4. Endomorphisms.

1 Introduction.

We saw in the previous chapter that an **endomorphism is a linear map from a vector space to itself:** $t : U \longrightarrow U$. If we fix a basis of U

$$B = \{\bar{u}_1, \ldots, \bar{u}_n\},$$

we can work with the matrix of t relative to this basis.

$$t(\bar{u}_i) = t^i_j \bar{u}_i, \qquad \Rightarrow \qquad T_{BB} = (t^i_j).$$

T_{BB} is called **matrix of the endomorphism** t **with respect to the basis** B. We will usually simply denote it T_B.

Given another basis U

$$B' = \{\bar{u}'_1, \ldots, \bar{u}'_n\},$$

we can consider the change-of-basis matrix $M_{B'B}$. It allows us to connect the matrices of t relative to both bases:

$$T_{B'B'} = M_{B'B} T_{BB} M_{BB'} = (M_{BB'})^{-1} T_{BB} M_{B'B}$$

We deduce that

Proposition 1.1 *Two matrices are associated with the same endomorphism if and only if they are similar.*

Let us call $End(U)$ the vector space $Hom(U, U)$ of endomorphisms of U. We saw in the previous chapter that given any basis B, the map

$$\begin{array}{cccc} \pi : & End(U) & \longrightarrow & M_{n\times n} \\ & t & \longrightarrow & T_B \end{array}$$

is an isomorphism. Therefore, the study of endomorphisms is equivalent to the study of square matrices and the similarity relation.

One of the fundamental goals of this chapter will be to find bases in which the matrix associated to a given endomorphism is as simple as possible. Equivalently, given a square matrix find a similar matrix as simple as possible.

2 Eigenvalues and eigenvectors.

2.1 Definition and properties.

Definition 2.1 *Given an endomorphism* $t : U \longrightarrow U$*, a scalar* λ *is said to be an* **eigenvalue** *of* t *if there exists a nonzero vector* $\bar{x}$ *such that:*

$$t(\bar{x}) = \lambda \bar{x}.$$

The scalar λ *satisfying this condition is also called* **proper value** *or* **characteristic value***; the vectors* $\bar{x}$ *satisfying this condition are called* **eigenvectors** *or* **proper vectors or characteristic vectors** *associated to* λ.

We next make some considerations concerning this definition.

1. *If* $\lambda = 0$ *is an eigenvalue of* t*, then the eigenvectors associated to* λ *are exactly the vectors in the kernel of* t.
2. *An eigenvector* $\bar{x} \neq 0$ *cannot be associated with two different eigenvalues.*

   Proof: If $\bar{x}$ is associated with the eigenvalues λ, μ we have:

   $$\lambda\bar{x} = t(\bar{x}) = \mu\bar{x} \quad \Rightarrow \quad (\lambda - \mu)\bar{x} = 0.$$

   Since $\bar{x} \neq \bar{0}$, we deduce that $\lambda - \mu = 0$ and therefore both eigenvalues are equal.
3. **Once fixed a basis** B **in** U**, the matrix expression of the eigenvalue condition is**

   $$t(\bar{x}) = \lambda\bar{x} \iff T_B(x) = \lambda(x).$$

   Thus one can assign eigenvalues and eigenvectors to a square matrix.
4. *Two similar matrices have the same eigenvalues.*

   Proof: We saw that any two similar matrices can be regarded as matrices associated to the same endomorphism relative to two different bases. But the eigenvalues of a matrix are the same as the eigenvalues of the associated endomorphism, and these do not depend on the basis (actually the definition of an eigenvalue given above does not involve any basis). ■

2.2 Characteristic subspaces.

Definition 2.2 *The union of the zero vector with the set of eigenvectors associated to an eigenvalue* λ *will be denoted by* S_λ *and called the* **characteristic subspace** *or* **eigenspace associated to** λ*:*

$$S_\lambda = \{\bar{x} \in U |\, t(\bar{x}) = \lambda\bar{x}\}.$$

Proposition 2.3 *The set* S_λ *of eigenvectors associated to an eigenvalue* λ *is a vector subspace.*

