

# Contenido del curso

## Contents

- Información del académico
- Prerequisitos
- Software requerido y herramientas
- Temario
- Para presentar examen
- Código de conducta
- Libros de apoyo

## Información del académico

- **Profesor:** Dr. Jesús Emmanuel Solís Pérez
- **Correo:** jsolisp@unam.mx
- [Página personal del profesor](#)

## Prerequisitos

1. Dinámica de Sistemas Físicos

## Software requerido y herramientas

1. Python & Jupyter Notebook.
2. Sistema operativo: Windows 10, Linux, o macOS.
3. Navegador: Google Chrome, Opera o Firefox.
4. Markdown para escribir documentación.
5. Sistema de control de versiones.

# Temario

Unidad	Tema	Contenido	Material
I	Introducción al análisis de la dinámica de sistemas	<ul style="list-style-type: none"><li>Ejemplos prácticos.</li></ul>	1. <a href="#">Código (Live Script)</a>
II	Antecedentes matemáticos	<ul style="list-style-type: none"><li>Ecuaciones diferenciales no lineales.</li><li>Análisis de la existencia, unicidad y continuidad de soluciones.</li></ul>	1. <a href="#">Código (Live Script)</a>
III	Fundamentos del análisis de estabilidad en el sentido de Lyapunov	<ul style="list-style-type: none"><li>Localización de los puntos de equilibrio.</li><li>Estabilidad de los puntos de equilibrio.</li><li>Linealización de sistemas en torno a un punto de equilibrio.</li><li>Análisis de la estabilidad de sistemas dinámicos linealizados.</li><li>Análisis de estabilidad por el método directo de Lyapunov.</li></ul>	1. <a href="#">Código (Live Script)</a>
IV	Controlabilidad de sistemas dinámicos	<ul style="list-style-type: none"><li>Concepto de controlabilidad.</li><li>Condiciones para la controlabilidad.</li><li>Formas canónicas de controlabilidad.</li></ul>	1. <a href="#">Código (Live Script)</a>
V	Observabilidad de sistemas dinámicos	<ul style="list-style-type: none"><li>Concepto de observabilidad.</li><li>Teorema de dualidad.</li></ul>	1. <a href="#">Código (Live Script)</a>

<b>Unidad</b>	<b>Tema</b>	<b>Contenido</b>	<b>Material</b>
		<ul style="list-style-type: none"> <li>• Fromas irreductibles de un sistema dinámico.</li> </ul>	
VI	Identificación de sistemas dinámicos	<ul style="list-style-type: none"> <li>• Objetivo del problema de identificación.</li> <li>• Representación de sistemas dinámicos en forma lineal con respecto a los parámetros.</li> <li>• Métodos de estimación de parámetros.</li> <li>• Diseño de experimentos para la identificación paramétrica.</li> </ul>	1. <a href="#"><u>Código (Live Script)</u></a>

## Para presentar examen

Tendrán derecho a presentar exámenes parciales aquellos estudiantes que hayan cubierto por lo menos el 80% de asistencia.

## Código de conducta

La conducta del profesorado y alumnado del curso será acorde con los principios y valores especificados en el Código de Ética de la Universidad Nacional Autónoma de México aprobado el 1 de julio del 2015 por el Consejo Universitario, en especial en lo referente a la integridad y honestidad académica. “La integridad y la honestidad académica implican: Citar las fuentes de ideas, textos, imágenes, gráficos u obras artísticas que se empleen en el trabajo universitario, y no sustraer o tomar la información generada por otros o por sí mismo sin señalar la cita correspondiente u obtener su consentimiento y acuerdo. No falsificar, alterar, manipular, fabricar, inventar o fingir la autenticidad de datos, resultados, imágenes o información en los trabajos académicos, proyectos de investigación, exámenes, ensayos, informes, reportes, tesis, audiencias, procedimientos de orden disciplinario o en cualquier documento inherente a la vida académica universitaria” (Gaceta UNAM, 30 de julio 2015).

# Libros de apoyo

1. Nise, N. S. (2020). Control systems engineering. John Wiley & Sons.
2. Khalil, H. K. (2002). Nonlinear systems third edition. Patience Hall, 115.
3. Slotine, J. J. E., & Li, W. (1991). Applied nonlinear control (Vol. 199, No. 1, p. 705). Englewood Cliffs, NJ: Prentice hall.
4. Johansson, R. (1993). System modeling and identification. Prentice-hall.

(Última modificación: 10 de julio de 2024)

# Introducción al análisis de la dinámica de sistemas

Print to PDF

## Contents

- Introducción al análisis de la dinámica de sistemas
- Índices de error

## Ejemplos prácticos

### Péndulo simple

Considere el péndulo simple mostrado en la siguiente figura

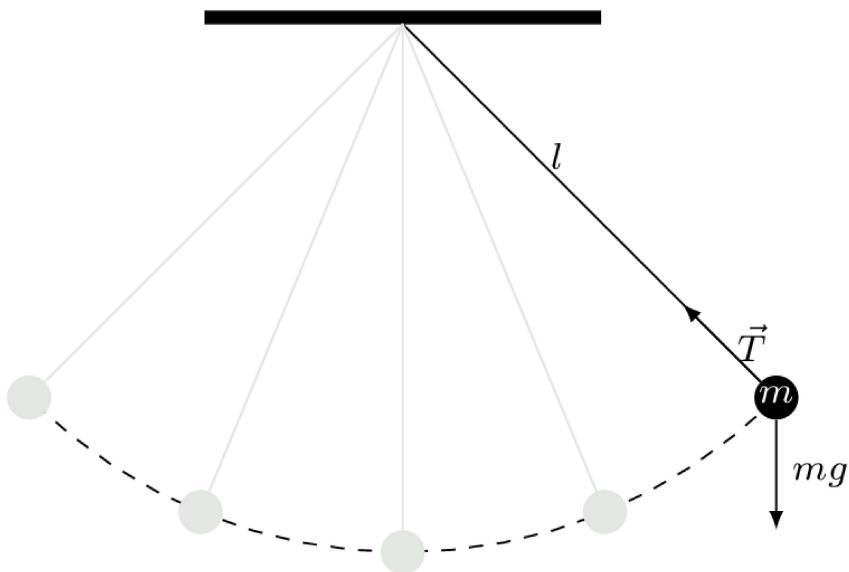


Fig. 1 Diagrama de péndulo simple sin fuerza externa.

donde  $l$  denota la longitud de la barra, y  $m$  la masa de la bola. Asuma que la barra es rígida y de masa despreciable. Suponga entonces,  $\theta$  como el ángulo entre la barra y el eje vertical a través

del pivote. Además, el péndulo se mueve libremente sobre un plano vertical formando un círculo de radio  $l$ . Para obtener la ecuación que describe el movimiento del péndulo es necesario identificar las fuerzas que interactúan sobre la bola o bien, utilizar el enfoque de Euler-Lagrange.

## Modelado a partir de leyes físicas

En un sistema mecánico de tipo rotacional, las leyes físicas que rigen el sistema están dadas por la Segunda ley de Newton, donde las fuerzas positivas están dadas hacia la derecha mientras que las negativas hacia el lado contrario. Si realizamos un diagrama de cuerpo libre para observar las fuerzas que interactúan con la masa como se muestra a continuación

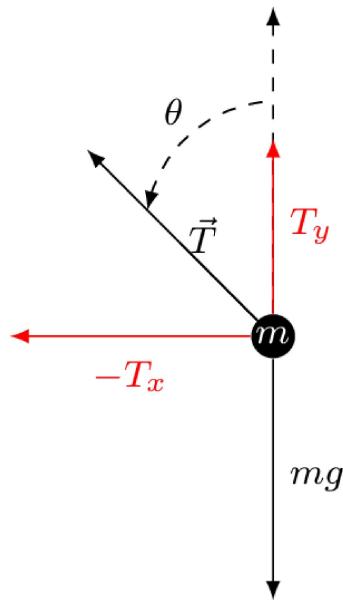


Fig. 2 Diagrama de cuerpo libre de la masa en el péndulo simple.

Podemos observar que hay una fuerza gravitacional  $mg$ , donde  $g$  es la aceleración de la gravedad. Así mismo, existe una fuerza de tensión  $\vec{T}$  con componentes  $T_x$  y  $T_y$ .

Podemos observar que la única fuerza que interactúa en  $x$  es  $T_x$ . Recordando que esta tensión tiene una fuerza negativa, por lo tanto la definimos como sigue

$$-T \sin(\theta) = m\ddot{x}, \quad (1)$$

donde  $T$  denota la magnitud de la tensión y  $\ddot{x} := \frac{d^2x}{dt^2}$ .

Para el análisis de las fuerzas que interactúan en el eje  $y$ , tenemos

$$-mg + T \cos(\theta) = m\ddot{y}. \quad (2)$$

Observamos entonces que tenemos dos variables que ecuaciones. Por ello, vamos a realizar el siguiente despeje

$$T = -\frac{m\ddot{x}}{\sin(\theta)}. \quad (3)$$

Sustituyendo (3) en (2), tenemos

$$-mg + \left(-\frac{m\ddot{x}}{\sin(\theta)}\right) \cos(\theta) = m\ddot{y}. \quad (4)$$

Simplificamos la expresión obtenida multiplicando ambos lados de la igualdad por el término  $\sin(\theta)$  y resulta

$$-mg \sin(\theta) - m\ddot{x} \cos(\theta) = m\ddot{y} \sin(\theta). \quad (5)$$

Observando el diagrama de la Fig. 1, tenemos que la posición de la masa está dada por sus componentes  $x$  e  $y$  dentro de un plano cartesiano. Así mismo, tenemos que  $l$  y  $\theta$  están en un sistema de coordenadas polares. Por lo tanto, podemos utilizar la relación que hay entre estos sistemas para establecer las siguientes relaciones

$$x = l \sin(\theta), \quad y = -l \cos(\theta). \quad (6)$$

Puesto que necesitamos la aceleración a lo largo del eje  $x$  e  $y$ , derivamos la expresión dada en la (6) recordando la regla de la cadena

$$\begin{aligned} \dot{x} &= l \cos(\theta) \dot{\theta}, \\ \dot{y} &= -l \sin(\theta) \dot{\theta}, \end{aligned} \quad (7)$$

$$\begin{aligned} \ddot{x} &= l \left[ \ddot{\theta} \cos(\theta) - \dot{\theta}^2 \sin(\theta) \right], \\ \ddot{y} &= -l \left[ \ddot{\theta} \sin(\theta) + \dot{\theta}^2 \cos(\theta) \right]. \end{aligned} \quad (8)$$

Sustituyendo (7) y (8) en (5)

$$\begin{aligned} -mg \sin(\theta) - ml\ddot{\theta} \cos^2(\theta) + ml\dot{\theta}^2 \sin(\theta) \cos(\theta) &= ml\ddot{\theta} \sin^2(\theta) + ml\dot{\theta}^2 \cos(\theta) \sin(\theta), \\ -mg \sin(\theta) - ml\ddot{\theta} \cos^2(\theta) &= ml\ddot{\theta} \sin^2(\theta), \\ -mg \sin(\theta) &= ml\ddot{\theta} [\sin^2(\theta) + \cos^2(\theta)], \end{aligned} \quad (9)$$

$$ml\ddot{\theta} = -mg \sin(\theta). \quad (10)$$

Suponiendo que hay una fuerza proporcional a la velocidad de la masa con un coeficiente  $k_f$  que resiste el movimiento, el modelo de *péndulo simple con fuerza de fricción* queda dado por la siguiente ecuación

$$ml\ddot{\theta} = -mg \sin(\theta) - k_f l\dot{\theta}. \quad (11)$$

Por lo tanto, la dinámica del péndulo simple con y sin fricción están dadas por ecuaciones diferenciales de segundo orden. Para resolver las Ecs. (10) y (11) realizamos un cambio de variable y definimos como variables de estado  $x_1 = \theta$  y  $x_2 = \dot{\theta}$ . Por consiguiente, el sistema queda representado como sigue

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -\frac{g}{l} \sin(x_1) - \frac{k_f}{m} x_2. \end{aligned} \quad (12)$$

Una variante del sistema de péndulo simple se puede obtener agregando una entrada exógena  $\tau$

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -\frac{g}{l} \sin(x_1) - \frac{k_f}{m} x_2 + \frac{\tau}{ml^2}. \end{aligned}$$

## Modelado a partir del enfoque Euler-Lagrange

Cualquier sistema debe satisfacer las ecuaciones de Lagrange dadas por

$$\frac{d}{dt} \left[ \frac{\partial \mathcal{L}}{\partial \dot{q}} \right] - \frac{\partial \mathcal{L}}{\partial q} + \frac{\partial J}{\partial \dot{q}} = 0, \quad (13)$$

donde  $\mathcal{L}$  es el Lagrangiano del sistema y se define como sigue

$$\mathcal{L}(q, \dot{q}) = U^*(\dot{q}) - T(q).$$

Además

- $U$ , representa la coenergía total en las reservas de flujo del sistema expresada como una función de las coordenadas de esfuerzo generalizado.
- $T$ , la energía total en los almacenes de esfuerzo del sistema expresada como una función de las coordenadas de acumulación de esfuerzo generalizadas.
- $J$ , el co-contenido total en los disipadores del sistema expresado como una función de las coordenadas de esfuerzo generalizado.
- $q$ , el vector de coordenadas generalizadas definida como el conjunto de coordenadas linealmente independientes que definen la configuración del sistema.

Para un sistema mecánico no conservativo, las ecuaciones de Lagrange están dadas como sigue

$$\frac{d}{dt} \left[ \frac{\partial \mathcal{L}}{\partial \dot{q}_j} \right] - \frac{\partial \mathcal{L}}{\partial q_j} + \frac{\partial J}{\partial \dot{q}_j} = \tau_j, \quad j = 1, \dots, l. \quad (14)$$

donde  $\tau_j$  denota las fuerzas generalizadas.

### 💡 Hint

En este contexto conservativo significa libre de dissipación y entradas de fuerzas externas.

El sistema mostrado en la [Fig. 1](#), sólo posee un grado de libertad denotado como  $\theta$ . Entonces, definimos las siguientes coordenadas generalizadas

$$q := \theta, \quad \dot{q} := \dot{\theta}.$$

Además, definimos como fuerza generalizada la variable  $u$ , i.e.

$$\tau_i = 0.$$

Ahora bien, podemos representar la Ec. [\(6\)](#) en forma vectorial como sigue

$$\mathbf{p}_m = \begin{bmatrix} x_m \\ y_m \end{bmatrix} = \begin{bmatrix} l \sin(\theta) \\ -l \cos(\theta) \end{bmatrix}, \quad (15)$$

y el vector de velocidad [\(7\)](#) como sigue

$$\mathbf{v}_m = \begin{bmatrix} \dot{x}_m \\ \dot{y}_m \end{bmatrix} = \begin{bmatrix} l \cos(\theta)\dot{\theta} \\ l \sin(\theta)\dot{\theta} \end{bmatrix}. \quad (16)$$

La energía co-cinética del sistema está dada por

$$\begin{aligned} U^* &= \frac{1}{2}mv^2 \equiv \frac{1}{2}m\|\mathbf{v}_m\|^2 \\ &= \frac{1}{2}m\mathbf{v}_m^T \mathbf{v}_m. \end{aligned} \quad (17)$$

mientras que la energía potencial del sistema por

$$T = mgh \equiv m\vec{g}^T \mathbf{p}_m, \quad (18)$$

donde  $\vec{g} = [0 \ g]^T$ , con  $g \in \mathbb{R}^2$ .

Sustituyendo [\(16\)](#) en [\(17\)](#), tenemos

$$\begin{aligned} U^* &= \frac{1}{2}m \begin{bmatrix} l \cos(\theta)\dot{\theta} & l \sin(\theta)\dot{\theta} \end{bmatrix} \begin{bmatrix} l \cos(\theta)\dot{\theta} \\ l \sin(\theta)\dot{\theta} \end{bmatrix} \\ &= \frac{1}{2}m \left( l^2 \cos^2(\theta)\dot{\theta}^2 + l^2 \sin^2(\theta)\dot{\theta}^2 \right), \\ &= \frac{1}{2}ml^2\dot{\theta}^2 (\cos^2(\theta) + \sin^2(\theta)), \\ &= \frac{1}{2}ml^2\dot{\theta}^2. \end{aligned}$$

Ahora, sustituyendo [\(15\)](#) en [\(18\)](#)

$$\begin{aligned} T &= m \begin{bmatrix} 0 & g \end{bmatrix} \begin{bmatrix} l \sin(\theta) \\ -l \cos(\theta) \end{bmatrix}, \\ &= -mgl \cos(\theta) \end{aligned}$$

Entonces, el Lagrangiano del sistema se define como sigue

$$\mathcal{L} = \frac{1}{2} ml^2 \dot{\theta}^2 + mgl \cos(\theta).$$

Obtenemos las derivadas parciales

$$\frac{\partial \mathcal{L}}{\partial \dot{\theta}} = ml^2 \dot{\theta}, \quad (19)$$

$$\frac{d}{dt} \left[ \frac{\partial \mathcal{L}}{\partial \dot{\theta}} \right] = ml^2 \ddot{\theta}, \quad (20)$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \theta} &= \frac{\partial}{\partial \theta} \left( \frac{1}{2} ml^2 \dot{\theta}^2 + mgl \cos(\theta) \right), \\ &= -mgl \sin(\theta), \end{aligned} \quad (21)$$

Sustituyendo, (19)-(21) en (14) y considerando que  $\tau = 0$  tenemos

$$ml^2 \ddot{\theta} + mgl \sin(\theta) = 0.$$

## Diodo tunel

El circuito de diodo tunel se muestra en la siguiente figura

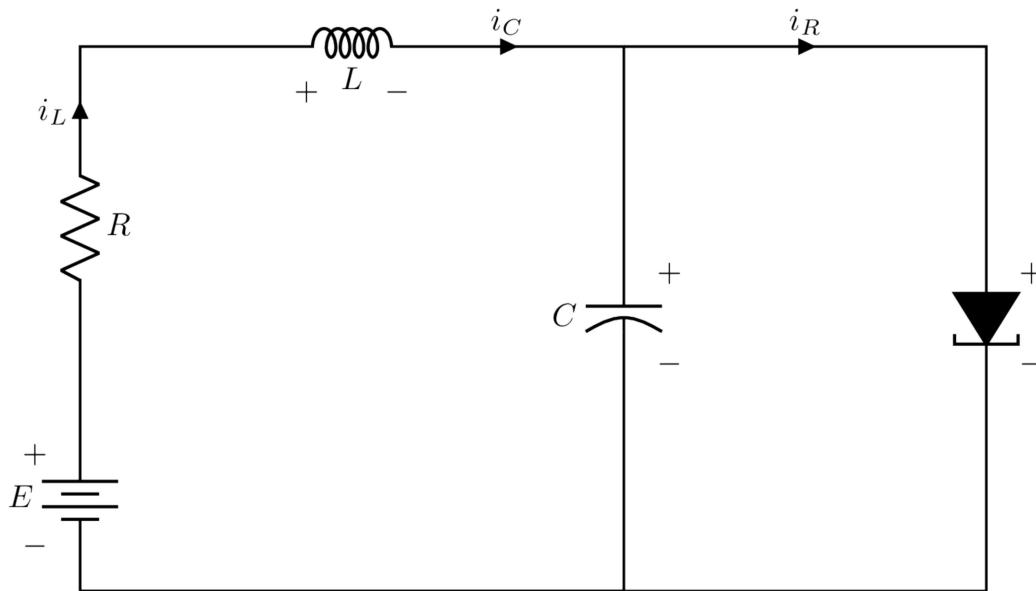


Fig. 3 Diodo túnel.

donde la relación constitutiva que caracteriza el tunel diodo está dada por  $i_R = h(v_R)$ . Los elementos almacenadores de energía son el inductor  $L$  y el capacitor  $C$ . Asumiendo que estos elementos son lineales e invariantes en el tiempo, podemos llegar a representar este sistema a partir del siguiente modelo

$$\begin{aligned} i_C &= C\dot{v}_C, \\ v_L &= L\frac{di_L}{dt}, \end{aligned} \tag{22}$$

donde  $v$  e  $i$  denotan el voltaje y corriente a través de un elemento. Además, el sub índice especifica el elemento.

Las ecuaciones del circuito de diodo tunel se pueden representar en forma de espacio-estados si consideramos  $u = E$  como una entrada constante,  $x_1 = v_C$  (tensión en el capacitor) y  $x_2 = i_L$  (corriente en el inductor).

Utilizando las *Leyes de Kirchhoff* y expresando  $i_C$  como una función de las variables de estado  $x_1$ ,  $x_2$  y la entrada  $u$ , tenemos lo siguiente

$$i_C + i_R - i_L = 0,$$

por consiguiente

$$i_C = -h(x_1) + x_2.$$

Del mismo modo, expresamos  $v_L$  como una función de las variables  $x_1$ ,  $x_2$ ,  $u$  y utilizamos las *Leyes de Kirchhoff* como sigue

$$v_C - E + Ri_L + v_L = 0.$$

donde  $v_L = -x_1 - Rx_2 + u$ .

Reescribiendo las ecuaciones del sistema (22), tenemos

$$\begin{aligned} \dot{x}_1 &= \frac{1}{C}(-h(x_1) + x_2), \\ \dot{x}_2 &= \frac{1}{L}(-x_1 - Rx_2 + u), \end{aligned} \tag{23}$$

# Círculo R-C

La relación que establece el flujo electromagnético  $\phi$  y la corriente  $i$  que lo produce está dada por la siguiente ecuación

$$\phi = Li,$$

donde  $L$  es una constante que depende de los factores geométricos y de entorno llamada inductancia.

Los cambios de flujo electromagnético originan potenciales eléctricos relacionados por la *Ley de Faraday*

$$u_L = -\dot{\phi},$$

donde  $u_L$  denota el voltaje en las terminales de la inductancia a razón del cambio de flujo. Por consiguiente, la *Ley de Faraday* se puede expresar como sigue

$$u_L = -L\dot{\phi}.$$

En elementos resistivos el voltaje  $u_R$  entre el componente y la corriente  $i$  que circula por él obedecen a la *Ley de Ohm* dada como siguiente

$$u_R = Ri, \tag{24}$$

donde  $R$  es una constante que depende del componente denominado resistencia.

