Matrix Theory, Math6304
Lecture Notes from September 6, 2012

taken by Nathaniel Hammen

Last Time (9/4/12)

Diagonalization: conditions for diagonalization

Eigenvalue Multiplicity: algebraic and geometric multiplicity

1 Further Review

1.1 Warm-up questions

1.1.1 Question. If $A \in M_n$ has only one eigenvalue λ (of multiplicity n) and is diagonalizable, what is A?

Answer. Because A is diagonalizable, there exists an invertible $S \in M_n$ such that $S^{-1}AS$ is diagonal. In fact, because all eigenvalues are λ, we have

$$A = S(S^{-1}AS)S^{-1} = S\lambda IS^{-1} = \lambda SS^{-1} = \lambda I$$

1.1.2 Question. Let $A \in M_n$ and $f(t) = det(I + tA)$. What is $f'(t)$ in terms of A?

Answer. The i, jth entry of $(I + tA)$ is $\delta_{i,j} + ta_{i,j}$, so

$$\begin{aligned} f(t) = det(I + tA) &= \sum_{\sigma \in S_n} sgn(\sigma) \prod_{j=1}^{n} (\delta_{\sigma(j),j} + ta_{\sigma(j),j}) \\ &= \sum_{\sigma \in S_n} sgn(\sigma) \left(\prod_{j=1}^{n} \delta_{\sigma(j),j} + t \sum_{j=1}^{n} a_{\sigma(j),j} \prod_{i \neq j} \delta_{\sigma(i),i} + o(t^2) \right) \end{aligned}$$

where S_n is the set of permutations of n elements and $sgn(\sigma)$ is $+1$ if σ is an even permutation and -1 if σ is an odd permutation. Differentiating $f(t)$ gives

$$\begin{aligned} f'(t) &= \sum_{\sigma \in S_n} sgn(\sigma) \left(\sum_{j=1}^{n} a_{\sigma(j),j} \prod_{i \neq j} \delta_{\sigma(i),i} + o(t) \right) \\ &= \sum_{j=1}^{n} a_{j,j} + o(t) = tr(A) + o(t) \end{aligned}$$

This last equality holds because $\prod_{i\neq j} \delta_{\sigma(i),i}$ is only nonzero when $\sigma(i) = i$ for all $i \neq j$, which only occurs when σ is the identity. When we plug 0 into the equation, we get that $f'(0) = tr(A)$.

1.4 Similarity (cont'd)

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

Proof. We begin by noting that if λ_j and λ_k are eigenvalues of A with $\lambda_j \neq \lambda_k$, then their eigenspaces intersect trivially; that is $E_{\lambda_j} \cap E_{\lambda_k} = \{0\}$. To show this, let $x \in E_{\lambda_j} \cap E_{\lambda_k}$. Then $\lambda_j x = Ax = \lambda_k x$, which implies $x = 0$ because $\lambda_j \neq \lambda_k$. Thus if $\{v_1, v_2, \ldots, v_{m_j}\}$ is a basis for E_{λ_j} and $\{u_1, u_2, \ldots, u_{m_k}\}$ is a basis for E_{λ_k}, then $\{v_1, v_2, \ldots, v_{m_j}\} \cup \{u_1, u_2, \ldots, u_{m_k}\}$ is a linearly independent set, and forms a basis for $E_{\lambda_j} + E_{\lambda_k}$. Inductively iterating this, we obtain a basis for $E_{\lambda_1} + \cdots + E_{\lambda_r}$, forming a subspace of dimension $\sum_{j=1}^{r} dim(E_{\lambda_j})$. If all algebraic and geometric multiplicities are equal, then this space has dimension

$$\sum_{j=1}^{r} dim(E_{\lambda_j}) = \sum_{j=1}^{r} m_j = n.$$

In this case, the subspace equals the entire space, and we have a basis of eigenvectors of A. Thus, A is diagonalizable.
If not all algebraic and geometric multiplicities are equal, then $\exists \lambda_k$ such that $dim(E_{\lambda_k}) < m_k$. Assume A is diagonalizable. Then there exists a set S of n linearly independent eigenvectors of A. For each eigenvalue λ_j, we know that at most $dim(E_{\lambda_j})$ elements in S are eigenvalues corresponding to λ_j. Then

$$n = |S| \leq \sum_{j=1}^{r} dim(E_{\lambda_j}) < \sum_{j=1}^{r} m_j = n$$

which is a contradiction. Thus, A is not diagonalizable. □

1.5 Simultaneous Diagonalization

1.5.28 Definition. Two matrices $A, B \in M_n$ are said to be *simultaneously diagonalizable* if $\exists S \in M_n$ such that S is invertible and both $S^{-1}AS$ and $S^{-1}BS$ are diagonal matrices.

