Algebra and Geometry

lecture 8: normal forms

Misha Verbitsky

Université Libre de Bruxelles

December 1, 2016

Jordan normal form

DEFINITION: Let G be a group (typically, a Lie group such as $GL(V)$ or $SO(V)$ acting on a set M. **Normal form** is a subset $Z \subset M$ which intersects each orbit in a finite, non-empty subset.

EXAMPLE: The group $GL(V)$ acts on the set $\operatorname{End}(V)$ by $g(A) = gAg^{-1}$ (this is called **adjoint action**). **Jordan normal form** is the set of matrices which have block form

$$\mathsf{JNF} := \begin{pmatrix} A_{\lambda_1} & & & 0 \\ & A_{\lambda_2} & & \\ & & \ddots & \\ 0 & & & A_{\lambda_k} \end{pmatrix}$$

with each block

$$A_{\lambda_i} = \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix}.$$

Indeed, for each $A \in \operatorname{End}(\mathbb{C}^n)$ there exists finitely many Jordan block matrices A' such that $A' = gAg^{-1}$ (finitely many because **the blocks are not ordered,** you can freely exchange A_{λ_i} and A_{λ_j}).

Therefore, **the set Z of matrices of form JNF intersects with each orbit of the adjoint action in a nonempty, finite set.**

Normal form of orthogonal matrix (even dimension)

THEOREM: Let V be a real vector space equipped with a positive definite scalar product. Consider the special orthogonal group $SO(V)$ acting on $SO(V)$ by $g(A) = gAg^{-1}$ ("adjoint action"; this action corresponds to a basis change from $x_1, ..., x_n$ to $g(x_1), ..., g(x_n)$). Then for $\dim V$ even, for each $A \in SO(V)$ **there is a basis $g(x_1), ..., g(x_n)$ in which the matrix A has the following block form**

$$gAg^{-1} = \begin{pmatrix} A_{\alpha_1} & & & 0 \\ & A_{\alpha_2} & & \\ & & \ddots & \\ 0 & & & A_{\alpha_k} \end{pmatrix}$$

with $\alpha_i \in [0, 2\pi[$ and A_{α_i} the corresponding rotation matrix,

$$A_{\alpha_i} = \begin{pmatrix} \cos \alpha_i & \sin \alpha_i \\ -\sin \alpha_i & \cos \alpha_i \end{pmatrix}$$

This is called **the normal form of orthogonal transform** (for even dimension).

Normal form of orthogonal matrix (odd dimension)

THEOREM: Let V be a real vector space equipped with a positive definite scalar product. Consider the orthogonal group $SO(V)$ acting on $SO(V)$ by $g(A) = gAg^{-1}$ ("adjoint action"; this action corresponds to a basis change from $x_1, ..., x_n$ to $g(x_1), ..., g(x_n)$). **For $\dim V$ odd, for some $g \in SO(V)$ one has**

$$gAg^{-1} = \begin{pmatrix} A_{\alpha_1} & & & & 0 \\ & A_{\alpha_2} & & & \\ & & \ddots & & \\ & & & A_{\alpha_k} & \\ 0 & & & & 1 \end{pmatrix}$$

(there is an extra 1-dimensional unit block). This is called **the normal form of orthogonal transform** (for odd dimension).

REMARK: Notice that eigenvalues of A_{α_i} are equal to $\cos \alpha_i \pm \sqrt{-1} \sin \alpha_i = e^{\pm\sqrt{-1}\,\alpha_i}$, hence the block form is determined by the matrix A uniquely up to permutation. Therefore, **it is indeed a normal form:** in each adjoint orbit, there are only finitely many matrices of this form.

Normal form of orthogonal matrix (proof, page 1)

Proof. Step 1: Let us construct the normal form of orthogonal matrix $A \in O(V)$ (orthogonal matrices, not necessarily preserving orientation). Consider an eigenvector $v_1 \in V \otimes_{\mathbb{R}} \mathbb{C}$, and let α_1 be the corresponding eigenvalue. Denote by h the scalar product h_0 on V, extended to $V \otimes_{\mathbb{R}} \mathbb{C}$ by $h(x,y) = h_0(\operatorname{Re} x, \operatorname{Re} y) + h_0(\operatorname{Im} x, \operatorname{Im} y)$. Clearly, h is $SO(V)$-invariant.

