UNIVERSIDAD DE EL SALVADOR

FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA

ESCUELA DE MATEMÁTICA

LICENCIATURA EN MATEMÁTICA

TRABAJO DE GRADUACIÓN TITULADO:

ESTABILIDAD E INTRODUCCIÓN A LA TEORÍA DE CONTROL

PREVIO A LA OBTENCIÓN DEL TÍTULO DE

LICENCIADO/A EN MATEMÁTICA

PRESENTADO POR:

Br. Maricela Elizabeth Martínez CARNÉ: MM06108

Br. Henrry José Soriano León CARNÉ: SL07015

Asesor:

Dr. Simón Alfredo Peña Aguilar

Ciudad Universitaria, 22 de enero de 2018

UNIVERSIDAD DE EL SALVADOR

RECTOR:

MSc. Roger Armando Arias Alvarado.

SECRETARIO GENERAL:

MSc. Cristobal Hernán Ríos Benítez

FISCAL GENERAL:

Lic. Rafael Humberto Peña Marín

FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA

DECANO:

Lic. Mauricio Hernán Lovo Córdoba.

VICE-DECANO:

Lic. Carlos Antonio Quintanilla Aparicio.

SECRETARIA:

Licda. Damaris Melany Herrera Turcios.

ESCUELA DE MATEMÁTICA

DIRECTOR DE ESCUELA:

Dr. José Nerys Funes Torres.

ASESOR:

Dr. Simón Alfredo Peña Aguilar

Dedicatoria

A nuestros padres Cristina, Felicita e Ignacio por su incondicional apoyo, amor y comprensión.

Agradecimientos

A Dios por la vida, sabiduría, salud y fuerzas para seguir adelante.

A nuestros padres por ser nuestro ejemplo de esfuerzo y dedicación y por animarnos a conseguir nuestras metas.

Al asesor, Dr. Simón Peña por su disponibilidad de tiempo, dedicación, esfuerzo y apoyo en cada aspecto de este trabajo.

A mis hermanos Fredy y Enmanuel por su alegría, ayuda y palabras de aliento.

Índice general

Índice de figuras

Resumen

El presente trabajo consiste de un tema que tiene mucha importancia en la mayoría de ramas de la Física, Ingeniería y Economía como es la estabilidad. Primero se presenta a modo de resumen algunos definiciones sobre autovalores y autovectores, así como la forma de Jordan que será de gran utilidad para decir cuando un sistema de ecuaciones diferenciales ordinarias o una matriz es estable, inestable o asintóticamente estable.

También se estudian las ecuaciones lineales y no lineales no autónomas, con coeficientes constantes para poder analizar que tipo de estabilidad presentan.

Pero principalmente se estudiarán las funciones de Lyapunov, las cuales nos ayudan a determinar si un sistema de ecuaciones diferenciales es estable, inestable o asintóticamente estable sin saber cual es su solución.

Introducción

Desde los orígenes de lo que conocemos como ciencia, el hombre ha tratado de entender y explicar su entorno, pero se ha encontrado con un mundo cambiante, donde todo está en movimiento, y se ha propuesto comprender el cómo y el por qué de esos movimientos. Una forma de explicar este comportamiento es a través de las Ecuaciones Diferenciales Ordinarias (EDO's), pero algunos fenómenos requieren de un sistema de éstas. La teoría de las ecuaciones diferenciales ordinarias intenta entender procesos en movimiento, es decir cambios o variaciones de un objeto con respecto al tiempo. Se entiende por sistema de ecuaciones diferenciales ordinarias a $x'(t) = f(t, x)$.

En este trabajo, estudiaremos el concepto de Estabilidad de Lyapunov el cual juega un papel fundamental en la Teoría de Sistemas y Control. La estabilidad es una propiedad cualitativa de las ecuaciones diferenciales a la que se le considera la más importante de todas.

La Teoría de Estabilidad es una materia antigua que aparece casi al mismo tiempo que la Teoría de Ecuaciones Diferenciales. El objetivo de la Teoría de Estabilidad en el sentido de Lyapunov es extraer conclusiones acerca del comportamiento de un sistema sin determinar sus trayectorias solución. Quizás el primero en estudiar estabilidad en el sentido "moderno"fue Lagrange (1788), quien analizó sistemas mecánicos usando lo que ahora llamamos Mecánica Lagrangiana. Una de sus conclusiones fue que en ausencia de fuerzas externas el estado de equilibrio de un sistema mecánico conservativo es estable siempre que este corresponda a un mínimo de la energía potencial del sistema. Varios investigadores siguieron los métodos de Lagrange, pero la mayor parte de su trabajo estuvo restringida a sistemas mecánicos conservativos descritos por ecuaciones de movimiento Lagrangiano. El avance significativo en la Teoría de Estabilidad que nos permite analizar ecuaciones diferenciales arbitrarias se debe al Matemático Ruso A.M. Lyapunov (1892). Lyapunov no únicamente introdujo las definiciones básicas de estabilidad que en la actualidad son usadas, sino que demostró muchos de los teoremas fundamentales. El trabajo de Lyapunov fue desconocido en el Oeste hasta alrededor de 1960 y todo el avance en la Teoría de Estabilidad hasta ese tiempo se debe a Matemáticos Rusos. Ahora, los fundamentos de la teoría están bien establecidos y se le considera como una herramienta indispensable en el análisis y síntesis de sistemas no lineales.

Este trabajo se puede resumir de la siguiente manera.

En el primer capítulo, se estudian las ecuaciones lineales. Donde la herramienta principal en el estudio de las ecuaciones lineales, es el análisis matricial, que se estudia en 1.1. Luego se exponen, en 1.2, los elementos básicos de la teoría espectral de operadores en espacios de dimensión finita. Esto nos permite presentar una demostración del teorema de descomposición de Jordan.

Las últimas secciones 1.3, 1.4 y 1.5 aplican los resultados de que disponemos al estudio general de las ecuaciones lineales, con coeficientes constantes y ecuaciones periódicas, respectivamente.

En el segundo capítulo, se dan las definiciones básicas de estabilidad, estabilidad asintótica y se estudia la estabilidad de ecuaciones lineales.

El estudio de la estabilidad en los sistemas lineales es sencillo, esto, al analizar las características de los autovalores y autovectores de la matriz asociada al sistema, pero en sistemas no lineales no se puede aplicar esta técnica directamente, para ello se procede a realizar una linealización y a partir del sistema linealizado se puede hacer el estudio de la estabilidad.

En el tercer capitulo, se exponen en primer lugar, en las secciones 3.1, 3.2 y 3.3, algunos teoremas sobre estabilidad, estabilidad asintótica, estabilidad uniforme e inestabilidad, que constituyen los resultados básicos de la teoría.

Presentamos a continuación en detalle la construcción de la función de Lyapunov para una ecuación lineal con coeficientes constantes en la sección 3.4.

En el cuarto capítulo, se expone en un brebe resumen la estabilidad en Ingeniería, se presentan las definiciones de la transformada de Lapalace y la función de transferencia que son las definiciones que se ocupan para estudiar la estabilidad y el criterio de Routh-Hurwitz, en particular la tabulación de Routh.

Metodología

A continuación se describen los aspectos importantes de la metodología de trabajo:

1. **Tipo de investigación:** Este proyecto de graduación es de carácter bibliográfico-descriptivo.

   Bibliográfico: Se ha hecho una extensa recopilación de libros impresos y de libros obtenidos por Internet para contar con el suficiente material que cubra las necesidades del estudio y de las que puedan surgir más adelante. El objetivo es compilar coherentemente la información más útil y destacada del tema.

   Descriptivo: Ya que se pretende estudiar con detalle las demostraciones de cada uno de los teoremas propuestos.

2. **Forma de Trabajo:** Se realizaron reuniones periódicas con el asesor del trabajo para tratar los diferentes aspectos de la investigación como estudiar y discutir la teoría, tratar los diferentes aspectos del trabajo escrito y de las presentaciones.

3. **Exposiciones:** Se tendrán dos exposiciones

   Primera Exposición (Pública): Presentación del Perfil del Proyecto de Investigación.

   Segunda Exposición (Pública): Presentación Final del Trabajo de Investigación: resumen de resultados y aplicaciones.

Capítulo 1

Ecuaciones Lineales

La teoría de ecuaciones diferenciales lineales es básica en multitud de aplicaciones a la Física, la Economía y la Ingeniería, etc. Muchos problemas reales vienen a plantear ecuaciones que son lineales o que admiten una aproximación lineal que es suficiente para muchos propósitos.
La herramienta principal en el estudio de las ecuaciones lineales, es el análisis matricial, cuyos resultados más útiles para nuestro propósito se exponen en 1.1, a modo de resumen. A continuación se exponen, en 1.2, los elementos básicos de la teoría espectral de operadores en espacios de dimensión finita. Esto nos permite presentar una demostración bastante sencilla y directa del teorema de descomposición de Jordan y al mismo tiempo nos da ocasión para desarrollar algunas de las ideas básicas subyacentes a la teoría espectral general de operadores, de tanta importancia en el análisis y en la Física actual.
Las secciones 1.3, 1.4 y 1.5 aplican los resultados de que disponemos al estudio general de las ecuaciones lineales, con coeficientes constantes y ecuaciones periódicas, respectivamente.

1.1. Elementos de análisis matricial

1.1.1. Isomorfismo entre operadores lineales y matriciales

Se puede establecer un isomorfismo entre el espacio vectorial de las matrices $n \times n$ de elementos complejos y el espacio de las aplicaciones (u operadores) lineales definidas en un espacio vectorial X complejo de dimensión n, hacia el mismo X. Esto nos permitirá hablar indiferentemente de operadores lineales o de matrices.
Lo primero que vamos a probar es que toda aplicación matricial es una aplicación lineal.
Sea A una matriz $n \times n$ de elementos de $\mathbf{C}$, $A = (a_{ij})$. Sea X un espacio vectorial sobre $\mathbf{C}$ de dimensión n, y sea $\mathfrak{B} = \{e_1, e_2, \ldots, e_n\}$ una base fija de X. Definimos la aplicación $\theta(A) : X \to X$ del siguiente modo:
Para e_j ponemos

$$\theta(A)e_j = \sum_{i=1}^{n} a_{ij}e_i$$

observemos que $\theta(A)e_j$ lo que nos esta dando son las columnas de la matriz A y, si $x \in X$ se puede expresar de manera única como una combinación lineal de los elementos de la base

$$x = \sum_{j=1}^{n} x_j e_j$$

entonces aplicando $\theta(A)$ a ambos lados de esta última igualdad se tiene:

$$\begin{aligned} \theta(A)x &= \sum_{j=1}^{n} x_j \theta(A) e_j \quad \textit{sustituyendo} \quad \theta(A)e_j \\ &= \sum_{i,j=1}^{n} a_{ij} x_j e_i = \sum_{i=1}^{n} x_i^* e_i \\ \theta(A)x &= x^* \end{aligned}$$

donde hemos llamado

$$x_i^* = \sum_{j=1}^{n} a_{ij} x_j$$

Ahora veamos que la aplicación así definida es lineal, para ello recordemos la definición de aplicación lineal.

Definición 1.1.1. *Sean X e Y dos espacios vectoriales sobre el mismo cuerpo $\mathbb{K}$, una aplicación $T : X \longmapsto Y$ se dice* ***lineal u homomorfismo*** *si se verifica:*

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y) \quad \textit{para} \quad \textit{cualquier} \quad \alpha, \beta \in \mathbb{K} \quad x, y \in X$$

si T es biyectivo, se dice que T es ***isomorfismo.***

Veamos que $\theta(A)$ es un operador lineal es, decir si se cumple la propiedad de la definición anterior

$$\theta(A)\left[\lambda x + \mu x'\right] = \lambda \theta(A) x + \mu \theta(A) x', \quad \lambda, \mu \in \mathbf{C}$$

Demostración:
Sea

$$x = \sum_{j=1}^{n} x_j e_j \qquad x' = \sum_{j=1}^{n} x'_j e_j$$

$$\begin{aligned} \theta(A)(\alpha x + \beta x') &= \theta(A)\left(\alpha \sum_{j=1}^{n} x_j e_j + \beta \sum_{j=1}^{n} x'_j e_j\right) = \theta(A)\left[\sum_{j=1}^{n}\left(\alpha x_j + \beta x'_j\right) e_j\right] \\ &= \sum_{j=1}^{n}\left(\alpha x_j + \beta x'_j\right)\theta(A)e_j = \sum_{j=1}^{n}\left(\alpha x_j + \beta x'_j\right)\sum_{i=1}^{n} a_{ij} e_i \\ &= \sum_{i,j=1}^{n}\left(\alpha x_j + \beta x'_j\right) a_{ij} e_i = \alpha \sum_{i,j=1}^{n} x_j a_{ij} e_i + \beta \sum_{i,j=1}^{n} x'_j a_{ij} e_i \end{aligned}$$

$$\begin{aligned}
&= \alpha \sum_{j=1}^{n} \left(\sum_{i=1}^{n} a_{ij} e_i \right) x_j + \beta \sum_{j=1}^{n} \left(\sum_{i=1}^{n} a_{ij} e_i \right) x'_j \\
&= \alpha \sum_{j=1}^{n} x_j \theta(A) e_j + \beta \sum_{j=1}^{n} x'_j \theta(A) e_j = \alpha\theta(A) \sum_{j=1}^{n} x_j e_j + \beta\theta(A) \sum_{j=1}^{n} x'_j e_j \\
&= \alpha\theta(A)x + \beta\theta(A)x'
\end{aligned}$$

así la aplicación es lineal y si escribimos las coordenadas de x y x^* como los vectores columnas

$$C(x) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad C(x^*) = \begin{pmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_n^* \end{pmatrix}$$

se tiene $C(x^*) = AC(x)$.

A continuación probaremos que toda aplicación lineal es una aplicación matricial.

Recíprocamente si se tiene el operador lineal $T : X \to X$, podemos definir una matriz $\psi(T)$ de elementos complejos de dimensión finita $n \times n$ del siguiente modo: Si

$$Te_j = \sum_{i=1}^{n} a_{ij} e_i$$

entonces la columna j-ésima de $\psi(T)$ será, precisamente:

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$$

Denominemos $M(n, \mathbf{C})$ al espacio vectorial sobre $\mathbf{C}$ de todas las matrices $n \times n$ de elementos complejos y $L(X, X)$ al espacio sobre $\mathbf{C}$ de todos los operadores lineales de X a X, siendo X el espacio vectorial sobre $\mathbf{C}$ de dimensión n en el que ha fijado la base $\mathfrak{B}$. Entonces, $\theta : M(n, \mathbf{C}) \to L(X, X)$ es un isomorfismo suprayectivo cuyo inverso es precisamente ψ.

Decimos que θ es un isomorfismo ya que por propiedad de isomorfismo tenemos que entre dos espacios de igual dimensión siempre existe un isomorfismo.

Si $X = \mathbf{C}^n$ y tomamos en X la base

$$\left\{ e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\}$$

entonces, dada una matriz $A = (a_{ij}) \in M(n, \mathbf{C})$ y un vector:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbf{C}^n$$

se tiene $\theta(A)x = A(x)$ y, en particular,

$$\theta(A)e_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$$

Es decir, $\theta(A)$ es precisamente el operador que transforma los elementos de la base $\mathfrak{B}$ en los correspondientes vectores columna de A.
Estas consideraciones permiten identificar operadores lineales y matrices de una forma natural. En lo que sigue hablaremos indistintamente del operador $T \in L(X, X)$ y de su matriz asociada $\psi(T)$, que también denominaremos a veces T, suponiendo que se ha elegido una base $\mathfrak{B}$ en X.

1.1.2. Equivalencias de normas en un espacio de dimensión finita

Las normas son herramientas básicas para definir y analizar la convergencia de una sucesión de vectores en espacios vectoriales. Como es habitual, una sucesión $\{x^k\} \subset X$ se dice que converge a x (y se escribe $x^k \mapsto x$) si la sucesión de números reales $\{\|x^k - x\|\}$ converge a cero, $\|x^k - x\| \mapsto 0$, siendo $\| \;\|$ una norma definida en X. Esta definición depende de la norma $\| \;\|$. En principio podría suceder que una sucesión de vectores convergiera para una norma pero no para otra. De hecho, esto es perfectamente posible en espacios de dimensión infinita, pero no en espacios de dimensión finita.
A continuación presentamos algunas definiciones de las cuales nos vamos a auxiliar para estudiar lo que es equivalencias de normas.

Definición 1.1.2. *Una **norma** en un espacio vectorial X sobre $\mathbb{K}$ es una función $\| \;\| : X \mapsto [0, \infty)$ tal que para cualesquiera $x, y \in X$ y $\lambda \in \mathbb{K}$:*

i) $\|x\| \geqslant 0$ *y* $\|x\| = 0$ *si y sólo si* $x = 0$

ii) $\|\lambda x\| = |\lambda| \|x\|$

iii) $\|x + y\| \leqslant \|x\| + \|y\|$

*Al par $(X, \| \;\|)$ se le llama **espacio normado.***

Definición 1.1.3. *Se dice que el operador lineal $T : X \to Y$ es isométrico o que es una isometría cuando $\|Tx\| = \|x\|$ $x \in X$. Si T es isométrico de la relación $\|Tx - Ty\| = \|x - y\|$ $(x, y \in X)$ se*

deduce que T es inyectivo. También es claro que toda isometría T es continua con $\|T\| = 1$. Si T es una isometría y además es biyectiva se dice que T es un isomorfismo isométrico.
Isometría: se conserva la distancia entre los puntos.

Sea X un espacio vectorial complejo de dimensión finita n. Sea

$$\mathfrak{B} = \{e_1, e_2, \dots, e_n\}$$

una base de X. Definimos una aplicación lineal $\alpha : X \to \mathbf{C}^n$ poniendo, $x = \sum_{i=1}^{n} x_i e_i \in X$

$$\alpha(x) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

La aplicación es trivialmente un isomorfismo suprayectivo que permite identificar X con $\mathbf{C}^n$, una vez fijada la base $\mathfrak{B}$.
Es claro que si en X se tiene una norma $\| \ \|$ y se define, para $a \in \mathbf{C}^n$:

$$p(a) = \|\alpha^{-1}(a)\| = \|\sum_{i=1}^{n} x_i e_i\|$$

entonces p es una norma en $\mathbf{C}^n$ y, así, α se convertirá en un isomorfismo suprayectivo de $(X, \| \ \|)$ a $(\mathbf{C}^n, p)$.

Definición 1.1.4. *Dos normas $\| \ \|$ y $\| \ \|^*$ sobre el espacio vectorial X son equivalentes si existen dos constantes $a, b > 0$ tales que*

$$a\|x\| \leqslant \|x\|^* \leqslant b\|x\|, \quad \forall x \in X$$

Demostraremos que en $\mathbf{C}^n$ una norma cualquiera p es equivalente a la norma:

$$|x|_1 = \left| \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right|_1 = \sum_{i=1}^{n} |x_i|$$

es decir debemos de encontrar constantes a y b tales que

$$ap(x) \leqslant \|x\|_1 \leqslant bp(x), \quad \forall x \in X$$

Para ello observemos en primer lugar que, si

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

y $N = \sup\{p(e_i) : i = 1, 2, \ldots, n\}$, se tiene:

$$\begin{aligned} p(x) &= p\left(\sum_1^n x_i e_i\right) \text{ \textit{por ser p lineal ya que α es lineal}} \\ &\leqslant \sum_1^n p(x_i e_i) \text{ \textit{por propiedad i) y ii) de la definición} 1.1.2} \\ &\leqslant \sum_1^n |x_i| p(e_i) \text{ \textit{por como se a definido} } N \\ &\leqslant N \sum_1^n |x_i| = N|x|_1 \end{aligned}$$

entonces

$$p(x) \leqslant N|x|_1 \tag{1.1.1}$$

Supongamos ahora que se tiene una sucesión $\{x^k\} \subset \mathbf{C}^n$ tal que $x^k \to 0$ en la norma $|\ |_1$. Si es que $p(x^k)$ no converge a 0, entonces existe $\eta > 0$ y una subsucesión de $\{x^k\}$ que denominaremos de nuevo $\{x^k\}$ tal que $x^k \to 0$, $x^k \neq 0$, y $p(x^k) \geqslant \eta$. Consideremos los puntos:

$$z^k = \frac{x^k}{|x^k|_1}$$

Se tiene, claramente, $|z^k|_1 = 1$ ya que así como esta definido es el vector unitario y, según la ecuación (1.1.1), se tiene

$$p(z^k) \leqslant N|z^k|_1 = N$$

Pero por otra parte,

$$\begin{aligned} p(z^k) &= p\left(\frac{x^k}{|x^k|_1}\right) \text{ \textit{por propiedad ii) de la definición} 1.1.2} \\ &= \left|\frac{1}{|x^k|_1}\right| p(x^k) = \frac{p(x^k)}{|x^k|_1} \\ &\geqslant \frac{\eta}{|x^k|_1} \end{aligned}$$

y, así, $p(z^k) \to \infty$ para $k \to \infty$. Esta contradicción demuestra que $p(x^k) \to 0$. Así, siendo p una norma en $\mathbf{C}^n$ resulta que es una aplicación continua de $\mathbf{C}^n$ a $[0, \infty)$ respecto de $|\ |_1$. Como el conjunto $\{x \in \mathbf{C}^n : |x|_1 = 1\}$ es compacto en $(\mathbf{C}^n, |\ |_1)$, existe z con $|z|_1 = 1$ tal que

$$p(z) = H = \min\{p(x) : |x|_1 = 1\}$$

Siendo p una norma en $\mathbf{C}^n$ se tiene $z \neq 0$, y así, $H > 0$. Por tanto, para todo $x \in \mathbf{C}^n$, $x \neq 0$, resulta:

$$\begin{aligned} p\left(\frac{x}{|x|}\right) &= \left|\frac{1}{|x|_1}\right| p(x) \quad \text{\textit{por propiedad} [\textit{ii})] \textit{de la definición} 1.1.2} \\ &= \frac{p(x)}{|x|_1} \geqslant H \end{aligned}$$

así

$$|x|_1 \leqslant \frac{1}{H}p(x) \tag{1.1.2}$$

de las ecuaciones (1.1.1) y (1.1.2) se tiene:

$$p(x)\frac{1}{N} \leqslant |x|_1 \leqslant \frac{1}{H}p(x)$$

Esto demuestra la equivalencia de las normas $|\ |_1$ y p en $\mathbf{C}^n$.

En particular, todo espacio normado de dimensión finita es un espacio de Banach. En general, en lo que sigue utilizaremos la norma euclídea de $\mathbf{C}^n$, es decir, si $x \in \mathbf{C}^n$

$$|x| = \left(\sum_{i=1}^{n}|x_i|^2\right)^{1/2}$$

Una ventaja importante de la norma euclídea consiste en que está asociada de una forma natural con el producto escalar definido del siguiente modo para $x, y \in \mathbf{C}^n$:

$$(x, y) = \sum_{i=1}^{n} x_i\bar{y}_i$$

Se tiene

$$|x| = (x, x)^{1/2}$$

La norma euclídea presenta también la ventaja de poseer una derivada continua respecto de las componentes en $\mathbf{C}^n - \{0\}$, lo cual no sucede con otras normas usuales como

$$|x|_1 = \sum_{i=1}^{n}|x_i| \quad o \quad bien \quad |x|_\infty = \max\{|x_i| : i = 1, 2, \ldots, n\}$$

Tiene, sin embargo, la desventaja de que la norma asociada a ella de los operadores $T \in L(X, X)$, norma que introducimos a continuación, es de una expresión de cálculo más complicada que en los otros casos.

1.1.3. Norma de un operador

Daremos primero la definición de norma para poder aplicar la definición al operador T.

Definición 1.1.5. *Se dice que una función $\|\ \| : M(n, \mathbf{C}) \to \mathbf{R}$ es una norma matricial si y sólo si para todo $A, B \in M(n, \mathbf{C})$ se cumplen las siguientes propiedades:*

1. $\|A\| > 0$ *si* $A \neq 0$ *y* $\|A\| = 0$ *si y sólo si* $A = 0$ *positiva*

2. $\|\alpha A\| = |\alpha|\|A\|$ *para todo* $\alpha \in \mathbb{R}$ *homogeneidad*

3. $\|A + B\| \leqslant \|A\| + \|B\|$ *desigualdad triangular*

estas propiedades expresan que las normas matriciales deben ser, en particular, normas vectoriales sobre $M(n, \mathbf{C})$ considerado como espacio vectorial.

Sea X un espacio vectorial complejo de dimensión n y sea $\mathfrak{B} = \{e_1, e_2, \ldots, e_n\}$ una base de X. Sea $\| \ \|$ una norma en X. Para $T \in L(X, X)$ definimos la norma matricial inducida por una norma vectorial:

$$\|T\| = \sup\{\|Tx\| : \|x\| \leqslant 1\}$$

Es claro que, por las consideraciones de la subsección 1.1.2, existe $M > 0$ tal que si $\|x\| \leqslant 1$, entonces

$$\sum_{i=1}^{n} |x_i| \leqslant M < \infty, \quad \textit{siendo} \quad x = \sum_{i=1}^{n} x_i e_i$$

Así, aplicando el operador T a x se tiene

$$\begin{aligned} \|Tx\| &= \left\| T\left(\sum_{i=1}^{n} x_i e_i\right)\right\| \textit{ por ser } T \textit{ lineal} \\ &= \left\|\sum_{i=1}^{n} x_i T e_i\right\| \textit{aplicando propiedades } [ii)] \textit{ y } [iii)] \textit{ de la definición } 1.1.2 \textit{ se tiene} \\ &\leqslant \sum_{i=1}^{n} |x_i| \|Te_i\| \leqslant M \sum_{i=1}^{n} \|Te_i\| < \infty \end{aligned}$$

para todo x con $\|x\| \leqslant 1$. Por tanto, la aplicación $\| \ \|$ definida así en $L(X, X)$ toma sus valores en $[0, \infty)$.

Nota 1. *La norma del operador (o matriz) T también la podemos definir como*

$$\|T\| = \sup\{\|Tx\| : \|x\| \leqslant 1\} = \sup_{\|x\| \leqslant 1} \{\|Tx\|\}$$

Ahora veamos que así como se ha definido T es una norma, y por la definición 1.1.5 y la subsección 1.1.1, se tiene para $A, B \in L(X, X)$.

1. $\|A\| > 0$ si $A \neq 0$.

2. $$\|\alpha A\| = \sup_{\|x\| \leqslant 1} \{\|\alpha A x\|\} = \sup_{\|x\| \leqslant 1} \{|\alpha| \|Ax\|\} = |\alpha| \sup_{\|x\| \leqslant 1} \{\|Ax\|\} = |\alpha| \|A\|$$

3. $$\begin{aligned} \|A + B\| &= \sup_{\|x\| \leqslant 1} \{\|(A + B)x\|\} = \sup_{\|x\| \leqslant 1} \{\|Ax + Bx\|\} \\ &\leqslant \sup_{\|x\| \leqslant 1} \{\|Ax\| + \|Bx\|\} \leqslant \sup_{\|x\| \leqslant 1} \{\|Ax\|\} + \sup_{\|x\| \leqslant 1} \{\|Bx\|\} \\ &= \|A\| + \|B\| \end{aligned}$$

Así, T es una norma y convierte a $L(X, X)$, que es espacio de dimensión $n \times n$, en un espacio de Banach. El espacio $(L(X, X), \| \ \|)$ admite, como se ha visto en la subsección 1.1.2, un isomorfismo isométrico con $(\mathbf{C}^{n \times n}, p)$, donde p es una norma en $\mathbf{C}^{n \times n}$ obtenida en $\mathbf{C}^{n \times n}$ por medio del isomorfismo suprayectivo α que existe entre $L(X, X)$ y $\mathbf{C}^{n \times n}$.
La norma $\| \ \|$ definida en $L(X, X)$ satisface, las siguientes propiedades.

Propiedad 1.1.1. *a)* $\|Ax\| \leqslant \|A\|\|x\|$

b) Consistencia $\|AB\| \leqslant \|A\|\|B\|$
para todo $x \in X$, $A \in L(X, X)$, $B \in L(X, X)$

Demostración. [a)]

Cuando $\|x\| \leqslant 1$, $\|A\| = \sup_{\|x\|\leqslant 1} \{\|Ax\|\} \geqslant \|Ax\| \Rightarrow \|A\|\|x\| \geqslant \|Ax\|$.

- Cuando $\|x\| = \alpha \neq 1$, sea $y = \dfrac{x}{\alpha}$ entonces $\|y\| \leqslant 1$ por lo tanto
$\|Ay\| \leqslant \|A\|\|y\| \Rightarrow \alpha\|Ay\| \leqslant \alpha\|A\|\|y\| \Rightarrow \|A\alpha y\| \leqslant \|A\|\|\alpha y\| \Rightarrow \|Ax\| \leqslant \|A\|\|x\|$

[b)] Para probar la propiedad de consistencia, ocupamos la propiedad a) así se tiene

$$\|ABx\| \quad \leqslant \quad \|A\|\|Bx\| \leqslant \|A\|\|B\|\|x\|$$

luego $\|AB\| \leqslant \|A\|\|B\|$. □

Es fácil interpretar en términos de las matrices $A \in M(n, \mathbf{C})$ asociadas a los operadores $A \in L(X, X)$, lo que significa $A_k \to 0$ en la norma asociada a una norma cualquiera $\|\ \|$ de X. Como una norma posible en $M(n, \mathbf{C})$ es

$$q(A) = \sum_{i,j=1}^{n} |a_{ij}|$$

y todas las normas son equivalentes, resulta que, para la sucesión $\{A_k\}$, $A_k = \left(a_{ij}^k\right)$, se tiene $A_k \to 0$ para $k \to \infty$ si, y sólo si, para cada par i, j, se tiene $a_{ij}^k \to 0$ para $k \to \infty$.
Es fácil ver que en $M(n, \mathbf{C})$ se tiene $T_k \to T$ si, y sólo si, para cada $x \in \mathbf{C}^n$ se tiene $T_k x \to Tx$. Para ello basta ver que $T_k \to 0$ es equivalente a que para cada $x \in \mathbf{C}^n$ se tenga $T_k x \to 0$. Supongamos que esta última condición se cumple. Sea $T_k = \left(t_{ij}^k\right)$ y tomemos

$$x = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Entonces

$$T_k x = \begin{pmatrix} t_{11}^k \\ t_{21}^k \\ \vdots \\ t_{n1}^k \end{pmatrix}$$

y así se tiene, para cada i,

$$t_{i1}^k \to 0 \quad \textit{para} \quad k \to \infty$$

Análogamente se obtiene para todo par i, j,

$$t_{ij}^k \to 0 \quad \textit{para} \quad k \to \infty$$

La implicación en el otro sentido es más sencilla.

1.1.4. Exponencial de una matriz

Antes de comenzar, recordemos un teorema importante sobre series. Se dice que la serie $\sum a_n$ converge absolutamente si converge la serie $\sum |a_n|$.

Teorema 1.1.1. *Supongamos que*

a) $\sum_{n=0}^{\infty} a_n$ *converge absolutamente*

b) $\sum_{n=0}^{\infty} a_n = A$,

c) $\sum_{n=0}^{\infty} b_n = B$,

d) $c_n = \sum_{k=0}^{n} a_n b_{n-k} \qquad (n = 0, 1, 2, \ldots)$

Se verificará que

$$\sum_{n=0}^{\infty} c_n = AB.$$

Esto es, el producto de dos series convergentes converge, haciéndolo además hacia el valor previsto, si al menos una de las dos series converge absolutamente.

La prueba de este teorema se pueden ver en los libros [3] y [4].

Sea, como hasta ahora, X un espacio vectorial complejo de dimensión n y sea $T \in L(X, X)$. Supongamos que

$$\sum_{k=0}^{\infty} a_k M^k$$

es una serie numérica con $a_k \geqslant 0$, $M > 0$, convergente. Si $\|T\| \leqslant M$, entonces la serie

$$\sum_{k=0}^{\infty} a_k T^k \quad donde \quad T^0 = I, \;\; identidad$$

satisface la condición de Cauchy, (recordemos que la condición de Cauchy nos dice que $\sum a_n$ converge si, y sólo si para cada $\varepsilon > 0$ existe un entero N tal que $|\sum_{k=n}^{m} a_k| \leqslant \varepsilon$ si $m \geqslant n \geqslant N$) es decir:

$$\left\| \sum_{k=j}^{m} a_k T^k \right\| \leqslant \sum_{k=j}^{m} a_k M^k \to 0 \; para \; j, m \to \infty, \quad j < m$$

y, así, define un operador lineal

$$S = \sum_{k=0}^{\infty} a_k T^k$$

En particular, $\sum_{k=0}^{\infty} \dfrac{M^k}{k!}$ converge a e^M para cualquier M real y, así, para todo $T \in L(X, X)$ se puede definir el operador

$$\exp(T) = e^T = \sum_{k=0}^{\infty} \frac{T^k}{k!}$$

que satisface las propiedades siguientes:

Propiedad 1.1.2. *Algunas propiedades importantes de* $\exp(T) = e^T$

1. $\|\exp(T)\| \leqslant \exp(\|T\|)$
2. *Si* $A, B \in L(X, X)$ *y* $AB = BA$ *entonces* $\exp(A+B) = \exp A \exp B$
3. $\exp A \exp(-A) = I = \exp(-A) \exp A$
4. $(\exp A)^{-1} = \exp(-A)$

Demostración. A continuación demostraremos estas propiedades

1. $$\|\exp(T)\| = \left\|\sum_{k=0}^{\infty} \frac{T^k}{k!}\right\| \leqslant \sum_{k=0}^{\infty} \frac{\|T^k\|}{k!} \leqslant \sum_{k=0}^{\infty} \frac{\|T\|^k}{k!} = \exp(\|T\|)$$

2. Por el binomio de Newton tenemos que

$$(A+B)^k = \sum_{j=0}^{k} \binom{k}{j} A^j B^{k-j} = \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} A^j B^{k-j}$$

y así

$$\begin{aligned} \exp(A+B) &= \sum_{k=0}^{\infty} \frac{(A+B)^k}{k!} = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} A^j B^{k-j} \\ &= \sum_{k=0}^{\infty} \sum_{j=0}^{k} \frac{1}{k!} \frac{k!}{j!(k-j)!} A^j B^{k-j} \\ &= \sum_{k=0}^{\infty} \sum_{j=0}^{k} \frac{1}{j!(k-j)!} A^j B^{k-j} \text{ sea } h = k-j \text{ y } 0 \leqslant j \leqslant k < \infty \\ &= \sum_{j=0}^{\infty} \sum_{h=0}^{\infty} \frac{1}{j!h!} A^j B^h = \left(\sum_{j=0}^{\infty} \frac{A^j}{j!}\right)\left(\sum_{h=0}^{\infty} \frac{B^h}{h!}\right) \text{ por teorema 1.1.1} \\ &= \exp(A)\exp(B) \end{aligned}$$

estando todos los cambios de orden de sumación justificados por la convergencia absoluta de las series $\exp(A)$, $\exp(B)$, es decir, la convergencia de las series $\exp(\|A\|)$, $\exp(\|B\|)$.

3. Es una consecuencia inmediata a partir de la propiedad anterior ya que

$$\exp(A)\exp(-A) = \exp(A-A) = \exp(0) = I = \exp(-A+A) = \exp(-A)\exp(A)$$

4. Por la propiedad anterior se tiene

$$\exp(A)\exp(-A) = I \Rightarrow \exp(-A) = \frac{I}{\exp(A)}$$

esto es $(\exp(A))^{-1} = \exp(-A)$, es decir, la $\exp(A)$ es invertible.

□

Por tanto, la matriz $\exp(A)$ nunca es singular.
Como veremos más adelante, el cálculo explícito de $\exp(A)$ es sencillo una vez que se conoce la forma canónica de Jordan de A.

1.1.5. Derivación

Consideremos ahora una función de $\mathbf{R}$ con valores en el espacio normado $L(X,X)$

$$A : t \in \mathbf{R} \to A(t) \in L(X,X)$$

Tiene sentido hablar de límites, continuidad, una vez que tenemos una topología en $L(X,X)$. En particular, si $t_0 \in \mathbf{R}$ podemos definir la derivada $A'(t_0)$ de $A(t)$ en t_0 del siguiente modo:

$$A'(t_0) = \lim_{h \to 0} \frac{A(t_0+h) - A(t_0)}{h}$$

cuando este límite existe.
La existencia de tal límite tiene lugar si, y sólo si, todas las componentes $a_{ij}(t)$ de $A(t)$ son derivables en t_0 y la derivada entonces es precisamente

$$A'(t_0) = (a'_{ij}(t_0))$$

Para verlo basta recordar que todas las normas en $L(X,X)$ son equivalentes y una de ellas es precisamente

$$\sum_{i,j=1}^{n} |a_{ij}| \quad para \quad A = (a_{ij})$$

La derivada tiene las propiedades siguientes. Si $A(t)$ y $B(t)$ admiten derivadas en t_0, entonces la función $S = A + B$ definida por $S(t) = A(t) + B(t)$ satisface

$$(A+B)'(t_0) = A'(t_0) + B'(t_0)$$

y la función $P = AB$ definida por $P(t) = A(t)B(t)$ satisface

$$P'(t_0) = A'(t_0)B(t_0) + A(t_0)B'(t_0)$$

La demostración es rutinaria.
Si $A'(t_0)$ existe, y también $(A(t_0))^{-1}$, entonces, siendo $A(t)$ continua en t_0 y, por tanto, $\det A(t)$ una función continua de $\mathbf{R}$ a $\mathbf{C}$ que no se anula en t_0, existe un intervalo $[t_0 - h, t_0 + h]$ de t_0 en el que $\det A(t) \neq 0$, y, así, se puede definir:

$$J : t \in [t_0 - h, t_0 + h] \to J(t) = (A(t))^{-1}$$

y se tiene, además,

$$\lim_{s \to 0} (A(t_o + s))^{-1} = (A(t_0))^{-1}$$

También se verifica:

$$J'(t_0) = -(A(t_0))^{-1}A'(t_0)(A(t_0))^{-1}$$

En efecto:

$$\begin{aligned}
J'(t_0) &= \lim_{s\to 0}\frac{J(t_0+s)-J(t_0)}{s}\ \textit{por como esta definida } J(t)\\
&= \lim_{s\to 0}\frac{(A(t_0+s))^{-1}-(A(t_0))^{-1}}{s}\ \textit{multiplico por la identidad } I=(A(t_0))^{-1}A(t_0)\\
&= \lim_{s\to 0}\frac{A(t_0))^{-1}A(t_0)(A(t_0+s))^{-1}-(A(t_0))^{-1}}{s}\ \textit{saco factor común } (A(t_0))^{-1}\\
&= \lim_{s\to 0}(A(t_0))^{-1}\frac{A(t_0)(A(t_0+s))^{-1}-I}{s}\ \textit{reescribimos } I=A(t_0+s)(A(t_0+s))^{-1}\\
&= \lim_{s\to 0}(A(t_0))^{-1}\frac{A(t_0)(A(t_0+s))^{-1}-A(t_0+s)(A(t_0+s))^{-1}}{s}\\
&= \lim_{s\to 0}(A(t_0))^{-1}\frac{A(t_0)-A(t_0+s)}{s}(A(t_0+s))^{-1}\\
&= -(A(t_0))^{-1}A'(t_0)(A(t_0))^{-1}
\end{aligned}$$

En general se tiene, para $A, B \in L(X,X)$,

$$\det(AB) = (\det A)(\det B),$$

y si se define la traza de A como

$$trA = \sum_{j=1}^{n} a_{jj}$$

entonces:

$$tr(AB) = tr(BA)$$

lo que se comprueba directamente mediante la definición de determinante y traza.

Es sencillo ver que si $A(t)$ es derivable, entonces:

$$\begin{aligned}
\frac{d}{dt}[detA(t)] &= \det\begin{pmatrix} a'_{11}(t) & a_{12}(t)\dots a_{1n}(t)\\ a'_{21}(t) & a_{22}(t)\dots a_{2n}(t)\\ \dots\dots & \dots\dots\\ a'_{n1}(t) & a_{n2}(t)\dots a_{nn}(t)\end{pmatrix} + \det\begin{pmatrix} a_{11}(t) & a'_{12}(t)\dots a_{1n}(t)\\ a_{21}(t) & a'_{22}(t)\dots a_{2n}(t)\\ \dots\dots & \dots\dots\\ a_{n1}(t) & a'_{n2}(t)\dots a_{nn}(t)\end{pmatrix}\\
&+ \quad \dots + \det\begin{pmatrix} a_{11}(t) & a_{12}(t)\dots a'_{1n}(t)\\ a_{21}(t) & a_{22}(t)\dots a'_{2n}(t)\\ \dots\dots & \dots\dots\\ a_{n1}(t) & a_{n2}(t)\dots a'_{nn}(t)\end{pmatrix}
\end{aligned}$$

es decir, la derivada del determinante de $A(t)$ es la suma de los determinantes de las matrices obtenidas sustituyendo en $A(t)$ una fila (columna) por su derivada.

1.1.6. Ejemplo

Ejemplo 1.1.1. *Se considera un espacio vectorial complejo V de base $\mathfrak{B} = \{e_1, e_2, \ldots, e_n\}$ y un operador lineal $T : V \to V$ definido por las siguientes relaciones:*

$$\begin{aligned} T(e_1 + e_2) &= 4e_3 \\ T(e_1 - e_2) &= 2e_2 \\ T(e_2 + e_3) &= 2e_3 - e_2 \end{aligned}$$

Hállese la expresión matricial de T respecto de la base $\mathfrak{B}$.

Solución. *Por la linealidad de T se tiene*

$$T(e_1 + e_2) = T(e_1) + T(e_2) = 4e_3 \tag{1.1.3}$$

$$T(e_1 - e_2) = T(e_1) - T(e_2) = 2e_2 \tag{1.1.4}$$

$$T(e_2 + e_3) = T(e_2) + T(e_3) = 2e_3 - e_2 \tag{1.1.5}$$

al sumar las ecuaciones (1.1.3) y (1.1.4) se tiene

$$2T(e_1) = 4e_3 + 2e_2 \Rightarrow T(e_1) = 2e_3 + e_2$$

y al restar las ecuaciones (1.1.3) y (1.1.4) se tiene

$$2T(e_2) = 4e_3 - 2e_2 \Rightarrow T(e_2) = 2e_3 - e_2$$

luego esta última ecuación junto con (1.1.5) nos da que $T(e_3) = 0$, entonces se tiene que

$$\begin{aligned} T(e_1) &= 2e_3 + e_2 \\ T(e_2) &= 2e_3 - e_2 \\ T(e_3) &= 0 \end{aligned}$$

así la matriz de T respecto a la base $\mathfrak{B} = \{e_1, e_2, \ldots, e_n\}$ es

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 2 & 2 & 0 \end{pmatrix}$$

de forma que

$$T(x) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 2 & 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

notemos que $T(e_i)$, $i = 1, 2, 3$ nos da las columnas de la matriz A como vimos en la subsección 1.1.1.

1.2. Teoría espectral elemental. Teorema de Jordan

Sea X un espacio vectorial complejo de dimensión n y sea $\mathfrak{B} = \{e_1, e_2, \ldots, e_n\}$ una base de X. Sea $T \in L(X, X)$ tal que $Te_1 = \lambda_1 e_1$, $\lambda_1 \in \mathbf{C}$. Entonces recordando que la expresión matricial $T_{\mathfrak{B}}$ de T respecto de la base $\mathfrak{B}$ según la subsección 1.1.1 es tal que $Te_j = \sum_{i=1}^{n} a_{ij} e_i$, es decir, que la columna j-ésima de $T_{\mathfrak{B}}$ está formada por las coordenadas de Te_j respecto de $\mathfrak{B}$ es de la forma:

$$T_{\mathfrak{B}} = \begin{pmatrix} \lambda_1 & a_{12} & a_{12} \ldots a_{1n} \\ 0 & a_{22} & a_{23} \ldots a_{2n} \\ \vdots & \vdots & \vdots \quad\quad \vdots \\ 0 & a_{n2} & a_{n3} \ldots a_{nn} \end{pmatrix}$$

Asimismo, si $Te_i = \lambda_i e_i$ para $i = 1, 2 \ldots, n$ entonces:

$$T_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & . \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

es decir la expresión de T respecto de $\mathcal{B}$ es extraordinariamente sencilla y hace transparente la estructura de T.

