EUCLIDEAN DISTANCE DEGREES OF REAL ALGEBRAIC GROUPS

JAN DRAISMA AND JASMIJN A. BAAIJENS

Abstract. We study the problem of finding, in a real algebraic matrix group, the matrix closest to a given data matrix. We do so from the algebro-geometric perspective of Euclidean distance degrees. We recover several classical results; and among the new results that we prove is a formula for the Euclidean distance degree of special linear groups.

1. The distance to a matrix group

Let V be an n-dimensional real vector space equipped with a positive definite inner product $(.|.)$, and write $\mathrm{End}(V)$ for the space of linear maps $V \to V$. The inner product gives rise to a linear map $\mathrm{End}(V) \to \mathrm{End}(V),\ a \mapsto a^t$ called transposition and determined by the property that $(av|w) = (v|a^t w)$ for all $v, w \in V$, and also to a positive definite inner product $\langle ., . \rangle$ on $\mathrm{End}(V)$ defined by $\langle a, b\rangle := \mathrm{tr}(a^t b)$. This inner product enjoys properties such as $\langle a, bc\rangle = \langle b^t a, c\rangle$. The associated norm $||.||$ on $\mathrm{End}(V)$ is called the Frobenius norm. If we choose an orthonormal basis of V and denote the entries of the matrix of $a \in \mathrm{End}(V)$ relative to this basis by a_{ij}, then $||a||^2 = \sum_{ij} a_{ij}^2$. We will use the words matrix and linear maps interchangeably, but we work without choosing coordinates because it allows for a more elegant statement of some of the results. For $a, b \in \mathrm{End}(V)$ and $u, v \in V$ we write $a \perp b$ for $\langle a, b\rangle = 0$, and $v \perp w$ for $(v|w) = 0$.

Let G be a Zariski-closed subgroup of the real algebraic group $\mathrm{GL}(V) \subseteq \mathrm{End}(V)$ of invertible linear maps. In other words, G is a subgroup of $\mathrm{GL}(V)$ characterised by polynomial equations in the matrix entries. Then G is a real algebraic group and in particular a smooth manifold. The problem motivating this note is the following.

Problem 1.1. *Given a general $u \in \mathrm{End}(V)$, determine $x \in G$ that minimises the squared-distance function $d_u(x) := ||u - x||^2$.*

Here, and in the rest of this note, *general* means that whenever convenient, we may assume that u lies outside some proper, Zariski-closed subset of $\mathrm{End}(V)$. Instances of this problem appear naturally in applications. For instance, the *nearest orthogonal matrix* plays a role in computer vision [Hor86], and we revisit its solution in Section 3. More or less equivalent to this is the solution to the orthogonal Procrustes problem [Sch66]. For these and other matrix nearness problems we refer to [Hig89, Kel75]. More recent applications include structured low-rank approximation, for which algebraic techniques are developed in [OSS13].

The bulk of this note is devoted to *counting* the number of critical points on G of the function d_u, in the general framework of the *Euclidean distance degree* (ED

JD is supported by a Vidi grant from the Netherlands Organisation for Scientific Research (NWO).

degree) [DHO$^+$13]. In Section 2 we specialise this framework to matrix groups. In Section 3 we discuss matrix groups preserving the inner product. In particular, we derive a conjecturally sharp upper bound on the ED degree of a compact torus preserving the inner product, revisit the classical cases of orthogonal and unitary groups, and express the ED degree as the algebraic degree of a certain matrix multiplication map. Then in Section 4 we discuss two classes of groups not preserving the inner product: the special linear groups, consisting of all determinant-one matrices, and the symplectic groups. For the former we determine the ED degree explicitly. We conclude the note with a conjecture for the latter.

Acknowledgments

We thank Pierre-Jean Spaenlehauer for his help with computing the ED degree of Sp_6, and Rob Eggermont, Emil Horobeţ, and Hanspeter Kraft for useful suggestions.

2. The ED degree and critical equations

As is common in the framework of ED degree computations, we aim to count the critical points of the function d_u over the complex numbers, as follows. A point $x \in G$ is critical for d_u if and only if $(u - x) \perp a$ for all a in the tangent space $T_x G \subseteq \mathrm{End}(V)$. As G is an algebraic group, we have $T_x G = x \cdot T_1 G = x\mathfrak{g}$, where $T_1 G = \mathfrak{g}$ is the tangent space of G at the identity element 1, i.e., the Lie algebra of G. So criticality means that

$$0 = \langle u - x, xb \rangle = \langle x^t(u - x), b \rangle$$

for all $b \in \mathfrak{g}$. Hence, given u, we look for the solutions of the *critical equations*

$$x^t(u - x) \perp \mathfrak{g} \text{ subject to } x \in G. \tag{1}$$

The number of solutions $x \in G$ to (1) can vary with $u \in \mathrm{End}(V)$. But if we set $V_\mathbb{C} := \mathbb{C} \otimes_\mathbb{R} V$, let $G_\mathbb{C} \subseteq \mathrm{GL}_\mathbb{C}(V_\mathbb{C}) \subseteq \mathrm{End}_\mathbb{C}(V_\mathbb{C})$ be the set of complex points of the algebraic group G, and extend $\langle .,. \rangle$ to a symmetric $\mathbb{C}$-linear form on $\mathrm{End}_\mathbb{C}(V_\mathbb{C})$ (and *not* to a Hermitian form!), then the number of solutions to (1) will not depend on u, provided that u is sufficiently general. Following [DHO$^+$13], we call this number the *Euclidean distance degree* (ED degree for short) of G. This number gives an algebraic measure for the complexity of writing down the solution to the minimisation problem 1.1. We now distinguish two classes of groups: those that preserve the inner product $(.|.)$ and those that do not.

