sebastião antônio josé filho.

" SÔBRE A REDUÇÃO DE OPERADORES LINEARES E DE MATRIZES A FORMAS CANÔNICAS. "

Dissertação apresentada ao Instituto de Matemática, Estatística e Ciência da Computação, da Universidade Estadual de Campinas, para obtenção do título de Mestre em Ciências (modalidade Matemática).

- CAMPINAS -

1 972.

UNICAMP
BIBLIOTECA CENTRAL

A meus pais.

PREFÁCIO

Nesta dissertação apresentamos algumas reduções de operadores lineares e de matrizes a formas normais ou canônicas. A menos que se especifique o contrário, V denota um K-espaço vetorial de dimensão finita e L, um operador linear sôbre V.

No capítulo I, apresentamos alguns conceitos básicos que nem sempre são vistos em um primeiro curso de Álgebra Linear. Nossa pretenção, ao introduzirmos êsse capítulo I, bem como os apêndices A e B, é que êste trabalho possa ser entendido por principiante. O teorema central do capítulo I é o teorema de Hamilton-Cayley.

No capítulo II, estudamos as formas triangular e diagonal, bem como um primeiro teorema de decomposição de V em soma direta de subespaços L-invariantes. Os teoremas centrais dêsse capítulo são o teorema da decomposição primária e o teorema de diagonalização, no qual enunciamos várias condições necessárias e suficientes para que um operador linear L seja diagonalizável.

O capítulo III trata de matrizes cujos elementos são polinômios de K[X] (anel dos polinômios na indeterminada X). A importante teoria dos divisores elementares é tratada, visando aplicá-la no capítulo IV desta dissertação. Métodos de obtenção de divisores elementares são considerados para certas matrizes polinomiais particulares. Terminamos êsse capítulo com um critério para semelhança de matrizes e um corolário, que usaremos no capítulo IV.

O capítulo IV é central nesta dissertação. Nêle obtemos métodos de redução de matrizes e de operadores lineares, às formas Racional e canônica de Jordan, independentes entre si. Métodos de determinação da matriz inversível C, tal que $C^{-1}AC$ esteja na forma canônica de Jordan também são tratados.

Quero expressar aqui meus sinceros agradecimentos aos bons colegas do IMECC da UNICAMP, os quais direta ou indiretamente me influenciaram na confecção desta dissertação. Um especial agradecimento ao Prof. Dr. Jayme Machado Cardoso que leu e fêz boas sugestões a grande parte dêste trabalho.

O meu maior muito obrigado ao orientador desta dissertação, Prof. Dr. Odelar Leite Linhares e também ao Prof. Dr. Mário Tourasse

Teixeira, que me abriu as portas para estudos de pós-graduação. Ambos muito fizeram por mim e por muitos de meus colegas, quer cientificamente, como grandes homens de ciência que são, quer como exemplos raros de seres humanos, para com tudo e para com todos.

Agradeço, ainda, ao Diretor em Exercício do IMECC, Prof. Dr. Ivan de Queiroz Barros, que deu condições materiais para a realização desta dissertação.

S. A. José Filho.

C O N T E Ú D O.

CAPÍTULO I - PRÉ-REQUISITOS. .... - 1 -

1.1 Resultados Iniciais .... - 1 -
1.2 O Polinômio Mínimo - Definições Iniciais .... - 5 -
1.3 Subespaços Invariantes .... - 7 -
1.4 Autovalores e Autovetores - Polinômio Característico... - 9 -
1.5 Núcleo do Operador Linear f(L) .... -24 -

CAPÍTULO II - FORMAS TRIANGULAR E DIAGONAL. .... -28 -

2.1 Introdução .... -28 -
2.2 Forma Triangular .... -30 -
2.3 Teorema da Decomposição Primária .... -38 -
2.4 Forma Diagonal .... -46 -

CAPÍTULO III - FORMA CANÔNICA DE UMA MATRIZ POLINOMIAL ....-58 -

3.1 Operações Elementares em -matrizes .... -58 -
3.2 Forma Canônica de -matrizes .... -61 -
3.3 Polinômios Invariantes e Divisores Elementares ....-71 -
3.4 Um Critério para Semelhança de Matrizes ....-79 -

CAPÍTULO IV - AS FORMAS RACIONAL E DE JORDAN ....-84 -

4.1 Redução às formas Racional e de Jordan através dos divisores elementares .... -84 -
4.2 O Teorema da Decomposição Racional. Redução à Forma Racional através de subespaços L-cíclicos .... -92 -
4.3 Redução à Forma de Jordan através dos teoremas da Decomposição Racional e Primária .... -100-
4.4 Redução à Forma de Jordan através de autovalores ......-107-
4.5 Um Algorítmo para a Redução à Forma Canônica de Jordan-113-
4.6 Alguns Métodos de Determinação da Matriz de Transformação à Forma Canônica de Jordan .... -124-
4.7 Algumas Aplicações .... -130-

APÊNDICE A - POLINÔMIOS SÔBRE UM CORPO .... -138-

A.1 Definições Iniciais .... -138-
A.2 Algoritmo da Divisão .... -141-
A.3 Ideal em K[X]. .... -142-
A.4 Máximo Divisor Comum .... -143-

A.5 Fatoração Única .... -145-
A.6 Mínimo Múltiplo Comum .... -146-

APÊNDICE B - MISCELÂNIA .... -148-

B.1 Relação de Equivalência .... -148
B.2 Somas Diretas .... -149-
B.3 Espaço Vetorial Quociente .... -152-
B.4 Subespaços Cíclicos e Anuladores .... -156-

BIBLIOGRAFIA .... -161-

CAPÍTULO I - PRÉ-REQUISITOS

Nêste capítulo , vamos introduzir alguns conceitos que podem escapar em um primeiro curso de Álgebra Linear, curso êsse que se torna como pré-requisito para esta dissertação.

1.1 RESULTADOS INICIAIS

Sejam V e W K-espaços vetoriais (K é um corpo comutativo) - de dimensões finitas n e m, respectivamente. Uma TRANSFORMAÇÃO T de V em W é uma correspondência que associa a cada vetor v de V um único vetor Tv de W.

Uma transformação L de V em W diz-se TRANSFORMAÇÃO LINEAR - se a seguinte condição fôr satisfeita:

$$\forall x,y \in V,\ \forall a \in K \Longrightarrow L(ax+y) = aLx + Ly\ .$$

Pode-se verificar, fàcilmente, que esta condição é equivalente às seguintes condições;

i) $\forall x,y \in V \Longrightarrow L(x+y) = Lx + Ly$

ii) $\forall x \in V,\ \forall a \in K \Longrightarrow L(ax) = aLx$

Ademais, vale para tôda transformação linear as seguintes - propriedades:

TL 1 $\quad Lo = o$

TL 2 $\quad \forall v_i \in V,\ \forall a_i \in K,\ i=1,\dots,r \Longrightarrow L\left(\sum_{i=1}^{r} a_i v_i\right) = \sum_{i=1}^{r} a_i Lv_i$

Uma transformação linear L de V em V denomina-se OPERADOR LINEAR sôbre V. Nesta dissertação, nos limitaremos ao estudo de operadores lineares sôbre um espaço vetorial V de dimensão finita.

EXEMPLO 1

Seja V um K-espaço vetorial de dimensão finita n(arbitrário) Fixemos a em K. Definimos;

$$Lv = av\ ,\ \forall v \in V\ .$$

Então, L é um operador linear sôbre V. De fato, L associa a cada v de V o único vetor a v de V e $\forall x,y \in V, \forall b \in K$, $L(bx+y) = a(bx+y) = b(ax) + ay = bLx + Ly$.

Nêste exemplo 1, se tomarmos a = 1 (a=0) temos que:

$\forall v \in V$, $Lv = 1\ v = v$ ($Lv = ov = o$).

Denotamos por I(ou O) êste operador linear sôbre V. I(ou O) denomina-se <u>OPERADOR IDENTIDADE</u> ou <u>IDÊNTICO</u> (ou <u>OPERADOR NULO</u>) de V.

Dada uma base ordenada $\mathcal{B} = \{e_1, \ldots, e_n\}$ de V e n-vetores arbitrários $f_1, \ldots, f_n$ de V, existe um único operador linear L sôbre V, tal que,

$$Le_j = f_j\ ,\ j = 1, \ldots, n \qquad \text{(ver } [6])$$

Sendo L um operador linear sôbre V, cada f_i, $i=1,\ldots,n$, é um vetor de V que se expressa de modo único como combinação linear da base ordenada $\mathcal{B} = \{e_1, \ldots, e_n\}$. Seja,

$$f_j = a_{1j}\ e_1 + a_{2j}\ e_2 + \cdots + a_{nj}\ e_n;\ j=1,\ldots,n.$$

Os elementos a_{ij}, $i,j = 1,\ldots,n$ formam uma matriz $[a_{ij}]$ nxn, que denominamos <u>MATRIZ DO OPERADOR LINEAR L RELATIVAMENTE À BASE</u> $\mathcal{B}$. Tal matriz é denotada por $\mathcal{L}$ ou por $[L]_{\mathcal{B}}$. Isto é, $\mathcal{L} = [L]_{\mathcal{B}} = [a_{ij}]$ $i,j = 1,2,\ldots,n$.

Assim sendo, fixada (arbitrariamente) uma base ordenada $\mathcal{B}$ de V, todo operador linear L sôbre V determina uma única matriz $\mathcal{L} = [L]_{\mathcal{B}}$, com elementos em K. Reciprocamente, tôda matriz nxn sôbre K, determina um único operador linear L, dado por:

1) $f_j = Le_j = \sum_{i=1}^{n} a_{ij} e_i$, $x = \sum_{i=1}^{n} b_i e_i$ e $Lx = \sum_{j=1}^{n} b_j\ f_j$.

Consequentemente, podemos descrever operadores lineares através de matrizes, e matrizes são os instrumentos para o estudo de operadores lineares.

<u>EXEMPLO 2</u>

Para o operador linear L do exemplos 1, temos que, relativamente a qualquer base ordenada $\mathcal{B} = \{e_1, \ldots, e_n\}$,

$$Le_i = a\ e_i\ ,\ i = 1, \ldots, n\ ,$$

logo

$$[L]_{\mathcal{B}} = \begin{bmatrix} a & 0 & \dots & 0 \\ 0 & a & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a \end{bmatrix} \quad \text{(matriz } n\text{x}n\text{)}$$

Em particular, a matriz do operador idêntico I, e do operador nulo 0, relativamente à qualquer base ordenada de V, são a matriz identidade E e a matriz nula 0 (de ordem n) respectivamente. Isto é,

$$[I]_{\mathcal{B}} = E = [\delta]_{ij} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix} \text{ e } [0]_{\mathcal{B}} = 0 = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

Observa-se que $\delta_{ij} = \begin{cases} 1 \text{ se } i=j \\ 0 \text{ se } i \neq j \end{cases}$ é o delta de Kronecker,

e que embora usemos a mesma notação 0 (ou o) para o operador nulo e matriz nula (vetor nulo e escalar zero), não há perigo de confusão. Ainda de (1), temos que se $x = \sum_{j=1}^{n} b_j e_j$ e $Lx = \sum_{i=1}^{n} c_i e_i$, então;

$$c_i = \sum_{j=1}^{n} a_{ij} e_j \ , \ i = 1,\dots,n \ ,$$

ou, em notação matricial;

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_n \end{bmatrix}$$

Concluimos assim que;

Sendo $[L]_{\mathcal{B}}$, as coordenadas do vetor Le_i são os elementos da coluna i de $\mathcal{L}$, i = 1,2,...,n. Por outro lado, a transformação das coordenadas de x para as Lx (na mesma base ordenada $\mathcal{B}$), envolve as linhas de $\mathcal{L}$.

Sendo operadores lineares transformações (ou funções) de V em V, é natural que se defina a adição $S=L+T$, de operadores lineares, e o produto $P = \lambda L$, de um operador linear por um escalar, ponto a ponto no domínio V. Sejam L e T operadores lineares sôbre V e

$\lambda \in K$. Define-se:

$$Sx = Lx + Tx \quad , \quad \forall x \in V.$$

$$Px = \lambda Lx \quad , \quad \forall x \in V.$$

Fàcilmente se verifica que S e P são operadores lineares sôbre V. Chamamos S de SOMA dos operadores lineares L e T e P de PRODUTO de L pelo escalar λ.

Denotaremos por $\mathcal{L}(V)$ o conjunto de todos os operadores lineares L sôbre V. Com a adição (+) e produto por escalar (.), definidos acima, pode-se verificar que $\{\mathcal{L}(V), + ; K, .\}$ é um K-espaço vetorial de dimensão n^2 (ver [6]), denominado ESPAÇO VETORIAL DOS OPERADORES LINEARES SÔBRE V.

A composição de dois operadores lineares L e T sôbre V, nessa ordem, é a transformação M de V em V, definida por:

$$Mx = L(Tx) \quad , \quad \forall x \in V.$$

É claro que M é um operador linear sôbre V, que denotamos por L T. Em particular podemos formar as potências de um operador linear L sôbre V, pondo-se:

$L^1 = L, L^2 = LL, \ldots, L^r = L^{r-1}L$. Define-se $L^0 = I$, desde que $L \neq 0$.

Valem as propriedades:

$L^{m+n} = L^m L^n$ e $(L^m)^n = L^{m.n}$, $\forall n, m \in N$.

Pode-se verificar ainda, que fixado uma base ordenada $\mathcal{B}$ de V, então:

$$[S]_{\mathcal{B}} = [L+T]_{\mathcal{B}} = [L]_{\mathcal{B}} + [T]_{\mathcal{B}}$$

$$[P]_{\mathcal{B}} = [\lambda L]_{\mathcal{B}} = \lambda [L]_{\mathcal{B}}$$

Chegamos assim, à importante conclusão, de que a correspondência 1 à 1, entre $\mathcal{L}(V)$ e o espaço vetorial $M_n(K)$ das matirzes nxn sôbre K a que nos referimos acima, preserva a adição e o produto por escalar. Então, essa correspondência é um "isomorfismo" entre $\mathcal{L}(V)$ e $M_n(K)$.

Seja $p(X) = a_0 + a_1 X + \ldots + a_r X^r$ um polinômio arbitrário de $K[X]$. Defini-se o polinômio p(L) no operador linear L, associa do ao polinômio $p(X)$, pondo-se:

$$p(L) = a_0 I + a_1 L + \dots + a_r L^r \quad (L \neq 0) .$$

Anàlogamente, o polinômio na matriz A, associado ao polinômio $p(X)$ é dado por:

$$p(A) = a_0 E + a_1 A + \dots + a_r A^r .$$

Observa-se que é possível definir-se outras funções na matriz A, tais como, exp. A, sen A, etc.

Sendo p e q polinômios de $K[X]$, $s = p + q$ e $m = p.q$, então têm-se que:

$$s(L) = p(L) + q(L) \quad e \quad m(L) = p(L).\, q(L).$$

Segue daí, que:

$$p(L)\, q(L) = q(L)\, p(L), \ \forall p, q \in K[X].$$

Sendo $\mathcal{B} = \{e_1, \dots, e_n\}$ uma base de V, os vetores $Le_1, \dots, Le_n$ geram um subespaço de V, denominado IMAGEM de V através de L, o qual denotamos por $L(V)$. A dim $L(V) = \rho(L)$ chamaremos PÔSTO ou CARACTERÍSTICA de L. O conjunto $N(L) = \{v \in V: Lv = 0\}$ é um subespaço de V, denominado NÚCLEO ou ESPAÇO NULO do operador linear L. A dim $N(L) = \nu(L)$ denomina-se NULIDADE de L. Pode-se verificar que:

$$\rho(L) + \nu(L) = \dim V = n. \qquad \text{(ver [6]).}$$

Finalizamos esta secção, com um resultado importante. Sejam $\mathcal{B} = \{e_1, \dots, e_n\}$ e $\mathcal{F} = \{f_1, \dots, f_n\}$ bases ordenadas de V. Seja C a matriz que muda a base $\mathcal{B}$ para a base $\mathcal{F}$. Então, para todo operador linear L sôbre V, têm-se que:

$$[L]_{\mathcal{F}} = C^{-1} [L]_{\mathcal{B}}\, C \qquad \text{(ver [6]).}$$

1.2 O POLINÔMIO MÍNIMO

Seja V um K-espaço vetorial de dimensão finita $n \geq 1$. Seja L um operador linear sôbre V. Existe algum polinômios não nulo, f em $K[X]$, tal que $f(L) = 0$?

Como citamos em 1.1, dim $\mathcal{L}(V) = n^2$, assim sendo os operadores lineares $I = L^0, L, L^2, \dots, L^{n^2}$, ($L \neq 0$, caso $L=0$ é trivial) são linearmente dependentes logo existe uma combinação linear:

$$a_0 I + a_1 L + \ldots + a_{n^2} L^{n^2} = 0 ,$$

onde nem todos os coeficientes $a_0, a_1, \ldots, a_{n^2}$ são nulos. Basta tomarmos $f(X) = a_0 + a_1 X + \ldots + a_{n^2} X^{n^2}$ e teremos que $f(L) = 0$ e f não é o polinômio nulo. Isto responde afirmativamente à pergunta proposta.

Quando $f(L) = 0$, diz-se que o operador linear L é um ZERO do polinômio f de $K[X]$.

Seja $\mathcal{Z}$ o conjunto de todos os polinômios de $K[X]$, dos quais L é zero. Já vimos que $\mathcal{Z} \neq \emptyset$ e contém polinômios não nulos. Verificaremos que $\mathcal{Z}$ é um ideal de $K[X]$.

i) $\forall f, g \in \mathcal{Z} \Longrightarrow f(L) = g(L) = 0$;
$f(L) + g(L) = 0 + 0 = 0 \Longrightarrow (f+g) \in \mathcal{Z}$.

ii) $\forall f \in \mathcal{Z}$ e $\forall h \in K[X]$, $f(L)h(L) = 0\, h(L) = 0 \Rightarrow f.h \in \mathcal{Z}$.

Sendo $\mathcal{Z}$ um ideal de $K[X]$, $\mathcal{Z}$ é um ideal principal (ver teorema 1 de A.3). Seja m o polinômio unitário de $K[X]$ que gera $\mathcal{Z}$, isto é, todo polinômio de $\mathcal{Z}$ é um múltiplo de m. Demonstrou-se assim o seguinte teorema:

Teorema 1

"Seja $L(L \neq 0)$ um operador linear sôbre V. Existe um único polinômio m em $K[X]$, tal que:

1. m é unitário, isto é, seu coeficiente dominante (coeficiente do têrmo de maior grau) é 1.

2. m é o polinômio de menor grau, tendo L por zero e é divisor de qualquer outro polinômio de $K[X]$, que tem L por zero".

O polinômio m, dado, pelo teorema 1, denomina-se POLINÔMIO MÍNIMO do operador linear L. Obrva-se que o polinômio mínimo de qualquer operador linear $L(L \neq 0)$, não é identicamente nulo, pois é unitário e seu grau deve ser maior que zero (isto é, $gr(m) > 0$), caso contrário $m(X) = 1 \Rightarrow m(L) = I \neq 0$.

Em virtude do isomorfismo entre os K-espaço vetoriais $M_n(K)$ e

$\mathcal{L}(V)$, temos que $f(A) = 0$ se, e sòmente se, $f(L) = 0$, onde $A = [L]_{\mathcal{B}}$ para alguma base ordenada $\mathcal{B}$ de V. Seja m o polinômio unitário, de menor grau, tendo A como um zero (isto é, $m(A) = 0$) e dividindo qualquer outro polinômio que tenha A por zero. Então, m é o POLINÔMIO MÍNIMO DA MATRIZ A. Temos assim, o seguinte teorema;

Teorema 2

"Um operador linear L sôbre V e sua matriz $[L]_{\mathcal{B}}$, relativa à qualquer base ordenada $\mathcal{B}$ de V, possuem o mesmo polinômio mínimo."

1.3 SUBESPAÇOS INVARIANTES

Seja L um operador linear sôbre V, e W um subespaço de V. É claro que o conjunto;

$$L(W) = \{v \in V : \exists\, w \in W : \; Lw = v\},$$

é um subespaço de V, denominado IMAGEM do subespaço W através de L. Diz-se que um subespaço W de V é INVARIANTE relativamente a L (ou é L-INVARIANTE), se $L(W) \subset W$, isto é, a imagem de todo vetor de W, através de L, pertence a W.

EXEMPLO 3

Os subespaços triviais, V e $\langle 0\rangle$, são invariantes relativamente a qualquer operador linear sôbre V. O mesmo acontece com o núcleo de L,

$$N(L) = \{v \in V : Lv = o\}.$$

EXEMPLO 4

Seja $V = \mathbb{R}^3$, o $\mathbb{R}$-espaço dos termos ordenados de números reais. Seja L uma rotação em tôrno de um eixo que passa pela origem. Os subespaços L-invariantes de V são:

i) os triviais, $\langle o\rangle$ e V

ii) o eixo de rotação (unidimensional)

iii) o plano perpendicular ao eixo de rotação e que passa pela origem (bidimensional).

Sendo W um subespaço de V L-invariante, é natural que L induza um operador linear L_W em W, definido por:

$$L_W w = L w \quad , \quad \forall w \in W. \tag{2}$$

A êsse operador linear, denominamos <u>RESTRIÇÃO</u> de L a W. Diz-se que L_W é o operador <u>INDUZIDO</u> em W por L. Da própria definição de L_W, vê-se que, em geral $L_W \neq L$, já que como transformações (funções) - êles possuem diferentes domínios.

Seja W um subespaço L-invariante de V. Seja $\{w_1,\ldots,w_r\}$ uma base ordenada de W, a qual completamos, formando uma base ordenada $\mathcal{B} = \{w_1,\ldots,w_r, v_1,\ldots,v_{n-r}\}$ de V. Então, temos;

$$Lw_i = L_W w_i = a_{i1} w_1 + a_{i2} w_2 + \ldots + a_{ir} w_r, \quad i = 1,2,\ldots,r.$$

$$Lv_j = b_{j1} w_1 + \ldots + b_{jr} w_r + c_{j1} v_1 + \ldots + c_{jn-r} v_{n-r}, \quad j=1,\ldots,n-r.$$

Logo, a matriz $[L]_{\mathcal{B}}$ tem a forma em "blocos" seguintes:

$$[L]_{\mathcal{B}} = \left[\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right].$$

onde A é a matriz de L_W relativamente à base ordenada $\{w_1, w_2,\ldots,w_r\}$

<u>PROPOSIÇÃO 1</u>

"Seja W um subespaço de V invariante relativamente à L. Então, para todo polinômio f de $K[X]$ e todo vetor $w \in W$, têm-se que:

(3) $\qquad f(L_W)(w) = f(L)(w)$" .

Demostração.

Usando-se (2) e fato de que $Lw \in W$, para todo $w \in W$, vem que:

$$L_W^2 w = L_W (L_W w) = L_W(Lw) = L(Lw) = L^2 w.$$

Analogamente, ou sendo maior preciso, usando-se indução, conclui-se que $L_W^r w = L^r w$, $\forall r \in N$

Desde que as igualdades acima, podem ser multiplicados por elementos de K e adicionadas membro a membro, obtêm-se (3).

Sendo $f(L_W)$ um operador linear de W, $f(L_W)(w) \in W$, para todo $w \in W$. Por (3), temos que $f(L)(w) \in W$, para todo $w \in W$, isto é, W é $f(L)$-invariante. Demonstrou-se assim, o seguinte teorema:

<u>TEOREMA 3</u>

"Seja W um subespaço de V, L-invariante.

Então, para todo polinômio f de $K[X]$, têm-se que W é também invariante, relativamente ao operador linear f(L) sôbre V".

1.4 AUTOVALORES E AUTOVETORES. POLINÔMIO CARACTERÍSTICO.

Na presente secção, subespaços L-invariantes de dimensão um desempenham papel de destaque.

Seja W um subespaço unidimensional de V gerado pelo vetor $v \neq 0$, isto é,

$$W = \langle v \rangle = \{a\,v : a \in K\} .$$

É claro que uma condição necessária e suficiente, para que W seja L-invariante, é que:

$$\forall y \in W \,,\ \exists \lambda \in K : \ Ly = \lambda y \,.$$

Dizemos que um vetor $v \neq o$ de V é um AUTOVETOR de L, se v satisfaz à relação.

$$L\,v = \lambda\, v,$$

para algum λ de K. O escalar λ denomina-se AUTOVALOR de L.

Outras denominações para autovalores são: valores característicos, raízes características, raízes latentes, valores próprios, valores espectrais, etc.

Nesta dissertação, usaremos as nomenclaturas "autovalores" e "valores característicos".

TEOREMA 4

"As seguintes condições são equivalentes:

1. λ é um autovalor de L.

2. $(L - \lambda I)$ é um operador linear não injetor.

3. $\det(\mathcal{L} - \lambda E) = o$, onde $\mathcal{L} = [L]_{\mathcal{B}}$ e $\mathcal{B}$ é uma base ordenada (arbitrária) pré-fixada de V".

Demonstração.

1. $\Longrightarrow$ 2.

Sendo λ um autovalor de L, existe um vetor $v \neq 0$, tal que

$$Lv = \lambda v \quad \text{ou} \quad (L - \lambda I)v = 0.$$

Então $L - \lambda I$ não é injetor pois $v \neq 0$ e $(L - \lambda I)v = (L - \lambda I)\,0 = 0$.

2. $\Longrightarrow$ 3.

Sendo $L - \lambda I$ não injetor, êle é um operador linear não inver sível. Logo, a matriz de $L-\lambda I$, relativamente a qualquer base ordenada $\mathcal{B}$ de V não é inversível, isto é,

$$\det(\mathcal{L}-\lambda[I]_{\mathcal{B}}) = \det(\mathcal{L}-\lambda E) = 0.$$

3. $\Longrightarrow$ 1.

Consideremos o sistema de equações lineares homogêneas:

$$(\mathcal{L}-\lambda E)\,[x] = 0\ ,$$

onde $[x]$ é a matriz coluna (nx1), formada pelas coordenadas de um vetor v, na base ordenada $\mathcal{B}$, pré-fixada.

Como $\det(\mathcal{L}-\lambda E) = 0$, o sistema acima tem uma solução $[\bar{x}]$ - não trivial. Pelo isomorfismo entre $\mathcal{J}(V)$ e $M_n(K)$, têm-se que:

$$(L-\lambda I)\,\bar{v} = 0 \Longrightarrow L\bar{v} = \lambda\bar{v} \quad e \quad \bar{v} \neq 0;$$

logo λ é um autovalor de L.

Fixada uma base ordenada $\mathcal{B}$ de V, o operador linear $L-\lambda I$ é inversível se, sòmente se, a matriz $\mathcal{L}-\lambda E$ o for.

Dêsse fato e do teorema anterior, podemos colocar a seguinte definição:

Seja A uma matriz nxn sôbre o corpo K.

Um AUTOVALOR de A é um escalar λ de K, tal que a matriz $(A-\lambda E)$ é singular, isso é, tal que $\det(A-\lambda E) = 0$.

Formemos a matriz XE-A, cujos elementos são polinômio de $K[X]$, de grau no máximo 1. Seja $A = [a_{ij}]$, nxn. Consideremos o polinômio:

$$(4) \qquad c(X)=\det(XE-A) = \begin{vmatrix} X-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & X-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & X-a_{nn} \end{vmatrix}$$

O polinômio c , dado por (4), denomina-se POLINÔMIO CARACTERÍSTICO da matriz A.

Como $\det(XE-A) = 0$ se, e somente se, $\det(A-XE) = 0$, vem que $\lambda \in K$ é um autovalor de A se, e somente se, $c(\lambda) = 0$, isto é, λ é um zero do polinômio característico c de A. Logo, o problema de se determinar todos os autovalores de uma matriz A, reduz-se ao problema de se encontrar tôdas as raízes do polinômio característico de A. Quando dim V = 1,2,3 ou 4 (veremos a seguir que gr(c) = dim V) existem fórmulas envolvendo radicais, que nos determinam as raízes da EQUAÇÃO CARACTERÍSTICA de A:

$$c\,(X) = \det\,(XE - A) = 0\,.$$

Tal não é o caso, para $n \geq 5$; pois não existe fórmulas gerais, para se resolver equações de grau $n \geq 5$. Existem métodos numéricos, para obtenção de raízes (ou aproximação de raízes), da equação $c(X) = 0$. Infelizmente, não temos espaço nessa dissertação, disponível ao desenvolvimento de tais métodos numéricos.

Da definição de determinante de uma matriz quadrada, têm-se que:

$$(5) \qquad c(X) = \det(XE-A) =$$

$$= \sum_{i_1,\dots,i_n} \operatorname{sn}(i_1,\dots,i_n)(X\delta_{1i_1} - a_{1i_1}) \dots (X\delta_{ni_n} - a_{ni_n})$$

onde $\delta_{ij} = \begin{cases} 1 \text{ se } i = j \\ 0 \text{ se } i \neq j \end{cases}$, é o delta de Kronecker

e $\operatorname{sn}(i_1,i_2,\dots,i_n)$ é o sinal da permutação:

$$\begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix} .$$

A soma do lado direito de (5), possui uma única parcela de grau n, que ocorre quando a permutação acima é a identidade, isto é, $i_1 = 1, \dots, i_n = n$. Nêsse caso, obtemos a parcela: $(X-a_{11})\dots(X-a_{nn}) = X^n + r(X)$, onde o polinômio $r(X)$ tem grau menor ou igual à $(n-1)$.

Isto nos permite afirmar, que o polinômio característico de uma matriz A, nxn sôbre um corpo K, é um polinômio unitário de grau n.

PROPOSIÇÃO 2

"Matrizes semelhantes, têm o mesmo polinômio característico"

Demonstração.

Sejam A e B matriz de $M_n(K)$, semelhantes. Isto é, existe uma matriz C inversível em $M_n(K)$, tal que:

$$B = C^{-1} A C.$$

Usando-se propriedades de determinantes e o fato de que a matriz XE comuta com qualquer matriz nxn, vem que:

$$\det(XE-B) = \det(XE-C^{-1}AC) = \det(C^{-1}XEC-C^{-1}AC) =$$
$$= \det[C^{-1}(XE-A)C] = \det C^{-1} \det(XE-A) \det C = \det(XE-A).$$

c.q.d.

Com base no último resultado de 1.1 e nessa proposição 2, define-se o <u>POLINÔMIO CARACTERÍSTICO</u> do operador linear L, como sendo o polinômio característico de uma das matrizes que representa L, relativamente à alguma base ordenada de V. Da mesma forma que para matrizes, $\lambda \in K$ é um autovalor de L se, e sòmente se, λ é um zero do polinômio característico ($c(X) = \det(XI-L)$) de L. Sendo $gr(c)=n$, sabe-se que $c(X)$ não tem mais do que n raízes. Logo L não pode ter mais do que n-autovalores. Concluimos ainda da proposição 2, que - sendo L e T, operadores lineares sôbre V, representáveis por matrizes semelhantes, então êles têm idênticos polinômios característicos.

Um corpo F, diz-se <u>ALGÈBRICAMENTE FECHADO</u>, se todo polinômio f de $F[X]$ com $gr(f) \geq 1$, puder ser colocado na forma:

$$f(X) = a(X - a_1)^{n_1} (X - a_2)^{n_2} \dots (X - a_r)^{n_r}$$

onde a, $a_1, a_2, \dots, a_r$ são elementos de F, e $n_1, n_2, \dots, n_r$ são inteiros positivos.

Muitas outras condições, podem ser usadas para se dizer que um corpo F é algèbricamente fechado. Por exemplo, se para todo f de $F[X]$ com $gr(f) \geq 1$, existe um escalar λ em F, que é um zero de f, isto é, tal que $f(\lambda) = 0$.

O Corpo $\mathbb{R}$ dos números reais não é algèbricamente fechado - pois $f(X) = X^2 + 1$ é um polinômio de $\mathbb{R}[X]$, $gr(f) = 2 \geq 1$ e não existe λ em $\mathbb{R}$, tal que $f(\lambda) = \lambda^2+1 = 0$. O teorema fundamental da Álgebra, afirma que o corpo $\mathbb{C}$, dos números complexos, é algèbricamente fechado.

TEOREMA 5

"Seja V um espaço vetorial de dimensão finita $n \geq 1$, sôbre - um corpo algèbricamente fechado F. Então, todo operador linear L sôbre V, possui no mínimo um autovetor (e um autovalor)".

Demonstração.

Como dim $V = n \geq 1$, temos que $gr(c) \geq 1$, onde c é o polinômio característico do operador linear L. Sendo F algèbricamente fechado existe um λ em F, tal que, $c(\lambda) = 0$. Então, λ é um autovalor de L. Logo L possui pelo menos um autovalor e um autovetor.

Observações.

1. Sendo V um espaço vetorial sôbre um corpo K, não algèbricamente fechado, pode ocorrer que um operador linear L sôbre V, não tenha nenhum autovalor em K (ou autovetor em V). Consideremos o $\mathbb{R}$-espaço vetorial $\mathbb{R}^2$. Seja L o operador linear sôbre $\mathbb{R}^2$, cuja matriz relativamente à base ordenada canônica $\mathcal{B} = \{e_i = (\delta_{i1}, \delta_{i2}),\ i=1,2\}$

$$\mathcal{A} = [L]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

O polinômio característico de L (ou de $\mathcal{A}$) é:

$$c(X) = \det(XE - \mathcal{A}) = X^2 + 1 ,$$

que não tem raiz em $\mathbb{R}$, consequentemente, L não possui nenhum autovalor em $\mathbb{R}$ (ou autovetor em $\mathbb{R}^2$).

2. A demonstração do teorema 5, permanece válida, para a restrição de L a um subespaço invariante de V, relativamente à L.

3. Seja $v \neq 0$ um autovetor de L. Então para todo $\mu \in K$, $\mu \neq 0$, o vetor μv também é um autovetor de L relativo ao mesmo autovalor. De fato, suponhamos que $Lv = \lambda v$, logo $L(\mu v) = \mu Lv = \mu(\lambda v) = \lambda(\mu v)$. É o subespaço gerado por $v(\langle v\rangle)$, que deve ser levado em considera - ção. Veremos mais adiante, que ao mesmo autovalor λ, podem estar

associados dois ou mais autovetores linearmente independentes.

4. Sendo λ um autovalor de L, o conjunto dos autovetores de L associados à λ, juntamente com o vetor nulo, é um subespaço de V. Para verificarmos tal fato, lembramos que por definição, v é um autovetor de L correspondente à λ ou $v = 0$ se, e sòmente se, v é solução de:

$$(6) \qquad (\mathcal{L} - \lambda E)\,[x] = [0],$$

onde $\mathcal{L} = [L]_{\mathcal{B}}$ e $\mathcal{B}$ é uma base ordenada, pré-fixada de V, $[x]$ é a matriz $n \times 1$ das coordenadas de v na base $\mathcal{B}$ e $[0]$ é a matriz nula $n \times 1$. O conjunto das soluções do sistema de equações lineares homogêneos (6), forma um subespaço vetorial de V. De fato, sejam x e y soluções de (6), então:

i) $(\mathcal{L} - \lambda E)\,[x+y] = (\mathcal{L} - \lambda E)\,[x] + (\mathcal{L} - \lambda E)\,[y] = [0]$.

ii) $\forall a \in K,\ (\mathcal{L} - \lambda E)\,[a\,x] = (\mathcal{L} - \lambda E)\,[x]\,a = [0]\,a = [0]$.

Segue ainda daí que a dimensão do subespaço de autovetores - (juntamente com o vetor nulo) associados ao autovalor λ de L é n menos a característica da matriz $(\mathcal{L} - \lambda E)$; ou equivalentemente da matriz $(\lambda E - \mathcal{L})$.

Sendo λ um autovalor de L, o operador linear $\lambda I - L$ não é injetor. Ao núcleo do operador linear $\lambda I - L$, denominamos ESPAÇO PRÓPRIO ou AUTOESPAÇO correspondente ao autovalor λ, o qual denotamos por V_λ. À dimensão do autoespaço V_λ (subespaço de V), denominamos MULTIPLICIDADE GEOMÉTRICA de λ. Como $\lambda I - L$ não é injetor, existe um vetor $v \neq 0$ em $N(\lambda I - L)$, logo a multiplicidade - geométrica de λ é ≥ 1. O subespaço de autovetores associados ao autovalor λ de L, juntamente com o vetor nulo, é justamente o espaço próprio correspondente à λ. Logo, sua dimensão (a multiplicidade geométrica de λ) é n menos a característica de $\mathcal{L}$. Observa-se ainda, que a multiplicidade geométrica de λ é igual ao maior número de autovetores linearmente independentes correspondentes ao autovalor λ.

Suponhamos que o polinômio característico $c(X) = \det(XI - L)$, de um operador linear L, possa ser fatorado na forma:

$$c(X) = \prod_{i=1}^{r} (X - \lambda_i)^{a_i},$$

com $\lambda_i \neq \lambda_j$ se $i \neq j$ e $a_1, \ldots, a_r$ inteiros extritamente positivos (isto sempre é possível quando V é um espaço vetorial sôbre um corpo F, algèbricamente fechado).

Chamamos <u>MULTIPLICIDADE ALGÉBRICA</u> do autovalor λ_i, ao inteiro - a_i, $i = 1,2,\ldots,r$. O conjunto $\sigma(L)$ dos autovalores de L, denomina-se <u>ESPECTRUM</u> de L.

<u>EXEMPLO 5</u>

Sejam $V = \mathbb{R}^2$, $\mathcal{B}_2 = \{e_i = (\delta_{i1}, \delta_{i2}),\ i = 1,2\}$ a base canônica de $\mathbb{R}^2$ e L o operador linear sôbre $\mathbb{R}^2$ dado por:

$$\mathcal{L} = [L]_{\mathcal{B}} = \begin{bmatrix} 0 & 6 \\ 1 & 1 \end{bmatrix}$$

Então, o polinômio característico de L é:

$$c(X) = \det \begin{bmatrix} X & -6 \\ -1 & X-1 \end{bmatrix} = X(X-1) - 6 = X^2 - X - 6 = (X+2)(X-3).$$

Logo os autovalores de L (zeros de $c(X)$) são 3 e - 2. A multiplicidade algébrica do autovalor 3 (ou -2) é 1.

Desde que $c(3) = 0$, $\det(3E - \mathcal{L}) = 0$ e a característica da matriz $(3E - \mathcal{L})$ é menor que 2. Como $\det[3] = 3$, a característica de $(3E - \mathcal{L})$ é 1. Então, a multiplicidade geométrica do autovalor 3 é $2 - 1 = 1$, isto é, o autoespaço associado ao autovalor 3 é gerado por um autovetor de L, correspondente ao autovalor 3. Procuremos - um tal autovetor. Seja $v = a_1 e_1 + a_2 e_2$. Então

$$\mathcal{L}[x] = \begin{bmatrix} 0 & 6 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 6a_2 \\ a_1 + a_2 \end{bmatrix}$$

e $\lambda[x] = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$. Pela definição, λ é um autovalor de L se, e sòmente se $\mathcal{L}[x] = \lambda[x]$ $([x] \neq [0])$. Como 3 é um autovalor de L vem

que: -
$$\begin{bmatrix} 6 & a_2 \\ a_1 + a_2 \end{bmatrix} = \begin{bmatrix} 3 & a_1 \\ 3 & a_2 \end{bmatrix} \Longleftrightarrow \begin{cases} a_1 = 2a_2 \\ a_1 = 2a_2 \end{cases}$$
e uma solução não trivial é $v = 2\ e_1 + e_2$.

Logo, o autoespaço associado ao autovalor 3 é
$$V_3 = \langle 2e_1 + e_2 \rangle .$$

Anàlogamente, conclui-se que o espaço próprio correspindente ao autovalor -2 é
$$V_{-2} = \langle 3e_1 - e_2 \rangle ,$$

e a multiplicidade geométrica do autovalor -2 é 1. Nêsse caso $\sigma(L) = \{3, -2\}$.

EXEMPLO 6

Seja L um operador linear sôbre $\mathbb{R}^3$, cuja matriz relativa à base canônica $\mathcal{B} = \{e_i = (\delta_{i1}, \delta_{i2}, \delta_{i3}),\ i = 1,2,3\}$ é
$$\mathcal{L} = [L]_{\mathcal{B}} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

O polinômio característico de L é:
$$c(X) = \det(X\mathbb{I} - \mathcal{L}) = (X-1)^2\ (X-2).$$

Então o espectrum de L é $\sigma(L) = \{1, 2\}$. A multiplicidade algébrica do autovalor 1 é 2 e do autovalor 2 é 1.

Os autovetores de L associados ao autovalor 1, tem por coordenadas relativamente à base canônica, as soluções não triviais do sistema
$$(1.\mathbb{I} - \mathcal{L})\ [x] = [0] \Longleftrightarrow \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Longleftrightarrow \begin{cases} x_2^{*} = 0 \\ x_3 = 0 \end{cases} .$$

Logo, $e_1 = (1,0,0)$ é um autovetor de L associado ao autovalor 1 e o autoespaço associado ao autovalor 1 é $V_1 = \langle (1,0,0) \rangle$.

Anàlogamente, pode-se concluir que o espaço próprio corresponde ao autovalor 2 é $V_2 = \langle (0,0,1) \rangle$.

Daí, a multiplicidade geométrica do autovalor 2 (ou 1) é 1.

EXEMPLO 7

Nas condições do exemplo 6, seja T o operador linear dado por:

$$\mathcal{T} = [T]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Então,

$$c(X) = \det(XE - \mathcal{T}) = (X-1)^2 (X-2).$$

$\sigma(T) = \{1,2\}$. A multiplicidade algébrica do autovalor 1 é 2 e do autovalor 2 é 1.

Os autovetores de T associados ao autovalor 1, são determinados por:

$$(1.E - \mathcal{T})\ [x] = [0] \Leftrightarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow \{x_3 = 0$$

Logo $e_1 = (1,0,0)$ e $e_2 = (0,1,0)$ são autovetores de T associados ao autovalor 1. O espaço próprio associado ao autovalor 1 é

$$V_1 = <(1,0,0),\ (0,1,0)>.$$

Vê-se fàcilmente, que $V_2 = <(0,0,1)>$.

Observação.

Nos exemplos 6 e 7 temos dois operadores lineares com mesmos polinômios característicos, logo mesmo espectrum e as multiplicidades algébricas de cada autovalor também iguais, mais êles são diferentes, isto é, êles não são semelhantes (mostraremos isso, no exemplo 8 de 2.4). Assim, a recíproca da proposição 2, não é válida.

TEOREMA 6

"Seja L um operador linear sôbre V, espaço vetorial sôbre um corpo F algèbricamente fechado. Seja λ um autovalor de L, com multiplicidade algébrica $\underline{a}$ e multiplicidade geométrica g.

Então, $g \leq a$".

Demonstração.

Seja $\{e_1, \ldots, e_g\}$ uma base ordenada para o autoespaço correspondente à λ, isto é, para o núcleo N do operador linear $(\lambda I - L)$, que por hipótese tem dimensão $g \leq n$, pois é um subespaço de V.

Completamos essa base ordenada de N, é uma base ordenada $\mathcal{B} = \{e_1, \ldots, e_g, e_{g+1}, \ldots, e_n\}$ de V. Temos que:

$$(7)\ \mathcal{L} = [L]_{\mathcal{B}} = \left[\begin{array}{cccc:c} \lambda & 0 & \cdots & 0 & \\ 0 & \lambda & \cdots & 0 & \\ \vdots & & \ddots\ {\scriptstyle g\text{-vêzes}} & \vdots & A \\ 0 & 0 & \cdots & \lambda & \\ \hdashline & & 0 & & B \end{array}\right],$$

onde 0 é a matriz nula $(n-g)\mathrm{x}g$. De fato, sendo $\{e_1, \ldots, e_g\}$ uma base para o núcleo N do operador linear $(\lambda I-L)$, temos que:

$$(\lambda I-L)\, e_i = 0 \iff Le_i = \lambda e_i\ , \quad i = 1, 2, \ldots, g.$$

De (7), vê-se que o polinômio característico de L, isto é, $\det(\lambda E-\mathcal{L})$, tem $(X-\lambda)$ como fator no mínimo g-vêzes (o desenvolvimento do determinante por Laplace, g-vêzes, sempre pela primeira coluna, mostra êsse fato). Então, da definição de multiplicidade algébrica, têm-se que $g \leq a$.

Observamos que no exemplo 6, ocorreu $g < a$, para o autovalor 1.

TEOREMA 7

"Sejam $\lambda_1, \lambda_2, \ldots, \lambda_r$, autovalores do operador linear L, distintos dois à dois, e $v_1, v_2, \ldots, v_r$ os autovetores respectivamente associados. Então, os vetores $v_1, v_2, \ldots, v_r$ são linearmente independentes".

Demonstração.

Consideremos uma combinação linear arbitrária dos vetores $v_1, \ldots, v_r$:

$$(8) \qquad \sum_{i=1}^{r} a_i v_i = 0.$$

Aplicando-se repetidas vêzes ($(r-1)$-vêzes) o operador linear L à (8), obtemos o seguinte sistema de equações lineares;

$$\begin{cases} \sum_{i=1}^{r} a_i v_i = 0 \\ L\left(\sum_{i=1}^{r} a_i v_i\right) = \sum_{i=1}^{r} a_i Lv_i = \sum_{i=1}^{r} a_i \lambda_i v_i = 0 \\ L^2\left(\sum_{i=1}^{r} a_i v_i\right) = L\left(\sum_{i=1}^{r} a_i \lambda_i v_i\right) = \sum_{i=1}^{r} a_i \lambda_i Lv_i = \sum_{i=1}^{r} a_i \lambda_i^2 v_i = 0 \\ \text{---} \\ \text{---} \\ L^{r-1}\left(\sum_{i=1}^{r} a_i v_i\right) = \sum_{i=1}^{r} a_i \lambda_i^{r-1} v_i = 0 \end{cases}$$

Ou escrevendo-se o sistema em notação matricial;

$$(10) \qquad A \begin{bmatrix} a_1 v_1 \\ a_2 v_2 \\ \vdots \\ a_r v_r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \text{ onde } A = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_r \\ \lambda_1^2 & \lambda_2^2 & & \lambda_r^2 \\ \vdots & \vdots & & \vdots \\ \lambda_1^{r-1} & \lambda_2^{r-1} & & \lambda_r^{r-1} \end{bmatrix}$$

A é uma matriz de Vandermonde, cujo determinante é:

$$\det A = \prod_{j>i} (\lambda_j - \lambda_i) \; ; \; i = 1,2,\ldots,r-1$$

Por hipótese $\lambda_j \neq \lambda_i$ se $i \neq j$, logo $\det A \neq 0$, isto é, A é inversível. Aplicando-se A^{-1} à esquerda em (10), obtém-se:

$$\mathcal{I} \begin{bmatrix} a_1 & v_1 \\ a_2 & v_2 \\ \vdots & \\ a_r & v_r \end{bmatrix} = \begin{bmatrix} a_1 & v_1 \\ a_2 & v_2 \\ \vdots & \\ a_r & v_r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \Longleftrightarrow a_i v_i = 0, \; i=1,\ldots,r$$

Sendo $v_i \neq 0$, pois é um autovetor, temos que $a_i = 0$, $i = 1,2,\ldots,r$.

c.q.d.

Uma consequência imediata do teorema 7, é o seguinte corolário:

COROLÁRIO 1

"Quando um operador linear L sôbre V (dim $V=n$, finita) possui n-autovalores dois à dois distintos, os autovetores correspondentes, formam uma base de V".

Nosso próximo teorema é uma espécie de generalização do teorema 7.

TEOREMA 8

"Sejam $\lambda_1, \lambda_2, \ldots, \lambda_r$ autovalores distintos dois à dois, de um operador linear L sôbre V e

$$e^{i1}, e^{i2}, \ldots, e^{in_i}, \quad i = 1,2,\ldots,r,$$

uma família de autovetores de L, linearmente independentes, associados à λ_i. Então, a família de todos os autovetores;

$$(11) \quad e^{11}, \ldots, e^{1n_1}, e^{21}, \ldots, e^{2n_2}, \ldots, e^{r1}, \ldots, e^{rn_r},$$

é linearmente independente"

Demonstração.

Consideremos uma combinação linear nula dos vetores (11), - isto é;

$$(12) \quad \sum_{i=1}^{r} \sum_{j=1}^{n_i} a_{ij}\, e^{ij} = 0 .$$

Queremos mostrar que $a_{ij} = 0$, $i = 1,\ldots,r$ e $j = 1,\ldots,n_i$. Seja,

$$x^i = \sum_{j=1}^{n_i} a_{ij}\, e^{ij}$$

então x^i é um vetor do espaço próprio correspondente ao autovalor λ_i, pois:

$$(\lambda_i I-L)\, x^i = \sum_{j=1}^{n_i} a_{ij}\, (\lambda_i I-L)\, e^{ij} = 0 ,$$

já que $(\lambda_i I-L)\, e^{ij} = 0$. Consequente, $x^i = 0$ ou x^i é um autovetor associado à λ_i. Suponhamos que alguns x^i sejam diferentes de 0. Sem perda de generalidade (renumera-se os índices, se necessário), podemos supor que $x^i \neq 0$, $i = 1,\ldots,s$, $s \leq r$. Pelo teorema 7, - não é possível que $\sum_{i=1}^{r} x^i = 0$, pois $\lambda_1,\ldots,\lambda_s$ são distintos dois à dois, por hipótese. De (12) vem $\sum_{i=1}^{r} x_i = 0$, portanto nenhum x^i é diferente de 0, isto é,

$$0 = x^i = \sum_{j=1}^{n_i} a_{ij}\, e^{ij}\ ,\ i = 1,2,\ldots,r.$$

Por hipótese $e^{i1}, e^{i2},\ldots,e^{in_i}$ é uma família linearmente independente, logo temos que

$$a_{ij} = 0,\ i = 1,\ldots,r \quad e \quad j = 1,\ldots,n_i$$

c.q.d.

A correspondente consequência dêsse teorema, é o seguinte corolário:

<u>COROLÁRIO 2</u>

"Sejam $\lambda_1, \lambda_2,\ldots,\lambda_r$ <u>todos</u> os autovalores de L distintos dois à dois e

$$e^{i1}\ ,\ e^{i2},\ldots,e^{in_i}\ ,\ i = 1,2,\ldots,r\ ,$$

a família de <u>todos</u> os autovetores linearmente independentes de L, associados ao autovalor λ_i, tais que

$n_1 + n_2 + \ldots + n_r = n$ (n=dim V). Então,

$$\mathcal{B} = \left\{ e^{11},\ldots,e^{1n_1},\ e^{21},\ldots,e^{2n_2},\ldots,\ e^{r1},\ldots,e^{rn_r} \right\},$$

é uma base ordenada de V".

Nossa próxima meta, é a demonstração do importante teorema de Hamilton-Cayley.

<u>LEMA 1</u>

"Consideremos um polinômio f de K[X] e uma matriz A de $M_n(K)$, tais que:

$$(13) \qquad \begin{cases} f(X) = a_0 X^m + a_1 X^{m-1} + \ldots + a_{m-1} X + a_m \\ f(X)E = (XE - A)\,\mathcal{C}(X)\ , \end{cases}$$

onde $\mathcal{C}(X)$ é um polinômio na indeterminada X, cujos coeficientes são matrizes de $M_n(K)$, mais precisamente,

$$\mathcal{C}(X) = -C_0 X^{m-1} - C_1 X^{m-2} - \ldots - C_{m-1}\ .$$

Então, $f(A) = 0$".

Demonstração.

