4.4.7 Lemma Let λ be an eigenvalue of $L: V \to V$. If $i \geq 1$ is a positive integer such that

$$\ker(L-\lambda \mathrm{Id})^{i} = \ker(L-\lambda \mathrm{Id})^{i+1}.$$

then

$$V_{(\lambda)} = \ker(L-\lambda \mathrm{Id})^{i}.$$

Proof. Homework. $\square$

4.4.8 Corollary Let λ be an eigenvalue of $L: V \to V$ and let m be the multiplicity of λ as a root of $m_L(x)$. Set $V_i = \ker(L-\lambda \mathrm{Id})^i$. Then we have an ascending chain of subspaces

$$\{0\} \subsetneq V_1 \subsetneq V_2 \subsetneq \cdots \subsetneq V_m = V_{(\lambda)}$$

where the inclusions are strict. $\square$

Remark: We have reduced the problem of understanding a linear operator $L: V \to V$ (of a finite-dimensional vector space over a field containing the eigenvalues of L) to the problem of understanding the operators $L|_{V(\lambda)}: V_{(\lambda)} \to V_{(\lambda)}$. The key to understanding $L|_{V_{(\lambda)}}$ is to consider the ascending chain of subspaces

$$\{0\} \subsetneq V_1 \subsetneq V_2 \subsetneq \cdots \subsetneq V_m = V_{(\lambda)}$$

from Corollary 4.4.8.

Idea: Use the ascending chain of subspaces to build a basis for $V_{(\lambda)}$ with respect to which the matrix for $L|_{V_{(\lambda)}}$ is as simple as possible. Below, we will sketch the procedure for finding such a basis.

Sketch: Let $L: V \to V$ be a linear operator and λ an eigenvalue. Set $V_i = \ker(L - \lambda \mathrm{Id})^i$. We have an ascending chain of subspaces

$$\{0\} := V_0 \underbrace{\subset}_{r_1} V_1 \underbrace{\subset}_{r_2} V_2 \subset \cdots \subset V_{m-1} \underbrace{\subset}_{r_m} V_m = V_{(\lambda)}$$

where m is the mult. of λ as a root of $m_L(x)$.

Set $r_i = \dim V_{i+1} - \dim V_i = \dim(V_{i+1}/V_i)$. Then

$$\begin{aligned}
\dim V_{(\lambda)} = \dim V_m &= \dim(\underbrace{V_m/V_{m-1}}_{r_m}) + \dim(V_{m-1}) \\
&= \dim(\underbrace{V_m/V_{m-1}}_{r_m}) + \dim(\underbrace{V_{m-1}/V_{m-2}}_{r_{m-1}}) + \dim(V_{m-2}) \\
&= \cdots \\
&= \sum_{i=1}^{m} \dim(V_i/V_{i-1}) = \sum_{i=1}^{m} r_i .
\end{aligned}$$

Since $V_{i-1} \subset V_i$ for each i, we can find a basis for V_i by finding a basis for V_{i-1} and adding to it a collection of vectors

$$v_1, \ldots, v_{r_i} \in V_i$$

with the property that the cosets

$$v_1 + V_{i-1}, \ldots, v_{r_i} + V_{i-1} \in V_i/V_{i-1}$$

form a basis for V_i/V_{i-1} (this fact is essentially the converse to Weekly 2.4). This fact is not difficult to prove. This implies the following:

Suppose for each $i = 1, \dots, m$ that

Key Fact #1

$$B_i = \{v_1, \dots, v_{r_i}\} \subseteq V_i$$

is chosen such that $v_1 + V_{i-1}, \dots, v_{r_i} + V_{i-1}$ is a basis for V_i/V_{i-1}. Then

$$B_1 \cup B_2 \cup \cdots \cup B_m$$

is a basis for V_m.

Our chain of subspaces

$$\{0\} := V_0 \underbrace{\subset}_{r_1} V_1 \underbrace{\subset}_{r_2} V_2 \subset \cdots \subset V_{m-1} \underbrace{\subset}_{r_m} V_m = V_{(\lambda)}$$

has a special feature:

$$r_1 \geq r_2 \geq \cdots \geq r_m .$$

Set $K = L - \lambda \mathrm{Id}$. Suppose $v_1, \dots, v_{r_i} \in V_i$ are chosen such that $v_1 + V_{i-1}, \dots, v_{r_i} + V_{i-i}$ is a basis for V_i/V_{i-1}. Then

Key Fact #2

$$K(v_1), \dots, K(v_{r_i}) \in V_{i-1}$$

and

$$K(v_1) + V_{i-2}, \dots, K(v_{r_i}) + V_{i-2} \in V_{i-1}/V_{i-2}$$

are independent.

