Geometry of orthogonally invariant matrix varieties

Dmitriy Drusvyatskiy
Mathematics, University of Washington

Joint work with
H.-L. Lee (UW), G. Ottaviano (Florence), and R.R. Thomas (UW)

UC Davis
Algebra & Discrete Mathematics

AFOSR YIP

"The Euclidean distance degree of orthogonally invariant matrix varieties", D-Lee-Ottaviani-Thomas,
To appear in **Israel J. Math.**, 2016.

"Counting real critical points of the distance to orthogonally invariant matrix sets", D-Lee-Thomas,
SIAM J. Matrix Anal. Applic. 36(3):1360-1380, 2015.

Absolute symmetry

Signed permutation group:

$$\Pi^n_\pm = \{\text{signed permutations}\}$$

Subset $S \subset \mathbf{R}^n$ is $\Pi^n_\pm$-invariant if

$$\pi S = S \qquad \text{for all} \quad \pi \in \Pi^n_\pm.$$

Absolute symmetry

Signed permutation group:

$$\Pi_{\pm}^{n} = \{\text{signed permutations}\}$$

Subset $S \subset \mathbf{R}^n$ is $\Pi_{\pm}^{n}$-invariant if

$$\pi S = S \qquad \text{for all} \quad \pi \in \Pi_{\pm}^{n}.$$

Real orthogonal group:

$$O^n = \{n \times n \text{ orthogonal matrices}\}$$

Subset $\mathcal{M} \subset \mathbf{R}^{n \times n}$ is O^n-invariant if

$$U\mathcal{M} = \mathcal{M}U = \mathcal{M} \qquad \text{for all} \quad U \in O^n$$

Universal examples

Singular values decomp.: any $X \in \mathbf{R}^{n\times n}$ can be written as

$$X = U \begin{bmatrix} \sigma_1(X) & & & \\ & \sigma_2(X) & & \\ & & \ddots & \\ & & & \sigma_n(X) \end{bmatrix} V^T$$

with $U, V \in O^n$ and $\sigma_1(X) \geq \sigma_2(x) \geq \ldots \geq \sigma_n(X) \geq 0$.

Universal examples

Singular values decomp.: any $X \in \mathbf{R}^{n\times n}$ can be written as

$$X = U \begin{bmatrix} \sigma_1(X) & & & \\ & \sigma_2(X) & & \\ & & \ddots & \\ & & & \sigma_n(X) \end{bmatrix} V^T$$

with $U, V \in O^n$ and $\sigma_1(X) \geq \sigma_2(x) \geq \ldots \geq \sigma_n(X) \geq 0$.

$$\boxed{\text{Singular values } \sigma_i(X) \quad = \quad \sqrt{\text{eigenvalues of } X^T X}}$$

Universal examples

Singular values decomp.: any $X \in \mathbf{R}^{n\times n}$ can be written as

$$X = U \begin{bmatrix} \sigma_1(X) & & & \\ & \sigma_2(X) & & \\ & & \ddots & \\ & & & \sigma_n(X) \end{bmatrix} V^T$$

with $U, V \in O^n$ and $\sigma_1(X) \geq \sigma_2(x) \geq \ldots \geq \sigma_n(X) \geq 0$.

Singular values $\sigma_i(X) \quad = \quad \sqrt{\text{eigenvalues of } X^T X}$

Examples of O^n-invariant sets:

$$\mathbf{R}_r^{n\times n} = \{X : \operatorname{rank} X \leq r\} \qquad \mathbf{B}_* = \{X : \sum_i \sigma_i(X) \leq 1\}$$

$$\mathcal{E}^{3\times 3} = \{X \in \mathbf{R}^{3\times 3} : \sigma_1(X) = \sigma_2(X), \sigma_3(X) = 0\}$$

