



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CENTRO DE CIÊNCIAS**  
**DEPARTAMENTO DE FÍSICA**  
**CURSO DE GRADUAÇÃO EM BACHARELADO EM FÍSICA**

**LARA DOMINGOS HISSA**

**ESTUDO DE SISTEMAS DINÂMICOS NÃO LINEARES**

**FORTALEZA**

**2022**

LARA DOMINGOS HISSA

ESTUDO DE SISTEMAS DINÂMICOS NÃO LINEARES

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Bacharelado em Física do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Bacharelado em Física.

Orientador: Prof. Dr. Saulo Davi Soares e Reis

FORTALEZA

2022

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Biblioteca Universitária  
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

H579e Hissa, Lara Domingos.

Estudo de sistemas dinâmicos não lineares / Lara Domingos Hissa. – 2022.  
40 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Ciências,  
Curso de Física, Fortaleza, 2022.

Orientação: Prof. Dr. Saulo Davi Soares e Reis .

1. Sistemas dinâmicos não lineares. 2. Equações governantes. 3. Aprendizado de máquina. 4. Equações de Lorenz. I. Título.

CDD 530

---

LARA DOMINGOS HISSA

ESTUDO DE SISTEMAS DINÂMICOS NÃO LINEARES

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Bacharelado em Física do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Bacharelado em Física.

Aprovada em:

BANCA EXAMINADORA

---

Prof. Dr. Saulo Davi Soares e Reis (Orientador)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Daniel Brito de Freitas  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Erneson Alves de Oliveira  
Universidade de Fortaleza (UNIFOR)

Ao meu amado pai José Salim Hissa Júnior  
(In memoriam)

## AGRADECIMENTOS

Agradeço à Deus e aos meus pais, José Salim Hissa Júnior e Márcia Domingos Saboia Hissa, e ao meu irmão, Vinicius Salim, por todo suporte, apoio e força que me deram neste percurso.

À minha família de Cascavel, principalmente ao meu tio Welligton e minha tia Katia por sempre me incentivarem.

Ao professor Saulo Davi Soares e Reis pela orientação durante a realização deste trabalho e por ter me amparado em vários momentos da graduação.

Aos meus amigos que me acompanharam nesta trajetória acadêmica: Isabel Castro, Carlos Miguel, Francisco Cassimiro, Paulo Cleber e Vasco Stacxak. Em especial aos meus amigos que me deram assistência durante os momentos mais difíceis que enfrentei nesta pandemia: André Domingos, Brena Casimiro, Bruno Bomfim (Nazaky), Gabriel Sousa (careca), Israel Lima (coala), Mateus Falcão, Raul Batista, Rômulo Sampaio e Victor Sales.

Agradeço também pela experiência que tive no Laboratório de Raios X e à todos ali que ajudaram no meu crescimento no início da graduação e pela participação na gestão de 2020 do Diretório Acadêmico.

Ao CNPq pela bolsa e à Universidade Federal do Ceará por ter me proporcionado todas essas vivências.

“Help, I need somebody

Help, not just anybody

Help, you know I need someone, help!”

(The Beatles)

## RESUMO

Neste trabalho apresentamos um estudo sobre sistemas dinâmicos não lineares e o processo qualitativo e quantitativo para encontrar as soluções de suas equações governantes. Encontramos soluções numéricas aproximadas usando o algoritmo Runge-Kutta de quarta ordem. Aplicamos métodos de aprendizado de máquina para encontrar e prever o comportamento do sistema a partir dos dados coletados sobre ele, supondo que a sua equação governante ainda não fosse conhecida. Os métodos utilizados foram as regressões Ridge e Lasso, as quais apresentam um termo de regularização sobre os coeficientes que regem as equações de movimento. Empregamos todos esses mecanismos nas equações de Lorenz. Obtemos pelo Runge-Kutta de quarta ordem as soluções numéricas e o comportamento por meio das equações governantes e dos parâmetros iniciais; pelas regressões obtemos o comportamento e produzimos previsões a partir do conjunto de dados das velocidades e dos estados coletados no sistema. Concluimos que as regressões utilizadas se aproximaram do comportamento previsto pela literatura, mas não foram o suficiente para extrair os coeficientes exatos que acompanham as equações governantes do sistema de tal forma que as identificassem.

**Palavras-chave:** Sistemas dinâmicos não lineares. Equações governantes. Aprendizado de máquina. Equações de Lorenz.



## ABSTRACT

In this work we present a study on nonlinear dynamical systems and the qualitative and quantitative process to find the solutions of their governing equations. We found approximate numerical solutions using the fourth-order Runge-Kutta algorithm. We apply machine learning methods to find and predict the system's behavior from the data collected about it, assuming that its governing equation was not yet known. The methods used were the Ridge and Lasso regressions which present a regularization term on the coefficients that govern the equations of motion. We apply all these mechanisms in the Lorenz equations. We obtain by the fourth order Runge-Kutta the approximate solutions and the behavior through the governing equations and the initial parameters; through regressions we obtain the behavior and produce predictions from the data set of velocities and states collected in the system. We conclude that the regressions used approached the behavior predicted by the literature but were not enough to extract the exact coefficients that accompany the governing equations of the system in such a way as to identify them.

**Keywords:** Nonlinear dynamic systems. Governing equations. Machine Learning. Lorenz equations.

## LISTA DE FIGURAS

Figura 1	– Exemplo da violação do Teorema 2.2.1. . . . .	17
Figura 2	– Retrato de fase da Equação de evolução 2.9. . . . .	20
Figura 3	– Série temporal de $x(t)$ do sistema das Equações de Lorenz. . . . .	23
Figura 4	– Atrator de Lorenz. . . . .	24
Figura 5	– Representação do algoritmo Gradiente Descendente inicializado por um valor aleatório. O algoritmo percorre a função custo a uma taxa de aprendizado ideal $\eta$ até encontrar o seu valor mínimo. . . . .	27
Figura 6	– À esquerda tem a representação do funcionamento do algoritmo Gradiente Descendente para uma taxa de aprendizado muito pequena, fazendo com que o algoritmo demore a chegar no mínimo e à direita para uma taxa de aprendizado muito alta fazendo com que o algoritmo divirja do valor mínimo. . . . .	27
Figura 7	– À direita representa o funcionamento da regressão Lasso. À esquerda representa o funcionamento da regressão Ridge. O diamante e o disco rosa representam a região de restrição das regressões Lasso e Ridge respectivamente. Os contornos elípticos são a função custo da regressão linear. . . . .	30
Figura 8	– Evolução temporal das soluções $x(t)$ , $y(t)$ e $z(t)$ do sistema das equações de Lorenz computadas pelo método Runge-Kutta de quarta ordem. . . . .	31
Figura 9	– Atrator de Lorenz a partir das soluções calculadas pelo método Runge-Kutta de quarta ordem. . . . .	32
Figura 10	– A matriz $\mathbf{V}$ e $\Theta$ são implementadas nos métodos de regressão Ridge e Lasso para encontrar os coeficientes que regem as equações governantes do sistema. . . . .	33
Figura 11	– Evoluções temporais de $\dot{x}$ , $\dot{y}$ e $\dot{z}$ pelos método Runge-Kutta de quarta ordem, regressão Ridge e Lasso juntamente com a ampliação do gráfico no intervalo de tempo $[0,200]$ . . . . .	34
Figura 12	– Comparação das soluções das velocidades para os métodos de regressão com o método Runge-Kutta de quarta ordem. À direita é comparado com a regressão Ridge e à esquerda com a regressão Lasso. . . . .	35
Figura 13	– Gráficos da previsão das evoluções temporais de $\dot{x}$ , $\dot{y}$ e $\dot{z}$ para um período maior. Os gráficos foram plotados para o método Runge-Kutta de quarta ordem e pelos coeficientes encontrados anteriormente pelos métodos de regressão Ridge e Lasso. A ampliação dos gráficos está no intervalo $[0,200]$ . . . . .	36

Figura 14 – Comparação da previsão de soluções das velocidades para os coeficientes computados anteriormente pelas regressões com o método Runge-Kutta de quarta ordem. À direita é comparado com os coeficientes da regressão Ridge e à esquerda pelos coeficientes da regressão Lasso. . . . . 37

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>12</b>
<b>2</b>	<b>SISTEMAS DINÂMICOS NÃO LINEARES</b>	<b>14</b>
<b>2.1</b>	<b>Equações Lineares e não lineares</b>	<b>14</b>
<b>2.2</b>	<b>Plano de Fase</b>	<b>16</b>
<b>2.2.1</b>	<i>Problema de Cauchy</i>	<b>17</b>
<b>2.2.2</b>	<i>Interpretação dos pontos críticos em um retrato de fase</i>	<b>18</b>
<b>2.3</b>	<b>Aspectos quantitativos de um sistema dinâmico não linear</b>	<b>20</b>
<b>2.4</b>	<b>Teoria do caos</b>	<b>21</b>
<b>2.4.1</b>	<i>Equações de Lorenz</i>	<b>22</b>
<b>2.4.1.1</b>	<i>Atrator estranho</i>	<b>22</b>
<b>3</b>	<b>APRENDIZADO DE MÁQUINA</b>	<b>25</b>
<b>3.1</b>	<b>Regressão Linear</b>	<b>25</b>
<b>3.1.1</b>	<i>Função Custo</i>	<b>26</b>
<b>3.1.2</b>	<i>Gradiente Descendente</i>	<b>26</b>
<b>3.1.3</b>	<i>Overfitting</i>	<b>28</b>
<b>3.1.4</b>	<i>Regressão Ridge</i>	<b>28</b>
<b>3.1.5</b>	<i>Regressão Lasso</i>	<b>29</b>
<b>4</b>	<b>RESULTADOS</b>	<b>31</b>
<b>5</b>	<b>CONCLUSÃO</b>	<b>38</b>
	<b>REFERÊNCIAS</b>	<b>39</b>

## 1 INTRODUÇÃO

O estudo da dinâmica começou em meados de 1600, quando Isaac Newton criou as equações diferenciais para descrever o movimento. Após descobrir as leis de movimento e da gravitação universal, ele conseguiu resolver problemas dinâmicos de dois corpos, exemplo disso, foi o cálculo do movimento da Terra ao redor do Sol. A geração seguinte de físicos e matemáticos estenderam os métodos analíticos de Newton para problemas mais abrangentes e complexos, aplicando para sistemas de três corpos, problema de difícil solução analítica. Após décadas, problemas desse tipo começaram a ser trabalhados de forma mais qualitativa. No final do século XIX, Poincaré, desenvolveu um método geométrico que analisava o comportamento do sistema sem precisar resolver analiticamente a equação. Essa resolução trouxe uma abordagem mais moderna da dinâmica (POINCARÉ; MAITLAND, 2003).

Em 1950, a invenção do computador de alta velocidade foi um marco na história da dinâmica. O computador permitiu que fosse possível desenvolver alguma intuição sobre esses sistemas. A experimentação desses sistemas por métodos computacionais levaram à descoberta de Edward Lorenz em 1963, sobre movimentos caóticos em um atrator estranho. Lorenz estudou um modelo simplificado de rolos de convecção na atmosfera para obter uma visão sobre a imprevisibilidade do tempo e descobriu que as soluções para suas equações nunca chegavam em um estado periódico (LORENZ, 1963)(LORENZ, 1995).

