Universidad Nacional Autónoma de Nicaragua, UNAN-León

"A la libertad, por la Universidad"

Facultad de Ciencias y Tecnología

Departamento de Matemática y Estadística.

Tema: **Interpolación Segmentaria, Aproximación de Funciones y Aplicaciones**

Monografía Presentada por: **María Victoria Chávez Esquivel**

Previo para optar al título de: **Licenciada en Matemática**

Tutor: MSc. W. Milton Carvajal Herradora

León, Nicaragua, Centroamérica, 2014

DEDICATORIA

La presente tesis está dedicada en primer lugar a mi bello hijo Manuel Esaú Chávez quien ha sido la razón y el motivo de superarme en la vida y seguir adelante en este camino. En segundo lugar se la dedico a mis padres Gregorio Chávez y Ada Luz Esquivel; mis hermanos que de una u otra manera me han alentado a seguir adelante a pesar de los obstáculos que se me han presentado en la vida y continuar en firme.

AGRADECIMIENTOS

Agradezco primeramente a Dios todo poderoso por darme la sabiduría y el entendimiento para realizar mi trabajo monográfico, a mis padres por darme el apoyo en todo lo que he necesitado y en especial a mi estimado profesor MSc. W. Milton Carvajal Herradora quien tomo a bien darme la mano, aceptar ser mi tutor y compartir parte de sus conocimientos con mi persona para llegar a finalizar mi trabajo y tener la oportunidad de optar a mi título universitario.

INDICE GENERAL

Introducción

En el subcampo matemático del análisis numérico, se denomina interpolación a la construcción de nuevos puntos partiendo del conocimiento de un conjunto discreto de puntos. En las ciencias e ingeniería es frecuente disponer de un cierto número de datos obtenidos por muestreo o a partir de un experimento y pretender construir una función que los ajuste para el análisis posterior, la predicción o ajuste sobre datos nuevos o la búsqueda de puntos óptimos, longitudes, áreas o superficies de respuesta; problemas que conlleva integrar o derivar aquella función. Otro problema estrechamente ligado con el de la interpolación es la aproximación de una función complicada por una más simple. Si tenemos una función cuyo cálculo resulta costoso, podemos partir de un cierto número de sus valores e interpolar dichos datos construyendo una función más simple o sencilla de cálculo.

En general, por supuesto, no obtendremos los mismos valores evaluando la función construida que si evaluásemos la función original, si bien dependiendo de las características del problema y del método de interpolación usado la ganancia en eficiencia puede compensar el error cometido. En todo caso, se trata de, a partir de n parejas de puntos $(X_k ; Y_k)$, obtener una función f que verifique $f(X_k) = Y_k$; k = 1, 2…, n. A f se la denomina función interpolante de dichos puntos. A los valores X_k se les llama nodos, y los Y_k, valores de respuesta. Algunas formas de interpolación que se utilizan con frecuencia son la interpolación polinómica, la interpolación lineal (la cual es un caso particular de la anterior) y la interpolación por medio de Splines cúbico.

Los Splines también permiten representaciones matemáticas de superficies de respuesta partiendo de información relativa a algunos puntos de la misma. Su construcción consiste en obtener una función de interpolación que pase por esos puntos. Para poder hacer esto es necesario contar con algunas habilidades matemáticas, por ejemplo: derivación, buen manejo del álgebra en la resolución de sistemas de ecuaciones lineales grandes y complejas. Los polinomios cúbicos ofrecen una buena relación entre flexibilidad y velocidad en los cálculos; comparando con polinomios de orden superior, los Splines cúbicos requieren menos cálculo y memoria, además de ser más estables en todo el dominio.

El procedimiento usado para este estudio, fue en primera instancia la Revisión bibliográfica de Análisis Numérico, escritos e investigaciones particulares acerca de la temática; las disponibles en artículos científicos, en Hemeroteca y en páginas web. Para la parte didáctica se trabajó en forma manual y por computadora la construcción de funciones de interpolación que aproximen funciones más complejas y aquellas que pudieran realizarse con pocos puntos y que requieran de poco esfuerzo computacional.

