



Universidade Federal do Pará
Instituto de Ciências Exatas e Naturais
Faculdade de Matemática

Ândson José Silva Almeida

Grupos Matriciais são Variedades Diferenciáveis

Belém
2023

Ândson José Silva Almeida

Grupos Matriciais são Variedades Diferenciáveis

Projeto de pesquisa apresentado como requisito para aprovação na disciplina de Trabalho de Conclusão de Curso na Faculdade de Matemática da Universidade Federal do Pará.

Orientador: Prof. Dr. Alex Cardenas

Coorientador: Prof. Dr. Ashish Mishra

Belém

2023

Resumo

O objetivo principal deste projeto é demonstrar que os grupos matriciais possuem a estrutura de variedades diferenciáveis. Nosso foco é desenvolver as ferramentas matemáticas necessárias para entender o enunciado desse teorema e sua demonstração, como por exemplo: a diferenciação em espaços reais de dimensão finita, sequências e séries de vetores e alguns resultados algébricos. Além disso, damos ênfase em demonstrar resultados importantes que são fundamentais no desenvolvimento da teoria dos grupos matriciais, tais como o Teorema da Função inversa e o da suavidade da exponencial de matrizes.

Palavras-chave: Exponencial de Matrizes; Grupos Matriciais; Variedades Diferenciais.

Abstract

The principal objective of this project is to demonstrate that matrix groups have the structure of differentiable manifolds. Our focus is to develop the mathematical tools necessary for the understanding of the statement of this theorem and its demonstration, for instance: the differentiation in finite dimensional real vector spaces, sequences and series of vectors and some algebraic results. Also, we put emphasis to prove the important results that are fundamental in the development of the theory of matrix groups, such as the Inverse Function Theorem and the smoothness of the matrix exponential.

Key-words: Matrix Exponential; Matrix Groups; Differentiable Manifolds.

Conteúdo

Introdução	1
1 Preliminares	3
1.1 Espaços Vetoriais	3
1.2 Análise	5
2 Sequências e Séries	15
2.1 Definições de sequência e de séries	15
2.2 A exponencial de matrizes	19
3 Diferenciação	21
3.1 Definição de diferenciabilidade e propriedades gerais	21
3.2 Diferenciabilidade da função exponencial	31
4 Grupos matriciais e suas Álgebras de Lie	37
4.1 Grupos Matriciais	37
4.1.1 Grupos compactos	37
4.2 Álgebras de Lie	39
5 Grupos Matriciais como variedades diferenciáveis	43
Referências	45

Introdução

Como o título deste trabalho sugere, o objetivo principal do texto é enunciar e demonstrar o teorema de que grupos matriciais (os subgrupos fechados de $GL_n(\mathbb{K})$) são variedades diferenciáveis. Este teorema é de grande interesse pois nos dá uma das classes mais importantes de grupos de Lie, além de ser uma boa introdução a teoria dos grupos de Lie. E por que a teoria dos grupos de Lie é importante? Ela é um ponto de interseção e conexão entre álgebra e geometria, e assim através dos grupos de Lie problemas em geometria podem ser traduzidos em problemas algébricos e vice-versa, se tornando uma poderosa ferramenta para demonstrações nessas áreas, bem como para a teoria de equações diferenciais e aplicações na física.

Como nos concentramos em grupos matriciais, cabe perguntar-se o quão restritivo é essa classe em relação aos grupos de Lie. Em primeiro lugar, na maioria das aplicações, os grupos de Lie que surgem em física, geometria diferencial e equações diferenciais acabam sendo grupos matriciais. Além disso, os grupos de Lie que são compactos são isomorfos a grupos matriciais, um resultado poderoso e cuja a demonstração é sofisticada, estando além deste trabalho. O leitor interessado pode consultar o Capítulo 12 de [2].

Assumi-se que o leitor possua conhecimento de análise na reta e cálculo em uma variável, além de conhecimentos sobre álgebra linear e álgebra abstrata, pouco mais que as definições de grupo e de anel. O motivo para isso é que, sendo uma tese de conclusão de curso da graduação em licenciatura em matemática, é desejado que o material aqui apresentado possa ser consultado por futuros estudantes que do curso de graduação. De todo modo, é feita uma revisão e apresentação dos resultados básicos mais importantes no Capítulo 1. Além disso, grande parte do material é voltada ao desenvolvimento de ferramentas como definições e teoremas prévios que servirão para o entendimento do enunciado e da prova do teorema principal. O Capítulo 3, por exemplo, é fundamental pois lá é demonstrado o Teorema da função inversa e o da suavidade da função exponencial.

De forma mais detalhada, o Capítulo 1 está dividido em duas partes: uma revisando resultados prévios de álgebra linear e outra apresentando resultados de análise em $\mathbb{R}^n$ e $\mathbb{C}^n$ que serão necessários no decorrer do trabalho. A maioria das demonstrações é feita de modo muito semelhante a teoria em $\mathbb{R}$ (que assumimos que o leitor já conheça), mas apresentamos a maioria das demonstrações mesmo assim.

No Capítulo 2 desenvolvemos a teoria das sequências e séries em espaços vetoriais de dimensão finita com vistas a dois objetivos: o primeiro é apresentar resultados que serão usadas na demonstração do Teorema da função inversa, a ser demonstrado no Capítulo 3; o segundo é definir a função exponencial e apresentar algumas de suas propriedades.

No Capítulo 3 desenvolvemos a teoria da diferenciação de funções de várias variáveis reais, provando suas propriedades mais importantes. Provamos no final do capítulo a que a função exponencial é infinitamente diferenciável.

No Capítulo 4 definimos grupos matriciais e definimos suas álgebras de Lie como o espaço tangente ao grupo no ponto $p = I$, com I a matriz identidade (aqui seguimos a terminologia de [7], que não menciona o colchete de Lie na definição de álgebra de Lie).

Por fim, no Capítulo 5 provamos o teorema que dá nome a este trabalho.

Havendo excelentes textos sobre a teoria dos grupos matriciais e como esse texto é sobre a parte inicial da teoria, demos atenção apresentar as demonstrações de forma detalhada, apresentando provas para resultados que costumam ser deixadas a cargo do leitor nos livros-texto (destacamos, por exemplo, o Teorema 3.29 que afirma que a função exponencial é diferenciável de classe C^∞).

1 Preliminares

1.1 Espaços Vetoriais

Ao longo de todo o texto trabalharemos com o corpo $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, onde $\mathbb{R}$ é o corpo dos números reais e $\mathbb{C}$ o corpo dos números complexos. O conjunto das matrizes $m \times n$ com entradas em $\mathbb{K}$ será representado por $M_{m \times n}(\mathbb{K})$. Quando $m = 1$ escreveremos $\mathbb{K}^n$ e quando $n = m$ escreveremos $M_n(\mathbb{K})$. O conjunto das funções $\mathbb{K}$ -lineares com domínio em um espaço vetorial E e contradomínio em um espaço vetorial F será representado por $\mathcal{L}(E, F)$, e quando $n = m$ escreveremos $\mathcal{L}(E)$. Definimos $GL(\mathbb{K}^n) \subset \mathcal{L}(\mathbb{K}^n)$ como o conjunto das transformações lineares que possuem inversa. Usaremos $\mathbf{0}$ para representar o vetor nulo de $\mathbb{K}^n$.

Foi consultado na elaboração desta seção os livros [6] e [1].

Proposição 1.1. $\mathcal{L}(\mathbb{K}^n)$ é um anel, onde a multiplicação é a composição de funções e a adição é a soma de funções.

Demonstração. Usaremos a definição de Anel dada em [1]. A composição de funções é associativa e temos $I \in \mathcal{L}(\mathbb{K}^n)$, que é o elemento neutro da composição de funções e portanto será a unidade do anel. Temos também que $\mathcal{L}(\mathbb{K}^n)$ é um espaço vetorial e portanto um grupo abeliano em relação a soma cujo o elemento neutro é a função identicamente nula, que irá desempenhar o papel de zero do anel. Resta apenas a verificação da distributividade, que vale pois $A \circ (B + C) = A \circ B + A \circ C$ vem da linearidade de A e $(A + B) \circ C = A \circ C + B \circ C$ vem da definição de soma de funções. $\square$

Teorema 1.2. $GL(\mathbb{K}^n)$, munido com a operação de composição de funções, é um grupo.

Demonstração. Usaremos a definição de grupo dada em [1]. A composição de funções lineares é uma função linear e a composição de duas funções bijetoras é uma função bijetora, assim $GL(\mathbb{K}^n)$ é fechado para composição de funções, que é uma operação associativa.

Seja $T \in GL(\mathbb{K}^n)$. Sendo T uma bijeção, sua inversa T^{-1} existe. Observemos que para todo $u \in \mathbb{K}^n$ se tem $u = T(T^{-1}u)$. Assim

$$\begin{aligned} T^{-1}(\alpha v + \beta u) &= T^{-1}(\alpha(T \circ T^{-1})(v) + \beta(T \circ T^{-1})(u)) \\ &= T^{-1}(T(\alpha T^{-1}(v) + \beta T^{-1}(u))) \\ &= \alpha T^{-1}(v) + \beta T^{-1}(u), \end{aligned}$$

desse modo $T^{-1}(\alpha v + \beta u) = \alpha T^{-1}(v) + \beta T^{-1}(u)$. Logo $T \in GL(\mathbb{K}^n) \Rightarrow T^{-1} \in \mathcal{L}(\mathbb{K}^n)$ e assim $T \in GL(\mathbb{K}^n) \Rightarrow T^{-1} \in GL(\mathbb{K}^n)$, sendo fechado para a operação de inversão. Desse modo $GL(\mathbb{K}^n)$ é um grupo. $\square$

Definição 1.3. Sejam os conjuntos $\{a_i | 1 \leq i \leq n\} \subset \mathbb{K}^n$ e $\{b_j | 1 \leq j \leq m\} \subset \mathbb{K}^m$ bases de seus respectivos espaços. Dado $T \in \mathcal{L}(\mathbb{K}^m, \mathbb{K}^n)$, definimos a *matriz da transformação linear* T nas bases $(a_i)_{1 \leq i \leq n}$ e $(b_j)_{1 \leq j \leq m}$ como a matriz A tal que $T(b_j) = \sum_{i=1}^n [A]_{i,j} a_i$.

Definição 1.4. Definimos a base canônica de $M_{m,n}(\mathbb{K})$ como as matrizes $e_{i,j}$ tais que $[e_{i,j}]_{p,q} = 1$ quando $(i,j) = (p,q)$, caso contrário $[e_{i,j}]_{p,q} = 0$. escreveremos apenas e_i para representar as matrizes $e_{1,i} \in \mathbb{K}^n$.

Definição 1.5. Sejam $A, B \in M_{m,n}(\mathbb{K})$. Definimos o produto interno $\langle A, B \rangle = \sum_{i,j=1}^n A_{i,j} \overline{B}_{i,j}$.

Uma importante propriedade do produto interno é que $v = \sum_{i=1}^n \langle v, e_i \rangle e_i$. Com e_i os vetores da base canônica de $\mathbb{R}^n$. Desse modo, é possível calcular a matriz de uma função linear na base canônica através do produto interno.

Proposição 1.6. *Seja a função $\mathfrak{z} : \mathcal{L}(\mathbb{K}^m, \mathbb{K}^n) \rightarrow M_{n \times m}(\mathbb{K})$ dada por $[\mathfrak{z}(T)]_{i,j} = \langle T(u_j), e_i \rangle$ onde temos $\{e_i | i \in \mathbb{N} \ 1 \leq i \leq n\}$ é a base canônica de $\mathbb{K}^n$ e $\{u_j | j \in \mathbb{N} \ 1 \leq j \leq m\}$ a base canônica de $\mathbb{K}^m$. Então $\mathfrak{z}(T)$ é a matriz da transformação linear T nas bases canônicas de $\mathbb{K}^m$ e $\mathbb{K}^n$.*

Definição 1.7. O produto de uma matriz $A \in M_{n \times p}(\mathbb{K})$ por uma matriz $B \in M_{p \times m}(\mathbb{K})$ é definido como a matriz $AB \in M_{n \times m}(\mathbb{K})$ tal que

$$[AB]_{i,j} = \sum_{k=1}^p [A]_{i,k} [B]_{k,j}$$

Proposição 1.8. *Seja $A \in M_{p \times m}(\mathbb{K})$ a matriz da transformação linear $T \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^m)$ e $B \in M_{n \times p}(\mathbb{K})$ a matriz da transformação linear $F \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^p)$, então $\mathfrak{z}(T \circ F) = AB$.*

Demonstração. Sejam $\{e_i | i \in \mathbb{N} \ 1 \leq i \leq n\}$ a base canônica de $\mathbb{K}^n$, $\{u_j | j \in \mathbb{N} \ 1 \leq j \leq m\}$ a base canônica de $\mathbb{K}^m$ e $\{v_k | k \in \mathbb{N} \ 1 \leq k \leq p\}$ a base canônica de $\mathbb{K}^p$. Temos

$$\begin{aligned} [\mathfrak{z}(T \circ F)]_{i,j} &= \langle (T \circ F)(u_j), e_i \rangle \\ &= \langle T(F(u_j)), e_i \rangle = \left\langle T \left(\sum_{k=1}^p \langle F(u_j), v_k \rangle v_k \right), e_i \right\rangle. \end{aligned}$$

Na igualdade acima, usamos $F(u_j) = \sum_{k=1}^p \langle F(u_j), v_k \rangle v_k$. Agora

$$\begin{aligned} \left\langle \sum_{k=1}^p T(\langle F(u_j), v_k \rangle v_k), e_i \right\rangle &= \left\langle \sum_{k=1}^p \langle F(u_j), v_k \rangle T(v_k), e_i \right\rangle \\ &= \sum_{k=1}^p \langle \langle F(u_j), v_k \rangle T(v_k), e_i \rangle = \sum_{k=1}^p \langle F(u_j), v_k \rangle \langle T(v_k), e_i \rangle \\ &= \sum_{k=1}^p [\mathfrak{z}(F)]_{k,i} [\mathfrak{z}(T)]_{j,k} = \sum_{k=1}^p [A]_{i,k} [B]_{k,j} = \sum_{k=1}^p [A]_{i,k} [B]_{k,j}, \end{aligned}$$

encerrando a demonstração. □

Corolário 1.9. *Com as operações de soma e multiplicação de matrizes, $M_n(\mathbb{K})$ é um anel, e além disso $\mathfrak{z}$ é um isomorfismo de anéis.*

Demonstração. Notemos que $\mathfrak{z}$ é uma bijeção e sua inversa $\mathfrak{z}^{-1} : M_{n \times m}(\mathbb{K}) \rightarrow \mathcal{L}(\mathbb{K}^m, \mathbb{K}^n)$ é definida como $\mathfrak{z}^{-1}A \left(\sum_{j=1}^m a_j u_j \right) = \sum_{j=1}^m a_j \sum_{i=1}^n [A]_{i,j} e_i$. Onde por conveniência escrevemos $\mathfrak{z}^{-1}A$ para representar $\mathfrak{z}^{-1}(A)$.

Para o caso $n = m$, que é o que estamos interessados, obtemos para matrizes $A, B, C \in M_n(\mathbb{K})$ que se tem

$$A(BC) = \mathfrak{z}(\mathfrak{z}^{-1}A) \left(\mathfrak{z}(\mathfrak{z}^{-1}B) \mathfrak{z}(\mathfrak{z}^{-1}C) \right) = \mathfrak{z}(\mathfrak{z}^{-1}A) \mathfrak{z} \left((\mathfrak{z}^{-1}B)(\mathfrak{z}^{-1}C) \right) = \mathfrak{z} \left((\mathfrak{z}^{-1}A) \left((\mathfrak{z}^{-1}B)(\mathfrak{z}^{-1}C) \right) \right).$$

A associatividade é válida para composição de funções, assim $(\mathfrak{z}^{-1}A) \left((\mathfrak{z}^{-1}B)(\mathfrak{z}^{-1}C) \right)$ é igual a $\left((\mathfrak{z}^{-1}A)(\mathfrak{z}^{-1}B) \right) (\mathfrak{z}^{-1}C)$ e desse modo obtemos

$$A(BC) = \mathfrak{z} \left(\left((\mathfrak{z}^{-1}A)(\mathfrak{z}^{-1}B) \right) (\mathfrak{z}^{-1}C) \right) = \left(\mathfrak{z}(\mathfrak{z}^{-1}A) \mathfrak{z}(\mathfrak{z}^{-1}B) \right) \mathfrak{z}(\mathfrak{z}^{-1}C) = (AB)C.$$

De onde concluímos que a multiplicação de matrizes em $M_n(\mathbb{K})$ é associativa. A prova de que $M_n(\mathbb{K})$ é um anel agora pode ser feita de maneira inteiramente análoga a demonstração da Proposição 1.1.

Sabendo agora que $M_n(\mathbb{K})$ é um anel, e através da Proposição 1.6 obtemos que $\mathfrak{z}$ é um homomorfismo de anéis, e por ser uma função bijetora é um isomorfismo. $\square$

Definição 1.10. Representaremos por $GL_n(\mathbb{K})$ o grupo das *matrizes inversíveis* $n \times n$ com o produto de matrizes. Note que $GL_n(\mathbb{K}) = \mathfrak{z}(GL(\mathbb{K}^n))$.

Ao longo deste trabalho iremos identificar uma transformação linear com sua matriz.

Definição 1.11. Dizemos que $\zeta : \mathbb{K}^n \times \mathbb{K}^m \rightarrow \mathbb{K}^r$ é uma função bilinear quando dados qualquer $\alpha \in \mathbb{K}$ e quaisquer vetores $v_1, v_2 \in \mathbb{K}^n$ e $u_1, u_2 \in \mathbb{K}^m$ se tem

- (i) $\zeta(\alpha v_1, u_1) = \zeta(v_1, \alpha u_1) = \alpha \zeta(v_1, u_1)$,
- (ii) $\zeta(v_1 + v_2, u_1) = \zeta(v_1, u_1) + \zeta(v_2, u_1)$,
- (iii) $\zeta(v_1, u_1 + u_2) = \zeta(v_1, u_1) + \zeta(v_1, u_2)$.

Os exemplos mais importantes das funções bilineares é a composição de transformações lineares (o que no presente trabalho equivale a multiplicação de matrizes, como já vimos), a multiplicação por escalar e o produto interno $\langle, \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ (cabe destacar que o produto interno em $\mathbb{C}^n$ não é bilinear).

1.2 Análise

Esta seção é um breve resumo dos principais resultados de análise. Recomendamos ao leitor interessado nos teoremas aqui apresentado consultar o livros [5], [3] e [4].

Definição 1.12. Chamamos de norma uma função $|\cdot|_N : \mathbb{K}^n \rightarrow \mathbb{R}$ satisfazendo, para todo $\alpha \in \mathbb{K}$ e $u, v \in \mathbb{K}^n$:

- (i) $|u + v|_N \leq |u|_N + |v|_N$;
- (ii) $|\alpha v|_N = |\alpha| |v|_N$;
- (iii) se $|u|_N = 0$ então $u = \mathbf{0}$.

Proposição 1.13. Se $|\cdot|_N$ em $\mathbb{K}^n$ é uma norma, então qualquer que seja $v \in \mathbb{K}^n$ nós temos:

- (i) $|\mathbf{0}|_N = 0$;
- (ii) $|v|_N = |-v|_N$;
- (iii) $0 \leq |v|_N$;
- (iv) $||v|_N - |u|_N| \leq |v - u|_N$.

Demonstração. Temos que $|0v|_N = 0|v|_N \Rightarrow |\mathbf{0}|_N = 0$. Temos também as igualdades $|-v|_N = |-1||v|_N = |v|_N$, e assim para o terceiro item

$$|v - v|_N \leq |v|_N + |v|_N \Rightarrow 0 \leq 2|v|_N \Rightarrow 0 \leq |v|_N.$$

para o item (iv), temos

$$|u|_N \leq |u - v|_N + |v|_N \Rightarrow |u|_N - |v|_N \leq |u - v|_N + |v|_N$$

$$|v|_N \leq |u - v|_N + |u|_N \Rightarrow |v|_N - |u|_N \leq |u - v|_N + |v|_N.$$

Através das desigualdades acima se conclui que $||v|_N - |u|_N| \leq |u - v|_N + |v|_N$. $\square$

Teorema 1.14. (a) São normas em $\mathbb{K}^n$ as seguintes funções:

(i) A norma euclidiana $\|A\| = \sqrt{\langle A, A \rangle}$.

(ii) A norma da soma $|A|_s = \sum_{i,j=1}^n |[A]_{i,j}|$.

(iii) A norma do máximo $|A|_m = \max \{|[A]_{i,j}| \mid 1 \leq i, j \leq n\}$.

(b) A função $|X|_e = \sup \{\|X(u)\| \mid \text{com } u \in \mathbb{K}^n, \|u\| = 1\}$ é uma norma em $\mathcal{L}(\mathbb{K}^m, \mathbb{K}^n)$, chamada de norma espectral.

Essas quatro normas e a notação usada para representa-las serão livremente usadas ao longo deste trabalho.

Demonstração. Mostraremos que $| \cdot |_s$ é uma norma. Dadas duas matrizes A, B nós temos:

$$\begin{aligned} |[A]_{i,j} + [B]_{i,j}| &\leq |[A]_{i,j}| + |[B]_{i,j}| \\ \Rightarrow \sum_{i,j=1}^n |[A]_{i,j} + [B]_{i,j}| &\leq \sum_{i,j=1}^n (|[A]_{i,j}| + |[B]_{i,j}|) \\ \Rightarrow \sum_{i,j=1}^n |[A + B]_{i,j}| &\leq \sum_{i,j=1}^n |[A]_{i,j}| + \sum_{i,j=1}^n |[B]_{i,j}| \\ \Rightarrow |A + B|_s &\leq |A|_s + |B|_s. \end{aligned}$$

Seja agora a matriz tA

$$|tA|_s = \sum_{i,j=1}^n |[tA]_{i,j}| = \sum_{i,j=1}^n |t[A]_{i,j}| = \sum_{i,j=1}^n |t| |[A]_{i,j}| = |t| \sum_{i,j=1}^n |[A]_{i,j}| = |t| |A|_s$$

Por último, se $|A| = 0$ então

$$\sum_{i,j=1}^n |[A]_{i,j}| \Rightarrow |[A]_{i,j}| = 0, \quad \forall i, j$$

isso ocorre pois em um somatório de termos não negativos a soma geral é 0 apenas se cada uma das parcelas também for 0, de modo que:

$$|[A]_{i,j}| = 0 \Rightarrow [A]_{i,j} = 0, \quad \forall i, j$$

assim A é a matriz nula, concluindo a prova. Com algumas modificações as mesmas ideias podem mostrar que as outras funções apresentadas também são normas. $\square$

Proposição 1.15. Sejam $A \in M_{n \times p}(\mathbb{K})$ e $B \in M_{p \times m}(\mathbb{K})$, então $|AB|_e \leq |A|_e |B|_e$.

