Advanced Linear Algebra Week 3 Day 1

2018/09/24 – Jonathan Hayase, updated by Prof. Weiqing Gu

1 Schur's Theorem

1.1 Statement

Given $A \in M_n$ and $\epsilon > 0$, then there exists a diagonal matrix $\tilde{A} \in M_n$ such that

$$\sum_{1 \le i,j \le n} \left| a_{ij} - \tilde{a}_{ij} \right|^2 < \epsilon$$

1.2 Proof

By Schur's Theorem, we have

$$A = U \begin{bmatrix} \lambda_1 & x & \cdots & x \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & x \\ & & & \lambda_n \end{bmatrix} U^*$$

Key $\tilde{\lambda}_i = \lambda_i + \; i, j \in 1, 2, \ldots, n$

$$\tilde{A} = U \begin{bmatrix} \lambda_1 & x & \cdots & x \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & x \\ & & & \lambda_n \end{bmatrix} U^*$$

$$\operatorname{tr}\left(\left(A - \tilde{A}\right)^T \left(A - \tilde{A}\right)\right) = \operatorname{tr}\left(U^* \begin{bmatrix} \lambda_1 & x & \cdots & x \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & x \\ & & & \lambda_n \end{bmatrix} UU^* \begin{bmatrix} \lambda_1 & x & \cdots & x \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & x \\ & & & \lambda_n \end{bmatrix} U \right)$$
$$= \sum |\lambda - \lambda|^2 = \sum i^2 \eta^2 < \epsilon$$

Let $\eta^2 = \frac{\epsilon}{\sum_{1 \le i \le n} i^2}$ We made $\tilde{A}$ diagonalizable since every eigenvalue is distinct. [1]

[1] Sorry this was kind of a mess. I definitely missed some things.

2 Euclidean Structure

We have $\left(V/\mathbb{F}\ \langle\cdot,\cdot\rangle\right)$ Where
$$\langle\cdot,\cdot\rangle : V \times V \to \mathbb{F}$$

1. Linear on the first fact
2. Conjugate symmetry $\langle x, y\rangle = \overline{\langle y, x\rangle}$
3. $\langle v, v\rangle \geq 0$ and $= 0$ iff $v = 0$

Two key facts:

1. Cauchy-Schwarz inequality
2. There exist an orthonormal basis

A Euclidean space is also called an inner product space.

2.1 Hilbert space

Definition: An inner product space X is called a Hilbert space if every Cauchy sequence converges to a vector.

Definition: $\{x_n\}$ is called a Cauchy sequence if for all $\epsilon > 0$, there exists an N such that whenever we have $m, n > N$, we also have $\|x_m - x_n\| < \epsilon$.

Normal linear operator/map.

Let $\ell(X) = \{T : X \to X \text{ linear map}\}$ where $\dim X = n$.

Q: Is it possible to give a norm to this T?

Let $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ be an orthonormal basis.

Then for all $x \in X$, $x = \sum a_i x_i$.
$$\begin{aligned} \|T\,x\| = \left\|\sum a_i\, T\, x_i\right\| &\leq \sum |a_i| \|T\,x_i\| \\ &\leq M \sum |a_i| && \text{let } m = \max_i \|T\,x_i\| \\ &\leq M \sum \sqrt{|a_i|^2}\sqrt{n} && \text{Cauchy-Schwarz inequality} \\ &= M\|x\|\sqrt{n} \end{aligned}$$

Thus, we have
$$\frac{\|T\,x\|}{\|x\|} \leq M$$
$$\left\|T\,\frac{x}{\|x\|}\right\| \leq M$$

Let $u = \frac{x}{\|x\| \in S^{n-1}}$. Then $\|T\,u\| \leq M$, so there exist a max on compact set S^2.

Definition: $\|T\| = \max\left\{\|T\,u\|\, m \mid \|u\| = 1\right\}$.

In general,
$$\|T\| = \sup_u \left\{\|T\,u\|\, m \mid \|u\| = 1\right\}.$$

3 Riesz Representation Theorem

Q: What does X^* $(= X')$ look like if we put Euclidean structure on $X/\mathbb{F}$? Let $\{x_1, \ldots, x_n\}$ be an orthonormal basis. Then, for any x, we can write $x = a_1 x_1 + \cdots + a_n x_n = \langle x, x_1\rangle x_1 + \cdots + \langle x, x_n\rangle x_n$.[2] Let $\ell \in X^*$ if $x = k_1 x_1 + \cdots + k_n x_n$. Then

$$\ell(x) = k_1 \underbrace{\ell(x_1)}_{a_1} + \cdots k_n \underbrace{\ell(x_n)}_{a_n} = a_1\langle x, x_1\rangle + a_n\langle x, x_n\rangle = \langle x, \bar{a}_1 x_1 + \cdots \bar{a}_n x_n\rangle.$$

Now, let $y = \bar{a}_1 x_1 + \cdots \bar{a}_n x_n$, then $\ell(x) = \langle x, y\rangle$.

