1 Review of basics

Fix an algebraically closed field F, that is, a field F such that for any polynomial $p(\lambda) \in P_n(F)$ of degree n (and not less), we have $p(\lambda) = (\lambda - \lambda_1)^{m_1} \cdot ... \cdot (\lambda - \lambda_k)^{m_k}$ with $m_1 + ... + m_k = n$ for some $\lambda_1, ..., \lambda_k \in F$.

1.1 The matrix of a linear map

Let V be a vector space over F throughout, and let $\alpha = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ a basis for V. Then for each $\mathbf{v} \in V$ there exists a unique list

$$x_1, ..., x_n \in F \quad \text{such that} \quad \mathbf{v} = x_1\mathbf{v}_1 + ... + x_n\mathbf{v}_n \qquad \text{and we write} \qquad [\mathbf{v}]^\alpha = [x_i] \ .$$

Let V, W be vector spaces and $T : V \to W$ a linear map. Given $\alpha = \{\mathbf{v}_1, ..., \mathbf{v}_n\}, \beta = \{\mathbf{w}_1, ..., \mathbf{w}_m\}$ a choice of bases for V and W, for each $j = 1, ..., n$, the vector $T(\mathbf{v}_j) \in W$ has a unique decomposition $T(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + ... + a_{mj}\mathbf{w}_m$ for some $a_{ij} \in F$ with $i = 1, ..., m$. Then, we write

$$[T]^\beta_\alpha = \big[T(\mathbf{v}_1)| \quad \cdots \quad |T(\mathbf{v}_n)\big]^\beta = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & a_{ij} & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in M_{m\times n}(F)$$

The j^{th} column of $[T]^\beta_\alpha$ describes $T(\mathbf{v}_j)$, the image under T of the j^{th} vector $\mathbf{v}_j$ in the basis α, in terms of coordinates defined by β.

1.2 Eigenspaces and diagonalization

Let $T : V \to V$ be a linear map throughout. For each $\lambda \in F$, we define the λ eigenspace $E_\lambda = \ker(T - \lambda I) = \{\mathbf{v} \in V | T(\mathbf{v}) = \lambda\mathbf{v}\} \subset V$ for T. A nonzero vector $\mathbf{v} \in E_\lambda$ is called an eigenvector, and any λ admitting such is called an eigenvalue.

Note that $\mathbf{w} \in E_\lambda$ is an eigenvector if and only if, for any basis $\alpha = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ such that $\mathbf{v}_i = \mathbf{w}$, the matrix $[T]^\alpha_\alpha$ has i^{th} column entirely zero except in the i^{th} row, which is λ:

$$[T]^\alpha_\alpha = \big[T(\mathbf{v}_1)| \quad \cdots \quad |T(\mathbf{w})| \quad \cdots \quad |T(\mathbf{v}_n)\big]^\alpha = \begin{bmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{i1} & \cdots & \lambda & \cdots & a_{in} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{m1} & \cdots & 0 & \cdots & a_{mn} \end{bmatrix} \in M_{n\times n}(F)$$

In particular, this implies that $[T]^\alpha_\alpha$ is diagonal with $a_{ii} = \lambda_i$ if and only if every element $\mathbf{v}_i$ of the basis α is an eigenvector with eigenvalue λ_i.

1.3 Invariant subspaces and blocks of a matrix

A subspace $W \subset V$ is called invariant under T if $T(W) \subset W$, or equivalently $T(\mathbf{w}) \in W$ for each $\mathbf{w} \in W$. Note that for W invariant, we can define $T|_W : W \to W$.

A matrix $A \in M_{n\times n}(F)$ has a block of size j if there are j columns of A, numbered by $i_1, ..., i_j$ say, such that each column is zero except in the rows $i_1, ..., i_j$; the remaining nonzero entries define a matrix $\tilde{A} \in M_{j\times j}(F)$, which is called the block.