Universal examples

Singular values decomp.: any $X \in \mathbf{R}^{n\times n}$ can be written as

$$X = U \begin{bmatrix} \sigma_1(X) & & & \\ & \sigma_2(X) & & \\ & & \ddots & \\ & & & \sigma_n(X) \end{bmatrix} V^T$$

with $U, V \in O^n$ and $\sigma_1(X) \geq \sigma_2(x) \geq \ldots \geq \sigma_n(X) \geq 0$.

$$\boxed{\text{Singular values } \sigma_i(X) \quad = \quad \sqrt{\text{eigenvalues of } X^T X}}$$

Examples of O^n-invariant sets:

$$\mathbf{R}_r^{n\times n} = \{X : \operatorname{rank} X \leq r\} \qquad \mathbf{B}_* = \{X : \sum_i \sigma_i(X) \leq 1\}$$

$$\mathcal{E}^{3\times 3} = \{X \in \mathbf{R}^{3\times 3} : \sigma_1(X) = \sigma_2(X), \sigma_3(X) = 0\}$$

Elementary fact:

$\mathcal{M}$ is O^n-invariant $\iff$ $\mathcal{M} = \sigma^{-1}(S)$ for $\Pi^n_{\pm}$-invariant S

Universal examples

Singular values decomp.: any $X \in \mathbf{R}^{n\times n}$ can be written as

$$X = U \begin{bmatrix} \sigma_1(X) & & & \\ & \sigma_2(X) & & \\ & & \ddots & \\ & & & \sigma_n(X) \end{bmatrix} V^T$$

with $U, V \in O^n$ and $\sigma_1(X) \geq \sigma_2(x) \geq \ldots \geq \sigma_n(X) \geq 0$.

$$\boxed{\text{Singular values } \sigma_i(X) \quad = \quad \sqrt{\text{eigenvalues of } X^T X}}$$

Examples of O^n-invariant sets:

$$\mathbf{R}_r^{n\times n} = \{X : \operatorname{rank} X \leq r\} \qquad \mathbf{B}_* = \{X : \sum_i \sigma_i(X) \leq 1\}$$

$$\mathcal{E}^{3\times 3} = \{X \in \mathbf{R}^{3\times 3} : \sigma_1(X) = \sigma_2(X), \sigma_3(X) = 0\}$$

Elementary fact:

$\mathcal{M}$ is O^n-invariant $\iff$ $\mathcal{M} = \sigma^{-1}(S)$ for $\Pi^n_\pm$-invariant S

Concisely $S = \{x \in \mathbf{R}^n : \operatorname{Diag}(x) \in \mathcal{M}\}$ is diagonal restriction

Guiding philosophy

Transfer Principle:

“Properties” of $\sigma^{-1}(S)$ and S are in one-to-one correspondence.

Guiding philosophy

Transfer Principle:

"Properties" of $\sigma^{-1}(S)$ and S are in one-to-one correspondence.

Examples:

- $\sigma^{-1}(S)$ convex $\iff$ S convex
  (von Neumann '37, Davis '57, Lewis '96)

Guiding philosophy

Transfer Principle:

"Properties" of $\sigma^{-1}(S)$ and S are in one-to-one correspondence.

Examples:

- $\sigma^{-1}(S)$ convex $\iff$ S convex
  (von Neumann '37, Davis '57, Lewis '96)
  Reason: $\max_{X \in \sigma^{-1}(S)} \langle V, X \rangle = \max_{x \in S} \langle \sigma(V), x \rangle$

Guiding philosophy

Transfer Principle:

"Properties" of $\sigma^{-1}(S)$ and S are in one-to-one correspondence.

Examples:

- $\sigma^{-1}(S)$ convex $\iff$ S convex
  (von Neumann '37, Davis '57, Lewis '96)
  Reason: $\max_{X \in \sigma^{-1}(S)} \langle V, X \rangle = \max_{x \in S} \langle \sigma(V), x \rangle$

Here $\langle V, X \rangle = \mathrm{tr}\,(V^T X)$ is the trace product.

Guiding philosophy

Transfer Principle:

"Properties" of $\sigma^{-1}(S)$ and S are in one-to-one correspondence.

