Lecture 19: Isometries, Positive operators, Polar and singular value decompositions; Unitary matrices and classical groups; Previews (1)

Travis Schedler

Thurs, Nov 18, 2010 (version: Wed, Nov 17, 2:15 PM)
Goals (2)

- Isometries
- Positive operators
- Polar and singular value decompositions
- Unitary matrices and classical matrix groups
- Preview of rest of course!
Characterization of isometries (3)

Correction: An isometry \( V \to W \) need not be invertible even if \( V \) and \( W \) are finite-dimensional: e.g., if \( V = 0 \) the zero map is an isometry! However, it must be injective so invertible if \( \dim V = \dim W < \infty \).

Proposition

If \( T^* \) exists (e.g., if \( V \) is finite-dimensional), then \( T \in \mathcal{L}(V, W) \) is an isometry iff \( T^* T = I_V \).

Proof.

- First, \( T \) is an isometry iff \( \langle T^* T v, v \rangle = \langle v, v \rangle \) for all \( v \).
- Next, \( T^* T \) and hence \( S := T^* T - I \) is always self-adjoint.
- \( T \) is an isometry iff \( \langle S v, v \rangle = 0 \) for all \( v \).
- Now, apply the skipped result from last time below. \( \square \)

Proposition (Proposition 7.4)

If \( T \) is self-adjoint, then \( \langle T v, v \rangle = 0 \) for all \( v \) iff \( T = 0 \).
Proof of the proposition (4)

Proposition (Proposition 7.4)

If $T$ is self-adjoint, then $\langle Tv, v \rangle = 0$ for all $v$ iff $T = 0$.

Proof.

- Assume $\langle Tv, v \rangle = 0$ for all $v$. Then
  $\langle T(v + u), v + u \rangle - \langle T(v), v \rangle - \langle T(u), u \rangle = 0$ for all $v, u$.
- Thus, $\langle T(v), u \rangle + \langle T(u), v \rangle = 0$ for all $u, v$.
- When $T$ is self-adjoint, this says $2\Re \langle T(v), u \rangle = 0$ for all $u, v$.
- Plugging in $iu$ for $u$, also $2\Im \langle T(v), u \rangle = 0$ for all $u, v$. Thus $T(v) = 0$ for all $v$, i.e., $T = 0$.

Note: From another skipped result on last week’s slides, when $V$ is finite-dimensional, $\langle Tv, v \rangle = 0$ for all $v$ iff $T = 0$ or $F = \mathbb{R}$ and $T$ is anti-self-adjoint.
Positive operators (5)

Definition
A positive operator $T$ is a self-adjoint operator such that $\langle Tv, v \rangle \geq 0$ for all $v$.

In view of the spectral theorem, a self-adjoint operator is positive iff its eigenvalues are nonnegative (part of Theorem 7.27).

Theorem (Remainder of Theorem 7.27)

(i) Every operator of the form $T = S^* S$ is positive.
(ii) Every positive operator admits a positive square root.

Proof.

- (i) First, $T^* = (S^* S) = S^* S$ is self-adjoint.
- Next, $\langle Tv, v \rangle = \langle Sv, Sv \rangle \geq 0$ for all $v$.
- (ii) For any orthonormal eigenbasis of $T$, let $\sqrt{T}$ be the operator with the same orthonormal eigenbasis, but with the nonnegative square root of the eigenvalues.
Polar decomposition (6)

After the spectral theorem, the second-most important theorem of Chapters 6 and 7 is:

Theorem (Polar decomposition: Theorem 7.41)

Every $T \in \mathcal{L}(V)$ equals $S\sqrt{T^*T}$ for some isometry $S$.

Main difficulty: $T$ need not be invertible!

Lemma

For all $v \in V$, $\|Tv\| = \|\sqrt{T^*T}v\|$.

