5374 Fall '22

Numerical Linear Algebra — Lecture 10

Today: A review / quick survey of linear algebra concepts and results

* The space of linear transformations
* Solvability of linear equations
* Inner products, orthogonality, adjoints, and orthogonal transformations

(see: first two or three "lectures" in Trefethen and Bau)

The space of linear transformation

Fix two dimensions n and m, and define

(when n and m are clear from context I will simply write this)

$$(\mathcal{L} =) \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) = \{ L : \mathbb{R}^n \to \mathbb{R}^m \mid L \text{ is linear} \}$$

More generally, if X and Y are normed vector spaces (and complete w.r.t. to their respective norms) we will consider

$$\mathcal{L}(X,Y) = \{ L : X \to Y \mid L \text{ is linear AND continuous} \}$$

Examples : (1) Any $m \times n$ matrix $A = (a_{ij})$ defines a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$

$$L_A(x) = A \cdot x ,$$

L_A is clearly linear from the properties of vector-matrix multiplication

(2) Let $X = C([0,1]) = Y$ (continuous, real valued function in $[0,1]$) (*).

$$L(f)(t) = \int_0^t f(s)\,ds , \quad \text{for } t \in [0,1].$$

Clearly if $f(t)$ is continuous in $[0,1]$, so will $L(f)(t)$

(*) $C([0,1])$ is defined as a normed vector space with the following norm

$$\left(\|f\|_{C([0,1])} \overset{\text{AKA}}{:=} \|f\|_{L^\infty} := \max_{0 \le t \le 1} |f(t)| \right)$$

That $L(f)$ is linear is immediate from the properties of the integral. Moreover, it is a

("Functional".)

continuous map. To see why, observe first that

$J(f) = \int_0^1 \sqrt{1+(f')^2}\,dt$ $\quad J : C([0,1]) \to \mathbb{R}$

function $f(x)$ $\quad f \quad C([0,1])$

$x \in [0,1] \quad x \quad [0,1]$

$$|Lf(t)| = \left| \int_0^t f(s)\,ds \right|$$

$$\leq \int_0^t |f(s)|\,ds$$

$$\leq \int_0^t \max_{0\leq s\leq 1} |f(s)|\,ds \quad , \quad 0 \leq t \leq 1$$

$$\leq \int_0^1 \max_{0\leq t\leq 1} |f(s)|\,ds \leq \max_{0\leq s\leq 1} |f(s)| = \|f\|_{L^\infty} \quad \forall t$$

$$\Longrightarrow \quad \max_{0\leq t\leq 1} |L(f)(t)| \leq \|f\|_{L^\infty}$$

$$\|L(f)\|_{L^\infty} \leq \|f\|_{L^\infty}$$

This implies $L : C([0,1]) \longrightarrow C([0,1])$ is continuous! Since

$$\|L(f_1) - L(f_2)\|_{L^\infty}$$

$$= \|L(f_1 - f_2)\|_{L^\infty} \quad \text{(by linearity)}$$

$$\leq \|f_1 - f_2\|_{L^\infty} \quad \text{(by the inequality above)}$$

so $\|L(f_1) - L(f_2)\|_{L^\infty} < \varepsilon$ if $\|f_1 - f_2\|_{L^\infty} < \varepsilon$

(3) Let $P_n = \{$ polynomials of degree $\leq n \}$

$= \{ p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \}$

Define $D : P_n \longrightarrow P_{n-1}$ as follows

$$D\left(a_n x^n + \cdots + a_1 x + a_0 \right)$$
$$= a_n n x^{n-1} + a_{n-1}(n-1) x^{n-2} + \cdots + 2a_2 x + a_1$$

(4) Let $T_n = \{$ trigonometric polynomials of degree $\leq n \}$

i.e. $T_n = \{ f(x) \mid f(x) = a_0 + \sum_{k=1}^{n} a_k \cos(kx) + b_k \sin(kx) \}$

Define $D : T_n \longrightarrow T_n$ by

$$D\left(a_0 + \sum_{k=1}^{n} a_k \cos(kx) + b_k \sin(kx) \right)$$
$$= \sum_{k=1}^{n} b_k k \cos(kx) - a_k k \sin(kx)$$

Linear transformations and bases

Suppose $L : X \longrightarrow Y$, where X and Y are finite dimensional. Concretely, we select some basis for X,

$$e_1, e_2, \ldots, e_n$$

and a basis for Y,

$$f_1, f_2, \ldots, f_m$$

Then given any $x \in X$, there must be (unique) coefficients $c_1, c_2, \ldots, c_n$ such that

$$x = c_1 e_1 + \cdots + c_n e_n$$

If I am given a linear transformation L from X to Y, I can compute $L(x)$ based alone on the values of L on the basis vectors $e_1, \ldots, e_n$,

$$\begin{aligned} L(x) &= L(c_1 e_1 + \cdots + c_n e_n) \\ &= c_1 L(e_1) + \cdots + c_n L(e_n) \end{aligned}$$

Thus, L is completely determined by its values on a basis of X

(at the matrix level, this corresponds to the statement that matrix-vector multiplication amounts to linear combination of the matrix columns using the entries of the vectors)

On the other hand, using the basis in Y we can write every vector $L(e_j)$ $(j=1,\dots,n)$ as follows

$$L(e_j) = a_{1j} f_1 + a_{2j} f_2 + \cdots + a_{mj} f_m$$

(i.e. $L(e_1) = a_{11} f_1 + a_{21} f_2 + \cdots + a_{m1} f_m$

$L(e_2) = a_{12} f_1 + a_{22} f_2 + \cdots + a_{m2} f_m$

$\vdots$ etc)