Proof: First, $S_\lambda \neq \emptyset$ because $\bar{0} \in S_\lambda$. Moreover, given $\bar{x}, \bar{y} \in S_\lambda$ and $\alpha, \beta \in \mathbb{K}$:

$$t(\alpha\bar{x} + \beta\bar{y}) = \alpha t(\bar{x}) + \beta t(\bar{y}) \underset{\substack{\uparrow \\ \bar{x}, \bar{y} \in S_\lambda \iff t(\bar{x}) = \lambda\bar{x}, \quad t(\bar{y}) = \lambda\bar{y}}}{=} \alpha\lambda\bar{x} + \beta\lambda\bar{y} = \lambda(\alpha\bar{x} + \beta\bar{y})$$

and then $\alpha\bar{x} + \beta\bar{y} \in S_\lambda$. ■

Remark 2.4 *A different way to prove the above result is as follows. Given an endomorphism t and an eigenvalue λ, we define the map*

$$t' : U \longrightarrow U; \qquad t'(\bar{x}) = t(\bar{x}) - \lambda\bar{x} \iff t' = t - \lambda Id.$$

This is a linear map because it is the difference of two linear maps. Furthermore

$$S_\lambda = \{\bar{x} \in U |\, t(\bar{x}) = \lambda\bar{x}\} = \{\bar{x} \in U |\, t(\bar{x}) - \lambda\bar{x} = \bar{0}\} = ker(t - \lambda Id).$$

We see that

$$\boxed{S_\lambda = ker(t - \lambda Id)}$$

is a vector subspace because it is the kernel of a linear map.

Proposition 2.5 *If λ_1 and λ_2 are different eigenvalues of an endomorphism then $S_{\lambda_1} \cap S_{\lambda_2} = \{\bar{0}\}$.*

Equivalently, the sum $S_{\lambda_1} + S_{\lambda_2}$ is a direct sum.

Proof: It follows from the fact that an eigenvector cannot be associated to different eigenvalues. ■

Proposition 2.6 *If $\lambda_1, \ldots, \lambda_k$ are pairwise different eigenvalues of an endomorphism then $(S_{\lambda_1} + \ldots + S_{\lambda_i}) \cap S_{\lambda_{i+1}} = \{0\}$ for $i = 1, \ldots, k-1$.*

Equivalently, the sum $S_{\lambda_1} + \ldots + S_{\lambda_k}$ is a direct sum.

Proof: We will prove it by induction. Note that the case $k = 2$ has already been proved.

Suppose it is true for i eigenspaces and let us prove it for $i + 1$.

Fix $\bar{x} \in (S_{\lambda_1} + \ldots + S_{\lambda_i}) \cap S_{\lambda_{i+1}}$. Then we have

$$\bar{x} = \bar{x}_1 + \ldots + \bar{x}_i, \qquad \text{with} \quad x_j \in S_{\lambda_j}, \quad j = 1, \ldots, i.$$

Applying t we obtain:

$$t(\bar{x}) = t(\bar{x}_1) + \ldots + t(\bar{x}_i) = \lambda_1\bar{x}_1 + \ldots + \lambda_i\bar{x}_i.$$

On the other hand, since $\bar{x} \in S_{\lambda_{i+1}}$, we have:

$$t(\bar{x}) = \lambda_{i+1}\bar{x} = \lambda_{i+1}\bar{x}_1 + \ldots + \lambda_{i+1}\bar{x}_i.$$

Comparing both expressions:

$$(\lambda_1 - \lambda_{i+1})\bar{x}_1 + \ldots + (\lambda_i - \lambda_{i+1})\bar{x}_i = \bar{0}$$

Since $S_{\lambda_1}, \ldots, S_{\lambda_i}$ is a direct sum, the decomposition of $\bar{0}$ as a sum of elements of $S_{\lambda_1} \ldots, S_{\lambda_i}$ is unique. Hence

$$(\lambda_1 - \lambda_{i+1})\bar{x}_1 = \ldots = (\lambda_i - \lambda_{i+1})\bar{x}_i = \bar{0}$$

Since all eigenvalues are different, $\bar{x}_1 = \ldots = \bar{x}_i = \bar{0}$ and $\bar{x} = \bar{0}$. ■

2.3 Characteristic polynomial.

Definition 2.7 *Let $t : U \longrightarrow U$ be an endomorphism, B a basis of U and T the matrix associated to t with respect to this basis. The* **characteristic polynomial** *of t is*

$$p_t(\lambda) = det(T - \lambda Id).$$

Proposition 2.8 *The characteristic polynomial is independent of the choice of the basis.*

