CS 598 mat Algebraic and Geometric Complexity Theory

Lecture 8 (2023-02-16)

logistics: lectures / pset 2 uploaded tonight

last lecture: - non-scalar vs total complexity
- bilinear rank vs multiplicative complexity

today: tensor rank

idea [Strassen]: better matrix mult also via minimizing non-scalar mult

prop: f computable w/ s non-scalar mult
$\Rightarrow f \xrightarrow{\quad O(s) \quad}$ w/ $O(s^2)$ ckt size

prop: $f_1, \dots, f_n$ bilinear forms over $\mathbb{F}[\bar{x}, \bar{y}]$ computable w/ s non-scalar $\Rightarrow$ $\bar{f}$ computable w/ $\leq 2s$ bilinear non-scalar mult

Q: what about addition?

prop: suppose $n \times n$ matrix mult has bilinear rank $\leq n^{\omega}$ for $n = m$, $\omega > 2$
$\Rightarrow$ $n \times n$ matrix mult has $O_m(n^{\omega})$ size ckt
depend on m

idea: recursion

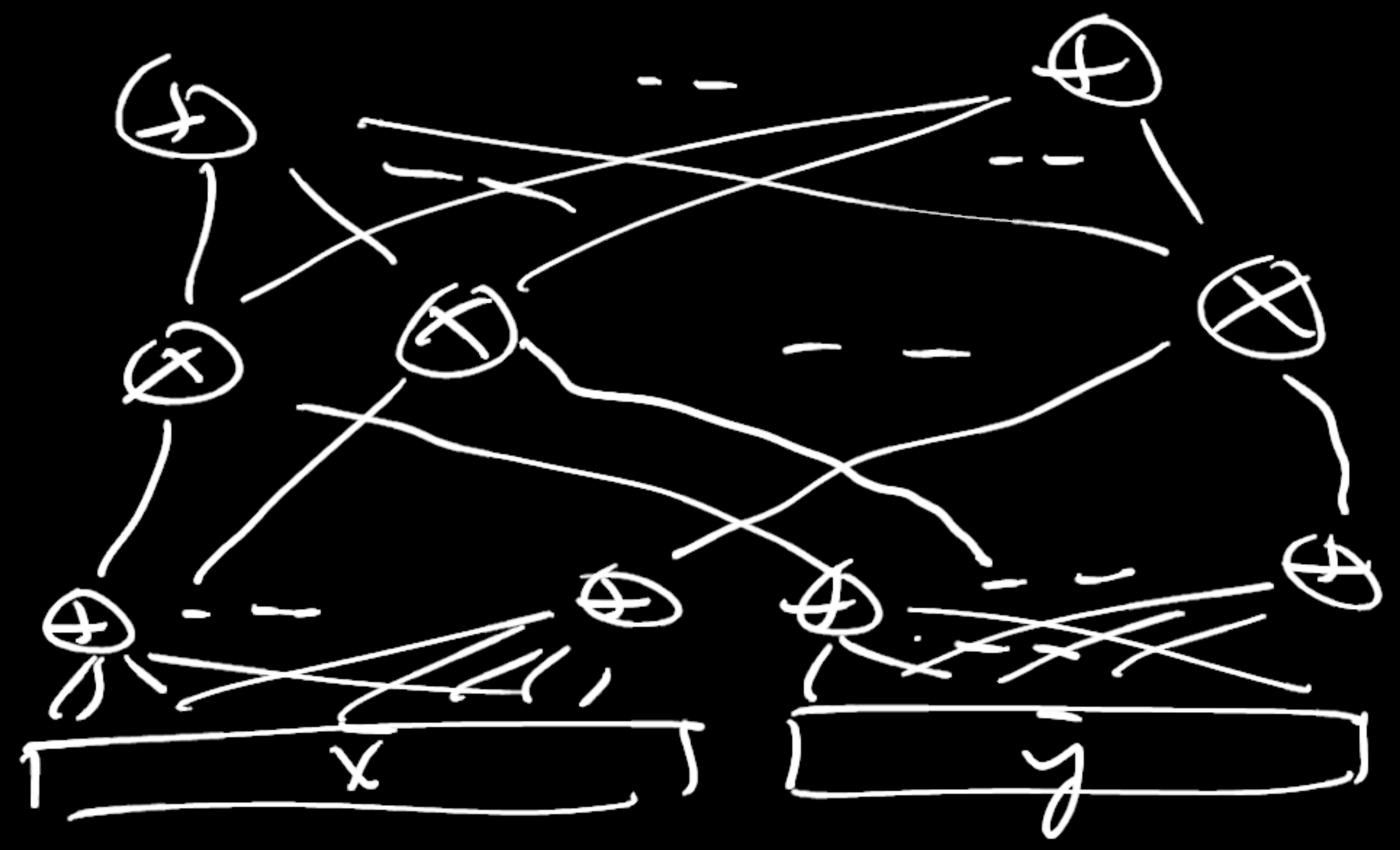
pf. multiplication of $n \times n$ matrices,

