

Teoría de Control

Roberto Cadena Vega

23 de octubre de 2014

Índice general

1. Criterio de Estabilidad de Routh-Hurwitz	9
Tabla de Routh	10
Casos Especiales	12
Aplicación del criterio de Routh	14
Ejemplo:	15
Acción Proporcional	16
Estabilidad	16
Error en el estado permanente al escalón unitario	16
Acción Integral	17
Estabilidad	17
Error en el estado permanente al escalón unitario	17
Acción Proporcional Integral	18
Estabilidad	18
Error en el estado permanente al escalón unitario	18
2. Lugar de las Raíces	19
3. Compensador Adelanto/Atraso (LR)	23
Compensador de adelanto de fase	23
Compensador de atraso de fase	23
Error estático de posición k_p	23
Error estático de velocidad k_v	23
4. Diagramas de Bode	25
Factor integral	25
Factor derivativo	25
Factores de primer orden	25
Factores de segundo orden	25
Frecuencia de resonancia ω_n y valor par de resonancia M_R	25
5. Diagramas de Nyquist	27
Factor integral	27
Factor derivativo	27
Factores de primer orden	27
Factores de segundo orden	27

6. Criterio de Estabilidad de Nyquist	29
Ejemplos	29
7. Estabilidad Relativa	31
Margen de Fase	31
Estable	31
Inestable	31
Margen de Ganancia	31
Estable	31
Inestable	31
8. Compensador de adelanto y atrase de fase (Frecuencia)	33
Compensador de adelanto de fase	33
Compensador de atraso de fase	33
Ejemplos	33
9. Controladores PID	35
Sintonización: Reglas de Ziegler-Nichols	35
Respuesta al escalón	35
Respuesta a oscilaciones sostenidas	35
Esquemas modificados	35
Controlador PID	35
Controlador PI-D	35
Controlador I-PD	35
10. Representación de estado	37
Solución temporal de la ecuación de estado	40
Función (Matriz) de transferencia de la ecuación de estado	43
Función de transferencia de la representación de estado	44
Matriz sistema	46
Propiedades de la Matriz A	47
11. Controlabilidad y asignación de polos	51
Alcanzabilidad y Controlabilidad	52
Asignación de polos	54
Propiedades de la matriz de controlabilidad	58
Formas canónicas	61
Forma canónica controlador	61
Forma canónica controlabilidad	61
12. Inobservabilidad y observador de estado	63
Observabilidad e inobservabilidad	65
Dualidad	68
Propiedades de la matriz de observabilidad	69
Formas canónicas	70
Forma canónica observador	70
Forma canónica observabilidad	70
Observador de estado	71
13. Principio de Separación	73

14. Estabilidad de Lyapunov	75
15. Introducción a la optimización de funcionales	81
Ecuación de Euler	83
Multiplicadores de Lagrange	85
Derivada direccional y vector gradiente	85
Multiplicadores de Lagrange	88

Todo list

Falta escribir apunte	23
Falta escribir apunte	23
Falta escribir apunte	23
Falta escribir apunte	23
Falta escribir apunte	25
Falta escribir apunte	25
Falta escribir apunte	25
Falta escribir apunte	25
Falta escribir apunte	25
Falta escribir apunte	27
Falta escribir apunte	27
Falta escribir apunte	27
Falta escribir apunte	27
Falta escribir apunte	29
Falta escribir apunte	31
Falta escribir apunte	31
Falta escribir apunte	31
Falta escribir apunte	31
Falta escribir apunte	31
Falta escribir apunte	31
Falta escribir apunte	33
Falta escribir apunte	33
Falta escribir apunte	33
Falta escribir apunte	35
Falta escribir apunte	35
Falta escribir apunte	35
Falta escribir apunte	35
Falta escribir apunte	35
Falta escribir apunte	35
Falta escribir apunte	35
Falta escribir apunte	35
Falta escribir apunte	70
Falta escribir apunte	70
Falta escribir apunte	70
Falta escribir apunte	71
Falta escribir apunte	73
Figure: Diagrama de trayectoria acotada en 2 dimensiones.	75
Resumen de matrices Hermitianas de Introduction to Matrix Computations - G.W. Stewart Cap. 4 y 6	75

Resumen de estabilidad de Lyapunov de Linear Systems - Thomas Kailath Cap. 2.6 . . .	76
Feedback Systems - Input/Output Properties - Desoer, Vidyasagar Ap. E	77
Figure: Función arbitraria acotada por una exponencial negativa.	77
Buscar demostración en The Theory of Matrices - F.R. Gantmacher, Vol 1, Cap. 8	78
Figure: Trayectorias diferentes con inicio y fin comunes.	81
Figure: Superficie con vector posición evaluado y vector dirección de la derivada unitario.	85
Figure: Curva cortada por una recta por la que pasa un vector restando los dos puntos evaluados de la curva.	85
Figure: Superficie con vector tangente.	87
Figure: Plano tangente a superficie de nivel y vector normal al plano.	87

Capítulo 1

Criterio de Estabilidad de Routh-Hurwitz

$$\frac{\hat{y}(s)}{\hat{R}(s)} = \frac{b_0s^m + b_1s^{m-1} + \dots + b_{m-1}s + b_m}{a_0s^n + a_1s^{n-1} + \dots + a_{n-1}s + a_n} = \frac{B(s)}{A(s)} \tag{1.1}$$

$$\frac{\hat{y}(s)}{\hat{R}(s)} = \sum \frac{k_{1,i}}{s + \alpha_i} + \sum \frac{k_{2,j} + k_{3,j} \cdot s}{(s + \beta_i)^2 + \gamma_i^2}; m \leq n \tag{1.2}$$

El criterio de Routh-Hurwitz determina si existen raíces en el semiplano complejo derecho cerrado.

Tabla de Routh

La tabla de Routh es un método para obtener el numero de raíces con parte real positiva que se encontraran en el polinomio característico del sistema (Ecuación 1.3) sin tener que calcular las raíces en cuestión. Se puede dividir en cuatro pasos que se enumeran a continuación.

$$A(s) = a_0s^n + a_1s^{n-1} + ... + a_{n-1}s + a_n = 0 \tag{1.3}$$

1. Hipótesis Si $a_0 = 0 \Rightarrow$ el polinomio es de orden menor a n .
Si $a_n = 0 \Rightarrow \exists$ una raíz que es $0 \Rightarrow A(s) = (\tilde{n}_0s^{\tilde{n}} + \tilde{a}_{\tilde{n}}s^{\tilde{n}-1} + ...)s^k$.
2. Si existen coeficientes nulos o de diferente (cambio de) signo, entonces existen raíces con parte real positiva.

3. Construir la tabla de Routh (Ver Cuadro 1.1).

s^n	a_0	a_2	a_4	a_6	...
s^{n-1}	a_1	a_3	a_5	a_7	...
s^{n-2}	b_1	b_3	b_5	b_7	...
s^{n-3}	c_1	c_3	c_5	c_7	...
s^{n-4}	d_1	d_3	d_5	d_7	...
\vdots					
s^2	e_1	e_2			
s^1	f_1				
s^0	g_1				

Cuadro 1.1: Ejemplo de tabla de Routh.

Donde:

$$\begin{aligned}
 b_1 &= \frac{a_1 a_2 - a_0 a_3}{a_1}, b_2 = \frac{a_1 a_4 - a_0 a_5}{a_1}, \dots \\
 c_1 &= \frac{b_1 a_3 - a_1 b_2}{b_1}, c_2 = \frac{b_1 a_5 - a_1 b_3}{b_1}, \dots \\
 d_1 &= \frac{c_1 b_2 - b_1 c_2}{c_1}, d_2 = \frac{c_1 b_3 - b_1 c_3}{c_1}, \dots \\
 &\vdots
 \end{aligned}$$

4. El número de raíces con parte real positiva es igual al numero de cambios de signo en la primera columna(Ver Cuadro 1.2).

s^n	a_0
s^{n-1}	a_1
s^{n-2}	b_1
s^{n-3}	c_1
s^{n-4}	d_1
\vdots	\vdots
s^2	e_1
s^1	f_1
s^0	g_1

Cuadro 1.2: Números en los que hay que revisar el cambio de signo.

Casos Especiales

1. En los casos en los que un coeficiente es 0 se puede intercambiar por un ϵ lo suficientemente pequeño para aproximar a 0 (Véase el Cuadro 1.3).

$$A(s) = s^3 + 2s^2 + s + 2 = 0$$

```
>> A = [1 2 1 2];
>> r = roots(A)
r =
-2.00000 + 0.00000i
-0.00000 + 1.00000i
-0.00000 - 1.00000i
```

$$\begin{array}{c|cc} s^3 & 1 & 1 \\ s^2 & 2 & 2 \\ s^1 & 0 \approx \epsilon & \\ s^0 & 2 & \end{array}$$

Cuadro 1.3: Caso Especial 1.

2. Cuando existen cambios en los coeficientes del polinomio característico se sabe que existirán raíces con parte real positiva (Véase el cuadro 1.4).

$$A(s) = s^3 - 3s + 2 = 0$$

```
>> A = [1 0 -3 2];
>> r = roots(A)
r =
-2.00000
1.00000
1.00000
```

$$\begin{array}{c|cc} s^3 & 1 & -3 \\ s^2 & 0 \approx \epsilon & 2 \\ s^1 & -\frac{2}{\epsilon} & 0 \\ s^0 & 2 & \end{array}$$

Cuadro 1.4: Caso Especial 2.

3. Cuando todos los coeficientes en una linea se eliminan se puede crear un nuevo polinomio auxiliar con la linea anterior, obtener su derivada e insertar en la siguiente linea para continuar calculando la tabla (Véase el Cuadro 1.5 y 1.6).

$$A(s) = s^5 + 2s^4 + 24s^3 + 48s^2 - 25s - 50 = 0$$

$$p_{aux}(s) = 2s^4 + 48s^2 - 50$$

$$\frac{d}{dx}p_{aux}(s) = 8s^3 + 96s$$

```
>> A = [1 2 24 48 -25 -50];
>> r = roots(A)
r =
-0.00000 + 5.00000i
-0.00000 - 5.00000i
 1.00000 + 0.00000i
-2.00000 + 0.00000i
-1.00000 + 0.00000i
```

$$\begin{array}{c|ccc} s^5 & 1 & 24 & -25 \\ s^4 & 2 & 48 & -50 \\ s^3 & 0 & 0 & 0 \\ s^2 & & & \\ s^1 & & & \\ s^0 & & & \end{array}$$

Cuadro 1.5: Caso Especial 3a.

$$\begin{array}{c|ccc} s^5 & 1 & 24 & -25 \\ s^4 & 2 & 48 & -50 \\ s^3 & 8 & 96 & 0 \\ s^2 & 24 & -50 & 0 \\ s^1 & 112,6 & 0 & \\ s^0 & -50 & & \end{array}$$

Cuadro 1.6: Caso Especial 3b.

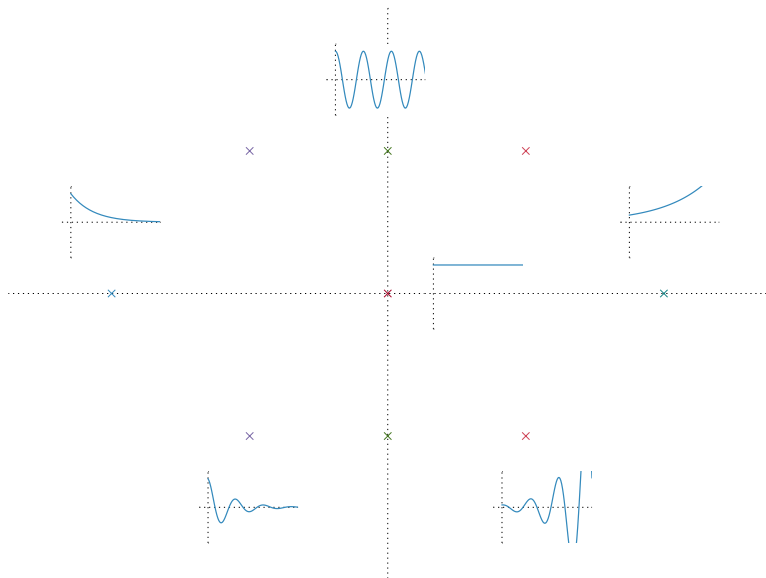


Figura 1.1: Polos en el plano complejo.

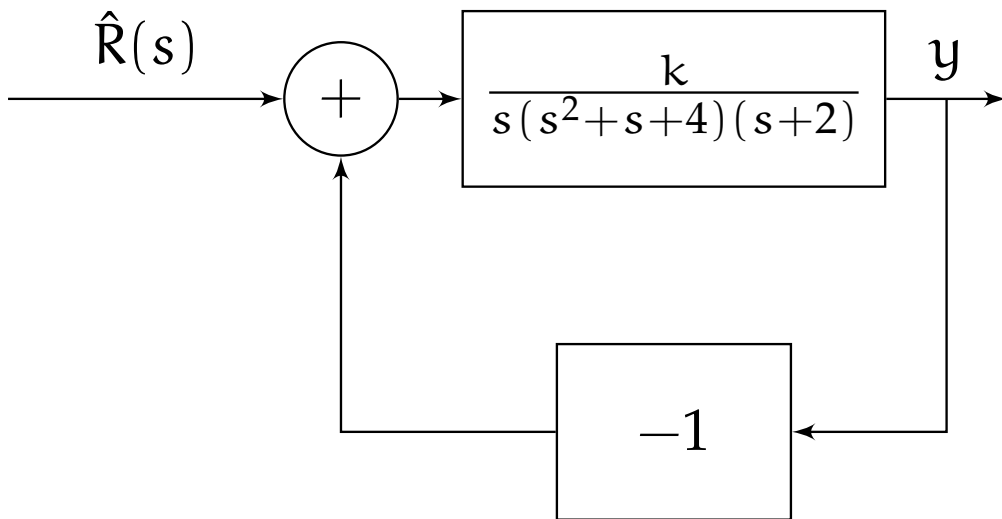
Aplicación del criterio de Routh

Si bien los sistemas numéricos actuales permiten el cálculo de las raíces de un sistema de manera más rápida y sencilla que con la aplicación de este método, aun existen aplicaciones prácticas en las que es de suma importancia el determinar el número de raíces positivas. Por ejemplo podemos tener ganancias en un sistema para las que queremos determinar de primera instancia, un rango de valores para los cuales el sistema no se volverá inestable.

