MTHSC 853: Linear Algebra

3. Linear Mappings.

The goal here is to abstract the concept of a matrix as a linear mapping between linear spaces. This has several advantages:

* Simple, transparent proofs
* Better handles infinite dimensional spaces.

Recall the concept of a linear mapping (or linear map, transformation, or operator) between vector spaces X and U is a function $T: X \longrightarrow U$ that is

(i) additive: $T(x+y) = T(x) + T(y) \quad \forall x, y \in X$, and

(ii) homogeneous: $T(ax) = aT(x) \quad \forall x \in X,\ a \in K$.

The space X is called the domain space, and U the target space. We assume both are over the same field, K.

We usually write Tx for $T(x)$, so the additative property just becomes the distributive law: $T(x+y) = Tx + Ty$.

Examples:

(1) Any isomorphism

(2) $X = U = \{$polynomials of degree $< n$ in $s\}$, $T = \frac{d}{ds}$.

(3) $X = U = \mathbb{R}^2$, T rotation about the origin

(4) X any linear space, $U = K$ (1-dimensional), T any linear function on X.

(5) $X = U = C_0(\mathbb{R})$ (continuous functions on $\mathbb{R}$).

$$(Tf)(x) = \int_{-1}^{1} f(y)\,(x-y)^2\,dy$$

(6) $X = \mathbb{R}^n$, $U = \mathbb{R}^m$, $u = Tx$, where $u_i = \sum_{j=1}^{n} t_{ij}\, x_j \quad i = 1, \dots, m$.

(7) $X = \{$ piecewise continuous functions $[0,\infty) \to \mathbb{R}$, of "exponential order" i.e, $|f(t)| \le c e^{-at}$ for suff. large t, $a, c > 0$ consts. $\}$.

$U = X \qquad (Tf)(s) = \int_0^\infty f(t)\, e^{-st}\, dt.$ "<u>Laplace transform</u>"

(8) X = absolutely integrable functions: $\int_{-\infty}^{\infty} |f(x)|\,dx < \infty$.

$U = X. \quad (Tf)(\xi) = \int_{-\infty}^{\infty} f(x)\, e^{i\xi x}\, dx.$ "<u>Fourier transform</u>"

<u>Theorem 3.1</u> (a) The <u>image</u> of a subspace of X under a linear map $T: X \to U$ is a subspace of U.

(b) The <u>inverse image</u> of a subspace of U is a subspace of X.

<u>Proof</u>: Exercise.

<u>Def</u>: The <u>range</u> of T is the image $T(X)$; we denote as R_T.

The <u>nullspace</u> of T is the inverse image of 0, i.e, $\{x \in X : Tx = 0\}$; we denote as N_T.

By Theorem 3.1, R_T and N_T are subspaces (of U and X, resp.).

<u>Theorem 3.2</u>: (Rank-nullity theorem). Let $T: X \to U$ be a linear map. Then $\dim N_T + \dim R_T = \dim X$.

<u>Proof</u>: Since T maps N_T to 0, $Tx_1 = Tx_2$ if $x_1 \equiv x_2 \mod N_T$.

Thus T acts on the quotient space X/N_T, i.e.,

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

Note that this is a bijection between X/N_T and R_T.

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

We have $\dim X = \dim N_T + \dim X/N_T = \dim N_T + \dim R_T$. $\square$

<u>Corollary A</u>: Suppose $\dim U < \dim X$. Then $Tx = 0$ for some $x \neq 0$.

<u>Proof</u>: We have $\dim R_T \leq \dim U < \dim X$, so by Thm 3.2, $\dim N_T > 0$. Thus, $\exists\, x \in N_T$ s.t. $x \neq 0$. $\square$

<u>Example A</u>: Take $X = \mathbb{R}^n$, $U = \mathbb{R}^m$, $m < n$. Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be any linear mapping (see Example 6). Since $m = \dim U < \dim X < n$, Corollary A implies that the system of equations

$\sum_{j=1}^{n} t_{ij} x_j \quad i = 1, \ldots, m$ has a non-trivial solution, i.e, some $x_j \neq 0$.