Proof: Suppose that T and T' are two matrices associated to the same endomorphism $t : U \longrightarrow U$, with respect to two different bases. We know that T and T' are similar, that is, there exists a regular matrix P satisfying that $T' = P^{-1}TP$. Now:

$$|T' - \lambda I| = |P^{-1}TP - \lambda P^{-1}P| = |P^{-1}(T - \lambda I)P| = |P^{-1}||T - \lambda I||P| = |T - \lambda I|.$$

■

Theorem 2.9 *Given any endomorphism $t : U \longrightarrow U$, λ is an eigenvalue of t if and only if $p_t(\lambda) = 0$.*

Proof: By definiton λ is an eigenvalue of t if and only if there is a vector $\bar{x} \neq 0$ with:

$$t(\bar{x}) = \lambda\bar{x}.$$

If we choose a basis B of U, we can write this condition as:

$$T_B(x) = \lambda(x) \qquad \text{or equivalently} \qquad (T_B - \lambda I)(x) = \bar{0}.$$

That is, λ is an eigenvalue of t if and only if the homogeneous system

$$(T_B - \lambda I)(x) = \bar{0}$$

has a non trivial solution. But this happens if and only if the matrix of the system is singular, that is:

$$|T - \lambda I| = 0 \iff p_t(\lambda) = 0.$$

■

2.4 Algebraic and geometric multiplicity of an eigenvalue.

Definition 2.10 *Given an endomorphism $t : U \longrightarrow U$ and an eigenvalue λ of t we define*

- **Algebraic multiplicity of** λ*: It is the multiplicity of λ as a root of the characteristic polynomial $p_t(\lambda)$. It is denoted by $m(\lambda)$.*

- **Geometric multiplicity of** λ*: It is the dimension of the characteristic space associated to λ. It is denoted by $d(\lambda)$.*

Let us see some properties of these multiplicities:

1. *The geometric multiplicity of an eigenvalue is always greater than* 0.

$$d(\lambda) \geq 1$$

Proof: It is sufficient to note that by definition, any eigenvalue has always a nonzero eigenvector.

2. *If* n *is the dimension of the vector space* U *and* T *is the associated matrix to* t *with respect to some basis, a frequent way to compute the geometric multiplicity is :*

$$\boxed{d(\lambda) = n - rank(T - \lambda I)}$$

Proof: If follows from:

$$d(\lambda) = dim(S_\lambda) = dim(ker(t - \lambda I)) = n - dim(im(t - \lambda I)) = n - rank(T - \lambda I).$$

3. $\boxed{\textit{Algebraic multiplicity} \geq \textit{Geometric multuplicity}}$ $\boxed{m(\lambda_i) \geq d(\lambda_i)}$

Proof: Suppose λ_i is an eigenvalue of the endomorphism t. Let $d = d(\lambda_i)$ be its geometric multiplicity. We consider a basis of the characteristic subspace S_{λ_i}:

$$\{\bar{u}_1, \ldots, \bar{u}_d\}$$

and we complete it up to a base B of U:

$$B = \{\bar{u}_1, \ldots, \bar{u}_d, \bar{u}_{d+1}, \ldots, \bar{u}_n\}.$$

Let us see what the matrix T associated to t with respect to this basis looks like. We know that

$$t(\bar{u}_1) = \lambda_i \bar{u}_1; \quad \ldots \quad t(\bar{u}_d) = \lambda_i \bar{u}_d.$$

Thus the matrix T has the form:

$$T = \left(\begin{array}{cccc|c} \lambda_i & 0 & \ldots & 0 & \\ 0 & \lambda_i & \ldots & 0 & \\ \vdots & \vdots & \ddots & \vdots & A \\ 0 & 0 & \ldots & \lambda_i & \\ \hline & \Omega & & & B \end{array}\right) \quad \text{with} \quad \begin{array}{l} A \in \mathcal{M}_{d\times(n-d)}(\mathbb{K}). \\ B \in \mathcal{M}_{(n-d)\times(n-d)}(\mathbb{K}). \end{array}$$

Now, if we compute the characteristic polynomial of t using this matrix, we get:

$$|T - \lambda I| = \left|\begin{array}{cccc|c} \lambda_i - \lambda & 0 & \ldots & 0 & \\ 0 & \lambda_i - \lambda & \ldots & 0 & \\ \vdots & \vdots & \ddots & \vdots & A \\ 0 & 0 & \ldots & \lambda_i - \lambda & \\ \hline & \Omega & & & B - \lambda I \end{array}\right| = (\lambda_i - \lambda)^d |B - \lambda I|$$

We see that λ_i is a root of the characteristic polynomial of t with **at least** multiplicity d, therefore:

$$m(\lambda_i) \geq d(\lambda_i).$$

■

4. *If* λ *is an eigenvalue with* $m(\lambda) = 1$*, then* $d(\lambda) = m(\lambda) = 1$.