Nessa trajetória histórica, compreender as propriedades das equações de movimento se tornou relevante, pois o que diferencia a dificuldade de resolução das equações de um sistema dinâmico é seu aspecto de linearidade. Para equações diferenciais não lineares é praticamente impossível resolver analiticamente, por isso apenas depois da invenção do computador foi possível entender melhor a dinâmica de diversos problemas, já que grande parte dos fenômenos naturais apresentam o aspecto não linear incluindo assuntos que ultrapassam a física, sendo vistos também na biologia e na química.

Dessa forma, um dos grandes desafios atuais de várias áreas da ciência e da engenharia é encontrar as equações governantes em um sistema com dados abundantes. Métodos de aprendizagem de máquina, como a regressão linear com regularização, estão sendo utilizados para descobrir essas equações de sistemas dinâmicos não lineares. Para a realização desse método se faz a suposição de que existem apenas alguns termos importantes que governam a dinâmica, o que é válido para muitos sistemas físicos. Isso resulta em modelos que evitam o *overfitting* sobre os dados (BRUNTON *et al.*, 2016).

A partir disso, neste trabalho estudamos o sistema dinâmico não linear de Lorenz obtendo suas soluções numéricas pelo método Runge-Kutta de quarta ordem e por meio somente desse conjunto de soluções retornamos as suas equações de movimento utilizando a regressão linear com regularização Ridge e Lasso como método de aprendizagem de máquina escolhido.

Dessa forma, este trabalho está organizado com apresentação dos conceitos básicos sobre o estudo de sistemas dinâmicos não lineares no Capítulo 2. Nele é abordado a diferença de equações lineares e não lineares e suas propriedades sendo exemplificado em sistemas físicos conhecidos; métodos para resolver sistemas não lineares no seu aspecto qualitativo e quantitativo; análise do comportamento do sistema dinâmico, se o sistema se comporta em ciclos, se há pontos estáveis ou instáveis; e uma apresentação sobre quando sistemas dinâmicos não lineares apresentam caos, abordando as Equações de Lorenz como um exemplo de sistema caótico.

No Capítulo 3 se trata sobre Machine Learning, trazendo como ponto principal o funcionamento de um dos seus métodos supervisionados: a regressão linear. Nele é mostrado o funcionamento do algoritmo Gradiente Descendente, desenvolvido por meio da função hipótese e função custo. Além disso, é apresentado o problema de *overfitting* que a regressão linear pode acarretar e as soluções para esse problema através das regularizações realizadas pelas regressões Ridge e Lasso. Também são abordadas o funcionamento dessas regressões e suas diferenças.

No Capítulo 4 o comportamento do sistema das equações de Lorenz é obtido pelos métodos Runge-Kutta de quarta ordem, regressão Ridge e Lasso. Comparações são realizadas entre os três métodos e previsões desse sistema não linear são produzidas para períodos maiores. No Capítulo 5 são comentadas as considerações finais sobre este trabalho e suas perspectivas futuras.

## 2 SISTEMAS DINÂMICOS NÃO LINEARES

O estudo da dinâmica trata sobre a mudança dos sistemas que evoluem no tempo, assim, é possível analisar o comportamento de qualquer sistema, podendo este sistema em questão estar estabilizando, se repetindo em ciclos ou até mesmo em situações mais complexas. A descrição da dinâmica de um sistema físico é realizada a partir das suas equações de movimento.

Essas equações podem ter formas lineares ou não lineares. Grande parte dos fenômenos naturais estudados na física não são lineares. Sempre que partes de um sistema interferem, cooperam ou competem, ocorrem interações não lineares e essa não linearidade é essencial para a compreensão de diversos problemas físicos, como em sistemas oscilatórios (POL; MARK, 1927), (MOON, 1980), em operação de um laser (ATMANSPACHER; SCHEINGRABER, 1986) e em órbitas planetárias (SUSSMAN; WISDOM, 1988).

O interessante das equações de movimento de um sistema dinâmico não linear é que elas podem exibir o fenômeno de *caos*. O comportamento de um sistema caótico pode ser muito complicado e a necessidade de descrevê-lo gerou novos caminhos para encontrar as soluções de suas equações governantes, exemplos disso são pela análise qualitativa realizada pelo plano de fase, discutida na Seção 2.2, e pela forma quantitativa produzida por métodos numéricos como Runge-Kutta, discutida na Seção 2.3.

Por isso, para uma melhor compreensão do efeito caótico que a não linearidade trás, é necessário um entendimento profundo sobre as equações diferenciais não lineares que muitas das vezes não são facilmente resolvidas analiticamente.

### 2.1 Equações Lineares e não lineares

Considera-se uma equação linear se ela envolve a variável, ou variáveis dependentes e suas derivadas, apenas linearmente. Um sistema linear bidimensional é um sistema representado pelas Equações 2.1 (TAYLOR, 2013)

$$\begin{aligned} \dot{x} &= ax + by, \\ \dot{y} &= cx + dy, \end{aligned} \tag{2.1}$$

onde a, b, c e d são constantes.

Sendo os vetores denotados por negrito, este sistema pode ser escrito na forma matricial apresentada pela Equação 2.2

$$\dot{\mathbf{x}} = A\mathbf{x}, \quad (2.2)$$

onde

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{e} \quad \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}. \quad (2.3)$$

Tal sistema é linear no sentido de que se  $\mathbf{x}_1$  e  $\mathbf{x}_2$  são soluções dessas equações, então qualquer combinação linear  $c_1\mathbf{x}_1 + c_2\mathbf{x}_2$  é solução. Esse resultado é chamado de princípio da superposição.

Um exemplo de um sistema linear são as vibrações de uma massa pendurada em uma mola governadas pela Equação diferencial linear 2.4

$$m\ddot{x} + kx = 0, \quad (2.4)$$

onde  $m$  é a massa,  $k$  é a constante da mola e  $x$  é o deslocamento da massa em relação a posição de equilíbrio.

A Equação 2.4 é facilmente resolvida e tem como solução  $x(t) = A \cos(\sqrt{\frac{k}{m}}t) + B \sin(\sqrt{\frac{k}{m}}t)$  em que  $A$  e  $B$  são constantes.

Ao contrário de um sistema com equações lineares, os sistemas não lineares consistem de equações formadas por combinação de funções algébricas, como polinômios, e funções transcendentais, como a função exponencial, a função logaritmo e as funções trigonométricas. Por causa da não linearidade, esses sistemas de equações não podem ser reduzidos na forma matricial da Equação 2.2.

Um exemplo de sistema físico não linear é a Equação de movimento 2.5 de um pêndulo simples, pois apresenta termos não lineares em  $\theta$  como o  $\sin\theta$ .

$$mL^2\ddot{\theta} + mgL \sin\theta = 0, \quad (2.5)$$

onde  $m$  é a massa,  $L$  é o comprimento do fio,  $g$  é a gravidade e  $\theta$  é o ângulo em que o pêndulo se encontra em relação a origem.

Muitas vezes, para resolver esse problema, usa-se a aproximação  $\sin\theta \approx \theta$  para transformar a Equação 2.5 em uma equação linear. Porém, essa aproximação somente se aplica para ângulos pequenos. Caso contrário, esse problema não é resolvido trivialmente. Isso não é



algo particular do problema do pêndulo. A maioria das equações não lineares não são possíveis de serem resolvidas analiticamente e para descrever melhor o comportamento desses sistema, geralmente é realizado uma análise mais qualitativa das equações governantes, retratando-as a partir do chamado *Plano de Fase*.

## 2.2 Plano de Fase

O plano de fase exhibe visualmente as soluções das equações diferenciais de um sistema. Nele, para o caso bidimensional, há um plano de coordenadas com eixos  $(x, y)$ . A forma geral de um campo vetorial bidimensional no plano de fase é mostrado na Equação 2.6

$$\begin{aligned} \dot{x} &= f(x, y), \\ \dot{y} &= g(x, y), \end{aligned} \tag{2.6}$$

onde  $f$  e  $g$  são funções.

Esse sistema pode ser escrito na forma compacta apresentada na Equação 2.7

$$\dot{\mathbf{p}} = \mathbf{F}(\mathbf{p}), \tag{2.7}$$

onde  $\mathbf{p} = (x, y)$  e  $\mathbf{F}(\mathbf{p}) = (f, g)$ .

De acordo com a Equação 2.7, pode-se considerar que  $\mathbf{p}$  representa um ponto no plano de fase e  $\dot{\mathbf{p}}$  é o vetor de velocidade nesse ponto. Um ponto de fase traça uma solução  $p(t)$  decorrendo ao longo do campo vetorial. Essa solução corresponde a uma trajetória através do plano de fase. Além disso, o plano de fase é preenchido por diversas trajetórias, pois cada ponto pode ser considerado uma condição inicial. (STROGATZ, 2018)

Quando consideramos todo o campo vetorial, a representação de todas as trajetórias de um sistema dinâmico no plano é chamado de *retrato de fase*. A construção das trajetórias em relação as condições iniciais serão discutidas na subseção 2.2.1.

Por meio do retrato de fase é possível determinar o comportamento qualitativo das soluções, o que facilita compreender melhor a evolução de sistemas dinâmicos não lineares.

Portanto, a partir das propriedades de  $\mathbf{F}(\mathbf{p})$  é possível encontrar diretamente uma enorme variedade de retratos de fase, mostrando assim como a dinâmica do sistema se comporta, se é um sistema estável ou não. Exemplo disso podem ser visto em sistemas oscilatórios como modelos predadores-presas (LOTKA, 1925).

### 2.2.1 Problema de Cauchy

O problema de Cauchy, também conhecido como problema de valor inicial, comprova a existência e unicidade das soluções e essas características são importantes para montar o retrato de fase de um sistema dinâmico.

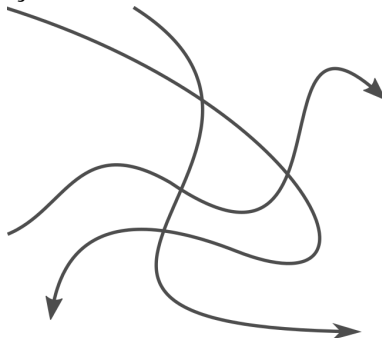
**Teorema 2.2.1 (Existência e unicidade local)** *Suponha que  $f$  seja contínuo em  $\Omega \subseteq \mathbb{R}^2$  e tem derivada parcial contínua em  $f_x$  com respeito a  $x$ . Seja  $(t_0, x_0)$  um dado ponto no interior de  $\Omega$ . Então existe um intervalo fechado  $I$  contendo  $t_0$  em seu interior tal que o problema de Cauchy*

$$\begin{cases} \dot{x} &= f(t, x) \\ x(t_0) &= x_0 \end{cases}$$

*tem uma solução única, definida em  $I$  (AHMAD, 2013).*

Ou seja, de acordo com esse teorema a existência e unicidade de soluções são garantidas caso  $f$  seja continuamente diferenciável. Desse teorema, também se deduz um corolário importante: *trajetórias de soluções diferentes nunca se cruzam*, pois se alguma trajetória se intersectar com outra vai existir mais de uma solução para aquele ponto de intersecção. Logo, violando a unicidade do Teorema 2.2.1 como apresentado na Figura 1.

