

Capítulo 3. Interpolación polinómica avanzada**Interpolación de Hermite**

Hasta ahora hemos construido polinomios interpolantes utilizando información de la forma:

$$(x_i, f(x_i))$$

Es decir, conocemos el valor que toma una función en determinadas posiciones y buscamos un polinomio que reproduzca exactamente esos valores.

Por ejemplo, si conocemos:

$$f(x_0) = y_0, f(x_1) = y_1,$$

podemos imponer:

$$P(x_0) = f(x_0), P(x_1) = f(x_1).$$

Estas condiciones nos indican por qué puntos debe pasar el polinomio interpolante. Pero en algunos problemas podemos disponer de información adicional sobre la función.

Supongamos que, además de conocer:

$$f(x_0), f(x_1),$$

conocemos también:

$$f'(x_0), f'(x_1).$$

Surge entonces una pregunta:

¿Podemos utilizar también esa información para construir el polinomio interpolante?

La respuesta conduce a la **interpolación de Hermite**.

¿Qué información adicional proporciona una derivada?

Aquí conviene observar simultáneamente el problema desde dos perspectivas.

Desde el álgebra	Desde la geometría
Conocer $f(x_i)$ permite imponer $H(x_i) = f(x_i)$.	Conocemos un punto por el que debe pasar la curva.
Conocer además $f'(x_i)$ permite imponer $H'(x_i) = f'(x_i)$.	Conocemos también la pendiente que debe tener la curva al pasar por ese punto.

Cada nueva condición aporta Ya no conocemos solamente **dónde** información que debe satisfacer el **está** la función, sino también **cómo se** polinomio. **comporta localmente** en ese punto.

Esta diferencia es fundamental. Si conocemos únicamente dos valores:

$$f(x_0), f(x_1),$$

la interpolación lineal puede construir una recta que pase por ambos puntos.

Pero esos dos valores, por sí solos, **no nos indican las pendientes que posee la función original en los extremos.**

En cambio, si conocemos también:

$$f'(x_0), f'(x_1),$$

disponemos de información sobre la dirección local de la función en esos puntos.

Por lo tanto, Hermite no busca solamente que el polinomio pase por los datos conocidos. Busca que reproduzca también la información disponible sobre sus derivadas.

¿Qué cambia algebraicamente?

Consideremos nuevamente dos puntos x_0 y x_1 . Si conocemos solamente sus valores funcionales tenemos dos condiciones:

$$\begin{aligned} P(x_0) &= f(x_0), \\ P(x_1) &= f(x_1). \end{aligned}$$

Dos condiciones permiten determinar, en el caso más sencillo, los dos coeficientes de un polinomio lineal:

$$P_1(x) = a_0 + a_1x.$$

Ahora supongamos que conocemos también las derivadas.

Queremos que el nuevo polinomio $H(x)$ satisfaga:

$$\begin{aligned} H(x_0) &= f(x_0), \\ H(x_1) &= f(x_1), \end{aligned}$$

pero, además:

$$\begin{aligned} H'(x_0) &= f'(x_0), \\ H'(x_1) &= f'(x_1). \end{aligned}$$

Tenemos ahora **cuatro condiciones independientes** que debe satisfacer el polinomio interpolante.

La pregunta entonces debería ser:

¿Cuál es el polinomio de menor grado que dispone de suficientes coeficientes para satisfacer estas cuatro condiciones?

Razonemos la respuesta de manera progresiva. Un polinomio de grado 1:

$$H_1(x) = a_0 + a_1x$$

tiene **2 coeficientes**. En general, no alcanza para imponer cuatro condiciones.

Uno de grado 2:

$$H_2(x) = a_0 + a_1x + a_2x^2$$

tiene **3 coeficientes**. Tampoco alcanza, en general.

El siguiente es:

$$H_3(x) = a_0 + a_1x + a_2x^2 + a_3x^3,$$

que posee **4 coeficientes**.

Ahora sí tenemos cuatro parámetros que pueden quedar determinados por las cuatro condiciones:

$$\begin{cases} H_3(x_0) = f(x_0) \\ H_3(x_1) = f(x_1) \\ H'_3(x_0) = f'(x_0) \\ H'_3(x_1) = f'(x_1) \end{cases}$$

Por lo tanto, las cuatro condiciones anteriores permiten determinar, en general, un único polinomio de grado **a lo sumo 3**. El grado 3 no se elige arbitrariamente: **surge de la cantidad de información que queremos reproducir**.

¿Y si tenemos más nodos?

El razonamiento anterior puede extenderse a cualquier cantidad de nodos.

En el caso de Hermite que estamos estudiando, en cada nodo x_i conocemos dos datos:

$$f(x_i) \text{ y } f'(x_i),$$

cada nodo aporta **dos condiciones** que debe satisfacer el polinomio interpolante:

- una asociada al valor de la función $H(x_i) = f(x_i)$;
- otra asociada al valor de su primera derivada $H'(x_i) = f'(x_i)$.

A medida que incorporamos nodos, aumenta la cantidad de condiciones que debe satisfacer el polinomio. Esas condiciones permiten determinar los coeficientes necesarios para construirlo.

Podemos seguir entonces esta relación:

nodos → condiciones disponibles → coeficientes → forma y grado del polinomio

Por ejemplo:

Cantidad de nodos	Condiciones disponibles	Coeficientes necesarios	Forma del polinomio en el caso general	Grado máximo
2	4	4	$a_0 + a_1x + a_2x^2 + a_3x^3$	3
3	6	6	$a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5$	5
4	8	8	$a_0 + a_1x + \dots + a_7x^7$	7

Por ejemplo, con **tres nodos** conocemos:

$$\begin{aligned} &f(x_0), f'(x_0), \\ &f(x_1), f'(x_1), \\ &f(x_2), f'(x_2). \end{aligned}$$

Disponemos entonces de seis condiciones.

Para poder satisfacerlas necesitamos, en el caso general, seis coeficientes:

$$a_0, a_1, a_2, a_3, a_4, a_5.$$

Estos coeficientes corresponden a un polinomio de la forma:

$$H_5(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5.$$

Por lo tanto, el grado máximo que puede ser necesario es 5.

La tabla permite reconocer el siguiente patrón:

2 nodos → 4 condiciones → 4 coeficientes → grado máximo 3,
3 nodos → 6 condiciones → 6 coeficientes → grado máximo 5,
4 nodos → 8 condiciones → 8 coeficientes → grado máximo 7.

Para expresar esta relación de manera general, llamemos:

$$N = \text{cantidad de nodos}$$

Como cada nodo aporta dos condiciones, con N nodos disponemos de:

$$2N \text{ condiciones}$$

En el caso general, esas condiciones permiten determinar $2N$ coeficientes.

