



Métodos Numéricos II

Ecuaciones diferenciales ordinarias

Método del trapecio

Resumen

En este texto puedes incluir un resumen del documento. Este informa al lector sobre el contenido del texto, indicando el objetivo del mismo y qué se puede aprender de él.

Andrés Herrera Poyatos
Javier Poyatos Amador
Rodrigo Raya Castellano
Universidad de Granada

Índice

1. Motivación: ecuaciones diferenciales ordinarias de primer orden	2
2. Definiciones y resultados previos.	4
3. Introducción al método del trapecio	7
4. Método del trapecio explícito	8
5. Método del trapecio iterativo	8
5.1. Estudio del error	9
6. Artículo de investigación	10
7. Conclusión	10

1. Motivación: ecuaciones diferenciales ordinarias de primer orden

Definición 1.1. Dada una función $f : \Omega \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ continua, un problema de valores iniciales de primer orden consiste en encontrar aquellas funciones $y : [a, b] \rightarrow \mathbb{R}$ de clase 1 que verifiquen $G(y) \subset \Omega$, $y'(t) = f(t, y(t)) \forall t \in [a, b]$ y la condición inicial $y(t_0) = y_0$, donde $t_0 \in [a, b]$.

De forma simplificada, un problema de valores iniciales se representa de la siguiente forma:

$$\begin{cases} y'(t) = f(t, y(t)) \\ y(t_0) = y_0 \\ t \in [a, b] \end{cases}$$

Resolver de forma exacta un problema de valores iniciales es muy difícil. Existen ecuaciones diferenciales como $y'(t)^2 + y(t)^2 + 1 = 0$ de las cuales no se conoce una solución exacta. Sin embargo, existen múltiples resultados que permiten asegurar la existencia y unicidad de soluciones de la ecuación diferencial incluso cuando no se puedan obtener soluciones explícitamente.

Uno de los objetivos de la teoría del Análisis Numérico en el campo de las ecuaciones diferenciales ordinarias es resolver de forma aproximada problemas de valores iniciales una vez se conoce la existencia y unicidad de soluciones. Para ello, una técnica habitual son los métodos de discretización que son cualquier método numérico que trata obtener valores aproximados de una sucesión y_n de la solución $y(x)$ en los diferentes nodos x_n que se obtienen al realizar la partición del intervalo en el que se trabaja. Se toman $n + 1$ puntos $t_0, t_1, \dots, t_n \in [a, b]$ con $a = t_0 < t_1 < \dots < t_n = b$ y se intenta aproximar la imagen de la solución del problema de valores iniciales en estos puntos. Es habitual denotar a las aproximaciones obtenidas w_i . En lo que sigue, se trabajará siempre con condiciones iniciales sobre $y(a)$ para simplificar la definición de los métodos de discretización.

Repasar estos dos párrafos, no sé si queda claro. La primera idea intuitiva para resolver este problema consiste en interpretar la primera ecuación como un campo vectorial aprovechando la definición de derivada como aproximación lineal de la función en un punto. Esto es, f le asigna a cada punto la dirección en la que varía cualquier solución del problema que pase por ese punto.

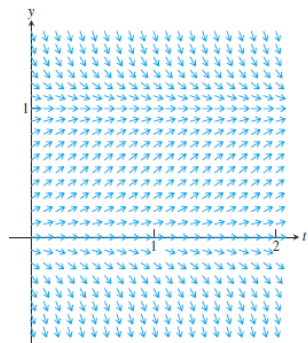


Figura 1: Representación del campo vectorial asociado a la ecuación logística $y'(t) = cy(t)(1 - y(t))$.

Si conocemos la imagen de la solución y en un punto t_{i-1} , entonces sabemos que en ese punto la función variará en la dirección dada por el campo vectorial comentado previamente. Nótese que esta es la

dirección de la recta tangente a y en t_{i-1} . Podemos utilizar la imagen de esta recta tangente en t_{i+1} para aproximar $y(t_{i+1})$. Repitiendo el proceso para aproximar $y(t_{i+2})$ a partir de w_{i+1} , se obtiene el método de Euler cuya expresión resumida es la siguiente:

$$\begin{cases} w_0 = y_0 \\ h_i = t_{i+1} - t_i \\ w_{i+1} = w_i + h_i f(t_i, w_i) \end{cases}$$

Los mejores resultados se obtienen mediante el uso de puntos equidistantes, esto es, $h = \frac{b-a}{n}$ y $t_i = a + ih \forall i = 0 \dots n$. En el resto del texto se trabajará siempre con puntos equidistantes. El estudio del método de Euler concluye que el error global de aproximación cometido es $O(h)$, esto es, existe $M \geq 0$ tal que $|y_i - w_i| \leq \frac{Mh}{2}$ para todo $i = 0 \dots n$.

