Parcial Ajuste de Curvas (Octubre 2023)

Nicole Cuervo Pérez
nicole.cuervo@utp.edu.co

Resumen - El desarrollo de este parcial, consistió en aplicar técnicas de interpolación polinomial y regresión por mínimos cuadrados para modelar datos. Luego, se llevó a cabo un análisis comparativo para identificar cuál de estas técnicas era más adecuada y en qué situaciones cada enfoque podía ofrecer resultados más precisos. Este proceso es esencial para la toma de decisiones informadas, en la modelización y análisis de datos. Los pormenores del desarrollo se detallarán a continuación.

Índice de Términos – interpolación polinomial, regresión por mínimos cuadrados, polinomio interpolante de Hermite, polinomio interpolante de Newton.

I. INTRODUCCION

En el apasionante mundo de la modelización y análisis de datos, la búsqueda de métodos efectivos para representar y entender conjuntos de datos es una constante. En este contexto, se ha llevado a cabo un procedimiento que combina dos enfoques fundamentales: la interpolación polinomial y la regresión por mínimos cuadrados. El objetivo era claro: encontrar la mejor manera de ajustar una curva a un conjunto de datos dado, considerando la precisión, la simplicidad y la idoneidad del modelo.

Este procedimiento no solo fue una oportunidad para explorar y aplicar técnicas matemáticas avanzadas, sino también para cuestionar la eficacia de cada enfoque en diferentes escenarios. En este artículo, se le invita a adentrarse en un viaje por los meandros de la interpolación y la regresión, dos pilares de la estadística y las matemáticas aplicadas, y a descubrir cómo estas técnicas pueden arrojar luz sobre la comprensión y predicción de datos en situaciones diversas.

A lo largo de las siguientes secciones, se detallarán los pasos concretos que se siguieron en el desarrollo, haciendo un especial énfasis en los resultados obtenidos y en el análisis comparativo, lo que permitió discernir cuál de estos enfoques demostró ser el más apropiado en función de las circunstancias.

En última instancia, este artículo no solo arroja luz sobre la potencia de estas herramientas analíticas, sino que también resalta la importancia de seleccionar cuidadosamente el método adecuado para el análisis de datos, considerando la naturaleza de los datos y los objetivos específicos del estudio.

II. OBJETIVOS

1. Utilizar técnicas de interpolación polinomial para encontrar una curva que se ajuste al conjunto de datos dado.
2. Utilizar regresión por mínimos cuadrados para encontrar un modelo polinomial que mejor describa el conjunto dado.
3. Realizar un análisis comparativo entre las técnicas de interpolación junto con la regresión.

III. MATERIALES

1. Computadora con software de análisis de datos (Python y/o MATLAB).
2. Conjunto de puntos.

IV. PROCEDIMIENTO

A partir de la función $f(x) = \sin\left(\frac{\pi}{2}x\right)(x^2+1)$, tomando el conjunto de puntos $\{x_i\}_{i=0}^{4} = \{-3, 1.5, 0, 1.5, 3\}$ construya:

a) Polinomio interpolante de Vandermonde, Newton o Lagrange.
b) Polinomio interpolante de Hermite.
c) Trazador Cúbico.
Ahora, a partir de conjunto:

x	-3.0	-2.5	-2	-1.5	-1	-0.5	0	0.5	1	1.5	2	2.5	3
y	10	5	-1	-2	-1	-1	0	3	1	2.3	-0.5	-7	-11

Construya un polinomio de cuarto grado ($f(x) = a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0$), donde los coeficientes a_i deben ser determinados usando regresión por mínimos cuadrados.

V. RESULTADOS Y ANÁLISIS

a) Polinomio interpolante de Newton.

Aunque hay uno y sólo un polinomio de n-ésimo grado que se ajusta a n + 1 puntos, existe una gran variedad de formas matemáticas en las cuales puede expresarse este polinomio. En este caso, se optó por una alternativa que es muy adecuada

Documento recibido el 11 de octubre de 2023.

para implementarse en computadora: los polinomios de Newton.

De manera generalizada, el polinomio interpolante de Newton está dado por la siguiente ecuación:

$$f_n(x) = b_0 + b_1(x - x_0) + \cdots + b_n(x - x_0)(x - x_1)\cdots(x - x_{n-1})$$

Donde los coeficientes b_i pueden obtenerse a partir de diferencias divididas. Las diferencias divididas pueden ser vistas como la variación promedio de un conjunto de puntos, dependiendo del número de puntos, recibirá su orden.

