
Linear Algebra

주재걸

고려대학교 컴퓨터학과



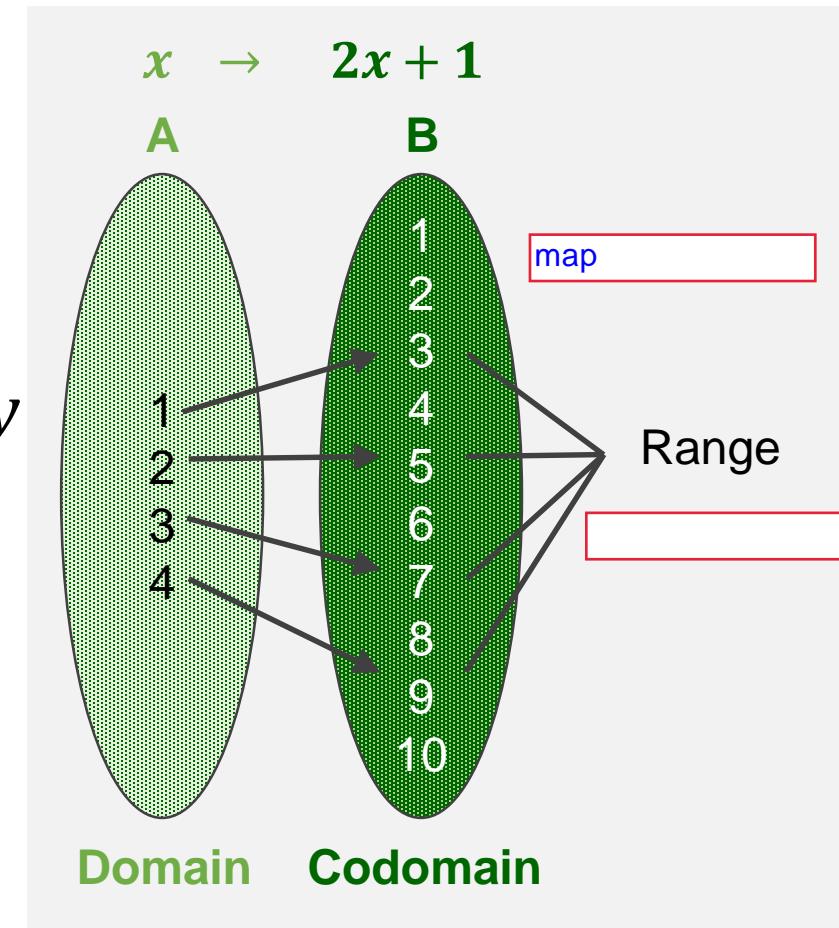


Lecture Overview

- Elements in linear algebra
- Linear system
- Linear combination, vector equation,
Four views of matrix multiplication
- Linear independence, span, and subspace
- Linear transformation
- Least squares
- Eigendecomposition
- Singular value decomposition

