SF1684 Algebra and Geometry

Lecture 4

Linear transformations, kernel and range

Bao Nguyen

Outline

Outline

1. Matrices and linear transformations
2. Kernel and range
3. Inverse linear transformations

Transformations

Definition

A **transformation** T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a function that maps each vector $\mathbf{x} \in \mathbb{R}^n$ into a vector $\mathbf{y} \in \mathbb{R}^m$. We can write either

$$\mathbf{y} = T(\mathbf{x}), \quad \mathbf{x} \xrightarrow{T} \mathbf{y} \quad \text{or} \quad T\colon \mathbb{R}^n \to \mathbb{R}^m.$$

- $\mathbb{R}^n$ is called the **domain** and $\mathbb{R}^m$ the **codomain** of T.
- The vector $T(\mathbf{x})$ is the image of $\mathbf{x}$ under T.
- The set of all images $\{T(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n\}$ is the **range** of T (may actually be smaller than the codomain of T).

Example

If $\mathbf{x}_0$ is a fixed vector, then the translation $T_{\mathbf{x}_0} : \mathbb{R}^n \to \mathbb{R}^n$ is defined as $T_{\mathbf{x}_0}(\mathbf{x}) = \mathbf{x} + \mathbf{x}_0$.

Matrix transformations

We are primarily interested in transformations that can be defined using matrices. Let A be an $m \times n$ matrix, we can define a transformation $T_A \colon \mathbb{R}^n \to \mathbb{R}^m$ such that T_A maps $\mathbf{x} \in \mathbb{R}^n$ into $A\mathbf{x} \in \mathbb{R}^m$. That is $T_A(\mathbf{x}) = A\mathbf{x}$.

Linear transformation

A function $T \colon \mathbb{R}^n \to \mathbb{R}^m$ is called a linear transformation if

1. $T(c\mathbf{u}) = cT(\mathbf{u})$ and
2. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and all scalars c.

In the case where $m = n$, T is called a **linear operator** on $\mathbb{R}^n$.

Some properties

If T is a linear transformation then

- i $T(\mathbf{0}) = \mathbf{0}$.
- ii $T(c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k) = c_1T(\mathbf{u}_1) + c_2T(\mathbf{u}_2) + \cdots + c_kT(\mathbf{u}_k)$.

Every matrix transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is linear. Conversely, all linear transformations can be performed by matrix multiplications.

Theorem

Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ be standard unit vectors in $\mathbb{R}^n$. If $\mathbf{x}$ is a vector in $\mathbb{R}^n$, then T can be expressed as

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

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

The $m \times n$ matrix A is called the **standard matrix of T** and T is the **transformation corresponding to A**.

Remark

It follows that $T(\mathbf{x}) = \mathbf{y}$ is a linear transformation if and only if the equations relating the components of $\mathbf{x}$ and $\mathbf{y}$ are linear equations.

Linear operators on $\mathbb{R}^2$

Rotation through θ about the origin

$$R_\theta = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Let L be a line through $\mathbf{0}$ and making an angle θ with the positive x-axis

Reflection through the line L

$$H_\theta = \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix}.$$

Projection onto the line L

$$P_\theta = \begin{bmatrix} \cos^2(\theta) & \sin(\theta)\cos(\theta) \\ \sin(\theta)\cos(\theta) & \sin^2(\theta) \end{bmatrix}.$$

Composition of linear transformations

Definition

Let $T_1 : \mathbb{R}^n \to \mathbb{R}^k$ and $T_2 : \mathbb{R}^k \to \mathbb{R}^m$ be linear transformations. Their composition is the linear transformation $T_2 \circ T_1 : \mathbb{R}^n \to \mathbb{R}^m$ defined by

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

Theorem

If A is the standard matrix for $T_A : \mathbb{R}^n \to \mathbb{R}^k$ and B is the standard matrix for $T_B : \mathbb{R}^k \to \mathbb{R}^m$ then BA is the standard matrix for $T_B \circ T_A$

$$T_B \circ T_A = T_{BA}.$$

Note that A has size $k \times n$, B has size $m \times k$, therefore BA has size $m \times n$.