Realizar este trabajo precisa herramientas de computación y el uso de software. Se exploró en software matemáticos y estadísticos la disponibilidad de rutinas ya programadas o bien la construcción de nuevos programas, las posibilidades en generación de resultados, de gráficos, y la rapidez de resolución. Disponibilidad de software fue SciLab, R y MatLab®.

La obtención de datos relacionales [X, Y] en correspondencia a magnitudes físicas o en cuantificadores subjetivos es algo que se sucede de forma cotidiana en cualquier fenómeno físico o de abstracción. Modelos matemáticos construidos a partir de funciones complejas o trascendentes, adolecen de la utilización de computadores o calculadoras para su aplicación práctica. La simple aplicación de operaciones aritméticas elementales (suma y multiplicación) en polinomios de grado menor, acentúan el interés por la interpolación o extrapolación en conjuntos de datos y la aproximación de funciones complejas. Tales polinomios, por interpolación segmentaria, ofrecen una versatilidad al definirse por segmentos adyacentes del dominio.

Compararemos diferentes métodos de interpolación, para interpolar las funciones seno y coseno, aplicando la "Proposición Ixim" para la bondad de ajuste; considerando la Entropía (medida de incertidumbre acerca de información recuperada o perdida) e Información de Fisher. Veremos en qué casos es adecuado aplicar Interpolación Segmentaria y algunas aplicaciones de la vida real.

Hipótesis:

Al utilizar Interpolación Segmentaria,
Es posible interpolar puntos y funciones muy complejas
A través de funciones más simples.

Objetivo general:

- Utilizar los métodos de interpolación segmentaria con enfoque didáctico construyendo modelos matemáticos tanto de forma manual como haciendo uso de software, en problemas abstractos y de aplicación.

Objetivos específicos:

- Comparar métodos de interpolación exactas con interpolación segmentaria con respecto a la precisión de cada método en sus bondades y deficiencias.

- Mostrar cuales son los mejores y peores escenarios del método de interpolación segmentaria “Spline cubico”.

- Aplicar Interpolación segmentaria a problemas reales.

Antecedentes

"Esencialmente todos los modelos son incorrectos, pero algunos son útiles"

Box, George E. P. (1987)

Un modelo matemático además de ser parsimonioso o sencillo, debe contribuir a nuestro entendimiento científico del fenómeno bajo estudio y debe, a veces, proveer una base para la interpolación y extrapolación de valores desconocidos dentro del ámbito de variables explicativas.

La Interpolación segmentaria, es un modelo matemático, que a diferencia de otros modelos de interpolación, su formulación no es tan sencilla, aunque podría presentarse de una forma más amigable para que pueda resolverse tanto de forma manual como con asistencia de software computacional.

Los errores para los métodos de interpolación generalmente requieren el cálculo de derivadas de órdenes superiores, y ya de por sí para funciones complejas, se debe acotar valores para aquella, haciendo aun más complicada la acotación de errores, por lo que nuestra propuesta intenta otras formas para afrontar esto.

De la naturaleza surgen y se construyen modelos y funciones complejas cuyo manejo podría no ser tan simple, pero que puede aproximarse por otros menos complicados de manejar. Una vez construido el modelo sencillo, su manejo suele ser más fácil que el de los modelos complejos que intentan reproducir, tal es el caso de la Interpolación Segmentaria.

Justificación

La fracción de libros de Análisis Numérico que abordan la temática concerniente a Interpolación Segmentaria, es exigua; el abordaje de la misma en el currículo universitario, también es no prioritario, aunque la aplicación de la misma sea importante, según nuestras exploraciones. Además al indagar en las tesis monográficas no se encontró abordaje sobre Interpolación Segmentaria, o de la más general que es la Interpolación Hermitiana.

El presente trabajo, basado en Interpolación segmentaria por Spline cúbicos, se puede comparar con otros métodos en su eficiencia para obtener resultados que permitan un mejor entendimiento de problemas reales que puedan presentarse en circunstancia de nuestra vida cotidiana, social y laboral. También está pensado como una forma didáctica para abordar temáticas que no son muy del agrado para profesionales de áreas diversas, e incluso para matemáticos, temas tales como geometría, trigonometría, álgebra matricial, acotación de errores, algoritmos, entre otros.

