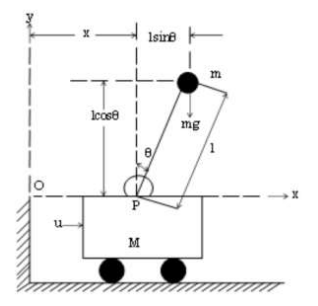
|  |
| --- |
| Explicación |

## Dinámica del Sistema

El modelo matemático de un péndulo invertido está gobernado por las ecuaciones diferenciales que describen el equilibrio de fuerzas. Estas ecuaciones se derivan del diagrama de cuerpo libre

Figura: Diagrama del Péndulo Invertido Sobre un Carro



Nota: El carro intenta mantener el péndulo recto

### Variables y Parámetros

### Ecuación de Movimientos

Estas ecuaciones describen el movimiento combinado del péndulo y el carrito. La fuerza actúa como un control que debe estabilizar el sistema.

## Características del Problema

**Estados:**

El estado del sistema puede representarse como un vector de 4 variables:

**Acciones:**

El espacio de acciones es discreto y representa la fuerza aplicada al carrito:

* : Fuerza aplicada hacia la izquierda, ninguna fuerza, o hacia la derecha.

**Espacio de Observación:**

* Observamos los estados , que se miden continuamente.
* Para algoritmos de aprendizaje, este espacio puede ser discretizado en rangos para convertirlo en un problema manejable.

**Estados Iniciales:**

* Generalmente, el carrito comienza en el centro
* El péndulo comienza inclinado .

**Estados Terminales:**

* El episodio termina cuando:
  + El péndulo cae
  + El carrito se sale de los límites

## Discretización del Espacio

Discretizar el espacio continuo permite usar algoritmos como *Q-Learning*, que son más efectivos en dominios discretos. Por ejemplo:

* Discretizar en intervalos:
  + , dividido en 10 pasos.
  + , dividido en 12 pasos.

Esto reduce el espacio de observación a un número finito de estados.

## Aprendizaje por Refuerzo

1. Q-Learning

El aprendizaje por refuerzo se puede emplear para aprender una política óptima que estabilice el péndulo.

* + Función Q: Representa el valor esperado de tomar una acción en un estado y seguir la política :
  + Recompensas:
    - asas

1. Filtro de Kalman

El filtro de Kalman puede utilizarse para estimar estados en tiempo real cuando las mediciones están contaminadas por ruido. No es directamente un método de control, pero puede mejorar el aprendizaje o el control al proporcionar estimaciones más precisas de los estados.

1. Deep Q-Learning

Cuando el espacio de estados es continuo o muy grande, se puede usar una red neuronal para aproximar

## Ejemplo de Simulación

Episodios:

* Cada episodio simula un intento de estabilizar el péndulo, terminando en un estado terminal.
* Por ejemplo, episodios de entrenamiento.

Pseudo-Código

1. Inicializa o la red neuronal.
2. Para cada episodio:
   1. Inicializar el estado .
   2. Realizar acciones usando una política .
   3. Actualizar o entrenar la red con experiencias.

Criterio de Éxito

* Reducir y a casi 0, manteniendo el péndulo estable.

## Matemática

Parámetros del sistema:

Datos:

* La varilla del péndulo no tiene masa.
* No hay fricción en la bisagra del péndulo ni en el carrito.

Fuerzas Relevantes:

* Fuerza gravitacional:
* Fuerza de inercia:
* Fuerza centrípeta:
* Fuerza aplicada:

Diagrama de Cuerpo Libre

Carrito:

La fuerza neta que actúa sobre el carrito en la dirección se describe como:

Péndulo:

El momento neto sobre el péndulo respecto a su pivote (debido a gravedad y aceleración angular) es:

Ecuaciones de movimiento:

1. Ecuación para el carrito :

Usando la segunda ley de Newton para el carrito, obtenemos:

Reordenando:

… (1)

1. Ecuación para el péndulo :

Usando la ecuación de momentos sobre el pivote del péndulo:

Reorganizando:

… (2)

Sustitución y Combinación

Sustituimos de la ecuación (1) en (2) para eliminar

1. Sustituimos :
2. Reorganizamos :
3. Simplificamos:

Linealización:

Para simplificar las ecuaciones (usadas en controladores lineales o Q-Learning), se hace la suposición de que:

* , para triángulos pequeños
* Los términos cuadráticos son despreciados.

Esto resulta en: