Linear Algebra

주재걸
고려대학교 컴퓨터학과

Lecture Overview

- Elements in linear algebra

- Linear system

- <u>Linear combination, vector equation,</u>
  <u>Four views of matrix multiplication</u>

- Linear independence, span, and subspace

- Linear transformation

- Least squares

- Eigendecomposition

- Singular value decomposition

Linear Combinations

- Given vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p$ in $\mathbb{R}^n$ and given scalars $c_1, c_2, \cdots, c_p$,

$$c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$$

  is called a **linear combination** of $\mathbf{v}_1, \cdots, \mathbf{v}_p$ with **weights or coefficients** $c_1, \cdots, c_p$.

- The weights in a linear combination can be any real numbers, including zero.

From **Matrix** Equation to **Vector** Equation

- Recall the matrix equation of a linear system:

Person ID	Weight	Height	Is_smoking	Life-span
1	60kg	5.5ft	Yes (=1)	66
2	65kg	5.0ft	No (=0)	74
3	55kg	6.0ft	Yes (=1)	78

$$\begin{bmatrix} 60 & 5.5 & 1 \\ 65 & 5.0 & 0 \\ 55 & 6.0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 66 \\ 74 \\ 78 \end{bmatrix}$$

$$A \qquad \mathbf{x} = \mathbf{b}$$

- A matrix equation can be converted into a vector equation:

$$\begin{bmatrix} 60 \\ 65 \\ 55 \end{bmatrix} x_1 + \begin{bmatrix} 5.5 \\ 5.0 \\ 6.0 \end{bmatrix} x_2 + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} x_3 = \begin{bmatrix} 66 \\ 74 \\ 78 \end{bmatrix}$$

$$\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \mathbf{a}_3 x_3 = \mathbf{b}$$

Existence of Solution for $A\mathbf{x} = \mathbf{b}$

- Consider its vector equation:

$$\begin{bmatrix} 60 \\ 65 \\ 55 \end{bmatrix} x_1 + \begin{bmatrix} 5.5 \\ 5.0 \\ 6.0 \end{bmatrix} x_2 + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} x_3 = \begin{bmatrix} 66 \\ 74 \\ 78 \end{bmatrix}$$

$$\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \mathbf{a}_3 x_3 = \mathbf{b}$$

- When does the solution exist for $A\mathbf{x} = \mathbf{b}$?

Span

- **Definition**: Given a set of vectors $\mathbf{v}_1, \cdots, \mathbf{v}_p \in \mathbb{R}^n$, Span $\{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$ is defined as the set of all linear combinations of $\mathbf{v}_1, \cdots, \mathbf{v}_p$.

- That is, Span $\{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$ is the collection of all vectors that can be written in the form
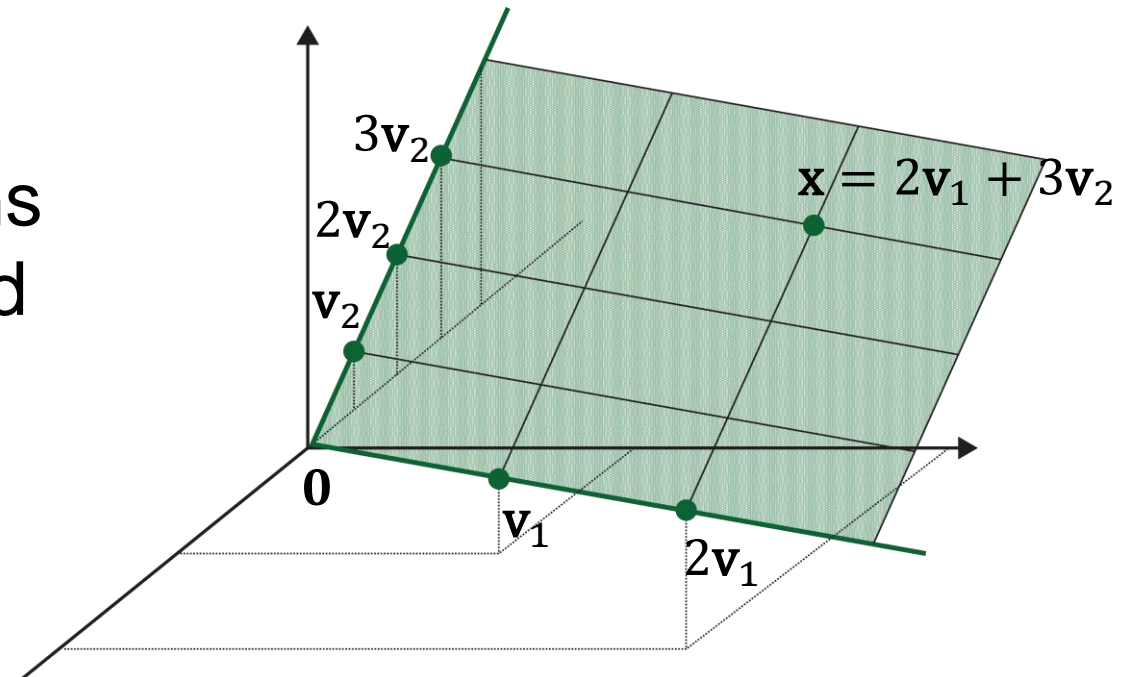
$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \cdots + c_p\mathbf{v}_p$$

  with arbitrary scalars $c_1, \cdots, c_p$.