5. *The maximum number of independent eigenvectors of an endomorphism is*

$$d(\lambda_1) + \ldots + d(\lambda_k)$$

where $\lambda_1, \ldots, \lambda_k$ *are all eigenvalues of* t.

Proof: Note that the sum of all spaces of characteristic vectors is a direct sum. In addition, we have:

$$d(\lambda_1) + \ldots + d(\lambda_k) \leq m(\lambda_1) + \ldots + m(\lambda_k) \leq n.$$

3 Diagonalization and similarity.

3.1 Diagonalizable endomorphisms.

Definition 3.1 *An endomorphism* $t : U \longrightarrow U$ *is said to be* **diagonalizable** *if there is a basis* B *of* U *such that the matrix associated to* t *with respect to* B *is diagonal:*

$$t \textit{ diagonalizable} \iff \exists \textit{ basis } B / \ T_B \textit{ diagonal.}$$

We can define the analogous concept for square matrices:

Definition 3.2 *A matrix* T *is said to be* **diagonalizable** *when it is a similar to a diagonal matrix:*

$$T \textit{ diagonalizable} \iff \exists D \textit{ diagonal and } P \textit{ regular }, \ D = P^{-1}TP.$$

Since two matrices are similar precisely when they are associated to the same endomorphism, the study of diagonalization of matrices is equivalent to that of diagonalization of endomorphisms.

3.2 Matrix of an endomorphism respect to a basis of eigenvectors.

Theorem 3.3 *The necessary and sufficient condition for the associated matrix to an endomorphism to be diagonal is that it is expressed with respect to a basis of eigenvectors.*

Proof:

$\Longrightarrow$: Suppose that $B = \{\bar{u}_1, \ldots, \bar{u}_n\}$ is a basis of U, such that T_B is diagonal:

$$T_B = \begin{pmatrix} t_1 & 0 & \ldots & 0 \\ 0 & t_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & t_n \end{pmatrix}.$$

This means that:

$$t(\bar{u}_1) = t_1\bar{u}_1; \quad \ldots \quad t(\bar{u}_n) = t_n\bar{u}_n.$$

and hence all vectors of B are eigenvectors.

$\Longleftarrow$: Suppose that $B = \{\bar{u}_1, \ldots, \bar{u}_n\}$ is a basis of eigenvectors of U, then we have:

$$t(\bar{u}_1) = \lambda_1\bar{u}_1; \quad \ldots \quad t(\bar{u}_n) = \lambda_n\bar{u}_n.$$

It follows that the associate matrix T_B is

$$T_B = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}.$$

3.3 Characterization of a diagonalizable endomorphism.

We have seen that

Proposition 3.4 *An endomorphism $t : U \longrightarrow U$ is diagonalizable if and only if there is a basis of U formed by eigenvectors of t.*

This result is usually reformulated as follows:

Theorem 3.5 *An endomorphism t of an n-dimensional space U is diagonalizable if and only if the sum of the algebraic multiplicities is equal to the dimension of U and the algebraic multiplicities are equal to the geometric multiplicities. Equivalently:*

$$t \text{ diagonalizable} \iff \begin{cases} m(\lambda_1) + \ldots + m(\lambda_k) = n. \\ \\ d(\lambda_1) = m(\lambda_1), \ldots, d(\lambda_k) = m(\lambda_k). \end{cases}$$

where $\lambda_1, \ldots, \lambda_k$ is the set of eigenvalues of t.

Proof: It is sufficient to note that if t is diagonalizable, then there is a basis of eigenvectors of U. In this basis the matrix T is diagonal and therefore

$$n = d(\lambda_1) + \ldots + d(\lambda_k) \leq m(\lambda_1) + \ldots + m(\lambda_k) = n$$

Reciprocally if $d(\lambda_1) + \ldots + d(\lambda_k) = n$, since the sum of characteristic subspaces $S_{\lambda_1} + \ldots + S_{\lambda_k}$ is a direct sum, we can choose a basis of U formed by eigenvectors. Therefore t is diagonalizable. ■

From the previous proof we deduce that the characterization can be written simply as follows:

Corollary 3.6 *If t is an endomorphism of a vector space U of dimension n:*

$$t \text{ diagonalizable} \iff d(\lambda_1) + \ldots + d(\lambda_k) = n$$

where $\lambda_1, \ldots, \lambda_k$ is the set of eigenvalues of t.

