The NC Waring problem and fast NC polynomial evaluation

Bill Helton UC San Diego
Jiawang Nie UC San Diego
Shiyuan Huang Columbia CS Grad Student

Your narrator Eric Evert KU Leuven

IWOTA 2019 Lisbon

Advertisement: **Try noncommutative computation**
NCAlgebra[1] **NCSoSTools**[2]

---

[1] Helton, de Oliveira (UCSD), Stankus (CalPoly SanLObispo), Miller
[2] Igor Klep

Ingredients of Talk: NC polynomials

$x = (x_1, \cdots, x_g)$ $\quad x^* = (x_1^*, \cdots, x_g^*)$ noncommuting variables

Noncommutative polynomials: $p(x)$:

$$Eg. \qquad p(x) = x_1^* x_2 + x_2^* x_1$$

An **analytic polynomial** contains no x_j^*.

Evaluate p: on matrices $X = (X_1, \cdots, X_g)$ a tuple of matrices.

Substitute matrices for variables

$x_1 \mapsto X_1$, $x_2 \mapsto X_2$ $x_1^* \mapsto X_1^*$, $x_2^* \mapsto X_2^*$

$$Eg. \qquad p(X) = X_1^* X_2 + X_2^* X_1.$$

Noncommutative inequalities: p is positive means:

$$p(X) \text{ is PSD for all } X$$

An homogeneous noncommutative degree d polynomial p has a t-term real Waring (resp. complex Waring) decomposition provided that $p(x)$ can be written as the sum of t terms of the d^{th}-power of linear functions of x, i.e.,

$$p(x) = \sum_{s=1}^{t} [A_1^s x_1 + A_2^s x_2 + \ldots A_g^s x_g]^d \tag{1}$$

with real (resp. complex) numbers A_j^s.

NC Waring Problem:
Determine if a noncommutative homogeneous degree d polynomial p has a t-term Waring decomposition.

Claim: This problem reduces to the classical commutative variable Waring problem, thereby effectively solving it over $\mathbb{C}$.

General NC Waring Problem

A NC homogeneous degree δd polynomial p,

$$p(x) = \sum_{s=1}^{t} \left(\sum_{|\beta|=\delta} A_\beta^s x^\beta \right)^d$$

is called a t-term (δ, d)-NC Waring decomposition.
Of course $\delta = 1$ is the Waring decomposition.

The Compatibility Condition

LEMMA

If a noncommutative homogeneous degree d polynomial $p(x)$,

$$p(x) = \sum_{|\alpha|=d} P_\alpha x^\alpha \qquad P_\alpha := P_{\alpha_1,\alpha_2,\dots,\alpha_d} \in \mathbb{R} \text{ or } \mathbb{C},$$

has a Waring decomposition, then its coefficients satisfy the compatibility condition

$$P_\alpha = P_{\tilde{\alpha}} \qquad \text{for all } \alpha \sim_c \tilde{\alpha}. \tag{2}$$

Here two index tuples α and $\tilde{\alpha}$ satisfy $\alpha \sim_c \tilde{\alpha}$, iff x^α and $x^{\tilde{\alpha}}$ have the same commutative collapse: eg.

$$x_1 x_2 x_1 x_3 x_1 x_2 \ \sim_c \ x_1^3 x_2^2 x_3$$

$$\alpha = (1,2,1,3,1,2) \ \sim_c \ \tilde{\alpha} = (1,1,1,2,2,3)$$

LEMMA

If p meets the compatibility condition, then p has a t-term NC Waring decomposition over the complex numbers (resp. real numbers) if and only if

$$P_\alpha = \sum_{s=1}^{t} \prod_{i=1}^{d} A_{\alpha_i}^s$$

has a solution $A_j^s \in \mathbb{C}$(resp. $A_j^s \in \mathbb{R}$).

