



UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
Instituto de Geociências e Ciências Exatas
Campus de Rio Claro

Modelos Matemáticos e Aplicações ao Ensino Médio

Luciano Aparecido Magrini

Dissertação apresentada ao Programa de Pós-Graduação – Mestrado Profissional em Matemática em Rede Nacional como requisito parcial para a obtenção do grau de Mestre

Orientadora
Profa. Dra. Marta Cilene Gadotti

2013

517.38 Magrini, Luciano Aparecido
M212m Modelos Matemáticos
e Aplicações ao Ensino Médio/ Luciano Aparecido Magrini- Rio
Claro: [s.n.], 2013.
119 f.: fig., tab.

Dissertação (mestrado) - Universidade Estadual Paulista, Insti-
tuto de Geociências e Ciências Exatas.

Orientadora: Marta Cilene Gadotti

1. Equações Diferenciais. 2. Equações Diferenciais Ordinárias.
3. Modelagem Matemática. 4. Estabilidade. I. Título

TERMO DE APROVAÇÃO

Luciano Aparecido Magrini
MODELOS MATEMÁTICOS
E APLICAÇÕES AO ENSINO MÉDIO

Dissertação APROVADA como requisito parcial para a obtenção do grau de Mestre no Curso de Pós-Graduação Mestrado Profissional em Matemática em Rede Nacional do Instituto de Geociências e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, pela seguinte banca examinadora:

Profa. Dra. Marta Cilene Gadotti
Orientadora

Profa. Dra. Suzinei Aparecida Siqueira Marconato
IGCE - UNESP/Rio Claro

Prof. Dr. Luiz Augusto da Costa Ladeira
ICMC - USP/São Carlos

Rio Claro, 15 de agosto de 2013

*Dedico a meus pais,
Terezinha e Alcides.*

Agradecimentos

Agradeço a Deus por ter permitido que eu concluísse mais esta etapa de minha vida e assim realizasse um sonho antigo.

Sou grato à minha orientadora, Profa. Dra. Marta Cilene Gadotti, pela acolhida e orientação, pela paciência, carinho e dedicação com que me orientou e tornou este trabalho possível; também pelas palavras de ânimo em cada e-mail trocado ou conversa e por suas valiosas sugestões diante das minhas inúmeras desatenções ao elaborar este texto.

À Profa. Dra. Suzinei Aparecida Siqueira Marconato por acreditar no PROFMAT, em mim e em cada um de meus colegas de turma. Pela presença constante ao nosso lado, evitando que desanimássemos a cada novo tropeço ou dificuldade; também pelo carinho, dedicação e respeito com nossas limitações matemáticas.

À Profa. Dra. Alice Kimie Miwa Libardi, por sua competência e conhecimento; pela oportunidade de ter aprendido muito com suas aulas durante o Verão de 2012.

Às minhas queridas Ana Cecília, Gláucia, Mariana, Sibeli e Patrícia, que dividiram suas vidas comigo ao longo dos últimos dois anos, oferecendo amizade, ânimo, horas de estudo juntos e muitas vezes suas casas para que eu estudasse aos fins de semana. Juntos estudamos, rimos (muito!) e dividimos muitas experiências e angústias ao longo desta jornada.

Ao meu querido amigo Ariosvaldo Trindade pelas caronas de São Paulo a Rio Claro toda semana; por sua amizade, carinho e longas conversas pelo caminho.

Aos funcionários da Biblioteca, Seção de Pós Graduação e à SBM pela iniciativa.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior(CAPES) pelo apoio financeiro.

*Se a "verdade" fosse aquilo que posso entender,
terminaria sendo apenas uma verdade pequena, do meu caminho.
(Clarice Lispector)*

Resumo

Neste trabalho apresentamos os fundamentos básicos da teoria das equações diferenciais ordinárias, instrumentos indispensáveis na modelagem matemática de problemas nas mais variadas áreas do conhecimento humano. Também discutimos alguns modelos clássicos na literatura e apresentamos critérios que asseguram a estabilidade das soluções de uma equação diferencial, assunto de grande interesse para todos que estudam este tema. Além disso, apresentamos duas propostas didáticas para os professores de matemática da educação básica que desejam incluir a modelagem em suas aulas como elemento motivador para um estudo significativo dos temas discutidos nos currículos do Ensino Médio.

Palavras-chave: Equações Diferenciais, Equações Diferenciais Ordinárias, Modelagem Matemática, Estabilidade.

Abstract

We present the fundamentals of the theory of ordinary differential equations, indispensable tools in mathematical modeling of problems in various areas of human knowledge. We also discuss some classic models in mathematical literature and present criteria to ensure the stability of the solutions of an equation differentiates subject of great interest to all who study the subject. Furthermore, we present two proposals for the teaching of mathematics teachers of basic education who want to include modeling in their classes as motivator for motivating and meaningful study of the topics discussed in the curricula of secondary education.

Keywords: Differential Equations, Ordinary Differential Equations, Mathematical Modeling, Stability.

Lista de Figuras

2.1	Curvas integrais e campo de direções de (2.6).	23
2.2	Gráfico das soluções do modelo de Malthus.	38
2.3	Gráfico das soluções do modelo de Verhulst	41
2.4	Retrato de fase de (2.63).	47
2.5	Retrato de fase de (2.64).	47
2.6	Retrato de fase de (2.65).	48
3.1	Nó Estável: $\lambda_1 < \lambda_2 < 0$.	63
3.2	Nó Instável: $\lambda_1 > \lambda_2 > 0$.	63
3.3	Ponto de Sela: $\lambda_1 < 0 < \lambda_2$.	64
3.4	Nó Estável: $\lambda_1 = \lambda_2 = \lambda < 0$.	64
3.5	Nó Estável: $\lambda_1 = \lambda_2 = \lambda > 0$.	65
3.6	Autovalores repetidos: multiplicidade geométrica menor que a algébrica.	65
3.7	Centro: $a = 0$.	66
3.8	Foco Instável: $a > 0$.	66
3.9	Foco Estável: $a < 0$.	66
4.1	Solução $\gamma(t, t_0, x_0)$ estável	68
4.2	Curvas integrais de (4.1).	68
4.3	Solução estável, assintoticamente estável e instável	70
5.1	Tela Inicial do Simulador do Sistema Massa-Mola	90
5.2	Tela Inicial do Simulador do Sistema Pêndulo Simples	90

Lista de Tabelas

2.1	Taxas de natalidade e mortalidade de São Paulo entre 2000 e 2005. . .	39
2.2	Erro no Modelo de Malthus - População do Município de São Paulo entre 2000-2010.	39
4.1	Estabilidade dos sistemas lineares $X' = AX$ de Ordem 2	72
5.1	População oficial do município de São Paulo entre 2000 e 2005.	94
5.2	Erro percentual do modelo malthusiano para o município de São Paulo.	95

Sumário

1	Introdução	19
2	Equações diferenciais ordinárias	21
2.1	Preliminares	24
2.2	Existência e unicidade de soluções	28
2.3	Algumas técnicas de resolução	31
2.3.1	Variáveis separáveis	31
2.3.2	Equações diferenciais lineares escalares de primeira ordem	32
2.3.3	Equações exatas	34
2.4	Modelos descritos por equações diferenciais	37
2.4.1	Modelo populacional de Malthus	37
2.4.2	O modelo de Verhulst	40
2.4.3	Propagação da podridão em maçãs	42
2.4.4	Disseminação de doenças contagiosas	44
2.5	Retratos de fase	46
3	Sistemas de equações diferenciais lineares	49
3.1	Soluções e propriedades	52
3.1.1	Autovalores reais e distintos	55
3.1.2	Autovalores complexos	56
3.1.3	Autovalores repetidos	59
3.2	Sistemas planares: retratos de fase	62
3.2.1	Autovalores reais e distintos	63
3.2.2	Autovalores repetidos	64
3.2.3	Autovalores complexos e conjugados	65
4	Introdução à estabilidade	67
4.1	Definições e exemplos	67
4.2	Estabilidade de equações lineares	70
4.3	Estabilidade de equações não lineares	76
4.4	Estabilidade segundo Lyapunov	82

5	Propostas didáticas	89
5.1	Pêndulo simples e massa-mola	91
5.2	Evolução populacional do município de São Paulo	93
A	Resultados de Álgebra Linear	97
A.1	Diagonalização e forma de Jordan	100
A.2	Exponencial de matrizes	109
	Referências	119

1 Introdução

Isaac Newton (1643-1727) e Gottfried Wilhelm von Leibniz (1646-1716) foram responsáveis pela invenção do Cálculo no século XVII ao descobrirem de maneira independente a conexão entre o problema da determinação da reta tangente ao gráfico de uma função f e o problema da determinação da área sob o gráfico de uma função *derivada* de f , estabelecendo deste modo o Teorema Fundamental do Cálculo.

Neste mesmo período, em grande parte devido aos trabalhos de Newton, os princípios da Mecânica Clássica foram descobertos e encontraram no Cálculo as ferramentas matemáticas apropriadas para seu correto tratamento. No século XVIII, a abordagem de problemas físicos e geométricos com os métodos do Cálculo se mostrou bastante profícua, surgindo assim as equações diferenciais.

Newton aplicou séries infinitas e assim resolveu algumas classes de equações diferenciais do tipo $y' = F(x, y)$ onde $F(x, y)$ é um polinômio nas variáveis x e y . Leibniz, por volta de 1691 apresentou a técnica das variáveis separáveis e na sequência obteve a solução geral das equações da forma

$$\frac{dy}{dx} + p(x)y = q(x),$$

onde p e q são funções de x .

Os métodos do Cálculo e as equações diferenciais começaram a ganhar força ao se mostrarem decisivos na resolução de problemas. Um destes problemas, proposto por Galileu Galilei (1564-1642), consistia em descrever matematicamente a forma da curva formada por um fio suspenso entre dois pontos e sujeito apenas à ação da gravidade e de seu próprio peso (erroneamente alguns pensavam ser a parábola a solução do problema). Em 1690, Johann Bernoulli (1667-1748) resolveu o problema mostrando que a curva procurada (conhecida como catenária) satisfaz à equação diferencial

$$\frac{d^2y}{dx^2} = \frac{pg}{h} \left(\sqrt{1 + \left(\frac{dy}{dx} \right)^2} \right),$$

onde p indica o peso do fio, g a força da gravidade e h a altura do ponto P mais baixo da curva, isto é, do ponto P pertencente à curva que está mais próximo do solo.

Por sua vez, Leonhard Euler (1707-1783) em um artigo de 1734 apresentou a teoria dos fatores integrantes e explicitou condições necessárias e suficientes para que uma equação diferencial fosse exata. Além disso, ele resolveu corretamente as equações da forma

$$a_2 \frac{d^2 y}{dx^2} + a_1 \frac{dy}{dx} + a_0 y(x) = f(x),$$

onde a_2, a_1 e a_0 são constantes. Coube a Euler também desenvolver um dos métodos numéricos empregados para a resolução de equações diferenciais do tipo

$$\frac{dy}{dx} = F(x, y),$$

sujeita à condição $y(x_0) = y_0$.

Até este ponto de desenvolvimento da teoria a preocupação era a de resolver explicitamente algumas classes de equações, de modo que as técnicas desenvolvidas não eram gerais; havia uma profusão de métodos particulares que não estavam devidamente fundamentados. Com o avanço da Análise, por volta das últimas décadas de 1800 houve a preocupação em se desenvolver uma teoria que respondesse às questões de existência e unicidade de soluções.

Em geral não é um processo simples determinar as soluções de uma equação diferencial e até hoje não existem métodos gerais. Na verdade a maior parte das equações diferenciais não pode ser resolvida analiticamente, isto é, não é possível determinar explicitamente as soluções. Jules Henri Poincaré (1854-1912) no início do século XX com uma sequência de artigos propõe o tratamento qualitativo das equações diferenciais, onde os esforços são dedicados a entender o comportamento da família de soluções, o que basta na maioria das aplicações.

Este breve relato histórico é baseado na referência [1].

Quanto à organização deste trabalho, registramos que no segundo capítulo estão apresentados os fundamentos da teoria das equações diferenciais, onde incluímos (dentre outros resultados) a demonstração da existência e unicidade de soluções sob hipóteses convenientes e alguns modelos matemáticos clássicos na literatura. No terceiro capítulo apresentamos a teoria dos sistemas lineares, caracterizando as soluções em termos de exponenciais de matrizes e apresentando uma descrição dos retratos de fase dos sistemas em dimensão dois.

O quarto capítulo faz uma introdução à teoria da estabilidade dos pontos singulares para equações lineares e não lineares e à teoria das Funções de Lyapunov. No quinto capítulo são apresentadas duas propostas didáticas com modelos matemáticos a serem utilizadas em sala de aula com alunos do Ensino Médio. Finalmente, no apêndice incluímos os resultados de Álgebra Linear utilizados.

2 Equações diferenciais ordinárias

De maneira pouco precisa, uma equação diferencial ordinária pode ser definida como uma equação que envolve uma "função incógnita" (de apenas uma variável) e suas derivadas.

A definição a seguir explicita a noção de equação diferencial ordinária que empregaremos neste texto.

Definição 2.1. *Uma equação diferencial ordinária é uma relação do tipo*

$$x^{(d)} = F(t, x, x^{(1)}, x^{(2)}, \dots, x^{(d-1)}), \quad (2.1)$$

onde $F : U \rightarrow \mathbb{R}$ é uma função contínua definida em um aberto U de \mathbb{R}^{1+d} com $t \in \mathbb{R}$ e onde $x^{(j)}$ é a j -ésima derivada da função x de uma variável real, com $j \in \{1, 2, \dots, d\}$.

O inteiro d que aparece na definição acima é chamado de **ordem** da equação diferencial ordinária.

No restante deste trabalho escreveremos simplesmente equação diferencial, devendo ficar subentendido que estamos nos referindo à equação diferencial ordinária. O exemplo a seguir ilustra a modelagem de um fenômeno físico.

Exemplo 2.1. (Lei de Hooke) Seja m a massa de uma partícula acoplada à uma mola de comprimento l em equilíbrio. Considere que a mola seja esticada (ou comprimida) de modo a sofrer um afastamento (ou deformação) x em relação ao equilíbrio, não estando sujeito a qualquer força externa. A Lei de Hooke, ao afirmar que a intensidade da força elástica F_t é proporcional à deformação x , permite-nos escrever

$$F_t = -Cx, \quad (2.2)$$

onde a constante C é uma constante física, dependendo por exemplo do material da mola, da estrutura dos anéis, etc. Pela Segunda Lei de Newton temos

$$F_t = m \frac{d^2x}{dt^2},$$

segue de (2.2) que

$$\frac{d^2x}{dt^2} = -kx, \quad (2.3)$$

desde que façamos $k = C/m$. A equação (2.3) modela o sistema massa-mola nas condições admitidas.

Note que em (2.3) a função $F(t, x) = -kx$ não depende explicitamente da variável t ; dizemos neste caso que a equação diferencial é uma **equação diferencial autônoma** e escrevemos por simplicidade apenas $F(x)$. Neste caso, interpretamos a aplicação $F : U \rightarrow \mathbb{R}$ como um **campo de vetores** definido no aberto $U \subset \mathbb{R}^d$ (e não em \mathbb{R}^{1+d}).

Observe que a função $x(t)$ que satisfaz a equação (2.3) deve ser tal que sua derivada segunda seja proporcional à ela mesma. Ora, as funções trigonométricas seno e cosseno possuem esta propriedade. Ajustando os argumentos de maneira conveniente, somos tentados a considerar as funções reais

$$x_1(t) = \alpha \cos(t\sqrt{k}) \quad (2.4)$$

e

$$x_1(t) = \alpha \sin(t\sqrt{k}) \quad (2.5)$$

como soluções para (2.3). Podemos verificar que de fato estas funções são soluções, pois de (2.4) podemos escrever

$$x_1'(t) = -\alpha\sqrt{k} \sin(t\sqrt{k})$$

e

$$x_1''(t) = -k\alpha \cos(t\sqrt{k})$$

e então é fácil ver que $x_1''(t) = -kx_1(t)$. Um raciocínio análogo permite mostrar que (2.5) também é solução de (2.3). Além disso, pelo mesmo argumento podemos ver que qualquer combinação linear de (2.4) e (2.5) também é solução de (2.3), o que em geral é verdadeiro para qualquer **equação diferencial linear**¹.

A seguir apresentamos a definição de solução da equação (2.1). Observe que os itens 1 e 2 da definição dada são exigências para que o item 3 faça sentido.

Definição 2.2. *Uma solução de (2.1) é uma curva $\gamma : I \rightarrow \mathbb{R}^d$ de classe C^d tal que:*

1. *O intervalo I é um aberto de \mathbb{R} ;*
2. *O vetor $\left(t, \gamma(t), \frac{d\gamma}{dt}(t), \dots, \frac{d^{d-1}\gamma}{dt^{d-1}}(t)\right)$ pertence a U para todo $t \in I$;*
3. *Para todo $t \in I$, tem-se $F\left(t, \gamma(t), \frac{d\gamma}{dt}(t), \dots, \frac{d^{d-1}\gamma}{dt^{d-1}}(t)\right) = \frac{d^d\gamma}{dt^d}(t)$.*

Os próximos exemplos ilustram alguns conceitos pertinentes à noção de solução de uma equação diferencial.

¹A definição de equação diferencial linear será apresentada no Capítulo 03.

Exemplo 2.2. Considere a equação diferencial

$$\frac{dx}{dt} = t, \quad (2.6)$$

cujas soluções são dadas explicitamente por

$$x(t) = \frac{t^2}{2} + k, \quad (2.7)$$

onde k é uma constante.

Variando a constante k determinamos infinitas curvas (chamadas de **curvas integrais** da equação) e portanto infinitas soluções da equação (2.6). A equação (2.7) é conhecida como **solução geral** de (2.6) e ao fixar a constante k obtemos uma **solução particular**.

Do ponto de vista geométrico, as soluções da equação $y' = f(x, y)$ são tais que em cada ponto (x, y) a reta tangente à curva integral passando pelo ponto tem coeficiente angular $m = f(x, y)$. Isto sugere um método geométrico para entender o comportamento das curvas integrais da equação: para isto, traçamos um pequeno segmento de reta em cada ponto (x, y) com coeficiente angular $f(x, y)$; ao conjunto destes segmentos é dado o nome de **campo de direções** da equação $y' = f(x, y)$. Com o uso de softwares gratuitos podemos desenhar o campo de direções e assim analisar o comportamento das soluções.

Abaixo encontram-se representadas algumas curvas integrais de (2.6). Compare a semelhança de comportamento entre as curvas integrais e o campo de direções plotado abaixo: ²

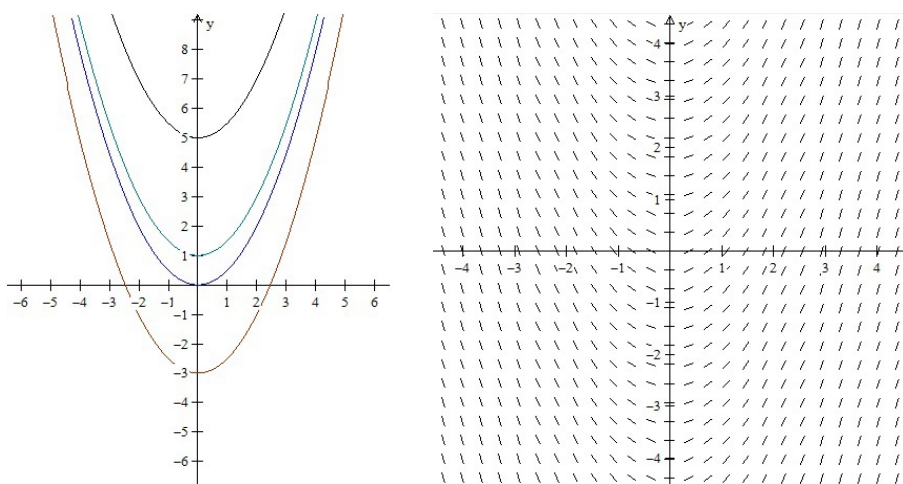


Figura 2.1: Curvas integrais e campo de direções de (2.6).

²O campo de vetores foi construído no software **WinPlot**, obtido gratuitamente na Internet.

2.1 Preliminares

Nesta seção vamos apresentar os fundamentos da teoria básica das equações diferenciais ordinárias e a desenvolveremos privilegiando as equações diferenciais de primeira ordem. Esta escolha não é arbitrária, pois uma equação diferencial de ordem qualquer pode convenientemente ser escrita como um sistema de equações de primeira ordem através de uma mudança de variáveis conveniente.

De fato, considere a equação (2.1). Para transformá-la em um sistema de equações de primeira ordem, basta definir uma mudança de variáveis do seguinte modo:

$$\begin{aligned} y_1 &= x, \\ y_2 &= \frac{dx}{dt}, \\ &\vdots \\ y_d &= \frac{d^{d-1}x}{dt^{d-1}}, \end{aligned}$$

de modo que

$$\begin{aligned} y_1' &= y_2, \\ y_2' &= y_3, \\ &\vdots \\ y_{d-1}' &= y_d, \end{aligned}$$

e a equação (2.1) se escreve como o sistema

$$\left\{ \begin{array}{l} y_1' = y_2 \\ y_2' = y_3 \\ \vdots \\ y_d' = F(t, y_1, \dots, y_{d-1}) \end{array} \right. \quad (2.8)$$

Na sequência apresentaremos e provaremos alguns resultados que usaremos ao longo de todo este trabalho e especialmente na seção seguinte, quando demonstraremos o Teorema de Existência e Unicidade de Soluções.

Definição 2.3. Uma função $f : U \subset \mathbb{R}^{1+d} \rightarrow \mathbb{R}^d$ é **lipschitziana** na segunda variável se existir $C \geq 0$ tal que

$$\|f(t, x_1) - f(t, x_2)\| \leq C\|x_1 - x_2\|,$$

para quaisquer dois pontos (t, x_1) e (t, x_2) de U cujas primeiras coordenadas são iguais. Neste caso, dizemos que C é uma constante de Lipschitz e que $f(t, x)$ satisfaz uma condição de Lipschitz.

Definição 2.4. Uma **métrica** num conjunto M é uma função $d : M \times M \rightarrow \mathbb{R}$ que associa a cada par ordenado de elementos $x, y \in M$ um número real $d(x, y)$ chamado de *distância de x a y* , de modo que são satisfeitas as seguintes propriedades para quaisquer $x, y, z \in M$:

1. $d(x, x) = 0$;
2. Se $x \neq y$ então $d(x, y) > 0$;
3. $d(x, y) = d(y, x)$;
4. $d(x, z) \leq d(x, y) + d(y, z)$

Exemplo 2.3. A função

$$d(x_1, x_2) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_d - b_d)^2},$$

onde $x_1 = (a_1, a_2, \dots, a_d)$ e $x_2 = (b_1, b_2, \dots, b_d)$ são elementos de \mathbb{R}^d define uma métrica em \mathbb{R}^d , chamada de **métrica euclidiana**.

Um **espaço métrico** é um par (M, d) onde M é um conjunto e d é uma métrica em M . Se (M, d) é um espaço métrico, então todo subconjunto $S \subset M$ pode ser considerado como um espaço métrico desde que seja considerada a restrição de d a $S \times S$. Quando se faz isto, a métrica de S é dita **induzida** pela de M . Os elementos de M ou S podem ser de natureza bastante distinta: podem ser números, funções, matrizes, conjuntos e são chamados simplesmente de **pontos**.

Uma sequência $(x_n)_{n \in \mathbb{N}}$ de pontos em um espaço métrico M é dita convergente se x se para toda bola aberta³ B tal que $x \in B$ existe um número finito de índices j tais que x_j não pertence a B . Uma sequência $(x_n)_{n \in \mathbb{N}}$ chama-se **sequência de Cauchy** se para todo $\epsilon > 0$ dado, existir $n_0 \in \mathbb{N}$ tal que $m, n > n_0 \Rightarrow d(x_m, x_n) < \epsilon$. Toda sequência de Cauchy é limitada. Um espaço métrico M é chamado de **completo** quanto toda sequência de Cauchy em M é convergente.

Exemplo 2.4. O conjunto \mathbb{R}^d é um espaço completo para qualquer $d \geq 1$. A demonstração pode ser encontrada em [2].

Dados quaisquer subconjuntos $E \subseteq \mathbb{R}^k$ e $F \subseteq \mathbb{R}^m$ dotados das respectivas métricas induzidas, denotamos por $B_0(E, F)$ o conjunto de todas as aplicações $f : E \rightarrow F$ que são limitadas, isto é, tal que a imagem $f(E) \subseteq F$ é um conjunto limitado de \mathbb{R}^m . Em $B_0(E, F)$ definimos a **métrica do supremo** ou **métrica uniforme** por

$$d(f, g) = \sup_{x \in E} \|f(x) - g(x)\|,$$

³Por definição, se o par (X, d) é um espaço métrico, então a **bola aberta** centrada em $x \in X$ e raio $r > 0$ é o conjunto $B(x, r) = \{y \in X | d(x, y) < r\}$.

para $f, g \in B_0(E, F)$. É possível mostrar que $B_0(E, F)$ com a métrica do supremo é um espaço métrico; mais ainda: trata-se de um espaço métrico completo, conforme a proposição abaixo.

Proposição 2.1. *Se $F \subseteq \mathbb{R}^m$ é um espaço métrico completo, então para qualquer $E \subseteq \mathbb{R}^k$ o espaço métrico $B_0(E, F)$ é completo com a métrica uniforme.*

Demonstração. Seja $(f_n)_{n \in \mathbb{N}}$ uma sequência de Cauchy (e portanto limitada) em $B_0(E, F)$. Fixando $x \in E$, a sequência (x_n) em que $x_n = f_n(x) \in F$ também é de Cauchy pois

$$\|f_n(x) - f_m(x)\| \leq d(f_n, f_m).$$

Como por hipótese F é completo, existe em F o limite $y = \lim f_n(x)$, que define uma função $y = f(x)$ de E em F . Vamos mostrar que f é limitada, ou seja, que f é um elemento de $B_0(E, F)$ e que a sequência f_n converge para f no espaço $B_0(E, F)$.

De fato, sendo (f_n) uma sequência limitada, podemos tomar $g_0 : E \rightarrow F$ constante qualquer e $r > 0$ tais que $d(f_n, g_0) \leq r$ para cada $n \in \mathbb{N}$. Dado $x \in E$ segue que

$$\|f(x) - g_0(x)\| = \lim \|f_n(x) - g_0(x)\| \leq \lim d(f_n, g_0) \leq r$$

e portanto, $d(f, g_0) \leq r$ e $f \in B_0(E, F)$. Para terminar, dado $\epsilon > 0$ tomemos N tal que $d(f_k, f_n) < \frac{\epsilon}{2}$ para quaisquer $k, n \geq N$ com $k, n, N \in \mathbb{N}$. Fixando agora $x \in E$, decorre que

$$\|f(x) - f_n(x)\| = \lim_{k \rightarrow +\infty} \|f_k(x) - f_n(x)\| \leq \frac{\epsilon}{2} < \epsilon.$$

Deste modo, $d(f, f_n) < \epsilon$ para cada $n \geq N$, ou seja, $\lim d(f_n, f) = 0$ quando $n \rightarrow \infty$. \square

Exemplo 2.5. Seja \mathcal{X} o conjunto das curvas $\gamma : I \rightarrow \mathbb{R}^d$ contínuas e limitadas com a métrica do supremo. Então (\mathcal{X}, d) é um espaço métrico completo conforme a Proposição 2.1 pois \mathbb{R}^d é completo.

Seja $f : M \rightarrow M$ uma aplicação definida num espaço métrico M qualquer. Um **ponto fixo** da aplicação f é um ponto x^* tal que $f(x^*) = x^*$.

Definição 2.5. *Uma aplicação $g : M \rightarrow M$ chama-se **contração** quando existe uma constante real $0 \leq c < 1$ tal que*

$$d(f(x), f(y)) \leq cd(x, y),$$

para quaisquer $x, y \in M$.

A seguir são enunciados resultados importantes da Análise Funcional que serão utilizados na prova de existência de soluções de uma EDO. As provas omitidas podem ser encontradas em [2].

Lema 2.1. *Sejam M um espaço métrico e $f : M \rightarrow M$ uma contração. Então f é (uniformemente) contínua.*

Teorema 2.1. *(do ponto fixo de Banach) Se M é um espaço métrico completo, toda contração $f : M \rightarrow M$ possui um único ponto fixo em M . Precisamente, escolhendo um ponto $x_0 \in M$ qualquer e colocando $x_1 = f(x_0)$, $x_2 = f(x_1)$, ..., $x_{n+1} = f(x_n)$, ... a sequência (x_n) converge em M e $x^* = \lim x_n$ é o único ponto fixo de f .*

Demonstração. Admita que a sequência (x_n) convirja para um ponto $x^* \in M$. Como f é uniformemente contínua (pelo Lema 2.1) devemos ter

$$f(x^*) = f(\lim(x_n)) = \lim f(x_n) = \lim f(x_n) = \lim x_{n+1} = x^*,$$

ou seja, $x^* \in M$ é um ponto fixo de f .

Para a prova da unicidade, suponha que f admita dois pontos fixos distintos, ou seja, suponha a existência de $x^*, x^{**} \in M$ tais que $f(x^*) = x^*$ e $f(x^{**}) = x^{**}$; então

$$d(x^*, x^{**}) = d(f(x^*), f(x^{**})) \leq cd(x^*, x^{**})$$

de onde $(1 - c)d(x^*, x^{**}) \leq 0$.

Como $1 - c > 0$ (de onde $0 < c < 1$) segue que $cd(x^*, x^{**}) = 0$ e portanto $x^* = x^{**}$.

Quanto à existência, note que (x_n) é uma sequência de Cauchy (e portanto convergente em M , pois M é completo); de fato, observe que

$$d(x_1, x_2) = d(f(x_0), f(x_1)) \leq cd(x_0, x_1),$$

$$d(x_2, x_3) = d(f(x_1), f(x_2)) \leq cd(x_1, x_2) \leq c^2d(x_0, x_1),$$

e segue indutivamente que para qualquer $n \in \mathbb{N}$,

$$d(x_n, x_{n+1}) \leq c^n d(x_0, x_1).$$

Portanto, para quaisquer $n, p \in \mathbb{N}$, tem-se

$$\begin{aligned} d(x_n, x_{n+p}) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) \dots + d(x_{n+p-1}, x_{n+p}) \\ &\leq [c^n + c^{n+1} + \dots + c^{n+p-1}]d(x_0, x_1), \end{aligned}$$

ou seja,

$$d(x_n, x_{n+p}) \leq \frac{c^n}{1 - c} d(x_0, x_1).$$

Como $\lim c^n = 0$ (pois $0 \leq c < 1$), a sequência (x_n) é de Cauchy em M e a demonstração está concluída. □

2.2 Existência e unicidade de soluções

Um problema de valor inicial (abreviadamente PVI) ou **problema de Cauchy** de primeira ordem consiste de uma equação diferencial de primeira ordem sujeita à uma condição inicial.

Deste modo, trata-se de encontrar uma solução particular $x(t)$ tal que em t_0 seu valor seja x_0 :

$$\begin{cases} \frac{dx}{dt} = F(t, x), \\ x(t_0) = x_0. \end{cases} \quad (2.9)$$

Demonstraremos a seguir um teorema que garante a existência e unicidade de solução para (2.9) mediante hipóteses convenientes, segundo a referência [3].

