3 Jordan Canonical Forms

Characteristic Polynomials and Root Spaces

Let $\mathcal{V}$ be a finite dimensional $\mathbb{K}$-vector space. We do not assume that $\mathcal{V}$ is equipped with any structure in addition to the structure of a $\mathbb{K}$-vector space. In this section, we study geometry of linear operators on $\mathcal{V}$. In other words, we study the problem of classification of linear operators $A : \mathcal{V} \to \mathcal{V}$ up to **similarity** transformations $A \mapsto C^{-1}AC$, where C stands for arbitrary invertible linear transformations of $\mathcal{V}$.

Let $n = \dim \mathcal{V}$, and let A be the matrix of a linear operator with respect to some basis of $\mathcal{V}$. Recall that

$$\det(\lambda I - A) = \lambda^n + p_1\lambda^{n-1} + \cdots + p_{n-1}\lambda + p_n$$

is called the **characteristic polynomial** of A. In fact it does not depend on the choice of a basis. Indeed, under a change $\mathbf{x} = C\mathbf{x}'$ of coordinates, the matrix of a linear operator $\mathbf{x} \mapsto A\mathbf{x}$ is transformed into the matrix $C^{-1}AC$ similar to A. We have:

$$\det(\lambda I - C^{-1}AC) = \det[C^{-1}(\lambda I - A)C] = \\ (\det C^{-1}) \det(\lambda I - A)(\det C) = \det(\lambda I - A).$$

Therefore, the characteristic polynomial of a linear operator is well-defined (by the geometry of A). In particular, ***coefficients of the characteristic polynomial do not change under similarity transformations.***

Let $\lambda_0 \in \mathbb{K}$ be a root of the characteristic polynomial. Then $\det(\lambda_0 I - A) = 0$, and hence the system of homogeneous linear equations $A\mathbf{x} = \lambda_0\mathbf{x}$ has a non-trivial solution, $\mathbf{x} \neq \mathbf{0}$. As before,we call any such solution an **eigenvector** of A, and call λ_0 the corresponding **eigenvalue**. All solutions to $A\mathbf{x} = \lambda_0\mathbf{x}$ (including $\mathbf{x} = \mathbf{0}$) form a linear subspace in $\mathcal{V}$, called the **eigenspace** of A corresponding to the eigenvalue λ_0.

Let us change slightly our point of view on the eigenspace. It is the null space of the operator $A - \lambda_0 I$. Consider powers of this operator and their null spaces. If $(A - \lambda_0 I)^k\mathbf{x} = 0$ for some $k > 0$, then $(A - \lambda_0 I)^l\mathbf{x} = 0$ for all $l \geq k$. Thus the null spaces are nested:

$$\operatorname{Ker}(A - \lambda_0 I) \subset \operatorname{Ker}(A - \lambda_0 I)^2 \subset \cdots \subset \operatorname{Ker}(A - \lambda_0 I)^k \subset \ldots$$

On the other hand, since $\dim \mathcal{V} < \infty$, nested subspaces must stabilize, i.e. starting from some $m > 0$, we have:

$$\mathcal{W}_{\lambda_0} := \operatorname{Ker}(A - \lambda_0 I)^m = \operatorname{Ker}(A - \lambda_0 I)^{m+1} = \cdots$$

We call the subspace $\mathcal{W}_{\lambda_0}$ a **root space** of the operator A, namely, the root space corresponding to the root λ_0 of the characteristic polynomial.

Note that if $\mathbf{x} \in \mathcal{W}_{\lambda_0}$, then $A\mathbf{x} \in \mathcal{W}_{\lambda_0}$, because $(A - \lambda_0 I)^m A\mathbf{x} = A(A - \lambda_0 I)^m \mathbf{x} = A\mathbf{0} = \mathbf{0}$. Thus a root space is A-invariant. Denote by $\mathcal{U}_{\lambda_0}$ the range of $(A - \lambda_0 I)^m$. It is also A-invariant, since if $\mathbf{x} = (A - \lambda_0 I)^m \mathbf{y}$, then $A\mathbf{x} = A(A - \lambda_0 I)^m \mathbf{y} = (A - \lambda_0 I)^m (A\mathbf{y})$.

`Lemma.` $\mathcal{V} = \mathcal{W}_{\lambda_0} \oplus \mathcal{U}_{\lambda_0}$.

`Proof.` Put $B := (A - \lambda_0 I)^m$, so that $\mathcal{W}_{\lambda_0} = \operatorname{Ker} B$, $\mathcal{U}_{\lambda_0} = B(\mathcal{V})$. Let $\mathbf{x} = B\mathbf{y} \in \operatorname{Ker} B$. Then $B\mathbf{x} = \mathbf{0}$, i.e. $\mathbf{y} \in \operatorname{Ker} B^2$. But $\operatorname{Ker} B^2 = \operatorname{Ker} B$ by the assumption that $\operatorname{Ker} B = \mathcal{W}_{\lambda_0}$ is the root space. Thus $\mathbf{y} \in \operatorname{Ker} B$, and hence $\mathbf{x} = B\mathbf{y} = \mathbf{0}$. This proves that $\operatorname{Ker} B \cap B(\mathcal{V}) = \{\mathbf{0}\}$. Therefore the subspace in $\mathcal{V}$ spanned by $\operatorname{Ker} B$ and $\mathcal{B}(\mathcal{V})$ is their direct sum. On the other hand, for any operator, $\dim \operatorname{Ker} B + \dim B(\mathcal{V}) = \dim \mathcal{V}$. Thus, the subspace spanned by $\operatorname{Ker} B$ and $B(\mathcal{V})$ is the whole space $\mathcal{V}$.

`Corollary 1.` ***For any $\lambda \neq \lambda_0$, the root space $\mathcal{W}_\lambda \subset \mathcal{U}_{\lambda_0}$.***

`Proof.` Indeed, $\mathcal{W}_\lambda$ is invariant with respect to $A - \lambda_0 I$, but contains no eigenvectors of A with eigenvalue λ_0. Therefore $A - \lambda_0 I$ and all powers of it are invertible on $\mathcal{W}_\lambda$. Thus $\mathcal{W}_\lambda$ lies in the range $\mathcal{U}_{\lambda_0}$ of $B = (A - \lambda_0 I)^m$.

`Corollary 2.` ***Suppose that $(\lambda - \lambda_1)^{m_1} \dots (\lambda - \lambda_r)^{m_r}$ is the characteristic polynomial of $A : \mathcal{V} \to \mathcal{V}$, where $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ are pairwise distinct roots. Then $\mathcal{V}$ is the direct sum of root spaces:***

$$\mathcal{V} = \mathcal{W}_{\lambda_1} \oplus \cdots \oplus \mathcal{W}_{\lambda_r}.$$

`Proof.` From Corollary 1, it follows by induction on r, that $\mathcal{V} = \mathcal{W}_{\lambda_1} \oplus \cdots \oplus \mathcal{W}_{\lambda_r} \oplus \mathcal{U}$, where $\mathcal{U} = \mathcal{U}_{\lambda_1} \cap \cdots \cap \mathcal{U}_{\lambda_r}$. In particular, $\mathcal{U}$ is A-invariant as the intersection of A-invariant subspaces, but contains no eigenvectors of A with eigenvalues $\lambda_1, \dots, \lambda_r$. Picking bases in each of the direct summands $\mathcal{W}_{\lambda_i}$ and in $\mathcal{U}$, we obtain a basis in $\mathcal{V}$, in which the matrix of A is block-diagonal. Therefore the characteristic polynomial of A is the product of the characteristic polynomials of A restricted to the summands. So far we haven't used the hypothesis that the characteristic polynomial of A factors into a product of $\lambda - \lambda_i$. Invoking this hypothesis, we see that the factor of the characteristic polynomial, corresponding to $\mathcal{U}$ must have degree 0, and hence $\dim \mathcal{U} = 0$.

`Remarks.` (1) We will see later that dimensions of the root spaces coincide with multiplicities of the roots: $\dim \mathcal{W}_{\lambda_i} = m_i$.

(2) The restriction of A to $\mathcal{W}_{\lambda_i}$ has the property that some power of $A - \lambda_i I$ vanishes. A linear operator some power of which vanishes is called **nilpotent**. Our next task will be to study the geometry of nilpotent operators.

