

Resumen Análisis Numérico
Docente: Ricardo Antonio Prato Torres

Índice

1. Parcial 1	4
1.1. Teorema del Punto Fijo de Banach	4
1.2. Proceso del Punto Fijo	5
1.2.1. Cantidad de Iteraciones necesarias	6
1.2.2. Punto Fijo Atractivo	6
1.2.3. Punto Fijo Repulsivo	8
1.3. Método de Bisección	8
1.3.1. Método de Bisección: Teorema de Convergencia	9
1.4. Método de Regula-Falsi	10
1.5. Método de Newton-Rhapson	11
1.6. Proceso de Newton-Rhapson	11
1.7. Método de la Secante	12
1.8. Definición de error de convergencia	12
2. Parcial 2	13
2.1. Norma Matricial	13
2.1.1. Norma de Frobenius (Hilbert-Schmidt)	13
2.2. Tópicos de Teoría Espectral	15
2.2.1. Radio Espectral	15
2.3. Método de Jacobi	19
2.3.1. Condiciones para Jacobi	19
2.4. Método de Gauss-Seidel	22
2.5. Método Sobrerelajado de Jacobi (JOR)	23
2.6. Método Sobrerelajado de Gauss-Seidel (SOR)	23
2.7. Resultados para la Convergencia	24
2.8. Teorema del Punto Fijo n-dimensional	24
3. Parcial 3	28
3.1. Interpolación de Lagrange	28
3.2. Interpolación de Newton	33
3.2.1. Método de Horner	35
3.3. Calculo del error de Interpolación	36

3.4.	Nodos de Chebyshev	36
3.4.1.	Error con los Nodos de Chebyshev	39
3.5.	Integración Numérica	40
3.5.1.	Integración de Newton-Cotes	40
3.5.2.	Método del Rectángulo (1 nodo)	41
3.5.3.	Método del Trapecio (2 nodos)	43
3.5.4.	Método de Simpson (3 nodos)	44
3.5.5.	Cambio de Intervalo	45
3.5.6.	Verificación del Método del Rectángulo	46
3.5.7.	Verificación del Método del Trapecio	46
3.5.8.	Verificación del Método de Simpson	47
3.6.	Integración Numérica (Métodos Compuestos)	49
4.	Parcial 4	50
5.	Bibliografía sugerida	51

Nota al lector

Se realizan estas notas, para llevar un compendio entre lo visto en clase, y los ejercicios realizados en el curso a cargo del docente Ricardo Antonio Prato Torres, con el apoyo gráfico y de algoritmos realizados para mayor efectividad del curso y de las notas, es claro que este curso sin el apoyo computacional sería muy corto, puesto que nuestra capacidad de realizar todos los cálculos queda pequeña.

Dejaremos consignado que los códigos donde estamos realizando los ejemplos de algoritmos serán en Python, dado que es el programa que usa el docente, y el de acceso libre actualmente.

1. Parcial 1

1.1. Teorema del Punto Fijo de Banach

Durante todo el curso usaremos el concepto de Punto Fijo, por lo que estaremos realizando procesos partiendo de la existencia y unicidad de este punto fijo, por lo que empezaremos definiendo en primera instancia qué es un punto fijo.

Definición 1: Punto Fijo. Sea $g \in \mathcal{C}(I)$, luego, un punto fijo es aquel $p \in I$ tal que $p = g(p)$.

Demos también la definición de contracción, o función contractiva:

Definición 2: Contracción (Función Contractiva). Sea $g \in C(I)$, decimos que g es contractiva si existe $K \in (0, 1)$, tal que:

$$||g(x) - g(y)|| \leq K||x - y||$$

Teorema 1: Teorema del Punto Fijo de Banach. Sea (X, d) un espacio métrico completo y $f : X \rightarrow X$ una aplicación. Se dice que si f es contractiva entonces existe un único punto fijo en el intervalo.

Esto nos será muy importante en el caso de querer poder solucionar funciones del tipo:

$$f(x) := x - g(x) = 0$$

Veamos de igual modo los estimativos de error, dado que con estos nos basaremos en el curso para encontrar los errores de diferencia entre la aproximación y solución exacta.

$$||x_n - x|| \leq \frac{q^n}{1 - q} ||x_1 - x_0|| \quad (1)$$

Esta expresión es nuestra estimativo a-priori, veamos ahora otra:

$$||x_n - x|| \leq \frac{q}{1 - q} ||x_n - x_{n-1}|| \quad (2)$$

Esta segunda estimación se llama estimación a-posteriori, en el curso nos enfocaremos en la estimación a-priori, dado que la a-posteriori es más de un uso avanzado, y estamos en un curso básico de Análisis Numérico, por lo que estamos fundamentando lo que vamos a realizar.

Cuándo estemos buscando el demostrar que una función es contractiva en un intervalo, generalmente será un problema grande encontrar esa q , por lo que usaremos una expresión equivalente a probar que una función es contracción:

1. $\forall x \in Domg \ (Rang \subseteq Domg)$
2. $\forall x \in Domg \ (|g'(x)| < 1)$

La primera condición nos asegura existencia del punto fijo, y la segunda nos asegura la unicidad del punto fijo.

1.2. Proceso del Punto Fijo

Tenemos la función $f(x) = x^3 - 4$, veamos dónde se hace 0, es claro que el lugar dónde se hace 0 es $\sqrt[3]{4}$, pero veremos como lo aproxima una reestructuración de la función, de la forma de punto fijo:

Es claro que esta función reestructurada es equivalente a:

$$x = \frac{2}{\sqrt{x}}$$

Por lo que procederemos a dar un intervalo dónde funcione el método, como primera regla debemos considerar un intervalo donde este punto de aproximación se encuentre, es decir dónde $\sqrt[3]{4} \in I$, por lo que consideremos el intervalo $I := [5/4, 2]$, ahora procedamos a probar que existe ese único punto fijo en el intervalo dado:

Para analizar la parte del rango, debemos analizar su derivada, por lo que obtenemos:

$$g'(x) = -\frac{1}{x^{3/2}}; g''(x) = \frac{3}{2x^{5/2}}$$

Es claro que la función g no tiene Puntos Críticos¹, y además es decreciente, por lo que bastará por fijarnos en sus extremos, así tendremos que:

$$Rang = \left[g(b) = \sqrt{2}, g(a) = \frac{4}{5}\sqrt{5} \right] \subseteq \left[\frac{5}{4}, 2 \right]$$

Así hemos garantizado la existencia del punto fijo. Veamos la segunda condición:

Podemos notar que necesitamos que todos los valores de x cumplan que al evaluarlos en la derivada sean menores que 1 (bajo valor absoluto), por lo que analizaremos la norma del máximo, es decir nos fijaremos en el mayor valor posible, dado que sí cumple el mayor valor bajo valor absoluto, cumplirá el resto. Por lo que procederemos a fijarnos en la segunda derivada, y de esta concluimos que no hay Puntos Críticos, por lo que bastará con fijarnos en los extremos.

$$\|g'\|_{\infty} = |g'(a)| \text{ ó } |g'(b)|$$

¹Basta fijarse en g' , y que nunca se hace 0.

$$|g'(a)| = |g'(5/4)| = 0,715542; |g'(b)| = |g'(2)| = 0,353553$$

De este modo es claro que $\|g'\|_\infty = 0,715542 = q < 1$, es decir aquí estamos garantizando que mi punto fijo es único, cosa que nos permite dar una mejor aproximación basándonos en el método del punto fijo es decir:

$$x_{n+1} = g(x_n)$$

Así iterando varias veces llegamos a aproximar el número que deseamos, con la precisión que deseamos, esto mediante la siguiente fórmula que genera una cota inferior teórica para la cantidad de iteraciones:

1.2.1. Cantidad de Iteraciones necesarias

Muchas veces queremos saber cuántas iteraciones necesitamos hacer para establecer una cota teórica²

$$n = \frac{\ln \left[\frac{\xi(1-q)}{\|x_1 - x_0\|} \right]}{\ln q} \quad (3)$$

Dónde ξ se refiere a la estimación deseada que sale de la estimación a-priori, es decir $\xi = \frac{q^n}{1-q} \|x_1 - x_0\|$, y es justo de esta relación de dónde nace la fórmula (3), es usar un manejo algebraico adecuado, y podemos llegar a la fórmula teórica de la estimación de la cota inferior de n , existe el caso dónde nos pidan hallar la cota superior de la aproximación para una iteración k del proceso, en ese caso usaremos la fórmula de estimación a-priori y lo calcularemos, es decir:

$$\text{Cota Superior} = \frac{q^k}{1-q} \|x_1 - x_0\|$$

Luego de ver como es el método de punto fijo, daremos dos clasificaciones del punto fijo, las cuáles son el punto atractivo, y el punto repulsivo.

1.2.2. Punto Fijo Atractivo

Realmente las condiciones necesarias para esto es lo que ya hemos venido haciendo, es decir que cumpla la segunda condición de que $|g'(x)| < 1$, aquí diremos que la iteración de punto fijo **converge** al punto fijo p , y podemos establecer mediante el *Diagrama de Cobweb* como va siendo la convergencia del método.

²Digo teórica dado que al final resulta que se necesitan menos iteraciones de las que la fórmula dice.

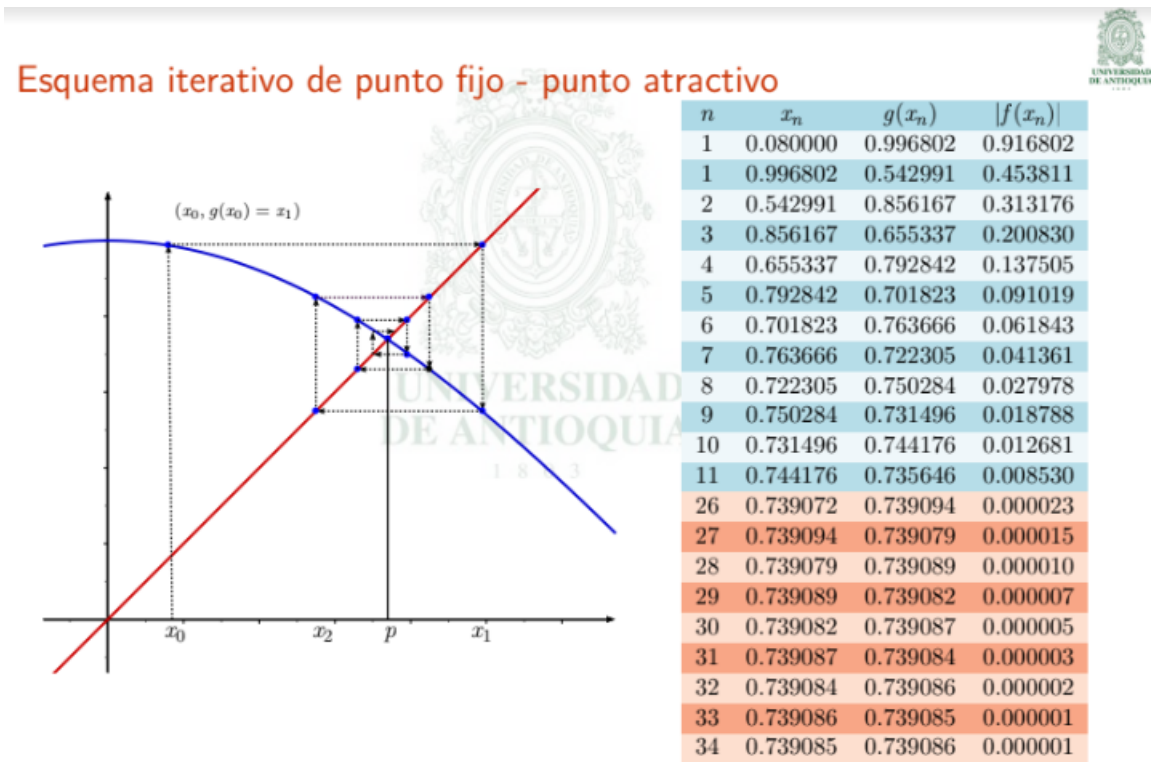


Figura 1: Esquema Iterativo de Punto Fijo (Punto Atractivo) Tomado de las Diapositivas del profesor Ricardo Prato.