Para ello calculamos la tabla de Routh de la misma manera en que lo hicimos anteriormente, pero teniendo en cuenta las ganancias a incluir en el cálculo de las raíces (Por ejemplo con una ganancia proporcional véase Cuadro 1.7).

s^n	a_0	a_2	a_4	a_6	...
s^{n-1}	a_1	a_3	a_5	a_7	k
s^{n-2}	b_1	b_3	b_5	b_7	...
s^{n-3}	c_1	c_3	c_5	c_7	...
s^{n-4}	d_1	d_3	d_5	d_7	...
\vdots					
s^2	e_1	e_2			
s^1	f_1				
s^0	g_1				

Cuadro 1.7: Aplicación del criterio de Routh.



Ejemplo:

Se toma el sistema $\frac{\hat{y}(s)}{\hat{R}(s)} = \frac{k}{s^4 + 3s^3 + 3s^2 + 2s + k}$, entonces el polinomio característico del sistema será $F(s) = s^4 + 3s^3 + 3s^2 + 2s + k$.

Construimos su tabla de Routh (Cuadro 1.8):

s^4	1	3	k
s^3	3	2	0
s^2	$\frac{7}{3}$	k	0
s^1	$2 - \frac{9}{7}k$	0	
s^0	k		

Cuadro 1.8: Ejemplo de Aplicación del criterio de Routh.

De lo anterior podemos concluir que, para que no existan cambios de signos, toda la primera columna tiene que ser positiva, por lo que $k > 0$ y $2 - \frac{9}{7}k > 0$, por lo que el rango de valores que puede ocupar la ganancia k es $0 < k < \frac{14}{9}$.

Si bien esto no nos aporta una ganancia específica para un comportamiento deseado, si nos da la pauta a los valores a tomar en cuenta, si no se desea que el sistema sea inestable.

Acción Proporcional

Tenemos un sistema de primer orden, al que le agregaremos un controlador de ganancia proporcional y una retroalimentación negativa, por lo que las ecuaciones que describen la salida y el error del sistema quedan:

$$\frac{\hat{y}(s)}{\hat{R}(s)} = \frac{k}{Ts + 1 + k} \tag{1.4}$$

$$\frac{\hat{e}(s)}{\hat{R}(s)} = \frac{R(s) - Y(s)}{R(s)} = \frac{Ts + 1}{Ts + 1 + k} \tag{1.5}$$

Estabilidad

El problema reside en encontrar un conjunto de ganancias k para las cuales el sistema es estable.

$$F(s) = s + \frac{1 + k}{T} \tag{1.6}$$

Aplicamos una tabla de Routh a este polinomio característico (Cuadro 1.9).

$$\begin{array}{c|c} s^1 & 1 \\ s^0 & \frac{1+k}{T} \end{array}$$

Cuadro 1.9: Tabla de Routh para acción proporcional.

Por lo que concluimos que la ganancia k debe de seguir: $k > -1$

Error en el estado permanente al escalón unitario

También es importante investigar el error que causara el controlador al introducirse. Si ponemos como señal de referencia al escalón unitario($R(s) = \frac{1}{s}$), podemos ver lo siguiente:

$$\lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} se(s) = \lim_{s \rightarrow 0} \frac{Ts + 1}{Ts + 1 + k} = \frac{1}{1 + k}$$

Acción Integral

Tenemos un sistema de primer orden, al que le agregaremos un controlador de ganancia integral y una retroalimentación negativa, por lo que las ecuaciones que describen la salida y el error del sistema quedan:

$$\frac{\hat{y}(s)}{\hat{R}(s)} = \frac{k}{s(Ts + 1) + k} \tag{1.7}$$

$$\frac{\hat{e}(s)}{\hat{R}(s)} = \frac{R(s) - Y(s)}{R(s)} = \frac{s(Ts + 1)}{s(Ts + 1) + k} \tag{1.8}$$

Estabilidad

El problema reside en encontrar un conjunto de ganancias k para las cuales el sistema es estable.

$$F(s) = s^2 + \frac{1}{T}s + \frac{k}{T} \tag{1.9}$$

Aplicamos una tabla de Routh a este polinomio característico (Cuadro 1.10).

s^2	1	$\frac{k}{T}$
s^1	$\frac{1}{T}$	0
s^0	$\frac{k}{T}$	

Cuadro 1.10: Tabla de Routh para acción integral.

Por lo que concluimos que la ganancia k debe de seguir: $k > 0$

Error en el estado permanente al escalón unitario

También es importante investigar el error que causara el controlador al introducirse. Si ponemos como señal de referencia al escalón unitario($R(s) = \frac{1}{s}$), podemos ver lo siguiente:

$$\lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} se(s) = \lim_{s \rightarrow 0} s \left(\frac{s(Ts + 1)}{s(Ts + 1) + k} \frac{1}{s} \right) = 0$$

Acción Proporcional Integral

Estabilidad

Error en el estado permanente al escalón unitario

Capítulo 2

Lugar de las Raíces

Si tenemos un sistema con retroalimentación, su polinomio característico es el siguiente:

$$F(s) = 1 + H(s)G(s) = 0 \quad (2.1)$$

Donde $G(s)$ es la planta y $H(s)$ es el elemento de retroalimentación. Las condiciones de ángulo y magnitud son las siguientes:

$$\angle H(s)G(s) = \pm 180^\circ (2R + 1) \mid R \in \mathbb{Z}^+ \quad (2.2)$$

$$|H(s)G(s)| = 1 \quad (2.3)$$

De aquí notamos que la condición de ángulo, nos da la forma del lugar de las raíces, y la condición de magnitud nos da su posición.

Pues bien, para trazar el lugar geométrico de las raíces seguimos una serie de pasos enumerados a continuación:

1. Determinar el lugar de las raíces en el eje real.

Ejemplo: $H(s) = 1, G(s) = \frac{k}{s(s+1)(s+2)}$

Sabemos, por una inspección visual, que los polos del sistema son 0, -1 , y -2 , y que este sistema no tiene ceros. Lo cual nos indica, por la condición de ángulo, que la suma de las interacciones de estas raíces, nos dará la interacción total del sistema:

$$\angle G(s) = -\angle s - \angle s + 1 - \angle s + 2 = \pm 180^\circ (2R + 1)$$

Notemos que cualquier lugar de las raíces en el semiplano derecho complejo (inestable), viene con un ángulo de 0° , por lo que las interacciones de cada polo serían:

$$\angle G(s) = -0^\circ - 0^\circ - 0^\circ = 0^\circ$$

lo cual obviamente no cumple con la condición de ángulo del sistema.

Si pasamos a la siguiente sección del eje real (creada por los mismos polos del sistema), tenemos que los ángulos de interacción de cada polo son:

$$\angle G(s) = -180^\circ - 0^\circ - 0^\circ = -180^\circ$$

lo cual cumple con la condición de ángulo del sistema.

En la siguiente sección (entre -1 y -2), tenemos lo siguiente:

$$\angle G(s) = -180^\circ - 180^\circ - 0^\circ = -360^\circ = 0^\circ$$

y esto no cumple con la condición de ángulo del sistema.

En la última sección (entre -2 y $-\infty$) tenemos:

$$\angle G(s) = -180^\circ - 180^\circ - 180^\circ = -540^\circ = -180^\circ$$

por lo que esta última sección también es parte del lugar geométrico de las raíces.

2. Determinar las asíntotas del lugar de las raíces.

El lugar de las raíces se aproxima a sus asíntotas, mientras $s \rightarrow \infty$, por lo que podemos hacer una simplificación:

$$\lim_{s \rightarrow \infty} G(s) = \lim_{s \rightarrow \infty} \frac{k}{s(s+1)(s+2)} = \lim_{s \rightarrow \infty} \frac{K}{s^3}$$

por lo que la condición de ángulo queda:

$$\angle G(s) = -3\angle s = \pm 180^\circ(2R+1) \implies \angle s = \pm 60^\circ(2R+1)$$

lo cual nos da que los ángulos de las asíntotas son 60° , -60° y 120° .

Por otro lado, si hacemos un proceso similar, pero con el polinomio característico desarrollado, podremos ver que hay términos más importantes que otros, en especial cuando hacemos $s \rightarrow \infty$, por lo que:

$$G(s) = \frac{k}{s(s+1)(s+2)} = \frac{k}{s^3 + 3s^2 + 2s} \approx \frac{k}{(s+1)^3}$$

por lo que podemos ver que las asíntotas tienen esa forma, y que podemos asegurar que parten del punto $-1 + 0i$.

3. Determinar el punto de ruptura o partida de las asíntotas en el eje real.

Para determinar el punto de ruptura del lugar de las raíces, tenemos que pensar en el polinomio característico como la suma de 2 polinomios diferentes $A(s)$ y $B(s)$, de tal manera que ninguno contenga a la ganancia k , entonces tendremos:

$$F(s) = B(s) + kA(s) = 0 \implies k = -\frac{B(s)}{A(s)}$$

implicando que estamos obteniendo las ganancias, para las cuales se tienen polos en el plano complejo.

De aqui podemos pensar en el punto maximo de esta función de ganancias, como el punto de ruptura buscado, es decir:

$$\frac{dk}{ds} = 0 \tag{2.4}$$

En nuestro ejemplo, esto nos da como resultado:

$$k = -s^3 - 3s^2 - 2s \implies \frac{dk}{ds} = -3s^2 + 6s + 2 = 0$$

de donde obtenemos un par de respuestas $s_1 = -0,423$ y $s_2 = -1,577$, con ganancias asociadas $k_1 = 0,385$ y $k_2 = -0,385$.

De aqui podemos descartar s_2 ya que no se encuentra en el lugar de las raices del eje real, y obviamente no puede partir de ahi, si no existe en ese lugar en especifico.

4. Determinar los puntos donde el lugar de las raíces atraviesa el eje imaginario.

Ya hemos visto que los polos sobre el eje real no cruzan el eje imaginario, ahora solo tenemos que encontrar las ganancias criticas, es decir, cuando los polos estan sobre el eje imaginario.

$$F(s) = s^3 + 3s^2 + 2s + k$$

s^3	1	2
s^2	3	k
s^1	$2 - k/3$	0
s^0	k	

De donde obtenemos que $k > 0$, lo cual ocurre en el polo del origen y $k < 6$, que es justo cuando cruza por el eje imaginario.

Ahora, tan solo tenemos que obtener las raices del polinomio caracteristico con la ganancia adecuada y obtendremos el punto de cruce, alternatively, podemos usar el polinomio auxiliar de la tabla de Routh, usaremos el correspondiente a s^2 .

$$P_{aux} = 3s^2 + k = 3s^2 + 6 = 0 \implies s = \pm\sqrt{2}j$$

Capítulo 3

Compensador Adelanto/Atraso (LR)

Compensador de adelanto de fase

Compensador de atraso de fase

Error estático de posición k_p

Error estático de velocidad k_v

Capítulo 4

Diagramas de Bode

Factor integral

Factor derivativo

Factores de primer orden

Factores de segundo orden

Frecuencia de resonancia ω_n y valor par de resonancia M_R

Capítulo 5

Diagramas de Nyquist

Factor integral

Factor derivativo

Factores de primer orden

Factores de segundo orden

Capítulo 6

Criterio de Estabilidad de Nyquist

Ejemplos

Capítulo 7

Estabilidad Relativa

Margen de Fase

Estable

Inestable

Margen de Ganancia

Estable

Inestable

Capítulo 8

Compensador de adelanto y atrase de fase (Frecuencia)

Compensador de adelanto de fase

Compensador de atraso de fase

Ejemplos

Capítulo 9

Controladores PID

Sintonización: Reglas de Ziegler-Nichols

Respuesta al escalón

Respuesta a oscilaciones sostenidas

Esquemas modificados

Controlador PID

Controlador PI-D

Controlador I-PD

Capítulo 10

Representación de estado

La siguiente funcion de transferencia es la Transformada de Laplace de la ecuacion diferencial ordinaria de orden n que describe al sistema.

$$\frac{\hat{y}(s)}{\hat{u}(s)} = \frac{b_0 s^m + b_1 s^{m-1} + \dots + b_{m-1} s + b_m}{s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n} = \frac{B(s)}{A(s)}, m \leq n \quad (10.1)$$

$$\begin{aligned} \frac{d^n}{dt^n} y(t) + a_1 \frac{d^{n-1}}{dt^{n-1}} y(t) + \dots + a_{n-1} \frac{d}{dt} y(t) + a_n y(t) \\ = b_0 \frac{d^m}{dt^m} u(t) + b_1 \frac{d^{m-1}}{dt^{m-1}} u(t) + \dots + b_{m-1} \frac{d}{dt} u(t) + b_m u(t) \end{aligned} \quad (10.2)$$

Haciendo la siguiente asignacion de variables:

$$\begin{aligned} x_1 &= z \\ x_2 &= \frac{d}{dt} x_1 = \frac{d}{dt} z \\ x_3 &= \frac{d}{dt} x_2 = \frac{d^2}{dt^2} z \\ &\vdots \\ x_{n-1} &= \frac{d}{dt} x_{n-2} = \frac{d^{n-2}}{dt^{n-2}} z \\ x_n &= \frac{d}{dt} x_{n-1} = \frac{d^{n-1}}{dt^{n-1}} z \end{aligned}$$

Donde:

$$\frac{d}{dt} x_n = -a_n x_1 - a_{n-1} x_2 - \dots - a_2 x_{n-1} - a_1 x_n + u(t) \quad (10.3)$$

$$y = b_m x_1 + b_{m-1} x_2 + \dots + b_1 x_{m-1} + b_0 x_m \quad (10.4)$$

Por lo que se obtiene:

$$\left(\frac{d^n}{dt^n} + a_1 \frac{d^{n-1}}{dt^{n-1}} + \cdots + a_{n-1} \frac{d}{dt} + a_n \right) z(t) = u(t)$$

$$y(t) = \left(b_m + b_{m-1} \frac{d}{dt} + \cdots + b_1 \frac{d^{m-1}}{dt^{m-1}} + b_0 \frac{d^m}{dt^m} \right) z(t)$$

es decir:

$$M \left(\frac{d}{dt} \right) z(t) = u(t) \quad (10.5)$$

$$y(t) = N \left(\frac{d}{dt} \right) z(t) \quad (10.6)$$

Lo cual implica $M \left(\frac{d}{dt} \right) y(t) = N \left(\frac{d}{dt} \right) u(t)$. Donde:

$$M \left(\frac{d}{dt} \right) = \left(\frac{d^n}{dt^n} + a_1 \frac{d^{n-1}}{dt^{n-1}} + \cdots + a_{n-1} \frac{d}{dt} + a_n \right)$$

$$N \left(\frac{d}{dt} \right) = \left(b_m + b_{m-1} \frac{d}{dt} + \cdots + b_1 \frac{d^{m-1}}{dt^{m-1}} + b_0 \frac{d^m}{dt^m} \right)$$