Teorema 2.2. (*Existência e Unicidade de Soluções*) *Seja $F : U \rightarrow \mathbb{R}^d$ onde U é um aberto de \mathbb{R}^{1+d} contínua e localmente lipschitziana na segunda variável. Então:*

1. *Para todo $(t_0, x_0) \in U$ existe $\gamma : I \subset \mathbb{R} \rightarrow \mathbb{R}^d$, solução de $x' = F(t, x)$ com $\gamma(t_0) = x_0$;*
2. *Se $\gamma_1 : I_1 \rightarrow \mathbb{R}^d$ e $\gamma_2 : I_2 \rightarrow \mathbb{R}^d$ são soluções de $x' = F(t, x)$ e existe $t_0 \in I_1 \cap I_2$ tal que $\gamma_1(t_0) = \gamma_2(t_0)$, então $\gamma_1(t) = \gamma_2(t)$ para todo $t \in I_1 \cap I_2$.*

Demonstração. Sejam $\delta > 0$ tal que⁴ $\overline{Q} = \overline{B_\delta(t_0)} \times \overline{B_\delta(x_0)} \subset U$, C a constante de lipschitz e $M_\delta = \sup_{(t,y) \in \overline{Q}} \|F\|$ em \overline{Q} . Seja ainda $\epsilon > 0$ tal que

$$\epsilon \leq \delta, \epsilon < \frac{1}{C}, \epsilon < \frac{\delta}{M_\delta}. \quad (2.10)$$

Definamos \mathcal{X} como o conjunto de todas as curvas contínuas do intervalo $(t_0 - \epsilon, t_0 + \epsilon)$ em \mathbb{R}^d cujas imagens pertencem à $\overline{B_\delta(x_0)}$ munido da métrica d do supremo, isto é,

$$\mathcal{X} = \{\gamma(t) \in \overline{B_\delta(x_0)} \mid \gamma : (t_0 - \epsilon, t_0 + \epsilon) \rightarrow \mathbb{R}^d \text{ cont\u00ednua, com } \gamma(t_0) = x_0\}.$$

Segue do Exemplo (1.1) que o conjunto \mathcal{X} , munido da norma d (denotado por (\mathcal{X}, d)) é um espaço métrico completo. Em (\mathcal{X}, d) vamos definir o operador $\mathcal{F} : \mathcal{X} \rightarrow \mathcal{C}$ por

$$\mathcal{F}(\gamma(t)) = x_0 + \int_{t_0}^t F(s, \gamma(s)) ds \quad (2.11)$$

e observar que não existem problemas de integração uma vez que $F(s, \gamma(s))$ é contínua e limitada para $s \in [t_0, t]$, o que garante a existência da integral.

Além disso, o operador \mathcal{F} está bem definido, isto é, $\mathcal{F}(\gamma) \in \mathcal{X}$, $\forall \gamma \in \mathcal{X}$, pois

⁴Por \overline{Q} estamos denotando o fecho do conjunto Q , isto é, a união de Q com todos os seus pontos de acumulação.

1. O operador $\mathcal{F}(\gamma(t))$ é contínuo para $t \in (t_0 - \epsilon, t_0 + \epsilon)$;
2. $\mathcal{F}(\gamma(t_0)) = x_0$, pois

$$\mathcal{F}(\gamma(t_0)) = x_0 + \int_{t_0}^{t_0} F(s, \gamma(s)) ds = x_0 + 0 = x_0.$$

3. $\mathcal{F}(\gamma(t)) \in \overline{B_\delta(x_0)}$, pois

$$\|\mathcal{F}(\gamma(t)) - x_0\| = \left\| \int_{t_0}^t F(s, \gamma(s)) ds \right\| \leq M_\delta |t - t_0| < M_\delta \epsilon < \delta.$$

O operador dado por (2.11) é uma contração em \mathcal{X} , isto é, existe $0 \leq \lambda < 1$ tal que

$$d(\mathcal{F}(\gamma_1(t)), \mathcal{F}(\gamma_2(t))) \leq \lambda d(\gamma_1(t), \gamma_2(t)),$$

para quaisquer $\gamma_1, \gamma_2 \in \mathcal{X}$ e $t \in (t_0, t_0 + \epsilon)$. De fato, de (2.10) temos

$$\begin{aligned} \|\mathcal{F}(\gamma_1(t)) - \mathcal{F}(\gamma_2(t))\| &= \left\| \int_{t_0}^t F(s, \gamma_1(s)) - F(s, \gamma_2(s)) ds \right\| \leq \\ &\int_{t_0}^t C \|\gamma_1(s) - \gamma_2(s)\| ds \leq C |t - t_0| d(\gamma_1(t), \gamma_2(t)) \leq \\ &C \epsilon d(\gamma_1(t), \gamma_2(t)) \end{aligned}$$

e basta tomar $\lambda = C\epsilon$.

Ora, sendo (2.11) uma contração num espaço métrico completo, segue do Teorema (1.1) que existe uma única curva $\gamma^* \in \mathcal{X}$ tal que $\mathcal{F}(\gamma^*(t)) = \gamma^*(t)$, ou seja,

$$\gamma^*(t) = x_0 + \int_{t_0}^t F(s, \gamma^*(s)) ds \quad \forall t \in (t_0 - \epsilon, t_0 + \epsilon).$$

Finalmente, notando que (2.2) é diferenciável, segue do Teorema Fundamental do Cálculo que

$$(\gamma^*)'(t) = F(t, \gamma^*(t)), \quad \forall t \in (t_0 - \epsilon, t_0 + \epsilon). \quad (2.12)$$

Isto conclui a demonstração do item 1.

Para a prova do item 2 sejam $\gamma_1 : I_1 \rightarrow \mathbb{R}^d$ e $\gamma_2 : I_2 \rightarrow \mathbb{R}^d$ soluções de $x' = F(t, x)$ e seja $t_0 \in I_1 \cap I_2$ com $\gamma_1(t_0) = \gamma_2(t_0) = x_0$.

Definindo

$$J = \{t \in I_1 \cap I_2 : \gamma_1(t) = \gamma_2(t)\},$$

vemos que:

1. J é não vazio, pois $t_0 \in J$;
2. J é fechado, pois o conjunto dos pontos onde duas funções contínuas coincidem é fechado (lembre-se de que $\gamma_1(t)$ e $\gamma_2(t)$ são contínuas);

3. J é aberto; para ver isso, basta tomar $s_0 \in J$ e considerar $y_0 = \gamma_1(s_0) = \gamma_2(s_0)$. Para (s_0, y_0) , repetindo o argumento usado na prova do item 1 deste teorema, podemos encontrar $\epsilon > 0$ tal que $\mathcal{F} : \mathcal{X} \rightarrow \mathcal{X}$ é uma contração e possui um único ponto fixo, de onde $\gamma_1(t) = \gamma_2(t)$ para $t \in (s_0 - \epsilon, s_0 + \epsilon)$ e portanto $(s_0 - \epsilon, s_0 + \epsilon) \in J$.

Pelos itens 1, 2 e 3 acima, $J = I_1 \cap I_2$ e isto prova o resultado. □

A demonstração acima é construtiva, pois (2.12) estabelece que a única solução $\gamma(t)$ da equação $x' = F(t, x)$ sujeita a $\gamma(t_0) = x_0$ é o ponto fixo do operador \mathcal{F} que pode ser obtido através de sucessivas iterações, como feito na prova do Teorema (1.1). Ilustramos esta afirmação no Exemplo (2.6).

Exemplo 2.6. Considere a equação diferencial $x' = 2tx$ sujeita à condição inicial $(t_0, x_0) = (0, 1)$.

Definindo o operador como em (2.11) para esta equação diferencial, temos

$$\mathcal{F}(\gamma(t)) = 1 + \int_0^t 2s\gamma(s)ds.$$

Tomando qualquer curva $\gamma(t)$ tal que $\gamma(0) = 1$ como primeira aproximação para a solução, temos garantida a convergência das sucessivas iteradas do operador \mathcal{F} para a solução de $x' = 2tx$. Consideraremos a mais simples das curvas com $\gamma(0) = 1$, a curva constante $\gamma_0(t) \equiv 1$ e então

$$\begin{aligned}\gamma_1(t) &= \mathcal{F}(\gamma_0(t)) = 1 + \int_0^t 2sds = 1 + t^2, \\ \gamma_2(t) &= \mathcal{F}(\gamma_1(t)) = 1 + \int_0^t 2s(1 + s^2)ds = 1 + t^2 + \frac{t^4}{2},\end{aligned}$$

e por indução temos

$$\gamma_n(t) = 1 + t^2 + \frac{t^4}{2} + \frac{t^6}{6} + \dots + \frac{t^{2n}}{n!} + \dots \quad (2.13)$$

Finalmente, fazendo $n \rightarrow \infty$ em (2.13) segue que

$$\gamma(t) = e^{t^2}$$

é a solução de $x' = 2tx$ sujeita à $(t_0, x_0) = (0, 1)$, como é fácil de verificar.

Observação 2.1. É comum substituímos a hipótese da função $F : U \rightarrow \mathbb{R}^d$ ser lipschitziana na segunda variável pela hipótese mais clássica da derivada parcial espacial $\frac{\partial F}{\partial x}$ ser contínua. Embora a continuidade dessas derivadas parciais seja mais restritiva, ela é de mais fácil verificação. A substituição das hipóteses não altera as conclusões do Teorema (2.2) pois a continuidade das derivadas parciais garante a condição de Lipschitz em conjuntos compactos de \mathbb{R}^d .

Observação 2.2. Exigindo no enunciado do Teorema (2.2) apenas a continuidade da função $F : U \rightarrow \mathbb{R}^d$ e relaxando a exigência de ser Lipschitz, ainda é possível provar a existência de soluções, mas se perde a propriedade de unicidade.

2.3 Algumas técnicas de resolução

Apesar de não existir um método geral para se encontrar a solução de equações diferenciais, para alguns tipos particulares é possível encontrá-la seja de forma explícita ou implícita.

Na sequência apresentaremos a solução de alguns tipos de equações diferenciais; em particular trataremos de alguns casos escalares, nos baseando na referência [4]. Apesar da variedade de técnicas empregadas observamos que a solução encontrada em cada caso é de fato a única solução existente o que pode ser verificado diretamente por derivação e justificado pelo Teorema (2.2).

2.3.1 Variáveis separáveis

Uma equação diferencial é chamada de **separável** se possuir a forma

$$x'(t) = \frac{f(t)}{g(x)} \quad \text{com } g(x) \neq 0, \quad (2.14)$$

onde as funções f e g são contínuas em algum intervalo de \mathbb{R} .

Tratando o símbolo $\frac{dx}{dt}$ como um quociente formal ⁵

$$g(x)dx = f(t)dt,$$

e portanto, a família de curvas integrais de (2.14) satisfaz

$$\int g(x)dx = \int f(t)dt.$$

Exemplo 2.7. Suponha que uma população de P animais viva isolada e que a taxa de variação do número de contaminados $x(t)$ por uma certa doença no instante t seja proporcional ao número de indivíduos sadios e ao número de indivíduos contaminados. É possível prever que, nestas condições, toda a população se contaminará pela doença para t suficientemente grande.

De fato, sendo P o tamanho da população (considerado constante) e $x(t)$ o número de indivíduos doentes, temos que $P - x(t)$ representa o número de indivíduos sadios. Supondo a taxa de variação $\frac{dx}{dt}$ proporcional a $x(t)$ e a $P - x(t)$ então a equação diferencial que modela o problema nas hipóteses dadas é

$$\frac{dx}{dt} = \alpha x(P - x), \quad (2.15)$$

⁵A justificativa para a validade de tal manipulação algébrica é a teoria das **formas diferenciais**. Uma forma diferencial em \mathbb{R}^2 é uma função $w : \Omega \rightarrow (\mathbb{R}^2)^*$ onde Ω é um aberto do plano e $(\mathbb{R}^2)^*$ é o **espaço dual** (dos funcionais lineares) de \mathbb{R}^2 . Uma introdução ao assunto pode ser encontrada em [5].

que é do tipo (2.14) e portanto pode ser reescrita na forma

$$\frac{dx}{x(P-x)} = \alpha dt.$$

Por integração, tem-se

$$\int \frac{dx}{x(P-x)} = \int \alpha dt,$$

de onde

$$\int \frac{1}{P} \left(\frac{1}{x} + \frac{1}{P-x} \right) dx = \alpha t + k_1,$$

$$\int \left(\frac{1}{x} + \frac{1}{P-x} \right) dx = P\alpha t + k_2,$$

$$\ln \frac{x}{P-x} = P\alpha t + k_2,$$

$$\frac{x}{P-x} = ke^{P\alpha t}.$$

Finalmente, isolando a função $x(t)$ na última das igualdades acima encontramos a solução de (2.15):

$$x(t) = \frac{P}{1 + \frac{1}{k}e^{-P\alpha t}}. \quad (2.16)$$

Tomando o limite de (2.16) quando $t \rightarrow \infty$ temos:

$$\lim_{t \rightarrow \infty} x(t) = \lim_{t \rightarrow \infty} \frac{P}{1 + \frac{1}{k}e^{-P\alpha t}} = P,$$

e portanto o número $x(t)$ de contaminados tende a P .

2.3.2 Equações diferenciais lineares escalares de primeira ordem

A forma geral de uma equação diferencial linear de primeira ordem é

$$x'(t) + p(t)x(t) = q(t), \quad (2.17)$$

onde $p, q : (a, b) \rightarrow \mathbb{R}$ são funções contínuas de variável independente t definidas no aberto (a, b) e $x(t)$ é não nula em (a, b) .

A solução geral $x(t)$ pode ser determinada pelo **método do fator integrante**, onde a equação diferencial (2.17) é multiplicada por uma função $\mu = \mu(t)$ apropriada, de maneira que a função obtida seja facilmente integrável. Seja portanto $\mu = \mu(t)$ um fator integrante para (2.17); então

$$\mu(t)x'(t) + \mu(t)p(t)x(t) = \mu(t)q(t). \quad (2.18)$$

Impondo a condição de que o primeiro membro de (2.18) seja a derivada do produto $\mu(t)x(t)$ (ou seja, que o primeiro membro seja igual a $\mu'(t)x(t) + \mu(t)x'(t)$), segue de (2.18) que

$$\mu'(t)x(t) + \mu(t)x'(t) = \mu(t)x'(t) + \mu(t)p(t)x(t),$$

de onde (uma vez que por hipótese $x(t)$ é não nula)

$$\mu'(t) = \mu(t)p(t). \quad (2.19)$$

Note agora que (2.19) é uma equação separável; resolvendo-a pelo método da seção anterior encontramos

$$\mu(t) = e^{\int_0^t p(s)ds}. \quad (2.20)$$

Agora, multiplicando a equação (2.17) por (2.20) temos

$$e^{\int_0^t p(s)ds}x'(t) + e^{\int_0^t p(s)ds}p(t)x(t) = e^{\int_0^t p(s)ds}q(t),$$

cujos primeiros membros são a derivada do produto $x(t)e^{\int_0^t p(s)ds}$. Integrando de ambos os lados podemos escrever

$$x(t)e^{\int_0^t p(s)ds} = \int q(s)e^{\int_0^s p(\xi)d\xi}ds + c.$$

e isolando a função $x(t)$ obtemos

$$x(t) = \frac{1}{e^{\int_0^t p(s)ds}} \int q(s)e^{\int_0^s p(\xi)d\xi}ds + c. \quad (2.21)$$

A equação (2.21) tem interesse principalmente teórico, pois indica que é possível obter as soluções da equação (2.17). Como as integrações existentes em (2.21) podem não ser expressas em termos de funções elementares em todos os casos, pode não ser possível obter uma expressão explícita para a solução.

Exemplo 2.8. A equação

$$\frac{dx}{dt} - 2x = 4 - t \quad (2.22)$$

é linear de primeira ordem com $p(t) = -2$ e $q(t) = 4 - t$.

Por (2.19), o fator integrante será

$$\mu(t) = e^{\int_0^t -2ds} = e^{-2t} \quad (2.23)$$

e multiplicando a equação (2.22) por (2.23) segue que

$$e^{-2t} \frac{dx}{dt} - 2e^{-2t}x(t) = 4e^{-2t} - te^{-2t}.$$

Por integração, temos

$$\int (x(t)e^{-2t})' dt = \int (4e^{-2t} - te^{-2t}) dt \Rightarrow$$

$$x(t)e^{-2t} = -2e^{-2t} + \frac{1}{2}te^{-2t} + \frac{1}{4}e^{-2t} + c,$$

e finalmente, ao isolarmos a função $x(t)$ na última das igualdades acima encontramos a solução geral:

$$x(t) = -\frac{7}{4} + \frac{1}{2}t + ce^{2t}, \quad \text{com } c \in \mathbb{R}.$$

2.3.3 Equações exatas

Como motivação para o que apresentaremos a seguir, considere a equação diferencial

$$2t + x^2 + 2txx' = 0, \quad (2.24)$$

que não é linear e tampouco separável e assim, não pode ser resolvida com os métodos desenvolvidos nas duas seções anteriores.

Note porém, que a função $\psi(t, x) = t^2 + tx^2$ tem derivadas parciais

$$\frac{\partial \psi}{\partial t} = 2t + x^2, \quad \frac{\partial \psi}{\partial x} = 2tx \quad (2.25)$$

e portanto, a equação (2.24) se escreve na forma

$$\frac{\partial \psi}{\partial t} + \frac{\partial \psi}{\partial x} \frac{dx}{dt} = 0.$$

Se supusermos x uma função da variável independente t segue que

$$\frac{\partial \psi}{\partial t} = \frac{d}{dt}(t^2 + tx^2) = 0.$$

e sendo c uma constante qualquer, obtemos

$$\psi(t, x) = t^2 + tx^2 = c,$$

que define implicitamente uma função que satisfaz (2.24).

Analisando o que fizemos acima, vemos que só foi possível resolver a equação (2.24) porque foi observada a existência de uma função ψ satisfazendo (2.25).

Mais geralmente, suponha que dada a equação

$$M(t, x) + N(t, x)x' = 0 \quad (2.26)$$

seja possível identificar uma função ψ tal que

$$\frac{\partial \psi}{\partial t} = M(t, x) \quad e \quad \frac{\partial \psi}{\partial x} = N(t, x) \quad (2.27)$$

e tal que $\psi(t, x) = c$ defina $x = \phi(t)$ implicitamente como uma função diferenciável de t . Então

$$M(t, x) + N(t, x)x' = \frac{\partial \psi}{\partial t} + \frac{\partial \psi}{\partial x} \frac{dx}{dt} = \frac{d}{dt}\psi(t, \phi(t)),$$

de modo que (2.26) se escreve como

$$\frac{d}{dt}\psi(t, \phi(t)) = 0. \quad (2.28)$$

Neste caso, a equação (2.26) é chamada de **diferencial exata** e suas soluções são definidas implicitamente por $\psi(t, x) = c$ onde c é uma constante arbitrária. Para reconhecer uma equação diferencial exata, usamos o teorema a seguir. Na sua demonstração usaremos o Lema (2.2), cuja prova pode ser encontrada em [6].

Lema 2.2. (Teorema de Schwarz) *Seja $f : U \rightarrow \mathbb{R}$ duas vezes diferenciável no ponto $c \in U \subset \mathbb{R}^d$. Para quaisquer $1 \leq i, j \leq 2$ tem-se*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(c) = \frac{\partial^2 f}{\partial x_j \partial x_i}(c). \quad (2.29)$$

Teorema 2.3. *Suponha que as funções M, N, M_x e N_t , onde os índices denotam derivadas parciais, sejam contínuas na região retangular*⁶

$$R = \{(x, y) \in \mathbb{R}^2; \alpha \leq x \leq \beta, \delta \leq y \leq \epsilon\}.$$

Então a equação (2.26) é exata em R se e somente se

$$M_x(t, x) = N_t(t, x) \quad (2.30)$$

em cada ponto de R . Isto é, existe uma função $\psi(x, y)$ satisfazendo (2.27) se e somente se M e N satisfizerem (2.29).

Demonstração. Faremos a demonstração em duas partes. Vamos inicialmente mostrar que se existe uma função ψ tal que (2.27) é válida então a equação (2.29) é satisfeita.

Calculando M_x e N_t das equações (2.27) obtemos

$$M_x(t, x) = \psi_{tx}(t, x) \quad e \quad N_t(t, x) = \psi_{xt}(t, x) \quad (2.31)$$

de onde (2.29) é satisfeita.

Provemos agora que, se M e N satisfizerem (2.29) então a equação (2.26) é exata. A prova envolve a construção de uma função ψ satisfazendo (2.27). Integrando a primeira das equações em (2.27) em relação a t e mantendo x constante, vemos que

$$\psi(t, x) = \int M(t, x)dt + g(x), \quad (2.32)$$

onde a função g é uma função arbitrária de x , fazendo o papel da constante de integração.

⁶Por simplicidade estamos supondo uma região retangular, o que não é necessário; basta que a região R seja simplesmente conexa. Em dimensão 2, isto significa que R não tem "buracos" em seu interior; assim regiões retangulares ou circulares são simplesmente conexas, mas uma coroa circular não é. Maiores detalhes podem ser encontrados em livros de Cálculo Avançado.

Vamos mostrar que sempre é possível escolher g de modo que $\psi_x = N$. Da equação (2.32), temos

$$\psi_x(t, x) = \frac{\partial}{\partial x} \int M(t, x) dt + g'(x) = \int M_x(t, x) dt + g'(x).$$

e fazendo $\psi_x = N$, para $g'(x)$ obtemos

$$g'(x) = N(t, x) - \int M_x(t, x) dt. \quad (2.33)$$

Para se determinar a função g em (2.33) é essencial que, independente da aparência, a expressão à direita da igualdade seja uma função apenas de t . Note que diferenciando (2.33) em relação a t obtemos

$$N_t(t, x) - M_x(t, x) = 0,$$

por (2.30). Logo, a expressão à direita da igualdade em (2.33) não depende de fato da variável t e com apenas uma integração obtemos g . Substituindo agora g em (2.32), obtemos como solução

$$\psi(t, x) = \int M(t, x) dt + \int \left[N_t(t, x) - \int M_x(t, x) dt \right] dx. \quad (2.34)$$

□

A demonstração acima é construtiva e fornece um método para o cálculo de $\psi(t, x)$ e portanto para a obtenção da solução de (2.26). Em geral, é mais prático fazer o processo inteiro a cada vez que for necessário do que lembrar a fórmula (2.34).

Exemplo 2.9. Vamos resolver a equação diferencial

$$(x \cos t + 2te^x) + (\sin t + t^2e^x - 1)x' = 0, \quad (2.35)$$

onde $M(t, x) = x \cos t + 2te^x$ e $N(t, x) = \sin t + t^2e^x - 1$.

Temos $M_x(t, x) = \cos t + 2te^x = N_t(t, x)$ e pelo Teorema (2.3) a equação (2.35) é exata, de modo que existe $\psi(t, x)$ tal que

$$\psi_t(t, x) = x \cos t + 2te^x \quad e \quad \psi_x(t, x) = \sin t + t^2e^x - 1.$$

Integrando a primeira das equações acima em relação a t , obtemos

$$\psi(t, x) = x \sin t + t^2e^x + g(x), \quad (2.36)$$

e fazendo $\psi_x(t, x) = N$ temos

$$\psi_x(t, x) = \sin t + t^2e^x + g'(x) = \sin t + t^2e^x - 1; \quad (2.37)$$

assim, $g'(x) = -1$ e portanto $g(x) = -x$ (a constante de integração é omitida, já que qualquer solução da equação $g'(x) = -1$ pode ser usada; não é necessário usar a mais geral possível). Finalmente, substituindo $g(x)$ na equação (2.36) temos

$$\psi(t, x) = x \operatorname{sen} t + t^2 e^x - x$$

e portanto as soluções de (2.35) são dadas implicitamente por

$$x \operatorname{sen} t + t^2 e^x - x = c, \quad (2.38)$$

onde c é uma constante real arbitrária.

2.4 Modelos descritos por equações diferenciais

Na Seção (2.3) deste Capítulo, resolvemos alguns tipos particulares de equações diferenciais. Os métodos estão longe de serem gerais e se aplicam a classes muito restritas de equações. Nesta Seção, analisaremos alguns modelos clássicos com as ferramentas desenvolvidas.

Retornaremos a alguns destes modelos no Apêndice A deste texto, quando apresentaremos propostas de trabalho para professores de matemática.

2.4.1 Modelo populacional de Malthus

O economista britânico Thomas Malthus (1766-1834) foi o primeiro a observar que algumas populações biológicas crescem a uma taxa proporcional à população e seu primeiro artigo apareceu em 1798.

A taxa de proporcionalidade r , chamada de **taxa de crescimento ou declínio** conforme seja positiva ou negativa, foi suposta constante por Malthus e tomada como a diferença entre o coeficiente de natalidade e coeficiente de mortalidade, representadas por m e n respectivamente. Portanto, $r = m - n$.

Em termos matemáticos, seja $x = x(t) > 0$ a população de uma espécie no instante de tempo t . Supondo que a taxa de crescimento em função do tempo, isto é, que $\frac{dx}{dt}$ seja proporcional à população atual $x(t)$ com constante de proporcionalidade r temos

$$\frac{dx}{dt} = rx(t). \quad (2.39)$$

Observe que a equação (2.39) é separável e usando o método apresentado na seção (2.3.1) obtemos a solução geral:

$$x(t) = ce^{rt}, \quad (2.40)$$

sendo c uma constante arbitrária.

Suponha que a população inicial seja x_0 , isto é, $x(0) = x_0$. Então a constante c pode ser calculada explicitamente; de (2.40) segue que

$$x_0 = ce^{r0}$$

e portanto, $c = x_0$.

Assim, a solução de (2.39) com condição inicial $x(0) = x_0$ é dada por

$$x(t) = x_0 e^{rt}, \quad (2.41)$$

onde $r = m - n$. O comportamento da solução é determinado pela constante r , pois:

1. Se $r > 0$ então o crescimento é exponencial e se verifica quando $m > n$, ou seja, quando a taxa de natalidade é maior que a taxa de mortalidade;
2. Se $r < 0$ então o decrescimento é exponencial sugerindo a extinção da população e se verifica quando $m < n$, ou seja, quando a taxa de natalidade é menor que a taxa de mortalidade;
3. Se $r = 0$ a população permanece constante, correspondendo à igualdade entre as taxas de natalidade e mortalidade.

No gráfico abaixo encontram-se ilustradas as soluções da equação(2.39) conforme o sinal da constante r .

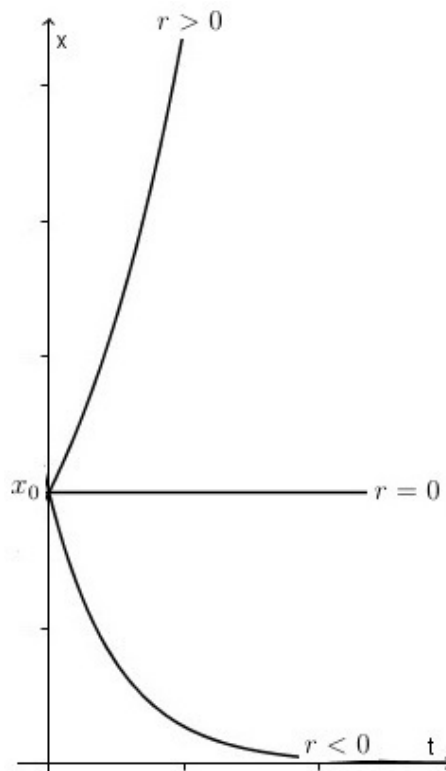


Figura 2.2: Gráfico das soluções do modelo de Malthus.

Para variações pequenas de tempo o modelo pode se mostrar bastante preciso. O exemplo a seguir mostra a adequabilidade do modelo para intervalos de tempo pequenos.

Exemplo 2.10. Segundo dados oficiais da Prefeitura Municipal de São Paulo⁷ as taxas de mortalidade e natalidade na cidade de São Paulo entre os anos de 2001 e 2005 são dadas pela Tabela 2.1.

Ano	Taxa de Natalidade	Taxa de Mortalidade
2000	0,0199	0,00655
2001	0,0182	0,00637
2002	0,0176	0,00638
2003	0,0173	0,00631
2004	0,0172	0,00631
2005	0,0167	0,00596

Tabela 2.1: Taxas de natalidade e mortalidade de São Paulo entre 2000 e 2005.

Na análise que faremos vamos considerar a média aritmética das taxas apresentadas, ou seja, tomaremos $m = 0,01781667$ e $n = 0,00631333$, de modo que a constante de proporcionalidade do modelo malthusiano que desejamos construir é $r = 0,01150333$. Assim, de acordo com (2.41) temos a seguinte equação para o modelo de Malthus:

$$x(t) = 10434252e^{0,01150333t}. \quad (2.42)$$

Abaixo são apresentados os dados oficiais da população do município de São Paulo e os valores calculados através de (2.42).

Ano	Dados Oficiais	Dados do Modelo de Malthus	Erro Percentual
2000	10434252	10434252	0
2001	10525367	10554974	0,28
2002	10613691	10677092	0,6
2003	10698381	10800623	0,96
2004	10782296	10925584	1,33
2005	10865573	11051990	1,72
2006	10944889	11179859	2,15
2007	11019484	11309207	2,63
2008	11093746	11440052	3,12
2009	11168194	11572410	3,62
2010	11244369	11706300	4,11

Tabela 2.2: Erro no Modelo de Malthus - População do Município de São Paulo entre 2000-2010.

Analisando os dados da tabela acima, vemos que o modelo de Malthus, tão simplista, que não considera as especificidades de uma cidade como São Paulo e nem fatores como

⁷Retirados de <http://www9.prefeitura.sp.gov.br/sempla/md/> em 03/02/2013.

imigração e migração se adapta tão bem à realidade. Contudo, como já mencionado, o crescimento exponencial de uma população não se mantém para longos períodos de tempo.

2.4.2 O modelo de Verhulst

A hipótese de que a taxa de crescimento de uma população r é constante não é razoável, pois ela não leva em conta que o crescimento de uma população aciona automaticamente mecanismos de controle visando reduzir a taxa de crescimento. Esta situação tem sido observada em várias populações, onde a superpopulação modifica o funcionamento fisiológico das espécies alterando seus hábitos sexuais e comportamento coletivo.

O modelo proposto por Verhulst consiste em supor que a taxa de crescimento de uma população decresce linearmente com ela, ou seja, que $r = a - bx$ com a, b sendo constantes positivas. Ainda não é a hipótese ideal porque não leva em consideração o fato de que indivíduos novos não contribuem de imediato para o aumento da espécie, mas em geral, modela com maior precisão o crescimento/decrescimento populacional se comparado com o modelo malthusiano.

Em termos de equações diferenciais, temos⁸

$$\frac{dx}{dt} = (a - bx)x \quad (2.43)$$

e este modelo foi chamado por Verhulst de **modelo do crescimento logístico**; também é costume se referir à equação (2.43) como **equação logística**.

Observe que (2.43) é uma equação autônoma cujas soluções constantes são dadas por $x(t) \equiv \frac{a}{b}$ e $x(t) \equiv 0$. Note ainda que $x(t) \equiv 0$ não se aplica ao problema, uma vez que não faz sentido considerar uma população nula.

A equação (2.43) é separável e portanto

$$\int \frac{1}{(a - bx)x} dx = \int dt.$$

Por frações parciais vemos que $\frac{1}{(a - bx)x} = \frac{1}{ax} + \frac{b}{a(a - bx)}$. Então,

$$\int \frac{1}{ax} dx + \int \frac{b}{a(a - bx)} dx = t + k.$$

de modo que

$$\ln \left| \frac{x}{a - bx} \right| = at + ak \Rightarrow |a - bx|e^{at}e^{ak} = |x|$$

⁸Na literatura esta equação é chamada de **equação de Verhulst-Pearl** por ter sido considerada pela primeira vez por Verhulst em 1834 e estudada por volta de 1920 por Pearl e Reed, no estudo da população norte americana.

