Linear Transformations

Math 240 — Calculus III

Summer 2015, Session II

Wednesday, July 15, 2015

Agenda

Motivation

In the $m \times n$ linear system

$$A\mathbf{x} = \mathbf{0},$$

we can regard A as transforming elements of $\mathbb{R}^n$ (as column vectors) into elements of $\mathbb{R}^m$ via the rule

$$T(\mathbf{x}) = A\mathbf{x}.$$

Then solving the system amounts to finding all of the vectors $\mathbf{x} \in \mathbb{R}^n$ such that $T(\mathbf{x}) = \mathbf{0}$.

Solving the differential equation

$$y'' + y = 0$$

is equivalent to finding functions y such that $T(y) = 0$, where T is defined as

$$T(y) = y'' + y.$$

Definition

Definition

Let V and W be vector spaces with the same scalars. A mapping $T : V \to W$ is called a **linear transformation** from V to W if it satisfies

1. $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ and
2. $T(c\,\mathbf{v}) = c\,T(\mathbf{v})$

for all vectors $\mathbf{u}, \mathbf{v} \in V$ and all scalars c. The space V is called the **domain** and W the **codomain** of T.

Examples

- $T : \mathbb{R}^n \to \mathbb{R}^m$ defined by $T(\mathbf{x}) = A\mathbf{x}$, where A is an $m \times n$ matrix
- $T : C^k(I) \to C^{k-2}(I)$ defined by $T(y) = y'' + y$
- $T : M_{m\times n}(\mathbb{R}) \to M_{n\times m}(\mathbb{R})$ defined by $T(A) = A^T$
- $T : P_1 \to P_2$ defined by $T(a+bx) = (a+2b) + 3ax + 4bx^2$

Examples

1. Verify that $T : M_{m\times n}(\mathbb{R}) \to M_{n\times m}(\mathbb{R})$, where $T(A) = A^T$, is a linear transformation.
   - The transpose of an $m \times n$ matrix is an $n \times m$ matrix.
   - If $A, B \in M_{m\times n}(\mathbb{R})$, then
   $$T(A + B) = (A + B)^T = A^T + B^T = T(A) + T(B).$$
   - If $A \in M_{m\times n}(\mathbb{R})$ and $c \in \mathbb{R}$, then
   $$T(cA) = (cA)^T = cA^T = c\,T(A).$$
2. Verify that $T : C^k(I) \to C^{k-2}(I)$, where $T(y) = y'' + y$, is a linear transformation.
   - If $y \in C^k(I)$ then $T(y) = y'' + y \in C^{k-2}(I)$.
   - If $y_1, y_2 \in C^k(I)$, then
   $$\begin{aligned} T(y_1 + y_2) = (y_1 + y_2)'' + (y_1 + y_2) &= y_1'' + y_2'' + y_1 + y_2 \\ &= (y_1'' + y_1) + (y_2'' + y_2) = T(y_1) + T(y_2). \end{aligned}$$
   - If $y \in C^k(I)$ and $c \in \mathbb{R}$, then
   $$T(cy) = (cy)'' + (cy) = cy'' + cy = c(y'' + y) = c\,T(y).$$

Specifying linear transformations

A consequence of the properties of a linear transformation is that they preserve linear combinations, in the sense that

$$T(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = c_1\,T(\mathbf{v}_1) + \cdots + c_n\,T(\mathbf{v}_n).$$

In particular, if $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis for the domain of T, then knowing $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)$ is enough to determine T everywhere.

Linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$

Let A be an $m \times n$ matrix with real entries and define $T : \mathbb{R}^n \to \mathbb{R}^m$ by $T(\mathbf{x}) = A\mathbf{x}$. Verify that T is a linear transformation.

- If $\mathbf{x}$ is an $n \times 1$ column vector then $A\mathbf{x}$ is an $m \times 1$ column vector.
- $T(\mathbf{x} + \mathbf{y}) = A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = T(\mathbf{x}) + T(\mathbf{y})$
- $T(c\mathbf{x}) = A(c\mathbf{x}) = cA\mathbf{x} = cT(\mathbf{x})$

Such a transformation is called a **matrix transformation**. In fact, *every* linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a matrix transformation.