Das hipóteses, vem que; -

$$(14) \qquad (XE - A)\, \mathcal{C}(X) = AC_{m-1} + (AC_{m-2} - C_{m-1})\, X + (AC_{m-3} - C_{m-2})\, X^2 + \dots - C_o\, X^m .$$

De (13) e (14), obtém-se o seguinte sistema:

$$(15) \begin{cases} AC_{m-1} = a_m E \\ AC_{m-2} - C_{m-1} = a_{m-1} E \\ AC_{m-3} - C_{m-2} = a_{m-2} E \\ \dots\dots\dots\dots \\ \dots\dots\dots\dots \\ AC_o - C_1 = a_1 E \\ - C_o = a_o E \end{cases} \qquad (16) \begin{cases} AC_{m-1} = a_m E \\ A^2C_{m-2} - AC_{m-1} = a_{m-1} A \\ A^3C_{m-3} - A^2C_{m-2} = a_{m-2} A^2 \\ \dots\dots\dots\dots \\ \dots\dots\dots\dots \\ A^m C_o - C_1 A^{m-1} = a_1 A^{m-1} \\ - A^m C_o = a_o A^m . \end{cases}$$

Multiplicando-se as equações de (15) à esquerda, respectivamente, por E, A, $A^2, \dots, A^m$ (ver (16)) e adicionando-se as equações resultantes, membro à membro, obtemos no primeiro membro a matriz nula 0 e no segundo obtém-se $f(A) = a_m E + a_{m-1} A + \dots + a_o A^m$, isto é,

$$f(A) = 0 \text{ , como queriámos demonstrar.}$$

__TEOREMA 9__ (__TEOREMA DE HAMILTON-CAYLEY__).

"Seja c o polinômio característico de uma matriz A de $M_n(K)$ (ou do operador linear L). Então,

$$c(A) = 0 \text{ (ou } c(L) = 0\text{)" .}$$

Demonstração.

Seja A uma matriz de $M_n(K)$ e seja c seu polinômio característico. Seja $\mathcal{C}(X)$ a adjunta clássica da matriz $XE-A$. Os elementos de $\mathcal{C}(X)$, são os complementos algébricos da matriz $XE-A$, logo são polinômios em X, de grau menor ou igual à $n-1$. Pela propriedade fundamental da adjunta clássica de A, que diz:

$$A . (\text{adj } A) = (\text{adj } A) . A = (\det A)\, E$$

(para uma demonstração dessa propriedade, ver [8] pag. 136); vem que:

(17) $\qquad (XE-A)\,\mathscr{B}(X) = \det(XE-A).\,E = c(X).\,E.$

De (17), com base no lema 1, conclui-se que $c(A) = 0$, o que demonstra o teorema.

Observação.

Da difinição de polinômio mínimo de um operador linear L sôbre V (teorema 1 de 1.2), podemos reenunciar o teorema de Hamilton-Cayley, pondo-se:

"O polinômio mínimo do operador linear L, é um divisor do polinômio característica de L".

Como uma consequência do teorema de Hamilton-Cauley, obteremos nosso próximo teorema:

<u>TEOREMA 10</u>

"Os zeros dos polinômios mínimo e característico, de todo operador linear L de V, são iguais".

Demosntração.

Observamõs inicialmente, que se o polinômio característico c não tiver nenhum zero em K, o polinômio mínimo m também não tem zeros em K. De fato, seja $\lambda \in K$ um zero de m. Como m divide c (teorema de Hamilton-Cayley), têm-se que;

$$c(X) = q(X)m(X) \Longrightarrow c(\lambda) = q(\lambda)m(\lambda) = q(\lambda).\,0 = 0$$

Isto é, λ sendo zero de m será também de c.

Suponhamos agora que $\lambda \in K$ é um zero de c. Então, λ é um autovalor de L, isto é, existe um vetor $v \neq 0$ em V, tal que $Lv = \lambda v$. Vamos mostrar que para todo polinômio f de $K[X]$, $f(\lambda)$ é um autovalor de operador linear $f(L)$. de fato,

$$Lv = v,\ Lv = v,\ L^2v = L(Lv) = L(\lambda v) = \lambda Lv = \lambda^2 v,$$

$$\ldots,\ L^r v = \lambda^r v\ ;\ \forall r \in \mathbb{N}.$$

logo se $f(X) = a_0 + a_1X + \ldots + a_s X^s$, temos que:

$$f(L)v = (a_0 + a_1\lambda + \ldots + a_s \lambda^s)v = f(\lambda)v.$$

Em particular, sendo $f(L) = 0$, $f(\lambda)$ é um autovalor do operador nulo, logo é o escalar. Assim sendo, no caso em que $f \equiv m$, temos que $m(L) = 0$ e def $m(\lambda) = 0$. O que demonstra o teorema .

1.5 NÚCLEO DO OPERADOR LINEAR $f(L)$.

Trataremos aqui, de uma classe especial de subespaços invariantes de V. Seja L um operador linear sôbre V e $f \in K[X]$. O conjunto;

(18) $N(f) = \{v \in V: f(L)v = 0\}$,

é o núcleo do operador linear $f(L)$, que como sabemos, é um subespaço vetorial de V.

Sejam, $f(L)$ e $g(L)$ polinômios no operador linear L, associados aos polinômios f e g de $K[X]$. Então, sabemos de 1.1 que:

(19) $f(L)\ g(L) = g(L)\ f(L)$.

De (19) com $g(X) = X$, para todo $v \in N(f)$, vem:

$$f(L)\ [Lv] = [f(L)L]\ v = [Lf(L)]\ v = L[f(L)v] = L\ 0 = 0 \ ;$$

o que demonstra o seguinte lema ;

LEMA 2.

"O núcleo $N(f)$ do operador linear $f(L)$ sôbre V, é invariante com relação à L" .

É claro que o subespaço $N(f)$ de V, depende do particular polinômio f de $K[X]$. A seguir estudaremos as relações entre núcleos de diferentes polinômios em L.

LEMA 3

"Sejam f e g polinômio de $K[X]$, tais que, g divide f. Então, para todo operador linear L sôbre V, têm-se ;

(20) $N(g) \subset N(f)$ " .

Demonstração.

Como g divide f, existe h em $K[X]$, tal que, $f = h.g$. Logo $f(L) = h(L)\ g(L)$. Seja $v \in N(g)$, então $g(L)\ v = 0$ e

$$f(L)\ v = [h(L)\ g(L)]\ v = h(L)\ [g(L)\ v] = h(L)\ 0 = 0 \ ,$$

isto é, $v \in N(f)$, o que demonstra (20) .

LEMA 4

"Sejam f_1, $f_2,\ldots,f_r$ polinômios de $K[X]$, não todos nulos e d o m.d.c entre êsses polinômios. Então, para todo operador linear L de V, têm-se:

(21) $$N(d) = N(f_1)\cap N(f_2)\cap \ldots \cap N(f_r)\text{"}.$$

Demonstração.

Sendo d o m.d.c. de f_1, f_2, $\ldots,f_r$, temos que d divide cada um dêsses polinômios. Consequentemente, pelo lema 3, temos ue:

$$N(d) \subset N(f_i)\ ,\ i = 1,2,\ldots,r,$$

isto é,

(22) $$N(d)\subset N(f_1)\cap N(f_2)\cap \ldots \cap N(f_r).$$

Por outro lado, sendo d o m.d.c. de $f_1,f_2,\ldots,f_r$, existem polinômios (ver teorema 2 de A.4) $g_1,g_2,\ldots,g_r$ de $K[X]$, tais que:

$$d = \sum_{i=1}^{r} g_i\, f_i\ .$$

Daí, vem que:

(23) $$d(L) = \sum_{i=1}^{r} g_i(L)\, f_i(L)\ .$$

Para cada $v \in N(f_1)\cap \ldots \cap N(f_r)$,

$f_i(L)\, v = 0$, $i = 1,2,\ldots,r$ e de (23), obtemos:

$$d(L)v = \left[\sum_{i=1}^{r} g_i(L)\, f_i(L)\right](v) = \sum_{i=1}^{r} g_i(L)\left[f_i(L)\,(v)\right] = 0\ .$$

isto é,

(24) $$N(f_1)\cap \ldots \cap N(f_r) \subset N(d)$$

De (22) e (24) concluimos a demonstração do lema.

Sejam f_1, $f_2,\ldots,f_r$ polinômios relativamente primos de $K[X]$. O m.d.c. entre êles é $d = 1$ e daí $N(d) = \langle 0\rangle$, pois $d(L)=I$.

Do lema 4, segue-se o seguinte lema:

LEMA 5

"Sejam f_1, $f_2,\ldots,f_r$ polinômios relativamente primos de $K[X]$. Então, para todo operador linear L sôbre V, têm-se:

$$N(f_1)\cap N(f_2)\cap \ldots \cap N(f_r) = \langle 0\rangle\text{"}.$$

Pela definição de polinômio mínimo m de um operador linear L de V, sabemos que $N(m) = V$. Segue do lema 4, o próximo lema:

<u>LEMA 6</u>

"Sejam L um operador linear qualquer de V e m seu polinômio mínimo. Sejam f um polinômio qualquer de $K[X]$, e d o m.d.c. entre f e m. Então,
$$N(d) = N(f) \text{ ".}$$

Observamos que o lema 6, é uma consequência imediata do lema 4, visto que nêsse caso $N(f)$ é um subespaço de $V = N(m)$, isto é, $N(f) \cap N(m) = N(f) = N(d)$. O lema 6, nos permitirá restringir nossas considerações aos divisores do polinômio mínimo, já que o caso geral pode ser reduzido à êsse.

<u>TEOREMA 11</u>

"Seja L um operador linear de V e m seu polinômio mínimo. Sejam $f, g \in K[X]$, tais que:

i) f divide m

ii) g divide f e $gr(g) < gr(f)$, isto é, g é um divisor próprio de f.

Então, $N(g)$ é um subespaço próprio de $N(f)$, isto é,
$\dim N(g) < \dim N(f)$".

Demonstração.

Pelo lema 3, $N(g) \subset N(f)$. Resta mostrar que existe um vetor em $N(f)$ que não pertence à $N(g)$. Da hipótese i), existe um h em $K[X]$, tal que:

(25) $\qquad m = f.h$

Da hipótese ii) existe $r \in K[X]$, tal que; $gr(r) \geq 1$ e $f = r\,g$, que levada em (25),
$$m = r\,g\,h \quad , \quad gr\,(r) \geq 1 \,.$$

Logo $p = g\,h$ é um divisor próprio de m, isto é, $gr(p) < gr(m)$. Da definição de polinômio mínimo, temos que o operador linear $p(L) = g(L)\,h(L)$ não é o operador nulo de V. Assim sendo, existe pelo menos um vetor $v \in V$, tal que
$$p(L)\;(v) = g(L)\;[h(L)\;(v)] \neq 0 \Longrightarrow h(L)\;(v) \notin N(g)\;.$$

Entretanto, de (25), têm-se:

$$f(L)\ [h(L)\ (v)] = m(L)\ (v) = 0 \Longrightarrow h(L)\ (v) \in N(f)\ ,$$

o que demonstra o teorema.

Procuraremos remover agora a hipótese i) do teorema 11 . Seja d o m.d.c. entre f e m. Quando $d \equiv 1$, $N(d) = \langle 0 \rangle$. No entanto, se $gr(d) \geq 1$, o polinômio constante $c \equiv 1$ é um divisor de d e $gr(c) < gr(d)$. Pelo teorema 11, $\langle 0 \rangle = N(c)$ é um subespaço próprio de $N(d)$, isto é, $\dim N(d) > 0$. Por outro lado, do lema 6, têm-se $N(f) = N(d)$, o que demonstra o seguinte teorema:

<u>TEOREMA 12</u>

"Seja m o polinômio mínimo de um operador linear L sôbre V. Seja f um polinômio qualquer de $K[X]$ e d o m.d.c. entre f e m. Então, $\dim N(f) > 0$ se, e sòmente se, $gr(d) > 0$ ".

<u>TEOREMA 13</u>

"Seja m o polinômio mínimo, de um operador linear L sôbre V. Seja $f \in K[X]$ um divisor qualquer de m, tal que f é, unitário. Então, f é o polinômio mínimo do operador linear L_W, restrição de L ao subespaço $W = N(f)$".

Demonstração.

Pelo lema 2, $W_{,} = N(f)$ é L-invariante. Daí, pelo teorema 2 de 1.3, W é f(L)-invariante. Por (3) da proposição 1 de 1.3, têm-se que para todo $v \in N(f)$,

$$f(L_W)\ (v) = f(L)\ (v) = 0\ ,$$

isto é, L_W é um zero de f. Seja m' o polinômio mínimo de L_W. Então, do teorema 1 de 1.2, sendo $f(L_W) = 0$, vem que m' divide f. Ademais,

$$\forall v \in N(f) \Longrightarrow m'(L)\ (v) = m'(L_W)\ (v) = 0\ .$$

Logo N(m') não é um subespaço próprio de N(f), e pelo teorema 11, m' não pode ter grau menor que o de f. Sendo f e m', unitários, do mesmo grau e m' um divisor de f, m' = f, isto é, f é o polinômio mínimo de N(f).

c.q.d.

CAPÍTULO II - FORMAS TRIANGULAR E DIAGONAL

2.1 INTRODUÇÃO.

Para entendermos "o que é" uma forma canônica ou normal, faremos uso do conceito de relação de equivalência (ver B.1).

Diz-se que a matriz A é SEMELHANTE à matriz B, A e B em $M_n(K)$ se, e sòmente se, existe uma matriz inversível C de $M_n(K)$, tal que:

$$B = C^{-1} A C.$$

Essa é uma relação de equivalência no conjunto $M_n(K)$. de fato,

(i) Reflexiva $\quad A = E^{-1} A E$

(ii) Simétrica Seja $B = C^{-1} A C$, multiplicando-se à direita por C^{-1} e à esquerda por C, obtemos:

$$C B C^{-1} = A = (C^{-1})^{-1} B C^{-1}.$$

(iii) Transitiva Seja $B = C^{-1} A C$ e $D = F^{-1} B F$, logo

$$D = F^{-1} (C^{-1} A C) F = (C F)^{-1} A (C F).$$

Outra relação de equivalência no conjunto das matrizes quadradas de ordem n, com elementos no corpo K, é dada pela relação de matrizes "associadas". Diz-se que as matrizes A e B de $M_n(K)$, são ASSOCIADAS se, e sòmente se, existem matrizes inversíveis C e D em $M_n(K)$, tais que:

$$B = D^{-1} A C.$$

Para uma fixada relação de equivalência, entre matriz quadradas de ordem n, uma FORMA NORMAL ou CANÔNICA é uma matriz de tipo especial, escolhida em cada classe de equivalência. Em matemática, os têrmos "canônico", frequentemente significam "stander" em algum sentido particular. Uma forma normal, é uma forma stander para representar uma classe de elementos equivalentes. As seguintes condições, devem selecionar uma forma canônica, no conjunto das matrizes quadradas de ordem n(M_n (K)):

1. Dada qualquer matriz A de M_n (K), deve ser possível, por métodos

razoàvelmente direto, encontrar a forma normal da classe de equivalência contendo a matriz A.

2. O método deve conduzir-nos à uma única forma normal.

Veremos que para algumas das formas "especiais" que obteremos, a condição 2. ou 1. não será totalmente satisfeita. Por exemplo, se a forma especial é uma "matriz diagonal" (aqui 1. pode não ser possível, para algumas matrizes A), não nos preocuparemos com a ordem dos elementos diagonais.

Para estudarmos um operador linear L sôbre um K- espaço vetotial V, de dimensão finita, procuraremos uma base ordenada de V, relativamente à qual, a matriz que representa L seja a mais simples possível(canônica). Nossa meta é "desmontar" o operador linear L sôbre V, para podermos estudá-lo, analizando suas "partes". Para se "desmontar" L, procura-se decompor V em uma soma direta (ver B.2):

$$V = W_1 \oplus W_2 \oplus \dots \oplus W_r,$$

onde cada W_i é um subespaço L- invariante de V. Assim teremos decomposto L em uma soma direta:

$$L = L_{W_1} \oplus L_{W_2} \oplus \dots \oplus L_{W_r},$$

e sabendo o comportamento das "partes" $L_{W_1}, \dots\dots\dots, L_{W_r}$, saberemos o comportamento de L.

Nêsse capítulo II veremos as primeiras formas especiais, que são bastante símples e por isso mesmo muito úteis. Aliás, é nêsse capítulo que estudaremos a forma diagonal(que não é canônica, pois não satisfaz nem 1. nem 2.), a mais simples das formas especiais de um operador linear ou de uma matriz. Infelizmente, não é todo operador linear que pode ser colocado nessa forma especial, isto é, a condição 1. , que deve ser satisfeita na seleção de uma forma normal, não se verifica. O mesmo acontece para a forma triangular,

2.2. FORMA TRIANGULAR

Diz-se que uma matriz A de $M_n(K)$ esta na FORMA TRIANGULAR SUPERIOR (respectivamente, INFERIOR), se todos os seus elementos abaixo (respectivamente, acima) da diagonal, são nulos. Isto é, se A tem a forma:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} \left(\text{ou } A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \right)$$

EXEMPLO 1

As matrizes identidade I e nula 0, são triangulares superior (e inferior). Tôda matriz 1x1 é triangular.

Vamos nos preocupar apenas com a forma triangular superior, isto é, da qui em diante, forma triangular significa forma triangular superior.

Dizemos que um operador linear L sôbre V, é TRIANGULÁVEL se existe uma base ordenada $\mathcal{B}$ de V, tal que $\mathcal{L} = [L]_{\mathcal{B}}$ tem a forma triangular.

Suponhamos, que L seja triangulável e seja $\mathcal{B}$ a base ordenada de V, tal que:

$$\mathcal{L} = [L]_{\mathcal{B}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n-1} & a_{2n} \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & \cdots & a_{n-1\,n-1} & a_{n-1\,n} \\ 0 & 0 & \cdots & 0 & a_{n\,n} \end{bmatrix}$$

Então, o polinômio característico de L é

(1) $c(X) = (X-a_{11})(X-a_{22}) \ldots (X-a_{n-1\,n-1})(X-a_{nn})$.

Isto é, c é um produto de fatores lineares.

Para se verificar (1), basta desenvolver o determinante $c(X) = \det(XE - \mathcal{L}_\varphi)$ por Laplace sempre através da primeira coluna, (n-1)-vêzes. Acabamos de mostrar que:

(2) "sendo L triangulável, então seu polinômio característico é um produto de fatores lineares de $K[X]$"

A recíproca de (2) é válida. Para demonstrarmos tal resultado, usaremos noções de espaços vetoriais quocientes, que descrevemos no apêndice B.4.

TEOREMA 1.

"Suponhamos que o polinômio característico c do operador linear L sôbre V ($\dim V = n \geq 1$), se fatora em um produto de polinômios lineares em $K[X]$. Então, existe uma base ordenada $\mathcal{B}$ de V, relativamente à qual, L é representado por uma matriz na forma triangular."

Observação. Como $gr(c) = n \geq 1$, se K for um corpo algèbricamente fechado (por exemplo, se $K = \mathbb{C}$), então c se fatora no produto de polinômios lineares em $K[X]$, para todo L. Nêsse caso, todo operador linear é triangulável.

Demonstração.

Vamos usar indução sôbre a dimensão de V.

(i) Quando $\dim V = 1$, para qualquer base ordenada $\mathcal{B}$ de V, a matriz $[L]_{\mathcal{B}}$ é 1×1, logo esta na forma triangular.

(ii) Suponhamos que $\dim V = n > 1$ e que o teorema vale para espaços vetoriais de dimensão n-1. Por hipótese, o polinômio característico c de L, pode ser fatorado na forma (1), isto é, a_{11} é um zero de c. Logo a_{11} é um autovalor de L.

Seja e_1 um autovetor de L associado à a_{11}. Seja $W = \langle e_1 \rangle$, subespaço de V gerado por e_1. Como $e_1 \neq 0$ (pois é um autovetor) e $L e_1 = a_{11} e_1$, W é unidimensional e L-invariante. Consideremos o espaço vetorial quociente $\overline{V} = V / W$. Então (ver teorema 2, de B. 3) vem que:

$\dim \overline{V} = \dim V - \dim W = n - 1;$

e ainda (ver teorema 3, de B.3), L induz um operador linear $\overline{L}$ em $\overline{V}$, cujo polinômio mínimo divide o polinômio mínimo de L. Pelo teorema 10 de 1.4, sabemos que os zeros de c são também zeros de m. Desde que, c se fatora no produto de polinômios lineares, o mesmo ocorre com m. Isto também ocorre, com os polinômios característico e mínimo de $\overline{L}$. Então, $\overline{V}$ e $\overline{L}$ satisfazem nossa hipótese de indução, isto é, existe uma base ordenada $\overline{\mathcal{B}} = \{\bar{e}_2, \ldots, \bar{e}_n\}$ de $\overline{V}$ tal que:

$$(3)\quad \begin{cases} \overline{L}\ \bar{e}_2 = a_{22}\ \bar{e}_2 \\ \overline{L}\ \bar{e}_3 = a_{32}\ \bar{e}_2 + a_{33}\ \bar{e}_3 \\ \ldots\ldots \\ \ldots\ldots \\ \overline{L}\ \bar{e}_n = a_{n2}\ \bar{e}_2 + a_{n3}\ \bar{e}_3 + \ldots + a_{nn}\ \bar{e}_n . \end{cases}$$

Agora, consideremos os elementos $e_2, e_3, \ldots, e_n$ de V, que pertencem as classes de equivalência $\bar{e}_2, \bar{e}_3, \ldots, \bar{e}_n$, respectivamente. Então (ver teorema 2 de B.3), temos que

$\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ é uma base ordenada de V. Da i-ésima equação de (3), vem que

$\overline{L}\ \bar{e}_i - a_{i2}\ \bar{e}_2 - a_{i3}\ \bar{e}_3 - \ldots - a_{ii}\ \bar{e}_i = 0$, e daí,

$(L\ e_i - a_{i2}\ e_2 - a_{i3}\ e_3 - \ldots - a_{ii}\ e_i) \in W = \langle e_1 \rangle$, isto é,

$L\ e_i = a_{i1}\ e_1 + a_{i2}\ e_2 + \ldots + a_{ii}\ e_i$, $i = 1, \ldots, n.$

Então,

$$[L]_{\mathcal{B}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ 0 & a_{22} & a_{23} & \ldots & a_{2n} \\ . & . & . & \ldots & . \\ . & . & . & \ldots & . \\ 0 & 0 & 0 & \ldots & a_{nn} \end{bmatrix}$$

c.q.d.

A forma alternativa do teorema 1 para matrizes é:

<u>TEOREMA 1'</u>

"Suponhamos que o polinômio característico c de uma matriz A de M_n (K), se fatora no produto de polinômios lineares em K $[X]$. Então, A é semelhante a uma matriz triangular, isto é, existe uma matriz inversível C em M_n (K), tal que C^{-1} AC é triangular".

Seja $\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ uma base ordenada de V, relativamente a qual a matriz $[L]_{\mathcal{B}} = \mathcal{L}$ é triangular. Já verificamos em (1), que os elementos diagonais de $\mathcal{L}$ são necessàriamente os valores característicos de L.

Para cada $i = 1, 2, \ldots, n$ seja

(4) $V_i = \langle e_1, e_2, \ldots, e_i \rangle$,

isto é, V_i é o subespaço gerado pelos vetores $\{e_1, \ldots, e_i\}$. Então, dim $V_i = i$, pois $\{e_1, e_2, \ldots, e_n\}$ é uma base ordenada de V. Ademais, $V_i \subset V_{i+1}$, $i = 1, \ldots, n-1$, isto é, V_i é um subespaço de V_{i+1}. Logo os subspaços $\{V_1, V_2, \ldots, V_n\}$, dados por (4), formam uma cadeia ascendente de subspaços de V. Ainda mais, cada subespaço de (4) é L-invariante, pois

$$\forall \; x \in V_i \Longrightarrow x = \sum_{j=1}^{i} a_j \; e_j \Longrightarrow L\,x = \sum_{j=1}^{i} a_j \; L\,e_j \, .$$

Sendo $\mathcal{L}$ triangular, as cordenadas de L e_j na base ordenada $\mathcal{B}$, são tais que $a_{j+1} = a_{j+2} = \ldots = a_n = 0$, isto é,

$L\,e_j \in V_j \subset V_i$, $j = 1, 2, \ldots, i$, o que mostra que $L\,x \in V_i$.

Seja V um K-espaço vetorial de dimenção finita $n \geq 1$ e L um operador linear de V. Denomina-se <u>LEQUE</u> de L(em V), à uma cadeia de subespaços

$\{V_1, V_2, \ldots, V_n\}$ de V, tais que, para cada

$i = 1, 2, \ldots, n$,

(i) V_i é um subspaço de V_{i+1}

(ii) V_i é L- invariante

(iii) dim $V_i = i$

Por uma <u>BASE ASSOCIADA</u> a um leque

$\{V_1, V_2, \ldots, V_n\}$, entende-se uma base ordenada

$\mathcal{B} = \{e_1, \ldots, e_n\}$ de V, tal que, $V_i = \langle e_1, e_2, \ldots, e_i \rangle$.

Existe sempre, uma base associada a um leque $\{V_1, V_2, \ldots, V_n\}$, dado. De fato, seja $V_1 = \langle e_1 \rangle$, então extende-se a base ordenada $\{e_1\}$ de V_1, à uma base ordenada $\{e_1, e_2\}$ de V_2, a seguir, extende-se $\{e_1, e_2\}$ a uma base ordenada $\{e_1, e_2, e_3\}$ de V_3, e por indução, obtém-se uma base ordenada $\{e_1, e_2, \ldots, e_n\}$ de $V_n \equiv V$, a qual é, òbviamente associada ao leque dado.

Do teorema 1, e do descrito acima, obtemos o seguinte corolário:

<u>COROLÁRIO 1.</u>

"Suponhamos que o polinômio característico c do operador linear L sôbre V (dim V = n $\geqslant$ 1) se fatora em um produto de polinômios lineares em K $[X]$. Então, existe um leque de L em V ".

<u>DEMOSTRAÇÃO</u>.

Pelo teorema 1, existe uma base ordenada

$\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ de V, tal que $[L]_{\mathcal{B}}$ é triangular.

Basta agora, considerarmos

$$V_i = \langle e_1, e_2, \ldots, e_i \rangle, \; i = 1, 2, \ldots, n.$$

<u>EXEMPLO 2</u>

Sejam $A = [a_{ij}]$ e $B = [b_{ij}]$ matrizes triangulares de $M_n(K)$, isto é $a_{ij} = b_{ij} = 0$ se $i > j$. Então $A + B$ e $A.B$, também são triangulares. De fato,

$A + B = [a_{ij} + b_{ij}] = [s_{ij}]$, logo se $i > j$.

$s_{ij} = a_{ij} + b_{ij} = 0+0 = 0$. Seja $A.B = [c_{ij}]$,

onde $c_{ij} = \sum_{K=1}^{n} a_{ik} \; b_{kj}$, queremos mostrar que $c_{ij} = 0$,

quando $i > j$. Seja $i > j$, então:

(i) quando $K > j \Longrightarrow b_{kj} = 0 \Rightarrow c_{ij} = \sum_{K > j} a_{ik} \; b_{kj} = 0$

(ii) quando $j \geqslant k \Rightarrow i > j \geqslant k \Rightarrow i > k \Rightarrow a_{ik} = 0 \Longrightarrow$

$\Longrightarrow c_{ij} = \sum_{k \leq j} a_{ik} \; b_{kj} = 0$. isto é, que $A.B = [c_{ij}]$ é triangular

Em particular, podemos concluir que se:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix},$$

então

$$A^r = \begin{bmatrix} a_{11}^r & * & * & \cdots & * \\ 0 & a_{22}^r & * & \cdots & * \\ 0 & 0 & a_{33}^r & \cdots & * \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & a_{nn}^r \end{bmatrix}.$$

para um inteiro $r \geqslant 1$, onde $*$ indica que naquela posição da matriz o elemento de K pode não ser zero

Consequentemente, quando A é triangular com autovalores

$a_{11}, a_{22}, \ldots, a_{nn}$,

A^r também é triangular com autovalores $a_{11}^r, a_{22}^r, \ldots, a_{nn}^r$ para um inteiro $r \geq 1$.

EXEMPLO 3

A demonstração do teorema 1, nos fornece um método prático, para obtermos a representação triangular, de uma matriz dada. O presente exemplo, ilustra êsse fato.

Seja a matriz $\mathcal{L}$ sôbre o corpo $\mathbb{R}$ (dos números reais),

$$\mathcal{L} = \begin{bmatrix} 8 & 9 & -9 & 0 \\ 0 & 2 & 0 & 2 \\ 4 & 6 & -4 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix},$$

representado um operador linear L sôbre $\mathbb{R}^4$, relativamente à base canônica $\mathcal{B} = \{e_i = (\delta_{i1}, \delta_{i2}, \delta_{i3}, \delta_{i4}), \ i = 1, 2, 3, 4,\}$.

O polinômio característico de L é;

$$c(X) = (X-2)^3 (X-3);$$

logo vale a hipótese do teorema 1.

Assim os autovalores de L são 2 e 3.

Como:

$$\mathcal{L}[x] = \begin{bmatrix} 8x_1 + 9x_2 - 9x_3 \\ 2x_2 + 2x_4 \\ 4x_1 + 6x_2 - 4x_3 \\ 3x_4 \end{bmatrix} \quad e \quad \lambda[x] = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \lambda x_3 \\ \lambda x_4 \end{bmatrix}$$

λ é um autovalor de L (ou de $\mathcal{L}$) se, e sòmente se,

$\mathcal{L}[x] = \lambda[x]$ ($[x] \neq [0]$). Para $\lambda = 2$, obtemos

$\mathcal{L}[x] = 2[x]$ ou

$$(*)\quad \begin{cases} 8x_1 + 9x_2 - 9x_3 & = 2x_1 \\ \quad\; 2x_2 + \quad 2x_4 & = 2x_2 \\ 4x_1 + 6x_2 - 4x_3 & = 2x_3 \\ \quad\quad\quad 3x_4 & = 2x_4 \end{cases}$$

Logo um autovetor de L associado ao autovalor 2 (as coordenadas dêsse autovetor na base ordenada $\mathcal{B}$, é uma solução não trivial, do sitema (*)) é $f_1 = (0, 1, 1, 0)$. Formemos , a seguir, uma base ordenada $\mathcal{F}$, qualquer de $\mathbb{R}^4$, cujo primeiro elemento é f_1. Seja, por exemplo,

$$\mathcal{F} = \{(0, 1, 1, 0), (1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 0, 1)\}$$

A matriz que muda a base canônica $\mathcal{B}$ para a base ordenada $\mathcal{F}$, é

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Logo, $[L]_{\mathcal{F}} = A^{-1} [L]_{\mathcal{B}} A = A^{-1} \mathcal{L} A$. Depòis de feitos os cálculos, vem que:-

$$[L]_{\mathcal{F}} = A^{-1}\mathcal{L} A = \begin{bmatrix} 2 & 4 & 6 & 0 \\ 0 & 8 & 9 & 0 \\ 0 & -4 & -4 & 2 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Seja $W = \langle f_1 \rangle = \langle (0, 1, 1, 0) \rangle$ e $\overline{V} = \mathbb{R}^4 / W$.

L induz o operador linear $\overline{L}$ em $\overline{V}$. O vetor $\overline{g}_2 = 3\,\overline{f}_2 - 2\,\overline{f}_3$ é um autovetor associado ao autovalor 2 de $\overline{L}$. Escolhemos uma base para $\mathbb{R}^4$, cujos dois primeiros vetores são

f_2 e $g_2 = (3, -2, 0, 0) = 3(1, 0, 0, 0) - 2(0, 1, 0, 0)$, digamos:

$$\mathcal{H} = \{(0, 1, 1, 0), (3, -2, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\} .$$

A matriz B que muda a base $\mathcal{F}$ para a base $\mathcal{H}$ é:

$$B = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & -2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Logo, $[L]_{\mathcal{H}} = B^{-1} \; [L]_{\mathcal{F}} \; B$, e dada a simplicidade deste exemplo, a matriz $[L]_{\mathcal{H}}$ já é triangular.

$$[L]_{\mathcal{H}} = B^{-1} \; [L]_{\mathcal{F}} \; B = \begin{bmatrix} 2 & 0 & -6 & 2 \\ 0 & 2 & -3 & 0 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Sejam, $V_1 = \langle (0, 1, 1, 0) \rangle$, $V_2 = \langle (0, 1, 1, 0), (3, -2, 0, 0) \rangle$,

$V_3 = \langle (0, 1, 1, 0), (3, -2, 0, 0), (0, 0, 1, 0) \rangle$ e

$V_4 = \langle (0, 1, 1, 0), (3, -2, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1) \rangle$

Então, $\{V_1, V_2, V_3, V_4\}$, é um leque de L (em $\mathbb{R}^4$) e a base $\mathcal{H} = \{(0,1,1,0), (3,-2,0,0), (0,0,1,0), (0,0,0,1)\}$ de $\mathbb{R}^4$, é uma base associada ao leque.

2.3 O TEOREMA DA DECOMPOSIÇÃO PRIMÁRIA.

TEOREMA 2.

" Sejam $f_1, f_2, \ldots, f_r$ polinômios de $K[X]$ e h o m.m.c. entre êsses polinômios. Então, para qualquer operador linear L sôbre V, têm-se que:

$$N(h) = N(f_1) + N(f_2) + \ldots + N(f_r),$$

isto é, todo vetor $v \in N(h)$ (núcleo do operador linear $h(L)$) é da forma:

(5) $v = v_1 + v_2 + \ldots + v_r$, onde $v_i \in N(f_i)$, $i = 1, \ldots, r$".

Demonstração.

Seja $W = N(f_1) + N(f_2) + \ldots + N(f_r)$, o subspaço de V, consistindo dos vetores da forma (5).

Como f_i divide h, $i = 1,2,\ldots,r$, pelo lema 3 de 1.5, tem-se que $N(f_i) \subset N(h)$, $i = 1,2,\ldots,r$. Seja v um vetor arbitrário de W, isto é, v é da forma (5) e como cada $v_i \in N(f_i) \subset N(h)$, $v_i \in N(h)$, $i = 1,2,\ldots,r$.

Sendo $N(h)$ um subespaço de V, $v \in N(h)$. Provamos assim que $W \subset N(h)$.

Resta-nos mostrar que $N(h) \subset W$.

Como h é um múltiplo comum de $f_1, \ldots, f_r$, temos que:

(6) $\qquad h = f_i \, p_i$, $i = 1, 2, \ldots, r$,

onde p_i é um polinômio de $K[X]$.

Afirmamos que os polinômios $p_1, \ldots, p_r$ que aparecem em (6), são relativamente primos.

Caso contrário, êles teriam um divisor comum, digamos d com $gr(d) > 0$. Daí, $\frac{h}{d}, \frac{p_1}{d}, \ldots, \frac{p_r}{d}$ seriam polinômios de $K[X]$ e ter-se-ia:

$$\frac{h}{d} = \left(\frac{p_i}{d}\right) f_i \quad , \quad i = 1,2,\ldots,r,$$

o que é contrário à definição de m.m.c., pois $\frac{h}{d} \in K[X]$, $gr\left(\frac{h}{d}\right) < gr(h)$ e $\frac{h}{d}$ é um múltiplo comum de $f_1, \ldots, f_r$.

Sendo os polinômios $p_1, p_2, \ldots, p_r$, relativamente primos, o m.d.c. entre êles é o polinômio constante $s \equiv 1$. Logo, pelo teorema 2 de A.4, existem r polinômios $g_1, g_2, \ldots, g_r$, tais que:

$$1 = g_1 \, p_1 + \ldots + g_r \, p_r.$$

Daí o operador linerar $g_1(L)\,p_1(L)+\ldots+g_r(L)\,p_r(L)$ é o operador unidade I de V, isto é, para todo $v \in V$, têm-se:

(7) $I\,v = v = g_1(L)\left[p_2(L)(v)\right]+\ldots+g_r(L)\left[p_r(L)(v)\right]$.

Em particular, se $v \in N(h)$, teremos que $h(L)(v) = 0$, e de (6) vem que $f_i(L)\left[p_i(L)(v)\right] = 0$, $i = 1,2,\ldots,r$, ou seja $p_i(L)(v) \in N(f_i)$, $i = 1,2,\ldots,r$.

Pelo lema 2 de 1.5, cada subespaço $N(f_i)$ é invariante com respeito à L, logo pelo teorema 3 de 1.3, cada um dos subespaços $N(f_i)$ é $g(L)$- invariantes. Consequentemente, o vetor $g_i(L)\left[p_i(L)(v)\right] \in N(f_i)$, e por (7), têm-se que para todo $v \in N(h)$, v é da forma (5), logo está em W, isto é, $N(h) \subseteq W$, como queríamos demonstrar.

Nas hipóteses do teorema 2 acima, consideremos o caso no qual os polinômios $f_1,\ldots,f_r$ são primos dois à dois. Assim teremos que o m.m.c. entre êles será o polinômios

$$h = f_1 \cdot f_2 \cdots f_r$$

Afirmamos que nêsse caso, todo vetor de V tem uma <u>única</u> representação na forma (5). Primeiramente, mostraremos que se

(8) $v_1+v_2+\ldots+v_r = 0$, com $v_i \in N(f_i)$, então $v_i = 0$, $i = 1, 2,\ldots, r$. É claro que (8) é equivalente à

(9) $v_i = -v_1-\ldots-v_{i-1}-v_{i+1}-\ldots-v_r$, $i=1,\ldots,r$

(quando $i = 1$ (ou $i = r$), considerá-se $v_{i-1}=0$ (ou $v_{i+1} = 0$) . Seja,

$p_i = \frac{h}{f_i} = f_1\cdots f_{i-1}\cdot f_{i+1}\cdots f_r$, $i= 1,2,\ldots,r$, isto é p é o m.m.c. entre os polinômios

$f_1,\ldots,f_{i-1}, f_{i+1},\ldots, f_r$. O m.d.c. entre p_i e f_i é o polinômio constante $\Delta \equiv 1$, pois os polinômios $f_1,\ldots,f_r$ são primos dois à dois. Do lema 4 de 1.5, vem:

$$N(p_i) \cap N(f_i) = N(\Delta) = \langle 0 \rangle$$

Como $v_i \in N(f_i)$, pelo teorema 2, o vetor do lado direito de (9), pertence à $N(p_i)$, isto é, $v_i = 0$.

Suponhamos agora, que um vetor $v \in V$ além da representação (5), possa ser expresso na forma:

$v = v_1' + v_2' + \ldots + v_r'$ onde $v_i' \in N(f_i)$, $i = 1, \ldots, r$.

Subtraindo-se membro à membro essa equação de (5), obtém-se:

$(v_1 - v_1') + (v_2 - v_2') + \ldots + (v_r - v_r') = 0$,

onde $(v_i - v_i') \in N(f_i)$. Essa equação está na forma (8) , logo $v_i - v_i' = 0$ ou $v_i = v_i'$, $i = 1, \ldots, r$, o que demonstra a afirmativa que fizemos.

Provamos assim o seguinte teorema:

<u>TEOREMA 3.</u>

" Sejam $f_1, f_2, \ldots, f_r$,polinômios dois à dois primos em $K[X]$ e h o m.m.c. entre êsses polinômios.

Então, para todo operador linear L sôbre V, têm-se que:

$N(h) = N(f_1) \oplus N(f_2) \oplus \ldots \oplus N(f_r)$,

isto é, $N(h)$ é a soma direta dos subespaços $N(f_1), \ldots, N(f_r)$."

Observamos que ao encontrarmos um polinômio (qualquer) f de $K[X]$, tal que $N(f) = V$, o teorema 3 nos dá uma "<u>decomposição de V em uma soma direta de subespaços L-invariantes</u> ". Em particular, sabemos que o polinômio mínimo m de L é tal que, $m(L) = 0$, logo $N(m) = V$. Fatoremos o polinômio m em $K[X]$, agrupando os fatores irredutíveis iguais, isto é,

$$(10) \qquad m = q_1^{a_1} \cdot q_2^{a_2} \cdots q_r^{a_r}$$

onde a_i é um inteiro extritamente positivo e $q_i \neq q_j$ se $i \neq j$. Sendo m unitário, podemos ter cada q_i , $i = 1, 2, \ldots, r$, também unitário (ver teorema 6 de A.5). Sejam $V_1, V_2, \ldots, V_r$ os núcleos dos operadores lineares,

$$\left[g_1(L)\right]^{a_1},\quad \left[g_2(L)\right]^{a_2},\ \ldots\ldots,\ \left[g_r(L)\right]^{a_r},$$

respectivamente. Então, pelo teorema 3, $V = N(m)$ é a soma direta de seus subespaços $V_1, V_2, \ldots, V_r$.

Suponhamos que $\dim V_i = n_i$, $i = 1,2, \ldots, r$ e seja $e^{i1}, e^{i2}, \ldots, e^{in_i}$ uma base para cada subespaço V_i.
Então (ver teorema 1 de B. 2), o conjunto de todos os vetores básicos para cada V_i, isto é,

(11) $\mathcal{B} = \left\{e^{11}, \ldots, e^{1n_1}, e^{21}, \ldots, e^{2n_2}, \ldots, e^{r1}, \ldots, e^{rn_r}\right\}$,

é uma base ordenada de V e

$$\dim V = n = n_1 + n_2 + \ldots + n_r .$$

Obteremos a seguir, a matriz $[L]_{\mathcal{B}} = \mathcal{L}$.
Lembramos que as colunas de $\mathcal{L}$ são os componentes dos vetores $L\,e^{ik}$, na base $\mathcal{B}$.
Fixemos um índice $i \in \{1,2,\ldots, r\}$. Cada vetor $L\,e^{i1},\ldots, L\,e^{in_i}$ está em V_i, pois V_i é invariante relativamente à L, e sendo $e^{i1},\ldots, e^{in_i}$ base de V_i, cada vetor $L\,e^{i1},\ldots, L\,e^{in_i}$ é uma combinação linear de $e^{i1}, e^{i2},\ldots, e^{in_i}$, digamos:

(12) $L\,e^{ik} = a_{1k}^{(i)}\, e^{i1} + a_{2k}^{(i)}\, e^{i2} + \ldots + a_{n_i k}^{(i)}\, e^{in_i}$, $K=1,\ldots,n_i$.

De (12), vê-se que as coordenadas de $L\,e^{ik}$, com respeito à base ordenada (11), são tais que; as primeiras $n_1 + n_2 + \ldots + n_{i-1}$ (quando $i = 1$, coloca-se $n_0 = 0$) coordenadas de $L\,e^{ik}$ são nulas, as seguintes são $a_{1k}^{(i)}, a_{2k}^{(i)},\ldots, a_{n_i k}^{(i)}$, e as restantes $n_{i+1} + n_{i+2} + \ldots + n_r$ (quando $i = r$, coloca-se $n_{r+1} = 0$) também são nulas.
Então, as coordenadas dos vetores $L\,e^{i1}, L\,e^{i2},\ldots, L\,e^{in_i}$, entram na formação da matriz $\mathcal{L}$, com um "bloco" quadrado da forma :-

$$(13)\quad \begin{vmatrix} a_{11}^{(i)} & a_{12}^{(i)} & \cdots & a_{1n_i}^{(i)} \\ a_{21}^{(i)} & a_{22}^{(i)} & \cdots & a_{2n_i}^{(i)} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a_{n_i1}^{(i)} & a_{n_i2}^{(i)} & \cdots & a_{n_in_i}^{(i)} \end{vmatrix}$$

o qual denotamos por $\mathcal{L}^{(i)}$, e que situa-se da $(n_1+n_2+\ldots+n_{i-1}+1)$-ésima linha e coluna, até a $(n_1+n_2+\ldots+n_i)$-ésima linha e coluna de $\mathcal{L}$. Logo, a diagonal do bloco (13) faz parte da diagonal de $\mathcal{L}$. Então, fazendo-se i variar de 1 à r, obtemos r dêsses blocos, $\mathcal{L}^{(1)}, \mathcal{L}^{(2)}, \ldots, \mathcal{L}^{(r)}$ da forma (13), os quais aparecem ao longo da diagonal de $\mathcal{L}$, sendo que fora dêsses blocos, aparecem apenas zeros, portanto têm-se:

$$(14)\ \mathcal{L} = \begin{bmatrix} \mathcal{L}^{(1)} & 0 & \cdots & 0 \\ 0 & \mathcal{L}^{(2)} & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \mathcal{L}^{(r)} \end{bmatrix}$$

Em (14), $\mathcal{L}$ tem a forma de blocos, desde que se interprete cada $\mathcal{L}^{(i)}$ como uma matriz $n_i \times n_i$ e os 0 representam matrizes retangulares nulas. Logo $\mathcal{L}$ é a "soma direta" (ver B.2) das matrizes $\mathcal{L}^{(1)}, \mathcal{L}^{(2)}, \ldots, \mathcal{L}^{(r)}$ e denota-se por:

$$\mathcal{L} = \mathcal{L}^{(1)} \oplus \mathcal{L}^{(2)} \oplus \ldots \oplus \mathcal{L}^{(r)}.$$

Consideremos agora o operador linear $L^{(i)}$, restrição de L ao subespaço V_i. A matriz de $L^{(i)}$, relativamente à base $e^{i1}, e^{i2}, \ldots, e^{in_i}$ de V_i, tem por colunas as coordenadas dos vetores $L e^{i1}, L e^{i2}, \ldots, L e^{in_i}$, relativamente à base $e^{i1}, e^{i2}, \ldots, e^{in_i}$. De (12) a matriz de $L^{(i)}$ procurada, é o bloco (13), considerando como uma matriz, ou seja, é a matriz $\mathcal{L}^{(i)}$, $i = 1,2, \ldots, r$.

Pelo teorema 13 de 1.5, se o polinômio mínimo de L, esta na forma (10), então o polinômio mínimo de $\mathcal{L}^{(i)}$ é

$$g_i^{a_i}, \quad i = 1,2, \ldots, r.$$

Resumiremos agora, os resultado obtidos em um teorema, para facilitar futuras referências:-

TEOREMA 4 (TEOREMA DA DECOMPOSIÇÃO PRIMÁRIA)

" Seja L um operador linear sôbre um K- espaço vetorial V, de dimensão finita $n \geq 1$. Seja m o polinômio mínimo de L, tal que:

$$m = g_1^{a_1} \cdot g_2^{a_2} \cdots g_r^{a_r},$$

onde os g_i são polinômios de $K[X]$, irredutíveis, unitários, e dois à dois distintos, e os a_i são números inteiros extritamente positivos. Seja V_i o núcleo do operador linear:

$$[g_i(L)]^{a_i}, \quad i = 1,2,\ldots,r.$$

Então;

1. $V = V_1 \oplus V_2 \oplus \cdots \oplus V_r$

2. Cada V_i é L- invariante.

3. Escolhida uma base $e^{i1}, \ldots, e^{in_i}$ de V_i, para $i = 1,2, \ldots, r$, os vetores:

$$\mathcal{B} = \{e^{11}, \ldots, e^{1n_1}, e^{21}, \ldots, e^{2n_2}, \ldots, e^{r1}, \ldots, e^{rn_r}\},$$

formam uma base de V.

4. Sendo $L^{(i)}$ o operador linear induzido por L em V_i, $i=1,2,\ldots,r$, o polinômio mínimo de $L^{(i)}$ é $g_i^{a_i}$

EXEMPLO 4.

Nêste exemplo, procuraremos aclarar os conceitos e resultados obtidos. Seja $V = \mathbb{R}^3$, espaço vetorial dos têrmos ordenados de números reais. Seja,

$\mathcal{C} = \{e_i = (\delta_{i1}, \delta_{i2}, \delta_{i3}), \ i = 1,2,3\}$ a base canônica de V. Seja L o operador linear sôbre V, cuja matriz $\mathcal{L}$ relativamente à base canônica é:

$$[L]_{\mathcal{C}} = \mathcal{L} = \begin{bmatrix} 0 & 2 & -1 \\ -2 & 5 & -2 \\ -4 & 8 & -3 \end{bmatrix}$$

Um cálculo direto nos mostra que $\mathcal{L}^2 = \mathcal{L}$, isto é, $\mathcal{L}$ é um zero do polinômio $p(X) = X^2 - X$ de $\mathbb{R}[X]$. Então, o polinômio mínimo m de $\mathcal{L}$, é um divisor de p em $\mathbb{R}[X]$. Logo $m(X)$ é $X-1$, X ou $X^2 - X$. Sendo $X-1$ e X polinômios mínimos das matrizes unidade e nula, respectivamente, segue-se que
$m(X) = X^2 - X = (X-1).X$. Assim sendo, m está na forma (10), com $g_1(X) = X-1$, $g_2(X) = X$ e $a_1 = a_2 = 1$.

Procuraremos agora os subespaços $V_1 = N(g_1)$ e $V_2 = N(g_2)$; núcleos dos operadores lineares.
$g_1(L) = L-I$ e $g_2(L) = L$, respectivamente. O subespaço V_1, que consiste de todos os vetores $v \in V$, tais que $(L-I)(v) = 0$, isto é, $Lv = v$, tem dimensão 2. De fato, a caracteristica da matriz $E - \mathcal{L}$ é 1, então $\dim V_1 = 2$. Anàlogamente, conclui-se que $\dim V_2 = 1$. Assim sendo, $n_1 = 2$ e $n_2 = 1$. Consideremos $\{e^{11}, e^{12}\}$ uma base ordenada de V_1 e $\{e^{21}\}$ uma base de V_2. Pelo teorema 4, item 3., $\mathcal{B}\{e^{11}, e^{12}, e^{21}\}$ é uma base ordenada de V.
Para escrevermos a matriz $\mathcal{L}^*$ de L, relativamente à base ordenada $\mathcal{B}$, observamos que o operador linear $L^{(1)}$, restrição de L à V_1, é o operador unidade de V_1, logo sua matriz $\mathcal{L}^{(1)}$, relativamente

à qualquer base ordenada (e em particular, à base $\{e^{11}, e^{12}\}$ de V_1, é a matriz identidade de ordem 2, isto é

$$\mathcal{L}^{(1)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Anàlogamente, a matriz $\mathcal{L}^{(2)}$, do operador, linear $L^{(2)}$ (nulo), restrição de L à V_2, é a matriz nula de ordem 1. Então, a matriz $\mathcal{L}^{*}$ é:

$$[L]_{\mathcal{B}} = \mathcal{L}^{*} = \mathcal{L}^{(1)} \oplus \mathcal{L}^{(2)} = \left[\begin{array}{c|c} \mathcal{L}^{(1)} & 0 \\ \hline 0 & \mathcal{L}^{(2)} \end{array}\right] = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & 0 \end{array}\right].$$

Agora, podemos descrever o operador linear L, de maneira mais simples. Seja $v = a_1 e^{11} + a_2 e^{12} + a_3 e^{21}$, então $L\, v = (a_1, a_2, 0)$, pois:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ 0 \end{bmatrix},$$

donde cocluimos que L é uma projeção do espaço $\mathbb{R}^3$ sôbre o plano gerado pelos vetores $\{e^{11}, e^{12}\}$, segundo a direção da reta que contém o terceiro vetor básico e^{21}

2.4. A FORMA DIAGONAL

Nessa seção, estudaremos sob quais condições, existe uma base ordenada de V, em relação à qual a matriz de um operador linear L é "diagonal ". Diz-se que uma matriz $A = [a_{ij}]$ de M_n (K), está na <u>FORMA DIAGONAL</u>, se para cada $i, j = 1, 2, \ldots, n$,

$$a_{ij} = a_i \, \delta_{ij} \quad , \text{ onde } \quad \delta_{ij} = \begin{cases} 1 \text{ se } i = j \\ 0 \text{ se } i \neq j. \end{cases}$$

Nêsse caso, denotamos a matriz A por:

$[a_1, a_2, \ldots, a_n]$, isto é,

$$A = \begin{bmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & a_n \end{bmatrix} = [a_1, a_2, \dots, a_n].$$

Diz-se que um operador linear L sôbre V, é <u>DIAGONALIZÁVEL</u> se, e sòmente se, existe uma base de V, formada por autovetores de L.