(3) Our assumption that the characteristic polynomial factors completely over $\mathbb{K}$ is automatically satisfied in the case $\mathbb{K} = \mathbb{C}$ due to the Fundamental Theorem of Algebra. Thus, we have proved for every linear operator on a finite dimensional complex vector space, that the space decomposes in a canonical fashion into the direct sum of invariant subspaces on each of which the operator differs from a nilpotent one by scalar summand.

EXERCISES

418. Let $A, B : \mathcal{V} \to \mathcal{V}$ be two commuting linear operators, and p and q two polynomials in one variable. Show that the operators $p(A)$ and $q(B)$ commute.

419. Prove that if A commutes with B, then root spaces of A are B-invariant.

420. Let λ_0 be a root of the characteristic polynomial of an operator A, and m its multiplicity. What are possible values for the dimension of the eigenspace corresponding to λ_0?

421. Let $\mathbf{v} \in \mathcal{V}$ be a non-zero vector, and $\mathbf{a} : \mathcal{V} \to \mathbb{K}$ a non-zero linear function. Find eigenvalues and eigenspaces of the operator $\mathbf{x} \mapsto \mathbf{a}(\mathbf{x})\mathbf{v}$.

Nilpotent Operators

`Example.` Introduce a nilpotent linear operator $N : \mathbb{K}^n \to \mathbb{K}^n$ by describing its action on vectors of the standard basis:

$$N\mathbf{e}_n = \mathbf{e}_{n-1}, \quad N\mathbf{e}_{n-1} = \mathbf{e}_{n-2}, \quad \dots, \quad N\mathbf{e}_2 = \mathbf{e}_1, \quad N\mathbf{e}_1 = \mathbf{0}.$$

Then $N^n = 0$ but $N^{n-1} \neq 0$. We will call N, as well as any operator similar to it, a **regular nilpotent** operator. The matrix of N in the standard basis has the form

$$\begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}.$$

It has the range of dimension $n-1$ spanned by $\mathbf{e}_1, \dots, \mathbf{e}_{n-1}$ and the null space of dimension 1 spanned by $\mathbf{e}_1$.

Proposition. ***Let $N : \mathcal{V} \to \mathcal{V}$ be a nilpotent operator on a $\mathbb{K}$-vector space of finite dimension. Then the space can be decomposed into the direct sum of N-invariant subspaces, on each of which N is regular.***

Proof. We use induction on $\dim \mathcal{V}$. When $\dim \mathcal{V} = 0$, there is nothing to prove. Now consider the case when $\dim \mathcal{V} > 0$.

The range $N(\mathcal{V})$ is N-invariant, and $\dim N(\mathcal{V}) < \dim V$ (since otherwise N could not be nilpotent). By the induction hypothesis, the space $N(\mathcal{V})$ can be decomposed into the direct sum of N-invariant subspaces, on each of which N is regular. Let l be the number of these subspaces, $n_1, \dots, n_l$ their dimensions, and $\mathbf{e}_1^{(i)}, \dots, \mathbf{e}_{n_i}^{(i)}$ a basis in the ith subspace such that N acts on the basis vectors as in Example:

$$\mathbf{e}_{n_i}^{(i)} \mapsto \cdots \mapsto \mathbf{e}_1^{(i)} \mapsto \mathbf{0}.$$

Since each $\mathbf{e}_{n_i}^{(i)}$ lies in the range of N, we can pick a vector $\mathbf{e}_{n_i+1}^{(i)} \in \mathcal{V}$ such that $N\mathbf{e}_{n_i+1}^{(i)} = \mathbf{e}_{n_i}^{(i)}$. Note that $\mathbf{e}_1^{(1)}, \dots, \mathbf{e}_1^{(l)}$ form a basis in $(\operatorname{Ker} N) \cap N(\mathcal{V})$. We complete it to a basis

$$\mathbf{e}_1^{(1)}, \dots, \mathbf{e}_1^{(l)}, \mathbf{e}_1^{(l+1)}, \dots, \mathbf{e}_1^{(r)}$$

of the whole null space $\operatorname{Ker} N$. We claim that *all the vectors $\mathbf{e}_j^{(i)}$ form a basis in $\mathcal{V}$,* and therefore the $l + r$ subspaces

$$\operatorname{Span}(\mathbf{e}_1^{(i)}, \dots, \mathbf{e}_{n_i}^{(i)}, \mathbf{e}_{n_i+1}^{(i)}), \quad i = 1, \dots, l, l+1, \dots, r,$$

(of which the last $r - l$ are 1-dimensional) form a decomposition of $\mathcal{V}$ into the direct sum with required properties.

To justify the claim, notice that the subspace $\mathcal{U} \subset \mathcal{V}$ spanned by $n_1 + \cdots + n_l = \dim N(\mathcal{V})$ vectors $\mathbf{e}_j^{(i)}$ with $j > 1$ is mapped by N onto the space $N(\mathcal{V})$. Therefore: (a) those vectors form a basis of $\mathcal{U}$, (b) $\dim \mathcal{U} = \dim N(\mathcal{V})$, and (c) $\mathcal{U} \cap \operatorname{Ker} N = \{\mathbf{0}\}$. On the other hand, vectors $\mathbf{e}_j^{(i)}$ with $j = 1$ form a basis of $\operatorname{Ker} N$, and since $\dim \operatorname{Ker} N + \dim N(\mathcal{V}) = \dim \mathcal{V}$, together with the above basis of $\mathcal{U}$, they form a basis of $\mathcal{V}$. □

Corollary 1. ***The matrix of a nilpotent operator in a suitable basis is block diagonal with regular diagonal blocks*** (as in Example) ***of certain sizes $n_1 \geq \cdots \geq n_r > 0$.***

The basis in which the matrix has this form, as well as the decomposition into the direct sum of invariant subspaces as described in Proposition, are not canonical, since choices are involved on each step of induction. However, the dimensions $n_1 \geq \cdots \geq n_r > 0$ of the subspaces turn out to be uniquely determined by the geometry of the operator.

To see why, introduce the following **Young tableaux** (Figure 46). It consists of r rows of identical square cells. The lengths of the rows represent dimensions $n_1 \geq n_2 \geq \cdots \geq n_r > 0$ of invariant subspaces, and in the cells of each row we place the basis vectors of the corresponding subspace, so that the operator N sends each vector to its left neighbor (and those of the leftmost column to $\mathbf{0}$).

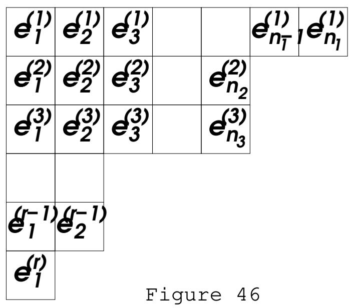

Figure 46

The format of the tableaux is determined by the **partition** of the total number n of cells (equal to $\dim \mathcal{V}$) into the sum $n_1 + \cdots + n_r$ of positive integers. Reading the same format *by columns*, we obtain another partition $n = m_1 + \cdots + m_d$, called **transposed** to the first one, where $m_1 \geq \cdots \geq m_d > 0$ are the heights of the columns. Obviously, two transposed partitions determine each other.

It follows from the way how the cells are filled with vectors $\mathbf{e}_j^{(i)}$, that the vectors in the columns 1 through k form a basis of the space $\operatorname{Ker} N^k$. Therefore

$$m_k = \dim \operatorname{Ker} N^k - \dim \operatorname{Ker} N^{k-1}, \quad k = 1, \ldots, d.$$

Corollary 2. ***Consider the flag of subspaces defined by a nilpotent operator*** $N : \mathcal{V} \to \mathcal{V}$**:**

$$\operatorname{Ker} N \subset \operatorname{Ker} N^2 \subset \cdots \subset \operatorname{Ker} N^d = \mathcal{V}$$

and the partition of $n = \dim V$ ***into the summands*** $m_k = \dim \operatorname{Ker} N^k - \dim \operatorname{Ker} N^{k-1}$***. The summands of the transposed***

partition $n = n_1 + \cdots + n_r$ ***are the dimensions of the regular nilpotent blocks of*** N (described in Proposition and Corollary 1).

`Corollary 3.` ***The number of equivalence classes of nilpotent operators on a vector space of dimension*** n ***is equal to the number of partitions of*** n***.***

EXERCISES

422. Let $\mathcal{V}_n \subset \mathbb{K}[x]$ be the space of all polynomials of degree $< n$. Prove that the differentiation $\frac{d}{dx} : \mathcal{V}_n \to \mathcal{V}_n$ is a regular nilpotent operator.