El voltaje  $u_C$  entre las terminales de una capacitancia y la carga  $q$  siguen la siguiente relación

$$u_C = \frac{Q}{C} \equiv \frac{1}{C} \int idt,$$

donde  $C$  es una constante que depende de la geometría y el entorno denominada capacitancia. Si consideramos que la corriente se define como una variación temporal de carga

$$i = \dot{Q}, \quad Q = \int idt,$$

entonces  $u_C$  se expresa en los siguientes términos

$$u_C = \frac{1}{C} \int i dt. \quad (25)$$

Considere el circuito mostrado en la [Fig. 4](#)

Aplicando la *Ley de tensiones de Kirchhoff*, obtenemos

$$u_R + u_C = V_{in}. \quad (26)$$

Sustituyendo [\(24\)](#) y [\(25\)](#) en [\(26\)](#), tenemos

$$Ri + \frac{1}{C} \int i dt = V_{in},$$

expresado la ecuación anterior en términos de la carga  $Q$ , tenemos la siguiente expresión

$$R\dot{Q} + \frac{1}{C}Q = V_{in}, \quad (27)$$

Aplicando la *Ley de corrientes de Kirchhoff*, obtenemos

$$i_R + i_C = i,$$

dado que el voltaje entre los componentes eléctricos es el mismo y lo denotamos por  $u$ , tenemos

$$\frac{u}{R} + C\dot{u} = i.$$

Considerando el modelo dado en la Ec. [\(27\)](#) y tomando  $V$  como la carga en el capacitor dividida por la capacitancia  $V := \frac{Q}{C}$  y  $\dot{V} := \frac{\dot{Q}}{C}$ , sustituyendo tenemos

$$\begin{aligned} R\dot{V}C + \frac{1}{C}VC &= V_{in}, \\ R\dot{V}C + V &= V_{in}. \end{aligned}$$

Entonces, el modelo del circuito RC mostrado en la [Fig. 4](#) está dado por la siguiente ecuación

$$\dot{V} + \frac{1}{RC}V = \frac{1}{RC}V_{in}.$$

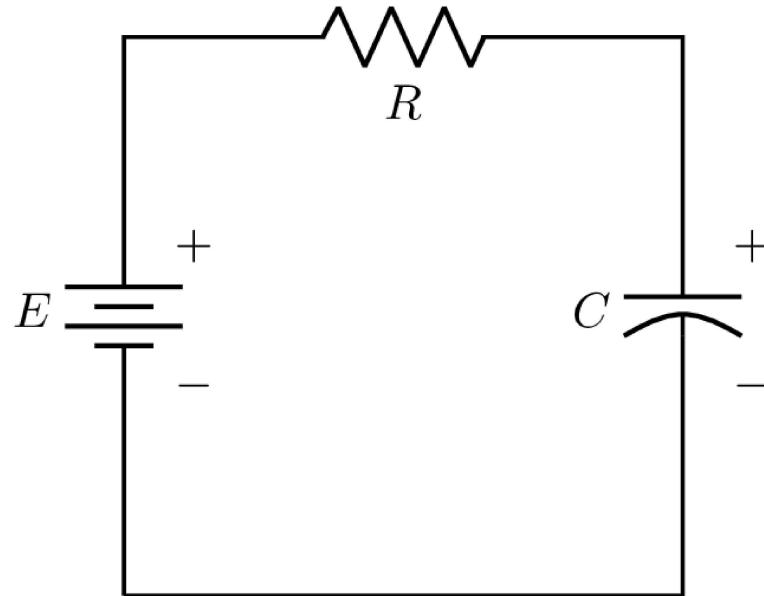


Fig. 4 Modelo de circuito R-C.

## Sistema masa-resorte

Considere una masa  $m$  que se desliza sobre una superficie horizontal y que se encuentra sujetada a una superficie vertical a través de un resorte con constante de fuerza  $k$  como se muestra en la siguiente figura.

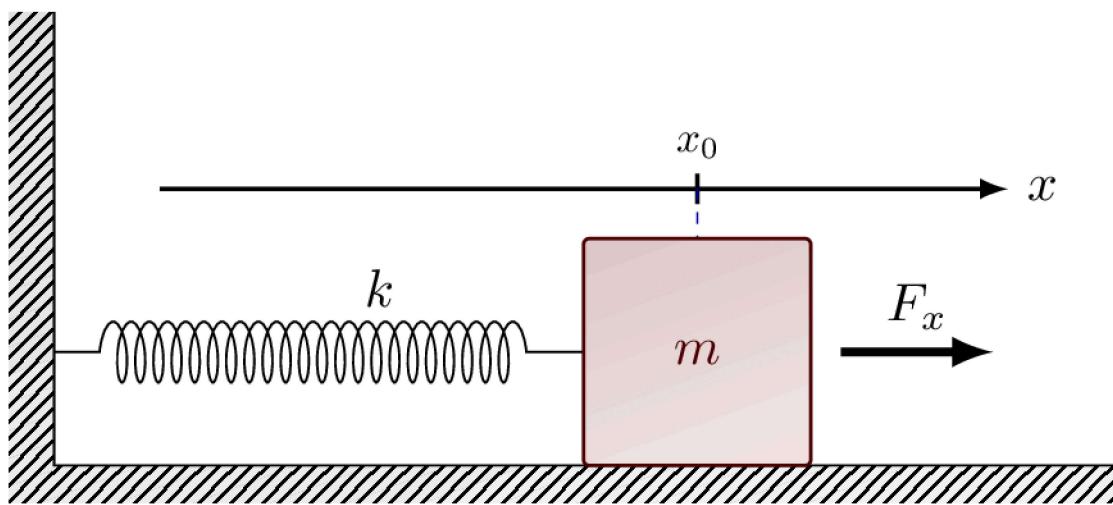


Fig. 5 Diagrama del sistema masa-resorte.

Si definimos  $x$  como el desplazamiento desde el punto de referencia  $x_0$ ; asumimos que la masa se desplaza sin fricción y aplicamos la Segunda ley de Newton así como la Ley de Hooke

obtenemos

$$ma + kx = F_x. \quad (28)$$

Dado que la aceleración  $a$  está definida como  $\ddot{x}$  y que la fuerza  $F_x$  está a nuestra disposición, sustituimos en la ecuación anterior para obtener el modelo del sistema masa-resorte como sigue

$$m\ddot{x} + kx = 0.$$

Para un desplazamiento largo, la fuerza restauradora ( $kx$ ) podría depender de manera no lineal de  $x$ , por ejemplo, expresada por la siguiente función

$$g(x) = k(1 - a^2x^2)x, \quad |ax| < 1,$$

representa al modelo llamado softening spring, donde un gran incremento de desplazamiento produce un incremento pequeño de fuerza.

Por otro lado, si tenemos la siguiente función

$$g(x) = k(1 + a^2y^2)x,$$

es posible obtener el modelo llamado hardening spring donde un incremento pequeño de desplazamiento produce un gran incremento de fuerza.

## Modelado a partir del enfoque Euler-Lagrange

En este caso, vamos a tratar de obtener el modelo dinámico del sistema masa-resorte utilizando las ecuaciones de Euler-Lagrange. Para ello, definimos la siguiente coordenada generalizada

$$q = x, \quad \dot{q} = \dot{x}.$$

Expresamos las energías en función de la coordenada generalizada para obtener el Lagrangiano del sistema. Donde, la energía potencial del sistema se obtiene a partir de la siguiente relación

$$A = \frac{\text{base} \cdot \text{altura}}{2}, \quad (29)$$

Sustituyendo la Ley de Hooke en (29), tenemos

$$T = \frac{qkq}{2} \equiv \frac{kq^2}{2},$$

mientras que la co-energía cinética está dada como sigue

$$U^* = \frac{m\dot{q}^2}{2}.$$

Por consiguiente, el Lagrangiano del sistema se define como sigue

$$\mathcal{L}(q, \dot{q}) = \frac{m\dot{q}^2}{2} - \frac{kq^2}{2}. \quad (30)$$

Calculando las ecuaciones de Lagrange (14) a partir de la ecuación (30), obtenemos

$$\frac{\partial \mathcal{L}}{\partial \dot{q}} = m\ddot{q}$$

$$\frac{d}{dt} \left[ \frac{\partial \mathcal{L}}{\partial \dot{q}} \right] = m\ddot{q} \quad (31)$$

$$\frac{\partial \mathcal{L}}{\partial q} = -kq \quad (32)$$

Finalmente, sustituimos las Ecs. (31), (32) en la Ec. (14)

$$m\ddot{x} + kx = F_x.$$

## Sistema masa-resorte-amortiguador

Considere el sistema masa-resorte-amortiguador mostrado en la siguiente Figura.

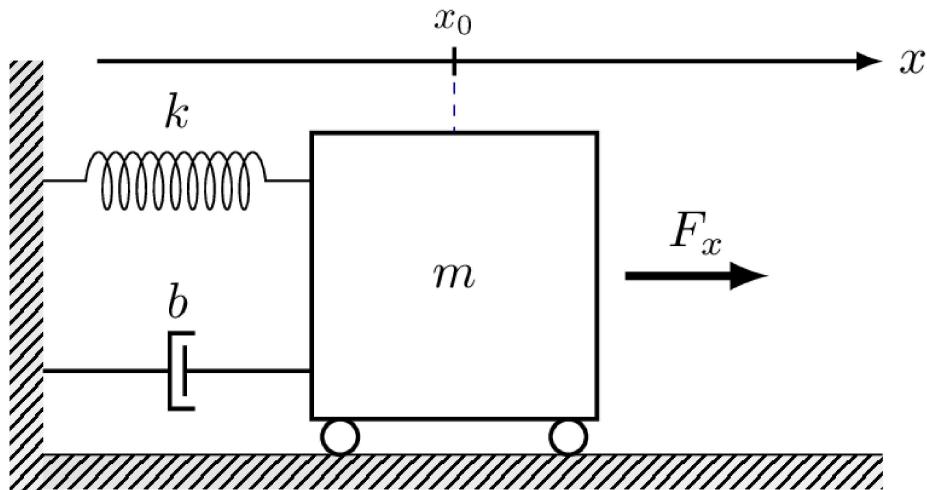


Fig. 6 Sistema masa-resorte-amortiguador.

Dicho sistema consiste en una masa  $m$  sujeta a un elemento de amortiguamiento con constante de viscosidad  $b$  y un resorte con rigidez  $k$ . Podemos observar que el sistema se desplaza sobre un plano horizontal a partir de una fuerza  $F_x$  aplicada.

Podemos partir del modelo dado en la Ec. [\(28\)](#) y modelando el amortiguador con la relación dada a continuación

$$F_k = b\dot{x},$$

donde  $b$  denota el coeficiente de fricción viscosa, tenemos

$$m\ddot{x} + b\dot{x} + kx = F_x.$$

De manera análoga, podemos obtener aplicar las ecuaciones de Lagrange [\(14\)](#) considerando que existe una fuerza de fricción viscosa como sigue

$$m\ddot{x} + kx = F_x - b\dot{x}.$$

## Criterios integrales

### Integral del error absoluto (IAE)

$$\text{IAE} = \int_0^{\infty} |e(t)| dt,$$

donde

$$e(t) = y(t) - \hat{y}(t).$$

- Fácil aplicación
- No se pueden optimizar sistemas altamente sub ni altamente sobre amortiguados
- Difícil de evaluar analíticamente

```
def IAE(y,yg,dt):
    return np.trapz(np.abs(y-yg))*dt
```

## Integral del tiempo por el error absoluto (ITAE)

- Los errores tardíos son más castigados
- Buena selectividad
- Difícil de evaluar analíticamente

```
def ITAE(y,yg,t,dt):
    return np.trapz(t*np.abs(y-yg))*dt
```

## Integral del error cuadrático (ISE)

$$\text{ISE} = \int_0^{\infty} e^2(t) dt.$$

- Da mayor importancia a los errores grandes
- No es un criterio muy selectivo
- Respuesta rápida pero oscilatoria, estabilidad pobre

```
def ISE(y,yg,dt):
    return np.trapz(np.abs(y-yg)**2)*dt
```

# Integral del tiempo por el error cuadrático (ITSE)

$$\text{ITSE} = \int_0^{\infty} t e^2(t) dt.$$

- Los grandes errores iniciales tienen poco peso pero los que se producen más tarde son fuertemente penados
- Mejor selectividad con respecto al ISE

```
def ITSE(y,yg,t,dt):
    return np.trapz(t*(y-yg)**2)*dt
```

## Criterios estadísticos

### Mean Square Error

$$\text{MSE} = \frac{1}{N} \sum_{k=0}^N e_k^2.$$

- No recomendable para estudiar modelos de predicción
- No tiene escala original el error porque está elevado al cuadrado
- No se mide en unidades de los datos experimentales

```
def MSE(y,yg):
    e = y - yg
    return np.mean(e**2)
```

### Root Mean Square Error

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{k=0}^N e_k^2}.$$

- Sensible a valores atípicos

- No se ajusta a la demanda (¿qué es demanda?)
- Se mide en unidades de los datos experimentales

```
def RMSE(y,yg):
    return np.sqrt(MSE(y,yg))
```

## Mean Absolute Error

$$\text{MAE} = \frac{1}{N} \sum_{k=0}^N |e_k|.$$

- Mide la precisión de los datos simulados
- Se mide en unidades de los datos experimentales
- No es sensible a valores atípicos
- Utilizado para analizar series temporales

```
def MAE(y,yg):
    return np.mean(np.abs(y-yg))
```

## Mean Absolute Percentage Error

$$\text{MAPE} = \frac{100\%}{N} \sum_{k=0}^N \frac{|e_k|}{y_k}$$

- Mide el error en porcentajes
- Indicador de desempeño
- Fácil interpretación
- Ampliamente utilizado para evaluar modelos de predicción

### Tabla de MAPE

- Si  $\text{MAPE} < 10$ , entonces el modelo es altamente preciso
- Si  $10 < \text{MAPE} < 20$ , entonces el modelo es bueno
- Si  $20 < \text{MAPE} < 50$ , entonces el modelo es razonable

- Si  $MAPE > 50$ , entonces el modelo es impreciso

## FIT

Obtiene el porcentaje de variación de salida que es explicado por un modelo

$$FIT = 100 \left( 1 - \frac{\|y - \hat{y}\|}{\|y - \bar{y}\|} \right)$$


---

## Métodos numéricos

### Método de Euler

Sea  $\phi(x)$  la solución exacta de la ecuación diferencial

$$\dot{y}(x) = f(x, y),$$

con condición iniciales

$$y(x_0) = y_0,$$

donde  $\phi(x)$  satisface la relación

$$\dot{\phi}(x) = f(x, \phi(x)), \quad \phi(x_0) = y_0.$$

La solución de una ecuación diferencial vía numérica es una solución aproximada del valor de la solución  $\phi(x)$  en un conjunto finito de puntos. Es decir,  $\phi(x_n) : y_n \approx \phi(x_n)$ .

Es común elegir los puntos  $x_n$  de forma equiespaciada, esto es  $h = x_{n+1} - x_n$ . En este caso,  $x_n = x_0 + nh$  donde  $h$  es el tamaño del paso.

Integramos la relación dada en la ecuación anterior entre  $x_0$  y  $x_1$

$$\int_{x_0}^{x_1} f(x, \phi(x)) dx = \int_{x_0}^{x_1} \dot{\phi}(x) dx = \phi(x_1) - \phi(x_0),$$

o bien

$$\phi(x_1) = \phi(x_0) + \int_{x_0}^{x_1} f(x, \phi(x)) dx.$$

Recordando que  $\phi(x_0) = y_0$ , entonces

$$\phi(x_1) = y_0 + \int_{x_0}^{x_1} f(x, \phi(x)) dx,$$

podemos hallar el valor de  $\phi(x_1)$  evaluando la integral anterior.

El método de Euler estima esta integral mediante la regla del rectángulo

$$\int_{x_0}^{x_1} f(x, \phi(x)) dx \approx f(x_0, y_0)(x_1 - x_0).$$

Es decir, aproxima el área que hay bajo la curva  $f(x, \phi(x))$  entre  $x_1$  y  $x_0$  por el área del rectángulo de ancho  $(x_1 - x_0)$  con altura igual a la ordenada de la curva en su extremo izquierdo  $f(x_0, y_0)$

$$\phi(x_1) \approx \phi(x_0) + f(x_0, \phi(x_0))h.$$

Para estimar el valor de  $\phi(x)$  en el siguiente punto  $x_2$ , integramos entre  $x_1$  y  $x_2$ , i.e.

$$\int_{x_1}^{x_2} \dot{\phi}(x) dx = \phi(x_2) - \phi(x_1) = \int_{x_1}^{x_2} f(x, \phi(x)) dx.$$

Aproximando la integral mediante la regla del rectángulo se tiene

$$\phi(x_2) \approx \phi(x_1) + f(x_1, \phi(x_1))(x_2 - x_1).$$

Dado que  $\phi(x_1)$  es desconocido, entonces lo aproximamos por  $y_1$  como sigue

$$\begin{aligned}\phi(x_2) &\approx \phi(x_1) + f(x_1, \phi(x_1))(x_2 - x_1), \\ \phi(x_2) &\approx y_1 + f(x_1, y_1)(x_2 - x_1),\end{aligned}$$

o bien

$$\phi(x_2) \approx y_1 + f(x_1, y_1)h.$$

Estimando  $\phi(x_2)$  por  $y_2$ , entonces cualquier estimación  $y_{n+1}$  de  $\phi(x_{n+1})$  puede hacerse por el método de Euler de acuerdo con la siguiente expresión

$$\phi(x_{n+1}) \approx y_{n+1} = y_n + f(x_n, y_n)h$$

## Método de Runge-Kutta

Sirve para buscar aproximaciones a la solución en puntos intermedios del intervalo  $[x_n, x_{n+1}]$  con una combinación lineal de los valores de la derivada en varias aproximaciones para obtener un valor de  $y_{n+1}$ .

Sea la ecuación diferencial

$$\dot{y}(x) = f(x, y), \quad y(y_0) = y_0,$$

y sea

$$x_n = x_0 + nh, \quad h > 0.$$

Para evaluar  $y(x_{n+1})$  conociendo el valor  $y_n$  y además  $0 \leq \alpha_1 \leq \alpha_2 \cdots \leq \alpha_r \leq 1$ , de modo que  $\sum_{n=1}^r \gamma_n = 1$ , el método de Runge-Kutta evalúa  $y_{n+1}$  como sigue

$$y_{n+1} = y_n + \sum_{n=1}^r \gamma_n k_n,$$

donde

$$k_n = h f \left( x_n + \alpha_n h, y_n + \sum_{j=1}^r \beta_{n,j} k_j \right), \quad \sum_{j=1}^r \beta_{n,j} = \alpha_n.$$

Los métodos de Runge-Kutta se clasifican en:

- **Explícitos.** Cuando los valores de  $k_n$  pueden ser evaluados en función de  $k_1, k_2, \dots, k_{n-1}$ .
- **Implícitos.** Cuando lo anterior no es posible.

Además, en los métodos explícitos se satisface la siguiente restricción

$$\beta_{n,j} = 0, \quad \forall n \leq j.$$

Mientras que para los implícitos se resuelve en cada paso un sistema de ecuaciones de la forma

$$\begin{aligned} k_1 &= f(x_n + \alpha_1 h, y_n + \beta_{1,1} k_1 + \beta_{1,2} k_2 + \dots + \beta_{1,p} k_p), \\ k_2 &= f(x_n + \alpha_2 h, y_n + \beta_{2,1} k_1 + \beta_{2,2} k_2 + \dots + \beta_{2,p} k_p), \\ &\vdots = \vdots \\ k_p &= f(x_n + \alpha_p h, y_n + \beta_{p,1} k_1 + \beta_{p,2} k_2 + \dots + \beta_{p,p} k_p). \end{aligned}$$

El método de Runge-Kutta de 4to orden es el más utilizado y evalúa la función  $f(x, y)$  en los puntos  $x_n, x_n + \frac{h}{2}$  y  $x_n + h$  de la forma

$$y_{n+1} = y_n + \alpha_1 k_1 + \alpha_2 k_2 + \alpha_3 k_3 + \alpha_4 k_4,$$

con

$$\begin{aligned} k_1 &= f(x_n, y_n)h, \\ k_2 &= f(x_n + a_2 h, y_n + b_{21} k_1)h, \\ k_3 &= f(x_n + a_3 h, y_n + b_{31} k_1 + b_{32} k_2)h, \\ k_4 &= f(x_n + a_4 h, y_n + b_{41} k_1 + b_{42} k_2 + b_{43} k_3)h. \end{aligned}$$

Eligiendo arbitrariamente  $\alpha_2 = \alpha_3 = \frac{1}{3}$  y hallando el resto de los coeficientes mediante un sistema algebráico de 11 ecuaciones, tenemos

$$\begin{aligned}\alpha_1 &= \frac{1}{6}, \\ \alpha_2 &= \frac{1}{3}, \quad \alpha_2 = \frac{1}{2}, \quad b_{21} = \frac{1}{2}, \\ \alpha_3 &= \frac{1}{3}, \quad \alpha_3 = \frac{1}{2}, \quad b_{31} = 0, \quad b_{32} = \frac{1}{2}, \\ \alpha_4 &= \frac{1}{6}, \quad \alpha_4 = 1, \quad b_{41} = 0, \quad b_{42} = 0, \quad b_{43} = 1.\end{aligned}$$

Tendríamos entonces

$$y_{n+1} = y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4),$$

con

$$\begin{aligned}k_1 &= f(x_n, y_n)h, \\ k_2 &= f\left(x_n + \frac{h}{2}, y_n + \frac{k_1}{2}\right)h, \\ k_3 &= f\left(x_n + \frac{h}{2}, y_n + \frac{k_2}{2}\right)h, \\ k_4 &= f(x_n + h, y_n + k_3)h.\end{aligned}$$

# Antecedentes matemáticos

Print to PDF

## Contents

- Ecuaciones diferenciales no lineales
- Análisis de la existencia, unicidad y continuidad de soluciones
- Existencia y unicidad

## Ecuaciones diferenciales no lineales

Un sistema dinámico puede ser modelado por un número finito de ecuaciones diferenciales ordinarias de primer orden acopladas como sigue

$$\begin{aligned}\dot{x}_1 &= f_1(t, x_1, \dots, x_n, u_1, \dots, u_p), \\ \dot{x}_2 &= f_2(t, x_1, \dots, x_n, u_1, \dots, u_p), \\ &\vdots = \vdots \\ \dot{x}_n &= f_n(t, x_1, \dots, x_n, u_1, \dots, u_p),\end{aligned}$$

donde  $\dot{x}_i$  denota la derivada de  $x_i$  con respecto al tiempo  $t$ ,  $u_1, u_2, \dots, u_p$  las variables de entrada o señales de control. Recordemos que las variables  $x_1, x_2, \dots, x_n$  son las variables de estado.