Veamos cómo varía $T_{\mathcal{B}}$ por un cambio de base.

Sea $P = (p_{ij})$ una matriz $n \times n$ no singular de elementos complejos y sea

$$\tilde{e}_i = \sum_{j=1}^{n} p_{ji} e_j$$

es decir, escribimos la nueva base como combinación lineal de los elementos de la base anterior de forma única.

La nueva base es $\tilde{\mathbf{B}} = \{\tilde{e}_1, \ldots, \tilde{e}_n\}$. Si la expresión de T respecto de $\tilde{\mathbf{B}}$ es $T_{\tilde{\mathcal{B}}} = \left(\tilde{t}_{ij}\right)$, esto quiere decir, según la subsección 1.1.1, que

$$T\tilde{e}_j = \sum_{i=1}^{n} \tilde{t}_{ij} \tilde{e}_i$$

es decir, teniendo en cuenta que $\tilde{e}_j = \sum_{l=1}^{n} p_{lj} e_l$:

$$T\tilde{e}_j = \sum_{j=1}^{n} p_{lj} Te_l = \sum_{i,k=1}^{n} \tilde{t}_{ij} p_{ki} e_k$$

Como $Te_l = \sum_{k=1}^{n} t_{kl} e_k$, resulta:

$$\sum_{l,k=1}^{n} t_{kl} p_{lj} e_k = \sum_{i,k=1}^{n} p_{ki} \tilde{t}_{ij} e_k$$

y siendo $\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ una base:

$$\sum_{l=1}^{n} t_{kl} p_{lj} = \sum_{i=1}^{n} p_{ki} \tilde{t}_{ij}$$

es decir, $T_{\mathcal{B}}P = PT_{\tilde{\mathcal{B}}}$ entonces

$$T_B = PT_{\tilde{\mathcal{B}}}P^{-1}$$

Ejemplo 1.2.1. *Si*

$$T_{\mathcal{B}} = \begin{pmatrix} 17 & 3 & -6 \\ -4 & 0 & 2 \\ 28 & 5 & -9 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 2 & 1 & 3 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 5 & 1 & -2 \\ 2 & 1 & -1 \\ -4 & -1 & 2 \end{pmatrix}$$

encontrar $T_{\tilde{\mathcal{B}}}$.

Solución.

$$T_{\mathcal{B}}P = \begin{pmatrix} 17 & 3 & -6 \\ -4 & 0 & 2 \\ 28 & 5 & -9 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 5 & 0 & 2 \\ 0 & 2 & 2 \\ 10 & 1 & 6 \end{pmatrix}$$

luego

$$P^{-1}T_{\mathcal{B}}P = \begin{pmatrix} 5 & 1 & -2 \\ 2 & 1 & -1 \\ -4 & -1 & 2 \end{pmatrix} \begin{pmatrix} 5 & 0 & 2 \\ 0 & 2 & 2 \\ 10 & 1 & 6 \end{pmatrix} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

entonces

$$T_{\tilde{\mathcal{B}}} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Así, un mismo $T \in L(X, X)$ que tiene una representación complicada respecto de una base $\mathcal{B}$ admite otra muy sencilla respecto de $\tilde{\mathcal{B}}$.

El problema que nos proponemos a continuación consiste en determinar bases respecto de las cuales el operador T tenga una expresión tan simple como sea posible. Como se ha visto, este problema es equivalente al siguiente: dada una matriz $T_{\mathcal{B}}$ hállese una matriz no singular P tal que $T_{\tilde{\mathcal{B}}} = P^{-1}T_{\mathcal{B}}P$ tenga una expresión simple.

Una primera respuesta sencilla la da el siguiente teorema. La ***forma triangular*** que se obtiene es muy útil para diversos propósitos.

1.2.1. Teorema de Schur

Teorema 1.2.1. *Sea* $A \in M(n, \mathbf{C})$ *una matriz cualquiera. Entonces existe una matriz* $P \in M(n, \mathbf{C})$ *no singular tal que* $P^{-1}AP$ *tiene la forma siguiente:*

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & a_{12} & a_{13} \dots a_{1n} \\ 0 & \lambda_2 & a_{23} \dots a_{2n} \\ 0 & 0 & \lambda_3 \dots a_{3n} \\ \vdots & \vdots & \vdots \quad \vdots \\ 0 & 0 & 0 \dots \lambda_n \end{pmatrix}$$

Demostración. La demostración procede por inducción sobre n.
Para $n = 1$. Efectivamente, $A = [a]$ basta elegir $P = [1]$, entonces

$$P^{-1}AP = [1]^{-1}[a][1] = [1][a][1] = [a]$$

Supongamos el teorema cierto para $n = h$ y sea ahora A de dimensión $h+1$. Existe un número $\lambda_1 \in \mathbf{C}$ y un vector no nulo $x \in \mathbf{C}^{h+1}$ tal que $Ax = \lambda_1 x$. En efecto, si escribimos esta ecuación $(A - \lambda_1 I)x = 0$ y λ_1 es tal que $\det(A - \lambda_1 I) = 0$, entonces la ecuación $(A - \lambda_1 I)x = 0$ tiene solución $x \neq 0$. Así, basta escoger λ_1 como una de las $h+1$ raíces del polinomio de grado $h+1$ del $\det(A - \lambda_1 I)$ y resolver dicha ecuación. Tomamos ahora una base en $\mathbf{C}^{h+1}$ que tenga x como primer elemento. Sea Q la matriz no singular de cambio de base. Entonces, por ser $Ax = \lambda_1 x$, la expresión en la nueva base del operador lineal cuya expresión era A en la base inicial es

$$\begin{aligned}
AQ &= A\,[x_1, q_2, q_3, \ldots, q_n] \\
&= [Ax_1, Aq_2, Aq_3, \ldots, Aq_n] \quad \textit{recordemos} \quad \textit{que} \quad Ax = \lambda_1 x \\
&= [\lambda_1 x_1, Aq_2, Aq_3, \ldots, Aq_n] \\
&= [x_1, q_2, q_3, \ldots, q_n] \begin{pmatrix} \lambda_1 & b_{12} & b_{13} & \ldots & b_{1n} \\ 0 & b_{22} & b_{23} & \ldots & b_{2n} \\ 0 & b_{32} & b_{33} & \ldots & b_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & b_{n2} & b_{n3} & \ldots & b_{nn} \end{pmatrix}
\end{aligned}$$

luego, aplicando Q^{-1} a la izquierda se tiene

$$Q^{-1}AQ = \begin{pmatrix} \lambda_1 & b_{12} & b_{13} & \ldots & b_{1n} \\ 0 & b_{22} & b_{23} & \ldots & b_{2n} \\ 0 & b_{32} & b_{33} & \ldots & b_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & b_{n2} & b_{n3} & \ldots & b_{nn} \end{pmatrix} = B$$

Consideremos ahora la matriz:

$$B_1 = \begin{pmatrix} b_{22} & b_{23} & \ldots & b_{2n} \\ b_{32} & b_{33} & \ldots & b_{3n} \\ \vdots & \vdots & & \vdots \\ b_{n2} & b_{n3} & \ldots & b_{nn} \end{pmatrix} \in M(h, \mathbf{C})$$

Por la hipótesis inductiva existe P_1 tal que:

$$P_1^{-1}B_1P_1 = \begin{pmatrix} \lambda_2 & c_{23} & c_{24} \ldots c_{2n} \\ 0 & \lambda_3 & c_{33} \ldots\ c_{3n} \\ 0 & 0 & \lambda_4 \ldots\ c_{4n} \\ \vdots & \vdots & \vdots \quad\quad \vdots \\ 0 & 0 & 0\ \ldots\ \lambda_n \end{pmatrix}$$

siendo $\lambda_1, \lambda_2, \cdots, \lambda_n$ los valores propios de B_1.
Sea

$$R = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P_1 \end{array}\right) \in M(h+1, \mathbf{C})$$

Se tiene:

$$R^{-1} = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P_1^{-1} \end{array}\right)$$

Definimos $P = QR$. Por se P invertible se tiene que $P^{-1} = (QR)^{-1} = R^{-1}Q^{-1}$ entonces $P^{-1} = R^{-1}Q^{-1}$, y se tiene que:

$$\begin{aligned}
(QR)^{-1}A(QR) &= R^{-1}Q^{-1}AQR \\
&= \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P_1^{-1} \end{array}\right) Q^{-1}AQ \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P_1 \end{array}\right) \\
&= \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P_1^{-1} \end{array}\right) \left(\begin{array}{c|c} \lambda_1 & a_1 \\ \hline 0 & B_1 \end{array}\right) \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P_1 \end{array}\right) \\
&= \left(\begin{array}{c|c} \lambda_1 & a_1 \\ \hline 0 & P_1^{-1}B_1 \end{array}\right) \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P_1 \end{array}\right) \\
&= \left(\begin{array}{c|c} \lambda_1 & a_1P_1 \\ \hline 0 & P_1^{-1}B_1P_1^{-1} \end{array}\right)
\end{aligned}$$

así se tiene que

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & a_{12} & a_{13}\dots a_{1n} \\ 0 & \lambda_2 & a_{23}\dots\ a_{2n} \\ 0 & 0 & \lambda_3\dots\ a_{3n} \\ \vdots & \vdots & \vdots \qquad \vdots \\ 0 & 0 & 0\ \dots\ \lambda_n \end{pmatrix}$$

con lo que se demuestra el teorema. $\square$

La forma triangular que hemos obtenido es interesante, pero no puede compararse en simplicidad con la forma diagonal que obtuvimos en la introducción con respecto a la base $\mathcal{B} = \{e_1, e_2, \dots, e_n\}$ tal que $Te_j = \lambda_j e_j$ para $j = 1, 2, \dots, n$. Surge en seguida la cuestión: ¿Será posible, dada $T \in L(X, X)$, hallar una base $\mathcal{B}$ que satisfaga la condición? En otras palabras, dada una matriz A, ¿se puede encontrar P no singular tal que $P^{-1}AP$ sea diagonal? La respuesta es negativa.
Supongamos, en efecto:

$$A = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}, \quad P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad P^{-1}AP = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} = \tilde{A}$$

Entonces habría de ser $AP = P\tilde{A}$, es decir:

$$\begin{pmatrix} 2a & 2b \\ a+2c & b+2d \end{pmatrix} = \begin{pmatrix} \alpha a & \beta b \\ \alpha c & \beta d \end{pmatrix}$$

así, tenemos que:

$$\begin{cases} 2a = a\alpha \\ 2b = b\beta \\ a + 2c = c\alpha \\ b + 2d = d\beta \end{cases}$$

lo que implica $\alpha = 2$, $\beta = 2$, $a = 0$, $b = 0$, y así, P sería singular, en contradicción con la hipótesis. Por tanto, A no admite una expresión diagonal.
Hemos visto el papel que desempeñan, tanto en la diagonalización de una matriz como en su forma triangular, los vectores $x \in X$ tales que $Tx = \lambda x$. Aun cuando no siempre se puede encontrar una base formada por tales vectores, es de esperar que sistemas de vectores de este tipo puedan ser útiles en la tarea de simplificar en lo posible la expresión de un operador mediante un cambio de base. Este es el significado del teorema de Jordan y de las definiciones que siguen.

1.2.2. Definiciones básicas

1. Sea $T \in L(X, X)$ y $\lambda \in \mathbf{C}$. Se dice que λ es ***autovalor*** de T si existe $x \in X$, $x \neq 0$ tal que $Tx = \lambda x$, es decir, si el operador $T - \lambda I$ no es invertible.

2. Si λ es autovalor de T, cualquier $x \in X$, $x \neq 0$ tal que $Tx = \lambda x$ se denomina ***autovector*** de T (correspondiente al autovalor λ).

3. El conjunto de todos los autovalores de T se denomina ***espectro*** de T y se denota $\sigma(T)$.

4. Los números $\lambda \in \mathbf{C} - \sigma(T)$ se denomina ***valores regulares*** de T, es decir, $\lambda \in \mathbf{C}$ es valor regular de T si $T - \lambda I$ es inversible.

5. El espectro $\sigma(T)$ de un operador lineal cualquiera T no es nunca vacío ni tiene más de n puntos distintos, siendo n la dimensión de X. En efecto, si $\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ es una base de X y designamos por T la matriz expresión del operador T respecto de $\mathcal{B}$, $\lambda \in \mathbf{C}$ será autovalor de T si, y sólo si, $\det(T - \lambda I) = 0$ por el teorema de Rouché, y siendo $\det(T - \lambda I)$ un polinomio de grado n (*polinomio característico de* T) resulta que existen a lo sumo n puntos distintos en $\sigma(T)$ (las n raices del polinomio característico) y que siempre existe alguno.

6. El conjunto de autovectores de T correspondiente al autovalor λ es un subespacio de X que se denomina ***autoespacio*** de T (correspondiente a λ) y lo denotaremos

   $$N(\lambda, I) = \{x \in X : (T - \lambda I)^1(x) = 0\}$$

   Un vector x puede ser tal que $(T - \lambda I)^1(x) = 0$, pero $(T - \lambda I)^2(x) \neq 0$. Por eso introducimos el siguiente subespacio de X para $k = 0, 1, 2, \ldots$ y cualquier $\lambda \in \mathbf{C}$

   $$N(\lambda, k) = \{x \in X : (T - \lambda I)^k x = 0\}$$

Se verifica claramente

$$N(\lambda, k) \subset N(\lambda, k+1)$$

dado que, son subespacios $N(\lambda, k+1)$ debe de tener una dimensión mayor que $N(\lambda, k)$. Además, si

$$N(\lambda, k) = N(\lambda, k+1)$$

entonces, para todo $h = 2, 3, \ldots,$ se tiene

$$N(\lambda, k) = N(\lambda, k+h)$$

Probaremos lo anterior por inducción sobre h. En efecto, supongamos $N(\lambda, k+1) = N(\lambda, k)$ y sea $x \in N(\lambda, k+2)$, es decir,

$$(T - \lambda I)^{k+2} x = 0$$

Entonces:

$$(T - \lambda I)^{k+1}((T - \lambda I)x) = 0.$$

Así, por ser $(T - \lambda I)x \in N(\lambda, k+1) = N(\lambda, k)$, se tiene:

$$(T - \lambda I)^{k}((T - \lambda I)x) = (T - \lambda I)^{k+1} x = 0.$$

y, por tanto,

$$x \in N(\lambda, k+1) = N(\lambda, k)$$

Así:

$$N(\lambda, k+2) = N(\lambda, k)$$

Del mismo modo,

$$N(\lambda, k+h) = N(\lambda, k)$$

Por tanto, la cadena de subespacios

$$\{0\} = N(\lambda, 0), N(\lambda, 1), N(\lambda, 2), \ldots$$

es una cadena estrictamente creciente, y, puesto que la dimensión del espacio es finita, los subespacios anteriores no pueden crecer de manera indefinida, sino que para algún k, $N(\lambda, k+1) = N(\lambda, k)$, entonces se estaciona en este $N(\lambda, k)$ se tendrá que

$$N(\lambda, 0) \subsetneqq N(\lambda, 1) \subsetneqq N(\lambda, 2) \subsetneqq \ldots \subsetneqq N(\lambda, n) = N(\lambda, n+1),$$

Ahora bien, es claro que para $\lambda \in \mathbf{C}$ es imposible que se tenga:

$$N(\lambda, 0) \subsetneqq N(\lambda, 1) \subsetneqq N(\lambda, 2) \subsetneqq \ldots \subsetneqq N(\lambda, n) \subsetneqq N(\lambda, n+1),$$

ya que si así fuera existirían $n+1$ vectores linealmente independientes en X.
Por tanto, existe para cada $\lambda \in \mathbf{C}$ un entero mínimo $\nu(\lambda)$, denominado ***índice de*** λ, tal que

$$N(\lambda, \nu(\lambda) + 1) = N(\lambda, \nu(\lambda))$$

Es claro que $\nu(\lambda) \leqslant n$ pues, no se puede pasar de la dimensión n por que si se pasa ya no sería de dimensión n. Obsérvese, además que $\lambda \in \sigma(T)$ es equivalente a $\nu(\lambda) > 0$ ya que

- Supongamos que $\lambda \in \sigma(T)$ y demostremos que $\nu(\lambda) > 0$, por la definición de autovalor se tiene que

  $$Tx = \lambda x \Rightarrow (T - \lambda I)x = 0$$

  y por la cadena

  $$\left\{\vec{0}\right\} = N(\lambda, 0) \subsetneqq N(\lambda, 1) \subsetneqq N(\lambda, 2) \subsetneqq \ldots \subsetneqq N(\lambda, n) = N(\lambda, n+1),$$

  lo cual significa que $\nu(\lambda) > 0$ por que al menos debe ser 1 para detenerse.

- Ahora suponemos que $\nu(\lambda) > 0$ y veamos que $\lambda \in \sigma(T)$. Por la definición de índice tenemos

  $$N(\lambda, \nu(\lambda)) = N(\lambda, \nu(\lambda) + 1) \;\; con \;\; \nu(\lambda) > 0$$

  veamos la cadena estricta

  $$\left\{\vec{0}\right\} = N(\lambda, 0) \subsetneqq N(\lambda, 1) \subsetneqq N(\lambda, 2) \subsetneqq \ldots \subsetneqq N(\lambda, \nu(\lambda)) = N(\lambda, \nu(\lambda) + 1),$$

  tenemos que si $\left\{\vec{0}\right\} = N(\lambda, \nu(\lambda))$, es decir, el último subespacio (núcleo) es sólo el vector cero, entonces todos son iguales. Pero esto no es posible, lo cual implica que existe $x \neq 0$ con $x \in N(\lambda, \nu(\lambda))$ tal que $(T - \lambda I)^{\nu(\lambda)}x = 0$.
  Ahora analicemos los siguientes casos para $\nu(\lambda)$

  - Si $\nu(\lambda) = 1$ se tendrá que $(T - \lambda I)x = 0$, pero esto es $Tx = \lambda x$, así $\lambda \in \sigma(T)$.
  - Si $\nu(\lambda) \neq 1$ tendremos entonces que

    $$\begin{aligned} (T - \lambda I)^{\nu(\lambda)}x &= 0 \\ (T - \lambda I)(T - \lambda I)^{\nu(\lambda)-1}x &= 0 \;\; sea \;\; y = (T - \lambda I)^{\nu(\lambda)-1}x \\ (T - \lambda I)y &= 0 \\ Ty &= \lambda y \end{aligned}$$

    de donde, $\lambda \in \sigma(T)$.

  Así, tenemos $\nu(\lambda) > 0$ entonces $\lambda \in \sigma(T)$.

1.2.3. Teorema de Cayley-Hamilton

Sea

$$P(\lambda) = \sum_{j=0}^{k} \alpha_j \lambda^j$$

un polinomio en λ con coeficientes complejos, y sea $T \in L(X, X)$. Podemos escribir:

$$P(T) = \sum_{j=0}^{k} \alpha_j T^j \in L(X, X)$$

Como veremos a continuación, dos polinomios $P(\lambda), Q(\lambda)$, aún sin ser iguales, pueden ser tales que los operadores $P(T), Q(T)$ coincidan, dependiendo esto de la estructura misma de T, esencialmente de su espectro. Este es el contenido del siguiente teorema.

Teorema 1.2.2 (Cayley-Hamilton)**.** *Sean $P(\lambda)$ y $Q(\lambda)$ dos polinomios y $T \in L(X,X)$ un operador lineal. Entonces $P(T) = Q(T)$ si, y sólo si, $P(\lambda)-Q(\lambda)$ tiene un cero de orden $\nu(\mu)$ para cada $\mu \in \sigma(T)$.*

Demostración. Sin pérdida de generalidad podemos suponer $Q(\lambda) \equiv 0$. Se trata de demostrar entonces que $P(T) = 0$ si, y sólo si, $P(\lambda)$ tiene un cero de orden $\nu(\mu)$ para cada $\mu \in \sigma(T)$.
Supongamos que $P(\lambda)$ tiene un cero de orden $\nu(\mu)$ para cada $\mu \in \sigma(T)$. Construiremos primero un polinomio $\bar{R}(\lambda) = 0$ tal que $\bar{R}(T) = 0$ y demostraremos que $P(\lambda)$ es múltiplo de $\bar{R}(\lambda)$. Así, $P(T) = 0$.
La construcción de $\bar{R}(\lambda)$ se realiza como sigue:
Tomamos una base $\{x_1, x_2, \ldots, x_n\}$ de X. Entonces los $n+1$ vectores

$$x_1, Tx_1, T^2x_1, \ldots, T^nx_1$$

son linealmente dependientes (puesto que tenemos un vector más que la dimensión en la que estamos trabajando, de lo cual resulta que uno de ellos es combinación lineal de los otros), es decir, existen a_j, $j = 0, 1, \ldots, n$, complejos no todos nulos tales que

$$\sum_{j=0}^{n} a_j T^j x_1 = 0$$

Así, si $S_1(\lambda)$ es el polinomio

$$S_1(\lambda) = \sum_{j=0}^{n} a_j \lambda^j$$

es claro que $S_1(T)x_1 = 0$ (de la observación anterior al teorema tenemos $S_1(T)x_1 = \sum_{j=0}^{n} a_j T^j x_1 = 0$).
Análogamente existen $S_j(\lambda)$, $j = 2, 3, \ldots, n$, polinomios no idénticamente nulos tales que $S_j(T)x_j = 0$ con $x_j \neq 0, x \in X$, para algún j. Tomemos ahora el polinomio

$$R(\lambda) = \prod_{j=1}^{n} S_j(\lambda)$$

Claramente, $R(T)x_j = 0$ para $j = 1, 2, \ldots, n$ y, así, $R(T)x = 0$ para todo $x \in X$, es decir, $R(T) = 0$.
Ahora reescribiremos $R(\lambda)$ como producto de la factorización de los polinomios $S_j(\lambda)$ con sus respectivas multiplicidades, de donde, resultara que algunos de ellos no son autovalores de $S_j(\lambda)$. Sea

$$R(\lambda) = a \prod_{h=1}^{m} (\lambda - \mu_h)^{\alpha_h}$$

Veamos que se pueden suprimir en $R(\lambda)$ algunos factores, conservándose aún la propiedad $R(T) = 0$.
En efecto, sea

$$\mu_1 \notin \sigma(T) \quad y \quad R^*(\lambda) = a \prod_{h=2}^{m} (\lambda - \mu_h)^{\alpha_h}$$

Entonces

$$(T-\mu_1 I)^{\alpha_1} R^*(T)x = 0 \quad para \quad todo \quad x \in X$$

Si existiera $x \in X$ tal que $y = R^*(T)x \neq 0$, entonces $(T-\mu_1 I)^{\alpha_1} y = 0$, y esto implica que $\mu_1 \in \sigma(T)$ contra lo supuesto. Así se ha de tener necesariamente $R^*(T)x = 0$. De esta forma se pueden suprimir en $R(\lambda)$ todos los factores $(\lambda - \mu_h)^{\alpha_h}$ con $\mu_h \notin \sigma(T)$ y se obtiene:

$$\tilde{R}(\lambda) = a \prod_{h=1}^{s} (\lambda - \theta_h)^{\beta_h} \quad con \quad \theta_h \in \sigma(T)$$

un polinomio tal que $\tilde{R}(T) = 0$. Con el desarrollo anterior se ha logrado factorizar cada $S_j(\lambda)$, de tal forma que sólo se tenga las raíces de estos polinomios.

Demostramos a continuación que los exponentes β_h se pueden todavía rebajar. Sabemos que para todo $x \in X$ se verifica

$$\tilde{R}(T) = a \prod_{h=1}^{s} (T - \theta_h I)^{\beta_h} x = 0$$

y, por tanto,

$$a(T-\theta_h I)^{\beta_1} \prod_{h=2}^{s} (T - \theta_h I)^{\beta_h} x = 0$$

Si $\beta_1 \geqslant \nu(\theta_1)$, entonces, en virtud de la definición de $\nu(\theta_1)$ se tiene también, para todo $x \in X$,

$$a(T-\theta_h I)^{\nu(\theta_1)} \prod_{h=2}^{s} (T - \theta_h I)^{\beta_h} x = 0$$

Por tanto, si ponemos $\eta_1 = \text{mín}(\beta_1, \nu(\theta_1))$, se tiene:

$$a(T-\theta_h I)^{\eta_1} \prod_{h=2}^{s} (T - \theta_h I)^{\beta_h} x = 0$$

para todo $x \in X$. Procediendo así con $\theta_2, \theta_3, \ldots, \theta_s$, resulta que si

$$\eta_h = \text{mín}(\beta_h, \nu(\theta_h)),$$

entonces el polinomio

$$\bar{R}(\lambda) = a \prod_{h=1}^{s} (\lambda - \theta_h)^{\eta_h}$$

es tal que $\bar{R}(T) = 0$.

Ahora bien, si $P(\lambda)$ tiene un cero de orden $\nu(\theta_h)$ para cada $\theta_h \in \sigma(T)$, entonces $P(\lambda)$ es múltiplo de $\bar{R}(\lambda)$ y, así, $P(T) = 0$.

Supongamos ahora que

$$P(\lambda) = c \prod_{j=1}^{p} (\lambda - \lambda_j)^{\alpha_j}$$

es tal que $P(T) = 0$. Demostraremos que cada autovalor $\mu \in \sigma(T)$, $(\lambda - \mu)^{\nu(\mu)}$ divide a $P(\lambda)$. Es decir que μ coincide con alguno de los λ_j y que el exponente correspondiente α_j es mayor o igual que $\nu(\mu)$.

En efecto, sea $\mu \in \sigma(T)$, $x \neq 0$ tal que $Tx = \mu x$. Entonces se puede escribir, aplicando P a ambos lados de la igualdad

$$P(T)x = P(\mu)x = 0;$$

y, así, siendo $x \neq 0$, ha de ser $P(\mu) = 0$, lo que demuestra que μ es raíz de $P(\lambda)$ y así coincide con algún λ_j, por ejemplo, con λ_1. Veamos que entonces $\alpha_1 \geqslant \nu(\lambda_1)$. Supongamos que fuese $\alpha_1 < \nu(\lambda_1)$. Entonces existiría, en virtud de la definición de $\nu(\lambda_1)$, un $x \in X$, $x \neq 0$ tal que

$$y = (T - \lambda_1 I)^{\alpha_1} x \neq 0,$$

y además,

$$(T - \lambda_1 I)^{\alpha_1 + 1} x = (T - \lambda_1 I)y = 0$$

Entonces podemos escribir, puesto que $Ty = \lambda_1 y$.
Como $y \neq 0$, esto implica $\lambda_1 = \lambda_j$ para algún $j \neq 1$, lo que es contradictorio con la factorización de $P(\lambda)$. Así, $\alpha_1 \geqslant \nu(\lambda_1)$ y lo mismo para los demás autovalores. Esto concluye la demostración del teorema. $\square$

Obsérvese que el teorema anterior se puede expresar de la forma siguiente, donde $P^m(\lambda)$ denota el polinomio obtenido a partir de $P(\lambda)$ derivando m veces.

Teorema 1.2.3. *Sea $T \in L(X, X)$ y $P(\lambda)$ un polinomio en λ con coeficientes complejos. Entonces $P(T) = 0$ si, y sólo si, $P^m(\mu) = 0$ para todo $\mu \in \sigma(T)$ y todo m, $0 \leqslant m \leqslant \nu(\mu) - 1$.*

Notemos que este teorema nos está diciendo el comportamiento del polinomio en el operador $T \in L(X, X)$ se reduce a estudiar el comportamiento del polinomio evaluado en un autovalor del espectro al derivar un cierto número de veces.
El teorema de Cayley-Hamilton permite obtener operadores con propiedades prefijadas construyendo polinomios adecuados. Presentamos a continuación un corolario del teorema que permite obtener una descomposición del espacio X en una suma directa sobre los que T actúa de una forma simple. Esta descomposición nos conducirá finalmente a la descomposición de Jordan. Anteponemos un lema algebraico que utilizaremos para dicho corolario.

Lema 1.2.1. *Se dan k puntos de $\mathbf{C}$, $a_1, a_2, \ldots, a_k$ y para cada $j = 1, 2, \ldots, k$ se dan $r_j + 1$ números complejos α_j^h, $h = 0, 1, \ldots, r_j$. Entonces existe un polinomio complejo $P(\lambda)$ tal que se tiene $P^h(a_j) = \alpha_j^h$.*

Demostración. Para simplificar la notación construimos el polinomio en un caso sencillo en el que queda clara la idea para el caso general. Se dan $a, \alpha_0, \alpha_1, \alpha_2, b, \beta_0, \beta_1, c, \gamma_0, \gamma_1, \gamma_2, \gamma_3$, todos ellos complejos, y se pide hallar $P(\lambda)$, polinomio tal que

$$\begin{aligned}
P(a) &= \alpha_0, \quad P'(a) = \alpha_1, \quad P''(a) = \alpha_2 \\
P(b) &= \beta_0, \quad P'(b) = \beta_1 \\
P(c) &= \gamma_0, \quad P'(c) = \gamma_1, \quad P''(c) = \gamma_2, \quad P'''(c) = \gamma_3
\end{aligned}$$

Tomamos el polinomio:

$$\begin{aligned} P(\lambda) &= c_0 + c_1(\lambda - a) + c_2(\lambda - a)^2 + (\lambda - a)^3 \left[c_3 + c_4(\lambda - b)\right] + \\ &+ (\lambda - a)^3(\lambda - b)^2 \left[c_5 + c_6(\lambda - c) + c_7(\lambda - c)^2 + c_8(\lambda - c)^3\right], \end{aligned}$$

donde los c_j se determinan del siguiente modo. Hacemos $\lambda = a$ y resulta $\alpha_0 = c_0$. Derivamos respecto de λ y hacemos $\lambda = a$ y se obtiene $\alpha_1 = c_1$. Derivamos dos veces y hacemos $\lambda = a$ y se obtiene $\alpha_2 = 2c_2$. Hacemos $\lambda = b$ y resulta

$$\beta_0 = c_0 + c_1(b - a) + c_2(b - a)^2 + c_3(b - a)^3;$$

es decir, c_3 se obtiene por recurrencia, conocidos c_0, c_1, c_2. Derivamos y hacemos $\lambda = b$ y resulta c_4 conocidos c_0, c_1, c_2, c_3, etc. $\square$

Corolario 1.2.1. *Sea $T \in L(X, X)$ y $\sigma(T) = \{\lambda_1, \lambda_2, \ldots, \lambda_p\}$. Para cada $i = 1, 2, \ldots, p$. Sea $E_i(\lambda)$ un polinomio en λ tal que:*

i) $E_i(\lambda_i) = 1, \quad E_i^{(m)}(\lambda_i) = 0$ *para* $1 \leqslant m \leqslant \nu(\lambda_i) - 1$

ii) $E_i^{(m)}(\lambda_j) = 0$ *para* $1 \leqslant m \leqslant \nu(\lambda_j) - 1, \quad j \neq i$

Entonces se tiene:

a) $E_i(T) \neq 0$.

b) $(E_i(T))^2 = E_i(T)$, $E_i(T)E_j(T) = 0$ *para* $i \neq j$.

c) $I = \sum_{i=1}^{p} E_i(T)$

Demostración.

a) Por el teorema 1.2.3 tenemos que $E(T) = 0$, si y sólo si, $E^{(m)}(\lambda_i) = 0$ para todo $\lambda_i \in \sigma(T)$ y todo m, $0 \leqslant m \leqslant \nu(\lambda_i) - 1$. Pero por la condición $i)$ del corolario se tiene que $E_i(\lambda_i) = 1, \quad E_i^{(m)}(\lambda_i) = 0$ *para* $1 \leqslant m \leqslant \nu(\lambda_i) - 1$. Por lo tanto $E_i(T) = 1 \neq 0$.

b) De $(E_i(T))^2 = E_i(T)$ tendremos que

$$(E_i(T))^2 - E_i(T) = 0 \tag{1.2.1}$$

Si no derivamos la ecuación anterior, por la condición $i)$ del corolario y teniendo en cuenta de nuevo el teorema 1.2.3 tendremos que $(E_i(T))^2 - E_i(T) = 1^2 - 1 = 0$.
Si derivamos la ecuación (1.2.1) tendremos $[(E_i(T))^2 - E_i(T)]' = 2E_i(T)E_i'(T) - E_i'(T)$ por $i)$

$$[(E_i(\lambda_i))^2 - E_i(\lambda_i)]' = 2E_i(\lambda_i)E_i'(\lambda_i) - E_i'(\lambda_i) = 0$$

para $m = 1$ se tiene $E_i'(\lambda_i) = 0$.
Luego $E_i(T)E_j(T) = 0$ para $i \neq j$ al derivar se tendrá $E_i'(T)E_j(T) + E_i(T)E_j'(T) = 0$ por el teorema 1.2.3 tenemos $E_i'(\lambda_i)E_j(\lambda_i) + E_i(\lambda_i)E_j'(\lambda_i) = 0$ y por la condición $ii)$ se cumple la igualdad, lo mismo sucede para λ_j.

c) Hacemos $I - \sum_{i=1}^{p} E_i(T) = 0$ pero de nuevo ocupando el teorema 1.2.3, puesto que dicho teorema nos dice que el comportamiento del polinomio en el operador $T \in L(X, X)$ se reduce a estudiar el comportamiento del polinomio evaluado en un autovalor del espectro al derivar un cierto número de veces. Así $I - \sum_{i=1}^{p} E_i(\lambda_i) = I - \sum_{i=1} 1 = 0$. Por tanto $I = \sum_{i=1}^{p} E_i(T)$.

□

Con las propiedades $a), b)$ y $c)$ del corolario anterior es de comprobación inmediata que si

$$A_j = E_j(T)X, \quad j = 1, 2, \ldots, p$$

entonces

$$X = A_1 \oplus A_2 \oplus \ldots \oplus A_p$$

Recordemos las propiedades de suma directa. Diremos que X es suma directa de los subespacios A_i cumplen

- $A_i \cap A_j = \varnothing$
- $X = A_1 + A_2 + \ldots + A_p$

En efecto, cualquier $x \in X$ se puede expresar como

$$x = \sum_{i=1}^{p} E_i(T)x$$

notemos que esto resulta de multiplicar por x el inciso $c)$ del corolario 1.2.1 y si z es de A_i y del espacio lineal engendrado por los A_j con $j \neq i$, entonces

$$z = E_i x_i, \quad z = \sum_{j \neq i} E_j(T)x_j$$

Así,

$$\begin{aligned} E_i(T)z &= E_i(T)E_i x_i \ \ \textit{se a sustituido} \ \ z = E_i x_i \\ &= (E_i(T))^2 x_i \ \ \textit{por} \ \ b) \\ &= E_i(T)x_i \\ &= z \end{aligned}$$

$$\begin{aligned} E_i(T)z &= E_i(T)\sum_{j \neq i} E_j(T)x_j \ \ \textit{se a sustituido} \ \ z = \sum_{j \neq i} E_j(T)x_j \\ &= \sum_{j \neq i} E_i(T)E_j(T)x_j \ \ \textit{como} \ \ E_i(T)z = z \\ &= z \end{aligned}$$

por lo anterior y por b)

$$E_i(T)z = (E_i(T))^2 x_i = E_i(T)x_i = z = \sum_{j\neq i} E_i(T)E_j(T)x_j = 0$$

También se verifica $TA_i \subset A_i$. En efecto, si $z \in TA_i$,

$$z = TE_i(T)y = E_i(T)Ty \in A_i$$

Por tanto, el estudio de T se reduce al estudio de cada una de las restricciones de T a A_i.
Si en X se toma una base formada por la unión de bases de cada uno de los subespacios A_i, entonces, recordando la subsección 1.1.1, resulta que la expresión de T es una matriz del tipo

$$\begin{pmatrix} \begin{array}{c|c|cc} M_1 & 0 & & \\ \hline 0 & M_2 & & \\ \hline & & \ddots & \multicolumn{1}{|c}{0} \\ \hline & & 0 & \multicolumn{1}{|c}{M_p} \end{array} \end{pmatrix}$$

siendo M_i una matriz cuadrada de dimensión igual a la dimensión de A_i.
El teorema siguiente identifica el espacio A_i con el espacio $N(\lambda_i, \nu(\lambda_i))$. Esto nos permitirá inmediatamente expresar T de una forma aún más simple, escogiendo en cada A_i una base adecuada.

Teorema 1.2.4. *Con las notaciones del corolario anterior y las observaciones precedentes se tiene para cada $\lambda_i \in \sigma(t)$:*

$$A_i = E_i(T)X = N(\lambda_i, \nu(\lambda_i))$$

Además,

$$\nu(\lambda_i) \leqslant \dim A_i$$

Demostración. Por el teorema de Cayley-Hamilton es inmediato que

$$(T - \lambda_i I))^{\nu(\lambda_i)} E_i(T) = 0$$

Esto demuestra que

$$A_i = E_i(T)X \subset N(\lambda_i, \nu(\lambda_i)) = \{x \in X : (T - \lambda_i I))^{\nu(\lambda_i)} x = 0\}$$

Como sabemos que $X = A_1 \oplus A_2 \oplus \ldots \oplus A_p$, a fin de demostrar que

$$N(\lambda_i, \nu(\lambda_i)) = A_i$$

bastará comprobar que:

$$P = N(\lambda_i, \nu(\lambda_i)) \bigcap l \left[\bigcup_{j\neq i} N(\lambda_i, \nu(\lambda_i)) \right] = \{0\}.$$

Supongamos que hubiese $x \neq 0$, $x \in P$. Sea α, con $0 \leqslant \alpha < \nu(\lambda_i)$ el entero mayor tal que $z = (T - \lambda_i I)^\alpha x \neq 0$. Entonces

$$(T - \lambda_i I)z = (T - \lambda_i I)^{\alpha+1}x = 0 \quad y \quad asi \quad Tz = \lambda_i z$$

Sea

$$x = \sum_{j \neq i} y_j \quad con \quad y_j \in N(\lambda_j, \nu(\lambda_j))$$

Entonces, si aplicamos a x el operador

$$(T - \lambda_i I)^\alpha \prod_{j \neq i} (T - \lambda_i I)^{\nu(\lambda_i)}$$

resulta

$$\prod_{j \neq i} (\lambda_i - \lambda_j) z = 0$$

lo que implica $z = 0$. Esta contradicción demuestra que $P = \{0\}$.
A fin de demostrar que se tiene

$$\dim A_i = \dim N(\lambda_i, \nu(\lambda_i)) \geqslant \nu(\lambda_i)$$

procederemos del siguiente modo. Sabemos que, por definición,

$$\{0\} \subsetneqq N(\lambda_i, 1) \subsetneqq N(\lambda_i, 2) \subsetneqq \ldots \subsetneqq N(\lambda_i, \nu(\lambda_i)) = N(\lambda_i, \nu(\lambda_i) + 1)$$

Así existen $\nu(\lambda_i)$ vectores $x_k \in N(\lambda_i, k)$, $k = 1, 2, \ldots, \nu(\lambda_i)$ linealmente independientes. Por tanto,

$$\dim A_i \geqslant \nu(\lambda_i)$$

Esto demuestra el teorema. $\square$

Observamos ahora que $(T - \lambda_i I)^{\nu(\lambda_i)} A_i = 0$ y llamemos $B_i = T - \lambda_i I$. Así resulta:

$$T|_{A_i} = B_i|_{A_i} + \lambda_i I|_{A_i} \quad con \quad B_i^{\nu(\lambda_i)} A_i = 0$$

Es decir, la restricción de T a A_i es la suma de un múltiplo de la identidad $\lambda_i I$ y de la restricción de B_i a A_i. Si logramos expresar $B_i|_{A_i}$ de una forma sencilla es claro que habremos logrado una expresión sencilla de T. Veamos ahora cómo se puede elegir en A_i una base para que $B_i|_{A_i}$ tenga una expresión sencilla (es decir buscamos una base de A_i para que nos quede la matriz T en la forma de Jordan). En lo que sigue, por comodidad de notación nos restringimos a considerar un A_i y suprimiremos los subíndices i.
Sabemos que se tiene

$$\{0\} \subsetneqq N(\lambda, 1) \subsetneqq N(\lambda, 2) \subsetneqq \ldots \subsetneqq N(\lambda, \nu(\lambda)) = A$$

Escogemos en $N(\lambda, \nu(\lambda))$ un espacio $H(\nu(\lambda)-1)$ complementario de $N(\lambda, \nu(\lambda)-1)$ en $N(\lambda, \nu(\lambda))$. Es decir, $H(\nu(\lambda)-1)$ es tal que

$$N(\lambda, \nu(\lambda)) = N(\lambda, \nu(\lambda)-1) \oplus H(\nu(\lambda)-1)$$

Sabemos que podemos encontrar una base de $N(\lambda, \nu(\lambda)-1) \in H(\nu(\lambda)-1)$la cual será $\{x_1, x_2, \ldots, x_p\}$ y otra de $N(\lambda, \nu(\lambda))$ la cual será $\{x_1, x_2, \ldots, x_p, x_{p+1}, \ldots x_q\}$ como son bases se puede escribir de forma única la sumatoria con $a \in A$, $a = \sum_{i=1}^{q} \alpha_i x_i = \sum_{j=1}^{p} \alpha_i x_i + \sum_{k=p+1}^{q} \alpha_k x_k$.
Entonces sea $\{x_1, x_2, \ldots, x_{h_1}\}$ una base de $H(\nu(\lambda)-1)$.
Veamos que

$$\{Bx_1, bx_2, \ldots, Bx_{h_1}\} \subset N(\lambda, \nu(\lambda)-1) - N(\lambda, \nu(\lambda)-2)$$

y además son linealmente independientes. En efecto, para lo primero se tiene, por ejemplo,

$$B^{\nu(\lambda)-1} Bx_1 = B^{\nu(\lambda)} x_1 = 0$$

pero, en cambio,

$$B^{\nu(\lambda)-2} Bx_1 = B^{\nu(\lambda)-1} x_1 \neq 0$$

ya que $x_1 \notin N(\lambda, \nu(\lambda)-1)$.
Para la independencia lineal se procede así. Si

$$\sum_1^N \alpha_i B x_i = 0, \quad \textit{entonces} \quad B^{\nu(\lambda)-1} \sum_1^N \alpha_i x_i = 0$$

(Podemos suponer $\nu(\lambda) > 1$, pues de otro modo todo lo que precede es trivial.) Como

$$\sum_1^{h_1} \alpha_i x_i \in H(\nu(\lambda)-1)$$

resulta que se ha de tener

$$\sum_1^{h_1} \alpha_i x_i = 0$$

Como $\{x_1, x_2, x_3, \ldots, x_{h_1}\}$ es base de $H(\nu(\lambda)-1)$, esto implica que $\alpha_i = 0$ para todo $i = 1, \ldots, h_1$. Así, $\{Bx_1, Bx_2, \ldots, Bx_{h_1}\}$ son vectores linalmente independientes de $N(\lambda, \nu(\lambda)-1) - N(\lambda, \nu(\lambda)-2)$. Por tanto, este conjunto de vectores se puede completar para obtener una base de un espacio $N(\lambda, \nu(\lambda)-2)$ complementario de $N(\lambda, \nu(\lambda)-2)$ en $N(\lambda, \nu(\lambda)-1)$, es decir, tal que

$$N(\lambda, \nu(\lambda)-1) = N(\lambda, \nu(\lambda)-2) \oplus H(\nu(\lambda)-2)$$

Sea esta base

$$\{Bx_1, Bx_2, \ldots, Bx_{h_1}, x_{h_1+1}, \ldots, x_{h_2}\}$$

Consideremos ahora el sistema de vectores:

$$\{B^2x_1, B^2x_2, \ldots, B^2x_{h_1}, Bx_{h_1+1}, \ldots, Bx_{h_2}\} \subset N(\lambda, \nu(\lambda)-2) - N(\lambda, \nu(\lambda)-3)$$

que son linealmente independientes. Procediendo análogamente, obtenemos finalmente una base de $A = N(\lambda, \nu(\lambda))$ formada por los vectores

$$\{x_1, x_2, \ldots, x_{h_1}, Bx_1, Bx_2, \ldots, Bx_{h_1}, x_{h_1+1}, \ldots, Bx_{h_2}, B^2x_1, B^2x_2, \ldots, B^2x_{h_1}, Bx_{h_1+1}, \ldots, Bx_{h_2}\}$$

Respecto de la base, ordenada de la forma siguiente:

$$\begin{array}{ll}
\{x_1, & Bx_1, B^2x_1, \ldots, B^{\nu(\lambda)-1}x_1, x_2, Bx_2, B^2x_2, \ldots, B^{\nu(\lambda)-1}x_2, \\
\ldots & \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \\
x_{h_1}, & Bx_{h_1}, B^2x_{h_1}, \ldots, B^{\nu(\lambda)-1}x_{h_1}, x_{h_1+1}, Bx_{h_1+1}, \ldots, B^{\nu(\lambda)-2}x_{h_1+1}, \\
\ldots & \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \\
x_{h_2}, & Bx_{h_2}, \ldots, B^{\nu(\lambda)-2}x_{h_2}, \\
\ldots & \ldots \ldots \ldots \ldots \ldots\}
\end{array}$$

el operador B tiene la expresión matricial siguiente:

$$\left(\begin{array}{c|c|c|c}
J_1 & 0 & & \\ \hline
0 & J_2 & & \\ \hline
 & \ddots & J_{H-1} & 0 \\ \hline
 & & 0 & J_H
\end{array}\right)$$

donde cada matriz J_i corresponde a cada sistema parcial $\{x_1, Bx_1, B^2x_1, \ldots, B^{\nu(\lambda)-1}x_1\}$ y tiene la forma:

$$J_i = \begin{pmatrix}
\lambda_i & 0 & 0 & 0 & 0 \\
1 & \lambda_i & 0 & 0 & 0 \\
0 & 1 & \lambda_i & 0 & 0 \\
0 & 0 & 1 & \lambda_i & 0 \\
\ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 0 & 0 & 1 & \lambda_i
\end{pmatrix}$$

Pero veamos como es que se ha obtenido la matriz J_i ella resulta de aplicar el operador T a cada elemento de la base ordenada, en particular para $\left\{x_1, Bx_1, B^2x_1, \ldots, B^{\nu(\lambda)-1}x_1\right\}$ para el resto de elementos resulta de manera análoga. Aplicando T a cada uno de ellos resulta

$$\begin{aligned}
T(x_1) &= \lambda x_1 + Bx_1 \\
T(Bx_1) &= B(Tx_1) = \lambda Bx_1 + B^2x_1 \\
T(B^2x_1) &= B^2(Tx_1) = \lambda B^2x_1 + B^3x_1 \\
T(B^3x_1) &= B^3(Tx_1) = \lambda B^3x_1 + B^4x_1 \\
&\ldots \\
T(B^{\nu(\lambda)-1}x_1) &= B^{\nu(\lambda)-1}(Tx_1) = \lambda B^{\nu(\lambda)-1}x_1 + B^{\nu(\lambda)}x_1
\end{aligned}$$

Por lo estudiado en la subsección 1.1.1, tenemos que la imagen de cada elemento se puede ver como un vector columna, de donde resulta que B es

$$
\begin{aligned}
B &= \begin{pmatrix} x_1 & Bx_1 & B^2x_1 & B^3x_1 & \dots & B^{\nu(\lambda)-1}x_1 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 & 0 & 0 \\ 1 & \lambda_1 & 0 & 0 & 0 \\ 0 & 1 & \lambda_1 & 0 & 0 \\ 0 & 0 & 1 & \lambda_1 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 1 & \lambda_1 \end{pmatrix} \\
&= \begin{pmatrix} x_1 & Bx_1 & B^2x_1 & \dots & B^{\nu(\lambda)-1}x_1 \end{pmatrix} \left[\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 1 & 0 \end{pmatrix} + \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix} \right] \\
&= \begin{pmatrix} x_1 & Bx_1 & B^2x_1 & B^3x_1 & \dots & B^{\nu(\lambda)-1}x_1 \end{pmatrix} [H_1 + \lambda_1 I]
\end{aligned}
$$

De todo esto resulta el teorema de descomposición de Jordan.