A subspace W is invariant under T if and only if for any basis $\tilde{\alpha}$ for W extended to a basis α for V, the matrix $[T]$ has a block of size $\dim W$ in the columns corresponding to the basis elements of W. In this case, we have that $[T|_W]^{\tilde{\alpha}}_{\tilde{\alpha}} = \tilde{A}$ is given by the corresponding block.

$$[T] = \begin{bmatrix} 1 & 0 & 4 & 0 \\ 5 & 4 & 3 & 1 \\ 4 & 0 & 7 & 8 \\ 9 & 0 & 5 & 9 \end{bmatrix} \qquad [T] = \begin{bmatrix} 1 & 3 & 0 & 0 & 6 \\ 3 & 4 & 0 & 0 & 3 \\ 0 & 2 & 4 & 2 & 0 \\ 5 & 4 & 3 & 4 & 1 \\ 4 & 6 & 0 & 0 & 8 \end{bmatrix} \qquad [T] = \begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 6 & 7 & 3 & 0 & 0 \\ 4 & 8 & 3 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 \\ 5 & 0 & 0 & 5 & 1 \end{bmatrix}$$

2 Jordan Canonical Form

2.1 Overview

A linear map $T : V \to V$ is called block diagonal if $V = W_1 \oplus ... \oplus W_k$ where each W_i is an invariant subspace. Equivalently, if we choose a basis for each W_i and adjoin them all to a basis α for V, then $[T]^\alpha_\alpha$ consists of k blocks of size $\dim W_j$:

$$[T] = \begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 7 & 3 & 0 & 0 \\ 0 & 8 & 3 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 5 & 1 \end{bmatrix} \qquad V = W_1 \oplus W_2 \oplus W_3 = \text{Span}(\mathbf{v}_1) \oplus \text{Span}(\mathbf{v}_2, \mathbf{v}_3) \oplus \text{Span}(\mathbf{v}_4, \mathbf{v}_5)$$

A matrix $A \in M_{l\times l}(F)$ is called a λ_i Jordan block of size l if it is of the form:

$$A = \begin{bmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & 0 \\ \vdots & \vdots & 0 & \lambda_i & 1 \\ 0 & \dots & 0 & 0 & \lambda_i \end{bmatrix} \qquad \text{e.g.} \qquad A = [\lambda_i] \qquad A = \begin{bmatrix} \lambda_i & 1 \\ 0 & \lambda_i \end{bmatrix} \qquad A = \begin{bmatrix} \lambda_i & 1 & 0 \\ 0 & \lambda_i & 1 \\ 0 & 0 & \lambda_i \end{bmatrix} \qquad A = \begin{bmatrix} \lambda_i & 1 & 0 & 0 \\ 0 & \lambda_i & 1 & 0 \\ 0 & 0 & \lambda_i & 1 \\ 0 & 0 & 0 & \lambda_i \end{bmatrix}$$

Now, we state the main theorem:

Theorem 2.1. Let V be an n dimensional vector space over F an algebraically closed field, and $T : V \to V$ be a linear map with characteristic polynomial

$$p_T(\lambda) = (\lambda - \lambda_1)^{m_1} ... (\lambda - \lambda_k)^{m_k} \qquad \text{so that} \qquad m_1 + ... + m_k = n\ .$$

Then

$$V = K_{\lambda_1} \oplus ... \oplus K_{\lambda_k} \qquad \text{with} \qquad \dim K_{\lambda_i} = m_i$$

and each K_{λ_i} is an invariant subspace; thus, T is block diagonal with respect to this decomposition.

Moreover, for each $i = 1, ..., k$, we have $E_{\lambda_i} \subset K_{\lambda_i}$ and there is a further decomposition

$$K_{\lambda_i} = C(\mathbf{v}_1) \oplus ... \oplus C(\mathbf{v}_{j_i}) \qquad j_i = \dim E_{\lambda_i}$$

such that for each $l = 1, ..., j_i$, $C(\mathbf{v}_l)$ is invariant, and there is a natural basis α_l for each $C(\mathbf{v}_l)$ such that $[T|_{C(\mathbf{v}_l)}]^{\alpha_l}_{\alpha_l}$ is a λ_i Jordan block.

Adjoining the bases α over all $i = 1, ..., k$ and all $l = 1, ..., j_i$, we obtain a basis for V such that $[T]$ is block diagonal with all blocks of Jordan type. A matrix of this type is said to be in *Jordan canonical form.*

Example 2.2. The following is an example of the Jordan decomposition guarenteed by the above Theorem.

$$[T] = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & \lambda_2 & 1 & 0 \\ 0 & 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix} \qquad \begin{cases} p_T(\lambda) = (\lambda - \lambda_1)^2 (\lambda - \lambda_2)^3 \\ E_{\lambda_1} = \text{Span}(\mathbf{v}_1) \quad \subset K_{\lambda_1} = \text{Span}(\mathbf{v}_1, \mathbf{v}_2) \\ E_{\lambda_2} = \text{Span}(\mathbf{v}_3, \mathbf{v}_5) \subset K_{\lambda_2} = \text{Span}(\mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5) \\ K_{\lambda_1} = C(\mathbf{v}_2) = \text{Span}(\mathbf{v}_1, \mathbf{v}_2) \\ K_{\lambda_2} = C(\mathbf{v}_4) \oplus C(\mathbf{v}_5) = \text{Span}(\mathbf{v}_3, \mathbf{v}_4) \oplus \text{Span}(\mathbf{v}_5) \end{cases}$$