Un polinomio con $2N$ coeficientes tiene la forma:

$$H(x) = a_0 + a_1x + a_2x^2 + \dots + a_{2N-1}x^{2N-1},$$

por lo que su grado máximo puede ser:

$$2N - 1$$

Por ejemplo, si:

$$N = 3,$$

tenemos tres nodos, seis condiciones y seis coeficientes. Por lo tanto:

$$2N - 1 = 5,$$

que coincide con el polinomio de grado máximo 5 obtenido anteriormente.

En la notación habitual, los nodos suelen escribirse como:

$$x_0, x_1, \dots, x_n.$$

En ese caso, la cantidad de nodos es:

$$N = n + 1.$$

Por esta razón:

$$2N - 1 = 2(n + 1) - 1 = 2n + 1.$$

Así, la relación anterior también puede expresarse de manera compacta como:

$$\deg H \leq 2n + 1$$

Esta última expresión no introduce una idea nueva: simplemente resume, utilizando la notación habitual $x_0, \dots, x_n$, la relación entre la cantidad de nodos, las condiciones disponibles, los coeficientes necesarios y el grado del polinomio de Hermite.

¿Cómo construimos el polinomio de Hermite?

Hasta aquí hemos establecido **qué condiciones debe satisfacer** el polinomio de Hermite. Pero todavía queda una pregunta fundamental:

¿Cómo utilizamos esos valores de las derivadas para construir efectivamente el polinomio?

Si partimos nuevamente del caso del polinomio cúbico, podríamos plantear el polinomio de esta manera:

$$H(x) = a_0 + a_1x + a_2x^2 + a_3x^3$$

y utilizar las cuatro condiciones anteriores para obtener sus coeficientes. Sin embargo, esto nos llevaría nuevamente a plantear y resolver un sistema de

ecuaciones, algo que ya buscamos evitar cuando estudiamos otras formas de construcción del polinomio interpolante.

Hermite puede construirse de una manera particularmente conveniente utilizando la forma de **Newton por diferencias divididas**, que ya conocemos.

La pregunta pasa a ser entonces:

¿Cómo podemos modificar la tabla de diferencias divididas para incorporar también la información de las derivadas?

El problema de incorporar una derivada a la tabla

Recordemos que una diferencia dividida de primer orden entre dos nodos diferentes se calcula como:

$$f[x_i, x_j] = \frac{f(x_j) - f(x_i)}{x_j - x_i}.$$

Por ejemplo:

$$f[x_0, x_1] = \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$

Pero en Hermite disponemos de una información que antes no teníamos:

$$f'(x_i).$$

Necesitamos encontrar una manera de introducir esa información dentro de la misma estructura de diferencias divididas.

Aquí aparece la idea fundamental del método de Hermite:

repetiremos cada nodo para poder incorporar en la tabla tanto el valor de la función como el valor de su derivada.

Por ejemplo, si originalmente tenemos x_0, x_1 , para construir la tabla de Hermite trabajaremos con una nueva sucesión de nodos:

$$z_0 = x_0, z_1 = x_0, z_2 = x_1, z_3 = x_1.$$

Es decir:

$$x_0, x_0, x_1, x_1$$

A primera vista esto parece generar un problema.

Si intentáramos calcular la diferencia dividida correspondiente a dos nodos repetidos utilizando la expresión habitual:

$$f[x_i, x_i] = \frac{f(x_i) - f(x_i)}{x_i - x_i},$$

obtendríamos:

$$\frac{0}{0}$$

que es una indeterminación.

Pero precisamente aquí aparece la relación con la derivada. Consideremos la diferencia dividida:

$$f[x_i, x] = \frac{f(x) - f(x_i)}{x - x_i}.$$

Si hacemos que x se aproxime a x_i :

$$\lim_{x \rightarrow x_i} \frac{f(x) - f(x_i)}{x - x_i},$$

reconocemos la definición de derivada:

$$f'(x_i).$$

Por lo tanto, cuando un nodo aparece repetido, definimos:

$$f[x_i, x_i] = f'(x_i)$$

Esta igualdad es la que permite incorporar la información de la derivada dentro de la tabla de diferencias divididas.

Mientras que en Newton establecemos que

$$f[x_i, x] = \frac{f(x) - f(x_i)}{x - x_i} \quad (x_i \neq x_j)$$

en Hermite, conservamos la misma estructura, pero cuando el nodo está repetido:

$$f[x_i, x] = f'(x_i)$$

Por eso Hermite no abandona las diferencias divididas de Newton. Las amplía para poder trabajar con información adicional sobre la función.

Ejemplo: construcción del polinomio de Hermite mediante diferencias divididas

Consideremos la función:

$$f(x) = \frac{1}{x+1}$$

de la cual conocemos, en los nodos $x_0 = 0$ y $x_1 = 1$, tanto el valor de la función como el de su primera derivada:

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

Nuestro objetivo es construir un polinomio $H(x)$ que no solamente pase por los dos puntos conocidos, sino que además tenga en ellos la misma pendiente que la función:

$$H(0) = 1, H'(0) = -1,$$

$$H(1) = \frac{1}{2}, H'(1) = -\frac{1}{4}.$$

1. Repetimos los nodos

Para incorporar la información de las derivadas a la tabla de diferencias divididas, repetimos cada nodo:

$$z_0 = x_0 = 0,$$

$$z_1 = x_0 = 0,$$

$$z_2 = x_1 = 1,$$

$$z_3 = x_1 = 1.$$

Por lo tanto, la sucesión con la que construiremos la tabla será:

$$\boxed{0, 0, 1, 1}$$

y los valores correspondientes de la función son:

$$f[z_0] = 1, f[z_1] = 1, f[z_2] = \frac{1}{2}, f[z_3] = \frac{1}{2}.$$

La repetición de los valores de f no significa que hayamos obtenido nuevos datos. Estamos preparando la estructura de la tabla para poder incorporar también las derivadas.

2. Calculamos las diferencias divididas de primer orden

Aquí aparece la particularidad de Hermite. Cuando encontramos un nodo repetido, no utilizamos:

$$\frac{f(x_i) - f(x_i)}{x_i - x_i},$$

porque produciría la indeterminación $0/0$. Utilizamos directamente la información de la derivada:

$$f[x_i, x_i] = f'(x_i).$$

Por lo tanto:

$$f[z_0, z_1] = f'(0) = -1.$$

Entre los dos nodos diferentes podemos aplicar la expresión habitual:

$$f[z_1, z_2] = \frac{f[z_2] - f[z_1]}{z_2 - z_1}.$$

Reemplazando:

$$f[z_1, z_2] = \frac{\frac{1}{2} - 1}{1 - 0} = -\frac{1}{2}.$$

Finalmente, como $z_2 = z_3 = 1$:

$$f[z_2, z_3] = f'(1) = -\frac{1}{4}.$$

Ya podemos observar algo importante:

$$\boxed{-1, -\frac{1}{2}, -\frac{1}{4}}$$

no surgieron todos de la misma manera. Los valores extremos provienen de las **derivadas conocidas**, mientras que el central se obtuvo mediante la diferencia dividida ordinaria.