A priori, puede parecer que el método de Euler es válido en cualquier aplicación simplemente reduciendo el valor de h , esto es, aproximando un mayor número de puntos. Sin embargo, a continuación estudiaremos un ejemplo para el cual el método de Euler requiere una excesiva cantidad de puntos para obtener un error de aproximación aceptable.

EJEMPLO DEL MÉTODO DE EULER. MEJORAR

EJEMPLO 1.1: Considérese el siguiente problema de valores iniciales

$$\begin{cases} y'(t) = -4t^3 y^2 \\ y(-10) = 1/10001 \\ t \in [-10, 0] \end{cases}$$

La solución exacta de este problema es $y(t) = \frac{1}{1+t^4}$. La Tabla 1 muestra los resultados de aproximación obtenidos por el método de Euler en $y(0)$ para distintos valores de h . Se observa que la aproximación obtenida deja mucho que desear a pesar de haber llegado a utilizar hasta un millón de puntos.

N	h	w_n
100	0.1	0.00390138
1000	0.01	0.03085162
5000	0.002	0.13282140
7500	0.0013	0.18614311
10000	0.001	0.23325153

Tabla 1: Ejemplo práctico del método de Euler.

El objetivo de este trabajo es introducir un método de discretización para aproximar soluciones de problemas de valores iniciales que presente menor error que el método de Euler. El método en cuestión se conoce como método del trapecio y presenta dos variantes denominadas explícita e iterativa.

El trabajo se organiza como sigue. En la Sección 2 se explican algunas definiciones y resultados sobre la existencia, unicidad y estabilidad de las soluciones que serán necesarios posteriormente. En la Sección 3 se muestra la idea a partir de la cual surgen las diferentes versiones del método del trapecio. Posteriormente, en las Secciones 4 y 5 se desarrollan los métodos del trapecio explícito e iterativo respectivamente. Ambos métodos se estudian desde una doble perspectiva: cálculo del error y estabilidad de las soluciones. Además, se muestra cómo los errores de redondeo afectan al comportamiento del

método. En la Sección 6 se resume el artículo de investigación "Nombre del artículo", que pone de manifiesto que la resolución de problemas de valores iniciales sigue siendo un tema abierto en la actualidad. Por último, en la Sección 7 se destacan las conclusiones obtenidas y las ventajas y desventajas del método del trapecio.

2. Definiciones y resultados previos.

En esta sección se proporcionan las definiciones y resultados que se necesitan para el estudio del método del trapecio. En primer lugar, una de las hipótesis con las que se suele trabajar para problemas de valores iniciales es que la función f sea lipschitziana en la segunda variable.

Definición 2.1. Sea $S = [a, b] \times [\alpha, \beta] \subseteq \mathbb{R}^2$. Se dice que una función $f(t, y)$ es de Lipschitz respecto a la segunda variable, y , si existe una constante L , llamada constante de Lipschitz de forma que $|f(t, y_1) - f(t, y_2)| \leq L * |y_1 - y_2|, \forall (t, y_1), (t, y_2) \in S$.

No tiene sentido aplicar un método numérico para resolver un problema de valores iniciales que no tenga solución. Por tanto, los resultados que garanticen la existencia de soluciones al problema son fundamentales en este contexto. Además, si el problema admitiese varias soluciones distintas, entonces el método puede no comportarse correctamente pues no se sabe cuál debe calcular. Por tanto, la unicidad de soluciones también es un concepto que se debe estudiar en profundidad. El resultado de este estudio se resume en el teorema de existencia y unicidad de soluciones, que utiliza como hipótesis fundamental el concepto de función lipschitziana en la segunda variable.

