

Jordan Decomposition Theorem: LinearAlg

Jonathan L.F. King
 University of Florida, Gainesville FL 32611-2082, USA
 squash@ufl.edu
 Webpage <http://squash.1gainesville.com/>
 9 November, 2023 (at 00:40)

ABSTRACT: Gives a home-grown proof of the Jordan Decomposition Theorem. (Some of the lemmas work in Hilbert space.) The “Partial-form JCF Theorem”, (26), needs to be reworked.

Prolegomenon

Our goal is to prove the “JCF” (Jordan Canonical Form) Theorem for a linear trn $T: \mathbf{H} \rightarrow \mathbf{H}$, where \mathbf{H} is a finite-dim’al vectorspace. Formally, we’ll assume that \mathbf{H} is $\mathbb{F}^{\times \mathcal{H}}$, where the field \mathbb{F} is either \mathbb{R} or \mathbb{C} .

For vectorspaces use

vectorspace: $\mathbf{H}, \mathbf{A}, \mathbf{B}, \mathbf{E}, \mathbf{V}$
 dimension: $\mathcal{H}, \mathcal{A}, \mathcal{B}, \mathcal{E}, \mathcal{V}$.

Use sans-serif font for matrices $\mathbf{A}, \mathbf{B}, \mathbf{G}, \mathbf{I}, \mathbf{M}$. For square matrices \mathbf{A}_e , let $\text{Diag}(\mathbf{A}_1, \dots, \mathbf{A}_\mathcal{E})$ be the partitioned matrix which has $\mathbf{A}_1, \dots, \mathbf{A}_\mathcal{E}$ along its diagonal, and zeros elsewhere.

1: Notation. A collection $\mathcal{C} := \{\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_L\}$ of subspaces of \mathbf{H} is *linearly independent* (abbreviation *lin-indep*) if the only soln to

$$\mathbf{v}_1 + \dots + \mathbf{v}_L = \mathbf{0}, \quad \text{with each } \mathbf{v}_\ell \in \mathbf{V}_\ell,$$

is the trivial soln $\mathbf{v}_1 = \mathbf{0}, \dots, \mathbf{v}_L = \mathbf{0}$.

Recall that a subspace $\mathbf{V} \subset \mathbf{H}$ is T -*invariant* if $T(\mathbf{V}) \subset \mathbf{V}$.

I’ll use *eVal*, *eVec* and *eSpace* for *eigenvalue*, *eigenvector* and *eigenspace*. \square

§1 Examining nilpotent case

In sections §1 and §2, “eVal” means the eigenvalue **zero**, and “eVec” means an eigenvector with eVal **zero**.

2: Defn. W.r.t T , a vector \mathbf{v} is *nilpotent* if $T^d(\mathbf{v}) = \mathbf{0}$ for some posint d . Indeed, the T -*depth* of a vector \mathbf{v} , written $T\text{-Depth}(\mathbf{v})$, is the infimum of all natnums n for which $T^n(\mathbf{v})$ is $\mathbf{0}$. The zero-vector has depth 0. An eVec for eVal=0 has depth 1. (A non-nilpotent vector has depth ∞ .)

Use $\text{Nil}(T)$ for the *nilspace* of T ; it comprises the set of finite-depth vectors. So

$$\text{Nil}(T) \stackrel{\text{def}}{=} \bigcup_{n=1}^{\infty} \text{Ker}(T^n) \stackrel{\text{note}}{\supset} \text{Ker}(T).$$

Transformation T is *nilpotent* if there exists a posint D such that $T^D = \mathbf{0}$. Since \mathbf{H} is finite dimensional, T is nilpotent iff $\text{Nil}(T) = \mathbf{H}$. \square

3: Depth Lemma (preliminary). Consider a sum

$$3': \quad \mathbf{v}_1 + \mathbf{v}_2 + \dots + \mathbf{v}_L$$

whose depths satisfy

$$d_1 > d_2 > \dots > d_L.$$

Then the depth of (3') is d_1 . **Proof.** Exercise. \diamond

A *downtup* (“down tuple”) $\vec{D} = (D_1, \dots, D_\mathcal{E})$ is a sequence of integers with

$$4: \quad D_1 \geq D_2 \geq \dots \geq D_\mathcal{E} \geq 1.$$

The *size* of \vec{D} is the sum $D_1 + \dots + D_\mathcal{E}$.

A posint D determines a $D \times D$ **Jordan Block** matrix

$$5: \quad \mathbb{J}\mathbf{B}(D) := \begin{bmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ & 0 & 0 & \ddots & \\ & & \ddots & \ddots & 0 \\ & & & 0 & 1 \end{bmatrix}$$

with zeros on the diagonal and ones on the first off-diagonal. Every undisplayed position is zero.

