MATRIX EXPONENTIAL AND NORMAL FORMS

1. Upper Triangular Form

Consider two $n \times n$ complex matrices A and B. We say that A and B are **similar** (and write $A \sim B$) if there is an invertible matrix C such that $A = C^{-1}BC$. An $n \times n$ matrix C is called **unitary** if $C^{-1} = C^*$, or, equivalently, if the columns of C form an orthonormal basis of C^n. The matrices A and B are called **unitarily equivalent** if there is a unitary matrix C such that $A = C^{-1}BC$.

Note that similarily is an equivalence relation. (That is: $A \sim A$; $A \sim B$ implies $B \sim A$; and $A \sim B$ and $B \sim C$ imply that $A \sim C$.) Likewise, unitary equivalence is an equivalence relation.

Lemma 1. *Let A be an $n \times n$ matrix of the form*

$$A = \begin{bmatrix} a & b \\ 0 & B \end{bmatrix}$$

where $a \in \mathbf{C}$, b is $1 \times (n-1)$ matrix, 0 is the 0 vector in $\mathbf{C}^{n-1}$, and B is an $(n-1) \times (n-1)$ matrix. Then

$$\det A = a \det B.$$

Proof. If $a = 0$, then $\mathbf{e}_1$ is the kernel of A so $\det A = 0 = 0 \det B$.

If $a \neq 0$, then we can add multiples of the first column of A to the subsequent columns to get the matrix

$$M = \begin{bmatrix} a & 0 \\ 0 & B \end{bmatrix}.$$

Then $\det A = \det M = a \det B$. ☐

Corollary 1. $\det(\lambda I - A) = (\lambda - a)\det(\lambda I - B)$.

Remark 2. More generally, suppose

$$A = \begin{bmatrix} P & Q \\ 0 & R \end{bmatrix},$$

where P is a $k \times k$ matrix, Q is a $k \times (n-k)$ matrix, 0 is the zero $(n-k) \times k$ matrix, and R is an $(n-k) \times (n-k)$ matrix. Then $\det A = \det P \det Q$. We won't need this more general fact. (But it is not hard to prove.)

Theorem 3. *Let A be an $n \times n$ complex matrix. Let $\lambda_1, \dots, \lambda_n$ be the roots of the characterstic polynomial. (That is, $\det(\lambda I - A) = \prod_{i=1}^n (\lambda - \lambda_i)$.) Then A is unitarily equivalent to an upper triangular matrix M with $M_{ii} = \lambda_i$ for each i.*

Thus, for example, if $n = 3$, then A is unitarily equivalent to a matrix of the form

$$\begin{bmatrix} \lambda_1 & * & * \\ 0 & \lambda_2 & * \\ 0 & 0 & \lambda_3 \end{bmatrix}.$$

Proof. Let $\mathbf{v}_1$ be a unit eigenvector with eigenvalue λ_1. Extend $\mathbf{v}_1$ to an orthonormal basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ of $\mathbf{C}^n$. (The $\mathbf{v}_2, \dots, \mathbf{v}_n$ need not be eigenvectors.)

Let S be the unitary matrix whose columns are $\mathbf{v}_1, \dots, \mathbf{v}_n$. Then

$$S^{-1}AS = S^*AS = \begin{bmatrix} \lambda_1 & b \\ 0 & B \end{bmatrix},$$

where r is a $1 \times (n-1)$ matrix, 0 is the zero $(n-1) \times 1$ vector, and B is an $(n-1) \times (n-1)$ matrix.

By the Lemma, $\det(\lambda I - B) = \prod_{i=2}^{n} (\lambda - \lambda_i)$.

By induction, there is a unitary $(n-1) \times (n-1)$ matrix C such that $C^{-1}BC$ is an upper triangular matrix T with entries $\lambda_2, \dots, \lambda_n$ along the diagonal:

$$T = \begin{bmatrix} \lambda_2 & * & * & \cdots & * \\ 0 & \lambda_3 & * & \cdots & * \\ 0 & 0 & \lambda_4 & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & * \\ 0 & 0 & 0 & 0 & \lambda_n \end{bmatrix}.$$