Matrix transformations

Example

Determine the matrix of the linear transformation $T : \mathbb{R}^4 \to \mathbb{R}^3$ defined by

$$T(x_1, x_2, x_3, x_4) = (2x_1 + 3x_2 + x_4,\ 5x_1 + 9x_3 - x_4,\\ 4x_1 + 2x_2 - x_3 + 7x_4).$$

Theorem

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then T is described by the matrix transformation $T(\mathbf{x}) = A\mathbf{x}$, where

$$A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$$

and $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ denote the standard basis vectors for $\mathbb{R}^n$. This A is called the **matrix of** *T.*

Kernel

Definition

Suppose $T : V \to W$ is a linear transformation. The set consisting of all the vectors $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \mathbf{0}$ is called the **kernel** of T. It is denoted

$$\operatorname{Ker}(T) = \{\mathbf{v} \in V : T(\mathbf{v}) = \mathbf{0}\}.$$

Example

Let $T : C^k(I) \to C^{k-2}(I)$ be the linear transformation $T(y) = y'' + y$. Its kernel is spanned by $\{\cos x, \sin x\}$.

Remarks

- The kernel of a linear transformation is a subspace of its domain.
- The kernel of a matrix transformation is simply the null space of the matrix.

Range

Definition

The **range** of the linear transformation $T : V \to W$ is the subset of W consisting of everything "hit by" T. In symbols,

$$\mathrm{Rng}(T) = \{T(\mathbf{v}) \in W : \mathbf{v} \in V\}.$$

Example

Consider the linear transformation $T : M_n(\mathbb{R}) \to M_n(\mathbb{R})$ defined by $T(A) = A + A^T$. The range of T is the subspace of symmetric $n \times n$ matrices.

Remarks

- The range of a linear transformation is a subspace of its codomain.
- The range of a matrix transformation is the column space of the matrix.

Rank-Nullity revisited

Suppose T is the matrix transformation with $m \times n$ matrix A.

We know

- $\mathrm{Ker}(T) = \mathrm{nullspace}(A)$,
- $\mathrm{Rng}(T) = \mathrm{colspace}(A)$,
- the domain of T is $\mathbb{R}^n$.

Hence,

- $\dim(\mathrm{Ker}(T)) = \mathrm{nullity}(A)$,
- $\dim(\mathrm{Rng}(T)) = \mathrm{rank}(A)$,
- $\dim(\text{domain of } T) = n$.

We know from the rank-nullity theorem that

$$\mathrm{rank}(A) + \mathrm{nullity}(A) = n.$$

This fact is also true when T is not a matrix transformation:

Theorem

If $T : V \to W$ is a linear transformation and V is finite-dimensional, then

$$\dim(\mathrm{Ker}(T)) + \dim(\mathrm{Rng}(T)) = \dim(V).$$

The function of bases

Theorem

Let V be a vector space with basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. Then every vector $\mathbf{v} \in V$ can be written in a unique way as a linear combination

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n.$$

In other words, picking a basis for a vector space allows us to give coordinates for points. This will allow us to give matrices for linear transformations of vector spaces besides $\mathbb{R}^n$.

The matrix of a linear transformation

Definition

Let V and W be vector spaces with *ordered* bases $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $C = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$, respectively, and let $T : V \to W$ be a linear transformation. The **matrix representation of T relative to the bases B and C** is

$$A = [a_{ij}]$$

where

$$T(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + a_{2j}\mathbf{w}_2 + \cdots + a_{mj}\mathbf{w}_m.$$

In other words, A is the matrix whose j-th column is $T(\mathbf{v}_j)$, expressed in coordinates using $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$.

