UMA REVISÃO DE ÁLGEBRA BÁSICA COMO INTRODUÇÃO À ALGEBRA LINEAR

MARCO BARONE

Conteúdo

1. Introdução

Esta breve apostila foi concebida (em horas que seria mais sábio dedicar ao descanso) para os estudantes do curso de álgebra linear de mestrado. Mais especificamente, é um material recomendado antes de começar o curso, por se tratar de uma revisão das noções de álgebra e da álgebra linear do bacharelado que serão utilizadas ao longo do curso e podem estar "enferrujadas". Contudo, a tratação apenas precisa das noções preliminares de teoria de conjuntos e funções, e portanto a apostila dirige-se também a qualquer aluno formado em matemática que queria revisar rapidamente conceitos algébricos espalhados ao longo de um curso de graduação e que podem ter caído no esquecimento, com um esforço relativamente exíguo e de forma concentrada e, consequentemente, cômoda. Requeremos o conhecimento das definições básicas dos fundamentos de matemática, tais como conjunto das partes, funções injetoras e sobrejetoras, imagem e pré-imagem. Na parte de lembretes de álgebra linear, também se supõe que o estudante tenha já pago um curso básico. Quando definiremos entidades, ao associarmos a um novo símbolo, até agora desprovido de significato, um significato novo, usaremos a notação ":=". O conjunto indicado como $\mathbf{N}$, segundo as nossa convenções, não incluirá o zero. Toda vez que quisermos incluir o zero, falaremos de $\mathbf{N}_0$ Resumindo,

$$\mathbf{N} := \{1, 2, 3, \ldots\}; \quad \mathbf{N}_0 := \{0, 1, 2, 3, \ldots\}.$$

Isto é mais por comodidade do que por convencimento filosófico. De fato, ao zero se estendem de forma "natural" muitos dos mecanismos que fazem sentido sobre os números naturais. Melhor dito, reconhecemos o estado de "natural" do zero como cardinal mas não como ordinal: quero falar de "conjuntos com zero elementos" mas não quero dar sentido à noção de "0-ésimo" elemento, negando assim sua existência. Quero começar a contar somente pelo 1, mas quero me deixar a possibilidade de "não começar a contar" como possibilidade concreta. Portanto saibam que, o conjunto dos "números naturais", quando assim denotados por meio de palavras, é por mim $\mathbf{N}_0$. Outra convenção, nem tão comum, mas muito útil, será denotarmos, por todo $n \in \mathbb{N}_0$, o conjunto finito dos primeiros n números naturais (entre 1 a n) por $\underline{n}$, ou seja

$$\underline{n} := \{i \in \mathbb{N} | 1 \leq i \leq n\} = \{1, 2, 3, \ldots, n\}$$

de forma que $\underline{n}$ contém n elementos, também nos casos $\underline{1} = \{1\}$ e $\underline{0} = \emptyset$

2. Conjuntos e operações

2.1. **Produtos de conjuntos.**

Definição 1. Seja I um conjunto (que chamaremos de "conjunto dos índices") e para cada $i \in I$, seja X_i um conjunto de maneira que faça sentido definir a união dos X_i e que esta seja um conjunto (por exemplo, quando os X_i são subconjuntos de um conjunto "ambiente" comum). Definamos o produto direto $\prod_{i \in I} X_i$ como o conjunto das funções $f : I \to \bigcup_{i \in I} X_i$ tais que $\forall i \in I$, $f(i) \in X_i$. No caso que $X_1 = \ldots = X_n = X$ sejam o mesmo conjunto, este produto é apenas o conjunto das funções de I a X, e será denotado por X^I:

$$X^I = \{f : I \to X\}$$

Observação 2. Observe que um elemento f de $\prod_{i \in I} X_i$ pode ser "naturalmente" [1] pensado com uma I-upla $(x_i)_{i \in I}$ de coordenadas em X, através da identificação $f(i) \leftrightarrow x_i$. Podemos pensar que para cada indice $i \in I$ existe uma posição i-ésima e

[1] Este conceito é melhor formalizado em uma área da matemática que se chama teoria de categorias. Mais ou menos, dizemos que a identificação apresentada é "natural" ou "canônica" porque se aplica a qualquer escolha dos conjuntos X_i sem que a definição da lei de associação dependa desses conjuntos

que f seja uma espécie de vetor com "tantas" coordenadas "quantos" são os índices (ou seja, os elementos de I) .

Consideraremos agora o caso em que I é um conjunto finito e nos restringiremos ao caso em que $I = \underline{n}$, lembrando a definição mais clássica de um produto cartesiano finito:

Definição 3. Sejam $X_1, \ldots X_n$ conjuntos. Então $X_1 \times \ldots \times X_n$ é definido como o conjunto das n-uplas $(x_1, \ldots x_n)$, com $x_i \in X_i$, $\forall i \in \underline{n}$. Se $X_1 = \ldots = X_n = X$ são o mesmo conjunto, denotamos $X_1 \times \ldots \times X_n = X \times \ldots \times X$ por X^n

Temos assim:

$$X^n := \{(x_1, \ldots, x_n) \,|\, x_1, \ldots, x_n \in X\}$$

Alternativamente, X^n pode ser definido, para $n \in \mathbf{N}_0$, como $X^{\underline{n}}$. Com efeito, uma n-upla de elementos de X pode ser pensada "naturalmente"[2] como uma função f de $\underline{n}$ para X, que manda todo i em x_i. Seguindo a mesma identificação, por exemplo, aprendemos que uma função f de $\mathbf{N}$ para X nada mais é do que uma sequência $a_n := f(n)$. Voltando ao produto cartesiano finito $X_1 \times \ldots \times X_n$, ele pode ser visto de modo natural como a execução repetida de várias operações de produto cartesiano entre conjuntos, $(\ldots((X_1 \times X_2) \times X_3) \times \ldots \times X_{n-1}) \times X_n$. Isto vale em qualquer ordem que se coloquem as parênteses. Por exemplo ao elemento (a, b, c, d) de X^4 associamos naturalmente $((a, (b, c)), d) \in (X \times (X \times X)) \times X$.

Mas qual é o significado disto para $n = 1$ ou $n = 0$? Diferentemente do caso $n \geq 2$, um produto cartesiano de 0 ou 1 conjuntos não se obtem como repetição de vários produtos de 2 cópias. Seja $I = \{1\}$ e X_1 um conjunto. É facil entender que $\prod_{i \in I} X_i = X_1 = X_1^1$. Seja como função, como como 1-upla, um elemento deste produto corresponde à escolha de hum (porque tal é o número de elementos de I) elemento de X_1.

No caso $n = 0$ é mais complicado entender esse conjunto. O que é (um)a 0-upla? O que é uma função de domínio vazio? Quantas são as possíveis escolhas de 0 elementos de certos conjuntos? Quando é que uma tal escolha é dada? Temos, neste caso $I = \emptyset$. Analisando a definição de função como "regra que associa a todo elemento do domínio um elemento do contradomínio", entendemos que "só **terminamos de definir** uma função **quando não houver mais nenhum elemento do domínio pelo qual a regra de associação não tenha sido definida".** Se reparamos atentamente, no caso $I = \emptyset$, uma tal eventualidade acontece sim e acontece antes mesmo de começar, uma vez que já no início, após definir a imagem de "todos os zero" elementos de I, ou seja, sem fazer nada, não resta mais nenhum elemento em I sobre o qual a função não seja definida. Portanto uma função de domínio I já está definida sem fazer nada, ou seja, fazendo zero associações. Chamamos esta de função vazia. Resta a pergunta, será que existem **outras**? Mas duas funções são distintas quando há pelo menos um elemento do domínio onde elas diferem: como não há elementos em I, tampouco pode haver dois elementos onde duas funções diferem, resultando que não existem duas funções distintas de domínio $\emptyset$, seja qual for o contradomínio, e portanto o produto vazio contém exatamente um elemento, a função vazia ou $\emptyset$-upla vazia. A confirmar esta intuição, podemos reparar que se X tem m elementos e I tem n elementos, X^I contém exatamente m^n elementos, porque existem exatamente m^n funções de I para X, correspondendo a n escolhas entre m elementos. Portanto, no caso $n = 0$, é razoável esperar que este número seja $m^0 = 1$. Note-se que

[2] O conceito de "natural", ou "canônico" que o estudante encontra desde o início do curso, começando pela disciplina de álgebra linear, e que somente vem a ser especificado em teoria de categorias significa, em palavras simples, que a correspondência é uniforme, no sentido que sua definição não menciona (e não depende de) X nem n.

este mesmo raciocínio funciona perfeitamente para $m = 0$, uma vez que nunca precisamos nos perguntar quem é o contradomínio, tornando intuitiva a igualdade $0^0 = 1$, enquanto a impossibilidade de definir uma função de contradomínio vazio só se dá **se existir** pelo menos um elemento do domínio ao qual não conseguimos associar um elemento desse contradomínio.

Outra identificação "natural" é, por exemplo, aquela que associa o elemento $(a, (b, c))$ de $X \times (X \times X)$ ao elemento (a, b, c) de X^3 e, ainda, à função $f \in X^{\underline{3}}$, tal que $f(1) = a$, $f(2) = b$, $f(3) = c$.

2.2. **Operações n-árias. Operações binárias internas.**

Definição 4. Seja X um conjunto e $n \in \mathbb{N}_0$. Uma operação n-ária interna em X, ou "sobre X" é uma função de X^n para X. Para n=1 também falamos de operação "unária", para n=2, de operação "binária", para $n = 3$, de operação "ternária", e assim por diante. Para uma operação binária interna costuma-se utilizar um símbolo colocado entre as duas coordenadas do elemento de X^2, por exemplo $\star$, falando em $a \star b$ para dizer $\star(a, b)$. Intuitivamente, a primeira escrita corresponde à noção de "maneira de combinar elementos de X **entre** si", e a segunda a uma "regra que associa a cada dupla um elemento" mas, formalmente, os conceitos se correspondem.

Exemplo 5. Uma operação 0-ária interna é uma função de X^0, o conjunto com um só elemento, em X, ou seja equivale à escolha de um elemento de X.

Exemplo 6. Uma operação 1-ária interna, ou "unária" interna, nada mais é que uma função de X em si.

Exemplo 7. Exemplos de operações 2-árias, também chamadas de "binárias" internas, são a soma e o produto em $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ ou $\mathbb{R}$. Como já observado, $a + b$ é a convenção que substitui a definição formal como função $+(a, b)$.

Exemplo 8. Considere uma operação binária interna $\star$ sobre X e defina a função $f : X^3 \to X$ como $f(a, b, c) := (a \star b) \star c = \star(\star(a, b), c)$. f é uma operção ternária interna sobre X definida por meio da repetição de $\star$. E assim é $(a, b, c) \mapsto a \star (b \star c)$. Usando repetição posso definir operações 4, 5-árias, e assim por diante. Porém nem toda operação n-ária para $n \geq 3$ é uma repetição de operações binárias, como mostram os exemplos seguintes.

Exemplo 9. $f(x, y, z) := y$ é uma operação ternária sobre $\mathbb{R}$ que, como muitas funções de $\mathbb{R}^3$ em $\mathbb{R}$, não é obtida como repetição de uma operação binária sobre X, nos sentidos apontados no parágrafo anterior. Com efeito, se fosse $y = f(x, y, z) = x \star (y \star z)$, então para todo $x, y \in \mathbb{R}$, teríamos:

$$x = x \star (x \star (y \star y)) = x \star y = x \star (y \star (y \star y)) = y,$$

que é obviamente falso.

Exemplo 10. Outro exemplo natural de operação n-ária que vem da álgebra linear é o produto vetorial. O produto vetorial em $\mathbb{R}^3$ é uma operação binária interna, mas nos é ensinado que se trata de uma peculiaridade da dimensão 3 e que, por exemplo, não podemos definir o produto vetorial em $\mathbb{R}^n$, para $n = 2$ ou $4, 5, \ldots$ A verdade é que em $\mathbb{R}^n$ tal operação existe sim, mas ela é $(n - 1)$-ária. Mais

exatamente, se $(v_1, \ldots, v_{n-1}) \in \mathbb{R}^3$:

$$v_1 \wedge \ldots \wedge v_{n-1} = f((v_1, \ldots, v_{n-1})) = \det \begin{pmatrix} \vec{e_1} & \ldots & \vec{e_n} \\ v_1^{(1)} & \ldots & v_1^{(n)} \\ v_2^{(1)} & \ldots & v_2^{(n)} \\ & \ldots & \\ v_2^{(1)} & \ldots & v_2^{(n)} \end{pmatrix}.^{3}$$

Com esta definição, por exemplo, faz sentido definir o "produto vetorial" de 3 vetores em $\mathbb{R}^4$, mas não de dois.

Exemplo 11. Outro exemplo de operação ternaria no conjunto dos pontos do plano ou do espaço euclidiano é aquela que associa a uma tripla de pontos (A, B, C) o baricentro (resp. incentro, circuncentro, ortocentro) do triângulo que os tem como vértices. É claro que tal ponto depende em medida igual, simultânea e significativa da informação dos três pontos de partida.

Operações "internas" são caracterizadas pelo fato que os elementos envolvidos estão todos em X, seja os que são combinados entre si, como o resultado. Este conceito se contrapõe a vários tipos de operação "externa". Por exemplo, em um espaço vetorial V o produto pelos escalares do corpo K, $K \times V \to V$, justamente chamado de produto externo, é operaçao externa porque um dos elementos envolvidos no produto λv é (possivelmente) externo a V (no caso, $\lambda \in K$). Outro exemplo é produto escalar $x, y \mapsto < x, y >$ de dois vetores, uma função de $V \times V$ para K, que é externa em outro sentido, a saber, porque é sim uma operação entre dois elementos de X, porém o resultado está em conjunto que não é X. As operações que mais estudaremos nesta introdução de álgebra são operações binárias internas.