3.4 Steps for the diagonalization of an endomorphism t.

Suppose we have a vector space U of dimension n and an endomorphism $t : U \longrightarrow U$ whose matrix with respect to a basis B is T_B. The steps to diagonalize it (if it is diagonalizable at all) are as follows:

1. Compute the characteristic polynomial $p_t(\lambda) = |T_B - \lambda I|$.
2. Find the roots of the characteristic polynomial, i. e. the eigenvalues of t. We will obtain values $\lambda_1, \ldots, \lambda_k$ with algebraic multiplicities $m_{\lambda_1}, \ldots, m_{\lambda_k}$.
   (a) If $m_{\lambda_1} + \ldots + m_{\lambda_k} < n$ the endomorphism is **not diagonalizable**.
   (b) If $m_{\lambda_1} + \ldots + m_{\lambda_k} = n$, we compute the geometric multiplicities of each eigenvalue of t:

   $$d(\lambda_i) = dim(S_{\lambda_i}) = n - rank(T - \lambda I) \qquad i = 1, \ldots, k.$$

   i. If some geometric multiplicity does not coincide with the corresponding algebraic one, then t is **not diagonalizable**.
   ii. If all algebraic multiplicities are equal to the corresponding geometric ones, then t is **diagonalizable**.
   The corresponding diagonal matrix has all the eigenvalues on its diagonal, repeated as many times as their algebraic multiplicities:

$$D = \left(\begin{array}{c|c|c} \begin{matrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_1 \end{matrix} & \ldots & \Omega \\ \hline \vdots & \ddots & \vdots \\ \hline \Omega & \ldots & \begin{matrix} \lambda_k & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_k \end{matrix} \end{array}\right)$$

so that

$$D = (M_{BB'})^{-1} T_B M_{BB'}$$

where $M_{BB'}$ is the change-of-basis matrix from the basis B' of eigenvectors of t to the initial basis B.
The eigenvectors of each eigenvalue λ_i of t are computed by solving the system

$$S_{\lambda_i} = ker(t - \lambda_i I) = \left\{ (x^1, \dots, x^n) \in U | \, (T - \lambda_i I) \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = \bar{0} \right\},$$

for each $i = 1, 2, \dots, k$.
In other words **the columns of $M_{BB'}$ are the coordinates, with respect to the initial basis B, of n linearly independent eigenvectors of t.** In addition, they **must be ordered consistently with the eigenvalues, so that if the eigenvalue λ_s appears on the j-th column of the diagonal matrix D, then an eigenvector associated with the eigenvalue λ_s must appear on the column j-th of $M_{BB'}$.**

3.5 Application to obtaining powers of matrices.

Suppose that A is a diagonalizable matrix and we are interested in finding A^p. We can proceed as follows.

Since A is diagonalizable, there is a regular matrix P such that

$$D = P^{-1}AP \quad \text{where} \quad D = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}.$$

Then $A = PDP^{-1}$ and

$$\begin{aligned} A^k &= \underbrace{(PDP^{-1}) \cdot (PDP^{-1}) \cdot \dots \cdot (PDP^{-1})}_{k \text{ times}} = \\ &= P \underbrace{(P^{-1}PD) \cdot (P^{-1}PD) \dots \cdot (P^{-1}P \cdot D)}_{k \text{ times}} P^{-1} = PD^kP^{-1}. \end{aligned}$$

That is,

$$\boxed{A^k = PD^kP^{-1}} \quad \text{with} \quad D^k = \begin{pmatrix} d_1^k & 0 & \dots & 0 \\ 0 & d_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n^k \end{pmatrix}.$$

4 Triangularization and similarity.

4.1 Triangularizable endomorphisms.

Definition 4.1 *An endomorphism $t : U \longrightarrow U$ is said to be* **triangularizable** *if there is a basis B of U such that the associated matrix T_B is triangular.*

$$t \text{ triangularizable} \iff \exists \text{ basis } B / \ T_B \text{ triangular.}$$

The analogous concept for square matrices would be

Definition 4.2 *A matrix T is said to be* **triangularizable** *when there is a triangular matrix similar to it:*

$$T \text{ triangularizable} \iff \exists J \text{ triangular and } P \text{ regular} \text{ , } J = P^{-1}TP.$$

