

UNIVERSIDAD DE EL SALVADOR.
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA.
ESCUELA DE MATEMÁTICA.



Universidad de El Salvador
Hacia la libertad por la cultura

PROYECTO DE GRADO:
**MÉTODOS NUMÉRICOS PARA PROBLEMAS DE VALOR
DE FRONTERA.**
PREVIO A LA OBTENCIÓN DEL TÍTULO DE:
LICENCIADOS EN MATEMÁTICA.

Estudiantes: David Omar Espinoza Cortez.
Walter Alexander Reyes Arias.

Asesor: MSc. Carlos Ernesto Gámez Rodríguez.

Ciudad Universitaria, 2 de junio de 2015.

AUTORIDADES UNIVERSITARIAS PERIODO 2011-2015

UNIVERSIDAD DE EL SALVADOR

RECTOR: ING. MARIO ROBERTO NIETO LOVO **VICERRECTORA**
ACADÉMICA: MAESTRA ANA MARÍA GLOWER DE ALVARADO
VICERRECTOR ADMINISTRATIVO: MAESTRO ÓSCAR NOÉ NAVARRETE
SECRETARIA GENERAL: DRA. ANA LETICIA DE AMAYA **DEFENSORA**
DE LOS DERECHOS UNIVERSITARIOS: LICDA. CLAUDIA MARÍA
MELGAR DE ZAMBRANA **FISCAL:**LIC. FRANCISCO CRUZ LETONA

FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA

DECANO: M.SC. MARTÍN ENRIQUE GUERRA CÁCERES **VICEDACANO:** LIC.
RAMÓN ARÍSTIDES PAZ SÁNCHEZ **SECRETARIO:** LIC. CARLOS ANTONIO
QUINTANILLA APARICIO

ESCUELA DE MATEMÁTICA

DIRECTOR: DR. JOSÉ NERYS FUNES TORRES **SECRETARIA:** LICDA. ALBA
IDALIA CÓRDOVA CUELLAR

Índice

1. Introducción.	7
2. Objetivos.	10
2.1. Objetivo General.	10
2.2. Objetivos Específicos.	10
3. Antecedentes y justificación.	11
3.1. Antecedentes.	11
3.2. Justificación.	15
4. Planteamiento del problema.	16
5. Metodología.	18
6. PRELIMINARES.	20
6.1. Cálculo.	20
6.2. Orden de aproximación $\mathcal{O}(h^n)$	22
6.3. Teoría de Álgebra Lineal.	23
6.4. Teoría de Ecuaciones Diferenciales.	23
6.5. Teoría de Análisis Funcional.	26
6.5.1. Espacios Normados.	26
6.5.2. Conjuntos Convexos.	31
6.6. Método de Newton Raphson y Runge Kutta.	31
6.6.1. Método de Newton.	31
6.6.2. Método de Newton para sistemas de ecuaciones no lineales.	34
6.6.3. Método de Runge-Kutta.	36
7. Eigenfunciones y problemas con valores en la frontera.	38
7.1. Problemas de Sturm-Liouville y desarrollo en eigenfunciones.	38
7.1.1. Problemas de Sturm-Liouville.	38
7.1.2. Eigenfunciones ortogonales.	40
7.1.3. Desarrollo en eigenfunciones.	41
7.2. Aplicaciones de las series de eigenfunciones.	42
7.2.1. Vibraciones longitudinales de barras.	42
7.2.2. Vibraciones transversales de barras.	44
7.3. Soluciones Periódicas estacionarias Y frecuencias naturales.	45
7.3.1. Vibraciones forzadas y resonancia.	46
7.3.2. Frecuencias naturales de la viga.	47
7.3.3. Oscilaciones de temperatura subterránea.	49
7.4. Problemas en coordenadas cilíndricas.	51
7.4.1. Problema de Sturm-Liouville singular.	53
7.4.2. Serie de Fourier-Bessel.	56
7.4.3. Coeficientes de Fourier-Bessel.	58

8. APROXIMACIÓN A LOS EIGENVALORES.	61
8.1. Álgebra Lineal y Eigenvalores.	61
8.2. Matrices Ortogonales y Transformaciones Similares.	62
8.3. Método de la potencia.	64
8.4. Método de Householders.	69
8.5. Método de deflación.	72
8.6. Descomposición en Valores Singulares.	78
9. MÉTODOS NUMÉRICOS DE PROBLEMAS DE VALOR DE FRONTERA.	85
9.1. Método del disparo lineal.	85
9.2. Método del disparo para problemas no lineales.	92
9.3. Métodos de diferencias finitas para los problemas lineales.	99
9.4. Métodos de diferencias finitas para los problemas no lineales.	106
9.5. El método de Rayleigh-Ritz.	111
9.6. B-Spline Básicas.	121
10. EXPERIMENTACIÓN NUMÉRICA.	131
11. Cronograma de Actividades.	141
12. Bibliografías.	142

1. Introducción.

En muchas áreas de las Ciencias Básicas e Ingeniería existen problemas donde es necesario encontrar la solución de una Ecuación Diferencial Ordinaria. Estas describen fenómenos que cambian frecuentemente.

Comúnmente, una solución de interés está determinada especificando los valores de todas sus componentes en un punto $x = a$. Esto es un Problema de Valor Inicial (PVI). Sin embargo, en muchas ocasiones, una solución está determinada en más de un punto. Un problema de este tipo es denominado Problema de Valor de Frontera (PVF). Un PVF muy trabajado en la actualidad son los de segundo orden, es decir, los PVF que se especifican en dos puntos:

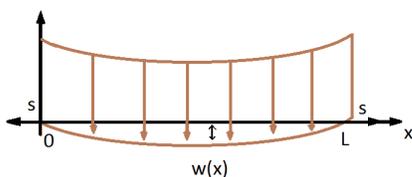
$$\begin{cases} y'' &= f(x, y, y') \\ y(a) &= \alpha \\ y(b) &= \beta. \end{cases}$$

Muchos de los problemas que realmente se presentan en la ingeniería no se pueden resolver directamente, puesto que sólo algunos tipos de ecuaciones diferenciales admiten soluciones en términos de funciones elementales.

Un problema común en ingeniería civil es el que se relaciona con la deflexión de una biga de sección transversal rectangular sujeta a una carga uniforme, mientras sus extremos están soportados de modo que no experimentan deflexión alguna.

La ecuación diferencial que aproxima la situación física es de la forma:

$$\frac{d^2w(x)}{dx^2} = \frac{S}{EI}w(x) + \frac{q}{2EI}x(x - l)$$



donde $w(x)$ es la deflexión a una distancia x desde el extremo izquierdo de la viga, y l , q , E , S e I representan, respectivamente, la longitud de la viga, la intensidad de la carga uniforme, el módulo de elasticidad, el esfuerzo en los extremos y el momento

central de inercia. Esta ecuación diferencial tiene asociadas dos condiciones de frontera dadas por la suposición de que no ocurre deflexión alguna en los extremos de la viga.

$$w(0) = w(l) = 0.$$

Cuando la viga tiene un espesor uniforme el producto EI es constante y la solución exacta se obtiene fácilmente. No obstante en muchas aplicaciones el espesor no es uniforme y, por tanto, el momento de inercia I es una función de x y se requieren métodos de aproximación.

Otros ejemplos de PVF se presentan en Física, las leyes de Newton y muchas otras se expresan como un problema de este tipo. En biología aparecen en el modelado de dinámica

de poblaciones, en la química surgen en la evaluación de las concentraciones de diversos reactivos durante una reacción, etc.

A lo largo de los años se han desarrollado técnicas para encontrar la solución analítica a Ecuaciones Diferenciales Ordinarias, y por ende, la solución de los PVF que surgen de los modelos anteriormente mencionados. Sin embargo, resulta habitual que en la mayoría de los casos no se conozca la solución analítica de la misma, y sólo estudios cualitativos de dichas ecuaciones son presentados en la literatura. Debido a esto, en la práctica, resulta impetuoso usar métodos numéricos para dar aproximaciones numéricas de la solución (aproximaciones tan buenas como se quiera).

Hoy en día, en la literatura, existen muchos métodos que ayudan a estimar la solución de un PVF de segundo orden, y es por esta razón, que este trabajo se concentrará en el estudio de algunos métodos.

Sabemos que los polinomios han sido durante mucho tiempo las funciones más utilizadas para aproximar otras funciones, principalmente a causa de sus propiedades matemáticas simples. Sin embargo, los polinomios de alto grado tienden a oscilar fuertemente y en muchos casos son susceptibles de producir aproximaciones muy pobres. Con una función spline combinamos polinomios de bajo grado por lo tanto, su oscilación es mínima de una manera tal como para obtener una función que es tan suave como sea posible en el sentido de que tiene continuidad máxima sin ser globalmente un polinomio. Las funciones spline pueden ser integradas y diferenciadas debido a ser polinomios a trozos y se pueden ser almacenados e implementado en la computadora con facilidad. Por lo tanto, las funciones spline se adaptan a métodos numéricos para obtener la solución de las ecuaciones diferenciales.

En general, el trabajo queda dividido de la siguiente manera:

Capítulo uno: se desarrollan los preliminares teóricos necesarios para abordar adecuadamente el estudio de los métodos numéricos. Se introducen resultados del Álgebra Lineal, Ecuaciones Diferenciales y Análisis Funcional. Así mismo, se introducen el método de Newton Raphson, tanto para una ecuación como para un sistema de ecuaciones no lineales, y el método de Runge Kutta, para resolver Problemas de Valor Inicial de Ecuaciones Diferenciales Ordinarias, estudiados en un curso básico de Análisis Numérico, necesarios al momento de definir los métodos numéricos.

Capítulo dos: se presentaran el estudio de forma analítica la resolución de problemas de valor de frontera, los cuales abordaremos temas como, problemas de Sturm-Liouville y desarrollo en eigenfunciones, aplicaciones de las series de eigenfunciones, soluciones periódicas estacionarias y frecuencias naturales.

Capítulo tres: se presentaran métodos numéricos para resolver PVF, como: El método del disparo para problemas lineal y no lineal, método de diferencias finitas para problemas

lineales y no lineales y el método de Rayleigh-Ritz.

Capítulo cuatro: desarrollara la construcción de las curvas spline, donde estudiaremos, splines lineales, splines cuadráticas y splines de grado superior.

2. Objetivos.

2.1. Objetivo General.

Estudiar e implementar los principales métodos numéricos para resolver Problemas de Valor en la Frontera para Ecuaciones Diferenciales Ordinarias.

2.2. Objetivos Específicos.

- Analizar métodos analíticos para resolver Problemas de Valor de Frontera.

- Estudiar e implementar el método del Disparo para ecuaciones diferenciales lineales y no lineales.

- Estudiar e implementar el método de Diferencias Finitas para problemas lineales y no lineales.

- Estudiar e implementar el método de Rayleigh-Ritz.

- Aplicación de Problemas de Valor de Frontera mediante B-splines

3. Antecedentes y justificación.

3.1. Antecedentes.

Gottfried Wilhelm Leibniz (1646-1716).



Filosofo, matemático alemán. Durante toda su vida se entregó a trabajos eruditos fue autodidacta, en el interés de la materia. En 1684 publicó la notación usual de la derivada de una variable dependiente y como: $(\frac{dy}{dx})$ y publicó así el símbolo de la integral de una función (\int). **Leibniz** descubre el método de separación de variable en 1691, la reducción de ecuaciones homogéneas a ecuaciones separables en 1691, y el procedimiento para resolver ecuaciones lineales de primer orden en 1694.

Jakob y Johann Bernoulli.

Hicieron mucho por llegar a métodos para resolver ecuaciones diferenciales y ampliar el alcance de sus aplicaciones.

Jakob Bernoulli (matemático suizo 1654-1705) resolvió la ecuación diferencial

$$y' = \sqrt[2]{\frac{a^3}{by - a^3}}$$

en 1690 y además publicó el término **integral** en el sentido moderno.

Johann Bernoulli (matemático, médico y filósofo suizo 1667-1748).

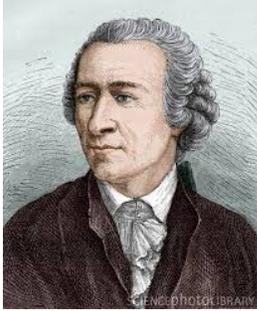


En 1694 **Johann Bernoulli** pudo resolver la ecuación diferencial $(\frac{dy}{dx} = \frac{y}{ax})$, a es una constante, aun cuando no se sabía que $d(\ln(x)) = \frac{dx}{x}$. Ambos resolvieron el problema de la **braquistocronia** (la curva de descenso más rápido) la ecuación no lineal de primer orden:

$$y[1 + (y')^2] = c$$

donde c es una constante, la cual fue resuelta por **Leibniz** y **Newton** posteriormente.

Leonhard Euler (matemático y filosofo suizo 1707-1783)



El matemático más grande del siglo *XVIII*, alumno de **Johann Bernoulli**, siendo el matemático más prolífico de todos los tiempos. Su interés particular fue el planteamiento de problemas matemáticos en mecánica. **Euler** identificó las condiciones para la exactitud de las **ecuaciones diferenciales de primer orden** en **1734-1735**, dando a la vez la solución general de las **ecuaciones lineales homogéneas con coeficientes constantes** en **1734**, de **1750 a 1751** extendió estos resultados a las **ecuaciones no homogéneas**. Aplicó con frecuencia las **series de potencias** para resolver ecuaciones diferenciales.

De 1786-1796 propuso un procedimiento numérico, hizo importantes contribuciones a las **Ecuaciones Diferenciales Parciales**, y dio el primer tratamiento sistemático del cálculo de variables.

Joseph-Louis de Lagrange (físico, matemático italiano 1736-1813).



Utilizó la cátedra dejada por **Euler**, demostró en **1762- 1765** la solución general de una ecuación diferencial homogénea de n -ésimo orden como una combinación lineal de n soluciones independientes, entre **1774 y 1775** dio un desarrollo completo del **método de Derivación de Parámetro**, publicó en **1788** un tratado de la mecánica newtoniana.

Pierre Simón de la Place destacó en el campo de la mecánica con su obra **Traité de Mecanique Celeste**.

La ecuación

$$u_{xx} + u_{yy} + u_{zz} = 0.$$

donde los subíndices denotan derivadas parciales, se conoce como: la **Ecuación de Laplace** o como la **Ecuación del Potencial**; la cual se le reconoció su utilidad en la resolución de ecuaciones diferenciales. A finales del siglo *XVIII* se habían descubierto muchos métodos elementales para resolver ecuaciones diferenciales ordinarias.

Las numerosas ecuaciones diferenciales que no podían ser resueltas por métodos analíticos originaron la investigación de métodos de aproximación numéricas, ya por **1900** se habían ideado métodos de integración numérica bastantes efectivos aunque su aplicación estaba restringida por la necesidad de ejecutar sus cálculos a mano o con equipos de computo bastante primitivos.

Durante los años **1950-1960** se desarrollan computadoras cada vez mas poderosas, para investigar la eficacia por medio de métodos numéricos.

Durante los últimos quince o veinte años se han unificado tendencias en especial las **gráficas por computadoras**, han dado un nuevo impulso al estudio de sistemas no lineales de ecuaciones diferenciales. Se han descubierto fenómenos inesperados a los que se le ha dado el nombre de **Atractores Extraños, Caos y Fractales**, los cuales se están estudiando intensamente y están dando orígenes a nuevas concepciones en diversas aplicaciones.

El rápido desarrollo de las funciones spline se debe principalmente a su gran utilidad en aplicaciones. Las clases de funciones spline poseen muchas propiedades estructurales agradables, así como excelentes poderes de aproximación. Splines tienen muchas aplicaciones en la solución numérica de una variedad de problemas en matemáticas aplicadas y la ingeniería.

Se acepta comúnmente que se hizo la primera referencia matemática para splines en el año 1946, en un interesante artículo de Schoenberg, que es probablemente el primer lugar que la palabra "spline" se utiliza en relación con suave, aproximación polinómica a trozos. Sin embargo, las ideas tienen sus raíces en la aeronave y las industrias de construcción naval. Splines son tipos de curvas, desarrolladas originalmente por la construcción de buques en los días previos a modelos informáticos. Los arquitectos navales necesitaban una forma de llamar la una curva suave a través de un conjunto de puntos. La solución fue colocar pesas de metal (llamado nudos) en los puntos de control, y doblar un metal delgado o viga de madera (llamado una spline) a través de los pesos. La física de la spline de flexión significaba que la influencia de cada peso fue mayor en el punto de contacto y disminuyó sin problemas aún más a lo largo de la spline. Para obtener más control sobre una determinada región de la spline, la dibujante simplemente añade más peso. Había una necesidad de forma matemática para describir la forma de la curva. A través de la llegada de las computadoras, las spline han adquirido más importancia.

Primero fueron utilizados como un sustituto de los polinomios en la interpolación y luego como herramienta para la construcción de formas suaves y flexibles en gráficos por ordenador.

En matemáticas, en el campo de las ecuaciones diferenciales, un problema de valor de frontera o contorno se lo denomina al conjunto de una ecuación diferencial y a las condiciones de frontera o contorno. Una solución de un problema de condiciones de frontera es una solución de una ecuación diferencial que también satisface condiciones de frontera. Un problema de condiciones de frontera aparece en muchos aspectos de la física, como en las

ecuaciones diferenciales que explican ciertos problemas físicos. Problemas que involucran la ecuación de onda son comúnmente problemas de condiciones de frontera. Muchas clases de problemas de valores de frontera importantes son los problemas de Sturm-Liouville. El análisis de estos problemas involucran funciones propias y operadores diferenciales. Muchos de los primeros problemas de valor de frontera han sido estudiados mediante los problemas de Dirichlet, o buscando una función armónica (solución de una ecuación de Laplace) cuya solución está dada por el principio de Dirichlet.

3.2. Justificación.

En el estudio de las EDO, se determinaron diferentes métodos de resolución de ciertas ecuaciones diferenciales como: EDO con coeficientes constantes, de forma analítica. Ahora analicemos las siguientes ecuaciones.

1. Ecuación de deflexión de una biga:

$$\frac{d^2w(x)}{dx^2} = \frac{S}{EI}w(x) + \frac{q}{2EI}x(x-l),$$

donde los coeficientes $\frac{S}{EI} > 0$ y $\frac{q}{2EI} > 0$ no son constantes.

2. Respuesta de Membrana de una capa Esférica.

$$y'' + \left(\frac{t^2}{32y^2} - \frac{\lambda^2}{8} \right) = 0, \quad 0 < t < 1, \text{ y } \lambda > 0.$$

Esta ecuación la modela la respuesta de membrana de una capa esférica.

3. Conductores Eléctricos Sólidos.

$$y'' = \lambda \exp(\mu y),$$

esta aplicación involucra la difusión de calor generada por la temperatura positiva dependiente de fuentes.

Al igual que estas EDO existe una gran variedad de ecuaciones las cuales no pueden ser resueltas con los métodos analíticos conocidos, ante ello surge la necesidad de estudiar métodos numéricos de solución de EDO, específicamente a las ecuaciones de problemas de valor de frontera.

La importancia de estudiar problemas de valor de frontera para ecuaciones diferenciales ordinarias se debe a que modelan muchos problemas de la realidad, los cuales son de gran interés en física, química, biología e ingeniería. Por esa razón necesitamos saber la solución de la EDO en el intervalo en estudio, por lo que es necesario hacer la implementación de una función que aproxime la solución de dicha ecuación diferencial, la cual la haremos mediante la función b-splines, dicha aproximación sera continua y diferenciable en dicho intervalo. Entonces nos interesa que dichos métodos que resuelven problemas de valor de frontera representen confiablemente los datos.

El estudio de los PVF para EDO, es una base para el estudio posterior de las **Soluciones Numéricas de las Ecuaciones Diferenciales Parciales**.

4. Planteamiento del problema.

Las ecuaciones diferenciales cuentan con un número infinito de soluciones. Con el objetivo de obtener una solución única debemos imponer condiciones adecuadas en la frontera de Ω denotada por $\partial\Omega$ y, para ecuaciones que dependen del tiempo, condiciones iniciales adecuadas en el tiempo $t = 0$.

En el caso unidimensional, una posibilidad de la solución es prescribir el valor de u en $x = a$ y $x = b$, obteniendo

$$\begin{aligned} -u''(x) &= f(x) & x &\in (a, b), \\ u(a) &= \alpha, & u(b) &= \beta, \end{aligned} \tag{1}$$

donde α y β son dos números reales dados. Este es un problema de valor de frontera de Dirichlet.

Realizando doble integración se ve fácilmente que si $f \in C^0([a, b])$, la solución u existe y es única; por otra parte f pertenece a $C^2([a, b])$.

Hay PVF que no es posible resolverlos de forma analítica, si no que se necesitan de métodos numéricos para dar una solución aproximada, entonces tenemos la necesidad de buscar métodos adecuados que nos den una solución aproximada tan cercana a la solución real. Nuestro objetivo es resolver PVF tales como:

- Sea

$$y'' = -\frac{2}{x}y' + \frac{2}{x^2}y + \frac{\sin(\ln(x))}{x^2}, \quad 1 \leq x \leq 2, \quad y(1) = 1, \quad y(2) = 2$$

un problema de valor de frontera, y la solución exacta es

$$y = c_1x + \frac{c_2}{x^2} - \frac{3}{10} \sin(\ln(x)) - \frac{1}{10} \cos(\ln(x)),$$

donde

$$c_2 = \frac{1}{10}[8 - 12 \sin(\ln(2)) - 4 \cos(\ln(2))] \approx -0.03920701320$$

y

$$c_1 = \frac{11}{10} - c_2 \approx 1.1392070132$$

- Considere el problema con valor de frontera

$$y'' = \frac{1}{8}(32 + 2x^3 - yy'), \quad 1 \leq x \leq 3, \quad y(1) = 17, \quad y(3) = \frac{43}{3},$$

que tiene la solución exacta

$$y(x) = x^2 + \frac{16}{x}.$$

- Considere el problema con valor en la frontera

$$-y'' + \pi^2 y = 2\pi^2 \sin(\pi x), \quad 0 \leq x \leq 1, \quad y(0) = y(1) = 0.$$

Cada uno de estos problemas no se pueden resolver de forma analítica, para ello estudiaremos métodos numéricos que ayuden y que nos den una aproximación con un error tan pequeño con respecto a la solución real, para poder identificar que método numérico nos ayuda con eficiencia a la resolución de nuestros problemas planteados, aremos una comparación entre los métodos numéricos para ver cual método nos da una mejor solución aproximada de la solución real.

5. Metodología.

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

1. Tipo de investigación.

Este proyecto de investigación es de carácter bibliográfico-descriptivo.

1.1. 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.

1.2. Descriptivo:

Ya que se pretende estudiar a detalle las demostraciones de los diversos teoremas.

2. Forma de Trabajo

Se tendrán reuniones periódicas con el asesor y del trabajo para tratar los diferentes aspectos de la investigación como estudiar y discutir la teoría, y tratar los diferentes aspectos del trabajo escrito.

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.

CAPITULO UNO.

PRELIMINARES.

6. PRELIMINARES.

En este capítulo se expondrán los requerimientos matemáticos necesarios para abordar adecuadamente el estudio de los métodos numéricos usados para aproximar la solución de problema de valor de frontera de segundo orden. Al mismo tiempo, se introducen el método de Newton Raphson, tanto para una ecuación como para un sistema de ecuaciones no lineales, y el método de Runge-Kutta para resolver Problemas de Valor Inicial de Ecuaciones Diferenciales Ordinarias, estudiados en un curso básico de Análisis Numérico, necesarios al momento de definir los métodos usados en el presente trabajo.

Es importante resaltar que los resultados que se desarrollan en el presente capítulo no serán tratados de forma detallada por completo, pues el objetivo es solamente dejar asentado un material de repaso, cuyo conocimiento es importante para entender el tema de interés.

6.1. Cálculo.

Los conceptos del límite y continuidad de una función son fundamentales en el estudio del cálculo diferencial.

Definición 1 : Una función f definida en un conjunto X de números reales tiene un **límite** L en x_0 denotado por

$$\lim_{x \rightarrow x_0} f(x) = L,$$

si dado cualquier número real $\epsilon > 0$, existe un número real $\delta > 0$ tal que $|f(x) - L| < \epsilon$, siempre que $x \in X$, y $0 < |x - x_0| < \delta$.

Definición 2 : Sea $\{x_n\}_{n=1}^{\infty}$ una sucesión de números reales o complejos. La sucesión $\{x_n\}_{n=1}^{\infty}$ tiene límite x (converge a x), si para cualquier $\epsilon > 0$, existe un entero positivo $N(\epsilon)$ tal que $|x_n - x| < \epsilon$, siempre que $n > N(\epsilon)$.

Notación $\lim_{n \rightarrow \infty} x_n = x$, o $x_n \rightarrow x$ cuando $n \rightarrow \infty$, significa que la sucesión $\{x_n\}_{n=1}^{\infty}$ converge a x .

Definición 3 : Sea f una función definida en un conjunto X de números reales y $x_0 \in X$. Entonces f es continua en x_0 si

$$\lim_{x \rightarrow x_0} f(x) = f(x_0).$$

La función f es continua en X si es continua en cada número en X .

Definición 4 : Sea f una función definida en un intervalo abierto (a, b) que contiene a x_0 , la función es diferenciable en x_0 si

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0},$$

existe. El número $f'(x_0)$ es la derivada de f en x_0 . Una función que tiene derivada en cada número de un conjunto X es diferenciable en X .

Teorema 1 : Si la función f es diferenciable en x_0 , entonces f es continua en x_0 .

El conjunto de todas las funciones que tiene n derivadas continuas en X se denota $C^n(X)$, y el conjunto de funciones que tiene derivadas de todos los ordenes en X se denota $C^\infty(X)$.

Teorema 2 : Teorema del valor medio. Si $f \in C^1[a, b]$ y f es diferenciable en (a, b) , entonces existe un número $c \in (a, b)$ tal que

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Teorema 3 : (Teorema del valor intermedio.) Si $f \in C^1[a, b]$ y k es cualquier número entre $f(a)$ y $f(b)$, entonces existe un número c en (a, b) tal que $f(c) = k$.

Teorema 4 : Teorema de punto fijo.

1. Si $g \in C^1[a, b]$ y $g(x) \in [a, b] \forall x \in [a, b]$, entonces g tiene un punto fijo en $[a, b]$,
2. y si además $g'(x)$ existe en (a, b) y existe una constante positiva $k < 1$ con

$$|g'(x)| \leq k, \forall x \in (a, b),$$

entonces el punto fijo en $[a, b]$ es único.

Prueba 1.

Si $f(a) = a$, ó $f(b) = b$, la prueba acaba. Supongamos que $f(a) > a$, y que $f(b) < b$, consideremos $f(x) = g(x) - x$, entonces tenemos que

$$f(a) = g(a) - a > 0$$

y

$$f(b) = g(b) - b < 0,$$

de aquí tenemos por teorema (3) que existe $p \in [a, b]$ tal que $f(p) = 0$ entonces $g(p) = p$, de aquí g tiene al menos un punto fijo.

Prueba 2.

Por 1) tenemos que g tiene al menos un punto fijo.

Supongamos que $x_1, x_2 \in [a, b]$ son puntos fijos de g , donde $x_1 \neq x_2$ por teorema (2) tenemos que existe un $c \in [a, b]$ tal que $|g'(c)| = \frac{|g(x_2) - g(x_1)|}{|x_2 - x_1|} = \frac{|x_2 - x_1|}{|x_2 - x_1|} = 1$ y esto es una contradicción ya que por hipótesis

$$|g'(x)| \leq k < 1,$$

por lo tanto g tiene un único punto fijo.

□

6.2. Orden de aproximación $\mathcal{O}(h^n)$.

Definición 5 : Se dice que $f(h)$ es de orden $g(h)$ cuando $h \rightarrow 0$ y $h \neq 0$, se denotará por $f(h) = \mathcal{O}(g(h))$, si existen números reales $M > 0$ y $k > 0$, tal que

$$|f(h)| \leq M|g(h)|.$$

Definición 6 : Sean p y f funciones, se dice que $p(h)$ aproxima a $f(h)$ con un orden de aproximación $\mathcal{O}(h^n)$, lo que se denota por $f(h) = p(h) + \mathcal{O}(h^n)$, si existe un número real $M > 0$ y un número natural n tal que,

$$\frac{|f(h) - p(h)|}{|h^n|} \leq M,$$

para h suficientemente pequeño.

Al considerar el caso en que $p(x)$ es la n -ésima aproximación por polinomios de Taylor a $f(x)$ alrededor de x_0 , es decir:

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1},$$

para algún ξ entre x y x_0 . Cuando $x - x_0 \rightarrow 0$, por la definición (5) se tiene que:

$$\mathcal{O}((x - x_0)^{n+1}) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1},$$

así $f(x) = p(x) + \mathcal{O}((x - x_0)^{n+1})$ es decir, $p(x)$ se aproxima a $f(x)$ con un orden de aproximación $\mathcal{O}((x - x_0)^{n+1})$. Luego, el Teorema de Taylor se puede enunciar de la siguiente forma:

Teorema 5 : (Teorema de Taylor.) Si $f \in C^{(n)}[a, b]$ y $f^{(n+1)}$ existe en $[a, b]$, para cada $x \in [a, b]$, existe un número $\xi(x)$ entre x_0 y x tal que:

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + \mathcal{O}(h^{(n+1)}),$$

donde $\mathcal{O}(h^n) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!} (x - x_0)^{n+1}$.

Teorema 6 : Teorema de Taylor con Resto Integral.

Sea \mathbf{I} un intervalo abierto que contenga al intervalo cerrado de extremos x_0 , y x . Consideremos una función f de clase $C^{(n+1)}(\mathbf{I})$, entonces

$$\int_{x_0}^x f'(t) dt = f(x) - f(x_0),$$

es decir,

$$f(x) = f(x_0) + \int_{x_0}^x f'(t) dt.$$

6.3. Teoría de Álgebra Lineal.

Definición 7 : Una matriz A de orden $n \times n$ es invertible, si existe una matriz A^{-1} de orden $n \times n$, tal que $AA^{-1} = A^{-1}A = I$, donde I es la matriz identidad de orden $n \times n$. La matriz A^{-1} denota la matriz inversa de A . Una matriz que no tiene inversa recibe el nombre de no invertible o singular.

Definición 8 : (Matriz tridiagonal.) Una matriz A de orden $n \times n$, se le llama **tridiagonal** si las componentes de la matriz $a_{i,j} = 0$, siempre que $|i - j| > 1$. Así

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & 0 & \dots & \dots & 0 & 0 \\ a_{2,1} & a_{2,2} & a_{2,2} & \dots & \dots & 0 & 0 \\ 0 & a_{3,2} & a_{3,3} & \dots & \dots & 0 & 0 \\ 0 & 0 & a_{4,3} & \dots & \dots & 0 & 0 \\ \vdots & \ddots & \ddots & \dots & \dots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \dots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \dots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & 0 & \dots & \dots & a_{n,n-1} & a_{n,n} \end{bmatrix}$$

es una matriz tridiagonal.

Ahora para que este sistema de ecuaciones tenga solución única estamos interesados en saber cuando esta matriz es invertible para que tenga solución única. Dichas condiciones nos los da el siguiente teorema.

Teorema 7 : Sea A una **matriz tridiagonal**, donde las componentes de A son $a_{i,j}$, para cada $i = 1, 2, 3, \dots, n$ y $j = 1, 2, 3, \dots, n$. Supongamos que para $i = 2, 3, \dots, n - 1$ se cumple $a_{i,i-1}a_{i,i+1} \neq 0$. Si $|a_{1,1}| > |a_{1,2}|$, $|a_{i,i}| \geq |a_{i,i-1}| + |a_{i,i+1}|$, para $i = 2, 3, \dots, n - 1$ y $|a_{n,n}| > |a_{n,n-1}|$, entonces A es invertible.

6.4. Teoría de Ecuaciones Diferenciales.

Estamos interesados en problemas en los que se busca una solución $y(x)$ de una ED de modo que $y(x)$ satisfaga condiciones adicionales prescritas, es decir, condiciones impuestas en la solución $y(x)$ desconocida o sus derivadas.

En el marco de las EDO estamos interesados en los siguientes conceptos:

Definición 9 : Problema de Valor Inicial. Lo denotaremos por (PVI) de n -ésimo orden es resolver

$$y^{(n)} = f(x, y, y', y'', \dots, y^{(n-1)}),$$

sujeto a:

$$y(x_0) = y_0, y'(x_0) = y_1, \dots, y^{(n-1)}(x_0) = y_{n-1}$$

Otro problema consiste en resolver una ED de orden dos o mayor en la cual la variable dependiente y sus derivadas se especifican en diferentes puntos.

Definición 10 : *Un Problema de Valor de Frontera.* Lo denotaremos (PVF), consiste en resolver

$$y'' = f(x, y, y')$$

sujeito a:

$$y(a) = y_0, y'(b) = y_1.$$

Los valores prescritos $y(a) = y_0, y'(b) = y_1$, se llaman condición de frontera.

En el siguiente teorema se establecen las condiciones generales que garantizan la existencia y la unicidad de la solución a dicho problema de segundo orden.

Pero antes necesitamos introducir lo siguiente.

Definición 11 : *Se dice que la función $f(t, y_1, y_2, \dots, y_m)$ definida en el conjunto*

$$D = \{(t, u_1, u_2, \dots, u_m) | a \leq t \leq b, -\infty < u_i < \infty, \text{ para cada } i = 1, 2, \dots, m\},$$

satisface una **condición de Lipschitz** sobre D en las variables u_1, u_2, \dots, u_m si existe una constante $L > 0$ con la siguiente propiedad:

$$|f(t, u_1, u_2, \dots, u_m) - f(t, z_1, z_2, \dots, z_m)| \leq L \sum_{j=1}^m |u_j - z_j|. \quad (2)$$

$$\forall (t, u_1, u_2, \dots, u_m), (t, z_1, z_2, \dots, z_m) \in D.$$

De aquí estamos en condiciones de enunciar el teorema:

Teorema 8 : *Supongamos que*

$$D = \{(t, u_1, u_2, \dots, u_m) | a \leq t \leq b, -\infty < u_i < \infty, \text{ para cada } i = 1, 2, \dots, m\},$$

y que $f_i(t, u_1, u_2, \dots, u_m)$, para $i = 1, 2, \dots, m$ es continua en D y cumple la condición de Lipschitz. El sistema de ecuaciones diferenciales de primer orden, sujeto a sus condiciones iniciales, tiene una solución única, $(u_1(t), u_2(t), \dots, u_m(t))$, para $a \leq t \leq b$.

Consideremos la siguiente ecuación:

$$y'' + p(x)y' + q(x)y = 0. \quad (3)$$

Teorema 9 : *Existen dos soluciones $y_1(x)$ e $y_2(x)$ de la ecuación (3) que son linealmente independiente en \mathcal{I} , es decir, no existe una constante c tal que $y_1(x) = cy_2(x)$ para todo $x \in \mathcal{I}$, donde $p, y q$ son continuas.*

Prueba: Supongamos que $y_1(x)$ e $y_2(x)$ son dos soluciones de la ecuación (3), teniendo así que:

$$\begin{aligned} y_1'' + py_1' + qy_1 &= 0 & y_2'' + py_2' + qy_2 &= 0 \\ y_1(0) = 1, y_1'(0) &= 0 & y_2(0) = 0, y_2'(0) &= 1. \end{aligned}$$

Ambas soluciones $y_1(x)$ e $y_2(x)$ son las soluciones únicas que satisfacen las condiciones iniciales indicadas.

Hay que demostrar que $y_1(x)$ e $y_2(x)$ son linealmente independientes.

Asumamos que $y_1(x)$ e $y_2(x)$ son linealmente dependientes, sin pérdida de generalidad supongamos que $y_1(x) = cy_2(x)$, derivando tenemos que $y_1'(x) = cy_2'(x)$, aplicando las condiciones iniciales $y_1'(0) = 0 = cy_2'(0) = c$ entonces $c = 0$, además tenemos que $y_1(0) = 1 = cy_2(0)$, donde tendríamos que $1 = c \times 0$, lo que es una contradicción. Por lo tanto $y_1(x)$ e $y_2(x)$ son linealmente independientes.

□

El wronskiano de dos funciones $y_1(x)$ e $y_2(x)$ cuando trabajamos en un intervalo \mathcal{I} se define como:

$$W(y_1, y_2) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = y_1 y_2' - y_2 y_1'.$$

Teorema 10 : Si $y_1(x)$, $y_2(x)$ son soluciones de la ecuación (3) en un intervalo abierto \mathcal{I} linealmente independientes entonces $W(y_1, y_2) \neq 0$ en cada punto $t \in \mathcal{I}$.

Teorema 11 : $y_1(x)$, $y_2(x)$ son soluciones de la ecuación (3) y c_1, c_2 constantes arbitrarias, entonces $c_1 y_1(x) + c_2 y_2(x)$ es también una solución de (3).

Prueba: Sabemos que

$$y_1'' + p(x)y_1' + q(x)y_1 = 0, \tag{4}$$

$$y_2'' + p(x)y_2' + q(x)y_2 = 0, \tag{5}$$

multiplicando por c_1 la ecuación (4) y por c_2 la ecuación (5) y luego sumando ambas ecuaciones tenemos:

$$\begin{aligned} c_1 y_1'' + p(x)c_1 y_1' + q(x)c_1 y_1 &= 0, \\ c_2 y_2'' + p(x)c_2 y_2' + q(x)c_2 y_2 &= 0 \\ \hline (c_1 y_1'' + c_2 y_2'') + p(x)(c_1 y_1' + c_2 y_2') + q(x)(c_1 y_1 + c_2 y_2) &= 0. \end{aligned}$$

De este resultado concluimos que $c_1y_1(x) + c_2y_2(x)$ es solución de la ecuación (3).

□

6.5. Teoría de Análisis Funcional.

6.5.1. Espacios Normados.

Definición 12 : Una norma de la matriz $\|\cdot\| : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ es una función real no negativo que satisface las siguientes propiedades:

- $\|A\| > 0$ donde $A \neq 0$ y $\|A\| = 0$ si y solo si $A = 0$;
- $\|\alpha A\| = |\alpha|\|A\|$ para todo $\alpha \in \mathbb{R}$;
- $\|A + B\| \leq \|A\| + \|B\|$;
- $\|AB\| \leq \|A\|\|B\|$.

Definición 13 : Supongamos que V es un espacio lineal sobre el campo \mathbb{R} de los números reales. La función no negativa $\|\cdot\|$ de valor real se llama una norma en el espacio si satisface las siguientes propiedades:

- $\|v\| > 0$ donde $v \neq 0$ y $\|v\| = 0$ si y solo si $v = 0$ en V .
- $\|\alpha v\| = |\alpha|\|v\|$ para todo $\alpha \in \mathbb{R}$ y v en V .
- $\|u + v\| \leq \|u\| + \|v\|$ para todo u y v en V . (Desigualdad triangular.)

Un espacio lineal V , dotado con una norma, se llama un espacio lineal normado.