1.5.29 Remark. If $A, B \in M_n$ are simultaneously diagonalizable, then $AB = BA$.

Proof. Because diagonal matrices commute, we have

$$AB = S(S^{-1}AS)(S^{-1}BS)S^{-1} = S(S^{-1}BS)(S^{-1}AS)S^{-1} = BA$$ □

1.5.30 Question. Is the converse true? That is, if $A, B \in M_n$ with A and B diagonalizable and $AB = BA$, are A and B simultaneously diagonalizable?

To answer this, we need a lemma.

1.5.31 Lemma. *Two matrices $A \in M_n$ and $B \in M_m$ are diagonalizable iff* $C = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \in M_{n+m}$ *is diagonalizable.*

Proof. First we assume $A \in M_n$ and $B \in M_m$ are diagonalizable. Then $\exists S_1 \in M_n, S_2 \in M_m$ such that S_1 and S_2 are invertible and $S^{-1}AS$ and $S^{-1}BS$ are diagonal matrices.

$$\text{Let } S = \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix}. \text{ Then } S^{-1} = \begin{bmatrix} S_1^{-1} & 0 \\ 0 & S_2^{-1} \end{bmatrix}$$

and

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

which is a diagonal matrix, so C is diagonalizable.
Conversely, assume that C is diagonalizable. Then $\exists S \in M_{n+m}$ such that S is invertible and $D = S^{-1}CS$ is a diagonal matrix. Write $S = [s_1\, s_2\, \ldots\, s_{n+m}]$ with each $s_j \in \mathbb{C}^{n+m}$. Then, because $SD = CS$, each s_j is an eigenvector for C. Each s_j may then be written as

$$s_j = \begin{bmatrix} x_j \\ y_j \end{bmatrix} \text{ with } x_j \in \mathbb{C}^n \text{ and } y_j \in \mathbb{C}^m$$

Then, the block form of C and the fact that $Cs_j = \lambda_j s_j$ implies that $Ax_j = \lambda_j x_j$ and $By_j = \lambda_j y_j$. If we let $X = [x_1\, x_2\, \ldots\, x_{n+m}]$ and $Y = [y_1\, y_2\, \ldots\, y_{n+m}]$ then $S = \begin{bmatrix} X \\ Y \end{bmatrix}$. The matrix S is invertible, so $rank(S) = n + m$. By the dimensions of X and Y, we also have $rank(X) \leq n$ and $rank(Y) \leq m$. When looking at row rank, we see that

$$n + m = rank(S) \leq rank(X) + rank(Y) \leq n + m$$

so $rank(X) + rank(Y) = n + m$. This can only occur if $rank(X) = n$ and $rank(Y) = m$. Thus X contains n linearly independent columns, each of which is an eigenvector for A, and Y contains m linearly independent columns, each of which is an eigenvector for B. Thus, we have bases of eigenvectors of A and B, and we conclude that A and B are diagonalizable. □

1.5.32 Theorem. *Let $A, B \in M_n$ be diagonalizable. Then $AB = BA$ if and only if A and B are simultaneously diagonalizable.*

Proof. We have already shown that if A and B are simultaneously diagonalizable then $AB = BA$. All that remains to show is the converse. Assume that $AB = BA$. Because A is diagonalizable, $\exists S \in M_n$ such that S is invertible and $D = S^{-1}AS$ is diagonal. We may multiply S by an invertible matrix to permute elements on the diagonal of D, so we may assume without loss of generality that

$$D = \begin{bmatrix} \lambda_1 I_{m_1} & 0 & \ldots & 0 \\ 0 & \lambda_2 I_{m_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_r I_{m_r} \end{bmatrix} \text{ with } \lambda_j \neq \lambda_k \ \ \forall j \neq k$$

