Jordan Decomposition Theorem: LinearAlg

Jonathan L.F. King
University of Florida, Gainesville FL 32611-2082, USA
squash@ufl.edu
Webpage http://squash.1gainesville.com/
20 March, 2018 (at *16:50*)

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 $\mathsf{T}:\mathbf{H}\to\mathbf{H}$, where $\mathbf{H}$ is a finite-dim'al vectorspace. Formally, we'll assume that $\mathbf{H}$ is $\mathbf{F}^{\times\mathcal{H}}$, where the field $\mathbf{F}$ is either $\mathbb{R}$ or $\mathbb{C}$.

For vectorspaces use

$$\begin{aligned}\text{vectorspace:}\quad & \mathbf{H},\mathbf{A},\mathbf{B},\mathbf{E},\mathbf{V}\\ \text{dimension:}\quad & \mathcal{H},\mathcal{A},\mathcal{B},\mathcal{E},\mathcal{V}\,.\end{aligned}$$

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

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

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

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

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

I'll use ***eval***, ***evec*** and ***e-space*** for *eigenvalue*, *eigenvector* and *eigenspace*. □

§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 $\mathsf{T}^d(\mathbf{v})=\mathbf{0}$ for some posint d. Indeed, the T-***depth*** of a vector $\mathbf{v}$, written $\mathsf{T}\text{-Depth}(\mathbf{v})$, is the infimum of all natnums n for which $\mathsf{T}^n(\mathbf{v})$ is $\mathbf{0}$. The zero-vector has depth 0. An evec for eval=0 has depth 1. (A non-nilpotent vector has depth ∞.)

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

$$\mathrm{Nil}(\mathsf{T}) \overset{\text{def}}{=\!=} \bigcup_{n=1}^{\infty}\mathrm{Ker}(\mathsf{T}^n) \overset{\text{note}}{\supset} \mathrm{Ker}(\mathsf{T})\,.$$

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

3: Depth Lemma (preliminary). *Consider a sum*

3′: $$\mathbf{v}_1+\mathbf{v}_2+\cdots+\mathbf{v}_L$$

whose depths satisfy

$$d_1 > d_2 > \ldots > d_L\,.$$

Then the depth of (3′) is d_1. *Proof.* Exercise. ◇

A ***downtup*** ("down tuple") $\overrightarrow{D} = (D_1,\ldots,D_\mathcal{E})$ is a sequence of integers with

4: $$D_1\geqslant D_2\geqslant\ldots\geqslant D_\mathcal{E}\geqslant 1\,.$$

The ***size*** of $\overrightarrow{D}$ is the sum $D_1+\cdots+D_\mathcal{E}$. A posint D determines a $D{\times}D$ ***Jordan Block*** matrix

5: $$\mathsf{JB}(D) := \begin{bmatrix} 0 & 1 & & & & \\ & 0 & 1 & & & \\ & & 0 & 1 & & \\ & & & \ddots & \ddots & \\ & & & & 0 & 1 \\ & & & & & 0 \end{bmatrix}$$

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

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

7: $$\mathsf{M} = \mathsf{M}(\overrightarrow{D}) := \mathit{Diag}\Big(\mathsf{JB}(D_1),\ldots,\mathsf{JB}(D_\mathcal{E})\Big)$$

is the matrix of T *w.r.t* *some* *basis. In particular,* $\mathrm{Size}(\overrightarrow{D})$ *equals* $\mathcal{H}$. ◇

Remark. In general, the above basis is not unique.

The theorem can be restated ITOf matrices. *A nilpotent* $\mathbf{F}$*-matrix* M' *determines a unique downtup* $\overrightarrow{D}$ *so that, with* M *from* (7),

$$\mathsf{M}' \;=\; \mathsf{G}^{-1} \cdot \mathsf{M}(\overrightarrow{D}) \cdot \mathsf{G}\,,$$

for <u>*some*</u> *invertible* $\mathbf{F}$*-matrix* G. □

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

8: $$\mathbf{0} \xleftarrow{\mathsf{T}} \mathbf{c}^1 \xleftarrow{\mathsf{T}} \mathbf{c}^2 \xleftarrow{\mathsf{T}} \ldots \xleftarrow{\mathsf{T}} \mathbf{c}^D\,.$$

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.