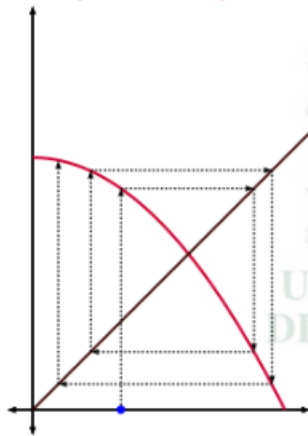
Esto nos muestra de manera gráfica como se evidencia la convergencia del método de punto fijo, mostrando que al hacer una elección adecuada del intervalo la convergencia parece ser rápida³ y llegar en un número mucho menor de pasos, dado que a pesar de que tengamos una cota mínima de convergencia, expresada por el número n en la fórmula (3), se ha observado experimentalmente que el método converge en muchas menos iteraciones de las determinadas allí, haciendo que la efectividad sea mucho mejor, es bueno aclarar que sí el método no va convergiendo en medida se acerca al número n , debemos preocuparnos en que realmente cinverja, lo que da paso al punto fijo repulsivo.

³Depende de la expresión de punto fijo, dado que existen infinitas funciones para expresar la convergencia.

1.2.3. Punto Fijo Repulsivo

En esta parte observamos que existe la posibilidad de que $|g'(x)| > 1$, es decir que no cumpla la condición (2) que nos implica unicidad, por lo que cuándo ocurra esto, se dirá que p es un punto fijo repulsivo, y la iteración en lugar irse acercando, se irá separando, y divergerá, como es el siguiente caso:

Punto fijo (repulsivo)



n	x_n	$g(x_n)$	$ f(x_n) $
0	0.350000	0.877500	0.527500
1	0.877500	0.229994	0.647506
2	0.229994	0.947103	0.717109
3	0.947103	0.102996	0.844107
4	0.102996	0.989392	0.886396
5	0.989392	0.021104	0.968288
6	0.021104	0.999555	0.978451
7	0.999555	0.000891	0.998664
8	0.000891	0.999999	0.999109
9	0.999999	0.000002	0.999998
10	0.000002	1.000000	0.999998
11	1.000000	0.000000	1.000000

Tiene punto fijo, pero no satisface el Teorema de punto fijo. El esquema iterativo de punto fijo no converge.

Figura 2: Esquema Iterativo de Punto Fijo (Punto Repulsivo) Tomado de las Diapositivas del profesor Ricardo Prato.

1.3. Método de Bisección

Veamos uno de los métodos más antiguos, el cuál es el método de bisección.

Sea f continua en $[a, b]$ y $f(a)f(b) < 0$, entonces por el TVI existe x^* tal que $f(x^*) = 0$.

Luego nuestro método parte de que:

$$C_0 = \frac{a+b}{2}$$

De esto se generan varios casos:

1. Caso 1: $f(C_0) = 0 \therefore x^* = C_0$.
2. Caso 2: $f(C_0) > 0$, si asumimos $f(a) > 0, f(b) < 0 \therefore f(C_0)f(b) < 0 \Rightarrow x^* \in (C_0, b)$.
3. Caso 3: $f(C_0) < 0$, si asumimos $f(a) > 0, f(b) < 0 \therefore f(C_0)f(a) < 0 \Rightarrow x^* \in (a, C_0)$.

En el primer caso, el método queda ahí, dado que hallamos la raíz, en los otros dos hay que seguir iterando, es decir, con el nuevo intervalo que generamos, ya sea $(a, C_0); (C_0, b)$ volvemos a analizar si el punto medio es decir:

$$C_1 = \frac{C_0 + a}{2}; C_1 = \frac{C_0 + b}{2}$$

Y realizamos nuevamente el método iterativo de ir seleccionando los intervalos, hasta que el número C_k , tenga la aproximación deseada, de este modo podemos generar otro método iterativo, mediante la aproximación y partición de intervalos encajados⁴ dado que sabemos que estos métodos al llevarlos al infinito convergen en su totalidad, cosa que un computador no puede realizar, he ahí la diferencia entre el Análisis Matemático y el Numérico, buscando traer a lo finito esa infinitud del Análisis.

1.3.1. Método de Bisección: Teorema de Convergencia

Sea $f \in C([a, b])$ con $f(a), f(b)$ de signos contrarios. Sea $\{c_n\}_{n=0}^{\infty}$ la sucesión de puntos medios de los intervalos generados por el método de bisección. Entonces existe un número $r \in [a, b]$ tal que $f(r) = 0$, y

$$|r - c_n| \leq \frac{b-a}{2^{n+1}} \quad \text{para } n = 0, 1, \dots \quad (4)$$

En particular tenemos:

$$\lim_{n \rightarrow \infty} c_n = r$$

Llamemos de este modo al $(b-a)/2^{n+1}$ como nuestra tolerancia, es decir ξ , así podemos despejar de este modo:

⁴Recordemos que un intervalo encajado, es aquel que cumple la siguiente condición: $I_1 \supset I_2 \supset \dots \supset I_k \supset \dots$.

$$\frac{b-a}{2^{n+1}} = \xi \Rightarrow n = \frac{\ln\left(\frac{b-a}{\xi}\right)}{\ln 2} - 1 \quad (5)$$

Quedando así determinado el número n que significa la cota inferior de convergencia de acuerdo a la tolerancia ξ .

1.4. Método de Regula-Falsi

Cabe el caso de que existan funciones dónde las raíces estén demasiado cerca, para estos casos el método de bisección al ser tan arcaico, no es demasiado confiable, por lo que daremos un nuevo método llamado *Regula-Falsi*, para apoyarnos en estos casos.

Inicialmente nos fijamos en unos puntos a, b tales que el punto dónde se hace 0 la función pertenezca, así calculemos la pendiente de esta secante, la cuál termina en:

$$m_s = \frac{f(b) - f(a)}{b - a}$$

Para obtener así:

$$y - f(a) = \frac{f(b) - f(a)}{b - a}(x - a)$$

Tomemos de este modo el punto $(C_0, 0)$, luego nos queda la función y como:

$$C_0 = a - f(a) \cdot \frac{b - a}{f(b) - f(a)} \quad (6)$$

Luego evaluamos $f(C_0)$ y se compara.

- Si $f(C_0) \cdot f(a) < 0$, el intervalo a considerar es $[a, C_0]$.
- Si $f(C_0) \cdot f(b) < 0$, el intervalo a considerar es $[C_0, b]$.
- Si $f(C_0) = 0$, se concluye que la raíz es $x = C_0$.

El proceso continua en subintervalos

$[C_p, C_q]$, y se detiene si

- $|C_{n+1} - C_n| < \xi$
- $|C_{n+1} - x^*| < \xi$
- $|f(C_{n+1})| < \xi$

Todos estos son conocidos como los criterios de parada

1.5. Método de Newton-Rhapson

El método de Newton (1D Newton-Rhapson) es un método que nos permite ser aplicado en varias dimensiones, lo mismo que Punto Fijo⁵, de este modo damos el método:

Sabemos que $m_t = f'(x_0)$, por lo que tendremos: $l_t : y - f(x_0) = f'(x_0)(x - x_0)$, lo cuál implica la linealización de f , así aplicando regula-falsi, calculamos $(x_1, 0)$:

$$-f(x_0) = f'(x_0)(x_1 - x_0) \Rightarrow x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \quad (7)$$

Construimos así una recta tangente en x_1 y el proceso continua, por lo que obtenemos:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = g(x_n) \quad (8)$$

Por lo que nuestro problema se reduce a un problema de punto fijo.

Mientras que los métodos pasados tenían una velocidad de convergencia de tipo lineal (incluso bisección es menor que lineal) Newton-Rhapson posee una velocidad de convergencia cuadrática, es decir que converge al punto a la velocidad que x^2 converge a 0.

1.6. Proceso de Newton-Rhapson

Resolvamos $f(x) = 0$ con $f(x) = a - \frac{1}{x}$, $a > 0$.

Sabemos que para poder aplicar el método de Newton por lo menos debemos encontrar un dominio D abierto en donde f sea continuamente diferenciable. Tenemos

$$f(x) = (x^2 - 2) \cos\left(\frac{\pi}{2}(x^2 - 3)\right)$$

$$f'(x) = 2x \cos\left(\frac{\pi(x^2 - 3)}{2}\right) - \pi x (x^2 - 2) \sin\left(\frac{\pi(x^2 - 3)}{2}\right)$$

A pesar de la efectividad de Newton-Rhapson, existe el problema con las funciones que no tienen derivada, es decir con aquellas que son $f \in C^6$

⁵El resto de métodos no se pueden llegar a aplicar, dado que en \mathbb{R}^n el concepto de orden no es posible establecerlo (de forma tan directa) en estos espacios.

⁶Es decir solo la función es continua, cuándo la derivada posee problemas.

1.7. Método de la Secante

Cuándo tengamos el problema descrito con anterioridad, lo más efectivo será coger dos puntos, y aproximar la derivada como si fuera un estilo de secante, por lo que el esquema de Secante nos quedaría muy parecido a Newton-Rhapson, solo con el siguiente detalle:

$$x_{i+2} = x_i - f(x_i) \frac{x_{i+1} - x_i}{f(x_{i+1}) - f(x_i)} \quad (9)$$

Siendo el paso base:

$$x_2 = x_0 - f(x_0) \frac{x_1 - x_0}{f(x_1) - f(x_0)}$$

Podemos tomar dos puntos $(x_0, f(x_0)); (x_1, f(x_1))$ cualesquiera, y veremos que el método sigue aproximando nuestra solución, la potencia de este método es el hecho de que usamos funciones de clase $f \in C(I)$, es decir funciones continuas, pagando el hecho de que necesitamos un parámetro extra para poder calcular la secante.

1.8. Definición de error de convergencia

Sea e_i el error después del paso i de un método iterativo (¡cualquier método iterativo!), diremos que el orden del error de convergencia será $\mathcal{O}(h^n)$, sí se cumple que:

$$M = \lim_{i \rightarrow \infty} \frac{e_{i+1}}{e_i^n} < \infty \quad (10)$$

Más aún tenemos que se cumple que:

Teorema de Convergencia de Newton-Rhapson. *El método de Newton-Rhapson, cumple la convergencia cuadrática, es decir:*

$$M = \lim_{i \rightarrow \infty} \frac{e_{i+1}}{e_i^2} < \infty \quad (11)$$

Es por esto que Newton-Rhapson es el que llega más rápido a la solución que cualquier otro método.

Entremos ahora al campo de la Introducción al Álgebra Lineal Numérica, definiendo normas matriciales.

2. Parcial 2

2.1. Norma Matricial

Recordemos que en \mathbb{R}^n , la norma del máximo y la norma euclidiana son equivalentes. Recordemos que:

$$\vec{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$$

$$\|\vec{x}\|_2^2 := \sum_{i=1}^n x_i^2$$

$$\|x\|_\infty := \max_{1 \leq i \leq n} |x_i|$$

Luego ser equivalentes significa que existen constantes c_1, c_2 tales que:

$$c_1 \|x\|_b \leq \|x\|_a \leq \|x\|_b$$

Con $a, b = 2, \infty$.

Luego, la aplicación:

$$\|\cdot\| : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$$

define una norma matricial si y solo si

- $\|A\| \geq 0$, para toda $A \in \mathbb{R}^{m \times n}$, y $\|A\| = 0 \Leftrightarrow A = 0$.
- $\|\alpha A\| = |\alpha| \|A\|$, para toda $\alpha \in \mathbb{R}, A \in \mathbb{R}^{m \times n}$.
- $\|A + B\| \leq \|A\| + \|B\|$, para toda $A, B \in \mathbb{R}^{m \times n}$.

Definamos además:

2.1.1. Norma de Frobenius (Hilbert-Schmidt)

Definiremos una nueva norma que nos permitirá realizar mejores cálculos:

$$\|A\|_F^2 = \sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2 = \text{tr}(AA^*) \quad (12)$$

Sea $A = (a_{ij}); i, j = 1, \dots, n$, y $\vec{x} = (x_1, \dots, x_n)$, luego definimos:

$$\|Ax\|_2^2 := \sum_{i=1}^n \left| \sum_{j=1}^n a_{ij}x_j \right|^2 \leq \sum_{i=1}^n \left(\sum_{j=1}^n |a_{ij}|^2 \sum_{j=1}^n |x_j|^2 \right) = \|A\|_F^2 \|x\|_F^2 \quad (13)$$

Es decir que la norma 2 va a estar acotada superiormente por la normas de Frobenius.

