



E

Nome do Candidato: Erwin Göpp Parades

• Theorem of Halm-Banasch

We start with some definitions. Let A be a set and \leq a partial order.

- 1) We say that, for $Q \subseteq A$, $a \in A$ is an upper bound for Q if $q \leq a$ for all $q \in Q$.
- 2) We say that A is inductive if for each $Q \subseteq A$ totally ordered set, Q has an upper bound.
- 3) We say that $\bar{a} \in A$ is maximal if, whenever there exists $b \in A$ s.t. $\bar{a} \leq b$, then $\bar{a} = b$.

The principal tool we use here is the following

- Zorn Lemma: If A is inductive, there exists a maximal element for A .

This is a deep result in mathematics and it is known that it is equivalent to the relevant axiom of choice, so we skip its proof. Now we are in position to state H-B theorem

- H-B theorem: Let E be a normed space, $E \neq \emptyset$, and $E_0 \subseteq E$ linear space. Let $g: E_0 \rightarrow \mathbb{R}$ be a linear functional. Assume there exists $p: E \rightarrow \mathbb{R}$ satisfying $g(x) \leq p(x) \forall x \in E_0$ and p such that

- 1) $p(\lambda x) = \lambda p(x) \forall \lambda > 0, x \in E$.
- 2) $p(x+y) \leq p(x) + p(y), \forall x, y \in E$.

Then, $\exists l: E \rightarrow \mathbb{R}$

linear, with $l|_{E_0} = g$ and $l \leq p$ in E .

FAVOR NÃO ESCREVER NO VERSO!
NÃO ULTRAPASSAR A LINHA DA MARGEM!



CONCURSO PARA PROFESSOR ADJUNTO - DM
Área: Matemática Pura

Proof: Consider the set $A = \left\{ l : D(l) \subseteq E \rightarrow \mathbb{R} \mid \begin{array}{l} l \text{ is linear, } l \leq p \text{ in } D(l) \\ E_0 \subseteq D(l) \\ l|_{E_0} = g \end{array} \right\}$

which is nonempty ~~since~~ since $g \in A$. On A , we consider the partial order $l_1 \leq l_2$ if $\left. \begin{array}{l} D(l_1) \subseteq D(l_2) \\ l_2|_{D(l_1)} = l_1 \end{array} \right\} (*)$

It is possible to see that A is an inductive set with this order. In fact, if $Q \subseteq A$ is totally ordered, we consider the functional $L : \bigcup_{l \in Q} D(l) \subseteq E \rightarrow \mathbb{R}$

$x \mapsto l(x)$ if $x \in D(l)$
which is well-defined because of $(*)$, is linear and $L \leq p$ in $D(L)$. Of course $L(x) = g(x)$ if $x \in E_0$. This is an upper bound for Q and A is inductive.

Then, using Zorn's Lemma, we conclude there exists a maximal element $\bar{l} : D(\bar{l}) \subseteq E \rightarrow \mathbb{R}$ in A . To conclude the proof, it remains to prove that $D(\bar{l}) = E$. Assume this does not hold. Then, there exists $x_0 \in E \setminus D(\bar{l})$ and we consider the functional $\bar{L} : \overbrace{D(\bar{l})}^{D(\bar{l})} + \langle \{x_0\} \rangle \rightarrow \mathbb{R}$

$x + tx_0 \mapsto \bar{l}(x) + t \cdot \alpha$
for some $\alpha \in \mathbb{R}$ to be defined later. It is linear, extends g and we will prove it satisfies $\bar{L} \leq p$ in $D(\bar{L})$.



Now we state a number of corollaries associated to this theorem. From now on, E always denotes a normed linear space.

- Cor 1: Let $l: E_0 \rightarrow \mathbb{R}$ ^{continuous} linear functional E_0 linear subspace of E . Then, there exists a continuous linear functional $\bar{l}: E \rightarrow \mathbb{R}$ extending l .

