Mathematics for AI

Lecture 4
linear maps, matrix multiplication, matrix rank

Functions

functions / maps / tranformations

$f: X \to Y$

$\mathrm{Domain}(f) = X$

$\mathrm{Codomain}(f) = Y$

$\mathrm{Range}(f) = \{ f(x) \mid x \in X \}$

f: one-to-one (injective) $\qquad f(x) = f(y) \Rightarrow x = y$

یک به یک

f: onto (surjective) $\qquad \mathrm{Range}(f) = Y$

پوشا

$\forall y \in Y \ \exists x \in X : f(x) = y$

f: one-to-one & onto (bijective) $\Rightarrow$ Inrertible

یک به یک و پوشا $\qquad$ معکوس پذیر

Functions

K. N. Toosi
University of Technology

f: one-to-one & onto (bijective) $\Longrightarrow$ Invertible

یک به یک و پوشا

معکوس پذیر

bijective $\exists g$ such that ~~$f \circ g$~~ $g(f(x)) = x \quad \forall x \in X$

$\exists g \; Y \to X \qquad f(g(y)) = y \quad \forall y \in Y$

$g = f^{-1}$

Functions in linear algebra

- Here, we are interested in functions from a vector space V to a vector space U

$$(f: U \rightarrow V)$$

Linear Transformations

A linear map $f: V \longrightarrow U$

1. $f(u+v) = f(u)+f(v) \quad \forall u,v \in V$
2. $f(\alpha u) = \alpha f(u) \quad \forall u \in V, \alpha \in \mathbb{R}$

$\mathbb{C}$

$$1, 2 \iff f(\alpha \underline{u} + \beta \underline{v}) = \alpha f(u) + \beta f(v) \quad \forall u,v \in V$$

$$f(\alpha u) + f(\alpha v) \quad \forall \alpha,\beta \in \mathbb{R}$$

A linear map preserves linear combinations

Linear Transformations

K. N. Toosi
University of Technology

$f(u+v) = f(u) + f(v)$

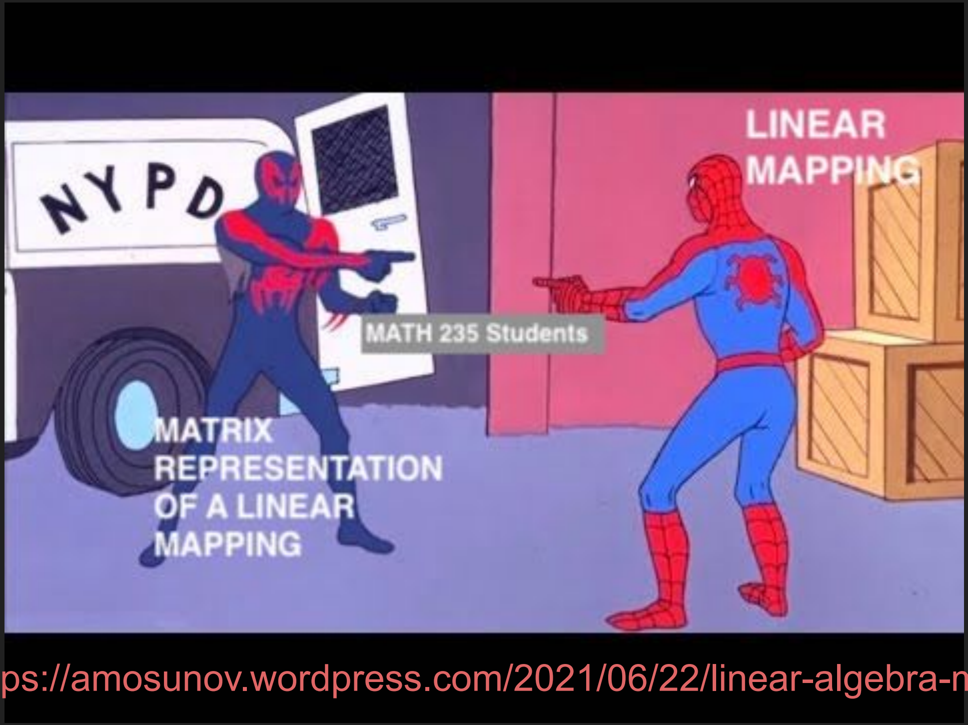
$f(a\,u) = a\,f(u)$

$\Leftrightarrow$

$f(a\,u + b\,v) = a\,f(u) + b\,f(v)$

does not matter if linear combination applied before or after transformation.