Definición 14 : El espacio lineal $L_2[a, b]$ es la colección de todas las funciones medibles $f(x)$ definida en $[a, b]$ tal que

$$\int_a^b [f(x)]^2 dx < \infty,$$

donde la integración es la integración de Lebesgue.

Definición 15 : Sea $f(x) \in C[a, b]$ (el espacio de todas las funciones continuas en $[a, b]$) o $f(x) \in L_2[a, b]$. La función real $\|\cdot\|_2$ definida por $\|f\|_2 = \left(\int_a^b [f(x)]^2 dx \right)^{\frac{1}{2}}$ es llamada L_2 la norma de f .

En este documento se utiliza la siguiente aproximación para norma de error L_2

$$L_2 = \left(\int_a^b [u - U]^2 dx \right)^{\frac{1}{2}} \approx \sqrt{h \sum_{i=1}^N (u_i - U_i)^2},$$

donde u y U son soluciones exactas y aproximadas respectivamente y h es el tamaño de paso.

Definición 16 : Una norma vectorial $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$ es una función no negativa satisfaciendo las siguientes propiedades:

- $\|\mathbf{v}\| > 0$ cuando $\mathbf{v} \neq 0$ y $\|\mathbf{v}\| = 0$ si y solo si $\mathbf{v} = 0$ en \mathbb{R}^n .
- $\|\alpha \mathbf{v}\| = |\alpha| \|\mathbf{v}\| \quad \forall \alpha \in \mathbb{R}$.
- $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\| \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.

Las siguientes normas vectoriales son de uso común en álgebra lineal numérica: la norma-1 $\|\cdot\|_1$. La norma-2 (norma euclidiana) $\|\cdot\|_2$, norma infinita $\|\cdot\|_\infty$. Si $\mathbf{v} = (v_1, v_2, \dots, v_n)^T \in \mathbb{R}^n$, estas normas se definen como siguen.

Definición 17 : La norma-1 $\|\cdot\|_1$, del vector \mathbf{v} es definida como

$$\|\mathbf{v}\|_1 = \sum_{i=1}^n |\mathbf{v}_i|.$$

Definición 18 : La norma-2 (norma euclidiana) $\|\cdot\|_2$, del vector \mathbf{v} esta definida

$$\|\mathbf{v}\|_2 = (\mathbf{v}^T \mathbf{v})^{\frac{1}{2}} = \left(\sum_{i=1}^n |\mathbf{v}_i|^2 \right)^{\frac{1}{2}}.$$

Definición 19 : La norma infinita $\|\cdot\|_\infty$, del vector \mathbf{v} esta definida por

$$\|\mathbf{v}\|_\infty = \max_{i=1}^n |\mathbf{v}_i|.$$

Definición 20 : Una matriz A es convergente a cero si la secuencia de matrices A_1, A^2, A^3, \dots converge a la matriz nula.

Teorema 12 : Para una matriz A , la condición necesaria y suficiente para decir que la sucesión de matrices I, A, A^2, A^3, \dots converge a la matriz cero, es que la magnitud de los eigenvalores de A deben ser menor que la unidad.

Teorema 13 : Una condición necesaria y suficiente para decir que la sucesión de matrices $I + A + A^2 + A^3 + \dots$ converge a $(I - A)^{-1}$ es que $\lim_{r \rightarrow \infty} A^r = 0$.

La prueba de este teorema se hará más adelante.

Lema 1 : Para cualquier norma matricial con $\|I\| = 1$, entonces se cumple la siguiente desigualdad

$$\|(I + A)^{-1}\| \leq (1 - \|A\|)^{-1}.$$

Demostración:

Por hipótesis sabemos que $\|I\| = 1$, entonces

$$\| \|I\| - \|A\| \| \leq \|(I + A)\|$$

$$\|(I + A)\|^{-1} \leq (1 - \|A\|)^{-1}.$$

□

Lema 2 : Una condición suficiente para la convergencia de la serie $I + A + A^2 + A^3 + \dots$ es que $\|A\| < 1$ para cualquier norma de una matriz A .

Lema 3 : Sea A alguna matriz; entonces si $\|A\| < 1$ para alguna norma, entonces la siguiente desigualdad se cumple

$$\|(I - A)^{-1} - (I + A + A^2 + A^3 + \dots + A^r)\| \leq \frac{\|A\|^{r+1}}{(1 - \|A\|)}.$$

Demostración:

Es fácil ver que del teorema (13) obtenemos

$$\begin{aligned} (I - A)^{-1} - (I + A + A^2 + A^3 + \dots + A^r) &= A^{r+1} + A^{r+2} + \dots \\ \|(I - A)^{-1} - (I + A + A^2 + A^3 + \dots + A^r)\| &\leq \|A^{r+1}\| + \|A^{r+2}\| + \dots \\ &\leq \|A\|^{r+1} + \|A\|^{r+2} + \dots \\ &= \|A\|^{r+1}(1 + \|A\| + \|A\|^2 + \dots) \\ &= \frac{\|A\|^{r+1}}{(1 - \|A\|)} \quad \text{si } \|A\| < 1. \end{aligned}$$

□

Lema 4 : Cada norma de la matriz $\|A\|$, es una función continua de n^2 elementos a_i , de A .

Teorema 14 : Para cada par de normas de una matriz, es decir $\|A\|$ y $\|A\|'$, existe constantes positivas m y M tal que para toda matriz A de orden n ,

$$m\|A\|' \leq \|A\| \leq M\|A\|'.$$

Corolario 1 Para cualquier matriz cuadrada A

$$\rho(A) = \inf_{\{N(\cdot)\}} \max_{N(x)=1} N(Ax)$$

donde el \inf se toma sobre todas las normas de vectores, $N(\cdot)$; o equivalentemente

$$\rho(A) = \inf_{\{\|\cdot\|\}} \|A\|,$$

donde el \inf se toma sobre todas las normas naturales.

Teorema 15 : Para cada matriz A de orden n y $\epsilon > 0$ arbitrario, una norma natural, $\|\cdot\|$ cumple que

$$\rho(A) \leq \|A\| \leq \rho(A) + \epsilon. \text{ Donde } \rho(A) \text{ denota el radio espectral de } A$$

Los resultados anteriores se pueden resumir en el siguiente teorema.

Teorema 16 : Las siguientes afirmaciones son equivalentes

1. A es matriz convergente.
2. $\lim_{r \rightarrow \infty} \|A^r\| = 0$ para alguna norma $\|\cdot\|$.
3. $\rho(A) < 1$ cuando $\rho(A)$ es el radio espectral de A .

Demostración:

Primero demostraremos que (1) y (2) son equivalentes.

Puesto que $\|\cdot\|$ es continua, por el lema (4), y $\|0\| = 0$, entonces (1) implica (2).

Pero si (2) se cumple por alguna norma, entonces el teorema (14) implica que existe un M tal que

$$\|A^m\|_\infty \leq M\|A^m\| \rightarrow 0,$$

por lo tanto se cumple (1).

Veamos que (2) y (3) son equivalentes.

Por el teorema (14), sin perdida de generalidad en si asumimos que es la norma natural.

Pero entonces por el lema (4) y el hecho de que

$$\lambda(A^m) = \lambda^m(A), \text{ donde } \lambda \text{ es eigenvalor}$$

nosotros tenemos

$$\|A^m\| \geq \rho(A^m) = \rho^m(A),$$

por lo tanto (2) implica (3).

Demostraremos que (3) y (2) son equivalentes.

Por otro lado, por el teorema (15) podemos encontrar un $\epsilon > 0$ y una norma natural, es decir $N(\cdot)$, tal que

$$N(A) \leq \rho(A) + \epsilon \equiv \theta < 1,$$

entonces

$$N(A^m) \leq [N(A)]^m \leq \theta^m,$$

de modo que $\lim_{m \rightarrow \infty} N(A^m) = 0$. Por tanto (3) implica (2).

Teorema 17 :

1. *La serie geométrica $I + A + A^2 + A^3 + \dots$ converge si y solo si A converge.*
2. *Si A es convergente, entonces $I - A$ es no singular y $(I - A)^{-1} = I + A + A^2 + \dots$*

Demostración:

Una condición necesaria para la serie en la parte (1) para que converja es que $\lim_{m \rightarrow \infty} A^m = 0$, es decir que A converge.

La condición suficiente seguirá de la parte 2.

Sea A convergente, de donde por el teorema 16 sabemos que $\rho(A) < 1$.

Ya que los eigenvalores de $I - A$ son $1 - \lambda(A)$ se deduce que el $\det(I - A) \neq 0$ y por lo tanto esta matriz es no singular. Consideremos la siguiente identidad

$$(I - A)(I + A + A^2 + A^3 + \dots) = I - A^{r+1},$$

que es válida para todos los enteros r . Entonces

$$\begin{aligned} I + A + A^2 + A^3 + \dots + A^r &= (I - A)^{-1}(I - A^{r+1}) \\ &= (I - A)^{-1} - (I - A)^{-1}A^{r+1} \\ &= (I - A)^{-1}, \end{aligned}$$

como $r \rightarrow \infty$, $A^{r+1} \rightarrow 0$.

Por lo tanto la serie $I + A + A^2 + A^3 + \dots + A^r$ converge a $(I - A)^{-1}$ si $A^{r+1} \rightarrow 0$ cuando $r \rightarrow \infty$.

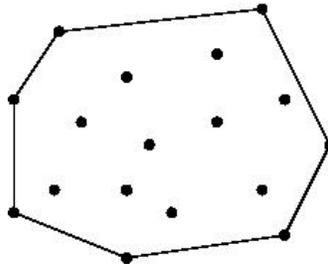
6.5.2. Conjuntos Convexos.

Definición 21 : Un conjunto A se llama convexo si juntamente con cada par de puntos $x \in A$, $y \in A$, el conjunto contiene a todo el segmento que une x con y .

Ejemplo 1 : Cada bola $B(a, r)$ es un conjunto convexo, pues si $z = (1 - \lambda)x + \lambda y$, donde $x, y \in B(a, r)$, $0 \leq \lambda \leq 1$ entonces $|z - a| = |(1 - \lambda)(x - a) + \lambda(y - a)| \leq |(1 - \lambda)(x - a)| + |\lambda(y - a)| = (1 - \lambda)|x - a| + \lambda|y - a| < (1 - \lambda)r + \lambda r = 1$. Por lo tanto $B(a, r)$ es un conjunto convexo.

Definición 22 : La intersección de todos los conjuntos convexos que contienen al conjunto A se llama **capsula convexa** de A .

Ejemplo 2 En \mathbb{R}^3 tenemos el siguiente conjunto de puntos, el rombo cerrado por los segmentos de rectas forman la **capsula convexa**.



Ejemplo 3 : Si $F = (a_0, a_1, \dots, a_n)$ es un conjunto de $n + 1$ puntos de \mathbb{R}^n tales que los vectores $a_k - a_0$ para $(k = 1, 2, 3, \dots, n)$, son linealmente independientes, la capsula convexa de F se le llama **simple** con vértices a_k , $(k = 1, 2, 3, \dots, n)$. Este conjunto consta de todos los puntos x de la forma $x = t_0 a_0 + t_1 a_1 + \dots + t_n a_n$, donde cada t_k es no negativo y

$$\sum_{k=0}^n t_k = 1.$$

6.6. Método de Newton Raphson y Runge Kutta.

En esta sección introduciremos dos métodos necesarios para el estudio de los métodos que se emplearan para aproximar la solución de un PVF de segundo orden. Dichos métodos son el **Método de Newton**, tanto para aproximar una ecuación o un sistema de ecuaciones no lineales, y el **método de Runge Kutta**, para resolver PVI de ecuaciones diferenciales ordinarias.

6.6.1. Método de Newton.

El método de **Newton** (o método de **Newton-Raphson**), es una de las técnicas numéricas para resolver un problema de búsqueda de raíces $f(x) = 0$, mas poderosas y conocidas.

El estudio de este método lo haremos por los polinomios de Taylor.

Supongamos que $f \in C^2[a, b]$. Sea $\bar{x} \in [a, b]$ una aproximación de p tal que $f'(\bar{x}) \neq 0$ y $|p - \bar{x}|$ es **pequeño**. Consideremos el primer polinomio de Taylor para $f(x)$ expandido alrededor de \bar{x} .

$$f(x) = f(\bar{x}) + (x - \bar{x})f'(\bar{x}) + \frac{(x - \bar{x})^2}{2}f''(\xi(x)),$$

donde $\xi(x)$ esta entre x y \bar{x} . Dado que $f(p) = 0$ esta ecuación, con $x = p$, da

$$0 = f(\bar{x}) + (p - \bar{x})f'(\bar{x}) + \frac{(p - \bar{x})^2}{2}f''(\xi(p)).$$

El **método de Newton** se obtiene suponiendo que, como $|p - \bar{x}|$ es tan pequeño, el termino que contiene $(p - \bar{x})^2$ es mucho menor y así

$$0 \approx f(\bar{x}) + (p - \bar{x})f'(\bar{x}),$$

despejando p de esta ecuación obtenemos

$$p \approx \bar{x} - \frac{f(\bar{x})}{f'(\bar{x})}.$$

Esto nos prepara para introducir el método de Newton, el cual comienza con una aproximación p_0 y genera la sucesión $\{p_n\}_{n=0}^{\infty}$ definida por

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})}, \quad \text{para } n \geq 1. \quad (6)$$

Teorema 18 : Convergencia del método de Newton.

Sea $f \in C^2[a, b]$. Si $p \in [a, b]$ es tal que $f(p) = 0$ y $f'(p) \neq 0$, entonces existe $\delta > 0$ tal que el método de de newton genera una sucesión $\{p_n\}_{n=0}^{\infty}$, que converge a p para cualquier aproximación inicial $p_0 \in [p - \delta, p + \delta]$.

Demostración: la prueba esta basada en analizar el método de Newton como un esquema de iteración funcional $p_n = g(p_{n-1})$, para $n \geq 1$, con

$$g(x) = x - \frac{f(x)}{f'(x)}. \quad (7)$$

Sea k un numero cualquiera en $(0, 1)$. Primero debemos encontrar un intervalo $[p - \delta, p + \delta]$, que g mapea en si mismo y en que $|g'(x)| \leq k$, $\forall x \in (p - \delta, p + \delta)$.

Como $f'(p) \neq 0$, y f' es continua, existe $\delta_1 > 0$ tal que $f'(x) \neq 0, \forall x \in [p - \delta_1, p + \delta_1] \subseteq [a, b]$, por tanto, g esta definida y es continua en $[p - \delta_1, p + \delta_1]$ ya que $f(x), f'(x)$ y por tanto $\frac{f(x)}{f'(x)}$ es continua por ser un composición de funciones continuas, ademas $x - \frac{f(x)}{f'(x)}$ por ser suma de funciones continuas, También,

$$g'(x) = 1 - \frac{f'(x)f'(x) - f(x)f''(x)}{[f'(x)]^2} = \frac{f(x)f''(x)}{[f'(x)]^2},$$

$\forall x \in [p - \delta_1, p + \delta_1]$ y como $f \in C^2[a, b]$, tendremos $g \in C^1[p - \delta_1, p + \delta_1]$. substituyendo $f(p) = 0$ tenemos

$$g'(p) = \frac{f(p)f''(p)}{[f'(p)]^2} = 0.$$

Como g' es continua, por definición de continuidad tenemos $\forall \epsilon > 0$ existe $\delta > 0$ tal que $|g'(x) - g'(p)| < \epsilon$ y como $g'(p) = 0$ tenemos que $|g'(x)| < \epsilon$ en particular tenemos que se cumple para $\delta < \delta_1$ y $\epsilon > k$ tenemos que

$$|g'(x)| \leq k, \forall x \in [p - \delta, p + \delta] \subseteq [p - \delta_1, p + \delta_1].$$

Ahora falta ver que $g : [p - \delta, p + \delta] \longrightarrow [p - \delta, p + \delta]$. Consideremos $x \in [p - \delta, p + \delta]$ por teorema (2) tenemos que para algún numero ξ entre x y p se tiene

$$\frac{g(x) - g(p)}{x - p} = g'(\xi),$$

$$g(x) - g(p) = g'(\xi)(x - p)$$

$$|g(x) - g(p)| = |g'(\xi)||x - p|,$$

por ecuación (7) tenemos que $g(p) = p$, de a qui

$$|g(x) - p| = |g(x) - g(p)| = |g'(\xi)||x - p|,$$

y como encontramos que $|g'(x)| \leq k$ de a qui que

$$|g(x) - p| = |g(x) - g(p)| = |g'(\xi)||x - p| \leq k|x - p| < |x - p|.$$

Puesto que $x \in [p - \delta, p + \delta]$, tenemos que $|x - p| < \delta$, y que $|g(x) - p| < \delta$, por lo que de a qui concluimos que $g : [p - \delta, p + \delta] \longrightarrow [p - \delta, p + \delta]$. De a qui tenemos que cumple con las condiciones del teorema (4), entonces la sucesión $\{p_n\}_{n=1}^{\infty}$ definida por

$$p_n = g(p_{n-1}) = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})}, \text{ para } n \geq 1,$$

converge a p para cualquier $p_0 \in [p - \delta, p + \delta]$.

6.6.2. Método de Newton para sistemas de ecuaciones no lineales.

Para la solución de sistemas no lineales utilizamos el método de Newton. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ suave; $f(x) = (f_1(x) \cdots f_n(x))^T$, $x \in \mathbb{R}^n$; $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$; encontramos $x^* \in \mathbb{R}^n$ tal que $f(x^*) = 0$.

Sea x^k una estimación de x^* . La aproximación localmente de primer orden de Taylor: tenemos la aproximación lineal

$$f(x) \approx f(x^k) + J(x^k)(x - x^k),$$

donde $J(x^k)$ es la matriz jacobiana de $f(x)$ en $x = x^k$, en otras palabras

$$J(x^k)_{ij} = \left(\frac{\partial f_i}{\partial x_j} \right) (x^k) \quad \text{con } i, j \in \{1, 2, \dots, n\}.$$

Sea x^{k+1} tal que $f(x^k) + J(x^k)(x^{k+1} - x^k) = 0$.

$J(x^k)$ no singular entonces, despejando x^{k+1} tenemos: La recursión definida por:

$$x^{k+1} = x^k - (J(x^k))^{-1} f(x^k).$$

Lema 5 : Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $f \in C^1(\mathbb{R}^n)$ con jacobiano J localmente Lipschitz y continua en el punto x^* , quiere decir que existe una bola abierta \mathcal{N} con radio positivo con centro en x^* tal que

$$\|J(x) - J(x^*)\| \leq L_* \|x - x^*\|, \quad \forall x \in \mathcal{N}. \quad (8)$$

donde L_* es una constante positiva. Entonces para $x \in \mathcal{N}$

$$f(x^*) = f(x) + J(x)(x^* - x) + \mathbf{E}(x^* - x), \quad (9)$$

donde el error $\mathbf{E}(x^* - x)$ satisface

$$\|\mathbf{E}(x^* - x)\| \leq \frac{3}{2} L_* \|x^* - x\|^2. \quad (10)$$

Prueba.

Por Teorema: (6), la expansión de Taylor de primer orden de f en x tenemos: $f(x+h) = f(x) + \int_0^1 J(x+th) h dt$.

Sea $h = x^* - x$, sustituyendo en la expansión de Taylor, deducimos que

$$\begin{aligned} f(x^*) &= f(x) + \int_0^1 J(x+t(x^*-x))(x^*-x) dt. \\ &= f(x) + J(x)(x^*-x) + \int_0^1 [J(x+t(x^*-x)) - J(x)](x^*-x) dt. \end{aligned}$$

De a qui igualando con la ecuación (9) tenemos:

$$\mathbf{E}(x^* - x) = \int_0^1 [J(x + t(x^* - x)) - J(x)](x^* - x) dt,$$

aplicando norma tenemos:

$$\begin{aligned} \|\mathbf{E}(x^* - x)\| &\leq \int_0^1 \| [J(x + t(x^* - x)) - J(x)](x^* - x) \| dt. \\ &\leq \int_0^1 \| [J(x + t(x^* - x)) - J(x)] \| \cdot \| (x^* - x) \| dt. \end{aligned}$$

Usando desigualdad triangular, tenemos que $\forall x \in \mathcal{N}$ y $t \in [0, 1]$,

$$\begin{aligned} \| J(x + t(x^* - x)) - J(x) \| &\leq \| J(x + t(x^* - x)) - J(x^*) \| + \| J(x) - J(x^*) \| \\ &\leq L_* \| x + t(x^* - x) - x^* \| + L_* \| x - x^* \| = (2-t)L_* \| x^* - x \|, \end{aligned}$$

utilizando la hipótesis que J es localmente lipschitz y $x + t(x^* - x) \in \mathcal{N}$, $\forall t \in [0, 1]$.
Teniendo así que

$$\begin{aligned} \|\mathbf{E}(x^* - x)\| &\leq \int_0^1 (2-t)L_* \| x^* - x \| \cdot \| x^* - x \| dt = \int_0^1 (2-t)L_* \| x^* - x \|^2 dt. \\ \|\mathbf{E}(x^* - x)\| &\leq L_* \| x^* - x \|^2 \int_0^1 (2-t) dt = \frac{3}{2}L_* \| x^* - x \|^2. \end{aligned}$$

□

Teorema 19 : (Convergencia local del método de Newton): Sea $f(x^*) = 0$ con $J(x^*)$ no singular y localmente Lipschitz y continua en x^* . Entonces existe un $\mathcal{N} = \mathcal{N}(x^*, \delta)$ una vecindad de x^* tal que si $x^{k_0} \in \mathcal{N}$ para algún $k_0 \geq 0$, entonces $x^k \in \mathcal{N}$ y para todo $k \geq k_0$. Entonces también $x^k \rightarrow x^*$ cuando $k \rightarrow \infty$, y la tasa de convergencia es cuadrática.

- En condiciones del teorema (19) x^k esta suficientemente cerca de x^* .

Prueba: Elijamos \mathcal{N} tal que $\forall x \in \mathcal{N}$, $J(x)$ es no singular $\| J(x)^{-1} \| \leq \gamma$ sujeto a (8) y $\| x - x^* \| \leq \frac{1}{3\gamma L_*}$.

Asumamos que $x^k \in \mathcal{N}$, de aquí tenemos que $J(x)^{-1}$ existe y también x^{k+1} es igualmente definido. En el Lema (5) con $x = x^k$ y $f(x^*) = 0$ tenemos:

$$0 = f(x^k) + J(x^k)(x^* - x^k) + \mathbf{E}(x^* - x^k).$$

multiplicando por $J(x^k)^{-1}$ y usando $x^{k+1} = x^k - J(x^k)^{-1} f(x^k)$,

$$0 = x^* - x^{k+1} + J(x^k)^{-1} \mathbf{E}(x^* - x^k).$$

entonces

$$\|x^{k+1} - x^*\| = \|J(x^k)^{-1} \mathbf{E}(x^* - x^k)\| \leq \|J(x^k)^{-1}\| \cdot \|\mathbf{E}(x^* - x^k)\|.$$

Del lema (5) y la elección de \mathcal{N} obtenemos la tasa de convergencia cuadrática

$$\|x^{k+1} - x^*\| \leq \frac{3}{2} \gamma L_* \|x^k - x^*\|^2. \quad (11)$$

Ahora bien, de la ecuación (11) y de la definición de \mathcal{N} , deducimos ya que $\|x - x^*\| \leq \frac{1}{3\gamma L_*}$ tenemos:

$$\|x^{k+1} - x^*\| \leq \frac{1}{2} \|x^k - x^*\| < \|x^k - x^*\|,$$

dado que $x^k \in \mathcal{N}$. De este resultado y de la ecuación (11) inductivamente tenemos que toda subsecuencia iterativa de x^l definidas, anteriormente de \mathcal{N} , convergen a x^* , cuadráticamente.

□

6.6.3. Método de Runge-Kutta.

Es un método de resolución numérica de ecuaciones diferenciales ordinarias. Para introducirlo definamos el PVI:

$$y' = f(t, y), \quad y(t_0) = y_0$$

Entonces el método de **Runge-Kutta de cuarto orden** para este problema está dado por la siguiente ecuación:

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

donde:

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

$\forall n \geq 0$. Por ser de cuarto orden, se tiene que el error de paso por teorema de Taylor es de $\mathcal{O}(h^5)$, mientras que el error acumulado tiene orden $\mathcal{O}(h^4)$.

CAPITULO DOS.

EIGENVALORES CON PROBLEMAS DE VALOR EN LA FRONTERA.

7. Eigenfunciones y problemas con valores en la frontera.

7.1. Problemas de Sturm-Liouville y desarrollo en eigenfunciones.

Diversos problemas con valores en la frontera dan como resultado **por separación de variables** la misma ecuación diferencial ordinaria.

$$X''(x) + \lambda X(x) = 0 \quad (0 < x < L), \quad (12)$$

con un eigenvalor λ , pero con diferentes condiciones de frontera, tales como

$$X(0) = X(L) = 0, \quad (13)$$

o

$$X'(0) = X'(L) = 0, \quad (14)$$

o

$$X(0) = X'(L) = 0, \quad (15)$$

dependiendo de las condiciones de frontera originales.

7.1.1. Problemas de Sturm-Liouville.

Para unificar y generalizar el método de separación de variables es útil formular un tipo general del problema del eigenvalor que incluya como casos especiales cada uno de los mencionados previamente. La ecuación (12), con y en vez de X como variable dependiente, puede escribirse en la forma

$$\frac{d}{dx} \left[p(x) \frac{dy}{dx} \right] - q(x)y + \lambda r(x)y = 0, \quad (16)$$

donde $p(x) = r(x) \equiv 1$ y $q(x) \equiv 0$. De hecho, cualquier ecuación diferencial lineal de segundo orden del tipo

$$A(x)y'' + B(x)y' + C(x)y + \lambda D(x)y = 0,$$

toma la forma dada en (16) después de multiplicarse por un factor adecuado.

Definición 23 : *Problema de Sturm-Liouville.*

Un **problema de Sturm-Liouville** es un problema con valores en la frontera de la forma

$$\frac{d}{dx} \left[p(x) \frac{dy}{dx} \right] - q(x)y + \lambda r(x)y = 0 \quad (a < x < b); \quad (17)$$

$$\alpha_1 y(a) - \alpha_2 y'(a) = 0, \quad \beta_1 y(b) - \beta_2 y'(b) = 0, \quad (18)$$

donde tanto α_1 y α_2 como β_1 y β_2 son diferentes de cero. El parámetro λ en (17) es el “**eigenvalor**” cuyos posibles valores (constantes) se buscan.

Nótese que el problema de Sturm-Liouville dado en (17)-(18) siempre tiene la solución trivial $y(x) \equiv 0$. Por consiguiente se buscan los valores de λ —los eigenvalores— para los cuales este problema tiene una solución real no trivial (una eigenfunción) y cada eigenvalor cuenta con su eigenfunción asociada (o eigenfunciones). Puede advertirse que cualquier constante (diferente de cero) múltiplo de una eigenfunción será también una eigenfunción. El siguiente teorema ofrece condiciones suficientes bajo las cuales el problema dado en (17)-(18) presenta una sucesión infinita $\{\lambda_n\}_1^\infty$ de eigenvalores no negativos, donde cada eigenvalor λ_n tiene (por la diferencia de un múltiplo constante) exactamente una eigenfunción asociada $y_n(x)$.

Teorema 20 : *Eigenvalores de Sturm-Liouville.*

Asumamos que las funciones $p(x)$, $p'(x)$, $q(x)$ y $r(x)$ de la ecuación (17) son continuas en el intervalo $[a, b]$ y que tanto $p(x) > 0$ como $r(x) > 0$ en cada punto de $[a, b]$. De este modo los eigenvalores del problema de Sturm-Liouville dado en (17)-(18) constituyen una sucesión creciente

$$\lambda_1 < \lambda_2 < \lambda_3 < \cdots < \lambda_{n-1} < \lambda_n < \cdots, \quad (19)$$

de números reales con

$$\lim_{n \rightarrow \infty} \lambda_n = +\infty. \quad (20)$$

Salvo por un factor constante, solo una eigenfunción $y_n(x)$ se asocia con cada eigenvalor λ_n . Además, si $q(x) \geq 0$ en $[a, b]$ y los coeficientes α_1 y α_2 , β_1 y β_2 en (18) son todos no negativos, entonces los eigenvalores son todos no negativos.

Observación 1 : *Es importante poner atención a los signos de (17) y (18) cuando se verifica la hipótesis del teorema (20). Algunas veces el problema de Sturm-Liouville dado en (17)-(18) se llama **regular** si satisface las hipótesis del teorema (20); en caso contrario, es **singular**.*

7.1.2. Eingenfunciones ortogonales.

El desarrollo de una función dada en términos de las eigenfunciones de un problema de Sturm-Liouville depende de una propiedad de ortogonalidad crucial de estas eigenfunciones. Diremos que las funciones $\phi(x)$ y $\psi(x)$ son ortogonales en el intervalo $[a, b]$ con respecto a la función de peso $r(x)$ siempre que

$$\int_a^b \phi(x) \psi(x) r(x) dx = 0. \quad (21)$$

El siguiente teorema implica que cualesquiera dos eigenfunciones en un problema de Sturm-Liouville regular asociadas con eigenvalores distintos son ortogonales con respecto a la función de peso $r(x)$ en la ecuación (17).

Teorema 21 : *Ortogonalidad de eigenfunciones.*

Supóngase que las funciones p , q y r en el problema de Sturm-Liouville de las ecuaciones (17)-(18) satisfacen las hipótesis del teorema 20 y que $y_i(x)$ y $y_j(x)$ son eigenfunciones asociadas con los eigenvalores distintos λ_i y λ_j . Entonces

$$\int_a^b y_i(x) y_j(x) r(x) dx = 0. \quad (22)$$

Demostración. *Se inicia con las ecuaciones*

$$\begin{aligned} \frac{d}{dx} \left[p(x) \frac{dy_i}{dx} \right] - q(x) y_i + \lambda_i r(x) y_i &= 0 \\ \frac{d}{dx} \left[p(x) \frac{dy_j}{dx} \right] - q(x) y_j + \lambda_j r(x) y_j &= 0 \end{aligned} \quad (23)$$

Estas ecuaciones implican que λ_i , y_i y λ_j , y_j son parejas eigenvalor-eigenfunción. Si se multiplica la primera ecuación por y_j y la segunda por y_i , y luego se restan los resultados, se obtiene

$$y_j \frac{d}{dx} \left[p(x) \frac{dy_i}{dx} \right] - y_i \frac{d}{dx} \left[p(x) \frac{dy_j}{dx} \right] + (\lambda_i - \lambda_j) r(x) y_i y_j = 0.$$

Por tanto,

$$\begin{aligned}
(\lambda_i - \lambda_j) r(x) y_i y_j &= y_i \frac{d}{dx} \left[p(x) \frac{dy_j}{dx} \right] - y_j \frac{d}{dx} \left[p(x) \frac{dy_i}{dx} \right], \\
&= \frac{d}{dx} \left[p(x) \left(y_i \frac{dy_j}{dx} - y_j \frac{dy_i}{dx} \right) \right],
\end{aligned}$$

las dos últimas igualdades pueden verificarse por derivación directa. Por consiguiente, integrando de $x = a$ a $x = b$ se llega a

$$\int_a^b (\lambda_i - \lambda_j) r(x) y_i y_j dx = \left[p(x) (y_i y_j' - y_j y_i') \right]_a^b. \quad (24)$$

Por la condición de frontera dada en (18) se tiene que

$$\alpha_1 y_i(a) - \alpha_2 y_i'(a) = 0 \quad \text{y} \quad \alpha_1 y_j(a) - \alpha_2 y_j'(a) = 0.$$

Debido a que α_1 y α_2 son diferentes de cero, se concluye que el determinante de los coeficientes deben ser cero:

$$y_i(a) y_j'(a) - y_j(a) y_i'(a) = 0.$$

De manera similar, la segunda condición de la frontera en (18) implica que

$$y_i(b) y_j'(b) - y_j(b) y_i'(b) = 0.$$

De este modo, el lado derecho en la ecuación (24) se anula. Puesto que $\lambda_i \neq \lambda_j$, se llega al resultado dado en (22) y se completa la prueba. ■

7.1.3. Desarrollo en eigenfunciones.

Supongamos ahora que la función $f(x)$ puede representarse en el intervalo $[a, b]$ por una serie en términos de eigenfunciones

$$f(x) = \sum_{m=1}^{\infty} c_m y_m, \quad (25)$$

donde las funciones y_1, y_2, y_3, \dots son las eigenfunciones del problema de Sturm-Liouville regular en las ecuaciones (17)-(18). Para determinar los coeficientes c_1, c_2, c_3, \dots , se

generaliza la técnica por medio de la cual se determinaron los coeficientes de Fourier. Primero se multiplica cada lado de la ecuación (25) por $y_n(x) r(x)$ y después se integra de $x = a$ a $x = b$. Bajo la consideración de que la integración por partes es válida, se obtiene

$$\int_a^b f(x) y_n(x) r(x) dx = \sum_{m=1}^{\infty} c_m \int_a^b y_m(x) y_n(x) r(x) dx. \quad (26)$$

Pero debido a la ortogonalidad en (22), el único término diferente de cero en el lado derecho de la ecuación (26) es aquel en el que $m = n$. Así, la ecuación (26) toma la forma

$$\int_a^b f(x) y_n(x) r(x) dx = c_n \int_a^b [y_n(x)]^2 r(x) dx,$$

y por tanto

$$c_n = \frac{\int_a^b f(x) y_n(x) r(x) dx}{\int_a^b [y_n(x)]^2 r(x) dx}. \quad (27)$$

De esta manera se define la serie en términos de eigenfunciones en (25) —representando $f(x)$ en términos de las eigenfunciones del problema de Sturm-Liouville dado— con base en la elección de los coeficientes especificados en la ecuación (27).

Teorema 22 : *Convergencia de la serie de la eigenfunción.*

Sean y_1, y_2, y_3, \dots las eigenfunciones de un problema de Sturm-Liouville regular en $[a, b]$. Si la función $f(x)$ es suave por tramos en $[a, b]$, entonces la serie en términos de eigenfunciones de la ecuación (25) converge en $a < x < b$ al valor $f(x)$ en cualquier punto donde f sea continua, así como en el promedio $\frac{1}{2} [f(x^+) + f(x^-)]$ de los límites de los lados derecho e izquierdo de cada punto de discontinuidad.

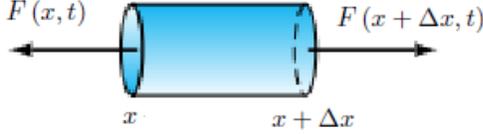
7.2. Aplicaciones de las series de eigenfunciones.

7.2.1. Vibraciones longitudinales de barras.

Supongamos que una barra uniforme elástica tiene longitud L , área de sección transversal A y densidad δ (masa/unidad de volumen), a la vez que está definida en el intervalo $0 \leq x \leq L$ cuando no se encuentra estirada. Se consideran vibraciones longitudinales de la barra aquellas en las cuales cada sección transversal (normal al eje x) se mueve únicamente en dirección x . Puede entonces describirse el movimiento de la barra en términos del desplazamiento $u(x, t)$ de la sección transversal en el tiempo t , cuya posición es x cuando la barra no está estirada (y en reposo); esta sección particular debe entenderse como la sección transversal x de la barra. Entonces la posición en el tiempo t de la sección transversal x es $x + u(x, t)$. Con base en la ley de Hooke y de la definición del módulo de Young E del material de la barra, se concluye en consecuencia que la fuerza $F(x, t)$ ejercida por la barra a la izquierda de la sección transversal x es

$$F(x, t) = -AEu_x(x, t), \quad (28)$$

donde el signo menos significa que F actúa hacia la izquierda cuando $u_x(x, t) > 0$. Para ver porqué esto es así, se considera el segmento de la barra ubicado entre la sección transversal x y la sección transversal $x + \Delta x$ como se muestra en la siguiente imagen.



Segmento pequeño de la barra.

En el tiempo t los extremos de este segmento están en $x + u(x, t)$ y $x + \Delta x + u(x + \Delta x, t)$, respectivamente, de tal manera que su longitud (originalmente $\Delta x > 0$) es ahora (usando el teorema de valor medio)

$$\Delta x + u(x + \Delta x, t) - u(x, t) = \Delta x + u_x(\hat{x}, t) \Delta x$$

para alguna \hat{x} entre x y $x + \Delta x$. De esta manera, si $u_x(x, t) > 0$ y Δx es suficientemente pequeño, entonces (por continuidad) $u_x(\hat{x}, t) > 0$. Así, el segmento está en realidad estirado a una longitud mayor que Δx , y por tanto sería necesario que las fuerzas $F(x, t)$ y $F(x + \Delta x, t)$ actuaran a la izquierda y a la derecha, respectivamente (como se indicó en la imagen anterior), para sustentar este estiramiento. Se toma la ecuación (28) como punto de partida para deducir la ecuación diferencial parcial que la función de desplazamiento $u(x, t)$ satisface cuando los desplazamientos son suficientemente pequeños para que pueda aplicarse la ley de Hooke.

Si se aplica la segunda ley de movimiento de Newton al segmento de la barra entre la sección transversal x y la sección transversal $x + \Delta x$, se obtiene

$$\begin{aligned} (\delta A \Delta x) u_{tt}(\bar{x}, t) &\approx -F(x + \Delta x, t) + F(x, t) \\ &= AE [u_x(x + \Delta x, t) - u_x(x, t)], \end{aligned} \quad (29)$$

donde \bar{x} representa el punto medio de $[x, x + \Delta x]$ porque este segmento tiene masa $\delta A \Delta x$ y aceleración aproximada $u_{tt}(\bar{x}, t)$. Cuando se divide la expresión en (29) entre $\delta A \Delta x$ y se toma el límite conforme $\Delta x \rightarrow 0$, el resultado es la ecuación de onda de una dimensión

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad (30)$$

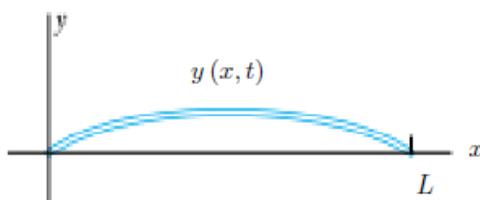
donde

$$a^2 = \frac{E}{\delta}. \quad (31)$$

Debido a que (30) es idéntica a la ecuación de la cuerda vibrando, se concluye de la solución de d'Alembert que las vibraciones (libres) longitudinales de una barra con extremos fijos están representadas por ondas de la forma $u(x, t) = g(x \pm at)$. La velocidad $a = \sqrt{E/\delta}$ con la cual estas ondas viajan es la velocidad del sonido en el material de la barra. De hecho, la ecuación de onda dada en (30) describe también las ondas del sonido ordinario de una dimensión en un gas dentro de un tubo. En este caso, en la ecuación (31) E representa los módulos de volumen (incremento fraccional en densidad por incremento unitario en presión) de gas y δ su densidad de equilibrio.

7.2.2. Vibraciones transversales de barras.

Ahora se estudian las vibraciones de una barra elástica uniforme en la cual el movimiento de cada punto no es longitudinal, sino perpendicular al eje x (el eje de la barra en su posición de equilibrio). Sea $y(x, t)$ el desplazamiento transversal de la sección x en el tiempo t , como se indica en la siguiente imagen.



