Invariant Theory with Applications

Jan Draisma and Dion Gijswijt

September 9 2009

Contents

Chapter 1

Lecture1: introducing invariant theory

The first lecture gives some flavor of the theory of invariants. Basic notions such as (linear) group representation, the ring of regular functions on a vector space and the ring of invariant functions are defined, and some instructive examples are given.

1.1 Polynomial functions

Let V be a complex vector space. We denote by $V^* := \{f : V \to \mathbb{C} \text{ linear map}\}$ the dual vector space. Viewing the elements of V^* as functions on V, and taking the usual pointwise product of functions, we can consider the algebra of all $\mathbb{C}$-linear combinations of products of elements from V^*.

Definition 1.1.1. The *coordinate ring* $\mathcal{O}(V)$ of the vectorspace V is the algebra of functions $F : V \to \mathbb{C}$ generated by the elements of V^*. The elements of $\mathcal{O}(V)$ are called *polynomial* or *regular* functions on V.

If we fix a basis $e_1, \ldots, e_n$ of V, then a dual basis of V^* is given by the coordinate functions $x_1, \ldots, x_n$ defined by $x_i(c_1e_1 + \cdots + c_ne_n) := c_i$. For the coordinate ring we obtain $\mathcal{O}(V) = \mathbb{C}[x_1, \ldots, x_n]$. This is a polynomial ring in the x_i, since our base field $\mathbb{C}$ is infinite.

Exercise 1.1.2. Show that indeed $\mathbb{C}[x_1, \ldots, x_n]$ is a polynomial ring. In other words, show that the x_i are algebraically independent over $\mathbb{C}$: there is no nonzero polynomial $p \in \mathbb{C}[X_1, \ldots, X_n]$ in n variables $X_1, \ldots, X_n$, such that $p(x_1, \ldots, x_n) = 0$. Hint: this is easy for the case $n = 1$. Now use induction on n.

We call a regular function $f \in \mathcal{O}(V)$ *homogeneous of degree* d if $f(tv) = t^d f(v)$ for all $v \in V$ and $t \in \mathbb{C}$. Clearly, the elements of V^* are regular of degree

1, and the product of polynomials f, g homogeneous of degrees d, d' yields a homogeneous polynomial of degree $d+d'$. It follows that every regular function f can be written as a sum $f = c_0 + c_1 f_1 + \cdots + c_k f_k$ of regular functions f_i homogeneous of degree i. This decomposition is unique (disregarding the terms with zero coefficient). Hence we have a direct sum decomposition $\mathcal{O}(V) = \bigoplus_{d\in\mathbb{N}} \mathcal{O}(V)_d$, where $\mathcal{O}(V)_d := \{f \in \mathcal{O}(V) \mid f \text{ homogeneous of degree } d\}$, making $\mathcal{O}(V)$ into a *graded algebra*.

Exercise 1.1.3. Show that indeed the decomposition of a regular function f into its homogeneous parts is unique.

In terms of the basis $x_1, \ldots, x_n$, we have $\mathcal{O}(V)_d = \mathbb{C}[x_1, \ldots, x_n]_d$, where $\mathbb{C}[x_1, \ldots, x_n]_d$ consists of all polynomials of total degree d and has as basis the monomials $x_1^{d_1} x_2^{d_2} \cdots x_n^{d_n}$ for $d_1 + d_2 + \cdots + d_n = d$.

1.2 Representations

Central objects in this course are linear representations of groups. For any vector space V we write $\mathrm{GL}(V)$ for the group of all invertible linear maps from V to itself. When we have a fixed basis of V, we may identify V with $\mathbb{C}^n$ and $\mathrm{GL}(V)$ with the set of invertible matrices $n \times n$ matrices $\mathrm{GL}(\mathbb{C}^n) \subset \mathrm{Mat}_n(\mathbb{C})$.

Definition 1.2.1. Let G be a group and let X be a set. An *action of G on X* is a map $\alpha : G \times X \to X$ such that $\alpha(1, x) = x$ and $\alpha(g, \alpha(h, x)) = \alpha(gh, x)$ for all $g, h \in G$ and $x \in X$.

If α is clear from the context, we will usually write gx instead of $\alpha(g, x)$. What we have just defined is sometimes called a *left action* of G on X; *right actions* are defined similarly.

Definition 1.2.2. If G acts on two sets X and Y, then a map $\phi : X \to Y$ is called *G-equivariant* if $\phi(gx) = g\phi(x)$ for all $x \in X$ and $g \in G$. As a particular case of this, if X is a subset of Y satisfying $gx \in X$ for all $x \in X$ and $g \in G$, then X is called *G-stable*, and the inclusion map is G-equivariant.