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

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

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

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

By the way, we call $\overrightarrow{\mathbf{C}}$ 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 $D_1 + \cdots + D_{\mathcal{E}}$ equals $\mathrm{Dim}(\mathbf{H})$.

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

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

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

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

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

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

†: $$\mathbf{0} \;=\; \mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_{\mathcal{E}-1} + \mathbf{v}_{\mathcal{E}}\,.$$

Let D be the max of $\mathrm{Depth}(\mathbf{v}_e)$, taken over $e = 1, \ldots, \mathcal{E}$. Replacing each $\mathbf{v}_e$ by $\mathsf{T}^{D-1}(\mathbf{v}_e)$ 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 $\mathrm{Spn}(\mathbf{C}_e)$ is 1-dim'al, so $\mathbf{v}_e$ is a multiple of $\mathbf{c}_e^1$. But, by hypothesis, these evecs are linearly independent, ※. (I use ※ for "contradiction".) ♦

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

12: Depth Corollary. *Let* D *be the max depth of some list of vectors* $\mathbf{v}_1, \ldots, \mathbf{v}_{\mathcal{E}}$. *Suppose that those vectors of depth* D *have linearly-independent penultimate vectors. Then the sum* $\mathbf{v}_1 + \ldots + \mathbf{v}_{\mathcal{E}}$ *has depth* D. ♢

The Construction

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

Pick a maximum-length chain $\mathbf{C}_1$. Look at the lengths of those chains whose evec is <u>not</u> in $\mathrm{Spn}(\mathbf{C}_1)$; 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 $\mathrm{Spn}(\mathbf{C}_1 \sqcup \mathbf{C}_2)$. Continuing, produces a sequence of some $\mathcal{E}$ many chains $\mathbf{C}_1, \ldots, \mathbf{C}_{\mathcal{E}}$. By construction, the eigenvectors $\mathbf{c}_1^1, \ldots, \mathbf{c}_{\mathcal{E}}^1$ are linearly indep., and their *lengths* satisfy $D_1 \geqslant D_2 \geqslant \ldots \geqslant D_{\mathcal{E}}$.

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

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

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

(This, since T is nilpotent.) Thus we can write $\mathsf{T}(\mathbf{b})$ 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 $\mathsf{T}(\mathbf{b})$, that each non-zero coeff is 1.

Suppose, for example, that

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

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 $[\mathbf{b} - \mathbf{c}_3^6]$ preserves (13) and arranges that the 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′: $$\mathsf{T}(\mathbf{b}) = \mathbf{c}_3^{D_3} + \mathbf{c}_8^{D_8} + \ldots + \mathbf{c}_{73}^{D_{73}}.$$

Corollary 12 tells us that vector $\mathsf{T}(\mathbf{b})$ has depth[♡2] $\mathrm{Max}(D_3, D_8, \ldots)$ —which is D_3. Consequently:

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

But all the vectors in $\mathrm{RhS}(14')$ were chosen, during "The Construction", at stages 3 and after. So D_3 was *not* in fact the length of the longest available chain. ※. ♦

[♡1] In this example, (14) and (14′) together tell us that $D_3 = 47$, $D_6 > 9$, $D_8 = 31$, and so on.

[♡2] Since $\mathbf{b} \notin \mathbf{V}$, our $\mathbf{b}$ is not an evec; so $\mathsf{T}(\mathbf{b})$ is not $\mathbf{0}$. Hence $\mathsf{T}(\mathbf{b})$ has *at least one* chain-end.

