ÁLGEBRA LINEAR. NOTAS DE AULAS (ICMC-USP SÃO CARLOS)

$2^{\underline{o}}$ SEMESTRE DE 2015

1. Notação de aplicações e conjuntos

Sejam A e B dois conjuntos de natureza qualquer. Uma *aplicação* $f : A \to B$ de A para B é uma lei pela qual a cada elemento $a \in A$ está associado um único elemento $b \in B$, denotado por $b = f(a)$ e chamado a *imagem* de a. Escreve-se também $A \xrightarrow{f} B$. Dizemos que A é o *domínio* e B é o *contra-domínio* de f. Note que, por definição, consideramos duas aplicações $f : A \to B$ e $g : A \to C$ como diferentes se os contra-domínios são diferentes, isto é, $B \neq C$, mesmo se a lei parece a mesma. Por exemplo, denotando[1] $\mathbb{R} := \{r \mid r \text{ é um número real}\}$ e $\mathbb{R}^+ := \{r \in \mathbb{R} \mid r \geq 0\}$, temos duas aplicações $f : \mathbb{R} \to \mathbb{R}$, $f : r \mapsto r^2$, e $g : \mathbb{R} \to \mathbb{R}^+$, $g : r \mapsto r^2$, que são diferentes embora dadas através da mesma lei $r \mapsto r^2$. (Escrevendo $f : a \mapsto b$ enfatizamos como f "age" sobre elementos; isto é equivalente a escrever $f(a) = b$.)

Utilizamos a seguinte notação de conjuntos. O símbolo $\varnothing$ denota o conjunto vazio, isto é, sem elementos. (É razoável imaginar um conjunto como um saco de coisas. Neste caso, o vazio é um saco vazio.) Quando A está contido em B, ou seja, quando B contém A, escrevemos $A \subset B$ ou $B \supset A$. Para verificar que $A \subset B$, precisa-se provar a implicação $a \in A \Longrightarrow a \in B$. Dois conjuntos A e B são considerados como iguais se eles têm os mesmos elementos. Em outras palavras, $A = B$ é equivalente a $A \subset B$ e $A \supset B$. Por exemplo, $p \neq \{p\}$ para qualquer conjunto p. Em particular, o conjunto $\{\varnothing\}$ não é vazio.

Sejam A e B conjuntos. Denotamos por $A \cap B$ a *interseção* de A e B, isto é, $A \cap B := \{x \mid x \in A,\ x \in B\}$. Denotamos por $A \cup B := \{x \mid x \in A \text{ ou } x \in B\}$ a *união* de A e B. Denotamos por $A \times B := \big\{(a, b) \mid a \in A, b \in B\big\}$ o produto *cartesiano* de A e B. Este produto é formado por todos os pares *ordenados* (a, b), onde $a \in A$ e $b \in B$. Não precisa saber o que é um par ordenado. É suficiente saber apenas a propriedade que caracteriza este conceito: $(a, b) = (a', b') \Longleftrightarrow a = a'$ e $b = b'$. De modo análogo, podemos definir o produto cartesiano $A_1 \times A_2 \times \dots \times A_n$ de conjuntos $A_1, A_2, \dots, A_n$. Seja A um conjunto e sejam $S, S' \subset A$ subconjuntos. Denotamos por $S \setminus S' := \{s \in S \mid s \notin S'\}$ o *complemento* de S' em S.

Seja $f : A \to B$ uma aplicação e sejam $A' \subset A$ e $B' \subset B$. Então $f(A') := \big\{f(a) \mid a \in A'\big\}$ é a *imagem* de A' por f e $f^{-1}(B') := \big\{a \in A \mid f(a) \in B'\big\}$ é a imagem *inversa* de B' por f. Definimos a *restrição* $f|_{A'} : A' \to B$ de f para A' pela regra óbvia $f|_{A'} : a' \mapsto f(a')$. A aplicação de *inclusão* $i : A' \hookrightarrow A$ é dada pela regra $i : a' \mapsto a'$ para todo $a' \in A'$.

Sejam $A \xrightarrow{f} B \xrightarrow{g} C$ duas aplicações dos formatos indicados. Definimos a aplicação *composta* ou a *composição* $g \circ f : A \to C$ pela regra $(g \circ f)(a) := g\big(f(a)\big)$ para todo $a \in A$. Essa operação é associativa: é fácil verificar que $(h \circ g) \circ f = h \circ (g \circ f)$ para aplicações $A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D$. Podemos observar também que a restrição $f|_{A'}$ da aplicação $f : A \to B$ para $A' \subset A$ é a composição $f \circ i$, isto é, $f|_{A'} = f \circ i$, onde $i : A' \hookrightarrow A$ é a aplicação de inclusão. Para qualquer conjunto A, temos a aplicação

[1] Este é nosso jeito de definir um conjunto. Seja dado um conjunto A e seja $P(x)$ uma propriedade de elementos. Então o conjunto $C = \big\{a \in A \mid P(a)\big\}$ é formado por todos os elementos $a \in A$ que satisfazem a propriedade $P(a)$.

idêntica $1_A : A \to A$ dada pela regra $1_A : a \mapsto a$. Essa aplicação satisfaz as identidades $f \circ 1_A = f$ e $1_A \circ g = g$ para quaisquer aplicações $f : A \to B$ e $g : C \to A$.

Uma aplicação $f : A \to B$ é dita *injetora* ou uma *injeção* se $f(a_1) = f(a_2) \implies a_1 = a_2$ para todos $a_1, a_2 \in A$. A aplicação de inclusão considerada acima é um exemplo de uma aplicação injetora. Uma aplicação $f : A \to B$ é dita *sobrejetora* ou uma *sobrejeção* se todo elemento de B é a imagem por f de um elemento de A, isto é, se, para todo $b \in B$, existe um $a \in A$ tal que $f(a) = b$. Uma aplicação $f : A \to B$ simultâneamente injetora e sobrejetora é dita *bijetora* ou uma *bijeção*. Um jeito equivalente de definir bijeção: uma aplicação $f : A \to B$ se chama *bijeção* se ela possui uma inversa de dois lados relativamente à composição; isto significa que existe uma aplicação (*inversa*) $g : B \to A$ tal que $f \circ g = 1_B$ e $g \circ f = 1_A$.

2. Escalares

Necessitamos fixar um conjunto $\mathbb{K}$ de escalares. Normalmente, este será $\mathbb{R}$ ou $\mathbb{C} := \{c \mid c$ é um número complexo$\}$. Mas quase tudo funcionará[2] tomando-se $\mathbb{K} = \mathbb{Q}$, onde $\mathbb{Q} := \{q \mid q$ é um número racional$\}$.

As propriedades que um conjunto numérico $\mathbb{K}$ deve ter para servir como conjunto de escalares são as seguintes:

E1. $1 \in \mathbb{K}$.
E2. $k_1, k_2 \in \mathbb{K} \implies -k_1, k_1 + k_2, k_1k_2 \in \mathbb{K}$.
E3. $0 \neq k \in \mathbb{K} \implies \dfrac{1}{k} \in \mathbb{K}$.

Em palavras: temos escalar 1 e podemos efetuar adição, subtração, multiplicação e divisão.[3] Em particular, $0 \in \mathbb{K}$ pois $0 = 1 - 1$.

3. Espaço vetorial

Seja V um conjunto munido de duas operações: $\mathbb{K} \times V \xrightarrow{\cdot} V$ denotada $(k, v) \mapsto k \cdot v$ e chamada *multiplicação* por escalar, e $V \times V \xrightarrow{+} V$ denotada $(v_1, v_2) \mapsto v_1 + v_2$ e chamada *adição*. Dizemos que V é um $\mathbb{K}$-*espaço vetorial* se

V1. $v_1 + v_2 = v_2 + v_1$ para todos $v_1, v_2 \in V$ (comutatividade da adição).

V2. $(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$ para todos $v_1, v_2, v_3 \in V$ (associatividade da adição).

V3. Existe um $n \in V$ tal que $n + v = v$ para todo $v \in V$ (existência do elemento *neutro* para a adição). Tal n é único: para elementos neutros n e n' temos $n' = n + n' = n' + n = n$. No que se segue, denotaremos este elemento por 0.

V4. Para todo $v \in V$, existe um $v' \in V$ tal que $v + v' = 0$. Para um dado $v \in V$, este v' é único: se $v + v'' = 0$, temos $v'' = 0 + v'' = (v + v') + v'' = (v' + v) + v'' = v' + (v + v'') = v' + 0 = 0 + v' = v'$. Em seguida, chamaremos v' *oposto* a v e o denotaremos por $(-v)$.

V5. $k \cdot (v_1 + v_2) = (k \cdot v_1) + (k \cdot v_2)$ para todos $v_1, v_2 \in V$ e $k \in \mathbb{K}$ (distributividade da multiplicação por escalar relativamente à adição em V).

V6. $(k_1 + k_2) \cdot v = (k_1 \cdot v) + (k_2 \cdot v)$ para todos $k_1, k_2 \in \mathbb{K}$ e $v \in V$ (distributividade da multiplicação por escalar relativamente às adições).

V7. $(k_1k_2) \cdot v = k_1 \cdot (k_2 \cdot v)$ para todos $k_1, k_2 \in \mathbb{K}$ e $v \in V$ (associatividade da multiplicação por escalar).

V8. $1 \cdot v = v$ para todo $v \in V$.

Daí segue $0 \cdot v = 0$. Realmente, $v = 1 \cdot v = (1 + 0) \cdot v = 1 \cdot v + 0 \cdot v = v + 0 \cdot v$. Agora $0 = (-v) + v = (-v) + (v + 0 \cdot v) = \big((-v) + v\big) + 0 \cdot v = 0 + 0 \cdot v = 0 \cdot v$. Também temos $(-1) \cdot v = -v$. Com efeito,

[2]Na computação usa-se uma escolha mais exótica de escalares, a de 16 elementos.
[3]A última operação exige que o escalar pelo qual dividimos não seja nulo.

$v+(-1)\cdot v=1\cdot v+(-1)\cdot v=\big(1+(-1)\big)\cdot v=0\cdot v=0$. Em seguida, vamos às vezes omitir $\cdot$ e utilizar a notação de subtração $v_1-v_2:=v_1+(-v_2)$.

3. Exemplos. 1. $\mathbb{R}$-espaço vetorial de vetores no plano.

3.2. $\mathbb{R}$-espaço vetorial de vetores no espaço.

3.3. Seja C um conjunto. Denotamos por $\operatorname{Func}(C,\mathbb{K}):=\{f:C\to\mathbb{K}\}$ o conjunto de todas as funções (= aplicações) de C para $\mathbb{K}$. Definimos operações. Para $f,f_1,f_2\in\operatorname{Func}(C,\mathbb{K})$ e $k\in\mathbb{K}$, façamos $(f_1+f_2)(c):=f_1(c)+f_2(c)$ e $(k\cdot f)(c):=kf(c)$ para qualquer $c\in C$. É fácil verificar que obtemos um $\mathbb{K}$-espaço vetorial. Realmente, para provar que $f_1+f_2=f_2+f_1$, precisamos apenas verificar que $(f_1+f_2)(c)=(f_2+f_1)(c)$ para todo $c\in C$: pela definição, $(f_1+f_2)(c)=f_1(c)+f_2(c)=f_2(c)+f_1(c)=(f_2+f_1)(c)$. O elemento neutro para a adição é a função identicamente nula, dada por $0(c):=0$ para todo $c\in C$. A função oposta a f é dada por $(-f)(c):=-f(c)$ para todo $c\in C$. Os outros axiomas se verificam de modo análogo.

3.4. Os próprios escalares $\mathbb{K}$ munidos das óbvias operações constituem um $\mathbb{K}$-espaço vetorial.

3.5. Sejam V_1,V_2 $\mathbb{K}$-espaços vetoriais. Em $V_1\times V_2$, definimos operações $(v_1,v_2)+(v_1',v_2'):=(v_1+v_1',v_2+v_2')$ e $k\cdot(v_1,v_2):=(k\cdot v_1,k\cdot v_2)$ para todos $v_1,v_1'\in V_1$, $v_2,v_2'\in V_2$ e $k\in\mathbb{K}$. É fácil verificar que obtemos um $\mathbb{K}$-espaço vetorial chamado *soma direta* de V_1 e V_2 e denotado por $V_1\oplus V_2$. Por exemplo, o elemento neutro para a adição é $(0,0)$ e o oposto a (v_1,v_2) é $(-v_1,-v_2)$. De maneira semelhante, podemos definir a soma direta $V_1\oplus V_2\oplus\dots\oplus V_n$ de n $\mathbb{K}$-espaços vetoriais $V_1,V_2,\dots,V_n$. Denotamos $\mathbb{K}^n:=\underbrace{\mathbb{K}\oplus\mathbb{K}\oplus\dots\oplus\mathbb{K}}_{n\text{ vezes}}$.

3.6. $\mathbb{C}$ é um $\mathbb{R}$-espaço vetorial. $\mathbb{R}$ e $\mathbb{C}$ são $\mathbb{Q}$-espaços vetoriais.

3.7. O conjunto $\mathbb{K}[x]:=\{k_nx^n+k_{n-1}x^{n-1}+\dots+k_1x+k_0\mid k_0,k_1,\dots,k_{n-1},k_n\in\mathbb{K},n\geq 0\}$ de todos os polinômios em uma variável x com coeficientes em $\mathbb{K}$ munido das óbvias operações (aqu, é melhor interpretar um polinômio como uma expressão formal, não como uma função) é um $\mathbb{K}$-espaço vetorial. Fazendo $\mathbb{K}[x]_{<n}:=\big\{f(x)\in\mathbb{K}[x]\mid\text{grau de }f(x)<n\big\}$ obtemos um $\mathbb{K}$-espaço vetorial que é de fato $\mathbb{K}^n$.

3.8. Seja C um conjunto e seja $C_0\subset C$. O conjunto $\big\{f\in\operatorname{Func}(C,\mathbb{K})\mid f|_{C_0}\equiv 0\big\}$ munido das mesmas operações como no Exemplo 3.3 é um $\mathbb{K}$-espaço vetorial, onde $f|_{C_0}\equiv 0$ significa que a função f é identicamente nula sobre C_0.

4. Subespaços. Dependência linear. Base e dimensão

Seja V um $\mathbb{K}$-espaço vetorial e seja $W\subset V$. Dizemos que W é um *subespaço* de V se

S0. $0\in W$.

S1. $k\in\mathbb{K},w\in W\Longrightarrow k\cdot w\in W$.

S2. $w_1,w_2\in W\Longrightarrow w_1+w_2\in W$.

Em palavras: W é um subconjunto em V fechado relativamente às operações de tomar-se o 0, adição e multiplicação por escalar. É fácil ver que W munido das operações induzidas é um $\mathbb{K}$-espaço vetorial. Denotamos $W\leq V$.

4.1. Exemplos. 1. Seja V um $\mathbb{K}$-espaço vetorial e sejam $W_1,W_2\leq V$. Então $W_1\cap W_2\leq V$ e $W_1+W_2:=\{w_1+w_2\mid w_1\in W_1,w_2\in W_2\}\leq V$. Realmente, de $0\in W_1$ e $0\in W_2$ segue que $0\in W_1\cap W_2$. Se $w\in W_1\cap W_2$ e $k\in\mathbb{K}$, então $w\in W_i$ para $i=1,2$. Sendo W_i um subespaço, $k\cdot w\in W_i$ para $i=1,2$. Em outras palavras, $k\cdot w\in W_1\cap W_2$. Se $w,w'\in W_1\cap W_2$, então $w,w'\in W_i$ e $w+w'\in W_i$ para $i=1,2$, pois W_i é um subespaço. Logo, $w+w'\in W_1\cap W_2$. Assim provamos que $W_1\cap W_2\leq V$. Obviamente, $W_1\cap W_2$ é o maior (no sentido de inclusão de conjuntos) subespaço contido em ambos W_1 e W_2. Vamos verificar que $W_1+W_2\leq V$. Sendo $0\in W_1$ e $0\in W_2$, temos $0=0+0\in W_1+W_2$. Sejam $w\in W_1+W_2$ e $k\in\mathbb{K}$. Então $w=w_1+w_2$ para alguns $w_1\in W_1$ e $w_2\in W_2$. Logo, $k\cdot w=k\cdot w_1+k\cdot w_2\in W_1+W_2$, pois $k\cdot w_1\in W_1$ e $k\cdot w_2\in W_2$. Finalmente, sejam $w,w'\in W_1+W_2$. Então, $w=w_1+w_2$ e $w'=w_1'+w_2'$ para $w_1,w_1'\in W_1$ e $w_2,w_2'\in W_2$ apropriados.

Daí, $w + w' = w_1 + w_1' + w_2 + w_2'$ com $w_1 + w_1' \in W_1$ e $w_2 + w_2' \in W_2$. Assim, $W_1 + W_2 \leq V$. Note que $W_1 + W_2$ contém ambos W_1 e W_2. Realmente, todo $w_1 \in W_1$ pode ser escrito como $w_1 = w_1 + 0$ com $0 \in W_2$. Logo, $W_1 \subset W_1 + W_2$. De modo semelhante, obtemos $W_2 \subset W_1 + W_2$. Ainda mais, o subespaço $W_1 + W_2$ é o menor (no sentido de inclusão de conjuntos) subespaço que contém ambos W_1 e W_2. Com efeito, seja $W \leq V$ um subespaço tal que $W_1, W_2 \subset W$. Então, para quaisquer $w_1 \in W_1$ e $w_2 \in W_2$, temos $w_1, w_2 \in W$. Daí, $w_1 + w_2 \in W$. Assim vemos que qualquer elemento de $W_1 + W_2$ pertence a W, ou seja, $W_1 + W_2 \subset W$. Da mesma maneira podemos definir o subespaço $W_1 + W_2 + \cdots + W_n := \{w_1 + w_2 + \cdots + w_n \mid w_1 \in W_1, w_2 \in W_2, \ldots, w_n \in W_n\}$ para subespaços $W_1, W_2, \ldots, W_n \leq V$. Este subespaço é o menor que contém todos os $W_1, W_2, \ldots, W_n$. Como a adição é associativa, não colocamos os parênteses em $W_1 + W_2 + \cdots + W_n$.

4.1.2. Seja V um $\mathbb{K}$-espaço vetorial e seja $v \in V$. Então é fácil ver que $\mathbb{K}v := \{k \cdot v \mid k \in \mathbb{K}\} \leq V$. Obviamente, $\mathbb{K}v$ é o menor subespaço que contém v. Caso $v \neq 0$, todo $w \in \mathbb{K}v$ admite a única forma $w = k \cdot v$ com $k \in \mathbb{K}$.

4.1.3. Seja V um $\mathbb{K}$-espaço vetorial e sejam $v_1, \ldots, v_n \in V$. Denotamos $[v_1, \ldots, v_n] := \mathbb{K}v_1 + \cdots + \mathbb{K}v_n$. Pelos Exemplos 4.1.1 e 4.1.2, $[v_1, \ldots, v_n]$ é o menor subespaço que contém todos os $v_1, \ldots, v_n$. Dizemos que $[v_1, \ldots, v_n]$ é o subespaço *gerado* por $v_1, \ldots, v_n$ e chamamos os elementos $v_1, \ldots, v_n$ *geradores* deste subespaço. É imediato que todo $v \in [v_1, \ldots, v_n]$ tem uma forma $v = k_1 \cdot v_1 + \cdots + k_n \cdot v_n$, onde $k_1, \ldots, k_n \in \mathbb{K}$. A expressão $k_1 \cdot v_1 + \cdots + k_n \cdot v_n$ se chama *combinação linear* de $v_1, \ldots, v_n$ com coeficientes $k_1, \ldots, k_n \in \mathbb{K}$. Nestes termos, um subconjunto $W \subset V$ não-vazio é um subespaço se, e só se, é fechado relativamente às combinações lineares de seus elementos com coeficientes arbitrários.

4.1.4. Temos subespaços $\mathbb{K}[x]_{<n} \leq \mathbb{K}[x]_{<m} \leq \mathbb{K}[x]$ para $n \leq m$.

4.1.5. Temos $\mathbb{R} \leq \mathbb{C}$ para o $\mathbb{R}$-espaço vetorial $\mathbb{C}$ e $\mathbb{Q} \leq \mathbb{R} \leq \mathbb{C}$ para o $\mathbb{Q}$-espaço vetorial $\mathbb{C}$.

4.1.6. O conjunto $W := \big\{f \in \operatorname{Func}(C, \mathbb{K}) \mid f|_{C_0} \equiv 0\big\}$ no Exemplo 3.8 é um subespaço do $\mathbb{K}$-espaço vetorial $\operatorname{Func}(C, \mathbb{K})$, $W \leq \operatorname{Func}(C, \mathbb{K})$.

4.1.7. Soluções de um sistema homogêneo de equações lineares em n variáveis formam um subespaço do $\mathbb{K}$-espaço vetorial $\mathbb{K}^n$.

