

Previsão da trajetória de um sistema massa-mola-amortecedor não-linear por meio de redes neurais lagrangianas[★]

Sara A. Dias* Icaro M. F. de Santana* Saulo O. D. Luiz*
Antonio M. N. Lima*

* Departamento de Engenharia Elétrica, Universidade Federal de
Campina Grande, PB, (e-mails: {sara.dias,
icaro.santana}@ee.ufcg.edu.br, {saulo, amnlima}@dee.ufcg.edu.br).

Abstract: In this work, Lagrangian Neural Networks (LNNs) were applied to determine the Lagrangian of a mechanical system and to predict its trajectory in the state space. Neural networks, Lagrangian mechanics, Newtonian mechanics, dynamic models, numerical integration, parametric optimization, and evaluation criteria were applied for the case study of a non-linear mass-spring-damper system. The trajectories predicted by means of the Lagrangian Neural Networks were compared to the data generated by means of computer simulations of Newton's second law. The main conclusion of this work is that it is feasible to predict the trajectory of this system by means of Lagrangian Neural Networks. Furthermore, the flexibility of the LNN allowed the modeling of non-linear dynamics in the system.

Resumo: Neste trabalho, foram aplicadas Redes Neurais Lagrangianas (RNLs) para determinar o Lagrangiano de um sistema mecânico e prever sua trajetória no espaço de estados. Foram aplicadas as redes neurais, mecânica Lagrangiana, mecânica Newtoniana, modelos dinâmicos, integração numérica, otimização paramétrica, e critérios de avaliação para o estudo de caso de um sistema massa-mola-amortecedor não-linear. As trajetórias previstas por meio das Redes Neurais Lagrangianas foram comparadas aos dados gerados por meio de simulações computacionais da segunda lei de Newton. A conclusão principal deste trabalho é a viabilidade de se prever a trajetória desse sistema por meio das Redes Neurais Lagrangianas. Além do mais, a flexibilidade da RNL permitiu a modelagem de dinâmicas não-lineares no sistema.

Keywords: Lagrangian mechanics; Newtonian mechanics; Lagrangian Neural Networks; modeling; system energy.

Palavras-chaves: mecânica Lagrangiana; mecânica Newtoniana; Redes Neurais Lagrangianas; modelagem; energia do sistema.

1. INTRODUÇÃO

Por meio das redes neurais é possível implementar modelos matemáticos que são similares a estruturas neurais biológicas. Também é possível adaptar os parâmetros de tais redes em decorrência das interações com os dados de treinamento e de teste, melhorando gradativamente seu desempenho na solução de uma determinada tarefa. A possibilidade de criar uma rede neural capaz de aprender o Lagrangiano do sistema é muito promissora, tendo em vista que os Lagrangianos impõem a conservação da energia total, além de que esse processo pode ser feito utilizando coordenadas arbitrárias.

Embora existam diversos modelos para determinar o movimento dos sistemas mecânicos, para este trabalho será utilizada a formulação lagrangiana. A utilização da formulação de Lagrange quase sempre permite oferecer um meio simples para descrever as equações de movimento de um sistema. No entanto, nem sempre as equações resultantes

são fáceis de resolver, visto que existem casos em que as equações de Euler-Lagrange podem não ter uma solução analítica (Taylor, 2005).

As Redes Neurais Lagrangianas (RNLs) não requerem coordenadas canônicas e, portanto, funcionam bem em situações onde os momentos canônicos são desconhecidos ou difíceis de calcular. Lutter et al. (2019b) apresentaram redes neurais Lagrangianas profundas, em que a estrutura de uma rede neural profunda é usada para aprender o a dinâmica inversa de um sistema mecânico por meio da mecânica Lagrangiana. Lutter et al. (2019a) também estenderam as redes neurais Lagrangianas profundas para assegurar a conservação da energia e a passividade no domínio do treinamento dado um erro de treinamento suficientemente baixo.

Cranmer et al. (2020) apresentaram RNLs por meio das quais é possível aprender Lagrangianos arbitrários, os quais dispensam uma solução analítica para as equações de Euler-Lagrange. Por sua vez, Allen-Blanchette et al. (2020) expressaram as equações de Euler-Lagrange explicitamente por meio de uma matriz de inércia, um termo

* Os autores agradecem o suporte financeiro provido pelo PPgE-E/UFCEG, CAPES e CNPq.

de Coriolis, o gradiente da energia potencial em relação à posição e as forças generalizadas; e em seguida aplicaram RNLs para a predição de vídeo.

Roehrl et al. (2020) e Liu et al. (2021) apresentaram a estimação da dinâmica de um sistema por meio da decomposição do campo de forças nas suas componentes conservativas e não conservativas, representadas respectivamente por uma rede neural lagrangiana e uma rede de aproximação universal. Aoshima et al. (2021) propuseram uma rede neural baseada numa representação no tempo discreto para a mecânica Lagrangiana. Uma vantagem da rede neural proposta é que ela é treinada por meio apenas dos dados de posição, dispensando a medição da velocidade. Bhattoo et al. (2021) apresentaram uma RNL com conservação do momento, por meio da qual é possível aprender o Lagrangiano de um sistema e preservar as simetrias translacional e rotacional.

Neste trabalho, foi inicialmente determinada a expressão analítica da aceleração de um sistema mecânico a partir das equações de Euler-Lagrange na forma vetorial incluindo-se forças de atrito e forças exógenas. A partir desta expressão analítica da aceleração foi apresentado o método para treinamento e teste de uma rede neural lagrangiana com o objetivo de aprender o Lagrangiano de um sistema mecânico e prever a sua trajetória. Esse método portanto, tem como resultado o modelo dinâmico do sistema, no qual não são utilizados parâmetros físicos do mesmo, mas sim o Lagrangiano aprendido a partir de dados experimentais e aplicado às equações de Euler-Lagrange. Esse método foi utilizado no estudo de caso de um sistema massa-mola-amortecedor não-linear. Os resultados obtidos foram discutidos e analisados.