Remark 4.3 *If the matrix of an endomorphism $t : U \longrightarrow U$ relative to a basis*

$$B = \{\bar{u}_1, \dots, \bar{u}_n\}$$

is **upper** *triangular, then there is basis B' such that the associated matrix is* **lower** *triangular and conversely. It is sufficient to take the vectors of B in the inverse order:*

$$B' = \{\bar{u}_n, \dots, \bar{u}_1\}.$$

We will look for **upper triangular** matrices similar to a given one. The basic result is the following:

Theorem 4.4 *Let U be an n-dimensional vector space. An endomorphism $t : U \longrightarrow U$ is triangularizable if and only if t has n eigenvalues (taking algebraic multiplicities into account). Namely*

$$t \text{ triangularizable} \iff m(\lambda_1) + \dots + m(\lambda_k) = n$$

where $\{\lambda_1, \dots, \lambda_k\}$ is the set of eigenvalues of t.

Proof:

$\Longrightarrow$: Suppose that t is triangularizable. Then there is a basis B such that the matrix T_B is triangular. If we calculate the characteristic polynomial using this basis, we get:

$$|T - \lambda I| = (t_{11} - \lambda)(t_{22} - \lambda) \dots (t_{nn} - \lambda)$$

and we see that there are exactly n eigenvalues (possibly not pairwise different).

$\Longleftarrow$: Suppose that t has n eigenvalues (possibly not pairwise different). Let us see that t is triangularizable. We will prove it by induction:

- For $n = 1$ it is true because a matrix of dimension 1×1 is always triangular.

- Suppose the result is true for $n - 1$ and prove it for n.

Let λ_1 be an eigenvalue. Then $d(\lambda_1) \geq 1$ and there is at least one eigenvector $\bar{x}_1 \neq 0$ associated to λ_1. We choose a basis B_1 whose first vector is $\bar{x}_1$:

$$B_1 = \{\bar{x}_1, \bar{u}_2, \dots, \bar{u}_n\}.$$

The associated matrix relative to this basis is

$$T_{B_1} = \left(\begin{array}{c|ccc} \lambda_1 & t_{12} & \dots & t_{1n} \\ \hline 0 & & & \\ \vdots & & T' & \\ 0 & & & \end{array}\right)$$

Note that

$$|T_{B_1} - \lambda I| = (\lambda_1 - \lambda)|T' - \lambda I|,$$

that is, the characteristic polynomial of t factorizes through the characteristic polynomial of T'. So if t has n eigenvalues, T' will have $n - 1$ eigenvalues.

Now by induction hypothesis, T' is triangularizable. We know that there is a regular matrix $P \in \mathcal{M}_{n-1\times n-1}(\mathbb{K})$ such that:

$$B = P^{-1}T'P \quad \text{where } B \text{ is upper triangular.}$$

Let Q be the matrix:

$$Q = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & P & \\ 0 & & & \end{array}\right)$$

Then:

$$\begin{aligned} Q^{-1}T_{B_1}Q &= \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & P^{-1} & \\ 0 & & & \end{array}\right) \left(\begin{array}{c|ccc} \lambda_1 & t_{12} & \dots & t_{1n} \\ \hline 0 & & & \\ \vdots & & T' & \\ 0 & & & \end{array}\right) \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & P & \\ 0 & & & \end{array}\right) = \\ &= \left(\begin{array}{c|ccc} \lambda_1 & t_{12} & \dots & t_{1n} \\ \hline 0 & & & \\ \vdots & & P^{-1}T' & \\ 0 & & & \end{array}\right) \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & P & \\ 0 & & & \end{array}\right) = \\ &= \left(\begin{array}{c|ccc} \lambda_1 & s_{12} & \dots & s_{1n} \\ \hline 0 & & & \\ \vdots & & P^{-1}T'P & \\ 0 & & & \end{array}\right) = \left(\begin{array}{c|ccc} \lambda_1 & s_{12} & \dots & s_{1n} \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right) = S \end{aligned}$$

where since B is upper triangular, S is upper triangular. We have proved that the matrix T_{B_1} associated with t is triangularizable by similarity and therefore t is triangularizable. ■

Remark 4.5 *If we are working on the field $\mathbb{K} = \mathbb{C}$ of complex numbers, the Fundamental Theorem of Algebra states that every polynomial of degree n has exactly n complex roots (same or different). Thus, every square matrix over the field of complex numbers is triangularizable.*

In the field of real numbers we do not have this result. There are polynomials of degree n that do not have n real solutions. Thus, there exist non-triangularizable square matrices.