4.1.8. Seja V um $\mathbb{K}$-espaço vetorial. Então $V \leq V$ e $0 \leq V$, onde 0 denota o subconjunto formado apenas pelo elemento 0. Estes dois subespaços são o maior e o menor possíveis, respectivamente. Podemos escrever $0 = [\varnothing]$, ou seja, o subespaço 0 é gerado por nada. Note também que todo subespaço de um subespaço é um subespaço: $V \geq W_1$ e $W_1 \geq W_2 \Longrightarrow V \geq W_2$.

4.2. Definição. Sejam $v_1, v_2, \ldots, v_n \in V$ elementos de um $\mathbb{K}$-espaço vetorial V. Dizemos que $v_1, v_2, \ldots, v_n$ são *linearmente independentes* (abreviando, LI) se, para quaisquer $k_1, k_2, \ldots, k_n \in \mathbb{K}$,

$$\sum_{i=1}^{n} k_i \cdot v_i = 0 \Longrightarrow k_i = 0 \text{ para todo } i = 1, 2, \ldots, n.$$

Em palavras: se a combinação linear de $v_1, v_2, \ldots, v_n$ é nula, então ela é *trivial,* isto é, todos os seus coeficientes são nulos. Se $v_1, v_2, \ldots, v_n$ não são LI, dizemos que eles são *linearmente dependentes* (abreviando, LD). Os elementos $v_1, v_2, \ldots, v_n \in V$ são LD se uma combinação não-trivial de tais elementos é nula, isto é, se existem $k_1, k_2, \ldots, k_n \in \mathbb{K}$, nem todos nulos, tais que $\displaystyle\sum_{i=1}^{n} k_i \cdot v_i = 0$. Note que, quando $v_1, v_2, \ldots, v_n$ são LD, o elemento 0 admite a forma de duas diferentes combinações lineares de $v_1, v_2, \ldots, v_n$. Realmente, $0 = \displaystyle\sum_{i=1}^{n} k_i \cdot v_i = \sum_{i=1}^{n} 0 \cdot v_i$.

O conceito de dependência linear é essencial quando procuramos uma coleção mínima de geradores de um subespaço:

4.3. Lema. *Sejam $v_1, v_2, \ldots, v_n \in V$ elementos de um $\mathbb{K}$-espaço vetorial e seja $\displaystyle\sum_{i=1}^{n} k_i \cdot v_i = 0$ uma* *dependência linear não-trivial, onde $k_1, k_2, \dots, k_n \in \mathbb{K}$. Assim, $k_j \neq 0$ para algum j. Então podemos excluir o gerador v_j da coleção de geradores $v_1, v_2, \dots, v_n$ não alterando o subespaço $[v_1, v_2, \dots, v_n]$, isto é, $[v_1, v_2, \dots, v_n] = [v_1, v_2, \dots, v_{j-1}, v_{j+1}, \dots, v_n]$.*

Demonstração. Multiplicando a dependência por k_j^{-1} e isolando v_j, obtemos $v_j = \sum\limits_{i \neq j}(-k_j^{-1}k_i) \cdot v_i$. Logo, $v_j \in [v_1, v_2, \dots, v_{j-1}, v_{j+1}, \dots, v_n]$. É claro que $[v_1, v_2, \dots, v_{j-1}, v_{j+1}, \dots, v_n] \subset [v_1, v_2, \dots, v_n]$. Lembrando que $[v_1, v_2, \dots, v_n]$ é o mínimo subespaço que contém todos os $v_1, v_2, \dots, v_n$ e observando que $v_l \in [v_1, v_2, \dots, v_{j-1}, v_{j+1}, \dots, v_n]$ para todo l, concluímos que $[v_1, v_2, \dots, v_{j-1}, v_{j+1}, \dots, v_n] \supset [v_1, v_2, \dots, v_n]$ ■

4.4. Observação. *Sejam $v_1, v_2, \dots, v_n \in V$ elementos LI de um $\mathbb{K}$-espaço vetorial. Então todo $v \in [v_1, v_2, \dots, v_n]$ admite uma única forma de combinação linear de $v_1, v_2, \dots, v_n$.*

Demonstração. O fato que todo $v \in [v_1, v_2, \dots, v_n]$ admite a forma de uma combinação linear de $v_1, v_2, \dots, v_n$ foi observado no Exemplo 4.1.3. Para a unicidade, suponhamos que $\sum\limits_{i=1}^{n} k_i \cdot v_i = \sum\limits_{i=1}^{n} k'_i \cdot v_i$, onde $k_1, k_2, \dots, k_n, k'_1, k'_2, \dots, k'_n \in \mathbb{K}$. Então $\sum\limits_{i=1}^{n}(k_i - k'_i) \cdot v_i = 0$. Sendo $v_1, v_2, \dots, v_n$ LI, concluímos que $k_i - k'_i = 0$ para todo $i = 1, 2, \dots, n$, ou seja, $k_i = k'_i$ ■

4.5. Observação. *Qualquer parte de uma coleção LI em um espaço vetorial é LI* ■

4.6. Lema. *Sejam $v_1, v_2, \dots, v_n \in V$ elementos LI de um $\mathbb{K}$-espaço vetorial. Então, para qualquer $k \in \mathbb{K}$ e quaisquer $1 \leq i, j \leq n$ tais que $i \neq j$, os elementos $v_1, v_2, \dots, v_{i-1}, v_i + k \cdot v_j, v_{i+1}, \dots, v_n$ são LI.*

Demonstração. Seja $k_i \cdot (v_i + k \cdot v_j) + \sum\limits_{l \neq i} k_l \cdot v_l = 0$ uma dependência linear de elementos $v_1, v_2, \dots, v_{i-1}, v_i + k \cdot v_j, v_{i+1}, \dots, v_n$. Então $k_i \cdot (v_i + k \cdot v_j) + k_j \cdot v_j + \sum\limits_{l \neq i,j} k_l \cdot v_l = 0$, ou seja, $k_i \cdot v_i + (k_j + k_i k) \cdot v_j + \sum\limits_{l \neq i,j} k_l \cdot v_l = 0$. Sendo $v_1, v_2, \dots, v_n$ LI, obtemos $k_l = 0$ para todo $l \neq i, j$, $k_i = 0$ e $k_j + k_i k = 0$. Daí, $k_j = 0$. Em outras palavras, todos os coeficientes da dependência linear acima são nulos ■

4.7. Corolário. *Sejam $v_1, v_2, \dots, v_n \in V$ elementos LI de um $\mathbb{K}$-espaço vetorial. Fixemos um índice $1 \leq j \leq n$ e tomemos quaisquer $k_l \in \mathbb{K}$ para $l \neq j$. Então os elementos*

$$v_1 + k_1 \cdot v_j, v_2 + k_2 \cdot v_j, \dots, v_{j-1} + k_{j-1} \cdot v_j, v_j, v_{j+1} + k_{j+1} \cdot v_j, \dots, v_n + k_n \cdot v_j$$

são LI.

Demonstração. Aplicando o Lema 4.6 com $i = 1$ e $k = k_1$, obtemos uma nova coleção LI

$$v_1 + k_1 \cdot v_j, v_2, \dots, v_j, \dots, v_n.$$

Aplicando o Lema 4.6 à nova coleção com $i = 2$ e $k = k_2$, obtemos a coleção LI

$$v_1 + k_1 \cdot v_j, v_2 + k_2 \cdot v_j, v_3, \dots, v_j, \dots, v_n.$$

Assim, chegamos à coleção LI

$$v_1 + k_1 \cdot v_j, v_2 + k_2 \cdot v_j, \dots, v_{j-1} + k_{j-1} \cdot v_j, v_j, v_{j+1}, \dots, v_n.$$

Aplicando o Lema 4.6 a essa última coleção com $i = j + 1$ e $k = k_{j+1}$, obtemos a coleção LI

$$v_1 + k_1 \cdot v_j, v_2 + k_2 \cdot v_j, \dots, v_{j-1} + k_{j-1} \cdot v_j, v_j, v_{j+1} + k_{j+1} \cdot v_j, v_{j+2}, \dots, v_n.$$

Continuando a agir deste modo, chegamos ao resultado desejado ■

4.8. Teorema. *Sejam $v_1, v_2, \dots, v_n \in V$ LI e seja $V = [g_1, g_2, \dots, g_m]$. Então $m \geq n$.*

Demonstração. Utilizamos a indução sobre m. Para $m = 0$, o fato é óbvio. Suponhamos que $m > 0$. Façamos $W := [g_1, g_2, \dots, g_{m-1}]$. Se $v_i \in W$ para todo i, pela hipótese de indução, obtemos $m - 1 \geq n$ e, portanto, $m \geq n$. Podemos supor que um dos v_i's não pertence a W. Sem perda de generalidade, este é v_n, $v_n \notin W$. Vamos provar que $\mathbb{K}v_n + W = V$. Para alguns $k_1, k_2, \dots, k_m \in \mathbb{K}$, temos $v_n = \sum_{j=1}^{m} k_j \cdot g_j$ com $k_m \neq 0$ pois, caso contrário, $v_n \in [g_1, g_2, \dots, g_{m-1}] = W$. Isolando g_m, obtemos $g_m = k_m^{-1} v_n - \sum_{j=1}^{m-1} (k_m^{-1} k_j) \cdot g_j$. Em outras palavras, $g_m \in \mathbb{K}v_n + W$. Assim, $g_j \in \mathbb{K}v_n + W$ para todo j. Agora, $V = [g_1, g_2, \dots, g_m]$ implica $\mathbb{K}v_n + W = V$. Portanto, para todo $1 \leq i < n$, existem $k_i \in \mathbb{K}$ e $w_i \in W$ tais que $v_i = k_i \cdot v_n + w_i$. Isto pode ser reescrito como $w_i = v_i + (-k_i) \cdot v_n$ para todo $i = 1, 2, \dots, n-1$. Pelo Corolário 4.7 aplicado aos v_i's, ao $j = n$ e aos $(-k_l)$'s, $l \neq n$, e pela Observação 4.5, concluímos que $w_1, w_2, \dots, w_{n-1}$ são LI. Aplicando a hipótese de indução para $w_1, w_2, \dots, w_{n-1} \in W = [g_1, g_2, \dots, g_{m-1}]$, obtemos $m - 1 \geq n - 1$. Portanto, $m \geq n$ ■

4.9. Definição. Uma coleção LI de geradores $b_1, b_2, \dots, b_n \in V$ se chama *base linear* do $\mathbb{K}$-espaço vetorial V. Enfatizamos que a base é uma coleção ordenada.

Do Lema 4.3 segue o

4.10. Corolário. *De qualquer coleção finita de geradores de um espaço vetorial, é possível escolher uma base linear* ■

4.11. Corolário. *Seja V um $\mathbb{K}$-espaço vetorial finitamente gerado. Então toda coleção finita LI de elementos de V pode ser completada a uma base linear de V.*

Demonstração. Seja $V = [g_1, g_2, \dots, g_m]$ e seja $v_1, v_2, \dots, v_n \in V$ uma coleção LI. Pelo Teorema 4.8, $n \leq m$. Portanto, aumentando a coleção $v_1, v_2, \dots, v_n$ se necessário, podemos supor que ela é uma coleção máxima (no sentido de inclusão de conjuntos) LI. Vamos mostrar que ela é uma base linear. Basta provar que $V = [v_1, v_2, \dots, v_n]$. Se $V \neq [v_1, v_2, \dots, v_n]$, tomamos qualquer $v_{n+1} \in V \setminus [v_1, v_2, \dots, v_n]$ e demonstramos que $v_1, v_2, \dots, v_n, v_{n+1}$ são LI, assim chegando a uma contradição com a maximalidade de $v_1, v_2, \dots, v_n$. Realmente, suponhamos que $\sum_{i=1}^{n+1} k_i \cdot v_i = 0$. Se $k_{n+1} = 0$, obtemos uma dependência linear dos $v_1, v_2, \dots, v_n$, donde concluímos que todos os k_i's são nulos. Se $k_{n+1} \neq 0$, podemos expressar v_{n+1} na forma de combinação linear dos $v_1, v_2, \dots, v_n$: $v_{n+1} = \sum_{i=1}^{n} (-k_{n+1}^{-1} k_i) \cdot v_i$. Isto contradiz $v_{n+1} \notin [v_1, v_2, \dots, v_n]$ ■

Do Teorema 4.8 segue imediatamente o

4.12. Corolário. *Duas bases lineares de um espaço vetorial têm a mesma cardinalidade* ■

4.13. Definição. A cardinalidade[4] de uma base linear de um $\mathbb{K}$-espaço vetorial V se chama *dimensão* de V sobre $\mathbb{K}$ e se denota por $\dim_{\mathbb{K}} V$. Dos Lema 4.3 e Corolário 4.12 segue que a dimensão pode ser caracterizada como a cardinalidade de qualquer coleção mínima (no sentido de inclusão de conjuntos) de geradores. Pela demonstração do Corolário 4.11, a dimensão também pode ser caracterizada como a cardinalidade de qualquer coleção máxima LI. Note que do Teorema 4.8 segue que qualquer subespaço de um espaço vetorial de dimensão finita tem dimensão finita.

[4]Consideramos aqui somente os espaços vetoriais finitamente gerados, mas não é difícil generalizar as considerações para os espaços vetoriais de dimensão infinita.

4.14. Corolário. *Seja $W \leq V$ um subespaço de um $\mathbb{K}$-espaço vetorial V finitamente gerado. Então existe um subespaço $W' \leq V$ tal que $W \cap W' = 0$ e $W + W' = V$.*

Demonstração. Tomamos uma base linear $b_1, b_2, \dots, b_n \in W$ de W e completâmo-la a uma base linear $b_1, b_2, \dots, b_n, b_{n+1}, \dots, b_m \in V$ de V, $m \geq n$. Façamos $W' := [b_{n+1}, \dots, b_m]$. Claramente, $W + W' = V$. Seja $w \in W \cap W'$. Então $w = \sum\limits_{i=1}^{n} k_i \cdot b_i = \sum\limits_{j=n+1}^{m} k_j \cdot b_j$ para $k_1, k_2, \dots, k_n, k_{n+1}, \dots, k_m \in \mathbb{K}$ apropriados. Daí obtemos uma dependência linear $\sum\limits_{i=1}^{n} k_i \cdot b_i + \sum\limits_{j=n+1}^{m} (-k_j) \cdot b_j = 0$, implicando que todos os k_i's são nulos. Logo, $w = 0$ ∎

4.15. Corolário. *Sejam $W_1, W_2 \leq V$ subespaços de um $\mathbb{K}$-espaço vetorial V finitamente gerado. Então $\dim_{\mathbb{K}} W_1 + \dim_{\mathbb{K}} W_2 = \dim_{\mathbb{K}}(W_1 \cap W_2) + \dim_{\mathbb{K}}(W_1 + W_2)$.*

Demonstração. Escolhemos uma base linear $b_1, b_2, \dots, b_n \in W_1 \cap W_2$ de $W_1 \cap W_2$. Pelo Corolário 4.11, podemos completá-la a uma base linear $b_1, b_2, \dots, b_n, a_1, a_2, \dots, a_m \in W_1$ de W_1. Também completâmo-la a uma base linear $b_1, b_2, \dots, b_n, c_1, c_2, \dots, c_l \in W_2$ de W_2. Basta mostrar que $b_1, b_2, \dots, b_n$, $a_1, a_2, \dots, a_m, c_1, c_2, \dots, c_l \in W_1 + W_2$ é uma base linear de $W_1 + W_2$, pois, neste caso, $\dim_{\mathbb{K}}(W_1 \cap W_2) = n$, $\dim_{\mathbb{K}} W_1 = n + m$, $\dim_{\mathbb{K}} W_2 = n + l$ e $\dim_{\mathbb{K}}(W_1 + W_2) = n + m + l$, implicando a fórmula desejada.

De $W_1 = [b_1, b_2, \dots, b_n, a_1, a_2, \dots, a_m]$ e $W_2 = [b_1, b_2, \dots, b_n, c_1, c_2, \dots, c_l]$ segue que $W_1 + W_2 = [b_1, b_2, \dots, b_n, a_1, a_2, \dots, a_m, c_1, c_2, \dots, c_l]$ (lembre-se que $W_1 + W_2$ é o menor subespaço que contém ambos W_1 e W_2).

Seja $\sum\limits_{i=1}^{n} k_i \cdot b_i + \sum\limits_{j=1}^{m} k'_j \cdot a_j + \sum\limits_{s=1}^{l} k''_s \cdot c_s = 0$ uma dependência linear entre os $b_1, b_2, \dots, b_n, a_1, a_2, \dots, a_m, c_1, c_2, \dots, c_l$. Então $-\sum\limits_{s=1}^{l} k''_s \cdot c_s = \sum\limits_{i=1}^{n} k_i \cdot b_i + \sum\limits_{j=1}^{m} k'_j \cdot a_j \in W_1 \cap W_2$, pois a parte direita da igualdade pertence a W_1 e a parte esquerda da igualdade pertence a W_2. Por outro lado, em termos da base $b_1, b_2, \dots, b_n, c_1, c_2, \dots, c_l$ de W_2, o elemento $-\sum\limits_{s=1}^{l} k''_s \cdot c_s$, sendo pertencente a $W_1 \cap W_2$, tem que ter coeficientes não-nulos somente na frente dos b_i's. Em outras palavras, todos os k''_s's são nulos. Agora temos $\sum\limits_{i=1}^{n} k_i \cdot b_i + \sum\limits_{j=1}^{m} k'_j \cdot a_j = 0$. Lembrando que os $b_1, b_2, \dots, b_n, a_1, a_2, \dots, a_m$ são LI, vemos que todos os k_i's e k'_j's são nulos ∎

4.16. Dicionário. Seja $\beta : b_1, b_2, \dots, b_n$ uma base linear de um $\mathbb{K}$-espaço vetorial V. Podemos associar a todo elemento $v \in V$ uma coluna de escalares $[v]_\beta := \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}$ que, pela Observação 4.4, é univocamente determinada pela igualdade $v = \sum\limits_{i=1}^{n} k_i \cdot b_i$. Reciprocamente, qualquer coluna de n escalares $\begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}$ define um elemento $v \in V$ dado pela mesma fórmula $v := \sum\limits_{i=1}^{n} k_i \cdot b_i$. Assim, quando uma base linear em V é fixa, temos um dicionário perfeito que interpreta os elementos de V como as colunas de escalares. Note que b_i corresponde à coluna cujo único coeficiente não-nulo é igual a 1 e está no i-ésimo lugar.

Mais ainda, este dicionário preserva a adição e a multiplicação por escalar. Com efeito, se $[v]_\beta = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}$ e $[v']_\beta = \begin{bmatrix} k'_1 \\ k'_2 \\ \vdots \\ k'_n \end{bmatrix}$, então $v = \sum\limits_{i=1}^{n} k_i \cdot b_i$ e $v' = \sum\limits_{i=1}^{n} k'_i \cdot b_i$. Portanto, $v + v' = \sum\limits_{i=1}^{n} (k_i + k'_i) \cdot b_i$, ou seja, $[v + v']_\beta = \begin{bmatrix} k_1 + k'_1 \\ k_2 + k'_2 \\ \vdots \\ k_n + k'_n \end{bmatrix}$. Em outras palavras, $[v + v']_\beta = [v]_\beta + [v']_\beta$. Para qualquer $k \in \mathbb{K}$, temos $k \cdot v = \sum\limits_{i=1}^{n} (kk_i) \cdot b_i$. Isto significa que $[k \cdot v]_\beta = \begin{bmatrix} kk_1 \\ kk_2 \\ \vdots \\ kk_n \end{bmatrix}$. Logo, $[k \cdot v]_\beta = k[v]_\beta$.

5. Aplicações (transformações) lineares. Matrizes

Uma aplicação $A : U \to V$ entre $\mathbb{K}$-espaços vetoriais é dita *linear* ou *transformação linear* se

A1. $A(u_1 + u_2) = Au_1 + Au_2$ para todos $u_1, u_2 \in U$.
A2. $A(ku) = k(Au)$ para todos $u \in U$ e $k \in \mathbb{K}$.

Em palavras: A preserva a adição e a multiplicação por escalar.

5.1. Observação. *Seja $A : U \to V$ uma aplicação linear entre $\mathbb{K}$-espaços vetoriais, sejam $u_1, u_2, \dots, u_n \in U$ e sejam $k_1, k_2, \dots, k_n \in \mathbb{K}$. Então $A\Big(\sum\limits_{i=1}^{n} k_i u_i\Big) = \sum\limits_{i=1}^{n} k_i A u_i$. Em outras palavras, toda aplicação linear preserva combinações lineares* ■

5.2. Exemplos 1. Seja $U \le V$ um subespaço de um $\mathbb{K}$-espaço vetorial V. Então a aplicação de inclusão $i : U \hookrightarrow V$ é linear. Se uma aplicação linear é injetiva, ela se chama *monomorfismo.* Assim, $i : U \hookrightarrow V$ é um exemplo de monomorfismo.

