Lecture 11: Orthogonal Groups

Orthogonal Groups

Definition

Suppose $(v, (\ ,\ ))$ is an inner product space. Let $S \in \mathrm{Hom}\,(V, V)$. Then S is said to be orthogonal if

$$(Sv, Sw) = (v, w)\,, \qquad \text{all } v, w \in V.$$

We let $O\,(V, (\ ,\ ))$ denote the set of orthogonal linear transformations. (We will often write $O(V)$.)

Proposition

$O(v)$ is a subgroup $\mathrm{Aut}(V)$.

Proof. We show $O(v)$ is closed under $\circ$ and inverse.

Closed under $\circ$:
Suppose S, $T \in O(V)$. Let v, $w \in V$. Then

$$\begin{aligned} ((S \circ T)v, (S \circ T)w) &= ((S(Tv), (S(Tw)) \text{ by definition } \circ \\ &= (Tv, Tw) \text{ using } S \in O(V) \\ &= (v, w) \text{ using } T \in O(V) \end{aligned}$$

Closed under inverse:
Let $S \in O(V)$. First we show S^{-1} exists in $\mathrm{Hom}\,(V, V)$, then we will show $S^1 \in O(V)$. To show S is an invertible linear transformation it suffices to show S is 1:1 because $S : V \longrightarrow V$ so 1:1 $\Longrightarrow$ onto. To show S is 1:1 it suffices to prove $N(S) = \{0\}$.
Suppose $v \in N(S)$. Then $Sv = 0$ and hence $(Sv, Sv) = 0$. Since S is orthogonal, this implies (v, v), hence $v = 0$. Thus $N(S) = 0$.

Now we have $S^{-1} \in \mathrm{Aut}(V)$, but is $S^{-1} \in O(V)$?
Let $v, w \in V$, we need to show

$$\left(S^{-1}v,\, S^{-1}w\right) = (v,\, w) \qquad (*)$$

Since S is onto, there are $v',\, w' \in V$, so that

$$v = Sv',\, w = Sw'.$$

Substituting in $(*)$, we need to show

$$\left(S^{-1}Sv',\, S^{-1}Sw'\right) = (Sv',\, Sw')$$

But $S^{-1}S = I_V$, so

$$(v',\, w') = \left(S^{-1}Sv',\, S^{-1}Sw'\right) = (Sv',\, Sw') \qquad \square$$

Now since $||\cdot||$ and $\measuredangle$ are defined in terms of $(\ ,\ )$, we have

$$S \in O(V) \Longrightarrow S \text{ preserves length and angles.}$$

Precisely, for v, $w \in V$, we have

$$\begin{aligned} ||Sv|| &= \sqrt{(Sv,\, Sv)} = \sqrt{(v,\, v)} = ||V|| \\ \measuredangle\,(Sv,\, Sw) &= \frac{(Sv,\, Sw)}{||Sv||\,||Sw||} = \frac{(v,\, w)}{||v||\,||w||} = \measuredangle\,(v,\, w)\,. \end{aligned}$$

There is a converse:

Proposition

Suppose $S \in \mathrm{Hom}\,(V,\, V)$ and S preserves lengths (i.e., $||Sv|| = ||v||$, for all $v \in V$). Then $S \in O(V)$.

Proof. We will use an extremely important formula, the **polarization** formula:

$$(u,\, v) = \frac{1}{2}\left(||u+v||^2 - ||u||^2 - ||v||^2\right)$$

Now observe

$$\begin{aligned}(Su,\, Sv) &= \frac{1}{2}\left(||Su+Sv||^2 - ||Su||^2 - ||Sv||^2\right)\\ &= \frac{1}{2}\left(||S(u+v)||^2 - ||Su||^2 - ||Sv||^2\right)\\ &= \frac{1}{2}\left(||u+v||^2 - ||Su||^2 - ||Sv||^2\right)\\ &= (u,\, v) \qquad \square\end{aligned}$$

Remark: It is not true that S preserve angles $\Longrightarrow S \in O(V)$.

Proposition (See page 131, # 12)

If $S \in \mathrm{Hom}\,(V, V)$ *preserves (right) angles then there exits* $\lambda \in \mathbb{R}$ *and* $T \in O(V)$ *so that*

$$S = \lambda T$$

Note: In this case S is said to be conformal (or a similitude).