Utilizando una notación vectorial, podemos representar el sistema anterior de la siguiente forma

$$\dot{x} = f(t, x, u).$$

El vector de mediciones se representa como

$$y = h(t, x, u),$$

donde  $h$  son mapeos (posiblemente no lineales) de sus argumentos a vectores de la dimensión apropiada.

A partir de la notación anterior, podemos modelar un sistema físico como se muestra a continuación

$$\dot{x} = f(t, x). \quad (33)$$

Si el sistema dado en Ec. (33) contiene funciones independientes de  $t$  se dice que el sistema es autónomo de lo contrario, se dice que es no autónomo.

## Análisis de la existencia, unicidad y continuidad de soluciones

### Espacio Euclídeo

Dado un conjunto de todos los vectores  $n$ -dimensionales  $x = (x_1, \dots, x_n)$  donde  $x_1, \dots, x_n$  son números reales, define el espacio Euclídeo  $n$ -dimensional denotado por  $R^n$ . El espacio Euclídeo unidimensional consiste de todos los números reales y es denotado por  $R$ .

#### Note

Los vectores en  $R^n$  se pueden sumar, multiplicar, multiplicar por un escalar y su producto interno, dados dos vectores  $x$  e  $y$  es  $x^T y = \sum_{i=1}^n x_i y_i$ .

### Norma vectorial y matricial

La norma  $\|x\|$  de un vector  $x$  es una función de valor real que tiene las siguientes propiedades:

- $\|x\| \geq 0 \forall x \in R^n$ , con  $\|x\| = 0$  si y sólo si  $x = 0$ .
- $\|x + y\| \leq \|x\| + \|y\| \forall x, y \in R^n$ .
- $\|\alpha x\| = |\alpha| \|x\| \forall \alpha \in R, x \in R^n$ .

La norma  $p$  está definida como sigue

$$\|x\|_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p}, \quad 1 \leq p < \infty,$$

y

$$\|x\|_\infty = \max_i |x_i|.$$

Las tres normas más utilizadas son  $\|x\|_1$ ,  $\|x\|_\infty$  y la norma Euclideana

$$\|x\|_2 = (|x_1|^2 + \cdots + |x_n|^2)^{1/2} = (x^T x)^{1/2}$$

## Funciones continuas

Una función  $f$  que mapea un conjunto  $S_1$  a un conjunto  $S_2$  se denota como  $f : S_1 \rightarrow S_2$ . Una función  $f : R^n \rightarrow R^m$  se dice *continua* en un punto  $x$  si  $f(x_k) \rightarrow f(x)$  cuando  $x_k \rightarrow x$ .

Equivalentemente,  $f$  es continua en  $x$  si, dado  $\varepsilon > 0$ , hay un  $\delta > 0$  tal que

$$\|x - y\| < \delta \Rightarrow \|f(x) - f(y)\| < \varepsilon.$$

Una función  $f$  es continua sobre un conjunto  $S$  si esta es continua en cada punto de  $S$ , y es *uniformemente continua* sobre  $S$  si, dado  $\varepsilon > 0$  hay  $\delta > 0$  (dependiente únicamente sobre  $\varepsilon$ ) tal que la desigualdad se mantiene para todo  $x, y \in S$ .

### Note

Continuidad uniforme está definida sobre un conjunto mientras que continuidad está definida sobre un punto.

## Funciones diferenciables

Una función  $f : R \rightarrow R$  es *diferenciable* en  $x$  si el límite

$$\dot{f}(x) = \lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h},$$

existe. El límite  $\dot{f}(x)$  es llamado la derivada de  $f$  en  $x$ .

Una función  $f : R^n \rightarrow R^m$  es *continuamente diferenciable* en un punto  $x_0$  si las derivadas parciales  $\partial f_i / \partial x_j$  existen y son continuas en  $x_0$  para  $1 \leq i \leq m, 1 \leq j \leq n$ .

Una función  $f$  es continuamente diferenciable sobre un conjunto  $S$  si esta es continuamente diferenciable en cada punto de  $S$ . Para una función continuamente diferenciable  $f : R^n \rightarrow R$ , el vector fila  $\partial f / \partial x$  está definido por

$$\frac{\partial f}{\partial x} = \left[ \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right].$$

El vector *gradiente*, denotado por  $\nabla f(x)$  es

$$\nabla f(x) = \left[ \frac{\partial f}{\partial x} \right]^T.$$

La *matriz Jacobiana*  $[\partial f / \partial x]$  es una matriz  $m \times n$  cuyo elemento en la  $i$ -ésima fila y la  $j$ -ésima columna es  $\partial f_i / \partial x_j$  para una función continuamente diferenciable  $f : R^n \rightarrow R^m$ .

## Teorema del valor medio

Si  $x$  y  $y$  son dos puntos distintos en  $R^n$ , entonces el *segmento de recta*  $L(x, y)$  que une  $x$  e  $y$  es

$$L(x, y) = \{z | z = \theta x + (1 - \theta)y, 0 < \theta < 1\}.$$

Asuma que  $f : R^n \rightarrow R$  es continuamente diferenciable en cada punto  $x$  de un conjunto abierto  $S \subset R^n$ . Sea  $x$  e  $y$  dos puntos de  $S$  tal que el segmento de recta  $L(x, y) \subset S$ , entonces existe un punto  $z$  de  $L(x, y)$  tal que

$$f(y) - f(x) = \frac{\partial f}{\partial x} \Big|_{x=z} (y - x).$$

## Existencia y unicidad

Una ecuación diferencial de la forma

$$\dot{x} = f(t, x), \quad x(t_0) = x_0,$$

podría tener varias soluciones. Por ejemplo, si consideramos la siguiente función

$$\dot{x} = x^{1/3}, \quad x(0) = 0, \quad (34)$$

si resolvemos la Ec. (34), utilizando Laplace por ejemplo, obtenemos que la siguiente solución

$$x(t) = \left( \frac{2t}{3} \right)^{3/2}.$$

Podemos observar que esta solución no es única debido a que  $x(t) \equiv 0$  es otra solución.

### Theorem 1 (Existencia local y unicidad)

Sea  $f(t, x)$  continuo a pedazos en  $t$  y que satisface la condición de Lipschitz

$$\|f(t, x) - f(t, y)\| \leq L\|x - y\|,$$

$\forall x, y \in B = \{x \in R^n | \|x - x_0\| \leq r\}$ ,  $\forall t \in [t_0, t_1]$ . Entonces, existe algún  $\delta > 0$  tal que la ecuación de estado

$$\dot{x} = f(t, x), \quad \text{con } x(t_0) = x_0,$$

tiene una solución única sobre  $[t_0, t_0 + \delta]$ .

### Ejemplo

- Considere la siguiente función

$$f(x) = \begin{bmatrix} -x_1 + x_1 x_2 \\ x_2 - x_1 x_2 \end{bmatrix}.$$

### Ejemplo

- Considere la siguiente función

$$f(x) = \begin{bmatrix} x_2 \\ -\text{sat}(x_1 + x_2) \end{bmatrix},$$

donde la función de saturación está definida por

$$\text{sat}(y) = \begin{cases} -1, & \text{para } y < -1, \\ y, & \text{para } |y| \leq 1, \\ 1, & \text{para } y > 1 \end{cases}$$

### Ejemplo

- Considere el siguiente sistema escalar

$$\dot{x} = -x^2, \quad x(0) = -1.$$

#### Lemma 1

[Khalil, 1992] Sea  $f : [a, b] \times D \rightarrow R^m$  continua para algún dominio  $D \subset R^n$ .

Suponga que  $[\partial f / \partial x]$  existe y es continua sobre  $[a, b] \times D$ . Si, para un subconjunto convexo  $W \subset D$ , hay una constante  $L \geq 0$  tal que

$$\left\| \frac{\partial f}{\partial x}(t, x) \right\| \leq L,$$

sobre  $[a, b] \times W$ , entonces

$$\|f(t, x) - f(t, y)\| \leq L \|x - y\|,$$

para todo  $t \in [a, b], x \in W$  e  $y \in W$ .

#### Lemma 2

[Khalil, 1992] Sea  $f(t, x)$  continua sobre  $[a, b] \times D$ , para algún dominio  $D \subset R^n$ . Si  $[\partial f / \partial x]$  existe y es continua sobre  $[a, b] \times D$ , entonces  $f$  es localmente Lipschitz en  $x$  sobre  $[a, b] \times D$ .

### Lemma 3

[[Khalil, 1992](#)] Sea  $f(t, x)$  continua sobre  $[a, b] \times R^n$ . Si  $[\partial f / \partial x]$  existe y es continua sobre  $[a, b] \times R^n$ , entonces  $f$  es globalmente Lipschitz en  $x$  sobre  $[a, b] \times R^n$  si y sólo si  $[\partial f / \partial x]$  es uniformemente acotado sobre  $[a, b] \times R^n$ .

### Theorem 2 (Existencia global y unicidad)

[[Khalil, 1992](#)] Suponga que  $f(t, x)$  es continua definida a pedazos sobre  $t$  y satisface

$$\begin{aligned}\|f(t, x) - f(t, y)\| &\leq L\|x - y\|, \\ \|f(t, x_0)\| &\leq h,\end{aligned}$$

$\forall x, y \in R^n, \forall t \in [t_0, t_1]$ . Entonces, la ecuación de estado

$$\dot{x} = f(t, x), \quad x(t_0) = x_0,$$

tiene una solución única sobre  $[t_0, t_1]$ .

### Ejemplo

- Considere el siguiente sistema lineal

$$\dot{x} = A(t)x + g(t) = f(t, x),$$

donde  $A(\cdot)$  y  $g(\cdot)$  son funciones continuas definidas a pedazos sobre  $t$ .

### Ejemplo

- Considere el siguiente sistema escalar

$$\dot{x} = -x^3 = f(x).$$

### Theorem 3

[[Khalil, 1992](#)] Sea  $f(t, x)$  continua a trozos sobre  $t$  y localmente Lipschitz en  $x$  para todo  $t \geq t_0$  y toda  $x$  en un dominio  $D \subset R^n$ . Sea  $W$  un subconjunto compacto de  $D$ ,  $x_0 \in W$ , suponiendo que cada solución de

$$\dot{x} = f(t, x), \quad x(t_0) = x_0,$$

yace enteramente en  $W$ . Entonces, hay una solución única que está definida para todo  $t \geq t_0$ .

### Theorem 4

[[Khalil, 1992](#)] Sea  $f(t, x)$  continua a pedazos sobre  $t$  y Lipschitz en  $x$  sobre  $[t_0, t_1] \times W$  con una constante Lipschitz  $L$ , donde  $W \subset R^n$  es un conjunton conectado. Sea  $y(t)$  y  $z(t)$  soluciones de

$$\dot{y} = f(t, y), \quad y(t_0) = y_0,$$

y

$$\dot{z} = f(t, z) + g(t, z), \quad z(t_0) = z_0,$$

tal que  $y(t), z(t) \in W$  para todo  $t \in [t_0, t_1]$ . Suponga que

$$\|g(t, x)\| \leq \mu, \quad \forall (t, x) \in [t_0, t_1] \times W,$$

para algún  $\mu > 0$ , y

$$\|y_0 - z_0\| \leq \gamma,$$

entonces

$$\|y(t) - z(t)\| \leq \gamma \exp [L(t - t_0)] + \frac{\mu}{L} \{ \exp [L(t - t_0)] - 1 \}, \quad \forall t \in [t_0, t_1].$$

### Theorem 5

[[Khalil, 1992](#)] Sea  $f(t, x, \lambda)$  continua en  $(t, x, \lambda)$  y localmente Lipschitz en  $x$  (uniformemente en  $t$  y  $\lambda$ ) sobre  $[t_0, t_1] \times D \times \{\|\lambda - \lambda_0 \leq c\}$  donde  $D \subset R^n$  es un conjunto abierto conectado. Sea  $y(t, \lambda_0)$  una solución de  $\dot{x} = f(t, x, \lambda_0)$  con  $y(t_0, \lambda_0) = y_0 \in D$ . Suponga que  $y(t, \lambda_0)$  está definida y pertenece a  $D$  para todo  $t \in [t_0, t_1]$ . Entonces, dado  $\varepsilon > 0$ , hay un  $\delta > 0$ , tal que si

$$\|z_0 - y_0\| < \delta, \quad \|\lambda - \lambda_0\| < \delta,$$

entonces hay una solución única  $z(t, \lambda)$  de  $\dot{x} = f(t, x, \lambda)$  definida sobre  $[t_0, t_1]$ , con  $z(t_0, \lambda) = z_0$  y  $z(t, \lambda)$  satisface

$$\|z(t, \lambda) - y(t, \lambda_0)\| < \varepsilon, \quad \forall t \in [t_0, t_1].$$

# Fundamentos del análisis de estabilidad en el sentido de Lyapunov

## Contents

- Localización de los puntos de equilibrio
- Linealización de sistemas en torno a un punto de equilibrio
- Análisis de la estabilidad de sistemas dinámicos linealizados
- Análisis de estabilidad por el método directo de Lyapunov
- Formas cuadráticas

Considere un sistema autónomo representado por la siguiente ecuación

$$\dot{x} = f(x), \quad (35)$$

donde  $f : D \leftarrow R^n$  es un mapa local de Lipschitz de un dominio  $D \subset R^n$  en  $R^n$ .

Para poder estudiar la estabilidad del sistema (35), es necesario obtener  $\bar{x} \in D$  que denota el punto de equilibrio

$$f(\bar{x}) = 0.$$

# Localización de los puntos de equilibrio

## 🔔 Definition 1

El punto de equilibrio  $x = 0$  de  $\dot{x} = f(x)$  es

estable si, para cada  $\varepsilon > 0$ , hay un  $\delta = \delta(\varepsilon) > 0$  tal que

$$\|x(0)\| < \delta \Rightarrow \|x(t)\| < \varepsilon, \quad \forall t \geq 0$$

- inestable si no es estable.
- asintóticamente estable si este es estable y  $\delta$  se escoge tal que.

$$\|x(0)\| < \delta \Rightarrow \lim_{t \rightarrow \infty} x(t) = 0.$$

## Ejemplo 1.

Considere las ecuaciones del péndulo simple y obtenga sus puntos de equilibrio

$$\begin{aligned}\dot{x}_1 &= x_2, \\ \dot{x}_2 &= -\frac{g}{l} \sin(x_1) - \frac{k}{m} x_2.\end{aligned}$$

## Linealización de sistemas en torno a un punto de equilibrio

Un sistema (LTI) de la forma

$$\dot{x} = Ax, \tag{36}$$

tiene un punto de equilibrio en el origen. Este punto de equilibrio está aislado si y sólo si  $\det(A) \neq 0$ . Si  $\det(A) = 0$ , la matriz  $A$  tiene un espacio nulo no trivial.

Las propiedades de estabilidad del origen pueden ser caracterizadas por las ubicaciones de los *eigenvalores* de la matriz  $A$ .

Recordemos que la solución al sistema  $\dot{x} = Ax$  dada una condición inicial  $x_0$  está dada como sigue

$$x(t) = \exp(At)x_0.$$

Además, para cualquier matriz  $A$  existe una matriz  $P$  no singular que transforma  $A$  en su forma de **Jordan**, i.e.

$$P^{-1}AP = J = \text{blockdiag} [J_1, J_2, \dots, J_r],$$

donde  $J_i$  es el bloque de Jordan asociado con el eigenvalor  $\lambda_i$  de  $A$ .

La forma generalizada del bloque de Jordan, considerando un orden  $m$ , tiene la forma siguiente

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_i & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & & & \vdots \\ \vdots & & & \ddots & & 0 \\ \vdots & & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 & \lambda_i \end{bmatrix}_{m \times m},$$

por consiguiente

$$\exp(At) = P \exp(Jt) P^{-1} = \sum_{i=1}^r \sum_{k=1}^{m_i} t^{k-1} \exp(\lambda_i t) R_{ik},$$

donde  $m_i$  es el orden del bloque de Jordan asociado con el eigenvalor  $\lambda_i$ .

### Theorem 6

El punto de equilibrio  $x = 0$  de  $\dot{x} = Ax$  es estable si y sólo si todos los eigenvalores de  $A$  satisfacen  $\operatorname{Re}(\lambda_i) \leq 0$  y cada eigenvalor con  $\operatorname{Re}(\lambda_i) = 0$  tiene un bloque de Jordan asociado de uno. El punto de equilibrio  $x = 0$  es (globalmente) asintóticamente estable sí y sólo sí todos los eigenvalores de  $A$  satisfacen  $\operatorname{Re}(\lambda_i) < 0$ .

La demostración de este teorema se puede consultar en Khalil (2002).

## Linearización

Suponga que las funciones  $f_1$  y  $f_2$  del sistema autónomo de segundo orden

$$\begin{aligned}\dot{x}_1 &= f_1(x_1, x_2), \\ \dot{x}_2 &= f_2(x_1, x_2),\end{aligned}\tag{37}$$

son continuamente diferenciables y que además  $p = (p_1, p_2)$  es un punto de equilibrio, el sistema (37) se puede expandir mediante serie de Taylor en torno a  $(p_1, p_2)$  como sigue

$$\begin{aligned}\dot{x}_1 &= f_1(p_1, p_2) + a_{11}(x_1 - p_1) + a_{12}(x_2 - p_2) + \text{H.O.T}, \\ \dot{x}_2 &= f_2(p_1, p_2) + a_{21}(x_1 - p_1) + a_{22}(x_2 - p_2) + \text{H.O.T},\end{aligned}$$

donde

$$\begin{aligned}a_{11} &= \left. \frac{\partial f_1(x_1, x_2)}{\partial x_1} \right|_{x_1=p_1, x_2=p_2}, & a_{12} &= \left. \frac{\partial f_1(x_1, x_2)}{\partial x_2} \right|_{x_1=p_1, x_2=p_2}, \\ a_{21} &= \left. \frac{\partial f_2(x_1, x_2)}{\partial x_1} \right|_{x_1=p_1, x_2=p_2}, & a_{22} &= \left. \frac{\partial f_2(x_1, x_2)}{\partial x_2} \right|_{x_1=p_1, x_2=p_2},\end{aligned}$$

y H.O.T. denota los términos de orden superior de la expansión de Taylor, es decir  $(x_1 - p_1)^2, (x_2 - p_2)^2, (x_1 - p_1) \times (x_2 - p_2)$ , y así sucesivamente.

Dado que  $(p_1, p_2)$  denota el punto de equilibrio, tenemos

$$f_1(p_1, p_2) = f_2(p_1, p_2) = 0.$$

Puesto que consideramos trayectorias cercanas a este punto, definimos

$$y_1 = x_1 - p_1, \quad y_2 = x_2 - p_2.$$

Por consiguiente, las ecuaciones de estado quedan de la siguiente forma

$$\begin{aligned}\dot{y}_1 &= \dot{x}_1 = a_{11}y_1 + a_{12}y_2 + \text{H.O.T}, \\ \dot{y}_2 &= \dot{x}_2 = a_{21}y_1 + a_{22}y_2 + \text{H.O.T}.\end{aligned}$$

Si consideramos una vecindad cercana al punto de equilibrio, entonces podemos decir que los términos H.O.T son despreciables y por consiguiente, realizamos una aproximación de la ecuación de estado no lineal por la ecuación de estado lineal como sigue

$$\begin{aligned}\dot{y}_1 &= \dot{x}_1 = a_{11}y_1 + a_{12}y_2, \\ \dot{y}_2 &= \dot{x}_2 = a_{21}y_1 + a_{22}y_2.\end{aligned}$$

Reescribiendo lo anterior en una forma vectorial, tenemos

$$\dot{y} = Ay,$$

donde

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \left[ \begin{array}{cc} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{array} \right]_{x=p} = \frac{\partial f}{\partial x} \Big|_{x=p}$$

### 1 Note

La matriz  $[\partial f / \partial x]$  es llamada la matriz Jacobiana de  $f(x)$  mientras que  $A$  es la matriz Jacobiana evaluada en el punto de equilibrio  $p$ .

### Ejemplo

- Considere el circuito de diodo tunel que se muestra a continuación

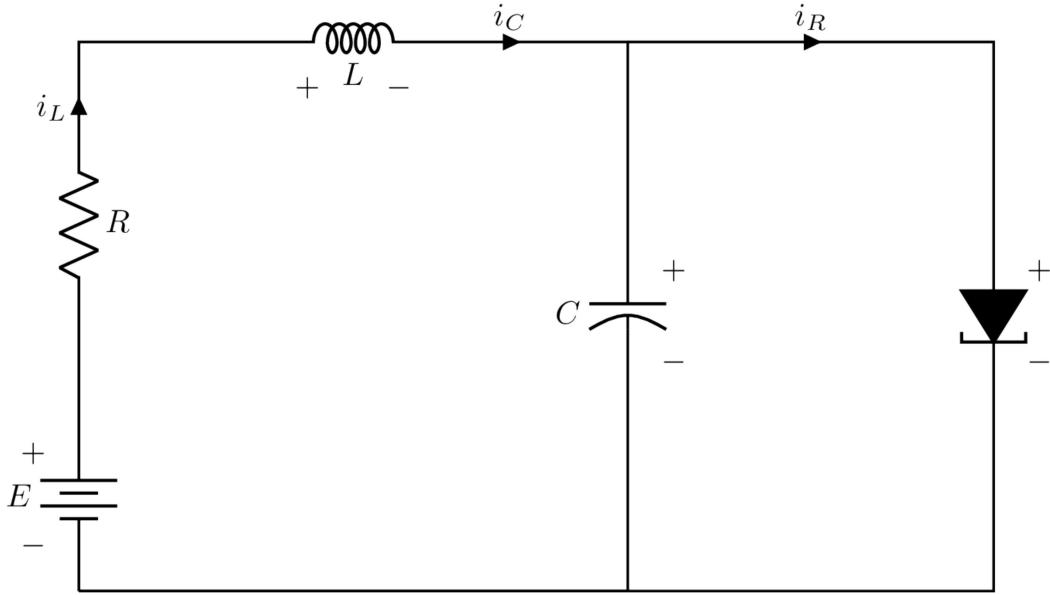


Fig. 7 Diodo túnel.

donde los elementos de almacenamiento de energía son el capacitor  $C$  y el inductor  $L$  mientras que el diodo tunel es representado por la relación  $i_R = h(v_R)$ . Asumiendo que los elementos del circuito son lineales e invariantes en el tiempo, el circuito de diodo tunel se puede representar matemáticamente como sigue

$$\begin{aligned} i_C &= C \frac{dv_C}{dt}, \\ v_L &= L \frac{di_L}{dt}, \end{aligned} \tag{38}$$

donde  $i, v$  representan la corriente y el voltaje a través de los elementos, respectivamente.