Its easy to prove the above fact. Thus,

$$r_{i-1} = \dim(V_{i-1}/V_{i-2}) \geq \dim(V_i/V_{i-1}) = r_i .$$

Key Facts 1 & 2 provide a procedure for building a basis for V_m. We obtain a basis as follows:

$$\begin{array}{cccc|cccc|cccc|c|cccc|l}
v_1 & v_2 & \cdots & v_{r_m} & & & & & & & & & & & & & & \in V_m \\
Kv_1 & Kv_2 & & Kv_{r_m} & v_{r_m+1} & v_{r_m+2} & \cdots & v_{r_{m-1}} & & & & & & & & & & \in V_{m-1} \\
K^2v_1 & K^2v_2 & & K^2v_{r_m} & Kv_{r_m+1} & Kv_{r_m+2} & \cdots & Kv_{r_{m-1}} & v_{r_{m-1}+1} & v_{r_{m-1}+2} & \cdots & v_{r_{m-2}} & & & & & & \in V_{m-2} \\
\vdots & \vdots & & \vdots & & & & & & & & & \ddots & & & & & \\
K^{m-1}v_1 & K^{m-1}v_2 & & K^{m-1}v_{r_m} & K^{m-2}v_{r_m+1} & K^{m-2}v_{r_m+2} & \cdots & K^{m-2}v_{r_{m-1}} & K^{m-3}v_{r_{m-1}+1} & K^{m-3}v_{r_{m-1}+2} & \cdots & K^{m-3}v_{r_{m-2}} & & v_{r_2+1} & v_{r_2+2} & \cdots & v_{r_1} & \in V_1
\end{array}$$

One begins at row 1 with vectors $v_1, \cdots, v_{r_m} \in V_m$ such that

$$v_1 + V_{m-1}, \cdots, v_{r_m} + V_{m-1}$$

is a basis for V_m / V_{m-1}. By Fact #2,

$$Kv_1, \cdots, Kv_{r_m} \in V_{m-1}$$

and

$$Kv_1 + V_{m-2}, \cdots, Kv_{r_m} + V_{m-2} \in V_{m-1}/V_{m-2}$$

are independent in V_{m-1}/V_{m-2}. Extend to a basis,

$Kv_1 + V_{m-2}, \cdots, Kv_{r_m} + V_{m-2}, v_{r_m+1} + V_{m-2}, \cdots, v_{r_{m-1}} + V_{m-2}$.
The vectors $Kv_1, \cdots, Kv_{r_m}, v_{r_m+1}, \cdots, v_{r_{m-1}}$ form the second row. Repeat!

By construction, the rows $1, \cdots, m$ satisfy the hypothesis of Fact #1:

1st row $\subseteq V_m$ & provides a basis of V_m / V_{m-1}

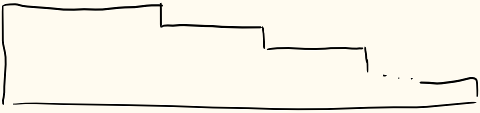
2nd row $\subseteq V_{m-1}$ & provides a basis of V_{m-1}/V_{m-2}

$\vdots$

mth row $\subseteq V_1$ & provides a basis of $V_1/V_0 = V_1 = E_\lambda$.

Thus, the collection of vectors in the diagram forms a basis for V_m.

The shape of the diagram



is important.

height = # of rows = m = multiplicity of λ in $m_L(x)$

width = # of columns = $\dim(E_\lambda)$ = geometric multiplicity of λ

area = $\dim(V(\lambda))$ = algebraic multiplicity of λ.

