Mostafa Sabri

Lecture 27

Extra: Jordan Form

جامعة نيويورك أبوظبي
NYU | ABU DHABI

Jordan Block

Let A be an $n \times n$ matrix. We know that if A is not symmetric, it may not be diagonalizable. We also know that in general, the lack of diagonalization is due to a "defect" in the matrix, where the algebraic and geometric multiplicities of some eigenvalue differ. An example of this is the matrix $J = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}$. As it is triangular, we easily see it has only the eigenvalue λ of algebraic multiplicity 4. When computing the eigenspace E_λ however, we see immediately that $t = z = y = 0$, so that $E_\lambda = \text{Span}\{(1, 0, 0, 0)\}$ and the geometric multiplicity is only 1.

Jordan Blocks

This works for any size of J that has this special triangular shape. We call an $r_i \times r_i$ matrix

$$J_i = \begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{bmatrix} \tag{1}$$

a *Jordan Block*. It is badly non-diagonalizable if $r_i > 1$. Here the convention is that a 1×1 Jordan block is the matrix $J = \begin{bmatrix} \lambda \end{bmatrix}$.

It turns out that these Jordan blocks are the building blocks of any square matrix.

Jordan Form

Jordan Form Thm: Any $n \times n$ matrix A on $\mathbb{C}^n$ is similar to a block diagonal matrix of the form

$$J = \begin{bmatrix} J_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & J_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & J_p \end{bmatrix}$$

where each J_i is an $r_i \times r_i$ Jordan Block of the form (1) with $\lambda_i \in \mathbb{C}$, and $\sum_{i=1}^{p} r_i = n$.

In the special case where $r_i = 1 \ \forall i$, the matrix J is actually diagonal and A is then diagonalizable.

This result has many applications. One immediate

Consequences of the Jordan Form

advantage over the SVD is that it actually captures the spectrum of A.

> Cor: If A is similar to the matrix J in the previous theorem, then the spectrum of A consists of the eigenvalues $\{\lambda_1, \ldots, \lambda_p\}$, where each λ_i has algebraic multiplicity r_i.

Proof. Since J is triangular, its eigenvalues are the elements on its diagonal, namely $\lambda_1, \lambda_1, \ldots, \lambda_2, \ldots, \lambda_p$. As A is similar to J, they have the same eigenvalues. $\square$

Consequences of the Jordan Form

Application: If A is an $n \times n$ matrix with eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$ counting multiplicity, and $q(x) = \sum_{k=1}^{m} c_k x^k$ is any polynomial, define the $n \times n$ matrix $q(A) = \sum_{i=1}^{m} c_k A^k$. Then the eigenvalues of $q(A)$ are $\{q(\lambda_1), \ldots, q(\lambda_n)\}$.

e.g. if A has spectrum $\{1, 1, 3\}$, then the spectrum of $q(A) = 2A^4 + 3A - 1$ is $\{q(1), q(1), q(3)\} = \{4, 4, 170\}$.

Proof. Let $J = P^{-1}AP$ be the Jordan form of A. Then $A = PJP^{-1}$, so $q(A) = \sum_{k=1}^{m} c_k A^k = \sum_{k=1}^{m} c_k (PJP^{-1})^k = \sum_{k=1}^{m} c_k (PJP^{-1})(PJP^{-1}) \cdots (PJP^{-1}) = P(\sum_{k=1}^{m} c_k J^k)P^{-1} = Pq(J)P^{-1}$. Thus, $q(A)$ is similar to $q(J)$ and they have the same eigenvalues. Now the good thing is that $q(J)$ is

Proving the Jordan Form

a triangular matrix with diagonal entries $\{q(\lambda_1), \dots, q(\lambda_n)\}$. This is not difficult to show, we skip the details. Since $q(A)$ is similar to $q(J)$, we are done. $\square$

We now move toward the proof of the Jordan form. There are many proofs in the literature, some very long but quite educative, check for example the book “Linear Algebra Done Right”, by Axler, Chapter 8. Several results in that chapter are interesting in their own right, but reading it is an investment.

We discuss in the following an approach by Gohberg and Goldberg[1]. Let us start with a preliminary remark.

---

[1] *A Simple Proof of the Jordan Decomposition Theorem for Matrices*, American Mathematical Monthly 1996.