Considere que los parámetros del modelo son  $u = 1.2V$ ,  $R = 1.5k\Omega = 1.5 \times 10^3\Omega$ ,  $C = 2pF = 2 \times 10^{-12}F$ ,  $L = 5\mu H = 5 \times 10^{-6}H$ . Además, suponga que  $h(\cdot)$  está dada por

$$h(x_1) = 17.76x_1 - 103.79x_1^2 + 229.62x_1^3 - 226.31x_1^4 + 83.72x_1^5.$$

1. Utilice las Leyes de Kirchhoff para reescribir las ecuaciones anteriores.
2. Encuentre los puntos de equilibrio.
3. Linealice el sistema.
4. Evalúe la matriz Jacobiana en los puntos de equilibrio.

## Ejemplo

- Considere el péndulo simple mostrado en la siguiente

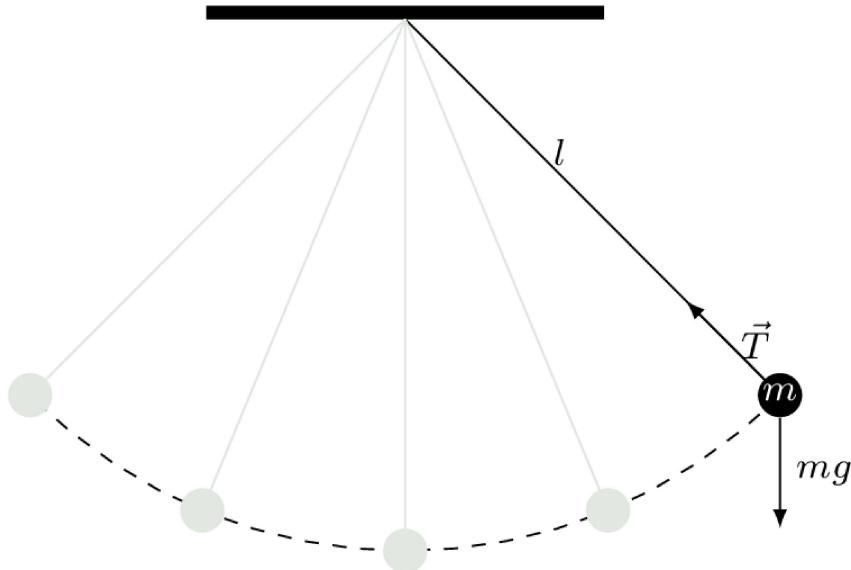


Fig. 8 Diagrama de péndulo simple sin fuerza externa.

donde  $l$  denota la longitud de la varilla, y  $m$  la masa de la bola.

Cuya representación en espacio de estados está dada como sigue

$$\begin{aligned}\dot{x}_1 &= x_2, \\ \dot{x}_2 &= -\frac{g}{l} \sin(x_1) - \frac{k}{m} x_2.\end{aligned}\tag{39}$$

1. Encuentre los puntos de equilibrio.
2. Linealice el sistema.
3. Evalúe la matriz Jacobiana en los puntos de equilibrio.

## Análisis de la estabilidad de sistemas dinámicos linealizados

Considere un sistema LTI como se muestra en la Ec. [\(36\)](#) donde  $A$  es una matriz real de dimensión  $2 \times 2$ . La solución a este sistema con una condición inicial  $x_0$  está dada por la siguiente ecuación

$$x(t) = M \exp(J_r t) M^{-1} x_0,$$

donde  $J_r$  es la forma real de Jordan de  $A$  y  $M$  es una matriz real no singular tal que  $M^{-1}AM = J_r$ . Dependiendo de los eigenvalores  $\lambda_i$  de la matriz  $A$ , la forma de Jordan puede tomar los siguientes casos

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \quad \begin{bmatrix} \lambda & k \\ 0 & \lambda \end{bmatrix}, \quad \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix},$$

donde  $k$  puede tomar el valor de 0 o 1.

## Caso 1. Ambos eigenvalores son reales: $\lambda_1 \neq \lambda_2 \neq 0$ .

El sistema tiene dos eigenvectores reales  $v_1$  y  $v_2$  asociados con  $\lambda_1$  y  $\lambda_2$ , respectivamente.

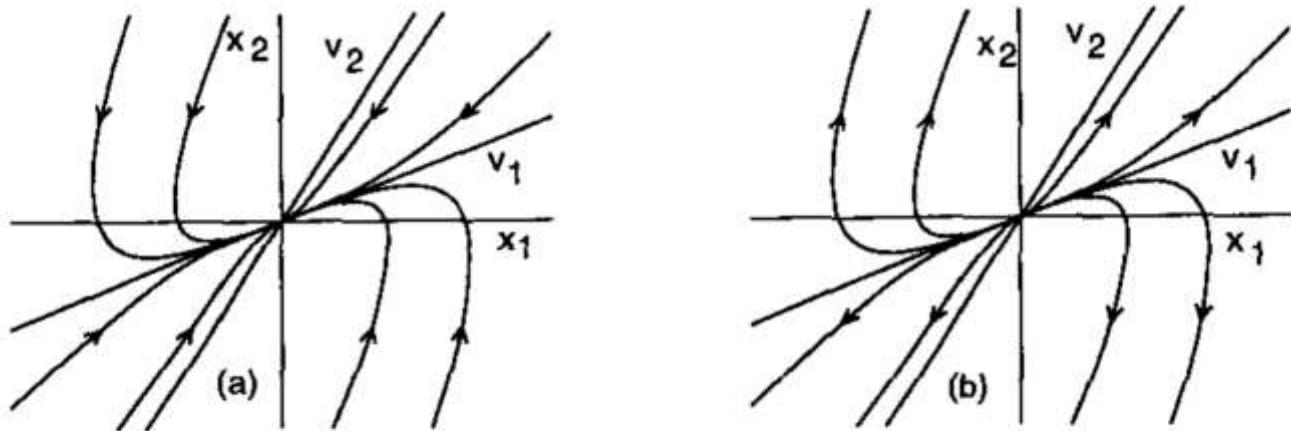
Donde el sistema es transformado en dos ecuaciones diferenciales de primer orden después de un cambio de coordenadas  $z = M^{-1}x$ , i.e.

$$\begin{aligned}\dot{z}_1 &= \lambda_1 z_1, \\ \dot{z}_2 &= \lambda_2 z_2,\end{aligned}$$

cuya solución está dada como sigue, para las condiciones iniciales  $(z_{10}, z_{20})$

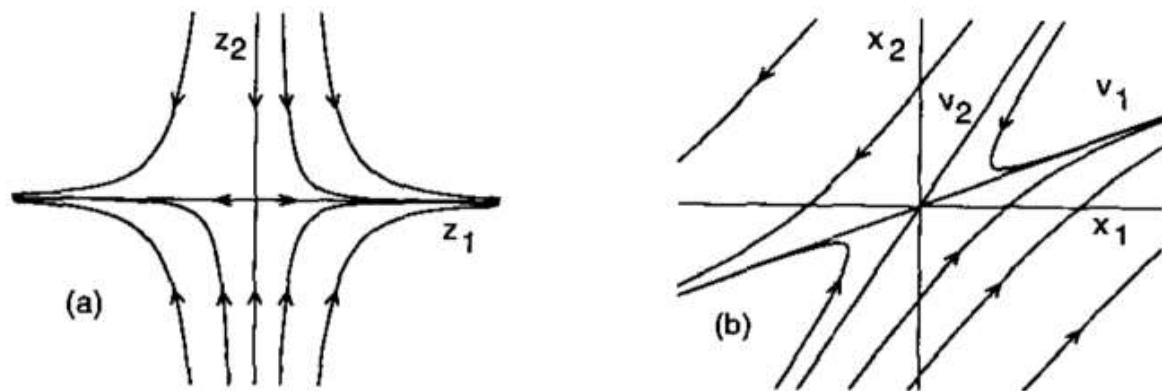
$$z_1(t) = z_{10} e^{\lambda_1 t}, \quad z_2(t) = z_{20} e^{\lambda_2 t}.$$

Cuando  $\lambda_2 < \lambda_1 < 0$ , el retrato fase tiene la forma de la Figura 1.15(a), donde  $x = 0$  es llamado *nodo estable*. Por otro lado, si  $\lambda_2 > \lambda_1 > 0$ ,  $x = 0$ , el retrato fase se muestra en la Figura 1.15(b) y se le conoce como *nodo inestable*.



**Figure 1.15: Phase portraits for (a) a stable node; (b) an unstable node.**

Si el sistema tiene eigenvalores con signos opuestos, esto es  $\lambda_2 < 0 < \lambda_1$ , decimos que  $\lambda_2$  es un eigenvalor estable y  $\lambda_1$  inestable. Por consiguiente, el retrato fase que corresponde a esta trayectoria se muestra en la Figura 1.16(a).



**Figure 1.16: Phase portrait of a saddle point (a) in modal coordinates; (b) in original coordinates.**

## Caso 2. Eigenvalores complejos: $\lambda_{1,2} = \alpha \pm j\beta$

El sistema  $\dot{x} = Ax$  después de la transformación de coordenadas  $z = M^{-1}x$  tiene la forma

$$\begin{aligned}\dot{z}_1 &= \alpha z_1 - \beta z_2, \\ \dot{z}_2 &= \beta z_1 + \alpha z_2.\end{aligned}$$

Puesto que la solución a este sistema es osculatoria, se puede expresar en coordenadas polares como sigue

$$r = \sqrt{z_1^2 + z_2^2}, \quad \theta = \tan^{-1} \left( \frac{z_2}{z_1} \right),$$

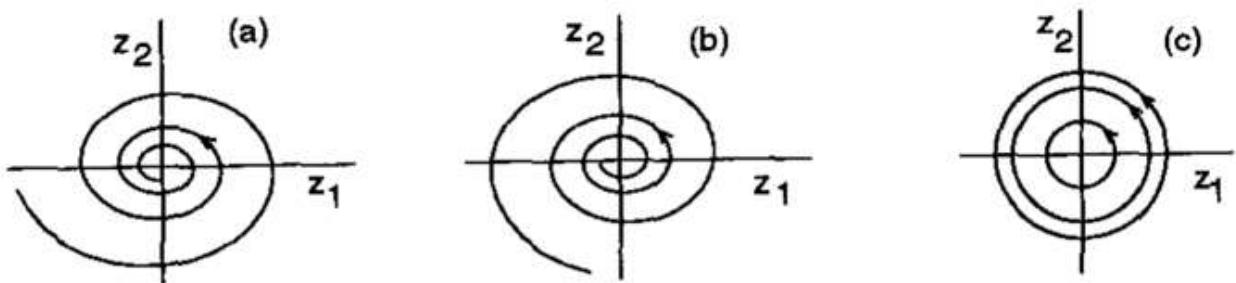
teniendo dos ecuaciones diferenciales de primer orden desacopladas

$$\begin{aligned}\dot{r} &= \alpha r, \\ \dot{\theta} &= \beta.\end{aligned}$$

La solución de este sistema, para una condición inicial  $(r_0, \theta_0)$  es

$$r(t) = r_0 e^{\alpha t}, \quad \theta(t) = \theta_0 + \beta t.$$

Dependiendo del valor de  $\alpha$ , el retrato fase puede tomar alguna de las siguientes formas

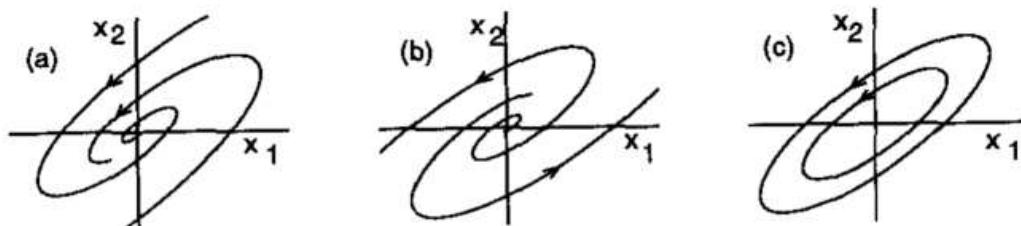


**Figure 1.17: Typical trajectories in the case of complex eigenvalues.**

(a)  $\alpha < 0$ ; (b)  $\alpha > 0$ ; (c)  $\alpha = 0$ .

### i Note

El punto de equilibrio  $x = 0$  es un *foco estable* si  $\alpha < 0$ , *foco inestable* si  $\alpha > 0$  y *centro* si  $\alpha = 0$  como se muestra a continuación



**Figure 1.18: Phase portraits for (a) a stable focus; (b) an unstable focus; (c) a center.**

### Caso 3. Eigenvalores múltiples diferentes de cero: $\lambda_1 = \lambda_2 = \lambda \neq 0$

Para este caso, el sistema  $\dot{x} = Ax$  después de un cambio de coordenadas toma la siguiente forma

$$\begin{aligned}\dot{z}_1 &= \lambda z_1 + kz_2, \\ \dot{z}_2 &= \lambda z_2,\end{aligned}$$

cuya solución, con c.i.  $(z_{10}, z_{20})$  está dada como sigue

$$z_1(t) = e^{\lambda t} (z_{10} + kz_{20}t), \quad z_2(t) = e^{\lambda t} z_{20}.$$

Los retrato fase para este caso, considerando  $k = 0$  y  $k = 1$  se muestran en la siguiente figura

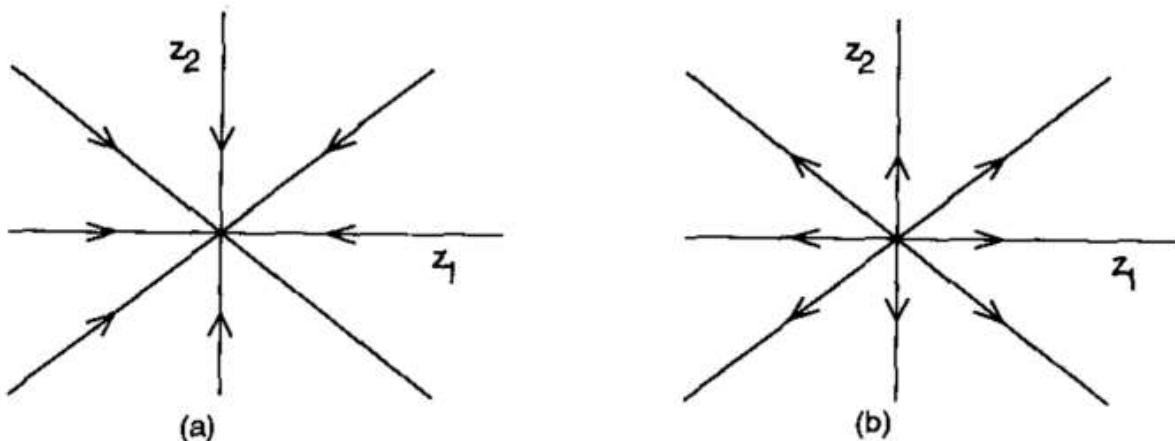
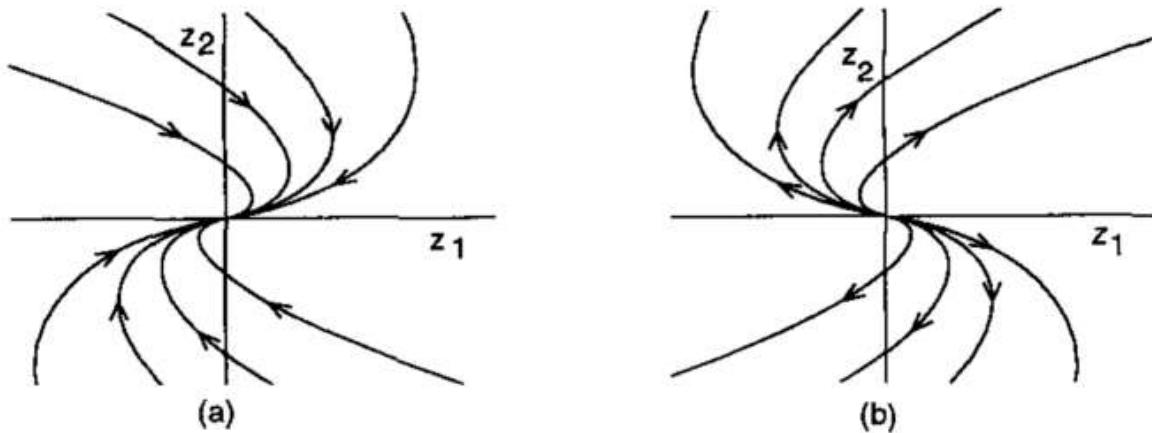


Figure 1.19: Phase portraits for the case of nonzero multiple eigenvalues when  $k = 0$ : (a)  $\lambda < 0$ ; (b)  $\lambda > 0$ .

Aquí, el punto de equilibrio  $x = 0$  se le conoce como *nodo estable* si  $\lambda < 0$  y *nodo inestable* si  $\lambda > 0$ . Las trayectorias se muestran en la siguiente representación



**Figure 1.20: Phase portraits for the case of nonzero multiple eigenvalues when  $k = 1$ :**  
**(a)  $\lambda < 0$ ; (b)  $\lambda > 0$ .**

## Caso 4. Uno o ambos autovalores son cero

Cuando  $\lambda_1 = 0$  y  $\lambda_2 \neq 0$ , el sistema después de una transformación de coordenadas tiene la siguiente representación

$$\begin{aligned}\dot{z}_1 &= 0, \\ \dot{z}_2 &= \lambda_2 z_2,\end{aligned}$$

cuya solución es

$$z_1(t) = z_{10}, \quad z_2(t) = z_{20}e^{\lambda_2 t}.$$

Cuando ambos eigenvalores se encuentran en el origen y se aplica una transformación del sistema  $\dot{x} = Ax$ , tenemos

$$\begin{aligned}\dot{z}_1 &= z_2, \\ \dot{z}_2 &= 0,\end{aligned}$$

cuya solución está dada como sigue

$$z_1(t) = z_{10} + z_{20}t, \quad z_2(t) = z_{20}.$$

Las trayectorias que forman el retrato fase para estos casos en particular se muestran en las siguientes figuras

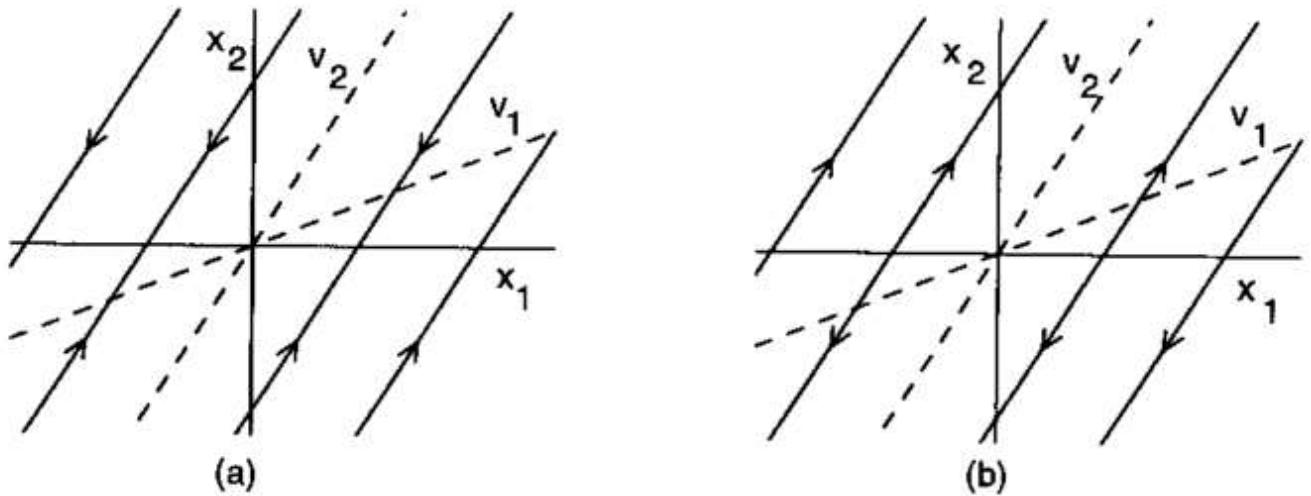


Figure 1.21: Phase portraits for (a)  $\lambda_1 = 0, \lambda_2 < 0$ ; (b)  $\lambda_1 = 0, \lambda_2 > 0$ .

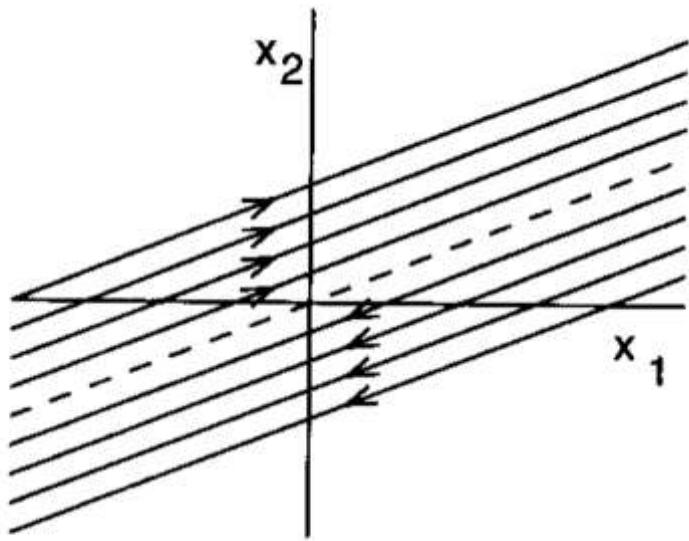


Figure 1.22: Phase portrait when  $\lambda_1 = \lambda_2 = 0$ .

### Ejemplo

- Clasifique los puntos de equilibrio del circuito de diodo tunel (38) y determine su estabilidad.

### Ejemplo

- Clasifique los puntos de equilibrio del sistema de péndulo simple (39) y determine su estabilidad.

# Análisis de estabilidad por el método directo de Lyapunov

Aleksandr Mikhailovich Lyapunov en 1892 presentó dos métodos para determinar la estabilidad de sistemas dinámicos representados en ecuaciones diferenciales ordinarias (ODEs).