Proof: It suffices to consider the function $p(x) = \|l_0\| \cdot \|x\|$ in H-B theorem. Notice that $\|\bar{l}\| \leq \|l_0\|$. ■

- Cor 2: Let $\{\alpha_i\}_{i=1}^n \subseteq \mathbb{R}$, $\{x_i\}_{i=1}^n \subseteq E$ linearly independent. Then, there exists a continuous linear functional $l: E \rightarrow \mathbb{R}$ such that $l(x_i) = \alpha_i$, $\forall i = 1, \dots, n$.

Proof: It suffices to consider the linear space $V = \langle \{x_i\}_{i=1}^n \rangle$ and the linear functional

$$l_0: V \rightarrow \mathbb{R} \\ \sum_{i=1}^n a_i x_i \mapsto \sum_{i=1}^n a_i \alpha_i$$

which is clearly linear, and since V is finite dimensional it is continuous.

Then, we apply Cor 1 to conclude the existence of l .

- Cor 3 [Existence of Support Functionals] Let $x_0 \neq 0$, $x_0 \in E$ and then, there exists $x_0^*: E \rightarrow \mathbb{R}$ linear, continuous, such that $\|x_0^*\|_* = 1$ and $\langle x_0^*, x_0 \rangle = \|x_0\|$.



CONCURSO PARA PROFESSOR ADJUNTO - DM
Área: Matemática Pura

OBS. The space of continuous linear functionals $l: E \rightarrow \mathbb{R}$ is the dual space of E , and it is denoted by E^* .

If $l \in E^*$, we write $\|l\|_* = \sup_{\substack{\|x\| \leq 1 \\ x \in E}} |l(x)|$ the usual norm

we also write $\langle l, x \rangle = \langle l, x \rangle_{E^*, E} = l(x)$ the duality pair.

Proof: [cor 3] we consider the 1-dimensional space $\langle \{x_0\} \rangle$ and the linear map $l_0: \langle \{x_0\} \rangle \rightarrow \mathbb{R}$

which is continuous. Notice that $\|l_0\| = 1$. $t x_0 \mapsto t \cdot \|x_0\|$

It is clear that $l_0(x) \leq \|x\|$ for $x \in \langle \{x_0\} \rangle$ and from here, taking $p(x) = \|x\|$ in H-B Theorem, we conclude that there exists a continuous (by cor 1 in fact) linear functional $l: E \rightarrow \mathbb{R}$ extending l_0 . We have $\|l\|_* = 1$ (since $\|l\|_* \leq 1$ and it is attained at $\langle \{x_0\} \rangle$) and by definition of l_0 we have $\langle l, x_0 \rangle = \|x_0\|$.

~~Now~~ Recalling that for $l \in E^*$ we have

$$\|l\|_* = \sup_{\|x\|=1} |\langle l, x \rangle|$$

we have the following interesting corollary concerning a characterization of the norm of $x \in E$ in terms of the linear functionals E^* .

- Cor 4: For all $x \in E$, we have $\|x\| = \sup_{\substack{x^* \in E^* \\ \|x^*\|_* \leq 1}} |\langle x^*, x \rangle|$

Proof: For all $\|x^*\|_* \leq 1$, we have

$$|\langle x^*, x \rangle| \leq \|x^*\|_* \|x\| \leq \|x\| \text{ and the inequality } \geq$$

holds. By taking the support functional (cor 3) we have the equality. ■

4



CONCURSO PARA PROFESSOR ADJUNTO - DM
Área: Matemática Pura

Cor 5: Let M be a closed linear subspace of E , $x_0 \notin M$.

Then, there exists a linear functional $l \in E^*$ s.t.
 $l(x) = 0, \forall x \in M, \quad l(x_0) = 1$ and $\|l\|_* = \frac{1}{\text{dist}(x_0, M)}$.

Proof: We consider the space $V = \langle \{x_0\} \rangle$ and the functional
 $l_0: V \rightarrow \mathbb{R}$ given by $l_0(tx_0) = t$.

Now, consider $p: E \rightarrow \mathbb{R}$ defined as $p(x) = \frac{\text{dist}(x, M)}{\text{dist}(x_0, M)}$,
where $\text{dist}(x, M) = \inf_{m \in M} \|x - m\|$ is the usual distance function.

