Matrix Theory, Math6304
Lecture Notes from January 28, 2016

taken by Nikolaos Mitsakos

1.7.12 Lemma. *If $A \in M_n$ has eigenvalue λ , and $p_A(t) = (t-\lambda)^m q(t)$ with $q(\lambda) \neq 0$ then*

$$r = nul(\lambda I - A) \leq m$$

Proof - continued from previous lecture. By the determinant property:

$$\begin{aligned} det[tI - A] &= det[tI - S^{-1}AS] \\ &= det[(t-\lambda)I_r]det[tI_{n-r} - C] \\ &= (t-\lambda)^r det[tI_{n-r} - C] \end{aligned}$$

So, the algebraic multiplicity of λ is at least r. □

1.7.13 Remark. By factorization of p_A,

$$p_A(t) = \prod_{j=1}^{r}(t-\lambda_j)^{m_j}$$

we have

$$\sum_{j=1}^{n} m_j = n$$

and

$$\sum_{j=1}^{n} dim E_{\lambda_j} \leq n$$

Recall: $E_{\lambda_j} = \{x \in \mathbb{C}^n : Ax = \lambda_j x\}$

1.7.14 Theorem. *A matrix $A \in M_n$ is diagonalizable if and only if the geometric and algebraic multiplicities for each eigenvalue are equal.*

Proof. First, we note that if two eigenvalues λ_j,λ_k are unequal ($\lambda_j \neq \lambda_k$), then the eigenspaces E_{λ_j},E_{λ_k} intersect trivially (i.e. $E_{\lambda_j} \bigcap E_{\lambda_k} = \{0\}$).[1] Thus, if $\{v_1, v_2, ..., v_{r_j}\}$ and $\{u_1, u_2, ..., u_{r_k}\}$

[1] Indeed, assume $x \in E_{\lambda_j} \bigcap E_{\lambda_k}$ for two eigenvalues $\lambda_j \neq \lambda_k$. Then $Ax = \lambda_j x$ and $Ax = \lambda_k x$, so $\lambda_j x = \lambda_k x \Longrightarrow (\lambda_j - \lambda_k)x = 0$ thus x=0.

are bases for E_{λ_j} and E_{λ_k}, then $\{v_1, v_2, ..., v_{r_j}, u_1, u_2, ..., u_{r_k}\}$ is a linearly independent set. Iterating this argument, we obtain from E_{λ_1},E_{λ_2},...,E_{λ_r} a basis for the subspace $E_{\lambda_1} + E_{\lambda_2} + ... + E_{\lambda_r}$ of dimension $\sum\limits_{j=1}^{r} dimE_{\lambda_j}$.
Now, assuming that matrix A is diagonalizable, there exists a basis of eigenvectors (by previous Thm) and we get

$$\sum_{j=1}^{r} dimE_{\lambda_j} = n$$

which means that $dimE_{\lambda_j} = m_j$, the algebraic multiplicity of λ_j. (Otherwise, if for some E_{λ_j} we had $dimE_{\lambda_j} < m_j$, this would result to $\sum\limits_{j=1}^{r} dimE_{\lambda_j} < n$, and then there could be no such basis of eigenvectors for $\mathbb{C}^n$).
Conversely, if the algebraic and geometric multiplicities are equal, then $E_{\lambda_1} + E_{\lambda_2} + ... + E_{\lambda_r}$ has dimension n and we obtain a basis of eigenvectors from the union of bases of the eigenspaces $\square$

Next, we investigate diagonalizing multiple matrices.

1.7.15 Definition. Two matrices $A, B \in M_n$ are called simultaneously diagonalizable if there exists some invertible $S \in M_n$ such that $S^{-1}AS = D_1$, $S^{-1}BS = D_2$ where both D_1, D_2 are diagonal.

