JORDAN FORM

BENEDICT MORRISSEY

Contents

These are supplementary notes for parts of Lectures 8,9 and 10 of Math 312 Summer II 2019.

1. Eigenvectors and Generalized Eigenvectors

Let $T : V \to V$ be a linear function, from a vector space V to itself.

Definition 1.1. An *eigenvector* of T (corresponding to the *eigenvalue* λ)is a vector (which we generally require to be non zero) $\vec{v}$ such that $T\vec{v} = \lambda\vec{v}$ for some scalar λ. We call a λ such that there is an eigenvector, an eigenvalue of T.

We sometimes call this a λ-eigenvector.

Equivalently $\vec{v} \in Ker(T - \lambda Id)$ (where Id is the identity matrix). We recall that the kernel of a linear transformation is zero precisely when the determinant is zero. Hence the *eigenvalues* of a linear transformation T are precisely the values of λ that are roots of the polynomial $p(\lambda) := det(A - \lambda Id)$. We call $p(\lambda)$ the *characteristic polynomial* of the linear transformation T.

Definition 1.2. A *generalized eigenvector* of T (corresponding to the eigenvalue λ) is a vector $\vec{v}$, such that there is some positive integer k such that $(T - \lambda Id)^k\vec{v} = 0$

We say that a generalized eigenvector of T is of *rank* k if $(T - \lambda Id)^k\vec{v} = 0$ and $(T - \lambda Id)^{k-1}\vec{v} \neq 0$.

We sometimes call this a generalized λ-eigenvector.

Definition 1.3. For λ an eigenvalue of T we let $V_\lambda \subset V$ be the vector subspace of V consisting of generalized eigenvectors of T corresponding to the eigenvalue λ (here we include the zero vector as a generalized eigenvector for the purposes of this definition).

Exercise 1.4. Show that V_λ is a vector subspace of V.

Example 1.5. Consider the matrix $A = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$. The vector $(1, 0)^T$ is an eigenvector of A corresponding to the eigenvalue 3, it is hence a generalized eigenvector of rank 1. The vectors $(0, 1)^T$ and $(1, 1)^T$ are generalized eigenvectors of rank 2.

Theorem 1.1. *If V is a complex vector space (and $T : V \to V$ a linear function), then there is an isomorphism*

$$V \cong V_{\lambda_1} \oplus V_{\lambda_2} \oplus ... \oplus V_{\lambda_n},$$

where λ_1,...,λ_n are the eigenvalues of T.

Note that the map corresponding to this isomorphism is the map

$$V_{\lambda_1} \oplus V_{\lambda_2} \oplus ... \oplus V_{\lambda_n} \xrightarrow{s} V,$$

$$(\vec{v}_1, \vec{v}_2, ...\vec{v}_n) \to \vec{v}_1 + ... + \vec{v}_n.$$

We do not cover the proof of the above theorem, and instead refer to Axler for a proof.

What does the above theorem "really mean?"

It means that if $\vec{v} \in V$ there is a unique way to write

$$\vec{v} = \vec{v}_1 + \vec{v}_2 + ... + \vec{v}_n,$$

where $\vec{v}_i \in V_{\lambda_i}$.

Example 1.6. In the case where the characteristic polynomial has no repeated roots Theorem 1.1 simply says that a set of one eigenvector for each eigenvalue, $(v_{\lambda_1}, ..., v_{\lambda_n})$, is a basis for V.

Exercise 1.7. Show that for V two dimensional the claim of the above example is true.

Proposition 1.8. *Suppose that we have a basis $\mathcal{B}_i = \{\vec{v}_{i1}, \vec{v}_{i2}, ..., \vec{v}_{ij_i}\}$ for each eigenspace V_{λ_i}, then*

$$\cup_i \mathcal{B}_i = \{\vec{v}_{11}, \vec{v}_{12}, ..., \vec{v}_{1j_1}, \vec{v}_{21}, \vec{v}_{22}, ..., \vec{v}_{2j_2}, ..., \vec{v}_{n1}, \vec{v}_{n2}, ..., \vec{v}_{nj_n}\}$$

is a basis of V.

We omit the proof.

2. Diagonalization

We first consider the case where we have a basis of eigenvectors $\mathcal{B} = \{\vec{v}_1, ...\vec{v}_m\}$, with $T\vec{v}_i = \lambda_i \vec{v}_i$. (In general we will need to instead use a basis of generalized eigenvectors).

In this basis we can write

$$D := [T]_{Basis\mathcal{B}} = diag(\lambda_1, ..., \lambda_m) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{pmatrix},$$

where we have adopted that the empty squares of the matrix should be read as containing the entry zero.

Supose we have some standard basis of $\mathcal{C}$ of V (e.g. if $V = \mathbb{R}^n$). Let A_T be the matrix representing is the matrix in this basis. Let P be the change of basis matrix from the basis $\mathcal{B}$ to the basis $\mathcal{C}$ (so P^{-1} is the change of basis matrix from the basis $\mathcal{C}$ to the basis $\mathcal{B}$). Then $P = [\vec{v}_1, ..., \vec{v}_m]$. Furthermore

$$A_T = PDP^{-1}.$$

We refer to such an equation as *diagonalizing* the matrix A.

