions**”) are the functions in the nullspace of Δ .

If we fix the value of u on the boundary of a region $G \subset \mathbb{R}^2$, the solution to the **boundary value problem** $\Delta u = 0$ is as “flat as possible”. [*Think*: plastic wrap stretched around ∂G .]

This models **steady-state solutions** to the heat equation PDE: $u_t = \Delta u$.

The **finite difference method** is a way to solve $\Delta u = 0$ numerically, using a square lattice with mesh spacing $h > 0$.

At a fixed lattice point O , let u_0 be the value of u at O , and u_W, u_E, u_N, u_S be the values at the neighbors.

We can approximate the derivatives with *centered differences*:

$$u_{xx} \approx \frac{u_W - 2u_0 + u_E}{h^2}, \quad u_{yy} \approx \frac{u_N - 2u_0 + u_S}{h^2}.$$

Plugging this back into $\Delta u = 0$ gives $u_0 = \frac{u_W + u_N + u_E + u_S}{4}$, i.e., u_0 is the average of its four neighbors.

Application to numerical analysis (cont.)

Recall that we are trying to solve an **inhomogeneous boundary value problem** for Laplace's equation

$$\Delta u = 0, \quad u|_{\partial G} = f(x, y) \neq 0.$$

Claim

The **homogeneous equation**: $\Delta u = 0$, where $u = 0$ on ∂G , has *only* the trivial solution $u_0 = 0$ for all $(x, y) \in G$.

Proof

Let \hat{O} be the lattice point at which u achieves its maximum value.

Since $u_0 = \frac{u_W + u_N + u_E + u_S}{4}$, then $u_0 = u_W = u_N = u_E = u_S$.

Repeating this, we see that *all* lattice points take the same value for u , and so $u = 0$.

By the result in Example B, the related **inhomogenous system** for $\Delta u = 0$, with arbitrary (non-zero) boundary conditions has a unique solution. \square

Algebra of linear mappings

Definition

Let $S, T: X \rightarrow U$ be linear maps. Define

- $T + S$ by $(T + S)(x) = Tx + Sx$ for each $x \in X$.
- aT by $(aT)(x) = T(ax)$ for each $x \in X$, $a \in K$.

Easy fact

The set of linear maps from $X \rightarrow U$, denoted $\mathcal{L}(X, U)$, or $\text{Hom}(X, U)$, is a vector space.

Theorem 3.3 (HW exercise)

If $T: X \rightarrow U$ and $S: U \rightarrow V$ are linear maps, then so is $(S \circ T): X \rightarrow V$.

Moreover, composition is distributive w.r.t. addition. That is, if $P, T: X \rightarrow U$ and $R, S: U \rightarrow V$, then

$$(R + S) \circ T = R \circ T + S \circ T, \quad S \circ (T + P) = S \circ T + S \circ P.$$

Remarks

- We usually just write $S \circ T$ as just ST .
- In general, $ST \neq TS$ (note that TS may not even be defined).

Invertibility

Definition

A linear map T is **invertible** if it is 1-1 and onto (i.e., if it is an **isomorphism**). Denote the inverse by T^{-1} .

Exercise

If T is invertible, then TT^{-1} is the identity.

Theorem 3.4 (exercise)

Let $T: X \rightarrow U$ be linear.

- (i) If T is linear, then so is T^{-1} .
- (ii) If S and T are invertible and ST defined, then it is invertible with $(ST)^{-1} = T^{-1}S^{-1}$.

Examples

- (ix) Take $X = U = V = \mathbb{R}[s]$, with $T = \frac{d}{ds}$ and $S =$ multiplication by s .
- (x) Take $X = U = V = \mathbb{R}^3$, with S a 90° -rotation around the x_1 axis, and T a 90° -rotation around the x_2 axis.

In both of these examples, S and T are linear with $ST \neq TS$. (Which are invertible?)

Transposes

Let $T: X \rightarrow U$ be linear and $\ell \in U'$ (recall: $\ell: U \rightarrow K$).

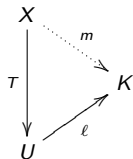
The composition $m := \ell T$ is a linear map $X \rightarrow K$, i.e., an element of X' .