Uniqueness of signature. A spanning chain-complex must have exactly $\mathcal{E} = \mathrm{Dim}(\mathbf{E})$ chains, where $\mathbf{E}$ is the eigenspace of 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 T, all spanning chain-complexes have the same signature.* ♢

Proof. Consider two spanning chain-complexes

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

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

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

Now write vector $\mathbf{c}_1^9$ over the $\vec{\mathbf{C}}^{\bullet}$-basis:

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

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

$$\mathbf{c}_1^1 \in \mathrm{Spn}\Big(\mathbf{q}_1^1, \mathbf{q}_2^1\Big).$$

Repeating the argument twice more gives

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

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 M is

$$\wp_{\mathsf{M}}(x) \;:=\; \operatorname{Det}(x\mathbf{I} - \mathsf{M})\,.$$

So $\wp_{\mathsf{M}}$ is a monic deg-$\mathcal{M}$ poly.

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

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

i.e, IFF $\wp_{\mathsf{B}}()$ 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 $\operatorname{Proj}^{\mathbf{A}}_{\mathbf{B}}$ 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 $\operatorname{Proj}^{\mathbf{A}}_{\mathbf{B}}$ is simply the orthogonal projection $\operatorname{Proj}_{\mathbf{B}}$. An arbitrary linear trn $\mathsf{T}{:}\mathbf{H}\circlearrowleft$ gives a composition

$$*: \qquad \mathbf{B} \xleftarrow{\operatorname{Proj}^{\mathbf{A}}_{\mathbf{B}}} \mathbf{H} \xleftarrow{\mathsf{T}} \mathbf{H}$$

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

$$\mathsf{T}^{\mathbf{A}}_{\mathbf{B}} \;:=\; [\operatorname{Proj}^{\mathbf{A}}_{\mathbf{B}} \circ \mathsf{T}]\lfloor\mathbf{B}\,.$$

17: Triangular Matrix Lemma. *Consider an upper-triangular partitioned matrix*

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

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

$$18': \qquad \wp_{\mathsf{M}} \;=\; \wp_{\mathsf{A}}\cdot\wp_{\mathsf{B}}\,.$$

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

$$19: \qquad \mathbf{H} \;=\; \mathbf{A}\oplus\mathbf{B}\,.$$

If subspace $\mathbf{A}$ is T-invariant then

$$19': \qquad \wp_{\mathsf{T}} \;=\; \wp_{\mathsf{T}\lfloor\mathbf{A}} \cdot \wp_{\mathsf{T}^{\mathbf{A}}_{\mathbf{B}}}\,.$$ ◊

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

WLOG this misguided transversal hits G. It therefore misses some row of A hence (since A is square) some *column* of 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 $\mathcal{B} = (\mathbf{a}_1,\dots,\mathbf{a}_{\mathcal{A}},\,\mathbf{b}_1,\dots,\mathbf{b}_{\mathcal{B}})$ be a basis for $\mathbf{H}$, with each $\mathbf{a}_i \in \mathbf{A}$ and $\mathbf{b}_j \in \mathbf{B}$. Then M, the $\mathcal{B}$-matrix of T, has form (18). Furthermore, the $(\mathbf{a}_1,\dots,\mathbf{a}_{\mathcal{A}})$-matrix of $\mathsf{T}\lfloor_{\mathbf{A}}$ is A and the $(\mathbf{b}_1,\dots,\mathbf{b}_{\mathcal{B}})$-matrix of $\mathsf{T}^{\mathbf{A}}_{\mathbf{B}}$ is B. Hence (18′)⇒(19′). ♦

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

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

$$\wp_{\mathsf{S}}(x) \;=\; [x-0]^{\mathcal{A}}\,.$$ ◊

Pf. The char-poly of a eval=0 Jordan-Block (5) is x^D. By the Triangular Matrix Lemma, the char-poly of $\mathit{Diag}\big(\mathbb{JB}(D_1),\dots,\mathbb{JB}(D_{\mathcal{E}})\big)$ is the product $x^{D_1}\cdots x^{D_{\mathcal{E}}}$. i.e $x^{D_1+\ldots+D_{\mathcal{E}}}$. ♦

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

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

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

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

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

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

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

Generalizing, a downtup $\overrightarrow{D} = (D_1, \ldots, D_{\mathcal{E}})$ engenders an $\lambda, \overrightarrow{D}$-***Jordan Block***

