A Lower Bound for the Waring Rank

Neriman Tokcan

University of Illinois at Urbana-Champaign

Midwest Algebraic Geometry Graduate Conference
Chicago

22 April 2017

Department of Mathematics, Statistics, and Computer Science
COLLEGE OF LIBERAL ARTS & SCIENCES

Main sources

[BS] G. Blekherman, R. Sinn, *Real rank with respect to varities*, *Linear Algebra and its Applications*, 505 (2016): 344-360.

[Re1] B. Reznick, *On the length of binary forms*, *Quadratic and higher degree forms*. Springer New York, 2013. 207-232.

[RT] B. Reznick and N. Tokcan, *Binary forms with three different relative ranks*, to appear in *Proceedings of the American Mathematical Society*.

[NT] N. Tokcan, *On the Waring rank of binary forms*, to appear in *Linear Algebra and its applications*.

Sylvester James Joseph Sylvester (1814–1897) made fundamental contributions to matrix theory, invariant theory, number theory, partition theory, and combinatorics. His works were republished in the 20th century (*The Collected Mathematical Papers of James Joseph Sylvester*, *Cambridge University Press*, 1904-12).

The K-rank of binary forms

Assume that f is a nonzero binary form of degree d with coefficients in a field $K \subseteq \mathbb{C}$. The K-rank of f, $L_K(f)$, is the smallest r for which there exist $\lambda_i, \alpha_i, \beta_i \in K$ such that

$$f(x,y) = \sum_{i=1}^{r} \lambda_i (\alpha_i x + \beta_i y)^d. \quad (1)$$

In case $K = \mathbb{C}$, the K-rank is commonly called the *Waring rank*, and for $K = \mathbb{R}$, it is called the *real Waring rank*.

Let E_f denote the field generated by the coefficients of f over $\mathbb{Q}$; $L_K(f)$ is defined for the fields satisfying $E_f \subseteq K \subseteq \mathbb{C}$.

Minimal representation of a binary form

A representation such as (1) is called *K-minimal* if $r = L_K(f)$.

1. Two linear forms are *distinct* if they (or their d-th powers) are not proportional.
2. A representation is *honest* if the summands are pairwise distinct.
3. Two honest representations are *different* if the ordered sets of summands are not rearrangements of each other.
4. We do not distinguish between ℓ^d and $(\zeta_d^k \ell)^d$ when $\zeta_d = e^{\frac{2\pi i}{d}}$.

$\mathbb{C}$-rank of forms

The computation of $\mathbb{C}$-rank is a huge and active subject, and challenging when $d \geq 3$. There are many open questions even for small $n, d \geq 3$.

- In 1995, Alexander and Hirschowitz proved that the $\mathbb{C}$-rank of a general form of degree d in n variables $(n, d \geq 3)$ is $\left\lceil \frac{1}{n}\binom{n+d-1}{n-1} \right\rceil$. Four exceptions are known since the 19-th century $(n, d) = (3,5), (4,3), (4,4), (4,5)$, where the rank is one more.
- In 1851, Sylvester explained how to compute $L_{\mathbb{C}}(f)$ for $n = 2$.

$\mathbb{C}$-rank of binary forms–Alexander and Hirschowitz

We identify $\mathbb{P}^d$ with the space $\mathbb{C}[x,y]_d$ of binary forms of degree d by associating the form

$$\sum_{i=1}^{d} \binom{d}{i} a_i x^{d-i} y^i$$

to the point $(a_0 : a_1 : \ldots : a_d) \in \mathbb{P}^d$. In this identification, the rational normal curve in $\mathbb{P}^d$ is sent to the set of powers of linear forms since $v_d(s:t) = (s^d : s^{d-1}t : \ldots : t^d)$ is identified with the binary form $(sx+ty)^d$.

The generic rank of binary forms of degree d is the smallest s such that $\sigma_s\left(v_d(\mathbb{P}^1)\right) = \mathbb{P}^d$ which is $\lceil \frac{d+1}{2} \rceil$.