<u>Corollary B</u>: Suppose $\dim U = \dim X$ and the only vector satisfying $Tx = 0$ is $x = 0$. Then $R_T = U$.

<u>Proof</u>: We have $N_T = \{0\} \Rightarrow \dim N_T = 0$. By Theorem 2.2, $\dim R_T = \dim X = \dim U$. By the corollary to Theorem 1.6, $R_T = U$.

Example B: Take $X = \mathbb{R}^n$, $U = \mathbb{R}^n$, T given by $\sum_{j=1}^{n} t_{ij} x_j = u_i$ for $i = 1, \ldots, n$. If the related homogeneous system of equations $\sum_{j=1}^{n} t_{ij} x_j = 0$, $i = 1, \ldots, n$ has only the trivial solution, i.e., $x_1 = \cdots = x_n = 0$, then the inhomogeneous system T has a unique solution for all $u_1, \ldots, u_n$. (Because T is an isomorphism from $\mathbb{R}^n$ to $\mathbb{R}^n$).

Application 1: Take $X = \{p(x) \in \mathbb{C}[x] : \deg p(x) < n\}$, $U = \mathbb{C}^n$.

Let $s_1, \ldots, s_n \in \mathbb{C}$ all be distinct. Define $T: X \to U$ by

$Tp = (p(s_1), \ldots, p(s_n))$.

Suppose $Tp = 0$ for some $p \in X$. Then $p(s_1) = \cdots = p(s_n) = 0$, which is impossible because p can have at most $n-1$ distinct roots. Thus, $N_T = \{0\}$.

By Cor. B, $R_T = U$.

Application 2: Take $X = \{p \in \mathbb{R}[x] : \deg p < n\}$, $U = \mathbb{R}^n$.

Let $S_1, \ldots, S_n$ be pairwise disjoint intervals on $\mathbb{R}$, and define

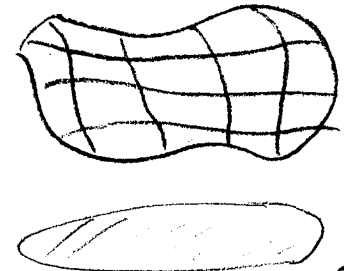
$\bar{p}_j = \frac{1}{|S_j|} \int_{S_j} p(s)\, ds$ (the average value of p over S_j).

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

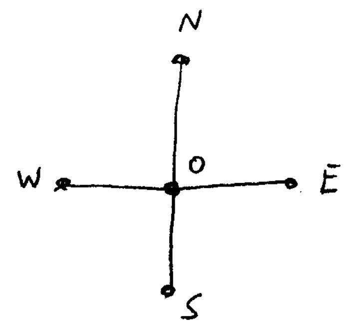
Suppose $Tp = 0$. Then $\bar{p}_j = 0$ $\forall j$, and so p (if nonzero) must change sign in S_j, i.e, p has a root in each S_j. But p has at most $n-1$ roots. Thus, $N_T = \{0\} \Rightarrow R_T = U$.

Application 3: Let $u(x,y)$ be twice differentiable in x & y.

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 $\Delta u = 0$ are precisely the functions in the nullspace of Δ. If we fix the value of u on a bounded region $G \subset \mathbb{R}^2$, then the solution to $\Delta u = 0$ ("harmonic functions") is as "flat as possible." (Think: plastic wrap stretched around ∂G). This makes sense, because they represent steady-state solutions to the heat equation PDE: $u_t = \Delta u$.



Suppose we want to solve $\Delta u = 0$ numerically, using a square lattice with mesh spacing $h > 0$, and finite difference methods.



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:

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

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

Claim: The homogeneous equation (boundaries fixed at zero) has only the trivial solution, $u_0 = 0$ for all points in G.