The shape is independent of the choices we made in the above procedure – it is completely determined by the dimensions of the subspaces .

$$\{0\} \subset V_1 \subset V_2 \subset \cdots \subset V_{m-2} \subset V_{m-1} \subset V_m$$

using that

$$r_m \le r_{m-1} \le r_{m-2} \le \cdots \le r_2 \le r_1$$

# of columns of height $m = \dim(V_m/V_{m-1}) = r_m$

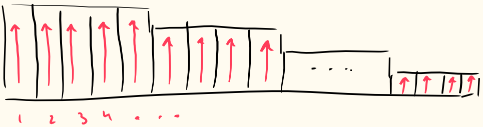
# of columns of height $m-1 = \dim(V_{m-1}/V_{m-2}) - \dim(V_m/V_{m-1}) = r_{m-1} - r_m$

$\vdots$

# of columns of height $2 = \dim(V_2/V_1) - \dim(V_3/V_2) = r_2 - r_3$

# of columns of height $1 = \dim(V_1/V_0) - \dim(V_2/V_1) = r_1 - r_2$.

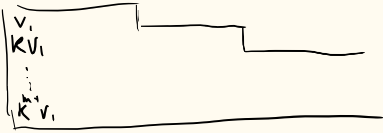
Now, we need to order our basis so that the matrix for $L|_{V(\lambda)}$ is in a simple form. One way is take the columns from left to right with the vectors in each column ordered upwards:



$$B = (1^{st}\text{ column}, 2^{nd}\text{ column}, \cdots\cdots, \text{last column})$$

$$= (K^{m-1}v_1, K^{m-2}v_1, \ldots, Kv_1, v_1, \cdots\cdots)$$

Consider the first column



We have

- $(L-\lambda Id)(K^{m-1}v_1) = K^m v_1 = 0$ (since $v_1 \in V_m$)

$$\Rightarrow L(K^{m-1}v_1) = \lambda K^{m-1}v_1$$

- $(L-\lambda Id)(K^{m-2}v_1) = K^{m-1}v_1$

$\Rightarrow$ $L(K^{m-2}v_1) = K^{m-1}v_1 + \lambda K^{m-2}v_1$

$\vdots$

- $(L-\lambda Id)(Kv_1) = K^2 v_1$

  $\Rightarrow$ $L(Kv_1) = K^2 v_1 + \lambda K v_1$

- $(L-\lambda Id)(v_1) = Kv_1$

  $\Rightarrow$ $L(v_1) = Kv_1 + \lambda v_1$.

Thus, we have computed the first several columns of the matrix:

$$[L|_{V(\lambda)}] = \left(\begin{array}{c|c} \begin{matrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & 0 & \lambda & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{matrix} & ? \\ \hline 0 & ? \end{array}\right)$$

In general, consider the jth column of height h

$$\begin{bmatrix} x_j \\ Kx_j \\ K^2x_j \\ \vdots \\ K^{h-1}x_j \end{bmatrix} \qquad (x_j \in V_h)$$

which consists of h-basis vectors. A similar computation shows that

$$[L|_{V(\lambda)}] = \left(\begin{array}{c|c|c} & 0 & \\ \hline & \begin{matrix}\lambda & 1 & & \\ & \lambda & 1 & \\ & & \lambda & \ddots \; 1 \\ & & & \lambda\end{matrix} & \\ \hline & 0 & \end{array}\right)$$

Notation:

$$J_h(\lambda) := \underbrace{\left.\begin{pmatrix} \lambda & 1 & & & & 0\\ & \lambda & 1 & & & \\ & & \lambda & 1 & & \\ & & & \lambda & \ddots & \\ & & & & \ddots & 1 \\ 0 & & & & & \lambda \end{pmatrix}\right\} h}_{h}$$

The matrix $J_h(\lambda)$ is called the Jordan block of size h for λ. E.g.

$$J_1(\lambda) = (\lambda)$$

$$J_2(\lambda) = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

$$J_3(\lambda) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

Ordering the basis $\mathcal{B}$ in the above w ayields

$$[L|_{V(\lambda)}] = \begin{pmatrix} J_m(\lambda) & & & & & & & & \\ & \ddots & & & & & & & \\ & & J_m(\lambda) & & & & & & \\ & & & J_{m-1}(\lambda) & & & & & \\ & & & & \ddots & & & & \\ & & & & & J_{m-1}(\lambda) & & & \\ & & & & & & \ddots & & \\ & & & & & & & J_2(\lambda) & \\ & & & & & & & & \ddots \\ & & & & & & & & & J_2(\lambda) \\ & & & & & & & & & & J_1(\lambda) \\ & & & & & & & & & & & \ddots \\ & & & & & & & & & & & & J_1(\lambda) \end{pmatrix}$$

a block diagonal matrix w/

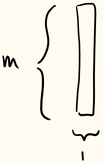
of $J_m(\lambda)$ = # of columns of height m