Esta es la misma ecuación diferencial con la que empezamos. Note que la escritura matricial de esta Ecuación Diferencial Ordinaria¹ es:

$$\frac{d}{dt} \vec{x} = A \vec{x}(t) + \vec{b} u(t) \quad (10.7)$$

$$\vec{y}(t) = \vec{c}^T \cdot \vec{x}(t) \quad (10.8)$$

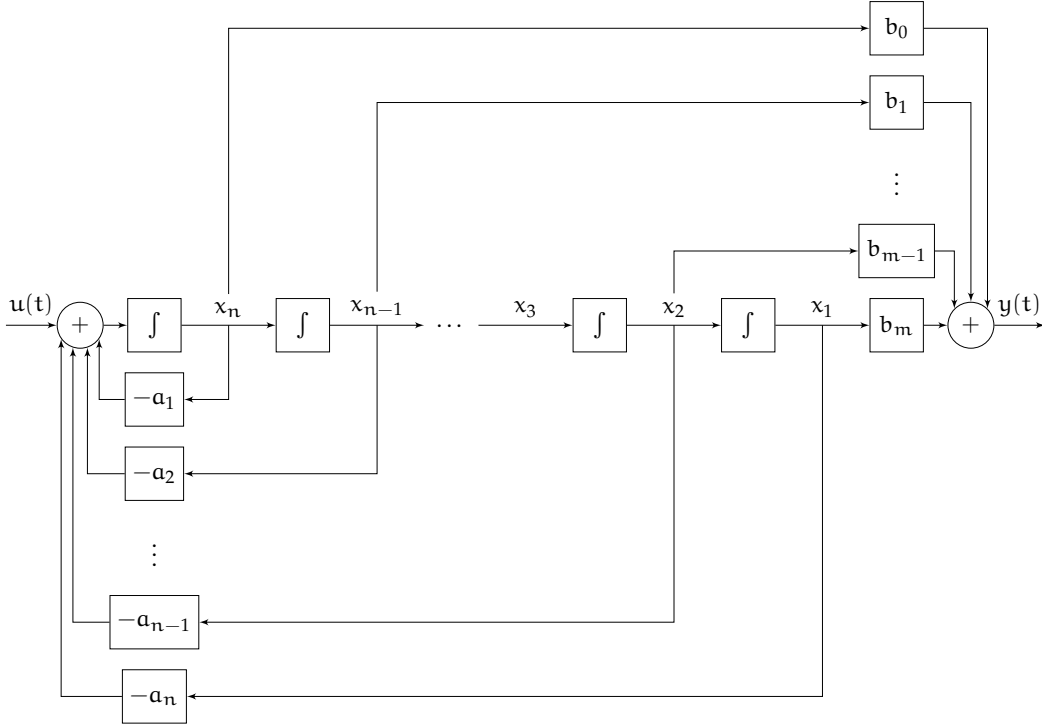
Donde:

$$\vec{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix} \quad (10.9)$$

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_2 & -a_1 \end{pmatrix} \quad (10.10)$$

$$\vec{b} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \quad (10.11)$$

¹Si bien la notación correcta es la que se utiliza justo ahora, al final del capítulo se dejará a un lado, para obviar el hecho de que son vectores y matrices, sin que por eso se entienda que ya no lo son.



$$\vec{c} = \begin{pmatrix} b_m \\ b_{m-1} \\ \vdots \\ b_1 \\ b_0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \tag{10.12}$$

Solución temporal de la ecuación de estado

1. Para el caso en que A es un escalar y la solución es homogénea se considera la siguiente Ecuación Diferencial Ordinaria:

$$\frac{d}{dt}x(t) = ax(t) \mid x(0) = x_0 \quad (10.13)$$

Suponga una solución de la forma:

$$x(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \dots + \alpha_k t^k + \dots \quad (10.14)$$

Entonces se tiene:

$$\begin{aligned} \alpha_1 + 2\alpha_2 t + 3\alpha_3 t^2 + \dots + k\alpha_k t^{k-1} + \dots \\ = a\alpha_0 + a\alpha_1 t + a\alpha_2 t^2 + \dots + a\alpha_k t^k + \dots \end{aligned}$$

Por lo que las α_i deben satisfacer:

$$\begin{aligned} \alpha_1 &= a\alpha_0 &= \frac{1}{1!} a^1 \alpha_0 \\ \alpha_2 &= a\alpha_1 &= \frac{1}{2!} a^2 \alpha_0 \\ \alpha_3 &= a\alpha_2 &= \frac{1}{3!} a^3 \alpha_0 \quad ; \quad \alpha_0 = x_0 \\ \vdots &= \vdots &= \vdots \\ \alpha_k &= a\alpha_{k-1} &= \frac{1}{k!} a^k \alpha_0 \end{aligned} \quad (10.15)$$

Esto es:

$$\begin{aligned} x(t) &= \left(\sum_{i=0}^{\infty} \frac{1}{i!} (at)^i \right) x_0 \\ x(t) &= e^{at} x_0 \end{aligned} \quad (10.16)$$

Notese que: $\frac{d}{dt}x(t) = \left(\sum_{i=0}^{\infty} \frac{1}{i!} \frac{d}{dt} (at)^i \right) x_0 = \left(\sum_{i=1}^{\infty} \frac{1}{(i-1)!} (at)^{i-1} \right) ax_0 = \left(\sum_{j=0}^{\infty} \frac{1}{j!} (at)^j \right) ax_0$

$$\frac{d}{dt}x(t) = ae^{at}x_0 = ax(t) \quad x(0) = x_0$$

2. Para el caso en que A es una matriz y la solución es homogénea se considera la siguiente Ecuación Diferencial Ordinaria:

$$\frac{d}{dt}\vec{x}(t) = A\vec{x}(t) \mid \vec{x}(0) = \vec{x}_0 \quad (10.17)$$

De la misma manera que en el caso escalar, se supone una solución de la forma:

$$\vec{x}(t) = \vec{\alpha}_0 + \vec{\alpha}_1 t + \vec{\alpha}_2 t^2 + \dots + \vec{\alpha}_k t^k + \dots \quad (10.18)$$

Entonces se tiene:

$$\begin{aligned}\vec{\alpha}_1 + 2\vec{\alpha}_2 t + 3\vec{\alpha}_3 t^2 + \dots + k\vec{\alpha}_k t^{k-1} + \dots \\ = A\vec{\alpha}_0 + A\vec{\alpha}_1 t + A\vec{\alpha}_2 t^2 + \dots + A\vec{\alpha}_k t^k + \dots\end{aligned}$$

Por lo que las $\vec{\alpha}_i$ deben satisfacer:

$$\begin{aligned}\vec{\alpha}_1 &= A\vec{\alpha}_0 = \frac{1}{1!}A^1\vec{\alpha}_0 \\ \vec{\alpha}_2 &= A\vec{\alpha}_1 = \frac{1}{2!}A^2\vec{\alpha}_0 \\ \vec{\alpha}_3 &= A\vec{\alpha}_2 = \frac{1}{3!}A^3\vec{\alpha}_0 \quad ; \quad \vec{\alpha}_0 = \vec{x}_0 \\ \vdots &= \vdots = \vdots \\ \vec{\alpha}_k &= A\vec{\alpha}_{k-1} = \frac{1}{k!}A^k\vec{\alpha}_0\end{aligned}\tag{10.19}$$

Esto es:

$$\vec{x}(t) = \left(\sum_{i=0}^{\infty} \frac{1}{i!} (At)^i \right) \vec{x}_0$$

En análisis real, se demuestra que esta serie es absolutamente convergente y se define como:

$$\exp(At) = \sum_{i=0}^{\infty} \frac{1}{i!} (At)^i\tag{10.20}$$

Notese que:

$$\frac{d}{dt} \exp(At) = \frac{d}{dt} \sum_{i=0}^{\infty} \frac{1}{i!} (At)^i = \left(\sum_{i=1}^{\infty} \frac{1}{(i-1)!} (At)^{i-1} \right) A = A \sum_{j=0}^{\infty} \frac{1}{j!} (At)^j = A \exp(At)$$

Por lo que:

$$\vec{x}'(t) = \exp(At)\vec{x}_0$$

$$\frac{d}{dt} \vec{x}(t) = A \exp(At)\vec{x}_0 = A\vec{x}(t) \quad \vec{x}(0) = \vec{x}_0$$

3. Para el caso en que A es escalar y la solución es forzada:

$$\frac{d}{dt} x(t) = ax(t) + bu(t) \mid x(0) = 0\tag{10.21}$$

La solución a esta ecuación es:

$$x(t) = \int_0^t e^{a(t-\tau)} bu(\tau) d\tau \quad (10.22)$$

$$\frac{d}{dt}x(t) = e^{a(t-t)}bu(t) + \int_0^t \frac{d}{dt}e^{a(t-\tau)}bu(\tau) d\tau = bu(t) + a \int_0^t e^{a(t-\tau)}bu(\tau) d\tau$$

$$\frac{d}{dt}x(t) = bu(t) + ax(t) \quad (10.23)$$

Por lo que la solución general (con $x(0) = x_0$):

$$x(t) = e^{at}x_0 + \int_0^t e^{a(t-\tau)}bu(\tau) d\tau \quad (10.24)$$

4. Para el caso en que A es una matriz y la solución es forzada:

$$\frac{d}{dt}\vec{x}(t) = A\vec{x}(t) + \vec{b}u(t) \mid \vec{x}(0) = 0 \quad (10.25)$$

La solución de esta ecuación es:

$$\vec{x}(t) = \int_0^t \exp(A(t-\tau))\vec{b}u(\tau) d\tau \quad (10.26)$$

En efecto, derivando tenemos:

$$\frac{d}{dt}\vec{x}(t) = \exp(A(t-t))\vec{b}u(t) + \int_0^t \frac{d}{dt}\exp(A(t-\tau))\vec{b}u(\tau) d\tau = \vec{b}u(t) + A \int_0^t \exp(A(t-\tau))\vec{b}u(\tau) d\tau$$

$$\frac{d}{dt}\vec{x}(t) = \vec{b}u(t) + A\vec{x}(t) \quad (10.27)$$

Por lo que la solución general (con $x(0) = x_0$):

$$\vec{x}(t) = \exp(At)\vec{x}_0 + \int_0^t \exp(A(t-\tau))\vec{b}u(\tau) d\tau \quad (10.28)$$

Función (Matriz) de transferencia de la ecuación de estado

1. Para el caso escalar, se tiene que la transformada de Laplace con coeficientes independientes nulos es:

$$\begin{aligned}sx(s) &= ax(s) + bu(s) \\(s - a)x(s) &= bu(s) \\x(s) &= (s - a)^{-1}bu(s) \\x(s) &= \frac{b}{s - a}u(s)\end{aligned}$$

Por lo que:

$$e^{at} = \mathcal{L}^{-1} \left\{ (s - a)^{-1} \right\} \quad (10.29)$$

2. Para el caso matricial, tenemos que la transformada de Laplace con coeficientes independientes nulos es:

$$\begin{aligned}s\vec{x}(s) &= A\vec{x}(s) + \vec{b}u(s) \\(sI - A)\vec{x}(s) &= \vec{b}u(s) \\x(s) &= (sI - A)^{-1}\vec{b}u(s)\end{aligned}$$

Por lo que:

$$\exp (At) = \mathcal{L}^{-1} \left\{ (sI - A)^{-1} \right\} \quad (10.30)$$

Función de transferencia de la representación de estado

Sea la siguiente Ecuación Diferencial Ordinaria:

$$M\left(\frac{d}{dt}\right)y(t) = N\left(\frac{d}{dt}\right)u(t)$$

donde:

$$M\left(\frac{d}{dt}\right) = \frac{d^n}{dt^n} + a_1 \frac{d^{n-1}}{dt^{n-1}} + \cdots + a_{n-1} \frac{d}{dt} + a_n$$
$$N\left(\frac{d}{dt}\right) = b_m + b_{m-1} \frac{d}{dt} + \cdots + b_1 \frac{d^{m-1}}{dt^{m-1}} + b_0 \frac{d^m}{dt^m}$$

La función de transferencia con coeficientes independientes nulos de las ecuaciones es:

$$F(s) = \frac{N(s)}{M(s)} = \frac{\frac{d^n}{dt^n} + a_1 \frac{d^{n-1}}{dt^{n-1}} + \cdots + a_{n-1} \frac{d}{dt} + a_n}{b_m + b_{m-1} \frac{d}{dt} + \cdots + b_1 \frac{d^{m-1}}{dt^{m-1}} + b_0 \frac{d^m}{dt^m}} \tag{10.31}$$

Ceros. Las raíces del polinomio $N(s)$.

Polos. Las raíces del polinomio $M(s)$.