El primer método utiliza la forma explícita de la solución de las ODEs para realizar el análisis de estabilidad mientras que el segundo método determina la estabilidad del sistema sin resolver las ecuaciones de estado.

Sea  $V : D \rightarrow \mathbb{R}$  una función continuamente diferenciable definida en un dominio  $D \subset \mathbb{R}^n$  que contiene el origen, entonces la derivada de  $V$  sobre la trayectoria del sistema autónomo  $\dot{x} = f(x)$ , denotada por  $\dot{V}(x)$  está dada como sigue

$$\begin{aligned}\dot{V}(x) &= \sum_{i=1}^n \frac{\partial V}{\partial x_i} \dot{x}_i = \sum_{i=1}^n \frac{\partial V}{\partial x_i} f_i(x), \\ &= \left[ \frac{\partial V}{\partial x_1}, \quad \frac{\partial V}{\partial x_2}, \quad \dots, \quad \frac{\partial V}{\partial x_n} \right] \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{bmatrix} = \frac{\partial V}{\partial x} f(x).\end{aligned}$$

Normalmente las funciones utilizadas son funciones reales de variable vectorial. También es posible utilizar funciones integrales pero nos vamos a centrar en las primeras.

## 🔔 Definition 2

Una función  $V$  es *definida positiva* si  $V(x) > 0 \forall x \neq 0$  y  $V(0) = 0$ .

## Ejemplos

- $V(x) = x_1^2 + 2x_2^2$
- $V(x) = x_1^2 + \frac{2x_2^2}{1+x_2^2}$

## 🔔 Definition 3

Una función  $V$  es *definida negativa* si  $V(x) < 0 \forall x \neq 0$  y  $V(0) = 0$ .

## Ejemplo

- $V(x) = -x_1^2 - (3x_1 + 2x_2)^2$

### 🔔 Definition 4

Una función  $V$  es *semidefinida positiva* si  $V(x) \geq 0 \forall x$  y  $V(x) = 0$ .

## Ejemplo

- $V(x) = (x_1 + x_2)^2$ .

### 🔔 Definition 5

Una función  $V$  es *semidefinida negativa* si  $-V(x)$  es semidefinida positiva, i.e.  
 $-V(x) \geq 0 \forall x$  y  $V(x) = 0$ .

## Ejemplo

- $V(x) = -(x_1 + x_2)^2$

### ℹ Note

$\dot{V}(x)$  es diferente para cada sistema y además, si  $\Phi(t; x)$  es la solución al sistema  $\dot{x} = f(x)$  con valor inicial  $x$  en  $t = 0$ , entonces

$$\dot{V}(x) = \frac{d}{dt} V(\Phi(t; x))|_{t=0}.$$

## Ejemplo

- Considere la siguiente función

$$V(x) = (x_1 + x_2)^2.$$

## Ejemplo

- Considere la siguiente función y determine si es definida positiva o semidefinida positiva

$$V(x) = x_1^2 + x_2^2 + 1.$$

### Ejemplo

- Considere la siguiente función y determine si es definida negativa

$$V(x) = -x_1^2 - (3x_1 + 2x_2)^2$$

### Ejemplo

- Considere la siguiente función

$$V(x) = x_1x_2 + x_2^2$$

## Formas cuadráticas

Una forma general de tener funciones  $V(x)$  es utilizando la forma cuadrática. Por ejemplo

$$V(x) = [x_1 \ x_2 \ \cdots \ x_n] \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{12} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1n} & p_{2n} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} =,$$

donde  $P$  es una matriz simétrica. Es decir  $P = P^T$ .

### Ejemplo

- Considere la siguiente función en forma cuadrática y determine si es definida positiva

$$\begin{aligned}
V(x) &= x^T P x \\
&= [x_1 \quad x_2] \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
&= [x_1 + 2x_2 \quad 2x_1 + 3x_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
&= x_1^2 + 2x_1x_2 + 2x_1x_2 + 3x_2^2 \\
&= x_1^2 + 4x_1x_2 + 3x_2^2
\end{aligned}$$

## Criterio de Sylvester

Una función  $V(x)$  en forma cuadrática o hermitiana es definida positiva si todas las menores principales sucesivas de  $P$  son positivas.

$$p_{11} > 0, \quad \begin{vmatrix} p_{11} & p_{12} \\ \bar{p}_{12} & p_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{12} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1n} & p_{2n} & \cdots & p_{nn} \end{vmatrix} > 0$$

### Ejemplo

- Considere la siguiente matriz  $P$

$$P = \begin{bmatrix} 1 & 6 & -1 \\ 2 & 4 & 0 \\ -2 & 3 & 7 \end{bmatrix}$$

### Ejemplo

- Considere la siguiente función y determine si es definida positiva

$$V(x) = 10x_1^2 + 4x_2^2 + x_3^2 + 2x_1x_2 - 2x_2x_3 - 4x_1x_3$$

# Principio del Teorema de Estabilidad de Lyapunov

Un sistema es estable si su  $E$ -energía- total (una función definida positiva) es continuamente decreciente (la derivada en el tiempo de la  $E$  total es definida negativa) hasta que alcanza un estado de equilibrio. En otras palabras, un sistema es estable si su  $E$  total se va perdiendo con el paso del tiempo.

## Theorem 7 (Teorema de estabilidad de Lyapunov)

Sea  $x = 0$  un punto de equilibrio para el sistema  $\dot{x} = f(x)$  y  $D \subset R^n$  un dominio que contiene  $x = 0$ , entonces  $V : D \rightarrow R$  es una función continuamente diferenciable tal que

$$V(0) = 0,$$

$$V(x) > 0 \text{ en } D - \{0\},$$

$$\dot{V}(x) \leq 0 \text{ en } D.$$

Entonces,  $x = 0$  es estable. Además, si

$$\dot{V}(x) < 0 \text{ en } D - \{0\},$$

entonces  $x = 0$  es asintóticamente estable.

La demostración de este teorema se puede consultar en Khalil (2002).

## Función candidata de Lyapunov

Es aquella función que se propone para poder realizar el análisis de estabilidad de Lyapunov. Recordemos que anteriormente trabajamos con funciones definidas positivas y cómo verificar que cumplieran con la definición mas no con funciones candidatas de Lyapunov.

En términos estrictos, una función candidata de Lyapunov es una función definida positiva. Esta función definida positiva puede ser de manera local o global. En nuestro caso, consideraremos que es local. Esto es

$$V(0) = 0$$

$$V(x) > 0 \quad \forall x \in U \setminus \{0\}$$

donde  $U$  denota un conjunto del espacio de estados y  $\setminus \{0\}$  significa **excepto en el origen**.

### **i Note**

Recordemos que el valor  $(0, 0)$  se refiere al origen del sistema que puede ser un punto de equilibrio. Cuando linealizamos un sistema no lineal en un punto de equilibrio que no sea el origen, entonces el origen del sistema lineal es el punto de equilibrio del sistema lineal.

### **Ejemplo**

- Considere el siguiente sistema, y verifique si es estable o no en el sentido de Lyapunov

$$\begin{aligned}\dot{x}_1 &= x_2 - x_1(x_1^2 + x_2^2), \\ \dot{x}_2 &= -x_1 - x_2(x_1^2 + x_2^2).\end{aligned}$$

### **Ejemplo**

- Considere el sistema de péndulo simple sin fricción

$$\begin{aligned}\dot{x}_1 &= x_2, \\ \dot{x}_2 &= -\frac{g}{l} \sin(x_1),\end{aligned}$$

$$\text{donde } x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix}.$$

### Theorem 8

Sea  $x = 0$  un punto de equilibrio para el sistema  $\dot{x} = f(x)$  y  $V : \mathbb{R}^n \rightarrow \mathbb{R}$  una función continuamente diferenciable tal que

$$V(0) = 0, \quad V(x) > 0, \quad \forall x \neq 0,$$

$$\|x\| \rightarrow \infty \Rightarrow V(x) \rightarrow \infty,$$

$$\dot{V}(x) < 0, \forall x \neq 0,$$

entonces  $x = 0$  es *globalmente asintóticamente estable*.

### Ejemplo

- Considere el siguiente sistema no lineal escalar

$$\dot{x} = \frac{2}{1+x} - x.$$

### Ejemplo

- Considere las ecuaciones que caracterizan un modelo para un péndulo colgante

$$\begin{aligned}\dot{x}_1 &= x_2, \\ \dot{x}_2 &= -\sin(x_1),\end{aligned}$$

donde  $x_1$  denota el ángulo entre el péndulo y la vertical con  $x_1 > 0$  corresponde a la rotación en sentido contrario a las manecillas del reloj.

# Controlabilidad de sistemas dinámicos

Print to PDF

## Contents

- Concepto de controlabilidad
- Formas canónicas de controlabilidad
- Estabilización por retroalimentación de estado
- Respuesta de estado estacionario

## Concepto de controlabilidad

Considere un sistema  $n$ -dimensional en representación de espacio-estado

$$\begin{aligned}\dot{x} &= Ax + Bu, \\ y &= Cx + Du,\end{aligned}\tag{40}$$

donde  $A, B, C$  son matrices de dimensión  $n \times n, n \times 1$ , y  $1 \times n$ , respectivamente. Además,  $D$  es escalar.

### 🔔 Definition 6

El sistema (40) es **controlable** si cualquier estado se puede transferir a cualquier otro estado en un tiempo finito aplicando una entrada.

### 🔔 Theorem 9

Para cada  $y$ , existe una solución  $x$  en  $Ax = y$  si y sólo si  $A$  es de rango completo.

Recordemos que la solución al sistema (40) dado una entrada  $u(t)$  y una c.i.  $x(0)$  está dado como sigue

$$x(t) = \underbrace{e^{At}x(0)}_{\text{Zero-Input Response}} + \underbrace{e^{At} \int_0^t e^{-A\tau} Bu(\tau) d\tau}_{\text{Zero-State Response}}, \quad (41)$$

y

$$y(t) = ce^{At}x(0) + ce^{At} \int_0^t e^{-A\tau} Bu(\tau) d\tau + Du(t).$$

Reescribiendo (41), tenemos

$$\begin{aligned} x(t) - e^{At}x(0) &= \int_0^t \left[ I + A(t-\tau) + A^2 \frac{(t-\tau)^2}{2} + \dots \right] Bu(\tau) d\tau, \\ &= B \int_0^t u(\tau) d\tau + AB \int_0^t (t-\tau)u(\tau) d\tau \\ &\quad + A^2B \int_0^t \frac{1}{2}(t-\tau)^2u(\tau) d\tau + \dots, \\ &= [B \quad AB \quad A^2B \dots] \begin{bmatrix} \int_0^t u(\tau) d\tau \\ \int_0^t (t-\tau)u(\tau) d\tau \\ \int_0^t \frac{1}{2}(t-\tau)^2u(\tau) d\tau \\ \vdots \end{bmatrix} \end{aligned}$$

Utilizando el [Theorem 9](#), tenemos que para cualquier  $x(0)$  y  $x(t)$ , una solución  $u(t)$  existe si y sólo si la matriz

$$[B \quad AB \quad A^2B \quad \dots \quad A^{n-1}B \quad \dots] \quad (42)$$

tiene rango  $n$ .

Utilizando el teorema de Cayley-Hamilton, la matriz (42) tiene rango  $n$  si y sólo si la matriz  $n \times n$

$$U = [B \quad AB \quad A^2B \quad \dots \quad A^{n-1}B] \quad (43)$$

tiene rango  $n$ .

En resumen, el sistema (40) es controlable si la matriz (43) tiene rango  $n$ , o bien,  $\det\{U\} \neq 0$ .

**i Note**

La matriz (43) es llamada **matriz de controlabilidad**.

**Ejemplo**

- Considere la siguiente función de transferencia

$$G(s) = \frac{N(s)}{D(s)} = \frac{2s - 1}{s^2 - 1.5s - 1} = \frac{2s - 1}{(s - 2)(s + \frac{1}{2})}$$

**Ejemplo**

- Considere la siguiente función de transferencia

$$G(s) = \frac{N(s)}{D(s)} = \frac{2s + 1}{s^2 - 1.5s - 1} = \frac{2s + 1}{(s - 2)(s + \frac{1}{2})}$$

**Ejemplo**

- Considere el sistema de equilibrio que se muestra en la siguiente Figura.

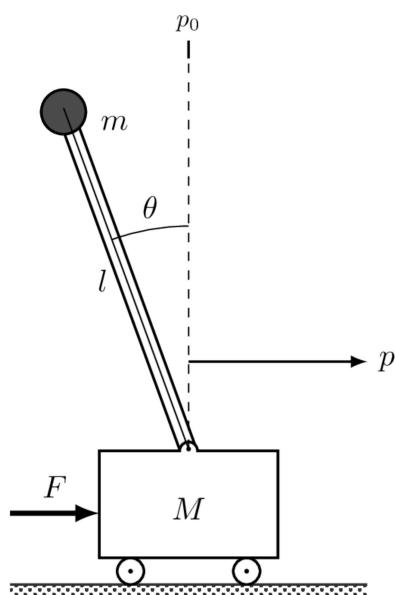


Fig. 9 Modelo del sistema carro-péndulo.

En este ejemplo, el centro de masa se equilibra por encima de un punto de pivot. Un ejemplo de ello se puede ver en la película "Guardia de supermercado".



Fig. 10 Paul Blart. Guardia de supermercado.

El modelo no lineal para el sistema de equilibrio se muestra a continuación

$$(M + m)\ddot{p} - ml \cos(\theta)\ddot{\theta} = -c\dot{p} - ml \sin(\theta)\dot{\theta}^2 + F,$$

$$(J + ml^2)\ddot{\theta} - ml \cos(\theta)\ddot{p} = -\gamma\dot{\theta} + mg\sin(\theta),$$

donde  $M$  es la masa de la base,  $m$  y  $J$  la masa y el momento de inercia del sistema a ser balanceado. Además,  $l$  denota la distancia de la base al centro de masa del cuerpo balanceado,  $c$  y  $\gamma$  son los coeficientes de fricción viscosa así como  $g$  representa la aceleración debido a la gravedad.

La dinámica del sistema en espacio de estado se puede obtener considerando como variables de estado  $x = (p, \theta, \dot{p}, \dot{\theta})$ , la entrada como  $u = F$  y salida como  $y = (p, \theta)$ .

Si definimos la masa total y la inercia total como sigue, tenemos

$$M_t = M + m, \quad J_t = J + ml^2.$$

Por consiguiente, podemos representar las ecuaciones de movimiento como sigue

$$\begin{bmatrix} \dot{p} \\ \dot{\theta} \\ \ddot{p} \\ \ddot{\theta} \end{bmatrix} = \begin{bmatrix} \dot{p} \\ \dot{\theta} \\ \frac{-mls_{\theta}\dot{\theta}^2 + mg(ml^2/J_t)s_{\theta}c_{\theta} - cp - \gamma lmc_{\theta}\dot{\theta} + u}{M_t - m(ml^2/J_t)c_{\theta}^2} \\ \frac{-ml^2s_{\theta}c_{\theta}\dot{\theta}^2 + M_tgls_{\theta} - clc_{\theta}\dot{p} - \gamma(M_t/m)\dot{\theta} + lc_{\theta}u}{J_t(M_t/m) - m(lc_{\theta})^2} \end{bmatrix},$$

$$y = \begin{bmatrix} p \\ \theta \end{bmatrix},$$

donde  $c_{\theta} = \cos(\theta)$  y  $s_{\theta} = \sin(\theta)$ .

En ocasiones, el ángulo  $\theta$  puede ser cercano a 0. Por consiguiente, podemos hacer las siguientes suposiciones:  $\sin(\theta) \approx \theta$  y  $\cos(\theta) \approx 1$ . Además, si  $\dot{\theta}$  es pequeño, entonces se puede menospreciar los términos cuadráticos y superiores en  $\dot{\theta}$ .

Reescribiendo el sistema a partir de estas suposiciones, tenemos

$$\begin{bmatrix} \dot{p} \\ \dot{\theta} \\ \ddot{p} \\ \ddot{\theta} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & m^2l^2g/\mu & -cJ_t/\mu & -\gamma J_tlm/\mu \\ 0 & M_tmgl/\mu & -clm/\mu & -\gamma M_t/\mu \end{bmatrix} \begin{bmatrix} p \\ \theta \\ \dot{p} \\ \dot{\theta} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ J_t/\mu \\ lm/\mu \end{bmatrix} u, \quad (44)$$

$$y = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} x,$$

donde  $\mu = M_tJ_t - m^2l^2$ . Además, considere  $c = \gamma = 0$ .

### Ejemplo

- El péndulo invertido es una variación del sistema mostrado en el ejemplo anterior. La diferencia radica en que la base  $p$  no necesita estar controlada y lo único que se desea es estabilizar la orientación vertical de la barra. Este sistema se puede apreciar en la siguiente Figura.

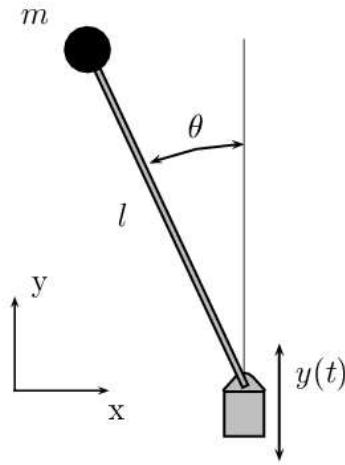


Fig. 11 Sistema de péndulo invertido.

El sistema de ecuaciones que describe la dinámica del péndulo simple está dado como sigue

$$\begin{bmatrix} \dot{\theta} \\ \ddot{\theta} \end{bmatrix} = \begin{bmatrix} \dot{\theta} \\ \frac{mgl}{J_t} \sin(\theta) - \frac{\gamma}{J_t} \dot{\theta} + \frac{l}{J_t} \cos(\theta) u \end{bmatrix}, \quad y = \theta,$$

donde  $\gamma$  es el coeficiente de fricción rotacional,  $J_t = J + ml^2$  y  $u$  es la fuerza aplicada a la base.

## Formas canónicas de controlabilidad

Considere el caso general representado en función de transferencia

$$G(s) = \frac{N(s)}{D(s)} = \frac{b_1 s^3 + b_2 s^2 + b_3 s + b_4}{s^4 + a_1 s^3 + a_2 s^2 + a_3 s + a_4}.$$

La forma de realización controlable del sistema anterior es

$$\begin{aligned} \dot{x} &= \begin{bmatrix} -a_1 & -a_2 & -a_3 & -a_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} u, \\ y &= [b_1 \ b_2 \ b_3 \ b_4] x, \end{aligned}$$

donde la matriz de controlabilidad puede ser obtenida como sigue

$$U = \begin{bmatrix} 1 & -a_1 & e_2 & e_3 \\ 0 & 1 & -a_1 & e_2 \\ 0 & 0 & 1 & -a_1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (45)$$

con  $e_2 := -a_2 + a_1^2$  y  $e_3 := -a_3 + 2a_1a_2 - a_1^3$ .

### i Note

La matriz (45) es una matriz triangular cuyo determinante siempre es 1. Por consiguiente, siempre es controlable y es la razón por la que se llama *forma controlable*.

## Estabilización por retroalimentación de estado

Supongamos que se desea controlar un sistema en representación de espacio de estados lineal. Por simplicidad consideramos que tiene una sola entrada y que la ley de control de retroalimentación parte de la idea de que el sistema en lazo cerrado tiene eigenvalores deseados.

El diagrama mostrado en la [Fig. 12](#) representa un sistema de control que usa retroalimentación de estado. Este sistema consta de un proceso representado por un modelo lineal, los elementos del control denotados por  $K$  y  $k_r$ , una señal de referencia  $r$  y perturbaciones  $d$ . El objetivo de control por retroalimentación es regular la salida del sistema.

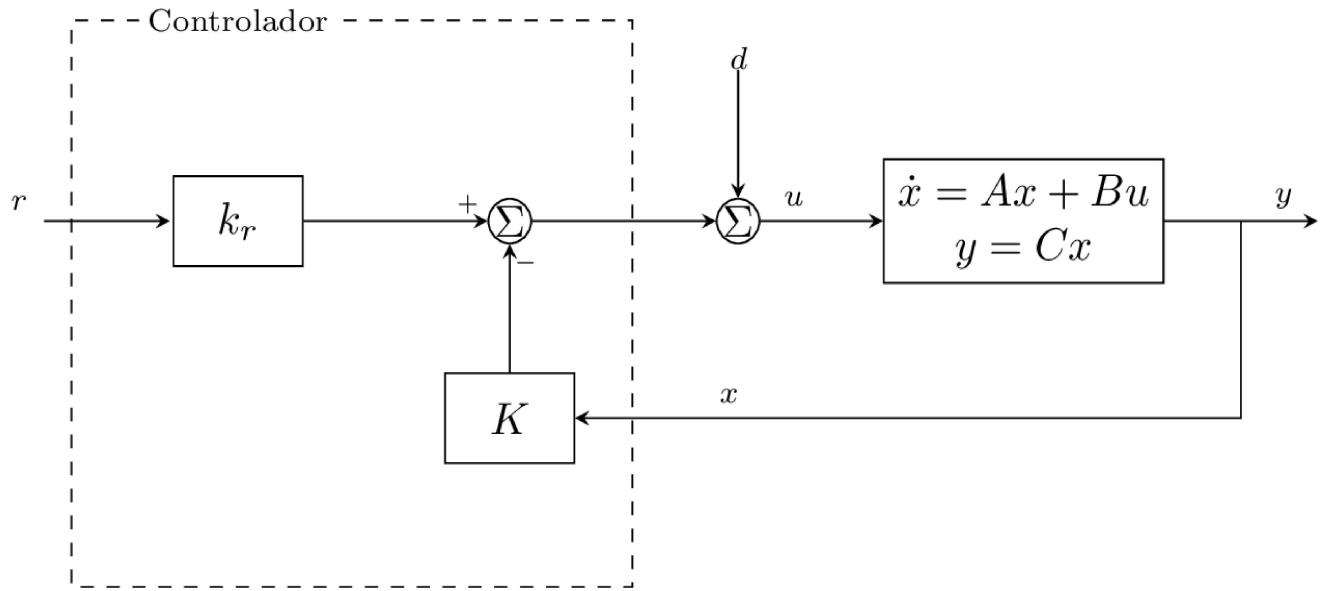


Fig. 12 Sistema de control con retroalimentación de estado donde el controlador utiliza el estado  $x$ , la entrada de referencia  $r$  y la entrada  $u$ . Aquí la perturbación está dada a través de  $d$ .