As etapas do método utilizado neste trabalho consistiram em:

- escolha do sistema físico que será modelado por meio da Rede Neural Lagrangiana;
- planejamento experimental para aquisição de dados do sistema físico para uso na RNL, considerando o os sinais de entrada (exógenos), condições iniciais, intervalo de amostragem dos sinais e a quantidade de dados;
- escolha dos parâmetros da rede neural, tais como dimensionamento do número de camadas, número de neurônios, número de iterações de treino, bem como a proporção de dados para treino e testes de validação; e
- após os treinos da RNL, realizar as etapas de teste e validação, com os parâmetros do sistema físico aprendidos pela RNL.

Na próxima seção são apresentados os fundamentos sobre a mecânica lagrangiana.

2. MECÂNICA LAGRANGIANA

Uma alternativa à formulação Newtoniana para modelagem de sistemas físicos é a formulação Lagrangiana, a qual faz uso das energias cinética T e potencial U . A formulação Lagrangiana oferece algumas vantagens em relação ao modelo mecânico de Newton, como por exemplo, o fato de o sistema de coordenadas utilizado não interferir na forma das equações Lagrangianas. Além disso, em sistemas com

restrições, tais como um pêndulo, nas equações Lagrangianas não é necessário considerar as forças de restrição, tais como a força normal na haste de um pêndulo, que mantém o objeto na extremidade do pêndulo numa trajetória circular, mas que é tipicamente desconhecida.

O Lagrangiano é definido como $\mathcal{L} = T - U$. Considerando um sistema de $n \in \mathbb{N}$ coordenadas generalizadas em que $q = [q_i]_{n \times 1}$ é a posição de uma partícula, onde $i = 1, 2, \dots, n$ é o índice da coordenada q_i , e $\dot{q} = [\dot{q}_i]_{n \times 1}$ a sua velocidade, o Lagrangiano é uma função da posição e da velocidade, ou seja, $\mathcal{L}(q, \dot{q}) = T(\dot{q}) - U(q)$. Num sistema de coordenadas inercial em que apenas forças conservativas atuam sobre uma partícula, a trajetória percorrida por ela satisfaz as equações de Euler-Lagrange (Taylor, 2005):

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = 0, \quad i = 1, 2, \dots, n. \quad (1)$$

Num sistema de coordenadas inercial em que forças conservativas e forças não-conservativas atuam sobre uma partícula, torna-se necessário incluir as forças não-conservativas nas equações de Euler-Lagrange. A energia potencial U deve corresponder apenas às forças conservativas. Sendo o somatório das forças não-conservativas representadas por $N = [N_i]_{n \times 1}$, as equações de Euler-Lagrange podem ser escritas como (Goldstein et al., 2002):

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = N_i, \quad i = 1, 2, \dots, n. \quad (2)$$

Quando forças de atrito são modeladas como proporcionais à velocidade de uma partícula, ou seja, o somatório das forças de atrito é $F = -[c_i \dot{q}_i]_{n \times 1}$, onde $i = 1, 2, \dots, n$, e c_i é o coeficiente de atrito viscoso na coordenada q_i , então define-se a função de dissipação de Rayleigh (Goldstein et al., 2002) como

$$\mathcal{F} = \frac{1}{2} \sum_{i=1}^n c_i \dot{q}_i^2. \quad (3)$$

Forças exógenas, tais como as produzidas por atuadores, podem ser aplicadas à partícula. O somatório das forças exógenas é $E = [E_i]_{n \times 1} = N - F$. Num sistema de coordenadas inercial, a trajetória percorrida por uma partícula satisfaz as seguintes equações de Euler-Lagrange (Goldstein et al., 2002):

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} + \frac{\partial \mathcal{F}}{\partial \dot{q}_i} = E_i, \quad i = 1, 2, \dots, n. \quad (4)$$

Neste trabalho, as equações de Euler-Lagrange (4) são escritas na forma vetorial:

$$\frac{d}{dt} \nabla_{\dot{q}} \mathcal{L} - \nabla_q \mathcal{L} + \nabla_{\dot{q}} \mathcal{F} = E, \quad (5)$$

onde ∇ é o operador gradiente. Usando a regra da cadeia:

$$\frac{d}{dt} \nabla_{\dot{q}} \mathcal{L}(q(t), \dot{q}(t)) = (\nabla_q \nabla_{\dot{q}}^T \mathcal{L}) \dot{q} + (\nabla_{\dot{q}} \nabla_{\dot{q}}^T \mathcal{L}) \ddot{q}. \quad (6)$$

Substituindo a expressão de $\frac{d}{dt} \nabla_{\dot{q}} \mathcal{L}$ de (6) em (5):

$$(\nabla_{\dot{q}} \nabla_{\dot{q}}^T \mathcal{L}) \ddot{q} + (\nabla_q \nabla_{\dot{q}}^T \mathcal{L}) \dot{q} - \nabla_q \mathcal{L} + \nabla_{\dot{q}} \mathcal{F} = E, \quad (7)$$

onde os termos

$$\begin{aligned} \nabla_{\dot{q}} \nabla_{\dot{q}}^T \mathcal{L} &= \left[(\nabla_{\dot{q}} \nabla_{\dot{q}}^T \mathcal{L})_{ij} \right]_{n \times n} = \left[\frac{\partial^2 \mathcal{L}}{\partial \dot{q}_j \partial \dot{q}_i} \right]_{n \times n} \\ \nabla_q \nabla_q^T \mathcal{L} &= \left[(\nabla_q \nabla_q^T \mathcal{L})_{ij} \right]_{n \times n} = \left[\frac{\partial^2 \mathcal{L}}{\partial q_j \partial q_i} \right]_{n \times n} \end{aligned} \quad (8)$$

são matrizes, $i = 1, 2, \dots, n$, e $j = 1, 2, \dots, n$. Sendo assim, se $\nabla_{\dot{q}} \nabla_{\dot{q}}^T \mathcal{L}$ é não-singular, por meio de (9) é possível estimar a aceleração $\hat{\ddot{q}}$, a partir dos dados medidos da posição q e velocidade \dot{q} .

