

UNIVERSIDADE ESTADUAL DE MARINGÁ

CENTRO DE CIÊNCIAS EXATAS

DEPARTAMENTO DE MATEMÁTICA

PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

(Mestrado)

FERNANDO CORDEIRO DE QUEIROZ

LINEARIZAÇÃO DE SISTEMAS DE CONTROLE

Maringá-PR

2016

UNIVERSIDADE ESTADUAL DE MARINGÁ
CENTRO DE CIÊNCIAS EXATAS
DEPARTAMENTO DE MATEMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

LINEARIZAÇÃO DE SISTEMAS DE CONTROLE

FERNANDO CORDEIRO DE QUEIROZ

Dissertação apresentada ao Programa de Pós-Graduação em Matemática do Departamento de Matemática, Centro de Ciências Exatas da Universidade Estadual de Maringá, como requisito para obtenção do título de Mestre em Matemática.

Área de concentração: Geometria/Topologia.

Orientador: Prof. Dr. Marcos André Verdi.

Maringá-PR

2016

LINEARIZAÇÃO DE SISTEMAS DE CONTROLE

FERNANDO CORDEIRO DE QUEIROZ

Dissertação apresentada ao Programa de Pós-Graduação em Matemática do Departamento de Matemática, Centro de Ciências Exatas da Universidade Estadual de Maringá, como requisito para obtenção do título de Mestre em Matemática.

BANCA EXAMINADORA

Prof. Dr. Marcos André Verdi

Universidade Estadual de Maringá - UEM - PR

Prof. Dr. Alexandre José Santana

Universidade Estadual de Maringá - UEM - PR

Prof. Dr. Hélio Vinicius Moreno Tozatti

Universidade Tecnológica Federal do Paraná - *campus* Campo Mourão - PR

Maringá, 17 de Junho de 2016.

*Este trabalho é dedicado aos meus pais,
Sr. Milton de Queiroz e,
Sr^{ta}. Maria Luzani Cordeiro de Queiroz.*

Agradecimentos

Primeiramente agradeço à Deus por ter me dado forças para persistir na realização deste trabalho.

Aos meus pais, Sr. Milton de Queiroz e Sr^a. Maria Luzani Cordeiro de Queiroz que sempre me apoiaram e incentivaram desde o início desta jornada.

Agradeço a todos os professores do departamento de matemática da Universidade Estadual de Maringá com quem tive contato no curso de graduação e mestrado que adicionaram muito para minha formação pessoal e profissional, à banca examinadora pelas correções e sugestões. Agradeço especialmente o Prof. Dr. Marcos André Verdi pela orientação no curso de mestrado, pelo aprendizado e por contribuir sempre com muita paciência e dedicação para o meu crescimento profissional.

Aos amigos que conquistei durante este percurso e que sempre me ajudaram quando precisei. Agradeço especialmente ao Me. João A. N. Cossich, Me. Bruno A. Rodrigues, Juniormar Organista, Me. Jonathan Prass e Fernando H. da Silva Vieira pela amizade, e pelas várias horas de estudos.

Finalmente, agradeço à CAPES pelo apoio financeiro que foi fundamental no período de realização deste trabalho.

Resumo

Nessa dissertação estudaremos alguns resultados clássicos da teoria de sistemas de controle. Demonstraremos o Critério de Kalman para a controlabilidade de sistemas lineares, o Teorema da Órbita e os Teoremas de Rashevsky-Chow e Krener. Também abordaremos a classificação dos sistemas de controle lineares controláveis, através da forma normal de Brunovský. Por fim, apresentamos o principal resultado desta dissertação: estabelecer condições para que um sistema de controle afim seja equivalente a um sistema linear controlável.

Palavras-chave: Sistemas de controle; Critério de Kalman; Forma Normal de Brunovský; linearização de sistemas afins.

Abstract

In this work we study some classical results of the control theory. We will demonstrate the Kalman Condition for controllability of linear systems, the Orbit Theorem and Rashevsky-Chow and Krener theorems. We will also discuss the classification of linear controllable systems, through of the Brunovský normal form. Finally, we present the main result in this dissertation: to establish conditions in order to affine control system be equivalent to a controllable linear system.

Keywords: Control systems; Kalman condition; Brunovský normal form; affine systems linearization.

CONTEÚDO

1	Preliminares	14
1.1	Variedades e campos de vetores diferenciáveis	14
1.2	Equações diferenciais ordinárias e fluxo	18
1.3	Sistemas de controle e colchete de Lie	27
1.4	Variedades e o espaço $\mathcal{C}^\infty(M)$	32
1.4.1	Pontos, difeomorfismos e campos de vetores	32
1.4.2	Ação de difeomorfismos sobre campos de vetores	41
2	Controlabilidade de Sistemas de Controle	47
2.1	Sistemas de controle lineares	47
2.2	Controlabilidade	49
3	Linearização de Sistemas de Controle	55
3.1	Linearização local	55
3.2	Linearização global	64
4	O Teorema da Órbita e Aplicações	75
4.1	Subvariedades imersas	75
4.2	O Teorema da Órbita	79
4.3	Caso analítico	92
4.4	Conjuntos atingíveis	97

5	State e Feedback Equivalência de Sistemas de Controle	101
5.1	State equivalência	101
5.2	Feedback equivalência	108
5.3	Sistemas lineares	111
5.3.1	Sistemas lineares com controle escalar	113
5.3.2	Sistemas lineares com controle vetorial	119
5.3.2.1	Índices de Kronecker	122
5.3.2.2	Forma Normal de Brunovský	126
6	Linearização por State-Feedback Equivalência	139

INTRODUÇÃO

A Teoria de Controle formalizada matematicamente como a conhecemos hoje é relativamente recente. Porém, a ideia de regulação de sistemas, que norteia a Teoria de Controle, tem suas origens na antiguidade, surgindo da necessidade de regular certos mecanismos físicos para uma finalidade cotidiana, como por exemplo, na construção de mecanismos de marcação de tempo e nas construções de aquedutos. Alguns historiadores afirmam que na antiga Mesopotâmia, em torno de 2000 anos a.C, já era praticado o controle sobre o sistema de irrigação.

Muito tempo depois, os trabalhos realizados por Ch. Huygens e R. Hooke, no final do século XVII, sobre as oscilações do pêndulo originaram um exemplo marcante da importância da Teoria de Controle. O objetivo era conseguir uma medição precisa de tempo e localização, que eram muito importantes para a navegação. Esse trabalho inspirou J. Watt a inventar, com ajustes, a máquina a vapor. De modo geral, o mecanismo consistia de um sistema com bolas rotativas em torno de um eixo de rotação, que com o aumento da velocidade de rotação se afastavam do eixo, ativando um controlador que abria válvulas permitindo que a pressão diminuísse e, conseqüentemente, diminuía a velocidade de rotação. Dessa forma, a máquina funcionava com uma velocidade próxima de uma constante. Mais tarde, o astrônomo G. Airy foi o primeiro cientista a analisar matematicamente o sistema de regulação inventado por J. Watt. Porém, uma descrição matemática definitiva foi dada apenas em 1868 por J. C. Maxwell.

Com a necessidade cada vez maior da automação de equipamentos, o interesse por essa teoria cresceu no início do século XX. No pós Segunda Guerra, ficou claro para os cientistas e engenheiros que os modelos considerados até então não eram su-

ficientemente precisos para descrever a complexidade das situações, criando maiores esforços em pesquisa de controlabilidade de sistemas não lineares, isto é, sistemas que são afetados por forças externas. Após 1960, os métodos e as ideias anteriores foram consideradas como parte da “Teoria de Controle Clássica”. Sugerimos o artigo [5] para o leitor interessado em obter mais notas históricas.

Os alicerces da Teoria de Controle Moderna foram estabelecidos por R. Bellman, R. Kalman e L. Pontryagin. A Teoria de controle pode ser considerada hoje em dia a partir de dois pontos de vistas diferentes e complementares: como suporte teórico para Engenharia de Controle e também como uma área de pesquisa matemática. A separação destas áreas é extremamente vaga e, na verdade, a Teoria de Controle é uma das mais interdisciplinares áreas de pesquisa hoje em dia.

O ponto de vista adotado neste trabalho é puramente matemático. De uma maneira um pouco informal, um sistema de controle pode ser visto como uma família de equações diferenciais ordinárias, parametrizadas por um *controle* u , sobre uma variedade diferenciável. A variedade é chamada de *espaço estado*. Para cada controle e estado inicial fixado, obtemos uma *trajetória do sistema*, que é um curva-solução do Problema de Valor Inicial obtido. Um problema interessante da Teoria de Controle é o da *controlabilidade*: partindo de um estado inicial x_0 e “caminhando” sobre as trajetórias do sistema, mudando de trajetória sempre que desejarmos, quais são os pontos que podemos atingir num certo tempo? Quais as propriedades dos conjuntos de pontos que podemos atingir? Veremos algumas respostas para essas perguntas nessa dissertação. Outros aspectos da Teoria de Controle, que não serão tratados aqui, são: A *estabilidade*, que analisa se um estado é estável para o sistema de controle; a *observabilidade*, que estuda a questão em que, conhecendo um controle e uma equação de observação, podemos descobrir o estado inicial do sistema; a *otimizabilidade*, que estuda a existência de controles que permitem partir de um estado e alcançar um outro num tempo mínimo.

Nessa dissertação estudaremos essencialmente aspectos da controlabilidade. Os sistemas de controle mais simples são os *sistemas lineares* em \mathbb{R}^n . Esses sistemas são da forma

$$\dot{x} = Ax + \sum_{i=1}^k u_i b_i,$$

onde x, b_1, \dots, b_k são vetores de \mathbb{R}^n e o controle $u = (u_1, \dots, u_k)$ está em \mathbb{R}^k . Condições relativamente simples para a *controlabilidade* do sistema, isto é, para que possamos atin-

gir qualquer ponto à partir de um estado inicial, são estabelecidas no chamado critério de Kalman. Por esta razão, é interessante saber quando um sistema não linear é, de alguma forma, equivalente a um linear. Como a controlabilidade do sistema depende de suas trajetórias, as boas transformações de equivalência de sistemas, sob o ponto de vista da controlabilidade, são aquelas nas quais as trajetórias de um sistema são levadas nas trajetórias do outro. Isso pode ser feito de duas formas: levar uma trajetória de um sistema em uma trajetória do outro, correspondendo ao mesmo controle; ou então levar a trajetória em outra correspondente a um controle diferente. Estes tipos de transformação são chamadas, respectivamente, de *state* transformações e *feedback* transformações, e são o principal objeto de estudo dessa dissertação. A principal referência deste trabalho é o livro [1], dos autores Andrei Agrachev e Yuri Sachkov.

Essa dissertação está organizada da seguinte forma: o Capítulo 1 tem caráter introdutório, contendo várias definições e resultados essenciais para o restante do trabalho. As Seções 3 e 4 têm maior relevância, pois nelas são definidas as principais ferramentas utilizadas aqui, como o colchete de Lie e o operador adjunto. No Capítulo 2 introduzimos formalmente o conceito de sistema de controle e provamos o Critério de Kalman para controlabilidade de sistemas lineares. No Capítulo 3, estudamos as condições para que um sistema não linear seja localmente e globalmente equivalente, por *state* transformações, a um sistema linear controlável. No Capítulo 4 apresentamos um importante resultado da teoria de controle, o Teorema da Órbita. A *órbita* de um sistema de controle através de um ponto é, de uma maneira informal, o conjunto dos pontos que podemos atingir através do ponto inicial, considerando tempos positivos e negativos. Num primeiro momento, a órbita poderia ser um conjunto qualquer, sem "boas" propriedades. No entanto, o Teorema da Órbita estabelece que esses conjuntos são na verdade subvariedades imersas, descrevendo também seus espaços tangentes. Com o auxílio deste resultado, descrevemos os conjuntos atingíveis para uma família de campos de vetores. No Capítulo 5 aprofundamos um pouco mais o conceito de *state* equivalência, que já foi tratada no Capítulo 3, e introduzimos formalmente a noção de *feedback* equivalência. Além disso, apresentamos uma classificação de sistemas de controle lineares. Isto é feito da seguinte maneira: definimos um tipo mais simples de sistemas lineares, chamado *forma normal de Brunovský*, e mostramos que todo sistema linear é *state-feedback* equivalente um sistema que é desta forma. Por fim, no Capítulo 6, estabelecemos condições para que um sistema de controle afim seja equivalente, por

state-feedback transformações, a um sistema linear controlável.

CAPÍTULO 1

PRELIMINARES

O objetivo principal deste capítulo é definir o ambiente e as ferramentas fundamentais com as quais vamos desenvolver este trabalho. As Seções 1 e 2 contêm várias definições básicas e os principais resultados sobre variedades diferenciáveis, campos vetoriais, E.D.O's e o fluxo gerado por um campo vetorial, que podem ser aprofundadas em [8], [14] e [1]. As Seções 3 e 4 possuem maior relevância. Ne-las definimos o conjunto dos pontos atingíveis do sistema de controle, o colchete de Lie, fornecemos exemplos e introduzimos uma correspondência entre variedades e o espaço das funções continuamente diferenciáveis. Esta correspondência será fundamental para todo o restante do texto.

1.1 Variedades e campos de vetores diferenciáveis

Esta seção introduz apenas a essência dos conceitos de variedades e campos de vetores diferenciáveis sem a preocupação com a classe de diferenciabilidade de ambos conceitos. Desta forma, assumiremos que são de classe C^∞ em todo o texto. Apresentaremos somente os principais resultados necessários ao nosso objetivo.

Definição 1.1 Dizemos que um subconjunto $M \subset \mathbb{R}^n$ é uma *subvariedade diferenciável* de \mathbb{R}^n de dimensão $k \leq n$ se, para qualquer ponto $q \in M$, existem vizinhanças abertas $O_q \subset \mathbb{R}^n$,

$O_0 \subset \mathbb{R}^n$ da origem em \mathbb{R}^n e uma aplicação diferenciável $\Phi : O_q \rightarrow O_0$ com $\Phi(q) = 0$ e

$$\dim(D_q\Phi \cdot \mathbb{R}^n) = n$$

tal que

$$\Phi(O_q \cap M) = O_0 \cap \mathbb{R}^k.$$

A seguir apresentaremos uma definição de variedade abstrata. Esta usa o conceito de *paracompacidade* que será omitida, mas pode ser encontrada em [7] página 284.

Definição 1.2 *Uma variedade diferenciável M de dimensão k é um espaço topológico de Hausdorff, paracompacto munido de uma estrutura diferenciável, isto é: M é coberto por um sistema de subconjuntos abertos*

$$M = \cup_{\alpha} O_{\alpha},$$

chamadas vizinhanças coordenadas, tal que para cada O_{α} existe um homeomorfismo

$$\Phi_{\alpha} : O_{\alpha} \rightarrow \mathbb{R}^k,$$

chamado sistema de coordenadas locais, tal que para quaisquer α, β as composições

$$\Phi_{\beta} \circ \Phi_{\alpha}^{-1} : \Phi_{\alpha}(O_{\alpha} \cap O_{\beta}) \subset \mathbb{R}^k \rightarrow \Phi_{\beta}(O_{\alpha} \cap O_{\beta}) \subset \mathbb{R}^k$$

são difeomorfismos.

Definição 1.3 *Sejam M, N variedades diferenciáveis de dimensões k e l respectivamente. Uma aplicação contínua $\varphi : M \rightarrow N$ é diferenciável se a **expressão** de φ em coordenadas locais é diferenciável, isto é, se*

$$M = \cup_{\alpha} O_{\alpha}, \quad N = \cup_{\beta} V_{\beta}$$

são cobertas por vizinhanças coordenadas de M, N respectivamente, e

$$\Phi_{\alpha} : O_{\alpha} \rightarrow \mathbb{R}^k \quad \Psi_{\beta} : V_{\beta} \rightarrow \mathbb{R}^l,$$

são sistemas de coordenadas locais, então

$$\Psi_\beta \circ \varphi \circ \Phi_\alpha^{-1} : \Phi_\alpha(O_\alpha \cap \varphi^{-1}(V_\beta)) \subset \mathbb{R}^k \rightarrow \Psi_\beta(\varphi(O_\alpha) \cap V_\beta) \subset \mathbb{R}^l$$

é diferenciável para quaisquer α, β .

Definição 1.4 *Sejam M, N variedades diferenciáveis. Dizemos que M é **difeomorfa** a N se existe um homeomorfismo*

$$\varphi : M \rightarrow N$$

tal que φ e φ^{-1} são diferenciáveis. A função φ é chamada de **difeomorfismo**.

Denotaremos por $\text{Diff}(M) = \{\varphi : M \rightarrow M \mid \varphi \text{ é difeomorfismo}\}$ o conjunto de todos os difeomorfismos de M em M .

Definição 1.5 *Dizemos que uma aplicação diferenciável $\varphi : M \rightarrow N$ é um **mergulho** de M em N se*

$$\varphi : M \rightarrow \varphi(M)$$

é um difeomorfismo. Uma aplicação $\varphi : M \rightarrow N$ é dita **própria** se $\varphi^{-1}(K)$ é compacto para todo subconjunto compacto $K \subset N$.

Teorema 1.6 (Whitney) *Toda variedade diferenciável conexa de dimensão k pode ser mergulhada propriamente em \mathbb{R}^{2k+1} .*

Da mesma forma que omitimos a classe de diferenciabilidade, assumiremos que todas as variedades e subvariedades mencionadas no restante do texto são diferenciáveis. Assim, omitiremos também a palavra diferenciável.

Definição 1.7 *Seja $M \subset \mathbb{R}^n$ uma subvariedade de dimensão $k \leq n$ e $q \in M$. O **espaço tangente** à M no ponto q é o subespaço vetorial de dimensão k dado por*

$$T_q M = (D_q \Phi)^{-1} \cdot \mathbb{R}^k,$$

onde Φ é a aplicação da definição 1.1.

Note nesta definição, que é fundamental o ambiente euclidiano que contém a sub-variedade. Quando nos referimos a uma variedade abstrata M , não temos esta referência euclidiana. Desta forma, também é necessário definir o espaço tangente à M de maneira abstrata. A maneira que vamos definir o espaço tangente abstrato é inspirado em [8], que consiste em escolher um ponto da variedade e construir um conjunto formado por todos os caminhos que passam pelo ponto escolhido no mesmo instante de tempo, e diferenciável neste instante. Em seguida, definiremos uma relação neste conjunto, que será uma relação de equivalência. O espaço tangente à variedade é então definido como o espaço quociente do conjunto construído por esta relação. Em outras palavras, um vetor será uma classe de equivalência.

Sejam M uma variedade abstrata de dimensão k e $q \in M$. Denote por \mathcal{C}_q o conjunto de todos os caminhos contínuos $\alpha : I \rightarrow M$, onde I é um intervalo contendo zero, com $\alpha(0) = q$ e diferenciável em $t = 0$.

Sejam $\alpha \in \mathcal{C}_q$ um caminho e $\Phi : O \rightarrow \mathbb{R}^k$ um sistema de coordenadas locais com $q \in O$. Ajustando, se necessário, o intervalo I , obtemos a composição

$$(\Phi \circ \alpha) : I \rightarrow \mathbb{R}^k.$$

Diremos que dois caminhos $\alpha, \beta \in \mathcal{C}_q$ são *equivalentes*, e escreveremos $\alpha \sim \beta$ se existe um sistema de coordenadas locais $\Phi : O \rightarrow \mathbb{R}^k$ em M com $q \in O$ tal que, $(\Phi \circ \alpha) : I_1 \rightarrow \mathbb{R}^k$ e $(\Phi \circ \beta) : I_2 \rightarrow \mathbb{R}^k$ têm o mesmo vetor-velocidade em $t = 0$, ou seja

$$(\Phi \circ \alpha)'(0) = (\Phi \circ \beta)'(0).$$

Note que a igualdade acima é válida para toda vizinhança coordenada O' contendo q e contida em O . Logo, esta igualdade valerá para todo sistema de coordenadas locais. Esta relação é claramente reflexiva e simétrica. Agora, se $\alpha \sim \beta$ e $\beta \sim \gamma$, então existem sistemas de coordenadas locais Φ_1 e Φ_2 , tais que

$$(\Phi_1 \circ \alpha)'(0) = (\Phi_1 \circ \beta)'(0), \quad (\Phi_2 \circ \beta)'(0) = (\Phi_2 \circ \gamma)'(0).$$

Como ambas igualdades são válidas para todo sistema de coordenadas locais, obtemos

$$(\Phi_1 \circ \alpha)'(0) = (\Phi_1 \circ \gamma)'(0),$$

segue que esta relação é transitiva $\alpha \sim \gamma$ e, portanto, uma relação de equivalência.

O espaço tangente à M no ponto q é definido como o conjunto quociente

$$\frac{\mathcal{C}_q}{\sim} = T_q M.$$

Desta forma, o vetor-velocidade $\dot{\alpha}$ do caminho $\alpha \in \mathcal{C}_q$, é a classe de equivalência de α .

Ou seja,

$$\dot{\alpha} = \{\beta \in \mathcal{C}_q \mid \beta \sim \alpha\} \in T_q M.$$

Definição 1.8 *Um campo de vetores diferenciável sobre uma variedade M é uma aplicação diferenciável*

$$V : M \rightarrow TM,$$

que associa a cada ponto q de M um vetor $V(q)$ de $T_q M$ do fibrado tangente TM .

Denotaremos por $\text{Vec}(M) = \{V : M \rightarrow TM \mid V \text{ é diferenciável}\}$ o conjunto de todos os campos de vetores diferenciáveis sobre M . Assumiremos também que todo campo de vetores mencionado no restante do texto é diferenciável, salvo menção contrária.

1.2 Equações diferenciais ordinárias e fluxo

Esta seção contém, além de uma breve introdução do conceito de E.D.O gerada por um campo de vetores sobre M , o importante Teorema de Existência e Unicidade de soluções para subvariedades. Definiremos também o fluxo gerado por um campo de vetores e finalizaremos tratando da completude de um campo de vetores.

Definição 1.9 *Uma equação diferencial ordinária, ou simplesmente E.D.O, sobre uma variedade M é uma equação da forma:*

$$\dot{q} = V(q), \tag{1.2-1}$$

ou

$$\frac{d}{dt}q(t) = V(q(t)),$$

onde $V \in \text{Vec}(M)$, $q \in M$ e $q(t)$ representa uma aplicação de \mathbb{R} em M .

Um problema de Cauchy consiste de uma E.D.O sujeita a uma condição inicial da forma $q(0) = q_0$, $q_0 \in M$. Uma solução de (1.2-1) é uma aplicação diferenciável (ou curva integral) definida em um intervalo aberto de \mathbb{R} em M :

$$\alpha : I \subset \mathbb{R} \rightarrow M,$$

tal que

$$\dot{\alpha}(t) = V(\alpha(t)) \quad \forall t \in I.$$

Definição 1.10 *Sejam M , N variedades e $\Phi : M \rightarrow N$ uma aplicação diferenciável. A diferencial de Φ em um ponto $q \in M$ é a transformação linear*

$$D_q\Phi : T_qM \rightarrow T_{\Phi(q)}N$$

definida por:

$$D_q\Phi \left(\left. \frac{d}{dt}\alpha(t) \right|_{t=0} \right) = \left. \frac{d}{dt}\Phi(\alpha(t)) \right|_{t=0},$$

onde $\alpha : I \subset \mathbb{R} \rightarrow M$ é uma aplicação diferenciável iniciando em q , isto é, $\alpha(0) = q$.

Observação 1.11 *Sejam*

$$\dot{q} = V(q) \quad \forall q \in M, \tag{1.2-2}$$

uma E.D.O e $\Phi : M \rightarrow N$ um difeomorfismo entre as variedades M e N , com $x = \Phi(q)$.

Consideremos $\alpha : I \subset \mathbb{R} \rightarrow M$ uma solução de (1.2-2) e seja $x(t) = \Phi(\alpha(t))$. Pela regra da cadeia obtemos:

$$\begin{aligned} \dot{x}(t) &= \frac{d}{dt}\Phi(\alpha(t)) \\ &= D_q\Phi(\dot{\alpha}(t)) \\ &= D_q\Phi(V(\alpha(t))) \\ &= D_{\Phi^{-1}(x)}\Phi \cdot V(\Phi^{-1}(x(t))). \end{aligned}$$

Com isso, obtemos a E.D.O sobre N , induzida pelo difeomorfismo Φ :

$$\dot{x}(t) = D_{\Phi^{-1}(x)}\Phi \cdot V(\Phi^{-1}(x)).$$

O campo de vetores em N induzido por V pelo difeomorfismo Φ é definido como

$$\Phi_*V(x) = D_{\Phi^{-1}(x)}\Phi \cdot V(\Phi^{-1}(x)). \quad (1.2-3)$$

Teorema 1.12 (Existência e unicidade de soluções) Considere o problema de Cauchy em uma subvariedade M de \mathbb{R}^n ,

$$\begin{cases} \dot{q} = V(q) \\ q(0) = q_0. \end{cases}$$

para qualquer $q_0 \in M$ fixado. Então existe uma única solução do problema acima

$$q : (a, b) \subset \mathbb{R} \rightarrow M$$

onde (a, b) é um intervalo suficientemente pequeno com $a < 0 < b$ tal que $q(0) = q_0$.

Demonstração: Pela Definição 1.1 existe uma aplicação diferenciável

$$\Phi : O_{q_0} \subset \mathbb{R}^n \rightarrow O_0 \subset \mathbb{R}^n$$

com $\Phi(q_0) = 0$ e

$$\dim(D_{q_0}\Phi \cdot \mathbb{R}^n) = n$$

tal que

$$\Phi(O_{q_0} \cap M) = O_0 \cap \mathbb{R}^k.$$

Como o fato de $D_{q_0}\Phi$ ser sobrejetora implica que tal aplicação é também injetora, temos que $D_{q_0}\Phi$ é um isomorfismo. Logo, Φ é um difeomorfismo local, ou seja, existem vizinhanças abertas $O'_{q_0} \subset O_{q_0}$ e $O'_0 \subset O_0$ tais que:

$$\Phi : O'_{q_0} \rightarrow O'_0 \subset \mathbb{R}^n$$

é difeomorfismo. Considere o campo de vetores em \mathbb{R}^n induzido pelo difeomorfismo Φ :

$$(\Phi_* V)(x) \quad x \in U_0$$

e o problema de Cauchy

$$\begin{cases} \dot{x} = (\Phi_* V)(x) \\ x(0) = 0. \end{cases}$$

Sabemos que existe uma única solução $x(t)$ em O'_0 passando por $0 \in \mathbb{R}^n$ (ver [13], Capítulo 1). Além disso existe uma única curva $q(t)$ em O'_{q_0} tal que $\Phi(q(t)) = x(t)$. Pela regra da cadeia,

$$\dot{x}(t) = D_q \Phi \cdot \dot{q}(t), \quad (1.2-4)$$

e por (1.2-3)

$$\dot{x} = (\Phi_* V)(x) = D_q \Phi \cdot V(q(t)). \quad (1.2-5)$$

De (1.2-4) e (1.2-5) obtemos que:

$$\begin{cases} \dot{q} = V(q) \\ q(0) = \Phi^{-1}(x(0)) = q_0. \end{cases}$$

Portanto $q(t)$ é solução local do problema de Cauchy em M . ■

Observação 1.13 *A aplicação q do teorema anterior*

$$\begin{aligned} q : (a_{q_0}, b_{q_0}) \times M &\rightarrow M \\ (t, q_0) &\mapsto q(t, q_0) \end{aligned}$$

é diferenciável. Em particular, o domínio da solução

$$q(\cdot, q_0) : (a_{q_0}, b_{q_0}) \rightarrow M$$

pode ser escolhido continuamente dependendo do ponto q_0 .

Teorema 1.14 *Sejam M uma subvariedade de \mathbb{R}^n e*

$$\dot{q} = V(q) \quad q \in \mathbb{R}^n \quad (1.2-6)$$

uma E.D.O tal que

$$q \in M \Rightarrow V(q) \in T_q M.$$

Então para qualquer ponto inicial $q_0 \in M$, existe $\epsilon > 0$ tal que a solução $q(t)$ de (1.2-6) com $q(0) = q_0$ pertence a M para $t \in (-\epsilon, \epsilon)$.

Demonstração: Considere a restrição do campo de vetores V sobre M :

$$f = V \Big|_M.$$

Obtemos assim o problema de Cauchy:

$$\begin{cases} \dot{q} = f(q) & q \in M \\ q(0) = q_0. \end{cases}$$

Pelo Teorema 1.12, existe uma única solução local $q(t)$ do problema acima para $t \in (a, b)$ com $a < 0 < b$. Tome $\epsilon > 0$ suficientemente pequeno tal que $(-\epsilon, \epsilon) \subset (a, b)$. Portanto, pela unicidade da solução, temos que $q(t)$ é também solução de (1.2-6), e $q(t) \subset M$ para todo $t \in (-\epsilon, \epsilon)$. ■

Definição 1.15 Dizemos que um campo de vetores $V \in \text{Vec}(M)$ é **completo** se, para todo $q_0 \in M$, a solução $q(t)$ do problema de Cauchy:

$$\begin{cases} \dot{q} = V(q) & q \in M \\ q(0) = q_0 \end{cases}$$

é definida para todo $t \in \mathbb{R}$.

Exemplo 1.16 O campo de vetores definido por

$$V(q) = q$$

é completo em \mathbb{R} , $\mathbb{R} \setminus \{0\}$, $(-\infty, 0)$, $(0, +\infty)$, $\{0\}$.

De fato, temos o problema de Cauchy:

$$\begin{cases} \dot{q} = q, \\ q(0) = q_0, \end{cases}$$

com solução $q(t) = e^t q_0$. Observe que este campo não é completo em todas as subvariedades de \mathbb{R} . De fato, para a subvariedade $S = (0, 1) \subset \mathbb{R}$ temos que a solução $q(t) = e^t q$, $q \in S$, não está definida para todo $t \in \mathbb{R}$ pois, tomando $t_0 > 0$ suficientemente grande obtemos $e^{t_0} q \notin (0, 1)$. Veja a Figura 1.1 abaixo. Para as outras subvariedades diferentes daquelas mencionadas no início, a análise é análoga.

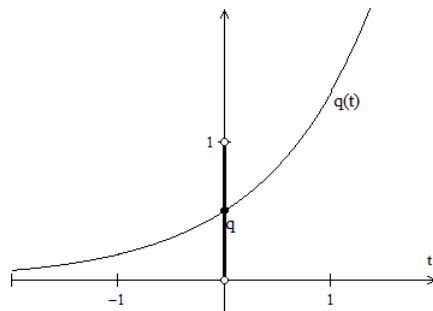


Figura 1.1: $e^t q$

Exemplo 1.17 O campo de vetores definido por

$$V(q) = q^2$$

não é completo em nenhuma subvariedade de \mathbb{R} , exceto $\{0\}$.

De fato, para $q_0 \neq 0$, temos o problema de Cauchy:

$$\begin{cases} \dot{q} = q^2, \\ q(0) = q_0, \end{cases}$$

com solução

$$x(t) = \frac{1}{k - t},$$

para todo $t \neq k$ e $k = \frac{1}{q_0}$ constante. Mas, se $M = \{0\}$ então $q_0 = 0$. Logo o problema de Cauchy possui única solução $q(t) = 0$ para todo $t \in \mathbb{R}$.

Observação 1.18 Considere o problema de Cauchy

$$\begin{cases} \dot{q} = V(q), \\ q(0) = q_0. \end{cases}$$

Pelo Teorema de Existência e Unicidade de soluções, sabemos que para cada ponto $q_0 \in M$ existe um intervalo suficientemente pequeno $(a_{q_0}, b_{q_0}) \subset \mathbb{R}$ no qual a solução está definida.

Suponha que exista um intervalo $(-\varepsilon, \varepsilon) \subset \mathbb{R}$, $\varepsilon > 0$, tal que a solução do problema de Cauchy com $q(0) = q_0$ está definido em $(-\varepsilon, \varepsilon)$ para todo $q_0 \in M$.

Definamos para cada $t \in (-\varepsilon, \varepsilon)$ a aplicação

$$\begin{aligned} P^t : M &\rightarrow M \\ q_0 &\mapsto P^t(q_0) = q(t, q_0), \end{aligned}$$

onde $q(t, q_0)$ é solução do problema de Cauchy com condição inicial $q(0) = q_0$.

Devido a unicidade e diferenciabilidade de $q(t)$ em cada ponto, temos que P^t está bem definida e é diferenciável.

Considere a família de aplicações diferenciáveis

$$\mathcal{P} = \{P^t \mid t \in (-\varepsilon, \varepsilon)\}.$$

Essa família tem as seguintes propriedades: se $t, s, t + s \in (-\varepsilon, \varepsilon)$ então

$$(i) P^t \circ P^s = P^s \circ P^t = P^{t+s};$$

$$(ii) P^0 = Id;$$

$$(iii) P^t \circ P^{-t} = Id;$$

$$(iv) P^t = (P^{-t})^{-1}.$$

Para mostrar o item (i), considere as curvas

$$\alpha(t) = P^t \circ P^s(q_0) = P^t(P^s(q_0)) = q(t, P^s(q_0))$$

$$\beta(t) = P^{t+s}(q_0) = q(t+s, q_0).$$

Temos que

$$\begin{aligned} \dot{\alpha}(t) &= \frac{d}{dt} P^t(P^s(q_0)) = \dot{q}(t, P^s(q_0)) = V(q(t, P^s(q_0))) \\ &= V(\alpha(t)) \end{aligned}$$

e também,

$$\begin{aligned} \dot{\beta}(t) &= \frac{d}{dt} P^{t+s}(q_0) = \dot{q}(t+s, q_0) = V(q(t+s, q_0)) \\ &= V(\beta(t)). \end{aligned}$$

Logo, $\alpha(t)$ e $\beta(t)$ são soluções do mesmo problema de Cauchy no ponto $P^s(q_0)$ pois, $\alpha(0) = P^s(q_0) = \beta(0)$. Pela unicidade de solução, $\alpha(t) = \beta(t)$ para todo $t \in (-\varepsilon, \varepsilon)$. Portanto $P^t \circ P^s = P^{t+s}$. O item (ii) segue direto da definição de P^t , enquanto que os itens (iii) e (iv) são consequências de (i). Como consequência, para cada $t \in (-\varepsilon, \varepsilon)$ temos que P^t é um difeomorfismo. A família \mathcal{P} é chamada de um **grupo a 1-parâmetro**.

Definição 1.19 Para um campo de vetores $V \in \text{Vec}(M)$ completo, a aplicação

$$t \mapsto P^t$$

é chamada de **fluxo gerado por** V e neste caso denotaremos $P^t(q_0) = e^{tV}(q_0)$.

Exemplo 1.20 Seja A uma matriz real de ordem $n \times n$ e considere o campo (linear) em \mathbb{R}^n definido por

$$V(x) = Ax.$$

A solução do problema de Cauchy

$$\begin{cases} \dot{x} &= V(x) \\ x(0) &= x_0, \end{cases}$$

é dada por $x(t) = e^{tA}x_0$. Neste caso, $P^t(x) = e^{tA}x_0$ satisfaz as propriedades (i)-(iv) da observação anterior, pois são as propriedades básicas da exponencial de matrizes.

Temos o seguinte resultado importante sobre campos de vetores completos.

Proposição 1.21 Suponha que exista $\varepsilon > 0$ tal que para qualquer ponto $q_0 \in M$ a solução do problema de Cauchy no ponto q_0 está definida para todo $t \in (-\varepsilon, \varepsilon)$. Então o campo de vetores é completo.

Demonstração: De acordo com a observação anterior, considere a família de aplicações

$$\mathcal{P} = \{P^t \mid t \in (-\varepsilon, \varepsilon)\}.$$

Para concluir que o campo de vetores é completo, basta mostrar que o intervalo $(-\varepsilon, \varepsilon)$ pode ser estendido a \mathbb{R} . De fato, para qualquer $t_0 \in \mathbb{R}$ fixado, existem $n \in \mathbb{N}$ e

$0 \leq r < \frac{\varepsilon}{2}$ tais que:

$$\pm t_0 = \pm \frac{\varepsilon}{2}n + r.$$

Logo,

$$P^{t_0}(q_0) = P^{\pm \frac{\varepsilon}{2}n+r}(q_0) = \underbrace{P^{\pm \frac{\varepsilon}{2}} \circ P^{\pm \frac{\varepsilon}{2}} \circ \dots \circ P^{\pm \frac{\varepsilon}{2}}}_{n\text{-vezes}} \circ P^r(q_0).$$

Portanto, $P^{t_0} \in \mathcal{P}$ para qualquer $t_0 \in \mathbb{R}$. Note que se $t_0 = 0$, então $n = r = 0$. Logo, $P^0(q_0) = q_0$. ■

Proposição 1.22 *Sejam $V \in \text{Vec}(M)$ e $K \subset M$ um subconjunto compacto. Então existe $\varepsilon_K > 0$ tal que, para todo $q_0 \in K$ a solução $q(t, q_0)$ do problema de Cauchy está definida para todo $t \in (-\varepsilon_K, \varepsilon_K)$.*

Demonstração: Pelo Teorema 1.12, para cada ponto $q \in M$ existe uma solução

$$q(t, q) : (a_q, b_q) \rightarrow M, \text{ com } a_q < 0 < b_q.$$

Pela Observação (1.13) temos que o domínio da aplicação q pode ser escolhido continuamente dependendo do ponto, isto é, a aplicação que associa a cada ponto $q \in M$ um domínio $(a_q, b_q) \subset \mathbb{R}$:

$$q \mapsto (a_q, b_q)$$

é contínua. Defina $\varepsilon_q = \min\{|a_q|, |b_q|\} > 0$ e a função:

$$\varphi : K \rightarrow \mathbb{R}$$

$$q \mapsto \varepsilon_q.$$

Temos que φ é contínua e definida em um subconjunto compacto K . Assim, φ assume um valor mínimo $\varepsilon_{q_0} > 0$ em algum $q_0 \in K$. Seja $\varepsilon_K = \varepsilon_{q_0}$. Isto significa que,

$$(-\varepsilon_K, \varepsilon_K) \subset (a_q, b_q) \quad \forall q \in K.$$

Portanto, para cada ponto $q \in K$, a solução do problema de Cauchy $q(t, q)$ está definida para todo $t \in (-\varepsilon_K, \varepsilon_K)$. ■

Corolário 1.23 *Se M é uma variedade compacta, então qualquer campo de vetores $V \in \text{Vec}(M)$*

é completo.

Corolário 1.24 *Suponha que $V \in \text{Vec}(M)$ tem suporte compacto, ou seja:*

$$\text{supp}(V) = \overline{\{q \in M \mid V(q) \neq 0\}}$$

é compacto. Então V é completo.

Demonstração: Como o suporte de V é compacto, pela Proposição 1.22, existe $\varepsilon > 0$ tal que para cada ponto $q \in \text{supp}(V)$ a solução $q(t, q)$ do problema de Cauchy está definida para todo $t \in (-\varepsilon, \varepsilon)$. Pela Proposição 1.21, V restrito ao $\text{supp}(V)$ é completo.

Para pontos

$$q \in M \setminus \text{supp}(V),$$

temos que $V(q) = 0$.

Consequentemente, o campo V é nulo fora do suporte de V . Logo, toda solução é constante. Portanto V é completo. ■

Observação 1.25 *Assumiremos no restante do texto que todos os campos de vetores considerados são completos.*

1.3 Sistemas de controle e colchete de Lie

Apresentamos até o momento conceitos básicos e resultados fundamentais a respeito de uma E.D.O determinada por um campo de vetores V sobre uma variedade M . Um resultado essencial apresentado foi o teorema de existência e unicidade, para cada E.D.O, ou seja, dado um ponto inicial, existe uma única solução em cada campo de vetores V definido. Com isto, podemos pensar em uma família de campos de vetores sobre M . O estudo das propriedades qualitativas de um sistema de controle depende das propriedades dos campos desta família e as interações entre eles. Neste contexto, surge de maneira natural uma ferramenta para entender essas interações, o colchete de Lie.

Definição 1.26 *Sejam M uma variedade e V_u uma família de campos de vetores completos*

sobre M . Um sistema de controle em M é um sistema dinâmico da forma:

$$\Sigma : \dot{q} = V_u(q). \quad (1.3-7)$$

A família V_u é parametrizada por controles $u \in \mathcal{U}$, onde \mathcal{U} é em geral um subconjunto de uma variedade, e chamada *espaço de parâmetros de controles*. Os pontos $q \in M$ são chamados de *estado* e M de *espaço estado* do sistema de controle.

Intuitivamente, podemos trocar a dinâmica do sistema em qualquer momento através dos controles u e, para cada u fixado, o campo V_u gera um fluxo, que denotaremos por P_u^t .

Uma questão que surge é: podemos, a partir de um estado inicial $q_0 \in M$, alcançar um outro estado final $q_1 \in M$ através dos fluxos, usando os controles $u \in \mathcal{U}$, mudando de um para o outro quando necessário?

