Dimensões de Espaços e Álgebras de Matrizes

Um Levantamento

Pedro Jorge Santos Freitas

Faculdade de Ciências da Universidade de Lisboa
Departamento de Matemática
1994

Não me invejo de quem tem
    Carros, parelhas e montes,
Só me invejo de quem bebe
    A água em todas as fontes.

Quadra popular portuguesa

Agradecimentos

Quero agradecer aqui ao meu orientador, o Professor José Perdigão Dias da Silva por me ter proposto o tema da dissertação, por toda a orientação científica e por todo o *extraordinário* apoio e amizade que me dedicou ao longo deste tempo. Queria agradecer também a todos aqueles que me deram algumas sugestões e esclarecimentos que possibilitaram a maior simplicidade e clareza do texto. Finalmente, quero agradecer aos meus familiares por todo o apoio, principalmente aos meus Pais e Avó, e a toda a gente amiga do Complexo II pelo excelente ambiente de trabalho, particularmente ao Fernando, pela companhia e interesse, ao Pedro Cristiano, e sua tertúlia, ao João Pedro Boto, compañero de desventuras, à João, sempre presente, e à Catarina, a testemunha perfeita.

Quero ainda agradecer ao Centro de Álgebra da Universidade de Lisboa, pelas magníficas condições de trabalho que me proporcionou; e à JNICT, que me concedeu uma bolsa de mestrado no âmbito do Programa CIÊNCIA.

Pedro Freitas

Índice

Prefácio

Esta dissertação foi redigida com o objectivo de obter o grau de Mestre em Matemática, pela Universidade de Lisboa.

As matrizes constituem um modo eficaz e engenhoso de guardar e apresentar informação. Pelo uso tão diversificado que elas têm, torna-se possível lançar diversos olhares quando se tenta desenvolver resultados que estudem a sua estrutura. Esta tese pretende ser uma recolha de alguns resultados, interessantes *de per si*, mas cujas demonstrações são de algum modo um exemplo de como é possível ter uma determinada abordagem quando se estuda o problema das dimensões de espaços vectoriais de matrizes. Apresentamos uma abordagem diferente em cada capítulo, e dois resultados que têm demonstrações diferentes em capítulos diferentes.

Pelo facto de vários dos problemas tratados estarem relacionados com comutatividade, apresenta-se na introdução um estudo da equação matricial $AX = XA$. De resto, não se pretende apresentar um estudo exaustivo de nenhum problema em particular, antes mostrar e comparar aproximações a determinados problemas. As lacunas e incorrecções que este trabalho contenha são, obviamente, da minha inteira responsabilidade.

Introdução

Notação

Sendo F um corpo, m, n números naturais, f e g um polinómios com coeficientes num corpo, e x um número real, vamos usar as seguintes convenções notacionais:

- δ_{mn} o símbolo de Kronecker, que representa 1 se $m = n$ e 0 caso contrário,
- $[n]$ o conjunto $\{1, \ldots, n\}$,
- $\mathrm{mdc}(f, g)$ o máximo divisor comum dos polinómios f e g,
- $\lceil x \rceil$ o menor inteiro maior ou igual a x,
- $\lfloor x \rfloor$ o maior inteiro menor ou igual a x,
- $\mathrm{gr}(f)$ o grau de f,
- $M_{m \times n}(F)$ o conjunto das matrizes do tipo $m \times n$ sobre o corpo F,
- $M_n(F)$ o conjunto das matrizes do tipo $n \times n$,
- $GL_n(F)$ o conjunto das matrizes invertíveis do tipo $n \times n$,
- $\mathrm{diag}(x_1, \ldots, x_n)$ a matriz diagonal $A = [a_{ij}]$, com $a_{ii} = x_i$,
- 0_{mn} e 0_n as matrizes nulas dos tipos $m \times n$ e $n \times n$, respectivamente,
- I_n a matriz identidade, do tipo $n \times n$,
- E_{ij} a matriz composta por zeros em todas as entradas, excepto na entrada (i, j), onde tem um 1.

Para matrizes $A, B, C \in M_n(F)$, em que C é uma matriz de zeros e uns, $i_1 \leq \ldots \leq i_r$, $j_1 \leq \ldots \leq j_s \in [n]$ e $(X_k : k \in [t])$, uma família de matrizes com $X_k \in M_{n_k}(F)$, $n_1 + \ldots + n_k = n$, notaremos por

- $\mathrm{diag}(X_1, \ldots, X_t)$ a matriz diagonal por blocos

$$X_1 \oplus \ldots \oplus X_t,$$

- $\mathrm{vec}(A)$ a matriz coluna, com as n^2 entradas da matriz A, dispostas por uma certa ordem, a especificar em cada ocasião,
- $\mathcal{C}(A)$ o *comutador de* A, ou seja, o conjunto das matrizes de $M_n(F)$ que comutam com A,
- $M_n[C](F)$ o subespaço de $M_n(F)$ constituído pelas matrizes $X = [x_{ij}]$ que verificam $x_{ij} = 0$ se $c_{ij} = 0$,
- $A[i_1 \ldots i_r | j_1 \ldots j_s]$ a matriz do tipo $r \times s$ formada pelas linhas $i_1, \ldots, i_r$ e pelas colunas $j_1, \ldots, j_s$ de A, isto é, cuja entrada (p, q) é a entrada (i_p, j_q) de A,
- $F[A]$ o conjunto $\{ f(A) : f \in F[x] \}$,
- A^T a transposta de A,
- $\mathrm{c}(A)$ a característica de A,
- $\mathrm{tr}(A)$ o traço de A,
- $\det(A)$ ou $|A|$ o determinante de A.

Quando falarmos do polinómio característico de A, estamos a considerar o polinómio $\det(\lambda I_n - A)$.

Sejam finalmente R um anel, M um módulo sobre R e $Y \subseteq M$. Sejam também T uma álgebra sobre F, e $X \subseteq T$. Notaremos por

- $\mathrm{alg} <X>$ a F-subálgebra de T gerada por X,
- $<X>$ o F-subespaço vectorial de T gerado por X e
- $<Y>_R$ o R-submódulo de M gerado por Y.

É claro que a notação $<X>$ pode ser usada quando T for apenas um espaço vectorial.

Para A e B conjuntos, usaremos $A \subset B$ para denotar a inclusão estrita de conjuntos, usando $A \subseteq B$ para permitir a igualdade.

Resultados Gerais sobre comutatividade

Dada uma matriz A, com entradas num corpo, é bem conhecido o resultado que afirma que ela é semelhante à soma directa de matrizes companheiras dos seus factores invariantes. Se todos os valores próprios de A pertencerem ao corpo, então A é também semelhante à soma directa de matrizes companheiras dos seus divisores elementares. Notaremos por J_k a matriz do tipo $k \times k$ com a forma

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & & & \ddots & 1 \\ 0 & & \cdots & & 0 \end{bmatrix}$$

com $J_1 = [0]$, e assim, a matriz companheira do divisor elementar $(x - \lambda)^k$ será $\lambda I_k + J_k$.

Vamos agora estudar a equação

$$AX = XA \tag{1}$$

em que A é uma matriz com entradas num corpo F algebricamente fechado. Aqui seguiremos de perto [Ga, pp. 215-223], onde se podem encontrar as demonstações dos resultados aqui apresentados.

É simples de verificar que, se A tiver forma normal de Jordan $\tilde{A}$, soma de matrizes companheiras dos divisores elementares, e se U for tal que $A = U\tilde{A}U^{-1}$ então as soluções da equação (1) terão a forma $X = U\tilde{X}U^{-1}$, com $\tilde{X}$ solução de

$$\tilde{A}\tilde{X} = \tilde{X}\tilde{A}$$

Assim, podemos estudar a equação (1) supondo que A está na forma normal de Jordan, com divisores elementares $(x - \lambda_1)^{p_1}, \ldots, (x - \lambda_t)^{p_t}$. Neste caso, o estudo das n^2 equações escalares leva ao seguinte resultado: particionando X em blocos

$$[X_{\alpha\beta} : \alpha, \beta \in [t]] = \begin{bmatrix} X_{11} & \cdots & X_{1t} \\ \vdots & & \vdots \\ X_{t1} & \cdots & X_{tt} \end{bmatrix},$$

em que $X_{\alpha\beta}$ é do tipo $p_\alpha \times p_\beta$, e sendo T_m a matriz triangular superior do tipo $m \times m$

$$\begin{bmatrix} a_0 & a_1 & \cdots & a_{m-1} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1 \\ 0 & \cdots & 0 & a_0 \end{bmatrix} \tag{2}$$

então, se $\lambda_\alpha = \lambda_\beta$, $X_{\alpha\beta}$ tem uma das seguintes formas:

$$\begin{aligned} X_{\alpha\beta} &= T_{p_\alpha} && \text{se } p_\alpha = p_\beta, \\ X_{\alpha\beta} &= \begin{bmatrix} 0_{p_\alpha\, p_\beta - p_\alpha} & T_{p_\alpha} \end{bmatrix} && \text{se } p_\alpha < p_\beta, \text{ ou} \\ X_{\alpha\beta} &= \begin{bmatrix} 0_{p_\alpha - p_\beta\, p_\alpha} \\ T_{p_\beta} \end{bmatrix} && \text{se } p_\alpha > p_\beta. \end{aligned}$$

Diremos de uma matriz que esteja numa destas três formas que é *triangular superior regular*. Se $\lambda_\alpha \neq \lambda_\beta$, então $X_{\alpha\beta} = 0_{p_\alpha\, p_\beta}$.

É simples de verificar que a matriz (2) também se pode obter como polinómio em J_m, da seguinte forma:

$$T_m = a_0 I + a_1 J_m + a_2 J_m^2 + \ldots + a_{m-1} J_m^{m-1}.$$

Sejam agora $i_1, \ldots, i_t$ os factores invariantes diferentes de 1 da matriz A, com $i_t \mid \ldots \mid i_1$, e $n_1 \geq \ldots \geq n_t > 0$ os seus respectivos graus. Ao fazer a contagem dos parâmetros que aparecem na matriz X, verificamos que estes são em número de

$$N = \sum_{i=1}^{t} (2i-1) n_i,$$

que é a dimensão do comutador de A. Como

$$n = n_1 + \ldots + n_t,$$

temos $N \geq n$, com igualdade se e só se $n_1 = n$, isto é, se o primeiro divisor elementar tiver grau n, ou seja, se a matriz for não derrogatória. Por outro lado, $F[A] \subseteq \mathcal{C}(A)$, e $\dim F[A] = n$ se e só se A for não derrogatória. Assim, obtivemos o seguinte resultado:

Proposição *Para $A \in M_n(F)$, temos que A é não derrogatória sse $\mathcal{C}(A) = F[A]$.*

Finalmente, enunciamos outro resultado que será útil para o que se segue.

Proposição *Se duas matrizes A e B comutam, e se A está na forma*

$$A = \mathrm{diag}(A_1, A_2),$$

com $A_1 \in M_{s_1}(F)$ e $A_2 \in M_{s_2}(F)$, $s_1 + s_2 = n$ e sem valores próprios comuns, então B tem a forma

$$B = \mathrm{diag}(B_1, B_2)$$

com $B_1 \in M_{s_1}(F)$ e $B_2 \in M_{s_2}(F)$.

□ □
□ □
□ □

Capítulo 1

Usando Geometria Algébrica

```
Listen very carefully,
I shall say this only once.
```
(Michelle, em *Allô Allô*,
de D. Croft e J. Lloyd)

1.1 Variedades de Matrizes

Neste capítulo vamos explorar a estrutura de espaço vectorial de $M_n(F)$ (e de outros espaços de matrizes), e a estrutura de variedade de alguns dos seus subconjuntos. Denotaremos por $\mathbf{A}^m(F)$ o espaço afim de dimensão m sobre F. Começamos com algumas definições elementares.

Definições 1.1 *Dizemos que um conjunto $V \in \mathbf{A}^m(F)$ é uma* Variedade *se for o conjunto de soluções de um sistema de equações polinomiais, com coeficientes em F, às quais chamaremos* equações definidoras *da variedade. Dizemos que uma variedade V é* irredutível *se, sempre que $V = V_1 \cup V_2$, com V_1 e V_2 variedades, tivermos $V = V_1$ ou $V = V_2$.*

Apresentamos a seguir alguns subconjuntos de $M_n(F)$ que se verifica sem dificuldade serem variedades.

Proposição 1.2 *Os seguintes conjuntos são variedades:*

1. *O conjunto das matrizes singulares,*

2. *O conjunto das matrizes derrogatórias,*

3. *O conjunto das matrizes com apenas um valor próprio, com F corpo algebricamente fechado,*

4. *O conjunto das matrizes com um valor próprio de multiplicidade algébrica maior que um.*

Demonstração. [1] Seja $f(x)$ o polinómio característico de A.

1. Basta observar que o cojunto em causa é

$$\{X \in M_n(F) : \det(X) = 0\}$$

e $\det(X)$ é polinomial nas entradas de X.

2. Afirmar que X é derrogatória é afirmar que os vectores $I, X, \ldots, X^{n-1}$ de $M_n(F)$ são linearmente dependentes, isto é, que todo o menor do tipo $n \times n$ da matriz

$$\left[\begin{array}{cccc} \mathrm{vec}(I) & \mathrm{vec}(X) & \cdots & \mathrm{vec}(X^{n-1}) \end{array} \right]$$

se anula, o que fornece $\binom{n^2}{n}$ equações polinomiais que definem a variedade.

3. Para que uma matriz tenha apenas um valor próprio λ, é necessário e suficiente que o seu polinómio característico tenha a forma

$$(x-\lambda)^n = \sum_{i=0}^{n} \binom{n}{i} (-\lambda)^{n-i} x^i = \sum_{i=0}^{n} a_i x^i. \tag{1}$$

Vamos considerar agora dois casos.

(i) Suponhamos que F tem característica p, com $p \nmid n$. Então temos

$$\frac{a_{n-1}}{n} = -\lambda,$$

e portanto, para $i = 0, \ldots, n-2$,

$$a_i = \binom{n}{i} \left(\frac{a_{n-1}}{n} \right)^{n-i} \tag{2}$$

As $n-1$ equações (2) definem a variedade pretendida, pois os coeficientes do polinómio característico são polinomiais nas entradas da matriz. De facto, se o polinómio característico de uma matriz A satisfizer as equações

[1] Em 3. e 4. vamos seguir ideias do Professor J. Dias da Silva.

(2), então, tomando $\lambda = -a_{n-1}/n$, o polinómio fica na forma (1), e A apenas tem um valor próprio.

(ii) Suponhamos agora que F tem característica p e $p \mid n$. Seja então $n = p^t k$, com $p \nmid k$. Temos então

$$\sum_{i=0}^{n} a_i x^i = (x-\lambda)^n = (x-\lambda)^{p^t k} = (x^{p^t} - \lambda^{p^t})^k$$

e fazendo $\Lambda = \lambda^{p^t}$ e $X = x^{p^t}$, obtemos, para o mesmo polinómio, a forma

$$(X-\Lambda)^k = \sum_{i=0}^{n} a'_i x^i \tag{3}$$

e $p \nmid k$. Podemos então aplicar o exposto em (i), e obtemos uma família de equações verificadas pelos coeficientes da expressão (3). Daqui se tiram as equações para os coeficientes da expressão (1), fazendo $a_{sp^t} = a'_s$, $s = 0, \ldots, k$, e $a_i = 0$ para os restantes índices, fazendo de novo $x^{p^t} = X$. Para verificar que estas equações definem de facto a variedade, basta ver que delas se obtém que $f(x)$ pode ter a forma $(x^{p^t} - \Lambda)^k$, ao fazer $\Lambda = -a_{n-1}/n$. Para se obter λ basta extrair a raiz índice-p^t de Λ, o que é possível por F ser algebricamente fechado. Temos assim sucessivamente

$$(x^{p^t} - \Lambda)^k = (x^{p^t} - \lambda^{p^t})^k = (x-\lambda)^n,$$

e A apenas tem um valor próprio.

4. Temos que A verifica a propriedade, por definição, se $f(x)$ tiver uma raiz multipla. Isto passa-se se e só se $f(x)$ e a sua derivada formal $f'(x)$ tiverem uma raiz comum. Ora, segundo [La, p. 202], isto é equivalente a afirmar que a resultante de $f(x)$ e $f'(x)$ é zero. Como a resultante é um polinómio nas entradas da matriz, temos o resultado. □

1.2 A Topologia de Zariski

Torna-se necessário, na demonstração dum teorema deste capítulo, utilizar o conceito de sucessão generalizada, que é a entidade que desempenha, no caso dos espaços topológicos que não são métricos, o papel desempenhado pelas sucessões nos espaços métricos. Assim, antes de começarmos a desenvolver a teoria relativa à topologia de Zariski (topologia que não provém de uma métrica, uma vez que nem sequer é separada), vamos estudar um

pouco o comportamento destas sucessões, seguindo [Ke]. A partir de agora, T será um espaço topológico qualquer.

Seja I um conjunto qualquer. Dizemos que $\preceq$ é uma *relação de ordem filtrante* em I se:

1. para $m \in I$, $m \preceq m$,
2. para $m, n, p \in I$, se $m \preceq n$ e $n \preceq q$, então $m \preceq p$ e
3. para $m, n \in I$, existe $p \in I$ tal que $m \preceq p$ e $n \preceq p$.

Escreveremos também $n \succeq m$ com o mesmo significado que $m \preceq n$. Seja então $\preceq$ uma relação de ordem filtrante em I. Uma *sucessão generalizada* $(x_i : i \in I)$ é uma aplicação de I em T. Diremos que (x_i) *converge para* x *em* T se, para toda a vizinhança V de x existir $p \in I$ tal que, para $i \in I, i \succeq p$ se tiver $x_i \in V$. Diremos também que X é *limite*[2] de (x_i). Note-se que, com esta definição, uma sucessão generalizada pode convergir para mais do que um ponto.

Sendo agora S outro espaço topológico, e $f : S \to T$ um função, dizemos que f *é contínua* se toda a pré-imagem de um aberto de T for um aberto de S. Apresentamos finalmente a proposição que nos vai ser útil. As demonstrações destas propriedades são simples e podem encontrar-se em [Ke], pp. 66 e 86.

Proposição 1.3 *Sejam S e T espaços topológicos quaisquer e $f : S \to T$ uma função. Temos que:*

1. *Um conjunto $A \subseteq T$ é fechado se e só se, para qualquer sucessão generalizada (x_i) de elementos de A, todos os limites de (x_i) pertencerem a A.*
2. *A função f é contínua num ponto $a \in S$ se e só se, para qualquer sucessão generalizada de elementos de S, (x_i), que convirja para a, a sucessão generalizada $f(x_i)$ convergir para $f(a)$.*

Vamos agora definir a topologia de Zariski, mostrando que as variedades em $\mathbf{A}^m(F)$ satisfazem as condições que definem a família de fechados de um espaço topológico. Começamos por alguns resultados gerais bem conhecidos. Sabemos que, dado um subcojunto qualquer de $\mathbf{A}^m(F)$, o conjunto dos

[2]Em [Ke] a palavra 'limite' não é usada nestas circunstâncias, mas é-o, por exemplo, em [Bo2].

polinómios que se anulam em todos os seus pontos é um ideal. Por outro lado, dado um ideal I de $F[X], X = (x_1, \ldots, x_m)$, pelo Teorema da Base de Hilbert, sabemos que ele admite um conjunto finito de geradores, e portanto, o conjunto de pontos de $\mathbf{A}^m(F)$ que anulam todos os polinómios de I é uma variedade.

Definição 1.4 (As correspondências $\mathcal{V}$ e $\mathcal{I}$) *Sejam I ideal de $F[X]$, e $A \subseteq \mathbf{A}^m(F)$. Então definimos $\mathcal{V}(I)$ e $\mathcal{I}(A)$ da seguinte forma:*

$$\mathcal{V}(I) = \{a \in \mathbf{A}^m(F) : \forall f \in I \; f(a) = 0\},$$

$$\mathcal{I}(A) = \{f \in F[X] : \forall a \in A \; f(a) = 0\}.$$

Temos que $\mathcal{V}(I)$ é uma variedade, e $\mathcal{I}(A)$ é um ideal de $F[X]$.

As seguintes propriedades são de verificação trivial.

Proposição 1.5 *Sejam U, V variedades, I, J ideais de $K[X]$. Temos que*

1. *$V \subseteq U \Rightarrow \mathcal{I}(V) \supseteq \mathcal{I}(U)$,*
2. *$I \subseteq J \Rightarrow \mathcal{V}(I) \supseteq \mathcal{V}(J)$,*
3. *$V = \mathcal{V}(\mathcal{I}(V))$,*
4. *$I \subseteq \mathcal{I}(\mathcal{V}(I))$, e a inclusão pode ser estrita.*

Observações. Como exemplo de um caso em que temos uma inclusão estrita em 4, podemos pensar num corpo F, que não seja algebricamente fechado, e no ideal I de $K[x]$ gerado por um polinómio f não constante que não tenha raízes em F. Neste caso, $\mathcal{V}(I) = \emptyset$, e $\mathcal{I}(\mathcal{V}(I)) = K[x] \supset I$.

Além disso, note-se que a aplicação $\mathcal{I}$ é injectiva no conjunto das variedades de $\mathbf{A}^m(F)$, e $\mathcal{I}$ sobrejectiva no conjunto dos ideais de $F[x]$, como se pode ver pela alínea 3.

Proposição 1.6 *Seja V variedade. Então V é irredutível se e só se $\mathcal{I}(V)$ for primo.*

Demonstração. Vamos demonstrar que V é redutível sse $\mathcal{I}(V)$ não é primo.

(i) Suponhamos que V é redutível e seja $V = V_1 \cup V_2$, com V_1 e V_2 variedades, $V_1, V_2 \subset V$. Como $V_i \subset V$, existem, pela injectividade de $\mathcal{I}$,

$f_i \in \mathcal{I}(V_i) \setminus \mathcal{I}(V), i = 1, 2$. Ora, $f_1 f_2$ anula-se em todos os pontos de V, logo pertence a $\mathcal{I}(V)$, que não é, portanto, primo.

(ii) Reciprocamente, suponhamos que $\mathcal{I}(V)$ não é primo; então existem $f_1, f_2 \notin \mathcal{I}(V)$ tais que $f_1 f_2 \in \mathcal{I}(V)$. Sejam então, para $i = 1, 2$,

$$I_i = (\mathcal{I}(V), (f_i)),\ V_i = \mathcal{V}(I_i),$$

e obtemos que $V_1, V_2 \subset V$ e $V \subseteq V_1 \cup V_2$, pois para $a \in V$, temos $f_1 f_2(a) = f_1(a) f_2(a) = 0$, o que implica que $f_1(a) = 0$ ou $f_2(a) = 0$. □

Proposição 1.7 *Temos as seguintes propriedades:*

1. $\mathcal{V}(0) = \mathbf{A}^m(F), \mathcal{V}(F[X]) = \emptyset$,
2. $\mathcal{V}(I_1 \cap I_2) = \mathcal{V}(I_1) \cup \mathcal{V}(I_2)$,
3. $\mathcal{V}(\sum_{\alpha \in A} I_\alpha) = \bigcap_{\alpha \in A} \mathcal{V}(i_\alpha)$.