To understand this theorem, we need to understand two main steps:

(1) How to define the decomposition $V = K_{\lambda_1} \oplus ... \oplus K_{\lambda_k}$

(2) For each λ_i, how to define the decomposition $K_{\lambda_i} = C(\mathbf{v}_1) \oplus ... \oplus C(\mathbf{v}_{j_i})$ and the basis α_l for each $C(\mathbf{v}_l)$ which makes $[T|_{C(\mathbf{v}_l)}]^{\alpha_l}_{\alpha_l}$ of Jordan type.

We will resolve the second item first, restricting to the case there is just a single eigenvalue, then resolve the first item and put it all together.

2.2 Nilpotent canonical form

Consider the special case of the above theorem when there is just a single eigenvalue, and for simplicity assume it is zero. We obtain:

Theorem 2.3. Let $T : V \to V$ with characteristic polynomial $p_T(\lambda) = \lambda^n$. Then there is a decomposition

$$V = K_{\lambda_0} = C(\mathbf{v}_1) \oplus ... \oplus C(\mathbf{v}_{j_0}) \qquad j_0 = \dim E_0 = \ker(T)$$

such that for each $l = 1, ..., j_0$, $C(\mathbf{v}_l)$ is invariant, and there is a natural basis α_l for each $C(\mathbf{v}_l)$ such that $[T|_{C(\mathbf{v}_l)}]_{\alpha_l}^{\alpha_l}$ is a Jordan block with zero diagonal entries.

We now explain how to define the subspaces $C(\mathbf{v}_l)$ and the basis α_l such that $[T|_{C(\mathbf{v}_l)}]_{\alpha_l}^{\alpha_l}$ is a Jordan block.

A linear map T with characteristic polynomial $p_T(\lambda) = \lambda^n$ is always nilpotent, that is, for each $\mathbf{v} \in V$ there exists k such that $T^k\mathbf{v} = 0$.

Fix $\mathbf{v} \in V$ and choose the minimal such k as above, the set $\{T^{k-1}\mathbf{v}, T^{k-2}\mathbf{v}, ..., \mathbf{v}\}$ is called the cycle of $\mathbf{v}$, and

$$C(\mathbf{v}) = \text{Span}\{T^{k-1}\mathbf{v}, T^{k-2}\mathbf{v}, ..., \mathbf{v}\} \subset V$$

is called the cyclic subspace. The integer $k = \dim C(\mathbf{v})$ is called the length of the cycle of $\mathbf{v}$. We showed that the elements of the cycle are always linearly independent, and thus form a basis for $C(\mathbf{v})$, which we denote α. Then we have:

Corollary 2.4. $[T|_{C(\mathbf{v})}]_{\alpha}^{\alpha}$ is a Jordan block with zero diagonal entries.

The above was defined for any vector $\mathbf{v} \in V$. We want a decomposition of V into cyclic subspaces as in the theorem.

A vector $\mathbf{v} \in V$ is called maximal if $\mathbf{v} \notin \text{im}(T)$. There can be at most $j_0 = \dim E_0 = \dim\ker(V)$ linearly independent maximal vectors, and we choose a list $\mathbf{v}_1, ..., \mathbf{v}_{j_0}$ of such. Then, we showed the resulting subspaces $C(\mathbf{v}_i)$ were mutually linearly independent, and that

$$V = C(\mathbf{v}_1) \oplus ... \oplus C(\mathbf{v}_{j_0})$$

as desired.

Example 2.5. Consider $V = F^2$ with basis $\alpha = \mathbf{v}_1, \mathbf{v}_2$. Then

$$[T]_\alpha^\alpha = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad \text{corresponds to} \qquad V = C(\mathbf{v}_2) = \text{Span}(\mathbf{v}_1, \mathbf{v}_2)$$

Note the dimension of the kernel $j_0 = 1$, and this agrees with the number of cyclic subspaces in the decomposition.

Example 2.6. Consider $V = F^3$ with basis $\alpha = \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$. Then

$$[T]_\alpha^\alpha = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad \text{corresponds to} \qquad V = C(\mathbf{v}_3) = \text{Span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$$

Again, the dimension of the kernel $j_0 = 1$, and this agrees with the number of cyclic subspaces in the decomposition.

