Linear Analysis
Lecture 12

Jordan Form

Let $T \in \mathbb{C}^{n\times n}$ be an upper triangular matrix in block diagonal form

$$T = \begin{bmatrix} T_1 & & 0 \\ & \ddots & \\ 0 & & T_k \end{bmatrix},$$

with $T_i \in \mathbb{C}^{m_i\times m_i}$ satisfying $T_i = \lambda_i I + N_i$, where $N_i \in \mathbb{C}^{m_i\times m_i}$ is strictly upper triangular, and $\lambda_1, \dots, \lambda_k$ are distinct. Then for $1 \le i \le k$, $N_i^{m_i} = 0$, so N is nilpotent. Recall that any nilpotent operator is a direct sum of shift operators in an appropriate basis. Therefore, the matrix N_i is similar to a direct sum of shift matrices

$$S_\ell = \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & 1 \\ 0 & & & 0 \end{bmatrix} \in \mathbb{C}^{\ell\times\ell}$$

of varying sizes ℓ. Thus each T_i is similar to a direct sum of **Jordan blocks**

$$J_\ell(\lambda) = \lambda I_\ell + S_\ell = \begin{bmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix} \in \mathbb{C}^{\ell\times\ell}$$

of varying sizes ℓ (with $\lambda = \lambda_i$).

Jordan Form

Definition. A matrix J is in **Jordan normal form** if it is the direct sum of finitely many Jordan blocks (with, of course, possibly different values of λ and ℓ).

Theorem. Every matrix $A \in \mathbb{C}^{n \times n}$ is similar to a matrix in Jordan normal form.

The Jordan form of A is not quite unique since the blocks may be arbitrarily reordered by a similarity transformation.

Proposition
(a) The Jordan form of A is unique up to reordering of the Jordan blocks.
(b) Two matrices in Jordan form are similar iff they can be obtained from each other by reordering the blocks.
(c) Two matrices are similar iff they are similar to the same Jordan normal form.

Jordan Form

In general, knowing the algebraic and geometric multiplicities of each eigenvalue of A is not sufficient to determine the Jordan form.

For example,

$$N_1 = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & 0 \end{bmatrix} \quad \text{and} \quad N_2 = \begin{bmatrix} 0 & 1 & & \\ 0 & 0 & & \\ & & 0 & 1 \\ & & 0 & 0 \end{bmatrix}$$

are not similar as $N_1^2 \neq 0 = N_2^2$, but both have 0 as the only eigenvalue with algebraic multiplicity 4 and geometric multiplicity 2.

Spectral Decomposition

An invariant formulation of Jordan Normal Form (i.e., basis-free).
Let $L \in \mathcal{L}(V)$ where $\dim V = n < \infty$ (and $\mathbb{F} = \mathbb{C}$). $\lambda_1, \ldots, \lambda_k$ distinct eigenvalues of L, with algebraic multiplicities $m_1, \ldots, m_k$. The **generalized eigenspaces** are

$$\widetilde{E}_i = \mathcal{N}(L - \lambda_i I)^{m_i}.$$

The eigenspaces are

$$E_{\lambda_i} = \mathcal{N}(L - \lambda_i I).$$

Then

$$\dim \widetilde{E}_i = m_i \ (1 \le i \le k) \quad \text{and} \quad V = \bigoplus_{i=1}^{k} \widetilde{E}_i$$

using basis representing L as a block-diagonal upper triangular matrix. If P_i $(1 \le i \le k)$ be the associated projections, then

$$I = \sum_{i=1}^{k} P_i \qquad \text{and} \quad P_i P_j = \delta_{ij} P_i.$$

Define

$$D = \sum_{i=1}^{k} \lambda_i P_i.$$

In the same basis, D is diagonal and the matrix for $N \equiv L - D$ is strictly upper triangular, so N is nilpotent, with $P_j N = N P_j \quad \forall\, j$, and so $N = \sum_{i=1}^{k} N_i$ where $N_i = P_i N P_i$.