3. Groups preserving the inner product

Assume that $(xv|xw) = (v|w)$ for all $x \in G$, so that G is a subgroup of the orthogonal group of $(.|.)$. Then all elements $x \in G$ satisfy $x^t x = I$ and hence $||x||^2 = n$, that is, G is contained in the sphere in $\mathrm{End}(V)$ of radius $\sqrt{n}$. As a consequence, $\mathfrak{g}$ is contained in the tangent space at 1 to that sphere, which equals $1^\perp$. Hence the critical equations simplify to

$$x^t u \perp \mathfrak{g} \text{ subject to } x \in G. \tag{2}$$

In other words, given the data matrix u we seek to find all $x \in G$ that satisfy a system of linear homogeneous equations. Alternatively, we can write the critical equations as $u \in x \cdot \mathfrak{g}^\perp$. This proves the following proposition.

Proposition 3.1. *If G preserves the inner product $(.|.)$, then the ED degree of G equals the degree of the multiplication map $G_{\mathbb{C}} \times \mathfrak{g}_{\mathbb{C}}^{\perp} \to \mathrm{End}_{\mathbb{C}}(V_{\mathbb{C}}),\ (x,s) \mapsto x \cdot s$. For general real u, among the* real *pairs (x,s) satisfying $xs = u$, the one with the largest value of* $\mathrm{tr}(s)$ *is the one that minimises $d_u(x)$.*

In other words, the ED degree counts the number of ways in which a general matrix u can be decomposed as a product of a matrix in $G_{\mathbb{C}}$ and a matrix in $\mathfrak{g}_{\mathbb{C}}^{\perp}$. The last statement follows from

$$d_u(x) = ||u - x||^2 = \mathrm{tr}(u^t u) - 2\,\mathrm{tr}(u^t x) + n = \mathrm{tr}(u^t u) - 2\,\mathrm{tr}(s) + n,$$

in which only the second term is not constant.

Orthogonal groups. If G is the full orthogonal group of $(.|.)$, then $\mathfrak{g}$ is the space of skew-symmetric matrices. Hence the decomposition of Proposition 3.1 boils down to the classical *polar decomposition*, where one writes a general matrix u as $u = xs$ with x orthogonal and s symmetric. If (x,s) is a solution, then

$$s^2 = s^t s = (x^{-1}u)^t (x^{-1}u) = u^t u,$$

a quadratic equation for s that has 2^n real solutions for general real u. Indeed, write

$$u^t u = y\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n) y^t$$

where y is orthogonal and the λ_i are the eigenvalues of $u^t u$ (which are positive and distinct for general u). Then any of the symmetric matrices

$$s = y\,\mathrm{diag}(\pm\sqrt{\lambda_1}, \ldots, \pm\sqrt{\lambda_n}) y^t$$

is a solution of the quadratic equation, and for each of these the matrix $x = us^{-1}$ is orthogonal, since

$$x^t x = s^{-1}(u^t u)s^{-1} = 1.$$

We summarise our findings in the following, well-known theorem (see, e.g., [Kel75]).

Theorem 3.2. *The ED degree of the orthogonal group of the n-dimensional inner product space V is 2^n. Moreover, for general real u, all 2^n critical points of the squared distance function d_u are real. The critical point that minimises d_u is*

$$x = us^{-1} \text{ with } s = y\,\mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}) y^t,$$

where y is an orthogonal matrix of eigenvectors of $u^t u$, and the λ_i are the corresponding eigenvalues.

The last statement follows since $\mathrm{tr}(s)$ is maximised by taking the *positive* square roots of the λ_i.

Remark 3.3. The closest orthogonal matrix to a general real u has determinant 1 if $\det(u) > 0$ and determinant -1 if $\det(u) < 0$. Half of the critical points of d_u on the orthogonal group have determinant 1, and half of the points have determinant -1, that is, the ED degree of the special orthogonal group is 2^{n-1}. To find the special orthogonal matrix closest to a matrix u with $\det(u) < 0$, one replaces the smallest $\sqrt{\lambda_i}$ in the construction above by $-\sqrt{\lambda_i}$.

