Lecture 2. Kernel and Image of a Linear Map
September 2022

Similarity

A particular case of Change-of-Basis Theorem (see Lecture 1) is so important that it is worth stating separately. It corresponds to the case in which $m = n$ and $B' = B$.

COROLLARY 2.1 (Similarity Theorem) Suppose that $f : V \to V$ is a linear operator on V and that $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ is a basis of V. Let A be the matrix corresponding to f with respect to basis E. Then, with respect to the basis B, the operator f is represented by the matrix

$$A_{B,B} = P^{-1}AP,$$

where P is a transition matrix from E to B.

The relationship between the matrices $A_{B,B}$ and A is a central one in the theory of linear algebra. The matrix $A_{B,B}$ performs the same linear map as the matrix A, but $A_{B,B}$ describes it in terms of the basis B. This effect inspires the following definition.

DEFINITION 2.2 (Similarity) We say that the square matrix C is **similar** to the matrix A if there is an invertible matrix P such that $C = P^{-1}AP$.

Similarity defines an equivalence relation on matrices. Recall that an equivalence relation satisfies three properties; it is reflexive, symmetric and transitive. For similarity, this means:

- a matrix A is similar to itself (reflexive),
- if C is similar to A, then A is similar to C (symmetric), and
- if D is similar to C, and C to A, then D is similar to A (transitive).

Because the relationship is symmetric, we usually just say that A and C are **similar matrices**, meaning one is similar to the other, and we can express this either as $C = P^{-1}AP$ or $A = Q^{-1}CQ$ for invertible matrices P and Q (in which case $Q = P^{-1}$).

Similar matrices share many properties, some of which are collected in the next theorem.

THEOREM 2.3 If A and C are similar $n \times n$-matrices, then A and C have the same determinant, rank, and trace.

Proof. 1) Let $C = P^{-1}AP$ for some invertible matrix P. Then we have

$$\det C = \det(P^{-1}) \det A \det P = (1/\det P) \det A \det P = \det A.$$

2) Since $\operatorname{rank}(AB) \leqslant \operatorname{rank} A$, it follows that $\operatorname{rank} C \leqslant \operatorname{rank} A$ and $\operatorname{rank} A \leqslant \operatorname{rank} C$, so they are equal.

3) Write $A = [a_{ij}]$ and $B = [b_{ij}]$. For each i, the (i,i)-entry of the matrix AB is $d_i = \sum_j a_{ij}b_{ji}$. Hence

$$\operatorname{tr}(AB) = d_1 + d_2 + \cdots + d_n = \sum_i \left(\sum_j a_{ij}b_{ji} \right) = \operatorname{tr}(BA).$$

Therefore, $\operatorname{tr} C = \operatorname{tr}(P^{-1}AP) = \operatorname{tr}(APP^{-1}) = \operatorname{tr} A$. □

Definition of Kernel and Image

Now we will introduce two important subspaces associated with a linear map $f : V \to W$.

DEFINITION 2.4 The **kernel** of f (denoted $\operatorname{Ker} f$) and the **image** of f (denoted $\operatorname{Im} f$ or $f(V)$) are defined by

$$\operatorname{Ker} f = \{\mathbf{v} \in V |\ f(\mathbf{v}) = \mathbf{0}\}, \qquad \operatorname{Im} f = f(V) = \{\mathbf{w} \in V |\ \exists \mathbf{v} \in V : \mathbf{w} = f(\mathbf{v})\}.$$

The kernel of f is often called the **nullspace** of f. It consists of all vectors $\mathbf{v}$ in V satisfying the condition $f(\mathbf{v}) = \mathbf{0}$. The image of f is often called the **range** of f and consists of all vectors $\mathbf{w}$ in W of the form $\mathbf{w} = f(\mathbf{v})$ for some $\mathbf{v}$ in V.

THEOREM 2.5 Let $f : V \to W$ be a linear map.