- Span $\{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$ is also called the **subset of** $\mathbb{R}^n$ **spanned** (or **generated**) by $\mathbf{v}_1, \cdots, \mathbf{v}_p$.

Geometric Description of Span

- If $\mathbf{v}_1$ are $\mathbf{v}_2$ nonzero vectors in $\mathbb{R}^3$, with $\mathbf{v}_2$ not a multiple of $\mathbf{v}_1$, then Span $\{\mathbf{v}_1, \mathbf{v}_2\}$ is the plane in $\mathbb{R}^3$ that contains $\mathbf{v}_1, \mathbf{v}_2$ and $\mathbf{0}$.

- In particular, Span $\{\mathbf{v}_1, \mathbf{v}_2\}$ contains the line in $\mathbb{R}^3$ through $\mathbf{v}_1$ and $\mathbf{0}$ and the line through $\mathbf{v}_2$ and $\mathbf{0}$.

Geometric Interpretation of **Vector** Equation

- Finding a linear combination of given vectors $\mathbf{a}_1, \mathbf{a}_2,$ and $\mathbf{a}_3$ to be equal to $\mathbf{b}$:

$$\begin{bmatrix} 60 \\ 65 \\ 55 \end{bmatrix} x_1 + \begin{bmatrix} 5.5 \\ 5.0 \\ 6.0 \end{bmatrix} x_2 + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} x_3 = \begin{bmatrix} 66 \\ 74 \\ 78 \end{bmatrix}$$

$$\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \mathbf{a}_3 x_3 = \mathbf{b}$$

- The solution exists only when $\mathbf{b} \in \text{Span } \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$.

Matrix Multiplications as Linear Combinations of Vectors

- **Recall**: we defined matrix-matrix multiplications as the inner product between the row on the left and the column on the right:

- e.g., $\begin{bmatrix} 1 & 6 \\ 3 & 4 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 13 & 5 \\ 11 & 1 \\ 9 & -3 \end{bmatrix}$

- Inspired by the vector equation, we can view $A\mathbf{x}$ as a linear combination of columns of the left matrix:

- $\begin{bmatrix} 60 & 5.5 & 1 \\ 65 & 5.0 & 0 \\ 55 & 6.0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \mathbf{a}_3 x_3$

Matrix Multiplications as **Column** Combinations

- Linear combinations of columns
  - Left matrix: bases, right matrix: coefficients

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} 1 + \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} 2 + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} 3$$

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{bmatrix} = [\mathbf{x} \ \mathbf{y}]$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} 1 + \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} 2 + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} 3$$

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} (-1) + \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} 0 + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} 1$$

Matrix Multiplications as Row Combinations

- Linear combinations of rows of the right matrix
  - Right matrix: bases, left matrix: coefficients

One row on the left

$$[1 \quad 2 \quad 3]\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix} = \begin{matrix} 1 \times [1 \quad 1 \quad 0] \\ +2 \times [1 \quad 0 \quad 1] \\ +3 \times [1 \quad -1 \quad 1] \end{matrix}$$

Multiple rows on the left

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix} = \begin{bmatrix} \mathbf{x}^{\mathrm{T}} \\ \mathbf{y}^{\mathrm{T}} \end{bmatrix}$$

$$\mathbf{x}^{\mathrm{T}} = [x_1 \quad x_2 \quad x_3] = 1[1 \quad 1 \quad 0] + 2[1 \quad 0 \quad 1] + \quad 3[1 \quad -1 \quad 1]$$

$$\mathbf{y}^{\mathrm{T}} = [y_1 \quad y_2 \quad y_3] = 1[1 \quad 1 \quad 0] + 0[1 \quad 0 \quad 1] + (-1)[1 \quad -1 \quad 1]$$

Matrix Multiplications as **Sum of (Rank-1) Outer Products**

- (Rank-1) outer product

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$$

- Sum of (Rank-1) outer products

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \begin{bmatrix} 4 & 5 & 6 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix} + \begin{bmatrix} 4 & 5 & 6 \\ -4 & -5 & -6 \\ 4 & 5 & 6 \end{bmatrix}$$

Matrix Multiplications
as Sum of (Rank-1) Outer Products

- Sum of (Rank-1) outer products is widely used in machine learning
  - Covariance matrix in multivariate Gaussian
  - Gram matrix in style transfer