Sea la siguiente representación de estado de la Ecuación Diferencial Ordinaria:

$$\begin{aligned} \frac{d}{dt}x &= Ax + bu \\ y &= c^T x + du \end{aligned}$$

La función de transferencia con coeficientes independientes nulos en esta representación es:

$$\begin{aligned} sx(s) &= Ax(s) + bu(s) \\ y(s) &= c^T x(s) + du(s) \\ &\vdots \\ (sI - A)x(s) &= bu(s) \\ y(s) &= c^T x(s) + du(s) \\ &\vdots \\ x(s) &= (sI - A)^{-1}bu(s) \\ y(s) &= c^T x(s) + du(s) \end{aligned}$$

$$\vdots$$

$$\mathbf{y}(s) = \mathbf{c}^T[(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}\mathbf{u}(s)] + \mathbf{d}\mathbf{u}(s) = [\mathbf{c}^T(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} + \mathbf{d}]\mathbf{u}(s)$$

$$\vdots$$

$$F(s) = \mathbf{c}^T(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} + \mathbf{d} \tag{10.32}$$

Matriz sistema

$$\Sigma(s) = \begin{pmatrix} sI - A & b \\ -c^T & d \end{pmatrix} \tag{10.33}$$

Note que:

$$\begin{aligned} \begin{pmatrix} (sI - A)^{-1} & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} sI - A & b \\ -c^T & d \end{pmatrix} \begin{pmatrix} I & -(sI - A)b \\ 0 & I \end{pmatrix} \\ = \begin{pmatrix} I & (sI - A)^{-1}b \\ -c^T & d \end{pmatrix} \begin{pmatrix} I & -(sI - A)b \\ 0 & I \end{pmatrix} \\ = \begin{pmatrix} I & 0 \\ -c^T & c^T(sI - A)^{-1}b + d \end{pmatrix} \end{aligned}$$

Por lo que:

$$\det \left((sI - A)^{-1} \right) \cdot \det (\Sigma(s)) \cdot I = c^T (sI - A)^{-1} b + d$$

$$F(s) = \frac{\det (\Sigma(s))}{\det (sI - A)} \tag{10.34}$$

Por lo que los polos coinciden con los valores propios de A y los ceros son los números complejos que hacen perder rango a la matriz sistema.

$$\text{Polos: } F(s) = \{s \in \mathbb{C} \mid \det (sI - A) = 0\} \tag{10.35}$$

$$\text{Ceros: } F(s) = \{s \in \mathbb{C} \mid \det (\Sigma(s)) = 0\} \tag{10.36}$$

Propiedades de la Matriz A

I) Definición de la matriz exponencial.

$$\exp (At)=\sum_{i=0}^{\infty} \frac{1}{i!}(At)^i \tag{10.37}$$

II) Derivada de la matriz exponencial.

$$\frac{d}{dt} \exp (At)=A \exp (At)=(\exp (At))A \tag{10.38}$$

III) Linealidad del operador matriz exponencial bajo escalar.

$$\begin{aligned} \exp (At) \exp (A\tau) &= \left(\sum_{i=0}^{\infty} \frac{1}{i!}(At)^i\right) \left(\sum_{j=0}^{\infty} \frac{1}{j!}(A\tau)^j\right) \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} A^{i+j} \frac{t^i \tau^j}{i!j!} = \sum_{k=0}^{\infty} A^k \sum_{i=0}^k \frac{t^i \tau^{k-i}}{i!(k-i)!} \\ &= \sum_{k=0}^{\infty} A^k \frac{(t+\tau)^k}{k!} = \exp (A(t+\tau)) \end{aligned} \tag{10.39}$$

IV) Linealidad del operador matriz exponencial bajo matriz.

$$\exp ((A+B)t)=\exp (At) \exp (Bt) \iff AB=BA \tag{10.40}$$

V) Cambio de base.

Sean dos matrices similares A y \bar{A} , esto es, dos matrices relacionadas por un cambio de base, T matriz invertible, esto es $\bar{A}=T^{-1}AT$.

a) Las matrices exponenciales asociadas a las matrices A y \bar{A} también son similares. En efecto:

$$\begin{aligned} T^{-1} \exp (At)T &= T^{-1} \left(\sum_{i=0}^{\infty} \frac{1}{i!}(At)^i\right) T = \sum_{i=0}^{\infty} \frac{1}{i!} T^{-1}A^i T t^i \\ &= \sum_{i=0}^{\infty} \frac{1}{i!} (T^{-1}AT)^i t^i = \exp \bar{A}t \end{aligned}$$

b) Los valores propios son invariantes bajo cambio de base. En efecto:

$$\begin{aligned} \det (sI-\bar{A}) &= \det (sI-T^{-1}AT) = \det (sT^{-1}T-T^{-1}AT) \\ &= \det (T^{-1}(sI-A)T) = \det (T^{-1}) \det (sI-A) \det (T) \\ &= \frac{1}{\det (T)} \det (sI-A) \det (T) = \det (sI-A) \end{aligned}$$

- c) Las raíces de la matriz sistema son invariantes bajo cambio de base. En efecto, sea el sistema representado por:

$$\begin{aligned}\frac{d}{dt}x &= Ax + bu \\ y &= c^T x + du\end{aligned}$$

Sea el cambio de variable $x = T\bar{x}$, T invertible. Entonces:

$$\begin{aligned}T \frac{d}{dt} \bar{x} &= AT\bar{x} + bu \\ y &= c^T T\bar{x} + du\end{aligned}$$

\vdots

$$\begin{aligned}\frac{d}{dt} \bar{x} &= T^{-1}AT\bar{x} + T^{-1}bu \\ y &= c^T T\bar{x} + du\end{aligned}$$

\vdots

$$\begin{aligned}\frac{d}{dt} \bar{x} &= \bar{A}\bar{x} + \bar{b}u \\ y &= \bar{c}^T \bar{x} + du\end{aligned} \tag{10.41}$$

donde $\bar{A} = T^{-1}AT$, $\bar{b} = T^{-1}b$, $\bar{c}^T = c^T T$. La matriz sistema se puede escribir de la siguiente manera:

$$\Sigma = \begin{pmatrix} sI - A & b \\ -c^T & d \end{pmatrix} \Rightarrow \bar{\Sigma} = \begin{pmatrix} sI - \bar{A} & \bar{b} \\ -\bar{c}^T & d \end{pmatrix} \tag{10.42}$$

Notese que:

$$\bar{\Sigma} = \begin{pmatrix} sI - \bar{A} & \bar{b} \\ -\bar{c}^T & d \end{pmatrix} = \begin{pmatrix} T^{-1} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} sI - A & b \\ -c^T & d \end{pmatrix} \begin{pmatrix} T & 0 \\ 0 & 1 \end{pmatrix}$$

Por lo que:

$$\det \bar{\Sigma} = \det T^{-1} \det \Sigma \det T = \det \Sigma$$

VI) Forma de Jordan

Dada una matriz A , existe una matriz de cambio de base T , tal que:

$$T^{-1}AT = J = D + N \quad (10.43)$$

donde D es una matriz diagonal (conteniendo los valores propios) y N es una matriz nilpotente ($\exists \gamma \in \mathbb{N} \mid N^\gamma = 0$) de la forma:

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

$$N = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Note que $DN = ND$, por lo que:

$$\exp((D + N)t) = \exp(Dt) \exp(Nt) \quad (10.44)$$

donde:

$$\exp(Dt) = \begin{pmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & e^{\lambda_n t} \end{pmatrix}$$

$$\exp(Nt) = \sum_{i=0}^{\infty} \frac{1}{i!} (Nt)^i = \sum_{i=0}^{\gamma-1} \frac{1}{i!} (Nt)^i$$

VII) Teorema de Cayley-Hamilton

Toda transformación lineal A satisface su polinomio característico.

$$\Pi(s) = \det(sI - A) = s^n + \pi_1 s^{n-1} + \dots + \pi_{n-1} s + \pi_n \quad (10.45)$$

$$\Pi(A) = A^n + \pi_1 A^{n-1} + \dots + \pi_{n-1} A + \pi_n I = 0 \quad (10.46)$$

Una implicación directa es que la n -ésima potencia de una transformación lineal A , es una combinación lineal de sus potencias predecesoras.

$$A^n = -\pi_n I - \pi_{n-1} A - \dots - \pi_1 A^{n-1}$$

A su vez, esto implica:

$$\exp (At)=\sum_{i=0}^{\infty} \frac{1}{i!} A^i t^i=\sum_{i=0}^{n-1} \varphi(t) A^i \tag{10.47}$$

donde:

$$\varphi_i(t)=\sum_{j=0}^{\infty} \varphi_{ij} t^j$$

Capítulo 11

Controlabilidad y asignación de polos

Sea un sistema para la Ecuación Diferencial Ordinaria:

$$M \left(\frac{d}{dt} \right) y(t) = N \left(\frac{d}{dt} \right) u(t) \quad (11.1)$$

Sea la siguiente representación de estado de esta Ecuación Diferencial Ordinaria:

$$\begin{aligned} \frac{d}{dt} x &= Ax + bu \\ y &= c^T x + du \end{aligned}$$

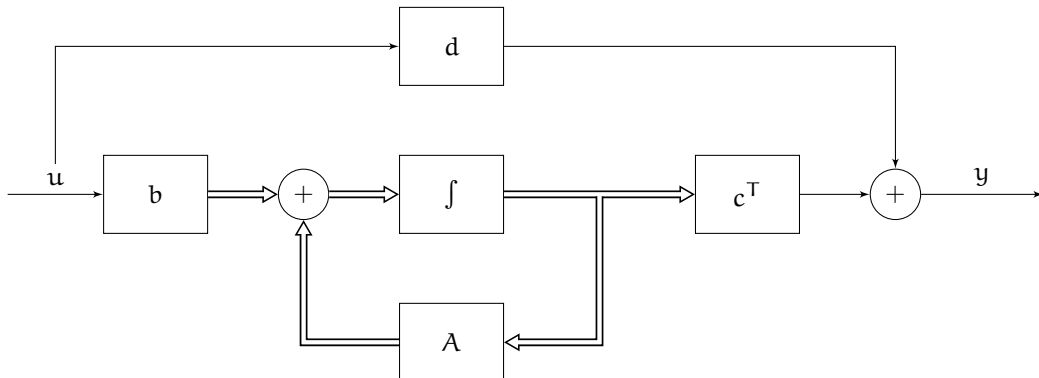
donde $x \in \mathbb{R}^n$ y $u, y \in \mathbb{R}$.

Problema. Se desea encontrar una ley de control $u = f(x)$, que nos permita asignar los polos a voluntad.

Sabemos que:

Polos. $\{s \in \mathbb{C} \mid M(s) = 0\} = \{s \in \mathbb{C} \mid \det(sI - A)\}$

Para resolver este problema hay que investigar el concepto estructural de la alcanzabilidad.



Alcanzabilidad y Controlabilidad

- Una representación de estado se dice controlable, si para cualquier condición inicial, $x(0) = x_0 \in \mathbb{R}^n$, existe una trayectoria, $x(\cdot)$, solución de la ecuación de estado, tal que en tiempo finito $t_f \in \mathbb{R}$ se llega al origen ($x(t_f) = 0$).
- Una representación de estado se dice alcanzable, si para cualquier punto $x \in \mathbb{R}$, existe una trayectoria, $x(\cdot)$, solución de la ecuación de estado, tal que en tiempo finito $t_f \in \mathbb{R}$ se llega a un punto cualquiera ($x(t_f) = x_f$) desde el origen.

En los sistemas lineales estas dos propiedades están mutuamente implicadas, por lo que se les trata indistinguiblemente. Pero en general:

$$\text{Alcanzabilidad} \implies \text{Controlabilidad} \tag{11.2}$$

La solución temporal de $\frac{dx}{dt} = Ax + bu$ con $x(0) = 0$ es:

$$x(t) = \int_0^t \exp(A(t-\tau))bu(\tau)d\tau \tag{11.3}$$

del teorema de Cayley-Hamilton se tiene:

$$\exp(At) = \sum_{i=0}^{\infty} \frac{1}{i!} A^i t^i = \sum_{i=0}^{n-1} \varphi_i(t) A^i \tag{11.4}$$

donde $\varphi_i(t) = \sum_{j=0}^{\infty} \varphi_{ij} t^j$, $\varphi_{ij} \in \mathbb{R}$, $n \in \mathbb{N}$, $j \in \mathbb{Z}^+$. Por lo anterior, tenemos:

$$\begin{aligned} x(t) &= \sum_{i=0}^{n-1} \psi_i(t) A^i b \\ x(t) &= (b \quad Ab \quad \dots \quad A^{n-1}b) \begin{pmatrix} \psi_0(t) \\ \psi_1(t) \\ \vdots \\ \psi_{n-1}(t) \end{pmatrix} \end{aligned} \tag{11.5}$$

donde $\psi_i(t) = \int_0^t \varphi_i(t-\tau)u(\tau)d\tau$.

Entonces una condición necesaria para que $x(t_f) = x_f \quad \forall x \in \mathbb{R}, \forall t_f \in \mathbb{R}, t_f > 0$, es que la matriz de controlabilidad

$$C_{(A,b)} = (b \quad Ab \quad \dots \quad A^{n-1}b) \tag{11.6}$$

sea de rango pleno por filas, de lo contrario existen componentes de $x(t)$ que siempre serán nulos. En nuestro caso particular ($y, u \in \mathbb{R}^n$):

$$\det C_{(A,b)} \neq 0 \tag{11.7}$$

Si la matriz de controlabilidad $C_{(A,b)}$ es de rango pleno por filas, entonces el gramiano de controlabilidad es invertible.¹

¹ Aquí se está abusando de la notación, ya que el gramiano de controlabilidad corresponde al caso en que $t \rightarrow \infty$

$$W = \int_0^t \exp(A\sigma)bb^t \exp(A^t\sigma)d\sigma \quad t > 0, \sigma = t - \tau \quad (11.8)$$

Entonces, con la siguiente ley de control se tiene:

$$u(t) = b^t \exp(A^t(t_f - t))W_{t_f}^{-1}x_f \quad (11.9)$$

Por lo que si sustituimos t_f en la solución para $x(t)$:

$$\begin{aligned} x(t_f) &= \int_0^{t_f} \exp(A(t - \tau))bu(\tau)d\tau \\ &= \int_0^{t_f} \exp(A(t - \tau))bb^t \exp(A^t(t_f - \tau))W_{t_f}^{-1}x_f d\tau \\ &= \int_{t_f}^0 \exp(A\sigma)bb^t \exp(A^t\sigma)d\sigma W_{t_f}^{-1}x_f \\ &= W_{t_f}W_{t_f}^{-1}x_f = x_f \quad (11.10) \end{aligned}$$

- Por lo que una condición suficiente y necesaria para que la ecuación de estado sea alcanzable (y por lo tanto controlable), es que su matriz de controlabilidad, $C_{(A,b)}$, sea de rango pleno por filas.
- Cuando la matriz de controlabilidad es de rango pleno por filas, se dice que el par (A, b) es controlable.