5.2.2. Seja U um $\mathbb{K}$-espaço vetorial e seja $k \in \mathbb{K}$. Então a multiplicação por k, dada pela regra $m_k : u \mapsto ku$, é uma aplicação linear $m_k : U \to U$. Se uma aplicação linear é bijetora, ela se chama *isomorfismo.* Quando dois espaços vetoriais são isomorfos, eles desfrutam as mesmas propriedades algébricas.[5] É possível ver que a aplicação inversa a um isomorfismo também é linear. Caso $k \neq 0$, a aplicação $m_k : U \to U$ é um exemplo de isomorfismo.

5.2.3. Seja V um $\mathbb{K}$-espaço vetorial e seja β uma base linear em V. Então o Dicionário 4.16 estabelece uma aplicação linear dada pela regra $v \mapsto [v]_\beta$. Essa aplicação é um isomorfismo entre V e $\mathbb{K}^n$, onde $n := \dim_{\mathbb{K}} V$.

5.2.4. Sejam $U \xrightarrow{A} V \xrightarrow{B} W$ aplicações lineares entre $\mathbb{K}$-espaços vetoriais. Então a composta $B \circ A$ é uma aplicação linear. Realmente, para todos $u_1, u_2 \in U$, temos, pela definição da composta, $(B \circ A)(u_1 + u_2) = B\big(A(u_1 + u_2)\big) = B(Au_1 + Au_2) = B(Au_1) + B(Au_2) = (B \circ A)u_1 + (B \circ A)u_2$, pois A e B são lineares. Para quaisquer $k \in \mathbb{K}$ e $u \in U$, temos $(B \circ A)(ku) = B\big(A(ku)\big) = B(kAu) = kB(Au) = k(B \circ A)u$ pelos mesmos motivos.

[5]Portanto, podemos considerá-los como os "mesmos". A única diferença entre tais espaços vetoriais é que os correspondentes conjuntos são diferentes, ou seja, as naturezas de elementos são diferentes. Mas, para os fins de nossos estudos, isto não importa. Qualquer que seja a "madeira" da qual é feito um espaço vetorial, o funcionamento do espaço não depende dessa particularidade.

5.2.5. Sejam U e V $\mathbb{K}$-espaços vetoriais. No conjunto $\operatorname{Lin}_{\mathbb{K}}(U,V) := \{A : U \to V \mid A \text{ é linear}\}$ de todas as aplicações lineares de U para V, definimos uma estrutura de $\mathbb{K}$-espaço vetorial. Para $A, A_1, A_2 \in \operatorname{Lin}_{\mathbb{K}}(U,V)$ e $k \in \mathbb{K}$, façamos $(A_1 + A_2)u := A_1u + A_2u$ e $(kA)u := kAu$ para todo $u \in U$. A aplicação 0 dada pela regra $0 : u \mapsto 0$ é obviamente linear e faz papel de um elemento neutro, pois $(A + 0)u = Au + 0u = Au + 0 = Au$ para todo $u \in U$. Seja $A \in \operatorname{Lin}_{\mathbb{K}}(U,V)$. Definimos $(-A)u := -Au$ para todo $u \in U$. A aplicação $(-A)$ assim definida é linear, pois $(-A)(u_1 + u_2) = -A(u_1 + u_2) = -(Au_1 + Au_2) = (-Au_1) + (-Au_2) = (-A)u_1 + (-A)u_2$ e $(-A)(ku) = -A(ku) = -kAu = k(-A)u$ para todos $u, u_1, u_2 \in U$ e $k \in \mathbb{K}$ (nestes cálculos, utilizamos que A é linear). Agora $\big(A + (-A)\big)u = Au + (-A)u = Au - Au = 0 = 0u$, ou seja, $A + (-A) = 0$. Os outros axiomas se verificam de modo análogo ou ainda mais fácil.

5.2.6. Sejam U, V e W $\mathbb{K}$-espaços vetoriais. Qualquer aplicação linear $A : U \to V$ define uma aplicação linear $-\circ A : \operatorname{Lin}_{\mathbb{K}}(V,W) \to \operatorname{Lin}_{\mathbb{K}}(U,W)$ dada pela regra $-\circ A : B \mapsto B\circ A$ (pelo Exemplo 5.4.2, $B \circ A \in \operatorname{Lin}_{\mathbb{K}}(U,W)$). Realmente, sejam $B, B_1, B_2 \in \operatorname{Lin}_{\mathbb{K}}(V,W)$ e seja $k \in \mathbb{K}$. Precisamos verificar que $(B_1 + B_2) \circ A = B_1 \circ A + B_2 \circ A$ e que $(kB) \circ A = k(B \circ A)$. As igualdades para verificar significam que $\big((B_1+B_2)\circ A\big)u = \big(B_1\circ A+B_2\circ A\big)u$ e $\big((kB)\circ A\big)u = \big(k(B\circ A)\big)u$ para todo $u \in U$. Essas verificações constituem um cálculo automático:

$$\big((B_1 + B_2) \circ A\big)u = (B_1 + B_2)(Au) = B_1(Au) + B_2(Au) = (B_1 \circ A)u + (B_2 \circ A)u = \big(B_1 \circ A + B_2 \circ A\big)u,$$
$$\big((kB) \circ A\big)u = (kB)(Au) = kB(Au) = k\big((B \circ A)u\big) = \big(k(B \circ A)\big)u.$$

Em palavras: a composição com A (à direita) é uma aplicação linear. De modo análogo, qualquer aplicação linear $B : V \to W$ define uma aplicação linear $B \circ - : \operatorname{Lin}_{\mathbb{K}}(U,V) \to \operatorname{Lin}_{\mathbb{K}}(U,W)$ dada pela regra $B\circ - : A \mapsto B\circ A$. Podemos resumir ambas propriedades afirmando que a composição $\circ$ é *bilinear*. Isto significa nada mais do que a linearidade da composição em cada um de seus argumentos quando o outro está fixo.

5.3. Lema. *Sejam U e V $\mathbb{K}$-espaços vetoriais, seja $b_1, b_2, \dots, b_n \in U$ uma base linear em U e sejam $v_1, v_2, \dots, v_n \in V$. Então existe uma única aplicação linear $A : U \to V$ tal que $Ab_j = v_j$ para todo $j = 1, 2, \dots, n$.*

Demonstração. Todo $u \in U$ admite uma única forma de combinação linear de elementos da base: $u = \sum_{j=1}^{n} k_j b_j$ para únicos $k_1, k_2, \dots, k_n \in \mathbb{K}$. Temos que definir a aplicação A pela regra $Au := \sum_{j=1}^{n} k_j v_j$, pois $Ab_j = v_j$ e $A\Big(\sum_{j=1}^{n} k_j b_j\Big) = \sum_{j=1}^{n} k_j Ab_j$ pela Observação 5.1. Assim, já temos a unicidade. Se $u' = \sum_{j=1}^{n} k'_j b_j$ com $k'_1, k'_2, \dots, k'_n \in \mathbb{K}$, então

$$Au = \sum_{j=1}^{n} k_j v_j, \qquad Au' = \sum_{j=1}^{n} k'_j v_j, \qquad A(u + u') = A\Big(\sum_{j=1}^{n}(k_j + k'_j)b_j\Big) = \sum_{j=1}^{n}(k_j + k'_j)v_j$$

pela regra acima. Concluímos que $A(u + u') = Au + Au'$. Para qualquer $k \in \mathbb{K}$, temos

$$A(ku) = A\Big(\sum_{j=1}^{n}(kk_j)b_j\Big) = \sum_{j=1}^{n} kk_j v_j = k\sum_{j=1}^{n} k_j v_j = kAu \ \blacksquare$$

5.4. Corolário. *Duas aplicações lineares que coincidem em uma base linear são iguais* $\blacksquare$

5.5. Exercício. Seja $A : U \to V$ uma aplicação linear entre $\mathbb{K}$-espaços vetoriais. Prove que $A0 = 0$ e $A(-u) = -Au$ para todo $u \in U$.

Seja $A : U \to V$ uma aplicação linear entre $\mathbb{K}$-espaços vetoriais. Chamamos $A^{-1}0 = \{u \in U \mid Au = 0\}$ *núcleo* de A.

5.6. Lema. *Seja $A : U \to V$ uma aplicação linear entre $\mathbb{K}$-espaços vetoriais e sejam $U' \leq U$ e $V' \leq V$ subespaços. Então a imagem AU' e a imagem inversa $A^{-1}V'$ são subespaços, $AU' \leq V$ e $A^{-1}V' \leq U$. Em particular, o núcleo de A é um subespaço de U.*

Demonstração. Todo elemento de AU' tem a forma Au' para algum $u' \in U'$. Sejam $Au', Au'_1, Au'_2 \in AU'$ quaisquer elementos de AU', onde $u', u'_1, u'_2 \in U'$, e seja $k \in \mathbb{K}$. Então $Au'_1 + Au'_2 = A(u'_1 + u'_2) \in AU'$ e $kAu' = A(ku') \in AU'$ pois $u'_1 + u'_2 \in U'$ e $ku' \in U'$. Pelo Exercício 5.5, $A0 = 0 \in AU'$.

Sejam $u, u_1, u_2 \in A^{-1}V'$ e seja $k \in \mathbb{K}$. Então $Au, Au_1, Au_2 \in V'$. Agora $A(u_1+u_2) = Au_1+Au_2 \in V'$ e $A(ku) = kAu \in V'$. Em outras palavras, $u_1 + u_2 \in A^{-1}V'$ e $ku \in A^{-1}V'$. Pelo Exercício 5.5, $A0 = 0$. Isto implica que $0 \in A^{-1}V'$ ■

O núcleo de uma aplicação linear é uma medida de até qual ponto a aplicação não é um monomorfismo:

5.7. Lema. *Seja $A : U \to V$ uma aplicação linear entre $\mathbb{K}$-espaços vetoriais. Então A é um monomorfismo se, e só se, o núcleo de A é nulo.*

Demonstração. Obviamente o núcleo de A é nulo se A é um monomorfismo. Suponhamos que $A^{-1}0 = 0$. Se $Au_1 = Au_2$, então $0 = Au_1 + (-Au_2) = Au_1 + A(-u_2) = A(u_1 - u_2)$ pelo Exercício 5.5. Sendo o núcleo nulo, $u_1 - u_2 = 0$, ou seja, $u_1 = u_2$ ■

5.8. Definição. Seja $W \leq V$ um subespaço de um $\mathbb{K}$-espaço vetorial V finitamente gerado. Pelo Corolário 4.14, existe um subespaço $W' \leq V$ chamado *complementar* a W tal que $W \cap W' = 0$ e $W + W' = V$. Neste caso, escrevemos $W \oplus W' = V$.

De fato, obtemos a *soma direta* de espaços vetoriais definida de uma outra forma, a interna. Vamos ver o porquê. Todo elemento $v \in V$ admite uma única decomposição $v = w + w'$ com $w \in W$ e $w' \in W'$. Realmente, para $w_1, w_2 \in W$ e $w'_1, w'_2 \in W'$, a igualdade $w_1 + w'_1 = w_2 + w'_2$ implica $w_1 - w_2 = w'_2 - w'_1 \in W \cap W'$. De $W \cap W' = 0$ segue $w_1 = w_2$ e $w'_1 = w'_2$.

Agora, para quaisquer $v, v_1, v_2 \in V$ e $k \in \mathbb{K}$, temos as decomposições $v = w + w'$, $v_1 = w_1 + w'_1$ e $v_2 = w_2 + w'_2$ para únicos $w, w_1, w_2 \in W$ e $w', w'_1, w'_2 \in W'$. Portanto, $kv = kw + kw'$ é a decomposição de kv e $v_1 + v_2 = (w_1 + w_2) + (w'_1 + w'_2)$ é a decomposição de $v_1 + v_2$. Em outras palavras, a adição e a multiplicação por escalar fazem-se pelas componentes na decomposição, isto é, do mesmo jeito como na soma direta definida no Exemplo 3.5.

Vemos também que as aplicações $\pi : V \to W$ e $\pi' : V \to W'$ dadas pelas regras $\pi : v \mapsto w$ e $\pi' : v \mapsto w'$, onde $v = w + w'$ é a decomposição de v com $w \in W$ e $w' \in W'$, chamadas *projeções* (relacionadas à soma direta), são lineares.

Juntando bases lineares de W e de W', é fácil ver que $\dim_{\mathbb{K}} W + \dim_{\mathbb{K}} W' = \dim_{\mathbb{K}} V$.

5.9. Lema. *Seja $A : U \to V$ uma aplicação linear entre $\mathbb{K}$-espaços vetoriais com U finitamente gerado e seja W um subespaço complementar ao núcleo $N := A^{-1}0$ de A, isto é, $U = N \oplus W$. Então $A|_W : W \to AU$ é um isomorfismo.*

Demonstração. De $AN = 0$ segue que $AU = A(N + W) = AW$. Portanto, a aplicação $A|_W : W \to AU$ é um epimorfismo (uma aplicação linear se chama *epimorfismo* se é sobrejetora). Se $w \in W$ está no núcleo de $A|_W : W \to AU$, então $Aw = 0$. Logo, $w \in N$. Assim obtemos $w \in N \cap W = 0$, ou seja, $w = 0$. Pelo Lema 5.7, $A|_W : W \to AU$ é um monomorfismo ■

5.10. Corolário. *Seja $A : U \to V$ uma aplicação linear entre $\mathbb{K}$-espaços vetoriais com U finitamente gerado. Então $\dim_{\mathbb{K}} A^{-1}0 + \dim_{\mathbb{K}} AU = \dim_{\mathbb{K}} U$* ■

A dimensão da imagem de uma aplicação linear $A : U \to V$ é dita *posto* de A, que se denota por $\operatorname{rk} A$, $\operatorname{rk} A := \dim_{\mathbb{K}} AU$.

5.11. Dicionário. Sejam U e V $\mathbb{K}$-espaços vetoriais, seja $\beta : b_1, b_2, \dots, b_n$ uma base linear de U e seja $\gamma : c_1, c_2, \dots, c_m$ uma base linear de V. Tomemos uma aplicação linear qualquer $A : U \to V$. Então, para todo $j = 1, 2, \dots, n$, temos $Ab_j = \sum_{i=1}^{m} a_{ij}c_i$ para únicos $a_{ij} \in \mathbb{K}$ (note o uso atípico de índices). Associamos à aplicação linear A a $(m \times n)$-matriz $[A]_\gamma^\beta := [a_{ij}]$ com coeficientes em $\mathbb{K}$. Reciprocamente, seja $[a_{ij}]$ uma $(m\times n)$-matriz arbitrária com coeficientes em $\mathbb{K}$. Pelo Lema 5.3, existe uma única aplicação linear $A : U \to V$ tal que $Ab_j = \sum_{i=1}^{m} a_{ij}c_i$ para todo $j = 1, 2, \dots, n$. Assim, quando bases lineares em U e V são fixas, temos um dicionário perfeito que interpreta as aplicações lineares de $\operatorname{Lin}_\mathbb{K}(U, V)$ como as $(m \times n)$-matrizes de escalares.

Mais ainda, este dicionário estabelece um isomorfismo entre os $\mathbb{K}$-espaços vetoriais $\operatorname{Lin}_\mathbb{K}(U, V)$ e $\operatorname{Matr}_{m\times n}\mathbb{K}$, onde $\operatorname{Matr}_{m\times n}\mathbb{K}$ é formado por todas as $(m \times n)$-matrizes sobre $\mathbb{K}$. Realmente, sejam $A, A' \in \operatorname{Lin}_\mathbb{K}(U, V)$ e seja $k \in \mathbb{K}$. Para todo $j = 1, 2, \dots, n$, temos $Ab_j = \sum_{i=1}^{m} a_{ij}c_i$ e $A'b_j = \sum_{i=1}^{m} a'_{ij}c_i$, onde $a_{ij}, a'_{ij} \in \mathbb{K}$. Então $(A + A')b_j = Ab_j + A'b_j = \sum_{i=1}^{m}(a_{ij} + a'_{ij})c_i$ e $(kA)b_j = kAb_j = \sum_{i=1}^{m}(ka_{ij})c_i$ para todo $j = 1, 2, \dots, n$. Traduzindo, obtemos $[A]_\gamma^\beta = [a_{ij}]$, $[A']_\gamma^\beta = [a'_{ij}]$, $[A + A']_\gamma^\beta = [a_{ij} + a'_{ij}]$ e $[kA]_\gamma^\beta = [ka_{ij}]$. Em outras palavras, $[A + A']_\gamma^\beta = [A]_\gamma^\beta + [A']_\gamma^\beta$ e $[kA]_\gamma^\beta = k[A]_\gamma^\beta$.

Este dicionário é compatível com o Dicionário 4.16, isto é, $[Au]_\gamma = [A]_\gamma^\beta[u]_\beta$ para todos $u \in U$ e $A \in \operatorname{Lin}_\mathbb{K}(U, V)$. Com efeito, os fatos que $[u]_\beta = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}$ e $[A]_\gamma^\beta = [a_{ij}]$ se expressam pelas igualdades $u = \sum_{j=1}^{n} k_jb_j$ e $Ab_j = \sum_{i=1}^{m} a_{ij}c_i$ para todo $j = 1, 2, \dots, n$. Então, pela Observação 5.1,

$$Au = \sum_{j=1}^{n} k_jAb_j = \sum_{j=1}^{n} k_j\Big(\sum_{i=1}^{m} a_{ij}c_i\Big) = \sum_{i=1}^{m} \Big(\sum_{j=1}^{n} a_{ij}k_j\Big)c_i.$$

Resta observar que $\sum_{j=1}^{n} a_{ij}k_j$ é o i-ésimo coeficiente da coluna $[a_{ij}]\begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}$.

5.12. Dicionário. Sejam U, V e W $\mathbb{K}$-espaços vetoriais, seja $\beta : b_1, b_2, \dots, b_n$ uma base linear de U, seja $\gamma : c_1, c_2, \dots, c_m$ uma base linear de V e seja $\delta : d_1, d_2, \dots, d_l$ uma base linear de W. Então $[B \circ A]_\delta^\beta = [B]_\delta^\gamma \cdot [A]_\gamma^\beta$ para todos $A \in \operatorname{Lin}_\mathbb{K}(U, V)$ e $B \in \operatorname{Lin}_\mathbb{K}(V, W)$. Realmente, as matrizes $[A]_\gamma^\beta = [a_{ij}]$ e $[B]_\delta^\gamma = [b_{si}]$ são dadas pelas igualdades $Ab_j = \sum_{i=1}^{m} a_{ij}c_i$ para todo $j = 1, 2, \dots, n$ e $Bc_i = \sum_{s=1}^{l} b_{si}d_s$ para todo $i = 1, 2, \dots, m$. Então

$$(B \circ A)b_j = B(Ab_j) = B\Big(\sum_{i=1}^{m} a_{ij}c_i\Big) = \sum_{i=1}^{m} a_{ij}Bc_i = \sum_{i=1}^{m} a_{ij}\Big(\sum_{s=1}^{l} b_{si}d_s\Big) = \sum_{s=1}^{l} \Big(\sum_{i=1}^{m} b_{si}a_{ij}\Big)d_s.$$

Assim, o sj-ésimo coeficiente da matriz $[B \circ A]_\delta^\beta$ é igual a $\sum_{i=1}^{m} b_{si}a_{ij}$, ou seja, $[B \circ A]_\delta^\beta = [B]_\delta^\gamma \cdot [A]_\gamma^\beta$. Deste modo, o Dicionário 5.11 traduz a composta de aplicações lineares para o produto de matrizes.

5.13. Definição. Seja V um $\mathbb{K}$-espaço vetorial e sejam β e γ duas bases lineares de V. A matriz

$M^\beta_\gamma := [1_V]^\beta_\gamma$ se chama matriz de *mudança* de base β para γ.