Demonstração. A propriedade 1. é elementar.

2. É trivial ver que $\mathcal{V}(I_1) \cup \mathcal{V}(I_2) \subseteq \mathcal{V}(I_1 \cap I_2)$ usando a proposição 1.5, alínea 2. Para ver a recíproca, tome-se $a \in \mathcal{V}(I_1 \cap I_2)$, e suponhamos que $a \notin \mathcal{V}(I_1) \cup \mathcal{V}(I_2)$. Existiriam então $f_1 \in I_1$ e $f_2 \in I_2$ com $f_1(a) \neq 0$ e $f_2(a) \neq 0$. Ora, $f_1 f_2 \in I_1 \cap I_2$ e $f_1 f_2(a) \neq 0$, donde $a \notin \mathcal{V}(I_1 \cap I_2)$, contradição.

3. Novamente pela proposição 1.5, a inclusão $\mathcal{V}(\sum_{\alpha \in A} I_\alpha) \subseteq \bigcap_{\alpha \in A} \mathcal{V}(I_\alpha)$ é trivial. Para verificar a outra, tome-se $a \in \bigcap_{\alpha \in A} \mathcal{V}(I_\alpha)$. Tal a é raiz de todos os polinómios de todos os ideais da família $(I_\alpha)_{\alpha \in A}$. Assim, a é raiz de todos os polinómios de $\sum_{\alpha \in A} I_\alpha$, e isto dá o resultado. □

Estas são as condições que o conjunto de fechados de uma topologia tem que verificar. Podemos assim definir uma topologia sobre $\mathbf{A}^m(F)$, em que os fechados são exactamente as variedades. Esta é a *Topologia de Zariski*. Se F for $\mathbb{R}$ ou $\mathbb{C}$, as variedades são conjuntos fechados para a topologia usual, portanto esta é mais fina que a topologia de Zariski; em termos de sucessões convergentes, podemos afirmar que se (X_i) é uma sucessão convergente para X em $\mathbb{R}^m$ ou $\mathbb{C}^m$, para a topologia usual, então também converge para X para a topologia de Zariski, embora possa ter outros limites, pois esta topologia não é separada.

Sendo agora T um espaço topológico qualquer, e $A \subseteq T$, A é um espaço topológico com a topologia induzida, e para $B \subseteq A$, notaremos por $\overline{B}_A$ o fecho de B em A. Se $A = T$, notaremos o fecho de B apenas por $\overline{B}$.

Apresentamos a seguir um lema e uma proposição com alguns resultados de topologia geral.

Lema 1.8 *Seja T um espaço topológico, e $A, B \subseteq T$ subespaços, munidos da topologia induzida, e $A \subseteq V$. Então temos:*

1. $\overline{A} \cup \overline{B} = \overline{A \cup B}$
2. *Se $A \subseteq B$, $\overline{A}_B = \overline{A} \cap B$*
3. *Se $f : T \to T$ for uma função contínua, com $f(A) \subseteq A$, então $f(\overline{A}) \subseteq \overline{A}$.*

Um espaço topológico, em geral, diz-se *irredutível* se não se puder decompor como união de dois fechados que sejam seus subconjuntos próprios. Observe-se que, segundo esta definição, afirmar que uma variedade é irredutível enquanto subespaço topológico de $\mathbf{A}^m(F)$ é o mesmo que afirmar que é irredutível enquanto variedade, segundo a definição já dada. A seguinte proposição caracteriza os subconjuntos irredutíveis de um subespaço topológico.

Proposição 1.9 *Seja T um espaço topológico, $A \subseteq T$ um subespaço topológico, com a topologia induzida. As seguintes afirmações são equivalentes:*

1. *A é irredutível,*
2. *Se U_1 e U_2 são abertos não vazios de V, então $U_1 \cap U_2 \neq \emptyset$,*
3. *Todo o subconjunto aberto não vazio de A é denso em A.*
4. *$\overline{A}$ é irredutível,*

Demonstração. A implicação 1. $\Rightarrow$ 2. é trivial, bastando pensar na definição de variedade irredutível. Para a implicação 2. $\Rightarrow$ 3, raciocinemos por absurdo. Se B fosse um aberto não vazio denso em A, o complementar de $\overline{B}_A$ seria outro aberto não vazio contido em A, disjunto de B, o que é impossível. Para 3. $\Rightarrow$ 1, veja-se que, se nestas condições $A = A_1 \cup A_2$, com A_1 e A_2 fechados em A, etc, o aberto $A \setminus A_1$ não seria denso em A, pois estaria contido em $A_2 \subset A$.

Vejamos agora que 1. $\Leftrightarrow$ 4. Suponhamos que A é irredutível, e $\overline{A}$ não. Então $\overline{A} = A_1 \cup A_2$, fechados em $\overline{A}$, e portanto também em T. Então

$A = (A_1 \cap A) \cup (A_2 \cap A)$, fechados em A, e distintos de A, pois se, por exemplo, $A_1 \cap A = A$, viria que A_1 era fechado, com $A \subseteq A_1 \subseteq \overline{A}$, e por definição de fecho, $A_1 = \overline{A}$, o que é falso. Da mesma forma se vê que $A_1 \cap A \neq A$. A recíproca demonstra-se análogamente. □

Observação. Do resultado anterior se pode concluir que um conjunto $A \subseteq \mathbf{A}^m(F)$ é irredutível se e só se $\mathcal{I}(A)$ for primo, tendo em conta que $\mathcal{I}(A) = \mathcal{I}(\overline{A})$.

Finalmente, duas últimas propriedades, que dizem respeito à continuidade de funções.

Lema 1.10 *Sejam $A \subseteq \mathbf{A}^m(F), B \subseteq \mathbf{A}^n(F)$ e $f: A \to B$ uma função, com*

$$f = (f_1, \ldots, f_n) \text{ e } f_i = \frac{r_i}{s_i} \quad i \in [n],$$

com r_i e s_i polinómios em m variáveis, $s_i(a) \neq 0$ para qualquer $a \in A$. Então f é contínua, para as topologias de Zariski de ambos os conjuntos.

Demonstração. Vamos ver que a pré-imagem de qualquer fechado é um fechado. Seja então $U \subseteq B$ um fechado, com $U = V \cap B$, V variedade de $\mathbf{A}^n(F)$. Sejam $(p_1, \ldots, p_t)$ os polinómios definidores de V, de graus $(\delta_1, \ldots, \delta_t)$ respectivamente. Então, para $a \in A$,

$$\begin{aligned} a \in f^{-1}(V) &\Leftrightarrow f(a) \in V \\ &\Leftrightarrow \forall j \in [t] \quad p_j(f(a)) = 0 \\ &\Leftrightarrow \forall j \in [t] \quad \frac{\overline{p}_j(a)}{(s_j)^{\delta_j}(a)} = 0 \\ &\Leftrightarrow \forall j \in [t] \quad \overline{p}_j(a) = 0 \end{aligned}$$

em que $\overline{p}_j$ é um polinómio. Tomando então V' a variedade de $\mathbf{A}^m(F)$ definida por $(\overline{p}_1, \ldots, \overline{p}_t)$, temos que $a \in f^{-1}(U)$ se e só se $a \in A \cap V'$, que é um fechado de A. □

Lema 1.11 *Sejam $A \subseteq \mathbf{A}^m(F)$ e $B \subseteq \mathbf{A}^n(F)$, e $f : A \to B$ uma função contínua sobrejectiva. Então se A é irredutível, B é irredutível.*

Demonstração. Se $B = B_1 \cup B_2$, então $= f^{-1}(B_1) \cup f^{-1}(B_2)$, fechados em A, e distintos de A, pois se, por exemplo, $f^{-1}(B_1) = A$, como f é sobrejectiva, viria $B = f(f^{-1}(B_1)) = B_1$, o que é falso. □

1.3 A variedade das matrizes que comutam

Vamos tomar agora F algebricamente fechado, e

$$\mathbf{A}^m(F) = M_n(F) \times M_n(F),$$

que é um espaço afim de dimensão $2n^2$ sobre F, e definimos

$$P := \{\, (A, B) \in M_n(F) \times M_n(F) : AB = BA \,\},$$

que é uma variedade, definida pelas n^2 equações fornecidas pela equação de comutatividade. Em [MT] mostra-se que esta variedade é irredutível, usando a densidade de um subconjunto, o dos pares de matrizes que comutam e são simultaneamente diagonalizáveis. Apresentamos aqui esse resultado, e a irredutibilidade, como consequência.

Teorema 1.12 *O conjunto dos pares $(D_1, D_2) \in P$ que são simultaneamente diagonalizáveis, P_D, é denso em P.*

Demonstração. Vamos tomar um par $(A, B) \in P$ e mostrar que ele pertence ao fecho de P_D. Como F é algebricamente fechado, podemos sempre supor que A está na sua forma normal de Jordan, pois o isomorfismo de conjugação é uma bijecção contínua. Vamos agora considerar três casos.

1. Suponhamos que uma das matrizes tem dois valores próprios distintos, e, sem perda de generalidade, suponhamos que é a matriz A. Então A tem a forma $\mathrm{diag}(A_1, A_2)$, em que A_1 e A_2 não têm valores próprios comuns. Como B comuta com A, pelo que ficou exposto na introdução, pg. 9, B tem que ter a forma $\mathrm{diag}(B_1, B_2)$, com $A_iB_i = B_iA_i$, $i = 1, 2$, e o resultado vem por indução em n, tendo em conta que a propriedade é trivial para $n = 1$.

2. Suponhamos agora que A apenas tem um valor próprio, λ, mas mais do que um bloco de Jordan, isto é, mais do que um vector próprio. Seja t o tamanho do bloco de Jordan que aparece em primeiro lugar, e seja $C = [c_{ij}]$ a matriz diagonal com $c_{11} = \ldots = c_{tt} \neq c_{t+1\,t+1} = \ldots = c_{nn}$. Então, para qualquer $x \in F$, A comuta com $C_x := B + x(C - B)$. Como vimos na proposição 1.2, as matrizes com apenas um valor próprio formam um fechado, U, e $V := \{C_x : x \in F\} \not\subseteq U$, pois para $x = 1$, $C_x = C$ tem mais que um valor próprio. Ora, V é um conjunto irredutível, por ser imagem, por meio de uma função contínua, de F, que é irredutível por ser infinito. Assim, $V \setminus U$ é denso em V, e como já vimos em 1. que todos os pares (A, C_x), com $C_x \in V \setminus U$ pertenciam ao fecho de P_D, o mesmo acontece para todos os pares com $C_x \in V$, pela densidade, em particular para $C_0 = B$.

3. Finalmente, suponhamos que $A = J_n$, isto é, tem apenas um bloco de Jordan. Então $B = [b_{ij}]$ tem que ser triangular superior regular. Tomando $b = b_{12} = b_{23} = \ldots = b_{n-1\,n}$, a matriz $B - bA$ fica com mais do que um vector próprio, pois as duas primeiras colunas ficam nulas. Assim, pela alínea 2, o par $(A, B - bA)$ pertence ao fecho de P_D. Considerando agora a função

$$f : (X, Y) \mapsto (X, Y + bX),$$

f é contínua em $M_n(F) \times M_n(F)$, e $f(P_D) \subseteq P_D$, logo $f(\overline{P_D}) \subseteq \overline{P_D}$, pelo lema 1.8. Assim $(A, B) = f(A, B - bA) \in \overline{P_D}$. □

Observação. Se soubessemos já que P era irredutível, poderíamos obter facilmente o resultado anterior. Notando que o conjunto das matrizes com n valores próprios distintos (que são todas diagonalizáveis) é um aberto, pela proposição 1.2, e tomando o conjunto dos pares $(A, B) \in P$, com A e B nestas condições, então é simples de ver que A e B vêm simultaneamente diagonalizáveis (aplicando, por exemplo, os resultados da introdução, referentes ao estudo da equação de comutatividade), e obteríamos um aberto denso em P, contido em P_D. Assim, obtemos a irredutibilidade usando o resultado anterior.

Proposição 1.13 *O conjunto P é uma variedade irredutível.*

Demonstração. Vejamos que $\mathcal{I}(P_D)$ é primo, e o resultado vem pela proposição 1.9. Sejam f e g polinómios em n^2 indeterminadas, tais que $fg = 0$ em P_D, e seja W a variedade irredutível dos pares de matrizes diagonais. Para $A, B, T \in M_n(F)$, T invertível, vamos notar por $T(A, B)T^{-1}$ o par (TAT^{-1}, TBT^{-1}). Assim,

$$P_D = \bigcup_{T \in GL_n(F)} TWT^{-1}$$

e, para cada T invertível, $f = 0$ ou $g = 0$ em TWT^{-1}, que também é irredutível, para cada T. Sejam então

$$F_f := \{T \in GL_n(F) : f(TWT^{-1}) = 0\}$$

e F_g definido analogamente. Verificaremos, usando sucessões generalizadas, que estes conjuntos são fechados de $GL_n(F)$. Ora, $GL_n(F) = F_f \cup F_g$ e $GL_n(F)$ é irredutível, porque o seu fecho, $M_n(F)$, o é. Assim, um dos

fechados tem que ser todo o espaço, $f(P_D) = 0$ ou $g(P_D) = 0$, o que dá o resultado pretendido.

Façamos agora a verificação. Tomemos I, um conjunto com uma relação de ordem filtrante, e $(T_i : i \in I)$ uma sucessão generalizada de elementos de F_f e T um seu limite (para F_g o raciocínio seria o mesmo). Vejamos que $T \in F_f$, isto é, que $f(TWT^{-1}) = 0$. Seja então $(A, B) \in W$ qualquer. A aplicação que a T_i faz corresponder $f(T_i(A, B)T_i^{-1})$ é contínua, e identicamente igual a zero. Neste caso, esta última sucessão de elementos de F tem apenas um limite (que é zero), pois dado qualquer elemento de F diferente de 0, existe uma sua vizinhança que não tem nenhum ponto da sucessão, que é o complementar da variedade definida pelo polinómio $x = 0$. Portanto, como $T(A, B)T^{-1}$ é um limite da sucessão $T_i(A, B)T_i^{-1}$, $f(T(A, B)T^{-1}) = 0$ e F_f é fechado. □

Em 1961, Gerstenhaber [Ge4] demonstra este resultado, construindo um ponto genérico para a variedade. Demonstra igualmente que a álgebra gerada em $M_n(F)$ por duas matrizes que comutam tem dimensão menor ou igual a n (enquanto espaço vectorial sobre F), construindo uma álgebra de dimensão exactamente n que a contém. A sua demonstração assenta no estudo de algumas propriedades combinatórias das matrizes e é bastante elaborada. Em 1992, Guralnick [Gu] demonstra este resultado de uma forma assaz simples, usando a densidade de um outro conjunto, o dos pares em que uma das matrizes é não derrogatória. Passamos a apresentar essa demonstração.

Teorema 1.14 *O conjunto dos pares de matrizes que comutam, em que uma das matrizes é não derrogatória, é denso em P.*

Demonstração. Observe-se primeiro que, dada uma matriz A, é sempre possível encontrar uma matriz não derrogatória que comute com ela. De facto, se a forma normal de Jordan de A for $\operatorname{diag}(\lambda_1 I_{i_1} + J_{i_1}, \ldots, \lambda_t I_{i_t} + J_{i_t}) = TAT^{-1}$, então $R := T^{-1} \operatorname{diag}(\mu_1 I_{i_1} + J_{i_1}, \ldots, \mu_t I_{i_t} + J_{i_t})T$, com $\mu_1, \ldots, \mu_t$ elementos distintos de F, é não derrogatória e comuta com A.

Seja então $(A, B) \in P$. Então, para cada $x \in F$, $(A, B + x(R - B)) \in P$. Além disso, o conjunto dos x para os quais $B + x(R - B)$ é não derrogatória é um aberto não vazio de F, por ser pré-imagem de um aberto, o conjunto das matrizes não derrogatórias — veja-se a proposição 1.2. Pelo facto de F ser infinito, é irredutível, e portanto o conjunto em questão é denso, e assim o par (A, B) está no fecho do conjunto dos pares em que uma das matrizes é não derrogatória, pois a função $x \mapsto B + x(R - B)$ é contínua. □

Deste resultado também se pode concluir a irredutibilidade de P. Considere-se a aplicação contínua $f : F_n[x] \times M_n[F] \to P$ definida por $f(g, A) = (g(A), A)$, em que $F_n[x]$ é o conjunto dos polinómios de grau menor que n. A imagem contém todos os pares em que a segunda componente é não derrogatória, pois as matrizes que comutam com uma matriz não derrogatória são exactamente os polinómios nessa matriz, e pelo teorema de Hamilton--Cayley e pelo algoritmo de divisão em $F[x]$, podemos concluir que para obter todas as matrizes que comutam com uma matriz não derrogatória, basta usar os polinómios de grau inferior a n. Assim, a imagem é densa, e é também irredutível, porque o domínio o é (Lema 1.11). Pela Proposição 1.9, o seu fecho, P, é irredutível.

Teorema 1.15 *Seja F um corpo algebricamente fechado. A álgebra gerada em $M_n(F)$ por duas matrizes que comutam tem dimensão menor ou igual a n.*

Demonstração. Para $(A, B) \in P$, seja $T \in M_{n^2}(F)$ a matriz cujas colunas são $\mathrm{vec}(A^i B^j)$, para $i, j \in [n]$, tomando uma ordem qualquer para as entradas da matriz $A^i B^j$. Então, a condição $\dim \mathrm{alg} <A, B> \leq k$ pode traduzir-se por $\mathrm{c}(T) \leq k$, e as matrizes que satisfazem esta última condição formam um fechado em $M_{n^2}(F)$, pois a condição pode traduzir-se como o anulamento dos menores de ordem $k + 1$. Então o conjunto

$$Q := \{(A, B) \in P : \dim \mathrm{alg} <A, B> \leq n\}$$

é uma variedade contida em P, e contém todos os pares em que A é não derrogatória, pois nesse caso B é um polinómio em A, e $\dim \mathrm{alg} <A, B> = \dim \mathrm{alg} <A> = n$, pelo teorema de Hamilton-Cayley. Assim, $Q = P$, e temos o resultado. □

Observação. Esta demonstração exige que o corpo F seja algebricamente fechado, mas de facto, daqui é possível deduzir o resultado para qualquer corpo, com o processo usado no teorema 2.5.

Note-se também que este resultado diz apenas respeito a propriedades algébricas das matrizes, apesar de se terem usado nas demonstrações propriedades geométricas. Usando argumentos similares se consegue provar também que, se $\mathrm{c}(AB - BA) \leq 1$, então $\dim \mathrm{alg} <A, B> \leq n + (n^2 - \epsilon)/4$, com $\epsilon = 0$ se n é par, e $\epsilon = 1$ se n é ímpar (cf. [Gu] e [Ne]), o que é também uma propriedade algébrica.

No capítulo seguinte retomaremos o problema da dimensão da álgebra gerada por duas matrizes que comutam, abordado de outra forma, e apresentaremos uma nova demonstração do teorema 1.15.

□ □
□ □
□ □

Capítulo 2

Usando Teoria de Matrizes

`Eu diria mesmo mais:`
`a dimensão é menor ou igual a` n.
(adaptado de Dupond e Dupont, de Hergé)

2.1 Uma generalização do Teorema de Hamilton--Cayley

Vamos neste capítulo desenvolver uma generalização do Teorema de Hamilton-Cayley, seguindo S. Lazarus [Lz], com algumas modificações. Esta generalização não só nos parece interessante em si mesma, como dela se podem tirar alguns resultados, que aqui apresentamos.

Tomemos uma matriz A, com apenas um valor próprio, que esteja na sua forma normal de Jordan, $aI_n+\text{diag}(J_{k_1},\ldots,J_{k_r})$, com $k_1 \geq \ldots \geq k_r \geq 1$, e B uma matriz que comute com A. Como vimos na introdução, a matriz B tem uma forma especial, pelo facto de comutar com A. Ao particionarmos B em blocos $[B_{ij}]$, em que B_{ij} é do tipo $k_i \times k_j$, cada bloco tem que ser triangular superior regular. Ora, já vimos também que uma matriz triangular superior regular quadrada, do tipo $k \times k$ pode ser encarada como um polinómio em J_k. Assim, passamos a introduzir alguma notação: sejam $f \in F[x]$ um polinómio, e X uma matriz qualquer, do tipo $p \times q$;

- para $k_i \leq p$, notaremos por X_{T_i} a matriz X *truncada a* k_i, ou seja, a matriz do tipo $k_i \times q$ constituida pelas k_i primeiras linhas da matriz X,

- para $k_i \geq p$, notaremos por X_{E_i} a matriz X *estendida a* k_i, ou seja,

a matriz do tipo $k_i \times q$ cujas primeiras p linhas coincidem com as da matriz X, e as restantes $k_i - p$ são zero.

Note-se que E_i e T_i gozam de algumas propriedades. Por exemplo, para $k_i \leq k_j$, $(X_{T_j})_{T_i} = X_{T_i}$ e $(X_{E_i})_{E_j} = X_{E_j}$.

Com a notação anterior, podemos escrever:

- para $k_i \leq k_j$, $B_{ij} = h_{ij}(J_{k_j})_{T_i}$, e
- para $k_i \geq k_j$, $B_{ij} = h_{ij}(J_{k_j})_{E_i}$,

para alguns polinómios $h_{ij} \in K[x]$. Note-se que para o caso $k_i \leq k_j$, o polinómio h_{ij} só tem termos de grau superior a $k_j - k_i$.

Apresentamos agora, num lema, o morfismo que será essencial para a demonstração do teorema principal.

Lema 2.1 *Nas condições descritas acima, a aplicação*

$$\begin{array}{rcl} \mathcal{C}(A) & \to & M_r(F[A]) \\ B & \mapsto & \overline{B} \end{array}$$

com $\overline{B} = [\overline{B_{ij}}]$, $i, j \in [r]$, *definida por* $\overline{B}_{ij} = h_{ij}(A)$, *com os polinómios* h_{ij} *definidos acima, é um monomorfismo de anéis que respeita igualmente as estruturas de módulo sobre* $F[A]$ *de ambos os conjuntos.*

Demonstração. Sejam $B, C \in \mathcal{C}(A)$. A propriedade $\overline{B+C} = \overline{B} + \overline{C}$ é simples de verificar. Quanto à injectividade da aplicação, basta observar que, se $B, C \in \mathcal{C}(A), B \neq C$, existe um bloco B_{ij} de B que é distinto do respectivo bloco de C, C_{ij}. Se suposermos que $i \leq j$ para fixar ideias, $B_{ij} = h_{ij}(J_{k_j})_{E_i}$ e $C_{ij} = g_{ij}(J_{k_j})_{E_i}$, em que h_{ij} e g_{ij} são polinómios, e $h_{ij}(J_{k_j}) \neq g_{ij}(J_{k_j})$. Daqui se pode concluir que as imagens são distintas, e que a aplicação é injectiva.

