LECTURE 13

Date of Lecture: March 9, 2020

As always, $\mathbf{K} \in \{\mathbf{R}, \mathbf{C}\}$.

The symbol ☡ is for flagging a cautionary comment or a tricky argument. It occurs in the margins and is Knuth's version of Bourbaki's "dangerous bend symbol".

An n-tuple $(x_1, \dots, x_n)$ of symbols (x_i not necessarily real or complex numbers) will also be written as a column vector when convenient. Thus

$$(x_1, \dots, x_n) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

A map $\boldsymbol{f}$ from a set S to a product set $T_1 \times \cdots \times T_n$ will often be written as an n-tuple $\boldsymbol{f} = (f_1, \dots, f_n)$, with f_i a map from S to T_i, and hence, by the above convention, as a column vector

$$\boldsymbol{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}.$$

(See Remark 2.2.2 of Lecture 5.)

The default norm on Euclidean spaces of the form $\mathbf{R}^n$ is the Euclidean norm $\| \, \|_2$ and we will simply denote it as $\| \, \|$.

☡ Note that $(x_1, \dots, x_n) \neq [x_1 \dots x_n]$. Each side is the transpose of the other.

1. Inner Product spaces

Throughout V is a finite dimensional inner product space over $\mathbf{K}$.

1.1. Orthonormal bases.

Recall that a basis $e_1, \dots, e_n$ of V is said to be orthonormal if $\langle e_i, e_j \rangle = \delta_{ij}$, $i, j \in \{1, \dots, n\}$. The following result shows that the standard inner product on $\mathbf{K}^n$ is natural.

Proposition 1.1.1. *Let $f_1, \dots, f_n$ be an orthonormal basis on V and let $\mathbf{e}_1, \dots, \mathbf{e}_n$ the standard basis on $\mathbf{K}^n$. Let $\Phi \colon V \xrightarrow{\sim} \mathbf{K}^n$ be the $\mathbf{K}$ linear isomorphism given by $f_i \mapsto \mathbf{e}_i$, $i = 1, \dots, n$. Then*

$$\langle v, w \rangle = \langle \Phi v, \Phi w \rangle, \qquad (v, w \in V)$$

where the inner product on $\mathbf{K}^n$ is the usual one.

Proof. If $v = \sum_i a_i f_i$ and $w = \sum_i b_i f_i$, then we have to show that $\langle v, w \rangle = \sum_i a_i \bar{b}_i$. This follows easily from the fact that $\langle f_i, f_j \rangle = \delta_{ij}$ and the fact that inner products are conjugate linear in the second argument. $\square$

1.2. **The adjoint of a linear transformation.** Fix a $\mathbf{K}$ linear transformation $T: V \to V$. According to Problem 4 of Homework 5 there is a unique map $T^*: V \to V$ such that $\langle Tv, w\rangle = \langle v, T^*w\rangle$ for $v, w \in V$.

Definition 1.2.1. Let T and T^* be as above. Then T^* is called the *adjoint* of T.

Lemma 1.2.2. $T^{**} = T$.

Proof. Let $v, w \in V$. Then

$$\langle T^*w, v\rangle = \overline{\langle v, T^*w\rangle} = \overline{\langle Tv, w\rangle} = \langle w, Tv\rangle$$

The assertion follows. □

Proposition 1.2.3. *Let $\mathcal{B} = \{e_i\}_{i=1}^n$ be an ordered orthonormal basis for V and A the matrix of T with respect to $\mathcal{B}$. Then the matrix of T^* with respect $\mathcal{B}$ is $\bar{A}^t$.*

Proof. Let a_{ij} be the $(i,j)^{\text{th}}$ entry of A, so that $Te_j = \sum_{i=1}^n a_{ij}e_i$. Then for $i, j \in \{1, \dots, n\}$,

$$\langle T^*e_j, e_i\rangle = \langle e_j, Te_i\rangle = \langle e_j, \sum_{k=1}^n a_{ki}e_k\rangle = \sum_{k=1}^n \bar{a}_{ki}\langle e_j, e_k\rangle = \sum_{k=1}^n \bar{a}_{ki}\delta_{jk} = \bar{a}_{ji}$$

