MATH 108B: NOTES ON JORDAN CANONICAL FORMS

WRITTEN BY: CINDY TSANG

Note: This is only a brief review and doesn't contain everything that you ought to know. You should read Chapter 7 in the textbook. Also, let me know if you find any mistakes.

Motivation. Let $T : V \to V$ be a linear operator on a finite dimensional vector space V. We say that T is *diagonalizable* if V has a basis β such that $[T]_\beta$ is a diagonal matrix. In this case, the vectors in the basis are necessarily eigenvectors. In particular, we have

$$T \text{ is diagonalizable } \iff V \text{ has a basis consisting of eigenvectors.}$$

Let $n = \dim(V)$. If the eigenspaces are "too small" and we can't find n linearly independent eigenvectors, then we cannot diagonalize T. But if the characteristic polynomial of T splits, then we can write T as a *Jordan canonical Form*, which is very close to a diagonal matrix.

Jordan Canonical Form. A matrix is a Jordan canonical form if it looks like

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix},$$

where each A_i is a *Jordan block*. In other words, we have

$$A_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}.$$

So it has λ_i, which will be an eigenvalue, on the diagonal and 1's on top of the diagonal. We want to find a basis β so that $[T]_\beta$ is a Jordan canonical form.

Let's consider one single Jordan block, and for simplicity let's say it has dimension 3. So

$$A_i = \begin{pmatrix} \lambda_i & 1 & \\ & \lambda_i & 1 \\ & & \lambda_i \end{pmatrix}.$$

Let v_1^i, v_2^i, v_3^i be the ordered basis with respect to which this matrix is written, and suppose that A_i is the matrix representation of T_i. We make the following important observations.

[Let's drop the index i for simplicity. So assume that A has one single block.]

(a) Since the j-th column corresponds to the vector $T(v_j)$ written in the basis v_1, v_2, v_3, to say that A has the above form is equivalent to saying that

$$T(v_1) = \lambda v_1 \qquad T(v_2) = v_1 + \lambda v_2 \qquad T(v_3) = v_2 + \lambda v_3.$$

(b) The above equations imply that v_1 is an eigenvector corresponding to λ and that v_2, v_3 are *not* eigenvectors corresponding to λ.

(c) The above equations also imply that

$$v_2 = (T - \lambda I)(v_3) \qquad v_1 = (T - \lambda I)(v_2) = (T - \lambda I)^2(v_3).$$

Hence, vectors in the basis $\{v_1, v_2, v_3\}$ are of the form $(T - \lambda I)^m(v_3)$, where $m = 2, 1, 0$.

(d) Since v_1 is an eigenvector corresponding to λ, we have $(T - \lambda I)^3(v_3) = 0$.

In general, suppose that A has dimension p and let $v_1, ..., v_p$ be the basis with respect to which A is written. Then, the above generalizes to the following.

(a) Among the basis vectors, v_1 and only v_1 is an eigenvector corresponding to λ.

(b) The basis $v_1, ..., v_p$ can be rewritten as

$$(T - \lambda I)^{p-1}(v_p), (T - \lambda I)^{p-2}(v_p), \cdots, (T - \lambda I)(v_p), v_p.$$

In particular, p is the smallest positive integer such that $(T - \lambda I)^p(v_p) = 0$.

These observations lead to the definitions of *generalized eigenvectors* and *generalized eigenspace corrsponding to an eigenvalue* λ, which we shall denote by K_λ.

Some Important Theorems/Facts/Observations. Let $T : V \to V$ be a linear operator on a finite dimensional vector space V. Assume that *the characteristic polynomial of* T *splits*. Let λ be an eigenvalue of T with algebraic multiplicity m.

(a) $\dim(K_\lambda) = m$. In particular, one can find a basis β_λ of K_λ for each λ, and the union of them will be a basis for V, since the multiplicities add up to $\dim(V)$.

(b) Let $d = \dim(E_\lambda)$. Recall that in each Jordan block, the basis vector corresponding to and only to the first column is an eigenvector. Hence, there should be d blocks corresponding to the eigenvalue λ in the matrix of T written as a Jordan canonical form.

(c) For example, if $d = 2$ and $m = 3$, then there should be a 1×1 block and a 2×2 block. If $d = m$, then there are enough eigenvectors and the"big block" corresponding to λ or the subspace K_λ is a diagonal matrix.

(d) Let r be the smallest positive integer such that $K_\lambda = N((T - \lambda I)^r)$. Then, first for any $v \in K_\lambda$ we have $(T - \lambda I)^r(v) = 0$ and so the dimension of each block is at most r. And second, by minimality of r there exists $v \in K_\lambda$ whose cycle has length r. In particular,

$$(T - \lambda I)^{r-1}(v), (T - \lambda I)^{r-2}, \cdots, (T - \lambda I)(v), v$$

are linearly independent (cf. Theorem 7.6), and there exists a block of dimension r.