Unitary groups. Assume that $n = 2m$ and let V be an m-dimensional complex vector space, regarded as an n-dimensional real vector space. Let h be a non-degenerate, positive definite, Hermitian form on V, where we follow the convention that $h(cv, w) = ch(v, w) = h(v, \overline{c}w)$. Define $(v|w) := \operatorname{Re} h(v, w)$. Then $(.|.)$ is a positive definite inner product on V regarded as a real vector space, and the norm on V coming from $(.|.)$ is the same as that coming from h. Let G be the unitary group of h, which consists of all $x : V \to V$ that are not only $\mathbb{R}$-linear but in fact $\mathbb{C}$-linear and that moreover preserve h. Such maps x also preserve $(.|.)$, so we are in the situation of this section. The converse is also true: if x is $\mathbb{C}$-linear and preserves $(.|.)$, then

$$\begin{aligned}\operatorname{Im} h(u, v) &= -\operatorname{Re} h(iv, w) = -(iv|w) = -(x(iv)|x(w)) = -(ix(v)|x(w)) \\ &= -\operatorname{Re} h(ix(v), x(w)) = \operatorname{Im} h(x(v), x(w)),\end{aligned}$$

so that x preserves both $\operatorname{Re} h$ and $\operatorname{Im} h$ and hence h.

Note that $\operatorname{End}_{\mathbb{R}}(V)$ has dimension $n^2 = 4m^2$, but G is contained in the real subspace $\operatorname{End}_{\mathbb{C}}(V) \subseteq \operatorname{End}_{\mathbb{R}}(V)$, which has real dimension $2m^2$. For a general data matrix $u \in \operatorname{End}_{\mathbb{R}}(V)$, the critical points of d_u on G will be the same as the critical points of $d_{u'}$ where u' is the orthogonal projection of u in $\operatorname{End}_{\mathbb{C}}(V)$. Hence in what follows we may assume that u already lies in $\operatorname{End}_{\mathbb{C}}(V)$, and we focus our attention entirely on the space $\operatorname{End}_{\mathbb{C}}(V)$. For a linear map u in the latter space, we write u^* for the $\mathbb{C}$-linear map determined by $h(v, uw) = h(u^*v, w)$ for all $v, w \in V$. This map will also have the property that $(v|uw) = (u^*v|w)$, i.e., u^* *coincides* with our transpose u^t relative to $(.|.)$. In what follows we follow the convention to write u^*.

The Lie algebra $\mathfrak{g}$ consists of all skew-Hermitian linear maps in $\operatorname{End}_{\mathbb{C}}(V)$, and its orthogonal complement $\mathfrak{g}^{\perp}$ inside $\operatorname{End}_{\mathbb{C}}(V)$ therefore consists of all Hermitian $\mathbb{C}$-linear maps. Again, the decomposition of Proposition 3.1 boils down to the polar decomposition. Here is the result (see, e.g., [Kel75]).

Theorem 3.4. *The ED degree of the unitary group of a non-degenerate Hermitian form $(.|.)$ on an m-dimensional complex vector space V equals 2^m. For a general data point $u \in \operatorname{End}_{\mathbb{C}}(V)$ the critical points are computed as follows. First write*

$$u^*u = y \operatorname{diag}(\lambda_1, \ldots, \lambda_m) y^*,$$

where y is a unitary map and the $\lambda_i \in \mathbb{R}_{\geq 0}$ are the eigenvalues of u^*u, then pick any of the 2^m square roots

$$s = y \operatorname{diag}(\pm\sqrt{\lambda_1}, \ldots, \pm\sqrt{\lambda_m}) y^*$$

of u^*u, and finally set $x := us^{-1}$. Choosing all square roots positive leads to the closest unitary matrix to u.

There is a slight subtlety in the last statement: to find the closest matrix, we have to maximise $\operatorname{tr}_{\mathbb{R}}(s)$, where we see s as an element of $\operatorname{End}_{\mathbb{R}}(V)$, while the sum of the m eigenvalues $\pm\sqrt{\lambda_i}$ equals $\operatorname{tr}_{\mathbb{C}}(s)$. But in fact, for any $z \in \operatorname{End}_{\mathbb{C}}(V)$, we have $\operatorname{tr}_{\mathbb{R}}(z) = \operatorname{tr}_{\mathbb{C}}(z) + \overline{\operatorname{tr}_{\mathbb{C}}(z)}$, so that $\operatorname{tr}_{\mathbb{R}}(s) = 2\operatorname{tr}_{\mathbb{C}}(s)$.

Compact tori. Assume that the real algebraic group G is a compact torus, i.e., that it is an abelian compact Lie group and abstractly isomorphic to a power $(S^1)^m$ of circle groups. We continue to assume that G preserves the inner product $(.|.)$. We will bound the ED degree of G, and unlike in the previous two examples we will make extensive use of the complexification $G_{\mathbb{C}}$ of G.