Teorema 1.2.5 (Teorema de descomposición de Jordan)**.** *Sea A una matriz $n \times n$ de elementos de $\mathbf{C}$. Existe entonces una matriz no singular P tal que $A = PJP^{-1}$, siendo J de la forma:*

$$
J = \left(\begin{array}{c|cc|c} J_1 & 0 & & \\ \hline 0 & J_2 & & \\ \hline 0 & & \ddots & 0 \\ \hline & & 0 & J_H \end{array}\right)
$$

donde cada J_i es de la forma:

$$
J_i = \begin{pmatrix} \alpha_i & & & 0 \\ 1 & & \alpha_i & \\ & \ddots & \ddots & \\ 0 & & 1 & \alpha_i \end{pmatrix}
$$

siendo α_i un autovalor de A.
La matriz J está unívocamente determinada salvo permutaciones de los bloques J_i. La matriz J se denomina la forma canónica de Jordan de A.

La determinación unívoca de J resulta de la unicidad del operador B estudiado anteriormente.

1.2.4. Forma canónica de una matriz real

Supongamos que T es una matriz real $n \times n$. Conviene a veces representar T mediante una descomposición *real* análoga a la de Jordan. Tal descomposición es posible en el sentido que se indica a continuación. La presentamos y demostramos en un caso particular a fin de evitar complicaciones de notación, pero el resultado es general y la marcha de la demostración es la misma.
Sea T una matriz real 5×5, representación de un operador lineal, que llamaremos también T, respecto de la base $\mathcal{B} = \{e_1, e_2, e_3, e_4, e_5\}$, del espacio X. Supongamos que la forma de Jordan de T es

$$J = \begin{pmatrix} a+bi & 0 & 0 & 0 & 0 \\ 1 & a+bi & 0 & 0 & 0 \\ 0 & 0 & a-bi & 0 & 0 \\ 0 & 0 & 1 & a-bi & 0 \\ 0 & 0 & 0 & 0 & c \end{pmatrix}$$

donde $a, b, c \in \mathbf{R}$. Veamos que entonces existe una matriz *real* no singular Q tal que $T = QKQ^{-1}$ donde:

$$K = \begin{pmatrix} a & b & 0 & 0 & 0 \\ -b & a & 0 & 0 & 0 \\ 1 & 0 & a & b & 0 \\ 0 & 1 & -b & a & 0 \\ 0 & 0 & 0 & 0 & c \end{pmatrix}$$

Para verlo volvamos a la demostración del teorema 1.2.5. Ante todo es claro que, siendo T real, si λ es un autovalor de T, entonces $\bar{\lambda}$ lo es también. Además, λ y $\bar{\lambda}$ tienen la misma multiplicidad, el mismo índice y las dimensiones de los espacios $N(\lambda, k)$, $N(\bar{\lambda}, k)$ son las mismas.
Al efectuar la elección de la base que conduce a la descomposición de Jordan, es claro asimismo que $x_1, x_2, \ldots, x_{h_1}$, con

$$x_j = \sum_{k=1}^{n} \alpha_{jk} e_k, \quad j = 1, \ldots, h_1$$

es base de $H(\nu(\lambda)-1)$, espacio complementario de $N(\lambda, \nu(\lambda)-1)$ en $N(\lambda, \nu(\lambda))$, y tomamos los vectores

$$\bar{x}_j = \sum_{k=1}^{n} \bar{\alpha}_{jk} e_k, \quad j = 1, \ldots, h_1$$

entonces $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{h_1}$, constituyen una base de un espacio $H(\nu(\bar{\lambda})-1)$ complementario de $N(\bar{\lambda}, \nu(\bar{\lambda})-1)$ en $N(\bar{\lambda}, \nu(\bar{\lambda}))$. Así, en la elección de la base de un espacio se puede proceder paralelamente, de la forma indicada con λ y $\bar{\lambda}$. Esta observación conduce con facilidad al resultado que buscamos.
Supongamos que hemos procedido así en el caso concreto que nos ocupa, y que hemos obtenido la base.

$$\{x_1, (T-(a+bi)I)x_1, \bar{x}_1, (T-(a+bi)I)\bar{x}_1, x_2\}$$

Llamemos

$$x_1 = y_1, (T-(a+bi)I)x_1 = y_2, \bar{x}_1 = \bar{y}_1, (T-(a+bi)I)\bar{x}_1 = \bar{y}_2, x_2 = y_3$$

y observamos que si

$$y_j = \sum_{k=1}^{5} \beta_{jk} e_k$$

entonces,

$$\bar{y}_j = \sum_{k=1}^{5} \bar{\beta}_{jk} e_k$$

La representación matricial del operador lineal T respecto de la base $\{y_1, y_2, \bar{y}_1, \bar{y}_2, y_3\}$ es

$$J = \begin{pmatrix} a+bi & 0 & 0 & 0 & 0 \\ 1 & a+bi & 0 & 0 & 0 \\ 0 & 0 & a-bi & 0 & 0 \\ 0 & 0 & 1 & a-bi & 0 \\ 0 & 0 & 0 & 0 & c \end{pmatrix}$$

Esto quiere decir:

$$\begin{aligned} Ty_1 &= (a+bi)y_1 + y_2 \\ Ty_2 &= (a+bi)y_2 \\ T\bar{y}_1 &= (a+bi)\bar{y}_1 + \bar{y}_2 \\ T\bar{y}_2 &= (a+bi)\bar{y}_2 \\ Ty_3 &= cy_3 \end{aligned}$$

Tomemos ahora los vectores

$$z_1 = \frac{y_1 + \bar{y}_1}{2}, \quad z_1^* = \frac{y_1 - \bar{y}_1}{2i}, \quad z_2 = \frac{y_2 + \bar{y}_2}{2}, \quad z_2^* = \frac{y_1 - \bar{y}_1}{2i}, z_3 = y_3$$

Es fácil comprobar que los vectores de $\{z_1, \bar{z}_1^*, y_2, \bar{y}_2^*, z_3\}$ tienen *coordenadas reales* respecto de $\{e_1, e_2, e_3, e_4, e_5\}$ y que constituyen una base Z. Asimismo es sencillo comprobar que

$$\begin{aligned} Tz_1 &= az_1 - bz_1^* + z_2 \\ Tz_1^* &= bz_1 + az_1^* + z_2^* \\ Tz_2 &= az_2 - bz_2^* \\ Tz_2^* &= bz_2 + az_2^* \\ Tz_3 &= cz_3 \end{aligned}$$

Es decir, la representación de T respecto de base Z es:

$$K = \begin{pmatrix} a & b & 0 & 0 & 0 \\ -b & a & 0 & 0 & 0 \\ 1 & 0 & a & b & 0 \\ 0 & 1 & -b & a & 0 \\ 0 & 0 & 0 & 0 & c \end{pmatrix}$$

Además, la matriz Q de cambio de la base $\mathcal{B}$ a la Z es real. Por tanto, $T = QKQ^{-1}$ con Q real, lo que demuestra la proposición inicial.

El resultado general puede enunciarse del modo siguiente.

Toda matriz $A \in M(n, \mathbf{R})$ admite una expresión $Q^{-1}J_1Q$, donde $Q \in M(n, \mathbf{R})$ y J_1 es una matriz de la forma:

$$J_1 = \left(\begin{array}{c|cccc|cc}
D & & 0 & & & & 0 \\ \hline
 & H_1 & & & 0 & & \\
0 & & H_2 & & & & 0 \\
 & & & \ddots & & & \\
 & 0 & & & H_p & & \\ \hline
 & & & & & L_1 & 0 \\
0 & & 0 & & & & \ddots \\
 & & & & & 0 & L_r
\end{array}\right)$$

donde D es una matriz diagonal real, cada H_i es una matriz real y de la forma:

$$J_i = \begin{pmatrix} \alpha_i & & 0 \\ & & \\ 1 & \alpha_i & \\ \ddots & & \ddots \\ 0 & 1 & \alpha_i \end{pmatrix}$$

y cada L_i es tambien real y de la forma:

$$L_i = \begin{pmatrix} \Lambda_i & & 0 \\ & & \\ I_2 & \Lambda_i & \\ \ddots & & \ddots \\ 0 & I_2 & \Lambda_i \end{pmatrix}$$

siendo

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad y \quad \Lambda_i = \begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix}$$

1.2.5. Cálculo de polinomios y series de matrices

Si $A = PJP^{-1}$, entonces $A^2 = PJ^2P^{-1}$, $A^k = PJ^kP^{-1}$ para todo k entero positivo. Así, si $Q(\lambda)$ es un polinomio en λ, se tiene:

$$Q(A) = PQ(J)P^{-1}$$

Asimismo,

$$\exp(A) = P\exp(J)P^{-1},$$

y la $\exp(J)$ se calcula muy sencillamente:

$$\exp(J) = \begin{pmatrix} \exp(J_1) & & 0 \\ & & \\ & \exp(J_2) & \\ & \ddots & \\ 0 & & \exp(J_h) \end{pmatrix}$$

siendo

$$\exp J_i = \exp(\alpha_i) \begin{pmatrix} 1 & & & 0 \\ 1/1! & 1 & & \\ 1/2! & 1/1! & 1 & \\ \vdots & \vdots & \ddots & \ddots \\ 1/(h-1)! & 1/(h-2)! & 1/1! & 1 \end{pmatrix}$$

1.2.6. Ejemplo

Ejemplo 1.2.2. *Calculo de la forma de Jordan de una matriz siguiendo el procedimiento de la demostración del teorema*

$$A = \begin{pmatrix} 10 & 4 & 13 \\ 5 & 3 & 7 \\ -9 & -4 & -12 \end{pmatrix}$$

Solución. *Se tiene* $\det(A - \lambda I) = -\lambda^3 + \lambda^2 + \lambda - 1 = -(\lambda - 1)^2(\lambda + 1)$.
Así, el espectro es $\sigma(T) = \{1, -1\}$ *y sus multiplicidades son 2 para* $\lambda = 1$ *y 1 para* $\lambda = -1$. *Estudiamos el índice*

a) $N(1,1) = \{x : (A-1)x = 0\}$

$$A - 1 = \begin{bmatrix} 9 & 4 & 13 \\ 5 & 2 & 7 \\ -9 & -4 & -13 \end{bmatrix}$$

Ahora, tenemos que encontrar un vector x *tal que* $(A-1)x = 0$

$$\begin{pmatrix} 9 & 4 & 13 \\ 5 & 2 & 7 \\ -9 & -4 & -13 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

notemos que el renglón 3 (que denotaremos R_3*) es el negativo del renglón 1* R_1*, entonces trabajando sólo con los renglones* R_1 *y* R_2 *obtenemos*

$$\left(\begin{array}{ccc|c} 9 & 4 & 13 & 0 \\ 5 & 2 & 7 & 0 \\ -9 & -4 & -13 & 0 \end{array}\right) R_1 \to R_1 - 2R_2 \quad \left(\begin{array}{ccc|c} -1 & 0 & -1 & 0 \\ 5 & 2 & 7 & 0 \\ -9 & -4 & -13 & 0 \end{array}\right) R_2 \to R_2 + 5R_1$$

$$\left(\begin{array}{ccc|c} -1 & 0 & -1 & 0 \\ 0 & 2 & 2 & 0 \\ -9 & -4 & -13 & 0 \end{array}\right) R_2 \to \frac{1}{2}R_2 \quad \left(\begin{array}{ccc|c} -1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ -9 & -4 & -13 & 0 \end{array}\right)$$

resulta que

$$\begin{aligned} -x_1 \quad - x_3 &= 0 \\ x_2 + x_3 &= 0 \end{aligned}$$

esto es, $x_1 = -x_3$ *y* $x_2 = -x_3$ *resulta que el vector* x *es* $(x_1, x_2, x_3) = (-x_3, -x_3, x_3) = x_3(-1, -1, 1)$. *El espacio es*

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} s$$

b) $N(1, 2) = \{x : (A-1)^2 x = 0\}$

$$\begin{pmatrix} 9 & 4 & 13 \\ 5 & 2 & 7 \\ -9 & -4 & -13 \end{pmatrix}^2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 9 & 4 & 13 \\ 5 & 2 & 7 \\ -9 & -4 & -13 \end{pmatrix} \left[\begin{pmatrix} 9 & 4 & 13 \\ 5 & 2 & 7 \\ -9 & -4 & -13 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \right] = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 9 & 4 & 13 \\ 5 & 2 & 7 \\ -9 & -4 & -13 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$$

$$\left(\begin{array}{ccc|c} 9 & 4 & 13 & -1 \\ 5 & 2 & 7 & -1 \\ -9 & -4 & -13 & 1 \end{array}\right) R_1 \to R_1 - 2R_2 \quad \left(\begin{array}{ccc|c} -1 & 0 & -1 & 1 \\ 5 & 2 & 7 & -1 \\ -9 & -4 & -13 & 1 \end{array}\right) R_2 \to R_2 + 5R_1$$

$$\left(\begin{array}{ccc|c} -1 & 0 & -1 & 1 \\ 0 & 2 & 2 & 4 \\ -9 & -4 & -13 & 1 \end{array}\right) R_2 \to \frac{1}{2}R_2 \quad \left(\begin{array}{ccc|c} -1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 2 \\ -9 & -4 & -13 & 1 \end{array}\right)$$

resulta que

$$\begin{aligned} -z_1 \quad - z_3 &= 1 \\ z_2 + z_3 &= 2 \end{aligned}$$

esto es, si $z_3 = 0$ *se tiene* $z_1 = -1$ *y* $z_2 = 2$*, por otro lado si* $z_3 = 2$ *entonces* $z_1 = -3$ *y* $z_2 = 0$. *El espacio es*

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} t + \begin{pmatrix} -3 \\ 0 \\ 2 \end{pmatrix} u$$

de manera análoga se hace para

c) $N(1,3) = \{x : (A-1)^3 x = 0\}$ *coincide con* $N(1,2)$*. Se tiene*

$$\nu(1) = 2$$

Ahora veamos para el autovalor $\lambda = -1$ *tenemos*

d) $N(-1,1) = \{x : (A+1)x = 0\}$
El espacio es

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix} v$$

e) $N(-1,2) = \{x : (A+1)^2 x = 0\}$ *coincide con* $N(-1,1)$*. Se tiene*

$$\nu(-1) = 1$$

Con esto es claro que la forma de Jordan es

$$J = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Para hallar la matriz P *se puede proceder como en la demostración. Tomamos*

$$\tilde{e}_1 = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} \in N(1,2) - N(1,1)$$

y, a continuación,

$$\tilde{e}_2 = (A-1) \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$$

Tomamos también

$$\tilde{e}_3 = \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix} \in N(-1,1)$$

Con esto, la matriz P *es*

$$P = \begin{array}{c} \begin{pmatrix} -1 & -1 & -2 \\ 2 & -1 & -1 \\ 0 & 1 & 2 \end{pmatrix} \\ \tilde{e}_1 \quad \tilde{e}_2 \quad \tilde{e}_3 \end{array}$$

y se tiene, efectivamente

$$A = PJP^{-1}$$

Ejemplo 1.2.3. *De una matriz se sabe que*

$$\sigma(A) = \{1, -1, 6\}$$

y que:

$$\begin{aligned}
\dim N(1,1) &= \dim N(1,2) = 3 \\
\dim N(-1,1) &= 3, \quad \dim N(-1,2) = 4 \quad \dim N(-1,3) = \dim N(-1,4) = 5 \\
\dim N(6,1) &= 2, \quad \dim N(6,2) = \dim N(6,3) = 3
\end{aligned}$$

¿Cuál es la forma de Jordan de A?

Solución. *Se da que $\sigma(A) = \{1, -1, 6\}$ y por teorema 1.2.4 sabemos que los correspondientes indices son*

$$\nu(1) = 1 \quad \nu(-1) = 3 \quad \nu(6) = 2$$

además se sabe que

$$\begin{aligned}
\dim N(1, \nu(1)) &= 3 \\
\dim N(-1, \nu(-1)) &= 5 \\
\dim N(6, \nu(6)) &= 3
\end{aligned}$$

según el teorema 1.2.4 y las consideraciones anteriores a dicho teorema se tiene que

$$X = N(1, \nu(1)) \oplus N(-1, \nu(1-)) \oplus N(6, \nu(6))$$

de donde

$$\dim X = 11$$

y así $A \in M(11, \mathbf{C})$. Ahora, se tiene que una base para A es

$$\{x_1, \ldots, B^{\nu(\lambda_1)-1}x_1, x_2, \ldots, B^{\nu(\lambda_2)-1}x_2, x_3, \ldots, B^{\nu(\lambda_3)-1}x_3\}$$

Para $\lambda_1 = 1$ tenemos que $\nu(\lambda_1) = 1$ entonces tenemos una matriz de 3×3 tal que en la diagonal principal van 1 y en las restantes posiciones van 0.
Para $\lambda_2 = -1$ se tiene $\nu(\lambda_2) = 3$ es decir que la matriz es de 5×5 tal que en las últimas dos filas es diagonal. Así

$$\begin{aligned}
T &= B + \lambda I = B - 1 \\
T(x_2) &= (B-1)x_2 = Bx_2 - x_2 \\
T(Bx_2) &= (B-1)Bx_2 = B^2x_2 - Bx_2 \\
T(B^2x_2) &= (B-1)B^2x_2 = B^3x_2 - B^2x_2
\end{aligned}$$

así se tiene la matriz

$$\left(\begin{array}{ccc|cc} -1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{array}\right)$$

Para $\lambda_3 = 6$ *tenemos que* $\nu(\lambda_3) = 2$ *entonces tenemos una matriz de* 3×3

$$\begin{aligned} T &= B + \lambda I = B - 6 \\ T(x_3) &= (B-6)x_3 = Bx_3 - 6x_3 \\ T(Bx_3) &= (B-6)Bx_3 = B^2x_3 - 6Bx_3 \end{aligned}$$

así se tiene la matriz

$$\left(\begin{array}{cc|c} 6 & 0 & 0 \\ 1 & 6 & 0 \\ \hline 0 & 0 & 6 \end{array}\right)$$

Por tanto la forma canónica de la matriz A *es*

$$J = \left(\begin{array}{c|c|c} \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} & 0 & 0 \\ \hline 0 & \begin{matrix} -1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{matrix} & 0 \\ \hline 0 & 0 & \begin{matrix} 6 & 0 & 0 \\ 1 & 6 & 0 \\ 0 & 0 & 6 \end{matrix} \end{array}\right)$$

1.3. Ecuaciones Lineales

Consideramos la ecuación

$$x'(t) = A(t)x(t) + b(t) \tag{1.3.1}$$

donde $A : [t_0, t_1] = J \to M_n(n, \mathbf{C})$ es una aplicación continua de $[t_0, t_1] = J \subset \mathbf{R}$ al espacio de matrices $n \times n$ de elementos complejos, y $b : [t_0, t_1] \to \mathbf{C}^n$ es una aplicación continua. Se busca una función $x : [t_0, t_1] \to \mathbf{C}^n$ (todos los vectores se escriben como vectores columna) que sea derivable en $[t_0, t_1]$ y su derivada satisfaga la ecuación (1.3.1). Tal función $x(t)$ se denomina ***solución*** de (1.3.1).

La ecuación (1.3.1) se llama ***homogénea*** si $b(t) = 0$, y afín o ***no homogénea*** si $b(t) \neq 0$.
El problema de Cauchy consistirá en obtener una solución $x(t)$ de (1.3.1) tal que

$$x(t_0) = \xi^0 \in \mathbf{C}^n$$

El teorema siguiente es una sencilla aplicación de los teoremas de existencia, unicidad y prolongación.

Teorema 1.3.1. *El problema de valores iniciales*

$$\begin{cases} x'(t) = A(t)x(t) + b(t), \quad t \in J \\ x(t_0) = \xi^0 \end{cases}$$

con la significación de A, b, x, J, ξ^0 dada arriba, tiene una única solución prolongable a todo el intervalo J.

Demostración. Obsérvese que siendo

$$f(t,\xi) = A(t)\xi + b(t), \quad t \in J, \quad \xi \in \mathbf{C}^n,$$

se tiene

$$|f(t,\xi^1) - f(t,\xi^2)| \leqslant (\max_{t \in J} |A(T)|)|\xi^1 - \xi^2|$$

y, así, se puede aplicar el teorema de Picard globalmente. (Aquí, $|\cdot|$ indicará una norma fija en $\mathbf{C}^n$ y la correspondiente norma en $M(n, \mathbf{C}^n)$). □

Definición 1.3.1. *Se dice que una matriz $\Phi(t)$ es una matriz fundamental de la ecuación*

$$x'(t) = A(t)x(t) \tag{1.3.2}$$

si sus n columnas son n soluciones linealmente independientes de dicha ecuación.

Si $\Phi(t)$ es una matriz fundamental de (1.3.2), el conjunto de soluciones de (1.3.2) se podrá representar por

$$\Phi(t)c \tag{1.3.3}$$

siendo $c \in \mathbf{C}^n$ un vector arbitrario. En efecto, la combinación lineal

$$c_1\Phi^1(t) + \cdots + c_n\Phi^n(t)$$

toma la forma (1.3.3) en notación matricial.
Resolver la ecuación (1.3.2) se reduce, pues, a encontrar una matriz fundamental.
El teorema que sigue se refiere primero a la estructura del conjunto de soluciones de la ecuación (1.3.1), y proporciona después la solución del problema de valores iniciales mediante la formula de Lagrange.

Teorema 1.3.2. *Sea la ecuación*

$$x'(t) = A(t)x(t) + b(t) \tag{1.3.4}$$

con $A : [t_0, t_1] = J \to M(n, \mathbf{C})$ continua y $b : [t_0, t_1] \to \mathbf{C}^n$ continua. Entonces:

a) ***Principio de Superposición.*** *Si $b(t) \equiv 0$, el conjunto de todas las soluciones de 1.3.4 es un espacio vectorial Z de dimensión n. Una base de este espacio se denomina un sistema fundamental de soluciones. (La definición de suma de soluciones y producto escalar por solución es natural.)*

b) Sea $b(t) \equiv 0$ y $\{x_1(t), x_2(t), ..., x_n(t)\} = H$ un sistema de soluciones. Entonces H es un sistema fundamental de soluciones si, y sólo si, la matriz $\Phi(t)$ de columnas $x_1(t), x_2(t), ..., x_n(t)$ es ***no singular*** *para algún $t \in J$. Cuando así sucede Φ se denomina* ***matriz fundamental de la solución*** *de la ecuación y se tiene, evidentemente,*

$$\Phi'(t) = A(t)\Phi(t)$$

Recíprocamente, si $X : [t_0, t_1] = J \to M(n, \mathbf{C})$ es derivable y verifica para $t \in [t_0, t_1]$

$$X'(t) = A(t)X(t), \quad detX \neq 0$$

para algún $t \in [t_0, t_1]$, entonces $X(t)$ es matriz fundamental de la ecuación

$$x'(t) = A(t)x(t)$$

c) Si $b(t) \neq 0$ el conjunto T de soluciones de (1.3.4) constituye un espacio afín de dimensión n. Si $x_p(t)$ es solución de la ecuación (1.3.4) con $b(t) \neq 0$ y Z es el conjunto de soluciones de $x'(t) = A(t)x(t)$, se tiene:

$$T = \{x_p(t) + y(t) : y(t) \in Z\}$$

d) ***Fórmula de Lagrange o de variación de las constantes.*** *El problema de valores iniciales*

$$\begin{cases} x'(t) = A(t)x(t) + b(t), & t \in J \\ x(t_0) = \xi^0 \end{cases}$$

tiene por solución única

$$\Phi(t)\xi^0 + \Phi(t)\int_{t_0}^{t} \Phi^{-1}(s)b(s)ds$$

donde $\Phi(t)$ es ***matriz fundamental*** *de $x'(t) = A(t)x(t)$ tal que $\Phi(t_0) = I$. Tal matriz siempre existe y es única.*

Demostración. a) La primera parte es sencilla. Demostraremos que existe una base de n soluciones $x_1, \ldots, x_n$.

Llamemos $e_1, e_2, \ldots, e_n$ a los siguientes vectores

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

y sea x_j la solución única del problema

$$P_j \begin{cases} x'(t) = A(t)x(t), \quad t \in J \\ x(t_0) = e_j \end{cases} \qquad j = 1, 2, \ldots, n$$

Las soluciones x_j forman una base del espacio vectorial Z. Para ver que son linealmente independientes, supongamos

$$\sum_{j=1}^{n} \alpha_j x_j(t) \equiv 0$$

Haciendo $t = t_0$, se tiene

$$\sum_{j=1}^{n} \alpha_j x_j(t_0) = 0$$

$$\sum_{j=1}^{n} \alpha_j e_j = 0$$

$$\alpha_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \alpha_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = 0$$

obtenemos

$$\begin{pmatrix} \alpha_1 \\ \alpha_n \\ \vdots \end{pmatrix} = 0$$

Además forman todo el espacio. En efecto, sea $z(t)$ una solución de $x'(t) = A(t)x(t)$, y sea

$$z(t_0) = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \beta$$

Formemos la solución

$$\bar{z}(t) = \sum_{j=1}^{n} \beta_j x_j(t)$$

Tanto $z(t)$ como $\bar{z}(t)$ son soluciones del problema

$$\begin{cases} x'(t) = A(t)x(t) \\ x(t_0) = \beta \end{cases}$$

así, por la unicidad,

$$z(t) \equiv \bar{z}(t) \equiv \sum_{j=1}^{n} \beta_j x_j(t)$$

b) Supongamos $\det \Phi(s) = 0$ para algún $s \in J$. Entonces se tendrá que las columnas de $\det \Phi(s)$ son linealmente dependientes, es decir:

$$\sum_{j=1}^{n} \alpha_j x_j(s) = 0 \quad \textit{con algún } \alpha_j \neq 0$$

con algún $\alpha_j \neq 0$. Consideramos

$$z(t) = \sum_{j=1}^{n} \alpha_j x_j(t)$$

Se tiene que $z(t)$ es solución del problema

$$\begin{cases} x'(t) = A(t)x(t), & t \in J \\ x(s) = 0 \end{cases}$$

Pero la función idénticamente nula es también solución de tal problema. Así, por la unicidad, $z(t) \equiv 0$. Por tanto,

$$\sum_{j=1}^{n} \alpha_j x_j(t) \equiv 0$$

y, por consiguiente,

$$\det \Phi(t) \equiv 0$$

De este modo, o bien $\det \Phi(t) \equiv 0$ o bien $\det \Phi(t) \neq 0$ para todo $t \in J$.

d) Cualquier solución de $x'(t) = A(t)x(t)$ es de la forma $\Phi(t)c$ con $c \in \mathbf{C}^n$ constante si $\Phi(t)$ es una matriz fundamental. Si se quiere hallar una solución de $x'(t) = A(t)x(t)$ se puede intentar variar c (variación de las constantes) y ensayar $z(t) = \Phi(t)c(t)$ (es decir, se supone que $z(t)$ es solución de la ecuación (1.3.4)). Se ha de tener al derivar y reemplazar $z(t)$ en (1.3.4):

$$z'(t) = A(t)z(t) + b(t) = \Phi'(t)c(t) + \Phi(t)c'(t)$$

Sabemos que $\Phi'(t) = A(t)\Phi(t)$, y, así, ha de resultar:

$$A(t)\Phi(t)c(t) + b(t) = A(t)\Phi(t)c(t) + \Phi(t)c'(t)$$

por lo que hay que escoger

$$c'(t) = \Phi^{-1}(t)b(t)$$

Si se quiere $z(t_0) = \xi^0$, entonces se ha de verificar simultáneamente

$$\begin{cases} c'(t) = \Phi^{-1}(t)b(t) \\ c(t_0) = \Phi^{-1}(t_0)\xi^0 \end{cases}$$

Si $\Phi(t)$ se ha escogido tal que $\Phi(t_0) = I$, entonces

$$c(t) = \xi^0 + \int_{t_0}^{t} \Phi^{-1}(s)b(s)ds$$

y, así

$$z(t) = \Phi(t)\xi^0 + \Phi(t)\int_{t_0}^{t} \Phi^{-1}(s)b(s)ds$$

□

Obsérvese que, con la notación de la demostración de a), la matriz $\Phi(t)$ tal que $\Phi(t_0)$ es la que tiene por columnas las soluciones de

$$P_j \begin{cases} x'(t) = A(t)x(t) \\ x(t_0) = e_j \end{cases} \qquad j = 1, 2, \ldots, n$$

El resultado siguiente se suele denominar *fórmula de Jacobi.*

Teorema 1.3.3 (Fórmula de Jacobi)**.** *Sea $\Phi(t)$ una matriz $n \times n$ cuyas columnas son todas solución de $x'(t) = A(t)x(t)$. Entonces se verifica:*

$$(\det \Phi)'(t) = trA(t) \det \Phi(t)$$

y, así:

$$\det \Phi(t) = \det \Phi(t_0) \exp \int_{t_0}^{t} trA(s)ds$$

Demostración. Sea

$$\Phi(t) = \begin{pmatrix} \psi_{11}(t) & \psi_{12}(t) \ldots \psi_{1n}(t) \\ \psi_{21}(t) & \psi_{22}(t) \ldots \psi_{2n}(t) \\ \ldots\ldots & \ldots\ldots \\ \psi_{n1}(t) & \psi_{n2}(t) \ldots \psi_{nn}(t) \end{pmatrix}$$

$$(\det \Phi)'(t) = \det \begin{pmatrix} \psi'_{11}(t) & \psi'_{12}(t) \ldots \psi'_{1n}(t) \\ \psi_{21}(t) & \psi_{22}(t) \ldots \psi_{2n}(t) \\ \ldots\ldots & \ldots\ldots \\ \psi_{n1}(t) & \psi_{n2}(t) \ldots \psi_{nn}(t) \end{pmatrix} + \ldots + \det \begin{pmatrix} \psi_{11}(t) & \psi_{12}(t) \ldots \psi_{1n}(t) \\ \psi_{21}(t) & \psi_{22}(t) \ldots \psi_{2n}(t) \\ \ldots\ldots & \ldots\ldots \\ \psi'_{n1}(t) & \psi'_{n2}(t) \ldots \psi'_{nn}(t) \end{pmatrix}$$

Como se tiene $\Phi'(t) = A(t)\Phi(t)$ el primer determinante del segundo miembro es

$$\det \begin{pmatrix} a_{11}\psi_{11} + a_{12}\psi_{21} + \ldots + a_{1n}\psi_{n1} & \ldots & a_{11}\psi_{1n} + a_{12}\psi_{2n} + \ldots + a_{1n}\psi_{nn} \\ \psi_{21} & \ldots & \psi_{2n} \\ \ldots\ldots & \ldots\ldots & \ldots\ldots \\ \psi_{n1} & \ldots & \psi_{nn} \end{pmatrix} = a_{11} \det \Phi(t)$$

con lo cual:

$$(\det \Phi)'(t) = trA(t) \det \Phi(t)$$

□

El estudio de una ecuación lineal de orden n del tipo:

$$x^{(n)}(t) + a_1(t)x^{(n-1)}(t) + \cdots + a_{n-1}(t)x'(t) + a_n(t)x(t) = f(t)$$

donde a_j son funciones escalares continuas de $[t_0, t_1]$ a $\mathbf{C}$, así como f, y la x que se busca es una función con n derivadas continuas en $[t_0, t_1]$, y con los valores en $\mathbf{C}$ se reduce al estudio realizado ya de la ecuación lineal vectorial mediante el cambio

$$\begin{pmatrix} x(t) \\ x'(t) \\ \vdots \\ x^{(n-1)}(t) \end{pmatrix} = \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix} = y(t)$$

Resulta así la ecuación

$$y'(t) = A(t)y(t) + f(t)$$

donde

$$\begin{pmatrix} 0 & 1 & & \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 \;\; 1 & \\ \vdots & \vdots & \ddots \;\ddots & 1 \\ -a_n(t) & -a_{n-1}(t) & & -a_1(t) \end{pmatrix}$$

Para la ecuación diferencial lineal de orden n se puede, por tanto, formular un teorema de existencia y de estructura paralelo a los teoremas 1.3.1 y 1.3.2. El teorema 1.3.3 da ahora un resultado especialmente sencillo, conocido como ***fórmula de Liouville***. Obsérvese que si $x_1, x_2, \ldots, x_n$ son soluciones de la ecuación

$$x^{(n)}(t) + a_1(t)x^{(n-1)}(t) + \cdots + a_n(t)x(t) = 0$$

entonces la matriz correspondiente a $\Phi(t)$ del teorema 1.3.3 es

$$\Phi(t) = \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ x_1' & x_2' & \ldots & x_n' \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \end{pmatrix}$$

y, así, su determinante es el wronskiano de las funciones $x_1, x_2, \ldots, x_n$. De este modo, si señalamos por

$$W(t) = W(x_1(t), \ldots, x_n(t)) = \det \Phi(t),$$

se tiene

$$W'(t) = -a_1(t)W(t)$$

y, por tanto

$$W(t) = W(t_0) \exp \int_{t_0}^{t} (-a_1(s))ds$$

que es la ***fórmula de Liouville***.

1.3.1. Ejemplos

Ejemplo 1.3.1. *Resuélvase el problema*

$$\begin{cases} x'''(t) + 4x''(t) - 5x(t) = 0 \\ x(0) = 1,\ x'(0) = 0,\ x''(0) = 0 \end{cases}$$

y determínese cómo se comporta la solución en el infinito, es decir, si es acotada o no, si tiende a cero para $t \to \infty$ o no, etc.

Solución. *La ecuación característica de la ecuación diferencial dada es*

$$\lambda^3 + 4\lambda^2 - 5 = 0$$

Encontramos las raíces de la ecuación anterior

$$\begin{aligned} \lambda^3 + 4\lambda^2 - 5 &= 0 \\ (\lambda - 1)(\lambda^2 + 5\lambda + 5) &= 0 \text{ resolviendo la cuadratica resulta} \\ (\lambda - 1)\left(\lambda - \left(\frac{-5 + \sqrt{5}}{2}\right)\right)\left(\lambda - \left(\frac{-5 - \sqrt{5}}{2}\right)\right) &= 0 \end{aligned}$$

de donde las raíces de la ecuación característica son

$$\lambda_1 = 1, \quad \lambda_2 = \frac{-5 + \sqrt{5}}{2}, \quad \lambda_3 = \frac{-5 - \sqrt{5}}{2}$$

Dado que las raíces son reales y distintas se tiene que la solución general de la ecuación diferencial dada es

$$x(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t} + c_3 e^{\lambda_3 t}$$

La derivación nos lleva a

$$\begin{aligned} x'(t) &= \lambda_1 c_1 e^{\lambda_1 t} + \lambda_2 c_2 e^{\lambda_2 t} + c_3 \lambda_3 e^{\lambda_3 t} \\ x''(t) &= \lambda_1^2 c_1 e^{\lambda_1 t} + \lambda_2^2 c_2 e^{\lambda_2 t} + \lambda_3^2 c_3 e^{\lambda_3 t} \end{aligned}$$

así, de las condiciones iniciales se obtienen las ecuaciones

$$\begin{aligned} x(0) &= c_1 + c_2 + c_3 = 1 \\ x'(0) &= \lambda_1 c_1 + \lambda_2 c_2 + \lambda_3 c_3 = 0 \\ x''(0) &= \lambda_1^2 c_1 + \lambda_2^2 c_2 + \lambda_3^2 c_3 = 0 \end{aligned}$$

para $\lambda_1 = 1$ el sistema anterior se convierte en

$$x(0) = c_1 + c_2 + c_3 = 1 \tag{1.3.5}$$

$$x'(0) = c_1 + \lambda_2 c_2 + \lambda_3 c_3 = 0 \tag{1.3.6}$$

$$x''(0) = c_1 + \lambda_2^2 c_2 + \lambda_3^2 c_3 = 0 \tag{1.3.7}$$

Restando 1.3.6 de 1.3.7 se tiene

$$\left[\left(\frac{-5+\sqrt{5}}{2}\right)^2-\left(\frac{-5+\sqrt{5}}{2}\right)\right]c_2+\left[\left(\frac{-5-\sqrt{5}}{2}\right)^2-\left(\frac{-5-\sqrt{5}}{2}\right)\right]c_3=0$$

resolviendo lo anterior se obtiene

$$c_3=\frac{-(10-3\sqrt{5})}{10-3\sqrt{5}}c_2$$

sustituyendo c_3 en 1.3.5

$$c_1+\frac{6\sqrt{5}}{10+3\sqrt{5}}c_2=1$$

y sustituyendo c_3, λ_2 y λ_3 en 1.3.6

$$c_1-\frac{5\sqrt{5}}{10+3\sqrt{5}}c_2=0$$

Simultaneando ambas ecuaciones se obtiene

$$c_1=\frac{5}{11}\quad c_2=\frac{3+2\sqrt{5}}{11}$$

Sustituyendo en 1.3.5 c_1 y c_2

$$c_3=\frac{3-2\sqrt{5}}{11}$$

así, se obtiene que la solución del problema inicial planteado es

$$x(t)=\frac{5}{11}e^{\lambda_1 t}+\frac{3+2\sqrt{5}}{11}e^{\lambda_2 t}+\frac{3-2\sqrt{5}}{11}e^{\lambda_3 t}$$

Es obvio que la solución no es acotada en $t\geqslant 0$, puesto que si bien los dos últimos sumandos tienden a cero cuando t tiende a infinito, el primero no es acotado.

1.4. Coeficientes Constantes

Definición 1.4.1. *Una ecuación diferencial $x'(t)=f(t,x(t))$ se denomina* ***autónoma*** *cuando la función $f(t,\xi)$ es solamente la función de ξ, es decir, es una ecuación de la forma $x'(t)=f(x(t))$. Si la ecuación es lineal y autónoma, será $x'(t)=Ax(t)+b$, con A,b constantes.*

Consideremos primero el problema

$$(P)\begin{cases} x'(t)=Ax(t) \quad en \quad [t_0,t_1]=J \\ x(t_0)=\xi^0 \end{cases}$$

El problema anterior posee una única solución que es la función definida por

$$x(t)=e^{A(t-t_0)}\xi^0$$

Veamos que en efecto es así, para ello nos auxiliamos del método de aproximaciones sucesivas (o iterantes de Picard). Sea $f(t,x(t))=Ax(t)$, por tanto, la ecuación integral equivalente a (P) es

$$x(t)=\xi^0+\int_{t_0}^{t}Ax(s)ds$$

La iterante incial asociada a (P) es la función constante definida por $x_0(t) = \xi^0$ las demás iterantes están definidas y vienen dadas así:

$$\begin{aligned}
x_1(t) &= \xi^0 + \int_{t_0}^{t} x_0(s)ds = \xi^0 + \int_{t_0}^{t} A\xi^0 ds = \xi^0 + A\xi^0 s|_{t_0}^{t} = \xi^0 + A\xi^0(t-t_0)\\
x_2(t) &= \xi^0 + \int_{t_0}^{t} x_1(s)ds = \xi^0 + \int_{t_0}^{t} A[\xi^0 + A\xi^0 s]ds = \xi^0 + \left[A\xi^0 s + A^2\xi^0\frac{s^2}{2}\right]_{t_0}^{t}\\
&= \xi^0 + A\xi^0(t-t_0) + A^2\xi^0\frac{(t-t_0)^2}{2!}\\
x_3(t) &= \xi^0 + \int_{t_0}^{t} x_2(s)ds = \xi^0 + \int_{t_0}^{t} A[\xi^0 + A\xi^0 s + A^2\xi^0\frac{s^2}{2}]ds\\
&= \xi^0 + \left[A\xi^0 s + A^2\xi^0\frac{s^2}{2} + A^3\xi^0\frac{s^3}{6}\right]_{t_0}^{t}\\
&= \xi^0 + A\xi^0(t-t_0) + A^2\xi^0\frac{(t-t_0)^2}{2!} + A^3\xi^0\frac{(t-t_0)^3}{3!}\\
&\vdots\\
x_n(t) &= \xi^0 + \int_{t_0}^{t} x_{n-1}(s)ds = \xi^0 + \int_{t_0}^{t} A\left[\xi^0 + A\xi^0 s + A^2\xi^0\frac{s^2}{2!} + \ldots + A^{n-1}\xi^0\frac{s^{n-1}}{(n-1)!}\right]ds\\
&= \xi^0 + \left[A\xi^0 s + A^2\xi^0\frac{s^2}{2!} + A^3\xi^0\frac{s^3}{3!} + \ldots + A^n\xi^0\frac{s^n}{n!}\right]_{t_0}^{t}\\
&= \xi^0 + A\xi^0(t-t_0) + A^2\xi^0\frac{(t-t_0)^2}{2!} + A^3\xi^0\frac{(t-t_0)^3}{3!} + \ldots + A^n\xi^0\frac{(t-t_0)^n}{n!}\\
&= \xi^0\left(I + A(t-t_0) + A^2\frac{(t-t_0)^2}{2!} + A^3\frac{(t-t_0)^3}{3!} + \ldots + A^n\frac{(t-t_0)^n}{n!}\right)
\end{aligned}$$

Se tiene entonces que la sucesión de funciones x_n así construida coincide con la sucesión de las sumas parciales de la serie $e^{A(t-t_0)}\xi^0 = \xi^0\sum_{k=0}^{\infty}\frac{A^k(t-t_0)^k}{k!}$ (definición de matriz exponencial $e^{At} = \sum_{k=0}^{\infty}\frac{(At)^k}{k!}$) para $t \in \mathbf{R}$

$$\lim_{n\to\infty} x_n(t) = \lim_{n\to\infty}\sum_{k=0}^{n}\frac{A^k\xi^0(t-t_0)^k}{k!} \doteq \xi^0\sum_{k=0}^{\infty}\frac{A^k(t-t_0)^k}{k!} = e^{A(t-t_0)}\xi^0$$

Por consiguiente, la sucesión de iterantes x_n converge puntualmente en $\mathbf{R}$ hacia la solución única $e^{A(t-t_0)}\xi^0$, del problema (P).