Proving the Jordan Form

Consider a linear map $\mathcal{A} : \mathbb{C}^n \to \mathbb{C}^n$. A subspace W of $\mathbb{C}^n$ is called *cyclic* if it has the form

$$W = \mathrm{Span}\{\mathbf{v}, (\mathcal{A} - \lambda I)\mathbf{v}, \ldots, (\mathcal{A} - \lambda I)^{m-1}\mathbf{v}\}$$

for some $\mathbf{v} \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$, such that $(\mathcal{A} - \lambda I)^{m-1}\mathbf{v} \neq \mathbf{0}$ and $(\mathcal{A} - \lambda I)^m \mathbf{v} = \mathbf{0}$.

Such a cyclic subspace W is invariant under $\mathcal{A}$, meaning that if $\mathbf{w} \in W$, then $\mathcal{A}\mathbf{w} \in W$. Indeed, if $\mathbf{w} \in W$, then $\mathbf{w} = \sum_{i=0}^{m-1} c_k (\mathcal{A} - \lambda I)^k \mathbf{v}$ and $\mathcal{A}\mathbf{w} = (\mathcal{A} - \lambda I)\mathbf{w} + \lambda \mathbf{w} = \sum_{i=1}^{m-1} c_k (\mathcal{A} - \lambda I)^k \mathbf{v} + \lambda \mathbf{w} \in W$. It also holds that $\dim W = m$. Indeed, to see that the spanning set is linearly independent, suppose $c_0 \mathbf{v} + c_1 (\mathcal{A} - \lambda I)\mathbf{v} + \cdots + c_{m-1}(\mathcal{A} - \lambda I)^{m-1}\mathbf{v} = \mathbf{0}$. Applying

Proving the Jordan Form

$(\mathcal{A}-\lambda I)^{m-1}$ to both sides, noting that $(\mathcal{A}-\lambda I)^m\mathbf{v}=\mathbf{0}$, all the terms on the LHS except the first one vanish and we get $c_0(\mathcal{A}-\lambda I)^{m-1}\mathbf{v}=\mathbf{0}$. Since $(\mathcal{A}-\lambda I)^{m-1}\mathbf{v}\neq\mathbf{0}$, we must have $c_0=0$. So we now get the equation $c_1(\mathcal{A}-\lambda I)\mathbf{v}+\cdots+c_{m-1}(\mathcal{A}-\lambda I)^{m-1}\mathbf{v}=\mathbf{0}$. This time we multiply both sides by $(\mathcal{A}-\lambda I)^{m-2}$ to get that $c_1=0$. Continuing this way, we see that $c_i=0\ \forall i$, so the set is linearly independent.

Suppose W is such a cyclic subspace and define $\mathbf{p}_1=(\mathcal{A}-\lambda I)^{m-1}\mathbf{v}$, $\mathbf{p}_2=(\mathcal{A}-\lambda I)^{m-2}\mathbf{v}$, $\ldots$, $\mathbf{p}_m=\mathbf{v}$. Then our argument above says that $\mathcal{B}=\{\mathbf{p}_1,\ldots,\mathbf{p}_m\}$ is a basis of W. A nice thing is that $(\mathcal{A}-\lambda I)\mathbf{p}_1=\mathbf{0}$, implying $\mathcal{A}\mathbf{p}_1=\lambda\mathbf{p}_1$. Next, $(\mathcal{A}-\lambda I)\mathbf{p}_2=\mathbf{p}_1$, implying $\mathcal{A}\mathbf{p}_2=\lambda\mathbf{p}_2+\mathbf{p}_1$. In general, we get

Proving the Jordan Form

$$\mathcal{A}\mathbf{p}_1 = \lambda\mathbf{p}_1 \quad \text{and} \quad \mathcal{A}\mathbf{p}_k = \lambda\mathbf{p}_k + \mathbf{p}_{k-1} \tag{2}$$

for $k = 2, \ldots, m$.

Reduction: To prove the Jordan form, it suffices to show that any linear $\mathcal{A} : \mathbb{C}^n \to \mathbb{C}^n$ induces a direct sum decomposition of $\mathbb{C}^n$ into cyclic subspaces.
In other words, it suffices to show that

(i) There are p cyclic subspaces $W_1, \ldots, W_p$ of $\mathbb{C}^n$, where $W_i = \mathrm{Span}(\mathcal{B}_i)$ and $\mathcal{B}_i = \{\mathbf{v}_i, (\mathcal{A} - \lambda_i)\mathbf{v}_i, \ldots, (\mathcal{A} - \lambda_i)^{m_i - 1}\mathbf{v}_i\}$,

(ii) The union $\mathcal{B} = \bigcup_{i=1}^{p} \mathcal{B}_i$ is a basis for $\mathbb{C}^n$.