$$\hat{\ddot{q}} = (\nabla_{\dot{q}} \nabla_{\dot{q}}^T \mathcal{L})^{-1} [\nabla_q \mathcal{L} - (\nabla_q \nabla_q^T \mathcal{L}) \dot{q} - \nabla_{\dot{q}} \mathcal{F} + E] \quad (9)$$

Nem sempre as equações de Euler-Lagrange são fáceis de resolver, visto que existem casos em que elas podem não ter uma solução analítica (Taylor, 2005), como por exemplo os casos em que o sistema é não conservativo e estão presentes forças de atrito não-lineares. Mas, Cranmer et al. (2020) introduziram uma nova classe de redes neurais, as Redes Neurais Lagrangianas (RNLs), por meio das quais é possível aprender Lagrangianos arbitrários, os quais dispensam uma solução analítica para as equações de Euler-Lagrange. As RNLs são apresentadas na próxima seção.

3. REDES NEURAIS LAGRANGIANAS

A rede neural lagrangiana é treinada para aprender parâmetros a partir das interações com os dados de treino, definindo assim um modelo de Lagrangiano capaz de descrever o sistema físico. A partir do resultado do Lagrangiano, pode-se calcular por meio das equações de Euler-Lagrange (9) a aceleração em determinado instante de tempo. Sendo assim, é possível prever a trajetória do sistema para o qual a rede neural foi treinada.

Tradicionalmente a expressão do Lagrangiano é escrita de forma analítica e então são definidas as equações de Euler-Lagrange em um sistema de equações diferenciais. No entanto, como o Lagrangiano representa uma caixa preta para a rede neural, deve-se admitir que pode não haver uma solução analítica da equação de Euler-Lagrange. Todavia, ainda é possível derivar uma expressão numérica para a aceleração \ddot{q} do sistema mecânico.

Para implementar a rede neural lagrangiana neste trabalho foi utilizado como referência o código implementado por Sam Greydanus, Miles Cranmer e Stephan Hoyer (2021)¹. Nesta implementação é utilizado o modelo físico do pêndulo duplo. Estes autores propuseram uma rede neural lagrangiana, que aproveita o poder computacional das redes neurais para aprender o Lagrangiano a partir de um conjunto de dados de trajetórias simuladas de sistemas mecânicos. Em relação a (9), as matrizes $\nabla_{\dot{q}} \nabla_{\dot{q}}^T \mathcal{L}$ e $\nabla_q \nabla_q^T \mathcal{L}$, e o Gradiente ($\nabla_q \mathcal{L}$ e $\nabla_{\dot{q}} \mathcal{F}$) são calculados por meio de diferenciação automática.

Na Figura 1 é apresentado o diagrama de blocos do modelo dinâmico do sistema por meio da Rede Neural Lagrangiana e das equações de Euler-Lagrange, como proposto neste

¹ CRANMER, Miles. **Lagrangian Neural Networks**. Disponível em: https://github.com/MilesCranmer/lagrangian_nns. Acesso em 30 mar. 2022.

trabalho. A RNL é formada por 4 camadas. A primeira camada (camada de entrada) possui dois neurônios, e nela são inseridos os dados de posição q e velocidade \dot{q} , por meio dos quais o algoritmo de treinamento da rede neural treina a RNL. As duas camadas seguintes, chamadas de camadas ocultas, apresentam 128 neurônios cada uma e são responsáveis por realizar cálculos intermediários que auxiliam a rede a determinar os valores finais. A última camada (camada de saída) apresenta apenas um neurônio, que fornece o Lagrangiano arbitrário \mathcal{L} aprendido pela RNL. A partir do Lagrangiano, a posição q , a velocidade \dot{q} , a função de dissipação de Rayleigh \mathcal{F} , e o somatório das forças exógenas E , calcula-se $\hat{\ddot{q}}$ por meio de (9). Ao integrar o valor de $\hat{\ddot{q}}$, encontra-se a velocidade $\hat{\dot{q}}$ do sistema físico, que integrada resulta na posição \hat{q} do mesmo.

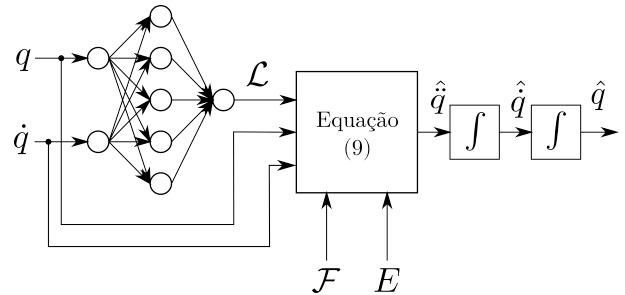


Figura 1. Diagrama de blocos do modelo dinâmico do sistema por meio da Rede Neural Lagrangiana e das equações de Euler-Lagrange.

