C. R. Acad. Sci. Paris, Ser. I 356 (2018) 52–55

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Lie algebras/Computer science

Plethysm and fast matrix multiplication

Pléthysme et multiplication rapide des matrices

Tim Seynnaeve

Max Planck Institute for Mathematics in the Sciences, Inselstrasse 22, 04103, Leipzig, Germany

ARTICLE INFO

Article history:
Received 10 October 2017
Accepted after revision 20 November 2017
Available online 6 December 2017

Presented by Michèle Vergne

ABSTRACT

Motivated by the symmetric version of matrix multiplication we study the plethysm $S^k(\mathfrak{sl}_n)$ of the adjoint representation $\mathfrak{sl}_n$ of the Lie group SL_n. In particular, we describe the decomposition of this representation into irreducible components for $k = 3$, and find highest-weight vectors for all irreducible components. Relations to fast matrix multiplication, in particular the Coppersmith–Winograd tensor, are presented.

RÉSUMÉ

Motivés par la version symétrique de la multiplication des matrices, nous étudions le pléthysme $S^k(\mathfrak{sl}_n)$ de la représentation adjointe $\mathfrak{sl}_n$ du groupe de Lie SL_n. En particulier, pour $k = 3$, nous décrivons la décomposition de cette représentation en composantes irréductibles, et nous trouvons les vecteurs de plus grand poids pour toutes ces dernières. Nous présentons les liens avec la multipliction rapide des matrices, notamment le tenseur de Coppersmith–Winograd.

1. Introduction

In 1969, Strassen [18] presented his celebrated algorithm for matrix multiplication breaking for the first time the naive complexity bound of n^3 for $n \times n$ matrices. Since then, the complexity of the optimal matrix multiplication algorithm is one of the central problems in computer science. In terms of algebra, we know that this question is equivalent to estimating the rank or the border rank of a specific tensor $M_{n,n,n} \in \mathbb{C}^{n^2} \otimes \mathbb{C}^{n^2} \otimes \mathbb{C}^{n^2}$ [1,8,9]. The current best lower and upper bounds are presented in [10,12–14,20].

We recall that the constant ω is defined as the smallest number such that, for any $\epsilon > 0$, the multiplication of $n \times n$ matrices can be performed in time $O(n^{\omega+\epsilon})$. Further, recall that the Waring rank of a homogeneous polynomial P of degree d is the smallest number r of linear forms $l_1, \ldots, l_r$ such that $P = \sum_{i=1}^{r} l_i^d$. Recently, Chiantini et al. [2] provided another equivalent interpretation of ω in terms of Waring (border) rank. Namely, let SM_n be a cubic in $S^3(\mathfrak{sl}_n^*)$ given by

E-mail address: tim.seynnaeve@mis.mpg.de.

https://doi.org/10.1016/j.crma.2017.11.012

$SM_n(A) = \mathrm{tr}(A^3)$. Then ω is the smallest number such that, for any $\epsilon > 0$, the Waring rank (or Waring border rank) of SM_n is $O(n^{\omega+\epsilon})$. This observation was the initial motivation for our study of the plethysm $S^3(\mathfrak{sl}_n)$.

The computations of plethysm are in general very hard, and explicit formulas are known only in specific cases [15]. For example, for a symmetric power $S^3(S^k)$, the decomposition was classically computed already in [17,19], but $S^4(S^k)$ and $S^5(S^k)$ were only recently explicitly obtained in [7]. As symmetric powers (together with exterior powers) are the simplest Schur functors, one could expect that the respective formulas for $S^d(\mathfrak{sl}_n)$ are harder. In principle, one could use the methods in [6,7,16] to decompose this plethysm, but this requires a lot of nontrivial character manipulations. Instead, we present a very easy proof of explicit decomposition based on the Cauchy formula and the Littlewood–Richardson rule in Theorem 1. In fact, using our method, one can inductively obtain the formula for $S^k(\mathfrak{sl}_n)$ for any k.