Barra con vibraciones transversales.

La idea es deducir una ecuación diferencial parcial que satisfaga la función de deflexión $y(x, t)$. De acuerdo con un principio dinámico general, se puede transformar la ecuación estática $EIy^{(4)} = F$ en una ecuación dinámica (sin fuerza externa) reemplazando F con una fuerza inercial inversa $F = -\rho y_{tt}$ —donde ρ es la densidad lineal (masa/longitud) de la barra—, y reemplazando también $y^{(4)}$ por y_{xxxx} . Esto resulta en

$$EI \frac{\partial^4 y}{\partial x^4} = -\rho \frac{\partial^2 y}{\partial t^2}$$

que puede escribirse en la forma

$$\frac{\partial^2 y}{\partial t^2} + a^4 \frac{\partial^4 y}{\partial x^4} = 0, \quad (32)$$

donde

$$a^4 = \frac{EI}{\rho} \quad (33)$$

7.3. Soluciones Periódicas estacionarias Y frecuencias naturales.

La solución

$$\begin{aligned} y(x, t) &= \sum_{i=1}^{\infty} \left[A_n \cos\left(\frac{n\pi at}{L}\right) + B_n \sin\left(\frac{n\pi at}{L}\right) \right] \sin\left(\frac{n\pi x}{L}\right) \\ &= \sum_{i=1}^{\infty} C_n \cos\left(\frac{n\pi at}{L} - \gamma_n\right) \sin\left(\frac{n\pi x}{L}\right) \end{aligned} \quad (34)$$

al problema de vibración de una cuerda

$$\frac{\partial^2 y}{\partial t^2} = a^2 \frac{\partial^2 y}{\partial x^2} \quad a = \sqrt{\frac{T}{\rho}}; \quad (35)$$

$$y(0, t) = y(L, t) = 0. \quad (36)$$

$$y(x, 0) = f(x), \quad y_t(x, 0) = g(x). \quad (37)$$

La solución de la ecuación (34) describe las vibraciones de una cuerda con longitud L y de densidad ρ , bajo una tensión T ; los coeficientes constantes en la ecuación (34) se determinan por las condiciones iniciales dadas en (37).

En particular, se observa de los términos de la ecuación (34) que las frecuencias naturales (angulares) de vibración (en rad/seg) de la cuerda están dadas por

$$\omega_n = \frac{n\pi a}{L} \quad (38)$$

para $n = 1, 2, 3, \dots$. Estos son los únicos valores de ω para los cuales la ecuación (35) tiene una solución periódica permanente de la forma

$$y(x, t) = X(x) \cos(\omega t - \gamma) \quad (39)$$

que satisface las condiciones de frontera dadas en (36). Pero si se sustituye la ecuación (39) en (35) y se cancela el factor $\cos(\omega t - \gamma)$, se encuentra que $X(x)$ debe satisfacer la ecuación

$$a^2 X''(x) + \omega^2 X(x) = 0$$

cuya solución general

$$X(x) = A \cos\left(\frac{\omega x}{a}\right) + B \sin\left(\frac{\omega x}{a}\right)$$

satisface las condiciones dadas en (36) sólo si $A = 0$ y $\omega = \frac{n\pi a}{L}$ para algún entero positivo n .

7.3.1. Vibraciones forzadas y resonancia.

Ahora asuma que la cuerda está bajo la influencia de una fuerza periódica externa $F(t) = F_0 \cos(\omega t)$ (fuerza/unidad de masa) que actúa de manera uniforme sobre la cuerda a lo largo de su longitud. Entonces, de acuerdo con la ecuación (34), el desplazamiento $y(x, t)$ de la cuerda satisface la ecuación diferencial parcial no homogénea

$$\frac{\partial^2 y}{\partial t^2} = a^2 \frac{\partial^2 y}{\partial x^2} + F_0 \cos(\omega t), \quad (40)$$

junto con condiciones de frontera dadas en (36) y (37). Por ejemplo, si la cuerda inicialmente está en reposo en posición de equilibrio cuando la fuerza externa inicia su acción, se tiene que encontrar una solución de la ecuación (40) que satisfaga las condiciones

$$y(0, t) = y(L, t) = y(x, 0) = y_t(x, 0) = 0, \quad (41)$$

donde $0 < x < L$. Para lograr esto, es suficiente con obtener primero una solución particular $y_p(x, t)$ de la ecuación (40) que satisfaga las condiciones de frontera dadas en (36), y encontrar después una solución $y_c(x, t)$ como la ecuación (34) del conocido problema dado en las ecuaciones (35)-(37) con

$$f(x) = -y_p(x, 0) \quad \text{y} \quad g(x) = D_t y_p(x, 0).$$

Evidentemente,

$$y(x, t) = y_c(x, t) + y_p(x, t),$$

deberá satisfacer las ecuaciones (40) y (41). De esta manera, la nueva tarea será buscar $y_p(x, t)$. Un examen de los términos individuales en la ecuación (40) sugiere intentar con

$$y_p(x, t) = X(x) \cos(\omega t). \quad (42)$$

La sustitución de esta propuesta en la ecuación (40), después de cancelar el factor común $\cos(\omega t)$, obtiene la ecuación diferencial ordinaria

$$a^2 X'' + \omega^2 X = F_0,$$

con la solución general

$$X(x) = A \cos\left(\frac{\omega x}{a}\right) + B \sin\left(\frac{\omega x}{a}\right) - \frac{F_0}{\omega^2}. \quad (43)$$

La condición $x(0) = 0$ requiere que $A = \frac{F_0}{\omega^2}$, y entonces $X(L) = 0$ necesita que

$$X(L) = \frac{F_0}{\omega^2} \left(\cos\left(\frac{\omega L}{a}\right) - 1 \right) + B \sin\left(\frac{\omega L}{a}\right) = 0. \quad (44)$$

Ahora asumamos que la frecuencia ω de la fuerza periódica externa no es igual a ninguna de las frecuencias naturales $\omega_n = n\pi a/L$ de la cuerda. Entonces $\sin(\omega x/a) \neq 0$, de tal modo que ahora podemos despejar B en la ecuación (44) y luego sustituir el resultado en la ecuación (43) para obtener

$$X(x) = \frac{F_0}{\omega^2} \left(\cos \left(\frac{\omega L}{a} - 1 \right) \right) - \frac{F_0 [\cos(\omega L/a)]}{\omega^2 \sin(\omega L/a)} \sin \left(\frac{\omega x}{a} \right). \quad (45)$$

Como consecuencia, con esta elección de $X(x)$ la ecuación (42) proporciona la solución particular deseada $y_p(x, t)$.

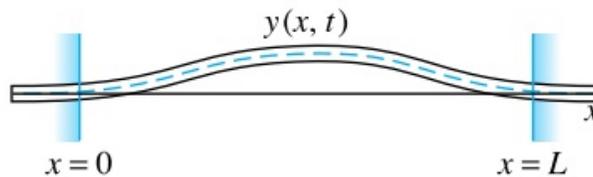
Nótese, sin embargo, que conforme el valor de ω se aproxima a $\omega_n = n\pi a/L$ con n impar, los coeficientes de $\sin(\omega x/a)$ en la ecuación (45) tiende a $\pm\infty$; así se presenta el fenómeno de la resonancia. Esto explica porqué cuando sólo se pulsa una de dos cuerdas idénticas cercanas, la otra también comenzará a vibrar debido a que su frecuencia fundamental es excitada a través del aire por una fuerza periódica externa. Obsérvese además que si $\omega = \omega_n = n\pi a/L$ con n par, entonces puede elegirse $B = 0$ en la ecuación (44) para que en este caso no haya resonancia.

La cuerda vibrando es típica de sistemas continuos con una secuencia infinita de frecuencias naturales de vibración. Cuando una fuerza externa periódica actúa sobre un sistema de este tipo, pueden ocurrir vibraciones de resonancia potencialmente destructivas si la frecuencia impuesta es muy cercana a una de las frecuencias naturales del sistema. En consecuencia, un aspecto importante de un diseño estructural adecuado es evitar este tipo de vibraciones de resonancia.

7.3.2. Frecuencias naturales de la viga.

En la figura se muestra una viga uniforme de longitud L , densidad lineal ρ y módulo de Young E , sujeta en cada extremo. Para $0 < x < L$ y $t > 0$, su función de deflexión $y(x, t)$ satisface la ecuación de cuarto orden

$$\frac{\partial^2 y}{\partial t^2} + a^4 \frac{\partial^4 y}{\partial x^4} = 0 \quad \left(a^4 = \frac{EI}{\rho} \right) \quad (46)$$



desarrollada en la sección (7.2), donde I representa el momento de inercia de la sección transversal de la viga alrededor de su eje horizontal de simetría. Dado a que tanto su

desplazamiento como su pendiente son cero en cada extremo fijo, las condiciones en la frontera son

$$y(0, t) = y_x(0, t) = 0 \quad \text{y} \quad y(L, t) = y_x(L, t) = 0. \quad (47)$$

Aquí no se consideran las condiciones iniciales porque únicamente se desea encontrar las frecuencias naturales de vibración de la viga. Las frecuencias naturales son los valores de ω para los cuales la ecuación (46) tiene una solución no trivial de la forma

$$y(x, t) = X(x) \cos(\omega t - \gamma), \quad (48)$$

que satisface las condiciones dadas en la ecuación (47). Al sustituir $y(x, t)$ de la ecuación (48) en (46) cancelando el factor común $\cos(\omega t - \gamma)$, se obtiene la ecuación diferencial ordinaria de cuarto orden $-\omega^2 X + a^4 X^{(4)} = 0$; esto es,

$$X^{(4)} - \frac{\omega^2}{a^4} X = 0. \quad (49)$$

Si se escribe $\alpha^4 = \frac{\omega^2}{a^4}$, puede expresarse la solución general de la ecuación (49) como

$$X(x) = A \cosh(\alpha x) + B \sinh(\alpha x) + C \cos(\alpha x) + D \sin(\alpha x),$$

con

$$X'(x) = \alpha(A \sinh(\alpha x) + B \cosh(\alpha x) - C \sin(\alpha x) + D \cos(\alpha x)).$$

Las primeras condiciones dadas en la ecuación (47) se obtienen

$$X(0) = A + C = 0 \quad \text{y} \quad X'(0) = \alpha(B + D) = 0$$

de tal manera que $C = -A$ y $D = -B$ por tanto, las otras condiciones de (47) consiguen

$$X(L) = A(\cosh(\alpha L) - \cos(\alpha L)) + B(\sinh(\alpha L) - \sin(\alpha L)) = 0$$

y

$$\frac{1}{\alpha} X'(L) = A(\sinh(\alpha L) + \sin(\alpha L)) + B(\cosh(\alpha L) - \cos(\alpha L)) = 0.$$

Para que estas dos ecuaciones lineales homogéneas en A y B tengan una solución no trivial, el determinante de los coeficientes debe ser cero:

$$(\cosh(\alpha L) - \cos(\alpha L))^2 - (\sinh^2(\alpha L) - \sin^2(\alpha L)) = 0$$

$$(\cosh^2(\alpha L) - \sinh^2(\alpha L)) + (\cos^2(\alpha L) - \sin^2(\alpha L)) - 2 \cosh(\alpha L) \cos(\alpha L) = 0$$

$$2 - 2 \cosh(\alpha L) \cos(\alpha L) = 0.$$

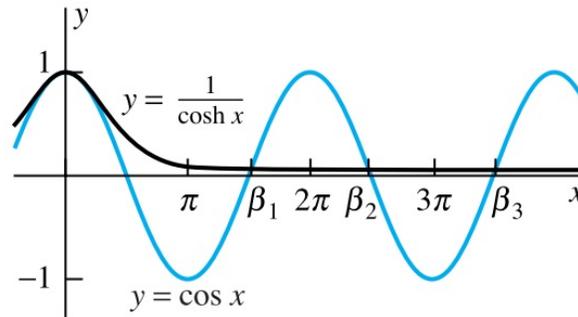
Entonces $\beta = \alpha L$ debe ser una raíz diferente de cero de la ecuación

$$\cosh(x) \cos(x) = 1. \quad (50)$$

En la figura se muestra que esta ecuación tiene una sucesión creciente de raíces positivas $\{\beta_n\}_1^\infty$. Como $\omega = \alpha^2 a^2 = \frac{\beta^2 a^2}{L^2}$ y $a^2 = \sqrt{\frac{EI}{\rho}}$, de esto se concluye que la frecuencia (angular) natural de vibración de la viga con extremos empotrados está dada por

$$\omega_n = \frac{\beta_n^2}{L^2} \sqrt{\frac{EI}{\rho}} \text{ rad/s} \quad (51)$$

Solución de $\cosh(x) \cos(x) = 1$



para $n = 1, 2, 3, \dots$. Entonces las raíces de la ecuación (50) son $\beta_1 \approx 4.73004$, $\beta_2 \approx 7.85320$, $\beta_3 \approx 10.99561$ y $\beta_n \approx \frac{(2n+1)\pi}{2}$ para $n \geq 4$ como se indica en la figura.

7.3.3. Oscilaciones de temperatura subterránea.

Considérese que la temperatura subterránea de algún lugar en particular es una función $u(x, t)$ dependiendo del tiempo t y de la profundidad x por debajo de la superficie. Entonces u satisface la ecuación de calor $u_t = ku_{xx}$, donde k es la difusividad térmica del suelo. Puede considerarse que la temperatura $u(0, t)$ en la superficie $x = 0$ se conoce a partir de los registros climatológicos. De hecho, la variación periódica de las temperaturas superficiales promedio mensuales, dependen de la estación del año, con un máximo en la mitad del verano (julio en el hemisferio norte) y un mínimo en la mitad del invierno (enero), es muy parecida a las oscilaciones del *seno* o *coseno*. Por tanto, se asumirá que

$$u(0, t) = T_0 + A_0 \cos(\omega t), \quad (52)$$

donde $t = 0$ corresponde a la mitad del verano. En este caso, T_0 es la temperatura anual promedio y A_0 la amplitud de variación de la temperatura de acuerdo con la estación, mientras que ω se elige de tal manera que el periodo de $u(0, t)$ sea exactamente de un año. (En unidades cgs, por ejemplo, ω sería 2π dividido entre el número de segundos 31, 557, 341 comprendidos en un año; así $\omega \approx 1.991 \times 10^{-7}$.)

Es razonable considerar que la temperatura a una profundidad fija también varía periódicamente en el tiempo t . Si se introduce $U(x, t) = u(x, t) - T_0$ por conveniencia, entonces interesa encontrar soluciones periódicas de la ecuación de calor de la forma

$$U(x, t) = A(x) \cos(\omega t - \gamma) = V(x) \cos(\omega t) + W(x) \sin(\omega t), \quad (53)$$

del problema:

$$\frac{\partial U}{\partial t} = k \frac{\partial^2 U}{\partial x^2} \quad (x > 0, t > 0) \quad (54)$$

$$U(0, t) = A_0 \cos(\omega t). \quad (55)$$

Para resolver este problema, considérese $U(x, t)$ de la ecuación (53) como la parte real de la función de valores complejos

$$\tilde{U}(x, t) = X(x) \exp(i\omega t). \quad (56)$$

Entonces se busca que $\tilde{U}(x, t)$ satisfaga las condiciones

$$\tilde{U}_t(x, t) = k \tilde{U}_{xx}(x, t), \quad (57)$$

$$\tilde{U}(0, t) = A_0 \exp(i\omega t). \quad (58)$$

Si se sustituye la ecuación (56) en (57), se obtiene $i\omega X = kX''$; esto es,

$$X'' - \alpha X = 0, \quad (59)$$

donde

$$\alpha = \pm \sqrt{\frac{i\omega}{k}} = \pm(1+i) \sqrt{\frac{\omega}{2k}}, \quad (60)$$

dado que $\sqrt{i} = \pm \frac{(1+i)}{\sqrt{2}}$. Por consiguiente, la solución general de la ecuación (59) es

$$X(x) = A \exp\left(- (1+i) \sqrt{\frac{\omega}{2k}} x\right) + B \exp\left((1+i) \sqrt{\frac{\omega}{2k}} x\right). \quad (61)$$

Asimismo para que $\tilde{U}(x, t)$, por tanto $X(x)$, sea acotada conforme $x \rightarrow +\infty$, es necesario que $B = 0$. Se observa también de las ecuaciones (58) y (56) que $A = X(0) = A_0$. De esta manera

$$X(x) = A_0 \exp\left(- (1+i) \sqrt{\frac{\omega}{2k}} x\right). \quad (62)$$

Finalmente, la solución del problema original de las ecuaciones (54)-(55) es:

$$\begin{aligned}
U(x, t) &= \operatorname{Re} \tilde{U}(x, t) = \operatorname{Re} X(x) \exp(i\omega t) \\
&= \operatorname{Re} \left[A_0 \exp(i\omega t) \exp \left(-(1+i)x \sqrt{\frac{\omega}{2k}} \right) \right] \\
&= \operatorname{Re} \left[A_0 \exp \left(-x \sqrt{\frac{\omega}{2k}} \right) \exp \left(i \left(\omega t - x \sqrt{\frac{\omega}{2k}} \right) \right) \right],
\end{aligned}$$

por lo tanto

$$U(x, t) = A_0 \exp \left(-x \sqrt{\frac{\omega}{2k}} \cos \left(\omega t - x \sqrt{\frac{\omega}{2k}} \right) \right). \quad (63)$$

Así, la amplitud $A(x)$ de la temperatura anual se amortigua exponencialmente como una función de la profundidad x :

$$A(x) = A_0 \exp \left(-x \sqrt{\frac{\omega}{2k}} \right). \quad (64)$$

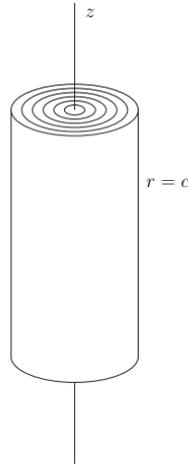
Existe además un retardo $\gamma(x) = x \sqrt{\frac{\omega}{2k}}$ en la profundidad x .

7.4. Problemas en coordenadas cilíndricas.

Cuando el laplaciano $\nabla^2 u = u_{xx} + v_{yy} + w_{zz}$ de una función $u = u(x, y, z)$ se transforma sustituyendo $x = r \cos(\theta)$, $y = r \sin(\theta)$, el resultado es el laplaciano en coordenadas cilíndricas:

$$\nabla^2 u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} + \frac{\partial^2 u}{\partial z^2}. \quad (65)$$

Para una aplicación típica, considérese un cilindro sólido uniforme muy largo de radio c , centrado a lo largo del eje z como se muestra en la siguiente imagen.



Cilindro sólido uniforme grande de radio $r = c$.

Supóngase que se calienta a una temperatura inicial **radialmente simétrica** ($t = 0$) que depende sólo de la distancia r desde el eje z (pero no de la coordenada angular θ o de la altura z de un punto). Considérese también que iniciando en el tiempo $t = 0$ la condición en la frontera

$$\left(\beta_1 u + \beta_2 \frac{\partial u}{\partial r} \right) \Big|_{r=c} = 0, \quad (66)$$

es impuesta en la superficie lateral $r = c$ del cilindro. Nótese los siguientes casos especiales para la ecuación (66):

- Se reduce a la condición $u = 0$ si $\beta_1 = 1$ y $\beta_2 = 0$.
- Se reduce a la condición de aislamiento $u_r = 0$ si $\beta_1 = 0$ y $\beta_2 = 1$.
- Se reduce a la condición de transferencia de calor $hu + u_r = 0$ si $\beta_1 = h$ y $\beta_2 = 1$.

Es razonable esperar que la temperatura u dentro del cilindro en el tiempo t depende sólo de r , por lo que puede escribirse $u = u(r, t)$. Entonces $u_{\theta\theta} = u_{zz} = 0$; así, la sustitución de la ecuación (65) en la ecuación de calor $u_t = k\nabla^2 u$ obtenemos el problema con valor en la frontera

$$\frac{\partial u}{\partial t} = k \left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} \right) \quad (r < c, t > 0); \quad (67)$$

$$\beta_1 u(c, t) + \beta_2 u_r(c, t) = 0, \quad (68)$$

$$u(r, 0) = f(r) \quad (\text{temperatura inicial}). \quad (69)$$

Para resolver este problema por separación de variables, se sustituye

$$u(r, t) = R(r)T(t),$$

en la ecuación (67); de este modo se obtiene

$$RT' = k \left(R''T + \frac{1}{r} R'T \right). \quad (70)$$

La división entre kRT resulta en

$$\frac{R'' + \frac{R'}{r}}{R} = \frac{T'}{kT} = -\lambda. \quad (71)$$

Por tanto, $R(r)$ debe satisfacer la ecuación

$$R'' + \frac{1}{r} R' + \lambda R = 0, \quad (72)$$

de la misma manera que lo hace la ecuación $\beta_1 R(c) + \beta_2 R'(c) = 0$, que se obtiene a partir de la ecuación (68). Además, la ecuación

$$T' = -\lambda kT,$$

implica que $T(t) = \exp(-\lambda kt)$, con algún múltiplo constante. Debido a que la difusividad k es positiva, se sigue que λ debe ser no negativa si $T(t)$ permanece acotada conforme $t \rightarrow +\infty$, como se requiere para el problema físico que representan las ecuaciones (67)-(69). Por tanto se escribe $\lambda = \alpha$, de tal manera que la ecuación (72) toma la forma

$$r^2 R'' + rR' + \alpha^2 r^2 R = 0. \quad (73)$$

7.4.1. Problema de Sturm-Liouville singular.

La ecuación (73) con $x = r$ y $y(x) = R(r)$ la ecuación paramétrica de Bessel

$$x^2 y'' + xy' + \alpha^2 x^2 y = 0, \quad (74)$$

de orden cero. En general, la ecuación paramétrica de Bessel

$$x^2 y'' + xy' + (\alpha^2 x^2 - n^2)y = 0, \quad (75)$$

tiene la solución general

$$y(x) = AJ_n(\alpha x) + BY_n(\alpha x), \quad (76)$$

si $\alpha = 0$. Después de dividir entre x , la ecuación de Bessel dada en (75) toma la forma de Sturm-Liouville

$$\frac{d}{dx} \left(x \frac{dy}{dx} \right) - \frac{n^2}{x} y + \lambda xy = 0, \quad (77)$$

con $p(x) = x$, $q(x) = \frac{n^2}{x}$, $r(x) = x$ y $\lambda = \alpha^2$. Se desea ahora determinar los valores no negativos de λ para los cuales existe una solución de la ecuación (77) en $(0, c)$ que es continua (al igual que su primer derivada) en el intervalo cerrado $[0, c]$ y que satisface la condición de frontera

$$\beta_1 y(c) + \beta_2 y'(c) = 0, \quad (78)$$

donde β_1 y β_2 son ceros.

El problema de Sturm-Liouville asociado con las ecuaciones (77) y (78) es singular debido a que $p(0) = r(0) = 0$ y $q(x) \rightarrow +\infty$ conforme $x \rightarrow 0$, mientras que en el teorema (20) se asumió que $p(x)$ y $r(x)$ eran positivas y que $q(x)$ era continua en todo el intervalo. Este problema tampoco puede ajustarse al patrón de la sección (7.1), donde no se impone una condición similar a la dada en (78) en la frontera izquierda $x = 0$. Sin embargo, el requisito de que $y(x)$ sea continua en $[0, c]$ juega el papel de una condición de este tipo. Debido a que $Y_n(x) \rightarrow -\infty$ conforme $x \rightarrow 0$, la solución en la ecuación (76) para $\alpha > 0$ puede ser continua en $x = 0$ sólo si $B = 0$, de tal manera que

$$y(x) = J_n(\alpha x),$$

con un múltiplo constante. Queda solamente por imponer la condición de la ecuación (78) en $x = c$.

Es conveniente distinguir los casos $\beta_2 = 0$ y $\beta_2 \neq 0$, pues si $\beta_2 = 0$, la ecuación (78) toma la forma simple

$$y(c) = 0. \quad (79)$$

Si $\beta_2 \neq 0$, se multiplica cada término en la ecuación (78) por $\frac{c}{\beta_2}$ y entonces se escribe $h = \frac{c\beta_1}{\beta_2}$ para obtener la condición equivalente

$$hy(c) + cy'(c) = 0. \quad (80)$$

Se asume de aquí en adelante que $h \geq 0$.

Caso 1 $\lambda = 0$. Primero considérese la posibilidad de un eigenvalor nulo $\lambda = 0$. Si $\lambda = 0$, $n = 0$, entonces la ecuación (77) se reduce a la ecuación $(xy')' = 0$ con solución general

$$y(x) = A \ln x + B,$$

y la continuidad en $[0, c]$ requiere que $A = 0$. Pero entonces la ecuación (79) implica que $B = 0$ al igual que en la ecuación (80), a menos que $h = 0$, caso en el cual $\lambda = 0$ es un eigenvalor con eigenfunción asociada $y(x) \equiv 1$.

Si $\lambda = 0$ pero $n > 0$, entonces (91) es simplemente la ecuación

$$x^2 y'' + xy' - n^2 y = 0,$$

con la solución general, sustituyendo la solución de prueba $y = x^k$ en

$$y(x) = Ax^n + Bx^{-n},$$

y la continuidad en $[0, c]$ requiere que $B = 0$. Pero es fácil verificar que $y(x) = Ax^n$ no satisface la ecuación (79) ni la ecuación (80) a menos que $A = 0$. De este modo, $\lambda = 0$ no es un eigenvalor si $n > 0$. Por tanto, se ha demostrado que $\lambda = 0$ es un eigenvalor del problema dado en las ecuaciones (77)-(78) si y sólo si $n = h = 0$ y la condición de frontera en $x = c$ es $y'(x) = 0$, en cuya circunstancia una eigenfunción asociada es $y(x) \equiv 1$. En este caso se escribe

$$\lambda = 0 \quad \text{y} \quad y_0(x) \equiv 1.$$

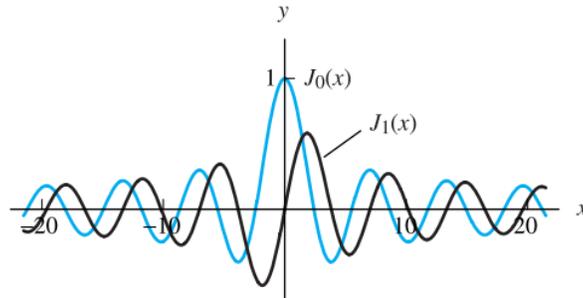
Caso 2 $\lambda > 0$ entonces $\lambda = \alpha^2 > 0$, y en este caso la única solución de la ecuación (77) que es continua en $[0, c]$ con un múltiplo constante,

$$y(x) = J_n(\alpha x).$$

Entonces la ecuación (79) implica que $J_n(\alpha c) = 0$; en otras palabras, αc debe ser una raíz positiva de la ecuación:

$$J_n(x) = 0. \quad (81)$$

Las gráficas de $J_0(x)$ y $J_1(x)$ son como las que se muestran en la siguiente gráfica.



Gráficas de la funciones de Bessel $J_0(x)$ y $J_1(x)$.

La gráfica de $J_n(x)$ para $n > 1$ se parece a las de $J_1(x)$, incluso en el caso en que $J_n(0) = 0$. En particular, para cada $n = 1, 2, 3, \dots$, la ecuación (81) tiene una sucesión creciente infinita $\{\gamma_{nk}\}_{k=1}^{\infty}$ de raíces positivas tales que $\gamma_{nk} \rightarrow +\infty$ conforme $k \rightarrow +\infty$. Estas raíces para $n \leq 8$ y $k \leq 20$.

Si $y(x) = J_n(\alpha x)$, de tal manera que $\frac{dy}{dx} = \alpha J'_n(\alpha x)$, entonces la ecuación (80) implica que

$$hJ_n(\alpha c) + \alpha c J'_n(\alpha c) = 0,$$

esto es, que αc es una raíz positiva de la ecuación

$$hJ_n(n) + xJ'_n(x) = 0. \quad (82)$$

Se sabe que esta ecuación también tiene una sucesión infinita creciente $\{\gamma_{nk}\}_{k=1}^{\infty}$ de raíces positivas que divergen a $+\infty$. Si $h = 0$, entonces la ecuación (82) se reduce a $J'_n(x) = 0$. Si cualquiera de las condiciones en la frontera en las ecuaciones (79) y (80) se cumple, entonces el k -ésimo eigenvalor positivo es

$$\lambda_k = \frac{\gamma_k^2}{c^2},$$

donde se escribe γ_k para la k -ésima raíz positiva correspondiente de las ecuaciones (81) o (82); la eigenfunción asociada es

$$y_k(x) = J_n\left(\frac{\gamma_k x}{c}\right).$$

En la siguiente tabla se resume lo anterior para una rápida referencia. El caso excepcional en que $n = h = 0$ corresponde a la condición en la frontera $y'(c) = 0$ se menciona por

separado. Se han analizado sólo eigenvalores no negativos, pero puede probarse que el problema en las ecuaciones (77)-(78) no tiene eigenvalores negativos.

Condiciones de frontera.	Eigenvalores.	Eigenfunciones asociadas.
Caso 1: $y(c) = 0$.	$\lambda_k = \gamma_k^2/c^2$; $\{\gamma_k\}_1^\infty$ las raíces positivas de $J_n(x) = 0$.	$y_k(x) = J_n\left(\frac{\gamma_k x}{c}\right)$.
Caso 2: $hy(c) + cy'(c) = 0$; $h \neq 0$ y $n \neq 0$.	$\gamma_k = \gamma_k^2/c^2$; $\{\gamma_k\}_1^\infty$ las raíces positivas de $hJ_n(x) + xJ'_n(x) = 0$	$y_k(x) = J_n\left(\frac{\gamma_k x}{c}\right)$.
Caso 3: $y'(c) = 0$, $n = 0$.	$\gamma_0 = 0$, $\gamma_k = \gamma_k^2/c^2$; $\{\gamma_k\}_1^\infty$ las raíces positivas de $J'_0(x) = 0$.	$y_0(x) = 1$. $y_k(x) = J_0\left(\frac{\gamma_k x}{c}\right)$.

Eigenvalores y eigenfunciones del problema singular de Sturm-Liouville.

$$\frac{d}{dx} \left[x \frac{dy}{dx} \right] - \frac{n^2}{x} y + \lambda xy = 0, \quad \beta_1 y(c) + \beta_2 y'(c) = 0 \text{ en } [0, c].$$

7.4.2. Serie de Fourier-Bessel.

Ahora que se sabe que el problema singular de Sturm-Liouville en las ecuaciones (77)-(78) tiene una secuencia infinita de eigenvalores y eigenfunciones asociadas similares a las del problema regular de Sturm-Liouville, se pueden analizar los desarrollos de la eigenfunción en términos de series. En los casos 1 o 2 de la tabla anterior se espera una función suave por tramos $f(x)$ en $[0, c]$ que tenga una serie de la eigenfunción de la forma

$$f(x) = \sum_{k=1}^{\infty} c_k y_k(x) = \sum_{k=1}^{\infty} c_k J_n\left(\frac{\gamma_k x}{c}\right), \quad (83)$$

mientras que en el caso de excepción 3 la serie tendrá también un término constante c_0 correspondiente a $\lambda_0 = 0$, $y_0(x) \equiv 1$. Si la conclusión del teorema (21) se ha de cumplir (a pesar de que sus hipótesis no se satisfagan), entonces las eigenfunciones

$$J_n\left(\frac{\gamma_k x}{c}\right), \quad k = 1, 2, 3, \dots$$

deben ser ortogonales en $[0, c]$ con función de peso $r(x) = x$. De hecho, si se sustituye

$$p(x) = r(x) = xy$$

$$y_k(x) = J_n\left(\frac{\gamma_k x}{c}\right), \quad y'_k(x) = \frac{\gamma_k}{c} J'_n\left(\frac{\gamma_k x}{c}\right),$$

en la ecuación (24), el resultado es

$$\begin{aligned} (\lambda_i - \lambda_j) \int_0^c x J_n\left(\frac{\gamma_i x}{c}\right) J_n\left(\frac{\gamma_j x}{c}\right) dx &= x \left[\frac{\gamma_j}{c} J_n\left(\frac{\gamma_i x}{c}\right) J'_n\left(\frac{\gamma_j x}{c}\right) - \frac{\gamma_i}{c} J_n\left(\frac{\gamma_j x}{c}\right) J'_n\left(\frac{\gamma_i x}{c}\right) \right] \\ &= \gamma_j J_n(\gamma_i) J'_n(\gamma_j) - \gamma_i J_n(\gamma_j) J'_n(\gamma_i). \end{aligned} \quad (84)$$

Es claro que la cantidad en la ecuación (84) es cero si γ_i y γ_j son ambas raíces de la ecuación (81) tenemos que $J_n(x) = 0$, mientras que si son raíces de la ecuación (82), ésta se reduce a

$$J_n(\gamma_i) [-h J_n(\gamma_j)] - J_n(\gamma_j) [-h J_n(\gamma_i)] = 0.$$

En cualquier caso, se observa que si $i \neq j$, entonces

$$\int_0^c x J_n\left(\frac{\gamma_i x}{c}\right) J_n\left(\frac{\gamma_j x}{c}\right) dx = 0. \quad (85)$$

Esta ortogonalidad con la función de peso $r(x) = x$ es lo que se necesita para determinar los coeficientes de la serie de la eigenfunción en la ecuación (83). Si se multiplica cada término de la ecuación (83) por $x J_n\left(\frac{\gamma_k x}{c}\right)$ y se integra término a término, se obtiene

$$\begin{aligned} \int_0^c x f(x) J_n\left(\frac{\gamma_k x}{c}\right) dx &= \sum_{j=1}^{\infty} c_j \int_0^c x J_n\left(\frac{\gamma_j x}{c}\right) J_n\left(\frac{\gamma_k x}{c}\right) dx \\ &= c_k \int_0^c x \left[J_n\left(\frac{\gamma_k x}{c}\right) \right]^2 dx \end{aligned}$$

por la ecuación (85). Entonces

$$c_k = \frac{\int_0^c x f(x) J_n\left(\frac{\gamma_k x}{c}\right) dx}{\int_0^c x \left[J_n\left(\frac{\gamma_k x}{c}\right) \right]^2 dx}. \quad (86)$$

Con estos coeficientes, la serie expresada en la ecuación (83) se denomina serie de Fourier-Bessel. Se sabe que una serie de Fourier-Bessel para una función suave por tramos $f(x)$ satisface la conclusión de convergencia del teorema (22). Esto es, converge en el valor promedio $\frac{1}{2} [f(x+) + f(x-)]$ en cada punto de $(0, c)$ y por tanto en el valor $f(x)$ en cada punto interior de continuidad.

7.4.3. Coeficientes de Fourier-Bessel.

A pesar de su apariencia, las integrales en el denominador de la ecuación (86) no son difíciles de evaluar. Supóngase que $y(x) = J_n(\alpha x)$, de tal manera que $y(n)$ satisface la ecuación paramétrica de Bessel de orden n ,

$$\frac{d}{dx} \left[x \frac{dy}{dx} \right] + \left(\alpha^2 x - \frac{n^2}{x} \right) y = 0. \quad (87)$$

Al multiplicar esta ecuación por $2xy'(x)$ e integrar por partes, puede obtenerse fácilmente la fórmula

$$2\alpha^2 \int_0^c x [J_n(\alpha x)]^2 dx = \alpha^2 c^2 [J'_n(\alpha c)]^2 + (\alpha^2 c^2 - n^2) [J_n(\alpha c)]^2. \quad (88)$$

Supóngase ahora que $\alpha = \frac{\gamma_k}{c}$, donde γ_k es una raíz de la ecuación $J_n(x) = 0$. Aplicamos la ecuación (88) así como la fórmula de recurrencia

$$xJ'_n(x) = nJ_n(x) - xJ_{n+1}(x),$$

lo que implica que $J'_n(\gamma_k) = -J_{n+1}(\gamma_k)$, el resultado es

$$\int_0^c x \left[J_n \left(\frac{\gamma_k x}{c} \right) \right]^2 dx = \frac{c^2}{2} [J'_n(\gamma_k)]^2 = \frac{c^2}{2} [J_{n+1}(\gamma_k)]^2. \quad (89)$$

Los otros datos de la siguiente tabla se obtienen de manera similar a partir de la ecuación (88). Las series de Fourier-Bessel con $n = 0$ son las más comunes.

Las formas que toman en los tres casos se presentan a continuación.

	$\{\gamma_k\}_1^\infty$ Raíces positivas de la ecuación.	Valor de $\int_0^c x \left[J_n \left(\frac{\gamma_k x}{c} \right) \right]^2 dx$.
Caso 1:	$J_n(x) = 0$.	$\frac{c^2}{2} [J_{n+1}(\gamma_k)]^2$.
Caso 2:	$hJ_n(x) + xJ'_n(x) = 0$. h y n ambas diferentes de cero.	$\frac{c^2 (\gamma_k^2 - n^2 + h^2)}{2\gamma_k^2} [J_n(\gamma_k)]^2$.
Caso 3:	$J'_0(x) = 0$.	$\frac{c^2}{2} [J_0(\gamma_k)]^2$.

Caso 1: Con $n = 0$: Si $\{\gamma_k\}_k^\infty$ son las raíces positivas de la ecuación $J_0(x) = 0$, entonces

$$f(x) = \sum_{k=1}^{\infty} c_k J_0\left(\frac{\gamma_k x}{c}\right); \quad (90)$$

$$c_k = \frac{2}{c^2 [J_1(\gamma_k)]^2} \int_0^c x f(x) J_0\left(\frac{\gamma_k x}{c}\right) dx.$$

Caso 2: Con $n = 0$: Si $\{\gamma_k\}_{k=1}^{\infty}$ son las raíces positivas de la ecuación $hJ_0 + J_0'(x) = 0$ con $h > 0$, entonces

$$f(x) = \sum_{k=1}^{\infty} c_k J_0\left(\frac{\gamma_k x}{c}\right); \quad (91)$$

$$c_k = \frac{2\gamma_k^2}{c^2 (\gamma_k^2 + h^2) [J_0(\gamma_k)]^2} \int_0^c x f(x) J_0\left(\frac{\gamma_k x}{c}\right) dx.$$

Caso 3: Si $\{\gamma_k\}_{k=1}^{\infty}$ son las raíces positivas de la ecuación $J_0'(x) = 0$, entonces

$$f(x) = c_0 + \sum_{k=1}^{\infty} J_0\left(\frac{\gamma_k x}{c}\right);$$

$$c_0 = \frac{2}{c^2} \int_0^{\infty} x f(x) dx, \quad (92)$$

$$c_k = \frac{2}{c^2 [J_0(\gamma_k)]^2} \int_0^c x f(x) J_0\left(\frac{\gamma_k x}{c}\right) dx.$$

CAPITULO TRES.

APROXIMACIÓN A LOS EIGENVALORES.

8. APROXIMACIÓN A LOS EIGENVALORES.

8.1. Algebra Lineal y Eigenvalores.

Cotas para aproximar los valores característicos.