(e) For example, say $m = 5$. If $r = 5$ then there is one single block. If $r = 1$ then $E_\lambda = K_\lambda$ and the "big block" corresponding to K_λ is diagonal. Suppose further that $d = 2$, then we can have a 1×1 block with a 4×4 block, or a 2×2 block with a 3×3 block. So, $r = 4$ or $r = 3$, which gives us the former and latter cases, respectively.

Example. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be given by the matrix

$$A = \begin{pmatrix} 3 & 1 & -2 \\ -1 & 0 & 5 \\ -1 & -1 & 4 \end{pmatrix}$$

with respect to the standard basis. The characteristic polynomial of A is $-(t-3)(t-2)^2$.

$\lambda_1 = 3$ with multiplicity $m_1 = 1$:

$$A - 3I = \begin{pmatrix} 0 & 1 & -2 \\ -1 & -3 & 5 \\ -1 & -1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}.$$

So $x_1 + x_3 = 0$ and $x_2 - 2x_3 = 0$, and x_3 is the only free variable. Hence, $\dim(E_\lambda) = 1$ and vectors in E_λ are of the form

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_1 \\ 2x_3 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} x_3.$$

We choose $\{(-1, 2, 1)^T\}$ to be a basis for $E_{\lambda_1} = K_{\lambda_1}$. Since $\dim(E_{\lambda_1}) = m_1$, we are done.

$\lambda_2 = 2$ with multiplicity $m_2 = 2$:

$$A - 2I = \begin{pmatrix} 1 & 1 & -2 \\ -1 & -2 & 5 \\ -1 & -1 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{pmatrix}.$$

So $x_1 + x_3 = 0$ and $x_2 - 3x_3 = 0$, and x_3 is the only free variable. Hence, $\dim(E_{\lambda_2}) = 1$, which is less than m_2. Hence, we will have one 2×2 block for the eigenvalue $\lambda_2 = 2$ (one block because $\dim(E_{\lambda_2}) = 1$) and we want a basis $\{v_1, v_2\}$ which satisfies

$$(A - 2I)^2(v_2) = 0 \qquad\qquad (A - 2I)(v_2) = v_1.$$

As there are only two vectors, we can first solve $(A - 2I)^2(v_2) = 0$ and pick any such v_2 with $(A - 2I)(v_2) \neq 0$. Then, define $v_1 = (A - 2I)(v_2)$ and $\{v_1, v_2\}$ will be a basis for K_{λ_2}) that gives a Jordan block (cf. Theorem 5.22 for linear independence of v_1, v_2).

So, write $v_2 = (a, b, c)$ and we solve the first equation. We have

$$\begin{pmatrix} 1 & 1 & -2 \\ -1 & -2 & 5 \\ -1 & -1 & 2 \end{pmatrix}^2 \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 1 & -1 \\ -4 & -2 & 2 \\ -2 & -1 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Since $R_2 = -2R_1$ and $R_3 = -R_1$, doing row reduction eliminates R_2 and R_3, and we obtain

$$\begin{pmatrix} 2 & 1 & -1 \\ -4 & -2 & 2 \\ -2 & -1 & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Hence, $2x_1 + x_2 - x_3 = 0$ and x_2 and x_3 are freen variables. The solution space is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} (x_3 - x_2)/2 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -0.5 \\ 1 \\ 0 \end{pmatrix} x_2 + \begin{pmatrix} 0.5 \\ 0 \\ 1 \end{pmatrix} x_3.$$

By trial and error, we see that $x_2 = 2$ and $x_3 = 0$ work. So, set $v_2 = (-1, 2, 0)^T$ and

$$v_1 = \begin{pmatrix} 1 & 1 & -2 \\ -1 & -2 & 5 \\ -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \\ -1 \end{pmatrix}.$$

Then, with respect to the basis

$$\beta = \left\{ \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -3 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} \right\},$$

the matrix A can be expressed as

$$[A]_\beta = \begin{pmatrix} 3 & & \\ & 2 & 1 \\ & & 2 \end{pmatrix}.$$

Remark. For example, suppose that λ has multiplicity $m = 3$ and $\dim(E_\lambda) = 2$. Since $\dim(E_\lambda) = 2$, there will be two blocks - a 1×1 block and 2×2 block. To find a basis such that the matrix becomes a Jordan canonical form, here is one approach.

(1) Solve $(T - \lambda I)^2(v_3) = 0$ and find such a v_3 such that $(T - \lambda I)(v_3) \neq 0$.
(2) Define $v_2 = (T - \lambda I)(v_3)$. Then $\{v_2, v_3\}$ will give you the 2×2 block.
(3) Solve $(T - \lambda)(v_1) = 0$ and find such a v_1 such that $v_1 \notin \text{span}(v_2, v_3)$.
(4) Then, you can use the basis $\{v_1, v_2, v_3\}$ and the matrix will have the form

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

Question. What do you do in general?
Question. Consider different values of m, d, and r (notation as under *Some Important....Observations* and think of what can happen in each case.