Los puntos asociados con datos se utilizan para evaluar los coeficientes b_0, b_1,..., b_n,. Para un polinomio de n-ésimo grado se requieren $n + 1$ puntos:

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

Para construir el polinomio de interpolación, se utilizaron estos datos y las siguientes ecuaciones para evaluar los coeficientes:

$$\begin{aligned} b_0 &= f(x_0) \\ b_1 &= f[x_1, x_0] \\ b_2 &= f[x_2, x_1, x_0] \\ &\vdots \\ b_n &= f[x_n, x_{n-1}, \cdots, x_1, x_0] \end{aligned}$$

Donde las evaluaciones de la función colocadas entre paréntesis son diferencias divididas finitas. Cabe puntalizar que estas ecuaciones son recursivas (es decir, las diferencias de orden superior se calculan tomando diferencias de orden inferior. Tal propiedad se puede aprovechar ampliamente a la hora de desarrollar un programa computacional eficiente

En este caso, para fines prácticos, estas diferencias divididas fueron evaluadas hacienda uso del software Excel. Los resultados se encuentran consignados en la tabla I.

Polinomio Interpolante de Newton						
i	**Xi**	**Yi**	**D.D1**	**D.D2**	**D.D3**	**D.D4**
0	-3	10				
1	-1.5	-2.298097039	-8.19873136			
2	0	0	1.53206469	3.24359868		
3	1.5	2.298097039	1.53206469	0	-0.72079971	
4	3	-10	-8.19873136	-3.24359868	-0.72079971	0

TABLA I
DATOS DEL POLINOMIO INTERPOLANTE DE NEWTON

Los datos resaltados en gris, corresponden a los coeficientes b_0, b_1,..., b_n, del polinomio interpolante, expresado a continuación:

$$P(x) = 10 - 8.19873136(x + 3) + 3.25359868(x + 3)(x + 1.5) - 0.72079971(x + 3)(x + 1.5)(x)$$

La gráfica correspondiente a este polinomio (línea verde punteada), y la función dada inicialmente (línea roja) se observa en la figura 1.

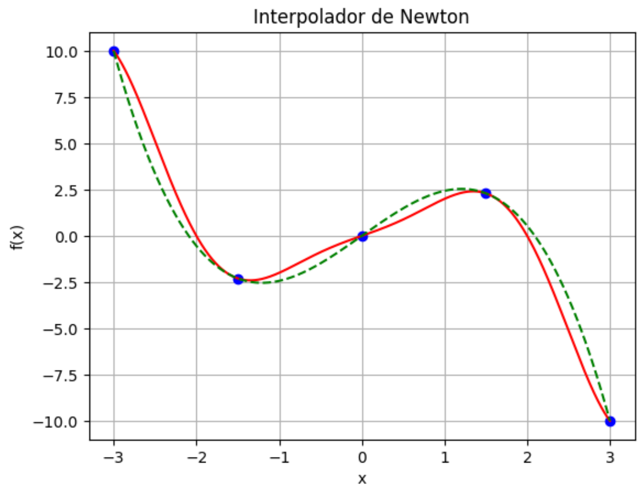

FIGURA 1
GRÁFICA DEL POLINOMIO INTERPOLADOR DE NEWTON

En coherencia con la gráfica y los cálculos, es factible afirmar que, cuando se elige un grado de polinomio apropiado y se seleccionan puntos de interpolación adecuados, el método de Newton en diferencias divididas puede proporcionar una interpolación muy precisa. Esto es útil en aplicaciones científicas y de ingeniería donde se requiere una aproximación cercana a la función real.

Además, ya que el cálculo de los coeficientes del polinomio de Newton se realiza de manera eficiente utilizando diferencias divididas, esto ahorra tiempo computacional en comparación con otros métodos de interpolación más complejos.

Otra ventaja de este polinomio interpolante, es que si se obtienen nuevos puntos de datos o se corrige un punto existente, es relativamente sencillo actualizar el polinomio de interpolación de Newton en diferencias divididas sin tener que recalcular todo el polinomio desde cero.