Esta es precisamente la adaptación que realiza Hermite sobre la tabla de Newton.

3. Continuamos con las diferencias divididas de orden superior

A partir de aquí, el procedimiento vuelve a ser el habitual.

Calculamos:

$$f[z_0, z_1, z_2] = \frac{f[z_1, z_2] - f[z_0, z_1]}{z_2 - z_0}.$$

Entonces:

$$f[z_0, z_1, z_2] = \frac{-\frac{1}{2} - (-1)}{1 - 0} = \frac{1}{2}.$$

Para la siguiente:

$$f[z_1, z_2, z_3] = \frac{f[z_2, z_3] - f[z_1, z_2]}{z_3 - z_1},$$

por lo que:

$$f[z_1, z_2, z_3] = \frac{-\frac{1}{4} - \left(-\frac{1}{2}\right)}{1 - 0} = \frac{1}{4}.$$

Finalmente:

$$f[z_0, z_1, z_2, z_3] = \frac{f[z_1, z_2, z_3] - f[z_0, z_1, z_2]}{z_3 - z_0},$$

y obtenemos:

$$f[z_0, z_1, z_2, z_3] = \frac{\frac{1}{4} - \frac{1}{2}}{1 - 0} = -\frac{1}{4}.$$

La tabla completa queda:

z_i	$f[z_i]$	1.ª diferencia	2.ª diferencia	3.ª diferencia
0	1	-1	$\frac{1}{2}$	$-\frac{1}{4}$
0	1	$-\frac{1}{2}$	$\frac{1}{4}$	
1	$\frac{1}{2}$	$-\frac{1}{4}$		
1	$\frac{1}{2}$			

4. Construimos el polinomio

Al igual que en Newton, los coeficientes que necesitamos se encuentran en la diagonal correspondiente de la tabla:

$$f[z_0] = 1,$$

$$f[z_0, z_1] = -1,$$

$$f[z_0, z_1, z_2] = \frac{1}{2},$$

$$f[z_0, z_1, z_2, z_3] = -\frac{1}{4}.$$

Por lo tanto:

$$H(x) = f[z_0] + f[z_0, z_1](x - z_0) + f[z_0, z_1, z_2](x - z_0)(x - z_1) + f[z_0, z_1, z_2, z_3](x - z_0)(x - z_1)(x - z_2).$$

Como:

$$z_0 = 0, z_1 = 0, z_2 = 1,$$

obtenemos:

$$H(x) = 1 - x + \frac{1}{2}x^2 - \frac{1}{4}x^2(x - 1).$$

Podemos dejarlo en esta forma de Newton-Hermite porque permite reconocer de dónde procede cada término. Si lo desarrollamos:

$$H(x) = 1 - x + \frac{3}{4}x^2 - \frac{1}{4}x^3$$

¿Qué logramos con el polinomio obtenido?

A partir de los datos conocidos en los dos nodos:

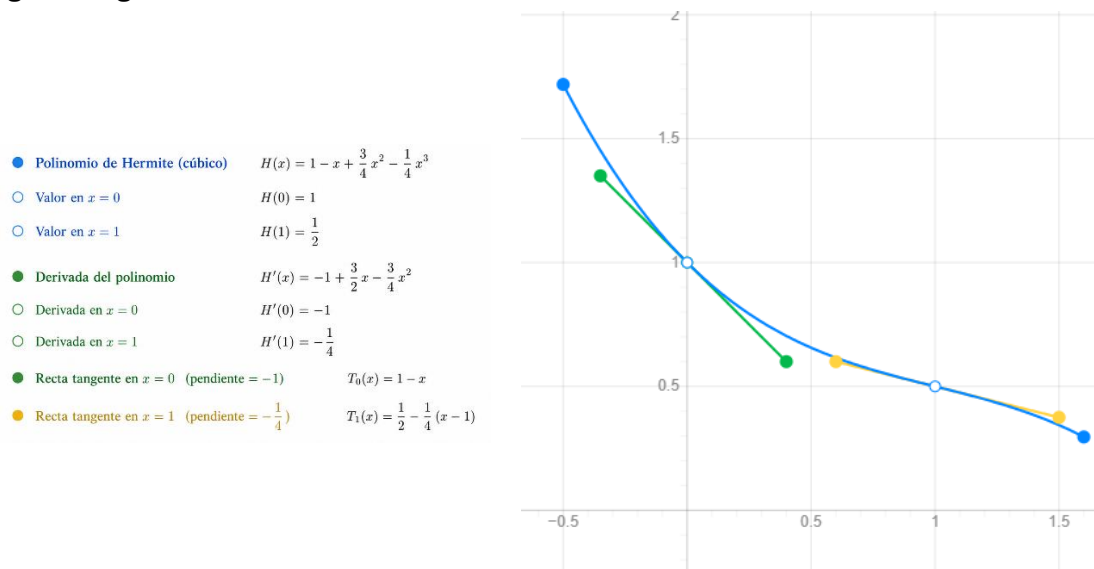
$$\begin{aligned} f(0) &= 1, \quad f'(0) = -1, \\ f(1) &= \frac{1}{2}, \quad f'(1) = -\frac{1}{4}, \end{aligned}$$

construimos el polinomio cúbico de Hermite.

Es interesante observar que, a partir de **cuatro condiciones**, hemos obtenido un polinomio cúbico capaz de satisfacerlas simultáneamente:

$$\begin{aligned} H(0) &= 1, \\ H'(0) &= -1, \\ H(1) &= \frac{1}{2}, \\ H'(1) &= -\frac{1}{4}. \end{aligned}$$

Es decir, el polinomio no solamente pasa por los dos puntos conocidos, sino que además reproduce en ellos las pendientes especificadas, tal como lo representa la siguiente gráfica



Pero esta observación conduce a una pregunta interesante.

Hasta ahora ya sabíamos construir un polinomio interpolante mediante Newton para este ejemplo, utilizando únicamente los valores:

$$(0, 1) \text{ y } \left(1, \frac{1}{2}\right).$$

Con esos mismos dos nodos, Newton produciría:

$$P_1(x) = 1 - \frac{1}{2}x.$$

Ahora disponemos entonces de dos interpolantes contruidos a partir de información diferente. El polinomio de Newton

$$P_1(x) = 1 - \frac{1}{2}x$$

utiliza solamente los valores de la función, mientras que el polinomio de Hermite

$$H_3(x) = 1 - x + \frac{3}{4}x^2 - \frac{1}{4}x^3$$

utiliza además las pendientes conocidas en los nodos. Por ello, Hermite está más condicionado: debe reproducir tanto los valores como las pendientes suministradas.

¿Más información implica necesariamente una mejor aproximación?

Hermite dispone de más información y debe satisfacer una mayor cantidad de condiciones.