PROOF OF LEMMA

p has a t-term Waring decomposition if and only if

$$\sum_{|\alpha|=d} P_\alpha x^\alpha = \sum_{s=1}^{t} [A_1^s x_1 + A_2^s x_2 + \ldots A_g^s x_g]^d = \sum_{s=1}^{t} \sum_{|\alpha|=d} \left(\prod_{i=1}^{d} A_{\alpha_i}^s \right) x^\alpha$$

$$= \sum_{|\alpha|=d} \left(\sum_{s=1}^{t} \prod_{i=1}^{d} A_{\alpha_i}^s \right) x^\alpha .$$

Comparing the coefficients of x^α on both sides, we get

$$P_\alpha = \sum_{s=1}^{t} \prod_{i=1}^{d} A_{\alpha_i}^s$$

This also implies $P_\alpha = P_{\tilde{\alpha}}$ if $\alpha \sim_c \tilde{\alpha}$

EXAMPLE

A NC homogeneous polynomial $p(x) = \sum_\alpha P_\alpha x^\alpha$ has the complex (resp. real) 2-term Waring decomposition

$$p(x) = (ax_1 + cx_2)^3 + (bx_1 + dx_2)^3$$

if and only if

$$\begin{aligned}
P_{1,1,1} &= a^3 + b^3 \\
P_{1,1,2} &= a^2c + b^2d = \frac{1}{6}((a+c)^3 + (b+d)^3 - (a-c)^3 - (b-d)^3) - \frac{1}{3}P_{2,2,2} \\
P_{1,2,2} &= ac^2 + bd^2 = \frac{1}{6}((a+c)^3 + (b+d)^3 + (a-c)^3 + (b-d)^3) - \frac{1}{3}P_{1,1,1} \\
P_{2,2,2} &= c^3 + d^3
\end{aligned} \tag{3}$$

has a solution $a, b, c, d \in \mathbb{C}$ (resp. $\mathbb{R}$).

THM

Suppose p is an NC homogeneous polynomial which satisfies the compatibility conditions. Then the commutative collapse p^c has the Waring decomposition

$$p^c(X) = \sum_{s=1}^{t} [A_1^s X_1 + A_2^s X_2 + \cdots + A_g^s X_g]^d \tag{4}$$

(with X_i being commuting variables) if and only if p has the nc Waring decomposition

$$p(x) = \sum_{s=1}^{t} [A_1^s x_1 + A_2^s x_2 + \cdots + A_g^s x_g]^d. \tag{5}$$

Note that the number of terms is the same and the real coefficients (resp. complex coefficients) A_j^s are the same.

COR

Suppose a NC homogeneous polynomial of degree d in g variables satisfies the compatibility condition.
Then p has an NC complex coefficient Waring decomposition.
The number of terms needed is no more than

$$\left\lceil \frac{\binom{g+d-1}{d}}{g} \right\rceil,$$

except in the cases

- $d = 2$, **where g terms are needed**
- $(d, g) = (3, 5), (4, 3), (4, 4), (4, 5)$ **where $\lceil \frac{1}{g}\binom{g+d-1}{d} \rceil + 1$ terms are needed.**

PROOF OF THM
DEFS

$\mathcal{R}$ **denotes a set consisting of one representative from each** $\sim_c$ **equivalence class.**

$\eta[\alpha] = \frac{d!}{\prod_{j=1}^{g}(\#\alpha_j \ in \ \alpha)!}$, **the number of members of the** α **equivalence class.**

Observe:

p^c **is the commutative collapse of a compatible NC homogeneous degree** d **polynomial** p

iff

$P^c_\alpha = \eta[\alpha] P_\alpha$ **for all index tuples** $\alpha \in \mathcal{R}$ **of length** d

Now we proceed to prove our theorem. Assume p has the NC Waring decomposition, we shall obtain a reversible formula for the Waring decomposition of p^c. By the above

$$p^c(X) = \sum_{\alpha\in\mathcal{R},\ |\alpha|=d} \eta[\alpha] P_\alpha X^\alpha$$

$$= \sum_{|\alpha|=d} P_\alpha X^\alpha = \sum_{|\alpha|=d} \sum_{s=1}^{t} \prod_{j=1}^{d} A^s_{\alpha_j} X^\alpha$$

Thus

$$p^c(X) = \sum_{s=1}^{t} \sum_{|\alpha|=d} \prod_{j=1}^{d} A^s_{\alpha_j} X^\alpha = \sum_{s=1}^{t} [A^s_1 X_1 + A^s_2 X_2 + \ldots A^s_g X_g]^d$$

so p^c has a Waring decomposition.