Figura 1 – Exemplo da violação do Teorema 2.2.1.



Fonte: Elaborada pelo autor (2021).

Esse teorema é usualmente trabalhado na física, uma vez que é constantemente buscado a solução da equação do sistema estudado a partir das suas condições iniciais dadas e por meio disso é sempre encontrado apenas uma solução para os pontos em que se deseja saber. Os osciladores harmônicos e o circuito RLC são exemplos de problemas físicos simples que mostram essa aplicação.

### 2.2.2 Interpretação dos pontos críticos em um retrato de fase

As consequências do Teorema 2.2.1 são perceptíveis em qualquer sistema físico e para montar o retrato de fase é importante definir os pontos críticos da equação governante do sistema.

Os pontos críticos são definidos encontrando primeiramente os pontos de equilíbrio do sistema, ou seja, os pontos em que as equações de movimento são iguais a zero. Após encontrar esses pontos, é necessário definir se o ponto de equilíbrio é um nó, espiral, centro (estável ou instável) ou um ponto de sela. Para descobrir a natureza desse ponto, usa-se a matriz jacobiana dada pela Equação 2.8

$$J(x,y) = \begin{pmatrix} \frac{\partial f(x,y)}{\partial x} & \frac{\partial f(x,y)}{\partial y} \\ \frac{\partial g(x,y)}{\partial x} & \frac{\partial g(x,y)}{\partial y} \end{pmatrix}, \quad (2.8)$$

onde  $f(x,y)$  e  $g(x,y)$  são funções.

Se o determinante da matriz Jacobiana for negativo e os autovalores forem reais com sinais opostos, o ponto fixo é um ponto de sela. Se o determinante da matriz Jacobiana for positivo e seus autovalores forem reais com o mesmo sinal, o ponto fixo será um nó. Caso os autovalores sejam complexos conjugados, o ponto fixo será espirais ou centro. A estabilidade dos nós e das espirais é determinada pela soma dos autovalores. Se a soma for negativa, então o ponto fixo é estável. Caso contrário, o ponto fixo é instável (STROGATZ, 2018).

Um exemplo disso é uma partícula percorrendo um plano bidimensional e suas equações de movimento são definidas pela Equação 2.9

$$\begin{aligned} \dot{x} &= -x + x^3, \\ \dot{y} &= -2y; \end{aligned} \quad (2.9)$$

O sistema é considerado não linear pois no sistema de equações 2.9,  $\dot{x}$  é não linear já que apresenta dependência em  $x^3$ , por esse motivo a análise qualitativa é útil para desenhar seu retrato de fase.

Para isso, primeiro se deve verificar quais são seus pontos de equilíbrio que devem ser quando  $\dot{x} = 0$  e  $\dot{y} = 0$  simultaneamente. Então para esses valores, tem-se os pontos em  $x$  e  $y$  iguais a  $(0,0)$ ,  $(-1,0)$  e  $(1,0)$ .

Para descobrir a natureza desse ponto, usa-se a matriz jacobiana dada pela Equação 2.8 e se obtém para o ponto crítico  $(0,0)$  a matriz Jacobiana igual a

$$J(0,0) = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}, \quad (2.10)$$

e para os pontos  $(\pm 1,0)$  a matriz Jacobiana é igual a

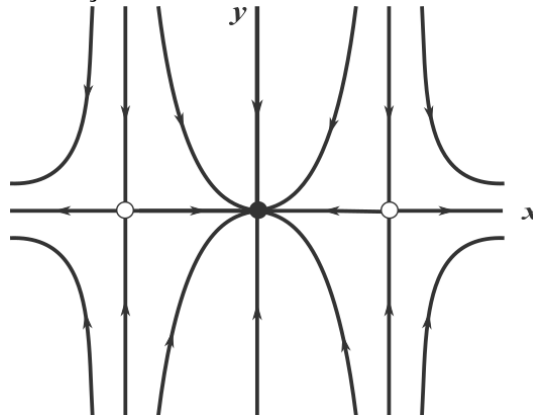
$$J = (\pm 1,0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}. \quad (2.11)$$

No ponto  $(0,0)$ , nota-se que seus autovalores  $\lambda_1 = -1$ ,  $\lambda_2 = -2$  são reais, negativos e a determinante da matriz é positiva, logo o ponto fixo é um nó. Para verificar a estabilidade desse nó, deve-se somar os autovalores, nesse caso  $\lambda_1 + \lambda_2 = -3$ . Portanto, o ponto  $(0,0)$  é um nó estável. Já para os pontos  $(-1,0)$  e  $(1,0)$ , como os autovalores  $\lambda_1 = 2$  e  $\lambda_2 = -2$  são reais com sinais opostos e a determinante da matriz é negativa, então os pontos serão pontos de sela.

Com isso, como as equações  $x$  e  $y$  são desacopladas, esse sistema representa dois sistemas independentes de primeira ordem perpendiculares entre si. Portanto, o desenho do retrato de fase é ilustrado na Figura 2 em que foi montado da seguinte forma: na direção  $y$ , todas as trajetórias decaem exponencialmente para  $y = 0$ . Na direção  $x$ , as trajetórias são atraídas para  $x = 0$  e repelidas para  $x = \pm 1$ , onde o ponto preto denota o nó estável e o branco denota o nó instável. As linhas verticais  $x = 0$  e  $x = \pm 1$  são invariantes, porque  $\dot{x} = 0$  neles; logo, qualquer trajetória que começa nessas linhas permanece nelas para sempre. Da mesma forma,  $\dot{y} = 0$  é uma linha horizontal invariável. Além disso, já que as equações são invariantes sob as transformações  $x \rightarrow -x$  e  $y \rightarrow -y$ , o retrato de fase deve ser simétrico em ambos os eixos  $x$  e  $y$ . Dessa forma colocando tudo isso junto é formado o retrato de fase apresentado na Figura 2.

Essa é uma abordagem mais qualitativa do problema da dinâmica não linear, uma vez que não foi necessário encontrar exatamente as soluções para se ter uma interpretação do sistema.

Figura 2 – Retrato de fase da Equação de evolução 2.9.



Fonte: Elaborada pelo autor (2021).

### 2.3 Aspectos quantitativos de um sistema dinâmico não linear

Como já visto, a maioria dessas equações não lineares não podem ser resolvidas analiticamente e por isso usar aproximações numéricas é útil para encontrar aspectos quantitativos. Um dos métodos numéricos bastante utilizado para encontrar soluções aproximadas do problema de valor inicial é o método Runge-Kutta.

O método Runge Kutta calcula o próximo valor  $x_{k+1}$  por meio de vários estágios com objetivo de obter melhores aproximações. Para este fim,  $x_{k+1}$  será  $x_k$  mais uma média ponderada de um número  $s$  de incrementos, o número  $s$  é fixo e denominado número de estágios. Cada incremento é apenas um produto do tamanho do passo  $h$  e uma inclinação estimada da curva de solução especificada pelo lado direito  $f(t, x)$  da equação do problema inicial da subseção 2.2.1(AHMAD, 2013).

Esse método em geral é resolvido para quatro estágios, popularmente chamado de método de Runge-Kutta de quarta ordem. Imagine um problema de valor inicial dependente da posição  $x$  e do instante  $t$ . Usando o método de Runge-Kutta de quarta ordem, a inclinação será estimada pela média ponderada de vários termos da forma  $\dot{x} = f(t_k, x_k)$  onde  $t_k$  ( $k = 1, 2, \dots$ ) com o passo sendo  $h = \Delta t$ , ou seja, os valores de  $t_k$  são escolhidos no intervalo  $[t, t + \Delta t]$ , e  $x_k$  serão obtidos usando alguma aproximação para  $x_k(t)$ . O método de Runge-Kutta de quarta ordem é definido pela Equação 2.12 (GIORDANO, 2005).

$$x(t + \Delta t) = x(t) + \frac{1}{6}[f(t_1, x_1) + 2f(t_2, x_2) + 2f(t_3, x_3) + f(t_4, x_4)]\Delta t, \quad (2.12)$$

onde

$$\begin{aligned}
 t_1 &= t, & x_1 &= x(t). \\
 t_2 &= t + \frac{1}{2}\Delta t, & x_2 &= x(t) + \frac{1}{2}f(t_1, x_1)\Delta t. \\
 t_3 &= t + \frac{1}{2}\Delta t, & x_3 &= x(t) + \frac{1}{2}f(t_2, x_2)\Delta t. \\
 t_4 &= t + \Delta t, & x_4 &= x(t) + f(t_3, x_3)\Delta t.
 \end{aligned}$$

## 2.4 Teoria do caos

Sempre que o caos é encontrado em um sistema dinâmico, ele é associado à não linearidade das equações que regem esse sistema. Nem todos sistemas não lineares são caóticos, para que eles possam ter tal aspecto o sistema deve ter mais de um grau de liberdade ou não ser autônomo, isto é, deve ser dependente do tempo. Assim, a teoria do caos trata de sistemas determinísticos que apresentam sensibilidade às condições iniciais, ou seja, o estado atual do sistema depende do estado anterior de forma rigidamente determinada, o que se torna imprevisível a observação ou medição ao longo prazo.

Essa sensibilidade dá a característica de instabilidade, por isso, mesmo se tratando de sistemas determinísticos, eles apresentam uma grande sensibilidade aos erros e às perturbações aparentando ser um sistema aleatório devido aos seus resultados imprevisíveis. Por conta disso, há muitos anos o caos era confundido com aleatoriedade, apesar de não ser. Um exemplo disso é retratado pelo matemático Norbert Wiener quando se referia ao caos como um hospedeiro de moléculas localizadas aleatoriamente que formavam um gás, ou a coleção desordenada de gotículas de água. (WIENER, 1938).

Portanto, pode-se dizer que o termo caos se refere coletivamente a processos determinísticos que aparentam proceder de acordo ao acaso. Exemplos como esse existem em inúmeros processos físicos, como o balanço de um pêndulo em um relógio, a queda de uma rocha na encosta de uma montanha ou o rompimento de ondas na costa de um oceano, em que variações de algum tipo ocorrem conforme o tempo avança. (LORENZ, 1995).

A dificuldade de se conhecer o estado presente com exatidão levou-se à necessidade de modelar esses sistemas caóticos e prever o que aparenta ser acaso, mas que na verdade são

representado por equações. Um exemplo de sistema caótico são as Equações de Lorenz que serão discutidas na subseção seguinte.

### 2.4.1 Equações de Lorenz

O matemático e meteorologista Edward Lorenz tentava descrever o comportamento da atmosfera a partir de um sistema de equações adequado. Buscando prever o tempo, ele percebeu que a atmosfera real não se comportava como um modelo simples. As temperaturas, os ventos e outras quantidades que entram na estimativa de clima não eram medidas com precisão e por conta disso as mudanças pequenas das condições iniciais ocasionava a falta de periodicidade das previsões climáticas.

