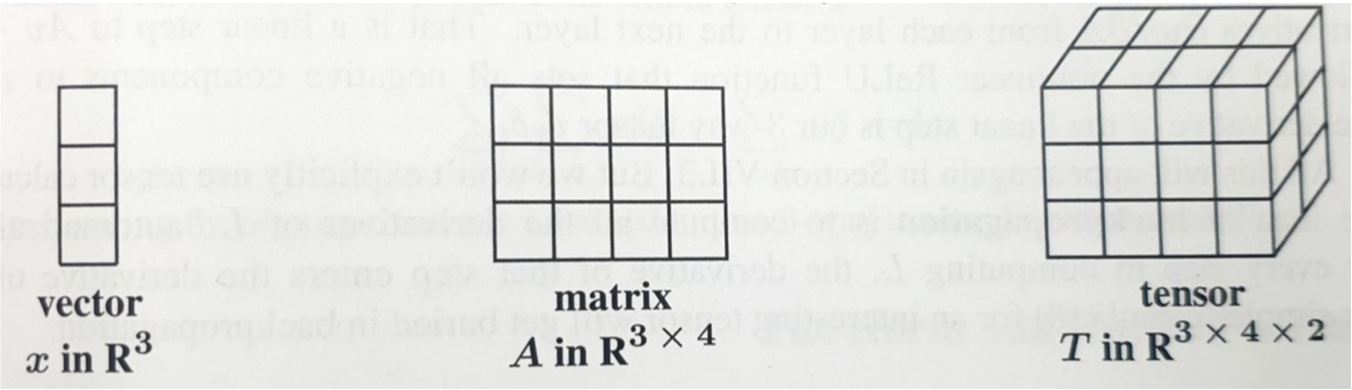
Linear Algebra, 선형대수(학), 線型代數學

- Algebra (대수학, 代數學)
  - 수를 대신하여 문자를 사용해 식을 전개하고 방정식을 푸는 방법을 연구하는 수학 분야
- Linear Algebra (선형대수학, 線型代數學)
  - 숫자(數)를 대신해서 선(線)의 형태(型)로 표현한다
  - 복잡한 비선형 문제를 간단한 선형 방정식으로 변환하여 문제를 해결
- Dimension of Data
  - Scalar, Vector, Matrix, Tensor



Highlights of Linear Algebra

- Goal: understanding even more than solving

  - $\mathbf{Ax} = \mathbf{b} \rightarrow$ Find $\mathbf{x}$ (Is the vector b in the column space of $\mathbf{A}$?)
  - $\mathbf{Ax} = \boldsymbol{\lambda}\mathbf{x} \rightarrow$ Find $\mathbf{x}$ and $\boldsymbol{\lambda}$
    - eigenvector directions so that $\mathbf{Ax}$ keeps the same direction as $\mathbf{x}$
    - solve anything linear when we know every $\mathbf{x}$ and $\boldsymbol{\lambda}$
  - $\mathbf{Av} = \boldsymbol{\sigma}\mathbf{u} \rightarrow$ Find $\mathbf{v}, \mathbf{u}$ and $\boldsymbol{\sigma}$
    - close but different, two vectors $\mathbf{u}$ and $\mathbf{v}$
    - $\mathbf{A}$: rectangular, full of data
    - what part of that data matrix is important?
    - Singular Value Decomposition: find its simplest pieces $\boldsymbol{\sigma}\mathbf{u}\mathbf{v}^T$
    - data science meets linear algebra in the SVD
    - Principal Component Analysis: find those pieces $\boldsymbol{\sigma}\mathbf{u}\mathbf{v}^T$
  - Minimize $\dfrac{\|\mathbf{Ax}\|^2}{\|\mathbf{x}\|^2}$
  - Factor the matix $\mathbf{A}$
  - $\rightarrow$ Factor $\mathbf{A} = (\text{columns})$ times (rows)

    $\rightarrow$
    - singular vectors and
    - compute
      - the best $\hat{\mathbf{x}}$ in the least squares
      - principal comonent $\mathbf{v}_1$ in PCA
    - fit the data

1. Multiplication Ax Using Columns of A

- Matrix-vector multiplication
  - Dot products: (row)·(column), computing, low level
  - Linear combination of the columns of A
- Independent columns = basis for the column space
- Combinations of the columns fill out the column space of A

$$A_1 = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix}, A_2 = \begin{bmatrix} 2 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 7 & 10 \end{bmatrix}, A_3 = \begin{bmatrix} 2 & 3 & 1 \\ 2 & 4 & 1 \\ 3 & 7 & 1 \end{bmatrix}$$

- rank of A = rank of C: number of independent columns
- = Dimension of the column space of A and C
  - Column rank = Row rank
- Meaning of Ax = b ?

