



UNIVERSIDADE ESTADUAL DA PARAÍBA
CENTRO DE CIÊNCIAS E TECNOLOGIA
DEPARTAMENTO DE MATEMÁTICA
CURSO DE LICENCIATURA EM MATEMÁTICA

TEORIA DO CONTROLE EM EQUAÇÕES
DIFERENCIAIS ORDINÁRIAS

JANDEILSON SANTOS DA SILVA

CAMPINA GRANDE

Dezembro de 2017

JANDEILSON SANTOS DA SILVA

**TEORIA DO CONTROLE EM EQUAÇÕES
DIFERENCIAIS ORDINÁRIAS**

Trabalho Acadêmico Orientado apresentado ao curso de Licenciatura em Matemática do Departamento de Matemática do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de licenciado em Matemática.

Orientador: Dr. Aldo Trajano Lourêdo

CAMPINA GRANDE

Dezembro de 2017

É expressamente proibido a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano do trabalho.

S586t Silva, Jandeilson Santos da.
Teoria do controle em equações diferenciais [manuscrito] :
/ Jandeilson Santos da Silva. - 2017.
114 p. : il. colorido.

Digitado.

Trabalho de Conclusão de Curso (Graduação em Matemática) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2017.

"Orientação : Prof. Dr. Aldo Trajano Lourêdo, Departamento de Matemática - CCT."

1. Teoria do controle. 2. Equações diferenciais ordinárias.
3. Controlabilidade. 4. Observabilidade.

21. ed. CDD 515.352

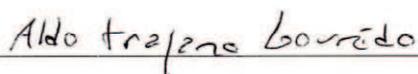
JANDEILSON SANTOS DA SILVA

TEORIA DO CONTROLE EM EQUAÇÕES
DIFERENCIAIS ORDINÁRIAS

Trabalho Acadêmico Orientado apresentado ao curso de Licenciatura e Matemática do Departamento de Matemática do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de licenciado em Matemática.

Aprovado em: 07 / 12 / 2017

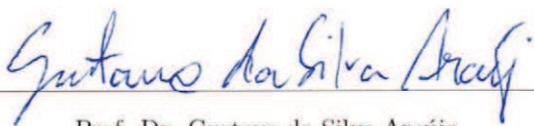
COMISSÃO EXAMINADORA



Prof. Dr. Aldo Trajano Lourêdo

Dpto. Matemática - CCT/UEPB

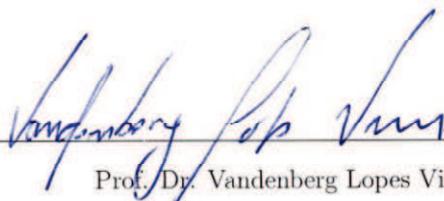
ORIENTADOR



Prof. Dr. Gustavo da Silva Araújo

Dpto. Matemática - CCT/UEPB

EXAMINADOR



Prof. Dr. Vandenberg Lopes Vieira

Dpto. Matemática - CCT/UEPB

EXAMINADOR

Dedicatória

Este trabalho é dedicado à minha família e em especial aos meus pais, João Batista Cândido da Silva e Tereza Bento dos Santos, por todo apoio, incentivo, direcionamento e educação a mim oferecidos, fatores externos pelos quais pude chegar até aqui. Dedico ainda à minha noiva, Silmara Moreira, por toda ajuda e incentivo na luta por mais este objetivo.

Agradecimentos

Estando no término de mais uma etapa, quero agradecer primeiramente ao Deus todo poderoso, que me concede constantemente força, sabedoria, oportunidades e perseverança no alcance dos meus objetivos. A Ele toda honra e glória.

Expresso aqui meu agradecimento à pessoas essenciais em toda esta caminhada, fundamentais para que eu tenha vencido inúmeros obstáculos chegando até aqui. Neste sentido, sou grato à toda minha família e de modo especial, ao meu pai João Batista e à minha mãe Tereza Santos por todo esforço e empenho na minha formação como cidadão, pelo incentivo e apoio nos estudos bem como nas minhas escolhas, dentre outras enormes parcelas de contribuições dos mesmos em todo sucesso que alcancei ou venha a alcançar. Meu muito obrigado pelo apoio, neste mesmo âmbito, às minhas irmãs Janyelle, Jandeilma, Jarlanny e Jardilane as quais estiveram sempre do meu lado e torcendo por minhas conquistas. Agradeço ainda à minha noiva Silmara Moreira por investir em mim tanto estímulo e motivação, sempre me apoiando e compreendendo.

Agradeço ao meu professor e orientador, Dr. Aldo Trajano Lourêdo, pela atenção, dedicação e compreensão a mim dirigidas, assim como pelos incentivos ao prosseguimento com os estudos e ao investimento em uma possível carreira acadêmica. Aproveito a oportunidade para agradecer a todos os professores do Curso de Licenciatura em Matemática que direta ou indiretamente contribuíram com minha formação, principalmente àqueles que estiveram a frente dos componentes curriculares por mim cursados. Em especial, registro meu agradecimento aos professores Vandenberg Lopes Vieira e José Elias da Silva (e a outros caso a memória tenha falhado) que sempre me incentivaram ao prosseguimento nos estudos apontando para uma possível pós-graduação, bem como aos professores Gustavo da Silva Araújo e Francisco Siberio Bezerra Albuquerque por aceitarem o convite para compor a banca examinadora deste trabalho.

Agradeço a todos os meus colegas de graduação que de uma forma ou outra con-

tribuíram com minha formação e no acúmulo de ricas experiências de convívio. Particularmente ao meu amigo Francinaldo Medeiros por todo apoio e colaboração ao longo do curso. Coloco ainda meu muito obrigado aos colegas de outras turmas que tive o prazer de conhecer e manter amizades, dos quais cito o amigo Cícero José.

Resumo

Neste trabalho estudamos conceitos relativos à Teoria do Controle em Equações Diferenciais Ordinárias, na qual sistemas de E.D.O são enxergados não como objetos estáticos, mas que podem ser conduzidos a estados desejáveis. Exploramos os conceitos de controlabilidade e observabilidade apresentando suas respectivas definições e estabelecendo resultados que caracterizam um sistema como controlável e/ou observável através de simples processos algébricos. Em seguida, algumas aplicações são exploradas evidenciando a riqueza e importância da temática.

Palavras-chave: Teoria do Controle, Sistemas de Equações Diferenciais Ordinárias, Controlabilidade e Observabilidade.

Abstract

In this work we study concepts related to the Control Theory in Ordinary Differential Equations, in which E.D.O systems are not seen as static objects, but which can be driven to desirable states. We explore the controllability and observability concepts presenting their respective definitions and establishing results that characterize a system as controllable and/or observable through simple algebraic processes. Next some applications are exploited, evidencing the richness and importance of the studied theory.

Keywords: Control Theory, Ordinary Differential Equations, Controllability and Observability.

Sumário

Lista de Símbolos	10
Introdução	12
1 Introdução à Teoria do Controle Linear	14
1.1 Controlabilidade	15
1.2 Observabilidade	43
2 Aplicações	62
2.1 O pêndulo invertido	62
2.2 Dinâmica de um satélite	76
2.2.1 Motivação	76
2.2.2 Modelagem Matemática	77
2.2.3 Controlabilidade e observabilidade do modelo linearizado	88
A Apêndice	99
A.1 Álgebra Linear e Teoria das Matrizes	99
A.2 Alguns resultados e definições sobre Derivadas, Integrais e Equações Diferenciais	107
Referências Bibliográficas	114

Lista de Símbolos

\mathbb{N}	O conjunto dos números naturais
\mathbb{N}_0	O conjunto dos números inteiros não-negativos: $\mathbb{N} \cup \{0\}$
\mathbb{Z}	O conjunto dos números inteiros
\mathbb{Q}	O conjunto dos números racionais
\mathbb{R}	O conjunto dos números reais
\mathbb{C}	O conjunto dos números complexos
\mathbb{R}_+	O conjunto dos números reais não negativos
\mathbb{C}_+	O semiplano complexo aberto à direita formado pelos 1º e 4º quadrantes do plano: $\{z \in \mathbb{C} ; \operatorname{Re}(z) > 0\}$
\mathbb{C}_-	O semiplano complexo aberto à esquerda formado pelos 2º e 3º quadrantes do plano: $\{z \in \mathbb{C} ; \operatorname{Re}(z) < 0\}$
\mathbb{F}	\mathbb{R} ou \mathbb{C}
\mathbb{F}^n	Espaço vetorial das n -uplas ordenadas de números de \mathbb{F}
$\mathbb{F}^{p \times n}$	Espaço vetorial das matrizes de ordem $p \times n$ com entradas em \mathbb{F} .
$GL(n, \mathbb{F})$	Grupo das matrizes invertíveis pertencentes a $\mathbb{F}^{n \times n}$ (grupo linear de ordem n).
*	(em sobrescrito) Hermitiano/Conjugado transposto de uma matriz
\oplus	Soma direta (de subespaços de \mathbb{F}^n)
S^\perp	Complemento ortogonal de um subespaço S do \mathbb{F}^n
Im	Imagem de um elemento pertencente a $\mathbb{F}^{n \times p}$
ker	Núcleo de um elemento de $\mathbb{F}^{n \times p}$

\det	Determinante de uma matriz quadrada
adj	Matriz adjunta de uma matriz quadrada
σ	Espectro de uma matriz quadrada. Dada uma matriz $A \in \mathbb{F}^{n \times n}$, o espectro de A é o conjunto dos seus autovalores.
$\ \cdot\ $	Símbolo genérico para uma norma
$\langle \cdot, \cdot \rangle$	Produto interno (sobre \mathbb{F}^n)
$PC(I, Y)$	Espaço vetorial das funções contínuas por partes $f : I \rightarrow Y$ onde $\mathbb{F}^{p \times n}$ e $I \subset \mathbb{R}$ um intervalo.
\star	Convolução de funções
\mathcal{L}	Transformada de Laplace

Introdução

Este manuscrito constitui-se em uma pesquisa bibliográfica tendo por principal objetivo um estudo introdutório à Teoria do Controle em sistemas de Equações Diferenciais Ordinárias. Em síntese, a preocupação está direcionada em conduzir um dado sistema a valores prescritos e desejáveis através da introdução de uma função, denominada *entrada* ou *controle*, no referido sistema.

É de certa forma difícil situar historicamente as origens da teoria em pauta uma vez que, não espantosamente, sempre foi desejável ao homem *controlar* determinadas variáveis de forma a mantê-las em valores fixos ou pelo menos *estáveis* nas proximidades destes. Podemos, entretanto, apontar primórdios de seu desenvolvimento moderno a partir do século XVII tendo origens em necessidades físicas como por exemplo, o problema de controlar a velocidade de relógios pendulares, proposto por Christiaan Huygens e Robert Hooke. Em todo o histórico seguinte, as necessidades desta teoria foram cada vez mais intensas surgindo em meios como a navegação, ainda no século mencionado e a utilização de moinhos de vento no século XVIII. Foi ainda neste último período que a Revolução Industrial, através da adaptação de James Watt, em 1769, de controladores de navios para máquinas a vapor, tornou os mecanismos de controle muito populares. No entanto, somente no século XIX, a Teoria do Controle teve seu primeiro tratamento puramente matemático proposto pelo físico escocês James Clerk Maxwell, através da publicação de um artigo seminal em 1868 intitulado “On Governors”. O trabalho de Maxwell impulsionou uma onda de pesquisa em Teoria do Controle com grandes avanços aqui omissos por uma questão de objetividade, mas que se mantiveram em expansão até o século XX.

Finalmente, durante a década de 1950 passaram a ser desenvolvidas técnicas mais gerais, sofisticadas e poderosas que permitiram o estudo de sistemas de várias variáveis bem como de sistemas não-lineares. Destacamos aí os trabalhos de L. Pontryagin (con-

trole ótimo não-linear) na União Soviética, Rudolf Kalman (filtragem, controle ótimo linear/quadrático e análise algébrica) e Richard Bellman (programação dinâmica), estes dois últimos nos Estados Unidos. Estas contribuições fomentaram uma enorme pesquisa na década seguinte, prosseguindo até os dias atuais.

O presente trabalho é um estudo sobre controle de sistemas lineares de Equações Diferenciais Ordinárias e está estruturado da seguinte maneira: no Capítulo 1, apresentamos os conceitos de *Controlabilidade* e *Observabilidade* expondo resultados que caracterizam algebricamente as definições de *sistema controlável* e *sistema observável* e ainda estabelecendo relações entre estes conceitos; no Capítulo 2 são estudadas duas aplicações físicas dos conhecimentos adquiridos no capítulo anterior e, por fim, no Apêndice A estão dispostos os resultados e definições advindos de diversos ramos da matemática e utilizados para estruturação formal da teoria que compõe o trabalho.

Capítulo 1

Introdução à Teoria do Controle Linear

Neste capítulo estudaremos sistemas com entrada u e saída y . Em particular, para $(0, x_0) \in J \times \mathbb{R}^n$, onde $J \subset \mathbb{R}$ é um intervalo contendo 0, estudaremos os sistemas da forma

$$\begin{cases} x' = Ax + Bu \\ y = Cx \end{cases} \quad (1.1)$$

com $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$ e $C \in \mathbb{R}^{p \times n}$. Pelo Teorema A.15 sabemos que para cada $u \in PC(J, \mathbb{R}^m)$ e $x_0 \in \mathbb{R}^n$, o sistema (1.1) possui solução única em J a qual denotaremos por $x(\cdot, x_0, u)$. Especificamente para cada $(x_0, u) \in \mathbb{R}^n \times PC(J, \mathbb{R}^m)$ a única solução de (1.1) em J é dada pela fórmula de variação dos parâmetros

$$x(\cdot, x_0, u) = e^{At}x_0 + \int_0^t e^{A(t-s)}Bu(s) ds, \quad \forall t \in J. \quad (1.2)$$

O valor $x(t, x_0, u) \in \mathbb{R}^n$ é o estado do sistema no momento t que corresponde ao estado inicial x_0 e entrada u . Por simplicidade, em situações em que os dados (x_0, u) são claros a partir do contexto, podemos escrever simplesmente $x(\cdot)$ ou x .

A saída $y(t, x_0, u) = Cx(t, x_0, u) = Cx(t)$ no instante t representa a informação sobre o estado atual do sistema disponível para o mundo exterior. Novamente, por simplicidade, podemos escrever $y(\cdot)$, desde que os dados (x_0, u) estejam claros.

O centro da Teoria da Controlabilidade abriga os seguintes dois problemas:

- a) **Problema de Controlabilidade:** Para um dado estado inicial (x_0) e através da escolha adequada de um controle (u) pode a evolução do estado de um sistema ser

determinada de modo a conduzi-lo a um estado alvo pré-especificado em um tempo finito?

- b) **Problema de Observabilidade:** É possível determinar o atual estado $x(t), t > 0$, a partir do conhecimento da entrada e da saída dados em $[0, t]$?

Estas questões serão discutidas nas próximas duas seções.

1.1 Controlabilidade

Apresentaremos agora as definições dos conceitos fundamentais de alcançabilidade e controlabilidade.

Definição 1.1 *Sejam $x_0, x_1 \in \mathbb{R}^n$. Se existe $T > 0$ e um controle $u \in PC([0, T], \mathbb{R}^m)$ tal que $x(T; x_0, u) = x_1$, então x_1 é dito ser alcançável a partir de x_0 . O sistema (1.1) é dito ser controlável se, quaisquer que sejam $x_0, x_1 \in \mathbb{R}^n$, x_1 é alcançável a partir de x_0 .*

Frequentemente no estudo da controlabilidade de (1.1), identificaremos tal sistema com o par de matrizes (A, B) . Esta notação deve-se ao fato de que, como se perceberá mais adiante, a controlabilidade do sistema (1.1) depende apenas das matrizes A e B . Neste contexto, é conveniente definir um par de matrizes $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$ como sendo controlável se o sistema (1.1) associado é controlável de acordo com a Definição 1.1. No que segue, a matriz de controlabilidade (também chamada matriz de alcançabilidade)

$$C(A, B) := (B, AB, A^2B, \dots, A^{n-1}B) \in \mathbb{R}^{n \times (mn)}$$

e o Gramian de controlabilidade (parametrizado por $T > 0$)

$$Q_T := \int_0^T e^{At} B B^* e^{A^*t} dt \in \mathbb{R}^{n \times n}$$

irão desempenhar papéis importantes na caracterização da controlabilidade de um sistema, uma vez que constituem-se elementos fundamentais em poderosas ferramentas algébricas para tal caracterização (Como veremos no Teorema 1.3, a controlabilidade do par (A, B) depende diretamente do posto da matriz controlabilidade e da invertibilidade do Gramian). Note que Q_T é simétrica, pois

$$\begin{aligned} Q_T^* &= \left(\int_0^T e^{At} B B^* e^{A^*t} dt \in \mathbb{R}^{n \times n} \right)^* = \int_0^T (e^{At} B B^* e^{A^*t})^* dt \\ &= \int_0^T (B^* e^{A^*t})^* (e^{At} B)^* dt = \int_0^T e^{At} B B^* e^{A^*t} dt = Q_T. \end{aligned}$$

Além disto, Q_T é positiva semi-definida. Com efeito, considere $z \in \mathbb{R}^n$. Temos

$$z^* Q_T z = z^* \left(\int_0^T e^{At} B B^* e^{A^*t} dt \right) z = \int_0^T z^* e^{At} B B^* e^{A^*t} z dt = \int_0^T (B^* e^{A^*t} z)^* B^* e^{A^*t} z dt .$$

Note que $B^* e^{A^*t} z \in \mathbb{R}^m$. Façamos $B^* e^{A^*t} z = (\alpha_1, \dots, \alpha_m)$. Assim,

$$(B^* e^{A^*t} z)^* B^* e^{A^*t} z = \begin{bmatrix} \alpha_1 & \cdots & \alpha_m \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix} = \alpha_1^2 + \alpha_2^2 + \cdots + \alpha_m^2 \geq 0$$

e então por propriedade da integral vem

$$\int_0^T (B^* e^{A^*t} z)^* B^* e^{A^*t} z dt \geq \int_0^T 0 dt = 0$$

ou seja,

$$z^* Q_T z \geq 0$$

mostrando que de fato Q_T é positiva semi-definida.

Lema 1.1 $z \in (\text{Im}C(A, B))^\perp$ se, e somente se, $z^* e^{At} B = 0$ para todo $t \in \mathbb{R}$.

Demonstração:(\Rightarrow) Se $z \in (\text{Im}C(A, B))^\perp$ então pelo Teorema A.7 $z \in \ker C(A, B)^*$ ou seja $C(A, B)^* z = 0$. Tomando a transposta em ambos os membros obtemos $z^* C(A, B) = 0$. Logo,

$$z^* A^k B = 0 \quad ; \quad k = 0, 1, \dots, n-1 .$$

Pelo Teorema de Cayley-Hamilton temos que

$$A^n \in [A^0, A, A^2, \dots, A^{n-1}] ,$$

isto é,

$$A^n = c_{n-1} A^{n-1} + c_{n-2} A^{n-2} + \cdots + c_1 A + c_0 A^0 .$$

com $c_i \in \mathbb{R}$ sendo $0 \leq i \leq n-1$. Multiplicando ambos os membros desta igualdade à esquerda por z^* e à direita por B temos

$$\begin{aligned} z^* A^n B &= c_{n-1} \underbrace{z^* A^{n-1} B}_0 + c_{n-2} \underbrace{z^* A^{n-2} B}_0 + \cdots + c_1 \underbrace{z^* A B}_0 + c_0 \underbrace{z^* A^0 B}_0 \\ &\Rightarrow z^* A^n B = 0 \end{aligned}$$

e assim, pelo princípio da indução, concluímos que $z^* A^k B = 0$ para todo $k \in \mathbb{N}_0$. Logo, vale para todo $k \in \mathbb{N}_0$ o seguinte

$$z^* A^k B = 0 \Rightarrow \frac{t^k}{k!} z^* A^k B = 0 \Rightarrow z^* \frac{t^k A^k}{k!} B = 0 .$$

Daí

$$\sum_{k=0}^{\infty} z^* \frac{t^k A^k}{k!} B = 0 \Rightarrow z^* \sum_{k=0}^{\infty} \frac{t^k A^k}{k!} B = 0$$

e pela definição de exponencial de matrizes, para todo $t \in \mathbb{R}$, temos

$$z^* e^{At} B = 0.$$

(\Leftarrow) Agora suponhamos que $z^* e^{At} B = 0$, para todo $t \in \mathbb{R}$. Podemos derivar tal expressão repetidas vezes obtendo

$$\begin{aligned} z^* e^{At} B &= 0 \\ \Rightarrow z^* A e^{At} B &= 0 \\ \Rightarrow z^* A^2 e^{At} B &= 0 \\ \Rightarrow z^* A^3 e^{At} B &= 0 \\ &\vdots \\ \Rightarrow z^* A^{n-1} e^{At} B &= 0 \end{aligned}$$

Concluimos que, $z^* A^k e^{At} B = 0$ para todo $t \in \mathbb{R}$ e $k = 1, 2, \dots, n-1$. Em particular, para $t = 0$ temos

$$z^* A^k B = 0 \quad ; \quad k = 0, 1, \dots, n-1$$

Daí

$$z^* C(A, B) = 0$$

e tomando a transposta em ambos os membros

$$C(A, B)^* z = 0$$

de modo que $z \in \ker C(A, B)^*$ e pelo Teorema A.7 $z \in (\text{Im} C(A, B))^\perp$. ■

Teorema 1.1 Para todo $T > 0$, temos, $\text{Im} C(A, B) = \text{Im} Q_T$.

Demonstração: A afirmação no teorema equivale à afirmação de que

$$(\text{Im} C(A, B))^\perp = (\text{Im} Q_T)^\perp$$

para todo $T > 0$. Provaremos esta última igualdade. Seja $T > 0$ arbitrário. Mostraremos de início que

$$(\text{Im} C(A, B))^\perp \subset (\text{Im} Q_T)^\perp.$$

Seja $z \in (\text{Im} C(A, B))^\perp$. Então pelo Lema 1.1

$$z^* e^{At} B = 0 \quad , \quad \forall t \in \mathbb{R}.$$

Daí, usando a Proposição A.11

$$\begin{aligned} z^* Q_T &= z^* \int_0^T e^{At} B B^* e^{A^* t} dt = \int_0^T \underbrace{z^* e^{At} B B^* e^{A^* t}}_0 dt = z^* Q_T = \int_0^T 0 dt \\ &\Rightarrow z^* Q_T = 0 . \end{aligned}$$

Então pelo Teorema A.7 $z \in (\text{Im} Q_T)^\perp$. Consequentemente,

$$(\text{Im} C(A, B))^\perp \subset (\text{Im} Q_T)^\perp . \quad (1.3)$$

Agora seja $z \in (\text{Im} Q_T)^\perp$. Logo, pelo Teorema A.7, $z^* Q_T = 0$. Como Q_T é simétrica, temos

$$z^* Q_T^* = 0 \Rightarrow (Q_T z)^* = 0 \Rightarrow Q_T z = 0$$

Por conseguinte, $\langle z, Q_T z \rangle = 0$, isto é,

$$\left\langle z, \int_0^T e^{At} B B^* e^{A^* t} dt z \right\rangle = 0 .$$

Usando a Proposição A.12 vem

$$\int_0^T \langle z, e^{At} B B^* e^{A^* t} z \rangle dt = 0 .$$

Recorde que $\langle Ax, y \rangle = \langle x, A^* y \rangle$ para todo $x \in \mathbb{R}^m$ e todo $y \in \mathbb{R}^n$. Com isto, por propriedades do conjugado transposto (que neste caso trata-se simplesmente da transposta) e do produto interno, temos

$$\begin{aligned} \int_0^T \left\langle \underbrace{z}_x, \underbrace{e^{At} B}_{A^*} \underbrace{B^* e^{A^* t} z}_y \right\rangle dt &= 0 \Rightarrow \int_0^T \langle (e^{At} B)^* z, B^* e^{A^* t} z \rangle dt = 0 \\ &\Rightarrow \int_0^T \langle B^* e^{A^* t} z, B^* e^{A^* t} z \rangle dt = 0 \Rightarrow \int_0^T \|B^* e^{A^* t} z\|^2 dt = 0 . \end{aligned}$$

Uma vez que a função $t \mapsto \|B^* e^{A^* t} z\|^2$ é contínua e não negativa, segue pelo Teorema A.14 que $\|B^* e^{A^* t} z\|^2 = 0$ e, consequentemente, $B^* e^{A^* t} z = 0$ para todo $t \in [0, T]$. Assim,

$$(B^* e^{A^* t} z)^* = 0 \Rightarrow z^* e^{At} B = 0$$

Logo, pelo Lema 1.1, $z \in (\text{Im} C(A, B))^\perp$ de forma que

$$(\text{Im} Q_T)^\perp \subset (\text{Im} C(A, B))^\perp . \quad (1.4)$$

Das inclusões (1.3) e (1.4) concluímos que $(\text{Im} C(A, B))^\perp = (\text{Im} Q_T)^\perp$ e, portanto,

$$\text{Im} C(A, B) = \text{Im} Q_T$$

qualquer que seja $T > 0$. ■

Corolário 1.1 Q_T é invertível para todo $T > 0 \Leftrightarrow \text{Posto } C(A, B) = n$.

Demonstração: (\Rightarrow) Se Q_T é invertível então pela proposição A.4 $\dim \text{Im}Q_T = n$ e como pelo Teorema 1.1 $\text{Im}Q_T = \text{Im}C(A, B)$, temos que $\dim \text{Im}C(A, B) = n$, ou seja, $\text{Posto } C(A, B) = n$.

(\Leftarrow) Já se $\text{Posto } C(A, B) = n$, então $\dim \text{Im}C(A, B) = n$ e pelo Teorema 1.1 $\dim \text{Im}Q_T = n$ de modo que, pela proposição A.4, Q_T é invertível. ■

Além disso, pelo Teorema A.8, temos que para cada $T > 0$, existe $Q_T^\# \in \mathbb{R}^{n \times n}$ tal que $Q_T Q_T^\# z = z$ para todo $z \in \text{Im}Q_T$. Daí, pelo Teorema 1.1, podemos reescrever tal afirmação da seguinte forma

$$\forall T > 0, \exists Q_T^\# \in \mathbb{R}^{n \times n} : Q_T Q_T^\# z = z, \forall z \in \text{Im}C(A, B). \quad (1.5)$$

Se Q_T possui posto completo, isto é, $\text{Posto } Q_T = \min\{n, n\} = n$, pelo Teorema 1.1, resulta que $\text{Posto } C(A, B) = n$. E ainda mais, se Q_T tem posto completo, pelo Teorema A.8 temos

$$Q_T Q_T^\# = I \Rightarrow Q_T^\# = Q_T^{-1} \text{ (a única inversa).}$$

Concluimos que se $\text{Posto } C(A, B) = n$, então $Q_T^\#$ é única e se $\text{Posto } C(A, B) < n$ então $Q_T^\#$ não é única em geral.

Introduzimos a aplicação controle-para-estado C_T (parametrizado por $T > 0$), a qual será importante na caracterização do espaço dos estados alcançáveis a partir de 0, bem como da controlabilidade do sistema (1.1) (ver Teorema 1.3).

$$\begin{aligned} C_T : PC([0, T], \mathbb{R}^m) &\longrightarrow \mathbb{R}^n \\ u &\longmapsto \int_0^T e^{A(T-t)} B u(t) dt. \end{aligned} \quad (1.6)$$

O mapa C_T é linear. De fato, dados $u, v \in PC([0, T], \mathbb{R}^m)$ e $\alpha \in \mathbb{R}$ tem-se:

$$C_T(\alpha u) = \int_0^T e^{A(T-t)} B \alpha u(t) dt = \alpha \int_0^T e^{A(T-t)} B u(t) dt = \alpha C_T(u) = \alpha C_T u$$

e

$$\begin{aligned} C_T(u + v) &= \int_0^T e^{A(T-t)} B (u + v)(t) dt = \int_0^T (e^{A(T-t)} B u(t) + e^{A(T-t)} B v(t)) dt \\ &= \int_0^T e^{A(T-t)} B u(t) dt + \int_0^T e^{A(T-t)} B v(t) dt = C_T(u) + C_T(v) = C_T u + C_T v \end{aligned}$$

onde, denotamos $C_T(u)$ simplesmente por $C_T u$. Note que, de

$$x(T; x_0, u) = e^{AT} x_0 + \int_0^T e^{A(T-t)} B u(t) dt,$$

segue que

$$x(T; x_0, u) = e^{AT} x_0 + C_T u .$$

Agora, observe que

$$\text{Im}C_T = \{x \in \mathbb{R}^n : x = C_T u , u \in PC([0, T], \mathbb{R}^m)\} .$$

Note que para $x = C_T u$, temos

$$x = C_T u = 0 + C_T u = e^{At} \cdot 0 + C_T u = x(T, 0, u) .$$

Logo,

$$\text{Im}C_T = \{x \in \mathbb{R}^n : x = x(T, 0, u) , u \in PC([0, T], \mathbb{R}^m)\} .$$

ou seja, $\text{Im}C_T$ é o conjunto de estados alcançáveis a partir de 0 no tempo $T > 0$.

Sejam $T > 0$ arbitrário e $Q_T^\#$ tal que (1.5) é satisfeito. Consideremos $C_T^\# : \mathbb{R}^n \rightarrow PC([0, T], \mathbb{R}^m)$ definida, para todo $t \in [0, T]$, por

$$(C_T^\# z)(t) := B^* e^{A^*(T-t)} Q_T^\# z . \quad (1.7)$$

Dados $z, w \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$, temos para todo $t \in [0, T]$:

$$(i) (C_T^\# \alpha z)(t) = B^* e^{A^*(T-t)} Q_T^\# \alpha z = \alpha B^* e^{A^*(T-t)} Q_T^\# z = \alpha (C_T^\# z)(t) .$$

$$(ii) (C_T^\# (z + w))(t) = B^* e^{A^*(T-t)} Q_T^\# (z + w) = B^* e^{A^*(T-t)} Q_T^\# z + B^* e^{A^*(T-t)} Q_T^\# w \\ = (C_T^\# z)(t) + (C_T^\# w)(t)$$

Portanto, $C_T^\#$ é linear e, para todo $z \in \mathbb{R}^n$

$$C_T C_T^\# z = C_T (C_T^\# z) = \int_0^T e^{A(T-t)} B C_T^\# z dt = \int_0^T e^{A(T-t)} B B^* e^{A^*(T-t)} Q_T^\# z dt$$

Recorde que Q_T é uma função de $T > 0$ e, de modo análogo, $Q_T^\#$ depende apenas de $T > 0$ e então podemos escrever pela Proposição A.11

$$\int_0^T e^{A(T-t)} B B^* e^{A^*(T-t)} dt Q_T^\# z = Q_T Q_T^\# z .$$

Logo, por 1.5, qualquer que seja $z \in \text{Im}C(A, B)$, tem-se

$$C_T C_T^\# z = z . \quad (1.8)$$

Em particular, se $\text{Im}C(A, B) = \mathbb{R}^n$, então C_T é invertível à direita (e $C_T^\#$ é uma inversa à direita). Com efeito, se $\text{Im}C(A, B) = \mathbb{R}^n$ então, pelo Teorema 1.1, $\text{Im}Q_T = \mathbb{R}^n$, isto é, Q_T possui posto completo, donde, pelo Teorema A.8 $Q_T Q_T^\# = I$. Logo,

$$C_T C_T^\# z = Q_T Q_T^\# z \Rightarrow C_T C_T^\# z = Iz \Rightarrow C_T C_T^\# = I .$$

Portanto,

$$\text{Posto } C(A, B) = n \Rightarrow C_T \text{ é invertível à direita para todo } T > 0 . \quad (1.9)$$

Seja R o conjunto dos estados do sistema (1.1) alcançáveis a partir de 0 em certo tempo $T > 0$, isto é,

$$R := \bigcup_{T>0} \text{Im}C_T = \{x(T; 0, u) = C_T u : T > 0, u \in PC([0, T], \mathbb{R}^m)\} . \quad (1.10)$$

No resultado abaixo será estabelecida a identidade $R = \text{Im}C(A, B)$ que nos dá uma descrição algébrica clara do subespaço R de estados alcançáveis a partir de 0, estabelecendo deste modo, um meio fácil e eficiente para o cálculo de R .

Teorema 1.2 *Para todo $T > 0$, temos, $R = \text{Im}C(A, B) = \text{Im}C_T$.*

Demonstração: Vamos mostrar inicialmente que $R \subset \text{Im}C(A, B)$. Seja $z \in R$ arbitrário. Então existem $T > 0$ e $u \in PC([0, T], \mathbb{R}^m)$ tal que $z = C_T u$. Recordamos que $C_T u \in \mathbb{R}^n$. Ainda mais, $\text{Im}C(A, B) \subset \mathbb{R}^n$. Pelo teorema da decomposição ortogonal

$$\mathbb{R}^n = \text{Im}C(A, B) \oplus (\text{Im}C(A, B))^\perp .$$

Assim, pela Proposição A.1, existem $z_1 \in \text{Im}C(A, B)$ e $z_2 \in (\text{Im}C(A, B))^\perp$ tais que

$$z = z_1 + z_2 . \quad (1.11)$$

Mas se $z_2 \in (\text{Im}C(A, B))^\perp$, pelo Lema 1.1, $z_2^* e^{At} B = 0$ para todo $t \in \mathbb{R}$ e em particular $z_2^* e^{A(T-t)} B = 0$ para todo $t \in [0, T]$. Multiplicando esta igualdade por $u(t)$ obtemos $z_2^* e^{A(T-t)} B u(t) = 0$ e agora, integrando ambos os membros de 0 a T obtemos:

$$\int_0^T z_2^* e^{A(T-t)} B u(t) dt = 0 .$$

Façamos $z_2 = (\alpha_1, \alpha_2, \dots, \alpha_n)$ e $e^{A(T-t)} B u(t) = (\lambda_1, \lambda_2, \dots, \lambda_n)$. Assim,

$$z_2^* e^{A(T-t)} B u(t) = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix}$$

$$= \alpha_1 \lambda_1 + \alpha_2 \lambda_2 + \cdots + \alpha_n \lambda_n = \langle z_2, e^{A(T-t)} B u(t) \rangle$$

e então

$$\int_0^T z_2^* e^{A(T-t)} B u(t) dt = 0 \Rightarrow \int_0^T \langle z_2, e^{A(T-t)} B u(t) \rangle dt = 0 .$$

Usando a Proposição A.12 temos da última igualdade acima

$$\left\langle z_2, \int_0^T e^{A(T-t)} B u(t) dt \right\rangle = 0 \Rightarrow \langle z_2, C_T u \rangle = 0 \Rightarrow \langle z_2, z_1 + z_2 \rangle = 0$$

e, por propriedades do produto interno, vem

$$\langle z_2, z_1 \rangle + \langle z_2, z_2 \rangle = 0 .$$

Como $z_2 \in (\text{Im}C(A, B))^\perp$ temos que $\langle z_2, z_1 \rangle = 0$. Logo,

$$\|z_2\|^2 = 0 \Rightarrow \|z_2\| = 0 \Rightarrow z_2 = 0 .$$

Resulta então daí e por (1.11) que $z = z_1 \in \text{Im}C(A, B)$. Portanto $R \subset \text{Im}C(A, B)$.