Demonstração. Seja $T : \mathbb{K}^k \rightarrow \mathbb{K}^l$ uma transformação linear e $u \in \mathbb{R}^k, u \neq \mathbf{0}$. Então vale

$$\left\| T \left(\frac{u}{\|u\|} \right) \right\| \leq |T|_e \Rightarrow \|T(u)\| \leq |T|_e \|u\|. \quad (1)$$

Notemos que $\|T(u)\| \leq |T|_e \|u\|$ é válido mesmo para $u = 0$.

Como k, l são genéricos, podemos fazer $(l, k) = (n, p)$ ou $(l, k) = (p, m)$. Portanto a relação (1) é válida para $T = A$ e também para $T = B$. Desse modo obtemos, para todo $u \in \mathbb{K}^m, \|u\| = 1$ que

$$\|A(B(u))\| \leq |A|_e \|B(u)\| \leq |A|_e |B|_e.$$

Desse modo $|AB|_e \leq |A|_e |B|_e$, como queríamos. $\square$

Definição 1.16. Sejam $u \in \mathbb{K}^n$ e $\varepsilon > 0$. Definimos o conjunto $B(u, \varepsilon)$, chamado “bola aberta de centro u e raio ε ” por $B(u, \varepsilon) = \{v \mid \|v - u\| < \varepsilon\}$. Definimos uma bola fechada $B[u, \varepsilon]$ como $B[u, \varepsilon] = \{v \mid \|v - u\| \leq \varepsilon\}$.

Definição 1.17. Dizemos que $U \subset \mathbb{K}^n$ é um conjunto limitado quando existir $r > 0$ tal que $U \subset B[0, r]$. Caso o contrário, diremos que U é ilimitado.

Definição 1.18. Dizemos que um conjunto $U \subset \mathbb{K}^n$ é aberto se para todo $a \in U$ existir $\varepsilon > 0$ tal que $B(a, \varepsilon) \subset U$. Sejam conjuntos $Y \subset X \subset \mathbb{K}^n$. Dizemos que Y é um conjunto aberto de X se existir um conjunto aberto $A \subset \mathbb{K}^n$ com $Y = A \cap X$.

Exemplo 1.19. O conjunto $\emptyset$ e $\mathbb{K}^n$ são conjuntos abertos. As bolas abertas $B(a, \delta)$ onde $a \in \mathbb{K}^n$ e $\delta > 0$ também são conjuntos abertos.

Proposição 1.20. Sejam $A, B \subset \mathbb{K}^n$ conjuntos abertos. Então $A \cap B$ é um conjunto aberto.

Demonstração. Se $a \in A \cap B$ então $a \in A$ e portanto existe $r_1 > 0$ com $B(a, r_1) \subset A$. Do mesmo modo existe $r_2 > 0$ com $B(a, r_2) \subset B$. Seja $r = \min\{r_1, r_2\}$, então $B(a, r) \subset A \cap B$, sendo a um elemento qualquer, $A \cap B$ é aberto, como havíamos afirmado. $\square$

Corolário 1.21. Sejam $\{A_1, \dots, A_k\} \subset \mathbb{K}^n$ uma família finita de conjuntos abertos. Então $\bigcap_{i=1}^k A_i$ é um conjunto aberto.

Proposição 1.22. Sejam $\mathcal{P}(\mathbb{K}^n)$ o conjunto das partes de $\mathbb{K}^n$ e $\mathcal{A} \subset \mathcal{P}(\mathbb{K}^n)$ tal que todo $X \in \mathcal{A}$ é um conjunto aberto, então $A = \bigcup_{X \in \mathcal{A}} X$ é um conjunto aberto.

Demonstração. Dado qualquer $x \in A$, existe algum $X \in \mathcal{A}$ tal que $x \in X$. Como X é aberto então $B(x, \delta) \subset X$ para algum $\delta > 0$. Temos que $X \subset A$ donde $B(x, \delta) \subset A$.

Concluimos então que para todo $x \in A$ existe δ positivo com $B(x, \delta) \subset A$, donde A é um conjunto aberto. $\square$

Definição 1.23. Seja $F \subset \mathbb{K}^n$, dizemos que F é fechado se $\mathbb{K}^n - F = F^C$ (o complementar de F em $\mathbb{K}^n$) for um conjunto aberto. Sejam conjuntos $Y \subset X \subset \mathbb{K}^n$. Dizemos que Y é um conjunto fechado de X se houver um conjunto fechado $F \subset \mathbb{K}^n$ com $Y = F \cap X$.

Exemplo 1.24. O conjunto $\emptyset$ e $\mathbb{K}^n$ são conjuntos fechados. As bolas fechadas $B[a, \delta]$ onde $a \in \mathbb{K}^n$ e $\delta > 0$ também são conjuntos fechados. Todo conjunto finito é fechado.

Definição 1.25. Seja $X \subset \mathbb{K}^n$ um conjunto qualquer e consideremos o conjunto

$$\mathfrak{X} = \{F \mid X \subset F \subset \mathbb{K}^n \text{ e } F \text{ é um conjunto fechado}\}.$$

Definimos então o conjunto $\overline{X}$, chamado de fecho de X como

$$\overline{X} = \bigcap_{F \in \mathfrak{X}} F.$$

Proposição 1.26. Para qualquer conjunto $X \subset \mathbb{K}^n$ o conjunto $\overline{X}$ é um conjunto fechado e $X \subset \overline{X}$.

Demonstração. Sabemos que para todo $F \in \mathfrak{X}$, F^C é um conjunto aberto, logo $\bigcup_{F \in \mathfrak{X}} F^C$ será um

conjunto aberto e assim $\left(\bigcup_{F \in \mathfrak{X}} F^C\right)^C = \bigcap_{F \in \mathfrak{X}} F = \overline{X}$ é um conjunto fechado. Seja agora $x \in X$.

Então $x \in F$ qualquer que seja $F \in \mathfrak{X}$, portanto $x \in \bigcap_{F \in \mathfrak{X}} F$, desse modo temos:

$$x \in X \Rightarrow x \in \overline{X}$$

o que implica em $X \subset \overline{X}$, como queríamos. $\square$

Definição 1.27. Seja um conjunto $U \subset \mathbb{K}^n$ qualquer, dizemos que $a \in U$ é um ponto isolado de U quando $a \notin \overline{U - \{a\}}$.

Proposição 1.28. *Seja $A \subset \mathbb{K}^n$ um conjunto aberto. Então A não possui pontos isolados.*

Demonstração. Seja $a \in A$. Suponha que a seja um ponto isolado de A , isto é, pertença ao complementar de $\overline{A - \{a\}}$. Como $\overline{A - \{a\}}$ é um conjunto fechado, seu complementar é aberto e portanto existe $r > 0$ com $B(a, r) \cap \overline{A - \{a\}} = \emptyset$. Temos

$$B(a, r) \cap \overline{A} = \{a\} \Rightarrow B(a, r) \cap A = \{a\}.$$

Como a intercessão de abertos é aberto, teríamos que $\{a\}$ é um conjunto aberto e portanto existiria $\varepsilon > 0$ como $B(a, \varepsilon) \subset \{a\}$, um absurdo. Obtemos portanto que A não possui pontos isolados. □

Definição 1.29. Seja $U \subset \mathbb{K}^m$ e uma função $f : U \rightarrow \mathbb{K}^n$. Dizemos que f é contínua em $a \in U$ quando para todo $x \in U$ e todo $\varepsilon > 0$ existir um $\delta > 0$ tal que

$$\|x - a\| < \delta \Rightarrow \|f(x) - f(a)\| < \varepsilon.$$

Chamamos de função contínua uma função que é contínua em todos os pontos de seu domínio. Tais funções também poderão ser chamadas de funções de classe C^0 .

Teorema 1.30. *Seja $U \subset \mathbb{K}^m$ e uma função $f : U \rightarrow \mathbb{K}^n$. Temos que se f é contínua então a imagem inversa $f^{-1}(A)$ é um conjunto aberto em U para todo conjunto aberto $A \subset \mathbb{K}^n$.*

Demonstração. A proposição é verdadeira para o caso $f^{-1}(A) = \emptyset$ pois o conjunto vazio é um conjunto aberto.

Suponha agora $a \in f^{-1}(A)$. Então $f(a) \in A$. Como A é aberto, irá existir $\epsilon > 0$ com $B(f(a), \epsilon) \subset A$. Do fato de f ser contínua existe $\delta > 0$ com $x \in B(a, \delta) \cap U \Rightarrow f(x) \in B(f(a), \epsilon)$, logo $x \in B(a, \delta) \cap U \Rightarrow f(x) \in A$, donde temos $(B(a, \delta) \cap U) \subset f^{-1}(A)$

Escolhamos para cada $a \in f^{-1}(A)$ um $\delta_a > 0$ com $(B(a, \delta_a) \cap U) \subset f^{-1}(A)$. Sabemos que $\{a\} \subset B(a, \delta_a) \cap U$ qualquer que seja o $a \in f^{-1}(A)$ e desse modo

$$\begin{aligned} \left(\bigcup_{a \in f^{-1}(A)} \{a\} \right) &\subset \left(\bigcup_{a \in f^{-1}(A)} (B(a, \delta_a) \cap U) \right) \subset f^{-1}(A) \\ \Rightarrow f^{-1}(A) &\subset \left(\bigcup_{a \in f^{-1}(A)} (B(a, \delta_a)) \right) \cap U \subset f^{-1}(A) \\ \Rightarrow U \cap \left(\bigcup_{a \in f^{-1}(A)} (B(a, \delta_a)) \right) &= f^{-1}(A). \end{aligned}$$

Logo $f^{-1}(A)$ é a intercessão de U e um conjunto aberto, como queríamos. □

Definição 1.31. Sejam um conjunto $U \subset \mathbb{K}^m$ e uma função $f : U \rightarrow \mathbb{K}^n$. Dizemos que f é lipschitziana se existir $c > 0$ tal que $\|f(x) - f(y)\| \leq c\|x - y\|$.

Proposição 1.32. *Toda função lipschitziana é contínua.*

Demonstração. Seja $\epsilon > 0$ e tomemos $\delta = \frac{\epsilon}{m}$, onde $m > 0$ é uma constante que satisfaz $\|f(x) - f(y)\| \leq m\|x - y\|$, temos então:

$$\|x - y\| < \frac{\epsilon}{m} \Rightarrow \|f(x) - f(y)\| \leq m\frac{\epsilon}{m} \Rightarrow \|f(x) - f(y)\| \leq \epsilon.$$

□

Teorema 1.33. *Seja $\zeta : \mathbb{K}^m \times \mathbb{K}^n \rightarrow \mathbb{K}^r$ bilinear, então ζ é contínua.*

Demonstração. Suponha $(a, b) \in \mathbb{K}^n \times \mathbb{K}^m$ e que $\{e_i | i \in \mathbb{N} \ 1 \leq i \leq n\}$ a base canônica de $\mathbb{K}^n$ e $\{u_j | j \in \mathbb{N} \ 1 \leq j \leq m\}$ a base canônica de $\mathbb{K}^m$. Então escrevendo $a_i = \langle a, e_i \rangle$ e $b_j = \langle b, u_j \rangle$ nós temos

$$\begin{aligned}\zeta(a, b) &= \zeta \left(\sum_{i=1}^n a_i e_i, \sum_{j=1}^m b_j u_j \right) \\ &= \sum_{i=1}^n \sum_{j=1}^m b_j a_i \zeta(e_i, u_j).\end{aligned}$$

Seja agora $w = \max\{\|\zeta(e_i, u_j)\| \mid i, j \in \mathbb{N}, \ 1 \leq i \leq n \text{ e } 1 \leq j \leq m\}$. Então

$$\begin{aligned}\|\zeta(a, b)\| &\leq w \left| \sum_{i=1}^n a_i \left(\sum_{j=1}^m b_j \right) \right| \Rightarrow \|\zeta(a, b)\| \leq w \sum_{i=1}^n |a_i| \left(\sum_{j=1}^m |b_j| \right) \\ &\Rightarrow \|\zeta(a, b)\| \leq w |b|_s \sum_{i=1}^n |a_i| \Rightarrow \|\zeta(a, b)\| \leq w |a|_s |b|_s.\end{aligned}$$

Sejam agora um novo par $(a', b') \in \mathbb{K}^n \times \mathbb{K}^m$. Se tem que

$$\begin{aligned}\|\zeta(a, b) - \zeta(a', b')\| &= \|\zeta(a, b) - \zeta(a, b') + \zeta(a, b') - \zeta(a', b')\| \\ &= \|\zeta(a, b - b') + \zeta(a - a', b')\| \\ &\leq \|\zeta(a, b - b')\| + \|\zeta(a - a', b')\| \\ &\leq w |a|_s |b - b'|_s + w |a - a'|_s |b'|_s.\end{aligned}$$

Seja agora $r = \max\{|(a, b)|_s, |(a', b')|_s\}$. Estamos identificando $\mathbb{K}^n \times \mathbb{K}^m \approx \mathbb{K}^{n+m}$, assim $x \in \mathbb{K}^n$ corresponde as primeiras n coordenadas do vetor $(x, y) \in \mathbb{K}^{n+m}$ e $y \in \mathbb{K}^m$ corresponde as m coordenadas seguintes. Portanto $|a|_s \leq |(a, b)|_s$ e $|b'|_s \leq |(a', b')|_s$ assim ambos são menores que r e temos

$$\begin{aligned}\|\zeta(a, b) - \zeta(a', b')\| &\leq wr(|b - b'|_s + |a - a'|_s) \\ &\leq wr \left(\sum_{j=1}^m |b_j - b'_j| + \sum_{i=1}^n |a_i - a'_i| \right) \\ &\leq wr |(a - a', b - b')|_s \\ &\leq wr |(a, b) - (a', b')|_s.\end{aligned}$$

Desse modo, ζ é lipschitziana em $B(0, K) \subset \mathbb{K}^{n+m}$, portanto é contínua em $B(0, r)$. Como r é arbitrário, todo ponto $(a, b) \in \mathbb{K}^n \times \mathbb{K}^m$ pertence a algum $B(0, r)$ e portanto ζ é contínua em todo o $\mathbb{K}^n \times \mathbb{K}^m$. $\square$

Definição 1.34. Seja uma função $f : U \rightarrow \mathbb{K}^n$ onde $U \subset \mathbb{K}^m$. A i -ésima função coordenada da função f , $f_i : U \rightarrow \mathbb{K}$, é definida como a única função tal que $f_i(v)$ é a i -ésima coordenada de $f(v)$. Note que temos a identidade $f_i(x) = \langle f(x), e_i \rangle$, onde e_i é o i -ésimo vetor da base canônica de $\mathbb{K}^n$.

As funções contínuas possuem as seguintes propriedades, cuja a demonstração pode ser encontrada em [5] e também em [4].

- (i) Sejam $U \subset \mathbb{K}^m$ e funções $f, g : U \rightarrow \mathbb{K}^n$. Se f e g são funções contínuas em $a \in U$, então $f + g$ é uma função contínua em a .

- (ii) Sejam os conjuntos $U \subset \mathbb{K}^m$ e $V \subset \mathbb{K}^m$. Consideremos as funções $f : U \rightarrow \mathbb{K}^m$ e $g : V \rightarrow \mathbb{K}^k$ e um ponto $a \in U$ tal que $f(a) \in V$, se f for contínua em a e g for contínua em $f(a)$, então $f \circ g$ é contínua em a .
- (iii) Seja uma função $f : U \rightarrow \mathbb{K}^n$ onde $U \subset \mathbb{K}^m$. Então f é contínua se e somente se cada uma das suas n funções coordenadas forem contínuas.

Definição 1.35. Definimos limite em duas situações:

- (a) Seja $U \subset \mathbb{K}^m$, uma função $f : U \rightarrow \mathbb{K}^n$ e a um ponto não isolado de $\overline{U}$. Dizemos que $y \in \mathbb{K}^n$ é o limite de f em a quando para todo $\varepsilon > 0$ e todo $x \in U$ existir $\delta > 0$ tal que

$$0 < \|x - a\| < \delta \Rightarrow \|y - f(x)\| < \varepsilon. \quad (2)$$

- (b) Seja $U \subset \mathbb{K}^m$ um conjunto ilimitado e uma função $f : U \rightarrow \mathbb{K}^n$. Dizemos que $y \in \mathbb{K}^n$ é o limite de f quando f vai ao infinito se para todo $\varepsilon > 0$ existir algum $\Delta > 0$ tal que se tenha:

$$\Delta < \|x\| \Rightarrow \|y - f(x)\| < \varepsilon.$$

Teorema 1.36. *Seja a um ponto não isolado de $U \subset \mathbb{K}^m$ e $f : U \rightarrow \mathbb{K}^n$, então f é contínua em a se e somente se f possuir limite em a e tivermos $f(a) = \lim_{x \rightarrow a} f(x)$.*

Demonstração. Da definição de continuidade temos que para todo $x \in U$ e $\varepsilon > 0$ existe $\delta > 0$ tal que

$$\|x - a\| < \delta \Rightarrow \|f(x) - f(a)\| < \varepsilon.$$

Para todo x , em particular para $x \in U$, nós temos

$$0 < \|x - a\| < \delta \Rightarrow \|x - a\| < \delta.$$

Assim obtemos

$$0 < \|x - a\| < \delta \Rightarrow \|f(x) - f(a)\| < \varepsilon,$$

e pela definição de limite, $f(a) = \lim_{x \rightarrow a} f(x)$. □

Em contraste com a proposição acima, temos que se $a \in U$ é um ponto isolado, então toda função definida em U deve ser contínua em a em decorrência da definição de continuidade, enquanto que os limites nunca existem em pontos isolados.

O limite possui as seguintes propriedades básicas, cuja as demonstrações podem ser encontradas em [5] e também em [4].

- (a) Seja $U \subset \mathbb{K}^m$ e $a \in \overline{U}$ um ponto não isolado. Se $f : U \rightarrow \mathbb{K}^n$ possuir um limite em a , então este limite é único. Quando $y \in \mathbb{K}^n$ for o limite de f em a então escrevemos $y = \lim_{x \rightarrow a} f(x)$.
- (b) Seja $U \subset \mathbb{K}^m$ ilimitado. Se $f : U \rightarrow \mathbb{K}^n$ possuir um limite no infinito, então este limite é único. Quando $y \in \mathbb{K}^n$ for o limite de f no infinito então escrevemos $y = \lim_{x \rightarrow \infty} f(x)$.
- (c) Seja $U \subset \mathbb{K}^n$, funções $f, g : U \rightarrow \mathbb{K}^n$ que possuam limite no ponto $a \in \overline{U}$, então vale:

$$\lim_{x \rightarrow a} (f + g)(x) = \lim_{x \rightarrow a} f(x) + \lim_{x \rightarrow a} g(x).$$

- (d) Seja $U \subset \mathbb{K}^n$ ilimitado e funções $f, g : U \rightarrow \mathbb{K}^n$ que possuam limite no no infinito, então vale:

$$\lim_{x \rightarrow \infty} (f + g)(x) = \lim_{x \rightarrow \infty} f(x) + \lim_{x \rightarrow \infty} g(x).$$

Proposição 1.37. *Valem as seguintes regras para a composição de limites.*

(a) *Sejam os conjuntos $U \subset \mathbb{K}^m$, $V \subset \mathbb{K}^n$, funções $f : U \rightarrow \mathbb{K}^n$ e $g : V \rightarrow \mathbb{K}^k$. Considere os ponto $a \in \overline{U}$ e $b \in \overline{V}$ tais que*

- (i) $\lim_{x \rightarrow a} f(x) = b$,
- (ii) g possui limite em b .

Então nesse caso temos

$$\lim_{x \rightarrow a} (g \circ f)(x) = \lim_{y \rightarrow b} g(y).$$

(b) *Sejam os conjuntos $U \subset \mathbb{K}^m$ e $V \subset \mathbb{K}^n$, com U um conjunto ilimitado, e funções $f : U \rightarrow \mathbb{K}^n$ e $g : V \rightarrow \mathbb{K}^k$. Considere um ponto $b \in \overline{V}$ tal que*

- (i) $\lim_{x \rightarrow \infty} f(x) = b$,
- (ii) g possui limite em b .

Então nesse caso temos

$$\lim_{x \rightarrow \infty} (g \circ f)(x) = \lim_{y \rightarrow b} g(y).$$

(c) *Sejam os conjuntos $U \subset \mathbb{K}^m$ e $V \subset \mathbb{K}^n$, com V um conjunto ilimitado e funções $f : U \rightarrow \mathbb{K}^n$ e $g : V \rightarrow \mathbb{K}^k$. Considere um ponto $a \in U$ tal que*

- (i) *Para todo $M > 0$ existe algum $\delta > 0$ tal que $0 < \|x - a\| < \delta \Rightarrow M < \|f(x)\|$,*
- (ii) g possui limite no infinito.

Então nesse caso temos

$$\lim_{x \rightarrow a} (g \circ f)(x) = \lim_{y \rightarrow \infty} g(y).$$

(d) *Sejam os conjuntos $U \subset \mathbb{K}^m$ e $V \subset \mathbb{K}^n$, ambos sendo conjuntos ilimitados, e funções $f : U \rightarrow \mathbb{K}^n$ e $g : V \rightarrow \mathbb{K}^k$ satisfazendo*

- (i) *Para todo $M > 0$ existe algum $\Delta > 0$ tal que $\Delta < \|x\| \Rightarrow M < \|f(x)\|$,*
- (ii) g possui limite no infinito.

Então nesse caso temos

$$\lim_{x \rightarrow \infty} (g \circ f)(x) = \lim_{y \rightarrow \infty} g(y).$$

Demonstração. Faremos a demonstração de (a). As outras se demonstram de maneira muito semelhante.