Example 1.2.3. The symmetric group S_4 acts on the set $\binom{[4]}{2}$ of unordered pairs of distinct numbers in $[4] := \{1, 2, 3, 4\}$ by $g\{i, j\} = \{g(i), g(j)\}$. Think of the edges in a tetrahedron to visualise this action. The group S_4 also acts on the set $X := \{(i, j) \mid i, j \in [4] \text{ distinct}\}$ of all ordered pairs by $g(i, j) = (g(i), g(j))$—think of directed edges—and the map $X \to \binom{[4]}{2}$ sending (i, j) to $\{i, j\}$ is S_4-equivariant.

Definition 1.2.4. Let G be a group and let V be a vector space. A *(linear) representation of G on V* is a group homomorphism $\rho : G \to \mathrm{GL}(V)$.

If ρ is a representation of G, then the map $(g, v) \mapsto \rho(g)v$ is an action of G on V. Conversely, if we have an action α of G on V such that $\alpha(g, .) : V \to V$ is a linear map for all $g \in G$, then the map $g \mapsto \alpha(g, .)$ is a linear representation.

As with actions, instead of $\rho(g)v$ we will often write gv. A vector space with an action of G by linear maps is also called a *G-module*.

Given a linear representation $\rho : G \to \mathrm{GL}(V)$, we obtain a linear representation $\rho^* : G \to \mathrm{GL}(V^*)$ on the dual space V^*, called the *dual representation* or *contragredient representation* and defined by

$$(\rho^*(g)x)(v) := x(\rho(g)^{-1}v) \text{ for all } g \in G,\ x \in V^* \text{ and } v \in V. \tag{1.1}$$

Exercise 1.2.5. Let $\rho : G \to \mathrm{GL}_n(\mathbb{C})$ be a representation of G on $\mathbb{C}^n$. Show that with respect to the dual basis, ρ^* is given by $\rho^*(g) = (\rho(g)^{-1})^T$, where A^T denotes the transpose of the matrix A.

1.3 Invariant functions

Definition 1.3.1. Given a representation of a group G on a vector space V, a regular function $f \in \mathcal{O}(V)$ is called *G-invariant* or simply *invariant* if $f(v) = f(gv)$ for all $g \in G, v \in V$. We denote by $\mathcal{O}(V)^G \subseteq \mathcal{O}(V)$ the subalgebra of invariant functions. The actual representation of G is assumed to be clear from the context.

Observe that $f \in \mathcal{O}(V)$ is invariant, precisely when it is constant on the orbits of V under the action of G. In particular, the constant functions are invariant.

The representation of G on V induces an action on the (regular) functions on V by defining $(gf)(v) := f(g^{-1}v)$ for all $g \in G, v \in V$. This way the invariant ring can be discribed as the set of regular functions fixed by the action of G: $\mathcal{O}(V)^G = \{f \in \mathcal{O}(V) \mid gf = f \text{ for all } g \in G\}$. Observe that when restricted to $V^* \subset \mathcal{O}(V)$, this action coincides with the action corresponding to the dual representation. In terms of a basis $x_1, \ldots, x_n$ of V^*, the regular functions are polynomials in the x_i and the action of G is given by $gp(x_1, \ldots, x_n) = p(gx_1, \ldots, gx_n)$ for any polynomial p. Since for every d, G maps the set of polynomials homogeneous of degree d to itself, it follows that the homogeneous parts of an invariant are invariant as well. This shows that $\mathcal{O}(V)^G = \bigoplus_d \mathcal{O}(V)^G_d$, where $\mathcal{O}(V)^G_d := \mathcal{O}(V)_d \cap \mathcal{O}(V)^G$.

Example 1.3.2. Consider the representation $\rho : \mathbb{Z}/3\mathbb{Z} \to \mathrm{GL}_2(\mathbb{C})$ defined by mapping 1 to the matrix $\left(\begin{smallmatrix} 0 & -1 \\ 1 & -1 \end{smallmatrix}\right)$ (and mapping 2 to $\left(\begin{smallmatrix} -1 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and 0 to the identity matrix). With respect to the dual basis x_1, x_2, the dual representation is given by:

$$\rho^*(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \rho^*(1) = \begin{smallmatrix} -1 & -1 \\ 1 & 0 \end{smallmatrix}, \qquad \rho^*(2) = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}. \tag{1.2}$$

The polynomial $f = x_1^2 - x_1x_2 + x_2^2$ is an invariant:

$$\rho^*(1)f = (-x_1 - x_2)^2 - (-x_1 - x_2)(-x_1) + (-x_1)^2 = x_1^2 - x_1x_2 + x_2^2 = f, \tag{1.3}$$

and since 1 is a generator of the group, f is invariant under all elements of the group. Other invariants are $x_1^2x_2 - x_1x_2^2$ and $x_1^3 - 3x_1x_2^2 + x_2^3$. These three invariants generate the ring of invariants, althought it requires some work to show that.