Agora seja $z \in \text{Im}C(A, B)$. Sejam $T > 0$ e $u = C_T^\# z$. Por (1.8)

$$z = C_T C_T^\# z = C_T u = x(T; 0, u) .$$

Logo, $z \in R$ e então $\text{Im}C(A, B) \subset R$. Das duas últimas inclusões verificadas resulta que

$$R = \text{Im}C(A, B) .$$

Além disso, $z \in \text{Im}C_T$ e então $\text{Im}C(A, B) \subset \text{Im}C_T$, isto é, $R \subset \text{Im}C_T$. Como $R :=$

$\bigcup_{T>0} \text{Im}C_T$, é imediato que $\text{Im}C_T \subset R$. Logo,

$$R = \text{Im}C_T$$

e portanto,

$$R = \text{Im}C(A, B) = \text{Im}C_T,$$

qualquer que seja $T > 0$. ■

Corolário 1.2 C_T é sobrejetora para todo $T > 0$ se, e somente se, $\text{Posto } C(A, B) = n$.

Demonstração: (\Rightarrow) De fato, se C_T é sobrejetora para todo $T > 0$, então $\text{Im}C_T = \mathbb{R}^n$, de modo que $\dim \text{Im}C_T = n$. Mas pelo Teorema 1.2, $\text{Im}C_T = \text{Im}C(A, B)$. Logo, $\dim \text{Im}C(A, B) = n$, ou seja, $\text{Posto } C(A, B) = n$.

(\Leftarrow) Agora, se $\text{Posto } C(A, B) = n$, temos $\dim \text{Im}C(A, B) = n$ e então pelo Teorema 1.2 $\dim \text{Im}C_T = n$ para todo $T > 0$. Como $\text{Im}C_T \subset \mathbb{R}^n$ e $\dim \mathbb{R}^n = n$ temos pelo Teorema A.1 que $\text{Im}C_T = \mathbb{R}^n$, mostrando que C_T é sobrejetora para todo $T > 0$. ■

Temos agora condições de formular várias caracterizações do conceito de controlabilidade do sistema linear (1.1).

Teorema 1.3 *As seguintes afirmações são equivalentes:*

- (1) O sistema 1.1 é controlável;
- (2) $\text{Posto } C(A, B) = n$;
- (3) Q_T é invertível para algum $T > 0$;
- (4) Q_T é invertível para todo $T > 0$;
- (5) C_T é sobrejetora para algum $T > 0$;
- (6) C_T é sobrejetora para todo $T > 0$;
- (7) C_T é invertível à direita para algum $T > 0$;
- (8) C_T é invertível à direita para todo $T > 0$.

Demonstração:

(1) \Rightarrow (2). Se o sistema 1.1 é controlável, então quaisquer que sejam $x_0, x_1 \in \mathbb{R}^n$, x_1 é alcançável a partir de x_0 . De modo particular, dados $0, x_1 \in \mathbb{R}^n$ tem-se x_1 alcançável a partir de 0. Logo $x_1 \in R$ e com isto $\mathbb{R}^n \subset R$. Como $R \subset \mathbb{R}^n$ por definição, concluímos que $R = \mathbb{R}^n$. Pelo Teorema 1.2 vem que $\text{Im}C(A, B) = \mathbb{R}^n$ e então $\dim \text{Im}C(A, B) = \dim \mathbb{R}^n$. Portanto $\text{Posto } C(A, B) = n$, como queríamos.

(2) \Rightarrow (3). Supondo $\text{Posto } C(A, B) = n$, pelo Corolário 1.1, Q_T é invertível para todo $T > 0$. Em particular, Q_T é invertível para algum $T > 0$.

(3) \Rightarrow (4). Supondo Q_T invertível para algum $T > 0$, temos, pela Proposição A.4, $\text{Posto } Q_T = n$ e então pelo Teorema 1.1 resulta que $\text{Posto } C(A, B) = n$. Segue-se pelo Corolário 1.1 que Q_T é invertível para todo $T > 0$.

(4) \Rightarrow (5). Supondo Q_T invertível para todo $T > 0$, pelo Corolário 1.1 resulta que $\text{Posto } C(A, B) = n$. Daí, pelo Corolário 1.2, C_T é sobrejetora para todo $T > 0$. Em particular, C_T é sobrejetora para algum $T > 0$.

(5) \Rightarrow (6). Seja $T > 0$ tal que C_T é sobrejetora. Então $\text{Im}C_T = \mathbb{R}^n$. Logo, pelo Teorema 1.2 resulta que $\text{Im}C(A, B) = \mathbb{R}^n$ e assim $\text{Posto } C(A, B) = \dim \text{Im}C(A, B) = n$ de modo que, pelo Corolário 1.2, temos que C_T é sobrejetora para todo $T > 0$.

(6) \Rightarrow (7). Suponha C_T sobrejetora para todo $T > 0$. Então, pelo Teorema A.4, C_T é invertível à direita para todo $T > 0$. Em particular, C_T é invertível à direita para algum $T > 0$.

(7) \Rightarrow (8). Se C_T é invertível à direita para algum $T > 0$ então, para o mesmo $T > 0$, C_T é sobrejetora. Como já temos (5) \Rightarrow (6), segue-se que C_T é sobrejetora para todo $T > 0$ e assim pelo Teorema A.4 temos que C_T é invertível à direita para todo $T > 0$.

(8) \Rightarrow (1). Assuma que C_T é invertível à direita para todo $T > 0$. Então pelo Teorema A.4, C_T é sobrejetora, ou seja, $\text{Im}C_T = \mathbb{R}^n$ para todo $T > 0$. Sendo assim, fixado $T > 0$ e tomando $z \in \mathbb{R}^n$ existe $u \in PC([0, T], \mathbb{R}^m)$ tal que $z = C_T u$. Considere pois $x_0, x_1 \in \mathbb{R}^n$. Claro que $e^{AT} x_0 \in \mathbb{R}^n$ e, por ser \mathbb{R}^n um espaço vetorial, temos que $x_1 - e^{AT} x_0 \in \mathbb{R}^n$. Existe então $u \in PC([0, T], \mathbb{R}^m)$ tal que

$$C_T u = x_1 - e^{AT} x_0 \Rightarrow e^{AT} x_0 + C_T u = x_1 .$$

Logo,

$$x(T; x_0, u) = x_1$$

mostrando que x_1 é alcançável a partir de x_0 . Como foram x_0 e x_1 tomados arbitrariamente, temos daí que o sistema (1.1) é controlável, concluindo a demonstração do teorema. ■

Das caracterizações acima da controlabilidade do sistema 1.1, o item (2) é o mais útil em aplicações pois fornece um simples teste de controlabilidade para o referido sistema através do posto da matriz de controlabilidade do mesmo.

Definição 1.2 Um subconjunto $X \subset \mathbb{R}^n$ é dito A -invariante quando, para todo $v \in X$, tem-se $Av \in X$.

Proposição 1.1 O subespaço $\text{Im}C(A, B) \subset \mathbb{R}^n$ é A -invariante e contém $\text{Im}B$. Mais ainda, este é o menor subespaço do \mathbb{R}^n com tais propriedades, isto é, se $S \subset \mathbb{R}^n$ é um

subespaço A -invariante contendo $\text{Im}B$, então $\text{Im}C(A, B) \subset S$.

Demonstração: Sejam $v \in \mathbb{R}^m$ e

$$w := \begin{bmatrix} v \\ 0 \end{bmatrix} \in \mathbb{R}^{mn}$$

Sabemos que

$$\text{Im}C(A, B) = \{C(A, B)x : x \in \mathbb{R}^{mn}\}$$

e

$$\text{Im}B = \{Bx : x \in \mathbb{R}^m\}.$$

Como $w \in \mathbb{R}^{mn}$ então $C(A, B)w \in \text{Im}C(A, B)$ e, já que $v \in \mathbb{R}^m$, temos que $Bv \in \text{Im}B$.

Mas

$$C(A, B)w = \begin{bmatrix} B & AB & A^2B & \dots & A^{n-1}B \end{bmatrix} \begin{bmatrix} v \\ 0 \end{bmatrix} = Bv$$

Isso nos diz que, qualquer que seja $Bv \in \text{Im}B$, podemos escrever

$$Bv = C(A, B)w$$

com $w \in \mathbb{R}^{mn}$ definido acima, isto é, $Bv \in \text{Im}C(A, B)$. Logo, $\text{Im}B \subset \text{Im}C(A, B)$. Para provar a A -invariância, seja $z \in \text{Im}C(A, B)$. Devemos mostrar que $Az \in \text{Im}C(A, B)$.

Como $z \in \text{Im}C(A, B)$ temos que $z = C(A, B)w$ para algum $w \in \mathbb{R}^{mn}$. Seja

$$w = \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_{n-1} \end{bmatrix} \in \mathbb{R}^{mn}$$

com $v_0, v_1, \dots, v_{n-1} \in \mathbb{R}^m$. Logo

$$\begin{aligned} z = C(A, B)w &= \begin{bmatrix} B & AB & A^2B & \dots & A^{n-1}B \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_{n-1} \end{bmatrix} \\ &= A^0Bv_0 + ABv_1 + A^2Bv_2 + \dots + A^{n-1}Bv_{n-1} \end{aligned}$$

e, portanto,

$$z = \sum_{k=0}^{n-1} A^k Bv_k. \tag{1.12}$$

É proveitoso enfatizar que o raciocínio acima nos diz que $z \in \text{Im}C(A, B)$ se, e somente se, existem n vetores do \mathbb{R}^m , digamos, v_0, v_1, \dots, v_{n-1} , tais que z é da forma dada em (1.12). Como consequência de (1.12) temos,

$$Az = \sum_{k=0}^{n-1} A^{k+1} B v_k = \sum_{k=1}^n A^k B v_{k-1} = \sum_{k=1}^{n-1} A^k B v_{k-1} + A^n B v_{n-1} .$$

Pelo Teorema de Cayley-Hamilton, existem $c_0, c_1, \dots, c_{n-1} \in \mathbb{R}$ tais que

$$A^n = c_{n-1} A^{n-1} + c_{n-2} A^{n-2} + \dots + c_1 A + c_0 A^0 = \sum_{k=0}^{n-1} c_k A^k .$$

Logo,

$$\begin{aligned} Az &= \sum_{k=1}^{n-1} A^k B v_{k-1} + \sum_{k=0}^{n-1} c_k A^k B v_{n-1} \\ &= \sum_{k=1}^{n-1} A^k B v_{k-1} + \sum_{k=1}^{n-1} c_k A^k B v_{n-1} + c_0 A^0 B v_{n-1} \\ &= \sum_{k=1}^{n-1} A^k B (v_{k-1} + c_k v_{n-1}) + A^0 B c_0 v_{n-1} . \end{aligned}$$

Defina $w_0 := c_0 v_{n-1} \in \mathbb{R}^m$ e para $1 \leq k \leq n-1$ defina $w_k := (v_{k-1} + c_k v_{n-1}) \in \mathbb{R}^m$.

Podemos então escrever a partir do exposto acima

$$Az = \sum_{k=1}^{n-1} A^k B w_k + A^0 B w_0 \Rightarrow Az = \sum_{k=0}^{n-1} A^k B w_k$$

com $w_k \in \mathbb{R}^m$. Daí,

$$\begin{aligned} Az &= B w_0 + A B w_1 + A^2 B w_2 + \dots + A^{n-1} B w_{n-1} \\ &= \begin{bmatrix} B & AB & A^2 B & \dots & A^{n-1} B \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_{n-1} \end{bmatrix} \end{aligned}$$

Definindo $\xi := (w_0, w_1, w_2, \dots, w_{n-1}) \in \mathbb{R}^{mn}$ temos da igualdade acima que

$$Az = C(A, B)\xi$$

e portanto $Az \in \text{Im}C(A, B)$.

Por fim, seja $S \subset \mathbb{R}^n$ um subespaço vetorial A -invariante contendo $\text{Im}B$ e seja $z \in \text{Im}C(A, B)$. Então existem n vetores $v_k \in \mathbb{R}^m$, com $1 \leq k \leq n - 1$, tais que z é da forma dada em (1.12), isto é,

$$z = \sum_{k=0}^{n-1} A^k B v_k .$$

Como $\text{Im}B \subset S$ e $B v_k \in \text{Im}B$ segue que $B v_k \in S$. Além disso, pela A -invariância de S

$$\begin{aligned} AB v_k &\in S, \\ A^2 B v_k &\in S, \\ &\vdots \\ A^k B v_k &\in S \quad \forall k \in \{0, 1, \dots, n-1\}. \end{aligned}$$

Sendo S um subespaço do \mathbb{R}^n , a soma finita de elementos de S pertence também a S . Desse modo, $z \in S$ mostrando que $\text{Im}C(A, B) \subset S$ e concluindo a demonstração. ■

Proposição 1.2 $\text{Im}C(A, B)$ é e^{At} -invariante, ou seja, $e^{At}v \in \text{Im}C(A, B)$ para todo $v \in \text{Im}C(A, B)$ e todo $t \in \mathbb{R}$.

Demonstração: Pela Proposição 1.1, $\text{Im}C(A, B)$ é A -invariante, isto é, $Av \in \text{Im}C(A, B)$ para todo $v \in \text{Im}C(A, B)$. Logo,

$$\begin{aligned} A^2 v &\in \text{Im}C(A, B) \\ A^3 v &\in \text{Im}C(A, B) \\ &\vdots \\ A^k v &\in \text{Im}C(A, B) \end{aligned}$$

ou seja, $\text{Im}C(A, B)$ é A^k -invariante, para todo $k \in \mathbb{N}_0$. Sejam $v \in \text{Im}C(A, B)$ e $t \in \mathbb{R}$. Então, $A^k v \in \text{Im}C(A, B)$ e como $\text{Im}C(A, B)$ é um subespaço \mathbb{R}^n , temos,

$$\frac{t^k}{k!} A^k v \in \text{Im}C(A, B),$$

para todo $k \in \mathbb{N}_0$. Daí

$$s_n = \sum_{k=0}^n \frac{t^k}{k!} A^k v \in \text{Im}C(A, B) .$$

Pela Proposição A.9, $\text{Im}C(A, B) \subset \mathbb{R}^n$ é fechado e, portanto,

$$\lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{t^k}{k!} A^k v = s \in \text{Im}C(A, B)$$

ou seja,

$$s = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k v \in \text{Im}C(A, B) .$$

Logo, pela Definição A.7, temos

$$e^{At}v \in \text{Im}C(A, B).$$

■

Uma consequência imediata da Proposição 1.2 é o fato de que, dado um par $x_0, x_1 \in \text{Im}C(A, B)$ e $T > 0$ temos $x_1 - e^{AT}x_0 \in \text{Im}C(A, B) = \text{Im}C_T$ (pelo Teorema 1.2) e então usando (1.8) o controle u^\dagger definido por

$$u^\dagger := C_T^\#(x_1 - e^{AT}x_0)$$

é tal que

$$C_T u^\dagger = C_T C_T^\#(x_1 - e^{AT}x_0) = x_1 - e^{AT}x_0 \Rightarrow x_1 = e^{AT}x_0 + C_T u^\dagger$$

ou, equivalentemente,

$$x(T; x_0, u^\dagger) = e^{AT}x_0 + C_T u^\dagger = x_1 .$$

Concluimos que para cada par $x_0, x_1 \in \text{Im}C(A, B)$ e $T > 0$, existe um controle u^\dagger conduzindo x_0 para x_1 no tempo T . Assim sendo, $\text{Im}C(A, B)$ é denominado *subespaço controlável*.

Dados $x_0, x_1 \in \text{Im}C(A, B)$, denotaremos por $\mathcal{U}_T(x_0, x_1)$ o conjunto dos controles que alcançam a evolução de x_0 para x_1 em um tempo $T > 0$, isto é,

$$\mathcal{U}_T(x_0, x_1) := \{u \in PC([0, T], \mathbb{R}^m) : x(T; x_0, u) = x_1\} .$$

Este conjunto é claramente não vazio pois, como vimos anteriormente, o controle $u^\dagger := C_T^\#(x_1 - e^{AT}x_0)$ está em $\mathcal{U}_T(x_0, x_1)$ já que

$$x(T; x_0, u^\dagger) = x_1 .$$

A medida da energia de um controle $u \in \mathcal{U}_T(x_0, x_1)$ é dada por $\int_0^T \|u(t)\|^2 dt$. Em termos de aplicação, ela nos diz o custo devido ao uso de um controle. No problema de colocar um satélite em órbita (o qual será estudado no Capítulo 2), por exemplo, possa ser que o uso contínuo de jatos acarrete num alto gasto de combustível. O total de combustível seria, neste caso, a energia do controle utilizado (os jatos).

O próximo resultado nos diz que o controle u^\dagger alcança a evolução de $x_0 \in \text{Im}C(A, B)$ para $x_1 \in \text{Im}C(A, B)$ em um tempo T com uma menor energia gasta.

Proposição 1.3 (Controle de Energia mínima) *Sejam $x_0, x_1 \in \text{Im}C(A, B)$ e $T > 0$.*

O controle u^\dagger definido por $u^\dagger := C_T^\#(x_1 - e^{AT}x_0)$ é tal que $u^\dagger \in \mathcal{U}_T(x_0, x_1)$ e

$$\int_0^T \|u^\dagger(t)\|^2 dt \leq \int_0^T \|u(t)\|^2 dt \quad \forall u \in \mathcal{U}_T(x_0, x_1).$$

Demonstração: Já mostramos que $u^\dagger \in \mathcal{U}_T(x_0, x_1)$. Agora, para todo $u \in \mathcal{U}_T(x_0, x_1)$ vale

$$x(T; x_0, u) = x_1 \Rightarrow e^{AT}x_0 + C_T u = x_1. \quad (1.13)$$

Desde que $u^\dagger \in \mathcal{U}_T(x_0, x_1)$, temos

$$x(T; x_0, u^\dagger) = x_1 \Rightarrow e^{AT}x_0 + C_T u^\dagger = x_1 \quad (1.14)$$

e subtraindo (1.13) de (1.14) obtemos

$$0 = C_T u^\dagger - C_T u.$$

Como C_T é linear, vem

$$\begin{aligned} 0 &= C_T(u^\dagger - u) \\ 0 &= \int_0^T e^{A(T-t)} B(u^\dagger(t) - u(t)) dt. \end{aligned}$$

Então

$$\left\langle \int_0^T e^{A(T-t)} B(u^\dagger(t) - u(t)) dt, Q_T^\#(x_1 - e^{AT}x_0) \right\rangle = 0.$$

Daí, pela Proposição A.12 aliada às demais propriedades do produto interno, temos para todo $u \in \mathcal{U}_T(x_0, x_1)$

$$\int_0^T \left\langle \underbrace{e^{A(T-t)} B}_M \underbrace{(u^\dagger(t) - u(t))}_x, \underbrace{Q_T^\#(x_1 - e^{AT}x_0)}_y \right\rangle dt = 0.$$

Sabemos por propriedades do conjugado transposto que dado $M \in \mathbb{R}^{n \times m}$, temos $\langle Mx, y \rangle = \langle x, M^*y \rangle$ para todos $x \in \mathbb{R}^m$ e $y \in \mathbb{R}^n$ e por propriedades do produto interno podemos ainda escrever $\langle Mx, y \rangle = \langle M^*y, x \rangle$. Logo,

$$\int_0^T \left\langle B^* e^{A^*(T-t)} Q_T^\#(x_1 - e^{AT}x_0), u^\dagger(t) - u(t) \right\rangle dt = 0.$$

Pela definição de $C_T^\#$ vem

$$\begin{aligned} &\int_0^T \left\langle C_T^\#(x_1 - e^{AT}x_0), u^\dagger(t) - u(t) \right\rangle dt = 0 \\ &\Rightarrow \int_0^T \langle u^\dagger(t), u^\dagger(t) - u(t) \rangle dt = 0. \end{aligned}$$

Usando novamente propriedades do produto interno bem como das integrais obtemos

$$\begin{aligned} & \int_0^T \langle u^\dagger(t), u^\dagger(t) \rangle dt - \int_0^T \langle u^\dagger(t), u(t) \rangle dt = 0 \\ \Rightarrow & \int_0^T \|u^\dagger(t)\|^2 dt - \int_0^T \langle u^\dagger(t), u(t) \rangle dt = 0 \\ \Rightarrow & \int_0^T \|u^\dagger(t)\|^2 dt = \int_0^T \langle u^\dagger(t), u(t) \rangle dt \end{aligned}$$

Assim temos para todo $u \in \mathcal{U}_T(x_0, x_1)$

$$\begin{aligned} 0 \leq \int_0^T \|u^\dagger(t) - u(t)\|^2 dt &= \int_0^T \|u^\dagger(t)\|^2 dt + \int_0^T \|u(t)\|^2 dt - 2 \int_0^T \langle u^\dagger(t), u(t) \rangle dt \\ &= \int_0^T \|u^\dagger(t)\|^2 dt + \int_0^T \|u(t)\|^2 dt - 2 \int_0^T \|u^\dagger(t)\|^2 dt \\ &= \int_0^T \|u(t)\|^2 dt - \int_0^T \|u^\dagger(t)\|^2 dt, \end{aligned}$$

donde

$$\int_0^T \|u^\dagger(t)\|^2 dt \leq \int_0^T \|u(t)\|^2 dt .$$

■

Consideraremos agora alguns fatos elementares sobre mudança de variáveis no espaço dos estados. Considere o sistema (1.1) controlado e seja $S \in GL(n, \mathbb{R})$, onde $GL(n, \mathbb{R})$ denota o grupo linear de ordem n definido por

$$GL(n, \mathbb{R}) := \{M \in \mathbb{R}^{n \times n} : M \text{ é invertível}\} = \{M \in \mathbb{R}^{n \times n} : \det M \neq 0\} .$$

Consideremos a mudança de variáveis $\tilde{x} = S^{-1}x$. Então multiplicando $x' = Ax + bu$ por S^{-1} temos

$$\begin{aligned} S^{-1}x' &= S^{-1}Ax + S^{-1}Bu \\ &= S^{-1}ASS^{-1}x + S^{-1}Bu. \end{aligned}$$

Fazendo

$$\tilde{x}' = S^{-1}x', \quad \tilde{A} := S^{-1}AS, \quad \tilde{B} := S^{-1}B,$$

segue que

$$\tilde{x}' = \tilde{A}\tilde{x} + \tilde{B}u .$$

Assim

$$\begin{aligned}
C(\tilde{A}, \tilde{B}) &= \begin{bmatrix} \tilde{B} & \tilde{A}\tilde{B} & \tilde{A}^2\tilde{B} & \dots & \tilde{A}^{n-1}\tilde{B} \end{bmatrix} \\
&= \begin{bmatrix} S^{-1}B & S^{-1}ASS^{-1}B & S^{-1}ASS^{-1}ASS^{-1}B & \dots \\ \underbrace{S^{-1}ASS^{-1}AS \dots S^{-1}AS}_{n-1 \text{ vezes}} S^{-1}B \end{bmatrix} \\
&= \begin{bmatrix} S^{-1}B & S^{-1}AB & S^{-1}A^2B & \dots & S^{-1}A^{n-1}B \end{bmatrix} = S^{-1}C(A, B).
\end{aligned}$$

Proposição 1.4 (A, B) é controlável se, e somente se, (\tilde{A}, \tilde{B}) é controlável.

Demonstração: (\Rightarrow) Suponha que o par (A, B) é controlável. Sejam $\tilde{x}_0, \tilde{x}_1 \in \mathbb{R}^n$ e $S \in GL(n, \mathbb{R})$. Existem então $x_0, x_1 \in \mathbb{R}^n$ tais que $\tilde{x}_0 = S^{-1}x_0$ e $\tilde{x}_1 = S^{-1}x_1$. Com efeito, basta tomar $x_0 = S\tilde{x}_0$ e $x_1 = S\tilde{x}_1$. Desse modo,

$$S^{-1}x_0 = S^{-1}S\tilde{x}_0 = \tilde{x}_0 \quad \text{e} \quad S^{-1}x_1 = S^{-1}S\tilde{x}_1 = \tilde{x}_1 .$$

Como o par (A, B) é controlável, por definição o sistema (1.1) é controlável. Assim, existem $T > 0$ e um controle $u \in PC([0, T], \mathbb{R}^m)$ devido ao qual x_1 é alcançável a partir de x_0 . Em símbolos,

$$x(T; x_0, u) = x_1 .$$

Logo

$$e^{AT}x_0 + \int_0^T e^{A(T-t)}Bu(t) dt = x_1$$

e multiplicando a igualdade acima por S^{-1} à esquerda, temos

$$\begin{aligned}
&S^{-1}e^{AT}x_0 + S^{-1} \int_0^T e^{A(T-t)}Bu(t) dt = S^{-1}x_1 \\
\Rightarrow &S^{-1}e^{AT}SS^{-1}x_0 + \int_0^T S^{-1}e^{A(T-t)}SS^{-1}Bu(t) dt = S^{-1}x_1 .
\end{aligned}$$

Pela Proposição A.7, obtemos

$$\begin{aligned}
&e^{(S^{-1}AT S)}S^{-1}x_0 + \int_0^T e^{[S^{-1}A(T-t) S]}S^{-1}Bu(t) dt = S^{-1}x_1 \\
\Rightarrow &e^{(S^{-1}A S)T}S^{-1}x_0 + \int_0^T e^{(S^{-1}A S)(T-t)}S^{-1}Bu(t) dt = S^{-1}x_1 .
\end{aligned}$$

Adotemos a mudança de variáveis antes definida, isto é,

$$S^{-1}A S = \tilde{A} \quad \text{e} \quad S^{-1}B = \tilde{B} .$$

Substituindo na equação acima, temos

$$e^{\tilde{A}}S^{-1}x_0 + \int_0^T e^{\tilde{A}(T-t)}\tilde{B}u(t) dt = S^{-1}x_1$$

e, como $S^{-1}x_0 = \tilde{x}_0$ e $S^{-1}x_1 = \tilde{x}_1$, temos ainda

$$e^{\tilde{A}T}\tilde{x}_0 + \int_0^T e^{\tilde{A}(T-t)}\tilde{B}u(t) dt = \tilde{x}_1$$

ou, equivalentemente,

$$\tilde{x}(T; \tilde{x}_0, u) = \tilde{x}_1,$$

mostrando que \tilde{x}_1 é alcançável a partir de \tilde{x}_0 . Como foram o $\tilde{x}_1, \tilde{x}_0 \in \mathbb{R}^n$ arbitrários, concluímos que o sistema

$$\tilde{x}'(t) = \tilde{A}\tilde{x}(t) + \tilde{B}u(t) \quad (1.15)$$

é controlável e portanto (\tilde{A}, \tilde{B}) é controlável.

(\Leftarrow) Suponha agora que (\tilde{A}, \tilde{B}) seja controlável. Sejam $x_0, x_1 \in \mathbb{R}^n$. Escolhamos $\tilde{x}_0 = S^{-1}x_0$ e $\tilde{x}_1 = S^{-1}x_1$. Pela controlabilidade de (1.15) resulta que \tilde{x}_1 é alcançável a partir de \tilde{x}_0 , isto é,

$$\tilde{x}(T; \tilde{x}_0, u) = \tilde{x}_1 .$$

Daí,

$$e^{\tilde{A}T}\tilde{x}_0 + \int_0^T e^{\tilde{A}(T-t)}\tilde{B}u(t) dt = \tilde{x}_1 .$$

Pelas definições de $\tilde{A}, \tilde{B}, \tilde{x}_0$ e \tilde{x}_1 vem

$$\begin{aligned} e^{(S^{-1}A S)T}S^{-1}x_0 + \int_0^T e^{(S^{-1}A S)(T-t)}S^{-1}Bu(t) dt &= S^{-1}x_1 \\ \Rightarrow e^{(S^{-1}AT S)}S^{-1}x_0 + \int_0^T e^{[S^{-1}A(T-t) S]}S^{-1}Bu(t) dt &= S^{-1}x_1 \end{aligned}$$

Pela Proposição A.7 isto equivale a

$$\begin{aligned} S^{-1}e^{AT}SS^{-1}x_0 + \int_0^T S^{-1}e^{A(T-t)}SS^{-1}Bu(t) dt &= S^{-1}x_1 \\ \Rightarrow S^{-1}e^{AT}x_0 + S^{-1} \int_0^T e^{A(T-t)}Bu(t) dt &= S^{-1}x_1 . \end{aligned}$$

Finalmente, multiplicando ambos os membros à esquerda por S obtemos

$$e^{AT}x_0 + \int_0^T e^{A(T-t)}Bu(t) dt = x_1$$

ou equivalentemente

$$x(T; x_0, u) = x_1.$$

Mostramos com isto que o sistema (1.1) é tal que, dados $x_0, x_1 \in \mathbb{R}^n$, existem $T > 0$ e um controle $u \in PC([0, T], \mathbb{R}^m)$ de tal modo que x_1 é alcançável a partir de x_0 , isto é, o sistema (1.1) é controlável ou equivalentemente o par (A, B) é controlável. ■

Lema 1.2 (Decomposição da Controlabilidade de Kalman) *Suponhamos que $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$ seja não controlável e $B \neq 0$, caso no qual*

$$0 < k := \text{Posto}C(A, B) = \dim \text{Im}C(A, B) < n.$$

Então, existe $S \in GL(n, \mathbb{R})$ tal que $\tilde{A} := S^{-1}AS$ e $\tilde{B} := S^{-1}B$ tem as seguintes estruturas

$$\tilde{A} = \begin{bmatrix} A_1 & A_2 \\ 0_A & A_3 \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} B_1 \\ 0_B \end{bmatrix} \quad (1.16)$$

onde $A_1 \in \mathbb{R}^{k \times k}$, $A_2 \in \mathbb{R}^{k \times (n-k)}$, $A_3 \in \mathbb{R}^{(n-k) \times (n-k)}$, $0_A = 0 \in \mathbb{R}^{(n-k) \times k}$, $B_1 \in \mathbb{R}^{k \times m}$ e $0_B = 0 \in \mathbb{R}^{(n-k) \times m}$.

Demonstração: Sendo $\text{Im}C(A, B) \subset \mathbb{R}^n$, pela proposição A.2 existe um subespaço $\mathcal{V} \subset \mathbb{R}^n$ tal que

$$\text{Im}C(A, B) \oplus \mathcal{V} = \mathbb{R}^n. \quad (1.17)$$

Sejam $\mathcal{D} = \{v_1, v_2, \dots, v_k\}$ e $\mathcal{E} = \{v_{k+1}, v_{k+2}, \dots, v_n\}$ bases para os subespaços $\text{Im}C(A, B)$ e \mathcal{V} , respectivamente. Definimos S por

$$S := \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}.$$

É claro que $S \in GL(n, \mathbb{R}^n)$, pois pelo Teorema A.2 o conjunto \mathcal{B} dado por

$$\mathcal{B} = \mathcal{D} \cup \mathcal{E} = \{v_1, v_2, \dots, v_k, v_{k+1}, v_{k+2}, \dots, v_n\}$$

é uma base para o \mathbb{R}^n . Logo, as colunas de S são todas *L.I.* e assim $\dim \text{Im} S = \text{Posto} S = n$ resultando, pela Proposição A.4 que S é invertível.

Agora, para qualquer $z = (z_1, z_2, \dots, z_n) \in \mathbb{R}^n$ temos que

$$Sz = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \cdot \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = v_1 z_1 + v_2 z_2 + \cdots + v_n z_n$$

e como $z_j \in \mathbb{R}$ com $j = 1, 2, \dots, n$, temos que

$$Sz = \sum_{j=1}^n z_j v_j = \underbrace{(z_1 v_1 + z_2 v_2 + \cdots + z_k v_k)}_{\in \text{Im}C(A, B)} + \underbrace{(z_{k+1} v_{k+1} + z_{k+2} v_{k+2} + \cdots + z_n v_n)}_{\in \mathcal{V}}.$$

Afirmação 1.1 Para todo $z = (z_1, \dots, z_k, z_{k+1}, \dots, z_n) \in \mathbb{R}^n$ têm-se

$$Sz \in \text{Im}C(A, B) \Leftrightarrow z_{k+1} = z_{k+2} = \dots = z_n = 0 .$$

(\Leftarrow) De fato, se $z_{k+1} = z_{k+2} = \dots = z_n = 0$, pela decomposição de Sz acima descrita decorre que

$$Sz = z_1 v_1 + z_2 v_2 + \dots + z_k v_k \in \text{Im}C(A, B).$$

(\Rightarrow) Agora, se $Sz \in \text{Im}C(A, B)$, existem $a_1, a_2, \dots, a_k \in \mathbb{R}$ tais que

$$Sz = a_1 v_1 + a_2 v_2 + \dots + a_k v_k ,$$

ou seja,

$$z_1 v_1 + z_2 v_2 + \dots + z_k v_k + z_{k+1} v_{k+1} + z_{k+2} v_{k+2} + \dots + z_n v_n = a_1 v_1 + a_2 v_2 + \dots + a_k v_k$$

$$\Rightarrow (z_1 - a_1) v_1 + (z_2 - a_2) v_2 + \dots + (z_k - a_k) v_k + z_{k+1} v_{k+1} + z_{k+2} v_{k+2} + \dots + z_n v_n = 0$$

e como $\mathcal{B} = \{v_1, v_2, \dots, v_k, v_{k+1}, v_{k+2}, \dots, v_n\}$ é *L.I.* (pois é uma base para o \mathbb{R}^n) segue-se que

$$(z_1 - a_1) = (z_2 - a_2) = \dots = (z_k - a_k) = z_{k+1} = z_{k+2} = \dots = z_n = 0 .$$

Em particular

$$z_{k+1} = z_{k+2} = \dots = z_n = 0,$$

comprovando a afirmação.