5.14. Lema. *Sejam U e V $\mathbb{K}$-espaços vetoriais, sejam β,β' bases lineares de U e sejam γ,γ' bases lineares de V. Então, para qualquer aplicação linear $A : U \to V$, temos $[A]^{\beta'}_{\gamma'} = M^\gamma_{\gamma'} \cdot [A]^\beta_\gamma \cdot M^{\beta'}_\beta$. Além disso, $M^\beta_{\beta'} = (M^{\beta'}_\beta)^{-1}$.*

Demonstração. Pelos Definição 5.13 e Dicionário 5.12,

$$M^\gamma_{\gamma'} \cdot [A]^\beta_\gamma \cdot M^{\beta'}_\beta = [1_V]^\gamma_{\gamma'} \cdot [A]^\beta_\gamma \cdot [1_U]^{\beta'}_\beta = [1_V \circ A \circ 1_U]^{\beta'}_{\gamma'} = [A]^{\beta'}_{\gamma'}.$$

É fácil ver que $[1_U]^\beta_\beta = 1$, onde 1 denota a matriz identidade. Portanto,

$$M^{\beta'}_\beta \cdot M^\beta_{\beta'} = [1_U]^{\beta'}_\beta \cdot [1_U]^\beta_{\beta'} = [1_U \circ 1_U]^\beta_\beta = [1_U]^\beta_\beta = 1,$$
$$M^\beta_{\beta'} \cdot M^{\beta'}_\beta = [1_U]^\beta_{\beta'} \cdot [1_U]^{\beta'}_\beta = [1_U \circ 1_U]^{\beta'}_{\beta'} = [1_U]^{\beta'}_{\beta'} = 1,$$

ou seja, $M^\beta_{\beta'} = (M^{\beta'}_\beta)^{-1}$ ■

5.15. Corolário. *Seja V um $\mathbb{K}$-espaço vetorial, sejam β e γ duas bases lineares de V, seja M a matriz de mudança de base β para γ e seja $A : V \to V$ uma aplicação linear. Então $[A]^\gamma_\gamma = M \cdot [A]^\beta_\beta \cdot M^{-1}$* ■

6. Espaço dual. Naturalidade

Seja V um $\mathbb{K}$-espaço vetorial. Uma aplicação linear do formato $f : V \to \mathbb{K}$ se chama *funcional* linear (lembre-se que, pelo Exemplo 3.4, $\mathbb{K}$ é um $\mathbb{K}$-espaço vetorial). O espaço *dual* V^* é formado por todos os funcionais lineares, $V^* := \operatorname{Lin}_\mathbb{K}(V, \mathbb{K})$.

Seja $A : U \to V$ uma aplicação linear entre $\mathbb{K}$-espaços vetoriais. Definimos a aplicação *dual* $A^* : V^* \to U^*$ pela regra $A^* : f \mapsto f \circ A$ para todo funcional linear $f : V \to \mathbb{K}$, ou seja, $A^*f := f \circ A$. Então $A^* : V^* \to U^*$ é uma aplicação linear. Realmente, para todos $f, f_1, f_2 \in V^*$ e $k \in \mathbb{K}$, temos

$$A^*(f_1 + f_2) = (f_1 + f_2) \circ A = f_1 \circ A + f_2 \circ A = A^*f_1 + A^*f_2, \quad A^*(kf) = (kf) \circ A = k(f \circ A) = kA^*f$$

pelo Exemplo 5.2.6 (bilinearidade de $\circ$).

Para aplicações lineares $U \xrightarrow{A,A'} V \xrightarrow{B} W$ entre $\mathbb{K}$-espaços vetoriais e $k \in \mathbb{K}$, temos

$$(A + A')^* = A^* + A'^* \qquad (kA)^* = kA^* \qquad (B \circ A)^* = A^* \circ B^*.$$

Com efeito, para todo $f \in V^*$, temos

$$(A + A')^*f = f \circ (A + A') = f \circ A + f \circ A' = A^*f + A'^*f = (A^* + A'^*)f,$$
$$(kA)^*f = f \circ (kA) = k(f \circ A) = kA^*f$$

pelo Exemplo 5.2.6 (bilinearidade de $\circ$). Para todo $g \in W^*$, temos

$$(B \circ A)^*g = g \circ (B \circ A) = (g \circ B) \circ A = (B^*g) \circ A = A^*(B^*g) = (A^* \circ B^*)g.$$

Em palavras: passar à aplicação dual é uma aplicação linear.

Temos uma aplicação linear natural $I_V : V \to V^{**}$ definida pela regra $I_V : v \mapsto (f \mapsto fv)$ para todos $v \in V$ e $f \in V^*$, ou seja, $(I_V v)f := fv \in \mathbb{K}$. Em outras palavras, interpretamos qualquer $v \in V$ como um funcional linear sobre V^* que manda $f \in V^*$ para $fv \in \mathbb{K}$. Precisamos fazer algumas verificações. O fato que $I_V v$ definido acima é linear segue de

$$(I_V v)(f_1 + f_2) = (f_1 + f_2)v = f_1v + f_2v = (I_V v)f_1 + (I_V v)f_2, \quad (I_V v)(kf) = (kf)v = k(fv) = k\big((I_V v)f\big),$$

onde $f, f_1 f_2 \in V^*$ e $k \in \mathbb{K}$. O fato que a aplicação I_V definida acima é linear reside no cálculo

$$\big(I_V(v_1 + v_2)\big)f = f(v_1 + v_2) = fv_1 + fv_2 = (I_V v_1)f + (I_V v_2)f = \big((I_V v_1) + (I_V v_2)\big)f,$$
$$\big(I_V(kv)\big)f = f(kv) = k(fv) = k\big((I_V v)f\big) = \big(k(I_V v)\big)f,$$

onde $v, v_1, v_2 \in V$, $f \in V^*$ e $k \in \mathbb{K}$.

$$\begin{tikzcd} U \arrow[r, "I_U"] \arrow[d, "A"'] & U^{**} \arrow[d, "A^{**}"] \\ V \arrow[r, "I_V"] & V^{**} \end{tikzcd}$$

A naturalidade de I_V em V significa o seguinte. Seja $A : U \to V$ uma aplicação linear. Então temos o diagrama de aplicações lineares à esquerda. A *naturalidade* diz que este diagrama é *comutativo,* isto é, $A^{**} \circ I_U = I_V \circ A$. Verifiquemos este fato. Para todos $u \in U$ e $f \in V^*$, temos $\big((I_V \circ A)u\big)f = \big(I_V(Au)\big)f = f(Au) = (f \circ A)u = (A^* f)u$ e $\big((A^{**} \circ I_U)u\big)f = \big((A^*)^*(I_U u)\big)f = \big((I_U u) \circ A^*\big)f = (I_U u)(A^* f) = (A^* f)u$. Isto implica que $(I_V \circ A)u = (A^{**} \circ I_U)u$ para todo $u \in U$, ou seja, que $I_V \circ A = A^{**} \circ I_U$. Intuitivamente, essa naturalidade expressa que, "deformando" ou "movendo" o espaço vetorial U através de A, temos uma correspondente "deformação" (natural) do espaço vetorial U^{**} tal que I_U naturalmente acompanha este processo. Um outro jeito de expressar a naturalidade de I_V é dizer que a definição de I_V não envolve nenhuma escolha arbitrária (tal como, por exemplo, uma escolha de base linear).

6.1. Proposição. *Sejam U, V, W $\mathbb{K}$-espaços vetoriais e sejam $V_1, V_2 \le V$ subespaços tais que $V = V_1 \oplus V_2$. Então temos os isomorfismos naturais*

$$i : \operatorname{Lin}_{\mathbb{K}}(U, V) \to \operatorname{Lin}_{\mathbb{K}}(U, V_1) \oplus \operatorname{Lin}_{\mathbb{K}}(U, V_2), \qquad j : \operatorname{Lin}_{\mathbb{K}}(V, W) \to \operatorname{Lin}_{\mathbb{K}}(V_1, W) \oplus \operatorname{Lin}_{\mathbb{K}}(V_2, W)$$

dados pelas regras $i : A \mapsto (\pi_1 \circ A, \pi_2 \circ A)$ e $j : B \mapsto (B \circ j_1, B \circ j_2)$, onde $\pi_l : V \to V_l$ é a projeção e $j_l : V_l \hookrightarrow V$ é a aplicação de inclusão, $l = 1, 2$.

Demonstração. A linearidade de i e j segue da bilinearidade de $\circ$ (vide o Exemplo 5.2.6). Se $iA = 0$, então $\pi_1 \circ A = 0$ e $\pi_2 \circ A = 0$ implicando $\pi_1(Au) = 0$ e $\pi_2(Au) = 0$ para todo $u \in U$. Daí, $Au = 0$ para todo $u \in U$. Logo, $A = 0$. Pelo Lema 5.7, i é um monomorfismo. Se $jB = 0$, temos $B \circ j_1 = 0$ e $B \circ j_2 = 0$ implicando $BV_1 = 0$ e $BV_2 = 0$. Daí, $BV = B(V_1 + V_2) = 0$ e $B = 0$. Pelo Lema 5.7, j é um monomorfismo.

Sejam $A_1 : U \to V_1$ e $A_2 : U \to V_2$ aplicações lineares. Definimos $A : U \to V$ pela fórmula $A = j_1 \circ A_1 + j_2 \circ A_2$. Então $\pi_l \circ A = A_l$ para $l = 1, 2$, pois

$$\pi_1 \circ j_1 = 1_{V_1}, \qquad \pi_1 \circ j_2 = 0, \qquad \pi_2 \circ j_1 = 0, \qquad \pi_2 \circ j_2 = 1_{V_2}. \tag{6.2}$$

Logo, $iA = (A_1, A_2)$. Em outras palavras, i é um epimorfismo. Concluímos que i é um isomorfismo.

Sejam $B_1 : V_1 \to W$ e $B_2 : V_2 \to W$ aplicações lineares. Definimos $B : V \to W$ pela regra $B : v_1 + v_2 \mapsto B_1 v_1 + B_2 v_2$, ou seja, $B(v_1 + v_2) := B_1 v_1 + B_2 v_2$ para todos $v_1 \in V_1$ e $v_2 \in V_2$. Uma verificação imediata mostra que B é linear. É fácil ver que $jB = (B_1, B_2)$. Em outras palavras, j é um epimorfismo. Concluímos que j é um isomorfismo ■

A Proposição 6.1 afirma que aplicações lineares para ou de uma soma direta são de fato pares de aplicações. Além das relações (6.2) utilizadas na demonstração da Proposição 6.1, é fácil verificar a relação

$$j_1 \circ \pi_1 + j_2 \circ \pi_2 = 1_V. \tag{6.3}$$

Na verdade, $I_V : V \to V^{**}$ é um monomorfismo. Verificamos este fato apenas para V finitamente gerado. Realmente, se $I_V v = 0$ para algum $v \in V$, então $fv = 0$ para todo $f \in V^*$. Pelo Corolário 4.14, podemos achar um subespaco $W \le V$ complementar a $\mathbb{K}v$, $V = \mathbb{K}v \oplus W$. Suponhamos que $v \ne 0$. Então v é uma base linear de $\mathbb{K}v$ e, pelo Lema 5.3, podemos encontrar um funcional linear $g : \mathbb{K}v \to \mathbb{K}$ tal que $gv = 1$. Pela Proposição 6.1, existe um funcional linear $f : V \to \mathbb{K}$ tal que $jf = (g, 0)$. Obviamente, $fv = gv = 1 \ne 0$. Uma contradição. Logo, $v = 0$ e, pelo Lema 5.7, I_V é um monomorfismo.

Seja $\beta : b_1, b_2, \dots, b_n$ uma base linear de V. Para todo $j = 1, 2, \dots, n$, pelo Lema 5.3, existe um único funcional linear $b_j^* : V \to \mathbb{K}$ tal que $b_j^* b_j = 1$ e $b_j^* b_i = 0$ para $i \ne j$. Então $\beta^* : b_1^*, b_2^*, \dots, b_n^*$ é uma base linear de V^* chamada *dual* a β. Realmente, se $\sum\limits_{j=1}^{n} k_j b_j^* = 0$, então $0 = \Big(\sum\limits_{j=1}^{n} k_j b_j^*\Big) b_i =$

$\sum_{j=1}^{n} k_j b_j^* b_i = k_i$ para todo i. Em outras palavras, os $b_1^*, b_2^*, \ldots, b_n^*$ são LI. Seja $f \in V^*$. Demonstramos que $f = \sum_{j=1}^{n} (fb_j)b_j^*$. Pelo Corolário 5.4, basta verificar que $fb_i = \Big(\sum_{j=1}^{n} (fb_j)b_j^*\Big)b_i$ para todo i. Isto é imediato: $\Big(\sum_{j=1}^{n} (fb_j)b_j^*\Big)b_i = \sum_{j=1}^{n} (fb_j)(b_j^* b_i) = fb_i$. Em particular, $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} V^* = \dim_{\mathbb{K}} V^{**}$ se V é finitamente gerado.

Daí concluímos que $I_V : V \to V^{**}$ é um isomorfismo natural para V finitamente gerado. Este isomorfismo pode ser visto como uma identificação. Isto significa que podemos pensar que $V = V^{**}$. Deste modo, a expressão fv para $v \in V$ e $f \in V^*$ pode ser lida de duas maneiras:

1. O escalar fv é o valor de f em v. Neste caso, interpretamos f como um funcional linear sobre V.
2. O escalar fv é o valor de v em f. Neste caso, interpretamos v como um funcional linear sobre V^*.

Seja $A : U \to V$ uma aplicação linear entre $\mathbb{K}$-espaços vetoriais finitamente gerados. Então, considerando I_U e I_V como aplicações idênticas, temos $A^{**} = A$ pela naturalidade de I_V.

6.4. Proposição. *Seja $A : U \to V$ uma aplicação linear entre $\mathbb{K}$-espaços vetoriais finitamente gerados. Então* $\operatorname{rk} A = \operatorname{rk} A^*$.

Demonstração. Denotamos por $N := A^{-1}0$ o núcleo de A e por $W' := AU$ a imagem de A. Seja $W \le U$ um subespaço complementar a N e seja $N' \le V$ um subespaço complementar a W', isto é, $U = N \oplus W$ e $V = N' \oplus W'$. Denotamos também as correspondentes injeções e projeções:

$$N \oplus W \overset{j_2}{\hookleftarrow} W, \qquad N' \overset{j_1'}{\hookrightarrow} N' \oplus W' \overset{j_2'}{\hookleftarrow} W', \qquad N' \xleftarrow{\pi_1'} N' \oplus W' \xrightarrow{\pi_2'} W'.$$

Note que a igualdade $\pi_1' \circ A = 0$ e a relação análoga à (6.3) implicam $A = 1_V \circ A = (j_1' \circ \pi_1' + j_2' \circ \pi_2') \circ A = j_2' \circ \pi_2' \circ A$. Pelo Lema 5.9, $\operatorname{rk} A = \dim_{\mathbb{K}} W$. Pela observação acima, $\dim_{\mathbb{K}} W = \dim_{\mathbb{K}} W^*$.

Basta mostrar que a aplicação $\varphi : A^*V^* \to W^*$ dada pela regra $\varphi : g \mapsto g \circ j_2$ para $g \in A^*V^* \le U^*$ é um isomorfismo. O fato que φ é linear segue da bilinearidade de $\circ$ (Exemplo 5.2.6).

Suponhamos que $\varphi g = 0$ para $g \in A^*V^*$. Então $g = A^*f$ para algum $f \in V^*$. Um elemento arbitrário $u \in U$ tem a forma $u = n + w$ com $n \in N$ e $w \in W$. Claramente, $An = 0$ e $j_2 w = w$. Logo,

$$(A^*f)u = (f \circ A)(n + w) = f\big(A(n + w)\big) = f(An + Aw) = f(Aw) =$$
$$= f\big(A(j_2 w)\big) = (f \circ A \circ j_2)w = \big((A^*f) \circ j_2\big)w = (\varphi g)w = 0.$$

Concluímos que $g = A^*f = 0$. Assim, φ é um monomorfismo.

Para todo $u \in U$, temos $Au \in W'$. Portanto, $(\pi_2' \circ A)u = \pi_2'(Au) = Au \in W'$ para todo $u \in U$. Pelo Lema 5.9, $I := \pi_2' \circ A \circ j_2 : W \to W'$ é um isomorfismo (lembre-se que $A|_W = A \circ j_2$). Denotamos por $I' : W' \to W$ o isomorfismo inverso a I.

Seja $h \in W^*$ um funcional linear. Então $h \circ I' \in \operatorname{Lin}_{\mathbb{K}}(W', \mathbb{K})$. Pela Proposição 6.1, existe um funcional linear $f \in \operatorname{Lin}_{\mathbb{K}}(V, \mathbb{K})$ tal que $jf = (0, h \circ I')$. Isto implica (vide a definição de j na Proposição 6.1) que $f \circ j_2' = h \circ I'$. De $I' \circ I = 1_W$ segue $f \circ j_2' \circ I = h$, ou seja, $h = f \circ j_2' \circ \pi_2' \circ A \circ j_2 = f \circ A \circ j_2 = \varphi(A^*f)$ pela relação $j_2' \circ \pi_2' \circ A = A$ demonstrada acima. Concluímos que φ é um epimorfismo ∎

6.5. Dicionário. Sejam U e V $\mathbb{K}$-espaços vetoriais, seja $\beta : b_1, b_2, \ldots, b_n$ uma base linear de U e seja $\gamma : c_1, c_2, \ldots, c_m$ uma base linear de V. Então, para qualquer aplicação linear $A : U \to V$, temos $[A^*]_{\beta^*}^{\gamma^*} = \big([A]_\gamma^\beta\big)^t$, onde M^t denota a matriz M transposta. Realmente, $[A]_\gamma^\beta = [a_{ij}]$ com os coeficientes $a_{ij} \in \mathbb{K}$ determinados pelas igualdades $Ab_j = \sum_{i=1}^{m} a_{ij} c_i$, $j = 1, 2, \ldots, n$. Para desenvolver o elemento

$A^*c_s^*$ na forma de uma combinação linear dos b_j^*'s, utilizamos a fórmula $g = \sum_{j=1}^{n}(gb_j)b_j^*$ demonstrada acima para qualquer $g \in U^*$. Assim, temos

$$A^*c_s^* = \sum_{j=1}^{n}\big((A^*c_s^*)b_j\big)b_j^* = \sum_{j=1}^{n}\big((c_s^* \circ A)b_j\big)b_j^* = \sum_{j=1}^{n}\big(c_s^*(Ab_j)\big)b_j^* =$$
$$= \sum_{j=1}^{n}\bigg(c_s^*\Big(\sum_{i=1}^{m}a_{ij}c_i\Big)\bigg)b_j^* = \sum_{j=1}^{n}\Big(\sum_{i=1}^{m}a_{ij}c_s^*c_i\Big)b_j^* = \sum_{j=1}^{n}a_{sj}b_j^*.$$

Isto significa que $[A^*]_{\beta^*}^{\gamma^*} = [a_{sj}]^t$.

6.6. Notação. Seja $M \in \operatorname{Matr}_{m\times n}\mathbb{K}$ uma matriz. Denotamos por $M_1, M_2, \dots, M_m$ todas as m sucessivas linhas de M e por $M^1, M^2, \dots, M^n$ todas as n sucessivas colunas de M. Os fatos que a matriz M está composta das suas próprias linhas e das suas próprias colunas podem ser agora escritos como

$$M = [M^1\, M^2 \dots M^n], \qquad M = \begin{bmatrix} M_1 \\ M_2 \\ \vdots \\ M_m \end{bmatrix}.$$

No mesmo estilo, para as matrizes $A \in \operatorname{Matr}_{m\times n}\mathbb{K}$ e $B \in \operatorname{Matr}_{m\times l}\mathbb{K}$, denotamos por $[A \mid B]$ a matriz $[A^1\, A^2 \dots A^n\, B^1\, B^2 \dots B^l] \in \operatorname{Matr}_{m\times(n+l)}\mathbb{K}$.

Seja $M \in \operatorname{Matr}_{m\times n}\mathbb{K}$ uma matriz. Consideramos o $\mathbb{K}$-espaço vetorial $\mathbb{K}^m$ como formado por colunas. A dimensão do subespaço de $\mathbb{K}^m$ gerado por todas as colunas de M se chama *posto* de M e é denotada por $\operatorname{rk} M$.

6.7. Dicionário. Seja $A : U \to V$ uma aplicação linear entre $\mathbb{K}$-espaços vetoriais, seja $\beta : b_1, b_2, \dots, b_n$ uma base linear de U e seja $\gamma : c_1, c_2, \dots, c_m$ uma base linear de V. Denotamos $M := [A]_\gamma^\beta$. Como foi observado no Dicionário 4.16, o j-ésimo elemento b_j da base linear β corresponde à coluna cujo único coeficiente não-nulo é igual a 1 e está no j-ésimo lugar. Daí, pela fórmula $[A]_\gamma^\beta[v]_\beta = [Av]_\gamma$, obtemos $M^j = M[b_j]_\beta = [Ab_j]_\gamma$. Em palavras: a j-ésima coluna da matriz $[A]_\gamma^\beta$ corresponde a Ab_j. Agora, pelo Dicionário 4.16, concluímos que a imagem AU corresponde ao subespaço de $\mathbb{K}^m$ gerado pelas colunas de $[A]_\gamma^\beta$. Em particular, $\operatorname{rk}[A]_\gamma^\beta = \operatorname{rk} A$. É fácil ver que o núcleo de A corresponde ao subespaço de $\mathbb{K}^n$ formado por todas as soluções do sistema homogêneo $MX = 0$ (vide o Exemplo 4.1.7).