Then the matrix

$$R = \begin{bmatrix} 1 & 0 \\ 0 & C \end{bmatrix}$$

is unitary and

$$\begin{aligned} R^{-1} \begin{bmatrix} \lambda_1 & b \\ 0 & B \end{bmatrix} R &= \begin{bmatrix} 1 & 0 \\ 0 & C^{-1} \end{bmatrix} \begin{bmatrix} \lambda_1 & b \\ 0 & B \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & C \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & C^{-1} \end{bmatrix} \begin{bmatrix} \lambda_1 & bC \\ 0 & BC \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & bC \\ 0 & C^{-1}BC \end{bmatrix} \\ (*) \qquad &= \begin{bmatrix} \lambda_1 & bC \\ 0 & T \end{bmatrix} \end{aligned}$$

We have shown that A is unitarily equivalent to $\begin{bmatrix} \lambda_1 & b \\ 0 & B \end{bmatrix}$, which in turn is unitarily equivalent to $\begin{bmatrix} \lambda_1 & bC \\ 0 & T \end{bmatrix}$. Thus A is unitarily equivalent to $\begin{bmatrix} \lambda_1 & bC \\ 0 & T \end{bmatrix}$, which has the required form. □

Equal Eigenvalues

Theorem 4. *Suppose A is an $n \times n$ matrix whose eigenvalues (i.e., the roots of the characteristic polynomial) are all equal to λ. Then $(A - \lambda I)^n = 0$.*

Proof. By Theorem 3, $A = S^{-1}MS$ for some upper triangular matrix M with $m_{ii} = \lambda$ for all i. That is, the matrix

$$U = M - \lambda I$$

is upper triangular and each of its diagonal elements U_{ii} is 0. Thus

(*) $$U_{ij} = 0 \text{ unless } j > i. \quad .$$

Claim: $U^n = 0$.

Note that

$$\begin{aligned}(U^2)_{ij} &= \sum_k U_{ik}U_{kj} \\ &= \sum_{k=i+1}^{j-1} U_{ik}U_{kj}\end{aligned}$$

The sum is 0 unless $i+1 \le j-1$, i.e., unless $j > i+2$, so

$$(U^2)_{ij} = 0 \text{ Unless } j > i+2.$$

Proceeding by induction, we see that $(U^p)_{ij} = 0$ unless $j > i+p$. Therefore $U^n = 0$, proving the claim.

Now

$$\begin{aligned}A - \lambda I &= S^{-1}MS - \lambda I \\ &= S^{-1}(M - \lambda I)S \\ &= S^{-1}US\end{aligned}$$

so

$$(A - \lambda I)^n = S^{-1}U^nS = 0.$$

□

Corollary 5. *If A is an $n \times n$ matrix whose only eigenvalue is λ, then*

$$e^{At} = e^{\lambda t} p_{n-1}(tN)$$

where $N = A - \lambda I$ and where $p_{n-1}(z) = \sum_{j=0}^{n-1} \frac{z^k}{k}$ is the degree $(n-1)$ Taylor polynomial for e^z.

Proof. Note that $\lambda t I$ commutes with every matrix. In particular, it commutes with tN. Thus

$$\begin{aligned}e^{At} &= e^{\lambda t I + tN} \\ &= e^{\lambda t I} e^{tN} \\ &= e^{\lambda t} I \sum_{k=0}^{\infty} \frac{1}{k} t^k N^k \\ &= e^{\lambda t} p_{n-1}(tN)\end{aligned}$$

since $N^n = 0$ and therefore $N^k = 0$ for all $k > n$. □

2. General Matrices

We proved in class:

Theorem 6. *Let A be an $n \times n$ matrix with characteristic polynomial*

$$\det(\lambda I - A) = \prod_{i=1}^{k} (\lambda - \lambda_i)^{\nu_i}$$

where $\lambda_1, \dots, \lambda_k$ are all distinct. Then $A = S^{-1}MS$ for some invertible matrix S and for a block diagonal matrix

$$M = \begin{bmatrix} M_1 & 0 & \dots & 0 \\ 0 & M_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & M_k \end{bmatrix}$$

$$M = \begin{bmatrix} M_1 & 0 & 0 & \dots & 0 \\ 0 & M_2 & 0 & \dots & 0 \\ 0 & 0 & M_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & M_k \end{bmatrix}$$

where each M_i is a $\nu_i \times \nu_i$ upper triangular matrix with diagonal elements all equal to λ_i.

Consequently,

$$e^{Mt} = \begin{bmatrix} e^{\lambda_1 t} p_{\nu_1 - 1}(tN_1) & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} p_{\nu_2 - 1}(tN_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{\lambda_k t} p_{\nu_k - 1}(tN_k). \end{bmatrix}$$

Of course this gives a formula for e^{At} since $e^{At} = S^{-1}e^{Mt}S$.