Some properties of the rank of binary forms

- The following relation is immediate:

$$K_1 \subseteq K_2 \Longrightarrow L_{K_1}(f) \geq L_{K_2}(f). \tag{2}$$

- If g is obtained from f by an invertible linear change of variables over K, then $L_K(f) = L_K(g)$.

Quick review - apolarity

Suppose $p(x,y) = \sum_{i=1}^{d} a_i x^{d-i} y^i \in H_d(\mathbb{C}^2)$. The differential operator associated to p is given by

$$p(D) = \sum_{i=1}^{d} a_i \frac{\partial^d}{\partial x^{d-i} \partial y^i} .$$

The *apolar ideal* of p, which is denoted by $p^{\perp}$, is the set of all binary forms whose differential operator kills p, that is,

$$p^{\perp} = \{h \in H(\mathbb{C}^2) \mid h(D)p = 0\}.$$

Apolarity-continued

The following are some basic properties of apolarity.

- Since $\frac{\partial}{\partial x_k}$ and $\frac{\partial}{\partial x_l}$ commute, then $(pq)(D) = p(D)q(D) = q(D)p(D)$ for any forms.
- If h is apolar to p, $h(D)p = 0$, then any polynomial multiple of h is also apolar to p.
- If $deg(h) > deg(p)$, then $h(D)p = 0$.

Bivariate apolar ideals

It is well known that any bivariate apolar ideal is a complete intersection.

Theorem

Let $p(x,y) \in H_d(\mathbb{C}^n)$. Then $p^\perp$ is a complete intersection ideal over $\mathbb{C}$, i.e. $p^\perp = \langle f, g\rangle$ such that $deg(f) + deg(g) = d + 2$ and $Z_{\mathbb{C}}(f,g) = \{(0,0)\}$. Also, any two such binary forms f, g generate an ideal $p^\perp$ for a binary form p of degree $deg(f) + deg(g) - 2$.

Example

Let n, m be positive integers. Then $(x^n y^m)^\perp = \langle x^{n+1},\ y^{m+1}\rangle$, $(x^n - y^n)^\perp = \langle xy,\ x^n + y^n\rangle$ and $((x+y)^n)^\perp = \langle x - y,\ y^{n+1}\rangle$.

Sylvester's 1851 Theorem

Theorem (Sylvester, Re1)

Suppose

$$f(x,y) = \sum_{j=0}^{d} \binom{d}{j} b_j x^{d-j} y^j \in K[x,y]. \tag{3}$$

and suppose $r \leq d, \alpha_j, \beta_j \in K$ *and*

$$h(x,y) = \sum_{t=0}^{r} c_t x^{r-t} y^t = \prod_{j=1}^{r} (-\beta_j x + \alpha_j y) \tag{4}$$

is a product of pairwise distinct linear factors. Then there exist $\lambda_j \in K$ *so that*

$$f(x,y) = \sum_{j=1}^{r} \lambda_j (\alpha_j x + \beta_j y)^d \tag{5}$$

if and only if

$$\begin{pmatrix} b_0 & b_1 & \cdots\cdots & b_r \\ b_1 & b_2 & \cdots\cdots & b_{r+1} \\ \vdots & \vdots & \ddots & \vdots \\ b_{d-r} & b_{d-r+1} & \cdots\cdots & b_d \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}; \qquad (6)$$

that is, if and only if

$$h(D)f = 0. \qquad (7)$$

The $(d - r + 1) \times (r + 1)$ Hankel matrix in (6) will be denoted by $H_r(p)$.

Sketch of the proof of Sylvester's 1851 Theorem

Different approaches to the proof of Sylvester's theorem have been discussed in several works. I will present the one which is based on apolarity.

Proof.