2.3. **Principais propriedades das operações binárias internas.** Listamos abaixo as mais importantes propriedades das operações binárias internas e seus nomes.

Definição 12. Diremos que uma operação binária interna $\star$ sobre um conjunto X é associativa, se para todo $x, y, z \in X$, $(x \star y) \star z = x \star (y \star z)$.

Observação 13. Soma e produto entre números, matrizes, somas de vetores, composições entre funções com mesmo domínio e contradomínio são operações associativas. Já o produto vetorial $v \times w$ no espaço $\mathbb{R}^3$ ou a função $(x, y) \mapsto x^y$ não possuem tal propriedade (verifique!).

Definição 14. Diremos que uma operação binária interna $\star$ sobre um conjunto X é comutativa, se para todo $x, y \in X$, $x \star y = y \star x$.

Observação 15. Soma e produto entre números são operações comutativas. Já a composição entre funções com mesmo domínio e contradomínio não é (verifique!).

Definição 16. Seja $\star$ uma operação binária interna sobre um conjunto X. Diremos que um elemento $e \in X$ é um elemento neutro para $\star$ se, para todo $x \in X$, $e \star x = x \star e = x$. (Para operações comutativas, a igualdade precisa ser testada apenas em um sentido).

Proposição 17. *Uma operaçao binária interna $\star$ sobre um conjunto X possui no máximo um elemento neutro, que portanto poderá ser chamado, quando existir, "o" elemento neutro de $\star$.*

[3] O leitor deve reparar que esta escritura do produto vetorial como determinante contém um abuso de notação, porque a primeira linha é composta por vetores e as outras por escalares. Mais especificamente, aplicando a definição formal de determinante como soma de produtos de elementos de linhas e colunas distintas, deve ser claro que cada um destes produtos é o produto externo de um escalar, obtido como produto de $n - 1$ escalares, com um vetor.

Demonstração 18. *Sejam e, e' elementos neutros de $\star$, então $e = e \star e' = e'$, onde as duas igualdades seguem das propriedades de elemento neutro de e' à direita e de e à esquerda, respectivamente.*

Observação 19. Note que, para operações não comutativas, é necessário que a definição de elemento neutro contenha a condição de igualdade nos dois sentidos ($e \star x = x$; $x \star e = x$). Por exemplo, a composição é uma operaçao binária interna associativa sobre o conjunto $X = \{f : \mathbb{N}_0 \to \mathbb{N}_0,\ 0 \notin im(f)\}$, que possui mais de um elemento neutro à esquerda e nenhum à direita. Com efeito, se definimos e_n como a identidade sobre $\mathbb{N}$ e $e_n(0) := n$, teremos que $e_n \circ f = f$, para todo $f \in X$, porque $im(f) \in \mathbb{N}$, mas se e for um candidato elemento neutro à esquerda, pondo $e(0) = n$, é fácil verificar que para qualquer elemento f com $f(0) \neq f(n)$ teremos $f \neq f \circ e$. Da mesma forma, se $A \subsetneq \mathbb{N}$ é um conjunto com pelo menos dois elementos e $X = \{f : \mathbb{N} \to \mathbb{N},\ \text{constante em } A\}$, então a composição é interna e é fácil ver que qualquer e que mande A em um elemento de A, e coincida com a identidade fora de A, é elemento neutro à direita para a composição, enquanto se f mandar dois elementos distintos de $\mathbb{N}$ em dois elementos distintos de A, é fácil verificar que $e \circ f \neq f$, para qualquer $e \in X$.

Definição 20. Seja $\star$ uma operação binária interna sobre um conjunto X que possui elemento neutro e seja e seu único (pela observação anterior) elemento neutro. Sejam $x, y \in X$; diremos que y é um elemento simétrico de x com respeito a $\star$ se $x \star y = y \star x = e$. (Para operações comutativas, apenas uma igualdade precisa ser testada). Diremos que a operação "possui elementos simétricos" se, para todo $x \in X$ existe $y \in X$ simétrico de x.

Proposição 21. *Seja $\star$ uma operaçao binária interna sobre um conjunto X,* **associativa** *e com elemento neutro e. Seja $x \in X$. Então x possui no máximo um elemento simétrico (com respeito a $\star$), que portanto poderá ser chamado, quando existir, "o" elemento simétrico de x (com respeito a $\star$).*

Demonstração 22. *Sejam y, y' elementos simétricos de x (com respeito a $\star$).Então $y = y \star e = y \star (x \star y') = (y \star x) \star y' = e \star y' = y'$, pelas propriedades associativa, as de elemento neutro de e, e as propriedades de y e y' de simétricos de x, à esquerda e à direita, respectivamente.*

Observação 23. Note que, assim como a de elemento neutro, a condição de elemento simétrico, para operações não comutativas, precisa da igualdade nos dois sentidos. Por exemplo, se X é o conjunto das funções de $\mathbb{N}_0$ em $\mathbb{N}_0$, a composição é uma operaçao binária interna associativa com elemento neutro dado pela função idêntica. Podemos notar que uma função injetora não sobrejetora, como $x \mapsto x + 1$ possui como inversas esquerdas todas as funções que valem $x - 1$ sobre $\mathbb{N}$ (seja qual for a imagem de zero), mas não possui inversa direita. Da mesma forma, qualquer função sobrejetora que seja constante sobre um subconjunto A com mais de um elemento (exemplo: $f(2k) := k$, $f(2k+1) := 0$) tem como inversas direitas todas as funções g com $g(n) \in f^{-1}(n)$, para todo n natural[4] (exiba uma para o exemplo mencionado!), mas não possui inversa esquerda.

Observação 24. Note que y é simétrico de x se e somente se x é simétrico de y e portanto diremos simplesmente que x e y são simétricos (um do outro).

Existem também propriedades que envolvem mais de uma operação. Vamos definir apenas a propriedade distributiva, por ser a única que usaremos:

[4]O leitor "cabuloso" observe que tais funções existem, sem precisarmos assumir o axioma da escolha, porque nosso exemplo $\mathbb{N}$ é bem ordenado: por exemplo, tome $g(n) := min\{f^{-1}(n)\}$

Definição 25. Seja X um conjunto com duas operações binárias internas, $\star$ e $\triangle$. Diremos que $\triangle$ é **distributiva com respeito a** $\star$ se para todos $a, b, c \in X$, temos que $a\triangle(b \star c) = (a\triangle b) \star (a\triangle c)$ e $(a \star b)\triangle c = (a\triangle c) \star (b\triangle c)$

Observação 26. Na definição anterior não podemos deixar de impor a igualdade em ambos os sentidos. Com efeito, se X for o conjunto das funções de $\mathbb{R}$ para $\mathbb{R}$, considerando a soma de funções no lugar de $\star$ e a composição $\circ$ no lugar de $\triangle$, podemos observar que, para todas $f, g, h \in X$, vale $(f + g) \circ h = f \circ h + g \circ h$ pela própria definição de soma de funções mas $f \circ (g + h) = f \circ g + f \circ h$ só vale, em geral, se f for linear. Este é um exemplo de uma dupla de operações, uma sendo distributiva com respeito à outra só pela esquerda. No outro sentido, podemos citar o exemplo de $X = \mathbb{R}^+$, $x \star y := x \cdot y$ e $x\triangle y := x^y$, obtendo o exemplo de uma operação distributiva com respeito à outra somente pela direita (o leitor pode verificar facilmente isto).

2.4. **Notação aditiva e multiplicativa e operações repetidas.** No parágrafo anterior, temos utilizado o símbolo $\star$ para denotar a genérica operação binária interna. Sabemos que muitos dos exemplos de operações conhecidas e trabalhadas na aritmética são chamadas de soma ou produto, e não somente entre números (ex. soma de funções, some de vetores, produto de matrices, etc.) Para operações associativas, a escolha dos símbolos "+" e "·" para indicar a operação está associada também a formas convencionais de denotar o elemento neutro, os elementos simétricos e a operação repetida entre cópias do mesmo elemento. Tais convenções costumam ser denotadas "notação aditiva" e "notação multiplicativa". Operações binárias internas definem estruturas algébricas: quando se trabalha em estruturas ou em contextos abstratos[5] onde apenas uma operação aparece, uma operação comutativa costuma ser chamada de "soma" e indicada pelo símbolo "+" (por exemplo na teoria dos grupos abelianos), enquanto o símbolo de produto, "·" costuma ser atribuído a uma operação que não é necessariamente comutativa, como por exemplo no contexto de teoria de grupos (não necessariamente abelianos). Na teoria dos anéis, onde se trabalha com duas operações na mesma estrutura, utilizaremos os dois símbolos, sendo o de soma reservado àquela operação de cada dupla que é necessariamente comutativa. Quando se trabalha com operações de várias estruturas envolvidas ao mesmo tempo podemos utilizar, ao invés dos de soma e produto, símbolos abstratos para as operações (ex. $\star$, $\triangle$), ou modificações do mesmo símbolo (ex. +, +', +" etc.) ou o nome do conjunto pode ser colocado como subíndice abaixo do símbolo da operação (ex. $+_G$, $+_H$). Operações em um contexto específico e não abstrato já possuem um nome, e também podem ser chamadas de soma (neste caso, são sempre comutativas) ou produto (podendo ou não ser comutativos, ex. produto de números reais vs produto de matrizes quadradas). Neste caso as notações aditiva e multiplicativa também seguem o nome da operação. Em conclusão, nem sempre existe uma regra específica de quando uma operação deve ser chamada de soma, produto ou com outro símbolo, mas há tendências, que foram descritas até agora. Porém, as associações entre o nome da operação e o nome de seus elementos neutros e (se associativas) simétricos, quando existem, costumam ser coerentes com os padrões descritos abaixo:

Definição 27. Seja X um conjunto com uma operação binária interna **comutativa**. Diremos que utilizamos **notação aditiva** para tal operação, se ela é chamada de "soma", indicada pelo símbolo "+", se seu elemento neutro (se ele existir) também é chamado de "zero da operação", e se denota mediante o símbolo "0" ou "0_X". Se + é associativa e possui elementos simétricos, o simétrico de um dado elemento $x \in X$, que já vimos ser único, será chamado de "oposto de x" e denotado por $-x$.

[5]No sentido que as operações são variáveis, não especificadas

Definição 28. Seja X um conjunto com uma operação binária interna. Diremos que utilizamos **notação de produto** para tal operação, se ela é indicada pelo símbolo "$\cdot$" (sem aspas, obviamente!), por vezes omitido entre elementos (xy no lugar de $x \cdot y$); neste caso o elemento neutro, se existe, também é chamado de "um" ou "identidade" da operação, e se denota mediante o símbolo "1" ou "1_X". Se $\cdot$ é associativa e possui elemento neutro, em notação de produto, o simétrico de um dado elemento $x \in X$, que já vimos ser único, será chamado de "inverso de x" e denotado por x^{-1}. Mais em geral, para uma operação associativa, com elemento neutro e simétricos, diremos que usamos **notação multiplicativa**, se os simétricos estão na supracitada notação de inversos (x^{-1}). Pode acontecer, em casos específicos, que a operação tenha um símbolo próprio diferente do de produto (ex. composição de funções, denotada por $\circ$) e que a identidade tenha um nome diferente de 1 (no caso, id_X), mas a inversa de uma função mantenha a notação multiplicativa f^{-1}; ou que apenas a identidade guarde seu nome próprio, diferente de 1 (ex, para produtos de matrizes $n \times n$, o símbolo 1_n ou $\mathbb{I}_n$). A título de curiosidade, outro exemplo, da topologia algébrica, é o produto de classes de laços em um grupo fundamental, $\alpha * \beta$, que mantém o nome da identidade $[c_{x_0}]$ mas usa notação multiplicativa para o inverso α^{-1}. Outro exemplo, da geometria algébrica (que não precisam entender), é o produto de feixes invertíveis, que usa o símbolo de produto tensorial $\mathcal{F} \otimes_{\mathcal{O}_X} \mathcal{G}$, mantém o nome da identidade $\mathcal{O}_X$ e notação multiplicativa para o inverso, $\mathcal{F}^{-1}$. Vamos introduzir agora o conceito de "operação repetida" e descobrir como as notações aditiva e multiplicativa se reconhecem também na operação repetida entre cópias de um mesmo elemento.

Se uma operação $\star$ é associativa e possui elemento neutro, se $n \geq 2$ e se $x_1, \ldots x_n \in X$, podemos definir $x_1 \star \ldots \star x_n$ como $(\ldots((x_1 \star x_2) \star x_3) \star \ldots \star x_{n-1}) \star x_n$. Assim como no caso dos produtos cartesianos, usando o fato que $\star$ é associativa, podemos provar que tal resultado não depende do modo de colocarmos as parênteses, ou seja, da ordem escolhida para executar as operações, desde que a ordem dos elementos seja mantida. Se $\star$ for comutativa, o resultado tampouco dependerá da ordem. De tal modo, por exemplo, $x_1 \star x_2 \star x_3 \star x_4 := (((x_1 \star x_2) \star x_3) \star x_4 = (x_1 \star x_2) \star (x_3 \star x_4) = x_1 \star ((x_2 \star x_3) \star x_4)$. Se $I = i_1, \ldots, i_n$ é um conjunto finito de índices, com pelo menos dois elementos, a convenção anterior define uma função $X^I \to X$, ou seja, no caso $I = \underline{n}$, uma operação n-ária interna sobre X: se $\{x_i\}_{i \in I}$ é uma n-upla de elementos de X (que pode ser pensada como uma função de I para X), ela será naturalmente ordenada como $(x_{i_1}, \ldots, x_{i_n})$, segundo a ordem de I, ou diretamente $(x_1, \ldots, x_n)$ se $I = \underline{n}$, ou utilizando a função composta com a função de $\underline{n}$ em I que fornece a ordem sobre I, ou seja $i \mapsto x_i$. Portanto podemos associar a esta I-upla o elemento $x_{i_1} \star \ldots \star x_{i_n}$ e definir a operação n-ária repetida originada da $\star$.