Teorema 2.1. (*Existencia y unicidad de soluciones*) Sea $f : [a, b] \times [\alpha, \beta] \rightarrow \mathbb{R}$ e $y_0 \in (\alpha, \beta)$. Entonces:

1. Si f es Lipschitz en $[a, b] \times [\alpha, \beta]$, entonces existe $c \in [a, b]$ tal que el problema de valores iniciales:

$$\begin{cases} y'(t) = f(t, y(t)) \\ y(a) = y_0 \\ t \in [a, c] \end{cases}$$

tiene exactamente una solución.

2. Si f es Lipschitz en $[a, b] \times]-\infty, \infty[$, entonces existe exactamente una solución en $[a, b]$

Demostración en caso de tener tiempo.

Nótese que el resultado es válido para cualquier condición inicial escogida. Esto es, la existencia y unicidad solamente depende de f . De aquí en adelante siempre se supondrá que el problema de valores iniciales a resolver tiene solución y que esta es única. En la práctica este hecho es algo que habrá que comprobar mediante el Teorema 2.1. Bajo hipótesis de existencia y unicidad se pueden definir y_i como los valores que toma la solución en los puntos $t_i = a + ih$ para todo $i = 0 \dots n$, donde $h = \frac{b-a}{n}$.

El estudio de los métodos numéricos para problemas de valores iniciales se centra en la acotación de los errores cometidos y en el análisis de la estabilidad de los métodos. Las demostraciones de resultados asociados a estos conceptos suelen requerir el uso de múltiples desigualdades. El siguiente resultado, basado en la desigualdad de Gronwall, proporciona una de las desigualdades con más aplicaciones en esta área.

Teorema 2.2. Sean dos soluciones $y(t), z(t)$ de los problemas de valores iniciales con ecuación $y'(t) = f(t, y(t))$ y condiciones iniciales $y(a)$ y $z(a)$, respectivamente. Supóngase que f es de Lipschitz en $[a, b] \times [\alpha, \beta]$. Entonces $|y(t) - z(t)| \leq e^{L(t-a)}|y(a) - z(a)|$ donde L es la constante de Lipschitz de f .

Un método será mejor que otro cuanto menor error presenten las aproximaciones obtenidas. Sin embargo, el concepto de error se puede ampliar introduciendo los errores locales y globales.

Definición 2.2. Sean w_i los valores estimados en los puntos t_i por cierto método de aproximación. Sea también z_i el valor de la solución exacta en t_i para el problema de valores iniciales

$$\begin{cases} y'(t) = f(t, y(t)) \\ y(t_{i-1}) = w_{i-1} \\ t \in [t_{i-1}, t_i] \end{cases}$$

Se definen los siguientes errores:

- Error global de truncatura o error acumulado en el nodo i -ésimo: $g_i = |y_i - w_i|$
- Error local de truncatura o error en un paso: $e_i = |z_i - w_i|$

El error global cometido en cualquier paso puede entenderse por la suma del error local y el error global del paso previo amplificado. Este hecho se puede visualizar en la Figura 2, que ejemplifica los conceptos de errores locales y globales.

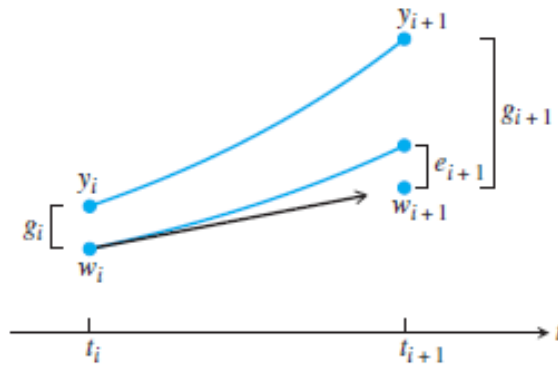


Figura 2: Representación gráfica de los errores locales y globales.

La relación entre errores locales y errores globales viene dada por el siguiente teorema:

Teorema 2.3. Supóngase que la función f es lipschitziana en la segunda variable con constante de Lipschitz L . Además, supóngase que existen $C \geq 0$ y $k \in \mathbb{N}$ tales que los errores locales verifican $e_i \leq Ch^{k+1}$ para todo $i = 0 \dots n$. Entonces, se verifica la siguiente desigualdad para los errores globales

$$g_i \leq \frac{Ch^k}{L} (e^{L(t_i - a)} - 1) \quad (1)$$

Demostración. En breve. □

La siguiente definición pone nombre a las acotaciones del Teorema 2.3.