Façamos $\tilde{B} := S^{-1}B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \in \mathbb{R}^{n \times m}$, $B_1 \in \mathbb{R}^{k \times m}$, $B_2 \in \mathbb{R}^{(n-k) \times m}$. Seja $v \in \mathbb{R}^m$ dado por $v = (\alpha_1, \dots, \alpha_m)$. Escrevendo

$$B_1 = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k1} & b_{k2} & \dots & b_{km} \end{bmatrix} \quad \text{e} \quad B_2 = \begin{bmatrix} b_{(k+1)1} & b_{(k+1)2} & \dots & b_{(k+1)m} \\ b_{(k+2)1} & b_{(k+2)2} & \dots & b_{(k+2)m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{bmatrix}$$

temos então,

$$\begin{aligned}
S^{-1}Bv &= \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{km} \\ \dots & \dots & \dots & \dots \\ b_{(k+1)1} & b_{(k+1)2} & \cdots & b_{(k+1)m} \\ b_{(k+2)1} & b_{(k+2)2} & \cdots & b_{(k+2)m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} \\
&= \begin{bmatrix} b_{11}\alpha_1 + b_{12}\alpha_2 + \cdots + b_{1m}\alpha_m \\ b_{21}\alpha_1 + b_{22}\alpha_2 + \cdots + b_{2m}\alpha_m \\ \vdots \\ b_{k1}\alpha_1 + b_{k2}\alpha_2 + \cdots + b_{km}\alpha_m \\ \dots \\ b_{(k+1)1}\alpha_1 + b_{(k+1)2}\alpha_2 + \cdots + b_{(k+1)m}\alpha_m \\ b_{(k+2)1}\alpha_1 + b_{(k+2)2}\alpha_2 + \cdots + b_{(k+2)m}\alpha_m \\ \vdots \\ b_{n1}\alpha_1 + b_{n2}\alpha_2 + \cdots + b_{nm}\alpha_m \end{bmatrix} = \begin{bmatrix} B_1v \\ B_2v \end{bmatrix},
\end{aligned}$$

ou seja,

$$S^{-1}Bv = \begin{bmatrix} B_1v \\ B_2v \end{bmatrix}$$

e multiplicando ambos os membros da igualdade acima por S , obtemos

$$S \begin{bmatrix} B_1v \\ B_2v \end{bmatrix} = Bv \in \text{Im}B .$$

Como, pela Proposição 1.1, $\text{Im}B \subset \text{Im}C(A, B)$, temos que

$$S \begin{bmatrix} B_1v \\ B_2v \end{bmatrix} \in \text{Im}C(A, B) , \forall v \in \mathbb{R}^m .$$

Sendo assim, escrevendo $B_1v = (z_1, z_2, \dots, z_k)$ e $B_2v = (z_{k+1}, \dots, z_n)$ temos

$$\begin{bmatrix} B_1v \\ B_2v \end{bmatrix} = z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_k \\ \dots \\ z_{k+1} \\ \vdots \\ z_n \end{bmatrix}$$

e então

$$S \begin{bmatrix} B_1v \\ B_2v \end{bmatrix} = Sz \in \text{Im}C(A, B), \forall v \in \mathbb{R}^m.$$

Disto e pela Afirmação 1.1 resulta que

$$z_{k+1} = z_{k+2} = \dots = z_n = 0,$$

ou seja, $B_2v = 0$ para todo $v \in \mathbb{R}^m$. Daí,

$$\begin{bmatrix} b_{(k+1)1} & b_{(k+1)2} & \dots & b_{(k+1)m} \\ b_{(k+2)1} & b_{(k+2)2} & \dots & b_{(k+2)m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} = 0$$

e em particular, para $v = e_1 = (1, 0, \dots, 0)$,

$$\begin{bmatrix} b_{(k+1)1} & b_{(k+1)2} & \dots & b_{(k+1)m} \\ b_{(k+2)1} & b_{(k+2)2} & \dots & b_{(k+2)m} \\ \vdots & \ddots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 0 \Rightarrow \begin{bmatrix} b_{(k+1)1} \\ b_{(k+2)1} \\ \vdots \\ b_{n1} \end{bmatrix} = 0.$$

Analogamente, para $v = e_2 = (0, 1, 0, \dots, 0)$ temos

$$\begin{bmatrix} b_{(k+1)1} & b_{(k+1)2} & \dots & b_{(k+1)m} \\ b_{(k+2)1} & b_{(k+2)2} & \dots & b_{(k+2)m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} = 0 \Rightarrow \begin{bmatrix} b_{(k+1)2} \\ b_{(k+2)2} \\ \vdots \\ b_{n2} \end{bmatrix} = 0.$$

Indutivamente, para cada $v = e_i$ com $i = 1, 2, \dots, m$, onde e_i denota o i -ésimo vetor da base canônica do \mathbb{R}^m , obtém-se

$$\begin{bmatrix} b_{(k+1)i} \\ b_{(k+2)i} \\ \vdots \\ b_{ni} \end{bmatrix} = 0,$$

mostrando que cada coluna de B_2 é nula e levando-nos à conclusão de que $B_2 = 0$.

Agora, seja

$$\tilde{A} := S^{-1}AS = \begin{bmatrix} A_1 & A_2 \\ A_4 & A_3 \end{bmatrix}$$

com $A_1 \in \mathbb{R}^{k \times k}$, $A_2 \in \mathbb{R}^{k \times (n-k)}$, $A_3 \in \mathbb{R}^{(n-k) \times (n-k)}$ e $A_4 \in \mathbb{R}^{(n-k) \times k}$. Vamos provar que $A_4 = 0$. Tomemos $w \in \text{Im}C(A, B) \subset \mathbb{R}^n$. Como pela Proposição 1.1, $\text{Im}C(A, B)$ é A -invariante, temos que $z := Aw \in \text{Im}C(A, B)$. Mais ainda,

$$S^{-1}AS = \tilde{A} \Rightarrow S^{-1}A = \tilde{A}S^{-1}.$$

Logo,

$$\begin{aligned} Aw = z &\Rightarrow \underbrace{S^{-1}A}_{\tilde{A}S^{-1}}w = S^{-1}z \\ &\Rightarrow \tilde{A}S^{-1}w = S^{-1}z. \end{aligned}$$

Façamos

$$S^{-1}w = \lambda$$

com $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{R}^n$. Multiplicando a igualdade acima por S à esquerda temos $w = S\lambda$ e como $w \in \text{Im}C(A, B)$ temos que $S\lambda \in \text{Im}C(A, B)$. Daí, pela Afirmação 1.1 segue-se que $\lambda_{k+1} = \lambda_{k+2} = \dots = \lambda_n = 0$ e, portanto,

$$S^{-1}w = \begin{bmatrix} w_1 \\ 0 \end{bmatrix}$$

com $w_1 = (\lambda_1, \lambda_2, \dots, \lambda_k) \in \mathbb{R}^k$. De modo análogo, escrevendo

$$S^{-1}z = \mu = (\mu_1, \mu_2, \dots, \mu_n) \in \mathbb{R}^n$$

obtemos $\mu_{k+1} = \mu_{k+2} = \dots = \mu_n = 0$ e então

$$S^{-1}z = \begin{bmatrix} z_1 \\ 0 \end{bmatrix}$$

com $z_1 = (\mu_1, \mu_2, \dots, \mu_k) \in \mathbb{R}^k$. Consequentemente,

$$\begin{aligned} \tilde{A}S^{-1}w = S^{-1}z &\Rightarrow \underbrace{\begin{bmatrix} A_1 & A_2 \\ A_4 & A_3 \end{bmatrix}}_{\tilde{A}} \cdot \underbrace{\begin{bmatrix} w_1 \\ 0 \end{bmatrix}}_{S^{-1}w} = \underbrace{\begin{bmatrix} z_1 \\ 0 \end{bmatrix}}_{S^{-1}z} \Rightarrow \begin{bmatrix} A_1w_1 + A_2 \cdot 0 \\ A_4w_1 + A_3 \cdot 0 \end{bmatrix} = \begin{bmatrix} z_1 \\ 0 \end{bmatrix} \\ &\Rightarrow \begin{bmatrix} A_1w_1 \\ A_4w_1 \end{bmatrix} = \begin{bmatrix} z_1 \\ 0 \end{bmatrix} \end{aligned}$$

Logo, $A_4w_1 = 0$.

Observe que para qualquer $w_1 = (a_1, \dots, a_k) \in \mathbb{R}^k$, tem-se que $w = (w_1, 0) \in \mathbb{R}^n$ é tal que $w \in \text{Im}C(A, B)$. De fato, considerando $\tau = S^{-1}w$ vem que $S\tau = w$, ou seja, $S\tau = (a_1, \dots, a_k, 0, \dots, 0) \in \mathbb{R}^n$ e como as $(n - k)$ -ésimas entradas de $S\tau$ são nulas, resulta pela Afirmação 1.1 que $S\tau \in \text{Im}C(A, B)$, ou seja, $w \in \text{Im}C(A, B)$.

Portanto, para cada $w_1 \in \mathbb{R}^k$ conseguimos um correspondente $w \in \text{Im}C(A, B)$ e, como o argumento que nos conduz ao fato de que $A_4w_1 = 0$ é válido para todo $w \in \text{Im}C(A, B)$, temos

$$A_4w_1 = 0,$$

para todo $w_1 \in \mathbb{R}^k$. Por raciocínio análogo àquele utilizado para o caso $B_2v = 0$, concluímos que $A_4 = 0$. ■

Proposição 1.5 *O par (A_1, B_1) dado no lema anterior é controlável.*

Demonstração: Pelo Teorema 1.3 basta mostrar que $\text{Posto } C(A_1, B_1) = k$. Relembre que:

$$\tilde{A} = \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix} \quad \text{e} \quad \tilde{B} = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}.$$

Assim, por indução temos para todo $k \in \mathbb{N}$

$$\tilde{A}^k \tilde{B} = \begin{bmatrix} A_1^k B_1 \\ 0 \end{bmatrix}.$$

De fato, para $k = 1$, temos

$$\tilde{A}\tilde{B} = \begin{bmatrix} A_1B_1 + A_2 \cdot 0 \\ 0 \cdot B_1 + A_3 \cdot 0 \end{bmatrix} = \begin{bmatrix} A_1B_1 \\ 0 \end{bmatrix}.$$

Suponhamos que seja válido para $k = n$, isto é,

$$\tilde{A}^n \tilde{B} = \begin{bmatrix} A_1^n B_1 \\ 0 \end{bmatrix}.$$

Então para $k = n + 1$, temos

$$\begin{aligned} \tilde{A}^{n+1} \tilde{B} &= \tilde{A} \cdot \tilde{A}^n \cdot \tilde{B} = \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix} \cdot \begin{bmatrix} A_1^n B_1 \\ 0 \end{bmatrix} = \begin{bmatrix} A_1 A_1^n B_1 + A_2 \cdot 0 \\ 0 \cdot A_1^n B_1 + A_3 \cdot 0 \end{bmatrix} \\ &= \begin{bmatrix} A_1^{n+1} B_1 \\ 0 \end{bmatrix}. \end{aligned}$$

Portanto, $\tilde{A}^k \tilde{B} = \begin{bmatrix} A_1^k B_1 \\ 0 \end{bmatrix}$ para todo $k \in \mathbb{N}$.

Agora, pelo teorema de Cayley-Hamilton a n -ésima potência de uma matriz pode ser escrita como combinação linear de todas as potências menores que a mesma. Em especial, na matriz

$$[B_1 \quad A_1 B_1 \quad \cdots \quad A_1^{n-1} B_1]$$

cada potência A_1^i com $k - 1 < i \leq n - 1$ é combinação linear das potências de A_1 menores que as mesmas e, portanto, os blocos $A_1^i B_1$ com $k - 1 < i \leq n - 1$ são LD. Daí,

$$\text{Posto } [B_1 \quad A_1 B_1 \quad \cdots \quad A_1^{n-1} B_1] = \text{Posto } [B_1 \quad A_1 B_1 \quad \cdots \quad A_1^{k-1} B_1] = \text{Posto } C(A_1, B_1).$$

Logo,

$$\begin{aligned} \text{Posto } C(A_1, B_1) &= \text{Posto } [B_1 \quad A_1 B_1 \quad \cdots \quad A_1^{n-1} B_1] \\ &= \text{Posto } \left[\begin{bmatrix} B_1 \\ 0 \end{bmatrix}, \begin{bmatrix} A_1 B_1 \\ 0 \end{bmatrix}, \begin{bmatrix} A_1^2 B_1 \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} A_1^{n-1} B_1 \\ 0 \end{bmatrix} \right] \\ &= \text{Posto } [\tilde{B} \quad \tilde{A} \tilde{B} \quad \tilde{A}^2 \tilde{B} \cdots \tilde{A}^{n-1} \tilde{B}] \\ &= \text{Posto } [S^{-1} B \quad S^{-1} A B \quad S^{-1} A^2 B \cdots S^{-1} A^{n-1} B] \\ &= \text{Posto } [S^{-1} [B \quad A B \quad A^2 B \cdots A^{n-1} B]] \\ &= \min\{\text{Posto } S^{-1}, \text{Posto } C(A, B)\}. \end{aligned}$$

Note que $\text{Posto } S^{-1} = n$ já que $S^{-1} \in GL(n, \mathbb{R})$ e como estamos nas hipóteses do Lema 1.2, $\text{Posto } C(A, B) = k < n$ de forma que $\text{Posto } C(A, B) < \text{Posto } S^{-1}$. Logo,

$$\min\{\text{Posto } S^{-1}, \text{Posto } C(A, B)\} = \text{Posto } C(A, B)$$

e então

$$\text{Posto } C(A_1, B_1) = \min\{\text{Posto } S^{-1}, \text{Posto } C(A, B)\} = \text{Posto } C(A, B) = k .$$

concluindo a demonstração. ■

O seguinte resultado é mais uma importante ferramenta algébrica para a caracterização da controlabilidade de um sistema. Sua utilidade pode ser notada, por exemplo, em situações nas quais o cálculo da matriz controlabilidade seja muito trabalhoso.

Teorema 1.4 (Critério de Hautus para controlabilidade). *O sistema (1.1) é controlável se, e somente se, $\text{Posto } (sI - A, B) = n$ para todo $s \in \mathbb{C}$.*

Demonstração:

(\Rightarrow) Provaremos o teorema usando contraposição. Para isto, suponhamos que $\text{Posto } (\lambda I - A, B) < n$ para algum $\lambda \in \mathbb{C}$. Mais especificamente, devemos ter $\lambda \in \sigma(A)$ pois para qualquer $\lambda \in \mathbb{C} \setminus \sigma(A)$ temos pela Proposição A.8, $\ker(\lambda I - A) = \{0\}$. Assim $\dim \ker(\lambda I - A) = 0$ e então, pelo Teorema do Núcleo e da Imagem, resulta que $\text{Posto } (\lambda I - A) = n$. Por definição, $\text{Posto } (\lambda I - A, B) \leq n$ e como as n colunas de $(\lambda I - A)$ já são *L.I.* segue-se que $\text{Posto } (\lambda I - A, B) = n$. Como $(\lambda I - A, B) \in \mathbb{C}^{n \times (n+m)}$, $(\lambda I - A, B)^* \in \mathbb{C}^{(n+m) \times n}$. Desde que $\text{Posto } (\lambda I - A, B)^* = \text{Posto } (\lambda I - A, B)$ temos que $\text{Posto } (\lambda I - A, B)^* < n$ e por conseguinte $n - \text{Posto } (\lambda I - A, B)^* > 0$. Pelo Teorema do Núcleo e da Imagem,

$$\text{Posto } (\lambda I - A, B)^* + \dim \ker(\lambda I - A, B)^* = n$$

e, portanto,

$$\dim \ker(\lambda I - A, B)^* = n - \text{Posto } (\lambda I - A, B)^* > 0.$$

Logo, existe $z \neq 0 \in \mathbb{C}^n$ tal que $z \in \ker(\lambda I - A, B)^*$ isto é,

$$(\lambda I - A, B)^* z = 0 .$$

Consequentemente, z também satisfaz

$$z^*(\lambda I - A, B) = 0 .$$

Então,

$$z^* \lambda I - z^* A = 0 \Rightarrow z^* A = \lambda z^* \text{ e } z^* B = 0 .$$

Daí, indutivamente, $z^* A^k = \lambda^k z^*$ para todo $k \in \mathbb{N}$. De fato, para $k = 1$ já temos $z^* A = \lambda z^*$. Suponhamos que seja válido para $k = n$, isto é,

$$z^* A^n = \lambda^n z^* .$$

Para $k = n + 1$ temos

$$z^* A^{n+1} = z^* A^n A = \lambda^n z^* A = \lambda^n \lambda z^* = \lambda^{n+1} z^*$$

e, portanto, $z^* A^k = \lambda^k z^*$ para todo $k \in \mathbb{N}$. Consequentemente,

$$z^* A^k B = \lambda^k z^* B, \quad \forall k \in \mathbb{N}.$$

Mas $z^* B = 0$. Logo,

$$z^* A^k B = 0, \quad \forall k \in \mathbb{N},$$

de onde concluímos que

$$z^* [B \quad AB \quad A^2 B \quad \dots \quad A^{n-1} B] = 0,$$

ou seja, $z^* C(A, B) = 0$. Mais ainda, $(C(A, B))^* z = 0$, isto é, $z \in \ker (C(A, B))^*$ e como $z \neq 0$, concluímos que $\dim \ker (C(A, B))^* > 0$. Pelo Teorema do Núcleo e da Imagem,

$$\text{Posto } (C(A, B))^* + \dim \ker (C(A, B))^* = n$$

$$\text{Posto } (C(A, B))^* = n - \dim \ker (C(A, B))^* < n$$

e como $\text{Posto } C(A, B) = \text{Posto } (C(A, B))^*$, concluímos que

$$\text{Posto } C(A, B) < n.$$

Resulta daí pelo Teorema 1.3 que o sistema (1.1) não é controlável.

(\Leftarrow) Suponhamos agora que o sistema (1.1) seja não controlável. Se $B = 0$, então $\text{Posto } (\lambda I - A, B) = \text{Posto } (\lambda I - A) < n$ para todo $\lambda \in \sigma(A)$. De fato, se $\lambda \in \sigma(A)$ então pela Proposição A.8, $\ker(\lambda I - A) \neq 0$. Logo,

$$\dim \ker(\lambda I - A) > 0 \Rightarrow -\dim \ker(\lambda I - A) < 0 \Rightarrow n - \dim \ker(\lambda I - A) < n.$$

Pelo Teorema do Núcleo e da Imagem

$$\text{Posto } (\lambda I - A) + \dim \ker(\lambda I - A) = n$$

$$\Rightarrow \text{Posto } (\lambda I - A) = n - \dim \ker(\lambda I - A) < n.$$

Portanto, $\text{Posto } (\lambda I - A, B) < n$ para todo $\lambda \in \sigma(A)$ e então o teorema está demonstrado.

Suponha $B \neq 0$. Então pelo Lema 1.2 existe $S \in GL(n, \mathbb{R})$, tal que

$$\tilde{A} := S^{-1} A S = \begin{bmatrix} A_1 & A_2 \\ 0_A & A_3 \end{bmatrix} \quad \text{e} \quad \tilde{B} := \begin{bmatrix} B_1 \\ 0_B \end{bmatrix}$$

onde $A_1 \in \mathbb{R}^{k \times k}$, $A_2 \in \mathbb{R}^{k \times (n-k)}$, $A_3 \in \mathbb{R}^{(n-k) \times (n-k)}$, $0_A = 0 \in \mathbb{R}^{(n-k) \times k}$, $B_1 \in \mathbb{R}^{k \times m}$, $0_B = 0 \in \mathbb{R}^{(n-k) \times m}$ e $k < n$. Seja $\lambda \in \mathbb{C}$ e $v \in \mathbb{C}^{n-k}$ um autopar de A_3^* . Então por definição, $v \neq 0$ e $v \in \ker(\lambda I - A_3^*)$, ou seja,

$$\begin{aligned} (\lambda I - A_3^*)v = 0 &\Rightarrow v^*(\lambda I - A_3^*)^* = 0 \Rightarrow v^*((\lambda I)^* - (A_3^*)^*) = 0 \Rightarrow v^*(I\lambda^* - A_3) = 0 \\ &\Rightarrow v^*(\bar{\lambda}I - A_3) = 0. \end{aligned}$$

Seja $w := \begin{bmatrix} 0 \\ v \end{bmatrix} \in \mathbb{C}^n$. Mostremos que $w^*(\bar{\lambda}I - \tilde{A}) = 0$ e $w^*\tilde{B} = 0$. Observe que:

(i) $w^* = \begin{bmatrix} 0 & v^* \end{bmatrix} \in \mathbb{C}^{1 \times n}$;

(ii)

$$\bar{\lambda}I = \bar{\lambda} \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = \begin{bmatrix} \bar{\lambda} & 0 & \dots & 0 \\ 0 & \bar{\lambda} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \bar{\lambda} \end{bmatrix}_{n \times n};$$

(iii)

$$\bar{\lambda}I - \tilde{A} = \begin{bmatrix} \bar{\lambda} & 0 & \dots & 0 \\ 0 & \bar{\lambda} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \bar{\lambda} \end{bmatrix} - \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix} = \begin{bmatrix} \bar{\lambda}I_k - A_1 & -A_2 \\ 0 & \bar{\lambda}I_{(n-k)} - A_3 \end{bmatrix}.$$

Logo,

$$\begin{aligned} w^*(\bar{\lambda}I - \tilde{A}) &= \begin{bmatrix} 0 & v^* \end{bmatrix} \cdot \begin{bmatrix} \bar{\lambda}I_k - A_1 & -A_2 \\ 0 & \bar{\lambda}I_{(n-k)} - A_3 \end{bmatrix} \\ &= \begin{bmatrix} 0 \cdot (\bar{\lambda}I_k - A_1) + v^* \cdot 0 & 0 \cdot (-A_2) + v^*(\bar{\lambda}I_{(n-k)} - A_3) \end{bmatrix} \\ &= \begin{bmatrix} 0 & v^*(\bar{\lambda}I_{(n-k)} - A_3) \end{bmatrix} \end{aligned}$$

Conforme verificado, $v^*(\bar{\lambda}I - A_3) = 0$ e então

$$w^*(\bar{\lambda}I - \tilde{A}) = 0. \quad (1.18)$$

Além disto

$$w^*B = \begin{bmatrix} 0 & v^* \end{bmatrix} \cdot \begin{bmatrix} B_1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \cdot B_1 + v^* \cdot 0 \end{bmatrix} = 0.$$

Obviamente $S^* \in GL(n, \mathbb{R})$ (pois $\det S = \det S^*$). Relembre ainda que $w = (0, v) \in \mathbb{C}^n$ e por construção, $v \neq 0$ de forma que $w \neq 0$. Sendo assim, $z := (S^*)^{-1}w \neq 0$. De fato, se $(S^*)^{-1}w = 0$, multiplicando ambos os membros desta igualdade por S^* obtemos $w = 0$, um absurdo. Desse modo, usando (1.18) temos

$$\begin{aligned} z^*(\bar{\lambda}I - A)S &= ((S^*)^{-1}w)^*(\bar{\lambda}I - A)S = w^*S^{-1}(\bar{\lambda}I - A)S = w^*(S^{-1}\bar{\lambda}IS - S^{-1}AS) \\ &= w^*(\bar{\lambda}S^{-1}IS - \tilde{A}) = w^*(\bar{\lambda}I - \tilde{A}) = 0 \end{aligned}$$

e multiplicando esta última igualdade à direita por S^{-1} , obtemos,

$$z^*(\bar{\lambda}I - A) = 0 .$$

Além disto,

$$z^*B = w^*S^{-1}B = w^*\tilde{B} = 0 .$$

Conseqüentemente, $(z^*(\bar{\lambda}I - A), z^*B) = 0$, ou melhor,

$$z^*(\bar{\lambda}I - A, B) = 0 .$$

Daí, $(\bar{\lambda}I - A, B)^*z = 0$, ou seja, $z \in \ker(\bar{\lambda}I - A, B)^*$ e sendo $z \neq 0$ concluímos que $\dim \ker(\bar{\lambda}I - A, B)^* > 0$. Pelo teorema do núcleo e da imagem,

$$Posto (\bar{\lambda}I - A, B)^* + \dim \ker(\bar{\lambda}I - A, B)^* = n$$

$$\Rightarrow Posto (\bar{\lambda}I - A, B)^* = n - \dim \ker(\bar{\lambda}I - A, B)^* < n .$$

Desde que $Posto (\bar{\lambda}I - A, B) = Posto (\bar{\lambda}I - A, B)^*$ temos

$$Posto (\bar{\lambda}I - A, B) < n ,$$

ou seja, $Posto (sI - A, B) < n$ para algum $s \in \mathbb{C}$, como queríamos demonstrar. ■

1.2 Observabilidade

Nesta seção trataremos do problema da observabilidade, introduzido de maneira informal no início do capítulo. Dados, $x_0 \in \mathbb{R}^n$, $T > 0$ e $u \in PC([0, T], \mathbb{R}^m)$, denotamos a saída do sistema (1.1) em $[0, T]$ por

$$y(t; x_0, u) = Cx(t; x_0, u) = Ce^{At}x_0 + C \int_0^t e^{A(t-s)}Bu(s)ds, \quad \forall t \in [0, T].$$

Note primeiramente que a expressão acima deixa clara dependência da saída em relação ao estado inicial x_0 e ao controle u . No problema da observabilidade, assumiremos os seguintes dados:

- a) (A,B,C) : O sistema 1.1;
- b) $u \in PC([0, T], \mathbb{R}^m)$: o controle;
- c) $y(\cdot; x_0, u) \in C([0, T], \mathbb{R}^p)$: a saída.

Como o estado inicial $x_0 \in \mathbb{R}^n$ é desconhecido, $x(t) = x(t; x_0, u)$, $t \in [0, T]$ pode não ser, em geral, calculado a partir dos dados fornecidos. A observabilidade investiga as circunstâncias sob as quais tal cálculo é possível. Em particular, ela procura determinar condições sob as quais o estado inicial x_0 é determinado de modo único a partir dos dados fornecidos. Mas se determinarmos x_0 , como já temos u , podemos enfim determinarmos o estado $x(t) = x(t; x_0, u)$. Portanto, podemos dizer que a observabilidade busca condições sob as quais o estado $x(t) = x(t; x_0, u)$, $t \in [0, T]$, é determinado a partir dos dados fornecidos.

Introduzimos o *mapa estado-para-saída* $O_T : \mathbb{R}^n \rightarrow C([0, T], \mathbb{R}^p)$, parametrizado por $T > 0$, dado por

$$(O_T x_0)(t) := C e^{At} x_0, \quad \forall t \in [0, T].$$

Mais adiante veremos que, para um dado $T > 0$, o estado inicial poderá ser determinado avaliando-se $O_T^\# O_T x_0$, onde $O_T^\#$ denota uma inversa à esquerda de O_T . Uma vez que x_0 tiver sido determinado, o estado $x(t; x_0, u)$, para $t \in [0, T]$, pode ser determinado através da fórmula de variação dos parâmetros. Claro que para isso é necessário que exista de fato uma inversa $O_T^\#$. Por este motivo, a observabilidade do sistema (1.1) será identificado com a invertibilidade à esquerda de O_T , qualquer que seja $T > 0$.

Resulta desta definição que

$$(O_T x_0)(t) = C e^{At} x_0 + \int_0^t e^{A(t-s)} B \cdot 0 \, ds = y(t; x_0, 0), \quad \forall t \in [0, T].$$

Ressaltamos o uso da notação $O_T x_0$ para indicar $O_T(x_0)$. Perceba ainda que o mapa O_T é linear. Com efeito, dados $x_0, x_1 \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$ temos

$$\begin{aligned} (O_T(\alpha x_0 + x_1))(t) &= C e^{At}(\alpha x_0 + x_1) = C e^{At} \alpha x_0 + C e^{At} x_1 = \alpha C e^{At} x_0 + C e^{At} x_1 \\ &= \alpha(O_T x_0)(t) + (O_T x_1)(t). \end{aligned}$$

Além disso, nota-se que para todo $t \in [0, T]$

$$\begin{aligned}
(O_T x_0)(t) &= C e^{At} x_0 = C e^{At} x_0 + \int_0^t e^{A(t-s)} B u(s) \, ds - \int_0^t e^{A(t-s)} B u(s) \, ds \\
&= C e^{At} x_0 + \int_0^t e^{A(t-s)} B u(s) \, ds - C e^{At} \cdot 0 - \int_0^t e^{A(t-s)} B u(s) \, ds \\
&= C e^{At} x_0 + \int_0^t e^{A(t-s)} B u(s) \, ds - \left(C e^{At} \cdot 0 + \int_0^t e^{A(t-s)} B u(s) \, ds \right)
\end{aligned}$$

e então

$$(O_T x_0)(t) = y(t; x_0, u) - y(t; 0, u) \quad (1.19)$$

Isto nos diz que, para qualquer $x_0 \in \mathbb{R}^n$, $O_T x_0$ é calculável a partir dos dados fornecidos.

Uma aplicação $O_T^\# : C([0, T], \mathbb{R}^p) \rightarrow \mathbb{R}^n$ é uma inversa à esquerda de O_T se,

$$O_T^\#(O_T x_0) = x_0, \quad \forall x_0 \in \mathbb{R}^n.$$

Decorre imediatamente da Proposição A.3 a seguinte afirmação:

Afirmção 1.2 *Uma inversa à esquerda de O_T existe se, e somente se, O_T é injetiva.*

Além do mais, da definição de O_T , temos

$$O_T x_0 = y(\cdot; x_0, 0).$$

Se existe uma inversa à esquerda $O_T^\#$ de O_T , resulta que

$$O_T^\#(O_T x_0) = O_T^\#(y(\cdot; x_0, 0)),$$

ou seja,

$$x_0 = O_T^\#(y(\cdot; x_0, 0))$$

e, uma vez que x_0 é determinado, o estado $x(t; x_0, u)$ pode ser calculado para todo $t \in [0, T]$ via fórmula de variação dos parâmetros A.15. Concluimos assim que, **a questão da observabilidade pode ser identificada com a da invertibilidade à esquerda ou, equivalentemente, a injetividade da aplicação O_T** . Observe que $y(\cdot; x_0, 0)$ é calculável a partir dos dados fornecidos, pois

$$y(\cdot; x_0, 0) = O_T x_0$$

e, por (1.19),

$$y(\cdot; x_0, 0) = y(\cdot; x_0, u) - y(\cdot; 0, u).$$

Proposição 1.6 *Sejam $T > 0$ e $x_0 \in \mathbb{R}^n$. Então, $O_T x_0 = 0$ se, e somente se, $Ce^{At}x_0 = 0$ para todo $t \in \mathbb{R}_+$.*

Demonstração:

(\Leftarrow) Com efeito, dado $x_0 \in \mathbb{R}^n$, se $Ce^{At}x_0 = 0$ para todo $t \in \mathbb{R}_+$, então em particular $Ce^{At}x_0 = 0$ para cada $t \in [0, T]$ qualquer que seja $T > 0$, ou seja, $O_T x_0 = 0$ qualquer que seja $T > 0$.

(\Rightarrow) Para mostrar a primeira implicação, suponha que $O_T x_0 = 0$. Daí, $Ce^{At}x_0 = 0$ para todo $t \in [0, T]^1$. Derivando esta expressão repetidas vezes, obtemos

$$\begin{aligned} CA e^{At}x_0 &= 0 \\ CA^2 e^{At}x_0 &= 0 \\ CA^3 e^{At}x_0 &= 0 \\ &\vdots \\ CA^k e^{At}x_0 &= 0 \end{aligned}$$

para $t \in [0, T]$ e $k \in \mathbb{N}$. Fazendo $t = 0$ em $Ce^{At}x_0 = 0$ e em suas derivadas sucessivas, obtemos,

$$\begin{aligned} 0 &= CA^0 x_0 \\ &= CA x_0 \\ &= CA^2 x_0 \\ &= CA^3 x_0 \\ &\vdots \\ &= CA^k x_0, \end{aligned}$$

isto é, $CA^k x_0 = 0$ para todo $k \in \mathbb{N}_0$. Ora, para $t \in \mathbb{R}$ usando a Definição A.7 obtemos,

$$Ce^{At}x_0 = C \sum_{k=0}^{\infty} \frac{(At)^k}{k!} x_0 = \sum_{k=0}^{\infty} \frac{t^k}{k!} \underbrace{CA^k x_0}_{=0} = 0$$

ou seja, $Ce^{At}x_0 = 0$ para todo $t \in \mathbb{R}$. ■

Proposição 1.7 *O_T é injetiva se, e somente se, a seguinte implicação ocorre*

$$O_T x_0 = 0 \Rightarrow x_0 = 0. \tag{1.20}$$

¹Até então temos $Ce^{At}x_0 = 0$ garantido para $t \in [0, T]$. Queremos assegurar para qualquer $t \in \mathbb{R}_+$.

