5

Direct Sum, Rank-Nullity Theorem Affine Map

★ The Rank-Nullity Theorem; Grassmann's Relation

Theorem 5.8. *(Rank-nullity theorem) Let $f : E \to F$ be a linear map with finite image. For any choice of a basis $(f_1, \ldots, f_r)$ of $\mathrm{Im}\, f$, let $(u_1, \ldots, u_r)$ be any vectors in E such that $f_i = f(u_i)$, for $i = 1, \ldots, r$. If $s : \mathrm{Im}\, f \to E$ is the unique linear map defined by $s(f_i) = u_i$, for $i = 1, \ldots, r$, then s is injective, $f \circ s = \mathrm{id}$, and we have a direct sum*

$$E = \mathrm{Ker}\, f \oplus \mathrm{Im}\, s$$

as illustrated by the following diagram:

$$\mathrm{Ker}\, f \longrightarrow E = \mathrm{Ker}\, f \oplus \mathrm{Im}\, s \underset{s}{\overset{f}{\rightleftarrows}} \mathrm{Im}\, f \subseteq F.$$

See Figure 5.2. As a consequence, if E is finite-dimensional, then

$$\dim(E) = \dim(\mathrm{Ker}\, f) + \dim(\mathrm{Im}\, f) = \dim(\mathrm{Ker}\, f) + \mathrm{rk}(f).$$

⇛ The dimension dim(Ker f) of the kernel of a linear map f is called the nullity of f . {Definition}

Proposition 5.9. *Given a vector space E, if U and V are any two finite-dimensional subspaces of E, then*

$$\dim(U) + \dim(V) = \dim(U + V) + \dim(U \cap V),$$

an equation known as Grassmann's relation.

⇛ The Grassmann relation can be very useful to figure out whether two subspace have a nontrivial intersection in spaces of dimension > 3.

Proposition 5.10. *If $U_1, \ldots, U_p$ are any subspaces of a finite dimensional vector space E, then*

$$\dim(U_1 + \cdots + U_p) \leq \dim(U_1) + \cdots + \dim(U_p),$$

and

$$\dim(U_1 + \cdots + U_p) = \dim(U_1) + \cdots + \dim(U_p)$$

iff the U_is form a direct sum $U_1 \oplus \cdots \oplus U_p$.

Proposition 5.11. *Let E and F be two vector spaces with the same finite dimension $\dim(E) = \dim(F) = n$. For every linear map $f : E \to F$, the following properties are equivalent:*

(a) f is bijective.

(b) f is surjective.

(c) f is injective.

(d) $\mathrm{Ker}\, f = (0)$.

Proposition 5.12. *Let E and F be vector spaces, and let $f\colon E \to F$ be a linear map. If $f\colon E \to F$ is injective, then there is a surjective linear map $r\colon F \to E$ called a retraction, such that $r \circ f = \mathrm{id}_E$. See Figure 5.3. If $f\colon E \to F$ is surjective, then there is an injective linear map $s\colon F \to E$ called a section, such that $f \circ s = \mathrm{id}_F$. See Figure 5.2.*

Proposition 5.13. *Given a linear map $f\colon E \to F$, the following properties hold:*

(i) $\mathrm{rk}(f) + \dim(\mathrm{Ker}\, f) = \dim(E)$.

(ii) $\mathrm{rk}(f) \leq \min(\dim(E), \dim(F))$.

Definition 5.5. Given a $m \times n$-matrix $A = (a_{i\,j})$, the *rank* $\mathrm{rk}(A)$ of the matrix A is the maximum number of linearly independent columns of A (viewed as vectors in $\mathbb{R}^m$).

★ Affine Map

⇒ Every linear map f must send the zero vector to the zero vector; that is:

→ f(0) = 0

→ Yet for any fixed nonzero vector u ∈ E (where E is any vector space), the function t_u given by

→ $t_u(x) = x + u, \quad \text{for all} \quad x \in E$

→ shows up in practice (for example, in robotics).