On the other hand if s_{ij} is the $(i,j)^{\text{th}}$ entry of the matrix of T^* with respect to $\mathcal{B}$ then

$$\langle T^*e_j, e_i\rangle = \langle\sum_{k=1}^n s_{kj}e_k, e_i\rangle = \sum_{k=1}^n s_{kj}\langle e_k, e_i\rangle = s_{ij}.$$

Thus $\bar{a}_{ji} = s_{ij}$ and we are done. □

1.2.4. If A is a matrix over K, then $\bar{A}^t$ is called the *conjugate transpose of* A. Proposition 1.2.3 can be rephrased by saying that if $\mathcal{B}$ is an orthonormal basis of V, then the matrix of T^* with respect to $\mathcal{B}$ is the conjugate transpose of the matrix of T with respect to $\mathcal{B}$.

2. Self-adjoint linear transformations

A $\mathbf{K}$-linear endomorphism on a $\mathbf{K}$ vector space is called a linear *operator*. We fix a $\mathbf{K}$ linear operator

$$T: V \to V$$

on our inner product space V.

2.1. Recall that a square matrix A over $\mathbf{K}$ is called *symmetric* if $A^t = A$ and *hermitian* if $A = \bar{A}^t$. If $\mathbf{K} = \mathbf{R}$, hermitian matrices are the same as symmetric matrices. One has, at least for inner product spaces, a generalisation to $\mathbf{K}$ endomorprhisms.

Definition 2.1.1. T is said to be *self-adjoint* if $T = T^*$. Let T be self-adjoint. T is said to be *symmetric* if $\mathbf{K} = \mathbf{R}$ and *hermitian* if $\mathbf{K} = \mathbf{C}$.

Proposition 2.1.2. *Let $\mathcal{B} = \{e_i\}_{i=1}^n$ be an ordered orthonormal basis for V and A the matrix of T with respect to $\mathcal{B}$.*

(a) *If $\mathbf{K} = \mathbf{R}$ then T is symmetric if and only if A is symmetric.*

(b) *If $\mathbf{K} = \mathbf{C}$ then T is hermitian if and only if A is hermitian.*

Proof. This is a direct consequences of the definitions and Proposition 1.2.3 $\square$

2.2. **Complexification.** We recall again the definition of complexification of a linear endomorphism of $\mathbf{R}^n$ given in §2.1 of Lecture 12. Let $V = \mathbf{R}^n$ and A the $n \times n$ matrix of T with respect to the standard basis on $\mathbf{R}^n$. The *complexification* $T_{\mathbf{C}}$ of T is the $\mathbf{C}$-linear map

$$T_{\mathbf{C}} \colon \mathbf{C}^n \longrightarrow \mathbf{C}^n \tag{2.2.1}$$

given by the formula

$$T_{\mathbf{C}}(\boldsymbol{x}) = A\boldsymbol{x}, \qquad (\boldsymbol{x} \in \mathbf{C}^n) \tag{2.2.2}$$

where the right side of the equality is matrix multiplication. It is obvious that the matrix of $T_{\mathbf{C}}$ with respect the standard basis of $\mathbf{C}^n$ is also A.