Demonstração: (\Rightarrow) Se O_T é injetiva devido à sua linearidade temos, pela Proposição A.3, $\ker O_T = \{0\}$ e então

$$O_T x_0 = 0 \Rightarrow x_0 \in \ker\{O_T\} \Rightarrow x_0 = 0 ,$$

ou seja,

$$O_T x_0 = 0 \Rightarrow x_0 = 0.$$

(\Leftarrow) Analogamente sendo O_T linear, se (1.20) ocorre temos que $\ker O_T = \{0\}$ e pela Proposição A.3 segue que O_T é injetiva. ■

Vimos anteriormente que, qualquer que seja $T > 0$, $O_T x_0 = 0$ equivale ao fato de que $Ce^{At}x_0 = 0$ para todo $t \in \mathbb{R}_+$. Daí a implicação (1.20), pode ser escrita da seguinte forma

$$y(t; x_0, 0) = Ce^{At}x_0 = 0 , \forall t \in \mathbb{R}_+ \Rightarrow x_0 = 0.$$

Toda a discussão acima sobre observabilidade serve de base para a próxima definição e para estabelecer a proposição posterior.

Definição 1.3 *O sistema (1.1) é dito observável se, para todo $x_0 \in \mathbb{R}^n$, têm-se*

$$\left(y(t; x_0, 0) = Ce^{At}x_0 = 0 , \forall t \in \mathbb{R}_+ \right) \Rightarrow x_0 = 0. \quad (1.21)$$

Proposição 1.8 *As seguintes sentenças são equivalentes:*

- (1) *O sistema (1.1) é observável.*
- (2) *O_T é injetiva para algum $T > 0$.*
- (3) *O_T é injetiva para todo $T > 0$.*
- (4) *O_T é invertível à esquerda para algum $T > 0$.*
- (5) *O_T é invertível à esquerda para todo $T > 0$.*

Demonstração:

(1) \Rightarrow (2). Se o sistema (1.1) é observável, então de acordo com a Definição 1.3 temos,

$$\left(y(t; x_0, 0) = Ce^{At}x_0 = 0 , \forall t \in \mathbb{R}_+ \right) \Rightarrow x_0 = 0.$$

Fixado $T > 0$, pela Proposição 1.6, a expressão

$$\left(y(t; x_0, 0) = Ce^{At}x_0 = 0 , \forall t \in \mathbb{R}_+ \right)$$

equivale a

$$\left(y(t; x_0, 0) = Ce^{At}x_0 = 0, \forall t \in [0, T] \right).$$

Assim, a implicação acima pode ser reescrita da seguinte maneira

$$\left(y(t; x_0, 0) = Ce^{At}x_0 = 0, \forall t \in [0, T] \right) \Rightarrow x_0 = 0.$$

Mas por definição, para $t \in [0, T]$, $Ce^{At}x_0 = O_T x_0$. Logo, podemos reescrever a última implicação da maneira abaixo

$$O_T x_0 = 0 \Rightarrow x_0 = 0,$$

qualquer que seja $T > 0$. Assim, do que vimos anteriormente, isto implica que O_T é injetiva para todo $T > 0$ e conseqüentemente, O_T é injetiva para algum $T > 0$.

(2) \Rightarrow (3). Seja $T > 0$ tal que O_T é injetiva. Então pela Proposição 1.7

$$O_T x_0 = 0 \Rightarrow x_0 = 0.$$

Pela Proposição 1.6, $O_T x_0 = 0$ equivale a

$$Ce^{At}x_0 = 0 \quad \forall t \in \mathbb{R}_+.$$

Logo

$$Ce^{At}x_0 = 0 \Rightarrow x_0 = 0, \forall t \in \mathbb{R}_+.$$

Assim, qualquer que seja $T > 0$ temos,

$$Ce^{At}x_0 = 0 \Rightarrow x_0 = 0 \quad \forall t \in [0, T].$$

Logo, pela definição de O_T , para todo $T > 0$, temos

$$O_T x_0 = 0 \Rightarrow x_0 = 0$$

e então pela Proposição 1.7, O_T é injetiva para todo $T > 0$.

(3) \Rightarrow (4). Se O_T é injetiva para todo $T > 0$, pela Afirmação 1.2 resulta que O_T é invertível à esquerda para todo $T > 0$. Imediatamente, O_T é invertível à esquerda para algum $T > 0$.

(4) \Rightarrow (5). Suponha que O_T seja invertível à esquerda para algum $T > 0$. Então, pela Afirmação 1.2, O_T é injetiva para algum $T > 0$. Como já temos (2) \Rightarrow (3), resulta

que O_T é injetiva para todo $T > 0$ e usando novamente a Afirmação 1.2, concluímos que O_T é invertível à esquerda qualquer que seja $T > 0$.

(5) \Rightarrow (1). Assuma que O_T seja invertível à esquerda para todo $T > 0$. Assim, pela Afirmação 1.2, O_T é injetiva para todo $T > 0$ e então pela Proposição 1.7 temos que

$$O_T x_0 = 0 \Rightarrow x_0 = 0, \quad \forall T > 0.$$

Mas, pela Proposição 1.6, $O_T x_0 = 0$ para algum $T > 0$ equivale a

$$C e^{At} x_0 = 0 \quad \forall t \in \mathbb{R}_+.$$

Logo, a implicação anterior pode ser reescrita da maneira que segue:

$$\left(C e^{At} x_0 = 0, \quad \forall t \in \mathbb{R}_+ \right) \Rightarrow x_0 = 0,$$

ou ainda,

$$\left(y(t; x_0, 0) = C e^{At} x_0 = 0, \quad \forall t \in \mathbb{R}_+ \right) \Rightarrow x_0 = 0$$

o que nos mostra, de acordo com a Definição 1.3, que o sistema (1.1) é observável. \blacksquare

Observe que na Definição 1.3 a matriz B não desempenha nenhum papel, isto é, a observabilidade do sistema (1.1) depende apenas das matrizes A e C . Assim, é conveniente definir um par de matrizes $(C, A) \in \mathbb{R}^{p \times n} \times \mathbb{R}^{n \times n}$ ser observável se, para todo $x_0 \in \mathbb{R}^n$, a implicação (1.21) acontece. O próximo resultado é uma consequência imediata da Definição 1.3.

Lema 1.3 *Defina,*

$$U := \{x_0 \in \mathbb{R}^n : C e^{At} x_0 = 0 \quad \forall t \in \mathbb{R}_+\}. \quad (1.22)$$

O par (C, A) é observável se, e somente se, $U = \{0\}$.

Com uma perspectiva para convenientes caracterizações do conjunto U , introduzimos a *Matriz Observabilidade* definida por

$$\mathcal{O}(C, A) := \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix} \in \mathbb{R}^{(pn) \times n}$$

e o *Gramian de Observabilidade* (parametrizado por $T > 0$) definido por

$$S_T := \int_0^T e^{A^*t} C^* C e^{At} dt \in \mathbb{R}^{n \times n}.$$

Note de início que S_T é simétrica e positiva semi-definida. De fato,

$$S_T^* = \left(\int_0^T e^{A^*t} C^* C e^{At} dt \right)^* = \int_0^T (e^{A^*t} C^* C e^{At})^* dt = \int_0^T e^{A^*t} C^* C e^{At} dt = S_T.$$

Agora, dado $x \in \mathbb{R}^n$ temos:

$$x^* S_T x = x^* \int_0^T e^{A^*t} C^* C e^{At} dt x = \int_0^T x^* e^{A^*t} C^* C e^{At} x dt = \int_0^T (C e^{At} x)^* C e^{At} x dt.$$

Recorde que $C \in \mathbb{R}^{p \times n}$, $e^{At} \in \mathbb{R}^{n \times n}$ e $x \in \mathbb{R}^{n \times 1}$ (lembre-se de que em todo o trabalho, um vetor do \mathbb{R}^n será representado ou por uma n -upla ordenada ou por uma matriz coluna de n linhas). Logo, $C e^{At} \in \mathbb{R}^{p \times n}$ e $C e^{At} x \in \mathbb{R}^{p \times 1}$. Façamos,

$$C e^{At} x = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix}.$$

Daí,

$$(C e^{At} x)^* = [c_1 \ c_2 \ \cdots \ c_p]$$

e por conseguinte,

$$(C e^{At} x)^* C e^{At} x = [c_1 \ c_2 \ \cdots \ c_p] \cdot \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix} = [c_1^2 + c_2^2 + \cdots + c_p^2] = k \geq 0.$$

Portanto,

$$x^* S_T x = \int_0^T k dt = kt \Big|_0^T = kT \geq 0$$

mostrando-nos que S_T é positiva semi-definida.

Teorema 1.5 Para todo $T > 0$, $\ker \mathcal{O}(C, A) = \ker S_T$.

Demonstração: De início, note que

$$(\mathcal{O}(C, A))^* = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}^* = \begin{bmatrix} C^* & (CA)^* & (CA^2)^* & \cdots & (CA^{n-1})^* \end{bmatrix} =$$

$$= \begin{bmatrix} C^* & A^*C^* & (A^*)^2C^* & \dots & (A^*)^{n-1}C^* \end{bmatrix}.$$

Recorde que, para $A \in \mathbb{R}^{n \times n}$ e $B \in \mathbb{R}^{n \times p}$ definimos a matriz de Controlabilidade para o par (A, B) por

$$C(A, B) = [B \ AB \ A^2B \ \dots \ A^{n-1}B]$$

e o Gramian de Controlabilidade por

$$Q_T = \int_0^T e^{At} B B^* e^{A^*t} dt.$$

Em nosso caso, tomando $A = A^* \in \mathbb{R}^{n \times n}$ e $B = C^* \in \mathbb{R}^{n \times p}$, concluímos que

$$(\mathcal{O}(C, A))^* = C(A^*, C^*)$$

e mais,

$$Q_T = \int_0^T e^{At} B B^* e^{A^*t} dt = \int_0^T e^{A^*t} C^* C e^{At} dt = S_T$$

ou seja, S_T é o Gramian de controlabilidade associado ao par (A^*, C^*) . Assim, pelo Teorema 1.1,

$$\text{Im } S_T = \text{Im } C(A^*, C^*)$$

ou, equivalentemente,

$$(\text{Im } S_T)^\perp = (\text{Im } C(A^*, C^*))^\perp. \quad (1.23)$$

Seja $z \in \mathbb{R}^n$. Suponha que $z \in \ker \mathcal{O}(C, A)$. Com isto,

$$\mathcal{O}(C, A)z = 0 \Rightarrow z^* (\mathcal{O}(C, A))^* = 0 \Rightarrow z^* C(A^*, C^*) = 0$$

o que nos diz, pelo Teorema A.7 que $z \in (\text{Im } C(A^*, C^*))^\perp$ e então pela igualdade (1.23) resulta que $z \in (\text{Im } S_T)^\perp$. Daí, usando novamente o Teorema A.7 temos,

$$z^* S_T = 0$$

ou, equivalentemente, $S_T^* z = 0$ e como S_T é simétrica, isto é, $S_T^* = S_T$ obtemos

$$S_T z = 0,$$

ou seja, $z \in \ker S_T$. Logo,

$$\ker \mathcal{O}(C, A) \subset \ker S_T.$$

Suponha agora que $z \in \ker S_T$. Então,

$$S_T z = 0 \Rightarrow z^* S_T^* = 0$$

e pela simetria de S_T temos ainda,

$$z^* S_T = 0 .$$

Logo, pelo Teorema A.7, $z \in (\text{Im } S_T)^\perp$ e assim pela igualdade (1.23), $z \in (\text{Im } C(A^*, C^*))^\perp$.

Daí, usando novamente o Teorema A.7 temos,

$$z^* C(A^*, C^*) = 0 \Rightarrow z^* (\mathcal{O}(C, A))^* = 0 \Rightarrow \mathcal{O}(C, A)z = 0,$$

ou seja, $z \in \ker \mathcal{O}(C, A)$ e então

$$\ker S_T \subset \ker \mathcal{O}(C, A).$$

Dessas duas últimas inclusões aqui verificadas, concluimos que

$$\ker \mathcal{O}(C, A) = \ker S_T.$$

■

Corolário 1.3 Para cada $T > 0$ existe $S_T^\# \in \mathbb{R}^{n \times n}$ tal que

$$S_T^\# S_T z = z \quad , \quad \forall z \in (\ker S_T)^\perp = (\ker \mathcal{O}(C, A))^\perp . \quad (1.24)$$

Demonstração: Com efeito, dado $T > 0$, $S_T \in \mathbb{R}^{n \times n}$ e como $n \geq n$, pelo Teorema A.8, existe $S_T^\# \in \mathbb{R}^{n \times n}$ tal que

$$S_T^\# S_T z = z \quad \forall z \in (\ker S_T)^\perp .$$

Mas pelo Teorema 1.5, $\ker S_T = \ker \mathcal{O}(C, A)$ o que resulta no fato de que $(\ker S_T)^\perp = (\ker \mathcal{O}(C, A))^\perp$. Logo, podemos reescrever a afirmação acima na forma dada em (1.24), isto é,

$$S_T^\# S_T z = z \quad , \quad \forall z \in (\ker S_T)^\perp = (\ker \mathcal{O}(C, A))^\perp .$$

■

Introduzimos a aplicação

$$\begin{aligned} O_T^\# : C([0, T], \mathbb{R}^p) &\longrightarrow \mathbb{R}^n \\ w &\longmapsto S_T^\# \int_0^T e^{A^* t} C^* w(t) dt . \end{aligned} \quad (1.25)$$

Recorde que para qualquer $x_0 \in \mathbb{R}^n$, $O_T x_0 \in C([0, T], \mathbb{R}^p)$. Então,

$$\begin{aligned} O_T^\#(O_T x_0) &= S_T^\# \int_0^T e^{A^*t} C^* O_T x_0 dt = S_T^\# \int_0^T e^{A^*t} C^* C e^{At} x_0 dt \\ &= S_T^\# \underbrace{\int_0^T e^{A^*t} C^* C e^{At} dt}_{S_T} x_0 \\ &= S_T^\# S_T x_0 . \end{aligned}$$

Em particular, tomando $x_0 \in (\ker S_T)^\perp \subset \mathbb{R}^n$, temos por (1.24) que

$$O_T^\#(O_T x_0) = S_T^\# S_T x_0 = x_0,$$

ou seja,

$$O_T^\#(O_T x_0) = x_0 \quad , \quad \forall x_0 \in (\ker S_T)^\perp = (\ker \mathcal{O}(C, A))^\perp . \quad (1.26)$$

Proposição 1.9 *Se Posto $\mathcal{O}(C, A) = n$, então O_T é invertível à esquerda para todo $T > 0$.*

Demonstração: Pelo Teorema do Núcleo e da Imagem, sendo $\mathcal{O}(C, A) \in \mathbb{R}^{(pn) \times n}$,

$$\text{Posto } \mathcal{O}(C, A) + \dim \ker \mathcal{O}(C, A) = n,$$

ou seja,

$$\dim \ker \mathcal{O}(C, A) = n - \text{Posto } \mathcal{O}(C, A).$$

Suponha que $\text{Posto } \mathcal{O}(C, A) = n$. Então,

$$\dim \ker \mathcal{O}(C, A) = n - n \Rightarrow \dim \ker \mathcal{O}(C, A) = 0.$$

Isto nos diz que,

$$\ker \mathcal{O}(C, A) = \{0\}.$$

Lembre-se que, $(\ker \mathcal{O}(C, A))^\perp \subset \mathbb{R}^n$. Agora, pelo Teorema A.10 podemos escrever

$$\mathbb{R}^n = (\ker \mathcal{O}(C, A))^\perp \oplus \ker \mathcal{O}(C, A) .$$

Seja $z \in \mathbb{R}^n$. Pela Proposição A.1, existem únicos z_1 e z_2 com $z_1 \in (\ker \mathcal{O}(C, A))^\perp$ e $z_2 \in \ker \mathcal{O}(C, A)$ tais que

$$z = z_1 + z_2 .$$

Mas se $z_2 \in \ker \mathcal{O}(C, A)$ resulta que $z_2 = 0$ e, portanto,

$$z = z_1 \in (\ker \mathcal{O}(C, A))^\perp,$$

mostrando que $\mathbb{R}^n \subset (\ker \mathcal{O}(C, A))^\perp$. Como já tínhamos a inclusão inversa, concluímos que

$$(\ker \mathcal{O}(C, A))^\perp = \mathbb{R}^n. \quad (1.27)$$

Pela igualdade (1.26), $O_T^\#(O_T x_0) = x_0$ para todo $x_0 \in (\ker \mathcal{O}(C, A))^\perp$. Logo, por (1.27),

$$O_T^\#(O_T x_0) = x_0, \quad \forall x_0 \in \mathbb{R}^n,$$

mostrando-nos que O_T é invertível à esquerda e $O_T^\#$ é uma inversa à esquerda, qualquer que seja $T > 0$. ■

Tendo em vista (1.26) no contexto da questão da observabilidade, podemos concluir que, se o estado inicial x_0 está em $(\ker \mathcal{O}(C, A))^\perp$, então x_0 pode ser calculado através da função conhecida $y := O_T x_0$, avaliando-se $O_T^\# y$. Neste sentido, $(\ker \mathcal{O}(C, A))^\perp$ é o *Subespaço Observável*. O próximo resultado estabelece que o conjunto U , definido por (1.22) coincide com $\ker \mathcal{O}(C, A)$, o complemento ortogonal do subespaço observável. No que segue, U é referido como o *Subespaço Não Observável*.

Teorema 1.6 *Seja U o conjunto definido em (1.22). Então,*

$$\ker \mathcal{O}(C, A) = U.$$

Demonstração: Seja $x_0 \in \mathbb{R}^n$. Se $x_0 \in \ker \mathcal{O}(C, A)$, então $\mathcal{O}(C, A)x_0 = 0$, isto é,

$$\begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix} x_0 = 0 \Rightarrow \begin{bmatrix} Cx_0 \\ CAx_0 \\ CA^2x_0 \\ \vdots \\ CA^{n-1}x_0 \end{bmatrix} = 0,$$

ou seja, $CA^k x_0 = 0$, para $k = 0, 1, 2, \dots, n-1$. Consideremos agora o caso em que $k = n$, isto é, vamos analisar a expressão $CA^n x_0$. Pelo Teorema de Cayley-Hamilton,

$$A^n = \alpha_{n-1}A^{n-1} + \alpha_{n-2}A^{n-2} + \dots + \alpha_1 A + \alpha_0 A^0, \quad \alpha_i \in \mathbb{R}, \quad i = 0, 1, 2, \dots, n-1.$$

Logo,

$$\begin{aligned} CA^n x_0 &= C(\alpha_{n-1}A^{n-1} + \alpha_{n-2}A^{n-2} + \dots + \alpha_1 A + \alpha_0)x_0 \\ &= \alpha_{n-1} \underbrace{CA^{n-1}x_0}_0 + \alpha_{n-2} \underbrace{CA^{n-2}x_0}_0 + \dots + \alpha_1 \underbrace{CAx_0}_0 + \alpha_0 \underbrace{Cx_0}_0. \end{aligned}$$

Desde que $CA^k x_0 = 0$ para $k = 0, 1, 2, \dots, n - 1$, resulta desta última expressão que,

$$CA^n x_0 = 0.$$

Uma vez que $CA^k x_0 = 0$ para $k = 0, k = 1, k = n - 1$ e $k = n$, podemos concluir pelo princípio da indução finita que,

$$CA^k x_0 = 0 \quad \forall k \in \mathbb{N}_0.$$

Daí, dado $t \in \mathbb{R}_+$ temos pela Definição A.7

$$Ce^{At} x_0 = C \sum_{k=0}^{\infty} \frac{(At)^k}{k!} x_0 = \sum_{k=0}^{\infty} C \frac{A^k t^k}{k!} x_0 = \sum_{k=0}^{\infty} \frac{t^k}{k!} \underbrace{CA^k x_0}_0 = 0.$$

Logo,

$$Ce^{At} x_0 = 0 \quad \forall t \in \mathbb{R}_+,$$

o que nos diz que $x_0 \in U$ assegurando o fato de que

$$\ker \mathcal{O}(C, A) \subset U.$$

Suponha agora que $x_0 \in U$. Então,

$$Ce^{At} x_0 = 0, \quad \forall t \in \mathbb{R}_+.$$

Derivando esta expressão repetidas vezes vem

$$\begin{aligned} CAe^{At} x_0 &= 0 \\ CA^2 e^{At} x_0 &= 0 \\ CA^3 e^{At} x_0 &= 0 \\ &\vdots \\ CA^k e^{At} x_0 &= 0. \end{aligned}$$

Analisando a primeira expressão e suas derivadas para $t = 0$, obtemos,

$$\begin{aligned} Cx_0 &= CA^0 x_0 = 0 \\ CAx_0 &= 0 \\ CA^2 x_0 &= 0 \\ CA^3 x_0 &= 0 \\ &\vdots \\ CA^k x_0 &= 0. \end{aligned}$$

Podemos então concluir indutivamente que $CA^k x_0 = 0$ para todo $k \in \mathbb{N}_0$. Consequentemente,

$$\begin{bmatrix} Cx_0 \\ CAx_0 \\ CA^2x_0 \\ \vdots \\ CA^{n-1}x_0 \end{bmatrix} = 0 \Rightarrow \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix} x_0 = 0$$

ou seja, $\mathcal{O}(C, A)x_0 = 0$ mostrando-nos que $x_0 \in \ker \mathcal{O}(C, A)$ e portanto,

$$U \subset \ker \mathcal{O}(C, A).$$

Uma vez que já verificamos a inclusão inversa, concluimos que

$$\ker \mathcal{O}(C, A) = U.$$

■

A significância deste resultado é o fato de que, o conjunto U pode ser determinado simplesmente por métodos algébricos.

Formularemos agora um resultado que, junto com a Proposição 1.8, fornece uma contrapartida da observabilidade ao Teorema de controlabilidade 1.3 e que, através da equivalência de (1) e (2), fornece condições algébricas do posto para a observabilidade facilmente verificáveis.

Teorema 1.7 *As seguintes sentenças são equivalentes:*

- (1) *O sistema (1.1) é observável.*
- (2) *Posto $\mathcal{O}(C, A) = n$.*
- (3) *S_T é invertível para algum $T > 0$.*
- (4) *S_T é invertível para todo $T > 0$.*

Demonstração:

(1) \Rightarrow (2). Pelo Lema 1.3, o sistema (1.1) é observável se, e somente se, $U = \{0\}$, ou seja, (1) equivale a $U = \{0\}$. Mas pelo Teorema 1.6, $\ker \mathcal{O}(C, A) = U$. Assim, supondo que (C, A) é observável (que equivale a dizer que o sistema 1.1 é observável) resulta que

$\ker \mathcal{O}(C, A) = \{0\}$. Logo, $\dim \ker \mathcal{O}(C, A) = 0$. Como, $\mathcal{O}(C, A) \in \mathbb{R}^{(pn) \times n}$, temos pelo Teorema do Núcleo e da Imagem que

$$\text{Posto } \mathcal{O}(C, A) + \dim \ker \mathcal{O}(C, A) = n,$$

ou seja,

$$\text{Posto } \mathcal{O}(C, A) = n - \dim \ker \mathcal{O}(C, A)$$

e como $\dim \ker \mathcal{O}(C, A) = 0$ resulta que

$$\text{Posto } \mathcal{O}(C, A) = n .$$

(2) \Rightarrow (3). Lembre de início que, $\text{Posto } \mathcal{O}(C, A) = \text{Posto } (\mathcal{O}(C, A))^*$. Sabemos da demonstração do Teorema 1.5 que:

$$i) (\mathcal{O}(C, A))^* = C(A^*, C^*),$$

ii) O Gramian de Observabilidade S_T para o par (C, A) coincide com o Gramian de Controlabilidade para o par (A^*, C^*) .

Se $\text{Posto } \mathcal{O}(C, A) = n$, resulta que $\text{Posto } C(A^*, C^*) = n$. Mas pelo Teorema 1.3 isto implica que o Gramian de Controlabilidade para o par (A^*, C^*) é invertível para algum $T > 0$. Portanto S_T , é invertível para algum $T > 0$.

(3) \Rightarrow (4). Suponha que S_T é invertível para algum $T > 0$. Como S_T é Gramian de Controlabilidade para o par (A^*, C^*) resulta, pelo Teorema 1.3, que S_T é invertível para todo $T > 0$.

(4) \Rightarrow (1) Se S_T é invertível para todo $T > 0$, existe $S_T^\# \in \mathbb{R}^{n \times n}$ tal que $S_T^\# S_T z = z$ para todo $z \in \mathbb{R}^n$. Considere $O_T x_0$ dado por

$$O_T x_0 = C e^{At} x_0 \quad \forall t \in [0, T], \quad T > 0$$

e $O_T^\#$ definido por (1.25), isto é,

$$\begin{aligned} O_T^\# : C([0, T], \mathbb{R}^p) &\rightarrow \mathbb{R}^n \\ w &\longmapsto S_T^\# \int_0^T e^{A^*t} C^* w(t) dt. \end{aligned}$$

Note que $O_T x_0 \in C([0, T], \mathbb{R}^p)$ e,

$$\begin{aligned} O_T^\#(O_T x_0) &= S_T^\# \int_0^T e^{A^*t} C^* O_T x_0 dt = S_T^\# \int_0^T e^{A^*t} C^* C e^{At} x_0 dt \\ &= S_T^\# \underbrace{\int_0^T e^{A^*t} C^* C e^{At} dt}_{S_T} x_0 \\ &= S_T^\# S_T x_0 = x_0 \end{aligned}$$

ou seja,

$$O_T^\#(O_T x_0) = x_0,$$

o que nos diz que O_T é invertível à esquerda para todo $T > 0$. Logo, pela Proposição 1.8 o sistema (1.1) é observável. ■

Temos agora o seguinte cenário. Assuma que o sistema (1.1) é observável e seja $T > 0$. Assuma ainda que o estado inicial x_0 é desconhecido mas os dados (A, B, C) são conhecidos e que, para todo $t \in [0, T]$ o controle $u(t)$ e a saída $y; tx_0, u)$ são conhecidos. Recorde que

$$O_T x_0 = y(t; x_0, u) - y(t; 0, u).$$

Assim, supondo $u(t)$ e $y(t; x_0, u)$ conhecidos, por consequência $O_T x_0$ é também conhecido. Pela observabilidade de (1.1), segue do Teorema 1.6 em conjunção com o Lema 1.3 que $\ker \mathcal{O}(C, A) = \{0\}$. Resulta ainda, pelo Teorema 1.5 que $\ker S_T = \{0\}$. Como

$$\ker S_T \oplus (\ker S_T)^\perp = \mathbb{R}^n,$$

resulta então que

$$(\ker S_T)^\perp = \mathbb{R}^n.$$

Então, por (1.26), $O_T^\# O_T x_0 = x_0$ para todo $x_0 \in \mathbb{R}^n$, isto é, $O_T^\#$ dado por (1.25) com $S_T^\# = S_T^{-1}$, é uma inversa à esquerda da aplicação O_T . Assim, o estado inicial é determinado avaliando-se $O_T^\# O_T x_0$. Uma vez que x_0 tiver sido determinado, o estado $x(t; x_0, u)$, para $t \in [0, T]$, pode ser determinado através da fórmula de variação dos parâmetros.

Corolário 1.4 *Seja $(A, B, C) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m} \times \mathbb{R}^{p \times n}$.*

- 1) *O par (C, A) é observável se, e somente se, o par (A^*, C^*) é controlável.*
- 2) *Similarmente, o par (A, B) é controlável se, e somente se, o par (B^*, A^*) é observável.*

Demonstração:

1) (\Rightarrow) Assuma que o par (C, A) é observável. Pelo Teorema 1.7 temos,

$$\text{Posto } \mathcal{O}(C, A) = n \Rightarrow \text{Posto } (\mathcal{O}(C, A))^* = n \Rightarrow \text{Posto } C(A^*, C^*) = n .$$

Logo, pelo Teorema 1.3 resulta que o par (A^*, C^*) é controlável.

(\Leftarrow) Suponha agora que, (A^*, C^*) seja controlável. Pelo Teorema 1.3,

$$\text{Posto } C(A^*, C^*) = n \Rightarrow \text{Posto } (C(A^*, C^*))^* = n \Rightarrow \text{Posto } \mathcal{O}(C, A) = n$$

e então pelo Teorema 1.7 temos que o par (C, A) é observável.

2) (\Rightarrow) Se o par (A, B) é controlável, então pela parte 1) é imediato que o par (B^*, A^*) é observável.

(\Leftarrow) Reciprocamente, se (B^*, A^*) é observável, pela parte 1) resulta que (A, B) é controlável. ■

O corolário acima (que algumas vezes é referido pela expressão “Controlabilidade e Observabilidade são conceitos duais”) pode ser usado para obter resultados referentes à Observabilidade a partir de resultados correspondentes em Controlabilidade e vice versa. A prova do seguinte teorema é um exemplo típico.

Teorema 1.8 (Critério de Hautus para Observabilidade) *O sistema (1.1) é observável se, e somente se,*

$$\text{Posto } \begin{bmatrix} sI - A \\ C \end{bmatrix} = n.$$

Demonstração: (\Rightarrow) Se o sistema (1.1) ou, equivalentemente, o par (C, A) é observável então pelo Corolário 1.4 o par (A^*, C^*) é controlável. Daí, pelo critério de Hautus para a controlabilidade (Teorema 1.4),

$$\text{Posto } \begin{bmatrix} sI - A^* & C^* \end{bmatrix} = n, \forall s \in \mathbb{C}.$$

Desse modo, dado $\bar{s} \in \mathbb{C}$ temos,

$$\text{Posto } \begin{bmatrix} \bar{s}I - A^* & C^* \end{bmatrix} = n, \forall \bar{s} \in \mathbb{C}.$$

Como $\text{Posto } \begin{bmatrix} \bar{s}I - A^* & C^* \end{bmatrix} = \text{Posto } \begin{bmatrix} \bar{s}I - A^* & C^* \end{bmatrix}^*$ temos

$$\text{Posto } \begin{bmatrix} \bar{s}I - A^* & C^* \end{bmatrix}^* = n \Rightarrow \text{Posto } \begin{bmatrix} (\bar{s}I - A^*)^* \\ (C^*)^* \end{bmatrix} = n \Rightarrow$$

$$\Rightarrow \text{Posto} \begin{bmatrix} I^*(\bar{s})^* - A \\ C \end{bmatrix} = n \Rightarrow \text{Posto} \begin{bmatrix} sI - A \\ C \end{bmatrix} = n .$$

(\Leftarrow) Reciprocamente, suponha que $\text{Posto} \begin{bmatrix} sI - A \\ C \end{bmatrix} = n$ para todo $s \in \mathbb{C}$. Como $\bar{s} \in \mathbb{C}$ qualquer que seja $s \in \mathbb{C}$ temos,

$$\text{Posto} \begin{bmatrix} \bar{s}I - A \\ C \end{bmatrix} = n , \quad \forall s \in \mathbb{C} .$$

Daí,

$$\begin{aligned} \text{Posto} \begin{bmatrix} \bar{s}I - A \\ C \end{bmatrix} = n &\Rightarrow \text{Posto} \left(\begin{bmatrix} \bar{s}I - A \\ C \end{bmatrix} \right)^* = n \Rightarrow \text{Posto} \begin{bmatrix} (\bar{s}I - A)^* & C^* \end{bmatrix} = n \Rightarrow \\ &\Rightarrow \text{Posto} \begin{bmatrix} I^*(\bar{s})^* - A^* & C^* \end{bmatrix} = n \Rightarrow \text{Posto} \begin{bmatrix} sI - A^* & C^* \end{bmatrix} = n , \end{aligned}$$

para todo $s \in \mathbb{C}$. Pelo critério de Hautus para a controlabilidade, resulta que o par (A^*, C^*) é controlável e pelo Corolário 1.4 segue que o par (C, A) é observável.

Como já era de se imaginar, o resultado acima pode ser obtido de maneira semelhante àquela usada para o correspondente resultado da controlabilidade. Para tanto, prova-se o seguinte *Lema de Decomposição da Observabilidade*, numa perspectiva de contrapartida ao resultado similar da controlabilidade.

Lema 1.4 (Decomposição da Observabilidade de Kalman) *Assuma que o par $(C, A) \in \mathbb{R}^{p \times n} \times \mathbb{R}^{n \times n}$ não é observável e $C \neq 0$, caso no qual*

$$0 < k := \text{Posto } \mathcal{O}(C, A) = \dim \text{Im } \mathcal{O}(C, A) < n .$$

Então, existe $S \in GL(n, \mathbb{R})$ tal que $\tilde{A} := S^{-1}AS$ e $\tilde{C} := CS$ tem as seguintes estruturas de bloco

$$\tilde{A} = \begin{bmatrix} A_1 & 0_A \\ A_2 & A_3 \end{bmatrix} \quad e \quad \tilde{C} = [C_1 \quad 0_c]$$

onde $A_1 \in \mathbb{R}^{k \times k}$, $A_2 \in \mathbb{R}^{(n-k) \times k}$, $A_3 \in \mathbb{R}^{(n-k) \times (n-k)}$, $0_A \in \mathbb{R}^{k \times (n-k)}$, $C_1 \in \mathbb{R}^{p \times k}$ e $0_c \in \mathbb{R}^{p \times (n-k)}$ e o par (C_1, A_1) é observável.

Demonstração: Se o par (C, A) é não observável, então pelo Corolário 1.4 o par $(A^*, C^*) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times p}$ é não controlável, com $C^* \neq 0$. Note que,

$$k := \text{Posto } \mathcal{O}(C, A) = \text{Posto } (\mathcal{O}(C, A))^* = \text{Posto } C(A^*, C^*) .$$

Pelo Lema 1.2, existe $Q \in GL(n, \mathbb{R})$ tal que $\tilde{A}^* = Q^{-1}A^*Q$ e $\tilde{C}^* = Q^{-1}C^*$ possuem as seguintes estruturas de bloco,

$$\tilde{A}^* = \begin{bmatrix} A_1^* & A_2^* \\ 0_A^* & A_3^* \end{bmatrix} \quad \text{e} \quad \tilde{C}^* = \begin{bmatrix} C_1^* \\ 0_c^* \end{bmatrix}$$

onde $A_1^* \in \mathbb{R}^{k \times k}$, $A_2^* \in \mathbb{R}^{k \times (n-k)}$, $A_3^* \in \mathbb{R}^{(n-k) \times (n-k)}$, $0_A^* \in \mathbb{R}^{(n-k) \times k}$, $C_1^* \in \mathbb{R}^{k \times p}$ e $0_c^* \in \mathbb{R}^{(n-k) \times p}$.