Conversely, suppose p's commutative collapse, p^c, has the commutative Waring decomposition, then the calculations above can be reversed. By comparing coefficients, this is equivalent to

$$P^c_\alpha = \eta[\alpha] \sum_{s=1}^{t} \prod_{j=1}^{d} A^s_{\alpha_j} X^\alpha$$

for all $\alpha \in \mathcal{R}$. Therefore p satisfies

$$P_\alpha = \sum_{s=1}^{t} \prod_{j=1}^{d} A^s_{\alpha_j} X^\alpha$$

for all index tuples α of length d. Hence by the Lemma p has the Waring decomposition.

General NC Waring Problem

A NC homogeneous degree δd polynomial p,

$$p(x) = \sum_{s=1}^{t} \left(\sum_{|\beta|=\delta} A_\beta^s x^\beta \right)^d$$

is called a t-term (δ, d)-NC Waring decomposition.

Idea: define new variables $z_\alpha := x^\alpha$ for $|\alpha| = \delta$.

THM There is a (necessary and sufficient) explicit reduction to commutative Waring in the z_α variables.
I will spare you the explicit (notation rich) statement.

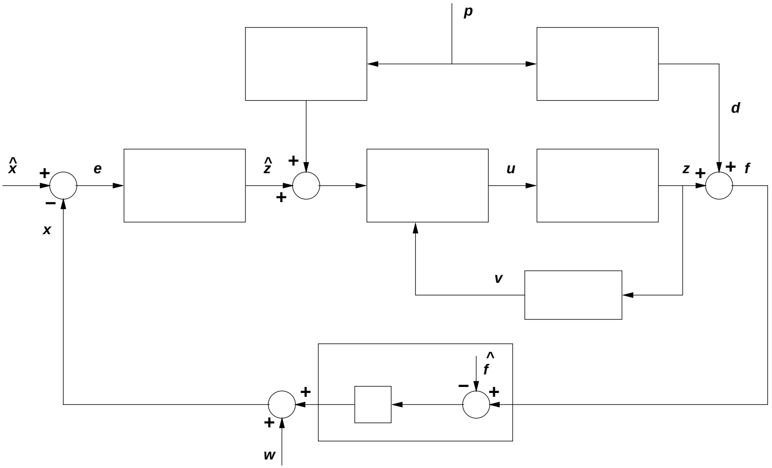
EVALUATING NC POLYNOMIALS ON MATRICES

1. MOTIVATION

2. USING POLYNOMIAL DeCOMPOSITIONS

MOTIVATION FROM LINEAR SYSTEMS ENGINEERING

Complicated systems give nc polynomials



Engineering problems defined entirely by signal flow diagrams and L^2 performance specs

are equivalent to

Polynomial Matrix Inequalities

Linear Systems and Algebra Synopsis

A Signal Flow Diagram with L^2 based performance, eg H^∞ gives precisely a nc polynomial

$$p(a, x) := \begin{pmatrix} p_{11}(a, x) & \cdots & p_{1k}(a, x) \\ \vdots & \ddots & \vdots \\ p_{k1}(a, x) & \cdots & p_{kk}(a, x) \end{pmatrix}$$

Such linear systems problems become exactly:

Given matrices A.
Find matrices X so that $P(A, X)$ is PosSemiDef.

BAD Typically p is a mess, until a hundred people work on it and maybe convert it to CONVEX Matrix Inequalities.

NUMERICAL OPTIMIZATION

Given matrices A.
Find matrices X so that $P(A, X)$ is PosSemiDef.
Always find X by numerical optimization.

Standard packages destroy the matrix structure and vectorize.

Maintain matrix (multiplication) structure:
Primal Interior Point method (Camino and Helton and Skelton)
Primal Dual(Helton and Mauricio deOliveira)

First compute second NC Directional derivatives with noncommutative computer algebra to produce the descent directions.