Aclaremos que $\|Ax\|_2$ es la norma del vector, más no de la matriz **¡Importante!**

Definamos también lo que usaremos como la Norma 2:

$$\|A\|_2 = \sqrt{\rho(A^*A)}$$

Es decir, la norma 2 es la raíz cuadrada del módulo del valor propio más grande de la matriz A^*A .

Teorema. Sea $\|\cdot\|$ una norma vectorial en \mathbb{R}^n . La función (estrictamente se llama operador o mucho más formal transformación):

$$\|A\| := \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

es una norma matricial, la cuál se dice **norma matricial inducida ó norma matricial natural**.

Es decir una norma matricial se amarra siempre a una norma vectorial.

Más aún tendremos:

$$\|A\| := \sup_{\|x\|=1} \|Ax\| \quad (14)$$

Veamos $\|A\|_1 = \max_{k=1, \dots, n} \sum_{j=1}^n |a_{jk}|$, es **El máximo de las sumas de las columnas**, es decir, voy

sumando con valor absoluto los elementos, por cada columna, y al final comparo estas sumas.

Veamos ahora que $\|A\|_\infty = \max_{j=1, \dots, n} \sum_{k=1}^n |a_{jk}|$, es **El máximo de las sumas de las filas**, es decir, voy

sumando en valor absoluto los elementos de cada fila, y al final comparo estas sumas.

Recordemos como se define $\|x\|_p$:

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}$$

Se puede demostrar que (de una manera no tan difícil):

$$\|Ax\|_1 \leq \|A\|_1 \|x\|_1$$

Y de una forma un poco más dura y de pensar que:

$$\|Ax\|_1 \geq \|A\|_1 \|x\|_1$$

Y queda la igualdad⁷. \square

2.2. Tópicos de Teoría Espectral

Recordemos que $\lambda \in \mathbb{C}$ se dice **valor propio** de A si y solo si existe un vector $x \in \mathbb{C} - \{0\}$ tal que

$$Ax = \lambda x$$

x se dice **vector propio** (asociado) del valor propio λ . El conjunto de todos los valores propios de la matriz A se dice **espectro** de A , y se denota $\sigma(A)$.

2.2.1. Radio Espectral

El radio espectral de la matriz A notado por $\rho(A)$ se define como:

$$\rho(A) := \max_{\lambda \in \sigma(A)} |\lambda| \quad (15)$$

Además tenemos la siguiente igualdad:

Teorema. Si $A \in \mathbb{C}^{n \times n}$, entonces $\|A\|_2 = \sqrt{\rho(AA^*)}$, y que si A es simétrica se da que $\|A\|_2 = \rho(A)$.

Recordemos que la solución del polinomio característico es λ si:

$$(1) \quad p_A(\lambda) = \det(A - \lambda I) = 0 \quad (16)$$

$$(2) \quad \det(A) = \prod_{i=1}^n \lambda_i; \quad \text{tr}(A) = \sum_{i=1}^n \lambda_i \quad (17)$$

⁷No incluyo la demostración por ser muy larga.

⁸Recordemos que el nombre/expresión de $\|\cdot\|$ hace referencia a la norma, en el caso de \mathbb{R} es el valor absoluto, y en el caso de \mathbb{C} es el módulo.

Lema. Sea $A \in \mathbb{K}^{n \times n}$, se define el siguiente producto interno:

$$(Ax, y) := \bar{y}^T Ax \in \mathbb{R}$$

en \mathbb{R}^n asociado a A .

Además tendremos que:

Si $A \in \mathbb{K}^{n \times n}$ se obtiene que:

$$A^* = (\bar{A})^T = \overline{(A^T)}, \text{ así obtenemos que: } (Ax, y) = (x, A^*y) \quad (18)$$

Si $A = (a_{ij})$, con $i, j = 1, \dots, n$; $\vec{x} = (x_1, x_2, \dots, x_n)$, $\vec{y} = (y_1, \dots, y_n)$

$$(Ax, y) = \sum_{j=1}^n \bar{y}_j^T (Ax)_j; (Ax)_j = A^{(j)}x = \sum_{i=1}^n a_{ji}x_i$$

$$(A\vec{x}, y) = \sum_{j=1}^n (Ax)_j \bar{y}_j = \sum_{j=1}^n \left(\sum_{i=1}^n a_{ji}x_i \right) \bar{y}_j$$

Consideremos los siguientes teoremas:

Lema (Kress, Th. 3.27). Para cada matriz A existe una matriz unitaria Q tal que Q^*AQ es triangular superior.

Teorema (Kress, Th. 3.29). Los valores de una matriz hermitiana $A \in \mathbb{C}^{n \times n}$ son reales y los valores propios forman una base ortogonal de \mathbb{C}^n .

Ahora el siguiente teorema nos dará una idea de lo que estaremos realizando de aquí en adelante para los métodos iterativos:

Teorema. Suponga que \tilde{x} es una aproximación de la solución $Ax = b$, donde A es una matriz no singular, y definimos: $r := b - Ax$ es el vector residual para \tilde{x} . Entonces para cualquier norma inducida, si $\|x\|, \|b\| \neq 0$ tendremos que:

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq \|A\| \cdot \|A^{-1}\| \frac{\|r\|}{\|b\|} \quad (19)$$

Realicemos una definición que será importante también para construir la teoría que estamos realizando, puede que esta parte aprezca muy teórica, pero es necesaria para poder definir correctamente los métodos de iteración en \mathbb{R}^n .

Definición de Condición de una Matriz. La condición de una matriz no singular relativa a una norma cualquiera $\|\cdot\|$ es:

$$\kappa(A) := \|A\| \cdot \|A^{-1}\|$$

Y tendremos las siguientes condiciones:

- Una matriz se dice bien condicionada si se da que $\kappa(A)$ es muy cercana a 1, o se dice mal condicionada si se da que $\kappa(A) \gg 1$ ⁹
- La condición de una matriz es una medida de la sensibilidad de un sistema de ecuaciones lineales a pequeños cambios en los datos del problema.
- Una matriz bien condicionada tiene una solución relativamente insensible a pequeños cambios, mientras que una matriz mal condicionada puede tener soluciones muy sensibles a pequeños cambios.

Definamos ahora conceptos de Álgebra Lineal:

Matriz Simétrica Definida Positiva. Las siguientes son equivalencias de una matriz simétrica definida positiva, si se cumple diremos que A es una matriz simétrica definida positiva:

1. La aplicación $(x, y) := x^T A y$ define un producto interno.
2. Todos los determinantes de las submatrices de A compuesta por las primeras i filas y columnas de A son positivos. **Esta es mi favorita, dado que es muy simple de realizar, el único inconveniente es cuándo se eleva el número n , y sacar los determinantes se vuelve un poco inviable sin remitirse a programación.**

La matriz A se puede expresar como $A = U^T U$, donde U es una matriz triangular superior con entradas positivas en la diagonal principal, además que $U^T U$ es una factorización única, a esto lo llamaremos **Factorización de Cholesky**.

Definamos ahora algo muy importante¹⁰ para el método iterativo, el cuál es el siguiente teorema:

⁹Recordemos que esta notación se usa mucho en Física para indicar que algo es mucho más pequeño que otra cosa, como por ejemplo, si $\theta \ll 1$ se da que $\sin \theta = \theta$, cosa que en Física tiene muchas aplicaciones.

¹⁰La importancia de esto cae en que cuándo nos quedamos sin argumentos, este es el argumento que nos puede salvar en el método iterativo, daremos una parte aquí, y cuándo definamos el método lo terminaremos de dar.

Teorema. Sea $A \in \mathbb{R}^{n \times n}$, y $\|\cdot\|$ una norma sobre $\mathbb{R}^{n \times n}$, entonces:

$$\rho(A) \leq \|A\| \quad (20)$$

Además para cada matriz A y para cada $\varepsilon > 0$ existe una norma sobre \mathbb{C}^n tal que

$$\|A\| \leq \rho(A) + \varepsilon$$

Definamos otro Lema que nos será muy útil:

Lema. Sea $B: \mathbb{K} \rightarrow \mathbb{K}$ una transformación lineal sobre \mathbb{K} , donde $\mathbb{K} = \mathbb{R}$, $\mathbb{K} = \mathbb{C}$ con $\|B\| < 1$, y sea I la transformación identidad. Entonces:

$$I - B$$

Es invertible, y para cada $z \in \mathbb{K}$ la ecuación

$$(I - B)x = x - Bx = z$$

Tiene solución única

Definiendo que existe esa solución única podemos entrar en el método como tal, hay que aclarar lo siguiente:

Las definiciones, teoremas, lemas que hemos mencionado no se anexa la demostración, y todos son necesarios para poder asegurar la siguiente existencia que tendremos, dado que sin eso no podríamos afirmar lo siguiente, por lo que es normal que sea tediosa esta parte por tanta teoría que tiene.

Procedamos a enunciar otro teorema acerca de los métodos iterativos:

Teorema de Convergencia. Sea $B \in \mathbb{C}^{n \times n}$. Las sucesivas aproximaciones

$$x_{v+1} := Bx_v + z, \quad v \in \mathbb{N}_0 \quad (21)$$

Converge para cada $z \in \mathbb{C}^n$ y cada $x_0 \in \mathbb{C}^n$ si y solo si

$$\rho(B) < 1$$

Si notamos esto es muy poderoso porque ya nos va diciendo que la iteración va a converger si se da esa característica con el radio espectral, ahora definido eso, podemos plantear nuestro primer método iterativo:

2.3. Método de Jacobi

Método Iterativo de Jacobi. Para la resolución de Sistemas de la forma $Ax = b$, tendremos la siguiente iteración:

$$x_{n+1} = -D^{-1}(A_R + A_L)x_n + D^{-1}b \quad (22)$$

que converge a la solución x que deseamos hallar.

Demostración. Más que una prueba será la forma de determinar este método, en virtud de los teoremas que ya probamos:

Inicialmente tenemos que A debe ser una matriz no singular (porque de serlo el método no sería efectivo), y podemos dividirla en las siguientes:

$$A = D + A_R + A_L$$

Dónde D hace referencia a la diagonal, y A_R, A_L son las matrices triangulares superior, y triangular inferior respectivamente, luego podremos ejecutar esto:

$$Ax = b \Leftrightarrow (D + A_R + A_L)x = b$$

$$Dx = -(A_R + A_L)x + b$$

$$x = -D^{-1}(A_R + A_L)x + D^{-1}b$$

Y en virtud del teorema anterior, concluimos la parte iterativa, es decir:

$$x_{n+1} = -D^{-1}(A_R + A_L)x_n + D^{-1}b$$

□

2.3.1. Condiciones para Jacobi

Nosotros para poder usar el método de Jacobi debemos de asegurar que alguna de las normas es menor que 1, ¿Cuál es el problema? Que no necesariamente las 4 normas definidas son las que den que la norma sea menor que 1, pues están son muy generales, pero puede existir alguna dentro de esas infinitas normas que cumplan lo pedido, para estos casos como hemos mencionado anteriormente nuestra última herramienta se llama *Radio Espectral*, la forma de usarlo va de la mano con el teorema de convergencia, es decir, tendremos que Jacobi converge si y solo si $\rho(A) < 1$, y si es mayor que 1, diremos que el método no converge a la solución.