Seja $k = \lim_{y \rightarrow b} g(y)$, então pela definição de limite temos que para todo ε_1 existe um δ_1 tal que

$$0 < \|y - b\| < \delta_1 \Rightarrow \|g(y) - k\| < \varepsilon_1.$$

Tomando $y = f(x)$, temos

$$0 < \|f(x) - b\| < \delta_1 \Rightarrow \|g(f(x)) - k\| < \varepsilon_1. \quad (3)$$

Também pela definição de limite temos que para todo ε_2 existe δ_2 tal que

$$0 < \|x - a\| < \delta_2 \Rightarrow \|f(x) - b\| < \varepsilon_2.$$

Tomando $\varepsilon_2 = \delta_1$ então existe um δ_3 tal que

$$0 < \|x - a\| < \delta_3 \Rightarrow \|f(x) - b\| < \delta_1. \quad (4)$$

Por (3) e por (4) temos que

$$0 < \|x - a\| < \delta_3 \Rightarrow \|g(f(x)) - k\| < \varepsilon_1.$$

Assim pela definição de limite temos $\lim_{x \rightarrow a} (g \circ f)(x) = k = \lim_{y \rightarrow b} g(y)$, como queríamos. $\square$

Proposição 1.38. *Sejam um subconjunto $V \subset \mathbb{K}^n$ e uma função $g : V \rightarrow \mathbb{K}^k$ contínua no ponto $b \in V$. Então temos que:*

(a) *Dada uma função $f : U \rightarrow V$ definida no conjunto $U \subset \mathbb{K}^m$ e um ponto $a \in \bar{U}$ tal que $b = \lim_{x \rightarrow a} f(x)$, $b \in V$. Nesse caso temos*

$$\lim_{x \rightarrow a} (g \circ f)(x) = g \left(\lim_{x \rightarrow a} f(x) \right).$$

(b) *Dada uma função $f : U \rightarrow V$ definida no conjunto $U \subset \mathbb{K}^m$, com U ilimitado e $b = \lim_{x \rightarrow \infty} f(x)$. Nesse caso temos*

$$\lim_{x \rightarrow \infty} (g \circ f)(x) = g \left(\lim_{x \rightarrow \infty} f(x) \right).$$

Demonstração. Mostraremos (a). A demonstração de (b) é análoga.

Se b for um ponto isolado, então existe $r > 0$ com $B(b, r) \cap U = \{b\}$. Então tomando $\varepsilon = r$ na fórmula (2) temos que existe $\delta > 0$ tal que $0 < \|x - a\| < \delta \Rightarrow f(x) \in B(b, r)$, ou seja, $0 < \|x - a\| < \delta \Rightarrow f(x) = b$ e assim

$$0 < \|x - a\| < \delta \Rightarrow g(f(x)) = g(b) \Rightarrow g(f(x)) - g(b) = 0$$

e é óbvio que $g(f(x)) - g(b) < \epsilon$ para todo $\epsilon > 0$. Assim $\lim_{x \rightarrow a} (g \circ f)(x) = g \left(\lim_{x \rightarrow a} f(x) \right)$.

Se b não for um ponto isolado, então usando o Teorema 1.36 e a Proposição 1.37 temos

$$\lim_{x \rightarrow a} (g \circ f)(x) = \lim_{y \rightarrow b} g(y) = g(b) = g \left(\lim_{x \rightarrow a} f(x) \right).$$

$\square$

Teorema 1.39. *Sejam os conjuntos $U \subset \mathbb{K}^m$ e $V \subset \mathbb{K}^n$, com V um conjunto fechado. Se uma função $f : U \rightarrow V$ possui limite em um ponto $a \in \bar{U}$, então $\lim_{x \rightarrow a} f(x) \in V$.*

Demonstração. Seja $b = \lim_{x \rightarrow a} f(x)$. Suponha por absurdo que $b \in \mathbb{K}^n - V$. Como V é fechado, seu complementar é aberto e existe $r > 0$ com $B(b, r) \subset \mathbb{K}^n - V$. Da definição de limite, existe um $\delta > 0$ com $0 < \|x - a\| \leq \delta \Rightarrow \|f(x) - b\| < r$. Desse modo, para $x \in B(a, \delta) - \{a\}$ se tem $f(x) \in \mathbb{K}^n - V$, contradizendo a hipótese de que f era uma função de U em V . $\square$

Definição 1.40. Dizemos que um subconjunto $K \subset \mathbb{K}^n$ é um conjunto compacto quando for um conjunto fechado e limitado.

Proposição 1.41 (Weierstrass). *Sejam $K \subset \mathbb{K}^m$ um conjunto compacto e $f : K \rightarrow \mathbb{R}$ uma função contínua. Então f possui pontos de máximo e de mínimo.*

Demonstração. Pode ser encontrada a demonstração desse resultado no capítulo 8 de [5] e também no capítulo 1 de [4]. $\square$

Definição 1.42. Dizemos que duas normas $|\cdot|_1, |\cdot|_2$ sobre $\mathbb{K}^n$ são equivalentes se existirem constantes $c, k > 0$ tais que $c|v|_1 \leq |v|_2 \leq k|v|_1$ para todo $v \in \mathbb{K}^n$.

Lema 1.43. A norma da soma e a norma euclidiana são equivalentes.

Demonstração. Inicialmente mostremos que $\|x\| \leq |x|_s$ para todo $x \in \mathbb{K}^n$.

$$\|x\|^2 = \sum_{i=1}^n |x_i|^2 \leq \sum_{i,j=1}^n |x_i||x_j| = |x|_s |x|_s,$$

e assim extraindo a raiz quadrada obtemos $\|x\| \leq |x|_s$, como queríamos.

Consideremos agora os vetores $\tilde{x}$ e ℓ dados por $\tilde{x}_i = |x_i|$ e $\ell_i = 1$ para todo i . Temos que $|x|_s = \langle \ell, \tilde{x} \rangle$, mas pela desigualdade de Cauchy [6] nós também temos

$$\langle \ell, \tilde{x} \rangle \leq \|\ell\| \|\tilde{x}\| \Rightarrow |x|_s \leq \sqrt{n} \|x\|,$$

como queríamos demonstrar. $\square$

Teorema 1.44. Quaisquer duas normas sobre $\mathbb{K}^n$ são equivalentes.

Demonstração. Tomemos uma norma $|\cdot|_N$ qualquer e seja $M = \max\{|e_i|_N | 1 \leq i \leq k\}$, então vale

$$|v|_N \leq M|v|_s, \forall v \in \mathbb{K}^n.$$

Em particular, $|v|_N \leq \sqrt{n}M\|v\|$ e assim toda norma é lipschitziana e portanto contínua.

Consideremos agora a bola fechada $B[\mathbf{0}, 1]$. Por definição, este é um conjunto compacto e assim $|\cdot|_N$ possui em $B[\mathbf{0}, 1]$ um valor mínimo, que chamaremos de C . Para todo $v \neq \mathbf{0}$ nós temos

$$C \leq \left| \frac{v}{\|v\|} \right|_N \Rightarrow C\|v\| \leq |v|_N.$$

A desigualdade acima é imediata para $\mathbf{0}$, e desta forma toda norma é equivalente com a norma da euclidiana. Sejam agora duas normas $|\cdot|_1$ e $|\cdot|_2$. Então vale

$$m_1\|v\| \leq |v|_1 \leq M_1\|v\|, \forall v \in \mathbb{K}^n$$

$$m_2\|v\| \leq |v|_2 \leq M_2\|v\|, \forall v \in \mathbb{K}^n.$$

Temos $\|v\| \leq \frac{|v|_2}{m_2}$ e $\frac{|v|_1}{M_1} \leq \|v\|$, donde obtemos $|v|_1 \leq \frac{M_1}{m_2}|v|_2$. Por raciocínio análogo obtemos $\frac{m_1}{M_2}|v|_2 \leq |v|_1$, provando o teorema. $\square$

Proposição 1.45. (a) Seja $A \subset \mathbb{K}^n$ aberto e $a \in A$. Para qualquer norma $|\cdot|_N$ de $\mathbb{K}^n$ existe um $\delta \in \mathbb{R}$ positivo tal que para todo $v \in \mathbb{K}^n$ se tem $|a - v|_N < \delta \Rightarrow v \in A$.

(b) Seja $B \subset \mathbb{K}^n$ e $|\cdot|_N$ uma norma em $\mathbb{K}^n$. Se para todo $b \in B$ existir um δ tal que para todo $v \in \mathbb{K}^n$ se tem $|b - v|_N < \delta \Rightarrow v \in B$, então B é aberto.

Demonstração. Como $a \in A$, então existe $\delta > 0$ com $B(a, \delta) \subset A$. Sabemos que existe $c > 0$ tal que $\|x - a\| \leq c|x - a|_N$, desse modo $|x - a|_N < \frac{\delta}{c}$ implica que $\|x - a\| < \delta$, de onde temos

$$B_N\left(a, \frac{\delta}{c}\right) \subset B(a, \delta) \subset A.$$

Por outro lado, se para todo $b \in B$ existir $\delta > 0$ com $|x - b|_N < \delta \Rightarrow x \in B$ e sabendo que existe $k > 0$ de modo que $|x - b|_N \leq k\|x - b\|$ obtemos que $\|x - b\| < \frac{\delta}{k}$ implica que $|x - a|_N < \delta$

e assim $B\left(b, \frac{\delta}{k}\right) \subset B_N(b, \delta) \subset B$, provando a proposição. $\square$

2 Sequências e Séries

2.1 Definições de sequência e de séries

Definição 2.1. Uma sequência $x : \mathbb{N} \rightarrow \mathbb{K}^n$, onde $\mathbb{N}$ é o conjunto dos números naturais, será representada por $(x_i)_{i \in \mathbb{N}}$ e os seus elementos $x(i)$ escritos como x_i , isto é, $x_i = x(i)$.

Exemplo 2.2. Seja $f : \mathbb{K}^n \rightarrow \mathbb{K}^n$ e $x \in \mathbb{K}^n$, podemos definir uma sequência $(x_n)_{n \in \mathbb{N}}$, chamada de sequência recursiva, da seguinte forma

$$x_0 = x \text{ e } x_{n+1} = f(x_n) \quad \forall n \in \mathbb{N}.$$

Definição 2.3. Uma função $f : U \rightarrow \mathbb{K}^n$, onde $U \subset \mathbb{R}^m$. Se existir uma constante $c \in (0, 1)$ com $\|f(x) - f(y)\| < c\|x - y\|$ quaisquer que sejam $x, y \in U$, chamamos f de c -contração.

Apresentamos a seguinte a versão do teorema do ponto fixo de Banach, e algumas de suas consequências que serão usadas na demonstração do teorema da função inversa, seguindo de perto a apresentação dos resultados em [4].

Teorema 2.4. *Seja $L \subset \mathbb{K}^n$ fechado e $f : L \rightarrow L$ uma c -contração. Então existe e é único um $p \in L$ tal que $f(p) = p$.*

Demonstração. Começemos observando o seguinte fato

$$\begin{aligned} \|x_{i+1} - x_i\| &= \|f(x_i) - f(x_{i-1})\| \leq c\|x_i - x_{i-1}\| \\ &\Rightarrow \|x_{i+1} - x_i\| \leq c\|x_i - x_{i-1}\|. \end{aligned}$$

Repetindo esse processo j vezes obtemos $\|x_{i+1} - x_i\| \leq c^j\|x_{i-j+1} - x_{i-j}\|$ e assim, fazendo $i = j$ se tem que $\|x_{i+1} - x_i\| \leq c^i\|x_1 - x_0\|$.

Tomando $x \in L$ arbitrário e definamos a sequência recursiva $x_0 = x$, $x_{n+1} = f(x_n)$. Para $m > n$ se tem

$$\|x_m - x_n\| = \left\| \sum_{i=n}^{m-1} x_{i+1} - x_i \right\| \leq \sum_{i=n}^{m-1} \|x_{i+1} - x_i\| \leq \sum_{i=n}^{m-1} c^i \|x_1 - x_0\| \leq \left(\sum_{i=n}^{m-1} c^i \right) \|x_1 - x_0\|$$

como $c > 0$ temos

$$\left(\sum_{i=n}^{m-1} c^i \right) \|x_1 - x_0\| \leq \left(\sum_{i=n}^{\infty} c^i \right) \|x_1 - x_0\| = \frac{\|x_1 - x_0\| c^n}{1 - c}.$$

Se $x_1 = x_0$ então f possui um ponto fixo (em particular $(x_i)_{i \in \mathbb{N}}$ é constante logo convergente). Se $x_1 \neq x_0$ então para $\varepsilon > 0$ temos

$$\frac{\|x_1 - x_0\| c^n}{1 - c} < \varepsilon \Leftrightarrow c^n < \frac{\varepsilon(1 - c)}{\|x_1 - x_0\|} \Leftrightarrow \ln(c^n) < \ln \left(\frac{\varepsilon(1 - c)}{\|x_1 - x_0\|} \right)$$

onde a segunda equivalência vem do fato da função logarítmica ser crescente. Então

$$\ln(c^n) < \ln \left(\frac{\varepsilon(1 - c)}{\|x_1 - x_0\|} \right) \Leftrightarrow n \ln(c) < \ln \left(\frac{\varepsilon(1 - c)}{\|x_1 - x_0\|} \right) \Leftrightarrow n > (\ln c)^{-1} \ln \left(\frac{\varepsilon(1 - c)}{\|x_1 - x_0\|} \right),$$

onde a última equivalência vem do fato de $\ln(c)$ ser negativo, uma vez que $c \in (0, 1)$. Assim $(x_i)_{i \in \mathbb{N}}$ é uma sequência de Cauchy [5] e portanto convergente. Seja p o limite desta sequência. Como L é fechado, $p \in L$.

Pela definição de limite, para todo $\varepsilon > 0$ existe $n \in \mathbb{N}$ tal que $m > n \Rightarrow \|p - x_m\| < \varepsilon \Rightarrow \|f(p) - x_{m+1}\| < \varepsilon c$, logo $f(p)$ também é limite dessa sequência e pela unicidade do limite $p = f(p)$.

Por fim, sejam p_1, p_2 pontos fixos de f . Então nós temos que

$$\|f(p_1) - f(p_2)\| \leq c\|p_1 - p_2\| \Rightarrow \|p_1 - p_2\| \leq c\|p_1 - p_2\| \Rightarrow (1 - c)\|p_1 - p_2\| \leq 0 \Rightarrow \|p_1 - p_2\| \leq 0$$

e portanto $p_1 = p_2$, provando a unicidade do ponto fixo, como queríamos. $\square$

Lema 2.5. *Seja $X \subset \mathbb{K}^n$ com $B[a, r] \subset X$ e $f : X \rightarrow \mathbb{K}^n$ uma c -contração tal que $\|f(a) - a\| \leq (1 - c)r$. Então f possui um ponto fixo em $B[a, r]$.*

Demonstração. Basta mostrar que $f(B[a, r]) \subset B[a, r]$ que então recaímos nas hipóteses do teorema anterior. Para $x \in B[a, r]$ temos

$$\|x - a\| \leq r \Rightarrow \|f(x) - f(a)\| \leq rc$$

$$\|f(a) - a\| + \|f(x) - f(a)\| \leq rc + (1 - c)r$$

$$\|a - f(x)\| \leq r \Rightarrow f(x) \in B[a, r]$$

donde temos $x \in B[a, r] \Rightarrow f(x) \in B[a, r]$ e assim $f(B[a, r]) \subset B[a, r]$. $\square$

Lema 2.6. *Seja $U \subset \mathbb{K}^m$ um conjunto aberto e $\varphi : U \rightarrow \mathbb{K}^m$ uma c -contração. Definindo $g : U \rightarrow \mathbb{K}^m$ como $g(x) = x + \varphi(x)$, temos que $Im(g) \subset \mathbb{K}^m$ é um conjunto aberto, g é injetiva e tanto g quanto $g^{-1} : Im(g) \rightarrow U$ são funções contínuas.*

Demonstração. Seja $a \in U$ um elemento qualquer e $b = g(a)$. Então $b \in Im(g)$.

Dizer que $Im(g)$ é aberto significa que existe $\delta > 0$ tal que $B(b, \delta) \subset Im(g)$, ou seja, se temos que $\|y - b\| < \delta$ então $y \in Im(g)$. É esta última condição que iremos mostrar.

Dado um y qualquer consideremos a função $\xi_y(x) = y - \varphi(x)$. Temos

$$x = \xi_y(x) \Rightarrow x = y - \varphi(x) \Rightarrow x + \varphi(x) = y \Rightarrow g(x) = y,$$

basta então encontrarmos um ponto fixo de ξ_y . Para podermos usar o Lema 2.5 devemos ter

$$\|\xi_y(a) - a\| < (1 - c)r \Rightarrow \|y - \varphi(a) - a\| < (1 - c)r \Rightarrow \|y - b\| < (1 - c)r,$$

onde $r > 0$ é tal que $B[a, r] \subset U$. Desse modo quando tomamos $\delta = (1 - c)r$, obtemos $B(b, \delta) \subset Im(g)$ e como b é um elemento arbitrário de $Im(g)$ temos que $Im(g)$ é um conjunto aberto.

É imediato que g é uma função contínua. Para mostrar a existência e a continuidade de g^{-1} notemos que

$$\begin{aligned} \|g(x) - g(y)\| &= \|x + \varphi(x) - y - \varphi(y)\| \geq \|x - y\| - \|\varphi(x) - \varphi(y)\| \Rightarrow \\ &\Rightarrow \|g(x) - g(y)\| \geq (1 - c)\|x - y\| \end{aligned}$$

donde temos que para $x \neq y$ se tem $g(x) \neq g(y)$ pois $c < 1$ por definição. Escrevendo $g(x) = w$ e $g(y) = z$ temos

$$\begin{aligned} \|w - z\| &\geq (1 - c)\|g^{-1}(w) - g^{-1}(z)\| \Rightarrow \\ &\Rightarrow \|g^{-1}(w) - g^{-1}(z)\| \leq \frac{\|w - z\|}{1 - c} \end{aligned}$$

desse modo g^{-1} é uma função lipschitziana e portanto contínua. $\square$

Definição 2.7. Seja $(x_i)_{i \in \mathbb{N}}$ uma sequência em $\mathbb{K}^n$. Iremos definir a série $\sum_{i=0}^{\infty} x_i$ como sendo a sequência a própria $(x_i)_{i \in \mathbb{N}}$. Entretanto, dizemos que a série $\sum_{i=0}^{\infty} x_i$ é convergente quando existir o limite $\lim_{n \rightarrow \infty} \sum_{i=0}^n x_i$ e divergente caso o contrário. Quando a série for convergente, definimos o valor da série como

$$\sum_{i=0}^{\infty} x_i = \lim_{n \rightarrow \infty} \sum_{i=0}^n x_i.$$

Quando $\sum_{i=0}^{\infty} \|x_i\|$ for uma série convergente, dizemos que $\sum_{i=0}^{\infty} x_i$ é uma série absolutamente convergente.

Teorema 2.8. *Se uma série for absolutamente convergente, então ela é uma série convergente.*

Demonstração. Mostraremos que $\sum_{i=0}^n x_i$ forma uma sequência de Cauchy e portanto convergem.

Seja $S = \sum_{i=0}^{\infty} \|x_i\|$. Tomemos $n, m \in \mathbb{N}$. Suponhamos sem perda de generalidade que $n < m$, assim

$$\left\| \sum_{i=0}^m x_i - \sum_{i=0}^n x_i \right\| = \left\| \sum_{i=n+1}^m x_i \right\| \leq \sum_{i=n+1}^m \|x_i\| \leq \sum_{i=n+1}^{\infty} \|x_i\| \leq \sum_{i=0}^{\infty} \|x_i\| - \sum_{i=0}^n \|x_i\|.$$

Portanto temos

$$0 \leq \left\| \sum_{i=0}^m x_i - \sum_{i=0}^n x_i \right\| \leq S - \sum_{i=0}^n \|x_i\|.$$

Tomando o limite quando $n \rightarrow \infty$ e usando o teorema do confronto está provado o que queríamos. $\square$

Proposição 2.9. *Se $\sum_{i=0}^{\infty} x_i$ e $\sum_{i=0}^{\infty} y_i$ são séries convergentes, então $\sum_{i=0}^{\infty} x_i + \sum_{i=0}^{\infty} y_i = \sum_{i=0}^{\infty} (x_i + y_i)$.*

Demonstração. Da definição temos

$$\sum_{i=0}^{\infty} x_i + \sum_{i=0}^{\infty} y_i = \lim_{n \rightarrow \infty} \sum_{i=0}^n x_i + \lim_{n \rightarrow \infty} \sum_{i=0}^n y_i = \lim_{n \rightarrow \infty} \left(\sum_{i=0}^n x_i + \sum_{i=0}^n y_i \right) = \lim_{n \rightarrow \infty} \sum_{i=0}^n (x_i + y_i) = \sum_{i=0}^{\infty} (x_i + y_i)$$

$\square$

Lema 2.10. *Seja $A \in M_{n \times p}(\mathbb{K})$ e $B \in M_{p \times m}(\mathbb{K})$. Então $\|AB\| \leq \|A\| \|B\|$.*

Apesar de já termos uma versão dessa desigualdade para a norma espectral, é interessante ter uma versão também para a norma euclidiana.

Demonstração. Temos

$$\|AB\|^2 = \sum_{i=1}^n \sum_{j=1}^m |[AB]_{i,j}|^2 = \sum_{i=1}^n \sum_{j=1}^m \left| \sum_{k=1}^p [A]_{i,k} [B]_{k,j} \right|^2.$$

Pela desigualdade de Cauchy-Schwarz, temos

$$\left| \sum_{k=1}^p [A]_{i,k} [B]_{k,j} \right|^2 \leq \left(\sum_{k=1}^p |[B]_{k,j}|^2 \right) \left(\sum_{k=1}^p |[A]_{i,k}|^2 \right).$$

Desse modo

$$\|AB\|^2 \leq \sum_{i=1}^n \sum_{j=1}^m \left(\sum_{k=1}^p |[B]_{k,j}|^2 \right) \left(\sum_{k=1}^p |[A]_{i,k}|^2 \right).$$

Calculando a expressão a direita

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^m \left(\sum_{k=1}^p |[A]_{i,k}|^2 \right) \left(\sum_{k=1}^p |[B]_{k,j}|^2 \right) &= \sum_{i=1}^n \left(\sum_{k=1}^p |[A]_{i,k}|^2 \right) \sum_{j=1}^m \left(\sum_{k=1}^p |[B]_{k,j}|^2 \right) \\ &= \sum_{i=1}^n \left(\sum_{k=1}^p |[A]_{i,k}|^2 \right) \|B\|^2 = \|B\|^2 \sum_{i=1}^n \left(\sum_{k=1}^p |[A]_{i,k}|^2 \right) \\ &= \|B\|^2 \|A\|^2. \end{aligned}$$

Assim $\|AB\|^2 \leq \|B\|^2 \|A\|^2$ e portanto $\|AB\| \leq \|B\| \|A\|$, o que conclui a demonstração. $\square$

O teorema a seguir e sua demonstração foram adaptados de [3], onde esse resultado é provado para séries numéricas.