O nome, operador diagonalizável, vem do fato de que se $\mathcal{B} = \{e_1, e_2, \dots, e_n\}$ é uma base ordenada de V, formada por autovetores de L, digamos que $L e_i = \lambda_i e_i$, $i = 1, 2, \dots, n$, então

(15) $\quad [L]_{\mathcal{B}} = [\lambda_1, \lambda_2, \dots, \lambda_n]$,

isto é, L é representado por uma matriz diagonal, relativamente à base de autovetores. Observamos que o corolário 1 de 1.4, nos informa que uma condição <u>suficiente</u> para que um operador linear L seja diagonalizável, é que L possua n-autovalores distintos dois à dois. Nêsse caso, além de sabermos que L é diagonalizável, (15) nos diz que a representação diagonal de L é dada pelos seus autovalores.

A condição acima, é suficiente, mas <u>não é necessária</u> para que L seja diagonalizável. Por exemplo se $L = \lambda I$, então o único autovalor de L é λ e $[L]_{\mathcal{B}} = [\lambda, \lambda, \dots, \lambda]$.

Passamos agora ao teorema central dessa secção, que nos dá condições necessárias e suficientes, para que um operador linear L seja diagonalizável.

<u>TEOREMA 5</u> (<u>TEOREMA DE DIAGONALIZAÇÃO</u>)

" Sejam V um espaço vetorial sôbre um corpo F, algèbricamente fechado, de dimensão finita n e L um operador linear sôbre V. Então, as seguintes condições são equivalentes:

1. L é diagonalizável.

2. O polinômio mínimo m de L, pode ser fatorado na forma.

$$m(X) = (X - \lambda_1)(X - \lambda_2) \dots (X - \lambda_r),$$

onde para cada $i = 1,2,\dots, r$, $\lambda_i \in F$ e $\lambda_i \neq \lambda_j$ se $i \neq j$.

3. Para cada $i = 1,2,\dots,r$, λ_i é um autovalor de L, e a multiplicidade algébrica de λ_i é igual a multiplicidade geométrica de λ_i.

4. Para cada λ_i, $i = 1,2,\dots, r$, o núcleo do operador linear $(L - \lambda_i I)^2$ esta contido no núcleo do operador linear $L - \lambda_i I$.

5. Seja v um autovetor de L, associado ao autovalor λ. Então, não existe nenhum vetor w em V, tal que $(L - \lambda_i I)w = v$.

6. Existe um inteiro r extritamente positivo, r-escalares distintos dois à dois, $\lambda_1, \lambda_2, \dots, \lambda_r$ e operadores lineares inversíveis $E_1, E_2, \dots, E_r$, tais que:

(i) $L = \sum_{j=1}^{r} \lambda_j E_j$.

(ii) $I = \sum_{j=1}^{r} E_j$.

(iii) $E_i E_j = 0$ se $i \neq j$."

Demonstração.

1. $\Longrightarrow$ 2.

Como L é diagonalizável, existe uma base ordenada $\mathcal{B} = \{e_1, e_2, \dots, e_n\}$ de V, formada por vetores caracteristicos de L. Isto é,

(16) $L e_i = \lambda_i e_i$, $i = 1,2,\dots,n$.

Sabemos que cada λ_i é uma raiz do polinômio característico de L, pois λ_i é um autovalor de L. Pelo teorema 10 de 1.4, cada λ_i é zero do polinômio mínimo m de L. Seja $g_i(X) = X - \lambda_i$ $i=1,2,\dots$

Então, f_i divide m, pois λ_i é um zero de m, $i=1,2,\ldots, n$.

Suponhamos que em (16), existem r escalares distintos λ_i. Sem perda de generalidade, podemos supo-los $\lambda_1, \lambda_2, \ldots, \lambda_r$ (reordena-se a base $\mathcal{B}$, se necessário). Seja,

$$g(X) = f_1(X) \cdot f_2(X) \ldots f_r(X) = (X-\lambda_1)(X-\lambda_2)\ldots(X-\lambda_r).$$

Como cada f_i divide m, temos que g também divide m, pois $r \leq n$. Como,

$$f_i(L)(e_i) = (L-\lambda_i I)e_i = 0,$$

g (L) leva e_i, $i = 1,2,\ldots, n$, no vetor nulo. Na verdade usamos o fato de que se e_k é um autovetor associado à λ_i, também $f_i(L)\ e_k = 0$, o que é óbvio. Assim sendo, $g(L) = 0$, logo o polinômio mínimo m de L, divide g. Desde que m divide g e g divide m, e ainda m e g sendo unitários, segue-se que m= g. O que mostra a validade de 2.

2. $\Rightarrow$ 3.

Como,

$$m(X) = (X-\lambda_1)(X-\lambda_2)\ldots(X-\lambda_r),\ \lambda_i \neq \lambda_j \text{ se } i \neq j,$$ cada λ_i é um autovetor de L. Pelo teorema da Decomposição Primária (teorema 4 de 2.3.), temos:

$$V = V_1 \oplus V_2 \oplus \ldots \oplus V_r,$$

onde se $f_i(X) = X - \lambda_i$, $i = 1,2,\ldots, r$, temos

$$V_i = \mathrm{Ker}\,(f_i(L)) = \{v \in V : Lv = \lambda_i v\},$$

isto é, V_i é o autoespaço associado à λ_i. Ainda, escolhida uma base $\{e^{i1}, e^{i2}, \ldots, e^{in_i}\}$ para V_i, a condição (3) do teorema da Decomposição Primária, nos diz que:

$$\{e^{11}, \ldots, e^{1n_1}, e^{21}, \ldots, e^{2n_2}, \ldots, e^{r1}, \ldots, e^{rn_r}\},$$

é uma base ordenada de V. Logo,

(17) $\quad n_1 + n_2 + \ldots + n_r = n,$

onde cada n_i é a multiplicidade geométrica do λ_i (n_i =dim V_i). Sendo a_i a multiplicidade algébrica de λ_i, temos que:

$$(18)\quad\begin{cases}(a)\ a_1 + a_2 + \dots + a_r = n \text{ (pois o grau do polinômio carac-} \\ \text{terístico de } L \text{ é } n \text{)} \\ (b)\ \ 1 \leq n_i \leq a_i\ ,\ i = 1,2,\dots,r \text{ (teorema 6 de 1.4).}\end{cases}$$

De (17) e (18) vem que, $n_i = a_i$, $i = 1,2,\dots,r$.

3. $\Longrightarrow$ 4.

Seja λ_i um zero de m, logo λ_i é um autovalor de L (teorema 10 de 1.4). De 3. sabemos que a multiplicidade geométrica n_λ de λ_i, coincide com a multiplicidade algébrica a_i de λ_i, $i = 1,2,\dots,r$. Logo, $n_1 + n_2 + \dots + n_r = n$, e pelo corolário 2 de 1.4, V tem uma base ordenada formada de autovetores de L, isto é, L é diagonalizável. Assim sendo, vale 1. e como já provamos que 1. $\Longrightarrow$ 2., vale 2. e $g(X) = (X - \lambda_i)^2$ não divide m, isto é, $f_i\ (X) = X - \lambda_i$ é o m.d.c. entre m e g. Do lema 6 de 1.5, concluímos que $N(f) = N\ (g)$. Demonstramos mais do que o exigido, mas na verdade, o fato de que o núcleo de $L - \lambda I$ esta contido no núcleo de $(L - \lambda I)^2$ é sempre válido.

4. $\Longrightarrow$ 5.

Temos que $L\,v = \lambda v$. Caso exista um w em V tal que $(L - \lambda I)\ w = v$, têm-se:

$$(L - \lambda I)^2\ w = (L - \lambda I)\left[(L - \lambda I)\,w\right] = (L - \lambda I)\ v = 0,$$

e por 4., temos que $(L - \lambda I)\ w = 0$, logo $v = 0$, o que é absurdo, pois v é um autovetor de L.

5. $\Rightarrow$ 6.

Suponhamos que m tenha um zero λ, de multiplicidade 2, isto é,

$$m\ (X) = f\ (X)\ (X - \lambda)^2\ ,\qquad f\ (\lambda) \neq 0.$$

Da definição de polinômio mínimo, temos que:

$$g\ (L) = f\ (L)\ (L - \lambda I\ ) \neq 0,\ \text{pois}\ \ gr\ (g) < gr\ (m).$$

Assim sendo, existe um vetor não nulo z em V, tal que: $g\ (L)\ z \neq 0$. Seja $v = g\ (L)\ z$, logo,

$$(L - \lambda I\ )\ v = f(L)\ (L - \lambda I)^2\ (z) = m\ (L)\ (z) = 0,$$

pois $m\ (L) = 0$. Concluímos assim que [illegible] λ [illegible] $f(L)(L$ [illegible] tal que $v = g(L)(z) = (L-\lambda I)w$ contrariando 5. . Consequentemente, m não tem raiz de multiplicidade maior que 1, isto é,

$$m(X) = (X-\lambda_1)\dots(X-\lambda_r),\ \lambda_i \neq \lambda_j \text{ se } i \neq j .$$

Definimos os polinômios p_i de $F[X]$, pondo-se:

$$m(X) = (X-\lambda_i)\,p_i(X)\ ,\ i = 1,2,\dots,r .$$

Então, $p_i(\lambda_j) = 0$ se, e sòmente se, $i \neq j$. Consideremos o polinômio p definido por:

$$(19)\qquad p(X) = 1 - \sum_{i=1}^{r} \left[p_i(\lambda_i)\right]^{-1} p_i(X)\ ,$$

observa-se que $p_i(\lambda_i) \neq 0$, $gr(p) < r$ e

$$p(\lambda_j) = 1 - \sum_{i=1}^{r} \left[p_i(\lambda_i)\right]^{-1} p_i(\lambda_j) = 1 - \left[p_j(\lambda_j)\right]^{-1}\left[p_j(\lambda_j)\right] = 0,$$

para $j = 1,2,\dots,r$, isto é, p tem por raízes $\lambda_1, \lambda_2,\dots, \lambda_r$, logo p é um múltiplo de $m(X) = (X-\lambda_1)(X-\lambda_2)\dots(X-\lambda_r)$, daí $p(L) = 0$. Como $gr(m) = r > gr(p)$, da definição de polinômio mínimo, concluimos que p é o polinômio identicamente nulo. Logo, para todo operador linear T sôbre V, $p(T) = 0$. Em particular, de (19) vem que:

$$(20)\qquad \sum_{i=1}^{r} \left[p_i(\lambda_i)\right]^{-1} p_i(X) = 1$$

Definimos, para cada $i = 1,2,\dots,r$,

$$(21)\qquad E_i = \left[p_i(\lambda_i)\right]^{-1} p_i(L) .$$

Como $gr(p_i) = (r-1) < gr(m)$, $E_i \neq 0$. De (20) e (21) vem:

$$(ii)\qquad \sum_{i=1}^{r} E_i = I .$$

Ainda, $m(L) = (L-\lambda_i I)\,p_i(L) = 0$, logo

$$\lambda_i\, p_i(L) = L\, p_i(L),$$

e de (ii), temos que:

$$(i)\quad \sum_{i=1}^{r} \lambda_i E_i = \sum_{i=1}^{r} \left[p_i(\lambda_i)\right]^{-1} L p_i(L) = L \sum_{i=1}^{r} E_i = L\, I = L.$$

Seja agora $i \neq j$, então,

$$E_i\, E_j = [p_i(\lambda_i)\cdot p_j(\lambda_j)]^{-1}\, p_i(L)\, p_j(L) =$$

$$= [p_i(\lambda_i)\, p_j(\lambda_j)]^{-1}\, h(L)\, m(L)\ ,$$

onde $h(L)$ é o produto dos $r-2$ operadores lineares $L-\lambda_k I$, onde $k\neq i$ e $k\neq j$. Logo:

(iii) $\qquad E_i\, E_j = 0 \quad$ se $\quad i \neq j$.

<u>6.</u> $\Longrightarrow$ <u>1.</u>

Mostraremos que para um vetor v qualquer de V, $E_i\ v\ (i=1,2,\ldots r)$ é um autovetor de L associado à λ_i ou é nulo.

$$L(E_i v) \overset{(i)}{=} (\sum_{j=1}^{r} \lambda_j\, E_j)(E_i v) = (\sum_{j=1}^{r} \lambda_j\, E_j\, E_i)\, v =$$

$$\overset{(iii)}{=} \lambda_i\, E_i^2\, v \overset{(*)}{=} \lambda_i\,(E_i\, v)$$

Na igualdade (*), usamos o fato $E_i^2 = E_i$ (isto é, cada E_i é idempotente ou uma projeção), vamos mostrar êsse fato. De (ii), $I = \sum_{j=1}^{r} E_j$, aplicando-se E_i a essa igualdade e usando (iii), vem que,

$$E_i = \sum_{j=1}^{r} E_j\, E_i = E_i^2 \ .$$

Para mostrarmos <u>1.</u> , basta mostrarmos que os autovetores de L, geram V. Seja $v\in V$, qualquer. Então, $v=Iv \overset{(ii)}{=} (\sum_{j=1}^{r} E_j)v =$ $= \sum_{j=1}^{r} (E_j\ v)$, e já vimos que $E_j\ v$ é autovetor de L associado à λ_i ou é o vetor nulo.

c.q.d.

Destas condições ou critérios de diagonalização, destacamos os <u>2.</u> e <u>3.</u> , para maior facilidade de cálculos. Quanto à condição <u>6.</u> , queremos destacar os seguintes fatos. Um operador linear L é diagonalizável se, e sòmente se, L é uma combinação linear ($L=\sum_{j=1}^{r}\lambda_j\, E_j$) de um conjunto de <u>PROJEÇÕES</u> ($E_j^2 = E_j$) <u>ORTOGONAIS</u> ($E_i\, E_j = 0$ se $i\neq j$) e <u>SUPLEMENTARES</u> ($I = \sum_{j=1}^{r} E_j$).

Além disso, os coeficientes λ_j desta combinação linear, são exatamen te os autovalores distintos dois à dois de L, como se pode ver na demonstração da última etapa do teorema 5.

Ademais, E_i é a projeção sôbre o autoespaço V_{λ_i}, associado ao autovalor λ_i. A decomposição,

$$L = \sum_{j=1}^{r} \lambda_j \, E_j ,$$

é muitas vêzes denominada <u>FORMA ESPECTRAL</u> de L.

Seja A uma matriz de M_n (K). Diz-se que A é <u>DIAGONALIZÁVEL</u> em K se, e sòmente se, A é semelhante a uma matriz diagonal. Isto é, A é diagonalizável em K se, e sòmente se, existe uma matriz inversível C em M_n (K), tal que C^{-1} A C é uma matriz diagonal.

Seja L o operador linear sôbre K^n, cuja matriz relativamente à base canônica é A. É claro que dizer que A é diagonalizável, é equivalente a se dizer que L é diagonalizável. Sabemos ainda do teorema 2 de 1.2, que o polinômio mínimo de L e de A, coincidem. "Mutatis-mutatis", obtém-se uma versão para matrizes do teorema de diagonalização (teorema 5, acima).

<u>EXEMPLO</u> <u>5</u>

Uma mesma matriz (ou operador linear) pode ser diagonalizável em um corpo e não o ser em um outro. Consideremos, por exemplo, a matriz:

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

cujo polinômio mínimo é $m(X) = X^2 + 1$, que é irredutível em $\mathbb{R}[X]$ e que em $\mathbb{C}[X]$ admite a fatorização $m(X) = (X+i)(X-i)$. Logo, A não é diagonalizável em $\mathbb{R}$, mas o é em $\mathbb{C}$.

Observamos, que basta sabermos que uma matriz A, de M_n (K), é semelhante a uma matriz diagonal D, a fim de podermos escrever uma tal D, simplesmente colocando os autovalores de A (sem nos preocupar com a ordem) como elementos diagonais, sendo que cada autovalor figura em D, tantas vêzes quanto for sua multiplicidade algébrica ou geométrica. Por outro lado, pode-se querer encontrar uma <u>MATRIZ DIAGONALIZADORA</u>, para A, isto é, uma matriz inversível C, de $M_n(K)$, tal que $D = C^{-1} AC$. Para tanto, seja T o operador linear sôbre K^n, cuja matriz relativamente à base canônica $\mathcal{B} = \{e_i = (\delta_{i1}, \ldots, \delta_{in}) ; i = 1,2,\ldots, n\}$, é A. Sendo A diagonalizável T também o é. Então existe uma base ordenada $\mathcal{F}$ de K^n, formada por autovetores de T e $[T]_{\mathcal{F}} = D$ (D é diagonal). Seja

$C = [c_{ij}]$ a matriz mudança da base ordenada $\mathcal{B}$ para a base $\mathcal{F}$.

Então, as colunas de C são as coordenadas dos autovetores da base ordenada $\mathcal{F}$, relativamente à base canônica $\mathcal{B}$, respectivamente. E temos:

$$D = [T]_{\mathcal{F}} = C^{-1} [T]_{\mathcal{B}} C = C^{-1} AC,$$

isto é, C é uma matriz diagonalizadora para A.

<u>EXEMPLO 6</u>

Seja
$$A = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 0 & 0 \\ -2 & 0 & 4 \end{bmatrix}$$

Verificamos se A é diagonalizável, e se for, encontraremos uma matriz diagonal semelhante à A, bem como uma matriz diagonalizadora C, para A.

O polinômio característico de A, é :

$$c(X) = \det(XE-A) = \begin{vmatrix} X-1 & 0 & 2 \\ 0 & X & 0 \\ 2 & 0 & X-4 \end{vmatrix} = X(X-1)(X-4)-4X,$$

ou

$$c(X) = X^2 (X-5).$$

Logo os autovalores de A são $\lambda_1 = 0$ e $\lambda_2 = 5$.

A multiplicidade algébrica de $\lambda_1 = 0$ é 2 e de $\lambda_2 = 5$ é 1. Do teorema 6 de 1.4, vem que a multiplicidade geométrica de $\lambda_1 = 0$ é 1 ou 2 e de $\lambda_2 = 5$ é 1. Pela condição 3. do teorema 5 dessa secção (na versão para matrizes), A será diagonalizável se, e sòmente se, a multiplicidade geométrica de $\lambda_1 = 0$ for 2. Calculemos os autovetores linearmente independentes associados à $\lambda_1 = 0$. Como

$$A\ [x] = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 0 & 0 \\ -2 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 - 2x_3 \\ 0 \\ -2x_1 + 4x_3 \end{bmatrix} \text{ e } \lambda[x] = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \lambda x_3 \end{bmatrix}$$

então λ é um autovalor de A se, e sòmente se,

$$(\mathrm{I}) \begin{cases} x_1 - 2x_3 = \lambda x_1 \\ 0 = \lambda x_2 \\ -2x_1 + 4x_3 = \lambda x_3 \end{cases}$$

Sabemos que $\lambda_1 = 0$ é um autovalor de A, que substituindo em (I) ($\lambda = \lambda_1 = 0$), vem:

$$\begin{cases} x_1 = 2x_3 \\ 0 = 0.x_2. \end{cases}$$

Portanto, dois autovetores linearmente independentes associados à $\lambda_1 = 0$, são : (2,0,1) e (0,1,0). Logo o autoespaço associado à $\lambda_1 = 0$ é

$$V_{\lambda_1} = \langle (2,0,1), (0,1,0) \rangle ,$$

e a multiplicidade geométrica de $\lambda_1 = 0$, que é a dimensão de V_{λ_1}, é 2. Então, A é diagonalizável.

Uma matriz D de $M_3(\mathbb{R})$, semelhante à A é:

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 5 \end{bmatrix} \left(\text{ou } D_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ ou } D_2 = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right)$$

Vamos encontrar uma matriz C, diagonalizadora para A. Precisamos determinar um autovetor associado à $\lambda_2 = 5$. Em (I), fazemos $\lambda = \lambda_2 = 5$, obtemos:

$$\begin{cases} x_3 = -2x_1 \\ 5x_2 = 0, \quad \text{logo } V_{\lambda_2} = \langle (1,0,-2) \rangle \end{cases}$$

Pelo corolário 2 de 1.4, $\beta = \{(2,0,1),(0,1,0),(1,0,-2)\}$ é uma base ordenada de $\mathbb{R}^3$ e pelo que observamos antes dêsse exemplo,

$$C = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -2 \end{bmatrix} \left(\text{ou } C_1 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -2 & 0 \end{bmatrix} \text{ ou } C_2 = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} \right)$$

é uma matriz diagonalizadora para A.

Verifiquemos por exemplo, que $D_1 = C_1^{-1} A C_1$.

$$C_1^{-1} = \frac{1}{5} \begin{bmatrix} 2 & 0 & 1 \\ 1 & 0 & -2 \\ 0 & 5 & 0 \end{bmatrix}, \quad \text{logo}$$

$$C_1^{-1} A C_1 = \frac{1}{5} \begin{bmatrix} 2 & 0 & 1 \\ 1 & 0 & -2 \\ 0 & 5 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 0 & 0 \\ -2 & 0 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -2 & 0 \end{bmatrix} =$$

$$= \frac{1}{5} \begin{bmatrix} 2 & 0 & 1 \\ 1 & 0 & -2 \\ 0 & 5 & 0 \end{bmatrix} \begin{bmatrix} 0 & 5 & 0 \\ 0 & 0 & 0 \\ 0 & -10 & 0 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 25 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 0 \end{bmatrix} = D_1.$$

EXEMPLO 7

Seja $B = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 0 & 2 \\ 1 & -2 & 3 \end{bmatrix}$

O polinômio característico de B é $c(X) = X(X-2)^2$.
As multiplicidades algébricas e geométrica do autovalor $\lambda_1 = 0$ é 1. A multiplicidade algébrica do autovalor $\lambda_2 = 2$ é 2. Como, $(2E-B)\ [x] = [0]$ ou seja o sistema:

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ -2x_1 + 2x_2 - 2x_3 = 0 \\ -x_1 + 2x_2 + x_3 = 0 \end{cases} \iff \begin{cases} x_1 = -x_3 \\ x_2 = 0, \end{cases}$$

vem que $V_{\lambda_2} = \langle (1,0,-1) \rangle$, logo a multiplicidade geométrica de λ_2 é 1 e B não é diagonalizável.

Êsse exemplo 6, motiva a continuação de nosso estudo sôbre formas canônicas, já que gostaríamos de saber, qual a matriz, de forma mais simples possível de $M_3(\mathbb{R})$, semelhante à B. Veremos que é a matriz (de "Jordan ")

$$J = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

EXEMPLO 8.

Consideremos os exemplos 6 e 7 de 1.r. Vamos mostrar que as matrizes $\mathcal{L}$ e $\mathcal{T}$ não são semelhantes . De fato, $\mathcal{L}$ não é diagonalizável, pois como vimos lá, a multiplicidade geométrica do autovalor 1 é 1 e a algébrica é 2. Logo não pode existir matriz inversível C de $M_3(\mathbb{R})$, tal que

$$\mathcal{T} = C^{-1} \mathcal{L} C,$$

pois $\mathcal{T}$ está na forma diagonal.

CAPÍTULO III - FORMA CANÔNICA DE UMA MATRIZ POLINOMIAL.

3.1 OPERAÇÕES ELEMENTARES EM MATRIZES POLINOMIAIS.

Uma MATRIZ POLINOMIAL ou uma λ-MATRIZ, é uma matriz

$$A(\lambda) = \left[a_{ik}(\lambda)\right], \; m \times p,$$

onde cada elemento $a_{ik}(\lambda)$ é um polinômio na indeterminada λ, isto é, $a_{ik}(\lambda) \in K[\lambda]$, onde K é um corpo. Diz-se que n é o grau da matriz polinomial $A(\lambda)$, se n for o grau do polinômio de mais alto grau, entre os elementos de $A(\lambda)$.

É fácil de se ver que uma λ-matriz de grau n,

$$A(\lambda) = \left[a_{ik}^{(0)} \lambda^{n} + a_{ik}^{(1)} \lambda^{n-1} + \dots + a_{ik}^{(n-1)} \lambda + a_{ik}^{(n)}\right],$$

pode ser colocada na forma :-

$$A(\lambda) = A_0 \lambda^{n} + A_1 \lambda^{n-1} + \dots + A_{n-1}\lambda + A_n,$$

onde cada A_j é uma matriz constante, isto é, com elementos em K.

Matrizes polinomiais, ocorrem em muitas áreas da matemática. Por exemplo, consideremos o sistema de m-equações diferenciais lineares e homogeneas de ordem n, com coeficientes constantes, nas funções incógnitas $x_1, x_2, \dots, x_p$, da variável independente t :-

$$(1) \quad \begin{cases} a_{11}(D)\, x_1 + a_{12}(D)\, x_2 + \dots + a_{1p}(D) x_p = 0 \\ a_{21}(D)\, x_1 + a_{22}(D)\, x_2 + \dots + a_{2p}(D) x_p = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}(D)\, x_1 + a_{m2}(D)\, x_2 + \dots + a_{mp}(D) x_p = 0, \end{cases}$$

onde $a_{ik}(D) = a_{ik}^{(0)} D^{n} + a_{ik}^{(1)} D^{n-1} + \dots + a_{ik}^{(n-1)} D + a_{ik}^{(n)}$,

é um polinômio no operador linear $D = \frac{d}{dt}$ (derivação), com coeficientes constantes, $i = 1,2, \dots, m$ e $k = 1,2, \dots, p$.

O estudo do sistema (1), geralmente é feito, utilizando-se a matriz polinomial :-

$$A(D) = \left[a_{ik}(D)\right] \qquad (i = 1,2,\dots,m \; ; \; k = 1,2,\dots,p).$$

Quisemos com êste exemplo acima, apenas observar, que as "ferramentas" que ora pretendemos conseguir, não são úteis apenas na obtenção de formas canônicas de matrizes constantes (ver 4.1), mas são, por si mesmas, importantes em algumas áreas da matemática.

Introduziremos agora, as chamadas "operações elementares" no conjunto das matrizes polinomiais $m \times p$.

Dada uma λ-matriz $A(\lambda)$, $m \times p$, denomina-se OPERAÇÃO ELEMENTAR À ESQUERDA à uma das seguintes operações :-

1. Multiplicação de uma linha qualquer de $A(\lambda)$, por um escalar c, não nulo, do corpo K.

2. Adição à uma linha de $A(\lambda)$, de outra linha de $A(\lambda)$, multiplicada por um polinômio arbitrário $f(\lambda)$ de $K[\lambda]$.

Observa-se aqui dois fatos :-

1º As operações elementares à esquerda em $A(\lambda)$, dadas por :-

(a) multiplicação da linha i de $A(\lambda)$, por $c \neq 0$, $c \in K$.

(b) adição à linha i de $A(\lambda)$, da linha j multiplicada por $f(\lambda) \in K[\lambda]$,

são equivalentes, à multiplicarmos a λ-matriz $A(\lambda)$, à esquerda, respectivamente pelas λ-matrizes quadradas de ordem m,

$$S_1 = \begin{array}{cc} \begin{array}{cccccc} & & & (i) & & \end{array} & \\ \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & c & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} & \begin{array}{c} \\ \\ \\ (i) \\ \\ \\ \end{array} \end{array} \quad e \quad S_2 = \begin{array}{cc} \begin{array}{cccccccc} & & & (i) & & (j) & & \end{array} & \\ \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 1 & \cdots & f(\lambda) & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} & \begin{array}{c} \\ \\ \\ (i) \\ \\ \\ \end{array} \end{array}$$

2º Apezar de não tomarmos a troca entre duas linhas como operação elementar à esquerda, vamos mostrar que tal operação pode ser obtida aplicando-se convenientemente, quatro vêzes, operações elementares à esquerda. Procede-se da seguinte maneira, para se obter a troca das linhas i e j de $A(\lambda)$:-

(a) Adiciona-se à linha j, a linha i multiplicada por $f(\lambda) = 1$.

(b) Adiciona-se à linha i, a linha j (nova) multiplicada por $f(\lambda) = -1$.

(c) Adiciona-se à linha j, a linha i multiplicada por $f(\lambda) = 1$.

(d) Multiplica-se a linha i por $c = -1 \neq 0$.

Graficamente, pode-se acompanhar o procedimento acima, como segue :-

$$A(\lambda) = \begin{bmatrix} & \vdots & \\ \cdots & a_{ik}(\lambda) & \cdots \\ & \vdots & \\ \cdots & a_{jk}(\lambda) & \cdots \\ & \vdots & \end{bmatrix} \xrightarrow{(a)} \begin{bmatrix} & \vdots & \\ \cdots & a_{ik}(\lambda) & \cdots \\ & \vdots & \\ \cdots & a_{ik}(\lambda) + a_{jk}(\lambda) & \cdots \\ & \vdots & \end{bmatrix} \xrightarrow{(b)}$$

$$\begin{bmatrix} & \vdots & \\ \cdots & -a_{jk}(\lambda) & \cdots \\ & \vdots & \\ \cdots & a_{ik}(\lambda) + a_{jk}(\lambda) & \cdots \\ & \vdots & \end{bmatrix} \xrightarrow{(c)} \begin{bmatrix} & \vdots & \\ \cdots & -a_{jk}(\lambda) & \cdots \\ & \vdots & \\ \cdots & a_{ik}(\lambda) & \cdots \\ & \vdots & \end{bmatrix} \xrightarrow{(d)} \begin{bmatrix} & \vdots & \\ \cdots & a_{jk}(\lambda) & \cdots \\ & \vdots & \\ \cdots & a_{ik}(\lambda) & \cdots \\ & \vdots & \end{bmatrix}$$

Pode-se de modo análogo definir <u>OPERAÇÃO ELEMENTAR À DIREITA</u> sôbre uma matriz polinomial,

$$A(\lambda) = \left[a_{ik}(\lambda)\right], \quad i = 1,2,\ldots,m; \ k = 1,2,\ldots,p,$$

pondo-se :-

<u>1'.</u> Multiplicação de uma coluna qualquer de $A(\lambda)$, por um escalar c, não nulo, do corpo K.

<u>2'.</u> Adição à uma coluna de $A(\lambda)$, de outra coluna de $A(\lambda)$, multiplicada por um polinômio $f(\lambda)$ de $K[\lambda]$.

Valem aqui observações idênticas às anteriores, sendo que as matrizes S_1 e S_2, são substituidas pelas matrizes quadradas de ordem p,

$$T_1 = \begin{array}{c} \begin{array}{cccccc} & & & (i) & & \end{array} \\ \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & c & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \end{array} \begin{array}{c} \\ \\ \\ \\ (i) \\ \\ \\ \end{array} \qquad T_2 = \begin{array}{c} \begin{array}{cccccc} & & & (i) & & \end{array} \\ \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & f(\lambda) & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \end{array} \begin{array}{c} \\ \\ \\ \\ (i) \\ \\ (j) \\ \\ \\ \end{array}$$

Isto é, o efeito das seguintes operações elementares à direita :-

1') multiplicação da coluna (i) de $A(\lambda)$, por $c \neq 0, c \in K$.

2') adição à coluna i de $A(\lambda)$, da coluna j multiplicada por $f(\lambda) \in K[\lambda]$

é equivalente a multiplicação à direita de $A(\lambda)$, por T_1 e T_2, respectivamente. Anàlogamente, a troca entre duas colunas pode ser realizada, através de quatro operações elementares à direita.

Denomina-se MATRIZ ELEMENTAR, a uma matriz do tipo:-
S_1, S_2, T_1 ou T_2.

LEMA 1.

" O determinante de uma matriz elementar, é constante (não depende de λ) e não nulo".

Demonstração.

É óbvio que $\det S_1 = \det T_1 = c \neq 0$. Vamos mostrar que $\det S_2 =$
$= \det T_2 = 1$.

Para S_2, desenvolve-se o determinante por Laplace, segundo a primeira linha, repetidas vêzes, até obtermos :-

$$\det S_2 = \begin{vmatrix} 1 & 0 & .. & f(\lambda) & .. & 0 \\ 0 & 1 & .. & 0 & .. & 0 \\ . & . & .. & . & .. & . \\ 0 & 0 & .. & 1 & .. & 0 \\ . & . & .. & . & .. & . \\ 0 & 0 & .. & 0 & .. & 1 \end{vmatrix} = 1+(-1)^{1+j} f(\lambda) \det \begin{bmatrix} 0 & 1 & 0 & .. & 0 \\ 0 & 0 & 1 & .. & 0 \\ . & . & . & .. & . \\ 0 & 0 & 0 & .. & 1 \\ 0 & 0 & 0 & .. & 0 \end{bmatrix} = 1.$$

De modo análogo, verifica-se que $\det T_2 = 1$.

c.q.d.

Do lema 1, concluimos que cada operação elementar à esquerda (ou à direita), possui uma operação inversa, que é também uma operação elementar à esquerda (ou à direita).

Duas matrizes polinomiais $m \times p$, $A(\lambda)$ e $B(\lambda)$, denominam-se:-
(1) EQUIVALENTES À ESQUERDA,
(2) EQUIVALENTES À DIREITA,
(3) EQUIVALENTES,
quando for possível, obter-se uma da outra, através de um número finito de :-
(1) operações elementares À ESQUERDA,
(2) operações elementares À DIREITA,
(3) operações elementares À ESQUERDA E À DIREITA,
respectivamente.

3.2 FORMAS CANÔNICAS DE λ-MATRIZES.

Inicialmente obteremos uma forma canônica de uma λ-matriz de grau n, $A(\lambda) = \left[a_{ik}(\lambda)\right]$, $i=1,2,\ldots,m$ e $k=1,2,\ldots,p$, que quando m=p, será a FORMA TRIANGULAR.

Suponhamos que os elementos (polinômios) da primeira coluna de $A(\lambda)$, não sejam todos nulos(caso contrário, passamos à segunda coluna no procedimento que estamos estabelecendo). Entre os $a_{i1}(\lambda)$, $i = 1,2, \ldots,m$, fixemos o de menor grau $a_{i_o1}(\lambda)$ e através de troca entre linhas de $A(\lambda)$, o colocamos na posição (1,1) da matriz $A(\lambda)$ (a posição (i,k), é na linha i e coluna k). A seguir, dividimos cada $a_{i1}(\lambda)$; $i=2,\ldots,m$ por $a_{i_o1}(\lambda)$. Seja :-

$a_{i1}(\lambda) = q_{i1}(\lambda)a_{i_o1}(\lambda) + r_{i1}(\lambda)$, onde $r_{i1}(\lambda) = o$ ou $gr(r_{i1}) < gr(a_{i_o1})$.

A seguir, adiciona-se à primeira linha, a linha i multiplicada por $q_{i1}(\lambda)$, $i = 2,\ldots,m$. Caso $r_{i1}(\lambda) = o$, $i =2,\ldots,m$, ótimo. Toma-se $b_{11}(\lambda) = a_{i_o1}(\lambda)$ e passamos à segunda coluna. Caso contrário, escolhemos entre os $r_{i1}(\lambda)$ o de menor grau, e o colocamos na posição (1,1), através de uma nova troca de linhas. Obtém-se assim, na posição (1,1) da matriz equivalente à esquerda à $A(\lambda)$, um polinômio de grau menor que o grau de $a_{i_o1}(\lambda)$. Repete-se êste proceder, até que os elementos das posições (i,1), $i = 2,\ldots,m$, sejam todos nulos, o que certamente acontecerá, pois o grau de $A(\lambda)$ é n, finito.

Passamos a seguir, a segunda coluna. Ao elemento $a'_{22}(\lambda)$ aplicamos o mesmo procedimento acima, para as linhas $2,3,\ldots,m$. Acabamos por obter os elementos das posições (i,2), $i=3,\ldots,m$, nulos. Observa-se ainda que com êste procedimento, a primeira coluna permanece inalterada. Seja $b_{22}(\lambda)$ o elemento da posição (2,2) desta última matriz. Caso $b_{22}(\lambda) \neq o$, dividimos o elemento $a'_{12}(\lambda)$ por $b_{22}(\lambda)$, obtemos :-

$a'_{12}(\lambda) = q_{12}(\lambda)b_{22}(\lambda) + r_{12}(\lambda)$, onde $r_{12}(\lambda)=o$ ou $gr(r_{12}) < gr(b_{22})$.

Sendo $r_{12}(\lambda) = o$, ótimo. Caso contrário, adiciona-se à primeira linha, a segunda multiplicada por $-q_{12}(\lambda)$. Obtemos na posição (1,2), $r_{12}(\lambda)$. Assim, o elemento (1,2) é identicamente nulo ou tem grau menor que o grau de $b_{22}(\lambda)$. Em particular, se $gr(b_{22}) = o$, o elemento (1,2) será nulo. Ainda esta última operação elementar à esquerda, que utilizamos, não alterou a primeira coluna, pois o elemento $(2,1)$ é nulo.

Continuando com êste proceder, mais precisamente por indução sôbre m, reduzimos a λ-matriz $A(\lambda)$, à uma sua equivalente à esquerda, na seguinte FORMA CANÔNICA :-

$$(2)\quad \begin{bmatrix} b_{11}(\lambda) & b_{12}(\lambda) & \cdots & b_{1m}(\lambda) & \cdots & b_{1p}(\lambda) \\ 0 & b_{22}(\lambda) & \cdots & b_{2m}(\lambda) & \cdots & b_{2p}(\lambda) \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & b_{mm}(\lambda) & \cdots & b_{mp}(\lambda) \end{bmatrix} \quad (\text{caso } m \leqslant p).$$

ou

$$(3)\quad \begin{bmatrix} b_{11}(\lambda) & b_{12}(\lambda) & \cdots & b_{1p}(\lambda) \\ 0 & b_{22}(\lambda) & \cdots & b_{2p}(\lambda) \\ 0 & 0 & \cdots & b_{3p}(\lambda) \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & b_{pp}(\lambda) \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}. \quad (\text{caso } p \leqslant m).$$

Sendo $m = p$, isto é, $A(\lambda)$ quadrada de ordem m (p), a forma canônica acima obtida, será a TRIANGULAR SUPERIOR . Acabamos de demonstrar o seguinte teorema :-

TEOREMA 1.

"Uma λ-matriz $A(\lambda) = \left[a_{ik}(\lambda)\right]$, $i = 1,2, \ldots,m$ e $k = 1,2, \ldots,p$, de grau n(finito), pode sempre ser colocada na forma canônica (2) ou (3), através de operações elementares à esquerda. Sendo que, os polinômios $b_{1j}(\lambda), \ldots, b_{j-1\ j}(\lambda)$ têm grau menor que o grau de $b_{jj}(\lambda)$, quando $b_{jj}(\lambda) \neq 0$, ou são todos nulos, quando $b_{jj}(\lambda) = \text{cte} \neq 0$, $j = 2,3,\ldots, \min\{m,p\}$".

Anàlogamente, poderíamos obter uma forma canônica (que seria TRIANGULAR INFERIOR, se $m=p$), equivalente à direita à $A(\lambda)$, utilizando apenas operações elementares à direita.

Como uma aplicação da forma canônica do teorema 1, mostraremos que as seguintes condições, são equivalentes :-

1. As matrizes polinomiais $m \times p$, $A(\lambda)$ e $B(\lambda)$, são equivalentes à esquerda.

2. Existe uma matriz polinomial $m \times m$, $P(\lambda)$, com determinante constante (independente de λ) e não nulo, tal que :-

$$B(\lambda) = P(\lambda)\, A(\lambda).$$

1. $\Longrightarrow$ **2.**

Sendo $B(\lambda)$ equivalente à esquerda à $A(\lambda)$, existem matrizes elementares $m \times m$, $S^{(1)}(\lambda), \dots, S^{(r)}(\lambda)$, tais que :-

$$B(\lambda) = S^{(r)}(\lambda)\, S^{(r-1)}(\lambda) \dots S^{(1)}(\lambda)\, A(\lambda).$$

Pelo lema 1, sabemos que $\det S^{(i)}(\lambda) = c_i \neq 0$, $i = 1, 2, \dots, r$. Toma--se :

$$P(\lambda) = S^{(r)}(\lambda)\, S^{(r-1)} \dots S^{(1)}(\lambda) ,$$

e obtemos **2.**, pois ;

$$\det P(\lambda) = \det S^{(r)}(\lambda) \det S^{(r-1)}(\lambda) \dots \det S^{(1)}(\lambda) = \prod_{i=1}^{r} c_i = c \neq 0.$$

2. $\Longrightarrow$ **1.**

Pelo teorema 1, acima, a matriz $P(\lambda)$ é equivalente à esquerda, a uma matriz da forma :-

$$(4) \qquad \begin{bmatrix} b_{11}(\lambda) & b_{12}(\lambda) & \dots & b_{1m}(\lambda) \\ 0 & b_{22}(\lambda) & \dots & b_{2m}(\lambda) \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & b_{mm}(\lambda) \end{bmatrix} .$$

Como $\det P(\lambda) = c \neq 0$ e a matriz (4) foi obtida, multiplicando-se à esquerda, $P(\lambda)$ por matrizes elementares, cujos determinantes são constantes e não nulos (lema 1), o determinante da matriz (4) é constante e não nulo. Isto é,

$$b_{11}(\lambda)\, b_{22}(\lambda) \dots b_{mm}(\lambda) = \text{cte} \neq 0,$$

donde se conclui que $b_{kk}(\lambda) = b_k = \text{cte} = \text{cte} \neq 0$, $k=1,2,\dots,m$. Pelo teorema 1, a matriz (4) é diagonal e constante, isto é, é da forma $\left[\, b_k \delta_{ik} \,\right]$, onde $b_k = \text{cte} \neq 0$ e $\delta_{ik} = \begin{cases} 1 \text{ se } i = k \\ 0 \text{ se } i \neq k \end{cases}.$

Assim sendo, multiplica-se cada linha i de (4) por b_i^{-1} (operação elementar à esquerda), para $i = 1,2,\ldots,m$, reduzimos (4) à matriz unidade constante E. Assim, $P(\lambda)$ é equivalente à esquerda, à matriz unidade E. Logo,

$$P(\lambda) = S^{(r)}(\lambda)\ S^{(r-1)}(\lambda)\ \ldots\ S^{(1)}(\lambda)\ E = S^{(r)}(\lambda)\ S^{(r-1)}(\lambda)\ldots S^{(1)}(\lambda),$$

ou seja, $P(\lambda)$ é um produto de um número finito de matrizes elementares à esquerda e como :-

$$B(\lambda) = P(\lambda)\ A(\lambda) = S^{(r)}(\lambda)\ S^{(r-1)}(\lambda)\ \ldots\ S^{(1)}(\lambda)\ A(\lambda),$$

temos que $B(\lambda)$ e $A(\lambda)$ são equivalentes à esquerda.

Observamos que na demonstração da equivalência <u>1.</u> $\Longleftrightarrow$ <u>2.</u>, demonstrou-se também que :-

<u>COROLÁRIO 1.</u>

"Tôda λ-matriz quadrada $P(\lambda)$ com determinante constante e não nulo, pode ser escrito como o produto de um número finito de matrizes elementares à esquerda".

É claro que poderíamos obter resultados análogos para matrizes equivalentes à direita ou matrizes equivalentes. Pode-se assim, dar as seguintes formas alternativas, para que as λ-matrizes $A(\lambda)$ e $B(\lambda)$ sejam :-

<u>1.</u> equivalentes à esquerda,

<u>2.</u> equivalentes à direita,

<u>3.</u> equivalentes,

se existirem -matrizes $P(\lambda)$ e $Q(\lambda)$, quadradas de ordens n e p, respectivamente, e com determinantes constantes e não nulos tais que :-

<u>1.</u> $B(\lambda) = P(\lambda)\ A(\lambda)$

<u>2.</u> $B(\lambda) = A(\lambda)\ Q(\lambda)$

<u>3.</u> $B(\lambda) = P(\lambda)\ A(\lambda)\ Q(\lambda)$,

respectivamente.

Vamos obter agora uma forma canônica, na qual tôda λ-matriz pode ser levada, através de operações elementares à direita e à esquerda, e que é a "FORMA CANÔNICA DIAGONAL".

Seja,

$$A(\lambda) = \left[a_{ik}(\lambda)\right],\ i = 1,2,\ldots,m \ \text{ e } \ k = 1,2,\ldots,p,$$

uma λ-matriz de grau n (finito). Quando todos os elementos de

λ-matriz $A(\lambda)$ são nulos, $A(\lambda)$ já está numa forma bem simples, é nula. Suponhamos que $A(\lambda)$ possua elementos não nulos. Pode-se supor que $a_{11}(\lambda) \neq o$ e $gr(a_{11}) \leqslant gr(a_{ik})$, se $a_{ik}(\lambda) \neq o$, $i=1,2,\ldots,m$ e $k=1,2,\ldots,p$. Pois se tal não ocorrer, colocamos na posição (1,1), um elemento não nulo de menor grau de $A(\lambda)$, através de troca entre linhas ou colunas. Consideremos a seguir, o quociente de cada elemento $a_{i1}(\lambda)$, $i=2,\ldots,m$ e $a_{1k}(\lambda), k = 2,3,\ldots,p$, por $a_{11}(\lambda)$. Seja :-

$$(5) \quad \begin{cases} a_{i1}(\lambda) = q_{i1}(\lambda)\ a_{11}(\lambda) + r_{i1}(\lambda),\ r_{i1}(\lambda) = o \text{ ou } gr(r_{i1}) < gr(a_{11}). \\ a_{1k}(\lambda) = q_{1k}(\lambda)\ a_{11}(\lambda) + r_{1k}(\lambda),\ r_{1k}(\lambda) = o \text{ ou } gr(r_{1k}) < gr(a_{11}). \end{cases}$$

Caso todos os polinômios $r_{i1}(\lambda)$ e $r_{1k}(\lambda)$ em (5) são nulos, ótimo. Admitamos que um dos restos, digamos $r_{1k_o}(\lambda)$, não é identicamente nulo. Adicionamos à coluna k_o, a primeira coluna multiplicada por $-q_{1k_o}(\lambda)$, dado por (5). Com esta operação elementar, o resto $r_{1k_o}(\lambda)$, ocupará a posição $(1,k_o)$. Troca-se as colunas 1 e k_o, fazendo com que $r_{1k_o}(\lambda)$, de grau menor que o grau de $a_{11}(\lambda)$, ocupe a posição (1,1). Repetimos o procedimento, até que todos os restos sejam nulos, o que certamente ocorrerá, pois o grau de $A(\lambda)$ é n, finito. Depois que todos os retos são nulos, adiciona-se à linha i (a coluna k), a primeira linha (coluna) multiplicada por $-q_{i1}(\lambda)$ ($-q_{1k}(\lambda)$), $i = 2,\ldots,m$ ($k=2,\ldots,p$), dados por (5). Obtemos assim $a_{21}(\lambda) = \ldots = a_{m1}(\lambda) = o$ ($a_{12}(\lambda) = \ldots = a_{1p}(\lambda) = o$). Isto é, chegamos à seguinte matriz equivalente à $A(\lambda)$:-

$$\begin{bmatrix} b_{11}(\lambda) & o & o & \ldots\ o & o \\ o & b_{22}(\lambda) & b_{23}(\lambda) & \ldots & b_{2p}(\lambda) \\ \circ & \circ & \circ & \circ\circ\circ & \circ \\ \circ & \circ & \circ & \circ\circ\circ & \circ \\ o & b_{m2}(\lambda) & b_{m3}(\lambda) & \ldots & b_{mp}(\lambda) \end{bmatrix} .$$

Caso algum $b_{ik}(\lambda)$, $i = 2,\ldots,m$ e $k = 2, \ldots ,p$, não seja divisível por $b_{11}(\lambda)$, adiciona-se à primeira coluna, a coluna que tem um tal elemento não divisível por $b_{11}(\lambda)$. Podemos com o procedimento anterior (através de divisão do tal elemento por $b_{11}(\lambda)$, etc) substituir o elemento $b_{11}(\lambda)$ por um polinômio de grau menor que o grau de $b_{11}(\lambda)$. Desde que o grau de $A(\lambda)$ é n, finito, a redução do grau do elemento na posição (1,1), não pode continuar indefinidamente. Assim, depois de um número finito de operações elementares, obtemos a λ-matriz $A(\lambda)$ na forma :-

$$(6)\qquad \begin{bmatrix} a_1(\lambda) & 0 & \cdots & 0 \\ 0 & c_{22}(\lambda) & \cdots & c_{2p}(\lambda) \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & c_{m2}(\lambda) & \cdots & c_{mp}(\lambda) \end{bmatrix},$$

onde todos os elementos $c_{ik}(\lambda)$, $i = 2,\ldots,m$ e $k = 2,\ldots,p$, são divisíveis por $a_1(\lambda)$.

Caso todos os elementos $c_{ik}(\lambda) = o$, ótimo. Caso contrário, continuamos o processo de redução para as linhas 2, ..., m e as colunas 2, ... ,p, reduzindo a matriz (6) à forma :-

$$(7)\qquad \begin{bmatrix} a_1(\lambda) & 0 & \cdots\ 0 & \cdots\ 0 \\ 0 & a_2(\lambda) & \cdots\ 0 & \cdots\ 0 \\ \cdot & & \cdots\ \cdot & \cdots\ \cdot \\ \cdot & \cdot & \cdots\ \cdot & \cdots\ \cdot \\ 0 & 0 & \cdots\ a_s(\lambda) & \cdots\ 0 \\ 0 & 0 & \cdots\ 0 & \cdots\ 0 \\ \cdot & \cdot & \cdots\ \cdot & \cdots\ \cdot \\ 0 & 0 & \cdots\ 0 & \cdots\ 0 \end{bmatrix},$$

onde os polinômios $a_1(\lambda)$, $a_2(\lambda)$, ... , $a_s(\lambda)$, $s \leq \min\{m,p\}$, não são nulos e $a_{i+1}(\lambda)$ é divisível por $a_i(\lambda)$, $i = 1,2, \ldots ,s-1$. Ademais podemos supor que em (7), cada $a_i(\lambda)$, $i = 1,2, \ldots ,s$ é um polinômio unitário, bastando para isso multiplicar cada linha i de (7) por convenientes constantes não nulas (recíprocos em K, do coeficiente dominante de $a_i(\lambda)$). Vê-se de (7), que <u>s</u> é a <u>CARACTERÍSTICA</u> de $A(\lambda)$, ou seja, <u>s</u> é a ordem do "maior subdeterminante" de $A(\lambda)$, não nulo. A forma (7) denomina-se <u>FORMA CANÔNICA DIAGONAL</u> de $A(\lambda)$. Acabamos de mostrar o seguinte teorema :-

<u>TEOREMA 2.</u>

" Tôda λ-matriz m×p $A(\lambda)$ de grau n (finito), é equivalente à uma λ-matriz na forma canônica diagonal".

EXEMPLO 1.

Vamos obter a forma canônica diagonal, para a λ-matriz :-

$$A(\lambda) = \begin{bmatrix} \lambda^2+5\lambda-2 & 4\lambda & 2\lambda-1 & -4\lambda+1 \\ \lambda+5 & 3 & 2 & -4 \\ -10\lambda^3-2 & -7\lambda^3 & \lambda^4-4\lambda^3-1 & 9\lambda^3+\lambda \\ -10 & -7 & \lambda-4 & 9 \end{bmatrix}.$$

Para facilidade de indicação das operações elementares, introduzimos a seguinte notação :-

1) $(i(c))$ (ou $[i(c)]$) significa que multiplicamos a linha (coluna) i por $c \neq 0$, $c \in K$.