Suponha que iniciemos de um ponto $q_0 \in M$ e considere a seguinte estratégia: escolhamos um controle $u_1 \in \mathcal{U}$ e depois outro controle $u_2 \in \mathcal{U}$. Para u_1 temos a E.D.O:

$$\dot{q} = V_{u_1}(q), \quad (1.3-8)$$

a qual gera um fluxo $P_{u_1}^t(q_0)$. O conjunto de pontos que podemos alcançar mediante o controle u_1 é:

$$\{P_{u_1}^t(q_0) \mid t \geq 0\},$$

que é uma semi trajetória. Para o controle u_2 , obtemos a E.D.O:

$$\dot{q} = V_{u_2}(q), \quad (1.3-9)$$

que fornece um fluxo $P_{u_2}^s(q')$ para cada $q' \in \{P_{u_1}^t(q_0) \mid t \geq 0\}$. Deste modo, o conjunto de pontos que podemos alcançar usando a nossa estratégia com os controles u_1 e u_2 é:

$$\{P_{u_2}^s \circ P_{u_1}^t(q_0) \mid s, t \geq 0\},$$

que é parte de uma superfície de dimensão 2.

Esta discussão motiva a definição seguinte:

Definição 1.27 O conjunto de pontos atingíveis do sistema de controle (1.3-7), com controles

constantes por partes, a partir de um ponto $q_0 \in M$ num tempo $t \geq 0$ é definido como:

$$\mathcal{A}_{q_0}(t) = \left\{ P_{u_k}^{t_k} \circ P_{u_{k-1}}^{t_{k-1}} \circ \dots \circ P_{u_2}^{t_2} \circ P_{u_1}^{t_1}(q_0); \sum_{i=1}^k t_i = t, t_i \geq 0, k \in \mathbb{N}, u_i \in \mathcal{U} \right\} \subset M. \quad (1.3-10)$$

O conjunto de pontos atingíveis a partir de q_0 é definido como

$$\mathcal{A}_{q_0} = \bigcup_{t \geq 0} \mathcal{A}_{q_0}(t). \quad (1.3-11)$$

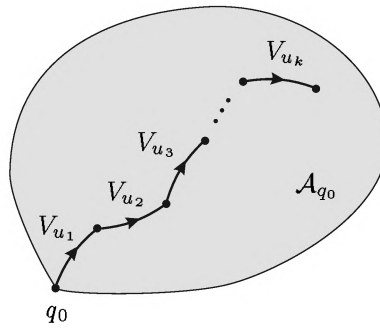


Figura 1.2: Conjunto atingível \mathcal{A}_{q_0} .

Observação 1.28 Suponha que o espaço dos parâmetros de controle \mathcal{U} consiste de dois controles, isto é, $\mathcal{U} = \{u_1, u_2\}$. Assim, o conjunto de pontos atingíveis para um tempo $t \geq 0$ arbitrário, torna-se:

$$\mathcal{A}_{q_0} = \left\{ P_{u_2}^{t_k} \circ P_{u_1}^{t_{k-1}} \circ \dots \circ P_{u_2}^{t_2} \circ P_{u_1}^{t_1}(q_0); \sum_{i=1}^k t_i = t, t_i \geq 0, k \in \mathbb{N}, u_i \in \mathcal{U} \right\}.$$

Aqui surge uma questão sobre a comutatividade dos fluxos $P_{u_2}^s, P_{u_1}^t(q_0)$. Quando os fluxos comutam, obtemos que: para $k \in \mathbb{N}$ e $u_i \in \mathcal{U}$,

$$\begin{aligned} \mathcal{A}_{q_0} &= P_{u_2}^{t_k} \circ P_{u_1}^{t_{k-1}} \circ \dots \circ P_{u_2}^{t_2} \circ P_{u_1}^{t_1}(q_0) \\ &= P_{u_2}^{t_k} \circ P_{u_2}^{t_{k-2}} \circ \dots \circ P_{u_1}^{t_3} \circ P_{u_1}^{t_1}(q_0) \\ &= P_{u_2}^{t_k+t_{k-2}+\dots+t_2} \circ P_{u_1}^{t_{k-1}+t_{k-3}+\dots+t_1}(q_0) \\ &= P_{u_2}^s \circ P_{u_1}^t(q_0), \end{aligned}$$

onde $s = t_k + t_{k-2} + \dots + t_2$, e $t = t_{k-1} + t_{k-3} + \dots + t_1$, que é parte de uma superfície de dimensão 2.

De modo geral, se temos $k \geq 2$ controles, e se os fluxos correspondentes a esses campos comutam, então \mathcal{A}_{q_0} é parte de uma variedade de dimensão k .

Esta discussão nos motiva a procurar propriedades comutativas dos campos de vetores sem conhecer seus fluxos explicitamente. Para tanto, sejam $V_1, V_2 \in \text{Vec}(M)$. Considere a aplicação,

$$\alpha(s, t) = P_1^{-t} \circ P_2^s \circ P_1^t(q),$$

para $q \in M$ fixado e $s, t \in \mathbb{R}$. Se os fluxos comutam, então obtemos uma curva

$$\alpha(s, t) = P_2^s(q),$$

que não depende de t .

É natural esperar que termos de baixa ordem na expansão de Taylor de $\alpha(s, t)$ possam dar pistas sobre a comutatividade dos fluxos dos campos de vetores V_1 e V_2 . Vejamos isso em um exemplo.

Exemplo 1.29 *Sejam A e B matrizes reais $n \times n$ e considere os campos em \mathbb{R}^n $V_1(x) = Ax$ e $V_2(x) = Bx$. Então $\alpha(s, t) = e^{-tA}e^{sB}e^{tA}x$. Um cálculo direto mostra que as derivadas de primeira ordem e as derivadas puras de segunda ordem, calculadas em $s = t = 0$, são dadas por*

$$\frac{\partial \alpha}{\partial s} = Bx = V_2(x), \quad \frac{\partial \alpha}{\partial t} = 0, \quad \frac{\partial^2 \alpha}{\partial t^2} = 0 \quad e \quad \frac{\partial^2 \alpha}{\partial s^2} = B^2x = V_2(P_2^s(x)),$$

as quais não fornecem qualquer pista sobre a comutatividade dos fluxos. Agora, a derivada mista de ordem 2

$$\frac{\partial^2 \alpha}{\partial t \partial s} = -Ae^{-tA}Be^{sB}e^{tA}\Big|_{(0,0)}x + e^{-tA}Be^{sB}Ae^{tA}\Big|_{(0,0)}x = (-AB + BA)x$$

dá uma ideia precisa da comutatividade dos fluxos.

O exemplo anterior motiva a seguinte definição.

Definição 1.30 *Sejam $V_1, V_2 \in \text{Vec}(M)$ com fluxos P_1^t, P_2^s , respectivamente. O campo de vetores denotado por*

$$[V_1, V_2] \in \text{Vec}(M),$$

definido por

$$[V_1, V_2](q) = \left. \frac{\partial^2}{\partial s \partial t} \right|_{s=t=0} P_1^{-t} \circ P_2^s \circ P_1^t(q) \in T_q M,$$

é chamado o **colchete de Lie** dos campos V_1 , e V_2 .

O colchete de Lie fornece informações sobre a comutatividade dos campos de vetores, no sentido que o campo $[V_1, V_2]$ é nulo se, e somente se, os campos V_1 e V_2 comutam. Um resultado relevante para este texto é

Proposição 1.31 $[V_1, V_2] = 0$ se, e somente se, seus fluxos correspondentes comutam, isto é, $P_1^t \circ P_2^s = P_2^{-s} \circ P_1^{-t}$.

Demonstração: Ver [12] Apêndice A. ■

A seguinte fórmula é muito útil para calcular o colchete de Lie em coordenadas locais.

Proposição 1.32 Para $V_1, V_2 \in \text{Vec}(\mathbb{R}^n)$, o colchete de Lie é dado por

$$[V_1, V_2](x) = \frac{dV_2}{dx} V_1(x) - \frac{dV_1}{dx} V_2(x),$$

onde $\frac{dV_1}{dx}$ e $\frac{dV_2}{dx}$ denotam respectivamente as matrizes jacobianas de V_1 e V_2 .

Demonstração: Ver [12] Apêndice A. ■

Exemplo 1.33 Sejam $f, g \in \text{Vec}(\mathbb{R}^2)$ definidos por

$$f(x) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad g(x) = \begin{bmatrix} 0 \\ x_1 \end{bmatrix}.$$

Pela Proposição 1.32, obtemos com um simples cálculo que,

$$[f, g](x) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

De maneira mais geral, para $f, g \in \text{Vec}(\mathbb{R}^n)$, se $f(x) = Ax$ e $g(x) = b$, onde A é uma matriz $n \times n$, e b constante. Temos que

$$[f, g](x) = [Ax, b](x) = 0 - Ab = -Ab.$$

1.4 Variedades e o espaço $\mathcal{C}^\infty(M)$

Introduziremos nesta seção uma importante relação entre uma variedade M e a álgebra $\mathcal{C}^\infty(M)$. Mais precisamente, exibiremos uma relação entre pontos, difeomorfismos e campos de vetores sobre M com funcionais, operadores e derivações lineares sobre $\mathcal{C}^\infty(M)$ (ver [1]). Isto nos permite estudar fluxos e sistemas não lineares através de seus representantes lineares em $\mathcal{C}^\infty(M)$.

1.4.1 Pontos, difeomorfismos e campos de vetores

Recordemos que $\mathcal{C}^\infty(M)$ representa o espaço de todas as aplicações diferenciáveis de M em \mathbb{R} . Na álgebra $\mathcal{C}^\infty(M)$ estão definidas as operações:

$$(a + b)(q) = a(q) + b(q)$$

$$(a \cdot b)(q) = a(q) \cdot b(q)$$

$$(\alpha \cdot a)(q) = \alpha \cdot a(q),$$

para quaisquer $a, b \in \mathcal{C}^\infty(M)$, $q \in M$, $\alpha \in \mathbb{R}$.

Começaremos com uma identificação entre M e $\mathcal{L}(\mathcal{C}^\infty(M), \mathbb{R})$, onde $\mathcal{L}(\mathcal{C}^\infty(M), \mathbb{R})$ denota o espaço de todos os funcionais lineares de $\mathcal{C}^\infty(M)$.

Associamos a cada ponto $q \in M$ um funcional linear

$$\begin{aligned} \hat{q}: \mathcal{C}^\infty(M) &\rightarrow \mathbb{R} \\ a &\mapsto \hat{q}(a) = a(q). \end{aligned}$$

Obtemos facilmente que \hat{q} está bem definido e único. Temos também que \hat{q} é um homomorfismo das álgebras $\mathcal{C}^\infty(M)$ e \mathbb{R} com as operações definidas em $\mathcal{C}^\infty(M)$:

$$\hat{q}(a + b) = (a + b)(q) = a(q) + b(q) = \hat{q}(a) + \hat{q}(b),$$

$$\hat{q}(a \cdot b) = (a \cdot b)(q) = a(q) \cdot b(q) = \hat{q}(a) \cdot \hat{q}(b),$$

$$\hat{q}(\alpha \cdot a) = (\alpha \cdot a)(q) = \alpha \cdot a(q) = \alpha \cdot \hat{q}(a),$$

para todo $a, b \in \mathcal{C}^\infty(M)$ e $\alpha \in \mathbb{R}$.

A proposição a seguir mostra que a correspondência entre pontos e funcionais lineares é biunívoca. Para demonstrá-la, é necessário o conceito de partições da unidade. Enunciaremos várias definições e um teorema, do qual omitiremos a demonstração. O leitor pode aprofundar mais os conceitos mencionados abaixo em [8] Capítulo 8, Seção 3.

Definição 1.34 Dizemos que uma família de subconjuntos $\mathcal{L} = (L_\alpha)_{\alpha \in \Lambda}$ de M é **localmente finita** se, para cada ponto $q \in M$, existe uma vizinhança $V_q \subset M$ contendo q tal que V_q intersecta um número finito de subconjuntos de \mathcal{L} .

Definição 1.35 Seja M uma variedade. Uma **partição da unidade** em M é uma família de funções $(\psi_\alpha)_{\alpha \in \Lambda}$ de $C^\infty(M)$, tais que

1. Para todo ponto $q \in M$ e $\alpha \in \Lambda$, $\psi_\alpha(q) \geq 0$;
2. A família $\mathcal{S} = (\text{supp}(\psi_\alpha))_{\alpha \in \Lambda}$ é localmente finita em M ;
3. Para todo $q \in M$ tem-se $\sum_{\alpha \in \Lambda} \psi_\alpha(q) = 1$.

Definição 1.36 Seja $\mathcal{C} = (C_\beta)_{\beta \in \Gamma}$ uma cobertura de M . Dizemos que uma partição da unidade $\sum_{\alpha \in \Lambda} \psi_\alpha(q) = 1$ está **subordinada** à cobertura \mathcal{C} se, para cada $\alpha \in \Lambda$, existe um $\beta \in \Gamma$ tal que $\text{supp}(\psi_\alpha) \subset C_\beta$. Dizemos também que uma partição $\sum_{\alpha \in \Lambda} \psi_\alpha(q) = 1$ da unidade é **estritamente subordinada** à cobertura $\mathcal{C} = (C_\beta)_{\beta \in \Gamma}$ se $\Gamma = \Lambda$ e, além disso, $\text{supp}(\psi_\alpha) \subset C_\alpha$ para todo $\alpha \in \Lambda$.

Teorema 1.37 Dada uma cobertura aberta $\mathcal{C} = (C_\beta)_{\beta \in \Gamma}$ de M , existe uma partição da unidade estritamente subordinada à cobertura \mathcal{C} .

Lema 1.38 Em qualquer variedade diferenciável M existe uma função $h \in C^\infty(M)$ tal que, para todo $\varepsilon > 0$, existe um subconjunto compacto $K \subset M$ no qual

$$h(q) > \varepsilon, \quad q \in M \setminus K.$$

Demonstração: Seja $\psi_k, k \in \mathbb{N}$, uma partição da unidade em M : $\psi_k \in C^\infty(M)$ para cada $k \in \mathbb{N}$, os suportes $\text{supp}(\psi_k) \subset M$ são compactos e formam uma cobertura localmente

finita de M e $\sum_{k=1}^{\infty} \psi_k \equiv 1$. Basta definir

$$h(q) = \sum_{k=1}^{\infty} k\psi_k(q).$$

■

Finalmente, estamos aptos para demonstrar o próximo resultado.

Proposição 1.39 *Seja $\varphi : C^\infty(M) \rightarrow \mathbb{R}$ um homomorfismo não trivial. Então existe um ponto $q \in M$ tal que $\hat{q} = \varphi$.*

Demonstração: Seja $\varphi : C^\infty(M) \rightarrow \mathbb{R}$ um funcional linear (homomorfismo não trivial e sobrejetor). O conjunto

$$\ker(\varphi) = \{f \in C^\infty(M) \mid \varphi(f) = 0\},$$

é um ideal maximal em $C^\infty(M)$. De fato, claramente este conjunto é um ideal de $C^\infty(M)$. Agora, pelo Teorema do Homomorfismo, temos que

$$\frac{C^\infty(M)}{\ker(\varphi)} \text{ é isomorfo a } \mathbb{R}.$$

Como \mathbb{R} é corpo, temos que $\frac{C^\infty(M)}{\ker(\varphi)}$ é corpo e, portanto $\ker(\varphi)$ é maximal. Considere o ideal de $C^\infty(M)$ definido por

$$I_q = \{f \in C^\infty(M) \mid f(q) = 0\},$$

para qualquer $q \in M$.

Afirmção: $\ker(\varphi) \subset I_q$ para algum $q \in M$. De fato, suponha que $\ker(\varphi) \not\subset I_q$ para todo q , isto é, para cada ponto $q \in M$ existe uma função $f_q \in \ker(\varphi)$ tal que $f_q \notin I_q$, implicando em $f_q(q) \neq 0$. Sem perda de generalidade, podemos assumir que $f_q(q) > 0$. Logo, pela continuidade de f_q obtemos:

$$\forall q \in M \quad \exists f_q \in \ker(\varphi), O_q \subset M \text{ tal que } f_q(q) > 0, q \in O_q. \quad (1.4-12)$$

Fixe a função $h \in C^\infty(M)$ satisfazendo o Lema 1.38. Denote $\varphi(h) = \alpha > 0$. Então

$\varphi(h - \alpha) = 0$. Logo,

$$(h - \alpha) \in \ker(\varphi). \quad (1.4-13)$$

Pelo Lema 1.38, temos que para este α existe um subconjunto compacto $K \subset M$ tal que

$$h(q) - \alpha > 0 \text{ para todo } q \in M \setminus K.$$

Como K é compacto, K admite uma subcobertura finita através dos abertos dados em (1.4-12),

$$K \subset \bigcup_{i=1}^n O_{q_i}.$$

Considere uma partição da unidade $\psi_0, \psi_1, \dots, \psi_n \in \mathcal{C}^\infty(M)$, estritamente subordinada à cobertura

$$M \setminus K, O_{q_1}, O_{q_2}, \dots, O_{q_n},$$

de M . Desta forma, definimos a função sobre M , pondo:

$$g = \psi_0(h - \alpha) + \psi_1 f_{q_1} + \dots + \psi_n f_{q_n}.$$

Como $\text{supp}(\psi_i) \subset O_{q_i}$ temos que ao menos um membro da soma de g é maior do que zero. Assim, $g > 0$. Como $(h - \alpha), f_{q_i} \in \ker(\varphi)$ para $i = 1, \dots, n$, e φ é homomorfismo, temos que $g \in \ker(\varphi)$. Porém, sendo $g > 0$, podemos escrever

$$1 = \varphi\left(g \frac{1}{g}\right) = \varphi(g) \varphi\left(\frac{1}{g}\right).$$

Segue que $\varphi(g) \neq 0$. Esta é uma contradição, obtida por supor (1.4-12). Portanto, $\ker(\varphi) \subset I_{q_0}$, $q_0 \in M$. Como $I_{q_0} \subsetneq \mathcal{C}^\infty(M)$ e $\ker(\varphi)$ é ideal maximal, obtemos a igualdade $\ker(\varphi) = I_{q_0}$. Defina $\widehat{q}_0 : \mathcal{C}^\infty(M) \rightarrow \mathbb{R}$, dada por $\widehat{q}_0(f) = f(q_0)$. Para mostrar que $\varphi = \widehat{q}_0$, basta mostrar a afirmação seguinte.

Afirmção: $\ker(\varphi - \widehat{q}_0) = \mathcal{C}^\infty(M)$. De fato, temos

$$\ker(\varphi) = I_{q_0} \subset \ker(\varphi - \widehat{q}_0), \quad (1.4-14)$$

assim, basta mostrar que a inclusão acima é estrita exibindo uma função $f \in \ker(\varphi - \widehat{q}_0)$ e $f \notin I_{q_0}$. Seja $f \in \mathcal{C}^\infty(M)$ dada por $f(q) = 1, \forall q \in M$. Claramente $f \notin I_{q_0}$. Temos

que, se $\varphi(f) = k \in \mathbb{R}$, então $k = 1$. Com efeito, se $\varphi(f) = k$ então $\varphi(f - k) = 0$, conseqüentemente $(f - k) \in \ker(\varphi) = I_{q_0}$, isto é, $(f - k)(q_0) = 0$, que implica em, $1 = f(q_0) = k$. Finalmente,

$$(\varphi - \widehat{q}_0)(f) = \varphi(f) - \widehat{q}_0(f) = \varphi(f) - f(q_0) = 1 - 1 = 0,$$

mostrando que $f \in \ker(\varphi - \widehat{q}_0)$. Segue que a inclusão (1.4-14) é estrita. Sendo $\ker(\varphi)$ um ideal maximal, concluímos que $\ker(\varphi - \widehat{q}_0) = \mathcal{C}^\infty(M)$, e portanto $\varphi = \widehat{q}_0$. ■

Apresentaremos agora uma identificação entre os conjuntos $\text{Diff}(M)$ e $\text{Aut}(\mathcal{C}^\infty(M))$, onde $\text{Aut}(\mathcal{C}^\infty(M))$ é o conjunto de todos os automorfismos de $\mathcal{C}^\infty(M)$. Para qualquer $P \in \text{Diff}(M)$ definimos:

$$\begin{aligned} \widehat{P} : \mathcal{C}^\infty(M) &\rightarrow \mathcal{C}^\infty(M) \\ a &\mapsto \widehat{P}(a) : M \rightarrow \mathbb{R} \\ q &\mapsto \widehat{P}(a)(q) = a(P(q)), \end{aligned}$$

para todo $a \in \mathcal{C}^\infty(M)$ e $q \in M$. Mostremos que \widehat{P} assim definido é de fato um automorfismo.

Para quaisquer $a, b \in \mathcal{C}^\infty(M)$, $q \in M$ e $\alpha \in \mathbb{R}$ temos:

$$\begin{aligned} \widehat{P}(a + b)(q) &= (a + b)(P(q)) = a(P(q)) + b(P(q)) = \widehat{P}(a)(q) + \widehat{P}(b)(q), \\ \widehat{P}(a \cdot b)(q) &= (a \cdot b)(P(q)) = a(P(q)) \cdot b(P(q)) = \widehat{P}(a)(q) \cdot \widehat{P}(b)(q), \\ \widehat{P}(\alpha \cdot a)(q) &= (\alpha \cdot a)(P(q)) = \alpha(P(q)) \cdot a(P(q)) = \alpha \cdot \widehat{P}(a)(q). \end{aligned}$$

Além disso,

$$\begin{aligned} \widehat{P}(a) = \widehat{P}(b) &\Rightarrow \widehat{P}(a)(q) = \widehat{P}(b)(q) \quad \forall q \in M \\ &\Rightarrow a(P(q)) = b(P(q)) \quad \forall q \in M \\ &\Rightarrow a = b. \end{aligned}$$

Naturalmente, a inversa de \widehat{P} é dada por \widehat{P}^{-1} .

Esta identificação é também sobrejetora, conforme mostra a

Proposição 1.40 *Para qualquer $A \in \text{Aut}(\mathcal{C}^\infty(M))$ existe $P \in \text{Diff}(M)$ tal que $A = \widehat{P}$.*

Demonstração: Seja $q \in M$ um ponto qualquer e considere o funcional linear correspondente

$$\widehat{q} : \mathcal{C}^\infty(M) \rightarrow \mathbb{R}.$$

Temos o funcional linear

$$\widehat{q} \circ A : \mathcal{C}^\infty(M) \rightarrow \mathbb{R},$$

onde A é o automorfismo dado. Pela Proposição 1.39, existe um ponto $p \in M$ tal que

$$\widehat{p} = \widehat{q} \circ A.$$

Por [4] Capítulo 2, Seção 2.2, existe um difeomorfismo $P \in \text{Diff}(M)$ tal que $P(q) = p$.

Logo,

$$\begin{aligned} \widehat{q} \circ A(a) &= \widehat{p}(a) = \widehat{P(q)}(a) = a(P(q)) = \widehat{P}(a)(q) = \widehat{q}(\widehat{P}(a)) \\ &= \widehat{q} \circ \widehat{P}(a), \end{aligned}$$

ou seja, $\widehat{q} \circ A(a) = \widehat{q} \circ \widehat{P}(a)$ para todo $q \in M$. Portanto, $A = \widehat{P}$. ■

Vamos agora identificar os vetores tangentes com certos funcionais lineares. Dados $q \in M$ e $v \in T_q M$, seja $\gamma : I \rightarrow M$ uma curva tal que $\gamma(0) = q$ e $\dot{\gamma}(0) = v$. Defina o funcional $\widehat{v} : \mathcal{C}^\infty(M) \rightarrow \mathbb{R}$ por

$$\widehat{v}a = \left. \frac{d}{dt} \right|_{t=0} a(\gamma(t)).$$

Note que, pela definição da diferencial de uma função, temos

$$\widehat{v}a = D_q a \cdot v.$$

Além disso, dados $a, b \in \mathcal{C}^\infty(M)$, temos:

$$\begin{aligned} \widehat{v}(a \cdot b) &= \left. \frac{d}{dt} \right|_{t=0} ((a \cdot b)(\gamma(t))) \\ &= \left. \frac{d}{dt} \right|_{t=0} (a(\gamma(t)) \cdot b(\gamma(t))) \\ &= \left. \frac{d}{dt} \right|_{t=0} a(\gamma(t)) \cdot b(\gamma(0)) + a(\gamma(0)) \cdot \left. \frac{d}{dt} \right|_{t=0} b(\gamma(t)) \\ &= \widehat{v}a \cdot b(q) + a(q) \cdot \widehat{v}b \end{aligned}$$

A igualdade acima é conhecida como a *regra de Leibniz* no ponto q . Qualquer funcional $\xi : C^\infty(M) \rightarrow \mathbb{R}$ satisfazendo a regra de Leibniz em q é chamada *derivação em q* . O conjunto dessas derivações é denotado por $\text{Der}_q(M)$.

Do que vimos acima, a um dado vetor $v \in T_qM$ podemos associar uma derivação $\hat{v} \in \text{Der}_q(M)$, a qual será chamada de *derivada direcional ao longo de v* . Para mostrar que esta correspondência é biunívoca, precisaremos do seguinte resultado:

Lema 1.41 (Lema de Hadamard) *Seja $f : U \rightarrow \mathbb{R}$ uma função de classe C^∞ , onde $U \subset \mathbb{R}^n$ é aberto convexo contendo 0. Então, existem funções $b_i : U \rightarrow \mathbb{R}$ de classe C^∞ , $i = 1, 2, \dots, n$ tais que*

$$f(x) = f(0) + \sum_{i=1}^n x_i b_i(x),$$

para todo $x = (x_1, x_2, \dots, x_n) \in U$. Além disso, as funções b_i são dadas por

$$b_i(x) = \int_0^1 \frac{\partial f}{\partial x_i}(tx) dt.$$

Demonstração: Ver [9] Capítulo 3, seção 3.3. ■

Proposição 1.42 *Para toda derivação $\xi \in \text{Der}_q(M)$ existe $v \in T_qM$ tal que $\hat{v} = \xi$.*

Demonstração: Primeiramente, note que o valor de uma derivação em q aplicada em uma função depende somente dos valores de a em vizinhanças arbitrariamente pequenas de q , isto é, se para alguma vizinhança $O|q$ de q tivermos $a|_{O_q} = \tilde{a}|_{O_q}$, então $\xi a = \xi \tilde{a}$. De fato, tomando $b \in C^\infty(M)$ tal que $b_{M-O_q} \equiv 1$ e $b(q) = 0$, temos que $(\tilde{a} - a)b = \tilde{a} - a$ e assim:

$$\xi(\tilde{a} - a) = \xi((\tilde{a} - a)b) = \xi(\tilde{a} - a) \cdot b(q) + (\tilde{a} - a)(q) \cdot \xi(b) = 0.$$

Desta forma, a proposição é de natureza local e será provada em coordenadas.

Sejam (x_1, \dots, x_n) coordenadas locais sobre M centradas no ponto q . Vamos mostrar que existem $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ tais que

$$\xi = \sum_{i=1}^n \alpha_i \frac{\partial}{\partial x_i} \Big|_0.$$

Para isso, observe primeiro que

$$\xi(1) = \xi(1) \cdot 1 + 1 \cdot \xi(1) = 2\xi(1).$$

Assim, $\xi(1) = 0$ e, por linearidade, ξ se anula sobre funções constantes. Agora, dado $a \in \mathcal{C}^\infty(M)$, usamos o Lema de Hadamard para escrever

$$a(x) = a(0) + \sum_{i=1}^n b_i(x)x_i.$$

Assim,

$$\xi(a) = \sum_{i=1}^n \xi(b_i x_i) = \sum_{i=1}^n ((\xi b_i)x_i(0) + b_i(0)(\xi x_i)) = \sum_{i=1}^n \alpha_i \frac{\partial a}{\partial x_i}(0),$$

onde $\alpha_i = \xi x_i$ e foi usada a igualdade $b_i(0) = \frac{\partial a}{\partial x_i}(0)$. ■

Por fim, identificaremos o conjunto dos campos de vetores diferenciáveis em M com o conjunto das *derivações* em M . Uma derivação em M é um operador linear $D : \mathcal{C}^\infty(M) \rightarrow \mathcal{C}^\infty(M)$ tal que, para quaisquer $a, b \in \mathcal{C}^\infty(M)$,

$$D(a \cdot b) = D(a) \cdot b + a \cdot D(b).$$

O conjunto das derivações em M será denotado por $\text{Der}(M)$.

Para cada $V \in \text{Vec}(M)$ definimos

$$\begin{aligned} \widehat{V} : \mathcal{C}^\infty(M) &\rightarrow \mathcal{C}^\infty(M) \\ a &\mapsto \widehat{V}a : M \rightarrow \mathbb{R} \\ q &\mapsto \widehat{V}a(q) = D_q a \cdot V(q). \end{aligned}$$

Note que

$$\widehat{V}a(q) = \widehat{V(q)}a.$$

De fato,

$$\widehat{V(q)}a \stackrel{\text{def}}{=} \left. \frac{d}{dt} a(q(t)) \right|_{t=0} = D_q a \cdot V(q) = \widehat{V}a(q),$$

onde $q(t)$ é uma curva em M com $q(0) = q$ e $\dot{q}(0) = V(q)$. Note que \widehat{V} é uma derivação:

$$\begin{aligned}\widehat{V}(a \cdot b)(q) &= \widehat{V(q)}(a \cdot b) \\ &= (\widehat{V(q)}a) \cdot b(q) + a(q) \cdot (\widehat{V(q)}b) \\ &= \widehat{V}a \cdot b(q) + a(q) \cdot \widehat{V}b.\end{aligned}$$

Chamamos \widehat{V} de **derivada de Lie ao longo do campo V** .

A proposição seguinte nos garante a recíproca da discussão precedente.

Proposição 1.43 *Para toda derivação $D \in \text{Der}(M)$ existe um campo de vetores $V \in \text{Vec}(M)$ tal que $\widehat{V} = D$.*

Demonstração: Seja

$$D : \mathcal{C}^\infty(M) \rightarrow \mathcal{C}^\infty(M)$$

uma derivação. Seja $q \in M$ um ponto qualquer e consideremos o funcional linear

$$\widehat{q} : \mathcal{C}^\infty(M) \rightarrow \mathbb{R}.$$

Defina o funcional linear $d_q = \widehat{q} \circ D$. Mostremos que d_q satisfaz a regra de Leibniz no ponto q :

$$\begin{aligned}d_q(a \cdot b) &= \widehat{q} \circ D(a \cdot b) = \widehat{q}(D(a \cdot b)) \\ &= \widehat{q}((Da)b(q) + a(q)(Db)) \\ &= \widehat{q}(Da)b(q) + \widehat{q}(a(q)(Db)) \\ &= \widehat{q}(Da) \cdot b(q) + a(q) \cdot \widehat{q}(Db) \\ &= (d_q a)b(q) + a(q)(d_q b).\end{aligned}$$

Pela Proposição 1.42, existe um vetor $v \in T_q M$ tal que $\widehat{v} = d_q$. Como q é arbitrário, obtemos um campo de vetores $V \in \text{Vec}(M)$. ■

Nesta subseção identificamos, pontos, difeomorfismos, vetores e campos de vetores com objetos do espaço dos funcionais e operadores. Com isso, obtemos uma versão do colchete de Lie para derivações.

Definição 1.44 *Sejam $\widehat{V}, \widehat{W} \in \text{Der}(M)$ derivações. O colchete de Lie dos operadores \widehat{V} e \widehat{W} é*

dado por

$$[\widehat{V}, \widehat{W}] = \widehat{V} \circ \widehat{W} - \widehat{W} \circ \widehat{V}. \quad (1.4-15)$$

1.4.2 Ação de difeomorfismos sobre campos de vetores

Nesta subseção usaremos as ferramentas apresentadas na seção anterior para introduzir um tipo especial de funcional e operador em $\text{Der}_q(M)$ e $\text{Der}(M)$ respectivamente, que são induzidos por um difeomorfismo, os quais serão fundamentais para o restante do texto.

Sejam M uma variedade, $v \in T_q M$ e $P \in \text{Diff}(M)$. Considere uma curva $q(t) \in M$ tal que

$$q(0) = q \quad \text{e} \quad \dot{q}(0) = v.$$

Com isso, obtemos uma curva $P(q(t)) \in M$ com vetor velocidade

$$\left. \frac{d}{dt} P(q(t)) \right|_{t=0} = D_q P \cdot v = P_* v \in T_{P(q)} M.$$

O vetor $P_* v$ tem um correspondente em $\text{Der}_{P(q)}(M)$.

Afirmamos que

$$\widehat{P_* v} = \widehat{v} \circ \widehat{P}. \quad (1.4-16)$$

De fato, a derivada direcional ao longo do vetor $\widehat{P_* v}$, foi definida por

$$\begin{aligned} \widehat{P_* v} a &= \left. \frac{d}{dt} \right|_{t=0} a(P(q(t))) \\ &= D_{P(q(t))}(a \circ P) \cdot \left. \frac{d}{dt} \right|_{t=0} q(t) \\ &= \left(\left. \frac{d}{dt} \right|_{t=0} q(t) \right) a \circ P \\ &= \widehat{v}(\widehat{P}a) \\ &= (\widehat{v} \circ \widehat{P})a, \quad \forall a \in \mathcal{C}^\infty(M). \end{aligned}$$

Portanto,

$$\begin{aligned} \widehat{P_* v} : \mathcal{C}^\infty(M) &\rightarrow \mathbb{R} \\ a &\mapsto \widehat{P_* v} a = \widehat{v} \circ \widehat{P} a. \end{aligned}$$

Agora, dado $V \in \text{Vec}(M)$ vamos determinar uma expressão para $P_* V$ como uma

derivação em M . Afirmamos que

$$\widehat{P}_* \widehat{V} = \widehat{P}^{-1} \circ \widehat{V} \circ \widehat{P}. \quad (1.4-17)$$

Para provar isso, vamos destacar algumas igualdades.

(i) $\widehat{v} \circ \widehat{P} = \widehat{P}_* \widehat{v}$, como já foi provado em (1.4-16).

(ii) $\widehat{q} \circ \widehat{P} = \widehat{P}(\widehat{q})$. De fato,

$$(\widehat{q} \circ \widehat{P})a = \widehat{q}(\widehat{P}(a)) = \widehat{P}(a)(q) = a(P(q)) = \widehat{P}(\widehat{q})a.$$

(iii) $\widehat{q} \circ \widehat{V} = \widehat{V}(\widehat{q})$. De fato,

$$(\widehat{q} \circ \widehat{V})a = \widehat{q}(\widehat{V}a) = \widehat{V}a(q) = D_q a(V(q)).$$

Por outro lado,

$$\widehat{V}(\widehat{q})a = \left. \frac{d}{dt} a(q(t)) \right|_{t=0} = D_q a(V(q)).$$

(iv) $P_* \widehat{V}(P(q)) = P_*(\widehat{V}(\widehat{q}))$. De fato, por (1.2-3) temos

$$P_* V(P(q)) = D_q P(V(q)).$$

Por outro lado,

$$P_*(\widehat{V}(\widehat{q})) = \left. \frac{d}{dt} P(q(t)) \right|_{t=0} = D_q P(V(q)).$$

Como $P_* V(P(q)) = P_*(\widehat{V}(\widehat{q}))$, então (iv) segue.

Mostraremos agora (1.4-17). Para todo $q \in M$, temos

$$\widehat{q} \circ \widehat{P} \circ \widehat{P}_* \widehat{V} \stackrel{(ii)}{=} \widehat{P}(\widehat{q}) \circ \widehat{P}_* \widehat{V} \stackrel{(iii)}{=} P_* \widehat{V}(P(q)) \stackrel{(iv)}{=} P_*(\widehat{V}(\widehat{q})) \stackrel{(i)}{=} \widehat{V}(\widehat{q}) \circ \widehat{P} \stackrel{(iii)}{=} \widehat{q} \circ \widehat{V} \circ \widehat{P},$$

de onde obtemos a igualdade desejada. Vemos assim que difeomorfismos agem em campos de vetores por semelhanças.

Observação 1.45 *Daqui em diante, omitiremos o símbolo circunflexo "ˆ" sobre os funcionais e*

operadores. Isto não causará ambiguidade, visto que $q \circ V, V \circ W$ por exemplo, só fazem sentido se forem operadores. O objetivo é simplificar a notação.

O operador P_*V tem a propriedade de preservar composições. De fato,

$$\begin{aligned} P_*(V \circ W)a &= (P^{-1} \circ V \circ W \circ P)a \\ &= (P^{-1} \circ V \circ (P \circ P^{-1}) \circ W \circ P)a \\ &= ((P^{-1} \circ V \circ P) \circ (P^{-1} \circ W \circ P))a \\ &= (P_*V \circ P_*W)a, \quad \forall a \in \mathcal{C}^\infty(M). \end{aligned}$$

Conseqüentemente, P_*V preserva o colchete de Lie:

$$\begin{aligned} P_*[V, W]a &= (P^{-1} \circ [V, W] \circ P)a \\ &= (P^{-1} \circ (V \circ W - W \circ V) \circ P)a \\ &= (P^{-1} \circ V \circ W \circ P - P^{-1} \circ W \circ V \circ P)a \\ &= (P_*V \circ P_*W - P_*W \circ P_*V)a \\ &= [P_*V, P_*W]a, \quad \forall a \in \mathcal{C}^\infty(M). \end{aligned} \tag{1.4-18}$$

Seja $T \in \text{Aut}(\mathcal{C}^\infty(M))$. A notação padrão para semelhanças é $\text{Ad}T$:

$$(\text{Ad}T)V = (T \circ V \circ T^{-1}). \tag{1.4-19}$$

Com essa notação, temos

$$P_* = \text{Ad}T^{-1} \quad \forall P \in \text{Diff}(M).$$

Um **campo vetorial não autônomo** sobre uma variedade diferenciável M é uma aplicação que associa à cada par $(t, p) \in J \times M$, onde J é um intervalo da reta real, um vetor $V(t, p) \in T_pM$. Tais campos são também chamados de **campos dependentes do tempo**. Dado um campo vetorial não autônomo V , também temos o teorema de existência e unicidade de solução do Problema de Valor Inicial

$$\begin{cases} \frac{d}{dt}q(t) = V(t, q(t)) \\ q(t_0) = q_0 \end{cases}$$

Para uma demonstração deste fato, veja [6], Teorema 17.15.

Seja P^t o fluxo de uma E.D.O determinada pelo campo de vetores não autônomo $V \in \text{Vec}(M)$. Vimos que P^t é um difeomorfismo e, portanto, pode ser considerado como um automorfismo de $C^\infty(M)$ que, por sua vez, é a única solução do *Operador Problema de Cauchy* (problema de Cauchy visto como operador):

$$\begin{cases} \dot{P}^t = P^t \circ V \\ P^0 = Id. \end{cases} \quad (1.4-20)$$

Como P^t é invertível, seja $Q^t = (P^t)^{-1}$. Desta forma, Q^t é a solução do Operador Problema de Cauchy:

$$\begin{cases} \dot{Q}^t = -V \circ Q^t \\ Q^0 = Id. \end{cases} \quad (1.4-21)$$

Vamos enunciar, sem demonstração, uma importante propriedade sobre famílias contínuas de operadores. A demonstração pode ser encontrada em [1], apêndice A.