First, we have p is positively homogeneous since: if $\lambda > 0$:

$$p(\lambda x) = \frac{\text{dist}(\lambda x, M)}{\text{dist}(x_0, M)} \quad \text{but} \quad \text{dist}(\lambda x, M) = \inf_{m \in M} \|\lambda x - m\| = \lambda \inf_{m \in M} \|x - \frac{m}{\lambda}\|$$

since M is linear space \searrow
 $= \lambda \inf_{m \in M} \|x - m\| = \lambda \text{dist}(x, M)$

and then $p(\lambda x) = \lambda p(x)$. On the other hand, for $x, y \in E$

we have: ~~proposition~~

$$\begin{aligned} \text{dist}(x+y, M) &= \inf_{m \in M} \|x+y-m\| \\ &= \inf_{m_1, m_2 \in M} \|x-m_1 + y-m_2\| \quad \text{since } M \text{ is a linear space} \\ &\leq \inf_{m_1, m_2} \|x-m_1\| + \inf_{m_1, m_2} \|y-m_2\| \\ &= \text{dist}(x, M) + \text{dist}(y, M) \end{aligned}$$

and p is sublinear. We also have $l_0(tx_0) \leq p(tx_0)$ since
for $t \leq 0$ we have $l_0(tx_0) \leq 0 \leq p(tx_0)$, meanwhile for
 $t > 0$ we have $l_0(tx_0) = t = t \frac{\text{dist}(x_0, M)}{\text{dist}(x_0, M)} = \frac{\text{dist}(tx_0, M)}{\text{dist}(x_0, M)} = p(tx_0)$

Then, we use H-B theorem to conclude the existence of
 $l: E \rightarrow \mathbb{R}$ continuous and linear such that

$$l(x_0) = l_0(x_0) = 1$$

$$\text{and } |l(x)| \leq p(x) = 0 \quad \text{for } x \in M.$$

5



CONCURSO PARA PROFESSOR ADJUNTO - DM
Área: Matemática Pura

Finally, we have that $\|l\|_F \leq \frac{1}{\text{dist}(x_0, M)}$ since $\|l_0\|_* \leq \frac{1}{\text{dist}(x_0, M)}$

because $\frac{|\langle l, x_0 \rangle|}{\|x_0\|} = \frac{1}{\|x_0\|} \leq \frac{1}{\text{dist}(x_0, M)}$, and by Cor 1 we have

the extension has a norm less or equal than the original functional, we have the equality by taking a sequence $(x_n) \subseteq M$ such that $\|x_n - x_0\| \rightarrow \text{dist}(x_0, M)$ and

$$\frac{|\langle l, x_n - x_0 \rangle|}{\|x_n - x_0\|} = \frac{|\langle l, x_n \rangle - \langle l, x_0 \rangle|}{\|x_n - x_0\|} = \frac{1}{\|x_n - x_0\|} \xrightarrow{n \rightarrow \infty} \frac{1}{\text{dist}(x_0, M)} \quad \square$$

We also have, as a corollary, a way to characterize dense subspaces for E .

Cor 6: Let E_0 be a ^{linear} subspace of E such that it satisfies the following property:

(*) $\left[\nexists x^* \in E^* \text{ such that } \langle x^*, x \rangle = 0 \ \forall x \in E_0, \text{ then } x^* = 0 \right]$
Then, E_0 is dense in E . (i.e. $\overline{E_0} = E$)

Proof: Assume there exists $x_0 \in E$, $x_0 \notin \overline{E_0}$. Then, we use Cor 5 to construct a linear functional $l \in E^*$ with $l|_{\overline{E_0}} = 0$ and $\langle l, x_0 \rangle = 1$. However, this functional contradicts the assumption (*). ($\Rightarrow \Leftarrow$) \square

We have the following characterization of ^{continuous} linear functionals:

Cor 7: Let $l: E \rightarrow \mathbb{R}$ linear. l is continuous iff $l^{-1}(\{0\})$ is closed.