Examples:

- $\sigma^{-1}(S)$ convex $\iff$ S convex
  (von Neumann '37, Davis '57, Lewis '96)
  Reason: $\max_{X \in \sigma^{-1}(S)} \langle V, X \rangle = \max_{x \in S} \langle \sigma(V), x \rangle$

Here $\langle V, X \rangle = \mathrm{tr}\,(V^T X)$ is the trace product.

- $\sigma^{-1}(S)$ is C^p-smooth $\iff$ S is C^p-smooth
  (Sylvester '85, Šilhavý '99, Daniilidis-D-Lewis '14)

Guiding philosophy

Transfer Principle:

"Properties" of $\sigma^{-1}(S)$ and S are in one-to-one correspondence.

Examples:

- $\sigma^{-1}(S)$ convex $\iff$ S convex
  (von Neumann '37, Davis '57, Lewis '96)
  Reason: $\max_{X \in \sigma^{-1}(S)} \langle V, X \rangle = \max_{x \in S} \langle \sigma(V), x \rangle$

Here $\langle V, X \rangle = \operatorname{tr}(V^T X)$ is the trace product.

- $\sigma^{-1}(S)$ is C^p-smooth $\iff$ S is C^p-smooth
  (Sylvester '85, Šilhavý '99, Daniilidis-D-Lewis '14)
  Reason: $\operatorname{dist}_{\sigma^{-1}(S)}(Y) = \operatorname{dist}_S(\sigma(Y))$

Guiding philosophy

Transfer Principle:

"Properties" of $\sigma^{-1}(S)$ and S are in one-to-one correspondence.

Examples:

- $\sigma^{-1}(S)$ convex $\iff$ S convex
  (von Neumann '37, Davis '57, Lewis '96)
  Reason: $\max_{X\in\sigma^{-1}(S)} \langle V, X\rangle = \max_{x\in S} \langle \sigma(V), x\rangle$

Here $\langle V, X\rangle = \mathrm{tr}\,(V^T X)$ is the trace product.

- $\sigma^{-1}(S)$ is C^p-smooth $\iff$ S is C^p-smooth
  (Sylvester '85, Šilhavý '99, Daniilidis-D-Lewis '14)
  Reason: $\mathrm{dist}_{\sigma^{-1}(S)}(Y) = \mathrm{dist}_S(\sigma(Y))$
- $\sigma^{-1}(S)$ is algebraic $\iff$ S is algebraic
  (Daniilidis-Malick-Sendov '09)

Guiding philosophy

Transfer Principle:

"Properties" of $\sigma^{-1}(S)$ and S are in one-to-one correspondence.

Examples:

- $\sigma^{-1}(S)$ convex $\iff$ S convex
  (von Neumann '37, Davis '57, Lewis '96)
  Reason: $\max_{X \in \sigma^{-1}(S)} \langle V, X \rangle = \max_{x \in S} \langle \sigma(V), x \rangle$

Here $\langle V, X \rangle = \mathrm{tr}\,(V^T X)$ is the trace product.

- $\sigma^{-1}(S)$ is C^p-smooth $\iff$ S is C^p-smooth
  (Sylvester '85, Šilhavý '99, Daniilidis-D-Lewis '14)
  Reason: $\mathrm{dist}_{\sigma^{-1}(S)}(Y) = \mathrm{dist}_S(\sigma(Y))$
- $\sigma^{-1}(S)$ is algebraic $\iff$ S is algebraic
  (Daniilidis-Malick-Sendov '09)
  Reason: $e_i(\sigma_1^2(X), \ldots, \sigma_n^2(X))$ coefficients of $\det(\lambda I - X^T X)$.