Indeed, $G_{\mathbb{C}}$ is now isomorphic to an algebraic torus $T := (\mathbb{C}^*)^m$. Let $X(T)$ be the set of all characters of T, i.e., of all algebraic group homomorphisms $\chi : T \to \mathbb{C}^*$. These are all of the form $\chi : (t_1, \ldots, t_m) \mapsto t_1^{a_1} \cdots t_m^{a_m}$, where $a_1, \ldots, a_m \in \mathbb{Z}$; and this gives an isomorphism $X(T) \cong \mathbb{Z}^m$ of finitely generated Abelian groups: $X(T)$ with respect to multiplication and $\mathbb{Z}^m$ with respect to addition. We will identify these groups, and accordingly write t^χ instead of $\chi(t)$ and write $+$ for the operation in $X(T)$, so $t^{\chi+\lambda} = t^\chi \cdot t^\lambda$ and $t^\chi = t_1^{a_1} \cdots t_m^{a_m}$ if $\chi = (a_1, \ldots, a_m)$.

The isomorphism $T \mapsto G_{\mathbb{C}} \subseteq \mathrm{GL}_{\mathbb{C}}(V_{\mathbb{C}})$ gives $V_{\mathbb{C}}$ the structure of a T-representation. As such, it splits as a direct sum of one-dimensional T-representations:

$$V_{\mathbb{C}} = \bigoplus_{\chi \in X(T)} V_\chi,$$

where, for any $\chi \in X(T)$, we let V_χ be the corresponding eigenspace (or *weight space*), defined by

$$V_\chi := \{v \in V_{\mathbb{C}} \mid \forall t \in T : tv = t^\chi v\}.$$

Of course, only finitely many of these spaces are non-zero, and their dimensions add up to n. Let $X_V \subseteq X(T) = \mathbb{Z}^m$ denote the set of characters χ for which V_χ is non-zero. The fact that the map $T \to G$ is an isomorphism implies that the lattice in $\mathbb{Z}^m$ generated by X_V has full rank. We will prove the following result.

Theorem 3.5. *The ED degree of the compact torus $G \cong (S^1)^m$ depends only on X_V, and is independent of the dimensions of the weight spaces $\dim V_\chi$, $\chi \in X_V$. Moreover, it is bounded from above by the normalised volume of the convex hull Δ of $X_V \subseteq \mathbb{Z}^m$. Here the normalisation is such that the simplex spanned by 0 and the standard basis vectors has volume one.*

For the proof, observe that the (complexified) bilinear form $(.|.)$ on V is preserved by T. For $v \in V_\chi$ and $w \in V_\lambda$ and $t \in T$ we therefore have

$$t^{\chi+\lambda} \cdot (v|w) = (t^\chi v | t^\lambda w) = (tv|tw) = (v|w).$$

Hence, if v and w are not perpendicular, then $\chi + \lambda$ is the trivial character sending all of T to 1 (so $\chi + \lambda = 0 \in \mathbb{Z}^m$). In other words, $(.|.)$ must pair each V_χ non-degenerately with the corresponding space $V_{-\chi}$, and is zero on all pairs $V_\chi \times V_\lambda$ with $\lambda \neq -\chi$. In particular, we have $\dim V_\chi = \dim V_{-\chi}$ for all $\chi \in X(T)$, and X_V is centrally symmetric.

We can now choose a basis $v_1, \ldots, v_n$ of $V_{\mathbb{C}}$ consisting of T-eigenvectors such that $(v_i|v_j) = \delta_{j,n+1-i}$. Let $\chi_i \in X_V$ be the character of v_i, i.e., we have $tv_i = t^{\chi_i} v_i$ for all $t \in T$. There will be repetitions among the χ_i if some of the weight spaces have dimensions greater than 1; and by the above we have $\chi_i = -\chi_{n+1-i}$. Relative to this basis we have

$$G_{\mathbb{C}} = \{(t^{\chi_1}, \ldots, t^{\chi_n}) \mid t \in (\mathbb{C}^*)^m\} \text{ and } \mathfrak{g}_{\mathbb{C}} = \{\mathrm{diag}(w \cdot \chi_1, \ldots, w \cdot \chi_n) \mid w \in \mathbb{C}^m\}$$

where $\chi_i \cdot w$ is the ordinary dot product of $w \in \mathbb{C}^m$ with $\chi_i \in \mathbb{Z}^m$. The bilinear form $\langle ., . \rangle$ on matrices takes the form

$$\langle E_{ij}, E_{kl} \rangle = \delta_{k,n+1-i}\delta_{l,n+1-j},$$

where E_{ij} is the map whose matrix relative to the basis $v_1, \ldots, v_n$ has a 1 at position (i, j) and zeroes elsewhere. Hence, for a data matrix $u \in \mathrm{End}(V)$, the critical equations (2) translate into the following equations for the pre-image $t \in T$ of $x \in G_{\mathbb{C}}$:

$$(t^{-\chi_1} u_{11})(w \cdot \chi_n) + \ldots + (t^{-\chi_n} u_{nn})(w \cdot \chi_1) = 0 \text{ for all } w \in \mathbb{C}^m.$$

Using $\chi_i = -\chi_{n+1-i}$ we may rewrite this as

$$\sum_{i=1}^{n} (t^{\chi_i} u_{ii})(w \cdot \chi_i) = 0 \text{ for all } w \in \mathbb{C}^m \tag{3}$$

The ED degree of G is the number of solutions $t \in (\mathbb{C}^*)^m$ to this system of m Laurent-polynomial equations for general values of the u_{ii}. Grouping the indices i for which the χ_i are equal and adding up the corresponding u_{ii} we find that the cardinality of the solution set is, indeed, independent of the dimensions $\dim V_{\chi_i}$, the first statement in the theorem.