Tomando as transpostas nas equações $\tilde{A}^* = Q^{-1}A^*Q$ e $\tilde{C}^* = Q^{-1}C^*$ obtemos

$$\tilde{A} = Q^*A(Q^*)^{-1} \quad \text{e} \quad \tilde{C} = C(Q^*)^{-1}.$$

Assim, basta tomar $S = (Q^*)^{-1}$, de onde $S^{-1} = Q^*$ e então obtemos

$$\tilde{A} = S^{-1}AS \quad \text{e} \quad \tilde{C} = CS,$$

que possuem as seguintes estruturas de bloco,

$$\tilde{A} = \begin{bmatrix} A_1 & 0_A \\ A_2 & A_3 \end{bmatrix} \quad \text{e} \quad \tilde{C} = [C_1 \quad 0_c]$$

as quais são obtidas tomando-se as transpostas na estruturas de \tilde{A}^* e \tilde{C}^* . Nestas últimas estruturas tem-se, $A_1 \in \mathbb{R}^{k \times k}$, $A_2 \in \mathbb{R}^{(n-k) \times k}$, $A_3 \in \mathbb{R}^{(n-k) \times (n-k)}$, $0_A \in \mathbb{R}^{k \times (n-k)}$, $C_1 \in \mathbb{R}^{p \times k}$ e $0_c \in \mathbb{R}^{p \times (n-k)}$. Além disto, pela Proposição 1.5 o par (A_1^*, C_1^*) é controlável e consequentemente, pelo Corolário 1.4 o par (C_1, A_1) é observável. ■

A cabo dos mais curiosos fica o exercício de demonstrar o *Critério de Hautus para a Observabilidade* via utilização do Lema 1.4.

Capítulo 2

Aplicações

2.1 O pêndulo invertido

Considere um pêndulo invertido e sem atrito com pivô P colocado no centro de um carrinho. Será possível exercer uma força sobre o carrinho de modo que o pêndulo fique equilibrado na vertical? Nesta seção iremos mostrar que a resposta a esta pergunta é sim. Assuma que o carrinho possui massa $M > 0$, está sujeito à uma força de controle horizontal u e pode se mover sem atrito sobre o plano horizontal. Seja $\theta(t)$ o ângulo entre a vertical e a haste (medido em sentido horário) no tempo $t \in J := \mathbb{R}$ e denotemos por $d(t)$ o deslocamento horizontal do carrinho (medido a partir de um ponto de referência fixado) no tempo t .

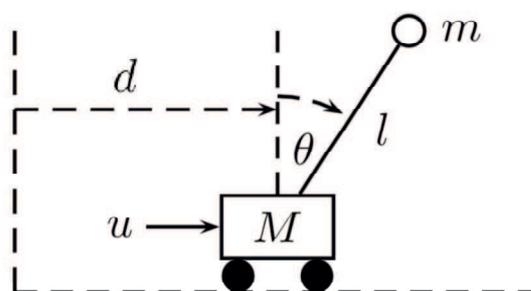


Figura 2.1: Pêndulo invertido

Tendo como referência a figura 2.1 podemos concluir que o deslocamento horizontal do pêndulo é dado por $d(t) + x_1$, ou melhor ainda por

$$d(t) + l \sin \theta,$$

onde l denota o comprimento da haste. Vamos expressar matematicamente a força horizontal resultante sobre o pêndulo.

(a) **Cálculo da força horizontal sobre o pêndulo devido ao movimento do carrinho**

Como o deslocamento horizontal do carrinho no tempo t é dado por $d(t)$, temos que a velocidade em tal tempo é $d'(t)$ e a aceleração é $d''(t)$. Indicando por $m > 0$ a massa do pêndulo, temos que a massa total sobre o carrinho é $M + m$. Disto resulta pela segunda lei de Newton que a força resultante horizontal sobre o carrinho é

$$(M + m)d''(t) .$$

(b) **Cálculo da força horizontal sobre o pêndulo devido ao seu deslocamento angular em torno do pivô P**

Como vimos, a componente horizontal do deslocamento angular do pêndulo no tempo t é dada por $l \sin \theta(t)$. Utilizando a regra da cadeia temos que a velocidade é dada por $l \cos \theta(t) \cdot \theta'(t)$. Agora aliando a regra da cadeia à regra do produto vem que a aceleração correspondente é

$$-l \sin \theta(t) \cdot \theta'(t) \cdot \theta'(t) + l \cos \theta(t) \cdot \theta''(t) = l(\theta''(t) \cos \theta(t) - [\theta'(t)]^2 \sin \theta(t))$$

e pela segunda lei de Newton temos que a força resultante correspondente é

$$ml(\theta''(t) \cos \theta(t) - [\theta'(t)]^2 \sin \theta(t)).$$

Podemos assim concluir que a força horizontal resultante sobre o pêndulo é

$$(M + m)d''(t) + ml(\theta''(t) \cos \theta(t) - [\theta'(t)]^2 \sin \theta(t)) .$$

É esta então, a força desempenhada pelo controle u no tempo t , isto é,

$$u(t) = (M + m)d''(t) + ml(\theta''(t) \cos \theta(t) - [\theta'(t)]^2 \sin \theta(t)) . \quad (2.1)$$

Por outro lado, o nosso objetivo é aplicar u ao sistema de modo a equilibrar o pêndulo na vertical. Desejamos assim que a aceleração horizontal $\vec{d}''(t)$ imposta ao carrinho produza uma aceleração inercial \vec{a}_i de mesma intensidade e sentido oposto sobre o pêndulo ocasionando no mesmo um deslocamento angular (no sentido anti-horário) com aceleração

tangencial de $l\theta''(t)$. Tendo como referência a figura 2.2, percebemos que nosso objetivo será alcançado se a aceleração $\vec{a}_{itg} + l\vec{\theta}''(t)$ for capaz de anular a componente tangencial da aceleração da gravidade o que ocorre se ambas as acelerações tiverem módulos iguais, isto é,

$$a_{itg} + l\theta''(t) = g_{tg} . \quad (2.2)$$

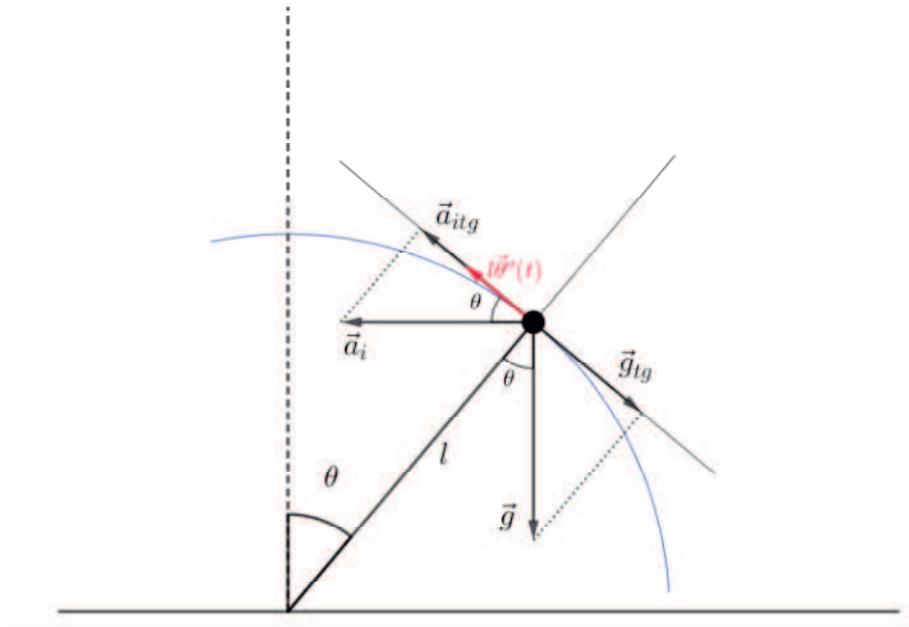


Figura 2.2: Dinâmica do pêndulo invertido

Recorde que $a = d''(t)$ e deste modo

$$\frac{a_{itg}}{a} = \cos \theta(t) \Rightarrow a_{itg} = a \cos \theta(t) \Rightarrow a_{itg} = d''(t) \cos \theta(t) .$$

Além disto,

$$\frac{g_{tg}}{g} = \text{sen} \theta(t) \Rightarrow g_{tg} = g \text{sen} \theta(t) .$$

Logo, a equação 2.2 pode ser reescrita como segue

$$d''(t) \cos \theta(t) + l\theta''(t) = g \text{sen} \theta(t) .$$

Devemos portanto ter

$$d''(t) \cos \theta(t) + l\theta''(t) - g \text{sen} \theta(t) = 0 . \quad (2.3)$$

Isolando $d''(t)$ em (2.3) obtêm-se

$$d''(t) = \frac{g \text{sen} \theta(t) - l\theta''(t)}{\cos \theta(t)} .$$

Substituindo isto em (2.1) vem

$$\begin{aligned}
u(t) &= (M + m) \left(\frac{g \operatorname{sen} \theta(t) - l \theta''(t)}{\cos \theta(t)} \right) + ml(\theta''(t) \cos \theta(t) - (\theta'(t))^2 \operatorname{sen} \theta(t)) \\
\Rightarrow u(t) \cos \theta(t) &= (M + m)g \operatorname{sen} \theta(t) - (M + m)l \theta''(t) + ml \theta''(t) \cos^2 \theta(t) \\
&\quad - ml(\theta'(t))^2 \cos \theta(t) \operatorname{sen} \theta(t) \\
\Rightarrow u(t) \cos \theta(t) &= \left[(M + m)g - ml(\theta'(t))^2 \cos \theta(t) \right] \operatorname{sen} \theta(t) \\
&\quad + \left[- (M + m) + m \cos^2 \theta(t) \right] l \theta''(t) \\
\Rightarrow \left[- (M + m) + m \cos^2 \theta(t) \right] l \theta''(t) &= - \left[(M + m)g - ml(\theta'(t))^2 \cos \theta(t) \right] \operatorname{sen} \theta(t) \\
&\quad + u(t) \cos \theta(t) \\
\Rightarrow \left[(M + m) - m \cos^2 \theta(t) \right] l \theta''(t) &= \left[(M + m)g - ml(\theta'(t))^2 \cos \theta(t) \right] \operatorname{sen} \theta(t) \\
&\quad - u(t) \cos \theta(t)
\end{aligned}$$

e portanto

$$\theta''(t) = \frac{\left[(M + m)g - ml(\theta'(t))^2 \cos \theta(t) \right] \operatorname{sen} \theta(t) - u(t) \cos \theta(t)}{l \left[(M + m) - m \cos^2 \theta(t) \right]} . \quad (2.4)$$

Logo,

$$d''(t) = \frac{g \operatorname{sen} \theta(t) - l \frac{\left[(M + m)g - ml(\theta'(t))^2 \cos \theta(t) \right] \operatorname{sen} \theta(t) - u(t) \cos \theta(t)}{l \left[(M + m) - m \cos^2 \theta(t) \right]}}{\cos \theta(t)} .$$

Vamos desenvolver o denominador da fração principal:

$$\begin{aligned}
&g \operatorname{sen} \theta(t) - l \frac{\left[(M + m)g - ml(\theta'(t))^2 \cos \theta(t) \right] \operatorname{sen} \theta(t) - u(t) \cos \theta(t)}{l \left[(M + m) - m \cos^2 \theta(t) \right]} \\
&= \frac{\left[(M + m) - m \cos^2 \theta(t) \right] g \operatorname{sen} \theta(t) - \left[(M + m)g - ml(\theta'(t))^2 \cos \theta(t) \right] \operatorname{sen} \theta(t)}{\left[(M + m) - m \cos^2 \theta(t) \right]} \\
&\quad + \frac{u(t) \cos \theta(t)}{\left[(M + m) - m \cos^2 \theta(t) \right]} \\
&= \frac{\left(\left[(M + m)g - mg \cos^2 \theta(t) \right] - \left[(M + m)g - ml(\theta'(t))^2 \cos \theta(t) \right] \right) \operatorname{sen} \theta(t)}{\left[(M + m) - m \cos^2 \theta(t) \right]} \\
&\quad + \frac{u(t) \cos \theta(t)}{\left[(M + m) - m \cos^2 \theta(t) \right]}
\end{aligned}$$

$$\begin{aligned}
&= \frac{\left((M+m)g - mg \cos^2 \theta(t) - (M+m)g + ml(\theta'(t))^2 \cos \theta(t) \right) \text{sen} \theta(t)}{\left[(M+m) - m \cos^2 \theta(t) \right]} \\
&+ \frac{u(t) \cos \theta(t)}{\left[(M+m) - m \cos^2 \theta(t) \right]} \\
&= \frac{\left(-mg \cos^2 \theta(t) + ml(\theta'(t))^2 \cos \theta(t) \right) \text{sen} \theta(t) + u(t) \cos \theta(t)}{\left[(M+m) - m \cos^2 \theta(t) \right]} \\
&= \frac{\left(l(\theta'(t))^2 - g \cos \theta(t) \right) m \cos \theta(t) \text{sen} \theta(t) + u(t) \cos \theta(t)}{\left[(M+m) - m \cos^2 \theta(t) \right]} .
\end{aligned}$$

Logo,

$$\begin{aligned}
d''(t) &= \frac{\left(l(\theta'(t))^2 - g \cos \theta(t) \right) m \cos \theta(t) \text{sen} \theta(t) + u(t) \cos \theta(t)}{\left[(M+m) - m \cos^2 \theta(t) \right]} \\
&= \frac{\left(l(\theta'(t))^2 - g \cos \theta(t) \right) m \cos \theta(t) \text{sen} \theta(t) + u(t) \cos \theta(t)}{\left[(M+m) - m \cos^2 \theta(t) \right] \cos \theta(t)} \\
&= \frac{\left(l(\theta'(t))^2 - g \cos \theta(t) \right) m \text{sen} \theta(t) + u(t)}{\left[(M+m) - m \cos^2 \theta(t) \right]} .
\end{aligned}$$

Portanto,

$$d''(t) = \frac{m \text{sen} \theta(t) \left(l(\theta'(t))^2 - g \cos \theta(t) \right) + u(t)}{M + m - m \cos^2 \theta(t)} . \quad (2.5)$$

Finalmente, o movimento do pêndulo sob a ação do controle u é governado pelo seguinte sistema de equações diferenciais de segunda ordem:

$$\left\{ \begin{aligned} \theta''(t) &= \frac{\left[(M+m)g - ml(\theta'(t))^2 \cos \theta(t) \right] \text{sen} \theta(t) - u(t) \cos \theta(t)}{l \left(M + m - m \cos^2 \theta(t) \right)} \\ d''(t) &= \frac{m \text{sen} \theta(t) \left(l(\theta'(t))^2 - g \cos \theta(t) \right) + u(t)}{M + m - m \cos^2 \theta(t)} . \end{aligned} \right. \quad (2.6)$$

No que segue, utilizaremos uma substituição de variáveis para escrever o sistema (2.6) como um sistema de equações diferenciais de segunda ordem. Fazemos pois

$$x_1(t) := \theta(t); \quad x_2(t) := \theta'(t); \quad x_3(t) := d(t) \quad \text{e} \quad x_4(t) = d'(t) . \quad (2.7)$$

Logo,

$$x'_1(t) = \theta'(t); \quad x'_2(t) = \theta''(t); \quad x'_3(t) = d'(t) \quad \text{e} \quad x'_4(t) = d''(t) .$$

Utilizando devidamente a substituição de variáveis definida, temos que

$$\left\{ \begin{array}{l} x'_1(t) = x_2(t) \\ x'_2(t) = \frac{\left[(M+m)g - mlx_2^2(t) \cos x_1(t) \right] \text{sen} x_1(t) - u(t) \cos x_1(t)}{l(M+m - m \cos^2 x_1(t))} \\ x'_3(t) = x_4(t) \\ x'_4(t) = \frac{m \text{sen} x_1(t) (lx_2^2(t) - g \cos x_1(t)) + u(t)}{M+m - m \cos^2 x_1(t)} \end{array} \right. \quad (2.8)$$

é o sistema de primeira ordem correspondente a (2.6) segundo a mencionada substituição. Considere as funções $f_2, f_4 : \mathbb{R}^3 \rightarrow \mathbb{R}$ dadas por

$$f_2(z_1, z_2, v) := \frac{\left[(M+m)g - mlz_2^2 \cos z_1 \right] \text{sen} z_1 - v \cos z_1}{l(M+m - m \cos^2 z_1)},$$

$$f_4(z_1, z_2, v) := \frac{m \text{sen} z_1 (lz_2^2 - g \cos z_1) + v}{M+m - m \cos^2 z_1}.$$

Seja $G := \mathbb{R}^4$. Para cada força u , associamos uma função $f_u : J \times G \rightarrow G$ definida por

$$f_u(t, z) = f_u(t, z_1, z_2, z_3, z_4) := \left(z_2, f_2(z_1, z_2, u(t)), z_4, f_4(z_1, z_2, u(t)) \right).$$

Definindo $x(t) := (x_1(t), x_2(t), x_3(t), x_4(t))$, com x_1, x_2, x_3 e x_4 dados em (2.7), o sistema (2.8) pode ser escrito da seguinte maneira

$$\left\{ \begin{array}{l} x'_1(t) = x_2(t) \\ x'_2(t) = f_2(x_1(t), x_2(t), u(t)) \\ x'_3(t) = x_4(t) \\ x'_4(t) = f_4(x_1(t), x_2(t), u(t)) \end{array} \right.$$

de onde tiramos que $x'(t) = \left(x_2(t), f_2(x_1(t), x_2(t), u(t)), x_4(t), f_4(x_1(t), x_2(t), u(t)) \right)$. Desse modo, o sistema (2.8) pode ser representado simplesmente por

$$x'(t) = f_u(t, x(t)). \quad (2.9)$$

Claramente cada escolha distinta de uma função u nos leva a uma distinta função f_u . Agora, é natural o surgimento da seguinte questão: é possível escolher uma função u tal que a solução do sistema (2.9) exibe um comportamento prescrito? Por exemplo, dada uma condição inicial

$$x(0) = (x_1(0), x_2(0), x_3(0), x_4(0))$$

do sistema (2.9), existe uma função u (um controle) tal que o pêndulo aproxima-se assintoticamente da posição vertical enquanto o carro se aproxima da posição de repouso 0? Simbolicamente, será que existe um controle u tal que

$$(\theta(t), \theta'(t), d(t), d'(t)) = (x_1(t), x_2(t), x_3(t), x_4(t)) = x(t) \rightarrow 0$$

quando $t \rightarrow \infty$? Este *problema de controle* pode ser enxergado naturalmente como o de equilibrar uma vassoura verticalmente sobre a palma da mão. Pode-se mostrar que, de fato existe um controle que cumpre estas condições. Para tanto, é útil considerar apenas “pequenos” desvios do pêndulo a partir da posição de repouso vertical ($(x_1(t), x_2(t), x_3(t), x_4(t)) = 0$ para todo t). Definindo $g : G \times \mathbb{R} \rightarrow G$ por

$$g(z, u) = f_u(t, z)$$

podemos escrever o sistema (2.9) da seguinte maneira

$$x'(t) = g(x(t), u(t)) . \tag{2.10}$$

A partir daqui, conduziremos nossos estudos com base nas aproximações lineares das componentes de $g(x, u)$ (que correspondem às aproximações lineares das componentes de f_u) calculadas nas origens de seus respectivos domínios. A saber, a *linearização* (caso venha a se desejar um melhor esclarecimento sobre linearização de sistemas não lineares, ficam como sugestões para tanto as referências [1] e [9]) sobre $(0, 0)$ do sistema (2.10) é o sistema linear $x' = Ax + Bu$ onde o par (A, B) é definido por

$$A := (D_1g)(0, 0) \in \mathbb{R}^{4 \times 4} \quad \text{e} \quad B := (D_2g)(0, 0) \in \mathbb{R}^{4 \times 1} ,$$

onde D_1g (respectivamente, D_2g) denota a matriz das derivadas parciais das componentes de g com relação às componentes do primeiro argumento (respectivamente do segundo argumento).

Cálculo de A: Sabemos que $D_1g =$

$$\begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \frac{\partial g_1}{\partial x_3} & \frac{\partial g_1}{\partial x_4} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \frac{\partial g_2}{\partial x_3} & \frac{\partial g_2}{\partial x_4} \\ \frac{\partial g_3}{\partial x_1} & \frac{\partial g_3}{\partial x_2} & \frac{\partial g_3}{\partial x_3} & \frac{\partial g_3}{\partial x_4} \\ \frac{\partial g_4}{\partial x_1} & \frac{\partial g_4}{\partial x_2} & \frac{\partial g_4}{\partial x_3} & \frac{\partial g_4}{\partial x_4} \end{bmatrix} .$$

E mais ainda, $A = D_1g(0, 0)$. Denotemos por A_i , $1 \leq i \leq 4$ a i -ésima linha de A e calculemos separadamente as entradas de cada uma destas linhas .

Linha 1:

$$\bullet \frac{\partial g_1}{\partial x_1} = \frac{dx_2}{dx_1} = 0; \quad \bullet \frac{\partial g_1}{\partial x_2} = \frac{dx_2}{dx_2} = 1; \quad \bullet \frac{\partial g_1}{\partial x_3} = \frac{dx_2}{dx_3} = 0; \quad \bullet \frac{\partial g_1}{\partial x_4} = \frac{dx_2}{dx_4} = 0.$$

Logo,

$$A_1 = \left[\begin{array}{cccc} \frac{\partial g_1}{\partial x_1}(0, 0) & \frac{\partial g_1}{\partial x_2}(0, 0) & \frac{\partial g_1}{\partial x_3}(0, 0) & \frac{\partial g_1}{\partial x_4}(0, 0) \end{array} \right] = \left[\begin{array}{cccc} 0 & 1 & 0 & 0 \end{array} \right].$$

Linha 2:

$$\begin{aligned} \bullet \frac{\partial g_2}{\partial x_1} &= \frac{\partial(f_2(x_1, x_2, u))}{\partial x_1} = \frac{\partial}{\partial x_1} \left(\frac{\left[(M+m)g - mlx_2^2 \cos x_1 \right] \text{sen} x_1 - u \cos x_1}{l(M+m - m \cos^2 x_1)} \right) \\ &= \frac{\left(mlx_2^2 \text{sen} x_1 \cdot \text{sen} x_1 + \left[(M+m)g - mlx_2^2 \cos x_1 \right] \cos x_1 \right) l(M+m - m \cos^2 x_1)}{l^2(M+m - m \cos^2 x_1(t))^2} \\ &\quad + \frac{u \text{sen} x_1 \cdot l(M+m - m \cos^2 x_1)}{l^2(M+m - m \cos^2 x_1(t))^2} \\ &\quad - \frac{\left(\left[(M+m)g - mlx_2^2 \cos x_1 \right] \text{sen} x_1 - u \cos x_1 \right) 2ml \cos x_1 \text{sen} x_1}{l^2(M+m - m \cos^2 x_1(t))^2} \\ &= \frac{\left(mlx_2^2 \text{sen} x_1 \cdot \text{sen} x_1 + \left[(M+m)g - mlx_2^2 \cos x_1 \right] \cos x_1 + u \text{sen} x_1 \right)}{l(M+m - m \cos^2 x_1(t))} \\ &\quad - \frac{\left(\left[(M+m)g - mlx_2^2 \cos x_1 \right] \text{sen} x_1 - u \cos x_1 \right) 2m \cos x_1 \text{sen} x_1}{l(M+m - m \cos^2 x_1(t))^2} \end{aligned}$$

Daí,

$$\frac{\partial g_2}{\partial x_1}(0, 0) = \frac{(M+m)g}{Ml} - 0 = \frac{(M+m)g}{Ml}.$$

$$\begin{aligned} \bullet \frac{\partial g_2}{\partial x_2} &= \frac{\partial(f_2(x_1, x_2, u))}{\partial x_2} = \frac{\partial}{\partial x_2} \left(\frac{\left[(M+m)g - mlx_2^2 \cos x_1 \right] \text{sen} x_1 - u \cos x_1}{l(M+m - m \cos^2 x_1)} \right) \\ &= \frac{(-2x_2 m l \text{sen} x_1 \cos x_1) l(M+m - m \cos^2 x_1)}{\left(l(M+m - m \cos^2 x_1) \right)^2} = \frac{(-2x_2 m l \text{sen} x_1 \cos x_1)}{l(M+m - m \cos^2 x_1)}. \end{aligned}$$

Logo,

$$\frac{\partial g_2}{\partial x_2}(0, 0) = 0.$$

- $\frac{\partial g_2}{\partial x_3} = \frac{\partial f_2(x_1, x_2, u)}{\partial x_3} = 0 \Rightarrow \frac{\partial g_2}{\partial x_3}(0, 0) = 0 ;$

- $\frac{\partial g_2}{\partial x_4} = \frac{\partial f_2(x_1, x_2, u)}{\partial x_4} = 0 \Rightarrow \frac{\partial g_2}{\partial x_4}(0, 0) = 0 .$

Desse modo,

$$A_2 = \left[\begin{array}{cccc} \frac{\partial g_2}{\partial x_1}(0, 0) & \frac{\partial g_2}{\partial x_2}(0, 0) & \frac{\partial g_2}{\partial x_3}(0, 0) & \frac{\partial g_2}{\partial x_4}(0, 0) \end{array} \right] = \left[\begin{array}{cccc} \frac{(M+m)g}{Ml} & 0 & 0 & 0 \end{array} \right] .$$

Linha 3:

- $\frac{\partial g_3}{\partial x_1} = \frac{dx_4}{dx_1} = 0;$ • $\frac{\partial g_3}{\partial x_2} = \frac{dx_4}{dx_2} = 0;$ • $\frac{\partial g_3}{\partial x_3} = \frac{dx_4}{dx_3} = 0;$ • $\frac{\partial g_3}{\partial x_4} = \frac{dx_4}{dx_4} = 1.$

Daí,

$$A_3 = \left[\begin{array}{cccc} \frac{\partial g_3}{\partial x_1}(0, 0) & \frac{\partial g_3}{\partial x_2}(0, 0) & \frac{\partial g_3}{\partial x_3}(0, 0) & \frac{\partial g_3}{\partial x_4}(0, 0) \end{array} \right] = \left[\begin{array}{cccc} 0 & 0 & 0 & 1 \end{array} \right] .$$

Linha 4:

- $$\begin{aligned} \frac{\partial g_4}{\partial x_1} &= \frac{\partial f_4(x_1, x_2, u)}{\partial x_1} = \frac{\partial}{\partial x_1} \left(\frac{m \operatorname{sen} x_1 (lx_2^2 - g \cos x_1) + u}{M + m - m \cos^2 x_1} \right) \\ &= \frac{[m \cos x_1 (lx_2^2 - g \cos x_1) + m \operatorname{sen} x_1 g \operatorname{sen} x_1] (M + m - m \cos^2 x_1)}{(M + m - m \cos^2 x_1)^2} \\ &\quad - \frac{[m \operatorname{sen} x_1 (lx_2^2 - g \cos x_1) + u] (2m \cos x_1 \operatorname{sen} x_1)}{(M + m - m \cos^2 x_1)^2} . \end{aligned}$$

Assim,

$$\frac{\partial g_4}{\partial x_1}(0, 0) = \frac{-mg}{M} .$$

- $$\begin{aligned} \frac{\partial g_4}{\partial x_2} &= \frac{\partial f_4(x_1, x_2, u)}{\partial x_2} = \frac{\partial}{\partial x_2} \left(\frac{m \operatorname{sen} x_1 (lx_2^2 - g \cos x_1) + u}{M + m - m \cos^2 x_1} \right) \\ &= \frac{2x_2 m l \operatorname{sen} x_1 (M + m - m \cos^2 x_1)}{(M + m - m \cos^2 x_1)^2} . \end{aligned}$$

Com isto temos

$$\frac{\partial g_4}{\partial x_2}(0, 0) = 0 .$$

- $\frac{\partial g_4}{\partial x_3} = \frac{\partial f_4(x_1, x_2, u)}{\partial x_3} = 0 \Rightarrow \frac{\partial g_4}{\partial x_3}(0, 0) = 0;$

- $\frac{\partial g_4}{\partial x_4} = \frac{\partial f_4(x_1, x_2, u)}{\partial x_4} = 0 \Rightarrow \frac{\partial g_4}{\partial x_4}(0, 0) = 0.$

Logo,

$$A_4 = \left[\begin{array}{cccc} \frac{\partial g_4}{\partial x_1}(0, 0) & \frac{\partial g_4}{\partial x_2}(0, 0) & \frac{\partial g_4}{\partial x_3}(0, 0) & \frac{\partial g_4}{\partial x_4}(0, 0) \end{array} \right] = \left[\begin{array}{cccc} \frac{-mg}{M} & 0 & 0 & 0 \end{array} \right] .$$

Finalmente, concluimos que

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ (M+m)g/Ml & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -mg/M & 0 & 0 & 0 \end{bmatrix}.$$

Cálculo de B

Do que sabemos, $D_2g = \left[\frac{\partial g_1}{\partial u} \quad \frac{\partial g_2}{\partial u} \quad \frac{\partial g_3}{\partial u} \quad \frac{\partial g_4}{\partial u} \right]^*$ e $B = (D_2g)(0,0)$. Temos:

- $\frac{\partial g_1}{\partial u} = \frac{dx_2}{du} = 0 \Rightarrow \frac{\partial g_1}{\partial u}(0,0) = 0;$
- $\frac{\partial g_2}{\partial u} = \frac{\partial(f_2(x_1, x_2, u))}{\partial u} = \frac{\partial}{\partial u} \left(\frac{[(M+m)g - mlx_2^2 \cos x_1] \text{sen} x_1 - u \cos x_1}{l(M+m - m \cos^2 x_1)} \right)$
 $= \frac{-\cos x_1}{l(M+m - m \cos^2 x_1)}.$

Logo,

$$\frac{\partial g_2}{\partial u}(0,0) = -\frac{1}{Ml};$$

- $\frac{\partial g_3}{\partial u} = \frac{dx_4}{du} = 0 \Rightarrow \frac{\partial g_3}{\partial u}(0,0) = 0;$
- $\frac{\partial g_4}{\partial u} = \frac{\partial f_4(x_1, x_2, u)}{\partial u} = \frac{\partial}{\partial u} \left(\frac{m \text{sen} x_1 (lx_2^2 - g \cos x_1) + u}{M+m - m \cos^2 x_1} \right)$
 $= \frac{1}{M+m - m \cos^2 x_1}.$

Daí,

$$\frac{\partial g_4}{\partial u}(0,0) = \frac{1}{M}.$$

De tudo isto temos que

$$B = \begin{bmatrix} 0 \\ -1/Ml \\ 0 \\ 1/M \end{bmatrix}.$$

Portanto o modelo para o pêndulo, linearizado, é dado por

$$x'(t) = Ax(t) + Bu(t), \tag{2.11}$$

onde,

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ (M+m)g/Ml & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -mg/M & 0 & 0 & 0 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 0 \\ -1/Ml \\ 0 \\ 1/M \end{bmatrix}. \quad (2.12)$$

Este sistema linear não homogêneo de equações diferenciais “aproximadamente” governa o comportamento do pêndulo invertido “próximo” a posição de repouso vertical. Este problema, comumente referido como problema de “equilíbrio da vassoura”, é uma simplificação de um modelo usado para estabilização de foguetes (neste caso o controle u corresponde à ação de jatos laterais).

Considere o sistema $x'(t) = Ax(t) + Bu(t)$, com A e B dados por (2.12), o qual modela a dinâmica do pêndulo invertido. Mostraremos que este sistema é controlável. Vamos calcular a matriz controlabilidade do sistema a qual, por definição, é dada por

$$C(A, B) = \begin{bmatrix} B & AB & A^2B & A^3B \end{bmatrix} \in \mathbb{R}^{4 \times 4}.$$

Por simplicidade adote $k_1 = (M+m)g/Ml$, $k_2 = -mg/M$, $k_3 = -1/Ml$ e $k_4 = 1/M$.