Metric comparison: (von Neumann '37, Fan '49)
Always

$$\|\sigma(X) - \sigma(Y)\|_2 \leq \|X - Y\|_F.$$

Metric comparison

Metric comparison: (von Neumann '37, Fan '49)
Always

$$\|\sigma(X) - \sigma(Y)\|_2 \leq \|X - Y\|_F.$$

Equality holds $\iff$ $\exists\, U, V \in O^n$ with

$$U^T X V = \operatorname{Diag} \sigma(X) \quad \text{and} \quad U^T Y V = \operatorname{Diag} \sigma(Y)$$

(simultaneous ordered singular value decomp.)

Metric comparison

Metric comparison: (von Neumann '37, Fan '49)
Always

$$\|\sigma(X) - \sigma(Y)\|_2 \leq \|X - Y\|_F.$$

Equality holds $\iff$ $\exists\, U, V \in O^n$ with

$$U^T X V = \operatorname{Diag}\sigma(X) \qquad \text{and} \qquad U^T Y V = \operatorname{Diag}\sigma(Y)$$

(simultaneous ordered singular value decomp.)

Special case: (Hardy-Littlewood-Pólya '52)

$$\|x^{\uparrow} - y^{\uparrow}\| \leq \|x - y\|$$

Euclidean Distance Degree

Consider variety

$$\mathcal{V} = \{x \in \mathbf{R}^n : f_1(x) = f_2(x) = \ldots = f_k(x) = 0\}.$$

with f_i polynomials.

Euclidean Distance Degree

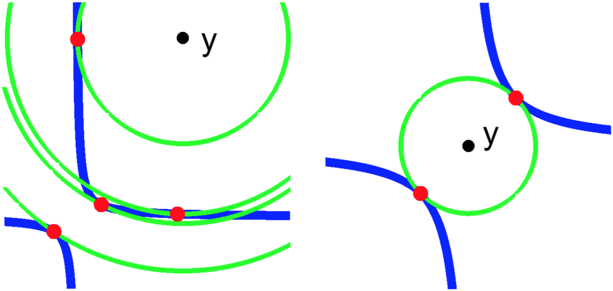
Consider variety

$$\mathcal{V} = \{x \in \mathbf{R}^n : f_1(x) = f_2(x) = \ldots = f_k(x) = 0\}.$$

with f_i polynomials.

Defn (ED critical point):

$$x \in \text{EDcrit}(y) \quad \Longleftrightarrow \quad x \text{ smooth on } \mathcal{V}, \quad y - x \perp \mathcal{T}_x\mathcal{V}_{\mathbb{C}}.$$



$\boxed{\text{EDdegree}(\mathcal{V}) = |\text{EDcrit}(y)|}$ (constant for general y!).

(Draisma-Horobeţ-Ottaviani-Sturmfels-Thomas '15)

Hilbert & Cohn-Vossen:
"Anschauliche Geometrie" (Geometry and the Imagination)
Springer-Verlag, Berlin 1932

"The simplest curves are the planar curves.
Among them the simplest one is the line.
The next simplest is the circle.
After that comes the parabola,
and finally, general conics."

Hilbert & Cohn-Vossen:
"Anschauliche Geometrie" (Geometry and the Imagination)
Springer-Verlag, Berlin 1932

"The simplest curves are the planar curves.
Among them the simplest one is the line (EDdegree 1).
The next simplest is the circle (EDdegree 2).
After that comes the parabola (EDdegree 3),
and finally, general conics (EDdegree 4).

ED degree of O^n-invariant varieties

Theorem: (D-Lee-Ottaviani-Thomas '15)

$$\boxed{\text{EDdegree}(\sigma^{-1}(S)) = \text{EDdegree}(S)}$$

ED degree of O^n-invariant varieties

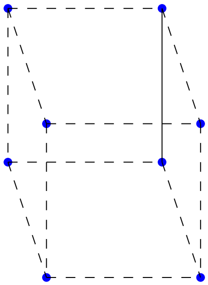
Theorem: (D-Lee-Ottaviani-Thomas '15)

$$\boxed{\text{EDdegree}(\sigma^{-1}(S)) = \text{EDdegree}(S)}$$

Examples:

Orthogonal Group: $O^n = \{X : X^T X = I\}$

$$\text{EDdegree}(O^n) = 2^n$$



(Draisma-Baaijens '14)

ED degree of O^n-invariant varieties

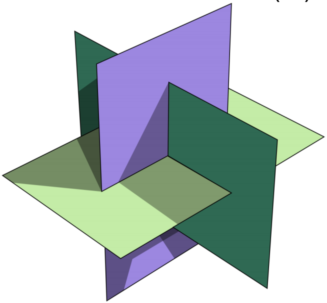
Theorem: (D-Lee-Ottaviani-Thomas '15)

$$\boxed{\mathrm{EDdegree}(\sigma^{-1}(S)) = \mathrm{EDdegree}(S)}$$

Examples:

Low rank matrices: $\mathbf{R}_r^{n\times n} = \{X : \operatorname{rank} X \le r\}$

$$\mathrm{EDdegree}(\mathbf{R}_r^{n\times n}) = \binom{n}{r}$$



(Draisma-Horobeţ-Ottaviani-Sturmfels-Thomas '15)

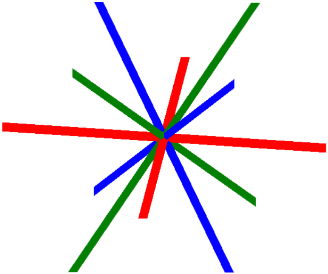
Theorem: (D-Lee-Ottaviani-Thomas '15)

$$\boxed{\text{EDdegree}(\sigma^{-1}(S)) = \text{EDdegree}(S)}$$

Examples:

Essential Variety: $\mathcal{E}^{3\times 3} = \{X : \sigma_1(X) = \sigma_2(X), \sigma_3(X) = 0\}$

$$\text{EDdegree}(\mathcal{E}^n) = 6$$



(D-Lee-Ottaviani-Thomas '15)

ED degree of O^n-invariant varieties

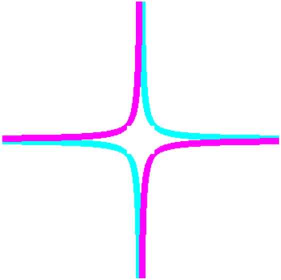
Theorem: (D-Lee-Ottaviani-Thomas '15)

$$\boxed{\text{EDdegree}(\sigma^{-1}(S)) = \text{EDdegree}(S)}$$

Examples:

Special Linear Group: $\quad \mathrm{SL}^n_{\pm} = \{X : \det(X) = \pm 1\}$

$$\text{EDdegree}(\mathrm{SL}^n_{\pm}) = n2^n$$



(Draisma-Baaijens '14)

ED degree of O^n-invariant varieties

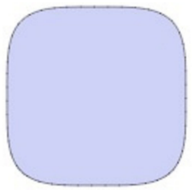
Theorem: (D-Lee-Ottaviani-Thomas '15)

$$\boxed{\mathrm{EDdegree}(\sigma^{-1}(S)) = \mathrm{EDdegree}(S)}$$

Examples:

Schatten hypersurface: $\mathcal{F}_{n,d} = \{X : \sum_i \sigma_i(X)^d = 1\}$

$$\mathrm{EDdegree}(\mathcal{F}_{n,d}) = d\sum_{i=1}^{n}(d-1)^i - \sum_{j=1}^{n-1}\binom{n}{j+1}\delta(j, d-2).$$



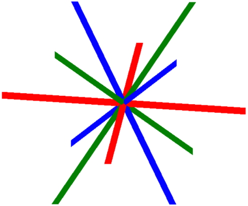
(Lee '14)

Why is this surprising?

Main challenges:

1) No explicit polynomial description of $\sigma^{-1}(S)$.

Eg. $\mathcal{E}^{3\times 3} = \{X : \det(X) = 0,\ 2XX^TX - \mathrm{tr}\,(XX^T)X = 0\}$



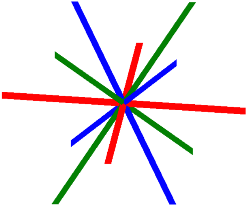
2)

Why is this surprising?