$n = m^i$

algo: decompose into $m \times m$ block matrix of size $\underbrace{m^{i-1} \times m^{i-1}}$

use m^ω rank also to multiply block matrices

- m^ω non-scalar mult $\Rightarrow$ m^ω recursive calls to mult
- $O_m(1)$ many additions/scalar mult on blocks



correctness: clear

complexity:

$$T(n) \le m^\omega T(n/m) + O_m(1) \cdot (m^{i-1})^2$$

recursive calls

per-call time

# of add/scalar mult

cost of add/scalar mult

$$\le \cdots \le O_m(n^\omega)$$

if $\omega > 2$

for $n \neq m^i$ can pad matrices w/ zeroes to size $m^{\lceil \log_m n \rceil}$ w/ no asymptotic penalty rmk: - constants may be poor

- rank $m \times m$ matrix mult to be $O(m^{\omega})$

$\Rightarrow$ rank $n \times n$ $O(n^{\omega+\varepsilon})$ any $\varepsilon > 0$

def: the exponent of matrix multiplication is $\omega := \inf \{ \log_n \text{rank}(n \times n \text{ matrix mult}) \}_n$

cor: $n \times n$ matrix mult has circuit size $O(n^{\omega+\varepsilon})$ any $\varepsilon > 0$

$\omega \leq 3$