423. Find all matrices commuting with a regular nilpotent one.

424. Is there an $n \times n$-matrix A such that $A^2 \neq 0$ but $A^3 = 0$: (a) if $n = 2$? (b) if $n = 3$?

425. Classify similarity classes of nilpotent 4×4-matrices.

426. An operator is called **unipotent**, if it differs from the identity by a nilpotent operator. Prove that the number of similarity classes of *unipotent* $n \times n$-matrices is equal to the number of partitions of n, and find this number for $n = 5$.

The Jordan Canonical Form Theorem

We proceed to classification of linear operators $A : \mathbb{C}^n \to \mathbb{C}^n$ up to similarity transformations.

`Theorem.` ***Every complex matrix is similar to a block-diagonal normal form with each diagonal block of the form:***

$$\begin{bmatrix} \lambda_0 & 1 & 0 & \dots & 0 \\ 0 & \lambda_0 & 1 & \dots & 0 \\ & & \dots & & \\ 0 & 0 & \dots & \lambda_0 & 1 \\ 0 & 0 & \dots & 0 & \lambda_0 \end{bmatrix}, \quad \lambda_0 \in \mathbb{C},$$

and such a normal form is unique up to permutations of the blocks.

The block-diagonal matrices described in the theorem are called **Jordan canonical forms** (or **Jordan normal forms**). Their diagonal blocks are called **Jordan cells**.

It is instructive to analyze a Jordan canonical form before going into the proof of the theorem. The characteristic polynomial of a Jordan cell is $(\lambda - \lambda_0)^m$ where m is the size of the cell. The characteristic polynomial of a block-diagonal matrix is equal to the product of characteristic polynomials of the diagonal blocks. Therefore the characteristic polynomial of the whole Jordan canonical form is the product of factors $(\lambda - \lambda_i)^{m_i}$, one per Jordan cell. Thus the diagonal entries of Jordan cells are roots of the characteristic polynomial. After subtracting the scalar matrix $\lambda_0 I$, Jordan cells with $\lambda_i = \lambda_0$ (and only these cells) become nilpotent. Therefore the root space $\mathcal{W}_{\lambda_0}$ is exactly the direct sum of those subspaces on which the Jordan cells with $\lambda_i = \lambda_0$ operate.

`Proof of Theorem.` Everything we need has been already established in the previous two subsections.

Thanks to the Fundamental Theorem of Algebra, the characteristic polynomial $\det(\lambda I - A)$ of a complex $n \times n$-matrix A factors into the product of powers of distinct linear factors: $(\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_r)^{n_r}$. According to Corollary 2 of Lemma, the space $\mathbb{C}^n$ is decomposed in a canonical fashion into the direct sum $\mathcal{W}_{\lambda_1} \oplus \cdots \oplus \mathcal{W}_{\lambda_r}$ of A-invariant root subspaces. On each root subspace $\mathcal{W}_{\lambda_i}$, the operator $A - \lambda_i I$ is nilpotent. According to Proposition, the root space $\mathcal{W}_{n_i}$ is represented (in a non-canonical fashion) as the direct sum of invariant subspaces on each of which $A - \lambda_i I$ acts as a regular nilpotent operator. Since the scalar operator $\lambda_i I$ leaves every subspace invariant, this means that $\mathcal{W}_{\lambda_i}$ is decomposed into the direct sum of A-invariant subspaces, on each of which A acts as a Jordan cell with the eigenvalue $\lambda_0 = \lambda_i$. Thus, existence of a basis in which A is described by a Jordan normal form is established.

To prove uniqueness, note that the root spaces $\mathcal{W}_{\lambda_i}$ are intrinsically determined by the operator A, and the partition of $\dim \mathcal{W}_{\lambda_i}$ into the sizes of Jordan cells with the eigenvalue λ_i is uniquely determined, according to Corollary 2 of Proposition, by the geometry of the operator $A - \lambda_i I$ nilpotent on $\mathcal{W}_{\lambda_i}$. Therefore the exact structure of the Jordan normal form of A (i.e. the numbers and sizes of Jordan cells for each of the eigenvalues λ_i) is uniquely determined by A, and only the ordering of the diagonal blocks remains ambiguous. □

`Corollary 1.` ***Dimensions of root spaces*** $\mathcal{W}_{\lambda_i}$ ***coincide with multiplicities of*** λ_i ***as roots of the characteristic polynomial.***

`Corollary 2.` ***If the characteristic polynomial of a complex matrix has only simple roots, then the matrix is*** **diagonalizable**, ***i.e. is similar to a diagonal matrix.***

`Corollary 3.` ***Every operator*** $A : \mathbb{C}^n \to \mathbb{C}^n$ ***in a suitable basis is described by the sum*** $D+N$ ***of two commuting matrices, of which*** D ***is diagonal, and*** N ***strictly upper triangular.***

`Corollary 4.` ***Every operator on a complex vector space of finite dimension can be represented as the sum $D+N$ of two commuting operators, of which D is diagonalizable and N nilpotent.***

`Remark.` We used that $\mathbb{K}=\mathbb{C}$ only to factor the characteristic polynomial of the matrix A into linear factors. Therefore the same results hold true over any field $\mathbb{K}$ such that all non-constant polynomials from $\mathbb{K}[\lambda]$ factor into linear factors. Such fields are called **algebraically closed**. In fact (see [8]) every field $\mathbb{K}$ is contained in an algebraically closed field. Thus every linear operator $A:\mathbb{K}^n\to\mathbb{K}^n$ can be brought to a Jordan normal form by transformations $A\mapsto C^{-1}AC$, where however entries of C and scalars λ_0 in Jordan cells may belong to a larger field $\mathbb{F}\supset\mathbb{K}$.[8] We will see how this works when $\mathbb{K}=\mathbb{R}$ and $\mathbb{F}=\mathbb{C}$.

EXERCISES

427. Find Jordan normal forms of the following matrices: ✓

$$(a)\begin{bmatrix}1&2&0\\0&2&0\\-2&-2&1\end{bmatrix},\ (b)\begin{bmatrix}4&6&0\\-3&-5&0\\-3&-6&1\end{bmatrix},\ (c)\begin{bmatrix}13&16&16\\-5&-7&-6\\-6&-8&-7\end{bmatrix},$$

$$(d)\begin{bmatrix}3&0&8\\3&-1&-6\\-2&0&-5\end{bmatrix},\ (e)\begin{bmatrix}-4&2&10\\-4&3&7\\-3&1&7\end{bmatrix},\ (f)\begin{bmatrix}7&-12&-2\\3&-4&0\\-2&0&2\end{bmatrix},$$

$$(g)\begin{bmatrix}-2&8&6\\-4&10&6\\4&-8&-4\end{bmatrix},\ (h)\begin{bmatrix}0&3&3\\-1&8&6\\2&-14&-10\end{bmatrix},\ (i)\begin{bmatrix}1&1&-1\\-3&-3&3\\-2&-2&2\end{bmatrix},$$

$$(j)\begin{bmatrix}1&-1&2\\3&-3&6\\2&-2&4\end{bmatrix},\ (k)\begin{bmatrix}-1&1&1\\-5&21&17\\6&-26&-21\end{bmatrix},\ (l)\begin{bmatrix}3&7&-3\\-2&-5&2\\-4&-10&3\end{bmatrix},$$

$$(m)\begin{bmatrix}8&30&-14\\-6&-19&9\\-6&-23&11\end{bmatrix},\ (n)\begin{bmatrix}9&22&-6\\-1&-4&1\\8&16&-5\end{bmatrix},\ (o)\begin{bmatrix}4&5&-2\\-2&-2&1\\-1&-1&1\end{bmatrix}.$$

428. Compute powers of Jordan cells. ϟ

429. Prove that if some power of a complex matrix is the identity, then the matrix is diagonalizable.

430. Prove that transposed square matrices are similar.

431. Prove that $\operatorname{tr} A=\sum\lambda_i$ and $\det A=\prod\lambda_i$, where $\lambda_1,\dots,\lambda_n$ are all roots of the characteristic polynomial (repeated according to their multiplicities).

---

[8]For this, $\mathbb{F}$ does not have to be algebraically closed, but only needs to contain all roots of $\det(\lambda I-A)$.