Ahora, veamos que la solución $\Phi(t) = e^{A(t-t_0)}$ del problema (P) es una matriz fundamental.

Tenemos que la matriz A es constante y escribamos

$$e^{A(t-t_0)} = \sum_{k=0}^{\infty}\frac{A^k(t-t_0)^k}{k!}$$

Teniendo en cuenta que

i) $\dfrac{\|A^k(t-t_0)^k\|}{k!} \leqslant \dfrac{\|A\|^k|(t-t_0)^k|}{k!}$

ii) $\sum_{k=0}^{\infty}\dfrac{\|A\|^k|(t-t_0)^k|}{k!} < \infty$, se sigue que $e^{A(t-t_0)}$ está bien definida. Más aún como $\sum_{n=0}^{\infty} \dfrac{A^n(t-t_0)^n}{n!}$ converge uniformemente sobre compactos, cada término de la serie es infinitamente diferenciable y la serie de las derivadas término a término también es uniformemente convergente, obtenemos que $\Phi(t) = e^{A(t-t_0)}$ satisface la ecuación matricial lineal homogénea

$$\Phi'(t) = \sum_{n=1}^{\infty} n \frac{A^n(t-t_0)^{n-1}}{n!} = A \sum_{n=1}^{\infty} \frac{A^{n-1}(t-t_0)^{n-1}}{(n-1)!} = A\Phi(t)$$

tal que $\Phi(t_0) = I$. Así, hemos mostrado que para sistemas lineales con coeficientes constantes la matriz fundamental principal viene dada por $\Phi(t) = e^{A(t-t_0)}$ con $\Phi(t_0) = I$. Y, así, se quiere resolver ahora el problema

$$\begin{cases} x'(t) = Ax(t) + b(t) \\ x(t_0) = \xi^0 \end{cases}$$

Aplicando la fórmula de Lagrange que esta en el teorema 1.3.2 d), se tiene que

$$\begin{aligned}
x(t) &= \Phi(t)\xi^0 + \Phi(t)\int_{t_0}^{t} \Phi^{-1}(s)b(s)ds \\
&= e^{A(t-t_0)}\xi^0 + e^{A(t-t_0)}\int_{t_0}^{t} [e^{A(s-t_0)}]^{-1}b(s)ds \\
&= e^{A(t-t_0)}\xi^0 + \int_{t_0}^{t} e^{A(t-t_0)}e^{-A(s-t_0)}b(s)ds \\
&= e^{A(t-t_0)}\xi^0 + \int_{t_0}^{t} e^{A(t-t_0-s+t_0)}b(s)ds \\
x(t) &= e^{A(t-t_0)}\xi_0 + \int_{t_0}^{t} e^{A(t-s)}b(s)ds
\end{aligned}$$

La expresión explícita de e^{At} es sencilla en términos de la forma canónica de Jordan.
Sea primero $L = \lambda I + H$, con

$$H = \begin{pmatrix} 0 & & & & 0 \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & \ddots & \\ 0 & & & 1 & 0 \end{pmatrix}$$

Entonces, puesto que $\lambda I + H$ conmutan, se tiene

$$H^2 = \begin{pmatrix} 0 & & & 0 \\ 0 & 0 & & \\ 1 & 0 & 0 & \\ \ddots & \ddots & \ddots & \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad H^3 = \begin{pmatrix} 0 & & & & 0 \\ 0 & 0 & & & \\ 0 & 0 & 0 & & \\ 1 & 0 & 0 & 0 & \\ \ddots & \ddots & \ddots & \ddots & \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix} \cdots H^{k-1} = \begin{pmatrix} 0 & & & 0 \\ 0 & 0 & & \\ 0 & 0 & 0 & \\ \vdots & \vdots & \vdots\ddots & \\ 1 & 0 & 0 \ldots 0 \end{pmatrix}$$

$H^k = 0$, resulta por lo estudiado en la subsección 1.1.4 que:

$$e^{Lt} = e^{(H+\lambda I)t} = e^{Ht}e^{\lambda t} = \sum_{k=0}^{\infty} \frac{(Ht)^k}{k!} e^{\lambda t}$$

en notación matricial, esto es:

$$e^{Lt} = \begin{pmatrix} 1 & & & & \\ \frac{t}{1!} & 1 & & & \\ \frac{t^2}{2!} & \frac{t}{1!} & 1 & & \\ \cdots & \ddots & & \ddots & \\ \frac{t^{k-1}}{(k-1)!} & \frac{t^{k-2}}{(k-2)!} \cdots & \frac{t}{1!} & & 1 \end{pmatrix} e^{\lambda t}$$

Así, si $A = PJP^{-1}$, entonces

$$e^{At} = Pe^{Jt}P^{-1}$$

Si J es la forma de Jordan de A:

$$J = \begin{pmatrix} L_1 & & 0 \\ & L_2 & \\ & & \ddots & \\ 0 & & & L_h \end{pmatrix}$$

donde cada $L_j = \lambda_j I + H_j$ es de la forma anterior, se tiene

$$e^{At} = Pe^{Jt}P^{-1} = P \begin{pmatrix} e^{L_1 t} & & 0 \\ & e^{L_2 t} & \\ & & \ddots & \\ 0 & & & e^{L_h t} \end{pmatrix} P^{-1}$$

donde cada $e^{L_j t}$ tiene la forma expresada como antes.

1.5. Ecuaciones periódicas

Sea la ecuación $x'(t) = A(t)x(t)$, $t \in \mathbf{R}$, donde $A(t+\tau) = A(t)$ para todo t, siendo τ un número real fijo. Si $x(t)$ satisface la ecuación y escribimos $y(t) = x(t+\tau)$ entonces se tiene:

$$y'(t) = x'(t+\tau) = A(t+\tau)x(t+\tau) = A(t)y(t)$$

y, así, también $y(t)$ es solución. Por lo tanto, $\Phi(t)$ es una matriz fundamental, también lo es $\Phi(t+\tau)$ y, por consiguiente, existe una matriz no singular constante C tal que

$$\Phi(t+\tau) = \Phi(t)C$$

haciendo $t = 0$, en la ecuación anterior y suponiendo que $\Phi(0) = I$, se tiene que

$$\Phi(\tau) = \Phi(0)C = C$$

A la matriz C se le llama *matriz de monodromía.*

Lema 1.5.1. *Sea C una matriz no singular, entonces existe una matriz L tal que $C = e^{\tau L}$, siendo L en general de elementos de* $\mathbf{C}$.

Entonces, si denominamos

$$B(t) = \Phi(t)e^{-tL}$$

se puede poner:

$$\begin{aligned} B(t+\tau) &= \Phi(t+\tau)e^{-\tau L}e^{-tL} = \Phi(t+\tau)C^{-1}e^{-tL} \\ &= \Phi(t)e^{-tL} = B(t) \end{aligned}$$

Así resulta el siguiente teorema.

Teorema 1.5.1 (Floquet 1883)**.** *Toda matriz fundamental $\Phi(t)$ para el sistema periódico $x'(t) = A(t)x(t)$ puede representarse en la forma*

$$\Phi(t) = B(t)e^{tL}$$

donde L es constante y $B(t)$ es periódica de periodo τ

Demostración de lema. Si C fuese un número distinto de 0, cualquier determinación de $\log C$ serviría como L. Es decir, se podría poner

$$L = \log C = \log(1 + (C-1)),$$

y, si $|C-1| < 1$, se tiene llamando $N = C - 1$:

$$L = N - \frac{N^2}{2} + \frac{N^3}{3} - \frac{N^4}{4} + \ldots$$

La expresión de la derecha puede tener sentido, según sabemos, para una matriz y, así tal vez podamos esperar que si ponemos $N = C - 1$ en el caso en que C sea una matriz, no singular, y obtenemos

$$L = N - \frac{N^2}{2} + \frac{N^3}{3} - \frac{N^4}{4} + \ldots$$

entonces sea $e^L = C$.

La presentación rigurosa de este razonamiento heurístico se puede realizar del modo siguiente: Sea en primer lugar N una matriz $n \times n$ de la forma

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

con $\alpha \in \mathbf{C}$. Demostraremos que, si

$$R(N) = N - \frac{N^2}{2} + \frac{N^3}{3} - \frac{N^4}{4} + \dots \quad y \quad T(U) = I + \frac{U}{1!} + \frac{U^2}{2!} + \frac{U^3}{3!} + \dots$$

entonces

$$T(R(N)) = I + N$$

(obsérvese que siendo $N^n = 0$, $R(N)$ es una matriz bien definida). Para ello procedemos por inducción sobre la dimensión de N. Si $n = 1$ nada hay que demostrar. Supongamos que la propiedad es cierta para toda matriz del tipo de la N de dimensión $n-1$. Entonces, escribiendo

$$A = \left(\begin{array}{c|cc} 0 & 0 & 0\cdots 0 \\ \hline \alpha & 0 & 0\cdots 0 \\ 0 & 0 & 0 \ \vdots \\ \vdots & \vdots & \ddots \vdots \\ 0 & 0 & \cdots 0 \end{array}\right) \qquad N^* = \left(\begin{array}{c|cc} 0 & 0 & 0\cdots 0 \\ \hline 0 & 0 & 0\cdots 0 \\ 0 & \alpha & 0\cdots 0 \\ \vdots & \vdots & \ddots \\ 0 & 0 & \alpha\cdots 0 \end{array}\right)$$

se puede poner, ya que $AN^* = N^*A = 0$, $A^2 = 0$:

$$R(N) = A + R(N^*)$$

y, por tanto, como $AR(N^*) = R(N^*) = 0$,

$$T(R(N)) = T(A)T(R(N^*)),$$

pero

$$T(A) = I + A \quad y \quad T(R(N^*)) = I + N^*$$

por la hipótesis inductiva. Así:

$$T(R(N)) = (I + A)(I + N^*) = I + N^* + A = I + N$$

Supongamos ahora que C es un bloque de Jordan, es decir, $C = \lambda I + H$ con $\lambda \neq 0$ y

$$H = \begin{pmatrix} 0 & & & & 0 \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & \ddots & \\ 0 & & & 1 & 0 \end{pmatrix}$$

Entonces se puede poner

$$\widetilde{C} = \frac{C}{\lambda} = I + \frac{H}{\lambda} = I + N$$

Si ponemos $\widetilde{L} = R(N)$, entonces $\exp \widetilde{L} = \widetilde{C}$ y, así,

$$C = \lambda \widetilde{C} = \lambda \exp \widetilde{L} = \exp((\log \lambda) I + \widetilde{L}) = \exp L$$

llamando $L = (\log \lambda) I + \widetilde{L}$, pudiéndose tomar en $\log \lambda$ cualquiera de las determinaciones del logaritmo de λ.

Sea ahora C una matriz arbitraria no singular. Entonces todos sus autovalores son distintos de cero y su forma canónica de Jordan J tiene bloques análogos a la C tratada antes. Es decir:

$$J = \begin{pmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_h \end{pmatrix}$$

y para cada J_j sabemos hallar L_j tal que $e^{L_j} = J_j$. Ponemos entonces:

$$M = \begin{pmatrix} L_1 & & & 0 \\ & L_2 & & \\ & & \ddots & \\ 0 & & & L_h \end{pmatrix}$$

y tendremos $e^M = J$. Ahora, si $C = P^{-1} J P$, ponemos $L = P^{-1} M P$ y tendremos $e^L = C$. □

Capítulo 2

Teoría de estabilidad. Método de la primera aproximación

El problema sobre estabilidad de soluciones de una ecuación diferencial ordinaria surge como continuación sobre la dependencia de las soluciones respecto de los valores iniciales y parámetros. Un proceso se desarrolla gobernado por una ecuación diferencial $x'(t) = f(t, x(t))$. Bajo condiciones x^0 en instante t_0 su funcionamiento viene dado por $x(t)$, pero por errores de medición o por una perturbación imprevista es posible que el valor real en el instante t_0 sea $\tilde{x}^0$ en lugar de ser x^0 y, por tanto, su funcionamiento real será dado por $\tilde{x}^0(t)$. ¿Cómo se comporta $\tilde{x}(t)$ con respecto a $x(t)$?
Cuando se considera un intervalo finito de tiempo proporciona bastante información suficiente. Pero en muchas ocasiones interesa especialmente lo que va a ocurrir cuando t tiende a ∞. Por ejemplo, y este es uno de los problemas clásicos que han originado la teoría, el sistema solar ¿es estable o llegará a una situación crítica, provocada por perturbaciones pequeñas, que produzca su desmembramiento?
El estudio de este tipo de problemas tiene su origen en LAGRANGE, DIRICHLET y sobre todo, más recientemente, en los trabajos de A.M LYAPUNOV Y POINCARÉ, formando lo que se denomina la teoría global o geométrica de ecuaciones diferenciales ordinarias. Constituye una rama muy amplia de la teoría de ecuaciones diferenciales. En este capítulo se exponen las nociones más importantes de estabilidad y algunos teoremas típicos de la teoría que se obtienen al estudiar ***las soluciones*** de la ecuación bajo estudio, siguiendo más o menos lo que se denomina el método de la primera aproximación. En el capítulo siguiente se expone el segundo método de Lyapunov o método directo que se basa en el estudio de la función $f(t, \xi)$ que aparece en la ecuación diferencial misma.

2.1. Noción de estabilidad

2.1.1. Definiciones

Consideremos la ecuación $x'(t) = f(t, x(t))$, donde f está definida en $[t_0, +\infty) \times \mathbf{R}^n$ en general, y satisface condiciones que permiten garantizar que existe solución única, al menos local, del problema

$$x'(t) = f(t, x(t)) \quad x(t_0) = x^0$$

Esta ***solución*** será denominada $x(t; t_0, x^0)$. Supongamos que $x(t; t_0, x^0)$ está definida en $[t_0, +\infty)$.

Definición 2.1.1. *Diremos que $x(t; t_0, x^0)$ es **estable** (a la derecha, en el sentido de Lyapunov) si para cada $\epsilon > 0$ existe $\delta = \delta(\epsilon, t_0) > 0$ tal que para todo x^1 con $|x^1 - x^0| \leqslant \delta$ se tiene que $x(t; t_0, x^1)$ existe y está definida en $[t_0, +\infty)$ y se verifica*

$$|x(t; t_0, x^1) - x(t; t_0, x^0)| \leqslant \epsilon \quad \textit{para} \quad \textit{todo} \quad t \geqslant t_0$$

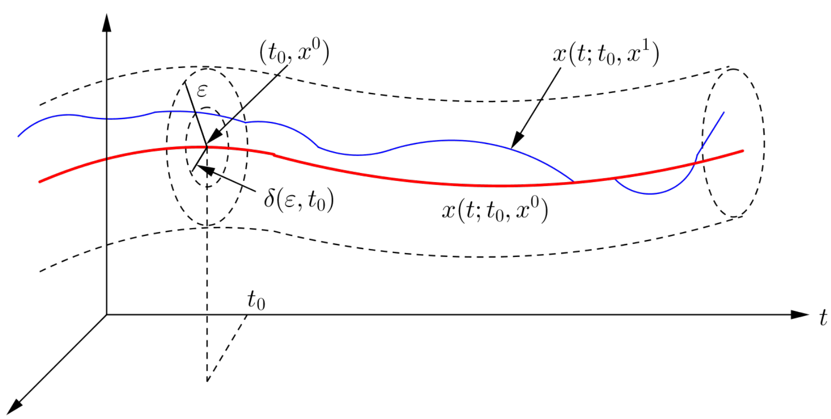


Figura 2.1: Estabilidad

Definición 2.1.2. *Se dice que $x(t; t_0, x^0)$ es **asintóticamente estable** (a la derecha, en el sentido de Lyapunov) si es **estable** y además existe $\eta > 0$ tal que si $|x^1 - x^0| \leqslant \eta$, entonces*

$$|x(t; t_0, x^1) - x(t; t_0, x^0)| \to 0 \quad \textit{para} \quad t \to +\infty$$

Definiciones análogas se pueden dar de la estabilidad (asintótica) a la izquierda, pero, en general, cuando nos refiramos a la estabilidad de una solución, sobrentenderemos que se trata de estabilidad a la derecha.

Normalmente, interesa estudiar la estabilidad a la derecha como parte del análisis del comportamiento a largo plazo (o sea, en el sentido de los tiempos crecientes) de las soluciones de un sistema y, por ello, el término "estabilidad "se referirá sistemáticamente a la estabilidad a la derecha (en sentido de Lyapunov). Como iremos viendo, en la práctica interesa estudiar la estabilidad de ciertas soluciones

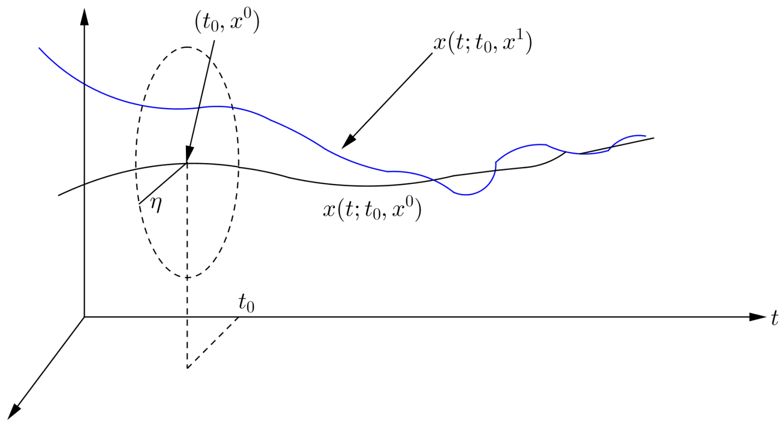


Figura 2.2: Estabilidad asintótica

que representen comportamientos significativos del sistema físico modelado por la ecuación diferencial, particularmente soluciones constantes (o estacionarias) y soluciones periódicas.
Cuando se tiene una ecuación de orden superior en forma normal:

$$x^{(n)}(t) = f(t, x(t), ..., x^{(n-1)})$$

la estabilidad (asintótica) de sus soluciones se define mediante la estabilidad (asintótica) de las soluciones correspondientes de la ecuación vectorial asociada.

2.1.2. Estabilidad de la solución trivial

El estudio de la estabilidad de las ecuaciones lineales admite una reducción importante.
Sea $x'(t) = A(t)x(t) + b(t)$ con $A(t)$ continua de $[t_0, +\infty)$ a $M_n(\mathbf{C})$ o $M_n(\mathbf{R})$ (matrices complejas o reales $n \times n$). La solución $x(t; t_0, x^0)$ existe y es única y está definida en todo $[t_0, +\infty)$ para cualquier x^0. Mediante la fórmula de Lagrange:

$$x(t; t_0, x^0) = \Phi(t)x^0 + \int_{t_0}^{t} \Phi(t)\Phi^{-1}(s)b(s)ds$$

donde $\Phi(t)$ es la matriz fundamental tal que $\Phi(t_0) = I$, resulta el siguiente teorema.

Teorema 2.1.1. *La solución $x(t; t_0, x^0)$ de*

$$x'(t) = A(t)x(t) + b(t) \tag{2.1.1}$$

es estable (asintóticamente estable) si, y sólo si, la solución $\tau(t) \equiv 0$ de $x'(t) = A(t)x(t)$ es estable (asintóticamente estable).

Demostración. Supongamos que el sistema (2.1.1) es estable. La estabilidad asintótica se prueba de manera análoga.

Si $x(t; t_0, x^0)$ es una solución fija y sea $x(t)$ cualquier otra solución de (2.1.1) y definimos $y(t) = x(t) - x(t; t_0, x^0)$ la cual es solución de $x'(t) = A(t)x(t)$. Como $x(t; t_0, x^0)$ es estable, se tiene que, dado $\epsilon > 0$, existe $\delta > 0$ tal que si $|x(t_0) - x^0| < \delta$ entonces $|x(t) - x(t; t_0, x^0)| < \epsilon$, $\forall t \geqslant t_0$ lo cual implica que:
$|y(t)| = |\tau(t) - y(t)| = |0 - y(t)| < \epsilon$, $\forall t \geqslant t_0$, si $|y(t_0)| = |0 - y(t_0)| < \delta$, es decir que la solución trivial de $x'(t) = A(t)x(t)$ es estable.
Ahora supongamos que la solución trivial es estable: Dado $\epsilon > 0$, existe $\delta > 0$ tal que si $|0 - y(t_0)| < \delta$ entonces $|0 - y(t)| < \epsilon$, $\forall t \geqslant t_0$, es decir $|x(t_0) - x^0| < \delta$, entonces $|x(t) - x(t; t_0, x^0)| < \epsilon$, $\forall t \geqslant t_0$ por lo cual se deduce que $x(t; t_0, x^0)$ es estable. □

Así pues, y en resumen, el estudio de la estabilidad de una solución cualquiera de un sistema lineal se reduce al estudio de la estabilidad de la solución $\tau(t) \equiv 0$ del sistema homogéneo asociado. El teorema anterior muestra que para la ecuación lineal $x'(t) = A(t)x(t) + b(t)$ todas sus soluciones tienen el mismo carácter de *estabilidad o inestabilidad*. Así, para una ecuación lineal tiene sentido hablar de *estabilidad (asintótica)* de la *ecuación* misma.

2.1.3. Ejemplos

Ejemplo 2.1.1. *La solución $y(t) \equiv 0$ de $x'(t) = (x(t))^2$ no es estable ni a la derecha ni a la izquierda.*

Solución. *Sabemos que las soluciones de la ecuación diferencial $x'(t) = (x(t))^2$ son las ramas de la hipérbola*

$$x(t) = \frac{1}{k - t} \quad k \text{ constante}$$

donde las soluciones positivas están dadas por $(-\infty, k)$ y las soluciones negativas por $(k, +\infty)$.

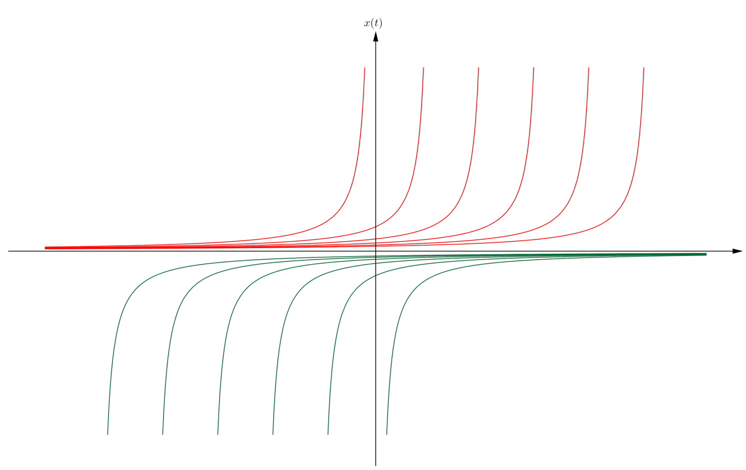


Figura 2.3: Curvas gráficas de las soluciones de la ecuación $x'(t) = (x(t))^2$ (Ejemplo 2.1.1)

De la gráfica podemos observar que si tomamos cualquier solución que este arriba (rojo) de $y(t) \equiv 0$ no se mantiene cerca de $y(t) \equiv 0$, es decir si nos tomamos un entorno bien pequeño que contenga a *estas soluciones para un tiempo dado, dichas soluciones no se mantendrán cercanas para un tiempo posterior. Por otro lado las soluciones que están por debajo (verde) de $y(t) \equiv 0$ si se parecen a ella, pero de la definición sabemos que la estabilidad es para toda solución que parta cerca de la solución que se está analizando. Un razonamiento análogo se puede hacer para ver la inestabilidad por la izquierda.*

Ejemplo 2.1.2. *Se considera la ecuación $x'(t) = 1-(x(t))^2$. La solución $y(t) \equiv +1$ es asintóticamente estable. La solución $z(t) \equiv -1$ no es estable.*

Solución. *La solución de la ecuación diferencial dada es*

$$x(t) = \frac{ke^{2t}-1}{ke^{2t}+1}$$

pero si imponemos la condición inicial $x(0) = x^0$ tendremos que la solución estará dada por

$$x(t) = \frac{(1+x^0)e^{2t}-(1-x^0)}{(1+x^0)e^{2t}+(1-x^0)}$$

Por lo que tendremos que analizar los siguientes casos, para poder graficar las curvas soluciones.

- *Si $-1 < x^0 < 1$ las curvas soluciones serán las que se encuentran entre las rectas $y = 1$ e $y = -1$ (curvas rosada).*
- *Si $x^0 > 1$ las curvas solución serán las curvas que se encuentran arriba de la recta $y = 1$ (curvas moradas).*
- *$x^0 < -1$ las curvas solución serán las curvas que se encuentran abajo de la recta $y = -1$ (curvas marrón).*

Veamos que $y(t) \equiv +1$ es asintóticamente estable, para ello nos tomamos alguna solución que se encuentre por encima o por abajo de la recta $y(t) \equiv +1$.
Sea $x^0 > 1$, diremos que $y(t) \equiv +1$ es asintóticamente estable si para todo $\epsilon > 0$ existe $\delta > 0$ tal que para todo x^0 con $|x^0-1| < \delta$ se tiene que la solución $x(t;t_0,x^0)$ está definida en $[t_0,\infty)$ y se verifica que $|x(t;t_0,x^1)-y(t)| \leqslant \epsilon$ y si, además, existe $\eta > 0$ tal que $|x^0-1| < \eta$ implica que $\lim_{t\to\infty} |x(t;t_0,x^0)-y(t)| = 0$

$$\begin{aligned}
|x(t;t_0,x^1)-y(t)| &= \left|\frac{(1+x^0)e^{2t}-(1-x^0)}{(1+x^0)e^{2t}+(1-x^0)} - 1\right| \\
&= \left|\frac{(1+x^0)e^{2t}-(1-x^0)-(1+x^0)e^{2t}-(1-x^0)}{(1+x^0)e^{2t}+(1-x^0)}\right| \\
&= \frac{|-2(1-x^0)|}{|(1+x^0)e^{2t}+(1-x^0)|} = \frac{2|(1-x^0)|}{|(1+x^0)e^{2t}+(1-x^0)|} \\
&\leqslant |(x^0-1)|
\end{aligned}$$

de donde se deduce que, dado $\epsilon > 0$, si x^0 es tal que $|(x^0-1)| \leqslant \delta = \epsilon$ entonces $x(t;t_0,x^0)$ está definida y se verifica que $|x(t;t_0,x^1)-y(t)| \leqslant \epsilon$ $\forall t \geqslant t_0$ y además $\lim_{t\to\infty} |x(t;t_0,x^0)-y(t)| = 0$.
Un razonamiento análogo se puede hacer para observar que $z(t) \equiv -1$ no es asintóticamente estable.

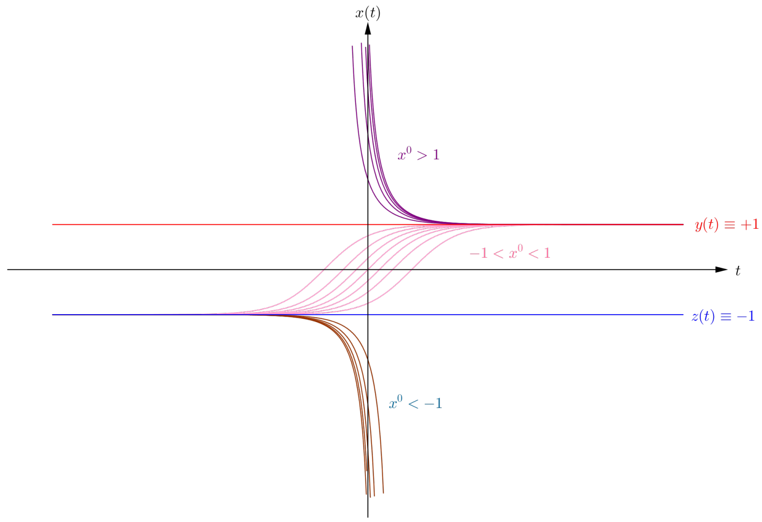


Figura 2.4: Curvas gráficas de las soluciones de la ecuación $x'(t) = 1 - (x(t))^2$ (Ejemplo 2.1.2)

Ejemplo 2.1.3. *Sea la ecuación $x'(t) = 0$. Todas sus soluciones son estable, pero ninguna es asintóticamente estable.*

Solución. *Las soluciones de la ecuación diferencial $x'(t) = 0$ son*

$$x(t) = a$$

donde a una constante real arbitraria.

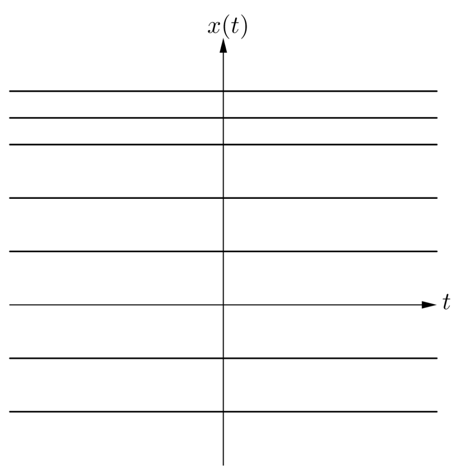


Figura 2.5: Curvas gráficas de las soluciones de la ecuación $x'(t) = 0$ (Ejemplo 2.1.3)

Notemos que si tomamos dos soluciones cualesquiera que estén cercanas, estas se matendrán siempre cercanas para todo tiempo, de ahí que se dice que las soluciones de $x'(t) = 0$ son estables, y no son asintóticamente estables ya que la estabilidad asintótica lo que nos dice es que una de las soluciones converge a la otra y en este caso no es así.

Ejemplo 2.1.4. *Las soluciones de $x'(t)+x(t)=0$ son asintóticamente estables a la derecha e inestable a la izquierda.*

Solución. *Sabemos que las soluciones de la ecuación diferencial $x'(t)+x(t)=0$ son dadas por $x(t)=ce^{-t}$.*
De la gráfica (ver Figura 2.6) podemos observar que cualesquiera dos soluciones que se tomen cercanas para un tiempo cualquiera inicial, estas siempre se manendrán cercanas para todo tiempo posterior y esto es lo que nos dice la definición de estabilidad, más aún para $t \to +\infty$ estas soluciones convergen una a la otra. Por tanto se dice que las soluciones $x(t)=ce^{-t}$ de $x'(t)+x(t)=0$ son asintóticamente estables.
Para ver la inestabilidad a la izquierda, veamos que si tomamos dos soluciones cualesquiera para un tiempo inicial que estén cercanas, estas soluciones para un tiempo posterior no se mantendrán cerca, es decir se alejan entre ellas, por lo que se dice que las soluciones son inestables.

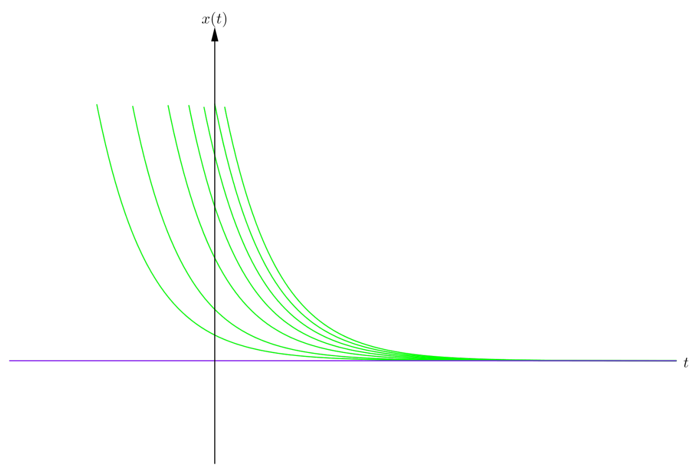


Figura 2.6: Curvas gráficas de las soluciones de la ecuación $x'(t)+x(t)=0$ (Ejemplo 2.1.4)

2.1.4. Acotación y estabilidad

Definición 2.1.3. *Consideramos la ecuación $x'(t)=f(t,x(t))$ y la solución $x(t;t_0,x^0)$ que suponemos definida para $t \geqslant t_0$. Decimos que $x(t;t_0,x^0)$ es* ***acotada*** *(a la derecha) si existe M, $0<M<\infty$, tal que $|x(t;t_0,x^0)| \leqslant M$ para $t \geqslant t_0$.*

Acotación y estabilidad son conceptos independientes en general. En efecto, toda solución $x(t)=t+c$ de $x'(t)=1$ es estable a ambos lados, pero no acotada a ningún lado. Asimismo, todas las soluciones de la ecuación

$$x'' = -\frac{1}{2}(x^2+(x^4+4x'^2)^{1/2})x$$

son de la forma $x=c\sin(ct+d)$ y así, son acotadas a la derecha y a la izquierda por c, pero, salvo la solución idénticamente nula, ninguna es estable ni a la derecha ni a la izquierda.

Sin embargo, para las ecuaciones lineales existe la siguiente relación.

Teorema 2.1.2. *La ecuación $x'(t) = A(t)x(t)$ es estable si, y sólo si, todas sus soluciones son acotadas.*

Demostración. Supongamos que la solución trivial $x(t) \equiv 0$ de $x'(t) = A(t)x(t)$ es estable. Entonces dado $\epsilon > 0$, existe $\delta > 0$ tal que $|z_i| \leqslant \delta$ implica $|x_i(t; t_0, z_i)| \leqslant \epsilon$ para todo $t \geqslant t_0$, si hacemos $z_i = \dfrac{\delta e_i}{2}$ tendremos

$$|x_i(t; t_0, \frac{\delta e_i}{2})| \leqslant \epsilon, \quad \forall t \geqslant t_0, \quad i = 1, \ldots, n$$

donde $\{e_1, e_2, \ldots, e_n\}$ es base canónica de $\mathbf{C}^n$ y δ es el número dado en la definición de estabilidad. Luego, si $\Phi(t)$ denota la matriz fundamental del sistema $x'(t) = A(t)x(t)$ cuyas columnas son las soluciones $x_i(t; t_0, \dfrac{\delta e_i}{2})$, entonces $\Phi(t)$ está acotada, lo cual a su vez implica la acotación de las soluciones del sistema $x'(t) = A(t)x(t)$.
Recíprocamente, supongamos que todas las soluciones de $x'(t) = A(t)x(t)$ estan acotadas en $[t_0, \infty)$ y sea $t \geqslant t_0$; entonces existe una constante M tal que $|\Phi(t)| \leqslant M$ para $t \in [t_0, \infty)$, siendo $\Phi(t)$ la matriz fundamental del sistema que verifica $\Phi(t_0) = I$ (las columnas de una matriz fundamental son soluciones linealmente independientes del sistema). Dado $\epsilon > 0$, $|x^0| \leqslant \dfrac{\epsilon}{M}$ implica que $|\Phi(t)x^0| \leqslant |\Phi(t)||x^0| \leqslant M\dfrac{\epsilon}{M} = \epsilon$ para $t \in [t_0, \infty)$ es decir la solución trivial es estable. □

Teorema 2.1.3. *Sea la ecuación lineal $x'(t) = A(t)x(t) + b(t)$.*

1. *Si sus soluciones son todas acotadas, entonces la ecuación es estable.*

2. *Si la ecuación es estable y una solución es acotada, entonces todas son acotadas.*

2.1.5. Ecuación variacional

Si deseamos estudiar la estabilidad de la solución $x(t; t_0, x^0)$ de la ecuación

$$x'(t) = f(t, x(t)) \tag{2.1.2}$$

se puede reducir al estudio de la estabilidad de la solución trivial de un sistema equivalente a (2.1.2) parece natural, observando las definiciones anteriores, introducir como nueva variable:

$$y(t) = x(t) - x(t; t_0, x^0)$$

con lo que la ecuación anterior resulta:

$$\begin{aligned}
y'(t) &= x'(t) - x'(t; t_0, x^0) \\
&= f(t, x(t)) - f(t, x(t; t_0, x^0)) \\
&= f(t, y(t) + x(t; t_0, x^0)) - f(t, x(t; t_0, x^0)) \\
y'(t) &= F(t, y(t)) \qquad (2.1.3)
\end{aligned}$$

Las propiedades de estabilidad de $x(t;t_0,x^0)$ como solución de (2.1.2) serán las mismas que las de la solución $y(t) \equiv 0$. Obsérvese que $y(t) \equiv 0$ es solución de esta ecuación (2.1.3) y que $x(t;t_0,x^0)$ es solución estable (asintóticamente estable) de (2.1.2) si, y sólo si, $y(t) \equiv 0$ es solución estable (asintóticamente estable) de (2.1.3). Si f tiene derivadas parciales continuas respecto de las componentes de x, la ecuación (2.1.3) se escribe:

$$y'(t) = f_x(t, x(t;t_0,x^0))y(t) + g(t, y(t)), \tag{2.1.4}$$

donde f_x representa la matriz jacobiana de f respecto de x y $g(t,u)$ es la función

$$g(t,u) = f(t, x(t;t_0,x^0) + u) - f(t, x(t;t_0,x^0)) - f_x(t, x(t;t_0,x^0))u,$$

Para t fijo se tiene $g(t,u) = o(|u|)$ para $|u| \to 0$, así, se puede conjeturar que el término $g(t,y(t))$ de (2.1.4) influya poco el comportamiento asintótico de la ecuación (2.1.4); es decir, se puede esperar que la solución $y(t) \equiv 0$ de (2.1.4) sea estable (asintóticamente estable) si, y sólo si, $y(t) \equiv 0$ es solución estable (asintóticamente estable) de la ecuación

$$y'(t) = f_x(t, x(t;t_0,x^0))y(t) \tag{2.1.5}$$

que es lineal y homogénea. Veremos que bajo ciertas condiciones resulta en efecto así, aunque esto no es cierto en general.

La ecuación (2.1.5) se denomina ***ecuación variacional de*** $x'(t) = f(t,x(t))$ ***respecto de la solución*** $x(t;t_0,x^0)$.

Por ejemplo, la solución $y(t) \equiv 0$ de $x'(t) = 0$ es estable, pero esta misma función como solución de $x'(t) = (x(t))^2$ no lo es. Obsérvese que la primera ecuación es la ecuación variacional de la segunda respecto de $y(t) \equiv 0$.

En cambio, $y(t) \equiv 0$ es solución inestable de $x'(t) = x(t)$, pero es asintóticamente estable como solución de $x'(t) = x(t) - e^t(x(t))^3$, cuyas soluciones tienen la forma

$$x(t) = \frac{Ce^t}{+\sqrt{1 + \frac{2}{3}C^2(e^{3t} - 1)}}$$

2.1.6. Estabilidad orbital

Sea $x'(t) = f(x(t))$ un sistema autónomo y supongamos que se tiene una solución $y(t)$ no constante que es periódica de periodo T. Entonces $y(t+h)$ es también solución de periodo T para cualquier $h \in \mathbf{R}$. Como $|y(t_0+h) - y(t_0)|$ puede hacerse tan pequeño como se quiera eligiendo h pequeño, y, sin embargo, $|y(t+h) - y(t)|$ con h fijo no tienden a cero para $t \to \infty$, resulta que $y(t)$ no es asintóticamente estable. Pero es muy distinto el comportamiento de $y(t)$ según que una solución $z(t)$ de la ecuación que pasa por un punto $(t_1, z(t_1))$ tal que $z(t_1)$ esté cercano a la curva γ (órbita) en $\mathbf{R}^n$ definida por $t \to y(t)$ se mantenga cercana a esta órbita para $t \geqslant t_1$, o que se aleje mucho de la curva. Esto motiva la definición siguiente.

Definición 2.1.4. *Sea γ la órbita en $\mathbf{R}^n$ definida por $t \to y(t)$, donde $y(t)$ es como se indica arriba. La solución $y(t)$ se denomina **orbitalmente estable** si para cualquier $\epsilon > 0$ existe $\delta > 0$ tal que si $x(t; t_0, x^0)$ es solución que satisface $d(x(t; t_0, x^0), \gamma) \leqslant \delta$, entonces $x(t; t_0, x^0)$ existe para $t \geqslant t_0$ y $d(x(t; t_0, x^0), \gamma) \leqslant \epsilon$ para todo $t \geqslant t_0$. (Aquí, $d(x, \gamma)$ significa la distancia euclídea del punto x al conjunto γ.)*

Definición 2.1.5. *La solución $y(t)$ se denomina **orbitalmente estable** (y la órbita γ se llama entonces **ciclo límite**) cuando $y(t)$ es orbitalmente estable y además existe $\eta > 0$ tal que si $d(x(t; t_0, x^0), \gamma) \leqslant \eta$ entonces $d(x(t; t_0, x^0), \gamma) \to 0$ para $t \to +\infty$.*

Nota: Todas las definiciones anteriores dan una idea de algunos de los problemas que se estudian en la teoría de la estabilidad de ecuaciones diferenciales ordinarias. Para muchos propósitos estas definiciones son insuficientes y es necesario afinar y modificar los conceptos aquí introducidos. En estas notas trataremos solamente de algunos problemas fundamentales relativos a las nociones introducidas.

2.2. Estabilidad de ecuaciones lineales

Al estudiar las propiedades de estabilidad de una solución $x(t; t_0, x^0)$ del sistema lineal $x'(t) = A(t)x(t) + b(t)$, hay que estimar $|x(t; t_0, x^1) - x(t; t_0, x^0)|$ para soluciones $x(t; t_0, x^1)$ tales que $|x^1 - x^0|$ sea pequeño; pero la función $x(t; t_0, x^1) - x(t; t_0, x^0)$ es solución del sistema homogéneo asociado $x'(t) = A(t)x(t)$, con las propiedades de estabilidad de una solución cualquiera del sistema lineal no homogéneo son exactamente las mismas que las de la solución trivial del sistema homogéneo asociado. En particular, cualquier solución no nula del sistema $x'(t) = A(t)x(t)$ tienen las mismas propiedades de estabilidad que la solución $y(t) \equiv 0$.

Como hemos visto, tiene sentido hablar de la estabilidad (asintótica) de una ecuación lineal, y nos podemos reducir, para estudiarla, a considerar el carácter de la solución $y(t) \equiv 0$ de $x'(t) = A(t)x(t)$. El teorema siguiente expresa una condición de estabilidad en términos de una matriz fundamental. Como siempre, suponemos A función continua de $[t_0, \infty)$ a $M(n, \mathbf{R})$.