Exercise 2.1. Go over why the above changes of basis are the correct changes of basis.

We do not always have a basis of eigenvectors. In example 1.5 we only had a one dimensional eigenspace, but a two dimensional generalized eigenspace.

In general we will *not* be able to diagonalize every matrix by changing the basis. We will instead be able to choose a basis that puts a matrix in what is called "Jordan form."

Definition 2.2 (Jordan Form)**.** We say that a matrix B is in Jordan form if B is block diagonal

$$B = \begin{pmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_l \end{pmatrix},$$

with each block being of the form

$$B_i = \begin{pmatrix} \lambda_j & 1 & & & \\ & \lambda_j & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_j & 1 \\ & & & & \lambda_j \end{pmatrix}.$$

The rest of this section is going to cover how to put a matrix in Jordan normal form. We are going to break this up into two algorithms, the first of which utilises the second algorithm.

Remark 2.3. The algorithm is numerically unstable so should not be implemented numerically.

2.0.1. *Algorithm 1.*

(1) Compute the characteristic polynomial and compute the roots of the characteristic polynomial (the eigenvalues).
(2) Compute the generalized eigenspaces V_λ for each eigenvalue λ.
(3) For each generalized eigenspace apply algorithm 2 to produce a basis of the generalized eigenspace V_λ.
(4) Combing the bases of each generalized eigenspace gives us a basis of V by proposition 1.8.

Proposition 2.4. *Let $\mathcal{B}$ be the basis produced by algorithm 2. Then writing T in the basis $\mathcal{B}$ gives us a matrix (which we denote $[T]_{Basis\mathcal{B}}$) which is in Jordan form.*

Proof. Note that $[T]_{Basis\mathcal{B}}$ will be block diagonal with a block for each generalized eigenspace. Hence this proposition is true if the blocks for each eigenspace are in Jordan form. We explain why this is true in proposition 2.5. □

2.0.2. *Algorithm 2.* This Algorithm produces a basis $\mathcal{B}_\lambda$ a generalized eigenspace V_λ such that when we write a matrix for $T|_{V_\lambda}$ in this basis the matrix is in Jordan form.

Set $S = T - \lambda Id$

(1) Pick the highest integer n such that there are generalized λ-eigenvectors of rank n. Pick a basis $\mathcal{C}$ of the vector space $V_{\lambda,<n}$ of generalized λ-eigenvectors of rank less than n. Pick a set of rank n eigenvectors $(\vec{v}_1, ..., \vec{v}_b)$ such that $\{\vec{v}_1, ..., \vec{v}_b\} \cup \mathcal{C}$ is a basis of V_λ. Set $\mathcal{B} = (S^{n-1}\vec{v}_1, S^{n-2}\vec{v}_1 ..., \vec{v}_1, S^{n-1}\vec{v}_2, ..., \vec{v}_2, ..., S^{n-1}\vec{v}_b, ..., \vec{v}_b)$.
(2) Iterate the following two steps until there is no such a on step (3):
(3) Find the highest integer a (smaller than n, and smaller than any a's used on previous iterations) such that $\mathcal{B}$ does not contain all generalized eigenvectors of rank a.
(4) Pick a basis $\mathcal{C}$ of $Span(\mathcal{B}, V_{\lambda,<a})$ Pick vectors $\vec{w}_1, ..., \vec{w}_c$ that extend $\mathcal{C}$ to a basis of V_λ. Add the list $(S^{a-1}\vec{w}_1, S^{a-2}\vec{w}_1, ..., \vec{w}_1, S^{a-1}\vec{w}_2, S^{a-2}\vec{w}_2, ..., \vec{w}_2, ..., A^{a-1}\vec{w}_c, ..., \vec{w}_c)$ to $\mathcal{B}$.

Proposition 2.5. *Consider the basis $\mathcal{B}_\lambda$ produced by algorithm 2. The matrix ($[T|_{V_\lambda}]_{Basis\mathcal{B}_\lambda}$) for $T|_{V_\lambda}$ in this basis is in Jordan form (with all the eigenvalues being λ).*

Proof. Note that $T(\vec{v}_1) = \lambda_1\vec{v}_1 + S\vec{v}_1$.

Hence the matrix for $T|_{Span(S^{n-1}\vec{v}_1, S^{n-2}\vec{v}_1 ..., \vec{v}_1)}$ with respect to the basis $(S^{n-1}\vec{v}_1, S^{n-2}\vec{v}_1 ..., \vec{v}_1)$ is

$$[T|_{Span(S^{n-1}\vec{v}_1, S^{n-2}\vec{v}_1 ..., \vec{v}_1)}]_{Basis\{S^{n-1}\vec{v}_1, S^{n-2}\vec{v}_1 ..., \vec{v}_1\}} == \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}$$

