

Control y Simulación por Métodos Numéricos de un Sistema de Péndulo Doble Invertido(SPDI)

Felipe T. Walter [†], Bardales R. Jordi [†], Casaico S. Jhonatan [‡]

[†]*Proyecto para la acreditación ABET*

[‡]*Escuela de Matemática*

Facultad de Ciencias, Universidad Nacional de Ingeniería

E-mail: wfelipet@uni.pe, jbardaler@uni.pe, jcasaicos@uni.pe

Entregado el 05 de Octubre del 2019

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 de 3 ecuaciones diferenciales, luego se establece las condiciones para la linealización en espacio de estados [2], se analiza la estabilidad, controlabilidad y observabilidad del sistema continuo, lineal e invariante en el tiempo y se procede a la simulación numérica en Simulink-MATLAB, usando el método de Runge-Kutta 4.5 y empleando el controlador LQR, obteniéndose la convergencia a lazo cerrado [3]. Finalmente se simula gráficamente el sistema en el lenguaje de programación Python usando la biblioteca Gym de OpenAI, con aprendizaje por refuerzo (RL), formulado como un proceso de decisión de Markov(MDP), y para la obtener la solución numérica usamos el método de Runge-Kutta [1].

Palabras claves: Sistema de Péndulo Doble Invertido (SPDI), Ecuaciones de Euler Lagrange, controlador LQR, Simulink, MATLAB, Runge-Kutta, Python-gym, OpenAI, Aprendizaje por refuerzo (RL).

This report is a practical integration of the works [1] and [2], taking from [3] the parameters for the simulation exercise, except any type of friction. The mathematical basis is given as an added value in each step of the modeling, control, simulation and SPDI learning. For modeling, the Euler-Lagrange equations are used, arriving at a non-linear system of 3 differential equations, then establishing the conditions for linearization in the state space, analyze the stability, controllability and observability of the linear and invariant system over time and proceed to the numerical simulation in Simulink-MATLAB, using the Runge-Kutta 4.5 method and using the LQR controller, obtaining the closed-loop convergence. Finally, the system is graphically simulated in the Python programming language using the OpenAI Gym library, with reinforcement learning (RL), formulated as a Markov decision process (MDP), and for numerical integration we use the Runge method. Kutta [1].

Keywords: Double inverted pendulum system (SPDI), Euler Lagrange equations, LQR controller, Simulink, MATLAB, Runge-Kutta, Python-gym, OpenAI, Reinforcement learning (RL).

1. Introducción

Desde la década de los 70, la dinámica y control de péndulos invertidos ha llamado mucho la atención desde el punto de vista de la teoría de control y la inteligencia artificial. Existe una bibliografía extensa sobre los péndulos simples, dobles y triples, desde el swing-up(balaceo-arriba) y la estabilización[3].

El SPDI permite estudiar e involucrarnos en diversas dificultades asociadas a los problemas de control y aprendizaje de los sistemas dinámicos. 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 en un punto de operación de equilibrio y el efecto que esta produce en la solución de control óptima, así como la elección del controlador para la realimentación, cuya solución numérica no llegue a ser prohibitiva computacionalmente.

Por otro lado el SPDI pertenece a la clase de sistemas mecánicos sub-actuados que consiste en tres sistemas interconectados (dos péndulos y un carrito) con un sólo actuador: una fuerza usada para controlar tres grados de libertad.

Nuestro trabajo presenta los siguientes 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.

2. Ecuaciones de Euler-Lagrange

Primero es necesario definir un sistema de coordenadas generalizadas que satisfacen las ligaduras del sistema.

Un sistema de coordenadas generalizadas $\{q_1, q_2, \dots, q_n\}$ usualmente se dice que representa los N grados de libertad del sistema y se pueden representar como

$$\forall j \in \{1, 2, \dots, N\} : q_j = q_j(r_1, r_2, \dots, r_n, t)$$

Donde r_j es el vector posición de la j -ésima partícula del sistema (o la posición del C.M de cada cuerpo ligado).

Llamamos ligadura a una restricción del movimiento de una partícula o sistemas de partículas, por ejemplo una partícula libre no tiene de ligaduras, por otro lado si imponemos que un cuerpo (sistema continuo de partículas) debe realizar su movimiento sobre una superficie, o siguiendo una determinada trayectoria (helicoidal o circular, por ejemplo) estamos imponiendo ligaduras al sistema.

Para nuestros propósitos nos interesa conocer un tipo especial de ligaduras, las holonomas, para estas existe una función $h : \mathbb{R}^{3 \times n} \times \mathbb{R} \rightarrow \mathbb{R}$, tal que la siguiente ecuación se pueda expresar en función de la posición de las partículas y el tiempo

$$h(r_1, r_2, r_3, \dots, r_n, t) = 0$$

Y decimos que $\{q_1, q_2, \dots, q_n\}$ satisfacen las ligaduras cuando definiendo la dinámica del sistema, usando estas coordenadas generalizadas, presenta las mismas ligaduras preservando el comportamiento de estas, ya sean holonomas o no.

