Research Article   **Open Access**

Orthogonal Matrix Invariants

Lopatin AA*

Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Russia

Abstract

The orthogonal group acts on the space of several n × n matrices by simultaneous conjugation. For an infinite field of characteristic different from two, relations between generators for the algebra of invariants are described. As an application, the maximal degree of elements of a minimal system of generators is described with deviation 3.

Keywords: Invariant theory; Classical linear groups; Polynomial identities; Generators

Introduction

All vector spaces, algebras, and modules are over an infinite field $\mathbb{F}$. By an algebra we always mean an associative algebra.

Let a linear group G be a subgroup of $GL(n)$ and

$$V = \mathbb{F}^{n \times n} \oplus \cdots \oplus \mathbb{F}^{n \times n}$$

be d-tuple of $n \times n$ matrices over $\mathbb{F}$. The group G acts on V by the diagonal conjugation, i.e.,

$$g \cdot (A_1, \ldots, A_d) = (gA_1 g^{-1}, \ldots, gA_d g^{-1}), \tag{1}$$

Where $g \in G$ and $A_1, \ldots, A_d \in \mathbb{F}^{n \times n}$.

The coordinate ring of V (i.e. the ring of polynomial functions $f : V \to \mathbb{F}$) is the ring of polynomials

$$\mathbb{F}[V] = \mathbb{F}[x_{ij}(k) \,|\, 1 \le i, j \le n, 1 \le k \le d],$$

where $x_{ij}(k)$ stands for the coordinate function on V that takes $(A_1, \ldots, A_d) \in V$ to the $(i,j)^{\text{th}}$ entry of the matrix A_k. Denote by

$$X_k = \begin{pmatrix} x_{11}(k) & \cdots & x_{1n}(k) \\ \vdots & & \vdots \\ x_{n1}(k) & \cdots & x_{nn}(k) \end{pmatrix}$$

the kth *generic* matrix $(1 \le k \le n)$.

The action of G on V induces the action on $\mathbb{F}[V]$ as follows: $(g \cdot f)(h) = f(g^{-1} \cdot h)$ for all $g \in G, f \in \mathbb{F}[V]$ $h \in V$. In other words,

$$g \cdot x_{ij}(k) = (i,j)^{th} \text{ entry of } g^{-1} X_k g.$$

The algebra of *matrix G-invariants* is

$$\mathbb{F}[V]^G = \{f \in \mathbb{F}[V] \,|\, g \cdot f = f \text{ for all } g \in G\}.$$

If G is a classical linear group, i.e., G belongs to the list $GL(n)$, $O(n)$, $Sp(n)$, $SO(n)$, $SL(n)$, then a generating set for the algebra $\mathbb{F}[V]^G$ is known for an arbitrary characteristic of $\mathbb{F}$ [1-3], where we assume that char $\ne 2$ in the case of $O(n)$ and $SO(n)$.

Denote coefficients in the characteristic polynomial of an $n \times n$ matrix X by $\sigma_t(X)$, i.e.,

$$\det(X + \lambda E) = \sum_{t=0}^{n} \lambda^{n-t} \sigma_t(X). \tag{2}$$

So, $\sigma_o(X) = 1$, $\sigma_1(X) = \text{tr}(X)$ and $\sigma_n(X) = \det(X)$. The following result holds [3].

Theorem 1.1. *If char $\mathbb{F} \ne 2$, then the algebra of matrix invariants*

$\mathbb{F}[V]^{O(n)}$ *is generated by $\sigma_t(B)$ $(1 \le t \le n)$, where B ranges over all monomials in $X_1, \ldots, X_d, X_1^T, \ldots, X_d^T$. Moreover, we can assume that B is primitive, i.e., is not equal to a power of a shorter monomial.*

Remark 1.2. In the case of a characteristic zero field it is enough to take traces instead of σ_t, $1 \le t \le n$, in the formulation of Theorem 1.1.

In characteristic zero case Procesi [4] described relations between generators for $\mathbb{F}[V]^G$ for $G \in \{GL(n), O(n), Sp(n)\}$. Zubkov [5] described relations for matrix $GL(n)$-invariants over a field of arbitrary characteristic. We have described relations for matrix $O(n)$-invariants (see Theorem 2.1) in the case of characteristic different from two. The proofs of Theorem 2.1 and Lemma 3.1 will be given in a separate paper.

