MA2101 L4 - Jordan Canonical Form

1 Jordan Block

- Suppose λ is any scalar.
- Then a *Jordan block* of size m is simply an $(m \times m)$ matrix of the form ...

$$\begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & \lambda \end{bmatrix}$$

(a) Generic form

$$\begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

(b) m=4

Figure 1: Jordan block structure

Note that a (1×1) Jordan block is simply a **scalar**

1.1 Properties of the Jordan block

<u>CLAIM</u>: Every Jordan block has only **one** eigenspace , and it is **one-dimensional**

1. Let's prove why the Jordan block has only one eigenspace. This is equivalent to saying that the Jordan block has only one eigenvalue (which already seems true by definition).
   1.1. Suppose we rewrite the diagonal λ's of the generic Jordan block in Fig1 as λ_1
   1.2. Consider the *characteristic polynomial* p_A of the 'rewritten' generic Jordan block A, defined as $p_A(\lambda) = det(A - \lambda I)$
   1.3. We know that the determinant of an upper triangular matrix is the product of the **main** diagonal entries, therefore $p_A = (\lambda_1 - \lambda)^m$, supposing that A is an $m \times m$ square matrix
   1.4. Thus, we have that the **only** eigenvalue is λ_1
2. Let's prove that the eigenspace corresponding to the eigenvalue λ_1 is one-dimensional
   2.1. The eigenspace is defined as the nullspace of (A-λI), since we are trying to solve $(A - \lambda I)v = 0$
   2.2. Solving for v above, we get $v = (1, 0, \ldots, 0)^T$
   2.3. Span of v is one dimensional, therefore the eigenspace must be one dimensional

Intuitively, the presence of Jordan blocks signals to us that we **do not have enough eigenvectors** to form a basis – which is why we cannot diagonalise

2 Jordan Basis

A *Jordan basis* is a basis st. the matrix (wrt to this Jordan basis) of some operator T consists of *Jordan blocks* J_i

$$\begin{pmatrix} J_1 & 0 & 0 \\ 0 & J_2 & 0 \\ 0 & 0 & J_3 \end{pmatrix}$$

Figure 2: Abstracted view

Since the entire matrix is upper triangular (and so are the Jordan blocks J_i), the eigenvalues of the matrix are along the **main** diagonal

$$\begin{pmatrix} \lambda_1 & 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_3 & 1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_3 & 1 \\ 0 & 0 & 0 & 0 & 0 & \lambda_3 \end{pmatrix}$$

Figure 3: 3 Jordan blocks in a 6×6 matrix

Ordering is not unique, it depends on how we order the basis vectors. Furthermore, the λ_i's do not have to be distinct

2.1 Properties of the Jordan basis

CLAIM: Every operator on a **complex** vector space has a Jordan basis

- This Jordan basis is **unique**, disregarding the possibility of changing the ordering of the basis. This is because a basis is strictly speaking an ordered set, therefore changing the ordering is changing the basis.
- Intuitively, this is saying that every **complex** matrix A is *similar* to a matrix B that is built up from Jordan blocks, similar in the sense that $A = P^{-1}BP$ for some *change-of-basis* matrix P

Proof is skipped for this, it is quite involved.

3 Jordan Canonical Form

When an operator is represented by a matrix wrt a Jordan basis, we say the matrix is in *Jordan Canonical Form.*

1. Suppose we agree that an upper triangular matrix A of the form $\begin{pmatrix} \lambda_1 & & & & \\ & \lambda_2 & & \textbf{*} & \\ & & \ddots & & \\ & \textbf{0} & & \lambda_{n-1} & \\ & & & & \lambda_n \end{pmatrix}_{\{b_1, b_2, \dots, b_n\}}$

   has the property that $Ab_i = \alpha_1 b_1 + \alpha_2 b_2 + \cdots + \lambda_i b_i$. (This is taken at face-value, it is not simple to prove.)

   1.1. Intuitively, every basis vector b_i is mapped to a linear combination of it's 'lower' basis vectors $b_j, j \leq i$

   1.2. We can see this from the matrix structure. For some i^{th} column, we have that the rows $\{A_{1i}, A_{2i}, \dots, A_{ii}\}$ are non-zero. Therefore, this means that the i^{th} basis vector is mapped to a linear combination of the previous *(i-1)* basis vectors.

