LA

Linear Algebra

LA.2 PROPERTIES OF LINEAR TRANSFORMATIONS

Consider the linear map $\mathcal{A} : \mathcal{R}^m \rightarrow \mathcal{R}^n$ described, with reference to specific bases in $\mathcal{R}^m$ and $\mathcal{R}^n$, by the $n \times m$ matrix A. The following notations refer both to maps and associated matrices:

- im $A := \{y : y = Ax,\ x \in \mathcal{R}^m\}$ denotes the *image* or *range* of A. The dimension of im A, i.e. the *rank* of A is indicated with $\rho(A)$.
- ker $A := \{x : Ax = 0,\ x \in \mathcal{R}^m\}$ denotes the *kernel* or *null space* of A. The dimension of ker A, i.e. the *nullity* of A, is indicated with $\nu(A)$.

The following lemmas describe some geometric properties of linear transformations.

Lemma LA.2.1 – im A and ker A are subspaces of $\mathcal{R}^n$ and $\mathcal{R}^m$.

Lemma LA.2.2 – The orthogonal complements of im A and ker A satisfy the following relations:

$$\left(\text{im } A\right)^{\perp} = \ker A^T \qquad \text{(LA.2.1)}$$
$$\left(\ker A\right)^{\perp} = \text{im } A^T. \qquad \text{(LA.2.2)}$$

Lemma LA.2.3 – $\rho(A) + \nu(A) = m$.

From properties (LA.2.1) and (LA.2.2) it follows that the Euclidean spaces $\mathcal{R}^m$ and $\mathcal{R}^n$ can be decomposed into the following direct sums:

$$\mathcal{R}^m = \text{im } A^T \oplus \ker A \qquad \text{(LA.2.3)}$$
$$\mathcal{R}^n = \text{im } A \oplus \ker A^T. \qquad \text{(LA.2.4)}$$

From (LA.2.2) and Lemma (LA.2.3), $\rho(A^T) + \nu(A) = m$, so that $\rho(A) = \rho(A^T)$; thus if A is not square, $\nu(A) \neq \nu(A^T)$. A basic property of linear maps can now be stated:

Theorem LA.2.1 – The restriction of the linear transformation $\mathcal{A} : \mathcal{R}^m \to \mathcal{R}^n$ to $\text{im}\, A^T \to \text{im}\, A$ is one to one.

Remark LA.2.1 – Theorem LA.2.1 allows an interesting geometric interpretation of some properties of linear maps: any vector $x \in \mathcal{R}^m$ can be univocally expressed as $x = x_1 + x_2$ with $x_1 \in \text{im}\, A^T$ and $x_2 \in \ker A$, so that $Ax = A(x_1 + x_2) = Ax_1$. As a consequence, the linear transformation decribed by A can be decomposed into the orthogonal projection on $\text{im}\, A^T$ and into the one to one map $A_r : \text{im}\, A^T \to \text{im}\, A$ defined as $A_r x = Ax$ for all $x \in \text{im}\, A^T$.

Consider now the system of equations

$$Ax = y \tag{LA.2.5}$$

where $x \in \mathcal{R}^m$ and $y \in \mathcal{R}^n$. Previous considerations establish that this system admits solutions only when $y \in \text{im}\, A$ and that this solution is unique if and only if $\ker A = \{0\}$.

The concept of invertibility of linear maps can be extended introducing the pseudoinverse of a linear transformation as follows.

Definition LA.2.1 – The pseudoinverse $\mathcal{A}^+$ of $\mathcal{A}$ is a linear transformation $\mathcal{A}^+ : \mathcal{R}^n \to \mathcal{R}^m$ that performs the orthogonal projection on $\text{im}\, A$ and the one to one transformation $A_r^{-1} : \text{im}\, A \to \text{im}\, A^T$.

From Definition LA.2.1 it follows that

$$x \in \text{im}\, A^T \Rightarrow A^+Ax = x \tag{LA.2.6}$$
$$x \in \ker A^T \Rightarrow A^+x = 0 \tag{LA.2.7}$$
$$\ker A^+ = \ker A^T \tag{LA.2.8}$$
$$\text{im}\, A^+ = \text{im}\, A^T. \tag{LA.2.9}$$

It can be noted that the pseudoinverse of a matrix is unique and that $(A^+)^+ = A$. Moreover when $\mathcal{A}$ is one to one, A is square and $A^+ = A^{-1}$.

The pseudoinverse of a matrix finds useful applications in solving the system of linear equations (LA.2.5) or in approximating its solution. Because of the properties of orthogonal projections it can be shown that when system (LA.2.5) admits infinite solutions, i.e. when $\ker A \neq \{0\}$, A^+y is the solution with minimal Euclidean norm. When, on the contrary, no solution exists, i.e. $y \notin \text{im}\, A$, $\hat{x} = A^+y$ is the pseudosolution with minimal Euclidean norm minimizing the Euclidean norm of the error $y - A\hat{x}$. It is also worth observing that when $y \in \text{im}\, A$, the set of all solutions of system (LA.2.5) is the linear variety

$$\mathcal{X} = \{A^+y\} + \ker A. \tag{LA.2.10}$$

Other properties of the pseudoinverse are the following:

- If $\rho(A) = m$, $A^+ = (A^T A)^{-1} A^T$ (LA.2.11)
- If $\rho(A) = n$, $A^+ = A^T (AA^T)^{-1}$ (LA.2.12)
- $A^+ A$ is the orthogonal projector of $\mathcal{R}^m$ onto im A^T (LA.2.13)
- $I - A^+ A$ is the orthogonal projector of $\mathcal{R}^m$ onto ker A (LA.2.14)
- AA^+ is the orthogonal projector of $\mathcal{R}^n$ onto im A (LA.2.15)
- $I - AA^+$ is the orthogonal projector of $\mathcal{R}^n$ onto ker A^T. (LA.2.16)