Vejamos agora que a aplicação respeita a multiplicação. Sejam g_{ij}, $i, j \in [r]$ os polinómios fornecidos pelos blocos de C. Temos

$$\overline{BC} = \left[\overline{\sum_{s=1}^{r} B_{is}C_{sj}}\right] = \left[\sum_{s=1}^{r} \overline{B_{is}C_{sj}}\right],\ i, j \in [r],$$

e portanto, basta provar que, para todos os valores i, j e s se tem

$$\overline{B_{is}C_{sj}} = \overline{B_{is}}\,\overline{C_{sj}}.$$

Notemos então, para já, que a equação de comutatividade de A com B fornece r^2 equações do tipo $J_{k_i}B_{ij} = B_{ij}J_{k_j}$. Destas equações podemos obter, para qualquer $f \in F[x]$,

$$f(J_{k_i})B_{ij} = B_{ij}f(J_{k_j}). \tag{1}$$

Vamos agora considerar seis casos, que correspondem às seis maneiras distintas de k_i, k_j e k_s estarem ordenados.

· $k_i \leq k_s \leq k_j$. Pela equação (1), pondo h_{is} no lugar de f, k_i no lugar de k_s e B_{sj} no lugar de B_{ij}, obtemos $h_{is}(J_{k_s})B_{sj} = B_{sj}h_{is}(J_{k_j})$. Como $B_{is} = h_{is}(J_{k_s})_{T_i}$, podemos concluir das primeiras k_i linhas desta última equação que

$$\begin{aligned} B_{is}C_{sj} &= (C_{sj})_{T_i}h_{is}(J_{k_j}) \\ &= (g_{sj}(J_{k_j})_{T_s})_{T_i}h_{is}(J_{k_j}) \\ &= g_{sj}(J_{k_j})_{T_i}h_{is}(J_{k_j}) \\ &= (g_{sj}(J_{k_j})h_{is}(J_{k_j}))_{T_i} \end{aligned}$$

e portanto,

$$\overline{B_{is}C_{sj}} = g_{sj}(A)h_{is}(A) = h_{is}(A)g_{sj}(A) = \overline{B_{is}}\,\overline{C_{sj}}.$$

Os casos $k_j \leq k_i \leq k_s$ e $k_i \leq k_j \leq k_s$ têm um estudo semelhante, pois em qualquer um deles $k_i \leq k_s$.

· $k_j \leq k_s \leq k_i$. Da equação (1) temos $h_{is}(J_{k_s})B_{sj} = B_{sj}h_{is}(J_{k_j})$, tal como acima, o que corresponde a dizer que as primeiras k_s linhas de $B_{is}C_{sj}$ são exactamente $C_{sj}h_{is}(J_{k_j})$. Como as restantes linhas de $B_{is}C_{sj}$ são zero, podemos escrever

$$\begin{aligned} B_{is}C_{sj} &= (C_{sj})_{E_i}h_{is}(J_{k_j}) \\ &= (g_{sj}(J_{k_j})_{E_s})_{E_i}h_{is}(J_{k_j}) \\ &= g_{sj}(J_{k_j})_{E_i}h_{is}(J_{k_j}) \\ &= (g_{sj}(J_{k_j})h_{is}(J_{k_j}))_{E_i} \end{aligned}$$

donde

$$\overline{B_{is}C_{sj}} = g_{sj}(A)h_{is}(A) = h_{is}(A)g_{sj}(A) = \overline{B_{is}}\,\overline{C_{sj}}.$$

Tal como anteriormente, os casos $k_s \leq k_i \leq k_j$ e $k_s \leq k_j \leq k_i$ seguem analogamente, pois aqui $k_s \leq k_i$. Isto mostra que a aplicação é de facto um monomorfismo de anéis.

No que diz respeito às estruturas de módulos, veja-se que

$$\overline{g(A)} = \overline{\mathrm{diag}(g(J_{k_1}), \ldots, g(J_{k_r}))} = \mathrm{diag}(\overbrace{g(A), \ldots, g(A)}^{\text{r vezes}}), \qquad (2)$$

para qualquer $g(x) \in K[x]$. Usando agora a propriedade multiplicativa, temos

$$\overline{g(A)B} = \overline{g(A)}\,\overline{B} = g(A)\overline{B},$$

pela relação (2). Isto prova o pretendido, e termina a demonstração. □

Apresentamos agora o resultado central deste capítulo.

Teorema 2.2 *Sejam A e B matrizes de $M_n(F)$, com $AB = BA$, e suponhamos que A tem apenas um valor próprio e está na sua forma normal de Jordan, $aI_n + \mathrm{diag}(J_{k_1}, \ldots, J_{k_r})$, com $k_1 \geq \ldots \geq k_r \geq 1$. Então B é raiz um polinómio mónico de grau r com coeficientes em $F[A]$.*

Demonstração. Considere-se o morfismo definido acima, e $\overline{B}$, imagem de B. Ora, $\overline{B}$ tem entradas em $F[A]$, anel comutativo, onde é válido o Teorema de Hamilton-Cayley (cf. por exemplo [Br]) isto é, se pusermos $p(x) = \det(xI_r - \overline{B})$[1], $p(x)$ tem coeficientes em $F[A]$ e grau r. Além disso, $p(\overline{B}) = 0$, e portanto, pelas propriedades referidas no lema, $p(B) = 0$, o que demonstra o resultado. □

Ao polinómio $p(x) = \det(xI_r - \overline{B})$ chamaremos o *A-polinómio característico* de B.

Podemos verificar facilmente que este teorema é de facto uma generalização do teorema de Hamilton-Cayley, tomando B uma matriz qualquer, e $A := 0_{nn}$. Neste caso, $r = n$ e os polinómios h_{ij} só têm termo independente, que coincide com a entrada (i, j) de B. Vamos obter, para $\overline{B}$ uma matriz formada por matrizes escalares do tipo $n \times n$, e ao calcular o seu polinómio característico, vamos obter um polinómio cujos coeficientes são novamente matrizes escalares, em que na diagonal principal aparecem os coeficientes do polinómio característico de B. Assim, o 0-polinómio característico de B pode identificar-se com o polinómio característico de B, e neste caso, o resultado anterior corresponde ao teorema de Hamilton-Cayley.

Este teorema, tal como está enunciado, exige que A tenha apenas um valor próprio e que esteja na sua forma normal de Jordan, mas de facto o resultado pode estabelecer-se num quadro mais geral.

[1]Estamos a notar por I_r a matriz identidade de $M_r(F[A])$.

Corolário 2.3 *Considerem-se $A, B \in M_n(F)$, com $AB = BA$, em que F é um corpo que contém todos os valores próprios de A, e suponhamos que A tem r divisores elementares. Então B é raiz um polinómio mónico de grau r com coeficientes em $F[A]$.*

Demonstração. Vamos para já ver que o resultado é válido se A tiver mais do que um valor próprio, estando ainda na sua forma normal de Jordan. Procedemos por indução no número de valores próprios, q, estando o caso $q = 1$ já demonstrado. Suponhamos então o resultado válido para um certo $q \geq 1$, e tomemos A com $q + 1$ valores próprios, ainda na sua forma normal de Jordan. Então $A = \mathrm{diag}(A_1, A_2)$, em que A_1 é soma directa dos blocos de Jordan associados a q dos valores próprios, e A_2 a soma directa dos restantes blocos, todos associados ao mesmo valor próprio. Sejam s e t o número de divisores elementares de A_1 e A_2, respectivamente. Claramente $r = s + t$. Como já vimos na Introdução pg. 9, temos que ter $B = \mathrm{diag}(B_1, B_2)$, com B_i do mesmo tipo que A_i, e $A_i B_i = B_i A_i$, $i = 1, 2$. Ora, A_1 e B_1 estão nas condições da hipótese de indução, e A_2 e B_2 nas condições do teorema inicial 2.2. Existem portanto o A_1-polinómio característico de B_1, e o A_2-polinómio característico de B_2, sejam respectivamente

$$\begin{aligned} g(x) &= x^s + g_{s-1}(A_1)x^{s-1} + \ldots + g_0(A_1), \\ h(x) &= x^t + h_{t-1}(A_2)x^{t-1} + \ldots + h_0(A_2). \end{aligned}$$

Pomos agora

$$\begin{aligned} \tilde{g}(x) &:= x^s + g_{s-1}(A)x^{s-1} + \ldots + g_0(A), \\ \tilde{h}(x) &:= x^t + h_{t-1}(A)x^{t-1} + \ldots + h_0(A), \end{aligned}$$

e tomamos

$$f(x) := \tilde{g}(x)\tilde{h}(x),$$

que é um polinómio de grau r, com coeficientes em $F[A]$. Resta verificar que $f(B) = 0$. Tendo então em conta que A e B são diagonais por blocos, e que, para qualquer $i \in [s]$, $g_i(A) = \mathrm{diag}(g_i(A_1), g_i(A_2))$, e o mesmo para os coeficientes de $\tilde{h}$,

$$\begin{aligned} f(B) &= \tilde{g}(B)\tilde{h}(B) \\ &= \mathrm{diag}(g(B_1), *)\,\mathrm{diag}(*, h(B_2)) \\ &= \mathrm{diag}(0, *)\,\mathrm{diag}(*, 0) = 0. \end{aligned}$$

Suponhamos finalmente que A não está na sua forma normal de Jordan. Como F contém todos os valores próprios de A, existe uma matriz invertível, T, tal que $\tilde{A} := TAT^{-1}$ é a forma normal de Jordan de A. Nestas condições, $\tilde{B} := TBT^{-1}$ comuta com $\tilde{A}$, e portanto, existe o $\tilde{A}$-polinómio característico de $\tilde{B}$, seja $\tilde{f} = \sum_{i=1}^{r} f_i(\tilde{A})x^i$, $f_r(x) \equiv 1$. Definimos então

$$f(x) := \sum_{i=1}^{r} f_i(A)x^i,$$

que é mónico, e tem o grau pretendido. Vejamos que $f(B) = 0$.

$$\begin{aligned} f(B) &= \sum_{i=1}^{r} f_i(T^{-1}\tilde{A}T)(T^{-1}\tilde{B}T)^i \\ &= \sum_{i=1}^{r} T^{-1} f_i(\tilde{A}) T T^{-1} \tilde{B}^i T \\ &= T^{-1}\left(\sum_{i=1}^{r} f_i(\tilde{A})\tilde{B}^i\right) T = 0, \end{aligned}$$

o que termina a demonstração. □

2.2 Dimensões de álgebras de matrizes que comutam

Vamos agora usar os resultados desenvolvidos na secção anterior para obter alguns limites superiores para as dimensões de álgebras geradas por matrizes que comutam, exibindo, em cada caso, uma base do espaço considerado. Assim, sempre que falarmos em 'família geradora' dum espaço, entenda-se que se está a considerar uma família de vectores que o gera enquanto espaço vectorial, ou seja, uma família que contém uma base.

Álgebras geradas por duas matrizes

Vamos apresentar aqui uma nova demonstração do resultado 1.15 — o Teorema 2.5.

Teorema 2.4 *Nas condições do Teorema 2.2, com a hipótese adicional de A ser nilpotente, existe uma base para* alg $<I_n, A, B>$ *com a forma*

$$\{\, A^i B^j : 0 \le j \le r-1 \text{ e, para cada j, } 0 \le i \le k'_{j+1} - 1 \,\}, \qquad (3)$$

com $k_1 = k'_1$, $k_i \ge k'_i$ para $i = 2, \ldots, r$ e $k'_1 \ge \ldots \ge k'_r$

Demonstração. Pelo Teorema 2.2, podemos escrever B^r como combinação linear, com coeficientes em $F[A]$, das potências inferiores de B. Então existe uma base de alg $<I_n, A, B>$ contida na família

$$(A^i B^j : 0 \leq i \leq k_1 - 1,\ 0 \leq j \leq r - 1),$$

uma vez que $A^{k_1} = 0$. Vamos provar agora que podemos melhorar esta família geradora, mostrando que

$$\begin{array}{l} (I_n, A, A^2, \ldots, A^{k_1-1}, \\ \ B, AB, A^2B, \ldots, A^{k_2-1}B, \\ \ \vdots \\ \ B^{r-1}, AB^{r-1}, \ldots, A^{k_r-1}B^{r-1}) = \end{array}$$

$$(A^i B^j : 0 \leq j \leq r - 1 \text{ e, para cada } j,\ 0 \leq i \leq k_{j+1} - 1) \qquad (4)$$

é ainda uma família geradora. Para o provar, basta ver que, para cada $j \in [r]$, $A^{k_j} B^{j-1}$ pode ser escrita como combinação linear dos elementos anteriores, isto é, dos elementos $A^i B^s$, com $0 \leq s \leq j$, e para cada s, $0 \leq i \leq k_{s+1} - 1$. Isto porque assim se pode garantir, por iteração, que para $k_{j+1} < i < k_1$, $A^i B^{j-1}$ é ainda combinação linear desses elementos. Demonstraremos isto usando indução em j.

Para $j = 1$ a afirmação é trivial, pois $A^{k_1} = 0$. Tomemos então $s > 1$ e suponhamos que temos o resultado para todos os valores de j inferiores a s. Consideremos B particionada por blocos

$$B = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$$

em que o bloco B_1 contém os primeiros s^2 blocos de B, isto é, $B_1 = [B_{it}]$, $i, t \in [s]$. Tome-se igualmente A na forma

$$A = \text{diag}(A_1, A_2),$$

com $A_1 = \text{diag}(J_{k_1}, \ldots, J_{k_s})$. Da equação $AB = BA$ podemos tirar que $A_1 B_1 = B_1 A_1$, e portanto existe o A_1-polinómio característico de B_1, seja

$$B_1^s = f_{s-1}(A_1) B_1^{s-1} + \ldots + f_1(A_1) B_1 + f_0(A_1).$$

Tomemos agora $t \leq s$, qualquer. Ao calcular B^t, verificamos que tem a forma

$$\begin{bmatrix} B_1^t + D_1 & D_2 \\ D_3 & D_4 \end{bmatrix}$$

em que cada parcela de $D_1, \ldots, D_4$ é um produto de t matrizes em que aparece B_2, B_3 ou B_4. Estas matrizes B_2, B_3 e B_4 são formadas por blocos que são matrizes triangulares superiores regulares, do tipo $k_\alpha \times k_\beta$, com $k_\alpha \leq k_{s+1}$ ou $k_\beta \leq k_{s+1}$. Em qualquer caso, não têm mais de k_{s+1} linhas não nulas, que serão sempre as primeiras k_{s+1} linhas. Ora, vemos facilmente que

$$A^{k_{s+1}} = \begin{bmatrix} A_1^{k_{s+1}} & 0 \\ 0 & 0 \end{bmatrix}$$

e que os blocos que constituem $A_1^{k_{s+1}}$ têm k_{s+1} colunas nulas, que são as primeiras também. Portanto, ao fazer a multiplicação $A^{k_{s+1}}B^t$, os blocos D_1, D_2, D_3 e D_4 são aniquilados, lembrando que A_1 comuta com B_1. Pondo assim

$$\tilde{B_1} := \begin{bmatrix} B_1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{e} \quad \tilde{A_1} := \begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix},$$

e, por questão de comodidade, $f_s \equiv -1$, temos sucessivamente:

$$\begin{aligned} & A^{k_{s+1}}(B^s - f_{s-1}(A)B^{s-1} - \ldots - f_1(A)B - f_0(A)) \\ & \quad = \sum_{i=0}^{s} -f_i(A)A^{k_{s+1}}{B_1}^i \\ & \quad = \sum_{i=0}^{s} -A^{k_{s+1}}f_i(A)\tilde{B}^i \\ & \quad = A^{k_{s+1}} \sum_{i=0}^{s} -f_i(\tilde{A_1}){\tilde{B_1}}^i = 0. \end{aligned}$$

Assim, $A^{k_{s+1}}B^s = \sum_{i=0}^{s-1} g_i(A)B^i, g_i = -f_i$. Para controlar agora o grau dos polinómios $g_0, \ldots, g_{s-1}$, usamos a hipótese de indução, e podemos tomar assim o grau de g_i menor que k_{i+1}. Isto prova que a família (4) é de facto geradora.

Finalmente, vamos escolher uma base com elementos desta família, fazendo o seguinte: para cada j, sucessivamente, retiramos o elemento A^iB^j da lista se ele puder ser escrito como combinação linear dos elementos anteriores. Isto fornece uma nova família,

$$\{\, A^iB^j : 0 \leq j \leq r-1 \text{ e, para cada j, } 0 \leq i \leq k'_{j+1} - 1 \,\},$$

com $k'_j \leq k_j$, $k'_1 = k_1$. Esta família é linearmente independente, pois se houvesse uma combinação linear nula dos seus elementos, em que nem todos

os coeficientes fossem nulos, então o último elemento com um coeficiente não nulo seria combinação linear dos anteriores, o que contraria a nossa construção. Isto mostra que, de facto, a família exibida é uma base. Além disso, se $A^{k'_i}B^{i-1}$ pode ser escrito como combinação linear dos elementos anteriores, multiplicando a expressão por B, vem que $A^{k'_i}B^i$ também pode ser escrito como combinação linear de elementos anteriores. Assim, $k'_{i+1} \leq k'_i$, e isto termina a nossa demonstração. □

Teorema 2.5 *Seja F um corpo, e $A, B \in M_n(F)$, duas matrizes que comutam. Então* $\dim \operatorname{alg} <A, B> \leq n$.

Demonstração. Seja $\mathcal{A} := \operatorname{alg} <I_n, A, B>$, e seja $\overline{F}$ o fecho algébrico de F. Tomando $\overline{\mathcal{A}} := \mathcal{A} \otimes_F \overline{F}$, temos $\dim_{\overline{F}} \overline{\mathcal{A}} = \dim_F \mathcal{A}$, o que nos permite supor que o corpo é algebricamente fechado.

Suponhamos agora que $\mathcal{A}$ é decomponível, isto é, que existem subálgebras $\mathcal{A}_1$ e $\mathcal{A}_2$ tais que $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$ $\mathcal{A}_1$ e $\mathcal{A}_2$ são ideais de $\mathcal{A}$ e $\mathcal{A}_1 \cap \mathcal{A}_2 = 0$ (portanto $\mathcal{A}_1\mathcal{A}_2 = \mathcal{A}_2\mathcal{A}_1 = 0$). Nestas condições, a menos de isomorfismo,

$$A = \operatorname{diag}(A_1, A_2) \text{ e } B = \operatorname{diag}(B_1, B_2),$$

com $A_i, B_i \in M_{n_i}(F)$, $i = 1, 2$ e $n_1 + n_2 = n$. Além disso $A_iB_i = B_iA_i$, e $\mathcal{A}_i = \operatorname{alg} <I_{n_i}, A_i, B_i>$. Usando indução, podemos supor que

$$\dim \mathcal{A}_i \leq n_i \qquad i = 1, 2,$$

e como $\mathcal{A} = \mathcal{A}_1 \oplus \mathcal{A}_2$, temos o resultado. Podemos portanto supor que $\mathcal{A}$ é indecomponível.

Finalmente, como consideramos F algebricamente fechado, suponhamos que A está na sua forma normal de Jordan. Se A tivesse dois valores próprios distintos, viria $B = \operatorname{diag}(B_1, B_2)$, pela equação de comutatividade, e $\mathcal{A}$ seria decomponível. Assim, podemos supor que A apenas tem um valor próprio, isto é, $A = aI_n + N$, em que N é nilpotente. Porém, como $I_n \in \mathcal{A}$, temos

$$\mathcal{A} = \operatorname{alg} <I_n, N, B>,$$

e portanto podemos supor que A é nilpotente.

Ora, nestas condições, já construímos uma base — o conjunto (3) — com $\sum_{i=1}^r k'_i \leq \sum_{i=1}^r k_i = n$ elementos, o que demontra o resultado. □

Álgebras geradas por três matrizes

Este teorema permite também obter alguns resultados para as álgebras geradas por três matrizes que comutam. Os raciocínios apresentados seguirão sempre a linha anterior: apresenta-se um conjunto gerador, contam-se os seus elementos e obtém-se um limite para a dimensão do espaço por eles gerado. Tomemos agora

$$\mathcal{A} := \text{alg} <I_n, A, B, C>,$$

em que $A, B, C \in M_n(F)$ e comutam duas a duas. Tal como acima, podemos supor, sem perda de generalidade, que F é algebricamente fechado, $\mathcal{A}$ é indecomponível, e vamos agora considerar que A, B e C são nilpotentes, e que A está na sua forma normal de Jordan, $A = \text{diag}(J_{k_1}, \ldots, J_{k_r})$, onde $k_1 \geq \ldots \geq k_r \geq 1$. Pelo Teorema 2.5, sabemos que $\dim \text{alg} <B, C> \leq n$. Logo, se o conjunto $\{ X_1, \ldots, X_s \}$, $s \leq n$, for uma base de $\text{alg} <B, C>$, então o conjunto

$$\{ A^i X_j : 0 \leq i \leq k_1 - 1, j = 1, \ldots, s \}$$

é um conjunto gerador para $\mathcal{A}$. Ora, podemos fazer este raciocínio começando por mudar o nome ás matrizes, de modo que k_1 seja o menor dos maiores blocos de Jordan de A, Be C, isto é, supondo que A tem o menor índice de nilpotência das três matrizes que geram a álgebra. Assim, temos imediatamente o seguinte resultado.

Proposição 2.6 *Suponhamos que A tem o menor dos índices de nilpotência de A, B e C, k_1. Então* $\dim \mathcal{A} \leq nk_1$.

A partir de agora teremos também que supor que F tem característica zero.

Lema 2.7 *Suponhamos que k_1 é o menor índice de nilpotência das três matrizes A, B e C. Então, o conjunto*

$$S_{r-1} := <\{ A^i B^j C^l : 0 \leq i \leq k_1 - 1,\ 0 \leq j + l \leq r - 1 \}>$$

contém uma base de $\mathcal{A}$.

Demonstração. Vamos mostrar, de modo análogo ao que já fizemos, que para $j = 0, \ldots, r$, $B^j C^{r-j} \in S_{r-1}$. Novamente, por iteração, podemos concluir disto que para quaisquer i, j, l, $A^i B^j C^l \in S_{r-1}$.