Definição 29. Seja X um conjunto com uma operação comutativa e associativa, em notação aditiva (melhor dito, um conjunto com "uma soma") e seja $n \geq 2$. A operação n-ária repetida definida anteriormente pode ser indicada, por brevidade, com o símbolo de **somatória**. Também podemos escrever diretamente:

$$\sum_{i \in I} x_i = x_{i_1} + \ldots + x_{i_n},$$

no caso que $I = \{i_1, \ldots, i_n\}$ é um conjunto finito, a função $f : I \to X$ representa uma I-upla (uma n-upla) de elementos de X, com $f(i) = x_i$. Podemos escrever:

$$\sum_{i=1}^{n} x_i = x_1 + \ldots + x_n$$

se $I = \underline{n}$, ou no caso que $I = \{i_1, \ldots, i_n\}$ seja um conjunto finito, $j : \underline{n} \to I$, $f : I \to X$ e $f \circ j(i) = x_i$ Em notação de produto (às vezes qualquer notação multiplicativa),

$$\prod_{i \in I} x_i \quad ou \quad \prod_{i=1}^{n} x_i$$

Sempre no caso em que "+" é associativa, para $n \geq 2$ e $x \in X$, podemos definir a operação repetida entre cópias de x como:

- $nx = \underbrace{x + \ldots + x}_{n \text{ cópias}}$ em notação aditiva
- $x^n = \underbrace{x \cdot \ldots \cdot x}_{n \text{ cópias}}$ ou $\underbrace{x \star \ldots \star x}_{n \text{ cópias}}$ em notação multiplicativa.

A primeira construção é chamada "múltiplo inteiro", a segunda "potência" e valem, para $m, n \geq 2$:

$$(m + n)x = mx + nx; \quad x^{m+n} = x^m \cdot x^n$$

Para $n = 0, 1$ o modo de definir a operação repetida é visto por alguns como uma obviedade, por outros como uma convenção: com efeito, se trata da única definição natural que mantém válidas as propriedades acima. De fato com 0 ou 1 cópias de um elemento a somar/multiplicar, nenhuma soma/produto "entre" duas delas pode ser realizado. Se $n = 1$ a soma/produto de 1 só cópia de x é definida como o próprio x, enquanto, se $n = 0$ a soma/produto de uma família vazia de cópias é definida como o elemento neutro da operação. Para melhor imaginar tal cenário podemos reescrever as operações como:

- $nx = 0 + 0 + \underbrace{x + \ldots + x}_{n \, vezes}$ ou
- $x^n = 1 \cdot 1 \cdot \underbrace{x \cdot \ldots \cdot x}_{n \, vezes}$ ou $e \star e \star \underbrace{x \star \ldots \star x}_{n \, vezes}$,

de modo que a definição fica invariada para $n \geq 2$ e mesmo para $n = 0$ ou 1 a expressão pode ser vista como resultado da operação repetida e executada realmente pelo menos alguma vez.

3. Estruturas Algébricas

Uma estrutura algébrica é um conjunto com certas operações, geralmente binárias, internas ou externas (com referência a outra estrutura pré-definida, neste caso), e certas propiedades satisfeitas por tais operações que, para as operações binárias internas, geralmente aparecem entre as propriedades descritas nos parágrafos anteriores. Segundo o número de operações e propriedades requisitadas, as estruturas se classificam em tipos, em padrões de estruturas, tais como grupos, anéis, espaços vetoriais, módulos, álgebras, e assim por diante. Resumiremos abaixo os tipos principais de estruturas que vão aparecer no texto.

Definição 30. Um grupo $(G, \star)$ é um conjunto G com uma operação binária interna $\star$ associativa, que possui elemento neutro e_G e simétricos de todos os elementos (únicos, como foi provado em Proposição 21 para toda operação associativa com elemento neutro). Há quem se refira a G e quem a $(G, \star)$ como sendo "o grupo". Se $\star$ é comutativa diremos que G é um grupo abeliano e usaremos notação aditiva, se não se usa notação multiplicativa e, salvo possivelmente casos de grupos concretos, notação de produto.

Definição 31. Um anel $(R, +, \cdot)$ é um conjunto com duas operações binárias internas, uma em notação aditiva (soma), que é comutativa, associativa, possui elemento

neutro 0_R e simétricos de todos os elementos, e outra em notação de produto, associativa, com elemento neutro 1_R e distributiva com respeito à primeira. Se o produto também é comutativo, o anel se dirá comutativo. Um elemento $x \in R$ se diz "divisor do zero" se $\exists y \in R, y \neq 0_R$, com $xy = 0$. Um elemento que não é divisor do zero é dito "regular". Um elemento de um anel se diz "invertível" ou "inversível" se possui inverso (ou seja, simétrico pelo produto).

Observação 32. Existe uma notação específica, mais clássica, para se referir a estruturas do tipo descrito acima como "anéis unitários", enquanto um "anel" não conteria necessariamente uma identidade do produto, ou seja 1_R. Contudo, por muito boas razões que entenderemos melhor ao definirmos subestruturas, eu vou chamar tais "anéis unitários" simplesmente de "anéis" e desconsiderar completamente anéis sem unidade. Alguns se referem a R, outros a $(R, +, \cdot)$ como sendo "o anel". Podemos adotar a filosofia que o anel é "o conjunto, onde estão definidas as tais operações", ou a tripla que contém como informações ao mesmo tempo conjunto e operações, ou ainda a 5-upla $(R, +, \cdot, 0_R, 1_R)$. Importante é entender que a identificação de um anel, (como também a igualdade entre anéis), depende da identificação de toda essas informações.

Proposição 33. *Seja R um anel, seja 0_R o elemento neutro da soma de R e 1_R o elemento neutro do produto. Temos que:*

a) $0_R \cdot x = x \cdot 0_R = 0_R$, *para todo* $x \in R$.
b) *Todo elemento invertível é regular.*
c) $R = \{0_R\}$ *se e somente se* $0_R = 1_R$ *(tal anel se diz "anel nulo"), portanto, a menos deste caso,* 0_R *nunca pode ser regular.*

Definição 34. Seja R um anel não nulo. Observe que aqui 0 é sempre um divisor de zero. Diremos que R é um domínio de integridade ou, simplesmente, domínio, se seu único divisor de zero é 0_R. Diremos que R é um corpo se todos os elementos diferentes de 0 são invertíveis. Quando um anel é um corpo, ele costuma ser denotado com a letra k ou K, ou F, ao invés que com a inicial R (do inglês *ring*="anel") e a seguir suas letras sucessivas.

Observação 35. Pela proposição anterior, todo corpo é um domínio. Observe bem que excluímos por definição o anel nulo da possibilidade de ser um corpo ou um domínio.

Definição 36. Seja $(R, +, \cdot)$ um anel. Um R-módulo é um conjunto M, com uma soma (operação binária interna em notação aditiva) tal que $(M, +)$ seja um grupo abeliano (ou seja a soma seja associativa, comutativa, tenha elemento neutro e simétricos de todo elemento) e uma operação externa, chamada produto por escalares:

$$R \times M \to M$$
$$(\lambda, v) \mapsto \lambda v$$

com as propriedades:

- $\forall \lambda, \mu \in R, v \in M, (\lambda + \mu)\, v = \lambda v + \mu v$
- $\forall \lambda \in R, v, w \in M, \lambda(v + w) = \lambda v + \lambda w$
- $\forall \lambda, \mu \in R, v \in M, (\lambda \cdot \mu)\, v = \lambda(\mu v)$
- $\forall v \in M, 1_R v = v$

Se $K = R$ é um corpo, um K-módulo V também é dito K-espaço vetorial, como o leitor deve ter notado ao observar as propriedades das operações, nas quais deixei de propósito a notação dos elementos, λ, v, etc., parecida à da álgebra linear, para maior familiaridade. Para espaços vetoriais, a última propriedade não precisa ser imposta, pois ela pode ser deduzida pelas outras. Em geral, na

teoria de módulos se utiliza mais frequentemente a notação $(r, m) \mapsto rm$, para o produto externo. Podemos observar que todo anel R é um R-módulo, definindo como produto externo o produto interno de R. Contudo devemos tomar cuidado com a diferente definição de suas subestruturas como anel ou como R-módulo. Todo grupo abeliano nada mais é que um $\mathbb{Z}$-módulo, cujo produto externo rm, $r \in \mathbb{Z}, m \in M$ é trivialmente definido mediante a função de múltiplo inteiro: a soma de r cópias de m, se r é não negativo, a de $-r$ cópias de $-m$ senão.

Definição 37. Se R é um anel comutativo, uma R-álgebra (associativa unitária) é um R-módulo M que também é um anel, com a compatibilidade dos produtos interno e externo e dos seus elementos neutros:

$$\forall r \in R, m, n \in M, (rm) \cdot_M n = r(m \cdot_M n) = m \cdot_M (rn).$$

R-álgebras costumam conter uma cópia do próprio R como subanel mediante a função $r \mapsto r1_M$, quando esta for injetora.

3.1. **Subestruturas.** Dada uma estrutura algébrica sobre um conjunto X, uma subestrutura de X é um subconjunto Y de X que é fechado com respeito às operações de X (ou seja, tal que as operações de X, restritas a Y, dão resultados em Y) e que forma uma estrutura do mesmo tipo com as mesmas operações (restritas), mesmos elementos neutros e, por consequência, elementos simétricos.

Exemplo 38. Um subgrupo de um grupo $(G, \star)$ é um subconjunto $H \subseteq G$ que forma um grupo com a operação restrita $\star|_{H\times H}$. Que o elemento neutro de H seja o mesmo de G, neste caso, segue automaticamente da existência do simétrico:

$$e_H = e_G \star e_H = ((e_H)^{-1} \star e_H) \star e_H = (e_H)^{-1} \star (e_H \star e_H) = (e_H)^{-1} \star e_H = e_G$$

onde $(e_H)^{-1}$, em notação multiplicativa, denota o simétrico de e_H em G e observando, na execução do produto $(e_H \star e_H)$ que a operação $\star$ de G, entre elementos de H, é a operação de H, e tem e_H como elemento neutro. Alternativamente, basta assumir que H seja um subconjunto fechado com respeito à operação "diferença" ou "quociente" (a depender da notação, ou seja a operação $\star$ executada entre o primeiro elemento e o simétrico do segundo) de G.

Exemplo 39. Um subanel de um anel $(R, +, \cdot)$ é um subconjunto $S \subseteq R$ que forma um anel com as operações restritas a $S \times S$ e com os mesmos elementos neutros de R. Enquanto a igualdade $0_R = 0_S$ pode ser deduzida pela existência do simétrico, $1_R = 1_S$ deve ser assumido: é por esta razão que escolhi evitar a distinção entre "anel" e "anel unitário": com a definição de anel (apenas) deveríamos dar uma definição geral de subanel (apenas) que portanto não pode, em geral, envolver unidades. Assim $\mathbb{Z}\times\{0\}$ (anel unitário), seria uma subestrutura, como anel simples, de $\mathbb{Z}\times\mathbb{Z}$ (que também é anel unitário), porém não seria sua subestrutura como anel unitário, porque sua unidade, $(1, 0)$ é diferente da unidade do anel maior, $(1, 1)$. Em outras palavras teríamos um subanel, unitário, de um anel também unitário, que não é um seu sub-(anel unitário), complicando imensamente as notações. Se R e S são corpos e S é um subanel de R, diremos que S é um "subcorpo" de R. Normalmente, quando anéis são corpos, ao invés da letra R (e suas sucessivas) os livros costumam usar a letra k, ou K, do alemão *körper*="corpo", e suas sucessivas, ou a letra F, do inglês *field*="corpo"(apenas em matemática).

Exemplo 40. Dado um anel R fixado, um R-submódulo N de um R-módulo M é um subconjunto de M que forma um subgrupo abeliano de M com sua soma e que é fechado com respeito ao produto externo $R \times M \to M$, em outras palavras a imagem de um elemento de $R \times N$ está em N, de modo que a restrição define um produto externo $R \times N \to N$ Da mesma forma, trocando R por um corpo k, se V é um k-espaço vetorial, um k-subespaço vetorial W de V é um seu subgrupo

abeliano tal que a restrição do produto por escalares a $k \times W$ tem imagem contida em W, definindo um produto por escalares sobre W.

Exemplo 41. Da mesma forma, uma subalgebra B de uma R-álgebra (resp. k-álgebra) A dada é um subanel de A que tambem é um seu R-submódulo (resp. k-subespaço vetorial) de modo que as operações sejam compatíveis, tornando B uma R-álgebra (resp. k-álgebra).

Definição 42. Dado um anel comutativo $(R, +, \cdot)$ daremos também a definição de "ideal". Se trata de um tipo especial de subconjunto de um anel, com respeito à estrutura, no sentido que ele "absorve em si todo R pelo produto", mas não se trata propriamente de uma sub-estrutura (ou seja, no caso, não é um subanel)[6] Um ideal de um anel comutativo R é definido como um subconjunto I de R que forma, com a soma de R, um subgrupo aditivo de $(R, +)$ e tal que, para todo $r \in R$ e $x \in I$, $rx \in I$. Em termos de estruturas algébricas, um ideal de R é nada mais que um R-submódulo do R-módulo R. Observe que um ideal sempre contém 0 (devendo ser um subgrupo aditivo) e que $\{0\}$ e R são sempre ideais de R. Se R não é comutativo a condição mencionada nesta definição para I apenas define um "ideal esquerdo", enquanto para obter um "ideal"ou "ideal bilateral", é preciso que para todo $r \in R$ e $x \in I$, também $rx \in I$. Não precisamos nos deter muito sobre ideais de anéis não comutativos.