Composing rotations

$$R_{\theta_2}R_{\theta_1} = \begin{bmatrix} \cos(\theta_2) & -\sin(\theta_2) \\ \sin(\theta_2) & \cos(\theta_2) \end{bmatrix} \begin{bmatrix} \cos(\theta_1) & -\sin(\theta_1) \\ \sin(\theta_1) & \cos(\theta_1) \end{bmatrix}$$

$$= \begin{bmatrix} \cos(\theta_1+\theta_2) & -\sin(\theta_1+\theta_2) \\ \sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2) \end{bmatrix} = R_{\theta_1+\theta_2}.$$

Composing reflections

$$H_{\theta_2}H_{\theta_1} = \begin{bmatrix} \cos(2\theta_2) & \sin(2\theta_2) \\ \sin(2\theta_2) & -\cos(2\theta_2) \end{bmatrix} \begin{bmatrix} \cos(2\theta_1) & \sin(2\theta_1) \\ \sin(2\theta_1) & -\cos(2\theta_1) \end{bmatrix}$$

$$= \begin{bmatrix} \cos(2\theta_2-2\theta_1) & -\sin(2\theta_2-2\theta_1) \\ \sin(2\theta_2-2\theta_1) & \cos(2\theta_2-2\theta_1) \end{bmatrix} = R_{2(\theta_2-\theta_1)}.$$

Warning

Composition of two transformations is not commutative.

Orthogonal operators

Definition

A linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^n$ is said to be orthogonal (or linear isometry) if $\|T(\mathbf{x})\| = \|\mathbf{x}\|$ (length preserving) for all $\mathbf{x} \in \mathbb{R}^n$.

Theorem

$T\colon \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal if and only if $T(\mathbf{x}) \cdot T(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$.

Since the angles between vectors are defined through the inner product, orthogonal operators also **preserve angles** between vectors.

Orthogonal matrices

Definition

A square matrix A is said to be orthogonal if $A^TA = I$, or equivalently, $A^{-1} = A^T$.

Theorem

If A and B are orthogonal matrices, then

i. A^T is orthogonal.
ii. A^{-1} is orthogonal.
iii. AB is orthogonal.
iv. $\det(A) = 1$ or $\det(A) = -1$.

Theorem

If A is an $n \times n$ matrix, then the following statements are equivalent

a. A is orthogonal.
b. $\|A\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{R}^n$.
c. $A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.
d. The column vectors of A are orthonormal.
e. The row vectors of A are orthonormal.

Since a linear operator T is orthogonal if and only if its standard matrix has the property $\|A\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{R}^n$, we have the result

Theorem

$T\colon \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal if and only if its standard matrix is orthogonal.

Orthogonal linear operators on $\mathbb{R}^2$

Rotations and reflections

The standard matrix of an orthogonal linear operator $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ is always expressed as

$$R_\theta = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \quad \text{or} \quad H_{\theta/2} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{bmatrix},$$

for some $\theta \in \mathbb{R}$. That is, T is either a rotation about the origin or a reflection about a line through the origin.

Remark

Since $\det(R_\theta) = 1$ and $\det(H_{\theta/2}) = -1$, we observe that a 2×2 orthogonal matrix A represents a rotation if $\det(A) = 1$ and a reflection if $\det(A) = -1$.

Example

See Example 4 on page 285 of the textbook.

Some important non-length-preserving linear operators on $\mathbb{R}^2$

Scaling operators

Scaling operator with factor k: $T(x, y) = (kx, ky)$ (called a contraction if $k \in [0, 1)$ and a dilation if $k > 1$).

Expansions and compressions

Expansion or compression with factor k in

- the x-direction: $T(x, y) = (kx, y)$.
- the y-direction: $T(x, y) = (x, ky)$.

Shears

Shear with factor k in

- the x-direction: $T(x, y) = (x + ky, y)$.
- the y-direction: $T(x, y) = (x, y + kx)$.