2) $(i(p(\lambda))+h)$ (ou $[i(p(\lambda))+h]$) significa que adicionamos à linha (coluna) h, $p(\lambda)$ multiplicado pela linha (coluna) i.

3) (i,h) (ou $[i,h]$) indica a troca, que é um número finito de operações elementares, entre as linhas (colunas) i e h.

$$A(\lambda) \underset{(4(-\lambda^3)+3)}{\overset{(2(-\lambda)+1)}{\sim}} \begin{bmatrix} -2 & \lambda & -1 & 1 \\ \lambda+5 & 3 & 2 & -4 \\ -2 & 0 & -1 & \lambda \\ -10 & -7 & \lambda-4 & 9 \end{bmatrix} \underset{[3(\lambda)+4]}{\overset{[3(-2)+1]}{\sim}}$$

$$\sim \begin{bmatrix} 0 & \lambda & -1 & -\lambda+1 \\ \lambda+1 & 3 & 2 & 2\lambda+4 \\ 0 & 0 & -1 & 0 \\ -2\lambda-2 & -7 & \lambda-4 & \lambda^2-4\lambda+9 \end{bmatrix} \overset{(2(2)+4)}{\sim} \begin{bmatrix} 0 & \lambda & -1 & -\lambda+1 \\ \lambda+1 & 3 & 2 & 2\lambda-4 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & \lambda & \lambda^2+1 \end{bmatrix}$$

$$\underset{(3(\lambda)+4)}{\overset{\substack{(3(-1)+1)\\(3(2)+2)}}{\sim}} \begin{bmatrix} 0 & \lambda & 0 & -\lambda+1 \\ \lambda+1 & 3 & 0 & 2\lambda-4 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & \lambda^2+1 \end{bmatrix} \quad \overset{2(\lambda^2+1)+4}{\sim}$$

$$\sim \begin{bmatrix} 0 & \lambda & 0 & \lambda^3+1 \\ \lambda+1 & 3 & 0 & 3\lambda^2+2\lambda-1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \underset{(4(3)+2)}{\overset{(4(\lambda)+1)}{\sim}} \begin{bmatrix} 0 & 0 & 0 & \lambda^3+1 \\ \lambda+1 & 0 & 0 & 3\lambda^2+2\lambda-1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}$$

$$\begin{array}{c}[1(-3\lambda+1)+4]\\ [2(-1)]\\ \sim\\ [3(-1)]\end{array}\begin{bmatrix} 0 & 0 & 0 & \lambda^3+1 \\ \lambda+1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}\begin{array}{c}(1,4)\\ \sim\\ (2,3)\end{array}$$

$$\sim\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \lambda+1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda^3+1 \end{bmatrix}\overset{[1,3]}{\sim}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \lambda+1 & 0 & 0 \\ 0 & 0 & 0 & \lambda^3+1 \end{bmatrix}\overset{[2,3]}{\sim}$$

$$\sim\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \lambda+1 & 0 \\ 0 & 0 & 0 & \lambda^3+1 \end{bmatrix}.$$

<u>EXEMPLO 2.</u>

Seja ,

$$A(\lambda) = \begin{bmatrix} \lambda^4+\lambda^2 & \lambda^6-\lambda^2 & \lambda^4+\lambda^2 \\ -\lambda^3+\lambda^2-\lambda+1 & 2\lambda^3-2\lambda^2+2\lambda-2 & \lambda^2+1 \\ \lambda^2+1 & -2\lambda^2-2 & \lambda^2+1 \end{bmatrix}.$$

Queremos encontrar matrizes $P(\lambda)$ e $Q(\lambda)$, quadradas de ordem 3, com determinantes constantes e não nulos, tais que :-

$$D(\lambda) = P(\lambda)\, A(\lambda)\, Q(\lambda),$$

esteja na forma canônica diagonal.

Pela aplicação que fizemos do teorema 1, (<u>1.</u> $\Longleftrightarrow$ <u>2.</u>), para encontrarmos $P(\lambda)$ (ou $Q(\lambda)$), nas condições pedidas, basta-nos colocar $A(\lambda)$ na forma canônica diagonal, enumerando as operações elementares (ou troca entre linhas (colunas)) à esquerda (à direita). Depois, aplica-se a mesma seqüência de operações elementares na matriz identidade E.

$$A(\lambda)\begin{array}{c}[1(2)+2]\\ \sim\\ [1(-1)+3]\end{array}\begin{bmatrix} \lambda^4+\lambda^2 & \lambda^6+2\lambda^4+\lambda^2 & 0 \\ -\lambda^3+\lambda^2-\lambda+1 & 0 & \lambda^3+\lambda \\ \lambda^2+1 & 0 & 0 \end{bmatrix}\begin{array}{c}(3(-\lambda^2)+1)\\ \sim\\ (3(\lambda-1)+2)\end{array}$$

$$\sim \begin{bmatrix} 0 & \lambda^6+2\lambda^4+\lambda^2 & 0 \\ 0 & 0 & \lambda^3+\lambda \\ \lambda^2+1 & 0 & 0 \end{bmatrix} \underset{[2,3]}{\overset{(1,3)}{\sim}} \begin{bmatrix} \lambda^2+1 & 0 & 0 \\ 0 & \lambda(\lambda^2+1) & 0 \\ 0 & 0 & \lambda^2(\lambda^2+1)^2 \end{bmatrix} = D(\lambda).$$

Para obtermos $P(\lambda)$, efetuamos na matriz identidade a mesma sequência de operações elementares à esquerda usadas na obtenção de $D(\lambda)$.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \underset{(3(\lambda-1)+2}{\overset{(3(-\lambda^2)+1)}{\sim}} \begin{bmatrix} 1 & 0 & -\lambda^2 \\ 0 & 1 & \lambda-1 \\ 0 & 0 & 1 \end{bmatrix} \overset{(1,3)}{\sim} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & \lambda-1 \\ 1 & 0 & -\lambda^2 \end{bmatrix} = P(\lambda).$$

Anàlogamente, para a obtenção da matriz $Q(\lambda)$, efetuamos na matriz identidade a mesma sequência de operações elementares à direita, usadas na obtenção de $D(\lambda)$.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \underset{[1(-1)+3]}{\overset{[1(2)+2]}{\sim}} \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \overset{[2,3]}{\sim} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = Q(\lambda).$$

A título de verificação, mostraremos que de fato :-

$$D(\lambda) = P(\lambda)\ A(\lambda)\ Q(\lambda),$$

e que $\det P(\lambda) = c \neq 0$ e $\det Q(\lambda) = c_1 \neq 0$.

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & \lambda-1 \\ 1 & 0 & -\lambda^2 \end{bmatrix} \begin{bmatrix} \lambda^4+\lambda^2 & \lambda^6-\lambda^2 & \lambda^4+\lambda^2 \\ -\lambda^3+\lambda^2-\lambda+1 & 2\lambda^3-2\lambda^2+2\lambda-2 & \lambda^2+1 \\ \lambda^2+1 & -2\lambda^2-2 & \lambda^2+1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} =$$

$$= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & \lambda-1 \\ 1 & 0 & -\lambda^2 \end{bmatrix} \begin{bmatrix} \lambda^4+\lambda^2 & 0 & \lambda^6+2\lambda^4+\lambda^2 \\ -\lambda^3+\lambda^2-\lambda+1 & \lambda^3+\lambda & 0 \\ \lambda^2+1 & 0 & 0 \end{bmatrix} =$$

$$= \begin{bmatrix} \lambda^2+1 & 0 & 0 \\ 0 & \lambda(\lambda^2+1) & 0 \\ 0 & 0 & \lambda^2(\lambda^2+1)^2 \end{bmatrix} = D(\lambda).$$

$$\det P(\lambda) = \begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & \lambda-1 \\ 1 & 0 & -\lambda^2 \end{vmatrix} \underset{=}{(\text{Laplace 1ª coluna})} (-1)^{3+1} \begin{vmatrix} 0 & 1 \\ 1 & \lambda-1 \end{vmatrix} = -1 \neq 0.$$

$$\det Q(\lambda) = \begin{vmatrix} 1 & -1 & 2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix} \underset{=}{(\text{idem})} \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1 \neq 0.$$

3.3 POLINÔMIOS INVARIANTES E DIVISORES ELEMENTARES.

Nesta secção, a menos que se explicite o contrário, $A(\lambda)$ é uma λ-matriz, $m \times p$ de característica s e grau n (finito). Para cada $k = 1, 2, \ldots, s$, todos os subdeterminantes de ordem k, da λ-matriz $A(\lambda)$, são polinômios em λ. Seja $D_k(\lambda)$ o m.d.c. (unitário) entre todos os subdeterminantes, não nulos, de ordem k, da λ-matriz $A(\lambda)$. Definimos ainda $D_0(\lambda) = 1$. Nesta nomenclatura, colocamos o seguinte lema :-

LEMA 2.

" Cada $D_k(\lambda)$ é invariante sob operações elementares na matriz polinomial $A(\lambda)$. Isto é, se $B(\lambda)$ é equivalente à $A(\lambda)$, então $A(\lambda)$ e $B(\lambda)$, possuem os mesmos $D_k(\lambda)$, $k = 1, 2, \ldots, s$".

Demonstração.

No caso de operações elementares do tipo 1 (ou 1'), multiplicação de uma linha (coluna) de $A(\lambda)$ por uma constante $c \neq 0$, os $D_k(\lambda)$ não se alteram. De fato, nêste caso, os subdeterminantes de ordem k de $B(\lambda)$ coincidem com os de $A(\lambda)$ à menos de uma constante (c ou 1) multiplicativa, não nula. Sendo $D_k(\lambda)$ um polinômio unitário, concluimos que $A(\lambda)$ e $B(\lambda)$ possuem os mesmos $D_k(\lambda)$.

Consideremos agora, $B(\lambda)$ obtida de $A(\lambda)$ através de uma operação elementar do tipo 2 (ou 2'). Especificamente, seja $B(\lambda)$ obtida de $A(\lambda)$, adicionando-se à coluna i, a coluna h multiplicada por $b(\lambda)$ (o caso de linha é análogo). Seja o caso, no qual um subdeterminante de ordem k de $A(\lambda)$, não contém nenhum elemento de ambas as colunas i e h. Então, o respectivo subdeterminante de ordem k, de $B(\lambda)$, coincide com o correspondente de $A(\lambda)$, ou pode ser expresso como uma soma dêste correspondente de $A(\lambda)$, com $b(\lambda)$ vêzes um determinante com as colunas i e h iguais, logo nulo. No caso restante, de um subdeterminante de $A(\lambda)$, no qual elementos da coluna i faz parte, mas da coluna h não, o correspondente subdeterminante de $B(\lambda)$, é uma combinação linear de subdeterminantes de ordem k de $A(\lambda)$. Da definição de m.d.c. o $D_k(\lambda)$ de $A(\lambda)$, divide o $D_k'(\lambda)$ (correspondente) de $B(\lambda)$. Adicionando-se à coluna i a coluna h multiplicada por $-b(\lambda)$, obtemos $A(\lambda)$ de $B(\lambda)$, logo com a mesma argumentação, $D_k'(\lambda)$ divide $D_k(\lambda)$. Sendo $D_k(\lambda)$ e $D_k'(\lambda)$ unitários, $D_k(\lambda) = D_k'(\lambda)$.

c.q.d.

Com base nêste Lema 2 e no teorema 2 de 3.2, pretendemos encontrar os $D_k(\lambda)$, $k = 1,2, \ldots ,s$, de uma λ-matriz $A(\lambda)$. Para tanto, reduzimos $A(\lambda)$ a forma canônica diagonal (7) de 3.2, e os $D_k(\lambda)$ de $A(\lambda)$ são, pelo Lema 2, os mesmos que os $D_k'(\lambda)$ desta forma canônica diagonal. Para cada $k = 1,2, \ldots ,s$, um subdeterminante $E_k(\lambda)$, não nulo, de ordem k de (7), é da forma :-

$$E_k(\lambda) = \begin{vmatrix} a_{j_1}(\lambda) & 0 & \cdots & 0 \\ 0 & a_{j_2}(\lambda) & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & a_{j_k}(\lambda) \end{vmatrix},$$

onde $1 \leq j_1 < j_2 < \ldots < j_k \leq s$. Isto é,

(8) $$E_k(\lambda) = a_{j_1}(\lambda)\, a_{j_2}(\lambda) \ \ldots \ a_{j_k}(\lambda) .$$

Um dêstes subdeterminantes não nulos de ordem k é :-

(9) $$a_1(\lambda)\, a_2(\lambda) \ \ldots \ a_k(\lambda) .$$

Como na forma canônica diagonal (7), $a_i(\lambda)$ divide $a_{i+1}(\lambda)$, $i = 1,2, \ldots, s-1$, vem que em (8), $a_h(\lambda)$ é um divisor de $a_{j_h}(\lambda)$, $h = 1,2, \ldots ,k$. Logo (9) é um divisor comum dos $E_k(\lambda)$, e sendo êle próprio um polinômio unitário($i = 1,2, \ldots$, s em (7) $a_i(\lambda)$ é unitário), concluimos que :-

(10) $$D_k(\lambda) = a_1(\lambda)\, a_2(\lambda) \ \ldots \ a_k(\lambda), \quad k = 1,2, \ldots ,s.$$

De (10), é fácil de se ver, que na sequência :-

$D_0(\lambda) = 1,\ D_1(\lambda) = a_1(\lambda),\ D_2(\lambda) = a_1(\lambda)a_2(\lambda), \ldots,\ D_s(\lambda) = a_1(\lambda)\ldots a_s(\lambda)$; cada polinômio, à partir do segundo, é divisível pelo seu precedente. Denomina-se POLINÔMIO INVARIANTE da matriz $A(\lambda)$, a cada um dos s-polinômios unitários, dados por :-

(11) $$i_k(\lambda) = \frac{D_k(\lambda)}{D_{k-1}(\lambda)}, \quad k = 1,2, \ldots ,s .$$

De (1o) e (11) é fácil de se concluir que :-

(12) $$i_k(\lambda) = a_k(\lambda), \quad k = 1,2, \ldots ,s.$$

(13) $i_h(\lambda)$ divide $i_{h+1}(\lambda)$, $h = 1,2, \ldots , s-1$.

Com base em (12), (13) e no teorema 2 de 3.2, obtemos o seguinte teorema :-

TEOREMA 3.

" Tôda λ-matriz $m \times p$, $A(\lambda)$ de característica s e grau n (finito), é equivalente à uma única matriz diagonal da forma :-

(14) $$\begin{bmatrix} i_1(\lambda) & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & i_2(\lambda) & \ldots & 0 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \ldots & i_s(\lambda) & 0 & \ldots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{bmatrix},$$

onde cada polinômio $i_h(\lambda)$ é unitário e $i_h(\lambda)$ divide $i_{h+1}(\lambda)$, $h=1,\ldots,\ldots, s-1$. Mais precisamente, $i_h(\lambda)$, $h = 1,2, \ldots ,s$ é um polinômio invariante de $A(\lambda)$".

COROLÁRIO 2.

" Duas matrizes polinomiais $m \times p$, $A(\lambda)$ e $B(\lambda)$, são equivalentes se , e sòmente se, elas têm os mesmos polinômios invariantes ou equivalentemente os mesmos $D_k(\lambda)$".

Demonstração.

Do lema 2, sendo $A(\lambda)$ e $B(\lambda)$ equivalentes, elas têm os mesmos $D_k(\lambda)$. Então, de (11) vem que $A(\lambda)$ e $B(\lambda)$ possuem os mesmos polinômios invariantes. Reciprocamente, suponhamos que $A(\lambda)$ e $B(\lambda)$ possuem a mesma sequência de polinômios invariantes (a ordem de tais polinômios invariantes é dada por (15)). Então, $A(\lambda)$ e $B(\lambda)$ são equivalentes à mesma forma diagonal (14) do teorema 3. Logo, $A(\lambda)$ e $B(\lambda)$ são equivalentes.

c.q.d.

Consequentemente, os $D_k(\lambda)$ ou $i_k(\lambda)$, formam um <u>SISTEMA COMPLETO DE INVARIANTES DA λ-MATRIZ</u> $A(\lambda)$, no sentido de que se os invariantes em questão, são os mesmos, para duas λ-matrizes, então elas são equivalentes.

Decompomos cada polinômio invariante $i_k(\lambda)$, $k = 1, 2, \ldots, s$, de $A(\lambda)$, em fatores irredutíveis de $K[\lambda]$, isto é, seja :-

$$(16) \quad \begin{cases} i_1(\lambda) = [p_1(\lambda)]^{a_1} \quad [p_2(\lambda)]^{a_2} \quad \ldots \quad [p_r(\lambda)]^{a_r}, \\ i_2(\lambda) = [p_1(\lambda)]^{b_1} \quad [p_2(\lambda)]^{b_2} \quad \ldots \quad [p_r(\lambda)]^{b_r}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ i_s(\lambda) = [p_1(\lambda)]^{g_1} \quad [p_2(\lambda)]^{g_2} \quad \ldots \quad [p_r(\lambda)]^{g_r}, \end{cases}$$

onde $0 \leq a_k \leq b_k \leq \ldots \leq g_k$, $k = 1, 2, \ldots, r$ e $p_1(\lambda), p_2(\lambda), \ldots, p_r(\lambda)$ são polinômios irredutíveis de $K[\lambda]$, dois à dois distintos e unitários. Denomina-se <u>DIVISOR ELEMENTAR</u> da λ-matriz $A(\lambda)$ no corpo K, a cada polinômio, não constante,

$$(17) \quad [p_1(\lambda)]^{a_1}, \ldots, [p_r(\lambda)]^{a_r}, [p_1(\lambda)]^{b_1}, \ldots, [p_r(\lambda)]^{b_r}, \ldots\ldots, [p_1(\lambda)]^{g_1}, \ldots, [p_r(\lambda)]^{g_r},$$

que aparecem em (16). Observa-se que se $a_1 = b_1$, por exemplo, $[p_1(\lambda)]^{a_1}$ aparece <u>duas</u> vêzes (em (17)) como divisor elementar de $A(\lambda)$.

De (16) vemos que os polinômios invariantes de $A(\lambda)$, determinam ùnivocamente todos os divisores elementares de $A(\lambda)$ em $K[\lambda]$. Reciprocamente, os divisores elementares (17), em qualquer ordem, juntamente com a característica s da λ-matriz $A(\lambda)$, determinam ùnivamente os polinômios invariantes de $A(\lambda)$. De fato, de (13), conclui-se que $i_s(\lambda)$ é o produto das mais altas potências de $p_1(\lambda)$, ..., $p_r(\lambda)$, que aparecem em (17). Analogamente, $i_{s-1}(\lambda)$ é definido no subconjunto restante de divisores elementares, e assim por diante, obtemos todos os polinômios invariantes. Por exemplo, se os divisores elementares de $A(\lambda)$, de característica 4, são :-

$$\lambda,\ \lambda^2,\ \lambda^2,\ (\lambda+1)^2,\ (\lambda+1)^4,$$

então os polinômios invariantes de $A(\lambda)$ são :-

$$i_4(\lambda) = \lambda^2(\lambda+1)^4,\ i_3(\lambda) = \lambda^2(\lambda+1)^2,\ i_2(\lambda) = \lambda \ \text{ e } \ i_1(\lambda) = 1.$$

Demonstramos acima o seguinte teorema :-

TEOREMA 4.

" Uma condição necessária e suficiente, para que duas λ-matrizes $A(\lambda)$ e $B(\lambda)$, tenham os mesmos polinômios invariantes, é que elas tenham a mesma característica e os mesmos divisores elementares".

TEOREMA 5.

" Suponhamos que a λ-matriz $A(\lambda)$ seja equivalente a uma λ-matriz $B(\lambda)$ da forma :-

$$B(\lambda) = \begin{bmatrix} f_1(\lambda) & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & f_2(\lambda) & \cdots & 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & f_s(\lambda) & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix},$$

onde cada $f_i(\lambda)$ é um polinômio unitário. Então, os divisores elementares de $A(\lambda)$, são as mais altas potências dos fatores irredutíveis de cada $f_i(\lambda)$, $i = 1,2, \ldots ,s$".

Demonstração.

Seja $p(\lambda)$ um fator irredutível, de algum $f_i(\lambda)$, $i = 1,2,\ldots,s$. Ordenemos os $f_i(\lambda)$, de acôrdo com as potências não negativas de $p(\lambda)$, isto é, sejam :-

$$f_{h_1}(\lambda) = g_1(\lambda)\,[p(\lambda)]^{a_1},\; f_{h_2}(\lambda) = g_2(\lambda)\,[p(\lambda)]^{a_2},\; \ldots, f_{h_s}(\lambda) = g_s(\lambda)[p(\lambda)]^{a_s};$$

onde $o \leq a_1 \leq a_2 \leq \ldots \leq a_s$ e cada $g_i(\lambda)$ é primo com $p(\lambda)$. Através de convenientes trocas de linhas e colunas, pode-se passar $B(\lambda)$, a uma λ-matriz $B'(\lambda)$, que lhe é equivalente (logo também é equivalente à $A(\lambda)$), tal que nas posições (j,j), $j = 1,2,\ldots,s$, apareçam $f_{h_j}(\lambda)$, e nas posições restantes zeros. É fácil de se ver que cada $D_k(\lambda)$, $k = 1,2, \ldots ,s$ de $B'(\lambda)$ (logo de $A(\lambda)$), são tais que o expoente da mais alta potência de $p(\lambda)$, que divide $D_k(\lambda)$, é $a_1 + a_2 + \ldots + a_k$. De (11), $i_k(\lambda) = D_k(\lambda) / D_{k-1}(\lambda)$, logo a maior potência de $p(\lambda)$ que divide o polinômio invariante $i_k(\lambda)$ de $A(\lambda)$ é :-

$$[p(\lambda)]^{a_k}, \qquad k = 1,2, \ldots ,s.$$

Então, para cada $k = 1,2, \ldots ,s$, tal que $a_k > o$, $[p(\lambda)]^{a_k}$ é um divisor elementar de $A(\lambda)$. Uma repetição dêstes argumentos, para cada fator irredutível de $K[\lambda]$, que divide algum $f_i(\lambda)$, $i = 1,2,\ldots,s$, mostra-nos que cada potência do mais alto grau, de cada fator irredutível de $f_i(\lambda)$, é um divisor elementar de $A(\lambda)$. Ainda mais, todos os divisores elementares de $A(\lambda)$, são assim obtidos. Isto demonstra o teorema.

<u>COROLÁRIO 3.</u>

" Seja a λ-matriz,

$$A(\lambda) = B(\lambda) \oplus C(\lambda) = \left[\begin{array}{c|c} B(\lambda) & 0 \\ \hline 0 & C(\lambda) \end{array}\right].$$

Então a coleção dos divisores elementares de $A(\lambda)$ é igual a coleção dos divisores elementares de $B(\lambda)$ mais os de $C(\lambda)$".

Demonstração.

Pelo teorema 3, as matrizes $B(\lambda)$ e $C(\lambda)$, são equivalentes a λ-matrizes $D'(\lambda)$ e $D''(\lambda)$, respectivamente, onde :-

$$
D'(\lambda) = \begin{bmatrix}
i'_1(\lambda) & 0 & \dots & 0 & \dots & 0 \\
0 & i'_2(\lambda) & \dots & 0 & \dots & 0 \\
\vdots & \vdots & \dots & \vdots & \dots & \vdots \\
0 & 0 & \dots & i'_t(\lambda) & \dots & 0 \\
0 & 0 & \dots & 0 & \dots & 0 \\
\vdots & \vdots & \dots & \vdots & \dots & \vdots \\
0 & 0 & \dots & 0 & \dots & 0
\end{bmatrix}
$$

e

$$
D''(\lambda) = \begin{bmatrix}
i''_1(\lambda) & 0 & \dots & 0 & \dots & 0 \\
0 & i''_2(\lambda) & \dots & 0 & \dots & 0 \\
\vdots & \vdots & \dots & \vdots & \dots & \vdots \\
0 & 0 & \dots & i''_v(\lambda) & \dots & 0 \\
0 & 0 & \dots & 0 & \dots & 0 \\
\vdots & \vdots & \dots & \vdots & \dots & \vdots \\
0 & 0 & \dots & 0 & \dots & 0
\end{bmatrix}
$$

Como $A(\lambda) = B(\lambda) \oplus C(\lambda)$, teremos que $A(\lambda)$ é equivalente à λ-matriz $D(\lambda) = D'(\lambda) \oplus D''(\lambda)$. Do teorema 5, concluimos nossa demonstração, pois os divisores elementares de $A(\lambda)$ são as maiores potências dos fatores irresutíveis dos polinômios :-

$i'_1(\lambda), i'_2(\lambda), \dots, i'_t(\lambda)$ (que são os divisores elementares de $B(\lambda)$),

$i''_1(\lambda), \dots, i''_v(\lambda)$ (que são os divisores elementares de $C(\lambda)$).

c.q.d.

Desde que os polinômios invariantes de $B(\lambda)$, não dividem necessàriamente os de $C(\lambda)$, não vale em geral, o resultado do corolário 3, para polinômios invariantes (no lugar dos divisores elementares). Com a condição adicional deque todos os polinômios invariantes de $B(\lambda)$, divide cada polinômio invariante de $C(\lambda)$, temos o seguinte corolário:-

<u>COROLÁRIO 4.</u>

" Seja a λ-matriz $A(\lambda) = B(\lambda) \oplus C(\lambda)$. Suponhamos ainda, que todo polinômio invariante de $B(\lambda)$ divide cada polinômio invariante de $C(\lambda)$. Então, o conjunto dos polinômios invariantes de $A(\lambda)$ é a união (nesta ordem), dos polinômios invariantes de $B(\lambda)$ com os de $C(\lambda)$".

<u>EXEMPLO 3.</u>

Suponhamos que uma λ-matriz $A(\lambda)$, seja reduzida através de operações elementares à forma :-

$$D(\lambda) = \begin{bmatrix} \lambda(\lambda+1)^2 & 0 & 0 & 0 & 0 \\ 0 & \lambda^2(\lambda+1) & 0 & 0 & 0 \\ 0 & 0 & \lambda^3(\lambda^2+1) & 0 & 0 \\ 0 & 0 & 0 & \lambda^2+1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Pelo teorema 5, os divisores elementares de $A(\lambda)$ são :-

$$\lambda,\ \lambda^2,\ \lambda^3,\ (\lambda+1),\ (\lambda+1)^2\ \ (\lambda^2+1) \text{ e } (\lambda^2+1),$$

no caso em que $K = R$ (corpo dos números reais).

Sendo 4, a característica de $A(\lambda)$, seus polinômios invariantes são :-

$i_4(\lambda) = \lambda^3(\lambda+1)^2(\lambda^2+1)$, $i_3(\lambda) = \lambda^2(\lambda+1)(\lambda^2+1)$, $i_2(\lambda) = \lambda$ e $i_1(\lambda) = 1$.

EXEMPLO 4.

Consideremos a λ-matriz $A(\lambda)$ do exemplo 1. Nêste exemplo 1, vimos que $A(\lambda)$ é equivalente à matriz diagonal :-

$$\left[1,\ 1,\ \lambda+1,\ \lambda^3+1\right].$$

Assim no corpo R, os divisores elementares de $A(\lambda)$ são :-

$$\lambda+1,\ \ \lambda+1,\ \ \lambda^2-\lambda+1.$$

No corpo C (dos números complexos), os divisores elementares de $A(\lambda)$ são :-

$$\lambda+1,\ \lambda+1,\ \lambda-(1-\sqrt{3}\,i)/2 \text{ e } \lambda-(1+\sqrt{3}\,i)/2.$$

Isto nos mostra, que os divisores elementares de $A(\lambda)$, dependem do corpo em questão. Vemos que os polinômios invariantes de $A(\lambda)$ (em $R[\lambda]$ ou $C[\lambda]$), são :-

$i_4(\lambda) = (\lambda+1)(\lambda^2-\lambda+1) = \lambda^3+1$, $i_3(\lambda) = \lambda+1$, $i_2(\lambda) = i_1(\lambda) = 1$.

Suponhamos agora que $A = \left[a_{ik}\right]$, é uma matriz quadrada de ordem n com elementos no corpo K, isto é, $A \in M_n(K)$. Formemos sua MATRIZ CARACTERÍSTICA $A(\lambda)$, que é λ-matriz, $n \times n$ de grau 1,

$$A(\lambda) = \lambda E - A = \begin{bmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{bmatrix}.$$

Os polinômios invariantes da matriz característica $A(\lambda)$ de A, são denominados <u>POLINÔMIOS INVARIANTES DA MATRIZ A</u>, e os divisores elementares de $A(\lambda)$ em $K[\lambda]$, denominam-se <u>DIVISORES ELEMENTARES DA MATRIZ A NO CORPO K.</u>

<u>EXEMPLO 5.</u>

Vamos calcular os divisores elementares e os polinômios invariantes, da matriz A, abaixo no corpo R,

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$A(\lambda) = \lambda E - A = \begin{bmatrix} \lambda-1 & -1 & 0 \\ 0 & \lambda-1 & 0 \\ 0 & 0 & \lambda-2 \end{bmatrix} \overset{[2(\lambda-1)+1]}{\sim} \begin{bmatrix} 0 & -1 & 0 \\ (\lambda-1)^2 & \lambda-1 & 0 \\ 0 & 0 & \lambda-2 \end{bmatrix}$$

$$\overset{(1(\lambda-1)+2)}{\sim} \begin{bmatrix} 0 & -1 & 0 \\ (\lambda-1)^2 & 0 & 0 \\ 0 & 0 & \lambda-2 \end{bmatrix} \overset{[1,2]}{\sim} \begin{bmatrix} -1 & 0 & 0 \\ 0 & (\lambda-1)^2 & 0 \\ 0 & 0 & \lambda-2 \end{bmatrix}.$$

Pelo teorema 5., os divisores elementares de $A(\lambda)$, logo de A, em $R[\lambda]$, são :-

$$(\lambda - 1)^2 \quad e \quad (\lambda - 2).$$

Assim os polinômios invariantes são :-

$$i_3(\lambda) = (\lambda - 1)^2(\lambda - 2), \; i_2(\lambda) = i_1(\lambda) = 1.$$

3.4 UM CRITÉRIO PARA SEMELHANÇA DE MATRIZES.

Nêste parágrafo, estudaremos λ-matrizes da forma $\lambda E - A$, isto é, matizes características de uma matriz constante A, $n \times n$.

Denomina-se matriz <u>BINOMIAL REGULAR</u>, à uma λ-matriz da forma

$A_0 + \lambda A_1$, onde A_0 e A_1 são matrizes constantes $n \times n$, e $\det A_1 \neq 0$. Tôda matriz binomial regular, é equivalente à uma λ-matriz da forma $\lambda E - A$. De fato,

$$A_0 + \lambda A_1 = A_1 (\lambda E + A_1^{-1} A_0),$$

e tomando $A = -A_1^{-1} A_0$, vem

$$A_0 + \lambda A_1 = A_1 (\lambda E - A).$$

A seguir, demonstraremos um lema, que é uma generalização do teorema de Bezout.

<u>LEMA 3.</u> (<u>TEOREMA DE BEZOUT GENERALIZADO</u>)

" Tôda λ-matriz :- $P(\lambda) = P_0 \lambda^n + P_1 \lambda^{n-1} + \dots + P_n$, pode ser dividida à esquerda, por uma matriz da forma $\lambda E - A$ ($A \in M_n(K)$), isto é, existem matrizes $S(\lambda)$ e R (R constante), tais que :-

$$P(\lambda) = (\lambda E - A) S(\lambda) + R \qquad e \qquad R = P(A) \text{ ".}$$

Demonstração.

Vê-se, fàcilmente, que a λ-matriz $P(\lambda) - (\lambda E - A) P_0 \lambda^{n-1}$, tem grau menor ou igual à n-1, pois :-

$$(P_0 \lambda^n + P_1 \lambda^{n-1} + \dots + P_n) - P_0 \lambda^n + AP_0 \lambda^{n-1} = (P_1 + AP_0) \lambda^{n-1} + P_2 \lambda^{n-2} + \dots$$

$$\dots + P_n = P_0' \lambda^{n-1} + P_1' \lambda^{n-2} + \dots + P_{n-1}' .$$

Anàlogamente, a λ-matriz $P(\lambda) - (\lambda E - A) P_0 \lambda^{n-1} - (\lambda E - A) P_0' \lambda^{n-2}$, tem grau menor ou igual à n-2. Continuando êste processo, ou mais precisamente por indução, obtemos uma λ-matriz :-

$$R = P(\lambda) - (\lambda E - A) P_0 \lambda^{n-1} - (\lambda E - A) P_0' \lambda^{n-2} - \dots - (\lambda E - A) P_0^r \lambda^{n-r+1},$$

de grau zero, isto é, constante. Logo,

$$(18) \qquad P(\lambda) = (\lambda E - A) S(\lambda) + R,$$

onde,

$$S(\lambda) = P_0 \lambda^{n-1} + P_0' \lambda^{n-2} + \dots + P_0^r \lambda^{n-r+1} .$$

De (18), temos que :-

$$P(A) = (AE - A) S(A) + R = R.$$

c.q.d.

Vale, òbviamente, um lema análogo, para divisão à direita, isto é, existem matrizes $S_1(\lambda)$ e R_1 (constante), tais que :-

$$P(\lambda) = S_1(\lambda)(\lambda E - A) + R_1 \quad \text{e} \quad R_1 = P(A).$$

Uma λ-matriz $P(\lambda)$, quadrada de ordem n e grau r (finito), diz-se INVERSÍVEL se, e sòmente se, existe uma λ-matriz $M(\lambda)$ (INVERSA DE $P(\lambda)$), tal que :-

$$P(\lambda)\, M(\lambda) = M(\lambda)\, P(\lambda) = E .$$

Uma definição alternativa para que $P(\lambda)$ seja inversível é que det $P(\lambda)$ = cte $\neq$ o. De fato, os elementos da matriz inversa são, à menos do sinal, subdeterminantes de ordem n-1, divididos pelo det $P(\lambda)$ que é constante e não nulo. Logo os elementos de $\left[P(\lambda)\right]^{-1}$ são polinômios. Reciprocamente, se $P(\lambda)$ é inversível, então det $P(\lambda)$ $\cdot$ det$\left[P(\lambda)\right]^{-1}$ = det E = 1. Sendo o produto entre dois polinômios iguais à 1, êstes polinômios são constantes e não nulos, logo det $P(\lambda)$ =c$\neq$ o.

Nosso próximo lema, além de estabelecer um critério para a equivalência das λ-matrizes $A(\lambda) = \lambda E - A$ e $B(\lambda) = \lambda E - B$, nos permitirá estabelecer um critério para semelhança das matrizes constantes A e B.

LEMA 4.

" Sejam $A(\lambda) = \lambda E - A$ e $B(\lambda) = \lambda E - B$, matrizes $n \times n$, binomiais regulares. Uma condição necessária e suficiente para $A(\lambda)$ e $B(\lambda)$ sejam equivalentes, é que existam matrizes constantes P e Q, $n \times n$ e não-singulares (inversíveis), tais que :-

(18) $\quad B(\lambda) = P\, A(\lambda)\, Q$".

Demonstração.

Como uma matriz $n \times n$, constante e não-singular, é um caso particular de uma λ-matriz inversível, é claro que (18) nos diz que $A(\lambda)$ e $B(\lambda)$ são equivalentes.

Sejam então $A(\lambda)$ e $B(\lambda)$ equivalentes. Então, existem λ-matrizes $P(\lambda)$ e $Q(\lambda)$, inversíveis, tais que :-

$$B(\lambda) = P(\lambda)\, A(\lambda)\, Q(\lambda),$$

ou se $\left[P(\lambda)\right]^{-1} = M(\lambda)$, temos que :-

(19) $\quad M(\lambda)\, B(\lambda) = A(\lambda)\, Q(\lambda)$ ou $M(\lambda)(\lambda E - B) = (\lambda E - A)\, Q(\lambda)$.

Com base no lema 3, dividimos $M(\lambda)$ à esquerda por $A(\lambda) = \lambda E - A$ e $Q(\lambda)$ à direita por $B(\lambda) = \lambda E - B$, obtemos :-

$$(20)\qquad \begin{cases} M(\lambda) = (\lambda E - A)\, S(\lambda) + M \\ Q(\lambda) = T(\lambda)\, (\lambda E - B) + Q\,, \end{cases}$$

onde $M = M(A)$ e $Q = Q(B)$ são matrizes constantes $n \times n$. De (19) e (20) vem que :-

$$\left[(\lambda E - A)\, S(\lambda) + M \right] (\lambda E - B) = (\lambda E - A) \left[T(\lambda)(\lambda E - B) + Q \right],$$

ou fazendo-se alguns cálculos :-

$$(21)\quad M(\lambda E - B) - (\lambda E - A)\, Q = (\lambda E - A) \left[T(\lambda) - S(\lambda) \right] (\lambda E - B).$$

De (21), temos que $T(\lambda) = S(\lambda)$, caso contrário, o grau da λ-matriz do segundo membro de (21) seria maior ou igual à 2, e do primeiro membro seria menor ou igual à 1. Então, $T(\lambda) = S(\lambda)$ e de (21), vem que :-

$$(22)\quad M(\lambda E - B) = (\lambda E - A) Q \quad \text{ou} \quad MB(\lambda) = A(\lambda)\, Q.$$

Resta demonstrar que M e Q são não-singulares. Pois, feito isto, toma-se $P = M^{-1}$ e conclui-se a demonstração do lema. Novamente com base na lema 3, dividimos $P(\lambda)$ à esquerda por $B(\lambda) = \lambda E - B$:-

$$(23)\quad P(\lambda) = (\lambda E - B) R(\lambda) + P \quad \text{e} \quad P(B) = P.$$

De (19), primeira equação de (20), (23) e sendo $M(\lambda) = \left[P(\lambda) \right]^{-1}$, vem que :-

$$E = M(\lambda) P(\lambda) = M(\lambda)\, (\lambda E - B)\, R(\lambda) + M(\lambda) P = (\lambda E - A)\, Q(\lambda)\, R(\lambda) +$$
$$+ (\lambda E - A)\, S(\lambda)\, P + MP\,, \quad \text{ou}$$

$$(24)\quad (\lambda E - A) \left[Q(\lambda)\, R(\lambda) + S(\lambda)\, P \right] + MP = E\,.$$

Sendo E uma λ-matriz de grau zero, de (24) concluimos que :-

$$Q(\lambda)\, R(\lambda) + S(\lambda)\, P = 0.$$

Assim sendo, (24) se reduz à : $MP = E$. Como $\det E = 1$, $\det M \neq 0$ e $P = M^{-1}$. Então, de (22), obtemos :-

$$(25)\quad B(\lambda) = P\, A(\lambda)\, Q \qquad \text{ou} \qquad \lambda E - B = P(\lambda E - A)\, Q.$$

De (25), têm-se :-

$$\lambda E - B = \lambda PQ - PAQ \quad \text{ou} \quad E = PQ,$$

logo $\det Q \neq 0$ e ainda $Q^{-1} = P$. c.q.d.

Observamos que de (25), segue ainda que $B = PAQ$, e como $P = Q^{-1}$, temos $B = Q^{-1}AQ$, isto é, A e B são semelhantes. Reciprocamente, se A e B são semelhantes, existe uma matriz inversível C, tal que :-

$$B = C^{-1} A C.$$

Assim, $\lambda E - B = \lambda E - C^{-1}AC = C^{-1} EC - C^{-1}AC = C^{-1}(\lambda E - A) C$, isto é, as matrizes características $A(\lambda) = \lambda E - A$ e $B(\lambda) = \lambda E - B$ são equivalentes, logo A e B possuem os mesmos polinômios invariantes e os mesmos divisores elementares(a característica de $A(\lambda)$ e $B(\lambda)$ é n). Demonstramos assim, o seguinte teorema :-

TEOREMA 6. (CRITÉRIO PARA SEMELHANÇA DE MATRIZES).

"Duas matrizes A e B de $M_n(K)$, são semelhantes se, e sòmente se, suas matrizes características $A(\lambda) = \lambda E - A$ e $B(\lambda) = \lambda E - B$, são equivalentes, ou o que é o mesmo, se elas têm os mesmos polinômios invariantes ou ainda os mesmos divisores elementares em $K[\lambda]$" .

Nas demonstrações do lema 4 e teorema 6, ficou ainda estabelecido o seguinte corolário, que usaremos no capítulo IV.

COROLÁRIO 5.

" Sejam A e B matrizes de $M_n(K)$, semelhantes, com $B = C^{-1}AC$. Então, pode-se tomar a matriz C, como sendo :-

(26) $C = Q(B) = \left[\hat{P}(B)\right]^{-1}$,

onde $P(\lambda)$ e $Q(\lambda)$ são λ-matrizes na indeterminada :-

$$\lambda E - B = P(\lambda)\ (\lambda E - A)\ Q(\lambda),$$

que nos dá a equivalência das matrizes características $A(\lambda) = \lambda E - A$ e $B(\lambda) = \lambda E - B$. Em (26), $Q(B)$ denota o valor à direita na substituição da indeterminada λ em $Q(\lambda)$, pela matriz B e $\hat{P}(B)$ denota o valor à esquerda da matriz polinomial $P(\lambda)$".

CAPÍTULO IV - AS FORMAS RACIONAL E DE JORDAN.

Êste é o capítulo central desta dissertação, visto que é nêle que descrevemos reduções às formas canônicas Racional e de Jordan.

4.1 REDUÇÃO ÀS FORMAS RACIONAL E DE JORDAN, ATRAVÉS DOS DIVISORES ELEMENTARES.

Dado um polinômio :

$$f(X) = X^n + a_1X^{n-1} + a_2X^{n-2} + \ldots + a_{n-1}X + a_n,$$

deomina-se MATRIZ COMPANHEIRA ou ASSOCIADA ao polinômio f, a matriz :

$$C_f = \begin{bmatrix} 0 & 0 & \ldots & 0 & -a_n \\ 1 & 0 & \ldots & 0 & -a_{n-1} \\ 0 & 1 & \ldots & 0 & -a_{n-2} \\ \cdot & \cdot & \ldots & \cdot & \cdot \\ \cdot & \cdot & \ldots & \cdot & \cdot \\ 0 & 0 & \ldots & 1 & -a_1 \end{bmatrix}.$$

Observa-se que a matriz companheira de $X + a_1$ é $-a_1$. Mostraremos a seguir, que o polinômio característico de C_f é f. De fato,

$$c(X) = \det(XE - C_f) = \begin{bmatrix} X & 0 & \ldots & 0 & a_n \\ -1 & X & \ldots & 0 & a_{n-1} \\ 0 & -1 & \ldots & 0 & a_{n-2} \\ \cdot & \cdot & \ldots & \cdot & \cdot \\ \cdot & \cdot & \ldots & \cdot & \cdot \\ 0 & 0 & \ldots & -1 & X + a_1 \end{bmatrix}$$

Multiplicamos as linhas 2, 3, ..., n, dêste determinante, por $X, X^2, \ldots, X^{n-1}$ respectivamente, e adicionamos à primeira linha. Assim, todos os elementos da primeira linha, tornam-se nulos, exceto o último que é $f(X)$. A seguir, desenvolve-se por Laplace segundo a primeira linha, temos:-

$$c(X) = (-1)^{n+1} f(X) \begin{bmatrix} -1 & X & \ldots & 0 & 0 \\ 0 & -1 & \ldots & 0 & 0 \\ \cdot & \cdot & \ldots & \cdot & \cdot \\ \cdot & \cdot & \ldots & \cdot & \cdot \\ 0 & 0 & \ldots & -1 & X \\ 0 & 0 & \ldots & 0 & -1 \end{bmatrix} = (-1)^{n+1} f(X)\,(-1)^{n-1} = f(X).$$

Concluimos ainda que o complemento algébrico (subdeterminante de ordem n-1) do elemento a_n, na matriz característica $A(X) = XE-C_f$ é ± 1, logo $D_{n-1}(X) = 1$, e de (1o) de 3.3, temos que $D_k(X) = 1$, para $k = 1,2, \ldots, n-1$, pois $D_k(X)$ divide $D_{n-1}(X) = 1$. Ainda temos que :-

$$i_n(X) = \frac{D_n(X)}{D_{n-1}(X)} = \frac{f(X)}{1} = f(X),\ i_{n-1}(X) = \ldots = i_2(X) = i_1(X)=1.$$

Consequentemente, C_f tem apenas um polinômio invariante (f) diferente de 1.

Consideremos uma matriz $A = \left[a_{ij}\right]$ de $M_n(K)$. Sejam :-

$$i_1 = i_2 = \ldots = i_r = 1,\ i_{r+1}, \ldots, i_n,$$

os polinômios invariantes de A. Estamos supondo que os polinômios invariantes :

(1) $\qquad i_{r+1}, i_{r+2}, \ldots, i_n,$

tem grau maior ou igual a 1. Sabemos que cada um dos polinômios de (1), à partir do segundo, é divisível pelo seu precedente.

Sejam $C_{i_{r+1}}, \ldots, C_{i_n}$, as matrizes companheiras dos polinômios de (1). Seja a matriz R_1, dada por :-

$$R_1 = C_{i_{r+1}} \oplus C_{i_{r+2}} \oplus \ldots \oplus C_{i_n}.$$

Pela extensão óbvia do corolário 4 de 3.3, os polinômios invariantes de R_1 são a união dos polinômios invariantes de $C_{i_{r+1}}, C_{i_{r+2}}, \ldots, C_{i_n}$. Os polinômios invariantes destas matrizes, distintos de 1, são seus polinômios carasterísticos (o que já verificamos acima) :

$$i_{r+1},\ i_{r+2}, \ldots, i_n.$$

Logo, os polinômios invariantes das matrizes A e R_1 são os mesmos .Pelo critério de semelhança (teorema 6 de 3.4), A e R_1 são semelhantes. Isto é, existe uma matriz inversível C de $M_n(K)$, tal que :

$$R_1 = C^{-1} A\ C.$$

A matriz R_1 é denominada <u>PRIMEIRA FORMA CANÔNICA NATURAL</u> da matriz A, ou <u>PRIMEIRA FORMA RACIONAL</u> da matriz A. Observa-se que essa forma canônica, é caracterisada pelos seguintes fatos:-

1) R_1 é a soma direta de matrizes companheiras $C_{i_{r+1}}, C_{i_{r+2}}, \ldots, C_{i_n}$, de polinômios unitários, não constantes, $i_{r+1}, i_{r+2}, \ldots, i_n$.

2) i_h divide i_{h+1}, $h = r+1, r+2, \ldots, n-1$.

3) i_h, $h = r+1, \ldots, n$, é o polinômio caraterístico da respectiva matriz C_{i_h}, na soma direta de R_1.

Consideremos agora os divisores elementares da matriz $A = [a_{ij}]$ de $M_n(K)$;

(2) $p_1, p_2, \dots, p_s$.

Sejam $C_{p_1}, C_{p_2}, \dots, C_{p_s}$, as matrizes companheiras dos polinômios (2) e

$$R_2 = C_{p_1} \oplus C_{p_2} \oplus \dots \oplus C_{p_s} .$$

Pela extensão óbvia do corolário 4 de 3.3, os divisores elementares de R_2 é a união (nessa ordem) dos divisores elementares de $C_{p_1}, \dots, C_{p_s}$. Como p_j é o único polinômio invariante de C_{p_j}, não constante, e sendo um divisor elementar de A, é uma potência de um polinômio irredutível e unitário de $K[X]$. Assim, p_j é o único divisor elementar de C_{p_j}, $j = 1,2, \dots, s$.

Então, as matrizes A e R_2 possuem os mesmos divisores elementares sôbre K. Logo, pelo teorema 6 de 3.4, A e R_2 são semelhantes, isto é, existe uma matriz inversível P de $M_n(K)$, tal que :-

$$R_2 = P^{-1} A P.$$

A matriz R_2, denomina-se <u>SEGUNDA FORMA CANÔNICA NATURAL</u> da matriz A, ou <u>SEGUNDA FORMA RACIONAL</u> da matriz A. Ela é caracterizada por:- R_2 é a soma direta das matrizes companheiras $C_{p_1}, C_{p_2}, \dots, C_{p_s}$, de polinômios unitários que são potências de um polinômio irredutível sôbre $K[X]$, não constantes. Êstes polinômios são exatamente, os divisores elementares de A.

As transpostas das matrizes R_1 e R_2, aparecem na literatura como a primeira e sgunda forma racional de A, respectivamente. Isto vem do fato de se considerar a transposta de C_f, como a matriz companheira de f. Esta forma canônica, foi introduzida por G. Kowalewski (Leipziger Berichte, 1 917, p. 325), que denominou <u>FORMA NORMAL NATURAL</u>.

Para melhor caracterizar a forma Racional, digamos a segunda, suponhamos que os polinômios de (2), sejam :-

$$\begin{cases} p_1(X) = X^{n_1} + a_{11} X^{n_1-1} + a_{12}X^{n_1-2} + \dots + a_{1n_1} \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ p_s(X) = X^{n_s} + a_{s1} X^{n_s-1} + a_{s2} X^{n_s-2} + \dots + a_{sn_s} . \end{cases}$$

Então, temos que:-

$$(3)\quad R_2 = \begin{bmatrix} \begin{array}{ccccc} 0 & 0 & \cdots & 0 & -a_{1n_1} \\ 1 & 0 & \cdots & 0 & -a_{1n_1-1} \\ 0 & 1 & \cdots & 0 & -a_{1n_1-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{11} \end{array} & & \bigcirc \\ & \ddots & \\ \bigcirc & & \begin{array}{ccccc} 0 & 0 & \cdots & 0 & -a_{sn_s} \\ 1 & 0 & \cdots & 0 & -a_{sn_s-1} \\ 0 & 1 & \cdots & 0 & -a_{sn_s-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{s1} \end{array} \end{bmatrix}$$

Os elementos não escritos são nulos. É claro que R_2, não é univocamente determinada, já que a ordem dos divisores elementares em (2), não foi especificada. Ordenando-se os divisores elementares através de seus graus digamos que :-

$$n_1 \geqslant n_2 \geqslant \dots \geqslant n_s,$$

no caso de igualdade, fixa-se qual polinômio vem primeiro nesta ordenação. Então R_2 fica ùnivocamente determinada. Vimos no exemplo 4 de 3.3, que os divisores elementares dependem do corpo K. Assim sendo a segunda forma Racional depende do K. Poderíamos fazer as mesmas considerações, para a primeira forma Raciona, que já é única dada a divisibilidade entre os polinômios invariantes. Isto justifica o fato de serem a primeira e segunda formas Racional, canônicas relativamente à semelhança.

A seguir obteremos a forma de Jordan. Agora necessitamos de impor uma condição adicional ao corpo K. Precisamos exigir que K contenha além dos elementos da matriz nxn $A = [a_{ij}]$, todos os seus autovalores. Isto sempre acontece, se K é algèbricamente fechado. Nêste caso, os divisores elementares de A, são da forma :-

$$(4)\qquad (X-\lambda_1)^{n_1},\ (X-\lambda_2)^{n_2},\ \dots,\ (X-\lambda_s)^{n_s},\ n_1+n_2+\dots+n_s = n,$$

podendo ocorrer $\lambda_i = \lambda_j$, mesmo para $i \neq j$.

Consideremos um dos divisores elementares de A, digamos,

$(X - \lambda)^r$.