Em 1963, ele derivou as equações de movimento de um sistema tridimensional vindas de um modelo de rolos de convecção na atmosfera (LORENZ, 1963). As Equações 2.13 são

$$\begin{aligned}\dot{x} &= \sigma(y - x), \\ \dot{y} &= \rho x - y - xz, \\ \dot{z} &= xy - \beta z,\end{aligned}\tag{2.13}$$

onde  $\sigma$ ,  $\rho$  e  $\beta$  são parâmetros e maiores que zero. Sendo  $\sigma$  o número de Prandtl,  $\rho$  o número de Rayleigh e  $\beta$  relacionado com o aspecto dos rolos nos problemas de convecção.

Nota-se que esse sistema é não linear, pois apresenta dependência em termos quadráticos  $xy$  e  $xz$ . Observa-se também que as equações possuem simetria, ou seja  $(x, y) \rightarrow (-x, -y)$ . Logo, todas as soluções são simétricas ou têm um parceiro simétrico. Além disso, o sistema de Lorenz é dissipativo, isto é, os volumes no espaço de fase se contraem sob o fluxo.

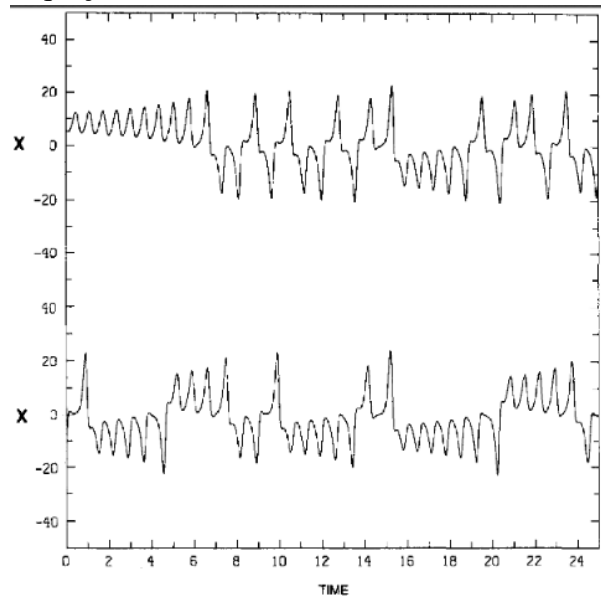
Com isso, Lorenz descobriu que esse sistema determinístico de aparência simples tem uma dinâmica estranha, pois com um número amplo de parâmetros, as soluções oscilavam irregularmente, nunca se repetindo exatamente, mas sempre permanecendo em uma região limitada do espaço de fase. Quando traçadas as trajetórias dessas equações em três dimensões, foi identificado um conjunto de soluções complicadas o que foi chamado de atrator estranho.

#### 2.4.1.1 Atrator estranho

As três equações demonstram uma das maneiras mais simples em que um sistema determinístico pode se comportar quando não se comporta periodicamente. Lorenz usou a

integração numérica para ver o que as trajetórias fariam ao longo prazo. Ele estudou o caso particular para  $\sigma = 10$ ,  $\rho = 28$  e  $\beta = \frac{8}{3}$ . Plotando gráficos para descrever como as soluções das variáveis se comportam ao passar do tempo conseguiu observar, por exemplo, o comportamento da evolução temporal das soluções de  $x(t)$ ,  $y(t)$  e  $z(t)$ . A Figura 3 mostra o comportamento da solução  $x(t)$  em relação ao tempo.

Figura 3 – Série temporal de  $x(t)$  do sistema das Equações de Lorenz.



Fonte: Retirada do livro (LORENZ, 1995), pg. 138

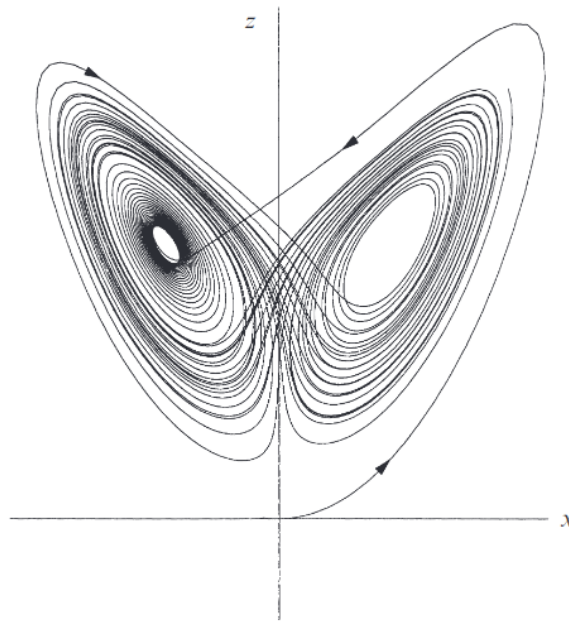
Pela Figura 3 nota-se que o sistema está em equilíbrio instável, pois ele sofre oscilações amplificadoras sobre um estado até que as oscilações se tornem muito fortes. Em seguida, ele passa oscilar sobre outro estado instável, até que essas oscilações se tornem muito fortes também, depois o sistema alterna continuamente entre esses dois modos de comportamento. A característica principal retratada na Figura 3 é como os números sucessivos de oscilações sobre um estado ao outro ocorrem em uma sequência irregular. A partir disso, conclui-se que uma solução não periódica é instável ou sensivelmente dependente, e assim mostra que se a solução geral de um sistema não é periódica, o seu comportamento geral é caótico.

Visualizando a evolução temporal das variáveis, Lorenz descobriu que uma estrutura diferente é criada se a solução for visualizada como uma trajetória no espaço de fase. Por exemplo, quando  $z(t)$  é plotado contra  $x(t)$ , um padrão de borboleta aparece. A Figura 4 mostra o contorno do atrator e uma curva de solução estendida dentro dele.

Na Figura 4 as trajetórias parecem se cruzar, isso acontece porque as trajetórias tridimensionais estão projetadas em um plano bidimensional. Quando desenhadas em três



Figura 4 – Atrator de Lorenz.



Fonte: Retirada do livro (STROGATZ, 2018), pg. 326.

dimensões nenhuma intersecção ocorre. Então, o teorema citado na subseção 2.2.1 é obedecido.

Pela Figura 4 a trajetória começa perto da origem, depois oscila para a direita e mergulha no centro de uma espiral à esquerda. Depois passa por uma espiral muito lenta para fora, a trajetória volta para o lado direito, espirala algumas vezes, e volta para a esquerda, espirala ao redor e assim por diante indefinidamente. O número de ciclos feitos em cada lado varia de forma imprevisível de um ciclo para o próximo.

Como modelo de convecção comum, esse sistema é deficiente porque impõe restrições excessivas ao movimento. Se você colocar uma panela com água no fogão e ligar o queimador, a água não vai subir em todo o lado esquerdo e afundar em todo o lado direito, ou subir apenas na direita e afundar apenas na esquerda, em um grande rolo, como as equações exigiriam que isso acontecesse; numerosos rolos menores serão desenvolvidos. Se estes pudessem ser suprimidos de alguma forma, o movimento restante poderia estar mais de acordo com as equações. (LORENZ, 1995)

Porém, a realidade física da circulação irregularmente alternada foi comprovada alguns anos depois, quando os matemáticos aplicados Willem Malkus e Louis Howard e Ruby Krishnamurti construíram rodas d'água que foram destinadas especificamente a executar o comportamento previsto pelas equações de Lorenz (MATSON, 2007).

### 3 APRENDIZADO DE MÁQUINA

Aprendizado de máquina estuda e constrói algoritmos capazes de aprender com os dados e fazer previsões sobre eles. Esses algoritmos são construídos baseados em um modelo a partir das entradas do sistema, chamadas de *características*, com objetivo de fazer previsões ou decisões conduzidas pelos dados.

O uso do aprendizado de máquina é útil para vários tipos de problemas complexos que possuem grande quantidades de dados e ajustes. Um exemplo simples é o filtro de spam do e-mail o qual é um programa que aprende a sinalizar e-mails indesejados e os classificam como spam, tais como propagandas. Para que o filtro consiga classificar o que é spam e o que não é, o sistema aprende a partir de um conjunto de treinamento com vários e-mails. Assim, o algoritmo treina com este conjunto para poder prever o que é spam nos e-mails futuros. A partir da previsão é possível observar se ela foi precisa e obteve êxito na sua classificação.

Outro exemplo de aplicação seria prever um valor numérico, chamado de *targets*, dado por um conjunto de características. Essa aplicação pode ser usada para prever preços de imóveis, havendo o conjunto de características como o número de quartos, localização, área e entre outras propriedades do imóvel.

Para esses problemas mencionados se usa métodos supervisionados, ou seja, os dados de treinamento fornecidos ao algoritmo, as características, incluem as soluções desejadas, os targets. Entre os métodos supervisionados, a regressão linear foi a utilizada neste trabalho e será discutida na próxima subseção.

#### 3.1 Regressão Linear

A regressão linear é um dos modelos mais simples para prever os valores de saída. Esse modelo usa uma função linear, chamada de hipótese, cuja a entrada são as características do conjunto de dados do sistema estudado. A hipótese  $h(\Theta, x^{(i)})$  é dada pela Equação 3.1

$$h(\Theta, x^{(i)}) = \Theta^T \cdot x^{(i)} = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_n x_n, \quad (3.1)$$

onde  $n$  é o número de características,  $x_i$  é o  $i$ -ésimo valor das características,  $\theta_j$  é o  $j$ -ésimo parâmetro do modelo.

Além disso, a regressão também pode ser não linear e escrita na forma polinomial. Por exemplo, pode-se substituir os termos  $x^{(i)}$  por polinômios tais que  $x_n = x^n$ , assim a função

hipótese é expressa pela equação 3.2

$$h(\Theta, x^{(i)}) = \Theta^T \cdot x^{(i)} = \theta_0 + \theta_1 x + \theta_2 x^2 + \dots + \theta_n x^n. \quad (3.2)$$

### 3.1.1 Função Custo

Para treinar o modelo de regressão, os parâmetros devem ser definidos de modo que o modelo se ajuste melhor ao conjunto de treinamento. Para isso, é necessário encontrar o valor de  $\Theta$  que minimiza o erro quadrático médio da hipótese  $h(\Theta^T, x^{(i)})$  em um conjunto de treinamento  $x^{(i)}$ . Essa minimização é chamada de função custo  $J(\Theta)$  e é calculada pela Equação 3.3

$$J(\Theta) = \frac{1}{2m} \sum_{i=1}^m (h(\Theta^T, x^{(i)}) - y^{(i)})^2, \quad (3.3)$$

onde  $m$  é o número de amostrar e  $y^{(i)}$  é o vetor target.