Dada uma condição inicial $x(t_0) = x_0$, segue desta última equação que

$$|x_0| = |a - bx_0|e^{at_0+ak},$$

e supondo soluções iniciais diferentes das soluções de equilíbrio temos

$$\left| \frac{x}{x_0} \right| = \left| \frac{a - bx}{a - bx_0} \right| e^{a(t-t_0)}.$$

Como pelo Teorema (2.2) as soluções de (2.43) não intersectam as soluções constantes e nem a si próprias, de modo que podemos retirar os valores absolutos e explicitar $x(t)$, solução de (2.43) sujeita à condição $x(t_0) = x_0$:

$$x(t) = \frac{ax_0}{bx_0 + (a - bx_0)e^{-a(t-t_0)}}. \quad (2.44)$$

Note que se $t \rightarrow +\infty$ então $x(t) \rightarrow \frac{a}{b}$ independente da condição inicial $x(t_0) = x_0 > 0$ e por isso $\frac{a}{b}$ é chamado de **valor assintótico da população**; temos portanto o curioso fato de que as soluções tenderão para a solução de equilíbrio $x(t) = \frac{a}{b}$ independente das condições iniciais.

Em alguns textos o valor $\frac{a}{b}$ é chamado de **nível de saturação** ou **capacidade de suporte**.

Observe abaixo o comportamento das soluções do modelo de Verhulst conforme a condição inicial x_0 esteja abaixo ou acima da capacidade de suporte:

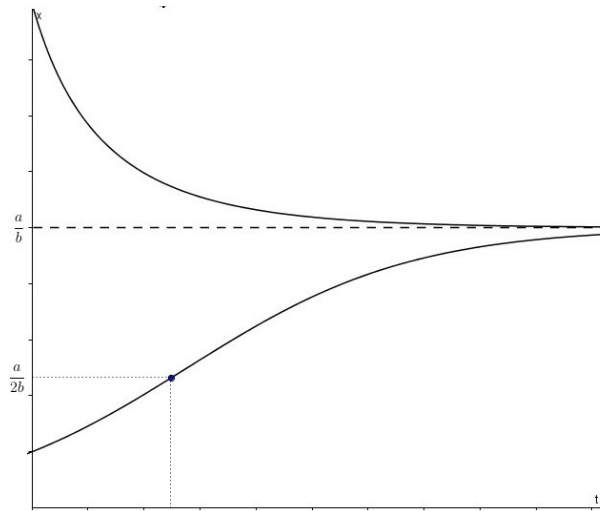


Figura 2.3: Gráfico das soluções do modelo de Verhulst

Dado $x_0 > \frac{a}{b}$, a população $x(t)$ decresce exponencialmente para $x(t) = \frac{a}{b}$; contudo, se $x_0 < \frac{a}{b}$, a população $x(t)$ cresce tendendo para $\frac{a}{b}$ e nesse caso o gráfico está entre $x = 0$ e $x = \frac{a}{b}$ e possui forma de S. Note que há um ponto de inflexão em $x = \frac{a}{2b}$ pois derivando (2.43) temos

$$x''(t) = (a - 2b)x'(t),$$

o que indica que o crescimento é mais rápido até a população atingir **metade** da capacidade de suporte e após este estágio passa a crescer de maneira mais lenta.

A equação logística se mostra aplicável à modelagem de problemas das mais variadas naturezas. Discutiremos dois destes modelos na sequência.

2.4.3 Propagação da podridão em maçãs

A armazenagem de maçãs se dá em caixas de madeira que comportam em média 3000 frutas. Experimentalmente, verifica-se que se uma fruta estiver apodrecida, em doze dias cerca de 80% das maçãs da caixa estarão podres⁹.

Suponha que a velocidade de propagação da podridão seja proporcional ao número de maçãs sadias $3000 - x(t)$ e ao número de maçãs podres $x(t)$ sendo m a razão de proporcionalidade; nestas condições, a equação (logística) para o modelo é

$$\frac{dx}{dt} = m(3000 - x)x, \quad (2.45)$$

sujeita à $x(0) = 1$, pois no instante $t = 0$ supomos a existência de apenas uma maçã podre.

Por separação de variáveis temos

$$\frac{1}{x(3000 - x)} dx = m dt,$$

de onde

$$\frac{1}{3000} \left(\int \frac{1}{x} dx + \int \frac{1}{3000 - x} dx \right) = \int m dt,$$

uma vez que, por frações parciais, segue que

$$\frac{1}{x(3000 - x)} = \frac{1}{x} + \frac{1}{3000 - x}.$$

Resolvendo as integrais encontramos

$$\begin{aligned} \ln x - \ln(3000 - x) &= 3000mt + D \\ \frac{x}{3000 - x} &= Ce^{3000mt} \end{aligned}$$

e isolando a variável $x(t)$ podemos determiná-la explicitamente:

$$x(t) = \frac{3000Ce^{3000mt}}{1 + Ce^{3000mt}}, \quad (2.46)$$

onde a constante C pode ser estimada usando a condição $x(0) = 1$. Substituindo $x = 1$ e $t = 0$ em (2.46) segue que

$$\frac{3000C}{1 + C} = 1 \Rightarrow 2999C = 1 \Rightarrow C \cong \frac{1}{3000}.$$

⁹O modelo deste problema com a curva logística foi apresentado no livro *Temas e Modelos* do Professor Rodney Bassanezi. Consulte a referência [7].

e a equação (2.46) se reescreve como

$$x(t) = \frac{e^{3000mt}}{1 + \frac{1}{3000}e^{3000mt}} = \frac{3000e^{3000mt}}{3000 + e^{3000mt}} = \frac{3000}{1 + 3000e^{-3000mt}}. \quad (2.47)$$

Como $x(12) = 2400$ a constante m de proporcionalidade também pode ser estimada por simples substituição em (2.47), ou seja,

$$\begin{aligned} 2400 &= \frac{3000}{1 + 3000e^{-36000m}} \Rightarrow 4 + 12000e^{-36000m} = 5 \\ \Rightarrow m &= \frac{\ln 12000}{36000} \Rightarrow m \cong 0,000261. \end{aligned}$$

Finalmente substituindo $m = 0,000261$ em (2.47) é possível escrever explicitamente a solução do modelo da propagação da podridão em maçãs:

$$x(t) = \frac{3000}{3000e^{-0,783t} + 1}. \quad (2.48)$$

O tempo (em dias) necessário para que a caixa toda com 3000 frutas esteja inteiramente podre não pode ser calculado a partir de (2.48), pois fazendo $x(t) = 3000$ encontraríamos

$$3000e^{-0,783t} + 1 = 1,$$

que não tem solução. Neste caso, uma boa estratégia é modificar (2.48) para realizarmos previsões de tempo necessário para que cada *porcentagem* de frutas na caixa apodreça; devemos portanto, expressar t em função de $x = 3000p$, ou seja,

$$\begin{aligned} 3000p &= \frac{3000}{3000e^{-0,783t} + 1}, \\ 3000pe^{-0,783t} + p &= 1, \\ e^{-0,783t} &= \frac{1-p}{3000p}. \end{aligned}$$

Usando logaritmos naturais isolamos a variável t , obtendo assim

$$t = -\frac{1}{0,783} \cdot \ln \frac{1-p}{3000p}. \quad (2.49)$$

que apesar de não estar definida para $p = 1$ (que equivale a 100% de maçãs podres), indica que uma boa estimativa pode ser conseguida tomando p suficientemente próximo de $p = 1$; fazendo $p = 0,99$ em (2.49) temos

$$t = -\frac{1}{0,783} \cdot \ln \frac{1-0,99}{3000 \cdot 0,99} \Rightarrow t \cong 16 \text{ dias.}$$

2.4.4 Disseminação de doenças contagiosas

A utilização de métodos matemáticos para estudar a disseminação de doenças contagiosas surgiu em 1760 quando Daniel Bernoulli (1700 – 1782) fez trabalhos sobre a varíola ¹⁰.

Um modelo simples¹¹ para o estudo das epidemias pode ser construído considerando que uma população esteja dividida em duas partes: os que tem a doença e podem infectar outros e os que não a tem, mas são suscetíveis a contraí-la.

Sejam y a proporção dos indivíduos suscetíveis e x a proporção dos infectados; então $x + y = 1$. Suponha que a doença se espalhe através do contato entre os doentes e os sãos, e que a taxa de disseminação $\frac{dx}{dt}$ seja proporcional ao número de tais contatos; suponha ainda que os indivíduos movem-se livremente entre si, de modo que o número de contatos seja proporcional ao produto xy .

Nestas condições, obtemos

$$\frac{dx}{dt} = \alpha x(1 - x), \quad (2.50)$$

sujeita à $x(0) = x_0$, onde α é a constante de proporcionalidade e x_0 é a população inicial de indivíduos contaminados.

Observe que a equação (2.50) é uma equação do tipo (2.43) com $\alpha = a = b = 1$ de modo que sua solução, por (2.44), é

$$x(t) = \frac{x_0}{x_0 + (1 + x_0)e^{-at}}. \quad (2.51)$$

Dentro das hipóteses assumidas, vê-se que fatalmente toda a população se contaminará, ainda que apenas um indivíduo esteja contaminado no instante $t = 0$, pois calculando o limite de (2.51) quando o tempo tende ao infinito vemos que

$$\lim_{t \rightarrow \infty} x(t) = \lim_{t \rightarrow \infty} \frac{x_0}{x_0 + (1 + x_0)e^{-at}} = 1. \quad (2.52)$$

Algumas variações do modelo acima são de análise interessante. Suponha agora a modelagem de doenças (como o tifo) que se disseminam através de portadores que, mesmo não exibindo os sintomas são capazes de transmití-la. Denotemos por x e y a proporção de portadores e suscetíveis da população, respectivamente.

Suponha que os portadores são identificados e removidos da população a uma taxa β , de modo que

$$\frac{dx}{dt} = -\beta x, \quad (2.53)$$

¹⁰Baseado no modelo descrito e usando os melhores dados sobre mortalidade disponíveis em sua época, Bernoulli calculou que se as mortes por varíola pudessem ser eliminadas, poder-se-ia adicionar aproximadamente três anos à expectativa de vida dos europeus, que em 1760 era de 26 anos e 7 meses. Isto fez com que ele apoiasse, na época, um programa controverso de vacinação.

¹¹Maiores detalhes sobre a modelagem de propagação de doenças usando equações diferenciais pode ser encontrada no livro de Bailey, listado na bibliografia, especialmente nos capítulos 5, 10 e 20. Consulte a referência [8].

suponha também que a doença se propague a uma taxa proporcional ao produto de x e y , ou seja, que

$$\frac{dy}{dt} = \alpha xy. \quad (2.54)$$

Resolvendo a equação (2.53) por variáveis separáveis encontramos

$$\begin{aligned} \int \frac{dx}{x} &= \int -\beta dt \\ \ln x &= -\beta t + c, \\ x &= ke^{-\beta t}. \end{aligned}$$

Impondo como condição inicial $x(0) = x_0$, a solução de (2.53) é

$$x = x_0 e^{-\beta t}. \quad (2.55)$$

Substituindo (2.55) na equação (2.54) encontramos

$$\frac{dy}{dt} = \alpha x_0 e^{-\beta t} y \quad (2.56)$$

e novamente por separação de variáveis, temos

$$\frac{dy}{y} = \alpha x_0 e^{-\beta t} dt.$$

Por integração,

$$\begin{aligned} \int \frac{dy}{y} &= \int \alpha x_0 e^{-\beta t} dt, \\ \ln y &= -\frac{\alpha x_0}{\beta} e^{-\beta t} + k, \end{aligned}$$

logo,

$$y(t) = e^{\left(\frac{-\alpha x_0}{\beta}\right) e^{-\beta t}} e^k. \quad (2.57)$$

Admitindo que no instante $t = 0$ fossem y_0 indivíduos suscetíveis (isto é, supondo que $y(0) = y_0$) é possível determinar o valor de e^k em (2.57):

$$y_0 = e^{\left(\frac{-\alpha x_0}{\beta}\right)} e^k$$

e então

$$e^k = y_0 e^{\frac{\alpha x_0}{\beta}}. \quad (2.58)$$

Agora, substituindo (2.58) em (2.57) encontramos

$$\begin{aligned} y(t) &= e^{\frac{-\alpha x_0}{\beta} e^{-\beta t}} y_0 e^{\frac{\alpha x_0}{\beta}}, \\ y(t) &= y(0) e^{\frac{\alpha x_0}{\beta} (1 - e^{-\beta t})}. \end{aligned} \quad (2.59)$$

O ponto curioso deste modelo é que fazendo t tender ao infinito em (2.59) é possível prever que existe uma parcela de indivíduos suscetíveis que escapa da contaminação:

$$\lim_{t \rightarrow \infty} y(t) = \lim_{t \rightarrow \infty} y(0) e^{\frac{\alpha x_0}{\beta} (1 - e^{-\beta t})} = y(0) e^{\frac{\alpha x_0}{\beta}}. \quad (2.60)$$

2.5 Retratos de fase

Consideraremos nesta seção somente equações diferenciais autônomas, isto é, equações diferenciais que independem da variável t (veja Exemplo 2.1) sujeita à condição inicial $x(0) = x_0$ e não exigiremos que f seja lipschitziana. Por simplicidade, escrevemos

$$x' = f(x) \tag{2.61}$$

onde $f : U \subset \mathbb{R}^d \rightarrow \mathbb{R}^d$ é contínua com derivada primeira também contínua.

Dizemos que $I \subset \mathbb{R}$ é um **intervalo máximo** da solução de (2.61) se dada qualquer solução $\gamma : J \rightarrow \mathbb{R}^d$ com $x(0) = x_0$ temos $J \subseteq I$. O intervalo máximo é único e nele está definida uma única solução de $x' = f(x)$. A solução definida no intervalo máximo é chamada de **solução maximal** de (2.61) com $x(0) = x_0$ ou simplesmente **trajetória** por x_0 .

Exemplo 2.11. Tome $U = \mathbb{R}^2$ e considere a equação diferencial

$$\frac{dx}{dt} = x^2$$

sujeita à condição inicial $x(t_0) = x_0$.

A solução, por separação de variáveis, é dada por

$$x(t) = \frac{x_0}{1 - x_0(t - t_0)} \tag{2.62}$$

e o domínio das soluções maximais dependem da análise do valor de x_0 . De fato, tomando $x_0 > 0$ a expressão (2.62) está definida em $I = \left(-\infty, t_0 + \frac{1}{x_0}\right)$; se $x_0 < 0$ então $I = \left(t_0 + \frac{1}{x_0}, +\infty\right)$ e se $x_0 = 0$ devemos ter $I = (-\infty, +\infty)$.

Note que apesar do campo de vetores $f(t, x) = x^2$ estar definido em todo o espaço euclidiano \mathbb{R}^2 nem todas as soluções maximais estão definidas para todo o tempo t .

Toda solução de uma equação diferencial pode ser estendida a uma solução maximal. Em termos de domínio de definição das soluções, a possibilidade de se estender uma solução a uma solução maximal implica que o domínio de uma solução maximal não pode ser ampliado a domínios maiores.

Definição 2.6. Dada $\gamma = \gamma(t)$ uma solução maximal de $x' = f(x)$ com intervalo de definição I , então o conjunto $\{\gamma(t) : t \in I\}$ é chamado de **órbita** e o conjunto U é o **espaço de fase**. Além disso, a representação das órbitas no conjunto U , juntamente com a indicação do sentido do movimento recebe o nome de **retrato de fase** de $x' = f(x)$.

Exemplo 2.12. Todo ponto de equilíbrio x_0 da equação $x' = f(x)$, isto é, todo ponto tal que $f(x_0) = 0$, é uma órbita desta equação, pois a curva constante $\gamma : \mathbb{R} \rightarrow \mathbb{R}^d$ tal que $\gamma(t) = x_0$ é solução de $x' = f(x)$.

Exemplo 2.13. Sem estabelecer condição inicial, de (2.62) podemos verificar que a equação $x' = x^2$ tem solução

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

cujos domínios das soluções maximais dependem da constante c . De fato, tomando $c > 0$ temos $I = (-\infty, c)$; se $c < 0$ então $I = (c, +\infty)$ e se $x_0 = 0$ devemos ter $I = (-\infty, +\infty)$. Portanto, obtemos destes intervalos de definição de soluções maximais as órbitas $\{0\}$, $\left\{ \frac{1}{c-t} : t \in (-\infty, c) \right\} = \mathbb{R}_+$ e $\left\{ \frac{1}{c-t} : t \in (c, +\infty) \right\} = \mathbb{R}_-$.

Exemplo 2.14. Vimos no Exemplo (2.13) que as órbitas da equação

$$x' = x^2 \tag{2.63}$$

são $\{0\}$, \mathbb{R}_+ e \mathbb{R}_- . Isto nos dá o seguinte retrato de fase:

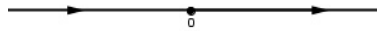


Figura 2.4: Retrato de fase de (2.63).

Exemplo 2.15. Vamos considerar o sistema

$$\begin{cases} x'(t) = y(t) \\ y'(t) = -x(t) \end{cases} \tag{2.64}$$

Trata-se de uma equação autônoma cujo único ponto crítico é a origem $(0, 0)$. Note agora que, se $x(t)$ e $y(t)$ são soluções da equação proposta então derivando a expressão $x^2 + y^2$ implicitamente temos

$$(x^2 + y^2)' = 2xx' + 2yy' = 2xy + 2y(-x) = 0,$$

ou seja, $x^2 + y^2 = k^2$, sendo $k \in \mathbb{R}$ e os retratos de fase são, portanto, circunferências com centro na origem do plano:

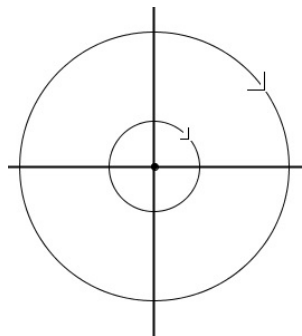


Figura 2.5: Retrato de fase de (2.64).

Exemplo 2.16. Considere agora a equação

$$\begin{cases} x'(t) = y(x^2 - 1), \\ y'(t) = x(x^2 - 1). \end{cases} \quad (2.65)$$

Os pontos críticos são determinados pela resolução do sistema

$$\begin{cases} y(x^2 - 1) = 0, \\ x(x^2 - 1) = 0, \end{cases}$$

cujas soluções são os pontos $(0, 0)$, $(1, y)$ e $(-1, y)$, com $y \in \mathbb{R}$. Por outro lado, ao longo das soluções devemos ter

$$(x^2 - y^2)' = 2xx' - 2yy' = 2xy(x^2 - 1) - 2xy(x^2 - 1) = 0,$$

ou seja, as órbitas estão contidas nas hipérbolas definidas por $x^2 - y^2 = k$, com $k \neq 0$ ou em uma das bissetrizes definidas por $x^2 = y^2$. A seguir, temos um esboço do retrato de fase, onde o sentido do movimento pode ser obtido determinando por exemplo, o sinal de $y' = x(x^2 - 1)$:

$$y' > 0 \iff x \in (-1, 0) \cup (1, +\infty),$$

$$y' < 0 \iff x \in (-\infty, -1) \cup (0, 1).$$

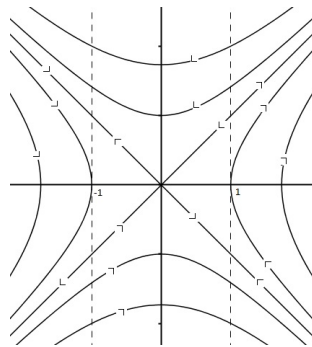


Figura 2.6: Retrato de fase de (2.65).

3 Sistemas de equações diferenciais lineares

Vimos no início do capítulo anterior que toda equação diferencial de ordem superior a um pode ser convenientemente escrita como um sistema de equações de primeira ordem através de uma mudança de variáveis do seguinte modo: dada a equação diferencial $y^{(d)} = a_{d-1}y^{(d-1)} + \dots + a_1y' + a_0y + b(t)$ definindo-se novas coordenadas através de $x_1 = y, x_2 = y', \dots, x_d = y^{(d-1)}$ segue que

$$\begin{aligned} x'_1 &= x_2, \\ x'_2 &= x_3, \\ &\vdots \\ x'_d &= a_{d-1}x_d + a_{d-2}x_{d-1} \dots + a_0x_1 + b(t), \end{aligned}$$

ou seja,

$$X' = AX + g(t)$$

onde $A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ a_0 & a_1 & a_2 & \dots & a_{n-1} \end{pmatrix} \in M_d, X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{d-1} \\ x_d \end{pmatrix}$ e $g(t) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b(t) \end{pmatrix}$ são matrizes

coluna com d linhas.

Para entender a maneira como a mudança de coordenadas age sobre uma equação diferencial de ordem maior que 1, vejamos um exemplo.

Exemplo 3.1. Considere a equação $y''' = -2y'' + y' - 4y$ e as mudanças de coordenadas $x_1 = y, x_2 = y'$ e $x_3 = y''$; segue que $x'_1 = x_2, x'_2 = x_3$ e $x'_3 = -2x_3 + x_2$ e portanto, a equação pode ser reescrita em notação matricial do seguinte modo:

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -4 & 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Neste capítulo vamos estudar a equação

$$X' = AX + g(t),$$

no caso em que a matriz A não depende da variável t sendo uma matriz com entradas constantes e consideraremos a função $g(t)$ como a função identicamente nula.

Trata-se do estudo da equação

$$X' = AX, \tag{3.1}$$

com $A \in M_d$, conhecida na literatura como **equação linear homogênea com coeficientes constantes**. Apresentaremos também uma descrição completa dos retratos de fase em dimensão 2.

Sendo a linguagem matricial a mais apropriada para tratar as questões relativas aos sistemas lineares, utilizaremos seguidamente alguns resultados e a linguagem da Álgebra Linear. Os principais resultados estão descritos no Apêndice B.

Caracterizaremos as soluções de (3.1) em termos da exponencial da matriz A . Conforme seção B2 do apêndice B, a exponencial de $A \in M_d$ é definida por

$$\sum_{n=0}^{\infty} \frac{A^n}{n!} = Id + A + \frac{A^2}{2!} + \dots + \frac{A^n}{n!} + \dots$$

e decorre das propriedades (5) e (6) do Lema (A.5) os seguintes teoremas:

Teorema 3.1. *Dada uma matriz $A \in M_d$ e $x_0 \in \mathbb{R}^d$ então a curva $\gamma(t) = e^{tA}X_0$ define a única solução de $X' = AX$ com condição inicial $X(0) = X_0$.*

Teorema 3.2. *A única solução do sistema $X' = AX$ com condição inicial $X(t_0) = X_0$ é dada por*

$$X(t) = e^{A(t-t_0)}X_0$$

com $t \in \mathbb{R}$.

Demonstração. O Teorema (2.2) garante a existência e unicidade de soluções. Basta notar que

$$X(t_0) = e^{A(t_0-t_0)}X_0 = e^0X_0 = IdX_0 = X_0$$

e que

$$X'(t) = Ae^{A(t-t_0)}X_0 = AX(t).$$

□

Terminamos esta seção mostrando que a equação homogênea $X' = AX$ sujeita à uma perturbação $g(t)$ também pode ter suas soluções explicitamente calculadas.

Teorema 3.3. (*Fórmula da Variação das Constantes*) Seja $g : \mathbb{R} \rightarrow \mathbb{R}^d$ uma função contínua. Supondo $(t_0, X_0) \in \mathbb{R}^{1+d}$ como condição inicial, a solução maximal da equação $X' = AX + g(t)$ é dada por

$$X(t) = e^{A(t-t_0)} X_0 + \int_{t_0}^t e^{A(t-s)} g(s) ds, \quad \text{com } t \in I \subset \mathbb{R}. \quad (3.2)$$

Demonstração. Pelo Teorema (2.2) o problema admite solução e ela é única. Basta verificar que $X(t)$ definida acima é tal que $X(t_0) = X_0$ e que $X'(t) = AX(t) + g(t)$:

$$X(t_0) = e^{A \cdot 0} X_0 + \int_{t_0}^{t_0} e^{A(t-s)} g(s) ds = Id \cdot X_0 = X_0$$

e por derivação

$$\begin{aligned} X'(t) &= A e^{A(t-t_0)} X_0 + \int_{t_0}^t A e^{A(t-s)} g(s) ds + e^{A(t-t)} g(t) \\ &= A \left(e^{A(t-t_0)} X_0 + \int_{t_0}^t e^{A(t-s)} g(s) ds \right) + g(t), \end{aligned}$$

de onde $X'(t) = AX(t) + g(t)$. □

Exemplo 3.2. Considere a equação diferencial

$$X' = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} X + \begin{pmatrix} e^{-t} \\ 0 \end{pmatrix},$$

sujeita à condição $X(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Pela fórmula (3.2), a solução neste caso é dada por

$$X(t) = e^{At} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \int_0^t e^{As} \begin{pmatrix} 1 \\ 0 \end{pmatrix} ds. \quad (3.3)$$

Conforme Exemplo B.16 segue que,

$$\begin{aligned} X(t) &= \begin{pmatrix} 2e^{2t} - e^{3t} & -e^{2t} + e^{3t} \\ 2e^{2t} - 2e^{3t} & -e^{2t} + 2e^{3t} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &\quad + \int_0^t \begin{pmatrix} 2e^{2s} - e^{3s} & -e^{2s} + e^{3s} \\ 2e^{2s} - 2e^{3s} & -e^{2s} + 2e^{3s} \end{pmatrix} \begin{pmatrix} e^{-s} \\ 0 \end{pmatrix} ds. \end{aligned}$$

e calculando a integral da expressão anterior,

$$X(t) = \begin{pmatrix} 2e^{2t} - e^{3t} \\ 2e^{2t} - 2e^{3t} \end{pmatrix} + \left[\begin{pmatrix} 2e^t - \frac{1}{2}e^{2t} \\ 2e^t - e^{2t} \end{pmatrix} - \begin{pmatrix} \frac{3}{2} \\ 1 \end{pmatrix} \right],$$

de modo que

$$X(t) = \begin{pmatrix} -e^{3t} + \frac{3}{2}e^{2t} + 2e^t - \frac{3}{2} \\ -2e^{3t} + e^{2t} + 2e^t - 1 \end{pmatrix}, \quad t \geq 0.$$

3.1 Soluções e propriedades

Já vimos na seção anterior que as soluções de $X' = AX$ são dadas por uma exponencial matricial e nesta seção vamos descrever mais detalhadamente as possibilidades de e^{At} .

Começaremos mostrando que o conjunto \mathcal{S} das soluções de (3.1) possui estrutura de espaço vetorial, o que torna indispensável os métodos da Álgebra Linear no estudo de \mathcal{S} .

Lema 3.1. *Se X_1, X_2, \dots, X_d são soluções da equação (3.1), então $c_1X_1 + c_2X_2 + \dots + c_dX_d$ é também solução, onde c_1, c_2, \dots, c_d são escalares; ou seja, a combinação linear de soluções da equação $X' = AX$ é ainda uma solução desta equação.*

Demonstração. Como por hipótese cada X_i , $i = 1, 2, \dots, d$ é solução de (3.1), segue em particular que

$$X'_1 = AX_1, \quad X'_2 = AX_2, \quad \dots, \quad X'_d = AX_d$$

e portanto, tomando uma combinação linear das funções X_i , $i = 1, 2, \dots, d$ segue que

$$\begin{aligned} (c_1X_1 + c_2X_2 + \dots + c_dX_d)' &= c_1X'_1 + c_2X'_2 + \dots + c_dX'_d = \\ c_1AX_1 + c_2AX_2 + \dots + c_dAX_d &= A(c_1X_1 + c_2X_2 + \dots + c_dX_d) \end{aligned}$$

o que prova o resultado. □

Proposição 3.1. *Considere $A \in M_d$. O conjunto \mathcal{S} das soluções da equação $X' = AX$ é um espaço vetorial de dimensão d .*

Demonstração. Para mostrar que \mathcal{S} tem estrutura de espaço vetorial, basta verificar diretamente que os axiomas que definem um espaço vetorial (veja apêndice B) são satisfeitos para quaisquer $X_1, X_2 \in \mathcal{S}$ e para quaisquer escalares $\alpha, \beta \in K$.

Mostraremos agora que $\dim \mathcal{S} = d$, ou seja, vamos exibir uma base de \mathcal{S} com d elementos. Para isto, vamos considerar as funções

$$\phi_j(t), \quad 1 \leq j \leq d \tag{3.4}$$

soluções de $X' = AX$ sujeitas à condição inicial $X(0) = e_j$, $1 \leq j \leq d$ que estão bem determinadas de acordo com o Teorema (2.2).

As d funções definidas por (3.4) são linearmente independentes; de fato, considere a combinação linear

$$c_1\phi_1 + c_2\phi_2 + \dots + c_d\phi_d = 0 \tag{3.5}$$

onde c_1, c_2, \dots, c_d são constantes e 0 indica o vetor nulo em \mathbb{R}^d . Fazendo $t = 0$ em (3.5), encontramos

$$c_1e_1 + c_2e_2 + \dots + c_de_d = 0$$

de onde $c_1 = c_2 = \dots = c_d = 0$.

Provemos agora que o conjunto $\mathcal{B} = \{\phi_j(t), 1 \leq j \leq d\}$ gera \mathcal{S} . Para isto, provemos que todo vetor em \mathcal{S} (isto é, toda solução $X(t)$ da equação $X' = AX$) pode ser escrito como combinação linear de $\phi_1, \phi_2, \dots, \phi_d$.

Seja X um elemento de \mathcal{S} , sujeito à condição $X(0) = c = (c_1, c_2, \dots, c_d) \in \mathbb{R}^d$ e tomando as constantes c_1, c_2, \dots, c_d construamos a função

$$\phi(t) = c_1\phi_1(t) + c_2\phi_2(t) + \dots + c_d\phi_d(t). \quad (3.6)$$

que satisfaz a equação $X' = AX$ por ser combinação linear das soluções $\phi_1, \phi_2, \dots, \phi_d$, conforme Lema 3.1. Além disso, $\phi(0) = c = X(0)$, pois

$$\begin{aligned} \phi(0) &= c_1\phi_1(0) + c_2\phi_2(0) + \dots + c_d\phi_d(0) = \\ &= c_1e_1 + c_2e_2 + \dots + c_de_d = (c_1, c_2, \dots, c_d) = c = X(0). \end{aligned}$$

Ora, pelo Teorema (2.2), as funções $\phi(t)$ e $X(t)$ devem ser idênticas, ou seja, $X(t)$ é combinação linear de $\phi_1, \phi_2, \dots, \phi_d$. \square

De acordo com a Proposição (3.1), para resolver completamente $X' = AX$ basta que se determine um conjunto com d soluções linearmente independentes, chamado de **um conjunto fundamental de soluções** ou em outras palavras, uma **base do espaço \mathcal{S}** .

Considere X_1, X_2, \dots, X_d um conjunto de d soluções de $X' = AX$ e considere a matriz $W(t)$ cujas colunas são os vetores X_1, X_2, \dots, X_d . Segue da observação (B.1) que os vetores coluna de $X(t)$ são linearmente independentes para um dado valor de t se $\det W \neq 0$ para esse valor de t . Esse determinante é chamado de **wronskiano** das d soluções X_1, X_2, \dots, X_d e será denotado por $W[X_1, X_2, \dots, X_d]$.

Lema 3.2. *Se X_1, X_2, \dots, X_d são soluções de $X' = AX$ no intervalo $\alpha < t < \beta$ então o wronskiano $W[X_1, X_2, \dots, X_d]$ ou é identicamente nulo ou nunca se anula neste intervalo.*

Demonstração. A demonstração pode ser encontrada em [4]. \square

O Lema (3.2) acima é importante pois indica que, para analisar a independência linear de d soluções definidas num intervalo aberto qualquer, basta escolher um ponto conveniente dentro deste intervalo.

Lema 3.3. *Considere a equação $X' = AX$ onde $A \in M_d$. Seja $v \in \mathbb{R}^d$ um autovetor de A associado ao autovalor $\lambda \in \mathbb{R}$. Então, dado t real qualquer,*

$$X(t) = e^{\lambda t}v$$

é solução de $X' = AX$ com $X(0) = v$.

Demonstração. Sendo v autovetor de A associado ao autovalor λ , segue que $Av = \lambda v$.