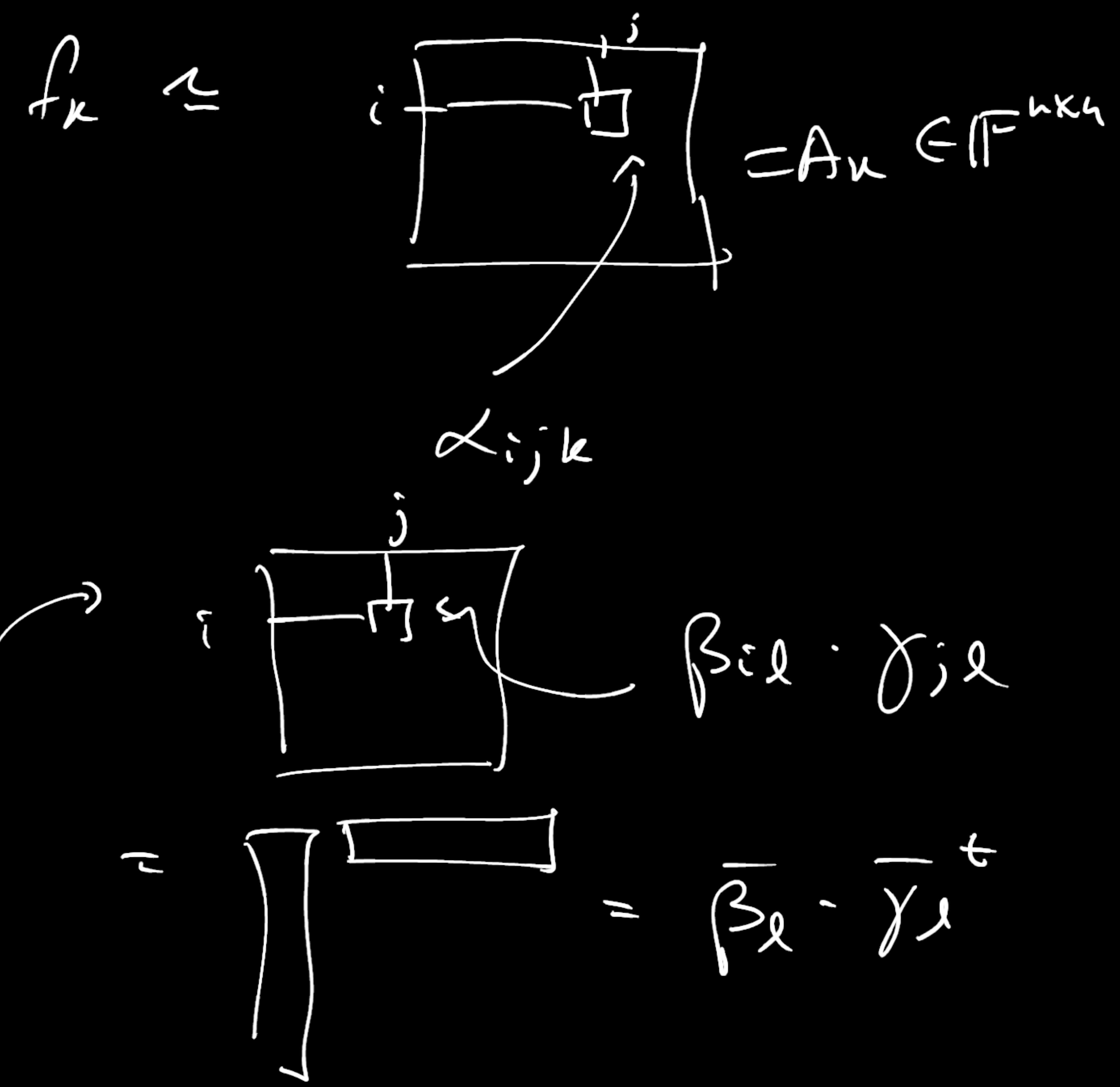
$\omega \leq 2.81\ldots$

fact: $\omega \leq 2.3728\ldots$

Q: what is ω? $\omega \geq 2$, $\omega = 2$?

Q: what is bilinear rank?

$$f_k = \sum_{ij} \alpha_{ijk} x_i y_j \in \text{span}\{ \underbrace{(\textstyle\sum_i \beta_{il} x_i)(\textstyle\sum_j \gamma_{jl} y_j)}_{\sum_{ij} \beta_{il}\gamma_{jl} x_i y_j} \}_{l \leq r}$$



$\Rightarrow$ rank$(\bar f) \leq r$ iff all A_k in span of a set of r outer products

rank r

fact: a matrix $M \in \mathbb{F}^{n\times n}$ is rank $\le r$

iff $\le r$ nonzero rows in reduced row echelon form

iff $\dim$ row-span $M \le r$

iff ——— col ———

iff ——— cols ——— col ———

iff all $(r+1)\times(r+1)$ submatrices have $\det = 0$

lem: $\bar\beta\,\bar\gamma^t$ is (a) zero $\Leftrightarrow$ rank $= 0$

(b) nonzero $\Leftrightarrow$ rank $= 1$

sketch: (a) easy

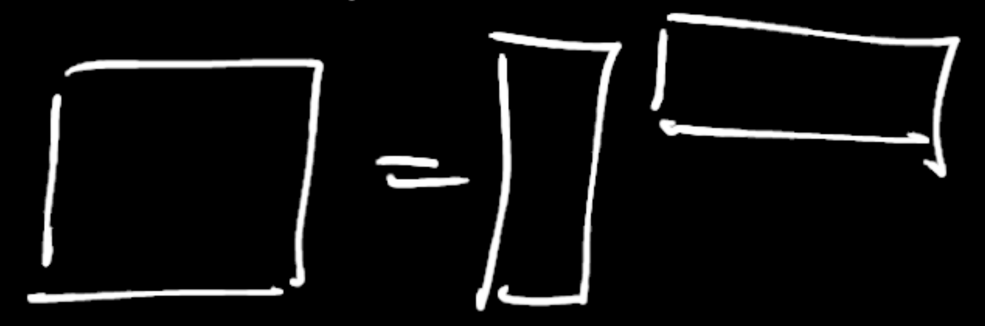
(b) $\bar\beta\,\bar\gamma^t$ nonzero $\Rightarrow$ all rows are multiples of $\bar\gamma^t$ $\Rightarrow$ rank ≤ 1