Proposição 43. *Seja R um anel comutativo e I um seu ideal. Se I contém uma unidade, então $I = R$. Consequentemente, se $R = k$ é um corpo, seus únicos ideais são $\{0\}$ e k*

Demonstração. Por definição $I \subseteq R$. Viceversa, se $r \in R$, seja a a unidade contida, por hipótese, em I e seja a^{-1} seu inverso multiplicativo, então $r = (ra^{-1})a$ é produto de um elemento de R, ra^{-1} e de um elemento de I, a, e pela propriedade de ideal deverá pertencer a I, provando também $R \subseteq I$, deonde $R = I$. Se $R = k$ é um corpo, um seu ideal I deve conter o 0 e portanto é igual a $\{0\}$ ou contém outro elemento, a saber, uma unidade, sendo dessa forma, pelo que acabamos de ver, todo k. □

O leitor pode verificar imediatamente o fato seguinte:

Proposição 44. *Seja R um anel. A interseção de uma família arbitrária de ideais de R é um ideal de R.*

Definição 45. Se $X \subseteq R$ é um subconjunto de R, posso definir o "ideal gerado por X", denotado por (X), como a interseção de todos os ideais que contêm X. Pela proposição anterior, se trata de um ideal.

Observação 46. Com as notações anteriores, é fácil verificar que (X) é o menor ideal que contém X, ou seja que (X) é um ideal, contém X, e qualquer outro ideal J que contiver X também deverá conter (X). (X) coincide com o conjunto $\{r_1x_1 + \ldots r_nx_n | n \in \mathbb{N}, r_i \in R, x_i \in X\}$, cujos elementos se chamam as "combinações lineares finitas de elementos de X a coeficientes em R".

Demonstração. A proposição anterior garante que (X) é um ideal. Se $\mathcal{F}$ denotar a família dos ideais de R que contêm I, (X), sendo a interseção de conjuntos que contêm X, também devera conter X. Se J é um ideal de R que contém X, $J \in \mathcal{F}$ e portanto $(X) \subseteq J$, uma vez que é interseção de uma família da qual J participa. Para a segunda parte do enunciado, é fácil verificar que $Y = \{r_1x_1 + \ldots r_nx_n | n \in \mathbb{N}, r_i \in R, x_i \in X\}$ é um ideal, pois o 0 é combinação linear finita (uma soma de zero termos), uma soma de duas combinações finitas é uma combinação finita e, se

[6] Para os livros que tratam de anéis não unitários a definição de subanel é diferente e portanto o ideal acaba sendo um subanel. Isto nunca vale para nós, a não ser no caso $I = R$, já que nossos subanéis devem conter 1_R (veja próxima proposição)

$r \in R$, $y = r_1x_1 + \ldots r_nx_n \in Y$ com os $x_i \in X$, temos que $ry = (rr_1)x_1 + \ldots (rr_n)x_n$ também é combinação linear dos mesmos elementos de X. Além de ser ideal, Y contém todo $x \in X$, sendo $n = 1$, $x_1 := x$ e $x = 1 \cdot x_1 \in Y$. Portanto $X \subseteq Y$ e $(X) \subseteq Y$ pela afirmação anterior. Mas todo elemento de F, por ser ideal e por conter todos os elementos de X, deve necessariamente conter todas suas combinações lineares, ou seja todo Y. Portanto (X) é interseção de elementos que contêm Y e por isto $Y \subseteq (X)$ também. □

Definição 47. Se $x_1, \ldots, x_n \in R$, o ideal $(\{x_1, \ldots, x_n\})$ também é denotado por $(x_1, \ldots, x_n)$ (sem as chaves) e se diz simplesmente "ideal gerado por $x_1, \ldots, x_n$ (ou seja, poderemos dizer que é gerado "pelos elementos" além de "pelo conjunto deles"). Um ideal gerado por um número finito de elementos se diz finitamente gerado. Observamos que todas as definições e considerações sobre subestruturas geradas e conjuntos (ou elementos) geradores que demos até agora no caso dos ideais se aplicam também aos submódulos de um módulo (e portanto espaços vetoriais, grupos abelianos). Anéis cujos ideais são sempre finitamente gerados se dizem **Noetherianos**. Um ideal gerado por um só elemento $x \in X$, ou seja $(x) := (\{x\})$ é chamado "principal". Se trata do conjunto dos "múltiplos de x em R". Domínios cujos ideais são principais são frequentemente abreviados no apelativo de **PIDs** (do inglês *principal ideal domains*=("domínios a ideais principais"). Se trata de uma classe muito importante de anéis.

3.2. **Morfismos ou funções preservando a estrutura.** Ao falarmos em estruturas algébricas, sempre consideraremos funções entre estruturas algébricas, com a propriedade que "preservam" a estrutura (as operaçoes, seus elementos neutros e, consequentemente, seus simétricos). Tais funções são chamadas geralmente de "morfismos", "homomorfismos" ou possuem um nome que depende do tipo de estrutura considerada (para espaços vetoriais, por exemplo, se chamam transformações lineares). Se $f : X \to Y$ é uma função e temos o mesmo tipo de estrutura algébrica sobre X e Y (podendo ser ambos grupos, anéis, R-módulos, k-espaços vetorias, R-álgebras, k-álgebras), diremos que a função preserva a estrutura quando f manda os elementos neutros de cada operação interna de X nos elementos neutros da respectiva operação interna de Y e quando f manda o resultado da operação de X executada entre dois elementos no resultado da operação de Y executada entre suas imagens. Ainda, em presença de uma operação externa, ao trasformarmos os elementos de X envolvidos na operação em elementos de Y mediante f, antes ou depois de executar a operação, devemos obter o mesmo resultado. Estas funções especiais têm normalmente a propriedade que uma sub-estrutura do mesmo tipo é dada também pela imagem de f e, exceto no caso dos anéis, pela pré-imagem do elemento neutro da (primeira) operação, também chamado de "núcleo". Mas vamos dando a definição caso por caso.

Definição 48. Se $(G, \star)$ e $(H, \triangle)$ são grupos, com elementos neutros e_G e e_H, respectivamente, um **homomorfismos de grupos** de G em H é uma função $f : G \to H$ tal que $f(e_G) = e_H$ e

$$\forall g_1, g_2 \in G, f(g_1 \star g_2) = f(g_1) \triangle f(g_2).$$

Dois grupos se dizem **isomorfos** se existe um homomorfismo bijetor entre eles (chamado um "isomorfismo de grupos", e cuja inversa, verifica-se, é sempre um homomorfismo).

Observação 49. É fácil demonstrar, usando a unicidade do simétrico, que a imagem do simétrico de um elemento, através de um homomorfismo de grupos f, é o simétrico da sua imagem (exercício). É fácil mostrar que são subgrupos, de H e de G, respectivamente, a imagem de f, imf e o núcleo de f, definido como

$Ker\ f := \{g \in G | f(g) = e_H\}$ (do alemão "Kern" ou do inglês "kernel"="caroço"). O leitor pode demonstrar que um homomorfismo de grupos f é uma função injetora (pré-imagens dos unitários são no máximo unitários) se e somente se $Ker\ f = \{e_G\}$ (ou seja, apenas requerendo que a pré-imagem do unitário $\{e_H\}$ seja unitaria). É fácil ver também que a propriedade se estende à operação repetida.

Definição 50. Se $(R, +_R, \cdot_R)$ e $(S, +_S, \cdot_S)$ são anéis, com elementos neutros $0_R, 1_R, 0_S$ e 1_S, respectivamente, um **homomorfismo de anéis** de R em S é uma função $f : R \to S$ tal que $f(1_R) = 1_S, \forall r_1, r_2 \in G, f(r_1 +_R r_2) = f(r_1) +_S f(r_2)$ e $\forall r_1, r_2 \in G, f(r_1 \cdot_R r_2) = f(r_1) \cdot_S f(r_2)$. Dois anéis se dizem isomorfos se existe um homomorfismo bijetor entre eles (chamado um "isomorfismo de anéis" e cuja inversa, verifica-se, é sempre um homomorfismo).

Observação 51. É fácil demonstrar, usando a propriedade distributiva e o item (a) da Proposição 33, que se $f : R \to S$ é um homomorfismo de anéis, então $f(0_R) = 0_S$ e, usando a unicidade do simétrico, que a imagem do simétrico de um elemento é o simétrico da sua imagem, relativamente a qualquer uma das duas operações (no caso do produto, quando o simétrico existir). Normalmente entre anéis, ao denotar somas e produtos, omitiremos doravante a indicação de se as operações são realizadas em R ou em S mediante subíndices, dado que isto se deduz sabendo a qual anel pertencem os elementos entre os quais se realiza cada operação. Neste caso, $im f$ é um subanel de S mas o núcleo de f, definido como $Ker\ f := \{r \in R | f(r) = 0_S\}$ é um ideal de R (exercício, segue da definição) e não, em geral, um seu subanel. É fácil ver também que a propriedade se estende às operações repetidas (somas e produtos finitos arbitrários).

Observação 52. O conjunto $\mathbb{Z}$ dos números inteiros forma um anel comutativo onde todo ideal é principal, gerado pelo seu menor elemento positivo (isto se vê por contradição usando o algoritmos de divisão de Euclides). É fácil ver que para todo R anel comutativo, existe um único homomorfismo de anéis comutativos $j_R : \mathbb{Z} \to R$ que manda $1_\mathbb{Z}$ em 1_R e se estende da única maneira possível como $j_R(1 + \ldots + 1) = j_R(1) + \ldots + j_R(1) = 1_R + \ldots + 1_R$ e $j_R((-1) + \ldots + (-1)) = j_R(-1) + \ldots + j_R(-1) = (-(1_R)) + \ldots + (-(1_R))$.

Definição 53. O núcleo do homomorfismo j_R definido acima é um ideal de $\mathbb{Z}$ e, portanto, é principal. O seu gerador não negativo é chamado "característica de R". Se este número for zero, j_R é injetor, e R contém uma cópia de $\mathbb{Z}$. A maioria dos anéis mais conhecidos, por exemplo $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}[i]$, $\mathbb{R}[x]$, possuem característica zero e contêm uma cópia de $\mathbb{Z}$. É impossível, neles, que a soma de um número finito de cópias de 1_R dê 0_R como resultado. Contudo, isto pode acontecer em outros anéis: por exemplo, o anel $\mathbb{Z}_m$ das classes de congruência (ou "de resto") módulo um inteiro m tem característica m ou seja é nula a soma $1_R + \ldots + 1_R$(m cópias), também denotada, na notação de mútliplo inteiro que introduzimos, de $m1_R$. É fácil provar que se R for um corpo, sua característica só pode ser zero ou um primo.

Observação 54. Observe que se S é subanel de R, então eles têm a mesma característica, pois a inclusão de S em R, composta com j_S, deve dar j_R. Observe também que todo homomorfismo de anéis entre dois corpos $f : k \to K$ é injetor (devendo mandar 1_k em 1_K, $Ker\ f$, que é um ideal, não pode ser todo k, e portanto deve ser $\{0\}$). Também verifica-se facilmente que qualquer corpo k de característica zero contém uma cópia isomorfa de $\mathbb{Q}$[7] (que se chama corpo de base de k) e se k tem característica p, então k contém uma cópia isomorfa de $\mathbb{Z}_p$ como subcorpo (que também se chama corpo de base de k).

[7]Isto é uma gíria dos algébristas para dizer "contém um corpo isomorfo a $\mathbb{Q}$".

Na álgebra linear estamos interessados em resultados válidos para espaços vetoriais sobre corpos de característica zero (em particular, sobre subcorpos de $\mathbb{C}$) mas muitas propriedades podem ser provadas indistintamente para espaços vetoriais sobre um corpo qualquer. Apenas a característica 2 fornece algum problema (o operador que permuta dois vetores, por exemplo, não é diagonalizável), mas em geral, se pode trabalhar em um contexto abstrato. Porém, para alguns tópicos, como a teoria de Jordan, se trabalha sobre uma classe mais restrita de corpos, os corpos algebricamente fechados, ou seja aqueles corpos k tais que todo polinômio não constante com coeficientes em k possui uma raiz em k. Os corpos algebricamente fechados são infinitos e o caso mais conhecido, em característica zero, é o corpo complexo $\mathbb{C}$, como afirma um notório resultado, o teorema fundamental da álgebra, cuja demonstração será omitida por complicada e pouco significativa.

3.3. **Um exemplo de grupos: os grupos de permutações.** Daremos abaixo um exemplo de grupo não abeliano que precisamos conhecer para poder dominar as definições e propriedades da teoria de determinantes, em particular a própria definição de determinante, e algumas matrizes especiais que se comportam como permutações. Seja n um número natural. Observo que a composição é uma operação interna ao conjunto $\underline{n}^{\underline{n}}$ das funções de $\underline{n}$ em si. Observo também que a composição de funções biunívocas é biunívoca. Portanto a composição é uma operação binária interna sobre $S_n := \{f \in \underline{n}^{\underline{n}} \mid f \text{ é bijetora}\}$.

Definição 55. Seja X um conjunto. Uma permutação é uma função bijetora de X em X. O conjunto das permutações de denota por $S(X)$. Se $X = \underline{n}$, também chamaremos $S(\underline{n})$ de S_n. Um elemento de S_n se chama "permutação de n elementos". Como observado, a composição é interna a $S(X)$.