Transformation

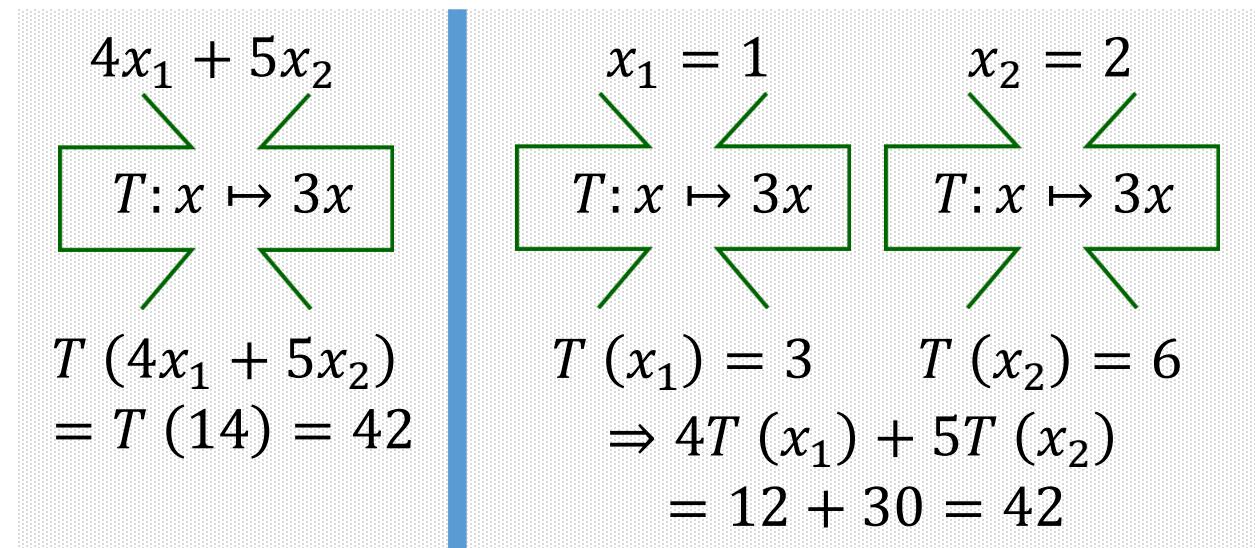
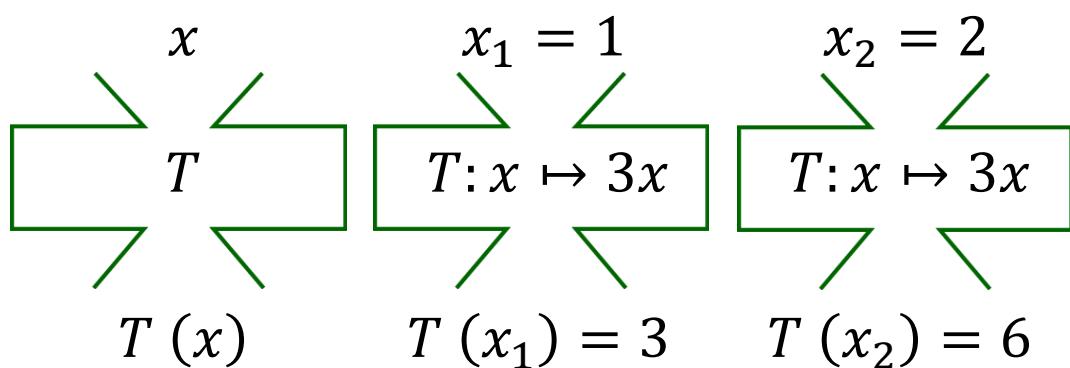
- A **transformation, function, or mapping, T** maps an input x to an output y
 - Mathematical notation: $T: x \mapsto y$
- **Domain:** Set of all the possible values of x
- **Co-domain:** Set of all the possible values of y
- **Image:** a mapped output y , given x
- **Range:** Set of all the output values mapped by each x in the domain
- **Note:** the output mapped by a particular x is **uniquely determined**.



Linear Transformation

- **Definition:** A transformation (or mapping) T is **linear** if:
 - $I. \quad T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$ for all \mathbf{u}, \mathbf{v} in the domain of T and for all scalars c and d

- Simple example: $T: x \mapsto y, T(x) = y = 3x$





Transformations between Vectors

- $T: \mathbf{x} \in \mathbb{R}^n \mapsto \mathbf{y} \in \mathbb{R}^m$: Mapping n -dim vector to m -dim vector
- Example:

$$T: \mathbf{x} \in \mathbb{R}^3 \mapsto \mathbf{y} \in \mathbb{R}^2 \quad \mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \in \mathbb{R}^3 \quad \mapsto \quad \mathbf{y} = T(\mathbf{x}) = \begin{bmatrix} 4 \\ 5 \end{bmatrix} \in \mathbb{R}^2$$



Matrix of Linear Transformation

- Example: Suppose T is a linear transformation from \mathbb{R}^2 to \mathbb{R}^3 such that

$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$ and $T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$. With no additional information,

find a formula for the image of an arbitrary \mathbf{x} in \mathbb{R}^2

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\Rightarrow T(\mathbf{x}) = T\left(x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = x_1 T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) + x_2 T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)$$

$$= x_1 \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ -1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$