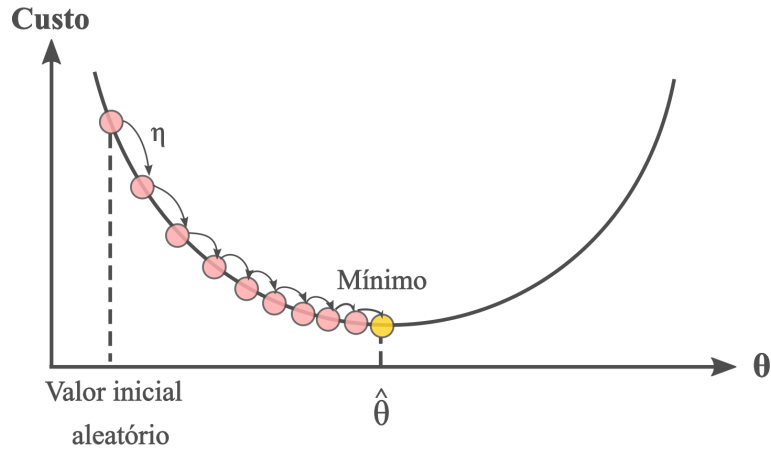
### 3.1.2 Gradiente Descendente

O Gradiente Descendente é um algoritmo de otimização capaz de encontrar soluções ótimas. Sua ideia geral é ajustar os parâmetros de forma iterativa para minimizar a função custo. Dessa forma, esse algoritmo, quando aplicado à 3.3, mede o gradiente local do erro da função em relação ao vetor de parâmetros  $\Theta$  e caminha na direção onde o gradiente decresce. Quando o gradiente é zero, significa que atingiu o mínimo (GÉRON, 2017).

O vetor  $\Theta$  começa sendo preenchido com valores aleatórios e vai sendo melhorado gradualmente até que consiga diminuir a função custo, convergindo-a ao mínimo. A Figura 5 ilustra o funcionamento do Gradiente Descendente.

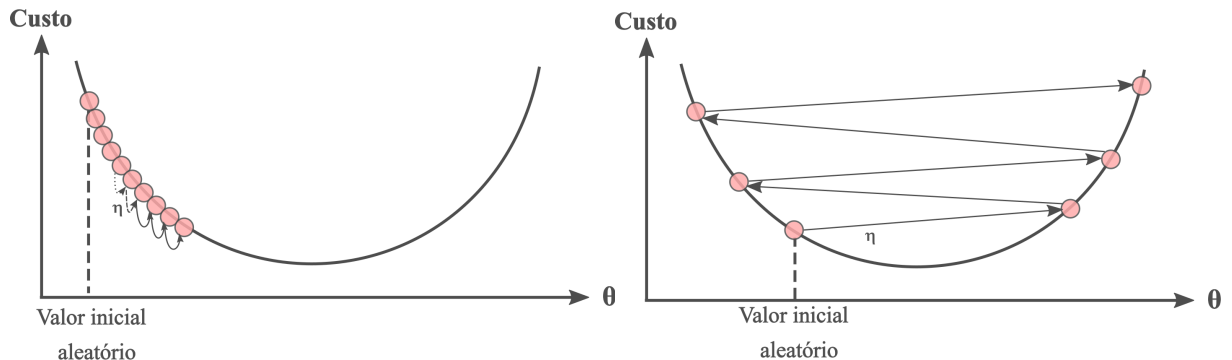
Para que essa convergência ocorra, é necessário determinar um hiperparâmetro  $\eta$  chamado de taxa de aprendizagem. Esse hiperparâmetro irá determinar o tamanho das etapas e deve ser escolhido atenciosamente, pois se a taxa for muito pequena o algoritmo terá que passar por muitas iterações para convergir o que pode levar muito tempo. No entanto, se a taxa de aprendizado for muito alta, o algoritmo pode pular o vale e acabar no outro lado da curva. Sendo possível esse outro lado ser mais alto do que estava antes e consequentemente divergir o resultado para valores cada vez maiores. A Figura 6 esquematiza o algoritmo para valores pequenos e grandes da taxa de aprendizagem.

Figura 5 – Representação do algoritmo Gradiente Descendente inicializado por um valor aleatório. O algoritmo percorre a função custo a uma taxa de aprendizado ideal  $\eta$  até encontrar o seu valor mínimo.



Fonte: Adaptada do livro (GÉRON, 2017), pg 156.

Figura 6 – À esquerda tem a representação do funcionamento do algoritmo Gradiente Descendente para uma taxa de aprendizado muito pequena, fazendo com que o algoritmo demore a chegar no mínimo e à direita para uma taxa de aprendizado muito alta fazendo com que o algoritmo divirja do valor mínimo.



Fonte: Adaptada do livro (GÉRON, 2017), pg 157.

Mesmo que a função custo apresente topologias irregulares como buracos e cristas, isso não será problema para esse modelo de regressão linear, pois a função custo é convexa, ou seja, não há presenças de mínimos locais, apenas um mínimo global, o que facilita a aplicação do algoritmo. Além disso, a função de custo é contínua com uma inclinação que não muda abruptamente. Dessa forma, esses dois fatores permite a garantia de que o algoritmo Gradiente Descendente se aproxime gradativamente do mínimo global quando tiver uma taxa de aprendizagem adequada (GÉRON, 2017).

Portanto, buscando calcular o quanto a função custo varia, na implementação do Gradiente Descendente deve-se computar o gradiente da função de custo em relação a cada parâmetro do modelo  $\theta_j$ . A equação 3.4 calcula o vetor gradiente da função custo

$$\nabla J(\Theta) = \begin{pmatrix} \frac{\partial}{\partial \theta_0} J(\Theta) \\ \frac{\partial}{\partial \theta_1} J(\Theta) \\ \vdots \\ \frac{\partial}{\partial \theta_n} J(\Theta) \end{pmatrix} = \frac{1}{m} \sum_{i=1}^m (x^{(i)})^T (x^{(i)} \cdot \Theta - y^{(i)}). \quad (3.4)$$

Tendo o vetor gradiente, o qual aponta para cima, basta ir na sua direção oposta para descer, ou seja é necessário subtrair  $\theta_j$  do vetor gradiente. Para uma determinada taxa de aprendizagem  $\eta$  a expressão é obtida pela equação 3.5

$$\theta_j = \theta_j - \eta \frac{1}{m} \sum_{i=1}^m [h(\Theta, x^{(i)}) - y^{(i)}] x_j^{(i)}. \quad (3.5)$$

Além disso, existem outras formas de implementar o Gradiente Descendente, ele pode descer em lotes (*Batch Gradient Descendent*) onde o gradiente da função de custo é calculada para cada parâmetro  $\theta_j$ , e pode ser calculado a cada passo (*Stochastic Gradient Descent*)(GÉRON, 2017).

### 3.1.3 Overfitting

Após usar a regressão linear em um conjunto de dados é possível verificar se o modelo está bem ajustado, ou seja, se o modelo prever de acordo com o que foi treinado. Porém, pode acontecer do modelo não estar bem ajustado, isto é, apresentar um overfitting. Isso acontece quando há muitas características, assim a hipótese aprendida pode se ajustar excepcionalmente ao conjunto de treino ( $J(\Theta) \approx 0$ ), mas falhando em generalizar para novas amostras (BHATTACHARYYA, 2018).

Dessa forma, para melhorar o ajuste da regressão, a função custo deverá ser regularizada com objetivo de reduzir a influência de determinados parâmetros  $\theta_j$  sobre ela. Para isso, a regressão Ridge e Lasso são técnicas bastante utilizadas para evitar o overfitting e apresentam um ótimo desempenho para um grande número de características, onde cada característica contribuí na previsão de  $y$ .

### 3.1.4 Regressão Ridge

A regressão Ridge é uma versão regularizada da regressão linear ou polinomial. Essa regularização é adicionada a função custo de forma que força o algoritmo a ajustar os dados

mantendo os pesos do modelo tão pequenos quanto possível, assim, contribuindo para a redução da influência dos parâmetros. A equação 3.6 mostra a função custo com a regularização ridge

$$J(\Theta) = \frac{1}{m} \sum_{i=1}^m (h(\Theta^T, x^{(i)}) - y^{(i)})^2 + \frac{\lambda}{2m} \sum_{j=1}^n \theta_j^2, \quad (3.6)$$

onde o  $\lambda$  é o hiperparâmetro que controla a regularização do modelo.

Quando  $\lambda = 0$ , a regressão Ridge se torna a regressão linear da Equação 3.3. Além disso, o termo de polarização  $\theta_0$  não é regularizado.

A regularização só deve ser adicionada na função custo durante o treinamento. Após ter treinado o modelo, o desempenho deve ser avaliado usando a medida de desempenho não regularizada.

O algoritmo de Gradiente Descendente para regressão Ridge é expresso pela Equação 3.7

$$\frac{\partial J_{\Theta}}{\partial \theta_j} = \frac{1}{m} \sum_{i=1}^m (x^{(i)})^T (x^{(i)} \cdot \Theta - y^{(i)}) + \frac{\lambda}{m} \sum_{n=1}^n \theta_j. \quad (3.7)$$

### 3.1.5 Regressão Lasso

A regressão Lasso (TIBSHIRANI, 1996) é outra versão regularizada da regressão que também adiciona um termo de regularização na função custo, mas dessa vez esse termo é o módulo do vetor de parâmetros  $\theta_j$ . A função custo na regressão Lasso é dada pela Equação

$$J(\Theta) = \frac{1}{m} \sum_{i=1}^m (h(\Theta^T, x^{(i)}) - y^{(i)})^2 + \lambda \sum_{j=1}^n |\theta_j|. \quad (3.8)$$

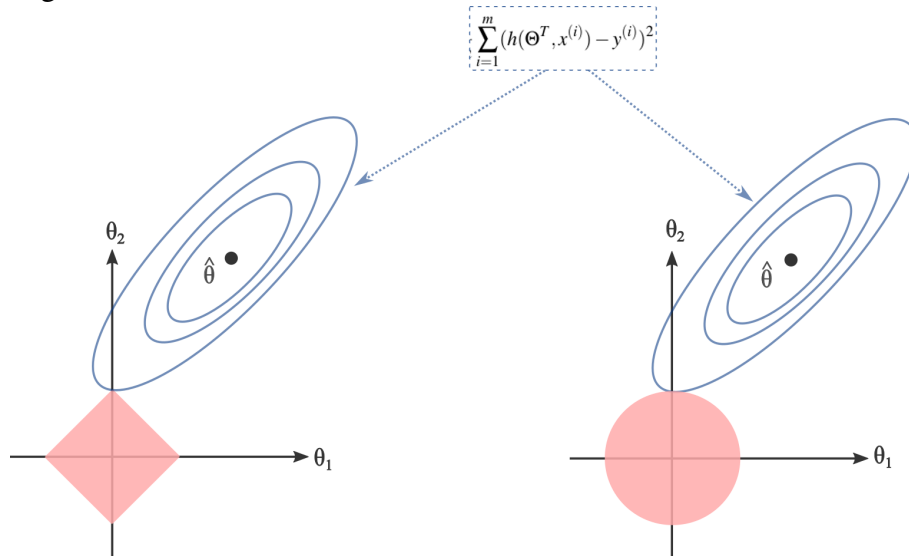
A função custo nesse tipo de regressão não é diferenciável em  $\theta_j = 0$ , mas o algoritmo do Gradiente Descendente ainda funciona quando usado um subgradiente para qualquer  $\theta_j = 0$ . Portanto, esse subgradiente é dado pela equação

$$\frac{\partial J_{\Theta}}{\partial \theta_j} = \frac{1}{m} \sum_{i=1}^m (x^{(i)})^T (x^{(i)} \cdot \Theta - y^{(i)}) + \lambda \text{sign}(\theta_j) \quad \text{onde} \quad \text{sign}(\theta_j) = \begin{cases} -1, & \text{se } \theta_j < 0 \\ 0, & \text{se } \theta_j = 0 \\ +1, & \text{se } \theta_j > 0 \end{cases} \quad (3.9)$$

A regressão Lasso tende a eliminar completamente os pesos das características menos importantes, ou seja, eles passam a ser definidos como zero diferentemente da regressão Ridge que apenas diminui os pesos.