Partiendo del principio dinámico de D'Alembert

$$\sum_{i=1}^N N(\dot{p}_i - F_i) \delta r_i = 0$$

y definiendo el Lagrangiano \mathcal{L} , como una cantidad escalar tal que sea la diferencia de la energía cinética y potencial

$$\mathcal{L} = T - V$$

obtenemos las ecuaciones de Euler-Lagrange

$$\forall j \in \{1, \dots, n\} : \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j} \right) + \frac{\partial \mathcal{L}}{\partial q_j} = Q_j \quad (1)$$

Y haciendo $q = (q_1, q_2, \dots, q_n)^t$ y $Q = (Q_1, Q_2, \dots, Q_n)^t$, representamos las ecuaciones de Euler Lagrange en su forma vectorial

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \right) + \frac{\partial \mathcal{L}}{\partial q} = Q \quad (2)$$

Donde Q se llama vector de fuerzas (o de momentos) generalizadas actuando en dirección del vector de coordenadas generalizadas.

3. Teoría de Control

3.1. Control Óptimo

La teoría del control clásico comprende dos enfoques principales para los sistemas de tiempo continuo y de tiempo discreto lineal invariante en el tiempo (LTI). El primero, constituido por métodos de dominio de frecuencia, se basa en los conceptos de la función de transferencia y la respuesta de frecuencia. El segundo enfoque surge de la concepto de espacio de estado y utiliza ecuaciones diferenciales o diferenciales para describir sistemas dinámicos [4].

Es necesario entender los siguientes conceptos, para la buena lectura del presente informe.

- **Variable controlada y variable manipulada.** La variable *controlada* es la cantidad o condición que se mide y controla. La variable *manipulada* (*señal de control*) es la cantidad o condición que el controlador modifica para afectar el valor de la variable controlada. *Controlar* significa medir el valor de la variable controlada del sistema y aplicar la variable manipulada al sistema para corregir o limitar la desviación del valor medido respecto del valor deseado.
- **Sistemas.** Un sistema es una combinación de componentes que actúan juntos y realizan un objetivo determinado.
- **Perturbaciones.** Una perturbación es una señal que tiende a afectar negativamente el valor de la salida de un sistema. Si esta se genera dentro del sistema se denomina *interna*, mientras que una perturbación *externa* se genera fuera del sistema y es una entrada.[6]

A continuación presentamos algunos métodos para modelar un sistema dinámico

3.2. Función de transferencia

Cualquier sistema físico (mecánico, eléctrico, etc.) se puede traducir a una serie de valores matemáticos a través de los cuales se conoce el comportamiento de estos sistemas frente a valores concretos. Esto es lo que permite la función de transferencia. Se trata de un modelo matemático que a través de un cociente relaciona la respuesta de un sistema (salida) a una señal de entrada o excitación. Por definición una función de transferencia se puede determinar según la expresión:

$$H(s) = \frac{Y(s)}{U(s)} \quad (3)$$

Donde $H(s)$ es la función de transferencia. $Y(s)$ es la transformada de Laplace de la respuesta y $U(s)$ es la transformada de Laplace de la señal de entrada.[J.Beltrán Alonso. Simulación de un péndulo invertido]

3.3. Espacio de Estados

Para trabajar en espacio de estados definamos algunos puntos importantes

- **Estado.** El estado de un sistema dinámico es el conjunto de variables más pequeño (llamado *variables de estado*), de forma que el conocimiento de estas variables en $t = t_0$, junto con el conocimiento de la entrada para $t \geq t_0$, determinan completamente el comportamiento del sistema en cualquier $t \geq t_0$.
- **Variables de estado.** Son el menor conjunto de variables que determinan el estado del sistema dinámico.
- **Vector de estados.** Es un vector x donde sus componentes son las variables de estado.

- **Espacio de estados.** El espacio n -dimensional cuyos ejes de coordenadas están formados por las n variables de estado.[Ogata]

El espacio de estados es otro método que permite modelar un sistema dinámico. Se representa por un conjunto de entradas, salidas y variables de estado relacionadas por ecuaciones diferenciales de primer orden que se combinan en una ecuación diferencial matricial de primer orden. A esta representación se le llama ecuación de estado. Una forma general de expresar la dinámica de un sistema lineal es:

$$\begin{aligned}\dot{x} &= Ax + Bu \\ y &= Cx + Du\end{aligned}\quad (4)$$

Donde:

- **x:** estado del sistema (contiene n elementos para un sistema de orden n)
- **u:** vector de entrada (contiene m elementos)
- **y:** vector de salida (contiene p elementos)
- **A:** matriz de estado ($n \times n$)
- **B:** matriz de entrada ($n \times m$)
- **C:** matriz de salida ($p \times n$)
- **D:** matriz de transmisión directa ($p \times m$)

Este tipo de representación tiene la ventaja de que permite conocer el comportamiento interno del sistema, además de que se puede trabajar con sistemas cuyas condiciones iniciales sean diferentes de cero.[J.Beltrán Alonso. Simulación de un péndulo invertido]

3.4. Controlabilidad

Definición. Considera el sistema en tiempo continuo, lineal e invariante en el tiempo (LTI)

$$\dot{x} = Ax + Bu \quad (5)$$

donde

- $x \in R^n$: vector de estados.
- $u \in R^m$: vector de entradas.
- A : matriz de estados de orden $n \times n$.
- B : matriz de entradas de orden $n \times m$.