Derivando $X(t) = e^{\lambda t}v$ em relação a t temos

$$X'(t) = \lambda e^{\lambda t}v,$$

$$X'(t) = e^{\lambda t}\lambda v,$$

$$X'(t) = e^{\lambda t}Av,$$

$$X'(t) = Ae^{\lambda t}v,$$

$$X'(t) = AX(t).$$

o que conclui a prova. □

Exemplo 3.3. Considere a equação diferencial

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = \begin{pmatrix} 0 & -2 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \quad (3.7)$$

O polinômio característico de $\begin{pmatrix} 0 & -2 \\ 3 & 5 \end{pmatrix}$ é $p(\lambda) = \lambda^2 - 5\lambda + 6$ cujas raízes são $\lambda_1 = 2$ e $\lambda_2 = 3$.

Para $\lambda_1 = 2$ encontramos o autovetor $v_1 = (1, -1)$, pois

$$\begin{pmatrix} 0 & -2 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$$

tem solução geral $(x, -x)$ e podemos considerar $v_1 = (1, -1)$.

De modo análogo, para $\lambda_1 = 3$ encontramos $v_2 = \left(1, -\frac{3}{2}\right)$.

Assim, pelo Lema 3.3 segue que

$$X_1 = e^{2t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} e^{2t} \\ -e^{2t} \end{pmatrix}$$

e

$$X_2 = e^{3t} \begin{pmatrix} 1 \\ -\frac{3}{2} \end{pmatrix} = \begin{pmatrix} e^{3t} \\ -\frac{3}{2}e^{3t} \end{pmatrix}$$

são soluções de (3.7).

Para verificar que X_1 e X_2 formam um conjunto fundamental de soluções definido para todo $t \in \mathbb{R}$, escolhemos $t = 0$, de modo que

$$W[X_1, X_2](0) = \det \begin{pmatrix} 1 & 1 \\ 1 & -\frac{3}{2} \end{pmatrix} \neq 0$$

e assim, a solução geral de (3.7) é dada por

$$X(t) = c_1 X_1 + c_2 X_2,$$

ou explicitamente por

$$X(t) = c_1 \begin{pmatrix} e^{2t} \\ -e^{2t} \end{pmatrix} + c_2 \begin{pmatrix} e^{3t} \\ -\frac{3}{2}e^{3t} \end{pmatrix}.$$

Usando o Lema 3.3, vemos que as soluções de $X' = AX$ são determinadas pelos autovalores e autovetores da matriz real A . Temos portanto três casos a analisar:

1. Todos os autovalores reais e distintos entre si;
2. Alguns autovalores complexos;
3. Alguns autovalores repetidos.

Passaremos agora à análise de cada um destes casos. As referências utilizadas foram [4], [3] e [9].

3.1.1 Autovalores reais e distintos

Se todos os d autovalores da matriz $A \in M_d$ são reais e distintos, existem exatamente d autovetores distintos.

Sejam $\lambda_1, \lambda_2, \dots, \lambda_d$ tais autovalores associados aos autovetores v_1, v_2, \dots, v_d respectivamente. Pelo Lema 3.3 cada função

$$X_i(t) = e^{\lambda_i t} v_i, \quad 1 \leq i \leq d \quad (3.8)$$

satisfaz à equação diferencial $X' = AX$, de modo que temos um candidato natural ao conjunto fundamental de soluções: o conjunto das d funções (3.8).

Basta verificar se tais funções são linearmente independentes, ou seja, se o wronskiano

$$W[X_1, X_2, \dots, X_d] = \det \begin{pmatrix} v_{11}e^{\lambda_1 t} & \dots & v_{1d}e^{\lambda_1 t} \\ \dots & \ddots & \dots \\ v_{d1}e^{\lambda_d t} & \dots & v_{dd}e^{\lambda_d t} \end{pmatrix} = e^{(\lambda_1 + \lambda_2 + \dots + \lambda_d)t} \det \begin{pmatrix} v_{11} & \dots & v_{1d} \\ \dots & \ddots & \dots \\ v_{d1} & \dots & v_{dd} \end{pmatrix}$$

é não nulo, o que é de fácil verificação pois a exponencial nunca se anula e o determinante do segundo membro da igualdade seguramente também não, uma vez que suas colunas são os autovetores de A , sempre linearmente independentes (veja observação (B.1)).

Assim, a solução geral de $X' = AX$, $A \in M_d$ no caso de autovalores reais e distintos é dada por

$$X(t) = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2 + \dots + c_d e^{\lambda_d t} v_d, \quad (3.9)$$

onde λ_i e v_i para $1 \leq i \leq d$ são os autovalores e autovetores da matriz A , respectivamente.

3.1.2 Autovalores complexos

Considere novamente a equação $X' = AX$ onde $A \in M_d$ é uma matriz com entradas reais e suponha que o complexo $\lambda = a + ib$ seja um autovalor de A , com $b \neq 0$ associado ao autovetor $w = u + iv$.

Para a discussão a seguir, usaremos os lemas abaixo:

Lema 3.4. *Seja $w \in \mathbb{C}^d$ e \bar{w} seu conjugado. O conjunto $\{w, \bar{w}\}$ é linearmente independente.*

Demonstração. Pode ser encontrada em [4]. □

Lema 3.5. *Sejam $A \in M_d$ uma matriz real e $w \in \mathbb{C}^d$ um autovalor complexo de A associado ao autovalor $a + ib \in \mathbb{C}$ com $b \neq 0$. Escrevendo $w = u + iv$ com $u, v \in \mathbb{R}^d$ dados por*

$$u = \frac{1}{2}(w + \bar{w})$$

e

$$v = \frac{1}{2i}(w - \bar{w})$$

segue que u, v são linearmente independentes em \mathbb{R}^d e mais:

$$Au = au - bv \tag{3.10}$$

e

$$Av = bu + av. \tag{3.11}$$

Demonstração. Seja w autovetor complexo de A e sejam $u, v \in \mathbb{R}^d$ tais que $w = u + iv$ como nas equações (3.10) e (3.11).

Suponhamos que $\{u, v\}$ seja linearmente dependente em \mathbb{R}^d , isto é, existe $\alpha \in \mathbb{R}$ tal que $v = \alpha u$. Então

$$w - \bar{w} = 2iv = i\alpha(w + \bar{w})$$

e portanto $(1 - i\alpha)w = (1 + i\alpha)\bar{w}$. Como $(1 - i\alpha)$ e $(1 + i\alpha)$ são não nulos, segue que $\{w, \bar{w}\}$ é linearmente dependente em \mathbb{C}^d , o que contraria o Lema 3.4. Logo, $\{u, v\}$ é linearmente independente.

Por outro lado, seja $z = a + ib$, com $b \neq 0$ o autovalor associado a w . As equações (3.10) e (3.11) seguem da igualdade

$$Au + iAv = A(u + iv) = Aw = zw = (a + ib)(u + iv) = (au - bv) + i(bu + av).$$

□

Sendo A uma matriz real, então os coeficientes de seu polinômio característico são reais e decorre que os autovalores complexos devem aparecer aos pares conjugados. Assim, sendo λ autovalor, segue que $\bar{\lambda}$ também é e mais que isso: os autovetores

também são conjugados. De fato, admitindo que λ é autovalor associado ao autovetor w , decorre que

$$(A - \lambda Id_d)w = 0$$

e então

$$\overline{(A - \lambda Id_d)w} = (A - \bar{\lambda} Id_d)\bar{w} = 0$$

Agora, observe que o Lema (3.3) permanece válido no caso complexo, uma vez para demonstrá-lo não usamos a hipótese de o autovalor ser real. Assim, são soluções de $X' = AX$:

$$X(t) = (u + iv)e^{(a+ib)t} \quad (3.12)$$

e

$$X(t) = \overline{u + iv}e^{\overline{(a+ib)t}}. \quad (3.13)$$

Usando a exponencial de Euler ¹, de (3.12) podemos escrever

$$X(t) = (u + iv)e^{at}e^{ibt},$$

$$X(t) = e^{at}(u + iv)(\cos bt + i \operatorname{sen} bt),$$

$$X(t) = e^{at}[u \cos bt + iu \operatorname{sen} bt + iv \cos bt - v \operatorname{sen} bt]$$

e então

$$X(t) = e^{at}[u \cos bt - v \operatorname{sen} bt] + ie^{at}[u \operatorname{sen} bt + v \cos bt].$$

Agora, escrevendo $X(t) = X_1(t) + iX_2(t)$, segue que

$$X_1(t) = e^{at}[u \cos bt - v \operatorname{sen} bt] \quad (3.14)$$

e

$$X_2(t) = e^{at}[u \operatorname{sen} bt + v \cos bt], \quad (3.15)$$

são funções reais que satisfazem $X' = AX$. Para provarmos isso, lembremos que pelo Lema (3.5) temos $Au = au - bv$ e $Av = bu + av$ e então, derivando (3.14) segue que

$$X_1'(t) = te^{at}[u \operatorname{sen} bt + v \cos bt] - e^{at}[vt \operatorname{sen} bt - bu \cos bt],$$

$$X_1'(t) = e^{at}(au - bv) \cos bt - e^{at}(bu + av) \operatorname{sen} bt,$$

de onde

$$X_1'(t) = A(e^{at}[u \cos bt - v \operatorname{sen} bt]) = AX_1(t).$$

De modo análogo se prova que (3.25) também é solução de $X' = AX$. Além disso, é possível provar que $X_1(t)$ e $X_2(t)$ sempre são linearmente independentes

Observe que se, ao invés de (3.12) tivéssemos tomado (3.13) o resultado se manteria e a discussão feita seria a mesma, de modo que basta escolher um autovalor complexo.

¹Prova-se em Análise Complexa que a exponencial natural de um número imaginário é igual à uma soma de funções trigonométricas; especificadamente vale que $e^{i\theta} = \cos \theta + i \operatorname{sen} \theta$.

Assim, se a matriz A admite o complexo $\lambda_1 = a + ib$ (e portanto $\lambda_2 = a - ib$) e os reais $\lambda_3, \lambda_4, \dots, \lambda_d$ reais como autovalores associados aos autovetores (3.14), (3.25), v_3, v_4, \dots, v_d a solução geral de $X' = AX$ se escreve como

$$X(t) = c_1 X_1(t) + c_2 X_2(t) + c_3 e^{\lambda_3 t} v_3 + \dots + c_d e^{\lambda_d t} v_d.$$

Exemplo 3.4. Vamos encontrar a solução geral da equação

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = \begin{pmatrix} -\frac{1}{2} & 1 \\ -1 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \quad (3.16)$$

O polinômio característico é $p(\lambda) = \lambda^2 + \lambda + \frac{5}{4}$ de onde os autovalores generalizados são $\lambda_1 = -\frac{1}{2} + i$ e $\lambda_2 = -\frac{1}{2} - i$. Escolhemos $\lambda_1 = -\frac{1}{2} + i$ de onde

$$\begin{pmatrix} -\frac{1}{2} & 1 \\ -1 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2}x + ix \\ -\frac{1}{2}x - ix \end{pmatrix}$$

e encontramos o autovetor generalizado $w = 1 + i$.

Logo, a solução de (3.16) é

$$X(t) = \begin{pmatrix} 1 \\ i \end{pmatrix} e^{(-\frac{1}{2}+i)t} \quad (3.17)$$

e portanto

$$X(t) = \begin{pmatrix} 1 \\ i \end{pmatrix} e^{(-\frac{1}{2}t)} e^{it},$$

$$X(t) = \begin{pmatrix} 1 \\ i \end{pmatrix} e^{(-\frac{1}{2}t)} (\cos t + i \operatorname{sen} t),$$

$$X(t) = \begin{pmatrix} 1 \\ i \end{pmatrix} e^{(-\frac{1}{2}t)} (\cos t + i \operatorname{sen} t),$$

$$X(t) = \begin{pmatrix} e^{-\frac{1}{2}t} \cos t \\ -e^{-\frac{1}{2}t} \operatorname{sen} t \end{pmatrix} + i \begin{pmatrix} e^{-\frac{1}{2}t} \operatorname{sen} t \\ e^{-\frac{1}{2}t} \cos t \end{pmatrix}$$

de modo que as soluções reais são

$$X_1(t) = \begin{pmatrix} 1 \\ e^{-\frac{1}{2}t} \cos t \\ -e^{-\frac{1}{2}t} \operatorname{sen} t \end{pmatrix}$$

e

$$X_2(t) = \begin{pmatrix} e^{-\frac{1}{2}t} \operatorname{sen} t \\ e^{-\frac{1}{2}t} \cos t \end{pmatrix}.$$

Assim, a solução geral de (3.16) é dada por

$$X(t) = c_1 \begin{pmatrix} e^{-\frac{1}{2}t} \cos t \\ -e^{-\frac{1}{2}t} \operatorname{sen} t \end{pmatrix} + c_2 \begin{pmatrix} e^{-\frac{1}{2}t} \operatorname{sen} t \\ e^{-\frac{1}{2}t} \cos t \end{pmatrix},$$

com c_1, c_2 constantes quaisquer.

3.1.3 Autovalores repetidos

Finalmente, vamos discutir o caso onde a matriz $A \in M_d$ apresenta autovalores repetidos. Este caso apresenta dificuldades maiores que os dois anteriores e que são contornadas com uma reformulação da definição de autovalor. Começamos com algumas definições e exemplos.

Definição 3.1. *Seja λ um autovalor de A com multiplicidade algébrica k . Se associado a λ existirem k autovetores linearmente independentes, então λ é um **autovalor completo**; se associado a λ existirem $p < k$ autovetores linearmente independentes, então λ é um **autovalor deficiente com grau de deficiência $d = k - p$** .*

Exemplo 3.5. Considere a matriz

$$A = \begin{pmatrix} 9 & 4 & 0 \\ -6 & -1 & 0 \\ 6 & 4 & 3 \end{pmatrix}$$

cujos polinômio característico é

$$p(\lambda) = (\lambda - 3)^2(\lambda - 5)$$

e cujos autovalores são $\lambda_1 = 3$ com multiplicidade 2 e $\lambda_2 = 5$ com multiplicidade 1. O autovalor λ_1 é completo; de fato, resolvendo

$$\begin{pmatrix} 9 & 4 & 0 \\ -6 & -1 & 0 \\ 6 & 4 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3x \\ 3y \\ 3z \end{pmatrix}$$

encontramos que a solução é dada por qualquer vetor $v = \left(x, -\frac{3}{2}x, z\right)$ de onde

$$v = x \left(1, -\frac{3}{2}, 0\right) + z(0, 0, 1)$$

de modo que existem dois autovalores linearmente independentes associados a λ_1 .

No caso de λ ser autovalor completo, o problema de achar um conjunto fundamental de soluções está resolvido, pois sempre é possível conseguir d vetores linearmente independentes e conseqüentemente d soluções linearmente independentes. A dificuldade aparece quando λ é deficiente, de modo que encontramos $p < d$ vetores independentes e portanto não é imediata a obtenção de um conjunto fundamental de soluções.

Proposição 3.2. *Seja λ um autovalor de $A \in M_d$ com multiplicidade algébrica k associado ao autovalor v_1 ; então os vetores v_1, v_2, \dots, v_k tais que*

$$(A - \lambda Id_d)v_2 = v_1,$$

$$\begin{aligned}(A - \lambda Id_d)v_3 &= v_2, \\ &\vdots \\ (A - \lambda Id_d)v_k &= v_{k-1}\end{aligned}$$

são linearmente independentes.

Demonstração. Vamos provar que a equação

$$a_1v_1 + a_2v_2 + \dots + a_kv_k = 0 \quad (3.18)$$

implica em $a_1 = a_2 = \dots = a_k = 0$.

Para isso, vamos multiplicar (3.18) por $(A - \lambda Id_d)$, de modo que

$$a_1(A - \lambda Id_d)v_1 + a_2(A - \lambda Id_d)v_2 + \dots + a_k(A - \lambda Id_d)v_k = 0$$

e como por hipótese $(A - \lambda Id_d)v_1 = 0$ e $(A - \lambda Id_d)v_k = v_{k-1}$ segue da equação anterior que

$$a_2v_1 + a_3v_2 + \dots + a_kv_{k-1} = 0. \quad (3.19)$$

Repetindo o argumento acima, ou seja, multiplicando (3.19) por $(A - \lambda Id_d)$ encontramos

$$a_3v_1 + a_4v_2 + \dots + a_kv_{k-2} = 0,$$

de modo que depois de sucessivas repetições deste mesmo argumento temos

$$a_kv_1 = 0,$$

de onde $a_k = 0$.

Como $a_k = 0$, a equação (3.18) se reescreve como

$$a_1v_1 + a_2v_2 + \dots + a_{k-1}v_{k-1} = 0$$

e o mesmo argumento mostra que $a_{k-1} = 0$.

Prosseguindo deste modo se mostra que todas as constantes a_i são nulas, o que prova o resultado. \square

A sequência de vetores linearmente independentes da proposição acima é importante para nosso estudo e justifica a seguinte definição:

Definição 3.2. *Seja λ um autovalor de $A \in M_d$ com multiplicidade algébrica k . Uma k -cadeia de autovetores generalizados associados a λ é uma sequência de k vetores linearmente independentes v_1, v_2, \dots, v_k tais que*

$$\begin{aligned}(A - \lambda Id_d)v_2 &= v_1, \\ (A - \lambda Id_d)v_3 &= v_2, \\ &\vdots \\ (A - \lambda Id_d)v_k &= v_{k-1}.\end{aligned}$$

Observação 3.1. Os k vetores generalizados satisfazem a equação

$$(A - \lambda Id_d)^k v_k = 0.$$

Para ver isto, note que $(A - \lambda Id_d)v_1 = 0$ e portanto da equação $(A - \lambda Id_d)v_2 = v_1$ temos

$$(A - \lambda Id_d)(A - \lambda Id_d)v_2 = (A - \lambda Id_d)v_1 \Rightarrow (A - \lambda Id_d)^2 v_2 = 0,$$

sendo a prova do caso geral feita de modo análogo.

Portanto, a ideia a ser utilizada para resolver o sistema $X' = AX$ no caso em que a matriz A apresenta autovalores repetidos é a de que seja encontrada uma k -cadeia de autovetores generalizados afim de que consigamos obter um sistema fundamental de soluções.

Exemplo 3.6. Suponha que a equação diferencial $X' = AX$ com $A \in M_3$ possua polinômio característico

$$p(\lambda) = (\lambda - a)^3,$$

com autovetor λ de multiplicidade algébrica 3.

Se associado a λ existir apenas um autovetor v_1 , ou seja, se a deficiência de λ for igual a 2, uma das soluções será

$$X_1 = v_1 e^{\lambda t},$$

de modo que ainda é necessário encontrar duas soluções linearmente independentes para escrever um conjunto fundamental de soluções. Essas duas soluções serão obtidas através dos autovetores generalizados de v_1 , ou seja, a solução geral será dada por

$$X(t) = c_1 v_1 e^{\lambda t} + c_2 (tv_1 + v_2) e^{\lambda t} + c_3 \left(\frac{t^2}{2!} v_1 + tv_2 + v_3 \right) e^{\lambda t},$$

onde os coeficientes da combinação linear entre v_1 , v_2 e v_3 são os mesmos do desenvolvimento da exponencial e^t em série de potências (a justificativa para este fato está no fato de que as colunas da exponencial da matriz A formam um conjunto fundamental de soluções para $X' = AX$, mas não a apresentaremos neste trabalho).

Por outro lado, se a λ existirem associados dois autovetores v_1 e v_2 , ou seja, se a deficiência de λ for igual a 1, duas das soluções serão dadas (conforme Lema (3.3)) por

$$X_1 = v_1 e^{\lambda t}$$

e

$$X_2 = v_2 e^{\lambda t},$$

de modo que a solução geral será dada por

$$X(t) = c_1 v_1 e^{\lambda t} + c_2 v_2 e^{\lambda t} + c_3 (tv_1 + v_2) e^{\lambda t}.$$

Exemplo 3.7. Vamos resolver a equação diferencial

$$X' = \begin{pmatrix} 0 & 1 & 2 \\ -5 & -3 & -7 \\ 1 & 0 & 0 \end{pmatrix} X, \quad (3.20)$$

cuja matriz $A = \begin{pmatrix} 0 & 1 & 2 \\ -5 & -3 & -7 \\ 1 & 0 & 0 \end{pmatrix}$ tem polinômio característico

$$p(\lambda) = (\lambda + 1)^3$$

e autovalor $\lambda = -1$ associado ao autovetor $v_1 = (1, 1, -1)$. Como a deficiência de v_1 é igual a 2, é necessário determinar dois autovetores generalizados de v_1 , que podemos obter através das equações da Observação 3.1, ou seja, impondo que

$$(A + Id_3)^2 v_2 = \begin{pmatrix} 1 & 1 & 2 \\ -5 & -2 & -7 \\ 1 & 0 & 1 \end{pmatrix}^2 \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -2 & -1 & -3 \\ -2 & -1 & -3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (3.21)$$

e

$$(A + Id_3)^3 v_3 = \begin{pmatrix} 1 & 1 & 2 \\ -5 & -2 & -7 \\ 1 & 0 & 1 \end{pmatrix}^3 \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

ou seja, o vetor v_3 é qualquer. Tomando $v_3 = (1, 0, 0)$ e calculando v_2 de (3.21) encontramos $v_2 = (1, -5, 1)$ e a solução geral de (3.20) é

$$X(t) = c_1 \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} e^{-t} + c_2 \left[c_1 \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} t + \begin{pmatrix} 1 \\ -5 \\ 1 \end{pmatrix} \right] e^{-t} + c_3 \left[\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \frac{t^2}{2} + \begin{pmatrix} 1 \\ -5 \\ 1 \end{pmatrix} t + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right] e^{-t}.$$

3.2 Sistemas planares: retratos de fase

Consideraremos nesta seção apenas as equações $X' = AX$ onde A é uma matriz quadrada de ordem 2. Vamos descrever completamente suas soluções e os retratos de fase, que dependem da análise dos autovalores de A , conforme visto na seção anterior. Por simplicidade de exposição, dividiremos o estudo a ser feito em três casos, conforme a matriz A possua autovalores distintos, autovalores complexos e conjugados ou autovalores iguais.

3.2.1 Autovalores reais e distintos

Suponha que a matriz A do sistema $X' = AX$ possua dois autovalores reais e distintos. Suponha que seus autovalores sejam λ_1 e λ_2 associados aos autovetores v_1 e v_2 respectivamente. Então, por (3.9) a solução geral será dada por

$$X(t) = c_1 v_1 e^{\lambda_1 t} + c_2 v_2 e^{\lambda_2 t}. \quad (3.22)$$

Consideremos agora vários casos conforme os sinais dos autovalores λ_1 e λ_2 :

1. Considere $\lambda_1 < \lambda_2 < 0$. Quando $c_2 \neq 0$, de (3.22) obtemos

$$\frac{X(t)}{c_2 e^{\lambda_2 t}} = \frac{c_1 v_1}{c_2} v_2 e^{(\lambda_1 - \lambda_2)t} + v_2 \quad (3.23)$$

e se $t \rightarrow \infty$ então $\frac{X(t)}{c_2 e^{\lambda_2 t}} \rightarrow v_2$ e portanto a solução $X(t)$ tem assintoticamente a direção de v_2 como no retrato de fase abaixo, onde a origem é chamada de **nó estável**

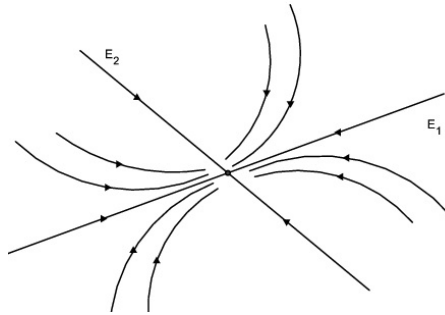


Figura 3.1: Nó Estável: $\lambda_1 < \lambda_2 < 0$.

2. Considere agora $\lambda_1 > \lambda_2 > 0$. De modo análogo, se $c_2 \neq 0$ temos $\frac{X(t)}{c_2 e^{\lambda_2 t}} \rightarrow v_2$ se $t \rightarrow -\infty$. Neste caso temos o seguinte retrato de fase e a origem é chamada de **nó instável**.

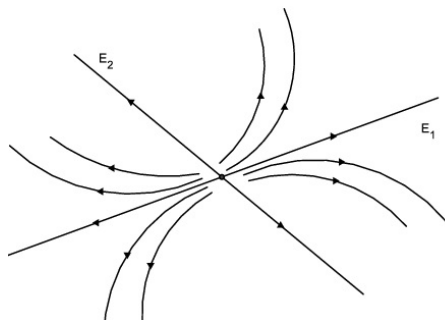


Figura 3.2: Nó Instável: $\lambda_1 > \lambda_2 > 0$.

3. Se $\lambda_1 < 0 < \lambda_2$ então há uma expansão ao longo da direção de v_2 e uma contração na direção de v_1 . O retrato de fase é o indicado na figura abaixo e recebe o nome de **ponto de sela**

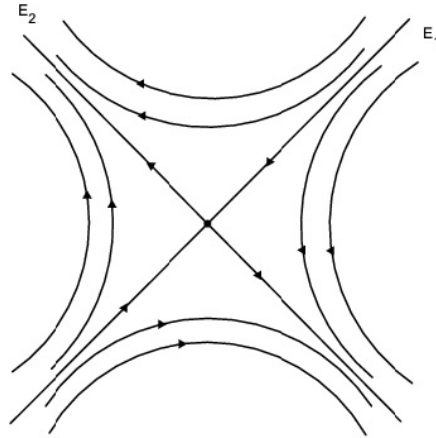


Figura 3.3: Ponto de Sela: $\lambda_1 < 0 < \lambda_2$.

3.2.2 Autovalores repetidos

Caso a matriz A apresente apenas o autovalor λ associado v_1 , por analogia com o Exemplo (3.22) a solução geral de $X' = AX$ se escreve como

$$X(t) = c_1 v_1 e^{\lambda t} + c_2 (v_1 + t v_2) e^{\lambda t}, \quad (3.24)$$

onde v_2 é um vetor generalizado de v_1 .

1. Se $\lambda_1 = \lambda_2 = \lambda < 0$ então a solução (3.24) tende à origem a partir de qualquer direção. Neste caso, a origem também é chamada de **nó estável**; o retrato de fase pode ser visto na Figura (3.4) abaixo.

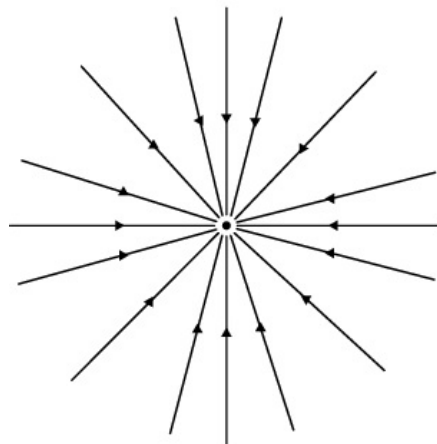


Figura 3.4: Nó Estável: $\lambda_1 = \lambda_2 = \lambda < 0$.

2. Se $\lambda_1 = \lambda_2 = \lambda > 0$ então a solução (3.24) afasta-se da origem em todas as direções, de modo que ela é chamada de **nó instável**; veja a Figura (3.5).

O comportamento descrito na Figura 3.4 tem caráter local; globalmente as trajetórias podem ter o comportamento descrito no exemplo a seguir.

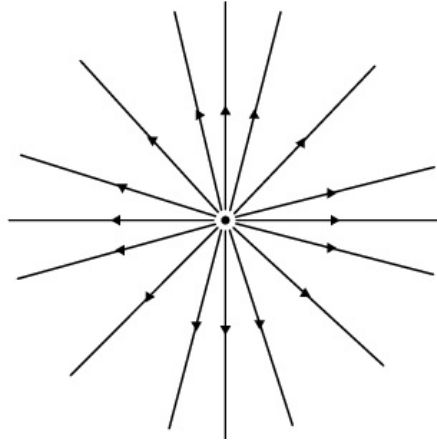


Figura 3.5: Nó Estável: $\lambda_1 = \lambda_2 = \lambda > 0$.

Exemplo 3.8. Considere o sistema de equações

$$\begin{cases} x' = 2x + y \\ y' = -2y \end{cases}$$

cujas soluções são $y(t) = ae^{2t}$ e $x(t) = ate^{2t} + be^{2t}$ com $a, b \in \mathbb{R}$. Como $(x(t), y(t)) \rightarrow (0, 0)$ quando $t \rightarrow -\infty$ e $(x(t), y(t)) \rightarrow +\infty$ quando $t \rightarrow +\infty$ as trajetórias apresentam o comportamento grafado abaixo (e que localmente é dado pelas Figuras 3.4 e 3.5).

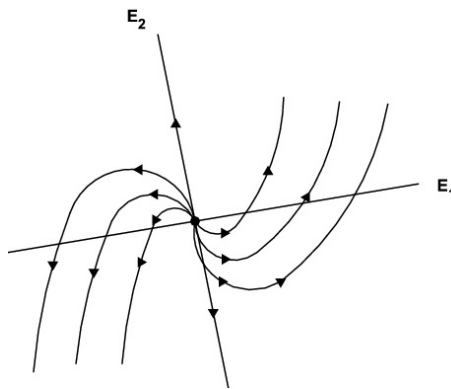


Figura 3.6: Autovalores repetidos: multiplicidade geométrica menor que a algébrica.

3.2.3 Autovalores complexos e conjugados

Considere agora que a equação $X' = AX$ onde $A \in M_d$ é uma matriz com entradas reais tal que o complexo $\lambda = a + ib$ seja autovalor (e consequentemente, também o complexo $\bar{\lambda} = a - ib$) com $b \neq 0$ associado ao autovetor $w = u + iv$ que são dadas por

$$X(t) = e^{at}[u \sin bt + v \cos bt]. \quad (3.25)$$

Temos três casos a considerar:

1. Se $a = 0$ a solução se escreve como $X(t) = u \sin(bt) + v \cos bt$ de onde o retrato de fase é uma elipse ou então àquele que se obtém invertendo o sentido do movimento. Dizemos que a origem é um **centro**. Observe que neste caso, as soluções são periódicas.

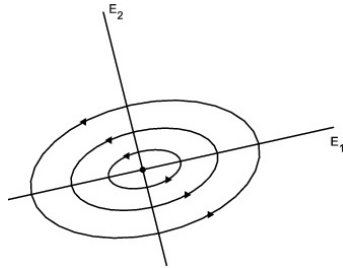


Figura 3.7: Centro: $a = 0$.

2. Se $a > 0$ o retrato de fase é o indicado na Figura 3.7 e a origem é chamada de **foco instável**.

Neste caso, qualquer solução se afasta da origem.

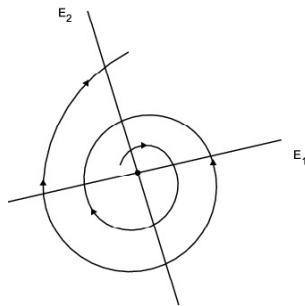


Figura 3.8: Foco Instável: $a > 0$.

3. Se $a < 0$ o retrato de fase toma a configuração indicada na Figura 3.8 e a origem é chamada de **foco estável**.

Neste caso, qualquer solução se aproxima da origem.

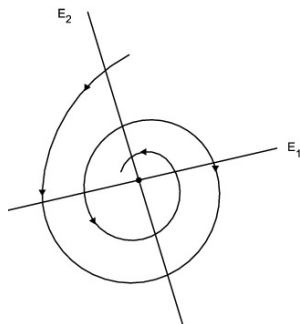


Figura 3.9: Foco Estável: $a < 0$.

4 Introdução à estabilidade

Faremos neste capítulo uma introdução ao estudo da estabilidade das singularidades de uma equação diferencial. Trata-se de investigar o comportamento das trajetórias no espaço de fase nas vizinhanças destes pontos. Apresentaremos critérios que permitem afirmar a existência ou não de estabilidade para as equações lineares através da análise dos autovalores associados à equação e a estabilidade segundo Lyapunov que permite estudar a estabilidade sem conhecimento algum dos autovalores do sistema. Este capítulo foi desenvolvido utilizando as referências [3], [9] [5] e [10].

