LECTURE VI: SELF-ADJOINT AND UNITARY OPERATORS
MAT 204 - FALL 2006
PRINCETON UNIVERSITY

ALFONSO SORRENTINO

1. Adjoint of a linear operator

Note: In these notes, $V$ will denote a $n$-dimensional euclidean vector space and we will denote the inner product by $\langle , \rangle$.

**Definition.** Let $(V, \langle , \rangle)$ be a $n$-dimensional euclidean vector space and $T : V \rightarrow V$ a linear operator. We will call the adjoint of $T$, the linear operator $S : V \rightarrow V$ such that:

$$\langle T(u), v \rangle = \langle u, S(v) \rangle, \quad \text{for all } u, v \in V.$$  

**Proposition 1.** Let $(V, \langle , \rangle)$ be a $n$-dimensional euclidean vector space and $T : V \rightarrow V$ a linear operator. The adjoint of $T$ exists and is unique. Moreover, if $E$ denotes an orthonormal basis of $V$ (with respect to $\langle , \rangle$) and $T$ has matrix $B$ with respect to $E$ (i.e., $T(E) = EB$), then the adjoint of $T$ is the linear operator $S : V \rightarrow V$, that has matrix $B^T$ with respect to $E$. For this reason, the adjoint of $T$ is sometimes called the “transpose operator” of $T$ and denoted $T^T$.

**Proof.** Let $E$ an arbitrary basis of $V$, let $T(E) = EB$, with $B \in M_n(R)$ and let $A$ be the matrix of $\langle , \rangle$ (w.r.t. $E$). Therefore:

$$\langle u, v \rangle = \langle E\alpha, E\beta \rangle = x^T A y \quad \text{for all } u = E\alpha, v = E\beta \in V$$

[observe that $A \in GL_n(R)$, since $\langle , \rangle$ is an inner product].

Let us denote with $S : V \rightarrow V$ an arbitrary linear operator with matrix $C \in M_n(R)$ (w.r.t. $E$), i.e., $S(E) = EC$. For any $u, v \in V$ we have:

$$T(u), v = (T(E\alpha), E\beta) = (EBx, E\beta) = (Bx)^T Ay = x^T (B^T A) y$$

$$\langle u, S(v) \rangle = \langle E\alpha, S(E\beta) \rangle = \langle E\alpha, EC\beta \rangle = x^T A(C\beta) = x^T (AC)y.$$  

Hence:

$$S \text{ is the adjoint of } T \iff B^T A = AC \iff C = A^{-1} B^T A.$$  

Therefore, the adjoint of $T$ exists and is unique (in fact, its matrix w.r.t. $E$ is uniquely determined by $A$ and $B$). Moreover, if $E$ is orthonormal, then $A = I_n$ and $C = B^T$. \qed

**Definition.** Let $(V, \langle , \rangle)$ be a $n$-dimensional euclidean vector space and $T : V \rightarrow V$ a linear operator. $T$ is said to be self-adjoint [or symmetric] if $T = T^T$; i.e.,

$$\langle T(u), v \rangle = \langle u, T(v) \rangle, \quad \text{for all } u, v \in V.$$  

**Remark.** Let $T : V \rightarrow V$ be a linear operator and $E$ an orthonormal basis of $V$. If $T(E) = EB$, then it follows from the previous proposition that:

$$T \text{ is self-adjoint } \iff B = B^T.$$
Hence, self-adjoint operators on $V$ (with $\dim V = n$) are in $1-1$ correspondence with symmetric matrices in $M_n(\mathbb{R})$. In particular, they form a vector subspace of $\text{End}(V)$, that is isomorphic to the vector subspace of symmetric matrices in $M_n(\mathbb{R})$. Notice that if $E$ is NOT orthonormal, then a self-adjoint operator might not have a symmetric matrix (w.r.t. that basis); viceversa, an operator with a symmetric matrix (w.r.t. a non-orthonormal basis) might not be self-adjoint.