Cada neurônio é caracterizado pelo peso, *bias* e função de ativação. Os pesos dos neurônios da RNL são inicializados de forma randomizada. Em cada neurônio de cada camada, multiplica-se o valor de entrada pelo peso do neurônio correspondente, soma-se com o *bias* associado ao neurônio, e, por fim, esse valor é passado por uma função de ativação, que agrega a capacidade não-linear ao processamento realizado pelas redes neurais. Sendo assim, cada camada apresenta uma função de ativação específica. Na RNL, utiliza-se a função de ativação *softplus*. Tendo em vista a necessidade de calcular as matrizes $\nabla_{\dot{q}} \nabla_{\dot{q}}^T \mathcal{L}$ e $\nabla_q \nabla_q^T \mathcal{L}$, e o Gradiente ($\nabla_q \mathcal{L}$ e $\nabla_{\dot{q}} \mathcal{F}$), a derivada de segunda ordem da função de ativação é importante. Por exemplo, a função de ativação ReLU é não-linear, ou seja, sua derivada de segunda ordem é zero, tornando-a uma escolha ruim para a RNL. Para encontrar uma melhor função de ativação, Cranmer et al. (2020) realizaram testes com diversas funções de ativação (*ReLU*², *ReLU*³, tanh, sigmoid, e *softplus*) e descobriram que a *softplus* obteve o melhor desempenho e, portanto, esta é a função de ativação aplicada na RNL.

As funções de perda são utilizadas para medir o quanto uma rede neural se aproximou ou se afastou do valor real durante o treino, e assim ajustar os pesos dos neurônios. A função de perda V_N utilizada neste trabalho foi o erro médio quadrático (10), que computa a média quadrática da diferença entre os dados experimentais $q(t_k)$ e $\dot{q}(t_k)$ e seus respectivos valores previstos $\hat{q}(t_k)$ e $\hat{\dot{q}}(t_k)$ pela RNL, onde os instantes de amostragem são $t_k = kh$, $k = 1, 2, \dots, N$, e h é o intervalo de amostragem. Além da função de perda, foram utilizados os otimizadores que ajudam a minimizar os erros mais rapidamente. Neste trabalho, utilizou-se o otimizador Adam para o treinamento da RNL.

$$V_N = \frac{1}{2N} \sum_{k=1}^N \left[(q(t_k) - \hat{q}(t_k))^2 + (\dot{q}(t_k) - \hat{\dot{q}}(t_k))^2 \right] \quad (10)$$

A cada iteração do treinamento, o seguinte procedimento se repete. Inicialmente, são inseridos como entrada na RNL os dados de posição e velocidade, a partir dos quais a rede calcula um Lagrangiano arbitrário. Em seguida, são utilizadas as equações de Euler-Lagrange para estimar a aceleração do sistema a partir do Lagrangiano aprendido. A velocidade e a aceleração calculadas são armazenadas em um vetor, e então é computada a média quadrática da diferença entre os valores da velocidade e aceleração gerados analiticamente pela segunda lei de Newton e os valores previstos pela RNL. Com base nesse erro médio quadrático, são recalculados os parâmetros da rede para a próxima iteração. Os códigos implementados neste trabalho estão disponíveis em um repositório no *GitHub*².

Cranmer et al. (2020) fizeram comparações entre a RNL e outros tipos de redes neurais especializadas em prever a dinâmica de sistemas físicos, e observaram que a RNL apresentou resultados melhores. Ao realizar os testes com a RNL, esses autores verificaram que o erro médio quadrático ao fim do treinamento convergiu para valores entre $7,3$ e $7,4 \cdot 10^{-2}$. Então, adotou-se neste trabalho que o erro médio quadrático deve assumir valores menores ou iguais a $7,4 \cdot 10^{-2}$ para considerar os resultados da RNL aceitáveis.

A grande vantagem da utilização da RNL é o fato de que, para treiná-la, são necessários apenas os dados de posição e velocidade do sistema. Ou seja, não é necessário um modelo analítico para o sistema, muito menos representá-lo em coordenadas canônicas. Além dessa vantagem, para calcular a trajetória do sistema pelo Lagrangiano ou pela segunda lei de Newton é necessário estimar parâmetros físicos, tais como a massa e a constante elástica. No entanto, esta necessidade não existe quando utiliza-se uma RNL para prever a trajetória do sistema.

Na seção a seguir é apresentado o modelo matemático do sistema massa-mola-amortecedor não-linear, que foi utilizado neste trabalho como estudo de caso. Em seguida, o método de treinamento da RNL e os resultados de simulação são apresentados e analisados.

4. SISTEMA MASSA-MOLA-AMORTECEDOR NÃO-LINEAR

Considerou-se um sistema mecânico em que a mola apresenta um comportamento não-linear. Neste cenário, também considerou-se a influência da força de atrito modelada como um amortecedor com comportamento linear. O diagrama deste sistema é apresentado Figura 2. Neste cenário, o movimento é translacional e unidimensional, e a posição q do sistema é representada por uma única coordenada x , ou seja, $q = x$. Na Figura 2(a), são representados a mola com característica não-linear parametrizada por β_1 e β_2 , e um amortecedor linear parametrizado pelo coeficiente de atrito viscoso c . Como representado na Figura 2(b), ao ser conectada uma carga de massa m às extremidades livres da

² DIAS, Sara Andrade. **Lagrangian-Neural-Networks**. Disponível em: <https://github.com/sarandrade/Lagrangian-Neural-Networks>. Acesso em: 2 abr. 2022.

mola e do amortecedor, o sistema passa por um transitório e chega a um ponto de equilíbrio em que a deformação da mola é x_0 , e as intensidades da força elástica e da força peso são iguais. Se a mola sofrer uma nova deformação por uma condição inicial não-nula de x ou uma força exógena E , o sistema entra em um novo regime transitório representado na Figura 2(c), em que a deformação em relação ao ponto de equilíbrio é x . Neste cenário, a relação entre a força elástica e a posição x é não-linear: $\beta_1(x_0 + x)^{\beta_2}$. Na Figura 3, é representado o diagrama de corpo livre da carga de massa m , onde as forças que influenciam a dinâmica de movimento translacional do sistema estão representadas por vetores. Os valores dos parâmetros utilizados para gerar os dados de treinamento são apresentados na Tabela 1.