Proposição 1.46 *Se A^t e B^t famílias contínuas de operadores em $C^\infty(M)$ que são diferenciáveis em t_0 , então a família de operadores $A^t \circ B^t$ é contínua e diferenciável em t_0 . Mais ainda, satisfaz a regra de Leibniz:*

$$\left. \frac{d}{dt} \right|_{t_0} (A^t \circ B^t) = \left(\left. \frac{d}{dt} \right|_{t_0} A^t \right) \circ B^{t_0} + A^{t_0} \circ \left(\left. \frac{d}{dt} \right|_{t_0} B^t \right). \quad (1.4-22)$$

Proposição 1.47 *Sejam P^t o automorfismo correspondente ao fluxo gerado pelo campo de vetores não autônomo V , e W uma derivação. Então,*

$$\frac{d}{dt} (\text{Ad}P^t) W = (\text{Ad}P^t) [V, W].$$

Demonstração: Basta usar as equações (1.4-19), (1.4-22), (1.4-21) e (1.4-20), para obter a expressão:

$$\frac{d}{dt} (\text{Ad}P^t) W = (\text{Ad}P^t) [V, W] + P^t \circ \left(\frac{d}{dt} W \right) \circ (P^t)^{-1}.$$

Como W não depende de t , temos que $\frac{d}{dt} W \equiv 0$. Portanto,

$$\frac{d}{dt} (\text{Ad}P^t) W = (\text{Ad}P^t) [V, W].$$



De acordo com a proposição acima, obtemos, em particular, o colchete de Lie

$$\begin{aligned} \left. \frac{d}{dt} \right|_{t=0} (\text{Ad}P^t)W &= \left. \frac{d}{dt} \right|_{t=0} (P^t \circ W \circ (P^t)^{-1}) \\ &= V \circ W - W \circ V \\ &= [V, W]. \end{aligned} \quad (1.4-23)$$

Definimos

$$\text{ad}V = \text{ad} \left(\left. \frac{d}{dt} P^t \right|_{t=0} \right) \stackrel{\text{def}}{=} \left. \frac{d}{dt} \right|_{t=0} \text{Ad}P^t. \quad (1.4-24)$$

Por (1.4-23), obtemos

$$(\text{ad}V)W = [V, W]. \quad (1.4-25)$$

Esta igualdade acima será bastante utilizada no restante do texto. Devido a (1.4-18) temos a igualdade

$$\text{Ad}P^t[X, Y] = [\text{Ad}P^t X, \text{Ad}P^t Y], \quad (1.4-26)$$

diferenciando esta igualdade e usando (1.4-24), obtemos

$$\begin{aligned} \left. \frac{d}{dt} \right|_{t=0} \text{Ad}P^t[X, Y] &= \left. \frac{d}{dt} \right|_{t=0} [\text{Ad}P^t X, \text{Ad}P^t Y] \\ &= \left. \frac{d}{dt} \right|_{t=0} (\text{Ad}P^t X \circ \text{Ad}P^t Y - \text{Ad}P^t Y \circ \text{Ad}P^t X) \\ &= \left(\left. \frac{d}{dt} \right|_{t=0} \text{Ad}P^t X \circ \text{Ad}P^t Y \right) - \left(\left. \frac{d}{dt} \right|_{t=0} \text{Ad}P^t Y \circ \text{Ad}P^t X \right) \\ &= \left(\left. \frac{d}{dt} \right|_{t=0} \text{Ad}P^t X \right) \circ Y + X \circ \left(\left. \frac{d}{dt} \right|_{t=0} \text{Ad}P^t Y \right) \\ &\quad - \left(\left. \frac{d}{dt} \right|_{t=0} \text{Ad}P^t Y \right) \circ X - Y \circ \left(\left. \frac{d}{dt} \right|_{t=0} \text{Ad}P^t X \right) \\ &= (\text{ad}V)X \circ Y + X \circ (\text{ad}V)Y - (\text{ad}V)Y \circ X - Y \circ (\text{ad}V)X \\ &= (\text{ad}V)X \circ Y - Y \circ (\text{ad}V)X + X \circ (\text{ad}V)Y - (\text{ad}V)Y \circ X \\ &= [(\text{ad}V)X, Y] + [X, (\text{ad}V)Y]. \end{aligned} \quad (1.4-27)$$

Por outro lado, ainda por (1.4-24)

$$\left. \frac{d}{dt} \right|_{t=0} \text{Ad}P^t[X, Y] = \text{ad}V[X, Y] = [V, [X, Y]]. \quad (1.4-28)$$

Por (1.4-27) e (1.4-28), temos

$$[V, [X, Y]] = [(\text{ad}V) X, Y] + [X, (\text{ad}V) Y]$$

ou de outra forma,

$$[V, [X, Y]] = [[V, X], Y] + [X, [V, Y]].$$

De maneira simétrica podemos escrever ainda como

$$[V, [X, Y]] + [Y, [V, X]] + [X, [Y, V]] = 0,$$

esta forma é conhecida com *Identidade de Jacobi*.

O colchete de Lie de campos de vetores satisfaz as seguintes propriedades:

Para todo $X, Y, Z \in \text{Vec}(M)$, e $\alpha, \beta \in \mathbb{R}$ valem

(i) Bilinear:

$$[\alpha X + \beta Y, Z] = \alpha[X, Z] + \beta[Y, Z];$$

(ii) Antissimétrica:

$$[X, Y] = -[Y, X];$$

(iii) Identidade de Jacobi:

$$[V, [X, Y]] + [Y, [V, X]] + [X, [Y, V]] = 0.$$

Com essas propriedades, temos que o espaço vetorial $(\text{Vec}(M), [,])$ munido com o colchete de Lie é uma *álgebra de Lie de campos de vetores*.

CAPÍTULO 2

CONTROLABILIDADE DE SISTEMAS DE CONTROLE

Estudaremos neste capítulo a definição formal de controlabilidade. Também introduziremos a classe mais simples de sistemas de controle: sistemas lineares. Apresentaremos a demonstração de um importante resultado da Teoria de Controle, conhecido como *critério de Kalman*, que fornece condições para a controlabilidade completa de sistemas de controle lineares.

2.1 Sistemas de controle lineares

Faremos uma apresentação bastante objetiva de sistemas lineares em \mathbb{R}^n , com os principais resultados. Esta seção tem fundamental importância para o próximo capítulo, uma vez que, a investigação sobre a "linearizabilidade" do sistema de controle não linear se baseia na existência de uma correspondência com um sistema linear.

Um *sistema de controle linear* em \mathbb{R}^n tem a forma:

$$\dot{x} = Ax + c + \sum_{i=1}^m u_i b_i, \quad x \in \mathbb{R}^n \quad u = (u_1, \dots, u_m) \in \mathbb{R}^m, \quad (2.1-1)$$

onde $c, b_i \in \mathbb{R}^n$, $i \in \{1, \dots, m\}$, e $A \in M_{n \times n}(\mathbb{R}^n)$ são constantes.

O sistema linear com a condição inicial

$$\begin{cases} \dot{x} = Ax + c + \sum_{i=1}^m u_i b_i, & x \in \mathbb{R}^n \\ x(0, x_0) = x_0 \end{cases}$$

tem uma única solução, que corresponde aos controles admissíveis $u(\tau) = (u_1(\tau), \dots, u_m(\tau))$ localmente integráveis. Esta solução é conhecida como, *Fórmula de Cauchy para sistemas lineares*, e descrita como:

$$x(t, x_0) = e^{tA} \left(x_0 + \int_0^t e^{-\tau A} \sum_{i=1}^m (u_i(\tau) b_i + c) d\tau \right) \quad (2.1-2)$$

onde $u(\tau) = (u_1(\tau), \dots, u_m(\tau))$ é localmente integrável e

$$e^{tA} = Id + tA + \frac{t^2}{2!} A^2 + \dots + \frac{t^n}{n!} A^n + \dots$$

De fato, reescrevendo (2.1-2) como:

$$x(t, x_0) = e^{tA} x_0 + e^{tA} \int_0^t \left(e^{-\tau A} \sum_{i=1}^m u_i b_i \right) d\tau + e^{tA} \int_0^t (e^{-\tau A} c) d\tau,$$

diferenciando obtemos:

$$\begin{aligned} \dot{x}(t, x_0) &= A(e^{tA} x_0) + A e^{tA} \int_0^t e^{-\tau A} \sum_{i=1}^m u_i b_i d\tau + e^{tA} e^{-tA} \sum_{i=1}^m u_i b_i \\ &+ A e^{tA} \int_0^t (e^{-\tau A} c) d\tau + e^{tA} e^{-tA} c \\ &= A(e^{tA} x_0) + A e^{tA} \int_0^t e^{-\tau A} \sum_{i=1}^m u_i b_i d\tau + \sum_{i=1}^m u_i b_i + A e^{tA} \int_0^t (e^{-\tau A} c) d\tau + c \\ &= A \left(e^{tA} x_0 + e^{tA} \int_0^t e^{-\tau A} \sum_{i=1}^m u_i b_i d\tau + e^{tA} \int_0^t (e^{-\tau A} c) d\tau \right) + \sum_{i=1}^m u_i b_i + c \\ &= A \left(e^{tA} \left(x_0 + \int_0^t e^{-\tau A} \sum_{i=1}^m u_i b_i d\tau + \int_0^t e^{-\tau A} c d\tau \right) \right) + \sum_{i=1}^m u_i b_i + c \\ &= A \left(e^{tA} \left(x_0 + \int_0^t e^{-\tau A} \left(\sum_{i=1}^m u_i b_i + c \right) d\tau \right) \right) + \sum_{i=1}^m u_i b_i + c \\ &= Ax(t, x_0) + c + \sum_{i=1}^m u_i b_i. \end{aligned}$$

Portanto, a fórmula de Cauchy é a solução de (2.1-1).

2.2 Controlabilidade

Vimos na seção anterior a fórmula de Cauchy para sistemas lineares (2.1-1). Denotando por \mathcal{U} o conjunto dos controles localmente integráveis,

$$\mathcal{U} = \{u : \mathbb{R} \rightarrow \mathbb{R}^n \mid u \text{ é localmente integrável}\},$$

obtemos uma aplicação

$$F : \mathcal{U} \rightarrow \mathbb{R}^n$$

definida por

$$u \mapsto x(t, u, x_0) = e^{tA} \left(x_0 + \int_0^t (e^{-\tau A} c) d\tau \right) + e^{tA} \int_0^t \left(e^{-\tau A} \sum_{i=1}^m u_i b_i \right) d\tau.$$

Por simplicidade, escrevemos

$$\sum_{i=1}^m u_i b_i = Bu$$

onde $B \in M_{m \times n}(\mathbb{R})$ é formada pelos vetores coluna constantes $b_i \in \mathbb{R}^n$ e $u = (u_1, \dots, u_m)$.

Com isso, obtemos

$$u \mapsto x(t, u, x_0) = e^{tA} \left(x_0 + \int_0^t (e^{-\tau A} c) d\tau \right) + e^{tA} \int_0^t (e^{-\tau A} Bu) d\tau.$$

Observe que F é uma aplicação afim. Segue que, o conjunto de pontos atingíveis partindo do estado x_0 para um tempo $t > 0$ fixado, é um *subespaço afim* de \mathbb{R}^n .

Daremos agora a definição geral de controlabilidade.

Definição 2.1 Dizemos que um sistema de controle Σ é **completamente controlável** para um tempo $t > 0$ sobre um espaço estado M se,

$$\mathcal{A}_{q_0}(t) = M, \quad \forall q_0 \in M.$$

Chamaremos tais sistemas simplesmente de **controláveis para um tempo** $t > 0$.

Esta definição nos diz que, se um sistema de controle Σ é controlável para um

tempo $t > 0$, então para quaisquer pontos $q_0, q_1 \in M$, existe um controle u (admissível), cuja solução $q(t, u, q_0)$ do problema de Cauchy

$$\dot{q} = V_u(q), \quad q(0, u, q_0) = q_0$$

satisfaz,

$$q(0, u, q_0) = q_0, \quad q(t, u, q_0) = q_1,$$

ou seja, é possível atingir q_1 à partir de q_0 em t unidades de tempo.

Observação 2.2 Para sistemas de controle lineares controláveis para um tempo $t > 0$ sobre \mathbb{R}^n , a aplicação afim

$$F : \mathcal{U} \rightarrow \mathbb{R}^n$$

torna-se sobrejetora. Ou seja,

$$u \mapsto \int_0^t (e^{-\tau A} B u) d\tau$$

é sobrejetora.

Isto deixa a investigação sobre controlabilidade de Σ facilitada pela análise da sobrejetividade de F . Em geral, dados espaços vetoriais V, W e uma transformação afim $F_1 : V \rightarrow W$ da forma $F_1 x = b_0 + T x$, com $b_0 \in W$ e $T : V \rightarrow W$ linear, temos que F_1 é sobrejetora se, e somente se, T é sobrejetora. Além disso, como e^{tA} é um isomorfismo linear de \mathbb{R}^n , a sobrejetividade da aplicação F é equivalente à sobrejetividade da aplicação

$$u \mapsto \int_0^t (e^{-\tau A} B u) d\tau.$$

Apresentaremos agora o principal teorema deste capítulo

Teorema 2.3 O sistema de controle linear em \mathbb{R}^n

$$\dot{x} = Ax + c + \sum_{i=1}^m u_i b_i$$

é completamente controlável para um tempo $t > 0$ se, e somente se,

$$\text{span}\{A^j b_i \mid j = 0, \dots, n-1, i = 1, \dots, m\} = \mathbb{R}^n. \quad (2.2-3)$$

A igualdade (2.2-3) é conhecida como **critério de Kalman** para controlabilidade de sistemas

lineares.

Demonstração: Suponha que

$$\text{span}\{A^j b_i \mid j = 0, \dots, n-1, i = 1, \dots, m\} \neq \mathbb{R}^n.$$

Isto significa que o conjunto dos elementos da forma $A^j b_i$, para $j = 0, \dots, n-1$ e $i = 1, \dots, m$, não geram o espaço \mathbb{R}^n . Logo, existe um funcional linear $\varphi \neq 0$

$$\varphi : \mathbb{R}^n \rightarrow \mathbb{R},$$

tal que

$$\varphi A^j b_i = 0, \quad j = 0, \dots, n-1, \quad i = 1, \dots, m. \quad (2.2-4)$$

Seja

$$p(\lambda) = \det(\lambda Id_n - A) = \lambda^n + \alpha_{n-1} \lambda^{n-1} + \alpha_{n-2} \lambda^{n-2} + \dots + \alpha_0 Id_n$$

o polinômio característico da matriz $A_{n \times n}$, onde $\alpha_0, \dots, \alpha_{n-1} \in \mathbb{R}$.

Pelo Teorema de Cayley-Hamilton, temos que:

$$\varphi(A) = A^n + \alpha_{n-1} A^{n-1} + \alpha_{n-2} A^{n-2} + \dots + \alpha_0 Id_n = 0,$$

logo,

$$A^n = \beta_{n-1} A^{n-1} + \beta_{n-2} A^{n-2} + \dots + \beta_0 Id_n = \sum_{k=0}^{n-1} \beta_k A^k, \quad (2.2-5)$$

com $\beta_0, \dots, \beta_{n-1} \in \mathbb{R}$.

Consequentemente, temos

$$A^r = \sum_{k=0}^{n-1} \beta_k^r A^k, \quad (2.2-6)$$

para todo $r \in \mathbb{N}$, e $\beta_k^r \in \mathbb{R}$.

De fato, para cada $r \in \{0, 1, \dots, n-1\}$ podemos escrever

$$A^r = 0 \cdot A^{n-1} + 0 \cdot A^{n-2} + \dots + 1 \cdot A^r + \dots + 0 \cdot A + 0 \cdot Id_n.$$

Logo, a igualdade (2.2-6) é válida para todo $r \in \{0, 1, \dots, n-1\}$. Se $r = n$ então a igualdade (2.2-6) é a própria (2.2-5). Para $r > n$ aplicamos (2.2-5) repetidamente. Por

exemplo, se $r = n + 1$ temos,

$$\begin{aligned}
A^{n+1} &= A \cdot A^n = A \cdot \sum_{k=0}^{n-1} \beta_k A^k = \sum_{k=0}^{n-1} \beta_k A^{k+1} = \beta_{n-1} \cdot A^n + \sum_{k=0}^{n-2} \beta_k A^k \\
&= \beta_{n-1} \cdot \left(\sum_{k=0}^{n-1} \beta_k A^k \right) + \sum_{k=0}^{n-2} \beta_k A^k \\
&= \sum_{k=0}^{n-1} \beta_{n-1} \cdot \beta_k A^k + \sum_{k=0}^{n-2} \beta_k A^k \\
&= \beta_{n-1} \cdot \beta_{n-1} A^{n-1} + \sum_{k=0}^{n-2} (\beta_k + \beta_{n-1} \cdot \beta_k) A^k \\
&= \sum_{k=0}^{n-1} \gamma_k^{n+1} A^k,
\end{aligned}$$

onde $\gamma_k^{n+1} = \beta_k + \beta_{n-1} \cdot \beta_k$, para $k \in \{0, 1, \dots, n-2\}$, e $\gamma_{n-1}^{n+1} = \beta_{n-1} \cdot \beta_{n-1}$. Portanto, a igualdade é válida para todo $r \in \mathbb{N}$.

Combinando as equações (2.2-4) e (2.2-6) temos

$$\varphi(A^r b_i) = \varphi\left(\sum_{k=0}^{n-1} \beta_k^r A^k b_i\right) = \sum_{k=0}^{n-1} \beta_k^r \varphi A^k b_i = 0, \quad \forall i \in \{1, \dots, m\}, r \in \mathbb{N},$$

o que implica

$$\begin{aligned}
\varphi e^{-\tau A} b_i &= \varphi\left(\text{Id} + (-\tau)A + \frac{(-\tau)^2}{2!}A^2 + \dots + \frac{(-\tau)^n}{n!}A^n + \dots\right) b_i \\
&= \varphi \text{Id} \cdot b_i + (-\tau)\varphi A b_i + \frac{(-\tau)^2}{2!}\varphi A^2 b_i + \dots + \frac{(-\tau)^n}{n!}\varphi A^n b_i + \dots \\
&= 0.
\end{aligned}$$

Assim,

$$\varphi \int_0^t e^{-\tau A} \sum_{i=1}^m u_i b_i d\tau = \int_0^t \sum_{i=1}^m u_i \varphi e^{-\tau A} b_i d\tau = 0.$$

Desta forma, definindo

$$T(u) = \int_0^t e^{-\tau A} \sum_{i=1}^m u_i(\tau) b_i d\tau,$$

temos que $\varphi(T(u)) = 0 \forall u \in \mathcal{U}$. Como existe $v_0 \in \mathbb{R}^n$ tal que $\varphi v_0 \neq 0$, segue que $T(u) \neq v_0$ para qualquer $u \in \mathcal{U}$. Portanto, T não é sobrejetora. Pela Observação (2.2), concluímos que $\mathcal{A}_{q_0}(t) \neq \mathbb{R}^n$.

Reciprocamente, supondo que $\mathcal{A}_{x_0}(t) \neq \mathbb{R}^n$ para algum $x_0 \in \mathbb{R}^n$, temos que a aplicação

$$u \in \mathcal{U} \xrightarrow{T} \int_0^t e^{-\tau A} \sum_{i=1}^m u_i b_i d\tau \in \mathbb{R}^n$$

não é sobrejetora. Logo existe um funcional linear $\varphi \in (\mathbb{R}^n)^*$, $\varphi \neq 0$, tal que

$$\varphi(T(u)) = 0, \quad \forall u \in \mathcal{U}. \quad (2.2-7)$$

Considere o controle $u(\tau) = (0, \dots, 0, u_i^s(\tau), 0, \dots, 0)$, onde a única componente não nula é dada por:

$$u_i^s(\tau) = \begin{cases} 1 & \text{se } 0 \leq \tau \leq s; \\ 0 & \text{se } \tau > s. \end{cases}$$

Deste modo a equação (2.2-7) fica sendo

$$\varphi \int_0^s e^{-\tau A} b_i d\tau = 0. \quad (2.2-8)$$

Como φ é linear, a primeira derivada fica

$$0 = \frac{d}{ds} \left(\varphi \int_0^s e^{-\tau A} b_i d\tau \right) = \varphi \left(\frac{d}{ds} \int_0^s e^{-\tau A} b_i d\tau \right) = \varphi e^{-sA} b_i.$$

Logo,

$$\varphi e^{-sA} b_i = 0, \quad \forall s \in \mathbb{R}, \quad i \in \{1, \dots, m\}.$$

Derivando sucessivamente a igualdade acima em $s = 0$, obtemos

$$\varphi(-A^k b_i) = 0, \quad i \in \{1, \dots, m\}, \quad k \in \mathbb{N}.$$

Portanto encontramos um funcional linear não-nulo φ tal que

$$\varphi A^j b_i = 0, \quad j \in \{0, 1, \dots, n-1\}, \quad i \in 1, \dots, m.$$

Isto implica que

$$\text{span}\{A^j b_i \mid j = 0, \dots, n-1, i = 1, \dots, m\} \neq \mathbb{R}^n.$$



Exemplo 2.4 O sistema linear em \mathbb{R}^2 :

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \underbrace{\begin{pmatrix} -2 & 2 \\ 1 & -1 \end{pmatrix}}_{=A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \underbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{=b} u,$$

com $(x_1, x_2) \in \mathbb{R}^2$, e $u \in \mathbb{R}$, é completamente controlável. De fato,

$$\text{span}\{b, Ab\} = \text{span}\{(1, 0)^T, (-2, 1)^T\} = \mathbb{R}^2.$$

Exemplo 2.5 O sistema linear em \mathbb{R}^2 :

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \underbrace{\begin{pmatrix} a_1 & a_2 \\ 0 & 1 \end{pmatrix}}_{=A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \underbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{=b} u,$$

com $(x_1, x_2) \in \mathbb{R}^2$, $a_1, a_2, u \in \mathbb{R}$, não é completamente controlável. De fato,

$$\text{span}\{b, Ab\} = \text{span}\{(1, 0)^T, (a_1, 0)^T\} = \text{span}\{(1, 0)^T\} \neq \mathbb{R}^2.$$

CAPÍTULO 3

LINEARIZAÇÃO DE SISTEMAS DE CONTROLE

Nosso objetivo neste capítulo é caracterizar os sistemas de controle não lineares que são equivalentes à sistemas de controle lineares controláveis num espaço euclidiano, localmente e globalmente. Vamos introduzir algumas ferramentas importantes para esta equivalência. Para tanto, sejam M uma variedade de dimensão n e f_0, f_1, \dots, f_m campos de vetores em M . Estudaremos sistemas de controle não lineares da forma

$$\dot{q} = f_0(q) + \sum_{i=1}^m u_i f_i(q) \quad q \in M, \quad (3.0-1)$$

onde $u = (u_1, \dots, u_m) \in \mathbb{R}^m$.

3.1 Linearização local

Para tratarmos da linearização, o conceito fundamental é o colchete de Lie, que nos fornece informações sobre como os campos de vetores interagem entre si e sobre a comutatividade ou não dos fluxos. Além disso, o colchete de Lie também é invariante por difeomorfismos:

$$\Phi_* [V, W] = [\Phi_* V, \Phi_* W], \quad \forall V, W \in \text{Vec}(M). \quad (3.1-2)$$

No caso em que os campos de vetores são da forma

$$f_0(x) = Ax + c, \text{ e } f_i(x) = b_i, \quad i = 1, \dots, m$$

temos, usando (1.4-25) e a Proposição 1.32, que

$$(\text{ad } f_0) f_i(x) = [f_0, f_i](x) = [Ax + c, b_i](x) = -Ab_i,$$

$$(\text{ad } f_0)^2 f_i = (\text{ad } f_0) (\text{ad } f_0) f_i = (\text{ad } f_0) [f_0, f_i] = [f_0, [f_0, f_i]] = (-A)^2 b_i$$

e, de modo geral, temos

$$(\text{ad } f_0)^j f_i = \underbrace{[f_0, [\dots [f_0, f_i] \dots]]}_{j\text{-vezes}} = (-A)^j b_i, \quad j = 0, 1, 2, \dots \quad (3.1-3)$$

Neste caso, novamente por 1.32 temos

$$\left[(\text{ad } f_0)^{j_1} f_{i_1}, (\text{ad } f_0)^{j_2} f_{i_2} \right] = \left[(-A)^{j_1} b_{i_1}, (-A)^{j_2} b_{i_2} \right] = 0, \quad (3.1-4)$$

para todo $j_1, j_2 \geq 0$ e $1 \leq i_1, i_2 \leq m$. Nas condições do Teorema 2.3 podemos escrever

$$\text{span}\{(\text{ad } f_0)^j f_i(x_0) \mid 0 \leq j \leq n-1, 1 \leq i \leq m\} = T_{x_0} \mathbb{R}^n. \quad (3.1-5)$$

O próximo resultado fornece condições que "retificam" os campos de vetores não lineares localmente. Para a demonstração, usaremos a notação da Definição 1.19.

Lema 3.1 *Sejam M uma variedade de dimensão n e $V_i \in \text{Vec}(M)$ $i \in \{1, \dots, k\}$, $k \leq n$. Existe um difeomorfismo local*

$$\Phi: O_0 \subset \mathbb{R}^n \rightarrow O_{q_0} \subset M$$

tal que

$$\Phi_* \left(\frac{\partial}{\partial x_i} \right) = V_i, \quad i = 1, \dots, k,$$

onde $\frac{\partial}{\partial x_i} \in \text{Vec}(\mathbb{R}^n)$ são campos de vetores constantes na direção das coordenadas se, e somente se,

$$[V_i, V_j] = 0 \quad i, j = 1, \dots, k,$$

e

$$\dim \text{span}\{V_1(q_0), \dots, V_k(q_0)\} = k,$$

isto é, os campos V_i , $i = 1, \dots, k$ formam um conjunto linearmente independente.

Demonstração: (\Rightarrow) Basta notar que o colchete de Lie de campos de vetores constantes é nulo e, juntamente com (3.1-2) temos o desejado. De fato,

$$[V_i, V_j] = \left[\Phi_* \left(\frac{\partial}{\partial x_i} \right), \Phi_* \left(\frac{\partial}{\partial x_j} \right) \right] = \Phi_* \left[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \right] = 0$$

pois, Φ_* é um isomorfismo uma vez que Φ é um difeomorfismo local. Por este fato, e pela igualdade

$$\Phi_* \left(\frac{\partial}{\partial x_i} \right) = V_i,$$

temos que V_i $i \in \{1, \dots, k\}$ são também linearmente independentes.

(\Leftarrow) Reciprocamente, suponha que os campos de vetores formam um conjunto V_1, \dots, V_k linearmente independente no ponto $q_0 \in M$. Vamos completar este conjunto para formar uma base para $T_{q_0}M$. Escolhemos $V_{k+1}, \dots, V_n \in \text{Vec}(M)$ tais que

$$\text{span}\{V_1(q_0), \dots, V_n(q_0)\} = T_{q_0}M.$$

Como os campos V_1, \dots, V_n são linearmente independentes no ponto q_0 , são também linearmente independentes em uma vizinhança de q_0 . Denote por U_{q_0} tal vizinhança. Assim,

$$\text{span}\{V_1(q), \dots, V_n(q)\} = T_qM, \quad \forall q \in U_{q_0}.$$

Considere a aplicação

$$\begin{aligned} \Phi : \quad U_0 \subset \mathbb{R}^n &\rightarrow U_{q_0} \subset M \\ (x_1, \dots, x_n) &\mapsto e^{x_n V_n} \circ \dots \circ e^{x_1 V_1}(q_0) \end{aligned}$$

onde U_0 é uma vizinhança da origem em \mathbb{R}^n . Vamos mostrar que Φ é o difeomorfismo que queremos. De fato, Φ é diferenciável pois os fluxos $e^{x_1 V_1}, \dots, e^{x_n V_n}$ são diferenciáveis.

veis. Além disso,

$$\begin{aligned}
\Phi_* \left(\frac{\partial}{\partial x_i} \Big|_{x=0} \right) &\stackrel{\text{def}}{=} \frac{\partial}{\partial x_i} \Big|_{x=0} \Phi(x) \\
&= \frac{\partial}{\partial x_i} \Big|_{x=0} e^{x_n V_n} \circ \dots \circ e^{x_1 V_1}(q_0) \\
&= e^{x_n V_n} \circ \dots \circ e^{x_i V_i} \circ V_i \circ \dots \circ e^{x_1 V_1}(q_0) \Big|_{x=0} \\
&= V_i(q_0) \\
&= \frac{\partial}{\partial \varepsilon} \Big|_{\varepsilon=0} e^{\varepsilon V_i}(q_0),
\end{aligned}$$

ou seja,

$$\Phi_* \left(\frac{\partial}{\partial x_i} \Big|_{x=0} \right) \stackrel{\text{def}}{=} \frac{\partial}{\partial \varepsilon} \Big|_{\varepsilon=0} e^{\varepsilon V_i}(q_0) = V_i(q_0). \quad (3.1-6)$$

Isto significa que

$$\Phi_{*0} : T_0 \mathbb{R}^n \rightarrow T_{q_0} M$$

leva a base canônica de \mathbb{R}^n na base de $T_{q_0} M$ em q_0 . Logo Φ_{*0} é um isomorfismo. Pelo Teorema da Função Inversa, existem abertos $O_0 \subset \mathbb{R}^n$ e $O_{q_0} \subset M$ tal que

$$\Phi : O_0 \rightarrow O_{q_0}$$

é um difeomorfismo local. Pela Proposição 1.31, a hipótese

$$[V_i, V_j] = 0, \quad \forall i, j = 1, \dots, k$$

implica que seus correspondentes fluxos comutam, isto é,

$$e^{x_i V_i} \circ e^{x_j V_j} = e^{x_j V_j} \circ e^{x_i V_i} = e^{x_i V_i + x_j V_j}, \quad i, j = 1, \dots, k.$$

Com isso podemos reescrever Φ na forma

$$\Phi(x_1, \dots, x_k, x_{k+1}, \dots, x_n) = e^{x_n V_n} \circ \dots \circ e^{x_{k+1} V_{k+1}} \circ e^{\sum_{i=1}^k x_i V_i}(q_0).$$

Assim, para $i = 1, \dots, k$ temos

$$\begin{aligned}
\Phi_* \left(\frac{\partial}{\partial x_i} \right) \Big|_{\Phi(x)} &= \frac{\partial}{\partial \varepsilon} \Phi(x_1, \dots, x_i + \varepsilon, \dots, x_n) \Big|_{\varepsilon=0} \\
&= \frac{\partial}{\partial \varepsilon} e^{x_n V_n} \circ \dots \circ e^{x_{k+1} V_{k+1}} \circ e^{\sum_{j=1}^k x_j V_j} \circ e^{\varepsilon V_i}(q_0) \Big|_{\varepsilon=0} \\
&= e^{x_n V_n} \circ \dots \circ e^{x_{k+1} V_{k+1}} \circ e^{\sum_{j=1}^k x_j V_j} \circ e^{\varepsilon V_i}(q_0) \circ V_i \Big|_{\varepsilon=0} \\
&= V_i(\Phi(x)).
\end{aligned}$$

Isto mostra que o difeomorfismo Φ retifica localmente os campos de vetores $V_i \in M$. ■

Teorema 3.2 (Linearização local) *Sejam M uma variedade de dimensão n , $q_0 \in M$, e $f_0, \dots, f_m \in \text{Vec}(M)$. Existe um difeomorfismo local*

$$\Phi: O_{q_0} \subset M \rightarrow O_0 \subset \mathbb{R}^n$$

tal que

$$(\Phi_* f_0)(x) = Ax + c, \quad x \in O_0 \tag{3.1-7}$$

$$(\Phi_* f_i)(x) = b_i, \quad x \in O_0, \quad i = 1, \dots, m \tag{3.1-8}$$

onde $c \in \mathbb{R}^n$, $A \in M_n(\mathbb{R})$ é uma matriz constante real e $b_i \in \mathbb{R}^n$ satisfazendo a condição de controlabilidade (2.2-3) se, e somente se,

$$\text{span}\{(\text{ad } f_0)^j f_i(q_0)\} = T_{q_0} M \tag{3.1-9}$$

$$\left[(\text{ad } f_0)^{j_1} f_{i_1}, (\text{ad } f_0)^{j_2} f_{i_2} \right](q) = 0, \quad \forall q \in U_{q_0} \tag{3.1-10}$$

com $0 \leq j, j_1, j_2 \leq n - 1$, e $1 \leq i, i_1, i_2 \leq m$.

Demonstração: (\Rightarrow) Por hipótese, a condição de controlabilidade (2.2-3) é satisfeita.

Então temos

$$\text{span}\{A^j b_i \mid j = 0, \dots, n - 1, i = 1, \dots, m\} = \mathbb{R}^n$$

e, por (3.1-3) e (3.1-5), podemos escrever

$$\text{span}\{(\text{ad}\Phi_*f_0)^j \Phi_*f_i(x)\} = T_x\mathbb{R}^n \quad (3.1-11)$$

para $j = 1, \dots, n-1$, $i = 1, \dots, m$. Assim, se $x \in O_0$, por (3.1-4) e (3.1-3) temos

$$\begin{aligned} 0 &= \left[(\text{ad}\Phi_*f_0)^{j_1} \Phi_*f_{i_1}, (\text{ad}\Phi_*f_0)^{j_2} \Phi_*f_{i_2} \right] (x) \\ &= \Phi_* \left[\left[f_0, [\dots [f_0, f_{i_1}] \dots] \right], \left[f_0, [\dots [f_0, f_{i_2}] \dots] \right] \right] (x) \\ &= \Phi_* \left[(\text{ad}f_0)^{j_1} f_{i_1}, (\text{ad}f_0)^{j_2} f_{i_2} \right] (x) \\ &= \Phi_{*\Phi^{-1}(x)} \left[(\text{ad}f_0)^{j_1} f_{i_1}, (\text{ad}f_0)^{j_2} f_{i_2} \right] (\Phi^{-1}(x)) \\ &= \Phi_{*q} \left[(\text{ad}f_0)^{j_1} f_{i_1}, (\text{ad}f_0)^{j_2} f_{i_2} \right] (q), \text{ onde } \Phi^{-1}(x) = q \in O_{q_0}. \end{aligned}$$

Como $\Phi_{*q} : T_qM \rightarrow T_x\mathbb{R}^n$ é isomorfismo para cada $q \in O_{q_0}$, temos que

$$\left[(\text{ad}f_0)^{j_1} f_{i_1}, (\text{ad}f_0)^{j_2} f_{i_2} \right] (q) = 0,$$

ou seja

$$\left[(\text{ad}f_0)^{j_1} f_{i_1}, (\text{ad}f_0)^{j_2} f_{i_2} \right] (q) = 0, \quad \forall q \in O_{q_0}.$$

Isto prova (3.1-10).

De (3.1-11) temos

$$\begin{aligned} T_0\mathbb{R}^n &= \text{span}\{(\text{ad}\Phi_*f_0)^j \Phi_*f_i(0)\} \\ &= \text{span}\{\Phi_* [f_0, [\dots [f_0, f_i] \dots]](0)\} \\ &= \text{span}\{\Phi_{*q_0} (\text{ad}f_0)^j f_i(q_0)\}. \end{aligned}$$

Visto isso, e novamente que $\Phi_{*q_0}^{-1} : T_0\mathbb{R}^n \rightarrow T_{q_0}M$ é isomorfismo, obtemos

$$\Phi_{*q_0}^{-1} \left(\Phi_{*q_0} (\text{ad}f_0)^j f_i(q_0) \right) = (\text{ad}f_0)^j f_i(q_0).$$

Consequentemente,

$$\text{span}\{(\text{ad}f_0)^j f_i(q_0)\} = T_{q_0}M,$$

par todo $j = 1, \dots, n-1$, e $i = 1, \dots, m$. Isto prova (3.1-9).

(\Leftarrow) Reciprocamente, da hipótese (3.1-9), podemos extrair uma base para $T_{q_0}M$ formada por vetores da forma $(\text{ad } f_0)^j f_i(q_0)$. Por simplicidade, denotaremos

$$(\text{ad } f_0)^{j_\alpha} f_{i_\alpha} = Y_\alpha \quad \alpha = 1, \dots, n.$$

Assim, obtemos

$$T_{q_0}M = \text{span}\{Y_\alpha(q_0) \mid \alpha = 1, \dots, n\},$$

segue que

$$\dim \text{span}\{Y_\alpha(q_0) \mid \alpha = 1, \dots, n\} = n.$$

Devido a hipótese (3.1-10) temos

$$[Y_{\alpha_1}, Y_{\alpha_2}](q) = 0, \quad \forall q \in U_{q_0} \quad 1 \leq \alpha_1, \alpha_2 \leq n.$$

Pelo Lema 3.1 existe um difeomorfismo local:

$$\Phi: O_{q_0} \subset M \rightarrow O_0 \subset \mathbb{R}^n$$

tal que

$$\Phi_*(Y_\alpha) = \frac{\partial}{\partial x_\alpha}, \quad \alpha = 1, \dots, n.$$

Vamos mostrar que Φ é o difeomorfismo desejado. De fato, como Φ é um difeomorfismo local e f_i para $i = 1, \dots, m$, são campos de vetores em M , temos que $\Phi_* f_i$ são campos de vetores induzidos em $O_0 \subset \mathbb{R}^n$. Assim, podemos escrevê-los em coordenadas locais. Para cada $i = 1, \dots, m$ temos

$$\Phi_* f_i(x) = \sum_{\alpha=1}^n b_\alpha^i(x) \frac{\partial}{\partial x_\alpha},$$

com $b_\alpha^i: O_0 \subset \mathbb{R}^n \rightarrow \mathbb{R}$ funções reais. Escrevemos também os campos de vetores constantes em coordenadas locais

$$\Phi_*(Y_\beta) = \frac{\partial}{\partial x_\beta} = \sum_{\alpha=1}^n a_\alpha^\beta(x) \frac{\partial}{\partial x_\alpha}, \quad \beta = 1, \dots, n,$$

com $a_\alpha^\beta : O_0 \subset \mathbb{R}^n \rightarrow \mathbb{R}$ funções reais dadas por

$$a_\alpha^\beta(x) = \begin{cases} 1 & \text{se } \alpha = \beta, \\ 0 & \text{se } \alpha \neq \beta. \end{cases}$$

Com isso obtemos, pela Proposição 1.32, que

$$\begin{aligned} \left[\frac{\partial}{\partial x_\beta}, \Phi_* f_i \right] &= \sum_{\alpha, \kappa=1}^n \left(a_\kappa^\beta \frac{\partial b_\alpha^i}{\partial x_\kappa} - b_\kappa^i \frac{\partial a_\alpha^\beta}{\partial x_\kappa} \right) \frac{\partial}{\partial x_\alpha} \\ &= \sum_{\alpha, \kappa=1}^n a_\kappa^\beta \frac{\partial b_\alpha^i}{\partial x_\kappa} \frac{\partial}{\partial x_\alpha} \\ &= \sum_{\alpha=1}^n \frac{\partial b_\alpha^i}{\partial x_\beta} \frac{\partial}{\partial x_\alpha}. \end{aligned} \tag{3.1-12}$$

Por outro lado

$$\begin{aligned} \left[\frac{\partial}{\partial x_\beta}, \Phi_* f_i \right] &= [\Phi_* Y_\beta, \Phi_* f_i] = \Phi_* [Y_\beta, f_i] \\ &= \Phi_* \left[(\text{ad } f_0)^{j_\beta} f_{i_\beta}, (\text{ad } f_0)^0 f_i \right] \\ &= \Phi_* 0 = 0. \end{aligned} \tag{3.1-13}$$

Agora, por (3.1-12) e (3.1-13), temos

$$\sum_{\alpha=1}^n \frac{\partial b_\alpha^i(x)}{\partial x_\beta} \frac{\partial}{\partial x_\alpha} = 0.$$

Logo

$$\frac{\partial b_\alpha^i(x)}{\partial x_\beta} = 0 \implies b_\alpha^i(x) = \text{constante}$$

para cada $i = 1, \dots, m$ e todo $\alpha = 1, \dots, n$. Isto implica que,

$$\Phi_* f_i(x) = b_i \quad \forall x \in O_0 \quad i = 1, \dots, m$$

com constantes $b_i \in \mathbb{R}^n$. Aqui fica provado (3.1-8). Analogamente, escrevemos

$$\Phi_* f_0(x) = \sum_{\alpha=1}^n b_\alpha(x) \frac{\partial}{\partial x_\alpha},$$

em coordenadas locais, onde $b_\alpha : O_0 \subset \mathbb{R}^n \rightarrow \mathbb{R}$ são funções reais para cada $\alpha = 1, \dots, n$.