linear map <=> matrix representation



https://amosunov.wordpress.com/2021/06/22/linear-algebra-memes/

Dot Product as matrix product

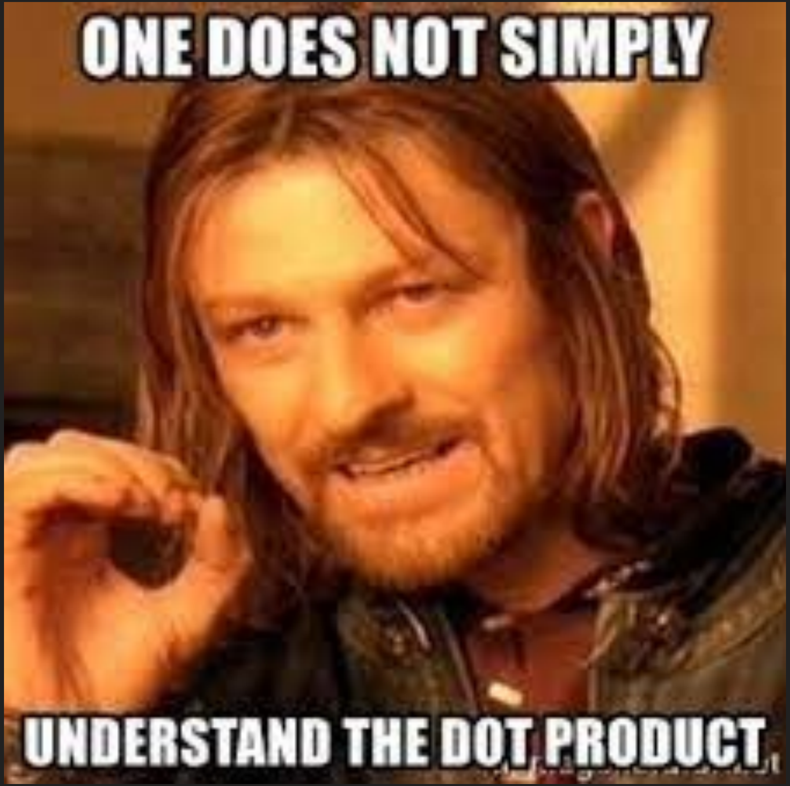
$$a = \underset{3\times 1}{\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}} \qquad b = \underset{3\times 1}{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}}$$

ضرب ماتریسی

$$a^T b = \overbrace{\underset{1\times 3}{\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}} \underset{3\times 1}{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}}} = \underset{1\times 1}{\langle a, b \rangle}$$

$$b^T a = a^T b = \langle a, b \rangle$$

Inner Product



K. N. Toosi
University of Technology

Inner Product

$$\langle u+w, v\rangle = \langle u, v\rangle + \langle w, v\rangle$$

$$\left\langle \begin{bmatrix} u_1+w_1 \\ u_2+w_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \right\rangle =$$

$$\langle \alpha u, v\rangle = \alpha \langle u, v\rangle$$

$$\langle u, v+w\rangle = \langle u, v\rangle + \langle u, w\rangle$$

$$\langle u, \alpha v\rangle = \alpha \langle u, v\rangle$$

$$\langle u, v\rangle = f(u, v)$$

$$\cancel{f(u,v)=} \quad f(u+w, v) = f(u, v) + f(w, v)$$

$$f(\alpha u, v) = \alpha f(u, v)$$

Inner Product

$$f(x,y) = f\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = xy$$

$$f\left(\alpha \begin{bmatrix} x \\ y \end{bmatrix}\right) = f\left(\begin{pmatrix} \alpha x \\ \alpha y \end{pmatrix}\right) = (\alpha x)(\alpha y) = \alpha^2 xy$$

$$\begin{bmatrix} u \\ v \end{bmatrix} \qquad f\left(\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} z \\ t \end{bmatrix}\right) = (x+z)(y+t) \neq xy + zt$$

$$f(u,v) = f\left(\begin{bmatrix} u \\ v \end{bmatrix}\right) = f\left(\begin{bmatrix} \alpha u \\ v \end{bmatrix}\right) = \alpha f\left(\begin{bmatrix} u \\ v \end{bmatrix}\right)$$

$$f\left(\begin{bmatrix} u+w \\ v \end{bmatrix}\right) = f\left(\begin{bmatrix} u \\ v \end{bmatrix}\right) + f\left(\begin{bmatrix} w \\ v \end{bmatrix}\right)$$

$f(x,y)$ is linear in x \
is linear in y $\Big\}$ bilinear \
is not linear in $\begin{pmatrix} x \\ y \end{pmatrix}$

$\langle u+w, v \rangle$

General vector spaces: Inner product space

$$\text{I}\quad \begin{cases} \langle \alpha u, v\rangle = \alpha \langle u, v\rangle \\ \langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle \end{cases}$$

$$\text{II}\quad \langle u, v\rangle = \langle v, u\rangle$$

$$\text{III}\quad \begin{cases} \langle u, u\rangle > 0 & u \neq 0 \\ \langle u, u\rangle = 0 & u = 0 \end{cases}$$