23: $$\lambda\text{-JB}(\overrightarrow{D}) := \mathit{Diag}\Big(\lambda\text{-JB}(D_1), \ldots, \lambda\text{-JB}(D_{\mathcal{E}})\Big)$$

24: Jordan Canonical Form Theorem. *Suppose that $\mathsf{T}:\mathbf{F}^{\times\mathcal{H}}\circlearrowleft$ has all of its eigenvalues $\lambda_1, \ldots, \lambda_L$ in $\mathbf{F}$. Then there is a unique list of downtups, $\overrightarrow{D}^1, \overrightarrow{D}^2, \ldots, \overrightarrow{D}^L$, so that*

25: $$\mathit{Diag}\Big(\lambda_1\text{-JB}(\overrightarrow{D}^1), \ldots, \lambda_L\text{-JB}(\overrightarrow{D}^L)\Big)$$

is the matrix of T relative to some basis.

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

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

is the $\mathbf{F}$-factorization of the char-poly of T. ♢

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

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

where g is an $\mathbf{F}$-poly with no roots in $\mathbf{F}$. (And $\lambda_1, \ldots, \lambda_L \in \mathbf{F}$ are distinct.) Then there is a unique list of downtups, $\overrightarrow{D}^1, \ldots, \overrightarrow{D}^L$, Unfinished: *as of 20Mar2018* ♢

For $\alpha \in \mathbf{F}$, let T_α abbreviate the $\mathsf{T} - \alpha\mathbf{I}$ transformation, and let $\mathbf{E}_\alpha^{\langle d\rangle}$ comprise the vectors of T_α-depth at most d. Evidently

$$\mathbf{E}_\alpha^{\langle d\rangle} = \text{Ker}(\mathsf{T}_\alpha{}^{\circ d})$$

is a subspace, and $\mathbf{E}_\alpha^{\langle 1\rangle}$ is the eigenspace (when α is an eigenvalue). Certainly

27: $$\{\mathbf{0}\} = \mathbf{E}_\alpha^{\langle 0\rangle} \subset \mathbf{E}_\alpha^{\langle 1\rangle} \subset \mathbf{E}_\alpha^{\langle 2\rangle} \subset \mathbf{E}_\alpha^{\langle 3\rangle} \subset \ldots \subset \mathbf{L}_\alpha,$$

where $\mathbf{L}_\alpha := \bigcup_{d=0}^\infty \mathbf{E}_\alpha^{\langle d\rangle}$ is the ***nilspace***.

28: Lemma. *Fix $\alpha, \beta \in \mathbb{C}$. For each $d = 0, 1, \ldots,$ the subspace $\mathbf{E}_\alpha^{\langle d\rangle}$ is forward-invariant under T_β. Therefore $\mathbf{L}_\alpha$ is T_β-forward-invariant.* ♢

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

29: Lemma. *Consider <u>distinct</u> scalars $\alpha \neq \beta$. For $d = 0, 1, 2, \ldots,$ the restricted operator*

$$\Big[\mathsf{T} - \beta\mathbf{I}\Big]\!\downharpoonright\!\mathbf{E}_\alpha^{\langle d\rangle}$$

has trivial kernel and so is a (linear) automorphism of $\mathbf{E}_\alpha^{\langle d\rangle}$ (since $\mathbf{E}_\alpha^{\langle d\rangle}$ is finite-dim'al). Taking a union, then, $\Big[\mathsf{T} - \beta\mathbf{I}\Big]\!\downharpoonright\!\mathbf{L}_\alpha$ is an automorphism of $\mathbf{L}_\alpha$. ♢

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 FTSOContradiction 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 $\mathsf{T} \leftrightarrows \mathsf{T}_\beta$. For $j = 0, 1, 2, \ldots,$ consequently, the vector $\mathsf{T}^j(\mathbf{v})$ is also in $\text{Ker}(\mathsf{T}_\beta)$. Consider <u>the</u> value of j for which $\mathsf{T}^j(\mathbf{v})$ is in $\mathbf{E}^{\langle 1\rangle} \smallsetminus \{\mathbf{0}\}$. Redefining $\mathbf{v}$ to be this $\mathsf{T}^j(\mathbf{v})$, we now have that