Entonces, considere el siguiente sistema  $n$ -dimensional en representación de espacio-estado

$$\begin{aligned} \dot{x} &= Ax + Bu, \\ y &= Cx + Du, \end{aligned} \tag{46}$$

con función de transferencia

$$G(s) = C(sI - A)^{-1}B.$$

El objetivo es llevar la salida  $y$  al valor de referencia  $r$  a partir de la medición de los estados  $x$ . Para ello, consideramos que la ley de control invariante en el tiempo definida como  $u$  es una función del estado y la referencia como se muestra a continuación

$$u = r - Kx, \tag{47}$$

donde  $K = [k_1 \ k_2 \ \cdots \ k_n]$  es un vector real de dimensión  $1 \times n$ .

La ecuación (47) es conocida como *retroalimentación de estado de ganancia constante* o simplemente *retroalimentación de estado* [[Åström and Murray, 2021](#)].

Sustituyendo (47) en (46), tenemos

$$\begin{aligned}\dot{x} &= Ax - BKx + Br, \\ &= (A - BK)x + Br, \\ y &= Cx,\end{aligned}$$

con función de transferencia

$$G_o(s) = C(sI - A + Bk)^{-1}B,$$

con polinomio característico

$$p(s) = s^n + a_1s^{n-1} + \cdots + a_{n-1}s + a_n. \quad (48)$$

El cálculo de la ganancia de retroalimentación  $K$  para que el sistema de lazo cerrado con polinomio (48) se conoce como *asignación de eigenvalores* o *localización de polos*.

Por otro lado, si consideramos el esquema de control mostrado en la [Fig. 12](#), tenemos la siguiente ley de control

$$u = -Kx + k_r r,$$

donde el sistema en lazo cerrado se expresa como sigue

$$\begin{aligned}\dot{x} &= (A - BK)x + Bk_r r, \\ y &= Cx.\end{aligned}$$

#### Note

Es evidente que  $k_r$  no afecta la estabilidad del sistema ya que esta está determinada por los eigenvalores del  $A - BK$ .

Los puntos de equilibrio del sistema en lazo cerrado están dadas por la siguiente expresión

$$x_e = -(A - BK)^{-1}Bk_r r, \quad y_e = Cx_e,$$

donde  $k_r$  es escogida tal que  $y_e = r$ . Además, dado que  $k_r$  es un escalar, podemos obtenerlo a partir de la siguiente expresión

$$k_r = -\frac{1}{C(A - BK)^{-1}B}$$

## Procedimiento para la asignación de eigenvalores

1. Calcule el polinomio característico de  $A : \Delta(\lambda) = \det(\lambda I - A)$ .
2. Calcule el polinomio característico deseado

$$\begin{aligned}\bar{\Delta}(\lambda) &= (\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3) \cdots (\lambda - \lambda_n), \\ &= \lambda^n + \bar{a}_1\lambda^{n-1} + \bar{a}_2\lambda^{n-2} + \cdots + \bar{a}_n.\end{aligned}$$

3. Calcule la ganancia de retroalimentación para la ecuación equivalente en la forma controlable

$$\bar{k} = [\bar{a}_1 - a_1 \quad \bar{a}_2 - a_2 \quad \cdots \quad \bar{a}_n - a_n].$$

4. Calcule la transformación equivalente

$$S := P^{-1} = [B \quad AB \quad A^2B \quad A^3B] \begin{bmatrix} 1 & a_1 & a_2 & \dots & a_{n-1} \\ 0 & 1 & a_1 & \dots & a_{n-2} \\ 0 & 0 & 1 & \dots & a_{n-3} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}.$$

5. Calcule la ganancia de retroalimentación  $k = \bar{k}P = \bar{k}S^{-1}$ .

### Note

Existen otras formas para obtener la ganancia de retroalimentación  $k$ , una de ellas es la *fórmula de Ackerman* dada por la siguiente ecuación

$$k = [0 \quad 0 \quad 0 \quad \cdots \quad 1][B \quad AB \quad A^2B \quad \cdots \quad A^nB]^{-1}\bar{\Delta}(A),$$

donde  $\bar{\Delta}(s)$  es el polinomio característico de  $(A - Bk)$ , por lo tanto  $\bar{\Delta}(A) \neq 0$ .

### Note

Si  $(A, B)$  es controlable, entonces los eigenvalores de  $(A - Bk)$  pueden ser asignados arbitrariamente escogiendo una ganancia de retroalimentación real  $k$ .

### Ejemplo

- Considere el siguiente sistema en representación de espacio de estados

$$\begin{aligned}\dot{x} &= \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}x + \begin{bmatrix} 0 \\ 10 \end{bmatrix}u, \\ y &= [1 \quad 0]x\end{aligned}$$

cuya función de transferencia es

$$G(s) = \frac{10}{s^2 + s}.$$

Encuentre la ganancia de retroalimentación  $k$  en  $u = r - kx$  tal que la ecuación resultante tenga sus eigenvalores en  $-2 \pm j2$ .

### Ejemplo

- Considere el problema de regular la población de un ecosistema mediante la modulación del suministro de alimentos. Para ello considere el modelo depredador-presa cuya dinámica está dada por las siguientes ecuaciones

$$\begin{aligned}\dot{H} &= (r + u)H \left(1 - \frac{H}{k}\right) - \frac{aHL}{c + H}, \quad H \geq 0, \\ \dot{L} &= b \frac{aHL}{c + H} - dL, \quad L \geq 0.\end{aligned}\tag{49}$$

En este sistema,  $r$  denota la tasa de crecimiento de las liebres,  $k$  la población máxima,  $a$  el término que describe la disminución de las liebres en función de la población,  $c$  representa la tasa de consumo de presas,  $b$  el coeficiente de crecimiento de los linces y finalmente  $d$  la tasa de mortalidad de los linces.

### Ejemplo

- Considere el problema del carro-péndulo dado por la Ec. [\(44\)](#)

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & m^2 l^2 g / \mu & -c J_t / \mu & -\gamma J_t l m / \mu \\ 0 & M_t m g l / \mu & -c l m / \mu & -\gamma M_t / \mu \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ J_t / \mu \\ l m / \mu \end{bmatrix}, \quad (50)$$

donde  $\mu = M_t J_t - m^2 l^2$  y  $\gamma = c \neq 0$ .

Para este ejercicio, utilice los parámetros presentados en la siguiente Tabla, demuestre que el sistema es controlable y que además, es posible utilizar la estrategia de *control por retroalimentación* de estado para estabilizar el sistema.

Variable	Valor	Unidades
M	10	kg
m	80	kg
c	0.1	N s/m
J	100	kg m <sup>2</sup> /s <sup>2</sup>
l	1	m
Y	0.01	N m s
g	9.81	m/s <sup>2</sup>

## Ejemplo

- Considere un vehículo de dos ruedas como se muestra en la [Fig. 13](#)

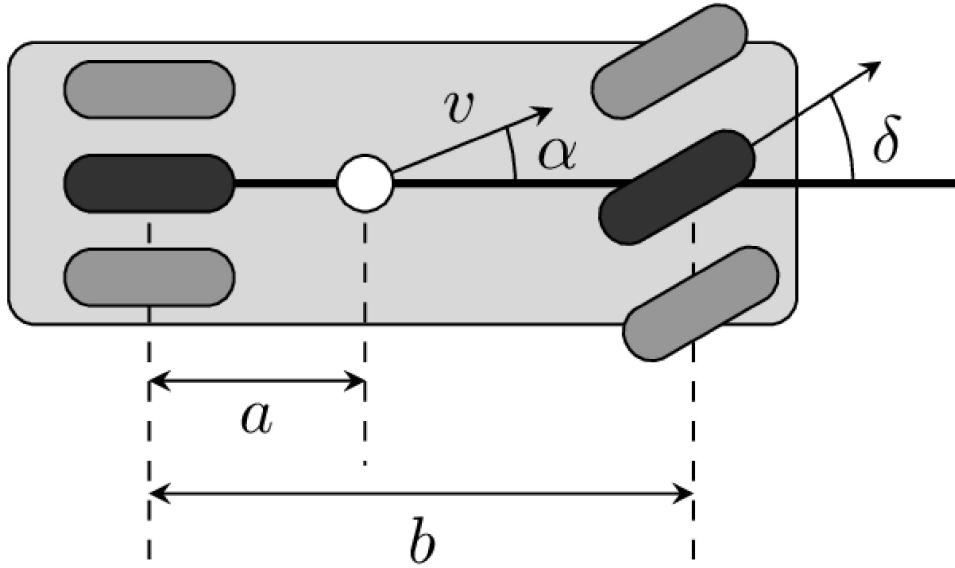


Fig. 13 Dirección del vehículo: modelo de bicicleta

Nuestro interés es analizar cómo la velocidad del vehículo depende del ángulo de dirección  $\delta$ . Para ello, consideramos que  $v$  representa la velocidad en el centro de la masa,  $a$  la distancia de la rueda trasera, y  $b$  la base de la rueda.

Sean  $x$  e  $y$  las coordenadas del centro de masa,  $\theta$  en ángulo de rumbo y  $\alpha$  el ángulo entre el vector de velocidad  $v$  y la línea central del vehículo, entonces la relación entre  $\alpha$  y el ángulo de dirección  $\delta$  está dada como sigue

$$\alpha(\delta) = \arctan\left(\frac{a \tan(\delta)}{b}\right),$$

donde  $a := r_a \tan(\alpha)$ ,  $b := r_a \tan(\delta)$ .

Asuma que las ruedas no patinan y que la velocidad de las ruedas traseras es  $v_0$ , entonces la velocidad del vehículo en su centro de masa es  $v = v_0 / \cos(\alpha)$ . Por consiguiente, el movimiento de este punto está dado como sigue

$$\begin{aligned} \dot{x} &= v \cos(\alpha + \theta) = v_0 \frac{\cos(\alpha + \theta)}{\cos(\alpha)}, \\ \dot{y} &= v \sin(\alpha + \theta) = v_0 \frac{\sin(\alpha + \theta)}{\cos(\alpha)}. \end{aligned} \tag{51}$$

El ángulo  $\theta$ , afectado por el ángulo de dirección está dado por la siguiente relación

$$\dot{\theta} = \frac{v_0}{r_a} = \frac{v_0}{b} \tan(\delta).$$

Asuma que  $\delta = 0$  y  $\dot{\theta} = 0$  para encontrar el punto de equilibrio del sistema y proceda a linealizar (51).

## Respuesta de estado estacionario

Considere un sistema entrada-salida lineal como se muestra en la siguiente expresión

$$\begin{aligned}\dot{x} &= Ax + Bu, \\ y &= Cx,\end{aligned}$$

donde su solución se puede obtener a partir de la ecuación de convolución

$$y(t) = Ce^{At}x_0 + \int_0^t Ce^{A(t-\tau)}Bu(\tau)d\tau. \quad (52)$$

Como se puede observar, la respuesta del sistema depende de una condición inicial  $x_0$  y una entrada  $u$ . Además, es posible observar que en esta expresión existen dos componentes: la *respuesta transitoria* y la *respuesta de estado estable*. La primera ocurre cuando se aplica una entrada y se observa un desajuste entre la condición inicial y la solución de estado estable. La segunda refleja el comportamiento del sistema bajo las entradas dadas. En la [Fig. 14](#) podemos ver estos dos componentes en respuesta ante una entrada de tipo escalón unitario.

### Note

En la práctica, se espera que si la entrada es periódica la respuesta también lo sea. Lo mismo con entradas constantes.

Un *escalón unitario*, *entrada escalón* o *escalón de Heaviside* es una función definida a pedazos como sigue

$$u(t) = \begin{cases} 0, & t = 0, \\ 1, & t > 0. \end{cases} \quad (53)$$

Definimos entonces como *respuesta escalonada* a la salida  $y(t)$  a partir de una condición inicial en el punto de equilibrio del sistema y una entrada  $u(t)$  de la forma (53).

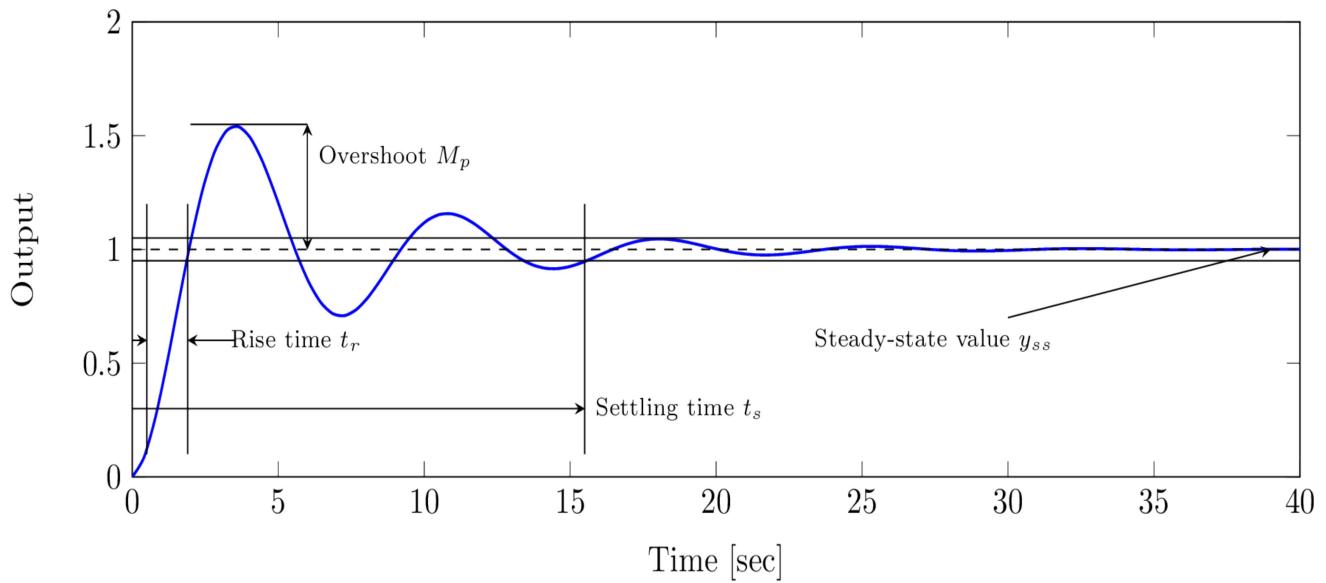


Fig. 14 Respuesta del sistema ante una entrada tipo escalón. Se observa el tiempo de subida (Rise time), sobretiro (Overshoot), tiempo de asentamiento (Settling time) y valor en estado estable (Steady-state value).

Utilizando la ecuación de convolución \eqref{eqn:conv\_eq}, podemos calcular la respuesta ante una entrada tipo escalón considerando  $x_0 = 0$ . Entonces, tenemos

$$\begin{aligned} y(t) &= \int_0^t C e^{A(t-\tau)} B u(\tau) d\tau = C \int_0^t e^{A(t-\tau)} B d\tau, \\ &= C \int_0^t e^{A\sigma} B d\sigma = C \left( A^{-1} e^{A\sigma} B \right) \Big|_{\sigma=0}^{\sigma=t}, \\ &= CA^{-1} e^{At} B - CA^{-1} B, \end{aligned}$$

o bien

$$y(t) = \underbrace{CA^{-1}e^{At}B}_{\text{transitorio}} - \underbrace{CA^{-1}B}_{\text{estado estable}}$$

# Velocidad de respuesta

Definimos como rendimiento transitorio a la velocidad de respuesta o la velocidad en la que el sistema alcanza al estado estable. Generalmente se especifica en términos de tiempo de levantamiento (*rise time*), tiempo de asentamiento (*settling time*) y sobretiro (*overshoot*). El *tiempo de levantamiento* lo definimos como el tiempo requerido para la respuesta pase de 0 al 90% del valor en estado estacionario como se muestra en la [Fig. 14](#). En otras palabras, buscamos el valor más pequeño  $t_r$  tal que

$$y(t_r) = 0.9y_{ss}$$

donde  $t_s$  denota el \myindex{tiempo de asentamiento}. Es decir, el tiempo que le toma a la respuesta del sistema alcanzar y mantenerse dentro del  $\pm 2\%$  de su valor en estado estable, o bien, el valor más pequeño  $t_s$  tal que

$$|y - y_{ss}| \leq 0.02y_{ss}, \quad \forall t \geq t_{ss}.$$

Sea  $y_{\max}$  el valor máximo de  $|y(t)|$ ,  $\forall t \geq 0$  o bien

$$y_{\max} := \max |y(t)|,$$

entonces el *sobretiro* se define como sigue

$$M_p := \frac{y_{\max} - y_{ss}}{y_{ss}} \times 100\%.$$

# Observabilidad de sistemas dinámicos

## Contents

- Observabilidad de sistemas dinámicos
- Teorema de dualidad
- Procedimiento para obtener la ganancia del observador via Lyapunov

## Concepto de observabilidad

Considere un sistema  $n$ -dimensional en representación de espacio-estado como sigue

$$\begin{aligned}\dot{x} &= Ax + Bu, \\ y &= Cx + Du,\end{aligned}\tag{54}$$

donde  $A, B, C$  son matrices de dimensión  $n \times n$ ,  $n \times 1$ , y  $1 \times n$ , respectivamente. Además,  $D$  es escalar.

Supongamos que queremos estimar el estado de un sistema dinámico a partir de sus entradas  $u$  y salidas  $y$  como se aprecia en la [Fig. 15](#).

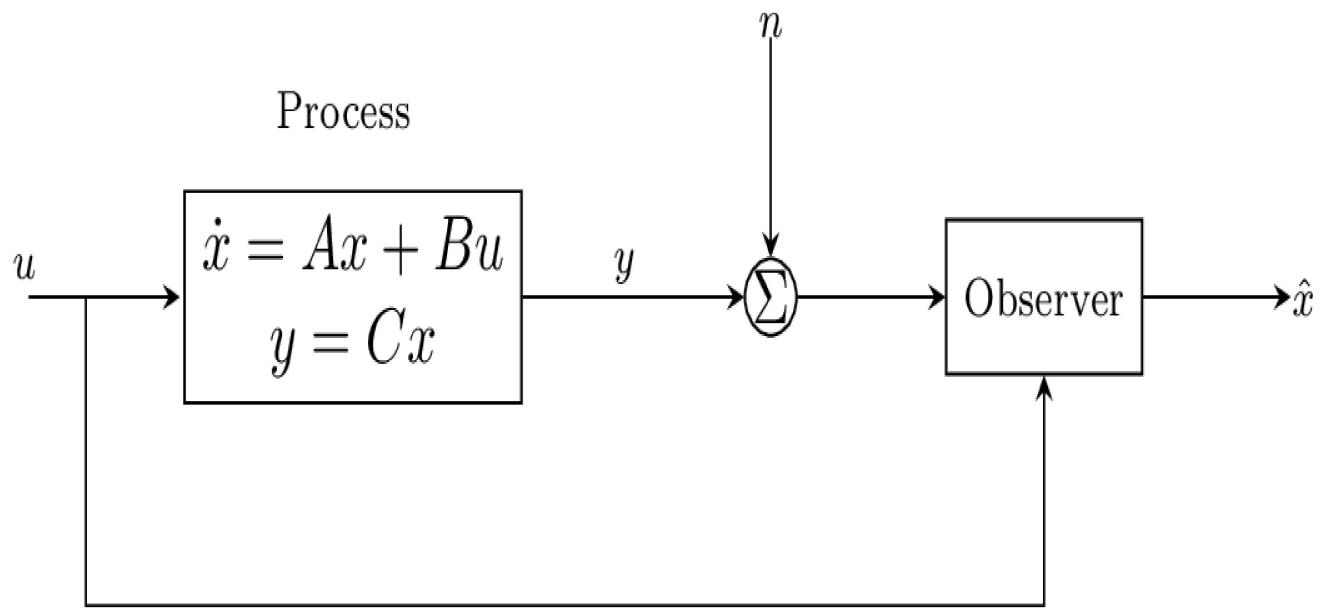


Fig. 15 Diagrama a bloques de un observador de estados. El observador utiliza la señal medida  $y$  y la entrada  $u$  para estimar el proceso denotado como  $\hat{x}$ .

Para ello, considere que tenemos sólo una señal medida  $y$  que pueda ser afectada por ruido  $n$ . Aquí, denotamos  $\hat{x}$  como el estado estimado dado por el observador.

### Definition 7

Un sistema LTI es *observable* si es posible determinar el estado de un sistema  $x$  a través de las mediciones de  $y$  y  $u$  sobre el intervalo  $[0, t]$ , para cualquier  $t > 0$ .

Entonces, definimos la matriz  $n \times n$

$$V = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix},$$

entonces el sistema (54) es observable si y sólo si la matriz  $V$  tiene rango  $n$  o bien,  $\det\{V\} \neq 0$ .

### **i Note**

La *controlabilidad* es una propiedad de  $(A, B)$  mientras que *observabilidad* es una propiedad de  $(A, C)$ .

### **Ejemplo**

- Considere la siguiente función de transferencia

$$G(s) = \frac{N(s)}{D(s)} = \frac{2s - 1}{s^2 - 1.5s - 1} = \frac{2s - 1}{(s - 2)(s + \frac{1}{2})},$$

y determine si el sistema es observable.

### **Ejemplo**

- Considere la siguiente función de transferencia

$$G(s) = \frac{N(s)}{D(s)} = \frac{2s + 1}{s^2 - 1.5s - 1} = \frac{2s + 1}{(s - 2)(s + \frac{1}{2})},$$

y determine si el sistema es observable.

### **Ejemplo**

- Considere un modelo de dos compartimientos como se muestra en la [Fig. 16](#)

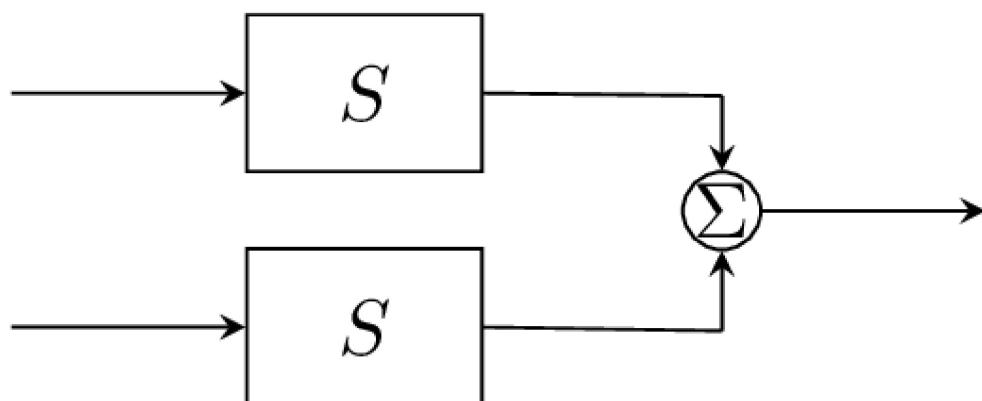


Fig. 16 Modelo compartmental.