Proving the Jordan Form

Proof of the reduction. In fact, suppose (i)-(ii) are true and define $\mathbf{p}_{i,k} = (\mathcal{A} - \lambda_i I)^{m_i - k}\mathbf{v}_i$. Then we showed in (2) that $\mathcal{B}_i = \{\mathbf{p}_{i,1}, \ldots, \mathbf{p}_{i,m_i}\}$ satisfies $\mathcal{A}\mathbf{p}_{i,1} = \lambda_i \mathbf{p}_{i,1}$ and $\mathcal{A}\mathbf{p}_{i,k} = \lambda_i \mathbf{p}_{i,k} + \mathbf{p}_{i,k-1}$ for $k > 1$. Now let $\mathbf{u} \in \mathbb{C}^n$. By (i)-(ii), $\mathbf{u}$ has a unique expansion

$$\mathbf{u} = \sum_{i=1}^{p} \sum_{k=1}^{m_i} \alpha_{i,k} \mathbf{p}_{i,k} \, .$$

So

$$\begin{aligned}
\mathcal{A}\mathbf{u} &= \sum_{i=1}^{p} \Big(\alpha_{i,k} \lambda_i \mathbf{p}_{i,1} + \sum_{k=2}^{m_i} \alpha_{i,k} (\lambda_i \mathbf{p}_{i,k} + \mathbf{p}_{i,k-1}) \Big) \\
&= \sum_{i=1}^{p} \Big(\sum_{k=1}^{m_i - 1} (\lambda_i \alpha_{i,k} \mathbf{p}_{i,k} + \alpha_{i,k+1} \mathbf{p}_{i,k}) + \lambda_i \alpha_{i,m_i} \mathbf{p}_{i,m_i} \Big) \, .
\end{aligned}$$

Proving the Jordan Form

So

$$[\mathbf{u}]_{\mathcal{B}} = \begin{bmatrix} \alpha_{1,1} \\ \alpha_{1,2} \\ \vdots \\ \alpha_{p,m_p} \end{bmatrix} \text{ and } [\mathcal{A}\mathbf{u}]_{\mathcal{B}} = \begin{bmatrix} \lambda_1\alpha_{1,1} + \alpha_{1,2} \\ \lambda_1\alpha_{1,2} + \alpha_{1,3} \\ \vdots \\ \lambda_1\alpha_{1,m_1-1} + \alpha_{1,m_1} \\ \lambda_1\alpha_{1,m_1} \\ \lambda_2\alpha_{2,1} + \alpha_{2,2} \\ \vdots \\ \lambda_m\alpha_{p,m_p} \end{bmatrix} = J[\mathbf{u}]_{\mathcal{B}},$$

where J is the Jordan Form matrix of the main theorem. This shows that the matrix of $\mathcal{A}$ in the basis $\mathcal{B}$ is J. If A is the standard matrix of $\mathcal{A}$ and P is the transition

Proving the Jordan Form

matrix from $\mathcal{B}$ to the standard basis $\mathcal{S}$, then this implies (Lec18) that $J = P^{-1}AP$, i.e. A is similar to J. Since any $n \times n$ matrix is the standard matrix of the linear map $\mathcal{A}(\mathbf{x}) = A\mathbf{x}$, this proves the Jordan Form Theorem. $\square$

So to prove the Jordan form, we now need to prove (i)-(ii) for any linear $\mathcal{A}$. This is what the paper of Gohberg-Goldberg does (see previous footnote) and you can check out their paper, it takes just 1.5 pages. [2]

For more information/applications of the Jordan form, check out the wikipedia article “Jordan normal form”.

---

[2] You will need this definition: Let U, W be subspaces of V. Then V is said to be the *direct sum* of U and W, and we write $V = U \oplus W$, if $V = U + W = \{\mathbf{u} + \mathbf{w} : \mathbf{u} \in U, \mathbf{w} \in W\}$ and $U \cap W = \{\mathbf{0}\}$.