El par (A, B) es controlable si, dado una duración $T \geq 0$ y dos puntos arbitrarios $x_0, x_T \in R^n$, existe una función continua a trozos $t \rightarrow \bar{u}(t)$ de $[0, T]$ a R^m , tal que la integral de la curva $\bar{x}(t)$ generado por \bar{u} con $\bar{x}(0) = x_0$, satisface $\bar{x}(T) = x_T$. En otras palabras

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

En otras palabras se dice que un sistema es controlable en el tiempo t_0 si se puede transferir desde cualquier estado inicial $x(t_0)$ a cualquier otro estado, mediante un vector de control sin restricciones, en un intervalo de tiempo finito. Si todos los estados son controlables, se dice que el sistema es de estado completamente controlable.[6] Además, este concepto nos permite determinar si es posible controlar un sistema para modificar su comportamiento (estabilización de un sistema inestable, modificación de las dinámicas propias del sistema).

El siguiente teorema muestra que la propiedad de controlabilidad solo depende de A y B :

Teorema (*Criterio de Controlabilidad de Kalman*). Una condición necesaria y suficiente para que (A, B) sea controlable es

$$\text{rang } \mathcal{C} = \text{rang}[B|AB|\dots|A^{n-1}B] = n$$

donde \mathcal{C} es llamado *Matriz de controlabilidad de Kalman* (de orden $n \times nm$).

Y, mientras mayor sea **m**, la matriz de Kalman tiene más vectores columnas y así es más fácil encontrar **n** vectores columnas que sean linealmente independientes. Esto intuitivamente diría que mientras más controles haya, más posibilidades hay de controlar el sistema.

3.5. Observabilidad

Definición. Se dice que el sistema es observable si el estado $x(t_0)$ se determina a partir de la observación de $y(t)$ durante un intervalo de tiempo finito, $t_0 \leq t \leq t_1$. Por tanto, el sistema es completamente observable si todas las transiciones del estado afectan eventualmente a todos los elementos del vector de salida.

Teorema (*Criterio de Observabilidad de Kalman*). Un sistema **LTI** representado por el sistema (4) es observable si, y solo si, la matriz de observabilidad \mathcal{O} tiene rango n :

$$\text{rang}(\mathcal{O}) = \begin{bmatrix} C \\ \vdots \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} = n$$

La controlabilidad y la observabilidad de un sistema son aspectos duales de un mismo problema.[5] Además, la observabilidad gobierna la existencia de una solución de control óptimo. Y, según Kalman, las condiciones de controlabilidad y observabilidad determinan la existencia de una solución completa para un problema de diseño de un sistema de control; así la

solución a este problema puede no existir si el sistema considerado no es controlable. [6,pg.675-676]

La controlabilidad y la observabilidad de un sistema son aspectos duales de un mismo problema.[5] Además, la observabilidad gobierna la existencia de una solución de control óptimo. Y, según Kalman, las condiciones de controlabilidad y observabilidad determinan la existencia de una solución completa para un problema de diseño de un sistema de control; así la solución a este problema puede no existir si el sistema considerado no es controlable. [6,pg.675-676]

3.6. Estabilidad

Al diseñar un sistema de control, se debe ser capaz de predecir su comportamiento dinámico a partir del conocimiento de los componentes. La característica más importante del comportamiento dinámico de un sistema de control es la estabilidad absoluta, es decir, si el sistema es estable o inestable. Un sistema de control está en *equilibrio* si, en ausencia de cualquier perturbación o entrada, la salida permanece en el mismo estado.

Definición (Estabilidad). Un sistema de control lineal e invariante con el tiempo es *estable* si la salida termina por regresar a su estado de equilibrio cuando el sistema está sujeto a una condición inicial. Un sistema de control lineal e invariante con el tiempo es *críticamente estable* si las oscilaciones de la salida continúan de forma indefinida. Es *inestable* si la salida diverge sin límite a partir de su estado de equilibrio cuando el sistema está sujeto a una condición inicial [6, pg.160].

Teorema (Criterio de Estabilidad). Consideremos el sistema lineal invariante en el tiempo con control

$$\begin{aligned}\dot{x} &= Ax + Bu \\ y &= Cx + Du\end{aligned}\quad (6)$$

, luego el sistema (6) es estable si, y solo si, la matriz A es diagonalizable y $\forall \lambda \in \Lambda(A) : \operatorname{Re}(\lambda) \leq 0$.

3.7. Sistemas de control en lazo abierto

Aquellos en los que la variable de salida (variable controlada) no tiene efecto sobre variable de control.