Letting w run over a basis of $\mathbb{C}^m$, we obtain a system of m Laurent polynomial equations for $t \in (\mathbb{C}^*)^m$ with fixed support X_V. The Bernstein-Kushnirenko-Khovanskii theorem [Ber76, Theorems A,B] ensures that the number of isolated solutions to this system is less than or equal to the normalised volume of Δ. This proves that the ED degree does not exceed that bound, and hence the second part of the theorem.

Example 3.6. Let the group $\mathrm{SO}_2 \cong S^1$ act on $V := (\mathbb{R}^2)^{\otimes d}$ via $g(v_1 \otimes \cdots \otimes v_d) = (gv_1) \otimes \cdots \otimes (gv_d)$, and let $(.|.)$ be the inner product on V induced from that on $\mathbb{R}^2$. Let $G \subseteq \mathrm{GL}(V)$ be the image of SO_2. Then the ED degree of G equals $2d$, computed as follows. All elements of SO_2 are diagonalised over $\mathbb{C}$ by the choice of basis $f_1 := e_1 + ie_2$ and $f_{-1} := e_1 - ie_2$ of $\mathbb{C}^2$. The complexification of SO_2 is the image of the one-dimensional torus $T = \mathbb{C}^*$ in its action on $\mathbb{C}^2$ via $(t, f_i) \mapsto t^i f_i$. The complexification $G_{\mathbb{C}}$ is the image of T in its induced action on $\mathbb{C}^{\otimes d}$ via

$$(t, f_{i_1} \otimes \cdots \otimes f_{i_d}) \mapsto t^{\sum_j i_j} f_{i_1} \otimes \cdots \otimes f_{i_d} \text{ for all } i_1, \ldots, i_d \in \{\pm 1\}.$$

We have $X_V = \{-d, -d+2, \ldots, d\} \subseteq \mathbb{Z}^1$, where $\mathbb{Z}^1$ is the character lattice of T. If d is odd, then the map $T \to G_{\mathbb{C}}$ is one-to-one, and Δ is a line segment of length $2d$. If d is even, then that map is two-to-one (since all exponents $\sum_j i_j$ above are then even), and the character lattice of $G_{\mathbb{C}}$ is then $2\mathbb{Z}^1 \subseteq \mathbb{Z}^1$. In this case, the normalised volume of Δ is d. The theorem says that the ED degree is at most $2d$ for odd d and at most d for even d. In this case equality holds: the system (3) reduces to

$$u'_{-d} t^{-d} + \cdots + u'_d t^d = 0,$$

where the u'_j are sums of u_{ii} corresponding to the same character; and for general values of the u'_j this equation has exactly $2d$ solutions for t, and exactly d solutions for t^2 when d is even.

We do not know if the system (3), for general choices of the u_{ii} is always sufficiently general for the BKK-bound to hold with equality.

Problem 3.7. *Is the ED degree of a torus $(S^1)^m \cong G \subseteq \mathrm{GL}(V)$ always equal to the normalised volume of the convex hull of the character set $X_V \subseteq \mathbb{Z}^m$ appearing in $V_{\mathbb{C}}$?*

Other reductive groups preserving the form. After this more or less satisfactory result for tori, it is tempting to hope that the ED degree of any group G preserving the bilinear form should be expressible in terms of the highest weights appearing in the complexification $V_{\mathbb{C}}$ as a $G_{\mathbb{C}}$-module. After all, $G_{\mathbb{C}}$ is then a reductive group and much is known about its representations. By Proposition 3.1 an upper bound to the ED degree is the degree of a general orbit of $G_{\mathbb{C}}$ in its action on $\mathrm{End}_{\mathbb{C}}(V_{\mathbb{C}})$ by left multiplication. A formula for this degree is known by [Kaz87]; see also [DK95, Theorem 8]. However, the space $\mathfrak{g}^{\perp}$ is in general not sufficiently general for that upper bound to be tight.

To test said hope, we have experimented with $G_{\mathbb{C}}$ equal to the image of $\mathrm{SL}_2(\mathbb{C})$ in its irreducible representation $V_{\mathbb{C}}$ of highest weight m with m even. Thus $V_{\mathbb{C}}$ is the m-th symmetric power $S^m\mathbb{C}^2$ where $\mathbb{C}^2$ is the standard representation of $G_{\mathbb{C}}$. Since m is even, the bilinear form on $S^m\mathbb{C}^2$ induced by the symplectic $\mathrm{SL}_2(\mathbb{C})$-invariant form on $\mathbb{C}^2$ is, indeed, an invariant symmetric bilinear form. In this case, the formula in [DK95, Theorem 8] evaluates to m^3. Below is a small table of ED degrees, in which we could not yet find a pattern.

m	0	2	4	6	8
ED-degree of SL_2 on $S^m\mathbb{C}^2$	1	4	40	156	400
m^3	0	8	64	216	512

Note that the formula m^3 does not apply for $m = 0$ since the representation does not have a finite kernel. Note also that the 4 is consistent with Theorem 3.2 and the fact that the map $\mathrm{SL}_2(\mathbb{C}) \to \mathrm{SL}(S^2\mathbb{C}^2)$ has image $G_{\mathbb{C}}$ equal to $\mathrm{SO}_3(\mathbb{C})$.