Proof.

\[ \|\sqrt{T^*T}v\|^2 = \langle \sqrt{T^*T}v, \sqrt{T^*T}v \rangle = \langle (\sqrt{T^*T})^*\sqrt{T^*T}v, v \rangle = \langle T^*Tv, v \rangle = \langle Tv, Tv \rangle = \|Tv\|^2. \]

Corollary: $\text{null}(T) = \text{null}(\sqrt{T^*T})$. We may thus define $S_1 : \text{range}(\sqrt{T^*T}) \sim \text{range}(T)$ by $S_1(\sqrt{T^*T}v) = Tv$.

Thus, for all $v \in V$, $S_1\sqrt{T^*T}v = Tv$.

Also, $\|S_1u\| = \|u\|$ for all $u(= \sqrt{T^*T}v)$ by the lemma.

So $S_1$ is an isometry.
Completion of proof (7)

- We only have to extend $S_1$ to an isometry on all of $V$.
- Note that $\text{range}(\sqrt{T^*T}) \oplus \text{range}(\sqrt{T^*T})^\perp = V = \text{range}(T) \oplus \text{range}(T)^\perp$.
- Thus, the extensions of $S_1 : \text{range}(\sqrt{T^*T}) \xrightarrow{\sim} \text{range}(T)$ to an isometry $S : V \xrightarrow{\sim} V$ are exactly $S = S_1 \oplus S_2$, where $S_2 : \text{range}(\sqrt{T^*T})^\perp \xrightarrow{\sim} \text{range}(T)^\perp$ is an isometry.
- Since these are inner product spaces of the same dimension, there always exists an isometry, by taking an orthonormal basis to an orthonormal basis.

Recall here that $T_1 \oplus T_2$ on $U_1 \oplus U_2$ means 

$$(T_1 \oplus T_2)(u_1 + u_2) = T_1(u_1) + T_2(u_2), \ \forall u_1 \in U_1, u_2 \in U_2.$$
Singular value decomposition (SVD) (8)

Let $V$ be finite dimensional and $T \in \mathcal{L}(V)$.

**Theorem**

There exist orthonormal bases $(e_1, \ldots, e_n)$ and $(f_1, \ldots, f_n)$ and nonnegative values $s_i \geq 0$ such that $Te_i = s_i f_i$.

The $s_i$ are called the *singular values*.

**Proof.**

- Let $(e_1, \ldots, e_n)$ be an orthonormal eigenbasis of the positive $\sqrt{T^* T}$. Let $s_1, \ldots, s_n$ be the nonnegative eigenvalues.
- Using polar decomposition, $T = S \sqrt{T^* T}$ for $S$ an isometry.
- Let $f_i := Se_i$. Then $Te_i = s_i f_i$.

**Corollary:** $\mathcal{M}_{(e_i), (f_i)}(T) = D$, where $D$ is diagonal with entries $s_i \geq 0$. *Improvement on normal form:* $(e_i), (f_i)$ orthonormal!

**Note:** To compute, first find eigenvalues $s_i^2$ of $T^* T$, then orthonormal eigenbasis $e_i$. This yields $f_i$, $S$, and $\sqrt{T^* T}$!
Unitary matrices and $A = U_1 D U_2^{-1}$ decomposition (9)

Corollary: If $A = \mathcal{M}_{(v_i)}(T)$ for any orthonormal basis, then $A = U_1 D U_2^{-1}$ for the change-of-basis matrices $U_1, U_2$. Moreover, the columns of $U_1$ and $U_2$ form orthonormal bases of $\text{Mat}(n, 1, F)$: such matrices are called unitary.

Proposition

The following are equivalent for $U \in \text{Mat}(n, n, F)$:

- The columns form an orthonormal basis;
- $UU^* = I = U^* U$;
- $U = \mathcal{M}_{(e_i)}(S)$ for some isometry $S \in \mathcal{L}(V)$ and some orthonormal basis $(e_i)$ of $V$.

Proof.

- The first two are immediately equivalent.
- In general, $\mathcal{M}_{(e_i)}(S)$ has columns equal to $Se_i$ in the basis $(e_i)$, so $\langle Se_i, Se_j \rangle = \text{the dot product of the } i\text{-th and } j\text{-th column of } \mathcal{M}_{(e_i)}(S)$. 

\[ \square \]
Example of SVD \((A = U_1 D U_2^{-1} \text{ decomposition})\)! (10)

Let \(A = \begin{pmatrix} 10 & -2 \\ -5 & 11 \end{pmatrix}\). Compute the decomposition \(A = U_1 D U_2^{-1}\).