Sin embargo, la precisión del polinomio dependerá en gran medida de la ubicación de los puntos de interpolación. Si los puntos están mal espaciados o elegidos de manera inadecuada, el polinomio puede ser una mala aproximación de la función real.

b) Polinomio interpolante de Hermite.

El interpolador de Hermite es un método de interpolación que se utiliza para aproximar una función desconocida a través de un conjunto de puntos conocidos, al igual que otros métodos de interpolación. Sin embargo, lo que distingue al interpolador de Hermite es su capacidad para manejar no solo los valores de la función en los puntos de interpolación, sino también las derivadas de la función en esos mismos puntos.

El interpolador de Hermite se utiliza cuando se necesita una interpolación más precisa y detallada que simplemente ajustar una curva a los valores de una función en puntos específicos. Permite representar de manera más precisa las características locales de la función y su comportamiento en los puntos de interpolación.

En este caso, el polinomio interpolante está dado por:

$$\begin{aligned} P(x) = b_0 + b_1(x - x_0) + b_2(x - x_0)^2 + b_3(x - x_0)^2(x - x_1) \\ + b_4(x - x_0)^2(x - x_1)^2 + \cdots \\ + b_{2n-1}(x - x_0)^2(x - x_1)^2 \ldots (x - x_n) \end{aligned}$$

Cuyo grado es $2n + 1$, y para este caso particular será de noveno grado.

Para calcular los coeficientes de este polinomio, se necesita el conjunto de datos $\{(x_i, y_i, y_i')\}$, y en esa ocasión, calcular una diferencia dividida en un mismo punto es como calcular la derivada en ese punto.

Siguiendo esto, nuevamente se empleó la herramienta de Excel para construir la tabla de diferencias divididas.

Polinomio Interpolante de Hermite											
i	Zi	Yi	D.D1	D.D2	D.D3	D.D4	D.D5	D.D6	D.D7	D.D8	D.D9
0	-3	10									
1	-3	10	-6								
2	-1.5	-2.298097039	-8.19873136	-1.46582091							
3	-1.5	-2.298097039	-0.1767767	5.34796978	4.54252712						
4	0	0	1.53206469	1.13922759	-1.40291406	-1.981813728					
5	0	0	1	-0.3547098	-0.99595826	0.135651935	0.70582189				
6	1.5	2.298097039	1.53206469	0.3547098	0.2364732	0.410810485	0.06114634	-0.14326123			
7	1.5	2.298097039	-0.1767767	-1.13922759	-0.99595826	-0.410810485	-0.27387366	-0.07444889	0.01529163		
8	3	-10	-8.19873136	-5.34796978	-1.40291406	-0.135651935	0.06114634	0.07444889	0.0248163	0.00158744	
9	3	-10	-6	1.46582091	4.54252712	1.981813728	0.70582189	0.14326123	0.01529163	-0.00158744	-0.00052915

TABLA II
DATOS DEL POLINOMIO INTERPOLANTE DE HERMITE.

Los datos resaltados en gris, corresponden a los coeficientes b_0, b_1,..., b_n, del polinomio interpolante, expresado a continuación:

$$\begin{aligned} P(x) = 10 + (-6)(x - (-3)) + (-1.4658)(x - (-3))(x \\ - (-3)) + (4.5425)(x - (-3))(x \\ - (-3))(x - (-1.5)) + (-1.9818)(x \\ - (-3))(x - (-3))(x - (-1.5))(x \\ - (-1.5)) + (0.70582)(x - (-3))(x \\ - (-3))(x - (-1.5))(x - (-1.5))(x \\ - (0)) + (-0.14326)(x - (-3))(x \\ - (-3))(x - (-1.5))(x - (-1.5))(x \\ - (0))(x - (0)) + (0.015294)(x \\ - (-3))(x - (-3))(x - (-1.5))(x \\ - (-1.5))(x - (0))(x - (0))(x - (1.5)) \\ + (0.0015867)(x - (-3))(x - (-3))(x \\ - (-1.5))(x - (-1.5))(x - (0))(x \\ - (0))(x - (1.5))(x - (1.5)) \\ + (-0.00052891)(x - (-3))(x \\ - (-3))(x - (-1.5))(x - (-1.5))(x \\ - (0))(x - (0))(x - (1.5))(x - (1.5)) \\ * (x - (3)) \end{aligned}$$

La gráfica correspondiente a este polinomio se observa en la figura 2.