Substituting the formulas for $L(e_j)$ in the expression for $L(x)$, we see that

$$\begin{aligned} L(x) &= c_1 L(e_1) + \cdots + c_n L(e_n) \\ &= \sum_{j=1}^{n} c_j L(e_j) \\ &= \sum_{j=1}^{n} \sum_{i=1}^{m} c_j a_{ij} f_i \end{aligned}$$

Now, if I exchange the summation order,

$$L(x) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} c_j a_{ij} \right) f_i \qquad (*)$$

What does this all mean? Any vector

in X can be codified as a unique vector in $\mathbb{R}^n$

$$\begin{array}{ccc} \mathbb{R}^n & \longleftrightarrow & X \\ \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} & & c_1 e_1 + \cdots + c_n e_n \end{array}$$

and this encoding happen via the basis $e_1, \dots, e_n$

Likewise for $\mathbb{R}^m$ and Y

$$\begin{array}{ccc} \mathbb{R}^m & \longleftrightarrow & Y \\ \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} & & b_1 f_1 + \cdots + b_m f_m \end{array}$$

The formula $(*)$ shows that if

$$L(x) = y$$

and

$$x = c_1 e_1 + \cdots + c_n e_n$$
$$y = b_1 f_1 + \cdots + b_m f_m$$

Then $\bar{c} \in \mathbb{R}^n$ and $\bar{b} \in \mathbb{R}^m$ are related by

$$\bar{b} = A\,\bar{c}$$

where A is the $m \times n$ matrix with entries a_{ij} as defined in $(*)$.

This shows how if X is n-dimensional and Y is m-dimensional then there is a clear bijection

$$\mathcal{L}(X, Y) \longleftrightarrow M_{m \times n} := \{ m \times n \text{ matrices} \}$$

The rank-nullity theorem

Given $L \in \mathcal{L}(X, Y)$ (henceforth X, Y will be of dimension n and m, respectively) we define two important vector spaces:

Image of L

$$\operatorname{Im}(L) = \{ y \in Y \mid \exists\, x \in X \text{ s.t. } L(x) = y \}$$

Kernel of L

$$\operatorname{Ker}(L) = \{ x \in X \mid L(x) = 0 \in Y \}$$

Clearly $\operatorname{Im}(L) \subseteq Y$, $\operatorname{Ker}(L) \subseteq X$.

The dimension of $\operatorname{Im}(L)$ is called **the rank of L** and it is denoted $\operatorname{rank}(L)$. While the dimension of $\operatorname{Ker}(L)$ is called **the nullity of L**, denoted $\operatorname{nullity}(L)$.

Remarks. That the map $L: X \to Y$ is surjective is the same as saying
$$\operatorname{Im}(L) = Y.$$
On the other hand, if $x_1, x_2 \in X$ and
$$L(x_1) = L(x_2)$$
By linearity, this means that
$$L(x_1 - x_2) = 0$$
so $x_1 - x_2 \in \operatorname{Ker}(L)$. It follows that L being injective is the same as
$$\operatorname{Ker}(L) = \{0\}.$$

There is a useful and elementary relationship between the dimension n of X, and the rank and nullity of a linear map $L: X \to Y$.

Theorem ("the rank-nullity theorem")

$$\boxed{n = \operatorname{rank}(L) + \operatorname{nullity}(L)}$$

Proof. First, let us build a basis for $\operatorname{Ker}(L)$,
$$e_1, \dots, e_k, \qquad k = \dim(\operatorname{Ker}(L)) = \operatorname{nullity}(L)$$

If this is a basis of X, then that means $\ker(L) = X$, so $L = 0$, and then $\operatorname{rank}(L) = 0$ and claim

$$n = 0 + n \quad \checkmark$$

Otherwise, we can complete this to a basis of X by adding vectors $e_{k+1}, \dots, e_n$.

I claim if $y \in \operatorname{Im}(L)$, then y can be written as

$$y = L\left(\sum_{j=k+1}^{n} c_j e_j\right)$$

why? well by definition $y \in \operatorname{Im}(L) \Rightarrow \exists\, x \in X$ s.t. $y = L(x)$, but

$$x = \sum_{j=1}^{n} c_j e_j \quad \text{for some } c_1, \dots, c_n$$

but

$$L(x) = \sum_{j=1}^{n} c_j L(e_j), \quad L(e_j) = 0 \text{ for } j = 1, \dots, k$$

$$= \sum_{j=k+1}^{n} c_j L(e_j)$$

$$= L\left(\sum_{j=k+1}^{n} c_j e_j\right),$$

As we claimed. Moreover, since $\{e_j\}_{j=k+1}^{n}$ are linearly independent from the basis of $\ker(L)$, it follows that $\{L(e_j)\}_{j=k+1}^{n}$ are a basis of $\operatorname{Im}(L)$, so

$$\operatorname{rank}(L) = \dim(\operatorname{Im}(L)) = n-k$$

$$\Rightarrow \quad n = (n-k) + k$$
$$= \operatorname{rank}(L) + \operatorname{nullity}(L)$$

$\blacksquare$

Corollary:

If $m<n$, there are no injective linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$.

If $m>n$, there are no surjective linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$.

Corollary:

If $m=n$, $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is injective if and only if it is surjective.

In particular, $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is invertible as soon as $\operatorname{Im}(L) = \mathbb{R}^n$ or $\ker(L) = \{0\}$.