Obsérvese que, si bien a veces resulta conveniente introducir números complejos para algún punto particular del estudio de las ecuaciones que se consideran, gracias a la forma de Jordan real para una matriz real.

Teorema 2.2.1. *Sea $X(t)$ una matriz fundamental de la ecuación*

$$x'(t) = A(t)x(t) \tag{2.2.1}$$

Entonces:

1) *La ecuación (2.2.1) es estable si, y sólo si, existe una constante M, $0 < M < \infty$, tal que $\|X(t)\| \leqslant M$ para todo $t \geqslant t_0$.*

2) La ecuación (2.2.1) es asintóticamente estable si, y sólo si, $\|X(t)\| \to 0$ *para* $t \to \infty$. *(Se considera, por ejemplo, la norma euclídea.)*

Demostración. Como $X(t)$ es matriz fundamental, podemos suponer, sin pérdida de generalidad, que $X(t_0) = I$, ya que para cualquier par de matrices fundamentales $X_1(t)$, $X_2(t)$ se tiene

$$X_1(t) = X_2(t)C,$$

con C matriz constante no singular.
La solución $x(t)$ tal que $x(t_0) = x^0$ es entonces $x(t) = X(t)x^0$.
Para demostrar 1), supongamos primero que $\|X(t)\| \leqslant M$. Se tiene entonces

$$|x(t)| = \|X(t)x^0\| = \|X(t)\||x^0| \leqslant M|x^0|$$

y, así, si $|x^0| \leqslant \epsilon M^{-1} = \delta$, se tiene $|x(t)| \leqslant \epsilon$ para $t \geqslant t_0$. Lo que demuestra la estabilidad de $y(t) \equiv 0$. Recíprocamente, si se tiene la estabilidad de $y(t) \equiv 0$, entonces, para todo $\epsilon > 0$ existe $\delta > 0$ tal que si $|x^0| \leqslant \delta$ se tiene $|x(t)| \leqslant \epsilon$, es decir, $|X(t)x^0| \leqslant \epsilon$. Fijemos un ϵ y un δ que satisfacen esta relación. Entonces se tiene:

$$\|X(t)\| = \sup\{|X(t)z| : |z| \leqslant 1\} = \sup\left\{|X(t)z\delta|\frac{1}{\delta} : |z\delta| \leqslant \delta\right\} \leqslant \frac{\epsilon}{\delta}$$

lo que demuestra

$$\|X(t)\| \leqslant \frac{\epsilon}{\delta} < \infty \quad para \quad todo \quad t \geqslant t_0$$

Para demostrar 2), supongamos primero que $\|X(t)\| \to 0$ para $t \to \infty$.
Entonces:

$$|x(t)| \leqslant \|X(t)\||x^0| \to 0 \quad para \quad t \to \infty$$

y, así, (2.2.1) es asintóticamente estable. Recíprocamente, supongamos que existe $\eta > 0$ tal que

$$|x(t)| \to 0 \quad para \quad t \to \infty \quad si \quad |x^0| \leqslant \eta$$

Entonces:

$$|X(t)x^0| \to 0 \quad para \quad t \to \infty \quad y \quad |x^0| \leqslant \eta,$$

o, lo que es lo mismo,

$$\left|X(t)\frac{x^0}{\eta}\right| \to 0 \quad para \quad t \to \infty \quad siempre \quad que \quad \left|\frac{x^0}{\eta}\right| \leqslant 1$$

es decir,

$$|X(t)y| \to 0 \quad para \quad t \to \infty \quad siempre \quad que \quad |y| \leqslant 1$$

pero esto implica

$$\|X(t)\| \to 0 \quad para \quad t \to \infty.$$

□

El teorema anterior nos asegura que el comportamiento de la norma de una matriz fundamental proporciona la información deseada sobre estabilidad de la ecuación lineal. En el caso de la ecuación lineal autónoma (coeficientes constantes), $A(t) \equiv A$, sabemos que una matriz fundamental es e^{At} y, así, el estudio de la estabilidad de tal ecuación se reduce al de la norma de esta matriz. Los resultados siguientes proporcionan desigualdades para la norma de esta matriz que simplifican el problema de la estabilidad.

Teorema 2.2.2. *Sea J una matriz $n \times n$ (real o compleja), $J = \lambda I + L$, siendo $\lambda \in \mathbf{C}$, I la identidad, $L^h = 0$ para $h \leqslant n$. Entonces existe un polinomio $p(t)$ de coeficientes positivos y de grado $h-1$ tal que*

$$\|e^{Jt}\| \leqslant p(t)e^{tRe\lambda}$$

Demostración. Se tiene, si $t \geqslant 0$,

$$\begin{aligned}
\|e^{Jt}\| &= \|e^{(\lambda I+L)t}\| \;\; \textit{sustituyendo } J \\
&= \|e^{\lambda It}e^{Lt}\| \;\; \textit{por propiedad } 1.1.2\ 2) \\
&= \|Ie^{\lambda t}e^{Lt}\| \;\; \textit{por ser } e^{\lambda It} = Ie^{\lambda t} \\
&\leqslant \|Ie^{\lambda t}\|\|e^{Lt}\| \;\; \textit{por propiedad } 1.1.1b) \\
&= |e^{\lambda t}|\|I\|\|e^{Lt}\| \;\; \textit{por propiedad } 2) \textit{ de la definición } 1.1.5 \\
&= |e^{\lambda t}|\|e^{Lt}\| \;\; \textit{por ser } \|I\| = 1 \\
&= e^{tRe\lambda}\|I + \frac{Lt}{1!} + \frac{L^2t^2}{2!} + ... + \frac{L^{h-1}t^{h-1}}{(h-1)!}\| \;\; \textit{ya que } |e^{\lambda t}| = e^{tRe\lambda} \\
&\leqslant e^{tRe\lambda} \underbrace{\left(1 + \frac{t\|L\|}{1!} + ... + \frac{t^{h-1}\|L^{h-1}\|}{(h-1)!}\right)}_{p(t)} \;\; \textit{por inciso } 3 \textit{ de la definición } 1.1.5
\end{aligned}$$

por lo tanto

$$\|e^{Jt}\| \leqslant p(t)e^{tRe\lambda}$$

lo que demuestra el teorema. □

Corolario 2.2.1. *Sea A una matriz real $n \times n$ con único autovalor λ y sea J su forma de Jordan, $A = P^{-1}JP$. Entonces existe un polinomio de coeficientes no negativos y de grado menor o igual a $n-1$ tal que*

$$\|e^{At}\| \leqslant p(t)e^{tRe\lambda}$$

para $t \geqslant 0$.

Además, si el índice de λ es $v(\lambda) = 1$, entonces se puede tomar $p(t) \equiv c > 0$ y, así:

$$\|e^{At}\| \leqslant ce^{tRe\lambda}$$

para $t \geqslant 0$.

Demostración.

$$\begin{aligned}
\|e^{At}\| &= \|P^{-1}e^{Jt}P\| \;\; \textit{ya que } e^{At} = P^{-1}e^{Jt}P \\
&\leqslant \|P^{-1}\|\|e^{Jt}\|\|P\| \;\; \textit{por propiedad } 1.1.1b) \\
&\leqslant \|P^{-1}\|\|P\|p_1(t)e^{tRe\lambda} \;\; \textit{aplicando el teorema anterior} \\
&= \underbrace{c_1p_1(t)}_{p(t)}e^{tRe\lambda} = p(t)e^{tRe\lambda} \;\; \textit{haciendo } \|P^{-1}\|\|P\| = c_1
\end{aligned}$$

por lo tanto

$$\|e^{At}\| \leqslant p(t)e^{tRe\lambda}$$

Demostración, para la segunda, si $v(\lambda) = 1$, es claro que $J = \lambda I$ y, así:

$$\begin{aligned}
\|e^{At}\| &= \|P^{-1}e^{Jt}P\| \;\; \textit{ya que } e^{At} = P^{-1}e^{Jt}P \\
&\leqslant \|P^{-1}\|\|e^{Jt}\|\|P\| \;\; \textit{por propiedad } 1.1.1b) \\
&\leqslant \|P^{-1}\|\|P\|p_1(t)e^{tRe\lambda} \;\; \textit{aplicando el teorema anterior} \\
&= \underbrace{c_1p_1(t)}_{p(t)}e^{tRe\lambda} \;\; \textit{haciendo } \|P^{-1}\|\|P\| = c_1 \\
&= p(t)e^{tRe\lambda} \;\; \textit{tomando } p(t) \equiv c \\
&= ce^{tRe\lambda}
\end{aligned}$$

por lo tanto

$$\|e^{At}\| \leqslant ce^{tRe\lambda}$$

□

Corolario 2.2.2. *a) Sea A una matriz real $n \times n$ y sea*

$$\mu = \max\{Re\lambda : \lambda \textit{ autovalor de } A\}$$

Entonces existe un polinomio $p(t)$ de grado menor o igual que $n-1$ y de coeficientes no negativos tal que

$$\|e^{At}\| \leqslant p(t)e^{\mu t}$$

para todo $t \geqslant 0$.

b) Sea A una matriz real $n \times n$ tal que

$$\mu = \max\{Re\lambda : \lambda \textit{ autovalor de } A\} = 0$$

y si $Re\lambda$=0 entonces $v(\lambda) = 1$. Entonces existe una constante $c > 0$ tal que

$$\|e^{At}\| \leqslant c$$

para todo $t \geqslant 0$.

Demostración. Para la parte a). Sea $A = P^{-1}JP$, J forma de Jordan, y sea

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_p \end{pmatrix}$$

y $J_i = \lambda_i I_{n_i \times n_i} + L_i$ donde $I_{n_i \times n_i}$ es la matriz identidad y L_i es nilpotente de grado n_i, $L^{n_i} = 0$. Se obtiene entonces

$$\begin{aligned} \|e^{At}\| &\leqslant \|P^{-1}\|\|e^{Jt}\|\|P\| \\ &\leqslant \|P^{-1}\| \left(\sum_{i=1}^{p} \|e^{J_i t}\| \right) \|P\| \end{aligned}$$

Ahora, para cada $i = 1, ..., p$

$$\|e^{J_i t}\| \leqslant e^{tRe\lambda_i} \left(1 + \frac{\|L_i\| t}{1!} + ... + \frac{\|L_i\|^{n_i - 1} t^{n_i - 1}}{(n_i - 1)!} \right)$$

Si se considera $\mu = máx \{Re\lambda_i : i = 1, ..., p\}$ se obtiene

$$\|e^{At}\| \leqslant \underbrace{\|P^{-1}\|\|P\| \left(\sum_{i=1}^{p} \sum_{j=1}^{n_i} \frac{\|L_i\|^j}{j!} t^j \right)}_{p(t)} e^{\mu t}$$

por lo tanto

$$\|e^{At}\| \leqslant p(t) e^{\mu t}$$

Ahora para la parte b)

$$\begin{aligned} \|e^{At}\| &\leqslant \|P^{-1}\| \left(\sum_{i=1}^{p} \|e^{J_i t}\| \right) \|P\| \\ &= \underbrace{\|P^{-1}\|\|P\|}_{c_1} \underbrace{\left(\sum_{i=1}^{p} \sum_{j=1}^{n_i} \frac{\|L_i\|^j}{j!} t^j \right)}_{p(t)} e^{\mu t} \\ &= c_1 p(t) e^{\mu t} \quad \mu = 0 \\ &= c_1 p(t) = c \end{aligned}$$

por tanto $\|e^{At}\| \leqslant c$. □

De los resultados anteriores se obtienen inmediatamente criterios de acotación o estabilidad. Obsérvese que estos resultados proporcionan además una estimación cualitativa útil sobre el comportamiento para $t \to \infty$ de las soluciones.

Teorema 2.2.3. *Se considera el problema*

$$(P)\begin{cases} x'(t) = Ax(t) \\ x(0) = x^0 \end{cases}$$

siendo A una matriz real $n \times n$ y sea $\sigma(A)$ el espectro de A.

a) Si $\mu = \text{máx}\,\{Re\lambda : \lambda \in \sigma(A)\} < 0$, entonces existe un polinomio $p(t)$ que depende sólo de A, de grado menor o igual que $n-1$ y de coeficientes no negativos, tal que la solución de (P), $x(t) = e^{At}x^0$, verifica:

$$|x(t)| \leqslant p(t)e^{\mu t}|x^0|$$

y, así, la ecuación $x'(t) = Ax(t)$ es asintóticamente estable.

b) Si $\mu = \text{máx}\,\{Re\lambda : \lambda \in \sigma(A)\} = 0$, y se verifica, para todo $\lambda \in \sigma(A)$ con $Re\lambda = 0$, que $v(\lambda) = 1$, entonces existe una constante $c > 0$ que depende sólo de A tal que $|x(t)| \leqslant c|x^0|$. En este caso, la ecuación $x'(t) = Ax(t)$ es estable, pero no asintóticamente estable.

c) Si $\mu = \text{máx}\,\{Re\lambda : \lambda \in \sigma(A)\} > 0$, entonces la ecuación $x'(t) = Ax(t)$ no es estable.

Demostración. La de a) y la primera parte de b) es una consecuencia directa del corolario 2.2.2. Para ver que en b) no hay estabilidad asintótica, sea λ un autovalor de A tal que $Re\lambda = 0$, es decir, sea $\lambda = \beta i$ con $\beta \in \mathbf{R}$ y sea $x^0 \neq 0$ autovector correspondiente a λ, es decir, $Ax^0 = \beta i x^0$. Entonces, claramente, $e^{At}x^0 = e^{\beta i t}x^0$ y, por tanto, la solución $x(t) = e^{At}x^0$ de (P) verifica $|x(t)| = |x^0|$ y, de ese modo, no hay estabilidad asintótica.
Para probar c) procedemos análogamente. Sea $\lambda = \alpha + \beta i$ con $\alpha > 0$ autovalor de A, y sea $x^0 \neq 0$ tal que $Ax^0 = (\alpha + \beta i)x^0$. Entonces

$$x(t) = e^{At}x^0 = e^{(\alpha+\beta i)t}x^0$$

y, por tanto, $|x(t)| = e^{\alpha t}|x^0| \to \infty$. □

2.2.1. Ejemplos

Ejemplo 2.2.1. *Sea*

$$J = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix}$$

Hállese un polinomio $p(t)$ tal que $\|e^{Jt}\| \leqslant p(t)e^{2t}$ para todo t.

Solución. *Por teorema 2.2.2 se tiene que $J = \lambda I + L$, donde*

$$J = \underbrace{2}_{\lambda}\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{I} + \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}}_{L}$$

se observa que $L^3 = 0$. *Por lo que*

$$p(t) = 1 + \frac{t\|L\|}{1!} + \frac{t^2\|L^2\|}{2!} + \frac{t^3\|L^3\|}{3!} = 1 + t + \frac{t^2}{2!}$$

Por lo tanto

$$\|e^{Jt}\| \leqslant (1 + t + \frac{t^2}{2!})e^{2t}$$

Ejemplo 2.2.2. *Estúdiese la estabilidad de los siguientes sistemas lineales:*

a)

$$x'(t) = \begin{pmatrix} -2 & 6 \\ -2 & 5 \end{pmatrix} x(t) + \begin{pmatrix} 1 \\ \cos(t) \end{pmatrix}$$

b)

$$x'(t) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -3 & 1 \\ 0 & 1 & -3 \end{pmatrix} x(t) + \begin{pmatrix} 1 \\ t \\ e^t \end{pmatrix}$$

Solución. *Para resolver los dos literales anteriores nos auxiliaremos de los teoremas 2.1.1 y 2.2.3 , por lo que estudiaremos la estabilidad de la solución nula de* $x'(t) = Ax(t)$ *para decir que tipo de estabilidad presenta la ecuación anterior* $x'(t) = Ax(t) + b(t)$.

a) Los valores propios de A *son* $\lambda = 1, 2$ *por lo que* $\sigma(A) = \{1, 2\}$, *es decir que los valores propios son ambos mayores que* 0 *así* $\mu > 0$, *por lo tanto aplicando el teorema 2.2.3 c) se concluye que la ecuación no es estable.*
Por teorema 2.1.1 la ecuación,

$$x'(t) = \begin{pmatrix} -2 & 6 \\ -2 & 5 \end{pmatrix} x(t) + \begin{pmatrix} 1 \\ \cos(t) \end{pmatrix}$$

no es estable.

b) Los autovalores de A *son* $\lambda = -1, -2, -4$ *de donde* $\sigma(A) = \{-1, -2, -4\}$, *así,* $\mu < 0$ *y aplicando el teorema 2.2.3 se tiene que la ecuación* $x'(t) = Ax(t)$ *es asintóticamente estable. Por teorema 2.1.1 se concluye que la ecuación*

$$x'(t) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -3 & 1 \\ 0 & 1 & -3 \end{pmatrix} x(t) + \begin{pmatrix} 1 \\ t \\ e^t \end{pmatrix}$$

es asintóticamente estable.

2.3. Algunos criterios de estabilidad asintótica

En la practica, cuando para un sistema que se rige en su funcionamiento por una ecuación lineal con coeficientes constantes es deseable que haya estabilidad, el tipo de estabilidad que se debe lograr es la

asintótica. En efecto, la estabilidad no asintótica, que se da cuando para algún autovalor λ se verifica $Re\lambda = 0$, $v(\lambda) = 1$, es muy lábil, ya que basta una perturbación pequeña de los coeficientes para que sea $Re\lambda > 0$ y así no haya estabilidad.
En cambio, si para todo λ es $Re\lambda < 0$, una perturbación pequeña de los coeficientes no variará esta propiedad y así seguirá habiendo estabilidad.
A continuación presentamos algunos criterios de estabilidad asintótica para una ecuación con coeficientes constantes reales, que son las de ordinario aparecen en la práctica. Obsérvese primero que el problema se ha reducido, por el teorema 2.2.3, a una cuestión puramente algebraica. Se trata simplemente de averiguar si todos los autovalores de una matriz, o, en forma equivalente, todos los ceros de un polinomio (el polinomio característico de tal matriz) tiene parte real estrictamente negativa.
El primer criterio es de muy sencilla aplicación, y, aunque no constituye una caracterización necesaria y suficientes de estabilidad, permite decidir en muchos casos la carencia de estabilidad.

Definición 2.3.1. *Diremos que $F(z) = z^n + a_1 z^{n-1} + ... + a_n$ con $a_j \in \boldsymbol{R}$ es un polinomio de Hurwitz, si todas sus raíces tienen parte real negativa.*

Teorema 2.3.1. *Si $F(z) = z^n + a_1 z^{n-1} + ... + a_n$ con $a_j \in \boldsymbol{R}$ tiene algún coeficiente menor o igual que cero, entonces $F(z)$ tiene alguna raíz que no tiene parte real estrictamente negativa.*

Demostración. En efecto, si $F(z)$ tuviese todas sus raíces con parte real estrictamente negativa, entonces

$$F(z) = \prod_{i=1}^{h}(z-\alpha_i)^{m_i} \prod_{j=1}^{r}\left[(z-\beta_j)^2 + \gamma_j^2\right]^{n_j}$$

con $\alpha_i < 0$, $\beta_j < 0$, $\gamma_j \in \mathbf{R}$. Por tanto, todos sus coeficientes son positivos, es decir, $a_j > 0$. Esto demuestra el teorema. □

El criterio anterior permite, por ejemplo, decidir que

$$z^8 + 17z^7 + 12z^6 + 13z^5 + 4z^4 + 21z^2 + 6z + 4$$

tiene alguna raiz con parte real mayor o igual que cero. Pero no nos permite decidir si

$$z^3 + 11z^2 + 6z + 6$$

por ejemplo, tiene todas sus raíces λ con $Re\lambda < 0$ o no.

Corolario 2.3.1. *Todo polinomio de segundo grado es de Hurwitz si y sólo si $a_i > 0$, $i = 1, 2$.*

Demostración. Cuando $n = 1$, $F_1(z) = z + a_1 \iff z = -a_1$. Para $n = 2$, sea el polinomio $F_2(z) = z^2 + a_1 z + a_2$ cuya ecucación característica es $F(\lambda) = \lambda^2 + a_1\lambda + a_0$, entonces

$$\lambda = \frac{a_1 \pm \sqrt{a_1^2 - 4a_0}}{2}$$

Si las raíces son reales, como $\sqrt{a_1^2 - 4a_0} < \sqrt{a_1^2} = a_1$ ya que $a_0 > 0$, se tiene entonces que ambas raíces tienen parte real negativa. Si las raíces son complejas se tiene que la parte real $-a_1$ es siempre negativa. $\square$

El criterio siguiente, basado en un caso particular sencillo de un teorema general de la teoría de la variable compleja, proporciona una condición necesaria y suficiente para determinar la estabilidad asintótica. Exponemos primero la noción de variación del argumento de $F(z)$ cuando z recorre un arco de curva y un lema que conduce fácilmente al teorema.

Definición 2.3.2. *Sea G un arco de curva orientada en $\mathbf{C}$ dado por una función continua* $\mathrm{g} : [0,1] \rightarrow \mathbf{C}$. *Sea $F : z \in\rightarrow F(z) \in \mathbf{C}$ una función continua en G tal que $f(z) \neq 0$ para todo $z \in G$. Consideramos una cualquiera de las determinaciones del argumento de $F(\mathrm{g}(0))$ en radianes que denominaremos $ArgF(\mathrm{g}(0))$, y para cada $t \in [0,1]$ tomamos la determinación $ArgF(\mathrm{g}(t))$ del argumento de $F(\mathrm{g}(t))$ que hace que la función $ArgF(\mathrm{g}(t))$ sea continua en $[0,1]$. Denominaremos variación del argumento de $F(z)$ sobre G al valor.*

$$v(F(z), G) = ArgF(\mathrm{g}(1)) - ArgF(\mathrm{g}(0))$$

Las propiedades siguientes que utilizamos son de demostración sencilla:

1) $v(F(z), G)$ no depende de la parametrización f de G ni de la determinación de $ArgF(\mathrm{g}(0))$ que se ha escogido.

2) Si G_1 y G_2 son dos curvas determinadas por

$$\mathrm{g}_1 : [0, 1/2] \rightarrow \mathbf{C} \quad y \quad \mathrm{g}_2 : [0, 1/2] \rightarrow \mathbf{C}$$

tales que

$$\mathrm{g}_1(1/2) = \mathrm{g}_2(1/2) \quad y \quad G = G_1 + G_2$$

es la curva determinada por

$$\mathrm{g}(x) = \begin{cases} \mathrm{g}_1(t) & si \quad 0 \leqslant t \leqslant 1/2 \\ \mathrm{g}_2(t) & si \quad 1/2 \leqslant t \leqslant 1 \end{cases}$$

entonces

$$v(F(z), G_1 + G_2) = v(F(z), G_1) + v(F(z), G_2)$$

3) Si F_1 y F_2 son dos funciones en las condiciones de la F de la definición, entonces

$$v(F_1(z)F_2(z), G) = v(F_1(z), G) + v(F_2(z), G)$$

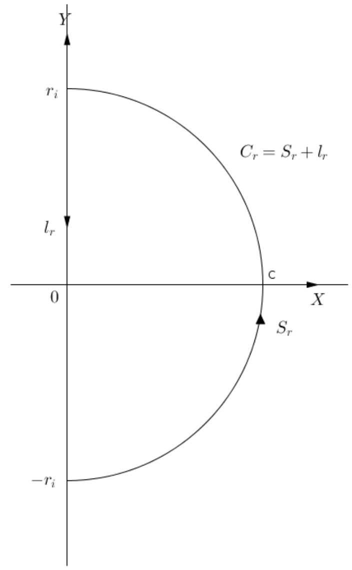


Figura 2.7: Representación gráfica del lema

Lema 2.3.1. *Sea $F(z) = z^n + a_1 z^{n-1} + ... + a_n$ un polinomio con coeficientes en* **C**. *Sea C_r para cada $r > 0$ fijo la curva cerrada formada por el segmento que va de ri a $-ri$ y la semicircunferencia de centro el origen que va de $-ri$ a ri situada en $Re\lambda \geqslant 0$. Supongamos $F(z) \neq 0$ siempre que $z \in C_r$. Entonces el número de ceros de $F(z)$ en D_r, interior del semicírculo determinado por C_r, contando cada cero tantas veces como indica su orden de multiplicidad, es igual a $v(F(z), C_r)/2\pi$*

Demostración. El polinomio $F(z)$ se puede escribir:

$$F(z) = (z - \lambda_1)^{\alpha_1}(z - \lambda_2)^{\alpha_2}.....(z - \lambda_p)^{\alpha_p}$$

no siendo λ_1, λ_2,... λ_p de C_r, ya que $F(z) \neq 0$, cuando $z \in C_r$. Utilizando la propiedad 2) de $v(F(z), G)$ se obtiene:

$$v(F(z), C_r) = \alpha v(z - \lambda_1, C_r) + ... + \alpha_p v(z - \lambda_p, C_r)$$

Es claro que si $\lambda_i \in D$ entonces

$$v(z - \lambda_i, C_r) = 2\pi,$$

y si $\lambda_i \notin D$ entonces

$$v(z - \lambda_i, C_r) = 0$$

Por tanto:

$$v(F(z), C_r) = 2\pi m$$

Siendo m el número de ceros en D_r, cada uno contando según su multiplicidad. $\square$

Supongamos ahora que $F(z)$ tiene coeficientes reales y no posee ningún cero imaginario puro, es decir:

$$F(iy) = A(y) + iB(y) = (a_n - a_{n-2}y^2 + a_{n-4}y^4...) + i(a_{n-1}y - a_{n-3}y^3 + ...) \neq 0$$

para todo $y \in \mathbf{R}$, lo que equivale a firmar que $A(y)$ y $B(y)$ no tienen ninguna raíz real común.
Llamemos I_r al segmento que va de ri a $-ri$, y S_r a la semicircunferencia que va de $-ri$ a $+ri$ con centro en el origen situada en $Rez \geqslant 0$. Entonces

$$C_r = I_r + S_r.$$

Supongamos que $r > 1$. Se tiene, usando la propiedad 3) de $v(F(z), G)$:

$$v(F(z), C_r) = v(F(z), I_r) + v(F(z), S_r)$$

para $v(F(z), S_r)$ se tiene:

$$\begin{aligned} v(F(z), S_r) &= v\left(\frac{F(z)}{(z-1)^n}(z-1)^n, S_r\right) \\ &= v\left(\frac{F(z)}{(z-1)^n}, S_r\right) + v((z-1)^n, S_r) \end{aligned}$$

Calculemos $v(F(z), C_r)$ cuando $r \to \infty$. Como para cualquier $\tau \in \mathbf{R}$ resulta que

$$\frac{F(re^{i\tau})}{(re^{i\tau}-1)^n} \to 1 \quad para \quad r \to \infty,$$

obtenemos de aquí que

$$\frac{F(z)}{(z-1)^n} \to 0 \quad para \quad r \to \infty,$$

ya que para r suficientemente grande $\dfrac{F(z)}{(z-1)^n}$ está muy próximo al punto 1 cuando $z \in S_r$. Por otra parte, se tiene claramente que

$$v((z-1)^n, S_r) = nv(z-1, S_r) = n\pi$$

para $r \to \infty$. Así, $v(F(z), S_r) \to n\pi$ para $r \to \infty$. También es sencillo ver que $v(F(z), I_r)$ coincide con la variación $-L_r$ del arco cuya tangente es $\dfrac{B(y)}{A(y)}$ (medida a lo largo de una determinación continua de tal arco) cuando y varía de $+r$ a $-r$. Como $v(F(z), C_r)$ es un múltiplo entero de 2π y es función continua de r en aquellos valores r en los que está definida ($F(z)$ ha de ser no nula sobre C_r), resulta que $v(F(z), C_r)$ tiene un límite para $r \to \infty$ y, así, $-L_r \to -L$ para $r \to \infty$. Por tanto :

$$\lim_{r\to\infty} v(F(z), C_r) = -L + n\pi$$

Obsérvese que L es la variación del arco tangente de $\dfrac{B(y)}{A(y)}$, medida en radianes a lo largo de una determinación continua de tal arco, cuando y varía de $-\infty$ a $+\infty$.
Obsérvese también que para r suficientemente grande todas las raíces de $F(z)$ con parte real positiva están en D_r. Así, el número de tales raíces con parte real positiva, según el lema, será cero si, y sólo si, $L = n\pi$. De este modo, resulta el siguiente teorema.

Teorema 2.3.2. *Sea $F(z)$ un polinomio con coeficientes reales de grado n y sea $F(iy) = A(y)+iB(y) \neq 0$ para $y \in \boldsymbol{R}$. Entonces todas las raíces de $F(z)$ tienen parte real negativa si, y sólo si, $L = n\pi$, donde L es la variación del arco tangente de $\dfrac{B(y)}{A(y)}$ cuando y varía de $-\infty$ a $+\infty$.*

Ejemplo 2.3.1. *Estúdiese la estabilidad de la ecuación lineal*

$$x''' + 11x'' + 6x' + 6x = 0 \tag{2.3.1}$$

Solución. *El polinomio característico es*

$$F(z) = z^3 + 11z^2 + 6z + 6$$

y se tiene:

$$F(iy) = A(y) + iB(y) = (6 - 11y^2) + i(6y - y^3)$$

la curva

$$\frac{B(y)}{A(y)} = \frac{y(y-\sqrt{6})(y+\sqrt{6})}{(\sqrt{11}y-\sqrt{6})(\sqrt{11}y+\sqrt{6})}$$

es de la forma siguiente:

$$\frac{B(y)}{A(y)}$$

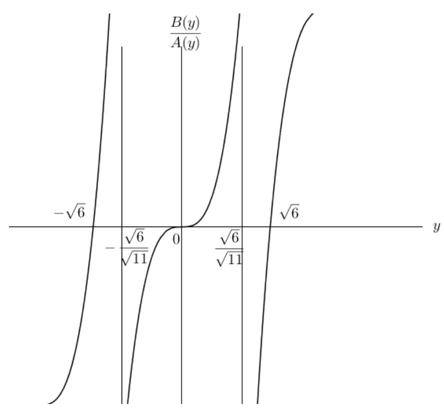


Figura 2.8: Representación gráfica del ejemplo

y, así, $L = 3\pi$ y, por tanto, todas las raíces de $F(z)$ tiene parte real negativa. Entonces la ecuación (2.3.1) es asintóticamente estable.

La conveniencia de la determinación de la estabilidad a la vista solamente de los coeficientes de $F(z)$ sin necesidad de calcular las raíces de $A(y)$ y $B(y)$ ha conducido al criterio de ***Routh-Hurwitz***. Obsérvese que en el ejemplo anterior los valores exactos de las raíces de $A(y)$ y $B(y)$ no interesan tanto como la situación relativa de estas raíces. La utilización del teorema clásico de ***Sturm*** sobre este aspecto da lugar al teorema siguiente.

Teorema 2.3.3 (Criterio de Routh-Hurwitz)**.** *Sea*

$$F(z) = z^n + a_1 z^{n-1} + ... + a_j z^{n-j} + ... + a_n$$

un polinomio con coeficientes reales. Entonces todos los ceros de $F(z)$ tienen parte real negativa si, y sólo si todos los determinantes de la matrices diagonales principales de la matriz

$$\begin{pmatrix} a_1 & 1 & 0 & 0 & 0 & \dots & 0 \\ a_3 & a_2 & a_1 & 1 & 0 & \dots & 0 \\ a_5 & a_4 & a_3 & a_2 & a_1 & \dots & 0 \\ a_7 & a_6 & a_5 & a_4 & a_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & a_n \end{pmatrix}$$

son positivos, es decir:

$D_1 = a_1 > 0,\ D_1 = \det\begin{pmatrix} a_1 & 1 \\ a_3 & a_2 \end{pmatrix} > 0,\ D_3 = \det\begin{pmatrix} a_1 & 1 & 0 \\ a_3 & a_2 & a_1 \\ a_5 & a_4 & a_3 \end{pmatrix} > 0, \dots$

Así, el criterio para $z^2+a_1z+a_2$ es $a_1 > 0$,$a_2 > 0$. Para $z^3+a_1z^2+a_2z+a_3$ es $a_1 > 0$, $a_1a_2-a_3 > 0$, $a_3 > 0$. Para $z^4 + a_1z^3 + a_2z^2 + a_3z + a_4$ es $a_1 > 0$, $a_1a_2 - a_3 > 0$, $(a_1a_2 - a_3) - a_1^2a_4 > 0$, $a_4 > 0$. Aplicando al ejemplo anterior, $F(z) = z^3 + 11z^2 + 6z + 6$.
Se tiene $a_1 = 11 > 0$, $a_1a_2 - a_3 = 11 \times 6 - 6 > 0$, $a_3 = 6 > 0$, y, así, obtenemos el mismo resultado que antes.

2.3.1. Ejemplos

Ejemplo 2.3.2. *Estúdiese la estabilidad de*

$$\frac{d^3}{dt^3}x + 8\frac{d^2}{dt^2}x + 14x + 24 = 0$$

Solución. *Reescribiendo la ecuación diferencial lineal dada se tiene que*

$$x''' + 8x'' + 14x + 24 = 0$$

el polinomio característico a esta ecuación es

$$F(z) = z^3 + 8z^2 + 14$$

si suponemos que $F(z)$ tiene coeficientes reales y no posee ningún cero imaginario puro, tenemos que

$$F(iy) = A(y) + iB(y) = (a_n - a_{n-2}y^2 + a_{n-4}y^4 - \cdots) + i(a_{n-1}y - a_{n-3}y^3 + \cdots) \neq 0$$

En nuestro caso $a_0 = 1$, $a_1 = 8$, $a_2 = 0$, $a_3 = 14$. Entonces

$$F(iy) = A(y) + iB(y) = (14 - 8y^2) + i(0 - 8y^3)$$

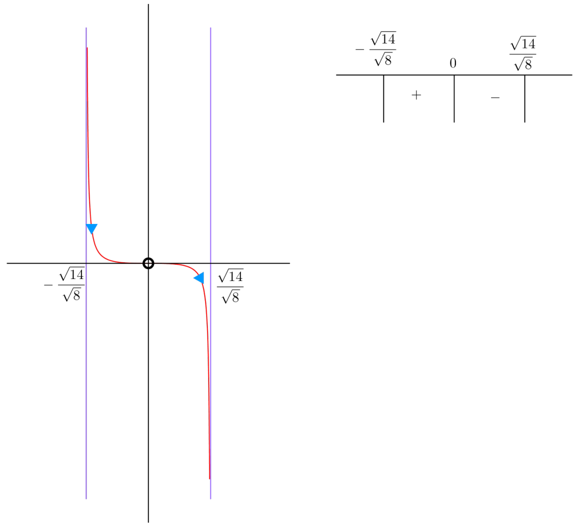


Figura 2.9: Curva de $L = \dfrac{B(y)}{A(y)}$

La curva

$$\frac{B(y)}{A(y)} = \frac{-8y^3}{14-8y^2} = \frac{-8y^3}{(\sqrt{14}-\sqrt{8}y)(\sqrt{14}+\sqrt{8}y)}$$

así, se puede notar que sólo hay una variación, entonces tenemos que $L = \pi$, *donde* $L = \dfrac{B(y)}{A(y)}$ *cuando* y *varía de* ∞ *a* $-\infty$. *Por lo que la ecuación diferencial lineal dada no es asintóticamente estable.*
También lo podemos ver a partir del criterio de Routh-Hurwitz. Ya que este criterio lo que nos dice es que todos los ceros de $F(z)$ *tienen parte real negativa si y sólo si todos los determinantes de las matrices diagonales principales son positivos.*
En efecto, para nuestro caso se tiene la matriz

$$\begin{pmatrix} 8 & 1 & 0 \\ 14 & 0 & 8 \\ 0 & 0 & 14 \end{pmatrix}$$

así, $a_1 = 8 > 0$, $a_1a_2 - a_3 = 0 - 14 = -14 < 0$ *por lo que no cumple con las condiciones del teorema. Por tanto no es asintóticamente estable.*

Ejemplo 2.3.3. *Para qué valores de* $k \in \mathbf{R}$ *es estable el sistema que tiene por ecuación característica*

$$\lambda^2 + k\lambda + 2k - 1 = 0$$

Solución. *Resolviendo la cuadrática en términos de* λ.

$$\lambda = \frac{-k \pm \sqrt{k^2 - 4(2k-1)}}{2}$$

$$\lambda_1 = \frac{-k + \sqrt{k^2 - 8k + 4}}{2}, \quad \lambda_2 = \frac{-k - \sqrt{k^2 - 8k + 4}}{2}$$

Resolviendo la cuadrática en términos de k

$$\begin{aligned} k &= \frac{8 \pm \sqrt{64-16}}{2} = \frac{8 \pm \sqrt{48}}{2} \\ k_1 &= 4 + 2\sqrt{3}, \quad k_2 = 4 - 2\sqrt{3} \end{aligned}$$

Veamos que al sustituir los valores de k_1 y k_2 en λ_1 y λ_2 se obtiene que las raices son

$$\begin{aligned} \lambda_1 &= \frac{-k_1}{2} = \frac{-4-2\sqrt{3}}{2} = -2 - \sqrt{3} < 0 \\ \lambda_1 &= \frac{-k_2}{2} = \frac{-4+2\sqrt{3}}{2} = -2 + \sqrt{3} < 0 \end{aligned}$$

lo mismo sucede para λ_2 em ambos casos.

Ahora bien, veamos que sucede para los valores que puede tomar k.

- *Si $k > k_1 = 4 + 2\sqrt{3} \approx 7.46$ entonces $\lambda_1 < 0$ y $\lambda_2 < 0$. Es decir las dos raíces son negativas.*
- *Si $k < k_2 = 4 - 2\sqrt{3} \approx 0.53$ entonces $\lambda_1 < 0$ y $\lambda_2 > 0$. Es decir en este caso tendremos raíces positivas y negativas.*
- *Si $k_2 < k < k_1$, entonces $\lambda_{1,2} \in \mathbf{C}$.*
  Por lo tanto, los valores de $k \in \mathbf{R}$ que hacen que el sistema sea estable son $k = 4 - 2\sqrt{3}$ y $k = [4 + 2\sqrt{3}, +\infty)$ y también los valores de $[0.5, 4 - 2\sqrt{3}]$ ya que si $k = 0.5$ tenemos que $\lambda_1 = 0$ y $\lambda_2 = -1$ y sabemos que si las raíces son cero es estable o si son negativas es asintóticamente estable.

2.4. Estabilidad de soluciones de ecuaciones no lineales

Como hemos visto, el estudio de la estabilidad de una ecuación lineal

$$x'(t) = A(t)x(t) + b(t)$$

Es problema particularmente simple cuando $A(t) \equiv A$ es matriz constante. El siguiente teorema, esencialmente debido a **LYAPUNOV**, se refiere a una reducción a este caso del estudio de la estabilidad de un problema todavía lineal, pero no de coeficientes constantes. El método por el cual aquí se resuelve, la utilización del lema de Gronwall, es muy fecundo y sirve también para tratar algunos casos de estabilidad de ecuaciones no lineales, como se indica a continuación.

Lema 2.4.1 (Lema de Gronwall)**.** *Sean u, v funciones continuas no negativas en el intervalo $J \subseteq \mathbf{R}$, $t_0 \in J$. Supongamos, que C es una constante real. Si se cumple*

$$u(t) \leqslant C + \left| \int_{t_0}^{t} u(s)v(s)ds \right|$$

Entonces

$$u(t) \leqslant C \exp\left(\left| \int_{t_0}^{t} v(s)ds \right| \right)$$

para cada $t \in J$.

Teorema 2.4.1. *Se considera la ecuación*

$$x'(t) = Ax(t) + C(t)x(t) \tag{2.4.1}$$

donde A es una matriz $n \times n$ constante de elementos reales y $C : [0, \infty) \to M(n, \boldsymbol{R})$ es una función matricial continua.
Sea $a = \max\{Re\lambda_i : \lambda_i \in \sigma(A)\}$ y sea $x(t)$ la solución de (2.4.1) en $[0, \infty)$ tal que $x(0) = x^0$.
Entonces, para todo $b > a$ existe una constante k que depende de b y de A solamente tal que para todo $t \geqslant 0$ se tiene:

$$|x(t)| \leqslant k|x^0| \exp\left(bt + k\int_0^t \|C(s)\|ds\right)$$

Demostración. Sabemos que la solución $x(t)$ de (2.4.1) tal que $x(0) = x^0$ existe y está definida en $[0, \infty)$. Llamemos $b(t) = C(t)x(t)$. Entonces $x(t)$ es la solución del problema

$$\begin{cases} x'(t) = Ax(t) + b(t) \\ x(0) = x^0 \end{cases}$$

y, así, aplicando la fórmula de Lagrange, resulta:

$$\begin{aligned}
x(t) &= \Phi(t)\xi^0 + \Phi(t)\int_{t_0}^t \Phi^{-1}(s)b(s)ds \quad \Phi(t) = e^{A(t-t_0)} \\
&= e^{A(t-t^0)}x^0 + \int_{t_0}^t e^{A(t-t_0)}[e^{A(s-t_0)}]^{-1}b(s)ds = e^{A(t-t^0)}x^0 + \int_{t_0}^t e^{A(t-t_0)}e^{A(t_0-s)}b(s)ds \\
&= e^{A(t-t^0)}x^0 + \int_{t_0}^t e^{A(t-s)}b(s)ds \quad t_0 = 0 \\
&= e^{At}x^0 + \int_0^t e^{A(t-s)}b(s)ds \quad b(t) = C(t)x(t) \\
&= e^{At}x^0 + \int_0^t e^{A(t-s)}C(s)x(s)ds
\end{aligned}$$

del corolario 2.2.2 resulta fácilmente que para $b > a$ existe k tal que

$$\|e^{At}\| \leqslant ke^{bt} \tag{2.4.2}$$

para todo $t \geqslant 0$. En efecto, sabemos que existe un polinomio $q(t)$ tal que

$$\|e^{At}\| \leqslant q(t)e^{at} = q(t)e^{(a-b)}e^{bt} \leqslant ke^{bt}$$

si se elige $k \geqslant q(t)e^{(a-b)t}$ para todo $t \geqslant 0$, lo que se puede hacer, ya que $q(t)e^{(a-b)t} \to 0$, para $t \to \infty$.

Por tanto, podemos escribir:

$$\begin{aligned}
|x(t)| &\leqslant \|e^{At}\||x^0| + \int_0^t \|e^{A(t-s)}\|\|C(s)\||x(s)|ds \quad \textit{por la ecuación } 2.4.2 \\
&\leqslant ke^{bt}|x^0| + \int_0^t ke^{b(t-s)}\|C(s)\||x(s)|ds \\
&= ke^{bt}|x^0| + \int_0^t e^{bt}e^{-bs}\|C(s)\||x(s)|ds \\
&= e^{bt}\left(k|x^0| + \int_0^t e^{-bs}\|C(s)\||x(s)|ds\right) \\
e^{-bt}|x(t)| &\leqslant k|x^0| + \int_0^t e^{-bs}\|C(s)\||x(s)|ds
\end{aligned}$$

ahora hacemos $u(t) = e^{-bt}|x(t)|$. Aplicando el lema de Gronwall, resulta la desigualdad del teorema.

$$\begin{aligned}
e^{-bt}|x(t)| &\leqslant k|x^0|e^{k\int_0^t \|C(s)\|ds} \\
|x(t)| &\leqslant k|x^0|e^{bt}e^{k\int_0^t \|C(s)\|ds} \\
|x(t)| &\leqslant k|x^0|e^{bt+k\int_0^t \|C(s)\|ds}
\end{aligned}$$

□

Del teorema anterior se obtiene inmediatamente el corolario siguiente.

Corolario 2.4.1. *1) Sea $x'(t) = Ax(t)$ asintóticamente estable y sea*

$$bt + k\int_0^t \|C(s)\|ds \to -\infty \quad para \quad t \to \infty \tag{2.4.3}$$

entonces la ecuación $x'(t) = (A + C(t))x(t)$ es también asintóticamente estable.

2) Sea $x'(t) = Ax(t)$ estable y sea

$$\int_0^\infty \|C(s)\|ds < \infty$$

Entonces la ecuación $x'(t) = (A + C(t))x(t)$ es estable.