Associamos à êle a seguinte matriz quadrada $r \times r$:-

$$(5)\quad J^r = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{bmatrix} \quad \left(\text{ou } J_r = \begin{bmatrix} \lambda & 0 & 0 & \cdots & 0 & 0 \\ 1 & \lambda & 0 & \cdots & 0 & 0 \\ 0 & 1 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \lambda \end{bmatrix} \right).$$

A matriz característica de J^r é a matriz $XE - J^r$, isto é,

$$(6)\quad \begin{bmatrix} X-\lambda & -1 & 0 & \cdots & 0 \\ 0 & X-\lambda & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \\ 0 & 0 & 0 & \cdots & X-\lambda \end{bmatrix} \quad \left(\text{ou } \begin{bmatrix} X-\lambda & 0 & \cdots & 0 & 0 \\ -1 & X-\lambda & \cdots & 0 & 0 \\ 0 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & X-\lambda \end{bmatrix} \right).$$

Desenvolvendo-se o determinante da matriz (6), pela primeira coluna (ou primeira linha), (r-1)-vêzes, vê-se que o polinômio característico de J^r (ou J_r) é

$$c(X) = (X - \lambda)^r.$$

Logo, $D_r(X) = c(X) = (X - \lambda)^r$.

Desde que o complemento algébrico do elemento da posição (r,1) é $(-1)^{r-1} = \pm 1$, temos que $D_{r-1}(X) = 1$. Consequentemente, o único divisor elementar de (5), é $(X - \lambda)^r$. A matriz J^r (ou J_r), dada por (5), denomina-se <u>BLOCO DE JORDAN</u> superior (ou inferior), associado ao divisor elementar $(X - \lambda)^r$.

Seja J a soma direta dos blocos de Jordan, associados aos divisores elementares (4) de A. Isto é,

$$(7)\quad J = J^{n_1} \oplus J^{n_2} \oplus \cdots \oplus J^{n_s} \quad (\text{ou } J = J_{n_1} \oplus J_{n_2} \oplus \cdots \oplus J_{n_s}).$$

Pela extensão óbvia do corolário 4 de 3.3, e de (7), vê-se que os divisores elementares de J, são justamente a união dos divisores elementares dos blocos de Jordan de (7). Assim sendo J e A possuem os mesmos divisores elementares, pois vimos acima, que o único divisor elementar do bloco de Jordan J^r (ou J_r) é $(X - \lambda)^r$. Pelo teorema 6 de 3.4, A e J são semelhantes, isto é, existe uma matriz inversível T em $M_n(K)$, tal que :-

$$(8)\qquad J = T^{-1} A\, T.$$

A matriz J, dada por (7), denomina-se FORMA CANÔNICA DE JORDAN superior (ou inferior) ou simplesmente FORMA DE JORDAN, da matriz A. A matriz T, dada por (8), denomina-se MATRIZ DE TRANSFORMAÇÃO de A à J.

A forma deJordan de uma matriz A, é caracterizada por :- J é a soma direta de blocos de Jordan J^{n_1} (ou J_{n_1}), ou seja, todo elemento de J que não esteja na diagonal ou imediatamente acima (abaixo) dela é nulo. Na diagonal de J, aparecem os s autovalores $\lambda_1, \lambda_2, \dots, \lambda_s$ de J. Desde que não se fixou a ordem dos divisores elementares (ou dos autovalores) em (4), a forma deJordan da matriz A, não esta univocamente determinada. Fixando-se tal ordem a forma de Jordan fica determinada, à menos da posição dos blocos de Jordan.

Para melhor caracterizar a forma de Jordan superior, a descrevemos abaixo, em relação aos divisores elementares (4), naquela ordem, lembrando que pode ocorrer $\lambda_i = \lambda_j$ mesmo para $i \neq j$.

$$J = \begin{bmatrix} \begin{matrix} \lambda_1 & 1 & o & \dots & o \\ o & \lambda_1 & 1 & \dots & o \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ o & o & o & \dots & 1 \\ o & o & o & \dots & \lambda_1 \end{matrix} & & & \\ & \begin{matrix} \lambda_2 & 1 & \dots & o \\ o & \lambda_2 & \dots & o \\ \vdots & \vdots & \ddots & \vdots \\ o & o & \dots & 1 \\ o & o & \dots & \lambda_2 \end{matrix} & & \\ & & \ddots & \\ & & & \begin{matrix} \lambda_s & 1 & \dots & o \\ o & \lambda_s & \dots & o \\ \vdots & \vdots & \ddots & \vdots \\ o & o & \dots & 1 \\ o & o & \dots & \lambda_s \end{matrix} \end{bmatrix}$$

Sempre é possível ordenarmos os divisores elementares de (4), pondo-se :-

$$(9)\begin{cases} (X-\lambda_1)^{n_{11}}, (X-\lambda_1)^{n_{12}}, \dots, (X-\lambda_1)^{n_{1k(1)}}, (X-\lambda_2)^{n_{21}}, \dots \\ (X-\lambda_2)^{n_{2k(2)}}, \dots \dots, (X-\lambda_r)^{n_{r1}}, \dots, (X-\lambda_r)^{n_{rk(r)}}, \\ \text{com } n_{i1} \geq n_{i2} \geq \dots \geq n_{ik(i)}, \; i = 1,2,\dots,r. \end{cases}$$

Assim, os elementos acima da diagonal de J (que são 1 ou o) ficam determinados de modo único pelos números n_{ij}, $i = 1,2, \ldots, r$ e $j = 1,2, \ldots, k(i)$. Consequentemente, temos uma única forma de Jordan associada à A, levando em conta (9), isto é, a ordem dos divisores elementares. Ademais, a cada autovalor λ_i, correspondem $k(i)$-blocos de Jordan de dimensões $n_{i1}, n_{i2}, \ldots, n_{ik(i)}$. Daí a afirmação de que para as matrizes de $M_n(K)$, com autovalores em K, a forma de Jordan é canônica relativamente à semelhança.

Os números n_{ij} de (9), são muitas vêzes escritos na forma :-

$$\{(n_{11}, n_{12}, \ldots, n_{1k(1)})(n_{21}, \ldots, n_{2k(2)}) \ldots (n_{r1}, \ldots, n_{rk(r)})\}$$

esta forma, que determina a forma de Jordan de A, é denominada <u>CARACTERÍSTICA DE SEGRE</u> de A.

Como uma consequência imediata da forma de Jordan, vem que :- " Uma matriz de $M_n(K)$ é diagonalizável se, e sòmente se, todos os seus divisores elementares são de grau 1 ".

Sabemos que matrizes semelhantes, representam o mesmo operador linear L sôbre V (K-espaço vetorial de dimensão finita n) (ver 1.1), relativamente à bases distintas de V. Assim, se $A = [L]_{\mathcal{B}}$, onde $\mathcal{B}$ é uma base ordenada de V, e se A é semelhante às matrizes R_1, R_2 ou J, existem bases ordenadas de V, $\mathcal{B}_1$, $\mathcal{B}_2$ e $\mathcal{B}_\gamma$, respectivamente, tal que :-

$$[L]_{\mathcal{B}_1} = R_1, \quad [L]_{\mathcal{B}_2} = R_2 \quad e \quad [L]_{\mathcal{B}_\gamma} = J.$$

Levando-se em conta que :

$$R_1 = C^{-1} A C, \quad R_2 = P^{-1} A P \quad e \quad J = T^{-1} A T,$$

vem que as bases ordenadas $\mathcal{B}_1$, $\mathcal{B}_2$ e $\mathcal{B}_\gamma$ de V, são tais que, C, P e T são as matrizes mudança da base ordenada $\mathcal{B}$ para as bases ordenadas $\mathcal{B}_1$, $\mathcal{B}_2$ e $\mathcal{B}_\gamma$, respectivamente.

<u>EXEMPLO 1</u>

Sejam a matriz $A \in M_3(R)$;

$$A = \begin{bmatrix} 3 & 1 & 0 \\ -4 & -1 & 0 \\ -5 & -4 & 2 \end{bmatrix}.$$

Sua matriz caracterítiva é :-

$$A(X) = XE - A = \begin{bmatrix} X-3 & -1 & 0 \\ 4 & X+1 & 0 \\ 5 & 4 & X-2 \end{bmatrix},$$

como det $A(X) = (X-2)(X-1)^2$, temos que $D_3(X) = (X-2)(X-1)^2$.
O complemento algébrico de $X-2$ e do 4 da primeira coluna, são respectivamente, $(X-1)^2$ e $-X+2$. Logo o m.d.c. unitário entre os subdeterminantes de ordem 2, é $D_2(X) = 1$. Como $D_1(X)$ divide $D_2(X)$, $D_1(X) = 1$. Conseqentemente, os polinômios invariantes de A são,

$i_1(X) = i_2(X) = 1$, $i_3(X) = (X-2)(X-1)^2 = X^3-4X^2+5X-2$,

e a primeira forma Racional R_1 de A, é dada por :-

$$R_1 = C_{i_3} = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 0 & -5 \\ 0 & 1 & 4 \end{bmatrix}.$$

Os divisores elementares de A, são :-

$p_1(X) = (X-1)^2 = X^2-2X+1$, $p_2(X) = X-2$,

e a segunda forma Racional R_2 de A, é dada por:-

$$R_2 = C_{p_1} + C_{p_2} = \left[\begin{array}{cc|c} 0 & -1 & 0 \\ 1 & 2 & 0 \\ \hline 0 & 0 & 2 \end{array}\right]$$

Como os divisores elementares de A em $R[X]$, são:

$(X-1)^2$ e $(X-2)$,

A é semelhante a uma matriz na forma de Jordan de $M_3(R)$. A característica de Segre de A é $\{(2)(1)\}$, isto é :-

$$J = \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & 2 \end{array}\right] \quad \left(\text{ou } J = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \hline 0 & 0 & 2 \end{array}\right] \right).$$

EXEMPLO 2.

Seja a matriz A, de $M_4(C)$, dada por :-

$$A = \begin{bmatrix} 1 & -1 & 1 & -1 \\ -3 & 3 & -5 & 4 \\ 8 & -4 & 3 & -4 \\ 15 & -10 & 11 & -11 \end{bmatrix},$$

Através da seguinte sequência de operações elementares, na matriz característica $XE - A$, (voltaremos à êste exemplo no exemplo 13 de 4.6),

$[4.(1)+3]\ [4.(-1)+2]\ [4.(1-X)+1]$, $[1,4]$, $(1.(4)+2)(1.(-4)+3)$ $(1.(-X-11)+4)$, $[3(-X-1)+2]\ [3(1-4X)+4]$, $[2,3]$, $(2(-X-1)+3)(2(-X)+4)$, $(3(-1))$, $[3(-5)+4]$, $(3,4)$, $[3,4]$, $[3(X+1)+4]$, $(3(X+2)+4)$, $(4.(-1))$,

obtemos a matriz na forma canônica diagonal (ver (7) de 3.2) seguinte:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & X+1 & 0 \\ 0 & 0 & 0 & (X+1)^3 \end{bmatrix}.$$

Logo, os polinômios invariantes de A são :-

$$i_1(X) = i_2(X) = 1,\ i_3(X) = X+1 \ \text{ e } \ i_4(X) = (X+1)^3,$$

e do teorema 5 de 3.3, os divisores elementares do A, em $C[X]$(ou em $R[X]$) são :

$$(X+1)^3 \quad \text{e} \quad X+1.$$

Assim, se R_1, R_2 e J são a primeira forma Raciona, segunda forma Racional e forma de Jordan de A, temos que :

$$R_1 = C_{i_3} \oplus C_{i_4} = \left[\begin{array}{c|ccc} -1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & -3 \end{array}\right]$$

$$R_2 = C_{i_4} \oplus C_{i_3} = \left[\begin{array}{ccc|c} 0 & 0 & -1 & 0 \\ 1 & 0 & -3 & 0 \\ 0 & 1 & -3 & 0 \\ \hline 0 & 0 & 0 & -1 \end{array}\right],$$

(a ordem dos divisores elementares, sendo no sentido não crescente de seus graus).

$$J = \left[\begin{array}{ccc|c} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ \hline 0 & 0 & 0 & -1 \end{array}\right].$$

4.2 O TEOREMA DA DECOMPOSIÇÃO RACIONAL. REDUÇÃO À FORMA RACIONAL ATRAVÉS DE SUBESPAÇOS L-CÍCLICOS

No apêndice B (B.4), escrevemos sôbre subespaços L-cíclicos e anuladores, visando esta secção. Assim, os conceitos, notação e teoremas citados aqui, estão em B.4. O principal escôpo desta secção, é demonstrar que para um operador linear L, arbitrário sôbre V, existem r-vetores $v_1, v_2, \ldots, v_r$ em V, tais que :-

$$V = C(v_1,L) \oplus C(v_2,L) \oplus \ldots \oplus C(v_r,L).$$

Em outras palavras, demonstraremos que V é soma direta de subespaços L- cíclicos $C(v_i,L)$. Poderemos depois concluir, que L é a soma direta de um número finito de operadores lineares ($L_{C(v_i,L)}$), cada um dos quais, possui um vetor cíclico.

TEOREMA 1 (TEOREMA DA DECOMPOSIÇÃO RACIONAL).

"Seja L um operador linear sôbre um K-espaço vetorial V dedimensão finita $n \geq 1$. Então, existem r-vetores, não nulos $v_1, v_2, \ldots, v_r$, em V, com respectivos L-anuladores, $m_1, m_2, \ldots, m_r$, tais que :

$$V = C(v_1,L) \oplus C(v_2,L) \oplus \ldots \oplus C(v_r,L),$$

onde m_1 coincide com o polinômio mínimo de L e cada m_i é divisível por m_{i-1}, $i = 2,3,\ldots,r$".

Demonstração.

Seja,

$$m(X) = m_1(X) = X^t + a_1X^{t-1} + \ldots + a_t,$$

o polinômio mínimo de V. Pelo teorema 5 de B.4, existe um vetor v_1 em V, cujo L-anulador é $m_1 = m$. Consideremos o subespaço L-cíclico $C(v_1) = C(v_1,L)$ gerado por v_1, com base $\{v_1, Lv_1, \ldots, L^{t-1}v_1\}$. Caso $n = t$, ótimo. Toma-se $V = C(v_1)$ e o teorema fica demonstrado. Caso $n > t$, seja :

$$m_2(X) = X^s + b_1X^{s-1} + \ldots + b_s,$$

o polinômio mínimo de V(mod.$C(v_1)$) (ver B.4). Sabe-se que m_2 é um divisor de m_1, isto é, o polinômio mínimo "relativo" m_2, é um divisor do polinômio mínimo m_1 de L. Logo, existe um polinômio q de $K[X]$, tal que :-

(10) $\qquad m_1 = m_2 q$.

Sabemos também (versão do teorema 5 de B.4, para o conceito de L-anulador relativo) que existe um vetor u em V, cujo L-anulador relativo é m_2. Assim, temos :-

(11) $\qquad m_2(L)u \equiv o \pmod{C(v_1)}$,

isto é, existe um polinômio g de grau menor ou iguál à $t-1$ (lembre que $\{v_1, Lv_1, \dots, L^{t-1}v_1\}$ é base de $C(v_1)$), tal que :-

(12) $$m_2(L)u = g(L)v_1 \ .$$

Apliquemos à (12) o operador linear $q(L)$ e levando-se (1o) em consideração, vem que :-

$$o = m_1(L)u = q(L)\ m_2(L)u = q(L)\ g(L)v_1 \ ,$$

pois $m_1 = m$ é o polinômio mínimo de L. Logo, $q\, g \in M(v_1, L) = \{p \in K[X] : p(L)v_1 = o\}$, e é portanto, divisível por $m_1 = q\, m_2$, o L-anulador de v_1. Como qg é divisível por qm_2, temos que g é divisível por m_2, ou seja, existe $q_1 \in K[X]$, tal que :

(13) $$g = q_1 m_2 \ .$$

Podemos reescrever (12), pondo-se :

(14) $$m_2(L)u = m_2(L)q_1(L)v_1 \quad \text{ou} \quad m_2(L)\left[u - q_1(L)v_1\right] = o.$$

Tomemos,

(15) $$v_2 = u - q_1(L)v_1 \ .$$

Então, (14) se escreve :-

(16) $$m_2(L)\ v_2 = o,$$

e daí, m_2 é o L-anulador de v_2 e observamos que (ver (1o)) o L-anulador m_1 do v_1 é divisível por m_2. De (15), vem que :

$$v_2 - u = -q_1(L)v_1 \in C(v_1),$$

pois $C(v_1)$ é L-invariante. Logo,

(17) $$v_2 \equiv u(\text{mod.}C(v_1)) \ .$$

Sendo m_2 o L-anulador relativo de u, é também de v_2. Concluimos que m_2 é o L-anulador e o Lanulador relativo de v_2, assim uma base do subespaço cíclico $C(v_2, L) = C(v_2)$ é :-

(18) $$\{v_2, Lv_2, \dots, L^{s-1}v_2\} \ ,$$

pois $gr(m_2) = s$. Como m_2 é o Lanulador relativo de v_2 mod.$C(v_1)$, os vetores de (18) são linearmente independentes módulo $C(v_1)$, isto é, se

$$a_1v_2 + a_2Lv_2 + \dots + a_sL^{s-1}v_2 \equiv o\ (\text{mod.}C(v_1)),$$

então, $a_1 = a_2 = \dots = a_s = o$. Consequentemente, todo vetor não nulo, de $C(v_2)$, não é uma combinação linear da base $\{v_1, Lv_1, \dots, L^{t-1}v_1\}$, de $C(v_1)$, ou seja, $C(v_1) \cap C(v_2) = \langle o \rangle$. Assim, os $(t+s)$-vetores :-

(19) $$\{v_1, Lv_1, \dots, L^{t-1}v, v_2, Lv_2, \dots, L^{s-1}v_2\},$$

são linearmente independentes, e formam uma base do subespaço L-invariante $C(v_1) \oplus C(v_2)$, de dimensão $t+s$.

Caso $n = t+s$, ótimo. Toma-se $V = C(v_1) \oplus C(v_2)$ e o teorema fica demonstrado. Caso $n > t+s$, consideramos V (mod. $C(v_1) \oplus C(v_2)$) e continuamos nosso procedimento de obtenção de subespaços L-cíclicos, análogo ao que fizemos para obter $C(v_2)$. É claro que o processo deve ter um fim, uma vez que em cada etapa $\dim (C(v_1) \oplus C(v_2) \oplus \dots \oplus C(v_i))$ aumenta, de pelo menos, uma unidade, e a dimensão de V é finita. Concluimos assim a demonstração.

Ressaltamos que da demonstração do teorema 1, vem que o primeiro vetor v_1, satisfazendo o teorema 1, é um vetor de V, cujo L-anulador é o polinômio mínimo de L.

Nas condições do teorema 1, denotaremos por L_i, o operador linear induzido em $C(v_i,L)$, por L. Então, temos que :-

(20) $$L = L_1 \oplus L_2 \oplus \dots \oplus L_r ,$$

onde cada L_i possui um vetor cíclico (v_i) e o polinômio mínimo de L_i é o L-anulador de v_i, $i = 1,2, \dots ,r$. Do teorema 4 de B.4, temos:

$$gr(m_i) = \dim C(v_i,L) \;,\; i = 1,2, \dots ,r.$$

Logo, o L-anulador de v_i é também o polinômio característico de L_i, pois m_i divide o polinômio característico c_i de L_i, e $gr(m_i) = \dim C(v_i,L) = gr(c_i)$ e ambos são unitários. Seja,

$$\mathcal{B}_i = \{ v_i, Lv_i, \dots , L^{n_i-1}v_i \} ,$$

a "BASE CÍCLICA" de $C(v_i,L)$, onde $n_i = \dim C(v_i,L) = gr(m_i)$, $i=1,\dots,r$. A matriz A_i de L_i, relativamente à base ordenada $\mathcal{B}_i$, é a matriz companheira (ver 4.1) do polinômio mínimo m_i. De fato, se enumerarmos a base $\mathcal{B}_i$, pondo-se :-

$$\mathcal{B}_i = \{ e_{1i}, e_{2i}, \dots , e_{n_i i} \} ,$$

teremos:-

$$Le_{ji} = L(L^{j-1}v_i) = L^j v_i = e_{j+1\, i} \;,\; j = 1,2, \dots ,n_i-1$$

e $Le_{n_i i} = L^{n_i}v_i = -a_{1i}L^{n_i-1}v_i - a_{2i}L^{n_i-2}v_i - \dots - a_{n_i i}v_i$.

De (20), considerando-se a base ordenada $\mathcal{B}$ de V, formada pela união das bases cíclicas $\mathcal{B}_1, \mathcal{B}_2, \dots , \mathcal{B}_r$, nessa ordem, vem que :-

(21) $$[L]_{\mathcal{B}} = A_1 \oplus A_2 \oplus \dots \oplus A_r .$$

Como o polinômio característico de cada A_i, é o L-anulador m_i de v_i, $i = 1,2, \ldots ,r$, a regra da "forma em blocos" (ver) , para se calcular determinantes, nos diz que o polinômio característico de L, é dado por :-

(22) $$c = m_1 m_2 \ldots m_r .$$

Ademais, pelo teorema 1, sabemos que m_{i+1} divide m_i para $i = 1,\ldots,r-1$ e cada m_i é unitário, logo (21) é uma matriz na FORMA CANÔNICA RACIONAL.

O análogo matricial do Teorema da Decomposição Racional, é dado por :-

TEOREMA 2.

" Seja B uma matriz de $M_n(K)$. Então, B é semelhante no corpo K, a uma matriz $n \times n$ na Forma Canônica Racional."

Demonstração.

Seja L o operador linear de K^n, tal que :-

$$[L]_{\mathcal{B}} = B,$$

onde $\mathcal{B} = \{ e_i = (\delta_{i1}, \delta_{i2}, \ldots , \delta_{in}),\ i = 1,2, \ldots ,n \}$ é a base canônica de K^n. Como vimos acima, existe uma base ordenada $\mathcal{B}'$ de K^n, em relação à qual, L é representado por uma matriz A sob a Forma Canônica Racional. Então, se P é a matriz que muda a base $\mathcal{B}$ para a base $\mathcal{B}'$, temos :-

$$[L]_{\mathcal{B}'} = P^{-1} [L]_{\mathcal{B}} P \qquad \text{ou } A = P^{-1} B P.$$

c.q.d.

COROLÁRIO 1.

"Seja L um operador linear sôbre um K-espaço vetorial V de dimensão finita $n \geq 1$ (ou seja A uma matriz de $M_n(K)$). Então, os fatores irredutíveis do polinômio mínimo de L (ou de A), coincidem com os fatores irredutíveis do polinômio característico de L (ou de A)".

Demonstração.

De (22), vem que o polinômio característico c de L, é um produto de polinômios, cada um dos quais divide o polinômio mínimo de L (teorema 1, acima). Logo os fatores irredutíveis de cada m_i, são fatores irredutíveis de m. Assim sendo, os fatores irredutíveis de c e de m coincidem, porém êles podem repetir-se mais vêzes em c do que em m.

c.q.d.

O próximo corolário, fornece um critério para se saber quando um espaço vetorial V é <u>L-CÍCLICO</u>, ou seja, possui um vetor L-cíclico.

<u>COROLÁRIO 2.</u>

" Uma condição necessária e suficiente, para que um K-espaço vetorial V seja L-cíclico (L é um operador linear sôbre V), é que a dimensão de V coincida com o grau do polinômio mínimo de L, ou equivalentemente, que os polinômios mínimo e característico de L coincidam".

Demonstração.

Seja V um espaço vetorial L-cíclico, n-dimensional, com polinômio mínimo :-

$$m(X) = X^t + a_1X^{t-1} + \dots + a_t.$$

Seja v_1, tal que, $V = C(v_1,L)$. Pelo teorema 1, o L-anulador de v_1 é $m_1 = m$. Do teorema 4 de B.4, temos que : $gr(m) = t = \dim C(v_1,L) = \dim V = n$. Reciprocamente, suponhamos que $\dim V = gr(m)$. Pelo Teorema da Decomposição Racional, obtemos :

$$V = C(v_1,L) \oplus C(v_2,L) \oplus \dots \oplus C(v_r,L).$$

Mas a $\dim C(v_1,L) = gr(m) = \dim V$, pois o L-anulador de v_1 é o polinômio mínimo de L. Daí,

$$V = C(v_1,L),$$

isto é, V é L-cíclico.

c.q.d.

<u>EXEMPLO 3.</u>

Seja V um K-espaço vetorial de dimensão 2 e L um operador linear sôbre V. Caso o grau do polinômio mínimo m de L é 2, pelo corolário 4, temos que V é L-cíclico, ou seja, existe um vetor $v \in V$, tal que, $\gamma = \{v, Lv\}$ é uma base ordenada de V e $V = C(v,L)$. Sendo

$$m(X) = X^2 + a_1X + a_2,$$

temos que :

$$[L]_{\gamma} = \begin{bmatrix} 0 & -a_2 \\ 1 & -a_1 \end{bmatrix}.$$

A possibilidade restante, é que o grau do polinômio mínimo de L é 1. Assim, $m(X) = X-a$, logo

$$0 = m(L) = L - aI, \text{ ou seja, } L = aI,$$

isto é, L é um múltiplo escalar do operador idêntico. Nêsse caso, se v_1 e v_2 são vetores linearmente independentes de V, temos :-

$$V = C(v_1, L) \oplus C(v_2, L),$$

ainda, $m_1(X) = m_2(X) = X - a$ e se $\mathcal{B} = \{v_1, v_2\}$ é a união das bases cíclicas $\{v_1\}$ e $\{v_2\}$, nesta ordem, temos :-

$$[L]_{\mathcal{B}} = \begin{bmatrix} a & o \\ o & a \end{bmatrix}.$$

Podemos concluir ainda, que qualquer matriz 2×2 sôbre o corpo K, é semelhante sôbre K, a uma matriz de um dos dois tipos :

$$\begin{bmatrix} o & -a_2 \\ 1 & -a_1 \end{bmatrix} \quad \text{ou} \quad \begin{bmatrix} a & o \\ o & a \end{bmatrix} \quad .$$

EXEMPLO 4.

Queremos determinar a Forma Canônica Racional, da seguinte matriz, sôbre o corpo R,

$$A = \begin{bmatrix} \cos\theta & \text{sen}\theta \\ -\text{sen}\,\theta & \cos\theta \end{bmatrix}$$

O polinômio característico de A é :

$$c(X) = \det(XE-A) = \begin{vmatrix} X - \cos\theta & -\text{sen}\,\theta \\ \text{sen}\,\theta & X-\cos\theta \end{vmatrix} = X^2 - 2X\cos\theta + 1.$$

Então, $\Delta = (-2\cos\theta)^2 - 4 = 4(\cos^2\theta - 1) \leqslant o$, pois $\cos^2\theta \leqslant 1$. Agora, caso $\theta = k\pi$, $k \in Z$, $\Delta = o$ e $m(X) = X \mp 1$ e a forma conônica Racional de A é :

$$\begin{bmatrix} \mp 1 & o \\ o & \pm 1 \end{bmatrix} \qquad (+ \text{ se } k \text{ é par, } - \text{ se } k \text{ é ímpar}).$$

Caso $\theta \neq k\pi$, $k \in Z$, $\Delta < o$ e $m(X) = c(X) = X^2 - 2X\cos\theta + 1$, sendo nêste caso, a Forma Canônica Racional de A dada por :

$$\begin{bmatrix} o & -1 \\ 1 & 2\cos\theta \end{bmatrix}$$

EXEMPLO 5.

Seja L o operador linear sôbre R^3, que é representado pela matriz A abaixo, relativamente à base canônica de R^3,

$$A = \begin{bmatrix} 3 & -4 & -4 \\ -1 & 3 & 2 \\ 2 & -4 & -3 \end{bmatrix}.$$

Temos que :

$$c(X) = \det(XE - A) = \begin{bmatrix} X-3 & 4 & 4 \\ 1 & X-3 & -2 \\ -2 & 4 & X+3 \end{bmatrix} = (X-1)^3.$$

Verifica-se através de um cálculo direto, que $(a - 1E)^2 = 0$. Logo, como $A \neq E$, $m(X) = (X-1)^2$. Assim, na decomposição cíclica de R^3, o L-anulador de v_1 tem que ser $m_1(X) = (X-1)^2$. Daí, $C(v_1,L)$ tem dimensão $2 = gr(m)$, e como $\dim R^3 = 3$, só pode existir mais um vetor v_2. Ademais, $\dim C(v_2,L) = 1$, logo v_2 tem que ser um autovetor de L (ver exemplos 3 de B.4), e seu L-anulador tem que ser $m_2(X) = X - 1$, uma vez que $c = m_1 m_2$ (ver (22)). Já temos, a matriz B na Forma Canônica Racional, semelhante à A:

$$B = \left[\begin{array}{cc|c} 0 & -1 & 0 \\ 1 & 2 & 0 \\ \hline 0 & 0 & 1 \end{array}\right],$$

que é a soma direta, das matrizes companheiras de $m_1(X) = (X-1)^2 = X^2 - 2X + 1$ e $m_2(X) = X-1$, respectivamente. Vamos querer determinar uma base ordenada de V, em relação à qual L é representado por B. Para tanto, devemos encontrar convenientes vetores v_1 e v_2. Como já sabemos que v_2 deve ser um autovalor de L, vamos encontrá-lo primeiro. O único autovalor de L é 1, logo para encontrarmos v_2, devemos resolver o sistema :

$$(a - 1E)[x] = 0 \iff \begin{bmatrix} 2 & -4 & -4 \\ -1 & 2 & 2 \\ 2 & -4 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff \{x_1 = 2x_2 + 2x_3,$$

logo um autovetor genérico de L é da forma ($2a+2b, a, b$) (com a ou b, não nulo). Vemos assim, que podemos determinar infinitos vetores v_2, satisfazendo o teorema 1. Seja $v_2 = (2,1,0)$. Para determinarmos v_1, queremos que $\{v_1, v_2\}$ sejam linearmente independente e que $\dim C(v_1,L) = 2$, logo v_1 não pode ser um autovetor de L. Tomemos, por exemplo, $v_1 = (1,0,0)$. Temos que $Lv_1 = (3,-1,2)$, que não é linearmente dependente com v_1. Assim sendo, $C(v_1,L)$ tem dimensão 2 e os vetores de $C(v_1,L)$ são da forma :

$$a(1,0,0) + b(3,-1,2) = (a + 3b, -b, 2b),$$

ou seja, todos os vetores (x_1,x_2,x_3) tais que : $+x_3 = -x_2$. Como $v_2 = (2,1,0) \notin C(v_1,L)$ e $C(v_2,L) = \langle v_2 \rangle$, temos que $C(v_1,L) \cap C(v_2,L) = \langle 0 \rangle$. Pode-se verificar diretamente, que a matriz representando L, relativamente à base ordenada $\{v_1, Lv_1, v_2\} = \{(1,0,0),(3,-1,2),(2,1,0)\}$ é exatamente a matriz B. Ademais, uma matriz inversível P de $M_3(K)$, tal que $P^{-1}AP$ esta sob a Forma Canônica Racional B, é :-

$$P = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \\ 0 & 2 & 0 \end{bmatrix}$$

4.3 REDUÇÃO À FORMA DE JORDAN ATRAVÉS DOS TEOREMAS DA DECOMPOSIÇÃO RACIONAL E PRIMÁRIA.

Um operador linear N sôbre um K-espaço vetorial V de dimensão finita n, diz-se NULIPOTENTE, se existe um inteiro positivo k, tal que :-

$$N^k = 0.$$

Ao menor inteiro positivo k (quando existe um), tal que, $N^k = 0$, denomina-se ÍNDICE DE NULIPOTÊNCIA de N.

Consideremos a "decomposição cíclica", do operador linear N, dada pelo teorema da Decomposição Racional (teorema 1 de 4.2). Sejam $v_1, v_2, \dots, v_r$, vetores não nulos de V com N-anuladóres $m_1, m_2, \dots, m_r$, tais que :-

$$V = C(v_1,N) \oplus C(v_2,N) \oplus \dots \oplus C(v_r,N) ,$$

e m_{i+1} divide m_i, para $i = 1,2, \dots, r-1$.

Sendo N nulipotente, seu polinômio mínimo divide X^k, pois $N^k = 0$. Logo, se o índice de nulipotência de N é k, seu polinômio mínimo é $m(X) = X^k$. O polinômio característico de N é $c(X) = X^n$, pois c e m têm os mesmos zeros em K (teorema 10 de 1.4). Assim sendo , cada N-anulador m_i é da forma :-

$$m_i(X) = X^{n_i}, \ i = 1,2, \dots r.$$

A condição de divisibilidade, dos N-anuladores, nos diz que :-

$$k = n_1 \geqslant n_2 \geqslant \dots \geqslant n_r \geqslant 1.$$

[illegible] do N-anulador $m_i(X) = X^{n_i}$, é a matriz $n_i \times n_i$.
[illegible] :-

$$(23) \qquad A_i = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

Pelo Teorema da Decomposição Racional, existe uma base ordenada de V, em relação à qual, N é representado pela soma direta das matrizes nulipotentes (23), cujas dimensões (n_i) <u>diminuem</u>, à medida que i <u>aumenta</u>. Vemos que a uma matriz "nulipotente" (definição óbvia) dada, está associado, um número inteiro r positivo e r-inteiros positivos $n_1 \geqslant n_2 \geqslant \ldots \geqslant n_r$, tais que : $n_1 + n_2 + \ldots + n_r = n$, e êstes inteiros positivos, determinam a forma Canônica Racional da matriz nulipotente dada.

Ressaltamos aqui, que a Forma Canônica Racional, existe sempre, para qualquer matriz, sem restrições no corpo K. Tal não é o caso, para a Forma de Jordan, pois devemos admitir que <u>todos os zeros</u> (autovalores de L) do polinômio característico de L (ou de A) estão em K, para podermos obter a Forma de Jordan para A.

Suponhamos que o polinômio característico c de L, possa ser colocado na forma :-

$$(24) \qquad \begin{cases} c(X) = \prod_{i=1}^{s}(X - \lambda_i)^{d_i}, \\ \text{onde para cada } i=1,2,\ldots,s,\ d_i \in Z,\ d_i \geqslant 1 \text{ e } \lambda_i \in K,\ \lambda_j \neq \lambda_i \text{ se} \\ i \neq j. \end{cases}$$

Sabemos que os zeros do polinômio mínimo m de L, coincidem com os zeros do polinômio característico c de L (teorema 10 de 1.4). De (24), vem que podemos fatorar m na forma :-

$$(25) \qquad \begin{cases} m(X) = \prod_{i=1}^{s}(X - \lambda_i)^{r_i} \\ \text{onde para cada } i = 1,2, \ldots ,s,\ r_i \in Z,\ 1 \leqslant r_i \leqslant d_i \text{ e} \\ \lambda_i \in K, \text{ ainda } \lambda_j \neq \lambda_i, \text{ se } j \neq i. \end{cases}$$

Seja V_i o núcleo do operador linear $(L - \lambda_i I)^{r_i}$. De (25), temos as hipóteses do teorema da Decomposição Primária (teorema 4 de 2.3) para L, satisfeitas. Tal teorema, nos diz que :-

UNICAMP
BIBLIOTECA CENTRAL

(26) $$V = V_1 \oplus V_2 \oplus \cdots \oplus V_s,$$

e que o operador linear L_i, induzido sôbre V_i, por L, possui polinômio mínimo :

(27) $$m_i(X) = (X - \lambda_i)^{r_i}.$$

Fixemos um índice i, arbitrariamente, em $\{1, 2, \ldots, s\}$. Definimos :-

$$N_i = L_i - \lambda_i I,$$

operador linear sôbre V_i (estamos cometendo o abuso de notação I, para V ou para V_i, é o operador idêntico). De (27), vem que :-

$$0 = m_i(L_i) = (L_i - \lambda_i I)^{r_i} = N_i^{r_i},$$

isto é, N_i é um operador linear nilpotente sôbre V_i, de índice de nilpotência igual à r_i. Logo, o polinômio mínimo de N_i é X^{r_i}.

Apliquemos agora, o teorema da Decomposição Racional (Teorema 1 de 4.2), para o operador linear N_i, obtemos assim, n_i-vetores, $v_{1i}, v_{2i}, \ldots, v_{n_i i}$, não nulos em V_i, com respectivos N_i-anuladores :-

$$m_{1i}, m_{2i}, \ldots, m_{n_i i},$$

tal que :-

(28) $$V_i = C(v_{1i}, N_i) \oplus C(v_{2i}, N_i) \oplus \cdots \oplus C(v_{n_i i}, N_i),$$

onde m_{1i} coincide com o polinômio mínimo X^{r_i} de N_i e $m_{j+1\,i}$ divide $m_{j\,i}$, $j = 1, 2, \ldots, n_i - 1$. Ademais, existe uma base ordenada $\mathcal{B}_i$ de V_i, relativamente à qual, a matriz de N_i é a soma direta das matrizes companheiras dos polinômios $m_{1i}, \ldots, m_{n_i i}$. Como $N_i = L_i - \lambda_i I$, temos que, $L_i = N_i + \lambda_i I$, e daí :

$$[L_i]_{\mathcal{B}_i} = [N_i]_{\mathcal{B}_i} + \lambda_i [I]_{\mathcal{B}_i},$$

ou seja, relativamente à essa base ordenada, a matriz de L_i é a soma direta de n_i-matrizes, J_{j,λ_i} da forma :-

(29) $$J_{j,\lambda_i} = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{bmatrix} \quad (j = 1, 2, \ldots, n_i).$$

Além disso, a dimensão das matrizes J_{j,λ_i} (blocos de Jordan), não aumentam (diminuem ou permanece) à medida que j aumenta .

De (26), podemos formar uma base ordenada de V, unindo as bases ordenadas $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_s$, nesta ordem, de $V_1, V_2, \dots, V_s$, respectivamente. Estamos interessados na matriz $[L]_{\mathcal{B}} = J$. De(26), temos :-

$$(30) \qquad J = A_1 \oplus A_2 \oplus \dots \oplus A_s,$$

onde $A_i = [L_i]_{\mathcal{B}_i}$, $i = 1,2, \dots, s$. De (28), temos que para cada $i = 1,2, \dots, s$,

$$(31) \qquad A_i = J_{n_i,\lambda_i} \oplus J_{n_i-1,\lambda_i} \oplus \dots \oplus J_{1,\lambda_i},$$

onde cada J_{j,λ_i} é um bloco de Jordan relativo ao autovalor λ_i. Além disso, as dimensões dessas matrizes, não aumentam à medida que j cresce. Substituindo (31) em (30), obtém-se :

$$(32) \quad J = J_{n_1,\lambda_1} \oplus \dots \oplus J_{1,\lambda_1} \oplus J_{n_2,\lambda_2} \oplus \dots \oplus J_{1,\lambda_2} \oplus \dots \oplus J_{n_s,\lambda_s} \oplus \dots \oplus J_{1,\lambda_s}$$

ou seja, $[L]_{\mathcal{B}} = J$ está na FORMA CANÔNICA DE JORDAN.

Acabamos de demonstrar, o seguinte teorema :-

TEOREMA 3. (FORMA CANÔNICA DE JORDAN).

" Seja L um operador linear sôbre um K-espaço vetorial V, de dimensão finita $n \geq 1$. Admitamos que todos os zeros do polinômio característico (autovalores) c de L, estão em K, ou seja, que c se fatora na forma (24) (isto sempre acontece, se K for um corpo algèbricamente fechado, por exemplo, para K = C). Então, existe uma base ordenada $\mathcal{B}$ de V, em relação à qual, a matriz J que representa L, esta na Forma Canônica de Jordan. Ainda mais, esta matriz J na forma de Jordan, esta associada de modo único, à menos da ordem dos λ_i na fatoração (24) de c, ao operador linear L."

Vale, òbviamente, um resultado análogo para uma matriz A de $M_n(K)$,

TEOREMA 4 (FORMA CANÔNICA DE JORDAN PARA MATRIZES).

" Seja A uma matriz de $M_n(K)$, cujo polinômio característico c, se fatora na forma (24). Então, A é semelhante sôbre K, à uma matriz J $n \times n$, na Forma de Jordan. Esta matriz J é univocamente determinada, à menos da ordem dos λ_i, na fatoração (24) de c."

Nas condições do teorema 4, a matriz J, diz-se a <u>FORMA DE JORDAN</u>, ou a <u>FORMA CANÔNICA DE JORDAN</u> da matriz A.

Vamos ressaltar aqui, fatos que caracterizam a matriz J na forma de Jordan, muito embora já fizemos algo nêste sentido em 4.1.

<u>1.</u> Todo elemento da matriz J, que não esta na diagonal, ou imediatamente acima dela é nulo. Na diagonal de J, aparecem os s-autovalores distintos dois à dois, $\lambda_1, \lambda_2, \ldots, \lambda_s$, de L. Ademais, cada autovalor λ_i se repete d_i-vêzes na diagonal, onde d_i é a multiplicidade algébrica do autovalor λ_i, isto é, d_i é a multiplicidade de λ_i como zero do polinómio característico de L.

<u>2.</u> Para cada $i = 1, 2, \ldots, s$, a matriz A_i de (30) é a soma direta de n_i blocos de Jordan J_{j,λ_i}, cujas dimensões não aumentam, quando j cresce. O número n_i é exatamente a multiplicidade geométrica do autovalor λ_i, ou seja, n_i é a dimensão do subespaço de V, gerado pelos vetores característicos de L, associados ao autovalor λ_i. De fato, n_i é o número de blocos nilpotentes do tipo (23), na Forma Racional de $N_i = L_i - \lambda_i I$, e como em cada bloco de Jordan, só existe um autovetor de L associado à λ_i sendo portanto, n_i o número de autovetores de L associados à λ_i. Em particular, L é diagonalizável se, e sòmente se, $n_i = d_i$, $i = 1, 2, \ldots, s$.

<u>3.</u> Para cada i, o primeiro bloco J_{n_{i1},λ_i} de A_i é uma matriz $n_{i1} \times n_{i1}$, onde n_{i1} é a multiplicidade de λ_i, como raiz do polinômio mínimo m de L. Isto porque, o polinômio mínimo do operador nilpotente $N_i = L_i - \lambda_i I$ é $x^{n_{i1}}$, e m_{Li} é o polinômio mínimo de L_i.

<u>4.</u> É claro, que fixada a ordem dos λ_i, em (24), temos a <u>CARACTERÍSTICA DE SEGRE</u> de L. Fixada a ordem, $\lambda_1, \lambda_2, \ldots, \lambda_s$ dos autovalores em (24), a característica de Segre é dada pelas dimensões de :-

$J_{n_1,\lambda_1}, J_{n_1-1,\lambda_1}, \ldots, J_{1,\lambda_1}, \ldots, J_{n_s,\lambda_s}, \ldots, J_{1,\lambda_s},$

respectivamente.

<u>5.</u> Dada uma matriz A de $M_n(K)$, é possível obtermos um processo sistemático (o que faremos em 4.6), para se determinar uma matriz inversível P, tal que $P^{-1}AP$ esteja na Forma de Jordan. Acontece, que em muitos casos particulares, pode-se obter J sem que se tenha P. Podemos proceder como segue. Determinamos os autovalores distintos

dois a dois, $\lambda_1, \lambda_2, \ldots, \lambda_s$; bem como suas multiplicidades algébricas $d_1, d_2, \ldots, d_s$. Caso $d_i = 1$, temos que :-

$$A_i = J_{1, \lambda_i} = [\lambda_i] .$$

Daí só pode haver ambigüidade sôbre a dimensão dos blocos de Jordan J_{j, λ_i} de A_i, se $d_i > 1$. Para os $d_i > 1$, calculamos a multiplicidade geométrica n_i de λ_i. Temos agora, a seguinte situação :

(i) $d_i = n_i$, então $A_i = \lambda_i E$, ou seja, a sub-matriz A_i esta na forma diagonal $[\lambda_i, \lambda_i, \ldots, \lambda_i]$.

(ii) $d_i > n_i$. Nêsse caso, para pequenos valores de d_i, há um número pequeno de formas possíveis para A_i, e nada podemos dizer para grandes valores de d_i. Por exemplo, seja $d_i = 3$, então o conhecimento de n_i, nos determina a forma de A_i, pois $d_i - n_i$ é o número dos 1 acima da diagonal de A_i. Consequentemente, se $n_i = 2$ (ou $n_i=1$), temos:

$$A_i = \begin{bmatrix} \lambda_i & 1 & 0 \\ 0 & \lambda_i & 0 \\ 0 & 0 & \lambda_i \end{bmatrix} \quad \left(\text{ou } A_i = \begin{bmatrix} \lambda_i & 1 & 0 \\ 0 & \lambda_i & 1 \\ 0 & 0 & \lambda_i \end{bmatrix} \right).$$

Já para $d_i = 4$, pode ocorrer ambiguidade, pois se $n_i=2$, podemos ter:

$$A_i = \begin{bmatrix} \lambda_i & 1 & 0 & 0 \\ 0 & \lambda_i & 1 & 0 \\ 0 & 0 & \lambda_i & 0 \\ 0 & 0 & 0 & \lambda_i \end{bmatrix} \quad \text{ou } A_i = \begin{bmatrix} \lambda_i & 1 & 0 & 0 \\ 0 & \lambda_i & 0 & 0 \\ 0 & 0 & \lambda_i & 1 \\ 0 & 0 & 0 & \lambda_i \end{bmatrix}.$$

Mesmo nêste caso ($d_i=4$), não há ambiguidade se $n_i=1$ (ou $n_i=3$), pois:

$$A_i = \begin{bmatrix} \lambda_i & 1 & 0 & 0 \\ 0 & \lambda_i & 1 & 0 \\ 0 & 0 & \lambda_i & 1 \\ 0 & 0 & 0 & \lambda_i \end{bmatrix} \left(\text{ou } A_i = \begin{bmatrix} \lambda_i & 1 & 0 & 0 \\ 0 & \lambda_i & 0 & 0 \\ 0 & 0 & \lambda_i & 0 \\ 0 & 0 & 0 & \lambda_i \end{bmatrix} \right).$$

Para matrizes A de ordem n pequena, estas informações são bastante úteis, para se determinar J, semelhante à A.

EXEMPLO 6.

Seja L um operador linear sôbre um C-espaço vetorial V de dimensão 2. Então, o polinômio característico de L pode ser colocado na forma (C é algèbricamente fechado):

a) $c(X) = (X-a_1)(X-a_2)$, $a_1 \neq a_2$.

Nêste caso, $n_i = d_i$, $i = 1, 2$, e L é diagonalizável, isto é, [illegible] uma base ordenada $\mathfrak{B}$ de V, tal que :-

$$[L]_{\mathfrak{B}} = \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix}.$$

b) $c(X) = (X - a)^2$. Nêste caso temos duas possibilidades, para o polinômio mínimo de L, que são :-

(i) $m(X) = X - a$. Daí, $L - aI = 0$ ou $L = aI$, e a matriz de L, relativamente à qualquer base de V é :

$$[L] = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}.$$

(ii) $m(X) = c(X) = (X - a)^2$. Daí, $n_1 = 1$ e $d_1 = 2$. Logo existe uma base ordenada $\mathfrak{B}$ de V, tal que :

$$[L]_{\mathfrak{B}} = \begin{bmatrix} a & 1 \\ 0 & a \end{bmatrix},$$

nêste caso, esta é a representação matricial mais simples possível para L.

Podemos concluir daí, que tôda matriz 2×2 sôbre C, é semelhante à uma matriz de um dos dois tipos abaixo :-

$$\begin{bmatrix} a & 1 \\ 0 & a \end{bmatrix} \quad \text{ou} \quad \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix} \text{(possivelmente } a_1 = a_2\text{)}$$

EXEMPLO 7.

Queremos determinar uma Forma de Jordan, para a matriz $A = [a_{ij}]$ de $M_6(R)$, que representa um operador linear L sôbre R^6, relativamente à base canônica :

$$\mathfrak{B} = \{ e_i = (\delta_{i1}, \delta_{i2}, \ldots, \delta_{i6}), \ i = 1, 2, \ldots, 6 \},$$

onde

$$[L]_{\mathfrak{B}} = A = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 \\ -1 & 0 & 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 0 & 0 \\ 1 & 1 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

Temos que :-

$$c(X) = \det(XE - A) = (X - 2)^5(X + 1).$$

As possibilidades para o polinômio mínimo são :

$(X - 2)(X + 1)$, $(X - 2)^2(X + 1)$, $(X - 2)^3(X + 1)$, $(X - 2)^4(X + 1)$ e $c(X)$.

Por um cálculo direto vê-se que $m(X) = (X-2)^4(X+1)$. Sejam V_1 e V_2 os núcleos dos operadores lineares : $(L-2I)^4$ e $(L+I)$, respectivamente. Pelo teorema da Decomposição Primária (teorema 4 de 2.3) vem :

$$R^6 = V_1 \oplus V_2 .$$

Sejam L_1 e L_2, operadores lineares induzidos sôbre V_1 e V_2, respectivamente, por L. Sejam $N_1 = L_1 - 2I$ e $N_2 = L_2 + I$, operadores lineares sôbre V_1 e V_2, respectivamente. O polinômio mínimo de N_1 é $m_{11}(X) = X^4$ e como seu polinômio característico é $c_1(X) = m_{11}(X)m_{21}(X) = X^5$ e $m_{11}(X) = X^4$, vem que $m_{21}(X) = X$. Logo,

$$J_{2,2} = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix} \quad \text{e} \quad J_{1,2} = \begin{bmatrix} 2 \end{bmatrix} ,$$

e

$$A_1 = J_{2,2} \oplus J_{1,2} = \left[\begin{array}{cccc|c} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ \hline 0 & 0 & 0 & 0 & 2 \end{array}\right]$$

O polinômio mínimo de N_2 é $m_{21}(X) = X$ e o polinômio característico é $c_2(X) = X$. Logo,

$$J_{1,-1} = \begin{bmatrix} -1 \end{bmatrix} = A_2 .$$

Fixando-se a ordem $\lambda_1 = 2$ e $\lambda_2 = -1$, para as raízes de c, a característica de Segre de L (ou de A) é $\{(4,1)\ (1)\}$ e então :

$$J = A_1 \oplus A_2 = J_{2,2} \oplus J_{1,2} \oplus J_{1,-1}, \text{ ou seja,}$$

$$J = \left[\begin{array}{cccc|c|c} 2 & 1 & 0 & 0 & & \\ 0 & 2 & 1 & 0 & & \\ 0 & 0 & 2 & 1 & & \\ 0 & 0 & 0 & 2 & & \\ \hline & & & & 2 & \\ \hline & & & & & -1 \end{array}\right]$$

4.4 REDUÇÃO À FORMA CANÔNICA DE JORDAN ATRAVÉS DE AUTOVALORES.

Apresentaremos aqui uma demonstração alternativa do teorema 3 de 4.3, bem como introduziremos alguns resultados que utilizaremos em 4.5. Começamos com um teorema que nos decompõe um operador linear arbitrário, na soma direta de um Nulipotente com um Não-singular. Antes dêste teorema, lembramos que os subespaços imagem de L, L(V), e núcleo de L, N(L), são dados por :-

$$L(V) = \{ w \in V : \exists v \in V, Lv = w \}$$

e

$$N(L) = \{ v \in V : Lv = 0 \}.$$

Vamos mostrar que $L^2(V) \subseteq L(V)$. De fato,

$\forall w \in L^2(V) \Rightarrow \exists v \in V : L^2 v = L(Lv) = w \Rightarrow \exists v_1 = Lv \in V : Lv_1 = w.$

Concluimos assim que $\dim L^2(V) \leqslant \dim L(V)$. Mais geralmente, os subespaços imagens das potências sucessivas de L, formam uma "cadeia", tal que :

(33) $\quad V \supsetneq L(V) \supsetneq L^2(V) \supsetneq \ldots \supsetneq L^a(V) = L^{a+1}(V) = \ldots ,$

pois dim V = n (finita) e $\dim L^i(V) > \dim L^{i+1}(V)$, $i = 1, 2, \ldots, a-1$.