Relations

For a vector $\underline{t} = (t_1, \ldots, t_u) \in \mathbb{N}^u$ we write $\#\underline{t} = u$, where $\mathbb{N}$ stands for the set of non-negative integers. In this paper we use the following notions:

The monoid $\mathcal{M}$ (without unity) freely generated by *letters* $x_1, \ldots, x_d, x_1^T, \ldots, x_d^T$, the vector space $\mathcal{M}_{\mathbb{F}}$ with the basis $\mathcal{M}$, and $\mathcal{N} \subset \mathcal{M}$ the subset of primitive elements, where the notion of a primitive element is defined as above;

The involution $^T : \mathcal{M}_{\mathbb{F}} \to \mathcal{M}_{\mathbb{F}}$ defined by $x^{TT} = x$ for a letter x and $(a_1 \ldots a_p)^T = a_p^T \ldots a_1^T$ for $a_1, \ldots, a_p \in \mathcal{M}$;

The equivalence $y_1 \ldots y_p \sim z_1 \ldots z_p$ that holds if there exists a cyclic permutation $\pi \in S_p$ such that $y_{\pi(1)} \ldots y_{\pi(p)} = z_1 \ldots z_p$ or $y_{\pi(1)} \ldots y_{\pi(p)} = z_p^T \ldots z_1^T$, where $y_1, \ldots, y_p, z_1, \ldots, z_p$ are letters;

$\mathcal{M}_\sigma$, the ring with unity of (commutative) polynomials over $\mathbb{F}$ freely generated by the "symbolic" elements $\sigma_t(\alpha)$, where $t > 0$ and $\alpha \in \mathcal{M}\mathbb{F}$;

$\mathcal{N}_\sigma$, a ring with unity of (commutative) polynomials over $\mathbb{F}$ freely generated by the "symbolic" elements $\sigma_t(\alpha)$, where $t > 0$ and $\alpha \in \mathbb{N}$ ranges over $\sim$-equivalence classes; note that $\mathcal{N}_\sigma \sim \mathcal{M}_\sigma / L$, where the ideal L is described in Lemma 3.1 (see below).

We will use the following notation:

*Corresponding author: Lopatin AA, Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Pevtsova street, 13, Omsk 644099, Russia, Tel: +73833332892; E-mail: artem_lopatin@yahoo.com

$tr(\alpha) = \sigma_t(\alpha)$

any $\alpha \in \mathcal{M}_{\mathbb{F}}$.

For a letter $\beta \in \mathcal{M}$ define

$$X_\beta = \begin{cases} X_i, \text{if } \beta = x_i \\ X_i^T, \text{if } \beta = x_i^T \end{cases}.$$

Given $\alpha = \alpha_1 \ldots \alpha_p \in \mathcal{M}$ where α_i is a letter, we assume that $X_\alpha = X_{\alpha_1} \ldots X_{\alpha_p}$. Consider a surjective homomorphism of algebras

$$\Psi_n : \mathcal{N}_\sigma \to \mathbb{F}[V]^{O(n)}$$

defined by $\sigma_t(\alpha)\ldots \sigma_t(X_\alpha)$, if $t \le n$, and $\sigma_t(\alpha) \to 0$ otherwise. Its kernel K_n is the ideal of *relations* for $\mathbb{F}[V]^{O(n)}$. Elements of $\bigcap_{i>0} K_i$ are called *free* relations. For $\alpha \beta, \gamma \in \mathcal{M}_{\mathbb{F}}$ and $t, r \in \mathbb{N}$, an element $t,r(\alpha, \beta, \gamma) \in \mathcal{N}_\sigma$ was introduced in "Invariants of mixed representations of quivers II: defining relations and applications" [6] (see Definition 3.3 below).

Theorem 2.1. *If char $\mathbb{F} \ne 2$, then the ideal of relations K_n for $\mathbb{F}[V]^{O(n)} \sim \mathcal{N}/K_n$ is generated by $\sigma_{tr}(\alpha, \beta, \gamma)$, where $t + 2r > n$ and $\alpha \beta, \gamma \in \mathcal{MF}$.*

The following result was known in characteristic zero case.

Corollary 2.2. *In case $G = O(n)$ there are no non-zero free relations.*

The Definition of $\sigma_{t,r}$

In this section we assume that $\mathcal{A}$ is a commutative unitary algebra over the field $\mathbb{F}$ and all matrices are considered over $\mathcal{A}$.

