Computing the non-commutative rank of linear matrices

Gábor Ivanyos
HUN-REN SZTAKI

RTCA 2023, Fundamental Algorithms and Algorithmic Complexity, 15-29 September 2023, Paris.

Based (mostly) on joint works with Youming Qiao and K. V. Subrahmanyam

Commutative and noncommutative rank

- linear n by n matrix: $A(x) = A(x_1, \ldots, x_k) = A_1x_1 + \ldots + A_kx_k$

  $A_1, \ldots, A_k \in F^{n \times n} (= M_n(F))$

  $\sim$ matrix space $\mathcal{A} = \text{Span}(A_1, \ldots, A_k)$;
- ordinary (commutative) rank $\text{rk}\, A(x)$: as a matrix over $F(x_1, \ldots, x_n)$

  max rank from $\mathcal{A}$ (if F is large enough)
- computational problem: determine $\text{rk}\, A(x)$ (Edmonds 1967)

  an instance of PIT, $\in RP$, not known to be in P

  "derandomization" would imply circuit lower bounds for NEXP

  (Kabanets, Impagliazzo 2003)
- noncommutative rank $\text{ncrk}\, A(x)$: as a matrix over the free skewfield

  $\max\{$max rank from $\mathcal{A} \otimes_F D$: D skewfield ext. of $F\}$

  $\mathcal{A} \otimes_F D =$ "D-span" of A_js

  Gaussian elim. and consequences to rank

  remain valid over skewfields

Commutative vs. noncommutative rank

- $\operatorname{rk} A(x) \leq \operatorname{ncrk} A(x)$
- Example for $<$: $\mathcal{A} =$ skew-symmetric 3 by 3 real matrices,

  A_1, A_2, A_3 a basis

  $\operatorname{rk} A(x) = 2$; $\operatorname{ncrk} A(x) = 3$ (over the quaternions)
- which one is easier to compute?
  - ncrk is a proper relaxation of rk
  - but its definition is more complicated

    uses a difficult object or a (possibly) infinite family of skewfields

    (can be pulled down to exp size)

    even computability in randomized poly time is not obvious
- ncrk is "easier":

  computable even in **deterministic polynomial** time!
  - Garg, Gurvits, Oliveira, Wigderson 2015-2016 $(\operatorname{char}(F) = 0)$;
  - IQS 2015-2018;
  - Hamada, Hirai 2021

The nc rank as a rank of a large matrix

- Can assume D is finite (d^2-)dimensional over its center C, where C is a fin. gen. (possibly transcendental) extension of F
  - $D \otimes_C L \cong L^{d\times d}$ explicitly for some field $L \geq C$
  - both D and $F^{d\times d}$ embedded in $L^{d\times d}$ as spanning subsets
- switching procedures

  $\mathcal{A} \otimes D \longleftrightarrow \mathcal{A} \otimes L^{d\times d} \longleftrightarrow \mathcal{A} \otimes F^{d\times d} \subseteq F^{nd\times nd}$

  rank r over $D \longrightarrow$ rank $\geq r \cdot d$ in $F^{nd\times nd}$

  rank R in $F^{nd\times nd} \longrightarrow$ rank $\geq \lceil R/d \rceil$ over D
  - round trip $\mathcal{A} \otimes F^{d\times d} \to \mathcal{A} \otimes D \to \mathcal{A} \otimes F^{d\times d}$

    rank R over $D \longrightarrow$ rank $\geq d\lceil R/d \rceil$ over F

  IQS 2015: can be done in deterministic poly time (for suitable D)
- Connection to invariant theory:

  determinants of matrices in $\mathcal{A} \otimes F^{d\times d}$

  $\sim$ invariants of $SL_n \times SL_n$ (degree dn homomgenous part)

Blowups of matrix spaces

- $\mathcal{A} \otimes F^{d\times d}$: *"blown up"* matrix space (d: blowup factor)
  n by n matrices with entries from $F^{d\times d}$
- based on the rounding, Derksen-Makam 2015-2017:

  a tool reducing d to $d-1$ if $d \geq n$

  preserving the "relative rank"

  matrix of rk $d$ncrk $\rightarrow$ matrix of rk $(d-1)$ncrk

$\text{ncrk}\, A(x) = \frac{1}{d}\max$ rank in $\mathcal{A} \otimes F^{d\times d}$ for some $d \leq n-1$.