Temos então

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ k_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ k_2 & 0 & 0 & 0 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 0 \\ k_3 \\ 0 \\ k_4 \end{bmatrix}.$$

Segue-se disto que:

$$\bullet AB = \begin{bmatrix} 0 & 1 & 0 & 0 \\ k_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ k_2 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ k_3 \\ 0 \\ k_4 \end{bmatrix} = \begin{bmatrix} k_3 \\ 0 \\ k_4 \\ 0 \end{bmatrix};$$

$$\bullet A^2B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ k_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ k_2 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ k_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ k_2 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ k_3 \\ 0 \\ k_4 \end{bmatrix} = \begin{bmatrix} k_1 & 0 & 0 & 0 \\ 0 & k_1 & 0 & 0 \\ k_2 & 0 & 0 & 0 \\ 0 & k_2 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ k_3 \\ 0 \\ k_4 \end{bmatrix};$$

$$\begin{aligned}
&= \begin{bmatrix} 0 \\ k_1 k_3 \\ 0 \\ k_2 k_3 \end{bmatrix}; \\
\bullet A^3 B &= \underbrace{\begin{bmatrix} k_1 & 0 & 0 & 0 \\ 0 & k_1 & 0 & 0 \\ k_2 & 0 & 0 & 0 \\ 0 & k_2 & 0 & 0 \end{bmatrix}}_{A^2} \begin{bmatrix} 0 & 1 & 0 & 0 \\ k_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ k_2 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ k_3 \\ 0 \\ k_4 \end{bmatrix} = \begin{bmatrix} 0 & k_1 & 0 & 0 \\ k_1^2 & 0 & 0 & 0 \\ 0 & k_2 & 0 & 0 \\ k_1 k_2 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ k_3 \\ 0 \\ k_4 \end{bmatrix} \\
&= \begin{bmatrix} k_1 k_3 \\ 0 \\ k_2 k_3 \\ 0 \end{bmatrix}.
\end{aligned}$$

Portanto

$$C(A, B) = \begin{bmatrix} 0 & k_3 & 0 & k_1 k_3 \\ k_3 & 0 & k_1 k_3 & 0 \\ 0 & k_4 & 0 & k_2 k_3 \\ k_4 & 0 & k_2 k_3 & 0 \end{bmatrix}.$$

Agora, usemos operações elementares sobre linhas para escrever a matriz $C(A, B)$ na forma escalonada reduzida:

$$\begin{aligned}
&\begin{bmatrix} 0 & k_3 & 0 & k_1 k_3 \\ k_3 & 0 & k_1 k_3 & 0 \\ 0 & k_4 & 0 & k_2 k_3 \\ k_4 & 0 & k_2 k_3 & 0 \end{bmatrix} \begin{array}{l} \xrightarrow{L_1 \leftrightarrow L_1/k_3} \\ \xrightarrow{L_2 \leftrightarrow L_2/k_3} \\ \xrightarrow{L_3 \leftrightarrow L_3/k_4} \\ \xrightarrow{L_4 \leftrightarrow L_4/k_4} \end{array} \begin{bmatrix} 0 & 1 & 0 & k_1 \\ 1 & 0 & k_1 & 0 \\ 0 & 1 & 0 & k_2 k_3/k_4 \\ 1 & 0 & k_2 k_3/k_4 & 0 \end{bmatrix} \begin{array}{l} \\ \xrightarrow{L_1 \leftrightarrow L_2} \\ \xrightarrow{L_3 \leftrightarrow L_4} \end{array} \\
&\begin{bmatrix} 1 & 0 & k_1 & 0 \\ 0 & 1 & 0 & k_1 \\ 1 & 0 & k_2 k_3/k_4 & 0 \\ 0 & 1 & 0 & k_2 k_3/k_4 \end{bmatrix} \begin{array}{l} \\ \xrightarrow{L_3 \leftrightarrow L_3 - L_1} \\ \xrightarrow{L_4 \leftrightarrow L_4 - L_2} \end{array} \begin{bmatrix} 1 & 0 & k_1 & 0 \\ 0 & 1 & 0 & k_1 \\ 0 & 0 & k_2 k_3/k_4 - k_1 & 0 \\ 0 & 0 & 0 & k_2 k_3/k_4 - k_1 \end{bmatrix}
\end{aligned}$$

$$\begin{array}{l} L_3 \leftrightarrow L_3 / \xrightarrow{(k_2 k_3 / k_4 - k_1)} \\ L_4 \leftrightarrow L_4 / \xrightarrow{(k_2 k_3 / k_4 - k_1)} \end{array} \begin{bmatrix} 1 & 0 & k_1 & 0 \\ 0 & 1 & 0 & k_1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{array}{l} L_1 \leftrightarrow L_1 - k_1 L_3 \\ L_2 \leftrightarrow L_2 - k_1 L_4 \end{array} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} .$$

Uma vez que o posto da matriz $C(A, B)$ é o número de linhas não nulas de sua forma escalonada reduzida temos que $\text{Posto } C(A, B) = 4$ e então, pelo Teorema 1.3 resulta que o sistema (2.11) é controlável.

Assumindo que d e θ são conhecidos, associamos ao sistema linearizado uma saída ou observação $y(t)$ valorizada em \mathbb{R}^2 , que representa informações sobre o estado do sistema, isto é, a posição do pêndulo no tempo t . Tal saída é definida por

$$y(t) := \begin{bmatrix} x_1(t) \\ x_3(t) \end{bmatrix} .$$

Observando que

$$y(t) = \begin{bmatrix} 1 \cdot x_1(t) + 0 \cdot x_2(t) + 0 \cdot x_3(t) + 0 \cdot x_4(t) \\ 0 \cdot x_1(t) + 0 \cdot x_2(t) + 1 \cdot x_3(t) + 0 \cdot x_4(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix}$$

podemos definir y por

$$y(t) := Cx(t) \tag{2.13}$$

onde

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} .$$

Representamos então o sistema (2.11) com sua respectiva saída como segue

$$\begin{cases} x'(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) \end{cases} . \tag{2.14}$$

Observação 2.1 É do nosso conhecimento que a controlabilidade de (2.14) não depende da matriz C (ou seja da saída) mas somente das matrizes A e B . Portanto, ela equivale à controlabilidade do sistema (2.11), a qual já foi verificada.

Supondo serem d e θ conhecidos, vamos mostrar que o sistema (2.14) é observável. Por definição, a matriz de observabilidade de (2.14) é dada por

$$\mathcal{O}(C, A) = \begin{bmatrix} C \\ CA \\ CA^2 \\ CA^3 \end{bmatrix} \in \mathbb{R}^{8 \times 4}.$$

Escrevendo A em termos de k_1 e k_2 definidos no exemplo anteriormente, temos:

$$\begin{aligned} \bullet CA &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ k_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ k_2 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \\ \bullet CA^2 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \underbrace{\begin{bmatrix} k_1 & 0 & 0 & 0 \\ 0 & k_1 & 0 & 0 \\ k_2 & 0 & 0 & 0 \\ 0 & k_2 & 0 & 0 \end{bmatrix}}_{A^2} = \begin{bmatrix} k_1 & 0 & 0 & 0 \\ k_2 & 0 & 0 & 0 \end{bmatrix}; \\ \bullet CA^3 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \underbrace{\begin{bmatrix} 0 & k_1 & 0 & 0 \\ k_1^2 & 0 & 0 & 0 \\ 0 & k_2 & 0 & 0 \\ k_1 k_2 & 0 & 0 & 0 \end{bmatrix}}_{A^3} = \begin{bmatrix} 0 & k_1 & 0 & 0 \\ 0 & k_2 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Portanto,

$$\mathcal{O}(C, A) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ k_1 & 0 & 0 & 0 \\ k_2 & 0 & 0 & 0 \\ 0 & k_1 & 0 & 0 \\ 0 & k_2 & 0 & 0 \end{bmatrix}.$$

Usando operações elementares sobre as linhas de $\mathcal{O}(C, A)$ vamos escrevê-la na forma escalonada reduzida:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ k_1 & 0 & 0 & 0 \\ k_2 & 0 & 0 & 0 \\ 0 & k_1 & 0 & 0 \\ 0 & k_2 & 0 & 0 \end{bmatrix} \xrightarrow{\begin{array}{l} L_5 \leftrightarrow L_5 - k_1 L_1 \\ L_6 \leftrightarrow L_6 - k_2 L_1 \\ L_7 \leftrightarrow L_7 - k_1 L_3 \\ L_8 \leftrightarrow L_8 - k_2 L_3 \end{array}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{L_2 \leftrightarrow L_3} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Uma vez que o número de linhas não nulas da forma escalonada reduzida de $\mathcal{O}(C, A)$ é 4, concluímos que $\text{Posto } \mathcal{O}(C, A) = 4$ e, portanto, pelo Teorema 1.7, o sistema (2.14) é observável.

2.2 Dinâmica de um satélite

O objetivo deste problema é ilustrar a modelagem da dinâmica de um satélite e a determinação de um equilíbrio desejado. Subsequentemente, o movimento é linearizado e estabilidade, controlabilidade e observabilidade são analisadas.

2.2.1 Motivação

A mais importante contribuição do programa espacial é sem dúvidas a comunicação via satélites. A ideia é simples: ao invés de transmitir uma mensagem através de um fio, ela é transmitida de um transmissor para um satélite e, a partir daí, a mensagem é encaminhada para um receptor. Por razões óbvias, é desejável que o satélite esteja localizado em uma posição no espaço que seja fixa para um observador na terra. Isto não somente evita ter que rastrear o satélite continuamente, mas também resulta no fato de que o satélite pode ser usado a todo momento, visto que nunca desaparece além do horizonte. Satélites que situam-se em uma posição fixa no céu são utilizados para comunicação por TV e telefone, na navegação de navios e aeronaves, para previsão do tempo e etc. Satélites como este são chamados satélites *geostacionários*.

A princípio, uma órbita geoestacionária pode ser alcançada pela aplicação de forças no satélite de modo que o mesmo permaneça na órbita desejada. Tais forças podem ser produzidas por meio de pequenos jatos que são montados nos satélites. No entanto, é indesejável exigir que estas forças seja exercidas continuamente. Estes jatos adquirem sua energia a partir do combustível que o satélite deve levar desde o lançamento (programas espaciais atuais visam satélites de reabastecimento) ou a partir de painéis solares. Entretanto, energia é um recurso escasso no céu e, desse modo, é desejável que os jatos não sejam usados para direcionar continuamente o satélite, mas apenas para correções orbitais inevitáveis e outras manobras.

Assim, surge a questão: Existe uma órbita de equilíbrio geoestacionário para um satélite quando a única força exercida sobre ele é o campo de força gravitacional da Terra? Essa órbita é estável? Se não, que controles são necessários para manter o satélite em sua órbita?

Se assumirmos que o satélite é influenciado apenas pelo campo gravitacional da Terra, então sua órbita obedece as Leis de Kepler. Desse modo, o satélite move-se em uma órbita elíptica com o centro da terra em um dos focos. Assim, como uma consequência do fato de que a Terra gira em torno de seu eixo Polo Norte/Polo Sul a uma taxa de 2π radianos/dia, uma órbita geoestacionária deve ser circular e no plano equatorial. As leis de Kepler também implicam que há uma relação entre o diâmetro da órbita circular e seu período de revolução que deve ser o mesmo que o da terra para que o satélite seja geoestacionário.

2.2.2 Modelagem Matemática

Iremos agora derivar as equações dinâmicas do movimento do satélite. Este último está sujeito a quatro tipos de forças:

- i)* A força inercial, \vec{F}_{in}
- ii)* A força gravitacional da Terra, \vec{F}_g
- iii)* Forças externas devidas aos jatos, \vec{F}_{jat} : estes são os nossos controles.
- iv)* Outras forças externas, como a atração gravitacional da lua e do sol e o vento solar. Denotaremos estas forças por \vec{F}_d . As mesmas tratam-se de distúrbios e o objetivo dos controles é exatamente compensar a influência imprevisível destas perturbações.

A posição do satélite pode ser descrita pelas coordenadas polares (r, ψ, θ) . A Figura 2.3(a) mostra a geometria a ser estudada. A princípio o movimento destas coordenadas é unificado (acoplado). Entretanto, conforme já comentado, a órbita desejada é circular e pertencente ao plano equatorial e portanto, é neste plano onde concentraremos nossos estudos. A geometria resultante da avaliação do movimento no plano equatorial é mostrada na figura 2.3(b).

Para um observador na terra, o satélite possui posição fixa. Tal observador tem então a percepção de uma força (Uma força *fictícia*, é claro, chamada de *força centrífuga*) que puxa o satélite para cima (o que para nós é na direção radial) anulando a atração gravitacional da terra e impedindo que o satélite caia. Esta força fictícia é aquela denominada por *força inercial*. Empenharemos agora esforços para expressar esta força matematicamente.

Denotaremos os vetores unitários nas direções radial e tangencial por $\vec{1}_r$ e $\vec{1}_\theta$ respectivamente. A velocidade do satélite em sua órbita é a variação de sua posição ao longo do tempo. Ela corresponde com a velocidade de rotação (trata-se portanto de uma velocidade angular) do vetor \vec{r} em torno da origem. Para determinar a mesma, observemos que

$$\vec{r} = r \cdot \vec{1}_r \quad (2.15)$$

De início, determinemos a velocidade de rotação do vetor $\vec{1}_r$. Sendo θ o ângulo entre $\vec{1}_r$ e o eixo das abscissas temos,

$$\left| \frac{d \vec{1}_r}{dt} \right| = \frac{d\theta}{dt}$$

e como a variação de θ está orientada na direção e no sentido de $\vec{1}_\theta$ obtemos,

$$\frac{d \vec{1}_r}{dt} = \frac{d\theta}{dt} \vec{1}_\theta \quad (2.16)$$

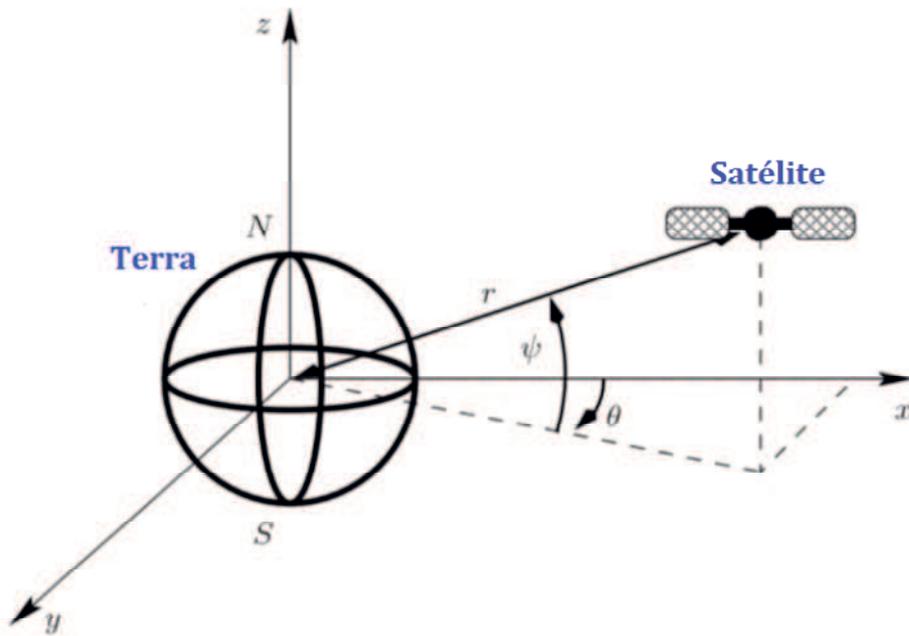
Desse modo, pela equação (2.15) temos

$$\frac{d \vec{r}}{dt} = \frac{d}{dt}(r \cdot \vec{1}_r) = \frac{dr}{dt} \vec{1}_r + r \frac{d \vec{1}_r}{dt}$$

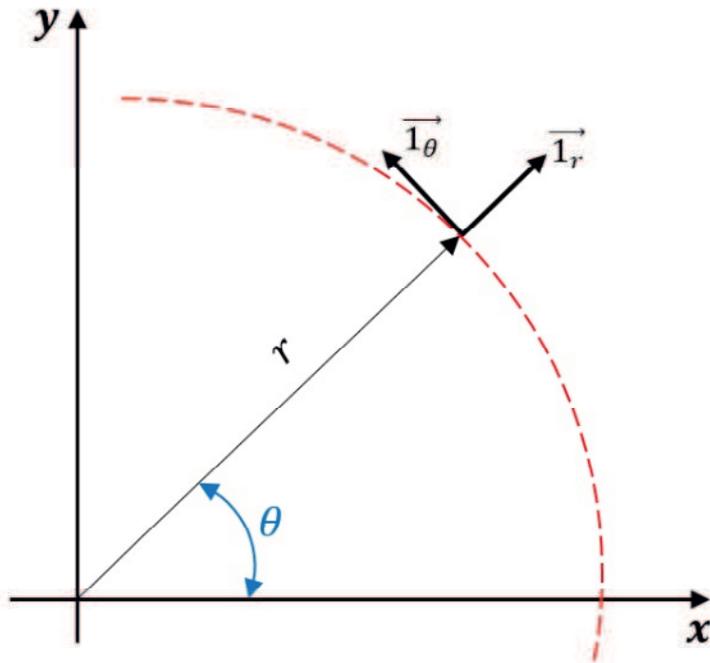
e pela equação (2.16) resulta

$$\frac{d \vec{r}}{dt} = \frac{dr}{dt} \vec{1}_r + r \frac{d\theta}{dt} \vec{1}_\theta \quad (2.17)$$

No sentido de determinar a aceleração da rotação de \vec{r} , devemos conhecer $\frac{d}{dt} \vec{1}_\theta$. Para isto, note que a velocidade de rotação de $\vec{1}_\theta$ é em módulo a mesma com que o ângulo θ varia,



(a) Satélite em órbita ao redor da Terra



(b) Movimento resultante no plano equatorial

Figura 2.3: Representação Geométrica do movimento do satélite.

já que o movimento de \vec{l}_r e \vec{l}_θ é unificado. Daí,

$$\left| \frac{d}{dt} \vec{l}_\theta \right| = \frac{d\theta}{dt}.$$

Observe ainda que, em cada ponto da trajetória a orientação da rotação de \vec{l}_θ é dada por $-\vec{l}_r$.

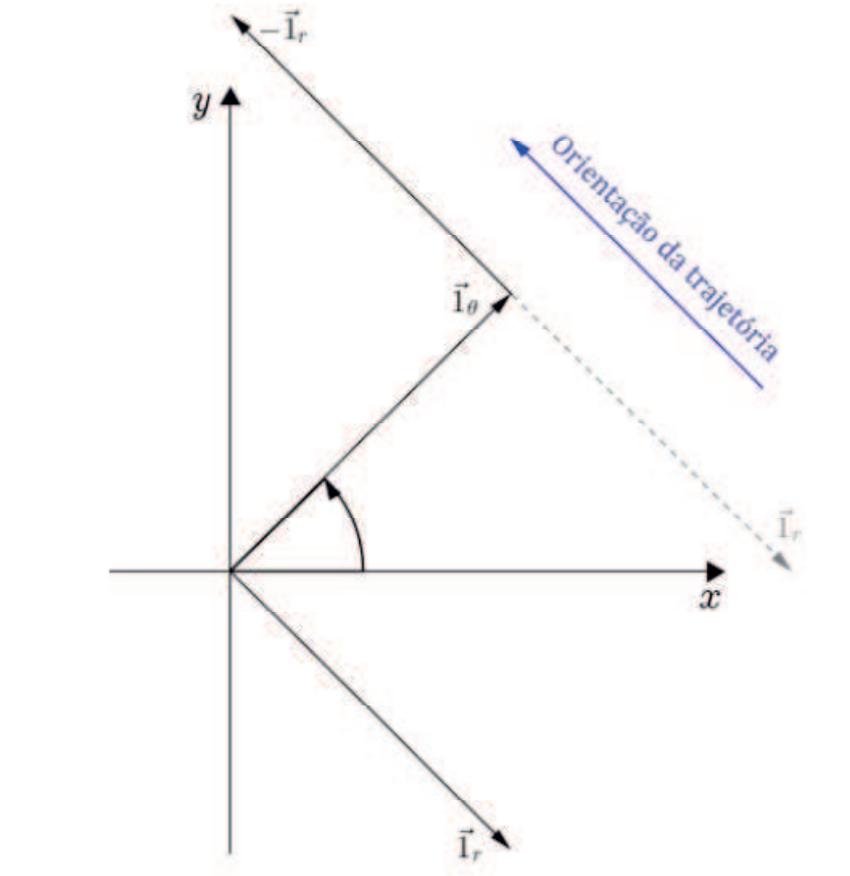


Figura 2.4: Representação de \vec{l}_θ e \vec{l}_r a partir da origem.

Logo, a direção de rotação de \vec{l}_θ é $-\vec{l}_r$ e portanto,

$$\frac{d}{dt} \vec{l}_\theta = \frac{d\theta}{dt} (-\vec{l}_r) = -\frac{d\theta}{dt} \vec{l}_r. \quad (2.18)$$

Assim, derivando a expressão (2.17) temos

$$\frac{d^2}{dt^2} \vec{r} = \frac{d^2 r}{dt^2} \vec{l}_r + \frac{dr}{dt} \frac{d}{dt} \vec{l}_r + \frac{dr}{dt} \frac{d\theta}{dt} \vec{l}_\theta + r \left(\frac{d^2 \theta}{dt^2} \vec{l}_\theta + \frac{d\theta}{dt} \frac{d}{dt} \vec{l}_\theta \right)$$

Usando as equações (2.16) e (2.18) obtemos,

$$\frac{d^2}{dt^2} \vec{r} = \frac{d^2 r}{dt^2} \vec{l}_r + \frac{dr}{dt} \frac{d\theta}{dt} \vec{l}_\theta + \frac{dr}{dt} \frac{d\theta}{dt} \vec{l}_\theta + r \left(\frac{d^2 \theta}{dt^2} \vec{l}_\theta + \frac{d\theta}{dt} \left(-\frac{d\theta}{dt} \vec{l}_r \right) \right)$$

$$= \frac{d^2 r}{dt^2} \vec{1}_r + 2 \frac{dr}{dt} \frac{d\theta}{dt} \vec{1}_\theta + r \frac{d^2 \theta}{dt^2} \vec{1}_\theta - r \left(\frac{d\theta}{dt} \right)^2 \vec{1}_r .$$

Portanto,

$$\frac{d^2}{dt^2} \vec{r} = \left(\frac{d^2 r}{dt^2} - r \left(\frac{d\theta}{dt} \right)^2 \right) \vec{1}_r + \left(r \frac{d^2 \theta}{dt^2} + 2 \frac{dr}{dt} \frac{d\theta}{dt} \right) \vec{1}_\theta \quad (2.19)$$

Fisicamente falando, a igualdade (2.19) fornece a aceleração da rotação de \vec{r} (e coincide com a aceleração que a força imaginada pelo observador impõe ao satélite) no tempo t .

Daí, pela segunda lei de Newton a força inercial é dada por

$$m \frac{d^2}{dt^2} \vec{r} = m \left(\frac{d^2 r}{dt^2} - r \left(\frac{d\theta}{dt} \right)^2 \right) \vec{1}_r + m \left(r \frac{d^2 \theta}{dt^2} + 2 \frac{dr}{dt} \frac{d\theta}{dt} \right) \vec{1}_\theta$$

onde m denota a massa do satélite. Do que sabemos, o módulo a força gravitacional da terra sobre o satélite é dada pela lei do inverso do quadrado

$$\left| \vec{F}_g \right| = \frac{G \cdot m_T \cdot m}{r^2}$$

onde G é a constante de gravitação universal e m_T é a massa da terra. Sabendo que $G = 5,973332 \cdot 10^{24} \text{ kg}$ e $m_T = 6,67 \cdot 10^{-11} \text{ N} \cdot \text{m}^2/\text{kg}^2$, fazamos

$$\begin{aligned} k = G \cdot m_T &= 39,84 \cdot 10^{13} \frac{\text{Nm}^2}{\text{kg}} = 3,984 \cdot 10^{14} \frac{(\text{kg} \cdot \text{m}/\text{s}^2)\text{m}^2}{\text{kg}} \Rightarrow \\ &\Rightarrow k = 3,984 \cdot 10^{14} \text{ m}^3/\text{s}^2 . \end{aligned}$$

Assim, podemos escrever

$$\left| \vec{F}_g \right| = k \frac{m}{r^2}$$

com k obtido acima. Como a força gravitacional “aponta” para o centro da terra, podemos concluir que a direção de \vec{F}_g é a de $-\vec{1}_r$. Logo

$$\vec{F}_g = -k \frac{m}{r^2} \vec{1}_r .$$

Assumimos que os jatos exercem uma força sobre o satélite denotada por \vec{F}_{jat} . Vamos decompor esta força em uma força de controle u_r na direção radial e uma força de controle u_θ na direção tangencial. A força \vec{F}_d , que corresponde às perturbações sofridas pelo satélite, será similarmente decomposta em p_r na direção radial e p_θ na direção tangencial. Desse modo,

$$\vec{F}_{jat} = u_r \vec{1}_r + u_\theta \vec{1}_\theta \quad \text{e} \quad \vec{F}_d = p_r \vec{1}_r + p_\theta \vec{1}_\theta .$$

Como sabemos, a força inercial e a força de atração gravitacional estão na mesma direção, em sentidos opostos em relação ao satélite e se anulam. Matematicamente,

$$\vec{F}_{in} - \vec{F}_g = 0 \quad .$$

Substituindo \vec{F}_{in} e \vec{F}_g pelas expressões obtidas para as mesmas obtemos

$$m \left(\frac{d^2 r}{dt^2} - r \left(\frac{d\theta}{dt} \right)^2 \right) \vec{1}_r + m \left(r \frac{d^2 \theta}{dt^2} + 2 \frac{dr}{dt} \frac{d\theta}{dt} \right) \vec{1}_\theta + k \frac{m}{r^2} \vec{1}_r = 0 \quad .$$

Além disto, já mencionamos que o papel de \vec{F}_{jat} é anular os efeitos de \vec{F}_d , ou seja,

$$\vec{F}_{jat} + \vec{F}_d = 0 \quad .$$

Disto concluímos que

$$m \left(\frac{d^2 r}{dt^2} - r \left(\frac{d\theta}{dt} \right)^2 \right) \vec{1}_r + m \left(r \frac{d^2 \theta}{dt^2} + 2 \frac{dr}{dt} \frac{d\theta}{dt} \right) \vec{1}_\theta + k \frac{m}{r^2} \vec{1}_r = \vec{F}_{jat} + \vec{F}_d \quad .$$

Daí,

$$\begin{aligned} & \left(\frac{d^2 r}{dt^2} - r \left(\frac{d\theta}{dt} \right)^2 \right) \vec{1}_r + \left(r \frac{d^2 \theta}{dt^2} + 2 \frac{dr}{dt} \frac{d\theta}{dt} \right) \vec{1}_\theta + \frac{k}{r^2} \vec{1}_r = \frac{1}{m} \left(\vec{F}_{jat} + \vec{F}_d \right) \Rightarrow \\ \Rightarrow & \left(\frac{d^2 r}{dt^2} - r \left(\frac{d\theta}{dt} \right)^2 + \frac{k}{r^2} \right) \vec{1}_r + \left(r \frac{d^2 \theta}{dt^2} + 2 \frac{dr}{dt} \frac{d\theta}{dt} \right) \vec{1}_\theta = \frac{1}{m} \left(u_r \vec{1}_r + u_\theta \vec{1}_\theta + p_r \vec{1}_r + p_\theta \vec{1}_\theta \right) \\ \Rightarrow & \left(\frac{d^2 r}{dt^2} - r \left(\frac{d\theta}{dt} \right)^2 + \frac{k}{r^2} \right) \vec{1}_r + \left(r \frac{d^2 \theta}{dt^2} + 2 \frac{dr}{dt} \frac{d\theta}{dt} \right) \vec{1}_\theta = \left(\frac{u_r}{m} + \frac{p_r}{m} \right) \vec{1}_r + \left(\frac{u_\theta}{m} + \frac{p_\theta}{m} \right) \vec{1}_\theta . \end{aligned}$$

Logo

$$\frac{d^2 r}{dt^2} - r \left(\frac{d\theta}{dt} \right)^2 + \frac{k}{r^2} = \frac{u_r}{m} + \frac{p_r}{m}$$

e

$$r \frac{d^2 \theta}{dt^2} + 2 \frac{dr}{dt} \frac{d\theta}{dt} = \frac{u_\theta}{m} + \frac{p_\theta}{m} \quad .$$

Portanto, as equações diferenciais que nos dão a ligação dinâmica das variáveis r e θ com as entradas de controle u_r e u_θ e com as entradas de perturbação d_r e d_θ são

$$\begin{cases} \frac{d^2}{dt^2} r(t) = r(t) \left(\frac{d}{dt} \theta(t) \right)^2 - \frac{k}{r(t)^2} + \frac{u_r(t)}{m} + \frac{p_r(t)}{m}, \\ \frac{d^2}{dt^2} \theta(t) = -\frac{2 \frac{d}{dt} r(t) \frac{d}{dt} \theta(t)}{r(t)} + \frac{u_\theta(t)}{mr} + \frac{p_\theta(t)}{mr} \quad . \end{cases} \quad (2.20)$$

No seguinte, escreveremos o sistema de equações de segunda ordem (2.20) como um sistema de equações diferenciais de primeira ordem. Para isto, tomemos $\sigma, \omega \in \mathbb{R}$ como anteriormente, isto é, $\sigma^3\omega^2 = k$. Sejam x_1, x_2, x_3 e x_4 dados por

$$x_1(t) = r(t) - \sigma \quad , \quad x_2(t) = r'(t) \quad , \quad x_3(t) = \theta(t) - \omega t \quad , \quad x_4(t) = \theta'(t) - \omega \quad ; \quad \forall t \geq 0$$

e definamos

$$u_1(t) = \frac{u_r(t) + p_r(t)}{m} \quad , \quad u_2(t) = \frac{u_\theta(t) + p_\theta(t)}{m} \quad ; \quad \forall t \geq 0 \quad .$$

Daí, $x_1'(t) = r(t) = x_2(t)$ e

$$x_2'(t) = r''(t) \quad .$$

Por (2.20) vem

$$x_2'(t) = r(t)(\theta'(t))^2 - \frac{k}{r(t)^2} + \frac{u_r(t)}{m} + \frac{p_r(t)}{m}$$

e usando as devidas substituições acima definidas vem

$$x_2'(t) = (x_1(t) + \sigma)(x_4(t) + \omega)^2 - \frac{\sigma^3\omega^2}{(x_1(t) + \sigma)^2} + u_1(t) \quad .$$

Temos ainda $x_3'(t) = \theta'(t) - \omega = x_4(t)$ bem como

$$x_4'(t) = \theta''(t)$$

e assim segue-se por (2.20)

$$x_4'(t) = -\frac{2r'(t)\theta'(t)}{r(t)} + \frac{u_\theta(t)}{mr(t)} + \frac{p_\theta(t)}{mr(t)}$$

e então pelas substituições impostas vem

$$x_4'(t) = -\frac{2x_2(t)(x_4(t) + \omega)}{x_1(t) + \sigma} + u_2(t) \cdot \frac{1}{r(t)} = -\frac{2x_2(t)(x_4(t) + \omega)}{x_1(t) + \sigma} + \frac{u_2(t)}{x_1(t) + \sigma} \quad .$$

Finalmente, o sistema (2.20) pode ser reescrito da seguinte forma

$$\left\{ \begin{array}{l} x_1'(t) = x_2(t) \\ x_2'(t) = (x_1(t) + \sigma)(x_4(t) + \omega)^2 - \frac{\sigma^3\omega^2}{(x_1(t) + \sigma)^2} + u_1(t) \\ x_3'(t) = x_4(t) \\ x_4'(t) = -\frac{2x_2(t)(x_4(t) + \omega)}{x_1(t) + \sigma} + \frac{u_2(t)}{x_1(t) + \sigma} \end{array} \right. \quad . \quad (2.21)$$

Consideremos a substituição de variáveis que origina (2.21) a partir de (2.20). Desde que $r(t) > 0$ para todo $t \geq 0$, temos que

$$x_1(t) = r(t) - \sigma > -\sigma$$

e assim $x_1 \in (-\sigma, \infty)$. A uma dada função de entrada $u = (u_1, u_2)$ que toma valores em \mathbb{R}^2 e definindo $G := (-\sigma, \infty) \times \mathbb{R}^3$, associamos uma função $f_u : \mathbb{R} \times G \rightarrow \mathbb{R}^4$ dada por

$$f_u(t, z) = f_u(t, z_1, z_2, z_3, z_4) := \left(z_2, f_2(z_1, z_4, u_1(t)), z_4, f_4(z_1, z_2, z_4, u_2(t)) \right),$$

onde as funções $f_2 : (-\sigma, \infty) \times \mathbb{R}^2 \rightarrow \mathbb{R}$ e $f_4 : G \rightarrow \mathbb{R}$ são definidas por

$$\begin{cases} f_2(z_1, z_4, v_1) := (z_1 + \sigma)(z_4 + \omega)^2 - \frac{\sigma^3 \omega^2}{(z_1 + \sigma)^2} + v_1 \\ f_4(z_1, z_2, z_4, v_2) := -\frac{2z_2(z_4 + \omega)}{z_1 + \sigma} + \frac{v_2}{z_1 + \sigma} \end{cases} \quad (2.22)$$

Definindo $x(t) := (x_1(t), x_2(t), x_3(t), x_4(t))$, temos por (2.21)

$$f_2(x_1, x_4, u_1) = (x_1 + \sigma)(x_4 + \omega)^2 - \frac{\sigma^3 \omega^2}{(x_1 + \sigma)^2} + u_1 = x_2'(t) ;$$

$$f_4(x_1, x_2, x_4, u_2) = -\frac{2x_2(x_4 + \omega)}{x_1 + \sigma} + \frac{u_2}{x_1 + \sigma} = x_4'(t)$$

e mais,

$$\begin{aligned} f_u(t, x(t)) &= f_u(t, x_1(t), x_2(t), x_3(t), x_4(t)) \\ &= \left(x_2(t), f_2(x_1(t), x_4(t), u_1(t)), x_4(t), f_4(x_1(t), x_2(t), x_4(t), u_2(t)) \right) \\ &= (x_1'(t), x_2'(t), x_3'(t), x_4'(t)) \\ &= x'(t) \quad . \end{aligned}$$

Portanto, o sistema (2.21) pode ser escrito da forma

$$x'(t) = f_u(t, x(t)) \quad (2.23)$$

Claro que cada escolha distinta de $u = (u_1, u_2)$ origina uma função distinta f_u . A partir daqui, conduziremos nossos estudos com base nas aproximações lineares das componentes de f_u calculadas nas origens de seus respectivos domínios. Uma vez que x_2 e x_4 já são termos lineares, iremos nos preocupar apenas em determinar as linearizações de $f_2(x_1, x_4, u_1)$ e $f_4(x_1, x_2, x_4, u_2)$. Recorde que a linearização¹ de uma função $f(x, y, z)$ (e analogamente para uma função de mais do que três variáveis) em um ponto $P_0(x_0, y_0, z_0)$ é a função

$$L(x, y, z) = f(P_0) + f_x(P_0)(x - x_0) + f_y(P_0)(y - y_0) + f_z(P_0)(z - z_0) \quad .$$

¹Para mais detalhes veja qualquer livro de cálculo (sugestão: ver referência [4]).