432. Prove that a square matrix satisfies its own characteristic equation; namely, if p denotes the characteristic polynomial of a matrix A, then $p(A) = 0$. (This identity is called the **Cayley–Hamilton equation**.)

The Real Case

Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be an $\mathbb{R}$-linear operator. It acts[9] on the complexification $\mathbb{C}^n$ of the real space, and commutes with the complex conjugation operator $\sigma : \mathbf{x} + i\mathbf{y} \mapsto \mathbf{x} - i\mathbf{y}$.

The characteristic polynomial $\det(\lambda I - A)$ has real coefficients, but its roots λ_i can be either real or come in pairs of complex conjugated roots (of the same multiplicity). Consequently, the complex root spaces $\mathcal{W}_{\lambda_i}$, which are defined as null spaces in $\mathbb{C}^n$ of sufficiently high powers of $A - \lambda_i I$, come in two types. If λ_i is real, then the root space is *real* in the sense that it is σ-invariant, and thus is the complexification of the real root space $\mathcal{W}_{\lambda_i} \cap \mathbb{R}^n$. If λ_i is not real, and $\bar{\lambda}_i$ is its complex conjugate, then $\mathcal{W}_{\lambda_i}$ and $\mathcal{W}_{\bar{\lambda}_i}$ are different root spaces of A, but they are transformed into each other by σ. Indeed, $\sigma A = A\sigma$, and $\sigma\lambda_i = \bar{\lambda}_i\sigma$. Therefore, if $\mathbf{z} \in \mathcal{W}_{\lambda_i}$, i.e. $(A - \lambda_i I)^d \mathbf{z} = \mathbf{0}$ for some d, then $\mathbf{0} = \sigma(A - \lambda_i I)^d \mathbf{z} = (A - \bar{\lambda}_i I)^d \sigma\mathbf{z}$, and hence $\sigma\mathbf{z} \in \mathcal{W}_{\bar{\lambda}_i}$. This allows one to obtain the following improvement for the Jordan Canonical Form Theorem applied to real matrices.

`Theorem.` ***A real linear operator*** $A : \mathbb{R}^n \to \mathbb{R}^n$ ***can be represented by the matrix in a Jordan normal form with respect to a basis of the complexified space*** $\mathbb{C}^n$ ***invariant under complex conjugation.***

`Proof.` In the process of construction bases in $\mathcal{W}_{\lambda_i}$ in which A has a Jordan normal form, we can use the following procedure. When λ_i is real, we take the real root space $\mathcal{W}_{\lambda_i} \cap \mathbb{R}^n$ and take in it a real basis in which the matrix of $A - \lambda_i I$ is block-diagonal with regular nilpotent blocks. This is possible due to Proposition applied to the case $\mathbb{K} = \mathbb{R}$. This real basis serves then as a σ-invariant complex basis in the complex root space $\mathcal{W}_{\lambda_i}$. When $\mathcal{W}_{\lambda_i}$ and $\mathcal{W}_{\bar{\lambda}_i}$ is a pair of complex conjugated root spaces, then we take a required basis in one of them, and then apply σ to obtain such a basis in the other. Taken together, the bases form a σ-invariant set of vectors. $\square$

Of course, for each Jordan cell with a non-real eigenvalue λ_0, there is another Jordan cell of the same size with the eigenvalue $\bar{\lambda}_0$.

[9]Strictly speaking, it is the complexification $A^{\mathbb{C}}$ of A that acts on $\mathbb{C}^n = (\mathbb{R}^n)^{\mathbb{C}}$, but we will denote it by the same letter A.

Moreover, if $\mathbf{e}_1, \dots, \mathbf{e}_m$ is the basis in the A-invariant subspace of the first cell, i.e. $A\mathbf{e}_k = \lambda_0 \mathbf{e}_k + \mathbf{e}_{k-1}$, $k = 2, \dots, m$, and $A\mathbf{e}_1 = \lambda_0 \mathbf{e}_1$, then the A-invariant subspace corresponding to the other cell comes with the complex conjugate basis $\bar{\mathbf{e}}_1, \dots, \bar{\mathbf{e}}_m$, where $\bar{\mathbf{e}}_k = \sigma \mathbf{e}_k$. The direct sum $\mathcal{U} := \mathrm{Span}(\mathbf{e}_1, \dots, \mathbf{e}_m, \bar{\mathbf{e}}_1, \dots, \bar{\mathbf{e}}_m)$ of the two subspaces is both A- and σ-invariant and thus is a complexification of the real A-invariant subspace $\mathcal{U} \cap \mathbb{R}^n$. We use this to describe a real normal form for the action of A on this subspace.

Namely, let $\lambda_0 = \alpha - i\beta$, and write each basis vector $\mathbf{e}_k$ in terms of its real and imaginary part: $\mathbf{e}_k = \mathbf{u}_k - i\mathbf{v}_k$. Then the real vectors $\mathbf{u}_k$ and $\mathbf{v}_k$ form a real basis in the real part of the complex 2-dimensional space spanned by $\mathbf{e}_k$ and $\bar{e}_k = \mathbf{u}_k + i\mathbf{v}_k$. Thus, we obtain a basis $\mathbf{u}_1, \mathbf{v}_1, \mathbf{u}_2, \mathbf{v}_2, \dots, \mathbf{u}_m, \mathbf{v}_m$ in the subspace $\mathcal{U} \cap \mathbb{R}^n$. The action of A on this basis is found from the formulas:

$$A\mathbf{u}_1 + iA\mathbf{v}_1 = (\alpha - i\beta)(\mathbf{u}_1 + i\mathbf{v}_1) = (\alpha\mathbf{u}_1 + \beta\mathbf{v}_1) + i(-\beta\mathbf{u}_1 + \alpha\mathbf{v}_1),$$
$$A\mathbf{u}_k + iA\mathbf{v}_k = (\alpha\mathbf{u}_k + \beta\mathbf{v}_k + \mathbf{u}_{k-1}) + i(-\beta\mathbf{u}_k + \alpha\mathbf{v}_k + \mathbf{v}_{k-1}),\ k > 1.$$

`Corollary 1.` ***A linear operator*** $A : \mathbb{R}^n \to \mathbb{R}^n$ ***is represented in a suitable basis by a block-diagonal matrix with the diagonal blocks that are either Jordan cells with real eigenvalues, or have the form*** ($\beta \neq 0$)***:***

$$\begin{bmatrix} \alpha & -\beta & 1 & 0 & 0 & & \dots & & 0 & 0 \\ \beta & \alpha & 0 & 1 & 0 & & \dots & & 0 & 0 \\ 0 & 0 & \alpha & -\beta & 1 & 0 & & \dots & & 0 \\ 0 & 0 & \beta & \alpha & 0 & 1 & & \dots & & 0 \\ & & & & & \dots & & & & \\ 0 & & & \dots & & 0 & \alpha & -\beta & 1 & 0 \\ 0 & & & \dots & & 0 & \beta & \alpha & 0 & 1 \\ 0 & 0 & & \dots & & & 0 & 0 & \alpha & -\beta \\ 0 & 0 & & \dots & & & 0 & 0 & \beta & \alpha \end{bmatrix}.$$

`Corollary 2.` ***If two real matrices are related by a complex similarity transformation, then they are related by a real similarity transformation.***

`Remark.` The proof meant here is that if two real matrices are similar over $\mathbb{C}$ then they have the same Jordan normal form, and thus they are similar over $\mathbb{R}$ to the same real matrix, as described in Corollary 1. However, Corollary 2 can be proved directly, without a reference to the Jordan Canonical Form Theorem. Namely, if two real matrices, A and A', are related by a complex similarity transformation: $A' = C^{-1}AC$, we can rewrite this as $CA' = AC$, and taking $C = B + iD$ where B and D are real, obtain: $BA' = AB$ and $DA' = AD$. The problem now is that neither B nor D is guaranteed to be invertible. Yet, there must exist an invertible linear combination $E = \lambda B + \mu D$, for if the polynomial $\det(\lambda B + \mu D)$ of λ and μ vanishes identically, then $\det(B + iD) = 0$ too. For invertible E, we have $EA' = AE$ and hence $A' = E^{-1}AE$.

EXERCISES

433. Classify all linear operators in $\mathbb{R}^2$ up to linear changes of coordinates.

434. Consider real traceless 2×2-matrices $\begin{bmatrix} a & b \\ c & -a \end{bmatrix}$ as points in the 3-dimensional space with coordinates a, b, c. Sketch the partition of this space into similarity classes.