4.2 Jordan canonical form.

Definition 4.6 *Given an scalar λ we call* **Jordan block** *or* **Jordan box** *associated to λ and of dimension m the $m \times m$ matrix*

$$J = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

Definition 4.7 *We call* **Jordan matrix** *to a matrix formed by several Jordan blocks associated with equal or different scalars, placed on the diagonal as follows:*

$$J = \left(\begin{array}{c|c|c|c} J_1 & \Omega & \dots & \Omega \\ \hline \Omega & J_2 & \dots & \Omega \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \Omega & \Omega & \dots & J_p \end{array}\right),$$

where $J_1, J_2, \dots, J_p$ are Jordan blocks.

We will assume without proof the validity of the following theorem:

Theorem 4.8 *Any triangularizable matrix is similar to a Jordan matrix.*

In the remainder of this section our goal is to describe how we can obtain the Jordan form of any given triangularizable matrix.

4.3 Obtaining the Jordan form.

Let us see the steps to calculate the Jordan form of a matrix $T \in \mathcal{M}_{n\times n}(\mathbb{K})$.

1. Compute the characteristic polynomial $p_t(\lambda) = |T_B - \lambda I|$.

2. Find the roots of the characteristic polynomial, i. e. the eigenvalues of t. We will obtain values $\lambda_1, \ldots, \lambda_k$ with algebraic multiplicities $m_{\lambda_1}, \ldots, m_{\lambda_k}$.

   (a) If $m_{\lambda_1} + \ldots + m_{\lambda_k} < n$ the matrix is **not triangularizable**.

   (b) If $m_{\lambda_1} + \ldots + m_{\lambda_k} = n$, the matrix is triangularizable. Let us see how to obtain its Jordan form. We will obtain **the Jordan block associated to each eigenvalue independently**. So, given an eigenvalue λ_i

   i. We compute the subspaces

$$S_{\lambda_i,p} = ker(T - \lambda I)^p$$

   for successive values of p. Note that with this notation $S_{\lambda_i,1} = S_{\lambda_i}$. We will stop at a p for which

$$dim(S_{\lambda_i,p+1}) = dim(S_{\lambda_i,p}).$$

   ii. The geometric multiplicity $d(\lambda_i)$ is equal to the **number of Jordan blocks relative to the eigenvalue λ_i:**

$$\boxed{\text{geometric multiplicity } d(\lambda_i) = \text{number of Jordan blocks}}$$

   iii. Consider the following values:

$$\begin{array}{ccl} dim(S_{\lambda_i,1}) & \longrightarrow & f_1 = d(\lambda_i) \\ & \longrightarrow & f_2 = dim(S_{\lambda_i,2}) - dim(S_{\lambda_i,1}) \\ dim(S_{\lambda_i,2}) & & \\ & \longrightarrow & f_3 = dim(S_{\lambda_i,3}) - dim(S_{\lambda_i,2}) \\ dim(S_{\lambda_i,3}) & & \\ \vdots & & \vdots \\ dim(S_{\lambda_i,p-1}) & & \\ & \longrightarrow & f_p = dim(S_{\lambda_i,p}) - dim(S_{\lambda_i,p-1}) \\ dim(S_{\lambda_i,p}) & & \end{array}$$

   It can be shown that always $f_k \geq f_{k+1}$.

   iv. We form a diagram with p columns, in such a way that the k-th column is formed by f_p elements:

$$\begin{array}{cccccccccc} n_1 & \longrightarrow & * & * & \ldots & * & \ldots & * & \ldots & * \\ n_2 & \longrightarrow & * & * & \ldots & * & \ldots & * & & \\ & & & \vdots & & & & & & \\ n_{f_1} & \longrightarrow & * & * & \ldots & * & & & & \\ & & \uparrow & \uparrow & & & & & & \uparrow \\ & & f_1 & f_2 & & & & & & f_p \end{array}$$

   where $d = f_1 = d(\lambda_i)$. Now **the number of elements in each row** of the diagram tells us **the dimension of each of the Jordan blocks relative to the eigenvalue λ_i**:

$$J(\lambda_i) = \left(\begin{array}{c|c|c|c} J_{n_1 \times n_1} & \Omega & \ldots & \Omega \\ \hline \Omega & J_{n_2 \times n_2} & \ldots & \Omega \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \Omega & \Omega & \ldots & J_{n_d \times n_d} \end{array}\right)$$