Teorema 23 : (Teorema del círculo de Gerschgorin.)

Sea A una matriz de $n \times n$ y denotemos con R_i el círculo en el plano complejo con centro a_{ii} y radio $\sum_{j=1, j \neq i}^n |a_{ij}|$; es decir,

$$R_i = \left\{ z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{j=1, j \neq i}^n |a_{ij}| \right\}.$$

Donde \mathbb{C} denota el plano complejo. Los valores característicos de A están contenidos dentro de $R = \bigcup_{i=1}^n R_i$. Más aun, la unión de cualquier k de estos círculos que no tenga una intersección con los $(n - k)$ restantes contendrá exactamente k (contando multiplicidades) de los valores característicos.

Demostración:

Supongamos que λ es un valor característico de A con el vector característico asociado \mathbf{x} , donde $\|\mathbf{x}\|_\infty = 1$. Dado que $A\mathbf{x} = \lambda\mathbf{x}$, la representación equivalente por componente es

$$\sum_{j=1}^n a_{ij}x_j = \lambda x_i, \quad \text{para cada } i = 1, 2, \dots, n.$$

Si k es un entero con $|x_k| = \|\mathbf{x}\|_\infty = 1$, esta ecuación, con $i = k$, implica que

$$\sum_{j=1}^n a_{kj}x_j = \lambda x_k$$

por tanto

$$\sum_{j=1, j \neq k}^n a_{kj}x_j = (\lambda - a_{kk})x_k.$$

y

$$|\lambda - a_{kk}| \cdot |x_k| = \left| \sum_{j=1, j \neq k}^n a_{kj}x_j \right| \leq \sum_{j=1, j \neq k}^n |a_{kj}| |x_j|$$

Como $|x_j| \leq |x_k| = 1$, para toda $j = 1, 2, \dots, n$.

$$|\lambda - a_{kk}| \leq \sum_{j=1, j \neq k}^n |a_{kj}|.$$

Por tanto, $\lambda \in R_k$, lo cual prueba la primera afirmación del teorema. La segunda parte requiere un argumento de continuidad ingenioso.

8.2. Matrices Ortogonales y Transformaciones Similares.

Definición 24 : Una matriz Q se dice **ortogonal** si sus columnas $\{q_1^t, q_2^t, \dots, q_n^t\}$ forman un conjunto ortonormal en \mathbb{R}^n .

Teorema 24 : Suponga que Q de $n \times n$ es una matriz ortogonal, entonces

1. Q es invertible con $Q^{-1} = Q^t$;
2. Para cualquier \mathbf{x} y \mathbf{y} en \mathbb{R}^n , $(Q\mathbf{x})^t Q\mathbf{y} = \mathbf{x}^t \mathbf{y}$;
3. Para cualquier \mathbf{x} en \mathbb{R}^n , $\|Q\mathbf{x}\|_2 = \|\mathbf{x}\|_2$.

Definición 25 : Diremos que dos matrices A y B son similares si existe una matriz no singular S con $A = S^{-1}BS$.

Teorema 25 : Suponga que A y B son matrices similares con $A = S^{-1}BS$ y λ es un eigenvalor de A con eigenvector asociado \mathbf{x} . Entonces λ es un eigenvalor de B con eigenvector asociado $S\mathbf{x}$.

Prueba. Sea $\mathbf{x} \neq \mathbf{0}$ es tal que

$$S^{-1}BS\mathbf{x} = A\mathbf{x} = \lambda\mathbf{x}.$$

multiplicando por la izquierda por la matriz S tenemos

$$BS\mathbf{x} = \lambda S\mathbf{x}.$$

Como $\mathbf{x} \neq \mathbf{0}$ y S es no singular, $S\mathbf{x} \neq \mathbf{0}$. Entonces, $S\mathbf{x}$ es un eigenvector de B correspondiente al eigenvalor λ .

Teorema 26 : Una matriz A de $n \times n$ a una matriz diagonal D si y solo si A tiene n eigenvectores linealmente independientes. En este caso, $D = S^{-1}AS$, donde las columnas de S consisten de los eigenvectores, y el i -ésimo elemento de la diagonal de D es el eigenvalor de A que corresponde a la i -ésima columna de S , donde las matrices S y D no son únicas.

Corolario 2 : Una matriz A de $n \times n$ que tiene n eigenvalores distintos es similar a una matriz diagonal.

Teorema 27 : (*Schur*)

Sea A una matriz arbitraria. Existe una matriz U no singular con la siguiente propiedad

$$T = U^{-1}AU.$$

donde T es una matriz triangular superior cuyas entradas diagonales consisten de los eigenvalores de A .

Teorema 28 : Una matriz A de $n \times n$ es simétrica si y solo si existe una matriz diagonal D y una matriz ortogonal Q con $A = QDQ^t$.

Prueba.

Primero supongamos que $A = QDQ^t$, donde Q es ortogonal y D es diagonal. Entonces

$$A^t = (QDQ^t)^t = (Q^t)^t D Q^t = QDQ^t = A,$$

de aquí A es simétrica.

Para probar que cada matriz simétrica A puede escribirse en la forma $A = QDQ^t$, primero consideremos los eigenvalores de A . Si $A\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ y $A\mathbf{v}_2 = \lambda_2\mathbf{v}_2$, con $\lambda_1 \neq \lambda_2$, entonces $A^t = A$ por lo que tenemos

$$(\lambda_1 - \lambda_2)\mathbf{v}_1^t \mathbf{v}_2 = (\lambda_1\mathbf{v}_1)^t \mathbf{v}_2 - \mathbf{v}_1^t (\lambda_2\mathbf{v}_2) = (A\mathbf{v}_1)^t \mathbf{v}_2 - \mathbf{v}_1^t (A\mathbf{v}_2) = \mathbf{v}_1^t A^t \mathbf{v}_2 - \mathbf{v}_1^t A \mathbf{v}_2 = 0,$$

también $\mathbf{v}_1^t \mathbf{v}_2 = 0$. Podemos seleccionar vectores ortonormales para distintos eigenvalores simplemente normalizando todos los eigenvectores ortogonales. Cuando los eigenvectores no son distintos, así tenemos sub-espacios de eigenvectores para cada múltiplo de eigenvalores, y con la ayuda del proceso de ortogonalización de Gram-Schmidt podemos encontrar un conjunto de n eigenvectores ortonormales.

Corolario 3 : Suponga que A de $n \times n$ es una matriz simétrica, entonces existen n eigenvectores de A que forman un conjunto ortonormal, y los eigenvalores de A son números reales.

Prueba. Si $Q = q_{ij}$ y $D = d_{ij}$ son las matrices especificadas en el teorema (28) entonces

$$D = Q^t A Q = Q^{-1} A Q \text{ implica que } A Q = Q D.$$

Sea $1 \leq i \leq n$ y sea $\mathbf{v}_i = (q_{1i}, \dots, q_{ni})^t$ la i -ésima columna de Q . Entonces

$$A \mathbf{v}_i = d_{ii} \mathbf{v}_i,$$

y d_{ii} es un eigenvalor de A con eigenvector, \mathbf{v}_i , de la i -ésima columna de Q . Las columnas de Q son ortonormales, también los eigenvectores de A son ortonormales.

Multiplicando esta ecuación por la izquierda por \mathbf{v}_i^t tenemos

$$\mathbf{v}_i^t A \mathbf{v}_i = d_{ii} \mathbf{v}_i^t \mathbf{v}_i.$$

Entonces $\mathbf{v}_i^t A \mathbf{v}_i$ y $\mathbf{v}_i^t \mathbf{v}_i$ son números reales y $\mathbf{v}_i^t \mathbf{v}_i = 1$, los eigenvalores $d_{ii} = \mathbf{v}_i^t A \mathbf{v}_i$ es un número real, para cada $i = 1, 2, \dots, n$.

Teorema 29 : Una matriz A es definida positiva si y solo si todos los eigenvalores de A son positivos.

Prueba. Primero supongamos que A es definida positiva y que λ es un eigenvalor de A con eigenvector asociado respectivo \mathbf{x} , con $\|\mathbf{x}\|_2 = 1$. Entonces

$$0 < \mathbf{x}^t A \mathbf{x} = \lambda \mathbf{x}^t \mathbf{x} = \lambda \|\mathbf{x}\|_2^2 = \lambda.$$

Ahora, supongamos que A es simétrica con eigenvalores positivos. Por corolario anterior, A tiene n eigenvectores, $\mathbf{v}^1, \mathbf{v}^2, \dots, \mathbf{v}^n$, que forman un conjunto ortonormal y, por tanto el conjunto es linealmente independiente. Entonces para cualquier $\mathbf{x} \neq 0$ existen n $\beta_1, \beta_2, \dots, \beta_n$ constantes únicas distintas de cero tal que

$$\mathbf{x} = \sum_{i=1}^n \beta_i \mathbf{v}^i,$$

Multiplicando por $\mathbf{x}^t A$ tenemos

$$\mathbf{x}^t A \mathbf{x} = \mathbf{x}^t \left(\sum_{i=1}^n \beta_i A \mathbf{v}^i \right) = \mathbf{x}^t \left(\sum_{i=1}^n \beta_i \lambda_i \mathbf{v}^i \right) = \sum_{j=1}^n \sum_{i=1}^n \beta_j \beta_i \lambda_i (\mathbf{v}^j)^t \mathbf{v}^i.$$

Pero los vectores $\mathbf{v}^1, \mathbf{v}^2, \dots, \mathbf{v}^n$ también forman un conjunto ortonormal

$$(\mathbf{v}^j)^t \mathbf{v}^i = \begin{cases} 0, & \text{si } i \neq j, \\ 1, & \text{si } i = j. \end{cases}$$

Esto resulta con el hecho que los λ_i son todos positivos, lo que implica que

$$\mathbf{x}^t A \mathbf{x} = \sum_{j=1}^n \sum_{i=1}^n \beta_j \beta_i \lambda_i (\mathbf{v}^j)^t \mathbf{v}^i = \sum_{i=1}^n \lambda_i \beta_i^2 > 0.$$

Entonces A es definida positiva.

8.3. Método de la potencia.

El método de la potencia es una técnica iterativa que permite determinar el valor característico dominante de una matriz, es decir, el valor característico con mayor magnitud. Un aspecto útil del método de la potencia es que no sólo produce un valor característico, sino un vector característico asociado.

Para aplicar el método de la potencia supondremos que la matriz A de $n \times n$ tiene n valores característico $\lambda_1, \lambda_2, \dots, \lambda_n$ con un conjunto asociado de vectores característicos linealmente independientes $\{v^{(1)}, v^{(2)}, \dots, v^{(n)}\}$. Más aún, supondremos que A tiene exactamente un valor característico, λ_1 , cuya magnitud es la mayor, por lo que $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n| \geq 0$.

Si \mathbf{x} es un vector cualquiera en \mathbb{R}^n , el hecho de que $\{v^{(1)}, v^{(2)}, \dots, v^{(n)}\}$ sea linealmente independiente implica que existe las constantes $\beta_1, \beta_2, \dots, \beta_n$ con

$$\mathbf{x} = \sum_{j=1}^n \beta_j v^{(j)}.$$

Al multiplicar ambos lados de la ecuación por $A, A^2, A^3, \dots, A^k, \dots$ obtenemos:

$$A\mathbf{x} = \sum_{j=1}^n \beta_j A v^{(j)} = \sum_{j=1}^n \beta_j \lambda_j v^{(j)},$$

$$A^2\mathbf{x} = \sum_{j=1}^n \beta_j \lambda_j A v^{(j)} = \sum_{j=1}^n \beta_j \lambda_j^2 v^{(j)}.$$

En general

$$A^k \mathbf{x} = \sum_{j=1}^n \beta_j \lambda_j^k v^{(j)}.$$

Si factorizamos λ_1^k en cada término de la derecha de la última ecuación, entonces

$$A^k \mathbf{x} = \lambda_1^k \sum_{j=1}^n \beta_j \left(\frac{\lambda_j}{\lambda_1} \right)^k v^{(j)}.$$

Como $|\lambda_1| > |\lambda_j|$, para cualquier $j = 2, \dots, n$, tenemos $\lim_{k \rightarrow \infty} (\lambda_j/\lambda_1)^k = 0$, y

$$\lim_{k \rightarrow \infty} A^k \mathbf{x} = \lim_{k \rightarrow \infty} \lambda_1^k \beta_1 v^{(1)}. \quad (93)$$

Esta sucesión converge a cero si $|\lambda_1| < 1$ y diverge si $|\lambda_1| > 1$, naturalmente a condición de que $\beta_1 \neq 0$.

De la relación expresada en la ecuación (93) puede obtenerse una ventaja al escalar en forma adecuada las potencias de $A^k \mathbf{x}$ para asegurarnos de que el límite de la ecuación (93) sea finito y no cero. El escalamiento comienza seleccionando \mathbf{x} como un vector unitario $\mathbf{x}^{(0)}$ en relación con $\|\cdot\|_\infty$ y un componente $x_{p_0}^{(0)}$ de $\mathbf{x}^{(0)}$ con

$$x_{p_0}^{(0)} = 1 = \|\mathbf{x}^{(0)}\|_\infty.$$

Sea $\mathbf{y}^{(1)} = A\mathbf{x}^{(0)}$ y definimos $\mu^{(1)} = y_{p_0}^{(1)}$. Con esta notación.

$$\mu^{(1)} = y_{p_0}^{(1)} = \frac{y_{p_0}^{(1)}}{x_{p_0}^{(0)}} = \frac{\beta_1 \lambda_1 v_{p_0}^{(0)} + \sum_{j=2}^n \beta_j \lambda_j v_{p_0}^{(j)}}{\beta_1 v_{p_0}^{(1)} + \sum_{j=2}^n \beta_j v_{p_0}^{(j)}} = \lambda_1 \left[\frac{\beta_1 v_{p_0}^{(0)} + \sum_{j=2}^n \beta_j (\lambda_j/\lambda_1) v_{p_0}^{(j)}}{\beta v_{p_0}^{(1)} + \sum_{j=2}^n \beta_j v_{p_0}^{(j)}} \right].$$

Entonces, sea p_1 el menor entero tal que

$$|y_{p_1}^{(1)}| = \|\mathbf{y}^{(1)}\|_\infty,$$

y definimos $\mathbf{x}^{(1)}$ por medio de

$$\mathbf{x}^{(1)} = \frac{1}{y_{p_1}^{(1)}} \mathbf{y}^{(1)} = \frac{1}{y_{p_1}^{(1)}} A \mathbf{x}^{(0)}.$$

Entonces

$$x_{p_1}^{(1)} = 1 = \|\mathbf{x}^{(1)}\|_\infty.$$

A continuación definimos

$$\mathbf{y}^{(2)} = A \mathbf{x}^{(1)} = \frac{1}{y_{p_1}^{(1)}} A^2 \mathbf{x}^{(0)}$$

y

$$\mu^{(2)} = y_{p_1}^{(2)} = \frac{y_{p_1}^{(2)}}{x_{p_1}^{(1)}} = \frac{\left[\beta_1 \lambda_1^2 v_{p_1}^{(1)} + \sum_{j=2}^n \beta_j \lambda_j^2 v_{p_1}^{(j)} \right]}{\left[\beta_1 \lambda_1 v_{p_1}^{(1)} + \sum_{j=2}^n \beta_j \lambda_j v_{p_1}^{(j)} \right]} = \lambda_1 \frac{\left[\beta_1 v_{p_1}^{(1)} + \sum_{j=2}^n \beta_j (\lambda_j / \lambda_1)^2 v_{p_1}^{(j)} \right]}{\left[\beta_1 v_{p_1}^{(1)} + \sum_{j=2}^n \beta_j (\lambda_j / \lambda_1) v_{p_1}^{(j)} \right]}.$$

Sea p_2 el entero más pequeño con

$$|y_{p_2}^{(2)}| = \|\mathbf{y}^{(2)}\|_\infty.$$

Y definimos

$$\mathbf{x}^{(2)} = \frac{1}{y_{p_2}^{(2)}} \mathbf{y}^{(2)} = \frac{1}{y_{p_2}^{(2)}} A \mathbf{x}^{(1)} = \frac{1}{y_{p_2}^{(2)} y_{p_1}^{(1)}} A^2 \mathbf{x}^{(0)}.$$

De modo semejante definimos las sucesiones de los vectores $\{\mathbf{x}^{(m)}\}_{m=0}^\infty$ y $\{\mathbf{y}^{(m)}\}_{m=0}^\infty$, inductivamente hacemos lo mismo con una sucesión de escalares $\{\mu^{(m)}\}_{m=1}^\infty$ mediante

$$\begin{aligned} \mathbf{y}^{(m)} &= A \mathbf{x}^{(m-1)} \\ \mu^{(m)} = y_{p_{m-1}}^{(m)} &= \lambda_1 \frac{\left[\beta_1 v_{p_{m-1}}^{(1)} + \sum_{j=2}^n (\lambda_j / \lambda_1)^m \beta_j v_{p_{m-1}}^{(j)} \right]}{\left[\beta_1 v_{p_{m-1}}^{(1)} + \sum_{j=2}^n (\lambda_j / \lambda_1)^{m-1} \beta_j v_{p_{m-1}}^{(j)} \right]} \end{aligned} \quad (94)$$

y

$$x^{(m)} = \frac{y^{(m)}}{y_{p_m}^{(m)}} = \frac{A^{(m)} x^{(0)}}{\prod_{k=1}^m y_{p_k}^{(k)}},$$

donde cada paso p_m sirve para representar el entero más pequeño para el cual

$$|y_{p_m}^{(m)}| = \|\mathbf{y}^{(m)}\|_\infty$$

Al examinar la ecuación (94) vemos que, como $|\lambda_j/\lambda_1| < 1$ para cualquier $j = 2, 3, \dots, n$, $\lim_{m \rightarrow \infty} \mu^{(m)} = \lambda_1$, siempre y cuando elijamos $x^{(0)}$ de modo que $\beta_1 \neq 0$. Más aún, la sucesión de vectores $\{\mathbf{x}^{(m)}\}_{m=0}^\infty$ converge al vector característico asociado a λ_1 , que tiene norma uno I_∞ .

Desventajas del método de la potencia.

La desventaja es que al inicio no se sabe si la matriz tiene o no un solo valor característico dominante. Tampoco se sabe cómo seleccionar $\mathbf{x}^{(0)}$ para estar seguro de que su representación mediante vectores característicos de la matriz contenga una contribución distinta de cero del vector característico asociado al valor característico dominante, en caso de que exista.

Observación:

- La rapidez con que $\{\mu^{(m)}\}_{m=1}^\infty$ converge a λ_1 se determina mediante las razones $|\lambda_j/\lambda_1|^m$, para $j = 2, 3, \dots, n$, y, particularmente, mediante $|\lambda_2/\lambda_1|^m$. La rapidez de convergencia es $O(|\lambda_2/\lambda_1|^m)$.
- No es necesario que la matriz tenga valores característicos distintos para que converja el método de la potencia. Es decir si el valor característico dominante y único, λ_1 , tiene una multiplicidad r mayor que 1, y $v^{(1)}, v^{(2)}, \dots, v^{(r)}$ son vectores característico linealmente independientes asociados a λ_1 , el procedimiento todavía convergerá a λ_1 . En este caso, la sucesión de vectores $\{\mathbf{x}^{(m)}\}_{m=0}^\infty$ convergerá a un vector característico de λ_1 de I_∞ con norma uno, que es una combinación lineal de $v^{(1)}, v^{(2)}, \dots, v^{(r)}$ y depende de la elección del vector inicial $\mathbf{x}^{(0)}$.

Cuando A es simétrica, podemos hacer una variación en la elección de los vectores $\mathbf{x}^{(m)}$, $\mathbf{y}^{(m)}$ y escalares $\mu^{(m)}$ para mejorar significativamente la razón de convergencia entre la sucesión $\{\mu^{(m)}\}_{m=1}^\infty$ y el valor característico dominante λ_1 . De hecho, aunque la razón de convergencia del método general de potencia es $O(|\lambda_2/\lambda_1|^m)$, la del método modificado para matrices simétricas es $O(|\lambda_2/\lambda_1|^{2m})$.

El siguiente teorema ofrece una cota de error para aproximar los valores característicos de una matriz simétrica.

Teorema 30 : Si A es una matriz simétrica de $n \times n$ con los valores característicos $\lambda_1, \lambda_2, \dots, \lambda_n$ y $\|A\mathbf{x} - \lambda\mathbf{x}\|_2 < \epsilon$ para algún vector \mathbf{x} con $\|\mathbf{x}\|_2 = 1$ y con el número real λ , entonces

$$\min_{1 \leq j \leq n} |\lambda_j - \lambda| < \epsilon.$$

Método de la potencia inversa.

Es una modificación del método de la potencia que ofrece una convergencia más rápida. Se usa para determinar el valor característico de A más cercano a un número q específico.

Supongamos que la matriz A tiene los valores característicos $\lambda_1, \lambda_2, \dots, \lambda_n$ con los vectores característicos linealmente independientes $v^{(1)}, v^{(2)}, \dots, v^{(n)}$. Consideremos la matriz $(A - qI)^{-1}$, donde $q \neq \lambda_i$ para $i = 1, 2, \dots, n$. Los valores característicos de $(A - qI)^{-1}$ son

$$\frac{1}{\lambda_1 - q}, \quad \frac{1}{\lambda_2 - q}, \quad \dots, \quad \frac{1}{\lambda_n - q},$$

con los vectores característicos $v^{(1)}, v^{(2)}, \dots, v^{(n)}$. Al aplicar el método de la potencia en $(A - qI)^{-1}$ obtenemos

$$\begin{aligned} \mathbf{y}^{(m)} &= (A - qI)^{-1} \mathbf{x}^{(m-1)}, \\ \mu^{(m)} &= y_{p_{m-1}}^{(m)} = \frac{\sum_{j=1}^n \beta_j \frac{1}{(\lambda_j - q)^m} v_{p_{m-1}}^{(j)}}{\sum_{j=1}^n \beta_j \frac{1}{(\lambda_j - q)^{m-1}} v_{p_{m-1}}^{(j)}}, \end{aligned} \quad (95)$$

y

$$\mathbf{x}^{(m)} = \frac{\mathbf{y}^{(m)}}{y_{p_m}^{(m)}},$$

donde p_m representa en cada paso el entero más pequeño para el cual $|y_{p_m}^{(m)}| = \|\mathbf{y}^{(m)}\|_\infty$. La sucesión $\{\mu^{(m)}\}$ de la ecuación (3) converge a $1/(\lambda_k - q)$, donde

$$\frac{1}{|\lambda_k - q|} = \max_{1 \leq i \leq n} \frac{1}{|\lambda_i - q|},$$

y $\lambda_k \approx q + 1/\mu^{(m)}$ es el valor característico de A más cercano a q .

Cuando k se conoce, la ecuación (95) puede escribirse así.

$$\mu^{(m)} = \frac{1}{\lambda_k - q} \left[\frac{\beta_k v_{p_{m-1}}^{(k)} + \sum_{j=1, j \neq k}^n \beta_j \left[\frac{\lambda_k - q}{\lambda_j - q} \right]^m v_{p_{m-1}}^{(j)}}{\beta_k v_{p_{m-1}}^{(k)} + \sum_{j=1, j \neq k}^n \beta_j \left[\frac{\lambda_k - q}{\lambda_j - q} \right]^{m-1} v_{p_{m-1}}^{(j)}} \right]. \quad (96)$$

Por tanto, la elección de q determina la convergencia siempre y cuando $1/(\lambda_k - q)$ sea un valor característico dominante y único de $(A - qI)^{-1}$ (aunque puede ser un valor característico múltiple). Cuando más se acerque q a un valor característico λ_k de A , más rápido será la convergencia, porque ésta es del orden de

$$O\left(\left|\frac{(\lambda - q)^{-1}}{(\lambda_k - q)^{-1}}\right|^m\right) = O\left(\left|\frac{(\lambda_k - q)}{(\lambda - q)}\right|^m\right),$$

donde λ representa el valor característico de A , que es el segundo más cercano a q .

El cálculo del vector $\mathbf{y}^{(m)}$ se obtiene de la ecuación

$$(A - qI)\mathbf{y}^{(m)} = \mathbf{x}^{(m-1)}.$$

En términos generales, para resolver este sistema se utiliza la eliminación gaussiana con pivote.

La elección de q puede tener como base el teorema del círculo de Gerschgorin.

Hay muchas técnicas para obtener aproximaciones a los otros valores característicos de una matriz, una vez calculada una aproximación al valor característico dominante.

8.4. Método de Householders.

Transformación Householders.

Definición 26 : Sea $\mathbf{w} \in \mathbb{R}^n$ con $\mathbf{w}^t \mathbf{w} = 1$. Entonces la matriz de $n \times n$

$$P = I - 2\mathbf{w}\mathbf{w}^t,$$

es llamada una **Transformación Householders**.

Teorema 31 : Una transformación de Householders, $P = I - 2\mathbf{w}\mathbf{w}^t$, es simétrica y ortogonal, también $P^{-1} = P$.

Prueba. Se sigue de

$$(\mathbf{w}\mathbf{w}^t)^t = (\mathbf{w}^t)^t \mathbf{w}^t = \mathbf{w}\mathbf{w}^t,$$

que

$$P^t = (I - 2\mathbf{w}\mathbf{w}^t)^t = I - 2\mathbf{w}\mathbf{w}^t = P.$$

Ademas, $\mathbf{w}^t \mathbf{w} = 1$, también

$$PP^t = (I - 2\mathbf{w}\mathbf{w}^t)(I - 2\mathbf{w}\mathbf{w}^t) = I - 2\mathbf{w}\mathbf{w}^t - 2\mathbf{w}\mathbf{w}^t + 4\mathbf{w}\mathbf{w}^t \mathbf{w}\mathbf{w}^t = I$$

y $P^{-1} = P^t = P$.

El método Householder's comienza con la determinación de una transformación $P^{(1)}$ tal que $A^{(2)} = P^{(1)}A^{(1)}$ donde

$$a_{j1}^{(2)} = 0 \text{ para } j = 3, \dots, n. \quad (97)$$

Por simetría también tenemos $a_{1j}^{(2)} = 0$.

Elegimos un vector $\mathbf{w} = (w_1, w_2, \dots, w_n)^t$ también que $\mathbf{w}^t \mathbf{w} = 1$, con la ecuación (97) y la matriz

$$A^{(2)} = P^{(1)} A P^{(2)} = (I - 2\mathbf{w}\mathbf{w}^t)A(I - 2\mathbf{w}\mathbf{w}^t),$$

tenemos $a_{11}^{(2)} = a_{11}$ y $a_{j1}^{(2)} = 0$ para $j = 3, \dots, n$ esta elección impone n condiciones a las n incógnitas w_1, \dots, w_n .

Seleccionando $w_1 = 0$ nos aseguramos que $a_{11}^{(2)} = a_{11}$ y queremos que

$$P^{(1)} = I - 2\mathbf{w}\mathbf{w}^t,$$

satisfaga

$$P^{(1)}(a_{11}, a_{21}, \dots, a_{n1})^t = (a_{11}, \alpha, 0, \dots, 0)^t, \quad (98)$$

donde α lo seleccionaremos más tarde, para simplificar la notación anterior denotamos por

$$\hat{\mathbf{w}} = (w_2, \dots, w_n)^t \in \mathbb{R}^{n-1}, \quad \hat{\mathbf{y}} = (a_{21}, \dots, a_{n1})^t \in \mathbb{R}^{n-1},$$

y

$$\hat{P} = I_{n-1} - 2\hat{\mathbf{w}}\hat{\mathbf{w}}^t.$$

Sustituyendo en la ecuación (98) obtenemos

$$P^{(1)} \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & & & & & & \\ \vdots & & \hat{P} & & & & \\ 0 & & & & & & \end{bmatrix} \cdot \begin{bmatrix} a_{11} \\ \hat{\mathbf{y}} \end{bmatrix} = \begin{bmatrix} a_{11} \\ \hat{P}\hat{\mathbf{y}} \end{bmatrix} = \begin{bmatrix} a_{11} \\ \alpha \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

con

$$\hat{P}\hat{\mathbf{y}} = (I_{n-1} - 2\hat{\mathbf{w}}\hat{\mathbf{w}}^t)\hat{\mathbf{y}} = \hat{\mathbf{y}} - 2(\hat{\mathbf{w}}^t \hat{\mathbf{y}})\hat{\mathbf{w}} = (\alpha, 0, \dots, 0)^t. \quad (99)$$

Sea $r = \hat{\mathbf{w}}^t \hat{\mathbf{y}}$ entonces $(\alpha, 0, \dots, 0)^t = (a_{21} - 2rw_2, a_{31} - 2rw_3, \dots, a_{n1} - 2rw_n)^t$, y podremos determinar todos los w_i , si logramos determinar α y r , igualando las componentes tenemos

$$\alpha = a_{21} - 2rw_2,$$

y

$$0 = a_{j1} - 2rw_j \quad \text{para } i = 3, \dots, n.$$

Entonces

$$2rw_2 = a_{21} - \alpha, \quad (100)$$

y

$$2rw_j = a_{j1} \quad \text{para } j = 3, \dots, n. \quad (101)$$

Elevando ambos lados de las ecuaciones anteriores y sumando los correspondientes términos tenemos

$$4r^2 \sum_{j=2}^n = (a_{21} - \alpha)^2 + \sum_{j=3}^n a_{j1}^2.$$

Como $\mathbf{w}\mathbf{w}^t = 1$ y $w_1 = 0$ tenemos que $\sum_{j=2}^n = 1$, de aquí

$$4r^2 = \sum_{j=2}^n a_{j1}^2 - 2\alpha a_{21} + \alpha^2. \quad (102)$$

De la ecuación (99) y del hecho que P es ortogonal implica que

$$\alpha^2 = (\alpha, 0, \dots, 0)(\alpha, 0, \dots, 0)^t = (\hat{P}\hat{\mathbf{y}})^t \hat{P}\hat{\mathbf{y}} = \hat{\mathbf{y}}^t \hat{P}^t \hat{P}\hat{\mathbf{y}}.$$

Entonces

$$\alpha^2 = \sum_{j=2}^n a_{j1}^2,$$

sustituyendo este resultado en la ecuación (102), tenemos

$$2r^2 = \sum_{j=2}^n a_{j1}^2 - \alpha a_{21}.$$

Haciendo $2r^2 = 0$ se cumple solo si $A_{21} = a_{31} = \dots = a_{n1} = 0$, eligiendo

$$\alpha = -\text{sgn}(a_{21}) \left(\sum_{j=2}^n a_{j1}^2 \right)^{1/2},$$

Lo que implica que

$$2r^2 = \sum_{j=2}^n a_{j1}^2 + |a_{21}| \left(\sum_{j=2}^n a_{j1}^2 \right)^{1/2}.$$

Con esta elección de α y $2r^2$, resolviendo la ecuaciones (100) y (101) obtenemos

$$w_2 = \frac{a_{21} - \alpha}{2r} \quad \text{y} \quad w_j = \frac{a_{2j}}{2r}, \quad \text{para } j = 3, \dots, n.$$

En resumen para la transformación de $P^{(1)}$ tenemos

$$\alpha = -\text{sign}(a_{21}) \left(\sum_{j=2}^n a_{j1}^2 \right)^{1/2},$$

$$r = \left(\frac{1}{2}\alpha^2 - \frac{1}{2}a_{21}\alpha \right)^{1/2},$$

$$w_1 = 0,$$

$$w_2 = \frac{a_{21} - \alpha}{2r},$$

y

$$w_j = \frac{a_{j1}}{2r}, \quad \text{para } j = 3, \dots, n.$$

Con esta elección

$$A^{(2)} = P^{(1)} A P^{(1)} = \begin{bmatrix} a_{11}^{(2)} & a_{12}^{(2)} & 0 & \cdots & 0 \\ a_{21}^{(2)} & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} \\ 0 & a_{32}^{(2)} & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & a_{n2}^{(2)} & a_{n3}^{(2)} & \cdots & a_{nn}^{(2)} \end{bmatrix}.$$

Habiendo encontrado $P^{(1)}$ y calculado $A^{(2)}$, el proceso se repite para $k = 2, 3, \dots, n-2$, como sigue

$$\alpha = -\text{sgn} \left(a_{k+1,k}^{(k)} \right) \left(\sum_{j=k+1}^n (a_{jk}^{(k)})^2 \right)^{1/2},$$

$$r = \left(\frac{1}{2} \alpha^2 - \frac{1}{2} a_{k+1,k}^{(k)} \alpha \right)^{1/2},$$

$$w_1^{(k)} = w_2^{(k)} = \cdots = w_k^{(k)} = 0,$$

$$w_{k+1}^{(k)} = \frac{a_{k+1,k}^{(k)} - \alpha}{2r},$$

$$w_j^{(k)} = \frac{a_{jk}^{(k)}}{2r}, \quad \text{para } j = k+2, k+3, \dots, n,$$

$$P^{(k)} = I - 2\mathbf{w}^{(k)} \cdot (\mathbf{w}^{(k)})^t,$$

y

$$A^{(k+1)} = P^{(k)} A^{(k)} P^{(k)}.$$

8.5. Método de deflación.

En el método de deflación debemos formar una matriz B nueva cuyos valores característicos sean iguales a los de A , salvo que el valor característico dominante de A se reemplaza con el valor característico 0 en B .

Teorema 32 : Supongamos que $\lambda_1, \lambda_2, \dots, \lambda_n$ son valores característicos de A con los vectores característicos asociados $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \dots, \mathbf{v}^{(n)}$ y que λ_1 tiene multiplicidad 1. Sea \mathbf{x} el vector con $\mathbf{x}^t \mathbf{v}^{(1)} = 1$. Entonces, la matriz

$$B = A - \lambda_1 \mathbf{v}^{(1)} \mathbf{x}^t,$$

tiene los valores característicos $0, \lambda_2, \lambda_3, \dots, \lambda_n$ con los vectores característicos asociados $\mathbf{v}^{(1)}, \mathbf{w}^{(2)}, \mathbf{w}^{(3)}, \dots, \mathbf{w}^{(n)}$, donde $\mathbf{v}^{(i)}$ y $\mathbf{w}^{(i)}$ están relacionados por la ecuación

$$\mathbf{v}^{(i)} = (\lambda_i - \lambda_1)\mathbf{w}^{(i)} + \lambda_1(\mathbf{x}^t \mathbf{w}^{(i)})\mathbf{v}^{(1)}, \quad (103)$$

para cualquier $i = 2, 3, 4, \dots, n$.

La deflación de Wielandt se inicia con la definición

$$\mathbf{x} = \frac{1}{\lambda_1 v_i^{(1)}} (a_{i1}, a_{i2}, \dots, a_{in})^t,$$

donde $v_i^{(1)}$ es una coordenada del vector característico $\mathbf{v}^{(1)}$ distinto de cero y los valores $a_{i1}, a_{i2}, \dots, a_{in}$ son los elementos del i -ésimo renglón de A .

Con esta definición veamos que satisface las condiciones del teorema

$$\mathbf{x}^t \mathbf{v}^{(1)} = \frac{1}{\lambda_1 v_i^{(1)}} [a_{i1}, a_{i2}, \dots, a_{in}] (\mathbf{v}_1^{(1)}, \mathbf{v}_2^{(1)}, \dots, \mathbf{v}_n^{(1)})^t = \frac{1}{\lambda_1 v_i^{(1)}} \sum_{j=1}^n a_{ij} v_j^{(1)},$$

donde la suma es la i -ésima coordenada del producto $A\mathbf{v}^{(1)}$. Dado que $A\mathbf{v}^{(1)} = \lambda_1 \mathbf{v}^{(1)}$, esto significa que

$$\sum_{j=1}^n a_{ij} v_j^{(1)} = \lambda_1 v_i^{(1)},$$

lo que implica que

$$\mathbf{x}^t \mathbf{v}^{(1)} = \frac{1}{\lambda_1 v_i^{(1)}} (\lambda_1 v_i^{(1)}) = 1.$$

Por lo tanto \mathbf{x} satisface las hipótesis del teorema. Más aun, el i -ésimo renglón de $B = A - \lambda_1 \mathbf{v}^{(1)} \mathbf{x}^t$ consta por completo de elementos de cero. Donde λ_1 es el valor característico más grande de A en valor absoluto, con $\mathbf{v}^{(1)}$ el vector asociado para λ_1 .

Si $\lambda \neq 0$ es un valor característico con vector característico asociado \mathbf{w} , la relación $B\mathbf{w} = \lambda\mathbf{w}$ implica que la i -ésima coordenada de \mathbf{w} también debe ser cero. En consecuencia, la columna de la matriz B no aporta nada al producto $B\mathbf{w} = \lambda\mathbf{w}$. Así, la matriz B puede ser reemplazada por una matriz B' de $(n-1) \times (n-1)$ obtenida al suprimir en B el i -ésimo renglón y la i -ésima columna. La matriz B' tiene los valores característicos $\lambda_2, \lambda_3, \dots, \lambda_n$. Si $|\lambda_2| > |\lambda_3|$, se aplica nuevamente el método de la potencia a la matriz B' para determinar este nuevo valor característico dominante y un vector característico $\mathbf{w}^{(2)}$ asociado a λ_2 respecto a la matriz B' .

Aunque este proceso de deflación puede servir para obtener aproximaciones a todos los valores y vectores característicos de una matriz, el proceso es susceptible al error de redondeo. Si lo empleamos para calcular todos los valores característicos de una matriz, las aproximaciones conseguidas deberán usarse como valores iniciales del método de la potencia inversa aplicado a la matriz original. Esto garantiza que las aproximaciones converjan

en los valores característicos de la matriz original, no a los de la matriz reducida, que probablemente contenga errores.

Algoritmo QR .

Los métodos de deflación generalmente no son adecuados para calcular todos los valores característicos de una matriz, debido al crecimiento del error de redondeo. El algoritmo QR , es una técnica de reducción matricial que permite determinar simultáneamente todos los valores característicos de una matriz simétrica.

Para aplicar el método QR , partimos de una matriz simétrica en forma tridiagonal; es decir, las únicas entradas no nulas de la matriz están en la diagonal o en las subdiagonales directamente arriba o abajo de la diagonal. Si está no es la forma de la matriz simétrica, el primer paso consiste en aplicar el método de Householder para calcular una matriz simétrica tridiagonal similar a la matriz dada.

Asumamos que la matriz simétrica A cuyos valores característicos calcularemos es tridiagonal.

Denotemos por A la matriz de este tipo, así podremos simplificar un poco la notación.

$$A = \begin{bmatrix} a_1 & b_2 & 0 & \cdots & \cdots & 0 \\ b_2 & a_2 & b_3 & \ddots & & \vdots \\ 0 & b_3 & a_3 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & \ddots & b_n \\ 0 & \cdots & \cdots & 0 & b_n & a_n \end{bmatrix}.$$

Si $b_2 = 0$ o $b_n = 0$, entonces la matriz 1×1 $[a_1]$ o bien $[a_n]$ produce inmediatamente un valor característico a_1 o a_n de A .