Transpose

We now introduce the important operation transpose.

Definition

Given $T \in \mathrm{Hom}\,(V, V)$, the transpose of T, denoted ${}^tT$, is the linear transformation that satisfies

$$\left({}^tT, v\right) = (u, Tv)\,.$$

We will see below that such a transformation exists (and it will be unique).
Given a matrix $A \in M_n(\mathbb{R})$, $A = (a_{ij})$, we define the transpose of A denoted ${}^tA$, to be the matrix obtained by interchanging the rows and columns of A (or reflextion in the diagonal).
Example:

$$\text{If } A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \text{ then } {}^tA = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$$

The two transposes agree. Precisely, we have the following proposition.

Transpose

Proposition

Given an ordered orthonormal basis $\mathscr{U} = (u_1, \ldots, u_n)$ *for* V *and* $T \in \mathrm{Hom}\,(V, V)$,

$$M({}^tT) = {}^tM(T)$$

Proof. Let

$$\begin{aligned} (a_{ij}) &= M({}^tT) \\ (b_{ij}) &= {}^tM(T) \end{aligned}$$

Then $a_{ij} = (Tu_{ij}, u_i)$ and $b_{ij} = ({}^tTuj, u_i)$. Since $(\,,\,)$ is symmetric,

$$a_{ij} = (Tu_{ij}, u_i) = (u_i, Tu_j) = \left({}^tTuj, u_i\right) = b_{ij}.$$

Thus $a_{ij} = b_{ij}$. □

Note: This proves existence and uniqueness: to determine ${}^tT$, choose an orthonormal basis $\mathscr{U}$ and let ${}^tT$ be the (unique) linear transformation given by ${}^tM(T)$.

Characterization of Orthogonal Transformations

We recall Proposition (2) from Lecture 6:

Proposition

Let V be a vector space and $T \in L(V\, V) = \mathrm{Hom}\,(V,\, V)$. Let $\mathscr{B} = \{b_1,\, \ldots,\, b_n\}$ be a basis for V. Let $v \in V$. Then

$$[T(v)\mathscr{B}]_{=\ \mathscr{B}}[T]_{\mathscr{B}}\,[v]_{\mathscr{B}}\,.$$

Lemma

Suppose $\{b_1,\, \ldots,\, b_n\}$ is an orthonormal basis for V. Let $v,\, w \in V$ and

$$v = \sum_{i=1}^{n} w_i u_i, \quad w = \sum_{i=1}^{n} y_i u_i.$$

Then $(v,\, w) = \sum_{i=1}^{n} x_i y_i$.

Proof. We have

$$\begin{aligned}(v,\, w) &= \left(\sum_{i=1}^{n} w_i u_i,\, \sum_{j=1}^{n} y_j u_j\right) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} (x_i u_i,\, y_j u_j) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j\, (u_i,\, u_j)\,.\end{aligned}$$

But

$$(u_i,\, u_j) = \begin{cases} 0 & \text{if} \quad i \neq j \\ 1 & \text{if} \quad i = j \end{cases}$$

So,

$$(v,\, w) = \sum_{i=1}^{n} x_i y_i\, (u_i,\, u_i) = \sum_{i=1}^{n} x_i y_i. \quad \square$$

Theorem (Text, Theorem 15.11)

Let $T \in \mathrm{Hom}\,(V, V)$. The following are equivalent.

(1) $T \in O(V)$.

(2) *For any orthonormal basis* $\mathscr{U} = \{u_1, \ldots, u_n\}$, *the set* $\mathscr{U}' = \{Tu, \ldots, Tu_n\}$ *is again an orthonormal basis.*

(3) *The matrix* $A = M(T)$ *satisfies*

$$^{t}AA = I$$

where $\mathscr{U} = (u_1, \ldots, u_n)$ *an orthonormal basis.*

(4) *The rown and columns of* $A = M(T)$ *are each orthonormal bases for* V.

Proof.