Definición 2.3. Considérese un método de discretización para problemas de valores iniciales. Entonces:

1. El método es localmente de orden k si existe una constante $C \geq 0$ tal que $e_i \leq Ch^k$ para todo $i = 0 \dots n$.

2. El método es de orden k si existe una constante $C \geq 0$ tal que $g_i \leq Ch^k$ para todo $i = 0 \dots n$.

Las constantes de la definición previa dependerán del problema de valores iniciales en cuestión. Nótese que el Teorema 2.3 está diciendo si un método es localmente de orden $k + 1$, entonces es de orden k . Como aplicación directa de este teorema se obtiene fácilmente el orden del método de Euler.

Teorema 2.4. *Supóngase que $f : [a, b] \times]-\infty, +\infty[\rightarrow \mathbb{R}$ es derivable y lipschitziana en la segunda variable. Entonces, el método de Euler es localmente de orden 2. Consecuentemente, el método de Euler es de orden 1.*

Demostración. Sea y la solución del problema de valores iniciales para $y(a) = y_0$. Fijemos $i = 1 \dots n$ y sea z la solución del problema de valores iniciales para $z(t_{i-1}) = w_{i-1}$. Por inducción, z es de clase infinito. El teorema de Taylor para orden 2 proporciona la siguiente igualdad para cualquier

$$z_i = w_{i-1} + hf(t_{i-1}, w_{i-1}) + \frac{h^2}{2} z''(\xi_i) = w_i + \frac{h^2}{2} z''(\xi_i) \quad (2)$$

donde $\xi_i \in [t_{i-1}, t_i]$. Por tanto, si se utiliza esta igualdad en la expresión del error local se tiene

$$e_i = |z_i - w_i| = \left| \frac{h^2}{2} z''(\xi_i) \right| \leq \frac{M_i}{2} h^2 \quad (3)$$

donde $M_i = \max\{z''(t) : t \in [t_{i-1}, t_i]\}$. Tomando $M = \max_{i=1 \dots n} M_i$, se tiene que el método de Euler es localmente de orden 2 como se quería. La prueba la cierra la aplicación del Teorema 2.3. \square

FORMATEAR EL SIGUIENTE RESULTADO CUANDO SE INTRODUZCA ESTABILIDAD.

Para la demostración del método de Euler probaremos la consistencia y la estabilidad para concluir que es convergente.

Teorema 2.5. *El método de Euler converge para cualquier PVI donde f satisface la condición de Lipschitz y la solución y es C^2 .*

Demostración. Por (??), probado en la demostración del error de truncamiento global, tenemos que:

$$\begin{aligned} |y(t_i) - y_i| &\leq e^{ihL} |y(t_0) - y_0| + \frac{e^{ihL} - 1}{hL} \frac{h^2 M}{2} = e^{ihL} |y(t_0) - y_0| + \frac{e^{ihL} - 1}{L} \tau_{i-1} \leq \\ &\leq^{TL} |y(t_0) - y_0| + \frac{e^{TL} - 1}{L} \max_{1 \leq i \leq n} |\tau_{i-1}| \end{aligned}$$

para $0 \leq t_i = ih + 0 \leq T$, con τ_i el residuo y asumiendo que y es C^2 e $|y''| \leq M$. Esto muestra la *estabilidad*, es decir, errores en la solución numérica están acotadas independientemente del tamaño de paso.

Por (??), τ_i satisface

$$|\tau_i| \leq \frac{hM}{2}$$

Esta condición se denomina *consistencia* (Control del residuo).

La consistencia da una cota local y la estabilidad nos permite concluir la *convergencia*:

$$|y(t_n) - y_n| \leq e^{LT} |y(t_0) - y_0| + \frac{e^{TL} - 1}{L} \frac{hM}{2}$$

□

3. Introducción al método del trapecio

El método del trapecio se basa en la siguiente proposición:

Proposición 3.1. *Considérese el problema de valores iniciales dado por la ecuación diferencial $y'(t) = f(t, y(t))$ sobre $[a, b]$ y la condición $y(t_0) = y_0$. Entonces, son equivalentes:*