Para qualquer $x \in F$, $B + xC \in \mathcal{A}$, portanto, pelo Teorema 2.2,

$$(B + xC)^r = \sum_{i=0}^{r-1} f_i(A)(B + xC)^i,$$

para alguns polinómios $f_i(x) \in F[x]$. Então $(B + xC)^r$ é combinação linear dos elementos de S_{r-1} para qualquer $x \in F$. Ora, tomando sucessivamente $x = i = 1, \ldots, r + 1$, temos

$$(B + iC)^r = \sum_{j=0}^{r} \binom{r}{j} i^j B^{r-j} C^j \ \in S_{r-1}, \tag{5}$$

para cada i. Seja agora M a matriz do tipo $(r + 1) \times (r + 1)$ definida por

$$M := \left[i^{j-1} \binom{r}{j} \right] = \begin{bmatrix} 1 & \binom{r}{1} & \cdots & \binom{r}{r} \\ 1 & 2\binom{r}{1} & \cdots & 2^r\binom{r}{r} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (r+1)\binom{r}{1} & \cdots & (r+1)^r\binom{r}{r} \end{bmatrix}.$$

Assim,

$$\det M = \binom{r}{1}\binom{r}{2}\cdots\binom{r}{r} \det[i^{j-1} \,:\, i, j \in [r+1]\,] \neq 0,$$

pois a última matriz é uma matriz de Vandermonde. Reescrevendo então as $r + 1$ equações da fórmula (5), vem

$$M \begin{bmatrix} B^r \\ B^{r-1}C \\ \vdots \\ BC^{r-1} \\ C^{r-1} \end{bmatrix} \in (S_{r-1})^{r+1},$$

sendo $(S_{r-1})^{r+1}$ a $(r + 1)$-ésima potência cartesiana de S_{r-1}. Finalmente,

$$\begin{bmatrix} B^r \\ B^{r-1}C \\ \vdots \\ BC^{r-1} \\ C^{r-1} \end{bmatrix} = M^{-1}M \begin{bmatrix} B^r \\ B^{r-1}C \\ \vdots \\ BC^{r-1} \\ C^{r-1} \end{bmatrix} \in M^{-1}(S_{r-1})^{r+1} \subseteq (S_{r-1})^{r+1},$$

donde se conclui que $B^iC^{r-i} \in S_{r-1}$, para $0 \leq i \leq r$, o que termina a verificação. Demonstrámos assim que o conjunto

$$\{\, A^iB^jC^l : 0 \leq i \leq k_1 - 1,\, 0 \leq j + l \leq r - 1 \,\}$$

contém uma base de $\mathcal{A}$. □

Proposição 2.8 *Suponhamos que k_1 é o menor índice de nilpotência das três matrizes A, B e C. Então* $\dim \mathcal{A} \leq k_1 r(r+1)/2$.

Demonstração. Basta ver que o número de pares (j, l) com $0 \leq j + l \leq r - 1$ é $r(r+1)/2$. □

Note-se que este limite superior depende de k_1 e de r, ao passo que o limite da proposição 2.6 dependia de k_1 e de n.

Vamos agora mudar o tipo de suposição que fazemos sobre k_1. Passamos a mostrar que podemos supor que A tem o maior índice de nilpotência de $\mathcal{A}$, isto é, para qualquer s tal que $A^s = 0$, temos $X^s = 0$, se $X \in \mathcal{A}$. Isto corresponde a dizer que A tem o maior bloco de Jordan associado a zero dentre as matrizes de $\mathcal{A}$, que são todas nilpotentes.

Lema 2.9 *Seja l o maior índice de nilpotência de $\mathcal{A}$. Então existe $D \in \mathcal{A}$, de índice de nilpotência, l, e $D = a_1A + a_2B + a_3C + a_4AB + \ldots$ uma sua expressão como combinação linear dos elementos geradores em que $a_1 \neq 0$.*

Demonstração. Vamos supor que $a_1 = 0$, com vista a um absudo. Então $D = a_2B + a_3C + a_4AB + \ldots$, e, para qualquer $a_1 \in F$, $a_1 \neq 0$, $a_1A + D$ tem um índice de nilpotência menor. Posto de outro modo, $D^l = 0$, $D^{l-1} \neq 0$ e $(a_1A + D)^{l-1} = 0$. Assim,

$$\begin{aligned} 0 &= (a_1A + D)^{l-1} \\ &= \sum_{i=0}^{l-1} (a_1A)^i D^{l-1-i} \\ &= D^{l-1} + AD^{l-2}a_1 + \ldots + A^{l-2}Da_1^{l-2} + A^{l-1}a_1^{l-1}. \end{aligned}$$

O somatório anterior pode ser interpretado como um polinómio de coeficientes matriciais na variável a_1, ou como uma matriz de polinómios. Ora, cada um destes polinómios tem grau $l - 1 \leq n - 1$, e anula-se para todos os valores de a_1 diferentes de zero. Assim, todos os polinómios viriam identicamente nulos, e logo, para todo o i, $A^iD^{l-1-i} = 0$, o que é falso, em

particular, para $i = 0$, pois viria $D^{l-1} = 0$. Portanto, é possível tomar $a_1 \neq 0$. □

Pelo lema anterior, $\text{alg} <I_n, A, B, C> = \text{alg} <I_n, B, C, D>$, e podemos assim supor, sem perda de generalidade, que A tem o maior índice de nilpotência de $\mathcal{A}$, k_1. Vamos apresentar uma relação semelhante às já apresentadas: assim como de $(B + xC)^r \in <S_{r-1}>$ pudemos concluir que $B^j C^{(r-j)} \in <S_{r-1}>$, agora, de $(A + xB + yC)^{k_1} = 0$ vamos poder concluir que $A^i B^j C^l = 0$ se $i + j + l = k_1$, e, por iteração, se $i + j + l \geq k_1$.

Lema 2.10 *Suponhamos que k_1 é o maior índice de nilpotência de $\mathcal{A}$. Temos $A^i B^j C^s = 0$ sempre que $i + j + s \geq k_1$. Assim, o conjunto*

$$\{\, A^i B^j C^s \,:\, 0 \leq i \leq k_1 - 1, \qquad 0 \leq j + s \leq \min\{r - 1, k_1 - i - 1\} \,\}$$

gera $\mathcal{A}$ como espaço vectorial, isto é, contém uma base de $\mathcal{A}$.

Demonstração. Estamos agora a supor que, para todos os valores de x e y, temos $(A+xB+yC)^{k_1} = 0$. Ao fazer o desenvolvimento deste trinómio, obtemos

$$\begin{aligned} 0 &= \sum_{\alpha=0}^{k_1} \sum_{\beta=0}^{j} \binom{k_1}{\alpha} \binom{\alpha}{\beta} x^\beta y^{\alpha-\beta} A^{k_1-\alpha} B^{\alpha-\beta} C^\beta \\ &= \sum_{i+j+s=k_1} \gamma_{ijs} x^j y^s A^i B^j C^s \end{aligned}$$

em que $\gamma_{ijs} \in F$, $\gamma_{ijs} \neq 0$, para cada terno (i, j, s) que aparece no somatório. Pondo agora $y := x^{k_1+1}$, obtemos um polinómio em x, de grau $k_1^2 + k_1$ (o termo $x^{k_1^2+k_1}$ obtém-se para $(i, j, s) = (0, 0, k_1)$). Além disso, para ternos distintos, vamos obter potências de x distintas, uma vez que $j \leq k_1$ e os expoentes têm a forma $j+s(k_1+1)$. Assim, os coeficientes do novo polinómio vão ser ou zero ou $\gamma_{ijs} A^i B^j C^s$. Novamente, podemos encarar este polinómio com coeficientes matriciais como uma matriz de polinómios, todos de grau $k_1^2 + k_1$, que se anulam para todos os valores de x. Assim, todos os seus coeficientes têm que ser nulos, o que demonstra o pretendido.

O conjunto gerador obtém-se conjugando o que se acabou de mostrar com o resultado do lema 2.7, tendo em conta que essa majoração ainda é válida por maioria de razão ao fazermos a suposição presente sobre k_1. □

Proposição 2.11 *Suponhamos que k_1 é o maior índice de nilpotência de $\mathcal{A}$. Temos que*

$$\dim \mathcal{A} \leq \frac{k_1(k_1+1)(k_1+2)}{6} \text{ se } k_1 \leq r, \text{ e}$$

$$\dim \mathcal{A} \leq \frac{r(r+1)(3k_1-2r-1)}{6} \text{ se } k_1 > r.$$

Demonstração. Para o caso $k_1 \leq r$, $\min\{r-1, k_1-i-1\} = k_1-i-1$, para qualquer i, com $0 \leq i \leq k_1$. Portanto o conjunto gerador pode escrever-se como:

$$\begin{aligned}
&\{\, B^jC^s : 0 \leq j+s \leq k_1-1 \,\} \cup \\
&\quad \cup \{\, AB^jC^s : 0 \leq j+s \leq k_1-2 \,\} \cup \\
&\quad \cup \cdots \cup \\
&\quad \cup \{\, A^{k_1-2}B^jC^s : 0 \leq j+s \leq 1 \,\} \cup \\
&\quad \cup \{A^{k_1-1}\},
\end{aligned}$$

e portanto[2]

$$\begin{aligned}
\dim \mathcal{A} &\leq \frac{k_1(k_1+1)}{2} + \frac{k_1(k_1-1)}{2} + \ldots + \frac{2\times 1}{2} + 1 \\
&= \frac{1}{2}\sum_{i=1}^{k_1} i(i+1) \\
&= \frac{1}{2}\sum_{i=1}^{r-1} i^2 + \frac{k_1(k_1+1)}{4} \\
&= \frac{k_1(k_1+1)(k_1+2)}{6}.
\end{aligned}$$

Se $k_1 > r$, com $k_1 - d = r$, nos primeiros d subconjuntos, temos $r-1 \leq k_1-i-1$, e portanto toma-se $j+s \leq r-1$. Para $i > d$, $k_1-i-1 < r-1$, e toma-se $j+s \leq k_1-i-1$. Assim, o conjunto fica

$$\begin{aligned}
&\{\, B^jC^s : 0 \leq j+s \leq r-1 \,\} \cup \\
&\quad \cup \{\, AB^jC^s : 0 \leq j+s \leq r-1 \,\} \cup \\
&\quad \cup \cdots \cup
\end{aligned}$$

[2]Aqui usamos, além da fórmula da soma dos termos de uma progressão aritmética, a fórmula $\sum_{i=1}^{m} i^2 = \frac{(2m+1)(m+1)m}{6}$.

$$
\begin{aligned}
&\cup\{\, A^d B^j C^s : 0 \le j+s \le r-1 \,\} \cup \\
&\cup\{\, A^{d+1} B^j C^s : 0 \le j+s \le k_1 - d - 2 \,\} \cup \\
&\cup \cdots \cup \\
&\cup\{\, A^{k_1-2} B^j C^s : 0 \le j+s \le 1 \,\} \cup \\
&\cup\{A^{k_1-1}\},
\end{aligned}
$$

o que dá

$$
\begin{aligned}
\dim \mathcal{A} &\le d\frac{(r+1)r}{2} + \frac{r(r-1)}{2} + \ldots + \frac{3\times 2}{2} + 1 \\
&= d\frac{(r+1)r}{2} + \sum_{i=1}^{r-1} i(i-1) \\
&= \frac{(k_1-r)(r+1)r}{2} + \frac{(r-1)(r+1)r}{6} \\
&= \frac{(3k_1-2r-1)(r+1)r}{6},
\end{aligned}
$$

o que demonstra o pretendido. □

No próximo capítulo, apresentaremos uma simplificação da demonstração do Teorema 2.5, usando teoria de módulos.

2.3 Espaços de matrizes nilpotentes

Vamos agora apresentar uma primeira demonstração de um resultado de Gerstenhaber [Ge1], que se pode encontrar em [MOR].

Nesta secção, F é um corpo qualquer. O lema que se segue apresenta alguns resultados bem conhecidos, aplicados a um caso particular.

Lema 2.12 *Seja $S \subseteq M_n(F)$ um subespaço vectorial, e ponhamos*

$$S^\perp := \{\, A : \operatorname{tr}(AB) = 0 \ \forall B \in S\}.$$

Temos as seguintes propriedades:

1. $\dim S + \dim S^\perp = n^2$,
2. $R \subseteq S \Leftrightarrow S^\perp \subseteq R^\perp$,
3. $(S^\perp)^\perp = S$,

e se $S = S_1 \oplus S_2$*, então*

4. $S^\perp = S_1^\perp \cap S_2^\perp$.

Demonstração. As afirmações são todas propriedades bem conhecidas, tendo em conta que $\mathrm{tr}(AB)$ define uma forma bilinear no espaço $M_n(F)$. Podem-se encontrar as demonstrações, por exemplo, em [GW, Cap. 5]. □

Lema 2.13 *Se* A*,* B *e* $A+B$ *são matrizes nilpotentes de* $M_n(F)$*, então* $\mathrm{tr}(AB) = 0$.

Demonstração. Suponhamos que B está na sua forma normal de Jordan

$$\begin{bmatrix} 0 & \epsilon_1 & & 0 \\ & 0 & \ddots & \\ & & 0 & \epsilon_{n-1} \\ 0 & & & 0 \end{bmatrix}$$

em que $\epsilon_i = 0$ ou 1, $i \in [n-1]$. Notemos por $s(M)$ a soma dos menores principais de ordem 2 da matriz M. Se M é nilpotente, então $s(M) = 0$, pois é o coeficiente do termo em t^{n-2} do polinómio caractrístico de M, ou o seu simétrico, se $n > 2$, e é o determinante de M se $n = 2$ (para $n = 1$ o resultado é trivial). Sendo $A = [a_{ij}]$, após alguns cálculos simples verificamos que

$$s(A) - s(A+B) = \sum_{i=1}^{n-1} \epsilon_i a_{i+1i} = \mathrm{tr}(AB),$$

e temos assim o pretendido. Para fazer a verificação, basta observar que, dada a forma de B, os únicos menores principais do tipo 2×2 de A que diferem dos respectivos menores de $A + B$ são os do tipo $|A[ii+1|ii+1]|$. Ao fazer a diferença, obtemos

$$|A[ii+1|ii+1]| - (a_{ii}a_{i+1i+1} - (a_{ii+1} + \epsilon_i)a_{i+1i}) = \epsilon_i a_{i+1i},$$

o que termina a demonstarção. □

Teorema 2.14 *Se* $W \subseteq M_n(F)$ *for um espaço de matrizes nilpotentes, então* $\dim(W) \leq n(n-1)/2$.

Demonstração. Seja T o espaço de todas as matrizes triangulares superiores nilpotentes (isto é, de diagonal principal nula), e $W_1 := W \cap T$.

Fixemos W_2 um espaço complementar de W_1 em W, isto é, $W_1 \cap W_2 = \{0\}$ e $W = W_1 + W_2$. Note-se agora que $T^\perp$ é o conjunto das matrizes triangulares superiores, e uma vez que W_2 é constituído por matrizes nilpotentes não triangulares, $W_2 \cap T^\perp = \{0\}$. Tomando agora $A \in W_1$, $B \in W_2$ e $C \in T^\perp$, observamos que $\mathrm{tr}(AC) = 0$ e $\mathrm{tr}(AB) = 0$ pelo lema anterior, e portanto $T^\perp \oplus W_2 \subseteq W_1^\perp$. Pela comparação das dimensões destes espaços, $\dim(W_2) + n(n+1)/2 \leq n^2 - \dim(W_1)$ temos o resultado pretendido. □

Apresentamos a seguir um resultado que diz respeito à possibilidade de triangularização de um espaço de matrizes nilpotentes.

Lema 2.15 *Se $A, B \in M_n(F)$ em que F é um corpo com mais do que dois elementos, e se toda a combinação linear de A e B for uma matriz nilpotente, então* $\mathrm{tr}(AB^2) = 0$.

Demonstração. Tal como no primeiro lema, ponhamos B na sua forma normal de Jordan. Notemos por $s'(M)$ a soma dos menores principais do tipo 3×3 da matriz M, que também é zero para qualquer matriz nilpotente. Temos assim $s'(A + xB) = 0$ para qualquer $x \in F$. Encaremos agora $s'(A+xB)$ como um polinómio em x. Este polinómio é quadrático, pelo facto de B estar na forma normal de Jordan, e anula-se em todos os elementos de F, que por hipótese são pelo menos três, portanto tem que ser o polinómio nulo. Calculemos agora o coeficiente de x^2. Observe-se que para que um menor do tipo 3×3 forneça uma contribuição para o coeficiente de x^2 é necessário que os índices das linhas e das colunas sejam consecutivos; se forem $i, i+1, i+2$, a contribuição é $\epsilon_i \epsilon_{i+1} a_{ii+2}$. Ao somar tudo, obtemos

$$\sum_{i=1}^{n} \epsilon_i \epsilon_{i+1} a_{i+2i} = \mathrm{tr}(AB^2),$$

considerando que B^2 é uma matriz constituida apenas por zeros, excepto nas entradas $(i, i+2)$, que são $\epsilon_i \epsilon_{i+1}$, $i \in [n-2]$. □

Apresentamos agora um teorema de Jacobson. Pode-se encontrar uma generalização interessante deste teorema em [Ra], com uma demonstração bastante simples.

Teorema 2.16 (Jacobson) *Seja $N \subseteq M_n(F)$ um conjunto de matrizes nilpotentes tal que para quaisquer duas matrizes $A, B \in N$ existe $c \in F$ tal que $AB - cBA \in N$. Então N é triangularizável, isto é, existe V, invertível tal que VNV^{-1} é um espaço de matrizes triangulares.*

Teorema 2.17 *Seja $W \subseteq M_n(F)$ um espaço de matrizes nilpotentes, e suponhamos que F tem mais do que dois elementos. Então, se a dimensão de W for $n(n-1)/2$, W é triangularizável.*

Demonstração. Seja $B \in W$ qualquer, e seja V uma matriz invertível tal que $B' := VBV^{-1}$ é triangular. Na notação do Teorema 2.14, definimos também W', W_1' e W_2' como os espaços transformados por conjugação. Dada a hipótese sobre as dimensões, temos $T^\perp \oplus W_2' = W_1'^\perp$, e portanto, pelos resultados do lema 2.12,

$$W_1' = W_2'^\perp \cap T \supseteq W'^\perp \cap T. \tag{6}$$

Temos agora $(B')^2 \in W'^\perp$, pelo lema 2.15, e além disso $(B')^2 \in T$. Assim $(B')^2 \in W_1' \subseteq W'$ e portanto $B^2 \in W$.

Tomando então $C, D \in W$, temos $CD + DC = (C+D)^2 - B^2 - C^2 \in W$, e pelo Teorema de Jacobson, o espaço é triangularizável. □

Terminamos com uma caracterização dos conjuntos de matrizes que geram espaços vectoriais de matrizes nilpotentes. Esta primeira proposição encontra-se demonstrada em [Bo1, pp. A.IV.70].

Proposição 2.18 (Relações de Newton) *Seja m um inteiro, e tomemos $X_1, \ldots, X_m$ uma família de indeterminadas. Sejam s_m o polinómio simétrico elementar de grau m em n variáveis, e p_m a soma das m-ésimas potências das indeterminadas:*

$$s_m = \sum_{H \subseteq [n], |H|=m} \left(\prod_{i \in H} X_i \right) \qquad p_m = \sum_{i=1}^{m} X_i^m.$$

Então

$$(-1)^m m s_m = \sum_{i=i}^{m-1} (-1)^{i-1} s_i p_{m-i} - p_m.$$

Por recusividade, o resultado anterior permite exprimir polinomialmente os polinómios simétricos elementares nas somas de potências das indeterminadas, se o corpo dos coeficientes tiver característica zero. Este facto vai ser usado no próximo teorema.

Teorema 2.19 *Seja $\mathcal{G} \subseteq M_n(F)$ e suponhamos que F tem característica zero. Então as seguintes afirmações são equivalentes:*

(i) O semigrupo aditivo gerado por $\mathcal{G}$ é constituido por matrizes nilpotentes.

(ii) O espaço vectorial gerado por $\mathcal{G}$ é constituido por matrizes nilpotentes.

(iii) Para qualquer família finita de matrizes de $\mathcal{G}$, $(G_1, \ldots, G_k)$,

$$\sum_{\sigma \in \mathcal{S}_k} \operatorname{tr}(G_{\sigma(1)} \ldots G_{\sigma(k)}) = 0.$$

Demonstração. A implicação (ii)⇒(i) é trivial. Para a recíproca, observemos que, na notação do lema anterior, o coeficiente em t^{n-j} do polinómio característico de uma matriz A com valores próprios $(\lambda_1, \ldots, \lambda_n)$ é igual a $s_j(\lambda_1, \ldots, \lambda_n)$, e este pode exprimir-se polinomialmente em $\sum_i \lambda_i^k, k \in [n]$, isto é, nos traços das potências de A. Em particular, A será nilpotente se e só se $\operatorname{tr}(A^m) = 0$ para qualquer inteiro positivo m. Assim, tomando (i) como hipótese, $A = \sum_{i=1}^k a_i G_i$, e fixando m, vamos encarar

$$\operatorname{tr}\left(\left(\sum_{i=1}^k a_i G_i\right)^m\right)$$

como um polinómio em k indeterminadas. Este polinómio anula-se para todas as famílias de k inteiros, portanto é o polinómio nulo e A é nilpotente.

Vamos agora provar a equivalência de (i) e (iii). Supondo que temos (i), seja $(G_i : i \in [k])$ uma família finita de elementos de $\mathcal{G}$. Para cada família $(m_i : i \in [k])$ de inteiros positivos temos por hipótese

$$\operatorname{tr}\left(\left(\sum_{i=1}^k m_i G_i\right)^k\right) = 0.$$

Encarando a expressão como um polinómio (nulo) em k indeterminadas, verificamos que o coeficiente de $m_1 \ldots m_k$ é exactamente

$$\sum_{\sigma \in \mathcal{S}_k} \operatorname{tr}(G_{\sigma(1)} \ldots G_{\sigma(k)}),$$

e obtemos (iii). Para ver a recíproca, tomemos $G_1, \ldots, G_k$ uma família finita qualquer de elementos de $\mathcal{G}$, e provemos que $\operatorname{tr}\left((\sum_i G_i)^m\right) = 0$ para qualquer inteiro m. Tomemos uma família de inteiros $r_1, \ldots, r_k$ não negativos, com $\sum_i r_i = m$ e definimos $B(r_1, \ldots, r_k)$ como sendo a soma dos produtos de

todas as famílias distintas de m matrizes em que r_i dos elementos são iguais a G_i. Obtemos assim

$$r_1! \dots r_k! B(r_1, \dots, r_k) = \sum_{\sigma \in \mathcal{S}_m} A_{\sigma(1)} \dots A_{\sigma(k)},$$

em que $A_1, \dots, A_m$ é uma família em que exactamente r_i dos elementos são iguais a G_i para cada $i \in [k]$. Pela nossa hipótese e pelo facto de F ter característica zero, temos $\operatorname{tr}(B(r_1, \dots, r_k)) = 0$ e observando finalmente que

$$\operatorname{tr}\left(\left(\sum_{i=1}^{k} G_i\right)^m\right) = \operatorname{tr}\left(\sum_{\substack{0 \leq r_i \leq m \\ r_1 + \dots + r_k = m}} B(r_1, \dots, r_k)\right) = 0,$$

concluimos a demonstração. □

□ □
□ □
□ □

Capítulo 3

Usando Teoria de Módulos

`And now, for something completely different.`
(Monty Python)

3.1 Ainda álgebras de matrizes que comutam

Nesta secção, começaremos por seguir [Wa], que surgiu como uma simplificação de algumas técnicas usadas no artigo [BH], que apresenta uma demonstração do teorema 1.15, usando técnicas semelhantes a [Lz]. No artigo [Wa] explora-se uma estrutura de módulo de F^n sobre $F[x]$, definida à custa da matriz A, para demonstrar o teorema 1.15. São estes resultados de teoria de módulos que aqui desenvolvemos.