- Factor A into C times R
  - R: row-reduced echelon form of A, rref(A)

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix} = \mathbf{CR}$$

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 5 \\ 1 & 2 & 5 \\ 1 & 2 & 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 5 \end{bmatrix}$$

subspaces of $\mathbb{R}^3$: independent column vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$

$$\begin{cases} r = 0: \text{ the zero vector } (0,0,0) \text{ itself} \\ r = 1: \text{ a line of all vectors } x_1\mathbf{a}_1 \\ r = 2: \text{ a plane of all vectors } x_1\mathbf{a}_1 + x_2\mathbf{a}_2 \\ r = 3: \text{ the whole } \mathbb{R}^3 \text{ with all vectors } x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 \end{cases}$$

2. Matrix-Matrix Multiplication AB

- AB = (m by n) times (n by p)
- Inner product: rows times columns
  - mp inner products, n multiplications each
- Outer product: columns times rows
  - n outer products, mp multiplications each

$$\mathbf{AB} = \begin{bmatrix} \cdot & \cdot & \cdot \\ a_{21} & a_{22} & a_{23} \\ \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot & \cdot & b_{13} \\ \cdot & \cdot & b_{23} \\ \cdot & \cdot & b_{33} \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & c_{23} \\ \cdot & \cdot & \cdot \end{bmatrix}$$

$$c_{23} = a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} = \sum_{k=1}^{3} a_{2k}b_{k3} \rightarrow c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

$$\mathbf{AB} = \begin{bmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{bmatrix} \begin{bmatrix} - & b_1^* & - \\ & \vdots & \\ - & b_n^* & - \end{bmatrix} = \underbrace{a_1 b_1^* + \cdots + a_n b_n^*}_{\text{sum of rank 1 matrices}}$$

$$\mathbf{u}\mathbf{v}^T = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 3 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 6 & 8 & 12 \\ 6 & 8 & 12 \\ 3 & 4 & 6 \end{bmatrix} = \text{rank one matrix}$$

$$\left(\mathbf{u}\mathbf{v}^T\right)^T = \begin{bmatrix} 6 & 8 & 12 \\ 6 & 8 & 12 \\ 3 & 4 & 6 \end{bmatrix}^T = \begin{bmatrix} 6 & 6 & 3 \\ 8 & 8 & 4 \\ 12 & 12 & 6 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 6 \end{bmatrix} \begin{bmatrix} 2 & 2 & 1 \end{bmatrix} = \mathbf{v}\mathbf{u}^T$$

row space of $\mathbf{u}\mathbf{v}^T$ is the line through $\mathbf{v}$

$$\begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \begin{bmatrix} 2 & 4 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 12 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 17 \end{bmatrix}$$

$(n \text{ by } n)$ times $(n \text{ by } n) \rightarrow n^3$ multiplications

$(m \text{ by } n)$ times $(n \text{ by } p) \rightarrow mnp$ multiplications

Insight from Column times Row

- Why is the outer product approach essential in data science?
  - We are looking for the important part of a matrix A
  - We don't usually want the biggest number in A
  - What we want more is the largest piece of A: those pieces are rank one matrices uv^T
- Factor A into CR and look at the pieces $c_k r_k^*$ of A=CR
- Five great factorizations

$$
\begin{cases}
\mathbf{A} = \mathbf{LU} \text{ from elimination} \\
\mathbf{A} = \mathbf{QR} \text{ from orthogonalization (Gram-Schmidt)} \\
\mathbf{S} = \mathbf{Q}\Lambda\mathbf{Q}^T \text{ from eigenvectors of a symmetric matrix} \\
\mathbf{A} = \mathbf{X}\Lambda\mathbf{X}^{-1} \text{ diagonalizes } \mathbf{A} \text{ by the eigenvector matrix } \mathbf{X} \\
\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T \text{ (orthogonal)(diagonal)(orthogonal)} \\
\quad = \text{Singular Value Decomposition}
\end{cases}
$$

3. Four Fundamental Subspaces m by n matrix A

- Column space C(A) in R^m, dim r
- Row space $C(A^T)$ in R^n, dim r
- Nullspace N(A), dim n-r
- Left nullspace $N(A^T)$, dim m-r