$\bar\beta\,\bar\gamma^t \ne 0 \Rightarrow$ rank > 0 ∎

prop: $\mathrm{rank}(M) \le r$ iff $M = UV$

$U \in \mathbb{F}^{n\times r}$

$V \in \mathbb{F}^{r\times n}$

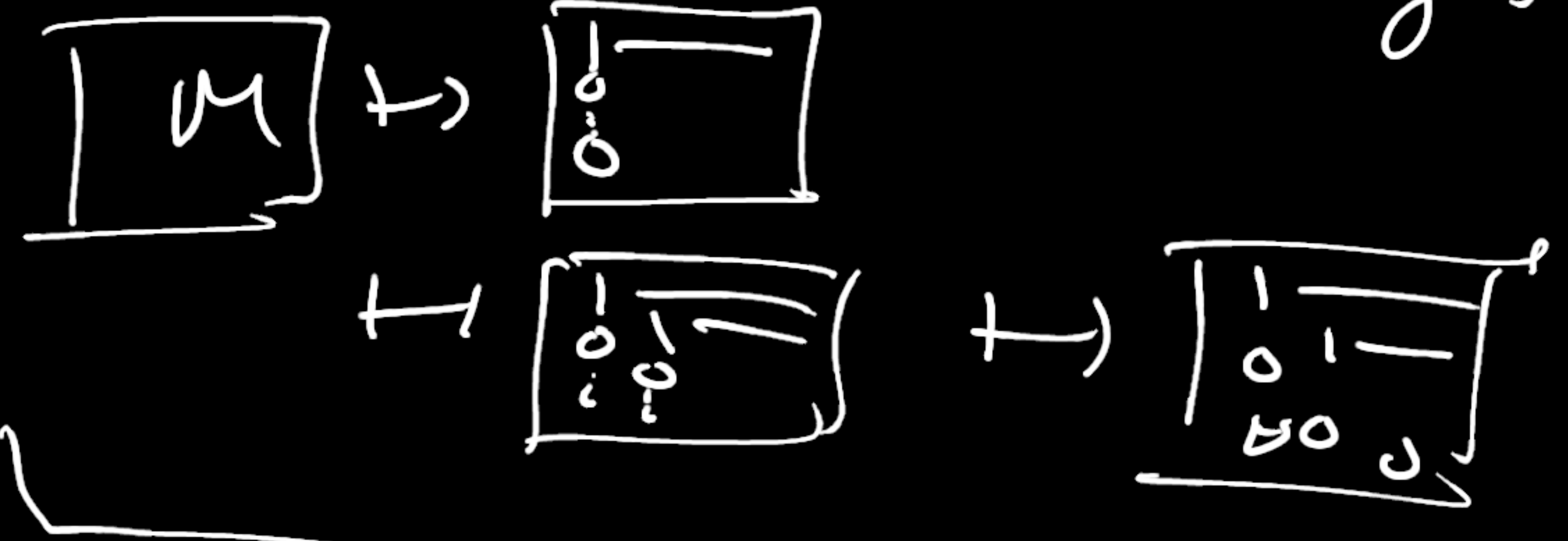


sketch: $\Leftarrow$: row span$(M) \subseteq$ rowspan(V)

view mult by U as row ops

dim $\le r$ by thm

$\Rightarrow$:



$M = UV$

row ops are invertible

$\le r$ nonzero rows is V

∎

$$\begin{bmatrix}1&2\\3&4\end{bmatrix} \xrightarrow{R_2 - 3R_1} \begin{bmatrix}1&2\\0&-2\end{bmatrix}$$

$$M = \begin{bmatrix}1&2\\3&4\end{bmatrix}, \qquad \begin{bmatrix}1&2\\0&-2\end{bmatrix} = \begin{bmatrix}1&0\\-3&1\end{bmatrix}\begin{bmatrix}1&2\\3&4\end{bmatrix}$$

Prop. $M = UV$, $U \in \mathbb{F}^{n\times r}$, $V \in \mathbb{F}^{r\times n}$

iff $M = \sum_{i=1}^{r} \bar{u}_i \cdot \bar{v}_i$

$\bar{u}_i$: ith col of U; $\bar{v}_i$: ith row of V

pf.

$$M_{ij} = \sum_k \underbrace{U_{ik} V_{kj}}_{(\bar{u}_k \cdot \bar{v}_k)_{ij}}, \qquad U_{ik} = (\bar{u}_k)_i, \quad V_{kj} = (\bar{v}_k)_j$$

$$= \Big(\sum_k \bar{u}_k \cdot \bar{v}_k\Big)_{ij}$$

outer product

Cor. M is rank $\le r$ iff

M is sum of $\le r$ rank-1 matrices

Ex. f bilinear form

$$f = \sum_{ij} A_{ij} x_i y_j$$

A: coeff matrix of f

$$\operatorname{rank}(f) = \operatorname{rank}(A)$$

$\operatorname{rank}(f)$: bilinear rank; $\operatorname{rank}(A)$: matrix rank

def. a (3-dimensional) tensor is a element of $\mathbb{F}^{n \times m \times p} \cong \mathbb{F}^{nmp}$

a tensor is simple if there exist $\bar{u} \in \mathbb{F}^n, \bar{v} \in \mathbb{F}^m, \bar{w} \in \mathbb{F}^p$ s.t.

$$T_{ijk} = u_i \cdot v_j \cdot w_k$$
$$T = \bar{u} \otimes \bar{v} \otimes \bar{w}$$

the rank of a tensor is min r s.t.

$$T = \sum_{\ell=1}^{r} T_\ell$$

simple tensor

rmk. – generalizes to $d > 3$ dim

– $d=2$ case is just matrix rank

lem. T simple tensor. T zero $\Leftrightarrow$ rank $=0$

T nonzero $\Leftrightarrow$ rank $=1$

prop. $f_1, \ldots, f_n$ bilinear forms on $\mathbb{F}[\bar{x}, \bar{y}]$

$$f_k = \sum_{ij} \alpha_{ijk} x_i y_j$$

define the structural tensor $T \in \mathbb{F}^{n \times n \times n}$ by $T_{ijk} = \alpha_{ijk}$

then $\operatorname{rank}(\bar{f}) = \operatorname{rank}(T)$

bilinear rank

tensor rank

pf. $\operatorname{rank}(\bar{f}) \le r$ iff all k $f_k = \sum \alpha_{ijk} x_i y_j \in \operatorname{span}\left\{ \left(\sum_i \beta_{i\ell} x_i\right) \cdot \left(\sum_j \gamma_{j\ell} y_j\right) \right\}_{\ell \le r}$

iff $T^k = (\alpha_{ijk})_{ij} \in \operatorname{span}\left(\bar{\beta}_\ell \cdot \bar{\gamma}_\ell^t\right)_\ell$

iff all k $\exists \delta$'s $T^k = \sum_\ell \delta_{k\ell} \bar{\beta}_\ell \cdot \bar{\gamma}_\ell^t$

iff $T_{ijk} = \sum_\ell \underbrace{\beta_{i\ell} \gamma_{j\ell} \delta_{k\ell}}_{T_\ell \text{ simple tensor}}$