E para cada $\gamma = 1, \dots, n$, os campos constantes em coordenadas locais ficam

$$\frac{\partial}{\partial x_\gamma}(x) = \sum_{\alpha=1}^n a_\alpha^\gamma(x) \frac{\partial}{\partial x_\alpha},$$

onde $a_\alpha^\gamma: O_0 \subset \mathbb{R}^n \rightarrow \mathbb{R}$ são funções reais dadas por:

$$a_\alpha^\gamma(x) = \begin{cases} 1 & \text{se } \alpha = \gamma, \\ 0 & \text{se } \alpha \neq \gamma. \end{cases}$$

Logo, pela Proposição 1.32, temos

$$\begin{aligned} \left[\frac{\partial}{\partial x_\gamma}, \left[\frac{\partial}{\partial x_\beta}, \Phi_* f_0 \right] \right] &= \left[\sum_{\alpha=1}^n a_\alpha^\gamma \frac{\partial}{\partial x_\alpha}, \sum_{\alpha=1}^n \frac{\partial b_\alpha}{\partial x_\beta} \frac{\partial}{\partial x_\alpha} \right] \\ &= \sum_{\alpha, \kappa=1}^n \left(a_\kappa^\gamma \frac{\partial^2 b_\alpha}{\partial x_\beta \partial x_\kappa} - \frac{\partial b_\kappa}{\partial x_\beta} \frac{\partial a_\alpha^\gamma}{\partial x_\kappa} \right) \frac{\partial}{\partial x_\alpha} \\ &= \sum_{\alpha, \kappa=1}^n a_\kappa^\gamma \frac{\partial^2 b_\alpha}{\partial x_\beta \partial x_\kappa} \frac{\partial}{\partial x_\alpha} \\ &= \sum_{\alpha=1}^n \frac{\partial^2 b_\alpha}{\partial x_\beta \partial x_\gamma} \frac{\partial}{\partial x_\alpha}. \end{aligned}$$

Por outro lado,

$$\begin{aligned} \left[\frac{\partial}{\partial x_\gamma}, \left[\frac{\partial}{\partial x_\beta}, \Phi_* f_0 \right] \right] &= [\Phi_*(Y_\gamma), [\Phi_*(Y_\beta), \Phi_* f_0]] \\ &= [\Phi_*(Y_\gamma), \Phi_* [Y_\beta, f_0]] \\ &= \Phi_* [Y_\gamma, [Y_\beta, f_0]] \\ &= \Phi_* \left[(\text{ad } f_0)^{j_\gamma} f_{i_\gamma}, [(\text{ad } f_0)^{j_\beta} f_{i_\beta}, f_0] \right] \\ &= \Phi_* \left[(\text{ad } f_0)^{j_\gamma} f_{i_\gamma}, -[f_0, (\text{ad } f_0)^{j_\beta} f_{i_\beta}] \right] \\ &= -\Phi_* \left[(\text{ad } f_0)^{j_\gamma} f_{i_\gamma}, (\text{ad } f_0) (\text{ad } f_0)^{j_\beta} f_{i_\beta} \right] \\ &= -\Phi_* \left[(\text{ad } f_0)^{j_\gamma} f_{i_\gamma}, (\text{ad } f_0)^{j_\beta+1} f_{i_\beta} \right] = 0, \end{aligned}$$

devido a hipótese (3.1-10). Consequentemente,

$$\sum_{\alpha=1}^n \frac{\partial^2 b_\alpha}{\partial x_\beta \partial x_\gamma} \frac{\partial}{\partial x_\alpha} = 0.$$

Isto significa que,

$$\frac{\partial^2 b_\alpha}{\partial x_\beta \partial x_\gamma} = 0, \quad \forall \alpha = 1, \dots, n$$

ou seja, para cada $\alpha = 1, \dots, n$ existem $A_\alpha = (A_{\alpha 1}, \dots, A_{\alpha n}) \in \mathbb{R}^n$ e $c_\alpha \in \mathbb{R}$ tais que $b_\alpha(x) = A_\alpha x + c_\alpha$. Logo, obtemos uma matriz constante real $A_{n \times n}$ formada por $A_{\alpha'}$ s e uma constante $c \in \mathbb{R}^n$ formada por $c_{\alpha'}$ s tais que

$$\Phi_* f_0(x) = Ax + c,$$

provando (3.1-7). Agora, por (3.1-7) e (3.1-8) obtemos o sistema de controle linear em $O_0 \subset \mathbb{R}^n$

$$\begin{aligned} \dot{x} &= (\Phi_* f_0)(x) + \sum_{i=1}^m u_i (\Phi_* f_i)(x) \\ &= Ax + c + \sum_{i=1}^m u_i b_i. \end{aligned}$$

Resta mostrarmos que o sistema acima satisfaz a condição de controlabilidade (2.2-3). Com efeito, temos que

$$T_{q_0} M = \text{span}\{(\text{ad} f_0)^{j_\alpha} f_{i_\alpha}(q_0)\}, \quad \forall \alpha = 1, \dots, n.$$

Como $\Phi_* : T_{q_0} M \rightarrow \mathbb{R}^n$ é isomorfismo, então

$$\Phi_* \left((\text{ad} f_0)^{j_\alpha} f_{i_\alpha}(q_0) \right) = (\text{ad} \Phi_* f_0)^{j_\alpha} \Phi_* f_{i_\alpha}(0)$$

é base para $T_0 \mathbb{R}^n = \mathbb{R}^n$, e por (3.1-3) temos

$$\text{span}\{(-A)^{j_\alpha} b_{i_\alpha} \mid \alpha = 1, \dots, n\} = \mathbb{R}^n.$$

■

3.2 Linearização global

Para provarmos a versão global do último resultado, precisaremos de alguns resultados auxiliares.

Proposição 3.3 *Seja $H \subset \mathbb{R}^n$ um subgrupo. Então H é discreto se, e somente se, a origem $0 \in \mathbb{R}^n$ é isolado.*

Demonstração: (\Leftarrow) Provaremos por contra-positiva. Suponha que H não é discreto. Então existe $h_0 \in H$ tal que

$$B(h_0, \epsilon) \cap (H \setminus \{h_0\}) \neq \emptyset \quad \forall \epsilon > 0.$$

Com isso, obtemos uma sequência $(h_n) \subset (H \setminus \{h_0\})$ que converge a h_0 , ou seja

$$h_n \rightarrow h_0 \quad \Rightarrow \quad (h_n - h_0) \rightarrow 0.$$

Repare que $(h_n - h_0) \in H$ pois, H é subgrupo. Portanto,

$$B(0, \epsilon) \cap (H \setminus \{h_0\}) \neq \emptyset \quad \forall \epsilon > 0,$$

ou seja, a origem $0 \in \mathbb{R}^n$ é não isolada.

(\Rightarrow) A recíproca é óbvia. ■

A demonstração do próximo resultado, de natureza puramente algébrica, pode ser encontrada em [10], Proposição 4.2.

Lema 3.4 *Se $H \subset \mathbb{R}^n$ é um subgrupo discreto não-trivial, então existem e_1, \dots, e_k em \mathbb{R}^n linearmente independentes tais que*

$$H = \left\{ \sum_{i=1}^k z_i e_i \mid z_i \in \mathbb{Z} \right\}.$$

Neste caso, o subgrupo H é chamado de **reticulado**.

No próximo resultado, denotaremos por \mathbb{T}^k o quociente de um espaço euclidiano \mathbb{R}^k por um reticulado $H \subset \mathbb{R}^k$. Com essa notação, estabelecemos um resultado sobre linearização global:

Teorema 3.5 (Linearização global) *Sejam M uma variedade conexa de dimensão n , e f_0, f_1, \dots, f_m campos de vetores em M , $m \leq n$. Existe um difeomorfismo global*

$$\Phi : M \rightarrow \mathbb{T}^k \times \mathbb{R}^{n-k} \quad k \leq n$$

tal que, para todo $x \in \mathbb{T}^k \times \mathbb{R}^{n-k}$, temos

$$(\Phi_* f_0)(x) = Ax + c, \quad (3.2-14)$$

$$(\Phi_* f_i)(x) = b_i, \quad i = 1, \dots, m \quad (3.2-15)$$

onde $A \in M_n(\mathbb{R})$ é uma matriz real constante, com as primeiras k -linhas nulas:

$$A e_i = 0 \quad i = 1, \dots, k \quad (3.2-16)$$

e $c, b_i \in \mathbb{R}^n$ são constantes satisfazendo as condições de Kalman se, e somente se,

$$0 = \left[(\text{ad } f_0)^{j_1} f_{i_1}, (\text{ad } f_0)^{j_2} f_{i_2} \right] (q), \quad (3.2-17)$$

$$T_q M = \text{span}\{(\text{ad } f_0)^j f_i(q)\}, \quad (3.2-18)$$

$$(\text{ad } f_0)^j f_i \quad \text{— são campos de vetores completos,} \quad (3.2-19)$$

para todo $0 \leq j, j_1, j_2 \leq n-1, 1 \leq i, i_1, i_2 \leq m$, e $q \in M$.

Demonstração: (\Leftarrow) Fixe $q_0 \in M$. Pela hipótese (3.2-18), podemos encontrar n vetores da forma $Y_\alpha = (\text{ad } f_0)^{j_\alpha} f_{i_\alpha}$ tais que

$$\text{span}\{Y_1(q_0), \dots, Y_n(q_0)\} = T_{q_0} M,$$

para $\alpha = 1, \dots, n$. Vamos mostrar que Y_α , para $\alpha = 1, \dots, n$, são linearmente independentes em todo ponto de M . Considere o conjunto

$$L = \{q \in M \mid \text{span}\{Y_1(q), \dots, Y_n(q)\} = T_q M\}.$$

Afirmamos que $L = M$. De fato, L é um conjunto aberto pois, dado um ponto $x \in L$, temos que $\text{span}\{Y_1(x), \dots, Y_n(x)\} = T_x M$. Como os vetores $Y_\alpha(x)$ para $\alpha = 1, \dots, n$ são linearmente independentes em x , são também linearmente independentes em uma vizinhança aberta V de x , ou seja,

$$\text{span}\{Y_1(y), \dots, Y_n(y)\} = T_y M \quad \forall y \in V,$$

logo $x \in V \subset L$.

Vamos mostrar que L é um conjunto fechado. Para todo i, j nas condições dadas e $q \in L$, temos que

$$(\text{ad} f_0)^j f_i(q) = \sum_{\alpha=1}^n a_{\alpha}^{ij} Y_{\alpha}(q),$$

onde $a_{\alpha}^{ij} \in \mathcal{C}^{\infty}(L)$. Mostraremos que a_{α}^{ij} são constantes. De fato, por (3.2-17), e a regra de Leibniz, temos

$$\begin{aligned} 0 &= \left[Y_{\beta}, \sum_{\alpha=1}^n a_{\alpha}^{ij} Y_{\alpha} \right] = Y_{\beta} \left(\sum_{\alpha=1}^n a_{\alpha}^{ij} Y_{\alpha} \right) + \sum_{\alpha=1}^n a_{\alpha}^{ij} [Y_{\beta}, Y_{\alpha}] = Y_{\beta} \left(\sum_{\alpha=1}^n a_{\alpha}^{ij} Y_{\alpha} \right) \\ &= \sum_{\alpha=1}^n (Y_{\beta} a_{\alpha}^{ij}) Y_{\alpha}. \end{aligned}$$

Como Y_{α} são linearmente independentes então,

$$Y_{\beta} a_{\alpha}^{ij} = 0, \quad \forall \alpha, \beta = 1, \dots, n.$$

Lembrando que $(Y_{\beta} a_{\alpha}^{ij})(q) = D_q a_{\alpha}^{ij} \cdot Y_{\beta}(q)$ (ver subseção 1.4.1), temos

$$Y_{\beta} a_{\alpha}^{ij} = 0 \quad \Rightarrow \quad D_q a_{\alpha}^{ij} = 0, \quad \forall q \in L,$$

que implica em

$$a_{\alpha}^{ij}(q) = K_{\alpha}^{ij}, \quad \forall q \in L$$

onde K_{α}^{ij} são constantes reais.

Agora, dado qualquer $x \in \bar{L}$, existe uma sequência $x_n \in L$ tal que $x_n \rightarrow x$. Como $x_n \in L$ então

$$(\text{ad} f_0)^j f_i(x_n) = \sum_{\alpha=1}^n K_{\alpha}^{ij} Y_{\alpha}(x_n)$$

como $x_n \rightarrow x$ e $(\text{ad} f_0)^j f_i$ são contínuas então

$$(\text{ad} f_0)^j f_i(x_n) \rightarrow (\text{ad} f_0)^j f_i(x)$$

e

$$\sum_{\alpha=1}^n K_{\alpha}^{ij} Y_{\alpha}(x_n) \rightarrow \sum_{\alpha=1}^n K_{\alpha}^{ij} Y_{\alpha}(x).$$

Como limite é único, obtemos

$$(\text{ad} f_0)^j f_i(x) = \sum_{\alpha=1}^n K_{\alpha}^{ij} Y_{\alpha}(x)$$

e, portanto, $x \in L$. Assim, o conjunto L é não vazio, pois $q_0 \in L$, aberto e fechado em M . Como M é conexo, segue que $L = M$ e os campos Y_{α} , para $\alpha = 1, \dots, n$, são linearmente independentes em M .

Como os campos $(\text{ad} f_0)^j f_i$ são completos, (hipótese (3.2-19)), podemos definir a aplicação

$$\begin{aligned} \Psi: \mathbb{R}^n &\rightarrow M, & x &= (x_1, \dots, x_n) \\ x &\mapsto & e^{x_1 Y_1} \circ \dots \circ e^{x_n Y_n}(q_0). \end{aligned}$$

Pela hipótese (3.2-17), temos que os campos de vetores comutam, logo

$$\Psi(x_1, \dots, x_n) = e^{\sum_{\alpha=1}^n x_{\alpha} Y_{\alpha}}(q_0).$$

Temos que Ψ é diferenciável e a diferencial é um isomorfismo. De fato, para cada $\alpha = 1, \dots, n$

$$\begin{aligned} \left. \frac{\partial \Psi}{\partial x_{\alpha}}(x) \right|_{x=0} &= \left. \frac{d}{d\epsilon} \Psi(x_1, \dots, x_{\alpha} + \epsilon, \dots, x_n) \right|_{\epsilon=0} \\ &= \left. \frac{d}{d\epsilon} e^{\sum_{\beta=1}^n x_{\beta} Y_{\beta}} \circ e^{\epsilon Y_{\alpha}}(q_0) \right|_{\epsilon=0} \\ &= \left. e^{\sum_{\beta=1}^n x_{\beta} Y_{\beta}} \circ e^{\epsilon Y_{\alpha}}(q_0) \circ Y_{\alpha} \right|_{\epsilon=0} \\ &= e^{\sum_{\beta=1}^n x_{\beta} Y_{\beta}}(q_0) \circ Y_{\alpha} \\ &= Y_{\alpha}(\Psi(x)). \end{aligned}$$

Desta forma, a diferencial $\Psi_{*x}: \mathbb{R}^n \rightarrow T_{q_0} M$ transforma a base canônica de \mathbb{R}^n na base $\{Y_1, \dots, Y_n\}$ de $T_{q_0} M$. Logo Ψ_{*x} é um isomorfismo em cada ponto $x \in \mathbb{R}^n$. Como consequência, Ψ é um difeomorfismo local em cada $x \in \mathbb{R}^n$.

Vamos mostrar que $\Psi(\mathbb{R}^n) = M$. De fato, $\Psi(\mathbb{R}^n)$ é aberto em M pois, dado qualquer $y \in \Psi(\mathbb{R}^n)$ existem $x \in \mathbb{R}^n$, abertos $U_x \subset \mathbb{R}^n$ e $V_y \subset M$ tais que

$$\Psi: U_x \rightarrow V_y$$

é um difeomorfismo local. Logo, $\Psi(U_x)$ é uma vizinhança aberta de y contida em $\Psi(\mathbb{R}^n)$. Temos também que $\Psi(\mathbb{R}^n)$ é fechado em M . De fato, dado qualquer $q \in \overline{\Psi(\mathbb{R}^n)}$ fixado, podemos definir a aplicação

$$\Phi: \mathbb{R}^n \rightarrow M,$$

dada por

$$\Phi(y_1, \dots, y_n) = e^{\sum_{\alpha=1}^n y_{\alpha} Y_{\alpha}}(q),$$

que satisfaz as mesmas propriedades que a aplicação Ψ . Como consequência, $\Phi(\mathbb{R}^n)$ é um subconjunto aberto contendo q . Como $q \in \overline{\Psi(\mathbb{R}^n)}$, então

$$\Phi(\mathbb{R}^n) \cap \Psi(\mathbb{R}^n) \neq \emptyset.$$

Segue que existe $y' \in \mathbb{R}^n$ tal que $\Phi(y') \in \Psi(\mathbb{R}^n)$, logo $\Phi(y') = \Psi(x')$ para algum $x' \in \mathbb{R}^n$, temos

$$\Phi(y') = e^{\sum_{\alpha=1}^n y'_{\alpha} Y_{\alpha}}(q) = e^{\sum_{\alpha=1}^n x'_{\alpha} Y_{\alpha}}(q) = \Psi(x').$$

Com isso, e novamente pelo fato dos campos de vetores serem completos,

$$q = e^{\sum_{\alpha=1}^n (x_{\alpha} - y_{\alpha}) Y_{\alpha}}(q_0) = \Psi(x - y).$$

Segue que existe $z = x - y \in \mathbb{R}^n$ tal que $\Psi(z) = q$, isto é, $q \in \Psi(\mathbb{R}^n)$ concluindo que $\Psi(\mathbb{R}^n)$ é fechado. Como $\Psi(\mathbb{R}^n)$ é não vazio, aberto e fechado em M , que é conexa, concluímos que $\Psi(\mathbb{R}^n) = M$.

Defina o conjunto

$$\mathbb{G}_0 = \Psi^{-1}(q_0) = \{x \in \mathbb{R}^n \mid \Psi(x) = q_0\} \subset \mathbb{R}^n.$$

Note que $\mathbb{G}_0 \neq \emptyset$ pois $0 \in \mathbb{G}_0$. Além disso \mathbb{G}_0 é um subgrupo de \mathbb{R}^n pois, se $x, y \in \mathbb{G}_0$ então,

$$\Psi(x) = e^{\sum_{\alpha=1}^n x_{\alpha} Y_{\alpha}}(q_0) = q_0 = e^{\sum_{\alpha=1}^n y_{\alpha} Y_{\alpha}}(q_0) = \Psi(y)$$

logo,

$$\Psi(x + y) = e^{\sum_{\alpha=1}^n (x_{\alpha} + y_{\alpha}) Y_{\alpha}}(q_0) = e^{\sum_{\alpha=1}^n x_{\alpha} Y_{\alpha}} \circ e^{\sum_{\alpha=1}^n y_{\alpha} Y_{\alpha}}(q_0) = e^{\sum_{\alpha=1}^n y_{\alpha} Y_{\alpha}}(q_0) = q_0.$$

Assim mostramos que $x + y \in \mathbb{G}_0$. Mais ainda, dado $x \in \mathbb{G}_0$ temos

$$\Psi(-x) = q_0 \circ e^{\sum_{\alpha=1}^n (-x_\alpha) Y_\alpha} = q_0 \circ e^{-\sum_{\alpha=1}^n x_\alpha Y_\alpha} = q_0$$

pois $q_0 \circ e^{-\sum_{\alpha=1}^n x_\alpha Y_\alpha}$ é a trajetória reversa de $q_0 \circ e^{\sum_{\alpha=1}^n x_\alpha Y_\alpha}$. Concluimos aqui que \mathbb{G}_0 é um subgrupo de \mathbb{R}^n . Além disso, \mathbb{G}_0 é discreto pois na origem $0 \in \mathbb{R}^n$ existem abertos $U_0 \subset \mathbb{R}^n$ e $V_{q_0} \subset M$ tais que $\Psi: U_0 \rightarrow V_{q_0}$ é um difeomorfismo local. Isso significa que não pode existir um ponto $x \in U_0 \cap \mathbb{G}_0$ com $x \neq 0$, caso contrario, $\Psi(0) = \Psi(x) = q_0$, violando a injetividade de Ψ em U_0 . Logo, a origem é um ponto isolado de \mathbb{G}_0 em \mathbb{R}^n . Portanto, pela Proposição 3.3, \mathbb{G}_0 é um subgrupo discreto de \mathbb{R}^n .

Definimos a aplicação

$$\begin{aligned} \bar{\Psi}: \frac{\mathbb{R}^n}{\mathbb{G}_0} &\rightarrow M \\ \bar{x} &\mapsto \Psi(x), \end{aligned}$$

onde $x \in \mathbb{R}^n$ é um representante da classe \bar{x} . Temos que $\bar{\Psi}$ é bem definida e injetora pois,

$$\begin{aligned} \bar{x} = \bar{y} &\Leftrightarrow \overline{x - y} = \bar{0} \\ &\Leftrightarrow x - y \in \mathbb{G}_0 \\ &\Leftrightarrow \Psi(x - y) = q_0 \\ &\Leftrightarrow e^{\sum_{\alpha=1}^n (x_\alpha - y_\alpha) Y_\alpha}(q_0) = q_0 \\ &\Leftrightarrow e^{\sum_{\alpha=1}^n x_\alpha Y_\alpha}(q_0) = e^{\sum_{\alpha=1}^n y_\alpha Y_\alpha}(q_0) \\ &\Leftrightarrow \Psi(x) = \Psi(y) \\ &\Leftrightarrow \bar{\Psi}(\bar{x}) = \bar{\Psi}(\bar{y}). \end{aligned}$$

Mais ainda, $\bar{\Psi}$ é um difeomorfismo local devido a Ψ . Por simplicidade, denote $\bar{\Psi} = \Psi$.

Temos assim

$$\Psi: \frac{\mathbb{R}^n}{\mathbb{G}_0} \rightarrow M$$

é uma aplicação injetora e um difeomorfismo local em cada ponto de $\frac{\mathbb{R}^n}{\mathbb{G}_0}$, consequentemente Ψ é um difeomorfismo global. Pelo Lema 3.4, \mathbb{G}_0 é reticulado. Logo, o quociente é um cilindro:

$$\frac{\mathbb{R}^n}{\mathbb{G}_0} = \mathbb{T}^k \times \mathbb{R}^{n-k}.$$

Finalmente, obtemos um difeomorfismo global:

$$\Phi = \Psi^{-1}: M \rightarrow \mathbb{T}^k \times \mathbb{R}^{n-k}.$$

Resta mostrar que Φ satisfaz as condições desejadas. De fato, a hipótese (3.2-18) mostra que a condição de controlabilidade (2.2-3) é satisfeita, com a mesma demonstração do Teorema 3.2.

Mostramos que Ψ_* é isomorfismo verificando as

$$\Phi_* \left(\frac{\partial}{\partial x_\alpha} \right) = Y_\alpha, \quad \alpha = 1, \dots, n.$$

Logo,

$$\Phi_*(Y_\alpha) = \Psi_*^{-1}(Y_\alpha) = \frac{\partial}{\partial x_\alpha}.$$

Visto isso, com a hipótese (3.2-17) mostramos (3.2-14) e (3.2-15) exatamente como foi feito na demonstração do Teorema 3.2.

Finalmente, o campo $A\bar{x} + c$ está bem definido no quociente $\frac{\mathbb{R}^n}{\mathbb{G}_0}$ pois Φ_* é um isomorfismo. Se $x \in \mathbb{G}_0$ então,

$$\bar{x} = \bar{0} \Rightarrow A\bar{x} + c = A\bar{0} + c = c \Rightarrow A\bar{x} = \bar{0}.$$

Como e_i formam uma base para o reticulado \mathbb{G}_0 , temos que

$$e_i \in \mathbb{G}_0 \Rightarrow Ae_i = \bar{0},$$

para todo $i = 1, \dots, k$. Isto mostra (3.2-16).

(\Rightarrow) Reciprocamente, temos o sistema linear

$$\begin{aligned} (\Phi_* f_0)(x) &= Ax + c, \\ (\Phi_* f_i)(x) &= b_i, \quad i = 1, \dots, m \end{aligned}$$

sobre o cilindro $\frac{\mathbb{R}^n}{\mathbb{G}_0}$. Note que

$$(-A^j b_i) = (\text{ad} \Phi_* f_0)^j \Phi_* f_i = \Phi_* (\text{ad} f_0)^j f_i \Rightarrow \Phi_*^{-1}(-A^j b_i) = (\text{ad} f_0)^j f_i.$$

Como os campos $(-A^j b_i)$ são constantes, não dependem do tempo t . Consequentemente $(\text{ad} f_0)^j f_i$ não dependem de t , logo são campos completos, mostrando (3.2-19). Além disso,

$$0 = \left[(\text{ad} \Phi_* f_0)^{j_1} \Phi_* f_{i_1}, (\text{ad} \Phi_* f_0)^{j_2} \Phi_* f_{i_2} \right] = \Phi_* \left[(\text{ad} f_0)^{j_1} f_{i_1}, (\text{ad} f_0)^{j_2} f_{i_2} \right].$$

Como Φ_* é isomorfismo, obtemos

$$\left[(\text{ad} f_0)^{j_1} f_{i_1}, (\text{ad} f_0)^{j_2} f_{i_2} \right] = 0, \quad \forall q \in M,$$

mostrando (3.2-17). Finalmente, de (2.2-3) temos que

$$\text{span}\{(\text{ad} \Phi_* f_0)^j \Phi_* f_i(x)\} = T_x \frac{\mathbb{R}^n}{\mathbb{G}_0}.$$

Como Φ_* é isomorfismo, podemos obter

$$\Phi_{*x}^{-1} : T_x \frac{\mathbb{R}^n}{\mathbb{G}_0} \rightarrow T_q M, \quad x = \Phi_*(q).$$

Portanto,

$$\text{span}\{(\text{ad} f_0)^j f_i(q)\} = T_q M,$$

mostrando (3.2-18). ■

Exemplo 3.6 Considere o sistema não linear sobre \mathbb{R}^3 :

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} x_2 - 2x_2x_3 + x_3^2 \\ x_3 \\ 0 \end{pmatrix} + \begin{pmatrix} 4x_2x_3 \\ -2x_3 \\ 1 \end{pmatrix} u, \quad (3.2-20)$$

com

$$f = \begin{pmatrix} x_2 - 2x_2x_3 + x_3^2 \\ x_3 \\ 0 \end{pmatrix}, \quad \text{e} \quad g = \begin{pmatrix} 4x_2x_3 \\ -2x_3 \\ 1 \end{pmatrix}.$$

Temos que

$$(\text{ad} f) g(x) = [f, g](x) = \text{Jac}(g(x))f(x) - \text{Jac}(f(x))g(x),$$

que é:

$$\begin{pmatrix} 0 & 4x_3 & 4x_2 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_2 - 2x_2x_3 + x_3^2 \\ x_3 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 - 2x_3 & -2x_2 + 2x_3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 4x_2x_3 \\ -2x_3 \\ 1 \end{pmatrix},$$

e fazendo os cálculos, obtemos

$$(\mathbf{ad}f)g(x) = \begin{pmatrix} 2x_2 \\ -1 \\ 0 \end{pmatrix}.$$

Temos também que $(\mathbf{ad}f)^2 g(x) = [f, [f, g]] = \text{Jac}([f, g](x))f(x) - \text{Jac}(f(x))[f, g](x)$, que é:

$$\begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_2 - 2x_2x_3 + x_3^2 \\ x_3 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 - 2x_3 & -2x_2 + 2x_3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2x_2 \\ -1 \\ 0 \end{pmatrix},$$

resolvendo, obtemos

$$(\mathbf{ad}f)^2 g(x) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Com cálculos simples, podemos mostrar que o conjunto $\{g, (\mathbf{ad}f)g, (\mathbf{ad}f)^2 g\}$ é um conjunto linearmente independente para todo $x \in \mathbb{R}^3$. Consequentemente,

$$\text{span}\{g, (\mathbf{ad}f)g, (\mathbf{ad}f)^2 g\} = \mathbb{R}^3,$$

para todo $x \in \mathbb{R}^3$. Analogamente, obtemos também que

$$[(\mathbf{ad}f)^j g, (\mathbf{ad}f)^l g](x) = 0,$$

para todo $x \in \mathbb{R}^3$ e $0 \leq j, l \leq 2$. Finalmente, como $g, (\mathbf{ad}f)g$, e $(\mathbf{ad}f)^2 g$ não dependem do tempo t , segue que são campos completos. Desta forma o Teorema 3.5 garante que existe um difeomorfismo global Φ que transforma o sistema não linear (3.2-20) em um sistema linear

controlável. De fato, defina $\Phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dada por

$$\Phi(x_1, x_2, x_3) = (x_1 + x_2^2, x_2 + x_3^2, x_3) = (z_1, z_2, z_3).$$

Assim, o sistema (3.2-20) com as novas coordenadas (z_1, z_2, z_3) têm a forma

$$\begin{cases} \dot{z}_1 = z_2 \\ \dot{z}_2 = z_3 \\ \dot{z}_3 = u, \end{cases}$$

que é um sistema linear em \mathbb{R}^3 . Podemos escrever este sistema como

$$\dot{z} = A(z) + bu,$$

onde

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

e obtemos $\text{span}\{b, Ab, A^2b\} = \text{span}\{(0, 0, 1)^T, (0, 1, 0)^T, (1, 0, 0)^T\} = \mathbb{R}^3$, satisfazendo a condição de Kalman.

CAPÍTULO 4

O TEOREMA DA ÓRBITA E APLICAÇÕES

Neste capítulo retornaremos ao estudo dos conjuntos atingíveis definido em 1.27, mas agora através de um conjunto com uma estrutura mais simples, a órbita de uma família $\mathcal{F} \subset \text{Vec}(M)$. Esta têm a propriedade bastante natural de ser subvariedade imersa de M , como será demonstrado no Teorema da Órbita. Primeiramente trataremos das subvariedades imersas, que são requisitos importantes para introduzirmos o Teorema da Órbita. Veremos também como alguns resultados ficam simplificados no caso analítico, e concluiremos este capítulo retornando ao estudo dos conjuntos atingíveis.

4.1 Subvariedades imersas

Esta seção tem o objetivo de fixar a definição de subvariedade que aparece naturalmente na descrição do Teorema da Órbita. O conceito de subvariedade tem interpretações distintas na literatura, dependendo do contexto ou necessidade. Para o nosso objetivo, trabalharemos com subvariedades imersas.

Definição 4.1 *Sejam M, N variedades de dimensões n, k respectivamente e $\varphi : M \rightarrow N$ uma aplicação diferenciável. Dizemos que φ é uma **imersão** se, para cada $q \in M$, a diferencial no*

ponto q é uma aplicação injetora, isto é

$$D_q\varphi : T_qM \rightarrow T_{\varphi(q)}N, \quad \text{com} \quad \ker D_q\varphi = \{0\}.$$

Definição 4.2 Dizemos que um subconjunto $W \subset M$ é uma **subvariedade imersa** de dimensão $k \leq n$ de M se existir uma imersão injetora de uma variedade N de dimensão k em M :

$$\Phi : N \rightarrow M,$$

tal que

$$W = \Phi(N).$$

De forma equivalente, uma subvariedade imersa W de M pode ser definida como uma variedade contida em M tal que a inclusão

$$i : W \rightarrow M, \quad q \mapsto q$$

é uma imersão. Desta forma, temos que se W é uma variedade contida em M e para cada $p \in W$ existe uma vizinhança V de p em W tal que V é subvariedade de M , então W é subvariedade de M . Usaremos este fato para mostrarmos o Teorema da Órbita.

Observação 4.3 Uma subvariedade imersa em geral herda uma topologia menos fina que a topologia induzida pelo espaço ambiente, logo não é necessariamente uma subvariedade mergulhada, de acordo com a Definição 1.1. Porém, vizinhanças apropriadas suficientemente pequenas em uma subvariedade imersa são subvariedades mergulhadas.

Exemplo 4.4 Seja $\Phi : \mathbb{R} \rightarrow \mathbb{R}^2$ uma imersão injetora tal que

$$\lim_{t \rightarrow +\infty} \Phi(t) = \Phi(0).$$

A imagem $\Phi(\mathbb{R})$ é uma subvariedade imersa de \mathbb{R}^2 . Os conjuntos $\Phi(-\epsilon, \epsilon)$ para todo $\epsilon > 0$, são abertos na topologia herdada de \mathbb{R} mas não são abertos na topologia induzida de \mathbb{R}^2 . Logo,

$$\Phi : \mathbb{R} \rightarrow \mathbb{R}^2$$

não é um mergulho. Portanto $\Phi(\mathbb{R})$ não é subvariedade no sentido da definição 1.1.

Um exemplo menos trivial e mais interessante é:

Exemplo 4.5 (Curva de Kronecker) *Considere o oscilador com 2 graus de liberdade descrito pelas equações:*

$$\begin{aligned} \dot{z}(t) &= i\alpha z(t), & z(t) &\in \mathbb{C} \\ \dot{w}(t) &= i\beta w(t), & w(t) &\in \mathbb{C}, \end{aligned} \quad (4.1-1)$$

com α e β constantes reais não-nulas. Este possui as seguintes soluções:

$$\begin{aligned} z(t) &= z(0)e^{i\alpha t}, \\ w(t) &= w(0)e^{i\beta t}. \end{aligned} \quad (4.1-2)$$

Repare que a norma de $z(t)$, é:

$$\begin{aligned} |z(t)| &= |z(0)e^{i\alpha t}| \\ &= |z(0)| \cdot |\cos(\alpha t) + i \sin(\alpha t)| \\ &= |z(0)| \cdot \sqrt{\cos^2(\alpha t) + \sin^2(\alpha t)} \\ &= |z(0)| = \text{constante}. \end{aligned}$$

Analogamente obtemos, $|w(t)| = |w(0)| = \text{constante}$.

Logo, toda solução de (4.1-1) pertence ao **Toro invariante**:

$$\mathbb{T}^2 = \{(z, w) \in \mathbb{C}^2; |z| = k_1 \text{ e } |w| = k_2, \quad k_1, k_2 \in \mathbb{R}\}.$$

Note que a recíproca é verdadeira, isto é, cada par $(z, w) \in \mathbb{T}^2$ corresponde a condições iniciais para (4.1-1), e uma solução da forma (4.1-2). Lembrando que este toro é parametrizado pelos argumentos de z e w módulo 2π , obtemos

$$\mathbb{T}^2 \simeq \frac{\mathbb{C}}{(2\pi\mathbb{Z})^2} \simeq \frac{\mathbb{R}^2}{(2\pi\mathbb{Z})^2}.$$

Introduzimos um novo parâmetro $\tau = \alpha t$ que implica em $\frac{\beta}{\alpha}\tau = \beta t$. Temos assim, a reta que

passa pela origem com ângulo $\frac{\beta}{\alpha}$

$$\begin{aligned}\lambda: \mathbb{R} &\rightarrow \mathbb{R}^2 \\ \tau &\mapsto \lambda(\tau) = \frac{\beta}{\alpha}\tau,\end{aligned}$$

com gráfico $\text{Graf}(\lambda) = \{(\tau, \frac{\beta}{\alpha}\tau) \mid \tau \in \mathbb{R}\} \subset \mathbb{R}^2$.

Defina a aplicação

$$\begin{aligned}\varphi: \text{Graf}(\lambda) &\rightarrow \frac{\mathbb{R}^2}{(2\pi\mathbb{Z})^2} \\ (\tau, \frac{\beta}{\alpha}\tau) &\mapsto (\tau + 2\pi\mathbb{Z}, \frac{\beta}{\alpha}\tau + 2\pi\mathbb{Z}),\end{aligned}$$

ou seja, as soluções $(z, w) \in \frac{\mathbb{R}^2}{(2\pi\mathbb{Z})^2}$ são imagens da reta $\lambda(\tau)$. Claramente φ é injetora, e uma imersão pois a matriz Jacobiana

$$\text{Jac}(\varphi) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{\beta}{\alpha} \end{pmatrix}$$

possui determinante não-nulo. Logo, as soluções de (4.1-1) são imagens de uma imersão injetora, e portanto são subvariedades imersas do Toro invariante.

Se $\frac{\beta}{\alpha}$ é irracional, as soluções são conhecidas como **Curvas de Kronecker** que, intuitivamente, são enrolamentos **densos** de retas $\lambda(t)$ no toro. Portanto, cada Curva de Kronecker é uma subvariedade imersa mas, não uma subvariedade.

Observação 4.6 As subvariedades imersas herdam muitas propriedades locais das subvariedades. Enunciaremos as que serão úteis posteriormente.

(i) Seja W uma subvariedade imersa de M . Assim, existe uma imersão injetora Φ de uma variedade N em M tal que $\Phi(N) = W$. Restringindo à imagem, a diferencial de Φ num ponto $p \in N$

$$D_p\Phi: T_pN \rightarrow T_qW, \quad \Phi(p) = q$$

é isomorfismo, logo

$$D_p\Phi(T_pN) = T_qW.$$

(ii) Seja $V \in \text{Vec}(M)$. Então,

$$V(q) \in T_qW, \quad \forall q \in W \iff e^{tV}(q) \in W, \quad (4.1-3)$$

para $|t|$ suficientemente pequeno.

A demonstração é análoga à demonstração do Teorema 1.14.

4.2 O Teorema da Órbita

Seja $\mathcal{F} \subset \text{Vec}(M)$ uma família de campos de vetores. Vamos estudar o conjunto atingível de \mathcal{F} a partir de um ponto $q_0 \in M$ num tempo $t \geq 0$,

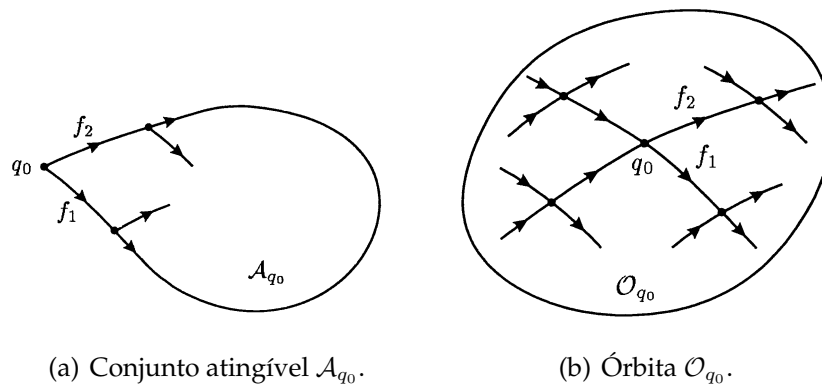
$$\mathcal{A}_{q_0}(t) = \left\{ e^{t_1 f_1} \circ \dots \circ e^{t_k f_k}(q_0) \mid f_i \in \mathcal{F}, \sum_{i=1}^k t_i = t, k \in \mathbb{N} \right\}$$

usando a órbita a partir de um ponto, que é definida a seguir:

Definição 4.7 A Órbita de uma família \mathcal{F} a partir de um ponto $q_0 \in M$ é definida por:

$$\mathcal{O}_{q_0} = \{ e^{t_1 f_1} \circ \dots \circ e^{t_k f_k}(q_0) \mid f_i \in \mathcal{F}, t_i \in \mathbb{R}, k \in \mathbb{N} \}. \quad (4.2-4)$$

Abaixo comparamos o conjunto de pontos atingíveis para uma família \mathcal{F} partindo de um ponto q_0 num tempo $t \geq 0$ (figura (a)), com a órbita partindo do mesmo ponto e para a mesma família (figura (b)).



Observação 4.8 É importante notar que no conjunto atingível temos $t \geq 0$, e para a órbita temos $t \in \mathbb{R}$. Intuitivamente, a órbita é um conjunto que abrange os pontos atingíveis, isto é, $\mathcal{A}_{q_0} \subset \mathcal{O}_{q_0}$. Repare que a igualdade $\mathcal{A}_{q_0} = \mathcal{O}_{q_0}$, depende da simetria da família \mathcal{F} . De modo formal, se

$$\mathcal{F} = -\mathcal{F},$$

então

$$\begin{aligned}\mathcal{O}_{q_0} \ni e^{-t_1 f_1} \circ \dots \circ e^{-t_k f_k}(q_0) &= e^{t_1(-f_1)} \circ \dots \circ e^{t_k(-f_k)}(q_0) \\ &= e^{t_1 f_1} \circ \dots \circ e^{t_k f_k}(q_0) \in \mathcal{A}_{q_0}.\end{aligned}$$

Assim, $\mathcal{O}_{q_0} = \mathcal{A}_{q_0}$. Defina em M uma relação (\sim):

$$q_1 \sim q_2 \Leftrightarrow q_1 \in \mathcal{O}_{q_2}.$$

A relação (\sim) é:

(i) Reflexiva pois, claramente $q_1 \in \mathcal{O}_{q_1}$, logo $q_1 \sim q_1$.

(ii) Simétrica pois, se $q_1 \in \mathcal{O}_{q_2}$, então

$$q_1 = e^{t_1 f_1} \circ \dots \circ e^{t_k f_k}(q_2), \quad t_i \in \mathbb{R}, f_i \in \mathcal{F}.$$

Logo,

$$e^{s_k f_k} \circ \dots \circ e^{s_1 f_1}(q_1) = q_2,$$

onde $s_i = -t_i \in \mathbb{R}$. Portanto, $q_2 \in \mathcal{O}_{q_1} \Rightarrow q_2 \sim q_1$.

(iii) Transitiva pois, se $q_1 \sim q_2$ e $q_2 \sim q_3$ então

$$q_1 \in \mathcal{O}_{q_2} \Rightarrow q_1 = e^{t_1 f_1} \circ \dots \circ e^{t_k f_k}(q_2)$$

e,

$$q_2 \in \mathcal{O}_{q_3} \Rightarrow q_2 = e^{s_1 g_1} \circ \dots \circ e^{s_k g_k}(q_3),$$

onde $s_i \in \mathbb{R}$ e $g_i \in \mathcal{F}$. Logo,

$$q_1 = e^{t_1 f_1} \circ \dots \circ e^{t_k f_k} \circ e^{s_1 g_1} \circ \dots \circ e^{s_k g_k}(q_3).$$

Portanto, $q_1 \in \mathcal{O}_{q_3} \Rightarrow q_1 \sim q_3$.

Concluimos que (\sim) é uma relação de equivalência em M . Consequentemente, as classes de equivalência de

$$\frac{M}{\sim} = \{\mathcal{O}_q \mid q \in M\}$$

são órbitas.

Introduziremos as seguintes notações

$$(\text{Ad}\mathcal{P})\mathcal{F} \stackrel{\text{def}}{=} \{(\text{Ad}P)f \mid P \in \mathcal{P}, f \in \mathcal{F}\},$$

onde

$$\mathcal{P} = \{e^{t_1 f_1} \circ \dots \circ e^{t_k f_k} \mid t_i \in \mathbb{R}, f_i \in \mathcal{F}, k \in \mathbb{N}\} \subset \text{Diff}(M)$$

é o subgrupo de difeomorfismos gerado pelos fluxos correspondentes aos campos $f_i \in \mathcal{F}$, e

$$\Pi_q \stackrel{\text{def}}{=} \text{span}\{(\text{Ad}\mathcal{P})\mathcal{F}(q)\} \subset T_q M, \quad q \in M.$$

Lema 4.9 Para $q_0 \in M$, $\dim \Pi_q = \dim \Pi_{q_0}$ para todo $q \in \mathcal{O}_{q_0}$.

Demonstração: Dado $q \in \mathcal{O}_{q_0}$ podemos escrever $q = Q(q_0)$, com $Q \in \mathcal{P}$.