Let us recall some formulas. In what follows $A, A_1, \ldots, A_p$ stand for $n \times n$ matrices and $1 \le t \; n$. Amitsur's formula states [7]:

$$\sigma_t(A_1 + \cdots + A_p) = \sum (-1)^{t-(j_1 + \cdots + j_q)} \sigma_{j_1}(\gamma_1) \cdots \sigma_{j_q}(\gamma_q), \qquad (3)$$

where the sum ranges over all pairwise different primitive cycles $\gamma_1, \ldots, \gamma_q$ in letters $A_1, \ldots, A_p$ and positive integers $j_1, \ldots, j_p$ with $\sum_{i=1}^{q} j_i \deg \gamma_i = t$. Denote the right hand side of (3) by $F_{t,p}(A_1, \ldots, A_p)$.

As an example,

$$\sigma_2(A_1 + A_2) = \sigma_2(A_1) + \sigma_2(A_2) + tr(A_1)\,tr(A_2) - tr(A_1 A_2). \qquad (4)$$

Note that for $a \in \mathcal{A}$ we have

$$\sigma_t(aA) = a^t \sigma_t(A). \qquad (5)$$

For $l \ge 2$ we have the following well-known formula:

$$\sigma_t(A^l) = \sum_{i_1, \ldots, i_{tl} \ge 0} b_{i_1, \ldots, i_{tl}}^{(t,l)} \sigma_1(A)^{i_1} \cdots \sigma_{tl}(A)^{i_{tl}}, \qquad (6)$$

where we assume that $\sigma_i(A) = 0$ for $i > n$. Denote the right hand side of (6) by $P_{t,l}(A)$. In (6) coefficients $b_{i_1, \ldots, i_{rl}}^{(t,l)} \in \mathbb{Z}$ do not depend on A and n. If we take a diagonal matrix $A = diag(a_1, \ldots, a_n)$, then $_i(A^l)$ is a symmetric polynomial in $a_1, \ldots, a_n$ and $\sigma_i(A)$ is the ith elementary symmetric polynomial in $a_1, \ldots, a_n$, where $1 \le i \le n$. Thus the coefficients $b_{i_1, \ldots, i_{tl}}^{(t,l)}$ with $tl \le n$ can easily be found. As an example,

$$tr(A^2) = tr(A)^2 - 2\sigma_2(A). \qquad (7)$$

Lemma 3.1. *We have $\mathcal{N}_\sigma \sim {}_\sigma/L$ for the ideal L generated by*

$(a)\, \sigma_t(\alpha_1 + \cdots + \alpha_p) - F_{t,p}(\alpha_1, \ldots, \alpha_p),$

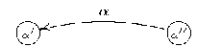$(b)\, \sigma_t(a\alpha) - a^t \sigma_t(\alpha),$

$(c)\, \sigma_t(\alpha^t) - P_{t,l}(\alpha),$

$(d)\, \sigma_t(a\beta) - \sigma_t(\beta\alpha),$

$(e)\, \sigma_t(\alpha) - \alpha_t(\alpha^T),$

where $p > 1$, $\alpha, \alpha_1 \ldots \alpha_p \in \mathcal{M}_{\mathbb{F}}$, $\alpha \in \mathbb{F}$, $t > 0$ and $l > 1$.

A quiver $\mathcal{Q} = (\mathcal{Q}_0, \mathcal{Q}_1)$ is a finite oriented graph, where $\mathcal{Q}_0$ is the set of vertices and $\mathcal{Q}_1$ is the set of arrows. Multiple arrows and loops in Q are allowed. For an arrow α, denote by α' its head and by α'' its tail, i.e.,

We say that $\alpha = \alpha_1 \ldots \alpha_s$ is a path in $\mathcal{Q}$ (where $\alpha_1 \ldots \alpha_s \in \mathcal{Q}_1$), if $\alpha_1'' = \alpha_2', \ldots, \alpha''_{s-1} = \alpha_s'$, i.e.,

The head of the path α is $\alpha' = \alpha_1'$ and the tail is $\alpha'' = \alpha_s''$. A path α is called closed if $\alpha' = \alpha''$.

Definition 3.2 (*of a mixed quiver*). A quiver $\mathcal{Q}$ is called mixed if there are two maps $^T : \mathcal{Q}_0 \to \mathcal{Q}_0$ and $^T : \mathcal{Q}_1 \to \mathcal{Q}_1$ satisfying

- $v^{TT} = v, \beta^{TT} = \beta;$

- $(\beta^T)' = (\beta'')^T, (\beta^T)'' = (\beta')^T$

for all $v \in \mathcal{Q}_0$ and $\beta \in \mathcal{Q}_1$.