Definição 3.1 *Seja R um anel. Se M for um módulo sobre R, diremos que M é um* módulo de torção *se, para cada $x \in M$ existir um $r \in R$ tal que $rx = 0$.*

A partir daqui, nesta secção suporemos que R é um domínio de ideais principais. Apresentamos aqui o teorema dos factores invariantes para módulos de torção finitamente gerados sobre domínios de ideais principais (que se pode encontrar, por exemplo, em [Bl]).

Proposição 3.2 *Sejam M e R nas condições anteriores. Então M é isomorfo a uma soma directa de módulos*

$$M \simeq \bigoplus_{i=1}^{t} R/Rh_i,$$

em que os ideais $Rh_1, \ldots Rh_t$ *estão univocamente determinados, verificam* $(0) \subset Rh_1 = \mathrm{An}_R(M) \subseteq \ldots \subseteq Rh_t \subset R$ *(isto é* $h_t \mid \ldots \mid h_1$*), e são chamados os* ideais factores invariantes *de* M*, em que notamos por* $\mathrm{An}_R(M)$ *o ideal aniquilador de* M*. Aos elementos* $(h_1, \ldots, h_t)$ *geradores destes ideais chamamos* factores invariantes *de* M*. Ao produto* $h_1 \times \ldots \times h_t$ *chamamos* ordem de M*, notada* $\mathrm{ord}(M)$*.*

Observação. Os elementos $h_1, \ldots, h_t$ e $\mathrm{ord}(M)$ estão bem determinados, a menos de produto por unidades de R. Quando falarmos, por abuso de linguagem, *dos* factores invariantes de M, estamos a pensar numa família completa de geradores dos ideais factores invariantes.

A partir de agora vamos considerar que $M = Rw_1 \oplus \ldots \oplus Rw_t$, para alguns $w_i \in M$, com $\mathrm{An}_R(Rw_i) = (h_i)$, e $h_t \mid \ldots \mid h_1$, pois os resultados que se seguem (lemas 3.3 e 3.4 e teorema 3.5) mantêm-se por isomorfismo. Vamos notar por $\mathrm{End}_R(M)$ o anel dos R-endomorfismos de M. Fixando b em $\mathrm{End}_R(M)$, vamos notar por T a R-subálgebra de $\mathrm{End}_R(M)$ gerado por B, isto é, o R-submódulo de $\mathrm{End}_R(M)$ gerado por $\{id, b, b^2, \ldots\}$. Apresentamos agora alguns resultados técnicos que dizem respeito à estrutura de T.

Lema 3.3 *Seja* t *o número de ideais factores invariantes de* M*. Então existe um polinómio mónico* $f \in R[x]$ *de grau* t *tal que* $f(b) = 0$*. Portanto,* T *é gerado como* R*-módulo por* $\{id, b, b^2, \ldots b^{t-1}\}$*.*

Demonstração. Sejam $z_1, \ldots, z_t$ elementos[1] de um R-módulo, tais que, para todo o i, $\mathrm{An}_R(z_i) = (h_1)$. Consideremos agora a seguinte aplicação

$$\begin{array}{ccc} M & \hookrightarrow & \displaystyle\bigoplus_{i=1}^{t} Rz_i \\ \displaystyle\sum_{i=1}^{t} a_i w_i & \mapsto & \displaystyle\sum_{i=1}^{t} (h_1/h_i) a_i z_i. \end{array}$$

É simples de verificar que está bem definida e que é um morfismo injectivo. Pondo $N := \bigoplus_{i=1}^{t} Rz_i$, vamos identificar então M com a sua imagem, isto é, considerar $M \subseteq N$.

Passamos a mostrar que existe $d \in \mathrm{End}_R(N)$ com $d|_M = b$. Seja, para cada w_i, $b(w_i) = \sum_j s_{ij} z_j$, com $s_{ij} \in R$. Ora, como $h_i w_i = 0$, temos

[1] De facto, podiam-se tomar todos iguais a w_1, mas é mais cómodo, por questões de notação, tomá-los assim.

$0 = h_i b(w_i) = \sum_j h_i s_{ij} z_j$, portanto $h_i s_{ij} \in \mathrm{An}_R(z_j) = (h_1)$. Sejam então $h_i s_{ij} = h_1 t_{ij}$, e portanto $s_{ij} = (h_1/h_i) t_{ij}$. Definimos então $d \in \mathrm{End}_R(N)$ por $d(z_i) = \sum_j t_{ij} z_j$. Recordando que $w_i = (h_1/h_i) z_i$ pela identificação, temos

$$d(w_i) = \sum_j (h_1/h_i) t_{ij} z_j = \sum_j s_{ij} z_j = b(w_i).$$

Temos também que N é um $R/(h_1)$-módulo livre, e portanto podemos aplicar o teorema de Hamilton-Cayley para anéis comutativos com identidade (cf. por exemplo [Br]) e obter que $\phi(b) = 0$ em que $\phi \in R/(h_1)[x]$ é o polinómio característico de b. Seja então f um polinómio mónico de $R[x]$ cujos coeficientes sejam pré-imagens dos coeficientes de ϕ. Então, f tem grau t, e

$$f(b) = \phi(b) = \phi(d|_M) = 0.$$

Isto conclui a demonstração. □

Lema 3.4 *Sejam R, M, b e T como acima, e suponhamos que M se decompõe como soma directa de dois submódulos, $M = P \oplus Q$, e mais ainda que $hQ = 0$, para algum $h \in R$. Seja $c \in \mathrm{End}_R(P)$ definido por*

$$c : p \mapsto \pi_1(b(p, 0)),$$

em que π_1 é a primeira projecção, com respeito à decomposição apresentada. Então, fazendo $S := <id, c, c^2, \ldots>_R$ a R-subálgebra de $\mathrm{End}_R(P)$ gerada por c, temos que $hT \simeq_R hS$, e o R-isomorfismo aplica hb^i em hc^i, para cada i.

Demonstração. Cada $d \in \mathrm{End}_R(M)$ pode exprimir-se como uma matriz

$$\begin{bmatrix} d_{pp} & d_{qp} \\ d_{pq} & d_{qq} \end{bmatrix},$$

com $d_{pp} \in \mathrm{End}_R(P)$, $d_{qp} \in \mathrm{Hom}_R(Q, P)$, $d_{pq} \in \mathrm{Hom}_R(P, Q)$, $d_{qq} \in \mathrm{End}_R(Q)$ e, para $a \in P$, $d_{pp}(a) = \pi_1(d(a, 0))$. Considere-se então

$$\begin{array}{rccc} g : \mathrm{End}_R(M) & \to & \mathrm{End}_R(P) \\ d & \mapsto & d_{pp}. \end{array}$$

A aplicação g é R-linear, mas não é um morfismo de anéis, pois não respeita a multiplicação. Note-se porém que, como $hQ = 0$,

$$hd = \begin{bmatrix} hd_{pp} & hd_{qp} \\ hd_{pq} & hd_{qq} \end{bmatrix} = \begin{bmatrix} hd_{pp} & 0 \\ 0 & 0 \end{bmatrix} \qquad (1)$$

e pondo

$$e = \begin{bmatrix} e_{pp} & e_{qp} \\ e_{pq} & e_{qq} \end{bmatrix},$$

vem $g(de) = d_{pp}e_{pp} + d_{qp}e_{pq}$ e obtém-se

$$g(hde) = hg(de) = hd_{pp}e_{pp} + hd_{qp}e_{pq} = hd_{pp}e_{pp} = hg(d)g(e).$$

Como c foi definida com $g(b) = c$, verifica-se por indução, usando a igualdade anterior, que $g(hb^i) = hc^i$ para todo o i. Portanto, g aplica hT em hS sobrejectivamente; e é também injectiva, pois pela equação (1) podemos ver que g é injectiva em $h\,\mathrm{End}_R(M)$. Isto termina a demonstração. □

Teorema 3.5 *Para R, M, b e T como acima, sendo $h_1, \ldots, h_t$ os factores invariantes de M, e $g_1, \ldots, g_k$ os factores invariantes de T, temos que $k \leq t$ e $g_i \mid h_i$ para $i \in [k]$. Daqui se pode concluir que T é isomorfo a um R-submódulo de M.*

Demonstração. O número k de factores invariantes de T é no máximo t, pois, pelo lema 3.3, T tem um conjunto gerador com t elementos[2]. Além disso, $(g_1) = \mathrm{An}_R(T) = \mathrm{An}_R(M) = (h_1)$, pois $id \in T$. Por outro lado, os factores invariantes de hT são $g'_1, \ldots, g'_m$, em que, para cada i, $g'_i = g_j/\operatorname{mdc}(h, g_i)$, para aqueles valores de i tais que $g_j \nmid h$. Fixemos agora $i \geq 1$ e vejamos que $g_{i+1} \mid h_{i+1}$. Aplicando o lema 3.4 com $P := Rw_1 \oplus \ldots \oplus Rw_i$, $Q := Rw_{i+1} \oplus \ldots \oplus Rw_t$ e $h := h_i$, pondo c e S como no lema, obtemos $hT \simeq hS$. Ora, hS tem no máximo i factores invariantes, pois, pelo lema 3.3 o conjunto $\{h.id, hc, \ldots, hc^{i-1}\}$ é gerador de hT e tem i elementos. Isto, conjugado com a expressão que já tínhamos para os factores invariantes de hT, mostra que T tem no máximo i factores invariantes que não dividem h_{i+1}, e portanto $g_{i+1} \mid h_{i+1}$. Isto é válido para todo o $i \geq 1$, e mostrámos o pretendido.

Vejamos finalmente que M vem isomorfo a um submódulo de T. Tomemos $T = Ry_1 \oplus \ldots \oplus Ry_t$, pondo $y_i := 0$ e $g_i := 1$ para $k+1 \leq i \leq t$. Sejam agora, para $1 \leq i \leq t$, $s_i := h_i/g_i$. Para obter o resultado, basta ver que a aplicação de T em M definida por $\sum_i r_i y_i \mapsto \sum s_i r_i w_i$ está bem definida e é um morfismo injectivo, o que é uma verificação trivial. □

[2] O número de factores invariantes é também o número mínimo de elementos necessário para gerar o módulo.

Corolário 3.6 *Pondo* $T_i := <id, b, b^2, \ldots, b^{i-1}>_R$, *temos* $h_i b^{i-1} \in T_{i-1}$, *para* $i \in [t]$.

Demonstração. Seguindo o percurso da demonstração anterior, tínhamos obtido que $c^{i-1} \in <id, c, \ldots, c^{i-2}>_R$, por aplicação do lema 3.3 a c e a P. Portanto, $h_i c^{i-1} \in <h_i id, h_i c, \ldots, h_i c^{i-2}>_R$. Agora, pelo isomorfismo do lema 3.4, temos $h_i b^{i-1} \in <h_i id, h_i b, \ldots, h_i b^{i-2}>_R \subseteq T_{i-1}$, o que conclui a demonstração. □

Vamos poder agora deduzir o resultado 1.15 a partir deste teorema.

Corolário 3.7 *Sejam* A *e* B *matrizes do tipo* $n \times n$ *que comutam. Então* $\dim \operatorname{alg} <A, B> \leq n$.

Demonstração. Começamos por pôr $M := M_{n\times 1}(F)$ e $R := F[x]$. Vamos agora criar em M uma estrutura de R-módulo, através de A: para $p \in F[x]$ e $m \in M$, define-se $p.m := p(A)m$. Então R pode identificar-se com $F[A]$, subálgebra de $M_n(F)$. Vem assim que $\operatorname{End}_R(M)$ é o centralizador de A, que contém B. Pondo agora $b := B$, o conjunto T dos teoremas anteriores é a F-subálgebra de $M_n(F)$ gerada por A e por B. Nestas condições, um R-módulo é também um F-espaço vectorial, um R-submódulo um F-subespaço vectorial, e um R-morfismo uma aplicação F-linear. Assim, obtemos que T é F-isomorfo a um subespaço vectorial de M, e portanto

$$\dim_F \operatorname{alg} <A, B>_F \leq \dim_F M = n,$$

que é o pretendido. □

Vamos finalmente apresentar um conjunto gerador, semelhante ao apresentado no teorema 2.4. A partir de agora, tomamos R e M como acabámos de definir, e supomos os polinómios h_i mónicos, para evitar ambiguidades.

Proposição 3.8 *Para* $i \in [t]$, *seja* $n_i = \operatorname{gr}(h_i)$. *Temos então que o conjunto*

$$\begin{aligned} &\{I, A, A^2, \ldots, A^{n_1-1}\} \cup \\ &\quad \cup \{B, AB, A^2B, \ldots, A^{n_2-1}B\} \cup \\ &\quad \cup \ldots \cup \\ &\quad \cup \{B^{t-1}, AB^{t-1}, \ldots, A^{n_t-1}B^{t-1}\} = \\ &= \{\, A^i B^j : 0 \leq j \leq t-1 \text{ e, para cada j, } 0 \leq i \leq n_{j+1} - 1 \,\} \end{aligned}$$

gera T *como espaço vectorial, isto é, contém uma base.*

Demonstração. Vamos aplicar o resultado do corolário 3.6. No nosso caso

$$T_i = <id, b, b^2, \ldots, b^{i-1}>_R = F[A] + F[A]B + F[A]B^2 + \ldots + F[A]B^{i-1},$$

e como $\mathrm{gr}(h_i) = n_i$, este resultado leva a concluir que

$$A^{n_i}B^{i-1} \in T_{i-1} + <B^{i-1}, AB^{i-1}, \ldots, A^{n_i-1}B^{i-1}>_F.$$

Multiplicando então $A^{n_i}B^{i-1}$ por A, repetidamente, concluimos que, para todo o j, $A^jB^{i-1} \in T_{i-1} + <B^{i-1}, AB^{i-1}, \ldots, A^{n_i-1}B^{i-1}>_F$, e portanto $T_i = T_{i-1} + <B^{i-1}, AB^{i-1}, \ldots, A^{n_i-1}B^{i-1}>_F$. Posto que $T = T_t$, pelo lema 3.3, o resultado fica demonstrado. □

Note-se que no capítulo anterior apresentámos primeiro o conjunto gerador (no teorema 2.4) e depois é que tirámos a conclusão sobre a dimensão, ao passo que aqui os resultados sairam paralelamente, à custa da demonstração do teorema 3.5. Além disso, essa base era construida à custa do estudo dos divisores elementares, e esta usou os factores invariantes. Em qualquer dos casos, obteve-se n como limite superior, pois a soma dos expoentes, quer dos factores invariantes, quer dos divisores elementares, é n.

3.2 Álgebras comutativas maximais

Vamos agora estudar a dimensão de álgebras comutativas maximais, isto é, subálgebras comutativas $R \subseteq M_n(F)$ tais que, se $S \supseteq R$, para S subálgebra de $M_n(F)$, então S não é comutativa. Uma condição equivalente é $\mathcal{C}(R) \subseteq R$. Para isso, vamos impor algumas condições sobre o radical da álgebra, que é o conjunto dos seus elementos nilpotentes (e também é uma álgebra) nomeadamente sobre o seu expoente, isto é, o número natural k tal que $(\mathrm{rad}\, R)^k = 0$ e $(\mathrm{rad}\, R)^{k-1} \neq 0$. Laffey [Lf] obteve o limite inferior

$$\dim R > (2n)^{2/3} - 1,$$

e, para o caso de o expoente de R ser três,

$$\dim R \geq \lceil 3n^{2/3} - 4 \rceil,$$

e este último limite é o melhor possível para um número infinito de valores de n. Quanto a limites superiores, existe um teorema de Schur (que se pode encontrar, por exemplo, em [ST]) afirma que

$$\dim R_n \leq \lfloor n^2/4 \rfloor + 1,$$

e, para o caso em que o expoente é três, Courter [Co] constrói uma sucessão de álgebras comutativas maximais, R_n, todas de expoente três, tais que $\lim(\dim R_n)/n = 0$. Vamos aqui seguir novamente [Lz], em que se faz uma adaptação dos raciocínios de Courter para o caso do expoente quatro. Assim, os resultados que aqui apresentamos destinam-se a serem aplicados no exemplo que iremos construir: vamos apresentar, para cada $\epsilon > 0$, uma álgebra comutativa maximal de $M_k(F)$, S_ϵ, tal que $\dim S_\epsilon/k < \epsilon$.

A construção geral

Tomemos então, para já, R anel comutativo com identidade, e M módulo sobre R. Escreveremos $\mathrm{Hom}_R(M) = R$ para afirmar que todo o R-endomorfismo de M é da forma $x \mapsto xr$, para algum $r \in R$. Denotaremos este endomorfismo por r^*. Vamos também considerar que R admite uma representação fiel sobre M, espaço vectorial sobre F de dimensão n. Vamos considerar também que $R \subseteq M_n(F)$, e fixamos uma base em M, de modo que cada elemento de $M_n(F)$ define uma aplicação de M em M.

Proposição 3.9 *Com R nas condições anteriores, $R = \mathrm{Hom}_R(M)$ se e só se R é uma álgebra comutativa maximal de $M_n(F)$.*

Demonstração. Suponhamos que $R = \mathrm{Hom}_R(M)$. Tomemos então $t \in M_n(F)$ com $t \in \mathcal{C}(R)$. Antes de mais, t define uma aplicação de M em M pelo facto de M ser espaço vectorial sobre F. Além disso, como $tr = rt$ para todo o $r \in R$, temos $t \in \mathrm{Hom}_R(M) = R$, e assim $\mathcal{C}(R) \subseteq R$, e isto equivale a afirmar que R é subálgebra comutativa maximal de $M_n(F)$.

Reciprocamente, supondo que R é subálgebra comutativa maximal de $M_n(F)$, vejamos que $R = \mathrm{Hom}_R(M)$. Imediatamente, $R \subseteq \mathrm{Hom}_R(M)$ por ser comutativo. Tomando $s \in \mathrm{Hom}_R(M) \subseteq M_n(F)$, obtemos, para quaisquer $x \in M$ e $r \in R$, $(xr)s = (x)sr$ e portanto, pela fidelidade da representação, $sr = rs$, $s \in \mathcal{C}(R) \subseteq R$, o que prova o pretendido. □

A partir de agora apenas usaremos módulos e anéis nas condições mencionadas, e portanto para ver que R é subálgebra comutativa maximal de $M_n(F)$ verificaremos se $\mathrm{Hom}_R(M) = R$.

Definição 3.10 *Considerem-se $x_1, \ldots, x_m,$, $y_1, \ldots, y_d$ elementos de M e $(r_{ij} : i \in [m], j \in [d])$ uma família de elementos de R. Dizemos que a família (r_{ij}) é* densa *para o par $((x_1, \ldots, x_m), (y_1, \ldots, y_d))$ se*

$$x_\sigma r_{ij} = \delta_{\sigma i} y_j.$$

Dizemos que R é denso *para o par se existir uma família (r_{ij}) de elementos de R que seja densa para o par.*

Proposição 3.11 *Se R é denso para o par $((x_i : i \in [m]), (y_j : j \in [d]))$ então*

$$\mathrm{Hom}_R(\sum_{i=1}^{m} x_i R, \sum_{j=1}^{d} y_j R) = R.$$

Demonstração. Como R é anel comutativo, temos imediatamente $R \subseteq \mathrm{Hom}_R(\sum x_i R, \sum y_j R)$. Para vermos a outra inclusão, vamos considerar $f \in \mathrm{Hom}_R(\sum x_i R, \sum y_j R)$ qualquer, e obtemos $(a_{ij} : i \in [m], j \in [d])$ uma família de elementos de R tal que, para todo o i e j,

$$f(x_i) = \sum_{i=1}^{d} y_i a_{ij}.$$

Seja (r_{ij}) uma família densa para o par $((x_i), (y_j))$. Pomos agora $s := \sum_{ij} r_{ij} a_{ij} \in R$, e vejamos que, para qualquer $x \in M$, $f(x) = xs$. Para tal, vejamos que os morfismos f e s^* coincidem sobre os elementos x_i. Sendo então x_t um desses elementos, temos

$$s^*(x_t) = \sum_{ij} x_t r_{ij} a_{ij} = \sum_{ij} \delta_{ti} y_j a_{ij} = \sum_{j} y_j a_{tj} = f(x_t),$$

o que termina a demonstração. □

Teorema 3.12 *Sejam R e M como acima, e N um R-submódulo de M. Sejam $x_1, \ldots, x_m$, $y_1, \ldots, y_d$ elementos de M tais que $M = \sum x_i R$ e $N = \sum y_j R$. Suponhamos que existe (r_{ij}), uma família de elementos de R densa para o par $((x_i), (y_j))$ e mais ainda, que para qualquer $t \in R$ tal que $y_1 t = y_2 t = \ldots = y_d t = 0$, se tem $Mt \subseteq N$. Nestas condições* $\mathrm{Hom}_R(M) = R$.

Demonstração. Tomemos $f \in \mathrm{Hom}_R(M)$, com $f(x_i) = \sum_j x_j c_{ij}$. Para $\sigma = 1, \ldots, d$ e $j = 2, \ldots, m$,

$$\begin{aligned} f(y_\sigma) &= f(x_1 r_{1\sigma}) = f(x_j r_{j\sigma}) \\ &= \left(\sum_i x_i c_{1i}\right) r_{1\sigma} = \left(\sum_i x_i c_{ji}\right) r_{j\sigma} \\ &= \left(\sum_i x_i r_{1\sigma} c_{1i}\right) = \left(\sum_i x_i r_{j\sigma} c_{ji}\right) \\ &= y_\sigma c_{11} = y_\sigma c_{jj}, \end{aligned}$$

donde podemos concluir que $y_\sigma(c_{11} - c_{jj}) = 0$.