Seja

$$(\text{Ad}P)f(q_0) \in \Pi_{q_0}$$

um elemento arbitrário. Lembrando a subseção 1.4.2, temos

$$\begin{aligned} Q_*((\text{Ad}P)f(q_0)) &= q_0 \circ (\text{Ad}P)f \circ Q \\ &= q_0 \circ P \circ f \circ P^{-1} \circ Q \\ &= (q_0 \circ Q) \circ (Q^{-1} \circ P \circ f \circ P^{-1} \circ Q) \\ &= q \circ (\text{Ad}(Q^{-1} \circ P)f) \\ &= (\text{Ad}(Q^{-1} \circ P)f(q)) \in \Pi_q, \end{aligned}$$

pois $Q^{-1} \circ P \in \mathcal{P}$. Isto mostra que $Q_*\Pi_{q_0} \subset \Pi_q$. Como R_* é isomorfismo, concluímos que $\dim \Pi_{q_0} \leq \dim \Pi_q$. Agora, podemos escrever $q_0 = Q^{-1}(q)$ e aplicar o argumento acima, invertendo os papéis de q e q_0 . Assim, obtemos a desigualdade oposta e, conseqüentemente, a igualdade. ■

Enunciamos agora o principal resultado deste capítulo.

Teorema 4.10 (Teorema da Órbita (Nagano-Sussmann)) *Sejam $\mathcal{F} \subset \text{Vec}(M)$ uma família de campos de vetores e $q_0 \in M$. Então*

- (i) A órbita \mathcal{O}_{q_0} é uma subvariedade conexa imersa de M ,
- (ii) $T_q\mathcal{O}_{q_0} = \text{span}\{(\text{Ad}P)f(q) \mid P \in \mathcal{P}, f \in \mathcal{F}\}$ para todo $q \in \mathcal{O}_{q_0}$.

A demonstração do Teorema da Órbita é longa e será dividida em diversos lemas, que terão natureza extremamente técnica. Antes de enunciarmos tais lemas, daremos uma ideia geral da demonstração do Teorema da Órbita: para cada $q \in M$, definiremos certas aplicações ξ_q tendo como domínio um espaço \mathbb{R}^k , para k menor ou igual a dimensão de M , e cuja imagem está contida na órbita através de q . Mostraremos que restringindo convenientemente seu domínio, ξ_q é uma imersão. Portanto, as imagens destas aplicações serão subvariedades imersas de M . Mostraremos que essas imagens dão origem a uma topologia em M , mais fina que a topologia original. Em relação a essa topologia, as órbitas serão as componentes conexas de M (e, conseqüentemente, serão conexas). À partir daí usamos novamente as aplicações ξ_q para munir cada órbita de uma estrutura de variedade diferenciável. Como as imagens das aplicações ξ_q são abertas na órbita (em relação à nova topologia), temos que cada ponto da órbita possui uma vizinhança que é uma variedade imersa, de onde concluímos que a própria órbita é subvariedade imersa conexa. A prova do item (ii) será consequência dessa construção.

Lema 4.11 Para cada $q \in M$, escolha $V_1, \dots, V_m \in (\text{Ad}\mathcal{P})\mathcal{F}$ tais que

$$\text{span}\{V_1(q), \dots, V_m(q)\} = \Pi_q$$

e defina

$$\begin{aligned} \xi_q : \quad \mathbb{R}^m &\rightarrow M \\ (t_1, t_2, \dots, t_m) &\mapsto e^{t_1 V_1} \circ \dots \circ e^{t_m V_m}(q). \end{aligned}$$

Então existe uma vizinhança O_0 da origem em \mathbb{R}^m tal que $\xi|_{O_0}$ é uma imersão e $\xi_q(O_0) \subset \mathcal{O}_q$.

Demonstração: Primeiramente, observe que, ξ_q é diferenciável e

$$\left. \frac{\partial}{\partial t_i} \right|_{t=0} \xi_q(t) = V_i(q), \quad i = 1, \dots, m.$$

Logo,

$$D_0\xi_q : \mathbb{R}^m \rightarrow \Pi_q \subset T_qM.$$

Pelo Teorema do Núcleo e da Imagem, temos que $D_0\xi_q$ é injetora. Pela Forma Local das Imersões, ξ_q é localmente injetora, isto é, existe uma vizinhança $W_0 \subset \mathbb{R}^m$ da origem tal que,

$$\xi_q : W_0 \rightarrow M \text{ é injetora.}$$

Temos que $V_i(q)$, para $i = 1, \dots, m$, são linearmente independentes em q , logo são também linearmente independentes em uma vizinhança $U_q \subset M$.

Seja $U_q \cap \xi_q(\mathbb{R}^m)$, que é aberto em $\xi_q(\mathbb{R}^m)$. Como ξ_q é contínua obtemos um subconjunto aberto

$$Z_0 = \xi_q^{-1}(U_q \cap \xi_q(\mathbb{R}^m)) \subset \mathbb{R}^m.$$

Considerando $S_0 = Z_0 \cap W_0 \subset \mathbb{R}^m$, temos uma imersão injetora:

$$\xi_q : S_0 \rightarrow M,$$

e, como toda imersão é localmente um mergulho, existe uma vizinhança suficientemente pequena O_0 tal que

$$\xi_q : O_0 \subset S_0 \rightarrow \xi_q(O_0)$$

é um homeomorfismo. Portanto $\xi_q(O_0)$ é subvariedade (*imersa*) de dimensão m de M .

Mostremos que $\xi_q(O_0) \subset \mathcal{O}_q$, $\forall q \in O_0$. De fato,

$$e^{t_i V_i} = e^{t_i (\text{Ad} P_i) f_i} = e^{t_i (P_i \circ f_i \circ P_i^{-1})} = P_i \circ e^{t_i f_i} \circ P_i^{-1} \in \mathcal{P},$$

o que implica

$$e^{t_i V_i} \in \mathcal{P}, \quad \forall i = 1, \dots, m.$$

Segue que,

$$\xi_q(t_i) = e^{t_i V_i}(q) \in \mathcal{O}_q, \quad t_i \in O_0.$$

Portanto,

$$\xi_q(t_1, \dots, t_m) = e^{t_1 V_1} \circ \dots \circ e^{t_m V_m}(q) \in \mathcal{O}_q, \quad \forall t \in O_0.$$

■

Lema 4.12 *Com as notações do lema anterior,*

$$D_t \xi_q(\mathbb{R}^m) = \Pi_{\xi_q(t)}, \quad \forall t \in O_0. \quad (4.2-5)$$

Demonstração: De fato, como $D_t \xi_q$ é injetora para todo $t \in O_0$ então,

$$\dim D_t \xi_q(\mathbb{R}^m) = m, \quad \forall t \in O_0. \quad (4.2-6)$$

Por outro lado, já vimos que $\xi_q(O_0) \subset \mathcal{O}_q$. Logo o Lema 4.9 nos dá

$$\dim \Pi_{q'} = \dim \Pi_q = m, \quad \forall q' \in \xi_q(O_0). \quad (4.2-7)$$

Por (4.2-6) e (4.2-7), segue que

$$\dim D_t \xi_q(\mathbb{R}^m) = \dim \Pi_{\xi_q(t)} = m.$$

Para concluirmos (4.2-5) basta verificar que $D_t \xi_q(\mathbb{R}^m) \subset \Pi_{\xi_q(t)}$. Para cada $i = 1, \dots, m - 1$ denote $Q_i = e^{t_{i+1}V_{i+1}} \circ \dots \circ e^{t_m V_m} \in \mathcal{P}$. Assim,

$$\begin{aligned} \frac{\partial}{\partial t_i} \xi_q(t) &= \frac{\partial}{\partial t_i} e^{t_1 V_1} \circ \dots \circ e^{t_m V_m}(q) \\ &= e^{t_1 V_1} \circ \dots \circ e^{t_i V_i} \circ V_i \circ \dots \circ e^{t_m V_m}(q) \\ &= e^{t_1 V_1} \circ \dots \circ e^{t_i V_i} \circ e^{t_{i+1} V_{i+1}} \circ \dots \circ e^{t_m V_m} \circ Q_i^{-1} \circ V_i \circ Q_i \\ &= \xi_q(t) \circ Q_i^{-1} \circ V_i \circ Q_i \\ &= \xi_q(t) \circ Q_i^{-1} \circ P_i \circ f_i \circ P_i^{-1} \circ Q_i \\ &= \xi_q(t) \circ (\text{Ad}_{Q_i^{-1}} \circ P_i) f_i \\ &= (\text{Ad}_{Q_i^{-1}} \circ P_i) f_i(\xi_q(t)) \in \Pi_{\xi_q(t)}. \end{aligned}$$

Portanto, $D_t \xi_q(\mathbb{R}^m) \subset \Pi_{\xi_q(t)}$. ■

Lema 4.13 *O conjunto*

$$\mathcal{B} = \{\xi_q(O_0) | q \in M\}$$

é uma base para uma topologia em M . A topologia gerada por esta base é mais fina que a topologia original de M .

Demonstração: Como $\xi_q(0) = q$, temos que dado $q \in M$ existe um elemento $\xi_q(O_0) \in \mathcal{B}$ tal que $q \in \xi_q(O_0)$. Para provar que \mathcal{B} é base para uma topologia, resta mostrar que dados $\xi_q(O_0)$ e $\xi_{\hat{q}}(\tilde{O}_0)$ em \mathcal{B} e \hat{q} em $\xi_q(O_0) \cap \xi_{\hat{q}}(\tilde{O}_0)$ existe um elemento $B \in \mathcal{B}$ com

$$\hat{q} \in B \subset \xi_q(O_0) \cap \xi_{\hat{q}}(\tilde{O}_0).$$

Para isto, é suficiente mostrar a seguinte afirmação.

Afirmação: dado $\hat{q} \in \xi_q(O_0)$ existe uma vizinhança \hat{O}_0 da origem em \mathbb{R}^m tal que $\xi_{\hat{q}}(\hat{O}_0) \subset \xi_q(O_0)$.

De fato, uma vez mostrado isso, como \hat{q} está em $\xi_q(O_0) \cap \xi_{\hat{q}}(\tilde{O}_0)$ podemos obter vizinhanças \hat{O}'_0 e \hat{O}''_0 da origem em \mathbb{R}^m tais que $\xi_{\hat{q}}(\hat{O}'_0) \subset \xi_q(O_0)$ e $\xi_{\hat{q}}(\hat{O}''_0) \subset \xi_{\hat{q}}(\tilde{O}_0)$. Tomando $\hat{O}_0 = \hat{O}'_0 \cap \hat{O}''_0$, obtemos o desejado.

Para mostrar a afirmação, observe que para $q' \in \xi_q(O_0)$ suficientemente próximo de \hat{q} , temos que

$$\{\widehat{V}_1(q'), \widehat{V}_2(q'), \dots, \widehat{V}_m(q')\}$$

é um conjunto linearmente independente. Assim, podemos considerar, como antes, a aplicação

$$\begin{aligned} \xi_{q'} : \quad \mathbb{R}^m &\rightarrow M \\ (t_1, t_2, \dots, t_m) &\mapsto e^{t_1 \widehat{V}_1} \circ \dots \circ e^{t_m \widehat{V}_m}(q'). \end{aligned}$$

Uma vez que

$$(D_0 \xi_{q'})(1, 0, \dots, 0) = \left. \frac{\partial \xi_{q'}}{\partial t_m} \right|_{t=0} = \widehat{V}_m(q'),$$

temos que $\widehat{V}_m(q') \in (D_0 \xi_{q'})(T_0 \mathbb{R}^m)$, o qual, por (4.2-5), coincide com $\Pi_{q'}$. Portanto, $\widehat{V}_m(q') \in \Pi_{q'}$. Agora, como $q' \in \xi_q(O_0)$, temos que $q' = \xi_q(x_0)$ para algum $x_0 \in O_0$. Usando (4.2-5) novamente e o fato de ξ_q ser imersão, concluímos que

$$\Pi_{q'} = \Pi_{\xi_q(x_0)} = (D_{x_0} \xi_q)(T_{x_0} \mathbb{R}^m) = T_{\xi_q(x_0)} \xi_q(O_0) = T_{q'} \xi_q(O_0).$$

Logo, $\widehat{V}_m \in T_{q'}(\xi_q(O_0))$. Por (4.1-3), concluímos que a curva $e^{t_m \widehat{V}_m}(\hat{q})$ está em $\xi_q(O_0)$ para $|t_m|$ suficientemente pequeno. Desta forma, podemos repetir o argumento acima para concluir que

$$e^{t_{m-1} \widehat{V}_{m-1}} \circ e^{t_m \widehat{V}_m}(\hat{q}) = e^{t_{m-1} \widehat{V}_{m-1}}(e^{t_m \widehat{V}_m}(\hat{q}))$$

está em $\xi_q(O_0)$ para $|t_{m-1}|, |t_m|$ suficientemente pequenos. Continuando esse processo, concluímos que

$$e^{t_1 \widehat{V}_1} \circ \dots \circ e^{t_m \widehat{V}_m}(\hat{q}) \subset \xi_q(O_0)$$

para (t_1, \dots, t_m) suficientemente próximo de $0 \in \mathbb{R}^n$.

Vamos verificar que a topologia gerada por \mathcal{B} é mais fina que a topologia original de M . Como ξ_q é diferenciável, então ξ_q é contínua na origem quando se considera em M a topologia original. Desta forma, dado um aberto $V \subset M$ e um ponto $q = \xi_q(0)$ em V , temos que existe uma vizinhança $O'_0 \subset O_0$ da origem em \mathbb{R}^m tal que $\xi_q(O'_0) \subset V$. ■

Denotaremos por $M^{\mathcal{F}}$ o conjunto M munido da topologia gerada pela base \mathcal{B} descrita no lema anterior. Mostraremos no próximo lema que as órbitas \mathcal{O}_q , $q \in M$ são exatamente as componentes conexas de $M^{\mathcal{F}}$.

Lema 4.14 *Para cada $q \in M$, a órbita \mathcal{O}_q é aberta, fechada e conexa em $M^{\mathcal{F}}$.*

Demonstração: Para mostrar que \mathcal{O}_q é conexa, mostraremos que a curva $t \mapsto e^{tf}q$, $f \in \mathcal{F}$ é contínua em cada $t_0 \in \mathbb{R}$, quando vista como uma aplicação de \mathbb{R} em $M^{\mathcal{F}}$. Se $f(q) = 0$ então $e^{tf}q = 0$ e não há nada para mostrar. Suponha então que $f(q) \neq 0$ e tome $V_2, \dots, V_m \in (\text{Ad}\mathcal{P})\mathcal{F}$ tais que

$$\text{span}\{f(q), V_2(q), \dots, V_m(q)\} = \Pi_q.$$

Considere, como nos lemas anteriores, a aplicação $\xi_q : \mathbb{R}^m \rightarrow M^{\mathcal{F}}$ definida por

$$\xi_q(t_1, \dots, t_m) = e^{t_1 f} \circ e^{t_2 V_2} \circ \dots \circ e^{t_m V_m} q.$$

Se a correspondente vizinhança O_0 contém algum ponto com a primeira coordenada igual a t_0 , então basta observar que

$$e^{t_0 f} q = e^{t_0 f} \circ e^{0V_2} \circ \dots \circ e^{0V_m} q$$

para concluir que $e^{tf}q \in \xi_q(O_0)$ para todo t em alguma vizinhança U de t_0 . Se O_0 não contém pontos com a primeira coordenada igual a t_0 , tome $a \in \mathbb{R}$ suficientemente pequeno de tal modo que O_0 contenha algum ponto com primeira coordenada igual a

at_0 . Considere o campo $g = \frac{1}{a}f$. Como

$$e^{t_0 f} = e^{t_0 a \frac{1}{a} f} = e^{at_0 g},$$

basta aplicar o argumento acima trocando f por g . Desta forma, as órbitas são conexas por caminhos e, conseqüentemente, conexas.

Vamos mostrar que as órbitas são abertas em $M^{\mathcal{F}}$. Tome $q, q_0 \in M$ com $q \in \mathcal{O}_{q_0}$. Como as órbitas são classes de equivalência, temos que $\mathcal{O}_q = \mathcal{O}_{q_0}$. Pelo Lema 4.11, $\xi_q(O_0) \subset \mathcal{O}_q$. Portanto, o aberto básico $\xi_q(O_0)$ é tal que

$$q \in \xi_q(O_0) \subset \mathcal{O}_{q_0},$$

mostrando que \mathcal{O}_{q_0} é aberta. Para concluir que cada órbita é fechada, basta observar que o complementar de uma órbita é uma união de órbitas, as quais são abertas pelo que acabamos de provar. ■

Finalmente, vamos provar o Teorema da Órbita.

Demonstração do Teorema da Órbita:

1. Para munir \mathcal{O}_{q_0} de uma estrutura de variedade diferenciável, tomamos como vizinhanças coordenadas os conjuntos $\xi_q(O_0)$, $q \in \mathcal{O}_{q_0}$, e as aplicações ξ_q^{-1} como sistemas de coordenadas locais. Como ξ_q são imersões em M , então cada órbita é uma subvariedade mergulhada de M . Além disso, como \mathcal{O}_{q_0} é conexa em $M^{\mathcal{F}}$ e essa topologia é mais fina que a topologia de M , segue que \mathcal{O}_{q_0} é conexa em M .
2. Como $T_q \mathcal{O}_{q_0} = D_0 \xi_q(\mathbb{R}^m)$, o resultado segue do Lema 4.12. ■

O Teorema da Órbita nos fornece importantes conseqüências. Primeiramente, lembremos que $(\text{Vec}(M), [,])$ é uma álgebra de Lie com a operação colchete $[,]$, como vimos em 1.4.2.

Sejam $\mathcal{F} \subset \text{Vec}(M)$ uma família de campos de vetores, \mathcal{O}_{q_0} a órbita de \mathcal{F} através do ponto $q_0 \in M$. Dado um ponto $q \in \mathcal{O}_{q_0}$ e um campo $f \in \mathcal{F}$, temos, por definição, que

$$e^{t f}(q) \in \mathcal{O}_{q_0}, \quad t \in \mathbb{R}.$$

Por (4.1-3), obtemos

$$\left. \frac{d}{dt} \right|_{t=0} e^{tf}(q) = f(q) \in T_q \mathcal{O}_{q_0}.$$

Continuando o raciocínio, dados $f_1, f_2 \in \mathcal{F}$ campos, temos que

$$e^{-tf_2} \circ e^{-tf_1} \circ e^{tf_2} \circ e^{tf_1}(q) \in \mathcal{O}_{q_0},$$

então

$$\left. \frac{d}{dt} \right|_{t=0} e^{-tf_2} \circ e^{-tf_1} \circ e^{tf_2} \circ e^{tf_1}(q) = [f_1, f_2](q) \in T_q \mathcal{O}_{q_0}.$$

Novamente, sejam agora $f_1, f_2, f_3 \in \mathcal{F}$ temos que

$$e^{-t[f_2, f_3]} \circ e^{-tf_1} \circ e^{t[f_2, f_3]} \circ e^{tf_1}(q) \in \mathcal{O}_{q_0},$$

e assim,

$$\left. \frac{d}{dt} \right|_{t=0} e^{-t[f_2, f_3]} \circ e^{-tf_1} \circ e^{t[f_2, f_3]} \circ e^{tf_1}(q) = [f_1, [f_2, f_3]](q) \in T_q \mathcal{O}_{q_0}.$$

De acordo com (1.4-25), podemos escrever

$$(\text{ad } f)^0(q) = f(q) \in T_q \mathcal{O}_{q_0},$$

$$(\text{ad } f_1)f_2(q) = [f_1, f_2](q) \in T_q \mathcal{O}_{q_0},$$

e

$$(\text{ad } f_1) \circ (\text{ad } f_2)f_3(q) = [f_1, [f_2, f_3]](q) \in T_q \mathcal{O}_{q_0}.$$

Neste contexto, apresentamos a

Definição 4.15 *Sejam \mathcal{O}_{q_0} a órbita da família $\mathcal{F} \subset \text{Vec}(M)$ de campos de vetores, e $q \in M$ qualquer. O subespaço vetorial denotado por*

$$\text{Lie}\mathcal{F} = \text{span}\{[f_1, [\dots [f_{k-1}, f_k] \dots]]; f_i \in \mathcal{F}, k \in \mathbb{N}\} \subset \text{Vec}(M)$$

é a Álgebra de Lie de campos de vetores gerada pela família \mathcal{F} .

Denotaremos

$$\text{Lie}_q \mathcal{F} = \text{span}\{V(q) \mid V \in \text{Lie}\mathcal{F}\} \subset T_q M.$$

Apesar da notação utilizada, não iremos introduzir em $\text{Lie}_q\mathcal{F}$ uma estrutura de álgebra de Lie. Veremos tal conjunto apenas como um subespaço vetorial de T_qM .

A álgebra de Lie de campos de vetores gerada pela família \mathcal{F} é em geral um espaço de dimensão infinita de $\text{Vec}(M)$. Porém, o subespaço tangente $\text{Lie}_q\mathcal{F}$ têm dimensão limitada pela dimensão da órbita \mathcal{O}_{q_0} , ou seja

$$\dim \text{Lie}_q\mathcal{F} \leq \dim \mathcal{O}_{q_0}.$$

Exemplo 4.16 Sejam $M = \mathbb{R}^3$, e $\mathcal{F} = \{X, Y\}$, onde X e Y são campos de lineares dados respectivamente pelas matrizes

$$A = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad e \quad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

O colchete de Lie de A e B é dado por $[A, B] = AB - BA$, fazendo os cálculos, obtemos

$$[A, B] = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

Neste caso,

$$\text{Lie}(\mathcal{F}) = \text{span} \left\{ \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \right\}$$

é formado por matrizes antissimétricas 3×3 .

Corolário 4.17

$$\text{Lie}_q\mathcal{F} \subset T_q\mathcal{O}_{q_0} \quad \forall q \in \mathcal{O}_{q_0}.$$

Demonstração: Segue da construção que fizemos anteriormente. ■

O Exemplo abaixo mostra que a inclusão acima pode ser estrita.

Exemplo 4.18 Considere $M = \mathbb{R}^2$ e $\mathcal{F} = \left\{ \frac{\partial}{\partial x}, a(x) \frac{\partial}{\partial y} \right\}$ para todo $z = (x, y) \in \mathbb{R}^2$, onde

$a \in C^\infty(\mathbb{R})$, $a \neq 0$, tem suporte compacto. A família

$$\mathcal{F} \cup -\mathcal{F} = \left\{ \pm \frac{\partial}{\partial x}, \pm a(x) \frac{\partial}{\partial y} \right\}$$

é completamente controlável em \mathbb{R}^2 . De fato, dados dois pontos quaisquer $z_0 = (x_0, y_0)$, $z_1 = (x_1, y_1)$ em \mathbb{R}^2 podemos a partir de z_0 alcançar z_1 :

Primeiro partimos de z_0 pelo campo $\pm \frac{\partial}{\partial x}$ até um ponto $z_2 = (x_2, y_0)$ tal que $a(x_2) \neq 0$, em seguida partimos do ponto $z_2 = (x_2, y_0)$ pelo campo $\pm a(x_2) \frac{\partial}{\partial y}$ até o ponto $z_3 = (x_2, y_1)$, e por fim, de $z_3 = (x_2, y_1)$ até $z_1 = (x_1, y_1)$ através do campo $\pm \frac{\partial}{\partial x}$. Portanto $\mathcal{O}_{(x,y)} = \mathbb{R}^2$. O Teorema da Órbita nos dá

$$T_{(x,y)}\mathcal{O}_{(x_0,y_0)} = \text{span}\{(\text{Ad}\mathcal{P})\mathcal{F}(x,y)\} = \mathbb{R}^2.$$

Por outro lado,

$$\dim \text{Lie}_{(x,y)}\mathcal{F} = \begin{cases} 1, & x \notin \text{supp}(a) \\ 2, & a(x) \neq 0, \end{cases}$$

ou seja,

$$\dim T_{(x,y)}\mathcal{O}_{(x_0,y_0)} \neq \dim \text{Lie}_{(x,y)}\mathcal{F} = 1, \text{ para todo } x \notin \text{supp}(a).$$

Neste caso,

$$\text{Lie}_{(x,y)}\mathcal{F} \subset T_{(x,y)}\mathcal{O}_{(x_0,y_0)}$$

é uma inclusão estrita.

Definição 4.19 Dizemos que a família $\mathcal{F} \subset \text{Vec}(M)$ é **completamente não-holonômica** se

$$\text{Lie}_q\mathcal{F} = T_qM \quad \forall q \in M.$$

Uma consequência do Teorema da Órbita é o seguinte teorema, que é largamente utilizada em teoria de controle.

Teorema 4.20 (Rashevsky-Chow) Sejam M uma variedade conexa e $\mathcal{F} \subset \text{Vec}(M)$ uma família de campos de vetores. Se \mathcal{F} é completamente não-holonômica, então

$$\mathcal{O}_{q_0} = M \quad q_0 \in M.$$

Demonstração: Por hipótese temos

$$\text{Lie}_q \mathcal{F} = T_q M \quad \forall q \in M,$$

e, pelo Corolário 4.17 temos

$$\text{Lie}_q \mathcal{F} \subset T_q \mathcal{O}_{q_0}.$$

Logo

$$T_q M = T_q \mathcal{O}_{q_0} \quad \forall q \in \mathcal{O}_{q_0}.$$

Assim, temos que

$$\dim \mathcal{O}_{q_0} = \dim M.$$

Como \mathcal{O}_{q_0} é uma subvariedade imersa, existe uma imersão injetora $\Phi : N \rightarrow M$ de uma variedade N de dimensão n em M , tal que $\Phi(N) = \mathcal{O}_{q_0}$. Assim, para qualquer $q \in \mathcal{O}_{q_0}$ temos um isomorfismo

$$D_p \Phi : T_p N \rightarrow T_q M, \quad \Phi(p) = q.$$

Consequentemente, Φ é um difeomorfismo local. Isto mostra que \mathcal{O}_{q_0} é um subconjunto aberto de M .

Considere a relação de equivalência definida na Observação 4.8, tal que

$$\frac{M}{\sim} = \{\mathcal{O}_q \mid q \in M\}.$$

Como M é conexa, e todas as órbitas são abertas e disjuntas, concluímos que existe uma única órbita. Portanto

$$M = \mathcal{O}_{q_0}.$$

■

Corolário 4.21 *Sejam M uma variedade conexa e $\mathcal{F} \subset \text{Vec}(M)$ uma família **simétrica** completamente não-holonômica. Então*

$$\mathcal{A}_{q_0} = M.$$

Demonstração: Como \mathcal{F} é não-holonômica, temos pelo teorema acima que

$$\mathcal{O}_{q_0} = M,$$

e por \mathcal{F} ser simétrica temos, como vimos na Observação 4.8, que

$$\mathcal{A}_{q_0} = \mathcal{O}_{q_0}.$$

Logo

$$\mathcal{A}_{q_0} = M.$$

■

A igualdade no Corolário 4.17 é obtida no caso em que M e \mathcal{F} são analíticas, como veremos na próxima seção.

4.3 Caso analítico

Como vimos, o espaço $\text{Vec}(M)$ munido com o colchete de Lie é uma álgebra de Lie de campos de vetores. Além disso, este espaço é um $\mathcal{C}^\infty(M)$ -módulo. Para ver isso, defina o produto de uma função $a \in \mathcal{C}^\infty(M)$ por um campo $V \in \text{Vec}(M)$ como o campo aV dado por

$$(aV)(q) = a(q)V(q).$$

É imediato verificar que este produto satisfaz as seguintes propriedades:

$$(i) (a_1 + a_2)V = a_1V + a_2V,$$

$$(ii) a(V_1 + V_2) = aV_1 + aV_2,$$

$$(iii) a_1(a_2, V) = a_1, a_2(V),$$

$$(iv) 1V = V,$$

para todo $a, a_1, a_2 \in \mathcal{C}^\infty(M)$, $V, V_1, V_2 \in \text{Vec}(M)$. Assim, $\text{Vec}(M)$ é um $\mathcal{C}^\infty(M)$ -módulo. Uma submódulo $\mathcal{V} \subset \text{Vec}(M)$ é um subgrupo de $\text{Vec}(M)$ fechado com respeito a multiplicação de elementos de $\mathcal{C}^\infty(M)$. Se o campo for considerado como uma derivação,

definimos o produto de uma função a por um campo V como sendo a derivação

$$aV : \mathcal{C}^\infty(M) \rightarrow \mathcal{C}^\infty(M)$$

pondo

$$(aV)b = a \cdot (Vb), \quad b \in \mathcal{C}^\infty(M). \quad (4.3-8)$$

Precisaremos do seguinte resultado:

Proposição 4.22 Para todo $X, Y \in \text{Vec}(M)$, $a \in \mathcal{C}^\infty(M)$ e $P \in \text{Diff}(M)$, vale:

$$(i) (\text{ad}X)(aY) = (Xa)Y + a(\text{ad}X)Y$$

$$(ii) (\text{Ad}P)(aX) = (Pa)(\text{Ad}P)X.$$

Demonstração: O item (i) segue da Proposição 5.3, página 27, de [Manfredo]. Para provar (ii), observe que, para qualquer $b \in \mathcal{C}^\infty(M)$, temos:

$$\begin{aligned} ((\text{Ad})(aX))(b) &= (P \circ aX \circ P^{-1})(b) \\ &= P(aX(P^{-1}b)) \\ &= P(a \cdot X(P^{-1}b)) \\ &= (Pa) \cdot P(X(P^{-1}b)) \quad (\text{pois } P \text{ é automorfismo}) \\ &= (Pa)(\text{Ad}PX)(b) \quad (\text{por (4.3-8)}) \end{aligned}$$

■

Definição 4.23 Dizemos que um submódulo $\mathcal{V} \subset \text{Vec}(M)$ é **finitamente gerado** sobre $\mathcal{C}^\infty(M)$ se, existe uma base finita global de campos de vetores, isto é, se existirem $V_1, \dots, V_k \in \mathcal{V}$ tais que

$$\mathcal{V} = \left\{ \sum_{i=1}^k a_i V_i \mid a_i \in \mathcal{C}^\infty(M) \right\}.$$

Definição 4.24 Dizemos que um submódulo $\mathcal{V} \subset \text{Vec}(M)$ é **localmente finitamente gerado** sobre $\mathcal{C}^\infty(M)$ se, para cada ponto $q \in M$ existe uma vizinhança aberta $O \subset M$ tal que $\mathcal{V}|_O$ é finitamente gerado sobre $\mathcal{C}^\infty(O)$, ou seja

$$\mathcal{V}|_O = \left\{ \sum_{i=1}^k a_i V_i \mid a_i \in \mathcal{C}^\infty(O), V_i \in \mathcal{V}|_O \right\}.$$

Lema 4.25 *Sejam $\mathcal{V} \subset \text{Vec}(M)$ um submódulo finitamente gerado sobre $\mathcal{C}^\infty(M)$ e $X \in \text{Vec}(M)$. Se*

$$[X, \mathcal{V}] = (\text{ad}X) \mathcal{V} = \{(\text{ad}X) V \mid V \in \mathcal{V}\} \subset \mathcal{V},$$

então $(\text{Ad } e^{tX}) \mathcal{V} = \mathcal{V}$, $t \in \mathbb{R}$.

Demonstração: Por hipótese, existem $V_1, \dots, V_k \in \text{Vec}(M)$ tais que

$$\mathcal{V} = \left\{ \sum_{i=1}^k a_i V_i \mid a_i \in \mathcal{C}^\infty(M) \right\}.$$

Como $[X, \mathcal{V}] \subset \mathcal{V}$, temos

$$[X, V_i] = \sum_{j=1}^k a_{ij} V_j. \quad (4.3-9)$$

Queremos mostrar que

$$(\text{Ad } e^{tX}) V_i = \sum_{j=1}^k b_{ij} V_j$$

para cada $i = 1, \dots, k$. Primeiramente, vale notar que $\mathcal{V} \subset (\text{Ad } e^{tX}) \mathcal{V}$ pois,

$$V_i = (\text{Ad } e^{0 \cdot X}) V_i \in (\text{Ad } e^{tX}) \mathcal{V}$$

para cada $i = 1, \dots, k$. Mostraremos agora que

$$(\text{Ad } e^{tX}) \mathcal{V} \subset \mathcal{V}.$$

Denotemos

$$V_i(t) = (\text{Ad } e^{tX}) V_i, \quad t \in \mathbb{R}, \quad (4.3-10)$$

obtendo a E.D.O:

$$\begin{aligned}
 \dot{V}_i(t) &= \frac{d}{dt} (\text{Ad } e^{tX}) V_i = (\text{Ad } e^{tX}) [X, V_i] \quad (\text{pela Proposição 1.47}) \\
 &= (\text{Ad } e^{tX}) \sum_{j=1}^k a_{ij} V_j \quad (\text{pela igualdade 4.3-9}) \\
 &= \sum_{j=1}^k (\text{Ad } e^{tX}) (a_{ij} V_j) \\
 &= \sum_{j=1}^k (e^{tX} a_{ij}) (\text{Ad } e^{tX}) V_j \quad (\text{pelo item (ii) da Proposição 4.22}) \\
 &= \sum_{j=1}^k (e^{tX} a_{ij}) V_j(t) \quad (\text{pela igualdade (4.3-10)}).
 \end{aligned}$$

Em resumo, obtemos para cada $i = 1, \dots, k$ uma E.D.O

$$\dot{V}_i(t) = \sum_{j=1}^k (e^{tX} a_{ij}) V_j(t). \quad (4.3-11)$$

Fixamos um ponto $q \in M$ e definimos a matriz $A(t)$ de ordem $k \times k$:

$$A(t) = (a_{ij}(t)) \quad \text{onde} \quad (a_{ij}(t)) = (e^{tX} a_{ij}(q)) \quad i, j = 1, \dots, k.$$

Obtemos assim, o sistema de E.D.O's lineares homogêneas

$$\dot{V}_i(t) = \sum_{j=1}^k a_{ij}(t) V_j(t). \quad (4.3-12)$$

Para resolvermos, escrevemos o sistema acima na forma de matrizes:

$$\dot{H} = A(t)H, \quad H(0) = \text{Id},$$

e encontramos a matriz fundamental H . Como $A(t)$ depende diferenciavelmente de q , então H depende diferenciavelmente de q . Obtemos,

$$H(t) = (b_{ij}(t)) \quad \text{com} \quad (b_{ij}(t)) \in C^\infty(M) \quad i, j = 1, \dots, k, \quad t \in \mathbb{R}.$$

Logo, a solução do sistema (4.3-12) é:

$$V_i(t) = \sum_{j=1}^k b_{ij}(t)V_j(0), \quad i = 1, \dots, k.$$

Por (4.3-10), temos $V_j(0) = V_j$. Portanto,

$$(\text{Ad } e^{tX}) V_i = V_i(t) = \sum_{j=1}^k b_{ij}(t)V_j \in \mathcal{V}, \quad i = 1, \dots, k,$$

como queríamos demonstrar. ■

Teorema 4.26 *Seja $\mathcal{F} \subset \text{Vec}(M)$ uma família. Suponha que o submódulo $\text{Lie}\mathcal{F}$ é localmente finitamente gerado sobre $\mathcal{C}^\infty(M)$. Então,*

$$T_q\mathcal{O}_{q_0} = \text{Lie}_q\mathcal{F}, \quad q \in \mathcal{O}_{q_0}$$

para qualquer órbita \mathcal{O}_{q_0} , $q_0 \in M$, da família \mathcal{F} .

Demonstração: Pelo Teorema da Órbita e o Corolário 4.17, temos que

$$\text{Lie}_q\mathcal{F} \subset T_q\mathcal{O}_{q_0} = \text{span}\{(\text{Ad}\mathcal{P})\mathcal{F}(q)\}, \quad q \in \mathcal{O}_{q_0}.$$

Assim, resta mostrar que

$$\text{span}\{(\text{Ad}\mathcal{P})\mathcal{F}(q)\} \subset \text{Lie}_q\mathcal{F} \quad q \in \mathcal{O}_{q_0}.$$

Por hipótese, para cada $q \in M$ existe uma vizinhança aberta $O \subset M$ contendo q , tal que o submódulo $\text{Lie}\mathcal{F}$ é finitamente gerado sobre $\mathcal{C}^\infty(O)$. Isto é,

$$\text{Lie}\mathcal{F} = \left\{ \sum_{i=1}^k a_i f_i \mid a_i \in \mathcal{C}^\infty(O), f_i \in \text{Lie}\mathcal{F}|_O \right\}.$$

Temos, devido a Definição 4.15, que

$$(\text{ad } f)\text{Lie}\mathcal{F} \subset \text{Lie}\mathcal{F}, \quad \forall f \in \text{Lie}\mathcal{F}|_O.$$

Como a afirmação acima é válida para todo campo $f \in \text{Lie}\mathcal{F}|_O$ temos, pelo Lema 4.25, que

$$(\text{Ad } e^{tf})\text{Lie}\mathcal{F} \subset \text{Lie}\mathcal{F}, \quad \forall e^{tf} \in \mathcal{P}.$$

Obtemos assim,

$$(\text{Ad}\mathcal{P})\mathcal{F} \subset (\text{Ad}\mathcal{P})\text{Lie}\mathcal{F} \subset \text{Lie}\mathcal{F}.$$

Portanto,

$$T_q\mathcal{O}_{q_0} \subset \text{Lie}_q\mathcal{F}, \quad q \in \mathcal{O}_{q_0}$$

e a igualdade segue. ■

Corolário 4.27 *Se M e \mathcal{F} são analíticas, então $T_q\mathcal{O}_{q_0} = \text{Lie}_q\mathcal{F}$, $q \in \mathcal{O}_{q_0}$.*

4.4 Conjuntos atingíveis

Nesta seção vamos estudar algumas propriedades gerais dos conjuntos atingíveis a partir de uma família de campos de vetores $\mathcal{F} \subset \text{Vec}(M)$ completamente não-holonômica, isto é,

$$\text{Lie}_q\mathcal{F} = T_qM \quad q \in M.$$

Teorema 4.28 (Krener) *Se $\mathcal{F} \subset \text{Vec}(M)$ é uma família completamente não-holonômica, então*

$$\mathcal{A}_{q_0} \subset \overline{\text{int}(\mathcal{A}_{q_0})}$$

para qualquer $q_0 \in M$.

Demonstração: Fixe um ponto $q_0 \in M$ e seja \mathcal{A}_{q_0} o conjunto atingível a partir de q_0 para tempo qualquer da família \mathcal{F} . Seja $q_1 \in \mathcal{A}_{q_0}$ um ponto qualquer. Vamos mostrar que toda vizinhança de q_1 intersecta o interior de \mathcal{A}_{q_0} , isto é,

$$q_1 \in \overline{\text{int}(\mathcal{A}_{q_0})}. \tag{4.4-13}$$

(1) Existe um campo de vetores $f_1 \in \mathcal{F}$ tal que $f_1(q_1) \neq 0$. Caso contrário, teríamos que $f_1(q_1) = 0$ para todo campo $f_1 \in \mathcal{F}$, logo $\dim \text{Lie}_{q_1}\mathcal{F} = 0$.

Considere a curva em M dada por:

$$s_1 \mapsto e^{s_1 f_1}(q_1), \quad 0 < s_1 < \varepsilon_1, \quad (4.4-14)$$

cuja imagem é uma subvariedade de dimensão 1 de M para $\varepsilon_1 > 0$ suficientemente pequeno. Se $\dim M = 1$, então $e^{s_1 f_1}(q_1) \in \text{int}(\mathcal{A}_{q_0})$ para $s_1 > 0$ suficientemente pequeno. Bem entendido, para qualquer aberto $U \subset M$ contendo o ponto q_1 , podemos obter um número $s_1 > 0$ suficientemente pequeno tal que

$$e^{s_1 f_1}(q_1) \in \text{int}(\mathcal{A}_{q_0})$$

e conseqüentemente,

$$U \cap \text{int}(\mathcal{A}_{q_0}) \neq \emptyset.$$

Portanto, (4.4-13) segue.

(2) Suponha que $\dim M > 1$. Podemos encontrar um ponto q_2 sobre a curva (4.4-14) *arbitrariamente próximo* de q_1 , isto é,

$$q_2 = e^{t_1^1 f_1}(q_1), \quad t_1^1 > 0 \quad \text{suficientemente pequeno,}$$

e um campo de vetores $f_2 \in \mathcal{F}$ tal que $f_2(q_1)$ não é tangente à curva (4.4-14). Repare que um tal campo f_2 existe pois, caso contrário teríamos $\dim \text{Lie}_{q_1} \mathcal{F} = \dim T_{q_1} M = 1$. Considere a aplicação

$$(s_1, s_2) \mapsto e^{s_1 f_1} \circ e^{s_2 f_2}(q_1) \quad (4.4-15)$$

onde

$$t_1^1 < s_1 < t_1^1 + \varepsilon_2 \quad \text{e} \quad 0 < s_2 < \varepsilon_2.$$

Esta aplicação é uma imersão em uma vizinhança suficientemente pequena da origem em \mathbb{R}^2 , logo sua imagem é uma subvariedade de dimensão 2 de M . Se $\dim M = 2$ então

$$e^{s_1 f_1} \circ e^{s_2 f_2}(q_1) \subset \text{int}(\mathcal{A}_{q_0})$$

para $\varepsilon_2 > 0$ suficientemente pequeno. Como o ponto q_2 é tomado *arbitrariamente próximo* de q_1 então a inclusão (4.4-13) segue.