Demos un ejemplo de este método para entender la iteración:

Ejemplo 1. Consideremos el sistema $Ax = b$, dónde tenemos que:

$$A = \begin{pmatrix} 4 & 7 & 8 \\ 7 & 6 & 8 \\ 5 & 7 & 2 \end{pmatrix} \quad b = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}$$

Verifiquemos las condiciones de Jacobi, es decir vamos a ver si se da que la matriz $D^{-1}(A_R + A_L)$ evaluada en alguna norma cumple ser menor que 1:

Paso 1: Dar la estructura de la matriz

$$B_J = -D^{-1}(A_R + A_L) = - \begin{pmatrix} \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{6} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 7 & 8 \\ 7 & 0 & 8 \\ 5 & 7 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\frac{7}{4} & -2 \\ -\frac{7}{6} & 0 & -\frac{4}{3} \\ -\frac{5}{2} & -\frac{7}{2} & 0 \end{pmatrix}$$

Paso 2: Sacar las normas matriciales

Ya que conocemos como es la matriz de iteración de Jacobi, veamos que ocurre con las normas matriciales:

Norma 1: Recordemos que la Norma 1 para matrices es sacar el máximo de la suma de las columnas, en este caso tenemos:

$$\|B_J\|_1 = \max \left\{ \frac{22}{6}, \frac{21}{4}, \frac{10}{3} \right\} = \frac{21}{4} > 1$$

Norma ∞ : Recordemos que la Norma infinito para matrices es sacar el máximo de la suma de las filas, en este caso tenemos:

$$\|B_J\|_\infty = \max \left\{ \frac{15}{4}, \frac{15}{6}, 6 \right\} = 6 > 1$$

Norma Frobenius: Si recordamos esta se saca sumando el cuadrado de las componentes, lo qué nos da:

$$\|B_J\|_F = \frac{49}{16} + 4 + \frac{49}{36} + \frac{16}{9} + \frac{25}{4} + \frac{49}{4} = \frac{875}{48} > 1$$

Para la norma dos se debe considerar solo la transpuesta, dado que todas las entradas son números reales, queda de tarea realizar la norma 2.

Nos da de resultado que:

$$\|B_J\|_2 = \sqrt{15,289} > 1$$

Si notamos, todas las normas que conocemos nos da que son mayores que 1, ahora nos queda de última herramienta verificar el radio espectral de B_J :

Paso 3 (Opcional)¹¹: Sacar las normas matriciales

Por el curso de Álgebra Lineal sacamos los valores propios, los cuáles son:

$$\rho(B_J) = \max\{|\lambda| : \lambda \text{ es un valor propio}\} = \max\{3,9101, 1,4634, 2,4467\} = 3,9101 > 1$$

Dado que el radio espectral da mayor que 1, concluimos rotundamente que

El Método no Converge a la Solución, pero si resolvemos el sistema, nos da que su solución x es:

$$x = \begin{pmatrix} \frac{39}{79} \\ \frac{35}{79} \\ -\frac{20}{79} \end{pmatrix}$$

Demos una equivalencia con la Norma infinito:

$$q_\infty < 1 \Leftrightarrow A \text{ es estrictamente diagonal dominante por filas.} \quad (23)$$

¿Qué significa es? Demos un ejemplo con una matriz para entenderlo:

$$\begin{pmatrix} 4 & -1 & 1 \\ 4 & -8 & 1 \\ -2 & 1 & 5 \end{pmatrix}$$

Esta matriz es estrictamente diagonal dominante por filas pues:

$$|1| + |1| < |4|; |4| + |1| < |-8|; |-2| + |1| < |5|$$

Pero Jacobi no es el único método disponible, también existe el *Método de Gauss- Seidel*.

¹¹Digo opcional, dado que si alguna norma nos da menor que 1, podemos concluir de inmediato que es convergente el método.

2.4. Método de Gauss-Seidel

Para el método de Gauss-Seidel tendremos que realizar esta modificación:

Método Iterativo de Gauss-Seidel. Para la resolución de Sistemas de la forma $Ax = b$, tendremos la siguiente iteración:

$$(D + A_L)x_{n+1} = -A_Rx_n + b \quad (24)$$

que converge a la solución x que deseamos hallar.

Demostración. Así como para Jacobi, aquí daremos el bosquejo del método y de la forma de llegar al proceso:

Tenemos un problema distinto aquí, dado que cabe la posibilidad de que la matriz $D + A_L$ no sea invertible, por lo que no siempre se podrá dar esa inversión tan alegremente, por lo que proponemos que quede de esta forma:

$$Ax = b \Leftrightarrow (D + A_R + A_L)x = b$$

$$(D + A_L)x = -A_Rx + b$$

Aquí se diferencian dos partes, dónde si esa matriz es no invertible, aquí termina el método, pero en caso de ser invertible se da que:

$$x = -(D + A_L)^{-1}A_Rx + (D + A_L)^{-1}b$$

Y en virtud del teorema anteriormente mencionado, concluimos la parte iterativa, es decir:

$$x_{n+1} = -(D + A_L)^{-1}A_Rx_n + (D + A_L)^{-1}b$$

Esto para el caso de invertibilidad. En el caso de que no sea invertible sería:

$$(D + A_L)x_{n+1} = -A_Rx_n + b$$

□

Ejemplo 2. Realizar el ejemplo anterior mediante el método de Gauss-Seidel.

Cabe aclarar que las condiciones de Gauss-Seidel se mantienen como las de Jacobi. Solo que está vez estaremos considerando la matriz a la que le sacaremos todas las normas como:

$$B_{GS} = -(D + A_L)^{-1}A_R$$

Pero a pesar de estos dos métodos podemos definir otros dos métodos basándonos en estos que ya pasaron

2.5. Método Sobrerelajado de Jacobi (JOR)

Daremos un cambio estructural a la fórmula de Jacobi, dónde usaremos lo siguiente:

Muchas veces queremos ver cuándo nuestro método puede converger más rápido, para esto haremos un producto de un factor ω que nos permitirá eso, determinemos ese factor:

$$x_{n+1} = x_n + \omega D^{-1}b - D^{-1}(A_R + A_L)x_n$$

Luego mi matriz de iteración (llamemosla B_ω) estaría definida por:

$$B_\omega := I + \omega D^{-1}(-A) = (1 - \omega)I - \omega D^{-1}(A_R + A_L) = (1 - \omega)I + \omega B; \quad B := -D^{-1}(A_R + A_L)$$

Ahora, un dato importante es que sí los valores propios de la matriz son reales, podemos encontrar el mejor y óptimo valor de ω :

$$\omega_{\text{opt}} = \frac{2}{2 - \lambda_{\max} - \lambda_{\min}} \quad (25)$$

Veamos el caso de Gauss-Seidel.

2.6. Método Sobrerelajado de Gauss-Seidel (SOR)

Para Gauss-Seidel también tendremos modificaciones, las cuáles serían:

$$x_{n+1} = x_n + \omega D^{-1}(b - A_L x_{n+1} - (D + A_R)x_n) \Leftrightarrow x_{n+1} = \omega D^{-1}(b - A_L x_{n+1} - A_R x_n) + (1 - \omega)x_n \quad (26)$$

Luego tendremos que la matriz de iteración definida por B_{SOR} es:

$$B_{SOR} = (D + \omega A_L)^{-1}((1 - \omega)D - \omega A_R)$$

Y el método iterativo nos queda en:

$$x_{n+1} = B_{SOR}x_n + \omega(D + \omega A_L)^{-1}b \quad (27)$$

Se deja de tarea realizar los métodos con la matriz anterior.

Condiciones Necesarias para SOR. Si tenemos que el esquema iterativo de Gauss-Seidel:

$$x_{n+1} = x_n + \omega D^{-1}(b - A_L x_{n+1} - (D + A_R)x_n)$$

converge entonces el valor de $\omega \in (0, 2)$

Se deja de tarea realizar los códigos en Python de los métodos SOR y JOR.

2.7. Resultados para la Convergencia

Dentro de todo este desarrollo, podemos dar los siguientes resultados para la convergencia de los métodos.

1. $\rho(D^{-1}(A_L + A_R)) < 1 \Leftrightarrow$ Jacobi converge.
2. $\rho((D + A_L)^{-1}A_R) < 1 \Leftrightarrow$ Gauss-Seidel converge.
3. $\rho(I - \omega D^{-1}A) < 1 \Leftrightarrow$ JOR converge.
4. Si A es diagonalmente dominante por filas entonces los métodos Jacobi, Gauss-Seidel convergen.
5. Si A es diagonalmente dominante por columnas entonces el método de Jacobi converge.
6. Si A es simétrica y definida positiva entonces el método de Gauss-Seidel converge.
7. Si A es tridiagonal entonces $\rho(B_{GS}) = \rho^2(B_J)$. Se puede concluir en este caso que ambos convergen, o ambos divergen.
8. Si el método de Jacobi converge, entonces JOR converge para $0 < \omega \leq 1$.
9. Si A es simétrica definida positiva entonces JOR converge si $0 < \omega < \frac{2}{\rho(D^{-1}A)}$.

Veamos ahora como se realiza punto fijo pero en \mathbb{R}^n .

2.8. Teorema del Punto Fijo n-dimensional

Tendremos que muchas veces nuestros sistemas de solución son de carácter no lineal, por lo que los métodos anteriormente mencionados empiezan a fallar, así que para estas ocasiones definiremos un nuevo modelo de solución, basado en algo que ya habíamos hecho, y es el modelo de punto fijo.

Recordemos que punto fijo nos exige que para su existencia y unicidad, el rango esté contenido en el dominio, y que el valor absoluto de la derivada sea menor que 1, extendamos estas definiciones.

Sea $F(x_1, \dots, x_n) := (f_i(x_1, \dots, x_n))_{i=1}^n$, con $f_i(x_1, \dots, x_n) : \mathbb{R}^n \rightarrow \mathbb{R}$, aquí nuestro problema es encontrar un elemento $x^* \in \text{Dom } F$ tal que $F(x^*) = 0$, es decir que nuestro problema se reduce a encontrar una función $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tal que $F(x^*) = x^* - G(x^*) = 0$. Así tenemos el siguiente teorema:

Teorema de Existencia y Unicidad (para \mathbb{R}^n). Sea $D = [a_1, b_1] \times \dots \times [a_n, b_n] \subset \mathbb{R}^n$.

- Si $G \in \mathcal{C}(D)$ con $G(D) \subset D$. Entonces G tiene un punto fijo en D .
- Si $\frac{\partial g_i}{\partial x_j}$, $i, j = 1, \dots, n$ tienen derivadas parciales continuas, y existe una constante $q < 1$ con

$$\left| \frac{\partial g_i(x)}{\partial x_j} \right| \leq \frac{q}{n}, \quad x \in D, \quad i, j = 1, \dots, n$$

Entonces la sucesión $\{x^{(k)}\}_{k=0}^{\infty}$ para un $x^{(0)} \in D$, y generada por:

$$x^{(k)} = G(x^{(k-1)}), \quad k \geq 1$$

Converge al único punto fijo $p \in D$.

Además tenemos que se da que:

$$\|x^{(k)} - p\| \leq \frac{q^k}{1-q} \|x^{(1)} - x^{(0)}\|$$

Veamos ahora un ejemplo de esto:

Ejemplo 3. Consideremos el sistema no lineal:

$$x_1^2 - 10x_1 + x_2^2 + 8 = 0$$

$$x_1x_2^2 + x_1 - 10x_2 + 8 = 0$$

$$\begin{cases} x_1^2 - 10x_1 + x_2^2 + 8 = 0 \\ x_1x_2^2 + x_1 - 10x_2 + 8 = 0 \end{cases}$$

Paso 1: Transformar el Sistema no lineal en uno de punto fijo

Luego este problema puede transformarse en:

$$x_1 = g_1(x_1, x_2) = \frac{x_1^2 + x_2^2 + 8}{10}$$

$$x_2 = g_2(x_1, x_2) = \frac{x_1x_2^2 + x_1 + 8}{10}$$

Ahora procederemos a mostrar que $G = (g_1, g_2)$ tiene un único punto fijo en $[0, 3/2]^2$

Paso 2: Calcular la matriz Jacobiana (Incluyendo dónde se anula)

Calculando el Jacobiano obtenemos:

$$J_G(x_1, x_2) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} \frac{2x_1}{10} & \frac{2x_2}{10} \\ \frac{x_2^2 + 1}{10} & \frac{2x_1 x_2}{10} \end{pmatrix}$$

Ahora hallemos dónde $J_G(x_1, x_2) = 0$:

Paso 2: Sacar las normas matriciales

Ya que conocemos como es la matriz de iteración de Jacobi, veamos que ocurre con las normas matriciales:

Norma 1: Recordemos que la Norma 1 para matrices es sacar el máximo de la suma de las columnas, en este caso tenemos:

$$\|B_J\|_1 = \max \left\{ \frac{22}{6}, \frac{21}{4}, \frac{10}{3} \right\} = \frac{21}{4} > 1$$

Norma ∞ : Recordemos que la Norma infinito para matrices es sacar el máximo de la suma de las filas, en este caso tenemos:

$$\|B_J\|_\infty = \max \left\{ \frac{15}{4}, \frac{15}{6}, 6 \right\} = 6 > 1$$

Norma Frobenius: Si recordamos esta se saca sumando el cuadrado de las componentes, lo qué nos da:

$$\|B_J\|_F = \frac{49}{16} + 4 + \frac{49}{36} + \frac{16}{9} + \frac{25}{4} + \frac{49}{4} = \frac{875}{48} > 1$$

Para la norma dos se debe considerar solo la transpuesta, dado que todas las entradas son números reales, queda de tarea realizar la norma 2.