6: Nilpotent JCF Theorem. A nilpotent $T: \mathbb{F}^{\times \mathcal{H}}$ has a unique *downtup* \vec{D} so that

$$7: \quad \mathbf{M} = \mathbf{M}(\vec{D}) := \text{Diag}(\mathbb{J}\mathbf{B}(D_1), \dots, \mathbb{J}\mathbf{B}(D_\mathcal{E}))$$

is the matrix of T w.r.t some basis. In particular, $\text{Size}(\vec{D})$ equals \mathcal{H} . \diamond

Remark. In general, the above basis is not unique. The theorem can be restated ITOf matrices. A nilpotent \mathbb{F} -matrix \mathbf{M}' determines a unique *downtup* \vec{D} so that, with \mathbf{M} from (7),

$$\mathbf{M}' = \mathbf{G}^{-1} \cdot \mathbf{M}(\vec{D}) \cdot \mathbf{G},$$

for some invertible \mathbb{F} -matrix \mathbf{G} . \square

Temporarily letting $\mathbf{c}^1, \dots, \mathbf{c}^D$ denote the standard basis, notice that the $D \times D$ jordan-blk (5) acts on the standard basis by sending $\mathbf{c}^D \rightarrow \dots \rightarrow \mathbf{c}^1 \rightarrow \mathbf{0}$. Let this motivate our definition of a *chain*: a sequence $\mathbf{C} = (\mathbf{c}^d)_{d=1}^D$ of vectors, with $D \geq 1$, fulfilling

$$8: \quad \mathbf{0} \xleftarrow{\mathbf{T}} \mathbf{c}^1 \xleftarrow{\mathbf{T}} \mathbf{c}^2 \xleftarrow{\mathbf{T}} \dots \xleftarrow{\mathbf{T}} \mathbf{c}^D.$$

Furthermore $\mathbf{c}^1 \neq \mathbf{0}$, i.e., \mathbf{c}^1 is an eigenvector. Equivalently, the depth of each \mathbf{c}^d is d .

Calling the (8)-chain \mathbf{C} , let $\text{Depth}(\mathbf{C}) := D$.

9: Lemma. Consider a chain \mathbf{C} as in (8), Then the eigenspace in $\text{Spn}(\mathbf{C})$ is just 1-dimensional. Further, \mathbf{C} is a basis for $\text{Spn}(\mathbf{C})$ and so the dimension of $\text{Spn}(\mathbf{C})$ is D . *Proof.* The Depth Lemma. \diamond

A *chain complex* $\vec{\mathbf{C}}$ for \mathbf{T} is a sequence $\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_\mathcal{E}$ of \mathbf{T} -chains such that \vec{D} is a down tuple, (4), where $D_e := \text{Depth}(\mathbf{C}_e)$. Furthermore

10: The list $\mathbf{c}_1^1, \mathbf{c}_2^1, \mathbf{c}_3^1, \dots, \mathbf{c}_\mathcal{E}^1$ of eigenvectors is linearly independent.

The *downtup* \vec{D} is called the *signature* of $\vec{\mathbf{C}}$.

By the way, we call $\vec{\mathbf{C}}$ a “*spanning* chain-complex” if $\bigsqcup_{e=1}^{\mathcal{E}} \mathbf{C}_e$ is a *basis* for \mathbf{H} . Courtesy the next lemma, the chain-complex spans iff $D_1 + \dots + D_\mathcal{E}$ equals $\text{Dim}(\mathbf{H})$.

11: Chain Independence Lemma. Suppose that $\mathbf{C}_1, \dots, \mathbf{C}_\mathcal{E}$ are chains (of possibly different lengths). Then TFAE equivalent.

a: The list of eigenvectors $\mathbf{c}_1^1, \mathbf{c}_2^1, \mathbf{c}_3^1, \dots, \mathbf{c}_\mathcal{E}^1$ is linearly independent.

b: The list $\text{Spn}(\mathbf{C}_1), \text{Spn}(\mathbf{C}_2), \dots, \text{Spn}(\mathbf{C}_\mathcal{E})$ of subspaces is linearly independent.

c: The disjoint union $\bigsqcup_{e=1}^{\mathcal{E}} \mathbf{C}_e$ is a lin-indep set. \diamond

Proof. That (b) \Rightarrow (c) follows from Lemma 9. The interesting implication is (a) \Rightarrow (b).

Were the subspaces dependent, then we could find vectors $\mathbf{v}_e \in \text{Spn}(\mathbf{C}_e)$, not all $\mathbf{0}$, so that

$$†: \quad \mathbf{0} = \mathbf{v}_1 + \mathbf{v}_2 + \dots + \mathbf{v}_{\mathcal{E}-1} + \mathbf{v}_\mathcal{E}.$$

Let D be the maximum of $\text{Depth}(\mathbf{v}_e)$, taken over $e = 1, \dots, \mathcal{E}$. Replacing each \mathbf{v}_e by $\mathbf{T}^{D-1}(\mathbf{v}_e)$ arranges that: Each \mathbf{v}_e is either an eigenvector or is $\mathbf{0}$, and not all are $\mathbf{0}$.