Portanto, o algoritmo da regressão Lasso executa uma seleção de características e produz um modelo esparso. A Figura 7 ilustra o funcionamento das regressões Lasso e Ridge mostrando a diferença dos pesos nessas duas regressões, onde o conjunto de dados hipotético é dado por dois parâmetros  $\theta_1$  e  $\theta_2$  e é usado a restrição  $|\theta_j|$  e  $\theta_j^2$  nos coeficientes para regressão Lasso e Ridge respectivamente.

Figura 7 – À direita representa o funcionamento da regressão Lasso. À esquerda representa o funcionamento da regressão Ridge. O diamante e o disco rosa representam a região de restrição das regressões Lasso e Ridge respectivamente. Os contornos elípticos são a função custo da regressão linear.



Fonte: Adaptada do artigo (TIBSHIRANI, 1996)

Na regressão Lasso, sempre que o contorno elíptico atinge algum ponto na região de restrição um dos parâmetros desaparece, como no diamante há cantos, se a solução ocorre em um canto, o parâmetro  $\theta_j$  se iguala a zero. Assim, quando estiver em um espaço de dimensão maior, o problema será reduzido para os parâmetros mais importantes. O que não acontece na regressão Ridge. Porém, quando se tem condições relaxadas nos coeficientes, essas regiões restritas podem ficar maiores e atingirão o centro da elipse, isso faz com que a regressão Lasso e Ridge se assemelhe com a regressão linear (FRIEDMAN *et al.*, 2001).

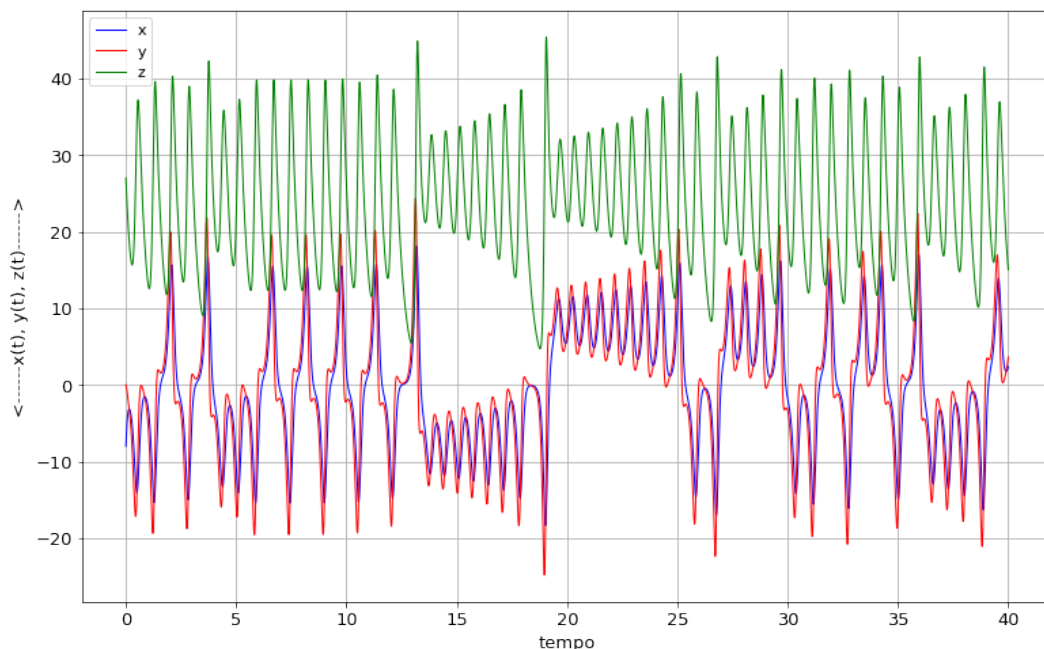
## 4 RESULTADOS

Neste trabalho foi construído o método de Runge-Kutta de quarta ordem, a regressão Ridge e Lasso pela linguagem de programação *Python*. Todos esses métodos foram realizados manualmente sem o uso de bibliotecas prontas.

Na primeira etapa deste trabalho foi construído o método de Runge-Kutta de quarta ordem para encontrar as soluções aproximadas de  $x(t)$ ,  $y(t)$  e  $z(t)$  das equações de Lorenz, Equação 2.13. Os valores usados para os parâmetros foram  $\sigma = 10$ ,  $\rho = 28$  e  $\beta = \frac{8}{3}$ . As variáveis  $x$ ,  $y$  e  $z$  foram inicializadas no instante  $t = 0$  com valores  $x(0) = -8.0$ ,  $y(0) = 0$  e  $z(0) = 27$  e as medidas de tempo foram escolhidas no intervalo de  $[0, 40]$  com passo de  $\Delta t = 0,001$ .

A partir dessa inicialização, o método calculou os valores numéricos aproximados dos estados para cada instante, resultando em 4000 soluções, e com eles foi possível visualizar o comportamento da evolução temporal de cada um dos estados do sistema, apresentada na Figura 8, e o atrator tridimensional do sistema, mostrado na Figura 9.

Figura 8 – Evolução temporal das soluções  $x(t)$ ,  $y(t)$  e  $z(t)$  do sistema das equações de Lorenz computadas pelo método Runge-Kutta de quarta ordem.

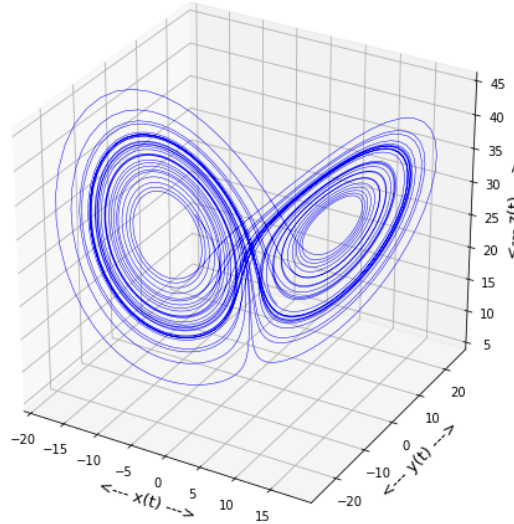


Fonte: Elaborada pelo autor (2022).

Nota-se a semelhança da Figura 8 com a Figura 3. O sistema está em equilíbrio instável como citado na Subseção 2.4.1, pois os números sucessivos de oscilações sobre um estado ao outro ocorrem em uma sequência irregular. A semelhança da Figura 9 com a Figura 4



Figura 9 – Atrator de Lorenz a partir das soluções calculadas pelo método Runge-Kutta de quarta ordem.



Fonte: Elaborada pelo autor (2022).

também é notória visto que o atrator apresenta o padrão borboleta como curva de solução. Pelo gráfico tridimensional, a visualização que a trajetória nunca se cruza, obedecendo o Teorema 2.2.1, é mais perceptível. Portanto, as soluções encontradas por esse método quantitativo corresponde com a teoria.

Além disso, com os valores de  $x(t)$ ,  $y(t)$  e  $z(t)$  obtidos, as velocidades  $\dot{x}(t)$ ,  $\dot{y}(t)$  e  $\dot{z}(t)$  do sistema de Lorenz foram calculadas, as quais resultaram também 4000 soluções, e a partir do conjunto completo de seus resultados foi criado uma matriz  $\mathbf{V}$ , de tamanho  $4000 \times 3$ , representada pela Equação 4.1, e com os resultados dos estados foi construído uma matriz biblioteca  $\Theta(x, y, z)$ , de tamanho  $4000 \times 56$ , com os termos da Série de Taylor apresentado na Equação 4.2

$$\mathbf{V} = \begin{pmatrix} \dot{\mathbf{x}}^T(t_1) \\ \dot{\mathbf{y}}^T(t_1) \\ \dot{\mathbf{z}}^T(t_1) \\ \vdots \\ \dot{\mathbf{x}}^T(t_n) \\ \dot{\mathbf{y}}^T(t_n) \\ \dot{\mathbf{z}}^T(t_n) \end{pmatrix} = \begin{pmatrix} \dot{x}(t_1) & \dot{y}(t_1) & \dot{z}(t_1) \\ \dot{x}(t_2) & \dot{y}(t_2) & \dot{z}(t_2) \\ \vdots & \vdots & \vdots \\ \dot{x}(t_n) & \dot{y}(t_n) & \dot{z}(t_n) \end{pmatrix}, \quad (4.1)$$

$$\Theta(\mathbf{x}, \mathbf{y}, \mathbf{z}) = (1 \quad \mathbf{x} \quad \mathbf{y} \quad \mathbf{z} \quad \mathbf{x}^2 \quad \mathbf{y}^2 \quad \mathbf{z}^2 \quad \mathbf{xy} \quad \mathbf{xz} \quad \mathbf{yz} \quad \dots \quad \mathbf{z}^5). \quad (4.2)$$

A biblioteca  $\Theta$  representa as possíveis funções não lineares candidatas a pertencerem as equações de movimento do sistema. Essa biblioteca pode consistir em constantes, polinômios

e em termos trigonométricos. No entanto, somente os termos das variáveis oriundos da série de Taylor foram utilizados por justamente já conhecer que as equações governantes do sistema tem dependência em termos polinomiais e por isso apenas alguns desses termos de  $\Theta$  estão presentes nas equações. Assim, surgiu um problema de regressão para encontrar os coeficientes  $\Omega = (\omega_1, \omega_2, \dots, \omega_n)$  que determinam quais funções não lineares estão ativas neste sistema. Esse problema é expresso pela Equação 4.3

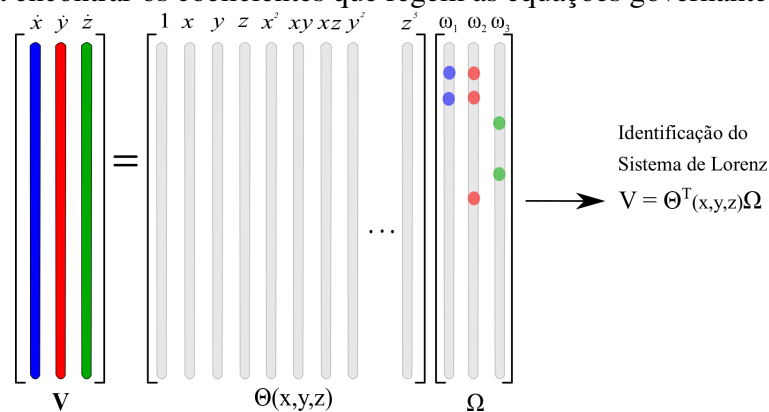
$$\mathbf{V} = \Theta(\mathbf{x}, \mathbf{y}, \mathbf{z})\Omega. \quad (4.3)$$

Portanto, a partir das matrizes  $\mathbf{V}$  e  $\Theta$  foi realizado o processo inverso do método Runge-Kutta de quarta ordem, isto é, o comportamento do sistema foi encontrado usando somente os resultados dos valores numéricos dos estados e das velocidades coletados.