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

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

30: Prop'n. *Let $\lambda_1, \ldots, \lambda_L$ be the <u>distinct</u> eigenvalues of T. Then the collection $\mathbf{L}_{\lambda_1}, \ldots, \mathbf{L}_{\lambda_L}$ of nilspaces is linearly independent.* ♢

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

To this end, pick a number D large enough that

$$\big[\mathsf{T}_{\lambda_\ell}\big]^{\circ D}(\mathbf{v}_\ell) \;=\; \mathbf{0}, \quad \text{for each } \ell = 2,3,\ldots,L.$$

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

$$\Lambda \;:=\; \big[\mathsf{T}_{\lambda_2}\circ\mathsf{T}_{\lambda_3}\circ\cdots\circ\mathsf{T}_{\lambda_L}\big]^{\circ D}$$

sends each of $\mathbf{v}_2,\ldots,\mathbf{v}_L$ to $\mathbf{0}$. And Lemma 29 assures us that Λ is an automorphism of $\mathbf{L}_{\lambda_1}$. ♦

Proof of JCF, (24). Apply the Nilpotent JCF to $\mathsf{T}_{\lambda_\ell}$ on $\mathbf{L}_{\lambda_\ell}$ to get a basis $\mathcal{B}_\ell$ for $\mathbf{L}_{\lambda_\ell}$ against which T has a matrix-block of form λ_ℓ-$\mathsf{JB}(\overrightarrow{D}^\ell)$. 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. ♦

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

$$\mathit{JCF}(\mathsf{T}^{\mathsf{t}}) \;=\; \mathit{JCF}(\mathsf{T})$$

for an arbitrary square matrix T. □

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

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

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

$$\overrightarrow{D}_{\mathsf{B}}(\lambda) \;=\; \overrightarrow{D}_{\mathsf{B}}(\overline{\lambda}),$$

for each complex number λ. ◊

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

$$\lambda^+ := c + \boldsymbol{i}s \;\text{ and }\; \lambda^- := c - \boldsymbol{i}s.$$

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

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

for $\mathbb{C}^{\times 6}$, 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 $\lambda^\pm$-$\mathsf{JB}(3)$. Thus for $j \in [1..3]$,

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

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

Define new vectors

$$\mathbf{u}_j := 1\cdot\mathbf{e}_j^+ + \boldsymbol{i}\cdot\mathbf{e}_j^-;$$
$$\mathbf{w}_j := \boldsymbol{i}\cdot\mathbf{e}_j^+ + 1\cdot\mathbf{e}_j^-;$$

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

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

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

The B***-images of vectors.*** Verify that

$$c\mathbf{u}_j + s\mathbf{w}_j \;=\; \lambda^+\mathbf{e}_j^+ + \boldsymbol{i}\lambda^-\mathbf{e}_j^- \quad\text{and}$$
$$\text{-}s\mathbf{u}_j + c\mathbf{w}_j \;=\; \boldsymbol{i}\lambda^+\mathbf{e}_j^+ + \lambda^-\mathbf{e}_j^-.$$

From (33) we compute:

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

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

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

This, together with similar elbow grease, yields

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

Since all the coefficients are real, we get that B is conjugate to a real matrix. ♦

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

$$\mathrm{Spn}(\mathbf{v}, \mathsf{T}\mathbf{v}, \mathsf{T}^2\mathbf{v}, \mathsf{T}^3\mathbf{v}, \ldots)\,.$$

And 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 T yields a T-cyclic decomposition of the vectorspace.

Yo! Look in source file, here. □

Filename: `Problems/Algebra/LinearAlg/jordan_decomp.latex`
As of: *Thursday 13Apr2006.* Typeset: *20Mar2018 at 16:50.*