Observação 56. $S(X)$ é um grupo, com a composição. Com efeito, vimos que a composição é interna, a identidade de X, bijetora, é o elemento neutro da composição e a simétrica de uma função bijetora de X em X com respeito à composição é sua função inversa, que é bijetora, de X em X. Se $X = \underline{n}$ podemos contar seus elementos, que serão $n! = n$(escolhas da imagem de 1) vezes (n-1)(escolhas da imagem de 2, excluindo a imagem de 1 já escolhida) vezes . . . até a última escolha, obrigada, da imagem de n.

Um dos símbolos mais frequentemente usados para indicar uma permutação é a letra grega σ. Existem três maneiras de identificar uma permutação $\sigma : \underline{n} \to \underline{n}$. A primeira consiste em elencar as imagens de todo elemento de $\underline{n}$, com a indicação explícita de quem é mandado em quem, como em:

$$\begin{array}{rccc} \sigma: & \underline{6} & \to & \underline{6} \\ & & & \\ & 1 & \mapsto & 4 \\ & 2 & \mapsto & 6 \\ & 3 & \mapsto & 1 \\ & 4 & \mapsto & 3 \\ & 5 & \mapsto & 5 \\ & 6 & \mapsto & 2 \end{array}$$

A segunda, mais prática, se expressa mediante uma tabela de duas linhas, a primeira listando os elementos de $\underline{n}$, na ordem, e onde na segunda aparece a imagem de cada elemento abaixo dele, como:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 6 & 1 & 3 & 5 & 2 \end{pmatrix}$$

Para a terceira notação, a notação **como produto de ciclos disjuntos**, precisamos lembrar algumas definições:

Definição 57. Um r-ciclo em S_n é uma permutação γ de S_n que troca circularmente os r elementos de uma sequência (subconjunto ordenado) de elementos distintos de $\underline{n}$, a saber, na ordem (de troca circular) $i_1, \ldots, i_r$ e deixa invariados os demais. Mais explicitamente $\gamma(i_1) = i_2, \ldots, \gamma(i_{r-1}) = i_r$ e, fechando o círculo, $\gamma(i_r) = i_1$. Além disso, se $i \notin \{i_1, \ldots, i_r\}$, então $\gamma(i) = i$. Escreveremos que $\gamma = (i_1 \ldots i_r)$. A permutação definida no exemplo anterior não é um ciclo. Já a permutação

$$\sigma' = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 2 & 1 & 3 & 5 & 6 \end{pmatrix}$$

é um 3-ciclo, com $r = 3$, $\sigma' = (i_1\, i_2\, i_3) = (1\,4\,3)$. Observe que tal escritura não é única, sendo também $\sigma' = (i_2\, i_3\, i_i) = (i_3\, i_1\, i_2)$ mas vira única se adotarmos a convenção que i_1 seja o menor dos elementos mexidos pelo ciclo.

Observação 58. Observe que duas permutações cujos respectivos conjuntos "mexidos" (que não ficam invariados) são disjuntos, necessariamente comutam: é como se cada uma se ocupasse de permutar um conjunto separado e exclusivo dela, sem interferir na atividade da outra. É um risultado clássico da teoria das permutações ,que toda permutação pode se escrever como produto (ou seja, composição) de cíclos disjuntos, onde a palavra "disjuntos" se refere aos conjuntos "mexidos", ao fato que as sequências $i_1, \ldots, i_r$ relativas aos vários ciclos são disjuntas. Por exemplo, a permutação σ mencionada anteriormente pode ser escrita como a composição do ciclo (1 4 3) com o ciclo (2 6): $\sigma = (1\,4\,3)(2\,6) = (2\,6)(1\,4\,3)$ (a composição é aqui omitida). Os conjuntos {1,3,4}, {2,6} e {5} são as três **órbitas** de σ, onde a palavra "órbita" provém da linguagem da teoria das ações de grupo (para quem tiver conhecimento, lembramos que S_n age sobre $\underline{n}$), sendo uma órbita mediante σ uma classe da relação de equivalência sobre $\underline{n}$ "ser obtido um do outro mediante uma potência, ou repetição, de σ". O método para obter a escritura única de σ como produto de ciclos disjuntos consiste em começar pelo primeiro número mexido por σ e calcular sua imagem, e a imagem desta, até fechar um círculo, correspondendo a um ciclo, que será escrito na fatoração, para depois recomeçar com o menor número dos que não foram tangidos e que é mexido, continuando até alcançar a fatoração desejada. Os números não mexidos, que terão portanto uma órbita unitária, não são mencionados em tal fatoração, ou podem alternativamente (e desnecessariamente) ser mencionados mediante 1-ciclos, todos correspondendo à identidade de σ. Tente rastrear a aplicação do método ao exemplo anterior até obter a escritura (1 4 3)(2 6) citada e, se tiver tempo, a outra permutação definida de modo casual. Na verdade, não há unicidade como produto de ciclos disjuntos, sendo $(1\,4\,3)(2\,6) = (6\,2)(3\,1\,4)$, por exemplo, mas se fixamos a convenção que o menor número de cada órbita deve estar em primeiro lugar e que os primeiros números das várias órbitas devem estar em ordem crescente, obteremos unicidade da escritura.

Observação 59. Um 2-ciclo se diz "transposição". Se é verdade que toda permutação se escreve em modo canônico como produto de ciclos disjuntos de forma única, é também verdade que toda permutação é produto de transposições (não disjuntas), porém tal maneira está longe de ser única. Por exemplo a permutação do exemplo anterior é

$$\sigma = (1\,3)(1\,4)(2\,6) = (1\,3)(1\,4)(2\,6)(2\,6)(2\,6) = (1\,5)(3\,6)(2\,4)(4\,6)(1\,2)(3\,5)(2\,5),$$

podendo se escrever como produto de um número distinto de transposições, este número podendo ser, pelo menos, 3, 5 ou 7. Se as primeiras duas escritas diferem

pela repetição de duas cópias consecutivas da mesma transposição (que sempre podem ser acrescentadas a um produto, sem alterá-lo, já que sua composição é a identidade), a terceira escrita mostra que σ é obtida por um caminho completamente diferente, mas curiosamente o número de permutações involvidas possui a mesma paridade dos anteriores, a saber, todos esses números são ímpar. O próximo teorema mostrará que isto não é um caso e nos ajudará a definir o sinal de uma permutação, indispensável na definição do determinante:

Proposição 60. *Suponhamos que $\sigma \in S_n$ pode ser obtida como produto de m transposições e, de outra maneira, como produto de n transposições. Então m é par se e somente se n é par, e neste caso σ dir-se-á uma "permutação par". Em caso contrário σ se dirá uma "permutação ímpar". Em outras palavras se σ é produto de um número par de transposições, ela não pode ser produto de um número ímpar de transposições, e viceversa.*

Demonstração. Defina $N(\sigma)$ como o número de duplas $(i, j) \in \underline{n} \times \underline{n}$ tais que $i < j$, porém $\sigma(i) > \sigma(j)$. Quero provar que se σ' se obtem multiplicando σ por uma permutação $\tau = (a\, b)$, (ponhamos $a < b$), então $N(\sigma)$ e $N(\sigma')$ possuem paridade distinta. Com efeito, repare bem que na contagem das duplas (i, j): as que mudarão entre $N(\sigma)$ e $N(\sigma')$ são exatamente as duplas (i, a) e as (i, b) (nesta ou outra ordem), para todos os $i \in \underline{n}$ tais que $\sigma(i)$ se encontre entre $\sigma(a)$ e $\sigma(b)$, ou seja dois conjuntos do mesmo número de duplas, mais a dupla (a, b), que fica solta e desempata a contagem, chegando a dar sempre um número ímpar de mudanças entre $N(\sigma)$ e $N(\sigma')$. O leitor pode verificar esta contagem nos exemplos anteriores, onde σ e σ' apenas diferem por uma multiplicação por (2 6). Como consequência disto, se σ pode se escrever como um produto de m transposições $\tau_1, \ldots \tau_m$, ela se obtém a partir da identidade através de uma sequência finita de m permutações $\sigma_i = \tau_1 \ldots \tau_i$ diferindo cada uma da seguinte pela multiplicação por uma transposição. De consequência, para cada i, $N(\sigma_i)$ tem paridade diferente de $N(\sigma_{i+1})$. Por indução provamos assim $(-1)^{N(\sigma_i)} = (-1)^i$, já que $(-1)^{N(\sigma_0)} = (-1)^{N(id_{\underline{n}})} = (-1)^0$, e $(-1)^{N(\sigma_{i+1})} = (-1)^{N(\sigma_i)} = (-1) \cdot (-1)^i$, concluindo que a paridade de m é exatamente a paridade de $N(\sigma)$: $(-1)^{N(\sigma)} = (-1)^{N(\sigma_m)} = (-1)^m$. □

Definição 61. Podemos assim definir o sinal de uma permutaçao σ como $|\sigma|$ ou $\mathrm{sig}(\sigma) = (-1)^N(\sigma)$, onde $N(\sigma) = \#\{(i, j) \in \underline{n} \times \underline{n} \mid i < j\,,\ \sigma(i) > \sigma(j)\}$. e observar que este sinal coincide com $(-1)^m$ sendo m quaquer inteiro tal que σ pode ser escrito como produto de m transposições.

Outro resultado importante diz respeito às permutações conjugadas: a permutação conjugada de σ mediante υ é a permutação $\sigma^\upsilon := \upsilon\sigma\upsilon^{-1}$. Se sigma é produto de k ciclos disjuntos de tamanhos $r_1, \ldots r_k$, respectivamente, o vetor $(r_1, \ldots, r_k)$ se diz estrutura de σ. Neste caso, uma simples observação garante que a estrutura de σ é igual à estrutura de σ^υ, para qualquer permutação υ. Mais especificamente, usando a escritura como produto de ciclos disjuntos, se $\sigma = (i_1^1 \ldots i_{r_1}^1) \cdot \ldots \cdot (i_1^k \ldots i_{r_k}^k)$, então $\sigma^\upsilon = (\upsilon(i_1^1) \ldots \upsilon(i_{r_1}^1)) \cdot \ldots \cdot (\upsilon(i_1^k) \ldots \upsilon(i_{r_k}^k))$. O leitor pode se convencer disto observando que, sendo υ bijetora, para definir a imagen dos elementos de $\underline{n}$ através de σ^υ basta definir a imagem das imagens de υ apenas, e calculando estas.

3.4. **Uma classe de álgebras: os polinômios com coeficientes em um anel.** Um polinômio em uma variável com coeficientes em um anel comutativo R pode ser definido como uma escritura formal $a_0x^0 + \ldots + a_nx^n$, onde x é um símbolo, $a_i \in R$, n natural arbitrário, dito grau se $a_n \neq 0$, ou de maneira (no meu ver mais prática!), como uma soma infinita $p(x) = p_0x^0 + \ldots + p_nx^n + \ldots$, onde x é sempre um símbolo e impomos que os p_i elementos de R sejam todos nulos

menos um número finito. Falaremos indistintamente de p ou de $p(x)$[8]. Se os p_i são todos nulos teremos o polinômio nulo, ou zero. Diferentemente, o maior índice n tal que $p_n \neq 0$ se dirá grau de p e se denotará por $\deg(p)$ (do inglês *degree*="grau") e p_n será dito "coeficiente líder" de p. Um polinômio com coeficiente líder igual a 1_R se diz "mônico". O grau do polinômio nulo será definido como $-\infty$ coerentemente com o entendimento que se $\deg(p) = \sup\{n \mid a_n \neq 0\}$, então faz certo sentido dizer que "$\sup \emptyset = -\infty$" e tal definição resultará muito útil na generalização das igualdades e disegualdades sobre graus. A segunda definição de polinômio nos permite de definir de forma mais prática soma e produto de polinômios, porque não devemos nos preocupar muito com o conjunto indicizador das somatórias, e tornamos desta maneira o conjunto dos polinômios em uma variável com coeficientes em R, denominado $R[x]$, um anel comutativo.

Definição 62. Se $p(x) = \sum p_k x^k$ e $q(x) = \sum q_k x^k$, então defino a soma de polinômios $p + q$ como $(p + q)(x) = \sum(r_k)x^k$ onde $r_k = p_k + q_k$. A expressão $p + q$ é evidentemente um polinômio, pois todos os r_k, menos um número finito, serão nulos. Da mesma forma deifno o produto pq como $pq(x) = \sum(s_k)x^k$ onde $s_k = \sum_{i+j=k} p_i q_j$. Novamente observo que os s_i não nulos não podem exceder o índice $\deg p + \deg q$ e portanto pq resulta bem definido. Com um pouco de paciência, se prova que $R[x]$ é um anel comutativo, com o polinômio nulo como elemento neutro da soma e o polinômio $1 = 1x^0 + 0x^1 + \ldots + 0x^n + \ldots$ como elemento neutro do produto. Ademais, identificando R com o subanel dos polinômios constantes (os polinômios de grau zero mais o polinômio nulo), R resulta ser também um R-módulo (um espaço vetorial se R for um corpo) e, finalmente, uma R-álgebra.

Observação 63. Atenção: não devemos confundir polinômios com funções polinomiais: todo polinômio em $R[x]$ produz uma função de R em R, substituindo $a \in R$ na variável e interpretando as operações formais de $p(x)$ como operações dentro de R aplicadas à expressão $p(a)$, onde a soma e o produto de R podem ser executados. Todavia, dois polinômios podem dar origem a funções iguais sem serem polinômios iguais. Por exemplo, se $R = \mathbb{Z}_3$, vemos que o polinômio $x^3 - x + 2$ e o polinômio constante 2 coincidem como funções de R em R mas o primeiro tem grau 3 e o segundo 0, portanto, eles diferem como elementos de R[x], ou seja, como objetos algébricos. Para que dois polinômios p e q sejam iguais (como elementos de R[x]) é necessário que $p_i = q_i$ para todo i natural.