6.8. Corolário. *Seja M uma matriz. Então $\operatorname{rk} M = \operatorname{rk} M^t$. Em palavras: o posto de uma matriz definido através de colunas e o definido através de linhas coincidem*[6] ∎

Fazer algo natural normalmente é contrário a um ato da escolha violenta, tal como a de base linear ou de coordenadas.[7] Há pessoas, (a maioria dos autores de livros de álgebra linear) que consideram o espaço $\mathbb{K}^n$ como o principal objeto de estudo na álgebra linear. Essa visão parece a tentativa de arrumar uma cama de Procrusto[8] retangular de matrizes para as aplicações lineares, obscurecendo assim a verdadeira natureza de tais aplicações. As matrizes naturalmente aparecem no estudo de aplicações lineares pois têm origem de somas diretas; mas mesmo a decomposição do espaço na soma direta dos unidimensionais é um ato de violência pois não é natural nem única.

Quando precisamos efetuar cálculos explícitos para obter um resultado numérico que é necessário numa aplicação prática, as matrizes podem ser realmente bem-vindas. Neste caso, sim, fazemos uma

[6] Vide também a Observação 8.13.
[7] Hermann Weyl (1885–1955): "The introduction of numbers as coordinates ... is an act of violence ..."
[8] Um bandido grego, dos antigos, famoso pela sua cama de ferro ... vide http://en.wikipedia.org/wiki/Procrustes

violência, mas essa pode ser comparada com a de um cirurgião e não tem nada a ver com a de um bandido. O dicionário serve para usar a língua mais adequada à situação. Não fale Alemão com Deus![9]

7. Sistemas de equações lineares. Matrizes elementares

Um sistema de equações lineares tem a forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m \end{cases}$$

e pode ser escrito na forma matricial $AX = B$, onde $A := [a_{ij}] \in \operatorname{Matr}_{m\times n}\mathbb{K}$ se chama matriz do sistema, $X := \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ é a coluna de variáveis e $B := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$. A matriz $[A\,B]$ se chama matriz *aumentada* do sistema. O sistema $AX = 0$ se chama sistema *homogêneo* associado ao sistema $AX = B$. Sabemos (vide Exemplo 4.1.7) que todas as soluções do sistema homogêneo formam um subespaço em $\mathbb{K}^n$.

7.1. Observação. *Seja S_0 uma solução particular do sistema $AX = B$. Então todas as soluções do sistema $AX = B$ constituem o conjunto $\{S_0 + S \mid AS = 0\}$* ■

7.2. Lema. *O sistema $AX = B$ admite uma solução se, e só se,* $\operatorname{rk} A = \operatorname{rk}[A\,B]$.

Demonstração. O sistema pode ser escrito na forma $x_1A^1 + x_2A^2 + \cdots + x_nA^n = B$. Portanto, $\operatorname{rk} A = \operatorname{rk}[A\,B]$ se existe uma solução. Suponhamos que $\operatorname{rk} A = \operatorname{rk}[A\,B]$. Então o subespaço gerado por $A^1, A^2, \dots, A^n$ contém B. Logo, $x_1A^1 + x_2A^2 + \cdots + x_nA^n = B$ para alguns $x_1, x_2, \dots, x_n \in \mathbb{K}$ ■

Recordemos o método de Gauss-Jordan de solução de um sistema de equações lineares. Fazendo as seguintes operações elementares com a matriz aumentada do sistema

- Trocar a posição de duas linhas da matriz.
- Multiplicar uma linha da matriz por um escalar diferente de zero.
- Somar a uma linha da matriz um múltiplo escalar de outra linha.

podemos conseguir a matriz *escalonada reduzida* que se caracteriza pelas propriedades

- Todas as linhas nulas ocorrem abaixo das não-nulas.
- O primeiro coeficiente não-nulo de cada linha não-nula, chamado *pivô,* é igual a 1.
- O pivô da $(i+1)$-ésima linha não-nula está à direita do da i-ésima.
- Na coluna de um pivô, todos os outros coeficientes são nulos.

(Omitindo a segunda exigência, caracterizamos uma matriz *escalonada semi-reduzida.* Para conseguir a escalonada semi-reduzida, as operações elementares do segundo tipo são desnecessárias.) Se o pivô de uma linha está na última coluna, o sistema não admite soluções. Caso contrário, chamamos *livres* as variáveis que não correspondem às colunas com pivôs. Essas servem como parâmetros da solução geral do sistema. Ainda mais, a solução geral obtida deste modo já providencia uma solução particular S_0 e uma base linear de soluções do sistema homogêneo associado. Um exemplo numérico: Seja $\begin{bmatrix} 1 & 0 & 3 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 & 0 & 3 \\ 0 & 0 & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ a matriz aumentada do sistema. Ela já está escalonada reduzida. As variáveis

[9]Imperador Carlos V (1500–1558) : "Eu falo Espanhol com Deus, Italiano com as mulheres, Francês com os homens e Alemão com meu cavalo".

livres são x_3 e x_5. O sistema correspondente tem a forma $\begin{cases} x_1 + 3x_3 &= 2 \\ x_2 + 2x_3 &= 3 \\ x_4 + 4x_5 &= 1 \end{cases}$. Considerando as variáveis livres como parâmetros, $x_3 := p_1$ e $x_5 := p_2$, obtemos a solução geral do sistema

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 0 \\ 1 \\ 0 \end{bmatrix} + p_1 \begin{bmatrix} -3 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + p_2 \begin{bmatrix} 0 \\ 0 \\ 0 \\ -4 \\ 1 \end{bmatrix} := S_0 + p_1 S_1 + p_2 S_2 \qquad p_1, p_2 \in \mathbb{K}.$$

Aqui S_0 é uma solução particular do sistema e S_1, S_2 formam uma base linear de soluções do sistema homogêneo associado.[10]

7.3. Observação. *Seja $A \in \operatorname{Matr}_{m\times n}\mathbb{K}$ e seja $B \in \operatorname{Matr}_{m\times l}\mathbb{K}$. Se a matriz $[A \mid B]$ é escalonada reduzida, então A é escalonada reduzida* ■

7.4. Observação. *Seja $A \in \operatorname{Matr}_{m\times m}\mathbb{K}$ uma matriz quadrada escalonada reduzida. Então, ou $A = 1_{m\times m}$, ou a última linha de A é nula, $A_m = 0$.*

Demonstração. Suponhamos que $A_m \neq 0$. Então A não possui linhas nulas. Se o pivô da i-ésima linha fica na posição ii-ésima para todo i, temos $A = 1_{m\times m}$. Caso contrário, um dos pivôs fica estritamente à direita da diagonal principal. Isto, lembrando-se que a matriz A é quadrada, não deixa nenhum espaço para o pivô da m-ésima linha ■

Podemos descrever o processo de escalonamento e pivotização utilizando matrizes elementares. Fixamos m. Denotamos por e_{ij} a $(m\times m)$-matriz cujo único coeficiente não-nulo está na ij-ésima posição e é igual a 1.

Sejam $1 \leq i, j \leq m$, $i \neq j$, dois índices distintos. Denotamos por E_{ij} a $(m\times m)$-matriz que difere da matriz identidade $1_{m\times m}$ somente nas posições ii, ij, ji e jj. Os correspondentes coeficientes de E_{ij} são 0, 1, 1 e 0. Podemos também definir $E_{ij} := 1_{m\times m} - e_{ii} + e_{ij} + e_{ji} - e_{jj}$. O leitor pode facilmente verificar que, para qualquer $(m\times n)$-matriz A, a matriz $E_{ij}A$ é a matriz A com as linhas A_i e A_j trocadas. Em outras palavras, a primeira operação elementar se realiza através da multiplicação à esquerda por uma matriz do tipo E_{ij}.

Seja $1 \leq i \leq m$ e seja $0 \neq k \in \mathbb{K}$. Denotamos por $E_i(k)$ a $(m\times m)$-matriz que difere da matriz identidade $1_{m\times m}$ somente na posição ii e tal que o ii-ésimo coeficiente de $E_i(k)$ é igual a k. Podemos também definir $E_i(k) := 1_{m\times m} + (k-1)e_{ii}$. É fácil ver que a multiplicação à esquerda por uma matriz do tipo $E_i(k)$ realiza a segunda operação elementar.

Sejam $1 \leq i, j \leq m$, $i \neq j$, dois índices distintos e seja $k \in \mathbb{K}$. Denotamos $E_{ij}(k) := 1_{m\times m} + ke_{ij}$. Para qualquer $(m\times n)$-matriz A, a matriz $e_{ij}A$ é a $(m\times n)$-matriz que possui uma única linha não-nula, a i-ésima, igual a A_j. Portanto, $E_{ij}(k)A$ é a matriz A com uma única mudança: sua i-ésima linha é igual a $A_i + kA_j$. Em outras palavras, a terceira operação elementar realiza-se através da multiplicação à esquerda por uma matriz do tipo $E_{ij}(k)$.

As matrizes dos três tipos descritos acima chamam-se *elementares.* Assim, para fazer uma matriz $A \in \operatorname{Matr}_{m\times n}\mathbb{K}$ escalonada reduzida, multiplicâmo-la à esquerda (sucessivamente) por algumas matrizes elementares $E_1, E_2, \dots, E_l \in \operatorname{Matr}_{m\times m}\mathbb{K}$ de modo que a matriz $E_l \dots E_2E_1A$ fique escalonada reduzida. Para conseguir uma matriz escalonada semi-reduzida precisamos apenas das matrizes elementares do primeiro e do terceiro tipos.

[10]O fato que S_1, S_2 são LI segue de uma óbvia observação, válida em geral: Seja x_i uma variável livre. Então os i-ésimos coeficientes das colunas-soluções do sistema homogêneo são todos nulos, exceto aquele correspondendo à própria variável x_i, que é igual a 1.

Por um cálculo direto, obtemos a

7.5. Observação. *Sejam $1 \le i, j \le m$, $i \ne j$, e seja $k \in \mathbb{K}$. Então*

$$E_{ij}E_{ij} = E_{ij}(k)E_{ij}(-k) = E_{ij}(-k)E_{ij}(k) = 1_{m\times m}.$$

Se $k \ne 0$, temos também $E_i(k)E_i(k^{-1}) = E_i(k^{-1})E_i(k) = 1_{m\times m}$ ■

Recordamos que uma matriz $A \in \operatorname{Matr}_{m\times n}\mathbb{K}$ é dita *inversível* se existe uma matriz $B \in \operatorname{Matr}_{n\times m}\mathbb{K}$, chamada *inversa* a A e denotada por A^{-1}, tal que $AB = 1_{m\times m}$ e $BA = 1_{n\times n}$. Note que tal inversa é única: se B' é uma outra inversa, obtemos $B' = 1_{n\times n}B' = (BA)B' = B(AB') = B1_{m\times m} = B$. Na verdade, pelos Dicionários 5.11 e 5.12, as matrizes inversíveis correspondem aos isomorfismos. Logo, tais matrizes são necessariamente quadradas (espaços vetoriais isomorfos têm a mesma dimensão). A Observação 7.5 diz que todas as matrizes elementares são inversíveis.

7.6. Observação. *Sejam $M_1, M_2, \dots, M_l$ $(m\times m)$-matrizes inversíveis. Então a matriz $M_1M_2\dots M_l$ é inversível e $(M_1M_2\dots M_l)^{-1} = M_l^{-1}\dots M_2^{-1}M_1^{-1}$.*

Demonstração. Temos

$$M_1M_2\dots M_{l-1}M_lM_l^{-1}M_{l-1}^{-1}\dots M_2^{-1}M_1^{-1} = M_1M_2\dots M_{l-1}M_{l-1}^{-1}\dots M_2^{-1}M_1^{-1} =$$
$$= \dots = M_1M_2M_2^{-1}M_1^{-1} = M_1M_1^{-1} = 1.$$

De modo semelhante,

$$M_l^{-1}\dots M_2^{-1}M_1^{-1}M_1M_2\dots M_l = M_l^{-1}\dots M_2^{-1}M_2\dots M_l = \dots = M_l^{-1}M_l = 1$$ ■

7.7. Observação. *Sejam $M \in \operatorname{Matr}_{s\times m}\mathbb{K}$, $A \in \operatorname{Matr}_{m\times n}\mathbb{K}$ e $B \in \operatorname{Matr}_{m\times l}\mathbb{K}$. Então $M[A \mid B] = [MA \mid MB]$.*

Demonstração. Basta observar que $M[A^1\,A^2\dots A^n] = [MA^1\,MA^2\dots MA^n]$ ■

7.8. Teorema. *Seja $A \in \operatorname{Matr}_{m\times m}\mathbb{K}$ uma matriz quadrada. Seja $M \in \operatorname{Matr}_{m\times m}\mathbb{K}$ o produto de matrizes elementares tal que $M[A \mid 1_{m\times m}]$ é uma matriz escalonada reduzida. Então A é inversível se, e só se, a matriz MA não possui linhas nulas. Neste caso, $M = A^{-1}$.*

Demonstração. Pela Observação 7.7, $M[A \mid 1_{m\times m}] = [MA \mid M]$. Pela Observação 7.3, a matriz MA é escalonada reduzida. Pelas Observações 7.5 e 7.6, a matriz M é inversível. Aplicamos agora a Observação 7.4 à matriz MA. Se $MA = 1_{m\times m}$, então, multiplicando essa igualdade à esquerda por M^{-1}, obtemos $A = M^{-1}$. Portanto, A é inversível e $A^{-1} = M$. Se a última linha de MA é nula, então $[0\dots01]MA = 0$. Neste caso, a matriz A não pode ser inversível. Caso contrário, multiplicando a última igualdade à direita pela inversa a MA, obtemos $[0\dots01] = 0$. Uma contradição ■

O leitor deve notar que o Lema 4.6 tem algo a ver com a terceira operação elementar.

8. Determinante

Seja $A = [a_{ij}] \in \operatorname{Matr}_{n\times n}\mathbb{K}$ uma matriz quadrada. Definimos o *determinante* de A pela fórmula

$$\det A = \sum_{(j_1j_2\cdots j_n)} (-1)^{I(j_1j_2\cdots j_n)}a_{1j_1}a_{2j_2}\dots a_{nj_n},$$

onde $(j_1j_2\dots j_n)$ percorre todas as permutações de $1, 2, \dots, n$ e o símbolo $I(j_1j_2\dots j_n)$ denota o número de inversões na permutação $(j_1j_2\dots j_n)$. Dizemos que $(j_1j_2\dots j_n)$ é uma *permutação* de $1, 2, \dots, n$ se $\{j_1, j_2, \dots, j_n\} = \{1, 2, \dots, n\}$, ou seja, se a seqüência $j_1, j_2, \dots, j_n$ representa os mesmos números $1, 2, \dots, n$ listados apenas em uma ordem (possivelmente) diferente. Uma *inversão* na permutação $(j_1j_2\dots j_n)$ ocorre quando $j_\alpha > j_\beta$, mas $\alpha < \beta$. Assim, $I(j_1j_2\dots j_n)$ é o número de todas as ocorrências de pares dos j_k's na ordem decrescente. Por exemplo, $I(12\dots n) = 0$. De fato, no cálculo do determinante, precisamos somente da paridade do número $I(j_1j_2\dots j_n)$. Em seguida, utilizaremos sem demonstração o seguinte

8.1. Fato. *Seja $(j_1j_2\dots j_n)$ uma permutação de $1,2,\dots,n$. Então, para $1\le\alpha<\beta\le n$, os números $I(j_1j_2\dots j_\alpha\dots j_\beta\dots j_n)$ e $I(j_1j_2\dots j_\beta\dots j_\alpha\dots j_n)$ têm paridades distintas.*

Este fato possibilita determinar a paridade de uma permutação $(j_1j_2\dots j_n)$ fazendo *transposições* de pares de índices (isto é, trocando como acima j_α e j_β com $\alpha\ne\beta$) até obter a permutação $(12\dots n)$ com $I(12\dots n)=0$.

Na definição de determinante, cada termo $a_{1j_1}a_{2j_2}\dots a_{nj_n}$ no somatório tem exatamente um fator em cada linha da matriz A. Sendo $(j_1j_2\dots j_n)$ uma permutação de $1,2,\dots,n$, neste mesmo termo temos exatamente um fator em cada coluna de A. Assim, podemos ver que os termos no somatório percorrem exatamente tais produtos de n coeficientes de A que contêm um fator em cada linha de A e um fator em cada coluna de A.

Consideremos o termo $a_{1j_1}a_{2j_2}\dots a_{nj_n}$. Ele não se altera se fizermos uma transposição de dois fatores. Fazendo tais transposições, podemos finalmente ordenar os índices j_α's. Em cada etapa, temos o termo escrito em uma forma do tipo $a_{s_1t_1}a_{s_2t_2}\dots a_{s_nt_n}$. Quando fazemos a transposição dos fatores $a_{s_kt_k},a_{s_lt_l}$, onde $1\le k<l\le n$, efetuamos de fato as transposições simultâneas de s_k,s_l na permutação $(s_1s_2\dots s_k\dots s_l\dots s_n)$ e de t_k,t_l na permutação $(t_1t_2\dots t_k\dots t_l\dots t_n)$. Pelo Fato 8.1, isto implica que, quando chegamos ao produto $a_{i_11}a_{i_22}\dots a_{i_nn}$, os sinais $(-1)^{I(i_1i_2\dots i_n)}$ e $(-1)^{I(j_1j_2\dots j_n)}$ são os mesmos, onde $(j_1j_2\dots j_n)$ é a permutação relacionada à forma original $a_{1j_1}a_{2j_2}\dots a_{nj_n}$ do termo em questão. Em outras palavras, provamos que

$$\sum_{(j_1j_2\cdots j_n)}(-1)^{I(j_1j_2\cdots j_n)}a_{1j_1}a_{2j_2}\dots a_{nj_n}=\sum_{(i_1i_2\cdots i_n)}(-1)^{I(i_1i_2\cdots i_n)}a_{i_11}a_{i_22}\dots a_{i_nn},$$

ou seja, chegamos ao seguinte

8.2. Corolário. *Para toda matriz $A\in\operatorname{Matr}_{n\times n}\mathbb{K}$, temos $\det A=\det A^t$* ■

Seja $A=[a_{ij}]\in\operatorname{Matr}_{n\times n}\mathbb{K}$, seja $1\le l\le n$ e seja $k\in\mathbb{K}$. Podemos escrever A como composta das suas linhas, $A=\left[\begin{smallmatrix}*\\A_l\\ *\end{smallmatrix}\right]$. Substituímos a l-ésima linha A_l pela linha $A_l+A'_l$, onde $A'_l=[a'_{l1}\,a'_{l2}\dots a'_{ln}]$. Temos

$$\det\left[\begin{smallmatrix}*\\A_l+A'_l\\ *\end{smallmatrix}\right]=\sum_{(j_1j_2\cdots j_l\cdots j_n)}(-1)^{I(j_1j_2\cdots j_l\cdots j_n)}a_{1j_1}a_{2j_2}\dots(a_{lj_l}+a'_{lj_l})\dots a_{nj_n}=$$

$$\sum_{(j_1j_2\cdots j_l\cdots j_n)}(-1)^{I(j_1j_2\cdots j_l\cdots j_n)}a_{1j_1}a_{2j_2}\dots a_{lj_l}\dots a_{nj_n}+\sum_{(j_1j_2\cdots j_l\cdots j_n)}(-1)^{I(j_1j_2\cdots j_l\cdots j_n)}a_{1j_1}a_{2j_2}\dots a'_{lj_l}\dots a_{nj_n}$$

$$=\det\left[\begin{smallmatrix}*\\A_l\\ *\end{smallmatrix}\right]+\det\left[\begin{smallmatrix}*\\A'_l\\ *\end{smallmatrix}\right].$$

Do mesmo modo,

$$\det\left[\begin{smallmatrix}*\\kA_l\\ *\end{smallmatrix}\right]=\sum_{(j_1j_2\cdots j_l\cdots j_n)}(-1)^{I(j_1j_2\cdots j_l\cdots j_n)}a_{1j_1}a_{2j_2}\dots(ka_{lj_l})\dots a_{nj_n}=$$

$$=k\sum_{(j_1j_2\cdots j_l\cdots j_n)}(-1)^{I(j_1j_2\cdots j_l\cdots j_n)}a_{1j_1}a_{2j_2}\dots a_{lj_l}\dots a_{nj_n}=k\det\left[\begin{smallmatrix}*\\A_l\\ *\end{smallmatrix}\right].$$

Chegamos à

8.3. Proposição. *Para todo l, o determinante* $\det A$ *é linear na l-ésima linha de A* (*fixando todas as outras linhas de A*) ■

Essa propriedade do determinante pode ser também assim expressa: o determinante é *multi-linear* nas linhas da matriz.