Proof (sketch): Let O be the lattice point at which u achieves its maximum value. Since $u_0 = \frac{u_W + u_N + u_E + u_S}{4}$, $u_0 = u_W = u_N = u_E = u_S$. Repeating this, we see that all lattice points take the same value for u, so $u = 0$.

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

Algebra of linear mappings

Let $S, T : X \longrightarrow U$ be linear maps. Define their sum $T+S$ by

$$(T+S)(x) = Tx + Sx \quad \text{for each } x \in X.$$

Similarly, define aT by $(aT)(x) = T(ax)$ for each $x \in X$.

Prop: The set of linear mappings from X to U, denoted $L(X, U)$ is a vector space.

Theorem 3.3: If $T : X \longrightarrow U$ and $S : U \longrightarrow V$ are linear maps between vector spaces, then so is their composition $(S \circ T) : X \longrightarrow V$. Moreover, composition is distributive w.r.t. addition, i.e, if $P, T : X \to U$ and $R, S : U \longrightarrow V$, then

$$(R+S) \circ T = R \circ T + S \circ T \quad \text{and} \quad S \circ (T+P) = S \circ T + S \circ P.$$

Proof: Exercise.

We usually write $S \circ T$ as just ST.

<u>Note</u>: In general, $ST \neq TS$ (in fact, TS may not even be defined!

<u>Examples</u>: (cont)

(9) Take $X = U = V = \mathbb{R}[s]$, $T = \frac{d}{ds}$, $S =$ mult. by s

(10) Take $X = U = V = \mathbb{R}^3$. S: $90°$ rotation around x_1-axis

T: $90°$ rotation around x_2-axis.

In both of these examples, S and T are linear, and $ST \neq TS$ (Check!).

<u>Def</u>: A linear map is called <u>invertible</u> if it is 1-1 and onto (i.e., if it is an isomorphism). Denote its inverse by T^{-1}.

<u>Prop</u>: If T is invertible, then TT^{-1} is the identity.

<u>Proof</u>: Exercise.

<u>Theorem 3.4</u>: (i) If T is linear, then so is T^{-1}.

(ii) If S & T are invertible, and ST is defined, then ST is also invertible, with inverse $(ST)^{-1} = T^{-1}S^{-1}$

<u>Proof</u>: Exercise.

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

Then, the product ℓT is a linear map $X \to K$, i.e., an element of X'.

$$\begin{tikzcd} X \arrow[dr, "m"] \arrow[dd, "T"'] & \\ & K \\ U \arrow[ur, "\ell"'] & \end{tikzcd}$$

Let's call this map m, i.e., $m(x) = \ell(Tx)$.

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

This is a linear map: $T': U' \longrightarrow X'$, $\ell \longmapsto m$ called the <u>transpose</u> of T.

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

The transpose was defined by $m = T'\ell$, i.e., $\boxed{(T'\ell, x) = (\ell, Tx)}$. $(*)$

<u>Prop</u>: Whenever meaningful, we have:

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

<u>Proof</u>: Exercise (HW).

<u>Example (11)</u>. Take $X = \mathbb{R}^N$, $U = \mathbb{R}^M$, T as in Example (6), i.e., $u_i = \sum_{j=1}^{N} t_{ij} x_j$.

Here, $X' = \mathbb{R}^N$ and $U' = \mathbb{R}^M$, so $T: \mathbb{R}^N \to \mathbb{R}^M$, $T': \mathbb{R}^M \to \mathbb{R}^N$

The maps T, l, and m are related by the following:

$$\begin{tikzcd} X = \mathbb{R}^N \arrow[d, "T"'] \arrow[dr, "m"] & \\ U = \mathbb{R}^M \arrow[r, "l"'] & \mathbb{R} = K \end{tikzcd} \qquad \begin{tikzcd} x \arrow[d, mapsto] \arrow[dr, mapsto] & \\ u \arrow[r, mapsto] & (l, u) = (m, x) \end{tikzcd}$$

By definition, $(l, u) = \sum_{i=1}^{M} l_i u_i$ for some $l_i \in K$

$$= \sum_{i=1}^{M} l_i \left(\sum_{j=1}^{N} t_{ij} x_j \right) = \sum_{i=1}^{M} \sum_{j=1}^{N} l_i t_{ij} x_j$$

$$= \sum_{j=1}^{N} \sum_{i=1}^{M} l_i t_{ij} x_j = \sum_{j=1}^{N} \left(\sum_{i=1}^{M} l_i t_{ij} \right) x_j$$

$$= \sum_{j=1}^{N} m_j x_j = (m, x).$$

Since $m = T'l$, we have $m_j = \sum_{i=1}^{M} l_i t_{ij}$.