Teorema 2.11. *Sejam $\sum_{i=0}^{\infty} A_i$ e $\sum_{i=0}^{\infty} B_i$ séries convergentes em $M_{n \times p}(\mathbb{K})$ e $M_{p \times m}(\mathbb{K})$, respectivamente, e ao menos uma delas é absolutamente convergente. Então $\left(\sum_{i=0}^{\infty} A_i \right) \left(\sum_{i=0}^{\infty} B_i \right) = \sum_{n=0}^{\infty} \left(\sum_{i=0}^n A_{n-i} B_i \right)$.*

Demonstração. Consideremos as somas parciais $\left(\sum_{n=0}^m \sum_{i=0}^n A_{n-i} B_i \right) = \sum_{n=0}^m A_{m-n} \left(\sum_{i=0}^n B_i \right)$. Por conveniência, definimos as matrizes $A = \sum_{i=0}^{\infty} A_i$ e $B = \sum_{i=0}^{\infty} B_i$. Dessa maneira

$$\begin{aligned} &\sum_{n=0}^m A_{m-n} \left(\sum_{i=0}^n B_i \right) - AB \\ &= \sum_{n=0}^m A_{m-n} \left(\sum_{i=0}^n B_i \right) - \left(\sum_{n=0}^m A_{m-n} \right) B + \left(\sum_{n=0}^m A_{m-n} \right) B - AB \\ &= \sum_{n=0}^m A_{m-n} \left(-B + \sum_{i=0}^n B_i \right) - \left(A - \sum_{n=0}^m A_{m-n} \right) B. \end{aligned}$$

Desse modo

$$\begin{aligned} &\left\| \left(\sum_{n=0}^m \sum_{i=0}^n A_{n-i} B_i \right) - AB \right\| \leq \left\| \sum_{n=0}^m A_{m-n} \left(-B + \sum_{i=0}^n B_i \right) \right\| + \left\| \left(\sum_{n=m+1}^{\infty} A_n \right) B \right\| \\ \Rightarrow &\left\| \left(\sum_{n=0}^m \sum_{i=0}^n A_{n-i} B_i \right) - AB \right\| \leq \sum_{n=0}^m \left(\|A_{m-n}\| \left\| B - \sum_{i=0}^n B_i \right\| \right) + \|B\| \left\| \sum_{n=m+1}^{\infty} A_n \right\|. \quad (5) \end{aligned}$$

Suponhamos, sem perda de generalidade, que a A seja absolutamente convergente, então podemos para cada $\varepsilon > 0$ encontrar um $K \in \mathbb{N}$ de modo que para todo $n \geq K$ temos

$$\sum_{i=n}^{\infty} \|A_i\| < \varepsilon \quad \text{e} \quad \left\| B - \sum_{i=0}^n B_i \right\| < \varepsilon.$$

Tomando $m > 2K$ e substituindo no lado direito de (5) obtemos

$$\begin{aligned}
& \left\| \left(\sum_{n=0}^m \sum_{i=0}^n A_{n-i} B_i \right) - AB \right\| \\
& \leq \sum_{n=0}^K \left(\|A_{m-n}\| \left\| B - \sum_{i=0}^n B_i \right\| \right) + \sum_{n=K+1}^m \left(\|A_{m-n}\| \left\| B - \sum_{i=0}^n B_i \right\| \right) + \|B\| \left\| \sum_{n=m+1}^{\infty} A_n \right\| \\
& = \sum_{j=m-K}^m \left(\|A_j\| \left\| B - \sum_{i=0}^{m-j} B_i \right\| \right) + \sum_{n=K+1}^m \left(\|A_{m-n}\| \left\| B - \sum_{i=0}^n B_i \right\| \right) + \|B\| \varepsilon \\
& \leq \sum_{j=m-K}^m C \|A_j\| + \sum_{n=K+1}^m \varepsilon \|A_{m-n}\| + \|B\| \varepsilon.
\end{aligned}$$

Justifica-se a passagem da penúltima para a última linha pois $B - \sum_{i=0}^n B_i$ é uma sequência convergente logo limitada e então existe uma constante $C > 0$ tal que para todo $n \in \mathbb{N}$ se tem $\left\| B - \sum_{i=0}^n B_i \right\| < C$ e no segundo somatório temos que $n > K$ e portanto $\left\| B - \sum_{i=0}^n B_i \right\| < \varepsilon$.
Agora

$$\begin{aligned}
\left\| \left(\sum_{n=0}^m \sum_{i=0}^n A_{n-i} B_i \right) - AB \right\| & \leq C \left(\sum_{j=m-K}^m \|A_j\| \right) + \varepsilon \left(\sum_{n=K+1}^m \|A_{m-n}\| \right) + \|B\| \varepsilon \\
& \leq C \left(\sum_{j=K}^{\infty} \|A_j\| \right) + \varepsilon \left(\sum_{j=0}^{m-K-1} \|A_j\| \right) + \|B\| \varepsilon.
\end{aligned}$$

Definamos $Q = \sum_{j=0}^{\infty} \|A_j\|$. Então

$$\left\| \left(\sum_{n=0}^m \sum_{i=0}^n A_{n-i} B_i \right) - AB \right\| \leq C\varepsilon + \varepsilon Q + \|B\| \varepsilon \leq \varepsilon(C + Q + \|B\|)$$

provando a convergência. □

2.2 A exponencial de matrizes

Definição 2.12. Chamamos de função exponencial a função $\exp : M_n(\mathbb{K}) \rightarrow M_n(\mathbb{K})$ definida por $\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!}$. Note que estamos considerando $A^0 = I$ para todo $A \in M_n(\mathbb{K})$.

Como $\sum_{n=0}^{\infty} \frac{\|A\|^n}{n!} = e^{\|A\|}$, então a série que define $\exp$ é absolutamente convergente e portanto a função está bem definida.

Outro fato importante é que $\exp(\mathbf{0}) = I + \sum_{n=1}^{\infty} \frac{\mathbf{0}^n}{n!} = I + \mathbf{0} = I$, ou seja $\exp(\mathbf{0}) = I$.

Estamos cometendo um certo abuso de notação aqui: seria mais correto escrever $\exp_n$, indicando o espaço onde a função está definida. Porém isso poucas vezes é necessário uma vez que costumeiramente usamos a função acompanhada de seu argumento (escrevendo $\exp(X)$ por exemplo), supondo saber em que espaço X reside. Com a função identidade, por exemplo, costumeiramente escrevemos apenas I sem nenhum argumento, e portanto contextos em que matrizes de várias dimensões estão sendo utilizadas faz-se necessário o uso de índices I_n .

Lema 2.13. *Seja $A \in GL_n(\mathbb{K})$. Então para qualquer $B \in M_n(\mathbb{K})$ se tem que $(ABA^{-1})^n = AB^n A^{-1}$ para todo inteiro $n \geq 1$.*

Demonstração. Para $n = 1$ a proposição é imediata. Suponhamos a afirmação válida para n , então temos:

$$(ABA^{-1})^{n+1} = ABA^{-1}(ABA^{-1})^n = ABA^{-1}AB^n A^{-1} = ABB^n A^{-1} = AB^{n+1} A^{-1}$$

$$(ABA^{-1})^{n+1} = AB^{n+1} A^{-1}.$$

□

Proposição 2.14. *Seja $A \in GL_n(\mathbb{K})$. Então para qualquer $B \in M_n(\mathbb{K})$ se tem que $A \exp(B) A^{-1} = \exp(ABA^{-1})$.*

Demonstração. Temos

$$A \left(\sum_{k=0}^{\infty} \frac{B^k}{k!} \right) A^{-1} = A \left(\lim_{m \rightarrow \infty} \sum_{k=0}^m \frac{B^k}{k!} \right) A^{-1}.$$

A função $f(X) = AXA^{-1}$ é contínua e portanto usando a Proposição 1.38 temos

$$\lim_{m \rightarrow \infty} A \left(\sum_{k=0}^m \frac{B^k}{k!} \right) A^{-1} = \lim_{m \rightarrow \infty} \sum_{k=0}^m \frac{AB^k A^{-1}}{k!} = \sum_{k=0}^{\infty} \frac{(ABA^{-1})^k}{k!} = \exp(ABA^{-1}).$$

□

Proposição 2.15. *Sejam $A, B \in M_n(\mathbb{K})$ tais que $AB = BA$. Então $\exp(A+B) = \exp(A) \exp(B)$ e em particular $\exp(A) \exp(B) = \exp(B) \exp(A)$.*

Demonstração. Temos que

$$\exp(A) \exp(B) = \left(\sum_{i=0}^{\infty} \frac{A^i}{i!} \right) \left(\sum_{i=0}^{\infty} \frac{B^i}{i!} \right).$$

Pelo Teorema 2.11 temos que

$$\left(\sum_{i=0}^{\infty} \frac{A^i}{i!} \right) \left(\sum_{i=0}^{\infty} \frac{B^i}{i!} \right) = \sum_{n=0}^{\infty} \sum_{i=0}^n \frac{A^{n-i}}{(n-i)!} \frac{B^i}{i!} = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{i=0}^n \frac{n!}{(n-i)!i!} A^{n-i} B^i.$$

Da hipótese que A e B comutam nós temos

$$\sum_{i=0}^n \frac{n!}{(n-i)!i!} A^{n-i} B^i = (A+B)^n \quad (\text{Identidade Binomial})$$

obtemos assim

$$\left(\sum_{i=0}^{\infty} \frac{A^i}{i!} \right) \left(\sum_{i=0}^{\infty} \frac{B^i}{i!} \right) = \sum_{n=0}^{\infty} \frac{(A+B)^n}{n!}.$$

Da definição de exponencial temos finalmente

$$\exp(A) \exp(B) = \exp(A+B).$$

Como $A+B = B+A$, então $\exp(A+B) = \exp(B+A) \Rightarrow \exp(A) \exp(B) = \exp(B) \exp(A)$. □

Corolário 2.16. *Para todo $A \in M_n(\mathbb{K})$ se tem $\exp(A) \in GL_n(\mathbb{K})$.*

Demonstração. Segue do fato de $A(-A) = -AA$ e assim $\exp(A) \exp(-A) = \exp(0) = I_n$ e portanto $\exp(-A) = \exp(A)^{-1}$. □

3 Diferenciação

3.1 Definição de diferenciabilidade e propriedades gerais

Definição 3.1. Seja $U \subset \mathbb{R}^m$ um conjunto aberto. Definimos a derivada direcional da função $f : U \rightarrow \mathbb{R}^n$ em $p \in U$ na direção $v \in \mathbb{R}^m$, $v \neq \mathbf{0}$, como sendo o limite

$$\frac{\partial f}{\partial v}(p) = \lim_{t \rightarrow 0} \frac{f(p + tv) - f(p)}{t}, \quad (6)$$

quando este existir.

Proposição 3.2. A função $\exp : M_n(\mathbb{R}) \rightarrow GL_n(\mathbb{R})$ possui derivada direcional na origem para todo $X \in M_n(\mathbb{R})$ não nulo e além disso temos $\frac{\partial \exp}{\partial X}(\mathbf{0}) = X$.

Demonstração. Usemos o limite (6) para $f = \exp$ e $p = 0$, assim

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{\exp(tX) - \exp(\mathbf{0})}{t} &= \lim_{t \rightarrow 0} \frac{1}{t} \left[\left(\sum_{n=0}^{\infty} \frac{(tX)^n}{n!} \right) - I \right] &&= \lim_{t \rightarrow 0} \frac{1}{t} \left(\sum_{n=1}^{\infty} \frac{t^n X^n}{n!} \right) \\ &= \lim_{t \rightarrow 0} \sum_{n=1}^{\infty} \frac{t^{n-1} X^n}{n!} &&= \lim_{t \rightarrow 0} \left(X + \sum_{n=2}^{\infty} \frac{t^{n-1} X^n}{n!} \right) \\ &= X + \lim_{t \rightarrow 0} \sum_{n=2}^{\infty} \frac{t^{n-1} X^n}{n!} &&= X + \lim_{t \rightarrow 0} \sum_{n=1}^{\infty} \frac{t^n X^{n+1}}{(n+1)!} \\ &= X + \lim_{t \rightarrow 0} \frac{tX^2}{2} + \lim_{t \rightarrow 0} \sum_{n=2}^{\infty} \frac{t^n X^{n+1}}{(n+1)!} &&= X + \lim_{t \rightarrow 0} t \sum_{n=2}^{\infty} \frac{t^{n-1} X^{n+1}}{(n+1)!} \end{aligned}$$

por se tratar de limite com $t \rightarrow 0$, podemos considerar $|t| < 1$. Desse modo temos que

$$\left\| \sum_{n=2}^{\infty} \frac{t^{n-1} X^{n+1}}{(n+1)!} \right\| \leq \sum_{n=2}^{\infty} \left\| \frac{t^{n-1} X^{n+1}}{(n+1)!} \right\| \leq \sum_{n=2}^{\infty} \frac{|t|^{n-1} \|X\|^{n+1}}{(n+1)!} \leq \sum_{n=2}^{\infty} \frac{\|X\|^{n+1}}{(n+1)!} \leq e^{\|X\|}.$$

Portanto temos que $\lim_{t \rightarrow 0} t \sum_{n=2}^{\infty} \frac{t^{n-1} X^{n+1}}{(n+1)!}$ é o produto de uma função limitada por uma função cujo limite é 0. Então

$$\lim_{t \rightarrow 0} t \sum_{n=2}^{\infty} \frac{t^{n-1} X^{n+1}}{(n+1)!} = \mathbf{0}.$$

Dessa maneira obtemos $\frac{\partial \exp}{\partial X}(\mathbf{0}) = X$. □

Proposição 3.3. Sejam $A, B \in M_n(\mathbb{R})$ tais que $AB = BA$. Então $\frac{\partial \exp}{\partial B}(A) = B \exp(A) = \exp(A)B$. Em particular, se A e B comutam então $\exp(A)$ e B comutam.

Demonstração. Se $AB = BA$ então $A(tB) = (tB)A$ para todo $t \in \mathbb{R}$. Usando a Proposição 2.15 se tem

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{\exp(A + tB) - \exp(A)}{t} &= \lim_{t \rightarrow 0} \frac{\exp(A) \exp(tB) - \exp(A)}{t} &&= \lim_{t \rightarrow 0} \frac{\exp(tB) \exp(A) - \exp(A)}{t} \\ &= \lim_{t \rightarrow 0} \frac{\exp(A)(\exp(tB) - I)}{t} &&= \lim_{t \rightarrow 0} \frac{(\exp(tB) - I) \exp(A)}{t} \\ &= \exp(A) \frac{\partial \exp}{\partial B}(\mathbf{0}) &&= \frac{\partial \exp}{\partial B}(\mathbf{0}) \exp(A) \\ &= \exp(A)B &&= B \exp(A). \end{aligned}$$

□

Exemplo 3.4. Seja $||_N$ uma norma sobre $\mathbb{R}^m$. Então $||_N$ não possui derivada direcional na origem qualquer que seja $v \in \mathbb{R}^m$ não nulo.

Demonstração. Seja $v \in \mathbb{R}^m$, $v \neq \mathbf{0}$. Então, para $t \neq 0$ nós temos

$$\frac{|\mathbf{0} + tv|_N - |\mathbf{0}|_N}{t} = \frac{|tv|_N}{t} = \frac{|t||v|_N}{t}.$$

Estamos então em dois casos:

$$\frac{|t||v|_N}{t} = |v|_N \quad t > 0,$$

$$\frac{|t||v|_N}{t} = -|v|_N \quad t < 0.$$

Como v é não nulo, $|v|_N \neq -|v|_N$ e portanto não existe $\lim_{t \rightarrow 0} \frac{|t||v|_N}{t}$. □

Definição 3.5. Seja a função $f : U \rightarrow \mathbb{R}$ definida em um conjunto aberto $U \subset \mathbb{R}^m$. Definimos a i -ésima derivada parcial de f no ponto $p \in U$, representada por $\frac{\partial f}{\partial x_i}(p)$, como a derivada direcional de f em p na direção na direção e_i , onde e_i é o i -ésimo vetor da base canônica de $\mathbb{R}^m$. Em símbolos

$$\frac{\partial f}{\partial x_i}(p) = \frac{\partial f}{\partial e_i}(p) = \lim_{t \rightarrow 0} \frac{f(p + te_i) - f(p)}{t}.$$

Definição 3.6. Seja $f : U \rightarrow \mathbb{R}^n$ uma função definida em um conjunto aberto $U \subset \mathbb{R}^m$. Dizemos que f é diferenciável em $p \in U$ se existir $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ tal que

$$\lim_{v \rightarrow \mathbf{0}} \left[\frac{f(p + v) - (f(p) + T(v))}{\|v\|} \right] = \mathbf{0}.$$

Dizemos que a transformação linear T é uma diferencial de f no ponto p . Mostraremos no Corolário 3.8 que uma tal transformação linear, se existir, é única e a partir de então diremos que T é a diferencial de f no ponto p .

Proposição 3.7. Sejam um conjunto aberto $U \subset \mathbb{R}^m$, uma função $f : U \rightarrow \mathbb{R}^n$ diferenciável em $a \in U$ e T uma diferencial de f nesse ponto. Então $T(v) = \frac{\partial f}{\partial v}(a)$.

Demonstração. A partir da definição de derivada direcional sabemos que

$$\frac{\partial f}{\partial v}(a) = \lim_{t \rightarrow 0} \left[\frac{f(a + tv) - f(a)}{t} \right].$$

Assim

$$\begin{aligned} \lim_{t \rightarrow 0} \left[\frac{f(a + tv) - f(a)}{t} \right] &= \lim_{t \rightarrow 0} \left[\frac{f(a + tv) - f(a) - T(tv) + T(tv)}{t} \right] \\ &= \lim_{t \rightarrow 0} \left[\frac{f(a + tv) - (f(a) + T(tv))}{t} + \frac{T(tv)}{t} \right] \\ &= \lim_{t \rightarrow 0} \left[\frac{\|tv\|}{t} \frac{f(a + tv) - (f(a) + T(tv))}{\|tv\|} + T\left(\frac{tv}{t}\right) \right] \\ &= \lim_{t \rightarrow 0} \left[\frac{\|tv\|}{t} \frac{f(a + tv) - (f(a) + T(tv))}{\|tv\|} \right] + \lim_{t \rightarrow 0} [T(v)]. \end{aligned}$$

Como f é diferenciável em a , temos

$$\lim_{t \rightarrow 0} \left[\frac{f(a + tv) - (f(a) + T(tv))}{\|tv\|} \right] = \mathbf{0},$$

e uma vez que $\frac{\|tv\|}{t}$ é limitada, temos que $\lim_{t \rightarrow 0} \left[\frac{\|tv\|}{t} \frac{f(a + tv) - (f(a) + T(tv))}{\|tv\|} \right] = \mathbf{0}$. Portanto $\frac{\partial f}{\partial v}(a) = T(v)$. $\square$

Corolário 3.8. *Seja a função $f : U \rightarrow \mathbb{R}^n$, definida em um conjunto aberto $U \subset \mathbb{R}^m$, diferenciável em $a \in U$. A diferencial de f no ponto a é única.*

Daqui em diante, dada uma função $f : U \rightarrow \mathbb{R}^n$ definida em um conjunto aberto $U \subset \mathbb{R}^m$ e diferenciável em $a \in U$, representaremos a diferencial de f em a por $df(a)$, de modo que a igualdade na Proposição 3.7 pode ser escrita como $df(a)(v) = \frac{\partial f}{\partial v}(a)$.

Exemplo 3.9. Através do Exemplo 3.4 e da Proposição 3.7, sabemos que nenhuma norma é diferenciável na origem.

Proposição 3.10. *Sejam as funções $f, g : U \rightarrow \mathbb{R}^n$ definidas em um conjunto aberto $U \subset \mathbb{R}^m$, diferenciáveis em $a \in U$. Então temos que $d(f + g)(a) = df(a) + dg(a)$.*

Demonstração. Para todo vetor $v \neq \mathbf{0}$ nós temos

$$\begin{aligned} & \frac{f(a + v) + g(a + v) - f(a) - g(a) - df(a)(v) - dg(a)(v)}{\|v\|} \\ &= \frac{f(a + v) - f(a) - df(a)(v)}{\|v\|} + \frac{g(a + v) - g(a) - dg(a)(v)}{\|v\|} \end{aligned}$$

o que nos dá

$$\begin{aligned} & \lim_{v \rightarrow \mathbf{0}} \left[\frac{(f + g)(a + v) - (f + g)(a) - (df(a)(v) + dg(a)(v))}{\|v\|} \right] \\ &= \lim_{v \rightarrow \mathbf{0}} \left[\frac{f(a + v) - f(a) - df(a)(v)}{\|v\|} + \frac{g(a + v) - g(a) - dg(a)(v)}{\|v\|} \right] = \mathbf{0}. \end{aligned}$$

Portanto $f + g$ é diferenciável e $d(f + g)(a) = df + dg(a)$. $\square$

Definição 3.11. Sejam $U \subset \mathbb{R}^m$ um conjunto aberto e $f : U \rightarrow \mathbb{R}$ uma função que possui todas as derivadas parciais em cada um dos pontos de seu domínio. Diremos que f é de classe C^1 se para cada $j, 1 \leq j \leq m$, ocorrer de $\frac{\partial f}{\partial x_j}$ ser uma função de classe C^0 , isto é, contínua. De

maneira recursiva, f é de classe C^k com $k \geq 1$ se todas as $\frac{\partial f}{\partial x_j}$ forem de classe C^{k-1} . No caso em que f for de classe C^k para todo $k \in \mathbb{N}$, dizemos que f é de classe C^∞ .

Sejam $U \subset \mathbb{R}^m$ um conjunto aberto e $f : U \rightarrow \mathbb{R}^n$. Diremos que f é de classe C^k quando cada uma das suas n funções coordenadas $f_1, \dots, f_n$ for uma função de classe C^k . Da mesma forma, f é de classe C^∞ quando cada uma das suas n funções coordenadas $f_1, \dots, f_n$ for uma função de classe C^∞ .