   We note that it is convenient to arrange the boxes starting with the largest and ending with the one with the smallest dimension.

   v. To obtain the basis relative to which this Jordan matrix is obtained, we look for a set vectors whose structure mirrors the one described above:

$$\begin{array}{cccccccccc} & & S_{\lambda_i,1} & S_{\lambda_i,2} & \ldots & S_{\lambda_i,n_{f_1}} & \ldots & S_{\lambda_i,n_2} & \ldots & S_{\lambda_i,n_1} \\ n_1 & \longrightarrow & \bar{x}_{1,1} & \bar{x}_{1,2} & \ldots & \bar{x}_{1,n_{f_1}} & \ldots & \bar{x}_{1,n_2} & \ldots & \boxed{\bar{x}_{1,n_1}} \\ n_2 & \longrightarrow & \bar{x}_{2,1} & \bar{x}_{2,2} & \ldots & \bar{x}_{2,n_{f_1}} & \ldots & \boxed{\bar{x}_{2,n_2}} & & \\ & & & \vdots & & & & & & \\ n_{f_1} & \longrightarrow & \bar{x}_{f_1,1} & \bar{x}_{f_1,2} & \ldots & \boxed{\bar{x}_{f_1,n_{f_1}}} & & & & \\ & & \uparrow & \uparrow & & & & & & \uparrow \\ & & f_1 & f_2 & & & & & & f_p \end{array}$$

   where

$$\boxed{\bar{x}_{j,k} = (T - \lambda_i I)\bar{x}_{j,k+1}}$$

   Therefore, to construct these vectors, it is enough to choose the vector at **the right end** of each row in such a way that the resulting **first** vectors of all rows turn out to be linearly independent:

$$\{\bar{x}_{1,1}, \bar{x}_{2,1}, \ldots, \bar{x}_{f_1,1}\}.$$

   The required basis will be formed by all these vectors **ordered from left to right and from top to bottom**:

$$\{\bar{x}_{1,1}, \bar{x}_{1,2}, \ldots, \bar{x}_{1,n_1}, \bar{x}_{2,1}, \bar{x}_{2,2}, \ldots, \bar{x}_{2,n_2}, \ldots, \bar{x}_{f_1,1}, \bar{x}_{f_1,2}, \ldots, \bar{x}_{f_1,n_{f_1}}\}$$

   vi. Finally, the Jordan form is formed by joining the Jordan blocks corresponding to each eigenvalue:

$$J = \left(\begin{array}{c|c|c|c} J(\lambda_1) & \Omega & \ldots & \Omega \\ \hline \Omega & J(\lambda_2) & \ldots & \Omega \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \Omega & \Omega & \ldots & J(\lambda_p) \end{array}\right),$$

   and the basis B' in which it is expressed will be formed by joining all the bases we have obtained for each eigenvalue, **preserving the order in which each vector has been placed**. Thus:

$$J = (M_{BB'})^{-1} T M_{BB'}$$

   where B is the initial basis.

4.4 Application to obtain powers of matrices.

Suppose A is a triangularizable matrix and we want to obtain A^k. The procedure is analogous to the one we have seen for diagonalizable matrices.

Since A is triangularizable, there is a regular matrix P such that

$$J = P^{-1}AP \quad \text{where} \quad J \text{ is a Jordan matrix.}$$

Then $A = P^{-1}JP$ and

$$\begin{aligned} A^k &= \underbrace{(PJP^{-1}) \cdot (PJP^{-1}) \cdot \ldots \cdot (PJP^{-1})}_{k \text{ times}} = \\ &= P\,\underbrace{(P^{-1}PJ) \cdot (P^{-1}PJ) \ldots \cdot (P^{-1}P \cdot J)}_{k \text{ times}}\,P^{-1} = PJ^kP^{-1}. \end{aligned}$$

That is,

$$\boxed{A^k = PJ^kP^{-1}}$$

One way to calculate J^k is the following: we decompose J as the sum of two matrices, one diagonal D and another upper triangular T with zeros on the diagonal:

$$J = D + T \quad \Rightarrow \quad J^k = (D+T)^k.$$

Although **in general Newton's binomial formula is not true for matrices, it is true in this case (since D and T commute)**. So:

$$(D+T)^k = \sum_{i=0}^{k} \binom{k}{i} D^i T^{k-i}.$$

The computation of the powers of D and T is very simple; in fact, from a sufficiently large exponent onwards the matrix T^{k-i} will vanish.

Glosary