Matrix of Linear Transformation

- In general, let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a **linear** transformation. Then T is always written as a matrix-vector multiplication, i.e.,

$$T(\mathbf{x}) = A\mathbf{x} \text{ for all } \mathbf{x} \in \mathbb{R}^n$$

- In fact, the j -th column of $A \in \mathbb{R}^{m \times n}$ is equal to the vector $T(\mathbf{e}_j)$, where \mathbf{e}_j is the j -th column of the identity matrix in $\mathbb{R}^{n \times n}$:

$$A = [T(\mathbf{e}_1) \quad \cdots \quad T(\mathbf{e}_n)]$$

- Here, the matrix A is called the **standard matrix** of the linear transformation T



Matrix of Linear Transformation

- **Example:** Find the standard matrix A of a linear transformation from \mathbb{R}^3 to \mathbb{R}^2 such that

$$T\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, T\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}, \text{ and } T\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}.$$

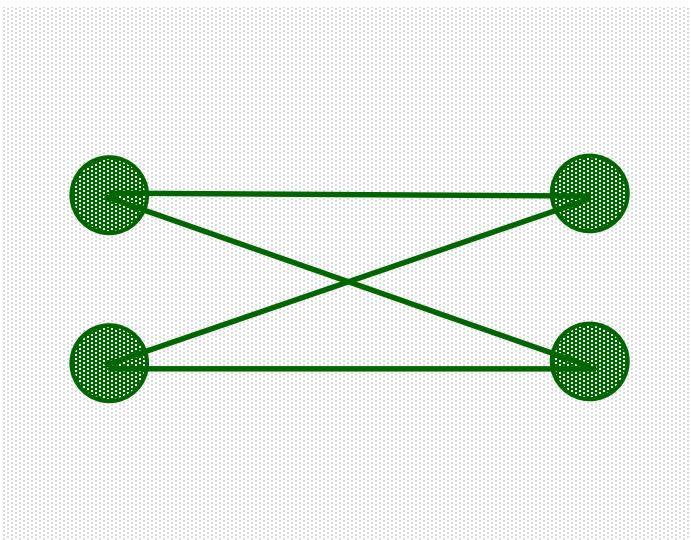
$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{aligned} \Rightarrow T(\mathbf{x}) &= T\left(x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = x_1 T\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 T\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3 T\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ &= x_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} 4 \\ 3 \end{pmatrix} + x_3 \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{bmatrix} 1 & 4 & 5 \\ 2 & 3 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = A\mathbf{x} \end{aligned}$$

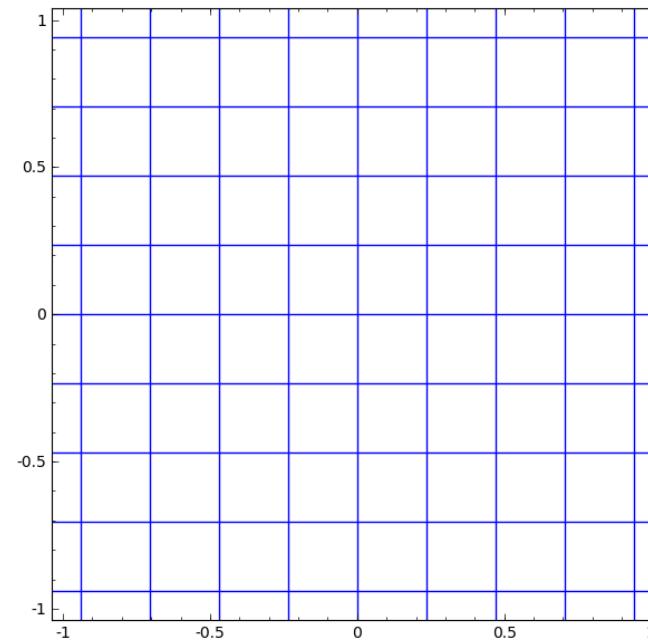


Linear Transformation in Neural Networks

- Fully-connected layers (linear layer)