Proof: The "only if" part is clear. Now we assume that $l^{-1}(\{0\}) = M$ is closed.



CONCURSO PARA PROFESSOR ADJUNTO - DM
Área: Matemática Pura

~~Then~~, Assume l is not identically zero, otherwise the result holds. Then, there exists $x_0 \notin M$, and by Cor 5 we have the existence of a functional $\bar{l} \in E^*$ such that $\bar{l}(x) = 0 \quad \forall x \in M$. Now, for each $x \in E$ we can write

$$\begin{aligned} x &= x - \frac{l(x)}{l(x_0)} x_0 + \frac{l(x)}{l(x_0)} x_0 \\ &= a + b \end{aligned}$$

and notice that $a \in \ker(l) = M$. Thus, $\bar{l}(a) = 0$ and we can write

$\bar{l}(x) = 0 + \frac{l(x)}{l(x_0)} \bar{l}(x_0)$, that is l is a \bar{l} up to a scalar factor $\frac{\bar{l}(x_0)}{l(x_0)} \neq 0$, hence continuous. \square

• Separation Theorems

Probably the most important application of H-B Theorem concerns the separation of convex sets. We recall that $C \subseteq E$ is convex if for all $x, y \in C$ and $\lambda \in [0, 1]$, we have $\lambda x + (1-\lambda)y \in C$.

An hyperplane in E is a set with the form

$$H = \{x \in E : \langle l, x - x_0 \rangle = 0\} \quad \text{for some } l \in E^* \text{ and } x_0 \in E.$$

or, equivalently $= \{x \in E : \langle l, x \rangle = \alpha\}$ for some $l \in E^*$ and $\alpha \in \mathbb{R}$.

We say that two sets $A, B \subseteq E$ (nonempty) are separated by the hyperplane H if $\langle l, x \rangle \leq \alpha \leq \langle l, y \rangle$ for all $x \in A, y \in B$.



CONCURSO PARA PROFESSOR ADJUNTO - DM
Área: Matemática Pura

The first geometric form of H-B is the following

Thm 1 [H-B, first geometric form] Let A, B be nonempty convex sets such that $A \cap B = \emptyset$, and A open. Then, A and B can be separated by a hyperplane.

To prove this, we require some preliminaries: let C be a convex, open, nonempty subset of E such that $0 \in C$. We define the Minkowski functional associated to C as the function

$$p(x) = \inf \{ \alpha > 0 \mid \alpha^{-1}x \in C \}.$$

Concerning this function, we can say that:

- 1) p is positively homogeneous ($p(\lambda x) = \lambda p(x)$ if $\lambda > 0$)
- 2) p is sublinear ($p(x+y) \leq p(x) + p(y)$)
- 3) $\exists M > 0$ such that $0 \leq p(x) \leq M \|x\|$ for all $x \in E$.
- 4) $C = \{x \in E \mid p(x) < 1\}$.

1) is clear from the definition. Now, for 3), we see that there exists $r > 0$ such that $B(0, r) \subseteq C$. Then, given $x \in E$, we have that $\frac{r}{2\|x\|} \cdot x \in B(0, r) \subseteq C$ and from here we see that $0 \leq p(x) \leq \frac{2}{r} \|x\|$ and 3) follows taking $M = \frac{2}{r}$.

For 4) if $p(x) < 1$ we have that for some $\varepsilon > 0$ small enough $(1-\varepsilon)^{-1}x \in C$. But x belongs to the line joining 0 and $(1-\varepsilon)^{-1}x$, from which x can be written as a convex combination of these points belonging to C , hence $x \in C$.



CONCURSO PARA PROFESSOR ADJUNTO - DM
Área: Matemática Pura

on the other hand, if $x \in C$ and C is open, we have $B(x, r) \subseteq C$ for some $r > 0$ small. Then, $(1+r/2)x \in C$ from which $p(x) \leq (1+r/2)^{-1} < 1$. This concludes 4) for 3)...

We require the following separation property with points:

- lemma: Let $C \subseteq E$ convex, nonempty open set containing the origin and let $x_0 \notin C$. Then, there exists a hyperplane separating C and $\{x_0\}$.