Proposição 3.12. *Sejam as funções $f, g : U \rightarrow \mathbb{R}^n$ definidas em um conjunto aberto $U \subset \mathbb{R}^m$, de classe C^k em U . Então $f + g$ é de classe C^k .*

Demonstração. Se as funções f, g forem diferenciáveis de classe C^1 então df e dg são contínuas. Como a soma de funções contínuas é contínua teremos que $d(f + g)$ será contínua e portanto $f + g$ também será de classe C^1 .

Suponha a afirmação verdadeira para k e as funções f, g de classe C^{k+1} , então df e dg são de classe C^k . Como estamos supondo que a soma de funções de classe C^k é uma função de classe C^k , então $d(f + g)$ será de classe C^k e assim $f + g$ será de classe C^{k+1} , provando a proposição. $\square$

Exemplo 3.13. As transformações lineares $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$ são de classe C^∞ e vale para todo $p \in \mathbb{R}^n$ que $dT(p) = T$.

Demonstração. Usando a definição de função diferenciável temos

$$\lim_{v \rightarrow \mathbf{0}} \frac{T(p+v) - T(p) - T(v)}{\|v\|} = \lim_{v \rightarrow \mathbf{0}} \frac{\mathbf{0}}{\|v\|} = \mathbf{0}.$$

As funções coordenadas $T_1, \dots, T_n$ nada mais são que os vetores linha da matriz de T , e assim $\frac{\partial T_i}{\partial x_j} = [T]_{i,j}$. Como sabemos que as funções constantes $f : \mathbb{R} \rightarrow \mathbb{R}$ são de classe C^∞ , isso nos dá que as transformações lineares são de classe C^∞ . $\square$

Exemplo 3.14. Dado $u \in \mathbb{R}^n$, a função translação $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida como $f(v) = v + u$ é de classe C^∞ e para todo $p \in \mathbb{R}^n$ nós temos que $df(p) = I$, onde I é a função identidade.

Demonstração. Novamente, pela definição temos

$$\lim_{v \rightarrow \mathbf{0}} \frac{f(p+v) - f(p) - I(v)}{\|v\|} = \lim_{v \rightarrow \mathbf{0}} \frac{(p+v+u) - (p+u) - v}{\|v\|} = \lim_{v \rightarrow \mathbf{0}} \frac{\mathbf{0}}{\|v\|} = \mathbf{0}.$$

Do mesmo modo que ocorre com as funções lineares, $\frac{\partial f_j}{\partial x_i}$ serão funções constantes e assim as translações são funções de classe C^∞ . $\square$

Proposição 3.15. Seja uma função $f : U \rightarrow \mathbb{R}^n$ definida em um conjunto aberto $U \subset \mathbb{R}^m$, diferenciável em $a \in U$. Então ela é contínua em a .

Demonstração. Do limite

$$\lim_{v \rightarrow \mathbf{0}} \frac{f(a+v) - (f(a) + df(a)(v))}{\|v\|} = \mathbf{0}$$

podemos concluir

$$\lim_{v \rightarrow \mathbf{0}} \frac{\|f(a+v) - f(a) - df(a)(v)\|}{\|v\|} = \mathbf{0}.$$

Isto nos diz que para todo $\epsilon > 0$ existe $\delta_\epsilon > 0$ tal que para $0 < \|v\| < \delta_\epsilon$ nós temos

$$\begin{aligned} \frac{\|f(a+v) - f(a) - df(a)(v)\|}{\|v\|} &< \epsilon \\ \Rightarrow \|f(a+v) - f(a) - df(a)(v)\| &< \epsilon \|v\| \\ \Rightarrow \|f(a+v) - f(a)\| - \|df(a)(v)\| &< \epsilon \|v\| \\ \Rightarrow 0 \leq \|f(a+v) - f(a)\| &< \|df(a)(v)\| + \epsilon \|v\| \\ \Rightarrow 0 \leq \lim_{v \rightarrow \mathbf{0}} \|f(a+v) - f(a)\| &\leq \lim_{v \rightarrow \mathbf{0}} \|df(a)(v)\| + \lim_{v \rightarrow \mathbf{0}} \epsilon \|v\| \\ \Rightarrow 0 \leq \lim_{v \rightarrow \mathbf{0}} \|f(a+v) - f(a)\| &\leq 0 \\ \Rightarrow \lim_{v \rightarrow \mathbf{0}} f(a+v) - f(a) &= \mathbf{0} \Rightarrow \lim_{v \rightarrow \mathbf{0}} f(a+v) = f(a). \end{aligned}$$

Obtendo que f é contínua. $\square$

Teorema 3.16. *Seja a função bilinear $\zeta : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^r$ e funções $f : \mathbb{R}^k \rightarrow \mathbb{R}^n$ e $g : \mathbb{R}^p \rightarrow \mathbb{R}^m$ diferenciáveis em $a \in \mathbb{R}^k$ e $b \in \mathbb{R}^p$, respectivamente. Tomemos a função $h : U \times V \rightarrow \mathbb{R}^r$ dada por $h(x, y) = \zeta(f(x), g(y))$ temos*

$$dh(a, b)(x, y) = \zeta(df(a)(x), g(b)) + \zeta(f(a), dg(b)(y)),$$

onde $(x, y) \in \mathbb{R}^k \times \mathbb{R}^p$.

Demonstração.

$$\begin{aligned} dh(a, b)(x, y) &= \lim_{t \rightarrow 0} \left[\frac{h((a + tx, b + ty) - h(a, b))}{t} \right] \\ &= \lim_{t \rightarrow 0} \left[\frac{\zeta(f(a + tx), g(b + ty)) - \zeta(f(a), g(b))}{t} \right] \\ &= \lim_{t \rightarrow 0} \left[\frac{\zeta(f(a + tx), g(b + ty)) - \zeta(f(a + tx), g(b))}{t} + \frac{\zeta(f(a + tx), g(b)) - \zeta(f(a), g(b))}{t} \right] \\ &= \lim_{t \rightarrow 0} \left[\frac{\zeta(f(a + tx), g(b + ty) - g(b))}{t} + \frac{\zeta(f(a + tx) - f(a), g(b))}{t} \right] \\ &= \lim_{t \rightarrow 0} \left[\zeta \left(f(a + tx), \frac{g(b + ty) - g(b)}{t} \right) + \zeta \left(\frac{f(a + tx) - f(a)}{t}, g(b) \right) \right]. \end{aligned}$$

Como toda forma bilinear é contínua, podemos usar a Proposição 1.38

$$\begin{aligned} dh(a, b)(x, y) &= \zeta \left(\lim_{t \rightarrow 0} f(a + tx), \lim_{t \rightarrow 0} \frac{g(b + ty) - g(b)}{t} \right) + \zeta \left(\lim_{t \rightarrow 0} \frac{f(a + tx) - f(a)}{t}, \lim_{t \rightarrow 0} g(b) \right) \\ &\Rightarrow dh(a, b)(x, y) = \zeta(f(a), dg(b)(y)) + \zeta(df(a)(x), g(b)). \end{aligned}$$

□

Proposição 3.17. *Sejam funções $f : \mathbb{R}^q \rightarrow \mathbb{R}^n$ e $g : \mathbb{R}^p \rightarrow \mathbb{R}^m$ diferenciáveis de classe C^k . Dada a função bilinear $\zeta : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^r$, a função $h : U \times V \rightarrow \mathbb{R}^r$ definida como $h(x, y) = \zeta(f(x), g(y))$ é de classe C^k .*

Demonstração. Se faz de maneira análoga a demonstração da Proposição 3.12. □

Teorema 3.18. *A função $\Upsilon : GL_n(\mathbb{R}) \rightarrow GL_n(\mathbb{R})$ definida por $\Upsilon(X) = X^{-1}$ é diferenciável e de classe C^∞ e temos para todo $A \in GL_n(\mathbb{R})$ que a diferencial $d\Upsilon$ aplicada em A , isto é, a função $d\Upsilon(A) : M_n(\mathbb{R}) \rightarrow M_n(\mathbb{R})$ é da forma $d\Upsilon(A)(X) = -A^{-1}XA^{-1}$.*

Demonstração. Na cláusula de definição diferenciável, exige-se que o domínio da função seja um conjunto aberto, então comecemos demonstrando que $GL_n(\mathbb{R})$ é um conjunto aberto: seja A invertível, então para todo $v \in \mathbb{R}^n$ temos

$$\|v\| = \|A^{-1}A(v)\| \leq |A^{-1}|_e \|A(v)\| \Rightarrow \frac{\|v\|}{|A^{-1}|_e} \leq \|A(v)\|.$$

Tomando $X \in M_n(\mathbb{R})$ com $|X|_e \leq \frac{1}{2|A^{-1}|_e}$, então $-\frac{\|v\|}{2|A^{-1}|_e} \leq -|X|_e \|v\|$. Assim

$$\begin{aligned} \frac{\|v\|}{|A^{-1}|_e} - \frac{\|v\|}{2|A^{-1}|_e} &\leq \|A(v)\| - \|X(v)\| \\ \Rightarrow \frac{\|v\|}{2|A^{-1}|_e} &\leq \|A(v) - X(v)\| = \|(A - X)(v)\|. \end{aligned}$$

Concluindo que $A - X$ é não singular. Logo é invertível e assim $B\left(A, \frac{1}{2|A^{-1}|_e}\right) \subset GL_n(\mathbb{R})$, portanto $GL_n(\mathbb{R})$ é um conjunto aberto.

Ainda do fato de $A - X$ ser não singular, podemos escrever $v = (A - X)^{-1}u$ para um certo u e obtemos

$$\begin{aligned} \frac{\|(A - X)^{-1}u\|}{2|A^{-1}|_e} &\leq \|(A - X)((A - X)^{-1}u)\| \\ \Rightarrow \|(A - X)^{-1}u\| &\leq 2|A^{-1}|_e \|u\|. \end{aligned}$$

Considerando $\|u\| = 1$

$$\|(A - X)^{-1}u\| \leq 2|A^{-1}|_e.$$

Assim obtemos que $\sup_{\|u\|=1} \|(A - X)^{-1}u\| \leq 2|A^{-1}|_e$ e portanto $|(A - X)^{-1}|_e \leq 2|A^{-1}|_e$.

Consideremos agora a função

$$r(X) = (A + X)^{-1} - (A^{-1} - A^{-1}XA^{-1}).$$

Temos que $d\Upsilon(A) = -A^{-1}XA^{-1}$ equivale a $\lim_{X \rightarrow 0} \frac{r(X)}{\|X\|} = \mathbf{0}$. Verifiquemos o valor deste limite.

$$\begin{aligned} (A + X)r(X) &= (A + X)(A + X)^{-1} - (A + X)(A^{-1} - A^{-1}XA^{-1}) \\ \Rightarrow r(X) &= I - A(A^{-1} - A^{-1}XA^{-1}) - X(A^{-1} - A^{-1}XA^{-1}). \end{aligned}$$

Onde I representa a matriz identidade.

$$\begin{aligned} \Rightarrow (A + X)r(X) &= I - I + XA^{-1} - XA^{-1} + XA^{-1}XA^{-1} \\ \Rightarrow (A + X)r(X) &= (XA^{-1})^2 \\ \Rightarrow r(X) &= (A + X)^{-1}(XA^{-1})^2 \\ \Rightarrow |r(X)|_e &= |(A + X)^{-1}(XA^{-1})^2|_e \\ \Rightarrow |r(X)|_e &\leq |(A + X)^{-1}|_e |X|_e^2 |A^{-1}|_e^2. \end{aligned}$$

Como estamos trabalhando com limite de X indo a $\mathbf{0}$, podemos considerar $|X|_e < \frac{1}{2|A^{-1}|_e}$, e nesse caso sabemos que $|(A - X)^{-1}|_e \leq 2|A^{-1}|_e$. Portanto

$$\begin{aligned} |r(X)|_e &\leq 2|A^{-1}|_e^3 |X|_e^2 \\ \Rightarrow \frac{|r(X)|_e}{|X|_e} &\leq 2|A^{-1}|_e^3 |X|_e \\ \Rightarrow 0 &\leq \frac{|r(X)|_e}{|X|_e} \leq 2|A^{-1}|_e^3 |X|_e \\ \Rightarrow 0 &\leq \lim_{X \rightarrow 0} \frac{|r(X)|_e}{|X|_e} \leq \lim_{X \rightarrow 0} 2|A^{-1}|_e^3 |X|_e. \end{aligned}$$

Pela continuidade e não negatividade da norma

$$\begin{aligned} \Rightarrow 0 &\leq \left| \lim_{X \rightarrow 0} \frac{r(X)}{|X|_e} \right|_e \leq 2|A^{-1}|_e^3 \lim_{X \rightarrow 0} |X|_e \\ \Rightarrow 0 &\leq \left| \lim_{X \rightarrow 0} \frac{r(X)}{|X|_e} \right|_e \leq 0 \Rightarrow \lim_{X \rightarrow 0} \frac{r(X)}{|X|_e} = \mathbf{0}. \end{aligned}$$

E isto encerra a demonstração. □

Teorema 3.19 (Regra da cadeia). *Seja uma função $f : U \rightarrow \mathbb{R}^k$ com $U \subset \mathbb{R}^n$ um conjunto aberto com f diferenciável em $a \in U$ e $g : V \rightarrow \mathbb{R}^m$, $V \subset \mathbb{R}^k$ um conjunto aberto com $f(a) \in V$ e g diferenciável em $f(a)$. Então vale a igualdade*

$$d(g \circ f)(a) = dg(f(a)) \circ df(a).$$

Demonstração. Pela definição de diferenciabilidade, o teorema equivale a dizer que o limite a seguir é igual a $\mathbf{0}$

$$\lim_{v \rightarrow \mathbf{0}} \left[\frac{(g \circ f)(a + v) - g(f(a)) - (dg(f(a)) \circ df(a))(v)}{\|v\|} \right].$$

Reescrevamos o limite e calculemos:

$$\begin{aligned} & \lim_{v \rightarrow \mathbf{0}} \left[\frac{g(f(a + v)) - g(f(a)) - dg(f(a))(df(a)(v))}{\|v\|} \right] \\ &= \lim_{v \rightarrow \mathbf{0}} \left[\frac{g(f(a + v)) - g(f(a))}{\|v\|} - \frac{dg(f(a))(df(a)(v))}{\|v\|} \right] \\ &= \lim_{v \rightarrow \mathbf{0}} \left[\frac{g(f(a + v)) - g(f(a))}{\|v\|} - dg(f(a)) \left(\frac{df(a)(v)}{\|v\|} \right) \right] \\ &= \lim_{v \rightarrow \mathbf{0}} \left[\frac{g(f(a + v)) - g(f(a))}{\|v\|} - dg(f(a)) \left(\frac{f(a + v) - f(a) + df(a)(v)}{\|v\|} \right) \right] \\ &= \lim_{v \rightarrow \mathbf{0}} \left[\frac{g(f(a + v)) - g(f(a))}{\|v\|} - dg(f(a)) \left(\frac{f(a + v) - f(a)}{\|v\|} + \frac{df(a)(v) + f(a) - f(a + v)}{\|v\|} \right) \right] \\ &= \lim_{v \rightarrow \mathbf{0}} \left[\frac{g(f(a + v)) - g(f(a))}{\|v\|} - \frac{dg(f(a))(f(a + v) - f(a))}{\|v\|} \right. \\ & \quad \left. - dg(f(a)) \left(\frac{df(a)(v) + f(a) - f(a + v)}{\|v\|} \right) \right] \\ &= \lim_{v \rightarrow \mathbf{0}} \left[\frac{g(f(a + v)) - g(f(a)) - dg(f(a))(f(a + v) - f(a))}{\|v\|} \right] \end{aligned} \tag{7}$$

$$+ \lim_{v \rightarrow \mathbf{0}} \left[dg(f(a)) \left(\frac{f(a + v) - (df(a)(v) + f(a))}{\|v\|} \right) \right]. \tag{8}$$

Calculemos cada um dos dois limites acima separadamente. Primeiro o limite em (8).

$$\lim_{v \rightarrow \mathbf{0}} \left[dg(f(a)) \left(\frac{f(a + v) - (df(a)(v) + f(a))}{\|v\|} \right) \right]$$

como $dg(f(a))$ é linear, logo é contínua e a expressão acima é igual a

$$dg(f(a)) \left(\lim_{v \rightarrow \mathbf{0}} \frac{f(a + v) - (df(a)(v) + f(a))}{\|v\|} \right)$$

como f é diferenciável por hipótese, o limite acima é $\mathbf{0}$ e assim

$$dg(f(a))(\mathbf{0}) = \mathbf{0}$$

Agora, sejam os conjuntos $U_0 = \{v | f(p) = f(a)\}$ e $U_1 = \{v | f(p) \neq f(a)\}$. Se $a \neq \bar{U}_1$ então o limite (7) é igual 0. Caso contrário, considere a função $\bar{f} : U_1 \rightarrow \mathbb{R}^k$, $\bar{f}(x) = f(x)$ para todo $x \in U_1$, ou seja, $\bar{f}$ é a restrição da função f aos pontos p tais que $f(p) \neq f(a)$. Podemos escrever $p = a + v$ para algum vetor v . Desse modo nós temos

$$\begin{aligned} & \lim_{v \rightarrow 0} \left[\frac{\|\bar{f}(a+v) - f(a)\|}{\|\bar{f}(a+v) - f(a)\|} \frac{g(\bar{f}(a+v)) - g(f(a)) - dg(f(a))(\bar{f}(a+v) - f(a))}{\|v\|} \right] \\ &= \lim_{v \rightarrow 0} \left[\frac{\|\bar{f}(a+v) - f(a)\|}{\|v\|} \frac{g(\bar{f}(a+v)) - g(f(a)) - dg(f(a))(\bar{f}(a+v) - f(a))}{\|\bar{f}(a+v) - f(a)\|} \right] \\ &= \lim_{v \rightarrow 0} \left[\frac{\|\bar{f}(a+v) - f(a)\|}{\|v\|} \right] \lim_{v \rightarrow 0} \left[\frac{g(\bar{f}(a+v)) - g(f(a)) - dg(f(a))(\bar{f}(a+v) - f(a))}{\|\bar{f}(a+v) - f(a)\|} \right] \\ &= \left\| \lim_{v \rightarrow 0} \frac{\bar{f}(a+v) - f(a)}{\|v\|} \right\| \lim_{v \rightarrow 0} \left[\frac{g(\bar{f}(a+v)) - g(f(a)) - dg(f(a))(\bar{f}(a+v) - f(a))}{\|\bar{f}(a+v) - f(a)\|} \right]. \end{aligned}$$

Consideremos agora uma mudança de variáveis $u(v) = \bar{f}(a+v) - f(a)$ e usando a Proposição 1.37 nós temos

$$= \|\mathbf{0}\| \lim_{v \rightarrow 0} \left[\frac{g(u(v) + f(a)) - g(f(a)) - dg(f(a))(u(v))}{\|u(v)\|} \right] = 0 \|\mathbf{0}\| = 0$$

Desse modo, se a for um ponto isolado de U_0 então o limite em (7) é igual 0.

Agora se tivermos $a \in \bar{U}_1 \cap \bar{U}_0 - \{a\}$, então pela definição de limite temos que para todo $\varepsilon > 0$ existirá um $\delta > 0$ tal que

$$p \in U_1 \text{ e } \|p - a\| < \delta \Rightarrow \|f(p) - f(a)\| < \varepsilon$$

por outro lado, temos também

$$p \in U_0 \text{ e } \|p - a\| < \delta \Rightarrow \|f(p) - f(a)\| = 0 < \varepsilon$$

como $B(a, \delta) \subset U_0 \cup U_1$, temos assim

$$\|p - a\| < \delta \Rightarrow \|f(p) - f(a)\| < \varepsilon$$

Desse modo, como o nosso limite original equivale a soma dos limites (7) e (8), ambos iguais a 0, ele próprio é igual 0 e assim $d(g \circ f)(a) = dg(f(a)) \circ df(a)$, como queríamos. $\square$

Corolário 3.20. *Sejam conjuntos abertos $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$ e funções $f : U \rightarrow V$, $g : V \rightarrow \mathbb{R}^p$ diferenciáveis de classe C^k . Então a função $g \circ f$ é diferenciável de classe C^k .*

Demonstração. A demonstração é semelhante as demonstrações das proposições 3.12 e 3.17. $\square$

Teorema 3.21 (Teorema da Função Inversa). *Sejam $U \subset \mathbb{R}^n$ um conjunto aberto e $f : U \rightarrow \mathbb{R}^n$ uma função de classe C^k com $k \geq 1$ em U e $a \in U$ com $df(a)$ invertível. Então existem conjuntos abertos V, W com $f(a) \in V$ tal que f^{-1} está definida em V e $W = f^{-1}(V)$, valendo a igualdade $df^{-1}(f(x)) = df(x)^{-1}$ para todo $x \in W$. Além disso f^{-1} é de classe C^k em V .*

Demonstração. Pela definição de diferenciabilidade e pelo fato de $k \geq 1$, para cada $\varepsilon > 0$ existe um $\delta > 0$ tal que $B(a, \delta) \subset U$ e verifica-se que para todo $x, y \in B(a, \delta)$ se tem

$$\|f(x) - (f(y) + df(y)(x - y))\| \leq \varepsilon \|x - y\| \text{ e } |df(a) - df(y)|_e \leq \varepsilon.$$

Desenvolvendo obtemos

$$|df(a) - df(y)|_e \|x - y\| \leq \varepsilon \|x - y\|$$

$$\begin{aligned}
&\Rightarrow \|df(a)(x - y) - df(y)(x - y)\| \leq \varepsilon \|x - y\| \\
&\Rightarrow \|f(x) - (f(y) + df(y)(x - y))\| + \|df(a)(x - y) - df(y)(x - y)\| \leq \varepsilon \|x - y\| + \varepsilon \|x - y\| \\
&\Rightarrow \|f(x) - f(y) - df(a)(x - y)\| \leq 2\varepsilon \|x - y\|.
\end{aligned}$$

Escolhendo $\varepsilon = \frac{1}{4|df(a)^{-1}|_e}$ obtemos

$$\begin{aligned}
&\|f(x) - (f(y) + df(a)(x - y))\| \leq \frac{\|x - y\|}{2|df(a)^{-1}|_e} \\
&\Rightarrow |df(a)^{-1}|_e \|f(x) - f(y) + df(a)(y - x)\| \leq \frac{\|x - y\|}{2} \\
&\Rightarrow \|df(a)^{-1}(f(x) - df(a)(x)) - df(a)^{-1}(f(y) - df(a)(y))\| \leq \frac{\|x - y\|}{2}.
\end{aligned}$$

Obtemos portanto que a função $\varphi(x) = df(a)^{-1}(f(x) - df(a)(x))$ é uma contração em $B(a, \delta_0)$, para algum $\delta_0 > 0$ associado a $\varepsilon = \frac{1}{4|df(a)^{-1}|_e}$. Usando o Lema 2.6 temos que $\Phi(x) = x + \varphi(x)$ definida em $B(a, \delta_0)$ é contínua, possui inversa contínua e $\Phi(B(a, \delta_0))$ é um conjunto aberto. Da continuidade de $df^{-1}(a)$ (pois é uma transformação linear) temos que $(df(a) \circ \Phi)(B(a, \delta_0))$ também é um conjunto aberto, desse modo

$$\begin{aligned}
(df(a) \circ \Phi)(x) &= df(a)[x + \varphi(x)] \\
&= df(a)[x + df(a)^{-1}(f(x) - df(a)(x))] \\
&= df(a)(x) + f(x) - df(a)(x) = f(x)
\end{aligned}$$

e obtemos $f(B(a, \delta_0))$ é um conjunto aberto. Chamaremos $B(a, \delta_0)$ de W e $f(B(a, \delta_0))$ de V . Da igualdade $f|_W = df(a) \circ \Phi|_W$ obtemos $f^{-1}|_V = \Phi^{-1}|_V \circ df(a)^{-1}$, mostrando que f^{-1} existem quando restrita a V e além disso ela é contínua pois é a composição de funções contínuas.