2. Suppose now that we have the Jordan Canonical Form matrix as seen below.

$$\begin{pmatrix} \lambda_1 & 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_3 & 1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_3 & 1 \\ 0 & 0 & 0 & 0 & 0 & \lambda_3 \end{pmatrix}$$

Figure 4: JCF Matrix

   2.1. Observe some i^{th} column. We have that there are **at most two** non-zero entries, therefore every basis vector is mapped to a linear combination of **at most two** basis vectors (including the original basis vector).

3.1 Properties of the Jordan Canonical Form

3.1.1 Multiplicity of λ

1. Let λ be an eigenvalue of a linear transformation
2. We define the *multiplicity* of λ as the sum of sizes of the Jordan blocks corresponding to that eigenvalue

   2.1. Simply put, multiplicity is the total number of times a given eigenvalue occurs down the diagonal

3.1.2 Characteristic polynomial of T

1. Let $\lambda_1, \lambda_2, \dots$ be the eigenvalues of a linear operator T
2. Let m_i be the multiplicity of some eigenvalue λ_i
3. We define the *characteristic polynomial of T* as the polynomial $\mathcal{X}_T(x) = (x - \lambda_1)^{m_1} \times (x - \lambda_2)^{m_2} \times \dots$

   3.1. By definition, we have that $\forall i,\ \mathcal{X}_T(\lambda_i) = 0$

   3.2. However, a more surprising claim is the *Cayley-Hamilton Theorem* in the subsequent section

4 Cayley-Hamilton Theorem

4.1 Definiton of Cayley-Hamilton Theorem

Cayley-Hamilton Theorem states that $\mathcal{X}_T(T) = \mathbf{0}$; (Note that $\mathbf{0}$ is the zero matrix)

- Simply put, it says that the transformation (equivalently, it's matrix relative to any basis) satisfies it's own characteristic equation
- Bear in mind that any matrix *similar* to the given one has the same eigenvalues (a property of similarity). Therefore, it must also satisfy any polynomial equation satisfied by the original matrix.

4.2 Proving Cayley-Hamilton Theorem

4.2.1 Proving CHT - Diagonal matrix

The theorem is easily seen to be true for a diagonal matrix $diag(a_1, a_2, \cdots, a_n)$. We note that the characteristic polynomial is simply $f(X) = (a_1 - X)(a_2 - X)\cdots(a_n - X)$. Then we see that

$$\begin{aligned} f(A) &= (a_1 I - A)(a_2 I - A)\cdots(a_n I - A) \\ &= diag(0, a_2, \cdots, a_n) diag(a_1, 0, a_3, \cdots, a_n) \cdots diag(a_1, a_2, \cdots, a_{n-1}, 0) \\ &= diag(0, 0, \cdots, 0) = 0 \end{aligned}$$

4.2.2 Proving CHT - Diagonalizable matrix

1. Suppose we have a *diagonalizable* matrix A
2. By definition of diagonalizablity, $A = PDP^{-1}$
3. Therefore, $(A - \lambda I) = (PAP^{-1} - \lambda I) = P(A - \lambda I)P^{-1}$
4. det$\Big(P(A - \lambda I)P^{-1}\Big)$ = det(P) * det(A-λI) * det(P^{-1}) = det(A-λI)
5. Therefore, we observe that *diagonalizable* matrix A and *diagonal* matrix D have the same characteristic polynomial

4.2.3 Proving CHT - Non-diagonalizable matrix (Using JCF)

The difficult part is in proving CHT for non-diagonalizable matrices (if we can prove this, then we can prove CHT is true for any general square matrix).

Tutorial question uses *Analysis* method of proving. Here, we will use JCF to prove.

1. Assume that we have used *similarity* to turn some square matrix into it's Jordan Canonical Form
2. Let's use a concrete example for this case. Consider the Jordan block B we saw in the *6×6* matrix of Fig3

$$B = \begin{bmatrix} \lambda_3 & 1 & 0 \\ 0 & \lambda_3 & 1 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

TO BE DONE.... MA2101 L4 Pg39

$$\frac{1}{1+e^{-x}} \xrightarrow{\text{Translate right}} \frac{1}{1+e^{-(x-3)}} \xrightarrow{\text{Horizontal Stretch}} \frac{1}{1+e^{-(\frac{6}{256}x-3)}} \xrightarrow{\text{Vertical Stretch}} \frac{256}{1+e^{-(\frac{6}{256}x-3)}}$$