Courtesy (9), the eigenspace in each $\text{Spn}(\mathbf{C}_e)$ is 1-dim’al, so \mathbf{v}_e is a multiple of \mathbf{c}_e^1 . But, by hypothesis, these evecs are linearly independent, \times . (I use \times for “contradiction”). \diamond

The above proof allows us to jazz up an earlier lemma. For a nilpotent vector \mathbf{v} of depth $d \geq 1$, call $\mathbf{T}^{d-1}(\mathbf{v})$ its *penultimate vector*. [This penultimate vector is an eVec.]

12: Depth Corollary. Let D be the max depth of some list $\mathbf{v}_1, \dots, \mathbf{v}_\mathcal{E}$ of vectors. For those vectors of depth D , suppose the set of their penultimate vectors is linearly-independent.

Then $\text{Depth}(\mathbf{v}_1 + \dots + \mathbf{v}_\mathcal{E}) = D$. \diamond

The Construction

To establish the existence part of Nilpotent JCF Theorem, we fabricate a spanning chain-complex for our nilpotent \mathbf{T} .

Pick a maximum-length chain \mathbf{C}_1 . Look at the lengths of those chains whose eVec is not in $\text{Spn}(\mathbf{C}_1)$; among those having the maximum length, take one such chain and call it \mathbf{C}_2 . Pick a maximum-length chain, call it \mathbf{C}_3 , from those whose eVec is not in $\text{Spn}(\mathbf{C}_1 \sqcup \mathbf{C}_2)$. Continuing, produces a sequence of some \mathcal{E} many

chains $\mathbf{C}_1, \dots, \mathbf{C}_{\mathcal{E}}$. By construction, the eigenvectors $\mathbf{c}_1^1, \dots, \mathbf{c}_{\mathcal{E}}^1$ are linearly indep., and their *lengths* satisfy $D_1 \geq D_2 \geq \dots \geq D_{\mathcal{E}}$.

Our goal is to show that $\text{Spn}(\mathbf{C}_1 \sqcup \dots \sqcup \mathbf{C}_{\mathcal{E}})$ is all of \mathbf{H} . [Corollarily, (11)&(12) will imply that \mathcal{E} is the dimension of the eigenspace of \mathbf{T} .]

Proof. If $\mathbf{V} := \text{Spn}(\mathbf{C}_1, \dots, \mathbf{C}_{\mathcal{E}})$ is not all of \mathbf{H} , then there exists a “bad” vector \mathbf{b} i.e:

$$13: \quad \mathbf{b} \notin \mathbf{V}, \text{ yet } \mathbf{T}(\mathbf{b}) \in \mathbf{V}.$$

[This, since \mathbf{T} is nilpotent.] Thus we can write $\mathbf{T}(\mathbf{b})$ as a lin.comb over the \mathbf{V} -basis $\bigsqcup_1^{\mathcal{E}} \mathbf{C}_e$. Because it will make no difference to the following argument, I will assume, in the expansion of $\mathbf{T}(\mathbf{b})$, that each non-zero coeff is 1.

Suppose, for example, that

$$14: \quad \mathbf{T}(\mathbf{b}) = [\mathbf{c}_3^5 + \mathbf{c}_3^9 + \mathbf{c}_3^{13} + \mathbf{c}_3^{47}] + [\mathbf{c}_6^1 + \mathbf{c}_6^4 + \mathbf{c}_6^9] + [\mathbf{c}_8^2 + \mathbf{c}_8^{31}] + \dots + [\mathbf{c}_{73}^2 + \mathbf{c}_{73}^{14}].$$

Consider the \mathbf{c}_3^5 term. It has a predecessor on its chain, since $5 < D_3$. (After all, D_3 is at least 47.) Hence *replacing* the bad vector \mathbf{b} by $[\mathbf{b} - \mathbf{c}_3^5]$ preserves (13) and arranges that the \mathbf{T} -image of this *new* \mathbf{b} has one fewer term in (14). Only the **chain-end** $\mathbf{c}_e^{D_e}$ of a chain \mathbf{C}_e cannot be so removed.

Continue this until there are only chain-ends. For example, ^① suppose that the new \mathbf{b} maps to

$$14': \quad \mathbf{T}(\mathbf{b}) = \mathbf{c}_3^{D_3} + \mathbf{c}_8^{D_8} + \dots + \mathbf{c}_{73}^{D_{73}}.$$

Corollary 12 tells us that vector $\mathbf{T}(\mathbf{b})$ has depth ^② $\text{Max}(D_3, D_8, \dots)$ —which is D_3 . Consequently:

The depth of \mathbf{b} is $1+D_3$.