Prop: "If X'' is canonically identified with X and U'' with U, then $T'' = T$.

Proof: Exercise.

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'.

Theorem 3.5: The annihilator of the range of T is the nullspace of its transpose, i.e., $R_T^{\perp} = N_{T'}$

Proof: By definition,

$$R_T^\perp = \{\ell \in U' : (\ell, u) = 0 \ \forall u \in R_T\}.$$

$$= \{\ell \in U' : (\ell, Tx) = 0 \ \forall x \in X\}.$$

$$= \{\ell \in U' : (T'\ell, x) = 0 \ \forall x \in X\}.$$

(by (*)).

$$\begin{tikzcd} X \arrow[d, "T"'] \arrow[dr, "m"] & \\ U \arrow[r, "\ell"'] & K \end{tikzcd} \qquad \begin{tikzcd} x \arrow[d, mapsto, "T"'] \arrow[dr, mapsto, dashed, "m"] & \\ u \arrow[r, mapsto, "\ell"'] & 0 \end{tikzcd}$$

Recall: $T' : U' \to X'$, $\ell \xmapsto{T'} m$

So $\ell \in N_{T'}$ iff $\ell \xmapsto{T'} 0 = m$

$\ell \in R_T^\perp$ iff $\ell : Tu \xmapsto{T'} 0 = m$

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

Corollary 3.5: The range of T is the annihilator of the nullspace of T', i.e., $R_T = N_{T'}^\perp$.

Proof: By Theorem 2.5, $(R^\perp)^\perp = R$. So apply $\perp$ to both sides: $R_T^\perp = N_{T'} \implies R_T = N_{T'}^\perp$. $\square$

Exercise: Give a direct algebraic proof of $R_T^{\perp\perp} = N_{T'}^\perp$, much like the proof of Theorem 3.5.

Theorem 3.6: For any linear mapping $T : X \to U$, we have $\dim R_T = \dim R_{T'}$.

Proof: Apply Theorem 2.4 to $R_T \subseteq U$:

$$\dim R_T^\perp + \dim R_T = \dim U.$$

Now apply Theorem 3.2' to $T' : U' \to X'$:

$$\dim N_{T'} + \dim R_{T'} = \dim U'.$$

Now put these together, using:

$\dim U = \dim U'$ (Theorem 2.2)

$\dim R_T^\perp = \dim N_{T'}$ ($R_T^\perp = N_{T'}$ by Theorem 3.5)

$\Rightarrow \dim R_T = \dim R_{T'}$. $\square$

Corollary 3.6: Let $T: X \to U$ be linear with $\dim X = \dim U$. Then $\dim N_T = \dim N_{T'}$.

Proof: Apply Theorem 3.2 to T and T':

$\dim N_T = \dim X - \dim R_T$

$\dim N_{T'} = \dim U' - \dim R_{T'}$

By assumption, $\dim X' = \dim X = \dim U = \dim U'$.

By Theorem 3.6, $\dim R_T = \dim R_{T'}$. $\square$

In the "language of matrices," Theorem 3.6 says that the row rank and column rank of any matrix are equal.

Def: An endomorphism of a vector space X is a linear map from X to itself. We denote the set of endomorphisms of X by $\mathcal{L}(X,X)$, or $\mathrm{End}(X)$.

Remark: $\mathcal{L}(X,X)$ is a vector space, but we can also "multiply" vectors (composition). Thus, $\mathcal{L}(X,X)$ is an algebra.