Por outro lado, para $k \neq l$ e $\sigma \in [d]$, temos $x_k r_{l\sigma} = 0$ e portanto

$$0 = f(x_k)r_{l\sigma} = \sum_i x_i r_{l\sigma} c_{ki} = x_l r_{l\sigma} c_{kl} = y_\sigma c_{kl},$$

e portanto $y_\sigma c_{kl} = 0$. Assim, pela hipótese do teorema,

$$Mc_{kl} \subseteq N \; k, l \in [d], k \neq l \text{ e } M(c_{11} - c_{jj}) \subseteq N, \; j \in [m].$$

Pondo então $c := c_{11}$ e $g := f - c^*$, temos $g \in \mathrm{Hom}_R(M, N)$:

$$g(x) = f - c^*(x) = \sum_j xc_{ij} - xc_{11} = \sum_{j \neq i} xc_{ij} - x(c_{11} - c_{ii}) \in N,$$

para $x \in M$ e $i \in [m]$. Então, pela proposição 3.11 temos $g = b^*$ para algum $b \in R$, e portanto $f = (b + c)^*$, o que prova o pretendido. □

Vamos agora construir os módulos M_q e os anéis R_q que vão servir para o nosso exemplo. Seja M um módulo sobre um anel R, com dois submódulos $M \supseteq M' \supseteq M'' \supseteq (0)$. Para cada $q \in \mathbb{N}$, vamos construir um módulo M_q da seguinte forma: tomamos todos os q-uplos de elementos de M, $(x_1, \ldots, x_q)$ com $x_i \equiv x_j \bmod M'$, para todos os $i, j \in [q]$ e identificamos os elementos $(x_1, \ldots, x_q)$ e $(y_1, \ldots, y_q)$ se $\sum x_i = \sum y_i$ e $x_i \equiv y_i \bmod M''$ para $i \in [q]$. A adição em M_q e o produto por um elemento de R definem-se componente a componente, e é simples de verificar que ficam bem definidas. Se N for um submódulo de M, notaremos por N_q o submódulo de M_q em que as componentes dos q-uplos são escolhidas em N. Seguem-se algumas propriedades simples destes módulos.

Lema 3.13 *Com as definições anteriores, temos:*

1. $M''_q \simeq M''$,
2. $M_q/M'_q \simeq M/M'$,
3. $M'_q/M''_q \simeq (M'/M'')^{(q)}$ *(a q-ésima potência directa).*

Demonstração. A afirmação 1 é de verificação trivial, se pensarmos na aplicação de M'' em M''_q definida por $x \mapsto (x, 0, \ldots, 0)$. Quanto à afirmação 2, observamos que os elementos de M_q podem ser postos na forma

$$\tilde{x} = (x_1, x_1 + x'_2, \ldots, x_1 + x'_q), \qquad x'_i \in M', 2 \leq i \leq q.$$

Isto permite definir um epimorfismo de M_q em M/M', pondo $\tilde{x} \mapsto x_1$, e obter o resultado pelo teorema do homomorfismo, observando que o núcleo é M'_q. Finalmente, para a afirmação 3, basta considerar o epimorfismo de M'_q na q-ésima potência directa de M'/M'' definido por $x \mapsto (x_1+M'', \ldots, x_q+M'')$ e verificar que o núcleo é M''_q. □

Definição 3.14 *Sejam M um módulo sobre o anel R, nas condições descritas. Vamos dizer que a* **Condição (A)** *é satisfeita se existirem submódulos de M, M' e M'', e ideais de R, R' e R'' tais que:*

- $R \supseteq R' \supseteq R'' \supseteq (0)$,
- $M \supseteq M' \supseteq M'' \supseteq (0)$,
- $MR' \subseteq M'$, $MR'' \subseteq M''$, $M'R'' = M''R' = (0)$ *e*
- $R'R'' = R''R' = (0)$.

Supondo válida a condição **(A)**, é simples de ver que o R-módulo R_q obtido a partir de R, considerado como módulo sobre si mesmo tem a estrutura de anel, se dotado da mutiplicação componente a componente, que é comutativo se R o for. Mais, M_q tem também uma estrutura de R_q-módulo, com a multiplicação definida também componente a componente: para $r = (r_1, \ldots, r_q) \in R_q$ e $m = (m_1, \ldots, m_q) \in M_q$, escrevendo, para cada i, $m_i = a + x_i$, com $x_i \in M'$ e $r_i = b + y_i$, com $y_i \in R'$, temos $mr = (m_i r_i : i \in [q])$ $m_i r_i = ab + ay_i + x_i(b + y_i)$ e $ay_i + x_i(b + y_i) \in M'$ porque $MR' \subseteq M'$. De forma análoga se via que $mr = 0$ se $r = 0$ ou se $m = 0$, usando outras propriedades incluidas na condição **(A)**. Se N é submódulo de M, N_q vem R_q-submódulo de M_q. Os cálculos envolvidos nestas verificações são um pouco morosos, mas são todos bastante simples.

Introduzimos agora alguma notação. Para $x \in M$ ou R, vamos notar por $\overline{x}$ o elemento de M_q ou de R_q cujas componentes são todas iguais a x. Para $y \in R'$ ou M', e $j \in [q]$, notaremos por $y^{(j)}$ o elemento de M_q ou R_q que tem y na sua j-ésima componente, e zero nas outras. Observe-se que para que este elemento pertença a M_q ou R_q (conforme o caso) é preciso exigir que y pertença a M' ou R'. Finalmente, se R for uma álgebra sobre um corpo F, então, por identificação, $F \subseteq R$ e os isomorfismos mencionados no lema 3.13 são F-isomorfismos.

Observação. Suponhamos que temos a condição **(A)**, e que (r_{ij}) é uma família de elementos de R' densa para o par $((x_i : i \in [m]), (y_j : j \in [d]))$,

com $x_i \in M$, $y_j \in M'$. Então

$$r_{ij}^{(\tau)} \overline{x}_\sigma = \delta_{\sigma i} y_j^{(\tau)},$$

o que mostra que R_q é denso para o par

$$((\overline{x}_i : i \in [m]), (y_1^{(1)}, \ldots, y_d^{(1)}, y_1^{(2)}, \ldots, y_d^{(2)}, \ldots, y_1^{(q)}, \ldots, y_d^{(q)})).$$

Teorema 3.15 *Seja R uma álgebra comutativa sobre um corpo F, M o espaço de representação de R, e suponhamos que a condição* **(A)** *é satisfeita e que R e M satisfazem as condições do teorema 3.12, com os referidos elementos r_{ij} pertencendo a R', e y_j pertencendo a M', para cada i, j (como na observação anterior). Suponhamos ainda que $M' = MR'$. Então*

$$\mathrm{Hom}_{R_q}(M_q) = R_q.$$

Demonstração. Vamos verificar que M_q e N_q satisfazem as hipóteses do teorema 3.12. Seja então $h = (h_1, \ldots, h_q) \in R_q$ tal que $y_i^{(\tau)} h = (0, \ldots, 0)$, com $\tau \in [q]$ e $i \in [d]$. Assim, temos, para cada τ e i, $y_i h_\tau = 0$. Ora, por hipótese (cf. teorema 3.12), nestas condições $Mh_\tau \subseteq N$, e portanto $M_q h \subseteq N_q$.

Vejamos agora a densidade de R_q para os conjuntos geradores. Vejamos primeiro que M_q é gerado como módulo sobre R_q pelos elementos $\overline{x}_1, \ldots, \overline{x}_m$. Seja $v \in M_q$, $v = (a+b_1, \ldots, a+b_q)$ em que $a \in M$ e cada $b_i \in M'$, e portanto $b_i = \sum_j x_j s_{ij}$ com $s_{ij} \in R'$. Temos então $b_i^{(i)} = \sum \overline{x}_j s_{ij}^{(i)}$. Por outro lado, se suposermos que $a = \sum x_j t_j$, temos evidentemente $\overline{a} = \sum \overline{x}_j \overline{t}_j$. Isto prova o pretendido, uma vez que

$$v = \overline{a} + (b_1, \ldots, b_q) = \overline{a} + b_1^{(1)} + \ldots + b_q^{(q)}.$$

Vendo agora que N_q é gerado pelos elementos $(y_i^{(\tau)} : i \in [d], \tau \in [q])$, a observação anterior prova a densidade pretendida. Isto termina a nossa demonstração. □

Teorema 3.16 *Seja R uma álgebra comutativa sobre o corpo F, com radical P. Suponhamos que $P^4 = 0$ mas $P^3 \neq 0$. Seja M o espaço de representação de R. Ponhamos $R' := P$, $R'' := P^3$, $M' := MP$ e $M'' := MP^3$ (neste caso a condição* **(A)** *é válida). Suponhamos que M_q é um R_q-módulo fiel e que $\mathrm{Hom}_{R_q}(M_q) = R_q$, para cada q. Então o expoente de R_q é quatro e se R e M tiverem dimensão finita sobre F, temos*

1. *R_q é uma subálgebra comutativa maximal de $M_{k(q)}(F)$, em que $k(q)$ é a dimensão de M_q sobre F e*

2. $\lim_{q\to\infty} \dfrac{\dim R_q}{k(q)} = \dfrac{\dim(P/P^3)}{\dim(MP/MP^3)}$.

Demonstração. Para ver que o expoente de R_q é quatro basta observar que o seu radical é P_q, que tem índice de nilpotência quatro. A afirmação (1) é também imediata, em vista da proposição 3.9. Quanto à afirmação (2), tendo em conta os isomorfismos do lema 3.13, e que $k(q) = \dim M_q$, podemos escrever:

$$\begin{aligned}\frac{\dim R_q}{\dim M_q} &= \frac{\dim(R_q/R'_q) + \dim(R'_q/R''_q) + \dim R''_q}{\dim(M_q/M'_q) + \dim(M'_q/M''_q) + \dim M''_q} \\ &= \frac{\dim(R/R') + q \times \dim(R'/R'') + \dim R''}{\dim(M/M') + q \times \dim(M'/M'') + \dim M''},\end{aligned}$$

e portanto

$$\lim_{q\to\infty} \frac{\dim R_q}{\dim M_q} = \frac{\dim(R'/R'')}{\dim(M'/M'')},$$

que é o resultado pretendido. □

O Exemplo

Vamos agora construir a álgebra inicial $R_1 = R$. Sejam s um inteiro, $s \geq 3$ e $n := s^3$. Sejam A a subálgebra de $M_s(F)$ gerada pelas matrizes $I_s, E_{12}, \ldots, E_{1s}$, e B a subálgebra de $M_s(F)$ gerada pelas transpostas destas matrizes.

Estudemos um pouco a álgebra A (a álgebra B tem evidentemente um estudo análogo). Para começar, vejamos que é comutativa maximal. As matrizes geradoras apresentadas são linearmente independentes e verificam $E_{1i}E_{1j} = 0_s$ (portanto comutam), para quaisquer $2 \leq i, j \leq s$. Temos portanto a propriedade

$$A = \text{alg} <I_s, E_{12}, \ldots, E_{1s}> = <I_s, E_{12}, \ldots, E_{1s}>.$$

Para ver a maximalidade, verifiquemos que $\mathcal{C}(A) \subseteq A$. Tomemos então $X = [x_{ij}] \in \mathcal{C}(A)$. Das $n-1$ equações do tipo $E_{1i}X = XE_{1i}$, $i \geq 2$, podemos deduzir as relações

$$x_{ii} = x_{11},\ x_{ij} = 0, \quad \text{para } i \geq 2 \text{ e } j \neq i.$$

Daqui se pode concluir imediatamente que X tem que ser combinação linear das matrizes apresentadas, e portanto, $X \in A$, como pretendíamos.

Pomos agora

$$R := A \otimes B \otimes A.$$

É simples de verificar que o radical de R, P, é gerado, como espaço vectorial, pelo conjunto

$$\begin{aligned}\{\, f \otimes g \otimes h : (f,g,h) \neq (I,I,I);\\ f,h = I, E_{12}, \ldots, E_{1s};\, g = I, E_{21}, \ldots, E_{s1}\,\},\end{aligned} \tag{2}$$

tendo portanto dimensão $n-1$ sobre F. O espaço P^2 admite como base o conjunto

$$\begin{aligned}\{\, f \otimes g \otimes h : \text{não mais do que um } f,g,h = I;\\ f,h = I, E_{12}, \ldots, E_{1s};\, g = I, E_{21}, \ldots, E_{s1},\,\}\end{aligned} \tag{3}$$

e P^3 admite como base o conjunto

$$\{\, f \otimes g \otimes h : f,h = E_{12}, \ldots, E_{1s}, g = E_{21}, \ldots, E_{s1}\,\}. \tag{4}$$

Daqui se pode concluir imediatamente que R tem expoente quatro, e que

$$\dim(P/P^3) = \dim P - \dim P^3 = s^3 - 1 - (s-1)^3 = 3s^2 - 3s. \tag{5}$$

Sejam agora U um espaço de representação de A, com base $(u_1, \ldots, u_s)$ e V o de B, com base $(v_1, \ldots, v_s)$. Então $M := U \otimes V \otimes U$ é um espaço de representação de R, de dimensão n. Tomamos M', M'', R' e R'' como no teorema 3.16. Verifiquemos agora que M e R satisfazem as condições dos teoremas 3.12, 3.15 e 3.16.

Notemos para já que $U = u_1 A$, pois $u_{i+1} = u_1 E_{1i+1}$, $i \in [s-1]$, e $V = \sum_{i=2}^{s} v_i B$, pois $v_1 = v_i E_{i1}$ para qualquer $i \geq 2$. Assim, M é gerado sobre R pelo conjunto

$$\{\, u_1 \otimes v_i \otimes u_1 : i = 2, \ldots, s\,\}. \tag{6}$$

Pomos então $x_i := u_1 \otimes v_{i+1} \otimes u_1$, $i \in [s-1]$ e com $y := u_1 \otimes v_1 \otimes u_1$ e $N := yR$, obtemos imediatamente que P é denso para o par $((x_1, \ldots, x_m), (y))$ (com $m := s-1$) uma vez que

$$x_i(I \otimes E_{j+1i} \otimes I) = u_1 \otimes \delta_{ij} v_1 \otimes u_1 = \delta_{ij} u_1 \otimes v_1 \otimes u_1 = \delta_{ij} y.$$

Tomemos agora $t \in R$ tal que $yt = 0$ e verifiquemos que $Mt \subseteq N$. Note-se antes de mais que os elementos

$$\{ f \otimes g \otimes h : f, h = I, E_{12}, \ldots, E_{1s};\ g = E_{21}, \ldots, E_{s1} \} \tag{7}$$

aniquilam y e aplicam M em N (é uma verificação simples, tendo em conta o conjunto gerador (6)). Considerando agora t como combinação linear dos elementos do conjunto gerador de R,

$$\begin{aligned} t &= a_1(I \otimes I \otimes I) + \sum_{i=2}^{s} a_i(I \otimes I \otimes E_{1i}) + \\ &\quad \sum_{i=2}^{s} b_i(E_{1i} \otimes I \otimes I) + \sum_{i,j=2}^{s} c_{ij}(E_{1i} \otimes I \otimes E_{1j}) + t' \end{aligned}$$

em que t' é uma combinação linear de elementos do conjunto (7). Assim

$$\begin{aligned} 0 = yt &= a_1(u_1 \otimes v_1 \otimes u_1) + \sum_{i=2}^{s} a_i(u_1 \otimes v_1 \otimes u_i) + \\ &\quad \sum_{i=2}^{s} b_i(u_i \otimes v_1 \otimes u_1) + \sum_{i,j=2}^{s} c_{ij}(u_i \otimes v_1 \otimes u_j), \end{aligned}$$

e como estes elementos são linearmente independentes, vem

$$a_1 = \ldots = a_s = b_2 = \ldots = b_s = d_{22} = d_{23} = \ldots = d_{ss} = 0,$$

e portanto $t = t'$, t está no espaço gerado pelo conjunto (7), e pelo que já observamos, $Mt \subseteq N$. Assim, acabámos de ver que se verificam as condições do teorema 3.12.

Passemos agora a verificar as condições do teorema 3.15. Pela construção de R, temos claramente a condição **(A)**. Além disso,

$$y = u_1 \otimes v_1 \otimes u_1 = (u_1 \otimes v_i \otimes u_1)(I \otimes E_{i1} \otimes I)$$

para qualquer $i = 2, \ldots, s$. Portanto $y \in MP$. Podemos então desde já concuir que $\mathrm{Hom}_{R_q}(M_q) = R_q$, e portanto, para qualquer q, R_q é subálgebra comutativa maximal de $M_n(F)$, com $n = \dim M_q$.

Resta agora ver que M_q é um R_q-módulo fiel, para completar as hipóteses do teorema 3.16. Notemos para já que, observando os conjuntos geradores (2), (3), (4) e (6), podemos ver que MP^3 é gerado pelos $(s-1)^2$ elementos

$$\{ u_i \otimes v_1 \otimes u_j : i, j = 2, \ldots, s \}$$

e MP é gerado pelos $(s^2-1)s+1$ elementos

$$\{u_1 \otimes v_1 \otimes u_1\} \cup \{\, u_i \otimes v_j \otimes u_k : (i,k) \neq (1,1),\ 1 \leq j \leq s \,\}.$$

Daqui se pode concluir imediatamente que

$$\dim(MP/MP^3) = (s^2-1)s + 1 - (s-1)^2 = s^3 - s^2 + s. \tag{8}$$

Para ver agora a fidelidade do módulo, vamos tomar $p = (s_1, \ldots, s_q)$ um elemento não nulo de R_q, e encontrar $t \in M$ tal que $\overline{t}p \neq 0$. Se cada s_j pertencer a R'', então forçosamente $\sum_i s_i \neq 0$ uma vez que tomamos $p \neq 0$. Portanto, para algum elemento t do R-módulo fiel M, $t\sum_i s_i \neq 0$, e então $\overline{t}p = (ts_1, \ldots, ts_q) \neq 0$, como pretendíamos. Suponhamos então que para um certo $l \in [q]$, $s_l \notin R''$. Afirmar que $\overline{t}p = 0$ é dizer que, para cada i, $ts_i \in M'' = MP^3$. Basta então encontrar um t em R tal que $ts_l \notin MP^3$. Suponhamos para já que s_l é tensor decomponível, $s_l = f \otimes g \otimes h$. Vamos apresentar, em cada caso possível para s_l, um vector de M tal que $ts_l \notin MP^3$. Como $s_l \notin P^3$, pelo menos um dos elementos f, g ou h tem que ser I. Então sete casos podem ocorrer:

s_l	t	$ts_l\ (\notin MP^3)$
$I \otimes I \otimes I$	$u_1 \otimes v_1 \otimes u_1$	$u_1 \otimes v_1 \otimes u_1$
$E_{1i} \otimes I \otimes I$	$u_1 \otimes v_1 \otimes u_1$	$u_i \otimes v_1 \otimes u_1$
$I \otimes E_{i1} \otimes I$	$u_1 \otimes v_i \otimes u_1$	$u_1 \otimes v_1 \otimes u_1$
$I \otimes I \otimes E_{1i}$	$u_2 \otimes v_2 \otimes u_1$	$u_2 \otimes v_2 \otimes u_i$
$E_{1i} \otimes E_{j1} \otimes I$	$u_1 \otimes v_j \otimes u_1$	$u_i \otimes v_1 \otimes u_1$
$E_{1i} \otimes I \otimes E_{1j}$	$u_1 \otimes v_2 \otimes u_1$	$u_i \otimes v_2 \otimes u_j$
$I \otimes E_{i1} \otimes E_{1j}$	$u_1 \otimes v_i \otimes u_1$	$u_1 \otimes v_1 \otimes u_j$

(9)

Observação. Não pode haver dois elementos da base de R que apliquem um elemento α da base de M induzida pelas bases de U e de V noutro elemento β da mesma base. Sendo $\alpha = u_i \otimes v_j \otimes u_k$ e $\beta = u_{i'} \otimes v_{j'} \otimes u_{k'}$, se existir um elemento $\gamma = f \otimes g \otimes h$ da base de R tal que $\alpha\gamma = \beta$, então este está bem determinado:

- se $i = i'$, $f = I$, se $i = 1 \neq i'$, $f = E_{1i'}$,
- se $j = j'$, $g = I$, se $j' = 1 \neq j$, $g = E_{j1}$,
- se $k = k'$, $h = I$, se $k = 1 \neq k'$, $h = E_{1k'}$,

caso contrário não existe γ nestas condições. Para concluir isto basta observar que $u_i E_{1j} = \delta_{i1} u_j$ e $v_i E_{j1} = \delta_{ij} v_1$.

Vamos finalmente considerar o caso em que s_l não é decomponível. Neste caso s_l terá que ter um coeficiente não nulo num certo elemento da primeira coluna da lista (9), quando expresso como combinação linear dos elementos da base de M induzida pelas bases de U e de V. Tomando então o vector correspondente na segunda coluna, e aplicando-lhe s_l, o resultado é uma combinação linear de vectores da base de M em que o coeficiente do vector que aparece na terceira coluna da lista é não nulo, pela observação anterior. Assim, a imagem não está em MP^3 como pretendido.

Mostrámos, em qualquer caso, que existe $t \in M$ tal que $\overline{t}p \neq 0$, e isto mostra que M_q é um R_q-módulo fiel.

Finalmente, o resultado 3.16 aplicado ao nosso caso dá

$$\lim_{q\to\infty} \frac{\dim R_q}{\dim M_q} = \frac{\dim(P/P^3)}{\dim(MP/MP^3)} = \frac{3s^2 - 3s}{s^3 - s^2 + s},$$

tendo em conta as dimensões apresentadas nas equações (5) e (8). Sendo então $\epsilon > 0$ qualquer, tomemos s tal que $(3s^2 - 3s)/(s^3 - s^2 + s) < \epsilon/2$, e fazendo toda a construção anterior usando este s no início. Finalmente, tomamos q de tal modo que

$$\left| \frac{\dim R_q}{\dim M_q} - \frac{3s^2 - 3s}{s^3 - s^2 + s} \right| \leq \frac{\epsilon}{2}.$$

Pondo então $S_\epsilon := R_q$ e portanto $k := \dim M_q$ obtemos o pretendido, que era $\dim S_\epsilon / k \leq \epsilon$, sendo S_ϵ subálgebra comutativa maximal de $M_k(F)$.

□ □
□ □
□ □

Capítulo 4

Usando Análise Combinatória

Se vuol venire nella mia scola
la capriola Le insegnerò.
(Fígaro, em *As Bodas de Fígaro*,
de L. da Ponte, segundo Beaumarchais).

Neste capítulo não iremos demonstrar de novo o resultado acerca da álgebra gerada por duas matrizes que comutam, mas iremos apresentar algumas demonstrações recentes de resultados de Gerstenhaber, nomeadamente de um resultado apresentado em [Ge5], último artigo de uma série de vários ([Ge1] – [Ge5]) onde se estudam as variedades de matrizes nilpotentes de um ponto de vista da geometria algébrica, com algumas relações interessantes com aspectos combinatórios (nomeadamente em [Ge4]).