Since T is fixed, this defines an assignment of each $m \in X'$ to $\ell \in U'$.

This defines the following linear map, called the **transpose** of T :

$$T': U' \rightarrow X', \quad T': \ell \mapsto m,$$

Using scalar product notation we can rewrite $m(x) = \ell(T(x))$ as $(m, x) = (\ell, Tx)$.



Key property

The transpose of $T: X \rightarrow U$ is the (unique) map $T': U' \rightarrow X'$ that satisfies $m = T'\ell$, i.e.,

$$(T'\ell, x) = (\ell, Tx), \quad \text{for all } x \in X, \ell \in U'.$$

Caveat: We are writing ℓT for $\ell \circ T$, but $T'\ell$ for $T'(\ell)$ (much like Tx for $T(x)$).

Properties (HW exercise)

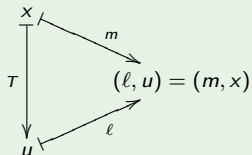
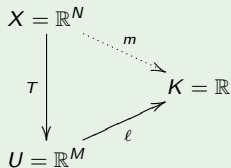
Whenever meaningful, we have

$$(ST)' = T'S', \quad (T + R)' = T' + R', \quad (T^{-1})' = (T')^{-1}.$$

Transposes

Examples (cont.)

(xi) Let $X = \mathbb{R}^N$, $U = \mathbb{R}^M$, and $Tx = u$, where $u_i = \sum_{j=1}^N t_{ij}x_j$.



By definition, for some $\ell_1, \dots, \ell_m \in K$,

$$(\ell, u) = \sum_{i=1}^M \ell_i u_i = \sum_{i=1}^M \ell_i \left(\sum_{j=1}^N t_{ij} x_j \right) = \sum_{i=1}^M \sum_{j=1}^N \ell_i t_{ij} x_j = \sum_{i=1}^M \left(\ell_i \sum_{j=1}^N t_{ij} x_j \right) = \sum_{j=1}^N m_j x_j$$

This gives us a formula for $m = (m_1, \dots, m_N)$, where $(\ell, u) = (m, x)$.

We'll see later that if we express T in matrix form, then T' is formed by making the rows of T the columns of T' .

Transposes

Proposition

If X'' and U'' are canonically identified with X and U , respectively, then $T'' = T$. \square

Theorem 3.5

The annihilator of the range of T is the nullspace of its transpose, i.e., $R_T^\perp = N_{T'}$.

Proof

$$\begin{aligned} \text{By definition,} \quad R_T^\perp &= \{l \in U' : (l, u) = 0 \quad \forall u \in R_T\} \\ &= \{l \in U' : (l, Tx) = 0 \quad \forall x \in X\} \\ &= \{l \in U' : (T'l, x) = 0 \quad \forall x \in X\} \\ &= N_{T'}. \end{aligned}$$

Thus, $l \in R_T^\perp$ iff $T'l = 0$, i.e., iff $l \in N_{T'}$. \square

Applying \perp to both sides of $R_T^\perp = N_{T'}$ (Theorem 3.5) yields the following:

Corollary 3.5

The range of T is the annihilator of the nullspace of T' , i.e., $R_T = N_{T'}^\perp$. \square

Transposes

Theorem 3.6

For any linear mapping $T: X \rightarrow U$, we have $\dim R_T = \dim R_{T'}$.

Proof

We can deduce the following easy facts:

- $\dim R_T^\perp + \dim R_T = \dim U$ (Theorem 2.4 applied to $R_T \subseteq U$);
- $\dim N_{T'} + \dim R_{T'} = \dim U'$ (R-N Theorem applied to $T': U' \rightarrow X'$);
- $\dim U = \dim U'$ (Theorem 2.2).

Now, $R_T^\perp = N_{T'}$ (Theorem 3.5) immediately yields the result. \square

Corollary 3.6

Let $T: X \rightarrow U$ be linear with $\dim X = \dim U$. Then $\dim N_T = \dim N_{T'}$.