(3) Suponha que $\dim M > 2$. Repetimos o processo anterior. Existem $t_2^1, t_2^2 > 0$ e um

campo de vetores $f_3 \in \mathcal{F}$ tal que f_3 não é tangente à superfície (4.4-15) no ponto

$$q_3 = e^{t_2^1 f_1} \circ e^{t_2^2 f_2}(q_1).$$

Caso contrário, teremos $\dim M = 2$. Desta maneira, a imagem da aplicação

$$(s_1, s_2, s_3) \mapsto e^{s_1 f_1} \circ e^{s_2 f_2} \circ e^{s_3 f_3}(q_1), \quad (4.4-16)$$

onde

$$t_2^i < s_i < t_2^i + \varepsilon_3 \quad i = 1, 2 \quad \text{e} \quad 0 < s_3 < \varepsilon_3,$$

é uma subvariedade de dimensão 3 de M , quando restringimos seu domínio a uma vizinhança suficientemente pequena da origem e \mathbb{R}^3 . Se $\dim M = 3$ então a inclusão (4.4-13) segue.

(4) Para $\dim M = n$, obtemos um ponto

$$q_{n-1} = e^{t_{n-1}^1 f_1} \circ \dots \circ e^{t_{n-1}^{n-1} f_{n-1}}(q_1)$$

com

$$(t_{n-1}^1, \dots, t_{n-1}^{n-1}) \in \mathbb{R}^{n-1}, \quad t_{n-1}^i > 0,$$

e campos de vetores $f_1, \dots, f_n \in \mathcal{F}$ tais que a aplicação

$$(s_1, \dots, s_n) \mapsto e^{s_1 f_1} \circ \dots \circ e^{s_{n-1} f_{n-1}} \circ e^{s_n f_n}(q_1), \quad (4.4-17)$$

onde

$$t_{n-1}^i < s_i < t_{n-1}^i + \varepsilon_n \quad i = 1, 2, \dots, n-1 \quad \text{e} \quad 0 < s_n < \varepsilon_n,$$

é uma imersão. A imagem desta imersão é uma subvariedade de dimensão n de M , logo um subconjunto aberto de M . Como este conjunto aberto está contido em \mathcal{A}_{q_0} e pode ser tomado arbitrariamente próximo de q_1 , obtemos que

$$q_1 \in \overline{\text{int}(\mathcal{A}_{q_0})}.$$

■

Corolário 4.29 *Seja $\mathcal{F} \subset \text{Vec}(M)$ uma família completamente não-holonômica. Se $\overline{\mathcal{A}_{q_0}(\mathcal{F})} = M$ para algum q_0 , então $\mathcal{A}_{q_0}(\mathcal{F}) = M$.*

Demonstração: Seja $q \in M$ um ponto arbitrário e consideremos a família

$$-\mathcal{F} = \{-f \mid f \in \mathcal{F}\}$$

não-holonômica. Pelo Teorema 4.28 temos

$$\mathcal{A}_q(-\mathcal{F}) \subset \overline{\text{int}(\mathcal{A}_q(-\mathcal{F}))}, \quad q \in M.$$

Seja $q' \in \text{int}(\mathcal{A}_q(-\mathcal{F}))$ e $U_{q'}$ o aberto contendo q' e contido em $\mathcal{A}_q(-\mathcal{F})$. Por hipótese, $\mathcal{A}_{q_0}(\mathcal{F})$ é denso em M , logo $U_{q'} \cap \mathcal{A}_{q_0}(\mathcal{F}) \neq \emptyset$. Ou seja,

$$\mathcal{A}_{q_0}(\mathcal{F}) \cap \mathcal{A}_q(-\mathcal{F}) \neq \emptyset.$$

Segue que existe q'' representado em ambos conjuntos por:

$$q'' = e^{t_1 f_1} \circ \dots \circ e^{t_k f_k}(q_0) = e^{-s_1 g_1} \circ \dots \circ e^{-s_l g_l}(q),$$

assim

$$q = e^{t_1 f_1} \circ \dots \circ e^{t_k f_k} \circ e^{s_l g_l} \circ \dots \circ e^{s_1 g_1}(q_0).$$

Portanto, $q \in \mathcal{A}_{q_0}(\mathcal{F})$, mostrando que $M \subset \mathcal{A}_{q_0}(\mathcal{F})$. A inclusão contrária é natural. ■

CAPÍTULO 5

STATE E FEEDBACK EQUIVALÊNCIA DE SISTEMAS DE CONTROLE

Neste capítulo discutiremos dois tipos de equivalência entre sistemas de controle: state-equivalência e feedback-equivalência. De modo geral, dois sistemas são state-equivalentes se existe uma transformação no espaço estado que “transforma” um sistema no outro, sem alterar os controles. Esse tipo de equivalência já foi tratado nos Teoremas 3.2 e 3.5 do Capítulo 3, quando exibimos condições para que um sistema não linear seja equivalente a um sistema linear controlável. Já a feedback-equivalência “transforma” os sistemas alterando os controles. Os detalhes serão vistos na seção 5.2.

5.1 State equivalência

Sejam M, N variedades e f_U, g_U famílias de campos de vetores sobre M, N respectivamente, parametrizadas por um mesmo conjunto de parâmetros de controle U :

$$\begin{aligned}f_U &= \{f_u \mid u \in U\} \subset \text{Vec}(M), \\g_U &= \{g_u \mid u \in U\} \subset \text{Vec}(N).\end{aligned}$$

Suponha que essas famílias sejam completamente não-holonômicas, isto é, para quaisquer $x_0 \in M$ e $y_0 \in N$ temos,

$$\text{Lie}_{x_0} f_U = T_{x_0} M, \quad \text{Lie}_{y_0} g_U = T_{y_0} N.$$

Definição 5.1 Dizemos que as famílias f_U e g_U são **localmente state-equivalente** em $x_0 \in M$ e $y_0 \in N$ se existe um difeomorfismo local

$$\Phi : O_{x_0} \subset M \rightarrow O_{y_0} \subset N$$

com $\Phi(x_0) = y_0$, tal que

$$\Phi_*(f_u) = g_u, \quad \forall u \in U.$$

Notação: $(f_U, x_0) \simeq (g_U, y_0)$ e diremos simplesmente que são **localmente S-equivalente**.

Vamos procurar condições necessárias para obtermos uma equivalência local dos sistemas f_U e g_U .

Suponha que $(f_U, x_0) \simeq (g_U, y_0)$. Lembramos que o colchete de Lie é invariante por isomorfismos, ou seja,

$$\Phi_*[f_{u_1}, f_{u_2}] = [\Phi_*f_{u_1}, \Phi_*f_{u_2}] = [g_{u_1}, g_{u_2}], \quad u_1, u_2 \in U.$$

Definimos os vetores tangentes para um conjunto de controles $u_1, \dots, u_k \in U$:

$$\xi_{u_1, \dots, u_k} = \left[f_{u_1}, \left[f_{u_2}, \left[\dots [f_{u_{k-1}}, f_{u_k}] \dots \right] \right] \right] (x_0) \in T_{x_0} M \quad (5.1-1)$$

$$\eta_{u_1, \dots, u_k} = \left[g_{u_1}, \left[g_{u_2}, \left[\dots [g_{u_{k-1}}, g_{u_k}] \dots \right] \right] \right] (y_0) \in T_{y_0} N. \quad (5.1-2)$$

Segue que,

$$\begin{aligned} \Phi_{*x_0}(\xi_{u_1, \dots, u_k}) &= \Phi_* \left[f_{u_1}, \left[f_{u_2}, \left[\dots [f_{u_{k-1}}, f_{u_k}] \dots \right] \right] \right] (x_0) \\ &= \left[g_{u_1}, \left[g_{u_2}, \left[\dots [g_{u_{k-1}}, g_{u_k}] \dots \right] \right] \right] (y_0) \\ &= \eta_{u_1, \dots, u_k}. \end{aligned}$$

Desta forma, obtemos que se $(f_U, x_0) \simeq (g_U, y_0)$ então existe um isomorfismo

$$\Phi_{*x_0} : T_{x_0}M \rightarrow T_{y_0}N$$

tal que

$$\Phi_{*x_0}(\xi_{u_1, \dots, u_k}) = \eta_{u_1, \dots, u_k}, \quad u_i \in U, \quad k \in \mathbb{N}.$$

A recíproca é verdadeira no caso analítico. Chamaremos o par (M, f_U) de *analítico* quando a variedade M e a família f_U são analíticas.

Teorema 5.2 *Sejam (M, f_U) e (N, g_U) analíticos tais que,*

$$\text{Lie}_{x_0} f_U = T_{x_0}M, \quad \text{Lie}_{y_0} g_U = T_{y_0}N,$$

para quaisquer $x_0 \in M$ e $y_0 \in N$. Então $(f_U, x_0) \simeq (g_U, y_0)$ se, e somente se, existe um isomorfismo

$$\Phi_{*x_0} : T_{x_0}M \rightarrow T_{y_0}N$$

tal que

$$\Phi_{*x_0}(\xi_{u_1, \dots, u_k}) = \eta_{u_1, \dots, u_k}, \quad u_i \in U, \quad k \in \mathbb{N}. \quad (5.1-3)$$

Demonstração: (\Leftarrow) Considere a variedade produto definida por M e N :

$$M \times N = \{(x, y) \mid x \in M \text{ e } y \in N\}.$$

Dados $f \in \text{Vec}(M)$, $g \in \text{Vec}(N)$ definimos o campo de vetores produto $f \times g \in \text{Vec}(M \times N)$ como a derivação:

$$\widehat{f \times g} : \mathcal{C}^\infty(M \times N) \rightarrow \mathcal{C}^\infty(M \times N)$$

por:

$$(\widehat{f \times g}) \cdot a(x, y) = D_{(x,y)} a(f \times g)(x, y) = D_{(x,y)} a \cdot (f(x), g(y)). \quad (5.1-4)$$

Repare que a é uma função $a : M \times N \rightarrow \mathbb{R}$. Logo

$$\begin{aligned} D_{(x,y)} a : T_x M \times T_y N &\rightarrow \mathbb{R} \\ (u, v) &\mapsto D_{(x,y)} a \cdot (u, v). \end{aligned}$$

Além disso,

$$\begin{aligned} D_{(x,y)}a \cdot (f(x), g(y)) &= \left[\frac{\partial}{\partial x}a(x, y) \quad \frac{\partial}{\partial y}a(x, y) \right] \cdot \begin{bmatrix} f(x) \\ g(y) \end{bmatrix} \\ &= \frac{\partial}{\partial x}a(x, y)f(x) + \frac{\partial}{\partial y}a(x, y)g(y). \end{aligned} \quad (5.1-5)$$

Definindo as funções $a_y^1 : M \rightarrow \mathbb{R}$ e $a_x^2 : N \rightarrow \mathbb{R}$ por

$$a_y^1(x) = a(x, y) \quad \text{e} \quad a_x^2(y) = a(x, y),$$

temos que

$$\frac{\partial}{\partial x}a(x, y) = D_x a_y^1 \quad \text{e} \quad \frac{\partial}{\partial y}a(x, y) = D_y a_x^2.$$

Assim, a igualdade (5.1-5) fica como:

$$\begin{aligned} D_{(x,y)}a \cdot (f(x), g(y)) &= D_x a_y^1(x)f(x) + D_y a_x^2(y)g(y) \\ &= (\widehat{f a_y^1})(x) + (\widehat{g a_x^2})(y), \end{aligned} \quad (5.1-6)$$

onde estamos com a notação utilizada ao identificarmos campos de vetores com derivações (ver Seção 1.4.1). De (5.1-4) e (5.1-6) obtemos:

$$(\widehat{f \times g}) \cdot a(x, y) = (\widehat{f a_y^1})(x) + (\widehat{g a_x^2})(y). \quad (5.1-7)$$

Definimos o produto direto das famílias f_U e g_U como

$$f_U \times g_U = \{f_u \times g_u \mid u \in U\} \subset \text{Vec}(M \times N).$$

Seja

$$\Phi_{*x_0} : T_{x_0}M \rightarrow T_{y_0}N$$

o isomorfismo tal que,

$$\Phi_{*x_0}(\xi_{u_1, \dots, u_k}) = \eta_{u_1, \dots, u_k}, \quad u_i \in U, \quad k \in \mathbb{N}.$$

Usando (5.1-7) e a fórmula $[\widehat{V}, \widehat{W}] = \widehat{V} \circ \widehat{W} - \widehat{W} \circ \widehat{V}$, temos que o colchete de Lie sobre a família $f_U \times g_U$ é calculado como

$$[f_{u_1} \times g_{u_1}, f_{u_2} \times g_{u_2}] = [f_{u_1}, f_{u_2}] \times [g_{u_1}, g_{u_2}], \quad (5.1-8)$$

consequentemente,

$$\begin{aligned} & \left[f_{u_1} \times g_{u_1}, \left[f_{u_2} \times g_{u_2}, \left[\cdots \left[f_{u_{k-1}} \times g_{u_{k-1}}, f_{u_k} \times g_{u_k} \right] \cdots \right] \right] \right] (x_0, y_0) = \\ & = \left[f_{u_1}, \left[f_{u_2}, \left[\cdots \left[f_{u_{k-1}}, f_{u_k} \right] \cdots \right] \right] \right] (x_0) \times \left[g_{u_1}, \left[g_{u_2}, \left[\cdots \left[g_{u_{k-1}}, g_{u_k} \right] \cdots \right] \right] \right] (y_0) \\ & = \xi_{u_1, \dots, u_k} \times \eta_{u_1, \dots, u_k} \\ & = \xi_{u_1, \dots, u_k} \times \Phi_{*x_0}(\xi_{u_1, \dots, u_k}), \end{aligned}$$

para $u_i \in U$, $i = 1, \dots, k$, e $k \in \mathbb{N}$. Assim, o espaço $\text{Lie}_{(x_0, y_0)} f_U \times g_U$ é gerado por elementos da forma $(\xi_{u_1, \dots, u_k} \times \Phi_{*x_0}(\xi_{u_1, \dots, u_k}))$. Logo,

$$\dim \text{Lie}_{(x_0, y_0)} f_U \times g_U = n = \dim M.$$

Pelo Teorema da Órbita, a órbita da família $f_U \times g_U$ através do ponto (x_0, y_0) , $\mathcal{O}_{(x_0, y_0)}$, é uma subvariedade imersa (*localmente subvariedade*) conexa de $M \times N$. Como $(M \times N, f_U \times g_U)$ é analítico temos pelo Corolário 4.27, que

$$\begin{aligned} T_{(x_0, y_0)} \mathcal{O}_{(x_0, y_0)} &= \text{Lie}_{(x_0, y_0)} f_U \times g_U \\ &= \text{span}(\xi_{u_1, \dots, u_k} \times \Phi_{*x_0}(\xi_{u_1, \dots, u_k})) \\ &= \text{span}\{(v, \Phi_{*x_0} v) \mid v \in T_{x_0} M\} \\ &\subset T_{x_0} M \times T_{y_0} N. \end{aligned}$$

Além disso,

$$\dim T_{(x_0, y_0)} \mathcal{O}_{(x_0, y_0)} = \dim \text{Lie}_{(x_0, y_0)} f_U \times g_U = n.$$

Desta forma, a órbita $\mathcal{O}_{(x_0, y_0)}$ é uma subvariedade imersa n -dimensional, e cujo espaço tangente em (x_0, y_0) é o gráfico do isomorfismo linear Φ_{*x_0} .

Considere as projeções π_1 e π_2 de $M \times N$ em M e N respectivamente:

$$\begin{aligned}\pi_1 : M \times N &\rightarrow M, & \pi_1(x, y) &= x \\ \pi_2 : M \times N &\rightarrow N, & \pi_2(x, y) &= y.\end{aligned}$$

As restrições

$$\begin{aligned}\pi_1 : \mathcal{O}_{(x_0, y_0)} &\rightarrow M \\ \pi_2 : \mathcal{O}_{(x_0, y_0)} &\rightarrow N,\end{aligned}$$

são difeomorfismos locais pois, as diferenciais são dadas por:

$$\begin{aligned}(\pi_1)_* : \text{Lie}_{(x_0, y_0)} f_U \times g_U &\rightarrow T_{x_0} M \\ (v, \Phi_{*x_0} v) &\mapsto v\end{aligned}$$

$$\begin{aligned}(\pi_2)_* : \text{Lie}_{(x_0, y_0)} f_U \times g_U &\rightarrow T_{y_0} N \\ (v, \Phi_{*x_0} v) &\mapsto \Phi_{*x_0} v,\end{aligned}$$

que são aplicações injetoras em espaços de mesma dimensão n . Logo, são isomorfismos. Seja

$$\Psi = \pi_2 \circ \pi_1^{-1} : \pi_1(\mathcal{O}_{(x_0, y_0)}) \subset M \rightarrow \pi_2(\mathcal{O}_{(x_0, y_0)}) \subset N,$$

que é difeomorfismo local de M em N . Note que o gráfico de Ψ é a órbita, isto é, $\text{Graf}(\Psi) = \mathcal{O}_{(x_0, y_0)}$. Finalmente,

$$\begin{aligned}\Psi_*(f_u(x_0)) &= (\pi_{2*}) \circ (\pi_{1*})^{-1}(f_u(x_0)) \\ &= (\pi_{2*}) (\pi_{1*}^{-1}(f_u(x_0))) \\ &= (\pi_{2*}) (f_u(x_0), \Phi_{*x_0}(f_u(x_0))) \\ &= \Phi_{*x_0}(f_u(x_0)) \\ &= g_u(y_0), \quad u \in U.\end{aligned}$$

Portanto, $(f_U, x_0) \simeq (g_U, y_0)$.

(\Rightarrow) Esta implicação foi provada na discussão que precede o teorema. ■

Observação 5.3 A condição (5.1-3) do teorema acima pode ser melhorada com argumentos

conhecidos da álgebra. Como as famílias f_U e g_U são completamente não-holonômicas, ou seja,

$$\text{Lie}_{x_0} f_U = T_{x_0} M, \quad e \quad \text{Lie}_{y_0} g_U = T_{y_0} N,$$

por (5.1-1) podemos escolher uma base para $T_{x_0} M$ formada por vetores do forma ξ_{u_1, \dots, u_k} com $u_i \in U$ e $k \in \mathbb{N}$. Sejam $\xi_{\bar{\alpha}_1}, \dots, \xi_{\bar{\alpha}_n} \in T_{x_0} M$ tais que

$$\text{span}\{\xi_{\bar{\alpha}_1}, \dots, \xi_{\bar{\alpha}_n}\} = T_{x_0} M \quad (5.1-9)$$

onde $\bar{\alpha}_i = (u_{1i}, \dots, u_{ki}), i = 1, \dots, n$. Desta forma, podemos escrever todos os vetores de $T_{x_0} M$ nesta base, isto é,

$$\xi_{u_1, \dots, u_k} = \sum_{i=1}^n c_{u_1, \dots, u_k}^i \xi_{\bar{\alpha}_i} \quad (5.1-10)$$

onde c_{u_1, \dots, u_k}^i são coordenadas do vetor ξ_{u_1, \dots, u_k} nesta base.

Se existe um isomorfismo

$$\Phi : T_{x_0} M \rightarrow T_{y_0} N$$

com a condição (5.1-3), então

$$\Phi(\xi_{\bar{\alpha}_i}) = \eta_{\bar{\alpha}_i}$$

e o conjunto $\{\eta_{\bar{\alpha}_1}, \dots, \eta_{\bar{\alpha}_n}\}$ é base para $T_{y_0} N$, ou seja,

$$\text{span}\{\eta_{\bar{\alpha}_1}, \dots, \eta_{\bar{\alpha}_n}\} = T_{y_0} N. \quad (5.1-11)$$

Assim, podemos expressar qualquer vetor η_{u_1, \dots, u_k} nesta base, com as mesmas coordenadas:

$$\eta_{u_1, \dots, u_k} = \sum_{i=1}^n c_{u_1, \dots, u_k}^i \eta_{\bar{\alpha}_i}. \quad (5.1-12)$$

A recíproca é verdadeira: dadas bases como em (5.1-9) e (5.1-11) tais que as decomposições (5.1-10) e (5.1-12) são satisfeitas com as mesmas coordenadas, então existe um isomorfismo $T : T_{x_0} M \rightarrow T_{y_0} N$. Logo, a condição (5.1-3) pode ser reformulada, bastando analisar sua validade nos elementos da base, ou seja,

$$\Phi_{*x_0}(\xi_{\bar{\alpha}_i}) = \eta_{\bar{\alpha}_i} \quad i = 1, \dots, n.$$

Exemplo 5.4 Considere o sistema não linear sobre $\mathbb{R}_+ \times \mathbb{R}_+$:

$$\Sigma : \begin{cases} \dot{x}_1 = x_1 \ln(x_2) \\ \dot{x}_2 = -x_2 \ln(x_1) + x_2 u, \end{cases} \quad (5.1-13)$$

com $u \in \mathbb{R}$. Introduzindo a S-transformação $T : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R} \times \mathbb{R}$ definida por:

$$T(x_1, x_2) = (\ln(x_1), \ln(x_2)) = (z_1, z_2),$$

temos

$$\begin{aligned} \dot{z}_1 &= \frac{\dot{x}_1}{x_1} = \frac{x_1 \ln(x_2)}{x_1} = \ln(x_2) = z_2 \\ \dot{z}_2 &= \frac{\dot{x}_2}{x_2} = \frac{-x_2 \ln(x_1) + x_2 u}{x_2} = -z_1 + u. \end{aligned}$$

Logo, o sistema (5.1-13) nas novas coordenadas (z_1, z_2) têm a forma:

$$\Sigma' : \begin{cases} \dot{z}_1 = z_2 \\ \dot{z}_2 = -z_1 + u \end{cases}$$

que é um sistema linear.

5.2 Feedback equivalência

Nesta seção, iremos considerar sistemas de controle da forma

$$\dot{q} = f(q, u), \quad q \in M, \quad u \in U. \quad (5.2-14)$$

Como em seções anteriores, M é uma variedade diferenciável. Aqui assumiremos que U também é uma variedade diferenciável. Supomos também que para cada $u \in U$ fixado, $f(q, u)$ é um campo de vetores diferenciável sobre M e que a aplicação

$$(q, u) \mapsto f(q, u)$$

é diferenciável. Os controles admissíveis serão aplicações mensuráveis localmente limitadas

$$t \mapsto u(t) \in U.$$

Se um tal controle $u(t)$ é substituído no sistema (5.2-14), obtemos uma E.D.O

$$\dot{q} = f(q, u(t)),$$

na qual o lado direito é diferenciável em q e mensurável e localmente limitada em t . Para tais E.D.O's, temos a Teoria de Equações de Carathéodory, que garante existência e unicidade de problemas de Cauchy (veja a Seção 1.2 de [3] para a situação em espaços euclidianos e a Seção 2.4.1 de [1] para a demonstração deste fato em variedades diferenciáveis).

Fixados essas notações e resultados, introduziremos agora o conceito de transformação de feedback, que leva trajetórias de um sistema em trajetórias de outro sistema, mas possivelmente com um novo controle. O espaço dos novos parâmetros de controle será denotado por \widehat{U} , o qual também assumiremos como uma variedade diferenciável.

Definição 5.5 *Seja $\varphi : M \times U \rightarrow \widehat{U}$ uma aplicação diferenciável. Uma transformação da forma*

$$\dot{q} = f(q, u) \longrightarrow \dot{q} = f(q, \varphi(q, u)) \quad (5.2-15)$$

é chamada **transformação de feedback**. Neste caso, chamaremos apenas por **F-transformação**.

Observação 5.6 *Note que uma F-transformação reparametriza os controles $u \in U$ dependendo do estado $q \in M$. Note também que, se $q(t)$ é uma trajetória admissível para o sistema $\dot{q} = f(q, u)$ com respeito a um controle u então a mesma trajetória é admissível para o sistema transformado $\dot{q} = f(q, \varphi(q, u))$ com respeito ao controle $\widehat{u} = \varphi(q, u)$. Porém, não há garantia que trajetórias admissíveis do sistema transformado sejam admissíveis para o sistema inicial.*

A fim de obtermos uma F-equivalência, consideramos transformações invertíveis de feedback com

$$U = \widehat{U} \quad \text{e} \quad \varphi_q \in \text{Diff}(U),$$

para cada $q \in M$ fixado. Desta forma, a aplicação

$$\varphi : M \times U \rightarrow U$$

geram transformações invertíveis de feedback:

$$\dot{q} = f(q, u) \longleftrightarrow \dot{q} = f(q, \varphi(q, u)). \quad (5.2-16)$$

Neste caso dizemos que os sistemas $\dot{q} = f(q, u)$ e $\dot{q} = f(q, \varphi(q, u))$ são *F-equivalente*.

Para sistemas de controle afim:

$$\dot{q} = f_0(q) + \sum_{i=1}^k u_i g_i(q), \quad u = (u_1, \dots, u_k) \in \mathbb{R}^k, \quad q \in M, \quad (5.2-17)$$

temos as seguintes F-transformações com controle afim:

$$\varphi = (\varphi_1, \dots, \varphi_k) : M \times \mathbb{R}^k \rightarrow \mathbb{R}^k$$

onde

$$\varphi_i(q, u) = c_i(q) + \sum_{j=1}^k u_j d_{ij}(q) \quad (5.2-18)$$

com funções $c_i, d_{ij} : M \rightarrow \mathbb{R}$ para cada $i = 1, \dots, k$. Observe que essas transformações levam um sistema de controle afim em outro sistema de controle afim. Para ver isso, considere o caso em que $k = 2$. O sistema dado pode ser escrito na forma

$$\dot{q} = F(q, u),$$

onde

$$F(q, u) = f_0(q) + u_1 g_1(q) + u_2 g_2(q).$$

Neste caso, temos

$$\varphi(q, u) = (c_1(q) + u_1 d_{11}(q) + u_2 d_{12}(q), c_2(q) + u_1 d_{21}(q) + u_2 d_{22}(q)).$$

A F-transformação correspondente

$$F(q, u) \mapsto F(q, \varphi(q, u)),$$

leva $F(q, u)$ em

$$f_0(q) + (c_1(q) + u_1 d_{11}(q) + u_2 d_{12}(q))g_1(q) + (c_2(q) + u_2 d_{21}(q) + u_2 d_{22}(q))g_2(q) =$$

$$[f_0(q) + c_1(q)g_1(q) + c_2(q)g_2(q)] + [d_{11}(q)u_1 + d_{12}(q)u_2]g_1(q) + [d_{21}(q)u_1 + d_{22}(q)u_2]g_2(q).$$

Nosso objetivo é simplificar os sistemas de controle afim que são localmente equivalente à sistemas de controle lineares controláveis com respeito a state e feedback transformações. Assim, começaremos por sistemas de controle lineares.

5.3 Sistemas lineares

Faremos nesta seção uma classificação dos sistemas lineares controláveis para controle escalar e vetorial, isto é, obteremos a sua forma mais simples (ou normal). Considere o sistema de controle linear controlável em \mathbb{R}^n :

$$\dot{x} = Ax + \sum_{i=1}^k u_i b_i, \quad x \in \mathbb{R}^n \tag{5.3-19}$$

$$\text{span}\{A^j b_i \mid j = 1, \dots, n-1, i = 1, \dots, k\} = \mathbb{R}^n$$

onde $u = (u_1, \dots, u_k) \in \mathbb{R}^k$, $b_i \in \mathbb{R}^n$, $i = 1, \dots, k$. Assumiremos que os vetores b_i , $i = 1, \dots, k$ são linearmente independentes, ou seja,

$$\dim \text{span}\{b_1, \dots, b_k\} = k.$$

Se não for o caso, eliminamos alguns vetores para obter um conjunto linearmente independente. Vamos encontrar a forma normal para os sistemas (5.3-19).

Para os sistemas (5.3-19) aplicamos as F-transformações (5.2-18) que preservam a estrutura linear. Desta forma, tomaremos as funções c_i como funcionais lineares e as funções d_{ij} constantes. Usando o fato que todo funcional linear $T : \mathbb{R}^n \rightarrow \mathbb{R}$ é da forma $T(x) = \langle v, x \rangle$ para algum $v \in \mathbb{R}^n$ (*Teorema da Representação de Riez*), usaremos a notação

$c_i(x) = \langle c_i, x \rangle$ para indicar o funcional c_i . Assim:

$$\begin{aligned} c_i : \mathbb{R}^n &\rightarrow \mathbb{R} \\ x &\mapsto \langle c_i, x \rangle \end{aligned} \tag{5.3-20}$$

$$\begin{aligned} d_{ij} : \mathbb{R}^n &\rightarrow \mathbb{R} \\ x &\mapsto d_{ij}(x) = d_{ij} \end{aligned}$$

para todo $i, j = 1, \dots, k$. Desta forma, obtemos

$$\varphi = (\varphi_1, \dots, \varphi_k) : \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^k$$

tal que, para cada $i = 1, \dots, k$, temos

$$\varphi_i(x, u) = \langle c_i, x \rangle + \sum_{j=1}^k u_j d_{ij}, \tag{5.3-21}$$

para todo $x \in \mathbb{R}^n, u \in \mathbb{R}^k$.

Defina o operador linear

$$D : \text{span}\{b_1, \dots, b_k\} \rightarrow \text{span}\{b_1, \dots, b_k\} \tag{5.3-22}$$

cujas matriz de D na base $\mathcal{B} = \{b_1, \dots, b_k\}$ é (d_{ij}) , ou seja,

$$D(b_j) = \sum_{i=1}^k d_{ij} b_i, \quad [D]_{\mathcal{B}} = (d_{ij}).$$

Com estas informações, obtemos a F-transformação:

$$\begin{aligned}
 \dot{x} = Ax + \sum_{i=1}^k u_i b_i &\quad \mapsto \quad \dot{x} = Ax + \sum_{i=1}^k \varphi_i(x, u) b_i \\
 &= Ax + \sum_{i=1}^k \left(\langle c_i, x \rangle + \sum_{j=1}^k u_j d_{ij} \right) b_i \\
 &= Ax + \sum_{i=1}^k \langle c_i, x \rangle b_i + \sum_{i,j=1}^k u_j d_{ij} b_i \\
 &= \left(Ax + \sum_{i=1}^k \langle c_i, x \rangle b_i \right) + \sum_{j=1}^k u_j D b_j.
 \end{aligned}$$

Em resumo, obtemos

$$\dot{x} = Ax + \sum_{i=1}^k u_i b_i \quad \mapsto \quad \dot{x} = \left(Ax + \sum_{i=1}^k \langle c_i, x \rangle b_i \right) + \sum_{j=1}^k u_j D b_j.$$

Como um sistema linear é unicamente determinado fornecendo a matriz A e os vetores b_1, \dots, b_k , indicamos a F-transformação na forma:

$$(Ax, b_1, \dots, b_k) \quad \longrightarrow \quad \left(Ax + \sum_{i=1}^k \langle c_i, x \rangle b_i, D b_1, \dots, D b_k \right). \quad (5.3-23)$$

Queremos que esta transformação seja invertível, assim, vamos assumir que o operador D em (5.3-22) seja invertível.

As S-transformações lineares em \mathbb{R}^n são exatamente as transformações lineares inversíveis. Assim, tais transformações agem em sistemas lineares como segue:

$$(Ax, b_1, \dots, b_k) \quad \longrightarrow \quad (\mathcal{C}A\mathcal{C}^{-1}x, \mathcal{C}b_1, \dots, \mathcal{C}b_k), \quad (5.3-24)$$

onde $\mathcal{C} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é um operador linear invertível.

5.3.1 Sistemas lineares com controle escalar

Considere inicialmente um modelo bastante simples de sistema de controle linear, chamado sistema escalar de alta ordem:

$$u = x^{(n)} + \alpha_{n-1}x^{(n-1)} + \dots + \alpha_1x^{(1)} + \alpha_0x \in \mathbb{R}, \quad (5.3-25)$$

com $\alpha_0, \dots, \alpha_{n-1} \in \mathbb{R}$ e $x^{(k)}$ indica a derivada de ordem k . Podemos escrever o sistema (5.3-25) na forma clássica, considerando a mudança de variáveis

$$x_i = x^{(i-1)}, \quad i = 1, \dots, n. \quad (5.3-26)$$

Obtemos assim,

$$\begin{cases} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 \\ &\vdots \\ \dot{x}_{n-1} &= x_n \\ \dot{x}_n &= -\sum_{i=0}^{n-1} \alpha_i x_{i+1} + u, \end{cases} \quad (5.3-27)$$

que pode ser escrito na forma $\dot{x} = Ax + bu$ como

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \cdots & -\alpha_{n-1} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} u. \quad (5.3-28)$$

Tomaremos $-\sum_{i=0}^{n-1} \alpha_i x_{i+1} + u = \hat{u}$ como um novo controle. Para isto, basta aplicar a F-transformação (5.3-21) com

$$k = 1, \quad c_i = c = (-\alpha_0, \dots, -\alpha_{n-1}), \quad \text{e} \quad d_{ij} = d = 1.$$

Com isso, a transformação (5.3-21) tem a forma

$$\varphi_i(x, u) = \varphi(x, u) = -\sum_{i=0}^{n-1} \alpha_i x_{i+1} + u \quad (5.3-29)$$

onde

$$-\sum_{i=0}^{n-1} \alpha_i x_{i+1} = \langle c, x \rangle = \langle (-\alpha_0, \dots, -\alpha_{n-1}), (x_1, \dots, x_n) \rangle,$$

e

$$u = \sum_{j=1}^k u_j d_{ij}.$$

Obtemos assim,

$$\begin{cases} \dot{x}_1 & = & x_2 \\ \dot{x}_2 & = & x_3 \\ & \vdots & \\ \dot{x}_{n-1} & = & x_n \\ \dot{x}_n & = & \hat{u} \end{cases}, \quad (5.3-30)$$

que pode ser escrito na forma escalar como:

$$x^{(n)} = \dot{x}_n = \hat{u}. \quad (5.3-31)$$

Iremos provar na próxima proposição que um sistema de controle linear controlável com controle escalar é S-equivalente ao sistema (5.3-28), o qual por sua vez é F-equivalente ao sistema (5.3-30). Assim, dizemos que o sistema inicial é *state-feedback equivalente* ao sistema (5.3-30). Por essa razão, diremos que os sistemas (5.3-27) e (5.3-30) são formas normais de sistemas lineares com controle escalar sob S-transformações e F-transformações, respectivamente.

Temos o seguinte resultado.

Proposição 5.7 *Todo sistema de controle linear controlável com controle escalar:*

$$\dot{x} = Ax + ub, \quad u \in \mathbb{R}, \quad x \in \mathbb{R}^n \quad (5.3-32)$$

$$\text{span}(b, Ab, \dots, A^{n-1}b) = \mathbb{R}^n,$$

é state equivalente ao sistema da forma (5.3-28), e state-feedback equivalente ao sistema da forma (5.3-30).

Demonstração: Vamos mostrar que existe um operador linear invertível $\mathcal{C} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tal que a transformação (5.3-24) leva o sistema (5.3-32) no sistema (5.3-28). Para isso, iremos construir uma base $\mathcal{B} = \{v_1, \dots, v_n\}$ de \mathbb{R}^n tal que, o sistema (5.3-32) escrito nesta base transforma-se no sistema (5.3-28). Considere o polinômio característico de A :

$$p(\lambda) = \lambda^n + \alpha_{n-1}\lambda^{n-1} + \dots + \alpha_1\lambda + \alpha_0.$$

Pelo Teorema de Cayley-Hamilton, sabemos que:

$$p(A) = A^n + \alpha_{n-1}A^{n-1} + \cdots + \alpha_1A + \alpha_0 = 0.$$

Segue que,

$$A^n b + \alpha_{n-1}A^{n-1}b + \cdots + \alpha_1Ab + \alpha_0b = 0. \quad (5.3-33)$$

Vamos agora extrair da igualdade acima, os vetores v_1, \dots, v_n . Reescrevemos esta igualdade como,

$$A(\underbrace{A^{n-1}b + \alpha_{n-1}A^{n-2}b + \cdots + \alpha_1b}_{v_1}) + \alpha_0b = 0.$$

Note que $v_1 \neq 0$, caso contrário teríamos que $A^{n-1}b = -\alpha_{n-1}A^{n-2}b - \cdots - \alpha_1b$, contradizendo a condição de controlabilidade. Extraímos agora o vetor v_2 de v_1 , de modo análogo, isto é,

$$\begin{aligned} v_1 &= A^{n-1}b + \alpha_{n-1}A^{n-2}b + \cdots + \alpha_1b \\ &= A(\underbrace{A^{n-2}b + \alpha_{n-1}A^{n-3}b + \cdots + \alpha_2b}_{v_2}) + \alpha_1b. \end{aligned}$$

Temos que $v_2 \neq 0$ pelo mesmo argumento usado para v_1 . Novamente,

$$\begin{aligned} v_2 &= A^{n-2}b + \alpha_{n-1}A^{n-3}b + \cdots + \alpha_2b \\ &= A(\underbrace{A^{n-3}b + \alpha_{n-1}A^{n-4}b + \cdots + \alpha_3b}_{v_3}) + \alpha_2b, \end{aligned}$$

com $v_3 \neq 0$. Repetimos esse procedimento até obtermos:

$$\begin{aligned} v_{n-1} &= Ab + \alpha_{n-1}b \\ v_n &= b. \end{aligned}$$

Observe que, podemos escrever esta penúltima igualdade acima como,

$$v_{n-1} = Av_n + \alpha_{n-1}b, \quad v_n = b,$$

ou de outra forma

$$Av_n = v_{n-1} - \alpha_{n-1}b. \quad (5.3-34)$$

Com isso, podemos recuperar recursivamente *de trás para frente* todos os vetores v_1, \dots, v_n , definidos anteriormente.

Em resumo, obtemos os vetores:

$$\begin{aligned}
 v_1 &= A^{n-1}b + \alpha_{n-1}A^{n-2}b + \dots + \alpha_3A^2b + \alpha_2Ab + \alpha_1b \\
 v_2 &= A^{n-2}b + \alpha_{n-1}A^{n-3}b + \dots + \alpha_3Ab + \alpha_2b \\
 &\vdots \\
 v_{n-1} &= Ab + \alpha_{n-1}b \\
 v_n &= b.
 \end{aligned} \tag{5.3-35}$$

Vendo v_i como um vetor coluna, consideramos a matriz $\mathcal{B} = (v_1, \dots, v_n)$. Observe que:

$$\mathcal{B} = \begin{pmatrix} b & Ab & \dots & A^{n-2}b & A^{n-1}b \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_{n-1} & 1 \\ \alpha_2 & \alpha_3 & \dots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha_{n-1} & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Temos que \mathcal{B} é invertível pois, o determinante de \mathcal{B} é o produto do determinante de uma matriz controlável com o determinante de uma matriz triangular. Daí concluímos que (v_1, \dots, v_n) é uma base de \mathbb{R}^n . Definimos $\mathcal{C} = \mathcal{B}^{-1}$. Temos que,

$$b = \mathcal{B} \cdot \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \Rightarrow \mathcal{C}b = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Aplicamos a matriz A na base \mathcal{B} , ou seja, $A\mathcal{B} = A(v_1 \dots v_n)$. Vamos proceder apli-

cando A nas colunas de \mathcal{B} . Usando (5.3-34), temos

$$Av_n = v_{n-1} - \alpha_{n-1}b = \mathcal{B} \cdot \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -\alpha_{n-1} \end{pmatrix}, \quad Av_{n-1} = v_{n-2} - \alpha_{n-2}b = \mathcal{B} \cdot \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ -\alpha_{n-2} \end{pmatrix}.$$

Repetimos esse processo até obtermos as primeiras colunas:

$$Av_2 = v_1 - \alpha_1b = \mathcal{B} \cdot \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -\alpha_1 \end{pmatrix}, \quad \text{e} \quad Av_1 = -\alpha_0b = \mathcal{B} \cdot \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -\alpha_0 \end{pmatrix}.$$

Logo, na forma matricial, obtemos:

$$A\mathcal{B} = \mathcal{B} \cdot \begin{pmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_{n-2} & -\alpha_{n-1} \end{pmatrix},$$

ou melhor,

$$\mathcal{C}A\mathcal{C}^{-1} = \mathcal{B}^{-1}A\mathcal{B} = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_{n-2} & -\alpha_{n-1} \end{pmatrix}.$$

Além disso, já obtivemos

$$Cb = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Desta forma, mostramos que existe um operador linear invertível C que transforma o sistema (5.3-32) no sistema (5.3-28). Logo, esses sistemas são S-equivalente. Note também que a base \mathcal{B} construída é única, o que significa que a transformação do sistema em sua forma normal é única. Para finalizarmos esta demonstração, consideremos o novo parâmetro de controle

$$\hat{u} = - \sum_{i=0}^{n-1} \alpha_i x_{i+1} + u.$$

Pela discussão feita nos parágrafos que antecedem a proposição, temos que o sistema (5.3-32) é state-feedback equivalente ao sistema (5.3-30).

■

5.3.2 Sistemas lineares com controle vetorial

Nosso objetivo nesta subseção é encontrar a forma normal de um sistema de controle linear controlável geral via state e feedback equivalência. Tal forma é chamada *forma normal de Brunovský*. Para tanto, é fundamental os *índices de Kronecker* do sistema linear inicial, (ver [2]) que são invariantes por state-feedback equivalência. Iniciemos a descrição de alguns desses invariantes.