É fácil verificar a seguinte

Proposição 64. *Se R é um domínio, então $R[x]$ também é. Se k é um corpo, não podemos exigir que $k[x]$ seja um corpo (os únicos polinômios invertíveis são as constantes não nulas) mas em compensação será um tipo de domínio com propriedades muito importantes das quais falaremos em breve: um domínio euclidiano.*

Lembremos também as seguintes propriedades sobre o grau:

Proposição 65. *Se R é um anel e $p, q \in R[x]$, temos que* $\deg(p+q) \leq max\{\deg p, \deg q\}$ *e* $\deg(pq) \leq \deg p + \deg q$. *Ademais, se R for um domínio, teremos que* $\deg(pq) = \deg p + \deg q$ *e o coeficiente líder de pq é o produto dos coeficientes líderes de p e q.*

Polinômios f, g em uma variável com coeficientes sobre um corpo podem ser divididos um com o outro. O procedimento é idêntico aos polinômios reais: se $d = \deg f < \deg g = e$, em cada passo da divisão de g por f pode-se diminuir o grau produzindo, no divisor, um monômio do grau adequado com coeficiente

[8]Toda vez que a expressão de sua estrutura interna, em termos da variável e dos coeficientes, for explicitada, somente o chamaremos de p(x)

igual à divisão dos coeficientes líderes, ou seja, $\frac{g_e}{f_d}x^{e-d}$, até obter um resto de grau menor de deg f. Aqui $\frac{g_e}{f_d}$ significa $g_e {f_d}^{-1}$, que faz sentido porque estamos em um corpo e $f_d \neq 0$ e portanto é invertível. Mais exatamente:

Proposição 66. *Seja k um corpo e $f, g \in k[x]$, $g \neq 0$, então existem $q, r \in k[x]$, com* $\deg r < \deg g$ *e* $f = qg + r$

Demonstração. Se $d = \deg f < \deg g = e$, tome $q = 0$ e $r = f$. Diferentemente, temos $d = \deg f \geq \deg g = e$. Se $d = 0$, também $e = 0$ e podemos tomar $q = f/g$ e $r = 0$, senão como na observação anterior, basta definir $q_1(x) := \frac{g_e}{f_d}x^{e-d}$ e observe, como na discussão anterior, que o grau de $r_1 := f - q_1 g$ é estritamente menor de d. Então, procedendo por indução sobre d, se deg r_1 continua sendo maior ou igual a e, temos $r_1 = q_2 g + r_2$, deg $r_2 < e$, chegando a $f = (q_1 + q_2)g + r_2$, com as propriedades desejadas. □

Lembremos que um domínio Euclidiano é um domínio onde vale o algoritmo de divisão de Euclides. Mais especificamente, um anel S se diz Euclidiano se é um domínio e existe uma "função-grau" $\delta : S\backslash\{0\} \to \mathbb{N}_0$ tal que se $f \in S$, $g \in S\backslash\{0\}$, $\exists q, r \in S$, tais que $f = qg + r$ e $r = 0$ ou $\delta(r) < \delta(g)$. Provamos que se k é um corpo, então $k[x]$ é Euclidiano, tomando δ como a função deg. Temos também os seguintes fatos da álgebra: domínios euclidianos são domínios a ideais principais e domínios a ideais principais são domínios de fatoração única. A primeira afirmação sobre um ideal I de um domínio euclidiano se prova observando que qualquer elemento não nulo g de I que atinja o valor mínimo de δ sobre I precisa gerar I: ao supor o contrário, tente deduzir uma contradição! O fato que domínios euclidianos são domínios a fatoração única não é imediato, dada a definição de domínio euclidiano que adotamos (a mais geral). Contudo, é bem mais simples provar a fatoração única para anéis de polinômios com coeficientes em um corpo, que é a única coisa que precisaremos no curso de álgebra linear. Os seguintes fatos são relevantes:

Proposição 67. *Se $p \in R[x]$ e $a \in R$, $p(a) = 0$ se e somente se p é múltiplo do polinômio $x - a$ em $R[x]$. Neste caso se dirá que a é uma raiz de p.*

Demonstração. Obviamente, se $p(x) = q(x)(x - a)$, teremos que $p(a) = 0$. Viceversa $p(x) = p(x) - p(a) + p(a) = \sum(p_i(x^i - a^i)) + p(a)$ e todos os termos do primeiro membro são múltiplos por (x-a), e por isto, se $p(a) = 0$, p é múltiplo de $(x - a)$. □

Note que a interpretação dos polinômios como funções e a possibilidade de avaliar eles em um valor de R, para todo $a \in R$ fornece um homomorfismo de anéis $ev_a : R[x] \to R$, chamado avaliação em a, dado por $ev_a(f) := f(a)$. Pela proposição anterior, seu núcleo é o ideal gerado pelo polinômio $(x - a)$

Definição 68. Em um anel comutativo qualquer R, para $a, b \in R$, faz sentido dizer que "a divide b", denotado $a|b$, quando existe $c \in R$, $b = ac$. Unidades dividem qualquer elemento e qualquer elemento divide o zero. Tomem cuidado para não confundir "dividir o zero" no sentido formal (a | 0) com "ser um divisor de zero" na definição convencional que foi dada na teoria de anéis (ou seja, ser um elemento não regular). Em um domínio, se $a|b$ e $b|a$, então a e b se dizem "associados": eles se obtêm um do outro pela multiplicação por uma unidade. Todo polinômio em $k[x]$ é associado a um único polinomio mônico. Em um domínio a ideais principais podemos definir o que é um máximo divisor comum de um conjunto finito de elementos $f_1, \ldots f_n$: como todo ideal é principal, o ideal $(f_1, \ldots f_n)$ das combinações lineares dos $f_1, \ldots f_n$ será igual ao ideal (d) gerado por um só elemento d, que será dito "um" máximo comum divisor de $f_1, \ldots f_n$. É crucial a observação que, desta forma, d é ele mesmo uma combinação linear dos f_i com coeficientes no anel. Note

que um tal d precisa dividir todo f_i e que se outro d' divide todos eles, então d' precisa dividir toda sua combinação linear, portanto também d, concluindo que d é máximo divisor comum no sentido clássico, com respeito à ordem parcial dada pela divisibilidade. Podemos provar que máximos divisores comuns distintos do mesmo conjunto de polinômios devem ser associados. Em $k[x]$, adotando a convenção que d seja mônico, posso chegar a uma definição convencional de "o máximo divisor comum de um conjunto de polinômios". Em geral, um máximo divisor comum existe sempre nos domínios a ideais principais e é combinação linear dos elementos; se 1 é um máximo comum divisor de elementos $a_1, \ldots, a_n$ escreveremos simplesmente $(a_1, \ldots, a_n) = 1$.

Definição 69. Se $f = \sum f_i x^i \in k[x]$ e $a \in k$, a multiplicidade de a como raiz de f, $\mu_f(a)$ é o maior natural n tal que $(x-a)^n$ divide f. Obviamente, a é uma raiz de f se e somente se $\mu_f(a) \geq 1$. Se $\mu_f(a) = 1$, a se diz raiz simples de f, se $\mu_f(a) > 1$, a se diz raiz múltipla de f e neste caso a será raiz também da "derivada formal de f", definida como o polinômio $f'(x) := \sum i f_i x^{i-1}$.

Definição 70. Um elemento a de um domínio R se diz irredutível se a não é invertível e se, toda vez que podemos escrever $a = bc$ em R, temos que ou b ou c deve ser invertível (e/ou, analogamente, ou c ou b deve ser associado a a).

Proposição 71. *Em um domínio a ideais principais, se $a|bc$ e $(a,b) = 1$, então $a|c$. Em particular, se um elemento irredutível divide um produto de n fatores, então ele divide um dos fatores.*

Demonstração. Escrevendo $1 = sa + tb$, é fácil ver que $c = sac + t(bc)$ é um múltiplo de a. Se a é irredutível, a condição $(a,b) = 1$ é equivalente a "a não divide b", fornecendo a segunda parte do enunciado para $n = 2$. O caso geral segue por indução. □

A observação seguinte vale em um domínio euclidiano qualquer, após provar que a função δ pode ser escolhida de modo a preservar a ordem parcial dada pela divisibilidade (que é um pouco laborioso). Provamos ela apenas em $k[x]$, que é o que serve para nosso propósito.

Proposição 72. *Todo polinômio $f \in k[x]$ não constante é divisível por algum polinômio irredutível p_1, ou seja $f = p_1 \hat{f}$.*

Demonstração. Seja $n = \deg f$. O leitor pode observar que, como todas as constantes não nulas são invertíveis em um corpo, polinômios de grau 1 são necessariamente irredutíveis. Observe que f não é invertível, pois $n > 0$ por hipótese. Se f é irredutível, terminamos, pegando $p_1 := f$, senão posso escrever f como produto de dois polinômios g e h, nenhum dos dois sendo invertível (e portanto nenhum dos dois sendo constante). Observo que $\deg g < n$ e repito o raciocínio com g, sabendo que precisa parar quando o grau alcança 1. Ou, simplesmente, uso uma hipótese de indução sobre g. □

É usando repetidamente o método anterior que posso transformar $f \in k[x]$ em um produto finito de polinômios irredutíveis, e provar a existência de uma fatoração irredutível, trabalhando indutivamente sobre n e usando a fatoração de $\hat{f}$. Para unicidade da fatoração, a menos de troca de fatores, devemos exigir que os fatores irredutíveis sejam mônicos e admitir a multiplicação por uma constante, que corresponde necessariamente ao coeficiente líder f_n de f. Se $f = f_n p_1^{\alpha_1} \cdot \ldots \cdot p_r^{\alpha_r} = f_n q_1^{\beta_1} \cdot \ldots \cdot q_s^{\beta_s}$, com os p_i e q_j irredutíveis mônicos, posso usar a segunda afirmação de Proposição 71 aplicada a qualquer um dos p_i para provar que ele deve dividir um dos q_j e portanto ser igual a este, pois ambos são irredutíveis mônicos. Assim,

os conjuntos $\{p_1, \ldots, p_r\}$ e $\{q_1, \ldots, q_s\}$ coincidem e portanto $r = s$. Ademais, tendo $p_i = q_j$, se fosse $\alpha_i \neq \beta_j$, por exemplo $\alpha_i < \beta_j$, ao dividirmos as duas expressões para f por $p_i^{\alpha_i}$ eu teria igualdade entre uma expressão não divisível por p_i e uma divisível por p_i, uma contradição.

Provamos a fatoração única dos elementos de $k[x]$ em polinômios irredutíveis. Todavia, para corpos esepcíficos, existe uma estrutura especial para a fatoração, devido à estrutura especial dos elementos irredutíveis. Já vimos que polinômios de grau 1 são irredutíveis. Se $k = \mathbb{R}$, é sabido que polinômios irredutíveis têm grau 1 ou grau 2, com discriminante negativo. Se $k = \mathbb{C}$, pelo contrário, os únicos polinômios irredutíveis têm grau 1. Esta é outra maneira de enunciar o teorema fundamental da álgebra. Um Equivalentemente, todo polinômio complexo tem uma raiz complexa. Equivalentemente, todo polinômio complexo tem uma expressão

$$f(x) = c \prod_{i=1}^{k} (x - a_i)^{\alpha_i}$$

onde c é o coeficiente líder de f e a_i são suas raízes, com multiplicidade α_i. Antes de encerrar a subsessão, torna-se necessário explicitar uma consequência deste fato, que nos tornará útil no desenvolvimento da teoria de Jordan:

Proposição 73. *Se $f = c \prod_{i=1}^{k} (x - a_i)^{\alpha_i}$ é um polinômio sobre um corpo algebricamente fechado, então os polinômios $\hat{f}_i = f/(x - a_i)^{\alpha_i}$ são coprimos e portanto o polinômio 1, seu máximo divisor comum, é combinação linear deles.*

4. Lembretes sobre matrizes e determinantes

Apesar de termos colocado um corte preciso, a próxima seção (que investiga objetos da álgebra linear) se abre com o uso de ferramentas relacionadas com os resultados da última subseçao, e estabelece um elo entre a teoria de anéis de polinômios e a teoria de matrizes. Matrizes formam k-álgebras e merecem ser tratadas com ferramentas algébricas gerais e sofisticadas, já que estão disponíveis.

4.1. Polinômios mínimos de matrizes.

Vamos introduzir o conceito de polinômio mínimo de uma matriz sobre seu corpo de coeficientes. Este polinômio está relacionado com o polinômio característico, no sentido que divide ele e coincide com ele quando a matriz é diagonalizável, mas estes resultados requerem uma pequena digressão na álgebra abstrata.

Proposição 74. *Se $S_1 \subseteq S_2$ é uma inclusão de anéis comutativos e J é um ideal de S_2, então $J \cap S_1$ é um ideal de S_1. Note que a afirmação é válida também se S_2 não é comutativo e J é um ideal esquerdo.*

Demonstração. A interseção de conjuntos fechados pela soma é fechado pela soma também. Se $r \in S_1$ e $x \in S_1 \cap J$, rx pertence a S_1 pela sua propriedade de subanel e a J, pela sua propriedade de ideal. □

Definição 75. Seja $R_1 \subseteq R_2$ uma inclusão de anéis comutativos. Se $S_i = R_i[x]$, para $i = 1, 2$, teremos uma inclusão $S_1 \subseteq S_2$. Seja a um elemento R_2. Todo elemento de S_1 pertence a S_2 e como tal, pode ser avaliado em a e o núcleo da restrição a S_1 do homomorfismo de avaliação em a, $ev_a|_{S_1} : S_1 \to R_2$ é exatamente a interseção de S_1 com o núcleo de ev_a e, portanto, um ideal de S_1. Se $R_1 = k$ é um corpo, S_1 tem ideais principais e posso encontrar um gerador do ideal mencionado. Se tal ideal não for nulo, o elemento a será dito algébrico sobre R_1 e o gerador mônico $m_k(a)$ (ou $m_{k,a}$) do ideal mencionado, será chamado "polinômio mínimo de a sobre k".