Características:

- No se compara la salida del sistema con la entrada de referencia.
- Para cada entrada de referencia le corresponde una condición de operación fija.
- La precisión de la salida del sistema depende de la calibración del controlador.
- En presencia de perturbaciones estos sistemas de control no cumplen su función adecuadamente.

En la práctica, el control en lazo abierto solo se usa si se conoce la relación entre la entrada y la salida y si no hay perturbaciones internas ni externas. Estos sistemas no son de control realimentado.

3.8. Sistemas de control en lazo cerrado

Los sistemas de *control en lazo cerrado* se denominan también sistemas de *control realimentados* son aquellos en los que la señal de salida del sistema (variable controlada) tiene efecto directo sobre la variable de control; es decir, en un sistema de control en lazo cerrado, se alimenta al controlador la señal de error de actuación, que es la diferencia entre la señal de entrada y la señal de realimentación (que puede ser la propia señal de salida o una función de la señal de salida y sus derivadas y/o integrales), con el fin de reducir el error y llevar la salida a un valor deseado. Estos sistemas solo tienen ventajas cuando se presentan perturbaciones y/o variaciones impredecibles en los componentes del sistema, además el uso de la realimentación vuelve la respuesta del sistema relativamente insensible a las perturbaciones externas y a las variaciones internas en los parámetros del sistema.[Ogata]

3.9. Regulador Lineal Cuadrático(LQR)

Un LQR es un controlador de retroalimentación cuya configuración gobierna la dinámica del sistema utilizando un algoritmo matemático que minimiza un funcional de costo.

Con las matrices de estado (6), es posible realizar los cálculos para los controladores óptimos LQR; esta técnica de control parte del modelo dinámico de cada sistema, para obtener una matriz de realimentación que minimice el funcional de costo:

$$J = \int_0^{\infty} (x^T Q x + u^T R u) dt \quad (7)$$

donde Q es una matriz hermitiana o simétrica real definida positiva (o semidefinida positiva), R es una matriz hermitiana o simétrica real definida positiva y u no está restringida, además Q y R son matrices definidas heurísticamente de acuerdo a una ponderación realizada por el diseñador para las variables de estado del sistema en ambos lazos; estas matrices buscan minimizar a partir de la ecuación (7) la energía empleada por las variables de estado (Q) y por su entrada (R); dando como resultado una realimentación óptima según la ley de control ($u = -Kx$), en donde la ganancia, al tener un comportamiento continuo en el tiempo, se vuelve una función, que se obtiene a partir de la derivación e igualación a cero de la ecuación (7), dando como resultado la función de la ecuación:

$$K(t) = -R^{-1} B^T P(t) \quad (8)$$

Donde en este caso, P es la solución de la ecuación de Riccati mostrada en la ecuación:

$$A^T P + PA - PBR^{-1}B^T P + Q = 0 \quad (9)$$

Además el LQR produce una solución óptima cuando la dinámica del sistema es lineal.

3.10. Métodos numéricos

Para la resolución de sistemas de ecuaciones diferenciales ordinarias (EDO's) el entorno de Matlab emplea la familia de métodos numéricos Runge-Kutta-Fehlberg, como Runge-Kutta de primer orden (Euler), Runge-Kutta de cuarto orden, Dormand-Prince, entre otros; los cuales son implementados por los diferentes solucionadores que posee, tales como ode45, ode12, ode113, entre otros. Sin embargo, el método principal que emplea Matlab es Dormand-Prince, el cual es implementado por el solucionador estándar ode45, este se basa en una fórmula explícita de Runge-Kutta (4,5), el par Dormand-Prince (i.e. que el método tiene orden 5 y el método incrustado tiene orden 4) [7], [8], [9], [10].

4. Modelamiento

Representamos gráficamente el sistema con la siguiente figura.

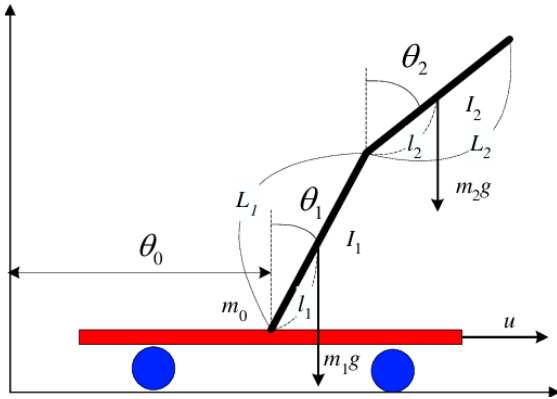


Figura 1: variables y constantes del SPDI

En la figura se intuye la denominación de algunas constantes, como la masa (m_0) del carro, ángulos, longitudes, entre otras. Cabe mencionar que también se debe considerar los momentos de inercia de cada péndulo I_1 e I_2 , para el cálculo posterior de la energía cinética rotacional.