1.7.16 Remark. It is not hard to see that if $A, B \in M_n$ are simultaneously diagonalizable, then they commute, i.e. $AB = BA$

Proof. Indeed

$$\begin{aligned} AB &= SD_1S^{-1}SD_2S^{-1} \\ &= SD_1D_2S^{-1} \\ &= SD_2D_1S^{-1} \\ &= SD_2S{-}1SD_1S^{-1} = BA \end{aligned}$$

For the third equality, recall that the product of diagonal matrices only involves the products of their diagonal elements, thus $D_1D_2 = D_2D_1$. $\square$

Question: Is the converse true?
Ans. We prepare our answer by proving the following:

1.7.17 Lemma (Lemma 1). *Two matrices $A \in M_n$, $B \in M_m$ are diagonalizable if and only if the matrix $C = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ is diagonalizable.*

Proof. First, assume S_1 diagonalizes A and S_2 diagonalizes B. Then for $S = \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix}$ we get

$$S^{-1} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} S = \begin{bmatrix} S_1^{-1}AS_1 & 0 \\ 0 & S_2^{-1}BS_2 \end{bmatrix}$$

which is diagonal.

Conversely, assume that S diagonalizes C. Write S in the form

$$S = [s_1 s_2 ... s_{n+m}]$$

where

$$s_j = \begin{bmatrix} x_j \\ y_j \end{bmatrix}$$

for some $x_j \in \mathbb{C}^n, y_j \in \mathbb{C}^m$. By the block diagonal structure of C we get that

$$Cs_j = \lambda_j s_j \Longrightarrow$$

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} x_j \\ y_j \end{bmatrix} = \lambda_j s_j$$

which implies that

$$Ax_j = \lambda_j x_j$$

and

$$By_j = \lambda_j y_j$$

Note: The last two results might for a moment lure us into concluding that A and B have the same eigenvalues. However, this is not the case, since x_j, y_j might as well be zero.

Next, write

$$S = \begin{bmatrix} X \\ Y \end{bmatrix}$$

for some $X \in \mathbb{C}^{n \times (n+m)}, Y \in \mathbb{C}^{m \times (n+m)}$ and observe that, for the (row) rank of S, we have:

$$rankS \leq rankX + rankY$$

as well as

$$rankX \leq n, \quad rankY \leq m$$

But, by invertibility of S we have

$$rankS = n + m$$

So, equality must hold, i.e.

$$rankX = n$$

$$rankY = m$$

Thus, matrix $X = [x_1 x_2 ... x_{n+m}]$ has n linearly independent column vectors while $Y = [y_1 y_2 ... y_{n+m}]$ has m linearly independent column vectors.

Going back to the expressions $Ax_j = \lambda_j x_j$ and $By_j = \lambda_j y_j$, this shows that: A has a basis of n eigenvectors and B has a basis of m eigenvectors. We conclude that A and B are (individually) diagonalizable. $\square$

1.7.18 Lemma (Lemma 2). *(Main ADDITION - Nikolaos Mitsakos) Assume $A \in M_n$, diagonalizable, and* $S^{-1}AS = D_1 = \begin{bmatrix} \lambda_{S_1} & & \mathit{0} \\ \mathit{0} & \ddots & \\ & & \lambda_{S_n} \end{bmatrix}$ *Then, for any rearrangement $[\lambda_{T_1}...\lambda_{T_n}]$ of the diagonal elements $[\lambda_{S_1}...\lambda_{S_n}]$ of D_1, there exists some $T \in M_n$, invertible, such that*

$$T^{-1}AT = \begin{bmatrix} \lambda_{T_1} & & \\ & \ddots & \\ & & \lambda_{T_n} \end{bmatrix}$$

Proof. Choose any rearrangement $[\lambda_{T_1}...\lambda_{T_n}]$ of the diagonal elements $[\lambda_{S_1}...\lambda_{S_n}]$ of D_1 and denote $D_2 = \begin{bmatrix} \lambda_{T_1} & & \\ & \ddots & \\ & & \lambda_{T_n} \end{bmatrix}$ Also, let P be the permutation matrix corresponding to the permutation $\binom{\lambda_{S_1}...\lambda_{S_n}}{\lambda_{T_1}...\lambda_{T_n}}$, i.e. $P \begin{bmatrix} \lambda_{S_1} \\ \vdots \\ \lambda_{S_n} \end{bmatrix} = \begin{bmatrix} \lambda_{T_1} \\ \vdots \\ \lambda_{T_n} \end{bmatrix}$ Finally, let's use $\pi(i)$ to denote the position of the $i-th$ element of $[\lambda_{S_1}...\lambda_{S_n}]$ in the permuted vector $[\lambda_{T_1}...\lambda_{T_n}]$.