Recall: $X^{**} = X$. Fix x to x_0, so then we have $\langle x_0, y\rangle$.

$$\ell_{x_0} = \langle x_0, y\rangle$$

Life analogy: Goats always eat cabbages and spit out numbers. Goats are linear functionals and cabbages are vectors. Now, if we want to view the difference, in life we know that cabbages are passive, but this cabbage gets angry and it wants to eat a goat! Life analogy: Cabbage becomes a poison cabbage and kills the goats! (And the number the cabbage returns is how sick the Goat gets). This is the idea of duality.

Riesz Representation Theorem: For $\ell \in X'$, there exists a unique $y \in X$ such that $\ell(x) = \langle x, y\rangle$.

Suppose we have $\mathbb{B} = \{v_1, \ldots, v_n\}$ a basis of V. Then we have $\mathbb{B}^* = \{\ell_1, \ldots, \ell_n\}$ dual basis of $\mathbb{B}$.

$$\ell_i(v_j) = \delta ij$$

Now V has an inner product, there exists an orthonormal basis $\{u_1, \ldots, u_n\}$ now, $\{\ell_1, \ell_n\} = \{u_1, \ldots, u_n\}$.

$$u_i(u_j) = \delta_{ij} \implies \mathrm{d}x_1\left(\frac{\partial}{\partial x_j}\right) = \delta_{ij}$$

Recall:

$$\operatorname{curl} \boldsymbol{F} = \begin{vmatrix} \hat{u} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix} = V \times \boldsymbol{F}$$

Similarly, the curl is represented by

$$\sum b_1 \frac{\partial}{\partial x_i}$$

4 Self-adjoint Map

Definition $T : X \to X$ is called a self-adjoint map if $T^* = T$. That is, $\langle x, Tx\rangle = \langle Tx, x\rangle$ for all $x \in X$.

Let M be a matrix representation of T with respect to some basis $\mathbb{B}$. For complex $n \times n$ matrix, M is self-adjoint if $M^* = \overline{M}^T = M$. M is also called Hermitian. If M is real, then $M^T = M$.

[2] $\langle\cdot,\cdot\rangle$ is an inner product

5 Adjoints of Transforms

Q: What is exactly T^* (or T')?

Adjoint of a linear map. Let X be an inner product space. Recall: For any x, $\ell(x) = \langle x, y\rangle$ so $\ell \leftrightarrow y$. So we identify $X' \cong X$. Consider a linear transform $T : X \to X$.

The dual of T is defined as $X' \to X'$ we know that $\ell : X \to \mathbb{F}$. If $X \xrightarrow{T} X \xrightarrow{\ell} \mathbb{F}$, then we can define $\ell \circ T \triangleq T'$.

$$\begin{tikzcd}
\ell \arrow[r, "\in"] \arrow[d, leftrightarrow] & X' \arrow[r, "T'"] \arrow[d, "\cong"] & X' \arrow[d, "\cong"] \\
x \arrow[r, "\in"] & X \arrow[r, "T^*"] & X
\end{tikzcd}$$ [3]

Define T' to be $\ell \circ T$.[4]

$X \cong X'$. We know there exists a y corresponding to ℓ, $\ell \leftrightarrow y$. Definition: $T^* : X \to X$ where $y \mapsto z_y$ is called the adjoint of T.

Claim: T^* is a linear map $X \to X$ and $\langle T, y\rangle = \langle x, T^*\, y\rangle$ for all $x, y \in X$. T^* is called the adjoint of T.

To get a more concrete feeling for the adjoint, let's take a look at its matrix representation.

Let $\{x_1, \ldots, x_n\}$ be an orthonormal basis of X. What is $T\, x_j = \sum_{i=1}^{n} m_{ij} x_i$ so we have $T \leftrightarrow \left(m_{ij}\right) = M$. Then $T^*\, x_j = \sum_{i=1}^{n} n_{ij} x_i$ so $T^* \leftrightarrow n_{ij} = N$.

Are are M and N related? Because we have an orthonormal basis, $\langle T\, x_j, x_i\rangle = m_{ij} = \langle x_j, T^*\, x_i\rangle$ and $\langle T^* x_j, x_i\rangle = n_{ij} = \overline{\langle x_i, T^*\, x_j\rangle} = \overline{m_{ij}}$ Thus, $n_{ij} = \overline{m_{ij}}$. So $N = \overline{\left(M^T\right)} = M^*$.

Definition: M^* is called the conjugate transpose of M. Thus, $T \leftrightarrow M$ and $T^* \leftrightarrow M^*$.