4 Linear Dynamical Systems

We consider here applications of Jordan Canonical Forms to dynamical systems, i.e. models of evolutionary processes. An evolutionary system can be described mathematically as a set, X, of *states* of the system, and a collection of maps $g_{t_1}^{t_2} : X \to X$ which describe the evolution of the states from the time moment t_1 to the time moment t_2, whereas the composition of $g_{t_2}^{t_3}$ with $g_{t_1}^{t_2}$ is required to coincide with $g_{t_1}^{t_3}$.

One says that the dynamical sustem is **time-independent** (or *stationary*), if the maps g_{t_1,t_2} depend only on the time increment $t = t_2 - t_1$, i.e. $g_{t_1}^{t_2} = g_0^{t_2-t_1}$. Omitting the subscript 0, we then have a family of maps $g^t : X \to X$ satisfying $g^t g^u = g^{t+u}$.

We will consider time-independent **linear dynamical systems**, i.e. such systems where the states form a vector space, $\mathcal{V}$, and the evolution maps are linear, and examine both cases: of **discrete time** or **continuous time**. In the latter case, time takes on integer values $t \in \mathbb{Z}$, and the evolution is described by **iterations** of an invertible linear map $G : \mathcal{V} \to \mathcal{V}$. Namely, if the state of the system at $t = 0$ is $\mathbf{x}(0) \in \mathcal{V}$, then the state of the system at the moment $n \in \mathbb{Z}$ is $\mathbf{x}(n) = G^n x(0)$. In the case of continuous time, a time-independent linear dynamical systems are described by a system of constant coefficient **linear** ordinary differential equations (ODE for short): $\dot{\mathbf{x}} = A\mathbf{x}$, and the evolution maps are found by solving the system.

We begin with the case of discrete time.

Iterations of Linear Maps

`Example 1:` *Fibonacci numbers.*. The sequence of integers

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \dots,$$

formed by adding the previous two terms to obtaine the next one, is well-known under the name of **Fibonacci sequence.**[10] It is the solution of the 2nd order **linear recursion relation** $f_{n+1} = f_n + f_{n-1}$, satisfying the intial conditions $f_0 = 0, f_1 = 1$. The sequence can be recast as a trajectory of a linear dynamical system on the plane as follows. Append the pair f_n, f_{n+1} of consecutive terms of the sequence into a 2-column $(x_n, y_n)^t$ and express the next pair as a

[10] After Leonardo **Fibonacci** (c. 1170 – c. 1250).

function of the previous one:

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix}, \quad \text{where} \quad \begin{bmatrix} x_n \\ y_n \end{bmatrix} := \begin{bmatrix} f_n \\ f_{n+1} \end{bmatrix}.$$

We will now derive the general formula for the numbers f_n (and for all solutions to the recursion relation).

The characteristic polynomial of the matrix $G := \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ is $\lambda^2 - \lambda - 1$. It has two roots $\lambda_\pm = (1 \pm \sqrt{5})/2$. The corresponding eigenvectors can be taken in the form $\mathbf{v}_\pm = \begin{bmatrix} 1 \\ \lambda_\pm \end{bmatrix}$. Let $\mathbf{v}(0)$ be the vector of initial conditions (it is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ for the Fibonacci sequence), and let it be written as a linear combination of the eigenvectors: $\mathbf{v}(0) = C_+\mathbf{v}_+ + C_-\mathbf{v}_-$. Then $\mathbf{v}(n) := G^n\mathbf{v}(0) = C_+\lambda_+^n\mathbf{v}_+ + C_-\lambda_-^n\mathbf{v}_-$. Taking the first component of this vector equality, we obtain the general formula, depending on two arbitrary constants, $C_\pm$, for solutions of the recursion relation $x_{n+1} = x_n + x_{n-1}$:

$$x_n = C_+\left(\frac{1+\sqrt{5}}{2}\right)^n + C_-\left(\frac{1-\sqrt{5}}{2}\right)^n.$$

For the Fibonacci sequence, from equations $\lambda_+C_+ + \lambda_-C_- = 1$, $C_+ + C_- = 0$, we find $C_\pm = \pm 1/\sqrt{5}$, and hence

$$f_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right].$$

Note that $|\lambda_-| < 1$, and $|\lambda_+| > 1$. Consequently, as n increases indefinitely, the second summand tends to 0, and the first to infinity. Thus asymptotically (for large n) $f_n \approx [(1+\sqrt{5})/2]^n/\sqrt{5}$, while consecutive ratios f_{n+1}/f_n tend to $\lambda_+ = (1+\sqrt{5})/2$, known as **golden ratio**. In fact the fractions $3/2, 5/3, 8/5, 13/8, 21/13, 34/21, \dots$, i.e. f_{n+1}/f_n, are the best possible rational approximations to the golden ratio with the denominators not eexceeding f_n.[11]

[11] The golden ratio often occurs in Nature. For example, on a pine cone, or a pineapple fruit, count the numbers of spirals (formed by the scales near the stem) going in clockwise and counter-clockwise directions. Most likely you'll find two consquitive Fibonacci numbers: 5 and 8, or 8 and 13. It is not surprising that such observations led to the mystical believes of our ancestors into magical properties of the golden ratio. In fact we shouldn't judge them too harshely: Inspite of some plausible scientific models, the causes for occurences of Fibonacci numbers and the golden ratio in various plants are still not entirely clear.

In general, an (invertible) linear map $G : \mathbb{R}^m \to \mathbb{R}^m$ defines a linear dynamical system: Given the initial state $\mathbf{v}(0)$, the state at the discrete time moment $n \in \mathbb{Z}$ is defined by $\mathbf{v}(n) = G^n\mathbf{v}(0)$. In order to efficiently compute the state $\mathbf{v}(n)$, we need therefore to compute powers of a linear map.

According to the general theory, there exists an invertible complex matrix C such that $G = CJC^{-1}$, where J is one of the Jordam Canonical Forms. Therefore

$$G^n = (CJC^{-1})(CJC^{-1})\dots(CJC^{-1}) = CJ^nC^{-1},$$

and the problem reduces to that for J. Recall that $J = D + N$, the sum of a diagonal matrix D (whose diagonal entries are eigenvalues of G), and of a nilpotent Jordan matrix N (in particular, $N^m = 0$), commuting with D. Thus, by the binomial formula, we have a finite sum:

$$J^n = (D+N)^n = D^n + nD^{n-1}N + \binom{n}{2}D^{n-1}N^2 + \binom{n}{3}D^{n-3}N^3 + \cdots.$$

`Example 2:` *Powers of Jordan cells.* In fact, to find explicitly the powers of J^n it suffcices to do this for each Jordan cell ceparately:

$$\begin{bmatrix} \lambda & 1 & 0 & 0 & \dots \\ 0 & \lambda & 1 & 0 & \dots \\ 0 & 0 & \lambda & 1 & \dots \\ 0 & 0 & 0 & \lambda & \dots \\ & \dots & & & \end{bmatrix}^n = \begin{bmatrix} \lambda^n & n\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} & \binom{n}{3}\lambda^{n-3} & \dots \\ 0 & \lambda^n & n\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} & \dots \\ 0 & 0 & \lambda^n & n\lambda^{n-1} & \dots \\ 0 & 0 & 0 & \lambda^n & \dots \\ & \dots & & & \end{bmatrix}$$

`Example 3.` Let us find the general formula for solutions of the recursive sequence $a_{n+1} = 4a_n - 6a_{n-1} + 4a_{n-2} - a_{n-3}$. Put $\mathbf{v}(n) = (a_{n-3}, a_{n-2}, a_{n-1}, a_n)^t$, Then $\mathbf{v}(n+1) = G\mathbf{v}(n)$, where

$$G = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 4 & -6 & 4 \end{bmatrix}.$$

The reader is asked to check that

$$\det(\lambda I - G) = \lambda^4 - 4\lambda^3 + 6\lambda^2 - 4\lambda + 1 = (\lambda - 1)^4.$$

Therefore G has only one eigenvalue: $\lambda = 1$, which is a root of multiplicity 4. Besides, the rank of $I - G$ is at least 3 due to the

3×3-identity matrix which occurs in the right upper corner of G. This shows that the eigenspace has dimension 1, and thus the Jordan canonical form of G has only one Jordan cell, J, of size 4, with the eigenvalue 1. Examining the matrix entries of J^n from Example 2 (with $\lambda = 1$, $m = 4$), we see that they all have the form $\binom{n}{k}$, where $k = 0, 1, 2, 3$. Note that as functions of n these binomial coefficients form a basis in the space of polynomilas of the form $C_0+C_1n+C_2n^2+C_3n^3$. We conclude that matrix entries of $G^n = CJ^nC^{-1}$ must be, as functions of n, polynomilas of degree ≤ 3. In particular, a_n must be a polynomial of n of degree ≤ 3. Since solution sequences $\{a_n\}$ must depend on 4 arbitrary parameters a_0, a_1, a_2, a_3, and thus form a 4-dimensional space (the same as the space of polynomials of degree ≤ 3), we finally conlude that *the general solution to the recursion equation has the form* $a_n = C_0 + C_1n + C_2n^2 + C_3n^3$, *where* C_i *are arbitrary constants.* Given specific values of a_0, a_1, a_2, a_3, one can find the corresponding values of C_0, C_1, C_2, C_3 by plugging $n = 0, 1, 2, 3$, and solving the system of 4 linear equations.