$$\langle 0, u\rangle = \langle v - v, u\rangle$$

$$f(u,v) \quad V \times V \longrightarrow \mathbb{R}$$

Matrix Multiplication in terms of inner products

K. N. Toosi
University of Technology

IV

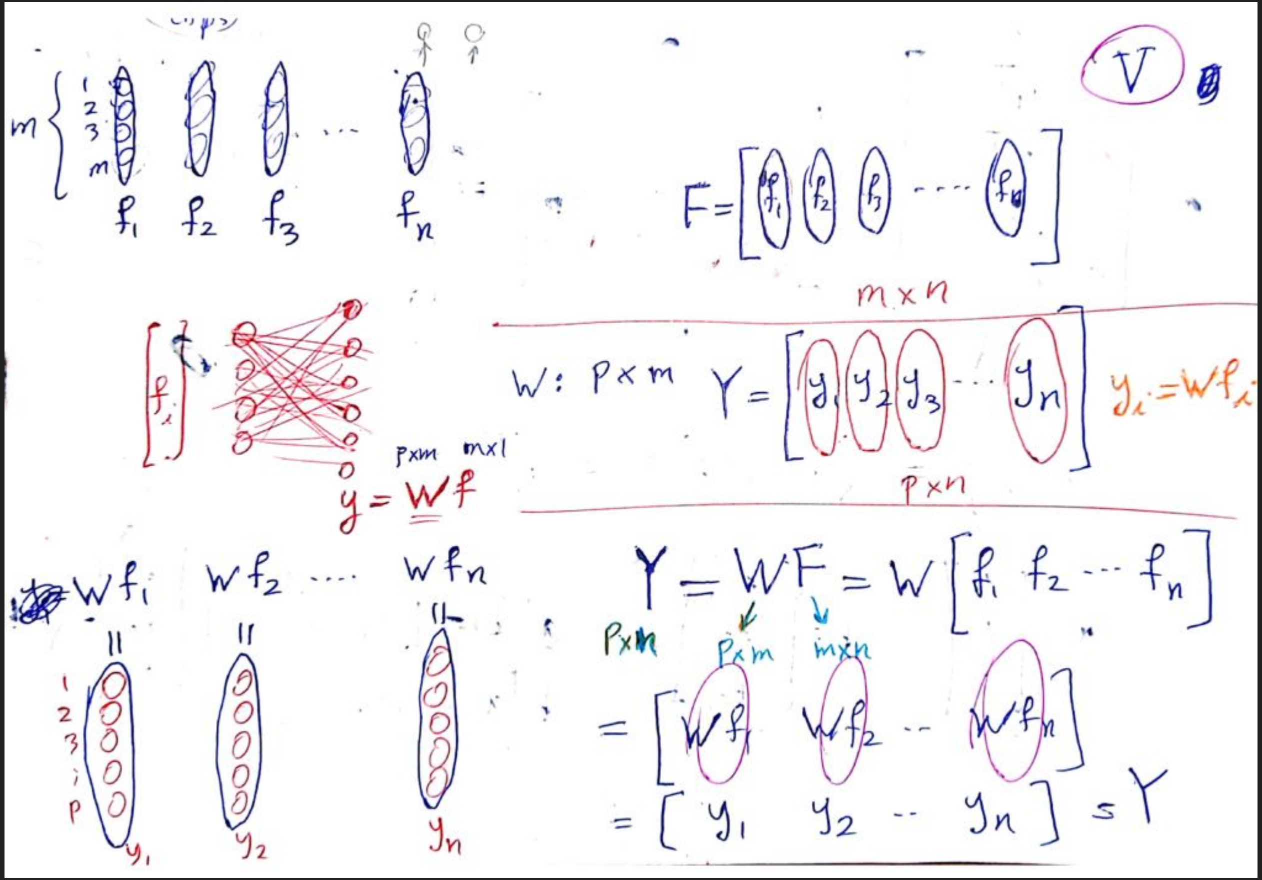
$$\underset{A}{\begin{array}{cc} & n \\ m & \begin{bmatrix} r_1^T \\ r_2^T \\ \vdots \\ r_m^T \end{bmatrix} \end{array}} \quad \underset{B}{\begin{array}{c} p \\ \begin{bmatrix} c_1 & c_2 & \cdots & c_p \end{bmatrix} \end{array}} = \underset{M}{\begin{array}{cc} & j \\ i & \begin{bmatrix} \; & r_i^T c_j & \; \end{bmatrix} \end{array}}$$

$$M_{ij} = \langle r_i , c_j \rangle = r_i^T c_j$$

ضرب ماتریس

Matrix multiplications in terms of matrix-vector product

K. N. Toosi
University of Technology



Transforming a bunch of points
data points as columns of a matrix

$$\underset{\substack{A \\ m \times p}}{\begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix}} \underset{\substack{B \\ p \times n}}{\begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix}} = \underset{\substack{C \\ m \times n}}{\begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}}$$