La obtención de datos relacionales [X, Y] en correspondencia a magnitudes físicas o en cuantificadores subjetivos es algo que se sucede de forma cotidiana en cualquier fenómeno físico o de abstracción.

Modelos matemáticos construidos a partir de funciones complejas o trascendentes, adolecen de la utilización de computadores o calculadoras científicas para su resolución, lo cual resulta impráctico para muchos construir modelos de forma manual. Se enfatiza en el abordaje y búsqueda de soluciones de forma manual así como también en la construcción de algoritmos y programas no muy complejos en lenguajes de computación amigable.

La simple aplicación de operaciones aritméticas elementales (suma y multiplicación) en funciones polinómicas, acentúan el interés por la interpolación o extrapolación de puntos y la aproximación de funciones complejas para resolverlas a mano. Las funciones polinómicas y los métodos de interpolación que las implementan, además ser adaptables, son dinámicas al definirse por segmentos adyacentes del dominio.

Métodos de Interpolación

El procedimiento usado para este estudio, fue en primera instancia la Revisión bibliográfica de Análisis Numérico, escritos e investigaciones particulares acerca de la temática; las disponibles en artículos científicos, en Hemeroteca y en páginas web.

Para la parte didáctica se trabajó en forma manual la construcción de funciones de interpolación que aproximen funciones más complejas y aquellas que pudieran realizarse con pocos puntos y que no requieran de mucho esfuerzo computacional.

Para realizar este trabajo se precisó contar con herramientas de computación y el uso de programas para aplicaciones en torno a resolver problemas reales. Se exploró en software matemáticos y estadísticos la disponibilidad de rutinas ya programadas o bien la construcción de nuevos programas, las posibilidades en generación de resultados, de gráficos, y la rapidez de resolución. La disponibilidad de software fue en software libre SciLab, R y MatLab®.

La simple aplicación a problemas reales y cotidianos es una manera sencilla de utilizar interpolación segmentaria. Al explorar e indagar sobre Aplicaciones de Interpolación Segmentaria se encontró una variedad de usos, siendo una de ellas en animación avanzada: Dibujo, figuras, Curvas y Superficies en 2D y 3D.

Veremos el abordaje de cada método según el siguiente esquema, y en lo posible se dará de forma explícita cada método con ejemplo y de ser posible en aplicación real

<table>
<tr><td colspan="2">Interpolación por Sistemas de Ecuaciones Lineales</td><td colspan="2">Interpolación Directa</td><td>Sistema de Ecuaciones Lineales Tridiagonal</td></tr>
<tr><td>Evaluación Directa</td><td>Mínimos Cuadrados</td><td>Diferencias dividas de Newton</td><td>Diferencias fijas de Lagrange</td><td>Interpolación Segmentaria</td></tr>
<tr><td colspan="5">Ejemplos comparativos y Aplicaciones</td></tr>
</table>

Previo al abordaje de métodos, nos referiremos brevemente a funciones de densidad y similitud entre funciones como propuesta de criterios de comparación entre ajustes y métodos. En lo posible, evitaremos deducir la acotación de errores por derivadas de orden superior sobre las funciones complejas que se afronten, planteando otra alternativa de enfoque a la calidad de ajuste ya sea de puntos dados o de funciones.

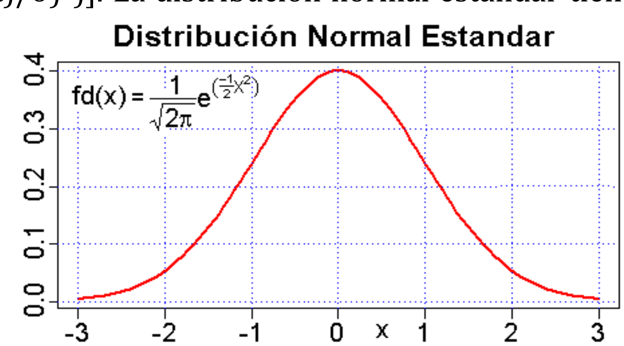
Funciones de Densidad, Momentos y Similitud entre funciones

Una función de densidad $f_d(X)$ de una variable aleatoria continua X, se define como una función de valor real $f_d(X) \geq 0$, de área $a = \int f_d(X)*dX = 1$, sobre el dominio de X, real. El modelo básico es el de la distribución normal que ahora veremos.