Assim, as linearizações de $f_2(x_1, x_4, u_1)$ e $f_4(x_1, x_2, x_4, u_2)$ nos pontos $(0, 0, 0)$ e $(0, 0, 0, 0)$, respectivamente, são dadas por

$$L_2(x_1, x_4, u_1) = f_2(0, 0, 0) + (f_2)_{x_1}(0, 0, 0)(x_1 - 0) + (f_2)_{x_4}(0, 0, 0)(x_4 - 0) + (f_2)_{u_1}(0, 0, 0)(u_1 - 0)$$

e

$$L_4(x_1, x_2, x_4, u_2) = f_4(0, 0, 0, 0) + (f_4)_{x_1}(0, 0, 0, 0)(x_1 - 0) + (f_4)_{x_2}(0, 0, 0, 0)(x_2 - 0) + (f_4)_{x_4}(0, 0, 0, 0)(x_4 - 0) + (f_4)_{u_2}(0, 0, 0, 0)(u_2 - 0) .$$

Note então que:

$$(i) \quad f_2(x_1, x_4, u_1) = x_1 x_4^2 + 2x_1 x_4 \omega + x_1 \omega^2 + \sigma x_4^2 + 2x_4 \omega \sigma + \sigma \omega^2 - \frac{\sigma^3 \omega^2}{(x_1 + \sigma)^2} + u_1 . \text{ Daí}$$

$$f_2(0, 0, 0) = \sigma \omega^2 - \sigma \omega^2 = 0$$

$$(ii) \quad (f_2)_{x_1}(x_1, x_4, u_1) = x_4^2 + 2x_4 \omega + \omega^2 - \left(-\frac{\sigma^3 \omega^2 2(x_1 + \sigma)}{(x_1 + \sigma)^4} \right) \text{ donde tiramos}$$

$$(f_2)_{x_1}(0, 0, 0) = \omega^2 + \frac{2\sigma^4 \omega^2}{\sigma^4} = 3\omega^2 .$$

$$(iii) \quad (f_2)_{x_4}(x_1, x_4, u_1) = 2x_1 x_4 + 2x_1 \omega + 2\sigma x_4 + 2\omega \sigma \text{ de modo que}$$

$$(f_2)_{x_4}(0, 0, 0) = 2\omega \sigma .$$

$$(iv) \quad (f_2)_{u_1}(x_1, x_4, u_1) = 1 \text{ de modo que}$$

$$(f_2)_{u_1}(0, 0, 0) = 1 .$$

De (i) à (iv) segue-se que

$$L_2(x_1, x_4, u_1) = 3\omega^2 x_1 + 2\omega \sigma x_4 + u_1 .$$

Observe agora que:

$$(v) \quad f_4(x_1, x_2, x_4, u_2) = -\frac{2x_2 x_4 + 2x_2 \omega}{x_1 + \sigma} + \frac{u_2}{x_1 + \sigma} . \text{ Daí tem-se que}$$

$$f_4(0, 0, 0, 0) = 0 .$$

$$(vi) \quad (f_4)_{x_1}(x_1, x_2, x_4, u_2) = -\left(-\frac{2x_2 x_4 + 2x_2 \omega}{(x_1 + \sigma)^2} \right) - \frac{u_2}{(x_1 + \sigma)^2} = \frac{2x_2 x_4 + 2x_2 \omega - u_2}{(x_1 + \sigma)^2} \text{ de modo que}$$

$$(f_4)_{x_1}(0, 0, 0, 0) = 0 .$$

$$(vii) \quad (f_4)_{x_2}(x_1, x_2, x_4, u_2) = -\frac{2x_4 + 2\omega}{x_1 + \sigma}. \quad \text{Logo}$$

$$(f_4)_{x_2}(0, 0, 0, 0) = -\frac{2\omega}{\sigma}.$$

$$(viii) \quad (f_4)_{x_4}(x_1, x_2, x_4, u_2) = -\frac{2x_2}{x_1 + \sigma} \text{ de onde vem}$$

$$(f_4)_{x_4}(0, 0, 0, 0) = 0.$$

$$(ix) \quad (f_4)_{u_2}(x_1, x_2, x_4, u_2) = \frac{1}{x_1 + \sigma} \text{ e então}$$

$$(f_4)_{u_2}(0, 0, 0, 0) = \frac{1}{\sigma}.$$

De (v) à (ix) segue-se que

$$L_4(x_1, x_2, x_4, u_2) = -\frac{2\omega}{\sigma}x_2 + \frac{1}{\sigma}u_2.$$

Portanto o sistema(2.23) linearizado é dado por

$$x'(t) = f_{u,L}(t, x(t))$$

onde $f_{u,L}(t, x(t))$ é dada por

$$f_{u,L}(t, x(t)) = \left(x_2(t), L_2(x_1(t), x_4(t), u_1(t)), x_4(t), L_4(x_1(t), x_2(t), x_4(t), u_2(t)) \right).$$

Considerando a representação de um vetor do \mathbb{R}^n (qualquer que seja n) na forma de matriz coluna podemos escrever

$$\begin{aligned} \begin{bmatrix} x'_1(t) \\ x'_2(t) \\ x'_3(t) \\ x'_4(t) \end{bmatrix} &= \begin{bmatrix} x_2(t) \\ L_2(x_1(t), x_4(t), u_1(t)) \\ x_4(t) \\ L_4(x_1(t), x_2(t), x_4(t), u_2(t)) \end{bmatrix} \Rightarrow \\ \Rightarrow \begin{bmatrix} x'_1(t) \\ x'_2(t) \\ x'_3(t) \\ x'_4(t) \end{bmatrix} &= \begin{bmatrix} x_2(t) \\ 3\omega^2 x_1 + 2\omega\sigma x_4 + u_1(t) \\ x_4(t) \\ -\frac{2\omega}{\sigma}x_2 + \frac{1}{\sigma}u_2(t) \end{bmatrix} = \begin{bmatrix} x_2(t) \\ 3\omega^2 x_1 + 2\omega\sigma x_4 \\ x_4(t) \\ -\frac{2\omega}{\sigma}x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ u_1(t) \\ 0 \\ \frac{1}{\sigma}u_2(t) \end{bmatrix} \Rightarrow \end{aligned}$$

$$\begin{aligned} \Rightarrow \begin{bmatrix} x'_1(t) \\ x'_2(t) \\ x'_3(t) \\ x'_4(t) \end{bmatrix} &= \begin{bmatrix} 0 \cdot x_1(t) + 1 \cdot x_2(t) + 0 \cdot x_3(t) + 0 \cdot x_4(t) \\ 3\omega^2 \cdot x_1(t) + 0 \cdot x_2(t) + 0 \cdot x_3(t) + 2\omega\sigma \cdot x_4(t) \\ 0 \cdot x_1(t) + 0 \cdot x_2(t) + 0 \cdot x_3(t) + 1 \cdot x_4(t) \\ 0 \cdot x_1(t) + \left(-\frac{2\omega}{\sigma}\right) \cdot x_2 + 0 \cdot x_3(t) + 0 \cdot x_4(t) \end{bmatrix} + \begin{bmatrix} 0 \cdot u_1(t) + 0 \cdot u_2(t) \\ 1 \cdot u_1(t) + 0 \cdot u_2(t) \\ 0 \cdot u_1(t) + 0 \cdot u_2(t) \\ 0 \cdot u_1(t) + \frac{1}{\sigma} \cdot u_2(t) \end{bmatrix} \Rightarrow \\ \Rightarrow \begin{bmatrix} x'_1(t) \\ x'_2(t) \\ x'_3(t) \\ x'_4(t) \end{bmatrix} &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega\sigma \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega/\sigma & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1/\sigma \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} . \end{aligned}$$

Definindo A e B por

$$A := \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega\sigma \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega/\sigma & 0 & 0 \end{bmatrix} \quad \text{e} \quad B := \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1/\sigma \end{bmatrix}$$

temos que o modelo linearizado é dado por

$$x'(t) = Ax(t) + Bu(t) . \quad (2.24)$$

Pontuamos que a linearização pode também ser obtida por procedimentos análogos àqueles adotados para o caso do pêndulo invertido definindo-se a função

$$\begin{aligned} g : \quad \mathbb{R}^4 \times \mathbb{R}^2 &\longrightarrow \mathbb{R}^4 \\ (x(t), u(t)) &\longmapsto f_u(t, x, (t)) , \end{aligned}$$

o que permite a escrita do sistema (2.23) na forma $x'(t) = g(x(t), u(t))$, e tomando $A = (D_1g)(0, 0)$ e $B = (D_2g)(0, 0)$. O sistema linear não homogêneo (2.24) de equações diferenciais “aproximadamente” governa o comportamento do satélite “próximo” à órbita circular. Assumindo que r e θ podem ser mensurados, associamos ao sistema linearizado uma saída ou observação $y(t)$ valorizada em \mathbb{R}^2 , que fornece informações sobre o estado do sistema (posição do satélite) no tempo (t) , definida por

$$y(t) := \begin{bmatrix} x_1(t) \\ x_3(t) \end{bmatrix} .$$

Observando que

$$y(t) = \begin{bmatrix} 1 \cdot x_1(t) + 0 \cdot x_2(t) + 0 \cdot x_3(t) + 0 \cdot x_4(t) \\ 0 \cdot x_1(t) + 0 \cdot x_2(t) + 1 \cdot x_3(t) + 0 \cdot x_4(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix}$$

e escrevendo $C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$, podemos definir y por

$$y(t) := Cx(t) . \quad (2.25)$$

Na próxima seção, mostraremos que o sistema linear controlado e observado dado por (2.24) e (2.25) é *controlável* e *observável*, ou seja, possui as seguintes propriedades da *controlabilidade* e *observabilidade*:

- (a) Para cada estado inicial $x(0) = x_0 \in \mathbb{R}^4$, todo estado pré-estabelecido $x^1 \in \mathbb{R}^4$ e todo tempo $T > 0$, existe uma função controle $u : [0, T] \rightarrow \mathbb{R}^2$ tal que $x(T) = x^1$.
- (b) Para cada tempo $T > 0$, o estado $x(T)$ pode ser determinado a partir do conhecimento do controle u e da saída y no intervalo $[0, T]$.

2.2.3 Controlabilidade e observabilidade do modelo linearizado

Considere o modelo do satélite linearizado dado por

$$\begin{cases} x'(t) = Ax(t) + Bu(t) \\ y = Cx(t) \end{cases} \quad (2.26)$$

com $A \in \mathbb{R}^{4 \times 4}$, $B \in \mathbb{R}^{4 \times 2}$ e $C \in \mathbb{R}^{2 \times 4}$ dados na seção anterior. Normalizando a constante σ para 1 temos

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{e} \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \quad (2.27)$$

Relembre que as componentes u_1 e u_2 de u correspondem às forças nas direções radial e tangencial, respectivamente, enquanto as componentes y_1 e y_2 de y correspondem às coordenadas radial e angular (do satélite) respectivamente.

Exemplo 2.1 Considere o modelo do satélite linearizado (e normalizado) com A e B dados por (2.27).

- (a) Calcule R (o conjunto dos estados alcançáveis a partir de 0).
- (b) Substitua B por $B_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}^*$, a transposta da primeira coluna de B (fisicamente isto significa que a entrada u_2 está inoperante, isto é, a força tangencial é inexistente). Calcule R neste caso.

Solução:

- (a) Usaremos o Teorema 1.2 pelo qual $R = \text{Im } \mathcal{C}(A, B)$. Por definição,

$$\mathcal{C}(A, B) = \begin{bmatrix} B & AB & A^2B & A^3B \end{bmatrix}.$$

Vamos calcular as colunas de $\mathcal{C}(A, B)$.

$$\begin{aligned} \bullet AB &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2\omega \\ 0 & 1 \\ -2\omega & 0 \end{bmatrix}; \\ \bullet A^2B &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 3\omega^2 & 0 & 0 & 2\omega \\ 0 & -\omega^2 & 0 & 0 \\ 0 & -2\omega & 0 & 0 \\ -6\omega^3 & 0 & 0 & -4\omega^2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 2\omega \\ -\omega^2 & 0 \\ -2\omega & 0 \\ 0 & -4\omega^2 \end{bmatrix}; \end{aligned}$$

$$\begin{aligned}
\bullet A^3 B &= \underbrace{\begin{bmatrix} 3\omega^2 & 0 & 0 & 2\omega \\ 0 & -\omega^2 & 0 & 0 \\ 0 & -2\omega & 0 & 0 \\ -6\omega^3 & 0 & 0 & -4\omega^2 \end{bmatrix}}_{A^2} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 0 & -\omega^2 & 0 & 0 \\ -3\omega^4 & 0 & 0 & -2\omega^3 \\ -6\omega^3 & 0 & 0 & -4\omega^2 \\ 0 & 2\omega^3 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -\omega^2 & 0 \\ 0 & -2\omega^3 \\ 0 & -4\omega^2 \\ 2\omega^3 & 0 \end{bmatrix}.
\end{aligned}$$

Desse modo,

$$\mathcal{C}(A, B) = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 2\omega & -\omega^2 & 0 \\ 1 & 0 & 0 & 2\omega & -\omega^2 & 0 & 0 & -2\omega^3 \\ 0 & 0 & 0 & 1 & -2\omega & 0 & 0 & -4\omega^2 \\ 0 & 1 & -2\omega & 0 & 0 & -4\omega^2 & 2\omega^3 & 0 \end{bmatrix}.$$

Agora, usemos operações elementares sobre linhas para escrever a matriz $\mathcal{C}(A, B)$ na forma escalonada reduzida.

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 2\omega & -\omega^2 & 0 \\ 1 & 0 & 0 & 2\omega & -\omega^2 & 0 & 0 & -2\omega^3 \\ 0 & 0 & 0 & 1 & -2\omega & 0 & 0 & -4\omega^2 \\ 0 & 1 & -2\omega & 0 & 0 & -4\omega^2 & 2\omega^3 & 0 \end{bmatrix} \xrightarrow{L_1 \leftrightarrow L_2}$$

$$\begin{bmatrix} 1 & 0 & 0 & 2\omega & -\omega^2 & 0 & 0 & -2\omega^3 \\ 0 & 0 & 1 & 0 & 0 & 2\omega & -\omega^2 & 0 \\ 0 & 0 & 0 & 1 & -2\omega & 0 & 0 & -4\omega^2 \\ 0 & 1 & -2\omega & 0 & 0 & -4\omega^2 & 2\omega^3 & 0 \end{bmatrix} \xrightarrow{L_2 \leftrightarrow L_3}$$

$$\begin{bmatrix} 1 & 0 & 0 & 2\omega & -\omega^2 & 0 & 0 & -2\omega^3 \\ 0 & 0 & 0 & 1 & -2\omega & 0 & 0 & -4\omega^2 \\ 0 & 0 & 1 & 0 & 0 & 2\omega & -\omega^2 & 0 \\ 0 & 1 & -2\omega & 0 & 0 & -4\omega^2 & 2\omega^3 & 0 \end{bmatrix} \xrightarrow{L_2 \leftrightarrow L_4}$$

$$\begin{bmatrix} 1 & 0 & 0 & 2\omega & -\omega^2 & 0 & 0 & -2\omega^3 \\ 0 & 1 & -2\omega & 0 & 0 & -4\omega^2 & 2\omega^3 & 0 \\ 0 & 0 & 1 & 0 & 0 & 2\omega & -\omega^2 & 0 \\ 0 & 0 & 0 & 1 & -2\omega & 0 & 0 & -4\omega^2 \end{bmatrix} \xrightarrow{L_2 \leftrightarrow L_2 + 2\omega L_3}$$

$$\begin{bmatrix} 1 & 0 & 0 & 2\omega & -\omega^2 & 0 & 0 & -2\omega^3 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 2\omega & -\omega^2 & 0 \\ 0 & 0 & 0 & 1 & -2\omega & 0 & 0 & -4\omega^2 \end{bmatrix} \xrightarrow{L_1 \leftrightarrow L_1 - 2\omega L_4}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 3\omega^2 & 0 & 0 & 6\omega^3 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 2\omega & -\omega^2 & 0 \\ 0 & 0 & 0 & 1 & -2\omega & 0 & 0 & -4\omega^2 \end{bmatrix} .$$

Uma vez que o posto da matriz $\mathcal{C}(A, B)$ é o número de linhas não nulas de sua forma escalonada reduzida, temos que $\text{Posto } \mathcal{C}(A, B) = 4$, isto é, $\dim \text{Im} \mathcal{C}(A, B) = 4 = \dim \mathbb{R}^4$ e como $\text{Im } \mathcal{C}(A, B) \subset \mathbb{R}^4$ concluímos que $\text{Im } \mathcal{C}(A, B) = \mathbb{R}^4$, ou seja,

$$R = \mathbb{R}^4 .$$

Isto significa que todo $x \in \mathbb{R}^4$ pode ser alcançado a partir de 0 em um tempo $T > 0$ através da escolha adequada de $u \in PC([0, T], \mathbb{R}^2)$.

(b) Substituindo B por B_1 temos

$$\mathcal{C}(A, B) = \begin{bmatrix} B_1 & AB_1 & A^2B_1 & A^3B_1 \end{bmatrix}.$$

Observe então que:

$$AB_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2\omega \end{bmatrix},$$

$$A^2B_1 = \begin{bmatrix} 3\omega^2 & 0 & 0 & 2\omega \\ 0 & -\omega^2 & 0 & 0 \\ 0 & -2\omega & 0 & 0 \\ -6\omega^3 & 0 & 0 & -4\omega^2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -\omega^2 \\ -2\omega \\ 0 \end{bmatrix},$$

$$A^3B_1 = \begin{bmatrix} 0 & -\omega^2 & 0 & 0 \\ -3\omega^4 & 0 & 0 & -2\omega^3 \\ -6\omega^3 & 0 & 0 & -4\omega^2 \\ 0 & 2\omega^3 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -\omega^2 \\ 0 \\ 0 \\ 2\omega^3 \end{bmatrix}.$$

Portanto,

$$\mathcal{C}(A, B_1) = \begin{bmatrix} 0 & 1 & 0 & -\omega^2 \\ 1 & 0 & -\omega^2 & 0 \\ 0 & 0 & -2\omega & 0 \\ 0 & -2\omega & 0 & 2\omega^3 \end{bmatrix}.$$

Sabemos que $\text{Im } \mathcal{C}(A, B_1) = \{\mathcal{C}(A, B_1)z : z \in \mathbb{R}^4\}$. Dado $z = (a, b, c, d) \in \mathbb{R}^4$ temos

$$\mathcal{C}(A, B_1)z = \begin{bmatrix} 0 & 1 & 0 & -\omega^2 \\ 1 & 0 & -\omega^2 & 0 \\ 0 & 0 & -2\omega & 0 \\ 0 & -2\omega & 0 & 2\omega^3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \cdot a + b + 0 \cdot c - \omega^2 d \\ a + 0 \cdot b - \omega^2 c + 0 \cdot d \\ 0 \cdot a + 0 \cdot b - 2\omega c + 0 \cdot d \\ 0 \cdot a - 2\omega b + 0 \cdot c + 2\omega^3 d \end{bmatrix} \Rightarrow$$

$$\Rightarrow \mathcal{C}(A, B_1)z = a \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2\omega \end{bmatrix} + c \begin{bmatrix} 0 \\ -\omega^2 \\ -2\omega \\ 0 \end{bmatrix} + d \begin{bmatrix} -\omega^2 \\ 0 \\ 0 \\ 2\omega^3 \end{bmatrix} \Rightarrow$$

$$\begin{aligned} \Rightarrow \mathcal{C}(A, B_1)z &= a \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2\omega \end{bmatrix} + c \begin{bmatrix} 0 \\ -\omega^2 \\ -2\omega \\ 0 \end{bmatrix} - \omega^2 d \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2\omega \end{bmatrix} \Rightarrow \\ \Rightarrow \mathcal{C}(A, B_1)z &= a \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + (b - \omega^2 d) \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2\omega \end{bmatrix} + c \begin{bmatrix} 0 \\ -\omega^2 \\ -2\omega \\ 0 \end{bmatrix}. \end{aligned}$$

Portanto, $R = \text{Im } \mathcal{C}(A, B_1) = [v_1, v_2, v_3]$ onde $v_1 = (0, 1, 0, 0)$, $v_2 = (1, 0, 0, -2\omega)$ e $v_3 = (0, -\omega^2, -2\omega, 0)$ ².

□

Exemplo 2.2 Conforme observado anteriormente $\text{Posto } \mathcal{C}(A, B) = 4$ e então, pelo Teorema 1.3, temos que o sistema (2.26) é controlável. Por outro lado, escrevendo $\mathcal{C}(A, B_1)$ forma escalonada reduzida tem-se

$$\begin{aligned} \begin{bmatrix} 0 & 1 & 0 & -\omega^2 \\ 1 & 0 & -\omega^2 & 0 \\ 0 & 0 & -2\omega & 0 \\ 0 & -2\omega & 0 & 2\omega^3 \end{bmatrix} &\xrightarrow{L_1 \leftrightarrow L_2} \begin{bmatrix} 1 & 0 & -\omega^2 & 0 \\ 0 & 1 & 0 & -\omega^2 \\ 0 & 0 & -2\omega & 0 \\ 0 & -2\omega & 0 & 2\omega^3 \end{bmatrix} \xrightarrow{L_4 \leftrightarrow L_4 + 2\omega L_2} \\ \begin{bmatrix} 1 & 0 & -\omega^2 & 0 \\ 0 & 1 & 0 & -\omega^2 \\ 0 & 0 & -2\omega & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} &\xrightarrow{L_3 \leftrightarrow -\frac{L_3}{2\omega}} \begin{bmatrix} 1 & 0 & -\omega^2 & 0 \\ 0 & 1 & 0 & -\omega^2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{L_1 \leftrightarrow L_1 + \omega^2 L_3} \\ &\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -\omega^2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \end{aligned}$$

de onde tiramos que $\text{Posto } \mathcal{C}(A, B_1) = 3 < 4$ de modo que, pelo Teorema 1.3, o par (A, B_1) não é controlável. Isto nos diz que se $u_2 = 0$, ou seja, na ausência de força na

² $[v_1, v_2, v_3]$ denota o subespaço gerado pelos vetores v_1 , v_2 e v_3 , isto é, o espaço formado por todos os vetores que são combinações lineares destes três.

direção tangencial o sistema não é controlável. Agora fazamos $u_1 = 0$ trocando B por $B_2 = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^*$. Temos então:

$$AB_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2\omega \\ 1 \\ 0 \end{bmatrix},$$

$$A^2B_2 = \begin{bmatrix} 3\omega^2 & 0 & 0 & 2\omega \\ 0 & -\omega^2 & 0 & 0 \\ 0 & -2\omega & 0 & 0 \\ -6\omega^3 & 0 & 0 & -4\omega^2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2\omega \\ 0 \\ 0 \\ -4\omega^2 \end{bmatrix},$$

$$A^3B_2 = \begin{bmatrix} 0 & -\omega^2 & 0 & 0 \\ -3\omega^4 & 0 & 0 & -2\omega^3 \\ -6\omega^3 & 0 & 0 & -4\omega^2 \\ 0 & 2\omega^3 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -2\omega^3 \\ -4\omega^2 \\ 0 \end{bmatrix}.$$

Logo,

$$\mathcal{C}(A, B_2) = \begin{bmatrix} 0 & 0 & 2\omega & 0 \\ 0 & 2\omega & 0 & -2\omega^3 \\ 0 & 1 & 0 & -4\omega^2 \\ 1 & 0 & -4\omega^2 & 0 \end{bmatrix}.$$

Escalonando $\mathcal{C}(A, B_2)$ temos:

$$\begin{bmatrix} 0 & 0 & 2\omega & 0 \\ 0 & 2\omega & 0 & -2\omega^3 \\ 0 & 1 & 0 & -4\omega^2 \\ 1 & 0 & -4\omega^2 & 0 \end{bmatrix} \xrightarrow{\substack{L_1 \leftrightarrow L_4 \\ L_2 \leftrightarrow L_3}} \begin{bmatrix} 1 & 0 & -4\omega^2 & 0 \\ 0 & 1 & 0 & -4\omega^2 \\ 0 & 2\omega & 0 & -2\omega^3 \\ 0 & 0 & 2\omega & 0 \end{bmatrix} \xrightarrow{\substack{L_3 \leftrightarrow L_3/2\omega \\ L_4 \leftrightarrow L_4/2\omega}} \begin{bmatrix} 1 & 0 & -4\omega^2 & 0 \\ 0 & 1 & 0 & -4\omega^2 \\ 0 & 1 & 0 & -\omega^2 \\ 0 & 0 & 1 & 0 \end{bmatrix} \xrightarrow{L_3 \leftrightarrow L_3 - L_2} \begin{bmatrix} 1 & 0 & -4\omega^2 & 0 \\ 0 & 1 & 0 & -4\omega^2 \\ 0 & 0 & 0 & 3\omega^2 \\ 0 & 0 & 1 & 0 \end{bmatrix} \xrightarrow{L_3 \leftrightarrow L_4}$$

$$\begin{array}{ccc}
\begin{bmatrix} 1 & 0 & -4\omega^2 & 0 \\ 0 & 1 & 0 & -4\omega^2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3\omega^2 \end{bmatrix} & \begin{array}{l} L_1 \leftrightarrow L_1 + 4\omega^2 L_3 \\ L_4 \leftrightarrow L_4 / 3\omega^2 \end{array} & \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -4\omega^2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
& & \begin{array}{l} L_2 \leftrightarrow L_2 + 4\omega^2 L_4 \end{array} \\
& & \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} .
\end{array}$$

Portanto, $\text{Posto } \mathcal{C}(A, B_2) = 4$ e pelo Teorema 1.3 resulta que o par (A, B_2) é controlável. Concluimos então que a ausência de força na direção radial não descaracteriza a controlabilidade. □

Exemplo 2.3 No exemplo 2.2 vimos, usando o Teorema 1.3 que o sistema (2.26), com A e B , dados por (2.27) é controlável. Verificaremos aqui que tal conclusão pode também ser alcançada pelo uso do Teorema 1.4. Dado $s \in \mathbb{C}$ arbitrário temos

$$\begin{bmatrix} sI - A & B \end{bmatrix} = \begin{bmatrix} s & -1 & 0 & 0 & 0 & 0 \\ -3\omega^2 & s & 0 & -2\omega & 1 & 0 \\ 0 & 0 & s & -1 & 0 & 0 \\ 0 & 2\omega & 0 & s & 0 & 1 \end{bmatrix} .$$

Sejam $a, b, c, d \in \mathbb{R}$ tais que

$$a(s, -1, 0, 0, 0, 0) + b(-3\omega^2, s, 0, -2\omega, 1, 0) + c(0, 0, s, -1, 0, 0) + d(0, 2\omega, 0, s, 0, 1) = (0, 0, 0, 0, 0, 0) .$$

Daí,

$$\begin{aligned}
(as, -a, 0, 0, 0, 0) + (-b3\omega^2, bs, 0, -b2\omega, b, 0) + (0, 0, cs, -c, 0, 0) + (0, d2\omega, 0, ds, 0, d) &= (0, 0, 0, 0, 0, 0) \Rightarrow \\
\Rightarrow (as - 3b\omega^2, -a + bs + 2d\omega, cs, -2b\omega - c + ds, b, d) &= (0, 0, 0, 0, 0, 0) .
\end{aligned}$$

Temos então o sistema

$$\begin{cases} as - 3b\omega^2 = 0 \\ -a + bs + 2d\omega = 0 \\ cs = 0 \\ -2b\omega - c + ds = 0 \\ b = 0 \\ d = 0 \end{cases}$$

de onde segue imediatamente que $b = d = 0$. Substituindo isto em $-2b\omega - c + ds = 0$ e em $-a + bs + 2d\omega = 0$ resulta, no primeiro caso, que $c = 0$ e no segundo que $a = 0$. Portanto as 4 linhas de $(sI - A, B)$ são *L.I.* o que nos diz que $\text{Posto}(sI - A, B) = 4$ e então pelo Teorema 1.4, o sistema (2.26) é controlável. Obviamente pelo mesmo teorema podem ser alcançadas àquelas conclusões sobre o sistema (2.26) para $u_1 = 0$ bem como para $u_2 = 0$. □

Exemplo 2.4 Verifique se o sistema (2.26) é observável.

Solução: Temos

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \end{bmatrix} \text{ e } C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} .$$

A matriz de observabilidade é dada por

$$\mathcal{O}(C, A) = \begin{bmatrix} C \\ CA \\ CA^2 \\ CA^3 \end{bmatrix} \in \mathbb{R}^{8 \times 4} .$$

Relembre do exemplo 2.1 que

$$A^2 = \begin{bmatrix} 3\omega^2 & 0 & 0 & 2\omega \\ 0 & -\omega^2 & 0 & 0 \\ 0 & -2\omega & 0 & 0 \\ -6\omega^3 & 0 & 0 & -4\omega^2 \end{bmatrix} \text{ e } A^3 = \begin{bmatrix} 0 & -\omega^2 & 0 & 0 \\ -3\omega^4 & 0 & 0 & -2\omega^3 \\ -6\omega^3 & 0 & 0 & -4\omega^2 \\ 0 & 2\omega^3 & 0 & 0 \end{bmatrix} .$$

Com isto, vamos calcular $\mathcal{O}(C, A)$.

$$\bullet CA = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} ;$$

$$\begin{aligned}
\bullet CA^2 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 3\omega^2 & 0 & 0 & 2\omega \\ 0 & -\omega^2 & 0 & 0 \\ 0 & -2\omega & 0 & 0 \\ -6\omega^3 & 0 & 0 & -4\omega^2 \end{bmatrix} = \begin{bmatrix} 3\omega^2 & 0 & 0 & 2\omega \\ 0 & -2\omega & 0 & 0 \end{bmatrix}; \\
\bullet CA^3 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -\omega^2 & 0 & 0 \\ -3\omega^4 & 0 & 0 & -2\omega^3 \\ -6\omega^3 & 0 & 0 & -4\omega^2 \\ 0 & 2\omega^3 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -\omega^2 & 0 & 0 \\ -6\omega^3 & 0 & 0 & -4\omega^2 \end{bmatrix}.
\end{aligned}$$

Logo,

$$\mathcal{O}(C, A) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & -2\omega & 0 & 0 \\ 0 & -\omega^2 & 0 & 0 \\ -6\omega^3 & 0 & 0 & -4\omega^2 \end{bmatrix}.$$

Agora, usemos operações elementares sobre as linhas de $\mathcal{O}(C, A)$ para colocá-la na forma escalonada reduzida.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & -2\omega & 0 & 0 \\ 0 & -\omega^2 & 0 & 0 \\ -6\omega^3 & 0 & 0 & -4\omega^2 \end{bmatrix} \xrightarrow{\substack{L_8 \leftrightarrow L_8 + 2\omega L_5 \\ L_7 \leftrightarrow L_7 - \omega L_5/2}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & -2\omega & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{\substack{L_6 \leftrightarrow L_6 + 2\omega L_3 \\ L_5 \leftrightarrow L_5 - 3\omega^2 L_1}}$$

$$\begin{array}{ccc}
\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & \begin{array}{c} L_5 \leftrightarrow L_5 \xrightarrow{-2\omega L_4} \\ L_3 \leftrightarrow L_2 \xrightarrow{\quad} \end{array} & \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}
\end{array}$$

Como o número de linhas não nulas da forma escalonada reduzida de $\mathcal{O}(C, A)$ é 4, concluímos que $\text{Posto } \mathcal{O}(C, A) = 4$ e portanto, pelo Teorema 1.7 que o sistema 2.26 é observável.

□

Apêndice A

A.1 Álgebra Linear e Teoria das Matrizes

Proposição A.1 (Caracterização da soma direta de dois subespaços) *Sejam V um espaço vetorial sobre \mathbb{R} e U e W subespaços de V . Então $V = U \oplus W$ se, e somente se, para todo $v \in V$ existem únicos $u \in U$ e $w \in W$ tais que $v = u + w$.*

Demonstração: Ver em [2]. ■

Teorema A.1 *Seja V um espaço vetorial sobre \mathbb{R} com $\dim V = n$ e W um subespaço de V . Então,*

(i) W é de dimensão finita e $\dim W \leq \dim V$;

(ii) $\dim W = \dim V$ se, e somente se, $W = V$.

Demonstração: Ver [2]. ■

Proposição A.2 *Seja V um espaço vetorial não nulo e W_1 um subespaço vetorial de V . Então existe um subespaço $W_2 \subset V$, tal que*

$$V = W_1 \oplus W_2 .$$

Demonstração: Ver em [2]. ■

Teorema A.2 *Sejam V um espaço vetorial e W_1, W_2, \dots, W_m subespaços vetoriais de V . Assuma que $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_m$ são bases de W_1, W_2, \dots, W_m respectivamente. Então V é soma direta dos W_i , com $1 \leq i \leq m$, se, e somente se, $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_m$ é base de V .*

Demonstração: Ver em [2]. ■

Teorema A.3 . Dados dois conjuntos V e W , uma função $T : V \rightarrow W$ tem uma inversa à esquerda se, e somente se, T é injetora.

Demonstração: Ver em [8]. ■

Teorema A.4 *Dados dois conjuntos V e W , uma função $T : V \rightarrow W$ tem uma inversa à direita se, e somente se, T é sobrejetora.*

Demonstração: Ver em [8]. ■

Teorema A.5 (Teorema do Núcleo e da Imagem) *Sejam V e W espaços vetoriais sobre \mathbb{R} e $T : V \rightarrow W$ uma transformação linear. Então,*

$$\dim \operatorname{Im}(T) + \dim \ker(T) = \dim V .$$

Demonstração: Ver [2]. ■

Sabemos que dadas as bases β e γ de V e W , respectivamente, com $\#\beta = n$ e $\#\gamma = m$, podemos escrever uma representação matricial para T a qual denotamos por $[T]_{\gamma}^{\beta}$ pertencente a $M_{m \times n}(\mathbb{R})$. Considerando o espaço,

$$\mathcal{L}(V, W) = \{T : V \rightarrow W; T \text{ é uma transformação linear}\}$$

o Teorema da Representação Matricial¹, nos garante que $\mathcal{L}(V, W)$ é isomorfo a $M_{m \times n}(\mathbb{R})$. Deste modo, dada a matriz $A \in M_{m \times n}(\mathbb{R})$ faz sentido os conceitos de imagem e núcleo de A . Consequentemente, é coerente a enunciação do Teorema do Núcleo e da Imagem para matrizes.

Definição A.1 *Seja $A \in \mathbb{F}^{n \times m}$. A **imagem** de A é o conjunto*

$$\operatorname{Im}A = \{Ax : x \in \mathbb{F}^m\}$$

e o **núcleo** de A é o conjunto

$$\ker A = \{x \in \mathbb{F}^m : Ax = 0\} .$$

Note que $\operatorname{Im} A \subset \mathbb{F}^n$ e $\ker A \subset \mathbb{F}^m$.