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} \xrightarrow{m=2,n=2,r=1} \begin{cases} \mathbf{Ax} = \mathbf{0} \to \mathbf{x} = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \to C\left(\mathbf{A}^T\right) \perp N\left(\mathbf{A}\right) \\ \mathbf{A}^T\mathbf{y} = \mathbf{0} \to \mathbf{y} = \begin{bmatrix} 3 \\ -1 \end{bmatrix} \to C\left(\mathbf{A}\right) \perp N\left(\mathbf{A}^T\right) \end{cases}$$

$$\mathbf{B} = \begin{bmatrix} 1 & -2 & -2 \\ 3 & -6 & -6 \end{bmatrix} \xrightarrow{m=2,n=3,r=1} \begin{Bmatrix} C\left(\mathbf{B}\right) \subset R^m \\ C\left(\mathbf{B}^T\right) \subset R^n \end{Bmatrix} \to \begin{cases} \mathbf{Bx} = \mathbf{0} : \left(n-r\right) \text{ independent solutions} \\ \mathbf{x}_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \to \text{orthogonal?} \end{cases}$$

The other two fundamental spaces come from the transpose matrix A^{T}. They are $\boldsymbol{N}(A^{\mathrm{T}})$ and $\boldsymbol{C}(A^{\mathrm{T}})$. We call $\boldsymbol{C}(A^{\mathrm{T}})$ the "row space of A" because the rows of A are the columns of A^{T}. What are those spaces for our 2 by 2 example?

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \quad \text{transposes to} \quad A^{\mathrm{T}} = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}.$$

Both columns of A^{T} are in the direction of $(1,2)$. The line of all vectors $(c,2c)$ is $\boldsymbol{C}(A^{\mathrm{T}})$ = row space of A. The nullspace of A^{T} is in the direction of $(3,-1)$:

Nullspace of A^{T} $$A^{\mathrm{T}}y = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} E \\ F \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{gives} \quad \begin{bmatrix} E \\ F \end{bmatrix} = \begin{bmatrix} 3c \\ -c \end{bmatrix}.$$

The four subspaces $\boldsymbol{N}(A), \boldsymbol{C}(A), \boldsymbol{N}(A^{\mathrm{T}}), \boldsymbol{C}(A^{\mathrm{T}})$ combine beautifully into the big picture of linear algebra. Figure A2 shows how the nullspace $\boldsymbol{N}(A)$ is perpendicular to the row space $\boldsymbol{C}(A^{\mathrm{T}})$. Every input vector x splits into a row space part x_r and a nullspace part x_n. Multiplying by A always(!) produces a vector in the column space. Multiplication goes from left to right in the picture, from x to $Ax = b$.

