UNIVERSIDAD NACIONAL DE INGENIERÍA
FACULTAD DE CIENCIAS
ESCUELA PROFESIONAL DE MATEMÁTICA

Control y simulación por métodos numéricos de un sistema de péndulo doble invertido (SPDI)

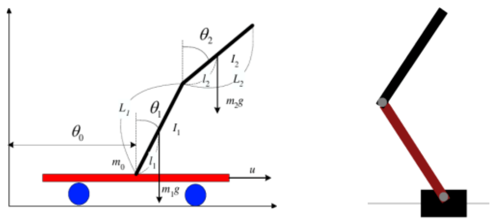
Walter J. Felipe Tolentino*, Jordi J. Bardales Rojas, Jhonatan J. Casaico Suarez, [1,2]Luis R. Roca Galindo
[1]Docente, Escuela Profesional de Matemática, Facultad de Ciencias
[2]Áreas: Ecuaciones Diferenciales - Análisis Numérico – Control – Mecánica Clásica - Machine Learning
*wfelipet@uni.pe *github.com/jesusuni/md4-spdi-openai

RESUMEN

El presente trabajo es una integración práctica de los trabajos **[1]** y **[2]**, tomando de **[3]** los parámetros para el ejercicio de simulación, exceptuando cualquier tipo de fricción. Se da como valor agregado el fundamento matemático en cada paso del proceso de modelación, control, simulación y aprendizaje del **SPDI**. Para la modelación se emplea las ecuaciones de **Euler-Lagrange**, llegando a un sistema no lineal **3-ED**, luego se establece las condiciones para la linealización en espacio de estados, se analiza la estabilidad, controlabilidad y observabilidad del sistema lineal y se procede a la simulación numérica en **Simulink-MATLAB(Runge-Kutta)** empleando **el controlador LQR**, obteniéndose la convergencia a lazo cerrado. Finalmente se simula gráficamente el sistema en **Python** usando la biblioteca **Gym-OpenAI**, con aprendizaje por refuerzo (**RL**), **MDP** y **RK.**

INTRODUCCIÓN

El problema de la estabilidad, controlabilidad y observabilidad del **SPDI**, como sistema dinámico no lineal, radica en establecer bajo que condiciones se puede linealizar el sistema y el efecto que esta produce, así como la elección del método óptimo de control cuya solución numérica no llegue a ser prohibitiva computacionalmente.



Objetivos

* Modelar matemáticamente el sistema no lineal y lineal.
* Estudio del controlador **LQR**.
* Analizar la estabilidad, controlabilidad Y observabilidad del sistema.
* Simular numéricamente el sistema en espacio de estados.
* Simulación gráfica por aprendizaje reforzado(**RL**) – **Open AI.**

MARCO TEÓRICO – MODELO MATEMÁTICO

SPDI : Es un sistema dinámico no lineal que presenta tres grados de libertad , es linealizable e invariante en el tiempo(**LTI**).

EC. EULER-LAGRANGE:

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}}\right) + \frac{\partial \mathcal{L}}{\partial \theta} = Q$$

$$\mathrm{D}(\theta)\ddot{\theta} + \mathrm{C}(\theta,\dot{\theta})\dot{\theta} + \mathrm{G}(\theta) = \mathrm{H}u$$

$$\mathrm{x} = (\theta\ \dot{\theta})^t$$

$$\dot{\mathrm{x}} = \begin{pmatrix} 0 & \mathrm{I} \\ 0 & -\mathrm{D}^{-1}\mathrm{C} \end{pmatrix}\mathrm{x} + \begin{pmatrix} 0 \\ -\mathrm{D}^{-1}\mathrm{G} \end{pmatrix} + \begin{pmatrix} 0 \\ \mathrm{D}^{-1}\mathrm{H} \end{pmatrix} u$$

LINEALIZACION (x=0) Y SISTEMA DE ESPACIO DE ESTADO :

$$\dot{\mathrm{x}} = \mathrm{Ax(t)} + \mathrm{B}u(t)$$
$$\mathrm{y} = \mathrm{Cx(t)} + \mathrm{D}u(t)$$

donde

$$\mathrm{A} = \begin{pmatrix} 0 & \mathrm{I} \\ -\mathrm{D}(0)^{-1}\partial_\theta \mathrm{G}(0) & 0 \end{pmatrix}$$

$$\mathrm{B} = \begin{pmatrix} 0 \\ \mathrm{D}(0)^{-1}\mathrm{H} \end{pmatrix}$$

$$\mathrm{C} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}$$

$$\mathrm{D} = 0$$

LQR **:** $J = \int_0^\infty x^T(t)Qx(t) + u^T(t)Ru(t)dt;$

El control que minimiza J, esta dado por la ley de realimentación :

$u(t) = -Kx(t)$ *con* $K = R^{-1}B^TP$ tal que $A^TP + PA - PBR^{-1}B^TP + Q = 0$

ESTABILIDAD, CONTROLABILIDAD Y OBSERVABILIDAD – KALMAN :

Como el **SPDI** es **LTI,** tenemos:

Teorema de Kalman - Controlabilidad **:** El sistema es controlable si y solo si la matriz de controlabilidad (**C***) es de rango completo.