Considere o sistema linear controlável com controle vetorial:

$$\dot{x} = Ax + \sum_{i=1}^k u_i b_i, \quad u = (u_1, \dots, u_k) \in \mathbb{R}^k, \quad x \in \mathbb{R}^n, \quad (5.3-36)$$

$$\text{span}\{A^j b_i \mid j = 1, \dots, n-1, i = 1, \dots, k\} = \mathbb{R}^n.$$

Assumiremos que o conjunto $\{b_1, \dots, b_k\}$ é linearmente independente. Definimos os

seguintes subespaços em \mathbb{R}^n :

$$D^m = \text{span}\{A^j b_i \mid j = 0, \dots, m-1; i = 1, \dots, k\}, \quad m = 1, \dots, n. \quad (5.3-37)$$

Note que as dimensões $\dim D^m$ são preservadas por S-transformações (5.3-24) pois, o operador \mathcal{C} é invertível. Desta forma, dizemos que os números $\dim D^m$ são *S-invariantes*, isto é, são invariantes por S-transformações (5.3-24). Vamos verificar agora que os subespaços D^m são *F-invariantes*, isto é, são invariantes por F-transformação. De fato, considere a F-transformação vista em (5.3-23)

$$(Ax, b_1, \dots, b_k) \longrightarrow \left(Ax + \sum_{i=1}^k \langle c_i, x \rangle b_i, Db_1, \dots, Db_k \right).$$

Podemos decompor esta transformação em duas:

$$(Ax, b_1, \dots, b_k) \longrightarrow \left(Ax + \sum_{i=1}^k \langle c_i, x \rangle b_i, b_1, \dots, b_k \right), \quad (5.3-38)$$

$$(Ax, b_1, \dots, b_k) \longrightarrow (Ax, Db_1, \dots, Db_k). \quad (5.3-39)$$

A transformação (5.3-39) é mais simples uma vez que o operador

$$D : \text{span}\{b_1, \dots, b_k\} \rightarrow \text{span}\{b_1, \dots, b_k\}$$

é invertível e $D^{m-1} \subset D^m$, $m = 2, \dots, n$. Em outras palavras, o operador D muda apenas os vetores $b_i \in D^1 \subset D^m$, $i = 1, \dots, k$. Portanto, os espaços D^m são preservadas por esta transformação.

Para mostrarmos que os espaços D^m são preservados pela transformação (5.3-38), vamos verificar primeiro a seguinte igualdade:

$$\widehat{A}^j x \equiv A^j x \text{ mod } D^j, \quad j = 1, \dots, n \quad (5.3-40)$$

onde

$$\widehat{A}x = Ax + \sum_{i=1}^k \langle c_i, x \rangle b_i. \quad (5.3-41)$$

De fato, para $j = 1$ a equação (5.3-41) acima, nos dá

$$\widehat{A}x - Ax = \sum_{i=1}^k \langle c_i, x \rangle b_i \in D^1.$$

Para mostrar que a equação (5.3-40) é verdadeira para $j > 1$ vamos proceder por iteração da equação (5.3-41) e usar que $D^{m-1} \subset D^m$, $m = 2, \dots, n$, para concluir o desejado.

Por simplicidade, denotaremos $x_1 = \widehat{A}x$, e obtemos,

$$\begin{aligned} \widehat{A}^2x = \widehat{A}(\widehat{A}x) &= \widehat{A} \left(Ax + \sum_{i=1}^k \langle c_i, x \rangle b_i \right) \\ &= A \left(Ax + \sum_{i=1}^k \langle c_i, x \rangle b_i \right) + \sum_{i=1}^k \langle c_i, x_1 \rangle b_i \\ &= A^2x + \sum_{i=1}^k \langle c_i, x \rangle Ab_i + \sum_{i=1}^k \langle c_i, x_1 \rangle b_i, \end{aligned}$$

segue que,

$$\widehat{A}^2x - A^2x = \underbrace{\sum_{i=1}^k \langle c_i, x \rangle Ab_i}_{\in D^2} + \underbrace{\sum_{i=1}^k \langle c_i, x_1 \rangle b_i}_{\in D^1} \in D^2.$$

Sejam $x_1 = \widehat{A}x$ e $x_l = \widehat{A}(x_{l-1})$, $l = 2, \dots, n$. De maneira geral, temos

$$\begin{aligned} \widehat{A}^n x &= \widehat{A} \left(A^{n-1}x + \sum_{i=1}^k \langle c_i, x \rangle A^{n-2}b_i + \sum_{i=1}^k \langle c_i, x_1 \rangle A^{n-3}b_i + \dots + \sum_{i=1}^k \langle c_i, x_{n-2} \rangle b_i \right) \\ &= A^n x + \underbrace{\sum_{i=1}^k \langle c_i, x \rangle A^{n-1}b_i}_{\in D^n} + \underbrace{\sum_{i=1}^k \langle c_i, x_1 \rangle A^{n-2}b_i}_{\in D^{n-1}} + \dots + \underbrace{\sum_{i=1}^k \langle c_i, x_{n-1} \rangle b_i}_{\in D^1}. \end{aligned}$$

Logo,

$$\widehat{A}^n x - A^n x \in D^n,$$

e concluímos que

$$\widehat{A}^j x \equiv A^j x \pmod{D^j}, \quad j = 1, \dots, n.$$

Vamos agora mostrar que os espaços D^m são preservados pela F-transformação, isto é, se

$$\widehat{D}^m = \text{span}\{\widehat{A}^j b_i \mid j = 0, \dots, m-1; i = 1, \dots, n\}$$

então

$$\widehat{D}^m = D^m, \quad m = 1, \dots, n.$$

A igualdade (5.3-40) implica em

$$\widehat{A}^j b_i = A^j b_i + v_i,$$

com $v_i \in D^j$. Desta forma, $\widehat{A}^j b_i \in D^{j+1} + D^j \subset D^{j+1}$. Como \widehat{D}^m é gerado por elementos da forma $\widehat{A}^j b_i$ com $0 \leq j \leq m-1$, segue que $\widehat{D}^m \subset D^m$. Vamos usar indução sobre m para mostrar a inclusão oposta. Como D^1 e \widehat{D}^1 são gerados somente pelos b_i 's, a inclusão $D^1 \subset \widehat{D}^1$ é trivial para $m = 1$. Suponha que $D^{m-1} \subset \widehat{D}^{m-1}$. Observe que obtemos D^m a partir de D^{m-1} unindo o conjunto $\{A^{m-1}b_i \mid i = 1, \dots, n\}$ ao conjunto de geradores de D^{m-1} . Pela igualdade (5.3-40), temos que

$$\widehat{A}^{m-1} b_i = A^{m-1} b_i + v_i,$$

com $v_i \in D^{m-1} \subset \widehat{D}^{m-1}$. Assim:

$$\begin{aligned} A^{m-1} b_i &= (A^{m-1} b_i + v_i) - v_i \\ &= \widehat{A}^{m-1} b_i - v_i \in \widehat{D}^m + \widehat{D}^{m-1} \subset \widehat{D}^m. \end{aligned}$$

Logo, todos os geradores de D^m pertencem a \widehat{D}^m , de onde segue a inclusão desejada. Portanto, cada espaço D^m é F-invariante e, visto que os números $\dim D^m$, $m = 1, \dots, n$ são S-invariantes, temos que as dimensões $\dim D^m$, $m = 1, \dots, n$, são state-feedback invariantes.

5.3.2.1 Índices de Kronecker

Vamos introduzir agora os *índices de Kronecker* de um sistema linear controlável e mostrar que tais índices são state-feedback invariante expressando-os através das dimensões de D^m , $m = 1, \dots, n$.

Considere a matriz com entradas em \mathbb{R}^n formada pelos geradores do subespaço D^n :

$$\begin{pmatrix} b_1 & b_2 & \cdots & b_k \\ Ab_1 & Ab_2 & \cdots & Ab_k \\ \vdots & \vdots & \vdots & \vdots \\ A^{n-1}b_1 & A^{n-1}b_2 & \cdots & A^{n-1}b_k \end{pmatrix}_{n \times k}. \quad (5.3-42)$$

Repare que D^1 é gerado pela primeira linha, D^2 é gerado pela primeira e segunda linha, ..., D^n é gerado por todas as linhas.

Vamos substituir cada entrada da matriz (5.3-42) acima por *circulo* (\circ) ou *estrela* (\star) de acordo com a seguinte regra. Primeiramente, formamos uma lista colocando cada linha da matriz (5.3-42) em sequencia numa mesma linha:

$$\Gamma = \{b_1, b_2, \dots, b_k; Ab_1, Ab_2, \dots, Ab_k; \dots; A^{n-1}b_1, A^{n-1}b_2, \dots, A^{n-1}b_k\}.$$

A substituição será feita mediante a seguinte regra: da esquerda para a direita, um vetor de Γ será substituído por (\star) se este for *linearmente independente em relação aos vetores anteriores*, e por (\circ) caso contrário. Este procedimento nos dá uma matriz da forma:

$$\Sigma = \begin{pmatrix} \star & \star & \star & \cdots & \star \\ \circ & \star & \star & \cdots & \circ \\ \circ & \star & \circ & \cdots & \circ \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \circ & \star & \circ & \cdots & \circ \end{pmatrix}_{n \times k}. \quad (5.3-43)$$

Observação 5.8 (1) O número de estrelas (\star) é exatamente n , devido a condição de controlabilidade.

(2) Toda a primeira linha é substituída por (\star) pois, assumimos que o conjunto $\{b_i \mid i = 1, \dots, k\}$ é linearmente independente.

(3) Se um vetor é substituído por (\circ) então, na mesma coluna, abaixo deste vetor, todos os vetores serão substituído por (\circ). De fato, se um vetor $A^j b_i \in \Gamma$ é substituído por (\circ)

então, $A^j b_i$ é combinação linear dos vetores anteriores a ele, isto é,

$$A^j b_i \in \text{span}\{A^\gamma b_\gamma \mid \gamma < i\} + \text{span}\{A^\beta b_\gamma \mid \beta < j, \gamma = 1, \dots, k\}.$$

Desta forma,

$$\begin{aligned} A \cdot A^j b_i &= A^{j+1} b_i \in \text{span}\{A^{j+1} b_\gamma \mid \gamma < i\} + \text{span}\{A^\beta b_\gamma \mid \beta < j+1, \gamma = 1, \dots, k\} \\ &\vdots \\ A \cdot A^{n-2} b_i &= A^{n-1} b_i \in \text{span}\{A^{n-1} b_\gamma \mid \gamma < i\} + \text{span}\{A^\beta b_\gamma \mid \beta < n-1, \gamma = 1, \dots, k\}. \end{aligned}$$

Portanto, cada coluna da matriz Σ é formada por uma coluna de estrelas (\star) sobre uma coluna de círculos (\circ).

Denote por n_1 a quantidade de estrelas da maior coluna de estrelas de Σ , por n_2 a quantidade de estrelas da segunda maior coluna de estrelas de Σ , ..., e por fim, denote por n_k a quantidade de estrelas da menor coluna de estrelas de Σ . Naturalmente teremos,

$$n_1 \geq n_2 \geq \dots \geq n_k, \quad \text{e} \quad \sum_{i=1}^k n_i = n,$$

pela Observação 5.8. Chamamos de *índices de Kronecker* do sistema (5.3-36) os números $n_1, n_2, \dots, n_k \in \mathbb{N}$ encontrados acima. Isto nos dá uma condição de controlabilidade simplificada:

$$\text{span}\{b_1, Ab_1, \dots, A^{n_1-1} b_1; \dots; b_k, Ab_k, \dots, A^{n_k-1} b_k\} = \mathbb{R}^n. \quad (5.3-44)$$

Resta mostrarmos que os índices de Kronecker são state-feedback invariantes. Para isso, mostraremos que podem ser expressados em termos das dimensões dos subespaços D^m , $m = 1, \dots, n$. Temos,

$$\begin{aligned} \dim D^1 &= k = \text{número de estrelas na primeira linha } \Sigma, \\ \dim D^2 &= \text{número de estrelas nas duas primeiras linhas de } \Sigma, \\ &\vdots \\ \dim D^n &= n. \end{aligned}$$

Defina

$$\begin{aligned} \Delta : \{1, \dots, n\} &\rightarrow \{1, \dots, k\} \\ i &\mapsto \dim D^i - \dim D^{i-1} \end{aligned}$$

Observe que $\Delta(i)$ é o número de estrelas na i -ésima linha de Σ . Vamos permutar as colunas da matriz Σ , colocando na primeira coluna a coluna com mais estrelas, na segunda coluna a próxima coluna com mais estrelas, ..., na k -ésima coluna a que tem menos estrelas. Desta forma, Σ torna-se uma matriz na forma de bloco triangular, que denotaremos por Σ' . A partir disso podemos constatar o seguinte: o número de linhas que possuem pelo menos uma estrela é exatamente a altura da maior coluna de estrelas. O número de linhas que possuem pelo menos duas estrelas é a altura da segunda maior coluna de estrelas. Em geral, o número de linhas que possuem pelo menos i estrelas é o índice de Kronecker n_i . Além disso, a quantidade de estrelas na j -ésima linha é $\Delta(j)$. Portanto, o índice de Kronecker n_i é a quantidade de números $\Delta(j)$ tais que $\Delta(j) \geq i$. Como $\Delta(j) = \dim D^j - \dim D^{j-1}$ e essas dimensões são invariantes por F-transformações, concluímos que os índices de Kronecker são state-feedback invariantes. Destacamos esse fato no seguinte teorema:

Teorema 5.9 *Os índices de Kronecker são state-feedback invariantes.*

Para fixar ideias, consideremos um exemplo.

Exemplo 5.10 *Considere o sistema de controle linear controlável $\dot{x} = Ax + Bu$ com $n = 5$ e $k = 3$, onde A é uma matriz $n \times n$ constante e B é gerado pelos vetores linearmente independentes b_1, b_2, b_3 , dados por:*

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad b_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad b_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Resta encontrar mais dois vetores para que o $\text{span}\{b_1, b_2, b_3, A^{j_1}b_{i_1}, A^{j_2}b_{i_2}\} = \mathbb{R}^5$ para alguns $j_1, j_2 = 1, 2, 3, 4$ e $i_1, i_2 = 1, 2, 3$. Fazendo os cálculos, obtemos que o conjunto $\{b_1, b_2, b_3, Ab_1, A^2b_1\}$ é uma base para \mathbb{R}^5 , isto é

$$\text{span}\{(0, 0, 1, 0, 0)^T, (0, 0, 0, 1, 0)^T, (0, 0, 0, 0, 1)^T, (0, 1, 0, 0, 0)^T, (1, 0, 0, 0, 0)^T\} = \mathbb{R}^5.$$

A matriz (5.3-42) junto com a regra definida para a substituição por estrelas e círculos, fica na forma

$$\Sigma' = \begin{pmatrix} \star & \star & \star \\ \star & \circ & \circ \\ \star & \circ & \circ \\ \circ & \circ & \circ \\ \circ & \circ & \circ \end{pmatrix}_{5 \times 3} .$$

Observe que a quantidade de estrelas são $5 = \dim \mathbb{R}^5$. Para $i = 1, 2, 3$, temos que os índices de Kronecker são definidos como

“ n_i é o número de linhas que possuem pelo menos i estrelas”,

obtemos assim $n_1 = 3$, $n_2 = 1$, e $n_3 = 1$. E para $j = 1, 2, 3$ temos que

$$“\Delta(j) = \dim D^j - \dim D^{j-1},”$$

segue que, $\Delta(1) = 3$, $\Delta(2) = 1$ e $\Delta(3) = 1$. Finalmente, para $i = 1, 2, 3$, temos que

“ n_i é a quantidade de números $\Delta(j)$ tais que $\Delta(j) \geq i$ ”,

portanto, $n_1 = 3$, $n_2 = 1$, e $n_3 = 1$, coincidindo com os índices que já obtivemos.

5.3.2.2 Forma Normal de Brunovský

Vamos encontrar a forma mais simples de um sistema de controle linear controlável com controle vetorial via S-transformação.

Seja E a matriz linha com entradas em \mathbb{R}^n formada precisamente pelos vetores do conjunto gerador de \mathbb{R}^n obtido na igualdade (5.3-44). Ou seja,

$$E = \left(b_1 \quad \dots \quad A^{n_1-1}b_1 \quad \dots \quad b_2 \quad \dots \quad A^{n_2-1}b_2 \quad \dots \quad b_k \quad \dots \quad A^{n_k-1}b_k \right).$$

Temos que E é invertível pela construção feita anteriormente. Seja $F = E^{-1}$ e escrevemos

$$F^T = \left(p_1 \quad p_2 \quad \dots \quad p_n \right), \quad p_i \in \mathbb{R}^n, \quad i = 1, \dots, n.$$

Vamos definir um matriz C da seguinte forma: em F^T selecionamos, da esquerda para direita, k vetores cujos índices são da seguinte forma: $n_1, n_1 + n_2, n_1 + n_2 + n_3, \dots, n_1 +$

$n_2 + \dots + n_k = n$, onde n_i é o índice de Kronecker. Isto é, os vetores selecionados são:

$$p_{n_1}, p_{n_1+n_2}, \dots, p_{n_1+n_2+\dots+n_k} = p_n.$$

Com esses vetores, definimos a matriz coluna \mathcal{C} com entradas em \mathbb{R}^n :

$$\mathcal{C} = \begin{pmatrix} p_{n_1} \\ p_{n_1}A \\ \vdots \\ p_{n_1}A^{n_1-1} \\ p_{n_1+n_2} \\ p_{n_1+n_2}A \\ \vdots \\ p_{n_1+n_2}A^{n_2-1} \\ \cdot \\ \vdots \\ \cdot \\ p_n \\ p_nA \\ \vdots \\ p_nA^{n_k-1} \end{pmatrix}_{n \times 1}. \quad (5.3-45)$$

Note que, a matriz \mathcal{C} é formada por k blocos de ordem $n_i \times 1$, $i = 1, \dots, k$.

O Lema a seguir é técnico e tem o objetivo de simplificar os cálculos que faremos.

Lema 5.11 *Com a notação acima, temos*

$$p_{n_1}A^{n_1-1}b_1 = p_{n_1+n_2}A^{n_2-1}b_2 = \dots = p_nA^{n_k-1}b_k = 1.$$

Demonstração: Segue diretamente da definição, temos $F \cdot E = I$, isto é,

$$I = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix} \cdot \left(b_1 \quad \dots \quad A^{n_1-1}b_1 \quad \dots \quad b_2 \quad \dots \quad A^{n_2-1}b_2 \quad \dots \quad b_k \quad \dots \quad A^{n_k-1}b_k \right).$$

Para fazermos esse cálculo, aplicamos F nas colunas de E , obtendo a matriz de ordem $n \times n$:

$$\begin{pmatrix} p_1b_1 & p_1Ab_1 & \dots & p_1A^{n_1-1}b_1 & \dots & p_1b_k & p_1Ab_k & \dots & p_1A^{n_k-1}b_k \\ p_2b_1 & p_2Ab_1 & \dots & p_2A^{n_1-1}b_1 & \dots & p_2b_k & p_2Ab_k & \dots & p_2A^{n_k-1}b_k \\ p_3b_1 & p_3Ab_1 & \dots & p_3A^{n_1-1}b_1 & \dots & p_3b_k & p_3Ab_k & \dots & p_3A^{n_k-1}b_k \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & p_{n_k}b_k & p_{n_k}Ab_k & \dots & p_{n_k}A^{n_k-1}b_k \\ p_{n_1}b_1 & p_{n_1}Ab_1 & \dots & p_{n_1}A^{n_1-1}b_1 & \dots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & p_{n-2}b_k & p_{n-2}Ab_k & \dots & p_{n-2}A^{n_k-1}b_k \\ p_{n-1}b_1 & p_{n-1}Ab_1 & \dots & p_{n-1}A^{n_1-1}b_1 & \dots & p_{n-1}b_k & p_{n-1}Ab_k & \dots & p_{n-1}A^{n_k-1}b_k \\ p_nb_1 & p_nAb_1 & \dots & p_nA^{n_1-1}b_1 & \dots & p_nb_k & p_nAb_k & \dots & p_nA^{n_k-1}b_k \end{pmatrix}.$$

Como esta é igual a matriz identidade, temos que todos os elementos da diagonal principal são 1, em particular temos:

$$p_{n_1}A^{n_1-1}b_1 = p_{n_1+n_2}A^{n_2-1}b_2 = \dots = p_nA^{n_k-1}b_k = 1.$$

Enfatizamos que todos elementos fora da diagonal são *nulos*. ■

Vamos mostrar que a matriz C é invertível. Para isso, aplicamos C nas colunas de E obtendo assim uma matriz P invertível. De fato, a matriz P obtida pela aplicação de C nas colunas de E é, em blocos

$$C \cdot E = P = \begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1k} \\ P_{21} & P_{22} & \cdots & P_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ P_{k1} & P_{k2} & \cdots & P_{kk} \end{pmatrix}_{k \times k},$$

onde cada bloco P_{ij} têm ordem $n_i \times n_j$, para $i, j = 1, \dots, k$. Os blocos P_{ii} têm a seguinte forma: para $i = 1$,

$$P_{11} = \begin{pmatrix} p_{n_1} b_1 & p_{n_1} A b_1 & \cdots & p_{n_1} A^{n_1-1} b_1 \\ p_{n_1} A b_1 & p_{n_1} A^2 b_1 & \cdots & p_{n_1} A^{n_1} b_1 \\ \vdots & \vdots & \ddots & \vdots \\ p_{n_1} A^{n_1-2} b_1 & p_{n_1} A^{n_1-1} b_1 & \cdots & p_{n_1} A^{2n_1+1} b_1 \\ p_{n_1} A^{n_1-1} b_1 & p_{n_1} A^{n_1} b_1 & \cdots & p_{n_1} A^{2n_1-2} b_1 \end{pmatrix}_{n_1 \times n_1},$$

para $i = 2$,

$$P_{22} = \begin{pmatrix} p_{n_1+n_2} b_2 & p_{n_1+n_2} A b_2 & \cdots & p_{n_1+n_2} A^{n_2-1} b_2 \\ p_{n_1+n_2} A b_2 & p_{n_1+n_2} A^2 b_2 & \cdots & p_{n_1+n_2} A^{n_2} b_2 \\ \vdots & \vdots & \ddots & \vdots \\ p_{n_1+n_2} A^{n_2-2} b_2 & p_{n_1+n_2} A^{n_2-1} b_2 & \cdots & p_{n_1+n_2} A^{2n_2+1} b_2 \\ p_{n_1+n_2} A^{n_2-1} b_2 & p_{n_1+n_2} A^{n_2} b_2 & \cdots & p_{n_1+n_2} A^{2n_2-2} b_2 \end{pmatrix}_{n_2 \times n_2},$$

por fim, para $i = k$,

$$P_{kk} = \begin{pmatrix} p_n b_k & p_n A b_k & \cdots & p_n A^{n_k-1} b_k \\ p_n A b_k & p_n A^2 b_k & \cdots & p_n A^{n_k} b_k \\ \vdots & \vdots & \ddots & \vdots \\ p_n A^{n_k-2} b_k & p_n A^{n_k-1} b_k & \cdots & p_n A^{2n_k+1} b_k \\ p_n A^{n_k-1} b_k & p_n A^{n_k} b_k & \cdots & p_n A^{2n_k-2} b_k \end{pmatrix}_{n_k \times n_k}.$$

Note que, em todos os blocos P_{ii} , temos a **anti-diagonal** (ou diagonal secundária) formada por 1 (um), e 0 (zero) em todas as outras entradas, devido ao Lema 5.11. Nos blocos restantes, P_{ij} com $i \neq j$, temos todas as entradas nulas pois, esses blocos não possuem entradas da forma $p_{n_1} A^{n_1-1} b_1 = p_{n_1+n_2} A^{n_2-1} b_2 = \cdots = p_n A^{n_k-1} b_k$. Logo, os

blocos de P têm as formas:

$$P_{ii} = \begin{pmatrix} 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \end{pmatrix}_{n_i \times n_i \ (i=j)}, \quad P_{ij} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix}_{n_i \times n_j \ (i \neq j)}.$$

Pelo cálculo do determinante

$$\det(\mathcal{C}) \cdot \det(E) = \det(\mathcal{C} \cdot E) = \det(P) = \det(P_{11}) \cdot \det(P_{22}) \cdots \det(P_{kk}) = 1,$$

temos que $\det(\mathcal{C}) \neq 0$. Portanto a matriz \mathcal{C} é invertível.

Enunciamos agora o principal resultado desta seção.

Teorema 5.12 *Todo sistema linear controlável com controle vetorial:*

$$\dot{x} = Ax + Bu, \quad u = (u_1, \dots, u_k) \in \mathbb{R}^k, \quad x \in \mathbb{R}^n, \quad (5.3-46)$$

$$\text{span}\{A^j b_i \mid j = 1, \dots, n-1, i = 1, \dots, k\} = \mathbb{R}^n,$$

com B na forma matricial

$$B = \begin{pmatrix} b_1 & b_2 & \cdots & b_k \end{pmatrix},$$

é S -equivalente a um sistema da forma:

$$\begin{cases} \dot{x}_1^i = x_2^i \\ \dot{x}_2^i = x_3^i \\ \vdots \\ \dot{x}_{n_i-1}^i = x_{n_i}^i \\ \dot{x}_{n_i}^i = \sum_{j=1}^{n_1} \alpha_{n_i j}^{i1} x_j^1 + \sum_{j=1}^{n_2} \alpha_{n_i j}^{i2} x_j^2 + \cdots + \sum_{j=1}^{n_k} \alpha_{n_i j}^{ik} x_j^k + u_i, \end{cases} \quad (5.3-47)$$

para cada $i = 1, \dots, k$, e state-feedback equivalente a um sistema da forma:

$$\left\{ \begin{array}{l} x^{1(n_1)} = \hat{u}_1 \\ x^{2(n_2)} = \hat{u}_2 \\ \vdots \\ x^{k-1(n_{k-1})} = \hat{u}_{n_{k-1}} \\ x^{k(n_k)} = \hat{u}_k \end{array} \right. \quad (5.3-48)$$

onde n_1, \dots, n_k são os índices de Kronecker do sistema (5.3-46). A forma em (5.3-48) é chamada **forma normal de Brunovský do sistema** (5.3-46).

Antes de provarmos esse resultado, vamos ver um exemplo que ajudará a compreender melhor a notação.

Exemplo 5.13 Tome um sistema linear $\dot{x} = Ax + Bu$ com $n = 5$, $k = 3$ e índices de Kronecker $n_1 = 2$, $n_2 = 2$ e $n_3 = 1$. Para simplificar um pouco a carregada notação do teorema, iremos escrever $x_1^1 = x_1$, $x_2^1 = x_2$, $x_1^2 = x_3$, $x_2^2 = x_4$ e $x_1^3 = x_5$. Também substituiremos os coeficientes $\alpha_{n_i j}^{ik}$ por α 's, β 's e γ 's de maneira conveniente. Por fim, escrevemos os sistemas para $i = 1, 2, 3$ num único sistema. A forma normal (5.3-47) para esse sistema fica

$$\left\{ \begin{array}{l} \dot{x}_1 = x_2 \\ \dot{x}_2 = (\sum_i^5 \alpha_i x_i) + u_1 \\ \dot{x}_3 = x_4 \\ \dot{x}_4 = (\sum_i^5 \beta_i x_i) + u_2 \\ \dot{x}_5 = (\sum_i^5 \gamma_i x_i) + u_3 \end{array} \right. , \quad (5.3-49)$$

que pode ser escrito na forma matricial $\dot{x} = \hat{A}x + \hat{B}u$ com

$$\hat{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 \\ 0 & 0 & 0 & 1 & 0 \\ \beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_5 \\ \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 & \gamma_5 \end{pmatrix} \quad e \quad \hat{B} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Para a forma normal (5.3-48), mantemos a notação do teorema e obtemos

$$\begin{cases} \ddot{x}^1 &= \hat{u}_1 \\ \ddot{x}^2 &= \hat{u}_2 \\ \dot{x}^3 &= \hat{u}_3 \end{cases} .$$

Podemos reescrever o sistema acima na forma $\dot{x} = \hat{A}x + \hat{B}u$ introduzindo as seguintes variáveis:

$$x_1 = x^1, \quad x_2 = \dot{x}^1, \quad x_3 = x^2, \quad x_4 = \dot{x}^2, \quad x_5 = x^3.$$

Assim o sistema envolvendo derivadas de ordem 2 é escrito como o seguinte sistema envolvendo somente derivadas de ordem 1:

$$\begin{cases} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \hat{u}_1 \\ \dot{x}_3 &= x_4 \\ \dot{x}_4 &= \hat{u}_2 \\ \dot{x}_5 &= \hat{u}_3 \end{cases} . \quad (5.3-50)$$

Comparando (5.3-49) e (5.3-50), percebemos que a F-transformação é dada por

$$\hat{u}_1 = \left(\sum_i^5 \alpha_i x_i \right) + u_1, \quad \hat{u}_2 = \left(\sum_i^5 \beta_i x_i \right) + u_2, \quad e \quad \hat{u}_3 = \left(\sum_i^5 \gamma_i x_i \right) + u_3,$$

pondo

$$\hat{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad e \quad \hat{B} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Obtemos a forma canônica na forma matricial. Ressaltamos que é sempre possível escrever a expressão (mais compacta) (5.3-48) na forma matricial, usando mudanças de variáveis como as deste exemplo. Também é possível obter diretamente a forma normal de Brunovský na forma matricial (para detalhes, ver [11]).

Demonstração: Vamos mostrar que o sistema (5.3-46) é S-equivalente ao sistema (5.3-47)

exibindo uma matriz invertível C satisfazendo (5.3-24), isto é,

$$\widehat{A} = CAC^{-1}, \quad \text{e} \quad \widehat{B} = CB,$$

onde \widehat{A} e \widehat{B} são matrizes nas formas normais. Escrevemos as matrizes \widehat{A} e \widehat{B} em blocos,

$$\widehat{A} = \begin{pmatrix} \widehat{A}_{11} & \widehat{A}_{12} & \cdots & \widehat{A}_{1k} \\ \widehat{A}_{21} & \widehat{A}_{22} & \cdots & \widehat{A}_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{A}_{k1} & \widehat{A}_{k2} & \cdots & \widehat{A}_{kk} \end{pmatrix}_{k \times k}, \quad \widehat{B} = \begin{pmatrix} \widehat{B}_1 \\ \widehat{B}_2 \\ \vdots \\ \widehat{B}_k \end{pmatrix}_{k \times 1},$$

onde \widehat{A}_{ij} têm ordem $n_i \times n_j$, e \widehat{B}_i têm ordem $n_i \times k$, para $i, j = 1, \dots, k$.

Seja C a matriz invertível definida em (5.3-45). Primeiramente vamos obter \widehat{B} . Para isso aplicamos C nas colunas de B . O resultado é a matriz:

$$C \cdot B = \begin{pmatrix} p_{n_1} b_1 & p_{n_1} b_2 & \cdots & p_{n_1} b_k \\ p_{n_1} A b_1 & p_{n_1} A b_2 & \cdots & p_{n_1} A b_k \\ \vdots & \vdots & \cdots & \vdots \\ p_{n_1} A^{n_1-1} b_1 & p_{n_1} A^{n_1-1} b_2 & \cdots & p_{n_1} A^{n_1-1} b_k \\ \cdot & p_{n_1+n_2} b_2 & \cdots & \cdot \\ \cdot & p_{n_1+n_2} A b_2 & \cdots & \cdot \\ \cdot & \vdots & \cdots & \cdot \\ \cdot & p_{n_1+n_2} A^{n_2-1} b_2 & \cdots & \cdot \\ \cdot & \vdots & \cdots & \cdot \\ \cdot & \vdots & \cdots & \cdot \\ p_n b_1 & p_n b_2 & \cdots & p_n b_k \\ p_n A b_1 & p_n A b_2 & \cdots & p_n A b_k \\ \vdots & \vdots & \cdots & \vdots \\ p_n A^{n_k-1} b_1 & p_n A^{n_k-1} b_2 & \cdots & p_n A^{n_k-1} b_k \end{pmatrix}_{n \times k}.$$

Pelo Lema 5.11, esta matriz torna-se:

$$C \cdot B = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}_{n \times k} = \widehat{B}. \quad (5.3-51)$$

Para encontrar \widehat{A} , vamos calcular $\widehat{A} \cdot C = CA$. Esta igualdade é precisamente,

$$\begin{pmatrix} \widehat{A}_{11} & \widehat{A}_{12} & \cdots & \widehat{A}_{1k} \\ \widehat{A}_{21} & \widehat{A}_{22} & \cdots & \widehat{A}_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{A}_{k1} & \widehat{A}_{k2} & \cdots & \widehat{A}_{kk} \end{pmatrix}_{k \times k} \cdot \begin{pmatrix} p_{n_1} \\ p_{n_1}A \\ \vdots \\ p_{n_1}A^{n_1-1} \\ p_{n_1+n_2} \\ p_{n_1+n_2}A \\ \vdots \\ p_{n_1+n_2}A^{n_2-1} \\ \cdot \\ \vdots \\ \cdot \\ p_n \\ p_nA \\ \vdots \\ p_nA^{n_k-1} \end{pmatrix}_{n \times 1} = \begin{pmatrix} p_{n_1} \\ p_{n_1}A \\ \vdots \\ p_{n_1}A^{n_1-1} \\ p_{n_1+n_2} \\ p_{n_1+n_2}A \\ \vdots \\ p_{n_1+n_2}A^{n_2-1} \\ \cdot \\ \vdots \\ \cdot \\ p_n \\ p_nA \\ \vdots \\ p_nA^{n_k-1} \end{pmatrix}_{n \times 1} \cdot A,$$

multiplicando os blocos de tamanhos adequados, obtemos:

$$\left(\begin{array}{c} \widehat{A}_{11} \begin{pmatrix} p_{n_1} \\ p_{n_1}A \\ \vdots \\ p_{n_1}A^{n_1-1} \end{pmatrix} + \cdots + \widehat{A}_{1k} \begin{pmatrix} p_n \\ p_nA \\ \vdots \\ p_nA^{n_k-1} \end{pmatrix} \\ \widehat{A}_{21} \begin{pmatrix} p_{n_1} \\ p_{n_1}A \\ \vdots \\ p_{n_1}A^{n_1-1} \end{pmatrix} + \cdots + \widehat{A}_{2k} \begin{pmatrix} p_n \\ p_nA \\ \vdots \\ p_nA^{n_k-1} \end{pmatrix} \\ \vdots \\ \widehat{A}_{k1} \begin{pmatrix} p_{n_1} \\ p_{n_1}A \\ \vdots \\ p_{n_1}A^{n_1-1} \end{pmatrix} + \cdots + \widehat{A}_{kk} \begin{pmatrix} p_n \\ p_nA \\ \vdots \\ p_nA^{n_k-1} \end{pmatrix} \end{array} \right)_{k \times 1} = \begin{pmatrix} p_{n_1}A \\ p_{n_1}A^2 \\ \vdots \\ p_{n_1}A^{n_1} \\ p_{n_1+n_2}A \\ p_{n_1+n_2}A^2 \\ \vdots \\ p_{n_1+n_2}A^{n_2} \\ \vdots \\ p_nA \\ p_nA^2 \\ \vdots \\ p_nA^{n_k} \end{pmatrix}_{k \times 1}$$

Desenvolvendo os cálculos (multiplicação e soma) na matriz à esquerda e comparando cada bloco com o bloco correspondente na matriz à direita, obteremos os coeficientes de cada bloco da matriz \widehat{A} , que são

$$\left(\begin{array}{cccccc} 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0 \\ \alpha_{n_1 1}^{11} & \alpha_{n_1 2}^{11} & \cdots & \cdots & \alpha_{n_1 n_1}^{11} & \alpha_{n_1 1}^{1k} & \alpha_{n_1 2}^{1k} & \cdots & \cdots & \alpha_{n_1 n_k}^{1k} \\ & & \vdots & & \ddots & & & \vdots & & \\ & & \vdots & & & & & \vdots & & \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 1 \\ \alpha_{n_k 1}^{k1} & \alpha_{n_k 2}^{k1} & \cdots & \cdots & \alpha_{n_k n_1}^{k1} & \alpha_{n_k 1}^{kk} & \alpha_{n_k 2}^{kk} & \cdots & \cdots & \alpha_{n_k n_k}^{kk} \end{array} \right) \tag{5.3-52}$$

onde $\alpha^{ij} \in \hat{A}_{ij}$. Esta forma (5.3-52) é a forma normal para sistemas (5.3-46), (S-equivalente), que por sua vez é equivalente a forma (5.3-27). De fato, para qualquer

$$x = (x_1^1, \dots, x_{n_1}^1, x_1^2, \dots, x_{n_2}^2, \dots, x_1^k, \dots, x_{n_k}^k) \in \mathbb{R}^n,$$

ou escrito na forma

$$x = (x^1, x^2, \dots, x^k) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \dots \times \mathbb{R}^{n_k}, \quad (5.3-53)$$

e $u = (u_1, \dots, u_k) \in \mathbb{R}^k$, temos que $\hat{A}x + \hat{B}u$ é, para cada $i = 1, \dots, k$

$$\left\{ \begin{array}{l} \dot{x}_1^i = x_2^i \\ \dot{x}_2^i = x_3^i \\ \vdots \\ \dot{x}_{n_i-1}^i = x_{n_i}^i \\ \dot{x}_{n_i}^i = \sum_{j=1}^{n_1} \alpha_{n_i j}^{i1} x_j^1 + \sum_{j=1}^{n_2} \alpha_{n_i j}^{i2} x_j^2 + \dots + \sum_{j=1}^{n_k} \alpha_{n_i j}^{ik} x_j^k + u_i, \end{array} \right.$$

que está na forma (5.3-47). Considerando novos parâmetros de controle:

$$\hat{u}_i = \sum_{j=1}^{n_1} \alpha_{n_i j}^{i1} x_j^1 + \sum_{j=1}^{n_2} \alpha_{n_i j}^{i2} x_j^2 + \dots + \sum_{j=1}^{n_k} \alpha_{n_i j}^{ik} x_j^k + u_i,$$

obtemos, para cada $i = 1, \dots, k$, o sistema

$$\left\{ \begin{array}{l} \dot{x}_1^i = x_2^i \\ \dot{x}_2^i = x_3^i \\ \vdots \\ \dot{x}_{n_i-1}^i = x_{n_i}^i \\ \dot{x}_{n_i}^i = \hat{u}_i, \end{array} \right.$$

Fazendo $x^i = x_1^i$, podemos juntar os k sistemas em um só na forma mais compacta

$$\left\{ \begin{array}{l} x^{1(n_1)} = \hat{u}_1 \\ x^{2(n_2)} = \hat{u}_2 \\ \vdots \\ x^{k-1(n_{k-1})} = \hat{u}_{u_{k-1}} \\ x^{k(n_k)} = \hat{u}_k. \end{array} \right. \quad (5.3-54)$$

Logo, o sistema (5.3-46) é state-feedback equivalente ao sistema (5.3-54).



CAPÍTULO 6

LINEARIZAÇÃO POR STATE-FEEDBACK EQUIVALÊNCIA

O objetivo principal deste capítulo é introduzir um teorema que nos dá condições para a linearizabilidade local via state-feedback equivalência de um sistema de controle afim sobre uma variedade M no caso de controle vetorial. Iremos apenas enunciar este resultado na sua forma mais geral, provando-o apenas para sistemas com controle escalar. Para isso, iremos apresentar uma série de definições, incluindo também o Teorema de Fröbenius, visando o nosso objetivo.

Considere o sistema de controle não linear:

$$\dot{q} = f(q) + \sum_{i=1}^k u_i g_i(q), \quad u = (u_1, \dots, u_k) \in \mathbb{R}^k, \quad q \in M. \quad (6.0-1)$$

Nosso objetivo é verificar quando é possível transformar o sistema acima no sistema:

$$\dot{x} = Ax + \sum_{i=1}^k u_i b_i, \quad u = (u_1, \dots, u_k) \in \mathbb{R}^k, \quad x \in \mathbb{R}^n, \quad (6.0-2)$$

$$\text{span}\{A^j b_i \mid j = 1, \dots, n-1, i = 1, \dots, k\} = \mathbb{R}^n.$$

Apresentaremos a seguir algumas definições que serão fundamentais para o desenvolvimento deste capítulo.

Definição 6.1 *Seja TM o fibrado tangente. Uma **distribuição** $\Delta \subset TM$ sobre uma variedade M é uma família de subespaços vetoriais $\Delta_q \subset T_qM$ diferenciável em relação a $q \in M$, com dimensão $\dim \Delta_q$ constante para todo q .*

Observação 6.2 *Na definição anterior, o significado de uma distribuição ser diferenciável é o seguinte: para cada $q \in M$ existem uma vizinhança O de q e campos de vetores diferenciáveis X_1, \dots, X_k , tais que $X_i(q') \in \Delta(q')$ para todo $q' \in O$, $i = 1, \dots, k$, e $\Delta(q') = \text{span}\{X_1(q'), \dots, X_k(q')\}$. Em geral pode-se também considerar distribuições que não sejam diferenciáveis ou tais que $\dim \Delta_q$ dependam de q . As distribuições como as consideradas aqui, com $\dim \Delta_q$ constantes, são chamadas na literatura de distribuições regulares. Como estamos interessados somente em distribuições regulares e diferenciáveis, continuaremos a usar somente a expressão distribuição, ficando sempre subentendido que são regulares e diferenciáveis.*

Definição 6.3 *Dizemos que uma distribuição Δ é **integrável** se para cada $q \in M$ existe uma subvariedade imersa $W_q \subset M$, $q \in W_q$ tal que*

$$T_{q'}W_q = \Delta_{q'} \quad q' \in W_q.$$

A subvariedade W_q é chamada de **variedade integral** da distribuição Δ através do ponto q .