Problem 3.8. *Determine the ED-degree of* $\mathrm{SL}_2(\mathbb{C})$ *on* $S^m\mathbb{C}^2$ *with* m *even. More generally, find a formula for that ED degree for any group* $G_{\mathbb{C}}$ *on a representation* $V_{\mathbb{C}}$ *with an invariant symmetric bilinear form.*

4. Groups not preserving the inner product

If G does not preserve the inner product $(.|.)$, then it is much harder to compute (or even estimate) the ED degree of G. The following two classes of groups illustrate this.

The special linear groups. Consider the group

$$G := \mathrm{SL}^{\pm}(V) = \{x \in \mathrm{End}(V) \mid \det x = \pm 1\}.$$

Its Lie algebra $\mathfrak{g}$ is the space of matrices with trace equal to zero. Given a real data matrix u, the critical equations for the nearest $x \in G$ become

$$\langle x^t(u - x), a\rangle = 0 \quad \forall a \in \mathrm{End}(V), \mathrm{tr}(a) = 0 \text{ subject to } \det x = 1.$$

Since this equation must hold for all a with $\mathrm{tr}(a) = 0$ we see that $x^t(u - x)$ must be of the form cI for some $c \in \mathbb{R}$, so that $u = cx^{-t} + x$. From this expression for u we find

$$u^t u = (cx^{-1} + x^t)(cx^{-t} + x) = c^2(x^t x)^{-1} + 2cI + x^t x.$$

Hence $s := x^t x$ must be a symmetric matrix of determinant 1 satisfying

$$u^t u = c^2 s^{-1} + 2cI + s = s^{-1}(cI + s)^2 \tag{4}$$

Conversely, if s is a symmetric determinant-1 matrix satisfying this equation, then we can set $x := u^{-t}(cI + s)$. This matrix then satisfies

$$x^t x = (cI + s)u^{-1}u^{-t}(cI + s) = (cI + s)^2(u^t u)^{-1} = s,$$

where we have used that s commutes with $u^t u$. This means that $\det x = \pm 1$. We also find $x^t(u-x) = (cI+s) - s = cI$. Thus to compute the ED degree of G it suffices to count the symmetric matrices s solving (4).

Let $\mu_1, \ldots, \mu_n$ be the eigenvalues of $u^t u$. Since u is general, these are all distinct, and (4) forces s to be simultaneously diagonalisable with $u^t u$. Thus we need only find the eigenvalues $\lambda_1, \ldots, \lambda_n$ of s, where λ_i corresponds to μ_i. The equation (4) translates into

$$\mu_i = c^2 \frac{1}{\lambda_i} + 2c + \lambda_i, \qquad i = 1, \ldots, n.$$

Multiplying by λ_i and adding the condition $\lambda_1 \cdots \lambda_n = 1$, the system to solve becomes

$$\begin{cases} f_i := c^2 + (2c - \mu_i)\lambda_i + \lambda_i^2 & = 0, \qquad i = 1, \ldots, n \\ \lambda_1 \cdots \lambda_n & = 1. \end{cases} \tag{5}$$

Substituting $\lambda_n := (\lambda_1 \cdots \lambda_{n-1})^{-1}$ into f_n, we find an ideal I generated by n equations $f_1, \ldots, f_n$ in the ring $\mathbb{R}[\lambda_1^{\pm 1}, \ldots, \lambda_{n-1}^{\pm 1}, \mu_1, \ldots, \mu_n, c]$. This ideal is prime, since the equations can be read as defining the graph of a map from the Cartesian product of an $(n-1)$-dimensional torus with coordinates $\lambda_1, \ldots, \lambda_{n-1}$ with the affine line with coordinate c to the affine space with coordinates $\mu_1, \ldots, \mu_n$. The ED degree is the degree of this map. To determine it, we determine the intersection $I \cap \mathbb{R}[\mu_1, \ldots, \mu_n, c]$. For this, we eliminate the λ_i successively, as follows. For $i = 0, \ldots, n-1$ define $\lambda_{(i)} := \lambda_1 \cdots \lambda_i$. Define

$$R_n := c^2 \lambda_{(n-1)}^2 + (2c - \mu_n)\lambda_{(n-1)} + 1,$$

which is just f_n multiplied by $\lambda_{(n-1)}^2$. Now recursively define, for $i = 1, \ldots, n-1$,