Sin embargo, es importante distinguir dos afirmaciones:

más información → más condiciones que debe satisfacer el interpolante

pero esto **no implica automáticamente**:

más información → menor error en todo el intervalo

La calidad de la aproximación depende también de la función, de los nodos utilizados y de la información disponible.

En nuestro ejemplo podemos comprobar qué ocurre porque conocemos la función que originó los datos:

$$f(x) = \frac{1}{x+1}$$

Podemos entonces comparar las tres expresiones:

$$f(x) = \frac{1}{x+1},$$

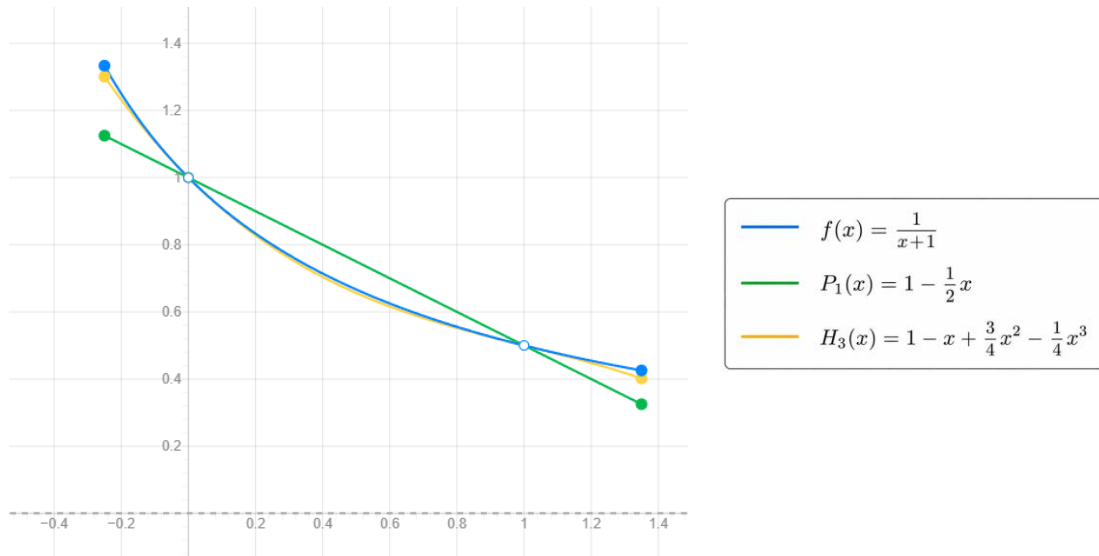
$$P_1(x) = 1 - \frac{1}{2}x,$$

$$H_3(x) = 1 - x + \frac{3}{4}x^2 - \frac{1}{4}x^3.$$

Comparación gráfica

Observa el gráfico que representa las tres curvas:

- función real $f(x)$;
- interpolante lineal de Newton $P_1(x)$;
- interpolante cúbico de Hermite $H_3(x)$;



En el gráfico también se observan los nodos

$$(0, 1), \left(1, \frac{1}{2}\right).$$

Ambos polinomios interpolantes coinciden con la función en los nodos, pero siguen trayectorias diferentes entre ellos.

En este caso particular, el polinomio de Hermite permanece mucho más próximo a la función original dentro del intervalo.

Comparación en un punto interior

Tomemos, por ejemplo $x = 0.5$.

El valor verdadero de la función es:

$$f(0.5) = \frac{1}{1.5} = \frac{2}{3} \approx 0.666667.$$

- **Interpolación de Newton**

$$P_1(0.5) = 1 - \frac{1}{2}(0.5) = 0.75.$$

Por lo tanto, el error absoluto es:

$$E_N = |f(0.5) - P_1(0.5)| = |0.666667 - 0.75|$$

$$E_N \approx 0.083333$$

- **Interpolación de Hermite**

$$H_3(0.5) = 1 - 0.5 + \frac{3}{4}(0.5)^2 - \frac{1}{4}(0.5)^3.$$

Calculando:

$$H_3(0.5) = 0.65625.$$

Entonces:

$$\begin{aligned} E_H &= |f(0.5) - H_3(0.5)| \\ &= |0.666667 - 0.65625| \\ E_H &\approx 0.010417 \end{aligned}$$

Podemos resumir:

Método	Aproximación en $x = 0.5$	Error absoluto
Newton	0.750000	0.083333
Hermite	0.656250	0.010417

y, por lo tanto, **Hermite produce una aproximación considerablemente más cercana al valor verdadero.**

Pero la conclusión correcta no es que Hermite sea siempre mejor.

Lo que podemos afirmar es:

En este ejemplo, la información adicional proporcionada por las derivadas permite construir un interpolante que se aproxima mejor a la función original dentro del intervalo.

La diferencia es importante. Hermite reproduce necesariamente más condiciones en los nodos, pero eso no constituye por sí solo una garantía universal de menor error.

En este caso particular, Hermite no solamente gana en $x = 0.5$; dentro de $0 < x < 1$, su error es menor que el de la recta de Newton. Incluso los errores máximos aproximados en el intervalo son:

$$E_{N,\max} \approx 0.0858,$$

mientras que:

$$E_{H,\max} \approx 0.0106.$$

Esto describe el comportamiento de este caso, no una propiedad general que garantice que Hermite siempre será superior.

Generalización de la interpolación de Hermite

La interpolación de Hermite no se limita necesariamente al conocimiento de $f(x_i)$ y $f'(x_i)$. Si se dispone de información adicional, como derivadas de orden superior, estas también pueden incorporarse como nuevas condiciones del polinomio interpolante.

La idea de los nodos repetidos se generaliza de modo que las diferencias divididas correspondientes permiten incorporar derivadas sucesivas. Por ejemplo:

$$f[x_i, x_i] = f'(x_i)$$

y, para órdenes superiores,

$$f \left[\underbrace{x_i, \dots, x_i}_{k+1 \text{ veces}} \right] = \frac{f^{(k)}(x_i)}{k!}.$$

De esta manera, Hermite permite construir interpolantes que reproduzcan distintos niveles de información disponible sobre la función en los nodos.

Importante: disponer de más información impone más condiciones al interpolante, pero **no garantiza por sí mismo una mejor aproximación de la función en todo el intervalo**. La calidad de la aproximación depende también de la función, de los nodos utilizados y de la información disponible.

Síntesis: Interpolación de Hermite

La interpolación de Hermite extiende el problema de interpolación cuando, además de conocer los valores de una función en determinados nodos, se dispone de información sobre sus derivadas.

Mientras que en la interpolación polinómica estudiada anteriormente se exige que

$$P(x_i) = f(x_i),$$

en Hermite podemos imponer simultáneamente

$$H(x_i) = f(x_i) \text{ y } H'(x_i) = f'(x_i).$$

Geométricamente, esto significa que el interpolante no solo debe pasar por los puntos conocidos, sino que también debe reproducir en ellos las pendientes especificadas.