Asuma que la concentración del primer compartimiento puede ser medido y que el sistema está descrito por el siguiente sistema lineal

$$\begin{aligned}\dot{c} &= \begin{bmatrix} -k_0 - k_1 & k_1 \\ k_2 & -k_2 \end{bmatrix} c + \begin{bmatrix} b_0 \\ 0 \end{bmatrix} u, \\ y &= [1 \ 0]x.\end{aligned}$$

El primer compartimiento corresponde a la concentración de fármaco en el plasma sanguíneo mientras que el segundo representa la concentración del fármaco en el tejido donde es activo.

Determine si el sistema es observable.

## Formas canónicas de observabilidad

Un sistema *SISO*, está en la forma canónica observable si sus dinámicas están dadas como sigue

$$\begin{aligned}\dot{z} &= \begin{bmatrix} -a_1 & 1 & 0 & \cdots & 0 \\ -a_2 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \\ -a_{n-1} & 0 & 0 & \cdots & 1 \\ -a_n & 0 & 0 & \cdots & 0 \end{bmatrix} z + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{bmatrix} u, \\ y &= [1 \ 0 \ 0 \ \cdots \ 0]z.\end{aligned}$$

Si la entrada  $u$  y la salida  $y$  están disponibles, entonces el estado  $z_1$  puede ser estimado.

## Estimación de estados

Considere el siguiente sistema  $n$ -dimensional en representación de espacio-estado

$$\begin{aligned}\dot{x} &= Ax + Bu, \\ y &= Cx.\end{aligned}\tag{55}$$

Una estimación de los estados del sistema se puede obtener simulando las ecuaciones (55) dada una entrada adecuada. Por lo tanto, el estimador de estados está dado por la siguiente expresión

$$\dot{\hat{x}} = A\hat{x} + Bu.$$

Para determinar las propiedades de esta estimación, es necesario definir el error de estimación  $\tilde{x} := x - \hat{x}$ , de modo que el sistema se reescribe como sigue

$$\dot{\tilde{x}} = A\tilde{x}.$$

 **Note**

Si todos los eigenvalores de la matriz  $A$  se encuentran en el semi plano izquierdo, entonces el error  $\tilde{x} \rightarrow 0$ .

Considere el siguiente observador

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x}). \quad (56)$$

Podemos observar que la retroalimentación de la salida se logra agregando el término  $L(y - C\hat{x})$  la cual es proporcional a la diferencia entre la salida observada y la salida estimada por el sistema (56). Por lo tanto

$$\dot{\tilde{x}} = (A - LC)\tilde{x}.$$

Si la matriz  $L$  se escoge de manera que la matriz  $A - LC$  tiene eigenvalores con parte real negativa, entonces el error  $\tilde{x} \rightarrow 0$ .

 **Note**

La tasa de convergencia está determinada por una adecuada selección de los eigenvalores.

# Diseño de observador via asignación de eigenvalores

## Theorem 10

Considere el siguiente sistema SISO

$$\begin{aligned}\dot{x} &= Ax + Bu, \\ y &= Cx.\end{aligned}\tag{57}$$

Sea  $\lambda_A(s) = s^n + a_1s^{n-1} + \cdots + a_{n-1}s + a_n$  el polinomio característico para  $A$ . Si el sistema es observable, entonces el sistema dinámico

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x}),\tag{58}$$

es un observador para el sistema, con  $L$  elegida como sigue

$$L = V^{-1}\tilde{V} \begin{bmatrix} p_1 - a_1 \\ p_2 - a_2 \\ \vdots \\ p_n - a_n \end{bmatrix},$$

donde las matrices  $V$  y  $\tilde{V}$  están dadas por

$$V = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}, \quad \tilde{V} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ a_1 & 1 & 0 & \cdots & 0 & 0 \\ a_2 & a_1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-2} & a_{n-3} & a_{n-4} & \cdots & 1 & 0 \\ a_{n-1} & a_{n-2} & a_{n-3} & \cdots & a_1 & 1 \end{bmatrix}.$$

El error del observador  $\tilde{x} = x - \hat{x}$  está dado por una ecuación diferencial que tiene como polinomio característico

$$p(s) = s^n + p_1s^{n-1} + \cdots + p_n.$$

El sistema dado en la Ec. (58) es llamado *observador de estados* del sistema (57) y su diagrama a bloques se muestra en la Fig. 17.

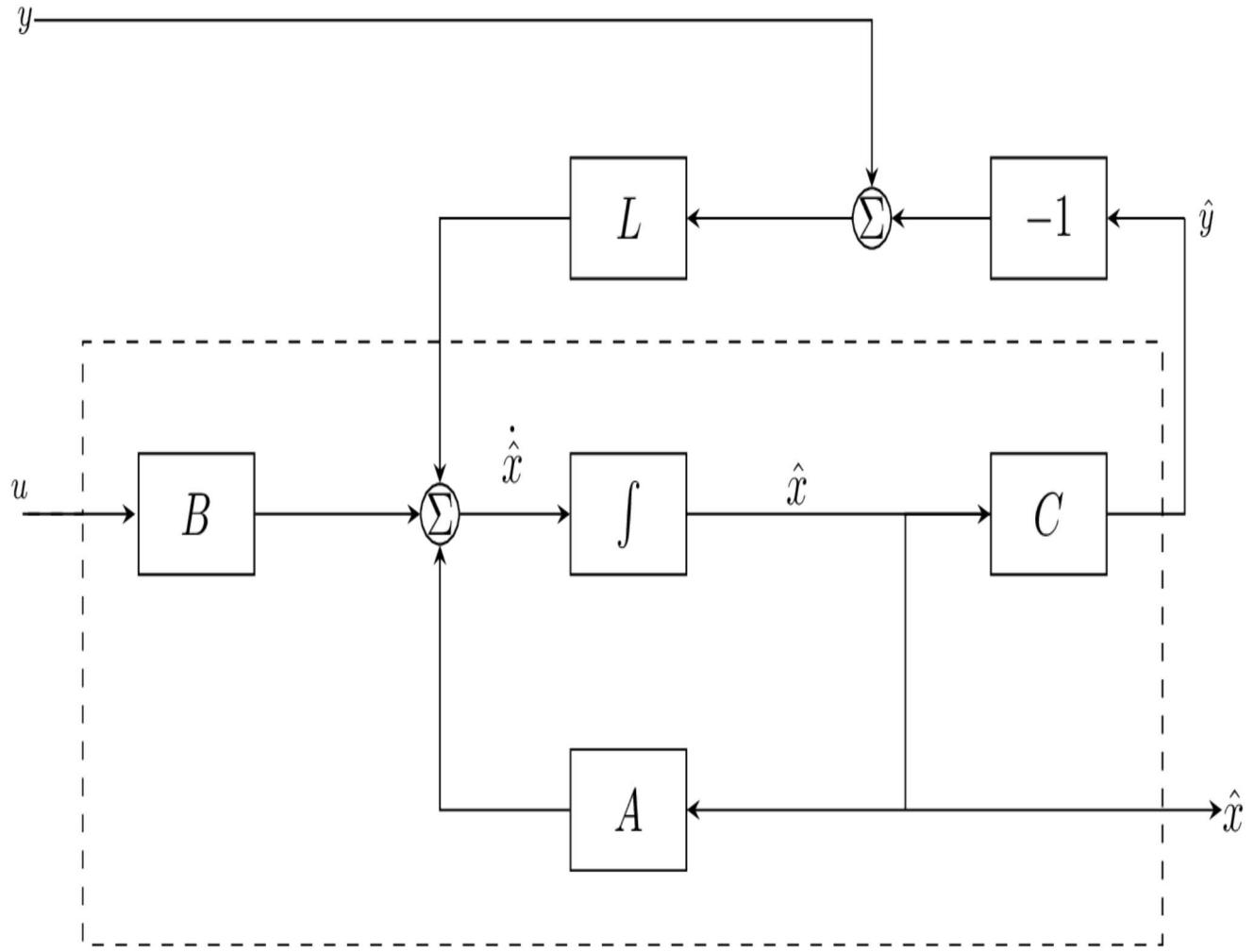


Fig. 17 Diagrama a bloques del observador de estados. Para producir una estimación  $\hat{x}$  se consideran las entradas  $u$  e  $y$  como entradas y  $L$  como la ganancia del observador.

### Ejemplo

- Considere un modelo de dos compartimientos como se muestra en la Fig. 16

$$\begin{aligned}\dot{c} &= \begin{bmatrix} -k_0 - k_1 & k_1 \\ k_2 & -k_2 \end{bmatrix} c + \begin{bmatrix} b_0 \\ 0 \end{bmatrix} u, \\ y &= [1 \quad 0] x,\end{aligned}$$

y diseñe un observador de estados.

# Cálculo de la ganancia del observador

## Ejemplo

- Considere el siguiente modelo lineal que representa la dirección de un vehículo, considerando la desviación de trayectoria lateral  $y$  con el ángulo de dirección  $u$

$$\begin{aligned}\dot{x} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}x + \begin{bmatrix} \gamma \\ 1 \end{bmatrix}u, \\ y &= [1 \quad 0]x.\end{aligned}$$

Diseñe un observador de estados y calcule su ganancia.

### Note

El estado  $x_1$  representa la desviación de la trayectoria lateral y  $x_2$  la velocidad de giro.

## Ejemplo

- Considere el siguiente modelo SISO en representación de espacio de estados

$$A = \begin{bmatrix} -1 & -0.75 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad C = [1 \quad 1]$$

Sea el sistema  $\Sigma_1$  representado por las siguientes ecuaciones

$$\Sigma_1 := \begin{cases} \dot{x} = Ax + Bu, \\ y = Cx, \end{cases}$$

y el sistema dual  $\Sigma_2$  definido como

$$\Sigma_2 := \begin{cases} \dot{z} = A^*z + C^*v, \\ n = B^*z, \end{cases}$$

donde  $z$  es el vector de estado,  $v$  el vector de control,  $n$  el vector de salida y  $*$  denota la transpuesta conjugada.

El *teorema de dualidad* menciona que el sistema  $\Sigma_1$  es completamente *controlable (observable)* si y sólo si el sistema  $\Sigma_2$  es completamente *observable (controlable)*.

# Condiciones necesarias

Para el sistema  $\Sigma_1$

- Condición de controlabilidad

$$\text{rank} \{ [B \ AB \ \dots \ A^{n-1}B] \} = n.$$

- Condición de observabilidad

$$\text{rank} \{ [C^* \ A^*C^* \ \dots \ (A^*)^{n-1}C^*] \} = n.$$

Para el sistema  $\Sigma_2$

- Condición de controlabilidad

$$\text{rank} \{ [C^* \ A^*C^* \ \dots \ (A^*)^{n-1}C^*] \} = n.$$

- Condición de observabilidad

$$\text{rank} \{ [B \ AB \ \dots \ A^{n-1}B] \} = n.$$

La ganancia  $L$  del observador de estados \eqref{eqn:observer}, se pude obtener a partir de la siguiente expresión [[Coronel-Escamilla et al., 2017](#)]

$$L = S^{-1}C^T,$$

donde  $S$  es la solución única de la *ecuación algebráica de Lyapunov*

$$-\lambda S - A^T S - SA + C^T C = 0, \tag{59}$$

donde  $\lambda$  denota el parámetro para sintonizar la convergencia del observador [\(56\)](#).

El siguiente fragmento de código permite encontrar la solución a la Ec. [\(59\)](#)

```

def lyapunov(lamb,N):
    S = np.zeros((N,N))

    for k in range(1,N+1):
        for l in range(1,N+1):
            c = nchoosek(l+k-2,k-1)
            S[l-1][k-1] = ((-1)**(l+k)*c)/lamb**(l+k-1)

    return S

```

donde la función nchoosek está dada como sigue

$${}_nC_k = \binom{n}{k} = \frac{n!}{(n-k)!k!},$$

y su representación en Python como

```

def nchoosek(n, k):
    return int(factorial(n)/(factorial(k)*factorial(n-k)))

```

## Control utilizando estimadores de estado

Considere un sistema en representación de espacio de estados de la forma

$$\begin{aligned}\dot{x} &= Ax + Bu, \\ y &= Cx.\end{aligned}$$

Este sistema sólo tiene una salida disponible y es necesario diseñar un *control por retroalimentación*. Para ello, asumimos que el sistema es controlable así como observable y que además,  $u$  e  $y$  son escalares.

Considerando la siguiente ley de control (retroalimentación), tenemos

$$u = -Kx + k_r r.$$

Bajo la suposición de que todos los estados son medibles, proponemos el diseño de un observador para estimar el estado  $\hat{x}$  a partir de las entradas y las salidas.

Si todos los estados no están disponibles o bien, no son medibles, entonces proponemos la siguiente *ley de control*

$$u = -K\hat{x} + k_r r,$$

donde  $\hat{x}$  es la salida del observador de estados

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x}).$$

Para analizar el sistema en lazo cerrado, definimos la siguiente variable

$$\tilde{x} := x - \hat{x},$$

y la reemplazamos en la variable de estado  $\hat{x}$ . Con esto obtenemos la siguiente expresión

$$\dot{x} = (A - BK)x + BK\tilde{x} + Bk_r r,$$

o en representación de espacio de estados como sigue

$$\begin{bmatrix} \dot{x} \\ \dot{\tilde{x}} \end{bmatrix} = \begin{bmatrix} A - BK & BK \\ 0 & A - LC \end{bmatrix} \begin{bmatrix} x \\ \tilde{x} \end{bmatrix} + \begin{bmatrix} Bk_r \\ 0 \end{bmatrix} r.$$

Por consiguiente, el polinomio característico del sistema está dado por la siguiente ecuación

$$\phi(\lambda) = |\lambda I - A + BK| |\lambda I - A + LC|$$

Un diagrama de un sistema de control basado en observador se presenta en la [Fig. 18](#).

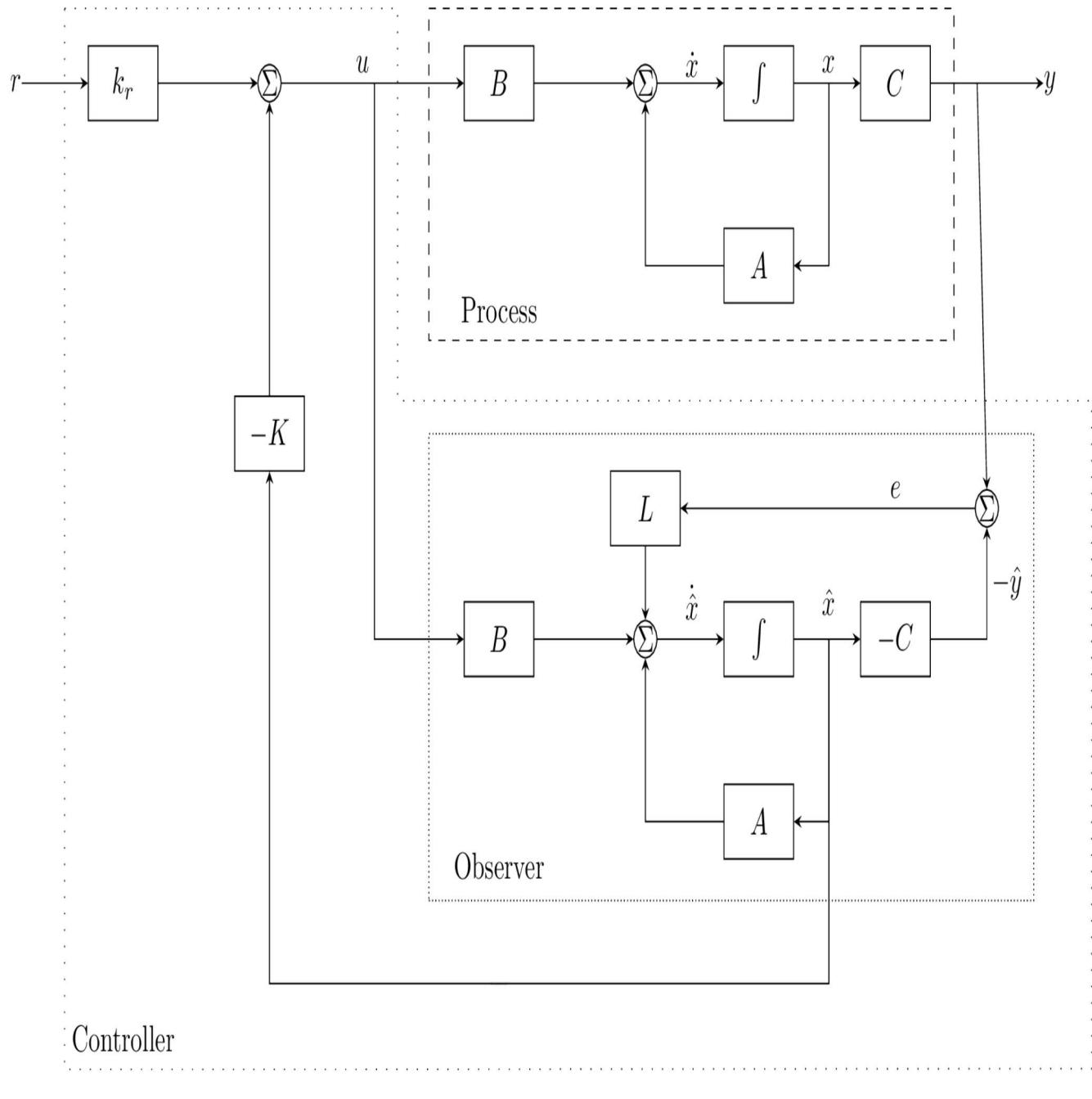


Fig. 18 Diagrama a bloques de un sistema de control basado en observador. El observador utiliza la salida estimada  $\hat{y}$  y la entrada  $u$  para obtener una aproximación de los estados  $x$ . Esta estimación se utiliza por un sistema de control por retroalimentación de estados para corregir la entrada.

# Identificación de sistemas dinámicos

## Contents

- Objetivo del problema de identificación.
- Variables aleatorias
- Representación de sistemas dinámicos en forma lineal con respecto a los parámetros.

## Objetivo del problema de identificación.

El objetivo de la Identificación de Sistemas (IS) consiste en construir modelos matemáticos de sistemas dinámicos a partir de datos experimentales. En la actualidad, la IS es importante en áreas como control automático, procesamiento de señales, física, economía, medicina, ecología, sismología, biología, entre otros.

Gauss y Legendre fueron los primeros en introducir el método de mínimos cuadrados (LS) para predecir el movimiento de los planetas y cometas a partir de mediciones de telescopios. En consecuencia, diversos trabajos se realizaron sobre problemas de estimación de parámetros.

Existen diferentes clases de modelos. Por ejemplo, modelos lineales determinísticos en tiempo discreto en representación de espacio de estados así como de la forma entrada-salida. También hay representaciones de sistemas a partir de modelos lineales estocásticos de tiempo discreto. Estos modelos pueden ser vistos como sigue:

- Filtros dinámicos lineales que permitan la generación, análisis y clasificación de señales aleatorias. Por ejemplo, autorregresivas (AR), modelos promedio (MA) y modelos ARMA.
- Modelos lineales con ruido aditivo aleatorio que son capaces de representar ruido, perturbaciones externas y errores de modelado. Por ejemplo, ARX, ARMAX y modelos ARARX.

En la práctica, es común encontrar dos clases de modelos para representar sistemas del mundo real; los modelos no lineales (NL) y los modelos lineales de parámetros variables (LPV).

Los modelos LPV son ideales para modelar sistemas lineales de tiempo variable (LTV) así como representar sistemas no lineales linealizados sobre una trayectoria  $p(t)$ . Este tipo de modelos pueden ser vistos como descripciones intermedias entre modelos lineales invariantes en el tiempo (LTI) y modelos no lineales variantes en el tiempo.

Por otro lado, los modelos NL son muy utilizados para describir fenómenos o procesos complejos por ejemplo: procesos bioquímicos en columnas de destilación, plantas hidráulicas, sistemas fisiológicos, vibraciones en estructuras, por mencionar algunos.

Los modelos NL orientados a bloques están compuestos por subsistemas dinámicos LTI y subsistemas estáticos NL donde la parte lineal son generalmente representaciones en funciones de transferencia, espacio de estados o I/O mientras que la parte NL pueden ser con memoria o sin memoria.

El proceso para realizar la identificación de sistemas, según Ljung [1998] consisten en seis pasos:

1. **Diseño del experimento.** Elección de la señal de excitación, periodo de muestreo, sensores para las señales de entrada y salida.
2. **Mediciones de entrada y salida.**
3. **Elección de la estructura del modelo.**
4. **Determinación de la estructura de los parámetros.** Criterio de información de Akaike [1974] (AIC) y el criterio de longitud de descripción mínima de Rissanen [1978].
5. **Estimación paramétrica del modelo.**
6. **Validación del modelo.** Medir el desempeño del modelo obtenido para representar los datos experimentales.

La elección de los algoritmos para la estimación de parámetros depende de dos factores:

1. La función costo a ser minimizada.
2. El algoritmo de optimización para encontrar la solución óptima.

De los métodos más utilizados para la estimación paramétrica podemos encontrar:

- Método de mínimos cuadrados ponderado (WLS).
  - Estimador Gauss-Markov.

- Mejor estimador lineal sin sesgo (BLUE).
- Método de mínimos cuadrados generalizado (GLS).
- Método de mínimos cuadrados extendido (ELS).
- Método de mínimos cuadrados total (TLS).
- Método de máxima verosimilitud (ML).
- Método de máximo a posteriori (MAP).
- Método de estimación de la mínima de la media del error al cuadrado (MMSE).
- Método de estimación M de Huber.
- Método de variable instrumental (IV).
- Método de subespacio como el algoritmo MUSIC (MULTiple SIgnal Classification).

## Variables aleatorias

Asuma que el valor verdadero de una cantidad de datos es 20, donde el conjunto de datos está dado como sigue:

1	2	3	4	5	6	7
20.13443	19.83828	20.01702	19.99835	19.94526	20.01415	19.96707

En términos estadísticos, se puede decir que los datos anteriormente mostrados son generados por una *variable aleatoria X* donde

 **Definition 8 (Variable aleatoria)**

Una variable aleatoria es una variable que tiene un valor numérico único, determinado por casualidad, para cada resultado de un procedimiento.