Example 2.7. Consider again $V = F^3$ with basis $\alpha = \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$. Then

$$[T]_\alpha^\alpha = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \text{corresponds to} \qquad V = C(\mathbf{v}_2) \oplus C(\mathbf{v}_3) = \text{Span}(\mathbf{v}_1, \mathbf{v}_2) \oplus \text{Span}(\mathbf{v}_3)$$

Here, the dimension of the kernel $j_0 = 2$, and this again agrees with the number of cyclic subspaces in the decomposition.

2.3 Jordan canonical form

Now, we return to the general case of $T : V \to V$ with $\dim V = n$, and

$$p_T(\lambda) = (\lambda - \lambda_1)^{m_1} ... (\lambda - \lambda_k)^{m_k} \qquad \text{so that} \qquad m_1 + ... + m_k = n \ .$$

We defined the generalized eigenspace of T corresponding to eigenvalue λ as the subspace

$$K_\lambda = \{\mathbf{v} \in V | (T - \lambda)^k \mathbf{v} = \mathbf{0}\} \ .$$

Moreover, we showed that

- K_{λ_i} is invariant under T
- $\dim K_{\lambda_i} = m_i$ is the multiplicity of the corresponding root of the characteristic polynomial
- $K_{\lambda_i} \subset V$ are mutually linearly independent

From this, we conclude that

$$V = K_{\lambda_1} \oplus ... \oplus K_{\lambda_k}$$

and that T is block diagonal with respect to this decomposition.

Now, it remains to apply our work in the nilpotent case to each K_{λ_i}. Note that by definition, we have:

Corollary 2.8. The restriction $T - \lambda_i I$ to K_{λ_i} is nilpotent.

Thus, we can apply our nilpotent normal form theorem to each $T - \lambda_i I$ restricted to K_{λ_i}: we choose maximal vectors $\mathbf{v}_1, ..., \mathbf{v}_{j_i}$ of lengths $k_1, ..., k_l$, and let

$$C(\mathbf{v}_l) = \text{Span}\{(T - \lambda_i I)^{k_l - 1}\mathbf{v}_l, (T - \lambda_i I)^{k_l - 2}\mathbf{v}_l, ..., (T - \lambda_i I)\mathbf{v}_l, \mathbf{v}_l\}$$

together with the basis

$$\alpha_l = \{(T - \lambda_i I)^{k_l - 1}\mathbf{v}_l, (T - \lambda_i I)^{k_l - 2}\mathbf{v}_l, ..., (T - \lambda_i I)\mathbf{v}_l, \mathbf{v}_l\} \ .$$

Then, by construction, we have:

Corollary 2.9. $[T|_{C(\mathbf{v}_l)}]_{\alpha_l}^{\alpha_l}$ is a λ_i Jordan block.

This essentially completes the proof of the theorem. Moreover, it provides the following algorithm for computing Jordan canonical form:

- Compute the characteristic polynomial $p_T(\lambda) = (\lambda - \lambda_1)^{m_1} ... (\lambda - \lambda_k)^{m_k}$
- For each eigenvalue λ_i, compute $\ker(T - \lambda_i I)^j$ for $j = 1, 2, ...$ until it is of dimension m_i.
- A vector is maximal if and only if it is in $\ker(T - \lambda_i)^j$ but not in the image of $(T - \lambda_i)$.

Recall our example from the overview:

$$[T] = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & \lambda_2 & 1 & 0 \\ 0 & 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix} \qquad \begin{cases} p_T(\lambda) = (\lambda - \lambda_1)^2 (\lambda - \lambda_2)^3 \\ E_{\lambda_1} = \text{Span}(\mathbf{v}_1) \quad \subset K_{\lambda_1} = \text{Span}(\mathbf{v}_1, \mathbf{v}_2) \\ E_{\lambda_2} = \text{Span}(\mathbf{v}_3, \mathbf{v}_5) \subset K_{\lambda_2} = \text{Span}(\mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5) \\ K_{\lambda_1} = C(\mathbf{v}_2) = \text{Span}(\mathbf{v}_1, \mathbf{v}_2) \\ K_{\lambda_2} = C(\mathbf{v}_4) \oplus C(\mathbf{v}_5) = \text{Span}(\mathbf{v}_3, \mathbf{v}_4) \oplus \text{Span}(\mathbf{v}_5) \end{cases}$$