Los Primeros cuatro momentos centrales (parámetros discriminantes) para un modelo de probabilidad son: Esperanza Matemática (μ), Varianza (σ^2), Sesgo (ς) y Curtosis (κ), los que se calculan sobre todo el dominio de X. El parámetro de localización μ es también llamado Media, y es una medida de centralidad. Al parámetro $\sigma = \sqrt{(\sigma^2)}$, se le llama Desviación estándar, y es una medida de dispersión al igual que la Varianza. Los coeficientes de sesgo y de curtosis, a ambos se los denomina como parámetros de forma para el modelo de probabilidad. Una variable estandarizada Z, es aquella a la que se ha restado la media y dividido por su desviación típica $Z=(X-\mu)/\sigma$. Observe de la tabla de integrales, que los Parámetros de Forma se calculan sobre variable estandarizada. Esta es una buena alternativa para trabajar con datos desescalados aunque los datos X no estén igualmente espaciados.

μ	$E(X) = \int x*f_d(x)*dx$
σ^2	$Var(X) = \int (x-\mu)^2 * f_d(x)*dx$
ς	$Sk(X) = \int ((x-\mu)/\sigma)^3 * f_d(x)*dx$
κ	$Kt(X) = \int ((x-\mu)/\sigma)^4 * f_d(x)*dx$

La Distribución Normal es un modelo probabilístico, fundamental en Estadística, tiene $E(X)=\mu$ y $Var(X)=\sigma^2$, con función de densidad definida para X en todo el dominio Real, $f_d(X)= 1/(\sqrt{(2\pi)}*\sigma)*\exp[-½*((x-\mu)/\sigma)^2)]$. La distribución normal estándar tiene $E(X)=0$, y $Var(X)=1$, con función de densidad $f_d(X)= 1/\sqrt{(2\pi)}*e\text{^}(-½*X^2)$, de igual forma definida para X Real. Observe del gráfico que casi toda el área de probabilidad se concentra entre -3 y 3, por lo cual se puede considerar el doble de este intervalo como $(-\infty, +\infty)$ para esta distribución. Los coeficientes de sesgo (skewness) y de curtosis (kurtosis) para la distribución normal estándar, son $\mu=0$, y $\sigma=1$, y se calculan por integración directa:



$$Sk(X) = \int x^3 *f_d(x)*dx = 0 \qquad Kt(X) = \int x^4 *f_d(x)*dx = 3.$$

Ahora, cualquier otra función de densidad f_d con sesgo $\varsigma > 0$, es asimétrica positiva ( ) y con sesgo $\varsigma < 0$, es asimétrica negativa ( ). La curtosis para f_d, si $\kappa > 3$, la normal tendría mayor área de cola que la función dada (f_d leptocúrtica), y si $\kappa < 3$, la normal tendría menor área de cola que aquella (f_d platicúrtica).

Información de Fisher: ¿Cuánta información puede una muestra $[X_1, X_2,\ldots, X_n]$ aportar acerca de un parámetro desconocido? En nuestro caso de interpolación no es de mucho interés el enfoque de variable aleatoria, aunque sí la utilización de los dispositivos utilizados para comparar modelos o funciones ajustadas.

Para el modelo Normal de media $E(X)=\mu$ y varianza $Var(X)= \sigma^2$, al suponer que μ es desconocida pero que el valor de la varianza es dado, la información de Fisher se calcula por el recíproco de la varianza $I(\mu) = 1/\sigma^2$.

Para una muestra aleatoria $[X_1, X_2,\ldots, X_n]$ extraída de un modelo normal con media μ desconocida pero con varianza dada σ^2, la información de Fisher contenida en la muestra es $I_n(\mu) = n*I(\mu) = n/\sigma^2$.