Dada uma função contínua $F : D \rightarrow \mathbb{R}^d$ num aberto $D \subset \mathbb{R} \times \mathbb{R}^d$ consideraremos a equação $x' = F(t, x)$ e no restante deste capítulo vamos supor que existe uma única solução $\gamma(t)$ da equação $x' = F(t, x)$ para cada condição inicial $(t_0, x_0) \in D$.

4.1 Definições e exemplos

Apresentamos na sequência a definição de estabilidade para a solução de uma equação diferencial. Na sequência, ilustramos com alguns exemplos este importante conceito.

Definição 4.1. *Uma solução $\gamma(t, t_0, x_0)$ da equação $x' = F(t, x)$ é chamada de **estável** se dado $\epsilon > 0$ existe $\delta > 0$ tal que se $\|x_0 - \bar{x}_0\| < \delta$ com $(t_0, \bar{x}_0) \in D$, então:*

1. *A solução $\gamma(t, t_0, x_0)$ da equação $x' = F(t, x)$ está definida para $t > t_0$;*
2. *$\|\gamma(t, t_0, x_0) - \gamma(t, t_0, \bar{x}_0)\| < \epsilon$ para $t > t_0$*

Caso uma solução não seja estável, dizemos que ela é **instável**.

Se a solução $\gamma(t, t_0, x_0)$ é estável e existe $\alpha > 0$ tal que se $\|x_0 - \bar{x}_0\| < \alpha$ então $\|\gamma(t, t_0, x_0) - \gamma(t, t_0, \bar{x}_0)\| \rightarrow 0$ quando $t \rightarrow \infty$ então a solução será chamada de **assintoticamente estável**.

A figura abaixo ilustra o conceito de estabilidade.

Para analisá-la, considere $\epsilon > 0$. Se a solução é estável é sempre possível encontrar um $\delta > 0$ tal que $\|x_0 - \bar{x}_0\| < \delta$ implica $\|\gamma(t) - \bar{\gamma}(t)\| < \epsilon$ por menor que seja $\epsilon > 0$. Em outras palavras, pequenos desvios na condição inicial não afetam significativamente o comportamento futuro da solução.

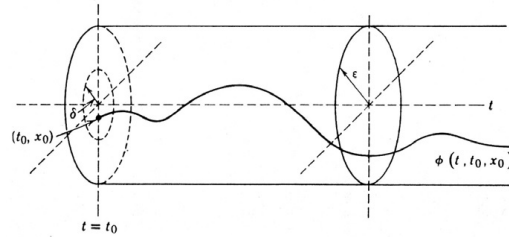


Figura 4.1: Solução $\gamma(t, t_0, x_0)$ estável

Exemplo 4.1. Considere a equação diferencial autônoma

$$\frac{dx}{dt} = x(x - 1), \quad (4.1)$$

onde $F(x) = x(x - 1)$.

Por definição, os valores \bar{x} tais que $F(\bar{x}) = 0$ são chamados de **pontos de equilíbrio ou singularidades** da equação diferencial (ou do campo de vetores) e as curvas constantes $\gamma(t) = \bar{x}$ são chamadas de **soluções de equilíbrio ou estacionárias**.

Observando a equação (4.1), vemos que as singularidades são $\bar{x}_1 = 0$ e $\bar{x}_2 = 1$ e portanto as soluções constantes são $\gamma(t) \equiv 0$ e $\gamma(t) \equiv 1$ para todo $t \in \mathbb{R}$. Note que para $x \in (0, 1)$ a expressão no segundo membro de (4.1) assume valores negativos, de modo que, neste intervalo as soluções devem ser funções contínuas decrescentes; de maneira análoga, nos intervalos $(-\infty, 0)$ e $(1, +\infty)$ a expressão (4.1) é crescente de modo que para eles as soluções devem ser funções contínuas crescentes. Segundo o Teorema de Existência e Unicidade, as curvas integrais não se intersectam. Portanto, é razoável concluir que as soluções no intervalo $(0, 1)$ sendo decrescentes tendam à solução de equilíbrio $\gamma(t) \equiv 0$ quando $t \rightarrow +\infty$. Analogamente concluímos que as soluções no intervalo $(-\infty, 0)$ sendo crescentes parecem tender à $\gamma(t) \equiv 0$ e no intervalo $(1, +\infty)$.

A análise simplista feita no parágrafo anterior nada prova de maneira rigorosa, mas é válida na medida em que nos dá uma ideia do comportamento das soluções. Observe a seguir algumas curvas integrais para a equação (4.1) e compare com a análise feita.

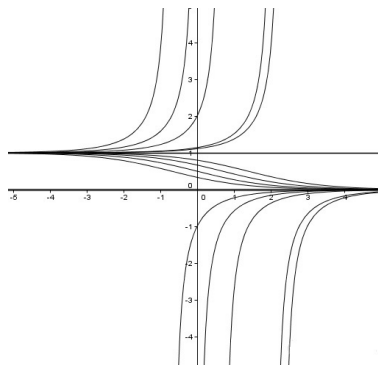


Figura 4.2: Curvas integrais de (4.1).

Exemplo 4.2. Considere a equação diferencial $x' = -x$ sujeita à condição inicial $x(0) = 1$ cuja solução é $\gamma(t) = e^{-t}$. A solução $\gamma(t)$ é assintoticamente estável ; de fato, considere \bar{x}_0 tal que $|\bar{x}_0 - x_0| = |\bar{x}_0 - 1| < \delta$. Neste caso:

$$|\gamma(t) - \bar{\gamma}(t)| = |x_0 e^{-t} - \bar{x}_0 e^{-t}| = |x_0 - \bar{x}_0| < \delta e^{-t} < \delta < \epsilon$$

e portanto para cada $\epsilon > 0$ basta escolher $\delta < \epsilon$. Como para qualquer δ_0 fazendo $t \rightarrow \infty$ temos

$$|\gamma(t) - \bar{\gamma}(t)| = |x_0 - \bar{x}_0| e^{-t} \rightarrow 0$$

segue que a estabilidade é assintótica.

Dizemos que uma equação diferencial é estável (respectivamente assintoticamente estável) se todas as soluções são estáveis (respectivamente assintoticamente estáveis).

Exemplo 4.3. Suponha que a população de uma certa espécie evolui de acordo com o modelo $x' = F(t, x)$. Num dado momento, acontece uma perturbação no meio, provocada pela introdução de novos indivíduos na população e que subitamente altera o valor desta população para x_0 . Seja t_0 o momento em que isto ocorre. Na ausência desta perturbação externa, o valor da população em t_0 seria \bar{x}_0 e sua trajetória futura seria $\gamma(t, t_0, \bar{x}_0)$ para $t \geq 0$. No entanto, devido à perturbação considerada, o comportamento da população passará a seguir a trajetória $\gamma(t, t_0, x_0)$. Se a solução é estável, o choque externo (supondo que ele seja pequeno) não irá alterar significativamente a evolução da população, pois supondo $\|x_0 - \bar{x}_0\|$ relativamente "pequeno" teremos $\|\gamma(t, t_0, x_0) - \gamma(t, t_0, \bar{x}_0)\|$ igualmente "pequeno". Se a solução é assintoticamente estável devemos esperar que $\|\gamma(t, t_0, x_0) - \gamma(t, t_0, \bar{x}_0)\| \rightarrow 0$ quando $t \rightarrow \infty$, ou seja, o efeito da perturbação é absorvido ao longo do tempo e desaparece assintoticamente. Finalmente, se a solução é instável, pequenas alterações nas condições iniciais afetam, para todo tempo t o comportamento da solução e o comportamento da população neste caso tenderia a desviar-se cada vez mais da solução $\gamma(t, t_0, x_0)$. A figura abaixo ilustra os três casos:

Como temos particular interesse na análise da estabilidade dos pontos de equilíbrio, isto é, na análise da estabilidade das soluções constantes de $x' = F(t, x)$ podemos reescrever a Definição (4.1) da seguinte maneira:

Definição 4.2. (*Ponto de equilíbrio estável, assintoticamente estável e instável*) Seja $\gamma(t) = \bar{x}$ um ponto de equilíbrio da equação diferencial $x' = F(t, x)$. Então:

1. O ponto de equilíbrio \bar{x} é **estável** se para cada $\epsilon > 0$ existe $\delta(t_0, \epsilon)$ tal que a desigualdade $\|x_0 - \bar{x}\| \leq \delta$ implica $\|\gamma(t, t_0, x_0) - \bar{x}\| \leq \epsilon$ para qualquer solução $\gamma(t, t_0, x_0)$;
2. O ponto de equilíbrio \bar{x} é **assintoticamente estável** se for estável e se existir $\delta_0 > 0$ tal que a desigualdade $\|x_0 - \bar{x}\| \leq \delta_0$ implica $\|\gamma(t, t_0, x_0) - \bar{x}\| \rightarrow 0$ quando $t \rightarrow \infty$;

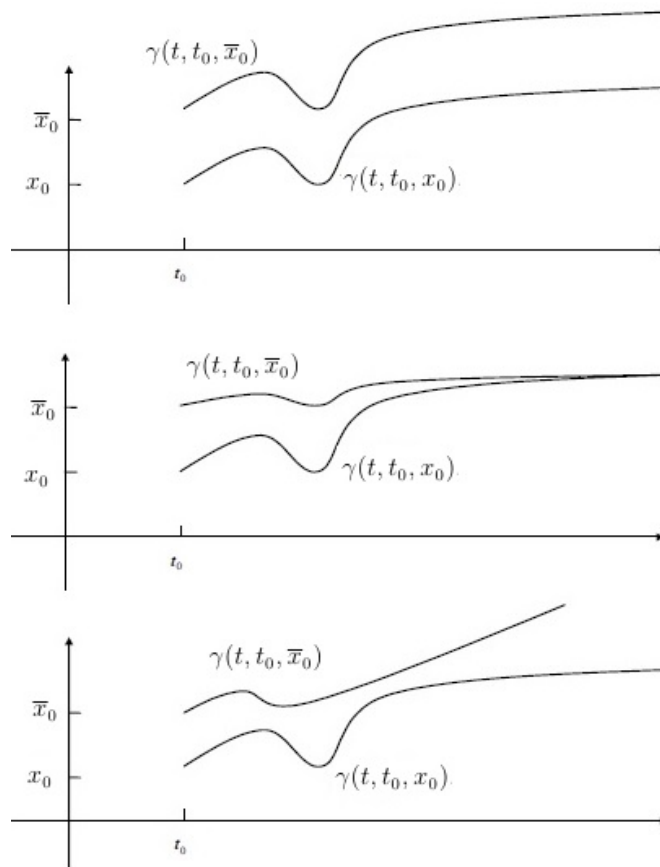


Figura 4.3: Solução estável, assintoticamente estável e instável

3. O ponto de equilíbrio \bar{x} é *instável* se não for estável.

4.2 Estabilidade de equações lineares

Provaremos nesta seção que a estabilidade (assintótica ou não) da equação linear homogênea $X' = AX$ é determinada pela análise dos autovalores da matriz A . O resultado é teoricamente simples, mas na prática pode se tornar de difícil uso, uma vez que o cálculo dos autovalores não é uma questão computacional fácil no caso de matrizes de ordem muito alta.

Para a demonstração do Teorema (4.1), resultado central desta seção precisaremos do lema abaixo:

Lema 4.1. *Se a matriz $A \in M_d$ tem apenas autovalores com parte real negativa, então existem constantes $k, \alpha > 0$ tais que para $t \geq 0$*

$$\|e^{At}\| \leq ke^{-\alpha t} \quad (4.2)$$

com a norma $\|A\| = \sup_{x \neq 0} \frac{\|A(X)\|}{\|X\|}$ definida a partir de qualquer norma $\|\cdot\|$ em \mathbb{R}^d .

Demonstração. Seja $M \in M_d$ tal que $M^{-1}AM = J$, onde J é a Forma de Jordan da matriz A (conforme a teoria desenvolvida no Apêndice B).

Então,

$$M^{-1}e^{At}M = e^{M^{-1}AMt} = e^{Jt},$$

de onde $e^{At} = Me^{Jt}M^{-1}$.

Logo,

$$\|e^{At}\| \leq \|M\| \|e^{Jt}\| \|M^{-1}\| \leq \sum_{j=1}^d \|e^{\lambda_j t}\| \|p_j(t)\| = \sum_{j=1}^d e^{Re(\lambda_j)t} |p_j(t)|,$$

onde p_j são polinômios.

Tomando agora $-\alpha_1 < -\alpha < 0$, para todo λ_j segue que $Re(\lambda_j) < -\alpha$ de onde

$$\|e^{At}\| \leq \sum_{j=1}^d e^{-\alpha_1 t} |p_j(t)| = e^{-\alpha_1 t} q(t) = e^{-\alpha t} e^{-(\alpha_1 - \alpha)t} q(t).$$

Como $\alpha_1 - \alpha > 0$, existe $k > 0$ tal que

$$e^{-\alpha t} e^{-(\alpha_1 - \alpha)t} q(t) \leq k, \quad t \geq 0$$

e portanto,

$$\|e^{At}\| \leq ke^{-\alpha t}.$$

□

Teorema 4.1. *Qualquer solução (incluindo as de equilíbrio) da equação $X' = AX$ linear e homogênea (inclusive a solução nula) é:*

1. **Assintoticamente Estável** se todos os autovalores de A possuírem parte real negativa;
2. **Estável** se todos os autovalores de A possuírem parte real negativa ou parte real nula e os autovalores de parte real nula possuírem multiplicidade algébrica igual à multiplicidade geométrica;
3. **Instável** se pelo menos um dos autovalores tiver parte real positiva ou se existir algum autovalor com parte real nula com multiplicidade algébrica maior que a multiplicidade geométrica.

Demonstração. Provemos apenas a afirmação 1. O restante da demonstração pode ser encontrada em [10].

Sem perda de generalidade vamos analisar a estabilidade da solução de equilíbrio $\bar{X} = 0$, pois para sistemas lineares a diferença de duas soluções é ainda uma solução. Conforme provado no capítulo anterior, toda solução $X(t) = (X_1(t), \dots, X_d(t))^T$ do sistema $X' = AX$ é da forma $X(t) = e^{At}X_0$. Seja $\phi_{ij}(t)$ o elemento ij da matriz e^{At} e seja $X_0 = (x_1^0, \dots, x_d^0)^T$; então

$$X_i(t) = \phi_{i1}(t)x_1^0 + \dots + \phi_{id}(t)x_d^0 \quad (4.3)$$

Suponha que todos os autovalores de A possuam parte real negativa, ou seja, que $\max_{i=1,\dots,d} \operatorname{Re}(\lambda_i) < 0$. Seja $-\alpha_1 = \max_{i=1,\dots,d} \operatorname{Re}(\lambda_i)$ o valor da maior parte real no conjunto dos autovalores. Então, existem (de acordo com o Lema (4.1)) $K > 0$ e $\alpha > 0$ tais que $-\alpha_1 < -\alpha < 0$ e $|\phi_{ij}(t)| \leq Ke^{-\alpha t}$ para todo $t \geq 0$. Consequentemente, para todo $i = 1, \dots, d$ temos

$$\begin{aligned} |X_i(t)| &= |\phi_{i1}(t)x_1^0 + \dots + \phi_{id}(t)x_d^0| \\ &\leq Ke^{-\alpha t}|x_1^0 + \dots + x_d^0| \\ &\leq Ke^{-\alpha t}(|x_1^0| + \dots + |x_d^0|) \\ &\leq ke^{-\alpha t} \max\{|x_1^0|, \dots, |x_d^0|\} = Ke^{-\alpha t} d \|X_0\| \end{aligned} \quad (4.4)$$

Assim,

$$\|X(t)\| \leq \max\{|X_1(t)|, \dots, |X_n(t)|\} \leq dKe^{-\alpha t} \|X_0\| \quad (4.5)$$

Dado $\epsilon > 0$, para garantir $\|X(t)\| < \epsilon$ é suficiente que ocorra $\|X_0\| < \frac{\epsilon}{dK}$, com $t = 0$. Neste caso, escolhemos $\delta > 0$ tal que $\delta < \frac{\epsilon}{dK}$, de onde $\|X_0\| \leq \delta$ e portanto, $\|X(t)\| < \epsilon$. Resulta também que para qualquer condição inicial temos $\|X(t)\| \rightarrow 0$ quando $t \rightarrow +\infty$, ou seja, a solução de equilíbrio é assintoticamente estável. \square

Exemplo 4.4. O Teorema (4.1) permite classificar totalmente os pontos críticos dos sistemas planares descritos na Seção 3.4 do capítulo anterior quanto à estabilidade através da análise dos autovalores.

Encontra-se resumida na tabela abaixo a classificação destes sistemas quanto à estabilidade de soluções.

Sinal dos Autovalores	Tipo de Ponto Crítico	Estabilidade
$\lambda_1 > \lambda_2 > 0$	Nó	Instável
$\lambda_1 < \lambda_2 < 0$	Nó	Assintoticamente Estável
$\lambda_1 < 0 < \lambda_2$	Ponto de Sela	Instável
$\lambda_1 = \lambda_2 > 0$	Nó	Instável
$\lambda_1 = \lambda_2 < 0$	Nó	Assintoticamente Estável
$\lambda_1, \lambda_2 = a \pm bi$ com $a > 0$	Ponto Espiral	Instável
$\lambda_1, \lambda_2 = a \pm bi$ com $a < 0$	Ponto Espiral	Assintoticamente Estável
$\lambda_1, \lambda_2 = \pm bi$	Centro	Estável

Tabela 4.1: Estabilidade dos sistemas lineares $X' = AX$ de Ordem 2

É importante destacar que vale a recíproca dos resultados apresentados acima, ou seja, o conhecimento do tipo de estabilidade apresentada pelo sistema linear $X' = AX$ permite que se deduzam as informações acerca do sinal dos autovalores associados a A .

Como enunciado e provado no teorema acima, o conhecimento do sinal da parte real dos autovalores da matriz A fornece a informação acerca do comportamento das soluções quando $t \rightarrow \infty$. Se for mais simples calcular os autovalores de A do que as soluções explícitas de $X' = F(t, X)$ o resultado é bastante útil.

Como os autovalores são as raízes do polinômio característico de A , no caso de dimensão 2 o problema está resolvido completamente, mas começa a ficar trabalhoso em dimensões 3 e 4 respectivamente e fica complicado em dimensão superior a 4 já que neste caso não existe solução por radicais.

Assim, torna-se vantajosa a existência de critérios que decidam sobre o sinal das raízes do polinômio característico sem o conhecimento explícito das mesmas, mas somente utilizando o conhecimento dos coeficientes do polinômio característico.

A busca destes critérios é conhecimento como **problema de Routh-Hurwitz** e para as dimensões 3 e 4 existe um resultado, cuja demonstração pode ser encontrada em *Pontriaguine* (1975), p.62, juntamente com critérios para qualquer grau do polinômio característico, dado por:

Proposição 4.1. *Seja $p(\lambda) = a_4\lambda^4 + a_3\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0$:*

1. *Se $a_4 = 1$ então as raízes de $p(\lambda)$ possuem (todas) parte real negativa se e somente se a_3, a_2, a_1 e a_0 são positivos e $a_3a_2a_1 > a_3^2a_0 + a_1^2$;*
2. *Se $a_4 = 0$ então as raízes de $p(\lambda)$ possuem (todas) parte real negativa se e somente se a_2, a_1 e a_0 são positivos e $a_2a_1 > a_0$.*

Exemplo 4.5. (Oscilador Harmônico sem Atrito) Conforme Exemplo (1.1) a equação que modela o movimento de uma mola sem a ação de forças externas é

$$x'' = -kx,$$

onde a constante $k = \frac{C}{m}$ é positiva e cujo ponto crítico é a origem.

Façamos $k = w^2$, de onde podemos reescrever a equação para o modelo do oscilador harmônico sem atrito como

$$x'' + w^2x = 0.$$

Escrevendo $y = x'$ e $y' = -w^2x$ a equação se transforma no sistema linear

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -w^2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

O polinômio característico de $\begin{pmatrix} 0 & 1 \\ -w^2 & 0 \end{pmatrix}$ é $p(\lambda) = \lambda^2 + w^2$ e suas raízes são $\lambda = \pm |w|i$. Segue do Exemplo (4.3) que o sistema é estável em torno da origem, o que é intuitivo do ponto de vista físico: não havendo rupturas da mola e nem ação de forças externas, o sistema deve oscilar de maneira independente das condições iniciais até parar completamente, o que acontece após a dissipação total da energia.

Exemplo 4.6. (Oscilador Harmônico com Atrito) Considere agora o problema anterior mas agora sujeito à ação de uma força de atrito que depende em geral da viscosidade do meio em que ocorre o movimento; pela Segunda Lei de Newton, sendo $\eta \geq 0$ o coeficiente de atrito (que independe do tempo) devemos ter $m\ddot{x} = -kx - \eta\dot{x}$ de onde podemos escrever

$$\ddot{x} + 2ax' + w^2x = 0,$$

desde que façamos $k = w^2 > 0$ e $a = \frac{\eta}{2m} > 0$. Em notação matricial a equação acima é equivalente a

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -w^2 & -2a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

O polinômio característico de $\begin{pmatrix} 0 & 1 \\ -w^2 & -2a \end{pmatrix}$ é $p(\lambda) = \lambda^2 + 2a\lambda + w^2$ onde sabemos ser a e w constantes positivas. Usando as Relações de Girard, as raízes λ_1 e λ_2 são tais que (i) : $\lambda_1\lambda_2 > 0$ e (ii) : $\lambda_1 + \lambda_2 = -a < 0$; por (i) segue que as raízes possuem o mesmo sinal e por (ii) segue que elas devem ser negativas. Pelo Teorema (4.1) segue que as soluções (e em particular a origem) são assintoticamente estáveis.

Observação 4.1. Para estudar a estabilidade da equação linear $X' = AX$ basta que se analise a solução de equilíbrio $X = 0$; o teorema a seguir afirma que a estabilidade desta solução é compartilhada por todas as demais soluções; em outras palavras, a solução nula $X = 0$ (com instante inicial t_0 arbitrário) é estável (respectivamente assintoticamente estável) se e somente se todas as soluções são estáveis (respectivamente assintoticamente estável).

Sob condições apropriadas, a estabilidade assintótica da equação $X' = AX$ se mantém mesmo que se introduza uma perturbação não linear, ou seja, quando passamos a considerar a equação $X' = AX + g(t, x)$. Precisaremos dos lemas abaixo:

Lema 4.2. (Lema de Gronwall) Sejam $u, v : [a, b] \rightarrow \mathbb{R}$ funções contínuas com $v \geq 0$ e seja $c \in \mathbb{R}$. Se

$$u(t) \leq c + \int_a^t u(s)v(s)ds$$

para qualquer $t \in [a, b]$, então

$$u(t) \leq ce^{\int_a^t v(s)ds} \tag{4.6}$$

para qualquer $t \in [a, b]$.

Demonstração. Considere as funções

$$R(t) = \int_a^t u(s)v(s)ds$$

e

$$V(t) = \int_a^t v(s)ds$$

Temos que $R(a) = 0$ e por hipótese $R'(t) = u(t)v(t) \leq (c + R(t))v(t)$ e portanto $R'(t) - v(t)R(t) \leq cv(t)$. Assim,

$$\frac{d}{dt}(e^{-V(t)}R(t)) = e^{-V(t)}(R'(t) - v(t)R(t)) \leq cv(t)e^{-V(t)}$$

Como $R(a) = 0$, concluímos que

$$e^{-V(t)}R(t) \leq \int_a^t cv(\xi)e^{-V(\xi)}d\xi = c(1 - e^{-V(t)})$$

e portanto $R(t) \leq ce^{V(t)} - c$, de onde

$$u(t) \leq c + R(t) \leq ce^{V(t)}$$

para todo $t \in [a, b]$ como queríamos. \square

Lema 4.3. *Seja $F : U \rightarrow \mathbb{R}^d$ uma função contínua e localmente lipschitz em X num conjunto aberto $U \subset \mathbb{R}^{1+d}$. Se uma solução $X(t)$ da equação $X' = F(t, x)$ tem intervalo máximo $]a, b[$ então para cada compacto $K \subset U$ existe $\epsilon > 0$ tal que $(t, X(t)) \in U$ para qualquer $t \in]a, a + \epsilon[\cup]b - \epsilon, b[$ (quando $a = -\infty$ temos $]a, a + \epsilon[= \emptyset$ e quando $b = +\infty$ temos $]b - \epsilon, b[= \emptyset$).*

Demonstração. Consulte [9]. \square

Teorema 4.2. *Considere a equação $X' = AX$ com $A \in M_d$ e suponha que ela seja assintoticamente estável. Seja $g : \mathbb{R}^{1+d} \rightarrow \mathbb{R}^d$ uma função contínua e localmente lipschitz em x com $g(t, 0) = 0$ para $t \in \mathbb{R}$ e tal que*

$$\limsup_{x \rightarrow 0} \sup_{t \in \mathbb{R}} \frac{\|g(t, x)\|}{\|x\|} = 0. \quad (4.7)$$

Então a solução nula da equação $X' = AX + g(t, x)$ é assintoticamente estável. Além disso, existem constantes $C, \lambda, \epsilon > 0$ tais que para cada $t_0 \in \mathbb{R}$ e cada solução $X(t)$ com $\|X(t_0)\| < \epsilon$ tem-se para $t \geq t_0$

$$\|X(t)\| \leq Ce^{-\lambda(t-t_0)}\|x(t_0)\|. \quad (4.8)$$

Demonstração. Como a equação $X' = AX$ é assintoticamente estável, a matriz A possui apenas autovalores com parte real negativa e pelo Lema (4.1) existem constantes $c, \mu > 0$ tais que

$$\|e^{At}\| \leq ce^{-\mu t},$$

para $t \geq 0$. Das hipóteses sobre a função g decorre que

$$\|g(t, x)\| \leq \epsilon\|x\|,$$

para quaisquer $t \in \mathbb{R}$ e $x \in \mathbb{R}^d$ com $\|x\| < \delta$. Considere agora $X(t)$ a solução de $X' = AX + g(t, x)$ com $X(t_0) = X_0$. Segue da Fórmula da Variação das Constantes (ver Teorema 3.3) temos

$$X(t) = e^{A(t-t_0)}X_0 + \int_{t_0}^t e^{A(t-s)}g(s, x(s))ds,$$

para t no intervalo máximo de solução. Notemos que dado $X_0 \in \mathbb{R}^d$ com $\|X_0\| < \delta$ a solução $X(t)$ satisfaz $\|X(t)\| < \delta$ para $t > t_0$ e suficientemente próximo de t_0 , ou seja, $t \in [t_0, t_1]$. Assim, para $t \in [t_0, t_1]$ segue que

$$\|X(t)\| \leq ce^{-\mu(t-t_0)}\|X_0\| + \int_{t_0}^t ce^{-\mu(t-s)}\epsilon\|x(s)\|ds$$

de onde

$$e^{\mu t}\|X(t)\| \leq ce^{\mu t_0}\|X_0\| + \int_{t_0}^t c\epsilon e^{\mu s}\|x(s)\|ds$$

e pelo Lema (4.2) (de Gronwall)

$$e^{\mu t}\|X(t)\| \leq ce^{\mu t_0}\|X_0\|e^{c\epsilon(t-t_0)}.$$

Assim,

$$\|X(t)\| \leq ce^{(-\mu+c\epsilon)(t-t_0)}\|X_0\|$$

Agora, supondo $c \geq 1$ e tomando $\epsilon > 0$ de modo que $-\mu + c\epsilon < 0$ resulta da última desigualdade acima que se $\|X_0\| < \frac{\delta}{c}$ então $\|X(t)\| < \delta$ para $t \in [t_0, t_1]$. Logo, existe $t_2 > t_1$ tal que a solução $X(t)$ está definida no intervalo $[t_0, t_2]$ e satisfaz $\|X(t)\| < \delta$ para $t \in [t_0, t_2]$. Podemos deste modo repetir sucessivamente o argumento anterior (sem mudar ϵ e δ) para concluir que existe uma sequência $(t_n)_{n \in \mathbb{N}}$ crescente tal que $X(t)$ está definida no intervalo $[t, t_n]$ e satisfazendo $\|X(t)\| < \delta$. Seja agora $b = \sup t_n$. Se $b < +\infty$ então teríamos $\|X(b^-)\| \leq \delta$ o que contradiria o Lema (4.3). Logo $b = +\infty$ e portanto a solução $X(t)$ está definida para $[t_0, +\infty)$. Isto prova que a solução nula é assintoticamente estável. \square

4.3 Estabilidade de equações não lineares

Nesta seção estenderemos para equações não lineares os resultados obtidos na seção anterior quando apresentamos um critério para a estabilidade das lineares. Em poucas palavras, veremos que para analisar os sistemas de equações não lineares, basta estudar sua *parte linear*, ou seja, estudar sua aplicação derivada.

Considere a equação $X' = X$ onde $X : U \subset \mathbb{R}^d \rightarrow \mathbb{R}^d$ é uma equação não linear com \bar{X} ponto de equilíbrio.

O lema abaixo, cuja demonstração essencialmente usa o Polinômio de Taylor, será fundamental para as considerações que faremos na sequência.

Lema 4.4. *Seja $F : U \subset \mathbb{R}^d \rightarrow \mathbb{R}^d$ uma função com derivadas parciais de segunda ordem contínuas com respeito a cada uma de suas variáveis x_1, x_2, \dots, x_d e seja \bar{X} um ponto qualquer do domínio de F . Então, dado $Z = (z_1, z_2, \dots, z_d) \in \mathbb{R}^d$ suficientemente próximo de \bar{X} , $F(\bar{X} + Z)$ pode ser escrita como*

$$F(\bar{X} + Z) = F(\bar{X}) + DF(\bar{X})Z + G(Z), \quad (4.9)$$

onde

$$\frac{G(Z)}{\max\{\|z_1\|, \|z_2\|, \dots, \|z_d\|\}}$$

é uma função contínua na variável Z e identicamente nula se $Z = 0$. e

$$DF(\bar{X}) = \begin{pmatrix} \frac{\partial F_1(\bar{X})}{\partial x_1} & \cdots & \frac{\partial F_1(\bar{X})}{\partial x_d} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_d(\bar{X})}{\partial x_1} & \cdots & \frac{\partial F_d(\bar{X})}{\partial x_d} \end{pmatrix} \quad (4.10)$$

é a **matriz jacobiana** de F calculada no ponto \bar{X} .

Demonstração. Consulte a referência [10].

□

Provaremos agora que a estabilidade do ponto de equilíbrio \bar{X} é determinada pela análise da matriz $DF(\bar{X})$, nos mesmos moldes do que fizemos na seção anterior para equações lineares.

Observação 4.2. Assumiremos sem perda de generalidade que $\bar{X} = 0$ é a solução de equilíbrio na demonstração abaixo. De fato, para um ponto \bar{X} qualquer de equilíbrio, considere $Z = X - \bar{X}$ e portanto $X = Z + \bar{X}$. Segue que $Z' = F(X) = F(Z + \bar{X})$ ou ainda $Z' = G(Z)$. Se \bar{X} é a solução de equilíbrio de $X' = F(X)$, isto é, se $F(\bar{X}) = 0$ então 0 é solução de equilíbrio de $Z' = G(Z)$ pois $G(0) = F(0 + \bar{X}) = 0$.