Proof: Consider the ~~linear~~ linear space $V = \langle \{x_0\} \rangle$ and the linear functional $l_0: V \rightarrow \mathbb{R}$. Notice that

$l_0 \leq p$, where p is the Minkowski functional associated to C . if $t \leq 0$, it is clear from the definition of l_0 and p that $l(t x_0) = t \leq 0 \leq p(t x_0)$. If $t > 0$, then $l(t x_0) = t \leq t p(x_0)$

since $x_0 \notin C$. Thus, by H-B Theorem, we can extend l_0 as a continuous linear functional in E . Now, if $x \in C$, we have: $l(x) \leq p(x) < 1 = l(x_0)$, and l separates C and x_0 .

Obs. A translation argument allows to drop the hypothesis that $0 \in C$ in Lemma 1.



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
Centro de Ciências Matemática e da Natureza
Instituto de Matemática

CONCURSO PARA PROFESSOR ADJUNTO - DM
Área: Matemática Pura

Now we continue with the proof of Thm 1: Let us write the set $C = A - B$ as $C = \{x - y : x \in A, y \in B\}$

$= \bigcup_{y \in B} A - y$. Since A is open, C is open. Since A, B are convex, C too. In addition, since $A \cap B = \emptyset$, the origin does not belong to C . Using Lemma 1 (in its translated version), there exist $l: E \rightarrow \mathbb{R}$ linear and continuous, and $\alpha \in \mathbb{R}$ s.t.

$$l(x - y) \leq \alpha \leq l(0) = 0, \text{ for all } x - y \in C$$

By linearity, we have

$$l(x) \leq l(y) \text{ for all } x \in A, y \in B$$

from which we get the separation of A and B .

We say that a hyperplane $H = \{l = \alpha\}$ strictly separates two nonempty sets A and B if there exist $\alpha < \beta$ such that $l(x) \leq \alpha < \beta \leq l(y)$, $\forall x \in A, \forall y \in B$.

We have the following stronger version of the separation theorem:

- Thm [H-B, second geometric form] Let $A, B \subseteq E$ convex, nonempty sets such that A is closed, B is compact. Then, there exists a hyperplane H that strictly separates A and B .

Proof: The set $C = A - B$ is convex and the origin does not belong to C . By the compactness of B , we have that C is closed: let $x_n - y_n \in C$ ($x_n \in A, y_n \in B$) such that

$$x_n - y_n \xrightarrow{n \rightarrow \infty} \bar{z}. \text{ By}$$

compactness, we have $y_n \xrightarrow{n \rightarrow \infty} \bar{y} \in B$ (up to subseq.)

FAVOR NÃO ESCREVER NO VERSO!
NÃO ULTRAPASSAR A LINHA DA MARGEM!



CONCURSO PARA PROFESSOR ADJUNTO - DM
Área: Matemática Pura

Nome do Candidato:

Then, $x_n = x_n - y_n + y_n$ is convergent, and since A is closed, $x_n \rightarrow \bar{x} \in A$. Thus $\bar{z} = \bar{x} - \bar{y} \in C$, and C is closed. This allows us to conclude that C^c is open and therefore there exists $r > 0$ such that $B(0, r) \cap C = \emptyset$.

We use the separation theorem 1 to conclude that there exists $l \in E^*$ such that

$$\alpha \in \mathbb{R} \quad l(x - y) \leq \alpha \leq l(z) \quad \text{for all } x \in A, y \in B \\ z \in B(0, r)$$

$$\Rightarrow l(x) - l(y) \leq \alpha \leq r l(w) \quad \text{for all } x \in A, y \in B$$

Now, we can take a sequence $w_n \in B(0, 1)$ $w \in B(0, 1)$ such that $l(w_n) \rightarrow -\|l\|_* < 0$, from which

$$l(x) - l(y) \leq -r \|l\|_* \quad \text{for all } x \in A, y \in B$$

and from here we conclude that A and B are strictly separated.



• Jordan Form.