Seja $A = [a_{ij}] \in \operatorname{Matr}_{n\times n}\mathbb{K}$ e sejam $1 \le p < q \le n$. Trocando as linhas p-ésima e q-ésima de A, pelo Fato 8.1, obtemos

$$\det\begin{bmatrix} * \\ A_q \\ * \\ A_p \\ * \end{bmatrix} = \sum_{(j_1 j_2 \dots j_p \dots j_q \dots j_n)} (-1)^{I(j_1 j_2 \dots j_p \dots j_q \dots j_n)} a_{1j_1} a_{2j_2} \dots a_{qj_p} \dots a_{pj_q} \dots a_{nj_n} =$$

$$= \sum_{(j_1 j_2 \dots j_p \dots j_q \dots j_n)} (-1)^{I(j_1 j_2 \dots j_p \dots j_q \dots j_n)} a_{1j_1} a_{2j_2} \dots a_{pj_q} \dots a_{qj_p} \dots a_{nj_n} =$$

$$= - \sum_{(j_1 j_2 \dots j_p \dots j_q \dots j_n)} (-1)^{I(j_1 j_2 \dots j_q \dots j_p \dots j_n)} a_{1j_1} a_{2j_2} \dots a_{pj_q} \dots a_{qj_p} \dots a_{nj_n} = -\det\begin{bmatrix} * \\ A_p \\ * \\ A_q \\ * \end{bmatrix}.$$

Chegamos ao

8.4. Corolário. *O determinante muda de sinal ao se trocar duas linhas da matriz* ■

Essa propriedade do determinante se expressa dizendo-se que o determinante é *anti-simétrico* nas linhas da matriz. A anti-simetricidade do determinante claramente implica que o determinante de uma matriz com duas linhas iguais é nulo.

Pelo Corolário 8.2, são válidas as afirmações envolvendo colunas no lugar de linhas, análogas à Proposição 8.3 e ao Corolário 8.4.

Sejam $d_1, d_2, \dots, d_n \in \mathbb{K}$. Denotamos por $\operatorname{diag}(d_1, d_2, \dots, d_n)$ a correspondente $(n\times n)$-matriz *diagonal,* isto é, a matriz cujos únicos coeficientes não-nulos estão na diagonal principal e são respectivamente iguais a $d_1, d_2, \dots, d_n$.

8.5. Observação. *Sejam* $1 \le i, j \le n$, $i \ne j$, *sejam* $d_1, d_2, \dots, d_n \in \mathbb{K}$, *seja* $k \in \mathbb{K}$ *e seja* $A \in \operatorname{Matr}_{n\times n}\mathbb{K}$. *Então* $\det(E_{ij}A) = -\det A$, $\det\big(E_{ij}(k)A\big) = \det A$ *e* $\det\big(\operatorname{diag}(d_1, d_2, \dots, d_n)A\big) = d_1 d_2 \dots d_n \det A$. *Em particular,* $\det E_{ij} = -1$, $\det E_{ij}(k) = 1$ *e* $\det\operatorname{diag}(d_1, d_2, \dots, d_n) = d_1 d_2 \dots d_n$. *Assim, se* $k \ne 0$, *temos* $\det E_i(k) = k$.

Demonstração. A igualdade $\det(E_{ij}A) = -\det A$ segue do Corolário 8.4 e das considerações acima da Observação 7.5. Por mesmas considerações, pela Proposição 8.3 e pelo Corolário 8.4, temos

$$\det\big(E_{ij}(k)A\big) = \det\begin{bmatrix} * \\ A_i + kA_j \\ * \\ A_j \\ * \end{bmatrix} = \det\begin{bmatrix} * \\ A_i \\ * \\ A_j \\ * \end{bmatrix} + k\det\begin{bmatrix} * \\ A_j \\ * \\ A_j \\ * \end{bmatrix} = \det A.$$

A igualdade $\det\big(\operatorname{diag}(d_1, d_2, \dots, d_n)A\big) = d_1 d_2 \dots d_n \det A$ segue imediatamente da multi-linearidade do determinante pois, multiplicando uma matriz A por uma matriz diagonal, multiplicamos de fato as linhas de A por correspondentes coeficientes da matriz diagonal. Para o resto, basta tomar $A = 1_{n\times n}$ ■

8.6. Exercício. Uma matriz quadrada T chama-se *triangular superior* (respectivamente, *inferior*) se todos os coeficientes abaixo (respectivamente, acima) da diagonal principal de T são nulos. Para qualquer matriz triangular T, prove que $\det T$ é o produto de todos os coeficientes da diagonal principal de T.

Pela Observação 8.5, podemos calcular o determinante de A simplesmente escalonando A. Devido ao Exercício 8.6, para os fins deste cálculo, basta conseguir no final uma matriz triangular.

8.7. Observação. *Seja* $A \in \operatorname{Matr}_{n\times n}\mathbb{K}$. *Então* A *é inversível se, e só se,* $\det A \neq 0$.

Demonstração. Pela Observação 8.5, podemos supor que A é escalonada reduzida. Basta observar que $\det A = 0$ caso a última linha de A seja nula e aplicar a Observação 7.4 (ou o Teorema 7.8) ■

8.8. Teorema. *Sejam* $A, B \in \operatorname{Matr}_{n\times n}\mathbb{K}$. *Então* $\det(AB) = (\det A)(\det B)$.

Demonstração. Pela Observação 8.5, multiplicando as matrizes A e AB à esquerda por uma mesma matriz elementar E temos as semelhantes mudanças dos determinantes $\det A$ e $\det(AB)$. Assim, podemos supor que A é escalonada reduzida. Se $A = 1_{n\times n}$, a igualdade $\det(AB) = (\det A)(\det B)$ é óbvia. Caso contrário, pela Observação 7.3, a última linha de A é nula e $\det A = 0$. Temos $[0\dots 01]A = 0$. Logo, $[0\dots 01]AB = 0$ e AB não pode ser inversível. Pela Observação 8.7, $\det(AB) = 0$ ■

Para $A \in \operatorname{Matr}_{n\times n}\mathbb{K}$, denotamos por A_{ij} a matriz A com as i-ésima linha e j-ésima coluna retiradas. Chamamos A_{ij} o ij-ésimo *menor* de A. O número $(-1)^{i+j}\det A_{ij}$ é dito o ij-ésimo *cofator* de A. A matriz transposta à matriz formada por todos os cofatores de A se chama matriz *adjunta* a A e é denotada por $\operatorname{adj} A := \big[(-1)^{i+j}\det A_{ij}\big]^t$.

Seja $(jj_2j_3\dots j_n)$ uma permutação de $1,2,3,\dots,n$. Então $(-1)^{I(jj_2j_3\cdots j_n)} = (-1)^{j+1}(-1)^{I(j_2j_3\cdots j_n)}$. Realmente, fazendo transposições dos $j_2, j_3, \dots, j_n$, podemos reduzir o problema ao caso da permutação $\big(j\,1\,2\dots(j-1)\,(j+1)\dots n\big)$ que tem $j-1$ inversões. Essa observação implica que

$$\det A = \sum_{(jj_2j_3\cdots j_n)} (-1)^{I(jj_2j_3\cdots j_n)} a_{1j}a_{2j_2}a_{3j_3}\dots a_{nj_n} =$$

$$= \sum_{j=1}^{n} a_{1j}(-1)^{j+1} \sum_{(j_2j_3\cdots j_n)} (-1)^{I(j_2j_3\cdots j_n)} a_{2j_2}a_{3j_3}\dots a_{nj_n} = \sum_{j=1}^{n} a_{1j}(-1)^{j+1}\det A_{1j}$$

para qualquer matriz $A = [a_{ij}] \in \operatorname{Matr}_{n\times n}\mathbb{K}$. Trocando as linhas, pelo Corolário 8.4, obtemos a fórmula semelhante $\det A = \sum\limits_{j=1}^{n} a_{ij}(-1)^{i+j}\det A_{ij}$ para qualquer i, chamada *desenvolvimento* de Laplace de determinante pela i-ésima linha. Pelo Corolário 8.2, obtemos a fórmula $\det A = \sum\limits_{i=1}^{n} a_{ij}(-1)^{i+j}\det A_{ij}$ para todo j, chamada *desenvolvimento* de Laplace de determinante pela j-ésima coluna.

8.9. Proposição. *Seja* $A \in \operatorname{Matr}_{n\times n}\mathbb{K}$. *Então* $A(\operatorname{adj} A) = (\operatorname{adj} A)A = (\det A)1_{n\times n}$.

Demonstração. Já sabemos que $\sum\limits_{j=1}^{n} a_{ij}(-1)^{i+j}\det A_{ij} = \det A$ e $\sum\limits_{j=1}^{n}(-1)^{j+i}\det A_{ji}a_{ji} = \det A$ para todo i. Isto implica que todos os coeficientes das diagonais principais de $A(\operatorname{adj} A)$ e $(\operatorname{adj} A)A$ são iguais a $\det A$. Resta mostrar que $\sum\limits_{j=1}^{n} a_{ij}(-1)^{l+j}\det A_{lj} = 0$ para $i \neq l$ e que $\sum\limits_{i=1}^{n}(-1)^{i+j}\det A_{ij}a_{il} = 0$ para $j \neq l$. Suponhamos que $i \neq l$. Pelo desenvolvimento de Laplace pela l-ésima linha para a matriz $\begin{bmatrix} * \\ A_i \\ * \\ A_i \\ * \end{bmatrix}$ que é a matriz A com a l-ésima linha substituída por A_i, temos $\sum\limits_{j=1}^{n} a_{ij}(-1)^{l+j}\det A_{lj} = \det\begin{bmatrix} * \\ A_i \\ * \\ A_i \\ * \end{bmatrix} = 0$.

Suponhamos que $j \neq l$. Pelo desenvolvimento de Laplace pela j-ésima coluna para a matriz A com a j-ésima coluna substituída por A_l, obtemos $0 = \det[*A^l * A^l*] = \sum\limits_{i=1}^{n} a_{il}(-1)^{i+j} \det A_{ij}$ ■

Utilizando as Observação 8.7 e Proposição 8.9, obtemos a fórmula explícita para a matriz inversa:

8.10. Corolário. *Seja* $A \in \operatorname{Matr}_{n\times n} \mathbb{K}$ *com* $\det A \neq 0$. *Então* $A^{-1} = \dfrac{1}{\det A} \operatorname{adj} A$ ■

Aplicando essa fórmula para sistemas de equações lineares, chegamos à *regra* de *Cramer*:

8.11. Corolário. *Seja* $AX = B$ *um sistema de* n *equações lineares em* n *variáveis tal que* $\det A \neq 0$. *Então o sistema admite uma única solução dada pelas fórmulas* $x_j = \dfrac{\det[A^1\, A^2 \dots A^{j-1}\, B\, A^{j+1} \dots A^n]}{\det A}$, $j = 1, 2, \dots, n$.

Demonstração. Multiplicando a igualdade $AX = B$ à esquerda pela inversa a A, obtemos a única solução $X = A^{-1}B$. Pelo Corolário 8.10, temos $x_j = \dfrac{1}{\det A} \sum\limits_{i=1}^{n} (-1)^{i+j} A_{ij} b_i$. Resta observar que $\sum\limits_{i=1}^{n} b_i(-1)^{i+j} A_{ij}$ é o desenvolvimento de Laplace de $\det[A^1\, A^2 \dots A^{j-1}\, B\, A^{j+1} \dots A^n]$ pela j-ésima coluna ■

8.12. Observação. *Sejam* $A, M \in \operatorname{Matr}_{n\times n} \mathbb{K}$ *com* M *inversível. Então* $\det(M^{-1}) = (\det M)^{-1}$ *e* $\det(MAM^{-1}) = \det A$.

Demonstração. Pelo Teorema 8.8, $1 = \det 1_{n\times n} = \det(MM^{-1}) = (\det M)(\det M^{-1})$. Portanto, $\det(M^{-1}) = (\det M)^{-1}$. Agora,

$$\det(MAM^{-1}) = (\det M)(\det A)(\det M^{-1}) = (\det M)(\det A)(\det M)^{-1} = \det A$$

pelo Teorema 8.8 ■

8.13. Observação. *O posto de qualquer matriz* A *é igual ao tamanho máximo de uma submatriz quadrada* Q *de* A *com* $\det Q \neq 0$.

Demonstração. Retirando de A colunas que são LD das outras, obtemos uma submatriz B com o mesmo posto. Retirando de B linhas que são LD das outras, não alteramos o posto e, pelo Corolário 6.8, chegamos a uma submatriz quadrada Q cujo posto coincide com seu tamanho. Logo, $\det Q \neq 0$.

Reciprocamente, se Q é uma submatriz quadrada de A com $\det Q \neq 0$, então as colunas de Q são LI. Portanto, as correspondentes colunas de A são LI, implicando que o posto de A é maior ou igual ao tamanho de Q ■

Seja $A : V \to V$ uma aplicação linear. Escolhendo qualquer base linear β de V, definimos $\det A := \det[A]_\beta^\beta$. Pelos Corolário 5.15 e Observação 8.12, esta definição é correta, ou seja, não depende da escolha de base linear. Parece que, durante toda essa seção, temos falado Alemão com Deus ...

9. Formas simétricas bilineares e 1.5-lineares

Em seguida, lidamos somente com espaços vetoriais de dimensão finita. Para $k \in \mathbb{K}$, o símbolo $\overline{k}$ tem duas variantes de leitura:

1. $\overline{k}$ denota o número (complexo) k *conjugado.*
2. $\overline{k} = k$.

9.1. Definição. Seja V um $\mathbb{K}$-espaço vetorial. Uma *forma simétrica* (*hermitiana*) é uma aplicação $\langle -,-\rangle : V\times V\to\mathbb{K}$, $(v_1,v_2)\mapsto\langle v_1,v_2\rangle$, linear em v_1 e tal que $\langle v_1,v_2\rangle=\overline{\langle v_2,v_1\rangle}$ para todos $v_1,v_2\in V$. Em outras palavras, a forma é *bilinear* (1.5-*linear*), pois $\langle kv_1,v_2\rangle=k\langle v_1,v_2\rangle$ e $\langle v_1,kv_2\rangle=\overline{k}\langle v_1,v_2\rangle$ para todo $k\in\mathbb{K}$. Se $W\le V$ é um subespaço, podemos restringir a forma $\langle -,-\rangle$ para W, obtendo o espaço vetorial W munido da forma *induzida.*

9.2. Definição. Seja V um espaço vetorial munido de forma e seja $W\le V$. Define-se W *ortogonal* por $W^\perp:=\big\{v\in V\mid \forall w\in W\ \langle v,w\rangle=0\big\}$. O *núcleo* da forma em V é $V^\perp$. Caso o núcleo seja nulo, dizemos que a forma é *não-degenerada.* Se a forma induzida no subespaço $W\le V$ é não-degenerada, dizemos que W é *não-degenerado.* Para $U,W\le V$, o ortogonal de W *relativo* a U é dado por $W^{\perp U}:=W^\perp\cap U$.

9.3. Lema. *Seja V um espaço com forma e sejam $W,W_1,W_2\le V$. Então $W^\perp\le V$, $W\subset W^{\perp\perp}$ e $(W_1+W_2)^\perp=W_1^\perp\cap W_2^\perp$.*

Demonstração. Podemos reescrever a definição de $W^\perp$ como $W^\perp:=\big\{v\in V\mid\langle v,W\rangle=0\big\}$. Sejam $v,v_1,v_2\in W^\perp$ e $k\in\mathbb{K}$. Então $\langle v,W\rangle=\langle v_1,W\rangle=\langle v_2,W\rangle=0$. Portanto, $\langle v_1+v_2,W\rangle\subset\langle v_1,W\rangle+\langle v_2,W\rangle=0$ e $\langle kv,W\rangle=k\langle v,W\rangle=0$. Logo, $W^\perp\le V$.

Claramente, $\langle W^\perp,W\rangle=0$. Aplicando $^-$, obtemos $\langle W,W^\perp\rangle=0$. Isto implica $W\subset W^{\perp\perp}$.

O fato que $v\in W_1^\perp\cap W_2^\perp$ é equivalente a $\langle v,W_1\rangle=\langle v,W_2\rangle=0$. Por outro lado, $\langle v,W_1+W_2\rangle=\langle v,W_1\rangle+\langle v,W_2\rangle$. Daí concluímos que $v\in W_1^\perp\cap W_2^\perp$ é equivalente a $v\in(W_1+W_2)^\perp$ ■

9.4. Observação. *Seja V um espaço com forma e seja $W\le V$ o subespaço complementar ao núcleo $V^\perp$, i.e., $V=V^\perp\oplus W$. Então W é não-degenerado e a forma sobre V é determinada pela forma sobre W.*

Demonstração. O fato que a forma sobre V é determinada pela forma sobre W segue imediatamente de $\langle V^\perp,V\rangle=0$. Se $w\in W$ está no núcleo da forma sobre W, então $\langle w,W\rangle=0$. Por outro lado, $\langle w,V^\perp\rangle=0$. Logo, $\langle w,V\rangle=\langle w,V^\perp\rangle+\langle w,W\rangle=0$, ou seja, $w\in V^\perp$. Resta observar que $w\in V^\perp\cap W=0$ ■

9.5. Lema. *Seja V um espaço com forma e seja $W\le V$. Então* $\dim_{\mathbb{K}}W+\dim_{\mathbb{K}}W^\perp\ge\dim_{\mathbb{K}}V$.

Demonstração. Utilizamos indução sobre $\dim_{\mathbb{K}}W$. Caso $\dim_{\mathbb{K}}W=0$, a desigualdade é óbvia. Suponhamos que $\dim_{\mathbb{K}}W\ge1$. Então temos um elemento não-nulo $0\ne w\in W$ e podemos decompor $W=\mathbb{K}w\oplus W'$. Definimos um funcional linear $f:W'^\perp\to\mathbb{K}$ pela regra $f:v\mapsto\langle v,w\rangle$. Então $(\mathbb{K}w)^\perp\cap W'^\perp$ é o núcleo de f, $f^{-1}0=(\mathbb{K}w)^\perp\cap W'^\perp$. Realmente, $v\in W'^\perp$ está no núcleo de f se, e só se, $\langle v,w\rangle=0$. Isto é equivalente a $\langle v,\mathbb{K}w\rangle=0$.

Pelo Corolário 5.10, $\dim_{\mathbb{K}}f^{-1}0\ge\dim_{\mathbb{K}}W'^\perp-1$. Pelo Lema 9.3, $(\mathbb{K}w)^\perp\cap W'^\perp=W^\perp$. Pela hipótese de indução, $\dim_{\mathbb{K}}W'+\dim_{\mathbb{K}}W'^\perp\ge\dim_{\mathbb{K}}V$. Portanto,

$$\dim_{\mathbb{K}}W+\dim_{\mathbb{K}}W^\perp=1+\dim_{\mathbb{K}}W'+\dim_{\mathbb{K}}f^{-1}0\ge$$

$$\ge1+\dim_{\mathbb{K}}W'+\dim_{\mathbb{K}}W'^\perp-1=\dim_{\mathbb{K}}W'+\dim_{\mathbb{K}}W'^\perp\ge\dim_{\mathbb{K}}V\ ■$$

9.6. Corolário. *Seja V um espaço com forma e seja $W\le V$. Então $W\cap W^\perp$ é o núcleo da forma induzida sobre W. Caso W seja não-degenerado, temos $V=W\oplus W^\perp$.*

Demonstração. Um elemento $w\in W$ está no núcleo da forma sobre W se, e só se, $\langle w,W\rangle=0$. Isto é equivalente a $w\in W^\perp$, ou seja, a $w\in W\cap W^\perp$. Suponhamos que $W\cap W^\perp=0$. Então $\dim_{\mathbb{K}}(W+W^\perp)=\dim_{\mathbb{K}}W+\dim_{\mathbb{K}}W^\perp\ge\dim_{\mathbb{K}}V$ pelo Lema 9.5 ■

9.7. Lema. *Seja V um espaço com forma não-degenerado e seja $W \leq V$ um subespaço não-degenerado. Então $W^\perp$ é não-degenerado e ${W^\perp}^\perp = W$.*

Demonstração. Pelo Corolário 9.6, $V = W \oplus W^\perp$. Basta mostrar que ${W^\perp}^\perp = W$, pois isto implica que o núcleo da forma induzida sobre $W^\perp$ é nulo pelo Corolário 9.6.

Seja $w + w' \in {W^\perp}^\perp$, onde $w \in W$ e $w' \in W^\perp$. Pelo Lema 9.3, $W \subset {W^\perp}^\perp$. Logo, $w' \in {W^\perp}^\perp$, ou seja, $\langle w', W^\perp\rangle = 0$. De $w' \in W^\perp$ segue $\langle w', W\rangle = 0$. Agora, $\langle w', V\rangle = \langle w', W + W^\perp\rangle = \langle w', W\rangle + \langle w', W^\perp\rangle = 0$. Sendo V não-degenerado, $w' = 0$ ∎

Um elemento $v \in V$ é dito *isotrópico* se $\langle v, v\rangle = 0$.