Teorema A.6 (Teorema do Núcleo e da Imagem para matrizes) *Seja $M \in \mathbb{F}^{m \times n}$. Então,*

$$\dim \operatorname{Im}M + \dim \ker M = n .$$

Demonstração: Ressaltamos que $\operatorname{Im}M \subset \mathbb{F}^n$ e $\ker M \subset \mathbb{F}^n$. Assim $\dim \ker M \leq n$. Se $\dim \ker M = n$ o resultado é imediato. Suponhamos $\dim \ker M = r < n$ e seja

$$\beta = \{x_1, \dots, x_r\}$$

¹Para uma maior compreensão sobre transformações lineares e representação matricial de uma transformação, fica como sugestão a referência [2].

uma base para $\ker M$. Desse modo, o conjunto β é *L.I.* e pelo Teorema do completamento (veja referência [2]) o mesmo pode ser completado de modo a formar uma base para \mathbb{F}^n . Seja então $\gamma = \{x_1, \dots, x_r, x_{r+1}, \dots, x_n\}$ a referida base. Vamos mostrar que

$$\alpha = \{Mx_{r+1}, \dots, Mx_n\}$$

é uma base para $\text{Im}M$. Com efeito, dado $z \in \text{Im}M$ temos $z = Mv$ com $v \in \mathbb{F}^n$. Existem então escalares $a_1, \dots, a_r, a_{r+1}, \dots, a_n$ tais que

$$v = a_1x_1 + \dots + a_rx_r + a_{r+1}x_{r+1} + \dots + a_nx_n$$

ou melhor

$$z = a_1Mx_1 + \dots + a_rMx_r + a_{r+1}Mx_{r+1} + \dots + a_nMx_n$$

e como $x_i \in \ker M$ para $1 \leq i \leq r$ temos $Mx_1 = Mx_2 = \dots = Mx_r = 0$ e então

$$z = a_{r+1}Mx_{r+1} + \dots + a_nMx_n .$$

Logo, $z \in [Mx_{r+1}, \dots, Mx_n]$, ou seja, $\text{Im}M \subset [Mx_{r+1}, \dots, Mx_n]$. Como já temos naturalmente $[Mx_{r+1}, \dots, Mx_n] \subset \text{Im}M$ resulta que

$$\text{Im}M = [Mx_{r+1}, \dots, Mx_n] .$$

Sejam c_{r+1}, \dots, c_n tais que $c_{r+1}Mx_{r+1} + \dots + c_nMx_n = 0$. Daí,

$$M(c_{r+1}x_{r+1} + \dots + c_nx_n) = 0$$

o que nos diz que $c_{r+1}x_{r+1} + \dots + c_nx_n \in \ker M$. Existem então escalares c_1, \dots, c_r satisfazendo

$$c_{r+1}x_{r+1} + \dots + c_nx_n = c_1x_1 + c_2x_2 + \dots + c_rx_r$$

de modo que

$$c_1x_1 + c_2x_2 + \dots + c_rx_r + (-c_{r+1})x_{r+1} + \dots + (-c_n)x_n = 0$$

e como γ é *L.I.* (pois é uma base para \mathbb{F}^n) segue-se que

$$c_1 = c_2 = \dots = c_r = -c_{r+1} = \dots = -c_n = 0$$

e particularmente que $c_{r+1} = \dots = c_n = 0$. Assim, $[Mx_{r+1}, \dots, Mx_n]$ é *L.I.* e portanto é uma base para $\text{Im}M$. Uma vez que $\#\{Mx_{r+1}, \dots, Mx_n\} = n - r$ resulta que $\dim \text{Im}M = n - r$ e como já tínhamos $\dim \ker M = r$ tem-se por conseguinte

$$\dim \text{Im}M + \dim \ker M = n - r + r$$

$$\dim \text{Im}M + \dim \ker M = n \quad .$$

■

Proposição A.3 *Seja $T : V \rightarrow W$ uma transformação linear. Então T é injetiva se, e somente se, $\ker(T) = \{0\}$.*

Demonstração: Ver em [2].

■

Definição A.2 . Seja $A = (a_{ij}) \in \mathbb{C}^{n \times m}$. Definimos o *Conjugado Transposto* ou *Transposto Hermitiano* da matriz A como sendo a matriz $A^* \in \mathbb{C}^{m \times n}$, cujas entradas são os correspondentes conjugados complexos das entradas da transposta de A .

Observe que se $A \in \mathbb{R}^{n \times m}$ o conjugado complexo pode ser ignorado e então A^* é simplesmente a transposta de A . Além disso, o conjugado transposto de uma matriz A agrega as seguintes propriedades:

(i)

$$\langle Ax, y \rangle = \langle x, A^*y \rangle \quad \forall x \in \mathbb{F}^m, \quad \forall y \in \mathbb{F}^n ; \quad (\text{A.1})$$

(ii) Se $n = m$, então $(e^A)^* = e^{A^*}$;

(iii) Se A é uma matriz invertível, então $(A^{-1})^* = (A^*)^{-1}$.

Definição A.3 . Uma matriz $A \in \mathbb{R}^{n \times n}$ é dita positiva semi-definida quando $x^*Ax \geq 0$ para todo $x \in \mathbb{R}^n$.

Teorema A.7 . Seja $M \in \mathbb{F}^{n \times m}$. Então

$$(\text{Im}M)^\perp = \ker M^* \quad .$$

Equivalentemente, $x \in (\text{Im}M)^\perp$ se, e somente se, $x^*M = 0$.

Demonstração: Se $x \in (\text{Im}M)^\perp$ então $\langle Mz, x \rangle = 0$ para todo $z \in \mathbb{F}^m$. Daí, por A.1 temos que $\langle z, M^*x \rangle = 0$ para todo $z \in \mathbb{F}^m$. Fazendo $z = (a_1, a_2, \dots, a_m)$ e $M^*x = (b_1, b_2, \dots, b_m)$ segue-se que

$$a_1b_1 + a_2b_2 + \dots + a_mb_m = 0 \quad , \quad \forall (a_1, a_2, \dots, a_m) = z \in \mathbb{F}^m .$$

Do caso particular em que $z = (1, 0, \dots, 0)$ obtemos $b_1 = 0$. Analogamente do caso especial $z = (0, 1, 0, \dots, 0)$ vem que $b_2 = 0$. Em geral, para cada caso em que

$$\begin{cases} a_k = 1 \\ a_i = 0, \quad i \neq k \end{cases} \quad k, i \in (1, 2, \dots, m)$$

obtemos $b_k = 0$. Assim, $b_1 = b_2 = \dots = b_m = 0$, isto é, $M^*x = 0$ de forma que $x \in \ker M^*$. Logo,

$$(\operatorname{Im}M)^\perp \subset \ker M^*.$$

Considere agora $x \in \ker M^*$. Sendo assim, $M^*x = 0$ e por conseguinte $\langle z, M^*x \rangle = 0$ para todo $z \in \mathbb{F}^m$. Daí, por (A.1) resulta $\langle Mz, x \rangle = 0$ para todo $z \in \mathbb{F}^m$, ou seja, $x \in (\operatorname{Im}M)^\perp$ mostrando que

$$\ker M^* \subset (\operatorname{Im}M)^\perp .$$

Das duas inclusões verificadas segue-se que

$$(\operatorname{Im}M)^\perp = \ker M^* .$$

■

Definição A.4 . Seja $A \in \mathbb{R}^{n \times m}$. O Posto de A , denotado por “*Posto A*” é definido por

$$\operatorname{Posto} A := \dim \operatorname{Im}A .$$

O posto de uma matriz pode também ser definido como o seu número de linhas (ou colunas) LI que coincide ainda com o número de linhas $L.I$ da sua forma escalonada, sendo demonstrada a equivalência das definições.

O posto de $A \in \mathbb{R}^{n \times m}$ é dito completo se $\operatorname{Posto} A = \min\{n, m\}$. Observe que

- Se $n \leq m$, então o posto de A é completo se, e somente se, $\operatorname{Im}A = \mathbb{R}^n$.
- Se $n \geq m$, então o posto de A é completo se, e somente se, $\ker A = \{0\}$.

Teorema A.8 .

1. Se $n \leq m$, então existe $A^\# \in \mathbb{R}^{m \times n}$ tal que $AA^\#x = x$ para todo $x \in \operatorname{Im}A$. Em particular, se A é de posto completo, então $AA^\# = I$, ou seja, A possui uma inversa à direita.
2. Se $n \geq m$, então existe $A^\# \in \mathbb{R}^{m \times n}$ tal que $A^\#Ax = x$ para todo $x \in (\ker A)^\perp$. Em particular, se A possui posto completo então $A^\#A = I$, isto é, A possui uma inversa à esquerda.

Demonstração: Ver em [1].

■

Proposição A.4 . Seja $A \in \mathbb{F}^{n \times n}$ uma matriz quadrada. A é invertível se, e somente se, $\dim \text{Im}A = n$.

Demonstração: Suponha que A é invertível. Então $\det A \neq 0$. Daí, pelo *teorema da combinação linear* das linhas (ou colunas) de uma matriz (veja em [5]) segue-se que todas as colunas de A são *L.I.* . Considere então o conjunto $\beta = \{v_1, v_2, \dots, v_n\}$; em que para cada $1 \leq i \leq n$, v_i denota o vetor do \mathbb{F}^n composto pela i -ésima coluna de A . Vamos mostrar que este conjunto é uma base para $\text{Im}A$. Note que $\beta \subset \text{Im}A$. Com efeito, considerando o vetor $x_i \in \mathbb{F}^n$ em que a i -ésima entrada é igual a 1 e todas as outras são iguais a 0, temos que $Ax_i = v_i$ mostrando que $v_i \in \text{Im}A$ para todo $1 \leq i \leq n$. Além disto é imediato de sua definição que β é *L.I.* . Resta-nos mostrar que β gera $\text{Im}A$. Para isto façamos

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} .$$

Para qualquer $x = (x_1, x_2, \dots, x_n) \in \mathbb{F}^n$ tem-se

$$\begin{aligned} Ax &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{bmatrix} \\ &= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix} = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n , \end{aligned}$$

mostrando-nos que $Ax \in [v_1, v_2, \dots, v_n]$ e conseqüentemente $\text{Im}A \subset [v_1, v_2, \dots, v_n]$. Como a inclusão $[v_1, v_2, \dots, v_n] \subset \text{Im}A$ é óbvia (pelo fato de que $\text{Im}A$ é um subespaço do \mathbb{F}^n) concluímos que

$$\text{Im}A = [v_1, v_2, \dots, v_n] ,$$

ou seja, β gera $\text{Im}A$ e portanto é uma base para este último. Uma vez que $\#\beta = n$, segue-se que $\dim \text{Im}A = n$.

(\Leftrightarrow) Assuma agora que $\dim \operatorname{Im} A = n$. Neste caso A é de posto completo e então pelo Teorema A.8 segue-se que A possui inversas à esquerda e à direita, ou seja, A é invertível. ■

Desde que $\dim \operatorname{Im} A + \dim \ker A = n$ temos equivalentemente à Proposição A.4 que, A é invertível se, e somente se, $\ker A = \{0\}$.

Proposição A.5 . Sejam $M \in \mathbb{R}^{n \times p}$ e $L \in \mathbb{R}^{p \times q}$. Então

$$\operatorname{Posto}(ML) \leq \min\{\operatorname{Posto} M, \operatorname{Posto} L\} .$$

Demonstração: Ver em [1]. ■

Definição A.5 . Seja $A \in \mathbb{R}^{n \times n}$. O polinômio característico de A é o polinômio mônico de grau n definido por

$$P_A(s) = \det(sI - A) .$$

Teorema A.9 (Cayley-Hamilton). Uma matriz $A \in \mathbb{R}^{n \times n}$ é um zero de seu polinômio característico, isto é

$$P_A(A) = A^n + a_{n-1}A^{n-1} + a_{n-2}A^{n-2} + \cdots + a_1A + a_0I = 0$$

Demonstração: Ver em [2]. ■

Definição A.6 . Dizemos que duas matrizes $A, B \in \mathbb{R}^{n \times n}$ são semelhantes quando existe uma matriz $S \in GL(n, \mathbb{R}) := \{M \in \mathbb{R}^{n \times n} : \det M \neq 0\}$, tal que $A = S^{-1}BS$.

Proposição A.6 . Matrizes semelhantes têm o mesmo polinômio característico.

Demonstração: Ver em [2]. ■

Teorema A.10 (Teorema da Decomposição ortogonal). Seja V um Espaço Euclidiano de dimensão finita e $U \subseteq V$ um subespaço de V . Então

$$V = U \oplus U^\perp$$

e

$$\|v\|^2 = \|u\|^2 + \|w\|^2$$

onde $v \in V$, $u \in U$, $w \in U^\perp$ e $v = u + w$.

Demonstração: Ver em [2]. ■

Definição A.7 . Dada uma matriz $A \in \mathbb{R}^{n \times n}$, a exponencial de A é definida por

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

Proposição A.7 . Seja S uma matriz invertível. Então, $e^{S^{-1}AS} = S^{-1}e^AS$.

Demonstração: Ver em [3]. ■

Definição A.8 . Seja $A \in \mathbb{C}^{n \times n}$. Um autovalor de A é um escalar $\lambda \in \mathbb{C}$ tal que $Av = \lambda v$ para algum $v \in \mathbb{C}^n$, $v \neq 0$. Um vetor v nestas condições é chamado *autovetor* de A associado a λ .

Proposição A.8 *Seja $A \in \mathbb{C}^{n \times n}$. Então,*

- (i) *Um escalar $\lambda \in \mathbb{C}$ é um autovalor de A se, e somente se, $\ker(\lambda I - A) \neq \{0\}$;*
- (ii) *Um autovetor de A associado a λ é um vetor não nulo do espaço $\ker(\lambda I - A)$.*

Demonstração:

- (i) (\Rightarrow) Se $\lambda \in \mathbb{C}$ é um autovalor de A então existe $v \in \mathbb{C}^n$, $v \neq 0$, tal que $Av = \lambda v$.

Logo,

$$\lambda v - Av = 0 \Rightarrow (\lambda I - A)v = 0 \Rightarrow v \in \ker(\lambda I - A)$$

e como $v \neq 0$ segue-se que $\ker(\lambda I - A) \neq \{0\}$.

(\Leftarrow) Já se $\ker(\lambda I - A) \neq \{0\}$, existe $v \in \mathbb{C}^n$, $v \neq 0$, tal que $v \in \ker(\lambda I - A)$, ou seja,

$$(\lambda I - A)v = 0 \Rightarrow \lambda v - Av = 0 \Rightarrow Av = \lambda v$$

e segue-se por definição que λ é um autovalor de A .

- (ii) Veja na primeira parte de (i). ■

Definição A.9 . Seja $A \in \mathbb{C}^{n \times n}$. Definimos o *espectro* de A denotado por $\sigma(A)$ como sendo o conjunto de todos os autovalores de A , ou seja, dado $\lambda \in \mathbb{C}$

$$\lambda \in \sigma(A) \Leftrightarrow Av = \lambda v,$$

para algum $v \in \mathbb{C}^n$ e $v \neq 0$.

Teorema A.11 . Se $M \in \mathbb{F}^{n \times n}$ é uma matriz da forma

$$M = \begin{bmatrix} M_1 & M_2 \\ 0 & M_3 \end{bmatrix}$$

onde $M_1 \in \mathbb{F}^{n_1 \times n_1}$, $M_2 \in \mathbb{F}^{n_1 \times n_2}$, $M_3 \in \mathbb{F}^{n_2 \times n_2}$ e $n_1 + n_2 = n$. Então

$$\sigma(M) = \sigma(M_1) \cup \sigma(M_3) \text{ .}$$

Demonstração: Ver em [1]. ■

Obviamente, o resultado acima é pode ser estendido para matrizes triangulares de blocos com p blocos na diagonal, para p arbitrário.

Teorema A.12 .(Regra de Cramer). Se $M \in \mathbb{F}^{n \times n}$, então

$$M(\text{adj}M) = (\text{adj}M)M = (\det M)I,$$

onde $\text{adj}M$ denota a matriz adjunta de M .

Demonstração: Ver em [2]. ■

Uma importante consequência deste Teorema é a seguinte fórmula para cálculo da matriz inversa de uma matriz inversa M :

$$M^{-1} = \frac{1}{\det M} \text{adj}M.$$

Proposição A.9 . Todo subespaço vetorial de um espaço de dimensão finita é fechado.

Demonstração: Ver em [7]. ■

A.2 Alguns resultados e definições sobre Derivadas, Integrais e Equações Diferenciais

Seja $X \subset \mathbb{R}^n$ um conjunto aberto e não-vazio. A função $f : X \rightarrow \mathbb{R}^m$ é diferenciável em $x \in X$ se existe uma matriz real $m \times n$, a qual denotamos por $(Df)(x) \in \mathbb{R}^{m \times n}$, tal que

$$\lim_{z \rightarrow 0} \frac{\|f(x+z) - f(x) - ((Df)(x))z\|}{\|z\|} = 0 \text{ .}$$

A função f é diferenciável de $(Df)(x)$ existe para todo $x \in X$. Quando f é diferenciável em $x \in X$, a matriz $(Df)(x)$ coincide com a matriz Jacobiana de f em x , isto é, é a matriz cujas entradas são as derivadas parciais das componentes de f com relação às

componentes de x . Escrevendo $x = x_1, x_2, \dots, x_n$ e $f(x) = (f_1(x), f_2(x), \dots, f_m(x))$, temos em símbolos

$$(Df)(x) = \begin{bmatrix} \left(\frac{\partial f_1}{\partial x_1}\right)(x) & \left(\frac{\partial f_1}{\partial x_2}\right)(x) & \cdots & \left(\frac{\partial f_1}{\partial x_n}\right)(x) \\ \left(\frac{\partial f_2}{\partial x_1}\right)(x) & \left(\frac{\partial f_2}{\partial x_2}\right)(x) & \cdots & \left(\frac{\partial f_2}{\partial x_n}\right)(x) \\ \vdots & \vdots & \ddots & \vdots \\ \left(\frac{\partial f_m}{\partial x_1}\right)(x) & \left(\frac{\partial f_m}{\partial x_2}\right)(x) & \cdots & \left(\frac{\partial f_m}{\partial x_n}\right)(x) \end{bmatrix}.$$

Teorema A.13 (Regra da cadeia). Sejam $X \subset \mathbb{R}^n$ e $Y \subset \mathbb{R}^m$ conjuntos abertos e não-vazios. Se $g : X \rightarrow Y$ é diferenciável em $x \in X$ e $f : Y \rightarrow \mathbb{R}^p$ é diferenciável em $g(x) \in Y$, então $f \circ g$ é diferenciável em x com derivada

$$(D(f \circ g))(x) = (Df)(g(x))(Dg)(x)$$

Demonstração: Ver em [6]. ■

O seguinte caso especial da regra da cadeia é usado livremente ao longo do trabalho.

Proposição A.10. Sejam I um intervalo e $Y \subset \mathbb{R}^m$ um conjunto não-vazio e aberto. Assuma que as funções $g : I \rightarrow Y$ e $f : Y \rightarrow \mathbb{R}$ são diferenciáveis. Então a composição $f \circ g$ é diferenciável e

$$\frac{d}{dt}(f \circ g)(t) = (f \circ g)'(t) = \sum_{i=1}^m (\partial_i f)(g(t))g'_i(t) = \langle (\nabla f)(g(t)), g'(t) \rangle \quad \forall t \in I,$$

onde $(\partial_i f)(g(t)) = \left(\frac{\partial f_i}{\partial g_i}\right)(g(t))$ e g_i denota a i -ésima componente de g .

Teorema A.14. Seja $f : [a, b] \rightarrow \mathbb{R}$ contínua e tal que $f(x) \geq 0 \quad \forall x \in [a, b]$ e $\int_a^b f(x)dx = 0$. Então $f \equiv 0$.

Demonstração: Usaremos contradição para provar o resultado. Assuma que exista $c \in [a, b]$ tal que $f(c) > 0$. Como f é contínua em c , para todo $\varepsilon > 0$ existe $\delta > 0$ tal que para $x \in [a, b]$,

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

Escolha $k \in \mathbb{R}$ com $k > 0$ e $k > 1$. Assim $0 < \frac{1}{k} < 1$ e $0 < \frac{f(c)}{k} < f(c)$. Existe então pela continuidade de f em c , um $\delta > 0$ para o qual dado $x \in [a, b]$ tem-se

$$|x - c| < \delta \Rightarrow |f(x) - f(c)| < \frac{f(c)}{k}.$$

Desse modo, dado $x \in [a, b]$ com $|x - c| < \delta$, isto é, $x \in (c - \delta, c + \delta) \cap [a, b]$ temos

$$|f(x) - f(c)| < \frac{f(c)}{k}.$$

Observe que

$$|f(x) - f(c)| = |f(c) - f(x)| \geq f(c) - f(x).$$

Logo, para $x \in (c - \delta, c + \delta) \cap [a, b]$ obtemos

$$f(c) - f(x) < \frac{f(c)}{k} \Rightarrow -f(c) + f(x) > -\frac{f(c)}{k} \Rightarrow f(x) > f(c) - \frac{f(c)}{k}$$

e como $f(c) > \frac{f(c)}{k}$ podemos escrever

$$f(x) > f(c) - \frac{f(c)}{k} > 0.$$

Sendo assim, pela hipótese de que $f(x) \geq 0$ e por propriedades da integral temos

$$\begin{aligned} \int_a^b f(x) dx &= \underbrace{\int_a^{c-\delta} f(x) dx}_{\geq 0} + \int_{c-\delta}^{c+\delta} f(x) dx + \underbrace{\int_{c+\delta}^b f(x) dx}_{\geq 0} \Rightarrow \\ &\Rightarrow \int_a^b f(x) dx \geq \int_{c-\delta}^{c+\delta} f(x) dx \geq \int_{c-\delta}^{c+\delta} \left(f(c) - \frac{f(c)}{k} \right) dx \Rightarrow \\ &\Rightarrow \int_a^b f(x) dx \geq \left(f(c) - \frac{f(c)}{k} \right) \int_{c-\delta}^{c+\delta} dx = \left(f(c) - \frac{f(c)}{k} \right) [c + \delta - (c - \delta)] \Rightarrow \\ &\Rightarrow \int_a^b f(x) dx \geq \underbrace{\left(f(c) - \frac{f(c)}{k} \right)}_{>0} \underbrace{2\delta}_{>0} \Rightarrow \int_a^b f(x) dx > 0 \end{aligned}$$

contrariando a hipótese. Portanto devemos ter $f(x) = 0$ para todo $x \in [a, b]$, ou seja, $f \equiv 0$. ■

Apresentaremos agora algumas noções sobre derivadas e integrais de funções matriciais. Seja \mathbb{F} um corpo qualquer e denotemos por $\mathbb{M}_{\mathbb{F}}$ o espaço das matrizes $\mathbb{F}^{p \times q}$. Derivadas e integrais de funções $f : I \rightarrow \mathbb{M}_{\mathbb{F}}$ podem ser interpretadas de maneira natural em termos de suas entradas, isto é,

$$f'(t) = \frac{df}{dt}(t) = \begin{bmatrix} f'_{11}(t) & \cdots & f'_{1q}(t) \\ \vdots & \ddots & \vdots \\ f'_{p1}(t) & \cdots & f'_{pq}(t) \end{bmatrix} \quad t \in I,$$

e

$$\int_a^b f(t) dt = \begin{bmatrix} \int_a^b f_{11}(t) dt & \cdots & \int_a^b f_{1q}(t) dt \\ \vdots & \ddots & \vdots \\ \int_a^b f_{p1}(t) dt & \cdots & \int_a^b f_{pq}(t) dt \end{bmatrix}, \quad a, b \in I, \quad a < b,$$

onde as funções $f_{ij} : I \rightarrow \mathbb{F}$ com $1 \leq i \leq p$ e $1 \leq j \leq q$ denotam as entradas de f .

Proposição A.11 .Sejam $C \in \mathbb{R}^{m \times p}$ uma matriz constante e $f \in PC(I, \mathbb{R}^{p \times q})$. Então,

$$\int_a^b C f(t) dt = C \int_a^b f(t) dt$$

onde $a, b \in I$ com $a < b$.

Demonstração: Temos $C = (c_{ki})_{m \times p}$, $f(t) = (f_{ij}(t))_{p \times q}$ e $C f(t) = (\alpha_{kj}(t))_{m \times p}$ e pela definição de integral de funções matriciais podemos escrever

$$\int_a^b f(t) dt = \left(\int_a^b f_{ij}(t) dt \right)_{p \times q}, \quad \int_a^b C f(t) dt = \left(\int_a^b \alpha_{kj}(t) dt \right)$$

de modo que $C \int_a^b f(t) dt = (d_{kj}(t))_{m \times q}$. Pela definição do produto de matrizes temos

$$\alpha_{kj}(t) = c_{k1} f_{1j}(t) + c_{k2} f_{2j}(t) + \cdots + c_{kp} f_{pj}(t)$$

e

$$d_{kj}(t) = c_{k1} \int_a^b f_{1j}(t) dt + c_{k2} \int_a^b f_{2j}(t) dt + \cdots + c_{kp} \int_a^b f_{pj}(t) dt.$$

Integrando a primeira igualdade ambos os membros de a a b temos

$$\int_a^b \alpha_{kj}(t) dt = \int_a^b \left(c_{k1} f_{1j}(t) + c_{k2} f_{2j}(t) + \cdots + c_{kp} f_{pj}(t) \right) dt$$

e pelas propriedades da integral resulta

$$\int_a^b \alpha_{kj}(t) dt = c_{k1} \int_a^b f_{1j}(t) dt + c_{k2} \int_a^b f_{2j}(t) dt + \cdots + c_{kp} \int_a^b f_{pj}(t) dt$$

mostrando que

$$\int_a^b \alpha_{kj}(t) dt = (d_{kj}(t))$$

para todo $k \in \{1, 2, \dots, m\}$ e todo $j \in \{1, 2, \dots, p\}$, ou seja,

$$\int_a^b C f(t) dt = C \int_a^b f(t) dt.$$

■

Por praticidade, no resultados seguinte, utilizaremos a representação de vetores do \mathbb{R}^n na forma de matriz coluna e por isto orientamos o leitor para a não confusão entre produto interno e multiplicação de matrizes.

Proposição A.12 . Seja $f \in PC(I, \mathbb{R}^{m \times n})$ e $a, b \in I$ com $a < b$. Dados $z, w \in \mathbb{R}^n$ têm-se

$$\left\langle z, \int_a^b f(t) dt w \right\rangle = \int_a^b \langle z, f(t) w \rangle dt .$$

Demonstração: Fazemos

$$z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}, \quad w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \quad \text{e} \quad f(t) = \begin{bmatrix} f_{11}(t) & \cdots & f_{1n}(t) \\ \vdots & \ddots & \vdots \\ f_{m1}(t) & \cdots & f_{mn}(t) \end{bmatrix} .$$

Daí,

$$\int_a^b f(t) dt = \begin{bmatrix} \int_a^b f_{11}(t) dt & \cdots & \int_a^b f_{1n}(t) dt \\ \vdots & \ddots & \vdots \\ \int_a^b f_{m1}(t) dt & \cdots & \int_a^b f_{mn}(t) dt \end{bmatrix}$$

e

$$f(t) w = \begin{bmatrix} f_{11}(t) w_1 + f_{12}(t) w_2 + \cdots + f_{1n}(t) w_n \\ \vdots \\ f_{m1}(t) w_1 + f_{m2}(t) w_2 + \cdots + f_{mn}(t) w_n \end{bmatrix} .$$

Usando propriedades da integral temos

$$\begin{aligned} \int_a^b f(t) dt w &= \begin{bmatrix} \int_a^b f_{11}(t) dt w_1 + \int_a^b f_{12}(t) dt w_2 + \cdots + \int_a^b f_{1n}(t) dt w_n \\ \vdots \\ \int_a^b f_{m1}(t) dt w_1 + \int_a^b f_{m2}(t) dt w_2 + \cdots + \int_a^b f_{mn}(t) dt w_n \end{bmatrix} \\ &= \begin{bmatrix} \int_a^b (f_{11}(t) w_1 + f_{12}(t) w_2 + \cdots + f_{1n}(t) w_n) dt \\ \vdots \\ \int_a^b (f_{m1}(t) w_1 + f_{m2}(t) w_2 + \cdots + f_{mn}(t) w_n) dt \end{bmatrix} . \end{aligned}$$

Note agora que

$$\langle z, f(t) w \rangle = \begin{bmatrix} z_1 (f_{11}(t) w_1 + f_{12}(t) w_2 + \cdots + f_{1n}(t) w_n) \\ \vdots \\ z_n (f_{m1}(t) w_1 + f_{m2}(t) w_2 + \cdots + f_{mn}(t) w_n) \end{bmatrix} .$$

Desse modo, mais uma vez por propriedades da integral

$$\begin{aligned}
 \left\langle z, \int_a^b f(t) dt w \right\rangle &= \begin{bmatrix} z_1 \int_a^b (f_{11}(t) w_1 + f_{12}(t) w_2 + \cdots + f_{1n}(t) w_n) dt \\ \vdots \\ z_n \int_a^b (f_{n1}(t) w_1 + f_{n2}(t) w_2 + \cdots + f_{nn}(t) w_n) dt \end{bmatrix} \\
 &= \begin{bmatrix} \int_a^b z_1 (f_{11}(t) w_1 + f_{12}(t) w_2 + \cdots + f_{1n}(t) w_n) dt \\ \vdots \\ \int_a^b z_n (f_{n1}(t) w_1 + f_{n2}(t) w_2 + \cdots + f_{nn}(t) w_n) dt \end{bmatrix} \\
 &= \int_a^b \left(\begin{bmatrix} z_1 (f_{11}(t) w_1 + f_{12}(t) w_2 + \cdots + f_{1n}(t) w_n) \\ \vdots \\ z_n (f_{n1}(t) w_1 + f_{n2}(t) w_2 + \cdots + f_{nn}(t) w_n) \end{bmatrix} \right) dt. \\
 &= \int_a^b \langle z, f(t) w \rangle dt
 \end{aligned}$$

e portanto

$$\left\langle z, \int_a^b f(t) dt w \right\rangle = \int_a^b \langle z, f(t) w \rangle dt .$$

■

Teorema A.15 . O problema de valor inicial

$$\begin{cases} x'(t) = Ax(t) + b(t) \\ x(t_0) = x_0 \in \mathbb{R}^n \end{cases}$$

onde $A \in \mathbb{R}^{n \times n}$ é uma matriz constante, $b : J \rightarrow \mathbb{R}^n$ é contínua por partes e $t_0 \in J$; $J \subset \mathbb{R}$ um intervalo; possui solução única dada pela função

$$\begin{aligned}
 x : J &\rightarrow \mathbb{R}^n \\
 t &\mapsto e^{A(t-t_0)} x_0 + \int_0^t e^{A(t-s)} b(s) ds
 \end{aligned}$$

para todo $t \in J$.

Demonstração: Ver em [1] ou ainda em [3].

■

Teorema A.16 .Sejam $A \in \mathbb{C}^{n \times n}$ e $\mu_A := \max\{\text{Re}(\lambda) : \lambda \in \sigma(A)\}$.

(1) $\mu_A < 0$ se, e somente se, $\|e^{At}\|$ decai exponencialmente quando $t \rightarrow \infty$.

(2) $\mu_A < 0$ se, e somente se, $\lim_{t \rightarrow \infty} e^{At}x_0 = 0$ para todo $x_0 \in \mathbb{C}^n$.

Demonstração: Ver em [1].



Referências Bibliográficas

- [1] Logemann, H. .*Ordinary Differential Equations: analysis, qualitative theory and control.*/Springer undergraduate mathematics series. Londres: Springer, 2014.
- [2] Lourêdo, Aldo Trajano. *Um primeiro curso de álgebra linear.* Campina Grande: EDUEPB, 2015.
- [3] Figueiredo, Djairo Guedes de. *Equações Diferenciais Aplicadas*, 3^a edição./Coleção matemática universitária. Rio de Janeiro: IMPA, 2015.
- [4] Thomas, George B.; Weir, Maurice D.; Hass, Joel. *Cálculo, volume 2*; tradução Carlos Scalici; revisão técnica Claudio Hirofume Asano. 12^a.ed. São paulo: Pearson Education do Brasil, 2012.
- [5] Iezzi, Gelson. *Fundamentos de matemática elementar-4.* 7^a. ed. São Paulo: Atual, 2004.
- [6] Lima, E.L. *Análise no espaço \mathbb{R}^n .* Rio de Janeiro: IMPA, 2004.
- [7] Machado, Luciana Bertholdi. *Análise Funcional e Aplicações.* 2012. 204f. Dissertação (mestrado) - Instituto de Geociências e Ciências Exatas, Universidade Estadual Paulista, Rio claro.
- [8] Lima, E.L. *Curso de Análise; V.1.* 14^a.ed./Projeto Euclides. Rio de Janeiro: Associação IMPA, 2016. 432p.
- [9] Kelley, G.W. and Peterson, C. A., *The Theory of Differential Equations: Classical and Qualitative* Second Edition, Springer, New York, 2010.
- [10] Sontag, E.D. *Mathematical Control Theory: Deterministic Finite Dimensional Systems.* 2nd ed., Springer Verlag, New York, 1998.

- [11] Polderman, J.W. and Willems, J.C. *Introduction to Mathematical Systems Theory*, Springer-Verlag, New York, 1998.
- [12] Hirsch, M. and Smale, S. *Differential Equations, Dynamical Systems and Linear Algebra*, Academic Press, San Diego, 1974.
- [13] Zabczyk, J. *Mathematical Control Theory An Introduction*, Birkhäuser Boston, 1995.
- [14] Hasegawa, Y. *Control Problems of Discrete-Time Dynamical Systems*, Springer New York 2015.
- [15] Coron, J.M. *Control and Nonlinearity*, American Mathematical Society 2000.
- [16] Micus, S. and Zuazua, E. *An Introduction to the Controllability of Partial Differential Equations*, 2002.