Asignación de polos

Sea la ecuación de estado controlable, es decir $\frac{dx}{dt} = Ax + bu$ con $b \neq 0$, $\det (C_{(A,b)}) \neq 0$, $\Pi(s)$ y $\alpha(s)$ los polinomios característico y mínimo de A respectivamente.

$$\Pi(s) = \det (sI - A) \quad \text{grado } \Pi(s) = n$$

$\alpha(s)$ es el polinomio de menor grado tal que $\alpha(A) = 0$

Sea $\kappa = \text{grado } \alpha(s)$, donde obviamente $1 \leq \kappa \leq n$.

$$\alpha(s) = s^\kappa + (a_\kappa + a_{\kappa-1}s + \cdots + a_1s^{\kappa-1})$$

$$\alpha(A) = A^\kappa + (a_\kappa + a_{\kappa-1}A + \cdots + a_1A^{\kappa-1})$$

Sean α_i con $i \in \{0, 1, \dots, \kappa\}$, los polinomios mónicos auxiliares tales que:

$$\begin{aligned} \alpha_0(s) &= \alpha(s) \\ \alpha_1(s) &= s^{\kappa-1} + (a_{\kappa-1} + a_{\kappa-2}s + \cdots + a_1s^{\kappa-2}) \\ &\vdots \\ \alpha_{\kappa-1}(s) &= s + a_1 \\ \alpha_\kappa(s) &= 1 \end{aligned}$$

en donde, por definición, $\alpha_1(A) \neq 0$ y $\alpha_0(A) = 0$.
 Sea $b \neq 0$, un vector en \mathbb{R}^n tal que su polinomio mínimo coincide con $\alpha(s)$.

$$\begin{aligned} \alpha_i(A)b &\neq 0 \quad i \in \{1, 2, \dots, \kappa\} \\ \alpha_0(A)b &= 0 \end{aligned}$$

$$\vdots$$

$$\begin{aligned} (A^{\kappa-1} + (a_{\kappa-1} + a_{\kappa-2}A + \cdots + a_1A^{\kappa-2}))b &\neq 0 \\ \alpha_0(A)b &= 0 \end{aligned}$$

$$\vdots$$

$$\begin{aligned} (b \quad Ab \quad \dots \quad A^{\kappa-1}b) \begin{pmatrix} a_{\kappa-1} \\ a_{\kappa-2} \\ \vdots \\ a_1 \end{pmatrix} &\neq 0 \\ \alpha_0(A)b &= 0 \end{aligned}$$

Suponga que el par (A, b) es controlable, por lo tanto $\det C_{(A,b)} \neq 0$, entonces:

$$(b \quad Ab \quad \dots \quad A^{n-1}b)v \neq 0 \quad \forall v \neq 0 \quad \therefore \kappa = n \quad (11.11)$$

Por lo que el polinomio mínimo y el polinomio característico coinciden cuando el par (A, b) es controlable. Definimos la base:

$$\begin{aligned} e_n &= \alpha_n(A)b = b \\ e_{n-1} &= \alpha_{n-1}(A)b = (A + a_1I)b = Ae_n + a_1e_n \\ e_{n-2} &= \alpha_{n-2}(A)b = (A^2 + (a_2I + a_1A))b = A(A + a_1I)b + a_2b = Ae_{n-1} + a_2e_n \\ &\vdots \\ e_1 &= \alpha_1(A)b = Ae_2 + a_{n-1}e_n \end{aligned} \quad (11.12)$$

Note que sustituyendo A , tenemos:

$$\alpha(A)b = (A^n + (a_nI + a_{n-1}A + \dots + a_1A^{n-1}))b = Ae_1 + a_ne_n = 0 \quad (11.13)$$

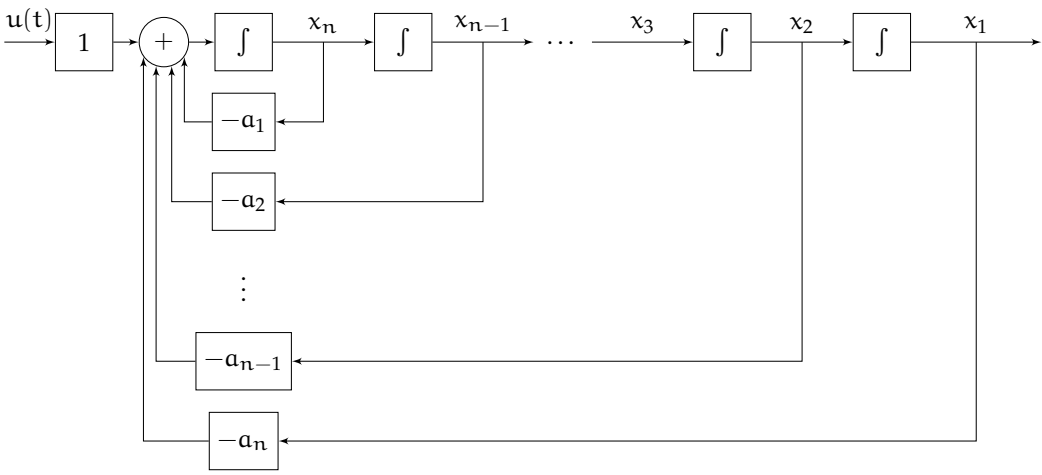
Por lo que:

$$\begin{aligned} e_n &= b \\ Ae_n &= e_{n-1} - a_1e_n \\ Ae_{n-1} &= e_{n-2} - a_2e_n \\ &\vdots \\ Ae_2 &= e_1 - a_{n-1}e_n \\ Ae_1 &= -a_ne_n \end{aligned} \quad (11.14)$$

Entonces, bajo la base definida, las transformaciones lineales tienen la siguiente forma:

$$A_c = [A]_{\{e_1, e_2, \dots, e_n\}} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \dots & -a_3 & -a_2 & -a_1 \end{pmatrix} \quad (11.15)$$

$$b_c = [b]_{\{e_1, e_2, \dots, e_n\}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad (11.16)$$



Por lo tanto:

$$\frac{d}{dt}x_c = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \dots & -a_3 & -a_2 & -a_1 \end{pmatrix} x_c + \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix} u \quad (11.17)$$

Observaciones.

1. Polinomio Característico

$$\det(sI - A_c) = s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n = \Pi(s)$$

2. Retroalimentación de Estado

Sea $u = f_c x_c + v$, donde $f_c = (a_n - \bar{a}_n)(a_{n-1} - \bar{a}_{n-1}) \dots (a_1 - \bar{a}_1)$. Entonces el sistema de lazo cerrado es:

$$\frac{d}{dt}x_c = A_{f_c} x_c + b_c v \quad (11.18)$$

donde $A_{f_c} = A_c + b_c f_c$, es decir:

$$A_{f_c} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \\ -\bar{a}_n & -\bar{a}_{n-1} & -\bar{a}_{n-2} & \dots & -\bar{a}_3 & -\bar{a}_2 & -\bar{a}_1 \end{pmatrix} \quad (11.19)$$

y su polinomio característico es $\det (sI - A_{f_c}) = s^n + \bar{a}_1 s^{n-1} + \cdots + \bar{a}_{n-1} s + \bar{a}_n$.

Propiedades de la matriz de controlabilidad

1. Matriz de controlabilidad del par (A_c, b_c) .

$$C_{(A_c, b_c)} = \begin{pmatrix} 0 & 0 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 1 & \dots & * \\ 0 & 1 & * & \dots & * \\ 1 & * & * & \dots & * \end{pmatrix} \quad (11.20)$$

lo que implica:

$$\det C_{(A_c, b_c)} = \pm 1 \quad (11.21)$$

2. Invarianza de la matriz de controlabilidad bajo cambio de base.

Sea el par (A, b) controlable, sea T una matriz de cambio de base y sean $A_1 = T^{-1}AT$ y $b_1 = T^{-1}b$ las matrices de nuestra nueva base.

$$\begin{aligned} C_{(A_1, b_1)} &= (b_1 \quad A_1 b_1 \quad \dots \quad A_1^{n-1} b_1) = \\ &= (T^{-1}b \quad T^{-1}AT T^{-1}b \quad \dots \quad (T^{-1}AT \dots T^{-1}AT) T^{-1}b) = \\ &= (T^{-1}b \quad T^{-1}Ab \quad \dots \quad T^{-1}A^{n-1}b) = \\ &= T^{-1} (b \quad Ab \quad \dots \quad A^{n-1}b) = T^{-1} C_{(A, b)} \end{aligned}$$

por lo que, podemos notar la siguiente correspondencia:

$$C_{(A_1, b_1)} = \frac{C_{(A, b)}}{T}$$

Mas notablemente podemos notar una manera de calcular la transformación lineal a una forma controlable.

$$T = C_{(A, b)} C_{(A_1, b_1)}^{-1} \quad (11.22)$$

3. Invarianza de la matriz de controlabilidad bajo retroalimentación de estado, $u = f^T x + v$.

Sea $A_f = A + bf^T$ la matriz A del sistema bajo la retroalimentación de estado $u = f^T x + v$. Tendremos que la matriz de controlabilidad de este sistema retroalimentado será:

$$\begin{aligned} C_{(A_f, b)} &= (b \quad A_f b \quad \dots \quad A_f^{n-1} b) = \\ &= (b \quad (A + bf^T) b \quad \dots \quad (A + bf^T)^{n-1} b) \end{aligned}$$

en donde podemos notar que los terminos van obteniendo la siguiente forma:

$$\begin{aligned}
(A + bf^T)b &= Ab + b(f^Tb) = Ab + k_1b \\
(A + bf^T)^2b &= A^2b + k_1Ab + b(f^TAb + k_1(f^Tb)) \\
&= A^2b + k_1Ab + k_2b \\
&\vdots \\
(A + bf^T)^{n-1}b &= A^{n-1}b + k_1A^{n-2}b + \dots + k_{n-2}Ab + k_{n-1}b
\end{aligned}$$

en donde los terminos k_i estan relacionados unicamente con f^T y b , y dejan de fuera a un termino b , por lo que es inmediato ver que lo podemos reescribir de la siguiente manera:

$$C_{(A_f, b)} = \begin{pmatrix} b & Ab & \dots & A^{n-1}b \end{pmatrix} \mathbb{X}$$

$$C_{(A_f, b)} = C_{(A, b)} \mathbb{X} \quad (11.23)$$

donde \mathbb{X} toma la forma:

$$\mathbb{X} = \begin{pmatrix} 1 & k_1 & k_2 & \dots & k_{n-1} \\ 0 & 1 & k_1 & \dots & k_{n-2} \\ 0 & 0 & 1 & \dots & k_{n-3} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

lo cual implica que $\det \mathbb{X} = \pm 1$, es decir:

$$\det C_{(A, b)} \neq 0 \implies \det C_{(A_f, b)} \neq 0$$

en particular nosotros tenemos que:

$$\det C_{(A, b)} = \det C_{(A_f, b)} \quad (11.24)$$

Dada la invarianza de la matriz de controlabilidad ($C_{(A, b)}$) bajo cambio de base y retroalimentación de estado, Brunovskii estudió la controlabilidad de los sistemas lineales con todos sus valores propios (polos) en el origen²:

$$(A_{Br} \quad b_{Br}) = \left(\begin{array}{cccc|ccc} 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{array} \right) \quad (11.25)$$

²El teorema de Brunovskii en realidad esta redactado para sistemas multientradas, y se expresa en matrices diagonales por bloques de tamaño $k_i \times (k_i + 1)$, con $\sum_{i=0}^n k_i = n$, donde k_i son los indices de controlabilidad

A los índices k_i de los polinomios mínimos se les denomina índices de controlabilidad. En nuestro caso particular, existe solamente un índice de controlabilidad; $k_i = n$.

4. Invarianza de los ceros del sistema bajo retroalimentación de estado.

Sea la matriz sistema del sistema en lazo abierto la siguiente:

$$\Sigma(s) = \begin{pmatrix} sI - A & b \\ -c^T & d \end{pmatrix}$$

entonces, la matriz sistema bajo la retroalimentación será:

$$\Sigma_{lc}(s) = \begin{pmatrix} sI - (A + bf^T) & b \\ -(c^T + df^T) & d \end{pmatrix}$$

notando que:

$$\begin{pmatrix} sI - A & b \\ -c^T & d \end{pmatrix} \begin{pmatrix} I & 0 \\ -f^T & 1 \end{pmatrix} = \begin{pmatrix} sI - (A + bf^T) & b \\ -(c^T + df^T) & d \end{pmatrix}$$

por lo que:

$$\det \Sigma(s) = \det \Sigma_{lc}(s) \quad (11.26)$$

Se concluye que la retroalimentación de estado no afecta a los ceros del sistema; solo puede modificar a los polos controlables.

Formas canónicas

Sea un sistema lineal invariante en el tiempo, una entrada, una salida (SISO), descrito por la siguiente Ecuación Diferencial Ordinaria (EDO):

$$\left(\frac{d^n}{dt^n} + a_1 \frac{d^{n-1}}{dt^{n-1}} + \dots + a_{n-1} \frac{d}{dt} + a_n \right) y(t) = \left(b_n + b_{n-1} \frac{d}{dt} + \dots + b_1 \frac{d^{n-1}}{dt^{n-1}} \right) u(t) \quad (11.27)$$

Forma canónica controlador

$$\frac{d}{dt} x_c = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \dots & -a_3 & -a_2 & -a_1 \end{pmatrix} x_c + \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix} u \quad (11.28)$$

$$y = (b_n \quad b_{n-1} \quad b_{n-2} \quad \dots \quad b_2 \quad b_1) x_c \quad (11.29)$$

Polos.

$$\det(sI - A_c) = s^n + a_1 s^{n-1} + a_2 s^{n-2} + \dots + a_{n-1} s + a_n \quad (11.30)$$

Ceros.

$$\det \Sigma(s) = b_1 s^{n-1} + \dots + b_{n-1} s + b_n \quad (11.31)$$

Matriz de Controlabilidad.