Obsérvese que si

$$\int_0^\infty \|C(s)\|ds < \infty$$

entonces, eligiendo $b < 0$, ya se tiene la condición (2.4.3). Asimismo, si $\|C(s)\| \to 0$ para $s \to \infty$ se tiene también (2.4.3), eligiendo $b < 0$. En efecto, veamos en primer lugar que si $\|C(s)\| \to 0$ para $s \to \infty$, entonces

$$(1/t)\int_0^t \|C(s)\|ds \to 0 \quad para \quad t \to \infty$$

Sea $\epsilon > 0$. Elegimos t tal que

$$\|C(s)\| \leqslant \frac{\epsilon}{2} \quad para \quad t \geqslant t_0$$

Entonces, si $t \geqslant t_0$, se puede poner:

$$\begin{aligned}
\frac{1}{t}\int_0^t \|C(s)\|ds &= \frac{1}{t}\left[\int_0^{t_0}\|C(s)\|ds + \int_{t_0}^t \|C(s)\|ds\right] = \frac{1}{t}\int_0^{t_0}\|C(s)\|ds + \frac{1}{t}\int_{t_0}^t \|C(s)\|ds \\
&\leqslant \frac{1}{t}\int_0^{t_0}\|C(s)\|ds + \frac{1}{t}\frac{\epsilon}{2}\int_{t_0}^t ds = \frac{1}{t}\int_0^{t_0}\|C(s)\|ds + \frac{t-t_0}{t}\frac{\epsilon}{2} \\
&\leqslant \frac{1}{t}\int_0^{t_0}\|C(s)\|ds + \frac{\epsilon}{2}1 \leqslant \frac{1}{t}\int_0^{t_0}\|C(s)\|ds + \frac{\epsilon}{2}
\end{aligned}$$

Si $t_1 \geqslant t_0$ es tal que

$$\frac{1}{t}\int_0^{t_0}\|C(s)\|ds \leqslant \frac{\epsilon}{2}$$

y tomamos $t \geqslant t_1$,entonces:

$$\frac{1}{t}\int_0^{t}\|C(s)\|ds \leqslant \epsilon$$

lo que demuestra que

$$\frac{1}{t}\int_0^{t}\|C(s)\|ds \to 0$$

para $t \to \infty$. Si se escribe

$$bt + k\int_0^t \|C(s)\|ds = t\left(b + k\frac{1}{t}\int_0^t \|C(s)\|ds\right)$$

con $b < 0$, se obtiene fácilmente (2.4.3).

El teorema siguiente se obtiene por el mismo procedimiento que el anterior. Las condiciones que se imponen son las adecuadas para poder aplicar el lema de Gronwall. Sin embargo, no siendo la ecuación que se estudia estrictamente lineal, la existencia de solución en $[0, \infty)$ no está asegurada de antemano con tanta facilidad como en el caso anterior.

Teorema 2.4.2. *Se condidera la ecuación*

$$x'(t) = A(t)x(t) + f(t, x(t)) \tag{2.4.4}$$

donde $A : [0, \infty) \to M(n, \boldsymbol{R})$ *es continua y* $f : [0, \infty) \times \bar{B}(0, a) \to \boldsymbol{R}^n$ *(donde* $\bar{B}(0, a)$ *es una bola euclídea cerrada de centro el origen y radio* a*) es también continua.*

Sea $Y(t)$ *matriz fundamental de* $y'(t) = A(t)y(t)$ *tal que* $Y(0) = I$*. Supongamos que existe* $k > 0$ *tal que*

$$\|Y(t)Y^{-1}(s)\| \leqslant k \quad \textit{para} \quad t \geqslant s \geqslant 0$$

Supongamos también que existe una función $\gamma(t)$ *continua y no negativa definida en* $[0, \infty)$ *tal que*

$$|f(t, \xi)| \leqslant \gamma(t)|\xi| \quad (t, \xi) \in [0, \infty) \times \bar{B}(0, a)$$

y tal que

$$\int_0^\infty \gamma(t)dt < \infty$$

Entonces existe una constante positiva $h \geqslant 1$ tal que si $t_1 \in [0,\infty)$ y $x(t)$ es solución local de (2.4.4) a la derecha de t_1 tal que

$$|x(t_1)| < h^{-1}a,$$

se verifica que $x(t)$ es prolongable a $[t_1,\infty)$, y cualquier prolongación a $[t_1,\infty)$ satisface

$$|x(t_1)| < h|x(t_1)| \quad \textit{para} \quad \textit{todo} \quad t \geqslant t_1$$

Si además se tiene $Y(t) \to 0$ para $t \to \infty$, entonces para la $x(t)$ anterior definida en $[t_1,\infty)$ se verifica $x(t) \to 0$ para $t \to \infty$.

Demostración. Sea $x(t)$ solución local de (2.4.4) a la derecha de t_1, y denotemos también por $x(t)$ no prolongable estrictamente a la derecha. Supongamos que ésta no está definida en $[t_1,\infty)$. El lema de Wintner excluye que esté definida en $[t_1,t_2)$ con $t_2 < \infty$. Así, podemos suponer que está definida $[t_1,t_2]$. En este intervalo la función $x(t)$ es la solución del problema lineal no homogéneo

$$(P)\begin{cases} y'(t) = A(t)x(t) + b(t) \\ y(t_1) = x(t_1) \end{cases}$$

siendo $b(t) = f(t,x(t))$. Por tanto, por la fórmula de Lagrange se obtiene:

$$x(t) = Y(t)Y^{-1}(t_1)x(t_1) + \int_{t_1}^{t} Y(t)Y^{-1}(s)f(s,x(s))ds$$

para $t \in [t_1,t_2]$.
Tomando normas resulta,

$$\begin{aligned} |x(t)| &= \left| Y(t)Y^{-1}(t_{-1})x(t_1) + \int_{t_1}^{t} Y(t)Y^{-1}(t_1)f(s,x(s))ds \right| \\ &\leqslant \left| Y(t)Y^{-1}(t_1)x(t_1) \right| + \left| \int_{t_1}^{t} Y(t)Y^{-1}(t_1)f(s,x(s))ds \right| \\ &\leqslant \underbrace{\|Y(t)Y^{-1}(t_1)\|}_{k}|x(t_1)| + \int_{t_1}^{t} \underbrace{\|Y(t)Y^{-1}(t_1)\|}_{k}\underbrace{|f(s,x(s))|}_{\gamma(s)|x(s)|}ds \quad \textit{por las hipótesis} \\ &\leqslant k|x(t_1)| + k\int_{t_1}^{t} \gamma(s)|x(s)|ds \end{aligned}$$

es decir,

$$|x(t)| \leqslant k|x(t_1)| + k\int_{t_1}^{t} \gamma(s)|x(s)|ds$$

para $t \in [t_1,t_2]$, y, por el lema de Gronwall:

$$|x(t)| \leqslant k|x(t_1)| \exp\left(k\int_{t_1}^{t} \gamma(s)ds\right)$$

Sea

$$h = \text{máx}\left(1, k\exp\left(k\int_{t_1}^{t} \gamma(s)ds\right)\right)$$

Entonces $|x(t)| \leqslant h|x(t_1)|$ para $t \in [t_1, t_2]$, y, por la condición dada también se tiene, $|x(t_2)| \leqslant h|x(t_1)|$ así, $|x(t_2)| < a$, lo que implica que $x(t)$ es estrictamente prolongable a la derecha de t_2, en contra de la hipótesis de no prolongabilidad. Así, $x(t)$ está definida en $[t_1, \infty)$. Para cualquier $t \geqslant t_1$ se puede repetir el razonamiento que nos ha conducido a

$$|x(t)| \leqslant h|x(t_1)|,$$

y esto concluye la primera parte del teorema.
Para la segunda parte se tiene, para $t \geqslant t_1$:

$$x(t) = Y(t)Y^{-1}(t_1)x(t_1) + \int_{t_1}^{t} Y(t)Y^{-1}(s)f(s, x(s))ds$$

Si se cumple que $Y(t) \to 0$ para $t \to \infty$ se verifica

$$Y(t)Y^{-1}(t_1)x(t_1) \to 0 \quad para \quad t \to \infty$$

También, escogiendo $t_1 \leqslant t_2 \leqslant t$ y aplicando normas se tiene para la integral:

$$\begin{aligned}
\left|\int_{t_1}^{t} Y(t)Y^{-1}(s)f(s, x(s))ds\right| &\leqslant \left|Y(t)\int_{t_1}^{t_2} Y^{-1}(s)f(s, x(s))ds\right| + \left|\int_{t_2}^{t} Y(t)Y^{-1}(s)f(s, x(s))ds\right| \\
&\leqslant \|Y(t)\|\left|\int_{t_1}^{t_2} Y^{-1}(s)f(s, x(s))ds\right| + \int_{t_1}^{t} \underbrace{\|Y(t)Y^{-1}(t_1)\|}_{k}|f(s, x(s))|ds \\
&\leqslant \|Y(t)\|\left|\int_{t_1}^{t_2} Y^{-1}(s)f(s, x(s))ds\right| + k\int_{t_2}^{\infty} \gamma(s)|x(s)|ds \\
&= S_1 + S_2
\end{aligned}$$

Por hipótesis se sabe que $\int_{t_2}^{\infty} \gamma(t)dt < \infty$ es decir que la integral converge, y que la función es continua y no negativa, es decir que se puede acotar superiormente.
Dado $\epsilon > 0$ podemos escoger t_2 grande de modo que

$$S_2 = k\int_{t_2}^{\infty} \gamma(s)|x(s)|ds \leqslant kh|x(t_1)|\int_{t_2}^{\infty} \gamma(s)ds$$

esto es cierto por lo demostrado en la primera parte que $|x(s)| \leqslant h|x(t_1)|$ para $s \geqslant t \geqslant 0$ así

$$S_2 \leqslant kh|x(t_1)|\int_{t_2}^{\infty} \gamma(s)ds \leqslant \frac{\epsilon}{2}$$

y, una vez fijado t_2, de este modo se tiene, para todo t suficientemente grande, que $S_1 \leqslant \epsilon/2$. Así $\left|\int_{t_1}^{t} Y(t)Y^{-1}(s)f(s, x(s))ds\right| < \epsilon$. Por tanto, se obtiene que $|x(t)| \to 0$ para $t \to \infty$. □

El corolario siguiente es consecuencia sencilla del teorema.

Corolario 2.4.2. *Se considera la ecuación (2.4.4) del teorema anterior. Se supone, como antes, que*

$$\|Y(t)Y^{-1}(s)\| \leqslant k < \infty \quad para \quad 0 \leqslant s \leqslant t$$

y que

$$|f(t,\xi)| \leqslant \gamma(t)|\xi| \quad con \quad \int_0^\infty \gamma(t)dt < \infty$$

Entonces la solución $x(t) \equiv 0$ de (2.4.4) satisface la siguiente condición de estabilidad (estabilidad uniforme). Para cada $\epsilon > 0$ existe $\delta = \delta(\epsilon) > 0$ tal que si $\bar{x}(t)$ es solución local de (2.4.4) a la derecha de $t_1 \geqslant 0$ con $|\bar{x}(t)| \leqslant \delta$ entonces $\bar{x}(t)$ es prolongable a $[t_1,\infty)$ y satisface $|\bar{x}(t)| \leqslant \epsilon$ para $t \in [t_1,\infty)$.

A fin de demostrar el teorema siguiente, relativo a la estabilidad asintótica de la solución nula de (2.4.4), presentamos primero un lema interesante también por sí mismo.

Lema 2.4.2. *Sea $Y : [0,\infty) \to M(n,\boldsymbol{R})$ continua y tal que $Y(t)$ es no singular para todo $t \in [0,\infty)$. Supongamos que existe $k > 0$ tal que*

$$\int_0^t \|Y(t)Y^{-1}(s)\|ds \leqslant K \quad para \quad todo \quad t \in [0,\infty)$$

Entonces existe una constante $h > 0$ tal que

$$\|Y(t)\| \leqslant h \exp\left(-\frac{t}{k}\right) \quad para \quad t \geqslant 1$$

Demostración. Ponemos $a(t) = \|Y(t)\|^{-1}$ y la identidad $Y(t)Y(t)^{-1} = I$. Y además

$$|a(t)| = |\|Y(t)^{-1}\|| = a(t)$$

entonces se tiene que $|a(t)| = \|Y(t)\|^{-1}$ y así $\|Y(t)\||a(t)| = I$, entonces $\|Y(t)\| = |a(t)|^{-1}$ y podemos escribir la identidad

$$Y(t)\int_0^t a(s)ds = \int_0^t Y(t)Y^{-1}(s)Y(s)a(s)ds$$

tomando normas y ocupando propiedades de normas resulta

$$\begin{aligned}
\left\|Y(t)\int_0^t a(s)ds\right\| &= \|Y(t)\|\left|\int_0^t a(s)ds\right| = \|Y(t)\|\int_0^t |a(s)|\,ds \\
&= |a(t)|^{-1}\int_0^t a(s)ds = \left\|\int_0^t Y(t)Y^{-1}(s)Y(s)a(s)ds\right\| \\
&\leqslant \int_0^t \|Y(t)Y^{-1}(s)\|\underbrace{\|Y(s)\|a(s)}_{I}ds = \int_0^t \|Y(t)Y^{-1}(s)\|ds \\
&\leqslant k \quad por \quad hipótesis
\end{aligned}$$

Así se tiene la siguiente desigualdad

$$|a(t)|^{-1}\int_0^t a(s)ds \leqslant k \tag{2.4.5}$$

para todo $t \in [0,\infty)$. Si escribimos

$$b(t) = \int_0^t a(s)ds$$

y derivando $b(t)$ se tiene:

$$b'(t) = \frac{d}{dt}\int_0^t a(s)ds = a(t)$$

entonces, sustituyendo $b(t)$ en la ecuación (2.4.5) se tiene:

$$\frac{b(t)}{|a(t)|} \leqslant k$$

ahora pasando $|a(t)|$ a multiplicar al otro lado de la desigualdad

$$b(t) \quad \leqslant \quad |a(t)|K = a(t)k$$

y sustituyendo $b'(t) = a(t)$ obtenemos

$$b(t) \leqslant b'(t)k$$

despejando k de la desigualdad

$$\frac{1}{k} \leqslant \frac{b'(t)}{b(t)}$$

o lo que es equivalente

$$\frac{b'(t)}{b(t)} \geqslant k^{-1}$$

integrando entre 1 y t se obtiene:

$$\int_1^t \frac{b'(s)}{b(s)}ds \geqslant \int_1^t k^{-1}ds \tag{2.4.6}$$

integrando cada una se tiene

$$\begin{aligned}\int_1^t \frac{b'(s)}{b(s)}ds &= \ln(b(t))\Big|_1^t \\ &= \ln(b(t)) - \ln(b(1)) = \ln\left(\frac{b(t)}{b(1)}\right)\end{aligned}$$

ahora con la otra integral

$$\int_1^t k^{-1}ds \quad = \quad k^{-1}t\Big|_1^t = k^{-1}\,(t-1)$$

volviendo a la ecuación (2.4.6) y sustituyendo los resultados de las integrales

$$\ln\left(\frac{b(t)}{b(1)}\right) \geqslant k^{-1}\,(t-1)$$

Aplicando la exponencial a ambos lados de la desigualdad se tiene:

$$\frac{b(t)}{b(1)} \geqslant \exp\left(k^{-1}\,(t-1)\right)$$

como teniamos que $ka(t) \geqslant b(t)$ entonces podemos concluir que

$$\frac{ka(t)}{b(1)} \geqslant \frac{b(t)}{b(1)} \geqslant \exp\left(k^{-1}\,(t-1)\right)$$

de aquí se tiene

$$\frac{ka(t)}{b(1)} \geqslant \exp\left(k^{-1}(t-1)\right)$$

$$\frac{k|a(t)|}{b(1)} \geqslant \exp\left(k^{-1}(t-1)\right)$$

$$\frac{k}{b(1)} \geqslant \frac{\exp\left(k^{-1}(t-1)\right)}{|a(t)|}$$

$$\frac{k}{b(1)} \geqslant |a(t)|^{-1}\exp\left(k^{-1}(t-1)\right)$$

sustituyendo $|a(t)|^{-1} = \|Y(t)\|$ resulta

$$\frac{k}{b(1)} \geqslant \|Y(t)\|\exp\left(k^{-1}(t-1)\right)$$

y que es equivalente

$$\|Y(t)\|\exp\left(k^{-1}(t-1)\right) \leqslant \frac{k}{b(1)}$$

$$\|Y(t)\| \leqslant \frac{k}{b(1)}\frac{1}{\exp\left(k^{-1}(t-1)\right)}$$

$$\|Y(t)\| \leqslant \frac{k}{b(1)}\exp\left(-\frac{t}{k}+\frac{1}{k}\right)$$

$$\|Y(t)\| \leqslant \frac{k}{b(1)}\exp\left(k^{-1}\right)\exp\left(-\frac{t}{k}\right)$$

tomando $h = \dfrac{k}{b(1)}\exp\left(k^{-1}\right)$ por lo tanto:

$$\|Y(t)\| \leqslant h\exp\left(-\frac{t}{k}\right)$$

□

Teorema 2.4.3. *Se considera nuevamente la ecuación*

$$x'(t) = A(t)x(t) + f(t, x(t)) \tag{2.4.7}$$

teniendo $A(t)$ *y* $f(t,\xi)$ *la significación del teorema (2.4.2).*
Supongamos que existe $k > 0$ *tal que*

$$\int_0^t \|Y(t)Y^{-1}(s)\|ds \leqslant k \quad para \quad t \in [0,\infty)$$

Sea $\gamma(t)$ *constante y menor que* k^{-1}. *Entonces la solución* $\bar{x}(t) \equiv 0$ *de (2.4.7) es asintóticamente estable.*

Demostración. Por el lema anterior $Y(t) \to 0$ para $t \to \infty$ y, así, $\|Y(t)\| \leqslant M$ para cierta $M < \infty$ y todo $t \in [0, \infty)$. Por la fórmula de Lagrange, si $x(t)$ es una solución local de (2.4.7) a la derecha de $t = 0$, se tiene:

$$x(t) = Y(t)x(0) + \int_0^t Y(t)Y^{-1}(s)f(s, x(s))ds$$

para t en el intervalo de definición de $x(t)$ a la derecha de $t = 0$. Así, por las hipótesis:

$$|x(t)| = M|x(0)| + \gamma k \sup |x(s)| : s \in [0, t]$$

por tanto :

$$\sup |x(s)| : s \in [0, t] \leqslant (1 - \gamma k)^{-1} M|x(0)|$$

con lo cual se demuestra fácilmente que la solución $x(t)$ se puede prolongar a $[0, \infty)$ si $|x(0)|$ es suficientemente pequeño, y que la solución $\bar{x}(t) \equiv 0$ es estable. Tratemos de ver que $x(t) \to 0$ para $t \to \infty$, con lo cual quedará probada la estabilidad asintótica de $\bar{x}(t) \equiv 0$. Sea

$$\mu = \overline{\lim_{t \to \infty}} \ |x(t)|$$

y escojamos τ tal que $\gamma k < \tau < 1$. Si fuese $\mu > 0$, entonces $\mu\tau^{-1} > \mu$, y, así, existe t_1 tal que $|x(t)| < \mu\tau^{-1}$ para todo $t \geqslant t_1$. De la fórmula de Lagrange utilizada antes, dividiendo la integral en dos partes y tomando normas, resulta:

$$|x(t)| \leqslant \|Y(t)\||x(0)| + \|Y(t)\| \left| \int_0^{t_1} Y^{-1}(s)f(s, x(s))ds \right| + \gamma k \mu \tau^{-1} \quad para \quad t \geqslant t_1$$

si hacemos $t \to \infty$ y tomamos el límite superior resulta, por ser $\|Y(t)\| \to 0$ para $t \to \infty$:

$$\mu \leqslant \gamma k \mu \tau^{-1} < \mu$$

lo que es contradictorio. Así, ha de ser $\mu = 0$, lo cual demuestra el teorema. □

Capítulo 3

Teoría de estabilidad. El método directo de Lyapunov

El problema de la estabilidad de una solución $\bar{y}(t)$ de una ecuación diferencial

$$y'(t) = g(t, y(t)) \tag{3.0.1}$$

queda reducido, mediante el cambio

$$y(t) = \bar{y}(t) + x(t) \tag{3.0.2}$$

al problema de la estabilidad de la solución trivial $x(t) \equiv 0$ de una ecuación diferencial

$$x'(t) = f(t, x(t)) \tag{3.0.3}$$

donde $f(t, 0) \equiv 0$.

Veamos como es que se pasa de la ecuación (3.0.1) a la ecuación (3.0.3). Como $\bar{y}(t)$ es solución de la ecuación diferencial $y'(t) = g(t, y(t))$, también se cumple que $\bar{y}'(t) = g(t, \bar{y}(t))$. Despejando $x(t)$ de la ecuación (3.0.2) se tiene que

$$\begin{aligned} x(t) &= y(t) - \bar{y}(t) \\ \Rightarrow x'(t) &= y'(t) - \bar{y}'(t) = g(t, y(t)) - g(t, \bar{y}(t)) \\ \Rightarrow x'(t) &= g(t, \bar{y}(t) + x(t)) - g(t, \bar{y}(t)) = f(t, x(t)) \end{aligned}$$

El método directo o segundo método de Lyapunov, permite resolver este problema sin necesidad de conocer la forma general de la solución del problema

$$\begin{cases} x'(t) = f(t, x(t)) \\ x(t_0) = \xi_0 \end{cases}$$

y de ahí proviene su nombre. Para ello, motivado por consideraciones mecánicas, Lyapunov introdujo una cierta función llamada ahora función de Lyapunov, que por sus relaciones con la función $f(t, \xi)$, ecuación diferencial que se estudia, proporciona información sobre la estabilidad de la solución trivial. En este capítulo se exponen en primer lugar, en las secciones 3.1, 3.2 y 3.3, algunos teoremas sobre estabilidad, estabilidad asintótica, estabilidad uniforme e inestabilidad, que constituyen los resultados básicos de la teoría. Presentamos aquí como muestra un resultado sobre lo que se denomina ***problema inverso en la teoría del método directo***, de gran actualidad, sobre todo en la escuela rusa, y que consiste en determinar hasta qué punto la estabilidad, estabilidad asintótica, etc., son equivalentes a la existencia de una función de Lyapunov adecuada.

La construcción de una función de Lyapunov para un problema dado no sigue ningún método universalmente válido, por eso presentamos a continuación en detalle la construcción de la función de Lyapunov para una ecuación lineal con coeficientes constantes en la sección 3.4.

El intenso interés que existe actualmente por el método directo de Lyapunov se debe sobre todo a sus aplicaciones importantes en la teoría de estabilidad de sistemas de control.

Muchos de los textos que se refieren a la estabilidad por segundo método de Lyapunov suelen exigir que en la ecuación

$$x'(t) = f(t, x(t))$$

la función $f(t, \xi)$ sea tal que la solución del problema

$$\begin{cases} x'(t) = f(t, x(t)) \\ x(t_0) = \xi_0 \end{cases}$$

cuando existe, sea única. Aquí se ha prescindido en general, de esta restricción y se ha impuesto solamente que $f(t, \xi)$ sea continua en un cierto dominio.

En el tratado de Lyapunov y en muchos de los textos con especial interés en la aplicación del método directo a los problemas de la teoría de control, la ecuación que se estudia es autónoma, es decir, del tipo

$$x'(t) = f(x(t))$$

3.1. Estabilidad y estabilidad uniforme

Definición 3.1.1. *La solución trivial $x(t) \equiv 0$ se dice estable en $[t_0, \infty)$ cuando para cada $\epsilon > 0$ existe $\delta > 0$ tal que cualquier solución local a la derecha del problema*

$$(P) \begin{cases} x'(t) = f(t, x(t)) \\ x(t_0) = \xi_0 \end{cases}$$

con $|\xi_0| \leqslant \delta$, esta definida en $[t_0, \infty)$ y verifica $|x(t)| \leqslant \epsilon$ para todo $t \geqslant t_0$.

Definición 3.1.2. *La solución trivial $x(t) \equiv 0$ se dice asintóticamente estable en $[t_0, \infty)$ cuando es estable en $[t_0, \infty)$ y además existe $\eta > 0$ tal que cualquier solución de (P) con $|\xi_0| \leqslant \eta$ verifica $|x(t)| \to 0$ para $t \to +\infty$.*

Consideremos en general en esta sección una ecuación $x'(t) = f(t, x(t))$, donde

$$f : [t_0, \infty) \times G \subset \mathbf{R} \times \mathbf{R^n} \to \mathbf{R^n} \tag{3.1.1}$$

(G entorno abierto del origen en $\mathbf{R^n}$) es una función continua en su dominio de definición y $f(t, 0) \equiv 0$ para todo $t \geqslant t_0$.
En el contexto del problema inverso en el método directo de Lyapunov es particularmente importante la noción de estabilidad uniforme que definimos a continuación.

Definición 3.1.3. *La solución trivial de $x'(t) = f(t, x(t))$ se dice **uniformemente estable** en $[t_0, \infty)$ cuando para cada $\epsilon > 0$ tal que $\bar{B}(0, \epsilon) \subset G$ existe $\delta > 0$ tal que si $x(t)$ es solución local a la derecha del problema*

$$(P) \begin{cases} x'(t) = f(t, x(t)) \\ x(t_1) = \xi_1 \end{cases}$$

con $t_1 \geqslant t_0$, $|\xi_1| \leqslant \delta$, se tiene entonces $|x(t)| \leqslant \epsilon$ para todo $t \geqslant t_1$.

La noción de estabilidad uniforme tiene también un claro interés práctico. De un modo vago significa que las perturbaciones pequeñas, no sólo en el instante inicial, sino también en cualquier instante posterior, no tiene mucha importancia.
Si la solución trivial es uniformemente estable en $[t_0, \infty)$ es claro que es estable, pero la implicación contraria es falsa.
El primer teorema que presentamos proporciona una condición suficiente de estabilidad para la solución trivial.

Teorema 3.1.1. *Consideramos la ecuación $x'(t) = f(t, x(t))$, donde f es una función tal como se ha indicado anteriormente.*
Supongamos que existe una función

$$V : [t_0, \infty) \times \bar{B}(0, r) \subset \mathbf{R} \times G \to [0, \infty)$$

tal que

a) $V(t_0, \xi) \to 0$, *para* $|\xi| \to 0$.

b) $V(t, \xi) \geqslant a\,(|\xi|)$ *para todo* $t \in [t_0, \infty)$, $|\xi| \leqslant r$, *siendo* $a : [0, r] \to [0, \infty)$ *una función continua, estrictamente creciente y tal que* $a(0) = 0$.

c) *Para toda solución local a la derecha $x(t)$ de $x'(t) = f(t, x(t))$ tal que $|x(t_0)| \leqslant r$ se verifica que la función $V^*(t) = V(t, x(t))$ es función no creciente de t allí donde esta definida.*

Entonces la solución trivial es estable en $[t_0, \infty)$.

Demostración. La idea geométrica de la demostración es sencilla para $n = 2$. El razonamiento analítico es el mismo para cualquier dimensión.
V es una función auxiliar y para $n = 2$ se tiene que

$$V : [t_0, \infty) \times \bar{B}(0, r) \subset \mathbf{R} \times G \subseteq \mathbf{R} \times \mathbf{R}^2 \to [0, \infty)$$

es decir que $\bar{B}(0, r) \subset G \subseteq \mathbf{R}^2$. Representemos gráficamente para t fijo (t no se muestra en la figura), $t \geqslant t_0$, los puntos $z = V(t, \xi)$, siendo $\xi = (x, y)$. Por la condición b) se tiene que $|\xi| = |(x, y)|$ es una circunferencia de radio menor o igual que r; también se nos dice que la función a se aplica a $|\xi|$ y esta aplicación lo que nos da son las imágenes de las circunferencias $|\xi|$ con su respectivo radio, además cada punto de la circunferencia tiene una imagen que también es un punto. Ahora bien, como decimos que cada punto de la circunferencia tiene un punto imagen podemos unir dicho punto con su respectiva imagen y así se obtiene el cilindro que se observa en la figura. Por otro lado dependiendo de como sea el radio de la circunferencia así es la imagen entonces todas estas imágenes vienen a ser como curvas de nivel, y así se obtiene el paraboloide que se muestra también en la figura y dicho paraboloide es lo que se obtiene de hacer $a(|\xi|)$. También podemos ver que a es continua pues no tiene salto y además que es estrictamente creciente, lo que nos dice $a(0) = 0$ es que al hacer $a(|\xi|)$ con $|\xi| = 0$ es el origen. Todos estos puntos $z = V(t, \xi)$ se mantienen, según la condición b), para cualquier $t \geqslant t_0$ por encima de la superficie de revolución $z = a(|\xi|)$ (ver la figura 3.1).
Dado $\epsilon > 0$ tal que $\bar{B}(0, \epsilon) \subset G$, sea $p = a(\epsilon) > 0$ (es decir en la figura tomamos que el radio de $|\xi|$ es ϵ y $p = a(\epsilon)$ es la distancia de la imagen de un punto de la circunferencia a dicho punto.) Como $V(t_0, \xi) \to 0$ para $|\xi| \to 0$ (esto lo que nos esta diciendo es que cuando se aplica el límite a $|\xi|$ se va acercando al origen entonces la función V también se acerca al origen), según la condición a), existe $\delta > 0$ tal que si $|\xi| \leqslant \delta$ se tiene

$$0 \leqslant V(t_0, \xi) < p$$

(Esto es cierto porque a es una función continua, entonces V no se puede salir de donde está definida la función.) Tomamos una solución local a la derecha de t_0 cualquiera $x(t)$ tal que $|x(t_0)| \leqslant \delta$. Si para algún $t_1 > t_0$ fuese $x(t_1) \geqslant \epsilon$, entonces

$$V^*(t_1) = V(t_1, x(t_1)) \geqslant a(|x(t_1)|) \geqslant p > V(t_0, x(t_0)) = V^*(t_0)$$

y, así, $V^*(t_1) > V^*(t_0)$ esto nos lleva a una contradicción ya que V^* es no creciente por c). Por tanto, $x(t) \leqslant \epsilon$ para todo $t \geqslant t_0$ en el que $x(t)$ está definida. Esto demuestra que x es prolongable a $[t_0, \infty]$ (porque al tomarnos un t que sea mayor que t_0 la solución no se sale siempre se mantiene en ese intervalo) y que la solución trivial es estable. □

Observación 1. *Si V es derivable respecto de t y ξ, entonces también lo es V^* respecto de t para una solución $x(t)$, y se tiene:*

$$\frac{dV^*}{dt}(t) = \frac{\partial V}{\partial t}(t, x(t)) + (gradV(t, x(t)), f(t, x(t)))$$

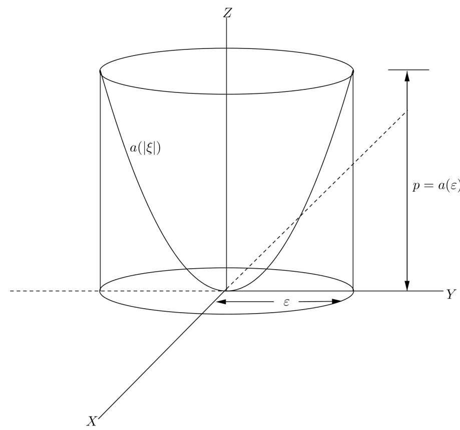


Figura 3.1: Representación geométrica del teorema

donde

$$gradV(t,x(t)) = \begin{pmatrix} \dfrac{\partial V}{\partial \xi_1}(t,x(t)) \\ \vdots \\ \dfrac{\partial V}{\partial \xi_n}(t,x(t)) \end{pmatrix}$$

y $(.,.)$ *representa el producto escalar en* $\mathbf{R^n}$*. La función del segundo miembro se suele denominar la derivada total de* V *respecto de la ecuación*

$$x'(t) = f(t,x(t))$$

Veamos que si el sistema es autónomo, es decir si $x'(t) = f(x(t))$*, podemos observemos que esta derivada puede pensarse como la derivada direccional de* $V(x)$ *en la dirección del campo* $f(x(t))$*, es decir*

$$V'(x) = \mathbf{D}V(x)\cdot x' = \sum_{i=1}^{n} \frac{\partial V}{\partial x_i}\cdot x' = \nabla V\cdot f(x)$$

Por lo tanto, si $V'(x)$ *es negativa,* V *decrecerá a lo largo de la solución de* $x'(t) = f(x(t))$*. Esto es lo que nos dice la condición c) del teorema.*

También la estabilidad uniforme se puede caracterizar de modo semejante, como se indica en el teorema que sigue.

Teorema 3.1.2. *Consideramos la ecuación* $x'(t) = f(t,x(t))$*, donde* f *es una función tal como se ha indicado anteriormente.*
Supongamos que existe una función

$$V : [t_0,\infty) \times \bar{B}(0,r) \subset \mathbf{R} \times G \to [0,\infty)$$

tal que

a) Existen dos funciones $a : [0, r] \to [0, \infty)$ *y* $b : [0, r] \to [0, \infty)$ *continuas, estrictamente crecientes, tales que*

$$a(0) = b(0), \quad a(|\xi|) \leqslant V(t, \xi) \leqslant b(|\xi|)$$

para todo $t \in [t_0, \infty)$, $|\xi| \leqslant r$.

b) Para toda solución local a la derecha $x(t)$ *de cada* $t \geqslant t_0$ *tal que* $|x(t_1)| \leqslant r$ *se verifica que la función* $V^*(t) = V(t, x(t))$ *es función no creciente de* t *a la derecha de* t_1 *allí donde está definida.*

Entonces, la solución trivial de $x'(t) = f(t, x(t))$ *es uniformemente estable en* $[t_0, \infty)$.

Demostración. La significación geométrica del razonamiento que sigue es clara con una imagen como la del teorema anterior. Para n=2. Como $\bar{B}(0, r) \subset G \subseteq \mathbf{R}^2$, y sea $\xi = (x, y)$ se tiene, entonces que $|\xi|$ es una circunferencia de radio menor o igual que r en el plano XY, $a(0) = b(0) = 0$ es el origen.
$a(|\xi|)$: representa la imagen de la circunferencia $|\xi| \leqslant r$, es un paraboloide, lo mismo que, $b(|\xi|)$.
$V(t, \xi) \geqslant a(|\xi|)$: significa que para $t \geqslant t_0$ los puntos $V(t, \xi)$ son mayores que el solido de revolución $a(|\xi|)$.
Para $b(|\xi|) \geqslant V(t, \xi)$: significa que para $t \geqslant t_0$ los puntos $V(t, \xi)$ estan por debajo del solido de revolución $b(|\xi|)$.
Sea $\epsilon > 0$ dado tal que

$$\bar{B}(0, \epsilon) \subset \bar{B}(0, r) \subset G$$

Tomamos $a(\epsilon) > 0$, y sea $\delta > 0$ tal que $a(\epsilon) = b(\delta)$ (Sabemos por a) que $a(|\xi|) \leqslant V(t, \xi) \leqslant b(|\xi|)$; es decir que $a(|\xi|) \leqslant b(|\xi|)$ esto es si elegimos una circunferencia de radio ϵ debe existir otra circunferencia de radio δ tal que en algún momento estas dos circunferencias son iguales).
Sea $t_1 \geqslant t_0$ y $x(t)$ una solución tal que $|x(t_1)| \leqslant \delta$, para ver que la función V^* es decreciente; suponemos que es estrictamente creciente para llegar a una contradicción con la condición b).
Si para $t_2 > t_1$ se verificase $|x(t_2)| > \epsilon$, entonces:

$$\begin{aligned} V^*(t_2) &= V(t_2, x(t_2)) \geqslant a(|x(t_2)|) \text{ por } a) \\ &> a(\epsilon) = b(\delta) \geqslant b(|x(t_1)|) \text{ por } |x(t_1)| \leqslant \delta \\ &\geqslant V(t_1, x(t_1)) = V^*(t_1) \end{aligned}$$

y, así, V^* no sería creciente para $x(t)$. Por tanto, ha de ser $|x(t)| \leqslant \epsilon$ para todo $t \geqslant t_1$. Esto demuestra que la solución x es prolongable a $[t_1, \infty)$ (porque al tomarnos un t_2 que sea mayor que t_1 la solución no se sale siempre se mantiene en ese intervalo) y que la solución trivial es uniformemente estable. □

El teorema siguiente es un ejemplo de los resultados sobre el **problema inverso** en la teoría del segundo método de Lyapunov. Afirma que la condición del teorema precedente sobre la estabilidad uniforme no es sólo suficiente, si no también necesaria.

Teorema 3.1.3. *Consideramos la ecuación $x'(t) = f(t, x(t))$, donde f es una función como se ha indicado en la introducción de esta sección. La solución trivial es uniformemente estable en $[t_0, \infty)$ si, y sólo si, existe una función $V(t, \xi)$ con las propiedades a) y b) anteriores.*

Demostración. En vista del teorema anterior es suficiente probar que si la solución trivial es uniformemente estable, entonces existe $V(t, \xi)$ con las propiedades a) y b).
Lo que se quiere es probar que existe una función $V(t, \xi)$ con las propiedades del teorema anterior. Para ello procedemos como sigue (es decir debemos de construir dicha función con las propiedades a) y b) del teorema).
Primero veamos que la función $V(t, \xi)$ verifica la propiedad a) para ello, por la definición de estabilidad asintótica nos tomamos en primer lugar $c > 0$ tal que $\bar{B}(0, c) \subset G$. Sea $0 < s \leqslant c$. Definimos $\delta(s)$ como el supremo de todos los números δ, con $0 < s \leqslant c$ tales que si $t_1 \geqslant t_0$, $|\xi| \leqslant \delta$ y $x(t)$ es solución local a la derecha de t_1 tal que $x(t_1) = \xi$, se verifica $|x(t)| \leqslant s$ para todo $t \geqslant t_1$. Según la definición de estabilidad uniforme se verifica $\delta(s) > 0$ para todo $s \in (0, c]$ (ya que si nos tomamos un radio bien pequeño tendremos que $\delta(s)$ es pequeño, pero a medida que el radio de la circunferencia es mayor $\delta(s)$ también es mayor es por eso que se dice que $\delta(s)$ es no decreciente). Además, $\delta(s)$ es no creciente. Así, por el lema 3.1.1 de variable real que enunciamos y demostramos a continuación de esta demostración, se puede elegir $\delta^* : (0, c] \to (0, c]$ tal que δ^* es continua, estrictamente creciente y tal que $\delta^*(s) \leqslant \delta(s)$ para todo $s \in (0, c]$. Además, se tiene $\delta^*(s) \leqslant s$, y, así, $\delta^* \to 0$ para $s \to 0$. Por consiguiente, existe una función b estrictamente creciente y continua inversa de δ^* definida en

$$(0, \delta^*] = (0, r]$$

Notemos que en el teorema 3.1.2 se tiene que $b : [0, r] \to [0, \infty)$ para lograr que cumpla esto debemos de poner $b(0) = 0$, entonces b es estrictamente creciente y continua en $[0, r]$.
Para $t \geqslant t_0$ y $\xi \in \mathbf{R}^\mathbf{n}$ tal que $|\xi| \leqslant r$ definimos $V(t, \xi)$ como el supremo de los valores absolutos de $|x(t + \sigma)|$ al recorrer x todas las soluciones tales que $x(t) = \xi$ y σ todos los valores de $[0, \infty)$ en x está definida.
Pongamos $a(s) = s$ para $s \in [0, r]$ y veamos que

$$a(|\xi|) \leqslant V(t, \xi) \leqslant b(|\xi|)$$

para $t \in [t_0, \infty)$, $\xi \in \mathbf{R}^\mathbf{n}$, $|\xi| \leqslant r$.
Que $V(t, \xi) \geqslant |\xi| = a(|\xi|)$ es consecuencia inmediata de la definición.
Para ver que $b(|\xi|) \geqslant V(t, \xi)$ procedemos de la forma siguiente. Si fuese

$$b(|\xi_1|) < V(t_1, \xi_1)$$

Para algún $t_1 \in [t_0, \infty)$ y ξ_1 con $|\xi_1| \leqslant r$, según la definición de $V(t_1, \xi_1)$, existiría una solución x_1 tal que

$$x_1(t_1) = \xi_1 \quad y \quad |x_1(t_1 + \sigma)| > b(|\xi_1|)$$

para algún $\sigma \geqslant 0$. Pero, según la definición de δ, esto quiere decir que

$$\delta(b(|\xi_1|)) < |\xi_1|,$$

lo que es claramente imposible si $\xi_1 = 0$. También lo es si $\xi \neq 0$, pues esta desigualdad implicaría:

$$|\xi_1| = \delta^*(b(|\xi_1|)) \leqslant \delta(b(|\xi_1|)) < |\xi_1|$$

lo que es contradictorio.
Por lo tanto la función $V(t, \xi)$ cumple con la propiedad a) del teorema anterior.
Veamos finalmente que la función $V(t, \xi)$ verifica la propiedad b). Sea $\bar{x}(t)$ una solución que verifica $\bar{x}(t_1) = \xi_1$ con $|\xi_1| \leqslant r$.
Consideramos $t_3 > t_2 > t_1$, y sea:

$$V^*(t_i) = V(t_i, \bar{x}(t_i)), \quad i = 2, 3$$

Queremos demostrar que

$$V^*(t_3) \leqslant V^*(t_2)$$

Esto resulta fácilmente de la misma definición de $V(t, \xi)$. En efecto, todas las soluciones $x(t)$ que intervienen en el cálculo de $V(t_3, \bar{x}(t_3))$ mediante el supremo de $x(t_3 + \sigma)$ indicado en la definición intervienen también en el cálculo de $V(t_2, \bar{x}(t_2))$, ya que los puntos $(t_2, \bar{x}(t_2))$ y $(t_3, \bar{x}(t_3))$ están sobre la trayectoria de la solución $\bar{x}(t)$ (véase la figura 3.2). □

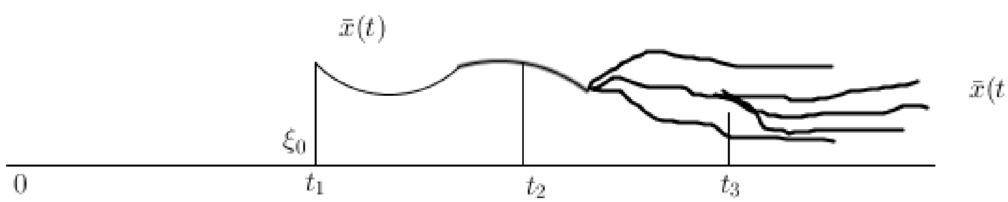


Figura 3.2: Representación gráfica del teorema 3.1.3

Lema 3.1.1 (Lema de la variable real.). *Dada una función* $p : [0, 1] \to [0, 1]$ *tal que* $p(0) = 0$, $p(x) > 0$ *para* $x \in (0, 1]$, *y* p *es no decreciente, existe otra función* $q : [0, 1] \to [0, 1]$ *estrictamente creciente, continua y tal que* $q(x) \leqslant p(x)$ *para todo* $x \in [0, 1]$.

Demostración. La función p tiene a lo sumo un conjunto numerable de saltos y se puede escribir $p(x) = s(x) + c(x)$, donde c es una función continua no decreciente y s es la función de saltos. Es decir, si los puntos de discontinuidad de p son $\{r_k\}$ y el salto en r_k es

$$m_k = \lim_{x \to r_k^+} p(x) - \lim_{x \to r_k^-} p(x)$$

Entonces:

$$s(x) = \sum_{r_k < x} m_k \quad y \quad c(x) = p(x) - s(x)$$

Si $c(x) > 0$ para $x(0)$, entonces podemos poner

$$q(x) = c(x)e^{-x}$$

Veamos que $q(x)$ cumple con las condiciones del enunciado.