→ Functions of this type are called translations.

→ They are not linear for u != 0, since t_u (0) = 0 + u = u.

→ More generally, functions combining linear maps and translations occur naturally in many applications (robotics, computer vision, etc.)

⇒ For any vector space E, given any family (u_i) i∈I of vectors u_i ∈ E, an affine combination of the family (u_i) i∈I is an expression of the form

$$\boxed{\sum_{i \in I} \lambda_i u_i \quad \text{with} \quad \sum_{i \in I} \lambda_i = 1,}$$

where (λ_i) i∈I is a family of scalars.

⇒ Affine combinations are also called barycentric combinations.

Proposition 5.14. *For any two vector spaces E and F, given any function $f\colon E \to F$ defined such that*

$$f(x) = h(x) + b, \quad \text{for all} \quad x \in E,$$

where $h\colon E \to F$ is a linear map and b is some fixed vector in F, for every affine combination $\sum_{i \in I} \lambda_i u_i$ (with $\sum_{i \in I} \lambda_i = 1$), we have

$$f\Big(\sum_{i \in I} \lambda_i u_i\Big) = \sum_{i \in I} \lambda_i f(u_i).$$

In other words, f preserves affine combinations.

Proposition 5.15. *For any two vector spaces E and F, let $f\colon E \to F$ be any function that preserves affine combinations, i.e., for every affine combination $\sum_{i\in I} \lambda_i u_i$ (with $\sum_{i\in I} \lambda_i = 1$), we have*

$$f\Big(\sum_{i\in I} \lambda_i u_i\Big) = \sum_{i\in I} \lambda_i f(u_i).$$

Then for any $a \in E$, the function $h\colon E \to F$ given by

$$h(x) = f(a + x) - f(a)$$

is a linear map independent of a, and

$$f(a + x) = h(x) + f(a), \quad \text{for all} \quad x \in E.$$

In particular, for $a = 0$, if we let $c = f(0)$, then

$$f(x) = h(x) + c, \quad \text{for all} \quad x \in E.$$

⇛ We should think of a as 'a' chosen origin in E.

⇛ The function f maps the origin 'a' in E to the origin f(a) in F .

⇛ Also, since

$$f(x) = h(x) + c, \quad \text{for all} \quad x \in E$$

for some fixed vector c ∈ F , we see that f is the composition of the linear map h with the translation t_c (in F).

⇛ The unique linear map h as above is called the linear map associated with f , and it is sometimes denoted by $\overrightarrow{f}$.

Definition 5.6. For any two vector spaces E and F, a function $f\colon E \to F$ is an *affine map* if f preserves affine combinations, *i.e.*, for every affine combination $\sum_{i\in I} \lambda_i u_i$ (with $\sum_{i\in I} \lambda_i = 1$), we have

$$f\Big(\sum_{i\in I} \lambda_i u_i\Big) = \sum_{i\in I} \lambda_i f(u_i).$$

Equivalently, a function $f\colon E \to F$ is an *affine map* if there is some linear map $h\colon E \to F$ (also denoted by $\overrightarrow{f}$) and some fixed vector $c \in F$ such that

$$f(x) = h(x) + c, \quad \text{for all} \quad x \in E.$$

Definition 5.7. An *affine space* is either the degenerate space reduced to the empty set, or a triple $\langle E, \overrightarrow{E}, +\rangle$ consisting of a nonempty set E (of *points*), a vector space $\overrightarrow{E}$ (of *translations*, or *free vectors*), and an action $+\colon E \times \overrightarrow{E} \to E$, satisfying the following conditions.

(A1) $a + 0 = a$, for every $a \in E$.

(A2) $(a + u) + v = a + (u + v)$, for every $a \in E$, and every $u, v \in \overrightarrow{E}$.

(A3) For any two points $a, b \in E$, there is a unique $u \in \overrightarrow{E}$ such that $a + u = b$.