We can write the system of equations (6) as

$$\sum_{i=0}^{r} b_{i+t}c_i = 0, \quad t = 0, 1, \cdots, d-r.$$

$$h(D) = \prod_{j=1}^{r}(-\beta_j \tfrac{\partial}{\partial x} + \alpha_j \tfrac{\partial}{\partial y}),$$ then

$$h(D)f = \sum_{t=0}^{d-r} \frac{d!}{(d-r-t)!t!}\left(\sum_{i=0}^{r} b_{i+t}c_i\right)x^{d-r-t}y^t.$$

Therefore, (6) is equivalent to $h(D)f = 0$. Main idea of the proof arguing that each linear factor in $h(D)$ kills a different summand in f. It gives rise to a representation of length r by dimension counting. □

Sylvester form

If (f, h) satisfies the criterion of the Sylvester's Theorem, we shall say that h is a *Sylvester form* for f. If the only Sylvester forms of degree r are λh for $\lambda \in \mathbb{C}$, we say that h is the (projectively) *unique* Sylvester form for f. It follows from apolarity that any square free multiple of a Sylvester form is also a Sylvester form.

The following is immediate from Sylvester's Theorem:

- Given $f \in K[x, y]$, $L_K(f)$ is the minimal degree of the Sylvester form for f which completely splits over K.

Apolarity Lemma

The following is a restatement of Sylvester's algorithm.

Theorem

Let $f \in H_d(\mathbb{C}^2)$. Then $f(x,y) = \sum_{i=1}^{r} \lambda_i(\alpha_i x + \beta_i y)^d$ if and only if the vanishing ideal of the set $\{(\alpha_i, \beta_i), 1 \le i \le r\}$ is contained in the apolar ideal $f^\perp$.

Example

Take $f(x, y) = 5x(x^2 + y^2)^2$.

In the process of finding Sylvester form of degree r, we should look for $r \times 1$ null vector of $6 - r \times r + 1$ Hankel matrix. Note that if $6 - r \geq r + 1$, then the null vector is trivial, unless the appopriate number of rows in the Hankel matrix are linearly dependent (It is not the case for this example).

Thus, we should look for Sylvester form of degree r, such that $r \geq 3$ (even without applying the whole process, we know that $L_{\mathbb{C}}(f) \geq 3$).

Continuing Example

$$f(x,y) = \binom{5}{0}.5\ x^5 + \binom{5}{1}.0\ x^4y + \binom{5}{2}.1\ x^3y^2 + \binom{5}{3}.0\ x^2y^3 + \binom{5}{4}.1\ xy^4 + \binom{5}{5}.0\ y^5$$

Since 3×4 matrix

$$\begin{pmatrix} 5 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Longleftrightarrow (c_0, c_1, c_2, c_3) = s(0,1,0,-1),$$

f has a unique Sylvester form $h(x,y) = x^2y - y^3$.
It directly implies that $L_{\mathbb{Q}}(f) = 3$, and so $L_K(f) = 3$ for any $\mathbb{Q} \subseteq K \subseteq \mathbb{C}$.

Continuing example - minimal representation

At this point, f can be represented over $\mathbb{Q}$ as a $\mathbb{Q}$-linear combination of 3 distinct, 5-th powers of linear forms.

$$h(x, y) \quad = \quad y \quad * \quad (x - y) \quad * \quad (x + y)$$

Continuing example - minimal representation

At this point, f can be represented over $\mathbb{Q}$ as a $\mathbb{Q}$-linear combination of 3 distinct, 5-th powers of linear forms.

$$\begin{array}{ccccccc} h(x,y) & = & y & * & (x-y) & * & (x+y) \\ & & \downarrow & & \downarrow & & \downarrow \\ f(x,y) & = & \lambda_1 x^5 & + & \lambda_2 (x+y)^5 & + & \lambda_3 (x-y)^5 \end{array}$$

Indeed, $\lambda_2 = \lambda_3 = \frac{1}{2}$ and $\lambda_1 = 4$.