Analogamente, $N(L^2) \supseteq N(L)$, pois

$\forall v \in N(L) \Rightarrow Lv = 0 \Rightarrow L^2 v = L(Lv) = L0 = 0 \Rightarrow v \in N(L^2).$

Daí, $\dim N(L^2) \geqslant \dim N(L)$. Com,

(34) $\qquad \dim N(L^i) + \dim L^i(V) = \dim V , \; i = 1, 2, \ldots$

(ver I.1), podemos concluir de (34) que :

$$N(L^a) = N(L^{a+p}) \Longleftrightarrow L^a(V) = L^{a+p}(V), \; \forall p \in \mathbb{N}.$$

Então de (33), temos a seguinte "cadeia"

(35) $\quad \langle 0 \rangle \subsetneq N(L) \subsetneq N(L^2) \subsetneq \ldots \subsetneq N(L^a) = N(L^{a+1}) = \ldots$

Ao inteiro não negativo a, dado por (33) ou (35) denominamos ASCENDENTE de L.

TEOREMA 5.

"Seja L um operador linear sôbre um K-espaço vetorial V, de dimensão finita n. Então existem subespaços V_1 e V_2 de V, tais que :-

(i) $V = V_1 \oplus V_2$

(ii) V_1 e V_2 são L-invariantes

(iii) L_1, operador linear induzido sôbre V_1 por L, é não-singular

L_2, operador linear induzido sôbre V_2 por L, é nulipotente."

Demonstração.

(i) Sejam $V_1 = L^a(V)$ e $V_2 = N(L^a)$, onde a é o ascendente de L. De (34), vemos que :

(36) $\quad \dim L^a(V) + \dim N(L^a) = \dim V_1 + \dim V_2 = \dim V = n.$

Seja $v \in V_1 \cap V_2$, qualquer. Então, existe $w \in V$, tal que, $L^a w = v$ e [illegible]

$L^a v = o$, temos $L^a v = L^a(L^a w) = L^{2a} w = o$. Assim, $w \in N(L^{2a}) \overset{(35)}{=\!=} N(L^a)$, logo $v = L^a w = o$. Isto prova que :

(37) $V_1 \cap V_2 = L^a(V) \cap N(L^a) = \{o\}$.

De (36) e (37), podemos concluir que $V = V_1 \oplus V_2$.

(ii) $L(L^a(V)) = L^{a+1}(V) \overset{(33)}{=\!=} L^a(V)$, isto é, $V_1 = L^a(V)$ é L-invariante. Por outro lado,

$\forall v \in N(L^a) \overset{(35)}{\Longrightarrow} v \in N(L^{a+1}) \Longrightarrow L^a(Lv) = o \Longrightarrow Lv \in N(L^a)$, isto é, $V_2 = N(L^a)$ é L-invariante .

(iii) $L(V_1) = L(L^a(V)) = L^{a+1}(V) \overset{(33)}{=\!=} L^a(V) = V_1$, ou seja, L aplica V_1 sôbre V_1. Sendo dim $V_1 \leqslant n$, logo finita, temos que L_1 (restrição de L à V_1) é injetora ou Não-singular, pois mostramos que L_1 é sôbre; Por outro lado,

$\forall v \in V_2 = N(L^a) \Longrightarrow L^a v = L_2^a v = o$,

logo existe um inteiro não negativo a, tal que $L_2^a v = o$, para todo v de V_2, ou seja, L_2 é nulipotente, c.q.d.

TEOREMA 6. (FORMA CANÔNICA DE JORDAN).

" Seja L um operador linear sôbre V, espaço vetorial de dimensão finita n, sôbre um corpo F algèbricamente fechado (por exemplo, F=C). Sejam $\lambda_1, \lambda_2, \ldots, \lambda_s$ os autovalores, dois à dois distintos de L, com multiplicidades algébricas $d_1, d_2, \ldots, d_s$, respectivamente. Então, V é uma soma direta de s subespaços :-

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_s:$$

tais que para cada $j = 1,2, \ldots, s$, têm-se :

(i) V_j é L-invariante

(ii) dim $V_j = d_j$

(iii) O operador linear L_j induzido sôbre V_j por L, se escreve na forma : $L_j = \lambda_j I + N_j$, onde N_j é nulipotente e I o operador idêntico sôbre V".

Demonstração.

Definimos o operador linear $L^{(1)}$ sôbre V, pondo :

(38) $$L^{(1)} = L - \lambda_1 I.$$

Apliquemos o teorema 5 para $L^{(1)}$, obtendo os subespaços $L^{(1)}$-invariantes V_1 e W_1, tais que :

(39) $$V = V_1 \oplus W_1:$$

onde $L^{(1)}$ é nulipotente sôbre V_1 e não-singular sôbre W_1. De (38), temos que :

$\forall v \in V_1$ (respectivamente, $v \in W_1$), $Lv = L^{(1)}v + \lambda Iv = (L^{(1)}v + \lambda v) \in V_1$ (ou $(L^{(1)}v + \lambda v) \in W_1$), pois V_1 (ou W_1) é $L^{(1)}$-invariante. Isto mostra que V_1 e W_1 são L-invariantes. Logo, $L_1 = L_1^{(1)} + \lambda_1 I$, onde $L_1^{(1)}$ é a restrição de $L^{(1)}$ à V_1, que é nulipotente sôbre V_1. Mostramos que vale (iii) para $j = 1$.

Mostraremos que $d_1 = \dim V_1$. Seja $\mathcal{B}_1$ e $\mathcal{B}_2$ bases ordenadas de V_1 e W_1, respectivamente. De (39) (ver teorema 1 de B.2), $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$, é uma base de V que ordenamos por $\mathcal{B}_1, \mathcal{B}_2$. Daí, temos a seguinte representação matricial :-

$$(40)\quad \mathfrak{D} = [L]_{\mathcal{B}} = \begin{bmatrix} [L_1]_{\mathcal{B}_1} & 0 \\ 0 & [L_2]_{\mathcal{B}_2} \end{bmatrix} = \begin{bmatrix} \mathfrak{D}_1 & 0 \\ 0 & \mathfrak{D}_2 \end{bmatrix},$$

onde L_1 e L_2, são os operadores lineares induzidos por L sôbre V_1 e V_2, respectivamente. De (38),

$$[L_2^{(1)}] = \mathfrak{D}_2 - \lambda_1 E.$$

Pela fórmula para se calcular o determinante de uma matriz em blocos de (40) vem que :-

$c(X) = \det(XI-L) = \det(XE - \mathfrak{D}_1).\det(XE - \mathfrak{D}_2)$,

e como $L^{(1)}$ é não-singular em W_1, temos que :

$$\det(\mathfrak{D}_2 - XE) \neq 0 \qquad \text{ou} \qquad \det(XE - \mathfrak{D}_2) \neq 0.$$

Logo, a dim V_1 é ao menos tão grande quanto a multiplicidade algébrica de λ_1 (a matriz $\mathfrak{D}_1$ pode, a princípio, ter autovalores distintos de λ_1), daí, $\dim V_1 \geq d_1$.

Por outro lado, $L^{(1)}$ é nulipotente sôbre V_1, e repetindo-se a argumentação que fizemos no início de 4.3, podemos escolher a base ordenada $\mathcal{B}_1$ para V_1, tal que, relativamente à $\mathcal{B}_1$, $L_1^{(1)}$ é representado por uma matriz com o ou 1 acima da diagonal e zero em qualquer outra posição. Desde que, $L^{(1)} = L_1^{(1)} + \lambda_1 I$, a matriz $[L_1]_{\mathcal{B}_1} = \mathfrak{D}_1$ tem λ_1 em cada posição da diagonal, o ou 1 acima dela, e zero nas posições restantes (isto é, é uma soma direta de blocos de Jordan). Daí, vemos que a matriz $\mathfrak{D}_1$ só tem um autovalor(λ_1). Donde, $\dim V_1 \leq d_1$. Isto mostra que vale (ii) para $j = 1$.

Todos os ítens do teorema estão estabelecidos para o caso $j = 1$. Consideremos agora, o operador linear L_2', induzido sôbre W_1 por L. Repetimos os argumentos usando :-

$$E^{(2)} = L_2' - \lambda_2 I.$$

Como a dim $V = n$, finita, o teorema resulta após s-etapas.

A versão dêste teorema 6 para matrizes é :-

<u>TEOREMA 7.</u>(<u>FORMA DE JORDAN PARA UMA MATRIZ</u>).

"Seja A, uma matriz de $M_n(F)$, onde F é um corpo algèbricamente fechado. Sejam $\lambda_1, \lambda_2, \ldots, \lambda_s$ os autovalores, distintos dois à dois, de A, com multiplicidades algébricas $d_1, d_2, \ldots, d_s$, respectivamente. Então, A é semelhante a uma matriz da forma :-

$$J = A_1 \oplus A_2 \oplus \ldots \oplus A_s,$$

onde cada sub-matriz A_i é $d_i \times d_i$, com λ_i em cada posição diagonal, o ou 1 acima dela, em uma certa ordem, e o nas posições restantes".

Tornaremos a fazer alguns comentários à forma de Jordan J, de uma matriz quadrada $n \times n$. Esta matriz J, na forma de Jordan e semelhante à A, tem exatamente s (que é o número de autovalores dois à dois distintos de A) "blocos principais" A_i. Exceto por possíveis permutações de blocos principais (que pode-se fixar, ordenando-se os autovalores de A), a matriz na forma de Jordan, semelhante à A é única. De fato, se

$J = A_1 \oplus A_2 \oplus \ldots \oplus A_s$ e $A_i = J_{p_{i1},\lambda_i} \oplus \ldots \oplus J_{1,\lambda_i}$, $i = 1, \ldots, s$, seja $p_{i1} \leq p_{i2} \leq \ldots \leq p_{in_i}$, as dimensões dos blocos de Jordan, $J_{1,\lambda_i}, J_{2,\lambda_i}, \ldots, J_{n_i,\lambda_i}$, respectivamente. Então a característica de Ségre de A, que determina univocamente J (depois de fixada uma ordem, para os autovalores dois à dois distintos de A) é :

$$\{ (p_{1n_1}, \ldots, p_{11}) (p_{2n_2}, \ldots, p_{21}) \ldots\ldots (p_{sn_s}, \ldots, p_{s1}) \}.$$

Para cada $i = 1, 2, \ldots, s$, temos que :

$$d_i = p_{i1} + p_{i2} + \ldots + p_{in_i}.$$

O número p_{ik}, $k = 1, 2, \ldots, n_i$, nos diz que acima da diagonal de A_i, aparecem $(p_{in_i}-1)$-números 1, depois um zero, depois vem $(p_{in_i-1}-1)$-números 1, seguido de um zero, e assim sucessivamente. Assim sendo, a forma de Jordan fica totalmente determinada pelos números :-

s (números de autovalores distintos dois à dois).

$\lambda_1, \lambda_2, \ldots, \lambda_s$ (autovalores distintos dois à dois).

$d_1, d_2, \ldots, d_s$ (multiplicidades algébricos de $\lambda_1, \lambda_2, \ldots, \lambda_s$).

$$\left.\begin{matrix} p_{11}, & \ldots & , p_{1m_1} \\ \vdots & & \vdots \\ p_{s1}, & \ldots & , p_{sm_s} \end{matrix}\right\} \quad \text{(dimensões dos blocos de Jordan de } A_1, \ldots, A_s \text{).}$$

EXEMPLO 8.

Neste exemplo, procuraremos recordar alguns conceitos de secções anteriores.

Seja J uma matriz na forma de Jordan 12×12, com elementos diagonais, na seguinte ordem

$$2,2,2,2,2,2,0,0,0,0,1,1$$

e cujos elementos imediatamente acima da diagonal são em ordem,

$$1,1,1,0,1,0,1,1,1,0,0.$$

Então temos que :

(i) J, escrita explicitamente é :

$$J = \left[\begin{array}{cccc|cc|cccc|c|c}
2 & 1 & 0 & 0 & & & & & & & & \\
0 & 2 & 1 & 0 & & & & & & & & \\
0 & 0 & 2 & 1 & & & & & & \mathbf{0} & & \\
0 & 0 & 0 & 2 & & & & & & & & \\
\hline
 & & & & 2 & 1 & & & & & & \\
 & & & & 0 & 2 & & & & & & \\
\hline
 & & & & & & 0 & 1 & 0 & 0 & & \\
 & & & & & & 0 & 0 & 1 & 0 & & \\
 & \mathbf{0} & & & & & 0 & 0 & 0 & 1 & & \\
 & & & & & & 0 & 0 & 0 & 0 & & \\
\hline
 & & & & & & & & & & 1 & \\
\hline
 & & & & & & & & & & & 1
\end{array}\right]$$

isto é, $J = A_1 \oplus A_2 \oplus A_3$, onde :

$A_1 = J_{2,2} \oplus J_{1,2}$, dim $J_{2,2} = 4 = p_{12}$ e dim $J_{1,2} = 2 = p_{11}$

$A_2 = J_{1,0}$ e dim $J_{1,0} = 4 = p_{21}$

$A_3 = J_{2,1} \oplus J_{1,1}$, dim $J_{2,1} = $ dim $J_{1,1} = 1 = p_{31} = p_{32}$.

Autovalores ($s = 3$)		Suas multiplicidades algébricas
2	....	$d_1 = 6$.
0	....	$d_2 = 4$.
1	....	$d_3 = 2$.

A característica de Segre é $\{(42)\ (4)\ (11)\}$

O polinômio característico é $c(X) = (X - 2)^6\, X^4\, (X - 1)^2$.

O polinômio mínimo é $m(X) = (X - 2)^4\, X^4\, (X - 1)$.

Os divisores elementares (na ordem dada) são :

$$(X - 2)^4,\ (X - 2)^2,\ X^4,\ (X - 1) \text{ e } (X - 1) \qquad \text{(ver 4.1)}$$

4.5 UM ALGORÍTMO PARA A REDUÇÃO À FORMA CANÔNICA DE JORDAN.

Neste parágrafo, A denota uma matriz $n \times n$, representando um operador linear L, sôbre um espaço vetorial V de dimensão finita n, sôbre um corpo F algèbricamente fechado. Sabemos que A é semelhante à uma matriz J, na forma de Jordan (ver 4.1, 4.3 ou 4.4). Nosso propósito nesta secção, é apresentar um algorítmo, que determina uma forma de Jordan J, para a matriz A.

Começaremos recordando algumas definições e introduzindo outras. Sabemos (ver (35) de 4.1) que existe um inteiro não-negativo a, tal que :

$$\langle 0 \rangle = N(L^0) \subsetneq N(L^1) \subsetneq N(L^2) \subsetneq \dots \subsetneq N(L^a) = N(L^{a+1}) = \dots,$$

ao qual denominaremos ASCENDENTE de L.

Para um escalar λ de F, denotaremos por L_λ o operador linear $L - \lambda I$ de V. O ESPECTRUM de L é o conjunto :

$$\sigma(L) = \{ \lambda \in F : L_\lambda \text{ não é injetor} \},$$

ou seja, $\sigma(L)$ é o conjunto dos autovetores distintos dois à dois de L. Seja $\lambda \in \sigma(L)$, denominamos AUTOVETOR GENERALIZADO de GRAU t (ou simplesmente AUTOVETOR DE GRAU t) associado ao autovalor λ, a um vetor v, não nulo, do subespaço $N(L_\lambda^t)$. Sendo $g \leq t$,

$$N(L^g) \subset N(L^t),$$

isto é, um autovetor de grau g é um autovetor de grau t ($g \leq t$). Seja $v \neq 0$, tal que :

$$v \in N(L_\lambda^t) \quad \text{e} \quad v \notin N(L_\lambda^{t-1}).$$

Neste caso, diz-se que v é um autovetor generalizado de grau ESSENCIALMENTE t. Um autovetor generalizado de grau essencialmente 1, é o que entendemos por um autovetor, no sentido de 1.4.

Seja $S = \{v_1, v_2, \dots, v_r\}$ um conjunto de vetores de V. Denomina-se DIMENSÃO do conjunto S, à dimensão do subespaço gerado por

S, isto é, *define-se* dim $S = \dim \langle S \rangle$.

Seja $v \in V$, tal que $L^{r-1}v \neq o$ e $L^r v = o$, para um inteiro positivo r. O <u>CICLO</u> de v com respeito à L, é definido como sendo o conjunto :-

$$C(L,v) = \{v, Lv, \ldots, L^{r-1}v\}.$$

O vetor $L^{r-1}v$, denomina-se <u>EXTREMO</u> do ciclo. A dimensão de $C(L,v)$ é $\underline{r}$, ou seja, os vetores de $C(L,v)$ são linearmente independentes.

Passemos aos resultados básicos do algorítmo.

<u>LEMA 1.</u>

" Seja $\sigma(L) = \{\lambda_1, \lambda_2, \ldots, \lambda_s\}$ o espectrum de L. Sejam $a_1, \ldots, a_s$ os ascendentes de $L_{\lambda_1}, \ldots, L_{\lambda_s}$, respectivamente. Então, o polinômio : $p(X) = (X-\lambda_1)^{a_1}(X-\lambda_2)^{a_2} \ldots (X-\lambda_s)^{a_s}$, é o polinômio mínimo de L".

Demonstração.

Seja $m(X) = (X-\lambda_1)^{r_1} \ldots (X-\lambda_s)^{r_s}$, o polinômio mínimo de L (esta fatoração é possível, por ser F algèbricamente fechado). É claramente suficiente, mostrarmos que : $r_i = a_i$, $i = 1,2,\ldots,s$.

Inicialmente mostraremos que $r_i \geq a_i$, $i = 1,2,\ldots,s$. Suponhamos, por absurdo, que exista um índice i em $\{1,2,\ldots,s\}$, tal que, $r_i < a_i$. Então, pela definição de ascendente a_i de $L_{\lambda_i} = L - \lambda_i I$, existe um vetor $v \neq o$, tal que :

$$L_{\lambda_i}^{r_i+1} v = o \quad \text{e} \quad w = L_{\lambda_i}^{r_i} v \neq o.$$

Logo, $L_{\lambda_i}^{r_i+1} v = L_{\lambda_i}(L_{\lambda_i}^{r_i} v) = L_{\lambda_i} w = o$ ou $Lw = \lambda_i w$.

O polinômio mínimo m de L, pode ser colocado na forma :

$$m(X) = q(X)\,(X-\lambda_i)^{r_i}, \text{ com } q(\lambda_i) \neq o,$$

pois λ_i, é uma raiz de multiplicidade r_i de m. De $Lw = \lambda_i w$, vem que, $q(L)w = q(\lambda_i)w \neq o$, pois o escalar $q(\lambda_i) \neq o$ e o vetor $w \neq o$. Ademais, $m(L)v = q(L)(L-\lambda_i I)^{r_i}v = q(L)w = q(\lambda_i)w \neq o$, o que é absurdo, pois $m(L) = O$. Logo, $r_i \geq a_i$, $\forall i \in \{1,2,\ldots,s\}$.

Resta mostrarmos que $a_i \geq r_i$, $i = 1,2,\ldots,s$.

Seja $\{e_1, e_2, \ldots, e_n\}$ uma base de V. Então, os $(n+1)$-vetores $e_i, Le_i, \ldots, L^n e_i$, são linearmente dependentes, $i = 1,2,\ldots,s$. Logo, existe um polinômio não nulo f_i ($gr(f_i) \leq n$), tal que $f_i(L)e_i = o$. Seja

$$g = f_1 f_2 \dots f_n .$$

Então,

$$g(L)e_i = 0, \ i = 1,2, \dots ,s, \text{ ou seja } g(L) = 0.$$

Pela definição de polinômio mínimo, temos que m divide g. Logo, $\lambda_1, \lambda_2, \dots , \lambda_s$, são raízes de g. Como o corpo F é algèbricamente fechado, podemos fatorar g na forma :-

$$g(X) = a \prod_{i=1}^{k} (X - \lambda_i)^{b_i}, \ (k \geqslant s).$$

Para todo λ_i, com $i > s$, isto é, $\lambda_i \notin \sigma(L)$, se $(L - \lambda_i I)v = 0$, então $v = 0$. Assim, o produto $g_1(X) = \prod_{i=1}^{s} (X - \lambda_i)^{b_i}$, dos fatores de g, com $\lambda_i \in \sigma(L)$, ainda satisfaz $g_1(L) = 0$. Pela mesma razão, o produto $g_2(X) = \prod_{i=1}^{s} (X - \lambda_i)^{c_i}$, com $\lambda_i \in \sigma(L)$ e $c_i = \min \{a_i, b_i\}$ (onde a_i é o ascendente de L_{λ_i}), satifaz $g_2(L) = 0$. Ademais, como $p(X) = (X - \lambda_1)^{a_1} \dots (X - \lambda_s)^{a_s}$, tem cada $\lambda_i \in \sigma(L)$, como um zero de multiplicidade $a_i \geqslant c_i$, $i = 1,2, \dots ,s$, g_2 divide p. Isto é, existe um $h \in F[X]$ tal que : $p = hg_2$. Então,

$$p(L) = h(L)g_2(L) = h(L)0 = 0.$$

Da definição do polinômio mínimo m de L, temos que m divide p, ou seja, $r_i \leqslant a_i$, $i = 1,2, \dots, s$. c.q.d.

Do lema 1 e do teorema da Decomposição Primária (teorema 4 de 2.3), obtemos nosso próximo teorema.

<u>TEOREMA 8.</u>

"Seja $\sigma(L) = \{\lambda_1, \dots , \lambda_s\}$ e a_i o ascendente de L_{λ_i}, $i = 1,2, \dots ,s$. Então,

$$V = N(L_{\lambda_1}^{a_1}) \oplus N(L_{\lambda_2}^{a_2}) \oplus \dots \oplus N(L_{\lambda_s}^{a_s}) \text{ ".}$$

<u>TEOREMA 9.</u>

"Seja $\lambda \in \sigma(L)$ e <u>a</u> (a ser determinado) o ascendente de L_λ. Então, existem <u>a</u> conjuntos,

$$E_i = \{v_{ij} : j = 1,2, \dots, s_i\}, \ i = 1,2, \dots, a,$$

tais que :

(i) Cada v_{ij} é um autovetor (generalizado) de grau essencialmente i, para $j = 1,2,\dots,s_i$;

(ii) $s_1 = \dim N(L_\lambda)$, isto é, s_1 é a multiplicidade geométrica de λ. Vale a seguinte fórmula de recorrência, para a obtenção de s_i, para $i = 2,3, \dots,a$:

$$s_i = \dim N(L_\lambda^i) - \dim N(L_\lambda^{i-1}) = \dim N(L_\lambda^i) - s_{i-1};$$

(iii) $E_1 \cup E_2 \cup \ldots \cup E_a$ é um conjunto linearmente independente;

(iv) $N(L_\lambda^b) = \langle E_1 \cup \ldots \cup E_b \rangle$, para $b = 1, 2, \ldots, a$".

Demonstração.

Seja $\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ uma base de V. Seja $\mathcal{B}_1 = L_\lambda(\mathcal{B}) = \{L_\lambda e_1, L_\lambda e_2, \ldots, L_\lambda e_n\}$. Como $\lambda \in \sigma(L)$, temos que $\dim(\mathcal{B}_1) < \dim(\mathcal{B})$. Consideremos :-

$$s_1 = \dim(\mathcal{B}) - \dim(\mathcal{B}_1) > 0.$$

Então, temos s_1-vetores em $\mathcal{B}_1$ que podem ser escritos como combinação linear dos restantes vetores de $\mathcal{B}_1$. Renumerando-se a base (se necessário for), não há perda de generalidade, supormos que :

$\{L e_{s_1+1}, \ldots, L e_n\}$, é linearmente independente e que :

$$(41) \quad \begin{cases} L e_1 = a_{1,s_1+1} L e_{s_1+1} + \ldots + a_{1,n} L e_n ; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ L e_{s_1} = a_{s_1,s_1+1} L e_{s_1+1} + \ldots + a_{s_1,n} L e_n . \end{cases}$$

Definamos agora, o conjunto E_1 dos vetores :-

$$(42) \quad v_{1j} = e_j - a_{j,s_1+1} e_{s_1+1} - \ldots - a_{j,n} e_n \ ; \ j = 1, 2, \ldots, s_1 ;$$

cujos coeficientes são dados em (41). Queremos mostrar que cada vetor v_{1j} de E_1 é um autovetor generalizado de grau essencialmente 1, associado ao autovalor λ. Isto é, queremos mostrar que cada v_{1j}, satisfaz $Lv_{1j} = \lambda v_{1j}$. Aplicando-se L_λ à (42) e levando-se em consideração (41), obtemos :-

$$L_\lambda v_{1j} = Le_j - Le_j = 0 \Rightarrow (L - \lambda I)v_{1j} = 0 \Rightarrow Lv_{1j} = \lambda v_{1j} .$$

Ademais, de (42), é óbvio que $E_1 \cup \{e_{s_1+1}, \ldots, e_n\}$ é um conjunto com n-vetores linearmente independentes de V, logo formam uma base para V.

Ficaram estabelecidas as partes (i), (ii) e (iii), para o caso i = 1. Vejamos (iv). Seja :-

$$(43) \quad v = c_1 v_{11} + \ldots + c_{s_1} v_{1s_1} + c_{s_1+1} e_{s_1+1} + \ldots + c_n e_n ,$$

um vetor arbitrário de $N(L_\lambda)$. Apliquemos L_λ à (43), obtemos :

$$0 = L_\lambda v = c_1 L_\lambda v_{11} + \ldots + c_{s_1} L_\lambda v_{1s_1} + c_{s_1+1} L_\lambda e_{s_1+1} + \ldots + c_n L_\lambda e_n ,$$

e como $L_\lambda v_{1j} = 0$, $j = 1,2, \ldots, s_1$, temos que :

$c_{s_1+1} L_\lambda e_{s_1+1} + \ldots + c_n L_\lambda e_n = 0$,

e como $\{L_\lambda e_{s_1+1}, \ldots, L_\lambda e_n\}$ é linearmente independente, temos que $c_{s_1+1} = \ldots = c_n = 0$ e (43) se reduz à :

$$v = c_1 v_{11} + \ldots + c_{s_1} v_{1s_1} ,$$

ou seja, $v \in \langle E_1 \rangle$. Logo $N(L_\lambda) = \langle E_1 \rangle$ e vale (iv), para $i = 1$.

Seja $\mathcal{B}_2 = L_\lambda^2(\mathcal{B}_1) = \{L_\lambda^2 e_{s_1+1}, \ldots, L_\lambda^2 e_n\}$. Caso $\dim \mathcal{B}_2 = \dim \mathcal{B}_1$, ótimo. Pois $a = 1$ e a demonstração fica concluida. Caso contrário, seja :-

$$s_2 = \dim(\mathcal{B}_1) - \dim(\mathcal{B}_2) > 0.$$

Daí existem s_2-vetores em $\mathcal{B}_2$ que são combinação linear dos restantes vetores de $\mathcal{B}_2$. Não há perda de generalidade, ao supormos que :

$\{L^2 e_{s_1+s_2+1}, \ldots, L^2 e_n\}$ são linearmente independentes e que:

$$\begin{cases} L^2 e_{s_1+1} = b_{1,s_1+s_2+1} L^2 e_{s_1+s_2+1} + \ldots + b_{1,n} L^2 e_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ L^2 e_{s_1+s_2} = b_{s_2,s_1+s_2+1} L^2 e_{s_1+s_2+1} + \ldots + b_{s_2,n} L^2 e_n. \end{cases}$$

Definimos agora, o conjunto E_2 dos vetores :-

(44) $v_{2j} = e_{s_1+j} - b_{j,s_1+s_2+1} e_{s_1+s_2+1} - \ldots - b_{j,n} e_n$,

para $j = 1,2, \ldots, s_2$. Anàlogamente ao que fizemos com os vetores de E_1, podemos concluir que :- $E_2 = \{v_{2j} ; j = 1,2, \ldots, s_2\}$ é um conjunto de s_2-autovetores de grau essencialmente 2 e linearmente independentes. De (42) e (44), concluimos que $E_1 \cup E_2 \cup \{e_{s_1+s_2+1}, \ldots, e_n\}$ é um conjunto de n vetores linearmente independentes de V, logo formam uma base de V.

Ademais, do mesmo modo que mostramos que um vetor de $N(L_\lambda)$ está em $\langle E_1 \rangle$, pode-se mostrar que todo vetor de $N(L_\lambda^2)$ está em $\langle E_1 \cup E_2 \rangle$. Daí $N(L_\lambda^2) = \langle E_1 \cup E_2 \rangle$. É trivial que :

$$s_2 = \dim N(L_\lambda^2) - \dim N(L_\lambda) = \dim N(L_\lambda^2) - s_1.$$

Fica assim demonstrado o teorema para $a = 2$.

Como dim $V = n$, finita, êste procedimento sendo repetido um número finito de vêzes (êsse número é $\underline{a}$), demonstra-se o teorema.

TEOREMA 10.

" Nas mesmas notações do teorema 9, seja :-

$d \leq c \leq \dots \leq a$, $w \in E_d$, $x \in E_c$, $\dots$, $z \in E_a$. Suponhamos que a união dos ciclos :

$C(L_\lambda, x) \cup \dots \cup C(L_\lambda, z)$, é linearmente independente, e que :-

$C(L_\lambda, w) \cup C(L_\lambda, x) \cup \dots \cup C(L_\lambda, z)$ é linearmente dependente.

Então, o extremo $(L_\lambda^{d-1} w)$ de $C(L_\lambda, w)$ é uma combinação linear dos extremos de $C(L_\lambda, x)$, $\dots$, $C(L_\lambda, z)$, isto é :-

(45) $L_\lambda^{d-1} w = \alpha_{c-1} L_\lambda^{c-1} x + \dots + \beta_{a-1} L_\lambda^{a-1} z$ ".

Demonstração.

Sendo $C(L_\lambda, w) \cup C(L_\lambda, x) \cup \dots \cup C(L_\lambda, z)$ linearmente dependente, existem escalares, não todos nulos,

$f_0, \dots, f_{d-1}, g_0, \dots, g_{c-1}, \dots\dots, h_0, \dots, h_{a-1}$

tais que :-

(46) $f_0 w + \dots + f_{d-1} L_\lambda^{d-1} w + g_0 x + \dots + g_{c-1} L_\lambda^{c-1} x + \dots + h_0 z + \dots + h_{a-1} L_\lambda^{a-1} z = 0.$

Digamos que f_k é o primeiro escalar não nulo. Depois de desconsiderarmos os $(k-1)$-primeiros vetore de (46) (caso $k > 0$), aplicamos L_λ^{d-k-1} à (46), obtendo :-

(47) $f_k L_\lambda^{d-1} w + g_0 L_\lambda^{d-k-1} x + \dots + g_{c-d+k} L_\lambda^{c-1} x + \dots + h_0 L_\lambda^{d-k-1} z + \dots + h_{a-d+k} L_\lambda^{a-1} z$

pois por hipótese $d \leq c \leq \dots \leq a$. Daí,

$L_\lambda^{d-1} w \in \langle C(L_\lambda, x) \cup \dots \cup C(L_\lambda, z) \rangle$.

Reescreve-se (47) na seguinte forma :-

(48) $L_\lambda^{d-1} w = \alpha_0 x + \dots + \alpha_{c-1} L_\lambda^{c-1} x + \dots\dots + \beta_0 z + \dots + \beta_{a-1} L_\lambda^{a-1} z.$

Aplicando-se L_λ à (48), e lembrando que $L^d w = L^c x = \dots = L^a z = 0$, obtemos:

$0 = \alpha_0 L_\lambda x + \dots + \alpha_{c-2} L_\lambda^{c-1} x + \dots + \beta_0 L_\lambda z + \dots + \beta_{a-2} L_\lambda^{a-1} z,$

e como por hipótese, $C(L_\lambda, x) \cup \dots \cup C(L_\lambda, z)$ é linearmente independente, temos que :-

$\alpha_0 = \dots = \alpha_{c-2} = \dots\dots \beta_0 = \dots = \beta_{a-2} = 0,$

e (48) se reduz à (45), o que demonstra o teorema.

TEOREMA 11.

" Com as mesmas notações dos teoremas 9 e 1o, supondo-se que (45) seja válido, têm-se que :-

$$\langle E_1 \cup \ldots \cup E_a \rangle = \langle E_1 \cup \ldots \cup E_a - [\{w\} \cup C(L_\lambda, x) \cup \ldots \cup C(L_\lambda, z)] \rangle".$$

Demonstração.

Consideremos o vetor v, tal que :-

$$(49) \qquad w = \alpha_{c-1} L_\lambda^{c-d} x + \ldots + \beta_{a-1} L_\lambda^{a-d} z + v .$$

Por hipótese, $L_\lambda^{d-1} v = o$ (aplique L_λ^{d-1} à (49), que obterá (45), só se $L_\lambda^{d-1} v = o$). Portanto, pelo teorema 9, $v \in \langle E_1 \cup \ldots \cup E_{d-1} \rangle$.

Assim,

$$\langle E_1 \cup \ldots \cup E_d \rangle = \langle E_1 \cup \ldots \cup E_d - [\{w\} \cup \{w - v\}] \rangle ,$$

para $w \in E_d$, $v \in \langle E_1 \cup \ldots \cup E_{d-1} \rangle$ e $E_1 \cup \ldots \cup E_d$ é um conjunto linearmente independente. De (49),

$$(w - v) \in \langle C(L_\lambda, x) \cup \ldots \cup C(L_\lambda, z) \rangle \text{ . Logo,}$$

$$N(L_\lambda^a) = \langle E_1 \cup \ldots \cup E_a \rangle = \langle E_1 \cup \ldots \cup E_a - [\{w\} \cup \{w - v\}] \rangle \subseteq$$

$$\subseteq \langle E_1 \cup \ldots \cup E_a - [\{w\} \cup C(L_\lambda, x) \cup \ldots \cup C(L_\lambda, z)] \rangle \subseteq N(L_\lambda^a),$$

c.q.d.

Agora estamos em condições de descrever as etapas do algorítimo, sendo que conservaremos as notações do lema 1 e teoremas 9, 1o e11.

Etapa o. Determina-se o spectrum $\sigma(L)$ e fixa-se uma ordem para seus elementos, digamos :

$$\sigma(L) = \{ \lambda_1, \lambda_2, \ldots, \lambda_s \} .$$

Etapa 1. Seja i = 1. Determinamos um conjunto $E_1 \cup \ldots \cup E_{a_i}$ (a_i é o ascendente de L_{λ_i}) de autovetores generalizados de L, linearmente independentes que geram $N(L_{\lambda_i}^{a_i})$ (ver demonstração do teorema 9). Além disso, também se determina o ascendente a_i.

Etapa 2. Fixa-se um vetor v_i em E_{a_i} e forma-se o ciclo $C(L_{\lambda_i}, v_i) = \mathcal{B}_i$

Etapa 3. Seja $t = \dim N(L_{\lambda_i}^{a_i}) - \dim \mathcal{B}_i$. Quando t = o, vamos para a etapa 4. Caso contrário, forma-se o conjunto $E_t = \{w_{t1}, \ldots, w_{tj}\}$.

<u>Etapa 3a.</u> Para k=1, verificamos se $L_{\lambda_i}^{t-1} w_{tk}$ é uma combinação linear dos extremos que ocorrem em $\mathcal{B}$. Em verificação afirmativa, desprezemos w_{tk}, não alteramos $\mathcal{B}$ e vamos para etapa 3b. Caso contrário, formamos o ciclo $C(L_{\lambda_i}, w_{tk})$ e substituimos $\mathcal{B}$ por $\mathcal{B} \cup C(L_{\lambda_i}, w_{tk})$.

<u>Etapa 3b.</u> Repete-se a etapa 3a, para k = 2,3, ... , j.

<u>Etapa 4.</u> Repete-se a etapa 3, até que dim $\mathcal{B}$ = dim $N(L_{\lambda_i}^{a_i})$.

<u>Etapa 5.</u> Seja i = i+1 e se $i \leq s$, voltamos à etapa 1 (isto é, repete-se as etapas 1,2,3 e4 para os autovalores restantes de L). Caso $i > s$, terminamos.

Realizadas estas etapas, estamos em condições de descrever uma forma de Jordan para a matriz A. Antes faremos um "diagrama em blocos" para o algorítimo, depois daremos um exemplo de uso do algorítmo, para encerrarmos com algumas observações gerais.

O diagrama em blocos está na página seguinte.

Visto o diagrama, vamos a um exemplo, no qual procuraremos ilustrar o procedimento.

<u>EXEMPLO 9.</u>

Seja F = C e

$$A = \begin{bmatrix} 2 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

representando um operador linear L de C^7, relativamente à base ordenada canônica :-

$$\mathcal{C} = \{e_i = (\delta_{i1}, \ldots, \delta_{i7}) \; ; \; i = 1,2, \ldots, 7\} .$$

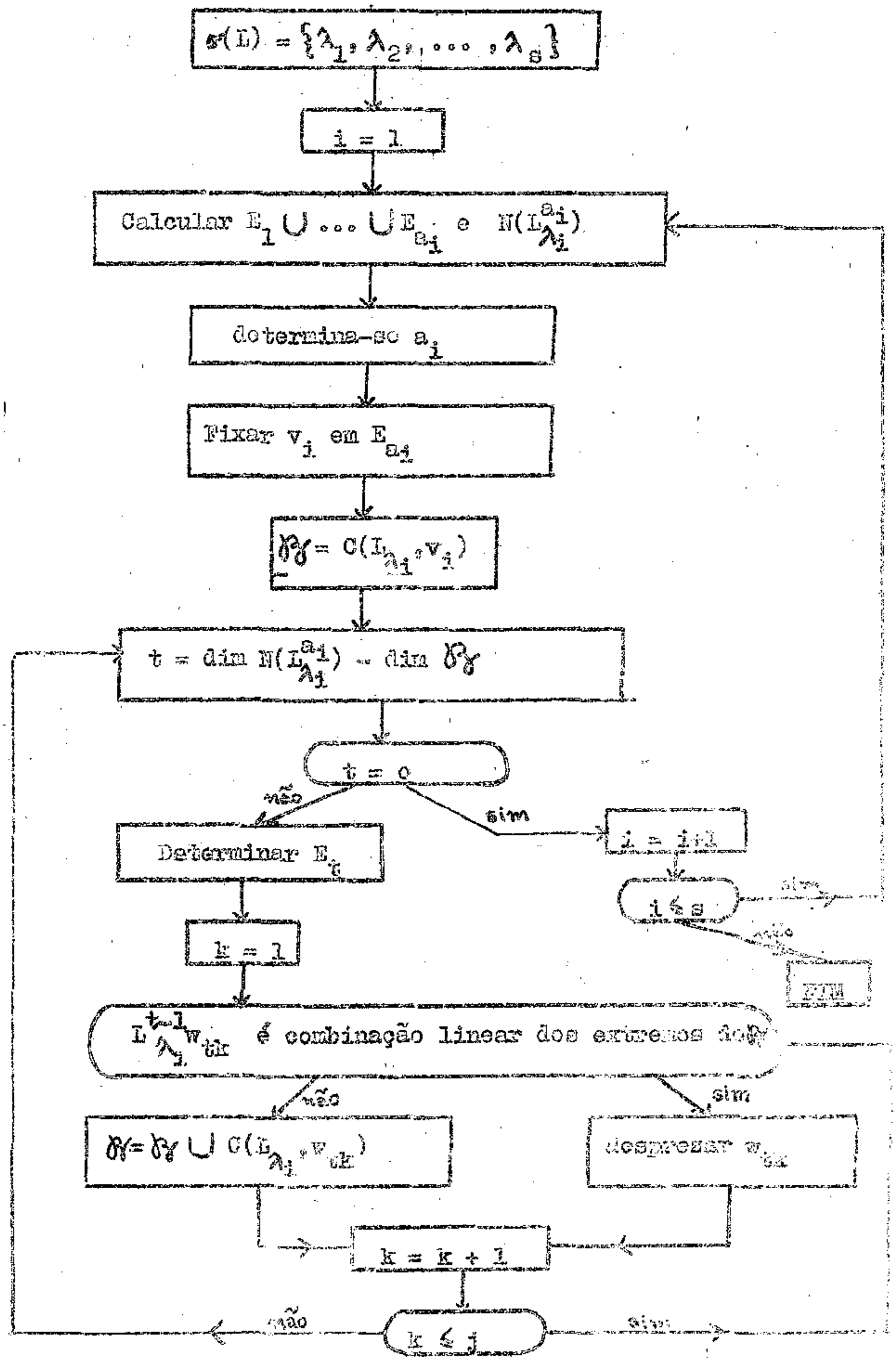
<u>Etapa 0.</u>

O polinômio característico de L (ou de A) é :

$$c(X) = (X - 2)^6 (X - 1),$$

Logo $\sigma(L) = \{\lambda_1 = 2, \lambda_2 = 1\}$

<u>Etapa 1.</u> Para i = 1 ($\lambda_1 = 2$). Aplica-se $L_2 = L - 2I$ à base canônica $\mathcal{C}$. Por razões de simplicidade, usaremos a notação auto-evidente seguinte (para indicar $[e_i]$ $[A - 2E]$) :-


σ(L) = {λ₁, λ₂, ..., λₛ}
i = 1
Calcular E₁ ∪ ... ∪ E_{aᵢ} e N(L^{aᵢ}_{λᵢ})
determina-se aᵢ
Fixar vᵢ em E_{aᵢ}
𝒱 = C(L_{λᵢ}, vᵢ)
t = dim N(L^{aᵢ}_{λᵢ}) - dim 𝒱
t = o
não
sim
i = i+1
i ≤ s
sim
não
FIM
Determinar Eₜ
k = 1
L^{t-1}_{λᵢ} w_{tk} é combinação linear dos extremos de 𝒱
não
sim
𝒱 = 𝒱 ∪ C(L_{λᵢ}, w_{tk})
desprezar w_{tk}
k = k + 1
k ≤ j
não
sim

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

Daí, vem que $E_1 = \{e_4-e_3, e_5, e_6\}$ e que $L_2(e_4-e_3) = L_2e_5 = L_2e_6 = 0$ (E_1 é o conjunto de todos os autovetores de L associados ao autovalor 2, ou seja, o autoespaço $V_2 = \langle E_1 \rangle$).

Assim sendo, os vetores linearmente independentes de $\mathcal{B}_1 = L_2(\mathcal{B})$ são :-

$\{(0 \ 0 \ -1 \ 0 \ 1 \ 0 \ 0), (0 \ 0 \ 0 \ 1 \ 1 \ 0 \ 0), (0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0), (0 \ 0 \ 0 \ 0 \ 0 \ 0 \ -1)\}$.

Aplicando-se L_2 à base $\mathcal{B}_1$, vem :-

$$\begin{pmatrix} 0 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix} \begin{bmatrix} 0 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Logo, $E_2 = \{e_2+e_1, e_3\}$. Então, os vetores linearmente independentes de $\mathcal{B}_2 = L_2^2(\mathcal{B})$ são :-

$\{(0 \ 0 \ 0 \ 0 \ -1 \ 0 \ 0), (0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1)\}$.

Seguindo vem :-

$$\begin{pmatrix} 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} 0 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

Daí $E_3 = \{e_1\}$. É óbvio que ;

$(0 \ 0 \ 0 \ 0 \ 0 \ 0 \ -1)(A - 2E) = (0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1)$.

Então o ascendente de $L_2 = L - 2I$ é $a_1 = 3$ e

$E_1 \cup E_2 \cup E_3 = \{e_4 - e_3, e_5, e_6, e_2 + e_1, e_3, e_1\}$;

formam uma base para $N(L_2^3)$ (esta base não é a que queríamos!).

<u>Etapa 2.</u> Fixemos $v_1 = e_1$ em E_3 e tomomos:
$$\mathcal{W} = C(L_2, e_1) = \{e_1, L_2 e_1, L_2^2 e_1\} = \{e_1, -e_3+e_5, -e_5\},$$
pois $L_2^3 e_1 = 0$.

<u>Etapa 3.</u>
$$t = \dim N(L_2^3) - \dim \mathcal{W} = 6 - 3 = 3.$$

Como não temos nenhum autovetor generalizado de grau 3, consideremos autovetores de grau 2. Fixemos $v_2 = e_3$ em E_2 e como o extremo de $C(L_2, e_2) = \{e_2, L_2 e_2\} = \{e_3, e_5\}$ é um múltiplo escalar do extremo de $C(L_2, e_1)$ ($e_5 = -1(-e_5)$), não alteramos $\mathcal{W}$. Por outro lado, o extremo de $C(L_2, e_2+e_1) = \{e_2+e_1, -e_3+e_4+2e_5\}$ não é um múltiplo escalar do extremo de $C(L_2, e_1)$. Então substituímos $\mathcal{W}$ por
$$\mathcal{W} \cup C(L_2, e_2+e_1) = \{e_1, -e_3+e_5, -e_5, e_2+e_1, -e_3+e_4+2e_5\}.$$

<u>Etapa 4.</u>

Agora $\dim N(L_2^3) - \dim \mathcal{W} = 1$ (aqui $\mathcal{W}$ já é o novo). Como e_5 e e_4-e_3 são combinações lineares dos extremos de $\mathcal{W}$ e e_6 não o é, substituímos $\mathcal{W}$ por $\mathcal{W} \cup C(L_2, e_6) = \{e_1, -e_3+e_5, -e_5, e_2+e_1, -e_3+e_4+2e_5, e_6\}$.

Agora, $\dim N(L_2^3) - \dim \mathcal{W} = 0$.

<u>Etapa 5.</u> Para $\lambda_2 = 1$, não há problema (pois a multiplicidade algébrica é 1) e temos :-
$$(0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 1)\ (A - 1I) = (0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0),$$
Logo substituindo-se $\mathcal{W}$ por :
$$\mathcal{W} = \{e_1, -e_3+e_5, -e_5, e_2+e_1, -e_3+e_4+2e_5, e_6, e_7\},$$
cuja ordem fixamos. Relativamente à esta (nova) base ordenada $\mathcal{W}$ de V, a matriz de L está na forma de Jordan :-
$$J = \left[\begin{array}{ccc|cc|c|c} 2 & 1 & 0 & & & & \\ 0 & 2 & 1 & & & & \\ 0 & 0 & 2 & & & & \\ \hline & & & 2 & 1 & & \\ & & & 0 & 2 & & \\ \hline & & & & & 2 & \\ \hline & & & & & & 1 \end{array}\right]$$

Faremos agora alguns comentários do que obtivemos. Pelo teorema 10, na etapa 3 o extremo de qualquer ciclo $C(L_{\lambda_1}, v_1)$ para $v_1 \in E_r$, com $r > t$, é linearmente dependente com os extremos de $\mathcal{W}$, e então pelo teorema 11, v_1 é substituído na nova base. Terminada a etapa 4, obtemos uma nova base $\mathcal{B}_{\lambda_1}$ para $N(L_{\lambda_1}^{e_1})$ consistindo de uma união de ciclos. Fica determinado um BLOCO DE JORDAN relativo ao autovalor λ_1.

Completada a etapa 5, temos tôdas as bases $\mathcal{B}_{\lambda_i}$, reunião de ciclos para $i = 1,2, \ldots ,s$, ou seja, todos os blocos de Jordan. Pelo teorema 6,

$V = N(L_{\lambda_1}^{e_1}) \oplus \ldots \oplus N(L_{\lambda_s}^{e_s})$, logo pelo teorema 1 de 2.2, a união destas bases $\mathcal{B}_{\lambda_1} \cup \ldots \cup \mathcal{B}_{\lambda_s}$ forma uma base de V.

Relativamente à esta base, a matriz que representa L está na forma de Jordan. A tais bases costuma-se chamar, BASE DE JORDAN para L.

Queremos ressaltar, que êste algorítmo, nos fornece também, além da forma de Jordan,

(i) Os expoentes e_i, de L_{λ_i}, $i = 1,2, \ldots ,s$.

(ii) O polinômio mínimo de L, que é :

$$m(X) = (X - \lambda_1)^{e_1} \ldots (X - \lambda_s)^{e_s}.$$

(iii) Uma base de Jordan para L.

Além disso, podemos concluir que :-

"L é diagonalizável se, e sòmente se, o expoente de L_{λ_i} é 1 para $i = 1,2, \ldots ,s$". (ver teorema de diagonalização, teorema 3 de 2.4).

4.6 ALGUNS MÉTODOS DE DETERMINAÇÃO DA MATRIZ DE TRANSFORMAÇÃO À FORMA CANÔNICA DE JORDAN.

Dada uma matriz A, $n \times n$ com elementos num corpo algèbricamente fechado F, denomina-se MATRIZ DE TRANSFORMAÇÃO de A à forma de Jordan J, a tôda matriz inversível P de $M_n(F)$, tal que ;

$$J = P^{-1} A P.$$

MÉTODO 1.

Seja A uma matriz de $M_n(F)$, representando um operador linear L sôbre um F-espaço vetorial V, relativamente à uma base ordenada $\mathcal{C}$ de V. Isto é, seja $A = [L]_{\mathcal{C}}$.

Suponhamos que é possível encontrarmos uma "base de Jordan" para L, ou seja, uma base ordenada $\mathcal{J}$ de V, tal que :-

$$[L]_{\mathcal{J}} = J.$$

Então, uma matriz de transformação, é a matriz P, que muda as coordenadas de um vetor na base ordenada $\mathcal{B}_c$ para suas coordenadas na base ordenada $\mathcal{B}_\gamma$. De fato, para esta P, temos que (ver 1.1) :-

$$J = [L]_{\mathcal{B}_\gamma} = P^{-1}\,[L]_{\mathcal{B}_c}\,P = P^{-1}\;A\;P.$$

Assim sendo, tôda redução de uma matriz A à sua forma de Jordan que dá uma base de Jordan, nos dá também uma matriz de transformação. Tal foi o caso na seção 4.5. Para o exemplo 9 de 4.5, relativamente as bases canônica $\mathcal{B}_c$ e

$\mathcal{B}_\gamma = \{e_1, -e_3+e_5, -e_5, e_2+e_1, -e_3+e_4+2e_5, e_6, e_7\}$, temos que,

$A = [L]_{\mathcal{B}_c}$ e $J = [L]_{\mathcal{B}_\gamma}$ e como (no [illegible]

$= \{f_1, f_2, \dots, f_7\}$) :

$$\begin{aligned}
f_1 &= e_1 &&= (1\ 0\ 0\ 0\ 0\ 0\ 0)\\
f_2 &= -e_3+e_5 &&= (0\ 0\ -1\ 0\ 1\ 0\ 0)\\
f_3 &= -e_5 &&= (0\ 0\ 0\ 0\ -1\ 0\ 0)\\
f_4 &= e_1+e_2 &&= (1\ 1\ 0\ 0\ 0\ 0\ 0)\\
f_5 &= -e_3+e_4+2e_5 &&= (0\ 0\ -1\ 1\ 2\ 0\ 0)\\
f_6 &= e_6 &&= (0\ 0\ 0\ 0\ 0\ 1\ 0)\\
f_7 &= e_7 &&= (0\ 0\ 0\ 0\ 0\ 0\ 1).
\end{aligned}$$

Logo a matriz P (mudança de base), é uma matriz de transformação de A para J, e

$$P = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0 & -1 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 1 & -1 & 0 & 2 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

MÉTODO 2.