9.8. Lema. *Seja V um espaço com forma não identicamente nula. Então V possui um elemento não-isotrópico.*

Demonstração. Suponhamos que $\langle v, v\rangle = 0$ para todo $v \in V$. Então, para todos $v_1, v_2 \in V$, temos $0 = \langle v_1 + v_2, v_1 + v_2\rangle = \langle v_1, v_2\rangle + \langle v_2, v_1\rangle = \langle v_1, v_2\rangle + \overline{\langle v_1, v_2\rangle}$. Se a operação $^-$ é idêntica, concluímos que a forma é identicamente nula em V. Caso contrário, temos $\mathrm{Re}\langle v_1, v_2\rangle = 0$ para todos $v_1, v_2 \in V$. Substituindo v_1 por iv_1, obtemos $\mathrm{Im}\langle v_1, v_2\rangle = 0$ para todos $v_1, v_2 \in V$ ∎

9.9. Lema. *Seja V um espaço com forma não-degenerado e seja $W \lneqq V$ um subespaço não-degenerado. Então existe um subespaço não-degenerado W' tal que $W \leq W'$ e $\dim_{\mathbb{K}} W' = \dim_{\mathbb{K}} W + 1$.*

Demonstração. Pelos Corolário 9.6, Lema 9.7 e Lema 9.8, existe um elemento não-isotrópico $w' \in W^\perp$. Façamos $W' := W + \mathbb{K}w'$. Suponhamos que $w + kw'$ está no núcleo da forma sobre W', onde $w \in W$ e $k \in \mathbb{K}$. Então $0 = \langle w + kw', W\rangle = \langle w, W\rangle$, pois $w' \in W^\perp$ implica $\langle w', W\rangle = 0$. Sendo W não-degenerado, temos $w = 0$. Agora concluímos, de $\langle w + kw', w'\rangle = 0$ e $\langle w', w'\rangle \neq 0$, que $k = 0$ ∎

9.10. Definição. Uma *bandeira* de subespaços é uma cadeia de subespaços $V_0 \leq V_1 \leq \cdots \leq V_n$ tal que $V_n = V$ e $\dim V_i = i$ para todo i. Caso V seja munido de uma forma, a bandeira é dita *não-degenerada* se todos os V_i's são não-degenerados.

Pelos Lemas 9.8 e 9.9, qualquer espaço com forma não-degenerado possui uma bandeira não-degenerada de subespaços.

9.11. Definição. Seja V um espaço com forma. Uma base linear $\beta : b_1, b_2, \ldots, b_n$ de V é dita *ortonormal*[11] se $\langle b_i, b_i\rangle \in \{-1, 0, 1\}$ e $\langle b_i, b_j\rangle = 0$ para todos i e j tais que $i \neq j$. Denotemos por $\beta_-, \beta_0, \beta_+$ as quantidades de elementos da base β tais que $\langle b_i, b_i\rangle = -1$, $\langle b_i, b_i\rangle = 0$, $\langle b_i, b_i\rangle = 1$, respectivamente. A tripla $(\beta_-, \beta_0, \beta_+)$ chama-se *assinatura* da base.

9.12. Lema. *Seja V um espaço com forma e seja $\beta : b_1, b_2, \ldots, b_n$ uma base linear ortonormal de V. Então $\beta_0 = \dim_{\mathbb{K}} V^\perp$.*

Demonstração. É imediato que todo elemento isotrópico de β pertence a $V^\perp$. Sejam $b_1, b_2, \ldots, b_m$ todos os elementos não-isotrópicos de β e seja $v = \sum_{i=1}^{m} k_i b_i \in V^\perp$, onde $k_i \in \mathbb{K}$. Então $\langle v, b_j\rangle = 0$ para todo $j = 1, 2, \ldots, m$. Logo, $0 = \Big\langle \sum_{i=1}^{m} k_i b_i, b_j \Big\rangle = \sum_{i=1}^{m} k_i \langle b_i, b_j\rangle = \pm k_j$ ∎

9.13. Ortogonalização de Gram-Schmidt. *Seja $V_0 \leq V_1 \leq \cdots \leq V_n$ uma bandeira não-degenerada de subespaços de V. Então existe $b_1, b_2, \ldots, b_n$, uma base ortonormal de V, tal que $b_1, b_2, \ldots, b_k$ é uma base de V_k para todo k.*

[11] Caso $\mathbb{K} = \mathbb{C}$ e a operação $^-$ seja idêntica, pedimos que $\langle b_i, b_i\rangle \in \{0, 1\}$, inserindo em seguida as correspondentes modificações.

Demonstração. Indução sobre n. Para $n=1$, simplesmente tomamos $0\neq c_1\in V_1$ e o normalizamos: $b_1=\dfrac{c_1}{\sqrt{|\langle c_1,c_1\rangle|}}$. (Sendo V_1 não-degenerado, $\langle c_1,c_1\rangle\neq 0$.) Suponhamos que, para $k<n$, já tenhamos encontrado uma base ortonormal $b_1,b_2,\dots,b_k$ de V_k tal que $b_1,b_2,\dots,b_i$ é uma base de V_i para todo $i\leq k$. Tomemos $c_{k+1}\in V_{k+1}\setminus V_k$ e façamos $c'_{k+1}:=c_{k+1}-\sum\limits_{i=1}^{k}\dfrac{\langle c_{k+1},b_i\rangle}{\langle b_i,b_i\rangle}b_i$. Sendo $b_1,b_2,\dots,b_k\in V_k$, temos $c'_{k+1}\in V_{k+1}\setminus V_k$. De $\dim_{\mathbb{K}}V_{k+1}=\dim_{\mathbb{K}}V_k+1$ segue que $V_{k+1}=V_k+\mathbb{K}c'_{k+1}$. Mostremos que $\langle c'_{k+1},b_j\rangle=0$ para todo $j\leq k$. Com efeito,

$$\langle c'_{k+1},b_j\rangle=\Big\langle c_{k+1}-\sum_{i=1}^{k}\frac{\langle c_{k+1},b_i\rangle}{\langle b_i,b_i\rangle}b_i,b_j\Big\rangle=\langle c_{k+1},b_j\rangle-\sum_{i=1}^{k}\frac{\langle c_{k+1},b_i\rangle}{\langle b_i,b_i\rangle}\langle b_i,b_j\rangle=$$

$$=\langle c_{k+1},b_j\rangle-\frac{\langle c_{k+1},b_j\rangle}{\langle b_j,b_j\rangle}\langle b_j,b_j\rangle=0.$$

Sendo $b_1,b_2,\dots,b_k$ geradores de V_k, temos $\langle c'_{k+1},V_k\rangle=0$. Se c'_{k+1} fosse isotrópico, ele pertenceria ao núcleo da forma sobre V_{k+1}, pois $V_{k+1}=V_k+\mathbb{K}c'_{k+1}$. Portanto, c'_{k+1} não é isotrópico e podemos normalizá-lo, obtendo o b_{k+1} desejado $\blacksquare$

9.14. Corolário. *Qualquer espaço com forma admite uma base ortonormal.*

Demonstração. Pela Observação 9.4, podemos supor que V é não-degenerado. Portanto, V possui uma bandeira não-degenerada de subespaços. Agora, o resultado segue de 9.13 $\blacksquare$

9.15. Definição. Seja V um espaço com forma e sejam $v_1,v_2,\dots,v_k\in V$. Façamos $g_{ij}:=\langle v_i,v_j\rangle$. A $(k\times k)$-matriz $G:=G(v_1,v_2,\dots,v_k):=[g_{ij}]$ se chama matriz de *Gram* de $v_1,v_2,\dots,v_k$. O fato que a forma é simétrica implica que $\overline{G}^t=G$, onde $\overline{M}$ denota a matriz M com todos os coeficientes "conjugados" por $^-$. Em palavras: G é *simétrica* (*hermitiana*).

9.16. Dicionário. Seja V um espaço com forma e seja $\beta:b_1,b_2,\dots,b_n$ uma base linear de V. Então a matriz de Gram $G^{\beta\beta}:=G(b_1,b_2,\dots,b_n)$ de β determina a forma de V, pois $\langle v,v'\rangle=[v]_\beta^tG^{\beta\beta}\overline{[v']}_\beta$ para quaisquer $v,v'\in V$. Realmente, se $v=\sum\limits_{i=1}^{n}k_ib_i$ e $v'=\sum\limits_{j=1}^{n}k'_jb_j$, então $\langle v,v'\rangle=\sum\limits_{i,j=1}^{n}k_i\langle b_i,b_j\rangle\overline{k'}_j=\sum\limits_{i,j=1}^{n}k_ig_{ij}\overline{k'}_j$, onde $g_{ij}=\langle b_i,b_j\rangle$.

Qualquer matriz simétrica (hermitiana) aparece como matriz de Gram de uma base de um certo espaço vetorial munido de uma forma apropriada. Com efeito, seja $G=[g_{ij}]\in\operatorname{Matr}_{n\times n}\mathbb{K}$ uma matriz simétrica (hermitiana), ou seja, $\overline{G}=G^t$. Seja V um $\mathbb{K}$-espaço linear com uma base linear $\beta:b_1,b_2,\dots,b_n$. Para $v_1,v_2\in V$, definimos $\langle v_1,v_2\rangle:=[v_1]_\beta^tG\overline{[v_2]}_\beta$. É imediato que $\langle v_1,v_2\rangle$ é linear em v_1. Resta observar que

$$\overline{\langle v_2,v_1\rangle}=\overline{[v_2]_\beta^tG\overline{[v_1]}_\beta}=\overline{[v_2]}_\beta^t\overline{G}[v_1]_\beta=\overline{[v_2]}_\beta^tG^t[v_1]_\beta=\big([v_1]_\beta^tG\overline{[v_2]}_\beta\big)^t=\langle v_1,v_2\rangle.$$

Seja $\gamma:c_1,c_2,\dots,c_n$ uma outra base de V e seja M_γ^β a matriz de mudança da base β para γ. Então $G^{\beta\beta}=(M_\gamma^\beta)^tG^{\gamma\gamma}\overline{M_\gamma^\beta}$. Realmente, temos $M_\gamma^\beta=[m_{ij}]$, onde $b_j=\sum\limits_{i=1}^{n}m_{ij}c_i$ para todo j. Então

$$g_{kl}=\langle b_k,b_l\rangle=\Big\langle\sum_{i=1}^{n}m_{ik}c_i,\sum_{j=1}^{n}m_{jl}c_j\Big\rangle=\sum_{i,j=1}^{n}m_{ik}\langle c_i,c_j\rangle\overline{m}_{jl}=\sum_{i,j=1}^{n}m_{ik}f_{ij}\overline{m}_{jl},$$

onde $G^{\gamma\gamma} = [f_{ij}]$.

9.17. Lema. *Seja V um espaço com forma e seja β uma base linear de V. Então V é degenerado se, e só se, $\det G^{\beta\beta} = 0$. Exceto no caso em que $\mathbb{K} = \mathbb{C}$ e a operação $^-$ é idêntica, o determinante $\det G^{\beta\beta}$ é um número real cujo sinal não depende da escolha de base.*

Demonstração. Uma base linear γ de V é ortonormal se, e só se, sua matriz de Gram $G^{\gamma\gamma}$ é diagonal com coeficientes $-1, 0, 1$. Pelo Lema 9.12, V é degenerado se, e só se, $\gamma_0 > 0$, o que é equivalente a $\det G^{\gamma\gamma} = 0$.

Pelo Corolário 9.14, podemos mudar da base original β para uma ortonormal γ. Pelo Dicionário 9.16, temos $G^{\beta\beta} = (M_\gamma^\beta)^t G^{\gamma\gamma} \overline{M_\gamma^\beta}$, onde M_γ^β é a matriz de mudança de β para γ. Pelo Teorema 8.8,

$$\det G^{\beta\beta} = \det(M_\gamma^\beta)^t \det G^{\gamma\gamma} \det \overline{M_\gamma^\beta} = \det M_\gamma^\beta \det G^{\gamma\gamma} \overline{\det M_\gamma^\beta} = \det M_\gamma^\beta \overline{\det M_\gamma^\beta} \det G^{\gamma\gamma},$$

pois a operação $^-$ comuta com as adições e multiplicações (que participam na expressão de $\det \overline{M_\gamma^\beta}$). Sabemos que $\det \overline{M_\gamma^\beta} \neq 0$. Assim obtemos a primeira afirmação. Para a segunda, basta observar que $\det M_\gamma^\beta \overline{\det M_\gamma^\beta} = |\det M_\gamma^\beta|^2 > 0$ ■

No que se segue, desconsideraremos o caso em que $\mathbb{K} = \mathbb{C}$ com a operação $^-$ idêntica, pois este é absolutamente trivial quanto aos aspectos que trataremos.

9.18. Exemplo. Seja V um espaço com forma e sejam $e, f \in V$ tais que $\langle e, e\rangle > 0 > \langle f, f\rangle$. Façamos $W := \mathbb{K}e + \mathbb{K}f$. Então $\dim_{\mathbb{K}} W = 2$ e qualquer base ortonormal em W tem a assinatura $(1, 0, 1)$. Além disso, W possui elementos não-isotrópicos (não-nulos).

Realmente, podemos supor que $W = V$. Se $0 \neq n \in V^\perp$, então $V = \mathbb{K}w + \mathbb{K}n$ para algum $w \in V$ pois $\dim_{\mathbb{K}} W \leq 2$. Caso $\langle w, w\rangle \geq 0$, concluímos que $\langle v, v\rangle \geq 0$ para todo $v \in V$. Caso $\langle w, w\rangle \leq 0$, temos $\langle v, v\rangle \leq 0$ para todo $v \in V$. Nenhum destes casos é possível pois V possui um elemento positivo e um negativo. Por motivo semelhante, $\dim_{\mathbb{K}} V = 2$. Seja β uma base ortonormal de V. É fácil ver que a assinatura de β é diferente de $(2, 0, 0)$ (pois V possui um elemento positivo) e de $(0, 0, 2)$ (pois V possui um elemento negativo). Pelo Lema 9.12, $\beta_0 = 0$. Logo, a assinatura de β só pode ser $(1, 0, 1)$. É imediato que a soma dos elementos de β é isotrópica.

9.19. Teorema da Inércia de Sylvester. *A assinatura não depende da escolha de base ortonormal.*

Demonstração. Utilizamos indução sobre $\dim_{\mathbb{K}} V$. Sejam $\beta : b_1, b_2, \dots, b_n$ e $\beta' : b'_1, b'_2, \dots, b'_n$ duas bases ortonormais de V. Pela demonstração do Lema 9.12, os elementos isotrópicos de β (mesmo como os elementos isotrópicos de β') geram o núcleo $V^\perp$. Os restantes elementos de β geram o subespaço W complementar ao núcleo, $V = V^\perp \oplus W$. Pela Observação 9.4, W é não-degenerado. Para todo não-isotrópico b'_i, temos $b'_i = n_i + w_i$, onde $n_i \in V^\perp$ e $w_i \in W$. Os elementos w_i's formam uma base linear de W. Realmente, seja $\sum k_i w_i = 0$ uma dependência linear não-trivial entre os w_i's. Então $0 \neq \sum k_i b'_i \in V^\perp$, onde os b'_i's são não-isotrópicos. Uma contradição. É fácil ver que $\langle b'_i, b'_j\rangle = \langle w_i, w_j\rangle$ para todos i, j. Logo, os w_i's formam uma base ortonormal de W cuja assinatura é $(\beta'_-, 0, \beta'_+)$. Em outras palavras, reduzimos o problema para o espaço W, ou seja, podemos supor que V é não-degenerado.

Se $\beta_- = 0$, então $\langle v, v\rangle \geq 0$ para todo $v \in V$. Com efeito, para $[v]_\beta = [k_i]$, temos $\langle v, v\rangle = \sum_{i=1}^n k_i \overline{k_i} = \sum_{i=1}^n |k_i|^2 \geq 0$ pelo Dicionário 9.16. Assim, $\beta_- = 0$ implica $\beta'_- = 0$ e, neste caso, nada há para se demonstrar. Do mesmo modo, $\beta_+ = 0$ implica $\beta'_+ = 0$. Portanto, podemos supor que $\langle b_n, b_n\rangle = 1$ e $\langle b'_n, b'_n\rangle = -1$.

Façamos

$$W := \mathbb{K}b_n + \mathbb{K}b'_n, \qquad U := (\mathbb{K}b_n)^\perp, \qquad U' := (\mathbb{K}b'_n)^\perp.$$

Claramente, $b_1, b_2, \dots, b_{n-1} \in U$ e $b'_1, b'_2, \dots, b'_{n-1} \in U'$. Pelo Corolário 9.6, $\dim_\mathbb{K} U = \dim_\mathbb{K} U' = n-1$. Logo, $U = \mathbb{K}b_1 + \mathbb{K}b_2 + \dots + \mathbb{K}b_{n-1}$ e $U' = \mathbb{K}b'_1 + \mathbb{K}b'_2 + \dots + \mathbb{K}b'_{n-1}$. Obviamente, as assinaturas das indicadas bases ortonormais de U e de U' são respectivamente $(\beta_-, 0, \beta_+ - 1)$ e $(\beta'_- - 1, 0, \beta'_+)$. Pelo Lema 9.3, $W^\perp = U \cap U'$. Pelos Exemplo 9.18 e Lema 9.12, W é não-degenerado. Pelo Lema 9.7, $U \cap U'$ é não-degenerado. Aplicando o Corolário 9.6 aos espaços U e U' e ao subespaço $U \cap U'$, obtemos as decomposições ortogonais $U = (U \cap U') \oplus (U \cap U')^{\perp U}$ e $U' = (U \cap U') \oplus (U \cap U')^{\perp U'}$ (para o ortogonal relativo, vide a Definição 9.2). Utilizando o Corolário 9.14, escolhemos uma base ortonormal γ de $U \cap U'$ e algumas bases ortonormais δ e δ' de $(U \cap U')^{\perp U}$ e de $(U \cap U')^{\perp U'}$, respectivamente. Então $\gamma \sqcup \delta$ e $\gamma \sqcup \delta'$ são bases ortonormais de U e de U', respectivamente. Calculando as assinaturas, obtemos

$$(\beta_-, 0, \beta_+ - 1) = \big((\gamma \sqcup \delta)_-, (\gamma \sqcup \delta)_0, (\gamma \sqcup \delta)_+\big) = (\gamma_-, \gamma_0, \gamma_+) + (\delta_-, \delta_0, \delta_+),$$

$$(\beta'_- - 1, 0, \beta'_+) = \big((\gamma \sqcup \delta')_-, (\gamma \sqcup \delta')_0, (\gamma \sqcup \delta')_+\big) = (\gamma_-, \gamma_0, \gamma_+) + (\delta'_-, \delta'_0, \delta'_+),$$

pois as assinaturas não dependem das escolhas de bases em U e em U' pela hipótese de indução. Resta provar que $(U \cap U')^{\perp U} = (\mathbb{K}b_n)^{\perp W}$ e que $(U \cap U')^{\perp U'} = (\mathbb{K}b'_n)^{\perp W}$, pois isto implica que $(\delta_-, \delta_0, \delta_+) = (1, 0, 0)$ e $(\delta'_-, \delta'_0, \delta'_+) = (0, 0, 1)$ devido ao Exemplo 9.18.