Applications to forms of general degree

The following results are useful to determine the K-rank of a binary form. Suppose $\mathbb{Q} \subseteq K \subseteq \mathbb{C}$.

1. If $f \in K[x, y]$, then $L_K(f) = 1$ if and only if $L_{\mathbb{C}}(f) = 1$ [Re1].
2. If $f \in K[x, y]$, then $L_K(f) \leq deg(f)$ [Re1].
3. Suppose f is a real binary form of degree d, and not a d-th power. If f has τ real linear factors, counting multiplicity, then $L_{\mathbb{R}}(f) \geq \tau$ [Sylvester].

Full Waring rank

Let f be real binary form of degree d in $K[x,y]$. If $L_K(f) = d$, we say that f has *full rank* over the field K. The case for $K = \mathbb{C}$ has been completely analyzed. In the last years, the case $K = \mathbb{R}$ has been considered in different works and a final result has been recently achieved by Blekherman and Sinn.

- If $d \geq 3$, then $L_{\mathbb{C}}(f) = d$ if and only if there are two distinct linear forms ℓ_0 and ℓ_1 so that $f = {\ell_0}^{d-1}\ell_1$.
- Suppose f is a real binary form of degree $d \geq 3$ and not a d-th power. Then $L_{\mathbb{R}}(f) = d$ if and only if f is hyperbolic, i.e, it completely splits over $\mathbb{R}$ [BS].

A lower bound for the Waring rank

The following theorem shows the relation between the $\mathbb{C}$-rank of a binary form and the factorization of the form over $\mathbb{C}$.

Theorem (NT)

Let $f(x, y)$ be a nonzero binary form of degree d with the factorization

$$f(x, y) = \prod_{i=0}^{r} \ell_i(x, y)^{m_i} \tag{8}$$

where $r \geq 1$ and the ℓ_i's are distinct linear forms and $m_0 \geq m_1 \geq \ldots \geq m_r$. Then $L_{\mathbb{C}}(f) \geq m_0 + 1$.

Proof.

We use the fact that rank is invariant under invertible linear change of variables. After a linear change of variables we may assume $\ell_0 = y$, then we have

$$\tilde{f}(x,y) = y^{m_0} g(x,y) \text{ such that } y \nmid g(x,y). \qquad (9)$$

The first m_0 coefficients of $\tilde{f}$ are zero, i.e. $a_0 = \ldots = a_{m_0-1} = 0$ and $a_{m_0} \neq 0$. Note that $deg(\tilde{f}) \geq m_0 + 1$, so by setting $r = m_0$, (6) becomes:

$$\begin{pmatrix} 0 & 0 & \ldots & 0 & a_{m_0} \\ 0 & 0 & \ldots & a_{m_0} & a_{m_0+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ \vdots \\ c_{m_0} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Hence, $a_{m_0}c_{m_0} = a_{m_0}c_{m_0-1} + a_{m_0+1}c_{m_0} = 0$. It follows that $c_{m_0-1} = c_{m_0} = 0$ and every apolar form of degree m_0 is divisible by x^2 and $L_{\mathbb{C}}(f) \geq m_0 + 1$. □

Rank of $\ell_0^{d-2}\ell_1\ell_2$

Corollary (NT)

Let $f(x,y) = \ell_0(x,y)^{d-2}\ell_1(x,y)\ell_2(x,y)$ such that $d \geq 3$ and the ℓ_i's are distinct binary linear forms. Then $L_{\mathbb{C}}(f) = d-1$.

Proof.

Recall that $L_{\mathbb{C}}(f) = d \iff f = x^{d-1}y$. Therefore, $L_{\mathbb{C}}(f) \leq d-1$. It also follows from the previous theorem that $d-1 \leq L_{\mathbb{C}}(f)$. □

We can generate forms with two different relative ranks by using the above corollary. For example, let $r \neq 0 \in \mathbb{R}$ and $d \geq 3$, then $L_{\mathbb{C}}(x^{d-2}y(x+ry)) = d-1$ and $L_{\mathbb{R}}(x^{d-2}y(x+ry)) = d$.