But all the vectors in RhS(14') were chosen, during “The Construction”, at stages 3 and after. So D_3 was *not* in fact the length of the longest available chain. ✖

Uniqueness of signature. A spanning chain-complex must have exactly $\mathcal{E} = \text{Dim}(\mathbf{E})$ chains, where \mathbf{E} is the eigenspace of \mathbf{T} . Although the spanning chain-complex itself is not unique nonetheless its *signature* is unique —the Nilpotent JCF Thm asserts this, so I’d better prove it!

15: Lemma. *Given a nilpotent \mathbf{T} , all spanning chain-complexes have the same signature.* ◊

Proof. Consider two spanning chain-complexes

$$\begin{aligned} \vec{\mathbf{C}} &= (\mathbf{C}_1, \dots, \mathbf{C}_{\mathcal{E}}), \\ \vec{\mathbf{C}}^\bullet &= (\mathbf{Q}_1, \dots, \mathbf{Q}_{\mathcal{E}}), \end{aligned}$$

with different signatures. For specificity, suppose that the two signatures differ in their third term as follows:

$$\begin{aligned} D_1 &\geq D_2 \geq 9 = D_3 \geq D_4 \geq \dots \geq D_{\mathcal{E}}; \\ D_1 &\geq D_2 \geq 8 \geq D_3^\bullet \geq D_4^\bullet \geq \dots \geq D_{\mathcal{E}}^\bullet. \end{aligned}$$

Now write vector \mathbf{c}_1^9 over the $\vec{\mathbf{C}}^\bullet$ -basis. This lin-comb must have a vector of depth 9 and none of greater depth. The only depth-9 vectors in $\vec{\mathbf{C}}^\bullet$ lie in chains \mathbf{Q}_1 and \mathbf{Q}_2 . So our lin-comb has form

$$\dagger: \quad \mathbf{c}_1^9 = \alpha_1 \mathbf{q}_1^9 + \alpha_2 \mathbf{q}_2^9 + \mathbf{u},$$

for some scalars α_1, α_2 and some vector \mathbf{u} [in $\text{Spn}(\vec{\mathbf{C}}^\bullet)$] whose depth is at most 8; this, courtesy (12). Applying \mathbf{T}^8 to (†) thus tells us that

$$\ddagger: \quad \mathbf{c}_1^1 \in \text{Spn}(\mathbf{q}_1^1, \mathbf{q}_2^1).$$

Repeating the argument twice more gives

$$\{\mathbf{c}_1^1, \mathbf{c}_2^1, \mathbf{c}_3^1\} \subset \text{Spn}(\mathbf{q}_1^1, \mathbf{q}_2^1).$$

But a 3-dim’al space won’t fit inside a 2-dim’al space. ◆

Having proved JCF in the nilpotent case, (6), we now develop the tools for the general case.

§2 Algebraic information

The *characteristic poly* of an $\mathcal{M} \times \mathcal{M}$ matrix \mathbf{M} is

$$\wp_{\mathbf{M}}(x) := \text{Det}(x\mathbf{I} - \mathbf{M}).$$

So $\wp_{\mathbf{M}}$ is a deg- \mathcal{M} poly. [Some textbooks define $\wp_{\mathbf{M}}(x)$ as $\text{Det}(\mathbf{M} - x\mathbf{I})$]

16: **Lemma.** *For a $\mathcal{B} \times \mathcal{B}$ matrix \mathbf{B} ,*

$$\text{Ker}(\mathbf{B}) \text{ is trivial} \iff \wp_{\mathbf{B}}(0) \neq 0,$$

i.e, IFF $\wp_{\mathbf{B}}()$ has a [non-zero] constant term. ◇

Suppose $\mathbf{A}, \mathbf{B} \subset \mathbf{H}$ is a lin-indep pair of subspaces, which jointly span \mathbf{H} . Indicate this by writing

$$\mathbf{A} \oplus \mathbf{B} = \mathbf{H}.$$

Let $\text{Proj}_{\mathbf{B}}^{\mathbf{A}}$ be projection, parallel to \mathbf{A} , from \mathbf{H} onto \mathbf{B} . Said differently, put an inner-product on \mathbf{H} making $\mathbf{A} \perp \mathbf{B}$. Then $\text{Proj}_{\mathbf{B}}^{\mathbf{A}}$ is simply the orthogonal projection $\text{Proj}_{\mathbf{B}}$.