Main challenges:

1) No explicit polynomial description of $\sigma^{-1}(S)$.

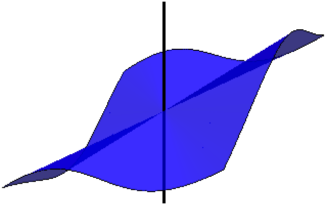
Eg. $\mathcal{E}^{3\times 3} = \{X : \det(X) = 0,\ 2XX^TX - \operatorname{tr}(XX^T)X = 0\}$



2) Smoothness is finicky:

Eg. Cartan Umbrella

$$\mathcal{V} = \{(x, y, z) : z(x^2 + y^2) - x^3 = 0\}$$



Strategy: step I (complexification)

Setting $\mathcal{M} = \sigma^{-1}(S)$, we have

$$\mathcal{M} = \bigcup_{U,V \in O^n} U \,\mathrm{Diag}\,(S)\, V^T$$

Strategy: step I (complexification)

Setting $\mathcal{M} = \sigma^{-1}(S)$, we have

$$\mathcal{M} = \bigcup_{U,V \in O^n} U \,\mathrm{Diag}\,(S)\, V^T$$

What about $\mathcal{M}_{\mathbb{C}}$?

Strategy: step I (complexification)

Setting $\mathcal{M} = \sigma^{-1}(S)$, we have

$$\mathcal{M} = \bigcup_{U,V \in O^n} U \operatorname{Diag}(S) V^T$$

What about $\mathcal{M}_{\mathbb{C}}$?

Obstruction: (Choudhury-Horn '87)
A matrix $X \in \mathbb{C}^{n \times n}$ admits an algebraic SVD

$$X = U\operatorname{Diag}(s)V^T \quad \text{with } U, V \in O^n_{\mathbb{C}} \text{ and } s \in \mathbb{C}^n$$

if and only if XX^T is diagonalizable and $\operatorname{rank} X = \operatorname{rank} XX^T$.

Strategy: step I (complexification)

Setting $\mathcal{M} = \sigma^{-1}(S)$, we have

$$\mathcal{M} = \bigcup_{U,V \in O^n} U\,\mathrm{Diag}\,(S)\,V^T$$

What about $\mathcal{M}_{\mathbb{C}}$?

Obstruction: (Choudhury-Horn '87)
A matrix $X \in \mathbb{C}^{n\times n}$ admits an algebraic SVD

$$X = U\mathrm{Diag}\,(s)V^T \quad \text{with } U, V \in O^n_{\mathbb{C}} \text{ and } s \in \mathbb{C}^n$$

if and only if XX^T is diagonalizable and $\mathrm{rank}\,X = \mathrm{rank}\,XX^T$.

Thm: (D-Lee-Ottaviani-Thomas '15)
Equalities hold:

$$\mathcal{S}_{\mathbb{C}} = \{x : \mathrm{Diag}\,(x) \in \mathcal{M}_{\mathbb{C}}\} \quad \text{and} \quad \mathcal{M}_{\mathbb{C}} = \overline{\bigcup_{U,V \in O^n_{\mathbb{C}}} U\,\mathrm{Diag}\,(S_{\mathbb{C}})\,V^T}$$

Strategy: step II (GIT perspective)

Suppose $\mathcal{M}_{\mathbb{C}}$ is invariant under $\mathcal{G} := O_{\mathbb{C}}^n \times O_{\mathbb{C}}^n$ as before.

Invariant ring:

$$\mathrm{Inv} = \{f \in \mathbb{C}[\mathcal{M}_{\mathbb{C}}] : f \text{ is } \mathcal{G}\text{-invariant}\}$$

Strategy: step II (GIT perspective)

Suppose $\mathcal{M}_{\mathbb{C}}$ is invariant under $\mathcal{G} := O_{\mathbb{C}}^n \times O_{\mathbb{C}}^n$ as before.