Cuando $b_j = 0$ para algún j , donde $2 < j < n$, el problema se puede minimizar considerando, en vez de A , las matrices más pequeñas

$$\begin{bmatrix} a_1 & b_2 & 0 & \cdots & \cdots & 0 \\ b_2 & a_2 & b_3 & \ddots & & \vdots \\ 0 & b_3 & a_3 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & \ddots & b_{j-1} \\ 0 & \cdots & \cdots & 0 & b_{j-1} & a_{j-1} \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} a_j & b_{j+1} & 0 & \cdots & \cdots & 0 \\ b_{j+1} & a_{j+1} & b_{j+2} & \ddots & & \vdots \\ 0 & b_{j+2} & a_{j+2} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & \ddots & b_n \\ 0 & \cdots & \cdots & 0 & b_n & a_n \end{bmatrix}.$$

Si ninguna b_j es cero, el método QR forma una sucesión de las matrices $A = A^{(1)}, A^{(2)}, A^{(3)}, \dots$, así;

1. $A^{(1)} = A$ se factoriza como un producto $A^{(1)} = Q^{(1)}R^{(1)}$, donde $Q^{(1)}$ es ortogonal y $R^{(1)}$ es triangular superior.
2. $A^{(2)}$ se define como $A^{(2)} = R^{(1)}Q^{(1)}$.

En general, $A^{(i)}$ se factoriza como un producto $A^{(i)} = Q^{(i)}R^{(i)}$ de una matriz ortogonal $Q^{(i)}$ y de una matriz triangular superior $R^{(i)}$. Después, definimos $A^{(i+1)}$ como el producto de $R^{(i)}$ y $Q^{(i)}$ en la dirección inversa $A^{(i+1)} = R^{(i)}Q^{(i)}$. Dado que $Q^{(i)}$ es ortogonal, $R^{(i)} = Q^{(i)t}A^{(i)}$ y

$$A^{(i+1)} = R^{(i)}Q^{(i)} = (Q^{(i)t}A^{(i)})Q^{(i)} = Q^{(i)t}A^{(i)}Q^{(i)}.$$

Y $A^{(i+1)}$ es simétrica con los mismos valores característicos que $A^{(i)}$. Por la forma en que definimos $R^{(i)}$ y $Q^{(i)}$, también podemos garantizar que $A^{(i+1)}$ es tridiagonal.

Continuando por inducción, $A^{(i+1)}$ tiene los mismos valores característicos que la matriz original A . El éxito del procedimiento se debe al hecho de que $A^{(i+1)}$ tiende a una matriz diagonal con los valores característicos de A a lo largo de la diagonal.

Para describir la construcción de las matrices de factor $Q^{(i)}$ y $R^{(i)}$, necesitamos manejar el concepto de matriz de rotación.

Definición 27 : Una matriz de rotación P difiere de la matriz identidad en cuatro elementos como máximo. Esto tiene la forma

$$P_{ii} = P_{jj} = \cos(\theta) \quad y \quad P_{ij} = -P_{ji} = \sin(\theta),$$

para algún θ e $i \neq j$.

En cualquier matriz de rotación P , la matriz AP difiere de A sólo en la i -ésima y j -ésima columnas y la matriz PA difiere de A sólo en la i -ésimo y j -ésimo renglones. Para cualquier $i \neq j$, podemos elegir el ángulo θ de modo que el producto PA tenga un elemento cero para $(PA)_{ij}$.

La factorización de $A^{(1)}$ en $A^{(1)} = Q^{(1)}R^{(1)}$ utiliza un producto de $n - 1$ matrices de rotación de este tipo para construir

$$R^{(1)} = P_n P_{n-1} \dots P_2 A^{(1)}.$$

Primero establecemos que la matriz de rotación P_2 tenga

$$P_{11} = P_{22} = \cos(\theta_2) \quad y \quad P_{12} = -P_{21} = \sin(\theta_2),$$

donde

$$\sin(\theta_2) = \frac{b_2}{\sqrt{b_2^2 + a_1^2}} \quad y \quad \cos(\theta_2) = \frac{a_1}{\sqrt{b_2^2 + a_1^2}},$$

entonces, la matriz

$$A_2^{(1)} = P_2 A^{(1)},$$

tiene un cero en la posición $(2, 1)$, ya que el elemento $(2, 1)$ de $A_2^{(1)}$ es

$$(-\sin(\theta_2))a_1 + (\cos(\theta_2))b_2 = \frac{-b_2 a_1}{\sqrt{b_2^2 + a_1^2}} + \frac{a_1 b_2}{\sqrt{b_2^2 + a_1^2}} = 0.$$

Como la multiplicación $P_2 A^{(1)}$ afecta a los renglones 1 y 2 de $A^{(1)}$, la nueva matriz no necesariamente conserva los elementos cero en las posiciones $(1, 3), (1, 4), \dots, (1, n)$. Sin embargo, $A^{(1)}$ es tridiagonal y, por tanto, los elementos $(1, 4), \dots, (1, n)$ de $A_2^{(1)}$ son cero. Sólo el elemento $(1, 3)$, el del primer renglón y tercera columna, puede hacerse distintos de cero.

En términos generales, seleccionamos la matriz P_k de modo que el $(k, k - 1)$ -ésimo elemento de $A_k^{(1)} = P_k A_{k-1}^{(1)}$ sea cero, lo cual hace que el $(k - 1, k + 1)$ elemento se convierta en uno distinto de cero. La matriz $A_k^{(1)}$ tiene la forma

$$A_k^{(1)} = \begin{bmatrix} z_1 & q_1 & r_1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & & & & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & \ddots & 0 & z_{k-1} & q_{k-1} & r_{k-1} & \ddots & & \vdots \\ \vdots & & & \ddots & 0 & x_k & y_k & 0 & \ddots & \vdots \\ \vdots & & & & \ddots & b_{k+1} & a_{k+1} & b_{k+1} & \ddots & 0 \\ \vdots & & & & & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & & & & \ddots & \ddots & \ddots & b_n \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & b_n & a_n \end{bmatrix}.$$

Y P_{k+1} tiene la forma

$$P_{k+1} = \left[\begin{array}{c|cc|c} I_{k-1} & & O & O \\ \hline O & c_{k+1} & s_{k+1} & O \\ \hline O & -s_{k+1} & c_{k+1} & O \\ \hline O & & O & I_{n-k-1} \end{array} \right],$$

donde O denota la matriz nula con la dimensión adecuada.

Elegimos las constantes $c_{k+1} = \cos(\theta_{k+1})$ y $s_{k+1} = \sin(\theta_{k+1})$ en P_{k+1} de modo que el $(k + 1, k)$ elemento de $A_{k+1}^{(1)}$ sea cero; es decir, $s_{k+1}x_k - c_{k+1}b_{k+1} = 0$. Puesto que $c_{k+1}^2 + s_{k+1}^2 = 1$, la solución a esta ecuación es

$$s_{k+1} = \frac{b_{k+1}}{\sqrt{b_{k+1}^2 + x_k^2}} \quad \text{y} \quad c_{k+1} = \frac{x_k}{\sqrt{b_{k+1}^2 + x_k^2}},$$

y $A_{k+1}^{(1)}$ tiene la forma

$$A_{k+1}^{(1)} = \begin{bmatrix} z_1 & q_1 & r_1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & & & & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & \ddots & 0 & z_k & q_k & r_k & \ddots & & \vdots \\ \vdots & & & \ddots & 0 & x_{k+1} & y_{k+1} & 0 & \ddots & \vdots \\ \vdots & & & & \ddots & b_{k+2} & a_{k+2} & b_{k+2} & \ddots & 0 \\ \vdots & & & & & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & & & & \ddots & \ddots & \ddots & b_n \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & b_n & a_n \end{bmatrix}$$

Al proseguir con esta construcción en la sucesión P_2, \dots, P_n obtenemos la matriz triangular superior

$$R^{(1)} \equiv A_n^{(1)} = \begin{bmatrix} z_1 & q_1 & r_1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & & & & \ddots & \ddots & \ddots & r_{n-2} \\ \vdots & & & & & & & \ddots & z_{n-1} & q_{n-1} \\ 0 & \cdots & 0 & x_n \end{bmatrix}$$

La otra mitad de la factorización QR es la matriz

$$Q^{(1)} = P_2^t P_3^t \dots P_n^t,$$

porque la ortogonalidad de las matrices de rotación implican que

$$Q^{(1)} R^{(1)} = (P_2^t P_3^t \dots P_n^t) \cdot (P_n \dots P_3 P_2) A^{(1)} = A^{(1)}.$$

La matriz $Q^{(1)}$ es ortogonal porque

$$(Q^{(1)})^t Q^{(1)} = (P_2^t P_3^t \dots P_n^t)^t (P_2^t P_3^t \dots P_n^t) = (P_n \dots P_3 P_2) (P_2^t P_3^t \dots P_n^t) = 1.$$

Además, $Q^{(1)}$ es una matriz Hessenberg superior. En consecuencia, $A^{(2)} = R^{(1)} Q^{(1)}$ es también una matriz Hessenberg superior, ya que la multiplicación de $Q^{(1)}$ de la izquierda por la matriz triangular superior $R^{(1)}$ no influye en los elementos del triángulo inferior. Lo anterior implica que $A^{(2)}$ se acerca más a ser una matriz diagonal que $A^{(1)}$. El proceso se repite para construir $A^{(3)}, A^{(4)}, \dots$.

8.6. Descomposición en Valores Singulares.

En esta sección consideramos la factorización de una matriz general A de $m \times n$ llamada Descomposición en Valores Singulares. Esta factorización toma la forma

$$A = USV^t.$$

Donde U de $m \times m$ es una matriz ortogonal, V de $n \times n$ una matriz ortogonal, y S de $m \times n$ una matriz cuyos elementos distintos de cero se encuentra a lo largo de la diagonal principal. Vamos a suponer que $m \geq n$.

Definición 28 : Sea A de $m \times n$ una matriz.

1. El rango de A denotado por $\text{Rango}(A)$ es el número de renglones linealmente independientes de A .
2. La nulidad de A , denotado por $\text{Nulidad}(A)$, es $n - \text{Rango}(A)$, y describe el conjunto mas grande de vectores linealmente independientes \mathbf{v} en \mathbb{R}^n para el cual $A\mathbf{v} = 0$.

Teorema 33 : El número de renglones linealmente independiente de una matriz A de $m \times n$ es el mismo que el número de columnas linealmente independientes de A .

Teorema 34 : Sea A de $m \times n$ una matriz.

1. Las matrices $A^t A$ y AA^t son simétricas.
2. $\text{Nulidad}(A) = \text{Nulidad}(A^t A)$.
3. $\text{Rango}(A) = \text{Rango}(A^t A)$.
4. Los eigenvalores de $A^t A$ y de AA^t son reales y no negativos.

5. Los eigenvalores no negativos de AA^t y A^tA son los mismos.

Prueba.

(1) Porque $(A^tA)^t = A^t(A^t)^t = A^tA$, esta matriz es simétrica y también es similar a AA^t .

(2) Sea $\mathbf{v} \neq \mathbf{0}$ un vector con $A\mathbf{v} = \mathbf{0}$. entonces

$$(A^tA)\mathbf{v} = A^t(A\mathbf{v}), \quad \text{tambien } Nulidad(A) \leq Nulidad(A^tA).$$

Ahora supongamos que \mathbf{v} es un vector con $A^tA\mathbf{v} = \mathbf{0}$. Entonces

$$0 = \mathbf{v}^t A^t A \mathbf{v} = (A\mathbf{v})^t A \mathbf{v} = \|A\mathbf{v}\|_2^2, \quad \text{eso implica que } A\mathbf{v} = \mathbf{0}.$$

Entonces la $Nulidad(A^tA) \leq Nulidad(A)$. como consecuencia, $Nulidad(A^tA) = Nulidad(A)$.

(3) Las matrices A y A^t ambas tiene n columnas y tienen el mismo grado de nulidad, también

$$Rango(A) = n - Nulidad(A) = n - Nulidad(A^tA) = Rango(A^tA).$$

(4) La matriz A^tA es simétrica entonces sus eigenvalores son números reales esto por Corolario (3). Supongamos que \mathbf{v} es un eigenvector de A^tA con $\|\mathbf{v}\|_2 = 1$ correspondiente al eigenvalor λ . Entonces

$$0 \leq \|A\mathbf{v}\|_2^2 = (A\mathbf{v})^t(A\mathbf{v}) = \mathbf{v}^t A^t A \mathbf{v} = \mathbf{v}^t (A^t A \mathbf{v}) = \mathbf{v}^t (\lambda \mathbf{v}) = \lambda \mathbf{v}^t \mathbf{v} = \lambda \|\mathbf{v}\|_2^2 = \lambda.$$

(5) Sea \mathbf{v} un eigenvector de un eigenvalor $\lambda \neq 0$ de A^tA . Entonces

$$A^t A \mathbf{v} = \lambda \mathbf{v} \quad \text{implica que } (AA^t)A\mathbf{v} = \lambda A\mathbf{v}.$$

Si $A\mathbf{v} = \mathbf{0}$, entonces $A^t A \mathbf{v} = A^t \mathbf{0} = \mathbf{0}$, contradiciendo la asunción que $\lambda \neq 0$. Entonces $A\mathbf{v} \neq \mathbf{0}$ y $A\mathbf{v}$ es un eigenvector de AA^t , asociado con λ . La conclusión reversa también se sigue de este argumento porque si $\lambda \neq 0$ es un eigenvalor de $AA^t = (A^t)^t$, entonces λ también es un eigenvalor de $A^t(A^t)^t = A^tA$.

□

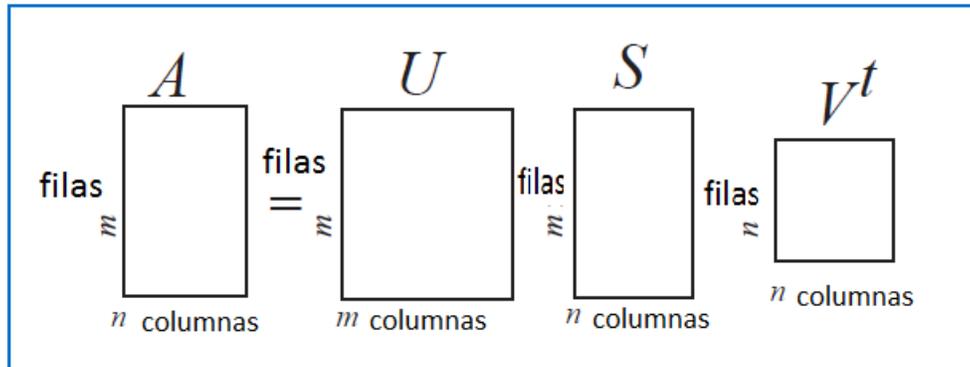
Construcción en Descomposición de Valores Singulares.

Una de las propiedades de la descomposición de valores singulares de una matriz general es que se permite una generalización de los eigenvalores y los eigenvectores en esta situación.

Nuestro objetivo es determinar una factorización de la matriz A de $n \times n$, donde $m \geq n$, en la forma

$$A = USV^t,$$

donde U es una matriz ortogonal de $m \times m$, V es una matriz ortogonal de $n \times n$ y S es una matriz diagonal, que sus entradas distintas de cero son $(S)_{ii} \equiv s_i \geq 0$, para $i = 1, \dots, n$. Como se muestra en la siguiente imagen.



Construcción de S en la Factorización $A = USV^t$.

Nosotros construimos la matriz S encontrando los eigenvalores de la matriz simétrica $A^t A$. Estos eigenvalores son números reales no negativos, y los ordenamos de mayor a menor denotados por

$$s_1^2 \geq s_2^2, \dots, s_k^2 = \dots = s_n = 0.$$

Denotando por s_k^2 al eigenvalor mas pequeño que sea distinto de cero de $A^t A$. La raíz cuadrada positiva de estos eigenvalores de $A^t A$ tiene como entradas en la diagonal de S . Ellos son llamados **Valores Singulares** de A . Entonces

$$S = \begin{bmatrix} s_1 & 0 & \dots & 0 \\ 0 & s_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & s_n \\ 0 & \dots & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \dots & 0 \end{bmatrix},$$

donde $s_i = 0$ cuando $k < i \leq n$.

Definición 29 : Los *valores singulares* de una matriz A de $m \times n$ son las raíces cuadradas positivas de los eigenvalores distintos de cero de la matriz simétrica $A^t A$.

Construcción de V en la Factorización $A = USV^t$.

La matriz $A^t A$ de $n \times n$ es simétrica, también por el teorema (28), nosotros tenemos una factorización

$$A^t A = V D V^t,$$

donde D es una matriz diagonal cuyas entradas en la diagonal son los eigenvalores de $A^t A$, y V es una matriz ortogonal cuya i -ésima columna es un eigenvector con norma 1 correspondiente al eigenvalor sobre la i -ésima entrada de la diagonal de D . La matriz diagonal específica depende del orden de los eigenvalores a lo largo de la diagonal. En la elección de la matriz D también son escritos en orden decreciente. Las columnas denotadas por $\mathbf{v}_1^t, \mathbf{v}_2^t, \dots, \mathbf{v}_n^t$, de la matriz ortogonal V de $n \times n$ son eigenvectores ortogonales correspondientes a los eigenvalores correspondientes. Múltiples eigenvalores de $A^t A$ permite múltiples elecciones de correspondientes eigenvectores, también D no está únicamente definida, del mismo modo para V . No hay problema con la elección de V , porque los eigenvalores de $A^t A$ son todos no negativos, teniendo así $D = S^2$.

Construcción de U en la Factorización de $A = USV^t$.

Una construcción de la matriz U de $m \times m$, consideramos los valores distintos de cero $s_1 \geq s_2 \geq \dots \geq s_k > 0$ y las correspondientes columnas en V dados por $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$. Definimos

$$\mathbf{u}_i = \frac{1}{s_i} A \mathbf{v}_i, \text{ para } i = 1, 2, \dots, k.$$

Usamos estos como los primeros k de m columnas de U . porque A es de $m \times n$ y cada \mathbf{v}_i es de $n \times 1$, el vector \mathbf{u}_i es de $m \times 1$, como se requiere. En adición, para cada $1 \leq i \leq k$ y $1 \leq j \leq k$, el hecho que los vectores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ son eigenvectores de $A^t A$ que forman un conjunto ortonormal implica que

$$\mathbf{u}_i^t \mathbf{u}_j = \left(\frac{1}{s_i} A \mathbf{v}_i \right)^t \frac{1}{s_j} A \mathbf{v}_j = \frac{1}{s_i s_j} \mathbf{v}_i^t A^t A \mathbf{v}_j = \frac{1}{s_i s_j} \mathbf{v}_i^t A s_j^2 \mathbf{v}_j = \frac{s_j}{s_i} \mathbf{v}_i^t \mathbf{v}_j = \begin{cases} 0 & \text{si } i \neq j, \\ 1 & \text{si } i = j. \end{cases}$$

También las primeras k columnas de U forman un conjunto de vectores ortonormales en \mathbb{R}^m . Mas aun necesitamos $m - k$ columnas adicionales de U . Para esto primero necesitamos encontrar primero $m - k$ vectores que unidos a los primeros k vectores columnas formen un conjunto linealmente independientes. Entonces podemos aplicar el método de Gram-Schmidt para obtener las columnas apropiadas adicionales.

Verificando la factorización $A = USV^t$.

Para verificar que el proceso da la Factorización $A = USV^t$, primero recalcamos que la transpuesta de una matriz ortogonal es también la inversa de la matriz. Entonces para mostrar que $A = USV^t$, podemos mostrar la relación equivalente $AV = US$.

Los vectores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ forman una base en \mathbb{R}^n , $A\mathbf{v}_i = s_i\mathbf{u}_i$, para $i = 1, \dots, k$ y $A\mathbf{v}_i = 0$, para $i = k + 1, \dots, n$. Solamente las primeras k columnas de U producen entradas distintas de cero en el producto US , también tenemos

$$\begin{aligned} A\mathbf{v} &= A[\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_n] \\ &= [A\mathbf{v}_1 \ A\mathbf{v}_2 \ \dots \ A\mathbf{v}_n] \\ &= [s_1\mathbf{u}_1 \ s_2\mathbf{u}_2 \ \dots \ s_k \ \mathbf{u}_k \ 0 \ \dots \ 0] \\ &= [\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_k \ 0 \ \dots \ 0] \begin{bmatrix} s_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & s_k & 0 & \dots & 0 \\ 0 & \dots & \dots & 0 & 0 & \dots & 0 \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \dots & \dots & 0 & 0 & \dots & 0 \end{bmatrix} = US. \end{aligned}$$

Esto completa la construcción de la descomposición en valores singulares de A .

Un método alternativo para encontrar U .

La parte (v) del teorema (34) indica que los eigenvalores de distintos de cero de A^tA y los de AA^t son los mismos. En adición a los correspondientes eigenvalores de las matrices simétricas A^tA y AA^t forman un subconjunto ortonormal completo de \mathbb{R}^n y \mathbb{R}^m , respectivamente.

En resumen, para determinar la descomposición en valores singulares de la matriz A de $m \times n$ podemos:

- Encontrar los eigenvalores $s_1^2 \leq s_2^2 \leq \dots \leq s_k^2 \leq s_{k+1} = \dots = 0$ para la matriz simétrica A^tA y colocar la raíz cuadrada positiva de s_i^2 en la entrada diagonal de $(S)_{ii}$ de $n \times n$ de la matriz S .
- Encontrar un conjunto ortonormal de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ correspondientes a los eigenvalores de A^tA y construir la matriz V de $n \times n$ con esos vectores como columnas.

- Encontrar un conjunto ortonormal de eigenvectores $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ correspondientes a los eigenvalores de AA^t y construir la matriz U de $m \times m$ con esos vectores como columnas.

Entonces A tiene la descomposición de valores singulares $A = USV^t$.

CAPITULO CUATRO.

MÉTODOS NUMÉRICOS DE PROBLEMAS DE VALOR DE FRONTERA.

9. MÉTODOS NUMÉRICOS DE PROBLEMAS DE VALOR DE FRONTERA.

9.1. Método del disparo lineal.

El siguiente teorema establece las condiciones generales que garantizan que exista la solución a un problema de valor de frontera de segundo orden y que dicha solución sea única.

Teorema 35 : Supongamos que la función f en el problema de valor de frontera

$$y'' = f(x, y, y'), \quad a \leq x \leq b, \quad y(a) = \alpha, \quad y(b) = \beta,$$

es continua en el conjunto

$$D = \{(x, y, y') \mid a \leq x \leq b, \quad -\infty < y < \infty, \quad -\infty < y' < \infty, \},$$

y que f_y y $f_{y'}$ también son continuas en D . Si

i) $f_y(x, y, y') > 0 \quad \forall (x, y, y') \in D, y$

ii) Existe una constante M , con

$$|f_{y'}(x, y, y')| \leq M, \quad \forall (x, y, y') \in D,$$

entonces el problema de valor de frontera tiene una solución única.

Ejemplo 4 : El problema de valor de frontera

$$y'' + \exp(-xy) + \sin(y') = 0, \quad 1 \leq x \leq 2, \quad y(1) = y(2) = 0,$$

tiene

$$f(x, y, y') = -\exp(-xy) - \sin(y').$$

Puesto que

$$f_y(x, y, y') = xe^{-xy} > 0, \quad \text{ya que } 1 \leq x \leq 2,$$

y

$$|f_{y'}(x, y, y')| = |-\cos(y')| \leq 1,$$

este problema tiene solución única, por el teorema (35).

Cuando $f(x, y, y')$ tiene la forma

$$f(x, y, y') = p(x)y' + q(x)y + r(x),$$

la ecuación diferencial

$$y'' = f(x, y, y'),$$

es una ecuación diferencial **lineal**. Este tipo de problemas ocurren frecuentemente, y en este caso el teorema (35) puede simplificarse de la siguiente forma.

Corolario 4 : Si el problema lineal de valor de frontera

$$y'' = p(x)y' + q(x)y + r(x), \quad a \leq x \leq b, \quad y(a) = \alpha, \quad y(b) = \beta, \quad (104)$$

satisface

(i) $p(x)$, $q(x)$ y $r(x)$ son continuas en $[a, b]$,

(ii) $q(x) > 0 \forall x \in [a, b]$,

entonces el problema tiene una única solución.

Para esta prueba nos vamos a auxiliar de los siguientes teoremas:

Teorema 36 : (**Teorema de alternativa**).

Si $p(x)$, $q(x)$ y $r(x)$ son continuas en $[a, b]$, se da una y solo una de las siguientes alternativas:

■ El problema (104) tiene solución única.

■ El problema homogéneo denotado por $L[y] = y'' - p(x)y' - q(x)y = 0$, $y(a) = 0$, $y(b) = 0$, tiene solución distinta de la trivial.

Prueba.

Si el problema homogéneo asociado tiene solución distinta de la trivial, la solución de (104) en caso de existir, no es única, pues se puede obtener otra distinta sumándole una solución del problema homogéneo por consiguiente, las dos alternativas no se pueden dar simultáneamente.

Veamos que siempre se da una de las dos alternativas. Sean y_1 y y_2 las soluciones de (104) teniendo así

$$\begin{aligned} L[y_1] &= r(x), & y_1(a) &= \alpha, & y_1'(a) &= 0, \\ L[y_2] &= 0, & y_2(a) &= 0, & y_2'(a) &= 1. \end{aligned} \quad (105)$$

De aquí con las hipótesis del corolario y el teorema **de existencia y unicidad de problemas de valor inicial** el sistema de ecuaciones (105) tiene solución única en $[a, b]$.

Ahora si $y_2(b) = 0$, entonces y_2 es una solución no trivial del problema homogéneo.

Si $y_2(b) \neq 0$, la función:

$$y(x) = y_1(x) + \frac{\beta - y_1(b)}{y_2(b)} y_2(x), \quad (106)$$

es solución de (104). Si esta solución no fuera única, restando dos soluciones distintas obtendríamos una solución no trivial del problema homogéneo.

□

Hasta ahora hemos probado que (104) tiene solución, falta ver que es única, para ello debemos probar que el problema homogéneo $L[y]$, tiene la solución trivial. Para ello haremos uso del principio del máximo.

Teorema 37 : (Principio del máximo).

Sea $u \in C^2(a, b) \cap C([a, b])$ y $q(x) \geq 0$ en $[a, b]$.

(i) Si $L[u] \leq 0$ en (a, b) , entonces $\max_{[a, b]} u \leq \max\{u(a), u(b), 0\}$.

(ii) Si $L[u] \geq 0$ en (a, b) , entonces $\min_{[a, b]} u \geq \min\{u(a), u(b), 0\}$.

Prueba. (i) Supongamos para empezar que $L[u] < 0$ en (a, b) . Sea $\hat{x} \in (a, b)$ un punto de máximo. De aquí tenemos que $u'(\hat{x}) = 0$ y $u''(\hat{x}) \leq 0$, si $u(\hat{x}) \geq 0$ se tiene que

$$L[u](\hat{x}) = -u''(\hat{x}) + p(\hat{x})u'(\hat{x}) + q(\hat{x})u(\hat{x}) \geq 0,$$

lo que es una contradicción.

Si la desigualdad no es estricta, consideremos

$$u_\epsilon(x) = u(x) + \epsilon e^{Mx},$$

con $\epsilon > 0$ pequeño y $M > 0$ grande. Se tiene que

$$L[u_\epsilon](x) = L[u](x) + \epsilon e^{Mx} (-M^2 + Mp(x) + q(x)).$$

Tomando M suficientemente grande (cuán grande, depende sólo de $\max_{[a, b]} |p|$ y $\max_{[a, b]} q$), se tiene que $L[u_\epsilon] < 0$. Así pues

$$\max_{[a, b]} u < \max_{[a, b]} u_\epsilon \leq \max\{u(a), u(b), 0\}.$$

Pasando al límite $\epsilon \rightarrow 0$ se tiene el resultado.

(ii) Se demuestra aplicando el apartado (i) a $-u$.

□

Con estos teoremas tenemos casi terminada la prueba, ya que por hipótesis $q(x) > 0$ en $[a, b]$. Aplicando el principio del máximo al problema homogéneo, tenemos que su única solución es la trivial, así por teorema (36) nos garantiza que el sistema (104) tenga solución única.

□

Probemos que la ecuación (106) es solución de la ecuación (104), primero calculemos la primera derivada de la ecuación (106) así:

$$y'(x) = y_1'(x) + \frac{\beta - y_1(b)}{y_2(b)} y_2'(x),$$

ahora derivemos nuevamente:

$$y''(x) = y_1''(x) + \frac{\beta - y_1(b)}{y_2(b)} y_2''(x),$$

Ahora sustituyamos $y_1''(x)$, $y_2''(x)$ y asociemos convenientemente de la siguiente manera:

$$\begin{aligned} y'' &= p(x)y_1' + q(x)y_1 + r(x) + \frac{\beta - y_1(b)}{y_2(b)}(p(x)y_2' + q(x)y_2), \\ &= p(x) \left(y_1' + \frac{\beta - y_1(b)}{y_2(b)} y_2' \right) + q(x) \left(y_1 + \frac{\beta - y_1(b)}{y_2(b)} y_2 \right) + r(x), \\ &= p(x)y' + q(x)y + r(x). \end{aligned}$$

Veamos que cumple con su valor de frontera:

$$\begin{aligned} y(a) &= y_1(a) + \frac{\beta - y_1(b)}{y_2(b)} y_2(a), \\ &= \alpha + \frac{\beta - y_1(b)}{y_2(b)} 0, \\ &= \alpha, \end{aligned}$$

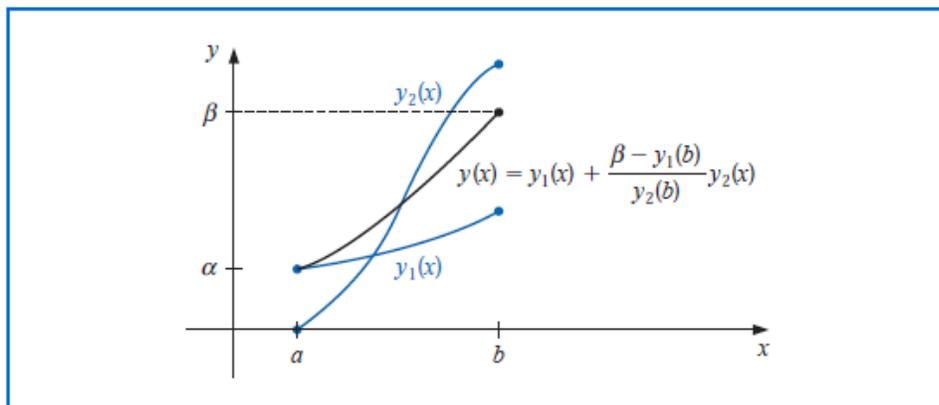
y

$$\begin{aligned} y(b) &= y_1(b) + \frac{\beta - y_1(b)}{y_2(b)} y_2(b), \\ &= y_1(b) + \frac{\beta - y_1(b)}{y_2(b)} y_2(b), \\ &= y_1(b) + (\beta - y_1(b)), \\ &= \beta. \end{aligned}$$

Por lo tanto, $y(x)$ es la solución única a nuestro problema con valor en la frontera, naturalmente a condiciones de que $y_2(b) \neq 0$. Si consideramos que $y_2(b) = 0$, tenemos una contradicción con las hipótesis del corolario (4) ya que por teorema (36) el problema homogéneo tendría solución distinta de la trivial, y ya se probó que eso no puede pasar.

El método del disparo para las ecuaciones lineales se basa en la sustitución del problema lineal con valor de frontera para dos problemas de valor inicial de la ecuación (105). Existen métodos para aproximar las soluciones $y_1(x)$ e $y_2(x)$, una vez que contamos con estas aproximaciones, la solución del problema de valor de frontera se aproxima por medio de la ecuación (104).

Desde el punto de vista gráfico, el método tiene el aspecto que se observa en la siguiente figura.



En el siguiente algoritmo se usa el método de Runge-Kutta de cuarto orden para obtener las aproximaciones de $y_1(x)$ e $y_2(x)$.

El algoritmo tiene la característica adicional de obtener aproximaciones para la derivada de la solución del problema con valor de frontera y también la solución del problema en sí mismo.

Algoritmo 1 : Método del Disparo Lineal.

Para aproximar la solución del problema de valor en la frontera

$$-y'' + p(x)y' + q(x)y + r(x), \text{ para } a \leq x \leq b, \\ \text{con las condiciones en la frontera } y(a) = \alpha, \text{ y } y(b) = \beta.$$

El sistema de ecuaciones (105) se escriben y se resuelven como un sistema de primer orden.

ENTRADA: extremos del intervalo a, b , las funciones continuas $p(x), q(x)$ y $r(x)$, condiciones en la frontera α, β y el número de subintervalos n .

SALIDA: aproximaciones W_1 a y ; W_2 a y' en los nodos de la partición del intervalo $[a, b]$.

Paso 1. Tome $h = (b - a) / n$;

$$\begin{aligned}
u_{1,0} &= \alpha; \\
u_{2,0} &= 0; \\
v_{1,0} &= 0; \\
v_{2,0} &= 1;
\end{aligned}$$

Paso 2. Para $i = 0, 1, 2, \dots, n - 1$ hacer los pasos 3 y 4.

(El método de Runge-Kutta para sistema es usado en los pasos 3 y 4.)

Paso 3. Tome $x = a + (i)h$;

Paso 4. Tome $k_{1,1} = hu_{2,i}$;

$$\begin{aligned}
k_{1,2} &= h[p(x)u_{2,i} + q(x)u_{1,i} + r(x)]; \\
k_{2,1} &= h[u_{2,i} + \frac{1}{2}k_{1,2}]; \\
k_{2,2} &= h[p(x + \frac{h}{2})(u_{2,i} + \frac{1}{2}k_{1,2}) + q(x + \frac{h}{2})(u_{1,i} + \frac{1}{2}k_{1,1}) + r(x + \frac{h}{2})]; \\
k_{3,1} &= h[u_{2,i} + \frac{1}{2}k_{2,2}]; \\
k_{3,2} &= h[p(x + \frac{h}{2})(u_{2,i} + \frac{1}{2}k_{2,2}) + q(x + \frac{h}{2})(u_{1,i} + \frac{1}{2}k_{2,1}) + r(x + \frac{h}{2})]; \\
k_{4,1} &= h[u_{2,i} + k_{3,2}]; \\
k_{4,2} &= h[p(x + h)(u_{2,i} + k_{3,2}) + q(x + h)(u_{1,i} + k_{3,1}) + r(x + h)]; \\
u_{1,i+1} &= u_{1,i} + \frac{1}{6}[k_{1,1} + 2k_{2,1} + 2k_{3,1} + k_{4,1}]; \\
u_{2,i+1} &= u_{2,i} + \frac{1}{6}[k_{1,2} + 2k_{2,2} + 2k_{3,2} + k_{4,2}]; \\
k'_{1,1} &= hv_{2,i}; \\
k'_{1,2} &= h[p(x)v_{2,i} + q(x)v_{1,i}]; \\
k'_{2,1} &= h[v_{2,i} + \frac{1}{2}k'_{1,2}]; \\
k'_{2,2} &= h[p(x + \frac{h}{2})(v_{2,i} + \frac{1}{2}k'_{1,2}) + q(x + \frac{h}{2})(v_{1,i} + \frac{1}{2}k'_{1,1})]; \\
k'_{3,1} &= h[v_{2,i} + \frac{1}{2}k'_{2,2}]; \\
k'_{3,2} &= h[p(x + \frac{h}{2})(v_{2,i} + \frac{1}{2}k'_{2,2}) + q(x + \frac{h}{2})(v_{1,i} + \frac{1}{2}k'_{2,1})]; \\
k'_{4,1} &= h[v_{2,i} + k'_{3,2}]; \\
k'_{4,2} &= h[p(x + h)(v_{2,i} + k'_{3,2}) + q(x + h)(v_{1,i} + k'_{3,1})]; \\
v_{1,i+1} &= v_{1,i} + \frac{1}{6}[k'_{1,1} + 2k'_{2,1} + 2k'_{3,1} + k'_{4,1}]; \\
v_{2,i+1} &= v_{2,i} + \frac{1}{6}[k'_{1,2} + 2k'_{2,2} + 2k'_{3,2} + k'_{4,2}];
\end{aligned}$$

Paso 5. Tome $w_{1,0} = \alpha$;

$$\begin{aligned}
w_{2,0} &= (\beta - u_{1,n}) / (v_{1,n}); \\
\text{SALIDA } &(a, w_{1,0}, w_{2,0}).
\end{aligned}$$

Paso 6 Para $i = 1, 2, \dots, n$

$$\begin{aligned}
\text{Tome } W_1 &= u_{1,i} + w_{2,0}v_{1,i}; \\
W_2 &= u_{2,i} + w_{2,0}v_{2,i}; \\
x &= a + ih; \\
\text{SALIDA } &(x, W_1, W_2).
\end{aligned}$$

Paso 7. Parar. (Proceso completado.)

□

Ejemplo 5 : *El problema de valor de frontera*

$$y'' = -\frac{2}{x}y' + \frac{2}{x^2}y + \frac{\sin(\ln(x))}{x^2}, \quad 1 \leq x \leq 2, \quad y(1) = 1, \quad y(2) = 2,$$

tiene solución exacta

$$y = c_1x + \frac{c_2}{x^2} - \frac{3}{10}\sin(\ln(x)) - \frac{1}{10}\cos(\ln(x)),$$

donde

$$c_2 = \frac{1}{10}[8 - 12\sin(\ln(2)) - 4\cos(\ln(2))] \approx -0.03920701320,$$

y

$$c_1 = \frac{11}{10} - c_2 \approx 1.1392070132.$$

Si queremos aplicar el algoritmo (1) a este problema es necesario aproximar las soluciones de los problemas con valor inicial

$$y_1'' = -\frac{2}{x}y_1' + \frac{2}{x^2}y_1 + \frac{\sin(\ln(x))}{x^2}, \quad 1 \leq x \leq 2, \quad y_1(1) = 1, \quad y_1'(1) = 0,$$

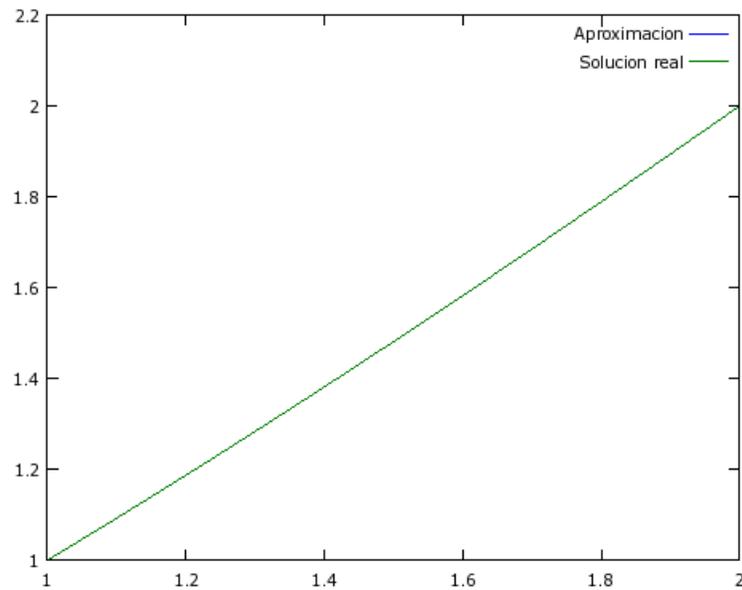
y

$$y_2'' = -\frac{2}{x}y_2' + \frac{2}{x^2}y_2, \quad 1 \leq x \leq 2, \quad y_2(1) = 0, \quad y_2'(1) = 1.$$

En la siguiente tabla se incluyen los resultados de los cálculos cuando se emplea el algoritmo (1), con un tamaño de paso $h = 0.1$. W_1 aproxima la solución del problema de valor de frontera dado, $f(x)$ es el valor real de función.