Cuando se conocen el valor y la primera derivada en N nodos, se tienen:

$$N \text{ condiciones de valor} + N \text{ condiciones de derivada} = 2N \text{ condiciones.}$$

Por lo tanto, en general puede ser necesario un polinomio de grado a lo sumo

$$\boxed{2N - 1}.$$

La construcción mediante diferencias divididas permite conservar la estructura que ya conocemos del método de Newton. La diferencia fundamental es que los nodos se repiten para poder incorporar la información proporcionada por las derivadas:

$$f[x_i, x_i] = f'(x_i).$$

De esta forma, la interpolación de Hermite puede interpretarse como una extensión de la interpolación polinómica que incorpora información adicional sobre el comportamiento local de la función en los nodos.

En el ejemplo desarrollado observamos además que la incorporación de las pendientes produjo una aproximación considerablemente mejor que la obtenida utilizando solamente los valores de la función. Sin embargo, esto no constituye una garantía general:

más información $\Rightarrow$ más condiciones para el interpolante

pero no necesariamente

más información $\Rightarrow$ menor error en todo el intervalo.

La calidad de la aproximación continúa dependiendo de la función considerada, de los nodos utilizados y de la información disponible.

Ideas principales

Idea	Interpolación de Hermite
¿Qué información puede utilizar?	Valores de la función y sus derivadas
¿Qué debe reproducir?	Valores y pendientes especificadas en los nodos
¿Cómo incorporamos $f'(x_i)$ en diferencias divididas?	Mediante nodos repetidos
Relación fundamental	$f[x_i, x_i] = f'(x_i)$
Con N nodos y f, f' conocidos	$2N$ condiciones
Grado máximo que puede ser necesario	$2N - 1$
¿Más información garantiza menor error?	No necesariamente

Fenómeno de Runge: ¿más puntos significan una mejor aproximación?

Hasta ahora hemos construido polinomios interpolantes que reproducen exactamente los valores conocidos de una función. Esto podría llevarnos intuitivamente a pensar que, si aumentamos la cantidad de puntos utilizados para construir el polinomio, obtendremos una aproximación cada vez mejor. Sin embargo, esto **no siempre ocurre**.

El fenómeno de Runge constituye un ejemplo clásico que muestra que la interpolación mediante un único polinomio global puede presentar dificultades cuando se aumenta progresivamente su grado utilizando nodos equidistantes.

Consideremos la función

$$f(x) = \frac{1}{1 + 25x^2} \quad -1 \leq x \leq 1$$

Esta función es continua y suave en todo el intervalo. Su representación gráfica tampoco parece anticipar ninguna dificultad particular.

Observemos qué sucede al aumentar los nodos

Tomemos inicialmente una pequeña cantidad de nodos equidistantes en el intervalo $[-1, 1]$ y construyamos el correspondiente polinomio interpolante. Luego repetiremos el procedimiento aumentando progresivamente la cantidad de nodos.

En las siguientes gráficas se compara la función de Runge

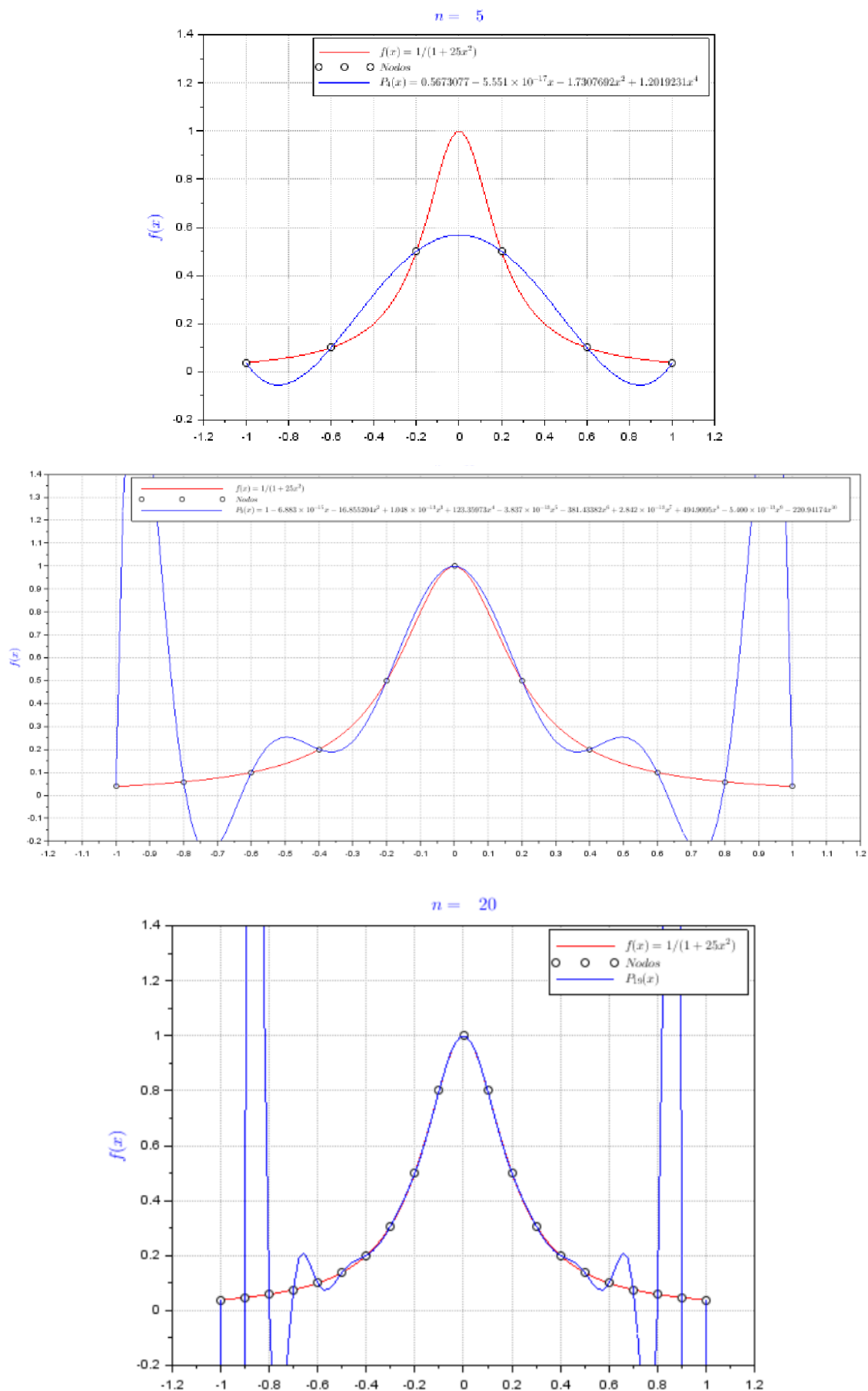
$$f(x) = \frac{1}{1 + 25x^2}$$

con los polinomios interpolantes obtenidos en tres situaciones diferentes:

Caso	Nodos equidistantes	Grado del polinomio obtenido
1	6	4
2	11	9
3	21	19

Recordemos que disponer de N nodos permite construir un polinomio interpolante de **grado a lo sumo** $N - 1$. El grado efectivo puede ser menor si, al obtener el polinomio, algunos de sus coeficientes de mayor grado resultan nulos.