Recall: P can easily be constructed from the identity matrix $I_n = \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix}$ by simply rearranging its rows, such that $P = \begin{bmatrix} e_{\pi(1)} \\ \vdots \\ e_{\pi(n)} \end{bmatrix}$.

Then

$$PD_1P^{-1} = D_2$$

and

$$D_1 = P^{-1}D_2P$$

(Recall: a permutation matrix is always invertible and $P^{-1} = P^T$).
So, we can write

$$S^{-1}AS = D_1 = P^{-1}D_2P$$

and thus

$$(PS^{-1})A(SP^{-1}) = D_2$$

Note that $(PS^{-1})^{-1} = SP^{-1}$, so the last result can be reformulated as

$$(SP^{-1})^{-1}A(SP^{-1}) = D_2$$

We conclude that $T = PS^{-1}$ is the desired invertible matrix.

□

1.7.19 Theorem. *(Main Result of the section) Let $A, B \in M_n$ be diagonalizable. Then A and B commute if and only if they are simultaneously diagonalizable.*

Proof. We have already seen (first Remark in this Lecture) that simultaneously diagonalizable matrices commute. It remains to show the forward direction of the proof.

Assume that A and B commute. Let $A' = S^{-1}AS$ (A' being diagonal) and define $B' = S^{-1}BS$. By the **previous Lemma - Lemma 2**, we can assume that A' is of the form:

$$A' = \begin{bmatrix} \lambda_1 & & & & & & & & & \\ & \lambda_1 & & & & & & & & \\ & & \ddots & & & & & & & \\ & & & \lambda_1 & & & & & & \\ & & & & \lambda_2 & & & & & \\ & & & & & \ddots & & & & \\ & & & & & & \lambda_2 & & & \\ & & & & & & & \lambda_r & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & \lambda_r \end{bmatrix}$$

(can also be written as $A' = \begin{bmatrix} \gamma_1 & & & \\ & \gamma_2 & & \\ & & \ddots & \\ & & & \gamma_n \end{bmatrix}$ where the entries have been indexed by row).

By assumption, $AB = BA$, so

$$\begin{aligned} A'B' &= S^{-1}ASS^{-1}BS \\ &= S^{-1}ABS \\ &= S^{-1}BAS \\ &= S^{-1}BSS^{-1}AS \\ &= B'A' \end{aligned}$$

Given that $B' = \left[b'_{ij}\right]_{i,j=1}^{n}$ and using the fact that A' is diagonal, we get

$$\gamma_i b'_{i,j} = b'_{i,j}\gamma_j \implies (\gamma_i - \gamma_j)b'_{i,j} = 0 \quad \forall i, j$$

This means that $b'_{i,j} = 0$ when $\gamma_i \neq \gamma_j$, in other words B' is block diagonal: $B' = \begin{bmatrix} B'_1 & & & \\ & B'_2 & & \\ & & \ddots & \\ & & & B'_r \end{bmatrix}$

Now, since B' is diagonalizable, by the previous Lemma (Lemma 1), each of its diagonal blocks B'_i is diagonalizable. Take matrices $T_1, T_2, ..., T_r$ which diagonize the blocks and write:

$$T = \begin{bmatrix} T_1 & & & \\ & T_2 & & \\ & & \ddots & \\ & & & T_r \end{bmatrix}$$

Then $T^{-1}BT = \begin{bmatrix} T_1^{-1}B_1T_1 & & & \\ & T_2^{-1}B_2T_2 & & \\ & & \ddots & \\ & & & T_r^{-1}B_rT_r \end{bmatrix} = \begin{bmatrix} D_1' & & & \\ & D_2' & & \\ & & \ddots & \\ & & & D_r' \end{bmatrix}$

But also $T_i^{-1}T_i = I$. So

$$T^{-1}AT = \begin{bmatrix} T_1^{-1}\lambda_1 IT_1 & & & \\ & T_2^{-1}\lambda_2 IT_2 & & \\ & & \ddots & \\ & & & T_r^{-1}\lambda_r IT_r \end{bmatrix} = A'$$

We conclude that

$$A' = T^{-1}S^{-1}A(ST) = (ST)^{-1}A(ST)$$

and

$$B'' = T^{-1}S^{-1}BST$$

are both diagonal.

$\square$