As it follows from Example 2 (and is illustrated by Examples 1 and 3), ***for any discrete dymanical system*** $\mathbf{v}(n+1) = G\mathbf{v}(n)$, ***components of the vector trajectories*** $\mathbf{v}(n) = G^n\mathbf{v}(0)$ ***as functions of*** n ***are linear combinations of the monomials*** $\lambda_i^n n^{k-1}$, ***where*** λ_i ***is a roots of the characteristic polynomial of the operator*** G, ***and*** k ***does not exceed the multiplicity of this root.*** In particular, if all roots are simple, the egneral solution has the form $\mathbf{v}(n) = \sum_{i=1}^m C_i\lambda_i^n\mathbf{v}_i$, where $\mathbf{v}_i$ is an eigenvector corresponding to the eigenvalue λ_i, and C_i are arbitrary constants.

EXERCISES

435. Find the general formula for the sequence a_n such that $a_0 = 2$, $a_1 = 1$, and $a_{n+1} = a_n + a_{n-1}$ for all $n \geq 1$.

436. Find all solutions to the recursion equation $a_{n+1} = 5a_n + 6a_{n-1}$.

437. For the recursion equation $a_{n+1} = 6a_n - 9a_{n-1}$, express the general solutions as a function of n, a_0 and a_1. ✓

438. Given two sequences $\{x_n\}, \{y_n\}$ satisfying $x_{n+1} = 4x_n + 3y_n$, $y_{n+1} = 3x_n + 2y_n$, and such that $x_0 = y_0 = 13$. Find the limit of the ratio x_n/y_n as n tends to: (a) $+\infty$; (b) $-\infty$. ✓

439. Prove that for a given $G : \mathcal{V} \to \mathcal{V}$, all trajectories $\{\mathbf{x}(n)\}$ of a given dynamical system $\mathbf{x}(n+1) = G\mathbf{x}(n)$ form a linear subspace of dimension $\dim \mathcal{V}$ in the space of all vector-valued sequences.

440.* For integer $n > 0$, prove that $\left(\frac{3+\sqrt{17}}{2}\right)^n + \left(\frac{3-\sqrt{17}}{2}\right)^n$ is an odd integer.

Linear ODE Systems

Let

$$\begin{aligned} \dot{x}_1 &= a_{11}x_1 + ... + a_{1n}x_n \\ &... \\ \dot{x}_2 &= a_{n1}x_1 + ... + a_{nn}x_n \end{aligned}$$

be a linear homogeneous system of ordinary differential equations with constant (possibly complex) coefficients a_{ij}. It can be written in the matrix form as $\dot{\mathbf{x}} = A\mathbf{x}$.

Consider the infinite matrix series

$$e^{tA} := I + tA + \frac{t^2A^2}{2} + \frac{t^3A^3}{6} + ... + \frac{t^kA^k}{k!} + ...$$

If M is an upper bound for the absolute values of the entries of A, then the matrix entries of t^kA^k are bounded by $n^kt^kM^k$. It is easy to deduce from this that the series converges (at least as fast as the series for e^{ntM}).

`Proposition.` ***The solution to the system*** $\dot{\mathbf{x}} = A\mathbf{x}$ ***with the initial condition*** $\mathbf{x}(0)$ ***is given by the formula*** $\mathbf{x}(t) = e^{tA}\mathbf{x}(0)$***.***

`Proof.` Differentiating the series $\sum_0^\infty t^kA^k/k!$ we find

$$\frac{d}{dt}e^{tA} = \sum_{k=1}^{\infty} \frac{t^{k-1}A^k}{(k-1)!} = \sum_{k=0}^{\infty} \frac{t^kA^{k+1}}{k!} = Ae^{tA}$$

and hence $\frac{d}{dt}e^{tA}\mathbf{x}(0) = A(e^{tA}\mathbf{x}(0))$. Thus $\mathbf{x}(t)$ satisfies the ODE system. At $t = 0$ we have $e^{0A}\mathbf{x}(0) = I\mathbf{x}(0) = \mathbf{x}(0)$ and therefore the initial condition is also satisfied. □

`Remark.` This Proposition uses the exponential fucntion $A \mapsto e^A$ of a matrix (or of a linear operator). The exponential fucntion satisfies $e^Ae^B = e^{A+B}$ ***provided that*** A ***and*** B ***commute***. This can be proved the same way as this is done in the supplement *Complex Numbers* for the exponential function of the complex argument.

The proposition reduces the problem of solving the ODE system $\dot{\mathbf{x}} = A\mathbf{x}$ to computation of the exponential function e^{tA} of a matrix. Notice that if $A = CBC^{-1}$ then

$$A^k = CBC^{-1}CBC^{-1}CBC^{-1}... = CBBB...C^{-1} = CB^kC^{-1},$$

and therefore $e^{tA} = Ce^{tB}C^{-1}$. This observation reduces computation of e^{tA} to that of e^{tB} where the Jordan normal form of A can be taken on the role of B.

`Example 4.` Let Λ be a diagonal matrix with the diagonal entries $\lambda_1, ..., \lambda_n$. Then Λ^k is a diagonal matrix with the diagonal entries $\lambda_1^k, ..., \lambda_n^k$ and hence

$$e^{t\Lambda} = I + t\Lambda + \frac{t^2}{2}\Lambda^2 + ... = \begin{bmatrix} e^{\lambda_1 t} & 0 & ... \\ & ... & \\ ... & 0 & e^{\lambda_n t} \end{bmatrix}.$$

`Example 5.` Let N be a nilpotent Jordan cell of size m. We have (for $m = 4$):

$$N = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad N^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad N^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and $N^4 = 0$. Generalizing to arbitrary m, we find:

$$e^{tN} = I + tN + \frac{t^2}{2}N^2 + \frac{t^3}{6}N^3 + ... = \begin{bmatrix} 1 & t & \frac{t^2}{2} & ... & \frac{t^{m-1}}{(m-1)!} \\ 0 & 1 & t & \frac{t^2}{2} & ... \\ 0 & 0 & 1 & t & ... \\ & & ... & & \\ & ... & & 0 & 1 \end{bmatrix}.$$

Let $\lambda I + N$ be the Jordan cell of size m with the eigenvalue λ. Then $e^{t(\lambda I+N)} = e^{t\lambda I}e^{tN} = e^{\lambda t}e^{tN}$. Here we use the multiplicative property of the matrix exponential function, valid since I and N commute.

Finally, let $A = \begin{bmatrix} B & \mathbf{0} \\ \mathbf{0} & D \end{bmatrix}$ be a block-diagonal square matrix. Then $A^k = \begin{bmatrix} B^k & \mathbf{0} \\ \mathbf{0} & D^k \end{bmatrix}$ and respectively $e^{tA} = \begin{bmatrix} e^{tB} & \mathbf{0} \\ \mathbf{0} & e^{tD} \end{bmatrix}$. Together with Examples 4 and 5, this shows how to compute the exponential function e^{tJ} for any Jordan normal form J: each Jordan cell has the form $\lambda I + N$ and should be replaced by $e^{\lambda t}e^{tN}$. Since any square matrix A can be reduced to one of the Jordan normal matrices J by similarity transformations, we conclude that $e^{tA} = Ce^{tJ}C^{-1}$ with suitable invertible C.