Theorem: Let X be a complex inner product space and let T be self-adjoint (that is, $T^* = T$). This means $\langle T\, x, y\rangle = \langle x, T\, y\rangle$. T has real eigenvalues and a set of eigenvectors that form an orthonormal basis for X.

Proof: Suppose $a + bi$ is an eigenvalue of T. Then there exists $x \neq 0$ such that $T\, x = (a + bi)\, x$. We want to show that $b = 0$. Then consider $(T - aI)\, x = ibx$. Now $(T - aI)^* = T^* - aI^* = T - aI$, so $T - aI$ is also self-adjoint.

$$\langle (T - aI)\, x, x\rangle = \langle x, (T - aI)\, x\rangle = \langle x, ibx\rangle = \overline{ib}\langle x, x\rangle = -ib\|x\|^2$$

But we also have

$$\langle (T - aI)\, x, x\rangle = \langle ibx, x\rangle = ib\langle x, x\rangle = ib\|x\|^2$$

Then we have $?ib\left\|x^2\right\| = 0$. If $x \neq 0$, then we must have $b = 0$.

Now, show "eigenvectors leading to distinct eigenvalues are orthogonal." Proof: Say $T\, x = \lambda_i x_i$ with $\lambda_i \in \mathbb{R}$ and $x \neq 0$. $T\, y = \lambda_j y_j$ for $\lambda_j \in \mathbb{R}$ $y \neq 0$.

Suppose T is self-adjoint.

$$\langle T\, x, y\rangle = \langle \lambda_i x, y\rangle = \lambda_i \langle x, y\rangle$$

[3] Having some trouble here w/ the diagram

[4] can't read

$$\langle x, T\,y\rangle = \langle x, \lambda_j y\rangle = \lambda_j \langle x, y\rangle$$

In each eigenspace using Gram-Schmidt[5] to get an orthonormal set. Then, putting them all together gives us an orthonormal space.

Theorem: Let M be a real self-adjoint matrix. (e.g. the Hessian matrix $\mathrm{H} = \nabla^2 f$). Then there exists an orthonormal matrix P such that. $P^* M P = D$ where D is a diagonal matrix and $P^* = P^{-1} = P^T$. Suppose we have $Mx = ax$. Then $x = \Re(x) + i\Im x$, so $M\left(\Re(x) + i\Im(x)\right) = a\Re(x) + ia\Im(x)$. So $M(\Re x) = a\Re x$.

[6]

6 Spectral resolution of a self-adjoint map

Let $T : X \to X$ a self-adjoint map. Now, let $\lambda_1 \leftrightarrow N_1, \ldots, \lambda_k \leftrightarrow N_k$ be the distinct eigenvalues of T, each corresponding to an eigenspace. Then
$$X = N_1 \oplus \cdots \oplus N_k$$
Then for all $x \in X$, $x = x^{(1)} + \cdots + x^{(k)} \in N_i$.

1. For $i = 1, 2, \ldots, k$ define projection $P_i : X \to X$.
$$Id = \sum_{i=1}^{k} P_i$$
This is called the identity resolution.

2. $P_i^2(x) = P_i(x^{(i)}) = x^{(i)} = P^{(i)}(x)$

3. $P_i P_j(x) = P_i\left(x^{(j)}\right) = 0$ for all x. So $P_i P_j = 0$.

4. $\langle x, p_j, y\rangle = \langle \sum_{i=1}^{k} P_i(x), P_j(y)\rangle = \sum_{i=1}^{k} \underbrace{\langle P_i(x), P_j(x)\rangle}_{0} = \langle P_j(x), P_i(x)\rangle$ Want to show that P_j is self-adjoint. But also $\langle P_j)x, y\rangle = \sum)i = 1^k \langle P_j(x), P_i(y)\rangle$. Thus,
$$\langle P_j(x), y\rangle = \langle x, P_j(y)\rangle$$
So P_j is self-adjoint.
$$T\,x = T\left(\sum_{i=1}^{k} P_i(x)\right) = \sum_{i=1}^{k} T(P_i(x)) = \sum_{i=1}^{k} T(x^{(i)}) = \sum_{i=1}^{k} \lambda_i \lambda^{(i)} = \sum_{i=1}^{k} \lambda_i P_i(x)$$

for all $x \in X$. Thus
$$T = \sum_{i=1}^{k} \lambda_i P_i$$
Called the spectral resolution of T.

Every eigenvalue deal with transformations for ???? large data?[7].

[5] spelling?
[6] Missed some things here.
[7] rip

So what?

$$T^n = \sum_{i=1}^{k} \lambda^n P_i$$

$$e^{sT} = \sum_{u=1}^{\infty} \frac{s^n T^n}{n!} = \sum_{n=1}^{\infty} \frac{s^n}{n!} \sum \lambda^i P_i{}^8$$

[8] rip