As we can write $V_\lambda = Span(S^{n-1}\vec{v}_1, S^{n-2}\vec{v}_1..., \vec{v}_1) \oplus Span(S^{n-1}\vec{v}_2, S^{n-2}\vec{v}_2..., \vec{v}_2) \oplus .... \oplus Span(S^{a-1}\vec{w}_c, S^{a-2}\vec{w}_c..., \vec{w}_c)$ (where a is the final value of a in the iteration of Algorithm 2), it follows that the matrix for $T|_{V_\lambda}$ with respect to the basis $\mathcal{B}$ is in Jordan form. □

2.1. **Extended Example – writing a matrix in Jordan Form.** Suppose we want to put the following matrix in Jordan Form:

$$A = \begin{pmatrix} 3 & 4 & 5 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & 1 \\ 0 & 0 & 0 & -2 & -4 & 6 \end{pmatrix}$$

The characteristic polynomial is

$$p(\lambda) = det(A - \lambda Id) = (\lambda - 4)^5(\lambda - 3).$$

Hence we have an eigenvalue $\lambda_1 = 3$ of algebraic multiplicity 1, and an eigenvalue $\lambda_2 = 4$ of algebraic multiplicity 5.

We can find the $(\lambda_1 = 3)$-eigenvector by solving $(A - 3Id)\vec{v}_{\lambda_1} = 0$, and we get the solution

$$\vec{v_{\lambda_1}} = (1, 0, 0, 0, 0, 0)^T$$

We can now solve for eigenvectors (generalized eigenvectors of rank 1) for the eigenvalue $\lambda = 4$.

The equation $(A - 4Id)\vec{v}_{\lambda_2} = 0$ has solutions:

$$\vec{v}_{\lambda_2} = a(0, 0, 0, 1, 0, 1)^T + b(0, 0, 0, -2, 1, 0)^T + c(4, 1, 0, 0, 0, 0)^T$$

for $a, b, c \in \mathbb{R}$. We can also solve $(A - 4Id)^2\vec{v}$, which has solutions:

$$\vec{v} = a(0, 0, 0, 1, 0, 1)^T + b(0, 0, 0, -2, 1, 0)^T + c(4, 1, 0, 0, 0, 0)^T + d(0, 0, 0, 0, 0, 1)^T + e(1, 0, 1, 0, 0, 0)^T$$

We can now apply algorithm 2 to the eigenvalue $\lambda_2 = 4$.

Two eigenvectors of rank 2, that together with a basis of $V_{4,<1}$ span V_4 are $(0, 0, 0, 0, 0, 1)^T$, and $(1, 0, 1, 0, 0, 0)^T$ (Note that we have made a choice here).

Hence in the first step of algorithm 2 (using the above choice) we get

$$\mathcal{B} = ((4, 1, 0, 0, 0, 0)^T, (1, 0, 1, 0, 0, 0)^T, (0, 0, 0, 0, 1, 2)^T, (0, 0, 0, 0, 0, 1)^T)$$

Note that $A(1, 0, 1, 0, 0, 0)^T = 4(1, 0, 1, 0, 0, 0)^T + (4, 1, 0, 0, 0, 0)^T$, $A(4, 1, 0, 0, 0, 0)^T = 4(4, 1, 0, 0, 0, 0)^T$, $A(0, 0, 0, 0, 0, 1)^T = 4(0, 0, 0, 0, 0, 1)^T + (0, 0, 0, 0, 1, 2)^T$, and $A(0, 0, 0, 0, 1, 2)^T = 4(0, 0, 0, 0, 1, 2)^T$.

Note that $\mathcal{B}$ does *not* span V_4. Hence there is another step to this algorithm. Namely we can pick a rank one generalized $(\lambda = 4)$-eigenvector that is not in this span, e.g. $(0, 0, 0, 1, 0, 1)^T$.

Hence in step 4 (with the above choice) we get the set

$$\mathcal{B} = ((4, 1, 0, 0, 0, 0)^T, (0, 0, 0, 0, 0, 1)^T, (1, 0, 1, 0, 0, 0)^T, (0, 0, 0, 0, 1)^T, (0, 0, 0, 1, 0, 1)^T)$$

This is a basis of V_4, and hence at this point the iteration stops, and algorithm 2 concludes.

Now with respect to the basis

$$\mathcal{D} = ((1, 0, 0, 0, 0, 0)^T, (4, 1, 0, 0, 0, 0)^T, (0, 0, 0, 0, 0, 1)^T, (1, 0, 1, 0, 0, 0)^T, (0, 0, 0, 0, 1)^T, (0, 0, 0, 1, 0, 1)^T)$$

(the $\mathcal{B}$ produced in the final iteration of the last two steps in algorithm 2, combined with the eigenvector for the eigenvalue $\lambda_1 = 3$) of $\mathbb{R}^6$ we have that in this basis the matrix for A is

$$[A]_{\mathcal{D}} = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{pmatrix}$$

Now let P be the change of basis matrix from $\mathcal{D}$ to the standard basis of $\mathbb{R}^6$. That is P is the matrix with columns the vectors of $\mathcal{D}$,

$$P = \begin{pmatrix} 1 & 4 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 1 \end{pmatrix}$$

We then have

$$A = P \begin{pmatrix} 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{pmatrix} P^{-1}.$$