While matrix multiplication is represented by the (unique) invariant in $S^3(\mathfrak{sl}_n)$ the aim of this article is to understand the other highest-weight vectors. A precise description of them is presented in Section 3. We plan to undertake a detailed study of ranks and border ranks of other highest-weight vectors in future work. Here we present just the first two nontrivial instances. It turns out that two of the highest-weight vectors are (isomorphic to) the (four and five dimensional) variants of the Coppersmith–Winograd tensor [4]. We recall that the best upper bounds for rank and border rank are based on a beautiful technique by Coppersmith and Winograd applied to a specific tensor T [20]. While T is extremely efficient for this technique, it is completely unclear which properties of T make it so useful and how to identify potentially better tensors. In fact, there are whole programs, see, e.g., [3], aimed at finding tensors similar to, but better than the Coppersmith–Winograd one. We hope that other highest-weight vectors will also reveal their importance.

2. The plethysm

In this section, we describe a general procedure to decompose $S^k(\mathfrak{gl}_n)$ and $S^k(\mathfrak{sl}_n)$ into irreducibles. Recall that the irreducible representations of SL_n are precisely the representations $\mathbb{S}_\lambda(\mathbb{C}^n)$, where $\lambda = [\lambda_1, \dots, \lambda_{n-1}]$ is a partition of length at most $n-1$, and $\mathbb{S}_\lambda$ is the Schur functor associated with the partition λ (consult, for example, [5]).

Theorem 1. *For $n \in \mathbb{N}$, it holds that*

$$S^k(\mathfrak{gl}_n) \cong \bigoplus_{\lambda \vdash k} \bigoplus_{\nu} N^{\nu}_{\lambda\bar{\lambda}} \mathbb{S}_\nu(\mathbb{C}^n) \tag{1}$$

as SL_n-representations. Here the second summation is over all partitions ν of length at most $n-1$, $N^{\nu}_{\lambda\mu}$ are the Littlewood–Richardson coefficients, and $\bar{\lambda} = [\lambda_1 - \lambda_{n-1}, \dots, \lambda_1 - \lambda_2]$.

Proof. Note that $\mathfrak{gl}_n \cong (\mathbb{C}^n) \otimes (\mathbb{C}^n)^*$ as SL_n-representations. So

$$\begin{aligned} S^k(\mathfrak{gl}_n) &\cong S^k\big((\mathbb{C}^n) \otimes (\mathbb{C}^n)^*\big) \cong \bigoplus_{\lambda \vdash k} \mathbb{S}_\lambda(\mathbb{C}^n) \otimes \mathbb{S}_\lambda(\mathbb{C}^n)^* \\ &\cong \bigoplus_{\lambda \vdash k} \mathbb{S}_\lambda(\mathbb{C}^n) \otimes \mathbb{S}_{\bar{\lambda}}(\mathbb{C}^n) \cong \bigoplus_{\lambda \vdash k} \bigoplus_{\nu} N^{\nu}_{\lambda\bar{\lambda}} \mathbb{S}_\nu(\mathbb{C}^n). \end{aligned}$$

The second isomorphism holds by Cauchy's formula; for the third one, see, for example, [5, 15.50]; the fourth isomorphism is the Littlewood–Richardson rule. □

To compute the decomposition of $S^k(\mathfrak{sl}_n)$, we simply note that

$$S^k(\mathfrak{gl}_n) \cong S^k(\mathfrak{sl}_n \oplus \mathbb{C}) \cong \mathbb{C} \oplus \bigoplus_{i=1}^{k} S^i(\mathfrak{sl}_n).$$

This allows us to compute the decomposition of $S^k(\mathfrak{sl}_n)$ inductively.

As a corollary we present an explicit decomposition in the case $k = 3$. Computing the Littlewood–Richardson coefficients in (1) gives us the decomposition of $S^3(\mathfrak{gl}_n)$ (resp. $S^3(\mathfrak{sl}_n)$) into irreducibles. We present these in Table 1: the first column lists the highest weights λ of the occurring irreducible representations $\mathbb{S}_\lambda(\mathbb{C}^n)$. To be more precise, the first column actually shows the highest weights when we view $S^3(\mathfrak{gl}_n)$ (resp. $S^3(\mathfrak{sl}_n)$) as a GL_n-representation. (Recall that weights of GL_n are n-tuples $[\lambda_1, \dots, \lambda_n] \in \mathbb{Z}^n$ with $\lambda_1 \geq \dots \geq \lambda_n$. The corresponding SL_n-weight is then $[\lambda_1 - \lambda_n, \dots, \lambda_{n-1} - \lambda_n]$.) The second and third column list the multiplicities of the irreducibles in $S^3(\mathfrak{gl}_n)$ resp. $S^3(\mathfrak{sl}_n)$. We also list the dimensions of the occurring irreducible representations $\mathbb{S}_\lambda(\mathbb{C}^n)$, as well as the dimensions of the projective homogeneous varieties contained in $\mathbb{P}(\mathbb{S}_\lambda(\mathbb{C}^n))$ (see Subsection 2.1).

