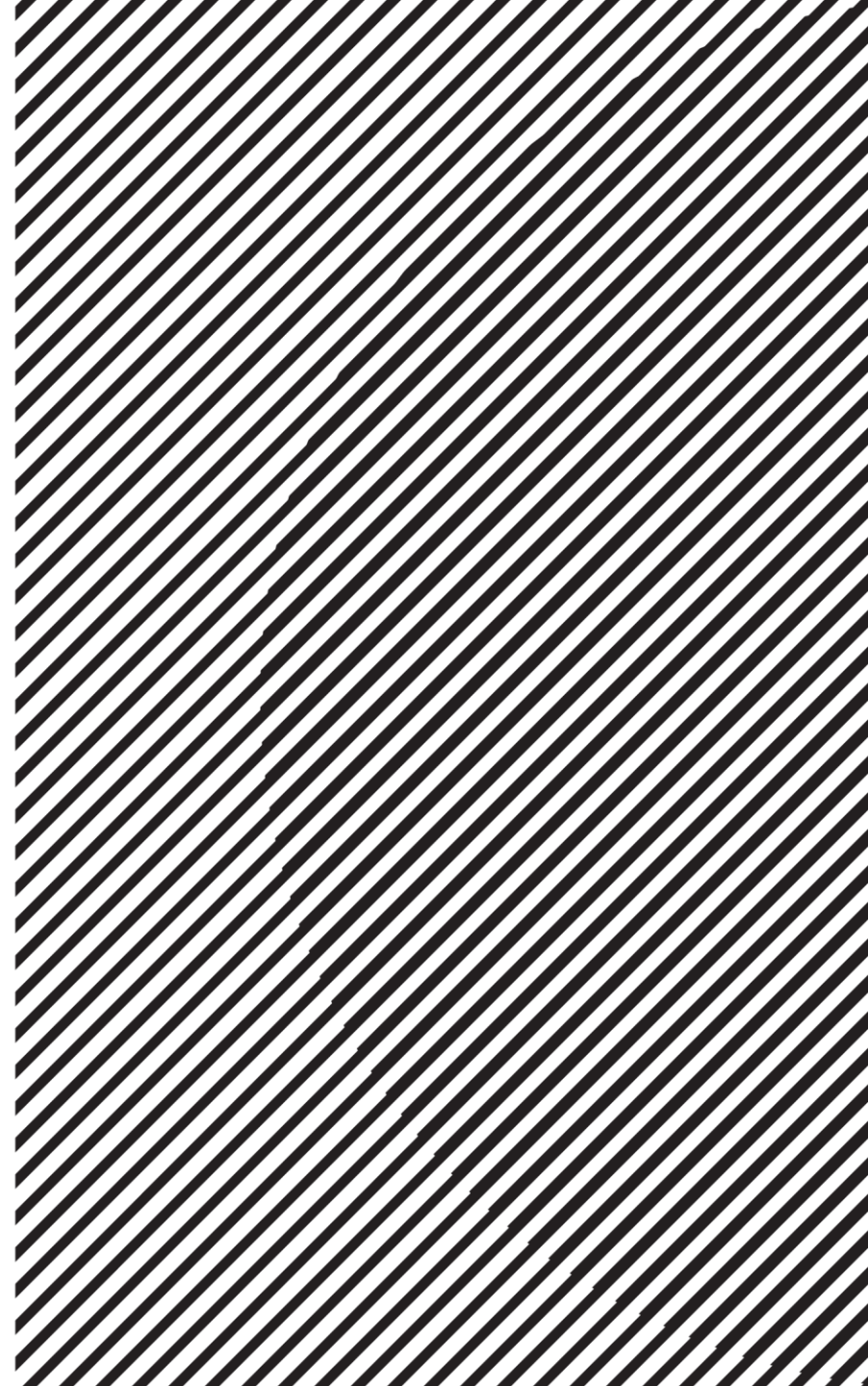# Linear Algebra

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

- <u>Linear transformation</u>

- 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.
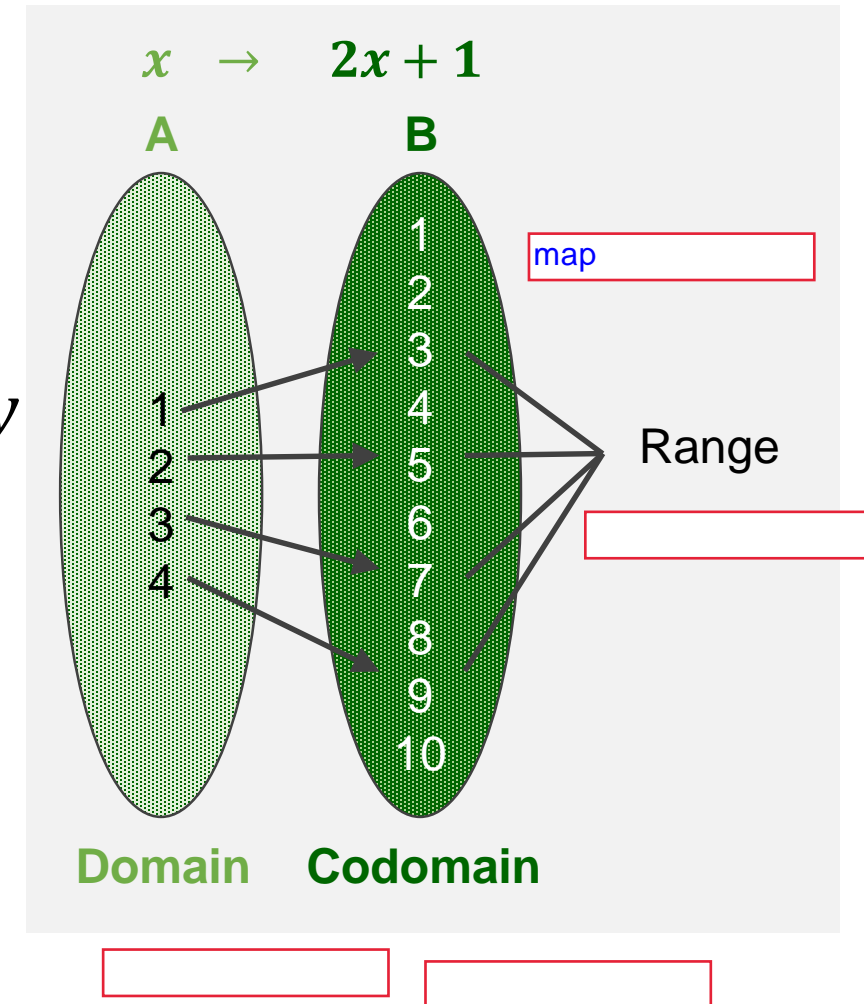
Linear Transformation

- **Definition**: A transformation (or mapping) T is **linear** if:

  I. $T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v}$ in the domain of T
  and for all scalars c and d

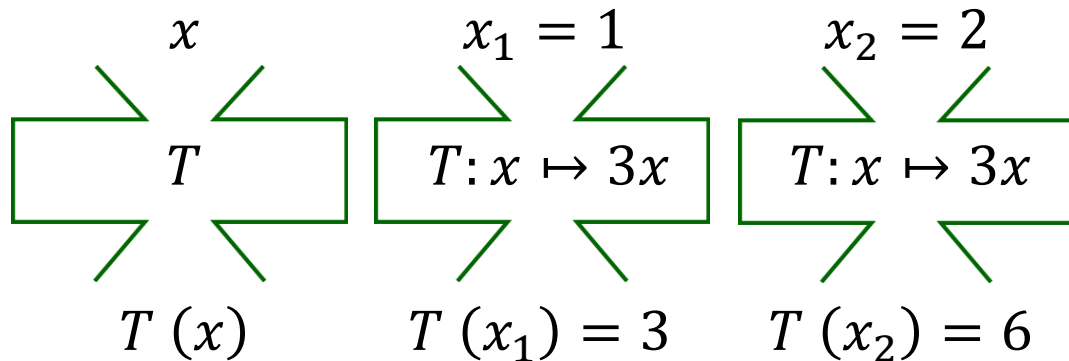- Simple example: $T: x \mapsto y,\ T(x) = y = 3x$

x $x_1 = 1$ $x_2 = 2$ $4x_1 + 5x_2$ $x_1 = 1$ $x_2 = 2$

T $T: x \mapsto 3x$ $T: x \mapsto 3x$ $T: x \mapsto 3x$ $T: x \mapsto 3x$ $T: x \mapsto 3x$

$T(x)$ $T(x_1) = 3$ $T(x_2) = 6$ $T(4x_1 + 5x_2)$ $T(x_1) = 3$ $T(x_2) = 6$
$= T(14) = 42$ $\Rightarrow 4T(x_1) + 5T(x_2)$
$= 12 + 30 = 42$

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:

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

Matrix of Linear Transformation

- Example: Suppose T is a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^3$ such that
$$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \text{ and } T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}. \text{ With no additional information,}$$
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

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 4 \\ 3 \end{bmatrix}, \text{ and } T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 5 \\ 6 \end{bmatrix}.$$

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

$$\Rightarrow T(\mathbf{x}) = T\left(x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right) = x_1 T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) + x_2 T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) + x_3 T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right)$$

$$= x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 5 \\ 2 & 3 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = A\mathbf{x}$$