**Example.** Let $V$ be a two-dimensional euclidean vector space, with inner product $\langle \cdot, \cdot \rangle$ defined by

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$

with respect to a fixed basis $E$. Let $T : V \rightarrow V$ a linear operator, defined by $T(E) = EB$, where

$$B = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$  

Let us verify that $T$ is not self-adjoint (although $B$ is symmetric) and find the matrix of its adjoint operator $S$ (w.r.t. $E$). Observe that, if $T$ were self-adjoint, we would have

$$\langle T(Ex), Ey \rangle = \langle Ex, T(Ey) \rangle; \quad \text{for all } Ex, Ey \in V;$$

from which

$$x^T(B^T A)y = x^T(AB)y \quad \text{for all } x, y \in M_{n,1}(\mathbb{R}).$$

Therefore, we should have $B^T A = AB$, but:

$$B^T A = \begin{pmatrix} 0 & 1 \\ 3 & -1 \end{pmatrix} \neq \begin{pmatrix} 0 & 3 \\ 1 & -1 \end{pmatrix}.$$  

The adjoint of $T$ is defined by:

$$S(E) = EC \quad \text{with } C = A^{-1} B^T A = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}.$$  

2. **Spectral theorem for self-adjoint operators**

The main theorem of this section states that if $T$ is a self-adjoint operator on an euclidean vector space, then there exists an orthonormal basis of $V$ formed by eigenvectors of $T$.

**Theorem 1 (Spectral theorem for self-adjoint operators).** Let $V$ be a $n$-dimensional euclidean vector space and $T : V \rightarrow V$ a self-adjoint linear operator. Then, there exists an orthonormal basis of $V$ formed by eigenvectors of $T$.

The proof of this theorem is based on the following preliminary results.

**Proposition 2.** Let $A \in M_n(\mathbb{R})$ be a symmetric matrix. Then, its characteristic polynomial $P = P_A$ has $n$ real roots (each counted with its algebraic multiplicity); hence, $P$ can be factorized in the product of $n$ linear polynomials in $\mathbb{R}[x]$.

**Proof.** From the Fundamental Theorem of Algebra, it follows that $P$ can be always linearly factorized in $\mathbb{C}[x]$. We need only to show that each root $\lambda \in \mathbb{C}$ of $P$ is indeed a real number; i.e., $\lambda \in \mathbb{R}$.

Let $T : \mathbb{C}^n \rightarrow \mathbb{C}^n$ be the linear operator with matrix $A$, with respect to the canonical basis $E$ of $\mathbb{C}^n$. Since $\lambda$ is an eigenvalue of $T$, then there exists a non-zero vector $z$ such that $T(z) = \lambda z$, i.e.,

$$Az = \lambda z.$$
Let us remember some definitions and simple results about the complex numbers. For any $\alpha = x + iy \in \mathbb{C}$, we define its conjugate as $\bar{\alpha} = x - iy$. In particular, the following holds:

$\alpha \bar{\alpha} = x^2 + y^2$ [it is called “norm” of $\alpha$];
$\alpha \bar{\alpha} \geq 0$ and $\alpha \bar{\alpha} = 0 \iff \alpha = 0$;
$\alpha = \bar{\alpha}$;
$\alpha + \beta = \bar{\alpha} + \bar{\beta}$ and $\alpha \beta = \bar{\alpha} \bar{\beta}$.

Let us consider a non-zero vector $z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \in \mathbb{C}^n$ and its conjugate $\bar{z} = \begin{pmatrix} \bar{z}_1 \\ \vdots \\ \bar{z}_n \end{pmatrix} \in \mathbb{C}^n$. One can easily verify that

$$z^T z = \sum_{i=1}^{n} z_i \bar{z}_i > 0 \quad \text{[at least one $z_i$ is different from zero];}$$

$$\overline{(z^T A z)} = z^T \overline{A} z = z^T A \overline{z} \quad \text{[since $A$ is real].}$$

If we multiply (1) by $z^T$ (on the left), we get

$$z^T A z = z^T \lambda z = \lambda (z^T z),$$

and consequently

$$\lambda = \frac{1}{z^T z} (z^T A z).$$

In order to show that $\lambda \in \mathbb{R}$, it suffices to verify that $z^T A z \in \mathbb{R}$ or equivalently:

$$\overline{(z^T A z)} = z^T A z = z^T A \overline{z}.$$

In fact, using that $A$ is symmetric:

$$\overline{(z^T A z)} = z^T A z = (z^T A \overline{z})^T = \overline{(z^T A z)}.$$

\[ \square \]

**Proposition 3.** Let $T$ be a self-adjoint operator on a $n$-dimensional euclidean vector space $V$ and $u$ an eigenvector. Then, $T(u^\perp) \subseteq u^\perp$; i.e., $T$ let the subspace $u^\perp$ fixed.