Aplicando a regra da cadeia e considerando $x \in W$ obtemos

$$\begin{aligned}
dI(x) &= d(f \circ f^{-1})(x) \\
I &= df^{-1}(x) \circ df(f(x)) \\
df^{-1}(x) &= df(f(x))^{-1}.
\end{aligned}$$

Por fim, sabemos que a função inversão de matrizes, representemos ela por Υ , é uma função de classe C^∞ . Como f^{-1} é diferenciável, sabemos que ela é de classe C_0 . Suponha que f^{-1} seja de classe C_p com $p \leq k - 1$. Como df é uma função C^{k-1} , logo também é de classe C^p , então

$$df^{-1} = \Upsilon \circ df \circ f^{-1}$$

também é de classe C_p e portanto f^{-1} é de classe $p + 1$. Desse modo df^{-1} é de classe C^p para todo $p \leq k - 1$, logo df^{-1} é de classe C^{k-1} . Assim f^{-1} é de classe C^k .

Uma observação importante é que não podemos ter f^{-1} de classe C^m com $m > k$ e f apenas de classe C^k , pois nesse caso poderíamos considerar a função $g = f^{-1}$ e então g^{-1} teria que ser ao menos de classe C^m . Como $g^{-1} = f$ e f não é de classe C^m , temos um absurdo. $\square$

Exemplo 3.22. O teorema da função inversa não dá condições suficientes para que uma função seja inversível na vizinhança de um ponto, mas essas condições não são necessárias, como podemos verificar através da função

$$\Psi(x) = \begin{cases} 4x + x^2 \sin\left(\frac{1}{x}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

no intervalo $(-1,1)$. Esta função é C^∞ em $\mathbb{R} - \{0\}$. Afirmamos que esta função possui derivada na origem e ela é igual a 4. Mostremos esse fato:

$$\lim_{x \rightarrow 0} \frac{[4x + x^2 \sin(x^{-1})] - \Psi(0)}{x} = \lim_{x \rightarrow 0} 4 + x \sin(x^{-1}) = 4.$$

Desse modo Ψ é diferenciável em toda reta e por isso sabemos que ela é C^0 . Então a partir do teorema do valor intermediário [3] temos que $[\Psi(-1), \Psi(1)] \subset \Psi[(-1, 1)]$. Mostremos que ela é uma função estritamente crescente: seja $x \in (-1, 1)$ diferente de 0, então:

$$\begin{aligned} \frac{d}{dx}[4x + x^2 \sin(x^{-1})] &= 4 + 2x \sin(x^{-1}) + x^2 \cos(x^{-1}) \frac{-1}{x^2} \\ \Rightarrow |\Psi(x)| &\geq 4 - |2x \sin(x^{-1})| - |\cos(x^{-1})| \\ \Rightarrow |\Psi(x)| &\geq 4 - 2|x| |\sin(x^{-1})| - 1 \\ \Rightarrow |\Psi(x)| &\geq 3 - 2 > 0. \end{aligned}$$

E portanto a derivada é estritamente positiva em $[-1,1]$, de modo que $x < y \Rightarrow \Psi(x) < \Psi(y)$ para $x, y \in [-1, 1]$. Obtemos que Ψ resita a $[-1,1]$ é uma função contínua, injetiva e sobrejetora sobre o intervalo $[\Psi(-1), \Psi(1)]$, ou seja, é uma bijeção contínua entre $[-1,1]$ e $[\Psi(-1), \Psi(1)]$, e mais do que isso: possui inversa contínua, pois

$$\frac{d\Psi^{-1}}{dt}(x) = \left(\frac{d\Psi}{dt}(\Psi^{-1}(x)) \right)^{-1}.$$

Portanto Ψ^{-1} é diferenciável e desse modo é contínua. Ou seja, Ψ possui quase todas as propriedades que estão na *conclusão* do teorema da função inversa, mas não cumpre uma das *hipóteses* mais importantes: a continuidade da função derivada.

Para ver isso, consideremos o conjunto $D = \left\{ \frac{1}{2k\pi} \mid k \in \mathbb{N} - \{0\} \right\}$. Calculando $\frac{d\Psi}{dx}$ nos pontos de D obtemos

$$\begin{aligned} \frac{d\Psi}{dx} \left(\frac{1}{2k\pi} \right) &= 4 + \frac{2 \sin(2k\pi)}{2k\pi} - \cos(2k\pi) \\ \frac{d\Psi}{dx} \left(\frac{1}{2k\pi} \right) &= 4 + \frac{0}{2k\pi} - 1 = 3. \end{aligned}$$

Desse modo, para todo $\lambda \in D$ temos que $\frac{d\Psi}{dx}(\lambda) = 3$. Entretanto, sabemos que para todo $\delta > 0$ existe $\lambda \in D$ com $0 < \lambda < \delta$. Logo $\frac{d\Psi}{dx}$ é descontínua em 0 uma vez que para qualquer $\delta > 0$ existe λ com $|\lambda - 0| < \delta$ e $\left| \frac{d\Psi}{dx}(\lambda) - \frac{d\Psi}{dx}(0) \right| = 3$.

Por outro lado, se considerarmos a função $\Lambda : \mathbb{R} \rightarrow \mathbb{R}$ dada por $\Lambda(x) = \Psi(x) - \frac{7x}{2}$. Também temos que Λ é diferenciável em toda a reta e sua derivada é descontínua em 0. Temos que $\Lambda(0) = 0$. Afirmamos que para $x \in \left(0, \frac{1}{2}\right)$ temos $\Lambda(x) > 0$, pois:

$$\Lambda(x) = \frac{x}{2} + x^2 \sin\left(\frac{1}{x}\right) = x \left(\frac{1}{2} + x \sin\left(\frac{1}{x}\right) \right),$$

como x é positivo, ele não altera o sinal da expressão dentro do parêntesis.

$$\frac{1}{2} + x \sin\left(\frac{1}{x}\right) \geq \frac{1}{2} - x > \frac{1}{2} - \frac{1}{2} = 0,$$

mostrando a positividade de $\Lambda(x)$ para $0 < x < \frac{1}{2}$. Verifiquemos o valor de $\frac{d\Lambda}{dx}$ nos elementos de D .

$$\begin{aligned}\frac{d\Lambda}{dx} \left(\frac{1}{2k\pi} \right) &= \frac{1}{2} + \frac{2 \sin(2k\pi)}{2k\pi} - \cos(2k\pi) \\ \frac{d\Lambda}{dx} \left(\frac{1}{2k\pi} \right) &= \frac{1}{2} + \frac{0}{2k\pi} - 1 = -\frac{1}{2}.\end{aligned}$$

Para todo $x \in D$, como $\frac{d\Lambda}{dx}(x) < 0$ existe um $\varepsilon_x > 0$ satisfazendo

$$\Lambda(t) > \Lambda(x) \quad \forall t \in (x - \varepsilon_x, x).$$

Observe que para $\delta \in (0, \varepsilon_x)$ temos $(x - \delta, x) \subset (x - \varepsilon_x, x)$, o que nos dá

$$\Lambda(t) > \Lambda(x) \quad \forall t \in (x - \delta, x).$$

Podemos considerar $\delta < x$. Desse modo existirá $t \in (0, x)$ com $\Lambda(t) > \Lambda(x)$. Se considerarmos $x \in D \cap \left(0, \frac{1}{2}\right)$, temos

$$\Lambda(0) < \Lambda(x) < \Lambda(t)$$

Pelo Teorema do Valor Intermediário existe $y \in (0, t)$ com $\Lambda(y) = \Lambda(x)$. Como $y < t < x \Rightarrow y \neq x$ e obtemos que Λ não é injetiva em nenhum intervalo contendo 0 e x , com $x \in D \cap \left(0, \frac{1}{2}\right)$. Porém para qualquer $r > 0$ temos que $(0, r) \cap D \neq \emptyset$, de modo que Λ não é inversível em nenhuma vizinhança de 0, mesmo com $\frac{d\Lambda}{dx}(0) = \frac{1}{2}$ sendo um valor inversível.

3.2 Diferenciabilidade da função exponencial

O objetivo desta seção é mostrar que a função $\exp$ é de classe C^∞ . O fato deste resultado receber uma seção própria já deve indicar ao leitor que não se trata de um resultado trivial. A demonstração que apresentamos foi adaptada de [2].

Lema 3.23. *Sejam $|\cdot|_1$ e $|\cdot|_2$ duas normas sobre os espaços $\mathbb{K}^m$ e $\mathbb{K}^n$, respectivamente. A função $|\cdot|_3 : \mathbb{K}^{m+n} \rightarrow \mathbb{R}$ definida por*

$$|(x, y)|_3 = |x|_1 + |y|_2, \quad x \in \mathbb{K}^m \text{ e } y \in \mathbb{K}^n.$$

é uma norma sobre $\mathbb{K}^{m+n}$.

Definição 3.24. Seja uma operação $\mathcal{A} : \mathbb{K}^n \times \mathbb{K}^n \rightarrow \mathbb{K}^n$. Diremos que $\mathcal{A}$ é uma álgebra unital quando $(\mathbb{K}^n, +, \mathcal{A})$ for um anel, onde $+$ é a soma de vetores usual de $\mathbb{K}^n$. Não havendo risco de confusão, usualmente escreveremos uv ou $u \cdot v$ no lugar de $\mathcal{A}(u, v)$.

Teorema 3.25. *Consideremos $\mathbb{R}^n$ munido de uma álgebra unital. Para cada $k \in \mathbb{N}$, consideremos a função $P_k : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida por $P_0(X) = \mathbf{U}$ e $P_k(X) = X \cdot P_{k-1}(X)$, onde $\mathbf{U}$ é o elemento unidade da álgebra. Para todo natural k temos que $\frac{\partial P_{k+1}}{\partial X}(A) = \sum_{i=0}^k P_i(A) X P_{k-i}(A)$.*

Demonstração. A afirmação é verdadeira para $k = 0$, pois nesse caso $P_{0+1}(X) = XU = I(X)$ e $I(X) = X = P_0(A)XP_{0-0}(A) = \sum_{i=0}^0 P_i(A)XP_{0-i}(A)$. Supondo a afirmação verdadeira para $k - 1$ temos $\frac{\partial P_k}{\partial X}(A) = \sum_{i=0}^{k-1} P_i(A)XP_{k-1-i}(A)$. Desse modo

$$\begin{aligned} P_{k+1}(A) &= I(A)P_k(A) \\ \Rightarrow dP_{k+1}(A)(X) &= I(A)dP_k(A)(X) + dI(A)(X)P_k(A) \\ \Rightarrow \frac{\partial P_{k+1}}{\partial X}(A) &= A \left(\sum_{i=0}^{k-1} P_i(A)XP_{k-1-i}(A) \right) + I(X)P_k(A) \\ &= \left(\sum_{i=0}^{k-1} P_{i+1}(A)XP_{k-1-i}(A) \right) + XP_k(A) \\ &= \left(\sum_{j=1}^k P_j(A)XP_{k-j}(A) \right) + XP_k(A) = \sum_{j=0}^k P_j(A)XP_{k-j}(A). \end{aligned}$$

Concluindo a demonstração. $\square$

Quando não houver risco de confusão, escreveremos X^n para nos referirmos a função $P_n(X)$.

Lema 3.26. *Seja o espaço $\mathbb{K}^n$ munido de uma álgebra unital. Definindo para cada $n \in \mathbb{N}$ os conjuntos $\hat{n} = \{i \in \mathbb{N} | 1 \leq i \leq n\}$ e $\mathcal{F}_n = \{\sigma | \sigma : \hat{n} \rightarrow \{0, 1\}\}$, temos:*

$$(A + B)^n = \sum_{\sigma \in \mathcal{F}_n} \prod_{i=1}^n A^{\sigma(i)} B^{1-\sigma(i)}.$$

Demonstração. Notemos inicialmente que $\mathcal{F}_n$ possui 2^n elementos e que para cada $\sigma \in \mathcal{F}_n$ existem duas funções $\sigma_0, \sigma_1 \in \mathcal{F}_{n+1}$ que são extensões de σ , dadas por:

$$\begin{aligned} \sigma_0(i) &= \sigma(i), \quad i \leq n & \text{e} & \quad \sigma_0(n+1) = 0 \\ \sigma_1(i) &= \sigma(i), \quad i \leq n & \text{e} & \quad \sigma_1(n+1) = 1. \end{aligned}$$

Temos que $\mathcal{F}_{n+1} = \bigcup_{\sigma \in \mathcal{F}_n} \{\sigma_0, \sigma_1\}$.

Para $n = 1$, $\mathcal{F}_1 = \{\alpha, \beta\}$ onde $\alpha(1) = 1$ e $\beta(1) = 0$. Sejam $A, B \in \mathbb{K}$. Temos que

$$\begin{aligned} (A + B)^1 &= A + B = A\mathbf{U} + \mathbf{U}B = A^1B^0 + A^0B^1 \\ &= A^{\alpha(1)}B^{1-\alpha(1)} + A^{\beta(1)}B^{1-\beta(1)} = \sum_{\sigma \in \mathcal{F}_1} \prod_{i=1}^1 A^{\sigma(i)} B^{1-\sigma(i)}. \end{aligned}$$

Suponha agora a fórmula válida para n . Então temos

$$(A + B)^{n+1} = (A + B)^n \cdot (A + B) = (A + B)^n A + (A + B)^n B.$$

Podemos então substituir

$$\begin{aligned} (A + B)^{n+1} &= \left(\sum_{\sigma \in \mathcal{F}_n} \prod_{i=1}^n A^{\sigma(i)} B^{1-\sigma(i)} \right) A + \left(\sum_{\sigma \in \mathcal{F}_n} \prod_{i=1}^n A^{\sigma(i)} B^{1-\sigma(i)} \right) B \\ &= \sum_{\sigma \in \mathcal{F}_n} \left(\prod_{i=1}^n A^{\sigma(i)} B^{1-\sigma(i)} \right) A + \sum_{\sigma \in \mathcal{F}_n} \left(\prod_{i=1}^n A^{\sigma(i)} B^{1-\sigma(i)} \right) B \end{aligned}$$

$$= \sum_{\sigma \in \mathcal{F}_n} \left(\prod_{i=1}^{n+1} A^{\sigma_1(i)} B^{1-\sigma_1(i)} \right) + \sum_{\sigma \in \mathcal{F}_n} \left(\prod_{i=1}^{n+1} A^{\sigma_0(i)} B^{1-\sigma_0(i)} \right).$$

Como $\mathcal{F}_{n+1} = \bigcup_{\sigma \in \mathcal{F}_n} \{\sigma_0, \sigma_1\}$, obtemos

$$(A + B)^{n+1} = \sum_{\sigma \in \mathcal{F}_{n+1}} \prod_{i=1}^{n+1} A^{\sigma(i)} B^{1-\sigma(i)},$$

encerrando a demonstração. $\square$

Teorema 3.27. *Seja $\mathbb{K}^n$ munido de uma álgebra unital $(A, B) \mapsto A \cdot B$. Podemos então munir $\mathbb{K}^{2n}$ com uma álgebra unital definida por:*

$$(A, B) \cdot (X, Y) = (A \cdot X, B \cdot X + A \cdot Y).$$

Além disso, se existir uma norma $|\cdot|_1$ sobre $\mathbb{K}^n$ tal que $|A \cdot B|_1 \leq |A|_1 |B|_1$ para todos A, B em $\mathbb{K}^n$, então existe uma norma $|\cdot|_2$ sobre $\mathbb{K}^n$ $|(A, B) \cdot (X, Y)|_2 \leq |(A, B)|_2 |(X, Y)|_2$ para todo A, B, X, Y em $\mathbb{K}_n$, onde $|\cdot|_2$ é definida por $|(A, B)|_2 = |A|_1 + |B|_1$.

Note que $\cdot$ está representando duas operações diferentes, mas como essas operações estão definidas em espaços diferentes isso não causará confusão.

Demonstração. A distributividade, tanto a esquerda quanto a direita, são consequências da função definida ser bilinear. Seja $\mathbf{U} \in \mathbb{K}^n$ o elemento neutro da álgebra sobre $\mathbb{K}^n$, temos:

$$(\mathbf{U}, 0) \cdot (X, Y) = (\mathbf{U} \cdot X, \mathbf{U} \cdot Y + 0 \cdot X) = (X, Y)$$

$$(X, Y) \cdot (\mathbf{U}, 0) = (X \cdot \mathbf{U}, X \cdot 0 + Y \cdot \mathbf{U}) = (X, Y),$$

para quaisquer elementos $X, Y \in \mathbb{K}^n$. Portanto, $(\mathbf{U}, 0)$ é o elemento neutro de $[\mathbb{K}^n, \cdot]$.

Resta mostrar a associatividade. Para melhorar a leitura dos cálculos, escreveremos AB para representar $A \cdot B$. Temos:

$$\begin{aligned} (X_1, X_2)((Y_1, Y_2)(Z_1, Z_2)) &= (X_1, X_2)(Y_1 Z_1, Y_1 Z_2 + Y_2 Z_1) \\ &= (X_1 Y_1 Z_1, X_1 Y_1 Z_2 + X_1 Y_2 Z_1 + X_2 Y_1 Z_1). \end{aligned}$$

Por outro lado

$$\begin{aligned} ((X_1, X_2)(Y_1, Y_2))(Z_1, Z_2) &= (X_1 Y_1, X_1 Y_2 + X_2 Y_1)(Z_1, Z_2) \\ &= (X_1 Y_1 Z_1, X_1 Y_1 Z_2 + X_1 Y_2 Z_1 + X_2 Y_1 Z_1). \end{aligned}$$

Donde obtemos $(X_1, X_2)((Y_1, Y_2)(Z_1, Z_2)) = ((X_1, X_2)(Y_1, Y_2))(Z_1, Z_2)$. Desse modo a associatividade é verdadeira e portanto se trata de uma álgebra unital.

Para mostrar a desigualdade no enunciado nós temos

$$\begin{aligned} |(A, B)(X, Y)|_2 &= |(AX, BX + AY)|_2 = |AX|_1 + |BX + AY|_1 \leq |AX|_1 + |BX|_1 + |AY|_1 \leq \\ &\leq |A|_1 |X|_1 + |B|_1 |X|_1 + |A|_1 |Y|_1 \leq |A|_1 |X|_1 + |B|_1 |X|_1 + |A|_1 |Y|_1 + |B|_1 |Y|_1 \\ &= (|A|_1 + |B|_1)(|X|_1 + |Y|_1) = |(A, B)|_2 |(X, Y)|_2. \end{aligned}$$

Desse modo $|(A, B)(X, Y)|_2 \leq |(A, B)|_2 |(X, Y)|_2$, concluindo a prova. $\square$

Teorema 3.28. *Seja uma sequência de números $(a_n)_{n \in \mathbb{N}}$ tal que $\sum_{n=0}^{\infty} |a_n| x^n$ é convergente para todo $x > 0$. Então dado um espaço $\mathbb{R}^n$ munido de uma álgebra unital e uma norma $|\cdot|_0$ tal que $|AB|_0 \leq |A|_0 |B|_0$. Então a função $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida como $f(X) = \sum_{n=0}^{\infty} a_n X^n$ é de classe C^∞ .*

Demonstração. Primeiro mostraremos que a f é diferenciável. Afirmamos que $df(A)(X) = \sum_{n=1}^{\infty} \left(\sum_{k=1}^n A^{k-1} X A^{n-k} \right)$. Para mostrar isto, consideremos o conjunto

$$\mathcal{F}_n^1 = \{\sigma \in \mathcal{F}_n \mid \text{existe um único elemento } i \in \hat{n} \text{ para o qual } \sigma(i) = 1\}.$$

Note que existe uma bijeção entre $J : \hat{n} \rightarrow \mathcal{F}_n^1$, com cada J_i definida por

$$\begin{aligned} J_i(j) &= 1 & j &= i \\ J_i(j) &= 0 & j &\neq i. \end{aligned}$$

Em particular, $\hat{n}$ e $\mathcal{F}_n^1$ são conjuntos com n . Assim obtemos as igualdades:

$$\sum_{k=1}^n A^{k-1} X A^{n-k} = \sum_{k=1}^n \prod_{i=1}^n X^{J_k(i)} A^{1-J_k(i)} = \sum_{\sigma \in \mathcal{F}_{n+1}^1} \prod_{i=1}^n X^{\sigma(i)} A^{1-\sigma(i)}.$$

Consideremos agora as funções $z_n : \hat{n} \rightarrow \{0, 1\}$ dada por $z_n(i) = 0$ para todo $i \in \hat{n}$. Temos a identidade $A^n = \prod_{i=1}^n X^{z_n(i)} A^{1-z_n(i)}$.