Antes de buscar una explicación matemática para lo que sucede, observemos las gráficas.



En todos los casos, el polinomio interpolante reproduce los valores de la función en los nodos utilizados para construirlo. Sin embargo, **aumentar la cantidad de nodos no produce necesariamente una mejor aproximación en todo el intervalo.**

En particular, las gráficas permiten observar que las diferencias entre la función y el polinomio se hacen especialmente importantes hacia los extremos del intervalo. Al aumentar la cantidad de nodos equidistantes aparecen oscilaciones cada vez más evidentes en esas regiones.

Esta situación resulta inicialmente llamativa:

incorporamos más puntos pero la aproximación no necesariamente mejora

¿Cómo puede ocurrir esto?

Para interpretar lo que estamos observando, resulta útil recuperar una expresión que presentamos con anterioridad al estudiar el error de interpolación:

$$E_n(x) = f(x) - P_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=0}^n (x - x_i).$$

En aquel momento nuestro objetivo era conocer una expresión que permitiera caracterizar el error. Ahora podemos utilizarla para **comprender qué factores intervienen en su comportamiento.**

Podemos reconocer dos partes importantes:

$$E_n(x) = \underbrace{\frac{f^{(n+1)}(\xi)}{(n+1)!}}_{\text{depende de la función}} \cdot \underbrace{\prod_{i=0}^n (x - x_i)}_{\text{depende de la ubicación de los nodos}}.$$

El primer factor contiene una derivada de orden $n + 1$:

$$\frac{f^{(n+1)}(\xi)}{(n+1)!}.$$

Esto nos recuerda que **el comportamiento del error depende también de la función que estamos intentando aproximar.**

Al aumentar el número de nodos generalmente podemos construir polinomios de mayor grado y, en la expresión del error, intervienen derivadas de órdenes cada vez mayores. Por lo tanto, no podemos afirmar únicamente a partir de la cantidad de nodos que el error necesariamente disminuirá.

En otras palabras:

más nodos $\nrightarrow$ necesariamente menor error

La ubicación de los nodos también importa

Consideremos ahora el otro factor:

$$\prod_{i=0}^n (x - x_i).$$

A diferencia del anterior, este término depende directamente de **la ubicación de los nodos** x_i .

Por lo tanto, no solamente importa cuántos nodos utilizamos. También importa **cómo se encuentran distribuidos dentro del intervalo**.

Esto es particularmente relevante en el ejemplo que estamos observando, ya que en las tres experiencias utilizamos nodos **equidistantes**.

Pero si existe error, ¿por qué el polinomio sigue pasando por todos los puntos?

La misma expresión nos permite responder una pregunta interesante.

Supongamos que evaluamos el error exactamente en uno de los nodos, por ejemplo $x = x_j$:

$$E_n(x_j) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=0}^n (x_j - x_i).$$

Dentro del producto necesariamente aparece:

$$(x_j - x_j) = 0.$$

Por lo tanto,

$$\prod_{i=0}^n (x_j - x_i) = 0$$

y entonces

$$E_n(x_j) = 0.$$

Esto explica matemáticamente algo que también podemos reconocer en las gráficas: el polinomio pasa exactamente por los nodos utilizados para construirlo.

Sin embargo, entre esos nodos el producto ya no tiene por qué ser cero y el polinomio puede alejarse de la función.

Por eso debemos distinguir dos ideas que pueden parecer similares, pero no lo son:

interpolarse exactamente los datos $\neq$ aproximar bien la función en todo el intervalo

¿Qué nos enseña entonces el fenómeno de Runge?

El fenómeno de Runge no significa que utilizar polinomios de grado elevado sea siempre incorrecto, ni que agregar nodos necesariamente empeore una interpolación.

Lo que este ejemplo pone en evidencia es que, para determinadas funciones, aumentar progresivamente la cantidad de nodos equidistantes y construir con ellos un único polinomio interpolante global no garantiza que la aproximación mejore en todo el intervalo.

En el ejemplo analizado, las dificultades se hacen especialmente visibles hacia los extremos.

Por lo tanto, debemos abandonar una idea aparentemente intuitiva:

más nodos equidistantes $\Rightarrow$ necesariamente una mejor aproximación

Pero la expresión del error nos dejó además una pista importante: uno de sus factores depende de **la ubicación de los nodos**.

Esto nos lleva a formular una nueva pregunta:

¿Qué ocurriría, si en lugar de utilizar nodos equidistantes, eligiéramos de manera más conveniente su ubicación?

Esta pregunta nos conduce al estudio de los **nodos de Chebyshev**.

Interpolación utilizando nodos de Chebyshev**¿Es necesario que los nodos sean equidistantes?**

En los ejemplos anteriores hemos utilizado frecuentemente nodos equidistantes. Esta elección resulta natural: si debemos seleccionar varios puntos dentro de un intervalo, distribuirlos uniformemente parece una decisión razonable.

Sin embargo, el fenómeno de Runge nos mostró que, para determinadas funciones, aumentar la cantidad de nodos equidistantes no garantiza una mejor aproximación mediante un polinomio interpolante global.

Al analizar la expresión del error observamos además que uno de sus factores depende directamente de la posición de los nodos:

$$E_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \underbrace{\prod_{i=0}^n (x - x_i)}_{\text{depende de la ubicación de los nodos}} .$$

Esto plantea una posibilidad diferente: en lugar de aumentar simplemente la cantidad de nodos