Table 1
Irreducible components of $S^3(\mathfrak{gl}_n)$ and $S^3(\mathfrak{sl}_n)$.

Highest weight	$S^3(\mathfrak{gl}_n)$	$S^3(\mathfrak{sl}_n)$	Dimension	Variety
$[0,\dots,0]$	3	1	1	0
$[1,0,\dots,0,-1]$	4	2	n^2-1	$2n-3$
$[2,0,\dots,0,-2]$	2	1	$\frac{(n-1)n^2(n+3)}{4}$	$2n-3$
$[3,0,\dots,0,-3]$	1	1	$\frac{(n-1)n^2(n+1)^2(n+5)}{36}$	$2n-3$
$[1,1,0,\dots,0,-1,-1]$	2	1	$\frac{(n-3)n^2(n+1)}{4}$	$4n-12$
$[2,0,\dots,0,-1,-1]$	1	1	$\frac{(n-2)(n-1)(n+1)(n+2)}{4}$	$3n-7$
$[1,1,0,\dots,0,-2]$	1	1	$\frac{(n-2)(n-1)(n+1)(n+2)}{4}$	$3n-7$
$[2,1,0,\dots,0,-1,-2]$	1	1	$\frac{(n-3)(n-1)^2(n+1)^2(n+3)}{9}$	$4n-10$
$[1,1,1,0,\dots,0,-1,-1,-1]$	1	1	$\frac{(n-5)(n-1)^2n^2(n+1)}{36}$	$6n-27$

Table 2
Highest-weight vectors of $S^3(\mathfrak{gl}_n)$.

Weight	Highest-weight vector
$[0,\dots,0]$	III
$[0,\dots,0]$	$\sum_{i,j} IE_{i,j}E_{j,i}$
$[0,\dots,0]$	$\sum_{i,j,k} E_{i,j}E_{j,k}E_{k,i}$
$[1,0,\dots,0,-1]$	$IIE_{1,n}$
$[1,0,\dots,0,-1]$	$\sum_i IE_{1,i}E_{i,n}$
$[1,0,\dots,0,-1]$	$\sum_{i,j} E_{1,n}E_{i,j}E_{j,i}$
$[1,0,\dots,0,-1]$	$\sum_{i,j} E_{1,i}E_{i,j}E_{j,n}$
$[2,0,\dots,0,-2]$	$IE_{1,n}E_{1,n}$
$[2,0,\dots,0,-2]$	$\sum_i E_{1,n}E_{1,i}E_{i,n}$
$[1,1,0,\dots,0,-2]$	$\sum_i E_{1,n}E_{2,i}E_{i,n} - E_{2,n}E_{1,i}E_{i,n}$
$[2,0,\dots,0,-1,-1]$	$\sum_i E_{1,n}E_{1,i}E_{i,n-1} - E_{1,n-1}E_{1,i}E_{i,n}$
$[1,1,0,\dots,0,-1,-1]$	$IE_{1,n}E_{2,n-1} - IE_{1,n-1}E_{2,n}$
$[1,1,0,\dots,0,-1,-1]$	$\sum_i E_{1,n}E_{2,i}E_{i,n-1} - E_{2,n}E_{1,i}E_{i,n-1} - E_{1,n-1}E_{2,i}E_{i,n} + E_{2,n-1}E_{1,i}E_{i,n}$
$[3,0,\dots,0,-3]$	$E_{1,n}E_{1,n}E_{1,n}$
$[2,1,0,\dots,0,-1,-2]$	$E_{1,n}E_{1,n-1}E_{2,n} - E_{1,n}E_{1,n}E_{2,n-1}$
$[1,1,1,0,\dots,0,-1,-1,-1]$	$\sum_{\sigma\in S_3} \operatorname{sgn}\sigma\, E_{\sigma(1),n}E_{\sigma(2),n-1}E_{\sigma(3),n-2}$

2.1. Homogeneous varieties

Let V be an irreducible representation of a semisimple Lie group G. Then $\mathbb{P}V$ has a unique closed G-orbit X, which is the orbit of the highest-weight vector in $\mathbb{P}V$ under the action of G. The projective variety X is isomorphic to G/P, where P is a parabolic subgroup. We call these varieties homogeneous varieties or partial flag varieties.