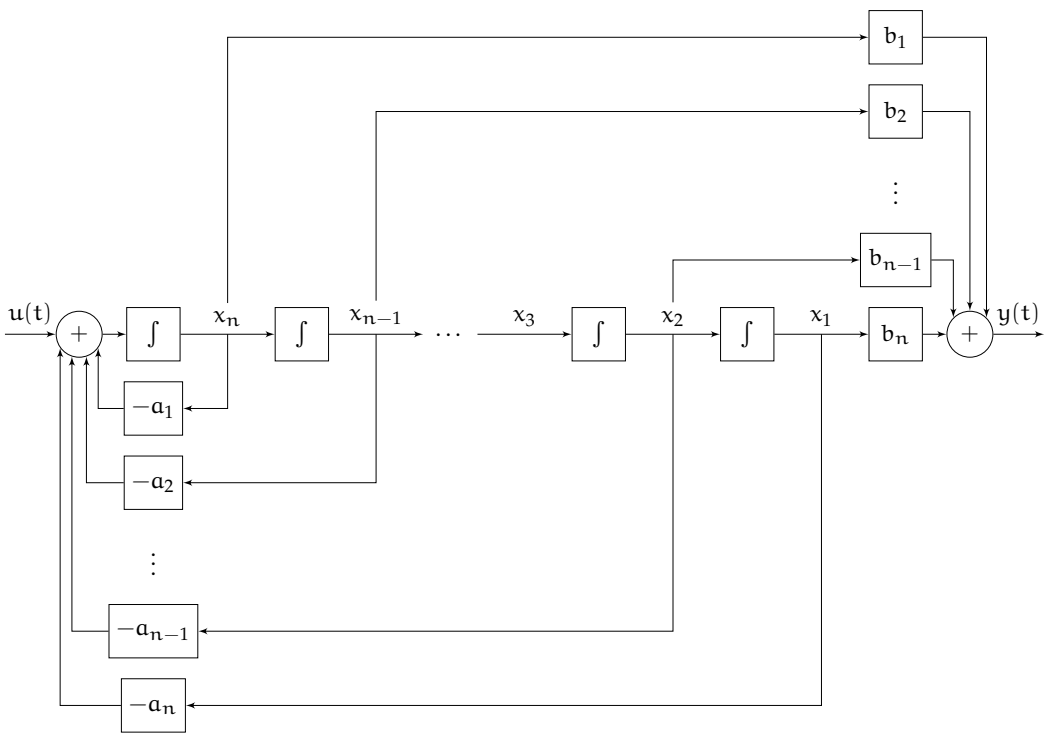
$$C_{(A_c, b_c)} = (b_c \quad A_c b_c \quad \dots \quad A_c^{n-1} b_c) \implies \det C_{(A_c, b_c)} = 1 \quad (11.32)$$

Forma canónica controlabilidad

$$\frac{d}{dt} x_c = \begin{pmatrix} -a_1 & 1 & 0 & \dots & 0 & 0 & 0 \\ -a_2 & 0 & 1 & \dots & 0 & 0 & 0 \\ -a_3 & 0 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ -a_{n-2} & 0 & 0 & \dots & 0 & 1 & 0 \\ -a_{n-1} & 0 & 0 & \dots & 0 & 0 & 1 \\ -a_n & 0 & 0 & \dots & 0 & 0 & 0 \end{pmatrix} x_c + \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix} u \quad (11.33)$$

$$y = (\beta_n \quad \beta_{n-1} \quad \beta_{n-2} \quad \dots \quad \beta_2 \quad \beta_1) x_c \quad (11.34)$$

donde:



$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_{n-1} \\ \beta_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ -a_1 & 1 & 0 & \dots & 0 & 0 \\ -a_2 & -a_1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ -a_{n-3} & -a_{n-4} & -a_{n-5} & \dots & 0 & 0 \\ -a_{n-2} & -a_{n-3} & -a_{n-4} & \dots & 1 & 0 \\ -a_{n-1} & -a_{n-2} & -a_{n-3} & \dots & -a_1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_{n-2} \\ b_{n-1} \\ b_n \end{pmatrix} \quad (11.35)$$

Polos.

$$\det(sI - A_{co}) = s^n + a_1 s^{n-1} + a_2 s^{n-2} + \dots + a_{n-1} s + a_n \quad (11.36)$$

Ceros.

$$\det \Sigma(s) = b_1 s^{n-1} + \dots + b_{n-1} s + b_n \quad (11.37)$$

Matriz de Controlabilidad.

$$C_{(A_{co}, b_{co})} = (b_{co} \quad A_{co} b_{co} \quad \dots \quad A_{co}^{n-1} b_{co}) \implies \det C_{(A_{co}, b_{co})} = \pm 1 \quad (11.38)$$

Capítulo 12

Inobservabilidad y observador de estado

Sea la siguiente representación de estado:

$$\begin{aligned}\frac{d}{dt}x &= Ax + bu \\ y &= c^T x + du\end{aligned}\tag{12.1}$$

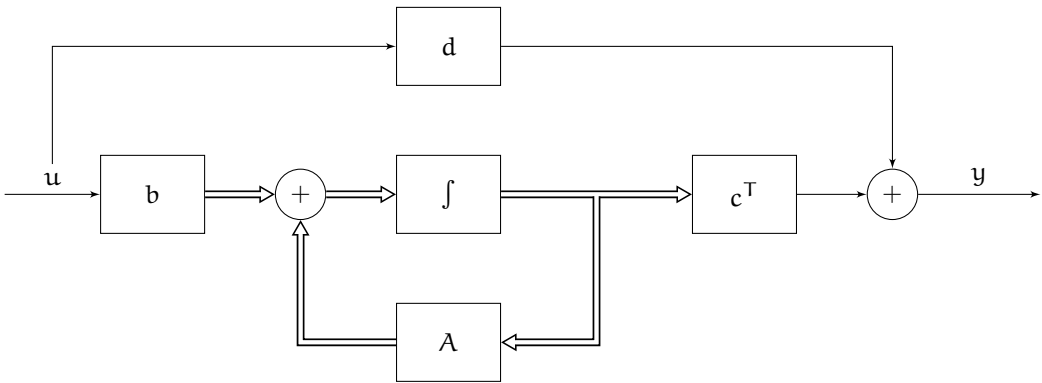
para el siguiente sistema:

donde $x \in \mathbb{R}^n$ es el estado, $u(t) \in \mathbb{R}^n$ y $y(t) \in \mathbb{R}^n$ sean la entrada y salida respectivamente; siendo la condición inicial del estado $x(0) = x_0 \in \mathbb{R}^n$. La solución esta descrita por:

$$x(t) = \exp (At)x_0 + \int_0^t \exp (A(t-\tau))bu(\tau)d\tau$$

Problema. Sea la representación de estado 12.1, donde el estado x no esta disponible. Se desea reconstruir el estado x , para poder aplicar una retroalimentación de estado.

$$u = f^T x + v\tag{12.2}$$



Para resolver este problema, hay que investigar el concepto estructural de la inobservabilidad.

Observabilidad e inobservabilidad

Una representación de estado se dice observable si dadas las trayectorias de salida, $y(t)$, y entrada, $u(t)$, en un horizonte de tiempo finito, $t_1 \in \mathbb{R}$, $t_1 > 0$, existe una función $\mathbb{F}(t, u, y)$, tal que:

$$\mathbb{F}(t_1, u(t), y(t)) = x(0) \quad t \in [0, t_1] \tag{12.3}$$

Una representación de estado se dice inobservable, si no es observable.

$$\begin{aligned} x(t) &= \exp (At)x_0 + \int_0^t \exp (A(t-\tau))bu(\tau) d\tau \\ y(t) &= c^T \exp (At)x_0 + \int_0^t c^T \exp (A(t-\tau))bu(\tau) d\tau + du(t) \end{aligned} \tag{12.4}$$

Sabemos del teorema de Caley-Hamilton que:

$$\exp (At) = \sum_{i=0}^{n-1} \varphi(t)A^i \tag{12.5}$$

donde:

$$\varphi_i(t) = \sum_{j=0}^{\infty} \varphi_{ij}t^j, \quad \varphi_{ij} \in \mathbb{R}, \quad i \in \{0, 1, \dots, n-1\}, \quad j \in \mathbb{Z}^+$$

Si juntamos las ecuaciones 12.4 y 12.5 obtendremos:

$$\begin{aligned} y(t_1) &= \sum_{i=0}^{n-1} \varphi_i(t_1)c^T A^i x_0 + \\ &\quad \int_0^{t_1} c^T \exp (A(t-\tau))bu(\tau) d\tau + du(t_1) = \\ &\quad (\varphi_0(t_1) \quad \varphi_1(t_1) \quad \dots \quad \varphi_{n-1}(t_1)) \begin{pmatrix} c^T \\ c^T A \\ \vdots \\ c^T A^{n-1} \end{pmatrix} x_0 + \\ &\quad \int_0^{t_1} c^T \exp (A(t-\tau))bu(\tau) d\tau + du(t_1) \\ \varphi^T(t_1)\mathcal{O}_{(c^T, A)}x(0) &= y(t_1) - \int_0^{t_1} c^T \exp (A(t-\tau))bu(\tau) d\tau + du(t_1) \end{aligned} \tag{12.6}$$

siendo el lado derecho, la función \mathbb{F} . Entonces, una condición necesai para que se pueda inferir cualquier condición inicial del estado $x(0) = x_0$, a partir de las trayectorias de salida, $y(t)$, y de entrada $u(t)$, en el horizonte de tiempo, es que la matriz de observabilidad:

$$\mathcal{O}_{(c^T, A)} = \begin{pmatrix} c^T \\ c^T A \\ \vdots \\ c^T A^{n-1} \end{pmatrix} \quad (12.7)$$

sea de rango pleno por columnas.

En nuestro caso particular, como la entrada y la salida estan en \mathbb{R} , esta condición es:

$$\det \mathcal{O}_{(c^T, A)} \neq 0$$

En efecto, si $\mathcal{O}_{(c^T, A)}$, no es de rango pleno por columna, existe una transformación T , invertible, tal que:

$$\mathcal{O}_{(c^T, A)} T^{-1} = (\mathbb{X} \quad 0)$$

siendo \mathbb{X} una matriz de rango plano por columnas.

Haciendo el cambio de base, $\bar{x} = Tx$, se obtiene de la ecuación 12.6:

$$\begin{aligned} \varphi^T(t_1) \mathcal{O}_{(c^T, A)} T^{-1} \bar{x}(0) &= \bar{\mathbb{F}}(t_1, u, y) \\ \varphi^T(t_1) (\mathbb{X} \quad 0) \begin{pmatrix} \bar{x}_1(0) \\ \bar{x}_2(0) \end{pmatrix} &= \bar{\mathbb{F}}(t_1, u, y) \\ \varphi^T(t_1) \mathbb{X} \bar{x}_1(0) &= \bar{\mathbb{F}}(t_1, u, y) \end{aligned}$$

por lo que no es posible determinar la segunda parte de componentes, $\bar{x}_2(0)$, a partir de $\bar{\mathbb{F}}(t_1, u, y)$.

Si la matriz de observabilidad es de rango pleno por columnas, entonces:

$$\ker \mathcal{O}_{(c^T, A)} = 0$$

y para nuestro caso particular, $u, y \in \mathbb{R}$:

$$\det \mathcal{O}_{(c^T, A)} \neq 0$$

es decir, $\mathcal{O}_{(c^T, A)}$ es invertible.

De la representación de estado en la ecuación 12.1 se tiene:

$$y = c^T x + du$$

$$\begin{aligned} \frac{dy}{dt} &= c^T \frac{dx}{dt} + d \frac{du}{dt} \\ &= c^T Ax + c^T bu + d \frac{du}{dt} \end{aligned}$$

$$\begin{aligned} \frac{d^2y}{dt^2} &= c^T A \frac{dx}{dt} + c^T b \frac{du}{dt} + d \frac{d^2u}{dt^2} \\ &= c^T A^2 x + c^T A bu + c^T b \frac{du}{dt} + d \frac{d^2u}{dt^2} \end{aligned}$$

\vdots

$$\frac{d^{n-1}y}{dt^{n-1}} = c^T A^{n-1}x + \sum_{i=0}^{n-2} c^T A^i b \frac{d^{n-2-i}u}{dt^{n-2-i}}$$

por lo que tendremos que:

$$\begin{pmatrix} 1 \\ \frac{d}{dt} \\ \vdots \\ \frac{d^{n-1}}{dt^{n-1}} \end{pmatrix} y = \begin{pmatrix} c^T \\ c^T A \\ \vdots \\ c^T A^{n-1} \end{pmatrix} x + \begin{pmatrix} 0 + d \\ b + d \frac{d}{dt} \\ \vdots \\ \sum_{i=0}^{n-2} c^T A^i b \frac{d^{n-2-i}}{dt^{n-2-i}} + d \frac{d^{n-1}}{dt^{n-1}} \end{pmatrix} u$$

o escrito de otra manera:

$$\Delta \left(\frac{d}{dt} \right) y = \mathcal{O}_{(c^T, A)} x + \Gamma \left(\frac{d}{dt} \right) u$$

lo cual implica:

$$x = \mathcal{O}_{(c^T, A)}^{-1} \left[\Delta \left(\frac{d}{dt} \right) y - \Gamma \left(\frac{d}{dt} \right) u \right] \tag{12.8}$$

Por lo que es una condición necesaria y suficiente, para que la representación de estado sea observable que su matriz de observabilidad, $\mathcal{O}_{(c^T, A)}$, sea de rango pleno por columnas.

Cuando la matriz de observabilidad es de rango pleno por columnas, se dice que el par (c^T, A) es observable.

Dualidad

La operación matricial "transpuesta", establece una dualidad entre la observabilidad y la controlabilidad. En efecto,

$$\mathcal{O}_{(c^T, A)} = \begin{pmatrix} c^T \\ c^T A \\ \vdots \\ c^T A^{n-1} \end{pmatrix} \tag{12.9}$$

$$\mathcal{O}_{(c^T, A)}^T = \begin{pmatrix} c^T \\ c^T A \\ \vdots \\ c^T A^{n-1} \end{pmatrix}^T = \begin{pmatrix} c & A^T c & \dots & (A^T)^{n-1} c \end{pmatrix} = C_{(A^T, c)} \tag{12.10}$$

$$C_{(A, b)}^T = \begin{pmatrix} b & Ab & \dots & A^{n-1}b \end{pmatrix}^T = \begin{pmatrix} b^T \\ b^T A^T \\ \vdots \\ b^T (A^T)^{n-1} \end{pmatrix} = \mathcal{O}_{(b^T, A^T)} \tag{12.11}$$

note tambien que la función de transferencia es una funcion continua en \mathbb{R}^1 , por lo que:

$$FT = FT^T$$

$$b^T \left(sI - A^T \right)^{-1} c + d = c^T (sI - A)^{-1} b + d \tag{12.12}$$

ademas, el polinomio caracteristico es el mismo:

$$\det (sI - A) = \det (sI - A^T) \implies \sigma(A) = \sigma(A^T) \tag{12.13}$$

Por lo que obtenemos la siguiente dualidad:

$$A \leftrightarrow A^T$$

$$b \leftrightarrow c$$

$$c^T \leftrightarrow b^T$$

$$d \leftrightarrow b$$

$$f^T \leftrightarrow k^T$$

Propiedades de la matriz de observabilidad

Dada la dualidad entre observabilidad y controlabilidad, todos los resultados de controlabilidad del par (A, b) son extrapolables a la observabilidad del par (c^T, A) .