Linear operators on $\mathbb{R}^3$

Orthogonal projections onto coordinate planes

Orthogonal projections on

- the xy-plane: $T(x, y, z) = (x, y, 0)$.
- the xz-plane: $T(x, y, z) = (x, 0, z)$.
- the yz-plane: $T(x, y, z) = (0, y, z)$.

Reflections about coordinate planes

Reflections about

- the xy-plane: $T(x, y, z) = (x, y, -z)$.
- the xz-plane: $T(x, y, z) = (x, -y, z)$.
- the yz-plane: $T(x, y, z) = (-x, y, z)$.

Rotations in $\mathbb{R}^3$

A three-dimensional rotation is determined by an axis $\mathbf{u}$ (through the origin), the direction of rotation about that axis, and an angle θ. The positive rotation direction satisfies the **right-hand rule**: If your right thumb points in the direction of $\mathbf{u}$, then your fingers curl in the direction of the positive rotation.

The standard matrix of rotation $R_{\mathbf{u},\theta}$ is given by Theorem 6.2.8 in the textbook.

Rotation matrix

A 3×3 matrix A represents a rotation if and only if A is orthogonal and $\det(A) = 1$.

All points lying on the axis of rotation must satisfy the linear system $A\mathbf{x} = \mathbf{x}$ (fixed points of the rotation).

Composing rotations

A composition of two rotations in R^3 (around axes through the origin) is another rotation (around some appropriate axis through the origin).

Determining the axis and angle of rotation from a matrix A

The steps are as follows

1. Show that A is orthogonal and $\det(A) = 1$.
2. Find the axis of rotation: The vector equation of the axis is the solution of the linear system $(I - A)\mathbf{x} = \mathbf{0}$.
3. Choose an orientation of the axis: Select a vector $\mathbf{w}$ perpendicular to the axis, then we can orient the axis using the vector $\mathbf{u} = \mathbf{w} \times A\mathbf{w}$.
4. The rotation angle θ, relative to $\mathbf{u}$, is the angle between $\mathbf{w}$ and $A\mathbf{w}$ (satisfying $0 \leq \theta \leq \pi$) and can be computed from the formula

$$\cos(\theta) = \frac{\mathbf{w} \cdot A\mathbf{w}}{\|\mathbf{w}\| \, \|A\mathbf{w}\|}.$$

Example

See Example 7 on page 292 of the textbook.

Rotations about the coordinate axes

Rotation about the positive x-axis through an angle θ

$$R_{x,\theta} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Rotation about the positive y-axis through an angle θ

$$R_{y,\theta} = \begin{bmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{bmatrix}.$$

Rotation about the positive z-axis through an angle θ

$$R_{z,\theta} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Outline

Kernel and range of a linear transformation

Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

Kernel of a linear transformation

The set of all vectors $\mathbf{x} \in \mathbb{R}^n$ for which $T(\mathbf{x}) = \mathbf{0}$ is a subspace of $\mathbb{R}^n$. It is called the kernel of T and is denoted by $\ker(T)$,

$$\ker(T) = \{\mathbf{x} \in \mathbb{R}^n \mid T(\mathbf{x}) = \mathbf{0}\}.$$

Image of a linear transformation

If S is a subspace of $\mathbb{R}^n$, then the image of S under the transformation T,

$$T(S) = \{T(\mathbf{x}) \mid \mathbf{x} \in S\},$$

is a subspace of $\mathbb{R}^m$. In particular, if $S = \mathbb{R}^n$ then the set $T(\mathbb{R}^n)$ is called the range of T and is denoted by $\operatorname{ran}(T)$.

Kernel and range of a matrix transformation

Let A be an $m \times n$ matrix and $T_A : \mathbb{R}^n \to \mathbb{R}^m$ be the corresponding transformation, that is $T_A(\mathbf{x}) = A\mathbf{x}$.