The quadratic form $x \mapsto h(x, \overline{x})$ is positive definite and A-invariant. Since $h(v, \overline{v}) = h(A(v_1), A(\overline{v}_1)) = \alpha_1 \overline{\alpha}_1 h(v, \overline{v})$, we have $|\alpha_1| = 1$: **all eigenvalues α_i of $A \in O(V)$ satisfy $|\alpha_i| = 1$**

Step 2: The eigenvector v_1 is real if and only if $\alpha_1 = \pm 1$. In the later case, consider the space $V_1 = v_1^\perp$ (orthogonal complement). For each $w \in V_1$, one has $0 = h(v_1, w) = h(A(v_1), A(w)) = \pm h(v_1, A(w)) = 0$, hence V_1 is A-invariant. Using induction, we may represent A in a block form as

$$gAg^{-1} = \begin{pmatrix} \pm 1 & 0 \\ 0 & A', \end{pmatrix}$$

where $A' := g' A\big|_{V_1} (g')^{-1}$ is a block matrix in the normal form.

Normal form of orthogonal matrix (proof, page 2)

Step 3: To finish the proof, it remains to treat the case $\alpha_1 \notin \mathbb{R}$. In this case, v_1 and $\overline{v}_1$ are eigenspaces, and A acts on a 2-dimensional real space $\langle \operatorname{Re} v, \operatorname{Im} v\rangle$ by rotation with eigenvalues $\alpha_1, \overline{\alpha}_1$. The matric of such a rotation is

$$A_{\alpha_i} = \begin{pmatrix} \cos\alpha_i & \sin\alpha_i \\ -\sin\alpha_i & \cos\alpha_i \end{pmatrix}.$$

To finish the proof, it would suffice to show that A preserves the space $V_1 := \langle \operatorname{Re} v_1, \operatorname{Im} v_1\rangle^\perp$. Then we would represent $A' = g'A\Big|_{V_1}(g')^{-1}$ as a block matrix in the normal form, using induction on $\dim V$, and obtain

$$gAg^{-1} = \begin{pmatrix} \cos\alpha_i & \sin\alpha_i & 0 \\ -\sin\alpha_i & \cos\alpha_i & 0 \\ 0 & & A' \end{pmatrix}$$

Step 4: For each $w \in V_1$, and each $v \in \langle \operatorname{Re} v_1, \operatorname{Im} v_1\rangle$ one has $A(v) \in \langle \operatorname{Re} v_1, \operatorname{Im} v_1\rangle$, hence $0 = h(A^{-1}(v), w) = h(v, A(w)) = 0$, and **this implies that $A(w) \in V_1$.**

Normal form of orthogonal matrix (proof, page 3)

Step 5: We proved that **any $A \in O(V)$ can be represented in block form with the blocks either 1-dimensional and equal to ± 1 or 2-dimensional and equal to $\begin{pmatrix} \cos\alpha_i & \sin\alpha_i \\ -\sin\alpha_i & \cos\alpha_i \end{pmatrix}$.** For $A \in SO(V)$, there is an even number of -1-blocks, which can be grouped together to block matrices $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = A_\pi$. **The number of 1-blocks is even when $\dim V$ is even and odd when it is odd.** We can group them together pairwise into matrices $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = A_0$. ∎

Orthogonal matrices are exponents

COROLLARY: Let $A \in SO(V)$ be a matrix written in block form as

$$A = \begin{pmatrix} A_{\alpha_1} & & & 0 \\ & A_{\alpha_2} & & \\ & & \ddots & \\ 0 & & & A_{\alpha_k} \end{pmatrix}$$

with

$$A_{\alpha_i} = \begin{pmatrix} \cos\alpha_i & \sin\alpha_i \\ -\sin\alpha_i & \cos\alpha_i \end{pmatrix}$$