- $q(x)$ es continua porque $c(x)$ es continua y la exponencial también lo es.
- $q(x)$ ya que $c(x)$ es no decreciente y así, como, se ha definido la exponencial vemos que es decreciente, pero cuando hacemos el producto se obtiene una función estrictamente creciente.
- $q(x) \leqslant p(x)$ por como hemos definido $p(x)$ se tiene que debe ser mayor que $q(x)$ ya que $p(x)$ cubre todo $[0,1] \times [0,1]$, en cambio, $q(x)$ no.

y así esta función satisface las condiciones del enunciado del teorema.
Supongamos $c(x) = 0$ para $x \in [0,r]$ y $c(x) > 0$ para $x \in (0,r]$. Consideramos entonces la función de saltos s en $[0,r]$. Se tiene $s(x) > 0$ para $x \in [0,r]$. Definimos la función $s^* : [0,1] \to [0,1]$ del siguiente modo:

$$s^*(x) = \begin{cases} s(s/2) & si \quad r/2 < x \leqslant r \\ s(s/4) & si \quad r/4 < x \leqslant r/2 \\ s(s/8) & si \quad r/8 < x \leqslant r/4 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots & \end{cases}$$

La función s^* es una función en escalera, menor o igual que s en todo punto de $[0,r]$, y con discontinuidades tan sólo en los puntos $r/2, r/4, r/8, ...$
Construimos a continuación la función $\bar{s} : [0,1] \to [0,1]$ indicada en la figura 3.3 (línea de puntos).
La función $q(s) = \bar{s}(x)e^{-x}$ satisface las condiciones del enunciado.
Veamos que

- $q(x)$ es continua: de la figura se observa que $\bar{s}$ es continua y además la exponecial es continua, por lo que producto de continuas es continua.
- $q(x)$ es estrictamente creciente: ya que $\bar{s}$ es estrictamente creciente y aunque e^{-x} sea no creciente cuando se realiza el producto de funciones se obtiene que es estrictamente creciente.
- $q(x) \leqslant p(x)$: como $c(x) = 0$ entonces $p(x) = s(x)$ para $x \in [0,r]$ con $s(x) > 0$ por otro lado $s^*(x) \leqslant s(x)$, por como esta definida en $[0,r]$. También de la figura se tiene que $\bar{s}(x) \leqslant s^*(x)$, entonces se tiene que $p(x) \geqslant \bar{s}(x)$ para $x \in [0,r]$, así $p(x) = \bar{s}(x) + c(x)$ para $x \in [0,1]$ por lo que $q(x) \leqslant p(x)$

$\square$

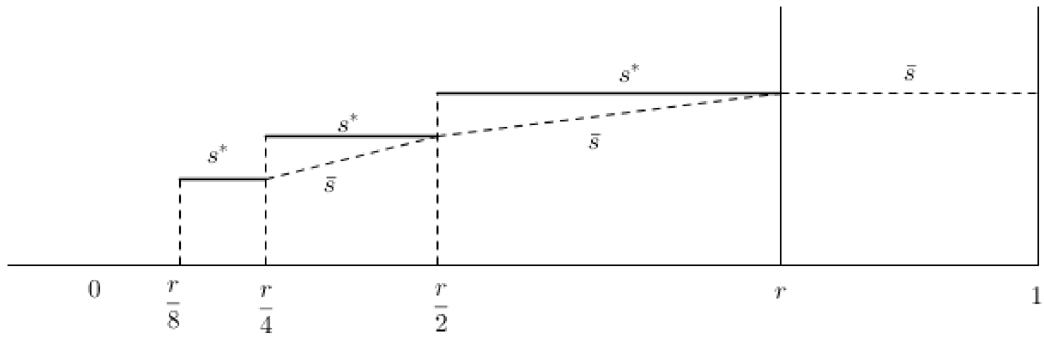


Figura 3.3: Representación gráfica del lema 3.1.1

3.1.1. Ejemplos

Ejemplo 3.1.1. *Estúdiese la estabilidad de la solución trivial del sistema autónomo*

$$\begin{cases} x'(t) = a(x(t))^3 + by(t) \\ y'(t) = -cx(t) + d(y(t))^3 \end{cases} \tag{3.1.2}$$

(Se puede seguir el método de separación de variables intentando obtener una función $V(\xi_1, \xi_2)$ de la forma

$$V(\xi_1, \xi_2) = V_1(\xi_1) + V_2(\xi_2)$$

e imponiendo además que la función V^* sea del mismo tipo que V.)

Solución. *Para simplificar los cálculos escribimos el sistema 3.1.2 en la forma*

$$\begin{cases} x' = ax^3 + by \\ y' = -cx + dy^3 \end{cases}$$

De la observación 1 se tiene por ser un sistema autónomo que

$$V' = \nabla V \cdot f(x) = \sum_{i=1}^{n} \frac{\partial V}{\partial x} \cdot f(x)$$

Denotemos $V' = V^*$. Como se nos dice que la función V' debe ser del mismo tipo que V tenemos

$$\begin{aligned} V' &= \sum_{i=1}^{n} \frac{\partial V}{\partial x} \cdot f(x) \\ &= V_1'(x)(ax^3 + by) + V_2'(y)(-cx + dy^3) \\ &= aV_1'(x)x^3 + bV_1'(x)y - cV_2'(y)x + dV_2'(y)y^3 \end{aligned} \tag{3.1.3}$$

Para que V' tenga la misma estructura funcional que V se requiere que

$$bV_1'(x)y - cV_2'(y)x = 0$$

Separando variables, se tiene

$$\frac{by}{V_2'(y)} = \frac{cx}{V_1'(x)}$$

donde cada una de estas fracciones pueden ser constantes, como por ejemplo, $\frac{1}{2}$. *Así,*

$$\begin{aligned} \frac{cx}{V_1'(x)} &= \frac{1}{2} \\ 2cx &= V_1'(x) \text{ integrando} \\ cx^2 &= V_1(x) \end{aligned}$$

$$\begin{aligned} \frac{by}{V_2'(y)} &= \frac{1}{2} \\ 2by &= V_2'(y) \quad \text{integrando} \\ by^2 &= V_1(y) \end{aligned}$$

se tiene, así que

$$V(x, y) = V_1(x) + V_2(y) = cx^2 + by^2$$

luego, sustituyendo $V_1'(x)$ *y* $V_2'(y)$ *en 3.1.3 resulta que*

$$V' = a(2cx)x^3 + b(2cx)y - c(2by)x + d(2by)y^3 = 2acx^4 + 2bdy^4$$

Cuando $a = 0$, $b > 0$, $c > 0$, $d < 0$ *se tiene que* $V' \leqslant 0$ *y* $V \to 0$, *además* V *es definida positiva, por lo que cumple las condiciones del teorema 3.1.1, así se concluye que el sistema dado es estable.*

Ejemplo 3.1.2. *Estúdiese la estabilidad de la solución trivial del sistema*

$$\begin{cases} x'(t) = -x(t)(y(t))^4 \\ y'(t) = (x(t))^4 y(t) \end{cases}$$

Solución. *Primero reescribamos el sistema anterior como sigue*

$$\begin{cases} x' = -xy^4 \\ y' = x^4 y(t) \end{cases}$$

Sea la función de Lyapunov de la forma $V(x, y) = V_1(x) + V_2(y)$. *Al igual que el ejemplo anterior se quiere que* V' *tenga la misma estructura que* V.

Así

$$V'(x) = V_1'(x)(-xy^4) + V_2'(y)x^4 y = 0 \tag{3.1.4}$$

entonces

$$\frac{-y^4}{V_2'(y)y} = \frac{x^4}{V_1'(x)x} \Longrightarrow \frac{-y^3}{V_2'(y)} = \frac{x^3}{V_1'(x)}$$

Cada una de estas fracciones pueden ser, constantes digamos $\frac{1}{4}$*. Se tiene, entonces que*

$$\begin{aligned} \frac{x^3}{V_1'(x)} &= \frac{1}{4} \\ V_1'(x) &= 4x^3 \quad \textit{integrando resulta que} \\ V_1(x) &= x^4 \end{aligned}$$

Haciendo lo mismo para y resulta $V_2'(y) = -4y^3$ *y* $V_2(y) = -y^4$ *de donde*

$$V(x,y) = x^4 - y^4$$

sustituyendo $V_1'(x)$ *y* $V_2'(y)$ *en 3.1.4*

$$V'(x) = 4x^3(-xy^4) + 4^3x^4y = -4x^4y^4 + 4x^4y^4 = 0 \tag{3.1.5}$$

Así, $V'(x) \leqslant 0$*,* $V \to 0$ *y además es definida positiva, entonces la solución trivial* $(x,y) = (0,0)$ *del sistema dado es estable, ya que cumple con las hipótesis del teorema 3.1.1.*

Ejemplo 3.1.3. ***Funciones de Lyapunov motivadas por la física del problema, péndulo simple.***

Consideremos el péndulo de masa unitaria, longitud l*, donde* θ *es el ángulo de desviación con respecto a la vertical y* g *es la aceleración gravitacional, entonces el movimiento del péndulo es gobernado por la ecuación diferencial*

$$ml\ddot{\theta} + mg\sin(\theta) = 0$$

Como $m = 1$*, y dividiendo entre* l *la ecuación anterior se convierte en*

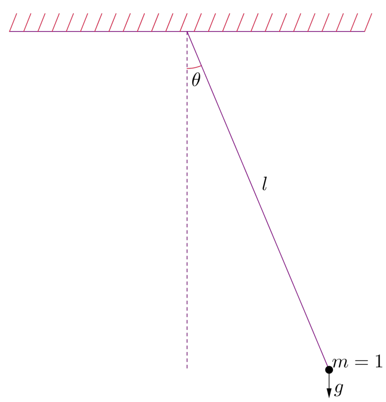


Figura 3.4: Péndulo simple

$$\ddot{\theta} + \frac{g}{l}\sin(\theta) = 0 \tag{3.1.6}$$

Al introducir las variables $x_1 = \theta$ *y* $x_2 = \dot{\theta}$ *la ecuación 3.1.6 es equivalente al sistema*

$$\begin{cases} x_1' = x_2 \\ x_2' = -\dfrac{g}{l}\sin(x_1) \end{cases}$$

Estudiar la estabilidad de la solución trivial del sistema anterior.

Solución. *Dado que se está tratando un modelo de un sistema físico y se conoce el significado de las variables, resulta natural tomar como función candidata de Lyapunov a la energía del mismo, que para este caso es la suma de la energía potencial y cinética de la masa.*
Definamos la energía del péndulo $V(x)$ *como la suma de sus energías cinética* $V(x_2) = \frac{x_2^2}{2}$ *y potencial* $V(x_1) = \int_0^{x_1} \frac{g}{l}\sin(x_1)dx = \frac{g}{l}(1-\cos(x_1))$*, con referencia de energía potencial elegida tal que* $V(0) = 0$*, es decir,*

$$V(x) = \int_0^{x_1} \frac{g}{l}\sin(x_1)dx + \frac{x_2^2}{2} = \frac{g}{l}(1-\cos(x_1)) + \frac{x_2^2}{2}$$

Definida sobre el conjunto $\{x \in \mathbf{R}^n : -\pi < x_1 < \pi\}$*. La derivada de* $V(x)$ *a lo largo de las trayectorias del sistema es*

$$V'(x) = \frac{g}{l}\sin(x_1)x_1' + x_2x_2' = \frac{g}{l}\sin(x_1)x_2 - x_2(\frac{g}{l}\sin(x_1)) = 0$$

Así, se muestra que la función $V(x)$ *cumple con las hipótesis del teorema que son* $V \to 0$*, es definida positiva y* $V' \leqslant 0$*. Por tanto la solución trivial es estable.*

3.2. Inestabilidad

El método directo de Lyapunov permite también dar condiciones que implican la inestabilidad de la solución trivial. Como en la sección anterior consideramos una ecuación $x'(t) = f(t, x(t))$, donde

$$f : [t_0, \infty) \times G \subset \mathbf{R} \times \mathbf{R^n} \to \mathbf{R^n} \tag{3.2.1}$$

(G entorno abierto del origen en $\mathbf{R^n}$) es una función continua y $f(t, 0) = 0$ para todo $t \geqslant t_0$. La solución trivial $x(t) \equiv 0$ es inestable en $[t_0, \infty)$ cuando no es estable en $[t_0, \infty)$.

Corolario 3.2.1. *Sea* $z : [a, b] \longrightarrow \mathbf{R^n}$ *una función continua. Las tres propiedades siguientes son equivalentes*

a) La función z *es no creciente en* $[a, b]$*.*

b) Para todo $t \in (a, b)$*,* $\overline{D}^{+}z(t) = \limsup_{h\longrightarrow 0^+} \dfrac{z(t+h) - z(t)}{h} \leqslant 0$

c) Para todo $t \in (a, b)$*,* $\overline{D}^{-}z(t) = \limsup_{h\longrightarrow 0^-} \dfrac{z(t+h) - z(t)}{h} \leqslant 0$

Teorema 3.2.1 (Teorema de Chetaiev)**.** *Sea* $x'(t) = f(t, x(t))$*, como se indica en (3.2.1), y supongamos que existe una función continua*

$$V : [t_0, \infty) \times \bar{B}(0, r) \subset \mathbf{R} \times G \to (-\infty, \infty)$$

tal que se verifica:

a) $|V(t, \xi)| \leqslant b(|\xi|)$ *para* $t \geqslant t_0$ *y* $|\xi| \leqslant r$*, siendo* $b : [0, r] \to [0, \infty)$ *estrictamente creciente, continua y tal que* $b(0) = 0$*.*

b) Para cada $\delta > 0$ *y* $t_1 \geqslant t_0$ *existe* ξ^**,* $|\xi^*| \leqslant \delta$*, tal que* $V(t_1, \xi^*) < 0$*.*

c) Si $x(t)$ *es una solución tal que* $x(t_1) = \xi_1$ *con* $t_1 \geqslant t_0$ *y* $|\xi_1| < r$*, se tiene:*

$$\limsup_{h \to 0^+} \frac{V(t_1 + h, x(t_1 + h)) - V(t_1, x(t_1))}{h} \leqslant -c(|\xi_1|)$$

siendo $c : [0, r] \to [0, \infty)$ *continua, no decreciente y tal que* $c(0) = 0$*.*
Entonces la solución trivial es inestable en $[t_0, \infty)$*.*

Demostración. La prueba la haremos por contradicción, es decir supondremos que la solución trivial es estable en $[t_0, \infty)$ y trataremos de llegar a una contradicción con la hipótesis. Por definición tenemos que si la solución trivial es estable en $[t_0, \infty)$, entonces para cada $\epsilon > 0$, $0 < \epsilon < r$, existe $\delta > 0$ tal que si $\eta_0 \leqslant \delta$ y $x(t)$ es una solución tal que $x(t_0) = \eta_0$, entonces $|x(t)| \leqslant \epsilon$ para todo $t \geqslant t_0$.
Por la condición *b*) podemos escoger ξ_0, $|\xi_0| \leqslant \delta$ tal que $V(t_0, \xi_0) < 0$. Si $x(t)$ es solución tal que $x(t_0) = \xi_0$, entonces $|x(t)| \leqslant \epsilon$ para todo $t \geqslant t_0$.
Por la condición *a*) se tiene, entonces

$$|V(t, x(t))| \leqslant b|x(t)| \leqslant b(\epsilon) \tag{3.2.2}$$

para todo $t \geqslant t_0$. En lo que sigue suponemos elegida esta solución $x(t)$ con $x(t_0) = \xi_0$.
La condición *c*) implica que $V(t, x(t))$ es no creciente en $[t_0, \infty)$ ya que una solución cualquiera $z(t)$ es no creciente en $[t_0, \infty)$ si, y sólo si $\bar{D}^+ z(t) \leqslant 0$ para $t \in [t_0, \infty)$ (esto es por corolario 3.2.1) y además como c es positiva de ahí resulta que $V(t, x(t))$ es no creciente. Así, para todo $t \geqslant t_0$ se tiene:

$$V(t, x(t)) \leqslant V(t_0, \xi_0) < 0$$

y, por tanto:

$$|V(t, x(t))| \geqslant |V(t_0, \xi_0)|$$

por la ecuación (3.2.2) se tiene que

$$b(|x(t)|) \geqslant |V(t_0, \xi_0)|$$

y, así:

$$|x(t)| \geqslant b^{-1}(|V(t_0, \xi_0)|)$$

para todo $t \geqslant t_0$ donde b^{-1} representa la función inversa de b.
La condición *c*) implica también que, si llamamos

$$z(t) = V(t, x(t)) - V(t_0, \xi_0) + \int_{t_0}^{t} c(|x(u)|)du$$

entonces

$$z(t + h) = V(t + h, x(t + h)) - V(t_0, \xi_0) + \int_{t_0}^{t} c(|x(u)|)du$$

luego se tiene que

$$z(t + h) - z(t) = V(t + h, x(t + h)) - V(t, x(t)) + \int_{t}^{t+h} c(|x(u)|)du$$

así que

$$\begin{aligned}
\limsup_{h\to 0^+}\frac{z(t+h)-z(t)}{h} &= \limsup_{h\to 0^+}\frac{V(t+h,x(t+h))-V(t,x(t))+\int_t^{t+h}c(|x(u)|)du}{h}\\
&= \limsup_{h\to 0^+}\frac{V(t+h,x(t+h))-V(t,x(t))}{h}+\limsup_{h\to 0^+}\frac{\int_t^{t+h}c(|x(u)|)du}{h}\\
&= \limsup_{h\to 0^+}\frac{V(t+h,x(t+h))-V(t,x(t))}{h}+c(|x(t)|)\ \textit{c es creciente}\\
&\leqslant \limsup_{h\to 0^+}\frac{V(t+h,x(t+h))-V(t,x(t))}{h}+c(\epsilon)\ \textit{por la condición c)}\\
&\leqslant 0
\end{aligned}$$

así

$$\limsup_{h\to 0^+}\frac{z(t+h)-z(t)}{h}\leqslant 0\ \ \textit{para todo}\ \ t\geqslant t_0$$

y, por tanto, por el corolario 3.2.1, $z(t)$ es una función no creciente.

Ahora, si evaluamos t_0 en la función z se tiene

$$\begin{aligned}
z(t_0) &= V(t_0,x(t_0))-V(t_0,\xi_0)+\int_{t_0}^{t_0}c(|x(u)|)du\ \textit{pero}\ x(t_0)=\xi^0\\
z(t_0) &= V(t_0,\xi^0)-V(t_0,\xi^0)+0\\
z(t_0) &= 0
\end{aligned}$$

$z(t_0)=0$, entonces $z(t)\leqslant z(t_0)=0$, es decir $z(t)\leqslant 0$, lo que implica que

$$V(t,x(t))-V(t_0,\xi_0)+\int_{t_0}^{t}c(|x(u)|)du\leqslant 0$$

entonces

$$V(t,x(t))\leqslant V(t_0,\xi_0)-\int_{t_0}^{t}c(|x(u)|)du$$

Como $|x(t)|\geqslant b^{-1}(|V(t_0,\xi_0)|)$, se sigue que

$$c(|x(t)|)\geqslant c(b^{-1}(|V(t_0,\xi_0)|))\quad \textit{por ser lafunción c no decreciente}$$

ahora, integrando se tiene

$$\int_{t_0}^{t}c(|x(u)|)du\geqslant\int_{t_0}^{t}c(b^{-1}(|V(t_0,\xi_0)|))=(t-t_0)c(b^{-1}(|V(t_0,\xi_0)|))$$

multiplicando por -1

$$-\int_{t_0}^{t}c(|x(u)|)du\leqslant-(t-t_0)c(b^{-1}(|V(t_0,\xi_0)|))$$

y, así:

$$\begin{aligned}
V(t,x(t) &\leqslant V(t_0,\xi_0)-\int_{t_0}^{t}c|x(u)|du\\
&\leqslant V(t_0,\xi_0)-(t-t_0)c(b^{-1}|(V(t_0,\xi_0))|))
\end{aligned}$$

por lo que

$$V(t, x(t)) = V(t_0, \xi_0) - (t - t_0)c(b^{-1}(|V(t_0, \xi_0)|))$$

de este modo,

$$V(t, x(t)) \to -\infty \ \ para \ \ t \to -\infty$$

lo cual contradice la relación hallada anteriormente:

$$|V(t, x(t))| \leqslant b(\epsilon)$$

□

Como caso particular del teorema de Chetaiev se obtiene el primer teorema de Lyapunov sobre inestabilidad. Aquí presentamos a continuación el llamado segundo teorema de Lyapunov sobre inestabilidad.

Teorema 3.2.2 (Segundo teorema de Lyapunov sobre inestabilidad)**.** *Sea $x'(t) = f(t, x(t))$, como se indica en (3.2.1), y supongamos que existe una función*

$$V : [t_0, \infty) \times \bar{B}(0, r) \to (-\infty, +\infty)$$

acotada, con derivadas continuas y tal que

a) Si $x(t)$ es una solución tal que $|x(t)| \leqslant r$ para todo $t \geqslant t_0$, se define $V^*(t) = V(t, x(t))$, entonces:

$$\frac{dV^*}{dt}(t) = \lambda V^*(t) + W(t, x(t))$$

con $\lambda > 0$, donde $W : [t_0, \infty) \times \bar{B}(0, r) \to [0, \infty)$ es una función continua no negativa.

b) Para todo $\delta > 0$, $0 < \delta \leqslant r$, existe ξ_0 con $|\xi_0| \leqslant \delta$ tal que $|V(t_0, \xi_0)| > 0$ y, si W no es idénticamente nula, se puede elegir ξ_0 tal que $V(t_0, \xi_0) > 0$.
Entonces la solución trivial es inestable.

Demostración. La prueba la haremos por contradicción, es decir supondremos que la solución trivial es estable y llegaremos a una contradicción. Si hay estabilidad, para $\epsilon > 0$, con $\epsilon \leqslant r$, existe $\delta > 0$ tal que una solución $x(t)$ tal que $|x(t_0)| \leqslant \delta$ verifica $|\xi| \leqslant \epsilon$ para todo $t \geqslant t_0$. Tomamos, de acuerdo con b), ξ_0 $|\xi_0| \leqslant \delta$, tal que $|V(t_0, \xi_0)| > 0$ y, si W no es idénticamente nula, $V(t_0, \xi_0) > 0$. Consideremos una solución $x(t)$ tal que

$$x(t_0) = \xi_0$$

Según las condiciones del teorema, $V(t, x(t))$ es acotada. Por otra parte, por a) podemos escribir:

$$\frac{dV^*}{dt}(t) = \lambda V^*(t) + W(t, x(t))$$

$$[V(t, x(t))]' = [V^*(t)]' = \lambda V^*(t) + W(t, x(t))$$

veamos que en esta última ecuación tenemos las hipótesis del teorema de la fórmula de Lagrange (teorema 1.3.2d) ya que $x'(t) = [V^*(t)]'$, $x(t) = V^*$, $A = \lambda$, $b(t) = W(t, x(t))$ $x(t_0) = V^*(0)$ y $\xi_0 = V(t_0, x(t_0))$ así se tiene el problema de valores iniciales

$$\begin{cases} [V^*(t)]' = \lambda V^*(t) + W(t, x(t)) \\ V^*(t_0) = V(t_0, x(t_0)) \end{cases}$$

que tiene solución única. Pero por ser λ una constante se tiene que $\Phi(t) = e^{\lambda(t)}$ así, se tiene que

$$V(t, x(t)) = e^{\lambda t} V(t_0, x(t_0)) + e^{\lambda t} \int_{t_0}^{t} e^{-\lambda s} W(s, x(s))$$

que es no acotada para $t \to +\infty$ y por hipótesis tenemos que $V(t, x(t))$ es acotada, lo que nos da una contradicción. Esto demuestra el teorema. □

Observese que en el teorema se puede imponer en a) la condición $W \leqslant 0$ en lugar de ser $W \geqslant 0$, e imponer en b), si W no es idénticamente nula, $V(t_0, \xi_0) < 0$.

3.2.1. Ejemplos

Ejemplo 3.2.1. *Se considera el sistema*

$$\begin{cases} x'(t) = a(x(t))^3 + by(t) \\ y'(t) = -cx(t) + d(y(t))^3 \end{cases} \tag{3.2.3}$$

estudiado en el ejemplo 3.1.1 de la sección anterior. Demuéstrese que si $bc < 0$, $ad < 0$, o bien si $a > 0$, $b > 0$, $c > 0$, $d > 0$, entonces la solución trivial es inestable.

Solución. *Del ejemplo 3.1.1, sabemos que*

$$V(x, y) = V_1(x) + V_2(y) = cx^2 + by^2$$

y

$$V'(x, y) = a(2cx)x^3 + b(2cx)y - c(2by)x + d(2by)y^3 = 2acx^4 + 2bdy^4$$

Veamos que V' se puede escribir como $V'(x) = \lambda V(x) + W(x)$. Así

$$V'(x, y) = 2(ax^2 + by^2)(cx^2 + dy^2) + [-2(ad + bc)x^2 y^2]$$

Si $bc < 0$, $ad < 0$, entonces $V'(x, y) > 0$, así, el sistema dado será inestable. Por otro lado, si $a > 0$, $b > 0$, $c > 0$, $d > 0$, tenemos $V'(x, y) > 0$, también es inestable. En ambos casos se cumplen las hipótesis del teorema 3.2.2.

Ejemplo 3.2.2. *Se considera el sistema:*

$$\begin{cases} x'(t) = (x(t))^5 + (y(t))^3 \\ y'(t) = (x(t))^3 - (y(t))^5 \end{cases}$$

Estúdiese la estabilidad de la solución trivial. (Utilícese $V(\xi_1, \xi_2) = \xi_1^4 - \xi_2^4$.)

Solución. *Primero reescribamos el sistema anterior de la siguiente forma*

$$\begin{cases} x' = -x^5 + y^3 \\ y' = x^3 - y^5 \end{cases}$$

Sabemos que

$$\frac{dV}{dx} = \sum_{i}^{n} \frac{\partial V}{\partial x_i} \cdot f(x).$$

de donde, resulta que

$$\begin{aligned} V'(x,y) &= V_1'(x)x' + V_2'(y)y' = V_1'(x)(x^5 + y^3) + V_2'(y)(x^3 - y^5) \\ &= V_1'(x)x^5 + V_1'(x)y^3 + V_2'(y)x^3 - V_2'(y)(y^5) \end{aligned}$$

Como, $V'(x,y)$ debe tener la misma estructura que $V(x,y) = V_1(x) + V_2(y)$, entonces se tendrá que

$$V_1'(x)y^3 + V_2'(y)x^3 = 0$$

separando variables

$$V_2'(y)x^3 = -V_1'(x)y^3 \Rightarrow \frac{x^3}{V_1'(x)} = -\frac{y^3}{V_2'(y)}$$

Cada una de las fracciones pueden ser constantes, digamos $\frac{1}{4}$. Así, se tendrá $V_1'(x) = 4x^3$ al integrar, resulta $V_1(x) = x^4$ y $V_2'(y) = -4y^3$ integrando $V_2(y) = -y^4$. Ahora

$$V(x,y) = x^4 - y^4$$

y

$$V'(x,y) = V_1'(x)x^5 + V_2'(y)y^5 = 4x^3x^5 + 4y^3y^5 = 4x^8 + 4y^8 = \underbrace{4}_{\lambda}(\underbrace{x^4 - y^4}_{V})^2 + (\underbrace{8x^4y^4}_{W})$$

Analizando para $x,\ y$ se tiene; si $x > 0,\ y > 0$, entonces por teorema 3.2.2 se tiene que $V'(x,y) > 0$, así la solución trivial del sistema dado es inestable. El mismo resultado se obtiene para $x < 0,\ y < 0$.

3.3. Estabilidad asintótica

El método directo de Lyapunov permite también estudiar la estabilidad asintótica de la solución trivial. Como en la sección anterior, consideramos una función continua

$$f : [t_0, \infty) \times G \subset \mathbf{R} \times \mathbf{R^n} \to \mathbf{R^n}$$

donde G es un entorno abierto del origen, y se tiene $f(t,0) = 0$ para todo $t \geqslant t_0$. Estudiamos la estabilidad asintótica de la solución trivial $x(t) \equiv 0$ de la ecuación $x'(t) = f(t, x(t))$. Recordamos que se dice que la solución trivial es asintóticamente estable en $[t_0, \infty)$ cuando es estable, y, además, existe $\delta > 0$ tal que si $x(t)$ es una solución que verifica $|x(t_0)| \leqslant \delta$, entonces $x(t)$ está definida para $t \geqslant t_0$ y verifica $x(t) \to 0$ para $t \to \infty$.

Teorema 3.3.1. *Consideramos la ecuación $x'(t) = f(t, x(t))$ que se indica arriba. Supongamos que existe la función*

$$V : [t_0, \infty) \times \bar{B}(0, r) \subset \mathbf{R} \times G \to \mathbf{R^n}$$

tal que

a) $V(t_0, \xi) \to 0$ *para todo* $|\xi| \to 0$.

b) $V(t, \xi) \geqslant a(|\xi|)$ *para todo* $t \geqslant t_0$, $|\xi| \leqslant r$ *siendo* $a : [0, r] \to [0, \infty)$ *una función continua, estrictamente creciente y tal que* $a(0) = 0$ *(es decir* $V(t, \xi)$ *es definida positiva).*

c) Para toda solución local $x(t)$ *a la derecha de* t_0 *tal que* $|x(t_0)| < r$ *se verifica:*

$$\limsup_{h \to 0^+} \frac{V(t + h, x(t + h)) - V(t, x(t))}{h} \leqslant -c(V(t, x(t)))$$

para todo t *donde este límite está definido, siendo* $c : [0, r] \to [0, \infty)$ *una función continua, no decreciente y tal que* $c(0) = 0$.

Entonces la solución trivial es asintóticamente estable.

Demostración. Por el corolario 3.2.1 tenemos que $V(t, x(t))$ es no creciente en t si para todo t

$$\overline{D}^+ V(t, x(t)) = \limsup_{h \to 0^+} \frac{V(t + h, x(t + h)) - V(t, x(t))}{h} \leqslant 0$$

y por la condición c) sabemos que c es positiva de ahí resulta que la condición c) implica junto con el corolario 3.2.1 que $V(t, x(t))$ es no creciente en t tal como existe el teorema 3.1.1 sobre estabilidad. Así, la solución trivial es estable. Por tanto, existe $\delta > 0$ tal que si $x(t)$ es una solución con $|x(t)| < r$ para $t \geqslant t_0$. De la condición c) se tiene que si llamamos

$$z(t) = V(t, x(t)) - V(t_0, x(t_0)) + \int_{t_0}^{t} c(V(s, x(s)))ds$$

entonces

$$z(t + h) = V(t + h, x(t + h)) - V(t_0, x(t_0)) + \int_{t_0}^{t+h} c(V(s, x(s)))ds$$

así que

$$z(t + h) - z(t) = V(t + h, x(t + h)) - V(t, x(t)) + \int_{t}^{t+h} c(V(s, x(s)))ds$$

luego

$$\begin{aligned}
\limsup_{h \to 0^+} \frac{z(t + h) - z(t)}{h} &= \limsup_{h \to 0^+} \left[\frac{V(t + h, x(t + h)) - V(t, x(t))}{h} + \frac{1}{h} \int_{t}^{t+h} c(V(s, x(s)))ds\right] \\
&\leqslant \limsup_{h \to 0^+} \frac{V(t + h, x(t + h)) - V(t, x(t))}{h} \\
&+ \limsup_{h \to 0^+} \frac{1}{h} \int_{t}^{t+h} c(V(s, x(s)))ds \text{ \textit{por c) se tiene}} \\
&\leqslant -c(V(t, x(t))) + c(V(t, x(t))) \\
&\leqslant 0
\end{aligned}$$

Así se tiene que

$$\limsup_{h\to 0^+}\left[\frac{V(t+h,x(t+h))-V(t,x(t))}{h}+\frac{1}{h}\int_t^{t+h}c(V(s,x(s)))ds\right]\leqslant 0 \tag{3.3.1}$$

Esto demuestra, en primer lugar, que $V(t,x(t))$ es no creciente, en virtud del corolario 3.2.1, y, así

$$\lim_{t\to\infty}V(t,x(t))=V_0\geqslant 0$$

Pues sabemos que al aplicar límite a una función no creciente está siempre es positiva.
Supongamos primero $V_0>0$. Entonces $c(V_0)>0$, y, siendo c no decreciente, se tiene:

$$c(V(t,x(t)))\geqslant c(V_0)>0 \;\; \textit{para todo } t\geqslant t_0$$

Sustituyendo V_0 en la ecuación (3.3.1) se tiene que

$$\begin{aligned}\limsup_{h\to 0^+}\left[\frac{V(t+h,x(t+h))-V(t,x(t))}{h}+\frac{1}{h}\int_t^{t+h}c(V_0))ds\right] &\leqslant 0\\ \limsup_{h\to 0^+}\left[\frac{V(t+h,x(t+h))-V(t,x(t))}{h}+\frac{1}{h}c(V_0)s|_t^{t+h}\right] &\leqslant 0\end{aligned}$$

así se tiene:

$$\limsup_{h\to 0^+}\left[\frac{V(t+h,x(t+h))-V(t,x(t))}{h}+c(V_0)\right]\leqslant 0$$

Ahora sustituyendo t_0 en $z(t)$ se obtiene

$$z(t_0)=V(t_0,x(t_0))-V(t_0,\xi_0)+\int_{t_0}^{t}c(V(s,x(s)))ds=V(t_0,\xi_0)-V(t_0,\xi_0)=0$$

Por ser $z(t)$ no creciente por el corolario 3.2.1 se tiene que $z(t)\leqslant z(t_0)=0$, así $z(t)\leqslant 0$. Entonces

$$z(t)=V(t,x(t))-V(t_0,x(t_0))+\int_{t_0}^{t}c(V(s,x(s)))ds\leqslant 0$$

de donde

$$\begin{aligned}V(t,x(t))-V(t_0,x(t_0)) &+ \int_{t_0}^{t}c(V(s,x(s)))ds\leqslant 0\\ V(t,x(t))-V(t_0,x(t_0)) &\leqslant -\int_{t_0}^{t}c(V(s,x(s)))ds\\ V(t,x(t))-V(t_0,x(t_0)) &\leqslant -\int_{t_0}^{t}c(V_0)ds\\ V(t,x(t))-V(t_0,x(t_0)) &\leqslant -c(V_0(t-t_0))\end{aligned}$$

Así,

$$V(t,x(t))\to-\infty \quad \textit{para} \quad t\to+\infty,$$

lo que contradice a). Por tanto,

$$V_0 = \lim V(t, x(t)) = 0 \quad para \quad t \to +\infty$$

Esto implica, según b), que

$$a(|x(t)|) \to 0 \quad para \quad t \to +\infty$$

y, de ese modo, $|x(t)| \to 0$ para $t \to +\infty$ lo que demuestra estabilidad asintótica. □

Un teorema importante sobre estabilidad asintótica es el siguiente.

Teorema 3.3.2. *Consideramos la ecuación autónoma*

$$x'(t) = f(x(t)), \quad x : \mathbf{R} \to \mathbf{R^n}$$

donde $f : \mathbf{R^n} \to \mathbf{R^n}$ *es de clase* $C^1(\mathbf{R^n})$ *y tal que* $f(0) = 0$. *Supongamos que existe una función* $V : \mathbf{R^n} \to [0, \infty)$, $V \in C^1(\mathbf{R^n})$ *tal que*

a) $V(\xi) > 0$ *si* $\xi \neq 0$, $V(0) = 0$.

b) $V(\xi) \to \infty$ *para* $|\xi| \to \infty$. *(Definición de radialmente no acotada.)*

c) Si $x(t)$ *es una solución cualquiera distinta de la trivial, y* $V^*(t) = V(x(t))$, *entonces*

$$\frac{dV^*}{dt}(t) < 0$$

para todo t.

Entonces, la solución trivial es globalmente asintóticamente estable, es decir, si $x(t)$ *es cualquier solución,* $x(t) \to 0$ *para* $t \to \infty$.
Si en lugar de c) se verifica la condición

c') $$\frac{dV^*}{dt}(t) = \left(\frac{\partial V}{\partial \xi}\,(x(t), f(x(t))\right) \leqslant 0$$

y el conjunto

$$M = \left\{\xi \in \mathbf{R^n} : \left(\frac{\partial V}{\partial \xi}(\xi), f(\xi)\right) = 0\right\}$$

es tal que para ninguna solución $x(t)$ *distinta de la trivial se verifica*

$$\{x(t) : t \geqslant t_0\} \subset M$$

para ningún t_0, *entonces la conclusión es también válida.*

Demostración. Sea $\xi_0 \in \mathbf{R^n}$ arbitrario. Sea $c = V(\xi_0)$. La condición de radialmente no acotada b) implica que para cualquier $c > 0$ existe $r > 0$ tal que $V(\xi) > V(\xi_0)$ cuando $\|\xi\| > r$. Sea $x(t)$ una solución tal que $x(0) = \xi_0$. Por tanto se sigue de la condición c) dado que:

$$\frac{dV^*}{dt}(t) \leqslant 0,$$

se tiene:

$$V^*(t) = V(x(t)) \leqslant V^*(0) = V(x(0)) = V(\xi_0)$$

para $t \geqslant t_0$. Por tanto se verifica que

$$B = \{\xi \in \mathbf{R^n} : V(\xi) \leqslant V(\xi_0)\}$$

es cerrado y acotado, existe algún punto $\eta \in B$ que es un punto ω de ξ_0, es decir, es tal que para cualquier entorno suyo $\mathbf{U}(\eta)$ y cualquier $\mathbf{T} \geqslant 0$ existe $t \geqslant \mathbf{T}$ tal que $x(t) \in \mathbf{U}$.
Por el lema que presentamos a continuación, el conjunto Ω de todos los puntos ω de ξ_0 es un conjunto cerrado, y, si $z(t)$ es una solución tal que $z(t_0) \in \Omega$ para algún t_0, entonces $z(t) \in \Omega$ para $t \geqslant t_0$.
Como $V(x(t))$ es no creciente, resulta que el conjunto de puntos ω de η se encuentra sobre

$$H = \{\xi \in \mathbf{R^n} : V(\xi) = V(\eta)\}$$

Supongamos primero $V(\eta) = 0$. Entonces $H = 0$, y, así,

$$\liminf x(t) = 0 \quad para \quad t \to \infty$$

Como de las condiciones dadas y del teorema 3.1.1 se deduce la estabilidad de la solución trivial, la condición $\liminf x(t) = 0$ conduce a $\lim x(t) = 0$ para $t \to \infty$.
Supongamos ahora $V(\eta) > 0$. Sobre H se encuentra el conjunto Ω de puntos ω de ξ_0 que tiene la propiedad anteriormente indicada. Si $z(t)$ es solución tal que $z(t_0) \in \Omega$ para algún t_0, entonces $z(t) \in \Omega$ para todo $t \geqslant t_0$. Si consideramos $V(z(t))$ para $t \geqslant t_0$, se tiene:

$$\left(\frac{\partial V}{\partial \xi}(z(t)), f(z(t))\right) = 0$$

y, así, $z(t)$ para $t \geqslant t_0$, se encuentra sobre el conjunto M y, por tanto, las condiciones c) o c') implican $z(t) \equiv 0$.
De este modo resulta, como antes, $\lim x(t) = 0$ para $t \to \infty$. □

Lema 3.3.1. *Consideramos la ecuación autónoma*

$$x'(t) = f(t, x(t)), \quad x : \mathbf{R} \to \mathbf{R^n}$$

donde $f : \mathbf{R^n} \to \mathbf{R^n}$ *es de clase* $C^1(\mathbf{R^n})$. *Sea* $\xi \in \mathbf{R^n}$ *y* $x(t)$ *solución tal que* $x(t_1) = \xi$ *para algún* t_1. *Decimos que un punto* $\eta \in \mathbf{R^n}$ *es punto* ω *de* ξ *cuando para entorno* $\boldsymbol{U}(\eta)$ *de* η *y todo* $\boldsymbol{T} \geqslant t_1$ *existe* $t \geqslant \boldsymbol{T}$ *tal que* $x(t) \in \boldsymbol{U}(\eta)$.
Con esta terminología el conjunto Ω *de puntos* ω *de* ξ *es cerrado y satisface la propiedad siguiente: Si* η *es punto* ω *de* ξ *y* $z(t)$ *es solución tal que* $z(t_0) = \eta$ *para algún* t_0, *entonces todos los puntos* $z(t)$ *para* $t \geqslant t_0$ *son puntos de* ω *de* ξ.

Demostración. Que Ω es cerrado se deduce fácilmente de la misma definición. Para demostrar la otra propiedad indicada, sean $\eta \in \Omega$ y $z(t)$ solución tal que $z(t_0) = \eta$. Fijemos $z(t_1)$ con $t_1 > t_0$. Sea G un entorno cualquiera de $z(t_1)$. Sea $\mathbf{U}(\eta)$ entorno de η tal que toda solución $y(t)$ con $y(t_0) \in \mathbf{U}(\eta)$ satisface $y(t_1) \in G$. La existencia de $\mathbf{U}(\eta)$ queda garantizada por la continuidad de la solución respecto de las condiciones iniciales. Sea ahora dado cualquier $\mathbf{T} \geqslant t_0$. Sabemos que existe $t^* \geqslant \mathbf{T}$ tal que $x(t^*) \in \mathbf{U}(\eta)$. Consideramos la función

$$u(t) = x(t + t^* - t_0)$$

Por ser la ecuación autónoma, resulta que $u(t)$ es solución y satisface

$$u(t_0) = x(t^*) \in \mathbf{U}(\eta)$$

Por tanto,

$$u(t_1) = x(t_1 + t^* - t_0) \in G,$$

lo que demuestra que $z(t_1)$ es punto ω de ξ. □

3.3.1. Ejemplos

Ejemplo 3.3.1. *Se considera el sistema*

$$\begin{cases} x'(t) = a(x(t))^3 + by(t) \\ y'(t) = -cx(t) + d(y(t))^3 \end{cases} \tag{3.3.2}$$

estudiado en el ejemplo 3.1.1 de la sección 3.1. Demuéstrese que si $b > 0$, $c > 0$ $a < 0$, $d < 0$, la solución trivial es asintóticamente estable.

Solución. *Sabemos por el ejemplo 3.1.1 que*

$$V(x, y) = V_1(x) + V_2(y) = cx^2 + by^2$$

y

$$V'(x, y) = a(2cx)x^3 + b(2cx)y - c(2by)x + d(2by)y^3 = 2acx^4 + 2bdy^4$$

Así, $V(x, y)$ cumple con $V(x, y) \to 0$ cuando $(x, y) \to 0$, además es definida positiva y al analizar $b > 0$, $c > 0$ $a < 0$, $d < 0$ en $V'(x, y)$ resulta que es menor que cero. Por tanto cumple las propiedades del teorema 3.1.1.

Ejemplo 3.3.2. *Estúdiese la estabilidad de la solución trivial del sistema*

$$\begin{cases} x'(t) = -2x(t)(y(t))^4 \\ y'(t) = (x(t))^4 y(t) \end{cases}$$

Solución. *Primero reescribamos el sistema anterior como sigue*

$$\begin{cases} x' = -2xy^4 \\ y' = x^4 y \end{cases}$$

Sea la función de Lyapunov de la forma $V(x,y) = V_1(x) + V_2(y)$. Al igual que el ejemplo anterior se quiere que V' tenga la misma estructura que V.
Así

$$V'(x) = V_1'(x)(-2xy^4) + V_2'(y)x^4 y = 0 \tag{3.3.3}$$

entonces

$$\frac{-y^4}{V_2'(y)y} = \frac{2x^4}{V_1'(x)x} \Longrightarrow \frac{-y^3}{V_2'(y)} = \frac{2x^3}{V_1'(x)}$$

Cada una de estas fracciones pueden ser, constantes digamos $\frac{1}{4}$. Se tiene, entonces que

$$\begin{aligned} \frac{2x^3}{V_1'(x)} &= \frac{1}{4} \\ V_1'(x) &= 8x^3 \quad \textit{integrando resulta que} \\ V_1(x) &= 2x^4 \end{aligned}$$

Haciendo lo mismo para y resulta $V_2'(y) = -4y^3$ y $V_2(y) = -y^4$ de donde

$$V(x,y) = 2x^4 - y^4$$

sustituyendo $V_1'(x)$ y $V_2'(y)$ en (3.3.3)

$$V'(x,y) = 8x^3(-2xy^4) + 4y^3x^4y = -16x^4y^4 + 4x^4y^4 = -12x^4y^4 \tag{3.3.4}$$

Así, $V'(x,y) < 0$, $V(x,y) \to 0$ y además es definida positiva, entonces la solución trivial $(x,y) = (0,0)$ del sistema dado es asintóticamente estable, ya que cumple con las hipótesis del teorema 3.3.1.