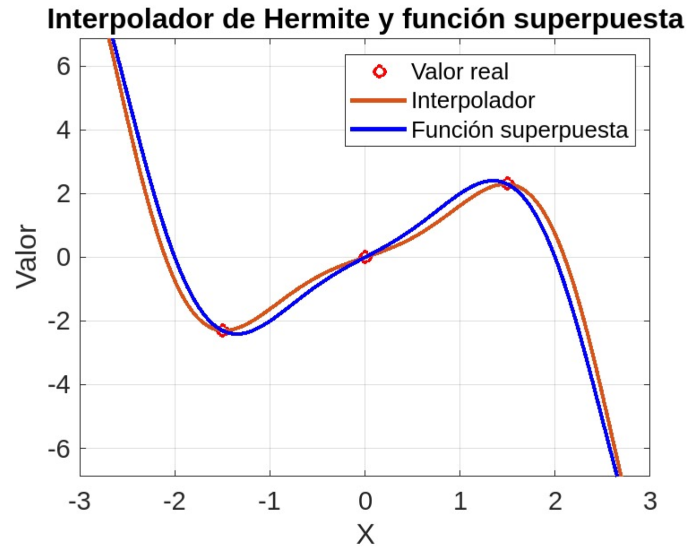

FIGURA 2
GRÁFICA DEL POLINOMIO INTERPOLADOR DE HERMITE

En este caso, en lugar de solo conocer los valores de la función en los puntos de interpolación, el interpolador de Hermite también requiere información sobre las derivadas de la función en esos puntos. Esto puede ser útil cuando se desea capturar cambios bruscos en la función o cuando se tienen condiciones específicas que deben satisfacerse en términos de derivadas.

Además de los valores de la función en los puntos de interpolación, también se deben proporcionar las derivadas de la función en esos puntos. Esto significa que el interpolador de Hermite requiere más información que otros métodos de interpolación, lo que puede ser una desventaja en situaciones en las que no se disponga de esa información adicional.

Debido a la inclusión de derivadas, los cálculos en el interpolador de Hermite pueden ser más complicados en comparación con otros métodos de interpolación, como el polinomio de Lagrange o el polinomio de Newton.

c) Trazador cúbico

Haciendo uso de este método, se buscó construir una función a trozos (polinomio de grado 3), en cada intervalo $[x_i, x_{i+1}]$, definido así:

$$\begin{aligned} T(x) = \{(t_0(x), -3 \le x \le -1.5); \\ (t_1(x), -1.5 \le x \le 0); \\ (t_2(x), 0 \le x \le 1.5); \\ (t_3(x), 1.5 \le x \le 3)\} \end{aligned}$$

Siguiendo este método, para este caso particular, es importante hallar el valor de los parámetros S_0, S_1, S_2, S_3 y S_4 y se determinan a partir de la siguiente ecuación:

$$h_{i-1} \cdot S_{i-1} + 2(h_{i-1} + h_i)S_i + h_i \cdot S_{i+1} = 6\left[\frac{y_{i+1} - y_i}{h_i} - \frac{y_i - y_{i-1}}{h_{i-1}}\right]$$

Como se trata de un trazador cúbico natural, asumimos $S_0 = S_4 = 0$

Además, el valor de h_i es constante y equivale a 1.5.

Por lo tanto, queda expresado el siguiente sistema de ecuaciones:

- $1.5S_0 + 6S_1 + 1.5S_2 = 58.38477631$
- $1.5S_1 + 6S_2 + 1.5S_3 = 0$
- $1.5S_2 + 6S_3 + 1.5S_4 = -58.38477631$

Planteándolo matricialmente, se tiene:

$$\begin{pmatrix} 6 & 1.5 & 0 \\ 1.5 & 6 & 1.5 \\ 0 & 1.5 & 6 \end{pmatrix} \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} 58.38477631 \\ 0 \\ -58.38477631 \end{pmatrix}$$

$$\begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} 9.73079 \\ 0 \\ -9.73079 \end{pmatrix}$$

Ahora, para encontrar los coeficientes de cada trozo del polinomio, se emplearon las siguientes fórmulas:

$$a_i = \frac{S_{i+1} - S_i}{6h_i}$$

$$b_i = \frac{S_i}{2}$$

$$c_i = \frac{Y_{i+1} - Y_i}{h_i} - \frac{S_{i+1} - 2S_i}{6} \cdot h_i$$

$$d_i = y_i$$

Una vez más, los cálculos fueron realizados empleando Excel y se encuentran consignados en la tabla III.