Nos da de resultado que:

$$\|B_J\|_2 = \sqrt{15,289} > 1$$

Si notamos, todas las normas que conocemos nos da que son mayores que 1, ahora nos queda de última herramienta verificar el radio espectral de B_J :

Paso 3 (Opcional)¹²: Sacar el radio espectral las normas matriciales

¹²Digo opcional, dado que si alguna norma nos da menor que 1, podemos concluir de inmediato que es convergente el método.

Por el curso de Álgebra Lineal sacamos los valores propios, los cuáles son:

$$\rho(B_J) = \max \{|\lambda| : \lambda \text{ es un valor propio}\} = \max \{3,9101, 1,4634, 2,4467\} = 3,9101 > 1$$

Dado que el radio espectral da mayor que 1, concluimos rotundamente que

El Método no Converge a la Solución, pero si resolvemos el sistema, nos da que su solución x es:

$$x = \begin{pmatrix} \frac{39}{79} \\ \frac{35}{79} \\ \frac{20}{79} \\ -\frac{20}{79} \end{pmatrix}$$

3. Parcial 3

Luego de haber analizado los conceptos básicos del Álgebra Lineal Numérica, nos preocuparemos por estudiar como sería el comportamiento polinomial entre ciertos puntos que de ahora en adelante llamaremos *nodos*.

Muchas veces tendremos funciones de distintos tipos¹³, y queremos modelar mediante funciones conocidas sus comportamientos, estas funciones conocidas son los polinomios, dado que con estos hay mayor manejo y facilidad¹⁴, existen muchos tipos de Interpolación, pero aquí nos centraremos en la interpolación de Lagrange, y en la de Newton.

3.1. Interpolación de Lagrange

Cómo hemos mencionado usaremos algo llamado nodos, los cuáles son puntos dentro del dominio dónde queremos hacer nuestra modelación con polinomios, podemos en este caso dar ciertas salvedades para algunas condiciones:

- Si tenemos $n + 1$ nodos. solo podemos asegurar un único polinomio de grado n . Podemos hacer la analogía con las lineales y las cuadráticas.
- La interpolación funcionará en un intervalo dado, hacerlo para todo \mathbb{R} , no es muy práctico, aquí ya usaríamos más bien Series de Taylor.
- Los nodos pueden ser equiespaciados (Lagrange y Newton) pero existe el caso dónde no lo sean (Método de Chebyshev).

Para realizar la interpolación de Lagrange daremos un preámbulo:

Tenemos el siguiente problema con los siguientes datos:

i	0	1	2	\dots	n
x_i	x_0	x_1	x_2	\dots	x_n
f_i	y_0	y_1	y_2	\dots	y_n

Cuadro 1: Tabla con datos para un problema.

Podemos preguntarnos si ¿existe una función que vaya de $[x_0, x_n]$ en \mathbb{R} que cumpla que al evaluarla en los nodos nos dé el valor de la función original, y además que haga un mínimo el error entre ambas?

¹³Buscaremos en inicio que sean continuas para mayor facilidad, ¿Podría definirse en no continuas en ciertos puntos? Si, pero la representación no sería la mejor comparándola con funciones uaves y continuas.

¹⁴No son los únicos, en el Kress, se menciona la interpolación trigonométrica, que puede llegar a ser muy útil, pero para el curso no es necesaria, dado que es lo básico.

La respuesta es que sí existe. y es a la que llamaremos *Función Interpoladora*.

Usaremos el concepto de la *Delta de Kronecker* para desarrollar la interpolación de Lagrange, dado que queremos que la función en los nodos valga el propio valor de la función, por lo que necesitamos que se haga 1, y luego multiplicamos por el valor de la función original evaluada en el nodo.

Usaremos la siguiente fórmula:

$$l_v(x) := \prod_{\substack{i=1 \\ i \neq v}}^n \frac{x - x_i}{x_v - x_i} \quad (28)$$

Dónde $v \in \{0, 1, \dots, n\}$ Es decir que recorre los $n + 1$ nodos dónde deseamos definir nuestra interpolación.

Definamos ψ_i como la función interpoladora. Ahora finalmente resulta en:

$$\psi_n(x) = \sum_{i=0}^n f_i l_i \quad (29)$$

Dónde f_i es el valor de la función en el nodo i , y l_i es el resultado de la ecuación anterior aplicado en el nodo i . Es decir al final nuestra función interpoladora es la suma de todos estos valores que pasan alrededor de nuestros nodos elegidos. Demos un ejemplo más claro:

Ejemplo 4. Sea $f(x) = \sin(x^2)$, y consideremos el intervalo $[0, 2]$, y consideremos los nodos $\left\{0, \frac{1}{2}, 1, \frac{3}{2}, 2\right\}$, es decir tenemos $4 + 1 = 5$ nodos, y con esto lograremos tener un polinomio de grado 4 **ÚNICO**, empecemos por hallar los f_i :

Paso 1: Calcular los Valores de f_i

Empecemos por evaluar en cada nodo:

- *Nodo 1* := $0 : f(0) = \sin(0) = 0$.
- *Nodo 2* := $\frac{1}{2} : f(0,5) = \sin(0,25) = 0,247404$.
- *Nodo 3* := $1 : f(1) = \sin(1) = 0,841471$.
- *Nodo 4* := $\frac{3}{2} : f(1,5) = \sin(1,5) = 0,778073$.
- *Nodo 5* := $2 : f(2) = \sin(2) = -0,756802$.

Paso 2: Calcular los Valores de l_v

De manera sistemática iremos calculando cada valor:

$$l_0 = \prod_{i=1}^4 \frac{x - x_i}{x_0 - x_i} = \frac{(x - 0,5)(x - 1)(x - 1,5)(x - 2)}{(-0,5)(-1)(-1,5)(-2)} = \frac{2}{3}(x - 0,5)(x - 1)(x - 1,5)(x - 2)$$

$$l_1 = \prod_{\substack{i=0 \\ i \neq 1}}^4 \frac{x - x_i}{x_1 - x_i} = \frac{(x - 0)(x - 1,5)(x - 2)}{(0,5)(-1)(-1,5)} = -\frac{8}{3}(x - 0)(x - 1,5)(x - 2)$$

$$l_2 = \prod_{\substack{i=0 \\ i \neq 2}}^4 \frac{x - x_i}{x_2 - x_i} = \frac{(x - 0)(x - 0,5)(x - 1)(x - 2)}{(1)(0,5)(-1)} = 4(x - 0)(x - 0,5)(x - 1)(x - 2)$$

$$l_3 = \prod_{\substack{i=0 \\ i \neq 3}}^4 \frac{x - x_i}{x_3 - x_i} = \frac{(x - 0)(x - 0,5)(x - 1)(x - 2)}{(1,5)(1)(0,5)(-0,5)} = -\frac{8}{3}(x - 0)(x - 0,5)(x - 1)(x - 2)$$

$$l_4 = \prod_{\substack{i=0 \\ i \neq 4}}^4 \frac{x - x_i}{x_4 - x_i} = \frac{(x - 0)(x - 0,5)(x - 1)(x - 1,5)}{(2)(1,5)(1)(0,5)} = \frac{2}{3}(x - 0)(x - 0,5)(x - 1)(x - 1,5)$$

Ahora para hallar el polinomio interpolador necesitamos multiplicar y sumar los l_i con los f_i :

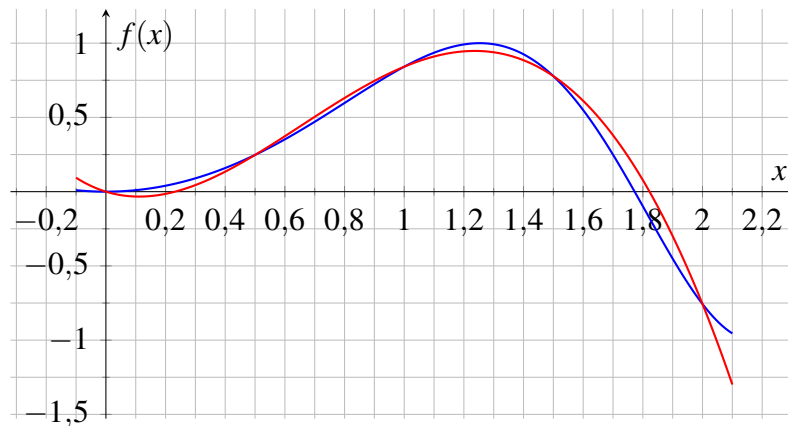
Paso 3: Calcular el polinomio interpolador

Recordemos que el polinomio interpolador cumple:

$$\begin{aligned} \psi_4(x) = \sum_{i=0}^4 f_i l_i = & 0 \cdot \frac{2}{3}(x - 0,5)(x - 1)(x - 1,5)(x - 2) - 0,247404 \cdot \frac{8}{3}(x - 0)(x - 1)(x - 1,5)(x - 2) \\ & + 0,841471 \cdot 4(x - 0)(x - 0,5)(x - 1)(x - 2) - 0,778073 \cdot \frac{8}{3}(x - 0)(x - 0,5)(x - 1)(x - 2) \\ & - 0,756802 \cdot \frac{2}{3}(x - 0)(x - 0,5)(x - 1)(x - 1,5) \end{aligned}$$

Es decir que con la suma de todas estas funciones nos da una interpolación muy interesante y chévere!

$$f_1(x) = \sin(x^2) \text{ y } \psi_4(x) = 0 - 0,616327x + 3,05011x^2 - 1,71906x^3 + 0,126741x^4$$



Ejemplo 5. Tarea: Calcular la interpolación con 9 nodos.

Pueden darse muchas veces los casos en los que no conozcamos la función sino solamente los puntos nodales, veamos que se hace en este caso:

Ejemplo 6. Tenemos la siguiente tabla de datos:

i	x	$f(x)$
0	-1	-0,5
1	0	0,2
2	0,5	0,6
3	2	-0,1
4	4	1

Cuadro 2: Tabla de Datos del Ejemplo

Veamos como sería la interpolación:

Sí notamos hay en total 5 nodos, es decir que la interpolación será única de grado 4, empecemos por calcular los l_v :

■

$$l_0 = \prod_{\substack{i=0 \\ i \neq 0}}^4 \frac{x - x_i}{x_0 - x_i} = \frac{(x - 0)(x - 0,5)(x - 2)(x - 4)}{(-1 - 0)(-1 - 0,5)(-1 - 2)(-1 - 4)} = \frac{2(x - 0)(x - 0,5)(x - 2)(x - 4)}{45}$$

■

$$l_1 = \prod_{\substack{i=0 \\ i \neq 1}}^4 \frac{x - x_i}{x_1 - x_i} = \frac{(x+1)(x-0,5)(x-2)(x-4)}{(0-(-1))(0-0,5)(0-2)(0-4)} = -\frac{(x+1)(x-0,5)(x-2)(x-4)}{4}$$

■

$$l_2 = \prod_{\substack{i=0 \\ i \neq 2}}^4 \frac{x - x_i}{x_2 - x_i} = \frac{(x+1)(x-0)(x-2)(x-4)}{(0,5+1)(0,5-0)(0,5-2)(0,5-4)} = \frac{16(x+1)(x-0)(x-2)(x-4)}{63}$$

■

$$l_3 = \prod_{\substack{i=0 \\ i \neq 3}}^4 \frac{x - x_i}{x_3 - x_i} = \frac{(x+1)(x-0)(x-0,5)(x-4)}{(2+1)(2-0)(2-0,5)(2-4)} = -\frac{(x+1)(x-0)(x-0,5)(x-4)}{18}$$