of $J_m(\lambda)$ = # of columns of height $m-1$

$\vdots$

of $J_1(\lambda)$ = # of columns of height 1

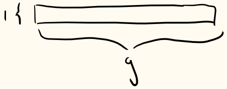
E.g. (some possible shapes)



geometric multiplicity = 1

m = algebraic multiplicity

$J_m(\lambda)$ single Jordan block

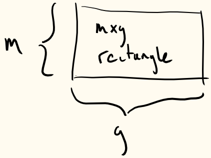


$m = 1$

geometric mult = algebraic mult

All Jordan block have size 1

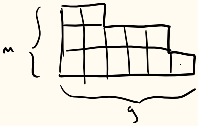
$$\begin{pmatrix} \lambda && \\ & \ddots & \\ && \lambda \end{pmatrix}$$ diagonal



All Jordan block have the same size

Algebraic mult = (geometric mult) x m.

$$\begin{pmatrix} J_m(\lambda) && \\ & \ddots & \\ && J_m(\lambda) \end{pmatrix}$$



$m = 3$

$g = 6$

d = area $= 13$

$$\begin{pmatrix} J_3(\lambda) &&&&& \\ & J_3(\lambda) &&&& \\ && J_2(\lambda) &&& \\ &&& J_2(\lambda) && \\ &&&& J_2(\lambda) & \\ &&&&& J_1(\lambda) \end{pmatrix} = \begin{pmatrix} \boxed{\begin{matrix}\lambda&1&0\\0&\lambda&1\\0&0&\lambda\end{matrix}} &&&&& \\ & \boxed{\begin{matrix}\lambda&1&0\\0&\lambda&1\\0&0&\lambda\end{matrix}} &&&& \\ && \boxed{\begin{matrix}\lambda&1\\0&\lambda\end{matrix}} &&& \\ &&& \boxed{\begin{matrix}\lambda&1\\0&\lambda\end{matrix}} && \\ &&&& \boxed{\begin{matrix}\lambda&1\\0&\lambda\end{matrix}} & \\ &&&&& \lambda \end{pmatrix}$$

In fact, the shape describes precisely the number and size of the Jordan blocks.

All of the above discussion was focused on finding a basis for $V_{(\lambda)}$ in which $T|_{V_{(\lambda)}}$ has a simple form. Putting all of this together, we have:

4.4.9 Theorem (Jordan Basis Theorem) Let V be a finite-dimensional vector space over a field F which contains the eigenvalues of $L: V \longrightarrow V$. Then there exists an ordered basis B of V such that

$$[L]_B = \begin{pmatrix} B_{\lambda_1} & & & \\ & B_{\lambda_2} & & \\ & & \ddots & \\ & & & B_{\lambda_n} \end{pmatrix} \quad \left(\begin{array}{c}\text{block}\\ \text{diagonal}\end{array}\right)$$

where $\lambda_1, \ldots, \lambda_n$ are the distinct eigenvalues of L, and where each matrix B_{λ_i} is a block diagonal matrix consisting of Jordan blocks $J_\ell(\lambda_i)$ (of possibly varying sizes ℓ). In particular, $[L]_B$ is a direct sum of Jordan blocks for its eigenvalues. This matrix is called the Jordan Canonical Form and it is unique up to permutation of the blocks. Thus, L is uniquely determined by

- the distinct eigenvalues $\lambda_1, \ldots, \lambda_n$; and
- the size and number of Jordan blocks for each eigenvalue.

$\square$

4.4.10 Definition A **Jordan matrix** is a square matrix $A \in F^{n\times n}$ which is block diagonal consisting of Jordan blocks $J_d(\lambda)$ for various $d \in \mathbb{N}$, $\lambda \in F$.

4.4.11 Corollary A matrix $A \in F^{n\times n}$ is similar to a Jordan matrix if and only if F contains the eigenvalues of A. Such a matrix is the **Jordan canonical form** of A and is unique up to permutation of the blocks. Two matrices A, B satisfying the above condition are similar if and only if they have the same Jordan form. ☐

4.4.12 Example

$$\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & & & \\ & & \lambda & 1 & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} \ \& \ \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & & & \\ & & \lambda & & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}$$ are not similar

but $$\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & & & \\ & & \lambda & 1 & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} \ \& \ \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}$$ are similar. ☐

How to Compute the Jordan form Let $A \in F^{n\times n}$.