1. $\operatorname{Ker} f$ is a subspace of V.
2. $\operatorname{Im} f$ is a subspace of W.

Proof. The fact that $f(\mathbf{0}) = \mathbf{0}$ shows that $\operatorname{Ker} f$ and $\operatorname{Im} f$ contain the zero vector of V and W respectively.

1. If $\mathbf{v}$ and $\mathbf{v}_1$ lie in $\operatorname{Ker} f$, then $f(\mathbf{v}) = \mathbf{0} = f(\mathbf{v}_1)$, so

$$\begin{aligned} f(\mathbf{v} + \mathbf{v}_1) &= f(\mathbf{v}) + f(\mathbf{v}_1) = \mathbf{0} + \mathbf{0} = \mathbf{0} \\ f(k\mathbf{v}) &= kf(\mathbf{v}) = k \cdot \mathbf{0} = \mathbf{0} \quad \text{for all } k \in \mathbb{R}. \end{aligned}$$

Hence $\mathbf{v}+\mathbf{v}_1$ and $k\mathbf{v}$ lie in $\operatorname{Ker} f$ (they satisfy the required condition), so $\operatorname{Ker} f$ is a subspace of V since it is closed under addition and scalar multiplication.

2. If $\mathbf{w}$ and $\mathbf{w}_1$ lie in $\operatorname{Im} f$, write $\mathbf{w} = f(\mathbf{v})$ and $\mathbf{w}_1 = f(\mathbf{v}_1)$ where $\mathbf{v}, \mathbf{v}_1 \in V$. Then

$$\begin{aligned} \mathbf{w} + \mathbf{w}_1 &= f(\mathbf{v}) + f(\mathbf{v}_1) = f(\mathbf{v} + \mathbf{v}_1) \\ k\mathbf{w} &= kf(\mathbf{v}) = f(k\mathbf{v}) \quad \text{for all } k \in \mathbb{R}. \end{aligned}$$

Hence $\mathbf{w}+\mathbf{w}_1$ and $k\mathbf{w}$ both lie in $\operatorname{Im} f$ (they have the required form), so $\operatorname{Im} f$ is a subspace of W. □

DEFINITION 2.6 Given a linear map $f : V \to W$:

- $\dim(\operatorname{Ker} f)$ is called the **nullity** of f and denoted as $\operatorname{nullity} f$
- $\dim(\operatorname{Im} f)$ is called the **rank** of f and denoted as $\operatorname{rank} f$

The rank of a matrix A was defined earlier to be the dimension of $\operatorname{col} A$, the column space of A. The two usages of the word *rank* are consistent in the following sense.

THEOREM 2.7 Let f_A be a matrix map induced by an $m \times n$ matrix A. Then

1. $\operatorname{Im} f_A = \operatorname{col} A$;
2. $\operatorname{rank} f_A = \operatorname{rank} A$.

Proof. Since $f_A(\mathbf{x}) = A\mathbf{x}$, then

$$\operatorname{Im} f_A = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\} = \{x_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n \mid x_i \in \mathbb{R}\} = \operatorname{col} A.$$

Therefore,

$$\operatorname{rank} A = \dim(\operatorname{col} A) = \dim(\operatorname{row} A). \quad \square$$

One-to-One and Onto Maps

DEFINITION 2.8 Let $f : V \to W$ be a linear map.

- f is said to be **onto** (or **surjective**) if $\operatorname{Im} f = W$.
- f is said to be **one-to-one** (or **injective**) if $f(\mathbf{v}) = f(\mathbf{v}_1)$ implies $\mathbf{v} = \mathbf{v}_1$.

EXAMPLE 2.9 The identity map $\operatorname{Id} : V \to V$ is both one-to-one and onto for any vector space V.

A vector $\mathbf{w}$ in W is said to be **hit** by f if $\mathbf{w} = f(\mathbf{v})$ for some $\mathbf{v}$ in V. Then f is onto if every vector in W is hit at least once, and f is one-to-one if no element of W gets hit twice. Clearly the onto maps f are those for which $\operatorname{Im} f = W$ is as large a subspace of W as possible. By contrast, the next theorem shows that the one-to-one maps f are the ones with $\operatorname{Ker} f$ as small a subspace of V as possible.