Comenzamos el análisis definiendo el siguiente sistema de coordenadas generalizadas $\{q_1 = \theta_0, q_2 = \theta_1, q_3 = \theta_2\}$ que satisfacen las ligaduras y representan los 3 grados de libertad del SPDI.

θ_0 : Traslación horizontal del actuador(carro)

θ_1 : Desviación vertical del péndulo 1

θ_2 : Desviación vertical del péndulo 2

Usamos las ecuaciones de Euler-Lagrange en su forma vectorial para obtener el sistema de ecuaciones diferenciales que modelan el movimiento del SPDI.

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}} \right) + \frac{\partial \mathcal{L}}{\partial \theta} = \mathcal{Q} \quad (10)$$

donde $\theta = (\theta_0, \theta_1, \theta_2)^t$ y $\mathcal{Q} = (u, 0, 0)^t$

Cálculamos el Lagrangiano $\mathcal{L} = T - V$, hallando la energía cinética y potencial del SPDI.

$$T = T_0 + T_1 + T_2$$

$$V = V_0 + V_1 + V_2$$

Donde

$$T_0 = \frac{1}{2} m_0 \dot{\theta}_0^2$$

$$T_1 = \frac{1}{2} m_1 \left[(\dot{\theta}_0 + l_1 \dot{\theta}_1 \cos \theta_1)^2 + (l_1 \dot{\theta}_1 \sin \theta_1)^2 \right] + \frac{1}{2} I_1 \dot{\theta}_1^2$$

$$T_2 = \frac{1}{2} m_2 (\dot{\theta}_0 + L_1 \dot{\theta}_1 \cos \theta_1 + l_2 \dot{\theta}_2 \cos \theta_2)^2 + \frac{1}{2} m_2 (L_1 \dot{\theta}_1 \sin \theta_1 + l_2 \dot{\theta}_2 \sin \theta_2)^2 + \frac{1}{2} I_2 \dot{\theta}_2^2$$

$$V_0 = 0$$

$$V_1 = m_1 g l_1 \cos \theta_1$$

$$V_2 = m_2 g (L_1 \cos \theta_1 + l_2 \cos \theta_2)$$

Derivando el Lagrangiano $\mathcal{L} = T_0 + T_1 + T_2 - (V_0 + V_1 + V_2)$ por $\dot{\theta}$ y θ y reemplazando en 10 tenemos el siguiente sistema no lineal de tres ecuaciones diferenciales de segundo orden

$$\begin{aligned} u &= (m_0 + m_1 + m_2) \ddot{\theta}_0 + (m_1 l_1 + m_2 L_1) \cos(\theta_1) \ddot{\theta}_1 \\ &+ m_2 l_2 \cos(\theta_2) \ddot{\theta}_2 - (m_1 l_1 + m_2 L_1) \sin(\theta_1) \dot{\theta}_1^2 \\ &- m_2 l_2 \sin(\theta_2) \dot{\theta}_2^2 \\ 0 &= (m_1 l_1 + m_2 L_1) \cos(\theta_1) \ddot{\theta}_0 + (m_1 l_1^2 + m_2 L_1^2 + I_1) \ddot{\theta}_1 \\ &+ m_2 L_1 l_2 \cos(\theta_1 - \theta_2) \ddot{\theta}_2 + m_2 L_1 l_2 \sin(\theta_1 - \theta_2) \dot{\theta}_2^2 \\ &- (m_1 l_1 + m_2 L_1) g \sin(\theta_1) \\ 0 &= m_2 l_2 \cos(\theta_2) \ddot{\theta}_0 + m_2 L_1 l_2 \cos(\theta_1 - \theta_2) \ddot{\theta}_1 \\ &+ (m_2 l_2^2 + I_2) \ddot{\theta}_2 - m_2 L_1 l_2 \sin(\theta_1 - \theta_2) \dot{\theta}_1^2 \\ &- m_2 l_2 g \sin(\theta_2) \end{aligned}$$

En forma matricial compacta:

$$\mathbf{D}(\theta) \ddot{\theta} + \mathbf{C}(\theta, \dot{\theta}) \dot{\theta} + \mathbf{G}(\theta) = \mathbf{H}u \quad (11)$$

Donde

$$\mathbf{D}(\theta) = \begin{pmatrix} \mathbf{D}_1(\theta) & \mathbf{D}_2(\theta) & \mathbf{D}_3(\theta) \end{pmatrix}$$

$$\mathbf{C}(\theta, \dot{\theta}) = \begin{pmatrix} \mathbf{0} & \mathbf{C}_2(\theta, \dot{\theta}) & \mathbf{C}_3(\theta, \dot{\theta}) \end{pmatrix}$$

$$\mathbf{G}(\theta) = \begin{pmatrix} 0 \\ -(\frac{1}{2} m_1 + m_2) L_1 g \sin(\theta_1) \\ -\frac{1}{2} m_2 L_2 g \sin(\theta_2) \end{pmatrix}$$