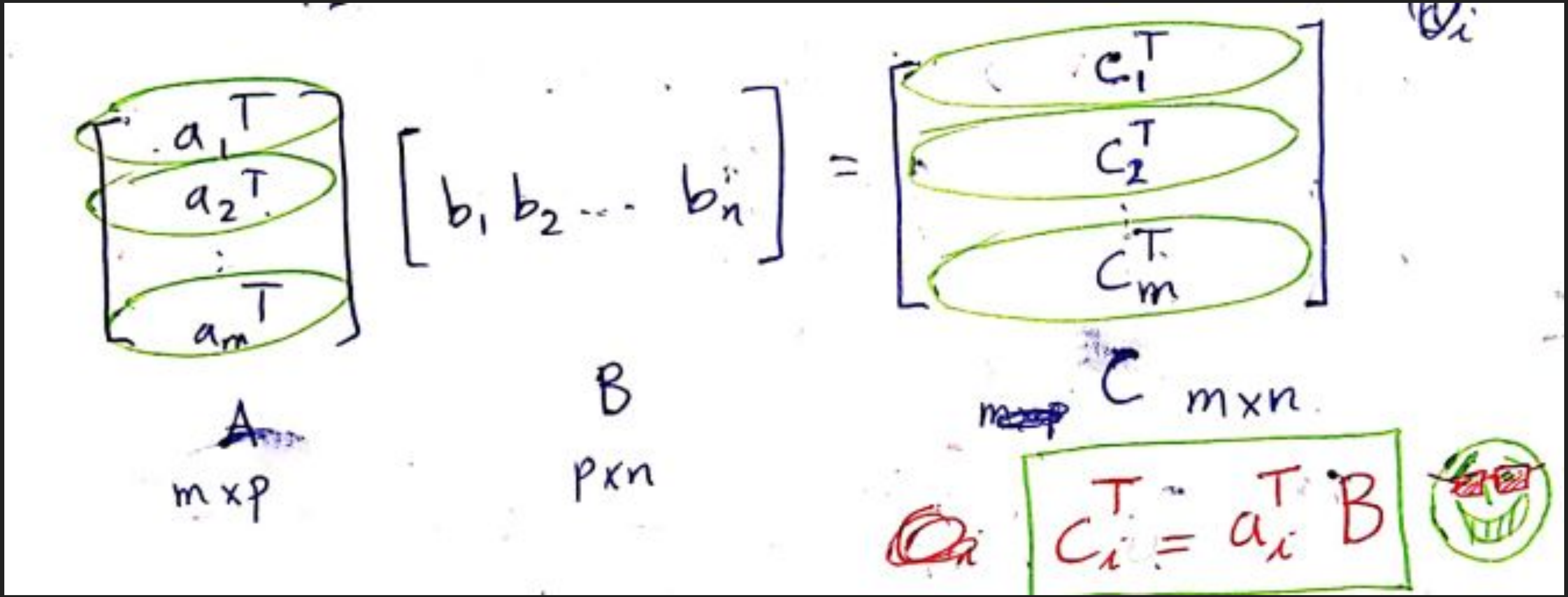
VI

$$\boxed{c_i = A b_i}$$

$$\begin{bmatrix} a_1^T \\ \vdots \end{bmatrix}$$

Transforming a bunch of points data points as rows of a matrix

K. N. Toosi
University of Technology

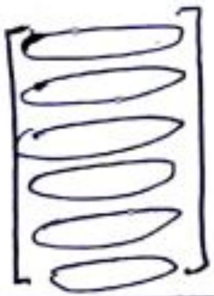


be careful about linear transformations on row vectors!

K. N. Toosi
University of Technology

$$R\begin{bmatrix}x\\ y\end{bmatrix} = \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}$$

$$R\left[\begin{bmatrix}x_1\\ y_1\end{bmatrix}\ \begin{bmatrix}x_2\\ y_2\end{bmatrix}\ \begin{bmatrix}x_3\\ y_3\end{bmatrix}\right]$$



W^T

$$\begin{bmatrix}x_1 & y_1\\ x_2 & y_2\\ \vdots & \\ x_n & y_n\end{bmatrix} R^T = \begin{bmatrix}x_1 & y_1\\ x_2 & y_2\\ \vdots & \\ x_n & y_n\end{bmatrix}\begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix}$$

be careful about linear transformations on row vectors!

$$\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_n & y_n \end{bmatrix} R^T = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_n & y_n \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

$$R \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$R \left[\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \begin{bmatrix} x_3 \\ y_3 \end{bmatrix} \right]$$

Outer Product

np.outer(u,v)

u @ v.T

How many independent columns?

How many independent rows?

outer(u,v) = outer(v,u).T

complex numbers?

Outer Product

K. N. Toosi
University of Technology

Outer product

VII

$$u = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \qquad v = \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix}$$

$u \in \mathbb{R}^3$ $\quad v \in \mathbb{R}^4$

$$3\ \begin{array}{c} \\ a \\ b \\ c \end{array} \begin{array}{c} \begin{array}{cccc} x & y & z & t \end{array} \\ \begin{bmatrix} ax & ay & az & at \\ bx & by & bz & bt \\ cx & cy & cz & ct \end{bmatrix} \end{array} = \begin{bmatrix} a(x,y,z,t) \\ b(x,y,z,t) \\ c(x,y,z,t) \end{bmatrix}$$

4

$W \in \mathbb{R}^{3\times 4}$

$$\boxed{W = u \otimes v}$$

$\in \mathbb{R}^{3\times 1}$

Outer Product

$$\boxed{W = u \otimes v}$$

$$u \in \mathbb{R}^{3\times 1}$$

$$\therefore\ v \otimes u = W^T$$

$$W = \left[\ x\begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad y\begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad z\begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad t\begin{bmatrix} a \\ b \\ c \end{bmatrix}\ \right] = \left[\ xu \quad yu \quad zu \quad tu\ \right]$$

$$W = \underset{3\times 1}{\begin{bmatrix} a \\ b \\ c \end{bmatrix}} \underset{1\times 4}{\begin{bmatrix} x & y & z & t \end{bmatrix}} = \overset{3\times 1}{u}\ \overset{1\times 4}{v^T} = \overset{3\times 4}{W}$$