Para que f seja diferenciável em $A \in \mathbb{R}^n$ devemos ter $\lim_{X \rightarrow 0} r(X) = 0$, onde $r : \mathbb{R}^n \rightarrow \mathbb{R}$ é a função:

$$\begin{aligned} r(X) &= \frac{\left| \left(\sum_{n=0}^{\infty} a_n (A + X)^n \right) - \left(\sum_{n=0}^{\infty} a_n A^n \right) - \left(\sum_{n=1}^{\infty} a_n \left(\sum_{k=1}^n A^{k-1} X A^{n-k} \right) \right) \right|_0}{|X|_0} \\ &= \frac{\left| \left(a_0 \mathbf{U} + \sum_{n=1}^{\infty} a_n (A + X)^n \right) - \left(a_0 \mathbf{U} + \sum_{n=1}^{\infty} a_n A^n \right) - \left(\sum_{n=1}^{\infty} a_n \left(\sum_{k=1}^n A^{k-1} X A^{n-k} \right) \right) \right|_0}{|X|_0} \\ &= \frac{\left| \left(\sum_{n=1}^{\infty} a_n (A + X)^n \right) - \left(\sum_{n=1}^{\infty} a_n A^n \right) - \left(\sum_{n=1}^{\infty} a_n \left(\sum_{k=1}^n A^{k-1} X A^{n-k} \right) \right) \right|_0}{|X|_0} \\ &= \frac{\left| \left(\sum_{n=1}^{\infty} a_n \sum_{\sigma \in \mathcal{F}_n} \prod_{i=1}^n X^{\sigma(i)} A^{1-\sigma(i)} \right) - \left(\sum_{n=1}^{\infty} a_n \left(A^n + \sum_{k=1}^n A^{k-1} X A^{n-k} \right) \right) \right|_0}{|X|_0} \\ &= \frac{\left| \left(\sum_{n=1}^{\infty} a_n \sum_{\sigma \in \mathcal{F}_n} \prod_{i=1}^n X^{\sigma(i)} A^{1-\sigma(i)} \right) - \left(\sum_{n=1}^{\infty} a_n \left(\prod_{i=1}^n X^{z_n(i)} A^{1-z_n(i)} + \sum_{\sigma \in \mathcal{F}_n^1} \prod_{i=1}^n X^{\sigma(i)} A^{1-\sigma(i)} \right) \right) \right|_0}{|X|_0}. \end{aligned}$$

Chamemos de H_1^n os conjuntos $\{z_n\} \cup \mathcal{F}_n^1$. Como $\{z_n\} \cap \mathcal{F}_n^1 = \emptyset$, nós obtemos que

$$r(X) = \frac{\left| \left(\sum_{n=1}^{\infty} a_n \sum_{\sigma \in \mathcal{F}_n} \prod_{i=1}^n X^{\sigma(i)} A^{1-\sigma(i)} \right) - \left(\sum_{n=1}^{\infty} a_n \left(\sum_{\sigma \in H_1^n} \prod_{i=1}^n X^{\sigma(i)} A^{1-\sigma(i)} \right) \right) \right|_0}{|X|_0}$$

$$= \frac{\left| \sum_{n=1}^{\infty} a_n \left(\sum_{\sigma \in \mathcal{F}_n} \prod_{i=1}^n X^{\sigma(i)} A^{1-\sigma(i)} - \sum_{\sigma \in H_1^n} \prod_{i=1}^n X^{\sigma(i)} A^{1-\sigma(i)} \right) \right|_0}{|X|_0}.$$

Temos que $\mathcal{F}_1 = H_1^1$, desse modo $\sum_{\sigma \in \mathcal{F}_1} \prod_{i=1}^1 X^{\sigma(i)} A^{1-\sigma(i)} - \sum_{\sigma \in H_1^1} \prod_{i=1}^1 X^{\sigma(i)} A^{1-\sigma(i)} = \mathbf{0}$. Dessa maneira a_1 multiplica um termo nulo, e assim

$$r(X) = \frac{\left| \sum_{n=2}^{\infty} a_n \left(\sum_{\sigma \in \mathcal{F}_n} \prod_{i=1}^n X^{\sigma(i)} A^{1-\sigma(i)} - \sum_{\sigma \in H_1^n} \prod_{i=1}^n X^{\sigma(i)} A^{1-\sigma(i)} \right) \right|_0}{|X|_0}.$$

Chamemos de O_n^2 os conjuntos $\mathcal{F}_n - H_1^n$. Como $H_1^n \subset \mathcal{F}_n$ para todo $n \in \mathbb{N}$ nós obtemos que

$$\begin{aligned} r(X) &= \frac{\left| \sum_{n=2}^{\infty} a_n \left(\sum_{\sigma \in O_n^2} \prod_{i=1}^n X^{\sigma(i)} A^{1-\sigma(i)} \right) \right|_0}{|X|_0} \\ \Rightarrow r(X) &\leq \frac{\sum_{n=2}^{\infty} |a_n| \left(\sum_{\sigma \in O_n^2} \prod_{i=1}^n |X^{\sigma(i)}|_0 |A^{1-\sigma(i)}|_0 \right)}{|X|_0}. \end{aligned} \quad (9)$$

Existe uma bijeção $\mathcal{R} : \mathcal{F}_n \rightarrow \mathcal{P}(\hat{n})$ definida como $\mathcal{R}(\sigma) = \{i \in \hat{n} | \sigma(i) = 1\}$. Em particular, quando aplicamos $\mathcal{R}$ em $O_n^2 = \mathcal{F}_n - (\{z_n\} \cup \mathcal{F}_n^1)$ obtemos os subconjuntos de $\hat{n}$ que possuem dois ou mais elementos. Além disso, como a multiplicação entre números reais é comutativa, temos que $\prod_{i=1}^n |X^{\sigma(i)}|_0 |A^{1-\sigma(i)}|_0 = \prod_{i=1}^n |X^{\varsigma(i)}|_0 |A^{1-\varsigma(i)}|_0$ sempre que $\mathcal{R}(\sigma)$ e $\mathcal{R}(\varsigma)$ possuírem o mesmo número de elementos.

A quantidade de subconjuntos de $\hat{n}$ que possuem k elementos é $\binom{n}{k}$. Portanto a desigualdade (9) implica em

$$r(X) \leq \frac{\sum_{n=2}^{\infty} |a_n| \left(\sum_{k=2}^n \binom{n}{k} |X|_0^k |A|_0^{n-k} \right)}{|X|_0} \Rightarrow r(X) \leq |X|_0 \sum_{n=2}^{\infty} |a_n| \left(\sum_{k=2}^n \binom{n}{k} |X|_0^{k-2} |A|_0^{n-k} \right).$$

Como estamos interessados no limite de X indo para $\mathbf{0}$, podemos considerar $|X|_0 < 1$, e assim

$$\sum_{k=2}^n \binom{n}{k} |X|_0^{k-2} |A|_0^{n-k} \leq \sum_{k=2}^n \binom{n}{k} |A|_0^{n-k} = (1 + |A|_0)^n - n|A|_0^{n-1} - |A|_0^n \leq (1 + |A|_0)^n$$

Assim obtemos

$$r(X) \leq |X|_0 \sum_{n=0}^{\infty} |a_n| (1 + |A|_0)^n.$$

Desse modo

$$\begin{aligned} \lim_{X \rightarrow \mathbf{0}} r(X) &= \lim_{X \rightarrow \mathbf{0}} \left(|X|_0 \sum_{n=0}^{\infty} |a_n| (1 + |A|_0)^n \right) \\ &= \left(\lim_{X \rightarrow \mathbf{0}} |X|_0 \right) \left(\sum_{n=0}^{\infty} |a_n| (1 + |A|_0)^n \right) = 0, \end{aligned}$$

portanto f é diferenciável em todo $A \in \mathbb{R}^n$ e em particular será de classe C^0 .

Demonstramos até o momento que para qualquer espaço $\mathbb{R}^n$ munido de uma álgebra unital, a função $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida por $f(X) = \sum_{n=0}^{\infty} a_n X^n$ é diferenciável, logo, de classe C^0 , onde consideramos que sequência de números $(a_n)_{n \in \mathbb{N}}$ tal que $\sum_{n=0}^{\infty} |a_n| x^n$ é convergente para todo $x > 0$.

Suponhamos então que para qualquer espaço $\mathbb{R}^n$ munido de uma álgebra unital, a função $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida por $f(X) = \sum_{n=0}^{\infty} a_n X^n$ é de classe C^k , para um certo $k \in \mathbb{N}$.

Fixado $\mathbb{R}^m$ e uma álgebra unital, podemos considerar a álgebra unital sobre $\mathbb{R}^{2m}$ dada por $(A_1, A_2)(B_1, B_2) = (A_1 B_1, A_1 B_2 + A_2 B_1)$. Sabemos pelo Teorema 3.27 que esta operação é de fato uma álgebra unital e escolhendo em $\mathbb{R}^{2m}$ a norma $|(A, B)|_1 = |A|_0 + |B|_0$ obtemos $|(A_1 B_1, A_1 B_2 + A_2 B_1)|_1 \leq |(A_1, A_2)|_1 |(B_1, B_2)|_1$, o que nos dá que a função $g : \mathbb{R}^{2m} \rightarrow \mathbb{R}^{2m}$ definida por $g(A, B) = \sum_{n=0}^{\infty} a_n (A, B)^n$ é uma função de classe C^k .

Afirmamos que $(A, B)^n = \left(A^n, \sum_{i=1}^n A^{i-1} B A^{n-i} \right)$ para todo $(A, B) \in \mathbb{R}^{2m}$. A afirmação é verdadeira para $n = 1$. Se for verdadeira para n , então temos

$$\begin{aligned} (A, B)^{n+1} &= (A, B) \left(A^n, \sum_{i=1}^n A^{i-1} B A^{n-i} \right) = \left(A^{n+1}, B A^n + \sum_{i=1}^n A^i B A^{n-i} \right) \\ &\Rightarrow (A, B)^{n+1} = \left(A^{n+1}, \sum_{i=0}^n A^i B A^{n-i} \right) = \left(A^{n+1}, \sum_{j=1}^{n+1} A^{j-1} B A^{n+1-j} \right) \end{aligned}$$

dessa maneira, podemos reescrever g como

$$\begin{aligned} g(A, B) &= \sum_{n=0}^{\infty} a_n \left(A^n, \sum_{i=1}^n A^{i-1} B A^{n-i} \right) = \left(\sum_{n=0}^{\infty} a_n A^n, \sum_{n=0}^{\infty} a_n \sum_{i=1}^n A^{i-1} B A^{n-i} \right) \\ &\Rightarrow g(A, B) = (f(A), df(A)(B)). \end{aligned}$$

Como a função $h : \mathbb{R}^{2n} \rightarrow \mathbb{R}^n$ dada por $h(A, B) = B$ é de classe C^∞ (h é linear), obtemos que df é C^k , pois $df = h \circ g$, donde obtemos as implicações

$$f \text{ é de classe } C^k \Rightarrow df \text{ é de classe } C^k \Rightarrow f \text{ é de classe } C^{k+1}.$$

Donde f é de classe C^∞ . □

Teorema 3.29 (Suavidade da exponencial). A função exponencial $\exp(X) = \sum_{n=0}^{\infty} \frac{X^n}{n!}$ é de classe C^∞ .

Demonstração. A multiplicação de matrizes é uma álgebra unital e a série $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ converge para todo $x > 0$, além do fato de que $\|AB\| \leq \|A\| \|B\|$. Dessa maneira recaímos nas hipóteses do Teorema 3.28. □

Algumas observações se fazem pertinentes: no caminho escolhido para demonstrarmos o Teorema 3.29 não podíamos simplesmente substituir o uso de álgebras unitais pela multiplicação de matrizes pois precisamos considerar o produto $(A, B)(X, Y) \mapsto (AX, AY + BX)$, que não é o produto matricial. Outro caminho possível seria considerar as funções $f_{i,j} : M_n(\mathbb{R}) \rightarrow \mathbb{R}$ dadas por $f_{i,j}(X) = \sum_{k=0}^{\infty} \frac{[X^k]_{i,j}}{k!}$, entretanto para poder tratar com $[X^k]_{i,j}$ precisaríamos lidar com polinômios envolvendo n^2 variáveis e usar resultados combinatórios mais sofisticados do que apenas o tratamento das funções $\sigma : \hat{n} \rightarrow \{0, 1\}$ e a contagem de subconjuntos.

4 Grupos matriciais e suas Álgebras de Lie

4.1 Grupos Matriciais

Definição 4.1. Seja $G \subset GL_n(\mathbb{K})$ um subgrupo. Dizemos que G é um grupo matricial se for um subconjunto fechado de $GL_n(\mathbb{K})$.

Exemplo 4.2. Seja $\det : M_n(\mathbb{K}) \rightarrow \mathbb{K}$ a função determinante. Afirmamos que o conjunto $SL_n(\mathbb{K}) = \{A \in M_n(\mathbb{K}) \mid \det(A) = 1\}$ é um grupo matricial.

Podemos definir $GL_n(\mathbb{K})$ como o conjunto das matrizes $A \in M_n(\mathbb{K})$ com $\det(A) \neq 0$, então $SL_n(\mathbb{K}) \subset GL_n(\mathbb{K})$. Do fato de $\det(AB) = \det(A)\det(B)$ obtemos

$$A, B \in SL_n(\mathbb{K}) \Rightarrow \det(AB) = 1 \times 1 = 1,$$

assim concluímos que $A, B \in SL_n(\mathbb{K}) \Rightarrow AB \in SL_n(\mathbb{K})$.

Temos também que $A \in SL_n(\mathbb{K}) \Rightarrow 1 \det(A^{-1}) = \det(I) \Rightarrow \det(A^{-1}) = 1$ e portanto $A \in SL_n(\mathbb{K}) \Rightarrow A^{-1} \in SL_n(\mathbb{K})$. Desse modo $SL_n(\mathbb{K})$ é um grupo.

Como consequência da definição $SL_n(\mathbb{K}) = \det^{-1}(\{1\})$, sabemos que $\det$ é uma função contínua e $\{1\}$ é um conjunto fechado, e assim $SL_n(\mathbb{K})$ é um conjunto fechado de $M_n(\mathbb{K})$. Como $SL_n(\mathbb{K}) = SL_n(\mathbb{K}) \cap GL_n(\mathbb{K})$, $SL_n(\mathbb{K})$ é um grupo matricial.

Exemplo 4.3. Consideremos o conjunto $M_1(\mathbb{R}) = \mathbb{R}$. A multiplicação de matrizes nesse caso é simplesmente a multiplicação de números reais. Definindo o conjunto $G \subset \mathbb{R}$ como

$$G = \{e^q \mid q \in \mathbb{Q}\},$$

onde $\mathbb{Q}$ é o conjunto dos números racionais. Do fato de $e^q e^r = e^{q+r}$ obtemos que G é fechado para a multiplicação, pois

$$q, r \in \mathbb{Q} \Rightarrow q + r \in \mathbb{Q}.$$

Além disso, é também fechado para a inversão de funções uma vez que

$$(e^q)^{-1} = e^{-q} \text{ e } q \in \mathbb{Q} \Rightarrow -q \in \mathbb{Q},$$

temos ainda que $0 \notin G$. Desse modo, G é um subgrupo de $GL_1(\mathbb{R})$.

Da análise real, sabemos que exponencial é uma função contínua e assim se tivermos uma sequência de números racionais $(a_i)_{i \in \mathbb{N}}$ convergindo para um número irracional a então a sequência e^{a_i} irá convergir para e^a . Entretanto, a função e^x é injetiva e portanto $e^a \neq e^q$ para todo número racional q , portanto $e^a \notin G$ e assim obtemos uma sequência de elementos de G , a sequência e^{a_i} , que converge para um elemento que não está em G mas que está em $GL_1(\mathbb{R})$, portanto G não é um grupo matricial.

4.1.1 Grupos compactos

Definição 4.4. Dizemos que um subgrupo $G \subset GL_n(\mathbb{K})$ é um subgrupo compacto quando for um conjunto compacto de $M_n(\mathbb{K})$. Note que todo subgrupo compacto de $GL_n(\mathbb{K})$ é um grupo matricial, pois sendo fechado em $M_n(\mathbb{K})$ então obrigatoriamente será fechado em $GL_n(\mathbb{K})$ (pois $G = G \cap GL_n(\mathbb{K})$).

Exemplo 4.5. Seja uma matriz $A \in M_n(\mathbb{K})$. Definimos a matriz A^* dada por $[A^*]_{i,j} = \overline{[A]_{j,i}}$. Chamamos A^* de matriz adjunta de A . Podemos então definir o conjunto

$$\mathcal{O}_n(\mathbb{K}) = \{A \in M_n(\mathbb{K}) \mid AA^* = I\}.$$

Notemos que $|x| = |\overline{x}|$. Desse modo

$$\|A\| = \sqrt{\sum_{i=1}^n \sum_{j=1}^n |[A]_{i,j}|^2} = \sqrt{\sum_{j=1}^n \sum_{i=1}^n |[A]_{j,i}|^2} = \sqrt{\sum_{j=1}^n \sum_{i=1}^n |\overline{[A]_{j,i}}|^2} = \|A^*\|.$$

Apesar de não ser linear, temos que $(A+B)^* = A^* + B^*$, portanto $\|A-B\| = \|A^* - B^*\|$ e assim a função $f(A) = A^*$ é lipschitziana, logo contínua. Consequentemente a função $g(A) = AA^*$ é contínua. Obtemos que $\mathcal{O}_n(\mathbb{K}) = g^{-1}(\{I\})$ é um conjunto fechado por $\{I\}$ ser fechado. Como $AA^* = I$, usando a definição de multiplicação de matrizes obtemos:

$$\begin{aligned} 1 = [AA^*]_{i,i} &= \sum_{j=1}^n [A]_{i,j} [A^*]_{j,i} = \sum_{j=1}^n [A]_{i,j} \overline{[A]_{i,j}} = \sum_{j=1}^n |[A]_{i,j}|^2 \Rightarrow \\ &\Rightarrow \sum_{j=1}^n |[A]_{i,j}|^2 = 1. \end{aligned}$$

Portanto

$$\|A\|^2 = \sum_{i=1}^n \sum_{j=1}^n |[A]_{i,j}|^2 = \sum_{i=1}^n 1 = n.$$

Desse modo $A \in \mathcal{O}_n(\mathbb{K})$ implica que $\|A\| = \sqrt{n}$ e portanto $\mathcal{O}_n(\mathbb{K})$ é um conjunto limitado. Sendo limitado e fechado, é um conjunto compacto.

Como todo $A \in \mathcal{O}_n(\mathbb{K})$ implica que $A^{-1} = A^*$, temos de imediato que todo A em $\mathcal{O}_n(\mathbb{K})$ é inversível, portanto $\mathcal{O}_n(\mathbb{K}) \subset GL_n(\mathbb{K})$. Resta mostrar que $\mathcal{O}_n(\mathbb{K})$ é um grupo.

Consideremos a matriz $(A^*)^*$. Temos que $[(A^*)^*]_{i,j} = \overline{[A^*]_{j,i}}$. Entretanto $[A^*]_{j,i} = \overline{[A]_{i,j}}$ e substituindo obtemos $[(A^*)^*]_{i,j} = \overline{\overline{[A]_{i,j}}} = [A]_{i,j}$ e concluímos que $(A^*)^* = A$. Desse modo $AA^* = I \Rightarrow A^*A = I \Rightarrow A^*A^{**} = I$ e assim $A \in \mathcal{O}_n(\mathbb{K}) \Rightarrow A^* \in \mathcal{O}_n(\mathbb{K})$. Obtemos portanto que $A \in \mathcal{O}_n(\mathbb{K}) \Rightarrow A^{-1} \in \mathcal{O}_n(\mathbb{K})$, pois nesse caso $A^* = A^{-1}$.

Sejam agora $A, B \in M_n(\mathbb{K})$, temos que

$$[(AB)^*]_{i,j} = \overline{[AB]_{j,i}} = \overline{\sum_{k=1}^n A_{j,k} B_{k,i}} = \sum_{k=1}^n \overline{A_{j,k}} \overline{B_{k,i}}.$$

Por outro lado, também temos que

$$[B^*A^*]_{i,j} = \sum_{k=1}^n [B^*]_{i,k} [A^*]_{k,j} = \sum_{k=1}^n \overline{[B]_{k,i}} \overline{[A]_{j,k}},$$

obtemos assim que $[B^*A^*]_{i,j} = [(AB)^*]_{i,j}$ e portanto $B^*A^* = (AB)^*$. Sejam agora $A, B \in \mathcal{O}_n(\mathbb{K})$, então

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

Assim $(AB)(AB)^* = I$ e desse modo $A, B \in \mathcal{O}_n(\mathbb{K}) \Rightarrow AB \in \mathcal{O}_n(\mathbb{K})$, concluindo a demonstração de que $\mathcal{O}_n(\mathbb{K})$ é um grupo. Quando $\mathbb{K} = \mathbb{R}$ representamos $\mathcal{O}_n(\mathbb{K})$ por $O(n)$ e o chamamos de grupo das matrizes ortogonais de ordem n , e quando $\mathbb{K} = \mathbb{C}$ representamos $\mathcal{O}_n(\mathbb{K})$ por $U(n)$ e o chamamos de grupo das matrizes unitárias de ordem n .

Exemplo 4.6. Quando $n > 1$, o grupo $SL_n(\mathbb{K})$ não é um conjunto compacto. Para ilustrar esse fato, para cada $t \neq 0$ seja a matriz A_t dada por

$$A_t = \begin{bmatrix} \frac{1}{t} & 0 \\ 0 & t \end{bmatrix}.$$

Temos que $A_t \in SL_2(\mathbb{K})$, entre tanto $\|A_t\| > t$. Portanto $SL_2(\mathbb{K})$ não é um conjunto limitado e não pode ser compacto. Pode-se adaptar a construção das matrizes A_t para todo $SL_n(\mathbb{K})$ com $n \geq 2$, e assim eles não serão conjuntos compactos.

4.2 Álgebras de Lie

Definição 4.7. Chamamos de *caminho* uma função $\gamma : l \rightarrow \mathbb{K}^n$, onde $l \subset \mathbb{R}$ é um intervalo. Se γ for uma função diferenciável (considerando a equivalência entre $\mathbb{C}^n$ e $\mathbb{R}^{2n}$ no caso $\mathbb{K} = \mathbb{C}$), dizemos que γ é um *caminho diferenciável* e definimos o *caminho tangente* a γ , representado por γ' , como o caminho $\gamma' : l \rightarrow \mathbb{K}^n$ definido por

$$\gamma'_i(a) = d\gamma(a)(e_i).$$

Definição 4.8. Seja $p \in U$ com $U \subset \mathbb{R}^n$. Definimos o *espaço tangente* a U no ponto p , representado por $T_p U$ como o conjunto dos vetores v para os quais existe algum $\varepsilon > 0$ e um caminho diferenciável $\gamma : (-\varepsilon, \varepsilon) \rightarrow U$ com $\gamma(0) = p$ e $\gamma'(0) = v$. Em símbolos nós temos

$$T_p U = \{v \mid \exists \gamma : (-\varepsilon, \varepsilon) \rightarrow U \text{ com } \gamma(0) = p \text{ e } \gamma'(0) = v\}.$$

Definição 4.9. A álgebra de Lie $\mathfrak{g}$ de um grupo matricial G é definida como $\mathfrak{g} = T_I G$, onde $T_I G$ é o espaço tangente de G em I .