Spectral Decomposition Theorem

Any $L \in \mathcal{L}(V)$ can be written as $L = D + N$ where D is diagonalizable, N is nilpotent, and $DN = ND$.
If P_i is the projection onto the λ_i-generalized eigenspace and $N_i = P_i N P_i$, then

$$D = \sum_{i=1}^{k} \lambda_i P_i \quad \text{and} \quad N = \sum_{i=1}^{k} N_i.$$

Moreover,

$$LP_i = P_i L = P_i L P_i = \lambda_i P_i + N_i \quad (1 \le i \le k),$$

$$P_i P_j = \delta_{ij} P_i, \quad P_i N_j = N_j P_i = \delta_{ij} N_j \quad (1 \le i \le k)(1 \le j \le k), \quad \text{and}$$

$$N_i N_j = N_j N_i = 0 \quad (1 \le i < j \le k),$$

where $\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ij} = 1$ if $i = j$.

Note: D and N are uniquely determined by L.

Spectral Decomposition Theorem for Normal Operators

If V has an inner product $\langle \cdot, \cdot \rangle$, and L is normal, then we know that L is diagonalizable, so $N = 0$. We also know that eigenvectors corresponding to different eigenvalues are orthogonal, so the subspaces $\widetilde{E}_i$ ($= E_{\lambda_i}$ here) are mutually orthogonal. Hence, the associated projections P_i are orthogonal projections.

Recall that P is called an **orthogonal projection** if $\mathcal{R}(P) \perp \mathcal{N}(P)$.

Proposition. A projection P is orthogonal iff it is self-adjoint (i.e., P is Hermitian: $P^* = P$, where P^* is the adjoint of P with respect to the inner product $\langle \cdot, \cdot \rangle$).

Proof: Let $P \in \mathcal{L}(V)$ be a projection. If $P^* = P$, then $\langle Px, y \rangle = \langle x, Py \rangle \;\forall\, x, y \in V$, so

$$y \in \mathcal{N}(P) \Leftrightarrow (\forall\, x \in V)\; \langle Px, y \rangle = \langle x, Py \rangle = 0 \Leftrightarrow y \in \mathcal{R}(P)^{\perp},$$

so P is an orthogonal projection.

Conversely, suppose $\mathcal{R}(P) \perp \mathcal{N}(P)$. We must show that

$$\langle Px, y \rangle = \langle x, Py \rangle \quad \forall\; x, y \in V.$$

Since $V = \mathcal{R}(P) \oplus \mathcal{N}(P)$, it suffices to check this separately in the four cases $x, y \in \mathcal{R}(P), \mathcal{N}(P)$. Each of these cases is straightforward since $Pv = v$ for $v \in \mathcal{R}(P)$ and $Pv = 0$ for $v \in \mathcal{N}(P)$. $\square$

Jordan Form over $\mathbb{R}$

In general, a real matrix is not similar to a real upper-triangular matrix via a real similarity transformation.

If it were, then its eigenvalues would be the real diagonal entries, but a real matrix need not have only real eigenvalues.

However, non-real eigenvalues are the only obstruction to carrying out our previous arguments.

Every real eigenvalue of a real matrix can be identified using only real eigenvectors (Why?).

If $A \in \mathbb{R}^{n \times n}$ has real eigenvalues, then A is orthogonally similar to a real upper triangular matrix, and A can be put into block diagonal an Jordan form using real similarity transformations, by following the same arguments as before.

If A does have some non-real eigenvalues, then there are substitute normal forms which can be obtained via real similarity transformations.

Jordan Form over $\mathbb{R}$

Let $p_A(t)$ be the characteristic polynomial of $A \in \mathbb{R}^{n\times n}$. Then $p_A(t)$ has real coefficients. If $\lambda = a + ib \in \sigma(A)$, then

$$0 = \overline{p_A(\lambda)} = p_A(\bar{\lambda}),$$

so $\bar{\lambda} = a - ib \in \sigma(A)$. That is, the non-real eigenvalues of A come in complex conjugate pairs.