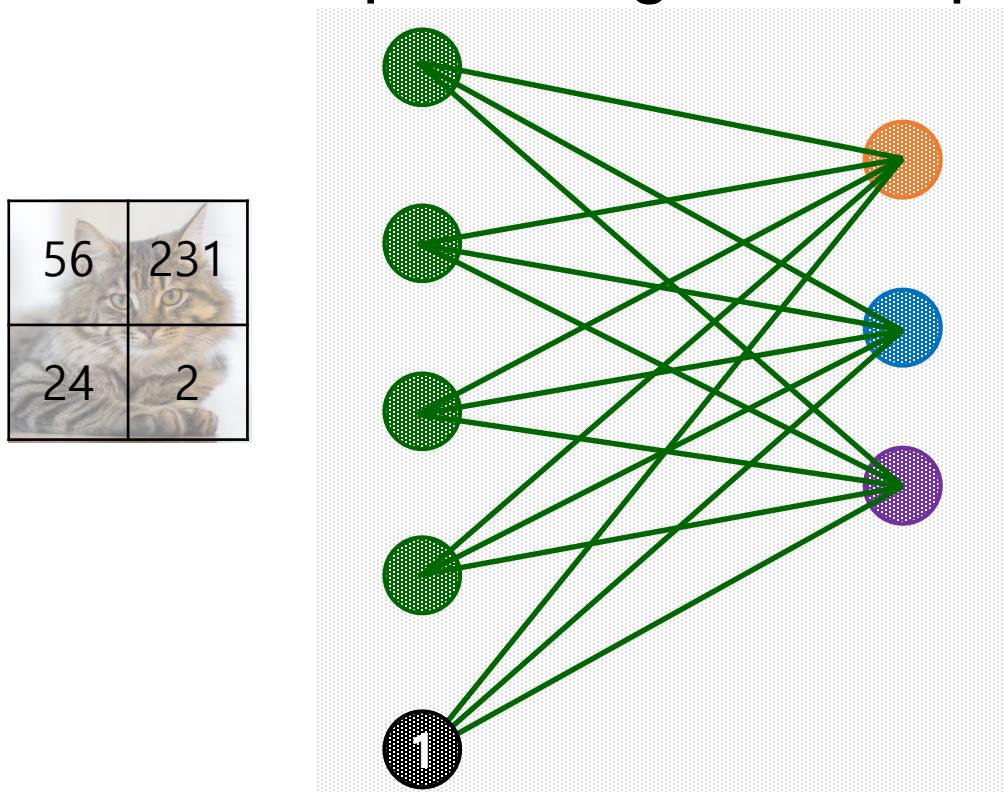
$$\mathbf{x} \rightarrow T_1 \mathbf{y}$$



<https://colah.github.io/posts/2014-03-NN-Manifolds-Topology/>

Affine Layer in Neural Networks

- Fully-connected layers usually involve a bias term. That's why we call it an affine layer, but not a linear layer.
- Example: Image with 4 pixels and 3 classes (**cat/dog/ship**)



$$\begin{array}{c} \begin{matrix} 0.2 & -0.5 & 0.1 & 2 \\ 1.5 & 1.3 & 2.1 & 1 \\ -0.2 & 0.3 & 0.7 & -1.3 \end{matrix} + \begin{matrix} 56 \\ 231 \\ 24 \\ 2 \end{matrix} = \begin{matrix} 1.1 \\ 3.2 \\ -1.2 \end{matrix} = \begin{matrix} -96.8 \\ 439.9 \\ 711 \end{matrix} \\ = 56 \begin{matrix} 0.2 \\ 1.5 \\ -0.2 \end{matrix} + 231 \begin{matrix} -0.5 \\ 1.3 \\ 0.3 \end{matrix} + 24 \begin{matrix} 0.1 \\ 2.1 \\ 0.7 \end{matrix} + 2 \begin{matrix} 2 \\ 1 \\ -1.3 \end{matrix} + 1 \begin{matrix} 1.1 \\ 3.2 \\ -1.2 \end{matrix} \\ = \begin{matrix} 0.2 & -0.5 & 0.1 & 2 & 1.1 \\ 1.5 & 1.3 & 2.1 & 1 & 3.2 \\ -0.2 & 0.3 & 0.7 & -1.3 & -1.2 \end{matrix} \begin{matrix} 56 \\ 231 \\ 24 \\ 2 \\ 1 \end{matrix} \end{array}$$