■

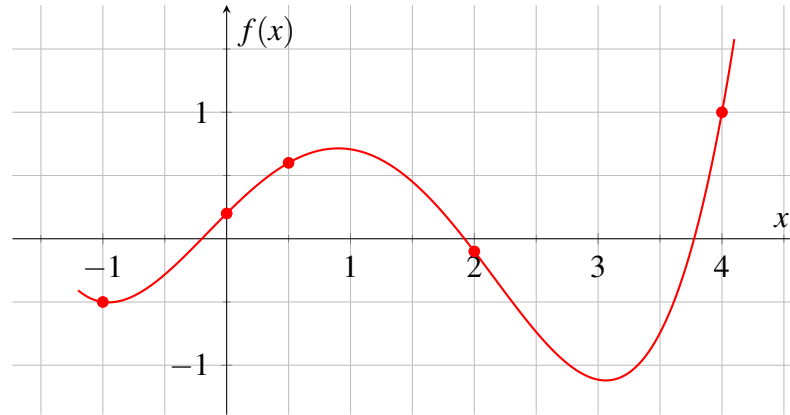
$$l_4 = \prod_{\substack{i=0 \\ i \neq 4}}^4 \frac{x - x_i}{x_4 - x_i} = \frac{(x+1)(x-0)(x-0,5)(x-2)}{(4+1)(4-0)(4-0,5)(4-2)} = \frac{(x+1)(x-0)(x-0,5)(x-2)}{140}$$

Así el polinomio interpolador será:

$$\begin{aligned} \psi_4(x) = \sum_{i=0}^4 f_i l_i = & -\frac{(x-0)(x-0,5)(x-2)(x-4)}{45} - \frac{(x+1)(x-0,5)(x-2)(x-4)}{20} \quad (30) \\ & + \frac{16(x+1)(x-0)(x-2)(x-4)}{105} + \frac{(x+1)(x-0)(x-0,5)(x-4)}{180} + \frac{(x+1)(x-0)(x-0,5)(x-2)}{140} \end{aligned}$$

Si lo deseamos observar gráficamente sería:

$$\psi_4(x) = 0,2 + 0,976191x - 0,189286x^2 - 0,372619x^3 + 0,0928572x^4$$



Es decir que nuestros datos pueden ser agrupados por esta función de grado 4, pero puede darse el caso dónde hallan más funciones, por ejemplo Splines Cúbicos, solo que debemos recordar que la interpolación reduce el error en lo mayor posible (y eso que falta una parte de interpolación que ¡mejora todo!)

Ejemplo 7. Considera la siguiente tabla de datos:

i	x	$f(x)$
0	-2	-0,8
1	-1,5	0,6
2	-0,5	0,3
3	-0,1	-0,1
4	0,1	0,1
5	0.5	-0.3
6	1.5	-0.6
7	2	0.8

Cuadro 3: Tabla de Datos del Ejemplo

Realice la interpolación, y grafique.

3.2. Interpolación de Newton

Consideremos que con $n + 1$ nodos solo se puede alcanzar un único polinomio interpolador, es decir que no hay otro para determinar, en este caso el método de Newton, me da otra alternativa de

como calcularlo, y en ocasiones mucho más efectiva. Para hacer esto, deberemos de definir algo llamado *Diferencias Divididas de Newton*, las cuáles implican:

Diferencias Divididas de Newton. *Tenemos una cantidad $n + 1$ de nodos distintos, enumerados como $x_0, \dots, x_n \in [x_0, x_n] \subset \mathbb{R}$, y también definamos $f_i := f(x_i)$, luego para $v, \kappa \geq 0$, $v + \kappa \leq n$ definimos las Diferencias Divididas de Newton como:*

$$f[x_v, \dots, x_{v+\kappa}] := \begin{cases} f[x_v] = f_v = f(x_v) & \kappa = 0 \\ \frac{f[x_{v+1}, \dots, x_{v+\kappa}] - f[x_v, \dots, x_{v+\kappa-1}]}{x_{v+\kappa} - x_v} & \kappa > 0 \end{cases}$$

Demos un ejemplo considerando la siguiente información:

i	0	1	2	3	4
x_i	-1	$-\frac{1}{2}$	0	$\frac{1}{2}$	1
$f(x_i)$	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0

Cuadro 4: Información

$$f[x_0] = 0; f[x_1] = \frac{1}{2}; f[x_2] = 0; f[x_3] = \frac{1}{2}; f[x_4] = 0$$

$$f[x_0, x_1] = \frac{f[x_0] - f[x_1]}{x_0 - x_1} = 1; f[x_1, x_2] = \frac{f[x_1] - f[x_2]}{x_1 - x_2} = -1; f[x_2, x_3] = \frac{f[x_2] - f[x_3]}{x_2 - x_3} = 1$$

$$f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0} = -2; f[x_1, x_2, x_3] = \frac{f[x_2, x_3] - f[x_1, x_2]}{x_3 - x_1} = 2$$

Veamos una nueva tabla dónde daremos el método ilustrativo. Inicialmente vamos a tener unos datos, e iremos descendiendo en su cantidad, es decir:

Para armar nuestro polinomio interpolador necesitaremos tomar ciertos elementos de la tabla 5, específicamente los elementos de arriba, es decir:

$$f[x_0], f[x_0, x_1], \dots, f[x_0, \dots, x_n]$$

Y además consideraremos lo que llamaremos *Polinomio nodal*, el cuál es el producto de los $x - x_i$, es decir:

$$w_n(x) = \prod_{i=0}^n (x - x_i) \quad (31)$$

Solo que aquí en el Método de Newton, el enésimo más 1 nodo no lo consideramos, es decir solo nos interesa saber el polinomio nodal hasta $n - 1$. Armemos ahora la interpolación de Newton:

i	x_i	$k = 0$	$k = 1$	$\dots k = n - 1$	$k = n$
0	x_0	$f[x_0]$			
			$f[x_0, x_1]$		
1	x_1	$f[x_1]$			
			$f[x_1, x_2]$		
2	x_2	$f[x_2]$			
				$f[x_0, x_1, \dots, x_{n-1}]$	
\vdots	\vdots	\vdots	\vdots		$f[x_0, \dots, x_n]$
				$f[x_1, x_2, \dots, x_n]$	
$n - 2$	x_{n-2}	$f[x_{n-2}]$			
			$f[x_{n-2}, x_{n-1}]$		
$n - 1$	x_{n-1}	$f[x_{n-1}]$			
			$f[x_{n-1}, x_n]$		
n	x_n	$f[x_n]$			

Cuadro 5: Diferencias Divididas de Newton

$$\rho_n(x) = f[x_0] + \sum_{i=1}^n f[x_0, \dots, x_i] \prod_{j=0}^{i-1} (x - x_j) = f_0 + \sum_{i=1}^n f[x_0, \dots, x_k] \omega_k(x)$$

Podemos verlo como:

$$\rho_n(x) = f_0 + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) + \dots + f[x_0, \dots, x_n](x - x_0) \dots (x - x_{n-1}) \quad (32)$$

Si observamos, debemos realizar muchas operaciones, en el paso n hemos realizado $\frac{n(n+1)}{2}$ productos, y n sumas, por lo que computacionalmente hay mucho esfuerzo, debido a esto podemos usar el método de Horner para resolver esto.

3.2.1. Método de Horner

Notemos inicialmente que todos tienen en común al inicio el valor $(x - x_0)$, saquémoslo de factor común, luego dentro de esos elementos, si quitamos el primero, los que quedan tienen como factor común a $(x - x_1)$, hacemos esto repetidamente y nos queda que se van reduciendo y así hasta que queda solo $(x - x_{n-1})$, con esto reducimos las multiplicaciones a solo n , si $n = 3$, resulta que:

$$\rho_3(x) = f_0 + (x - x_0)(f[x_0, x_1] + (x - x_1)(f[x_0, x_1, x_2] + (x - x_2)(f[x_0, x_1, x_2, x_3])))$$

Cualquiera de las dos formas es aceptable, ahora veamos los siguientes ejercicios:

Ejemplo 8. *Hacer la interpolación con Newton y con Horner, del ejemplo 4,6,7.*

3.3. Cálculo del error de Interpolación

Cómo en los parciales y temas pasados nosotros podemos determinar el error de la función cometido con cierta cantidad de nodos, la fórmula estará determinada por:

$$\mathcal{E}_n := \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \prod_{i=1}^n (x - x_i) \quad (33)$$

Ejemplo 9. *Calcular el error de todos los ejemplos y puntos de tarea¹⁵.*

Pero hay una pregunta latente dentro de este tema y es la idea de cómo hacer la **mejor** interpolación, en este caso Chebyshev logró dar con la respuesta, y es tomar nodos muy específicos los cuáles no son equidistantes.

3.4. Nodos de Chebyshev

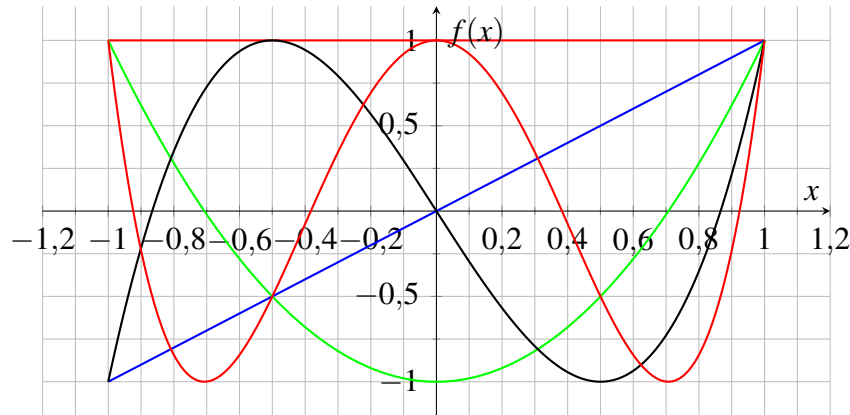
Muchas veces desearemos encontrar ciertos puntos dónde nuestra interpolación se reduzca cada vez más (error) y se parezca lo más posible a la original, esto lo hacemos mediante una fórmula y ciertas transformaciones, como lo es la que propone Chebyshev. Usaremos entonces una estructura de polinomios que conoceremos como: *Polinomios de Chebyshev*

En este sentido podemos calcular los polinomios recursivamente de esta forma:

$$T_0(x) = 1; T_1(x) = x; T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x); n > 1 \quad (34)$$

Demos por ejemplo los primeros polinomios:

¹⁵Pista: Recordemos usar la norma del máximo para hallar cotas del error, recordemos que nunca se puede hallar con exactitud el error, sino se usan cotas para el mismo.



Con estos polinomios, vamos encontrando cierto tipo de raíces, las cuáles nos van a ayudar mucho para poder desarrollar lo que deseamos.

T_n tiene ceros en los siguientes puntos de x :

$$x_k = \cos\left(\frac{2k+1}{n} \frac{\pi}{2}\right), \quad x_k \in [-1, 1], \quad k = 0, 1, \dots, n-1$$

Solo que hay una peculiaridad antes de usar los nodos de Chebyshev, y es el hecho de que solo funciona en el intervalo simétrico $[-1, 1]$, por lo que para convertirlo en un intervalo cualquiera deberemos multiplicar los $n+1$ nodos de Chebyshev, es decir:

$$x_k = \frac{b-a}{2} \cos\left(\frac{2k+1}{n+1} \frac{\pi}{2}\right) + \frac{b+a}{2}$$

Ahora, el paso clave dentro de los nodos de Chebyshev es saber usar estos ceros, y luego de hallar los nodos de Chebyshev usaremos alguna de las interpolaciones vistas con anterioridad, hagamos nuevamente el ejemplo 4, pero usando nodos de Chebyshev:

Ejemplo 10. Sea $f(x) = \sin(x^2)$, usaremos la transformación de los 5 nodos, es decir:

$$x_k = \frac{2-0}{2} \cos\left(\frac{2k+1}{5} \frac{\pi}{2}\right) + \frac{2+0}{2} = \cos\left(\frac{\pi(2k+1)}{10}\right) + 1$$