If $u + iv$ (with $u, v \in \mathbb{R}^n$) is an eigenvector of A for λ, then

$$A(u - iv) = A\overline{(u + iv)} = \overline{A(u + iv)} = \overline{\lambda(u + iv)} = \bar{\lambda}(u - iv),$$

so $u - iv$ is an eigenvector of A for $\bar{\lambda}$. It follows that $u + iv$ and $u - iv$ are linearly independent over $\mathbb{C}$. Thus

$$u = \frac{1}{2}(u + iv) + \frac{1}{2}(u - iv) \quad \text{and} \quad v = \frac{1}{2i}(u + iv) - \frac{1}{2i}(u - iv)$$

are linearly independent over $\mathbb{C}$, and consequently also over $\mathbb{R}$. Since

$$A(u + iv) = (a + ib)(u + iv) = (au - bv) + i(bu + av),$$

$$Au = au - bv \quad \text{and} \quad Av = bu + av.$$

Thus, $\mathrm{Span}\{u, v\}$ is a 2-dimensional real invariant subspace of $\mathbb{R}^n$ for A, and the matrix of A restricted to the subspaces $\mathrm{Span}\{u, v\}$ with respect to the basis $\{u, v\}$ is

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix}.$$

This 2×2 matrix has eigenvalues $\lambda, \bar{\lambda}$.

Jordan Form over $\mathbb{R}$

Over $\mathbb{R}$, the best one can do is to have such 2×2 diagonal blocks instead of upper triangular matrices with $\lambda, \bar{\lambda}$ on the diagonal. The real Jordan blocks for $\lambda, \bar{\lambda}$ are

$$J_\ell(\lambda, \bar{\lambda}) = \begin{bmatrix} \begin{bmatrix} a & b \\ -b & a \end{bmatrix} & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & & 0 \\ & \ddots & \ddots & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ 0 & & & \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \end{bmatrix} \in \mathbb{R}^{2\ell \times 2\ell}.$$

The real Jordan form of $A \in \mathbb{R}^{n \times n}$ is a direct sum of such blocks, with the usual Jordan blocks for the real eigenvalues.

Eigenvalues of Products

Proposition. Let $A \in \mathbb{C}^{m \times n}$ and $B \in \mathbb{C}^{n \times m}$ with $m \leq n$.

Then the eigenvalues of BA (counting multiplicity) are the eigenvalues of AB, together with $n - m$ zeroes.

Corollary. Let $A \in \mathbb{C}^{m \times n}$ with $m \leq n$. The the eigenvalues of $A^{\mathrm{H}} A$ and AA^{H} differ by $|n - m|$ zeroes.

Singular Values of Non-Square Matrices

Definition. Let $p = \min(m, n)$ and let

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \ \geq 0$$

be the joint eigenvalues of $A^{\mathrm{H}}A$ and AA^{H}. The **singular values** of A are the numbers

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0,$$

where $\sigma_i = \sqrt{\lambda_i}$.

It is a fundamental result that one can choose orthonormal bases for $\mathbb{C}^n$ and $\mathbb{C}^m$ so that A maps one into the other, scaled by the singular values.

In what follows let $A \in \mathbb{C}^{m\times m}$, with $\|A\|$ denoting the operator norm induced by the Euclidean norm, and $\|A\|_F$ denotes the Frobenius norm of A. As usual, the inner product on $\mathbb{C}^{m\times m}$ is

$$\langle Ax, y\rangle_{\mathbb{C}^m} = y^H Ax = \langle x, A^H y\rangle_{\mathbb{C}^n} \quad \text{for} \quad x \in \mathbb{C}^n,\ y \in \mathbb{C}^m.$$

Singular Value Decomposition (SVD)

If $A \in \mathbb{C}^{m \times n}$, then there exists unitary matrices

$$U \in \mathbb{C}^{m \times m} \quad \text{and} \quad V \in \mathbb{C}^{n \times n}$$

such that

$$A = U \Sigma V^H,$$

where $\Sigma \in \mathbb{C}^{m \times n}$ is the diagonal matrix of singular values.
In particular, if

$$\sigma_1 \geq \sigma_2 \geq \dots \sigma_p \qquad (p = \min(m, n))$$

are the singular values of A with

$$\operatorname{diag}(\Sigma) = [\sigma_1, \sigma_2, \sigma_3, \dots]$$

and

$$U = [u_1,\ u_2, \dots,\ u_m] \quad \text{and} \quad V = [v_1,\ v_2, \dots,\ v_n]\ ,$$

then

$$\sigma_j u_j = A v_j \quad j = 1, 2, \dots, p\ .$$