Linear Transformation in Neural Networks
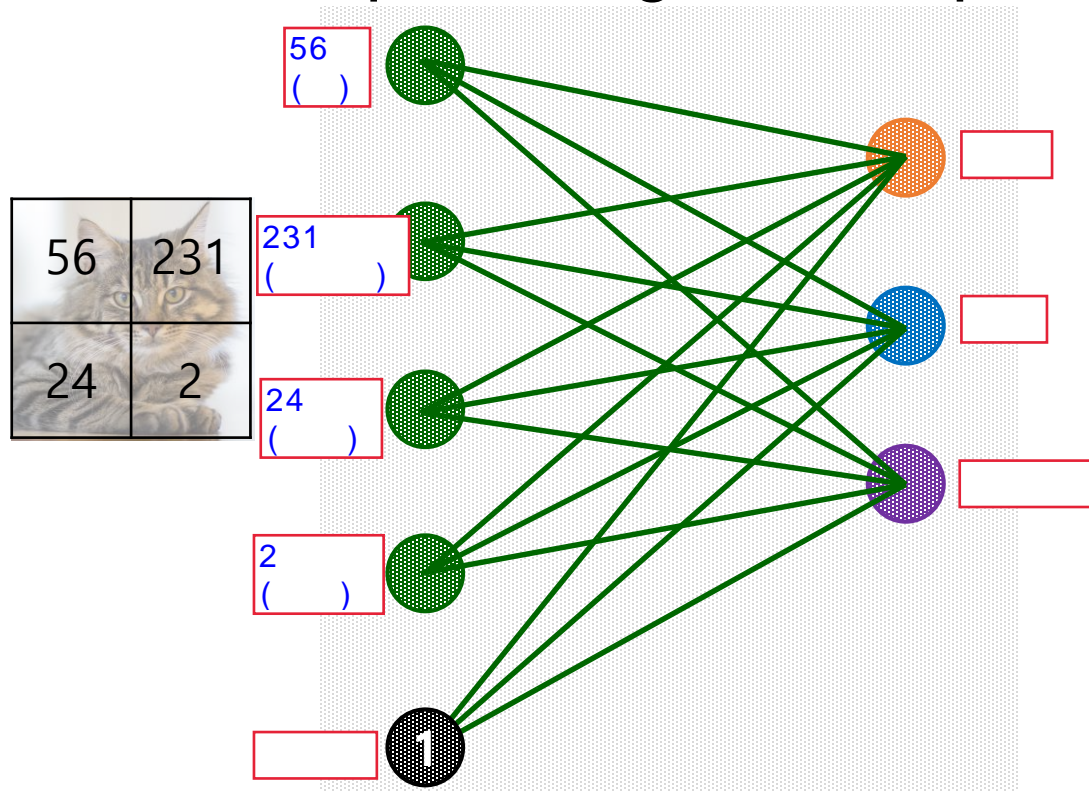
- Fully-connected layers (linear layer)

$$\mathbf{x} \xrightarrow[T_1]{} \mathbf{y}$$
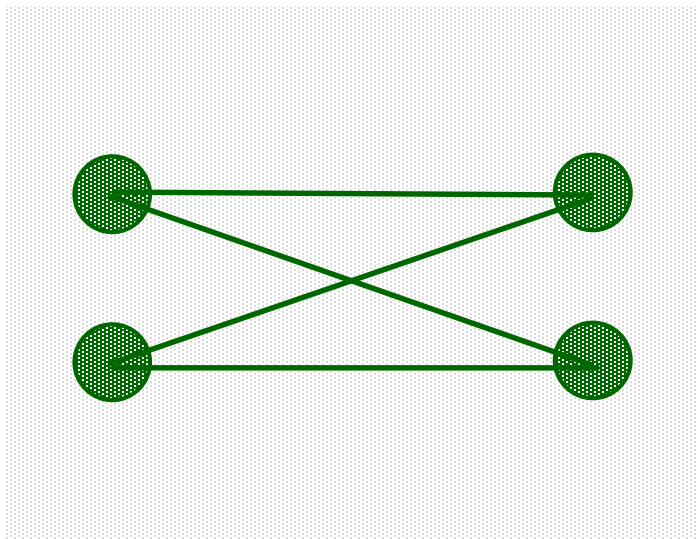
https://colah.github.io/posts/2014-03-NN-Manifolds-Topology/
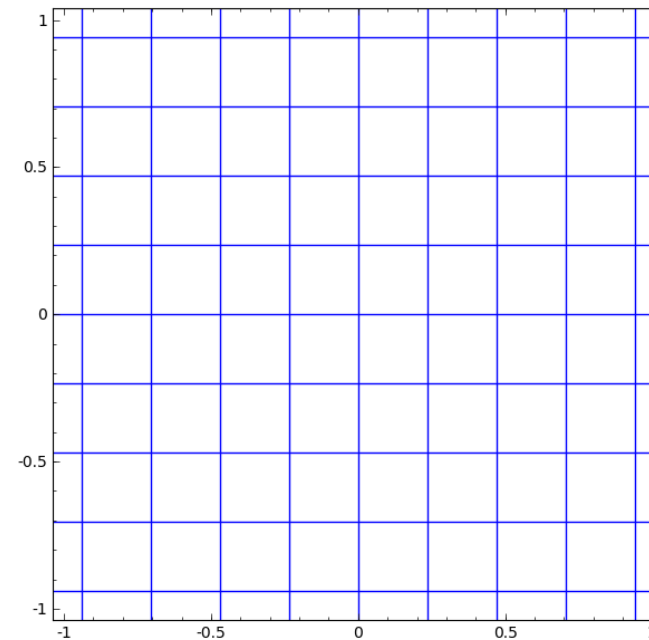
Affine Layer in Neural Networks

- 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)