Assume that $\mathcal{Q}$ is a mixed quiver. Denote by $\mathcal{M}(\mathcal{Q})$ the set of all closed paths in $\mathcal{Q}$ and denote by $\mathcal{N}(\mathcal{Q}) \subset \mathcal{M}(\mathcal{Q})$ the subset of primitive paths. Given a path α in $\mathcal{Q}$, we define the path α^T and introduce $\sim$-equivalence on $\mathcal{M}(\mathcal{Q})$ in the same way as in Section 2. Moreover, we define $\mathcal{MF}(\mathcal{Q})$, $\mathcal{M}_\sigma(\mathcal{Q})$, and $\mathbb{N}(\mathcal{Q})$ in the same way as $\mathcal{MF}$, $\mathcal{M}_\sigma$, and $\mathcal{N}_\sigma$ have been defined in Section 2.

Let $t, r \in \mathbb{N}$. In order to define $_{t,r}$, we consider the quiver $\mathcal{Q}$
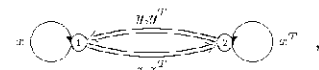
where there are two arrows from vertex 1 to vertex 2 and there are two arrows in the opposite direction. We define $1^T = 2$ for vertex 1 to turn $\mathcal{Q}$ into a mixed quiver. Denote the degree of $\alpha \in \mathcal{M}(\mathcal{Q})$ by $\deg \alpha$, the degree of α in a letter β by $\deg_\beta \alpha$, and the multidegree of α by

$$mdeg(\alpha) = (\deg_x \alpha + \deg_{x^T} \alpha, \deg_y \alpha + \deg_{y^T} \alpha, \deg_z \alpha + \deg_{z^T} \alpha).$$

Definition 3.3 (of $\sigma_{t,r}(x, y, z)$). Denote by $\mathcal{I} = \mathcal{I}_{t,r}$ the set of pairs $(\underline{j}, \underline{\alpha})$ such that

- $\#\underline{j} = \#\underline{\alpha} = p$ for some p;

$\alpha_1, \ldots, \alpha_p \in \mathcal{N}(\mathcal{Q})$ belong to pairwise different $\sim$-equivalence classes and $j_1, \ldots, j_p \ge 1$;

- $j_1\, mdeg(\alpha_1) + \ldots + j_p\, mdeg(\alpha_p) = (t, r, r).$

Then we define $\sigma_{t,r}(x, y, z) \in \mathcal{N}_\sigma(\mathcal{Q})$ by

$$\sigma_{t,r}(x, y, z) = \sum_{(\underline{j}, \underline{\alpha}) \in \mathcal{I}} (-1)^\xi \sigma_{j_1}(\alpha_1) \cdots \sigma_{j_p}(\alpha_p), \qquad (8)$$

where $p = \#\underline{j} = \#\underline{\alpha}$ and $\xi = \xi_{\underline{j}, \underline{\alpha}} = t + \sum_{i=1}^{p} j_i (\deg_y \alpha_i + \deg_z \alpha_i + 1)$. For t

$= r = 0$ we define $\sigma_{0,0}(x, y, z) = 1$. Moreover, if $\alpha, \beta, \gamma \in \mathcal{MF}$, then we define $\sigma_{t,r}(\alpha, \beta, \gamma) \in \mathcal{N}_\sigma$ as the result of substitution $x \to \alpha, y \to \beta, z \to \gamma$ in (8).

Example 3.4.

1. If $t = 0$ and $r = 1$, then $\sim$-equivalence classes on $\mathcal{N}(\mathcal{Q})$ are yz, $yz^T,...$ Hence, $\sigma_{0,1}(x, y, z) = - \mathrm{tr}(yz) + \mathrm{tr}(yz^T)$.

2. If $t = r = 1$, then $\sim$-equivalence classes on $\mathcal{N}(\mathcal{Q})$ are

$x, yz, yz^T, xyz, xyz^T, xy^T z, xy^T z^T, \ldots$

and we can see that $\sigma_{1,1}(x, y, z) =$

$- \mathrm{tr}(x)\,\mathrm{tr}(yz) + \mathrm{tr}(x)\,\mathrm{tr}(yz^T) + \mathrm{tr}(xyz) - \mathrm{tr}(xyz^T) - \mathrm{tr}(xy^T z) + \mathrm{tr}(xy^T z^T).$

Remark 3.5. $\sigma_{t,0}(x, y, z) = \sigma_t(x)$.