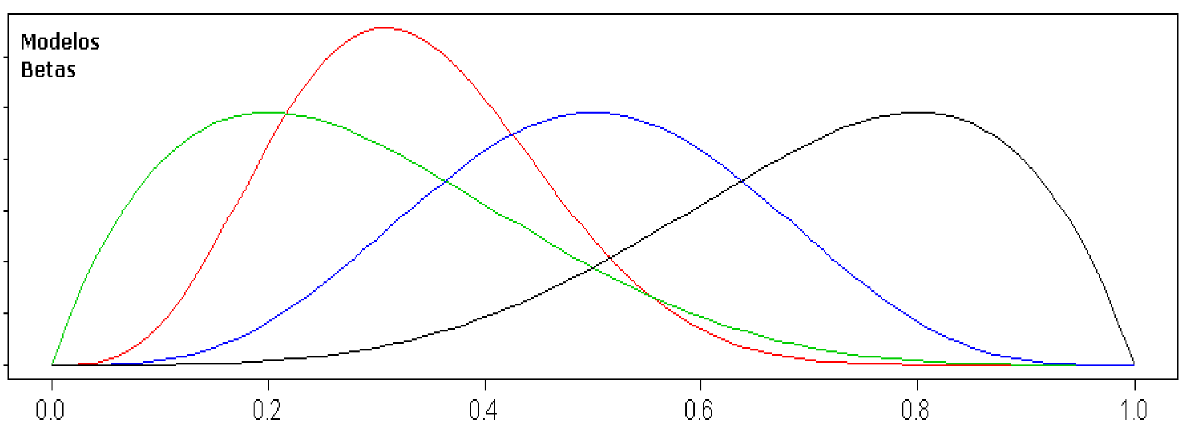
Proposición Ixim (Similitud entre funciones): En caso de funciones no necesariamente de densidad, para comparar entre ellas, se usa como conjunto mínimo de valores, los parámetros discriminantes (μ σ^2 ς κ) para la función normalizada y la Información de Fisher $I_n(\mu)$ para un modelo que incluya los valores de parámetros.

Para esta Proposición, se deberá re-escalar o redefinir la función para que cumpla la condición $f_d(x) \geq 0$, luego se debe obtener el valor de la integral $a = \int f(x)*dx$ en el dominio de x y tomar el recíproco de ese valor (1/a) como factor de "normalización" de la función re-escalada o redefinida. Luego veremos ejemplos para cada caso.

Como condición adicional a la Proposición Ixim, se podría construir un contraejemplo de dos funciones disímiles que cumplan con la proximidad o exactitud de los cuatro parámetros discriminantes, pero que en sus colas halla diferencia y en sus centros modales, pues el parámetro de curtosis no discrimina muy bien hacia los centros probabilísticos, aunque sí hacia las colas. El contraejemplo puede servir como pauta para agregar otros parámetros o condiciones que discriminen mejor entre funciones: momentos de orden superior, otras integrales especiales o derivadas de orden superior, cálculo de la integral de línea de cada función interpolante, etc.

Recordemos que el modelo Normal es simétrico y acampanado, por lo cual usar el criterio de información $I_n(\mu) = n/\sigma^2$ será adecuado cuando la función normalizada sea al menos simétrica. Para un caso más general, puede construirse el modelo Beta no estándar sobre los parámetros de forma u otro modelo.

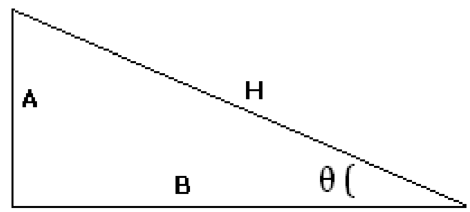
Observe de los cuatro parámetros discriminantes, que al trabajar sobre una función f(X) muy compleja, el cálculo de las integrales para determinar aquellos, podría ser infructuoso de forma analítica, pero bastante fácil para los polinomios interpolantes. Luego, al tomar muchos puntos de f(X) compleja, se estima al menos tres polinomios de interpolación, tomando puntos de forma sistemática o aleatoria, se comparan sus momentos centrales y se toma como mejor polinomio interpolante aquel con varianza mínima. Por último, al disponer de computador, el cálculo de integrales sobre funciones muy complejas se vuelve rutinario, pues los software actuales tienen bibliotecas y rutinas para integrales en una, dos o más variables.