An arbitrary linear trn $\mathbf{T}: \mathbf{H} \circ$ gives a composition

$$* : \mathbf{B} \xleftarrow{\text{Proj}_{\mathbf{B}}^{\mathbf{A}}} \mathbf{H} \xleftarrow{\mathbf{T}} \mathbf{H}$$

Let “ $\mathbf{T}_{\mathbf{B}}^{\mathbf{A}}$ ” denote the restriction of $(*)$ to \mathbf{B} , i.e, the mapping $\mathbf{B} \circ$ by

$$\mathbf{T}_{\mathbf{B}}^{\mathbf{A}} := [\text{Proj}_{\mathbf{B}}^{\mathbf{A}} \circ \mathbf{T}] \downarrow \mathbf{B}.$$

17: **Block-UT-matrix Lemma.** *Consider an upper-triangular partitioned matrix*

$$18: \quad \mathbf{M} = \begin{bmatrix} \mathbf{A}_{\mathcal{A} \times \mathcal{A}} & \mathbf{G}_{\mathcal{A} \times \mathcal{B}} \\ \mathbf{0}_{\mathcal{B} \times \mathcal{A}} & \mathbf{B}_{\mathcal{B} \times \mathcal{B}} \end{bmatrix}$$

Then $\text{Det}(\mathbf{M}) = \text{Det}(\mathbf{A}) \cdot \text{Det}(\mathbf{B})$. In consequence, the char-poly $\wp_{\mathbf{M}}$ factors as

$$18': \quad \wp_{\mathbf{M}} = \wp_{\mathbf{A}} \cdot \wp_{\mathbf{B}}.$$

Restated, suppose $\mathbf{T}: \mathbf{H} \circ$ has subspaces \mathbf{A}, \mathbf{B} st.

$$19: \quad \mathbf{H} = \mathbf{A} \oplus \mathbf{B}.$$

with subspace \mathbf{A} being \mathbf{T} -invariant. Then

$$19': \quad \wp_{\mathbf{T}} = \wp_{\mathbf{T} \downarrow \mathbf{A}} \cdot \wp_{\mathbf{T}_{\mathbf{B}}^{\mathbf{A}}}.$$
 ◇

Proof of $\text{Det}(\mathbf{M}) = \text{Det}(\mathbf{A}) \cdot \text{Det}(\mathbf{B})$. Since $\text{Det}(\mathbf{M})$ is a sum of products taken over all transversals of \mathbf{M} , ISTS that a transversal straying from the \mathbf{A}, \mathbf{B} blocks necessarily has product zero.

WLOG this misguided transversal hits \mathbf{G} . It therefore misses some row of \mathbf{A} hence (since \mathbf{A} is square) some column of \mathbf{A} . In this column, then, the transversal must hit the $\mathbf{0}_{\mathcal{B} \times \mathcal{A}}$ block.

Exer: Why do the signs of the permutations work out correctly? ◆

Proof of (19'). Let $\mathbf{B} = (\mathbf{a}_1, \dots, \mathbf{a}_{\mathcal{A}}, \mathbf{b}_1, \dots, \mathbf{b}_{\mathcal{B}})$ be a basis for \mathbf{H} , with each $\mathbf{a}_i \in \mathbf{A}$ and $\mathbf{b}_j \in \mathbf{B}$. Then \mathbf{M} , the \mathbf{B} -matrix of \mathbf{T} , has form (18). Furthermore, the $(\mathbf{a}_1, \dots, \mathbf{a}_{\mathcal{A}})$ -matrix of $\mathbf{T} \downarrow \mathbf{A}$ is \mathbf{A} and the $(\mathbf{b}_1, \dots, \mathbf{b}_{\mathcal{B}})$ -matrix of $\mathbf{T}_{\mathbf{B}}^{\mathbf{A}}$ is \mathbf{B} . Hence $(18') \Rightarrow (19')$. ◆

Recall from (2) the defn of *nilpotent* and $\text{Nil}(\mathbf{T})$.

20: **Lemma.** *Consider a nilpotent $\mathbf{S}: \mathbf{F}^{\mathcal{A}} \circ$. Then*

$$\wp_{\mathbf{S}}(x) = [x - 0]^{\mathcal{A}}.$$
 ◇

Pf. The char-poly of a eVal=0 Jordan-Block (5) is x^D . By the Block-UT-matrix Lemma, the char-poly of $\text{Diag}(\mathbf{J}\mathbf{B}(D_1), \dots, \mathbf{J}\mathbf{B}(D_{\mathcal{E}}))$ is the product $x^{D_1} \dots x^{D_{\mathcal{E}}}$. i.e $x^{D_1 + \dots + D_{\mathcal{E}}}$. ◆

21: **Multiplicity Theorem.** *Let $\mathbf{A} := \text{Nil}(\mathbf{T})$. Then $\mathcal{A} := \text{Dim}(\mathbf{A})$ is the multiplicity of 0 in the characteristic poly $\wp_{\mathbf{T}}$, i.e,*

$$\wp_{\mathbf{T}}(x) = [x - 0]^{\mathcal{A}} \cdot g(x),$$

where g is a poly with a constant term. ◇

Proof. Let \mathbf{B} be a complementary subspace $\mathbf{B} \oplus \mathbf{A} = \mathbf{H}$. Then (19') and (20) tell us that

$$\wp_{\mathbf{T}}(x) = [x - 0]^{\mathcal{A}} \cdot \wp_{\mathbf{T}_{\mathbf{B}}^{\mathbf{A}}}(x).$$

Consequently, courtesy (16), ISTProve that T_B^A has no kernel. So fix a $\mathbf{v} \in B$ sent to $\mathbf{0}$ by T_B^A .