**Proof.** Let $T(u) = \lambda u$. For any $v \in u^\perp$ [i.e., $\langle v, u \rangle = 0$] we have:

$$\langle T(v), u \rangle = \langle v, T(u) \rangle = \langle v, \lambda u \rangle = \lambda \langle v, u \rangle = 0.$$

Therefore, $T(v) \in u^\perp$. \[ \square \]

We can now prove the Spectral Theorem.

**Proof.** [Spectral Theorem] We proceed by induction on $n = \dim (V)$. If $\dim V = 1$, then the assertion is evident (it is sufficient to choose any basis $E = \{e_1\}$ with $\|e_1\| = 1$). Suppose that $n \geq 2$ and assume that the theorem is true for self-adjoint operators on euclidean vector spaces of dimension $n - 1$.

According to proposition 2, $T$ has at least one real eigenvalue $\lambda$ and call $e_1$ one of the corresponding eigenvectors. We can assume that $\|e_1\| = 1$ (otherwise, we just divide this vector by its norm).
One can observe that \( e_1^\perp \) is an euclidean vector subspace of \( V \) of dimension \( n - 1 \) (see Lecture V, § 2). From proposition 3, \( T(e_1^\perp) \subseteq e_1^\perp \), therefore we can restrict \( T \) to the subspace \( e_1^\perp \). We obtain a new linear operator \( T' : e_1^\perp \rightarrow e_1^\perp \), that is still self-adjoint (since it acts like \( T \) on the vectors in \( e_1^\perp \)).

Using the inductive hypothesis, \( T' \) has an orthonormal basis \( \{e_2, \ldots, e_n\} \), formed by eigenvectors. Since \( e_1 \perp e_i \) (for all \( i = 1, \ldots, n \)), then the vectors \( e_1, \ldots, e_n \) are pairwise orthogonal and therefore linearly independent. These vectors form the desired basis. □

**Remark.** Let \( T \) be a self-adjoint operator and \( F \) an orthonormal basis of eigenvectors of \( T \). We want to point out that, with respect to this basis, \( T \) is represented by a diagonal matrix:

\[
D = \begin{pmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \lambda_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \lambda_n
\end{pmatrix},
\]

where \( \lambda_1, \ldots, \lambda_n \) are the eigenvalues of \( T \). They are not necessarily distinct: each of them appears \( h_{\lambda_i} \) times on the diagonal (where \( h_{\lambda_i} \) is its algebraic multiplicity = geometric multiplicity ).

**Proposition 4.** Let \( T : V \rightarrow V \) be a self-adjoint operator. If \( u, v \) are eigenvectors corresponding to distinct eigenvalues, then they are orthogonal. Therefore, the eigenspaces of \( T \) are pairwise orthogonal.

**Proof.** Let \( T(u) = \lambda u \) and \( T(v) = \mu v \), with \( \lambda, \mu \in \mathbb{R} \) and \( \lambda \neq \mu \). We have:

\[
\langle T(u), v \rangle = \langle \lambda u, v \rangle = \lambda \langle u, v \rangle \quad \text{and} \quad \langle u, T(v) \rangle = \langle u, \mu v \rangle = \mu \langle u, v \rangle.
\]

Since \( \langle T(u), v \rangle = \langle u, T(v) \rangle \), then \( \lambda \langle u, v \rangle = \mu \langle u, v \rangle \) and therefore (since \( \lambda \neq \mu \)) \( \langle u, v \rangle = 0 \).