X_i	Aprox (X_i)	$Y(X_i)$	error
1.00000	1.00000	1.00000	0.00000
1.10000	1.09263	1.09263	0.00000
1.20000	1.18708	1.18708	0.00000
1.30000	1.28338	1.28338	0.00000
1.40000	1.38145	1.38145	0.00000
1.50000	1.48116	1.48116	0.00000
1.60000	1.58239	1.58239	0.00000
1.70000	1.68501	1.68501	0.00000
1.80000	1.78890	1.78890	0.00000
1.90000	1.89393	1.89393	0.00000
2.00000	2.00000	2.00000	0.00000

En la siguiente imagen se muestra la gráfica de la solución y la aproximación a la ecuación diferencial



9.2. Método del disparo para problemas no lineales.

El método del disparo para problema no lineal con valor de frontera de segundo orden

$$y'' = f(x, y, y'), \quad a \leq x \leq b, \quad y(a) = \alpha, \quad y(b) = \beta, \quad (107)$$

se parece al caso lineal, excepto la solución del problema no lineal no puede expresarse como una combinación lineal de las soluciones a los problemas de dos valores iniciales. Necesitamos en cambio, utilizar las soluciones de una sucesión de problemas de valor inicial de forma que contenga un parámetro t , para aproximar la solución del problema de valor de frontera.

$$y'' = f(x, y, y'), \quad a \leq x \leq b, \quad y(a) = \alpha, \quad y'(a) = t. \quad (108)$$

Esto lo hacemos escogiendo los parámetros $t = t_k$, de tal forma que

$$\lim_{k \rightarrow \infty} y(b, t_k) = y(b) = \beta,$$

donde $y(x, t_k)$ denota la solución del problema con valor inicial de la ecuación (108) con $t = t_k$ y $y(x)$ denota la solución al problema con valor en la frontera de la ecuación (107).

Esta técnica se conoce con el nombre **del disparo**, por la analogía con el procedimiento de dispararles a los objetos situados en un blanco fijo, gráficamente tenemos la siguiente figura.

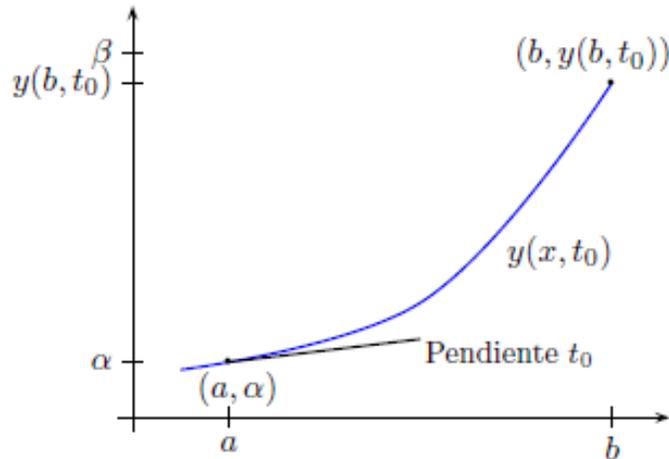
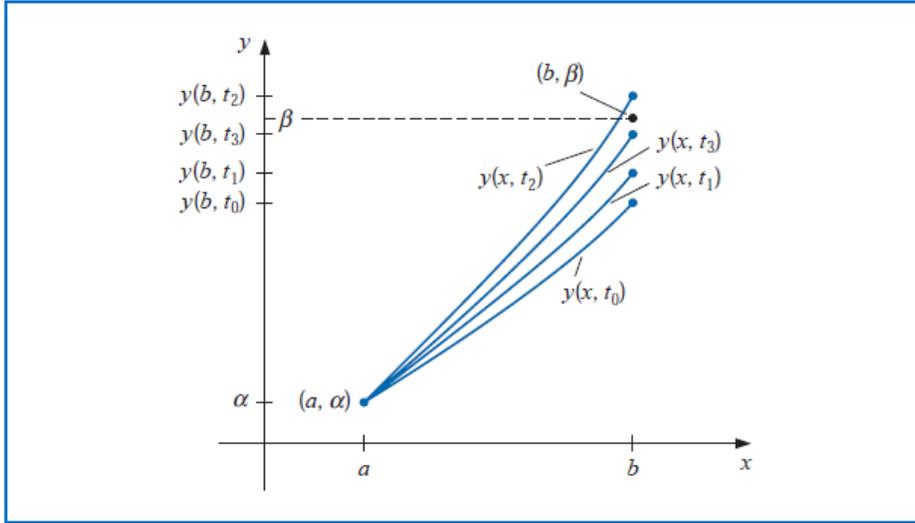


Figura 2.2: Método del Disparo para el caso no lineal.

Comenzamos con un parámetro t_0 que determina la elevación inicial a la cual se le dispara al objetivo desde el punto (a, α) y a lo largo de la curva descrita por la solución al problema de valor inicial:

$$y'' = f(x, y, y'), \quad a \leq x \leq b, \quad y(a) = \alpha, \quad y'(a) = t_0.$$

Si $y(b, t_0)$ no está suficientemente cerca de β , corregimos la aproximación seleccionando las elevaciones t_1, t_2 , y así sucesivamente hasta que $y(b, t_k)$ este bastante cerca de **acertar el blanco** β , como se muestra en la siguiente gráfica:



Para determinar los parámetros t_k , supongamos que un problema de valor de frontera de la ecuación (107) satisface las condiciones del teorema (35). Si $y(x, t)$ denota la solución del problema de valor inicial de la ecuación (108), el problema consistirá en determinar t tal que

$$y(b, t) - \beta = 0. \quad (109)$$

Esta es una ecuación no lineal en t de los cuales hay diferentes métodos para calcular dicha aproximación.

Si queremos emplear el método de la secante para resolver el problema, necesitamos elegir las aproximaciones iniciales t_0, t_1 y luego generar los términos restantes de la sucesión mediante

$$t_k = t_{k-1} - \frac{(y(b, t_{k-1}) - \beta)(t_{k-1} - t_{k-2})}{y(b, t_{k-1}) - y(b, t_{k-2})}, \text{ para } i = 1, 2, \dots$$

Para generar la sucesión $\{t_k\}$ con el método de Newton, que es más poderoso, solo necesitamos una aproximación inicial t_0 . Sin embargo la iteración tiene la forma

$$t_k = t_{k-1} - \frac{y(b, t_{k-1}) - \beta}{\frac{dy}{dt}(b, t_{k-1})}, \quad (110)$$

y requiere conocer $\frac{dy}{dt}(b, t_{k-1})$. Esto es un problema porque no se conoce una representación explícita de $y(b, t_{k-1})$; conocemos solo los valores de $y(b, t_0), y(b, t_1), \dots, y(b, t_{k-1})$.

Ahora supongamos que reescribimos el problema de valor inicial de la ecuación (108) haciendo énfasis en que la solución se basa tanto en x como en el parámetro t :

$$y''(x, t) = f(x, y(x, t), y'(x, t)), \quad a \leq x \leq b, \quad y(a, t) = \alpha, \quad y'(a, t) = t. \quad (111)$$

Hemos conservado la notación prima para indicar la derivada respecto a x , puesto que necesitamos determinar $\frac{dy}{dt}(b, t)$, cuando $t = t_{k-1}$, primero tomamos la derivada parcial de la ecuación (111) respecto a t . Esto significa que:

$$\begin{aligned}\frac{\partial y''}{\partial t}(x, t) &= \frac{\partial f}{\partial t}(x, y(x, t), y'(x, t)) \\ &= \frac{\partial f}{\partial x}(x, y(x, t), y'(x, t)) \frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}(x, y(x, t), y'(x, t)) \frac{\partial y}{\partial t}(x, t) + \\ &\quad \frac{\partial f}{\partial y'}(x, y(x, t), y'(x, t)) \frac{\partial y'}{\partial t}(x, t)\end{aligned}$$

Dado que x y t , son independientes, $\frac{\partial x}{\partial t} = 0$, sustituyendo en la ecuación anterior tenemos:

$$\frac{\partial y''}{\partial t}(x, t) = \frac{\partial f}{\partial y}(x, y(x, t), y'(x, t)) \frac{\partial y}{\partial t}(x, t) + \frac{\partial f}{\partial y'}(x, y(x, t), y'(x, t)) \frac{\partial y'}{\partial t}(x, t), \quad (112)$$

con $a \leq x \leq b$. Las condiciones iniciales dan

$$\frac{\partial y}{\partial t}(a, t) = 0 \quad \text{y} \quad \frac{\partial y'}{\partial t}(a, t) = 1.$$

Si simplificamos la notación usando $z(x, t)$ para denotar $\frac{\partial y}{\partial t}(x, t)$ y si suponemos que el orden de derivación de x y t pueden invertirse, con las condiciones iniciales dadas en la ecuación (112) se convierten en el problema de valor inicial

$$z''(x, t) = \frac{\partial f}{\partial y}(x, y, y')z(x, t) + \frac{\partial f}{\partial y'}(x, y, y')z'(x, t), \quad a \leq x \leq b, \quad z(a, t) = 0, \quad z'(a, t) = 1. \quad (113)$$

Así pues el método de Newton requiere que los dos problemas de valor inicial sean resueltos en cada iteración de las ecuaciones (111) y (113), entonces de la ecuación (110) tenemos,

$$t_k = t_{k-1} - \frac{y(b, t_{k-1}) - \beta}{z(b, t_{k-1})}. \quad (114)$$

Desde luego estos problemas de valor inicial no pueden resolverse exactamente; en el siguiente algoritmo se aproximan por el método de Runge-Kutta de cuarto orden para aproximar las dos soluciones que requiere el método de Newton.

Algoritmo 2 : Método del Disparo no Lineal con el método de Newton.

Para aproximar la solución del problema no lineal con valor en la frontera

$$y'' = f(x, y, y'), \quad a \leq x \leq b, \quad y(a) = \alpha, \quad y(b) = \beta;$$

Las ecuaciones (112) y (113) se escriben y se resuelven como sistema de primer orden.

ENTRADA: extremos del intervalo a y b ; condiciones en la frontera α , β ; número de subintervalos $n \geq 2$; tolerancia Tol ; número máximo de iteración M .

SALIDA: aproximaciones $w_{1,i}$ a $y(x_i)$; $w_{2,i}$ a $y'(x_i)$ para $i = 0, 1, 2, \dots, n$ o un mensaje que el número máximo de iteraciones fue excedido.

Paso 1. Tome $h = (b - a)/n$;

$$k = 1;$$

$$TK = (\beta - \alpha)/(b - a).$$

Paso 2. Mientras ($k \leq M$) hacer los Pasos 3-10.

Paso 3. Tome $w_{1,0} = \alpha$;

$$w_{2,0} = TK;$$

$$u_1 = 0;$$

$$u_2 = 1;$$

Paso 4. Para $i = 1, \dots, n$ hacer los Pasos 5-6.

(El método de Runge-Kutta para sistema es usado en los Pasos 5 y 6.)

Paso 5. Tome $x = a + (i - 1)h$;

Paso 6. Tome $k_{1,1} = hw_{2,i-1}$;

$$k_{1,2} = hf(x, w_{1,i-1}, w_{2,i-1});$$

$$k_{2,1} = h(w_{2,i-1} + \frac{1}{2}k_{1,2});$$

$$k_{2,2} = hf(x + \frac{h}{2}, w_{1,i-1} + \frac{1}{2}k_{1,1}, w_{2,i-1} + \frac{1}{2}k_{1,2});$$

$$k_{3,1} = h(w_{2,i-1} + \frac{1}{2}k_{2,2});$$

$$k_{3,2} = hf(x + \frac{h}{2}, w_{1,i-1} + \frac{1}{2}k_{2,1}, w_{2,i-1} + \frac{1}{2}k_{2,2});$$

$$k_{4,1} = h(w_{2,i-1} + k_{3,2});$$

$$k_{4,2} = hf(x + h, w_{1,i-1} + k_{3,1}, w_{2,i-1} + k_{3,2});$$

$$w_{1,i} = w_{1,i-1} + (k_{1,1} + 2k_{2,1} + 2k_{3,1} + k_{4,1})/6;$$

$$w_{2,i} = w_{2,i-1} + (k_{1,2} + 2k_{2,2} + 2k_{3,2} + k_{4,2})/6;$$

$$k'_{1,1} = hu_2;$$

$$k'_{1,2} = h[f_y(x + \frac{h}{2}, w_{1,i-1}, w_{2,i-1})(u_1 + \frac{1}{2}k'_{1,1}) + f'_y(x + \frac{h}{2}, w_{1,i-1}, w_{2,i-1})(u_2 + \frac{1}{2}k'_{1,2})];$$

$$k'_{3,1} = h(u_2 + \frac{1}{2}k'_{2,2});$$

$$k'_{3,2} = h[f_y(x + \frac{h}{2}, w_{1,i-1}, w_{2,i-1})(u_1 + \frac{1}{2}k'_{2,1})$$

$$\begin{aligned}
& + f_{y'}(x + \frac{h}{2}, w_{1,i-1}, w_{2,i-1})(u_2 + \frac{1}{2}k'_{2,2}); \\
k'_{4,1} &= h(u_2 + k'_{3,2}); \\
k'_{4,2} &= h[f_y(x + h, w_{1,i-1}, w_{2,i-1})(u_1 + k'_{3,1}) \\
& + f_{y'}(x + h, w_{1,i-1}, w_{2,i-1})(u_2 + k'_{3,2})] \\
u_1 &= u_1 + \frac{1}{6}[k'_{1,1} + 2k'_{2,1} + 2k'_{3,1} + k'_{4,1}]; \\
u_2 &= u_2 + \frac{1}{6}[k'_{1,2} + 2k'_{2,2} + 2k'_{3,2} + k'_{4,2}];
\end{aligned}$$

Paso 7. Si $|w_{1,n} - \beta| \leq TOL$ entonces hacer los Pasos 8-9.

Paso 8. Para $i = 1, \dots, n$
tome $x = a + ih$;
SALIDA $(x, w_{1,i}, w_{2,i})$.

Paso 9. (El procedimiento es completado.)
PARAR.

Paso 10. Tome $TK = TK - \frac{w_{1,n} - \beta}{u_1}$;
(El metodo de Newton es usado en el calculo de TK)
 $k = k + 1$;

Paso 11. SALIDA ('Numero máximo de iteraciones se ha excedido');
PARAR.

□

Ejemplo 6 Considere el problema con valor de frontera

$$y'' = \frac{1}{8}(32 + 2x^3 - yy'), \quad 1 \leq x \leq 3, \quad y(1) = 17, \quad y(3) = \frac{43}{3},$$

que tiene la solución exacta

$$y(x) = x^2 + \frac{16}{x}.$$

Si queremos aplicar el algoritmo (2), a este problema hay que aproximar los problemas de valor inicial

$$y'' = \frac{1}{8}(32 + 2x^3 - yy'), \quad 1 \leq x \leq 3, \quad y(1) = 17, \quad y'(1) = t_k,$$

y

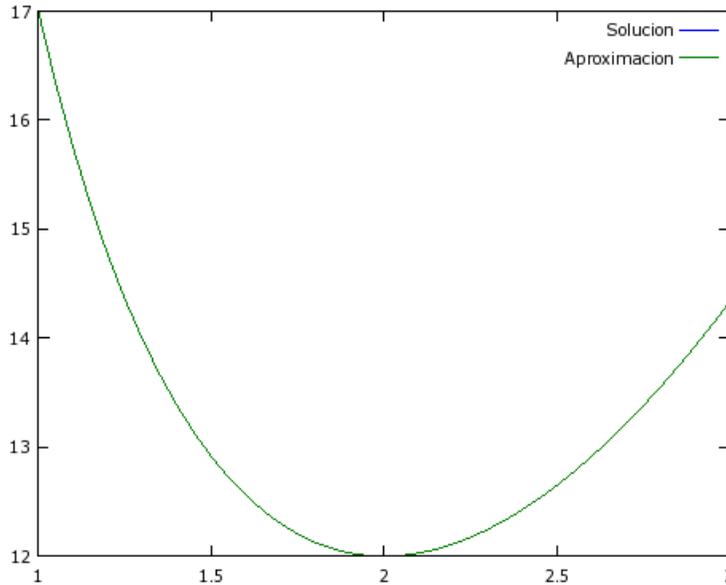
$$z'' = \frac{\partial f}{\partial y}z + \frac{\partial f}{\partial y'}z' = -\frac{1}{8}(y'z + yz'), \quad 1 \leq x \leq 3, \quad z(1) = 0, \quad z'(1) = 1,$$

en cada paso de la iteración.

Aplicando el algoritmo (2) obtenemos la siguiente tabla:

X_i	<i>Aprox</i> $y(X_i)$	<i>Aprox</i> $dy/dx(y(X_i))$	$y(X_i)$	<i>error</i>
1.00000	17.00000	-14.00001	17.00000	0.00000
1.05000	16.34060	-12.41249	16.34060	0.00000
1.10000	15.75546	-11.02315	15.75545	0.00000
1.15000	15.23555	-9.79831	15.23554	0.00000
1.20000	14.77334	-8.71112	14.77333	0.00000
1.25000	14.36250	-7.74001	14.36250	0.00000
1.30000	13.99770	-6.86747	13.99769	0.00000
1.35000	13.67435	-6.07916	13.67435	0.00000
1.40000	13.38857	-5.36328	13.38857	0.00000
1.45000	13.13699	-4.71000	13.13698	0.00000
1.50000	12.91667	-4.11112	12.91667	0.00000
1.55000	12.72508	-3.55974	12.72508	0.00000
1.60000	12.56000	-3.05001	12.56000	0.00000
1.65000	12.41947	-2.57696	12.41947	0.00000
1.70000	12.30177	-2.13634	12.30176	0.00000
1.75000	12.20536	-1.72450	12.20536	0.00000
1.80000	12.12889	-1.33828	12.12889	0.00000
1.85000	12.07115	-0.97495	12.07115	0.00000
1.90000	12.03105	-0.63214	12.03105	0.00000
1.95000	12.00763	-0.30776	12.00763	0.00000
2.00000	12.00000	-0.00000	12.00000	0.00000
2.05000	12.00738	0.29274	12.00738	0.00000
2.10000	12.02905	0.57188	12.02905	0.00000
2.15000	12.06436	0.83867	12.06436	0.00000
2.20000	12.11273	1.09421	12.11273	0.00000
2.25000	12.17361	1.33950	12.17361	0.00000
2.30000	12.24652	1.57542	12.24652	0.00000
2.35000	12.33101	1.80276	12.33101	0.00000
2.40000	12.42667	2.02222	12.42667	0.00000
2.45000	12.53311	2.23444	12.53311	0.00000
2.50000	12.65000	2.44000	12.65000	0.00000
2.55000	12.77701	2.63941	12.77701	0.00000
2.60000	12.91384	2.83313	12.91385	0.00000
2.65000	13.06023	3.02161	13.06024	0.00000
2.70000	13.21592	3.20521	13.21593	0.00000
2.75000	13.38068	3.38430	13.38068	0.00000
2.80000	13.55428	3.55918	13.55429	0.00000
2.85000	13.73653	3.73016	13.73654	0.00000
2.90000	13.92724	3.89750	13.92724	0.00000
2.95000	14.12623	4.06145	14.12623	0.00000
3.00000	14.33333	4.22222	14.33333	0.00000

En la siguiente imagen se muestra las gráficas de la solución real y de la aproximación:



9.3. Métodos de diferencias finitas para los problemas lineales.

Aunque el método del disparo puede emplearse en los problemas lineales y no lineales de valor de frontera, a menudo presentan problemas de inestabilidad. Los métodos expuestos a continuación tienen mejores características de estabilidad.

Los métodos para resolución de problema de valor de frontera que contiene diferencias finitas reemplazan las derivadas en la ecuación diferencial mediante una aproximación de cocientes de diferencia adecuada. Se selecciona el cociente de diferencia para mantener un orden especificado de error de truncamiento. Pero por la inestabilidad de las aproximaciones de diferencias finitas a las derivadas, no podemos escoger un parámetro h demasiado pequeño. El método de diferencias finita para el problema de valor de frontera de segundo orden

$$y'' = p(x)y' + q(x)y + r(x), \quad a \leq x \leq b, y(a) = \alpha, y(b) = \beta. \quad (115)$$

requiere utilizar las aproximaciones del cociente de diferencias para aproximar tanto a y' como a y'' . Primero seleccionamos un $N > 0$ y dividimos el intervalo $[a, b]$ en $N + 1$ subintervalos iguales cuyo extremos son los puntos de maya: $x_i = a + ih$ para $i = 0, 1, 2, \dots, N+1$, donde $h = \frac{b-a}{N+1}$. En los puntos de red del interior, x_i para $i = 1, 2, \dots, N$, la ecuación diferencial a aproximar es:

$$y'' = p(x_i)y' + q(x_i)y + r(x_i). \quad (116)$$

Al desarrollar y en el tercer polinomio de Taylor alrededor de x_i , evaluando en x_{i+1} y x_{i-1}

tenemos, suponiendo que $y \in C^4[x_{i-1}, x_i]$,

$$y(x_{i+1}) = y(x_i + h) = y(x_i) + hy'(x_i) + \frac{h^2}{2}y''(x_i) + \frac{h^3}{6}y'''(x_i) + \frac{h^4}{24}y^{(4)}(\xi_i^+),$$

para algún ξ_i^+ en (x_i, x_{i+1}) , y

$$y(x_{i-1}) = y(x_i - h) = y(x_i) - hy'(x_i) + \frac{h^2}{2}y''(x_i) - \frac{h^3}{6}y'''(x_i) + \frac{h^4}{24}y^{(4)}(\xi_i^-),$$

para algún ξ_i^- en (x_{i-1}, x_i) , sumando estas dos ultimas ecuaciones tenemos

$$y(x_{i+1}) + y(x_{i-1}) = 2y(x_i) + h^2y''(x_i) + \frac{h^4}{24}[y^{(4)}(\xi_i^+) + y^{(4)}(\xi_i^-)],$$

y al despejar $y''(x_i)$ se obtiene

$$y''(x_i) = \frac{1}{h^2}[y(x_{i+1}) - 2y(x_i) + y(x_{i-1})] - \frac{h^2}{24}[y^{(4)}(\xi_i^+) + y^{(4)}(\xi_i^-)],$$

Podemos aplicar el teorema de valor intermedio para simplificar a un mas esta expresión y transformarla en:

$$y''(x_i) = \frac{1}{h^2}[y(x_{i+1}) - 2y(x_i) + y(x_{i-1})] - \frac{h^2}{12}y^{(4)}(\xi_i), \quad (117)$$

para algún $\xi_i \in (x_{i-1}, x_{i+1})$. A esta ecuación le llamamos **fórmula de las diferencias centradas** para $y''(x_i)$.

De manera análoga se obtiene una formula de este tipo para $y'(x_i)$, que da por resultado

$$y'(x_i) = \frac{1}{2h}[y(x_{i+1}) - y(x_{i-1})] - \frac{h^2}{6}y''(\eta_i), \quad (118)$$

para algún $\eta_i \in (x_{i-1}, x_{i+1})$.

La utilización de las formulas de las diferencias centradas en la ecuación (116) generan la ecuación

$$\begin{aligned} \frac{y(x_{i+1}) - 2y(x_i) + y(x_{i-1}))}{h^2} &= p(x_i) \left[\frac{y(x_{i+1}) - y(x_{i-1}))}{2h} \right] + q(x_i)y(x_i) + \\ &r(x_i) - \frac{h^2}{2}[2p(x_i)y'''(\eta_i) - y^{(4)}(\xi_i)]. \end{aligned}$$

Un método de diferencias finitas con error de truncamiento de orden $\mathcal{O}(h^2)$, se obtiene empleando esta ecuación junto con las condiciones de frontera $y(a) = \alpha$ y $y(b) = \beta$ para definir

$$w_0 = \alpha, \quad w_{N+1} = \beta,$$

y

$$\left(\frac{-w_{i+1} + 2w_i - w_{i-1}}{h^2}\right) + p(x_i) \left(\frac{w_{i+1} - w_{i-1}}{2h}\right) + q(x_i)w_i = -r(x_i), \quad (119)$$

para $i = 1, 2, \dots, N$.

En forma que consideramos, la ecuación (119) y se reescribe como

$$-\left(1 + \frac{h}{2}p(x_i)\right)w_{i-1} + (2 + h^2q(x_i))w_i - \left(1 + \frac{h}{2}p(x_i)\right)w_{i+1} = -h^2r(x_i),$$

y el sistema de ecuaciones resultantes se expresa en forma de matriz tridiagonal de $N \times N$

$$Aw = b \quad (120)$$

$$A = \begin{bmatrix} 2 + h^2q(x_1) & -1 + \frac{h}{2}p(x_1) & 0 & \dots & \dots & 0 & 0 \\ -1 - \frac{h}{2}p(x_2) & 2 + h^2q(x_2) & -1 + \frac{h}{2}p(x_2) & \dots & \dots & 0 & 0 \\ 0 & -1 - \frac{h}{2}p(x_3) & 2 + h^2q(x_3) & \dots & \dots & 0 & 0 \\ 0 & 0 & -1 - \frac{h}{2}p(x_4) & \dots & \dots & 0 & 0 \\ \vdots & \ddots & \ddots & \dots & \dots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \dots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \dots & 2 + h^2q(x_{N-1}) & -1 + \frac{h}{2}p(x_{N-1}) \\ 0 & 0 & 0 & \dots & \dots & -1 - \frac{h}{2}p(x_N) & 2 + h^2q(x_N) \end{bmatrix} \quad (121)$$

$$w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_{N-1} \\ w_N \end{bmatrix}$$

y

$$b = \begin{bmatrix} -h^2 r(x_1) + \left(1 + \frac{h}{2} p(x_1)\right) w_0 \\ -h^2 r(x_2) \\ \vdots \\ -h^2 r(x_{N-1}) \\ -h^2 r(x_N) + \left(1 + \frac{h}{2} p(x_1)\right) w_{N+1} \end{bmatrix}$$

El siguiente teorema establece las condiciones bajo las cuales el sistema (120) tiene solución única.

Teorema 38 : Suponga que p , q y r son continuas en $[a, b]$. Si $q(x) \geq 0$ en $[a, b]$, entonces el sistema tridiagonal (120) tiene una solución única siempre y cuando $h < \frac{2}{L}$, donde $L = \max_{a \leq x \leq b} |p(x)|$.

Prueba.

Por hipótesis tenemos que $L = \max_{a \leq x \leq b} |p(x)|$ y $h < \frac{2}{L}$ entonces $\forall x_i \in [a, b]$ tenemos que

$$\begin{aligned} -L &\leq p(x_i) \leq L & , \\ -\frac{h}{2}L &\leq \frac{h}{2}p(x_i) \leq \frac{h}{2}L & , \\ -1 &< -\frac{h}{2}L \leq \frac{h}{2}p(x_i) \leq \frac{h}{2}L < 1 & , \\ -1 &< \frac{h}{2}p(x_i) < 1 & , \end{aligned} \tag{122}$$

de la matriz A las componentes $a_{i,i-1} = -1 - \frac{h}{2}p(x_i) \neq 0$ y $a_{i,i+1} = -1 + \frac{h}{2}p(x_i) \neq 0$ por resultado para $i = 2, 3, \dots, n-1$. Por lo tanto:

$$a_{i,i-1}, a_{i,i+1} \neq 0 \text{ para } i = 2, 3, \dots, n-1. \tag{123}$$

$|a_{1,1}| = |2 + h^2 q(x_1)|$ como $q(x) \geq 0$ entonces $|a_{1,1}| \geq 2$ y $|a_{1,2}| = |-1 + \frac{h}{2}p(x_1)| < 2$ por (122) se tiene que $|a_{1,2}| < 2$.

Por lo tanto tenemos

$$|a_{1,1}| > |a_{1,2}|. \tag{124}$$

Ahora $|a_{i,i}| = |2 + h^2q(x_i)| \geq 2$ ya que $q(x) \geq 0$, además

$|a_{i,i-1}| = |-1 - \frac{h}{2}p(x_i)|$, si $\frac{h}{2}p(x_i) \geq 0$ tenemos que $|a_{i,i-1}| \leq 1$

Y $|a_{i,i+1}| = |-1 + \frac{h}{2}p(x_i)| \leq 1$ así que

$$|a_{i,i}| \geq |a_{i,i-1}| + |a_{i,i+1}|. \quad (125)$$

Si $\frac{h}{2}p(x_i) < 0$ tenemos que $|a_{i,i-1}| = 1 + \frac{h}{2}p(x_i)$ y $|a_{i,i+1}| = |-1 + \frac{h}{2}p(x_i)| = 1 - \frac{h}{2}p(x_i)$ entonces $|a_{i,i-1}| + |a_{i,i+1}| = 1 + \frac{h}{2}p(x_i) + 1 - \frac{h}{2}p(x_i) = 2$ y como $|a_{i,i}| = |2 + h^2q(x_i)| \geq 2$ tenemos que

$$|a_{i,i}| \geq |a_{i,i-1}| + |a_{i,i+1}|. \quad (126)$$

De las ecuaciones (125) y (126) tenemos que si $\frac{h}{2}p(x_i) < 0$ o $\frac{h}{2}p(x_i) \geq 0$ se cumple que:

$$|a_{i,i}| \geq |a_{i,i-1}| + |a_{i,i+1}| \quad \text{para } i = 2, 3, \dots, n-1. \quad (127)$$

Luego tenemos por un razonamiento similar

$$|a_{n,n}| > |a_{n,n-1}|. \quad (128)$$

Por las ecuaciones (123), (124), (127) y (128) tenemos que se cumplen las hipótesis del teorema (7), teniendo que la matriz de la ecuación (121) es invertible y por lo tanto tiene solución única.

□

Conviene señalar que las hipótesis del teorema (38), garantizan una solución única al problema de valor de frontera (115), pero no que $y \in C^4[a, b]$. Para asegurarnos que el error de truncamiento tiene el orden $\mathcal{O}(h^2)$, debemos establecer que $y^{(4)}$ es continua en $[a, b]$.

Algoritmo 3 Método Diferencias Finitas Lineal.

Aproxima la solución del problema de valor de frontera lineal

$$y'' = p(x)y' + q(x)y + r(x), \quad \text{para } a \leq x \leq b,$$

con las condiciones en la frontera $y(a) = \alpha$, $y(b) = \beta$.

ENTRADA: extremos del intervalo a, b ; condiciones en la frontera α, β ; un entero $N \geq 2$.

SALIDA: aproximación w_i a $y(x_i)$ para cada $i = 0, 1, \dots, N+1$.

Paso 1 Tome $h = (b - a)/(N + 1)$;
 $x = a + h$;
 $a_1 = 2 + h^2q(x)$;
 $b_1 = -1 + (h/2)p(x)$;
 $d_1 = -h^2r(x) + (1 + (h/2)p(x))\alpha$;

Paso 2 Para $i = 2, \dots, N - 1$
tome $x = a + ih$;
 $a_i = 2 + h^2q(x)$;
 $b_i = -1 + (h/2)p(x)$;
 $c_i = -1 - (h/2)p(x)$;
 $d_i = -h^2r(x)$;

Paso 3 Tome $x = b - h$;
 $a_N = 2 + h^2q(x)$;
 $c_N = -1 - (h/2)p(x)$;
 $d_N = -h^2r(x) + (1 - (h/2)p(x))\beta$;

Paso 4 Tome $l_1 = a_1$;
 $u_1 = b_1/a_1$;
 $z_1 = d_1/l_1$;

Paso 5 Para $i = 2, \dots, N - 1$
tome $l_i = a_i - c_i u_{i-1}$;
 $u_i = b_i/l_i$;
 $z_i = (d_i - c_i z_{i-1})/l_i$;

Paso 6 Tome $L_N = a_N - c_N u_{N-1}$;
 $z_N = (d_N - c_N z_{N-1})/L_N$;

Paso 7 Tome $w_0 = \alpha$;
 $w_{N+1} = \beta$;
 $w_N = z_N$;

Paso 8 Para $i = N - 1, \dots, 1$
tome $w_i = z_i - u_i w_{i+1}$;

Paso 9 Para $i = 0, \dots, N + 1$
 tome $x = a + ih$;
 SALIDA (x, w_i)

Paso 10 Procedimiento completado.

□

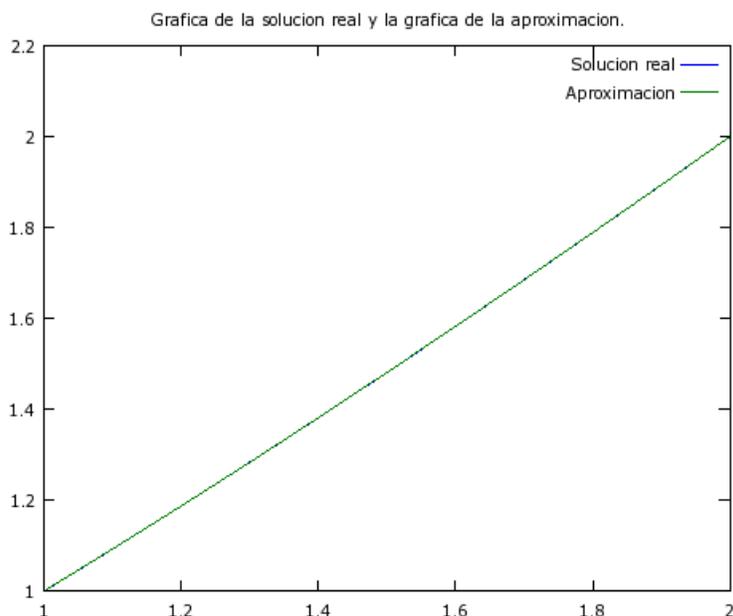
Ejemplo 7 Utilizaremos el algoritmo (3) para aproximar la solución al problema lineal con valor de frontera

$$y_1'' = -\frac{-2}{x}y_1' + \frac{2}{x^2}y_1 + \frac{\sin(\ln(x))}{x^2}, \quad 1 \leq x \leq 2, \quad y(1) = 1, \quad y(2) = 2$$

que también se aproxima mediante el método del disparo. En este ejemplo usaremos $N = 9$, de modo que $h = 0.1$, aplicando el algoritmo (3) obtenemos los siguientes resultados.

X_i	Aprox (X_i)	$Y(X_i)$	error
1.00000	1.00000	1.00000	0.00000
1.10000	1.09260	1.09263	0.00737
1.20000	1.18704	1.18708	0.01292
1.30000	1.28334	1.28338	0.01662
1.40000	1.38140	1.38145	0.01855
1.50000	1.48112	1.48116	0.01884
1.60000	1.58236	1.58239	0.01761
1.70000	1.68499	1.68501	0.01499
1.80000	1.78888	1.78890	0.01110
1.90000	1.89392	1.89393	0.00607
2.00000	2.00000	2.00000	0.00000

En la siguiente imagen se muestran las gráficas de la solución real y la aproximación a la ecuación diferencial



9.4. Métodos de diferencias finitas para los problemas no lineales.

Para el caso del problema no lineal general con valor en la frontera

$$y'' = f(x, y, y'), \quad a \leq x \leq b, \quad y(a) = \alpha, \quad y(b) = \beta$$

el método de diferencias se parece al que se aplicó a los problemas lineales. Sin embargo aquí el sistema de ecuaciones no será lineal, y por lo mismo, se requiere un proceso iterativo para resolverlo.

Para el desarrollo del procedimiento supondremos que f satisface las siguientes condiciones:

- * f y las derivadas parciales f_y y $f_{y'}$ son continuas en

$$D = (x, y, y') \mid a \leq x \leq b, \quad -\infty < y < \infty, \quad -\infty < y' < \infty.$$

- * $f_y(x, y, y') \geq \delta$ en D , para algún $\delta > 0$.

- * Existen las constantes K y L , con

$$K = \max_{(x,y,y') \in D} |f_y(x, y, y')| \quad \text{y} \quad L = \max_{(x,y,y') \in D} |f_{y'}(x, y, y')|.$$

Esto garantiza que, conforme con el teorema (35), exista una solución única. Al igual que en el caso de la ecuación lineal, dividamos $[a, b]$ en $(N + 1)$ subintervalos iguales cuyos extremos se encuentran en $x_i = a + ih$, para $i = 0, 1, 2, \dots, N + 1$. Supongamos que la

solución exacta tiene una cuarta derivada acotada nos permite reemplazar $y''(x_i)$ y $y'(x_i)$ en cada una de las ecuaciones

$$y''(X_i) = f(x_i, y(X_i), y'(X_i)),$$

por la formula adecuada de diferencias centradas que se incluyo en las ecuaciones (117) y (118). Esto nos da para todo $i = 0, 1, 2, \dots, N$

$$\frac{y(x_{i+1}) - 2y(x_i) + y(x_{i-1}))}{h^2} = f\left(x_i, y(x_i), \frac{y(x_{i+1}) - y(x_{i-1}))}{2h} - \frac{h^2}{6}y'''(\eta_i)\right) + \frac{h^2}{12}y^{(4)}(\xi_i),$$

para algún ξ_i y η_i en el intervalo (x_{i-1}, x_{i+1}) .

Como en el caso de la ecuación lineal, los resultados del método de diferencias se emplea cuando se eliminan los términos de error y las condiciones de frontera:

$$w_0 = \alpha, \quad w_{N+1} = \beta,$$

y

$$\frac{w_{i+1} - 2w_i + w_{i-1}}{h^2} + f\left(x_i, w_i, \frac{w_{i+1} - w_{i-1}}{2h}\right) = 0,$$

para todo $i = 1, 2, 3, \dots, N$.

El sistema no lineal de $N \times N$ obtenido con este método es:

$$\begin{aligned} 2w_1 - w_2 + h^2 f\left(x_1, w_1, \frac{w_2 - \alpha}{2h}\right) - \alpha &= 0, \\ -w_1 + 2w_2 - w_3 + h^2 f\left(x_2, w_2, \frac{w_3 - w_2}{2h}\right) &= 0, \\ &\vdots \\ -w_{N-2} + 2w_{N-1} - w_N + h^2 f\left(x_{N-1}, w_{N-1}, \frac{w_N - w_{N-2}}{2h}\right) &= 0, \\ -w_{N-1} + 2w_N + h^2 f\left(x_N, w_N, \frac{\beta - w_{N-1}}{2h}\right) - \beta &= 0, \end{aligned} \tag{129}$$

tiene una solución única siempre y cuando $h < \frac{2}{L}$.