The bottleneck is typically "the linear subproblem".
The linear Subproblem always has the same form.

LINEAR SUBPROBLEM

Solve linear equations of the following form in unknown H:

$$\sum_{j}^{\Gamma} A_j(Z)\ H\ B_j(Z) + B_j^T(Z)\ H\ A_j^T(Z) = Q(Z) \qquad H^T = H$$

Here $A_j(z), B_j(z), Q(z)$ NC rational functions which are computed symbolically in advance of the numerics (long story). Γ is small < 10.

At k^{th} iteration of opt algorithm plug in current Z_k, then

$$A_j := A_j(Z_k) \quad B_j := B_j(Z_k) \qquad , \ Q := Q(Z_k)$$

are matrices and we solve for H.

Big open “question” find an effective numerical linear solver?

Issue for talk: evaluate NC rationals or polys more quickly.

NC Polys and Tensors

The coefficients of a a homogeneous degree d NC polynomial

$$p(x) = \sum_{|\alpha|=d} P_\alpha x^\alpha \tag{6}$$

in g variables $x = (x_1, \ldots, x_g)$ can be thought of as depth d array in g dimensions. That is, get a tensor

$$T = (P_\alpha)_{|\alpha|=d}.$$

T has a rank r decomposition

$$T = \sum_{s=1}^{r} A^s(1) \otimes A^s(2) \otimes \cdots \otimes A^s(d) \quad \textit{where} \;\; A^s(i) = \begin{pmatrix} A_1^s(i)) \\ \vdots \\ A_g^s(i) \end{pmatrix} \in \mathbb{C}^g$$

is equivalent to p having a Waring like decomposition

$$p(x) = \sum_{s=1}^{r} \prod_{i=1}^{d} \Big(A_1^s(i) x_1 + A_2^s(i) x_2 + \cdots + A_g^s(i) x_g \Big)$$

T has rank r symmetric tensor decomposition means

$$T = \sum_{s=1}^{r} A^s \otimes \cdots \otimes A^s$$

where d copies of $A^s = (A_1^s, \ldots, A_g^s)$ appear in each term.
Equivalantly

$$p(x) = \sum_{s=1}^{r} \left(A_1^s x_1 + A_2^s x_2 + \cdots + A_g^s x_g \right)^d.$$

That is p has a Waring decomposition.

Consider the homogeneous noncommutative polynomial

$$\begin{aligned} p(x) \;=\; & x_1^3 - 4x_2^3 - 4x_3^3 + 5x_1x_1x_2 + 5x_1x_2x_1 + 5x_2x_1x_1 \\ & -3x_1x_1x_3 - 3x_1x_3x_1 - 3x_3x_1x_1 + 7x_2x_2x_1 + 7x_2x_1x_2 \\ & +7x_1x_2x_2 - 11x_2x_2x_3 - 11x_2x_3x_2 - 11x_3x_2x_2 + 6x_3x_3x_1 \\ & +6x_3x_1x_3 + 6x_1x_3x_3 - 6x_3x_3x_2 - 6x_3x_2x_3 - 6x_2x_3x_3 \\ & +x_1x_2x_3 + x_1x_3x_2 + x_2x_1x_3 + x_2x_3x_1 + x_3x_1x_2 + x_3x_2x_1. \end{aligned}$$

Note the compatibility condition holds.
We want to use TENSORLAB to find its Waring decomposition.