Let \mathbb{K} be a field (usually $\mathbb{K} = \mathbb{R}$ or \mathbb{C}), and let A be an $n \times n$ matrix with entries in \mathbb{K} . (Denote $A \in M_{n \times n}(\mathbb{K})$)

An eigenvalue of A is a scalar $\lambda \in \mathbb{K}$ such that there exists $x \in \mathbb{K}^n$ not zero such that

$$Ax = \lambda x. \quad (1)$$

In that case, we say that x is an eigenvector of A associated to λ , and it is direct to see that the sets of eigenvectors associated to λ is a linear subspace of \mathbb{K}^n .

The way to find the eigenvalues of A is a consequence of the Fundamental Theorem of Algebra. In fact, nontrivial solutions x for (1) exists if

$$\det(A - \lambda I) = 0, \text{ where } I \text{ is the identity matrix.}$$

It is easy to see that the expression $\lambda \mapsto p(\lambda) = \det(A - \lambda I)$ is a polynomial of degree n , and therefore we have the existence of n roots in \mathbb{C} . The polynomial $p(\lambda)$ is called the characteristic polynomial associated to A .

Thus, $p(\lambda)$ can be written as:

$$p(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_r)^{m_r}$$

for some $r \in \mathbb{N}$, and such that $\{\lambda_i\}_{i=1}^r$ are the roots of p , and $m_1 + \dots + m_r = n$. (12)

In this way, we have that the different λ_i 's are the eigenvalues of A



CONCURSO PARA PROFESSOR ADJUNTO - DM
Área: Matemática Pura

and for each $i=1, \dots, r$, the number m_i is called the algebraic multiplicity of λ_i . We denote it as $m_a(\lambda_i)$.

If V_i is the linear space of ^{vector} eigenvalues associated to λ_i , then we call the geometric multiplicity of λ_i the dimension of V_i subspace. We denote it as $m_g(\lambda_i)$.

It is possible to see that eigen vectors associated to different eigenvalues are linearly independent.

We also have that for each eigenvalue λ_i ,

$$m_g(\lambda_i) \leq m_a(\lambda_i).$$

In the case of equality, we will have a ~~full~~ set of linearly independent eigenvectors generating the space (hence, a basis). In ~~an~~ that case, we say that A is a diagonalizable matrix. This name is not a coincidence, since if we arrange the set of eigenvectors in a matrix $P = [v_1 | \dots | v_n]$ then it is possible to implement a change of basis in \mathbb{R}^n to conclude that

$$A = P \cdot D \cdot P^{-1}, \quad (2)$$

where $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \lambda_n \end{bmatrix}$ is a diagonal matrix whose entries are the eigenvalues of A . The "factorization" (2) has important consequences in the matrix analysis. For instance, It is easy to compute the determinant of A ^{eg} using the rule of the determinants: $\det(A) = \det(PDP^{-1}) = \det(P) \cdot \det(D) \cdot \det(P)^{-1}$

$= \det(D) = \prod_{i=1}^n \lambda_i$, as well as the trace: $\text{Tr}(A) = \text{Tr}(PDP^{-1})$
 $= \text{Tr}(DP^{-1}P) = \text{Tr}(D)$
 $= \sum_{i=1}^n \lambda_i$.

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
 Centro de Ciências Matemática e da Natureza
 Instituto de Matemática

CONCURSO PARA PROFESSOR ADJUNTO - DM
 Área: Matemática Pura

Moreover, given an analytic function $f: \mathbb{C} \rightarrow \mathbb{C}$, it is possible to define the matrix $f(A)$

as
$$f(A) = \sum_{i=0}^{+\infty} \frac{f^{(i)}(0)}{i!} A^i = \sum_{i=0}^{+\infty} \frac{f^{(i)}(0)}{i!} P D^i P^{-1}$$

$$= P \left(\sum_{i=0}^{+\infty} \frac{f^{(i)}(0)}{i!} D^i \right) P^{-1} = P f(D) P^{-1}$$

where $f(D) = \text{diag}(f(\lambda_1), \dots, f(\lambda_n))$. Of particular interest in the analysis of linear systems of ODE is the case of $f(z) = e^z$, giving rise to the exponential matrix.

