Lecture 10: Linear Transformations and Subspaces

February 15, 2024

Lecture 10: Linear Transformations and Subspaces

Linear Transformations

Linear transformations are the primary functions between vector spaces that are of interest in linear algebra. They are special because they cooperate with the algebraic structure.

Definition Let m and n be positive integers. A linear transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a function $T : \mathbb{R}^n \to \mathbb{R}^m$ that satisfies

- $T(u + v) = T(u) + T(v)$, and
- $T(\lambda v) = \lambda T(v)$,

for all u and v in $\mathbb{R}^n$, and all scalars $\lambda \in \mathbb{R}$.

The Matrix of a Linear Transformation

Suppose that $T : \mathbb{R}^3 \to \mathbb{R}^2$ is a linear transformation, with

$$T\begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}2\\-3\end{bmatrix},\ \begin{bmatrix}0\\1\\0\end{bmatrix} = \begin{bmatrix}1\\4\end{bmatrix},\ T\begin{bmatrix}0\\0\\1\end{bmatrix} = \begin{bmatrix}-6\\7\end{bmatrix}$$

Then for the vector in $\mathbb{R}^3$ with any entries a, b, c

$$T\begin{bmatrix}a\\b\\c\end{bmatrix} = aT\begin{bmatrix}1\\0\\0\end{bmatrix} + bT\begin{bmatrix}0\\1\\0\end{bmatrix} + cT\begin{bmatrix}0\\0\\1\end{bmatrix} = \underbrace{\begin{bmatrix}2 & 1 & -6\\-3 & 4 & 7\end{bmatrix}}_{M_T}\begin{bmatrix}a\\b\\c\end{bmatrix}.$$

and the 2×3 matrix M_T is called the (standard) matrix of A.

The matrix of a linear transformation

- A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is represented by a $m \times n$ matrix M_T. The columns of M_T are the images under T of the standard basis vectors $e_1, \dots, e_n$.
- If v is any vector in $\mathbb{R}^n$, we can calculate $T(v)$ by multiplying the column vector v on the left by the matrix M_T. Matrix-vector multiplication is evaluating linear transformations.
- On the other hand, if A is any $m \times n$ matrix, then A determines a linear transformation $\mathbb{R}^n \to \mathbb{R}^m$ by $v \to Av$, for $v \in \mathbb{R}^n$. So, in a sense, matrices are linear transformations.
- Examples of linear transformations include rotations, reflections and scaling, but not translations.
- If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then in order to evaluate T at any point/vector, we only need mn pieces of information, just the m coordinates of each of the n images of the standard basis vectors. This is very different for example from continuous functions from $\mathbb{R}$ to $\mathbb{R}$ - we cannot know all about them just by knowing their values at a few points.

Matrix multiplication is composition

Suppose that $T : \mathbb{R}^n \to \mathbb{R}^p$ and $S : \mathbb{R}^p \to \mathbb{R}^m$ are linear transformations. Then $S \circ T$ (S after T) is the linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ defined for $v \in \mathbb{R}^n$ by

$$S \circ T(v) = S(T(v)).$$

Question How does the $(m \times n)$ matrix $M_{S \circ T}$ of $S \circ T$ depend on the $(m \times p)$ matrix M_S of S and the $(p \times n)$ matrix M_T of T?
To answer this we have to think about the definition of $M_{S \circ T}$.

- Its first column has the coordinates of $S \circ T(e_1) = S(T(e_1))$.
- $T(e_1)$ is the first column of M_T.
- Then $S(T(e_1))$ is the matrix-vector product M_S[first column of M_T]. This is the first column of the matrix product $M_S M_T$.
- Same for all the other columns: the conclusion is $M_{S \circ T} = M_S M_T$.

Matrix multiplication is composition of linear transformations.

Corollary Matrix multiplication is associative.

The Image and Kernel of a Linear Transformation

$T : \mathbb{R}^3 \to \mathbb{R}^3$ is the linear transformation with $M_T = \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 5 \\ 1 & 1 & 1 \end{bmatrix}$.

The image of T is the subset of $\mathbb{R}^3$ consisting of all elements $T(v)$, where $v \in \mathbb{R}^3$. This is the set of all vectors of the form

$$a \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + b \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} + c \begin{bmatrix} 0 \\ 5 \\ 1 \end{bmatrix}.$$

In matrix terms, this is the column space of M_T.

The kernel of T is the set of all vectors v in $\mathbb{R}^3$ with $T(v) = 0$. This is the set of all column vectors whose entries a, b, c satisfy

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 5 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = a \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + b \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} + c \begin{bmatrix} 0 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

In matrix terms this is the (right) nullspace of M_T.

Example: The kernel is a **line** and the image is a **plane**

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 5 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 0 & 0 \\ 2 & -1 & 5 & 0 \\ 1 & 1 & 1 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

The kernel/nullspace is $\boxed{\{(-2,1,1)t,\ t \in \mathbb{R}\}}$ a line in $\mathbb{R}^3$.

That $(-2,1,1)$ is in the kernel of T means that (for example) Column 3 of M_T is a linear combination of Columns 1 and 2.

$$-2\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + 1\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Longrightarrow \begin{bmatrix} 0 \\ 5 \\ 1 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$$

It follows that every linear combination of all three columns of M_T is actually a linear combination just of Columns 1 and 2.

The column space of M_T is $\left\{ a\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + b\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} : a, b \in \mathbb{R} \right\}$, a plane in $\mathbb{R}^3$.