In our case, $G = SL_n$, we can compute the dimension of X in the following way. Consider the Dynkin diagram of $\mathfrak{sl}_n$, which consists of $n-1$ dots marked from 1 to $n-1$, and the Young diagram λ associated with the representation V. For every $j \in \{1,\dots,n-1\}$, if the Young diagram has at least one column of length j, we remove the dot j from the Dynkin diagram. After removing these dots, the Dynkin diagram splits in connected components of size k_i. The dimension of our variety X is then given by

$$\frac{1}{2}\left(n^2 - n - \sum_i (k_i^2 + k_i)\right).$$

This gives us the last column of Table 1.

3. Highest-weight vectors

We now describe highest-weight vectors for all irreducible components of $S^3(\mathfrak{gl}_n)$. We write $E_{i,j} \in \mathfrak{gl}_n$ for the $n \times n$ matrix with as only nonzero entry a 1 on position (i, j). Note that the vector $E_{i,j}E_{i',j'}E_{i'',j''} \in S^3(\mathfrak{gl}_n)$ has weight $e_i + e_{i'} + e_{i''} - e_j - e_{j'} - e_{j''}$, where e_i is the weight $[0,\dots,1,\dots,0]$ with a 1 on the i-th position. Furthermore, to check that a weight vector v in some representation V of SL_n is a highest-weight vector, it suffices to view V as a representation of the Lie algebra $\mathfrak{sl}_n$ and check that every matrix $E_{i,i+1}$ acts by zero. Using this, it is straightforward to check that the vectors listed in Table 2 are indeed highest-weight vectors.

3.1. Waring rank and border Waring rank

As explained in the introduction (see also [2]), estimating the (border) Waring rank of the highest-weight vector $\sum_{i,j,k} E_{i,j}E_{j,k}E_{k,i}$ is equivalent to determining the exponent ω of matrix multiplication. We will analyze the (border) Waring ranks of other highest-weight vectors. We start with the following surprising observation.

Observation 1. *Every highest-weight vector with weight different than* $[0,\ldots,0]$ *has Waring rank* $O(n^2)$*. Furthermore, the weight space of* $[0,\ldots,0]$ *is 3-dimensional: it has a basis consisting of two vectors of Waring rank* $O(n^2)$*, and the vector* $\sum_{i,j,k} E_{i,j}E_{j,k}E_{k,i}$*.*

Proof. Every highest-weight vector in Table 2, except $\sum_{i,j,k} E_{i,j}E_{j,k}E_{k,i}$, is a sum of at most n^2 monomials, and every degree 3 monomial has Waring rank at most 4. □

We now study the highest-weight vectors $IE_{1,n}E_{2,n-1} - IE_{1,n-1}E_{2,n}$ and $E_{1,n}E_{1,n-1}E_{2,n} - E_{1,n}E_{1,n}E_{2,n-1}$, which we will rewrite as $xyz - xwt$ and $xzt - x^2y$.

Proposition 1. *The cubics* $f_1 = xyz - xwt$ *and* $f_2 = xzt - x^2y$ *are two variants of the Coppersmith–Winograd tensor. Their ranks and border ranks (equal to Waring rank resp. Waring border rank) are given by* $\mathrm{rk}(f_1) = 9$, $\underline{rk}(f_1) = 6$, $\mathrm{rk}(f_2) = 7$, $\underline{rk}(f_2) = 4$.

Proof. After the change of basis $x = x_0$, $y = x_1 + \mathrm{i}x_2$, $z = x_1 - \mathrm{i}x_2$, $w = x_3 + \mathrm{i}x_4$, $t = -x_3 + \mathrm{i}x_4$, our cubic f_1 becomes $x_0x_1^2 + x_0x_2^2 + x_0x_3^2 + x_0x_4^2$, which is precisely the Coppersmith–Winograd tensor $T_{4,CW}$ (here we use the notation from [11, Section 7]). For f_2, we can do a similar change of basis, or alternatively we can use the geometric characterization of the Coppersmith–Winograd tensor form [11, Theorem 7.4]. We find that f_2 is isomorphic to $\tilde{T}_{2,CW}$.