The decomposition formula from "On block partial linearizations of the pfaffian"[8] implies that for $n \times n$ matrices A_i, $i = 1, 2, 3$, with $n = t_0 + 2r$, $t_0 \geq 0$ we have

$$\mathrm{DP}_{r,r}(A_1 + \lambda E, A_2, A_3) = \sum_{t=0}^{t_0} \lambda^{t_0 - t}\sigma_{t,r}(A_1, A_2, A_3),$$

where $\mathrm{DP}_{r,r}(A_1, A_2, A_3)$ stands for the determinant-pfaffian [9] and $\sigma_{t,r}(A_1, A_2, A_3)$ is defined as the result of the substitution $a_i \to A_i$, $a_i^T \to A_i^T$ in $\sigma_{t,r}(a_1, a_2, a_3)$. Thus $\mathrm{DP}_{r,r}$ relates to $\sigma_{t,r}$ in the same way as the determinant relates to σ_t.

Application

Given an $\mathbb{N}$-graded algebra $\mathcal{A}$, denote by $\mathcal{A}^+$ the subalgebra generated by elements of $\mathcal{A}$ of positive degree. It is easy to see that a set $\{a_i\} \subseteq \mathcal{A}$ is a minimal (by inclusion) homogeneous system of generators (m.h.s.g.) for $\mathcal{A}$ if and only if $\{\overline{a_i}\}$ is a basis for $\overline{\mathcal{A}} = \mathcal{A} / (\mathcal{A}^+)^2$ and $\{a_i\}$ are homogeneous. Let us recall that an element $a \in \mathcal{A}$ is called *decomposable* if it belongs to the ideal $(\mathcal{A}^+)^2$. Therefore the least upper bound for the degrees of elements of a m.h.s.g. for $\mathbb{F}[V]^{O(n)}$ is equal to the maximal degree of indecomposable invariants and we denote it by $D_{\max}$.

As an application of Theorem 2.1 we obtained the following result in "On minimal generating systems for matrix O(3)-invariants"[10].

Theorem 4.1. *Let* $n = 3$ *and* $d \geq 1$. *Then*

- *If char* $\mathbb{F} = 3$, *then* $2d + 4 \leq D_{\max} \leq 2d + 7$.

- *If char* $\mathbb{F} \neq 2, 3$, *then* $D_{\max} = 6$.

Moreover, in "Orthogonal invariants of skew-symmetric matrices" [11] we described a m.h.s.g. for orthogonal invariants of d-tuples of 3×3 skew-symmetric matrices. As about matrix $GL(n)$-invariants in case $n = 3$, its minimal system of generators was explicitly calculated [12,13].

Acknowledgements

References

1. Donkin S (1992) Invariants of several matrices. Invent Math 110: 389-401.

2. Lopatin AA (2009) Invariants of quivers under the action of classical groups. J Algebra 321: 1079-1106.

3. Zubkov AN (1999) Invariants of an adjoint action of classical groups. Algebra and Logic 38: 299-318.

4. Procesi C (1976) The invariant theory of n × n matrices. Adv Math 19: 306-381.

5. Zubkov AN (1996) On a generalization of the Razmyslov-Procesi theorem. Algebra and Logic 35: 241-254.

6. Zubkov AN (2005) Invariants of mixed representations of quivers II: defining relations and applications. J Algebra Appl 4: 287-312.

7. Amitsur SA (1980) On the characteristic polynomial of a sum of matrices. Linear and Multilinear Algebra 8: 177-182.

8. Lopatin AA (2007) On block partial linearizations of the pfaffian. Linear Algebra Appl 426: 109-129.

9. Lopatin AA, Zubkov AN (2007) Semi-invariants of mixed representations of quivers. Transform Groups 12: 341-369.

10. Lopatin AA (2011) On minimal generating systems for matrix O(3)-invariants. Linear Mult. Algebra 59: 87-99.

11. Lopatin AA (2011) Orthogonal invariants of skew-symmetric matrices. Linear Mult. Algebra 59: 851-862.

12. Lopatin AA (2004) The Algebra of Invariants of 3×3 Matrices Over a Field of Arbitrary Characteristic. Comm. Algebra 32: 2863-2883.

13. Lopatin AA (2006) Relatively Free Algebras with the Identity $x^3 = 0$. Comm. Algebra 33: 3583-3605.