- First step: Find eigenvalues and eigenbasis of 
  \[A^* A = \begin{pmatrix} 125 & -75 \\ -75 & 125 \end{pmatrix}\].

- Char. poly of \(A = x^2 - (\text{tr } A)x + \det A = x^2 - 250x + 10000 = (x - 200)(x - 50).\)
  Eigenvalues: \(s_1^2 = 200, s_2^2 = 50\).

- Nullspace of \(A^* A - 200I = \begin{pmatrix} -75 & -75 \\ -75 & -75 \end{pmatrix}\):
  spanned by \(e_1 := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}\).

- Same computation for \(e_2\): get \(e_2 := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}\).
  Alternative: \(e_2\) determined up to scaling so that \(e_1 \perp e_2\).
Recall: \( A = \begin{pmatrix} 10 & -2 \\ -5 & 11 \end{pmatrix} \),
\[ s_1^2 = 200, \ s_2^2 = 50, \ e_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \ e_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \]

Now, \( f_1 = s_1^{-1} A e_1 = \frac{1}{20} \begin{pmatrix} 12 \\ -16 \end{pmatrix} = \begin{pmatrix} 3/5 \\ -4/5 \end{pmatrix} \).

Similarly, \( f_2 = s_2^{-1} A e_2 = \frac{1}{10} \begin{pmatrix} 8 \\ 6 \end{pmatrix} = \begin{pmatrix} 4/5 \\ 3/5 \end{pmatrix} \).

Caution: \( f_2 \perp f_1 \), but this only determines \( f_2 \) up to scaling by absolute value one (here \( \pm 1 \))!

Now, \( U_1 = (f_1 f_2) = \begin{pmatrix} 3/5 & 4/5 \\ -4/5 & 3/5 \end{pmatrix} \),
\[ D = \begin{pmatrix} 10\sqrt{2} & 0 \\ 0 & 5\sqrt{2} \end{pmatrix}, \]
\[ \text{and } U_2 = (e_1 e_2) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}. \] Check: \( A = U_1 D U_2^{-1} \)!
Classical matrix groups (12)

Definition
A group is a set $G$ with an associative operation $G \times G \to G$ with an identity and inverses (not necessarily commutative).

Definition
$GL(n, F) \subseteq \text{Mat}(n, n, F)$ is the group of invertible matrices “general linear group,” under multiplication.
$SL(n, F) \subseteq \text{Mat}(n, n, F)$ is the subgroup of matrices of det $= 1$ “special linear group.”
(We haven’t defined determinant yet. But det$(AB) = \det(A) \det(B)$ so $SL(n, F)$ is closed under mult.)

Definition
$O(n, \mathbb{R}) = \text{the group of orthogonal (=unitary) matrices}$:
$O$ such that $O^t O = I = OO^t$.
$U(n, \mathbb{C}) = \text{the group of unitary matrices}$:
$U$ such that $U^* U = I = UU^*$.

Note: Also have $O(n, F)$: matrices such that $O^t O = I = OO^t$ for any $F$. *Hence* $U(n, \mathbb{C}) \neq O(n, \mathbb{C})$!
Finite subgroups; platonic solids (13)

- \( O(n, \mathbb{R}) = \) group generated by reflections (and rotations)!
- **Big question:** What are the finite subgroups \( G < O(n, \mathbb{R}) \)?
- In case \( n = 2 \): Just “cyclic” groups
  \[ C_m = \left\{ \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} : \theta = \frac{2\pi k}{m}; 1 \leq k \leq m \right\}, \]
  “dihedral” groups \( C_m \cup TC_m \), where \( T \) is a reflection.
- In case \( n = 3 \): Groups of rotations only are: cyclic and dihedral groups on the plane, and *platonic solid rotation groups*: groups of rotation fixing the platonic solids!
- The other groups are \( G \cup TG \) where \( G \) is as above, and \( T \) is a single reflection.
- This gives a classification of the platonic solids: the only three finite rotation groups in \( O(3, \mathbb{R}) \) which don’t fix a plane!
Classical groups of $V$; bilinear and quadratic forms (14)