Ahora hagamos el proceso con cada nodo, es decir $k = 0, 1, 2, 3, 4$:

$$1. \quad x_0 = \cos\left(\frac{\pi}{10}\right) + 1 = 1,951057, f(x_0) = -0,617083.$$

$$2. \quad x_1 = \cos\left(\frac{3\pi}{10}\right) + 1 = 1,587785, f(x_1) = 0,581468.$$

$$3. x_2 = \cos\left(\frac{5\pi}{10}\right) + 1 = 1, f(x_2) = 0,841471.$$

$$4. x_3 = \cos\left(\frac{7\pi}{10}\right) + 1 = 0,412215, f(x_3) = 0,169105.$$

$$5. x_4 = \cos\left(\frac{9\pi}{10}\right) + 1 = 0,048943, f(x_4) = 0,002395.$$

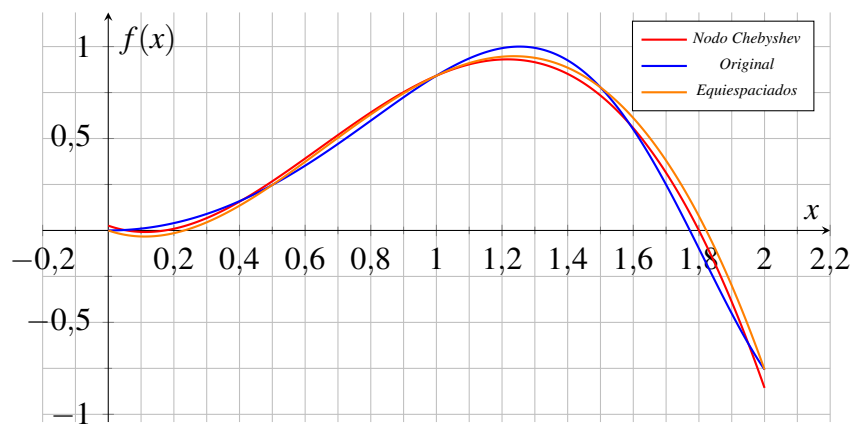
Conociendo estos datos, nos queda hacer la interpolación, ya sea Lagrange o Newton, luego el resultado es:

i	x_i	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$
0	0,048943	0,002395				
1	0,412215	0,169105	0,458912			
2	1	0,841471	1,143898	0,720237		
3	1,587785	0,581468	-0,442344	-1,349339	-1,344892	
4	1,951057	-0,617083	-3,299321	-3,004002	-1,075265	0,141751

Cuadro 6: Diferencias Divididas de Newton del problema planteado.

Luego nuestro polinomio interpolador será:

$$\tau_4(x) = 0,002395 + (x - 0,048943)(0,458912 + (x - 0,412215)(0,720237 + (x - 1)(-1,344892 + 0,141751(x - 1,587785))))$$



Con esto concluimos lo dicho anteriormente, es decir lo de que los nodos de Chebyshev dejan hacer mejor aproximación, dado que son los mejores nodos, y los adecuados para realizar una interpolación polinomial¹⁶, pero con estos el error mejor de una forma increíble, veamos como queda:

3.4.1. Error con los Nodos de Chebyshev

El objetivo inicial de Chebyshev era el poder manejar de una forma más tranquila el polinomio nodal, por lo que se lograron establecer las siguientes cotas:

$$|\omega_n| = \frac{|T_n|}{2^n} \leq \frac{1}{2^n}$$

Dado que los nodos de Chebyshev nos otorga la mejor interpolación, procederemos a ver como el error se estipula de esta nueva forma:

$$\mathcal{E} \leq \frac{1}{2^n} \frac{1}{(n+1)!} \|f^{(n+1)}\|_\infty \quad (35)$$

Lo cuál si analizamos, a mayor cantidad de nodos, nos resulta mucho mejor dado que ese exponencial con ese factorial dominan abajo, y hacen que el error se vaya mucho más rápido a cero, lo cuál concuerda con el ser los mejores nodos de elección para realizar este método, solo que se tiene el paso extra de tener que calcular los nodos mediante la expresión trigonométrica.

Hasta este punto llega la parte de Interpolación, sí la persona desea profundizar más o tener mayor conocimiento en este tema, queda invitado a estudiar Splines Lineales, y Splines Cúbicos, dado que estos proporcionan una ayuda de aproximación entre comillas más simple, pero que carga con más errores.

La importancia de los Splines Lineales es que estos inicialmente ayudan a analizar estructuras de FEM¹⁷ en 1D, tanto para EDO's como para EDP's, de una forma fácil simple y confiable, dado que funcionan inicialmente de buena forma, entre más subamos el grado de las funciones mejor se aproxima.

Entre este tema y el que sigue fue de los que personalmente más me disfruté por la idea y los polinomios y datos que se usan, tales como los polinomios de Chebyshev, ahora usaremos nuevas ideas para aproximar áreas bajo la curva, es decir para realizar Integrales.

El problema radica en que muchas veces estas integrales no son exactas como quisieramos en matemáticas hallar una expresión siempre, pero nos ayuda y facilita en ver comportamientos.

¹⁶Aclaremos Polinomial, dado que la Interpolación Trigonométrica va de la mano con otras cosas.

¹⁷Por las siglas en inglés significa *Finite Element Method*.

3.5. Integración Numérica

Como hemos dicho anteriormente y por las experiencias con cursos pasados sabemos que integrar en general no es fácil de realizar, por lo que daremos ciertos métodos para poder realizar este trabajo, pasando por ciertos nodos equiespaciados y luego llegando a las ideas de Gauss para realizar con mayor potencia las integrales propuestas, y como con ayuda de otros matemáticos se desarrollaron más estas ideas.

3.5.1. Integración de Newton-Cotes

Veamos ahora la idea de hacer estas Integrales:

Al realizar la integración obtenemos ciertos problemas, los cuáles son:

1. ¿ $\mathcal{I}(f) := \int_a^b f(x) dx$ existe?
2. ¿ $\mathcal{I}(f)$ es fácil de calcular?

Inicialmente podemos plantear como solución el aproximar nuestra solución con ciertas sucesiones que cumplan:

1. $\mathcal{I}(f_n) := \mathcal{I}_n(f) := \int_a^b f_n(x) dx$.
2. $\mathcal{I}_n(f)$ sea fácil de calcular.

Con esto surge la necesidad de buscar puntos y lugares conocidos de tal forma que podamos aproximar esa área, es decir

- 3.

$$\mathcal{I}_n(f) = \sum_{i=0}^n a_i f(x_i)$$

Dónde los x_i determinan los puntos de cuadratura, y los $f(x_i)$ los pesos de cuadratura.

Enunciemos entonces un teorema:

Cuadratura de Orden n por interpolación polinómica. Si $f_n := \mathcal{P}_n \in \mathbb{P}_n$ define el único polinomio interpolador de f en los $n+1$ nodos x_0, x_1, \dots, x_n entonces:

$$\mathcal{I}_n(f) := \int_a^b f_n(x) dx = \sum_{i=0}^n a_i f(x_i)$$

Dónde

$$a_i = \int_a^b \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j} dx, \quad i = 0, \dots, n$$

En este punto es bueno recordar ese proceso que realizábamos en la parte de interpolación, dado que ese a_i es como considerarlo en esta parte. Analicemos ahora otro teorema poderoso que nos permitirá entender mejor este tema de la Integración Numérica.

Cuadratura de Orden n por interpolación polinómica (2). Sean los $n + 1$ nodos distintos, ahora llamados **Puntos de Cuadratura** $x_0, \dots, x_n \in [a, b]$, la fórmula de cuadratura de orden n está únicamente determinada por la propiedad de integrar todo $p \in \mathbb{P}_n$ exactamente, es decir:

$$\int_a^b p(x) dx = \sum_{i=0}^n a_i p(x_i) \quad \forall p \in \mathbb{P}_n \quad (36)$$

Este resultado es muy poderoso porque nos dice que dependiendo de la cantidad de nodos podemos encontrar el valor exacto de una integral de polinomios, y tiene sentido, dado que sabemos que $n + 1$ puntos de cuadratura generan un único polinomio de grado n , y por la forma en que definimos los puntos de cuadratura, y la fórmula de integración (es decir mediante la interpolación), dado que integramos los polinomios de Lagrange, y con estos procedemos a encontrar esa área bajo la curva.

Analicemos los distintos métodos de Integración y qué tan efectivos son:

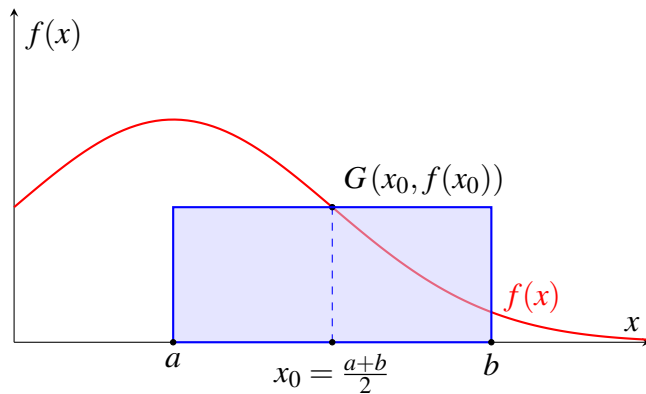
3.5.2. Método del Rectángulo (1 nodo)

La fórmula más fácil que podemos tener es usar solo un nodo de cuadratura, (en general consideramos siempre el punto medio), y con este generamos idealmente un rectángulo (dado que solo necesitamos altura, puesto que el intervalo es dado, es como que tenemos la base, pero necesitamos la ubicación del nodo para conocer la altura).

Tomemos por ejemplo la función:

$$f(x) = 2e^{(-\frac{1}{2}(x-1)^2)}$$

Y el intervalo $[1, 3]$, entonces el punto medio será: $x_0 = 2$ el cuál nos servirá de punto de cuadratura, luego tendremos que el rectángulo formado estará dado por:



Luego evaluando en la función el valor $x_0 = 2$ obtenemos que $f(x_0) = \frac{2}{\sqrt{e}} \approx 1,213061$, luego el área del rectángulo es:

$$(3 - 1)(1,213061) = 2,426122$$

Sí realizamos el proceso de integración, nos damos cuenta que:

$$\int_1^3 f(x) dx = 2,392576$$

Ahora vemos que este método no es muy efectivo que digamos, dado que ni una cifra decimal tiene de precisión.

Cómo fórmula general tenemos que:

$$\int_a^b f(x) dx \approx (b - a)f\left(\frac{b+a}{2}\right)$$

Ahora como lo dijimos en el teorema, al haber $0 + 1 = 1$ puntos de cuadratura, esta fórmula integra de forma exacta polinomios de grado a lo sumo 0, es decir en este caso solo integra de forma exacta constantes.

Tarea:

Considere las siguientes funciones con sus respectivos intervalos:

- $f(x) = \frac{e^x \sin(x^2)}{\sqrt{x+1}}, \quad x \in [0, 2]$
- $f(x) = \ln(1+x^3) \cdot \cos\left(\frac{1}{x+0,1}\right), \quad x \in [0, 1]$
- $f(x) = \frac{\sinh(x)}{1+x^4}, \quad x \in [-2, 2]$

- $f(x) = \sqrt{\ln(x+2)} \cdot \sin(x^2), \quad x \in [0, 2]$
- $f(x) = \frac{\ln(x^2+1)}{x^2 + \cos(x)}, \quad x \in [-2, 2]$
- $f(x) = \sin\left(\frac{1}{x+0,01}\right) \cdot e^{-x^2}, \quad x \in [0, 1]$
- $f(x) = \frac{1}{1+x^2 \ln(1+x)}, \quad x \in [1, 5]$
- $f(x) = \arctan\left(\frac{\sin(x)}{x+1}\right) \cdot \ln(x+2), \quad x \in [0, 3]$
- $f(x) = e^{-e^x} \cdot \sin(x), \quad x \in [0, 2]$
- $f(x) = \sqrt{|\cos(x^3)|} \cdot \frac{e^{-x}}{\ln(x+2)}, \quad x \in [0, 3]$

Usando regla del trapecio, halle la aproximación de la integral en cada intervalo.