$\Rightarrow$ ncrk computable in randomized poly time

Deterministic polynomial time algorithms

- Garg, Gurvits, Oliveira, Wigderson 2015-2016:

  operator scaling for over fileds of zero characteristic

- IQS 2015-2018: a constructive algorithm
  - computes a matrix of rank $d \cdot \text{ncrk}\, A(x)$ in $\mathcal{A} \otimes F^{d \times d}$

    $d \leq n-1$ (or $d \leq n \log n$ if F is too small)
  - computes an ("upper") witness for that ncrk cannot be larger
  - uses analogues of the alternating paths for matchings if graphs $+$ an efficient implementation of the DM reduction tool

- Franks, Soma, Goemans 2023:

  a version of GGOW that also finds an upper witness

- Hamada, Hirai 2021:

  convex optimization (based on finding an upper witness)

The upper witnesses: shrunk subspaces (Hall-like obstacles)

- ℓ-shrunk subspace: $U \leq F^n$ mapped to a subspace of dimension $\dim U - \ell$ by $\mathcal{A}$

$$\mathcal{A} \leq \begin{pmatrix} * & * & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \\ * & * & * & * & * \end{pmatrix} \text{alias} \begin{pmatrix} * & * & & & \\ * & * & & & \\ * & * & & & \\ * & * & * & * & * \end{pmatrix}$$

$\exists$ ℓ-shrunk subsp. $\Rightarrow$ the max rank in $\mathcal{A}$ is at most $n - \ell$

- Inheritance: $U \otimes F^{d \times d}$ mapped to a subspace of dim less by $\ell \cdot d$ $\Rightarrow$ max rank in $\mathcal{A} \otimes F^{d \times d}$ is at most $nd - \ell d$.
- $\Rightarrow$ ncrk $\leq n - \ell$
- $\sim$ a characterization of the nullcone of invariants $SL_n \times SL_n$ (by Hilbert-Mumford)

Main tool of IQS: the Wong sequence

- Idea: attempt to find a shrunk subspace

  (used in spec. commutative cases: Fortin, Reutenauer 2004; I., Karpinski, Saxena 2010; I., Karpinski, Qiao, Santha 2015)

- Assume we have $B \in \mathcal{A}$ with $\operatorname{rk} B = \operatorname{ncrk}$, $\ell = n - \operatorname{ncrk}$, U ℓ-shrunk. Then

$$U \geq \ker B \text{ and } \mathcal{A}U = \operatorname{Im} B.$$

- Wong sequence ($\sim$ alternating forest in bipartite graph matching):

  $U_1 = \ker B;\ U_{i+j} = B^{-1}(\mathcal{A}U_j)$ (inverse image under B)

  - Either stabilizes inside $\operatorname{Im} B$: gives an ℓ-shrunk subspace
  - or "escapes" : $\mathcal{A}U_j \not\subseteq \operatorname{Im} B$: ($\sim \exists$ augmenting path)

Escaping Wong sequence $\sim$ augmenting path

- sequence $i_1, \ldots, i_s$ – with s smallest – s.t.

$$A_{i_s} B^{-1}(A_{i_{s-1}} B^{-1}(\ldots B^{-1}(A_{i_1} \ker B))) \not\subseteq \operatorname{Im} B$$

- Key fact: ncrk $=$ rk if $\dim \mathcal{A} \leq 2$ (Atkinson, Stephens 1978)

  if $A^j \ker B \not\subseteq \operatorname{Im} B$ for some j, then

  $\operatorname{rk}(B + \lambda A) > \operatorname{rk} B$ for some λ (if F is large enough)
- Idea: try $A = \sum \lambda_i A_i$
- Why ncrk $\neq$ rk in general: escaping "paths" may cancel out
- Workaround let $d \geq s$;
  - Put $A'_1 = B' = B \otimes I_d$, $A'_2 = \sum A_{i_j} \otimes E_{j,j+1} \in \mathcal{A} \otimes F^{d \times d}$; $\mathcal{A}' = \langle A'_1, A'_2 \rangle$
  - Then the Wong seq. escapes $\operatorname{Im} B'$ and $C' = B' + \lambda A'_2$ has rank $> d \cdot \operatorname{rk} B$ for some λ
  - Round up the rank of C' in $\mathcal{A} \otimes F^{d \times d}$ to a multiple of d