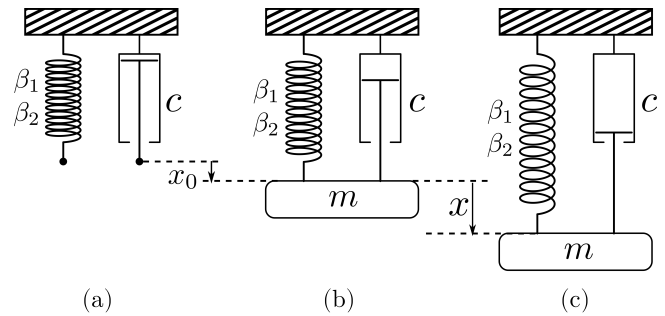


Figura 2. Diagrama do sistema massa-mola-amortecedor: (a) sistema sem carga acoplada; (b) sistema com carga acoplada e em estado de equilíbrio; (c) sistema com carga acoplada e em regime transitório.

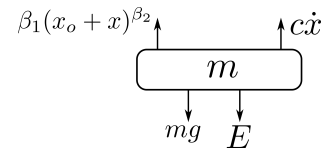


Figura 3. Diagrama de corpo livre da carga de massa m .

Tabela 1. Parâmetros do sistema massa-mola-amortecedor não-linear.

Parâmetro	Valor
β_1	161,9783
β_2	1,9099
Massa (m)	0,15 kg
Aceleração gravitacional (g)	9,8 m/s ²
Coefficiente de atrito viscoso (c)	1,2732 Ns/m

No ponto de equilíbrio em que a deformação da mola é x_0 , as intensidades da força elástica e da força peso são iguais. Portanto:

$$\beta_1(x_0)^{\beta_2} = mg \Rightarrow x_0 = \beta_2 \sqrt{\frac{mg}{\beta_1}} \quad (11)$$

Em torno do ponto de equilíbrio, a posição é representada pela coordenada x . Por meio da segunda lei de Newton, a função analítica da aceleração do sistema massa-mola-amortecedor não-linear é:

$$\begin{aligned} m\ddot{x} &= -\beta_1(x_0 + x)^{\beta_2} - c\dot{x} - mg \Rightarrow \\ \Rightarrow \ddot{x} &= \frac{-\beta_1(x_0 + x)^{\beta_2} - c\dot{x}}{m} - g \end{aligned} \quad (12)$$

4.1 Método

Linearizando-se o modelo do sistema representado em (12) em torno do ponto de operação $x = 0$ e $\dot{x} = 0$, por meio da linearização de pequenos sinais, encontra-se uma aproximação de $1,8/\sqrt{(\beta_1\beta_2x_0^{\beta_2-1})/m}$ para o tempo de subida t_r . Utilizando-se a aproximação $\frac{t_r}{10}$ para o intervalo de amostragem h do modelo em tempo discreto do sistema, encontra-se o valor de 0,031 s para este intervalo de amostragem. Foram testados os resultados da RNL para três diferentes intervalos de amostragem do sistema: 0,01 s, 0,005 s e 0,001 s.

A expressão analítica do Lagrangiano do sistema é apresentada em (13). Tal expressão pode ser usada nas equações de Euler-Lagrange (9) para a determinação da aceleração.

$$\mathcal{L} = \frac{1}{2}m\dot{x}^2 - \left(\frac{1}{\beta_2 + 1}\beta_1(x_0 + x)^{\beta_2+1} - mgx + \frac{1}{\beta_2 + 1}\beta_1x_0^{\beta_2+1} \right) \quad (13)$$

O estado inserido na rede neural é um vetor contendo a posição e a velocidade do sistema (x, \dot{x}) . O estado inicial do sistema é (0,012 m, 0 m/s). Integrando a velocidade e a aceleração, que são calculadas analiticamente pela segunda lei de Newton a cada iteração, obtém-se, assim, a posição e a velocidade do sistema no decorrer do tempo. Ou seja, tem-se a trajetória do sistema massa-mola-amortecedor não-linear calculada por meio da segunda lei de Newton. Devido ao estado inicial não-nulo, ocorre o transitório, e em consequência da força de atrito, a amplitude da oscilação diminui até que toda energia cinética e potencial sejam dissipadas e a massa converge para o ponto de equilíbrio $x = 0$ m, $\dot{x} = 0$ m/s.

Para inicializar o processo de treinamento da RNL, foram gerados 2.000 dados de treino no formato (x, \dot{x}) , que são os dados de entrada da rede, e 2.000 dados de treino no formato (\dot{x}, \ddot{x}) , utilizados para calcular o erro médio quadrático em relação aos dados previstos. O mesmo é feito para gerar os 2.000 dados de teste. Para o caso em que o intervalo de amostragem considerado foi de 0,001s, houve um aumento no número de dados de treino e de teste, passando de 2.000 dados para 3.000 dados. Esses dados foram obtidos por meio da mecânica Newtoniana. A partir da segunda lei de Newton, foi possível calcular a aceleração do sistema, e integrando este valor pelo método Runge-Kutta, foi possível obter o valor da velocidade. Sendo assim, dado um estado inicial e um intervalo de tempo de simulação, obtém-se a trajetória do sistema por meio da segunda lei de Newton.

Outras informações importantes sobre o treinamento da RNL são o número de lotes utilizados no treinamento e o tamanho destes lotes. Para todos os experimentos realizados, foram utilizados 1.500 lotes com o tamanho de 100 unidades cada. Estes valores foram utilizados especialmente para parametrizar a função que implementa o otimizador Adam. Além disso, estes valores também foram usados para calcular o número de iterações realizadas pela RNL, que foi igual à multiplicação entre o número de lotes

e seu tamanho, ou seja, a RNL realizou 150.000 iterações antes de finalizar o seu treinamento.