Nevertheless, it is possible to have $m_g(\lambda_i) < m_a(\lambda_i)$ and therefore the decomposition (2) is no longer available, as from which several of the mentioned interesting properties are not at hand anymore.

A way to overcome this difficulty goes along the so-called Jordan decomposition of a non diagonalizable matrix. The idea is to complete a basis of the space \mathbb{K}^n in a "wise way". We give some informal treatments of this procedure, starting with the following fundamental result:

- Thm [Schur Decomposition Thm] Let $A \in M_{n \times n}(\mathbb{K})$. Then, there is an unitary matrix U (i.e. $UU^* = I$) and a upper triangular matrix T with the eigenvalues of A in its diagonal, such that $A = U T U^*$.

FAVOR NÃO ESCREVER NO VERSO!
 NÃO ULTRAPASSAR A LINHA DA MARGEM!



CONCURSO PARA PROFESSOR ADJUNTO - DM
Área: Matemática Pura

This result is a consequence of the well-known Gauss algorithm. Then, how "good" can T be? If this cannot be diagonal, let say that it is of the form

$$T = \begin{bmatrix} \lambda_1 & 1 & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} \quad \text{that is, the only non zero elements in the upper diagonal are 1's,}$$

and they are located in the first upper diagonal. For simplicity, let us assume that $A \in M_{3 \times 3}(\mathbb{K})$ and it has only one eigenvalue (with ^{alg} multiplicity 3). If we can write

$$(4) \quad A = P \cdot \overbrace{\begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}}^J P^{-1} \quad \text{what shall the columns of } P \text{ be?}$$

If we write $P = [v_1 | v_2 | v_3]$ equality (4) means that (after multiplying (4) by P by the right)

$$[Av_1 | Av_2 | Av_3] = [v_1 | v_2 | v_3] \cdot \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}, \quad \text{that is}$$

$$Av_1 = \lambda v_1, \quad Av_2 = v_1 + \lambda v_2, \quad Av_3 = v_2 + \lambda v_3.$$

This tells us that v_1 is an eigenvector of A . On the other hand, $(A - \lambda I)v_2 = v_1$ and applying $A - \lambda I$ again we have that v_2 must satisfy $(A - \lambda I)^2 v_2 = 0$, that is, v_2 is in the kernel of $(A - \lambda I)^2$, and so forth for v_3 .

Decomposition (4) is useful as an alternative for (2) in the diagonalizable case, since the inner matrix J in (4) can be decomposed in the form

$$J = D + N, \quad \text{where } D \text{ is diagonal and } N \text{ is nilpotent}$$



CONCURSO PARA PROFESSOR ADJUNTO - DM
Área: Matemática Pura

that is, $N^K = 0$ for some K .

However, the algorithm to construct the decomposition (4) requires further analysis that we discuss now.

Def [Index of a matrix] Let $A \in M_{n \times n}(K)$. The index of A is the ^{smallest} ~~first~~ integer K such that $\ker(A^K) = \ker(A^{K+1})$.

It is easy to see that $\ker(A^K) \subseteq \ker(A^{K+1})$. By the null-range ~~then~~ theorem, we have that if k is the index of A , then $R(A^k) = R(A^{k+1})$ where $R(A)$ is the range of A . Moreover, $R(A^{k+i}) = R(A^k)$ and $\ker(A^{k+i}) = \ker(A^k)$ for all $i \geq 0$.

Let us start the analysis in the case of nilpotent matrices. It is clear that if L is nilpotent, then the index of L coincides with the order of nilpotency. Then, we are going to consider the following family of subspaces of K^n : For $i = 0, \dots, K$, we write

$$\mu_i = \ker(L) \cap R(L^i).$$

~~It is easy to~~ notice that $\mu_0 = \ker(L)$, (since $L^0 = I$) meanwhile $\mu_K = \{0\}$ (since $L^K = 0$).



CONCURSO PARA PROFESSOR ADJUNTO - DM
Área: Matemática Pura

~~Let S_i be a basis of M_i .~~ We also have $M_{i+1} \subseteq M_i$.