Summary of the IQS algorithm

- iterate the above "scaled" rank incrementation procedure (with iteratively blowing up $\mathcal{A}$)
- combine with the reduction tool to control blowup factor
- Result: $A \in \mathcal{A} \otimes F^{d \times d}$ of rank $d \cdot \mathrm{ncrk}$; and a maximally (by $(n-d)\mathrm{ncrk}$) shrunk subspace (of F^{nd}) for $\mathcal{A} \otimes F^{d \times d}$
- Use converse of inheritance to obtain a maximally (by $n - \mathrm{ncrk}$) shrunk subspace of F^n for $\mathcal{A}$.
- Remarks:

  (1) Actually, *the smallest* maximally shrunk subspace found. ((0) if $\mathrm{ncrk} = n$.)

  (2) The largest one can also be found (duality)

Applications I.: Module isomorphism

- Module data (over m-generated algebras)

  $B_1, \ldots, B_m \in \mathbb{F}^{n \times n}$ $\sim$ action of generators

- Space of homomorphisms

  V, V' with data $B_1, \ldots, B_m, B'_1, \ldots, B'_m$

$$\mathrm{Hom}(V, V') = \{A \in \mathbb{F}^{n \times n} : AB_i = B'_i A\}$$

  Isomorphism: nonsingular element

- Blowups of Hom-spaces

$$\mathrm{Hom}(V, V') \otimes \mathbb{F}^{d \times d} = \mathrm{Hom}(V^{\oplus d}, V'^{\oplus d})$$

Module isomorphism II.

- Krull-Schmidt
  - Unique direct decomposition into indecomposables
  - $V^{\oplus d} \cong V'^{\oplus d} \iff V \cong V'$
  - $V \cong V' \iff \operatorname{ncrk} \operatorname{Hom}(V, V') = n$
- deciding $\cong$: a simple application of ncrank computation
- can be made constructive

  using a "lazy" constructive Krull-Schmidt
- Unpublished, $\exists$ several more direct approaches, e.g.,
  - Brooksbank, Luks (2008)
  - I., Karpinski, Saxena (2010)

    based on Chistov, I., Karpinski (1997) (for the semisimple case)
  - Ciocănea-Teodorescu (2015)

Applications II. (Invariant theory and related)

- Orbit closure separation for left-right action of SL
  - Derksen, Makam 2018
    Compute a separating invariant (if $\exists$)
- Brascamp-Lieb inequalities

$$\int_{x\in\mathbb{R}^n} \prod_i (f_i(B_i x))^{p_i}\, dx \le C \prod_i \left(\int_{y_i\in\mathbb{R}^{n_i}} f_i(y_i \quad dy_i \right)^{p_i}$$

$\forall\ 0 \le f_i :\in L^1(\mathbb{R}^{n_i})$
$0 < C \le \infty$ (the BL-constant)
depending on $B_i \in \mathbb{R}^{n_i \times n}$, $p_i \ge 0$.

  - capture many known inequalities, e.g., Hölder's
  - Garg, Gurvits, Oliveira, Wigderson 2018
    Operator scaling for a related matix space computes C
  - $C < \infty$ iff full ncrk

Applications III.

Block triangularization in the full ncrk case

- $\sim$ finding flag of 0-shrunk subspaces U ($\dim \mathcal{A}U = \dim U$)
- If $I \in \mathcal{A}$ then (as $\mathcal{A}W \geq W$) equivalent to $\mathcal{A}U = U$.
  - U: a submodule for the enveloping algebra of $\mathcal{A}$,
  - over many F, $\exists$ good algorithms
- If $A \in \mathcal{A}$ of full rank found, $I \in A^{-1}\mathcal{A}$

$$\mathcal{A} \leftarrow A^{-1}\mathcal{A}$$

- In the general case,
  - Find $A \in \mathcal{A} \otimes F^{d\times d}$ of full rank,
  - Block triangularize $\mathcal{A} \otimes F^{d\times d}$ as above
  - Pull back by "reverse inheritance"

    Blowup as a "magnifier"

Applications of block triangularization

- Effective orbit closure intersection
  - I., Qiao 2023
  - Compute one-parameter subgroups driving from orbits to the intersection of orbit closures
- In multivariate cryptography
  - based on hardness of solving polynomial systems
  - Sometimes: secret $\sim$ block triang. strucure
  - e.g, Patarin's balanced Oil and Vinegar scheme