$$R_i := \mathrm{Res}_{\lambda_i}(R_{i+1}, f_i),$$

where Res is the *resultant* given by the determinant of a suitable Sylvester matrix. The first two are as follows:

$$\begin{aligned} R_{n-1} &= \det \begin{bmatrix} c^2\lambda_{(n-2)}^2 & (2c-\mu_n)\lambda_{(n-2)} & 1 & 0 \\ 0 & c^2\lambda_{(n-2)}^2 & (2c-\mu_n)\lambda_{(n-2)} & 1 \\ 1 & (2c-\mu_{n-1}) & c^2 & 0 \\ 0 & 1 & (2c-\mu_{n-1}) & c^2 \end{bmatrix} \\ &= c^8\lambda_{(n-2)}^4 + \cdots + 1, \end{aligned}$$

where the dots stand for terms of degrees strictly between 0 and 4 in $\lambda_{(n-2)}$; and similarly

$$\begin{aligned} R_{n-2} &= \det \begin{bmatrix} c^8\lambda_{(n-3)}^4 & . & . & . & 1 & 0 \\ 0 & c^8\lambda_{(n-3)}^4 & . & . & . & 1 \\ 1 & (2c-\mu_{n-2}) & c^2 & 0 & 0 & 0 \\ 0 & 1 & (2c-\mu_{n-2}) & c^2 & 0 & 0 \\ 0 & 0 & 1 & (2c-\mu_{n-2}) & c^2 & 0 \\ 0 & 0 & 0 & 1 & (2c-\mu_{n-2}) & c^2 \end{bmatrix} \\ &= c^{24}\lambda_{(n-3)}^8 + \cdots + 1. \end{aligned}$$

By induction, we find $R_i = c^{m_i}\lambda_{(i-1)}^{2^{n-i+1}} + \cdots + 1$ where the remaining terms have $\lambda_{(i-1)}$-degree strictly between zero and 2^{n-i+1}. The exponents m_i satisfy the recursion

$$m_i = 2m_{i+1} + 2 \cdot 2^{n-i}$$

and $m_n = 2$. This is solved by $m_i = (n-i+1)2^{n-i+1}$. In particular, we find that $m_1 = n2^n$. By induction one can prove that $f_1, \ldots, f_{i-1}, R_i$ generate the intersection

$$I_i := I \cap \mathbb{R}[\lambda_1^{\pm 1}, \ldots, \lambda_{i-1}^{\pm 1}, \mu_1, \ldots, \mu_n, c]$$

and that, modulo I_i, the variable λ_{i-1} can be expressed as a $\mathbb{Q}$-rational function of $\lambda_1, \ldots, \lambda_{i-2}, \mu_1, \ldots, \mu_n, c$. This can be used to show that for generic choices of the μ_i the degree-$n2^n$ equation R_1 in c lifts to as many distinct solutions to the system (5). Thus we have proved the following theorem.

Theorem 4.1. *The ED degree of* $\mathrm{SL}^{\pm}(V)$ *equals* $n2^n$*, and the ED degree of* $\mathrm{SL}(V)$ *equals* $n2^{n-1}$*.*

The last statement follows from the fact that there exists an orthogonal transformation of $\mathrm{End}(V)$ that takes the matrices with determinant 1 into the matrices with determinant -1 and vice versa (e.g., in matrix terms, multiplying the first column by -1). Hence the two connected components of $\mathrm{SL}^{\pm}(V)$ have the same ED degree.

The proof gives rise to the following algorithm for finding the closest matrix in G to a given real data matrix u: first diagonalise $u^t u$ as

$$u^t u = T \operatorname{diag}(\mu_1, \ldots, \mu_n) T^t,$$

where T is a real orthogonal transformation and the μ_i are positive. Then successively eliminate $\lambda_n, \ldots, \lambda_1$ as above, using Sylvester matrices for the resultants R_i. Compute all real roots c of R_1. For each of these, compute the corresponding $\lambda_1, \ldots, \lambda_n$ from the kernels of the Sylvester matrices: since the data is sufficiently general, each of those kernels will be one-dimensional and spanned by a vector of powers of the relevant λ_i. Since all λ_i are $\mathbb{Q}$-rational functions of $\mu_1, \ldots, \mu_n, c$, the λ_i are, indeed, real. Then construct s by

$$s = T \operatorname{diag}(\lambda_1, \ldots, \lambda_n) T^t.$$

Finally, construct x by

$$x = u^{-t}(cI + s).$$

We have already verified that x satisfies $x^t x = s$, so that the λ_i are necessarily positive, but this can also be seen directly from (5).

It would be useful to know in advance which real root c corresponds to the closest matrix x. Experiments with the algorithm above suggests that it may be the real root that is smallest in absolute value.