Teorema de Kalman - Observabilidad **:** El sistema es controlable si y solo si la matriz de observabilidad (**O***) es de rango completo.

***Teorema de Estabilidad*:** El sistema es estable si y sólo si A es diagonalizable y para todo **λ ∈ Λ(A) : Re(λ) ≤ 0.**

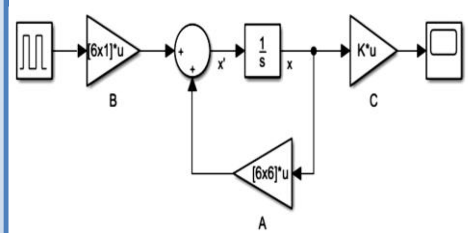
**** El número de condición de la matriz C* mide cuán sensible resulta el sistema a las perturbaciones.***

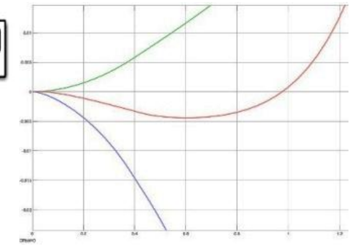
**** Para la simulación se emplea el método de Runge-Kutta.***

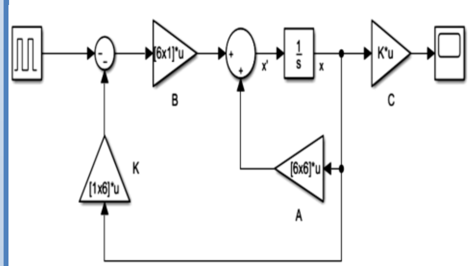
RL:

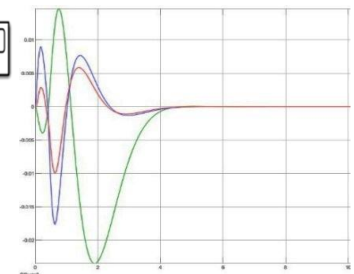
$$Q(s_t,a_t) \leftarrow \underbrace{Q(s_t,a_t)}_{\text{old value}} + \underbrace{\alpha}_{\text{learning rate}} \cdot \left(\overbrace{\underbrace{r_t}_{\text{reward}} + \underbrace{\gamma}_{\text{discount factor}} \cdot \underbrace{\max_a Q(s_{t+1},a)}_{\text{estimate of optimal future value}}}^{\text{learned value}} - \underbrace{Q(s_t,a_t)}_{\text{old value}} \right)$$

RESULTADOS Y DISCUSIONES

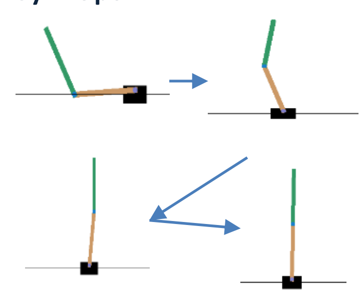








Gym-OpenAI-RL



La presencia de valores propios positivos de **A** índica que el sistema es inestable

C* y O* son de rango completo, el **SPDI** es Controlable y Observable, todos los estados son medibles.

k(C*) = 4,3028E10

CONCLUSIONES

* Se obtuvo el modelo no lineal del **SPDI** y se linealizo alrededor del pto de eq.
* El **SPDI** es inestable y, aunque es controlable presenta una dificultad inherente, debido al número de condición de la matriz de controlabilidad.
* Se observa que el **SPDI** - **LQR** para pequeñas perturbaciones se estabiliza cerca de los cuatro segundos, en aprox. el 70% de ensayos.
* Controlador **LQR** aprendido capaz de equilibrar más de 200 pasos de tiempo.
* ***El trabajo futuro estudiará algoritmos de Deep Learning y el controlador Gaussiano con redes neuronales.***

BIBLIOGRAFÍA

[1] Gustafsson Fredrik (2016) . **Control of an Inverted Double Pendulum using Reinforcement Learning**. CS 229 Final Project, Autumn

[2] Bogdanov Alexander(2004)**. Optimal Control of a Double Inverted Pendulum on a Cart**. Technical Report CSE-04- 006.

[3] Velandia German (n.d.). **Modelamiento, Diseño y Simulación del SPDI.** Universidad Autonoma de Colombia.

AGRADECIMIENTOS

A los profesores Roca e Hidalgo por su apoyo en la base teórica.