Corollary (NT)

Suppose $f(x, y) = \ell(x, y)^{d-2}q(x, y)$ is a real binary form of degree $d \geq 3$ where $\ell(x, y)$ is a real linear form and $q(x, y)$ is an irreducible real quadratic form. Then $L_{\mathbb{R}}(f) = d - 1$.

Notice that if $f(x, y) = \ell(x, y)^{d-2}q(x, y)$ as in above corollary, then the real rank and complex rank of f coincide, i.e, $L_{\mathbb{R}}(f) = L_{\mathbb{C}}(f) = d - 1$.

A lower bound for the real Waring rank

Corollary (NT)

Let $f(x, y)$ be a nonzero real binary form of degree d and not a d-th power with the factorization

$$f(x,y) = \prod_{i=0}^{r} \ell_i(x,y)^{m_i} \prod_{k=0}^{s} q_k(x,y)^{n_k} \tag{10}$$

where the ℓ_i's are distinct real binary linear forms and the q_k's are distinct irreducible real quadratics. Then

$$L_{\mathbb{R}}(f) \geq \max\Big(\sum_{i=0}^{r} m_i, \max(m_0, \ldots, m_r, n_0, \ldots, n_s) + 1\Big).$$

$\mathbb{C}$-rank of binary quartics: Assume that the ℓ_i's are distinct binary linear forms. The following table gives the rank of binary quartics based on their factorization over $\mathbb{C}$.

$p(x,y)$	$L_{\mathbb{C}}(p(x,y))$
$\ell_0(x,y)^4$	1
$\ell_0(x,y)^3\ell_1(x,y)$	4
$\ell_0(x,y)^2\ell_1(x,y)^2$	3
$\ell_0(x,y)^2\ell_1(x,y)\ell_2(x,y)$	3
$\ell_0(x,y)\ell_1(x,y)\ell_2(x,y)\ell_3(x,y)$	2,3

We provide supporting examples for the last case:

- $L_{\mathbb{C}}(x^4+y^4) = L_{\mathbb{C}}(8x^3y+36x^2y^2+36xy^3) = 2.$
- $L_{\mathbb{C}}(4x^3y+6x^2y^2+4xy^3) = L_{\mathbb{C}}(x^4+4x^2y^2+y^4) = 3.$

$\mathbb{C}$-rank of binary quintics

$\mathbb{C}$-rank of binary quintics: Assume that the ℓ_i's are distinct binary linear forms. The following table gives the rank of binary quintics based on their factorization over $\mathbb{C}$.

$p(x,y)$	$L_{\mathbb{C}}(p(x,y))$
$\ell_0(x,y)^5$	1
$\ell_0(x,y)^4\ell_1(x,y)$	5
$\ell_0(x,y)^3\ell_1(x,y)^2$	4
$\ell_0(x,y)^3\ell_1(x,y)\ell_2(x,y)$	4
$\ell_0(x,y)^2\ell_1(x,y)^2\ell_2(x,y)$	3
$\ell_0(x,y)^2\ell_1(x,y)\ell_2(x,y)\ell_3(x,y)$	3,4
$\ell_0(x,y)\ell_1(x,y)\ell_2(x,y)\ell_3(x,y)\ell_4(x,y)$	2,3,4

$\mathbb{C}$-rank of binary quintics-continued

We leverage Sylvester's algorithm and provide examples for the last two cases:

- $L_{\mathbb{C}}(x^2(x^3+10y^3)) = 3$, $L_{\mathbb{C}}(x^2(x^3+5x^2y+10xy^2+10y^3)) = 4$.
- $L_{\mathbb{C}}(x^5+y^5) = 2$, $L_{\mathbb{C}}(3x^5+20x^3y^2+10xy^4) = 3$, and $L_{\mathbb{C}}(x^5-5xy^4) = 4$.

Thanks!