It follows also that \( E_\lambda \subseteq E_\mu^\perp \) and \( E_\mu \subseteq E_\lambda^\perp \) (where \( E_\lambda \) and \( E_\mu \) are the associated eigenspaces).

**Remark.** From the previous proposition, it follows that in order to compute an orthonormal basis \( F \) of eigenvectors of a self-adjoint operator, it is enough to find the basis of each eigenspace \( E_\lambda \) and orthonormalize it (using Gram-Schmidt, for instance). The union of such bases provides the desired one.

**Example.** In \( \mathbb{R}^4 \) with the canonical inner product, consider the linear operator defined (w.r.t. the canonical basis of \( \mathbb{R}^4 \)) by:

\[
A = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0
\end{pmatrix}.
\]

Determine an orthonormal basis \( F \) of eigenvectors of \( T \) and write the matrix of \( T \) with respect to \( F \).

**Solution:** \( T \) has characteristic polynomial

\[
P = (x - 1)^2(x + 1)^2
\]

and therefore its spectrum is \( \Lambda(T) = \{1, -1\} \), with multiplicities \( h_1 = 2 \) and \( h_{-1} = 2 \).

The eigenspace \( E_1 \) has equations:

\[
\begin{align*}
-2x_2 &= 0 \\
-x_3 + x_4 &= 0
\end{align*}
\]
and therefore: \( E_1 = \langle (1,0,0,0), (0,0,1,1) \rangle \). This basis is already orthogonal, but we need to normalize the vectors, dividing by their norm:
\[
E_1 = \left\langle \left( 1, 0, 0, 0 \right), \left( 0, 0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) \right\rangle.
\]
The eigenspace \( E_{-1} \) has equations:
\[
\begin{align*}
2x_1 &= 0 \\
x_3 + x_4 &= 0
\end{align*}
\]
and therefore: \( E_{-1} = \langle (0,1,0,0), (0,0,1,-1) \rangle \). This basis is already orthogonal, but we need to normalize the vectors, dividing by their norm:
\[
E_{-1} = \left\langle \left( 0, 1, 0, 0 \right), \left( 0, 0, 1, -\frac{1}{\sqrt{2}} \right) \right\rangle.
\]
Concluding, an orthonormal basis \( F \) for \( T \) is given by:
\[
F = EC, \quad \text{where} \quad C = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\
0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}}
\end{pmatrix} \in O_4(\mathbb{R}).
\]
The matrix of \( T \) with respect to this basis is:
\[
T(F) = FD, \quad \text{where} \quad D = C^T AC = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & -1
\end{pmatrix}.
\]

3. Unitary operators

Definition. Let \((V, \langle \cdot, \cdot \rangle)\) be a \(n\)-dimensional Euclidean vector space and \(T : V \rightarrow V\) a linear operator. We will say that is unitary if:
\[
\langle T(u), T(v) \rangle = \langle u, v \rangle, \quad \text{for all } u, v \in V
\]
[i.e., \(T\) preserves the inner product \(\langle \cdot, \cdot \rangle\) in \(V\)].

Proposition 5. Let \(T\) be a linear operator on \((V, \langle \cdot, \cdot \rangle)\). We have:
\[
T \text{ is unitary } \iff \text{ \(T\) is invertible and } T^{-1} = T^T.
\]

Proof. (\(\implies\)) It is sufficient to verify that \(T^T \circ T = \text{Id}\), that is equivalent to
\[
T^T(T(v)) = v \quad \text{for all } v \in V.
\]
In fact, from the definition above and that of adjoint of \(T\), one can conclude:
\[
\langle u, v \rangle = \langle T(u), T(v) \rangle = \langle u, T^T(T(v)) \rangle;
\]
therefore,
\[
\langle u, v - T^T(T(v)) \rangle = 0 \quad \text{for all } u \in V.
\]
We can deduce from this (using the non-degeneracy of the inner product) that \(T^T(T(v)) = v\), for any \(v \in V\).

(\(\impliedby\)) We have that, for any \(u, v \in V\):
\[
\langle T(u), T(v) \rangle = \langle u, T^T(T(v)) \rangle = \langle u, T^{-1}(T(v)) \rangle = \langle u, v \rangle.
\]

\[\square\]
Let us try to deduce some information about the matrices of these unitary operators.