3.4. La función de Lyapunov para ecuaciones lineales con coeficientes constantes

En esta sección nos ocuparemos de la construcción de las funciones de Lyapunov, para sistemas lineales a coeficientes constantes.
Primero tenemos algunas clasificaciones de las formas cuadráticas. Clasificación a través de la forma canónica:

- Dados los elementos de la diagonal de la matriz asociada a la forma canónica, es decir, los coeficientes de los términos cuadrados: $a_{11}, a_{22}, \ldots, a_{nn}$ la forma cuadrática es:

- **Definida positiva,** si todos los $a_{ij}, i = j$ son positivos $n = r = p$. n : es dimensión de la matriz. r : es el número de valores de la diagonal diferente de cero. p : es el número de valores estrictamente positivos de la diagonal.

- **Definida negativa,** si todos los $a_{ij}, i = j$ son negativos $n = r$, $p = 0$.
- **Semidefinida positiva,** si algunos $a_{ij}, i = j$ son positivos y otros nulos $r < n, p = r$.
- **Semidefinida negativa,** si algunos $a_{ij}, i = j$ son negativos $r < n, p = 0$. **No definida,** si algunos $a_{ij}, i = j$ son negativos y otros positivos.

Consideremos la ecuación $x'(t) = Ax(t)$, donde $x : \mathbf{R} \to \mathbf{R^n}$ y A es una matriz $n \times n$ con elementos reales.

El estudio de la estabilidad de la solución trivial se puede llevar a cabo por el método directo de Lyapunov de la siguiente forma. Construiremos una función

$$V(\xi) = (\xi, B\xi)$$

para $\xi \in \mathbf{R^n}$ y B una matriz real simétrica $n \times n$ aún desconocida, tal que si $x(t)$ es una solución de $x'(t) = Ax(t)$ se verifique, si $V^*(t) = V(x(t)) = (x(t), Bx(t))$:

$$\frac{dV^*}{dt}(t) = -(x(t), Cx(t))$$

donde C es una matriz real dada definida positiva. Derivando V se tiene

$$\begin{aligned} V'(x(t)) &= (x'(t), Bx(t)) + (x(t), Bx'(t)) \\ &= (Ax(t), Bx(t)) + (x(t), BAx(t)) \\ &= (x(t), (A^*B + BA)x) \\ &= -(x(t), Cx(t)) \end{aligned}$$

basta demostrar que dada C existe B tal que $A^*B + BA = -C$, donde A^* representa la traspuesta de A.

Para ello consideramos el operador lineal

$$F : M(n, \mathbf{R}) \to M(n, \mathbf{R})$$

(donde $M(n, \mathbf{R})$ representa el espacio vectorial de dimensión $n \times n$ de las matrices $n \times n$ de elementos reales) definido por

$$F(B) = A^*B + BA \quad \textit{para cada } B \in M(n, \mathbf{R})$$

A fin de demostrar que $A^*B + BA = -C$ tiene solución, cualquiera que sea la matriz C bastará demostrar que el operador F es inversible, es decir, que ningún autovalor de F es nulo.

$$\det F(B) = \prod_{i=1}^{n} \mu_i \tag{3.4.1}$$

Sea μ un autovalor de F y sea $B \neq 0$ tal que

$$F(B) = \mu B,$$

es decir,

$$A^*B + BA = \mu B$$

o, lo que es lo mismo,

$$(A^* - \mu I)B = -BA$$

El lema general que sigue demuestra que $A^* - \mu I$ y $-A$ tienen que tener un autovalor común.

Lema 3.4.1. *Sean B, P, $Q \in M(n, \mathbf{R})$ tales que $B \neq 0$ y $PB = BQ$. Entonces P y Q tienen un autovalor común.*

Demostración. Supongamos que P y Q no tienen un autovalor común. Luego los polinomios característicos $p(\lambda) = \det(P - \lambda I)$ de P y $q(\lambda) = \det(Q - \lambda I)$ de Q son primos entre sí, por tanto, existen dos polinomios $a(\lambda)$ y $b(\lambda)$ tales que

$$a(\lambda)p(\lambda) + b(\lambda)q(\lambda) = 1$$

Sea $h(\lambda) = a(\lambda)p(\lambda)$. Se tiene

$$1 - h(\lambda) = b(\lambda)q(\lambda)$$

Entonces, por el teorema de Cayley-Hamilton 1.2.2, resulta al evaluar P en $h(P) = a(P)p(P) = 0$, y Q en $1 - h(Q) = b(Q)q(Q)$ entonces $h(Q) = 1$.
Por otro lado, como

$$\begin{aligned} p(\lambda) &= r_n\lambda^n + \cdots + r_1\lambda + r_0 \qquad a(\lambda) = s_m\lambda^n + \cdots + s_1\lambda + s_0 \\ p(P) &= r_nP^n + \cdots + r_1P + r_0 \qquad a(Q) = s_mQ^n + \cdots + s_1Q + s_0 \end{aligned}$$

y por hipótesis tenemos que $PB = BQ$, entonces al multiplicar $h(P) = a(P)p(P)$ por B se tiene

$$\begin{aligned} h(P) &= a(P)p(P)B \\ &= a(P)(r_nP^nB + \cdots + r_1PB + r_0B) \quad PB = BQ \\ &= a(P)(r_nBQ^n + \cdots + r_1BQ + r_0B) \\ &= a(P)BP(Q) \\ &= Ba(Q)P(Q) \\ &= Bh(Q) \end{aligned}$$

así

$$h(P)B = Bh(Q),$$

resulta $B = 0$, lo que es contradictorio con la hipótesis. □

El lema nos asegura, por tanto, que $A^* - \mu I$ y $-A$ tienen que tener un autovalor común. El teorema siguiente nos garantiza mediante ciertas condiciones que todo autovalor de $\mathbf{F}$ es distinto de 0 y, entonces, $\mathbf{F}$ es inversible.

Teorema 3.4.1. *Sea $A \in M(n, \mathbf{R})$ y sea $\sigma(A) = \lambda_1, ..., \lambda_p$ su espectro, es decir, el conjunto de sus autovalores. Supongamos que para dos autovalores cualesquiera λ_i,λ_k, no necesariamente distintos, se verifica $\lambda_i + \lambda_k \neq 0$. Entonces, la ecuación en $B \in M(n, \mathbf{R})$*

$$A^*B + BA = -C$$

donde C es una matriz dada de $M(n, \mathbf{R})$, tiene solución única. Si $C = C^*$, entonces $B = B^*$. Además, si $C = C^*$ es definida positiva (esto es, $(Cx, x) > 0$ para todo $x \neq 0$) y los autovalores de A tienen todos parte real negativa, entonces $B = B^*$ es también definida positiva.

Demostración. Para cualquier autovalor μ del operador F definido antes se verifica que $A^* - \mu I$ y $-A$ tienen algún autovalor común, según hemos visto. Los autovalores de $A^* - \mu I$ son $\lambda_i - \mu$ para $\lambda_i \in \sigma(A)$, y los de $-A$ son $-\lambda_k$ para $\lambda_k \in \sigma(A)$. Así, las condiciones del teorema garantizan que μ no es nulo, y, por tanto, **F** es inversible y por ser la inversa única se tiene la unicidad de la solución.
Si $C = C^*$ entonces B^* es también solución de $B^*A + A^*B = -C^* = -C$ y, por la unicidad de solución, resulta $B = B^*$. □

Recordemos lo estudiado en los capítulos anteriores, sobre la forma de Jordan en 1.1.4, 1.2, 1.4 y 2.1. Sabemos que si A es una matriz $n \times n$, y existe una matriz P tal que $A = P^{-1}JA$, entonces $e^{At} = P^{-1}e^{Jt}P$ donde J tiene la forma de Jordan de A.

$$\begin{aligned}
e^{At} &= P^{-1}e^{Jt}P \quad e^{Jt} = \sum_{i=1}^{p} e^{J_i t} \\
&= P^{-1}\sum_{i=1}^{p} e^{J_i t}P \quad e^{J_i t} = e^{\lambda_i t}\sum_{j=0}^{n-1}\frac{L^j t^j}{j!} \\
&= P^{-1}\left(\sum_{i=1}^{p}\sum_{j=0}^{n-1}\frac{L^j t^j}{j!}\right)e^{\lambda_i t}P \\
&= P^{-1}P\left(\sum_{i=1}^{p}\sum_{j=0}^{n-1}\frac{L^j}{j!}t^j\right)e^{\lambda_i t} \\
&= \sum_{i=1}^{p}\sum_{j=0}^{n-1} b_{ij}(A)t^j e^{\lambda_i t} \qquad (3.4.2)
\end{aligned}$$

Teorema 3.4.2. *Sea A una matriz $n \times n$. Las siguientes afirmaciones son equivalentes:*

1. *Todos los valores propios de la matriz A tienen parte real negativa.*

2. *Para todas las matrices $C = C^T > 0$ existe una solución única $B = B^T > 0$ a la siguiente ecuación de Lyapunov:*

$$A^TB + BA = -C \qquad (3.4.3)$$

Demostración. Solo probaremos la parte solo sí.
Supongamos que A tiene valores propios con parte real negativa. Definamos

$$B = \int_0^{+\infty} e^{A^T t} C e^{At} dt \tag{3.4.4}$$

notemos que la integral en (3.4.4) está bien definida, ya que los elementos de la matriz integrando son todas combinaciones lineales de funciones de la forma $t^k e^{\alpha t}$ (esto por (3.4.2)) donde α tiene una parte real negativa, la integral existe y es finita.
Ahora verifiquemos B así como se ha definido es una solución a (3.4.3). En efecto

$$\begin{aligned}
A^T B + BA &= \int_0^{+\infty} A^T e^{A^T t} C e^{At} dt + \int_0^{+\infty} e^{A^T t} C e^{At} A dt \\
&= \int_0^{+\infty} \left(A^T e^{A^T t} C e^{At} + e^{A^T t} C e^{At} A \right) dt \\
&= \int_0^{+\infty} \frac{d}{dt} \left(e^{A^T t} C e^{At} \right) dt \\
&= \left[e^{A^T t} C e^{At} \right]_0^{+\infty} \\
&= -C
\end{aligned}$$

(ya que $\lim\limits_{t\to\infty} e^{A^T t} C e^{At} = 0$). Lo cual muestra que existe una matriz B satisfaciendo (3.4.3). Además, la solución es única; en efecto, supongamos que existen dos soluciones B_1 y B_2 cualesquiera. Entonces:

$$A^T B_1 + B_1 A + C = 0 \tag{3.4.5}$$

$$A^T B_2 + B_2 A + C = 0 \tag{3.4.6}$$

restando (3.4.6) de (3.4.5) obtenemos

$$A^T (B_1 - B_2) + (B_1 - B_2) A = 0$$

multiplicando la ecuación anterior a la izquierda por $e^{A^T t}$ y a la derecha por e^{At}, se tiene:

$$\begin{aligned}
e^{A^T t} \left[A^T (B_1 - B_2) + (B_1 - B_2) A \right] e^{At} &= 0 \\
e^{A^T t} A^T (B_1 - B_2) e^{At} + e^{A^T t} (B_1 - B_2) A e^{At} &= 0 \\
\frac{d}{dt} \left(e^{A^T t} (B_1 - B_2) e^{At} \right) &= 0
\end{aligned}$$

lo cual además implica al integrar lo anterior que

$$\begin{aligned}
\int_0^{\infty} \frac{d}{dt} \left(e^{A^T t} (B_1 - B_2) e^{At} \right) dt &= 0 \\
\left[e^{A^T t} (B_1 - B_2) e^{At} \right]_0^{\infty} &= 0 \\
-(B_1 - B_2) &= 0 \\
B_1 &= B_2
\end{aligned}$$

así, existe una única solución que satisface (3.4.3) □

De las consideraciones anteriores resulta fácilmente el siguiente teorema.

Teorema 3.4.3. *Consideremos la ecuación*

$$x'(t) = Ax(t) \tag{3.4.7}$$

siendo $A \in M(n, \mathbf{R})$ tal que todos sus autovalores tienen parte real negativa. Entonces la solución trivial es asintóticamente estable.

Demostración. Dado que los autovalores de A tienen parte real negativa, existe una solución definida positiva B a la ecuación de Lyapunov

$$A^T B + BA + I = 0$$

Escogemos, como una función de Lyapunov, la forma cuadrática

$$V(x) = (x, Bx) = x^T Bx$$

que es por supuesto una función definida positiva. Su derivada es

$$\begin{aligned} V'(x) &= (x', Bx) + (x, Bx') \\ &= (Ax, Bx) + (x, BAx) = x^T A^T Bx + x^T BAx \\ &= x^T (A^T B + BA)x \\ &= (x, (A^T B + BA)x) \\ &= (x, -Ix) \\ &= -\|x\|^2 \end{aligned}$$

por tanto $V'(x)$ es definida negativa. Así el sistema (3.4.7) es asintóticamente estable. □

Los otros resultados sobre la estabilidad de la solución trivial de $x'(t) = Ax(t)$, A constante, que conocemos del capitulo 2, pueden también obtenerse a través del método directo de Lyapunov, pero se llega a ellos de una forma algo más artificiosa que allí.

3.4.1. Ejemplos

Ejemplo 3.4.1. *Dada la ecuación*

$$\begin{cases} x'(t) = ax(t) + by(t) \\ y'(t) = cx(t) + dy(y) \end{cases}$$

obténgase explícitamente una función de Lyapunov siguiendo la idea del teorema 3.4.3.

Solución. *Sea*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

El sistema anterior será asintóticamente estable, si los valores propios de la matriz A tienen parte real negativa y además $V(x) > 0$ y $V'(x) < 0$.
Primero determinemos los valores propios de A.

$$\begin{aligned} \det(A - \lambda I) &= 0 \quad \textit{es decir} \\ \lambda^2 - (a+d)\lambda + (ad - bc) &= 0 \quad \textit{resulta entonces que} \\ \lambda_{1,2} &= \frac{(a+d) \pm \sqrt{(a-d)^2 + 4bc}}{2} \end{aligned}$$

Para que $\lambda_{1,2}$ tengan parte real negativa se tiene que cumplir algunos de los siguiente casos:

- $(a-d)^2 + 4bc > 0$ *y* $(a+d) < 0$
- $(a-d)^2 + 4bc = 0$ *y* $(a+d) < 0$
- $(a-d)^2 + 4bc < 0$ *y* $(a+d) < 0$

Ahora bien, se debe de encontrar una función $V(x) = (x, Bx) = x^T Bx > 0$ tal que $V'(x) = (x, (A^T B + BA)x) = x^T(A^T B + BA)x < 0$. Sabemos que la ecuación de Lyapunov es dada por $A^T B + BA = -C$, donde C simétrica, definida positiva. Tomando $C = I$ y

$$B = \begin{pmatrix} p & q \\ q & r \end{pmatrix}$$

entonces

$$A^T B + BA = -C = -I$$

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} p & q \\ q & r \end{pmatrix} + \begin{pmatrix} p & q \\ q & r \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

al hacer el producto y suma de matrices se obtiene

$$\begin{pmatrix} 2ap + 2cq & aq + bp + cr + dq \\ aq + bp + cr + dq & 2bq + 2dr \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

de donde, se tienen las ecuaciones

$$2ap + 2cq = -1 \tag{3.4.8}$$

$$aq + bp + cr + dq = 0 \tag{3.4.9}$$

$$2bq + 2dr = -1 \tag{3.4.10}$$

despejando q de (3.4.8) se tiene

$$q = -\frac{1 + 2ap}{2c}$$

sustituyendo q en (3.4.10) resulta que

$$r = \frac{b - c + 2abp}{2cd}$$

sustituyendo q y r en (3.4.9) se obtiene

$$p = \frac{ad - bc + c^2 + d^2}{2(abc + bcd - a^2 d - ad^2)}$$

ahora, sustituyendo p en q y en r obtenemos

$$q = \frac{-(ac + bd)}{2(abc + bcd - a^2 d - ad^2)} \quad r = \frac{a^2 + ad + b^2 c - bc^2}{2c(abc + bcd - a^2 d - ad^2)}$$

denotemos por

$$\begin{aligned} k &= ad - bc + c^2 + d^2 \qquad l = -(ac + bd) \\ m &= a^2 + ad + b^2 c - bc^2 \qquad n = abc + bcd - a^2 d - ad^2 \end{aligned}$$

así,

$$p = \frac{k}{n} \quad q = \frac{l}{n} \quad r = \frac{m}{2cn}$$

por lo que la matriz B que se busca es

$$B = \begin{pmatrix} k/n & l/n \\ l/n & m/2cn \end{pmatrix}$$

resulta, entonces que

$$V(x) = (x, Bx) = x^T Bx = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} k/n & l/n \\ l/n & m/2cn \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$V(x) = \begin{pmatrix} \frac{k}{n}x + \frac{l}{n}y & \frac{l}{n}x + \frac{m}{2cn}y \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$V(x) = \frac{k}{n}x^2 + 2\frac{l}{n}xy + \frac{m}{2cn}y^2 > 0$$

Veamos que $\det(B - \lambda I) = 0$ *para calcular los valores propios de la matriz B.*

$$\begin{aligned} \det(B - \lambda I) &= 0 \quad \textit{es decir} \\ \lambda^2 - \left(\frac{l}{2k} + \frac{n}{2ck}\right)\lambda + \left(\frac{nl}{4ck^2} - \frac{m^2}{4k^2}\right) &= 0 \quad \textit{resulta entonces que} \\ \lambda_{1,2} &= \frac{\left(\frac{l}{2k} + \frac{n}{2ck}\right) \pm \sqrt{\left(\frac{l}{2k} - \frac{n}{2ck}\right)^2 + \frac{m^2}{k^2}}}{2} \end{aligned}$$

si los autovalores de la matriz B son positivos quiere decir que la matriz B es simétrica y definida positiva, por tanto el sistema es estable. Y además se cumple que $V'(x) < 0$ ya que como definimos $A^T B + BA = -C = -I$ entonces $V'(x) = -(x, x) = -x^T x = -(x^2 + y^2) < 0$.

Por lo tanto, el sistema dado es asintóticamente estable.

Ejemplo 3.4.2. *Dado el siguiente sistema*

$$\begin{aligned} x'(t) &= y(t) \\ y'(t) &= -4x(t) - 5y(t) \end{aligned}$$

Estudiar que tipo de estabilidad presenta.

Solución. *Sea*

$$A = \begin{pmatrix} 0 & 1 \\ -4 & -5 \end{pmatrix}$$

El sistema dado será asintóticamente estable , si los valores propios de la matriz A tienen parte real negativa y si $V(x) > 0$ y $V'(x) < 0$. Determinando los autovalores de la matriz A.

$$\begin{aligned} \det(A - \lambda I) &= 0 \\ \lambda^2 + 5\lambda + 4 &= 0 \end{aligned}$$

$$\lambda_1 = -1 \qquad \lambda_2 = -4$$

Debemos encontrar una función de Lyapunov $V(x) = (x, Bx) = x^T Bx > 0$ y $V'(x) = (x, (A^T B + BA)x) = x^T(A^T B + BA)x < 0$. Sea $C = I$, donde $A^T B + BA = -C$.
Definamos

$$B = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

así

$$\begin{pmatrix} 0 & -4 \\ 1 & -5 \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} + \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -4 & -5 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\begin{pmatrix} -8b & a - 5b - 4c \\ a - 5b - 4c & 2b - 10c \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

de donde se obtienen las siguientes ecuaciones

$$\begin{aligned} -8b &= -1 \\ a - 5b - 4c &= 0 \\ 2b - 10c &= -1 \end{aligned}$$

así $a = \frac{9}{8}$, $b = \frac{1}{8}$, $c = \frac{1}{8}$ resulta entonces que

$$B = \begin{pmatrix} 9/8 & 1/8 \\ 1/8 & 1/8 \end{pmatrix}$$

luego

$$V(x) = (x, (A^T B + BA)x) = x^T(A^T B + BA)x = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 9/8 & 1/8 \\ 1/8 & 1/8 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \frac{9}{8}x^2 + \frac{1}{4}xy + \frac{1}{8}y^2 > 0$$

si B tiene los autovalores positivos entonces estable.
Calculando los autovalores de la matriz B tenemos que

$$\begin{aligned} \det(B-\lambda I) &= 0 \\ \lambda^2 - \frac{5}{4}\lambda + \frac{1}{8} &= 0 \\ \lambda_{1,2} &= \frac{5 \pm \sqrt{17}}{8} \end{aligned}$$

como los autovalores son positivos, quiere decir que la matriz B es simétrica y definida positiva por tanto el sistema dado es estable y asintóticamente estable, ya que $V'(x) < 0$.

Capítulo 4

Estabilidad en Ingeniería

4.1. ¿Por qué un sistema estable es útil en ingeniería?

Supongamos que tenemos un sistema de ecuaciones diferenciales lineales no homogéneo

$$x'(t) = Ax(t) + b(t)$$

donde $A \in M(\mathbb{R}, n)$ y supongamos que el sistema homogéneo asociado

$$x'(t) = Ax(t)$$

es asintóticamente estable. Entonces toda solución del sistema homogéneo

$$\Phi(t) = x_h(t) = e^{At}C, \quad C \in \mathbb{R}$$

verifica que

$$\lim_{t\to\infty} x_h(t) = 0$$

Ahora bien, toda solución del sistema no homogéneo es de la forma

$$x(t) = x_h(t) + x_p(t)$$

donde $x_p(t)$ es una solución particular del sistema no homogéneo. Si tomamos límites cuando t a infinito, tenemos que

$$x(t) \backsimeq x_p(t)$$

es decir, para tiempos grandes (aquí lo de grande depende de cada sistema) la solución del sistema no autónomo es básicamente la solución particular del mismo y la parte de la solución correspondiente al sistema homogéneo se va reduciendo con el tiempo. En ingeniería a la función $b(t)$ se le llama entrada del sistema y $x_p(t)$ es la salida del mismo. Si el sistema es estable, al variar la entrada, varía la salida sin que la parte homogénea intervenga en el proceso. Esto es lo que ocurre en la mayoría de los sistemas lineales utilizados en las ciencias experimentales, como en circuitos eléctricos o vibraciones mecánicas. Muchos de los problemas de los modelos de aplicación en ingeniería estudian la estabilidad ocupando la transformada de Laplace.

4.2. Transformadas de Laplace

Definición 4.2.1 (La transformada de Laplace). *Dada una función $f(t)$ definida para toda $t \geqslant 0$ la transformada de Laplace de f es la función F definida como sigue:*

$$F(s) = \mathcal{L}f(t) = \int_0^{\infty} e^{-st} f(t)dt$$

para todo valor de s en los cuales la integral impropia converge.

Recuérdese que una integral impropia en un intervalo infinito está definida como el límite de la integral en el intervalo acotado; esto es,

$$\int_a^{\infty} g(t)dt = \lim_{b\to\infty} \int_a^b g(t)dt$$

Si el límite en la integral existe, entonces se dice que la integral impropia converge; de otra manera diverge o no existe.

Definición 4.2.2 (Función de transferencia). *La función de transferencia de un sistema descrito mediante una ecuación diferencial lineal e invariante con el tiempo se define como el cociente de la transformada de Laplace de la salida (función de respuesta del sistema) y la transformada de Laplace de la entrada (función excitación), bajo la suposición de que todas las condiciones iniciales son cero, es decir, se considera que el sistema bajo estudio está en reposo.*

$$R(s) \longrightarrow \boxed{G(s)} \longrightarrow C(s)$$

Figura 4.1: Representación de la función de transferencia (o ganancia) de un sistema

Para el sistema ilustrado en la Figura 4.1, la salida $C(s)$ es el producto de la ganancia $G(s)$ y la entrada $R(s)$, lo que implica que $C(s) = R(s)G(s)$; la ganancia del sistema es entonces $G(s) = \dfrac{C(s)}{R(s)}$; para sistemas descritos por ecuaciones diferenciales lineales e invariantes en el tiempo, tal como:

$$a_0y^n + a_1y^{n-1} + \cdots + a_{n-1}y' + a_ny = b_0x^m + b_1x^{m-1} + \cdots + b_{m-1}x' + b_mx \qquad (n \geqslant m)$$

donde y es la salida del sistema y x es la entrada. La función de transferencia (o ganancia) viene dada por:

$$Funci\acute{o}n\ de\ transferencia = G(s) = \frac{\mathcal{L}[salida]}{\mathcal{L}[entrada]} = \frac{Y(s)}{X(s)} = \frac{b_0s^m + b_1s^{m-1} + \cdots + b_{m-1}s + b_m}{a_0s^n + a_1s^{n-1} + \cdots + a_{n-1}s + a_n}$$

4.2.1. Función de transferencia en lazo cerrado

Para el sistema que aparece en la Figura 4.2, la salida $C(s)$ y la entrada $R(s)$ se relacionan del modo siguiente:

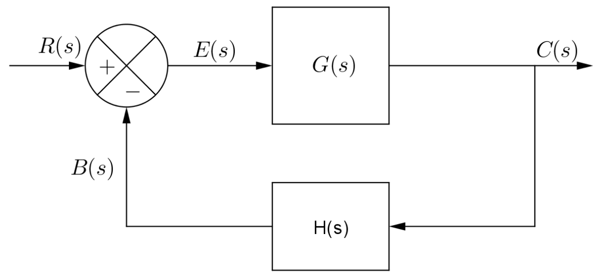


Figura 4.2: Sistema en lazo cerrado.

$$\begin{aligned} C(s) &= G(s)E(s) \\ E(s) &= R(s) - B(s) \\ &= R(s) - H(s)C(s) \end{aligned}$$

Si se elimina $E(s)$ de estas ecuaciones, se obtiene

$$C(s) = G(s)[R(s) - H(s)C(s)]$$

o bien,

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 + G(s)H(s)} \tag{4.2.1}$$

La función de transferencia que relaciona $C(s)$ con $R(s)$ se denomina función de transferencia en lazo cerrado. Esta función de transferencia relaciona la dinámica del sistema en lazo cerrado con la dinámica de los elementos de las trayectorias directa y de realimentación.

A partir de la Ecuación anterior, $C(s)$ se obtiene mediante

$$C(s) = \frac{G(s)}{1 + G(s)H(s)} R(s)$$

Por tanto, la salida del sistema en lazo cerrado depende claramente tanto de la función de transferencia en lazo cerrado como de la naturaleza de la entrada.

La ecuación característica de la ecuación 4.2.1 queda definida como:

$$f(s) = 1 + G(s)H(s) = 0 \quad G(s)H(s) = \frac{kQ(s)}{P(s)} \Rightarrow f(s) = P(s) + kQ(s)$$

donde $Q(s)$ es un polinomio de grado n de la ecuación característica en s y sus raíces son llamadas ceros, $P(s)$ es un polinomio de grado m y sus raíces son llamadas polos, el cual analizaremos con detalle a continuación.

Un sistema lineal invariante en el tiempo es estable si:

a) Ante una entrada acotada responde con unan salida acotada.

b) Si todos los polos de la función de transferencia están en el semiplano negativo de s, es decir, tienen la parte real negativa.

4.3. Métodos para determinar la estabilidad

Nuestro problema ahora es determinar la estabilidad de un sistema de control, para ello existen varios métodos para determinar la estabilidad de un sistema realimentado, estos involucran las raíces de la ecuación característica. Los métodos más utilizados para estudiar la estabilidad de sistemas de control son:

1. Criterio de Routh–Hurwitz.

2. Criterio de Nyquist.

3. Método de Diagrama de Bode.

Es evidente que para el análisis de los sistemas de control, se presentan métodos alternativos que resuelven el mismo problema, el diseñador simplemente selecciona el método a utilizar que considere que es la mejor herramienta, dependiendo de la situación particular que enfrenta. En lo particular, preferimos los dos primeros, sin desmeritar y quitar la importancia al diagrama de Bode.

Criterio de Routh–Hurwitz

Es un método algebraico que ofrece información sobre la estabilidad absoluta de un sistema lineal invariante en el tiempo que tiene una ecuación característica con coeficientes constantes.

Tabulación de Routh

Para construir la tabulacion de Routh se basa en ordenamiento de los coeficientes de la ecuación característica $f(s) = a_0 s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n = 0$ tomando una lista o arreglo como sigue a continuación:

$$\begin{array}{cl}
s^n & a_0 \;\; a_2 \;\; a_4 \;\; \ldots \\
s^{n-1} & a_1 \;\; a_3 \;\; a_5 \ldots \\
s^{n-2} & b_1 \;\; b_2 \;\; b_3 \ldots \\
s^{n-3} & c_1 \;\; c_2 \;\; c_3 \ldots \\
\vdots & \\
s^0 & d_1
\end{array}$$

Una regla nemotécnica consiste en tomar el inicial y luego saltar un coeficiente y seleccionar el otro que sigue hasta que se agoten los coeficientes, luego se empieza por el siguiente que no fue seleccionado en el paso anterior y se repite el mismo proceso para completar las dos filas principales. Los coeficientes $b_1,\ b_2,\ b_3$, etc., se evalúan del modo siguiente

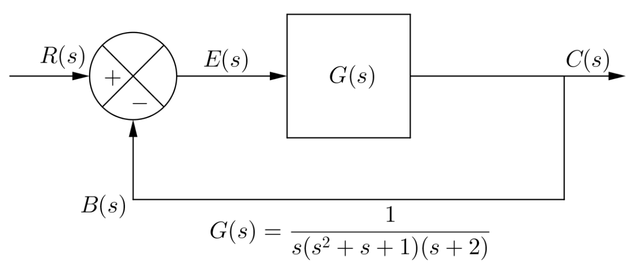
$$b_1 = \frac{a_1a_2 - a_0a_3}{a_1} \quad b_2 = \frac{a_1a_4 - a_0a_5}{a_1} \quad b_3 = \frac{a_1a_6 - a_0a_7}{a_1}$$

La evaluación de las b continúa hasta que todas las restantes son cero. Se sigue el mismo patrón de multiplicación cruzada de los coeficientes de los dos renglones anteriores al evaluar las c, las d, las e, etc. Las raíces de la ecuación característica estarán en el semiplano izquierdo, si todos los elementos de la primera columna tienen el mismo signo, así mismo el número de cambios de signos en los elementos de la primera columna equivale al número de raíces con parte real positiva o en el semiplano derecho. El criterio establece que para que un sistema sea estable, requiere que no haya cambios de signos en la primera columna de la tabulación, éste es un requisito necesario y suficiente.

4.3.1. Ejemplo

Ejemplo 4.3.1. *Sea el sistema, determine si es estable o no.*

Primero se debe hallar la función de transferencia de lazo cerrado:



$$\frac{C(s)}{R(s)} = \frac{G(s)}{1+G(s)H(s)} = \frac{1}{s^4+3s^3+3s^2+2s+1}$$

La ecuación característica del sistema es:

$$s^4+3s^3+3s^2+2s+1=0$$

Se construye la tabla, con los coeficientes: $\{a_0 = 1,\ a_1 = 3\ a_2 = 3\ a_3 = 2\ a_4 = 1\}$.

$$\begin{array}{ll}
s^4 & 1\ 3\ 1 \\
s^3 & 3\ 2 \\
s^2 & b_1\ b_2 \\
s^1 & c_1 \\
s^0 & d_1
\end{array}$$

$$b_1 = -\frac{3 \times 3 - 2 \times 1}{3} = \frac{7}{3} \qquad b_2 = -\frac{3 \times 1 - 1 \times 0}{3} = 1$$

$$c_1 = \frac{2b_1 - 3b_2}{b_1} = \frac{2\left(\frac{7}{3}\right) - 3}{\frac{7}{3}} = \frac{5}{7} \qquad d_1 = \frac{c_1 b_2 - c_2 b_1}{c_1} = \frac{\frac{5}{7}}{\frac{5}{7}} = 1$$

La tabla queda de la siguiente forma:

$$\begin{array}{ll}
s^4 & 1 \ 3 \ 1 \\
s^3 & 3 \ 2 \\
s^2 & \frac{7}{3} \ 1 \\
s^1 & \frac{5}{7} \\
s^0 & 1
\end{array}$$

Y como en la primera columna no hay cambios de signos en los coeficientes, se puede asegurar que el sistema es "estable", debido a que no posee polos positivos o con parte real positiva.

El método del Criterio de Nyquist y el método de Diagrama de Bode son menos utilizados para analizar la estabilidad. Por lo que en este trabajo no se estudiarón sólo nos enfocamos a estudiar el criterio de Routh–Hurwitz que es el más utilizado para determinar la estabilidad de un lazo cerrado o si una ecuación polinómicas posee raíces positivas sin resolverla para determinar támbien su estabilidad. Según el Lic. Ricardo Cortes en Ingeniería no estudian la estabilidad por medio del método de Lyapunov puesto que es mucho más complicado, así prefieren estudiar la estabilidad de una forma más sencilla.

Conclusiones

Se presentan las principales conclusiones, a las que se ha llegado después del desarrollo de este proyecto y que resumen los aspectos mas sobresalientes de la teoría estudiada:

1. Estudiar la Teoría de Estabilidad de sistemas de ecuaciones diferenciales ordinarias es muy importante, ya que es un tema que no sólo se aplica en matemáticas sino que en diversas áreas como la Física, la Economía y la Ingeniería, etc., por lo que requiere de conocimientos vistos en cálculo, algebra lineal y topología. Y así mejorar la comprensión y aplicación de muchos conceptos;, este tema una opción para incluirse en los estudios de licenciatura.

2. La estabilidad es una propiedad cualitativa de los sistemas de ecuaciones diferenciales a la que se le considera la más importante de todas.

3. La forma canónica de Jordan, simplifica los cálculos para encontrar la exponencial de una matriz, ya que esta es de gran utilidad para determinar la estabilidad de un sistema lineal, conociendo los autovalores.

4. Para determinar la estabilidad de un sistema lineal no homogéneo basta con estudiar la estabilidad del sistema lineal homogéneo.

5. El criterio de Routh-Hurwitz permite determinar la estabilidad asintótica de un sistema lineal con coeficientes constantes; aunque, como se requiere calcular el determinante esto puede resultar muy complicado para matrices de gran dimensión.

6. El teorema de Lyapunov se refiere a una reducción del estudio de la estabilidad de un problema todavía lineal pero no de coeficientes constantes, y así se requiere de un lema que es muy fecundo y sirve también para tratar algunos casos de estabilidad de ecuaciones no lineales. Además requiere de ciertas condiciones especiales adicionales, generalmente relacionadas con acotación y convergencia.

7. Después de realizar el análisis de sistemas de ecuaciones diferenciales no lineales mediante el uso de las funciones de Lyapunov, se observa que, se requiere de una gran habilidad para la obtención de una función de Lyapunov.

8. En el trabajo se ha mostrado en general el uso de las funciones de Lyapunov, aunque fueron poco los ejemplos utilizados. Sin embargo, el tema es muy amplio y existen diversas aplicaciones para estas.

9. Un trabajo a futuro, podría ser estudiar la Teoría de Control que requiere haber estudiado antes la Teoría de Estabilidad.

10. Se puede ver como trabajo a futuro desarrollar funciones para sistemas más especializados, no sólo sistemas físicos, sino también en ciencias en las que se pueda utilizar.

Bibliografía

[1] De Guzman, M. Ecuaciones Diferenciales Ordinarias. Teoría de estabilidad y control. Ed. Alhambra, Madrid, 1975.

[2] Lizana, M. Ecuaciones Diferenciales Ordinarias. Ed. Mérida – Venezuela, Junio de 2000.

[3] Rudin, Walter. PRINCIPIOS de ANÁLISIS MATEMÁTICO. 3a. Edición 1980

[4] Piza Volio, Eduardo. Introducción al análisis real en una variable. 1. ed. San José, C.R. : Editorial de la Universidad de Costa Rica, 2003.

[5] Imaz, Carlos. Ecuaciones Diferenciales Ordinarias. Ed. Limusa-Wiley, México.

[6] Fernández Pérez, Carlos(1992). ecuaciones diferenciales-I: ecuaciones lineales. Madrid: Ediciones Pirámide.

[7] Fernández Pérez, Carlos(1996). ecuaciones diferenciales-II: ecuaciones no lineales. Madrid: Ediciones Pirámide.

[8] Dórame Velásquez, Maximino. TEORIA Y APLICACION DE SISTEMAS REGULADORES A ESTABLES AUTO-AJUSTADOS. Hermosillo, Sonora, México, 1996

[9] BARBASHIN, E. A, INTRODUCTION TO THE THEORY. PRINTED IN THE NETHERLANDS BY NEDERLA.NDSE BOEKDRUK INDUSTRIE N.V. - DEN BOSCH OF STABILITY.

[10] Wassim M. Haddad VijaySekhar Chellaboina (2007). Nonlinear Dynamical Systems and Control. A Lyapunov-Based Approach. United States of América: Princeton University Press.

[11] Kuo, Benjamin C. SISTEMAS DE CONTROL AUTOMÁTICO. 7a. Edición.

[12] Ogata, Katsuhiko. Ingeniería de control moderna 5a. Edición. PEARSON EDUCACIÓN, S.A., Madrid, 2010.

[13] Hernández Gaviño, Ricardo. Introducción a los sistemas de control: Conceptos, aplicaciones y simulación con MATLAB. Primera edición. PEARSON EDUCACIÓN, México, 2010. **Fuentes de Internet**

[14] ECUACIONES DIFERENCIALES AUTÓNOMAS Y ESTABILIDAD DE LOS PUNTOS DE EQUILIBRIO
https://www.fing.edu.uy/ẽleonora/Recopilacion/Archivos/NotasEnsenanza/Calcu2EcDifComplEstabilidad.pdf

[15] Capítulo 1 Matriz fundamental
https://eva.fing.edu.uy/pluginfile.php/95279/Cap3_Matriz_fundamental.pdf

[16] Estabilidad de Lyapunov
http://dea.unsj.edu.ar/snl/snl_cap6_1.pdf

[17] Funciones de Lyapunov y Algunas Aplicaciones http://cdigital.uv.mx/bitstream/123456789/35838/1/yepezriveramario.pdf

Apéndice A

Existencia, unicidad y prolongabilidad de soluciones

A.1. El problema de Cauchy. Solución global

El primer teorema de existencia se suele denominar teorema de Picard-Lindelöf. Antes de presentar este teorema veremos cómo se reduce el problema que nos ocupa a la determinación de un punto fijo para una cierta transformación integral.

El *problema de valores iniciales o problema de Cauchy* consiste en lo siguiente:

Se da una función $f : [t_0, t_1] \times \mathbf{R}^n \to \mathbf{R}^n$ *continua y se da un punto* $\xi^0 \in \mathbf{R}^n$*. Se pide hallar, si es posible, una función*

$$x : [t_0, t_1] \to \mathbf{R^n} \ \ de \ \ \mathcal{C}^1([t_0, t_1], \mathbf{R^n})$$

tal que

$$x'(t) = f(t, x(t)) \ \ para \ \ t \in [t_0, t_1]$$

tal que

$$x(t_0) = \xi^0$$

Se comprueba de inmediato que el problema de Cauchy es equivalente a resolver la ecuación integral

$$x(t) = \xi^0 + \int_{t_0}^{t} f(s, x(s)) ds$$

A continuación presentamos dos teoremas. El primero de ellos es el teorema de Picard-Lindelöf resuelve el problema de la unicidad de la solución si la fución f verifica la condición de Lipschitz.

Teorema A.1.1 (Picard-Lindelöf). *Consideremos el problema*

$$\begin{cases} x'(t) = f(t, x(t)) \quad para \quad t \in [t_0, t_1] \\ x(t_0) = \xi^0 \end{cases}$$

Sea $f : [t_0, t_1] \times \mathbf{R}^n \to \mathbf{R}^n$ continua y tal que satisface la condición de Lipschitz siguiente:

$$|f(t, x^1) - f(t, x^2)| \leqslant L|x^1 - x^2|$$

para $t \in [t_0, t_1]$ y $x^1, x^2 \in \mathbf{R^n}$, siendo L una constante. Entonces el problema propuesto tiene una única solución.

Demostración. En el espacio vectorial $X = \mathcal{C}([t_0, t_1], \mathbf{R^n})$ definimos la norma siguiente:
Para $x \in X$ ponemos:

$$\|x\| = \sup\{e^{-K(t-t_0)}|x(t)| : t \in [t_0, t_1]\}$$

siendo $K > L$ una constante fija. Con esta norma el espacio X es un espacio de Banach.
Definimos, para $x \in X$:

$$Tx(t) = \xi^0 + \int_{t_0}^{t} f(s, x(s))ds$$

Entonces, si $x_1, x_2 \in X$, se tiene:

$$\begin{aligned} e^{-K(t-t_0)}|Tx_1(t) - Tx_2(t)| &\leqslant \int_{t_0}^{t} e^{-K(t-t_0)}|f(s, x_1(s)) - f(s, x_2(s))|ds \leqslant \\ &\leqslant L\|x_1 - x_2\| \int_{t_0}^{t} e^{-K(t-s)}ds \leqslant \\ &\leqslant \left(\frac{L}{K}\right)\|x_1 - x_2\| \end{aligned}$$

Por tanto

$$\|Tx_1(t) - Tx_2(t)\| \leqslant \frac{L}{K}\|x_1 - x_2\|$$

siendo $L/K < 1$. Así, T es una transformación contractiva y existe un único punto fijo x que es la única solución del problema. □

A.2. Prolongabilidad de soluciones

Para terminar se considera el problema de la prolongación de una solución de un problema de Cauchy. Los teoremas anteriores aseguran la existencia y unicidad o la existencia de la solución.
Tratamos de resolver la siguiente cuestión. ¿De qué formas puede suceder que una solución no sea estrictamente prolongable a la derecha?
Es decir, nos proponemos el problema (a la derecha)

$$(P)\begin{cases} x'(t) = f(t, x(t)) \\ x(t_0) = \xi_0 \end{cases}$$

donde f está definida en un cierto conjunto A de $\mathbf{R} \times \mathbf{R^n}$, $(t_0, \xi^0) \in A$ y suponemos que existe alguna solución a la derecha y, por consiguiente, alguna solución (x, I) no prolongable estrictamente a la derecha. ¿Cómo es entonces el comportamiento de $x(t)$ cuando t se acerca al extremo derecho del intervalo I?

Teorema A.2.1. *Se considera el problema*

$$(P)\begin{cases} x'(t) = f(t, x(t)) \\ x(t_0) = \xi_0 \end{cases}$$

donde f está definida y es continua en un conjunto K compacto de $\mathbf{R} \times \mathbf{R^n}$, $(t_0, \xi^0) \in K$ y suponemos que existe alguna solución a la derecha de t_0 y, por tanto, alguna solución (x, I) no prolongable estrictamente a la derecha. Entonces, $I = [t_0, t_1]$ con $t_1 < \infty$, y en este caso $(t_1, x(t_1))$ pertenece a la frontera $Fr(K)$ de K.

Es decir, se excluye $I = [t_0, t_1)$ con $t_1 < \infty$ y, además, resulta que la gráfica de x acaba en un punto de la frontera de K.

Cuando el conjunto de definición de f es un abierto D, la situación es más complicada.

Teorema A.2.2. *Sea f definida y continua en un conjunto abierto D de $\mathbf{R} \times \mathbf{R^n}$. Sea $(t_0, \xi^0) \in D$ y se considera el problema:*

$$(P)\begin{cases} x'(t) = f(t, x(t)) \\ x(t_0) = \xi_0 \end{cases}$$

Sea x una solución no prolongable estrictamente a la derecha, y sea I su intervalo de definición. Entonces $I = [t_0, t_1)$ con $t_1 \leqslant \infty$ y x es tal que su gráfica, que está contenida en D, tiene todos sus puntos límites para $t \uparrow t_1$ en la frontera de D. Si D no es acotado se considera que el punto del infinito es de la frontera de D. (Se considera $\mathbf{R} \times \mathbf{R^n} \cup \{\infty\}$ con la topología de la compactificación por adición de un punto.)