Applying Example 3 to ODE systems $\dot{\mathbf{x}} = A\mathbf{x}$ we arrive at the following conclusion. Suppose that the characteristic polynomial of A has n distinct roots $\lambda_1, ..., \lambda_n$. Let C be the matrix whose columns are the corresponding n complex eigenvectors $\mathbf{v}_1, ..., \mathbf{v}_n$. Then the solution with the initial condition $\mathbf{x}(0)$ is given by the formula

$$\mathbf{x}(t) = C \begin{bmatrix} e^{\lambda_1 t} & 0 & ... \\ & ... & \\ ... & 0 & e^{\lambda_n t} \end{bmatrix} C^{-1}\mathbf{x}(0).$$

Notice that $C^{-1}\mathbf{x}(0)$ here (as well as $\mathbf{x}(0)$) is a column $\mathbf{c} = (c_1, ..., c_n)^t$ of arbitrary constants, while the columns of $Ce^{t\Lambda}$ are $e^{\lambda_i t}\mathbf{v}_i$. We conclude that the general solution formula reads

$$\mathbf{x}(t) = c_1 e^{\lambda_1 t}\mathbf{v}_1 + ... + c_n e^{\lambda_n t}\mathbf{v}_n.$$

This formula involves eigenvalues λ_i of A, the corresponding eigenvectors $\mathbf{v}_i$, and the arbitrary complex constants c_i, to be found from the system of linear equations $\mathbf{x}(0) = c_1\mathbf{v}_1 + ... + c_n\mathbf{v}_n$.

`Example 6.` The ODE system

$$\begin{array}{rcl} \dot{x}_1 & = & 2x_1 + x_2 \\ \dot{x}_2 & = & x_1 + 3x_2 - x_3 \\ \dot{x}_3 & = & 2x_2 + 3x_3 - x_1 \end{array} \quad \text{has the matrix } A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & -1 \\ -1 & 2 & 3 \end{bmatrix}.$$

The characteristic polynomial of A is $\lambda^3 - 8\lambda^2 + 22\lambda - 20$. It has a real root $\lambda_0 = 2$, and factors as $(\lambda - 2)(\lambda^2 - 6\lambda + 10)$. Thus is has two complex conjugate roots $\lambda_\pm = 3 \pm i$. The eigenvectors $\mathbf{v}_0 = (1, 0, 1)^t$ and $\mathbf{v}_\pm = (1, 1 \pm i, 2 \mp i)^t$ are found from the systems of linear equations:

$$\begin{array}{rcl} 2x_1 + x_2 & = & 2x_1 \\ x_1 + 3x_2 - x_3 & = & 2x_2 \\ -x_1 + 2x_2 + 3x_3 & = & 2x_3 \end{array} \quad \text{and} \quad \begin{array}{rcl} 2x_1 + x_2 & = & (3 \pm i)x_1 \\ x_1 + 3x_2 - x_3 & = & (3 \pm i)x_2 \\ -x_1 + 2x_2 + 3x_3 & = & (3 \pm i)x_3 \end{array}.$$

Thus the general complex solution is a linear combination of the solutions

$$e^{2t}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \; e^{(3+i)t}\begin{bmatrix} 1 \\ 1+i \\ 2-i \end{bmatrix}, \; e^{(3-i)t}\begin{bmatrix} 1 \\ 1-i \\ 2+i \end{bmatrix}$$

with arbitrary complex coefficients. Real solutions can be extracted from the complex ones by taking their real and imaginary parts:

$$e^{3t}\begin{bmatrix} \cos t \\ \cos t - \sin t \\ 2\cos t + \sin t \end{bmatrix},\ e^{3t}\begin{bmatrix} \sin t \\ \cos t + \sin t \\ 2\sin t - \cos t \end{bmatrix}.$$

Thus the general real solution is described by the formulas

$$\begin{array}{rclllll} x_1(t) & = & c_1e^t & + & c_2e^{3t}\cos t & + & c_3e^{3t}\sin t \\ x_2(t) & = & & & c_2e^{3t}(\cos t - \sin t) & + & c_3e^{3t}(2\cos t + \sin t) \\ x_3(t) & = & c_1e^t & + & c_2e^{3t}(\cos t + \sin t) & + & c_3e^{3t}(2\sin t - \cos t) \end{array},$$

where c_1, c_2, c_3 are arbitrary *real* constants. At $t = 0$ we have

$$x_1(0) = c_1 + c_2,\ x_2(0) = c_2 + 2c_3,\ x_3(0) = c_1 + c_2 - c_3.$$

Given the initial values $(x_1(0), x_2(0), x_3(0))$ the corresponding constants c_1, c_2, c_3 can be found from this system of linear algebraic equations.

In general, even if the characteristic polynomial of A has multiple roots, our theory (and Examples 4, 5) show that ***all components of solutions to the ODE system*** $\dot{\mathbf{x}} = A\mathbf{x}$ ***are expressible as linear combinations of functions*** $t^ke^{\lambda t}$, $t^ke^{at}\cos bt$, $t^ke^{at}\sin bt$, ***where*** λ ***are real eigenvalues,*** $a \pm ib$ ***are complex eigenvalues, and*** $k = 0, 1, 2, ...$ ***is to be smaller than the multiplicity of the corresponding eigenvalue.*** This observation suggests to approach the ODE systems with multiple eigenvalues in the following way avoiding explicit similarity transformation to the Jordan normal form: look for the general solution in the form of linear combinations of these functions with arbitrary coefficients by substituting them into the equations, and find the relations between the arbitrary constants from the resulting system of linear algebraic equations.

`Example 7.` The ODE system

$$\begin{array}{rcl} \dot{x}_1 & = & 2x_1 + x_2 + x_3 \\ \dot{x}_2 & = & -3x_1 - 2x_2 - 3x_3 \\ \dot{x}_3 & = & 2x_1 + 2x_2 + 3x_3 \end{array} \quad \text{has the matrix } A = \begin{bmatrix} 2 & 1 & 1 \\ -3 & -2 & -3 \\ 2 & 2 & 3 \end{bmatrix}$$

with the characteristic polynomial $\lambda^3 - 3\lambda^2 + 3\lambda - 1 = (\lambda - 1)^3$. Thus we can look for solutions in the form

$$x_1 = e^t(a_1 + b_1t + c_1t^2),\ x_2 = e^t(a_2 + b_2t + c_2t^2),\ x_3 = e^t(a_3 + b_3t + c_3t^2).$$

Substituting into the ODE system, using the notation $\sum a_i =: A$, $\sum b_i =: B$, $\sum c_i =: C$, and omitting the factor e^t, we get:

$$\begin{aligned}(a_1+b_1)+(b_1+2c_1)t+c_1t^2 &= (a_1+A)+(b_1+B)t+(c_1+C)t^2,\\ (a_2+b_2)+(b_2+2c_2)t+c_2t^2 &= -(3A-a_2)-(3B-b_2)t-(3C-c_2)t^2,\\ (a_3+b_3)+(b_3+2c_3)t+c_3t^2 &= (2A+a_3)+(2B+b_3)t+(2C+c_3)t^2.\end{aligned}$$

Introducing the notation $P := A+Bt+Ct^2$, we rewrite the system of 9 linear equations in 9 unknowns in the form

$$b_1+2c_1t = P,\ b_2+2c_2t = -3P,\ b_3+2c_3t = 2P.$$

This yields $b_1 = A, b_2 = -3A, b_3 = 2A$, and (since $B = A-3A+2A = 0$), we find $c_1 = c_2 = c_3 = 0$. Thus, the general solution, depending on 3 arbitrary constants a_1, a_2, a_3, reads:

$$\begin{aligned}x_1 &= e^t(a_1+(a_1+a_2+a_3)t)\\ x_2 &= e^t(a_2-3(a_1+a_2+a_3)t)\\ x_3 &= e^t(a_3+2(a_1+a_2+a_3)t).\end{aligned}$$

In particular, since t^2 does not occur in the formulas, we can conclude that the Jordan form of our matrix has only Jordan cells of size 1 or 2 (and hence one of each, for the total size must be 3):

$$J = \begin{bmatrix} 1 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

EXERCISES

441.* Give an example of (non-commuting) 2×2 matrices A and B for which $e^Ae^B \neq e^{A+B}$. ↯