Decompose its image as $T(\mathbf{v}) = \mathbf{b} + \mathbf{a}$, with $\mathbf{b} \in B$ and $\mathbf{a} \in A$. Then $\mathbf{0} = T_B^A(\mathbf{v}) \stackrel{\text{note}}{=} \mathbf{b}$. Hence $T(\mathbf{v}) = \mathbf{a}$. So $T(\mathbf{v})$ is nilpotent. Thus \mathbf{v} too is nilpotent. So $\mathbf{v} \in A \cap B$ and is therefore $\mathbf{0}$. \diamond

26: Partial-form JCF Theorem. Given linear $T: F^{\times H} \circlearrowright$, factor its char-poly over F as

$$\wp_T(x) = [x - \lambda_1]^{S_1} \cdots [x - \lambda_L]^{S_L} \cdot g(x),$$

where g is an F -poly with no roots in F . (And $\lambda_1, \dots, \lambda_L \in F$ are distinct.) Then there is a unique list of downtups, $\vec{D}^1, \dots, \vec{D}^L$. Unfinished: as of 9Nov2023 \diamond

§3 Using all the eigenvalues

For $\lambda \in \mathbb{C}$, we now return to using “ λ -eVec” to mean an eigenvector with eigenvalue λ , and we extend our defns to other evals.

An λ -**Jordan Block** is a $D \times D$ matrix

$$22: \quad \lambda\text{-JB}(D) := \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & 1 & \\ & & & \ddots & \ddots \\ & & & & \lambda & 1 \\ & & & & & \lambda \end{bmatrix}.$$

Generalizing, a downtuple $\vec{D} = (D_1, \dots, D_\varepsilon)$ engenders an λ, \vec{D} -**Jordan-Block-System**

$$23: \quad \lambda\text{-JB}(\vec{D}) := \text{Diag}(\lambda\text{-JB}(D_1), \dots, \lambda\text{-JB}(D_\varepsilon))$$

24: Jordan Canonical Form Theorem. Suppose that $T: F^{\times H} \circlearrowright$ has all of its eigenvalues $\lambda_1, \dots, \lambda_L$ in F . Then there is a unique list of downtups, $\vec{D}^1, \vec{D}^2, \dots, \vec{D}^L$, so that

$$25: \quad \text{Diag}(\lambda_1\text{-JB}(\vec{D}^1), \dots, \lambda_L\text{-JB}(\vec{D}^L))$$

is the matrix of T relative to some basis.

In particular, $\text{Size}(\vec{D}^\ell) = S_\ell$, where

$$\wp_T(x) = [x - \lambda_1]^{S_1} \cdot [x - \lambda_2]^{S_2} \cdots [x - \lambda_L]^{S_L}.$$

is the F -factorization of the char-poly of T . \diamond

For $\alpha \in F$, let T_α abbreviate the $T - \alpha I$ transformation, and let $E_\alpha^{(d)}$ comprise the vectors of T_α -depth at most d . Evidently

$$E_\alpha^{(d)} = \text{Ker}(T_\alpha^{\circ d})$$

is a subspace, and $E_\alpha^{(1)}$ is the eigenspace (when α is an eigenvalue). Certainly

$$27: \quad \{\mathbf{0}\} = E_\alpha^{(0)} \subset E_\alpha^{(1)} \subset E_\alpha^{(2)} \subset E_\alpha^{(3)} \subset \dots \subset L_\alpha,$$

where $L_\alpha := \bigcup_{d=0}^{\infty} E_\alpha^{(d)}$ is the *nilspace*.