Vamos agora considerar um exemplo de álgebra, que provém da álgebra linear. Uma matriz $m \times n$, a coeficientes em um anel comutativo R, pode ser expressada como uma tabela retangular de elementos de R,

$$(a_{i,j}) = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}.$$

Mais formalmente, trata-se de uma função do produto $\underline{m} \times \underline{n}$ a R que manda o par (i, j) em $a_{i,j}$. O conjunto $M_{m,n}(R)$ das matrizes $m \times n$ sobre um corpo k é certamente um R-módulo (e, se $R = k$ for um corpo, um k-espaço vetorial), com a soma de matrizes como soma interna e o produto de um escalar por uma matriz, $\lambda(a_{i,j}) := (\lambda a_{i,j})$ como produto por escalares. Ademais, a multiplicação entre matrizes, como definida no curso básico de álgebra linear, é uma operação associativa, ou seja $(AB)C = A(BC)$, todas as vezes que as matrizes A, B e C têm tamanhos compatíveis para serem multiplicadas. As matrizes identidades I_n, do tamanho oportuno, também atuam como elemento neutro do produto de matrizes. Portanto o conjunto $S = M_{n,n}(R)$ das matrizes quadradas $n \times n$ sobre um anel comutativo, onde o produto de matrizes é uma operação interna, será um anel (não comutativo) e portanto uma R-álgebra não comutativa. Podemos considerar o anel não comutativo $S[x] = (M_{n,n}(k))[x]$ dos polinômios a coeficientes em S, que também resultará uma R-álgebra nao comutativa, sua não comutatividade provindo do próprio produto em S. Provamos que, para anéis comutativos, a avaliação dos polinômios em um elemento fixado produz um homomorfismo de $S[x]$ em S. Este fato não é mais verdadeiro se S não é comutativo, uma vez que, enquanto no produto de polinômios abstratos assumimos que todos os coeficientes comutam com a variável, recolhendo aqueles à esquerda desta, na avaliação efetiva em S tal comutatividade com a matriz que substitui o símbolo x não procede mais. Apesar disto, podemos provar o seguinte fato (por simplicidade, vamos nos restringir ao caso em que $k = R$ é um corpo):

Proposição 76. *Seja k um corpo, A um elemento fixado de $S = M_{n,n}(k)$ e considere o subconjunto Σ de $S[x]$ formado pela expressões $\{p(A) \mid p \in R[x]\}$. Então Σ é uma subálgebra comutativa de S, que coincide com a imagem do homomorfismo que se obtém restringindo a $k[x]$ a função (que não é homomorfismo!) $ev_A : S[x] \to S$.*

Demonstração. É claro que Σ é fechado pela soma, pelo produto, e e pelo produto por escalares, uma vez que $k[x]$ é, e que os elementos neutros das operações provém dos polinômios constantes 0 e 1 calculados em A, e como tais pertencem a Σ. Obviamente potências de A comutam: $A^i A^j = A^{i+j} = A^j A^i$. Agora, como os elementos de k comutam entre si e com A, o produto $(\sum p_i A^i)(\sum q_j A^j)$ será uma combinação dos A_k cujo k-ésimo coeficiente será $\sum_{i+j=k} p_i q_j = \sum_{i+j=k} p_j q_i$ e tal observação da simetria prova a comutatividade em Σ. Observe que a dimensão de S como k-espaço vetorial é n^2 e portanto Σ terá dimensão finita. □

Definição 77. Considerando o homomorfismo de anéis comutativos $ev_A : k[x] \to \Sigma$ definido acima, e lembrando que $k[x]$ é um domínio a ideais principais, o ideal $I := Ker(av_A|_{k[x]})$ deverá ser principal. Como Σ tem dimensão finita $\leq n^2$, os elementos A^i, $i \in \underline{n^2 + 1}$ devem ser linearmente dependentes, resultando em uma combinação não trivial $p(A) = \sum p_i A^i = 0$, com $p \neq 0$. Portanto, $I = Ker(av_A|_{k[x]})$ não é nulo, e o seu gerador mônico se dirá "o polinômio mínimo de A em k" e será denotado por m_A, $m_{k,A}$ ou $m_k(A)$.

Voltaremos a tratar do polinômio mínimo de uma matriz, após estabelecer alguns resultados sobre determinantes.

Observação 78. Seja $B = S^{-1}AS$ uma matriz semelhante com $A \in M_{n,n}(k)$ e $p \in k[x]$ um polinômio. Então

$$p(B) = \sum p_i (S^{-1}AS)^i = \sum p_i S^{-1} A^i S = S^{-1} p(A) S$$

e em particular $p(B) = 0 \leftrightarrow p(A) = 0$, que prova a coincidência dos polinômios mínimos de duas matrizes semelhantes sobre o mesmo corpo.

4.2. **Determinantes.** Lembraremos nesta unidade algumas propriedades básicas dos determinantes que se estudam nos cursos básicos de álgebra linear. Lembrando que S_n é o grupo das permutações de n elementos, definimos o determinante de uma matriz quadrada $n \times n$ sobre um anel R (utilizaremos de preferência anéis comutativos) como:

$$\det(A) := \sum_{\sigma \in S_n} (-1)^{\mathrm{sig}(\sigma)} \prod_{k \in \underline{n}} a_{k,\sigma(k)}.$$

Em outras palavras podemos dizer que o determinante é uma soma, a menos de sinais, de todos os possíveis produtos de n elementos obtidos pegando somente um elemento de cada linha e um de cada coluna. Dito desta maneira alternativa, se, em um termo fixado de tal soma, o elemento escolhido na linha k-ésima está na coluna $\sigma(k)$-ésima, a função σ precisa ser bijetora e o termo considerado é o termo associado à permutação σ na linha acima. Lembremos a escritura de uma matriz A $m \times n$ na notação que explicita seus vetores-linha ou seus vetores-coluna, como em:

$$A = (a_{i,j}) = \begin{pmatrix} \ell_1 \\ \ell_2 \\ \dots \\ \ell_m \end{pmatrix} = (c_1\, c_2\, \dots\, c_n).$$

Em outros livros (mais avançados) verão uma definição de determinante mais sofisticada como "a única forma multilinear sobre as linhas e colunas, alternante, que vale 1 na matriz identidade", que requer algumas definições e alguns teoremas de existência e unicidade. Tal definição de um lado complica, de outro simplifica (supostamente) e interpreta conceitualmente as propriedades dos determinantes. Nós preferimos a definição clássica e a demonstração manual dos teoremas, que ajudará a ganhar disciplina no uso das somatórias e familiaridade com as permutações.

Proposição 79. *Utilizando a notação anterior e lembrando que no nosso contexto, $m = n$, é fácil mostrar os seguintes fatos para matrizes sobre um anel comutativo R:*

a) *O determinante é uma função n-linear sobre as linhas e sobre as colunas, isto é, para qualquer índice i ou j:*

$$\det \begin{pmatrix} \ell_1 \\ \dots \\ \lambda \ell_i + \mu \ell_i' \\ \dots \\ \ell_n \end{pmatrix} = \lambda \det \begin{pmatrix} \ell_1 \\ \dots \\ \ell_i \\ \dots \\ \ell_n \end{pmatrix} + \mu \det \begin{pmatrix} \ell_1 \\ \dots \\ \ell_i' \\ \dots \\ \ell_n \end{pmatrix}, \quad e$$

$$\det(c_1 \dots [\lambda c_j + \mu c_j'] \dots c_n) = \lambda \det(c_1 \dots c_j \dots c_n) + \mu \det(c_1 \dots c_j' \dots c_n)$$

b) *O determinante de uma matriz com uma linha nula ou uma coluna nula é zero*

c) *Trocando duas linhas, o determinante muda de sinal (este procedimento é definido como um tipo de operação elementar e esta propriedade do determinante se chama alternância)*

d) *O determinante de uma matriz com duas linhas iguais ou com duas colunas iguais é zero (esta propriedade é equivalente à alternância em característica diferente de 2)*

e) *O determinante resulta multiplicado por uma constante c se substituo uma linha (ou coluna) pela sua multiplicação por c (este procedimento é definido como um tipo de operação elementar)*
f) *O determinante não muda se acrescento a uma linha (ou coluna) um múltiplo de outra (este procedimento é definido como um tipo de operação elementar)*
g) *O determinante da matriz identidade é* 1
h) *Se* $A^T = (a^T_{i,j})$ *é a matriz transposta de* A, $a^T_{i,j} = a_{j,i}$, *então* $\det A^T = \det A$.

Demonstração. Para a) basta colocar em evidência os escalares na expressão seguinte do determinante do lado esquerdo da igualdade almejada:

$$\sum_{\sigma\in S_n}(-1)^{\mathrm{sig}(\sigma)}(\lambda a_{k,\sigma(k)} + \mu a'_{k,\sigma(k)})\prod_{k\neq i} a_{k,\sigma(k)}.$$

e no caso das colunas se faz o mesmo com

$$\sum_{\sigma\in S_n}(-1)^{\mathrm{sig}(\sigma)}(\lambda a_{k,\sigma(k)} + \mu a'_{k,\sigma(k)})\prod_{\sigma(k)\neq j} a_{k,\sigma(k)}.$$

b) se obtém de a) pondo λ, $\mu = 0$ e matrices com qualquer escolha da linha i-ésima (ou coluna j-ésima) e as restantes linhas (ou colunas) iguais à da matriz dada.
Para c), se a transposição $\tau = (i\ j)$ troca os índices das linhas que são trocadas ao passar de A para A', de modo que $a'_{i,s} = a_{j,s}$ e $a'_{j,s} = a_{i,s}$ teremos:

$$\det A' = \sum_{\sigma\in S_n}(-1)^{\mathrm{sig}(\sigma)}a'_{i,\sigma(i)}a'_{j,\sigma(j)}\prod_{k\neq i,j} a_{k,\sigma(k)} = \sum_{\sigma\in S_n}(-1)^{\mathrm{sig}(\sigma)}a_{j,\sigma\tau(j)}a_{i,\sigma\tau(i)}\prod_{k\neq i,j} a_{k,\sigma\tau(k)}$$

e, observando que $\varphi = \sigma\tau$ descreve S_n a medida que σ o descreve, e também que $\mathrm{sig}(\varphi) = -\mathrm{sig}(\sigma)$, a expressão final pode ser reescrita como

$$\sum_{\varphi\in S_n}(-1)\cdot(-1)^{\mathrm{sig}(\varphi)}\prod_{k\in\underline{n}} a_{k,\varphi(k)} = -\det A.$$

O argumento para colunas é análogo.
d) segue do fato que, se A tem duas linhas iguais, trocando estas linhas obtenho novamente A, mas o item anterior mostrou que $\det A = -\det A$, que, em característica diferente de 2, dá $\det A = 0$. Em característica 2 temos que usar um argumento parecido com a prova de c) para dividir as permutações em dois conjuntos cujos termos associados se cancelam (o leitor pode tentar e, se não conseguir sozinho, imitar a segunda demonstração do teorema de Binet que será dada mais adiante). De fato *d*) pode ser usado para provar *c*, mediante a aplicação da linearidade por linhas (ou colunas) do determinante da matriz que substitui, a ambas as linhas (ou colunas) a serem trocadas, a soma delas.
e) é um caso especial de a), onde na segunda matriz a linha (coluna) a somar é posta como nula
f) é um caso de a), onde a segunda matriz tem duas linhas iguais e $\lambda = 1$
g) é óbvio
h) Na definição de determinante podemos substituir $a_{k,\sigma(k)}$ com $a_{\sigma(k),k}$, que podemos trocar pelo termo genérico $a_{\ell,\sigma^{-1}(\ell)}$, sendo que as σ são bijetoras, mostrando que isto não altera o valor da expressão total, pois a medida que σ descreve S_n, σ^{-1} também o descreve. Isto mostra o papel simétrico das linhas e colunas na definição de determinante e o resultado segue imediatamente. Este item poderia ser usado formalmente para obter uma prova das versões "por colunas" de todos os enunciados anteriores partindo de suas versões "por linhas". □

O próximo resultado, muito notório, é conhecido como teorema de Binet. Daremo dele duas provas.

Teorema 80. *Seja R um anel e $A, B \in \mathcal{M}_{n\times n}(R)$. Então* $\det(AB) = \det(A)\det(B)$.

Demonstração. (PRIMEIRA DEMONSTRAÇÃO) Podemos proceder por indução inversa sobre o numero k das colunas de B que são vetores da base canônica (que podemos supor todas distintas, senão B e AB têm duas colunas iguais e o teorema vale trivialmente) Para $k = n$, B é uma matriz de permutação, associada a σ, ou seja $b_{i,j} = \delta_{j,\sigma(i)}$ e AB será uma permutação segundo σ das linhas de A, portanto, aplicando c) repetidamente, A e AB têm o mesmo determinante a menos do sinal $\mathrm{sig}(\sigma) = \det B$ e o teorema vale. Agora, se a j-ésima coluna de B não for um vetor da base canônica, por linearidade por colunas, o determinante $\det(AB)$ é combinação linear (com elementos de B, $b_{k,j}$, sendo os coeficientes) das matrices obtidas de AB substituindo sua j-ésima coluna pela k-ésima coluna de A, ou seja, os produtos AB', onde B' é obtida substituindo sua j-ésima coluna pelo k-ésimo vetor canônico. É imediato ver, neste momento, que a afirmação segue por indução. □

Vamos dar agora uma prova mais direta.