Teorema 4.3. (*Lyapunov-Perron*) *Considere a equação diferencial $X' = F(X)$ e suponha que F possua derivadas de segunda ordem contínuas. Então:*

1. *O ponto de equilíbrio \bar{X} de $X' = F(X)$ é assintoticamente estável se todos os autovalores da matriz $DF(\bar{X})$ possuem parte real negativa;*
2. *O ponto de equilíbrio \bar{X} de $X' = F(X)$ é instável se pelo menos um dos autovalores de $DF(\bar{X})$ possuir parte real positiva.*

Demonstração. Conforme Observação (4.2) vamos supor que $\bar{X} = 0$ é ponto de equilíbrio de $X' = F(X)$. Como $X' = AX + G(X)$ e G é contínua segue Fórmula da Variação das Constantes e de $X(0) = X_0$ que

$$X(t) = e^{At}X_0 + \int_0^t e^{A(t-s)}G(X(s))ds. \quad (4.11)$$

Precisamos provar que, sob as hipóteses assumidas no item (1), não somente $\|X(t)\|$ é limitada para qualquer t como também que $\|X(t)\| \rightarrow 0$ quando $t \rightarrow +\infty$. Pela construção feita na demonstração do Teorema (3.1) sabemos que, sendo os autovalores todos negativos, existem $K > 0$ e $k > 0$ tais que

$$\|e^{At} X_0\| \leq K e^{-kt} \|X_0\|,$$

de onde resulta também

$$\|e^{A(t-s)} G(X(s))\| \leq K e^{-k(t-s)} \|G(X(s))\|.$$

Além disso, existe $\sigma > 0$ tal que ¹

$$\|X\| < \sigma \rightarrow \|G(X)\| \leq \frac{k}{2K} \|X\|. \quad (4.12)$$

Portanto, podemos agora escrever para $\|X\| < \sigma$:

$$\begin{aligned} \|X(t)\| &\leq \|e^{At} X_0\| + \int_0^t \|e^{A(t-s)} G(X(s))\| ds \\ \|X(t)\| &\leq K e^{-kt} \|X_0\| + \int_0^t K e^{-k(t-s)} \|G(X(s))\| ds \\ \|X(t)\| &\leq K e^{-kt} \|X_0\| + \int_0^t K e^{-k(t-s)} \frac{k}{2K} \|X(s)\| ds \\ \|X(t)\| &\leq K e^{-kt} \|X_0\| + \frac{k}{2K} \int_0^t K e^{-k(t-s)} \|X(s)\| ds. \end{aligned} \quad (4.13)$$

Multiplicando a equação anterior por e^{kt} segue que

$$e^{kt} \|X(t)\| \leq K \|X_0\| + \frac{k}{2K} \int_0^t K e^{ks} \|X(s)\| ds \quad (4.14)$$

e fazendo $Z(t) = e^{kt} \|X(t)\|$ a desigualdade acima pode ser colocada na forma

$$Z(t) \leq K \|X_0\| + \frac{k}{2} \int_0^t Z(s) ds \quad (4.15)$$

e pelo Lema (4.2) (de Gronwall) segue que

$$Z(t) \leq K \|X_0\| e^{\frac{tk}{2}} \quad (4.16)$$

e portanto $Z(t) = e^{kt} \|X(t)\| \leq K \|X_0\| e^{\frac{tk}{2}}$, ou seja,

$$\|X(t)\| \leq K \|X_0\| e^{-kt} e^{\frac{tk}{2}} = K \|X_0\| e^{-\frac{k}{2}t}, \quad (4.17)$$

¹Note que na demonstração do Teorema (4.1) foi usada a norma do máximo e que deste modo, $\|G(X)\|$ é uma forma quadrática e $\|X\|$ é uma forma linear. Para valores X_i próximos da solução de equilíbrio, isto é, próximos do zero, $\|G(X)\|$ é "pequeno" se comparado com $\|X\|$. Assim, numa vizinhança do zero, certamente se tem $\|G(X)\| \leq c \|X\|$, onde c é uma constante positiva e que depende do conjunto onde os X_i estão variando. Por conveniência, escolhemos $c = \frac{k}{2K}$. Basta admitir que os X_i variam num conjunto respeitando a restrição $\|x\| < \sigma$ para $\sigma > 0$.

desde que $\|X(s)\| \leq \sigma$, $0 \leq s \leq t$. Mas se $\|X_0\| = \|X(0)\| < \frac{\sigma}{K}$ a desigualdade $\|X(t)\| \leq K\|X_0\|e^{-\frac{k}{2}t}$ garante que $\|X(t)\| < \sigma$ se verifica para todo t e consequentemente ela é verdadeira para todo t se o valor inicial estiver sujeito à $\|X(0)\| < \frac{\sigma}{K}$, condição na qual se verifica também $\|X(t)\| \rightarrow 0$.

Para a demonstração do item 2, consulte [?]. \square

Observação 4.3. A estabilidade estudada pelo Teorema acima tem caráter essencialmente local. Para ver isso, considere por exemplo a equação $x' = -x^2 + 5x + 6$ cujos pontos de equilíbrio são $\bar{x} = -1$ e $\bar{x} = 6$. Na linguagem do teorema acima, a "matriz" $DF(-1) = f'(-1) = 2 + 5 = 7$ é tal que o único autovalor é positivo, de onde segue que a solução $\bar{x} = -1$ é instável. O ponto $\bar{x} = 6$ é assintoticamente estável, uma vez que $DF(6) = f'(6) = -12 + 5 = -7 < 0$, ou seja, qualquer solução iniciada numa vizinhança de $\bar{x} = 6$ tende para $\bar{x} = 6$ quando $t \rightarrow \infty$; no entanto, o valor inicial x_0 não pode ser escolhido livremente em \mathbb{R} , isto é, não se garante $x(t, 0, x_0) \rightarrow \bar{x} = 6$ para todo $x_0 \in \mathbb{R}$ (basta ver que para $x_0 = 1$ a solução é instável).

Exemplo 4.7. Considere um pêndulo simples num plano coordenado (vertical) sob a ação da força da gravidade. Suponha uma haste rígida de tamanho $l > 0$ e massa desprezível tenha um extremo fixado na origem do plano e uma partícula de massa $m > 0$ no outro extremo. Vê-se portanto que as posições possíveis desta partícula estão sobre a circunferência de centro $(0, 0)$ e raio l . Sendo θ a posição da partícula e supondo ainda a existência de uma força de atrito agindo sobre o pêndulo a equação deste sistema físico é

$$\theta'' + k\theta' + g \operatorname{sen} \theta = 0$$

Como se trata de uma equação diferencial de segunda ordem, introduzimos a variável $\omega = \theta'$ a velocidade angular, de modo que obtemos um sistema de primeira ordem, com $k > 0$:

$$\begin{cases} \theta' = \omega \\ \omega' = -g \operatorname{sen} \theta - k\omega \end{cases}$$

cujos pontos de equilíbrio são $(\theta, \omega) = (n\pi, 0)$ para $n \in \mathbb{N}$. É intuitivo que para n par os pontos de equilíbrio são assintoticamente estáveis, já que correspondem à posição da haste "voltada para baixo". Para provar a estabilidade assintótica destes pontos, observe que a matriz jacobiana $DF(n\pi, 0)$ é dada por

$$DF(n\pi, 0) = \begin{pmatrix} 0 & 1 \\ -g \cos n\pi & -k \end{pmatrix}$$

e para n par

$$DF(n\pi, 0) = \begin{pmatrix} 0 & 1 \\ -g & -k \end{pmatrix}$$

cujo polinômio característico é $p(\lambda) = \lambda^2 + k\lambda + g$ de raízes $\lambda = -\frac{1}{2}k \pm \frac{1}{2}\sqrt{k^2 - 4g}$ com parte real negativa, o que prova portanto a estabilidade assintótica para os pontos $(\theta, \omega) = (n\pi, 0)$ com n par. Para n ímpar, temos

$$DF(n\pi, 0) = \begin{pmatrix} 0 & 1 \\ g & -k \end{pmatrix}$$

com polinômio característico $p(\lambda) = \lambda^2 + k\lambda - g$ de raízes $\lambda = -\frac{1}{2}k \pm \frac{1}{2}\sqrt{k^2 + 4g}$ sempre reais, mas de sinais contrários; em particular, existe um autovalor com parte real positiva, o que é suficiente para garantir a instabilidade dos pontos $(\theta, \omega) = (n\pi, 0)$

Observação 4.4. O Teorema (4.3) é de aplicação limitada, uma vez que ele nada afirma nos casos em que a matriz jacobiana $DF(\bar{X})$ apresenta autovalores com parte real nula.

Exemplo 4.8. (Modelo presa-predador) Em termos biológicos, a interação entre duas espécies distintas é classificada de acordo com os efeitos que uma espécie causa sobre a outra. Esta interação é classificada como **presa-predador** quando uma das espécies (predadora) é beneficiada com a presença da segunda (presa) e esta última é diretamente afetada pela presença da primeira.

Um dos modelos clássicos em dinâmica populacional que descreve a relação presa-predador foi apresentado por volta de 1926 por Lotka e Volterra. Neste exemplo, descreveremos tal modelo e analisaremos seus pontos de singularidade.

Suponha que duas espécies $x(t), y(t)$ dividam um ambiente isolado de outras interações que não a interação presa-predador; representando por $x(t)$ a população de presas e $y(t)$ a população de predadores e suponha ainda que as duas populações variem de acordo com o tempo e que o crescimento ou decréscimo de cada uma delas dependa das taxas de natalidade e de mortalidade.

Dentro de um modelo biológico como estes e das observações naturais, também é válido supor que na ausência dos predadores, a população de presas crescerá ilimitadamente e que na ausência das presas, a população de predadores decairá até extinguir-se. Deste modo, a taxa de mortalidade das presas deve ser proporcional ao produto xy que representa o encontro entre as duas espécies.

Supondo ainda que a taxa de natalidade dos predadores é proporcional ao número de indivíduos da espécie e que o tamanho desta população seja proporcional à quantidade de alimento disponível (número de presas) concluímos que a taxa de natalidade dos predadores é também proporcional ao produto xy .

As considerações acima são suficientes para a construção do sistema

$$\begin{cases} \frac{dx}{dt} = mx - \alpha xy = x(m - \alpha y) \\ \frac{dy}{dt} = -ny + \beta xy = y(-n + \beta x) \end{cases} \quad (4.18)$$

onde as constantes m, n, α, β são positivas e:

- m é a taxa de crescimento da população de presas;
- n é taxa de mortalidade da população de predadores;
- α, β são medidas do efeito da interação entre as duas espécies.

As equações (4.19) são chamadas de **equações de Lotka-Volterra**. Vamos na sequência analisar o tipo de estabilidade em torno dos pontos de equilíbrio de (4.19).

Os pontos de equilíbrio correspondem às soluções de

$$\begin{cases} x(m - \alpha y) = 0 \\ y(-n + \beta x) = 0 \end{cases} \quad (4.19)$$

ou seja, são os pontos $A = (0, 0)$ e $B = \left(\frac{n}{\beta}, \frac{m}{\alpha}\right)$.

Vamos analisar agora a estabilidade dos pontos A e B . Como (4.19) é um sistema não linear, de acordo com o Teorema (4.3) devemos analisar a matriz jacobiana $JF(X)$ associada a este sistema; portanto, vamos analisar os autovalores da matriz

$$JF \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} m - \alpha y & -\alpha x \\ \beta y & -n + \beta x \end{pmatrix} \quad (4.20)$$

No ponto $A = (0, 0)$, (4.20) se reescreve como

$$JF \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} m & 0 \\ 0 & -n \end{pmatrix},$$

cujos polinômio característico é

$$p(\lambda) = (\lambda - m)(\lambda + n)$$

e portanto os autovalores são $\lambda_1 = m$ e $\lambda_2 = -n$.

De acordo com o Exemplo 4.3, sendo os autovalores de sinais distintos o ponto de equilíbrio $A = (0, 0)$ é ponto de sela e portanto instável.

No ponto $B = \left(\frac{n}{\beta}, \frac{m}{\alpha}\right)$, (4.20) se torna

$$JF \begin{pmatrix} \frac{n}{\beta} \\ \frac{m}{\alpha} \end{pmatrix} = \begin{pmatrix} 0 & -n\frac{\alpha}{\beta} \\ m\frac{\beta}{\alpha} & 0 \end{pmatrix},$$

cujos polinômio característico é

$$p(\lambda) = \lambda^2 + mn, \quad (4.21)$$

de raízes $\lambda = \pm i\sqrt{mn}$. Novamente pelo Exemplo 4.3 o ponto de equilíbrio $B = \left(\frac{n}{\beta}, \frac{m}{\alpha}\right)$ é um centro e portanto, estável.

Observe que ao redor do ponto B as trajetórias do sistema (4.19) são elipses (conforme a descrição apresentada na Seção 2.5 e este fato pode ser interpretado do seguinte modo: para populações de presa e predador próximas aos valores de B , o tamanho de cada uma destas sofrerá oscilações devido às interações, ora tornando-se um pouco maior e ora um pouco menor, mas sempre coexistirão; neste caso não há perigo de extinção ou superpopulação de qualquer uma das duas espécies.

4.4 Estabilidade segundo Lyapunov

A desvantagem ao se empregar os teoremas das duas seções anteriores ao estudo da estabilidade da solução de uma equação diferencial linear ou não se deve sobretudo ao fato de que eles dependem essencialmente de se conhecer os autovalores de uma matriz, o que passa pela resolução de um polinômio de grau arbitrário e portanto, bastante trabalhoso e muitas vezes inviável.

Nesta seção discutiremos critérios que são suficientes para saber se um ponto de equilíbrio é estável ou assintoticamente estável mas que não depende do conhecimento dos autovalores da parte linear do campo de vetores.

Em toda essa seção deixaremos de usar a notação vetorial e escreveremos simplesmente $x' = F(x)$ para designar uma equação diferencial mesmo que em dimensões maiores que 1. Além disso, consideraremos a função $F : U \subset \mathbb{R}^d \rightarrow \mathbb{R}^d$ localmente lipschitz e portanto contínua. Denotaremos também por $\varphi(t) = x(t, x_0)$ a solução da equação $x' = F(x)$ sujeita à condição $x(0) = x_0$.

Dada uma função diferenciável $V : U \rightarrow \mathbb{R}$ e denotando por $\nabla V(x)$ seu gradiente, definimos uma nova função $\dot{V} : U \rightarrow \mathbb{R}$ por

$$\dot{V} = \nabla V(x)F(x)$$

e notamos que

$$\dot{V} = \left(d_{\varphi(t)}V \frac{\partial}{\partial t} \varphi_t(x) \right) |_{t=0} = \frac{\partial}{\partial t} V(\varphi(t)) |_{t=0}$$

Definição 4.3. (*Função de Lyapunov*) Dado $x_0 \in U$ com $F(x_0) = 0$ dizemos que uma função diferenciável $V : U \rightarrow \mathbb{R}$ é uma função de Lyapunov para x_0 se existe um aberto $D \subset U$ contendo x_0 tal que

1. $V(x_0) = 0$ e $V(x) > 0$ para todo $x \in D - \{x_0\}$;
2. $\dot{V}(x) \leq 0$ para $x \in D$.

Diremos que uma função de Lyapunov é uma **função de Lyapunov estrita** se pudermos substituir a segunda condição por $\dot{V}(x) < 0$ para $x \in D$.

Exemplo 4.9. Considere o sistema de equações diferenciais

$$\begin{cases} x' = -x + y \\ y' = -x - y^3 \end{cases}$$

que tem $(0, 0)$ como ponto crítico. Então a função $V : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por

$$V(x, y) = x^2 + y^2$$

é uma função de Lyapunov estrita para $(0, 0)$. De fato, além de termos $V(0, 0) = 0$ e $V(x, y) > 0$ para $(x, y) \neq (0, 0)$ temos também

$$\dot{V}(x, y) = (2x, 2y)(-x + y, -x - y^3) = -2(x^2 + y^4) < 0$$

para $(x, y) \neq (0, 0)$

A existência de uma função de Lyapunov (respectivamente função de Lyapunov estrita) para um ponto de equilíbrio da equação $x' = F(x)$ garante a estabilidade (respectivamente estabilidade assintótica) neste ponto; essencialmente este é o conteúdo do teorema a seguir.

Teorema 4.4. *Considere a equação diferencial $x' = F(x)$ localmente Lipschitz em um conjunto aberto $D \subset \mathbb{R}^d$ e seja $x_0 \in D$ um ponto de equilíbrio, isto é, um ponto tal que $F(x_0) = 0$. Então*

1. *Se existe uma função de Lyapunov para x_0 então x_0 é estável.*
2. *Se existe uma função de Lyapunov estrita para x_0 então x_0 é assintoticamente estável.*

Demonstração. (1). Supomos primeiro que exista uma função de Lyapunov para x_0 em algum aberto $E \subset D$ contendo x_0 . Tome $\epsilon > 0$ tal que $B(x_0, \epsilon) \subset E$ e considere².

$$m = \min\{V(x) : x \in \partial B(x_0, \epsilon)\}.$$

Como V é contínua (uma vez que é localmente lipschitz) e $V > 0$ em $B(x_0, \epsilon) - \{x_0\}$, existe $\delta \in]0, \epsilon[$ tal que

$$0 < \min\{V(x) : x \in \overline{B(x_0, \delta)}\} < m$$

Por outro lado, segue que a função $t \mapsto V(\varphi(t))$ é não crescente (no intervalo maximal I de solução); de fato, notando que $\varphi_t = \varphi_{t-s} \circ \varphi_s^3$ para t suficientemente próximo de s então

$$\frac{\partial}{\partial t} V(\varphi_t(x))|_{t=s} = \frac{\partial}{\partial t} V(\varphi_{t-s}(\varphi_s(x)))|_{t=s} = \dot{V}(\varphi_s(x)) \leq 0 \quad (4.22)$$

²Aqui, o símbolo ∂ indica a fronteira do conjunto.

³Aqui a noção de fluxo de uma equação diferencial se torna indispensável. De maneira geral, chamamos de **fluxo** a qualquer família de transformações $\varphi_t : \mathbb{R}^d \rightarrow \mathbb{R}^d$ tal que $\varphi_0 = Id$ e $\varphi_{t+s} = \varphi_t \circ \varphi_s$, com $s, t \in \mathbb{R}$. Além disso, prova-se que as soluções de uma equação diferencial $x' = f(x)$ com condição inicial $x(0) = x_0$ definidas por $\varphi_t(x_0) = x(t, x_0)$ é um fluxo.

para $s \in I$. Portanto, qualquer solução $\varphi_t(x)$ da equação diferencial $x' = F(x)$ dada com a condição inicial $x(0) = x_0$ com $x \in \overline{B(x_0, \delta)}$ está contida em $B(x_0, \epsilon)$ para qualquer $t > 0$. Isso garante ainda que cada solução $\varphi_t(x)$ com $x \in \overline{B(x_0, \delta)}$ está definida para qualquer $t > 0$ e portanto o ponto crítico é estável.

(2). Suponha agora que exista uma função estrita de Lyapunov para x_0 . Resta mostrar que $\varphi_t(x) \rightarrow x_0$ quando $t \rightarrow +\infty$ para qualquer $x \in B(x_0, \alpha)$ com α suficientemente pequeno. Repetindo o argumento empregado na demonstração da parte (1) deste teorema, concluímos que dado $x \in E - \{x_0\}$ a função $t \mapsto V(\varphi(t))$ é decrescente (no intervalo maximal da solução). Seja agora (t_n) uma sequência de reais com $t_n \rightarrow +\infty$ tal que $(\varphi_{t_n})(x)$ é convergente para um certo y . Temos

$$V(\varphi_{t_n})(x) \rightarrow V(y),$$

quando $n \rightarrow +\infty$, pois $t \mapsto V(\varphi_t(x))$ é decrescente. Temos também

$$V(\varphi_t(x)) > V(y),$$

para todo $t > 0$. Para concluir a demonstração, suponha que $y \neq x_0$. Temos então $V(\varphi_s(y)) < V(y)$ para qualquer $s > 0$. Tomando n suficientemente grande, podemos fazer com que $\varphi_{t_n+s}(x) = \varphi(s)(\varphi_{t_n}(x))$ esteja tão próximo quanto quisermos de $V(\varphi_s(y)) < V(y)$ de onde

$$V(\varphi_{t_n+s}(x)) < V(y)$$

o que contradiz o fato de que

$$V(\varphi_t(x)) > V(y)$$

e somos levados a considerar $y = x_0$ de onde concluímos que $\varphi_t(x) \rightarrow x_0$ quando $t \rightarrow +\infty$.

□

Exemplo 4.10. Considere o sistema de equações diferenciais

$$\begin{cases} x' = y - xy^2 \\ y' = -x^3 \end{cases}$$

que tem $(0, 0)$ como ponto crítico. Então a função $V : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por

$$V(x, y) = x^4 + 2y^2 \tag{4.23}$$

é uma função de Lyapunov para $(0, 0)$. Além de termos $V(0, 0) = 0$ e $V(x, y) > 0$ para $(x, y) \neq (0, 0)$ temos também

$$\dot{V}(x, y) = (4x^3, 4y)(y - xy^2, -x^3) = 4x^3y - 4x^4y^2 - 4x^3y = -4x^4y^2 \leq 0$$

para $(x, y) \neq (0, 0)$ e pelo Teorema acima resulta que a origem é estável.

Exemplo 4.11. Considere a equação $x'' + f(x) = 0$ onde $f : \mathbb{R} \rightarrow \mathbb{R}$ é uma função de classe C^1 com $f(0) = 0$ tal que para $x \neq 0$

$$xf(x) > 0,$$

de modo que $f(x)$ tem sempre o mesmo sinal que x . Isto corresponde a aplicar uma força $-f(x)$ numa partícula de massa 1 que se desloca sem atrito. A condição $xf(x) > 0$ corresponde a supor que a força aponta sempre para a origem. A equação $x'' + f(x) = 0$ se escreve como

$$\begin{cases} x' = y \\ y' = -f(x) \end{cases}$$

que possui a origem como ponto crítico.

Vamos mostrar que a origem é um ponto estável usando funções de Lyapunov. Nomeadamente vamos considerar a função

$$V(x, y) = \frac{1}{2}y^2 + \int_0^x f(s)ds,$$

que corresponde à soma da energia cinética $\frac{1}{2}y^2$ (lembre-se que a partícula tem massa $m = 1$) com a energia potencial $\int_0^x f(s)ds$. Note que $V(0, 0) = 0$ e $V(x, y) > 0$ se $(x, y) \neq (0, 0)$ pois $xf(x) > 0$. Além disso,

$$\dot{V}(x, y) = (f(x), y)(y, -f(x)) = 0$$

de onde vemos que V é uma função de Lyapunov para $(0, 0)$. Logo, pelo Teorema (4.4) a origem é estável.

Teorema 4.5. (*Critério de Instabilidade*) Seja $x' = F(x)$ uma equação diferencial onde $F : U \subset \mathbb{R}^d \rightarrow \mathbb{R}^d$ é de classe C^1 num aberto $E \subset \mathbb{R}^d$ e seja x_0 um ponto de equilíbrio. Seja também $V : E \rightarrow \mathbb{R}^d$ uma função de classe C^1 numa vizinhança $E \subset U$ de x_0 tal que :

1. $V(x_0) = 0$ e $\dot{V} > 0$ para $x \in E - \{x_0\}$;
2. V toma valores positivos em qualquer vizinhança de x_0 .

Então, nestas condições o ponto de equilíbrio x_0 é instável.

Demonstração. Consulte [?]. □

Exemplo 4.12. Considere o sistema

$$\begin{cases} x' = 3x + y^3 \\ y' = -2y + x^2 \end{cases}$$

que tem a origem $(0, 0)$ como ponto crítico e a função

$$V(x, y) = x^2 - y^2$$

Note que $V(0, 0) = 0$ e é claro que V toma valores positivos em qualquer vizinhança de $(0, 0)$.

Além disso,

$$\dot{V}(x, y) = (2x, -2y)(3x + y^3, -2y + x^2) = 6x^2 + 4y^2 + 2xy^3 - 2x^2y$$

e como

$$\frac{\dot{V}(x, y)}{6x^2 + 4y^2} \rightarrow 1$$

quando $(x, y) \rightarrow (0, 0)$, concluímos que $\dot{V}(x, y) > 0$ para $(x, y) \neq (0, 0)$ suficientemente próximo da origem de modo que pelo teorema acima, a origem é instável.

Observação 4.5. Infelizmente não existem métodos para se determinar uma função de Lyapunov apropriada. Uma primeira tentativa consiste em analisar a função $V(x) = c_1(x_1 - \bar{x}_1)^2 + \dots + c_d(x_d - \bar{x}_d)$ onde as constantes c_i são positivas e os valores \bar{x}_i são as coordenadas do vetor x_0 , ponto de equilíbrio da equação $x' = f(x)$. Observe o próximo exemplo:

Exemplo 4.13. Vamos analisar a origem $(0, 0, 0)$, ponto de equilíbrio do sistema de equações

$$\begin{cases} x' = 2y(z - 1) \\ y' = -x(z - 1) \\ z' = -z \end{cases}$$

Se olharmos para a matriz jacobiana $DF(x, y, z)$ nada poderemos concluir, uma vez que

$$\begin{pmatrix} 0 & 2z - 2 & 2y \\ 1 - z & 0 & -x \\ 0 & 0 & -1 \end{pmatrix}$$

de onde no ponto $(x, y, z) = (0, 0, 0)$ encontramos

$$DF(0, 0, 0) = \begin{pmatrix} 0 & -2 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

cujos polinômio característico $p(\lambda) = \lambda^3 + \lambda^2 + 2\lambda + 2$ possui raízes $\lambda_1 = -1$, $\lambda_2 = i\sqrt{2}$ e $\lambda_3 = -i\sqrt{2}$.

Assim, conforme a Observação acima, vamos tentar obter uma função de Lyapunov do tipo

$$V(x, y, z) = c_1x^2 + c_2y^2 + c_3z^2$$

com $c_i > 0$ para $i = 1, 2, 3$. Note que $V(0, 0, 0) = 0$ e $V(x, y, z) > 0$ para todo $x \neq (0, 0, 0)$.

Notando que $\nabla V(x, y, z) = (2c_1x, 2c_2y, 2c_2z)$ segue que

$$\nabla V(x, y, z)F(x, y, z) = (2c_1x, 2c_2y, 2c_2z)(2y(z-1), -x(z-1), -z)$$

de onde, após simplificar as multiplicações encontramos

$$\nabla V(x, y, z)F(x, y, z) = (2c_2 - 4c_1)xy + (4c_1 - 2c_2)xyz - 2c_3z^3$$

e tomando $c_1 = 1, c_2 = 2c_1 = 2$ e $c_3 = 1$ finalmente segue que

$$\nabla V(x, y, z)F(x, y, z) = -2z^3 \leq 0$$

para $V(x, y, z) = x^2 + 2y^2 + z^3$, de modo que o ponto $(0, 0, 0)$ é de equilíbrio estável para o sistema dado.

5 Propostas didáticas

Neste capítulo apresentaremos duas propostas didáticas aos professores de Matemática do Ensino Médio que desejam trabalhar com modelagem matemática em suas aulas. Historicamente falando, as equações diferenciais ordinárias, assunto central deste trabalho, surgiram pela necessidade de modelagem de fenômenos nas diversas áreas. Assim, ao apresentar o trabalho com modelos matemáticos, temos um duplo propósito:

1. Despertar no professor de Matemática da Educação Básica a vontade de estudar modelos matemáticos contínuos ou discretos, de modo que este seja motivado a criar seus próprios modelos.
2. Motivar o aluno da Educação Básica para o estudo da Matemática, invertendo a ordem tradicional das aulas de modo que este possa ter papel significativo na construção da própria aprendizagem.

A primeira atividade discute, ainda que informalmente, as noções de estabilidade e instabilidade de um modelo matemático através de dois exemplos de fenômenos físicos.

Sugerimos inclusive, que esta proposta didática seja desenvolvida em parceria com o professor de Física, de modo que os alunos entendam os dois fenômenos abordados nesta proposta em toda a sua extensão.

Existem diversos simuladores dos sistemas massa-mola e pêndulo simples disponíveis na Internet.

Após alguns testes, optamos pelos simuladores dos sistema massa-mola e pêndulo simples disponíveis no portal do Ministério da Educação¹ nos links

<http://objetoseducacionais2.mec.gov.br/handle/mec/19475>

<http://objetoseducacionais2.mec.gov.br/handle/mec/3726>

No caso do simulador massa-mola, é permitido ao aluno escolher diversos pesos disponíveis para as molas, esticar ou comprimir a mola e analisar os efeitos, inclusive em "câmera lenta".

Para o simulador do pêndulo simples, é possível escolher a massa do pêndulo, a angulação de onde ele será abandonado, dois pêndulos ao mesmo tempo e inclusive

¹Acessado em 08/06/2014.

se o movimento se dará nas condições da Terra, na Lua, em outros planetas ou com ausência total de gravidade.

Abaixo estão apresentadas as telas iniciais dos simuladores sugeridos:

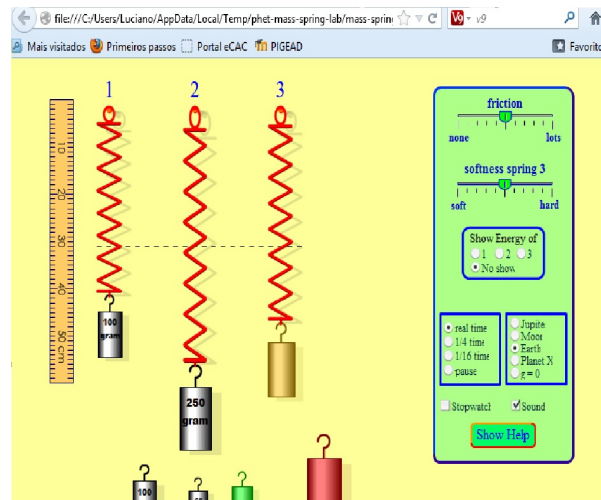


Figura 5.1: Tela Inicial do Simulador do Sistema Massa-Mola

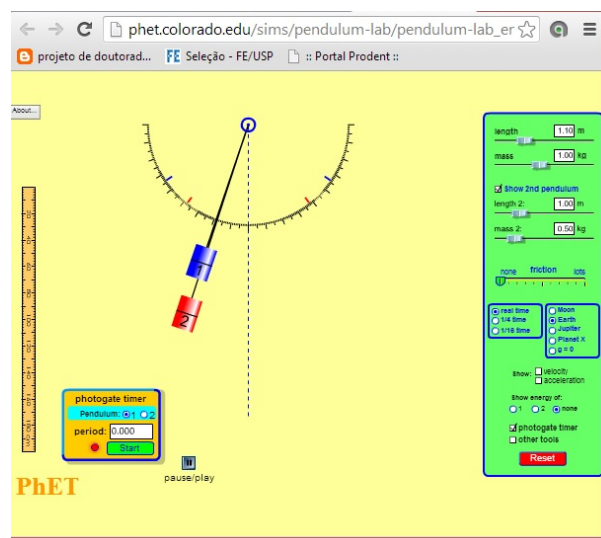


Figura 5.2: Tela Inicial do Simulador do Sistema Pêndulo Simples

A segunda das atividades, baseada no Exemplo (2.11) deste trabalho, objetiva discutir com os alunos a evolução populacional da cidade de São Paulo a partir do modelo malthusiano, que apesar de ser impreciso nas projeções para valores de t muito grandes, apresenta fundamentação matemática bastante simples e acessível a qualquer aluno do Ensino Médio que tenha estudado funções exponenciais.

5.1 Pêndulo simples e massa-mola

Esta proposta didática tem como objetivo analisar dois sistemas físicos quanto à noção de estabilidade. Claro que com alunos do Ensino Médio não cabe uma discussão formal do tema; assim, no plano de aula abaixo a discussão se dará no campo informal das ideias.

A informática se mostra uma importante aliada para a discussão a ser feita, uma vez que existem simulações/animações na Internet dos modelos que serão estudados, de modo que a compreensão dos alunos é favorecida.

Público Alvo: Alunos do Ensino Médio.

Recursos Pedagógicos: Laboratório de Informática.

Objetivo Geral: Apresentar aos alunos noções informais dos conceitos de estabilidade e instabilidade de sistemas físicos como o massa-mola e o pêndulo simples.

Objetivos Específicos:

1. Discutir a estabilidade do sistema massa-mola sem atrito ou forças externas com os alunos;
2. Discutir o sistema pêndulo simples com os alunos em condições ideais e a possibilidade da existência de "posições estáveis" e "posições instáveis";
3. Levar os alunos a concluir que, na modelagem de qualquer problema (seja físico, biológico, econômico, etc.), deve ser levada em consideração a estabilidade do fenômeno em estudo.
4. Incentivar a pesquisa por parte dos alunos, sugerindo que estes façam a pesquisa de modelos matemáticos discretos na Internet e tragam para a classe afim de que, em grupos, seja analisada a estabilidade ou não do problema modelado.