THEOREM 2.10 If $f : V \to W$ is a linear map, then f is one-to-one if and only if $\operatorname{Ker} f = \{\mathbf{0}\}$.

Proof. If f is one-to-one, let $\mathbf{v}$ be any vector in $\operatorname{Ker} f$. Then $f(\mathbf{v}) = \mathbf{0}$, so $f(\mathbf{v}) = f(\mathbf{0})$. Hence $\mathbf{v} = \mathbf{0}$ because f is one-to-one. Hence $\operatorname{Ker} f = \{\mathbf{0}\}$.

Conversely, assume that $\operatorname{Ker} f = \{\mathbf{0}\}$ and let $f(\mathbf{v}) = f(\mathbf{v}_1)$ with $\mathbf{v}$ and $\mathbf{v}_1$ in V. Then $f(\mathbf{v} - \mathbf{v}_1) = f(\mathbf{v}) - f(\mathbf{v}_1) = \mathbf{0}$, so $\mathbf{v} - \mathbf{v}_1$ lies in $\operatorname{Ker} f = \{\mathbf{0}\}$. This means that $\mathbf{v} - \mathbf{v}_1 = \mathbf{0}$, so $\mathbf{v} = \mathbf{v}_1$, proving that f is one-to-one. $\square$

EXAMPLE 2.11 Consider the linear maps

- $f : \mathbb{R}^3 \to \mathbb{R}^2$ given by $f(x, y, z) = (x + y + z, x - y)$
- $g : \mathbb{R}^2 \to \mathbb{R}^3$ given by $g(x, y) = (x, x + y, x - y)$

Show that f is onto but not one-to-one, whereas g is one-to-one but not onto.

Solution. The verification that f and g are linear is omitted.

Map f is not one-to-one because $(1, 1, -2)$ lies in $\operatorname{Ker} f$. But every element (s, t) in $\mathbb{R}^2$ lies in $\operatorname{Im} f$ because $(s, t) = (x+y+z, x-y) = f(x, y, z)$ for some x, y, and z (for example, $x = \frac{1}{2}(s + t)$, $y = \frac{1}{2}(s - t)$, and $z = 0$). Hence f is onto.

Map g is one-to-one because

$$\operatorname{Ker} g = \{(x, y) \mid x + y = x - y = x = 0\} = \{(0, 0)\}.$$

However, it is not onto. For example $(1, 0, 0)$ does not lie in $\operatorname{Im} g$ because if $(1, 0, 0) = (x, x+y, x-y)$ for some x and y, then $x+y = 0 = x-y$ and $x = 1$, which is contradictory.

The next theorem gives conditions under which a linear map $f : V \to W$ of finite-dimensional spaces is onto or one-to-one.

THEOREM 2.12 Let $f : V \to W$ be a linear map of finite-dimensional spaces, which is represented by an $m \times n$-matrix A with respect to some bases of V and W, i.e., $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x}$ in V. Then the following statements hold.

1. f is onto if and only if $\operatorname{rank} A = m$.
2. f is one-to-one if and only if $\operatorname{rank} A = n$.

Proof. 1. We have that $\operatorname{Im} f$ is the column space of A, so f is onto if and only if the column space of A is $\mathbb{R}^m$. Because the rank of A is the dimension of the column space, this holds if and only if $\operatorname{rank} A = m$.

2. $\operatorname{Ker} f = \{\mathbf{x} \in \mathbb{R}^n |\ A\mathbf{x} = \mathbf{0}\}$, so f is one-to-one if and only if $A\mathbf{x} = \mathbf{0}$ implies $\mathbf{x} = \mathbf{0}$. This is equivalent to $\operatorname{rank} A = n$. $\square$