ضرب تانسری

$$u \in \mathbb{R}^{3\times 1} \quad v \in \mathbb{R}^{4\times 1}$$

Inner product vs outer product

inner product $u^T v \in \mathbb{R}$

$u, v \in \mathbb{R}^{n\times 1}$

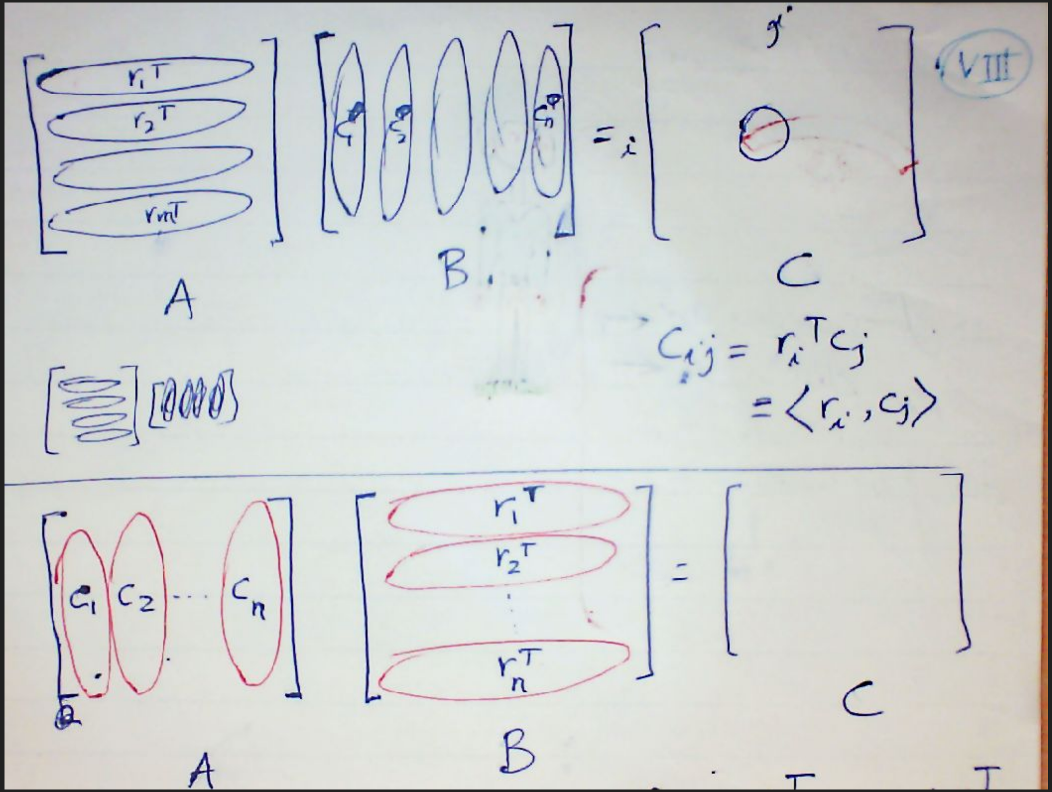
$u \cdot v = \langle u, v \rangle = u^T v$

outer prodct $uv^T \in \mathbb{R}^{m\times n}$

$u \in \mathbb{R}^{m\times 1}$ $v \in \mathbb{R}^{n\times 1}$

$u \otimes v = u v^T$;

Two ways of looking at matrix product



Matrix Multiplication in terms of outer products

$$\begin{bmatrix} a & b \\ d & e \end{bmatrix} \begin{bmatrix} x & y \\ z & t \end{bmatrix} = \boxed{\begin{bmatrix} ax+bz & ay+bt \\ dx+ez & dy+et \end{bmatrix}}$$

$$\begin{bmatrix} ax & ay \\ dx & dy \end{bmatrix} + \begin{bmatrix} bz & bt \\ ez & et \end{bmatrix} = \begin{bmatrix} a \\ d \end{bmatrix} \begin{bmatrix} x & y \end{bmatrix} + \begin{bmatrix} b \\ e \end{bmatrix} \begin{bmatrix} z & t \end{bmatrix}$$

$$\begin{bmatrix} a \\ d \end{bmatrix} \otimes \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} b \\ e \end{bmatrix} \otimes \begin{bmatrix} z \\ t \end{bmatrix}$$

