Given V is a finite-dimensional vector space over any field $\mathbb{F}$ and $T \in \mathrm{End}(V)$, it would be great to have a *nice* matrix representation $[T]_\beta$ for T.

If T is diagonalisable, we can take a basis of eigen vectors as β to obtain a diagonal matrix.

The happy result is that there is a way to maximise the number of zeroes in the matrix so that the matrix is in a block form.

If a acharacteristic polynomial splits, we can find a matrix in a Jordan Canonical Form.

In general, any matrix can be represented in a Rational Canonical Form.

1 Jordan Canonical Form

Definition 1.1. A **Jordan block** is a n $n \times n$ matrix of the form

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

Definition 1.2. A sqquare matrix is in **Jordan Canonical Form** (JCF) if it is of the form

$$\begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & 0 & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & J_k \end{pmatrix}, \tag{2}$$

where each J_i is a Jordan block.

e.g. any diagonal matrix is in JCF with 1×1 blocks.

Our goal is to prove that, if a characteristic polynomial of T splits, there exists an ordered basis β such that $[T]_\beta$ is in JCF. Moreover, JCF is unique, up to the reordering of blocks.

e.g. Over $\mathbb{Z}_2$, possible matrices in JCF when $n = 3$ are

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

$$\begin{pmatrix} 1 & 1 & \\ & 1 & 1 \\ & & 1 \end{pmatrix}, \begin{pmatrix} 0 & & \\ & 1 & \\ & & 1 \end{pmatrix}, \begin{pmatrix} 0 & & \\ & 1 & 1 \\ & & 1 \end{pmatrix}, \begin{pmatrix} 1 & & \\ & 0 & 1 \\ & & 0 \end{pmatrix}, \begin{pmatrix} 1 & & \\ & 0 & \\ & & 0 \end{pmatrix}.$$

Suppose $[T]_\beta$ is a $n \times n$ Jordan block. Then the characteristic polynomial is $(-1)^n(t-\lambda)^n$.

Thus, by the Cayley-Hamilton Theorem $(T - \lambda I)^n = 0$, and hence $(T - \lambda I)^n v = 0$ for all $v \in V$. Moreover, n is a minimal positive integer with such a property.

Definition 1.3. A nonzero vector $v \in V$ is a **generalised eigenvector** corresponding to λ if $(T - \lambda I)^n v = 0$ for some $n \geq 1$.

e.g. If v is an eigenvector with the corresponding eigenvalue λ, then it is a generalised egienvector with $n = 1$.

Definition 1.4. The **generalised eigenspace** of T corresponding to λ is

$$K_\lambda = \{v \in V \mid \exists n \geq 1.(T - \lambda I)^n v = 0\}.$$

Note that each $K_\lambda$is an eigenspace of all generalised eigenvectors and 0. Thus, $E_\lambda \subseteq K_\lambda$.

Theorem 1.5 a) K_λ is a T-invariant subspace.
b) $T - \mu I \in \mathrm{End}(K_\lambda)$ is injective for all $\mu \neq \lambda$.

Proof.

a) Note first that $0 \in K_\lambda$.

Suppose $x, y \in K_\lambda$, $c \in \mathbb{F}$.

Therefore, $(T - \lambda I)^r(x) = 0$ and $(T - \lambda I)^s(y)$ for some $r, s \geq 1$.

Let $n = \max(r, s)$. If $x, y \in \ker(T - \lambda I)^n$, then $cx + y \in \ker(T - \lambda I)^n \subseteq K_\lambda$.

If $x \in K_\lambda$, then $(T - \lambda I)^n x = 0$ for some n.

Therefore, $(T - \lambda I)^n Tx = T(T - \lambda I)^n x = 0$. Hence, $Tx \in K_\lambda$.

b) By a), we know that K_λ is $(T - \mu I)$-invariant.

Thence, $(T - \mu I) \in \mathrm{End}(K_\lambda)$ is well-defined.