Problem 4.2. *Is it true that the real root* c *of* R_1 *of smallest absolute value gives rise to the matrix* $x \in \mathrm{SL}^{\pm}(V)$ *that is closest to* u*?*

5. The symplectic groups

As a final case in our quest for ED degree of real algebraic groups we fix an even $n = 2m \in \mathbb{N}$ and study the symplectic group

$$\mathrm{Sp}_n := \{x \in \mathbb{R}^{n \times n} \mid x^t J x = J\},$$

where J has the block structure

$$J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

In the case of $\mathrm{SL}(V)$ the ED degree did not depend on the choice of an inner product on V, because $\mathrm{SL}(V)$ acts transitively on inner products (up to positive scalars). But for Sp_n the ED degree may well depend on the relative position of the symplectic form given by J and the inner product. A general study of orbits of pairs of a symplectic and a symmetric form is performed in [LR05, Die46], based on classical work by Kronecker. We choose the standard inner product. This choice is rather special in the sense that the complexified group $\mathrm{Sp}_n(\mathbb{C})$ intersects the complexified group $\mathrm{O}_n(\mathbb{C})$ in a large group, containing a copy of the group $\mathrm{GL}_m(\mathbb{C})$. This is not immediately clear from the chosen coordinates, but relative to the basis

$$v_1 := \frac{e_1 + ie_{m+1}}{\sqrt{2}}, \ldots, v_m := \frac{e_m + ie_{2m}}{\sqrt{2}}, v_{m+1} := \frac{ie_1 + e_{m+1}}{\sqrt{2}}, \ldots, v_{2m} := \frac{ie_m + e_m}{\sqrt{2}}$$

of $\mathbb{C}^n$ the symplectic form still has Gram matrix J, while the standard symmetric bilinear form on $\mathbb{C}^n$ has Gram matrix

$$\begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}. \tag{6}$$

Now all complex matrices that relative to the basis of the v_i have the block structure

$$\begin{bmatrix} g & 0 \\ 0 & g^{-T} \end{bmatrix}$$

lie both in $\mathrm{Sp}_n(\mathbb{C})$ and in $\mathrm{O}_n(\mathbb{C})$. This shows that we could have chosen the symmetric form with Gram matrix (a scalar multiple of) that in (6), without changing the ED degree.

We have implemented the equations 1 and computed the ED degree for very small values of n. The resulting table is as follows:

n	2	4	6
ED degree of Sp_n	4	24	544.

The pattern might be that the answer is $2^{m^2} + 2^{2m-1}$, but we do not know how to prove this.

Problem 5.1. *Determine the ED degree of* Sp_n *for general even* n.

References

[Ber76] D.N. Bernstein. The number of roots of a system of equations. *Funct. Anal. Appl. 9*, pages 183–185, 1976.

[DHO+13] Jan Draisma, Emil Horobet, Giorgio Ottaviani, Bernd Sturmfels, and Rekha R. Thomas. The euclidean distance degree of an algebraic variety. 2013. Preprint available from `http://arxiv.org/abs/1309.0049`.

[Die46] Jean Dieudonné. Sur la réduction canonique des couples de matrices. *Bull. Soc. Math. France*, 74:130–146, 1946.

[DK95] Harm Derksen and Hanspeter Kraft. Constructive invariant theory. In *Algèbre non commutative, groupes quantiques et invariants. Septième contact Franco-Belge, Reims, France, June 26–30, 1995*, pages 221–244. Paris: Société Mathématique de France, 1995.

[Hig89] N.J. Higham. Matrix nearness problems and applications. In *Applications of matrix theory, Proc. Conf., Bradford/UK 1988*, volume 22 of *Inst. Math. Appl. Conf. Ser., New. Ser.*, pages 1–27, 1989.

[Hor86] B.K.P. Horn. *Robot Vision*. MIT Press and McGraw-Hill, 1986.

[Kaz87] B. Ya. Kazarnovskiĭ. Newton polyhedra and Bezout's formula for matrix functions of finite-dimensional representations. *Funktsional. Anal. i Prilozhen.*, 21(4):73–74, 1987.

[Kel75] Joseph B. Keller. Closest unitary, orthogonal and hermitian operators to a given operator. *Math. Mag.*, 48:192–197, 1975.

[LR05] P. Lancaster and L. Rodman. Canonical forms for symmetric/skew-symmetric real matrix pairs under strict equivalence and congruence. *Linear Algebra Appl.*, 406:1–76, 2005.

[OSS13] Giorgio Ottaviani, Pierre-Jean Spaenlehauer, and Bernd Sturmfels. Algebraic methods for structured low-rank approximation. 2013. Preprint, available from `http://arxiv.org/abs/1311.2376`.

[Sch66] P.H. Schonemann. A generalized solution of the orthogonal procrustes problem. *Psychometrika*, 31:1–10, 1966.

(Jan Draisma) Department of Mathematics and Computer Science, Technische Universiteit Eindhoven, P.O. Box 513, 5600 MB Eindhoven, The Netherlands; and Centrum voor Wiskunde en Informatica, Amsterdam, The Netherlands

E-mail address: `j.draisma@tue.nl`

(Jasmijn A. Baaijens) Department of Mathematics and Computer Science, Technische Universiteit Eindhoven, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

E-mail address: `j.a.baaijens@student.tue.nl`