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 $T : V \longrightarrow V$ a linear operator. We will call the *adjoint* of T, the linear operator $S : V \longrightarrow V$ 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 $T : V \longrightarrow V$ a linear operator. The adjoint of T exists and is unique.*
Moreover, if $\mathbb{E}$ denotes an orthonormal basis of V (with respect to $\langle\ ,\ \rangle$) and T has matrix B with respect to $\mathbb{E}$ (i.e., $T(\mathbb{E}) = \mathbb{E}B$), then the adjoint of T is the linear operator $S : V \longrightarrow V$, that has matrix B^T with respect to $\mathbb{E}$.
For this reason, the adjoint of T is sometimes called the "transpose operator" of T and denoted T^T.

Proof. Let $\mathbb{E}$ an arbitrary basis of V, let $T(\mathbb{E}) = \mathbb{E}B$, with $B \in \mathcal{M}_n(\mathbb{R})$ and let A be the matrix of $\langle\ ,\ \rangle$ (w.r.t. $\mathbb{E}$). Therefore:

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

[observe that $A \in \mathrm{GL}_n(\mathbb{R})$, since $\langle\ ,\ \rangle$ is an inner product].
Let us denote with $S : V \longrightarrow V$ an arbitrary linear operator with matrix $C \in \mathcal{M}_n(\mathbb{R})$ (w.r.t. $\mathbb{E}$), *i.e.*, $S(\mathbb{E}) = \mathbb{E}C$. For any $u, v \in V$ we have:

$$\begin{aligned}
\langle T(u), v\rangle &= \langle T(\mathbb{E}x), \mathbb{E}y\rangle = \langle \mathbb{E}Bx, \mathbb{E}y\rangle = (Bx)^T Ay = x^T(B^T A)y \\
\langle u, S(v)\rangle &= \langle \mathbb{E}x, S(\mathbb{E}y)\rangle = \langle \mathbb{E}x, \mathbb{E}Cy\rangle = x^T A(Cy) = x^T(AC)y\,.
\end{aligned}$$

Hence:

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

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

Definition. Let $(V, \langle\ ,\ \rangle)$ be a n-dimensional euclidean vector space and $T : V \longrightarrow V$ 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 $T : V \longrightarrow V$ be a linear operator and $\mathbb{E}$ an orthonormal basis of V. If $T(\mathbb{E}) = \mathbb{E}B$, then it follows from the previous proposition that:

$$T \text{ is self-adjoint} \quad \Longleftrightarrow \quad B = B^T\,.$$

Hence, self-adjoint operators on V (with $\dim V = n$) are in $1-1$ correspondence with symmetric matrices in $\mathcal{M}_n(\mathbb{R})$. In particular, they form a vector subspace of $\mathrm{End}(V)$, that is isomorphic to the vector subspace of symmetric matrices in $\mathcal{M}_n(\mathbb{R})$. Notice that if $\mathbb{E}$ 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 $\langle\,,\,\rangle$ defined by

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

with respect to a fixed basis $\mathbb{E}$. Let $T : V \longrightarrow V$ a linear operator, defined by $T(\mathbb{E}) = \mathbb{E}B$, 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. $\mathbb{E}$).
Observe that, if T were self-adjoint, we would have

$$\langle T(\mathbb{E}x), \mathbb{E}y\rangle = \langle \mathbb{E}x, T(\mathbb{E}y)\rangle\,, \quad \text{for all } \mathbb{E}x,\, \mathbb{E}y \in V\,;$$

from which

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

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

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

The adjoint of T is defined by:

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

2. Spectral theorem for self-adjoint operators

The main theorem of this section states that if T is a self-adjoint operator on an eucliden 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 $T : V \longrightarrow V$ 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 $A \in \mathcal{M}_n(\mathbb{R})$ 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 $T : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ be the linear operator with matrix A, with respect to the canonical basis $\mathbb{E}$ of $\mathbb{C}^n$. Since λ is an eigenvalue of T, then there exists a non-zero vector z such that $T(z) = \lambda z$, *i.e.*,

$$Az = \lambda z\,. \tag{1}$$

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 $\overline{\alpha} = x - iy$. In particular, the following holds:

$$\begin{aligned}
&\alpha\overline{\alpha} = x^2 + y^2 \qquad \text{[it is called “norm” of } \alpha\text{]};\\
&\alpha\overline{\alpha} \geq 0 \text{ and } \alpha\overline{\alpha} = 0 \iff \alpha = 0\,;\\
&\alpha = \overline{\overline{\alpha}}\,;\\
&\alpha = \overline{\alpha} \quad \iff \quad \alpha \in \mathbb{R}\,;\\
&\overline{\alpha + \beta} = \overline{\alpha} + \overline{\beta} \quad \text{and} \quad \overline{\alpha\beta} = \overline{\alpha}\,\overline{\beta}\,.
\end{aligned}$$

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

$$\begin{aligned}
&\overline{z}^T z = \sum_{i=1}^n \overline{z_i} z_i > 0 \quad \text{[at least one } z_i \text{ is different from zero]}\,;\\
&\overline{\left(\overline{z}^T A z\right)} = \overline{\overline{z}^T}\,\overline{A}\,\overline{z} = z^T A \overline{z} \quad \text{[since } A \text{ is real]}\,.
\end{aligned}$$