(1) $\Longrightarrow$ (2)

$$(Tu_i,\, Tu_j) = (u_i,\, u_j) = \begin{cases} 0 & \text{if} \quad i \neq j \\ 1 & \text{if} \quad i = j \end{cases}$$

(2) $\Longrightarrow$ (3)

$$A = M(T) = \begin{pmatrix} [Tu_1]_{\mathscr{U}} & \dots & [Tu_n]_{\mathscr{U}} \\ \downarrow & \dots & \downarrow \\ & & \end{pmatrix}$$

Then,

$${}^{t}AA = \begin{pmatrix} [Tu_1]_{\mathscr{U}} & \longrightarrow & \\ & \vdots & \\ [Tu_n]_{\mathscr{U}} & \longrightarrow & \end{pmatrix} \begin{pmatrix} [Tu_1]_{\mathscr{U}} & \dots & [Tu_n]_{\mathscr{U}} \\ \downarrow & \dots & \downarrow \\ & & \end{pmatrix}$$

The ij^{th} entry of the resulting matrix is

$$\begin{aligned}
\left([Tu_i]_{\mathscr{U}} \longrightarrow\right)\left([Tu_j]_{\mathscr{U}} \downarrow\right) &= [Tu_i]_{\mathscr{U}} \cdot [Tu_j]_{\mathscr{U}} \\
&= (Tu_i,\, Tu_j) = (u_i,\, u_j) = \begin{cases} 0 & \text{if} \quad i \neq j \\ 1 & \text{if} \quad i = j \end{cases}
\end{aligned}$$

Thus the resulting matrix is the identity matrix.

(3) $\Longrightarrow$ (1) Since ${}^tM(T)M(T) = I$, the identity matrix, we have ${}^tTT = I$, the identity transformation. Thus

$$(Tu,\, Tv) = \left({}^tTTu,\, v\right) = (u,\, v)\,,$$

and hence $T \in O(V)$.

Characterization of Orthogonal Transformations

(2) $\implies$ (4)

$$A = M(T) = \begin{pmatrix} [Tu_1]_{\mathscr{U}} & \dots & [Tu_n]_{\mathscr{U}} \\ \downarrow & \dots & \downarrow \\ & & \end{pmatrix}$$

Hence the columns are an orthonormal basis. Also, if $T \in O(V)$, then ${}^tT = T^{-1} \in O(V)$ and thus since the columns of ${}^tT$ are an orthonormal basis, so are the rows of T.

(4) $\implies$ (2) Since the columns of $A = M(T)$ are an orthonormal basis, $\{Tu_1 \ \dots, Tu_n\}$ is an orthonormal basis. $\square$

Orthogonal Matrices

Definition

A matrix $A \in M_n(\mathbb{R})$ is said to be an **orthogonal** matrix if

$$ {}^tAA = I $$

The set of orthogonal matrices is denoted $O(n)$.

Proposition

A is orthogonal $\Longrightarrow {}^tA = A^{-1}$.

Proof.
($\Longrightarrow$) We know A orthogonal $\Longrightarrow A^{-1}$ exists.

$$ {}^tAA = I \Longrightarrow {}^tA = A^{-1} $$

where $\Longrightarrow$ means right multiplications by A^{-1}.
($\Longleftarrow$) Suppose ${}^tA = A^{-1}$. Then ${}^tAA = I$. $\square$

Orthogonal Matrices

Let $GL_n(\mathbb{R})$ denote the set of invertible n by n matrices.

$GL_n(\mathbb{R})$ is a group and $(AB)^{-1} = B^{-1}A^{-1}$. We've shown

Proposition

$O(n)$ is a subgroup of $GL_n(\mathbb{R})$.

The group $O(2)$

$$
\begin{aligned}
O(2) &= \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad 0 \leq \theta \leq 2\pi \right\} \\
&\cup \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}, \quad 0 \leq \theta \leq 2\pi \right\}
\end{aligned}
$$

Proof.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in O(2) \iff \begin{aligned} a^2 + c^2 &= 1 \\ b^2 + d^2 &= 1 \\ ab + cd &= 0. \end{aligned}$$

$\iff$ $(a,\, c)$ is on circle, $(b,\, d)$ is on the circle and $(a,\, c)$ is orthogonal to $(b,\, d)$.