Let S_{k-1} be a basis of M_{k-1} , S_{k-2} such that $S_{k-1} \cup S_{k-2}$ is a basis of M_{k-2} and so forth.

Then, $S_0 \cup S_1 \cup \dots \cup S_{k-1} \cup \beta$ is a basis of $\ker(L)$. Let $b \in \beta$, and then there exists a unique i such that $b \in S_i \subseteq M_i = \ker(L) \cap R(L^i)$. Then, by definition, there exists $x \in \mathbb{K}^n$ such that $L^i x = b$. We have that $L^{i+1} x = Lb = 0$ since $b \in \ker(L)$, and the set

$J_b := \{x, Lx, L^2x, \dots, L^{i-1}x, \underbrace{L^i x}_{=b}\}$ is linearly independent

This is what we call a Jordan chain associated to $b \in \beta$. Notice that the length of J_b is $i+1$ for $b \in S_i$.

It is possible to prove that: the union of the Jordan chains associated to $b \in \beta$ form a basis of \mathbb{K}^n , and allows us to decompose L in a Jordan form similar to (4).



CONCURSO PARA PROFESSOR ADJUNTO - DM
Área: Matemática Pura

In fact, if $b \in \beta$ and $J_b = \{b, L^{i-1}x, \dots, L^i x, x\}$ is its Jordan chain, we see that:

$$L \cdot \underbrace{\begin{bmatrix} b & | & L^{i-1}x & | & L^{i-2}x & | & \dots & | & Lx & | & x \end{bmatrix}}_{P_b} \quad \begin{matrix} n \times n \\ n \times (i+1) \end{matrix}$$

$$= \begin{bmatrix} Lb & | & L^i x & | & L^{i-1}x & | & \dots & | & L^2 x & | & Lx \end{bmatrix}_{n \times (i+1)}$$

$$= \begin{bmatrix} 0 & | & b & | & L^{i-1}x & | & \dots & | & Lx \end{bmatrix}_{n \times (i+1)} = \begin{bmatrix} \text{diagonal blocks} \end{bmatrix}_{n \times n}$$

$$= P_b \cdot \begin{bmatrix} 0 & 1 & 0 \\ & \ddots & \ddots \\ 0 & & 0 \end{bmatrix}_{(i+1) \times (i+1)}$$

, that is $L \cdot P_b = P_b N$

Since 0 is the only eigenvalue of L .

Performing this analysis block by block depending on the chains created with each $b \in \beta$, we conclude that for L nilpotent matrix, we have the existence of a matrix $P = [P_{b_1} | \dots | P_{b_r}]$ nonsingular (r is the number of l.i. eigenvectors of L) such that

$$L = P \cdot N \cdot P^{-1}, \text{ where } N \quad (18)$$

is a block matrix with the form $N = \begin{bmatrix} N_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & N_r \end{bmatrix}$

with each $N_i = \begin{bmatrix} 0 & 1 & 0 \\ & \ddots & \ddots \\ 0 & & 0 \end{bmatrix}$. The size of each N_i depends on the



CONCURSO PARA PROFESSOR ADJUNTO - DM
Área: Matemática Pura

S_i where $b \in S_i$ is considered, and then it is of size $(i+1) \times (i+1)$.

Since for a general matrix A (not necessarily nilpotent) the eigenspaces are invariant ~~through~~ with the action of A , we can proceed by considering the nilpotent matrix $A - \lambda_i I$ over the space $\text{Ker}(A - \lambda_i I)$ for each eigenvalue λ_i of A , from which we can decompose A in its Jordan canonical form, given by:

$$A = P \cdot J \cdot P^{-1}$$

where $J = \begin{bmatrix} N_1 & & & \\ & N_2 & & \\ & & \ddots & \\ \emptyset & & & N_r \end{bmatrix}$

where each N_i is associated to the Jordan form given by the action of

$A - \lambda_i I : \text{Ker}(A - \lambda_i I) \rightarrow \text{Ker}(A - \lambda_i I)$ as in the nilpotent case.