Proof

Apply the R-N Theorem to $T: X \rightarrow U$ and $T': U' \rightarrow X'$:

- $\dim N_T = \dim X - \dim R_T$;
- $\dim N_{T'} = \dim U' - \dim R_{T'}$.

Now apply $\dim X = \dim U = \dim U'$ (assumption), and $\dim R_T = \dim R_{T'}$ (Theorem 3.6). \square

Algebra of linear mappings, revisited

Definition

An **endomorphism** of a vector space X is a linear map from X to itself. Denote the set of endomorphisms of X by $\mathcal{L}(X, X)$ or $\text{Hom}(X, X)$ or $\text{End}(X)$.

Remarks

$\mathcal{L}(X, X)$ is a vector space, but we can also “multiply” vectors; it is an **algebra**.

It is an **associative** but **noncommutative** algebra, with **unity** I , satisfying $Ix = x$.

$\mathcal{L}(X, X)$ contains **zero divisors**: pairs S, T such that $ST = 0$ but neither S nor T is zero.

Proposition

If $A \in \mathcal{L}(X, X)$ is a left inverse of $B \in \mathcal{L}(X, X)$ [i.e., $AB = I$], then it is also a right inverse [i.e., $BA = I$]. \square

Definition

The **invertible** elements of $\mathcal{L}(X, X)$ forms the **general linear group**, denoted $\text{GL}(n, K)$, where $n = \dim X$.

Every $S \in \text{GL}(n, K)$ defines a **similarity transformation** of $\mathcal{L}(X, X)$, sending $M \mapsto M_S := SMS^{-1}$, for each $M \in \mathcal{L}(X, X)$. We say M and M_S are **similar**.

Similarity

Theorem 3.7

Every similarity transform is an **automorphism** [“structure-preserving bijection”] of $\mathcal{L}(X, X)$:

$$(kM)_S = kM_S, \quad (M + N)_S = M_S + N_S, \quad (MN)_S = M_S N_S.$$

Moreover, the set of similarity transforms forms a group under $(M_S)_T := M_{TS}$, called the **inner automorphism** group of $GL(n, K)$.

Proof

Verification of $(kM)_S = kM_S$, and $(M + N)_S = M_S + N_S$ is trivial.

Next, observe that $M_S N_S = (SMS^{-1})(SNS^{-1}) = SMNS^{-1} = (MN)_S$.

Finally, $(M_S)_T = T(SMS^{-1})T^{-1} = (TS)M(TS)^{-1} = M_{TS}$.

Checking the group axioms is a straight-forward exercise. □

Theorem 3.8 (exercise)

Similarity is an **equivalence relation**, i.e., it is:

- (i) Reflexive: $M \sim M$;
- (ii) Symmetric: $L \sim M$ implies $M \sim L$;
- (iii) Transitive: $L \sim M$ and $M \sim N$ implies $L \sim N$. □

Algebra of linear mappings

Theorem 3.9 (HW exercise)

If either A or B in $\mathcal{L}(X, X)$ is invertible, then AB and BA are similar. \square

Given any $A \in \mathcal{L}(X, X)$ and polynomial $p(s) = a_N s^N + \cdots + a_1 s + a_0$, consider the polynomial $p(A) = a_N A^N + \cdots + a_1 A + a_0 I$.

The set of polynomials in A is a **commutative subalgebra** of $\mathcal{L}(X, X)$. [to be revisited]

Miscellaneous definitions

- A linear map $P: X \rightarrow X$ is a **projection** if $P^2 = P$.
- The **commutator** of $A, B \in \mathcal{L}(X, X)$ is $[A, B] := AB - BA$, which is 0 iff A and B commute.

Examples (cont.)

(xii) If $X = \{f: \mathbb{R} \rightarrow \mathbb{R}, \text{contin.}\}$, then the following maps $P, Q \in \mathcal{L}(X, X)$ are projections:

- $(Pf)(x) = \frac{f(x) + f(-x)}{2}$; this is the **even part** of f .
- $(Qf)(x) = \frac{f(x) - f(-x)}{2}$; this is the **odd part** of f .

Note that $f = Pf + Qf$ for any $f \in X$.