The unique vector $u \in \overrightarrow{E}$ such that $a + u = b$ is denoted by $\overrightarrow{ab}$, or sometimes by $\mathbf{ab}$, or even by $b - a$. Thus, we also write

$$b = a + \overrightarrow{ab}$$

(or $b = a + \mathbf{ab}$, or even $b = a + (b - a)$).

⇛ If E and F are finite dimensional vector spaces with dim(E) = n and dim(F) = m, then it is useful to represent an affine map with respect to bases in E in F .

⇛ There is a standard trick to do this which amounts to viewing an affine map as a linear map between spaces of dimension n + 1 and m + 1.

⇛ Let (u 1 , . . . , u n) be a basis of E, (v 1 , . . . , v m) be a basis of F , and let a ∈ E and b ∈ F be any two fixed vectors viewed as origins.

⇛ Our affine map f has the property that if v = f (u), then

$$v - b = f(a + u - a) - b = f(a) - b + h(u - a),$$

⇛ If we let y = v − b, x = u − a, and d = f (a) − b, then

$$y = h(x) + d, \quad x \in E.$$

⇛ Over the basis $\mathcal{U} = (u_1, \ldots, u_n)$, we write

$$x = x_1 u_1 + \cdots + x_n u_n,$$

and over the basis $\mathcal{V} = (v_1, \ldots, v_m)$, we write

$$\begin{aligned} y &= y_1 v_1 + \cdots + y_m v_m, \\ d &= d_1 v_1 + \cdots + d_m v_m. \end{aligned}$$

⇛ If we let A be the m × n matrix representing the linear map h, that is, the jth column of A consists of the coordinates of h(u_j) over the basis (v 1 , . . . , v m), then we can write

$$y_{\mathcal{V}} = A x_{\mathcal{U}} + d_{\mathcal{V}}.$$

where $x_{\mathcal{U}} = (x_1, \ldots, x_n)^\top$, $y_{\mathcal{V}} = (y_1, \ldots, y_m)^\top$, and $d_{\mathcal{V}} = (d_1, \ldots, d_m)^\top$.

⇛ The above is the matrix representation of our affine map f with respect to (a, (u 1 , . . . , u n)) and (b, (v 1 , . . . , v m)).

⇛ When $E = F$, if there is some $a \in E$ such that $f(a) = a$ (a is a *fixed point* of f), then we can pick $b = a$. Then because $f(a) = a$, we get

$$v = f(u) = f(a + u - a) = f(a) + h(u - a) = a + h(u - a),$$

that is

$$v - a = h(u - a).$$

With respect to the new origin a, if we define x and y by

$$\begin{aligned} x &= u - a \\ y &= v - a, \end{aligned}$$

then we get

$$y = h(x).$$

⇛ Therefore, f really behaves like a linear map, but *with respect to the new origin* a (not the *standard origin* 0). This is the case of a rotation around an axis that does not pass through the origin.

⇛ A pair (a, (u_1 , . . . , u_n)) where (u_1 , . . . , u_n) is a basis of E and a is an origin chosen in E is called an affine frame.

We now describe the trick which allows us to incorporate the translation part d into the matrix A.

We define the $(m+1) \times (n+1)$ matrix A' obtained by first adding d as the $(n+1)$th column and then $(\underbrace{0, \ldots, 0}_{n}, 1)$ as the $(m+1)$th row:

$$A' = \begin{pmatrix} A & d \\ 0_n & 1 \end{pmatrix}.$$

It is clear that

$$\begin{pmatrix} y \\ 1 \end{pmatrix} = \begin{pmatrix} A & d \\ 0_n & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix}$$

iff

$$y = Ax + d.$$

This amounts to considering a point $x \in \mathbb{R}^n$ as a point $(x, 1)$ *in the (affine) hyperplane* H_{n+1} *in* $\mathbb{R}^{n+1}$ *of equation* $x_{n+1} = 1$.

The idea of considering $\mathbb{R}^n$ as an hyperplane in $\mathbb{R}^{n+1}$ can be used to define *projective maps*.