Conteúdo: Estabilidade de modelos matemáticos; análise de modelos existentes.

Desenvolvimento do Tema: Esta proposta didática tem a duração de duas aulas. Na primeira aula, os simuladores serão apresentados aos alunos. Em grupos pequenos, eles deverão simular com pesos distintos o sistema massa-mola e observar o que acontece com a mola no decorrer do tempo.

Com relação ao sistema do pêndulo simples, além de observar o que acontece com o decorrer do tempo nas diversas simulações que farão, eles deverão procurar posições nas quais o pêndulo fica em repouso.

As questões a seguir podem ser apresentadas aos alunos, de modo a guiar o estudo e observação dos grupos. Ressaltamos que cabe ao professor ampliar este repertório de questões, de acordo com o nível da turma.

1. Faça diversas simulações no sistema massa-mola. O que todas elas apresentam em comum?

2. Comprima a mola e solte-a. O que você observa? Repita esta simulação com diversos pesos e anote suas conclusões.
3. Estique a mola e solte-a. O que você observa no decorrer do tempo? Repita a simulação para diversos pesos e anote as conclusões.
4. Observe o simulador do pêndulo simples. Faça algumas simulações com pesos distintos e verifique se a massa influi no movimento circular.
5. Encontre duas posições no sistema do pêndulo simples onde este fica em repouso. Simule o abandono do pêndulo próximo destas duas regiões. O que você observa para cada uma destas posições?
6. Porque uma das posições que você encontrou no item anterior nunca é verificada quando se observa um sistema de pêndulo simples num Laboratório de Física?

Após a exploração das questões por parte dos grupos, sugerimos ao professor que discuta o que é um "modelo matemático" com os alunos e termine a aula pedindo que os alunos pesquisem exemplos modelos matemáticos discretos e os traga para a próxima aula.

Na segunda das aulas desta proposta, sugerimos ao professor que faça uma discussão com os alunos acerca das questões propostas no laboratório de informática e conclua com os alunos que a posição de repouso no sistema massa-mola é **estável**, no sentido de que, independente do massa do objeto, da mola, da força empregada na compressão ou estiramento da mola, sempre o sistema retorna para a posição de equilíbrio.

No caso do pêndulo simples, existem duas posições de repouso: o pêndulo voltado para baixo e voltado para cima, apesar de a segunda ser totalmente improvável na prática, apesar de possível teoricamente. Explique aos alunos que enquanto uma destas posições "atrai" qualquer trajetória do pêndulo, a outra "repele"; enquanto todas as simulações "tendem" para a posição do pêndulo voltado para baixo, qualquer perturbação destrói o equilíbrio do pêndulo voltado para cima.

Explique que estes fenômenos se relacionam com a estabilidade do sistema e que bons modelos matemáticos devem ser capazes de prever estas situações. Termine a aula analisando com a turma os modelos que foram pesquisados; sugira aos alunos que representem com softwares gratuitos as equações dos modelos pesquisados e a partir destes gráficos, analise junto com a turma a estabilidade ou não das soluções. Inclusive é provável que apareçam na pesquisa o modelo de Malthus, objeto de estudo da próxima proposta didática.

Em todas as aulas é importante que sejam reservados momentos para que o aluno exponha suas dúvidas, hipóteses e conclusões.

Avaliação: Durante todo o desenvolvimento da proposta os alunos serão avaliados pelos registros produzidos, pela participação das discussões e envolvimento com as atividades propostas.

5.2 Evolução populacional do município de São Paulo

Conforme já registramos, esta atividade apresenta a versão discreta do Exemplo (2.11), apropriada para alunos do Ensino Médio.

Público Alvo: Alunos do Ensino Médio.

Recursos Pedagógicos: Laboratório de Informática.

Objetivo Geral: Apresentar o modelo populacional de Malthus para os alunos em sua versão discreta.

Objetivos Específicos:

1. Criar um modelo populacional malthusiano para o Município de São Paulo e verificar sua adequabilidade;
2. Investigar se este modelo é instável ou estável de maneira informal com os alunos, através da análise gráfica do problema.

Conteúdo: Modelo de Malthus, Exponencial, Proporcionalidade.

Desenvolvimento do Tema: A proposta apresentada possui duração de quatro aulas e deve ser desenvolvida após o trabalho com a proposta "Pêndulo Simples e Massa-Mola".

Na primeira das aulas, explique o modelo malthusiano discreto aos alunos. Para a completude deste trabalho, transcrevemos abaixo em linhas gerais tal modelo:

A hipótese do modelo malthusiano é a de que a população P no instante de tempo t é proporcional à população no instante $t - 1$, sendo a constante de proporcionalidade α tomada fixa.

Assim, considere uma população inicial P_0 no instante t_0 e suponha que a constante de proporcionalidade seja α . Então, após uma unidade de tempo, ou seja, para $t = 1$ teremos

$$P_1 = (1 + \alpha)P_0$$

para os instantes $t = 2, t = 3$ e $t = 4$ segue das hipóteses que

$$P_2 = (1 + \alpha)P_1 \Rightarrow P_2 = (1 + \alpha)^2 P_0$$

$$P_3 = (1 + \alpha)P_2 \Rightarrow P_3 = (1 + \alpha)^3 P_0$$

e

$$P_4 = (1 + \alpha)P_3 \Rightarrow P_4 = (1 + \alpha)^4 P_0$$

e de maneira indutiva

$$P_n = (1 + \alpha)P_{n-1} = (1 + \alpha)^n P_0 \quad (5.1)$$

onde a taxa α é simplesmente a diferença entre a taxa de natalidade e de mortalidade da população.

Faça ainda nesta aula, uma pesquisa com os alunos como a realizada no Exemplo (2.11) sobre a população do Município de São Paulo entre os anos de 2000 e 2005 e

elabore uma tabela com os dados. Considere como população inicial a população no ano de 2000, de modo a estabelecer a variável t :

Tempo	Ano	Dados Oficiais
0	2000	10434252
1	2001	10525367
2	2002	10613691
3	2003	10698381
4	2004	10782296
5	2005	10865573

Tabela 5.1: População oficial do município de São Paulo entre 2000 e 2005.

Na segunda das aulas, discuta com os alunos os dados coletados e levante questões acerca da população como: (1) "A população pode crescer acima de qualquer limite?", (2) "Se existe limitação para o crescimento, qual seriam os fatores determinantes dela?". Após esta discussão com os alunos, volte à tabela e elabore o modelo malthusiano, explicando aos alunos que isso é possível se conhecermos a constante α na equação (5.1), que pode ser calculada do seguinte modo:

Suponha $t = 0$ no ano 2000 e então $P_0 = 10434252$; notando que para $t = 5$ (ano 2005) a população passou a ser $P = 10865573$, por substituição em (5.1) segue que

$$P_5 = (1 + \alpha)^5 P_0$$

$$10865573 = (1 + \alpha)^5 \cdot 10434252$$

de onde, isolando α , decorre que

$$\alpha = \sqrt[5]{\frac{10865573}{10434252}} - 1$$

ou seja,

$$\alpha = 0,008134$$

o que significa que entre os anos de 2000 e 2005 a população em estudo cresceu certa de 0,8134% ao ano.

Agora, substituindo (5.2) em (5.1) juntamente com $P_0 = 10434252$, obtem-se a equação do modelo malthusiano para o problema em estudo:

$$P_n = 1,008134^n \cdot 10434252 \tag{5.2}$$

Termine esta segunda aula oferecendo um espaço para os alunos manifestarem suas dúvidas sobre as manipulações matemáticas. Trata-se de uma ótima oportunidade de recuperação paralela para eles.

Na terceira das aulas, retome o modelo malthusiano construído e solicite aos alunos que em grupos façam pesquisas acerca da população do município de São Paulo entre os anos de 2006 e 2010 em sites oficiais² e comparem os valores pesquisados com os previstos pela equação (5.2), calculando inclusive os erros percentuais. Sugira um modelo de tabela a ser preenchido pelos grupos e estimule o uso das planilhas eletrônicas. É possível que os alunos apresentem dúvidas quanto ao cálculo do erro percentual; portanto, é importante dar um exemplo deste cálculo na planilha eletrônica.

Abaixo estão apresentados os dados oficiais (retirados do site do SEADE) com as respectivas projeções via equação (5.2) e os erros percentuais (os dados foram arredondados para o inteiro mais próximo no caso da projeção da população e para duas casas decimais no caso do erro percentual).

t	Ano	Projeção por (5.2)	Dados Oficiais	Erro Percentual
6	2006	10967390	10944889	0,21
7	2007	11058858	11019484	0,36
8	2008	11151089	11093746	0,52
9	2009	11244089	11168194	0,68
10	2010	11337865	11244369	0,83

Tabela 5.2: Erro percentual do modelo malthusiano para o município de São Paulo.

Após os grupos terminarem a pesquisa e o preenchimento da tabela, discuta com os alunos a adequabilidade dos dados projetados. Algumas questões a serem levantadas:

1. Os dados projetados pelo modelo construído se adequam bem à realidade?
2. O que se percebe quanto ao erro percentual?
3. Continue a tabela usando os dados para os anos de 2011, 2012 e 2013. Se possível use também as projeções de 2015, 2020 e 2030. Nestes casos, o erro se torna aceitável?

Encerre esta terceira aula comentando com os alunos o fato de que o IBGE considera aceitável um erro percentual até o limite de 5%.

Para a quarta e última aula desta proposta didática, retome com os alunos a noção de estabilidade discutida na primeira das propostas didáticas. Estimule-os a pensar na estabilidade do modelo malthusiano com as seguintes questões:

1. Se a taxa de proporcionalidade α calculada for negativa, o que acontece com a população P **independente** da população inicial P_0 ?

²Sugira os sites www.seade.sp.gov.br e www.ibge.gov.br, que seguramente são boas fontes de pesquisa.

2. Se a taxa de proporcionalidade α calculada for positiva, o que acontece com a população P **independente** da população inicial P_0 ?

Termine a aula sugerindo que construam gráficos para analisar o comportamento da equação (5.1) em cada um dos casos ($\alpha > 0$ ou $\alpha < 0$) e discuta com eles a estabilidade da população P_n . Diga a eles que esta é uma das principais críticas ao modelo de Malthus e que outros modelos foram criados de modo a corrigir o crescimento exponencial de P_n (no caso $\alpha > 0$) que não se verifica na prática.

Avaliação: Durante todo o desenvolvimento da proposta os alunos serão avaliados pelos registros escritos, pela participação das discussões e envolvimento com as atividades propostas.

A Resultados de Álgebra Linear

Um conjunto V não vazio é chamado de **espaço vetorial** sobre um corpo K (consideraremos $K = \mathbb{R}$ ou \mathbb{C}) se em seus elementos (que são chamados de vetores) estiverem definidas as operações de *soma* e *multiplicação por escalar* de modo que as seguintes propriedades sejam satisfeitas para quaisquer $u, v, w \in V$ e $\alpha, \beta \in K$:

1. (Comutativa) $u + v = v + u$;
2. (Associativa da Soma) $(u + v) + w = u + (v + w)$;
3. (Elemento Neutro) Existe um vetor, chamado de *vetor nulo*, tal que $0 + v = v$;
4. (Simétrico) Para cada vetor $v \in V$ existe $-v \in V$ tal que $v + (-v) = 0$;
5. (Associativa da Multiplicação por Escalar) $(\alpha\beta)v = \alpha(\beta v)$;
6. (Elemento Identidade) $1v = v$;
7. (Distributiva I) $\alpha(u + v) = \alpha u + \alpha v$;
8. (Distributiva II) $(\alpha + \beta)v = \alpha v + \beta v$.

Exemplo A.1. Os conjuntos M_d , \mathbb{R}^d , \mathbb{C}^d , $\mathcal{F}(\mathbb{R})$ e $\mathcal{F}(\mathbb{C})$ das matrizes quadradas de ordem d , das d -*uplas* de números reais, das d -*uplas* de números complexos, das funções contínuas com domínio real e das funções contínuas com domínio complexo respectivamente são espaços vetoriais sobre \mathbb{R} fundamentais para toda a teoria desenvolvida neste trabalho.

Um vetor $v \in V$ é uma **combinação linear** dos vetores $v_1, v_2, \dots, v_n \in V$ se existirem escalares tais que

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = \sum_{i=1}^n \alpha_i v_i.$$

Um conjunto $\mathcal{B} \subset V$ é chamado de **conjunto gerador** de V se todo elemento v do espaço V se escreve como uma combinação linear de um número finito de elementos de \mathcal{B} .

Somente nos referiremos neste texto a espaços vetoriais que sejam finitamente gerados, ou seja, espaços vetoriais que admitem um conjunto finito de geradores.

Em geral um espaço vetorial possui vários conjuntos geradores e é importante para as aplicações que encontremos um conjunto gerador com o menor número de elementos possíveis. A situação ideal é que exista um conjunto gerador onde cada elemento de V se escreva de modo **único** como combinação linear dos elementos deste conjunto gerador. Por trás dessa unicidade desejada está o importante conceito de conjunto linearmente independente.

Definição A.1. *Seja V um espaço vetorial e \mathcal{B} um subconjunto de V .*

1. *O conjunto $\mathcal{B} = \{v_1, v_2, \dots, v_d\}$ é linearmente independente se qualquer combinação linear*

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = 0$$

com $\alpha_1 \in K$ implicar $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$.

2. *O conjunto \mathcal{B} é linearmente dependente se não for linearmente independente.*

Exemplo A.2. O conjunto $\mathcal{B} = \{(0, 0, 1), (0, 1, 0), (0, 0, 1)\}$ é linearmente independente em \mathbb{R}^3 ; de fato, tomando (1) temos

$$(\alpha_1, 0, 0) + (0, \alpha_2, 0) + (0, 0, \alpha_3) = (0, 0, 0)$$

o que implica $\alpha_1 = \alpha_2 = \alpha_3 = 0$.

Observação A.1. A verificação da independência linear entre d vetores em \mathbb{R}^d pode ser feita através da verificação do valor do determinante cujas colunas sejam as coordenadas destes d vetores do seguinte modo:

Sejam d vetores $v_1 = (a_1, a_2, \dots, a_d)$, $v_2 = (b_1, b_2, \dots, b_d)$, ..., $v_d = (z_1, z_2, \dots, z_d)$ em \mathbb{R}^d . O conjunto $\mathcal{B} = \{v_1, v_2, \dots, v_d\}$ é linearmente independente se e somente se

$$\det \begin{pmatrix} a_1 & b_1 & \cdots & z_1 \\ a_2 & b_2 & \cdots & z_2 \\ \dots & \dots & \ddots & \vdots \\ a_d & b_d & \cdots & z_d \end{pmatrix} \neq 0$$

Definição A.2. *Seja V um espaço vetorial. Um conjunto $\mathcal{B} \in V$ é uma base do espaço V se*

1. *\mathcal{B} é um conjunto gerador de V ;*
2. *\mathcal{B} é um conjunto linearmente independente.*

Exemplo A.3. O conjunto $\mathcal{B} = \{(0, 0, 1), (0, 1, 0), (0, 0, 1)\}$ é uma base de \mathbb{R}^3 . Mais geralmente, qualquer d -upla de reais com entrada 1 na d -ésima posição e 0 nas demais é uma base do espaço \mathbb{R}^d . Esta base geralmente é chamada de **base canônica** de \mathbb{R}^d e seus elementos são designados por **vetores canônicos**.

Provaremos a seguir que toda base \mathcal{B} de um espaço vetorial possui a mesma quantidade de elementos; precisaremos do lema abaixo:

Lema A.1. *Seja V um espaço vetorial e assuma que $\{v_1, v_2, \dots, v_m\}$ seja um conjunto gerador de V . Então todo conjunto linearmente independente de vetores em V possui no máximo m elementos.*

Proposição A.1. *Duas bases quaisquer \mathcal{B} e \mathcal{B}' de um espaço vetorial V possuem o mesmo número de elementos.*

Demonstração. Suponha que as bases \mathcal{B} e \mathcal{B}' possuam respectivamente m e m' elementos.

Como \mathcal{B} gera V e \mathcal{B}' é linearmente independente (já que por hipótese os dois conjuntos são bases do espaço V), segue do Lema (A.1) que o número de elementos de \mathcal{B}' é menor ou igual ao número de elementos de \mathcal{B} , ou seja,

$$m' \leq m$$

Analogamente, \mathcal{B}' também gera V e \mathcal{B} é linearmente independente (novamente por serem bases), de modo que o número de elementos de \mathcal{B} é menor ou igual ao número de elementos de \mathcal{B}' :

$$m \leq m'$$

Das duas desigualdades acima, segue o resultado. □

Ao número de elementos de uma base \mathcal{B} de um espaço V dá-se o nome de **dimensão** de V . Simbolicamente: $\dim V$.

Exemplo A.4. Os espaços \mathbb{R}^d , \mathbb{C}^d e M_d sobre os corpos $\mathbb{R}, \mathbb{C}, \mathbb{R}$ possuem respectivamente dimensões d , d e $2d$.

Um conjunto $W \subset V$ de um espaço vetorial V é chamada de **subespaço vetorial de V** se com a restrição das operações de V é ele próprio um espaço vetorial. A proposição a seguir estabelece condições necessárias e suficientes para que um subconjunto W qualquer de V seja um subespaço vetorial de V :

Proposição A.2. *Seja W um subconjunto de um espaço vetorial V qualquer. Então W é subespaço vetorial de V se e somente se satisfaz as seguintes propriedades:*

1. $0 \in W$;
2. $v_1, v_2 \in W \Rightarrow v_1 + v_2 \in W \quad \forall v_1, v_2 \in W$;
3. $\lambda v \in W \quad \forall \lambda \in \mathbb{R}, v \in W$.

Exemplo A.5. O subconjunto

$$X = \{(x, 0), \quad \forall x \in \mathbb{R}\}$$

é um subespaço vetorial de \mathbb{R}^2 . De fato,

1. $(0, 0) \in X$;
2. $v_1 = (x_1, 0), v_2 = (x_2, 0) \in X \Rightarrow v_1 + v_2 = (x_1 + x_2, 0) \in X \quad \forall v_1, v_2 \in X$;
3. $\lambda \cdot (x, 0) = (\lambda x, 0) \in X, \quad \forall \lambda \in \mathbb{R}, v \in X$.

A.1 Diagonalização e forma de Jordan

Tem interesse particular em Álgebra Linear o estudo das funções entre espaços vetoriais, chamadas de **transformações lineares**. Formalmente:

Definição A.3. *Sejam U e V espaços vetoriais sobre \mathbb{R} . Uma função $T : U \rightarrow V$ é uma transformação linear se:*

1. $T(u_1 + u_2) = T(u_1) + T(u_2) \quad \forall u_1, u_2 \in U$;
2. $T(\lambda u) = \lambda T(u) \quad \forall \lambda \in K, u \in U$.

Observe que toda transformação linear leva o elemento neutro do espaço U no elemento neutro do espaço V . De fato,

$$0_V + T(0_U) = T(0_U) = T(0_U + 0_U) = T(0_U) + T(0_U)$$

e pela lei do cancelamento, $T(0_U) = 0_V$

Caso $U = V$ a transformação T será chamada de **operador linear**. No restante deste apêndice, restringiremos nossa abordagem aos operadores lineares, que desempenham papel importante na teoria das equações diferenciais.

Definição A.4. *Seja T um operador linear sobre o espaço vetorial V . Então:*

1. O **Núcleo de T** é o conjunto

$$Nuc_T = \{v \in V | T(v) = 0\}.$$

2. A **Imagem de T** é o conjunto

$$Im_T = \{v \in V | \exists u \in V : T(u) = v\}.$$

Os conjuntos Nuc_T e Im_T são subespaços vetoriais do espaço V . A demonstração deste fato é simples, bastando que se verifiquem as condições da Proposição (A.2). As dimensões destes subespaços se relacionam com a dimensão do espaço V conforme indicado na proposição a seguir.

Proposição A.3. *Sejam T um operador linear sobre o espaço vetorial V . Então*

$$\dim V = \dim \text{Nuc}_T + \dim \text{Im}_T.$$

Exemplo A.6. Considere o operador linear

$$T : (x, y) = (x - y, x + y).$$

Vamos determinar Nuc_T e Im_T e as respectivas dimensões destes subespaços.

$$\text{Nuc}_T = \{u = (x, y) \in \mathbb{R}^2 : T(x, y) = (x - y, x + y) = 0\},$$

de onde encontramos o sistema

$$\begin{cases} x - y = 0 \\ x + y = 0 \end{cases}$$

cuja solução é apenas o vetor $(0, 0)$. Assim, $\text{Nuc}_T = \{(0, 0)\}$ e

$$\dim \text{Nuc}_T = 0.$$

Ainda por definição de imagem de um operador linear, segue que

$$\text{Im}_T = \{(a, b) \in \mathbb{R}^2 : \exists u = (x, y) \in \mathbb{R}^2 \rightarrow T(x, y) = (a, b)\}$$

de onde encontramos o sistema

$$\begin{cases} x - y = a \\ x + y = b \end{cases}$$

cuja solução é $(x, y) = \left(\frac{a+b}{2}, \frac{b-a}{2}\right)$. Observando que

$$\left(\frac{a+b}{2}, \frac{b-a}{2}\right) = \frac{1}{2}a(1, -1) + \frac{1}{2}b(1, 1)$$

encontramos

$$\dim \text{Im}_T = 2.$$

Note a validade da proposição (A.3).

O que torna a estrutura dos operadores lineares tão simples é o fato de que, se for conhecida a ação do operador sobre a base do espaço V então se conhece a ação do operador sobre todo o espaço V .

De fato, admita que $V = \mathbb{R}^d$ e considere $\mathcal{B} = \{v_1, v_2, \dots, v_d\}$ uma base de V . Então, dado um vetor $w \in \mathbb{R}^d$ existem escalares reais $\alpha_1, \alpha_2, \dots, \alpha_d$ tais que

$$w = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_d v_d.$$

Aplicando agora o operador T (definido em \mathbb{R}^d) no vetor w podemos escrever da equação acima

$$T(w) = \alpha_1 T(v_1) + \alpha_2 T(v_2) + \dots + \alpha_d T(v_d).$$

Seja $\mathcal{B} = \{v_1, v_2, \dots, v_d\}$ uma base do espaço V . Dado o operador linear $T : V \rightarrow V$, definimos a **matriz de T com relação à base \mathcal{B}** e escrevemos $T_{[\mathcal{B}]}$ como a matriz cujas colunas são formadas pelos escalares $\alpha_{1j}, \alpha_{2j}, \dots, \alpha_{dj}$ que aparecem quando escrevemos a imagem do vetor $v_j \in \mathcal{B}$ como combinação linear dos elementos de \mathcal{B} .

Exemplo A.7. Considere o operador linear $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definido por

$$T(x, y, z) = (2x + y - z, y - z, x + 7z)$$

e seja \mathcal{B} a base canônica de \mathbb{R}^3 . Então, como

$$T(1, 0, 0) = (2, 0, 1)$$

$$T(0, 1, 0) = (1, 1, 0)$$

$$T(0, 0, 1) = (-1, -1, 7)$$

a matriz $T_{[\mathcal{B}]}$ do operador T é

$$\begin{pmatrix} 2 & 1 & -1 \\ 0 & 1 & -1 \\ 1 & 0 & 7 \end{pmatrix}.$$

Considere agora duas bases, \mathcal{B}_1 e \mathcal{B}_2 do espaço vetorial V . Faremos a seguir, uma construção que permite relacionar as coordenadas de um vetor v em relação à base \mathcal{B}_1 com as coordenadas do mesmo vetor v em relação à base \mathcal{B}_2 .

Considere $\mathcal{B}_1 = \{v_1, v_2, \dots, v_d\}$ e $\mathcal{B}_2 = \{w_1, w_2, \dots, w_d\}$ bases de V . Sendo $v \in V$ um vetor qualquer, este será escrito como combinação linear dos vetores de \mathcal{B}_1 e de \mathcal{B}_2 :

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_d v_d \tag{A.1}$$

e

$$v = \beta_1 w_1 + \beta_2 w_2 + \dots + \beta_d w_d \tag{A.2}$$

Por outro lado, os vetores da base \mathcal{B}_1 podem ser escritos em relação à base \mathcal{B}_2 , isto é,

$$v_1 = a_{11} w_1 + a_{21} w_2 + \dots + a_{d1} w_d; \tag{A.3}$$

$$v_2 = a_{12} w_1 + a_{22} w_2 + \dots + a_{d2} w_d; \tag{A.4}$$

⋮

$$v_d = a_{1d} w_1 + a_{2d} w_2 + \dots + a_{dd} w_d; \tag{A.5}$$

Substituindo (A.3), (A.4) e (A.5) em (A.1), temos:

$$v = \alpha_1 (a_{11} w_1 + a_{21} w_2 + \dots + a_{d1} w_d) + \alpha_2 (a_{12} w_1 + a_{22} w_2 + \dots + a_{d2} w_d) + \dots + \alpha_d (a_{1d} w_1 + a_{2d} w_2 + \dots + a_{dd} w_d)$$

e portanto,

$$v = (a_{11}\alpha_1 + a_{12}\alpha_2 + a_{1d}\alpha_d)w_1 + (a_{21}\alpha_1 + a_{22}\alpha_2 + \dots + a_{2d}\alpha_d)w_2 + \dots + (a_{d1}\alpha_1 + a_{d2}\alpha_2 + \dots + a_{dd}\alpha_d)w_d$$

e comparando (??) com (A.2) segue que

$$\begin{aligned} \beta_1 &= a_{11}\alpha_1 + a_{12}\alpha_2 + \dots + a_{1d}\alpha_d; \\ \beta_2 &= a_{21}\alpha_1 + a_{22}\alpha_2 + \dots + a_{2d}\alpha_d; \\ &\vdots \\ \beta_d &= a_{d1}\alpha_1 + a_{d2}\alpha_2 + \dots + a_{dd}\alpha_d \end{aligned}$$

ou, na forma matricial,

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_d \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1d} \\ a_{21} & a_{22} & \dots & a_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ a_{d1} & a_{d2} & \dots & a_{dd} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_d \end{pmatrix},$$

que mais simplesmente se escreve como

$$[v]_{\mathcal{B}_2} = [T]_{\mathcal{B}_2}^{\mathcal{B}_1} [v]_{\mathcal{B}_1}$$

sendo a matriz

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1d} \\ a_{21} & a_{22} & \dots & a_{2d} \\ \dots & \dots & \ddots & \dots \\ a_{d1} & a_{d2} & \dots & a_{dd} \end{pmatrix}$$

chamada de **matriz de mudança de base** de \mathcal{B}_1 para \mathcal{B}_2 .

Observação A.2. Apesar de ser possível construir bases distintas para um mesmo espaço V , a noção de matriz de um operador $T : V \rightarrow V$ está bem definida. De fato, dadas $T_{[\mathcal{B}]}$ e $T_{[\mathcal{A}]}$, duas representações matriciais distintas do operador V , existe uma única matriz $P \in M_d$ inversível tal que

$$T_{[\mathcal{B}]} = PT_{[\mathcal{A}]}P^{-1}.$$

A matriz P da Observação (A.2) é chamada de matriz de mudança da base A para a base B . As matrizes $T_{[\mathcal{B}]}$ e $T_{[\mathcal{A}]}$ são chamadas de matrizes semelhantes.

Definição A.5. *Seja $T : V \rightarrow V$ um operador linear. Então definimos:*

1. Um **autovalor** de T é um elemento de $\lambda \in \mathbb{C}$ tal que existe um vetor $v \in V$ não nulo com $T(v) = \lambda v$;

2. Seja λ um autovalor do operador T ; então todo vetor não nulo $v \in V$ tal que $T(v) = \lambda v$ é chamado de **autovetor de T associado a λ** . Denotamos por $\text{Aut}_T(\lambda)$ o subespaço de V gerado por todos os autovetores associados a λ ;
3. O operador T é **diagonalizável** se existe uma base \mathcal{B} de V formada por autovetores de T .

Considere $T_{[\mathcal{B}]}$ a representação matricial do operador T em relação à base \mathcal{B} . Então, se λ é autovalor de T , segue que

$$T_{[\mathcal{B}]}v = \lambda v,$$

onde v é um autovetor associado. Desta equação, deduzimos que

$$T_{[\mathcal{B}]}v - \lambda v = 0,$$

ou ainda

$$(T_{[\mathcal{B}]} - \lambda Id_d)v = 0$$

e então este sistema tem solução não nula (já que pela definição $v \neq 0$) se e somente se

$$\det(T_{[\mathcal{B}]} - \lambda Id_d) = 0,$$

onde Id_d é a matriz identidade em M_d .

Portanto, temos o seguinte resultado:

Lema A.2. *Seja $T : V \rightarrow V$ um operador linear tal que sua representação em uma base \mathcal{B} qualquer seja $T_{[\mathcal{B}]}$. Os autovalores λ de T são as raízes do polinômio*

$$p_T(\lambda) = \det(T_{[\mathcal{B}]} - \lambda Id_d) = 0 \tag{A.6}$$

Ao polinômio (A.6) dá-se o nome de **polinômio característico** do operador T .

O polinômio $p_T(\lambda)$ independe da escolha da base \mathcal{B} do espaço V . De fato, sejam \mathcal{B} e \mathcal{B}' duas bases de V ; pela observação (A.2) existe $P \in M_d$ tal que $T_{[\mathcal{B}]} = PT_{[\mathcal{B}']}P^{-1}$ e portanto,

$$\begin{aligned} \det(T_{[\mathcal{B}]} - \lambda Id_d) &= \det(PT_{[\mathcal{B}']}P^{-1} - \lambda PId_dP^{-1}) = \\ \det P(T_{[\mathcal{B}']} - \lambda Id_d)P^{-1} &= \det P \cdot \det(T_{[\mathcal{B}']} - \lambda Id_d) \cdot \det P^{-1} = \\ \det(T_{[\mathcal{B}']} - \lambda Id_d). \end{aligned}$$

Lema A.3. *Autovetores v_1, v_2, \dots, v_d associados a autovalores distintos são linearmente independentes.*

Demonstração. Por definição, todo autovetor é não nulo e portanto o resultado do lema é trivial para apenas um autovetor. Por contrapositiva, basta provar que dois ou mais autovetores linearmente dependentes têm pelo menos um autovalor igual.

Portanto, sejam $v_1, v_2 \in \mathbb{R}^d$ autovetores linearmente dependentes de $A \in M_d$ com autovalores associados λ_1, λ_2 respectivamente. Assim, existe $a \in \mathbb{R}$ tal que $a \neq 0$ e $v_2 = av_1$ e então

$$\lambda_2 v_2 = Av_2 = aAv_1 = a\lambda_1 v_1 = \lambda_1 av_1 = \lambda_1 v_2,$$

ou seja, $(\lambda_2 - \lambda_1)v_2 = 0 \in \mathbb{R}^d$ e como v_2 é não nulo segue que $\lambda_1 = \lambda_2$ e a afirmação está provada para dois vetores.

Por indução, vamos supor que o resultado seja verdadeiro para $k - 1$ autovetores, com $1 < k - 1 < d$ fixado. Considere $v_1, v_2, \dots, v_{k-1} \in \mathbb{R}^d$ autovetores linearmente dependentes de A com autovalores $\lambda_1, \lambda_2, \dots, \lambda_{k-1}$ respectivamente.