Kernel of a matrix transformation

The kernel of T_A is the solution space of $A\mathbf{x} = \mathbf{0}$. It is called the null space of A and is denoted by $\operatorname{null}(A)$

$$\ker(T_A) = \operatorname{null}(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}.$$

Range of a matrix transformation

The range of T_A is the column space of A

$$\operatorname{ran}(T_A) = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\} = \operatorname{span}\big\{\mathbf{c}_1(A), \mathbf{c}_2(A), \ldots, \mathbf{c}_n(A)\big\}.$$

Onto and one-to-one transformations

Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

Onto transformation

T is onto if its range is the entire codomain $\mathbb{R}^m$; that is, the equation $T(\mathbf{x}) = \mathbf{b}$ has **at least** one solution for every $\mathbf{b} \in \mathbb{R}^m$.

One-to-one transformation

T is one-to-one if T maps distinct vectors in $\mathbb{R}^n$ into distinct vectors in $\mathbb{R}^m$; that is, the equation $T(\mathbf{x}) = \mathbf{b}$ has **at most** one solution for every $\mathbf{b} \in \mathbb{R}^m$.

Theorem

The following statements are equivalent

a. T is one-to-one.
b. $\ker(T) = \{\mathbf{0}\}$.

Onto and one-to-one matrix transformations

Let A be an $m \times n$ matrix and $T_A : \mathbb{R}^n \to \mathbb{R}^m$ be the corresponding transformation.

Onto matrix transformation

T_A is onto if and only if the linear system $A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b} \in \mathbb{R}^m$.

One-to-one matrix transformation

T_A is one-to-one if and only if the linear system $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

One-to-one linear operator

If $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear operator on $\mathbb{R}^n$, then T is one-to-one if and only if it is onto.

Fundamental theorem of invertible matrices (cont.)

If A is an $n \times n$ matrix, and if T_A is a linear operator on $\mathbb{R}^n$ corresponding to A, then the following statements are equivalent

a. The reduced echelon form of A is I_n.
b. A is a product of elementary matrices.
c. A is invertible.
d. $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
e. $A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b} \in \mathbb{R}^n$.
f. $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b} \in \mathbb{R}^n$.
g. The column vectors of A are linearly independent.
h. The row vectors of A are linearly independent.
i. $\det(A) \neq 0$.
j. $\lambda = 0$ is not an eigenvalue of A.
k. T_A is one-to-one.
l. T_A is onto.

Outline

Inverse linear transformations

Definition

If $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is one-to-one, then it has an inverse transformation T^{-1}. For each vector $\mathbf{u} \in \operatorname{ran}(T)$, the value of $T^{-1}(\mathbf{u})$ is the unique vector $\mathbf{x}$ in the domain of T for which $T(\mathbf{x}) = \mathbf{u}$. Thus

$$T^{-1}(\mathbf{u}) = \mathbf{x} \quad \text{if and only if} \quad T(\mathbf{x}) = \mathbf{u}.$$

Theorem

$T^{-1}\colon\ \operatorname{ran}(T) \to \mathbb{R}^n$ is a one-to-one linear transformation.

Inverse linear operators

If T is a one-to-one linear operator on $\mathbb{R}^n$, then T is onto and hence that the domain of T^{-1} is all of $\mathbb{R}^n$. Thus T^{-1} is also a one-to-one linear operator on $\mathbb{R}^n$.

Standard matrix for inverse operator

If T_A is a one-to-one linear operator on $\mathbb{R}^n$ with standard matrix A, then A is invertible and its inverse is the standard matrix for T_A^{-1}, that is

$$T_A^{-1} = T_{A^{-1}}.$$

A one-to-one linear operator is also called an **invertible linear operator**.

Decomposition of linear transformations on $\mathbb{R}^2$

An invertible 2×2 matrix A can be expressed as a product of elementary matrices which are corresponding to basic operators on $\mathbb{R}^2$. Thus, we have

Theorem

Any invertible linear operator on $\mathbb{R}^2$ can be decomposed into shears, compressions, and expansions in the coordinate directions, and reflections about the coordinate axes and about the line $y = x$.

Transforming with a diagonal matrix

Since we can factor $A = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ as $\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & 1 \end{bmatrix}$, the geometric effect of multiplication by A is to scale by a factor of λ_1 in the x-direction and scale by a factor of λ_2 in the y-direction.

Example

See Example 8 on page 311 of the textbook.