Levando em consideração que a equação governante do sistema não fosse conhecida, para realizar esse processo inverso, os métodos de regressão Ridge e Lasso calcularam os coeficientes  $\Omega$  e com eles foi possível verificar o comportamento e realizar previsões desse sistema não linear.

Assim, a matriz  $\mathbf{V}$  representa o conjunto de *targets*, a biblioteca  $\Theta$  o conjunto das característica e após a implementação delas nas regressões foi encontrado a matriz de coeficientes  $\Omega$  cuja o tamanho era de  $4000 \times 3$ . A Figura 10 ilustra a esquematização desse processo.

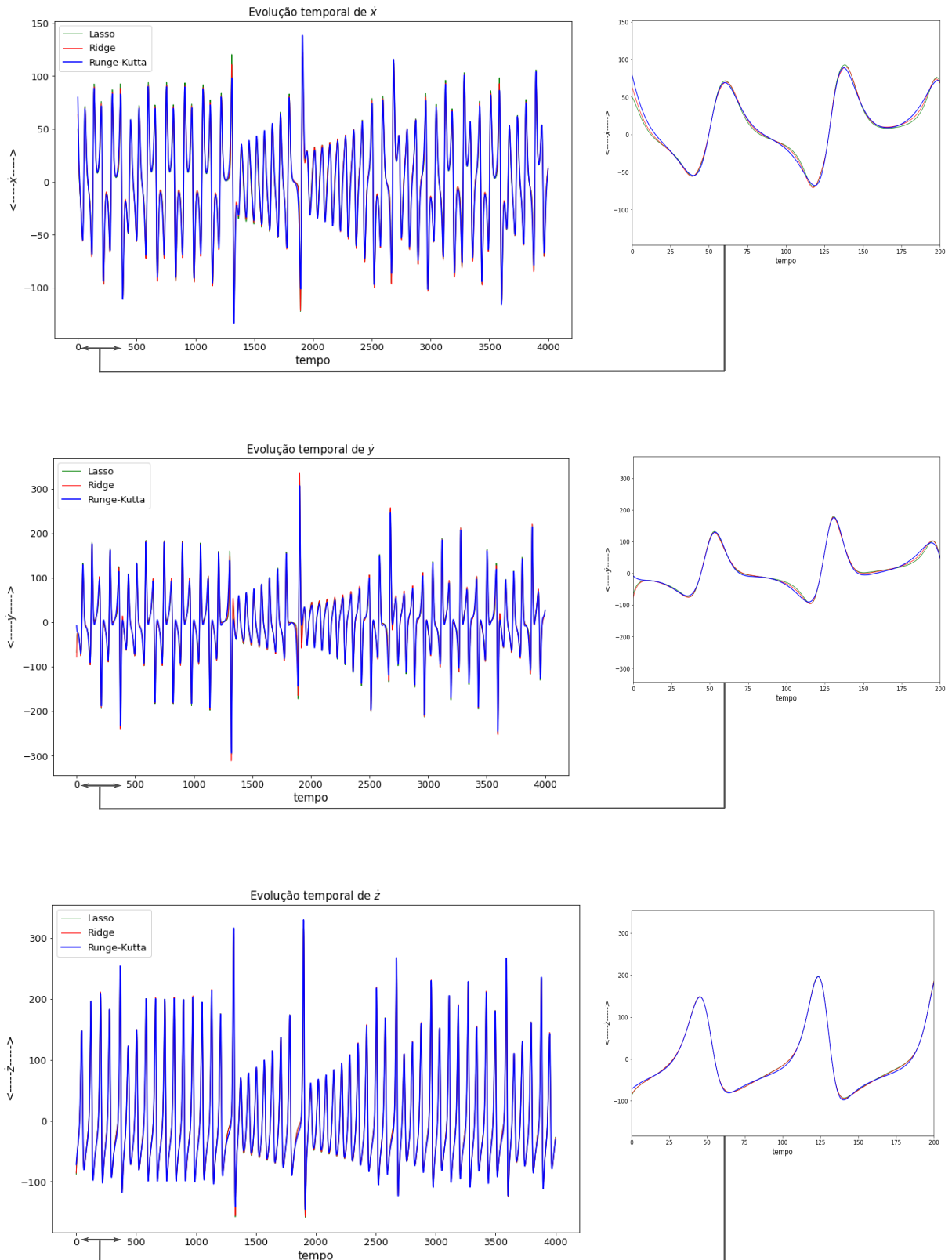
Figura 10 – A matriz  $\mathbf{V}$  e  $\Theta$  são implementadas nos métodos de regressão Ridge e Lasso para encontrar os coeficientes que regem as equações governantes do sistema.



Fonte: Adaptada do artigo (BRUNTON *et al.*, 2016).

Na implementação, foi usado o parâmetro  $\lambda = 10$  e após encontrar os valores de  $\Omega$ , eles foram usados para visualizar o comportamento das velocidades do sistema, a partir da Equação 4.3. A Figura 11 apresenta a comparação da evolução temporal resultantes dos três métodos utilizados neste trabalho.

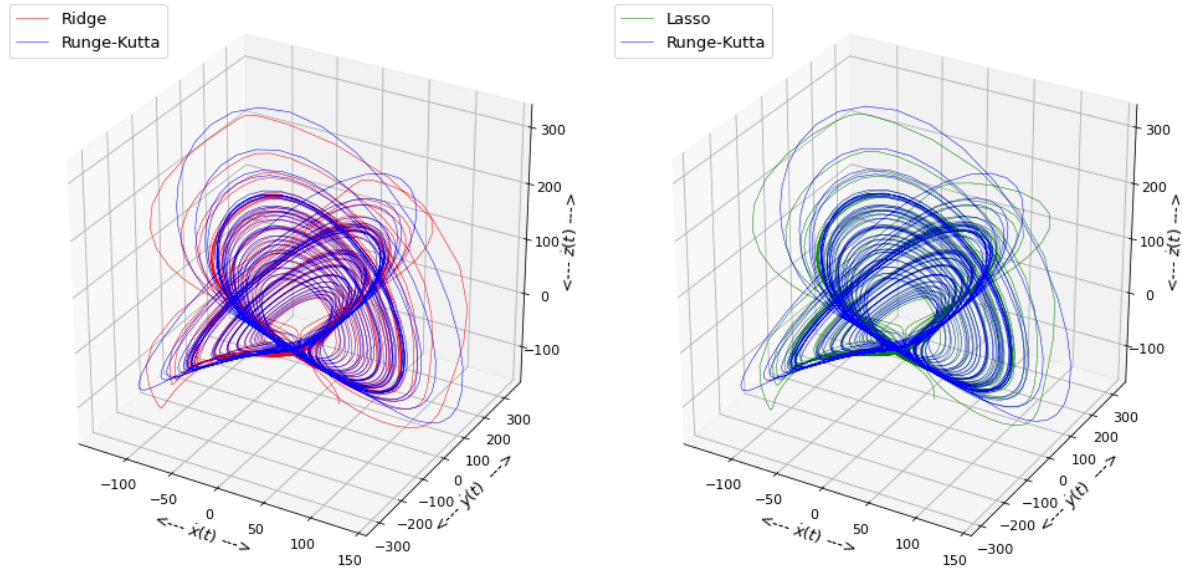
Figura 11 – Evoluções temporais de  $\dot{x}$ ,  $\dot{y}$  e  $\dot{z}$  pelos método Runge-Kutta de quarta ordem, regressão Ridge e Lasso juntamente com a ampliação do gráfico no intervalo de tempo  $[0,200]$ .



Fonte: Elaborada pelo autor (2022).

Pode-se visualizar também o comportamento das velocidades de forma tridimensional como apresentada na Figura 14.

Figura 12 – Comparação das soluções das velocidades para os métodos de regressão com o método Runge-Kutta de quarta ordem. À direita é comparado com a regressão Ridge e à esquerda com a regressão Lasso.



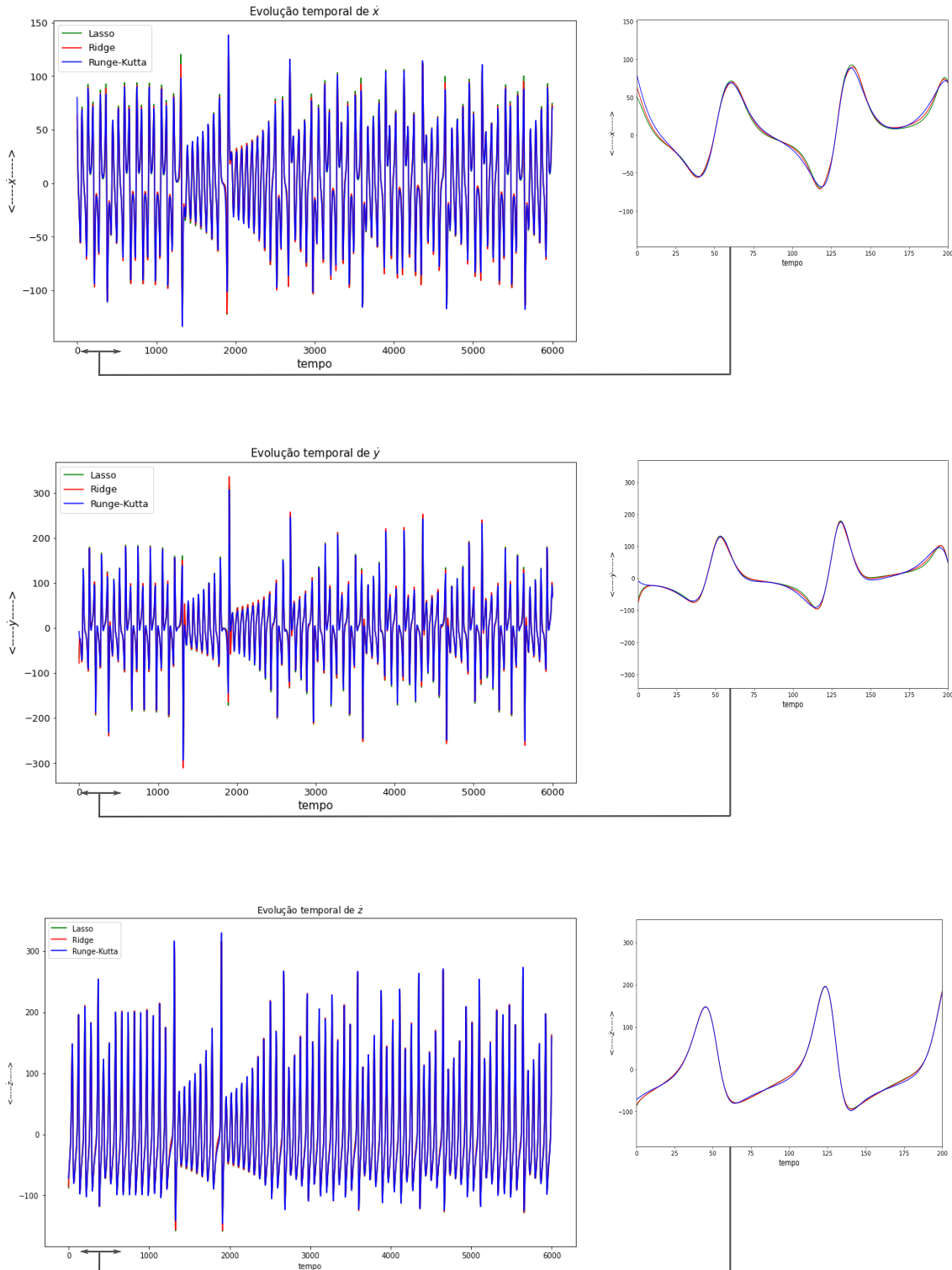
Fonte: Elaborado pelo autor (2022).