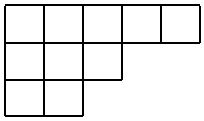
4.1 Generalidades sobre partições

Seja n um inteiro positivo. Diremos que o n-uplo de inteiros não negativos $\alpha = (a_1, \ldots, a_n)$ é uma *partição* de n se

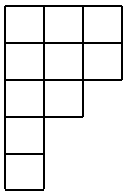
$$a_1 \geq \ldots \geq a_n \text{ e } a_1 + \ldots + a_n = n.$$

$a_1, \ldots, a_n$ são chamados os *termos* de α. Quando dissermos que um certo k-uplo de inteiros em ordem decrescente, com $k \leq n$ é uma partição de n, estamos a pensar no n-uplo obtido daquele juntando-lhe $n - k$ zeros finais. Supondo que a_k é o último termo não nulo de α, podemos associar à partição

um *quadro de Young*, que é uma figura que tem k linhas, tendo a_i caixas na i-ésima linha. Por exemplo, para $(5,3,2)$, que é uma partição de 10, o quadro de Young será:



Definimos a *conjugada* da partição α como sendo $\alpha^* = (a_1^*, \ldots, a_n^*)$ definida por $a_j^* = |\{i : a_i \geq j\}|$, $j = 1, \ldots, n$. Se pensarmos nos quadros de Young, obtemos a partição dual de α lendo o quadro por colunas, isto é, a_i^* é exactamente o número de caixas na i-ésima coluna do quadro de α. No exemplo dado, a partição dual de (5,3,2) é (3,3,2,1,1), e o respectivo quadro é



É simples de verificar que $(\alpha^*)^* = \alpha$. Seja $\beta = (b_1, \ldots, b_n)$ uma partição de n. Diremos que β *majora* α, que notaremos por $\alpha \preceq \beta$ se

$$a_1 + \ldots + a_j \leq b_1 + \ldots + b_j \qquad j = 1, \ldots, n.$$

A majoração assim definida é uma ordem parcial no conjunto das partições de n. Finalmente, para a inteiro, vamos notar $a^+ := \max\{0, a\}$. Note-se que, para b inteiro não negativo,

$$a^+ + b \geq (a+b)^+. \tag{1}$$

Terminamos esta secção com algumas propriedades simples desta relação, que serão úteis na secção seguinte.

Proposição 4.1 *Sejam $\alpha = (a_1, \ldots, a_n)$ e $\beta = (b_1, \ldots, b_n)$ partições de n, Então $\alpha \preceq \beta$ se e só se $\beta^* \preceq \alpha^*$.*

Demonstração. Sejam $\alpha^* = (a_1^*, \ldots, a_n^*)$ e $\beta^* = (b_1^*, \ldots, b_n^*)$ e suponhamos que $\beta^* \preceq \alpha^*$. Sabemos que $\sum_{i=1}^n a_i^* = \sum_{i=1}^n b_i^* = n$, e portanto, podemos dizer que $\beta^* \preceq \alpha^*$ se e só se $b_{j+1}^* + \ldots + b_n^* \geq a_{j+1}^* + \ldots + a_n^*$ para $j \in [n-1]$. Ora, pelos quadros de Young, é simples verificar que

$$\sum_{i=1}^{n} (a_i - j)^+ = a_{j+1}^* + \ldots + a_n^* \qquad j \in [n-1],$$

e o mesmo para β. Assim, estamos a supor que $\sum_{i=1}^{n}(b_i-j)^+ \geq \sum_{i=1}^{n}(a_i-j)^+$ para qualquer $j \in [n]$ (para $j = n$, temos a igualdade, com zero em ambos os membros). Tomando então $j := b_m$, obtemos

$$\sum_{i=1}^{n}(b_i - b_m)^+ \geq \sum_{i=1}^{n}(a_i - b_m)^+ \geq \sum_{i=1}^{m}(a_i - b_m)^+.$$

Adicionando agora mb_m ao primeiro e ao terceiro somatórios, obtemos, para o primeiro

$$\sum_{i=1}^{n}(b_i - b_m)^+ + mb_m = \sum_{i=1}^{m}(b_i - b_m) + mb_m = \sum_{i=1}^{m} b_i,$$

usando o facto de β ser decrescente, e para o terceiro

$$\sum_{i=1}^{m}(a_i - b_m)^+ + mb_m = \sum_{i=1}^{m}\left((a_i - b_m)^+ + b_m\right) \geq \sum_{i=1}^{m} a_i,$$

usando a relação (1). Obtivemos assim que $\sum_{i=1}^{m} b_i \geq \sum_{i=1}^{m} a_i$, para $m \in [n]$, ou seja, $\alpha \preceq \beta$. Observemos agora que $\alpha = (\alpha^*)^*$ e $\beta = (\beta^*)^*$. Assim, pondo β no lugar de α^* e α no lugar de β^* no início da demonstração, obtemos a recíproca, o que termina a demonstração da equivalência pretendida. □

A proposição que se segue é equivalente à afirmação que, dadas duas partições α e β, $\alpha \preceq \beta$ que difiram em apenas dois termos, e se γ é tal que $\alpha \preceq \gamma \preceq \beta$, então $\gamma = \alpha$ ou $\gamma = \beta$ (pode ver-se este resultado demonstrado em [JK]).

Proposição 4.2 *Sejam $\alpha = (a_1, \ldots, a_n)$ e $\beta = (b_1, \ldots, b_n)$ partições de n, com $\alpha \preceq \beta$, $\alpha \neq \beta$. Então existem $\gamma_1, \ldots, \gamma_h$, partições de n, tais que*

$$\alpha = \gamma_h \preceq \ldots \preceq \gamma_1 \preceq \gamma_0 = \beta,$$

e γ_i e γ_{i+1} apenas diferem em dois termos, para $0 \leq i \leq h-1$.

Demonstração. Visto que $\alpha \neq \beta$, seja p' o primeiro índice tal que $a_{p'} \neq b_{p'}$. Pela majoração, temos que ter $a_{p'} < b_{p'}$. Ora, como a soma de todos os termos de qualquer uma das partições é n, tem que existir pelo menos um termo de α maior que o respectivo termo de β. Seja q o menor índice tal que $a_q > b_q$. Como temos $q > p'$, tomemos p o maior índice menor

que q tal que $a_p < b_p$. Com esta construção garantimos que, para j entre p e q,

$$a_j = b_j \text{ e } a_q, b_q \leq a_j = b_j \leq a_p, b_p.$$

Seja agora c o mínimo entre $b_p - a_p$ e $a_q - b_q$. Pelas propriedades apresentadas, a sucessão

$$\gamma_1 := (b_1, \ldots, b_{q-1}, b_q - c, b_{q+1}, \ldots, b_{p-1}, b_p + c, b_{p+1}, \ldots, b_n)$$

é ainda uma sucessão decrescente de inteiros, é majorada por β, difere dela em apenas dois termos e majora α. Se $\gamma_1 \neq \alpha$, repetimos o processo, obtendo γ_2, Esta iteração terminará ao fim de um número finito de passos, uma vez que o número de termos em que γ_i é diferente de α vai decrescendo, em sentido estrito — em γ_1, por exemplo, ou $b_q - c = a_q$ ou $b_p + c = a_p$, e estes elementos não voltam a ser alterados. Isto termina a demonstração. □

Proposição 4.3 *Sejam $\alpha = (a_1, \ldots, a_n)$ e $\beta = (b_1, \ldots, b_n)$ partições de n, com $\alpha \preceq \beta$. Então*

$$\sum_{i=1}^{n} a_i^2 \leq \sum_{i=1}^{n} b_i^2,$$

com desigualdade estrita se $\alpha \neq \beta$.

Demonstração. Pela proposição anterior, podemos demonstrar o resultado supondo que α e β apenas diferem em dois termos, e sejam p e q os índices desses termos, $1 \leq p < q \leq n$. Então $a_p < b_p$ e $a_q > b_q$, e seja $c > 0$ tal que $c = b_p - a_p = a_q - b_q$. Ora $b_p - b_q > c$, pois caso contrário viria $b_p \leq b_q + c$ e $a_p < b_p \leq b_q + c = a_q$, contradizendo o facto de α ser decrescente. Assim

$$\begin{aligned}
\sum_{i=1}^{n} a_i^2 &= \sum_{i \neq p,q}^{n} a_i^2 + a_p^2 + a_q^2 \\
&= \sum_{i \neq p,q}^{n} b_i^2 + (b_p - c)^2 + (b_q + c)^2 \\
&= \sum_{i \neq p,q}^{n} b_i^2 + b_p^2 + b_q^2 + \underbrace{2c(c - (b_p - b_q))}_{<0} < \sum_{i=1}^{n} b_i^2,
\end{aligned}$$

o que termina a demonstração. □

4.2 Espaços de matrizes nilpotentes

Vamos apresentar agora dois resultados de majoração de dimensões de espaços de matrizes nilpotentes, seguindo [BC], onde se podem encontrar as novas demonstrações dos resultados de Gerstenhaber, no nosso parecer menos elaboradas que as iniciais.

Seja F um corpo, e C uma matriz de zeros e uns. Recorde-se que notamos por $M_n[C](F)$ o subespaço de $M_n(F)$ constituído pelas matrizes $X = [x_{ij}]$ que verificam $x_{ij} = 0$ se $c_{ij} = 0$. Sejam A uma matriz do tipo $m \times n$ e k um inteiro, $-m+1 \leq k \leq n-1$. Chamaremos *k-ésima diagonal* (ou *diagonal de ordem k*) da matriz A ao conjunto das posições (i, j) com $j = i + k$. Assim, por exemplo, a $(n-1)$-ésima diagonal da matriz será constituída apenas pela posição $(1, n)$, a diagonal de ordem 0 será a diagonal principal, a de ordem -1 será o conjunto das posições imediatamente abaixo desta, e assim sucessivamente. Dentro da mesma diagonal, diremos que uma posição (i, j) está acima de uma outra (p, q) se $i < p$, e abaixo se $i > p$.

Sejam agora $W \subseteq M_n(F)$ um espaço de matrizes, e $(A_1, \ldots, A_m)$ uma base de W. Seja $\sigma = (s_1, \ldots, s_{n^2})$ uma listagem das n^2 posições (i, j), $1 \leq i, j \leq n$. Consideremos agora os m vectores linha $(\operatorname{vec} A_i)^T, 1 \leq i \leq m$, com as entradas ordenadas segundo a ordem σ acima definida, e notemos por $M_\sigma(A_1, \ldots, A_m)$ a matriz do tipo $m \times n^2$ em que as linhas são exactamente esses vectores:

$$M_\sigma(A_1, \ldots, A_m) = \begin{bmatrix} (\operatorname{vec} A_1)^T \\ \vdots \\ (\operatorname{vec} A_m)^T \end{bmatrix}.$$

Condensando então esta matriz, sem trocar linhas, obtemos uma outra matriz $M_\sigma(B_1, \ldots, B_m)$, em que cada matriz B_i tem um 1 que era um dos 1's que apareceram, precedidos de zeros, na matriz $M_\sigma(B_1, \ldots, B_m)$. A esse 1 chamaremos o *1 principal* de B_i. Suponhamos que esses 1's apareciam nas posições $s_{t_1}, \ldots, s_{t_m}$, $t_1 \leq \ldots \leq t_m$. Como W é um espaço vectorial, as matrizes B_i continuam a pertencer a W e formam uma base.

Diremos que as matrizes $B_1, \ldots, B_m$ são obtidas de $A_1, \ldots, A_m$ por *condensação na ordem σ*. À matriz de 1's e 0's $\mathcal{B}^\sigma(A_1, \ldots, A_m)$ (que abreviaremos por $\mathcal{B}^\sigma$ quando isto não levar a confusão) do tipo $n \times n$ que tem 1's nas posições $s_{t_1}, \ldots, s_{t_m}$ e zeros nas outras chamaremos *matriz característica* do espaço W, relativamente à ordem σ e à base $(A_1, \ldots, A_m)$. O número de 1's em $\mathcal{B}^\sigma$ é exactamente a dimenão de W. A ideia principal das demonstrações que se seguem é escolher, em cada caso, uma ordem σ que confira a $\mathcal{B}^\sigma$ uma

estrutura combinatória adequada ao que se pretende.

Vamos agora apresentar uma nova demonstração do resultado 2.14, usando estas técnicas.

Proposição 4.4 *Se $W \subseteq M_n(F)$ é um espaço de matrizes nilpotentes, então* $\dim W \leq n(n-1)/2$.

Demonstração. Seja $(A_1, \ldots, A_m)$ uma base de W. Definimos então a ordem σ da seguinte forma: tomamos as diagonais na ordem $n-1, \ldots, -n+1$ e ordenamos as entradas em cada diagonal da seguinte forma: nas diagonais $n-1, \ldots, 0$ de cima para baixo, nas diagonais $-1, \ldots, -n-1$, de baixo para cima. Numa matriz do tipo 4×4, se colocarmos em cada entrada o lugar que a respectiva posição ocupa na ordem σ, obtemos

$$\begin{bmatrix} 7 & 4 & 2 & 1 \\ 13 & 8 & 5 & 3 \\ 15 & 12 & 9 & 6 \\ 16 & 14 & 11 & 10 \end{bmatrix}.$$

Seja $(B_1, \ldots, B_m)$ a base obtida a partir de $(A_1, \ldots, A_m)$ por condensação na ordem σ e seja $\mathcal{B}^\sigma$ a respectiva matriz característica. Suponhamos que para a matriz B_k, o 1 principal aparece na posição (i, j). Pela escolha de σ, temos as seguintes propriedades:

1. se $i \leq j$, as diagonais $n-1, \ldots, j-i+1$ apenas contêm zeros, e a diagonal $j-i$ contem zeros acima da posição (i, j), e

2. se $i > j$, as diagonais $n-1, \ldots, 0, \ldots j-i+1$ apenas contêm zeros, e a diagonal $j-i$ contem zeros abaixo da posição (i, j).

Se B_t tivesse um 1 principal na diagonal principal, pela propriedade 1, teríamos que qualquer sua potência teria um 1 na mesma posição, o que não pode ser, pois B_k é nilpotente. Portanto, $\mathcal{B}^\sigma$ apenas tem zeros na diagonal principal. Suponhamos agora que B_k tem o seu 1 principal na posição (i, j) e que existe B_l que tem o 1 principal na posição (j, i), com $i < j$. Notemos então por $s(X)$ a soma dos menores principais de ordem 2 da matriz X. Se X é nilpotente, então $s(X) = 0$, como já vimos no capítulo 2, lema 2.15. Vamos verificar que, nestas condições, teríamos

$$s(B_k + B_l) = s(B_k) - 1,$$

o que é impossível, pois contradiz a nilpotência das matrizes. Isto fornece imediatamente o resultado, pois neste caso o máximo de 1's que $\mathcal{B}^\sigma$ pode ter é exactamente $n(n-1)/2$.

Verifiquemos então a relação. Vamos mostrar que todos os menores principais de ordem 2 de B_k coincidem com os de $B_k + B_l$, à excepção de um, que é $|B_k[ij|ij]|$. As matrizes B_k, B_l e $B = [b_{ij} : i, j \in [n]] := B_k + B_l$ têm o seguinte aspecto:

$$B_k = \begin{bmatrix} 0 & & & 0 \\ & 0 & & \\ & & 1 & \\ * & & & * \end{bmatrix} \quad B_l = \begin{bmatrix} * & & & 0 \\ & * & & \\ & & 1 & \\ * & & & 0 \end{bmatrix} \quad B = \begin{bmatrix} & 0 & & & 0 \\ * & & 0 & & \\ & * & & 1 & \\ & & b_{ji} & & * \\ * & & & * & \end{bmatrix}$$

em que $b_{ji} = (b_k)_{ij} + 1$, sendo $B_k = [(b_k)_{ij} : i, j \in [n]]$. Observe-se para já que nas diagonais $i - j + 1, \ldots, 0, \ldots, n - 1$, as entradas das matrizes B_k e B coincidem. Tomemos então $p, q \in [n]$,e suponhamos que $q - p > j - i$ ou $q - p < j - i$ (informalmente, isto significa que a posição (p, q) está nas 'diagonais dos cantos', bem como a posição (q, p)). Então, supondo que $q > p$, as entradas (p, q) quer da matriz B_k quer da matriz B são zero, e portanto o menor resume-se a $(b_k)_{pp}(b_k)_{qq}$ em qualquer uma das duas matrizes, e os menores coincidem. Se é $q < p$, é a entrada (q, p) que é zero, e o raciocínio segue analogamente. Suponhamos agora que $i-j < q-p < j-i$ (aqui, quer a posição (p, q) quer a (q, p) estão nas 'diagonais do meio'). Nestas condições, as entradas das matrizes B_k e B são iguais, e portanto os respectivos menores coincidem. Vamos então supor que $q - p = j - i$, pois o caso em que $q - p = i - j$ teria um tratamento análogo, uma vez que nesse caso seria a a posição (q, p) que estaria na diagonal $j - i$. Se a posição (p, q) estiver acima da posição (i, j), então o menor fica novamente $(b_k)_{pp}(b_k)_{qq}$, pois $b_{pq} = (b_k)_{pq} = 0$. Se a posição (p, q) estiver abaixo da posição (i, j), então $b_{pq} = (b_k)_{pq}$ e $b_{qp} = (b_k)_{qp}$, uma vez que B_l tem entradas nulas nesta zona, e o menor resulta igual. Finalmente, se $(p, q) = (i, j)$, o menor $|B[ij|ij]|$ fica:

$$|B[ij|ij]| = b_{ii}b_{jj} - b_{ij}b_{ji} = (b_k)_{ii}(b_k)_{jj} - 1(1 + (b_k)_{ji}) = |B_k[ij|ij]| - 1,$$

o que conclui a verificação. □

Vamos terminar, apresentando uma demonstração de um resultado de [Ge5], que dá uma majoração da dimensão de um espaço de matrizes nilpo-

tentes à custa de um estudo combinatório das estruturas de blocos de Jordan das matrizes de W. Começamos com alguns resultados técnicos.

A partir de agora, vamos supor que o corpo F é infinito. Seja A uma matriz nilpotente de $M_n(F)$, e sejam $k_1 \geq \ldots \geq k_r \geq 1$ os tamanhos dos blocos de Jordan de A. Definimos a *partição de Jordan de* A como

$$\mathrm{prt}(A) := (k_1, \ldots, k_r, \underbrace{0, \ldots, 0}_{n-r}),$$

que fica sendo, portanto, uma partição de n, que notaremos por vezes apenas por $(k_1, \ldots, k_r)$. Assim, a majoração de partições induz uma ordem parcial nas classes de semelhança das matrizes nilpotentes de $M_n(F)$.

Lema 4.5 *Sejam A e B duas matrizes nilpotentes de $M_n(F)$. Temos que*

1. $\mathrm{prt}(A)^* = (n - \mathrm{c}(A), \mathrm{c}(A) - \mathrm{c}(A^2), \ldots, \mathrm{c}(A^{n-1}) - \mathrm{c}(A^n))$
2. $\mathrm{prt}(A) \preceq \mathrm{prt}(B)$ *se e só se* $\mathrm{c}(A^p) \leq \mathrm{c}(B^p)$, *para qualquer* $p \in [n]$.

Demonstração. Pondo

$$\alpha = (a_1, \ldots, a_n) := \mathrm{prt}(A)^* \text{ e } \beta = (b_1, \ldots, b_n) := p(B)^*,$$

basta observar que $\mathrm{c}(A^p) = n - (a_1 + \ldots + a_p)$ e $\mathrm{c}(B^p) = n - (b_1 + \ldots + b_p)$ para qualquer $p \in [n]$. Daqui se conclui imediatamente a afirmação 1. Além disso,

$$\mathrm{prt}(A) \preceq \mathrm{prt}(B) \Leftrightarrow \beta \preceq \alpha \Leftrightarrow \mathrm{c}(A^p) \leq \mathrm{c}(B^p) \qquad \forall p \in [n],$$

o que conclui a demonstração. □

Seja W um espaço de matrizes nilpotentes. Definimos a *partição de Jordan* de W como sendo o supremo, para a ordem parcial induzida, das partições de Jordan das matrizes de W. Sendo F infinito, o lema seguinte assegura que existe em W uma matriz que admite esta partição como partição de Jordan.

Lema 4.6 *Seja $W \subseteq M_n(F)$ um espaço de matrizes nilpotentes, com F infinito, e x uma indeterminada. Então temos que:*

1. *existe uma matriz Q em W tal que* $\mathrm{prt}(Q) = \mathrm{prt}(W)$ *e*

2. para toda a matriz $A \in W$, $\mathrm{prt}(Q - xA) = \mathrm{prt}(Q)$.

Além disso, para $A \in W$, existe uma matriz $P \in M_n(F)$ tal que

3. $A = QP - PQ$.

e mais ainda, se $R \in M_n(F)$ comutar com Q, então

4. $\mathrm{tr}(AR) = 0$ e $\mathrm{tr}(APR) = 0$.

Demonstração. 1. Seja p o maior índice de uma matriz em W, isto é o tamanho do maior bloco de Jordan de uma matriz de W, e seja

$$r_k := \max\{\mathrm{c}(A^k) : A \in W,\ k \in [p]\}.$$

Pelo lema anterior, basta mostrar que existe uma matriz $Q \in W$ tal que $\mathrm{c}(Q^k) = r_k$, para $k \in [p]$. Seja $(A_1, \ldots, A_m)$ uma base de W. Então W é constituído por todas as matrizes A que se podem escrever como

$$A = A(x_1, \ldots, x_m) = x_1A_1 + \ldots + x_mA_m \qquad x_1, \ldots, x_m \in F.$$

Ora, para cada $k \in [p]$, é possível escolher $v_1, \ldots, v_m \in F$ de tal modo que a matriz $A(v_1, \ldots, v_m)$ tenha um menor de ordem r_k não nulo, e portanto o respectivo menor de $A(x_1, \ldots, x_m)$ é um polinómio não nulo. Multiplicando estes p menores polinomiais, obtemos um polinómio não nulo $f(x_1, \ldots, x_m)$, e como F é infinito, existem $t_1, \ldots, t_m \in F$ tais que $f(t_1, \ldots, t_m) \neq 0$. Pondo agora $Q := A(t_1, \ldots, t_m)$, temos o pretendido. Para demonstrar a afirmação 2, basta ver que cada menor de Q se anula se e só se o respectivo menor de $Q - xA$ também se anula, o que é simples de verificar, tendo em conta a definição de Q.