Êste método, embora fácil de ser concluído (teòricamente), não é útil na prática, desde que envolve uma grande quantidade de cálculos. Dada a matriz A, de M_n(·) e conhecendo-se uma matriz J na forma de Jordan (ou na forma Racional), a equação matricial :-

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

cuja incógnita é a matriz P, pode ser colocada na forma :-

$$PJ - AP = 0.$$

Esta equação matricial em P, é equivalente a um sistema de n^2-equações lineares e homogêneas nos n^2 elementos da matriz P. Consequentemente, determina-se P, resolvendo-se tal sistema de n^2-equações à n^2-incógnitas, o que pode envolver uma grande quantidade de cálculos, mesmo para pequenos valores de n. Ainda mais, devemos escolher as soluções de modo que $\det P \neq 0$. A existência de uma tal solução é garantida pelo fato de que A e J possuem os mesmos polinômios invariantes, logo são semelhantes (ver teorema de 3.4), ou seja, existe uma matriz inversível P, de $M_n(C)$, tal que:

$$J = P^{-1} A P .$$

Ressaltamos aqui, que embora a forma de Jordan J, é unìvocamente determinada por A (a menos da ordem dos autovalores de A), a matriz de transformação P, não é única. De fato, seja B uma matriz de $M_n(C)$, tal que:-

(51) $\qquad A B = B A \quad$ e $\det B \neq 0$.

(existe uma tal B, por exemplo, $B = cE$, $c \neq 0$). Então a toda inversível $P_1 = BP$ é tal que :-

$$P_1^{-1} A P_1 = (BP)^{-1} A (BP) = P^{-1} (B^{-1} A B) P \overset{(51)}{=} P^{-1}(B^{-1}BA)P =$$
$$= P^{-1} A P = J.$$

Consequentemente, P_1 é também uma matriz de transformação de A à forma de Jordan J.

<u>EXEMPLO 12</u>.

Seja

$$A = \begin{bmatrix} -4 & -9 \\ 4 & 8 \end{bmatrix} .$$

Então $c(X) = (X - 2)^2$ e como $A \neq 2E$, $m(X) = (X - 2)^2$ e

$$J = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \qquad \text{(ver exemplo 6 de 4.3).}$$

Queremos encontrar uma matriz inversível P,

$$P = \begin{bmatrix} x & y \\ z & t \end{bmatrix} , \text{ tal que :-}$$

$$PJ - AP = 0 \qquad \text{ou} \qquad J = P^{-1} A P .$$

Temos assim o seguinte sistema :-

$$\begin{bmatrix} x & y \\ z & t \end{bmatrix} \begin{bmatrix} 2 & 1 \\ o & 2 \end{bmatrix} - \begin{bmatrix} -4 & -9 \\ 4 & 8 \end{bmatrix} \begin{bmatrix} x & y \\ z & t \end{bmatrix} = \begin{bmatrix} o & o \\ o & o \end{bmatrix}$$

ou $\begin{bmatrix} 6x+9z & x+6y+9t \\ -4x-6z & -4y+z-6t \end{bmatrix} = \begin{bmatrix} o & o \\ o & o \end{bmatrix}$ ou $\begin{cases} x = -3/2\ z \\ 4y+6t = z \end{cases}$.

Uma solução do sistema, tal que $\det P \neq o$, é dada por :-
$x = -3$, $y = -1$, $z = 2$ e $t = 1$.

Assim,

$$P = \begin{bmatrix} -3 & -1 \\ 2 & 1 \end{bmatrix}$$

Verificação :

$$P^{-1} = \begin{bmatrix} -1 & -1 \\ 2 & 3 \end{bmatrix}$$

e $P^{-1} A\, P = \begin{bmatrix} -1 & -1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} -4 & -9 \\ 4 & 8 \end{bmatrix} \begin{bmatrix} -3 & -1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} -6 & -5 \\ 4 & 4 \end{bmatrix} =$

$= \begin{bmatrix} 2 & 1 \\ o & 2 \end{bmatrix} = J$.

<u>MÉTODO 3</u>.

Êste método tem por base o corolário 5 de 3.4, que garante que a matriz de transformação P, tal que :-

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

é dada por :-

$$P = Q(J),$$

onde a matriz $Q(X)$ é uma matriz polinomial inversível (isto é, com determinante constante e não nulo) obtida da equação que dá a "equivalência" (ver 3.1) entre as equações características das matrizes A e J. Mais precisamente, Q é obtida da equação :-

$$XE - J = R(X)(XE - A)Q(X) .$$

Sendo que $Q(J)$, denota o valor à direita na substituição da indeterminada X de $Q(X)$, pela matriz J.

Para encontrarmos $Q(X)$, reduzimos as matrizes características $XE - A$ e $XE - J$, à forma canônica diagonal (14) do teorema 3 de 3.3, por meio de transformações elementares, digamos que :-

(52) $\quad D = \left[i_1(X) , \ldots , i_n(X) \right] = R_1(X)\ (XE - A)\ Q_1(X)$

(53) $\quad D = \left[i_1(X) , \ldots , i_n(X) \right] = R_2(X)\ (XE - J)\ Q_2(X)$.

onde :-

(54) $Q_1(X) = T_1\, T_2 \cdots T_{r_1}$ e $Q_2(X) = T'_1\, T'_2 \cdots T'_{r_2}$,

sendo $T_1, T_2, \ldots, T_{r_1}$ (ou $T'_1, T'_2, \ldots, T'_{r_2}$) as matrizes elementares (ver 3.1) correspondentes as operações elementares da redução da matriz polinomial $XE - A$ (ou $XE - J$) à D. De (52) e (53) vem

$$XE - J = R_2^{-1}(X)\, D\, Q_2^{-1}(X) = R_2^{-1}(X)\, R_1(X)\, (XE - A)\, Q_1(X)\, Q_2^{-1}(X) =$$

$$= R(X)\, (XE - A)\, Q(X).$$

Daí e de (54), resulta que :-

(55) $Q(X) = Q_1(X)\, Q_2^{-1}(X) = T_1\, T_2 \cdots T_{r_1} \left[T'_{r_2}\right]^{-1} \cdots \left[T'_1\right]^{-1}$.

Assim calcula-se $Q(X)$, aplicando-se sucessivamente, às colunas da matriz unidade E, as operações elementares dadas, respectivamente, pelas matrizes

$$T_1, T_2, \ldots, T_{r_1}, \left[T'_{r_2}\right]^{-1}, \ldots, \left[T'_1\right]^{-1} .$$

Como $P = Q(J)$, resta-nos substituir (à direita), X de $Q(X)$ por J.

EXEMPLO 13.

Seja a matriz A, 4×4 sôbre C, do exemplo 2 de 4.1. Por aquêle exemplo, temos a sequência que reduz $XE - A$ à forma canônica diagonal (teorema 3 de 3.3). Logo, temos :-

$$Q_1(X) = [4(1)+3]\, [4(-1)+2]\, [4(1-X)+1]\, [1,4]\, [3(-X-1)+2]\, [3(1-4X)+4]$$
$$[2,3]\, [3(-5)+4]\, [3,4]\, [3(X+1)+4].$$

e

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

Consideremos agora, a matriz característica $XE - J$, e a colocaremos na forma canônica diagonal (teorema 3 de 3.3), através da seguinte sequência de operações elementares :-

$$[3(X+1)+2]\, [2(X+1)+1]\, (1(X+1)^2+3)\, (2(X+1)+3)\, (1(-1))\, (2(-1))\, (3,4)$$
$$[1,2]\, [2,3]\, [3,4] .$$

Consequentemente,

$$Q_2(X) = [3(X+1)+2]\, [2(X+1)+1]\, [1,2]\, [2,3]\, [3,4] .$$

Sendo $Q(X) = Q_1(X)\, Q_2^{-1}(X)$, e como,

$$Q_2^{-1} = [3,4]\, [2,3]\, [1,2]\, [2(-X-1)+1]\, [3(-X-1)+2] ,$$

então,

$Q(X) = [4(1)+3]\ [4(-1)+2]\ [4(-1-X)+1]\ [1,4]\ [3(-X-1)+2]\ [3(1-4X)+4]$
$[2,3]\ [3(-5)+4]\ [3,4]\ [3(X+1)+4]\ [3,4]\ [2,3]\ [1,2]\ [2(-X-1)+1]$
$[3(-X-1)+2]$.

Aplicando estas operações, nesta ordem, à matriz unidade E 4 X 4, a reduzimos à forma :-

$$\begin{bmatrix} X+1 & 0 & 0 & 1 \\ -5X-4 & 0 & 0 & -5 \\ X^2+6X+5 & -X-1 & 1 & X+6 \\ 10X+9 & -X & 1 & 12 \end{bmatrix} = Q(X) = Q_1(X)\, Q_2^{-1}(X) .$$

Logo, com coeficientes matriciais, temos :-

$$Q(X) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} X^2 + \begin{bmatrix} 1 & 0 & 0 & 0 \\ -5 & 0 & 0 & 0 \\ 6 & -1 & 0 & 1 \\ 10 & -1 & 0 & 0 \end{bmatrix} X + \begin{bmatrix} 1 & 0 & 0 & 1 \\ -4 & 0 & 0 & -5 \\ 5 & -1 & 1 & 6 \\ 9 & 0 & 1 & 12 \end{bmatrix}$$

É fácil de se verificar que :-

$$J^2 = J\,J = \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} .$$

Assim sendo, a matriz de transformação é $P = Q(J)$ (à direita), isto é,

$$P = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ -5 & 0 & 0 & 0 \\ 6 & -1 & 0 & 1 \\ 10 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} +$$

$$+ \begin{bmatrix} 1 & 0 & 0 & 1 \\ -4 & 0 & 0 & -5 \\ 5 & -1 & 1 & 6 \\ 9 & 0 & 1 & 12 \end{bmatrix} \quad \text{ou seja} \quad P = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & -5 & 0 & -5 \\ 0 & 4 & 1 & 5 \\ -1 & 11 & 0 & 12 \end{bmatrix} .$$

4.7 ALGUMAS APLICAÇÕES.

1. POTÊNCIAS INTEIRAS DE UMA MATRIZ.

Seja $A \in M_n(K)$. Pretendemos um método prático para se calcular as potências inteiras positivas da matriz A. Sucessivas multiplicações é um método simples, desde que o expoente r, seja pequeno, e impraticável para grandes valores de r. Vamos admitir que A é semelhante à uma matriz J na forma canônica de Jordan, isto é, que o polinômio característico de A se fatora num produto de fatores lineares no corpo K, o que sabemos que sempre é possível quando o corpo K é algebricamente fechado (por exemplo, para K = C). Assim sendo, existe uma matriz $P \in M_n(K)$, inversível, tal que :-

$$J = P^{-1} A P \qquad \text{ou} \qquad A = P J P^{-1}$$

Da definição de potência de uma matriz, vem que :-

$$(56) \quad A^r = \underbrace{A\,A \ldots A}_{r\text{-vezes}} = \underbrace{PJP^{-1}}_{A}\,\underbrace{PJP^{-1}}_{A} \ldots \underbrace{PJP^{-1}}_{A} = P\,J^r\,P^{-1}$$

Sabemos calcular a matriz de transformação P da matriz A à J (ver 4.6), logo resta-nos apenas aprendermos a calcular J^r, isto é, reduzimos o problema de se calcular a potência r-ésima de uma matriz qualquer A (que seja semelhante à uma matriz J, na forma de Jordan), ao problema de se calcular a potência r-ésima de uma matriz J na forma de Jordan.

A situação fica muito cômoda, quando A é uma matriz diagonalizável (ver 2.4), isto é, quando A é semelhante à uma matriz $D = [\lambda_1, \lambda_2, \ldots, \lambda_n]$ (que é uma particular matriz na forma de Jordan), pois nêste caso, é fácil de se ver que :-

$$D^r = [\lambda_1^r, \lambda_2^r, \ldots, \lambda_n^r],$$

para todo $r \in N$ (r finito).

Para estudarmos o caso geral, recorremos à regra do produto de matrizes em blocos (ver [3], pág.40 e seguintes), [illegible] interessa as potências inteiras (positivas) de uma matriz da forma :-

$$J = [J_1, J_2, \ldots, J_s],$$

onde cada J_i é um bloco de Jordan, logo

$$(57) \qquad J^r = [J_1^r, J_2^r, \ldots, J_s^r],$$

e reduzimos o problema ao cálculo da potência r-ésima de um bloco de Jordan. Na realidade vamos fazer mais que o necessário, [illegible] vez que nossa preocupação é aplicarmos nossos conhecimentos sôbre formas canônicas. Dado qualquer polinômio :-

$$p(X) = a_0 + a_1X + \dots + a_nX^n \ , \ gr(p) = n,$$

calcularemos $p(J_i)$, onde J_i é um bloco de Jordan de ordem t.
Seja :-

$$J_i = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{bmatrix} = \lambda I + N$$

onde N é a matriz nilpotente (de índice de nilpotência t):-

$$N = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

Seja $\mathcal{B} = \{e_1, e_2, \dots, e_t\}$ a base ordenada de K^t e L o operador linear sôbre K^t, cuja matriz relativa à base $\mathcal{B}$ é N. Então :

$Le_1 = 0, \ Le_2 = e_1, \ Le_3 = e_2, \ \dots \ , \ Le_t = e_{t-1}$

$L^2e_1 = 0, \ L^2e_2 = 0, \ L^2e_3 = e_1, \ \dots \ , \ L^2e_t = e_{t-2}$

. . . .

. . . .

$L^{t-1}e_1 = 0, \ L^{t-1}e_2 = 0, \ \dots \ , \ L^{t-1}e_{t-1} = 0, \ L^{t-1}e_t = e_1$

$L^t e_1 = 0, \ L^t e_2 = 0, \ \dots \ , \ L^t e_t = 0,$

isto é,

$$N^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}, \ \dots, \ N^{t-1} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

$$N^t = N^{t+1} = \dots = 0.$$

Aplicando a fórmula de Taylor para o polinômio p $(gr(p)=m)$ numa vizinhança de λ, obtemos :-

$$p(X) = p(\lambda) + p'(\lambda)\ (X-\lambda) + \frac{p''(\lambda)}{2!}(X-\lambda)^2 + \dots + \frac{p^{(m)}(\lambda)}{m!}\ (X-\lambda)^m$$

Então ,

$$p(J_i) = p(\lambda)E + (J_i - \lambda E)p'(\lambda) + \frac{p''(\lambda)}{2!}(J_i - \lambda E)^2 + \dots + \frac{p^{(m)}(\lambda)}{m!}\ (J_i - \lambda E)^m$$

e como $J_i = \lambda E + N$, $N = J_i - \lambda E$. Consequentemente,

$$p(J_i) = p(\lambda)E + p'(\lambda)\ N + \frac{p''(\lambda)}{2!}N^2 + \dots + \frac{p^{(m)}(\lambda)}{m!}\ N^m .$$

Desde que $N^t = N^{t-1} = \dots = 0$, temos que :-

$$p(J_i) = \begin{bmatrix} p(\lambda) & \frac{p'(\lambda)}{1!} & \frac{p''(\lambda)}{2!} & \dots & \frac{p^{(t-1)}(\lambda)}{(t-1)!} \\ 0 & p(\lambda) & \frac{p'(\lambda)}{1} & \dots & \frac{p^{(t-2)}(\lambda)}{(t-2)!} \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & \dots & p(\lambda) \end{bmatrix}$$

Assim, para calcularmos $p(J_i)$, onde J_i é um bloco de Jordan de ordem t, devemos conhecer os valores de $p(\lambda)$ e das primeiras (t-1) derivadas de p em λ, onde λ é o autovalor de J_i. Podemos assim calcular $p(J)$, $J = [J_1, J_2, \dots, J_s]$, pois é fácil ver que $p(J) = [p(J_1), p(J_2), \dots, p(J_s)]$.

Agora, o que nos interessa, é calcularmos as potências J_i^r, e assim basta tomarmos o polinômio $p(X) = X^r$ e teremos que :-

$$p(J_i) = J_i^r = \begin{bmatrix} \lambda^r & r\lambda^{r-1} & \frac{r(r-1)}{2}\lambda^{r-2} & \dots & \frac{r(r-1)\dots(r-t+2)}{(t-1)}\lambda^{r-t+1} \\ 0 & \lambda^r & r\lambda^{r-1} & \dots & \frac{r(r-1)\dots(r-t+3)}{(t-2)}\lambda^{r-t+2} \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & \dots & \lambda^r \end{bmatrix}$$

EXEMPLO 14.

Vamos calcular A^{251}, onde

$$A = \begin{bmatrix} -1 & 2 \\ 4 & 1 \end{bmatrix}$$

O polinômio característico de A é :-

$$c(X) = (X+1)(X-1) - 8 = X^2 - 9 = (X+3)(X-3),$$

Então os autovalores de A são $\lambda_1 = -3$ e $\lambda_2 = 3$, e A é diagonalizácel (ver teorema 5, do 2.4). Ademais A é semelhante a matriz :-

$$D = \begin{bmatrix} -3 & 0 \\ 0 & 3 \end{bmatrix}.$$

Devemos encontrar a matriz P, que transforma A em D, isto é, a matriz P é tal que :-

$$D = P^{-1} A P .$$

Sabemos qe a primeira coluna de P são as coordenadas de um autovetor de A associado à -3, e a segunda, as coordenadas de um autovetor associado ao autovalor 3. Calculando os autovetores temos por exemplo,

$$P = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}, \text{ e daí } \quad P^{-1} = 1/3 \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}$$

Como $D = P^{-1}AP$ ou $A = P D P^{-1}$ então,

$$A^{251} = P D^{251} P^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} (-3)^{251} & 0 \\ 0 & 3^{251} \end{bmatrix} 1/3 \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} =$$

$$= 3^{250} \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} = 3^{250} \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix} = 3^{250} \begin{bmatrix} -1 & 2 \\ 4 & 1 \end{bmatrix}.$$

<u>EXEMPLO 15.</u>

Vamos considerar a matriz A do exemplo 13. de 4.6, isto é, a matriz :-

$$A = \begin{bmatrix} 1 & -1 & 1 & -1 \\ -3 & 3 & -5 & 4 \\ 8 & -4 & 3 & -4 \\ 15 & -10 & 11 & -11 \end{bmatrix},$$

e queremos calcular A^{1000}, por exemplo.

Como vimos naquêle exemplo,

$$A = P J P^{-1}, \text{ onde,}$$

$$J = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix} \text{ e } P = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & -5 & 0 & -5 \\ 0 & 4 & 1 & 5 \\ -1 & 11 & 0 & 12 \end{bmatrix}.$$

Assim, $A^{1000} = P\, J^{1000}\, P^{-1}$, e para calcularmos J^{1000}, temos:-

$$J^{1000} = \left[J_1^{1000} \; , \; J_2^{1000} \right], \text{ onde :-}$$

$$J_1 = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix} \quad \text{e} \quad J_2 = \begin{bmatrix} -1 \end{bmatrix}.$$

Tomamos $p(X) = X^{1000}$ e daí

$$p(J_1) = J_1^{1000} = \begin{bmatrix} (-1)^{1000} & 1000(-1)^{999} & 1/2\ 1000\ (999)(-1)^{998} \\ 0 & (-1)^{1000} & 1000\ (-1)^{999} \\ 0 & 0 & (-1)^{1000} \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & -1000 & 499500 \\ 0 & 1 & -1000 \\ 0 & 0 & 1 \end{bmatrix}$$

Como J_2 é uma submatriz diagonal $J_2^{1000} = \left[(-1)^{1000}\right] = [1]$.

Agora podemos escrever a matriz J^{1000}, que é :-

$$J^{1000} = \left[\begin{array}{ccc|c} 1 & -1000 & 499500 & 0 \\ 0 & 1 & -1000 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right]$$

Efetuando os cálculos, obtemos :-

$$P^{-1} = \begin{bmatrix} 5 & 1 & 0 & 0 \\ 7 & -1 & 0 & -1 \\ 2 & -1 & 1 & -1 \\ -6 & 1 & 0 & 1 \end{bmatrix},$$

$$A^{1000} = \begin{bmatrix} -1999 & 1000 & -100 & 1000 \\ 8895 & -503499 & 500000 & 493500 \\ -8000 & 4000 & -399 & 4000 \\ -1014000 & 505500 & -500600 & -487499 \end{bmatrix}.$$

2. DEMONSTRAÇÃO DO TEOREMA DE Hamilton-Cayley, USANDO A FORMA DE JORDAN.

Lembramos o enunciado do teorema de Hamilton-Cayley, teorema 9 de 1.4, no caso em que o corpo em questão é algebricamente fechado.

TEOREMA DE HAMILTON-CAYLEY.

" Seja c o polinômio característico de uma matriz A, n x n, sôbre um corpo F algebricamente fechado. Então, c(A) = 0" .

Demonstração .

Sendo F algebricamente fechado, podemos colocar o polinômio característico c de A na forma :-

$$(58) \qquad c(X) = \prod_{i=1}^{n} (X - \lambda_i),$$

onde para cada $i=1,2, \ldots, n$, λ_i é um autovalor de A (os λ_i não são necessariamente distintos). Sabemos que A é semelhante à uma matriz J na forma de Jordan. Digamos que :-

$$J = \begin{bmatrix} \lambda_1 & * & 0 & \ldots & 0 & 0 \\ 0 & \lambda_2 & * & \ldots & 0 & 0 \\ . & . & . & \ldots & . & . \\ 0 & 0 & 0 & \ldots & \lambda_{n-1} & * \\ 0 & 0 & 0 & \ldots & 0 & \lambda_n \end{bmatrix},$$

onde $*$ = 0 ou 1. De (58), sendo $A_i = J - \lambda_i E$, $i = 1,2, \ldots, n$, temos que :-

$$c(J) = \prod_{i=1}^{n} A_i .$$

Como a linha i de A_i tem $*$ (0 ou 1) na coluna i+1 e zeros nas posições restantes, a linha n de A_n é uma linha só de zeros, as duas últimas linhas de $A_{n-1}A_n$ são formadas só de zeros, as últimas três linhas de $A_{n-2}A_{n-1}A_n$, são formadas só de zeros e assim sucessivamente. Logo, a multiplicação de todos os A_i, $i=1,2,\ldots,n$, dá a matriz nula, isto é,

$$c(J) = \prod_{i=1}^{n} A_i = 0.$$

Sendo A semelhante à J, c(A) é semelhante à c(J) = 0, logo existe uma matriz inversível $P \in M_n(F)$, tal que :-

$$c(A) = P^{-1} c(J)\, P = P^{-1}\, 0\, P = 0.$$

c.q.d.

3. CÁLCULO DA INVERSA DE UMA MATRIZ NÃO-SINGULAR A.

Seja $A \in M_n(F)$, não-singular, onde F é um corpo algebricamente fechado. Então, A é semelhante à uma matriz J na forma de Jordan e o polinômio característico de A e de J é o mesmo. Os autovalores de J situam-se sôbre sua diagonal, e sendo J semelhante à A, J também é não-singular ($\det J = \det P^{-1} \det A \det P = \det A \neq 0$). Ademais, J é triangular e daí $\det J$ é o produto de seus autovalores. Concluimos assim que nenhum autovalor de J (ou de A) é zero, ou seja, que $c(0) \neq 0$. Seja,

$$c(X) = X^n + a_1 X^{n-1} + \dots + a_{n-1} X + a_n ,$$

o polinômio característico de A. De $c(0) \neq 0$, vem que $a_n \neq 0$. Do teorema de Hamilton-Cayley, temos que :-

$$c(A) = A^n + a_1 A^{n-1} + \dots + a_{n-1} A + a_n E = 0, \text{ ou}$$

$$(59) \qquad (A^n + a_1 A^{n-1} + \dots + a_{n-1} A) = -a_n E .$$

Sendo A não-singular, existe A^{-1} e como $a_n \neq 0$, multiplicando-se (59) por $-a_n^{-1} A^{-1}$, obtemos :-

$$(60) \qquad A^{-1} = -a_n^{-1} (A^{n-1} + a_1 A^{n-2} + \dots + a_{n-2} A + a_{n-1} E).$$

De (60) vê-se que a inversa A^{-1} de A pode ser calculada, conhecendo-se o polinômio característico de A, bem como as potências inteiras não negativas de A, até a ordem $n-1$. Infelizmente, para grandes valores de n, os coeficientes a_i do polinômio característico de A podem ser muito difíceis de se calcular diretamente.

Colocamos sem demonstrar (ver [2], pág. 188 e seguintes), um método para se calcular os coeficientes do polinômio característico.

O TRAÇO de uma matriz $A \in M_n(F)$, é dado por :-

$$\operatorname{tr} A = \sum_{i=1}^{n} \lambda_i , \text{ onde } \lambda_1, \lambda_2, \dots, \lambda_n \text{ são os autovalores}$$

de A.

Desde que matrizes semelhantes possuem idênticos autovalores, elas possuem também idênticos traços. Pode-se então definir o traço de um operador linear, sôbre um F-espaço vetorial V (de dimensão finita n), como sendo o traço de qualquer matriz que o represente. Uma definição alternativa para o traço de uma matriz $A = [a_{ij}]$ é :-

$$\operatorname{tr} A = \sum_{i=1}^{n} a_{ii} .$$

As fórmulas que nos fornecem os coeficientes do polinômio característico de A ($c(X) = X^n + a_1X^{n-1} + \dots + a_{n-1}X + a_n$), são:-

$$
(61) \begin{cases}
a_1 = -\mathrm{tr}\, A. \\
a_2 = -1/2 \ \left[a_1 \mathrm{tr}\, A + \mathrm{tr}\, A^2\right]. \\
a_3 = -1/3 \ \left[a_2 \mathrm{tr}\, A + a_1 \mathrm{tr}\, A^2 + \mathrm{tr}\, A^3\right]. \\
\dots\dots\dots\dots\dots\dots\dots\dots \\
\dots\dots\dots\dots\dots\dots\dots\dots \\
a_n = -1/n \ \left[a_{n-1} \mathrm{tr}\, A + a_{n-2} \mathrm{tr}\, A^2 + \dots + a_1 \mathrm{tr}\, A^{n-1} + \mathrm{tr}\, A^n\right].
\end{cases}
$$

Com êste processo além de se calcular A^{-1}, obtemos também o polinômio característico de A.

EXEMPLO 16.

Seja $A = \begin{bmatrix} 1 & -1 \\ 2 & 2 \end{bmatrix}$ uma matriz de $M_2(C)$. Como o polinômio característico de A é $c(X) = X^2 - 3X + 4$, logo $n = 2$, $a_1 = -3$ e $a_2 = 4$, logo :-

$$A^{-1} = -1/4 \begin{bmatrix} -2 & -1 \\ 2 & -1 \end{bmatrix}.$$

EXEMPLO 17.

Consideremos agora a matriz A dada por :-

$$A = \begin{vmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{vmatrix}.$$

Fazendo alguns cálculos com base nas formulas (61), vem que:- precisamos calcular as potências de A até a ordem 3 e a diagonal de A^4, que são :-

$$A^2 = \begin{bmatrix} 2 & -3 & 0 & -1 \\ 0 & 2 & 0 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 2 \end{bmatrix} \text{ e } A^3 = \begin{bmatrix} 2 & -6 & 2 & -4 \\ 0 & 2 & 0 & 2 \\ 2 & -2 & -2 & 0 \\ 0 & 2 & 0 & -2 \end{bmatrix},$$

e a diagonal de A^4 é 4,4,4,4 , assim
$n=4$, $a_1 = 0$, $a_2 = -4$, $a_3 = 0$, $a_4 = 4$, logo $c(X) = X^4 - 4X^2 + 4$ e

$$A^{-1} = 1/2 \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}.$$

APÊNDICE A - POLINÔMIOS SÔBRE UM CORPO.

A.1 DEFINIÇÕES INICIAIS.

Seja K um corpo comutativo com elemento unidade 1. Consideremos o conjunto :-

$$S(K) = \{(a_i)_{i\in N} : a_i \in K\} = \{(a_i) : a_i \in K\} ;$$

de tôdas as sequências de elementos de K. Definimos em S(K) uma operação de adição, pondo :-

$$+ : S(K)\times S(K) \longrightarrow S(K)$$
$$((a_i),(b_i)) \longmapsto (a_i)+(b_i) = (a_i+b_i).$$

É imediato de se verificar, que S(K) é um grupo comutativo em relação à esta operação. O elemento zero é (a_i) com $a_i = o$, $i = o,1,2,\ldots$ O inverso (aditivo) de (a_i) é $(-a_i)$. Diz-se que uma sequência (a_i) de S(K) é <u>QUASE-NULA</u> se, e sòmente se, $a_i \neq o$ sòmente para um número finito de índices $i\in N$.

É fácil de se ver, que uma definição alternativa de sequência quase-nula $(a_i)\in S(K)$ é que existe $n\in N$, tal que $a_i = o$, para todo $i>n$. Denotaremos por E o conjunto de tôdas as sequências quase-nula de S(K). Exemplos de elementos de E, são $O' = (a_i) : a_i = o,\ i\in N$ e $I' = (a_i) :\ a_o = 1$ e $a_i = o$, para $i = 1,2,\ \ldots$. Sejam (a_i) e (b_i) pertencentes à E, logo existem $n_1, n_2 \in N$, tais que $a_i = o$ para $i>n_1$ e $b_i = o$ para $i>n_2$. Então, $(a_i)+(b_i) = (a_i+b_i)\in E$, pois se tomarmos $n = \text{máx}\{n_1,n_2\}$, temos que $a_i+b_i = o$ para $i>n$. Vemos assim, que E é fechado em relação a adição, e é trivial que essa operação induz uma estrutura de "grupo comutativo" sôbre o conjunto E.

Definiremos uma multiplicação em E, pondo:-

$$\cdot : E\times E \longrightarrow E$$
$$((a_i),(b_j)) \longmapsto (c_k) = (a_i).(b_j)\ , \quad \text{onde :-}$$

$$(1)\qquad c_k = \sum_{i+j=k} a_i b_j = \sum_{i=o}^{k} a_i b_{k-i} = \sum_{j=o}^{k} a_{k-j} b_j =$$

$$= a_o b_k + a_1 b_{k-1} + a_2 b_{k-2} + \ldots + a_{k-1} b_1 + a_k b_o .$$

Devemos mostrar que a sequência $(c_h) = (a_i).(b_j)$ é uma sequência de E. De fato, como $(a_i),(b_j)\in E$, existem $m,n\in N$, tais que:

$$a_i = o,\ \text{para } i>m \quad \text{e} \quad b_j = o,\ \text{para } j>n.$$

Logo, de (1), temos que :-

$$c_{m+n} = a_m b_n \quad \text{e} \quad c_k = o, \text{para todo } k>m+n,$$

isto é, a sequência (c_k) é quase-nula.

Temos assim duas operações, adição e multiplicação, no conjunto E. Pode-se verificar que E é um "anel comutativo" com elemento unidade (1'), relativamente às operações que introduzimos.

Os elementos do anel E, que acabamos de construir, denominam-se POLINÔMIOS COM COEFICIENTES EM K e $(E,+,.)$ diz-se ANEL DE POLINÔMIOS COM COEFICIENTES EM K.

Passaremos a denotar os elementos de E por f,g,h,p,etc. O elemento zero de E será denotado por 0 e o elemento unidade por 1.

Seja $f \neq 0$, $f \in E$. Então, existe um índice $i_0 \in N$, tal que $a_{i_0} \neq 0$, ($f = (a_i)$), e por outro lado, existe $m \in N$ tal que $a_i = 0$ para todo $i > m$. Assim o conjunto :-

$$\{ i \in N : a_i \neq 0 \},$$

é não vazio e finito. Denomina-se GRAU do polinômio $f \in E$, $f \neq 0$, ao natural :-

$$n = \text{máx} \{ i \in N : a_i \neq 0 \}.$$

É óbvio que o grau de todo polinômio $f \neq 0$, é maior ou igual à zero. Nêsse caso, o escalar $a_n \neq 0$ é denominado COEFICIENTE LIDER ou DOMINANTE do polinômio f e quando $a_n = 1$, diz-se que f é um POLINÔMIO UNITÁRIO ou MÔNICO.

Denota-se o grau do polinômio $f \neq 0$, por gr(f). Assim, gr(f) = n se e, e sòmente se, $a_n \neq 0$ e $a_i = 0$ para todo $i > n$. Nenhum grau se atribui ao polinômio nulo 0.

Seja,

$$\bar{K} = \{ (a_i) \in E : a_i = 0, \forall i > 0 \}.$$

$\bar{K}$ é um subanel unitário do anel E, na verdade $\bar{K}$ é um corpo, pois se $(a_i) = (a,0,0, \ldots ,0, \ldots) \in \bar{K}$, $a \neq 0$, temos que $(b_i) = (a^{-1},0,0,\ldots)$ é um elemento de $\bar{K}$, tal que $(a_i).(b_i) = (1,0,0, \ldots ,0, \ldots)$. Consideremos a aplicação :-

$$\varphi : K \longrightarrow \bar{K}$$
$$a \longmapsto \varphi(a) = (a_i) = (a,0, \ldots ,0, \ldots) .$$

Fàcilmente se verifica que φ é bijetora e que :

$$\varphi (a+b) = \varphi (a) + \varphi (b)$$

e

$$\varphi (ab) = \varphi (a) \varphi (b),$$

para todo $a,b \in K$. Isto é, φ é um isomorfismo do corpo K no corpo $\bar{K}$. Assim sendo, identificamos o corpo K com o corpo $\bar{K}$, através dêsse isomorfismo, e podemos colocar $a = (a,0, \ldots ,0, \ldots)$, para todo $a \in K$. Os elementos de K são denominados POLINÔMIOS CONSTANTES.

Consideramos o polinômio ;

(2) $p_r = (\delta_{r,i})_{i \in N}$, onde $r \in N$ é fixo e $\delta_{r,r} = 1$ e $\delta_{r,i} = 0$ se $i \neq r$.

Assim sendo, temos que :-

$$p_0 = (1,0, \ldots ,0, \ldots)$$
$$p_1 = (0,1, \ldots ,0, \ldots)$$
$$\ldots\ldots\ldots$$

e de um modo geral para um r fixo em N temos que :-

$$p_r = (0,0, \ldots ,0,1,0, \ldots)$$
$$\quad\quad\quad\quad (r)$$

Para quaisquer $r,s \in N$, de (1) e (2) vem que :-

(i) $p_r p_s = (c_k)$, onde $c_k = \sum_{i+j=k} \delta_{r,i}\delta_{s,j}$, e $c_k \neq 0$ para $k \neq r+s$ e $c_{r+s} = 1$, logo $p_r p_s = p_{r+s}$.

(ii) Para todo $a \in K$, temos que:-

$$a\, p_r = (c_k),$$

e como $\underline{a}$ se identifica com o polinômio $(a\,\delta_{0,i})$ temos de (1) que:-

$$c_k = \sum_{i+j=k} a\,\delta_{0,i}\,\delta_{r,j} = a\,\delta_{r,k}\text{ , logo}$$

$$a\,p_r = (a\,\delta_{r,i})_{i \in N}\text{, para todo } r \text{ fixado em } N.$$

Denotaremos por $X^0 = 1, X, X^2, X^3, \ldots$, os polinômios $p_0, p_1, p_2, p_3, \ldots$, dados por (2), respectivamente.

De acôrdo com a definição de soma de polinômios e a fórmula (ii), temos que se $f = (a_i) \in E$ tem grau n, logo:-

$$f = (a_0,a_1,a_2, \ldots ,a_n,0,0, \ldots),\text{ com } a_n \neq 0$$

temos que:-

$$f = a_0p_0 + a_1p_1 + a_2p_2 + \ldots + a_np_n,$$

ou na notação que acabamos de introduzir,

(3) $\qquad f = a_0X^0 + a_1X^1 + a_2X^2 + \ldots + a_nX^n = \sum_{i=0}^{n} a_iX^i.$

Cada polinômio $a_ip_i = a_iX^i$, diz-se <u>MONÔMIO DE GRAU i</u>, pois $gr(p_i) = i$. O elemento a_i em (3) denomina-se <u>COEFICIENTE DO MONÔMIO a_iX^i EM f</u>. O monômio $p_1 = X = (0,1,0, \ldots ,0, \ldots)$ é denominado <u>INDETERMINADA</u> e diz-se que (3) <u>É UM POLINÔMIO NA INDETERMINADA X, COM COEFICIENTES EM K</u>. Isto é, (3) é a expressão usual de um polinômio em X, com coeficientes em K.

Dois polinômios, não nulos, f e g são <u>IGUAIS</u> se, e sòmente se, $gr(f) = gr(g)$ e se $f = a_0X^0 + a_1X + \ldots + a_nX^n$ e $g = b_0X^0 + b_1X + \ldots + b_nX^n$ então $a_i = b_i$, para todo $i = 0,1,2, \ldots ,n$.

Dados os polinômios $f = a_0 + a_1X + a_2X^2 + \ldots + a_nX^n$, $gr(f) = n$ e $g = b_0 + b_1X + B_2X^2 + \ldots + b_mX^m$, $gr(g) = m$, temos que :-

(4)

(i) $\quad f+g = \sum_{i=0}^{r} (a_i + b_i)\, X^i$, onde $r = \text{máx}\{n,m\}$.

(ii) $\quad -f = \sum_{i=0}^{n} (-a_i)\, X^i.$

(iii) $f.g = \sum_{k=0} c_k X^k$, onde $c_k = \sum_{i+j=k} a_i b_j$.

Essas fórmulas nos dão os processos usuais para determinarmos a soma, o inverso aditivo e o produto de polinômios dados (não nulos). Obsevamos que o produto entre f e g é obtido multiplicando-se cada monômio de f, por todos os monômios de g, somando-se em seguida, os produtos assim obtidos. Isto é, multiplicamos como se faz usualmente no corpo K, usando a lei distributiva e agrupando os monômios que têm o mesmo grau.

É imediato que o subanel de E, gerado por $K \cup \{X\} = K[X]$ coincide com E. Denotaremos o anel de polinômios E pela notação $K[X]$, o qual denominaremos <u>ANEL DE POLINÔMIOS NA INDETERMINADA X, COM COEFICIENTES EM K.</u>

Enunciaremos aqui uma proposição que não demonstraremos (para uma demonstração ver [6]).

<u>PROPOSIÇÃO 1</u>

" Sejam $f, g \in K[X]$, não nulos. Então:-

(a) $f.g \in K[X]$ é não nulo, isto é, $K[X]$ é um <u>domínio de integridade.</u>

(b) $gr(f.g) = gr(f) + gr(g)$.

(c) Sendo f e g unitários, f.g também o é,

(d) $f.g \in K$ se, e sòmente se, $f \in K$ e $g \in K$.

(e) quando $f + g \neq 0$, temos que :-

$$gr(f+g) \leq máx\{gr(f),\ gr(g)\} .$$

(f) Sendo $f.g = f.h$, para algum $h \in K[X]$, temos que $g = h$."

Sejam $f, g \in K[X]$, tais que $f = hg$, para algum $h \in K[X]$. Então, diz-se que f é um <u>MÚLTIPLO</u> de g ou f é <u>DIVISÍVEL</u> por g, ou ainda, que g <u>DIVIDE</u> f, em $K[X]$. É imediato que se g divide f ($f \neq 0$), então $gr(g) \leq gr(f)$. Com efeito, existe $h \in K[X]$, tal que :-

$$f = hg \quad (h \neq 0, \text{ caso contrário } f = 0),$$

e daí $gr(f) = gr(g) + gr(h)$ e como $gr(h) \geq 0$, logo $gr(f) \geq gr(g)$.

Ressaltamos por fim, que a construção que fizemos do anel $K[X]$, pode ser feita de modo totalmente análogo, para o caso de K ser apenas um anel comutativo com unidade. Sempre que houver perigo de confusão, nas notas que descrevemos, costumamos denotar um polinômio f na indeterminada X, por $f(X)$, tendo em vista que falamos também em polinômio em um operador linear ou numa matriz.

<u>A.2 ALGORITMO DA DIVISÃO.</u>

Vamos agora ver um algoritmo de divisão para polinômios, que é análogo ao algoritmo de Euclides para divisão de números inteiros.

<u>ALGORITMO DA DIVISÃO</u>

" Sejam $f = a_0 + a_1X + \dots + a_nX^n$ e $g = b_0 + b_1X + \dots + b_mX^m$, polinômios de $K[X]$, tais que $gr(f) = n$ ou $f = 0$ e $gr(g) = m$ e $g \neq 0$. Então existem polinômios q e r em $K[X]$, tal que :-

(4) $f = q\,g + r$,

onde ou $r = 0$ ou $gr(r) < gr(g)$. Ademais, os polinômios q e r são ùnicamente determinados, pelas condições acima".

DEMONSTRAÇÃO.

EXISTÊNCIA.

Sendo $n < m$, toma-se $q = 0$, e daí $f = r$ e $gr(r) = gr(f) = n < m = gr(g)$. Consideremos $n \geq m$, e definimos :-

(5) $f_1 = f - a_n b_m^{-1} X^{n-m} g = a_0 + a_1 X + \dots + a_n X^n - (a_n X^n + a_n b_m^{-1} b_{m-1} X^{n-1} + \dots + a_n b_m^{-1} b_0 X^{n-m}) = (a_{n-1} - a_n b_m^{-1} b_{m-1}) X^{n-1} + \dots + a_0$,

logo, $gr(f_1) \leq n-1$ ou $f_1 = 0$. Continuando nêsse processo, ou mais formalmente por indução sôbre n, pode-se encontrar polinômios q_1 e r_1 tais que :-

$$f_1 = q_1\, g + r_1,\ \text{com } gr(r_1) < gr(g) \text{ ou } r_1 = 0.$$

Vemos que de (5), teremos que

$$f = a_n b_m^{-1} X^{n-m} g + f_1 = a_n b_m^{-1} X^{n-m} g + q_1\, g + r_1 = (a_n b_m^{-1} X^{n-m} + q_1) g + r_1$$

se tomarmos $q = a_n b_m^{-1} X^{n-m} + q_1$ e $r = r_1$, temos f na forma (4).

UNICIDADE.

Suponhamos que $f = q_1 g + r_1 = q_2 g + r_2$, onde $r_1 = 0$ ou $gr(r_1) < gr(g)$ e $r_2 = 0$ ou $gr(r_2) < gr(g)$.Então temos que :-

$$(*)\ [q_1 - q_2]\, g = r_2 - r_1.$$

Daí g divide $r_2 - r_1$. Caso $r_2 - r_1 \neq 0$, teremos $gr(g) \leq gr(r_2 - r_1)$. Por outro lado, o grau de $[q_1 - q_2]g$ é maior ou igual ao grau de g (à menos que $q_2 - q_1 = 0$), e teremos que $gr(r_2 - r_1) \leq \text{máx}\{gr(r_2), gr(r_1)\} < gr(g)$. Assim a única possibilidade é que ambos os membros da igualdade(*)sejam nulos, ou seja $r_2 = r_1$ e $q_2 = q_1$.

c.q.d.

A.3 IDEAL EM $K[X]$.

Um IDEAL num domínio polinomial $K[X]$, é um conjunto $\mathfrak{G}$ não vazio, de polinômios que satisfazem as seguintes condições :-

(i) $f \in \mathfrak{G} \Longrightarrow \forall\, h \in K[X],\ h.f \in \mathfrak{G}$.

(ii) $f, g \in \mathfrak{G} \Rightarrow (f + g) \in \mathfrak{G}$.

É fácil de se verificar que o domínio polinomial $K[X]$ e o conjunto unitário $\langle 0 \rangle$, (ideal nulo), são ideais no domínio polinomial $K[X]$, êsses são os ideais triviais de $K[X]$; Um outro exemplo de ideal em $K[X]$ é o conjunto dos polinômios que são múltiplos de um polinômio dado g, o qual denomina-se GERADOR do ideal $\mathfrak{G}$.Um ideal dêsse tipo é denominado IDEAL PRINCIPAL.

TEOREMA 1.

" Todo ideal do domínio polinomial $K[X]$ é um ideal pricipal".

Demonstração.

Seja $\mathcal{T}$ um ideal qualquer de $K[X]$. Caso $\mathcal{T}=\langle 0\rangle$, tomamos o gerador como sendo o polinômio nulo. Podemos supor então, $\mathcal{T}\neq\langle 0\rangle$. Logo existe pelo menos um polinômio em $\mathcal{T}$ não nulo, e podemos selecionar em $\mathcal{T}$ um polinômio de grau mínimo, digamos $g\neq 0$. Nós afirmamos que todo polinômio de $\mathcal{T}$ é um múltiplo de g. De fato, fixemos arbitràriamente um polinômio $f\in\mathcal{T}$. Pelo algoritmo da divisão, existem $q,r\in K[X]$, tai q ue : -

$$r = f - qg,\ \text{onde}\ r = 0\ \text{ou}\ gr(r)\quad gr(g).$$

Desde que $g\in\mathcal{T}$, da condição (i) da definição de ideal, temos que $-qg\in\mathcal{T}$. Agora f e $-qg$ estão em $\mathcal{T}$, e da condição (ii) da definição de ideal obtemos $f - qg = r\in\mathcal{T}$. Sendo g 1 polinômio de menor grau em $\mathcal{T}$, temos que $gr(g)\leqslant gr(r)$, logo $r = 0$, e daí

$$f = qg.$$

Da arbitrariedade na escolha de f em $\mathcal{T}$, vem que todo polinômio de $\mathcal{T}$ é um múltiplo de g, isto é, g é um gerador do ideal $\mathcal{T}$ que é então principal, como queriamos demonstrar.

Observação. O polinômio g usado na demonstração acima é ùnivocamente determinado, a menos de produto por um fator constante. De fato, se $g'\neq 0$ é qualquer outro polinômio de grau mínimo no ideal $\mathcal{T}\neq\langle 0\rangle$ de $K[X]$, então pelo provado acima, g' é um múltiplo de g, logo

$$g' = hg \quad \text{e assim}\quad gr(g') = gr(h) + gr(g)\ \text{e como}$$

$gr(g') = gr(g)$, temos que $gr(h) = 0$, isto é, h é um polinômio constante não nulo. Então, se exigirmos que g tenha grau mínimo e coeficiente lider igual à 1, teremos que g é ùnivocamente determinado, já que não pode existir dois polinômios distintos, que diferem apenas por um fator constante, tendo ambos coeficientes lideres iguais à 1.

A.4 <u>MÁXIMO DIVISOR COMUM</u>

Sejam $g_1, g_2, \ldots, g_n$ polinômios arbitrários de $K[X]$. Todo polinômio de $K[X]$ que é um divisor de cada um dêsses g_i denomina-se um <u>DIVISOR COMUM</u> de $g_1, g_2, \ldots, g_n$. Examinaremos o conjunto de todos os divisores comuns. Caso $g_i = 0$ para todo $i = 1,2,\ldots,n$, então qualquer polinômio de $K[X]$ é um divisor comum, desde que 0 é múltiplo de todo polinômio, já que $0.f = 0$, para todo f de $K[X]$. Nós necessitamos então de discutir apenas o caso no qual nem todos os g_i são nulos. Consideremos o conjunto $\mathcal{T}$ de todos os polinômios da forma :-

(6) $h_1g_1 + h_2g_2 + \cdots + h_ng_n$,

onde cada h_i, $i = 1,2,\ldots,n$, é um polinômio arbitrário de $K[X]$. Mostraremos que êsse conjunto $\mathcal{T}$ é um ideal de $K[X]$, o qual se denomina <u>IDEAL GERADO</u> pelos polinômios $g_1, g_2, \ldots, g_n$. É claro que $\mathcal{T}\neq\emptyset$.

(i) Para todo f de $K[X]$, se multiplicarmos f por um polinômio da forma (6) obtemos,

$$[f\,h_1]g_1 + [f\,h_2]g_2 + \cdots + [f\,h_n]g_n$$

que é ainda um polinômio da forma (6).

(ii) Finalmente, se f e p são elementos de $\mathcal{T}$, digamos que
$f = h_1 g_1 + h_2 g_2 + \dots + h_n g_n$ e $p = h_1' g_1 + h_2' g_2 + \dots + h_n' g_n$,
então $f + p = (h_1 + h_1')\, g_1 + (h_2 + h_2')\, g_2 + \dots + (h_n + h_n')\, g_n \in \mathcal{T}$.

Sendo $\mathcal{T}$ um ideal de $K[X]$, segue do teorema 1 que $\mathcal{T}$ é um ideal principal. Logo o ideal $\mathcal{T}$ consiste de todos os múltiplos de um polinômio fixo d, e $d \neq 0$ já que $\mathcal{T}$ contém polinômios $\neq 0$, pois nem todos os g_i são nulos e podemos tomar $h_i = 1$, $i = 1, 2, \dots, n$. Ademais d é um divisor comum de $g_1, g_2, \dots, g_n$, pois cada um dêsses polinômios estão em $\mathcal{T}$ e d é um gerador do ideal $\mathcal{T}$. Por outro lado d é também um polinômio de $\mathcal{T}$, logo deve ser da forma (6), digamos que:-

(7) $$d = f_1 g_1 + f_2 g_2 + \dots + f_n g_n .$$

Afirmamos que todo divisor comum de $g_1, g_2, \dots, g_n$, é também um divisor de d. De fato, seja p um divisor comum de $g_1, g_2, \dots, g_n$ escolhido arbitrariamente. Então para todo i, uma equação da forma $g_i = q_i\, p_i$ deve ser válida. Substituindo-se em (7) e fatorando p, vem

$$d = (f_1 q_1 + f_2 q_2 + \dots + f_n q_n)\, p ;$$

logo p divide d. Vemos assim que d satisfaz às seguintes propriedades

(a) d é um divisor comum de $g_1, g_2, \dots, g_n$.

(b) Sendo p um divisor comum de $g_1, g_2, \dots, g_n$, então p divide d.

Êsse polinômio d é denominado <u>MÁXIMO DIVISOR COMUM</u> (m.d.c.) dos polinômios $g_1, g_2, \dots, g_n$. Da observação que fizemos logo abaixo do teorema 1, vem que o m.d.c. é ùnivocamente determinado, à menos de um fator constante. Demonstramos o seguinte teorema :-

<u>TEOREMA 2.</u>

" Sejam $g_1, g_2, \dots, g_n$ polinômios de $K[X]$, não todos nulos, então êles têm um m.d.c. d, que é um gerador do ideal gerado por êsses polinômios. Ademais, d é ùnivocamente determinado, à menos de um fator constante, e pode ser escrito na forma :-

(7) $$d = f_1 g_1 + f_2 g_2 + \dots + f_n g_n , \quad \text{com } f_i \in K[X]."$$

Quando o m.d.c. dos polinômios $g_1, g_2, \dots, g_n$ é uma constante, os polinômios $g_1, g_2, \dots, g_n$ são denominados <u>RELATIVAMENTE PRIMOS</u>. Nêsse caso êles não têm fatores comuns de grau maior (estritamente) que zero.

Os polinômios constantes $c \in K$, não nulos, são divisores de todo polinômios $f \in K[X]$. De fato $f = c(c^{-1} f)$. Ainda mais, para todo $f \in K[X]$ temos que o polinômio cf, $c \neq 0$, é um divisor de f. Os polinômios constantes e não nulos e os polinômios cf, $c \neq 0$, são denominados <u>DIVISORES IMPRÓPRIOS</u> de f. Por outro lado, um divisor g de f tal que $0 < gr(g) < gr(f)$, é denominado um <u>DIVISOR PRÓPRIO</u> de f. Um polinômio f, com $gr(f) \geq 1$, que não possui divisores próprios em $K[X]$ é denominado <u>POLINÔMIO IRREDUTÍVEL</u> em $K[X]$. Por exemplo, todo polinômio de grau 1 é irredutível em $K[X]$ (para todo K), já que todo divisor de tal polinômio, possui grau menor ou igual a 1, isto é, [illegible] logo será sempre um divisor impróprio. Polinômios que possuem divisores próprios são denominados <u>REDUTIVEIS</u>.