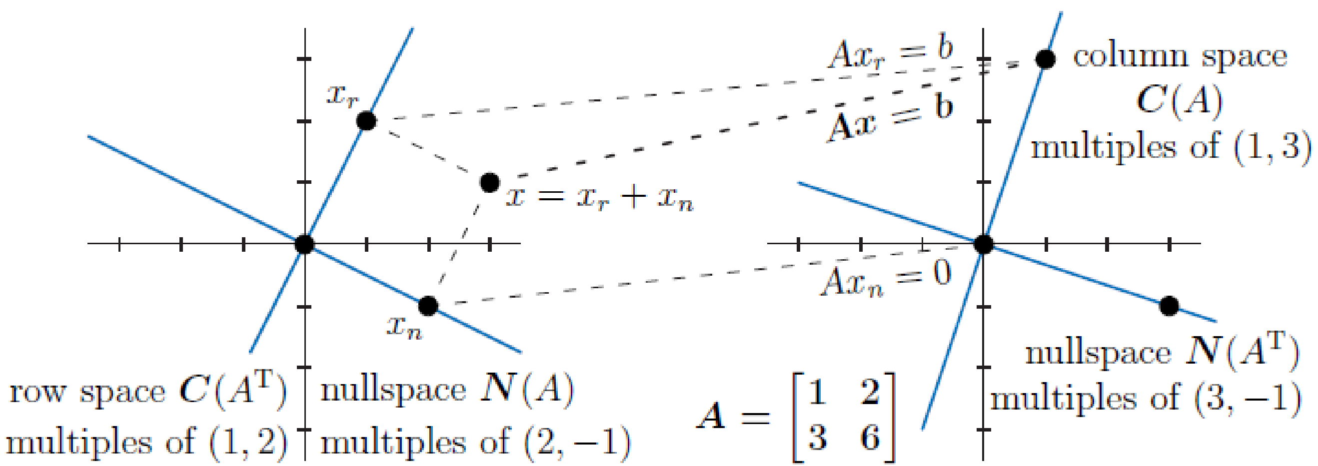
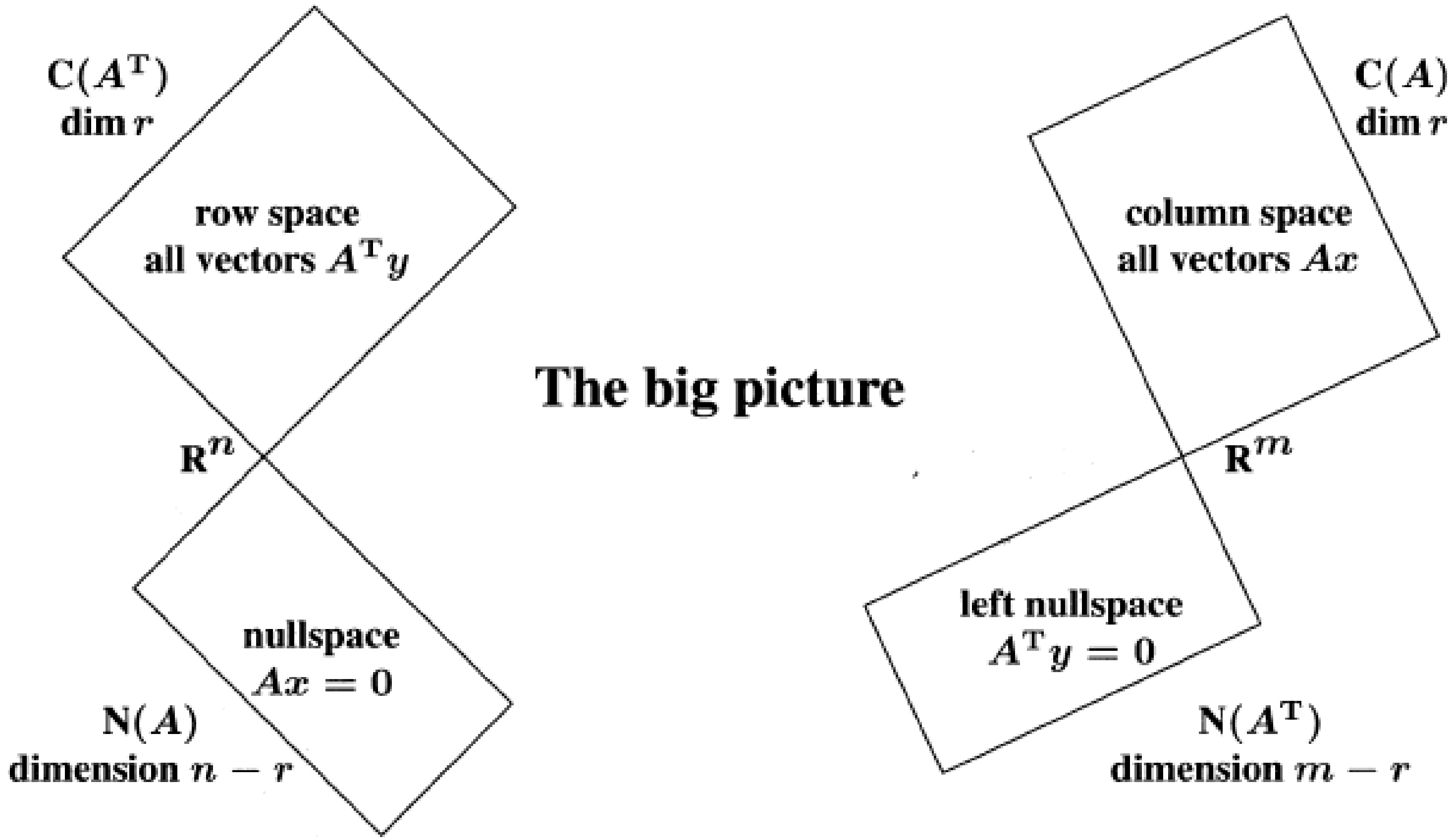


Figure A2: The four fundamental subspaces (lines) for the singular matrix A.


$\mathbf{C}(A^{\mathrm{T}})$
dim r
row space
all vectors $A^{\mathrm{T}}y$
The big picture
$\mathbf{R}^n$
nullspace
$Ax = 0$
$\mathbf{N}(A)$
dimension $n - r$
$\mathbf{C}(A)$
dim r
column space
all vectors Ax
$\mathbf{R}^m$
left nullspace
$A^{\mathrm{T}}y = 0$
$\mathbf{N}(A^{\mathrm{T}})$
dimension $m - r$