Sendo W e V não degenerados, $(U \cap U')^{\perp U} = (U \cap U')^\perp \cap U = W^{\perp\perp} \cap U = W \cap (\mathbb{K}b_n)^\perp = (\mathbb{K}b_n)^{\perp W}$ pelo Lema 9.7. Pelo mesmo motivo, $(U \cap U')^{\perp U'} = (\mathbb{K}b'_n)^{\perp W}$ ■

Assim, podemos falar sobre a *assinatura* (s_-, s_0, s_+) do espaço. Como medí-la? Pelo Lema 9.12, $s_0 = \dim_\mathbb{K} V^\perp$. Utilizando a Observação 9.4, podemos reduzir o problema ao caso de V não-degenerado. Seja $\gamma : c_1, c_2, \dots, c_n$ uma base de V e suponhamos que seja conhecida a matriz de Gram $G^{\gamma\gamma}$. Queremos calcular a assinatura de V em termos de $G^{\gamma\gamma}$. Fazendo $V_i := \mathbb{K}c_1 + \mathbb{K}c_2 + \dots + \mathbb{K}c_i$ para todo $0 \le i \le n$, obtemos uma bandeira de subespaços. É óbvio que a matriz de Gram da base $c_1, c_2, \dots, c_i$ de V_i é uma $(i \times i)$-submatriz $G_i^{\gamma\gamma}$, chamada *principal,* relacionada às primeiras i linhas e às primeiras i colunas da matriz $G^{\gamma\gamma} = G_n^{\gamma\gamma}$. Suponhamos que a bandeira é não-degenerada. Pelo Lema 9.17, isto é equivalente ao fato que $\det G_i^{\gamma\gamma} \ne 0$ para todo $1 \le i \le n$. Apliquemos à bandeira a Ortogonalização 9.13 e obtemos uma base ortonormal $\beta : b_1, b_2, \dots, b_n$ de V. Para todo $1 \le i \le n$, temos as bases $c_1, c_2, \dots, c_i$ e $b_1, b_2, \dots, b_i$ de V_i. Pelo Lema 9.17, os sinais de $\det G_i^{\gamma\gamma}$ e $\det G_i^{\beta\beta}$ são os mesmos. Resta observar que, para uma base ortonormal, podemos medir a assinatura como se segue:

9.20. Critério de Sylvester. *Seja V um espaço com forma e seja γ uma base linear de V. Suponhamos que $\det G_i^{\gamma\gamma} \ne 0$ para todo $1 \le i \le n$. Então a assinatura de V é igual a $(s_-, 0, s_+)$, onde s_- é a quantidade de números negativos na seqüência*

$$\det G_1^{\gamma\gamma}, \qquad \frac{\det G_2^{\gamma\gamma}}{\det G_1^{\gamma\gamma}}, \qquad \frac{\det G_3^{\gamma\gamma}}{\det G_2^{\gamma\gamma}}, \qquad \dots, \qquad \frac{\det G_n^{\gamma\gamma}}{\det G_{n-1}^{\gamma\gamma}}$$

e s_+ é a quantidade de números positivos nessa mesma seqüência ■

Um $\mathbb{R}$-espaço vetorial V munido de uma forma de assinatura $(0, 0, s_+)$ chama-se *Euclidiano.* Neste caso, a forma é referida como produto *interno.*

10. Autovalores e autovetores. Diagonalização

Seja $A : V \to V$ uma aplicação linear e seja $0 \ne v \in V$. Dizemos que v é um *autovetor* de A se $Av = kv$ para algum $k \in \mathbb{K}$. Para o autovetor v, o escalar k é único e se chama *autovalor* de A que corresponde ao autovetor v. Caso V admita uma base linear formada por autovetores de A, dizemos que A é *diagonalizável.* Em termos de matrizes, isto significa que a matriz $[A]^{\gamma}_{\gamma}$ é diagonal para alguma base linear γ de V. Quando a aplicação A é descrita através da matriz $B := [A]^{\beta}_{\beta}$ em uma base linear β de V, o fato que A é diagonalizável é equivalente, pelo Corolário 5.15, à existência de uma matriz inversível M tal que MBM^{-1} é uma matriz diagonal.

Em outras palavras, o escalar $k \in \mathbb{K}$ é um autovalor de A se, e só se, o núcleo da aplicação linear $k1_V - A$ não é nulo. Pelos Lema 5.7 e Corolário 5.10, isto é equivalente ao fato que a aplicação $k1_V - A$ não é um isomorfismo. Pela Observação 8.7, $k1_V - A$ é um isomorfismo se, e só se, $\det(k1_V - A) \neq 0$. Resumindo: para um escalar $k \in \mathbb{K}$ ser um autovalor de A, é necessário e suficiente que $\det(k1_V - A) = 0$, ou seja, em termos matriciais, que $\det\big(k1_{n\times n} - [A]^{\beta}_{\beta}\big) = 0$, onde β é uma (qualquer) base linear de V e $n := \dim_{\mathbb{K}} V$.

10.1. Definição. Seja $A : V \to V$ uma aplicação linear e seja β uma base linear de V. O polinômio $P_A(x) := \det\big(x1_{n\times n} - [A]^{\beta}_{\beta}\big)$ se chama *polinômio característico* de A, onde $n := \dim_{\mathbb{K}} V$. Pelos Observação 8.12 e Corolário 5.15, essa definição é correta, pois $Mx1_{n\times n}M^{-1} = x1_{n\times n}$ para qualquer $(n \times n)$-matriz inversível M.

10.2. Lema. *Seja $A : V \to V$ uma aplicação linear. Então $k \in \mathbb{K}$ é um autovalor de A se, e só se, k é raiz do polinômio característico de A, isto é, $P_A(k) = 0$* ∎

10.3. Proposição. *Seja $A : V \to V$ uma aplicação linear. Suponhamos que o polinômio característico $P_A(x)$ possui $n := \dim_{\mathbb{K}} V$ raízes distintas em $\mathbb{K}$. Então A é diagonalizável.*

Demonstração. Sejam $r_1, r_2, \dots, r_n \in \mathbb{K}$ raízes distintas de $P_A(x)$ e sejam $v_1, v_2, \dots, v_n \in V$ os autovetores correspondentes. Basta provar que $v_1, v_2, \dots, v_n$ são LI. Seja $\sum\limits_{i=1}^{m} k_i v_i = 0$, $1 \leq m \leq n$, uma dependência linear mínima entre os v_i's com todos os k_i's não-nulos. Por definição, autovetores são não-nulos. Logo, $m > 1$. Sendo $m > 1$ e sendo os r_i's distintos, podemos supor que $r_m \neq 0$. Aplicando A à dependência linear acima, obtemos $\sum\limits_{i=1}^{m} k_i r_i v_i = 0$. Daí, $\sum\limits_{i=1}^{m} k_i r_i r_m^{-1} v_i = 0$ e $\sum\limits_{i=1}^{m} k_i (r_i r_m^{-1} - 1) v_i = 0$. O coeficiente de v_m na última dependência linear é nulo. Pela minimalidade da dependência linear $\sum\limits_{i=1}^{m} k_i v_i = 0$, temos $k_i(r_i r_m^{-1} - 1) = 0$ para todo $1 \leq i \leq m$. Em particular, $k_1(r_1 r_m^{-1} - 1) = 0$. De $k_1 \neq 0$ segue $r_1 = r_m$. Uma contradição ∎

Substituindo $x = 0$ em $P_A(x)$, podemos ver que o termo constante de $P_A(x)$ é igual a $(-1)^n \det A$, onde $n := \dim_{\mathbb{K}} V$. É fácil ver que x^n é o termo de grau n em x no polinômio característico $P_A(x)$. Realmente, o determinante é a soma de produtos de n coeficientes da matriz $x1_{n\times n} - [A]^{\beta}_{\beta}$, um coeficiente em cada linha e um em cada coluna da matriz. Obviamente, apenas o produto dos coeficientes na diagonal principal pode contribuir ao termo de grau n em x, produzindo assim x^n. Este mesmo raciocínio possibilita calcular o termo de grau $n - 1$: ele é igual a $\big(-\operatorname{tr}[A]^{\beta}_{\beta}\big)x^{n-1}$, onde $\operatorname{tr} M$ denota o *traço* da matriz M, isto é, a soma de todos os coeficientes na diagonal principal de M. Com efeito, os termos de grau $n-1$ em x vêm somente de tais produtos de coeficientes da matriz $x1_{n\times n} - [A]^{\beta}_{\beta}$ que têm pelo menos $n - 1$ coeficientes na diagonal principal. Mas então o fator restante tem que estar também na diagonal principal para que o produto contenha um coeficiente em cada linha e um em cada coluna. Em outras palavras, o termo procurado é o termo de grau $n-1$ em x no produto $(x - a_{11})(x - a_{22})(x - a_{nn})$, isto é, $\Big(-\sum\limits_{i=1}^{n} a_{ii}\Big)x^{n-1}$, onde $[A]^{\beta}_{\beta} = [a_{ij}]$.

Na verdade, do mesmo modo como fizemos no caso do determinante, podemos definir o *traço* de uma aplicação linear $A : V \to V$ tomando $\operatorname{tr} A := \operatorname{tr}[A]^{\beta}_{\beta}$ para qualquer base linear β de V. Para provar que essa definição é correta, basta observar que $\operatorname{tr} B = \operatorname{tr}(MBM^{-1})$ para $B, M \in \operatorname{Matr}_{n\times n}\mathbb{K}$ e M inversível. Por sua vez, este fato segue da identidade $\operatorname{tr}(BC) = \operatorname{tr}(CB)$, válida para todos $B, C \in \operatorname{Matr}_{n\times n}\mathbb{K}$, pois tal identidade implica $\operatorname{tr}(MBM^{-1}) = \operatorname{tr}(M^{-1}MB) = \operatorname{tr} B$. Resta provar a identidade. Sejam $B = [b_{ij}]$ e $C = [c_{ij}]$. Então

$$\operatorname{tr}(BC) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}c_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{n} c_{ji}b_{ij} = \operatorname{tr}(CB).$$

10.4. Observação. *Seja $A : V \to V$ uma aplicação linear e seja $n := \dim_{\mathbb{K}} V$. Então $P_A(x) = x^n - (\operatorname{tr} A)x^{n-1} + \cdots + (-1)^n \det A$* ■

A partir deste momento, omitimos o sinal $\circ$ na composição de aplicações lineares escrevendo simplesmente BA no lugar de $B \circ A$. Em particular, A^i denotará a composta $\underbrace{A \circ A \circ \cdots \circ A}_{i \text{ vezes}}$ para qualquer aplicação linear do formato $A : V \to V$ para $i \geq 0$ (por convenção, $A^0 := 1_V$).

10.5. Definição. Seja $f(x) \in \mathbb{K}[x]$ um polinômio, $f(x) = k_m x^m + k_{m-1}x^{m-1} + \cdots + k_1 x + k_0$, e seja $A : V \to V$ uma aplicação linear. Definimos a aplicação linear $f(A) : V \to V$ pela fórmula $f(A) := k_m A^m + k_{m-1}A^{m-1} + \cdots + k_1 A + k_0 A^0$.

10.6. Observação. *Sejam $f_1(x), f_2(x), f(x), g_1(x), g_2(x), g(x) \in \mathbb{K}[x]$ polinômios tais que*

$$f_1(x) + f_2(x) = f(x), \qquad g_1(x)g_2(x) = g(x) \tag{10.7}$$

e seja $A : V \to V$ uma aplicação linear. Então

$$f_1(A) + f_2(A) = f(A), \qquad g_1(A)g_2(A) = g(A).$$

Demonstração. A validade das relações (10.7) é o resultado de uma manipulação formal com as expressões formais chamadas polinômios, utilizando a distributividade da adição e as comutatividade e associatividade das adição e multiplicação. Essas mesmas propriedades valem para as expressões envolvendo A no lugar da variável x pois, por exemplo, $A^iA^j = A^{i+j} = A^jA^i$ ■

De modo semelhante, podemos definir a matriz $f(B)$ para $f(x) \in \mathbb{K}[x]$ e $B \in \operatorname{Matr}_{n\times n}\mathbb{K}$.

Sejam $B, C \in \operatorname{Matr}_{n\times n}\mathbb{K}$ e seja $n = n_1 + n_2$, $n_1, n_2 > 0$. Podemos considerar as matrizes B e C como formadas por blocos, isto é, $B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$ e $C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$, onde $B_{11}, C_{11} \in \operatorname{Matr}_{n_1\times n_1}\mathbb{K}$, $B_{12}, C_{12} \in \operatorname{Matr}_{n_1\times n_2}\mathbb{K}$, $B_{21}, C_{21} \in \operatorname{Matr}_{n_2\times n_1}\mathbb{K}$ e $B_{22}, C_{22} \in \operatorname{Matr}_{n_2\times n_2}\mathbb{K}$. É fácil verificar as igualdades

$$B + C = \begin{bmatrix} B_{11}+C_{11} & B_{12}+C_{12} \\ B_{21}+C_{21} & B_{22}+C_{22} \end{bmatrix}, \qquad BC = \begin{bmatrix} B_{11}C_{11}+B_{12}C_{21} & B_{11}C_{12}+B_{12}C_{22} \\ B_{21}C_{11}+B_{22}C_{21} & B_{21}C_{12}+B_{22}C_{22} \end{bmatrix}.$$

Em palavras: as regras para operações de matrizes funcionam também para as matrizes exibidas na forma de blocos.

10.8. Observação. *Seja $B = \begin{bmatrix} B_{11} & * \\ 0 & B_{22} \end{bmatrix} \in \operatorname{Matr}_{n\times n}\mathbb{K}$ uma matriz exibida na forma de blocos e seja $f(x) \in \mathbb{K}[x]$. Então $f(B) = \begin{bmatrix} f(B_{11}) & * \\ 0 & f(B_{22}) \end{bmatrix}$* ■

Em seguida, precisaremos do seguinte

10.9. Fato. *Todo polinômio $f(x) \in \mathbb{C}[x]$ de grau > 0 admite uma raiz $r \in \mathbb{C}$, isto é, $f(r) = 0$.*

10.10. Teorema de Cayley-Hamilton. *Seja $A : V \to V$ uma aplicação linear. Então $P_A(A) = 0$.*

Demonstração. Inicialmente consideramos o caso $\mathbb{K}=\mathbb{C}$. Utilizamos indução sobre $n := \dim_{\mathbb{C}} V$. (O caso $n=1$ é trivial.) Pelo Fato 10.9, o polinômio $P_A(x)$ possui uma raiz $r\in\mathbb{C}$. Pelo Lema 10.2, existe $0\neq b_1\in V$ tal que $Ab_1=rb_1$. Pelo Corolário 4.11, podemos incluir b_1 em uma base linear $\beta : b_1, b_2, \dots, b_n$ de V. Pelo Dicionário 6.7, $[A]^\beta_\beta = \left[\begin{smallmatrix} r & * \\ 0 & B\end{smallmatrix}\right]$, onde $B\in \operatorname{Matr}_{(n-1)\times(n-1)}\mathbb{C}$. Pelo desenvolvimento de Laplace de determinante pela primeira coluna, obtemos

$$P_A(x) = \det\big(x1_{n\times n} - [A]^\beta_\beta\big) = \det\begin{bmatrix} x-r & * \\ 0 & x1_{(n-1)\times(n-1)}-B\end{bmatrix} = (x-r)P_B(x).$$

Pela hipótese de indução, $P_B(B)=0$. Pela Observação 10.6 aplicada a matrizes e pela Observação 10.8,

$$P_A\big([A]^\beta_\beta\big) = \big([A]^\beta_\beta - r1_{n\times n}\big)P_B\big([A]^\beta_\beta\big) = \begin{bmatrix} 0 & * \\ 0 & *\end{bmatrix}\begin{bmatrix} P_B(r) & * \\ 0 & P_B(B)\end{bmatrix} = \begin{bmatrix} 0 & * \\ 0 & *\end{bmatrix}\begin{bmatrix} * & * \\ 0 & 0\end{bmatrix} = 0.$$

O caso $\mathbb{K}=\mathbb{R}$ se reduz facilmente ao anterior. Com efeito, basta observar que $P_A\big([A]^\beta_\beta\big)=0$ para alguma base linear β de V, onde $P_A(x) := \det\big(x1_{n\times n} - [A]^\beta_\beta\big)$. De fato, isto já foi demonstrado: podemos interpretar a matriz $[A]^\beta_\beta$ (com coeficientes reais) como uma matriz com coeficientes complexos ■

10.11. Definição. Seja V um espaço com forma não-degenerado. Uma aplicação linear $A : V\to V$ é dita *ortogonal* (*unitária*) se A preserva a forma: $\langle Av_1, Av_2\rangle = \langle v_1, v_2\rangle$ para todos $v_1, v_2\in V$. Uma aplicação linear $A : V\to V$ é dita *autoadjunta* se $\langle Av_1, v_2\rangle = \langle v_1, Av_2\rangle$ para todos $v_1, v_2\in V$.

Toda aplicação ortogonal $A : V\to V$ é um isomorfismo. Realmente, se $Av=0$, então $0=\langle Av, Av'\rangle = \langle v, v'\rangle$ para todo v', implicando que $v\in V^\perp = 0$.

10.12. Proposição. *Seja V um $\mathbb{C}$-espaço vetorial com forma não-degenerado e seja $A : V\to V$ uma aplicação linear autoadjunta que não possui autovetores isotrópicos. Então V possui uma base ortonormal de autovetores de A.*

Demonstração. Indução sobre $n := \dim_{\mathbb{C}} V$. (Para $n=1$, o fato é trivial.) Como na demonstração do Teorema de Cayley-Hamilton 10.10, podemos encontrar um autovetor $0\neq b_1\in V$ de A, isto é, $Ab_1 = rb_1$ para algum $r\in\mathbb{C}$. Pela hipótese da Proposição 10.12, o subespaço $W := \mathbb{C}b_1$ é não-degenerado pois b_1 não é isotrópico. Seja $w'\in W^\perp$. Então $\langle Aw', b_1\rangle = \langle w', Ab_1\rangle = \langle w', rb_1\rangle = 0$. Daí, $Aw'\in W^\perp$. Assim, $AW^\perp\subset W^\perp$. Pelo Lema 9.7, $W^\perp$ é não-degenerado. Pelo Corolário 9.6, temos a decomposição ortogonal $V = W\oplus W^\perp$. A aplicação $A|_{W^\perp} : W^\perp\to W^\perp$ é autoadjunta e não possui autovetores isotrópicos. Resta aplicar a hipótese de indução ■

10.13. Dicionário. Seja V um $\mathbb{K}$-espaço com forma de assinatura $(0,0,n)$, seja β uma base ortonormal de V e seja $A : V\to V$ uma aplicação linear. Denotamos $B := [A]^\beta_\beta$.

A aplicação A é ortogonal se, e só se, a matriz B é *ortogonal* (*unitária*), isto é, se $B\overline{B}^t = \overline{B}^tB = 1_{n\times n}$, onde $n := \dim_{\mathbb{K}} V$. Realmente, pelo Dicionário 9.16, $\langle v_1, v_2\rangle = [v_1]^t_\beta\overline{[v_2]}_\beta$. Pelo Dicionário 5.11, $\langle Av_1, Av_2\rangle = \big(B[v_1]_\beta\big)^t\overline{B[v_2]}_\beta = [v_1]^t_\beta B^t\overline{B}\,\overline{[v_2]}_\beta$. Portanto, a identidade $\langle Av_1, Av_2\rangle = \langle v_1, v_2\rangle$ para todos $v_1, v_2\in V$ é equivalente a $B^t\overline{B} = 1_{n\times n}$, ou seja, a $\overline{B}^tB = 1_{n\times n}$. Para a matriz quadrada B, a igualdade $\overline{B}^tB = 1_{n\times n}$ implica $B^{-1} = \overline{B}^t$ e, portanto, $B\overline{B}^t = 1_{n\times n}$.

A aplicação A é autoadjunta se, e só se, a matriz B é *simétrica* (*hermitiana*), isto é, $B^t = \overline{B}$. Com efeito, pelos Dicionários 9.16 e 5.11, temos

$$\langle Av_1, v_2\rangle = \big(B[v_1]_\beta\big)^t\overline{[v_2]}_\beta = [v_1]^t_\beta B^t\overline{[v_2]}_\beta, \qquad \langle v_1, Av_2\rangle = [v_1]^t_\beta\overline{B[v_2]}_\beta = [v_1]^t_\beta\overline{B}\,\overline{[v_2]}_\beta.$$

Portanto, a identidade $\langle Av_1, v_2\rangle = \langle v_1, Av_2\rangle$ para todos $v_1, v_2\in V$ é equivalente a $B^t = \overline{B}$.

Em seguida, $^-$ denota a conjugação complexa.

10.14. Corolário. *Seja $S \in \operatorname{Matr}_{n\times n}\mathbb{K}$ uma matriz simétrica (hermitiana). Então, para alguma $(n \times n)$-matriz ortogonal (unitária) U, a matriz USU^{-1} é diagonal com coeficientes reais.*

Demonstração. No caso $\mathbb{K} = \mathbb{C}$, o fato segue dos Dicionário 10.13 e Proposição 10.12 pois, quando a assinatura é $(0, 0, n)$, todo elemento isotrópico é nulo e a matriz de mudança de uma base ortonormal para uma outra base ortonormal é sempre unitária. (A matriz USU^{-1} é real pois $U^{-1} = \overline{U}^t$ e $\overline{USU^{-1}} = \overline{U}\,\overline{S}U^t = (U^{-1})^t S^t U^t = (USU^{-1})^t$.)

No caso $\mathbb{K} = \mathbb{R}$, precisamos apenas de um análogo da Proposição 10.12 para o $\mathbb{R}$-espaço V com forma de assinatura $(0, 0, n)$. Neste caso, a demonstração da Proposição 10.12 funciona perfeitamente se encontrarmos um autovetor de A. Seja β uma base ortonormal de V. A matriz $[A]_\beta^\beta$ é real e simétrica, portanto, hermitiana. Pelo fato demonstrado no caso $\mathbb{K} = \mathbb{C}$, a matriz $[A]_\beta^\beta$ possui um autovalor real. Daí segue que A possui um autovalor real ∎