Definição 6.4 *Dizemos que uma distribuição Δ é **involutiva** se para quaisquer campos de vetores $X, Y \in \text{Vec}(M)$ e qualquer $q \in M$ tem-se:*

$$X(q), Y(q) \in \Delta_q \quad \Rightarrow \quad [X, Y](q) \in \Delta_q.$$

A implicação acima é também conhecida como **condição de Fröbenius**.

Enunciaremos aqui o importante teorema sobre uma distribuição, cuja demonstração pode ser encontrada em [1], Teorema 5.4.

Teorema 6.5 (Fröbenius) *Uma distribuição Δ sobre M é involutiva se, e somente se, é integrável.*

Definição 6.6 *Dizemos que o sistema (6.0-1) é **localmente state-feedback equivalente** ao sistema linear (6.0-2) na vizinhança de um ponto $q_0 \in M$, se existir um difeomorfismo local (S -transformação):*

$$\Phi : U_{q_0} \subset M \rightarrow V \subset \mathbb{R}^n$$

onde U_{q_0} é um aberto contendo o ponto q_0 , e uma F -transformação:

$$\varphi : U_{q_0} \times \mathbb{R}^k \rightarrow \mathbb{R}^k$$

da forma

$$\varphi(q, u) = \begin{pmatrix} a_1(q) \\ \vdots \\ a_k(q) \end{pmatrix} + D(q)u, \quad (6.0-3)$$

onde D é o operador invertível análogo ao definido em (5.3-22), ou seja,

$$D(q) : \text{span}\{g_1(q), \dots, g_k(q)\} \rightarrow \text{span}\{g_1(q), \dots, g_k(q)\},$$

cuja matriz de $D(q)$ na base $\mathcal{B} = \{g_1(q), \dots, g_k(q)\}$,

$$[D(q)]_{\mathcal{B}} = (d_{ij}(q))$$

é invertível com, $d_{ij}, a_i \in C^\infty(M)$ e $u = (u_1, \dots, u_k) \in \mathbb{R}^k$ para todo $i, j = 1, \dots, k$, que transformam o sistema (6.0-1) restrito a U_{q_0} no sistema (6.0-2) restrito a V .

Observação 6.7 Podemos escrever a F -transformação acima em detalhes:

$$\varphi(q, u) = \begin{pmatrix} \varphi_1(q, u) \\ \vdots \\ \varphi_k(q, u) \end{pmatrix} = \begin{pmatrix} a_1(q) \\ \vdots \\ a_k(q) \end{pmatrix} + \begin{pmatrix} d_{11}(q) & \cdots & d_{1k}(q) \\ \vdots & \ddots & \vdots \\ d_{k1}(q) & \cdots & d_{kk}(q) \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_k \end{pmatrix}$$

ou seja,

$$\begin{pmatrix} \varphi_1(q, u) \\ \vdots \\ \varphi_k(q, u) \end{pmatrix} = \begin{pmatrix} a_1(q) + \sum_{j=1}^k d_{1j}(q)u_j \\ \vdots \\ a_k(q) + \sum_{j=1}^k d_{kj}(q)u_j \end{pmatrix}.$$

Logo, para cada $i = 1, \dots, k$ temos,

$$\varphi_i(q, u) = a_i(q) + \sum_{j=1}^k d_{ij}(q)u_j \in \mathbb{R}.$$

Aplicando no sistema (6.0-1) obtemos,

$$\begin{aligned}
 \dot{q} &= f(q) + \sum_{i=1}^k u_i g_i(q) \longrightarrow \dot{q} = f(q) + \sum_{i=1}^k \varphi_i g_i(q) \\
 &= f(q) + \sum_{i=1}^k \left(a_i(q) + \sum_{j=1}^k d_{ij}(q) u_j \right) g_i(q) \\
 &= f(q) + \sum_{i=1}^k a_i(q) g_i(q) + \sum_{i,j=1}^k d_{ij}(q) u_j g_i(q) \\
 &= \left(f(q) + \sum_{i=1}^k a_i(q) g_i(q) \right) + u_i \sum_{i=1}^k D_i(q) g_i(q).
 \end{aligned}$$

Assim, podemos separar uma F-transformação em dois tipos:

$$(f, g_1, \dots, g_k) \longrightarrow (f + a_i g_i, g_1, \dots, g_k)$$

$$(f, g_1, \dots, g_k) \longrightarrow (f, Dg_1, \dots, Dg_k).$$

Exemplo 6.8 (Pêndulo) Considere um pêndulo consistindo de um comprimento l , massa m e com controle de torque u . Veja a figura abaixo.

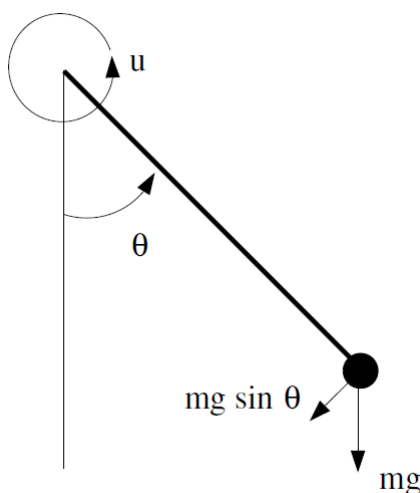


Figura 6.1: Pêndulo

A evolução do pêndulo é descrita pela equação de Euler-Lagrange com força externa

$$ml^2\ddot{\theta} + mgl \sin(\theta) = u.$$

Reescrevemos a equação acima como

$$\begin{aligned}\dot{\theta} &= \omega \\ \dot{\omega} &= -\frac{g}{l} \sin(\theta) + \frac{u}{ml^2}.\end{aligned}$$

Denote $x_1 := \theta$, $x_2 := \omega$ e considere a evolução do pêndulo sobre o espaço \mathbb{R}^2 , ou seja, $x = (x_1, x_2) \in \mathbb{R}^2$. Obtemos a sistema

$$\Sigma : \begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -\frac{g}{l} \sin(x_1) + \frac{u}{ml^2}. \end{cases}$$

Substituindo o controle u por

$$u = ml^2 \hat{u} + mlg \sin(x_1),$$

que pode ser interpretado como uma transformação no espaço dos parâmetros de controle \mathcal{U} com dependência do estado $x \in \mathbb{R}^2$ obtemos o sistema de controle linear

$$\hat{\Sigma} : \begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = \hat{u}. \end{cases}$$

Para colocar esse exemplo na notação do exemplo anterior, tome $q = (x_1, x_2)$, $f(q) = (x_2, -\frac{g}{l} \sin(x_1))$ e $g(q) = (0, \frac{1}{ml^2})$. Assim, o sistema afim fica na forma $\dot{q} = f(q) + ug(q)$. A F-transformação aplicada é dada por $\varphi : \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}$,

$$\varphi(x, y) = mlg \sin(x_1) + ml^2 u,$$

a qual transforma o sistema afim no sistema linear $\dot{x} = Ax + ub$, com

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad e \quad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Já definimos anteriormente, para sistemas da forma (6.0-2), a família de subespaços

$$D^m = \text{span}\{A^j b_i \mid j = 0, \dots, m-1, i = 1, \dots, k\}, \quad m = 1, \dots, n. \quad (6.0-4)$$

Esta família têm as seguintes propriedades:

1. Para cada $m = 1, \dots, n$ temos que $\dim D^m$ é constante para todo $x \in \mathbb{R}^n$. Logo, D^1, \dots, D^n são distribuições.
2. As distribuições D^m são integráveis pois, são geradas por campos de vetores constantes.
3. Devido a condição de controlabilidade, temos que $D^n = \mathbb{R}^n$.

Podemos generalizar a construção dos subespaços (6.0-4) para sistemas não lineares: para cada $m = 1, \dots, n$, definimos a família de subespaços

$$D_q^m = \text{span}\{(\text{ad} f)^j g_i(q) \mid j = 0, \dots, m-1, i = 1, \dots, k\} \subset T_q M. \quad (6.0-5)$$

Temos naturalmente que

$$D_q^1 \subset D_q^2 \subset \dots \subset D_q^n.$$

Mais ainda,

$$\begin{aligned} D^1 &= \text{span}\{g_1, \dots, g_k\} \\ D^2 &= \text{span}\{[f, g_1], \dots, [f, g_k]\} + D^1 \\ &\vdots \\ D^n &= \text{span}\{[f, X] \mid X \in D^{n-1}\} + D^{n-1}. \end{aligned}$$

Assim,

$$D^m = \text{span}\{[f, X] \mid X \in D^{m-1}\} + D^{m-1}, \quad m = 1, \dots, n. \quad (6.0-6)$$

Em geral, $\dim D_q^m$ é constante numa vizinhança suficientemente pequena do ponto q . De fato, podemos extrair uma base do conjunto de vetores geradores de D_q^m , sendo linearmente independentes no ponto q , são também em uma vizinhança do ponto q . Porém, não há garantia que tais vetores sejam linearmente independentes globalmente, conseqüentemente, $\dim D_q^m$ não é constante para cada m . Portanto, D_q^m não são distribuições.

A condição para a linearização de sistemas de controle não linear será dada em termos da família D_q^m . Estas famílias, com uma condição adicional, têm a propriedade de serem F-invariantes, isto é, para cada m , D_q^m contém as transformações por feedback de seus elementos. Isto é descrito no lema a seguir.

Lema 6.9 *Se as famílias D^m , $m = 1, \dots, n$, são involutivas, então elas são F-invariantes.*

Demonstração: Considere a decomposição da F-transformação vista na Observação 6.7:

$$(f, g_1, \dots, g_k) \longrightarrow (f + a_i g_i, g_1, \dots, g_k), \quad (6.0-7)$$

$$(f, g_1, \dots, g_k) \longrightarrow (f, Dg_1, \dots, Dg_k). \quad (6.0-8)$$

A prova será feita por indução sobre m . Para $m = 1$ temos,

$$D^1 = \text{span}\{g_1, \dots, g_k\}.$$

Como a transformação (6.0-7) não altera os campos g_1, \dots, g_k , temos que D^1 é invariante por esta transformação trivialmente. Como $D : D^1 \rightarrow D^1$ é invertível, então $Dg_i \in D^1$, $i = 1, \dots, k$. Logo, D^1 é também preservado pela transformação (6.0-8) e, portanto, D^1 é F-invariante.

Suponha que D^{m-1} é F-invariante e vamos provar que D^m também o é. Pela igualdade (6.0-6) temos,

$$D^m = \text{span}\{[f, X] \mid X \in D^{m-1}\} + D^{m-1}, \quad m = 1, \dots, n.$$

- D^m é preservado por (6.0-8). De fato, $Dg_i \in D^1 \subset D^{m-1}$, $i = 1, \dots, k$. Logo,

$$[f, Dg_i] \in D^m, \quad i = 1, \dots, k.$$

- D^m é preservado por (6.0-7). De fato, para $X \in D^{m-1}$ temos pela regra de Leibniz:

$$\begin{aligned} [f + a_i g_i, X] &= [f, X] + [a_i g_i, X] \\ &= [f, X] - [X, a_i g_i] \\ &= [f, X] - (X a_i) g_i - a_i [X, g_i]. \end{aligned}$$

Agora, como $X \in D^{m-1}$ temos que $[f, X] \in D^m$. Uma vez que $X a_i$ é um escalar, naturalmente $(X a_i) g_i \in D^1 \subset D^m$. Por fim, como g_i e X pertencem a D^{m-1} , segue da involutividade que:

$$[X, g_i] \in D^{m-1} \subset D^m.$$

Logo,

$$[f + a_i g_i, X] \in D^m.$$

Assim, D^m é preservado pelas F-transformações (6.0-7) e (6.0-8). Portanto D^m é F-invariante. ■

Enunciaremos sem demonstrar um resultado geral sobre state-feedback equivalência para sistemas afins. Posteriormente iremos demonstrar um resultado análogo para sistemas com controle escalar (Teorema 6.14).

Teorema 6.10 *O sistema (6.0-1) é localmente state-feedback equivalente ao sistema (6.0-2) se, e somente se,*

(i) *para cada $m = 1 \dots, n$ temos que, D^m são distribuições,*

(ii) $D_q^n = T_q M$, $q \in M$,

(iii) *as distribuições D^m são involutivas.*

Exemplo 6.11 *Considere o sistema não linear em \mathbb{R}^5 :*

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \end{pmatrix} = \underbrace{\begin{pmatrix} x_2 + x_2^2 \\ x_3 - x_1 x_4 + x_4 x_5 \\ x_2 x_4 + x_1 x_5 - x_5^2 \\ x_5 \\ x_2^2 \end{pmatrix}}_{=f(x,u)} + \underbrace{\begin{pmatrix} 0 \\ 0 \\ \cos(x_1 - x_5) \\ 0 \\ 0 \end{pmatrix}}_{=g_1(x,u)} u_1 + \underbrace{\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}}_{g_2(x,u)} u_2. \quad (6.0-9)$$

Neste sistema, a distribuição $D^1 = \text{span}\{g_1, g_2\}$ têm dimensão constante igual a 2 em uma vizinhança suficientemente pequena $O_0 \subset \mathbb{R}^5$ contendo a origem. Além disso, temos

$$[g_1, g_2](x) = 0 \in D^1,$$

para todo $x \in O_0$. Logo, D^1 é involutiva. Para obter a distribuição D^2 , usamos a fórmula da

Proposição 1.32 no Capítulo 1 para o cálculo de colchetes e obtemos

$$[f, g_1] = \begin{pmatrix} 0 \\ -\cos(x_1 - x_5) \\ -x_2 \sin(x_1 - x_5) \\ 0 \\ 0 \end{pmatrix}, \quad [f, g_2] = \begin{pmatrix} 0 \\ -1 \\ -(x_1 - x_5) \\ -1 \\ 0 \end{pmatrix}.$$

Logo $D^2 = \text{span}\{g_1, g_2, (\text{ad} f)g_1, (\text{ad} f)g_2\}$ têm dimensão constante igual a 4 em O_0 . Além disso,

$$[g_1, [f, g_1]] = [g_1, [f, g_2]] = [g_2, [f, g_1]] = [g_2, [f, g_2]] = 0 \in D^2,$$

para todo $x \in O_0$. Temos também

$$[[f, g_1], [f, g_2]](x) = \sin(x_1 - x_5)g_1(x) \in D^2.$$

Logo, D^2 é involutiva. Finalmente, com cálculos análogos, obtemos a distribuição

$$D^3 = \text{span}\{g_1, g_2, (\text{ad} f)g_1, (\text{ad} f)g_2, (\text{ad} f)^2 g_1, (\text{ad} f)^2 g_2\},$$

que tem dimensão constante igual a 5 para todo $x \in O_0$ e é involutiva. Como $D^{m-1} \subset D^m$ para $m \geq 2$ e D^3 tem dimensão máxima, temos que D^4 e D^5 são necessariamente involutivas. Portanto, D^m , $m = 1, \dots, 5$, possui dimensão constante em O_0 . Além disso, essas distribuições são involutivas e claramente $D_0^5 = T_0\mathbb{R}^5$. Pelo Teorema 6.10, o sistema (6.0-9) é state-feedback equivalente a um sistema linear controlável em O_0 .

Exemplo 6.12 Considere o sistema não linear sobre $\mathbb{R}^2 \times S^1$:

$$\begin{cases} \dot{x}_1 = u_1 \cos(\theta) \\ \dot{x}_2 = u_2 \sin(\theta) \\ \dot{\theta} = u_2, \end{cases} \quad (6.0-10)$$

com $((x_1, x_2), \theta) \in \mathbb{R}^2 \times S^1$. O sistema acima na forma de matrizes é:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{\theta} \end{pmatrix} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \\ 0 \end{pmatrix} u_1 + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} u_2,$$

que nos dá

$$g_1 = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \\ 0 \end{pmatrix}, \quad \text{e} \quad g_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

A distribuição $D^1 = \text{span}\{g_1, g_2\}$ não é involutiva pois,

$$[g_1, g_2] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \\ 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & -\sin(\theta) \\ 0 & 0 & \cos(\theta) \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Logo,

$$[g_1, g_2] = \begin{pmatrix} \sin(\theta) \\ -\cos(\theta) \\ 0 \end{pmatrix} \notin D^1.$$

Pelo Teorema 6.10, o sistema (6.0-10) não é localmente state-feedback equivalente a sistema linear controlável.

Para provarmos o resultado sobre state-feedback equivalência para controle escalar, precisaremos do seguinte resultado de natureza geral, que diz que variedades integráveis podem ser parametrizadas de maneira diferenciável.

Lema 6.13 *Sejam M uma variedade de dimensão n e $X_1, \dots, X_k \in \text{Vec}(M)$, $k \leq n$. Se $\Delta = \text{span}\{X_1, \dots, X_k\}$ é uma distribuição integrável sobre M , $\dim \Delta_q = k$, então para cada $q_0 \in M$ existem, uma vizinhança $O_{q_0} \subset M$ contendo q_0 e uma aplicação diferenciável*

$$\varphi : O_{q_0} \rightarrow \mathbb{R}^{n-k}$$

tal que:

$$(1) \dim \varphi_{*q}(T_q M) = n - k, \quad q \in O_{q_0},$$

(2) para todo $y \in \varphi(O_{q_0})$, $\varphi^{-1}(y)$ é uma variedade integral de Δ , ou, equivalentemente,

$$(2') \ker(\varphi_{*q}) = \Delta_q, \quad q \in O_{q_0}.$$

Demonstração: Seja $q_0 \in M$ um ponto qualquer. Escolha campos de vetores $Y_1, \dots, Y_{n-k} \in \text{Vec}(M)$ tais que

$$\text{span}\{X_1(q_0), \dots, X_k(q_0), Y_1(q_0), \dots, Y_{n-k}(q_0)\} = T_{q_0}M.$$

Da igualdade acima segue que

$$\text{span}\{X_1(q), \dots, X_k(q), Y_1(q), \dots, Y_{n-k}(q)\} = T_qM, \quad q \in U_{q_0},$$

onde U_{q_0} é uma vizinhança suficientemente pequena de q_0 . Considere a seguinte aplicação diferenciável:

$$\Psi : \mathbb{R}^k \times \mathbb{R}^{n-k} \rightarrow M$$

dada por

$$(t, s) \mapsto e^{t_1 X_1} \circ \dots \circ e^{t_k X_k} \circ e^{s_1 Y_1} \circ \dots \circ e^{s_{n-k} Y_{n-k}}(q_0),$$

com derivadas parciais na origem dadas por

$$\left. \frac{\partial}{\partial t_i} \right|_{t=s=0} \Psi(t, s) = X_i(q_0), \quad i = 1, \dots, k$$

$$\left. \frac{\partial}{\partial s_i} \right|_{t=s=0} \Psi(t, s) = Y_i(q_0), \quad i = 1, \dots, n-k.$$

Logo,

$$\Psi_{*(0,0)} : \mathbb{R}^k \times \mathbb{R}^{n-k} \rightarrow T_{q_0}M$$

é um isomorfismo e portanto, pelo Teorema da Função Inversa, existem abertos $O_{(0,0)} \subset \mathbb{R}^k \times \mathbb{R}^{n-k}$ e $O_{q_0} \subset U_{q_0}$ tais que

$$\Psi : O_{(0,0)} \rightarrow O_{q_0}$$

é um difeomorfismo. Seja

$$\begin{aligned} \pi_2 : \mathbb{R}^k \times \mathbb{R}^{n-k} &\rightarrow \mathbb{R}^{n-k} \\ (t, s) &\mapsto s \end{aligned}$$

a projeção na segunda coordenada. Com isso, definimos a aplicação diferenciável φ pondo:

$$\varphi := \pi_2 \circ \Psi^{-1} : O_{q_0} \rightarrow \mathbb{R}^{n-k}, \quad \varphi(t, s) = s.$$

Para $q \in O_{q_0}$, temos:

$$D_q \varphi = D_{(t_1, s_1)} \pi_2 \circ D_q \Psi^{-1} : T_q M \rightarrow \mathbb{R}^{n-k},$$

onde $\Psi^{-1}(q) = (t_1, s_1)$. Como $D_q \Psi^{-1}$ é isomorfismo e $D_{(t_1, s_1)} \pi_2$ é sobrejetora (pois π_2 é submersão), obtemos $\dim D_q \varphi(T_q M) = n - k$. Isto mostra o item (1).

Seja $s_0 = (s_1^0, \dots, s_{n-k}^0) \in \mathbb{R}^{n-k}$ fixado. O conjunto

$$N = \{\Psi_{s_0}(t) \mid t \in \mathbb{R}^k\}$$

é uma variedade integral de Δ . De fato, fixado s_0 , a aplicação

$$\Psi_{s_0} : \mathbb{R}^k \rightarrow M$$

é dada por

$$\Psi_{s_0}(t) = e^{t_1 X_1} \circ \dots \circ e^{t_k X_k}(q_1)$$

onde $q_1 = e^{s_1^0 Y_1} \circ \dots \circ e^{s_{n-k}^0 Y_{n-k}}(q_0)$. A diferencial, para $t \in \mathbb{R}^k$ qualquer,

$$D_t \Psi_{s_0} : \mathbb{R}^k \rightarrow T_{q'} M, \quad q' = \Psi_{s_0}(t)$$

é injetora. Logo Ψ_{s_0} é uma imersão. Consequentemente a imagem

$$N = \Psi_{s_0}(\mathbb{R}^k)$$

é uma subvariedade imersa de dimensão k em M . Como Δ é integrável, para cada $q \in N$ temos que

$$T_q N = \Delta_q$$

desta forma N é uma variedade integral de M . Observe que, para cada $s \in \varphi(O_{q_0})$ temos que $\varphi^{-1}(s) = N$. Em resumo, para qualquer ponto $q \in M$ podemos construir, usando a aplicação Ψ definida para q , uma variedade integral N_q contendo q , e como a

distribuição Δ é integrável, obtemos a igualdade acima. O item (2) segue. ■

O sistema (6.0-1) com $k = 1$, isto é, com controle escalar, têm a forma:

$$\dot{q} = f(q) + ug(q), \quad u \in \mathbb{R}, \quad q \in M. \quad (6.0-11)$$

Os subespaços D_q^m ficam definidos por:

$$D_q^m = \text{span}\{(\text{ad}f)^j g(q) \mid j = 0, \dots, m-1\} \subset T_q M, \quad q \in M, \quad m = 1, \dots, n.$$

Teorema 6.14 *O sistema (6.0-11) é localmente state-feedback equivalente ao sistema (5.3-32) se, e somente se,*

$$(1) D_q^n = T_q M,$$

(2) a distribuição D^{n-1} é involutiva.

Demonstração: (\Rightarrow) Segue da discussão que precede o Lema 6.9.

(\Leftarrow) Reciprocamente, vamos construir coordenadas adequadas que simplificarão o sistema para obter a transformação local desejada.

Seja $q_0 \in M$ qualquer. Como a distribuição D^{n-1} é integrável (Teorema de Frobenius), temos pelo Lema 6.13 que existem uma vizinhança $O_{q_0} \subset M$ de q_0 e uma aplicação diferenciável

$$\varphi_1 : O_{q_0} \rightarrow \mathbb{R}^{n-(n-1)} = \mathbb{R}$$

tal que:

$$(i) D_q \varphi_1(T_q M) = \mathbb{R} \quad \Rightarrow \quad D_q \varphi_1 \neq 0,$$

$$(ii) D_q \varphi_1(D_q^{n-1}) = 0.$$

Temos também,

$$(iii) D_q^{n-1} = \text{span}\{g(q), (\text{ad}f)g(q), \dots, (\text{ad}f)^{n-2}g(q)\},$$

$$(iv) T_q M = \text{span}\{D_q^{n-1}, (\text{ad}f)^{n-1}g(q)\}.$$

Considerando o campo f como derivação, construímos as seguintes funções definidas em O_{q_0} :

$$\begin{aligned}\varphi_2 &= f\varphi_1, \\ \varphi_3 &= f\varphi_2 = f^2\varphi_1, \\ &\vdots \\ \varphi_n &= f\varphi_{n-1} = f^{n-1}\varphi_1.\end{aligned}$$

Devido ao item (ii) temos, para cada $j = 0, \dots, n-2$,

$$0 = D_q\varphi_1(\mathbf{ad}f)^j g(q) = (\widehat{\mathbf{ad}f})^j g(q)\varphi_1(q).$$

Logo,

$$0 \neq D_q\varphi_1(\mathbf{ad}f)^{n-1}g(q) = (\widehat{\mathbf{ad}f})^{n-1}g(q)\varphi_1(q),$$

caso contrário $D_q\varphi_1(T_qM) = 0$, violando o item (i).

Com isso, provaremos que

$$(\mathbf{ad}f)^j g\varphi_l = \begin{cases} 0, & \text{se } j+l < n \\ \pm(\mathbf{ad}f)^{n-1}g\varphi_1 \neq 0, & \text{se } j+l = n, \end{cases} \quad (6.0-12)$$

onde $l = 1, \dots, n$ e $j = 0, \dots, n-1$. Denotamos $b = \pm(\mathbf{ad}f)^{n-1}g\varphi_1 \neq 0$.

Vamos provar por indução que a igualdade (6.0-12) é válida para todo $l \in \mathbb{N}$. De fato, se $l = 1$, já vimos anteriormente que

$$(\mathbf{ad}f)^{n-1}g\varphi_1 \neq 0 \quad \text{e} \quad (\mathbf{ad}f)^j g\varphi_1 = 0,$$

para $j = 0, \dots, n-2$. Suponha que (6.0-12) vale para $l-1$ e vamos mostrar que é

também válido para l . Com efeito,

$$\begin{aligned}
 (\text{ad} f)^j g \varphi_l &= (\text{ad} f)^j g (f \varphi_{l-1}) = ((\text{ad} f)^j g \circ f) \varphi_{l-1} \\
 &= ((\text{ad} f)^j g \circ f - f \circ (\text{ad} f)^j g + f \circ (\text{ad} f)^j g) \varphi_{l-1} \\
 &= ([(\text{ad} f)^j g, f] + f \circ (\text{ad} f)^j g) \varphi_{l-1} \\
 &= (-[f, (\text{ad} f)^j g] + f \circ (\text{ad} f)^j g) \varphi_{l-1} \\
 &= -(\text{ad} f)^{j+1} g \varphi_{l-1} + f \circ (\text{ad} f)^j g \varphi_{l-1} \\
 &= -(\text{ad} f)^{j+1} g \varphi_{l-1} + f \circ (\text{ad} f)^j g \varphi_{l-1}.
 \end{aligned}$$

Se $(j+1) + (l-1) = j+l = n$, então $j+l-1 < n$. Pela hipótese de indução, temos

$$-(\text{ad} f)^{j+1} g \varphi_{l-1} = -(\text{ad} f)^{n-1} g \varphi_1 \neq 0, \quad \text{e} \quad f \circ (\text{ad} f)^j g \varphi_{l-1} = 0,$$

e obtemos, $(\text{ad} f)^j g \varphi_l = -(\text{ad} f)^{j+1} g \varphi_{l-1}$.

Se $(j+1) + (l-1) = j+l < n$, então $j+l-1 < n$. Novamente por hipótese, temos

$$-(\text{ad} f)^{j+1} g \varphi_{l-1} = f \circ (\text{ad} f)^j g \varphi_{l-1} = 0,$$

e obtemos $(\text{ad} f)^j g \varphi_l = 0$. Portanto, a igualdade (6.0-12) é verdadeira para todo l .

Definimos a aplicação

$$\Phi = \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_n \end{pmatrix} : O_{q_0} \subset M \rightarrow \mathbb{R}^n.$$

As coordenadas de um ponto $q \in O_{q_0}$ ficam definidas como:

$$x_l = \varphi_l(q), \quad l = 1, \dots, n.$$

Escrevemos agora o sistema (6.0-11) com respeito a essas coordenadas:

$$\frac{d}{dt} x_l = \frac{d}{dt} \varphi_l(q(t)) = D_{q(t)} \varphi_l(\dot{q}(t)) = \dot{q}(t)(\varphi_l(q(t))) = (f + ug) \varphi_l = f \varphi_l + ug \varphi_l,$$

obtemos assim,

$$\frac{d}{dt} x_l = f \varphi_l + ug \varphi_l \quad l = 1, \dots, n. \quad (6.0-13)$$

Para $l < n$, a igualdade (6.0-12) nos dá, $g\varphi_l = (\text{ad}f)^0 g\varphi_l = 0$. Consequentemente a equação (6.0-13) acima fica:

$$\frac{d}{dt}x_l = f\varphi_l = \varphi_{l+1} = x_{l+1}, \quad l = 1, \dots, n-1. \quad (6.0-14)$$

Para $l = n$, a igualdade (6.0-12) nos dá,

$$g\varphi_n = (\text{ad}f)^0 g\varphi_n = \pm(\text{ad}f)^{n-1} g\varphi_1 = b \neq 0,$$

e a equação (6.0-13) torna-se,

$$\frac{d}{dt}x_n = f\varphi_n + ug\varphi_n = f\varphi_n \pm ub. \quad (6.0-15)$$

As equações (6.0-14) e (6.0-15) dão a forma do sistema (6.0-11) nas coordenadas x_1, \dots, x_n :

$$\left\{ \begin{array}{l} \dot{x}_1 = x_2 \\ \vdots \\ \dot{x}_{n-1} = x_n \\ \dot{x}_n = f\varphi_n \pm ub. \end{array} \right. \quad (6.0-16)$$

Finalmente, considerando a F-transformação:

$$u \mapsto \mp \frac{f\varphi_n - \hat{u}}{b},$$

obtemos,

$$\dot{x}_n = f\varphi_n \pm \left(\mp \frac{f\varphi_n - \hat{u}}{b} \right) b = \hat{u}.$$

Isto transforma (6.0-16) no sistema

$$\left\{ \begin{array}{l} \dot{x}_1 = x_2 \\ \vdots \\ \dot{x}_{n-1} = x_n \\ \dot{x}_n = \hat{u}. \end{array} \right. \quad (6.0-17)$$

Podemos escrever o sistema linear acima na forma $\dot{x} = Ax + \hat{u}b$ com

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \quad \text{e} \quad b = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Um cálculo direto mostra que $A^j b = e_{n-j}$, onde e_i denota o i -ésimo elemento do base canônica de \mathbb{R}^n . Portanto, o sistema linear obtido é controlável. ■

Exemplo 6.15 Considere o modelo de um braço mecânico acionado por um motor através de uma mola de torção.

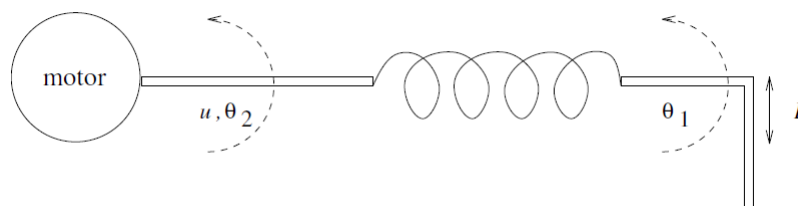


Figura 6.2: Braço mecânico

As equações diferenciais para as posições angulares θ_1 e θ_2 do braço e do eixo, respectivamente, são dadas por:

$$\begin{cases} I\ddot{\theta}_1 + mgL \sin(\theta_1) + k(\theta_1 - \theta_2) = 0 \\ J\ddot{\theta}_2 - k(\theta_1 - \theta_2) = u, \end{cases} \quad (6.0-18)$$

onde I e J são momentos de inércia, k é uma constante molar, m é a massa, L é a distancia da junta até o centro de massa do braço, e u é o controle de torque inicial aplicado pelo motor.

Sejam $x_1 := \theta_1$, $x_2 := \dot{\theta}_1$, $x_3 := \theta_2$ e $x_4 := \dot{\theta}_2$. Desta forma, o sistema (6.0-18) é considerado no espaço estado \mathbb{R}^4 . Por simplicidade escrevemos $a := mgL/I$, $b := k/I$, $c := k/J$, e

$d := 1/J$. O sistema (6.0-18) fica da seguinte forma:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -a \sin(x_1) - b(x_1 - x_3) \\ \dot{x}_3 = x_4 \\ \dot{x}_4 = c(x_1 - x_3) + du, \end{cases}$$

e escrito na forma estendida

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{pmatrix} = \underbrace{\begin{pmatrix} x_2 \\ -a \sin(x_1) - b(x_1 - x_3) \\ x_4 \\ c(x_1 - x_3) \end{pmatrix}}_{=f(x,u)} + \underbrace{\begin{pmatrix} 0 \\ 0 \\ 0 \\ d \end{pmatrix}}_{=g(x,u)} u.$$

Fazendo os cálculos obtemos,

$$[f, g](x, u) = \begin{pmatrix} 0 \\ 0 \\ -d \\ 0 \end{pmatrix}, \quad [f, [f, g]](x, u) = \begin{pmatrix} 0 \\ bd \\ 0 \\ -cd \end{pmatrix}, \quad [f, [f, [f, g]]](x, u) = \begin{pmatrix} -bd \\ 0 \\ cd \\ 0 \end{pmatrix}.$$

Segue que $D^4 = \text{span}\{g, (\text{ad}f)g, (\text{ad}f)^2g, (\text{ad}f)^3g\}$ tem dimensão constante igual a 4, pois os vetores obtidos são linearmente independentes. Finalmente, $D^3 = \text{span}\{g, (\text{ad}f)g, (\text{ad}f)^2g\}$ é uma distribuição involutiva. Pelo Teorema 6.14, o sistema (6.0-18) é localmente state-feedback equivalente a um sistema linear controlável.

Exemplo 6.16 Considere o sistema não linear em \mathbb{R}^4 :

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{pmatrix} = \underbrace{\begin{pmatrix} x_2 - x_3^2 \\ x_3 + 2x_1^2x_3 \\ x_1^2 \\ x_1 + x_3^2 \end{pmatrix}}_{=f(x,u)} + \underbrace{\begin{pmatrix} 0 \\ 2x_3 \\ 1 \\ 0 \end{pmatrix}}_{=g(x,u)} u. \quad (6.0-19)$$

Fazendo os cálculos, obtemos

$$D^3 = \text{span}\{g, (\text{ad}f)g, (\text{ad}f)^2g\},$$

onde $g = (0, 2x_3, 1, 0)^T$, $(\text{ad}f)g = (0, -1, 0, -2x_3)^T$ e $(\text{ad}f)^2g = (1, 0, 0, -2x_1^2)^T$. Porém,

$$\begin{aligned} [g, (\text{ad}f)g] &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 2x_3 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \\ 0 \\ -2x_3 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 0 \\ 0 \\ -2 \end{pmatrix}. \end{aligned}$$

Como $[g, (\text{ad}f)g]$ é linearmente independente em relação ao conjunto $\{g, (\text{ad}f)g, (\text{ad}f)^2g\}$ temos que, $[g, (\text{ad}f)g] \notin D^3$. Segue que D^3 não é involutiva. Portanto pelo Teorema 6.14, o sistema (6.0-19) não é state-feedback equivalente a um sistema linear.

Exemplo 6.17 Considere o sistema não linear em \mathbb{R}^3 :

$$\begin{cases} \dot{x} = y + yz \\ \dot{y} = z \\ \dot{z} = \sin(x) + u. \end{cases} \quad (6.0-20)$$

Assumimos $z \neq -1$ e denotaremos por $p = (x, y, z) \in \mathbb{R}^3$. Na forma estendida, temos

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = \underbrace{\begin{pmatrix} y + yz \\ z \\ \sin(x) \end{pmatrix}}_{=f(x,u)} + \underbrace{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}}_{=g(x,u)} u.$$

Fazendo os cálculos obtemos,

$$[f, g] = \begin{pmatrix} -y \\ -1 \\ 0 \end{pmatrix}, \quad [f, [f, g]] = \begin{pmatrix} 1+z \\ 0 \\ y \cos(x) \end{pmatrix}.$$

Por cálculos simples podemos mostrar que $\{g, (\text{ad}f)g, (\text{ad}f)^2g\}$ é um conjunto linearmente independente. Logo, $\dim \text{span}\{g, (\text{ad}f)g, (\text{ad}f)^2g\} = 3$. Finalmente, a distribuição $D^2 = \text{span}\{g, (\text{ad}f)g\}$ é involutiva. De fato,

$$[g, [f, g]] = [[f, g], g] = 0 \in D^2.$$

Pelo Teorema 6.14, o sistema (6.0-20) é localmente state-feedback equivalente a um sistema linear controlável. Pelo Teorema de Fröbenius, D^2 é uma distribuição integrável, e possui dimensão constante igual a 2. Para linearizar este sistema, vamos aplicar o Lema 6.13 para obter uma aplicação diferenciável

$$\varphi : O_p \subset \mathbb{R}^3 \rightarrow \mathbb{R}^{3-(2)} = \mathbb{R},$$

tal que $\varphi_{*p}(D_p^2) = 0$. Note que a distribuição $D^2 = \text{span}\{g, (\text{ad}f)g\}$ pode ser escrita como

$$D^2 = \text{span}\left\{ \frac{\partial}{\partial z}, y \frac{\partial}{\partial x} + \frac{\partial}{\partial y} \right\}.$$

Desta forma, obter uma aplicação diferenciável φ tal que $\varphi_{*p}(D_p^2) = 0$ equivale a encontrar a solução das equações diferenciais

$$\begin{cases} \frac{\partial \varphi}{\partial z} = 0 \\ y \frac{\partial \varphi}{\partial x} + \frac{\partial \varphi}{\partial y} = 0. \end{cases} \quad (6.0-21)$$

Neste caso, a solução é dada por $\varphi(x, y, z) = x - \frac{y^2}{2}$. Pondo $\tilde{x} = x - \frac{y^2}{2}$, definimos \tilde{y} e \tilde{z} recursivamente como:

$$\begin{aligned} \tilde{y} &= \dot{\tilde{x}} = \dot{x} - y\dot{y} = y + yz - yz = y \\ \tilde{z} &= \dot{\tilde{y}} = \dot{y} = z. \end{aligned}$$

Por fim, defina $\tilde{u} = u + \sin(x)$. Com isso, obtemos o sistema linear controlável em O_p :

$$\begin{cases} \dot{\tilde{x}} = \tilde{y} \\ \dot{\tilde{y}} = \tilde{z} \\ \dot{\tilde{z}} = \tilde{u}. \end{cases}$$

De outra forma,

$$\dot{\tilde{p}} = A(\tilde{p}) + b\tilde{u},$$

onde

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Com uma simples verificação, obtemos

$$\text{span}\{b, Ab, A^2b\} = \text{span}\{(0, 0, 1)^T, (0, 1, 0)^T, (1, 0, 0)^T\} = \mathbb{R}^3,$$

satisfazendo a condição de Kalman.

BIBLIOGRAFIA

- [1] A. A. Agrachev and Y. Sachkov. *Control theory from the geometric viewpoint*, volume 87. Springer Science, 2013.
- [2] P. Brunovský. A classification of linear controllable systems. *Kybernetika (Institute of Information Theory and Automation AS CR)*, 6(3):173–188, 1970.
- [3] J. A. N. Cossich. *Controlabilidade de sistemas de controle lineares (Dissertação de Mestrado)*. Universidade Estadual de Maringá, 2014.
- [4] W. de Melo. *Topologia das Variedades*. Instituto de Matemática Pura e Aplicada, 2014.
- [5] E. Fernández-Cara and E. Zuazua. Control theory: History, mathematical achievements and perspectives. *Boletín SEMA (Sociedad Española de Matemática Aplicada)*, 26(12):79–140, 2003.
- [6] J. M. Lee. *Introduction to Smooth Manifolds*. Springer Science Graduate Texts in Mathematics 218, 2002.
- [7] E. L. Lima. *Espaços métricos*, volume 4. Instituto de Matemática Pura e Aplicada, 1983.
- [8] E. L. Lima. *Variedades diferenciáveis*. Publicações Matemáticas - Instituto de Matemática Pura e Aplicada, 2011.
- [9] F. Manfio. *Variedades Diferenciáveis (Notas de Aula)*. ICMC – USP, 2011.

-
- [10] J. Neukirch. *Algebraic number theory*. Springer Science, 1999.
- [11] B. A. Rodrigues. *Classificação topológica de sistemas de controle lineares (Dissertação de Mestrado)*. Universidade Estadual de Maringá, 2014.
- [12] L. A. San Martin. *Grupos de Lie*. Editora da Unicamp, 2015.
- [13] J. Sotomayor. *Lições de equações diferenciais ordinárias*, volume 11. Instituto de Matemática Pura e Aplicada, CNPq, 1979.
- [14] F. W. Warner. *Foundations of differentiable manifolds and Lie groups*, volume 94. Springer Science & Business Media, 2013.