$$\mathbf{H} = (1 \ 0 \ 0)^t$$

Detallando

$$\mathbf{D}_1(\theta) = \begin{pmatrix} m_0 + m_1 + m_2 \\ (\frac{1}{2}m_1 + m_2)L_1 \cos(\theta_1) \\ \frac{1}{2}m_2 L_2 \cos(\theta_2) \end{pmatrix}$$

$$\mathbf{D}_2(\theta) = \begin{pmatrix} (\frac{1}{2}m_1 + m_2)L_1 \cos(\theta_1) \\ (\frac{1}{3}m_1 + m_2)L_1^2 \\ \frac{1}{2}m_2 L_2 \cos(\theta_2) \end{pmatrix}$$

$$\mathbf{D}_3(\theta) = \begin{pmatrix} \frac{1}{2}m_2 L_2 \cos(\theta_2) \\ \frac{1}{2}m_2 L_1 L_2 \cos(\theta_1 - \theta_2) \\ \frac{1}{3}m_2 L_2^2 \end{pmatrix}$$

$$\mathbf{C}_2(\theta, \dot{\theta}) = \begin{pmatrix} -(\frac{1}{2}m_1 + m_2)L_1 \sin(\theta_1)\dot{\theta}_1 \\ 0 \\ \frac{1}{2}m_2 L_1 L_2 \sin(\theta_1 - \theta_2)\dot{\theta}_2 \end{pmatrix}$$

$$\mathbf{C}_3(\theta, \dot{\theta}) = \begin{pmatrix} -\frac{1}{2}m_2 L_2 \sin(\theta_2)\dot{\theta}_2 \\ \frac{1}{2}m_2 L_1 L_2 \sin(\theta_1 - \theta_2)\dot{\theta}_2 \\ 0 \end{pmatrix}$$

Notar que la matriz $\mathbf{D}(\theta)$ es simétrica e invertible.

Para el diseño de la ley de control la ecuación 11 debe ser reformulada en seis ecuaciones no lineales de primer orden. Esto se hace definiendo un vector $\mathbf{x} \in \mathbb{R}^6$ como

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

Luego se tiene

$$\dot{\mathbf{x}} = \begin{pmatrix} 0 & \mathbf{I} \\ 0 & -\mathbf{D}^{-1}\mathbf{C} \end{pmatrix} \mathbf{x} + \begin{pmatrix} 0 \\ -\mathbf{D}^{-1}\mathbf{G} \end{pmatrix} + \begin{pmatrix} 0 \\ \mathbf{D}^{-1}\mathbf{H} \end{pmatrix} u \quad (12)$$

Representamos esta ecuación por una ecuación vectorial, con una función de dos variables vectoriales y una de control $f : \mathbb{R}^7 \rightarrow \mathbb{R}$, que en sistemas continuos es de la forma

$$f(\dot{\mathbf{x}}, \mathbf{x}, u) = 0 \quad (13)$$

4.1. Linealización

Procedemos a la linealización del sistema, este proceso brinda información acerca del comportamiento dinámico de los sistemas no lineales, para pequeñas perturbaciones de la condición de operación.

Por lo que el sistema se comporta linealmente si lo operamos en un punto apenas perturbado cerca del punto de equilibrio \mathbf{x}_{eq} ($u \approx 0$) tal que satisfaga la ecuación 13.

Truncamos la expansión de Taylor al primero orden en el punto de equilibrio, es decir consideramos pequeñas perturbaciones, se tiene en cuenta lo siguiente para cada grado de libertad angular.

$$\cos(\theta_j) \approx 1 \quad y \quad \sin(\theta_j) \approx 0$$

Con estas aproximaciones en 12 se obtiene el modelo lineal deseado, que en ecuaciones de espacio de estados, característico en sistemas modernos de control, se representa por

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t) \\ \mathbf{y} &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}u(t)\end{aligned}$$

Donde

$$\begin{aligned}\mathbf{A} &= \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{D}(0)^{-1}\partial_{\theta}\mathbf{G}(0) & \mathbf{0} \end{pmatrix} \\ \mathbf{B} &= \begin{pmatrix} \mathbf{0} \\ \mathbf{D}(0)^{-1}\mathbf{H} \end{pmatrix} \\ \mathbf{C} &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{pmatrix} \\ \mathbf{D} &= \mathbf{0}\end{aligned}$$

5. Estabilidad, Controlabilidad y Observabilidad

En este avance, para la simulación se ha usado los parametros de [3] .

$$\begin{aligned}m_0 &= 1,0Kg, \quad m_1 = 0,4Kg, \quad m_2 = 0,2Kg, \quad L_1 = 0,2m, \\ L_2 &= 0,4m, \quad g = 9,8m/s^2\end{aligned}$$

Para las matrices que resultan en las ecuaciones de espacio de estados, del modelo lineal, se obtienen los siguientes valores.