Tanto o método de regressão Ridge quanto a regressão Lasso se aproximaram com as curva de solução calculada pelo método Runge-Kutta, mas sofreram ajustes nos valores iniciais. A aproximação das curvas no intervalo de tempo  $[0,200]$  pela Figura 11 evidencia isso. Provavelmente as regressões estão com *overfitting*. Na Figura 14 se tem uma apresentação tridimensional do comportamento entre os três métodos. O comportamento das velocidades mostrado é o mesmo, ela se comporta em forma de espirais até mudar de direção e espiralar novamente. Assim, as regressões resultaram em um comportamento bem aproximado da realidade. Porém, a regressão Lasso não retornou  $\Omega$  de forma esparsa, assim, os coeficientes que descrevem as equações do sistema não foram exatamente encontrados e apenas foi possível se obter o comportamento temporal do sistema.

A partir dos coeficientes computados, a previsão para um sistema das equações de Lorenz foi realizado para períodos maiores. Os parâmetros  $\sigma$ ,  $\rho$ ,  $\beta$  e valores iniciais  $x(0)$ ,  $y(0)$  e  $z(0)$  foram mantidos, o intervalo de tempo foi mudado para  $[0,60]$  permanecendo o passo de  $\Delta t = 0,001$ . Com esses valores, os estados e velocidades foram calculados novamente pelo método Runge-Kutta e usamos diretamente os valores dos estados em  $\Theta$  ( $6000 \times 56$ ) e  $\Omega$  ( $6000 \times 3$ ) calculado anteriormente com objetivo de obter a previsão de  $\mathbf{V}(6000 \times 3)$ .

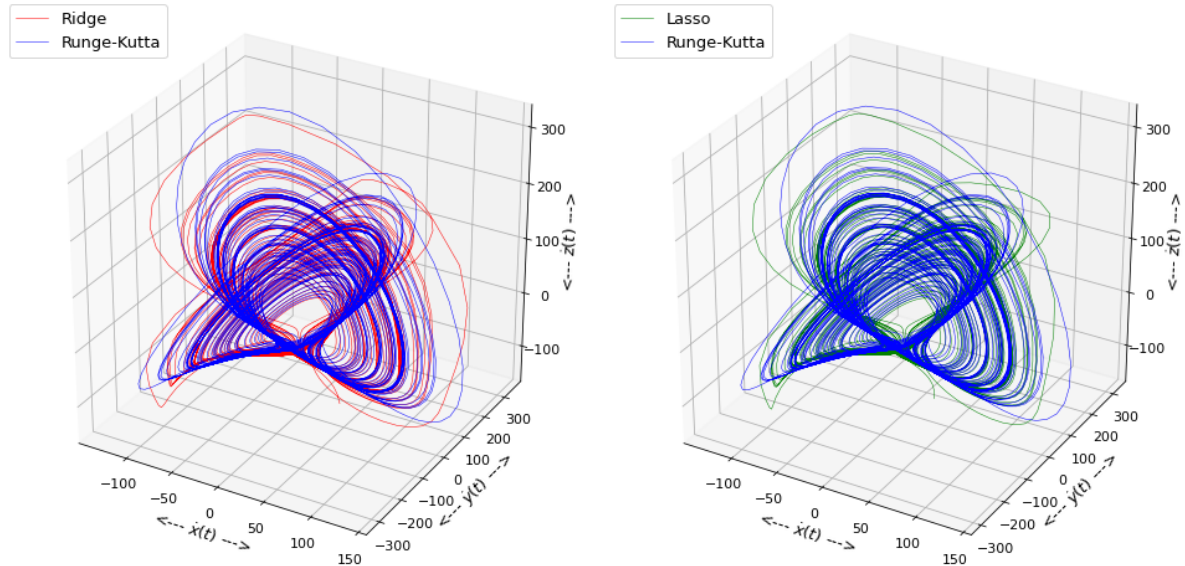
O comportamento da velocidade foi comparado para todos os métodos utilizados e essa previsão das evoluções temporais das velocidades são apresentadas na Figura 13 e o seu gráfico tridimensional é ilustrado na Figura 14.

Figura 13 – Gráficos da previsão das evoluções temporais de  $\dot{x}$ ,  $\dot{y}$  e  $\dot{z}$  para um período maior. Os gráficos foram plotados para o método Runge-Kutta de quarta ordem e pelos coeficientes encontrados anteriormente pelos métodos de regressão Ridge e Lasso. A ampliação dos gráficos está no intervalo  $[0,200]$



Fonte: Elaborada pelo autor (2022).

Figura 14 – Comparação da previsão de soluções das velocidades para os coeficientes computados anteriormente pelas regressões com o método Runge-Kutta de quarta ordem. À direita é comparado com os coeficientes da regressão Ridge e à esquerda pelos coeficientes da regressão Lasso.



Fonte: Elaborado pelo autor (2022).

Ambas as regressões realizaram previsões que coincidiram com a curva de solução do método Runge-Kutta de quarta ordem, mas o ajuste realizado nas condições iniciais ainda é aparente como mostrado na ampliação da imagem no intervalo  $[0,200]$ . Portanto, mesmo a regressão Lasso não retornando uma matriz esparsa de  $\Omega$ , ela conseguiu ajustar de forma que os coeficientes conseguem realizar previsões. Assim, tanto a regressão Lasso quanto a regressão Ridge conseguiram prever o comportamento para períodos futuros a partir de  $\Omega$  e condizeram com o que é previsto pelo Runge-Kutta de quarta ordem havendo apenas uma mudança nas condições iniciais.

## 5 CONCLUSÃO

O objetivo deste trabalho foi encontrar as equações governantes de um sistema não linear por meio das informações que são coletadas sobre ele. O sistema não linear utilizado foi baseado no problema de convecção atmosférica estudado por Edward Lorenz o qual apresenta um comportamento caótico. Métodos de aprendizado de máquina, como os métodos de regressão Ridge e Lasso, foram abordados com a intenção de obter os coeficientes dessas equações.

Por meio disso, foi realizado um estudo sobre o comportamento deste sistema a partir da evolução temporal das velocidades usando os métodos de regressão e o comparamos com o método Runge-Kutta de quarta ordem, o qual é um método quantitativo utilizado para obter soluções numéricas aproximadas de equações diferenciais ordinárias conhecidas. A matriz de coeficientes geradas pelas regressões não retornaram matrizes esparsas e portanto não resultou os coeficientes exatos das equações de Lorenz.

A partir dos valores calculados para os coeficientes ainda foi possível visualizar o comportamento do sistema e realizar previsões para períodos maiores, porém houve um ajuste nas condições iniciais. Esse ajuste não era esperado, uma vez que as soluções devem ser inicializadas no mesmo ponto.

A intenção futura deste trabalho é obter as equações de movimento de um sistema dinâmico não linear qualquer usando como meio somente o conjunto abundante de dados coletados sobre o sistema estudado.

## REFERÊNCIAS

- AHMAD, A. A. S. **A Textbook on Ordinary Differential Equations**. [S. l.]: Spring, 2013. Second Edition.
- ATMANSPACHER, H.; SCHEINGRABER, H. Deterministic chaos and dynamical instabilities in a multimode cw dye laser. **Physical Review A**, APS, v. 34, n. 1, p. 253, 1986.
- BHATTACHARYYA, S. **Ridge and Lasso Regression: L1 and L2 Regularization**. 2018. Disponível em: <https://towardsdatascience.com/ridge-and-lasso-regression-a-complete-guide-with-python-scikit-learn-e20e34bcbf0b>. Acesso em: 29 dez. 2021.
- BRUNTON, S. L.; PROCTOR, J. L.; KUTZ, J. N. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. **Proceedings of the National Academy of Sciences**, National Academy of Sciences, v. 113, n. 15, p. 3932–3937, 2016. ISSN 0027-8424. Disponível em: <https://www.pnas.org/content/113/15/3932>.
- FRIEDMAN, J.; HASTIE, T.; TIBSHIRANI, R. *et al.* **The elements of statistical learning**. [S. l.]: Springer series in statistics New York, 2001. v. 1.
- GIORDANO, H. N. N. **Computational Physics**. [S. l.]: Pearson, 2005. Second edition.
- GÉRON, A. **Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems**. [S. l.]: O'Reilly Media, 2017.
- LORENZ, E. N. Deterministic nonperiodic flow. **Journal of atmospheric sciences**, v. 20, n. 2, p. 130–141, 1963.
- LORENZ, F. L. E. N. **The Essence of Chaos**. [S. l.]: University of Washington Press, 1995.
- LOTKA, A. J. **Elements of Physical Biology**. [S. l.]: BALTIMORE, WILLIAMS WILKINS COMPANY, 1925. First Edition.
- MATSON, L. E. The malkus–lorenz water wheel revisited. **American Journal of Physics**, American Association of Physics Teachers, v. 75, n. 12, p. 1114–1122, 2007.
- MOON, F. C. Experiments on chaotic motions of a forced nonlinear oscillator: strange attractors. 1980.
- POINCARÉ, H.; MAITLAND, F. **Science and method**. [S. l.]: Courier Corporation, 2003.
- POL, B. Van der; MARK, J. V. D. Frequency demultiplication. **Nature**, Nature Publishing Group, v. 120, n. 3019, p. 363–364, 1927.
- STROGATZ, S. H. **Nonlinear dynamics and chaos with student solutions manual: With applications to physics, biology, chemistry, and engineering**. [S. l.]: CRC press, 2018.
- SUSSMAN, G. J.; WISDOM, J. Numerical evidence that the motion of pluto is chaotic. **Science**, American Association for the Advancement of Science, v. 241, n. 4864, p. 433–437, 1988.
- TAYLOR, J. R. **Mecânica Clássica**. [S. l.]: Bookman, 2013. Primeira Edição.



TIBSHIRANI, R. Regression shrinkage and selection via the lasso. **Journal of the Royal Statistical Society: Series B (Methodological)**, Wiley Online Library, v. 58, n. 1, p. 267–288, 1996.

WIENER, N. The homogeneous chaos. **American Journal of Mathematics**, JSTOR, v. 60, n. 4, p. 897–936, 1938.