Se v_1, v_2, \dots, v_{k-1} são linearmente dependentes então existem dois autovalores iguais, por hipótese de indução e não há nada a demonstrar. Vamos portanto supor que v_1, v_2, \dots, v_{k-1} seja linearmente independentes, de onde decorre que v_k é uma combinação linear não nula deles e portanto,

$$v_k = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_{k-1} v_{k-1}$$

para algum $\alpha_i \neq 0$. Por outro lado, multiplicando esta equação por λ_k obtemos

$$\lambda_k v_k = \lambda_k \alpha_1 v_1 + \lambda_k \alpha_2 v_2 + \dots + \lambda_k \alpha_{k-1} v_{k-1}$$

e aplicando A segue que

$$\lambda_k v_k = Av_k = \alpha_1 Av_1 + \alpha_2 Av_2 + \dots + \alpha_{k-1} Av_{k-1}.$$

Subtraindo as duas últimas igualdades termo a termo e reordenando as parcelas resulta

$$0 = \alpha_1(\lambda_k - \lambda_1)v_1 + \alpha_2(\lambda_k - \lambda_2)v_2 + \dots + \alpha_{k-1}(\lambda_k - \lambda_{k-1})v_{k-1}.$$

Como estamos supondo que v_1, v_2, \dots, v_{k-1} são linearmente independentes decorre que $\alpha_i(\lambda_k - \lambda_i) = 0$ para cada $1 \leq i \leq k - 1$ e como algum $\alpha_i \neq 0$ resulta que $\lambda_k = \lambda_i$ para algum $1 \leq i \leq k - 1$, o que conclui a demonstração. \square

Definição A.6. *Seja λ um autovalor do operador $T : V \rightarrow V$.*

1. *A multiplicidade algébrica de λ é a multiplicidade de λ enquanto raiz do polinômio característico $p_T(\lambda)$;*
2. *A multiplicidade geométrica de λ é a dimensão do subespaço de V gerado pelos autovetores associados a λ .*

O resultado a seguir é bastante notável. Toda matriz A anula seu polinômio característico:

Teorema A.1. *(Cayley-Hamilton) Um operador linear $T : V \rightarrow V$ é um zero de seu polinômio característico, isto é*

$$p_T(T) = 0.$$

Exemplo A.8. Considere o operador $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ cuja matriz em relação a uma certa base é

$$A = \begin{pmatrix} 3 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 3 \end{pmatrix}$$

seu polinômio característico é dado por

$$p_T(\lambda) = \det \begin{pmatrix} 3 - \lambda & -1 & 1 \\ -1 & 5 - \lambda & -1 \\ 1 & -1 & 3 - \lambda \end{pmatrix}$$

que após o desenvolvimento do determinante e redução de termos se escreve como

$$p_T(\lambda) = \lambda^3 - 11\lambda^2 + 36\lambda - 36$$

cujas raízes são

$$\lambda_1 = 2,$$

$$\lambda_2 = 3,$$

$$\lambda_3 = 6.$$

Os autovetores associados a λ_i , com $i = 1, 2, 3$ são dados pelos vetores v_i soluções dos sistemas

$$(A - \lambda_i Id_d)v_i = 0 \quad i = 1, 2, 3,$$

ou seja,

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},$$

$$v_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

$$v_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

Observe que $\mathcal{P} = \{v_1, v_2, v_3\}$ é uma base de \mathbb{R}^3 , de modo que o operador T é diagonalizável sendo a matriz

$$P = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -2 \\ -1 & 1 & 1 \end{pmatrix}$$

cuja inversa é

$$P^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6} & -\frac{1}{3} & \frac{1}{6} \end{pmatrix}$$

a matriz de mudança de base para a base P na qual o operador T admite a forma diagonal.

As colunas de P são formadas pelos autovetores do operador T e portanto, de acordo com a Observação (A.2) podemos determinar a forma diagonal D do operador T a partir da equação

$$D = PAP^{-1}$$

ou seja

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -2 \\ -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 3 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 3 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6} & -\frac{1}{3} & \frac{1}{6} \end{pmatrix}$$

Nem todo operador T é diagonalizável, uma vez que nem sempre é possível determinar uma base \mathcal{B} formada por autovetores de T . Nestes casos contudo, é possível encontrar uma forma de escrever a matriz A do operador T de modo que ao longo da diagonal apareçam matrizes quadradas, menores e mais simples que A . Trata-se da **Forma de Jordan** do operador T .

Definição A.7. *Seja λ um autovalor do operador T . Então:*

1. *O bloco de Jordan real do autovalor λ e tamanho l é*

$$J_{\lambda}l = \begin{pmatrix} \lambda & 0 & 0 & \dots & 0 & 0 \\ 1 & \lambda & 0 & \dots & 0 & 0 \\ 0 & 1 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & \lambda \end{pmatrix}$$

2. *O bloco de Jordan complexo do autovalor $\lambda = a + ib$, $b \neq 0$ e tamanho $2l$ é*

$$J_{a,b}(2l) = \begin{pmatrix} J_{a,b} & 0 & 0 & \dots & 0 & 0 \\ Id_2 & J_{a,b} & \dots & 0 & 0 & \\ 0 & Id_2 & J_{a,b} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & Id_2 & J_{a,b} \end{pmatrix}$$

onde 0 e Id_2 indicam respectivamente a matriz nula e a matriz identidade em M_2

e $J_{a,b} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.

O teorema a seguir garante a existência da Forma de Jordan, mas não o demonstraremos¹

Teorema A.2. (*Forma Canônica de Jordan*) Se $A \in M_d$ então A pode ser escrita como a matriz real

$$J = \text{diag}(J_1, J_2, \dots, J_r) \in M_d$$

onde cada J_i é um bloco de Jordan real ou complexo. A matriz J é única, a menos da ordem dos blocos da diagonal.

Todos os autovalores de A necessariamente estão representados nos blocos, mas o mesmo autovalor pode repetir em mais de um bloco, de modo que o número de blocos varia desde o número de autovalores distintos de A até no máximo d , caso em que A é diagonalizável.

A construção das possíveis Formas de Jordan de um operador é bastante simplificada com a noção de **polinômio minimal**. O Teorema (A.1) estabelece que $p_T(T) = 0$; contudo, este pode não ser o polinômio de menor grau que anula T .

Definição A.8. O polinômio minimal do operador $T : V \rightarrow V$ é o polinômio mônico de menor grau que possui as mesmas raízes do polinômio característico de T , tal que

$$m_T(T) = 0$$

A definição acima garante que o polinômio minimal é um divisor do polinômio característico. Veja o exemplo:

Exemplo A.9. Considere o operador linear $T : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ tal que

$$T(x_1, x_2, x_3, x_4) = (8x_1 - x_2, 4x_1 + 12x_2, 9x_3 + 2x_4, 2x_3 + 6x_4),$$

cuja matriz em relação à base canônica é

$$\begin{pmatrix} 8 & -1 & 0 & 0 \\ 4 & 12 & 0 & 0 \\ 0 & 0 & 9 & 2 \\ 0 & 0 & 2 & 6 \end{pmatrix}$$

e cujo polinômio característico é

$$p_T(\lambda) = (\lambda - 10)^3(\lambda - 5). \quad (\text{A.7})$$

As possibilidades para o polinômio minimal $m_T(\lambda)$ são

$$\begin{aligned} m_T(\lambda) &= (\lambda - 10)(\lambda - 5), \\ m_T(\lambda) &= (\lambda - 10)^2(\lambda - 5), \end{aligned} \quad (\text{A.8})$$

ou

$$m_T(\lambda) = (\lambda - 10)^3(\lambda - 5).$$

Verificando as equações acima, vemos que o minimal é dado por (A.8).

¹A demonstração pode ser encontrada na referência [11].

É possível mostrar que os expoentes do polinômio minimal do operador T determinam o tamanho dos blocos de Jordan que aparecem na Forma de Jordan de T .

No exemplo acima, o polinômio minimal encontrado foi $m_T(\lambda) = (\lambda - 10)^2(\lambda - 5)$, de modo que a Forma de Jordan deste operador terá um bloco de Jordan $J_2(10)$ e um bloco de Jordan $J_1(5)$. Portanto, neste caso temos

$$J = \begin{pmatrix} 10 & 0 & 0 & 0 \\ 1 & 10 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}.$$

A.2 Exponencial de matrizes

Seja M_d o espaço das matrizes quadradas de ordem d , de dimensão $\dim(M_d) = d^2$.

Para o que segue, seja $\|\cdot\|$ a norma sobre M_d definida por

$$\|A\| = \sup_{\|v\| \leq 1} \|A(v)\|.$$

Dados $A, B \in M_d$, segue que:

$$\|AB\| = \sup_{\|v\| \leq 1} \|A(B(v))\| = \sup_{\|v\| \leq 1} \left\| A \left(\frac{B(v)}{\|B(v)\|} \right) \right\| \cdot \|B(v)\|$$

ou seja,

$$\|AB\| \leq \sup_{\|v\| \leq 1} \left\| A \left(\frac{B(v)}{\|B(v)\|} \right) \right\| \cdot \sup_{\|v\| \leq 1} \|B(v)\|$$

de onde

$$\|AB\| \leq \|A\| \cdot \|B\| \tag{A.9}$$

A desigualdade (A.9) em particular significa que $\left\| \frac{A^n}{n!} \right\| \leq \frac{\alpha^n}{n!}$ com $\alpha = \|A\|$ e com ela podemos mostrar o resultado a seguir:

Lema A.4. A série $\sum_{n=0}^{\infty} \frac{A^n}{n!}$ é absolutamente convergente para todo $A \in M_d$.

Demonstração. Basta notar que o termo geral de $\sum_{n=0}^{\infty} \frac{A^n}{n!}$ é dominado em norma pelo

termo geral da série $\sum_{n=0}^{\infty} \frac{\alpha^n}{n!}$ que converge para a exponencial real e^α ; portanto, pelo

critério de comparação², a série $\sum_{n=0}^{\infty} \frac{A^n}{n!}$ é convergente. \square

²Suponha que $\sum a_n$ e $\sum b_n$ sejam séries de termos positivos. Se existem $c > 0$ e $n_0 \in \mathbb{N}$ tais que $a_n \leq c \cdot b_n$ para todo $n > n_0$ então a convergência de $\sum b_n$ implica a de $\sum a_n$ enquanto a divergência de $\sum a_n$ implica a de $\sum b_n$.

Pelas propriedades que apresenta, a série $\sum_{n=0}^{\infty} \frac{A^n}{n!}$ é chamada de **exponencial da matriz** A e denotada por e^A . No Lema a seguir estão listadas algumas delas:

Lema A.5. A série $e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}$ possui as seguintes propriedades:

1. $\|e^A\| \leq e^{\|A\|}$, para qualquer matriz $A \in M_d$.
2. Dada a matriz P invertível em M_d com inversa P^{-1} e qualquer matriz $A \in M_d$ então $e^{PAP^{-1}} = Pe^AP^{-1}$.
3. $e^AA = Ae^A$, para qualquer matriz $A \in M_d$.
4. Se $A, B \in M_d$ comutam, isto é, se $AB = BA$, então $e^{A+B} = e^{B+A}$.
5. Sendo $0 \in M_d$ a matriz nula de ordem d , temos $e^0 = Id$, onde Id é a matriz identidade de M_d .
6. A aplicação $t \in \mathbb{R} \mapsto e^{tA}$ é derivável em relação a t e:

$$\frac{d}{dt} e^{tA} = Ae^{tA},$$

ou seja, $\gamma(t) = e^{tA}$ é solução da equação diferencial $X' = AX$ ($X \in M_d$).

Demonstração. Demonstraremos apenas as afirmações 2, 3 e 6: (2) Por definição,

$$e^{PAP^{-1}} = \sum_{n=0}^{\infty} \frac{A^{PAP^{-1}}}{n!} = Id + PAP^{-1} + \frac{(PAP^{-1})^2}{2} + \frac{(PAP^{-1})^3}{3!} + \dots + \frac{(PAP^{-1})^n}{n!} + \dots$$

e notando que

$$\begin{aligned} (PAP^{-1})^2 &= PAP^{-1}PAP^{-1} = PA^2P^{-1}, \\ (PAP^{-1})^3 &= PAP^{-1}PAP^{-1}PAP^{-1} = PA^2P^{-1} = PA^3P^{-1}, \\ &\vdots \\ (PAP^{-1})^n &= PAP^{-1}PAP^{-1} = PA^nP^{-1}, \end{aligned}$$

temos

$$e^{PAP^{-1}} = Id + PAP^{-1} + \frac{(PA^2P^{-1})}{2} + \frac{(PA^3P^{-1})}{3!} + \dots + \frac{(PA^nP^{-1})}{n!} + \dots$$

de onde, colocando em evidência à esquerda a matriz P e à direita a matriz P^{-1} segue que

$$e^{PAP^{-1}} = P \left(Id + A + \frac{A^2}{2} + \dots + \frac{A^n}{n!} + \dots \right) P^{-1}$$

ou seja,

$$e^{PAP^{-1}} = Pe^AP^{-1}.$$

Para provar a afirmação (3), basta notar que

$$e^A A = A + A^2 + \frac{A^3}{2} + \frac{A^4}{3!} + \dots + \frac{A^{n+1}}{n!} + \dots,$$

de onde, colocando A em evidência do lado esquerdo da última igualdade segue que

$$e^A A = A \left(Id + A + \frac{A^2}{2} + \frac{A^3}{3!} + \dots + \frac{A^n}{n!} + \dots \right) = A e^A.$$

Finalmente, para a demonstração da afirmação (6) observamos que é válida a derivação termo a termo³ para $\sum_{n=0}^{\infty} \frac{A^n}{n!}$ (que pelo Lema B.4 é absolutamente convergente), e portanto

$$\begin{aligned} \frac{d}{dt} e^{tA} &= \frac{d}{dt} \left(Id + tA + \frac{t^2 A^2}{2} + \frac{t^3 A^3}{3!} + \dots + \frac{t^n A^n}{n!} + \dots \right) = A + tA^2 + \frac{t^2 A^3}{2} + \dots + \frac{t^{n-1} A^n}{(n-1)!} + \dots \\ &= A \left(Id + tA + \frac{t^2 A^2}{2} + \dots + \frac{t^{n-1} A^{n-1}}{(n-1)!} \right) = A e^{tA}. \end{aligned}$$

□

Os próximos exemplos ilustram o cálculo da exponencial e^A para algumas classes de matrizes.

Exemplo A.10. (Exponencial de matrizes diagonais⁴). Considere a matriz

$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

. Calculando D^2 e D^3 temos:

$$D^2 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}^2 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{pmatrix}$$

$$D^3 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}^3 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} \lambda_1^3 & 0 \\ 0 & \lambda_2^3 \end{pmatrix}$$

e de maneira indutiva

$$D^n = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}^n = \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix},$$

de onde é fácil obter a exponencial de D :

³Se a série de potências $f(x) = \sum a_n x^n$ converge absolutamente para $|x| < R$ então $g(x) = \sum n a_n x^{n-1}$ também é absolutamente convergente em $|x| < R$ e $f'(x) = g(x)$ para todo x no intervalo $-R < x < R$.

⁴Este exemplo pode ser generalizado para matrizes em blocos; dada a matriz M em cuja diagonal aparecem os blocos de matrizes A_1, A_2, \dots, A_d , ou seja, se $M = \text{diag}(A_1, A_2, \dots, A_d)$, então de maneira análoga ao caso das matrizes diagonais segue que $e^M = \text{diag}(e^{A_1}, e^{A_2}, \dots, e^{A_d})$, de modo que a exponencial de M é obtida tomando a exponencial de cada bloco individualmente ao longo da diagonal.

$$e^D = Id + D + \frac{D^2}{2} + \frac{D^3}{3!} + \dots + \frac{D^n}{n!} + \dots =$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} + \begin{pmatrix} \frac{\lambda_1^2}{2} & 0 \\ 0 & \frac{\lambda_2^2}{2} \end{pmatrix} + \dots + \begin{pmatrix} \frac{\lambda_1^n}{n!} & 0 \\ 0 & \frac{\lambda_2^n}{n!} \end{pmatrix} = \begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix}.$$

De maneira geral, dada uma matriz $A \in M_d$ diagonal com $\lambda_i = a_{ii}$, para $i = 1, 2, 3, \dots, d$ não necessariamente distintos, isto é, supondo $A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_d)$ segue que

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n = \sum_{n=0}^{\infty} \frac{1}{n!} \text{diag}(\lambda_1^n, \lambda_2^n, \dots, \lambda_d^n) = \text{diag}(e^{\lambda_1}, e^{\lambda_2}, \dots, e^{\lambda_d}).$$

Exemplo A.11. (Exponencial de matrizes nilpotentes) Uma matriz $N \in M_d$ (não nula) é dita nilpotente se existe um inteiro $k \geq 2$ tal que $N^k = 0$. As matrizes nilpotentes são aquelas onde apenas a subdiagonal inferior ou superior são não nulas; no restante desta dissertação usaremos a notação

$$N_c(d) = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ c & 0 & \dots & 0 & 0 \\ 0 & c & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & c & 0 \\ 0 & 0 & \dots & 0 & c \end{pmatrix}.$$

para nos referirmos à matriz nilpotente de ordem d em cuja subdiagonal inferior ou apareça a constante c real ou complexa.

O menor inteiro k tal que $N^k = 0$ é chamado de índice de nilpotência da matriz N . Neste caso, apenas as matrizes N, N^2, \dots, N^{k-1} são não nulas e a exponencial e^D se torna uma soma finita:

$$e^{N_c(d)} = Id + \frac{1}{2} N_c(d)^2 + \dots + \frac{1}{k-1} N_c(d)^{k-1},$$

cujo cálculo não é complicado se o índice de nilpotência for pequeno.

A exponencial de uma matriz nilpotente pode ser obtida por indução sobre a ordem d da matriz. Segue que

$$e^{N_c(d)} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ c & 1 & 0 & \dots & 0 & 0 \\ \frac{c^2}{2!} & c & 1 & \dots & 0 & 0 \\ \frac{c^3}{3!} & \frac{c^2}{2!} & c & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{c^{d-1}}{(d-1)!} & \frac{c^{d-2}}{(d-2)!} & \frac{c^{d-3}}{(d-3)!} & \dots & c & 1 \end{pmatrix}.$$

Exemplo A.12. (Exponencial de Blocos de Jordan Reais) De acordo com (A.7) o bloco de Jordan real do autovalor $\lambda \in R$ de tamanho l é a matriz

$$J_{\lambda} l = \begin{pmatrix} \lambda & 0 & 0 & \dots & 0 & 0 \\ 1 & \lambda & 0 & \dots & 0 & 0 \\ 0 & 1 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & \lambda \end{pmatrix},$$

que pode ser decomposta como a soma de uma matriz diagonal λI_d e uma matriz nilpotente $N_1(d)$ que comutam, ou seja, são tais que $\lambda I_d N_1(d) = N_1(d) \lambda I_d$, de modo que pelo item 4 do Lema (A.5)

$$e^{J_{\lambda}(l)} = e^{\lambda I_d + N_1(d)} = e^{\lambda I_d} e^{N_1(d)} = e^{\lambda} e^{N_1(d)}. \quad (\text{A.10})$$

Segue dos Exemplos B.10 e B.11 que

$$e^{J_{\lambda}(l)} = e^{\lambda} \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 1 & 0 & \dots & 0 & 0 \\ \frac{1}{2!} & 1 & 1 & \dots & 0 & 0 \\ \frac{1}{3!} & \frac{1}{2!} & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{(l-1)!} & \frac{1}{(l-2)!} & \frac{1}{(l-3)!} & \dots & 1 & 1 \end{pmatrix}. \quad (\text{A.11})$$

Exemplo A.13. Considere a matriz $A = \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} \in M_2$ sendo b um real qualquer.

Observe que

$$A^2 = \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} = \begin{pmatrix} -b^2 & 0 \\ 0 & -b^2 \end{pmatrix}$$

e

$$A^3 = \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} = \begin{pmatrix} 0 & -b^3 \\ b^3 & 0 \end{pmatrix}.$$

Por indução, podemos provar que para potências pares

$$\begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}^{2j} = (-1)^j \begin{pmatrix} b^{2j} & 0 \\ 0 & b^{2j} \end{pmatrix} \quad (\text{A.12})$$

e para as ímpares

$$\begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}^{2j+1} = (-1)^j \begin{pmatrix} 0 & b^{2j+1} \\ -b^{2j+1} & 0 \end{pmatrix}. \quad (\text{A.13})$$

Segue de (A.12) e de (A.13) que

$$e^A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} b^{2j} & 0 \\ 0 & b^{2j} \end{pmatrix} + \dots = \begin{pmatrix} \cos b & \operatorname{sen} b \\ -\operatorname{sen} b & \cos b \end{pmatrix}.$$

Exemplo A.14. Vamos generalizar o exemplo anterior de modo que consigamos calcular agora a exponencial de

$$J_{a,b} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \in M_2,$$

sendo a, b reais quaisquer.

Observe que

$$J_{a,b} = \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} + \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$$

e as matrizes que figuram nas parcelas desta soma comutam, isto é:

$$\begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} = \begin{pmatrix} 0 & ab \\ -ab & 0 \end{pmatrix}$$

e

$$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} = \begin{pmatrix} 0 & ab \\ -ab & 0 \end{pmatrix}$$

de modo que novamente pelo item 4 do Lema (A.5) e Exemplos (B.10) e (B.12) temos

$$e^{J_{a,b}} = \begin{pmatrix} e^a & 0 \\ 0 & e^a \end{pmatrix} \begin{pmatrix} \cos b & \operatorname{sen} b \\ -\operatorname{sen} b & \cos b \end{pmatrix} = e^a \begin{pmatrix} \cos b & \operatorname{sen} b \\ -\operatorname{sen} b & \cos b \end{pmatrix},$$

que corresponde à uma matriz de rotação $\theta = b$, cuja notação usual é $e^{J_{a,b}} = e^a R_b$.

Exemplo A.15. (Exponencial de Blocos de Jordan Associados a Autovetores Complexos) Conforme a Definição (A.7), o bloco de Jordan complexo do autovalor $\lambda = a + ib \in \mathbb{C}$ de tamanho $2l$ é

$$J_{a,b}(2l) = \begin{pmatrix} J_{a,b} & 0 & 0 & \dots & 0 & 0 \\ Id_2 & J_{a,b} & \dots & 0 & 0 & \\ 0 & Id_2 & J_{a,b} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & Id_2 & J_{a,b} \end{pmatrix},$$

onde 0 e Id_2 indicam respectivamente a matriz nula e a matriz identidade em M_2 e

$$J_{a,b} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

De maneira análoga ao que foi apresentado no exemplo (B.12), vamos escrever

$$J_{a,b}(l) = J_{a,b}^0(l) + G_{1,Id_2}(l),$$

onde

$$J_{a,b}^0(l) = \begin{pmatrix} J_{a,b} & 0 & 0 & \dots & 0 & 0 \\ 0 & J_{a,b} & \dots & 0 & 0 & \\ 0 & 0 & J_{a,b} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & J_{a,b} \end{pmatrix}$$

e

$$G_{1,Id_2}(l) = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ Id_2 & 0 & 0 & \dots & 0 & 0 \\ 0 & Id_2 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & Id_2 \end{pmatrix}.$$

Como $G_{1,Id_2}(l)$ é nilpotente, segue do Exemplo (B.11) que

$$e^{G_{1,Id_2}(l)} = \begin{pmatrix} Id_2 & 0 & 0 & \dots & 0 & 0 \\ Id_2 & Id_2 & 0 & \dots & 0 & 0 \\ \frac{1}{2!}Id_2 & Id_2 & Id_2 & \dots & 0 & 0 \\ \frac{c^2}{3!}Id_2 & \frac{c^2}{2!}Id_2 & Id_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{c^{l-1}}{(l-1)!}Id_2 & \frac{c^{l-2}}{(l-2)!}Id_2 & \frac{c^{l-3}}{(l-3)!}Id_2 & \dots & Id_2 & Id_2 \end{pmatrix}$$

e dos Exemplos (B.10) e (B.12) que

$$e^{J_{a,b}^0(l)} = \text{diag}(e^{J_{a,b}(l)}, e^{J_{a,b}(l)}, \dots, e^{J_{a,b}(l)}) = e^a \text{diag}(R_b, R_b, \dots, R_b).$$

Como $J_{a,b}^0(l)G_{1,Id_2}(l) = G_{1,Id_2}(l)J_{a,b}^0(l)$, segue do item 4 do Lema (A.5) que

$$e^{J_{a,b}(l)} = e^{J_{a,b}^0(l)} e^{G_{1,Id_2}(l)}$$

e então

$$e^{J_{a,b}(l)} = \begin{pmatrix} R_b & 0 & 0 & \dots & 0 & 0 \\ R_b & R_b & 0 & \dots & 0 & 0 \\ \frac{1}{2!}R_b & R_b & R_b & \dots & 0 & 0 \\ \frac{c^2}{3!}R_b & \frac{c^2}{2!}R_b & R_b & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{c^{l-1}}{(l-1)!}R_b & \frac{c^{l-2}}{(l-2)!}R_b & \frac{c^{l-3}}{(l-3)!}R_b & \dots & R_b & R_b \end{pmatrix}.$$

Exemplo A.16. Considere a matriz

$$A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$$

cujos autovalores são 2 e 3 e estão associados aos autovetores $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ e $v_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ respectivamente.

Como existe uma base de autovetores, A é diagonalizável, isto é, $A = BDB^{-1}$ onde $D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ e $B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. Portanto, pelo item 2 do Lema (A.5) a exponencial e^A pode ser calculada do seguinte modo:

$$e^A = Be^DB^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} e^2 & 0 \\ 0 & e^3 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 2e^2 - e^3 & -e^2 + e^3 \\ 2e^2 - 2e^3 & -e^2 + 2e^3 \end{pmatrix}.$$

Exemplo A.17. Considere a matriz

$$A = \begin{pmatrix} 1 & 0 & -2 \\ -5 & 6 & 11 \\ 5 & -5 & -10 \end{pmatrix}, \quad (\text{A.14})$$

cujo polinômio característico é

$$p(\lambda) = \lambda^3 - 5\lambda^2 + 9\lambda - 5 \quad (\text{A.15})$$

de raízes $\lambda_1 = 1$, $\lambda_2 = 2 + i$ e $\lambda_3 = 2 - i$.

Observe que o polinômio minimal coincide com o característico, de modo que a forma de Jordan de A é:

$$J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & -1 & -2 \end{pmatrix}.$$

Encontremos agora a matriz P tal que $A = PJP^{-1}$.

Vamos determinar um autovetor associado a $\lambda_1 = 1$, ou seja, resolver a equação

$$\begin{pmatrix} 1 & 0 & -2 \\ -5 & 6 & 11 \\ 5 & -5 & -10 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

cuja solução é

$$S = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y, z) = (x, x, 0) = x(1, 1, 0), x \in \mathbb{R}\},$$

de onde encontramos o autovetor $v_1 = (1, 1, 0)$.

Sendo $\lambda_2 = 2 + i$ um autovalor complexo de A , encontramos um autovalor v complexo associado resolvendo

$$Av - \lambda_2 v = 0,$$

que equivale ao sistema linear

$$\begin{cases} (3-i)x + 0y - 2z = 0 \\ -5x + (8-i)y + 11z = 0 \\ 5x - 5y - (8+i)z = 0 \end{cases}$$

Tomando $x = 2$, segue da primeira das equações que

$$6 - 2i - 2z = 0 \Rightarrow z = 3 - i$$

e substituindo $x = 2$ e o valor de z na terceira das equações encontramos

$$10 - 5y - (8+i)(3-i) = 0 \Rightarrow 10 - 5y - (25 - 5i) = 0$$

e portanto

$$y = -3 + i$$

Deste modo, a solução do sistema é

$$S = \{(x, y, z) \in \mathbb{C}^3 \mid (x, y, z) = (2, -3 + i, 3 - i) = (2, -3, 3) + i(0, 1, -1)\} \quad (\text{A.16})$$

Portanto, de (A.16) obtemos os vetores $v_2 = (2, -3, 3)$ e $v_3 = (0, 1, -1)$, linearmente independentes, que juntamente com $v_1 = (1, 1, 0)$ fornecem as colunas da matriz P :

$$P = \begin{pmatrix} 1 & 2 & 0 \\ 1 & -3 & 1 \\ 0 & 3 & -1 \end{pmatrix},$$

cuja inversa é

$$P^{-1} = \begin{pmatrix} 1 & 1 & -1 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{3}{2} & -\frac{3}{2} & \frac{1}{2} \end{pmatrix}.$$

Portanto,

$$\begin{pmatrix} 1 & 0 & -2 \\ -5 & 6 & 11 \\ 5 & -5 & -10 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 \\ 1 & -3 & 1 \\ 0 & 3 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & -1 & -2 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{3}{2} & -\frac{3}{2} & \frac{1}{2} \end{pmatrix} \quad (\text{A.17})$$

de modo que, pelo item 2 do Lema (A.5) encontramos

$$e^A = \begin{pmatrix} 1 & 2 & 0 \\ 1 & -3 & 1 \\ 0 & 3 & -1 \end{pmatrix} \begin{pmatrix} e & 0 & 0 \\ 0 & e^{-2} \cos 1 & e^{-2} \sin 1 \\ 0 & -e^{-2} \sin 1 & e^{-2} \cos 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{3}{2} & -\frac{3}{2} & \frac{1}{2} \end{pmatrix},$$

de onde é possível obter explicitamente e^A :

$$e^A = \begin{pmatrix} e^{-2} \cos 1 + 3e^{-2} \sin 1 & e - e^{-2} \cos 1 - 3e^{-2} \sin 1 & -e + e^{-2} \cos 1 + 3e^{-2} \sin 1 \\ -5e^{-2} \sin 1 & e + 5e^{-2} \sin 1 & -e - e^{-2} \cos 1 - 2e^{-2} \sin 1 \\ 5e^{-2} \sin 1 & -5e^{-2} \sin 1 & e^{-2} \cos 1 + 2e^{-2} \sin 1 \end{pmatrix}.$$

Observação A.3. Uma vez que toda matriz A possui Forma de Jordan J e que, de acordo com os exemplos (B.12) e (B.14), é possível calcular explicitamente a exponencial dos Blocos de Jordan (reais ou complexos), o problema de se determinar a exponencial de uma matriz A está completamente resolvido, apesar de os cálculos serem longos e trabalhosos como o Exemplo acima ilustra.

Referências

- [1] BOYER, C. B. *História da Matemática*. 11. ed. São Paulo: Edgard Blucher, 1974.
- [2] LIMA, E. L. *Espaços Métricos*. 5. ed. Rio de Janeiro: IMPA, 2011.
- [3] LOPES, A. e outro. *Equações Diferenciais Ordinárias*. 5. ed. Rio de Janeiro: IMPA, 2012.
- [4] BOYCE, W. e outro. *Equações Diferenciais Elementares e Problemas de Valores de Contorno*. 8. ed. Rio de Janeiro: Livros Técnicos e Científicos Editora S.A, 2006.
- [5] FIGUEIREDO, D. outro. *Equações Diferenciais Aplicadas*. 3. ed. Rio de Janeiro: IMPA, 2012.
- [6] LIMA, E. L. *Curso de Análise volume 2*. 11. ed. Rio de Janeiro: IMPA, 2012.
- [7] BASSANEZI, R. C. *Temas e Modelos*. 1. ed. Santo André: Universidade Federal do ABC, 2012.
- [8] BAILEY, N. *The Mathematical Theory of Infectious Diseases and Its Applications*. 2. ed. New York: Hafner Press, 1975.
- [9] BARREIRA, L. e outro. *Equações Diferenciais Ordinárias: Teoria Qualitativa*. 1. ed. Lisboa: Editorial Portugal, 2012.
- [10] BRAUN, M. *Differential Equations and their Applications*. 3. ed. New York: Springer-Verlag, 1993.
- [11] LIMA, E. L. *Álgebra Linear*. 1. ed. Rio de Janeiro: IMPA, 1996.
- [12] SOTOMAYOR, J. *Lições de Equações Diferenciais Ordinárias*. 2. ed. Rio de Janeiro: IMPA, 1979.
- [13] COELHO, F. e outro. *Um Curso de Álgebra Linear*. 2. ed. São Paulo: EDUSP, 2005.
- [14] BASSANEZI, R. C. *Ensino Aprendizagem com Modelagem Matemática*. 3. ed. São Paulo: Editora Contexto, 2011.