So associate $p(x)$ to the symmetric tensor T defined by its frontal slices

$$T(:,:,1) = \begin{pmatrix} 1 & 5 & -3 \\ 5 & 7 & 1 \\ -3 & 1 & 6 \end{pmatrix} \quad \text{and} \quad T(:,:,2) = \begin{pmatrix} 5 & 7 & 1 \\ 7 & -4 & -11 \\ 1 & -11 & -6 \end{pmatrix}$$

and

$$T(:,:,3) = \begin{pmatrix} -3 & 1 & 6 \\ 1 & -11 & -6 \\ 6 & -6 & -4 \end{pmatrix},$$

Using TENSORLAB[3] Can compute that T is a rank 4 tensor and has symmetric tensor decomposition

$$T = v_1 \otimes v_1 \otimes v_1 + v_2 \otimes v_2 \otimes v_2 + v_3 \otimes v_3 \otimes v_3 + v_4 \otimes v_4 \otimes v_4$$

where

$$v_1 \approx \begin{pmatrix} -0.081 \\ -0.409 \\ 1.890 \end{pmatrix} \qquad v_2 \approx \begin{pmatrix} 3.165 \\ -3.910 \\ -3.654 \end{pmatrix}$$

and

$$v_3 \approx \begin{pmatrix} -3.273 \\ 3.727 \\ 3.397 \end{pmatrix} \qquad v_4 \approx \begin{pmatrix} 1.636 \\ 1.581 \\ -1.051 \end{pmatrix}.$$

[3] A matlab script which computes such decompositions using TENSORLAB is avaliable

It follows that p has the rank 4 NC Waring decomposition

$$\begin{aligned} p(x) \approx\ & (-0.081x_1 - 0.409x_2 + 1.890x_3)^3 \\ & +(3.165x_1 - 3.910x_2 - 3.654x_3)^3 \\ & +(-3.273x_1 + 3.727x_2 + 3.397x_3)^3 \\ & +(1.636x_1 + 1.581x_2 - 1.051x_3)^3. \end{aligned} \tag{7}$$

This allows p to be evaluated with only 8 matrix multiplications. In contrast, naively using the original expression for p requires 54 matrix multiplies.

Example: A general NC polynomial

Consider the noncommutative polynomial

$$\begin{aligned} p(x) \;=\; & 20x_1^3 + 50x_1x_2x_1 + 20x_1x_3x_1 - 30x_2x_1x_1 - 75x_2x_2x_1 \\ & -30x_2x_3x_1 - 10x_3x_1x_1 - 25x_3x_2x_1 - 10x_3x_3x_1 - 8x_1x_1x_2 \\ & -62x_1x_2x_2 - 35x_1x_3x_2 + 46x_2x_1x_2 + 59x_2x_2x_2 + 10x_2x_3x_2 \\ & +26x_3x_1x_2 + 9x_3x_2x_2 - 10x_3x_3x_2 + 44x_1x_1x_3 + 26x_1x_2x_3 \\ & -10x_1x_3x_3 + 2x_2x_1x_3 - 107x_2x_2x_3 - 70x_2x_3x_3 + 22x_3x_1x_3 \\ & -57x_3x_2x_3 - 50x_3x_3x_3. \end{aligned} \tag{8}$$

TENSORLAB produces a rank 2 Waring Like Decomposition.

Thus p can be evaluated with only 4 matrix multiplications. Compared to 54 using equation (8) naively.

Computational savings

Waring-like vs. naive

The maximum rank of a tensor $T \in (\mathbb{C}^g)^{\otimes d}$ is not known, but it is conjectured that outside of a set of measure zero, the rank equals

$$\frac{g^d}{dg - d + 1} \approx \frac{g^d}{d(g-1)}.$$

The conjecture plus counting implies

$$\frac{\text{\# matrix mults used in the Waring-like method}}{\text{\# matrix mults used in naive method}} \approx \frac{1}{(g-1)(d-1)}.$$

Waring vs. Waring-like

Note that the rank of a symmetric tensor is necessarily less than or equal to the symmetric rank of a tensor. This plus counting implies

$$\frac{\#\text{ matrix mults used in the Waring method}}{\#\text{ matrix mults used in Waring-like method}} \geq \frac{2\lfloor \log_2 d \rfloor}{(d-1)}$$

with equality if the rank of T equals the symmetric rank of T.
We expect the ratio is what a practioner will invariably see.

Because:
An example of symmetric tensors whose rank is strictly less than its symmetric rank was not found until 2018 (by Shitov). The corresponding NC polynomial p is a homogeneous degree $d = 3$ polynomial in $g = 800$ variables.

Thanks for your attention