Matrix Multiplication in terms of outer products

K. N. Toosi
University of Technology

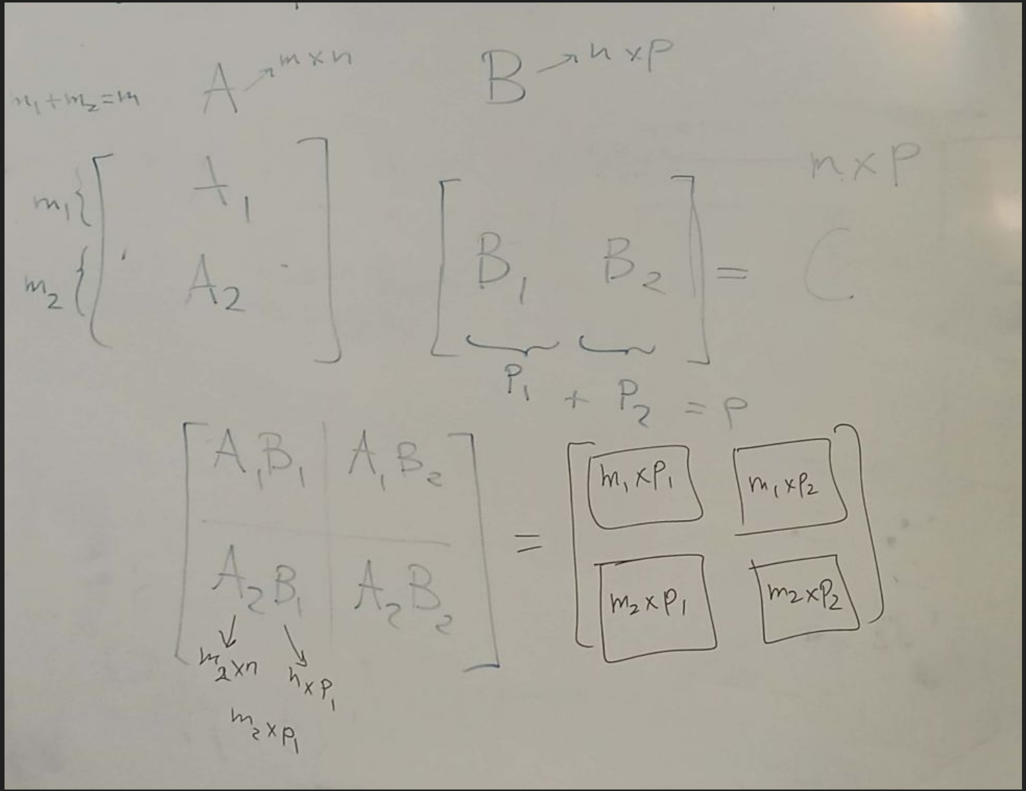
$$\underbrace{\begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} r_1^T \\ r_2^T \\ \vdots \\ r_n^T \end{bmatrix}}_{B} = \underbrace{\begin{bmatrix} \quad \end{bmatrix}}_{C}$$