Aplicando el método de Newton para sistemas no lineales, para aproximar la solución de este sistema. Se genera una sucesión de iteraciones $\{(w_1^{(k)}, w_2^{(k)}, w_3^{(k)}, \dots, w_N^{(k)})^t\}$, que converge al sistema de ecuaciones (129), a condiciones que la aproximación inicial $(w_1^{(0)}, w_2^{(0)}, w_3^{(0)}, \dots, w_N^{(0)})^t$ se acerque lo suficiente a la solución $(w_1, w_2, w_3, \dots, w_N)^t$ y

que la matriz jacobiana con el ij -ésimo elemento,

$$J(w_1, w_2, w_3, \dots, w_N)_{ij} = \begin{cases} -1 + \frac{h}{2} f_{y'} \left(x_i, w_i, \frac{w_{i+1} - w_{i-1}}{2h} \right), & \text{para } i = j - 1 \text{ y } j = 2, 3, \dots, N, \\ 2 + h^2 f_y \left(x_i, w_i, \frac{w_{i+1} - w_{i-1}}{2h} \right), & \text{para } i = j \text{ y } j = 1, 2, \dots, N, \\ -1 - \frac{h}{2} f_{y'} \left(x_i, w_i, \frac{w_{i+1} - w_{i-1}}{2h} \right), & \text{para } i = j + 1 \text{ y } j = 1, 2, \dots, N - 1, \end{cases}$$

donde $w_0 = \alpha$ y $w_{N+1} = \beta$.

El método de Newton para los sistemas no lineales requieren que en cada iteración del sistema lineal de $N \times N$,

$$\begin{aligned} & J(w_1, w_2, w_3, \dots, w_N)(v_1, v_2, v_3, \dots, v_N)^t \\ &= -\left(2w_1 - w_2 + h^2 f \left(x_1, w_1, \frac{w_2 - \alpha}{2h} \right) - \alpha, \right. \\ & \quad \left. -w_1 + 2w_2 - w_3 + h^2 f \left(x_2, w_2, \frac{w_3 - w_1}{2h} \right), \dots, \right. \\ & \quad \left. -w_{N-2} + 2w_{N-1} - w_N + h^2 f \left(x_{N-1}, w_{N-1}, \frac{w_N - w_{N-2}}{2h} \right), \right. \\ & \quad \left. -w_{N-1} + 2w_N + h^2 f \left(x_N, w_N, \frac{\beta - w_{N-1}}{2h} \right) - \beta \right)^t, \end{aligned}$$

se despejen $v_1, v_2, v_3, \dots, v_N$, porque

$$w_i^{(k)} = w_i^{(k-1)} + v_i, \quad \text{para cada } i = 1, 2, \dots, N.$$

Puesto que J es tridiagonal, esto no representa ningún problema tan difícil como podría parecer a primera vista. Podemos aplicar el algoritmo de la factorización de Crout para los sistemas tridiagonales.

Algoritmo 4 Método Diferencias Finitas No Lineal.

Aproxima la solución del problema de valor de frontera no lineal

$$\begin{aligned} & y'' = f(x, y, y'), \quad \text{para } a \leq x \leq b, \\ & \text{con las condiciones en la frontera } y(a) = \alpha, \text{ y } y(b) = \beta. \end{aligned}$$

ENTRADA: extremos del intervalo a, b ; condiciones de frontera α, β ; un entero $N \geq 2$; tolerancia TOL ; número máximo de iteraciones M .

SALIDA: aproximación w_i a $y(x_i)$ para cada $i = 0, 1, \dots, N + 1$ o un mensaje que el número máximo de iteraciones fue excedido.

Paso 1 Tome $h = (b - a)/(N + 1)$;

$$w_0 = \alpha;$$

$$w_{N+1} = \beta;$$

Paso 2 Para $i = 1, \dots, N$

$$w_i = \alpha + i \left(\frac{\beta - \alpha}{b - a} \right) h;$$

Paso 3 Tome $k = 1$;

Paso 4 Mientras $k \leq M$ (Hacer los Pasos 5-16)

Paso 5

$$\text{Tome } x = a + h;$$

$$t = (w_2 - \alpha)/(2h);$$

$$a_1 = 2 + h^2 f_y(x, w_1, t);$$

$$b_1 = -1 + (h/2) f_{y'}(x, w_1, t);$$

$$d_1 = -(2w_1 - w_2 - \alpha + h^2 f(x, w_1, t));$$

Paso 6 Para $i = 2, \dots, N - 1$

$$\text{tome } x = a + ih;$$

$$t = (w_{i+1} - w_{i-1})/(2h);$$

$$a_i = 2 + h^2 f_y(x, w_i, t);$$

$$b_i = -1 + (h/2) f_{y'}(x, w_i, t);$$

$$c_i = -1 - (h/2) f_{y'}(x, w_i, t);$$

$$d_i = -(2w_i - w_{i+1} - w_{i-1} + h^2 f(x, w_i, t));$$

Paso 7 Tome $x = b - h$;

$$t = (\beta - w_{N-1})/(2h);$$

$$a_N = 2 + h^2 f_y(x, w_N, t);$$

$$c_N = -1 - (h/2) f_{y'}(x, w_N, t);$$

$$d_N = -(2w_N - w_{N-1} - \beta + h^2 f(x, w_N, t));$$

Paso 8 Tome $l_1 = a_1$;
 $u_1 = b_1/a_1$;
 $z_1 = d_1/l_1$;

Paso 9 Para $i = 2, \dots, N - 1$
 $l_i = a_i - c_i u_{i-1}$;
 $u_i = b_i/l_i$;
 $z_i = (d_i - c_i z_{i-1})/l_i$;

Paso 10 Tome $l_N = a_N - c_N u_{N-1}$;
 $z_N = (d_N - c_N z_{N-1})$;

Paso 11 Tome $v_N = z_N$;
 $w_N = w_N + v_N$;

Paso 12 Para $i = N - 1, \dots, 1$
tome $v_i = z_i - u_i v_{i+1}$;
 $w_i = w_i + v_i$;

Paso 13 Si $\|v\| \leq TOL$ (Hacer los Pasos 14-15)

Paso 14 Para $i = 0, \dots, N + 1$
tome $x = a + ih$;
SALIDA (x, w_i) ;

Paso 15 ALTO (El procedimiento fue completado.)

Paso 16 Tome $k = k + 1$;

Paso 17 SALIDA ('Número máximo de iteraciones fue excedido')
('El procedimiento no fue completado.')
ALTO.

□

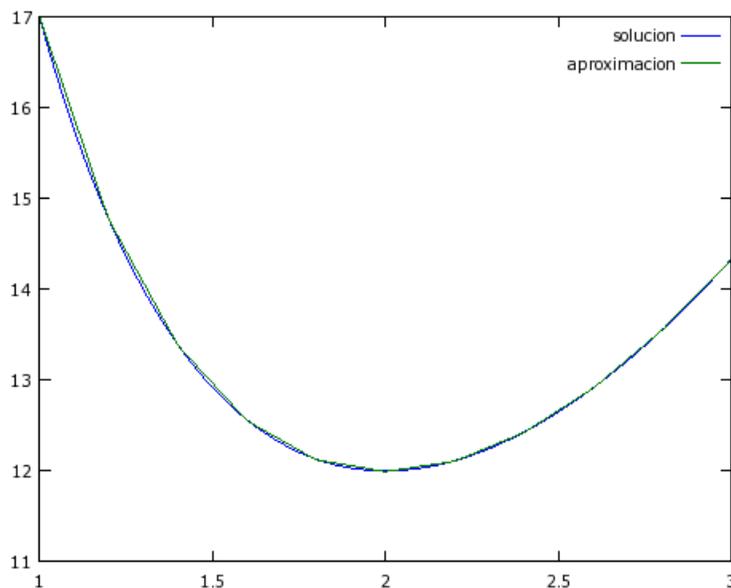
Ejemplo 8 : Al aplicar el algoritmo (4), con $N = 9$; con una tolerancia de 10^{-5} al problema no lineal de valor de frontera

$$y'' = \frac{1}{8}(32 + 2x^3 - yy'), \quad 1 \leq x \leq 3, \quad y(1) = 17, \quad y(3) = \frac{43}{3},$$

se obtienen los resultados:

X_i	Aprox $y(X_i)$	$y(X_i)$	error
1.00000	17.00000	17.00000	0.00000
1.20000	14.76689	14.77333	0.00645
1.40000	13.37937	13.38857	0.00920
1.60000	12.55005	12.56000	0.00995
1.80000	12.11935	12.12889	0.00954
2.00000	11.99157	12.00000	0.00843
2.20000	12.10583	12.11273	0.00690
2.40000	12.42150	12.42667	0.00516
2.60000	12.91048	12.91385	0.00337
2.80000	13.55267	13.55429	0.00162
3.00000	14.33333	14.33333	0.00000

En la siguiente imagen se obtienen las gráficas de la solución real y la aproximación:



9.5. El método de Rayleigh-Ritz.

El método del disparo para aproximar la solución de un problema de valor de frontera sustituyó dicho problema con un par de problemas de valor inicial. El enfoque de diferencias finitas reemplaza la operación continua de diferenciación con la operación discreta de

diferencias finitas. El método Rayleigh-Ritz es una variante que aborda el problema desde una tercera perspectiva. Primero se reformula el problema con valor de frontera como un problema que consiste en seleccionar, el conjunto de todas las funciones suficientemente derivables que satisfacen las condiciones de frontera, aquella que reduzca al mínimo determinada integral. Después, el tamaño del conjunto de todas las funciones factibles se disminuyen, obteniéndose así una aproximación a la solución al problema de minimización y, en consecuencia, una aproximación a la solución del problema con valor de frontera.

Para describir el método de Rayleigh-Ritz consideremos la aproximación de una solución al problema lineal con valor de frontera en dos puntos, a partir del análisis del esfuerzo de una viga. Este problema con valor en frontera se describe mediante la ecuación diferencial.

$$-\frac{d}{dx} \left(p(x) \frac{dy}{dx} \right) + q(x)y = f(x), \quad \text{para } 0 \leq x \leq 1, \quad (130)$$

con las condiciones de frontera

$$y(0) = y(1) = 0. \quad (131)$$

Esta ecuación describe la deflexión la función $y(x)$ de la viga de longitud 1 que tiene una sección transversal variable que está representada por $q(x)$. La deflexión se debe a los esfuerzos agregados $p(x)$ y $f(x)$.

En el análisis que sigue, supondremos que $p \in C^1[0, 1]$ y que $q, f \in C[0, 1]$. Más aún supondremos que existe una constante $\delta > 0$ tal que

$$p(x) \geq \delta, \text{ tal que } q(x) \geq 0, \quad \text{para cada } x \in [0, 1].$$

Las suposiciones anteriores son suficientes para garantizar que el problema de valor de frontera de la ecuación (131), tiene una única solución.

Problemas de Variación.

Como en el caso de los problemas con valor en frontera que describen fenómenos físicos, la solución a la ecuación de la viga satisface la propiedad **variacional**. En el caso de la ecuación de la viga, el principio variacional resulta indispensable para desarrollar el método de Rayleigh-Ritz y caracteriza la solución de esa ecuación como la función que reduce al mínimo cierta integral sobre las funciones en $C_0^2[0, 1]$, el conjunto de esas funciones u en $C^2[0, 1]$ con la propiedad de que $u(0) = u(1) = 0$.

El siguiente teorema establece la caracterización.

Teorema 39 : Sea $p \in C^1[0, 1]$, $q, f \in C[0, 1]$, y

$$p(x) \geq \delta > 0, \quad q(x) \geq 0, \quad \text{para } 0 \leq x \leq 1.$$

La función $y \in C_0^2[0, 1]$ es la solución única de la ecuación diferencial

$$-\frac{d}{dx} \left(p(x) \frac{dy}{dx} \right) + q(x)y = f(x), \quad \text{para } 0 \leq x \leq 1, \quad (132)$$

si y solo si y es la solución única en $C_0^2[0, 1]$, que reduce al mínimo la integral

$$I[u] = \int_0^1 \{p(x)[u'(x)]^2 + q(x)[u(x)]^2 - 2f(x)u(x)\} dx. \quad (133)$$

La demostración consta de tres pasos.

Primero se demuestra que cualquier solución y de (133) satisface también la ecuación

$$\int_0^1 f(x)u(x)dx = \int_0^1 p(x) \frac{dy}{dx}(x) \frac{du}{dx}(x) + q(x)y(x)u(x)dx, \quad (134)$$

para cada $u \in C_0^2[0, 1]$.

En el segundo paso se demuestra que $y \in C_0^2[0, 1]$ es solución de la ecuación (134) si y solo si la ecuación (134) se cumple para cada $u \in C_0^2[0, 1]$.

En el último paso se muestra que (134) tiene una única solución, la que además será solución de (134) y de (133), de modo que las soluciones de (133) y de (134) son idénticas.

Reduciendo al mínimo la integral, el método de Rayleigh-Ritz aproxima la solución y , no sobre todas las funciones en $C_0^2[0, 1]$, sino que sobre un conjunto más pequeño de las que contienen combinaciones lineales de ciertas funciones básicas $\phi_1, \phi_2, \dots, \phi_n$, las funciones básicas linealmente independientes y satisfacen:

$$\phi_i(0) = \phi_i(1) = 0, \quad \text{para cada } i = 1, 2, 3, \dots, n.$$

Después de encontrar las constantes c_1, c_2, \dots, c_n que reducen al mínimo $I[\sum_{i=1}^n c_i \phi_i]$, se obtiene una aproximación $\phi(x) = \sum_{i=1}^n c_i \phi_i(x)$ a la solución $y(x)$ de la ecuación (133).

Conforme a la ecuación (134)

$$\begin{aligned} I[\phi] &= I \left[\sum_{i=1}^n c_i \phi_i(x) \right], \\ &= \int_0^1 \left\{ p(x) \left[\sum_{i=1}^n c_i \phi_i'(x) \right]^2 + q(x) \left[\sum_{i=1}^n c_i \phi_i(x) \right]^2 - 2f(x) \sum_{i=1}^n c_i \phi_i(x) \right\} dx, \end{aligned} \quad (135)$$

y cuando se considera I como una función c_1, c_2, \dots, c_n , para que ocurra un mínimo es necesario tener

$$\frac{\partial I}{\partial c_j} = 0, \quad \text{para } j = 1, 2, \dots, n \quad (136)$$

Al derivar la ecuación (135) se obtiene

$$\frac{\partial I}{\partial c_j} = \int_0^1 \left\{ 2p(x) \sum_{i=1}^n c_i \phi'_i(x) \phi'_j(x) + 2q(x) \sum_{i=1}^n c_i \phi_i(x) \phi_j(x) - 2f(x) \phi_j(x) \right\} dx,$$

y al sustituir en la ecuación (136) se obtiene

$$0 = \sum_{i=1}^n \left[\int_0^1 \{ p(x) \phi'_i(x) \phi'_j(x) + q(x) \phi_i(x) \phi_j(x) dx \} \right] c_i - \int_0^1 f(x) \phi_j(x) dx, \quad (137)$$

para todo $j = 1, 2, \dots, n$.

Las ecuaciones descritas en anteriormente dan como resultado un sistema lineal de $n \times n$, digamos $Ac = b$ en las variables c_1, c_2, \dots, c_n , donde la matriz A es simétrica y esta dada por

$$a_{ij} = \int_0^1 [p(x) \phi'_i(x) \phi'_j(x) + q(x) \phi_i(x) \phi_j(x)] dx,$$

y \mathbf{b} se define por medio de

$$b_i = \int_0^1 f(x) \phi_i(x) dx.$$

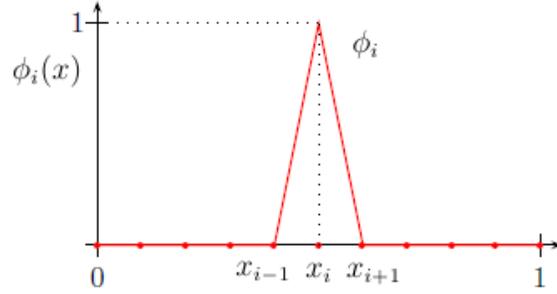
La elección más elemental de las funciones básicas requiere la intervención de polinomios lineales seccionados. El primer paso consiste en formar una partición de $[0, 1]$ al escoger los puntos x_0, x_1, \dots, x_{n+1} con

$$0 = x_0 < x_1 < \dots < x_n < x_{n+1} = 1.$$

Al utilizar $h_i = x_{i+1} - x_i$, para todo $i = 0, 1, 2, \dots, n$, definimos las funciones básicas $\phi_1(x), \phi_2(x), \dots, \phi_n(x)$, mediante

$$\phi_i(x) = \begin{cases} 0, & \text{si } 0 \leq x \leq x_{i-1}, \\ \frac{1}{h_{i-1}}(x - x_{i-1}), & \text{si } x_{i-1} < x \leq x_i, \\ \frac{1}{h_i}(x_{i+1} - x), & \text{si } x_i < x \leq x_{i+1}, \\ 0, & \text{si } x_{i+1} < x \leq 1, \end{cases} \quad (138)$$

para todo $i = 1, 2, \dots, n$, gráficamente tenemos:



Las funciones ϕ_i son lineales y seccionadas; por ello, aunque las derivadas ϕ'_i no son continuas, son constantes en el subintervalo abierto (x_j, x_{j+1}) para todo $j = 0, 1, 2, \dots, n$. Por tanto tenemos:

$$\phi'_i(x) = \begin{cases} 0, & \text{si } 0 < x < x_{i-1}, \\ \frac{1}{h_{i-1}}, & \text{si } x_{i-1} < x < x_i, \\ -\frac{1}{h_i}, & \text{si } x_i < x < x_{i+1}, \\ 0, & \text{si } x_{i+1} < x < 1, \end{cases} \quad (139)$$

para todo $i = 1, 2, \dots, n$.

Como ϕ_i y ϕ'_i , son distintos de cero sólo en (x_{i-1}, x_{i+1}) ,

$$\phi_i(x)\phi_j(x) \equiv 0, \quad \text{y} \quad \phi'_i(x)\phi'_j(x) \equiv 0,$$

excepto cuando j es $i-1, i, i+1$. En consecuencia, el sistema lineal dado por la ecuación (138) se reduce a un sistema lineal tridiagonal de $n \times n$. Los elementos distintos de cero de A son:

$$\begin{aligned} a_{ii} &= \int_0^1 \left\{ p(x) [\phi'_i(x)]^2 + q(x) [\phi_i(x)]^2 \right\} dx, \\ &= \left(\frac{1}{h_{i-1}} \right)^2 \int_{x_{i-1}}^{x_i} p(x) dx + \left(-\frac{1}{h_i} \right)^2 \int_{x_i}^{x_{i+1}} p(x) dx, \\ &\quad + \left(\frac{1}{h_{i-1}} \right)^2 \int_{x_{i-1}}^{x_i} (x - x_{i-1})^2 q(x) dx + \left(\frac{1}{h_i} \right)^2 \int_{x_i}^{x_{i+1}} (x_{i+1} - x)^2 q(x) dx, \end{aligned}$$

para todo $i = 1, 2, \dots, n$;

$$\begin{aligned} a_{i,i+1} &= \int_0^1 \{p(x)\phi'_i(x)\phi'_{i+1}(x) + q(x)\phi_i(x)\phi_{i+1}(x)\} dx, \\ &= -\left(\frac{1}{h_i}\right)^2 \int_{x_i}^{x_{i+1}} p(x)dx + \left(\frac{1}{h_i}\right)^2 \int_{x_i}^{x_{i+1}} (x_{i+1} - x)(x - x_i)q(x)dx, \end{aligned}$$

para todo $i = 1, 2, \dots, n - 1$; y

$$\begin{aligned} a_{i,i-1} &= \int_0^1 \{p(x)\phi'_i(x)\phi'_{i-1}(x) + q(x)\phi_i(x)\phi_{i-1}(x)\} dx, \\ &= -\left(\frac{1}{h_{i-1}}\right)^2 \int_{x_{i-1}}^{x_i} p(x)dx + \left(\frac{1}{h_{i-1}}\right)^2 \int_{x_{i-1}}^{x_i} (x_i - x)(x - x_{i-1})q(x)dx, \end{aligned}$$

para cada $i = 1, 2, \dots, n$. Las entradas en **b** son:

$$b_i = \int_0^1 f(x)\phi_i(x)dx = \frac{1}{h_{i-1}} \int_{x_{i-1}}^{x_i} (x - x_{i-1})f(x)dx + \frac{1}{h_i} \int_{x_i}^{x_{i+1}} (x_{i+1} - x)f(x)dx,$$

para todo $i = 1, 2, \dots, n$.

Hay seis tipos de integrales a evaluar :

$$Q_{1,i} = \left(\frac{1}{h_i}\right)^2 \int_{x_i}^{x_{i+1}} (x_{i+1} - x)(x - x_i)q(x)dx \quad \text{para } i = 1, 2, \dots, n - 1,$$

$$Q_{2,i} = \left(\frac{1}{h_{i-1}}\right)^2 \int_{x_{i-1}}^{x_i} (x - x_{i-1})^2 q(x)dx \quad \text{para } i = 1, 2, \dots, n,$$

$$Q_{3,i} = \left(\frac{1}{h_i}\right)^2 \int_{x_i}^{x_{i+1}} (x_{i+1} - x)^2 q(x)dx \quad \text{para } i = 1, 2, \dots, n,$$

$$Q_{4,i} = \left(\frac{1}{h_{i-1}}\right)^2 \int_{x_{i-1}}^{x_i} p(x)dx \quad \text{para } i = 1, 2, \dots, n + 1,$$

$$Q_{5,i} = \left(\frac{1}{h_{i-1}}\right) \int_{x_{i-1}}^{x_i} (x - x_{i-1})f(x)dx \quad \text{para } i = 1, 2, \dots, n,$$

$$Q_{6,i} = \left(\frac{1}{h_i}\right) \int_{x_i}^{x_{i+1}} (x_{i+1} - x)f(x)dx \quad \text{para } i = 1, 2, \dots, n,$$

La matriz A y el vector \mathbf{b} del sistema lineal $A\mathbf{c} = \mathbf{b}$ contienen los elementos

$$\begin{aligned} a_{i,i} &= Q_{4,i} + Q_{4,i+1} + Q_{2,i} + Q_{3,i}, & \text{para } i = 1, 2, \dots, n, \\ a_{i,i+1} &= -Q_{4,i+1} + Q_{1,i}, & \text{para } i = 1, 2, \dots, n-1, \\ a_{i,i-1} &= -Q_{4,i} + Q_{1,i-1}, & \text{para } i = 2, 3, \dots, n, \end{aligned}$$

y

$$b_i = Q_{5,i} + Q_{6,i}, \quad \text{para } i = 1, 2, \dots, n.$$

Los coeficientes de \mathbf{c} son desconocidos c_1, c_2, \dots, c_n , a partir de los cuales se construye la aproximación de Rayleigh-Ritz ϕ , dada por

$$\phi(x) = \sum_{i=1}^n c_i \phi_i.$$

Una dificultad práctica de este método es la necesidad de evaluarlas $6n$ integrales. Pueden evaluarse directamente o mediante una fórmula de cuadratura, como el método de Simpson. Un método alternativo para la evaluación de la integral consiste en aproximar las funciones p , q y f , con su polinomio interpolante lineal seccionado, e integrar luego la aproximación.

Supongamos por ejemplo la integral $Q_{1,i}$. La interpolación lineal segmentaria de $q(x)$ es

$$P_q(x) = \sum_{i=0}^{n+1} q(x_i) \phi_i(x),$$

donde ϕ_1, \dots, ϕ_n se definen en la ecuación (139) y

$$\phi_0(x) = \begin{cases} \frac{x_1 - x}{x_1}, & \text{si } 0 \leq x \leq x_1, \\ 0, & \text{en otro caso,} \end{cases}$$

y

$$\phi_{n+1}(x) = \begin{cases} \frac{x - x_n}{1 - x_n}, & \text{si } x_n \leq x \leq 1, \\ 0, & \text{en otro caso.} \end{cases}$$

Dado que el intervalo de integración es $[x_i, x_{i+1}]$, $P_q(x)$ se reduce a

$$P_q(x) = q(x_i) \phi_i(x) + q(x_{i+1}) \phi_{i+1}(x).$$

Este es el polinomio interpolante de primer grado, así:

$$|q(x) - P_q(x)| = \mathcal{O}(h_i^2), \quad \text{para } x_i \leq x \leq x_{i+1},$$

si $q \in C^2[x_i, x_{i+1}]$. Para todo $i = 1, 2, \dots, n-1$, la aproximación a $Q_{1,i}$ se obtiene al integrar la aproximación de la siguiente manera:

$$\begin{aligned} Q_{1,i} &= \left(\frac{1}{h_i}\right)^2 \int_{x_i}^{x_{i+1}} (x_{i+1} - x)(x - x_i)q(x)dx, \\ &\approx \left(\frac{1}{h_i}\right)^2 \int_{x_i}^{x_{i+1}} (x_{i+1} - x)(x - x_i) \left[\frac{q(x_i)(x_{i+1} - x)}{h_i} + \frac{q(x_{i+1})(x - x_i)}{h_i} \right] dx, \\ &= \frac{h_i}{12}[q(x_i) + q(x_{i+1})]. \end{aligned}$$

Más aún, si $q(x) \in C^2[x_i, x_{i+1}]$ entonces

$$\left| Q_{1,i} - \frac{h_i}{12}[q(x_i) + q(x_{i+1})] \right| = \mathcal{O}(h_i^3).$$

Las aproximaciones a las otras integrales se derivan de manera parecida y están dadas por:

$$\begin{aligned} Q_{2,i} &= \frac{h_{i-1}}{12}[3q(x_i) + q(x_{i-1})], & Q_{3,i} &= \frac{h_i}{12}[3q(x_i) + q(x_{i+1})], \\ Q_{4,i} &= \frac{1}{2h_{i-1}}[p(x_i) + p(x_{i-1})], & Q_{5,i} &= \frac{h_{i-1}}{6}[2f(x_i) + f(x_{i-1})], \\ Q_{6,i} &= \frac{h_i}{6}[2f(x_i) + f(x_{i+1})]. \end{aligned}$$

En el siguiente algoritmo se establece el sistema lineal tridiagonal, y se incorpora el algoritmo de la factorización de Crout para resolver el sistema. Las integrales $Q_{1,i}, Q_{2,i}, \dots, Q_{6,i}$ pueden calcularse por cualquiera de los métodos interpolantes o Simpson.

Algoritmo 5 *Método Segmentario Lineal de Rayleigh-Ritz.*

Aproximación a la solución del problema de valor de frontera

$$-\frac{d}{dx} \left(p(x) \frac{dy}{dx} \right) + q(x)y = f(x), \quad \text{para } 0 \leq x \leq 1, \quad \text{con } y(0) = 0 \text{ y } y(1) = 1;$$

con la función segmentaria lineal

$$\phi(x) = \sum_{i=1}^n c_i \phi_i(x)$$

ENTRADA: un entero $n \geq 1$; puntos $x_0 = 0 < x_1 < \cdots < x_n < x_{n+1} = 1$.

SALIDA: coeficientes c_1, c_2, \dots, c_n .

Paso 1 Para $i = 0, \dots, n$
tome $h_i = x_{i+1} - x_i$;

Paso 2 Para $i = 1, \dots, n$ defina las bases segmentarias lineales por

$$\phi_i(x) = \begin{cases} 0, & 0 \leq x \leq x_{i-1}, \\ \frac{x - x_{i-1}}{h_{i-1}}, & x_{i-1} \leq x \leq x_i, \\ \frac{x_{i+1} - x}{h_i}, & x_i \leq x \leq x_{i+1}, \\ 0, & x_{i+1} \leq x \leq 1. \end{cases}$$

Paso 3 Para cada $i = 1, \dots, n-1$ calcule
 $Q_{1,i}, Q_{2,i}, Q_{3,i}, Q_{4,i}, Q_{5,i}, Q_{6,i}$;
Calcule $Q_{2,n}, Q_{3,n}, Q_{4,n}, Q_{5,n}, Q_{6,n}$.

Paso 4 Para cada $i = 1, 2, \dots, n-1$
tome $\alpha_i = Q_{4,i} + Q_{4,i+1} + Q_{2,i} + Q_{3,i}$;
 $\beta_i = Q_{1,i} - Q_{4,i+1}$;
 $b_i = Q_{5,i} + Q_{6,i}$;

Paso 5 Tome $\alpha_n = Q_{4,n} + Q_{4,n+1} + Q_{2,n} + Q_{3,n}$;
 $b_n = Q_{5,n} + Q_{6,n}$;

Paso 6 Tome $a_1 = \alpha_1$; (Los pasos 6-10 resuelven un sistema tridiagonal lineal.)
 $\xi_i = \beta_i / \alpha_i$;
 $z_1 = b_1 / a_1$;

Paso 7 Para $i = 2, \dots, n-1$
Tome $a_i = \alpha_i - \beta_{i-1} \xi_{i-1}$;
 $\xi_i = \beta_i / a_i$;
 $z_i = (b_i - \beta_i z_{i-1}) / a_i$;

Paso 8 Tome $a_n = \alpha_n - \beta_{n-1}\xi_{n-1}$;
 $z_n = (b_n - \beta_{n-1}z_{n-1})/a_n$;

Paso 9 Tome $c_n = z_n$;
 SALIDA (c_n).

Paso 10 Para $i = n - 1, \dots, 1$
 tome $c_i = z_i - \xi_i c_{i+1}$;
 SALIDA (c_i).

Paso 11 Alto. (Procedimiento completo.)

□

En el siguiente ejemplo se utiliza el algoritmo (5).

Ejemplo 9 Considere el problema con valor en la frontera

$$-y'' + \pi^2 y = 2\pi^2 \sin(\pi x), \quad 0 \leq x \leq 1, \quad y(0) = y(1) = 0.$$

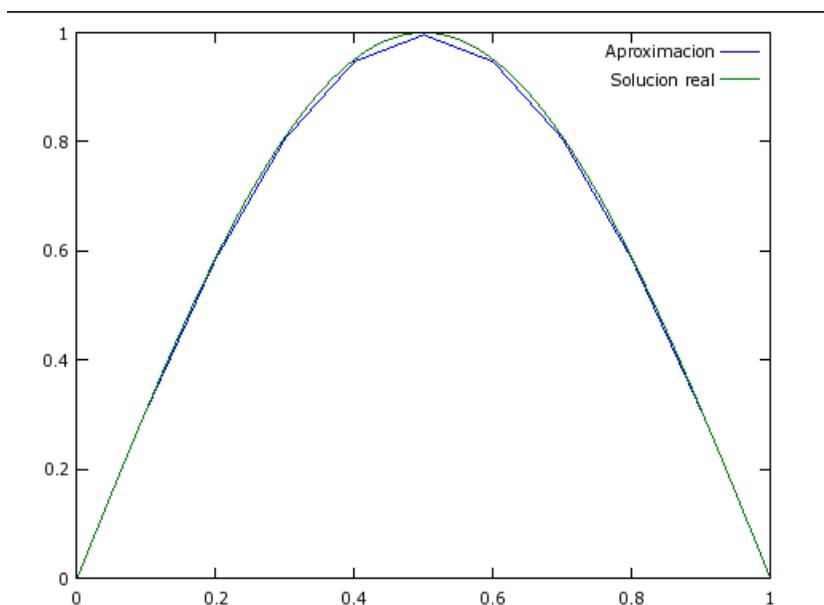
Sea $h_i = 0.1$ tal que $x_i = 0.1(i - 1)$ para $i = 1, 2, 3, \dots, 10$.

Aplicando el algoritmo (5), donde $p(x) = 1$, $q(x) = \pi^2$, $f(x) = 2\pi^2 \sin(\pi x)$ obtenemos la siguiente tabla, donde la aproximación lineal seccionada es

$$\phi(x) = \sum_{i=1}^9 C_i \phi_i$$

x_i	$\phi(x_i)$	$y(x_i)$	$ y(x_i) - \phi(x_i) $
0.1000000	0.3077473	0.3090170	0.0012697
0.2000000	0.5853702	0.5877853	0.0024151
0.3000000	0.8056929	0.8090170	0.0033241
0.4000000	0.9471488	0.9510565	0.0039077
0.5000000	0.9958912	1.0000000	0.0041088
0.6000000	0.9471488	0.9510565	0.0039077
0.7000000	0.8056929	0.8090170	0.0033241
0.8000000	0.5853702	0.5877853	0.0024151
0.9000000	0.3077473	0.3090170	0.0012697

En la siguiente imagen se muestra la gráfica de la solución real y de la aproximación:



9.6. B-Spline Básicas.

La utilización de las funciones lineales seccionadas básicas produce una solución aproximada a las ecuaciones (130) y (131), que es continua pero no diferenciable en $[0, 1]$. Se requiere un conjunto más complicado de funciones básicas para construir una aproximación que pertenezca a $C_0^2[0, 1]$. Dichas funciones se parecen a los trazadores cúbicos interpolantes.

Recordemos que el trazador cúbico interpolante S en los cinco nodos x_0, x_1, x_2, x_3 y x_4 para una función f está definido por:

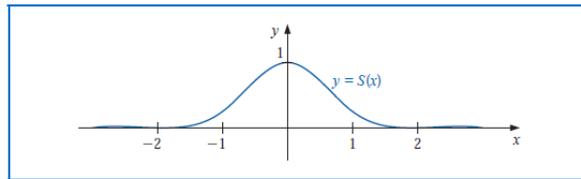
- a. S es un polinomio cúbico, denotado por S_j , en $[x_j, x_{j+1}] \forall j = 0, 1, 2, \dots$
- b. $S(x_j) = f(x_j)$, para $j = 0, 1, 2, 3, 4$,
- c. $S_{j+1}(x_{j+1}) = S_j(x_{j+1})$ para $j = 0, 1, 2$,
- d. $S'_{j+1}(x_{j+1}) = S'_j(x_{j+1})$ para $j = 0, 1, 2$,
- e. $S''_{j+1}(x_{j+1}) = S''_j(x_{j+1})$ para $j = 0, 1, 2$,
- f. Se satisface una de las siguientes condiciones de frontera:
 - (i) Libre: $S''(x_0) = S''(x_4) = 0$
 - (ii) Sujeto: $S'(x_0) = f'(x_0)$ y $S'(x_4) = f'(x_4)$

La unicidad de la solución requiere que el número de constantes en (a) que son 16, sea igual al de las condiciones en (b) a (f) por lo cual solo una de las condiciones de frontera en (f), puede especificarse para los trazadores cúbicos interpolantes.

Las funciones de trazadores cúbicos que utilizaremos en nuestras funciones básicas reciben el nombre de **trazadores B** ó *trazadores en forma de campana*. Estos difieren en los trazadores interpolantes en que satisfacen ambos conjuntos de las condiciones de frontera en (f). Para ello se debe flexibilizar dos de las condiciones de (b) a (e). Puesto que el trazador debe tener dos derivadas continuas en $[x_0, x_4]$, en la descripción de los trazadores interpolantes eliminamos dos de las condiciones de interpolación. En particular, modificamos la condición en (b) y la transformamos en:

$$(b) S(x_j) = f(x_j), \text{ para } j = 0, 2, 4,$$

El trazador **B** básico, S , que se define a continuación y que se muestra en la siguiente gráfica, usa los nodos uniformemente espaciados $x_0 = -2$, $x_1 = -1$, $x_2 = 0$, $x_3 = 1$, $x_4 = 2$. Satisface las condiciones de interpolación



$$(b) S(x_0) = 0, S(x_2) = 1, S(x_4) = 0.$$

Y también ambos conjunto de condiciones

$$(i) S'''(x_0) = S'''(x_4) = 0, \text{ y } (ii) S'(x_0) = S'(x_4) = 0.$$

En consecuencia, $S \in C_0^2(-\infty, \infty)$, y

$$S(x) = \begin{cases} 0, & \text{si } x \leq -2, \\ \frac{1}{4}(2+x)^3, & \text{si } -2 \leq x \leq -1, \\ \frac{1}{4}[(2+x)^3 - 4(1-x)^3], & \text{si } -1 < x \leq 0, \\ \frac{1}{4}[(2-x)^3 - 4(1-x)^3], & \text{si } 0 < x \leq 1, \\ \frac{1}{4}(2-x)^3, & \text{si } 1 < x \leq 2, \\ 0, & \text{si } 2 < x. \end{cases} \quad (140)$$

Para construir las funciones básicas ϕ_i en $C_0^2[0, 1]$, primero dividimos $[0, 1]$ seleccionando un entero positivo n y definiendo $h = \frac{1}{n+1}$. Así obtenemos los nodos uniformemente espaciados $x_i = hi$, para todo $i = 0, 1, 2, \dots, n+1$. Después definimos las funciones básicas $\{\phi_i\}_{i=0}^{n+1}$ como:

$$\phi_i(x) = \begin{cases} S\left(\frac{x}{h}\right) - 4S\left(\frac{x+h}{h}\right), & \text{si } i = 0, \\ S\left(\frac{x-h}{h}\right) - S\left(\frac{x+h}{h}\right), & \text{si } i = 1, \\ S\left(\frac{x-ih}{h}\right), & \text{si } 2 \leq i \leq n-1, \\ S\left(\frac{x-nh}{h}\right) - S\left(\frac{x-(n+2)h}{h}\right), & \text{si } i = n, \\ S\left(\frac{x-(n+1)h}{h}\right) - 4S\left(\frac{x-(n+2)h}{h}\right), & \text{si } i = n+1. \end{cases} \quad (141)$$

Donde $\{\phi_i\}_{i=0}^{n+1}$ es un conjunto de trazadores cúbicos linealmente independientes que satisfacen $\phi_i(0) = \phi_i(1) = 0$ para todo $i = 0, 1, 2, \dots, n, n+1$.

Problemos que $\{\phi_i\}$ son linealmente independientes.

La prueba la haremos por inducción: Caso base $n = 0$ tenemos que $a_0\phi_0 + a_1\phi_1 = 0$ con $a_0, a_1 \neq 0$, teniendo a si que $\phi_0 = -\frac{a_1}{a_0}\phi_1$ luego de evaluar ϕ_i tenemos que por los sistemas (140) y (141), por como estan construidas se llega a una contradicción por lo tanto ϕ_0, ϕ_1 son linealmente independientes.

Hipótesis inductiva.