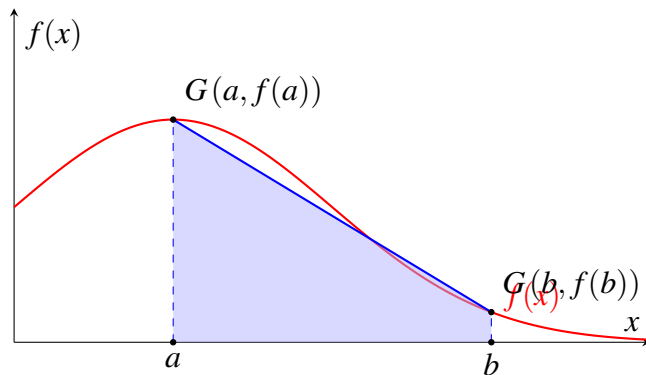
3.5.3. Método del Trapecio (2 nodos)

Ahora ¿Qué pasará si le aumentamos 1 nodo? En ese caso obtenemos una nueva regla, que llamaremos **regla del trapecio**, aquí ya se complica un poco más la geometría puesto que ya pasamos a una figura más elaborada, dónde sí vemos el siguiente gráfico, necesitaremos formar dos bases y considerar una altura¹⁸, usemos la misma función $f(x)$ para hacerle el análisis detallado de esta técnica de integración:

Aquí al tener dos nodos, usaremos los de los extremos (por pura facilidad y que recordemos que necesitamos que sean equiespaciados)¹⁹ para definir nuestra integración, podemos llamar $x_0 = a, x_1 = b$ como nuestros puntos de cuadratura. Procedamos entonces con el trabajo:

¹⁸Consideremos un trapecio acostado para próximamente ver la fórmula.

¹⁹Aquí estamos como en la misma idea de Interpolación, es decir estamos usando nodos equiespaciados inicialmente para definir la teoría, pero luego veremos que buscando otros nodos, las cosas van a funcionar mucho mejor.



Si recordamos la formula del área del trapecio, tendremos que:

$$\text{Área} = \frac{f(a) + f(b)}{2} (b - a)$$

Así nuestro método resultará en:

$$\text{Área} = \frac{\frac{2}{e^2} + 2}{2} (3 - 1) = \frac{2 + 2e^2}{e^2} \approx 2,270671$$

Graciasamente la elección de la función resultó en que el método del rectángulo es mejor que el del trapecio, ya es cuestión de Geometría del problema, pero se da que este método mejora la calidad de la integral, hay casos dónde este método puede llegar a dar 0, o menores valores, son casos muy específicos, o familia de funciones, fue una elección curiosa de función esta.

Ahora el teorema me garantiza que integra polinomios de a lo sumo 1 grado de forma exacta, es decir lineales y constantes, aquí evidenciamos el poder de este método comparado con el otro.

La fórmula para este método sería:

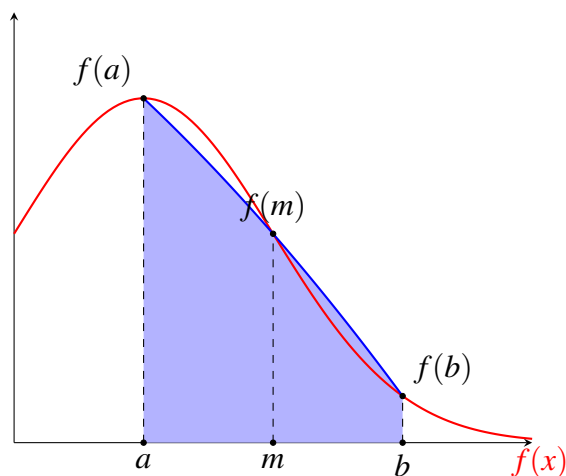
$$\int_a^b f(x) dx \approx \frac{(b-a)}{2} (f(a) + f(b))$$

Tarea:

Dejamos como ejercicios hacer los mismos anteriores, pero con la regla del trapecio.

3.5.4. Método de Simpson (3 nodos)

Cómo hemos venido anteriormente hablando, consideremos que pasa si tenemos ahora 3 nodos:



Ahora, geométricamente hablando es más difícil ver la figura y forma exacta, por lo que daremos la fórmula y la usaremos:

$$\int_a^b f(x) dx \approx \frac{b-a}{6} (f(a) + 4f(m) + f(b))$$

Dónde $m = \frac{a+b}{2}$, calculemos cuánto da dicha integral:

$$\int_a^b f(x) dx \approx \frac{3-1}{6} (f(1) + 4f(2) + f(3)) = \frac{1}{3} (2 + 4,85225 + 0,270671) = 2,374307$$

Ahora sí evidenciamos que mejoró la fórmula, dado que ya tiene una cifra decimal de exactitud, recordemos que en virtud del teorema, este método integra de forma exacta polinomios de a lo sumo grado 2, es decir constantes, lineales y cuadráticos.

Ahora nos podemos realizar una pregunta más profunda:

¿Cómo calculamos los pesos y los puntos de cuadratura?

Dado que hemos estado dando alegremente las fórmulas, pero como buenos matemáticos debemos entender todo.

3.5.5. Cambio de Intervalo

Haremos un paso extra que nos facilitará las cosas en términos de Cálculo, dado que se puede volver complejo todo de aquí en adelante.

Busquemos formar un isomorfismo entre espacios, para que de esta forma todo salga más fácil, buscaremos pasar del intervalo $[-1, 1]$ al intervalo $[a, b]$, todo eso mediante la siguiente transformación lineal:

$$x(u) = \frac{1}{2}(a+b) + \frac{1}{2}(b-a)u, \quad u \in [-1, 1]$$

Así nos resulta que:

$$\int_a^b f(x) dx = \frac{b-a}{2} \int_{-1}^1 f(x(u)) du = \frac{b-a}{2} \int_{-1}^1 g(u) du \approx \frac{b-a}{2} \sum_{i=0}^n a_i g(u_i)$$

3.5.6. Verificación del Método del Rectángulo

Recordemos de antemano que los puntos de cuadratura ya los conocemos²⁰ Por lo que nos faltan los pesos de cuadratura, así que procederemos a calcularlos.

$$\int_{-1}^1 p(u) du = \sum_{i=0}^0 a_i p(u_i) = a_0 p(u_0)$$

Lo de la integración exacta de polinomios nos ayuda mucho para lograr determinar la regla, porque sabemos que se debe de cumplir si o si en tal grado de los polinomios, consideremos entonces los de grado 0, es decir constantes:

Recordemos por Álgebra Lineal que: $\langle 1 \rangle = \mathbb{P}_0$, por lo que nos va a interesar conocer cómo se comporta el generador, así que:

$$2 = \int_{-1}^1 du = a_0$$

Así luego reemplazando en la fórmula resultante del cambio de intervalo queda:

$$\int_a^b f(x) dx \approx \frac{b-a}{2} \cdot 2 \cdot f\left(\frac{a+b}{2}\right) = (b-a)f\left(\frac{a+b}{2}\right)$$

Y recordemos que solo es exacta cuándo $f(x) \in \mathbb{P}_0$.

Con esto termina la verificación de la fórmula del rectángulo.

3.5.7. Verificación del Método del Trapecio

Para el caso de la fórmula del trapecio, recordemos que $\mathbb{P}_1 = \langle 1, x \rangle$, por lo que procederemos a evaluar la integral en los generadores, además es muy importante tener en cuenta la fórmula (36), dado que nos va a dar una idea de lo que hay que realizar. Los pesos van a estar dados por:

²⁰Se asumen equiespaciados.

$$\int_{-1}^1 p(u) du = \sum_{i=0}^1 a_i p(u_i) = a_0 p(u_0) + a_1 p(u_1)$$

- $p(u) = 1: 2 = \int_{-1}^1 du = a_0 + a_1$
- $p(u) = u: 0 = \int_{-1}^1 u du = -a_0 + a_1$

Resolviendo el sistema de ecuaciones obtenemos que $a_0 = a_1 = 1$, así la fórmula de integración resultará en:

$$\int_a^b f(x) dx \approx \frac{b-a}{2} \sum_{i=0}^1 a_i f(x_i) = \frac{b-a}{2} (1 \cdot f(a) + 1 \cdot f(b)) = \frac{b-a}{2} (f(a) + f(b))$$

Y recordemos que solo es exacta cuando $f(x) \in \mathbb{P}_1$.

Con esto termina la verificación de la fórmula del trapecio.

3.5.8. Verificación del Método de Simpson

Cómo hemos venido haciendo, usaremos la fórmula:

$$\int_{-1}^1 p(u) du = \sum_{i=0}^2 a_i p(u_i) = a_0 p(u_0) + a_1 p(u_1) + a_2 p(u_2)$$

Recordemos además que $\mathbb{P}_2 = \langle 1, x, x^2 \rangle$, por lo que procederemos a evaluar en estos elementos de la base:

- $p(u) = 1: 2 = \int_{-1}^1 du = a_0 + a_1 + a_2$
- $p(u) = u: 0 = \int_{-1}^1 u du = -a_0 + a_2$
- $p(u) = u^2: \frac{2}{3} = \int_{-1}^1 u^2 du = a_0 + a_2$

Resolviendo el sistema obtenemos que: $a_0 = a_2 = \frac{1}{3}, a_1 = \frac{4}{3}$, así la fórmula de integración resultará en:

$$\int_a^b f(x) dx \approx \frac{b-a}{2} \sum_{i=0}^2 a_i f(x_i) = \frac{b-a}{6} \left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right)$$

Y recordemos que solo es exacta cuándo $f(x) \in \mathbb{P}_2$.

Con esto termina la verificación de la fórmula de Simpson.

Veamos ahora más fórmulas de Newton-Cotes para ir aumentando la exactitud:

n	h	a_0	a_1	a_2	a_3	a_4	a_5
0	$b-a$	1					
1	$b-a$	$\frac{1}{2}$	$\frac{1}{2}$				
2	$\frac{b-a}{2}$	$\frac{1}{3}$	$\frac{4}{3}$	$\frac{1}{3}$			
3	$\frac{b-a}{3}$	$\frac{3}{8}$	$\frac{9}{8}$	$\frac{9}{8}$	$\frac{3}{8}$		
4	$\frac{b-a}{4}$	$\frac{14}{45}$	$\frac{64}{45}$	$\frac{24}{45}$	$\frac{64}{45}$	$\frac{14}{45}$	
5	$\frac{b-a}{5}$	$\frac{95}{288}$	$\frac{375}{288}$	$\frac{250}{288}$	$\frac{250}{288}$	$\frac{375}{288}$	$\frac{95}{288}$

Cuadro 7: Fórmulas de Pesos de Newton-Cotes

Con todo esto definido, veamos en acción varias de estas fórmulas:

Nuestro objetivo será calcular mediante aproximaciones a: $\int_0^2 \cos(x^2) dx \approx 0,461461$

1. Regla del Rectángulo:

$$\int_0^2 \cos(x^2) dx \approx (2-0)f(1) = 2 \cdot 0,540302 = 1,080604$$

.

2. Regla del Trapecio:

$$\int_0^2 \cos(x^2) dx \approx \frac{2-0}{2}(f(0) + f(2)) = 1 - 0,653644 = 0,346356$$

.

3. Regla de Simpson:

$$\int_0^2 \cos(x^2) dx \approx \frac{2-0}{6}(f(0) + f(1) + f(2)) = \frac{1}{3}(1 + 0,540302 - 0,653644) = 0,295553$$

4. Regla de 3/8 de Simpson:

$$\int_0^2 \cos(x^2) dx \approx \frac{1}{24}(3f(0) + 9f(2/3) + 9f(4/3) + 3f(2)) = 0,304798$$

En general estos números se deben a que luego del número 1, entra a una parte negativa el área de la integral, por lo que los números tienden a desacomodarse, pero sí seguimos la cadencia obtenemos que el siguiente número, es decir con $n + 1 = 5$ da: 0,440261, y con $n + 1 = 6$ obtenemos: 0,448470, es decir cada vez va mejorando el proceso del resultado de Integración.

Ahora veremos unas fórmulas muy parecidas a estas que permiten con estos métodos ir mejorando aún más el resultado de la integración, analizaremos solo Trapecio y Simpson, que son los más usados.

3.6. Integración Numérica (Métodos Compuestos)

4. Parcial 4

5. Bibliografía sugerida

Durante el desarrollo del curso vamos a ir usando dos libros en su mayoría, los cuáles son:

- Kress, Reiner. *Numerical Analysis*.
- Burden, Richard *Numerical Analysis*.

De igual modo existen más libros que pueden servirnos para orientar nuestro estudio en el campo del Análisis Numérico:

- Sauer, Timothy. *Análisis Numérico*.
- Quarteroni, Alfio M. and Salieri, Fausto and Sacco, Riccardo *Numerical Mathematics*.
- Trefethen L; Bau David *Numerical Linear Algebra*.
- Mathews, John H. Fink, Kurtis D. *Métodos Numéricos con MATLAB*.

Estos son unos de los libros sugeridos por el profesor, de modo que nos permitirán además de los talleres que él deja, expandir nuestro desarrollo por el campo del Análisis Numérico.