1. Asignación de los valores propios (polos), mediante la inyección de salida.

a) Si $\det C_{(A,b)} \neq 0$, entonces dado un conjunto simétrico con respecto al eje real de n números complejos, Λ , existe un vector $f \in \mathbb{R}^n$, tal que:

$$\sigma(A + bf^T) = \Lambda$$

b) Si $\det \mathcal{O}_{(c^T,A)} \neq 0$, entonces dado un conjunto simétrico con respecto al eje real de n números complejos, Λ , existe un vector $k \in \mathbb{R}^n$, tal que:

$$\sigma(A + kc^T) = \Lambda$$

$$\sigma(A^T + ck^T) = \Lambda$$

2. Invarianza de la matriz de observabilidad bajo cambio de base

Sea el siguiente cambio de base:

$$A_1 = T^{-1}AT \tag{12.14}$$

$$c_1^T = c^T T \tag{12.15}$$

dado un cambio de base T invertible; entonces tendremos lo siguiente:

$$\mathcal{O}_{(c_1^T,A_1)} = \mathcal{O}_{(c^T,A)} T^{-1} \tag{12.16}$$

de donde podemos notar que:

$$T = \mathcal{O}_{(c_1^T,A_1)}^{-1} \mathcal{O}_{(c^T,A)} \tag{12.17}$$

siempre que el par (c^T, A) sea observable.

3. Invarianza de la matriz de observabilidad bajo inyección de salida

4. Invarianza de los ceros del sistema bajo inyección de salida

Formas canónicas

Forma canónica observador

Forma canónica observabilidad

Observador de estado

Capítulo 13

Principio de Separación

Capítulo 14

Estabilidad de Lyapunov

Dada la siguiente representación de estado:

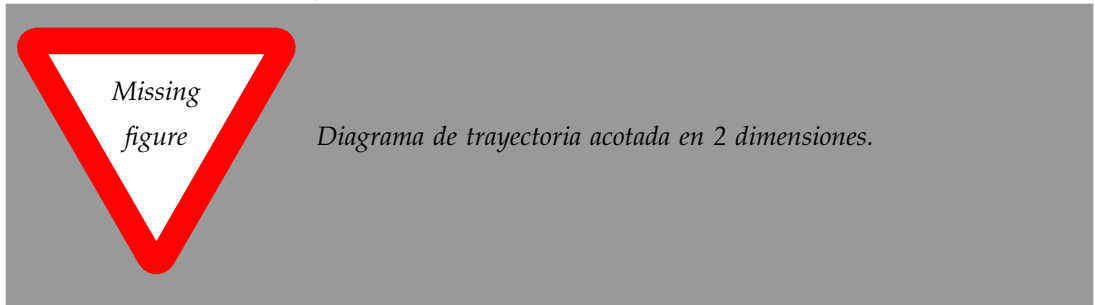
$$\frac{dx(t)}{dt} = Ax(t) + bu(t) \tag{14.1}$$

con $x(0) = x_0 \in \mathbb{R}^n$.

Definición 14.1. El sistema representado por la ecuación 14.1 es estable si para cada $\epsilon > 0$, existe un $\delta = \delta(\epsilon)$ tal que:

$$\|x_0\| < \delta(\epsilon) \implies \|x(t)\| < \epsilon \quad \forall t \geq 0$$

Si no es estable, decimos que es inestable.



Teorema 14.1. Una matriz A es Hurwitz estable, es decir $\Re\{\lambda(A)\} < 0$, si y solo si para cualquier matriz simétrica definida positiva dada, Q , existe una matriz simétrica definida positiva, P , que satisface:

$$A^T P + PA = -Q \tag{14.2}$$

Nota 14.1. Una matriz $H \in \mathbb{C}^{n \times n}$, se dice Hermitiana si su transpuesta conjugada es ella misma, $H^* = H$. Si $H \in \mathbb{R}^{n \times n}$, se dice simétrica si su transpuesta es ella misma, $H^T = H$.

De estas matrices, podemos notar ciertas propiedades:

1. Todos sus valores propios son reales.
2. Cuando los valores propios de H son todos positivos o negativos, se dice que H es definida positiva o negativa, y se escribe $H > 0$ o $H < 0$ respectivamente.

3. Cuando los valores propios de H son todos no negativos o no positivos, se dice que H es semidefinida positiva o semidefinida negativa y se escribe $H \geq 0$ o $H \leq 0$ respectivamente.

4. Desigualdad de Raleigh

Dada H Hermitiana:

$$\lambda_{\min}(H)x^*x \leq x^*Hx \leq \lambda_{\max}(H)x^*x \quad \forall x \in \mathbb{C}^n$$

Dada H simétrica:

$$\lambda_{\min}(H)x^T x \leq x^T Hx \leq \lambda_{\max}(H)x^T x \quad \forall x \in \mathbb{C}^n$$

5. H es semidefinida positiva, si y solo si, puede escribirse de la forma factorizada:

$$H = G^* G$$

para alguna matriz G , conocida como raíz cuadrada de H , también denotada por $H_{1/2}$, \sqrt{H} , $H^{1/2}$, por lo que la factorización queda como sigue:

$$H = H_{1/2}^* H_{1/2}$$

Cuando H es definida positiva, $H_{1/2}$ es una matriz de rango pleno.

Demostración. Sea la función de Lyapunov:

$$V(x(t)) = x^T P x(t) \quad \forall t \geq 0 \quad (14.3)$$

con $P = P^T > 0$.

Derivando a la ecuación 14.3 con respecto del tiempo, a lo largo de las trayectorias solución de la ecuación 14.1, con $u = 0$, se tiene:

$$\begin{aligned} \frac{dV}{dt} &= \frac{dx(t)}{dt}^T P x(t) + x^T(t) P \frac{dx(t)}{dt} \\ &= x^T(t) A^T P x(t) + x^T(t) P A x(t) \\ &= x^T(t) (A^T P + P A) x(t) \\ &= -x^T(t) Q x(t) \end{aligned} \quad (14.4)$$

Por otro lado, de la ecuación 14.3 se tiene:

$$\lambda_{\min}(P)x^T(t)x(t) \leq V(x(t)) \leq \lambda_{\max}(P)x^T(t)x(t)$$

por lo que:

$$0 \leq \frac{V(x(t))}{\lambda_{\min}(P)} \leq x^T(t)x(t) \leq \frac{V(x(t))}{\lambda_{\max}(P)} \quad (14.5)$$

Entonces, de las ecuaciones 14.4 y 14.5 se obtiene:

$$\frac{dV(x(t))}{dt} \leq -\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)} V(x(t)) \quad (14.6)$$

De manera análoga

$$\lambda_{\min}(Q)x^T(t)x(t) \leq x^T(t)Qx(t) \leq \lambda_{\max}(Q)x^T(t)x(t)$$

Si integramos la ecuación 14.6 tendremos:

$$\int_0^t \frac{dV(x(\tau))}{d\tau} d\tau \leq -\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)} \int_0^t V(x(\tau)) d\tau$$

para lo cual necesitamos el lema de Bellman - Grönwall.

Nota 14.2.

$$u(t) \leq c + \int_0^t K(\tau)u(\tau) d\tau \implies u(t) \leq c \exp\left(\int_0^t K(\tau) d\tau\right) \quad \forall t \geq 0$$

por lo tanto, podemos ver que:

$$V(x(t)) \leq V(x(0)) - \frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)} \int_0^t V(x(\tau)) d\tau \quad (14.7)$$

y aplicando el lema de Bellman - Grönwall aquí:

$$V(x(t)) \leq V(x(0)) \exp\left(-\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)} t\right) \quad \forall t \geq 0 \quad (14.8)$$

de la ecuación 14.5 y 14.8, obtenemos finalmente:

$$0 \leq x^T(t)x(t) \leq \frac{\lambda_{\max}(P)}{\lambda_{\min}(P)} x^T(0)x(0) \exp\left(-\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)} t\right)$$

es decir:

$$\|x(t)\|^2 \leq \frac{\lambda_{\max}(P)}{\lambda_{\min}(P)} \|x(0)\|^2 \exp\left(-\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)} t\right)$$

Missing
figure

Función arbitraria acotada por una exponencial negativa.

Dado que A es Hurwitz estable tenemos que $\Re \lambda(A) < 0$, sea la siguiente matriz definida positiva:

$$P = \int_0^\infty \exp(A^T t) Q \exp(At) dt$$

con $Q = Q^T > 0$. Entonces tendremos:

$$\begin{aligned}
 A^T P + P A &= \int_0^\infty \left(A^T \exp(A^T t) Q \exp(At) + \exp(A^T t) Q \exp(At) A \right) dt \\
 &= \int_0^\infty \frac{d}{dt} \left(\exp(A^T t) Q \exp(At) \right) dt \\
 &= \exp(A^T t) Q \exp(At) \Big|_0^\infty \\
 &= \lim_{t \rightarrow \infty} \exp(A^T t) Q \exp(At) - \exp(A^T \cdot 0) Q \exp(A \cdot 0) \\
 &= 0 - Q = -Q
 \end{aligned}$$

□

Lema 14.1. La ecuación matricial $AX = XB$, tiene únicamente la solución trivial, $X = 0$, si y solo si, A y B no tienen valores propios en común, es decir:

$$\sigma(A) \cap \sigma(B) \neq \emptyset$$

Demostración.

□

Corolario 14.1. En el teorema 14.1 se puede elegir a Q como una matriz semidefinida positiva, $Q \geq 0$, bajo la condición inicial de que $x^T(t)Qx(t)$ no sea idénticamente nula a lo largo de cualquier trayectoria no nula, solución de:

$$\frac{dx(t)}{dt} = Ax(t)$$

El requerimiento de este corolario se reduce a la condición de que el par $(Q_{1/2}, A)$ sea observable:

$$Q \geq 0 \exists Q_{1/2} \text{ tal que } Q = Q_{1/2}^T Q_{1/2}$$

Corolario 14.2. Si A es una matriz Hurwitz estable entonces la ecuación de Lyapunov, $A^T P + P A = -Q$, tiene una única solución para cada Q .

Demostración. Suponga que existen dos soluciones, P_1 y P_2 , de la ecuación de Lyapunov, es decir:

$$\begin{aligned}
 A^T P_1 + P_1 A &= -Q \\
 A^T P_2 + P_2 A &= -Q
 \end{aligned}$$

lo cual implica:

$$\begin{aligned}
 A^T (P_2 - P_1) + (P_2 - P_1) A &= 0 \\
 A^T (P_2 - P_1) - (P_2 - P_1)(-A) &= 0
 \end{aligned}$$

pero sabemos que el espectro de una matriz, no cambia debido a la transposición, $\sigma(A^T) = \sigma(A)$, y por otro lado tenemos que la matriz A es Hurwitz estable, es decir $\Re\{\lambda(A)\} < 0$, lo cual implica que:

$$\Re\{\lambda(-A)\} > 0$$

por lo que podemos concluir que:

$$\sigma(A^T) \cap \sigma(-A) = \emptyset$$

por lo tanto, solo hay una solución y es la trivial.

□

Capítulo 15

Introducción a la optimización de funcionales

El problema que tratamos de resolver es el siguiente; determinar $v(t)$, tal que la siguiente ecuación:

$$J(v(t), v'(t)) = \int_{t_1}^{t_2} f_1(v(t), v'(t), t) dt \tag{15.1}$$


sea un extremo.

Primero definamos una función de vecindad $v(t) \rightarrow v(\alpha, t)$ tal que si $\alpha = 0 \implies v^* = v(0, t) = v(t)$, es decir, v^* es la función que extremiza a $J(v(t), v'(t))$.

Proponemos una solución en forma lineal:

$$v(\alpha, t) = v(0, t) + \alpha \eta(t) \tag{15.2}$$

pero $v(t)$ y $u(t)$ deben ser idénticos en los puntos extremos.



Missing figure

Trayectorias diferentes con inicio y fin comunes.

es decir $\eta(t_1) = \eta(t_2) = 0$, pero además $\eta(t) \in e^1$.¹

Con esta parametrización en α :

$$v(t) \rightarrow v(\alpha, t) = v(0, t) + \alpha \eta(t)$$

tenemos que la ecuación 15.1 nos queda:

¹Si $f \in e^1$, f es una función diferenciable al menos una vez.

$$J(\alpha) = \int_{t_1}^{t_2} f_1(v(\alpha, t), v'(\alpha, t), t) dt$$

donde tenemos que $\alpha = 0$ implica que J es un extremo y $\alpha \neq 0$ implica que J no es un extremo.