Demonstração. (SEGUNDA DEMONSTRAÇÃO) Ponha $A = (a_{i,j}), B = (b_{i,j})$. Notamos que o produto

$$\det(A)\det(B) = (\sum_{\sigma\in S_n}(-1)^{\mathrm{sig}(\sigma)}\prod_{i\in\underline{n}}a_{i,\sigma(i)})(\sum_{\varphi\in S_n}(-1)^{\mathrm{sig}(\varphi)}\prod_{j\in\underline{n}}b_{j,\phi(j)})$$

pode ser reescrito como

$$\sum_{\sigma\in S_n}\sum_{\varphi\in S_n}-1^{\mathrm{sig}(\sigma\varphi)}(\prod_{i\in\underline{n}}a_{i,\sigma(i)}\prod_{j\in\underline{n}}b_{j,\varphi(j)})$$

cujo termo genérico é um produto que não depende de nenhuma relação entre i e j enquanto eles variam e descrevem todo $\underline{n}$. Em outras palavras podemos pôr, para escrever um produto só, $j = \sigma(i)$, e deixar j (ou, equivalentemente, i), variar em $\underline{n}$. Poderíamos ter posto $j = i$ ou usado qualquer outra permutação, mas a escolha σ é a que vai ajudar na demonstração. Finalmente, dado que para cada σ fixado, a função $\mu_\sigma : \varphi \mapsto \psi = \varphi\circ\sigma$ é uma bijeção de S_n, podemos deixar variar, na segunda soma $\psi = \varphi\circ\sigma$ ao invés de φ, obtendo dessa forma:

$$\sum_{\sigma\in S_n}\sum_{\psi\in S_n}-1^{\mathrm{sig}(\psi)}\prod_{i\in\underline{n}}a_{i,\sigma(i)}b_{\sigma(i),\psi(i)}.$$

Agora, considerando $C = AB = (c_{i,j})$, observamos que $c_{i,j} = \sum_{k\in\underline{n}}a_{i,k}b_{k,j}$ e portanto

$$\det AB = \sum_{\psi\in S_n}-1^{\mathrm{sig}(\psi)}\prod_{i\in\underline{n}}\sum_{k\in\underline{n}}a_{i,k}b_{k,\psi(i)} = \sum_{\psi\in S_n}-1^{\mathrm{sig}(\psi)}\sum_{\vec{k}:\underline{n}\to\underline{n}}\prod_{i\in\underline{n}}a_{i,\vec{k}(i)}b_{\vec{k}(i),\psi(i)}.$$

que por sua vez pode ser reescrito como:

$$\sum_{\vec{k}:\underline{n}\to\underline{n}}\sum_{\psi\in S_n}-1^{\mathrm{sig}(\psi)}\prod_{i\in\underline{n}}a_{i,\vec{k}(i)}b_{\vec{k}(i),\psi(i)}$$

onde $\vec{k}$ está variando entre todas as funções de $\underline{n}$ em si (possivelmente não bijetoras). Para concluir e restringir a soma a $\vec{k} = \alpha \in S_n$, portanto, basta mostrar que, toda vez que $\vec{k}$ não é bijetora (ou, equivalentemente, injetora), sua controbuição ao termo de direita é zero. Suponha que existam $j, \ell \in \underline{n}$ tais que $\vec{k}(j) = \vec{k}(\ell)$ e seja $\tau = (j\,\ell)$ a transposição que troca i e j. Composição com τ define uma bijeção entre as permutações par de S_n e as ímpar, de modo que a somatória interna indicizada em $S_n = A_n \cup (S_n\backslash A_n)$ pode ser dividida em duas partes,

e podemos comparar cada termo $-1^{\text{sig}(\psi)} \prod_{i\in\underline{n}} a_{i,\vec{k}(i)} b_{\vec{k}(i),\psi(i)}$, para cada $\psi \in A_n$ com o termo correspondente na segunda soma, associado à permutação ímpar $\psi \circ \tau$, ou seja, $-(-1^{\text{sig}(\psi)}) \prod_{i\in\underline{n}} a_{i,\vec{k}(i)} b_{\vec{k}(i),\psi(\tau(i))}$ que afirmamos ser seu oposto, e desta maneira todos os termos se cancelam. Com efeito, além do sinal negativo inicial, para $i \neq j, \ell$, todos os fatores são iguais e, além deles, o primeiro produto exibe um fator $a_{j,\vec{k}(j)} b_{\vec{k}(j),\psi(j)} a_{\ell,\vec{k}(\ell)} b_{\vec{k}(\ell),\psi(\ell)}$ enquanto o segundo possui um fator $a_{j,\vec{k}(j)} b_{\vec{k}(j),\psi(\ell)} a_{\ell,\vec{k}(\ell)} b_{\vec{k}(\ell),\psi(j)} = a_{j,\vec{k}(j)} b_{\vec{k}(\ell),\psi(\ell)} a_{\ell,\vec{k}(\ell)} b_{\vec{k}(j),\psi(j)}$, ou seja, os mesmos fatores. □

O próximo resultado é conhecido como "desenvolvimento por linhas ou por colunas". Antes de enunciá-lo, precisamos definir os cofatores de uma matriz. Precisamos introduzir, nesta altura, uma notação mais breve para o determinante de uma matriz A, ou seja:

$$|A| := \det A$$

.

Definição 81. Se $A = (a_{i,j})$ é uma matriz quadrada, seja $\hat{A}_{i,j}$ a matriz $(n-1) \times (n-1)$ obtida cancelando de A a i-ésima linha e a j-ésima coluna. O "cofator" ou "complemento algébrico" de $a_{i,j}$ é o escalar $A_{i,j} := (-1)^{i+j} |\hat{A}_{i,j}|$. Colocando na i-ésima linha e j-ésima coluna o escalar $(\hat{A})_{i,j} := A_{i,j}$, se forma a chamada matriz dos cofatores $\hat{A}$, cuja tranposta, $adj(A)$ é chamada matriz adjunta de A. A seguinte fórmula é conhecida como fórmula ou expansão de Laplace.

Proposição 82. *(EXPANSÃO DO DETERMINANTE POR LINHAS/COLUNAS) Para $A \in M_{n,n}(R)$, para todo i fixado em $\underline{n}$ temos:*

$$\det A = \sum_{j=1}^{n} a_{i,j} A_{i,j}.$$

Da mesma forma, para todo j fixado em $\underline{n}$, temos:

$$\det A = \sum_{i=1}^{n} a_{i,j} A_{i,j}$$

Demonstração. Vamos provar o enunciado para linhas:

$$\sum_{\sigma\in S_n} (-1)^{\text{sig}(\sigma)} a_{i,\sigma(i)} \prod_{k\neq i} a_{k,\sigma(k)} = \sum_{j\in\underline{n}} a_{i,j} \sum_{\sigma(i)=j} (-1)^{\text{sig}(\sigma)} \prod_{k\neq i} a_{k,\sigma(k)}$$

Podemos estabelecer uma bijeção entre as permutações σ de S_n que mandam i em j e as permutações φ de S_{n-1}, mandando σ na composição φ assim obtida:

$$\underline{n-1} \xrightarrow{a} \underline{n}\backslash\{i\} \xrightarrow{\sigma} \underline{n}\backslash\{j\} \xrightarrow{b} \underline{n-1},$$

onde a e b são as (únicas) bijeções crescentes. O leitor pode verificar que $(-1)^{\text{sig}(\sigma)} = (-1)^{i+j}(-1)^{\text{sig}(\varphi)}$ e que se $\varphi = b\sigma a$ e $k = a(r)$, temos $A_{k,\sigma(k)} = (\hat{A}_{i,j})_{r,\varphi(r)}$ e portanto a expressão a direita da igualdade inicial se torna:

$$\sum_{j=1}^{n} a_{i,j} (-1)^{i+j} \sum_{\varphi\in S_{n-1}} (-1)^{\text{sig}(\varphi)} \prod_{r\in\underline{n-1}} (\hat{A}_{i,j})_{r,\varphi(r)},$$

como desejado. □

Proposição 83. *Para toda matriz quadrada A sobre um anel, $A \cdot adj(A) = adj(a) \cdot A = \det A \cdot 1_n$. Consequentemente A é invertível se e somente se $\det A \neq 0$ e sua inversa é dada por $\frac{1}{\det A} adj(A)$.*

Demonstração. Provamos que $\det A = \sum_{k=1}^{n} a_{i,k}A_{i,k}$. Porém, se $i \neq j$, a expressão $\sum_{k=1}^{n} a_{i,k}A_{j,k}$ é nula, por ser a expansão, segundo a linha j-ésima, da matriz obtida de A copiando a linha i-ésima no lugar da j-ésima, ou sejam de uma matriz com duas linhas iguais. Portanto, se $A \cdot adj(A) = (b_{i,j})$, temos que:

$$b_{i,j} = \sum a_{i,k}(adj(A))_{k,j} = \sum a_{i,k}A_{j,k} = \delta_{i,j}det(A),$$

como era desejado. A outra igualdade se obtém usando expansões por coluna, e é deixada como exercício. Para a segunda parte da proposição, se A é invertível, então $AA^{-1} = 1_n$ e pelo teorema de Binet $\det A \det A^{-1} = 1$, tornando impossível $\det A = 0$. Viceversa se $\det A \neq 0$ temos $A[\frac{1}{\det A}adj(A)] = [\frac{1}{\det A}adj(A)]A = 1_n$. □

4.3. **O teorema de Hamilton-Cayley.** Antes de começar esta seção é oportuno observar que a multiplicação de matrizes quadradas sobre um anel R pode ser definida mesmo quando R não é comutativo e resulta, mesmo assim, associativa. No caso que $R = M_{n,n}(k)$, onde k é um corpo, consideremos uma matriz $A \in R$ e seja Σ a subálgebra de R formada pelos polinômios em A com coeficientes em k. Já provamos que Σ é comutativo, contém A como elemento e k como sub-corpo (polinômios constantes correspondendo a matrizes escalares). Σ é uma k-subálgebra comutativa (de matrizes) da k-álgebra não comutativa (de matrizes) R. Agora, posso construir o anel $S_2 = M_{n,n}(R)$ que portanto contém $S_1 = Mn, n(\Sigma)$ como subanel. Se trata de considerar "matrizes de matrizes". O produto é associativo em S_2, porém as ferramentas teóricas desenvolvidas até agora somente podem ser usadas para matrizes sobre anéis comutativos e portanto não em S_2 mas sim em S_1. Em particular, se $B \in S_1$ é verdade que $adj(B)B = \det B$, onde $\det B$ será um elemento de Σ.

Teorema 84. *(Teorema de Hamilton-Cayley) O polinômio mínimo de uma matriz quadrada com coeficientes em um corpo divide seu polinômio característico.*

Demonstração. Utilizaremos as notações definidas acima. Seja $A = (a_{i,j}) \in M_{n,n}(k)$ e seja $p_A(x) = \det(A - x1_n) \in k[x]$ seu polinômio característico. Como o polinômio mínimo $m = m_A$ gera o núcleo da avaliação em A, a afirmação $m_a|p_a$ é equivalente a dizer que $p_A(A) = 0$ em Σ. Sejam S_1 e S_2 definidos como acima e note que seus elementos são matrizes de matrizes, podendo conter em cada posição matrizes, assim como escalares, identificados com matrizes escalares. Considere o seguinte elemento de S_1:

$$B = \begin{pmatrix} a_{1,1} - A & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} - A & \cdots & a_{2,n} \\ \vdots & & \ddots & \vdots \\ a_{n,1} & \cdots & & a_{n,n} - A \end{pmatrix}$$

Temos que $p_a(A) = \det B^T = \det B$. Por outro lado, como S_1 é uma álgebra de matrizes sobre o anel comutativo Σ, temos também $\det B 1_n = adj(B) \cdot B$ em S_1. Se trata de provar que tal matriz é nula. Esta matriz é um elemento de S_1 e portanto de $S_2 = M_{n,n}(R)$ e estabelece uma transformação lineares sobre os vetores de R^n. Considere a base de R formada pelas matrizes $E_{i,j}$, onde Ei, j tem 1 na posição (i, j) e zero nas outras. Podemos pensar $E_{i,j}$ também como matriz cuja j-ésima coluna é o vetor canônico e_i e as outras são nulas. Vamos formar os vetores-coluna v_j de R^n da seguinte forma:

$$v_j = \begin{pmatrix} E_{1,j} \\ E_{2,j} \\ \vdots \\ E_{n,j} \end{pmatrix} = \begin{pmatrix} ((0)\ldots(0)(e_1)(0)\ldots(0)) \\ ((0)\ldots(0)(e_2)(0)\ldots(0)) \\ \vdots \\ ((0)\ldots(0)(e_n)(0)\ldots(0)) \end{pmatrix}$$

onde os vetores em parênteses $(0) = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ e os $(e_i) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ devem ser pensados como colocados por colunas para formar as matrizes $E_{i,j}$. Podemos calcular, agora, que pelo jeito que foram construídas essas matrizes, para todo $j \in \underline{n}$ teremos $Bv_j = 0$ e portanto $\det(B)1_n v_j = adj(B)Bv_j = 0$ onde $\det(B)1_n$ é uma matriz "escalar"de matrizes em $S_1 \in S_2$, ou seja:

$$\det B 1_n = \begin{pmatrix} \det B & 0 & \cdots & 0 \\ 0 & \det B & \cdots & 0 \\ \vdots & & \ddots & 0 \\ 0 & \cdots & & \det B \end{pmatrix}$$

Mas o leitor pode observar que $\det Bv_1$ é um vetor, por um lado nulo, por outro lado, formado de matrizes cujas colunas são as próprias colunas de $\det B$, resultando em $\det B = 0$ e portanto $p_A(A) = 0$ □

Departamento de Matemática, Universidade Federal de Pernambuco, Avenida Jornalista Aníbal Fernandes, S/N - Cidade Universitária, Recife/PE - Brasil - 50740-560

E-mail address, M. Barone: marco@dmat.ufpe.br