Lemma 2.2.3. *If $V = \mathbf{R}^n$ then T is self-adjoint if and only if $T_{\mathbf{C}}$ is self-adjoint.*

Proof. Follows from Proposition 2.1.2 and the fact that a matrix with real entries is hermitian if and only if it is symmetric. $\square$

Remark 2.2.4. Let F be a field, V a finite dimensional vector space over F, and $S \colon V \to V$ an F-linear endomorphism. Then S may or may not have an eigenvector. If F is algebraically closed it does, for the characteristic polynomial has all its roots in F. In particular, if F is algebraically closed, the set of eigenvectors for S is non-empty. On the other hand, if F is not algebraically closed, then the set of eigenvectors for S could be empty, for its characteristic polynomial may have no roots in F. However, if $F = \mathbf{R}$ and S is symmetric, then the situation is positive even though $\mathbf{R}$ is not algebraically closed, as the following theorem shows.

Theorem 2.2.5. *If T is self-adjoint, then all its eigenvalues are real, i.e all the roots of the characteristic polynomial of T are real. In particular, if $\mathbf{K} = \mathbf{R}$ (i.e. T is symmetric), then T has an eigenvector in V.*

Proof. First assume $\mathbf{K} = \mathbf{C}$ so that T is hermitian. Let λ be an eigenvalue of T and v an eigenvector with eigenvalue λ. Then, using the fact that $T = T^*$, we have

$$\lambda\|v\|^2 = \lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Tv, v\rangle = \langle T^*v, v\rangle = \langle v, Tv\rangle = \langle v, \lambda v\rangle = \bar{\lambda}\|v\|^2.$$

Since $\|v\| \neq 0$, this means $\lambda = \bar{\lambda}$. Thus λ is real.

Now suppose $\mathbf{K} = \mathbf{R}$ so that T is symmetric. Pick an orthonormal basis $\{f_i\}_{i=1}^n$ of V, where $n = \dim V$. Let $\Phi \colon V \xrightarrow{\sim} \mathbf{R}^n$ be the $\mathbf{R}$-linear isomorphism given by $f_i \mapsto \mathbf{e}_i$, $i = 1, \ldots, n$, where $\{\mathbf{e}_i\}$ is the standard basis on $\mathbf{R}^n$.

Let $S \colon \mathbf{R}^n \to \mathbf{R}^n$ be the $\mathbf{R}$-linear map given by $S = \Phi \circ T \circ \Phi^{-1}$ (see commutative diagram below).

$$\begin{array}{ccc} V & \xrightarrow{\;T\;} & V \\ {\scriptstyle\Phi}\downarrow\wr & & \wr\downarrow{\scriptstyle\Phi} \\ \mathbf{R}^n & \xrightarrow[\;S\;]{} & \mathbf{R}^n \end{array}$$

Note that the matrix of T with respect to $\{f_i\}$ is the same as the matrix of S with respect to $\{\mathbf{e}_i\}$. Call this common matrix A. Since T is self-adjoint, by Proposition 2.1.2 A is symmetric, and by applying *loc.cit.* once again, S is self-adjoint. By Lemma 2.2.3, $S_{\mathbf{C}}$ is hermitian, and hence by what we proved above (in the first few lines of the proof of this theorem), all its eigenvalues are real. It follows that the roots of the characteristic polynomial of the matrix of $S_{\mathbf{C}}$ with respect to the standard basis on $\mathbf{C}^n$ are all real. Now, clearly, the matrix of $S_{\mathbf{C}}$ (with respect to the standard basis of $\mathbf{C}^n$) is A. Thus the roots of the characteristic polynomial of A, and hence those of T, are all real. □

2.3. **Diagonalisation.** It turns out that self-adjoint operators are diagonalisable. In fact, if T is self-adjoint, one can find an orthonormal basis of V consisting of eigenvectors of T. One consequence is that symmetric matrices A over $\mathbf{R}$ are diagnolisable and $A = \Gamma\Delta\Gamma^{-1}$ where Δ is a diagonal matrix and Γ a so-called *orthogonal matrix*, i.e. Γ is a real matrix and $\Gamma^t = \Gamma^{-1}$. The same statement holds for hermitian matrices, with Γ now a *unitary matrix*, i.e. $\overline{\Gamma}^t = \Gamma^{-1}$. We will prove the (seemingly) stronger statements in the next lecture, but today we prove the essentially equivalent Theorem 2.3.2.