1. *y es una solución del problema de valores iniciales.*

2. $y(t) = y_0 + \int_{t_0}^t f(s, y(s)) ds \quad \forall t \in [a, b]$

Demostración. Es consecuencia directa del Teorema Fundamental del Cálculo. □

Utilizando la Proposición 3.1, si un PVI con condición inicial $t_0 = a$, $y(t_0) = y_0$ tiene solución única, entonces esta es la única solución de la siguiente ecuación

$$y(t) = y_0 + \int_{t_0}^t f(s, y(s)) ds \quad (4)$$

En este contexto se pueden aplicar los métodos de integración numérica para aproximar la integral que aparece en la segunda igualdad. Para ello supóngase que f es diferenciable. En tal caso una obvia inducción concluye que y es de clase infinito. Por tanto, se puede utilizar la fórmula del trapecio para integración numérica, obteniendo la siguiente igualdad

$$y(t_1) = y_0 + \frac{h}{2} [f(t_0, y_0) + f(t_1, y(t_1))] - \frac{h^3}{12} y^{(3)}(\xi) \quad (5)$$

donde $\xi \in [t_0, t_1]$. Ignorando el último sumando se obtiene la aproximación dada en (6), que tiene error $-\frac{h^3}{12} y^{(3)}(\xi)$.

$$y(t_1) \approx w_1 = w_0 + \frac{h}{2} [f(t_0, w_0) + f(t_1, y(t_1))] \quad (6)$$

El problema reside en que para aproximar el valor de y en t_1 se debe conocer previamente dicho valor. En este contexto se plantean dos soluciones diferentes obteniendo dos métodos, denominados método del trapecio explícito e iterativo respectivamente. En el resto del texto se desarrollan sendos métodos, proporcionando el error teórico cometido y resultados de convergencia y estabilidad.

4. Método del trapecio explícito

Recuérdese en este punto el método de Euler para ecuaciones diferenciales ordinarias que se comentó en la Sección 1. Denotemos w'_i a las aproximaciones obtenidas por este método. El valor de la solución en cada punto se aproxima a partir del anterior mediante la siguiente expresión:

$$y(t_{k+1}) \approx w'_{i+1} = w'_i + hf(t_i, w'_i))$$

Se comentó previamente que el problema de la aproximación (6) reside en que el valor a aproximar aparece en el segundo miembro de la expresión. Para solventar este hecho se puede utilizar la aproximación dada por el método de Euler en su lugar. De esta forma se obtiene la siguiente aproximación:

$$y(t_{k+1}) \approx w_{i+1} = w_i + \frac{h}{2} [f(t_i, w_i) + f(t_i + h, w_i + hf(t_i, w_i))] \quad (7)$$

Sean $S_L = hf(t_i, w_i)$ y $S_R = hf(t_{i+1}, w'_{i+1})$. El método de Euler obtiene (t_{i+1}, w'_{i+1}) sumándole S_L a (t_i, w_i) . Por su parte, el método del trapecio explícito obtiene (t_{i+1}, w_{i+1}) como (t_i, w_i) más la media de S_L y S_R . La Figura 3 muestra este hecho de forma visual.

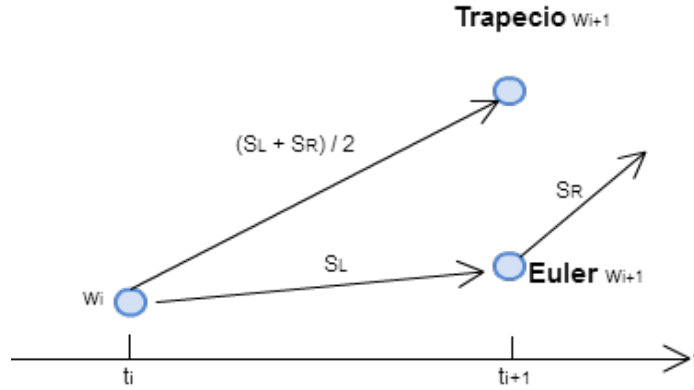


Figura 3: Esquema visual del método del trapecio explícito.

5. Método del trapecio iterativo

Supóngase que se han calculado las i primeras aproximaciones de la función y y que, además, se ha obtenido una aproximación inicial para y_i . Esta aproximación se denota $w_i^{(0)}$. Si la aproximación es buena, la expresión (6) sugiere evaluar el miembro de la derecha utilizando $w_i^{(0)}$ en lugar de y_i para obtener una nueva aproximación. De esta forma se define

$$w_i^{(j+1)} = y_{i-1} + \frac{h}{2} [f(t_{i-1}, y_{i-1}) + f(t_i, w_i^{(j)})] \quad (8)$$

La primera pregunta que surge es si la sucesión $\{w_i^{(j)}\}$ converge. Este hecho se estudia en la Sección ??, donde se probará la convergencia bajo determinadas condiciones. En tal caso, se toma $w_i = \lim w_i^{(j)}$ como aproximación de y_i . Además, se probará que el error cometido utilizando esta aproximación es $O(h^3)$.