TRAZADOR CÚBICO						
i	**Xi**	**Yi**	**Ai**	**Bi**	**Ci**	**Di**
0	-3	10	1.08119889	0	-10.6314289	10
1	-1.5	-2.298097039	-1.08119889	4.865395	-3.33333031	-2.298097039
2	0	0	-1.08119889	0	3.96476219	0
3	1.5	2.298097039	1.08119889	-4.865395	-13.0641264	2.298097039
4	3	-10				-10

TABLA III
DATOS DEL TRAZADOR CÚBICO

De esta manera, los trozos de la función son los siguientes:

$T_0 = 1.08119889(x + 3)^3 - 10.6314289(x + 3) + 10$
$T_1 = -1.08119889(x + 1.5)^3 + 4.865395(x + 1.5)^2 - 3.33333031(x + 1.5) - 2.298097039$
$T_2 = -1.08119889(x)^3 + 3.96476219(x)$
$T_3 = 1.08119889(x - 1.5)^3 - 4.865395(x - 1.5)^2 - 13.0641264(x - 1.5) - 2.298097039$

En la figura 3, se observa el gráfico de la función dada inicialmente (color negro), en contraste con el trazador cúbico (cada color corresponde a un trozo de la función).

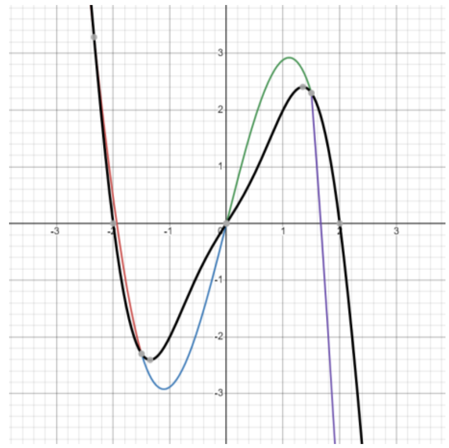

FIGURA 3
GRÁFICA DEL TRAZADOR CÚBICO

Tal como se observa, en un trazador cúbico, se divide el dominio de interpolación en segmentos o intervalos, y en cada intervalo se ajusta una función cúbica. Cada función cúbica se elige de manera que sea suave y continua en los puntos de interpolación, lo que significa que la primera y segunda derivadas son continuas a lo largo de toda la curva. Esto proporciona una transición suave entre los segmentos adyacentes y evita cambios bruscos en la interpolación.

A pesar de su suavidad, los trazadores cúbicos se pueden calcular de manera eficiente y su aplicación no suele ser computacionalmente costosa.

No obstante, los trazadores cúbicos tienden a producir curvas suaves, pero no garantizan que la función resultante sea convexa o cóncava en todo el intervalo de interpolación. En algunos casos, esto puede llevar a interpolaciones que no reflejan la naturaleza subyacente de los datos.

Regresión por mínimos cuadrados

Con el conjunto de datos a evaluar, la función que mejor describe el comportamiento del conjunto está dada por:

$$y_i = b_0 x_i^0 + b_1 x_i + b_2 x_i^2 + b_3 x_i^3 + b_4 x_i^4$$

Para el conjunto de puntos, se tiene el siguiente sistema 13x2:

-3	10
-2.5	5
-2	-1
-1.5	-2
-1	-1
-0.5	-1
0	0
0.5	3
1	1
1.5	2.3
2	-0.5
2.5	-7
3	-11

TABLA IV
SISTEMA 13X2

Evaluando matricialmente, tenemos:

$$\begin{pmatrix} x0^0 & x0^1 & x0^2 & x0^3 & x0^4 \\ x1^0 & x1^1 & x1^2 & x1^3 & x1^4 \\ x2^0 & x2^1 & x2^2 & x2^3 & x2^4 \\ x3^0 & x3^1 & x3^2 & x3^3 & x3^4 \\ \dots & \dots & \dots & \dots & \dots \\ x13^0 & x13^1 & x13^2 & x13^3 & x13^4 \end{pmatrix} \begin{pmatrix} b0 \\ b1 \\ b2 \\ b3 \\ b4 \end{pmatrix} = \begin{pmatrix} y0 \\ y1 \\ y2 \\ y3 \\ y4 \end{pmatrix}$$