Outro uso do tamanho de cada lote com o valor igual a 100 unidades foi indicar que, a cada 100 iterações da RNL, realiza-se o cálculo do erro médio quadrático em relação aos dados de treino e de teste. Esses erros são armazenados em vetores, e ao fim do treinamento é possível exibir em um gráfico o decaimento deste erro com o passar das iterações.

A seguir, realizou-se o treinamento da RNL com os dados de treino gerados a partir da segunda lei de Newton. Foram calculados os erros médios quadráticos em relação aos dados de treino e aos dados de teste a cada 100 iterações. Nas Figuras 4, 5 e 6 são apresentados estes erros em função do número de iterações para os casos em que o intervalo de amostragem é, respectivamente, 0,01 s, 0,005 s e 0,001 s. Para os tempos de amostragem de 0,01 s e 0,005 s, a ordem de grandeza do erro médio quadrático foi de 10^{-2} ao fim do treinamento. Já para o intervalo de amostragem de 0,001 s, a ordem de grandeza do erro médio quadrático foi de 10^{-3} ao fim do treinamento.

A partir do Lagrangiano aprendido pela RNL, é possível, então, realizar a previsão da trajetória do sistema massa-mola-amortecedor não-linear por meio do modelo dinâmico do sistema apresentado na Figura 1. A posição e a velocidade da trajetória prevista são avaliadas neste trabalho por meio das seguintes métricas: as médias (14) e valores máximos (15) do valor absoluto das diferenças entre as posições calculadas pela segunda lei de Newton e previstas pela RNL; e as médias (16) e valores máximos (17) do valor absoluto das diferenças entre as velocidades calculadas pela segunda lei de Newton e previstas pela RNL.

$$\frac{1}{N} \sum_{k=1}^N |q(t_k) - \hat{q}(t_k)| \quad (14)$$

$$\max_k |q(t_k) - \hat{q}(t_k)| \quad (15)$$

$$\frac{1}{N} \sum_{k=1}^N |\dot{q}(t_k) - \hat{\dot{q}}(t_k)| \quad (16)$$

$$\max_k |\dot{q}(t_k) - \hat{\dot{q}}(t_k)| \quad (17)$$

4.2 Resultados

Os gráficos de posição e velocidade para os intervalos de amostragem de 0,01 s, 0,005 s e 0,001 s são apresentados respectivamente nas Figuras 7 e 8. Os resultados previstos pela RNL para o intervalo de amostragem de 0,001s são mais próximos dos calculados pela segunda lei de Newton, em comparação aos resultados previstos pela RNL para os intervalos de amostragem de 0,01s e 0,005s.

Na Figura 9 são apresentadas as diferenças entre as posições calculadas pela segunda lei de Newton e previstas pela RNL para os intervalos de amostragem de 0,01s, 0,005s e 0,001s. Já na Figura 10 são apresentadas as diferenças para a velocidade do sistema. Verifica-se novamente que os resultados previstos pela RNL para o intervalo de amostragem de 0,001s são mais próximos dos calculados

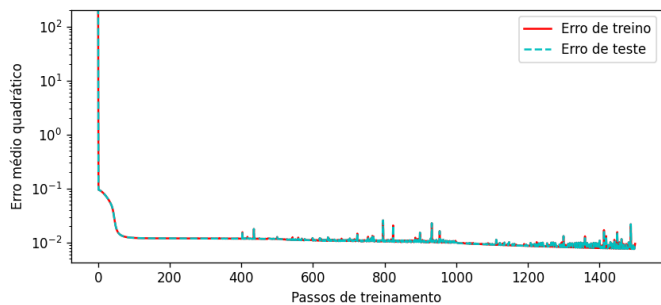


Figura 4. Erros de treino e teste calculados durante o treinamento para o sistema massa-mola-amortecedor não-linear com intervalo de amostragem $h = 0,01$ s.

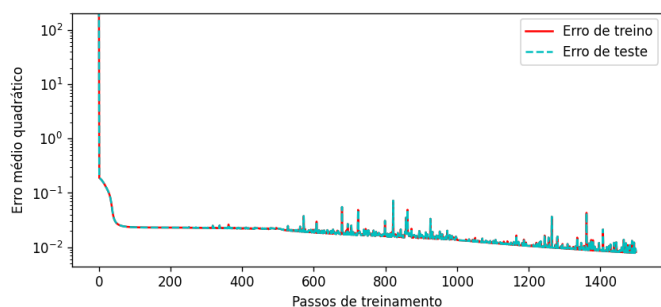


Figura 5. Erros de treino e teste calculados durante o treinamento para o sistema massa-mola-amortecedor não-linear com intervalo de amostragem $h = 0,005$ s.

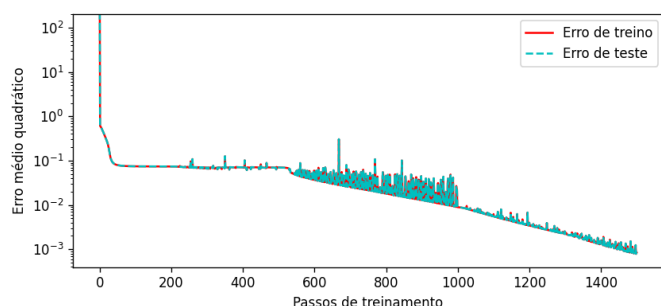


Figura 6. Erros de treino e teste calculados durante o treinamento para o sistema massa-mola-amortecedor não-linear com intervalo de amostragem $h = 0,001$ s.

pela segunda lei de Newton, em comparação aos resultados previstos pela RNL para os intervalos de amostragem de 0,01 s e 0,005 s.