where I_m is the $m \times m$ identity and m_j is the multiplicity of λ_j. Because $AB = BA$, we have

$$(S^{-1}AS)(S^{-1}BS) = S^{-1}ABS = S^{-1}BAS = (S^{-1}BS)(S^{-1}AS).$$

If we let $C = (S^{-1}BS)$, then the above gives that $DC = CD$. If we denote $C = [c_{i,j}]_{i,j=1}^n$ and $D = [d_{i,j}]_{i,j=1}^n$, then by the diagonal structure of D we have $d_{i,i}c_{i,j} = c_{i,j}d_{j,j}$. Then $(d_{i,i} - d_{j,j})c_{i,j} = 0$ implies that if $d_{i,i} \neq d_{j,j}$ then $c_{i,j} = 0$. By the block structure of D, this implies that C is block diagonal with blocks of the same size. That is

$$C = \begin{bmatrix} C_1 & 0 & \dots & 0 \\ 0 & C_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & C_r \end{bmatrix} \text{ with } C_j \in M_{m_j} \ \ \forall j$$

Because B is diagonalizable, $\exists R \in M_n$ such that R is invertible and $R^{-1}BR$ is diagonal. Then

$$R^{-1}SCS^{-1}R = R^{-1}SS^{-1}BSS^{-1}R = R^{-1}BR$$

so C is diagonalizable. From the previous lemma, we deduce that each block C_j is diagonalizable. Thus, for each $j, \exists T_j \in M_{m_j}$ such that each T_j is invertible and $T_j^{-1}C_jT_j$ is diagonal.

$$\text{Let } T = \begin{bmatrix} T_1 & 0 & \dots & 0 \\ 0 & T_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & T_r \end{bmatrix}. \text{ Then } T^{-1} = \begin{bmatrix} T_1^{-1} & 0 & \dots & 0 \\ 0 & T_2^{-1} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & T_r^{-1} \end{bmatrix}$$

and

$$T^{-1}S^{-1}BST = T^{-1}CT = \begin{bmatrix} T_1^{-1}C_1T_1 & 0 & \dots & 0 \\ 0 & T_2^{-1}C_2T_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & T_r^{-1}C_rT_r \end{bmatrix}$$

which is a diagonal matrix. Also

$$\begin{aligned} T^{-1}S^{-1}AST &= T^{-1}DT \\ &= \begin{bmatrix} T_1^{-1}\lambda_1 I_{m_1}T_1 & 0 & \dots & 0 \\ 0 & T_2^{-1}\lambda_2 I_{m_2}T_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & T_r^{-1}\lambda_r I_{m_r}T_r \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 I_{m_1} & 0 & \dots & 0 \\ 0 & \lambda_2 I_{m_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_r I_{m_r} \end{bmatrix} \\ &= D. \end{aligned}$$

Thus, $T^{-1}S^{-1}AST$ and $T^{-1}S^{-1}BST$ are both diagonal matrices, so A and B are simultaneously diagonalizable by $ST \in M_n$. $\square$

We want to make this equivalence more general.

1.5.33 Remark. Let $A \in M_n$ and consider polynomials

$$p(A) = p_0 I + p_1 A + \cdots + p_r A^r \text{ and}$$
$$q(A) = q_0 I + q_1 A + \cdots + q_s A^s \text{ with } r, s \in \mathbb{N}.$$

Then $p(A)q(A) = q(A)p(A)$. Thus the family of all polynomials of A is commuting. If $S \in M_n$ is invertible, then $S^{-1}A^k S = (S^{-1}AS)^k$ for any integer value of k, so

$$\begin{aligned} S^{-1}p(A)S &= S^{-1}(p_0 I + p_1 A + \cdots + p_r A^r)S \\ &= p_0 S^{-1}IS + p_1 S^{-1}AS + \cdots + p_r S^{-1}A^r S \\ &= p_0 I + p_1 S^{-1}AS + \cdots + p_r (S^{-1}AS)^r \\ &= p(S^{-1}AS) \end{aligned}$$

Thus, if S diagonalizes A, then $p(S^{-1}AS)$ is a polynomial of a diagonal matrix, which is diagonal. Thus, if S diagonalizes A, then $S^{-1}p(A)S$ is diagonal, so S diagonalizes all polynomials of A simultaneously.