Supongamos que se cumple para $n = k$, y probemos que se cumplen para $n = k + 1$. De lo que tenemos lo siguiente

$$a_0\phi_0 + a_1\phi_1 + \dots, a_k\phi_k + a_{k+1}\phi_{k+1} = 0,$$

y por hipótesis inductiva tenemos que:

$$a_0\phi_0 + a_1\phi_1 + \dots, a_k\phi_k = 0,$$

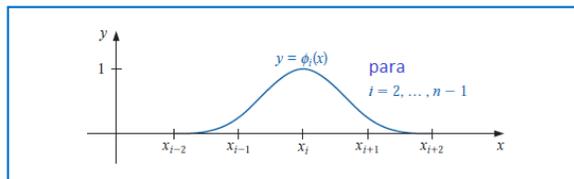
entonces

$$a_{k+1}\phi_{k+1} = 0 \Leftrightarrow a_{k+1} = 0,$$

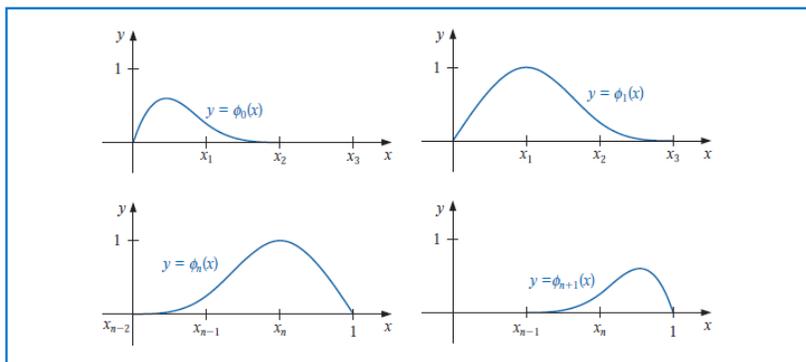
por lo tanto $\{\phi_i\}_{i=0}^{n+1}$ son linealmente independientes.

□

Las gráficas de ϕ_i para $2 \leq i \leq n - 1$ se muestran en la siguiente gráfica:



Aquí se muestran las gráficas de $\phi_0, \phi_1, \phi_n, \phi_{n+1}$.



Puesto que $\phi_i(x)$ y $\phi'_i(x)$ son distintas de cero sólo para $x \in [x_{i-2}, x_{i+2}]$, la matriz de aproximación de Rayleigh-Ritz es una matriz de banda con un ancho máximo de banda de siete:

$$A = \begin{bmatrix} a_{0,0} & a_{0,1} & a_{0,2} & a_{0,3} & 0 & \dots & \dots & \dots & 0 \\ a_{1,0} & a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} & \ddots & & & \vdots \\ a_{2,0} & a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} & a_{2,5} & \ddots & & \vdots \\ a_{3,0} & a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} & a_{3,5} & a_{3,6} & \ddots & \vdots \\ a_{4,0} & a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} & a_{4,5} & a_{4,6} & \ddots & \vdots \\ 0 & a_{5,1} & a_{5,2} & a_{5,3} & a_{5,4} & a_{5,5} & a_{5,6} & \ddots & 0 \\ \vdots & \ddots & a_{n-2,n+1} \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & a_{n-1,n+1} \\ \vdots & & & \ddots & \ddots & \ddots & \ddots & \ddots & a_{n,n+1} \\ 0 & \dots & \dots & \dots & 0 & a_{n+1,n-2} & a_{n+1,n-1} & a_{n+1,n} & a_{n+1,n+1} \end{bmatrix} \quad (142)$$

donde

$$a_{i,j} = \int_0^1 \{p(x)\phi'_i(x)\phi'_j(x) + q(x)\phi_i(x)\phi_j(x)\}dx.$$

para todo $i, j = 0, 1, 2, \dots, n, n+1$. El vector b tiene los términos

$$b_i = \int_0^1 f(x)\phi_i(x)dx.$$

Veamos que la matriz A es definida positiva.

Primero veamos que es simétrica de la siguiente manera:

$$\begin{aligned} a_{i,j} &= \int_0^1 \{p(x)\phi'_i(x)\phi'_j(x) + q(x)\phi_i(x)\phi_j(x)\}dx. \\ &= \int_0^1 \{p(x)\phi'_j(x)\phi'_i(x) + q(x)\phi_j(x)\phi_i(x)\}dx. \\ &= a_{j,i}, \end{aligned}$$

para todo $i \neq j$ por lo tanto la matriz A es simétrica.

Ahora tenemos que para $C = (c_0, c_1, c_2, \dots, c_n, c_{n+1})^t$ y $\phi(x) = \sum_{i=0}^{n+1} c_i\phi_i(x)$, de aquí tenemos que:

$$C^t A C = \int_0^1 \{p(x)[\phi'(x)]^2 + q(x)[\phi(x)]^2\}dx.$$

Como estamos considerando que $p(x) > 0$, además $q(x) \geq 0$ de aquí tenemos que $q(x)[\phi(x)]^2 \geq 0$, si consideramos que $C^t AC = 0$ para todo vector X distinto del vector cero, si y solo si $\phi'(x) \equiv 0$ en el intervalo $[0, 1]$. Como teníamos que $\phi'_0, \phi'_1, \dots, \phi'_{n+1}$ forman una base linealmente independientes, también tenemos que $\phi'(x) \neq 0$ en el intervalo $[0, 1]$ y así que $C^t AC = 0$ si y solo si C es igual al vector cero.

Por lo tanto tenemos que la matriz A es definida positiva, y por teorema que dice: **Si una matriz es definida positiva entonces es invertible.** Esto nos garantiza que nuestro sistema tiene solución única. \square

También el sistema lineal $Ac = \mathbf{b}$ puede ser resuelto por el algoritmo de **Cholesky's** o por **Eliminación Gaussiana**. En el siguiente algoritmo detalla la construcción de la aproximación cubica spline $\phi(x)$ por el método de **Rayleigh-Ritz** para los problema de valor de frontera lineales.

Algoritmo 6 Método Cubico Spline de Rayleigh-Ritz.

Aproximación a la solución del problema de valor de frontera

$$-\frac{d}{dx} \left(p(x) \frac{dy}{dx} \right) + q(x)y = f(x), \text{ para } 0 \leq x \leq 1, \text{ con } y(0) = 0 \text{ y } y(1) = 1;$$

con la suma de cúbicos spline

$$\phi(x) = \sum_{i=1}^n c_i \phi_i(x)$$

ENTRADA: un entero $n \geq 1$.

SALIDA: coeficientes c_0, c_1, \dots, c_{n+1} .

Paso 1 Tome $h = 1/(n + 1)$;

*Paso 2 Para $i = 0, 1, \dots, n + 1$
tome $x_i = ih$;
 $x_{-2} = x_{-1} = 0$;
 $x_{n+2} = x_{n+3} = 1$;*

Paso 3 Defina la función S por

$$S(x) = \begin{cases} 0, & x \leq -2 \\ \frac{1}{4}(2+x)^3 & -2 < x \leq -1, \\ \frac{1}{4}[(2+x)^3 - 4(1+x)^3], & -1 < x \leq 0, \\ \frac{1}{4}[(2-x)^3 - 4(1-x)^3], & 0 < x \leq 1, \\ \frac{1}{4}(2-x)^3, & 2 < x \end{cases}$$

Paso 4 Defina las bases cúbicas spline $\{\phi_i\}_{i=0}^{n+1}$ por

$$\phi_0(x) = S\left(\frac{x}{h}\right) - 4S\left(\frac{x+h}{h}\right);$$

$$\phi_1(x) = S\left(\frac{x-h}{h}\right) - S\left(\frac{x+h}{h}\right);$$

$$\phi_i(x) = S\left(\frac{x-ih}{h}\right); \text{ para } i = 2, \dots, n-1;$$

$$\phi_n(x) = S\left(\frac{x-nh}{h}\right) - S\left(\frac{x-(n+2)h}{h}\right);$$

$$\phi_{n+1}(x) = S\left(\frac{x-(n+1)h}{h}\right) - 4S\left(\frac{x-(n+2)h}{h}\right);$$

Paso 5 Para $i = 0, \dots, n+1$ hacer los pasos 6-9

Paso 6 Para $j = i, i+1, \dots, \min\{i+3, n+1\}$

$$\text{tome } L = \max\{x_{j-2}, 0\};$$

$$U = \min\{x_{i+2}, 1\};$$

$$a_{i,j} = \int_L^U [p(x)\phi'_i(x)\phi'_j(x) + q(x)\phi_i(x)\phi_j(x)] dx;$$

Si $i \neq j$, tome $a_{j,i} = a_{i,j}$. (Por ser A simétrica)

Paso 7 Si $i \geq 4$ entonces para $j = 0, \dots, i - 4$ tome $a_{i,j} = 0$.

Paso 8 Si $i \leq n - 3$ entonces para $j = i + 4, \dots, n + 1$ tome $a_{i,j} = 0$.

Paso 9 Tome $L = \max\{x_{i-2}, 0\}$;
 $U = \min\{x_{i+2}, 1\}$;
 $b_i = \int_L^U f(x)\phi_i(x)dx$.

Paso 10 Resolver el sistema lineal $Ac = b$, donde $A = (a_{i,j})$, $b = (b_0, \dots, b_{n+1})^t$
 $c = (c_0, \dots, c_{n+1})^t$.

Paso 11 Para $i = 0, \dots, n + 1$
 SALIDA (c_i);

Paso 12 Parar (Procedimiento completado.)

□

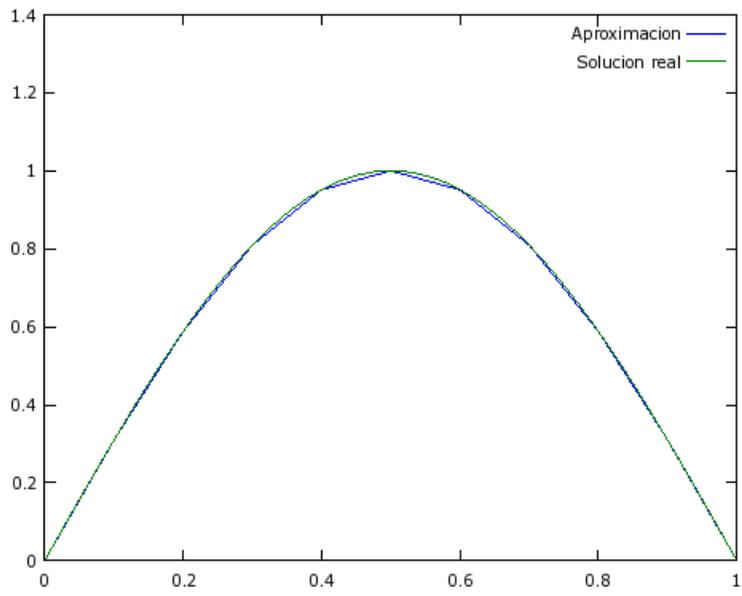
Ejemplo 10 Considere el problema de valor de frontera

$$-y'' + \pi^2 y = 2\pi^2 \sin(x), \text{ para } 0 \leq x \leq 1, \text{ con } y(0) = y(1) = 0.$$

En la siguiente ilustración del algoritmo (6) tomamos $h = 0.1$ obtenemos:

x_i	c_i	$\phi(x_i)$	$y(x_i)$	error
0.00000	0.00001	0.00000	0.00000	0.00000
0.10000	0.20943	0.30902	0.30902	0.00000
0.20000	0.39836	0.58779	0.58779	0.00001
0.30000	0.54830	0.80903	0.80902	0.00001
0.40000	0.64456	0.95107	0.95106	0.00001
0.50000	0.67773	1.00001	1.00000	0.00001
0.60000	0.64456	0.95107	0.95106	0.00001
0.70000	0.54830	0.80903	0.80902	0.00001
0.80000	0.39836	0.58779	0.58779	0.00001
0.90000	0.20943	0.30902	0.30902	0.00000
1.00000	0.00001	0.00000	0.00000	0.00000

En la siguiente imagen se muestra la gráfica de la solución real y de la aproximación:



CAPITULO CINCO.

EXPERIMENTACIÓN NUMÉRICA.

10. EXPERIMENTACIÓN NUMÉRICA.

En esta sección se hace un estudio en cuanto a las aproximaciones que haces los diferentes métodos numéricos para aproximar las diferentes ecuaciones diferenciales.

Ejemplo 11 *Demuestre que el problema con valor en la frontera*

$$\frac{d}{dx} \left(p(x) \frac{dy}{dx} \right) + q(x)y = f(x), \quad (143)$$

y

$$0 \leq x \leq 1, \quad y(0) = \alpha, \quad y(1) = \beta,$$

puede transformarse con un cambio de variable, en la forma

$$-\frac{d}{dx} \left(p(x) \frac{dz}{dx} \right) + q(x)z = F(x), \quad 0 \leq x \leq 1, \quad z(0) = z(1) = 0.$$

Prueba:

Hacemos un cambio de variable

$$z(x) = y(x) - \beta x - (1-x)\alpha,$$

donde

$$\begin{aligned} z(0) &= y(0) - \beta(0) - (1-0)\alpha = 0 \\ z(1) &= y(1) - \beta(1) - (1-1)\alpha = 0. \end{aligned}$$

Además $z'(x) = y'(x) - \beta + \alpha$.

Sustituyendo a $y(x) = z(x) + \beta x + (1-x)\alpha$ y $y'(x) = z'(x) + \beta - \alpha$, en (143) tenemos.

$$\begin{aligned} -\frac{d}{dx} (p(x)(z' + \beta - \alpha)) + q(x)(z + \beta x + (1-x)\alpha) &= f(x) \\ -\frac{d}{dx} (p(x)z') - \frac{d}{dx} (p(x)(\beta - \alpha)) + q(x)(z) + q(x)(\beta x + (1-x)\alpha) &= f(x) \\ -\frac{d}{dx} (p(x)z') + q(x)(z) &= f(x) + p'(x)(\beta - \alpha) - q(x)(\beta x + (1-x)\alpha) \\ -\frac{d}{dx} (p(x)z') + q(x)(z) &= F(x). \end{aligned}$$

Donde $F(x) = f(x) + p'(x)(\beta - \alpha) - q(x)(\beta x + (1-x)\alpha)$. Entonces

$$-\frac{d}{dx} \left(p(x) \frac{dz}{dx} \right) + q(x)z = F(x), \quad 0 \leq x \leq 1, \quad z(0) = z(1) = 0.$$

□

Para finalizar, se presenta un conjunto de gráficas y los datos correspondientes de los métodos en estudio, para hacer una comparación de los métodos lineales y no lineales.

Análisis de los métodos lineales

Ejemplo: Consideremos el siguiente problema de valor de frontera lineal.

$$y'' = -27x + \frac{28}{3}, \quad (144)$$

$$y(0) = 1, \quad y(1) = 0, \quad n = 10. \quad (145)$$

el cual tiene por solución exacta

$$y(x) = -\frac{9}{2}x^3 + \frac{14}{3}x^2 - \frac{7}{3}x + 1. \quad (146)$$

Solución:

Para aplicar el método de Rayleigh-Ritz y el método de trazadores cúbicos de Rayleigh-Ritz debemos escribir la E.D (144) y las condiciones de frontera (145) en la forma

$$-\frac{d}{dx} \left(p(x) \frac{dy}{dx} \right) + q(x)y = r(x),$$

y

$$y(0) = y(1) = 0.$$

Para ello vamos a ocupar el ejemplo (11), haciendo un cambio de variable

$$z(x) = y(x) - (1 - x), \quad (147)$$

obtenemos la E.D

$$-\frac{d}{dx} \left(-\frac{dz}{dx} \right) = -27x + \frac{28}{3}, \quad 0 \leq x \leq 1, \quad (148)$$

y las condiciones de frontera

$$z(0) = z(1) = 0. \quad (149)$$

Entonces la solución para la E.D (148) se transforma en

$$z(x) = -\frac{9}{x}x^3 + \frac{14}{3}x^2 - \frac{x}{6}.$$

Método del disparo.

x_i	$z(x_i) \approx y$	$y(x_i)$	$ y(x_i) - z(x_i) $
0.1	1.00000	1.00000	0
0.2	0.92550	0.92550	0
0.3	0.91733	0.91733	0
0.4	0.94850	0.94850	0
0.5	0.99200	0.99200	2.2204×10^{-16}
0.6	1.02083	1.02083	0
0.7	1.00800	1.00800	4.4409×10^{-16}
0.8	0.92650	0.92650	1.1102×10^{-16}
0.9	0.74933	0.74933	5.5511×10^{-16}
1.0	0.44950	0.44950	4.4409×10^{-16}

z : solución aproximada, y : solución real

Método de diferencias finitas.

x_i	$z(x_i) \approx y$	$y(x_i)$	$ y(x_i) - z(x_i) $
0.1	1.00000	1.00000	0
0.2	0.92550	0.92550	0
0.3	0.91733	0.91733	0
0.4	0.94850	0.94850	2.2204×10^{-16}
0.5	0.99200	0.99200	4.4409×10^{-16}
0.6	1.02083	1.02083	2.2204×10^{-16}
0.7	1.00800	1.00800	8.8818×10^{-16}
0.8	0.92650	0.92650	4.4409×10^{-16}
0.9	0.74933	0.74933	9.9920×10^{-16}
1.0	0.44950	0.44949	8.3267×10^{-16}

z : solución aproximada, y : solución real

Método de Rayleigh-Ritz.

x_i	$\psi(x_i) \approx z$	$z(x_i)$	$y(x_i) = z(x_i) - (x - 1)$	$ y(x_i) - \phi(x_i) $
0.1	0.02500	0.02500	0.92550	2.4286×10^{-17}
0.2	0.11733	0.11733	0.91733	4.1633×10^{-17}
0.3	0.248500	0.248500	0.94850	8.3267×10^{-17}
0.4	0.392000	0.392000	0.99200	1.6653×10^{-16}
0.5	0.520833	0.520833	1.02083	1.1102×10^{-16}
0.6	0.608000	0.608000	1.00800	$0.3.3307 \times 10^{-16}$
0.7	0.626500	0.626500	0.92650	0
0.8	0.549333	0.549333	0.74933	5.5511×10^{-16}
0.9	0.349500	0.349500	0.44950	6.1062×10^{-16}

ψ : solución aproximada de la E.D (148), z : solución real de la E.D (148)
 y : solución real de la E.D (150), ϕ : solución aproximada de la E.D (150).

Método de trazadores cúbicos splin de Rayleigh-Ritz.

x_i	C_i	$\psi(x_i) \approx z$	$z(x_i)$	$y(x_i)$	$ y(x_i) - \phi(x_i) $
0	-0.00857	0	0	1	0
0.09090	0.00703	0.02003	0.02003	0.92912	2.6160×10^{-7}
0.18181	0.06055	0.09691	0.09691	0.91510	4.7127×10^{-7}
0.27272	0.13843	0.21036	0.21036	0.93764	9.4429×10^{-7}
0.36363	0.22717	0.34009	0.34009	0.97645	1.4009×10^{-6}
0.45454	0.31324	0.46581	0.46581	1.01126	1.9062×10^{-6}
0.54545	0.38311	0.56724	0.56724	1.02178	2.4817×10^{-6}
0.63636	0.42326	0.62408	0.62408	0.98772	2.8805×10^{-6}
0.72727	0.42017	0.61607	0.61607	0.88880	2.8911×10^{-6}
0.81818	0.36032	0.52291	0.52291	0.70473	2.4731×10^{-6}
0.90909	0.23017	0.32431	0.32431	0.41522	1.1161×10^{-6}
1	0.01622	0	0	0	$3.48541.1161 \times 10^{-16}$

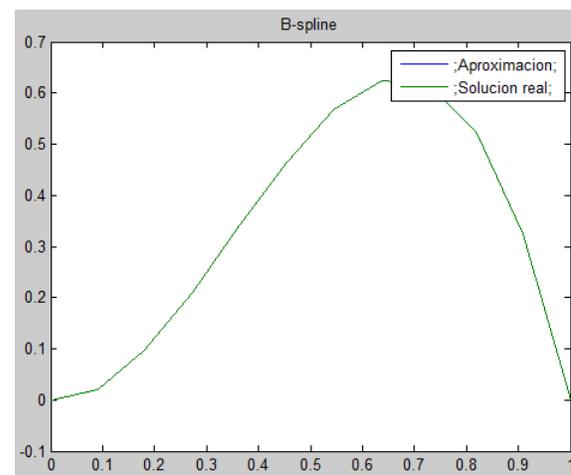
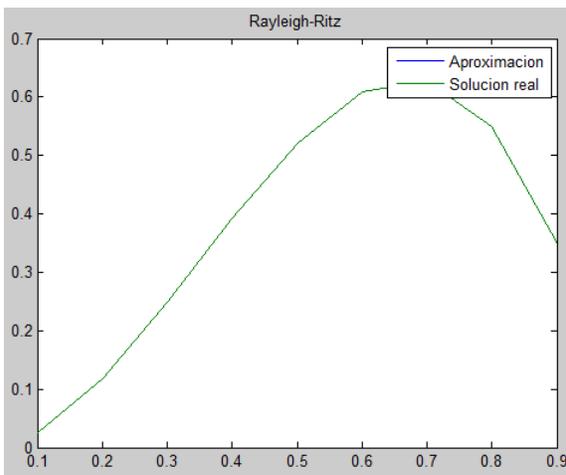
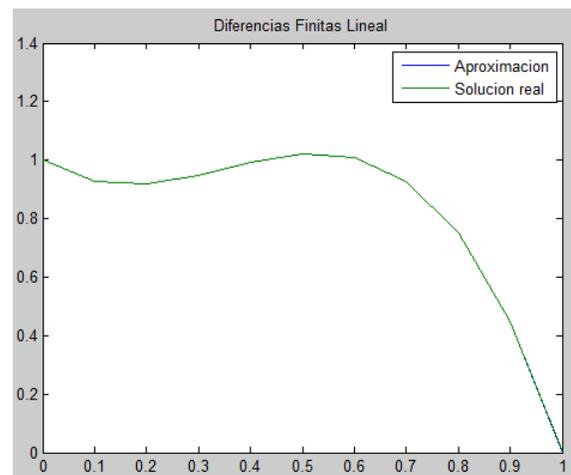
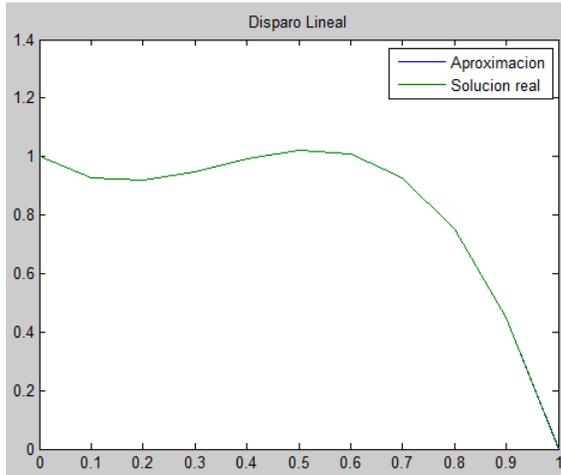
C_i : constantes de la función base, ψ : solución aproximada de la ecuación (148), z :

solución real de la ecuación (148), y : solución real de la ecuación (150), ϕ : solución aproximada de la ecuación (150).

Se particiona el intervalo $[0, 1]$ en diez subintervalos de igual longitud y en virtud de las tablas de cada método y basados en los errores respecto a la solución real y aproximada de los métodos del Disparo, Diferencias Finitas, Rayleigh-Ritz y Trazadores cúbicos de Rayleigh-Ritz, para este caso particular el método del disparo aproxima mejor a la E.D (150).

Se muestra en las gráficas las soluciones reales y aproximadas, y aunque gráficamente se observa que los cuatro métodos aproximan de la mejor manera a la solución real, podemos ver mediante el error que el método del disparo es mejor para modelar la solución. Aunque esto no quita el hecho que los demás métodos den una buena aproximación.

Gráfica de la solución real y la solución aproximada.



Ejemplo: Consideremos el siguiente problema de valor de frontera lineal

$$-\frac{d}{dx}(e^x y') + e^x y = x + (2+x)e^x, \quad y(0) = y(1) = 0, \quad h = 0.1$$

y la solución real

$$y(x) = (x-1)(e^{-x} - 1).$$

Método del disparo.

x_i	$w(x_i) \approx y$	$y(x_i)$	$ y(x_i) - w(x_i) $
0.	0	0	0
0.1	0.0856457	0.0856463	6.1030×10^{-7}
0.2	0.1450144	0.1450153	9.9604×10^{-7}
0.3	0.1814260	0.1814272	1.2025×10^{-6}
0.4	0.1978067	0.1978079	1.2668×10^{-6}
0.5	0.1967334	0.1967346	1.2191×10^{-6}
0.6	0.1804742	0.1804753	1.0835×10^{-6}
0.7	0.1510235	0.1510244	8.7987×10^{-7}
0.8	0.1101335	0.1101342	6.2358×10^{-7}
0.9	0.0593427	0.0593430	3.2707×10^{-7}
1.0	0	0	0

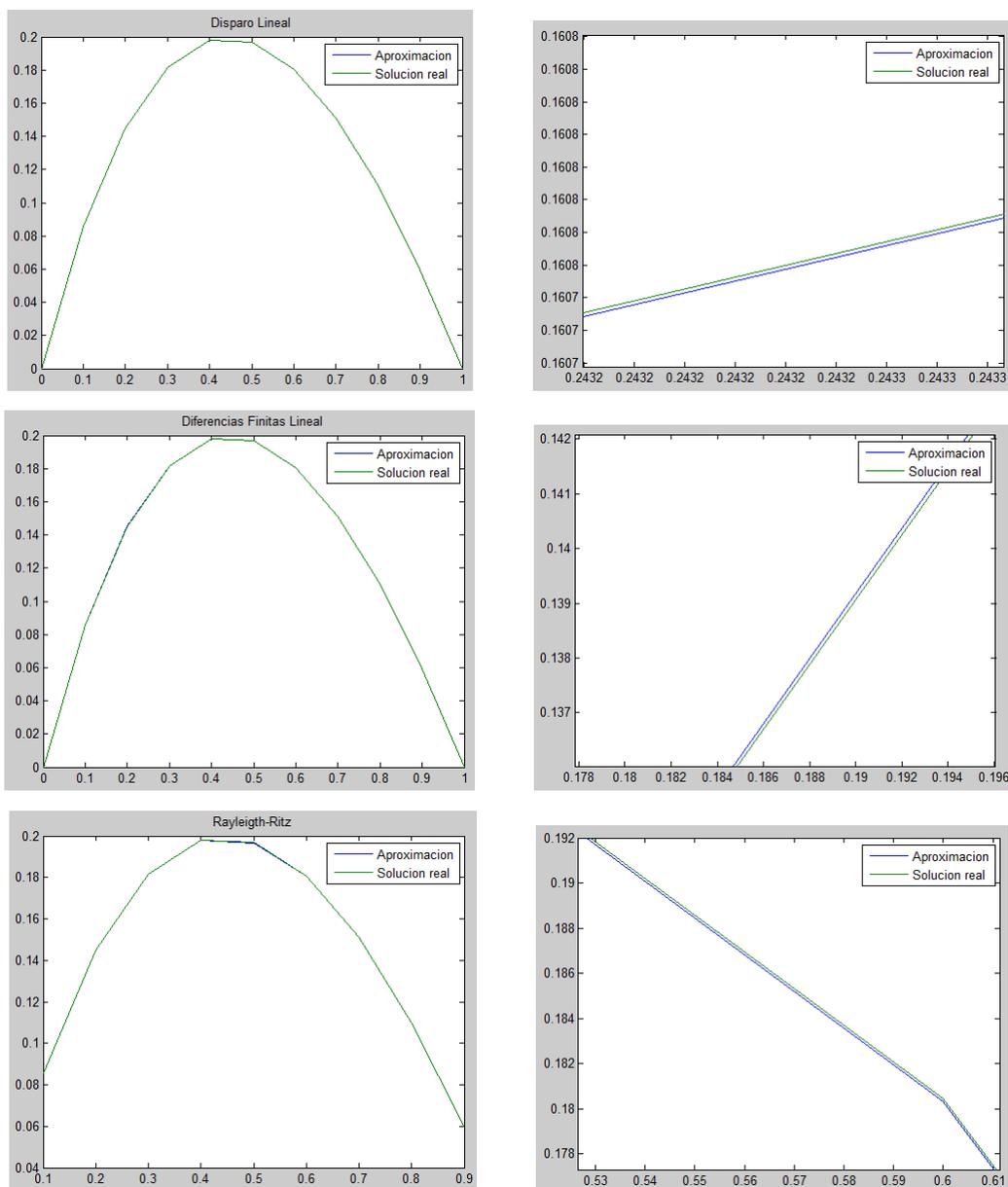
Método de diferencias finitas.

x_i	$w(x_i) \approx y$	$y(x_i)$	$ y(x_i) - w(x_i) $
0.	0	0	0
0.1	0.08571	0.08564	7.2152×10^{-5}
0.2	0.14513	0.14501	1.1726×10^{-4}
0.3	0.18156	0.18142	1.4098×10^{-4}
0.4	0.19795	0.19780	1.4789×10^{-4}
0.5	0.19687	0.19673	1.4170×10^{-4}
0.6	0.18060	0.18047	1.2540×10^{-4}
0.7	0.15112	0.15102	1.0138×10^{-4}
0.8	0.11020	0.11013	7.1534×10^{-5}
0.9	0.05938	0.05934	3.7354×10^{-5}
1.0	0	0	0

Método de Rayleigh-Ritz.

x_i	$\phi(x_i)$	$y(x_i)$	$ y(x_i) - \phi(x_i) $
0.1	0.085605	0.085646	4.1015×10^{-5}
0.2	0.144940	0.145015	7.5580×10^{-5}
0.3	0.181325	0.181427	1.0211×10^{-4}
0.4	0.197689	0.197808	1.1945×10^{-4}
0.5	0.196608	0.196735	1.2679×10^{-4}
0.6	0.180352	0.180475	1.2358×10^{-3}
0.7	0.150915	0.151024	1.0948×10^{-4}
0.8	0.110050	0.110134	8.4247×10^{-5}
0.9	0.059295	0.059343	4.7775×10^{-5}

Gráfica de la solución real y la solución aproximada.



Análisis de los métodos no lineales.

Ejemplo: Consideremos el siguiente problema de valor de frontera no lineal

$$y'' = \frac{1}{8}(32 + 2x^3 - yy'), \quad 1 \leq x \leq 3 \quad (150)$$

$$y(1) = 17, \quad y(3) = \frac{43}{3}, \quad \text{usando } N = 20, M = 10, TOL = 10^5 \quad (151)$$

el cual tiene por solución exacta

$$y(x) = x^2 + \frac{16}{x}. \quad (152)$$

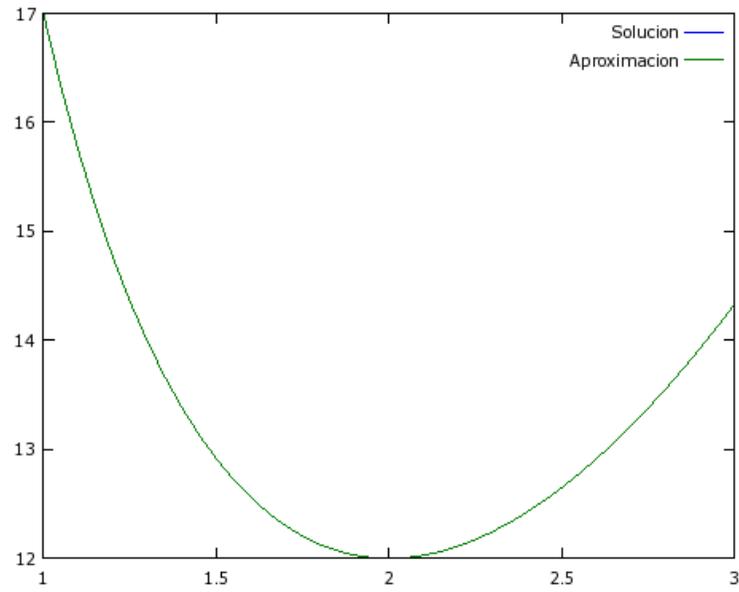
Método Disparo no lineal.

X_i	Aprox $y(X_i)$	Aprox $dy/dx(y(X_i))$	$y(X_i)$	error
1.00000	17.00000	-14.00001	17.00000	0.00000
1.05000	16.34060	-12.41249	16.34060	0.00000
1.10000	15.75546	-11.02315	15.75545	0.00000
1.15000	15.23555	-9.79831	15.23554	0.00000
1.20000	14.77334	-8.71112	14.77333	0.00000
1.25000	14.36250	-7.74001	14.36250	0.00000
1.30000	13.99770	-6.86747	13.99769	0.00000
1.35000	13.67435	-6.07916	13.67435	0.00000
1.40000	13.38857	-5.36328	13.38857	0.00000
1.45000	13.13699	-4.71000	13.13698	0.00000
1.50000	12.91667	-4.11112	12.91667	0.00000
1.55000	12.72508	-3.55974	12.72508	0.00000
1.60000	12.56000	-3.05001	12.56000	0.00000
1.65000	12.41947	-2.57696	12.41947	0.00000
1.70000	12.30177	-2.13634	12.30176	0.00000
1.75000	12.20536	-1.72450	12.20536	0.00000
1.80000	12.12889	-1.33828	12.12889	0.00000
1.85000	12.07115	-0.97495	12.07115	0.00000
1.90000	12.03105	-0.63214	12.03105	0.00000
1.95000	12.00763	-0.30776	12.00763	0.00000
2.00000	12.00000	-0.00000	12.00000	0.00000
2.05000	12.00738	0.29274	12.00738	0.00000
2.10000	12.02905	0.57188	12.02905	0.00000
2.15000	12.06436	0.83867	12.06436	0.00000
2.20000	12.11273	1.09421	12.11273	0.00000
2.25000	12.17361	1.33950	12.17361	0.00000
2.30000	12.24652	1.57542	12.24652	0.00000
2.35000	12.33101	1.80276	12.33101	0.00000
2.40000	12.42667	2.02222	12.42667	0.00000
2.45000	12.53311	2.23444	12.53311	0.00000
2.50000	12.65000	2.44000	12.65000	0.00000
2.55000	12.77701	2.63941	12.77701	0.00000
2.60000	12.91384	2.83313	12.91385	0.00000
2.65000	13.06023	3.02161	13.06024	0.00000
2.70000	13.21592	3.20521	13.21593	0.00000
2.75000	13.38068	3.38430	13.38068	0.00000
2.80000	13.55428	3.55918	13.55429	0.00000
2.85000	13.73653	3.73016	13.73654	0.00000
2.90000	13.92724	3.89750	13.92724	0.00000
2.95000	14.12623	4.06145	14.12623	0.00000
3.00000	14.33333	4.22222	14.33333	0.00000

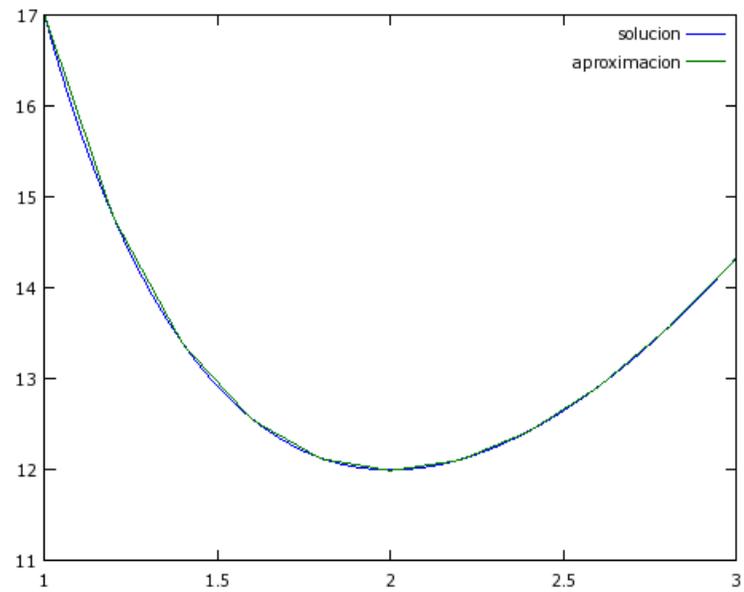
Método de diferencias finitas

x_i	$w(x_i) \approx y$	$y(x_i)$	$ y(x_i) - w(x_i) $
1	17.0000	17.0000	0
1.1	15.8042	15.7555	4.8725×10^{-2}
1.2	14.7429	14.7733	3.0454×10^{-2}
1.3	13.8739	13.9977	1.2378×10^{-1}
1.4	13.2226	13.3886	1.6599×10^{-1}
1.5	12.7548	12.9167	1.6182×10^{-1}
1.6	12.4196	12.5600	1.4037×10^{-1}
1.7	12.1839	12.3018	1.1784×10^{-1}
1.8	12.0306	12.1289	9.833×10^{-2}
1.9	11.9491	12.0311	8.1943×10^{-2}
2.0	11.9319	12.0000	6.8146×10^{-2}
2.1	11.9726	12.0290	5.6443×10^{-2}
2.2	12.0663	12.1127	4.6423×10^{-2}
2.3	12.2088	12.2465	3.7763×10^{-2}
2.4	12.3965	12.4267	3.0208×10^{-2}
2.5	12.6264	12.6500	2.3558×10^{-2}
2.6	12.8962	12.9138	1.7664×10^{-2}
2.7	13.2035	13.2159	1.2416×10^{-2}
2.8	13.5465	13.5543	7.7423×10^{-3}
2.9	13.9236	13.9272	3.6041×10^{-3}
3	14.3333	14.3333	2×10^{-15}

Gráfica según Método diferencias finitas no lineal.



Gráfica según Método Disparo no lineal.



11. Cronograma de Actividades.

Cronograma de Actividades Julio-Diciembre de 2014

Actividades	Julio				Agosto				Septiembre				Octubre				Noviembre				Diciembre											
	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4					
Reunión con asesor.																																
Establecimiento del tema del T. I.																																
Inscripción del T. I.																																
Búsqueda de Bibliografía																																
Elaboración del Perfil del T. I.																																
Entrega del Perfil del T.I. a los asesores																																
Primeras Correcciones al Perfil del T.I.																																
Entrega del Perfil del T.I. a un evaluador externo																																
Envío del Perfil del T. I. a Junta Directiva																																
Presentación del Perfil del T. I.																																
Desarrollo del primer avance																																
Presentación del primer avance																																
Desarrollo del segundo avance																																
Presentación del segundo avance																																
Desarrollo de aplicaciones																																
Presentación Final: Resultados y Aplicaciones																																
Elaboración del Documento Final																																

T.I. : Trabajo de Investigación

12. Bibliografías.

Referencias

- [1] Eugene Isaacson y Herbert Bishop Keller. *Analysis Numerical Methods: Ecuaciones Diferenciales*. New York: Ediciones Dover.
- [2] Richard L. Burden y J. Douglas Faires. *Numerical Analysis: Problemas de Valor de Frontera y Metodos Numericos de Problemas de Valor de Frontera*. Boston: Ediciones Shaylin Walsh.
- [3] Ward Cheney y David Kincaid. *Numerical Mathematics and Computing: Metodos Numericos de Problemas de Valor de Frontera*.
- [4] David L. Powers. *Problemas de Valor de Frontera y ecuaciones Diferenciales Parciales: Metodos Numericos*. Estados Unidos de America: Ediciones Elsevier.
- [5] Fazal y Haq. *Solución Numérica de Valor de Frontera y Problemas de valor de Frontera iniciales usando función Spline (Paper): Problemas de Valor de Frontera*.
- [6] C. Henry Edwards y David E. Penney *Ecuaciones Diferenciales y Problemas con valor en la Frontera: Eigenvalores y Problemas con valor de Forntera*. Mexico: Ediciones Pearson.