Suponga un sistema simple como un dado en el cual necesitamos conocer la probabilidad de obtener un "3" al lanzar el dado. La probabilidad de obtener este valor es 1/6 o bien, 16.7%. En estadística, podemos escribir esto como

$$P(X_1 = 3) = \frac{1}{6},$$

donde  $P$  es una función de probabilidad que representa la probabilidad del "evento"  $X_1 = 3$  como un número entre 0 y 1. Para tal efecto, definimos

#### 🔔 Definition 9 (Eventos y espacio muestral)

1. Un *evento* es cualquier colección de resultados o resultados de un procedimiento.
2. Un *evento simple* es un resultado o un evento que no se puede dividir en componentes más simples.
3. El *espacio muestral* para un procedimiento se conforma de todos los eventos simples posibles.

Por ejemplo, el espacio muestral para el ejemplo del dado está dado por el siguiente conjunto

$$S = \{1, 2, 3, 4, 5, 6\}$$

y todos los subconjuntos  $A \subset S$  como  $A_1 = \{1, 2\}$  o  $A_2 = \{1, 3, 5\}$  son eventos que pertenecen a un espacio muestral.

La función de probabilidad  $P$  parte de su definición en axiomas. Sin embargo, podemos definir probabilidad como sigue:

#### 🔔 Proposition 1 (Probabilidad)

Dado un espacio muestral  $S$ , la *función de probabilidad*  $P$  que asigna a cada evento  $A \subset S$  un número  $P(A) \in [0, 1]$ , llamada \*probabilidad\* del evento  $A$ , que dará una medida precisa de la posibilidad de que ocurra  $A$ .

#### 🔔 Proposition 2 (Enfoque clásico de probabilidad)

Un procedimiento dado tiene  $n$  eventos simples diferentes y que cada uno de esos eventos simples tiene las mismas posibilidades de ocurrir. Si el evento  $A$  puede ocurrir de estas  $n$  maneras, entonces

$$P(A) = \frac{s}{n}.$$

La aplica para el ejemplo del dado así como para otras *variables discretas aleatorias* similares que involucran un número finito de resultados posibles igualmente probables. No obstante, no aplica para *variables aleatorias continuas* con un número infinito de posibles resultados similares a la variable aleatoria  $X_2$ .

Para el espacio muestral dado por el conjunto

$$S = \{x \in \mathbb{R} | 0 \leq x < 15\}$$

que involucra un número infinito de posibles resultados continuamente distribuidos entre 0 y 15, se puede utilizar la siguiente expresión

 **Proposition 3 (Aproximación de frecuencia relativa)**

Asuma que dado que un procedimiento es repetido  $n$  veces, y  $f_n(A)$  denota la frecuencia relativa con la cual un evento  $A$  ocurre, entonces

$$P(A) = \lim_{n \rightarrow \infty} f_n(A)$$

## Representación de sistemas dinámicos en forma lineal con respecto a los parámetros.

En los modelos de regresión (MR) se involucra el análisis de una variable dependiente en términos de una o más variables independientes. En estos modelos, los parámetros son ajustados para describir a los datos experimentales. Algunos autores consideran que los modelos de regresión son sistemas entrada-salida ya que contienen los elementos característicos de estos sistemas. Es decir, una entrada  $x$ , una salida  $y$ , un sistema tipo caja negra que transforma  $x$  en  $y$ .

## Modelo de regresión lineal

Suponga que se tiene un conjunto de datos  $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$  ( $x_i, y_i \in \mathbb{R}, i = 1, \dots, m, m \in \mathbb{N}$ ), entonces es posible describir datos experimentales utilizando una *función de regresión* o *función de modelo* de la forma

$$\hat{y}(x) = \beta_0 x + \beta_1, \quad (60)$$

donde los coeficientes  $\beta_0$  y  $\beta_1$  son llamados *coeficientes de regresión* o *parámetros* del modelo de regresión. Además,  $x$  es usualmente llamada *variable explicativa*, *variable predictora* o *variable independiente*, mientras que  $\hat{y}$  es conocida como *variable dependiente* o *variable respuesta*.

 **Note**

La expresión dada en la Ec. (60) es llamada *función de regresión lineal* debido a que los coeficientes  $\beta_0, \beta_1$  dependen linealmente de esta función.

Se dice que el modelo (60) ajusta los datos experimentales si la diferencia  $y_i - \hat{y}(x_i)$  ( $i = 1, \dots, m$ ) es pequeña. Para lograr este objetivo, es necesario definir la siguiente expresión

$$\text{RSQ} = \sum_{i=1}^m (y_i - \hat{y}(x_i))^2.$$

Esta expresión es llamada *suma residual de cuadrados (RSQ)*. RSQ mide la distancia entre los datos experimentales y el modelo; en un problema de regresión estos parámetros pueden ser encontrados planteando un problema de optimización

$$\min_{\beta_0, \beta_1 \in \mathbb{R}} \{\text{RSQ}\}.$$

La solución de este problema de optimización se puede obtener utilizando el procedimiento de minimización de una función de varias variables visto en cálculo de varias variables. Entonces, obteniendo las derivadas parciales de la función  $\text{RSQ}(\beta_0, \beta_1)$  con respecto  $\beta_0$  y  $\beta_1$  a cero.

$$\beta_0 = \frac{\sum_{i=1}^m x_i y_i - m \bar{x} \bar{y}}{\sum_{i=1}^m x_i^2 - m \bar{x}^2},$$

$$\beta_1 = \bar{y} - \beta_0 \bar{x}$$

### **i** Note

RSQ es utilizado ya que realiza *estimaciones de máxima verosimilitud* de los parámetros del modelo,  $\beta_0$  y  $\beta_1$ .

## Ejemplo

Considere los siguientes datos experimentales y encuentre un modelo de regresión lineal.

	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<i>x</i>	10	20	30	40	50
<i>y</i>	3	5	11	12	16

### **i** Note

Las funciones de regresión pueden ser utilizadas para predecir valores de la variable de respuesta a partir de las variables explicativas. Las buenas predicciones se alcanzan sólo si las funciones de regresión se ajustan lo suficientemente bien a los datos experimentales.

## Coeficiente de determinación

El coeficiente de determinación mide la calidad de ajuste entre los datos experimentales y el modelo. Dicho valor se encuentra en una escala entre 0 y 100%, donde 0% representa un ajuste pobre mientras que el 100% implica un ajuste perfecto entre el modelo y los datos. El coeficiente de determinación está definido como

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2},$$

donde  $\hat{y}_i = \hat{y}(x_i)$ . Para el caso de los modelos de regresión lineal, esto se puede escribir como

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2},$$

### Note

Si se obtienen valores negativos de  $R^2$ , entonces el modelo funciona peor que un modelo que produciría el valor medio.

## Modelo de regresión no lineal

Debido a que algunos datos experimentales siguen algún patrón curvo no lineal que no se puede describir utilizando un modelo lineal, una alternativa es utilizar una *función de regresión polinomial* de la forma

$$\hat{y}(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_s x^s. \quad (61)$$

Si bien la Ec. (61) representa una función que puede ser altamente no lineal, si  $x$  es dado y además  $\hat{y}$  es calculado como una combinación lineal de los coeficientes de regresión  $\beta_0, \beta_1, \dots, \beta_s$ . Por lo tanto, todas las funciones de regresión se pueden llevar a la forma

$$\hat{y}(x) = \beta_0 + \beta_1 f_1(x) + \beta_2 f_2(x) + \cdots + \beta_s f_s(x).$$

## Regresión lineal múltiple

Suponga un problema en el que se desea predecir la variable de interés  $y$  a partir de diversas variables independientes  $x_1, x_2, \dots, x_n$  ( $n \in \mathbb{N}$ ). A este problema se le conoce como *problema de regresión lineal múltiple (MRLM)*. La expresión que representa al MRLM está dada como sigue

### Definition 10 (Regresión múltiple)

Las funciones de regresión múltiple  $\hat{y}(\mathbf{x})$  calculan una variable de respuesta  $y$  a partir de variables explicativas  $\mathbf{x} = (x_1, \dots, x_n)$  ( $n > 1$ ) y los coeficientes de regresión  $\beta_0, \beta_1, \dots, \beta_s$ . Si  $\hat{y}(\mathbf{x})$  depende linealmente de  $\beta_0, \beta_1, \dots, \beta_s$  entonces los datos experimentales pueden ser ajustados utilizando un modelo de regresión lineal múltiple.

Entonces, la forma general de un modelo de regresión lineal múltiple que involucra un número arbitrario de  $n \in \mathbb{N}$  variables explicatorias es

$$\hat{y}(\mathbf{x}) = \beta_0 + \beta_1 f_1(\mathbf{x}) + \beta_2 f_2(\mathbf{x}) + \cdots + \beta_s f_s(\mathbf{x}),$$

donde  $\mathbf{x} = (x_1, x_2, \dots, x_n)^t$  y  $f_i$  son funciones reales arbitrarias.

## Redes Neuronales Artificiales

En los casos en los que la función de regresión no puede obtenerse de la teoría o en la que es difícil obtener una representación gráfica de los datos se pueden utilizar las redes neuronales artificiales (ANN). Este tipo de modelos pueden ser considerados como funciones de regresión no lineales con un gran número de parámetros a ajustar para aproximar a cualquier función suave.

## Redes neuronales de alimentación hacia adelante

Este tipo de redes involucra una capa de entrada, una capa de salida y *capas ocultas* entre las capas de entrada y salida. En estos modelos se asume que la información viaja de izquierda a derecha y por esta razón reciben este nombre. Además, este tipo de representaciones matemáticas son suficientemente complejas para aproximar funciones suaves arbitrarias.

Asuma que se tienen  $n \in \mathbb{N}$  nodos de entrada que corresponden a las cantidades de entrada  $x_1, \dots, x_n$ , con  $H \in \mathbb{N}$  nodos ocultos y  $m \in \mathbb{N}$  nodos de salida que corresponden a las cantidades de salida  $y_1, \dots, y_m$ , si se multiplica cada entrada con una constante y se realiza la suma de todas las entradas y se agrega una constante se obtiene la siguiente expresión

$$\sum_{k=1}^n w_{ik;h1}x_k + b_{h1}, \quad (62)$$

donde los llamados *pesos*  $w_{ik;h1}$  representan el coeficiente real utilizado por el nodo oculto para multiplicar a la  $k$ -ésima entrada ( $ik$ ) mientras que los sesgos o "bias"  $b_{h1}$  son sumados por el nodo oculto.

Para evitar que la expresión (62) sea considerada como una forma compleja de un modelo de regresión lineal múltiple es necesario que en el nodo oculto exista una función real no lineal llamada *función de activación*

$$\varphi_h \left( \sum_{k=1}^n w_{ik;h1}x_k + b_{h1} \right). \quad (63)$$

La función de activación más utilizada es la función logística

$$\varphi(\xi) = \frac{e^\xi}{1 + e^\xi}, \quad (64)$$

donde el estado de los nodos ocultos  $l = 1, \dots, H$  después del procesamiento de las entradas se puede representar como sigue

$$\varphi_h \left( \sum_{k=1}^n w_{ik;hl} \cdot x_k + b_{hl} \right), \quad l = 1, \dots, H. \quad (65)$$

Asumiendo que la capa de salida procesa esta entrada de la misma manera que la capa oculta utilizando diferentes coeficientes y una función no lineal diferente, es posible obtener una salida a partir de

$$y_j = \varphi_o \left( b_{oj} + \sum_{k=1}^n w_{ik,oj} \cdot x_k + \sum_{l=1}^H w_{hl,oj} \cdot \varphi_h \left( \sum_{k=1}^n w_{ik;hl} \cdot x_k + b_{hl} \right) \right), \quad j = 1, \dots, m.$$

# Máquina de aprendizaje extrema

De acuerdo con Huang *et al.* [2006], los algoritmos de entrenamiento para redes neuronales son lentos y pueden llevar horas, días o inclusive más tiempo para encontrar los parámetros del modelo.

En la práctica, las redes neuronales se entran en un conjunto de entrenamiento finito.

Partiendo de esta suposición, Huang and Babri [1998] mostraron que una red neuronal de retroalimentación (SLFN) con una capa oculta con máximo  $N$  nodos ocultos y casi cualquier función de activación no lineal puede aprender exactamente  $N$  observaciones distintas.

Para  $N$  muestras distintas arbitrarias  $(\mathbf{x}_i, \mathbf{t}_i)$  donde  $\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in \mathbb{R}^n$  y  $\mathbf{t}_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T \in \mathbb{R}^m$ , la red feedforward de capa oculta única (SLFN) con nodos ocultos aleatorios y función de activación  $\varphi$  se modelada matemáticamente como

$$\sum_{i=1}^{\tilde{N}} \beta_i \varphi_i(\mathbf{x}_j) = \sum_{i=1}^{\tilde{N}} \beta_i \varphi_i(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = \mathbf{o}_j, \quad j = 1, \dots, N,$$

donde  $\mathbf{w}_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$  es el vector de pesos conectados al  $i$ -ésimo nodo oculto y los nodos de entrada,  $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$  es el vector de pesos conectados al  $i$ -ésimo nodo oculto y los nodos de salida, además  $b_i$  es el umbral de el  $i$ -ésimo nodo oculto.

Entonces, la red SLFN con  $\tilde{N}$  nodos ocultos y función de activación  $\varphi$  puede aproximar las  $N$  muestras con cero error  $\sum_{j=1}^N \|\mathbf{o}_j - \mathbf{t}_j\| = 0$ , i.e. que existe  $\beta_i$ ,  $\mathbf{w}_i$  y  $b_i$  tal que

$$\sum_{i=1}^{\tilde{N}} \beta_i \varphi_i(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = \mathbf{t}_j, \quad j = 1, \dots, N.$$

Las  $N$  ecuaciones se pueden escribir como

$$\mathbf{H}\beta = \mathbf{T},$$

donde

$$\mathbf{H}(\mathbf{w}_1, \dots, \mathbf{w}_{\tilde{N}}, b_1, \dots, b_{\tilde{N}}, \mathbf{x}_1, \dots, \mathbf{x}_N) = \begin{bmatrix} \varphi(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \cdots & \varphi(\mathbf{w}_{\tilde{N}} \cdot \mathbf{x}_1 + b_{\tilde{N}}) \\ \vdots & \ddots & \vdots \\ \varphi(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \cdots & \varphi(\mathbf{w}_{\tilde{N}} \cdot \mathbf{x}_N + b_{\tilde{N}}) \end{bmatrix}_{N \times \tilde{N}},$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{\tilde{N}}^T \end{bmatrix}_{\tilde{N} \times m}, \quad \mathbf{T} = \begin{bmatrix} t_1^T \\ \vdots \\ t_{\tilde{N}}^T \end{bmatrix}_{N \times m}.$$

Para pesos de entrada fijos  $w_i$  y sesgos de capa oculta  $b_i$ , el entrenamiento de una red tipo SLFN es equivalente a encontrar una solución de mínimos cuadrados  $\hat{\beta}$  del sistema lineal  $H\beta = T$

$$\|H(w_1, \dots, w_{\tilde{N}}, b_1, \dots, b_{\tilde{N}})\hat{\beta} - T\| = \min_{\beta} \|H(w_1, \dots, w_{\tilde{N}}, b_1, \dots, b_{\tilde{N}})\beta - T\|$$

Si  $\tilde{N} = N$ , entonces la matriz  $H$  es cuadrada y por consiguiente tiene inversa cuando el vector de peso de entrada  $w_i$  y los sesgos ocultos  $b_i$  son escogidos aleatoriamente.

En la práctica,  $\tilde{N} \ll N$ . Por lo tanto,  $H$  es rectangular y podría no existir  $w_i, b_i, \beta_i$  ( $i = 1, \dots, \tilde{N}$ ) tal que  $H\beta = T$ . Por consiguiente, la solución del sistema lineal está dada como sigue

$$\hat{\beta} = H^\dagger T,$$

donde  $H^\dagger$  es la *inversa generalizada Moore-Penrose* de la matriz  $H$  la cual satisface las siguientes condiciones:

1.  $HH^\dagger H = H$ ,
2.  $H^\dagger HH^\dagger = H^\dagger$ ,
3.  $(HH^\dagger)^T = HH^\dagger$ ,
4.  $(H^\dagger H)^T = H^\dagger H$ .

### **Note**

En el caso especial cuando la matriz  $H$  es una matriz cuadrada no singular, la pseudoinversa de  $H$  es simplemente su inversa, i.e.  $H^\dagger = H^{-1}$ .

La pseudoinversa también está definida como

$$\mathbf{H}^\dagger = \lim_{\nu \rightarrow 0} (\mathbf{H}^T \mathbf{H} + \nu^2 I)^{-1} \mathbf{H}^T = \lim_{\nu \rightarrow 0} \mathbf{H}^T (\mathbf{H} \mathbf{H}^T + \nu^2 I)^{-1}$$

o bien

$$\mathbf{H}^\dagger = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T = \mathbf{H}^T (\mathbf{H} \mathbf{H}^T)^{-1},$$

si  $\mathbf{H}^T \mathbf{H}$  o  $\mathbf{H} \mathbf{H}^T$  es no singular.

Finalmente, el método de entrenamiento para la SLFN llamado máquina de aprendizaje extrema (ELM) se presenta a continuación:

#### Algorithm 1 (Algoritmo ELM)

**Entradas** Dado un conjunto de entrenamiento

$\mathbb{N} = \{(\mathbf{x}_i, \mathbf{t}_i) | \mathbf{x}_i \in \mathbb{R}^n, \mathbf{t}_i \in \mathbb{R}^m, i = 1, \dots, N\}$ , una función de activación  $\varphi(\xi)$ , y nodo oculto  $\tilde{N}$ .

1. Asignar aleatoriamente pesos de entrada  $w_i$  y sesgos  $b_i$ .
2. Calcular la matriz de salida de la capa oculta  $\mathbf{H}$ .
3. Calcular los pesos de salida  $\beta$ .

$$\beta = \mathbf{H}^\dagger \mathbf{T},$$

donde  $\mathbf{T} = [\mathbf{t}_1, \dots, \mathbf{t}_N]^T$ .

## Identificación de sistemas utilizando modelos NARMAX

Los modelos de media móvil autorregresiva no lineal con entrada exógena (NARMAX) fueron propuestos por Billings [1981], Chen and Billings [1989], Leontaritis and Billings [1985]. Este modelo puede ser descrito como sigue

$$y_k = f(y_{k-1}, \dots, y_{k-n_y}, x_{k-d}, x_{k-d-1}, \dots, x_{k-d-n_x} + e_{k-1}, \dots, e_{k-n_e}) + e_k$$

donde  $n_y \in \mathbb{N}^*$ ,  $n_x \in \mathbb{N}$ ,  $n_e \in \mathbb{N}$  son los retrasos máximos para la salida y entrada del sistema;  $x_k \in \mathbb{R}^{n_x}$  es la entrada del sistema y  $y_k \in \mathbb{R}^{n_y}$  es la salida del sistema en tiempo discreto

$k \in \mathbb{N}^n$  mientras que  $e_k \in \mathbb{R}^{n_e}$  representa posible ruido e incertidumbres en el tiempo  $k$ . Además,  $f$  es una función no lineal de los regresores de entrada y salida;  $d$  es un retardo de tiempo típicamente establecido en  $d = 1$ .

Aunque se pueden hacer diversas aproximaciones a la función  $f$ , por ejemplo mediante redes neuronales, funciones de base radial, el modelo NARMAX polinomial de potencia es el más utilizado

$$y_k = \sum_{i=1}^p \Theta_i \times \prod_{j=0}^{n_x} u_{k-j}^{b_i,j} \prod_{l=1}^{n_e} e_{k-l}^{d_i,l} \prod_{m=1}^{n_y} y_{k-m}^{a_i,m},$$

donde  $p$  es el numero de regresores,  $\Theta_i$  los parámetros del modelo,  $a_i, m, b_i, j$  y  $d_i, l \in \mathbb{N}$  son los exponentes de la salida, la entrada y los términos de ruido, respectivamente.

Encontrar los parámetros de un modelo NARMAX es simple si la estructura del modelo se conoce a priori. Sin embargo, en la práctica no hay información sobre los términos que deben ser incluidos en el modelo final. No obstante, utilizar modelos NARMAX implica:

1. Seleccionar el número adecuado de regresores.
2. Estimar los parámetros del modelo.

# Referencias

- [Aka74] Hirotugu Akaike. A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723, 1974.
- [Bil81] SA Billings. Identification of nonlinear systems using parameter estimation techniques. In *Institute of Electrical Engineers Conference*, 183–190. 1981.
- [CB89] Sheng Chen and Steve A Billings. Representations of non-linear systems: the narmax model. *International journal of control*, 49(3):1013–1032, 1989.
- [CEGomezAT+17] Antonio Coronel-Escamilla, José Francisco Gómez-Aguilar, L Torres, Martin Valtierra-Rodriguez, and Ricardo Fabricio Escobar-Jiménez. Design of a state observer to approximate signals by using the concept of fractional variable-order derivative. *Digital Signal Processing*, 69:127–139, 2017.
- [HB98] Guang-Bin Huang and Haroon A Babri. Upper bounds on the number of hidden neurons in feedforward networks with arbitrary bounded nonlinear activation functions. *IEEE transactions on neural networks*, 9(1):224–229, 1998.
- [HZZS06] Guang-Bin Huang, Qin-Yu Zhu, and Chee-Kheong Siew. Extreme learning machine: theory and applications. *Neurocomputing*, 70(1-3):489–501, 2006.
- [Kha92] H.K. Khalil. *Nonlinear Systems*. Macmillan Publishing Company, 1992. ISBN 9780023635410. URL: <https://books.google.com.mx/books?id=RVHvAAAAMAAJ>.
- [LB85] I J̄ Leontaritis and Stephen A Billings. Input-output parametric models for non-linear systems part i: deterministic non-linear systems. *International journal of control*, 41(2):303–328, 1985.
- [Lju98] L. Ljung. *System Identification: Theory for the User*. Pearson Education, 1998. ISBN 9780132440530. URL: <https://books.google.com.mx/books?id=fYSrk4wDKPsC>.
- [Ris78] Jorma Rissanen. Modeling by shortest data description. *Automatica*, 14(5):465–471, 1978.
- [AAstromM21] Karl Johan Åström and Richard M Murray. *Feedback systems: an introduction for scientists and engineers*. Princeton university press, 2021.