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & -4,3377 & 0,4820 & 0 & 0,9836 & -0,2459 \\ 0 & 96,3934 & -43,3770 & 0 & -13,5246 & 22,1311 \\ 0 & -115,6721 & 110,8525 & 0 & 1,2295 & -56,5574 \end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0,9016 \\ -6,1475 \\ 7,3770 \end{pmatrix}$$

Para analizar la estabilidad y controlabilidad del sistema, hallamos primero los valores propios de la matriz \mathbf{A} .

$$[3,43 \quad 1,80 \quad 0 \quad 0 \quad -15,3 \quad -60,0]$$

La presencia de dos valores propios positivos de la matriz \mathbf{A} indica que el sistema es inestable. La medida de controlabilidad la podemos conocer a partir del número de condición de la matriz de controlabilidad, la cual es

$$\begin{bmatrix} 0 & 9.0e^{-1} & -7.8 & 3.8e^{+2} & -2.1e^{+4} & 1.2e^{+6} \\ 0 & -6.1 & 2.5e^{+2} & -1.4e^{+4} & 7.9e^{+5} & -4.7e^{+7} \\ 0 & 7.4 & -4.2e^{+2} & 2.6e^{+4} & -1.6e^{+6} & 9.3e^{+7} \\ 9.0e^{-1} & -7.8 & 3.8e^{+2} & -2.1e^{+4} & 1.2e^{+6} & -7.4e^{+7} \\ -6.1 & 2.5e^{+2} & -1.4e^{+4} & 7.9e^{+5} & -4.7e^{+7} & 2.8e^{+9} \\ 7.4 & -4.2e^{+2} & 2.6e^{+4} & -1.5e^{+6} & 9.3e^{+7} & -5.6e^{+9} \end{bmatrix}$$

Figura 2: Matriz de Controlabilidad \mathbf{C}^*

el cual se obtiene por la razón entre el mayor y el menor valor singular. El número de condición de controlabilidad obtenido es **4,3028E10**. Este número es una indicación de la dificultad inherente a la estabilización del sistema, sin embargo como la matriz de controlabilidad y observabilidad son de rango completo, el SPDI es totalmente controlable y observable.

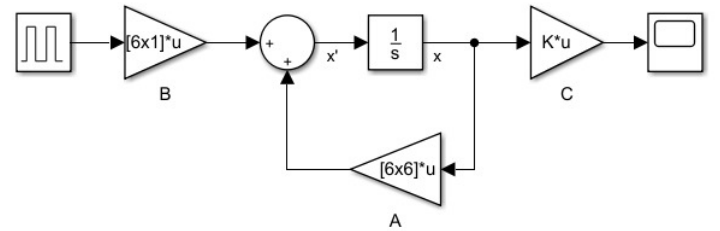
6. Simulación y Control

La simulación se hizo en Matlab-Simulink con diagrama de bloques, usando las ecuaciones en espacio de estados. Se ha empleado un pulso de dos segundos de duración como señal de entrada al sistema.

6.1. Control en lazo abierto

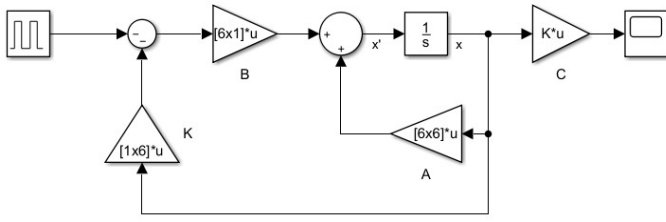
6.1.1. Lineal

Figura 3: Modelo del control a lazo cerrado - Simulink



6.2. Control en lazo cerrado

6.2.1. Lineal



Se presentan las gráficas de las variables del SPDI linealizado en lazo cerrado y abierto respectivamente