Xi^0	Xi^1	Xi^2	Xi^4	Xi^4
1	-3	9	-27	81
1	-2.5	6.25	-15.625	39.0625
1	-2	4	-8	16
1	-1.5	2.25	-3.375	5.0625
1	-1	1	-1	1
1	-0.5	0.25	-0.125	0.0625
1	0	0	0	0
1	0.5	0.25	0.125	0.0625
1	1	1	1	1
1	1.5	2.25	3.375	5.0625
1	2	4	8	16
1	2.5	6.25	15.625	39.0625
1	3	9	27	81

TABLA IV
POTENCIAS DE XI

Siguiendo la propiedad $\vec{B} = (A^T A) A^T \vec{Y}$, y habiendo encontrado los coeficientes b_i:

$$y = 0.7081 + 2.555x - 0.4982x^2 - 0.6956x^3 + 0.0396x^4$$

Graficando, se obtiene lo siguiente:

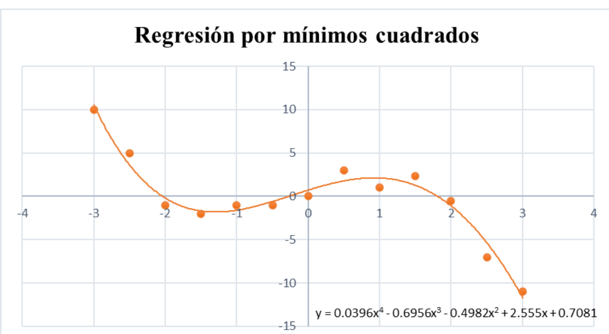

FIGURA 4
GRÁFICA DE LA REGRESIÓN LINEAL

En concordancia con el gráfico, y los cálculos realizados, es factible afirmar que, los trazadores cúbicos son ideales para datos que siguen una tendencia suave y continua. Para datos que presentan cambios bruscos o discontinuidades, los trazadores cúbicos pueden no ser la mejor opción, ya que tienden a suavizar demasiado esas transiciones.

Calcular trazadores cúbicos y gestionar las condiciones de borde puede requerir una implementación cuidadosa y detallada. Errores en la implementación pueden llevar a interpolaciones incorrectas.

VI. CONCLUSIONES

1. Gracias los métodos de interpolación se logró aproximar a una función deseada a partir de un conjunto de puntos dados.
2. En términos de costo computacional, el método de Newton resulta ser el más eficiente.
3. El método de Hermite, resulta muy eficiente para ajustarse al polinomio esperado. El interpolador de Hermite es especialmente útil en aplicaciones en las que se necesita una interpolación altamente precisa que conserve las características locales y cumpla con condiciones específicas en términos de derivadas. Ejemplos de aplicaciones incluyen la interpolación de datos científicos o ingenieriles, la representación de trayectorias en gráficos por computadora y la solución numérica de ecuaciones diferenciales.
4. Hermite tiene un coste computacional relativamente alto, puede llegar a ser problemático en polinomios de grados elevados.
5. El método del trazador cúbico es una herramienta poderosa y versátil para la interpolación de datos, especialmente cuando se necesita una interpolación suave y continua. Es ampliamente utilizado en aplicaciones donde la estética y la continuidad de las curvas son importantes, como gráficos por computadora, diseño de curvas y análisis de datos científicos.

REFERENCIAS

Numerical Analysis (Burden, Richard L. & Faires, J. Douglas, 10th edition, 2016, Cengage Learning, Boston, MA) - Capítulo 3: "Interpolation and Polynomial Approximation

Numerical Mathematics (Quarteroni, Alfio & Saleri, Fausto & Gervasio, Paola, 2010, Springer Science & Business Media, New York, NY)

"Newton interpolation: a comparison of divided difference formulae" (Cavaretta, AS & Fabiano, RC, Mathematics and Computers in Simulation, 56(5), 2001, Elsevier)

"The method of Hermite interpolation in numerical analysis" (Davis, Philip J, SIAM Review, 21(2), 1979, Society for Industrial and Applied Mathematics)