A simpler example in which the complete ring of invariants can be computed is the following.

Example 1.3.3. Let D_4 be the symmetry group of the square, generated by a rotation r, a reflection s and the relations $r^4 = e, s^2 = e$ and $srs = r^3$, where e is the identity. The representation ρ of D_4 on $\mathbb{C}^2$ is given by

$$\rho(r) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \rho(s) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{1.4}$$

the dual representation is given by the same matrices:

$$\rho^*(r) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \rho^*(s) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{1.5}$$

It is easy to check that $x_1^2+x_2^2$ and $x_1^2x_2^2$ are invariants, and so are all polynomial expressions in these two invariants. We will show that in fact $\mathcal{O}(()\mathbb{C}^2)^{D_4} = \mathbb{C}[x_1^2 + x_2^2, x_1^2x_2^2] =: R$. It suffices to show that all homogeneous invariants belong to R.

Let $p \in \mathbb{C}[x_1, x_2]$ be a homogeneous invariant. Since $sp = p$, only monomials having even exponents for x_1 can occur in p. Since r^2s exchanges x_1 and x_2, for every monomial $x_1^a x_2^b$ in p, the monomial $x_1^b x_2^a$ must occur with the same exponent. This proves the claim since every polynomial of the form $x_1^{2n}x_2^{2m} + x_1^{2m}x_2^{2n}$ is an element of R. Indeed, we may assume that $n \leq m$ and proceed by induction on $n + m$, the case $n + m = 0$ being trivial. If $n > 0$ we have $q = (x_1^2x_2^2)^n(x_2^{2m-2n}+x_1^{2m-2n})$ and we are done. If $n = 0$ we have $2q = 2(x_1^{2m} + x_2^{2m}) = 2(x_1^2+x_2^2)^m-\sum_{i=1}^{m-1}\binom{m}{i}(x_1^{2i}x_2^{2m-2i})$ and we are done by induction again.

1.4 Conjugacy classes of matrices

In this section we discuss the polynomial functions on the square matrices, invariant under conjugation of the matrix variable by elements of $\mathrm{GL}_n(\mathbb{C})$. This example shows some tricks that are useful when proving that certain invariants are equal. Denote by $M_n(\mathbb{C})$ the vectorspace of complex $n \times n$ matrices. We consider the action of $G = \mathrm{GL}_n(\mathbb{C})$ on $M_n(\mathbb{C})$ by conjugation: $(g, A) \mapsto gAg^{-1}$ for $g \in \mathrm{GL}_n(\mathbb{C})$ and $A \in M_n(\mathbb{C})$. We are interested in finding all polynomials in the entries of $n \times n$ matrices that are invariant under G. Two invariants are given by the functions $A \mapsto \det A$ and $A \mapsto \mathrm{trace}A$.

Let $\chi_A(t) := \det(tI-A) = t^n - s_1(A)t^{n-1}+s_2(A)t^{n-2}-\cdots+(-1)^n s_n(A)$ be the characteristic polynomial of A. Here the s_i are polynomials in the entries of A. Clearly, $\chi_{gAg^{-1}}(t) = \det(g(tI - A)g^{-1}) = \det(tI - A) = \chi_A(t)$ for all $t \in \mathbb{C}$. It follows that the functions $s_1, \ldots, s_n$ are G-invariant. Observe that $s_1(A) = \mathrm{trace}A$ and $s_n(A) = \det A$.

Proposition 1.4.1. *The functions $s_1, \ldots, s_n$ generate $\mathcal{O}(\mathrm{Mat}_n(\mathbb{C}))^{\mathrm{GL}_n(\mathbb{C})}$ and are algebraically independent.*

Proof. To each $c = (c_1, \ldots, c_n \in \mathbb{C}^n$ we associate the so-called *companion matrix*

$$A_c := \begin{pmatrix} 0 & \cdots & \cdots & 0 & -c_n \\ 1 & \ddots & & \vdots & -c_{n-1} \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & c_2 \\ 0 & \cdots & 0 & 1 & c_1 \end{pmatrix} \in M_n(\mathbb{C}). \tag{1.6}$$

A simple calculation shows that $\chi_{A_c}(t) = t^n + c_{n-1}t^{n-1} + \cdots + c_1 t + c_0$.

Exercise 1.4.2. Verify that $\chi_{A_c}(t) = t^n + c_{n-1}t^{n-1} + \cdots + c_1 t + c_0$.