Subspaces

Definition A (non-empty) subset V of $\mathbb{R}^n$ is a subspace if

- It is closed under addition: $u + v \in V$ whenever $u \in V$ and $v \in V$.
- It is closed under scalar multiplication: $ku \in V$ whenever $u \in V$ and $k \in \mathbb{R}$.

Examples

1. $\{(x, y, z) \in \mathbb{R}^3 : x + y + z = 1\}$ is not a subspace of $\mathbb{R}^3$. The vectors $(1, 0, 0)$ and $(0, 1, 0)$ belong to this set but their sum $(1, 1, 0)$ does not.
2. $\{(x, y, z) \in \mathbb{R}^3 : (x, y, z) \cdot (1, 2, 3) = 0\}$ is a subspace of $\mathbb{R}^3$.
3. $\{(x, y, z) \in \mathbb{R}^3 : (x, y, z) \cdot (1, 2, 3) \neq 0\}$ is not a subspace of $\mathbb{R}^3$. For example, $(1, 4, 1)$ and $(-5, -2, -1)$ belong to this set but their sum $(-4, 2, 0)$ does not.
4. The kernel of any linear transformation is a subspace.
5. The image of any linear transformation is a subspace.

Exercise Prove these last two points.

How to make subspaces

Let $S = \{v_1, \dots, v_k\}$ be any (finite) subset of $\mathbb{R}^n$.

The subset of $\mathbb{R}^n$ consisting of all linear combinations of the elements of S is a subspace of $\mathbb{R}^n$, denoted by $\langle S \rangle$ or $\langle v_1, v_2, \dots, v_k \rangle$ and called the linear span (or just span) of S.

Proof (that $\langle S \rangle$ is a subspace).
Closed under addition: let $u, v \in \langle S \rangle$. Then $u = a_1 v_1 + a_2 v_2 + \cdots + a_k v_k$, and $v = c_1 v_1 + c_2 v_2 + \cdots + c_k v_k$, where the a_i and b_i are scalars. We need to show that $u + v \in \langle S \rangle$, which means showing that it is a linear combination of $v_1, \dots, v_k$. This is straightforward after everything has been set up, since $u + v = (a_1 + c_1)v_1 + (a_2 + c_2)v_2 + \cdots + (a_k + c_k)v_k$. So S is closed under addition.

Closed under scalar multiplication: let $u \in \langle S \rangle$ and $c \in \mathbb{R}$. We need to show that cu is a linear combination of $v_1, \dots, v_k$. We know that $u = a_1 v_1 + a_2 v_2 + \cdots + a_k v_k$, for scalars $a_1, \dots, a_k$. Then $cu = ca_1\, v_1 + ca_2\, v_2 + \cdots + ca_k\, v_k$, so $cu \in \langle S \rangle$.

Spanning Sets

Let V be a subspace of $\mathbb{R}^n$ (possibly V is all of $\mathbb{R}^n$). A subset S of V is called a spanning set of V if $\langle S \rangle = V$.
This means that every element of V is a linear combination of the elements of S.
Example The set $\{e_1, e_2, e_3\}$ is a spanning set of $\mathbb{R}^3$, where (as usual) $e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. This is saying that every element of $\mathbb{R}^3$ is a linear combination of e_1, e_2, e_3. For example

$$\begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix} = 2e_1 - 3e_2 + 4e_3.$$

Remark A set S of three column vectors in $\mathbb{R}^3$ is a spanning set of $\mathbb{R}^3$ if and only if each of e_1, e_2, e_3 is a linear combination of elements of S.
This occurs if and only if the 3×3 matrix whose columns are the vectors in S has an *inverse*.

Questions about Spanning Sets

1. Does $\mathbb{R}^3$ have a spanning set with fewer than three elements?
2. Does every spanning set of $\mathbb{R}^3$ have exactly three elements?
   NO (why not?)
3. Does every spanning set of $\mathbb{R}^3$ contain one with exactly three elements?
4. If V is a subspace of $\mathbb{R}^3$, does V have a spanning set with at most three elements?
5. If V is a proper subspace of $\mathbb{R}^3$ (i.e. not all of $\mathbb{R}^3$) does V have a spanning set with fewer than three elements?

Note A pair of vectors in $\mathbb{R}^3$ (if they are not scalar multiples of each other) span a plane. Adding a third vector (if it does not lie in this plane) gives a spanning set for all of $\mathbb{R}^3$.

Linear Dependence and Linear Independence

For a subset $\{v_1, \ldots, v_k\}$ of $\mathbb{R}^n$, suppose that v_k is a linear combination of $v_1, \ldots, v_{k-1}$. Then every linear combination of $v_1, \ldots, v_k$ is "already" a linear combination of $v_1, \ldots, v_{k-1}$ and

$$\langle v_1, \ldots, v_k \rangle = \langle v_1, \ldots, v_{k-1} \rangle.$$

If we are interested in the span of $\{v_1, \ldots, v_k\}$ we could throw away v_k and this would not change the span.

Definition A set of (at least two) vectors in R^n is linearly dependent if one of its elements is a linear combination of the others.
A set of vectors in R^n is linearly independent if it is not linearly dependent.[1]

Linear independence means that throwing away any element results in shrinking the span.

[1] Small print: a set with just one vector is linearly independent, unless this vector is the zero vector. Any set that contains the zero vector is linearly dependent.