Na Tabela 2 são apresentadas as médias (14) e valores máximos (15) do valor absoluto das diferenças entre as posições calculadas pela segunda lei de Newton e previstas pela RNL. Já na Tabela 3, são apresentados as médias (16) e valores máximos (17) do valor absoluto das diferenças entre as velocidades calculadas pela segunda lei de Newton e previstas pela RNL. Sendo assim, confirma-se que, para o menor intervalo de amostragem $h = 0,001$ s, ocorrem as menores diferenças entre a trajetória calculada pela segunda lei de Newton e a prevista pela RNL.

Para examinar as diferenças entre os Lagrangianos, o aprendido pela RNL e o gerado analiticamente, foram calculadas as diferenças entre a posição e a velocidade calculadas por Euler-Lagrange e as previstas pela RNL,

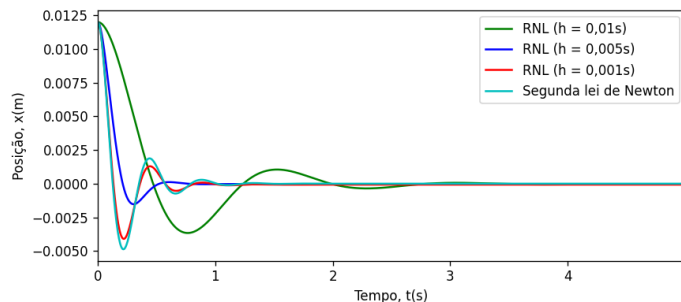


Figura 7. Posições calculadas pela segunda lei de Newton e previstas pela RNL para o sistema massa-mola-amortecedor não-linear.

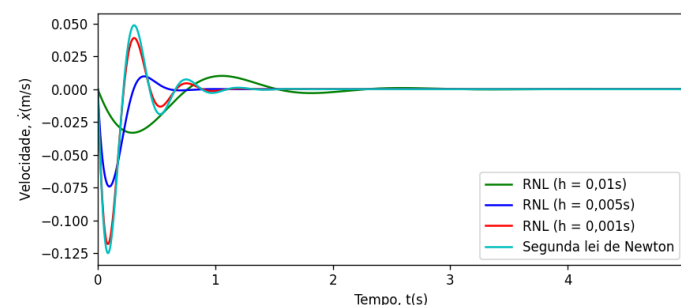


Figura 8. Velocidades calculadas pela segunda lei de Newton e previstas pela RNL para o sistema massa-mola-amortecedor não-linear.

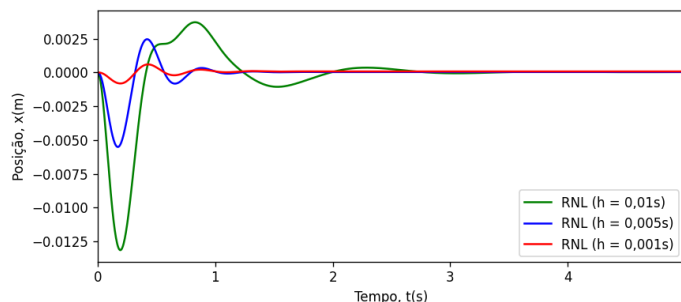


Figura 9. Diferença entre a posição calculada pela segunda lei de Newton e as posições previstas pela RNL para o sistema massa-mola-amortecedor não-linear.

com o intervalo de amostragem que apresentou os melhores resultados $h = 0,001$ s. Os resultados obtidos são apresentados nas Figuras 11 e 12.

Assim como nas Figuras 9 e 10, as diferenças observadas, para o intervalo de amostragem de 0,001 s, nas Figuras 11 e 12 apresentam uma ordem de grandeza de 10^{-4} para

Tabela 2. Valor absoluto das diferenças entre a posição calculada pela segunda lei de Newton e as posições previstas pela RNL.

Intervalo de amostragem h (s)	Diferença absoluta média (m)	Diferença absoluta máxima (m)
0,01	$1,09 \times 10^{-3}$	$1,31 \times 10^{-2}$
0,005	$3,26 \times 10^{-4}$	$5,52 \times 10^{-3}$
0,001	$1,17 \times 10^{-4}$	$8,16 \times 10^{-4}$

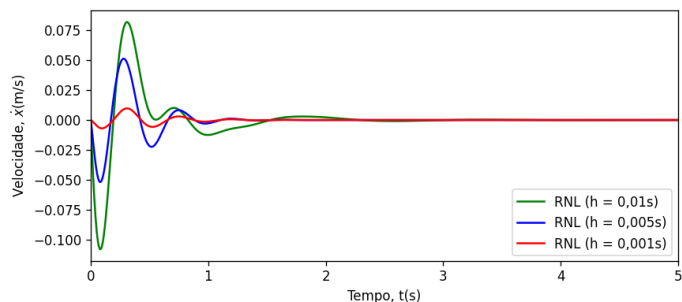


Figura 10. Diferença entre a velocidade calculada pela segunda lei de Newton e as velocidades previstas pela RNL para o sistema massa-mola-amortecedor não-linear.

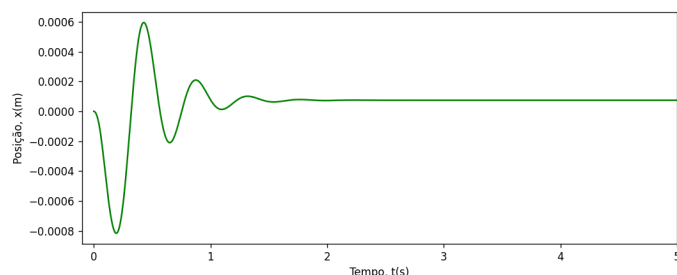


Figura 11. Diferença entre a posição calculada por meio das equações de Euler-Lagrange e a prevista pela RNL para o sistema massa-mola-amortecedor não-linear com intervalo de amostragem $h = 0,001$ s.