3. Visto que Q e $Q - xA$ são semelhantes, existe $C = C(x) \in M_n(F(x))$, invertível, tal que

$$(Q - xA)C = CQ, \tag{2}$$

em que $F(x)$ é o corpo de fracções de $F[x]$. Podemos agora supor que $C \in M_n(F[x])$, tomando o mínimo múltiplo comum dos denominadores das entradas de C, e multiplicando C por esse polinómio. A nova matriz

$$C = C(x) = C_0 + xC_1 + \ldots + x^kC_k$$

continua a satisfazer a condição (2) e a ser invertível, isto é, o seu determinante é um polinómio não nulo. Assim, existe em F um elemento a, tal que $\det C(a) \neq 0$. Pondo agora $C(x - a)$ no lugar de C, a nova matriz ainda

satisfaz a condição (2) e portanto podemos supor que $C(0) = C_0$ é invertível. Além disso, C_0 comuta com Q, bastando tomar $x = 0$ na equação (2) para verificar isto. Finalmente, multiplicando a equação (2) por C_0^{-1} à direita, obtemos que CC_0^{-1} ainda satisfaz (2). Podemos então supor $C_0 = I_n$. Desenvolvendo a equação, obtemos

$$Q + x(QC_1 - A) + x^2(QC_2 - AC_1) + \ldots = Q + xC_1Q + x^2C_2Q + \ldots \quad (3)$$

e portanto $A = QC_1 - C_1Q$. Pondo agora $P := C_1$ obtemos o resultado. Para os resultados de 4, basta ver que

$$\begin{aligned} \mathrm{tr}(AR) &= \mathrm{tr}(QPR - PQR) \\ &= \mathrm{tr}(Q(PR)) - \mathrm{tr}((PR)Q) \\ &= \mathrm{tr}(QPR) - \mathrm{tr}(QPR) = 0, \end{aligned}$$

e o outro resultado sai similarmente, observando que

$$AP = AC_1 = QC_2 - C_2Q,$$

da comparação dos coeficientes de x^2 da equação (3). □

A proposição que se segue apresenta já o limite superior para a dimensão de um espaço de matrizes nilpotentes, afirmando que ele é alcançado.

Proposição 4.7 *Seja F um corpo infinito, e W um espaço de matrizes nilpotentes de $M_n(F)$, η uma partição de n e $\gamma = (c_1, \ldots, c_n) := \eta^*$. Então existe um subespaço $W \subseteq M_n(F)$ de matrizes nilpotentes tal que* $\mathrm{prt}(W) = \eta$ *e além disso*

$$\dim(W) = \frac{1}{2}\left(n^2 - \sum_{i=1}^{n} c_i^2\right).$$

Temos ainda a igualdade

$$\frac{1}{2}\left(n^2 - \sum_{i=1}^{n} c_i^2\right) = \sum_{p=1}^{r} \frac{k_p(k_p - 1)}{2} + \sum_{1 \le p < q \le r} k_q(k_p + 1), \quad (4)$$

sendo $k_1 \geq \ldots \geq k_r$ os termos diferentes de zero de η.

Demonstração. Vamos para já mostrar que existe um espaço W com a dimensão pretendida. Sejam $c_1 \geq \ldots \geq c_l$ os termos diferentes de zero de

γ. Vamos notar por U_{rs} uma matriz do tipo $r \times s$ com as entradas todas iguais a 1. Tomemos agora

$$C = \begin{bmatrix} 0_{c_1} & U_{c_1c_2} & \cdots & U_{c_1c_l} \\ 0 & 0_{c_2} & \cdots & U_{c_2c_l} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots 0 & 0_{c_l} \end{bmatrix}$$

e seja $W := M_n[C](F)$. Usando agora a alínea 1. do lema 4.5, podemos verificar que $\mathrm{prt}(W)^* = \gamma$ e temos a igualdade pretendida com respeito à dimensão, pois o número de 1's em C é exactamente $1/2(n^2 - \sum_i c_i^2)$.

Mostremos agora a outra igualdade. Seja $W' := M_n[C'](F)$, em que $C' := [V_{ij} : i, j \in [r]]$ é uma matriz por blocos, em que $V_{pq} \in M_{k_q \times k_q}$, e tem zeros nas diagonais de ordem inferior ou igual a $n_q - n_p$ e uns nas restantes posições. São matrizes do tipo

$$\begin{bmatrix} & & 1 \\ 1 & & \\ 0 & 1 & \\ & 0 & 1 \\ 0 & & 0 \end{bmatrix} \text{ e } \begin{bmatrix} & 0 & 1 & & 1 \\ & & 0 & 1 & \\ & & & 0 & 1 \\ 0 & & & & 0 \end{bmatrix}.$$

Daqui se pode ver que, para $p < q$, a matriz $V_{pq} + V_{qp}^T$ tem apenas zeros nas posições da diagonal de ordem $n_p - n_q$, e uns nas restantes posições. Podemos ver assim que o número de 1's na matriz C' é

$$\sum_{p=1}^{r} \frac{k_p(k_p - 1)}{2} + \sum_{1 \le p < q \le r} k_q(k_p + 1),$$

sendo cada parela do primeiro somatório relativa a um bloco principal e cada parcela do outro a um par de blocos V_{pq} e V_{qp} com $p < q$. Vamos agora reordenar as linhas e as colunas de C'. Escolhemos primeiro a última linha de cada linha de blocos, e pomo-las na parte de cima da matriz, o que fornece c_1 linhas nulas. Depois fazemos o mesmo com as segundas linhas, colocando-as por baixo das anteriores, o que fornece mais c_2 linhas. Ficamos no fim com c_l linhas na parte de baix da matriz. Executamos um processo análogo com as colunas, escolhendo as últimas e colocando-as à esquerda. A matriz obtida no fim deste processo é C^T. Assim, provámos que a dimensão de W' é igual à de W, o que é exactamente a igualdade (4). □

Apresentamos agora o teorema principal.

Teorema 4.8 *Seja F um corpo infinito, e W um espaço de matrizes nilpotentes de $M_n(F)$. Seja γ a conjugada da partição de Jordan de W,*

$$\gamma = (c_1, \ldots, c_n) = \mathrm{prt}(W)^*.$$

Então, temos que

$$\dim(W) \leq \frac{1}{2}\left(n^2 - \sum_{i=1}^{n} c_i^2\right).$$

Demonstração. Tomemos Q uma matriz satisfazendo

$$\mathrm{prt}(Q) = \mathrm{prt}(W) = (k_, \ldots, k_r),$$

que existe, pelo lema 4.6. Seja T uma matriz invertível de $M_n(F)$ tal que TQT^{-1} é a forma normal de Jordan de Q, e tomemos TWT^{-1} em lugar de W, de forma a podermos supor que Q está na sua forma normal de Jordan,

$$Q = \mathrm{diag}(Q_1, \ldots, Q_r), \qquad Q_i \in M_{k_i}(F),$$

e Q_i é um bloco de Jordan. Seja agora $(A_1, \ldots, A_m)$ uma base de W. Particionamos agora cada matriz A da base em blocos identicos em tamanho aos de Q, isto é, pomos $A = [A_{ij}]$, com $A_{ij} \in M_{k_i \times k_j}(F)$. Vamos agora definir uma ordem para as entradas de A. Em cada bloco A_{pq}, tomamos as diagonais por ordem decrescente, $k_q - 1, \ldots, 0, \ldots, -k_p + 1$, e dentro de cada diagonal ordenamos as posições da seguinte forma:

- para $p < q$, de cima para baixo,
- para $p = q$ de cima para baixo nas diagonais de ordem não negativa ($k_q - 1$ até 0), e de baixo para cima nas restantes, e
- para $p > q$, de baixo para cima.

Ordenamos agora os blocos. Os primeiros serão os blocos A_{pq} em que $p \leq q$, isto é, sobre e acima da diagonal principal, que serão ordenados considerando as colunas de blocos ordenadas da direita para a esquerda, e dentro de cada coluna, os blocos ordenados de cima para baixo. Formalmente, será A_{pq} precede A_{rs} sse $s > q$ ou então $(s = q)$ e $(r > s)$. Para os blocos abaixo da diagonal principal $(p > q)$, consideramos as linhas de blocos ordenadas de cima para baixo, e cada linha ordenada da esquerda para a direita. Formalmente, A_{pq} precede A_{rs} sse $r > p$ ou então $(r = p)$ e $(s > q)$. Fica assim definida a ordem, que chamaremos σ.

Seja agora $(B_1, \ldots, B_m)$ a base obtida de $(A_1, \ldots, A_m)$ por condensação na ordem σ, e $\mathcal{B} = \mathcal{B}^\sigma$ a respectiva matriz característica. Vamos particionar $\mathcal{B}$ em blocos idênticos aos de Q, $\mathcal{B} = [\mathcal{B}_{pq} : p, q \in r]$. Vamos agora verificar que

(a) Cada bloco $\mathcal{B}_{pp}$ tem no máximo $k_p(k_p - 1)/2$ 1's,

(b) Cada par de blocos simetricamente colocados $\mathcal{B}_{pq}$ e $\mathcal{B}_{qp}$ $q < p$ têm, conjuntamente, no máximo $k_q(k_p - 1)$ 1's.

Vamos considerar que também os B's estão particionados por blocos, $B_t = [(B_t)_{pq} : p, q \in [r]]$, $t \in [m]$. Suponhamos que o 1 principal de B_t aparece no seu bloco $(B_t)_{pq}$. Pela escolha da ordem σ, temos as duas propriedades seguintes:

(i) Para $p \leq q$, $(B_t)_{rs} = 0$ se $s > q$ e $s \geq r$ ou se $s = q$ e $r < p$. Supondo agora $p < q$, então o bloco $(B_t)_{pq}$ tem apenas zeros nas diagonais $k_q - 1, \ldots, j - i + 1$, e na diagonal $j - i$ acima da posição (i, j).

(ii) Para $p > q$, $(B_t)_{rs} = 0$ se $r \leq s$ ou se $r < p$ ou se $r = p$ e $s > q$. Além disso, o bloco $(B_t)_{pq}$ tem apenas zeros nas diagonais $k_q - 1, \ldots, j - i + 1$, e na diagonal $j - i$ abaixo da posição (i, j).

Vamos agora ver o que se passa com os blocos principais. Seja $p \in [r]$ fixo, e considere-se o espaço W_p de matrizes de M_{k_p} gerado pelas matrizes $(B_t)_{pp}$ que contêm o 1 principal de B_t. Pelo facto de W ser um espaço de matrizes nilpotentes, e pela propriedade (i), podemos concluir que W_p é um espaço de matrizes nilpotentes (triangulares inferiores). Pela proposição 4.4, a dimensão de W_p (que é o número de zeros em $\mathcal{B}_{pp}$) é no máximo $k_p(k_p - 1)/2$, o que prova a afirmação (a).

Sejam agora $p, q \in [r]$, $p < q$. Usando um raciocínio anaálogo ao da proposição 4.4, verificamos que

(iii) Se a entrada (i, j) de $\mathcal{B}_{pq}$ for 1, a entrada (j, i) de $\mathcal{B}_{qp}$ é zero.

Vejamos agora que

(iv) Se a entrada (i, j) de $\mathcal{B}_{pq}$ for 1, a entrada $(j, i + 1)$ de $\mathcal{B}_{qp}$ é zero, para $i < k_p$.

Suponhamos então que a entrada (i, j) de $\mathcal{B}_{pq}$ é 1, e que a entrada $(j, i + 1)$ de $\mathcal{B}_{qp}$ é zero. Sejam B_u e B_v as duas matrizes cujos 1's principais estão nas

posições (i,j) e $(j,i+1)$, respectivamente. Pelo lema 4.6 existe uma matriz P tal que

$$B_u = QP - PQ. \tag{5}$$

Particionamos também P em blocos, $P = [P_{rs} : r,s \in [r]]$, da mesma forma que as matrizes de W. Pela equação (5), temos

$$(B_u)_{rs} = Q_r P_{rs} - P_{rs} Q_s \qquad r,s \in [r]. \tag{6}$$

Assim, se $(B_u)_{rs} = 0$, a matriz S que se obtém de P pondo todos os blocos diferentes de P_{rs} iguais a zero comuta com Q. Pondo então $P - S$ no lugar de P, a condição (5) ainda é satisfeita. Portanto, podemos supor que $P_{rs} = 0$ sempre que $(B_u)_{rs} = 0$. Tomamos agora $r := p$ e $s := q$, obtendo $(B_u)_{pq} = Q_p P_{pq} - P_{pq} Q_q$. Ponhamos agora

$$P_{pq} = P_{pq}^{(1)} + P_{pq}^{(2)},$$

em que $P_{pq}^{(2)}$ tem entradas iguais às de P_{pq} nas diagonais $-k_p+1, \ldots, j-i-2$ e na diagonal $j-i-1$ abaixo da posição $(i+1,j)$, um 1 na posição $(i+1,j)$ e zeros nas restantes posições. Observando agora que Q_p é um bloco de Jordan, é simples de verificar, com alguns cálculos, que $Q_p P_{pq}^{(1)} - P_{pq}^{(1)} Q_q = 0$, e portanto podemos supor que $P_{pq} = P_{pq}^{(2)}$.

Vamos agora aplicar os resultados do lema 4.6 à matriz $A = B_u + B_v$: existe uma matriz $\tilde{P}$ tal que $A = Q\tilde{P} - \tilde{P}Q$. Ora, pela propriedade (ii), os blocos principais de B_v são nulos, bem como os que estão acima da diagonal principal. Assim, A e B_u coincidem em todos os blocos sobre e acima da diagonal principal, e podemos assim supor que $\tilde{P}$ coincide com P em todos os blocos sobre e acima da diagonal principal (observe-se a equação (6)). Pomos agora

$$R := 0_{k_1} \oplus \ldots \oplus 0_{k_{q-1}} \oplus I_{k_q} \oplus 0_{k_{q+1}} \oplus \ldots \oplus 0_{k_r}$$

e vamos verificar que $\mathrm{tr}(A\tilde{P}R) = \mathrm{tr}(B_u P R) + 1$, o que contradiz o lema 4.6, uma vez que Q comuta com R, provando assim, por absurdo, que é válida a propriedade (iii). Façamos então a verificação.

A matriz PR é constituída apenas por zeros, excepto na q-ésima coluna de blocos, que é igual à de P, e passa-se algo análogo para a matriz $\tilde{P}R$. Assim, para os traços, basta fazer o produto da q-ésima linha de blocos de A (respectivamente, de B_u) pela q-ésima coluna de blocos de $\tilde{P}$ (respectivamente, de P), e ver quais são os traços das matrizes assim calculadas. As

matrizes ficam

$$\sum_{l=1}^{n}((B_v)_{ql} + (B_u)_{ql})\tilde{P}_{lq} \qquad \text{e} \qquad \sum_{l=1}^{n}(B_u)_{ql}P_{lq}$$

respectivamente. Vamos dividir cada somatório em quatro partes.

· $l < p$. Nestas condições $(B_u)_{lq} = 0$, e portanto podemos tomar $P_{lq} = 0$, e também $\tilde{P}_{lq} = 0$, pois a posição (l, q) fica acima da diagonal principal. Ambos os somatórios têm parcelas nulas.

· $p < l \leq q$. Como ainda estamos acima ou sobre a diagonal principal, $P_{lq} = \tilde{P}_{lq}$ e $(B_v)_{ql} = 0$. As parcelas ficam, em ambos os casos, $(B_u)_{ql}P_{lq}$.

· $l > q$ Aqui $(B_v)_{ql} = (B_u)_{ql} = 0$, portanto ambos os somatórios têm parcelas nulas.

· $l = p$ Ainda temos $\tilde{P}_{pq} = P_{pq}$. Vejamos que $\mathrm{tr}((B_v)_{qp}P_{pq}) = 1$. Observando as matrizes, $(B_v)_{qp}$ tem as diagonais $i-j+2, \ldots, k_p - 1$ apenas com entradas nulas, e na diagonal $i-j+1$ tem um 1 na posição $(j, i+1)$, e zeros abaixo dessa posição (pela ordenação das entradas e pela definição de B_v); P_{pq} tem zeros nas diagonais $j - i, \ldots, k_q - 1$, um 1 na posição $(i+1, j)$ e zeros na diagonal $j - i - 1$ acima desta posição, pelo que já vimos. Assim, todas as entradas principais de $(B_v)_{qp}P_{pq}$ são zero, excepto a entrada (j, j) que é 1. Temos assim provado o que pretendíamos.

Vamos agora ver que

(v) Os blocos de $\mathcal{B}$ acima da diagonal principal têm apenas zeros na última linha, e

(vi) Os blocos de $\mathcal{B}$ abaixo da diagonal principal têm apenas zeros na primeira coluna.

Suponhamos que havia uma matriz B_t que tinha o seu 1 principal na posição (k_p, j) do bloco $(B_t)_{pq}$, em que $1 \leq p < q \leq r$. Seja R uma matriz por blocos em que o único bloco diferente de zero é R_{qp}, que tem zeros em todas as diagonais, excepto na de ordem $k_p - j$, que tem apenas 1's (em particular, tem um 1 na posição (j, k_p)). Considerando a estrutura de B_t mencionada em (i), obtemos $\mathrm{tr}(B_t R) = 1$, contradizendo o lema 4.6, uma vez que R comuta com Q (é uma matriz triangular superior regular — veja-se a Introdução). Demonstrámos assim (v), e (vi) tem uma demonstração análoga.

Vamos agora completar a demonstração de (b). Seja $p < q$ e considere-se $K := \mathcal{B}_{pq} + \mathcal{B}_{qp}^T$. Por (iii), K é uma matriz de zeros e uns. Suponhamos,

com vista a absurdo, que uma certa coluna v de K tem as entradas todas iguais a 1. Por (v), o primeiro 1 vem de $\mathcal{B}_{pq}$, e por (vi), o último 1 vem de $\mathcal{B}_{qp}^T$. Suponhamos que o primeiro 1 de v que provém de $\mathcal{B}_{qp}^T$ está na linha i. Então $2 \leq i \leq n_p$ e o 1 na linha $i-1$ de v provém de $\mathcal{B}_{pq}$, o que é falso, pela afirmação (iv). Assim, cada coluna de K tem pelo menos um zero, e isto termina a demonstração de (b).

Visto que $\dim(W)$ é igual ao número de uns na matriz $\mathcal{B}$, segue-se de (a) e de (b) que

$$\dim(W) \leq \sum_{p=1}^{r} \frac{k_p(k_p-1)}{2} + \sum_{1 \leq p < q \leq r} k_q(k_p+1),$$

e pela igualdade (4), obtemos

$$\dim(W) \leq \frac{1}{2}\left(n^2 - \sum_{i=1}^{n} c_i^2\right),$$

que é o pretendido. □

Terminamos com dois corolários deste teorema.

Corolário 4.9 *Seja $p \in [n-1]$ um inteiro, e seja $W \subseteq M_n(F)$ um espaço de matrizes nilpotentes de característica menor ou igual a p. Então, se F for infinito,*

$$\dim(W) \leq np - \frac{p(p-1)}{2}.$$

Demonstração. A partição de Jordan de qualquer matriz em W é majorada por $(p+1, 1, \ldots, 1)$ (o que se pode ver facilmente, por exemplo, por indução, pensando na forma normal de Jordan das matrizes de W), e portanto o mesmo acontece com a partição de Jordan de W. A conjugada desta partição é $(n-p, 1, \ldots, 1)$, e neste caso o limite superior fornecido pelo teorema é $np - (p(p-1))/2$. □

Corolário 4.10 *Sejam α e β duas partições de n, com $\alpha \preceq \beta$, $\alpha \neq \beta$. Então, se F for infinito, o limite superior para a dimensão de um espaço de matrizes nilpotentes de $M_n(F)$ com partição de Jordan α é menor do que este limite para um espaço de matrizes nilpotentes com partição de Jordan β.*

Demonstração. Basta aplicar o resultado da proposição 4.3. □

□ □
□ □
□ □

Bibliografia

[BC] R. Brualdi e K. Chavey, Linear spaces of Toeplitz and Nilpotent matrices, *Jour. of Comb. Theory* 63 (1), 65–78; 1993.

[BH] J. Barría e P. R. Halmos, Vector basis for two commuting matrices, *Linear and Mult. Algebra* 27, 147–157; 1990.

[Bl] T. S. Blyth, *Module theory · An approach to linear algebra* · Second edition, Clarendon Press · Oxford; 1990.

[Bo1] N. Bourbaki, *Elements of Mathematics · Algebra II · Chapters 4–7*, Springer Verlag; 1988.

[Bo2] N. Bourbaki, *Éléments de Mathématique · Topologie Générale*, Hermann · Paris; 1971.

[Br] William C. Brown, *Matrices over commutative rings*, Marcel Dekker Inc; 1993.

[Co] R. C. Courter, Maximal commutative subalgebras of K_n at exponent three, *Linear Algebra and Appl.* 6, 1–11; 1973.

[Ga] F. R. Gantmacher, *The Theory of Matrices* · Vol. 1, Chelsea Publshing Company; 1960.

[Ge1] M. Gerstenhaber, On nilalgebras and linear varieties of nilpotent matrices I, *Amer. J. Math.* 80, 614–622; 1958.

[Ge2] M. Gerstenhaber, On nilalgebras and linear varieties of nilpotent matrices II, *Duke Math. J.* 27, 21–31; 1960.

[Ge3] M. Gerstenhaber, On nilalgebras and linear varieties of nilpotent matrices III, *Ann. of Math.* 70, 167–205; 1959.

[Ge4] M. Gerstenhaber, On dominance and varieties of commuting matrices, *Annals of Math.* 73, 324–348; 1961.

[Ge5] M. Gerstenhaber, On nilalgebras and linear varieties of nilpotent matrices IV, *Ann. of Math.* 75, 382–418; 1962.

[Gu] R. Guralnick, A note on commuting pairs of matrices, *Linear and Mult. Algebra* 31, 71–75; 1992.

[GW] K. Gruenberg e A. Weir, *Linear Geometry*, Springer Verlag; 1977.

[JK] G. James e A. Kerber, The Representation Theory of the Symmetric Group, Encyclopedia of Mathematics and its Applications · Vol 16 · Addison-Wesley Publishing Company; 1981.

[Ke] J. L. Kelley, *General Topology*, D. Van Nostrand Company, Inc; 1955.

[La] S. Lang, *Algebra* · Third edition, Addison-Wesley Publishing Company; 1993.

[Lf] T. J. Laffey, The minimal dimension of maximal commutative matrix algebras, *Linear Algebra And Appl.* 71, 199–212; 1985.

[Lz] Susan Lazarus, *Dimensions of commutative matrix algebras*, Ph.D. Thesis; 1991.

[MOR] B. Mathes, M. Omladič e H. Radjavi, Linear spaces of nilpotent matrices, *Linear Algebra and Appl.* 149, 215–225; 1991.

[MT] T. Motzkin e O. Taussky, Pairs of matrices with property L. II, *Trans. Amer. Math. Soc.* 80, 387–401; 1965.

[Ne] M. Neubauer, The variety of pairs of matrices with rank $(AB - BA) \leq 1$, *Proc. Amer. Math. Soc.* 105, 787–792; 1989.

[Ra] H. Radjavi, The Engel–Jacobson theorem revisited, *Journal of Algebra* 111, 427–430; 1987.

[Re] M. Reid, *Undergraduate Algebraic Geometry*, Cambridge University Press; 1988.

[Si] J. P. Dias da Silva, *Notas do curso de Álgebra Comutativa e Introdução à Geometria Algébrica*; 1989.

[ST] D. Suprunenko e R. Tyshkevich, *Commutative matrices*, Academic Press · New York; 1968.

[Wa] A. R. Wadsworth, The algebra generated by two commuting matrices, *Linear and Mult. Algebra* 27, 159–162; 1990.

□ □
□ □
□ □