Similarly:

- $\text{GL}(V), \text{SL}(V) \subseteq \mathcal{L}(V)$ are the groups of invertible, $\det = 1$ linear transformations ($V =$ vector space);
- $\text{O}(V), \text{U}(V)$ are the groups of isometries in the cases $F = \mathbb{R}, \mathbb{C}$ ($V =$ inner product space);
- For general $F$, we can define groups $\text{O}(V)$ if we generalize inner product spaces:
  - A bilinear form $\langle -, - \rangle : V \times V \to F$ is a map which is additive and homogeneous (in both slots!).
  - It is symmetric if $\langle u, v \rangle = \langle v, u \rangle$.
  - It is nondegenerate if $\langle u, v \rangle = 0$ for all $u \in V$ implies $v = 0$, and also $\langle v, u \rangle = 0$ for all $u \in V$ implies $v = 0$.

Definition

Let $V$ have a (nondegenerate symmetric) bilinear form. Then $O(V) \subseteq \mathcal{L}(V) = \{ T : \langle u, v \rangle = \langle Tu, Tv \rangle \text{ for all } u, v \in V \}$.

Quadratic form: $q(v) := \langle v, v \rangle$, satisfying $q(\lambda v) = \lambda^2 q(v)$, parallelogram identity; uniquely determines $\langle -, - \rangle$ if $1/2 \in F$. 

Characteristic polynomial: write upper-triangular matrix $A = M(T)$. If the diagonal entries are $\lambda_1, \ldots, \lambda_N$, then char. poly := $(x - \lambda_1) \cdots (x - \lambda_N)$.

Theorem 1: This does not depend on basis.
Call it $\chi_T(x) = x^N + a_{N-1}x^{N-1} + \cdots + a_0$. 
Definition: $\text{tr}(T) = -a_{N-1} = \sum \lambda_i$; $\text{det}(T) = a_0 = \prod \lambda_i$.

Theorem 2 (Cayley-Hamilton): $\chi_T(T) = 0$.

Theorem 3: $\text{tr}(T) = \text{tr}(A) = \text{sum of diagonal entries}$. 
Corollary: $\text{tr}(T) + \text{tr}(S) = \text{tr}(T + S)$ (sum of sums of eigenvalues = sum of eigenvalues of sum!)

Theorem 4: $a_0 = \text{det}(A) = \text{an explicit formula we will study, satisfying } \text{det}(AB) = \text{det}(A) \text{det}(B)$.

Case $F = R$: $|\text{det}(A)| = \text{volume of } A(\text{unit } n\text{-cube})$.

Corollary: $\text{det}(TS) = \text{det}(T)\text{det}(S)$. 
Corollary: $\chi_T(T) = \text{det}(M(T) - xI)$. So the $a_i$ are polynomial functions in entries of $A$!
Definition: $\chi_T(x) = \det(M(T) - xI)$.

- Theorem 1’: Does not depend on basis.
- Theorem 2’ (C-H): $\chi_T(T) = 0$.
- Corollary: The roots of $\chi_T$ are exactly the eigenvalues.
Finally, back to $F = \mathbb{C}$. The following only requires Theorem 1:

**Theorem**

*For* $F = \mathbb{C}$, *in some basis*, $T$ has block diagonal form with blocks

\[
\begin{pmatrix}
\lambda & 1 \\
\vdots & \ddots & \ddots \\
\vdots & \ddots & 1 \\
& & & \lambda
\end{pmatrix}.
\]

More generally (unfortunately we don’t have time to prove):

- For $F = \mathbb{R}$, we can say the same thing if we also allow $\lambda$ to be

  \[
  \begin{pmatrix}
a & -b \\
b & a
\end{pmatrix},
\]

  and then 1 becomes $I$ (alternatively, \[
  1
\]).

- For general $F$, we can say the same if we allow $\lambda$ to be a matrix with *irreducible* characteristic polynomial, and replace 1 by \[
  1
\].