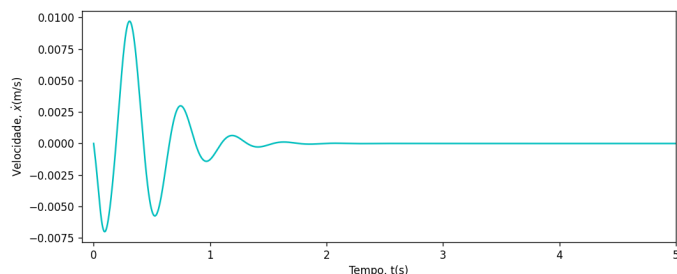


Figura 12. Diferença entre a velocidade calculada por meio das equações de Euler-Lagrange e a prevista pela RNL para o sistema massa-mola-amortecedor não-linear com intervalo de amostragem $h = 0,001$ s.

a posição e de até 10^{-3} para a velocidade. Apesar de as diferenças entre as velocidades apresentarem um pico com o valor próximo a $1 \cdot 10^{-2}$ m/s, este valor ainda é aproximadamente cinco vezes menor do que o valor obtido no caso em que o intervalo de amostragem é igual a 0,005 s. Já as ordens de grandeza da diferença de posição também são menores do que nos casos em que o intervalo

Tabela 3. Diferenças entre a velocidade calculada pela segunda lei de Newton e as velocidades previstas pela RNL.

Intervalo de amostragem h (s)	Diferença absoluta média (m/s)	Diferença absoluta máxima (m/s)
0,01	$7,36 \times 10^{-3}$	$1,08 \times 10^{-1}$
0,005	$3,72 \times 10^{-3}$	$5,18 \times 10^{-2}$
0,001	$7,61 \times 10^{-4}$	$9,71 \times 10^{-3}$

de amostragem é 0,01 s e 0,005 s. Além disso, para o caso em que o intervalo de amostragem considerado foi de 0,001s, o erro médio quadrático de teste e treino ao fim do treinamento foi bem menor que o erro obtido por Cranmer et al. (2020).

5. CONCLUSÃO

Neste trabalho, verificou-se a viabilidade da rede neural lagrangiana de aprender o Lagrangiano a partir das interações com os dados de treino gerados pela segunda lei de Newton. A partir do Lagrangiano aprendido pela RNL, foi possível realizar previsões da trajetória de um sistema mecânico massa-mola-amortecedor não-linear em um determinado intervalo de tempo. Os resultados previstos apresentaram erros com uma ordem de grandeza entre 10^{-2} e 10^{-3} , ou seja, a RNL teve desempenho satisfatório tanto nos dados de treino quanto nos dados de teste, indicando que a RNL conseguiu aprender satisfatoriamente a dinâmica do sistema usado no estudo de caso. Concluiu-se também que a escolha do intervalo de amostragem é relevante no projeto do experimento para coleta de dados para treinamento e teste da RNL.

Os próximos passos deste trabalho são: (i) a análise da influência da escolha do intervalo de amostragem sobre o treinamento e teste da rede neural, e a proposta de um método para a escolha adequada do intervalo de amostragem; (ii) a determinação de uma técnica para o dimensionamento da RNL, por exemplo, as escolhas do número de camadas, o número de neurônios por camada, e a função de ativação; e (iii) realizar experimentos para obtenção de dados de treino e teste reais, ou seja, que não sejam gerados analiticamente pela segunda lei de Newton, com o objetivo de avaliar a influência de ruído de medição e fenômenos não-lineares desconhecidos sobre o treinamento e teste da RNL.

REFERÊNCIAS

- Allen-Blanchette, C., Veer, S., Majumdar, A., and Leonard, N.E. (2020). Lagnetvip: A lagrangian neural network for video prediction. *CoRR*, abs/2010.12932. URL <https://doi.org/10.48550/arXiv.2010.12932>.
- Aoshima, T., Matsubara, T., and Yaguchi, T. (2021). Deep discrete-time lagrangian mechanics. In *Conf. Rec. of ICLR'2021 SimDL Workshop*.
- Bhattoo, R., Ranu, S., and Krishnan, N.M.A. (2021). Lagrangian neural network with differentiable symmetries and relational inductive bias. *CoRR*, abs/2110.03266. URL <https://doi.org/10.48550/arXiv.2110.03266>.
- Cranmer, M.D., Greydanus, S., Hoyer, S., Battaglia, P.W., Spergel, D.N., and Ho, S. (2020). Lagrangian neural networks. *CoRR*, abs/2003.04630. URL <https://doi.org/10.48550/arXiv.2003.04630>.
- Goldstein, H., Poole, C., and Saffko, J. (2002). *Classical mechanics*. Addison Wesley.
- Liu, Z., Wang, B., Meng, Q., Chen, W., Tegmark, M., and Liu, T. (2021). Machine-learning non-conservative dynamics for new-physics detection. *CoRR*, abs/2106.00026. URL <https://doi.org/10.48550/arXiv.2106.00026>.
- Lutter, M., Listmann, K., and Peters, J. (2019a). Deep lagrangian networks for end-to-end learning of energy-based control for under-actuated systems. In *Conf. Rec.*

- of *IROS'2019*, 7718–7725. URL <https://doi.org/10.1109/IROS40897.2019.8968268>.
- Lutter, M., Ritter, C., and Peters, J. (2019b). Deep lagrangian networks: Using physics as model prior for deep learning. *CoRR*, abs/1907.04490. URL <http://arxiv.org/abs/1907.04490>.
- Roehrl, M.A., Runkler, T.A., Brandtstetter, V., Tokic, M., and Obermayer, S. (2020). Modeling system dynamics with physics-informed neural networks based on lagrangian mechanics. *IFAC-PapersOnLine*, 53(2), 9195–9200. URL <https://doi.org/10.1016/j.ifacol.2020.12.2182>. 21st IFAC World Congress.
- Taylor, J.R. (2005). *Classical mechanics*. University Science Books.