$$C = c_1 r_1^T + \cdots + c_n r_n^T$$

$$C = \sum_{i=1}^{n} c_i r_i^T$$

$$= \sum_{i=1}^{n} c_i \otimes r_i$$

Block-wise multlipication



Block-wise multlipication

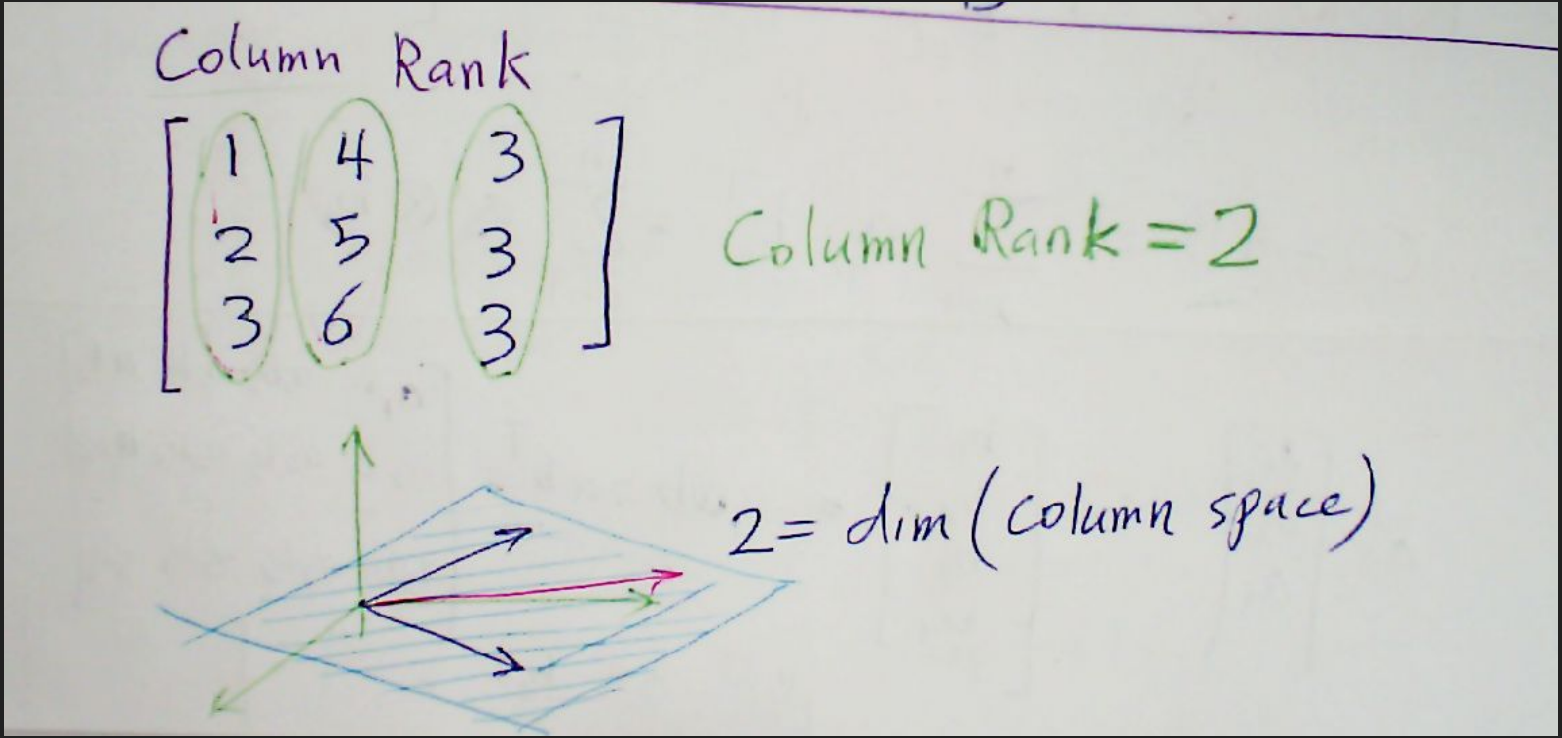
$m_1 \times n_1$

A

$$\begin{bmatrix} A_1 & A_2 \\ \underbrace{A_3}_{n_1} & \underbrace{A_4}_{n_2} \end{bmatrix} \begin{array}{c} n_1\{ \\ n_2\{ \end{array} \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$$

$$\begin{bmatrix} A_1B_1 + A_2B_3 & A_1B_2 + A_2B_4 \\ A_3B_1 + A_4B_3 & A_3B_2 + A_4B_4 \end{bmatrix}$$

Column Rank


Column Rank
$\begin{bmatrix} 1 & 4 & 3 \\ 2 & 5 & 3 \\ 3 & 6 & 3 \end{bmatrix}$
Column Rank = 2
2 = dim (column space)

Column Rank and Row Rank

K. N. Toosi
University of Technology

$$\begin{bmatrix} 1 & -2 \\ 2 & -4 \\ 3 & -6 \\ 4 & -8 \end{bmatrix}$$

column rank = 1

row rank = 1

$$\begin{bmatrix} 1 & 0 & 4 \\ 2 & 0 & 5 \\ 3 & \underline{\underline{0}} & 6 \end{bmatrix}$$

column rank = 2

row rank = 2

Column Rank and Row Rank

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

column rank = 3

row rank = 3

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

colum rank = 0

row rank = 0

Column Rank and Row Rank

K. N. Toosi
University of Technology

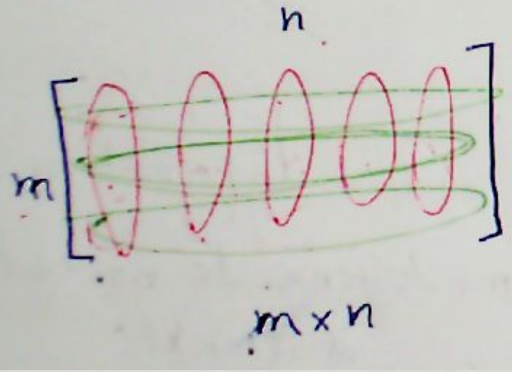
$$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \qquad \vec{a} \otimes \vec{b} = \vec{a}\,\vec{b}^T = \begin{bmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 & a_1 b_4 \\ a_2 b_1 & a_2 b_2 & a_2 b_3 & a_2 b_4 \\ a_3 b_1 & a_3 b_2 & a_3 b_3 & a_3 b_4 \end{bmatrix}$$

$$a \otimes b = a\,b^T = \begin{bmatrix} b_1 \vec{a} & b_2 \vec{a} & b_3 \vec{a} & b_4 \vec{a} \end{bmatrix} = \begin{bmatrix} a_1\, b^T \\ a_2\, b^T \\ a_3\, b^T \end{bmatrix}$$

(at most) 1 independent column 1 independent row

$$\text{RowRank}(a\ b^T) = \text{ColumnRank}(a\ b^T) \leq 1$$

Column Rank and Row Rank



$\leqslant m$

column rank = dim(column space) $\leqslant n$

row rank = dim(row space) $\leqslant m$

$\leqslant n$

column rank $\leqslant \min(m, n)$

row rank $\leqslant \min(m, n)$

Column Rank and Row Rank

K. N. Toosi
University of Technology

$$\begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 3 & 5 & 1 \\ 3 & 4 & 7 & 2 \end{bmatrix} = \underset{3\times 2}{\begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix}} \underset{2\times 4}{\begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & -1 \end{bmatrix}}$$