Debido a esto, podemos concluir que J tambien esta parametrizada de esta manera:

$$J \rightarrow J(\alpha)$$

La condición necesaria para que J tenga un valor estacionario (extremo), es que J sea independiente de α en primer orden (que este relacionado linealmente), a lo largo de la trayectoria que otorga el extremo ($\alpha = 0$), es decir:

$$\left. \frac{\partial J}{\partial \alpha} \right|_{\alpha=0} = 0 \quad \forall \eta \in \mathbf{e}^1 \quad (15.3)$$

Nota 15.1. *Observe que solo es una condición necesaria, es decir:*

$$J \text{ es extremo} \implies \left. \frac{\partial J}{\partial \alpha} \right|_{\alpha=0} = 0$$

Ecuación de Euler

La condición necesaria es:

$$\left. \frac{\partial J}{\partial \alpha} \right|_{\alpha=0} = 0$$

entonces hay que seguir los siguientes pasos:

1. Calcular $\frac{\partial J}{\partial \alpha}$.
2. Hacer $\alpha = 0$.

Empecemos calculando la derivada parcial de J:

$$\begin{aligned} \frac{\partial J}{\partial \alpha} &= \frac{\partial}{\partial \alpha} \int_{t_1}^{t_2} f_1(v(\alpha, t), v'(\alpha, t), t) dt = \\ &= \int_{t_1}^{t_2} \left(\frac{\partial f_1(\dots)}{\partial v(\alpha, t)} \frac{\partial v(\alpha, t)}{\partial \alpha} + \frac{\partial f_1(\dots)}{\partial v'(\alpha, t)} \frac{\partial v'(\alpha, t)}{\partial \alpha} \right) dt \end{aligned}$$

en este punto aparecen términos reducibles:

$$\begin{aligned} \frac{\partial v(\alpha, t)}{\partial \alpha} &= \frac{\partial v(t)}{\partial \alpha} + \frac{\partial (\alpha \eta(t))}{\partial \alpha} = \eta(t) \\ \frac{dv'(\alpha, t)}{d\alpha} &= \frac{\partial}{\partial \alpha} \left(\frac{dv(\alpha, t)}{dt} \right) = \frac{\partial}{\partial \alpha} \left(v'(t) + \alpha \frac{d\eta(t)}{dt} \right) = \frac{d\eta(t)}{dt} \end{aligned}$$

lo que nos deja:

$$\frac{\partial J}{\partial \alpha} = \int_{t_1}^{t_2} \left(\frac{\partial f_1(\dots)}{\partial v(\alpha, t)} \eta(t) + \frac{\partial f_1(\dots)}{\partial v'(\alpha, t)} \frac{d\eta(t)}{dt} \right) dt$$

la segunda parte de esta integral es integrable por partes, si hacemos $u = \frac{\partial f_1(\dots)}{\partial v'(\alpha, t)}$, $dv = \frac{d\eta(t)}{dt} dt$, $du = \frac{d}{dt} \left(\frac{\partial f_1(\dots)}{\partial v'(\alpha, t)} \right)$ y $v = \eta(t)$:

$$\frac{\partial J}{\partial \alpha} = \int_{t_1}^{t_2} \left(\frac{\partial f_1(\dots)}{\partial v(\alpha, t)} \eta(t) - \frac{d}{dt} \left(\frac{\partial f_1(\dots)}{\partial v'(\alpha, t)} \right) \eta(t) \right) dt + \left. \frac{\partial f_1(\dots)}{\partial v'(\alpha, t)} \eta(t) \right|_{t_1}^{t_2}$$

pero recordemos que $\eta(t_1) = \eta(t_2) = 0$, por lo que el ultimo termino se elimina y nos queda:

$$\begin{aligned} \frac{\partial J}{\partial \alpha} &= \int_{t_1}^{t_2} \left(\frac{\partial f_1(\dots)}{\partial v(\alpha, t)} \eta(t) - \frac{d}{dt} \left(\frac{\partial f_1(\dots)}{\partial v'(\alpha, t)} \right) \eta(t) \right) dt = \\ &= \int_{t_1}^{t_2} \left(\frac{\partial f_1(v(\alpha, t), v'(\alpha, t), t)}{\partial v(\alpha, t)} - \frac{d}{dt} \left(\frac{\partial f_1(v(\alpha, t), v'(\alpha, t), t)}{\partial v'(\alpha, t)} \right) \right) \eta(t) dt \quad (15.4) \end{aligned}$$

Si ahora, en la ecuación 15.4 sustituimos $\alpha = 0$, obtendremos:

$$\left. \frac{\partial J}{\partial \alpha} \right|_{\alpha=0} = \int_{t_1}^{t_2} \left(\frac{\partial f_1(v(t), v'(t), t)}{\partial v(t)} - \frac{d}{dt} \left(\frac{\partial f_1(v(t), v'(t), t)}{\partial v'(t)} \right) \right) \eta(t) dt$$

Por lo que la condición necesaria es:

$$\int_{t_1}^{t_2} \left(\frac{\partial f_1(v(t), v'(t), t)}{\partial v(t)} - \frac{d}{dt} \left(\frac{\partial f_1(v(t), v'(t), t)}{\partial v'(t)} \right) \right) \eta(t) dt = 0 \quad \forall \eta(t) \in e^1$$

lo cual implica que:

$$\frac{\partial f_1(v(t), v'(t), t)}{\partial v(t)} - \frac{d}{dt} \left(\frac{\partial f_1(v(t), v'(t), t)}{\partial v'(t)} \right) = 0 \quad (15.5)$$

esta es la que conocemos como ecuación de Euler.

Multiplicadores de Lagrange

Deseamos resolver el siguiente problema:

Minimizar la función $f(v) : \mathbb{R}^n \rightarrow \mathbb{R}$ sujeta a la restricción $\mathcal{G}(v) = 0$, donde $\mathcal{G}(v) = (\mathcal{G}_1(v) \ \mathcal{G}_2(v) \ \dots \ \mathcal{G}_m(v)) \in \mathbb{R}^m$ y $\mathcal{G}_i(v) : \mathbb{R}^n \rightarrow \mathbb{R}$ con $i \in \{1, 2, \dots, m\}$.


Para resolver este problema haremos uso de los multiplicadores de Lagrange, los cuales estan basados en el concepto de la derivada direccional²

Derivada direccional y vector gradiente

Definición 15.1. La derivada direccional de f_1 en $v_0 \in \mathbb{R}^n$ en la dirección del vector unitario $\eta \in \mathbb{R}^n$ es:


$$D_{\eta} f_1(v_0) = \lim_{\alpha \rightarrow 0} \frac{f_1(v_0 + \alpha) - f_1(v_0)}{\alpha} \tag{15.6}$$

donde α es un escalar, es decir, $\alpha \in \mathbb{R}$, en el caso de que este limite exista.



Missing figure

Superficie con vector posición evaluado y vector dirección de la derivada unitario.



Missing figure

Curva cortada por una recta por la que pasa un vector restando los dos puntos evaluados de la curva.

Esta derivada nos da la razon de cambio de f_1 en el punto v_0 y en la dirección η , siempre y cuando $\alpha \rightarrow 0$.

Teorema 15.1. Si f_1 es una función diferenciable de $v \in \mathbb{R}^n$, entonces f_1 tiene una derivada direccional en la dirección de cualquier vector unitario $\eta \in \mathbb{R}^n$ y por lo tanto:

$$D_{\eta} f_1(v) = \sum_{i=1}^n \frac{\partial f_1(v_i)}{\partial v_i} \eta_i \tag{15.7}$$

²Esta derivada es formalmente conocida como la derivada de Fréchet, la cual se relaciona linealmente con la diferencial de Fréchet.

$$\text{donde } \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n, \boldsymbol{\eta} = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} \in \mathbb{R}^n \text{ con } \|\boldsymbol{\eta}\| = 1.$$

Definición 15.2. Dada una base ortonormal de \mathbb{R}^n , $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$, entonces el gradiente de t_1 es la función vectorial, ∇f_1 , definida por:

$$\nabla_{\mathbf{v}} f_1(\mathbf{v}) = \sum_{i=1}^n \frac{\partial f_1}{\partial v_i} \mathbf{e}_i \quad (15.8)$$

$$\text{donde } v_i = \sum_{i=1}^n v_i \mathbf{e}_i.$$

Con esta notación gradiente, la ecuación 15.19 de la derivada direccional se escribe:

$$D_{\boldsymbol{\eta}} f_1(\mathbf{v}) = (\nabla f_1(\mathbf{v}), \boldsymbol{\eta}) \quad (15.9)$$

donde $(\mathbf{R}, \mathbf{S}) : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ es el producto punto, es decir, la proyección escalar del vector \mathbf{R} sobre la dirección del vector \mathbf{S} .

Teorema 15.2. Suponga que $f_1 : \mathbb{R}^n \rightarrow \mathbb{R}$ es una funcional diferenciable. El máximo valor de la derivada $D_{\boldsymbol{\eta}} f(\mathbf{v})$ es $\|D_{\boldsymbol{\eta}} f(\mathbf{v})\|$ y se obtiene cuando la dirección de $\boldsymbol{\eta}$ coincide con el vector gradiente $\nabla_{\mathbf{v}} f_1(\mathbf{v})$.

Sea la superficie, \mathcal{S} , definida por la funcional $\mathcal{G} : \mathbb{R}^n \rightarrow \mathbb{R}$ de la siguiente manera:

$$\mathcal{S} = \{\mathbf{v} \in \mathbb{R}^n \mid \mathcal{G}(\mathbf{v}) = k\} \quad (15.10)$$

Sea \mathcal{C} una curva contenida en \mathcal{S} y que pase por el punto $\mathbf{v}_0 \in \mathbb{R}$, la cual está definida por la función vectorial $\mathbf{R} : \mathbb{R} \rightarrow \mathbb{R}^n$, esto es:

$$\mathcal{C} = \{\mathbf{v} \in \mathbb{R}^n, t \in \mathbb{R} \mid \mathbf{v} = \mathbf{R}(t)\} \quad (15.11)$$

Sea $t_0 \in \mathbb{R}$ tal que $\mathbf{R}(t_0) = \mathbf{v}_0$.

Como $\mathcal{C} \subset \mathcal{S}$, entonces cualquier punto, $\mathbf{v} \in \mathcal{C}$, satisface:

$$\mathcal{G}(\mathbf{v}) = k$$

lo cual implica (asumiendo que \mathbf{v} es una función diferenciable y que también \mathcal{G} lo es):

$$\sum_{i=1}^n \frac{\partial \mathcal{G}}{\partial v_i} \frac{dv_i}{dt} = 0$$

es decir:

$$(\nabla_{\mathbf{i}} \mathcal{G}, \mathbf{R}'(t)) = 0$$

$$\text{donde } \mathbf{R}'(t) = \frac{d\mathbf{R}(t)}{dt} = \begin{pmatrix} \frac{dv_1(t)}{dt} \\ \frac{dv_2(t)}{dt} \\ \vdots \\ \frac{dv_n(t)}{dt} \end{pmatrix} \in \mathbb{R}^n.$$

En particular, cuando $t = t_0$, se tiene:

$$\mathbf{R}(\mathbf{t}_0) = \mathbf{v}_0$$

$$\left(\nabla_{\mathbf{v}} \mathcal{G}(\mathbf{v}_0), \frac{d\mathbf{R}(\mathbf{t}_0)}{dt} \right) = 0$$

Esta ecuación indica que el vector gradiente en \mathbf{v}_0 , $\nabla_{\mathbf{v}} \mathcal{G}(\mathbf{v}_0)$, es perpendicular al vector tangente $\frac{d\mathbf{R}(\mathbf{t}_0)}{dt}$ a cualquier $\mathcal{C} \subset \mathcal{S}$ que pase por \mathbf{v}_0 .

Missing
figure

Superficie con vector tangente.

Missing
figure

Plano tangente a superficie de nivel y vector normal al plano.

Si $\nabla_{\mathbf{v}} \mathcal{G}(\mathbf{v}_0) \neq 0$, entonces se define el plano tangente a la superficie de nivel \mathcal{S} , en el punto $\mathcal{G}(\mathbf{v}_0)$ y tiene un vector normal $\nabla_{\mathbf{v}} \mathcal{G}(\mathbf{v}_0)$, esto es, el plano tangente τ que esta definido por:

$$\tau = \{\mathbf{v} \in \mathbb{R}^n \mid (\nabla_{\mathbf{v}} \mathcal{G}(\mathbf{v}_0), \mathbf{v} - \mathbf{v}_0) = 0\} \quad (15.12)$$

Multiplicadores de Lagrange

Procederemos ahora a resolver el problema original. Suponga que la funcional f_1 tiene un extremo en el punto v_0 en la superficie:

$$\mathcal{S}_m \{v \in \mathbb{R}^n, i \in \{1, 2, \dots, m\} \mid \mathcal{G}(v) = 0\} \tag{15.13}$$

donde $\mathcal{G}(v) = \begin{pmatrix} \mathcal{G}_1(v) \\ \mathcal{G}_2(v) \\ \vdots \\ \mathcal{G}_m(v) \end{pmatrix} \in \mathbb{R}^m$ y $\mathcal{G}_i(v) : \mathbb{R}^n \rightarrow \mathbb{R}$ con $i \in \{1, 2, \dots, m\}$.

Sea una curva:

$$\mathcal{C} = \{v \in \mathbb{R}^n \mid t \in \mathbb{R} \mid v = R(t)\} \subset \mathcal{S}_m \tag{15.14}$$

tal que $v_0 \in \mathcal{C}$.

Sea $t_0 \in \mathbb{R}$ el paramtero correspondiente a v_0 , es decir $v_0 = R(t_0)$.

La funcional compuesta:

$$\mathcal{H} = f_1(R(t)) \tag{15.15}$$

representa a los valores de f que también está en \mathcal{C} .

Como f_1 tiene un extremo en v_0 , entonces \mathcal{H} tiene un extremo en t_0 , por lo que:

$$\left. \frac{d\mathcal{H}(t)}{dt} \right|_{t=t_0} = 0 \tag{15.16}$$

Pero si f es diferenciable, se deduce por la regla de la cadena:

$$0 = \left. \frac{d}{dt} \mathcal{H} \right|_{t=t_0} = \sum_{i=1}^n \left. \frac{\partial f_1}{\partial v_i} \frac{v_i(t)}{dt} \right|_{t=t_0, v=v_0} = \left(\nabla_v f_1(v_0), \frac{d}{dt} R(t_0) \right) \tag{15.17}$$

esto nos indica que el vector gradiente $\nabla_v f_1(v_0)$, es ortogonal al vector tangente, $\frac{dR(t_0)}{dt}$, para cada una de estas curvas.

Pero sabemos que los vectores gradientes de las coordenadas, $(\mathcal{G}_i, \mathcal{G}_i(v_0))$, son tambien ortogonales a $\frac{dR(t_0)}{dt}$, por lo que los vectores gradiente $\nabla_v f_1(v_0)$ y $\nabla_v \mathcal{G}_i(v_0)$ con $i \in \{1, 2, \dots, m\}$ necesariamente son paralelos.

Entonces, si $\nabla_v \mathcal{G}_i(v_0)$ con $i \in \{1, 2, \dots, m\}$, existen $\lambda_1, \lambda_2, \dots, \lambda_m$, tales que:

$$\nabla_v f_1(v_0) = \sum_{i=1}^m \lambda_i \nabla_v \mathcal{G}_i(v_0) \tag{15.18}$$

y al vector $\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix} \in \mathbb{R}^m$ se le conoce como multiplicadores de Lagrange.

Entonces para resolver el problema original se procede como sigue:

1. Se construye la siguiente funcional aumentada:

$$f_a(v, \lambda) = f_1(v) - (\lambda, \mathcal{G}(v)) \quad (15.19)$$

2. Se encuentran los puntos estacionarios de la ecuación 15.19:

$$\begin{aligned} \nabla_v f_a(v, \lambda) &= 0 \\ \nabla_\lambda f_a(v, \lambda) &= 0 \end{aligned} \quad (15.20)$$

El par (v_0, λ_0) que satisface la ecuación 15.20 es la solución del problema original.

Observe que con este método, se está transformando un problema de optimización con restricciones, en uno sin restricciones.