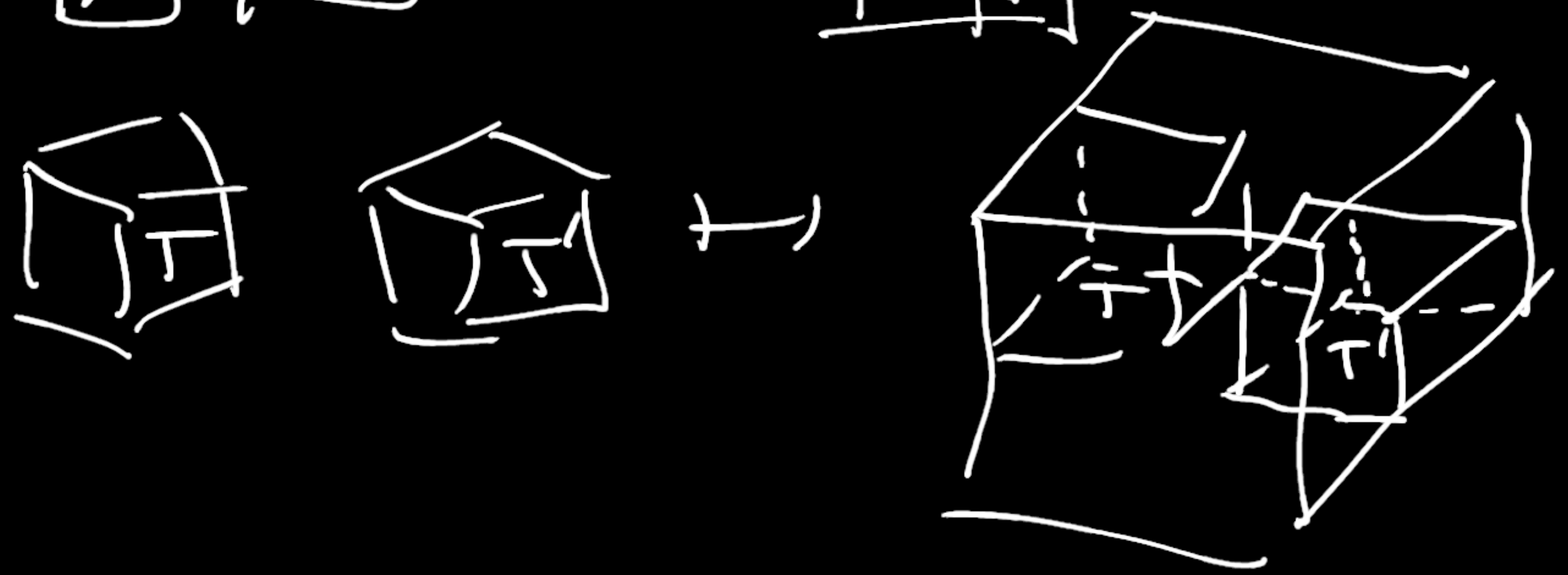
∎

Q: understand tensor rank?

def: $T, T' \in \mathbb{F}^{n\times n\times n}$

the direct sum $T \oplus T' \in \mathbb{F}^{2n\times 2n\times 2n}$ is

$$\boxed{A}, \boxed{B} \mapsto \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$



disjoint coord

the tensor product $T \otimes T' \in \mathbb{F}^{n^2\times n^2\times n^2} \cong \mathbb{F}^{(n\times n)\times(n\times n)\times(n\times n)}$

is $(T\otimes T')_{ii'jj'kk'} = T_{ijk}\cdot T'_{i'j'k'}$

fact: for matrices A, B – $\operatorname{rank}(A\oplus B) = \operatorname{rank}(A) + \operatorname{rank}(B)$

– $\operatorname{rank}(A\otimes B) = \operatorname{rank}(A)\operatorname{rank}(B)$

rank is efficiently computable

the naive ub on rank is n, achieved by obvious example,

identity matrix

fact: tensor T, T'

$-\operatorname{rank}(T\oplus T') \le \operatorname{rank}(T) + \operatorname{rank}(T')$