Teorema 4.10. A álgebra de Lie $\mathfrak{g}$ de um grupo matricial $G \subset GL_n(\mathbb{K})$ é um subespaço vetorial real de $M_n(\mathbb{K})$.

Demonstração. Se $A \in \mathfrak{g}$, então existe um caminho diferenciável $\gamma : (-\varepsilon, \varepsilon) \rightarrow G$ com $\gamma(0) = I$ e $\gamma'(0) = A$. Para cada $k \in \mathbb{R}$, $k \neq 0$ seja o caminho $\beta_k : \left(-\frac{\varepsilon}{|k|}, \frac{\varepsilon}{|k|}\right) \rightarrow G$ definido como $\beta_k(t) = \gamma(kt)$. Temos que $\beta'_k(0) = kA$, $\beta_k(0) = I$ e assim $kA \in \mathfrak{g}$.

Para ver que $0A = \mathbf{0} \in \mathfrak{g}$ basta considerar o caminho $f : (-\varepsilon, \varepsilon) \rightarrow G$ constante $f(t) = I$.

Seja agora $A, B \in G$. Podemos tomar caminhos $\gamma, \lambda : (-\varepsilon, \varepsilon) \rightarrow G$ com $\gamma(0) = \lambda(0) = I$, $\gamma'(0) = A$ e $\lambda'(0) = B$. Definamos o caminho $\mu : (-\varepsilon, \varepsilon) \rightarrow G$ dado por $\mu(t) = \gamma(t)\lambda(t)$. Temos $\mu(0) = I$ e $\mu'(0) = \gamma'(0)\lambda(0) + \gamma(0)\lambda'(0) = AI + IB = A + B$, desse modo $A + B \in \mathfrak{g}$. $\square$

Teorema 4.11. Seja $G \subset GL_n(\mathbb{K})$ um grupo matricial e $\mathfrak{g}$ sua álgebra de Lie. Então $\exp(\mathfrak{g}) \subset G$.

Demonstração. Seja $\{X_i \mid i \in \mathbb{N}, 1 \leq i \leq k\}$ uma base de $\mathfrak{g}$. Então podemos escolher caminhos diferenciáveis $\alpha_i : \mathbb{R} \rightarrow G$ tais que $\alpha_i(0) = I_n$ e $\alpha'_i(0) = X_i$. Definimos a função $F_{\mathfrak{g}} : \mathfrak{g} \rightarrow G$ dada por

$$F_{\mathfrak{g}} \left(\sum_{i=1}^k c_i X_i \right) = \prod_{i=1}^k \alpha_i(c_i).$$

Temos que $F_{\mathfrak{g}}(\mathbf{0}) = I_n$. O cálculo de $dF_{\mathfrak{g}}(\mathbf{0})$ nos dá

$$dF_{\mathfrak{g}}(\mathbf{0})(X_i) = \lim_{t \rightarrow 0} \left[\frac{F_{\mathfrak{g}}(tX_i) - F_{\mathfrak{g}}(\mathbf{0})}{t} \right] = \lim_{t \rightarrow 0} \left[\frac{\alpha_i(t) - I_n}{t} \right] = \alpha'_i(0) = X_i,$$

$$\text{isto é, } dF_{\mathfrak{g}}(\mathbf{0})(X_i) = X_i.$$

Usando a linearidade do operador diferencial temos que $dF_{\mathfrak{g}}(\mathbf{0})(X) = X$ para todo $X \in \mathfrak{g}$.

Escolhamos agora um subespaço vetorial $\mathfrak{p}$ satisfazendo $M_n(\mathbb{K}) = \mathfrak{g} \oplus \mathfrak{p}$ para poder definir a função $F_{\mathfrak{p}} : \mathfrak{p} \rightarrow M_n(\mathbb{K})$ dada por $F_{\mathfrak{p}}(Y) = Y + I_n$.

Seja a função $F : M_n(\mathbb{K}) \rightarrow M_n(\mathbb{K})$ dada por $F(X + Y) = F_{\mathfrak{g}}(X)F_{\mathfrak{p}}(Y)$ com $X \in \mathfrak{g}$ e $Y \in \mathfrak{p}$. Temos $F(\mathbf{0}) = I_n$ e calculando $dF(\mathbf{0})$ em $X + Y$ obtemos

$$\begin{aligned} dF(\mathbf{0})(X + Y) &= \lim_{t \rightarrow 0} \left[\frac{F_{\mathfrak{g}}(tX)F_{\mathfrak{p}}(tY) - I_n}{t} \right] \\ &= \lim_{t \rightarrow 0} \left[\frac{F_{\mathfrak{g}}(tX)F_{\mathfrak{p}}(tY) - F_{\mathfrak{p}}(tY) + F_{\mathfrak{p}}(tY) - I_n}{t} \right] \\ &= \lim_{t \rightarrow 0} \left[\frac{(F_{\mathfrak{g}}(tX) - I_n)F_{\mathfrak{p}}(tY)}{t} \right] + \lim_{t \rightarrow 0} \left[\frac{F_{\mathfrak{p}}(tY) - I_n}{t} \right] \\ &= dF_{\mathfrak{g}}(\mathbf{0})(X)I_n + dF_{\mathfrak{p}}(\mathbf{0})(Y) \\ &= X + Y, \end{aligned}$$

concluindo que $dF(\mathbf{0})$ é o operador identidade de $M_n(\mathbb{K})$.

Cabe uma observação aqui: I_n é o operador identidade de $\mathbb{K}^n$, enquanto que $dF(\mathbf{0})$ é o operador identidade de $M_n(\mathbb{K})$, ou seja, $I_n \neq dF(\mathbf{0})$.

Como $dF(\mathbf{0})$ é bijetiva, então pelo Teorema da Função Inversa existe uma vizinhança U de $I_n = F(\mathbf{0})$ onde F é invertível. Seja $F^{-1} : U \rightarrow M_n(\mathbb{K})$ essa inversa. Definimos as funções u e v dadas por $F^{-1}(a) = u(a) + v(a)$, $u(a) \in \mathfrak{g}$ e $v(a) \in \mathfrak{p}$ para todo $a \in U$.

Um fato importante é que $v(a) = \mathbf{0} \Rightarrow a \in G$ pois $v(a) = \mathbf{0} \Rightarrow a = F(X)$ para algum $X \in \mathfrak{g}$ e portanto $F(X) = F_{\mathfrak{g}}(X)$, porém $\text{Im}(F_{\mathfrak{g}}) \subset G$ e assim $F(X) \in G \Rightarrow a \in G$.

Seja $E \in \mathfrak{g}$. Queremos obter $\exp(sE) \in G$ para todo s suficiente pequeno, e obteremos isso mostrando $v(\exp(sE)) = \mathbf{0}$. Essa última igualdade nos diz que essa função seria constante em uma vizinhança de 0, portanto queremos obter para s pequeno

$$\left. \frac{d}{dt} \right|_{t=s} v(\exp(tE)) = \mathbf{0} \Rightarrow dv(\exp(sE))(E \exp(sE)) = \mathbf{0}. \quad (10)$$

Como s está fixo, simplificaremos a notação tomando $a = \exp(sE)$. Para s pequeno $a \in U$ e assim

$$a = F(X + Y) = F_{\mathfrak{g}}(X)F_{\mathfrak{p}}(Y) \text{ com } X \in \mathfrak{g} \text{ e } Y \in \mathfrak{p} \text{ e desse modo } v(a) = Y.$$

Seja então $W \in \mathfrak{g}$ e t pequeno tal que $F(X + tW + Y) \in U$, então $v(F(X + tW + Y)) = Y$, donde obtemos que em uma vizinhança de 0 essa função é constante e assim

$$\begin{aligned} \left. \frac{d}{dt} \right|_{t=0} v(F(X + tW + Y)) &= \mathbf{0} \\ dv(F(X + Y)) \left(\left. \frac{d}{dt} \right|_{t=0} F(X + tW + Y) \right) &= \mathbf{0} \\ dv(a) \left(\left. \frac{d}{dt} \right|_{t=0} (F_{\mathfrak{g}}(X + tW)F_{\mathfrak{p}}(Y)) \right) &= \mathbf{0} \\ dv(a) \left(\left(\left. \frac{d}{dt} \right|_{t=0} (F_{\mathfrak{g}}(X + tW)) \right) F_{\mathfrak{p}}(Y) \right) &= \mathbf{0} \\ dv(a) (dF_{\mathfrak{g}}(X)(W)F_{\mathfrak{p}}(Y)) &= \mathbf{0}. \end{aligned}$$

Como $a = F_{\mathfrak{g}}(X)F_{\mathfrak{p}}(Y)$ então $(F(X))^{-1}a = F_{\mathfrak{p}}(Y)$

$$dv(a) \left(dF_{\mathfrak{g}}(X)(W) (F(X))^{-1} a \right) = \mathbf{0}.$$

Comparando essa última expressão com (10), obtemos que se mostrarmos $E = dF_{\mathfrak{g}}(X)(W) (F(X))^{-1}$ então teremos $dv(a)(Ea) = \mathbf{0}$.

Inicialmente tomemos o caminho $\beta : \mathbb{R} \rightarrow G$ dado por $\beta(t) = F_{\mathfrak{g}}(X + tW)(F_{\mathfrak{g}}(X))^{-1}$. Temos $\beta(0) = I_n$ e $\beta'(0) = dF_{\mathfrak{g}}(X)(W) (F(X))^{-1}$, então por definição $dF_{\mathfrak{g}}(X)(W) (F(X))^{-1} \in \mathfrak{g}$.

Segundo, podemos definir uma função contínua $\Gamma : \mathfrak{g} \rightarrow \mathcal{L}(\mathfrak{g})$ tal que $\Gamma(X)$ é a função linear dada por

$$\Gamma(X)(W) = dF_{\mathfrak{g}}(X)(W) (F(X))^{-1}$$

é tal que $\Gamma(\mathbf{0})$ é a função identidade de $\mathfrak{g}$ e portanto $\det(\Gamma(\mathbf{0})) = 1$. Dessa maneira, em uma vizinhança V de $\mathbf{0}$ temos que, para todo $X \in V$, $\det(\Gamma(X)) > 0$ e assim $\Gamma(X)$ é um isomorfismo de $\mathfrak{g}$ em si mesmo, logo sobrejetor e podemos então escolher W de modo que $E = \Gamma(X)(W)$.

Então se escolhermos s suficientemente pequeno de modo que $\exp(sE) \in U$ e $u(\exp(sE)) \in V$ obtemos $dv(\exp(sE))(E \exp(sE)) = \mathbf{0}$ e portanto $v(\exp(tE))$ é uma função constante em uma vizinhança de $\mathbf{0}$, e como $v(\exp(0E)) = v(I_n) = \mathbf{0}$, temos que $\exp(tE) \in G$ para t suficientemente pequeno.

Seja $\varepsilon > 0$ tal que $|t| < \varepsilon \Rightarrow \exp(tE) \in G$. Sabemos que existe $n \in \mathbb{N}$ tal que $\frac{1}{n} < \varepsilon$. Desse modo

$$\begin{aligned} \exp(1E) &= \exp \left(\left(n \frac{1}{n} \right) E \right) = \exp \left(n \left(\frac{1}{n} E \right) \right) \\ &= \exp \left(\sum_{i=1}^n \left(\frac{1}{n} E \right) \right) = \prod_{i=1}^n \exp \left(\frac{1}{n} E \right). \end{aligned}$$

Como $\exp \left(\frac{1}{n} E \right) \in G$ e G é fechado para multiplicação de matrizes, $\exp(E) \in G$ para todo $E \in \mathfrak{g}$, encerrando a demonstração. $\square$

Corolário 4.12. *Seja $G \subset GL_n(K)$ com $\mathfrak{g}$ sua álgebra de Lie. Dado $A \in M_n(\mathbb{K})$, definimos o caminho $\gamma(t) = \exp(tA)$. Então $A \in \mathfrak{g}$ se e somente se $\gamma(t) \subset G$.*

5 Grupos Matriciais como variedades diferenciáveis

Definição 5.1. Uma variedade diferenciável de dimensão n é um conjunto M e uma família de funções $X_\alpha : U_\alpha \rightarrow M$ com índices em um conjunto $\mathfrak{A}$ onde cada $U_\alpha \subset \mathbb{R}^n$ é um conjunto aberto e são satisfeitas as seguintes condições:

- (i) $\bigcup_{\alpha \in \mathfrak{A}} X_\alpha(U_\alpha) = M$
- (ii) Para quaisquer α e β da família de índices definimos o seguinte conjunto:
 $W_{\alpha,\beta} = X_\alpha(U_\alpha) \cap X_\beta(U_\beta)$. Temos que os conjuntos $X_\alpha^{-1}(W_{\alpha,\beta})$ e $X_\beta^{-1}(W_{\alpha,\beta})$ são abertos do $\mathbb{R}^n$ e as aplicações $X_\alpha^{-1} \circ X_\beta : X_\beta^{-1}(W_{\alpha,\beta}) \rightarrow \mathbb{R}^n$ e $X_\beta^{-1} \circ X_\alpha : X_\alpha^{-1}(W_{\alpha,\beta}) \rightarrow \mathbb{R}^n$ são funções C^∞
- (iii) A família $\mathfrak{A} = \{(U_\alpha, X_\alpha) \mid \alpha \in \mathcal{A}\}$ é maximal em relação as condições (i) e (ii), isto é, se uma família $\mathfrak{B} = \{(U_\beta, X_\beta) \mid \beta \in \mathcal{B}\}$ satisfaz as condições (i) e (ii) e além disso $\mathfrak{A} \subset \mathfrak{B}$, então $\mathfrak{A} = \mathfrak{B}$.

A demonstração de que grupos matriciais são variedades diferenciáveis é a intenção deste capítulo.

Lema 5.2. *Seja G um grupo matricial e $\mathfrak{g}$ sua álgebra de Lie. Existe um $r > 0$ tal que $B(I_n, r) \cap G \subset \exp(\mathfrak{g})$.*

Demonstração. Inicialmente escolhamos um subespaço $\mathfrak{p}$ tal que $M_n(\mathbb{K}) = \mathfrak{g} \oplus \mathfrak{p}$ e projeções $p : M_n(\mathbb{K}) \rightarrow \mathfrak{p}$ e $g : M_n(\mathbb{K}) \rightarrow \mathfrak{g}$ tais que para todo $A \in M_n(\mathbb{K})$ se tenha $A = p(A) + g(A)$.

Definimos a função $\Phi : M_n(\mathbb{K}) \rightarrow M_n(\mathbb{K})$ como $\Phi(A) = \exp(g(A)) \exp(p(A))$. Note que nós temos $\Phi|_{\mathfrak{g}} = \exp|_{\mathfrak{g}}$ e $\Phi|_{\mathfrak{p}} = \exp|_{\mathfrak{p}}$. Pela linearidade da diferencial, temos

$$\begin{aligned} d\Phi(\mathbf{0})(A) &= d\Phi(\mathbf{0})(p(A)) + d\Phi(\mathbf{0})(g(A)) \\ &= \lim_{t \rightarrow 0} \left[\frac{\Phi(p(tA)) - \Phi(p(\mathbf{0}))}{t} \right] + \lim_{t \rightarrow 0} \left[\frac{\Phi(g(tA)) - \Phi(g(\mathbf{0}))}{t} \right] \\ &= \lim_{t \rightarrow 0} \left[\frac{\exp(p(tA)) - I_n}{t} \right] + \lim_{t \rightarrow 0} \left[\frac{\exp(g(tA)) - I_n}{t} \right] \\ &= d\exp(\mathbf{0})(p(A)) + d\exp(\mathbf{0})(g(A)) = d\exp(\mathbf{0})(A). \end{aligned}$$

Como A é arbitrário, isto implica que $d\Phi(\mathbf{0}) = d\exp(\mathbf{0})$, que é a função identidade de $M_n(\mathbb{K})$.

Pelo Teorema da Função Inversa existe um $\delta > 0$ tal que Φ é invertível em $B(\Phi(\mathbf{0}), \delta) = B(I_n, \delta)$. Caso exista um r , $0 < r < \delta$, tal que $\Phi^{-1}(B(I_n, r) \cap G) \subset \mathfrak{g}$ teremos $B(I_n, r) \cap G \subset \Phi(\mathfrak{g})$ e como $\Phi|_{\mathfrak{g}} = \exp|_{\mathfrak{g}}$, então $B(I_n, r) \cap G \subset \exp(\mathfrak{g})$.

Se o lema for falso, então não existe nenhum $r > 0$ com $\Phi^{-1}(B(I_n, r) \cap G) \subset \mathfrak{g}$, ou seja, $\Phi^{-1}(B(I_n, r) \cap G) \cap (M_n(\mathbb{K}) - \mathfrak{g}) \neq \emptyset$ e portanto podemos escolher elementos X em G que estão arbitrariamente próximos de I mas $\Phi^{-1}(X)$ não está em $\mathfrak{g}$. Note que como Φ^{-1} é contínua $\Phi^{-1}(X)$ está cada vez mais perto de (0) . Tomemos então uma sequência A_i tal que $A_i \rightarrow (0)$, tal que para todo i , $\Phi(A_i) \in G$ e $A_i \notin \mathfrak{g}$. Temos

$$\Phi(A_i) \in G \Rightarrow \exp(g(A_i)) \exp(p(A_i)) \in G.$$

Sabemos pelo Teorema 4.11 que $\exp(\mathfrak{g}) \subset G$. Assim

$$\exp(-g(A_i)) \in G \Rightarrow \exp(-g(A_i)) \exp(g(A_i)) \exp(p(A_i)) \in G \Rightarrow \exp(p(A_i)) \in G \quad (11)$$

e assim a sequência $\exp(p(A_i))$ é uma sequência de elementos em G .

Consideremos agora a sequência $\frac{p(A_i)}{|p(A_i)|}$. Pela compacidade da esfera unitária de $\mathfrak{p}$ essa sequência possui uma subsequência convergente, digamos $\frac{p(B_i)}{|p(B_i)|}$ com $B = \lim_{i \rightarrow \infty} \frac{p(B_i)}{|p(B_i)|}$. Como a esfera unitária de $\mathfrak{p}$ é um conjunto fechado, B é um elemento da esfera unitária de $\mathfrak{p}$ e em particular $B \in \mathfrak{p}$.

Agora, como a sequência $p(B_i) \rightarrow (0)$ podemos construir para qualquer número real t uma sequência $n(t)_i$ de inteiros tais que $n(t)_i p(B_i) \rightarrow tB$ (a sequência $n(t)_i$ pode ser definida por exemplo como o menor inteiro m , dado i , que satisfaz $|m p(B_i) - tB| \leq \left| \frac{p(B_i)}{2} \right|$).

Por (11) sabemos que $\exp(p(B_i)) \in G$ e assim, como G é um grupo, temos então que para todo t real $\exp(n(t)_i p(B_i)) \in G$.

G é um conjunto fechado e $\exp(n(t)_i p(B_i)) \rightarrow \exp(tB)$ então $\exp(tB) \in G$ para todo t . Isto significa que $\gamma(t) = \exp(tB)$ é um caminho em G tal que $\gamma(0) = I$ e $d\gamma(0) = B$. Portanto $B \in \mathfrak{g}$.

$B \in \mathfrak{p}$ então $B \in \mathfrak{g} \cap \mathfrak{p} \Rightarrow B = 0$, entretanto B está na esfera unitária de $\mathfrak{p}$ e portanto $|B| = 1$, o que é um absurdo e logo a hipótese que fizemos era falsa, existindo então um r tal que $\Phi^{-1}(B(I_n, r) \cap G) \subset \mathfrak{g}$ e assim $B(I_n, r) \cap G \subset \exp(\mathfrak{g})$, portanto o lema é verdadeiro. $\square$

Teorema 5.3. *Todo grupo matricial G é uma variedade diferencial de dimensão $\dim(\mathfrak{g})$, onde $\mathfrak{g}$ é a álgebra de Lie de G .*

Demonstração. Sabemos pelo Teorema da Função Inversa e pelos lemas 4.11 e 5.2 que existem vizinhanças U de $\mathbf{0}$ e V de I tais que $\exp : U \cap \mathfrak{g} \rightarrow V \cap G$ é um difeomorfismo de classe C^∞ .

Consideremos agora para cada $X \in G$ e consideremos o conjunto $X(V \cap G) = \{XA \mid A \in V \cap G\}$ (que é uma vizinhança de X em G) e a função $X \exp : U \cap \mathfrak{g} \rightarrow X(V \cap G)$. Como $X \in GL_n(\mathbb{K})$, então a função $A \mapsto XA$ é uma bijeção linear em um espaço de dimensão finita, por isso é um difeomorfismo C^∞ . Assim $X \exp : U \cap \mathfrak{g} \rightarrow X(V \cap G)$ também é um difeomorfismo para cada $X \in G$. Como para cada $X \in G$ a vizinhança $X(V \cap G)$ é difeomorfa ao aberto $U \cap \mathfrak{g}$ de $\mathfrak{g}$, o grupo G é uma variedade diferencial de dimensão $\dim(\mathfrak{g})$. $\square$

Referências

- [1] Arnaldo Garcia e Yves Lequain. *Elementos de Álgebra*. Rio de Janeiro: IMPA, 2015.
- [2] Joachim Hilgert e Karl-Hermann Neeb. *Structure and Geometry of Lie Groups*. New York: Springer, 2012.
- [3] Elon Lages Lima. *Curso de análise*. Vol. 1. Rio de Janeiro: IMPA, 2019.
- [4] Elon Lages Lima. *Curso de análise*. Vol. 2. Rio de Janeiro: IMPA, 2020.
- [5] Elon Lages Lima. *Espaços métricos*. Rio de Janeiro: IMPA, 2013.
- [6] Elon Lages Lima. *Álgebra linear*. Rio de Janeiro: IMPA, 2014.
- [7] Kristopher Tapp. *Matrix groups for undergraduates*. Providence, Rhode Island: American Mathematical Society, 2005.