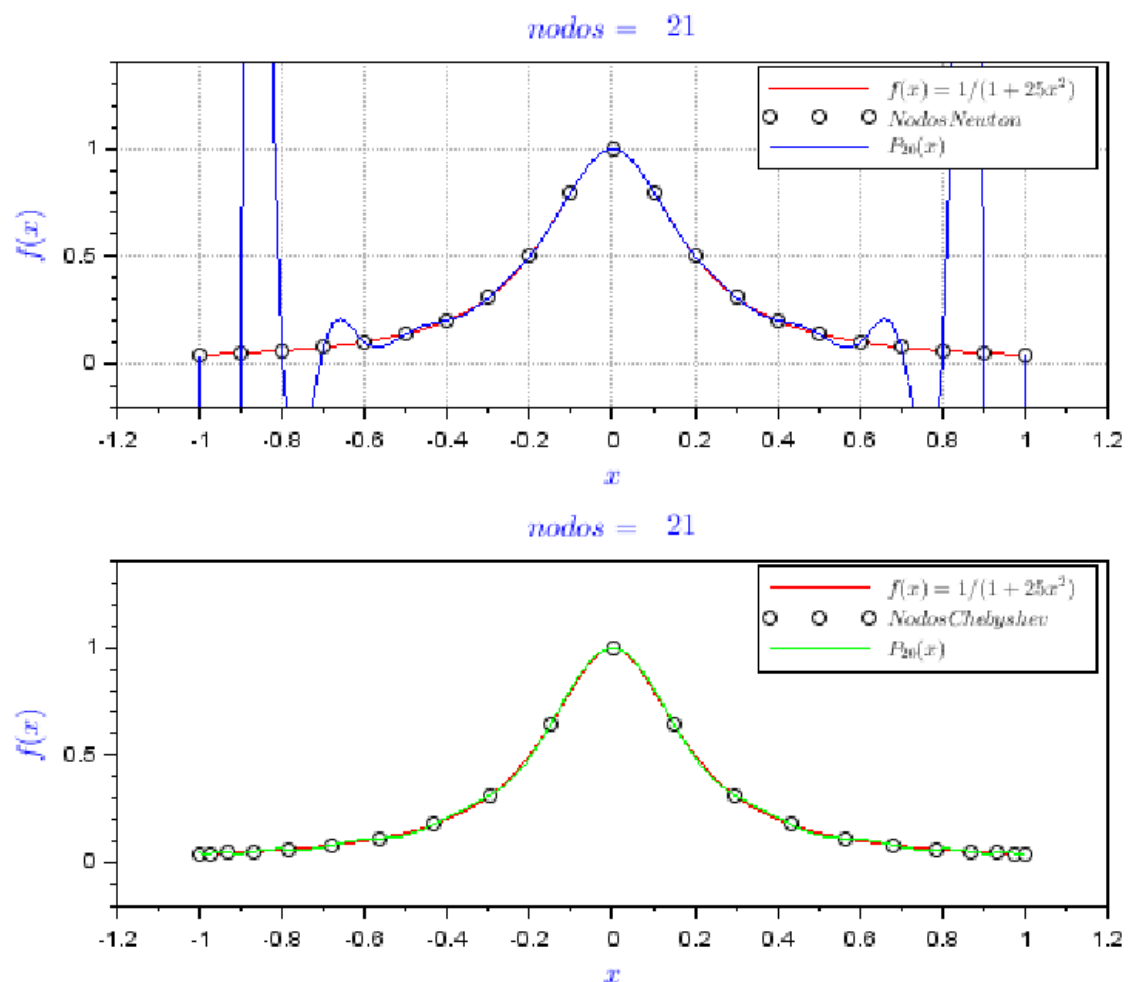
¿podemos elegir mejor dónde ubicarlos?

Una respuesta a este problema se obtiene mediante los **nodos de Chebyshev**.

La idea detrás de los nodos de Chebyshev

Los nodos de Chebyshev **no se distribuyen uniformemente** dentro del intervalo.

Los nodos de Chebyshev no se distribuyen uniformemente dentro del intervalo. A diferencia de los nodos equidistantes, presentan una mayor concentración hacia los extremos y una mayor separación hacia la zona central, como puede observarse en la siguiente figura.



La diferencia es importante: no hemos aumentado necesariamente la cantidad de nodos; hemos cambiado su ubicación. En lugar de repartirlos a distancias iguales, los nodos de Chebyshev redistribuyen los puntos, concentrándolos progresivamente hacia ambos extremos del intervalo.

Esta distribución no es arbitraria. Está relacionada precisamente con el término del error que depende de la posición de los nodos:

$$\omega_{n+1}(x) = \prod_{i=0}^n (x - x_i).$$

Si podemos elegir libremente los x_i , podemos buscar una distribución que controle mejor el máximo valor que alcanza este producto dentro del intervalo.

Es decir, buscamos que

$$\max_{x \in [-1, 1]} \left| \prod_{i=0}^n (x - x_i) \right|$$

sea lo más pequeño posible.

Esta es la conexión fundamental entre el fenómeno de Runge y los nodos de Chebyshev.

No cambiamos la función que queremos aproximar ni dejamos de utilizar interpolación polinómica. Lo que modificamos es **la ubicación de los puntos utilizados para construir el polinomio**.

¿De dónde surgen los nodos de Chebyshev?

Hasta aquí hemos establecido que la ubicación de los nodos influye en el comportamiento del error de interpolación. Los nodos de Chebyshev proponen una distribución particular de estos puntos: **más concentrados hacia los extremos del intervalo y más separados en su zona central**.

Pero ¿de dónde surge esta distribución?

Los nodos que utilizaremos se obtienen a partir de las **raíces de los polinomios de Chebyshev**.

Los polinomios de Chebyshev de primera especie constituyen una familia de polinomios que puede definirse, en el intervalo $[-1, 1]$, mediante:

$$T_n(x) = \cos(n \arccos(x)) \quad -1 \leq x \leq 1.$$

Aunque la expresión contiene funciones trigonométricas, $T_n(x)$ es efectivamente un **polinomio de grado n** .

Por ejemplo, los primeros polinomios de esta familia son:

$$\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \end{aligned}$$

$$T_2(x) = 2x^2 - 1$$

$$T_3(x) = 4x^3 - 3x.$$

Para nuestro propósito no necesitamos estudiar en profundidad toda la familia de polinomios de Chebyshev. Lo que nos interesa particularmente son **los puntos donde estos polinomios se anulan**, es decir, sus raíces.

Obtención de las raíces de los polinomios de Chebyshev

Los nodos de Chebyshev se obtienen a partir de las raíces de los polinomios de Chebyshev de primera especie. Estos polinomios se definen en el intervalo $[-1, 1]$ mediante:

$$T_n(x) = \cos(n \arccos(x)).$$

Para determinar sus raíces se buscan aquellos valores de x para los cuales:

$$T_n(x) = 0.$$

Utilizando la definición anterior:

$$\cos(n \arccos(x)) = 0.$$

La función coseno toma el valor cero para los ángulos

$$\frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}, \dots$$

que pueden expresarse de manera general como:

$$\frac{(2k+1)\pi}{2}, k = 0, 1, 2, \dots$$

Por lo tanto, para obtener las raíces de $T_n(x)$ se plantea:

$$n \arccos(x_k) = \frac{(2k+1)\pi}{2}.$$

Despejando:

$$\arccos(x_k) = \frac{(2k+1)\pi}{2n},$$

y aplicando la función coseno en ambos miembros:

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

De esta manera se obtienen las n raíces del polinomio $T_n(x)$, que utilizaremos como **nodos de Chebyshev** para realizar la interpolación.

A diferencia de los nodos equidistantes, estos puntos no mantienen una separación constante: se encuentran más concentrados hacia los extremos del intervalo $[-1, 1]$ y más separados en su zona central.