This implies that $s_i(A_c) = (-1)^i c_i$ and therefore

$$\{(s_1(A_c), s_2(A_c), \ldots, s_n(A_c) \mid A \in M_n(\mathbb{C})\} = \mathbb{C}^n. \tag{1.7}$$

It follows that the s_i are algebraically independent over $\mathbb{C}$. Indeed, suppose that $p(s_1, \ldots, s_n) = 0$ for some polynomial p in n variables. Then $0 = p(s_1, \ldots, s_n)(A) = p(s_1(A), \ldots, s_n(A))$ for all A and hence $p(c_1, \ldots, c_n) = 0$ for all $c \in \mathbb{C}^n$. But this implies that p itself is the zero polynomial.

Now let $f \in \mathcal{O}(\mathrm{Mat}_n(\mathbb{C}))^G$ be an invariant function. Define the polynomial p in n variables by $p(c_1, \ldots, c_n) := f(A_c)$, and $P \in \mathcal{O}(\mathrm{Mat}_n(\mathbb{C}))^G$ by $P(A) := p(-s_1(A), s_2(A), \ldots, (-1)^n s_n(A))$. By definition, P and f agree on all companion matrices, and since they are both G-invariant they agree on $W := \{gA_c g^{-1} \mid g \in G, c \in \mathbb{C}^n\}$. To finish the proof, it suffices to show that W is dense in $\mathrm{Mat}_n(\mathbb{C})$ since $f - P$ is continuous and zero on W. To show that W is dense in $\mathcal{O}(\mathrm{Mat}_n(\mathbb{C}))$, it suffices to show that the set of matrices with n distinct nonzero eigenvalues is a subset of W and is itself dense in $\mathcal{O}(\mathrm{Mat}_n(\mathbb{C}))$. This we leave as an exercise.

Exercise 1.4.3. Let $A \in \mathrm{Mat}_n(\mathbb{C})$ have n distinct nonzero eigenvalues. Show that A is conjugate to A_c for some $c \in \mathbb{C}^n$. Hint: find $v \in \mathbb{C}^n$ such that $v, Av, A^2v, \ldots, A^{n-1}v$ is a basis for $\mathbb{C}^n$. You might want to use the fact that the Vandermonde determinant

$$\det \begin{pmatrix} 1 & \ldots & 1 \\ c_1 & \ldots & c_n \\ c_1^2 & \ldots & c_n^2 \\ \vdots & \ddots & \vdots \\ c_1^{n-1} & \ldots & c_n^{n-1} \end{pmatrix} \tag{1.8}$$

is nonzero if $c_1, \ldots, c_n$ are distinct and nonzero.

Exercise 1.4.4. Show that the set of matrices with n distinct nonzero eigenvalues is dense in the set of all complex $n \times n$ matrices. Hint: every matrix is conjugate to an upper triangular matrix.

□

1.5 Exercises

Exercise 1.5.1. Let G be a finite group acting on $V = \mathbb{C}^n$, $n \geq 1$. Show that $\mathcal{O}(V)^G$ contains a nontrivial invariant. That is, $\mathcal{O}(V)^G \neq \mathbb{C}$. Give an example of an action of an infinite group G on V with the property that only the constant functions are invariant.

Exercise 1.5.2. Let $\rho : \mathbb{Z}/2\mathbb{Z} \to \mathrm{GL}_2(\mathbb{C})$ be the representation given by $\rho(1) := \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Compute the invariant ring. That is, give a minimal set of generators for $\mathcal{O}(\mathbb{C}^2)^{\mathbb{Z}/2\mathbb{Z}}$.

Exercise 1.5.3. Let $U := \{\left(\begin{smallmatrix} 1 & a \\ 0 & 1 \end{smallmatrix}\right) \mid a \in \mathbb{C}\}$ act on $\mathbb{C}^2$ in the obvious way. Denote the coordinate functions by x_1, x_2. Show that $\mathcal{O}(\mathbb{C}^2)^U = \mathbb{C}[x_2]$.

Exercise 1.5.4. Let $\rho : \mathbb{C}^* \to \mathrm{GL}_3(\mathbb{C})$ be the representation given by $\rho(t) = \begin{pmatrix} t^{-2} & 0 & 0 \\ 0 & t^{-3} & 0 \\ 0 & 0 & t^4 \end{pmatrix}$. Find a minimal system of generators for the invariant ring.

Exercise 1.5.5. Let $\pi : \mathrm{Mat}_n(\mathbb{C}) \to \mathbb{C}^n$ be given by $\pi(A) := (s_1(A), \ldots, s_n(A))$. Show that for every $c \in \mathbb{C}^n$ the fiber $\{A \mid \pi(A) = c\}$ contains a unique conjugacy class $\{gAg^{-1} \mid g \in \mathrm{GL}_n(\mathbb{C})\}$ of a diagonalizable (semisimple) matrix A.