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

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

and consequently

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

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

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

In fact, using that A is symmetric:

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

□

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, $T(v) \in u^\perp$; □

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 $\mathbb{E} = (e_1)$ 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 λ 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 $T' : e_1^\perp \longrightarrow e_1^\perp$, 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 $\mathbb{F}$ 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:

$$D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix},$$

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

Proposition 4. *Let $T : V \longrightarrow V$ 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_λ and E_μ are the associated eigenspaces). □

Remark. From the previous proposition, it follows that in order to compute an orthonormal basis $\mathbb{F}$ of eigenvectors of a self-adjoint operators, it is enough to find the basis of each eigenspace E_λ 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 $\mathbb{F}$ of eigenvectors of T and write the matrix of T with respect to $\mathbb{F}$.

<u>Solution:</u> T has characteristic polynomial

$$P = (x-1)^2(x+1)^2$$

and therefore its specturm is $\Lambda(T) = \{1, -1\}$, with multiplicities $h_1 = 2$ and $h_{-1} = 2$.
The eigenspace E_1 has equations:

$$\begin{cases} -2x_2 = 0 \\ -x_3 + x_4 = 0 \end{cases}$$

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:

$$E_1 = \left\langle (1,0,0,0)\,,\ \left(0,0,\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)\right\rangle .$$

The eigenspace E_{-1} has equations:

$$\begin{cases} 2x_1 = 0 \\ x_3 + x_4 = 0 \end{cases}$$

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:

$$E_{-1} = \left\langle (0,1,0,0)\,,\ \left(0,0,\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\right)\right\rangle .$$

Concluding, an orthonormal basis $\mathbb{F}$ for T is given by:

$$\mathbb{F} = \mathbb{E}C, \quad \text{where } C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \end{pmatrix} \in O_4(\mathbb{R})\,.$$

The matrix of T with respect to this basis is:

$$T(\mathbb{F}) = \mathbb{F}D, \quad \text{where } D = C^T A C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} .$$

3. Unitary operators

Definition. Let $(V, \langle\,,\,\rangle)$ be a n-dimensional euclidean vector space and $T : V \longrightarrow V$ 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 $\langle\,,\,\rangle$ in V].

Proposition 5. *Let T be a linear operator on $(V, \langle\,,\,\rangle)$. We have:*

$$T \textit{ is unitary} \quad \Longleftrightarrow \quad T \textit{ is invertible and } T^{-1} = T^T\,.$$

Proof. $(\Longrightarrow)$ It is sufficient to verify that $T^T \circ T = \mathrm{Id}$, 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:

$$\langle u, v\rangle = \langle T(u), T(v)\rangle = \left\langle u, T^T(T(v))\right\rangle ;$$

therefore,

$$\left\langle u, v - T^T(T(v))\right\rangle = 0 \quad \text{for all } u \in V\,.$$

We can deduce from this (using the non-degeneracy of the inner product) that $T^T(T(v)) = v$, for any $v \in V$.

$(\Longleftarrow)$ We have that, for any $u, v \in V$:

$$\langle T(u), T(v)\rangle = \left\langle u, T^T(T(v))\right\rangle = \left\langle u, T^{-1}(T(v))\right\rangle = \langle u, v\rangle\,.$$

□

Let us try to deduce some information about the matrices of these unitary operators. Let $\mathbb{E}$ be a basis for $(V, \langle\ ,\ \rangle)$ and suppose that $\langle \mathbb{E}x, \mathbb{E}y\rangle = x^T Ay$. If T is a linear operator on V, such that $T(\mathbb{E}) = \mathbb{E}B$, then:

$$\begin{aligned}
T \text{ is unitary} &\iff \langle T(\mathbb{E}x), T(\mathbb{E}y)\rangle = \langle \mathbb{E}x, \mathbb{E}y\rangle\,, \quad \text{for all } \mathbb{E}x, \mathbb{E}y \in V\\
&\iff \langle \mathbb{E}Bx, \mathbb{E}By\rangle = \langle \mathbb{E}x, \mathbb{E}y\rangle\,, \quad \text{for all } \mathbb{E}x, \mathbb{E}y \in V\\
&\iff (Bx)^T A(By) = x^T Ay, \quad \text{for all } x, y \in \mathcal{M}_{n,1}(\mathbb{R})\\
&\iff x^T(B^T AB)y = x^T Ay, \quad \text{for all } x, y \in \mathcal{M}_{n,1}(\mathbb{R})\\
&\iff B^T AB = A\,.
\end{aligned}$$

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

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

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

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

□

Remark.

i) Let T be a unitary operator. If $\mathbb{E}$ is an orthonormal basis, then also $T(\mathbb{E})$ 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 $T(\mathbb{E}) = \mathbb{E}B$, with $\mathbb{E}$ 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
$$R_{\frac{\pi}{2}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathrm{SO}_2(\mathbb{R}) \subset O_2(\mathbb{R})\,.$$
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 $\langle u, v\rangle = 0$.

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

Department of Mathematics, Princeton University
E-mail address: `asorrent@math.princeton.edu`