Uma consequência imediata da definição de polinômios irredutíveis é que :- " dois polinômios irredutíveis são relativamente primos ou diferem por um fator constante." Os polinômios irredutíveis se comportam em $K[X]$, semelhantemente aos números primos na teoria dos números.

A.5 FATORAÇÃO UNICA.

<u>TEOREMA 3.</u>

" Quando o produto fg entre dois polinômios é divisível por um polinômio irredutível p, então no mínimo um dos fatores f ou g é divisível por p ".

Demonstração.

Suponhamos que f não seja divisível por p, e mostraremos que g é divisível por p. Desde que p é irredutível e p não divide f, temos que os únicos divisores comuns de p e f, são polinômios constantes não nulos. Logo p e f são relativamente primos , e o m.d.c. entre êles é uma constante, que podemos tomar como sendo 1, já que o m.d.c. é determinado à menos de uma constante. Então por (7), 1 pode ser representado por :-

(8) $$1 = h_1 f + h_2 p$$

Multiplicando-se (8) por g obtemos ;

(9) $$g = h_1 fg + h_2 gp$$

Da hipótese de fg ser divisível por p, temos que $fg = qp$, que levado em (9) nos dá,

$$g = (h_1 q + h_2 g) p ,$$

ou seja, g é divisível por p.

c.q.d.

O resultado do teorema anterior pode fàcilmente ser extendido, por indução, para o produto de n polinômios, isto é, temos o seguinte teorema.

<u>TEOREMA 4.</u>

" Quando o produto $f_1 . f_2 . \cdots . f_n$ entre polinômios de $K[X]$, é divisível por um polinômio irredutível p de $K[X]$, então no mínimo um dos polinômios fatores deve ser divisível por p".

<u>TEOREMA 5.</u> (<u>FATORAÇÃO UNICA</u>)

" Todo polinômio f de $K[X]$, com $gr(f) \geqslant 1$, pode ser fatorado no produto de polinômios irredutíveis em $K[X]$. Ademais, se tivermos duas tais fatorações, digamos :-

(10) $$f = p_1 p_2 \cdots p_r$$

(11) $$f = q_1 q_2 \cdots q_s ;$$

então $r = s$, isto é, ambas as decomposições contém o mesmo número de fatores (à menos de fatores constantes), e renumerando-se convenientemente (se necessário) os polinômios q_i, teremos que :-

$$p_i = c_i q_i \text{ ,com } c_i \in K, \text{ para todo } i."$$

Demonstração.

Caso f seja irredutível, nada temos a demonstrar. Seja f redutível e usaremos indução sôbre $gr(f)$. Para $gr(f) = 1$, temos que f é irredutível e temos o desejado. Suponhamos então f na forma :-

$$f = gh, \quad \text{com } gr(g) < gr(f) \text{ e } gr(h) < gr(f) .$$

Pela hipótese de indução g e h se exprimem como produto de polinômios irredutíveis, e assim f também o faz. Provamos a existência da fatoração.

Passemos agora a unicidade. de (10) e (11) temos que :-

$$p_1 p_2 \ldots p_r = q_1 q_2 \ldots q_s .$$

Renumerando-se os q_i (se for necessário), podemos supor que $p_1 = c_1 q_1$ para algum $c_1 \in K$. Cancelando q_1, obtemos :-

$$c_1 p_2 \ldots p_r = q_2 \ldots q_s .$$

Repetindo-se essa argumentação indutivamente, concluímos que existem constantes $c_i \in K$, tais que,

$$p_i = c_i q_i \text{ , para todo } i .$$

c.q.d.

Observação. Na decomposição de um polinômio em fatores irredutíveis, os fatores são determinados à menos de um fator constante. Essa ambiguidade pode ser removida, exigindo-se que cada fator tenha coeficiente líder igual à 1. Temos assim a seguinte extensão do teorema 5.

TEOREMA 6.

" Seja f um polinômio unitário com $gr(f) \geq 1$, de $K[X]$. Então f pode ser escrito como um produto de fatores irredutíveis, cada um dos quais é unitário. Essa decomposição de f é univocamente determinada à menos da ordem dos fatores."

Na decomposição (10) pode ocorrer que $p_i = c\, p_j$, mesmo para $i \neq j$. Combinando os fatores idênticos, podemos escrever f (que por simplicidade suporemos f unitário) na forma :-

$$(12) \qquad f = q_1^{a_1} . q_2^{a_2} \ldots q_t^{a_t} ,$$

onde $q_i \neq q_j$, se $i \neq j$. Os números naturais a_i indicam o número de vêzes que o fator q_i aparece na fatoração de f.

A.6 MÍNIMO MÚLTIPLO COMUM.

Vamos agora investigarmos os múltiplos comuns de dois (ou mais) polinômios não nulos f e g de $K[X]$, isto é, polinômios que são divisíveis simultaneamente por f e g. Seja $\mathcal{Z}$ o conjunto de todos polinômios [illegible]. Claramente $\mathcal{Z}$ é um ideal de $K[X]$. [illegible]

[illegible] $\mathcal{Z}$ [illegible]

polinômio denomina-se MÍNIMO MÚLTIPLO COMUM (m.m.c.), de f e g. Como $fg \in \mathcal{C}$, temos que

(13) $fg = md$, para algum de $K[X]$.

Desde que $m \in \mathcal{C}$, êle é um múltiplo comum de f e g. Então nós temos,

(14) $m = f h_1$ e $m = g h_2$.

De (13) e (14) vemos que :-

$$fg = f h_1 d \text{ ou } f.(g - h_1 d) = 0.$$

como $f \neq 0$ e $K[X]$ é um domínio de integridade, obtemos

(15) $g = h_1 d$.

Anàlogamente, obtemos :-

(16) $f = h_2 d$.

As equações (15) e (16) mostram que d é um divisor comum de g e f. Da definição de m.m.c. m é um polinômio de menor grau em $\mathcal{C}$. Por outro lado, se d'é o m.d.c. de f e g, então o polinômio

$$fgd' \quad \frac{fg}{d'} = \left(\frac{f}{d'}\right) g = f\left(\frac{g}{d'}\right),$$

é um múltiplo comum de g ef, logo está em $\mathcal{C}$. Agora se d for um divisor próprio d', então o polinômio $\frac{fg}{d'}$ teria grau menor que o grau de $m = \frac{fg}{d}$. Isso é impossível. Então d é o m.d.c. de f e g.

Provamos o seguinte teorema.

TEOREMA 7.

" Seja d o m.d.c. de f e g, então o polinômio $\frac{fg}{d}$ é o m.m.c. de f e g.

APÊNDICE B - MISCELÂNIA

B.1 RELAÇÃO DE EQUIVALÊNCIA.

Dado um conjunto A, não vazio, indicamos por:

$$A \times A = \{(a,b) : a,b \in A\},$$

o PRODUTO CARTESIANO de A por A. Todo subconjunto R de $A \times A$, é uma RELAÇÃO (binária) em A. Sendo R uma relação em A, se $(a,b) \in R$, diz-se que " a está relacionado com b " e denota-se por aRb. Diz-se que uma relação R em A, é uma RELAÇÃO DE EQUIVALÊNCIA em A, se as seguintes propriedades são satisfeitas:

(i) REFLEXIVA : $\forall a \in A$, aRa.

(ii) SIMÉTRICA : $\forall a,b \in A$, $aRb \rightarrow bRa$.

(iii) TRANSITIVA : $\forall a,b,c \in A$, $(aRb \wedge bRc) \Rightarrow aRc$.

Sendo R uma relação de equivalência em A e aRb, diz-se que " a é EQUIVALENTE à b ".

EXEMPLOS I

Entre números racionais,

$$\frac{a}{b} \; R \; \frac{c}{d} \iff ad=bc,$$

é uma relação de equivalência. Essa relação, é a definição usual de "igualdade " entre números racionais.

Em geometria, não se costuma dizer, que uma reta é paralela à ela mesma. Mas se assim fizermos, o conceito de "paralelismo", é uma relação de equivalência, entre retas no plano ou no espaço usual. Existem várias relações de equivalência em geometria. Por exemplo, congruência de triângulos, semelhança de triângulos etc.

Seja R uma relação de equivalência em A e $a \in A$. Denomina-se CLASSE DE EQUIVALÊNCIA determinada pelo elemento $a \in A$, ao subconjunto

$$R_a = \bar{a} = \{b \in A : aRb\}.$$

Uma relação de equivalência, é quase uma generalização de "igualdade". Sendo que no caso da igualdade, cada classe de equivalência possui um único elemento.

As propriedades destacáveis das classes de equivalência, são:

1. $\bar{a} \neq \emptyset$. Isto é óbvio, pois aRa.

2. $\bar{a} \cap \bar{b} = \emptyset$ ou $\bar{a}=\bar{b}$, isto é, duas classes de equivalência são disjuntas ou coincidentes. De fato, suponhamos que $\bar{a} \cap \bar{b} \neq \emptyset$. Seja $c \in \bar{a} \cap \bar{b}$. Fixemos, arbitrariamente, um x em $\bar{a}$. Então, xRa, aRc e cRb, usando-se duas vêzes a propriedade transitiva da relação R, obtemos xRb. Mostramos assim, que $\bar{a} \subset \bar{b}$. Logo, com raciocínio análogo se verifica que $\bar{b} \subset \bar{a}$, o que prova 2.

3. $\bigcup_{a \in A} \bar{a} = A$. De fato,

(a) $\forall x \in A$, $x \in \bar{x} \subset \bigcup_{a \in A} \bar{a} \Rightarrow A \subset \bigcup_{a \in A} \bar{a}$.

[illegible]

Com essas três propriedades das classes de equivalência, vê-se fàcilmente, que uma relação de equivalência em um conjunto A, determina uma " partição " (as classes de equivalência) no conjunto A, no seguinte sentido:

1) Todo elemento de A, está em alguma classe de equivalência.
2) Dois elementos de A, estão na mesma classe de equivalência se, e sòmente se, êles são equivalentes.
3) Classes de equivalência diferentes, são disjuntas.

Por outro lado, uma partição de um conjunto A, em subconjuntos disjuntos, pode ser usada para se definir uma relação de equivalência R em A. Basta considerarmos aRb, quando a e b são elementos do mesmo subconjunto da partição.

B.2 SOMAS DIRETAS.

Seja V um K-espaço vetorial de dimensão finita n. Sejam V_1 e V_2 subespaços vetorias de V. Então,

$$V_1 + V_2 = \{ v_1 + v_2 : v_1 \in V_1 \text{ e } v_2 \in V_2 \},$$

é um subespaço de V, denominado SOMA dos subespaços V_1 e V_2. Quando V_1 e V_2 forem disjuntos, diremos que o subespaço $W = V_1 + V_2$, é a SOMA DIRETA de V_1 e V_2 e denotamos $W = V_1 \oplus V_2$. Nêsse caso, um vetor $w \in W$, se escreve de modo único na forma:

$$(1) \qquad w = v_1 + v_2 \; , \; v_1 \in V_1 \quad \text{e} \quad v_2 \in V_2 \; .$$

De fato, suponhamos que $w = v_1' + v_2'$, $v_1' \in V_1$ e $v_2' \in V_2$.

De (1), temos que :

$$v_1 + v_2 = v_1' + v_2' ,$$

ou

$$v_1 - v_1' = v_2' - v_2 ,$$

como $(v_1 - v_1') \in V_1$, $v_2' - v_2 \in V_2$ e V_1 e V_2 só têm em comum o vetor nulo temos que :

$$v_1 - v_1' = v_2' - v_2 = 0, \text{ ou } v_1 = v_1' \text{ e } v_2 = v_2' .$$

Queremos extender a definição de soma direta, para vários subespaços de V. Sejam, $V_1, V_2, \ldots, V_r$, subespaços de V e seja

$$(2) \qquad W = V_1 + V_2 + \ldots + V_r .$$

Dizemos que W, subespaço de V dado por (2), é a SOMA DIRETA de $V_1, V_2, \ldots, V_r$ e escreve-se

$$W = V_1 \oplus V_2 \oplus \ldots \oplus V_r ,$$

se para cada $j = 2, 3, \ldots, r$, o subespaço V_j é disjunto da soma $(V_1 + V_2 + \ldots + V_{j-1})$.

Pode-se verificar, que W dado por (2) é a soma direta de V_1 , V_2 , ... , V_r se, e sòmente se, cada $v \in W$, se escreve de modo único na forma:

$$v = v_1 + v_2 + \dots + v_r, \quad v_i \in V_i \ , \ i = 1,2, \dots ,r.$$

De fato, suponhamos que cada $v \in W$ se escreve de modo único na forma, $v = v_1 + v_2 + \dots + v_r$, $v_i \in V_i$, $i = 1,2, \dots ,r$. Seja $v \in V_j \cap (V_1 + V_2 + \dots + V_{j-1})$. Então, $v \in (V_1 + V_2 + \dots + V_{j-1})$, logo $v = v_1 + v_2 + \dots + v_{j-1}$ e $v \in V_j$, então a única expressão para v como uma soma de vetores, um em cada V_i , $i = 1,2, \dots ,r$, é :

$$v = 0 + 0 + \dots + 0 + \underset{(j)}{v} + 0 + \dots + 0.$$

Assim sendo, $v_1 = v_2 = \dots = v_{j-1} = 0$, e $v = v_1 + v_2 + \dots + v_{j-1} = 0$.

Para a recíproca, usaremos indução sôbre r. Para $r = 2$, já vimos que a soma direta, implica na unicidade de se es crever um vetor v de W como s oma de vetores de V_1 e V_2. Suponhamos válido para $r-1$, $r \geqslant 3$, e vamos mostrar para r. Seja :

$$v = v_1 + v_2 + \dots + v_{r-1} + v_r = w_1 + w_2 + \dots + w_{r-1} + w_r.$$

Logo,

$$(3) \quad v_1 + v_2 + \dots + v_{r-1} - w_1 - w_2 - \dots - w_{r-1} = w_r - v_r$$

e como o primeiro membro de (3) é um vetor de $V_1 + V_2 + \dots + V_{r-1}$ e o segundo é de V_r, e $V_r \cap (V_1 + V_2 + \dots + V_{r-1}) = \langle 0 \rangle$, temos que $v_r = w_r$ .Logo v fica :

$$v = v_1 + v_2 + \dots + v_{r-1} = w_1 + w_2 + \dots + w_{r-1}$$

e da hipótese de indução, $v_i = w_i$, $i = 1,2, \dots ,r-1$.

Assim, uma definição alternativa do fato de que $W = V_1 + \dots + V_r$, é a soma direta dos subespaços $V_1, \dots, V_r$, é que cada vetor $v \in W$, se escreve de modo único, na forma :

$$v = v_1 + v_2 + \dots + v_r \ , \ v_i \in V_i, \ i = 1,2, \dots ,r.$$

TEOREMA 1

" Seja V um K-espaço vetorial de dimensão finita n. Sejam $V_1, \dots, V_r$ s ubespaços de V. As afirmativas abaixo, são equivalentes:

(i) $V = V_1 \oplus V_2 \oplus \dots \oplus V_r$.

(ii) Sendo $\mathcal{B}_i$ base de V_i , $i = 1,2, \dots ,r$, então $\mathcal{B} = \bigcup_{i=1}^{r} \mathcal{B}_i$ é base de V".

Demonstração.

(i) $\Longrightarrow$ (ii).

Seja $v \in V$, qualquer. Então,

$$(4) \quad v = v_1 + v_2 + \dots + v_r, \quad v_i \in V_i, \ i = 1,2, \dots ,r,$$

de modo único. Seja $\mathcal{B}_i = \{e_{i1}, e_{i2}, \dots, e_{in_i}\}$ base de V_i, $i = 1,2, \dots, r$. Seja $\mathcal{B} = \bigcup_{i=1}^{r} \mathcal{B}_i$. De (4), temos que :-

$$v = \sum_{j=1}^{n_1} a_{1j}e_{1j} + \sum_{j=1}^{n_2} a_{2j}e_{2j} + \dots + \sum_{j=1}^{n_r} a_{rj}e_{rj},$$

isto é, $\mathcal{B}$ gera V.

Seja,

$$\sum_{j=1}^{n_1} a_{1j}e_{1j} + \sum_{j=1}^{n_2} a_{2j}e_{2j} + \dots + \sum_{j=1}^{n_r} a_{rj}e_{rj} = o = o + o + \dots + o.$$

Como o vetor nulo de V, se exprime de modo único como soma de vetores $v_1, v_2, \dots, v_r$, $v_i \in V_i$, $i = 1,2, \dots, r$, temos que :-

$$\sum_{j=1}^{n_i} a_{ij}e_{ij} = o, \; i = 1,2, \dots, r.$$

Sendo $\mathcal{B}_i = \{e_{i1}, e_{i2}, \dots, e_{in_i}\}$ base de V_i, $i = 1,2, \dots, r$, vem que $a_{ij} = o$, $i = 1,2, \dots, r$ e $j = 1,2, \dots, n_i$. Isto é, os vetores de $\mathcal{B}$ são linearmente independentes.

(ii) $\Longrightarrow$ (i)

Seja $v \in V$, qualquer. Como $\mathcal{B} = \bigcup_{i=1}^{r} \mathcal{B}_i$ é uma base de V, temos que : -

$v = a_{11}e_{11} + a_{12}e_{12} + \dots + a_{1n_1}e_{1n_1} + \dots + a_{r1}e_{r1} + \dots + a_{rn_r}e_{rn_r} =$
$v_1 + v_2 + \dots + v_r$, $v_i = (a_{i1}e_{i1} + \dots + a_{in_i}e_{in_i}) \in V_i$, $i = 1,2,\dots,r$.

Mostraremos que tal soma é única. Suponhamos que :-

$$v = w_1 + w_2 + \dots + w_r, \; w_i \in V_i, \; i = 1,2, \dots, r.$$

Desde que $\{e_{i1}, e_{i2}, \dots, e_{in_i}\}$ é uma base de V_i, $w_i = b_{i1}e_{i1} + \dots + \dots + b_{in_i}e_{in_i}$ e então :-

$$v = b_{11}e_{11} + \dots + b_{1n_1}e_{1n_1} + \dots + b_{r1}e_{r1} + \dots + b_{rn_r}e_{rn_r}.$$

Sendo $\mathcal{B}$ uma base de V, $a_{ij} = b_{ij}$, para cada i e cada j. Então, $v_i = w_i$, $i = 1,2, \dots, r$ e v se escreve de modo único como soma de vetores $v_1, \dots, v_r$, $v_i \in V_i$; isto é,

$$V = V_1 \oplus V_2 \oplus \dots \oplus V_r.$$

c.q.d.

Pode-se concluir do teorema 1, que se :-

$$V = V_1 \oplus V_2 \oplus \dots \oplus V_r,$$

então $\dim V = \dim V_1 + \dim V_2 + \dots + \dim V_r$.

Seja L um operador linear de V. Durante tôda essa dissertação, destacou-se " decomposições " de V em soma direta :-

$$V = V_1 \oplus V_2 \oplus \dots \oplus V_r ,$$

onde cada V_i é L-invariante. Com uma tal decomposição, a ação de L, fica determinada, pelas ações de L_{V_1} , L_{V_2} , ... , L_{V_r} , restrições de L aos subespaços V_1, V_2, ... , V_r. Pois, se $v \in V$,

$$v = v_1 + v_2 + \dots + v_r ,$$

de modo único, isto é, os $v_i \in V_i$ são bem determinados e

$$Lv = L_{V_1}v_1 + L_{V_2}v_2 + \dots + L_{V_r}v_r .$$

Nêsse caso, diz-se que L é a SOMA DIRETA dos operadores $L_{V_1}, L_{V_2}, \dots, L_{V_r}$ dos subespaços $V_1, V_2, \dots, V_r$, respectivamente. Denota-se êsse fato pc

$$L = L_{V_1} \oplus L_{V_2} \oplus \dots \oplus L_{V_r} .$$

O análogo dessa situação para matrizes, pode ser considerado da seguinte maneira. Seja $\mathcal{B}_i$ uma base ordenada para V_i , $i = 1,2, \dots ,r$. O teorema 1, nos afirma que se $V = V_1 \oplus V_2 \oplus \dots \oplus V_r$, $\mathcal{B} = \bigcup_{i=1}^{r} \mathcal{B}_i$ é uma base ordenada de V. Assim é fácil de ver-se que, sendo :-

$$A = [L]_{\mathcal{B}} \quad e \quad A_i = [L_{V_i}]_{\mathcal{B}_i} , \; i = 1,2, \dots , r,$$

então A tem a seguinte " forma em blocos " :-

$$A = \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & A_r \end{bmatrix}$$

Nêsse caso, diz-se que A é a SOMA DIRETA DAS MATRIZES $A_1, A_2, \dots, A_r$; e denota-se :-

$$A = A_1 \oplus A_2 \oplus \dots \oplus A_r = [A_1, A_2, \dots , A_r]$$

B.3 ESPAÇO VETORIAL QUOCIENTE.

Seja W um subespaço vetorial de V. Diz-se que $u,v \in V$, são CONGRUOS MODULO W se, e sòmente se, $(v-u) \in W$. Denota-se êsse fato por:-

$$u \equiv v \pmod{W}.$$

Observa-se a analogia dessa definição, com a congruência módulo m, no conjunto $\mathbb{Z}$ dos números inteiros.

Vamos verificar que congruência módulo W, é uma relação de equivalência, em V.

(i) Reflexiva. $v \equiv v$ (mod. W), pois $v-v = 0 \in W$.
(ii) Simétrica. Seja $v \equiv u$ (mod. W), isto é, $(u-v) \in W$, logo $-(u-v) = (v-u) \in W$, ou seja, $u \equiv v$ (mod. W).
(iii) Transitiva. Sejam $u \equiv v$ (mod. W) e $v \equiv z$ (mod. W), isto é, $(v-u)$, $(z-v) \in W$, sendo W um subespaço vetorial, $(v-u)+(z-v) = (z-u) \in W$, logo $u \equiv z$ (mod. W).

Assim, congruência módulo W, determina uma partição de V em classes de equivalência, sendo que dois vetores $u, v \in V$, estão na mesma classe de equivalência se, e sòmente se, $u \equiv v$ (mod. W). Denota-se à classe de equivalência contendo o vetor v, por:

$\bar{\bar{v}}$ ou por $v+W$, e têm-se que:

$$\bar{\bar{v}} = v+W = \{v+w : w \in W\} .$$

Seja, $V/W = \{v+W : v \in V\}$. Define-se em V/W, as operações de adição e multiplicação por escalar, pondo-se:-

(i) Adição: $(u+W)+(v+W) = (u+v)+W$ ou $\bar{\bar{u}}+\bar{\bar{v}} = \overline{\overline{u+v}}$

(i) Multiplicação por escalar: $a(u+W) = (au)+W$ ou $a\bar{\bar{u}} = \overline{\overline{au}}$.

Deve-se verificar que essas operações estão bem definidas, isto é, independem do particular representante de cada classe de equivalência. De fato, seja $\bar{\bar{u}} = \bar{\bar{u}}_1$ e $\bar{\bar{v}} = \bar{\bar{v}}_1$. Então, $(u-u_1) \in W$ e $(v-v_1) \in W$, e sendo W subespaço, $(u-u_1)+(v-v_1) = (u+v) - (u_1+v_1) \in W$, isto é, $\overline{\overline{u+v}} = \overline{\overline{u_1+v_1}}$. Ainda, de $(u-u_1) \in W$, temos que, $a(u-u_1) = (au-au_1) \in W$, istoé, $\overline{\overline{au}} = \overline{\overline{au_1}}$.

É fácil de se verificar, que com essas operações, V/W é um K-espaço vetorial, denominado <u>ESPAÇO VETORIAL QUOCIENTE</u> de V por W, cujo elemento zero é $W = \bar{\bar{0}}$.

As operações no espaço vetorial V/W, foram definidas, de modo que a <u>APLICAÇÃO NATURAL</u> :-

$$\theta : V \longrightarrow V/W$$

$$v \longmapsto \theta v = v+W = \bar{\bar{v}} ,$$

é linear. De fato, $\theta(av+u) = \overline{\overline{av+u}} = \overline{\overline{av}}+\bar{\bar{u}} = a\,\bar{\bar{v}}+\bar{\bar{u}} = a\,\theta v+\theta u$.

<u>TEOREMA 2.</u>

" Seja V um K-espaço vetorial, de dimensão finita n, e W um subespaço vetorial de V. Seja $\{w_1, w_2, \ldots, w_r\}$ uma base ordenada de W e $\{\bar{\bar{v}}_1, \bar{\bar{v}}_2, \ldots, \bar{\bar{v}}_s\}$ uma base ordenada de V/W. Então, $\{\bar{\bar{v}}_1, \bar{\bar{v}}_2, \ldots, \bar{\bar{v}}_s, w_1, \ldots, w_r\}$ é uma base ordenada de V. Em particular,

$$\dim V = \dim W + \dim (V/W)."$$

Demonstração.

Seja $v \in V$, qualquer. Desde que, $\{\bar{\bar{v}}_1, \bar{\bar{v}}_2, \ldots, \bar{\bar{v}}_s\}$ é uma base ordenada de V/W, temos que :-

$$\bar{\bar{v}} = a_1\bar{\bar{v}}_1 + a_2\bar{\bar{v}}_2 + \ldots + a_s\bar{\bar{v}}_s ,$$

daí, $v = a_1v_1 + a_2v_2 + \ldots + a_sv_s + w$, onde $w \in W$. Sendo $\{w_1, w_2, \ldots, w_r\}$ base ordenada de W, temos:-

$v = a_1v_1 + a_2v_2 + \dots + a_sv_s + b_1w_1 + b_2w_2 + \dots + b_rw_r$;
isto é, $\{v_1, \dots, v_s, w_1, \dots, w_r\}$ geram V. Suponhamos que:-

(5) $a_1v_1 + a_2v_2 + \dots + a_sv_s + b_1w_1 + b_2w_2 + \dots + b_rw_r = o$,

então,

$$a_1\overline{v}_1 + a_2\overline{v}_2 + \dots + a_s\overline{v}_s + b_1\overline{w}_1 + b_2\overline{w}_2 + \dots + b_r\overline{w}_r = \overline{o} = W ,$$

e como $w_i \in W$, $i = 1,2, \dots, r$, $\overline{w}_i = \overline{o}$, e daí:-

(6) $a_1\overline{v}_1 + a_2\overline{v}_2 + \dots + a_s\overline{v}_s = \overline{o} = W$.

Sendo $\{\overline{v}_1, \overline{v}_2, \dots, \overline{v}_s\}$ linearmente independentes, pois é uma base de V/W, de (6) vemos que :-

$$a_1 = a_2 = \dots = a_s = o.$$

Daí (5) se reduz à :-

$$b_1w_1 + b_2w_2 + \dots + b_rw_r = o,$$

e sendo $\{w_1, w_2, \dots, w_r\}$ base de W, temos que:-

$$b_1 = b_2 = \dots = b_r = o .$$

Isto é, $\{v_1, v_2, \dots, v_s, w_1, w_2, \dots, w_r\}$ são linearmente independentes.

c.q.d.

TEOREMA 3.

" Sejam V um K-espaço vetorial de dimensão finita n, L um operador linear de V e W um subespaço de V, L-invariante. Então, L induz um operador linear $\overline{L}$ em V/W, definido por $\overline{L}\,\overline{v} = \overline{Lv}$. Sendo L, zero de algum polinômio, então $\overline{L}$ também será zero dêsse polinômio. Em particular, o polinômio minimal de $\overline{L}$, divide o polinômio minimal de L."

Demonstração.

Devemos mostrar, de início, que $\overline{L}$ esta bem definida. Isto é, se $\overline{u} = \overline{v}$, então $\overline{L}\,\overline{u} = \overline{L}\,\overline{v}$. Sendo, $\overline{u} = \overline{v}$, temos que, $(u-v) \in W$, e como W é L-invariante, L $(u-v) = (Lu-Lv) \in W$, ou seja, $Lv \equiv Lu$ (mod. W) ou $\overline{Lu} = \overline{Lv}$. Daí, $\overline{L}\,\overline{u} = \overline{Lu} = \overline{Lv} = \overline{L}\,\overline{v}$.

Mostraremos agora que $\overline{L}$ é linear. De fato, $\overline{L}(a\overline{u} + \overline{v}) = \overline{L(au + v)} = \overline{(aLu + Lv)} = a\overline{Lu} + \overline{Lv} = a\,\overline{L}\,\overline{u} + \overline{L}\,\overline{v}$.

Passemos a asserção final. Para todo $\overline{u} \in V/W$, temos que :-
$\overline{L^2}\,\overline{u} = \overline{L^2 u} = \overline{L(Lu)} = \overline{L}(\overline{Lu}) = (\overline{L})^2u$. Então, $\overline{L^2} = (\overline{L})^2$. Anàlogamente, ou sendo mais preciso, por indução, conclui-se que $\overline{L^n} = (\overline{L})^n$, para todo inteiro positivo n. Assim, para qualquer polinômio --

$$f(X) = a_nX^n + a_{n-1}X^{n-1} + \dots + a_1X + a_o = \sum_{i=1}^{n} a_iX^i,$$

temos que :-

$$\overline{f(L)}\,\overline{u} = \overline{f(L)u} = \overline{\sum_{i=o}^{n} a_iL^iu} = \sum_{i=o}^{n} a_i\overline{L^iu} = \sum_{i=o}^{n} a_i\overline{L^i}\,\overline{u} = \sum_{i=o}^{n} a_i(\overline{L})^i\,\overline{u} =$$

$$\left[\sum_{i=o}^{n} a_i(\overline{L})^i\right]\overline{u} = f(\overline{L})\,\overline{u}.$$ Assim, concluímos que :- $\overline{f(L)} = f(\overline{L})$.

Portanto, se L é zero de $f(X)$, temos que:-

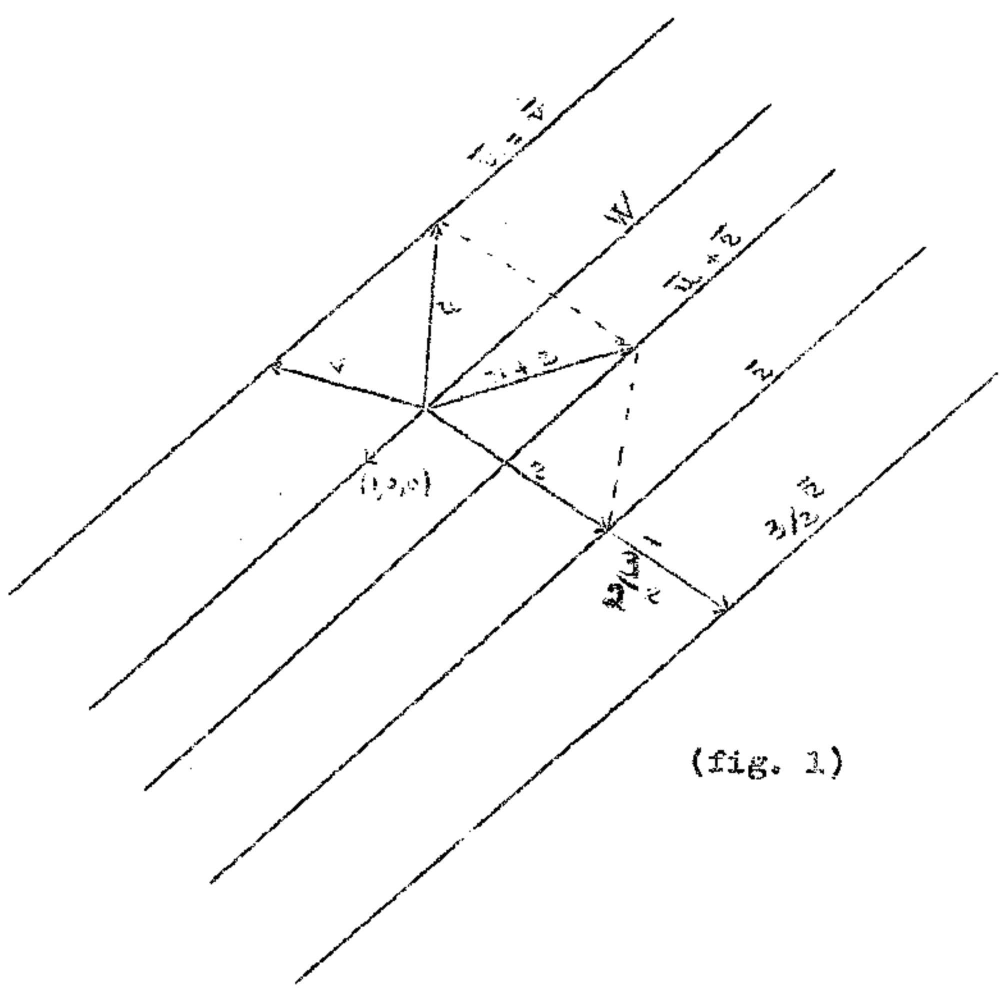
$$\overline{f(L)} = \overline{0} = W = \bar{f}(\bar{L}),$$

isto é, $\bar{L}$ também é zero de $f(X)$, como queríamos monstrar.

Vamos ver um exemplo, que ilustrará os concitos instroduzidos.

EXEMPLO 2.

Seja $V = R^3$ e $W = \langle(1,0,0)\rangle$, isto é, W é a reta passando pelos pontos $(0,0,0)$ e $(1,0,0)$ de R^3. Para melhor clareza geométrica, representaremos o elemento (a,b,c) de R^3, pelo segmento orientado com início no ponto $(0,0,0)$ e extremidade em (a,b,c) (ver fig.1). A congruência, $u \equiv v$ (mod. W), significa que u e v diferem por elemento de W, isto é, o segmento de pontos finais u e v, é paralelo à W. A classe $\bar{v}$ é representada pela reta paralela à W, passando pelo ponto v (mais precisamente, pelo "feixe" de vetores aplicados na origem, e com extremos nessa reta). Feixes podem ser adicionados e também pode-se multiplicar um feixe por um escalar real, para tanto, basta adicionármos vetores dos feixes, ou multiplicarmos um vetor do feixe pelo escalar. Êsses feixes, são exatamente os elementos de V/W. Pelo teorema 1, $\dim(V/W) = \dim V - \dim W = 3 - 1 = 2$, e $\{\overline{(0,1,0)}, \overline{(0,0,1)}\}$ é uma base ordenada de V/W. A aplicação natural, associa a cada (a,b,c) de R^3 a reta de R^3, paralela a W passando por (a,b,c).



(fig. 1)

Lembramos que, a menos que se explicite o contrário, V denota um K-espaço vetorial de dimensão finita n ($n \geq 1$) e L um operador linear fixo (mas arbitrário) de V.

Seja v um vetor, não nulo, arbitrário de V. Consideremos a seqüência de vetores :-

(7) $v, Lv, L^2v, L^3v, \ldots, L^nv?$ ......

Sendo dim V = n finita, existe um número finito e inteiro r, $o \leq r \leq n$, tal que os vetores $v, Lv, L^2v, \ldots, L^{r-1}v$, são linearmente independentes, e L^rv é uma combinação linear dêsses vetores com coeficientes em K, isto é,

(8) $$L^rv = -a_1L^{r-1}v - a_2L^{r-2}v - \ldots - a_{r-1}Lv - a_rv.$$

Consideremos o polinômio unitário :-

(9) $$p(X) = X^r + a_1X^{r-1} + \ldots + a_{r-1}X + a_r .$$

Então, pode-se reescrever (8), pondo-se :-

(1o) $$p(L)v = o .$$

Consideremos o conjunto :-

$$M(v,L) = \{ f \in K[X] : f(L)v = o \} .$$

É claro que M(v,L) é não vazio e tem polinômio não nulo, pois o polinômio minimal m(X) do operador linear L é tal que m(L)v = o. Ademais, M(v,L) é um ideal de K[X], pois para todo $f,g \in M(v,L)$ e todo $a \in K$, tem-se que :-

$$[(af+g)(L)]v = af(L)v + g(L)v = ao + o = o ,$$

isto é, $(af+g) \in M(v,L)$. Como todo ideal de K[X] é pricipal, existe em M(v,L) um único polinômio unitário $m_v(X)$ de menor grau, gerador de M(v,L). O polinômio m_v, é denominado o L-ANULADOR do vetor v.

Como ressaltamos acima, o polinômio minimal m do operador linear L, está em M(v,L), logo m_v divide m, pois todo polinômio $f \in M(v,L)$ é divisível por m_v.

Queremos mostrar, que o L-anulador de v, é o polinômio unitário p dado por (9). Basta mostrarmos que o polinômio p é o de menor grau, tal que p(L)v = o. Suponhamos que exista um polinômio não nulo $f(X) = X^s + b_1X^{s-1} + \ldots + b_s$, com $gr(f) = s < r = gr(p)$ e $f(L)v = o$. Então,

$$L^sv + b_1L^{s-1}v + \ldots + b_{s-1}Lv + b_sv = o,$$

o que é absurdo, pois $s < r$ logo $v, Lv, \ldots, L^sv$ são linearmente independentes (tôda subcoleção da coleção $v, Lv, \ldots, L^{r-1}v$ é linearmente independente, pois essa coleção o é).

Em particular, temos que $gr(m_v) > o$, isto porque na formação de (7), tomamos v não nulo.

Seja $\{e_1, e_2, \ldots, e_n\}$ uma base de V. Sejam $m_1, m_2, \ldots, m_n$ os L-anuladores dêsses vetores básicos $e_1, e_2, \ldots, e_n$, respectivamente. Seja m_V o m.m.c. (unitário) dos L-anuladores $m_1, m_2, \ldots, m_n$. Então, $m_V \in M(e_i, L)$, $i = 1, 2, \ldots, n$. De fato, para cada $i = 1, \ldots, n$ existe $h_i \in K[X]$, tal que, $m_V = h_i m_i$, logo $m_V(e_i) = h_i(e_i)\, m_i(e_i) = h_i(e_i).o = o$. Desde que para todo $v \in V$, $v = a_1e_1 + a_2e_2 + \ldots + a_ne_n$, temos que :-

(11) $\qquad m_V(L)v = a_1m_V(L)e_1 + \ldots + a_nm_V(L)e_n = o.$

Seja g em $K[X]$, tal que:-

$$g(L)v = o, \text{ para todo } v \in V.$$

Então, $g \in M(e_i, L)$, $i = 1, 2, \ldots, n$. Daí, g é um múltiplo comum dos L-anuladores $m_1, \ldots, m_n$. Logo m_V divide g, pois m_V é o m.m.c. entre $m_1, \ldots, m_n$. Em particular, como o polinômio minimal m, é tal que, $m(L)v = o$, para todo v de V, temos que m_V divide m. Por outro lado, de (11) e da definição de polinômio minimal, temos que m divide m_V. Como m e m_V são unitários, temos que êles coincidem.

Seja $m_v(X) = X^r + a_1X^{r-1} + \ldots + a_r$ o L-anulador de $v \in V$. Logo os vetores :-

(12) $\qquad \{v, Lv, L^2v, \ldots L^{r-1}v\}$,

são linearmente independentes e

(13) $\qquad L^rv = -a_rv - a_{r-1}Lv - \ldots - a_1L^{r-1}v.$

Quando $v \neq o$, os vetores (12) são linearmente independentes, logo geram um subespaço de V de dimensão r, o qual denominamos <u>SUBESPAÇO L-CICLICO GERADO</u> por v, e o denotamos por $C(v,L)$ ou simplesmente por $C(v)$. Quando $C(v) = V$, diz-se que o vetor v é um <u>VETOR CICLICO</u> relativamente à L. Ademais, para $v = o$, coloca-se $C(o) = \langle o \rangle$. A nomenclatura é sugerida por (12) e (13). Os vetores de (12) formam uma base para $C(v)$, $v \neq o$. Daí, temos que $C(v)$ é L-invariante, pois L leva o primeiro vetor básico, no segundo, o segundo no terceiro, etc. O último vetor básico ($L^{r-1}v$) é levado por L em L^rv, que por (13) é uma combinação linear de (12). Em particular, qualquer vetor $w \in C(v)$, é tal que : $w = g(L)v$, para algum $g \in K[X]$ e $gr(g) \leq r-1$.

<u>TEOREMA 4</u>.

" Seja v um vetor não nulo e arbitrário de V, cujo L-anulador é m_v. Então, temos que :-

(i) $\dim C(v) = gr(m_v)$

(ii) seja $gr(m_v) = r$. Então os r-vetores; $\{v, Lv, \ldots, L^{r-1}v\}$, formam uma base para V.

(iii) m_v é o polinômio mínimo do operador linear $T = L_{C(v)}$, induzido em $C(v)$ por L ".

Demonstração.

(i) e (ii) são óbvias, e já as observamos acima.

Passemos à (iii). Seja $g \in K[X]$, arbitrário. Então,

$m_v(T)g(L)v = m_v(L)g(L)v = g(L)m_v(L)v = g(L)o = o.$

Assim, $m_v(T)$ leva todo vetor de $C(v)$ no vetor nulo, pois levará, cada vetor básico, $\{v, Lv, L^2v, \dots, L^{r-1}v\}$, no vetor nulo. Basta tomar g como sendo os polinômio $1, X, X^2, \dots, X^{r-1}$, respectivamente. Além disso, se f é tal que, $gr(f) < gr(m_v)$, não se pode ter $f(T) = 0$. De fato, se $f(T) = 0$, $f(T)v = f(L)v = o$ e $gr(f) < gr(m_v)$ contrariando a definição de L-anulador. Isto mostra que m_v é o polinômio mínimo de T.

c.q.d.

Uma consequência do teorema 4, é que se v é um vetor cíclico de L, então $C(v) = V$, e daí $gr(m_v) = n$ e $m_v = m = c$, isto é, os polinômios minimal e característico coincidem. Demonstraremos a seguir que existe um vetor v em V, cujo L-anulador é o polinômio mínimo de L. Logo, teremos que : " L possui um vetor cíclico se , e sòmente se, os polinômios mínimo e característico de L são idênticos ".

LEMA 1

" Suponhamos que o L-anulador dos vetores v_1 e v_2 são relativamente primos. Então, o L-anulador do vetor $v = v_1 + v_2$ é igual ao produto dos L-anuladores de v_1 e v_2 ".

Demonstração.

Sejam m_1 e m_2 os L-anuladores de v_1 e v_2 ,respetivamente. Seja $g \in M(v,L) = M(v_1 + v_2, L)$, arbitrário. Logo, $g(L)v = g(L)(v_1 + v_2) = o$

Então, $v_1 = v - v_2$ e temos que :-

$$m_2(L)g(L)v_1 = m_2(L)g(L)v - g(L)m_2(L)v_2 = o,$$

isto é, $m_2g \in M(v_1, L)$. Daí, m_2g é divisísivel por m_1, e como m.d.c. entre m_1 e m_2 é 1, g é divisível por m_1. Com os mesmos argumentos, invertemdo-se os papéis de m_1 e m_2, chega-se que g é divisível por m_2. Sendo m_1 e m_2 relativamente primos, e g divisível por m_1 e m_2, temos que g é divisível por m_1m_2. Provamos que todo $g \in M(v,L)$ é divisível por m_1m_2, em particular, m_v é divisível por m_1m_2. Como

$[m_1(L)\,m_2(L)]\,v = m_1(L)\,m_2(L)v_2 + m_2(L)\,m_1(L)v_1 = o + o = o,$

m_1m_2 divide m_v. Sendo m_1, m_2 e m_v unitários, temos que $m_v = m_1m_2$.

É claro que o lema 1, se extende para o caso de s-vetores, $v_1, v_2, \ldots, v_s$, com L-anuladores $m_1, m_2, \ldots m_s$, primos dois à dois.

LEMA 2.

" Suponhamos que o polinômio mínimo m de L, é uma potência de um polinômio irredutível p em $K[X]$. Então, existe um vetor v em V, cujo L-anulador é m".

Demonstração.

Seja $m = p^s$. Seja $\{e_1, e_2, \ldots, e_n\}$ uma base de V. Sejam $m_1, \ldots, m_n$, os L-anuladores de $e_1, \ldots, e_n$, respectivamente. Sendo o L-anulador m_i de e_i, um divisor dem e $gr(m_i) > o$, tem-se que :-

$$m_i = p^{s_i} \quad , \text{ onde } o < s_i \leq s \; , \; i = 1,2, \ldots, n.$$

Mas m é o m.m.c. entre $m_1, \ldots, m_n$, logo $m = p^{s_j}$, onde $\max\{s_1, \ldots, s_n\} = s_j$, e o polinômio mínimo deL, é o L-anulador de e_j, isto é, $s_j = s$.

TEOREMA 5.

" Sejam V um K-espaço vetorial de dimensão finita $n \geq 1$ e L um operador linear (arbitrário) de V. Existe um vetor de V, cujo L-anulador é o polinômio mínimo de V ".

Demonstração.

Consideremos o polinômio mínimo m de L, fatorado na forma :-

$$m = p_1^{a_1} \, p_2^{a_2} \, \ldots \, p_r^{a_r} \, ,$$

onde os p_i são polinômios de $K[X]$ irredutíveis, unitários e dois à dois distintos, e os a_i são inteiros extritamente positivos. Então, pelo Teorema da Decomposição Primária (teorema 4 de 2.3), temos que-

$$V = V_1 \oplus V_2 \oplus \ldots \oplus V_r \, ,$$

onde cada V_i é o núcleo de $[p_i(L)]^{a_i}$, $i = 1,2, \ldots, r$, e p_i é o polinômio mínimo de L_{V_i}. Pelo lema 2, para cada $i = 1,2, \ldots, r$, existem vetores $v_i \in V_i$ cujo L-anulador é $[p_i(L)]^{a_i}$. Seja $v = v_1 + \ldots + v + \ldots + v_r$, pela extensão óbvia do lema 1, temos que

$$m_v = p_1^{a_1} \, p_2^{a_2} \, \ldots \, p_r^{a_r} = m$$

c.q.d.

EXEMPLO 3.

Seja L um operador linear arbitrário de V. O subespaço $C(v,L)$ é unidimensional se, e sòmente se, v é um autovetor de L. De fato, $\dim C(v,L) = 1$ se, e sòmente se, v e Lv são linearmente dependentes ($v \neq o$), ou seja, $Lv = \lambda v$, para algum $\lambda \in K$. Nêsse caso, $C(v,L) = \langle v \rangle$. Temos assim, muitos exemplos de subespaços L-cíclicos. Também podemos concluir, que se L é um múltiplo escalar ($\lambda \neq 0$) do operador idêntico de V e $\dim V > 1$, então L não possui nenhum vetor cíclico, pois $C(o,L) = \langle o \rangle$ e sendo $v \neq o$, $Lv = \lambda v$ ($L = \lambda I$, $\lambda \neq o$) logo $\dim C(v,L) = 1$, isto é, $C(v,L) \neq V$.

Consideremos $L : K^2 \longrightarrow K^2$, tal que

$$[L]_{\mathcal{B}} = \begin{bmatrix} o & o \\ 1 & o \end{bmatrix}, \text{ onde } \mathcal{B} = \{(1,o), (o,1)\} = \{e_1, e_2\}.$$

Então $e_1 = (1,o) \in K^2$ é um vetor cíclico de L, pois para todo $v = ae_1 + be_2$ de K^2, $v = g(L)e_1$ onde $g(X) = a + bX$, daí $v \in C(e_1)$. Como $Le_2 = oe_2$, temos que $C(e_2) = \langle e_2 \rangle$.

Seja L um operador linear de V e W um subespaço de V, L-invariante. Consideremos o espaço vetorial quociente V/W. Diz-se que, os vetores $v_1, v_2, \ldots, v_r$ de V são LINEARMENTE DEPENDENTES MÓDULO W se, e sòmente se, existem escalares $a_1, a_2, \ldots, a_r$, não todos nulos em K, tal que :-

$$a_1v_1 + a_2v_2 + \ldots + a_rv_r \equiv o \pmod{W}.$$

Anàlogamente, pode-se introduzir os conceitos de L-anulador de um vetor módulo W, ou de polinômio mínimo de L módulo W. Chamaremos à êsses conceitos, " L-anulador relativo " e " polinômio mínimo relativo" respectivamente. Seja $M'(v,L) = \{g \in K[X] : g(L)v \equiv o \pmod{W}\}$, para $v \neq o$, arbitrário em V. É fácil de se verificar que $M'(v,L)$ é um ideal de $K[X]$. Denominaremos L-ANULADOR RELATIVO ou L-ANULADOR DE v MÓDULO W ao polinômio unitário m'_v, gerador do ideal $M'(v,L)$. Denomina-se POLINÔMIO MÍNIMO RELATIVO de L ou POLINÔMIO MÍNIMO de L MÓDULO W, ao polinômio unitário m' de menor grau, tal que, para todo $v \in V$:-

$$m'(L)v \equiv o \pmod{W},$$

e m' divide todo polinômio $f \in K[X]$, tal que, para todo $v \in V$:-

$$f(L)v \equiv o \pmod{W} \quad \text{(ver teorema 1, de 1.2).}$$

Ressaltamos que o L-anulador relativo de um vetor v (ou o polinômio mínimo relativo) é um divisor do L-anulador de v (ou do polinômio mínimo). Seja m_v o L-anulador de v e m'_v o L-anulador relativo de v. Então, $m_v(L)(v) = o$ e daí,

$$m_v(L)(v) \equiv o \pmod{W} \text{ , isto é,}$$

$$m_v \in M'(v, L) = \{g \in K[X] : g(L)(v) \equiv o \pmod{W}\},$$

e sendo m'_v o gerador do ideal $M'(v,L)$, m'_v divide m_v.

Todos os resultados válidos para os conceitos acima, continuam válidos para os conceitos relativos, inclusive o teorema 5, cuja versão é :- " Em todo espaço vetorial V, existe um vetor cujo L-anulador relativo coincide com o polinômio mínimo relativo de L".

Os resultados continuam válidos para os conceitos relativos, pois que as operação com congruência módulo W, são essencialmente com igualdades em V/W.

B I B L I O G R A F I A.

[1] Ficken, F. A., Linear Transformations and Matrices - Prentice-Hall, Inc., Englewood Cliffs - 1 967.

[2] Finkbeiner II, D. T., Introduction to Matrices and Linear Transformations - W. H. Freeman and Company, San Francisco and London

[3] Gantmacher, F. R., The Theory of Matrices - Chelsea Publishing Company, New York - 1 959.

[4] Gel'fand, I. M., Lectures on Linear Algebra - Interscience Publishers, Inc., New York - 1 950.

[5] Halmos, P., Finite-Dimensional Vector Spaces - D. Van Nostrand Co., Princeton - 1 958.

[6] Hoffman, K. and Kunze, R., Linear Algebra - Prentice-Hall Inc., Englewood Cliffs, N. J. - 1 961.

[7] Lang, S., Algebraic Structures - Addison Wesley Publishing Co., New York - 1 967.

[8] Lipschutz, S., Linear Algebra - Schaum's Outline Series, McGraw-Hill Book Company - 1 968.

[9] Nering, E. D., Linear Algebra and Matrix Theory - John Wiley & Sons, Inc., New York - 1 963.

[10] Schreier, O. and Sperner, E., Modern Algebra and Matrix Theory - Chelsea Publishing Co., New York - 1 955.

[11] Stoll, R. R., Linear Algebra and Matrix Theory - McGraw-Hill Book Company, Inc. - 1 952.

Artigos em revistas :-

[1] Brand, L., The Companion Matrix and Its Properties - The American Mathematical Montly, vol. 71 (1964) pág. 629-634.

[2] Su, C. K., An Algorithm for Putting a Matrix Into a Jordan Form - Pre-Print.