Figura 4: Modelo del control a lazo abierto - Simulink

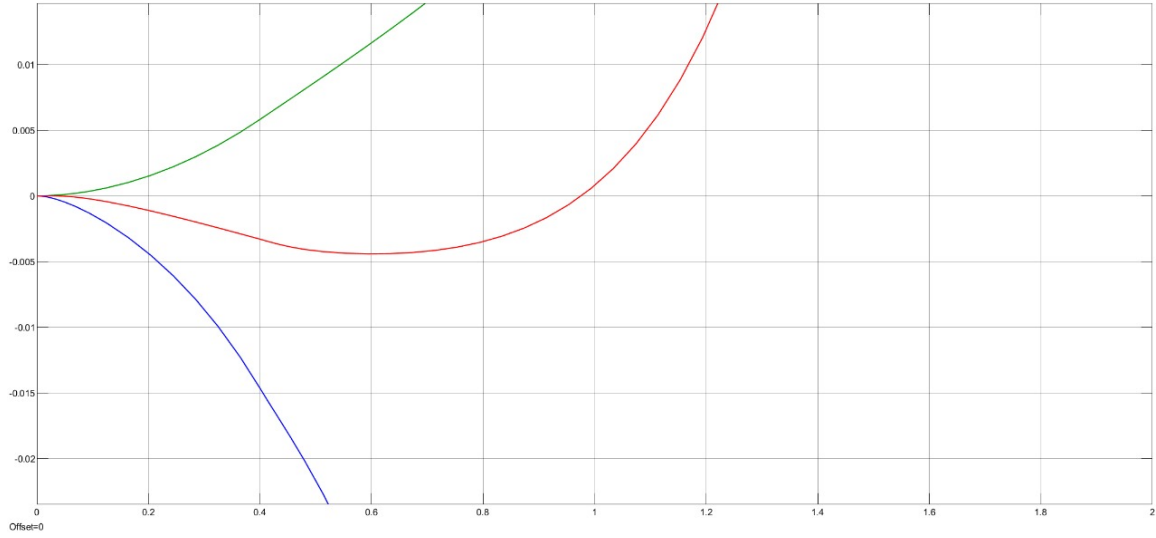


Figura 5: Modelo del control a lazo abierto - Simulink

Se observa que el sistema es completamente inestable, dado que a medida que avanza el tiempo las trayectorias generalizadas de los dos péndulos y el carro divergen mutuamente. Por el contrario en lazo cerrado realimentamos el sistema con una matriz de ganancias \mathbf{K} tal que $\mathbf{u}(t) = \mathbf{K}\mathbf{y}(t)$

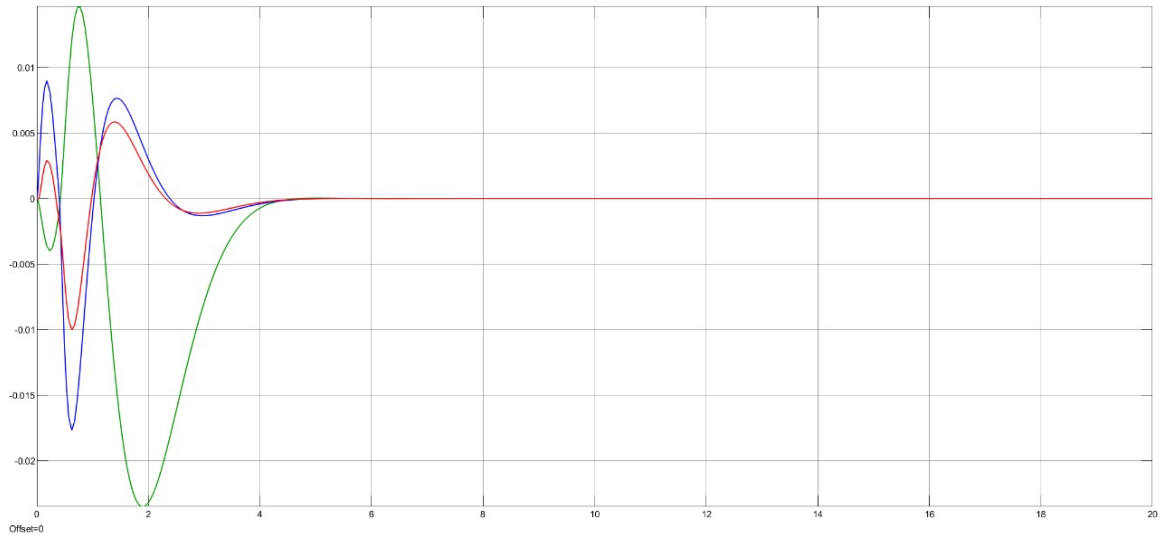


Figura 6: Modelo del control a lazo abierto - Simulink

Se observa que el SPDI para pequeñas perturbaciones las trayectorias generalizadas convergen, es decir comienza a estabilizarse pasado los cuatro segundos.

N. Conclusiones

- Se obtuvo el modelo no lineal del SPDI y se linealizó 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 por ciento 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.

Agradecimientos

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

-
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 Germán (n.d.). **Modelamiento, Diseño y Simulación del SPDI**. Universidad Autónoma de Colombia.
 4. Efim N. Rosenwasser and Bernhard P. Lampe Multi-variable Computer-controlled Systems A Transfer Function Approach
 5. Depto. de Ingenieria de Sistemas y Automática APUNTES DE INGENIERIA DE CONTROL ANÁLISIS Y CONTROL DE SISTEMAS EN ESPACIO DE ESTADO IDENTIFICACIÓN DE SISTEMAS CONTROL ADAPTATIVO CONTROL PREDICTIVO Daniel Rodríguez Ramírez Carlos Bordóns Alba Rev. 5/05/2005.pg.19
 6. K. Ogata, Ingeniería de Control Moderna (Pearson Educación, S.A., Madrid, 2010)
 7. Dormand, J. R. and P. J. Prince, "A family of embedded Runge-Kutta formulae," J. Comp. Appl. Math., Vol. 6, 1980, pp. 19–26.
 8. Shampine, L. F. and M. W. Reichelt, "The MATLAB ODE Suite," SIAM Journal on Scientific Computing, Vol. 18, 1997, pp. 1–22.
 9. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.619.551rep=rep1type=pdf>
 10. Dormand, John R. (1996), Métodos numéricos para ecuaciones diferenciales: un enfoque computacional , Boca Raton: CRC Press , pp. 82-84, ISBN 0-8493-9433-3.