28: Lemma. Fix $\alpha, \beta \in \mathbb{C}$. For each $d = 0, 1, \dots$, the subspace $E_\alpha^{(d)}$ is forward-invariant under T_β . Therefore L_α is T_β -forward-invariant. \diamond

Proof. WLOG $\alpha = 0$ (replace T by $T - \alpha I$ and β by $\beta - \alpha$). Fix an order d , say $d=3$, and fix a vector $\mathbf{v} \in E^{(3)}$. Automatically $T(\mathbf{v}) \in E^{(2)} \subset E^{(3)}$. Hence $[T - \beta I]\mathbf{v} = T(\mathbf{v}) - \beta\mathbf{v} \in E^{(3)}$, as desired. \diamond

29: Lemma. Consider distinct scalars $\alpha \neq \beta$. For $d = 0, 1, 2, \dots$, the restricted operator

$$[T - \beta I] \downarrow E_\alpha^{(d)}$$

has trivial kernel and so is a (linear) automorphism of $E_\alpha^{(d)}$ (since $E_\alpha^{(d)}$ is finite-dim'el). Taking a union, then, $[T - \beta I] \downarrow L_\alpha$ is an automorphism of L_α . \diamond

Proof. WLOG $\alpha = 0$; so $\beta \neq 0$. By the preceding lemma, T_β maps $\mathbf{E}^{\langle d \rangle}$ into $\mathbf{E}^{\langle d \rangle}$. So FTSOContra-diction we may suppose that there is a non-zero $\mathbf{v} \in \mathbf{E}^{\langle d \rangle}$ which is sent to $\mathbf{0}$ by T_β .

Evidently $T \leq T_\beta$. For $j = 0, 1, 2, \dots$, consequently, the vector $T^j(\mathbf{v})$ is also in $\text{Ker}(T_\beta)$. Consider the value of j for which $T^j(\mathbf{v})$ is in $\mathbf{E}^{\langle 1 \rangle} \setminus \{\mathbf{0}\}$. Redefining \mathbf{v} to be this $T^j(\mathbf{v})$, we now have that

\mathbf{v} is a non-zero vector simultaneously in $\text{Ker}(T)$ and $\text{Ker}(T_\beta)$.

But then $\mathbf{0} = T_\beta(\mathbf{v}) = T(\mathbf{v}) - \beta\mathbf{v} = -\beta\mathbf{v}$. And this latter is not zero, since $\beta \neq 0$. \diamond

30: Prop'n. Let $\lambda_1, \dots, \lambda_L$ be the distinct eigenvalues of T . Then the collection $\mathbf{L}_{\lambda_1}, \dots, \mathbf{L}_{\lambda_L}$ of nilspaces is linearly independent. \diamond

Proof. Consider a sum $\mathbf{v}_1 + \dots + \mathbf{v}_L = \mathbf{0}$, with each $\mathbf{v}_\ell \in \mathbf{L}_{\lambda_\ell}$. ISTShow that $\mathbf{v}_1 = \mathbf{0}$. So ISTConstruct a linear $\Lambda: \mathbf{H} \rightarrow \mathbf{H}$ sending each of $\mathbf{v}_2, \dots, \mathbf{v}_L$ to $\mathbf{0}$, so that $\Lambda|_{\mathbf{L}_{\lambda_1}}$ is an automorphism of \mathbf{L}_{λ_1} .

To this end, pick a number D large enough that

$$[T_{\lambda_\ell}]^{\circ D}(\mathbf{v}_\ell) = \mathbf{0}, \quad \text{for each } \ell = 2, 3, \dots, L.$$

Since all the operators $(T_\alpha)_{\alpha \in \mathbb{C}}$ commute, it follows that the composition

$$\Lambda := [T_{\lambda_2} \circ T_{\lambda_3} \circ \dots \circ T_{\lambda_L}]^{\circ D}$$

sends each of $\mathbf{v}_2, \dots, \mathbf{v}_L$ to $\mathbf{0}$. And Lemma 29 assures us that Λ is an automorphism of \mathbf{L}_{λ_1} . \diamond

Proof of JCF, (24). Apply the Nilpotent JCF to T_{λ_ℓ} on $\mathbf{L}_{\lambda_\ell}$ to get a basis \mathcal{B}_ℓ for $\mathbf{L}_{\lambda_\ell}$ against which T has a matrix-block of form $\lambda_\ell \cdot \mathbf{J}\mathbf{B}(\vec{D}^\ell)$. Then $\bigsqcup_{\ell=1}^L \mathcal{B}_\ell$ is a basis against which T looks like (25). That the downtup sequence is unique follows from the uniqueness in Nilpotent thm and that T uniquely determines its nilspaces. \diamond

End Notes

See `cayley_hamilton.latex` for several applications of JCF, and the minimum polynomial of a matrix.

Transposes. Let J be a jordan block. Its transpose J^t is conjugate to J simply by reversing the order of the vectors in the basis. It follows that

$$JCF(T^t) = JCF(T)$$

for an arbitrary square matrix T . \square

Remark. Fix a square matrix T . Given a scalar λ , let $\vec{D}_T(\lambda)$ be the corresponding downtup in $JCF(T)$; if λ is not a T -eVal, then the downtup is empty.

The *complex-conjugate* of a JCF is a JCF. So $JCF(\bar{T}) = \bar{J}$. This gives the (\Rightarrow) direction below.