By way of conradiction, suppose that $(T - \mu I)x = 0$, where $x \in K_\lambda$ and $x \neq 0$.

Since $x \in K_\lambda$, $(T - \lambda I)^n(x) = 0$ for some $n \geq 1$.

By Well-Ordering Principle, We may assume that $n \geq 1$ is the smallest integer satisfying the conditions.

Let $y = (T - \lambda I)^{n-1}x \neq 0$. Note that $(T - \lambda I)y = 0$, and thus $y \in E_\lambda$.

Moreover,

$$\begin{aligned}(T - \mu I)y &= (T - \mu I)(T - \lambda I)^{n-1}x &(3)\\ &= (T - \lambda I)^{n-1}(T - \mu I)x = 0. &(4)\end{aligned}$$

Therefore, $y \in E_\mu$.

So $y \in E_\lambda \cap E_\mu = \{0\}$, because $\lambda \neq \mu$, and thus $y = 0$, which is a contradiction.

Therefore, $T - \mu I \in \mathrm{End}(K_\lambda)$ is injective.

$\square$

Theorem 1.6

Suppose the characteristic polynomial $f(t)$ of T splits.

a) $\dim K_\lambda \leq m_\lambda$, where m_λ is the algebraic multiplicity
b) $K_\lambda = \ker(T - \lambda I)^{m_\lambda}$

Proof.

a) Let $W = K_\lambda$. Then W is T-invariant by Theorem 1.5.

Therefore, the characteristic polynomial of T_W, $f_W(t)$, divides the characteristic polynomial of T by Theorem 5.21. Therefore, $f_W(t)$ splits.

From Theorem 1.5 (b) we know that the only eigenvalue of T_W can be λ.

Thus, $f_W(t) = (-1)^d(1-\lambda)^d | f(t)$, where $d = \dim W \leq m_\lambda$.

b) The fact that $\ker(T-\lambda I)^{m_\lambda} \subseteq K_\lambda$ follows by the definition of K_λ.

We prove that $K_\lambda \subseteq \ker(T-\lambda I)^{m_\lambda}$.

By the Cayley-Hamilton Theorem, $(T_W - \lambda I)^d = 0$ for all $w \in W$.

Therefore, $(T-\lambda I)^{m_\lambda} w = 0$ for all $w \in W$ by part a).

Thus, $W \subseteq \ker(T-\lambda I)^{m_\lambda}$.

□

Theorem 1.7

Suppose that the characteristic polynomial Tsplits.

a) $V = \bigoplus_{i=1}^r K_{\lambda_i}$, where $\lambda_1, \ldots, \lambda_r$ are distinct eigenvalues.

b) $\dim K_\lambda = m_\lambda$ for all eigenvalues λ.

Proof.

We first show that $V = \sum_{i=1}^r K_{\lambda_i}$ by induction on r.

Suppose first that $r = 0$. Since $f(t)$ splits, then $\dim V = 0$.

Assume the claim holds for $r-1$ eigenvalues.

Claim. Let $W = \operatorname{im}(T-\lambda_1 I)^{m_1}$, where $m_1 = m_{\lambda_1}$. Show that $V = K_{\lambda_1} \oplus W$.

We know that $K_{\lambda_1} = \ker(T-\lambda_1 I)^{m_1}$ by Theorem 7.2. Moreover, $\dim K_{\lambda_1} + \dim W = \dim V$.

If $x \in K_{\lambda_1} \cap W$, then $x = (T-\lambda_1 I)^{m_1} y$ for some $y \in V$ and $(T-\lambda_1 I)^{m_1} x = 0$.

Therefore, $(T-\lambda_1 I)^{2m_1} y = 0$ and hence $y \in K_{\lambda_1} = \ker(T-\lambda_1)^{m_1}$. Therefore, $(T-\lambda_1 I)^{m_1} y = x = 0$.

□