— — $\otimes$ — $\cdot$ —

Strassen's additivity conj: is =

is false [2017]

tensor rank is NP-hard to compute

lem: every $n\times n\times n$ tensor is rank $\le n^2$

pf attempt:

$$T = \sum_{ijk} T_{ijk}\, e_i \otimes e_j \otimes e_k$$

n^3

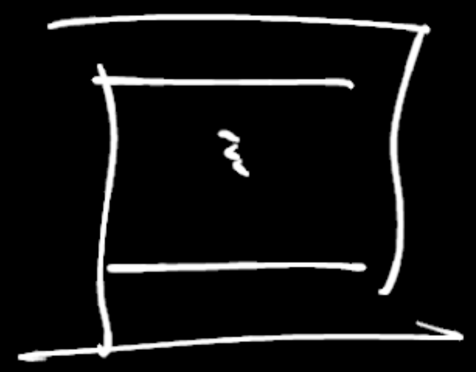
standard basis

pf: $T = \sum_i \vec{e}_i \otimes T_{i,:,:}$

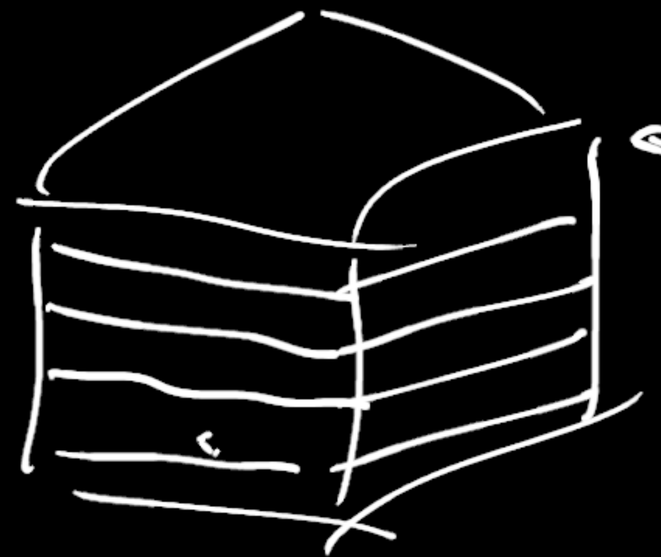
n

ith slice matrix

rank $\le n$



$\longleftarrow$ n rows each rank 1



$\longleftarrow$ matrix rank $\le n$

fact: every $n \times n \times n$ ~~matrix~~ tensor is rank $\le \left(\frac{2}{3} + o(1)\right) n^2 < n^2$

tensors of rank $\ge n^2/3$ exist

no explicit tensors known with rank $\gg n$

rank: - not obvious from bilinear perspective

- $\boxed{\text{is}}$ used in theory of fast matrix mult

Q: tensor rank of MM?

$A, B, C \in \mathbb{F}^{n \times n}$, $C = AB$ $\qquad C_{k,k'} = \sum_l A_{kl} \cdot B_{l,k'}$

def: the matrix mult tensor $MM^n \in \mathbb{F}^{n^2 \times n^2 \times n^2}$

is $(MM^n)_{ii',jj',kk'} = \begin{cases} 1 & i=k,\ i'=j,\ k'=j' \\ 0 & \end{cases}$

how $C_{kk'}$ depends on $A_{ii'} \cdot B_{jj'}$

$\equiv \begin{cases} 1 & i'=j,\ j'=k,\ k'=i \\ 0 & \end{cases}$

reordering

thm: $\mathrm{rank}(MM^{n,m,p}) = \mathrm{rank}(MM^{m,p,n})$

$(n \times m) \times (m \times p)$ $\qquad$ $(m \times p) \times (p \times n)$

sketch: def is symmetric wrt matrices

lem: $MM^{n,m,p} \otimes MM^{N,M,P} = MM^{nN,mM,pP}$

idea: block matrix mult in tensor lang

pf: $\left(MM^{nN,mM,pP}\right)_{ii'II',\, jj'JJ',\, kk'KK'}$

$$= \begin{cases} 1 & i'I' = jJ,\ j'J' = kK,\ k'K' = iI \\ 0 & \text{else} \end{cases}$$

$\underbrace{i' = j,\ j' = k,\ k' = i,}\ \underbrace{I' = J,\ J' = K,\ K' = I}$

today: - ex. pres. of MM via bilinear rank

- ————————— tensor rank

next lecture: beating Strassen

logistics: materials uploaded soon