442. Solve the following ODE systems. Find the solution satisfying the initial condition $x_1(0) = 1,\ x_2(0) = 0,\ x_3(0) = 0$:

(a) $\lambda_1 = 1$
$\dot x_1 = 3x_1 - x_2 + x_3$
$\dot x_2 = x_1 + x_2 + x_3$
$\dot x_3 = 4x_1 - x_2 + 4x_3$

(b) $\lambda_1 = 1$
$\dot x_1 = -3x_1 + 4x_2 - 2x_3$
$\dot x_2 = x_1 + x_3$
$\dot x_3 = 6x_1 - 6x_2 + 5x_3$

(c) $\lambda_1 = 1$
$\dot x_1 = x_1 - x_2 - x_3$
$\dot x_2 = x_1 + x_2$
$\dot x_3 = 3x_1 + x_3$

(d) $\lambda_1 = 2$
$\dot x_1 = 4x_1 - x_2 - x_3$
$\dot x_2 = x_1 + 2x_2 - x_3$
$\dot x_3 = x_1 - x_2 + 2x_3$

(e) $\lambda_1 = 1$
$\dot x_1 = -x_1 + x_2 - 2x_3$
$\dot x_2 = 4x_1 + x_2$
$\dot x_3 = 2x_1 + x_2 - x_3$

$$
\begin{array}{ll}
(f)\ \ \lambda_1 = 2 & (g)\ \ \lambda_1 = 1 \\
\dot{x}_1 = 4x_1 - x_2 & \dot{x}_1 = 2x_1 - x_2 - x_3 \\
\dot{x}_2 = 3x_1 + x_2 - x_3 & \dot{x}_2 = 2x_1 - x_2 - 2x_3 \\
\dot{x}_3 = x_1 + x_3 & \dot{x}_3 = -x_1 + x_2 + 2x_3
\end{array}.
$$

443. Find the Jordan canonical forms of the 3×3-matrices of the ODE systems (a–g) from the previous exercise.

444.* Prove the following identity: $\det e^A = e^{\operatorname{tr} A}$. ↯

Higher Order Linear ODE

A linear homogeneous constant coefficient n-th order ODE

$$\frac{d^n}{dt^n}x + a_1 \frac{d^{n-1}}{dt^{n-1}}x + ... + a_{n-1}\frac{d}{dt}x + a_n x = 0$$

can be rewritten, by introducing the notations

$$x_1 := x,\ x_2 := \dot{x},\ x_2 := \ddot{x}, ..., x_{n-1} := d^{n-1}x/dt^{n-1},$$

as a system $\dot{\mathbf{x}} = A\mathbf{x}$ of n first order ODEs with the matrix

$$A = \begin{bmatrix} 0 & 1 & 0 & ... & 0 \\ 0 & 0 & 1 & 0 & ... \\ & & ... & & \\ 0 & ... & 0 & 0 & 1 \\ -a_n & -a_{n-1} & ... & -a_2 & -a_1 \end{bmatrix},$$

Then our theory of linear ODE systems applies. There are however some simplifications which are due to the special form of the matrix A. First, computing the characteristic polynomial of A we find

$$\det(\lambda I - A) = \lambda^n + a_1\lambda^{n-1} + ... + a_{n-1}\lambda + a_n.$$

Thus the polynomial can be easily read off the high order differential equation. Next, let $\lambda_1, ..., \lambda_r$ be the roots of the characteristic polynomial, and $m_1, ..., m_r$ their multiplicities ($m_1 + ... + m_r = n$). Then it is clear that the solutions $x(t) = x_1(t)$ must have the form

$$e^{\lambda_1 t}P_1(t) + ... + e^{\lambda_r t}P_r(t),$$

where $P_i = a_0 + a_1 t + ... + a_{m_i - 1}t^{m_i - 1}$ is a polynomial of degree $< m_i$. The total number of arbitrary coefficients in these polynomials equals $m_1 + ... + m_r = n$. On the other hand, the general solution to the n-th order ODE must depend on n arbitrary initial values $(x(0), \dot{x}(0), ..., x^{(n-1)}(0))$. This can be justified by the general Existence and Uniqueness Theorem for solutions of ODE systems. We conclude that ***each of the functions***

$$e^{\lambda_1 t},\ e^{\lambda_1 t}t,\ ...,\ e^{\lambda_1 t}t^{m_1-1},\ ...,\ e^{\lambda_r t},\ e^{\lambda_r t}t,\ ...,\ e^{\lambda_r t}t^{m_r-1}$$

must satisfy our differential equation, and any (complex) solution is uniquely written as a linear combination of these functions with suitable (complex) coefficients. (In other words, these functions form a *basis* of the space of complex solutions to the differential equation.)

`Example 8:` $x^{(xii)} - 3x^{(viii)} + 3x^{(iv)} - x = 0$ has the characteristic polynomial $\lambda^{12} - 3\lambda^8 + 3\lambda^4 - 1 = (\lambda-1)^3(\lambda+1)^3(\lambda-i)^3(\lambda+i)^3$. The following 12 functions form therefore a complex basis of the solution space:

$$e^t,\ te^t,\ t^2e^t,\ e^{-t},\ te^{-t},\ t^2e^{-t},\ e^{it},\ te^{it},\ t^2e^{it},\ e^{-it},\ te^{-it},\ t^2e^{-it}.$$

Of course, a basis of the space of real solutions is obtained by taking real and imaginary parts of the complex basis:

$$e^t,\ te^t,\ t^2e^t,\ e^{-t},\ te^{-t},\ t^2e^{-t},$$
$$\mathbf{cos}\, t,\ \mathbf{sin}\, t,\ t\,\mathbf{cos}\, t,\ t\,\mathbf{sin}\, t,\ t^2\,\mathbf{cos}\, t,\ t^2\,\mathbf{sin}\, t.$$

`Remark.` The fact that the functions $e^{\lambda_i t}t^k,\ k < m_i,\ i = 1, ..., r$, form a basis of the space of solutions to the differential equation with the characteristic polynomial $(\lambda - \lambda_1)^{m_1}...(\lambda - \lambda_r)^{m_r}$ is not hard to check directly, without a reference to linear algebra and Existence and Uniqueness Theorem. However, it is useful to understand how this property of the equation is related to the Jordan structure of the corresponding matrix A. In fact the Jordan normal form of the matrix A consists of exactly r Jordan cells — one cell of size m_i for each eigenvalue λ_i. This simplification can be explained as follows. For every λ, the matrix $\lambda I - A$ has rank $n-1$ *at least* (due to the presence of the $(n-1) \times (n-1)$ identity submatrix in the right upper corner of A). This guarantees that the eigenspaces of A have dimension 1 *at most*, and hence that A cannot have more than one Jordan cell corresponding to the same root of the characteristic polynomial. Using this property, the reader can check now that the formulation of the Jordan Theorem in terms of differential equations given in the Introduction is indeed equivalent to the matrix formulation given in the previous section.

EXERCISES

445. Solve the following ODE, and find the solution satisfying the initial condition $x(0) = 1, \dot{x}(0) = 0, ..., x^{(n-1)}(0) = 0$: $(a)\ \ x^{(3)} - 8x = 0,$

$$(b)\ \ x^{(4)} + 4x = 0, \quad (c)\ \ x^{(6)} + 64x = 0, \quad (d)\ \ x^{(5)} - 10x^{(3)} + 9x = 0,$$
$$(e)\ \ x^{(3)} - 3\dot{x} - 2x = 0, \quad (f)\ \ x^{(5)} - 6x^{(4)} + x^{(3)} = 0,$$
$$(g)\ \ x^{(5)} + 8x^{(3)} + 16\dot{x} = 0, \quad (h)\ \ x^{(4)} + 4\ddot{x} + 3x = 0.$$

446. Rewrite the ODE system

$$\begin{array}{l} \ddot{x}_1 + 4\dot{x}_1 - 2x_1 - 2\dot{x}_2 - x_2 = 0 \\ \ddot{x}_1 - 4\dot{x}_1 - \ddot{x}_2 + 2\dot{x}_2 + 2x_2 = 0 \end{array}$$

of two 2-nd order equations in the form of a linear ODE system $\dot{\mathbf{x}} = A\mathbf{x}$ of four 1-st order equations and solve it.

447. Show that every polynomial $\lambda^n + a_1\lambda^{n-1} + \cdots + a_n$ is the charactersitic polynomial of an $n \times n$-matrix.