Example

Let $T : P_1 \to P_2$ be the linear transformation defined by

$$T(a+bx) = (2a-3b) + (b-5a)x + (a+b)x^2.$$

Use bases $\{1, x\}$ for P_1 and $\{1, x, x^2\}$ for P_2 to give a matrix representation of T. We have

$$T(1) = 2 - 5x + x^2 \quad \text{and} \quad T(x) = -3 + x + x^2,$$

so

$$A_1 = \begin{bmatrix} 2 & -3 \\ -5 & 1 \\ 1 & 1 \end{bmatrix}.$$

Example

Let $T : P_1 \to P_2$ be the linear transformation defined by

$$T(a + bx) = (2a - 3b) + (b - 5a)x + (a + b)x^2.$$

Now instead use the bases $\{1, x + 5\}$ for P_1 and $\{1, 1 + x, 1 + x^2\}$ for P_2. We have

$$T(1) = 2 - 5x + x^2 = 6(1) - 5(1 + x) + (1 + x^2)$$

and

$$T(x + 5) = 7 - 24x + 6x^2 = 25(1) - 24(1 + x) + 6(1 + x^2),$$

so

$$A_1 = \begin{bmatrix} 2 & -3 \\ -5 & 1 \\ 1 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 6 & 25 \\ -5 & -24 \\ 1 & 6 \end{bmatrix}.$$

Composition of linear transformations

Definition

Let $T_1 : U \to V$ and $T_2 : V \to W$ be linear transformations. Their **composition** is the linear transformation $T_2 \circ T_1$ defined by

$$(T_2 \circ T_1)(\mathbf{u}) = T_2(T_1(\mathbf{u})).$$

Theorem

Let T_1 and T_2 be as above, and let B, C, and D be ordered bases for U, V, and W, respectively. If

- *A_1 is the matrix representation for T_1 relative to B and C,*
- *A_2 is the matrix representation for T_2 relative to C and D,*
- *A_{21} is the matrix representation for $T_2 \circ T_1$ relative to B and D,*

then $A_{21} = A_2 A_1$.

The inverse of a linear transformation

Definition

If $T : V \to W$ is a linear transformation, its **inverse** (if it exists) is a linear transformation $T^{-1} : W \to V$ such that

$$\left(T^{-1} \circ T\right)(\mathbf{v}) = \mathbf{v} \quad \text{and} \quad \left(T \circ T^{-1}\right)(\mathbf{w}) = \mathbf{w}$$

for all $\mathbf{v} \in V$ and $\mathbf{w} \in W$.

Theorem

Let T be as above and let A be the matrix representation of T relative to bases B and C for V and W, respectively. T has an inverse transformation if and only if A is invertible and, if so, T^{-1} is the linear transformation with matrix A^{-1} relative to C and B.

Example

Let $T : P_2 \to P_2$ be defined by

$$T(a + bx + cx^2) = (3a - b + c) + (a - c)x + (4b + c)x^2.$$

Using the basis $\{1, x, x^2\}$ for P_2, the matrix representation for T is

$$A = \begin{bmatrix} 3 & -1 & 1 \\ 1 & 0 & -1 \\ 0 & 4 & 1 \end{bmatrix}.$$

This matrix is invertible and

$$A^{-1} = \frac{1}{17} \begin{bmatrix} 4 & 5 & 1 \\ -1 & 3 & 4 \\ 4 & -12 & 1 \end{bmatrix}.$$

Thus, T^{-1} is given by

$$T^{-1}(a + bx + cx^2) = \tfrac{4a+5b+c}{17} + \tfrac{-a+3b+4c}{17}x + \tfrac{4a-12b+c}{17}x^2.$$

Kernel and Range

Theorem

Let $T : V \to W$ be a linear transformation and A be a matrix representation of T relative to some bases for V and W.

- $\mathrm{Ker}(T) = \{c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \in V : (c_1, \ldots, c_n) \in \mathrm{nullspace}(A)\}$,
- $\mathrm{Rng}(T) = \{c_1\mathbf{w}_1 + \cdots + c_m\mathbf{w}_m \in W : (c_1, \ldots, c_m) \in \mathrm{colspace}(A)\}$.