Then $A = e^B$, where B is written in the same basis as

$$B = \begin{pmatrix} B_{\alpha_1} & & & 0 \\ & B_{\alpha_2} & & \\ & & \ddots & \\ 0 & & & B_{\alpha_k} \end{pmatrix}$$

with

$$B_{\alpha_i} = \begin{pmatrix} 0 & \alpha_i \\ -\alpha_i & 0 \end{pmatrix}$$

■

Normal form of a bilinear symmetric form

Recall that **the group $GL(V)$ acts on bilinear forms by**

$$g(h)(\cdot,\cdot) = h(g^{-1}(\cdot), g^{-1}(\cdot)),$$

where $g \in GL(V)$, $h \in \operatorname{Bil}(V) = V^* \otimes V^*$. This action corresponds to a basis change.

We consider an action of $GL(V)$ on pairs of bilinear symmetric forms, and **find a normal form of this action on the set of pairs of forms.**

DEFINITION: Let $h \in \operatorname{Sym}^2 V^*$ be a bilinear symmetric form, and $x_1, ..., x_n$ a basis. This basis is called **orthogonal** if $h(x_i, x_j) = 0$ for $i \neq j$, and **orthonormal** if in addition $h(x_i, x_i) = \pm 1$.

REMARK: Previously, we proved that **any non-degenerate bilinear symmetric form on $\mathbb{R}^n$ admits an orthonormal basis.** This result can be understood as **providing the normal form of a non-degenerate bilinear symmetric form.**

Normal form for a pair of bilinear symmetric forms (1)

The following result is proven at the end of this lecture.

Theorem 1: Let $V = \mathbb{R}^n$, and $h, h' \in \operatorname{Sym}^2 V^*$ be two bilinear symmetric forms, with h positive definite. **Then there exists a basis $x_1, ..., x_n$ which is orthonormal with respect to h, and orthogonal with respect to h'.**

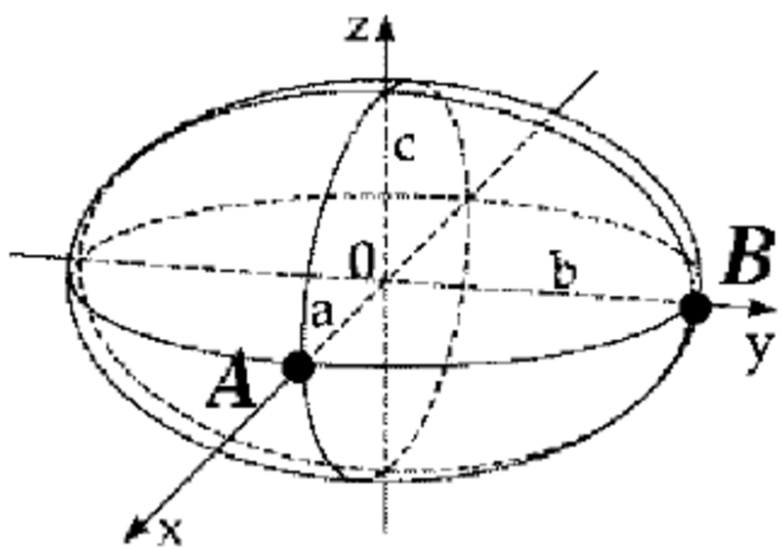
REMARK: In this basis, h' is written as diagonal matrix, with eigenvalues $\alpha_1, ..., \alpha_n$ independent from the choice of the basis. Indeed, consider h, h' as maps from V to V^*, $h(v) = h(v, \cdot)$. Then $h_1 h^{-1}$ is an endomorphism with eigenvalues $\alpha_1, ..., \alpha_n$. **This implies that Theorem 1 gives a normal form of the pair h, h'.**

Finding principal axes of an ellipsoid

REMARK: Theorem 1 implies the following statement about ellipsoids: for any positive definite quadratic form q in $\mathbb{R}^n$, consider the ellipsoid

$$S = \{v \in V \mid q(v) = 1\}.$$

The group $SO(n)$ acts on $\mathbb{R}^n$ preserving the standard scalar product. **Then for some $g \in SO(n)$, $g(S)$ is given by equation $\sum a_i x_i^2 = 1$, where $a_i > 0$.** This is called **finding principal axes of an ellipsoid**.