31: JCF-of-real Theorem. A complex JCF B is the JCF of some real matrix IFF

$$\vec{D}_B(\lambda) = \vec{D}_B(\bar{\lambda}),$$

for each complex number λ . \diamond

Proof of (\Leftarrow) . ISTProve this when B consists of a jordan block and its complex-conjugate. For specificity suppose that each jordan block has dimension $D=3$.

Fix reals c and s , and let

$$\lambda^+ := c + is \quad \text{and} \quad \lambda^- := c - is.$$

We show that $B := \text{Diag}(\lambda^+ \cdot \mathbf{J}\mathbf{B}(3), \lambda^- \cdot \mathbf{J}\mathbf{B}(3))$ is conjugate to a real matrix by producing a basis

$$32: \quad \{\mathbf{u}_j, \mathbf{w}_j\}_{j=1}^3$$

for $\mathbb{C}^{6 \times 6}$, against which B acts using only real coefficients.

For each choice of “+” and “-”, let $\{\mathbf{e}_j^\pm\}_{j=1}^3$ be the std basis for $\lambda^\pm \cdot \mathbf{J}\mathbf{B}(3)$. Thus for $j \in [1..3]$,

$$33: \quad B(\mathbf{e}_j^\pm) = \lambda^\pm \cdot \mathbf{e}_j^\pm + 1 \cdot \mathbf{e}_{j-1}^\pm,$$

where \mathbf{e}_0^\pm are two other names for $\mathbf{0}$.

Define new vectors

$$\begin{aligned}\mathbf{u}_j &:= 1 \cdot \mathbf{e}_j^+ + \mathbf{i} \cdot \mathbf{e}_j^-; \\ \mathbf{w}_j &:= \mathbf{i} \cdot \mathbf{e}_j^+ + 1 \cdot \mathbf{e}_j^-;\end{aligned}$$

so \mathbf{u}_0 and \mathbf{w}_0 are each $\mathbf{0}$. Check that

$$\frac{1}{2} \cdot [\mathbf{u}_j \mp \mathbf{i}\mathbf{w}_j] = \mathbf{e}_j^\pm,$$

so (32) spans all the \mathbf{e} 's. Thus (32) indeed is a basis for $\mathbb{C}^{6 \times 6}$.

The B-images of vectors. Verify that

$$\begin{aligned}c\mathbf{u}_j + s\mathbf{w}_j &= \lambda^+ \mathbf{e}_j^+ + \mathbf{i}\lambda^- \mathbf{e}_j^- \quad \text{and} \\ -s\mathbf{u}_j + c\mathbf{w}_j &= \mathbf{i}\lambda^+ \mathbf{e}_j^+ + \lambda^- \mathbf{e}_j^-.\end{aligned}$$

From (33) we compute:

$$\begin{aligned}\mathbf{B}(\mathbf{u}_j) &= 1 \cdot [\lambda^+ \cdot \mathbf{e}_j^+ + 1 \cdot \mathbf{e}_{j-1}^+] + \\ &\quad \mathbf{i} \cdot [\lambda^- \cdot \mathbf{e}_j^- + 1 \cdot \mathbf{e}_{j-1}^-].\end{aligned}$$

Grouping terms by subscript, our $\mathbf{B}(\mathbf{u}_j)$ equals

$$[\lambda^+ \mathbf{e}_j^+ + \mathbf{i}\lambda^- \mathbf{e}_j^-] + [1 \cdot \mathbf{e}_{j-1}^+ + \mathbf{i} \cdot \mathbf{e}_{j-1}^-].$$

This, together with similar elbow grease, yields

$$\begin{aligned}\mathbf{B}(\mathbf{u}_j) &= [c\mathbf{u}_j + s\mathbf{w}_j] + \mathbf{u}_{j-1}; \\ \mathbf{B}(\mathbf{w}_j) &= [-s\mathbf{u}_j + c\mathbf{w}_j] + \mathbf{w}_{j-1}.\end{aligned}$$

Since all the coefficients are real, we get that \mathbf{B} is conjugate to a real matrix. ◆

Cyclic decompositions. The (forward) *cyclic subspace* generated by \mathbf{v} is

$$\text{Spn}(\mathbf{v}, \mathbf{T}\mathbf{v}, \mathbf{T}^2\mathbf{v}, \mathbf{T}^3\mathbf{v}, \dots).$$

And \mathbf{T} is a *cyclic operator* if there is a \mathbf{v} whose cyclic subspace is all of \mathbf{V} .

Easily, a jordan-block is a cyclic operator on its space. So the jordan decomposition of \mathbf{T} yields a \mathbf{T} -cyclic decomposition of the vectorspace.

Yo! Look in source file, here. □

Filename: Problems/Algebra/LinearAlg/jordan_decomp.latex
As of: Thursday 13Apr2006. Typeset: 9Nov2023 at 00:40.