Let $E$ be a basis for $(V, \langle \cdot, \cdot \rangle)$ and suppose that $\langle Ex, Ey \rangle = x^T Ay$. If $T$ is a linear operator on $V$, such that $T(E) = EB$, then:

$T$ is unitary $\iff \langle T(Ex), T(Ey) \rangle = \langle Ex, Ey \rangle$, for all $Ex, Ey \in V$

$\iff \langle EBx, EBy \rangle = \langle Ex, Ey \rangle$, for all $Ex, Ey \in V$

$\iff (Bx)^T A(By) = x^T Ay$, for all $x, y \in M_{n,1}(\mathbb{R})$

$\iff x^T(B^T AB)y = x^T Ay$, for all $x, y \in M_{n,1}(\mathbb{R})$

$\iff B^T AB = A$.

**Corollary 1.** Let $(V, \langle \cdot, \cdot \rangle)$ be a $n$-dimensional euclidean vector space, with an orthonormal basis $E$ and let $T : V \rightarrow V$ a linear operator, such that $T(E) = EB$. Then,

$T$ is unitary $\iff B \in O_n(\mathbb{R})$ [i.e., $B$ is an orthogonal matrix].

**Proof.** For what observed above: $T$ is unitary if and only if $B^T AB = A$. Since $E$ is orthonormal, then $A = I_n$ and we get:

$T$ is unitary $\iff B^T B = I_n$ $\iff B \in O_n(\mathbb{R})$.

\[\square\]

**Remark.**

i) Let $T$ be a unitary operator. If $E$ is an orthonormal basis, then also $T(E)$ is an orthonormal basis (see Lecture V, prop. 8). Therefore, unitary operators send orthonormal bases into orthonormal bases.

ii) If $T$ is a unitary operator, then $\det T = \pm 1$. In fact, if $T(E) = EB$, with $E$ orthonormal, then $B \in O_n(\mathbb{R})$ and $\det T = \det B = \pm 1$.

In particular, unitary operators with $\det T = 1$ are called *special unitary operators* (or *rotations*) of $V$.

iii) If $T$ is unitary, then its spectrum $\Lambda_T \subseteq \{1, -1\}$. In fact, if $T(u) = \lambda u$, then:

$\langle u, u \rangle = \langle T(u), T(u) \rangle = \langle \lambda u, \lambda u \rangle = \lambda^2 \langle u, u \rangle$;

therefore $\lambda^2 = 1$, that implies $\lambda = \pm 1$.

iv) Let us point out that, in general, a unitary linear operator might not be diagonalizable. Consider, for instance, $\mathbb{R}^2$ with the canonical inner product and the operator that is defined (w.r.t. the canonical basis) by

$R_{\frac{\pi}{2}} = \left(\begin{array}{cc}0 & -1 \\1 & 0 \end{array}\right) \in SO_2(\mathbb{R}) \subset O_2(\mathbb{R})$.

This is a unitary operator and it has no eigenvalues, hence it is not diagonalizable.

v) If $T$ is unitary and $\Lambda_T = \{1, -1\}$, then the eigenspaces $E_1$ and $E_{-1}$ are orthogonal to each other. In fact, if $T(u) = u$ and $T(v) = -v$, then:

$\langle u, v \rangle = \langle T(u), T(v) \rangle = \langle u, -v \rangle = -\langle u, v \rangle$;

this implies that $\langle u, v \rangle = 0$ and therefore $(u, v) = 0$.

vi) If $T$ is unitary and $u$ is one of its eigenvectors, then $T(u^\perp) \subseteq u^\perp$. In fact, for any $v \in u^\perp$ (remember that $\lambda = \pm 1$):

$\langle T(v), u \rangle = \frac{1}{\lambda} \langle T(v), \lambda u \rangle = \frac{1}{\lambda} \langle T(v), T(u) \rangle = \frac{1}{\lambda} \langle v, u \rangle = 0$;

hence, $T(v) \in u^\perp$.