La segunda cuestión a tratar es cómo aplicar este nuevo método para obtener aproximaciones para todos los t_i . Nótese que en la definición previa se suponía haber calculado y_{i-1} previamente. La idea es simple, utilizar la aproximación w_{i-1} como y_{i-1} en la expresión (8). De esta forma se obtiene el siguiente método

$$w_i^{(j+1)} = w_{i-1} + \frac{h}{2} [f(t_{i-1}, y_{i-1}) + f(t_i, w_i^{(j)})] \quad (9)$$

En tal caso se está cometiendo un nuevo error de aproximación.

5.1. Estudio del error

El resultado del estudio de esta cuestión viene dado en la siguiente proposición.

Proposición 5.1. *Si la función f es Lipschitziana en la segunda variable con constante de Lipschitz L y se verifica $\frac{hL}{2} < 1$, entonces la sucesión $\{w_i^{(j)}\}$ converge cuando $j \rightarrow +\infty$. Denotando $w_i = \lim w_i^{(j)}$, se tiene además $|y_i - w_i| = O(h^3)$.*

Demostración. La prueba se centra en comprobar que $\{w_i^{(j)}\}$ es una sucesión de Cauchy. En primer lugar, se estudia $\{w_i^{(j+1)} - w_i^{(j)}\}$. Utilizando la condición de Lipschitz

$$\left| w_i^{(j+1)} - w_i^{(j)} \right| \leq \frac{h}{2} \left| f(t_i, w_i^{(j)}) - f(t_i, w_i^{(j-1)}) \right| \leq \frac{hL}{2} \left| w_i^{(j)} - w_i^{(j-1)} \right| \leq \dots \leq \left(\frac{hL}{2} \right)^j \left| w_i^{(1)} - w_i^{(0)} \right|$$

Posteriormente se utiliza la desigualdad obtenida para conseguir acotar $\left| w_i^{(j+q)} - w_i^{(j)} \right|$.

$$\left| w_i^{(j+q)} - w_i^{(j)} \right| \leq \sum_{k=j}^{j+q} \left| w_i^{(k)} - w_i^{(k-1)} \right| \leq \sum_{k=j}^{j+q} \left(\frac{hL}{2} \right)^{k-1} \left| w_i^{(1)} - w_i^{(0)} \right| = \left| w_i^{(1)} - w_i^{(0)} \right| \sum_{k=j}^{j+q} \left(\frac{hL}{2} \right)^{k-1}$$

Como $\frac{hL}{2} < 1$, se tiene que la serie $\sum_{j \geq 0} \left(\frac{hL}{2} \right)^j$ converge. Por consiguiente, es de Cauchy. Aplicando este hecho en la expresión previa se obtiene que $\{w_i^{(j)}\}$ es una sucesión de Cauchy.

Nótese que tomando límites en (8) se obtiene la siguiente igualdad

$$w_i = y_{i-1} + \frac{h}{2} [f(t_{i-1}, y_{i-1}) + f(t_i, w_i)] \quad (10)$$

Por último, se estudia $|y_i - w_i|$. Restando las expresiones (5) y (10) se obtiene

$$|y_i - w_i| = \left| \frac{h}{2} [f(t_i, y_i) - f(t_i, w_i)] - \frac{h^3}{12} y^{(3)}(\xi) \right| \leq \frac{h}{2} |f(t_i, y_i) - f(t_i, w_i)| + \frac{h^3}{12} |y^{(3)}(\xi)| \leq$$

$$\frac{hL}{2} |y_i - w_i| + \frac{h^3}{12} \left| y^{(3)}(\xi) \right|$$

Por tanto, agrupando los $|y_i - w_i|$ se tiene la siguiente desigualdad

$$|y_i - w_i| \leq \frac{2}{2 - HL} \frac{h^3}{12} \left| y^{(3)}(\xi) \right| = O(h^3)$$

□

6. Artículo de investigación

7. Conclusión