$\mathcal{L}(X,X)$ is an <u>associative</u>, but not <u>commutative</u> algebra, with <u>unity</u>. The <u>unit</u> is the identity map I, where $Ix = x$.

The <u>zero map</u> 0 satisfies $0x = 0$.

$\mathcal{L}(X,X)$ contains <u>zero divisors</u>, i.e, pairs S, T such that $ST = 0$ but neither S nor T is zero.

<u>Prop</u>: If $A \in \mathcal{L}(X,X)$ is a left inverse of $B \in \mathcal{L}(X,X)$, that is, $AB = I$, then it is also a right inverse: $BA = I$.

<u>Proof</u>: Exercise.

The set of <u>invertible</u> elements of $\mathcal{L}(X,X)$ forms a <u>group</u> under multiplication, called the <u>general linear group</u>, and denoted $GL(n,K)$, where $n = \dim X$.

Every $S \in GL(n,K)$ defines a <u>similarly transformation</u> of $\mathcal{L}(X,X)$, sending $M \longmapsto SMS^{-1}$ for each $M \in \mathcal{L}(X,X)$.

We denote SMS^{-1} as M_S for short, and say that M and M_S are <u>similar</u>.

<u>Theorem 3.7</u>: Every similarity transform is an <u>automorphism</u> of $\mathcal{L}(X,X)$, that is, it maps sums to sums, products to products, and scalar multiples to scalar multiples. Said differently, it preserves the algebraic structure 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 $(M_S)_T := M_{TS}$. (often called the <u>inner automorphism group</u> of $GL(n,K)$).

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

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

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

We leave the details of checking the group axioms as an exercise. □

<u>Theorem 3.8</u>: Similarity is an equivalence relation, i.e, it is

(i) Reflexive: $M \sim M$.

(ii) Symmetric: $L \sim M \Rightarrow M \sim L$

(iii) Transitive: $L \sim M$ and $M \sim N \Rightarrow L \sim N$

(Here, $\sim$ denotes similarity.)

Proof: (i) $M_I = IMI^{-1} = M$ ✓

(ii) If $M = SLS^{-1}$, then $L = S^{-1}MS$. ✓

(iii) $L \sim M \Rightarrow M = SLS^{-1}$ for some $S \in GL(n,K)$.

$M \sim N \Rightarrow N = TMT^{-1}$ for some $T \in GL(n,K)$.

Thus, $N = T(SLS^{-1})T^{-1} = (TS)L(TS)^{-1}$. ✓

□

Theorem 3.9: If either A or B in $\mathcal{L}(X,X)$ is invertible, then AB and BA are similar.

Proof: Exercise.

Given any $A \in \mathcal{L}(X,X)$, and polynomial $p(s) = a_N s^N + \dots + a_1 s + a_0$, we can consider the polynomial $p(A) = a_N A^N + a_{N-1} A^{N-1} + \dots + a_1 A + a_0$. The set of all polynomials in A is a subalgebra of $\mathcal{L}(X,X)$, and is clearly commutative. More on this later (Re. Spectral theory).

Def: A linear mapping $P: X \to X$ is a projection if $P^2 = P$.

Examples: (12) Put $X = \mathbb{R}^n$ and $Px = (0, 0, a_3, \dots, a_n)$, where $x = (a_1, \dots, a_n)$. Here, P is a projection.

(13) Let $X = \{f: \mathbb{R} \to \mathbb{R}, \text{ continuous}\}$.

Take Pf to be the even part of f, i.e, $(Pf)(x) = \dfrac{f(x) + f(-x)}{2}$

Take Qf to be the odd part of f, i.e, $(Qf)(x) = \dfrac{f(x) - f(-x)}{2}$

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

Def: The commutator of two mappings $A, B \in \mathcal{L}(X,X)$ is $[A,B] := AB - BA$. Note that A and B commute iff $[A,B] = 0$.

Def: The rank of a linear mapping is the dimension of its range.