Proposition 2.3.1. *Let T be self-adjoint, λ an eigenvalue of T with eigenspace V_λ, and let $W_\lambda = V_\lambda^\perp$.*

(a) *W_λ is stable under T, i.e. $T(W_\lambda) \subset W_\lambda$.*
(b) *Let $T_\lambda^\perp \colon W_\lambda \to W_\lambda$ be the restriction of T to W_λ. Then $T_\lambda^\perp$ is also self-adjoint.*

Proof. Let $v \in V_\lambda$ and $w \in W_\lambda$. We have

$$0 = \lambda\langle v, w\rangle = \langle \lambda v, w\rangle = \langle Tv, w\rangle = \langle v, T^*w\rangle = \langle v, Tw\rangle,$$

since $T = T^*$. Hence $Tw \in V_\lambda^\perp = W_\lambda$, giving (a).

Next suppose $w, w' \in W_\lambda$. Then $\langle T_\lambda^\perp w, w'\rangle = \langle Tw, w'\rangle = \langle w, T^*w'\rangle = \langle w, Tw'\rangle = \langle w, T_\lambda^\perp w'\rangle$. The last two equalities have used the fact that $T^* = T$ and the fact that $T_\lambda^\perp$ is the restriction of T to W_λ. □

Theorem 2.3.2. *Let T be self-adjoint. Then V has an orthonormal basis consisting of eigenvectors of T. In particular T is diagonalisable.*

Proof. Let λ be a root of the characteristic polynomial of T. Since λ is real (according to Theorem 2.2.5), there is at least one eigenvector of T with eigenvalue λ, even if $\mathbf{K} = \mathbf{R}$, i.e. there is a *non-zero* vector $v \in V$ with $Tv = \lambda v$. If $\dim V = 1$ we are clearly done, since, in that case, V is the eigenspace for λ and any unit vector in V will act as a singleton member of an orthonormal basis for V. We proceed by induction on $\dim V$. So suppose $\dim V = n > 1$ and assume the theorem is true for self-adjoint operators on non-zero inner product spaces over $\mathbf{K}$ of dimension strictly less than n.

As we observed above, T has an eigenvalue λ and let V_λ be the corresponding eigenspace of T. Then $V_\lambda \neq 0$. Pick an orthornormal basis $e_1^\lambda, \dots, e_{d(\lambda)}^\lambda$ of V_λ. Each member of this basis is clearly an eigenvector for T, since all non-zero elements of V_λ are (by definition of V_λ). We know such an orthonormal basis exists by Gram-Schmidt. As in the proof of Proposition 2.3.1, let $W_\lambda = V_\lambda^\perp$. If $W_\lambda = 0$ then $V = V_\lambda$ and we have found an orthonormal basis of eigenvectors for V. So assume $\dim W_\lambda \neq 0$. Continuing with the notations used in the proof of Proposition 2.3.1, the map $T_\lambda^\perp \colon W_\lambda \to W_\lambda$ is self-adjoint. Moreover, $0 < \dim W_\lambda < n$. By our induction hypothesis W_λ has an orthonormal basis $e_1', \dots, e_k'$ $(k = n - d(\lambda))$ consisting of eigenvectors of $T_\lambda^\perp$. It is clear that e_j' are eigenvectors for T, and that $e_1, \dots, e_{d(\lambda)}, e_1' \dots, e_k'$ is an orthonormal basis of V. □

About these notes. These course notes are a reasonably faithful record of the lectures given at the Chennai Mathematical Institute (CMI) in the January-April semester of 2019-20. The course is Analysis II, a core course for first year undergraduates at CMI. For more material related to this course, visit `https://www.cmi.ac.in/~pramath/teaching.html#ANA2`. If you have comments on these notes, or on related course material, please send an email to pramath@cmi.ac.in.