Ejemplo: obtención de cinco nodos de Chebyshev

Supongamos que se desean obtener cinco nodos de Chebyshev en el intervalo $[-1, 1]$.

En este caso:

$$n = 5$$

y la expresión anterior resulta:

$$x_k = \cos\left(\frac{(2k+1)\pi}{10}\right), \quad k = 0, 1, 2, 3, 4.$$

Evaluando para cada valor de k :

$$x_0 = \cos\left(\frac{\pi}{10}\right) \approx 0.9511$$

$$x_1 = \cos\left(\frac{3\pi}{10}\right) \approx 0.5878$$

$$x_2 = \cos\left(\frac{5\pi}{10}\right) = 0$$

$$x_3 = \cos\left(\frac{7\pi}{10}\right) \approx -0.5878$$

$$x_4 = \cos\left(\frac{9\pi}{10}\right) \approx -0.9511.$$

Por lo tanto, ordenando los nodos de izquierda a derecha:

$$[-0.9511, -0.5878, 0, 0.5878, 0.9511]$$

Puede observarse que estos puntos **no están igualmente separados**. Las distancias entre nodos disminuyen al aproximarnos a los extremos del intervalo.

Esta distribución constituye una diferencia fundamental respecto de los nodos equidistantes estudiados anteriormente.

Relación con el error de interpolación

La elección de esta distribución no es arbitraria. Recordemos que en la expresión del error de interpolación aparece el producto:

$$\omega_n(x) = \prod_{i=0}^{n-1} (x - x_i),$$

cuyo comportamiento depende directamente de la ubicación de los nodos.

La elección de las raíces de Chebyshev permite minimizar, entre los polinomios mónicos de grado n , el máximo valor absoluto de este producto en el intervalo $[-1, 1]$. En términos del error de interpolación, esto permite controlar de manera favorable el factor asociado con la distribución de los nodos.

Por esta razón, los nodos de Chebyshev constituyen una alternativa a la distribución equidistante cuando se trabaja con interpolación polinómica global.

Es importante observar que **no se está introduciendo un nuevo método para construir el polinomio interpolante**. Una vez determinados los nodos de Chebyshev x_i y calculados los correspondientes valores $f(x_i)$, el polinomio puede construirse utilizando alguno de los métodos ya estudiados, como Lagrange o Newton. Lo que cambia es la elección de los datos utilizados para interpolar:

nodos equidistantes $\rightarrow$ nodos seleccionados según Chebyshev

Nodos de Chebyshev en un intervalo $[a, b]$

La expresión obtenida anteriormente genera nodos en el intervalo estándar:

$$[-1, 1].$$

Sin embargo, los problemas de interpolación pueden estar definidos sobre un intervalo cualquiera:

$$[a, b].$$

Para trasladar un punto $x \in [-1, 1]$ al intervalo $[a, b]$ se utiliza la transformación lineal:

$$z = \frac{a+b}{2} + \frac{b-a}{2}x.$$

Sustituyendo en ella la expresión de los nodos de Chebyshev:

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

se obtiene:

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

con

$$k = 0, 1, \dots, n-1.$$

Esta expresión permite obtener directamente n nodos de Chebyshev distribuidos en cualquier intervalo $[a, b]$.

La transformación conserva la característica principal de la distribución: **los nodos permanecen más concentrados hacia los extremos del intervalo y más separados hacia su zona central.**

Aplicación: nodos de Chebyshev para la función de Runge

Consideremos nuevamente la función utilizada para estudiar el fenómeno de Runge:

$$f(x) = \frac{1}{1 + 25x^2}, \quad -1 \leq x \leq 1.$$

Anteriormente observamos el comportamiento de su interpolación utilizando nodos equidistantes. Utilicemos ahora, a modo de ejemplo, **cinco nodos de Chebyshev.**

Para $n = 5$, los nodos están dados por:

$$x_k = \cos\left(\frac{(2k+1)\pi}{10}\right), \quad k = 0, 1, 2, 3, 4.$$

Como obtuvimos anteriormente, ordenados de menor a mayor resultan:

$$x_0 \approx -0.9511, \quad x_1 \approx -0.5878, \quad x_2 = 0, \quad x_3 \approx 0.5878, \quad x_4 \approx 0.9511.$$

El siguiente paso consiste en evaluar la función de Runge en cada uno de estos nodos:

$$f(x_i) = \frac{1}{1 + 25x_i^2}.$$

Por ejemplo, para el primer nodo:

$$f(-0.9511) = \frac{1}{1 + 25(-0.9511)^2} \approx 0.0424.$$

Debido a que la función es simétrica respecto del eje y ,

$$f(-x) = f(x),$$

los valores correspondientes a nodos opuestos serán iguales.

De esta manera se obtiene la siguiente tabla:

i	x_i	$f(x_i)$
0	-0.9511	0.0424

1	-0.5878	0.1038
2	0	1
3	0.5878	0.1038
4	0.9511	0.0424

Por lo tanto, ya disponemos de los cinco pares de datos:

$$(x_i, f(x_i))$$

necesarios para construir el polinomio interpolante.

Construcción del polinomio

Una vez obtenidos los nodos de Chebyshev, **el procedimiento de interpolación no cambia**. Podemos utilizar cualquiera de los métodos estudiados anteriormente.

Por ejemplo, mediante la forma de Lagrange:

$$P_4(x) = \sum_{i=0}^4 f(x_i) L_i(x),$$

donde

$$L_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^4 \frac{x - x_j}{x_i - x_j}.$$

Al sustituir los cinco pares de datos anteriores y simplificar, se obtiene aproximadamente:

$$P_4(x) \approx 1 - 4.8780x^2 + 4.4037x^4$$

Los términos de grado impar desaparecen como consecuencia de la simetría de los datos, por lo que el polinomio resultante también es una función par:

$$P_4(-x) = P_4(x).$$

Evaluación del interpolante

Una vez construido el polinomio podemos utilizarlo para estimar valores de la función dentro del intervalo.

Por ejemplo, para

$$x = 0.5,$$

el valor exacto de la función es:

$$f(0.5) = \frac{1}{1 + 25(0.5)^2} = \frac{1}{7.25} \approx 0.13793.$$

Utilizando el polinomio obtenido:

$$P_4(0.5) = 1 - 4.8780(0.5)^2 + 4.4037(0.5)^4$$
$$P_4(0.5) \approx 0.0557.$$

Por lo tanto, el error absoluto en este punto es aproximadamente:

$$|E(0.5)| = |f(0.5) - P_4(0.5)| \approx 0.0822.$$

Este resultado también permite realizar una observación importante: **utilizar nodos de Chebyshev no implica que el error sea pequeño en cualquier punto cuando se utiliza un número reducido de nodos**. Su ventaja está relacionada con una distribución más favorable de los nodos y con el control del comportamiento global del error, no con la obtención automática de una aproximación exacta entre ellos.