The ranks and border ranks of Coppersmith–Winograd tensors are known: consult for example [4] for the border ranks and [11, Proposition 7.1] for the ranks. □

Remark 1. The highest-weight vectors that are monomials are easily understood: III and $E_{1,n}E_{1,n}E_{1,n}$ trivially have Waring rank equal to 1; $IIE_{1,n}$ and $IE_{1,n}E_{1,n}$ agree with the Coppersmith–Winograd tensor $T_{1,CW}$, hence they have Waring rank 3 and border Waring rank 2.

Acknowledgement

The author would like to thank his advisor, Mateusz Michałek, for the many helpful comments and discussions.

References

[1] P. Bürgisser, M. Clausen, M.A. Shokrollahi, Algebraic Complexity Theory, Grundlehren der Mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences), vol. 315, Springer-Verlag, Berlin, 1997. With the collaboration of Thomas Lickteig.
[2] L. Chiantini, J.D. Hauenstein, C. Ikenmeyer, J.M. Landsberg, G. Ottaviani, Polynomials and the exponent of matrix multiplication, arXiv preprint, arXiv:1706.05074, 2017.
[3] H. Cohn, C. Umans, A group-theoretic approach to fast matrix multiplication, in: Foundations of Computer Science, 2003. Proceedings. 44th Annual IEEE Symposium on, IEEE, 2003, pp. 438–449.
[4] D. Coppersmith, S. Winograd, Matrix multiplication via arithmetic progressions, J. Symb. Comput. 9 (3) (1990) 251–280.
[5] W. Fulton, J. Harris, Representation Theory: A First Course, vol. 129, Springer Science & Business Media, 2013.
[6] R. Howe, $(\mathrm{GL}_n, \mathrm{GL}_m)$-duality and symmetric plethysm, Proc. Indian Acad. Sci. Math. Sci. 97 (1–3) (1988) 85–109, 1987.
[7] T. Kahle, M. Michałek, Plethysm and lattice point counting, Found. Comput. Math. 16 (5) (2016) 1241–1261.
[8] J.M. Landsberg, Tensors: Geometry and Applications, Graduate Studies in Mathematics, vol. 128, American Mathematical Society, Providence, RI, 2012.
[9] J.M. Landsberg, Geometry and Complexity Theory, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2017.
[10] J.M. Landsberg, M. Michałek, On the geometry of border rank decompositions for matrix multiplication and other tensors with symmetry, SIAM J. Appl. Algebra Geom. 1 (1) (2017) 2–19.
[11] J.M. Landsberg, M. Michałek, Abelian tensors, J. Math. Pures Appl. 108 (3) (2017) 333–371.
[12] J.M. Landsberg, M. Michałek, A lower bound for the border rank of matrix multiplication, Int. Math. Res. Not. (2017), rnx025.
[13] J.M. Landsberg, G. Ottaviani, New lower bounds for the border rank of matrix multiplication, Theory Comput. 11 (2015) 285–298.
[14] F. Le Gall, Powers of tensors and fast matrix multiplication, in: ISSAC 2014—Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation, ACM, New York, 2014, pp. 296–303.
[15] I.G. Macdonald, Symmetric Functions and Hall Polynomials, Oxford University Press, 1998.
[16] L. Manivel, M. Michałek, Secants of minuscule and cominuscule minimal orbits, Linear Algebra Appl. 481 (2015) 288–312.
[17] S.P.O. Plunkett, On the plethysm of S-functions, Can. J. Math. 24 (1972) 541–552.
[18] V. Strassen, Gaussian elimination is not optimal, Numer. Math. 13 (4) (1969) 354–356.
[19] R.M. Thrall, On symmetrized Kronecker powers and the structure of the free Lie ring, Amer. J. Math. 64 (1942) 371–388.
[20] V.V. Williams, Multiplying matrices faster than Coppersmith–Winograd [extended abstract], in: STOC'12—Proceedings of the 2012 ACM Symposium on Theory of Computing, ACM, New York, 2012, pp. 887–898.