Compactness

DEFINITION: Recall that a subset $Z \subset \mathbb{R}^n$ is called **(sequentially) compact** if any sequence $x_1, ..., x_n, ... \subset Z$ has a converging subsequence.

THEOREM: A subset $Z \subset \mathbb{R}^n$ is **sequentially compact if and only if Z is closed and bounded** (that is, contained in a ball of finite diameter).

EXERCISE: Prove this theorem.

EXERCISE: Let f be a continuous function on a compact Z. **Prove that Z is bounded and attains its supremum on Z.**

COROLLARY: Let f be a continuous function on a sphere $S^n \subset \mathbb{R}^{n+1}$. **Then f is bounded, and attains its supremum.**

Further on, we need the following lemma.

LEMMA: Let $V = \mathbb{R}^n$, and $h, h' \in \operatorname{Sym}^2 V^*$ be two bilinear symmetric forms, h positive definite, and $q(v) = h(v,v), q'(v) = h'(v,v)$ the corresponding quadratic forms. Consider q' as a function on a sphere $S = \{v \in V \mid q(v) = 1\}$, and let $x \in S$ be the point where q' attains maximum. Denote by $x^{\perp_h}$ and $x^{\perp_{h'}}$ the orthogonal complement with respect to h, h'. **Then $x^{\perp_h} = x^{\perp_{h'}}$.**

Maximum of a quadratic form on a sphere

LEMMA: Let $V = \mathbb{R}^n$, and $h, h' \in \operatorname{Sym}^2 V^*$ be two bilinear symmetric forms, h positive definite, and $q(v) = h(v,v), q'(v) = h'(v,v)$ the corresponding quadratic forms. Consider q' as a function on a sphere $S = \{v \in V \mid q(v) = 1\}$, and let $x \in S$ be the point where q' attains maximum. Denote by $x^{\perp_h}$ and $x^{\perp_{h'}}$ the orthogonal complements with respect to h, h'. **Then $x^{\perp_h} = x^{\perp_{h'}}$.**

Proof: Let us rescale q, q' in such a way that $q \geqslant q'$, with equality on x. Suppose that $v \in x^{\perp_h}$. Then $q(x + \varepsilon v) = q(x) + \varepsilon^2 q(v)$. However, $q'(x + \varepsilon v) = q(x) + \varepsilon^2 q'(v) + 2\varepsilon h'(v,x)$. This gives

$$q(x) + \varepsilon^2 q(v) \geqslant q(x) + \varepsilon^2 q'(v) + 2\varepsilon h'(v,x)$$

cancelling $q(x)$ and dividing by $\varepsilon > 0$, obtain

$$\varepsilon(q(v) - q'(v)) \geqslant 2h'(v,x).$$

for all $\varepsilon > 0$. This implies that $0 \geqslant 2h'(v,x)$ for all $v \in x^{\perp_h}$. **Since $v \mapsto h'(v,x)$ is a linear form on v, inequality $0 \geqslant h'(v,x)$ implies that $h'(v,x) = 0$.** ■

Normal form for a pair of bilinear symmetric forms (2)

Theorem 1: Let $V = \mathbb{R}^n$, and $h, h' \in \operatorname{Sym}^2 V^*$ be two bilinear symmetric forms, with h positive definite. **Then there exists a basis $x_1, ..., x_n$ which is orthonormal with respect to h, and orthogonal with respect to h'.**

Proof: Let $q(v) = h(v,v), q'(v) = h'(v,v)$ the corresponding quadratic forms. Consider q' as a function on a sphere $S = \{v \in V \ \ | \ \ q(v) = 1\}$, and let $x_1 \in S$ be the point where q' attains maximum. Then $x_1^{\perp_h} = x_1^{\perp_{h'}}$. Using induction, we may assume that on $x_1^{\perp_h}$, Theorem 1 is already proven, and there exists a basis $x_2, ..., x_n$ orthonormal for h and orthogonal for h'. **Then $x_1, x_2, ..., x_n$ is orthonormal for h and orthogonal for h'.** ∎