① Compute the eigenvalues $\lambda_1, \dots, \lambda_n$ by finding $cp_A(x)$.

② For each eigenvalue λ, compute the dimensions

$$\dim(\ker(A-\lambda I)) < \dim(\text{Ker}(A-\lambda I)^2) < \cdots < \dim(\text{Ker}(A-\lambda I)^m)$$

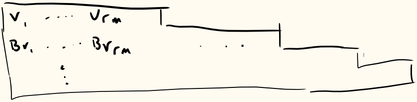
when m is the smallest integer such that

$$\dim(\text{Ker}(A-\lambda I)^m) = V_{(\lambda)} = \text{algebraic mult. of } \lambda.$$

(Note: we know m = mult of λ in $m_A(x)$, so it could be helpful to compute $m_A(x)$ first.)

The dimensions tell you the number and sizes of the Jordan blocks.

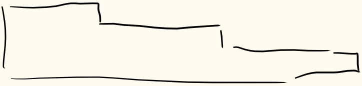
③ If you want to find P such that $P^{-1}AP = J$, then you must follow the procedure for constructing the basis



This procedure provides a basis for $V_{(\lambda)}$. □

Remark A basis of V putting A into Jordan form consists of generalized eigenvectors, but not every basis of generalized eigenvectors will work. One must construct the basis in a special way, as discussed above.

Remark The characteristic and minimal polynomials of A put restrictions on the possible shapes



for each eigenvalue — but, in general, knowledge of the characteristic and minimal polynomials is not enough to determine the Jordan form. Basically, the polynomials determine the height and the area, but not the shape.

4.4.13 Example Let $A = \begin{pmatrix} -12 & -13 & -17 \\ 27 & 29 & 36 \\ -13 & -14 & -17 \end{pmatrix} \in \mathbb{R}^{3\times 3}$.

Then $cp_A(x) = (-1-t)^2(2-t)$. The eigenvalues are $\lambda_1 = -1$ w/ alg. mult. equal to 2, and $\lambda_2 = 2$ w/ alg. mult. equal to 1. One can check that $m_A(x) = cp_A(x)$.

$\underline{\lambda_1 = -1}$ We compute

$$\dim(\ker(A+I)) < \dim(\ker(A+I)^2) = \dim(\ker(A+I)^3)$$

$\text{Rref}(A+I) = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$

$\dim E_{\lambda_1} = \dim(\text{Ker}(A+I)) = 1$

This tells us that the width of the shape is 1.

$\text{Rref}(A+I)^2 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

$\dim V_{(\lambda_1)} = \dim(\ker(A+I)^2) = 2$

This tells us that the area of the shape is 2

The height of the shape is 2 since λ_1 has mult. in $m_A(x) = cp_A(x)$. One could also see this w/out computing $m_A(x)$ by checking that

$$\text{Rref}((A+I)^2) = \text{Rref}((A+I)^3).$$

The shape for λ_1 is

$\rightsquigarrow$ single block $J_2(\lambda_1) = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$

To find the corresponding basis, choose any vector $v_1 \in \ker(A+I)^2$, e.g. $b_1 = (1,0,-1)$. Then $b_2 = (A+I)b_1 = (6,-9,3)$ is the second vector in the basis for $V_{(\lambda_1)}$.

$\lambda_2 = 2$ We know now that $\dim V_{(\lambda_2)} = 1$ since $\dim V_{(\lambda_1)} = 2$ and $\dim \mathbb{R}^3 = 3$ (or because the mult of λ_2 in $m_A(x)$ is equal to 1.). Thus, the shape has area equal to 1. The only possibility is

$\rightsquigarrow$ single block $J_1(\lambda_2) = (2)$

A basis is any non zero vector in $V_{(\lambda_2)} = \ker(A-2I)$. We have

$$\mathrm{Rref}(A-2I) = \begin{pmatrix} 1 & 0 & -1/3 \\ 0 & 1 & 5/3 \\ 0 & 0 & 0 \end{pmatrix}$$

So we could take $b_3 = (1,-5,3)$.

Thus, a basis for V will be $B = (v_1, v_2, v_3)$ and we have $J = P^{-1}AP$ where

$$J = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \qquad P = T^E_B = \begin{pmatrix} 6 & 1 & 1 \\ -9 & 0 & -5 \\ 3 & -1 & 3 \end{pmatrix}.$$

Note: J can be determined from the shapes alone! One does not need to compute P in order to determine J.

□