Invariant ring:

$$\text{Inv} = \{f \in \mathbb{C}[\mathcal{M}_{\mathbb{C}}] : f \text{ is } \mathcal{G}\text{-invariant}\}$$

Then $\text{Inv} \cong \mathbb{C}[\mathcal{M}_{\mathbb{C}} /\!\!/ \mathcal{G}]$ where $\mathcal{M}_{\mathbb{C}} /\!\!/ \mathcal{G}$ is the *GIT quotient*

$$\mathcal{M}_{\mathbb{C}} /\!\!/ \mathcal{G} = \{y \in \mathbb{C}^n : y_i = e_i(x_1^2, \ldots, x_n^2) \quad \text{for some } x \in S_{\mathbb{C}}\}$$

(Geometric Invariant Theory (GIT) developed by Mumford '65)

Strategy: step II (GIT perspective)

Suppose $\mathcal{M}_{\mathbb{C}}$ is invariant under $\mathcal{G} := O^n_{\mathbb{C}} \times O^n_{\mathbb{C}}$ as before.

Invariant ring:

$$\mathrm{Inv} = \{f \in \mathbb{C}[\mathcal{M}_{\mathbb{C}}] : f \text{ is } \mathcal{G}\text{-invariant}\}$$

Then $\mathrm{Inv} \cong \mathbb{C}[\mathcal{M}_{\mathbb{C}} /\!\!/ \mathcal{G}]$ where $\mathcal{M}_{\mathbb{C}} /\!\!/ \mathcal{G}$ is the *GIT quotient*

$$\mathcal{M}_{\mathbb{C}} /\!\!/ \mathcal{G} = \{y \in \mathbb{C}^n : y_i = e_i(x_1^2, \ldots, x_n^2) \quad \text{for some } x \in S_{\mathbb{C}}\}$$

(Geometric Invariant Theory (GIT) developed by Mumford '65)

Then the quotient map is

$$\begin{aligned} \pi \colon \mathcal{M}_{\mathbb{C}} &\to \mathcal{M}_{\mathbb{C}} /\!\!/ \mathcal{G} \\ X &\mapsto (e_1(XX^T), \ldots, e_n(XX^T)) \end{aligned}$$

Strategy: step II (GIT perspective)

Suppose $\mathcal{M}_{\mathbb{C}}$ is invariant under $\mathcal{G} := O_{\mathbb{C}}^n \times O_{\mathbb{C}}^n$ as before.

Invariant ring:

$$\mathrm{Inv} = \{f \in \mathbb{C}[\mathcal{M}_{\mathbb{C}}] : f \text{ is } \mathcal{G}\text{-invariant}\}$$

Then $\mathrm{Inv} \cong \mathbb{C}[\mathcal{M}_{\mathbb{C}} /\!\!/ \mathcal{G}]$ where $\mathcal{M}_{\mathbb{C}} /\!\!/ \mathcal{G}$ is the *GIT quotient*

$$\mathcal{M}_{\mathbb{C}} /\!\!/ \mathcal{G} = \{y \in \mathbb{C}^n : y_i = e_i(x_1^2, \ldots, x_n^2) \quad \text{for some } x \in S_{\mathbb{C}}\}$$

(Geometric Invariant Theory (GIT) developed by Mumford '65)

Then the quotient map is

$$\begin{aligned} \pi \colon \mathcal{M}_{\mathbb{C}} &\to \mathcal{M}_{\mathbb{C}} /\!\!/ \mathcal{G} \\ X &\mapsto (e_1(XX^T), \ldots, e_n(XX^T)) \end{aligned}$$

- The fiber $\pi^{-1}(y)$ is a union of orbits, and
- $\pi^{-1}(y)$ is a single orbit $\Leftrightarrow$ $\pi^{-1}(y)$ contains a diagonal matrix.

Conclusion

- Classical ideas in algebra/geometry/matrix analysis come together to explain behavior of O^n-invariant varieties.

Thank you.