Matrix Factorization

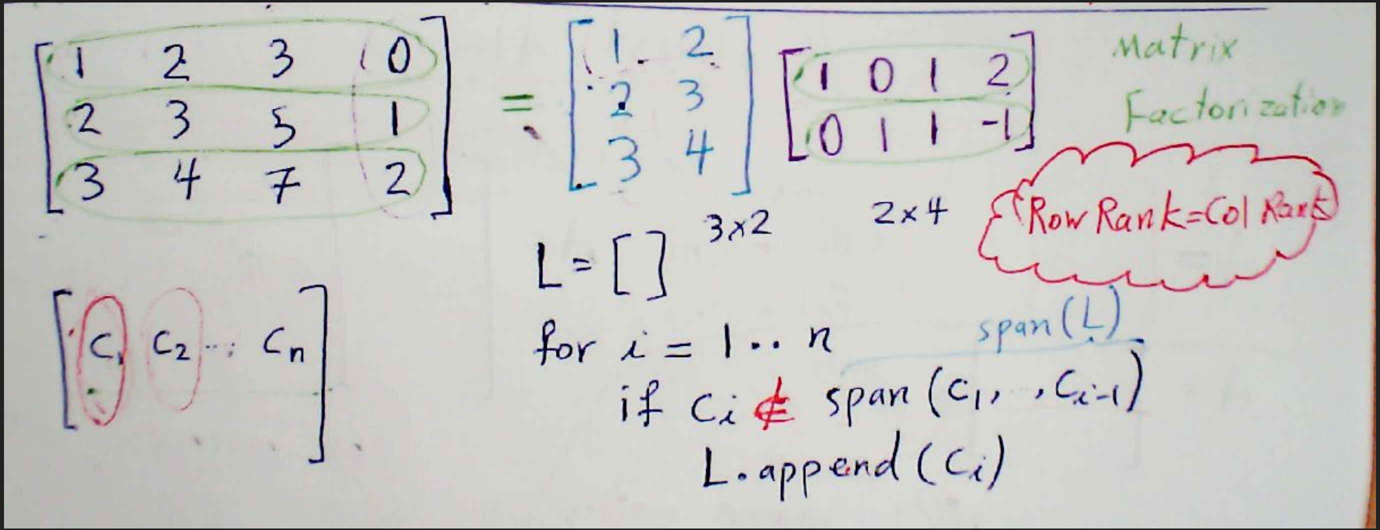
Row Rank = Col Rank

$$\begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}$$

```
L = []
for i = 1..n
    if c_i ∉ span(c_1, ..., c_{i-1})      span(L)
        L.append(c_i)
```

* Column Rank and Row Rank

K. N. Toosi
University of Technology



Column Rank = Row Rank

Column Rank and Row Rank

$$\begin{array}{cc} & n \\ m & \begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \end{array} = \begin{array}{cc} & r \\ m & \begin{bmatrix} \, \cdot \, \end{bmatrix} \end{array} \begin{array}{cc} & n \\ r & \begin{bmatrix} \, \cdot \, \end{bmatrix} \end{array}$$

a basis for column space. a basis for row space

IV

column rank $= r \leq \min(m,n)$

$$\boxed{\begin{aligned} \text{Row Rank}(A) &= \text{Column Rank}(A) \\ &\overset{\text{def}}{=} \text{Rank}(A) \end{aligned}}$$

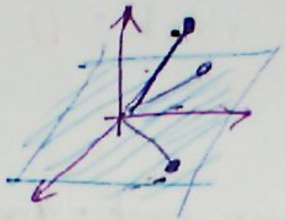
رتبه .

$$\text{Rank}(A) = \dim(\text{Colspace}(A)) = \dim(\text{RowSpace}(A))$$

"Most" matrices have full rank

$$\begin{array}{cc} & 3 \\ 3 & \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix} \end{array}$$

A



if A is a random 3×3 matrix (uniform distribution / nondegenerate normal distribution)

$\text{rank}(A) = 3$ with probability $= 1$

almost Always

thin and fat matrices

$A = \begin{bmatrix} \; \end{bmatrix} \in \mathbb{R}^{m \times n}$ $\quad m < n$

Fat matrix چاق

$A = \begin{bmatrix} \; \end{bmatrix} \in \mathbb{R}^{m \times n}$ $\quad m > n$

thin matrix لاغر

$A = \begin{bmatrix} \; \end{bmatrix} \in \mathbb{R}^{m \times n}$

$m = n$

مربع Square matirx

full-rank and rank-deficient

K. N. Toosi
University of Technology

V 9

$A \in \mathbb{R}^{m \times n}$ $\quad \text{rank}(A) \leq \min(m,n)$

رتبه کامل

$$\begin{cases} \text{rank}(A) = \min(m,n) & \text{full-rank} \\ \text{rank}(A) < \min(m,n) & \text{rank-deficient} \end{cases}$$

رتبه ناقص

A is full-rank

A has full rank

full-rank and rank-deficient

$A \in \mathbb{R}^{m \times n}$ $\qquad$ $\text{rank}(A) \leq \min(m,n)$

$$\begin{cases} \text{rank}(A) = \min(m,n) & \text{full-rank} \\ \text{rank}(A) < \min(m,n) & \text{rank-deficient} \end{cases}$$

رتبه کامل

رتبه ناقص

A is full-rank

A has full rank

$A \in \mathbb{R}^{n \times n}$

$\text{rank}(A) = n$ $\quad$ full-rank

$\text{rank}(A) < n$ $\quad$ rank-deficient