Funciones trigonométricas Seno, Coseno y Tangente; valores más frecuentes

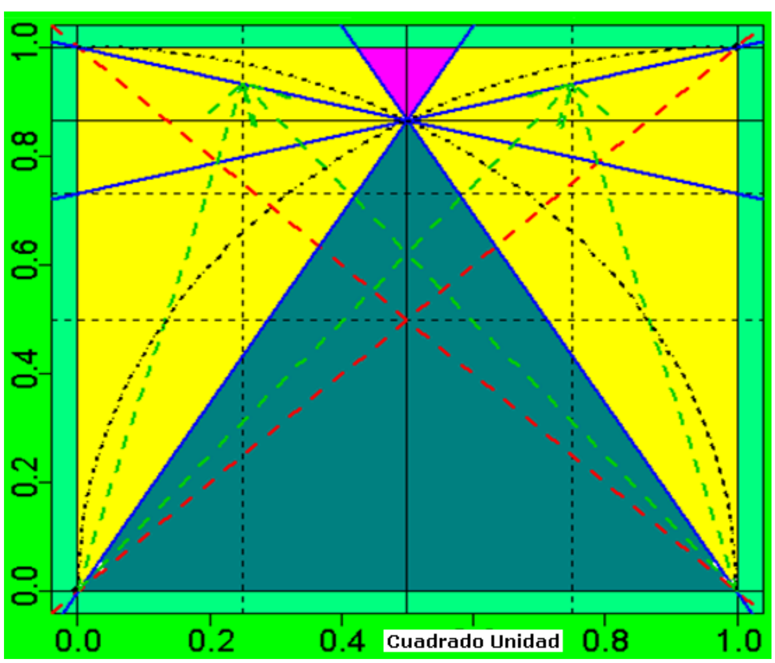
La definición de "Cuerda Media" o "Cuerda inscrita" fue el término adoptado por matemáticos hindús, para la función **Seno**, en sus estudios astronómicos y del círculo unitario, que se malinterpretó como "bahía" o en<u>sena</u>da, al traducirla del sánscrito al árabe y luego al latín. Así la triada (seno, coseno, tangente) corresponde en sánscrito a (jyā, koti-jyā, utkrama-jyā).

En un Triángulo rectángulo de base B, altura A, hipotenusa H y ángulo de base $\theta < 90°$:
$H^2 = A^2 + B^2$; $\text{seno}(\theta) = A/H$; $\text{coseno}(\theta) = B/H$;
$\text{tangente}(\theta) = A/B$; $\text{secante}(\theta) = H/B$.



Veamos abajo en el gráfico, algunos valores de la función Seno a partir del Cuadrado Unidad con Triángulos inscritos: equiláteros, rectángulos e isósceles. Cálculos fáciles de realizar para ángulos sexagesimales (dedúzcase además $\sqrt{(2-\sqrt{3})} = \frac{1}{2}(\sqrt{6}+\sqrt{2})$):

Ángulo θ	0°	15°	30°	60°	75°	45°	90°
seno(θ)	0	$\frac{1}{2}*\sqrt{(2-\sqrt{3})}$	½	$\frac{1}{2}*\sqrt{3}$	$\frac{1}{2}/\sqrt{(2-\sqrt{3})}$	$\frac{1}{2}*\sqrt{2}$	1



Las computadoras evalúan las funciones trigonométricas en radianes, siendo el valor π la razón natural de la circunferencia a su radio ($C=2\pi r$), así $\pi^{Rad} \equiv 180^{O} \equiv 200^{Grad}$; es la relación entre grados centesimales (^{Grad}), grados sexagesimales (^{O}) y radianes (^{Rad}).

Interpolación Polinómica: Evaluación Directa y Mínimos Cuadrados

Existen dos tipos de problemas que se pueden considerar dentro de este concepto; primero el problema de interpolación que involucra el encontrar valores intermedios cuando los valores están dados como un conjunto finito de datos y, segundo el problema de aproximarse a una función dentro de un intervalo, por medio de una función simple, por ejemplo un polinomio.

Claramente, la función aproximada debe ser capaz de reducir el error de aproximación tanto como sea posible; por supuesto, diferentes formas de definir el error corresponden a diferentes métodos.

En el caso de interpolación por el método de sustitución directa, la función aproximada es un polinomio que tiene valores iguales a los dados en un conjunto finito de puntos. Denotando con f_i (i = 0, 1,..., n) los valores de $f(x)$ en los puntos x_i (i = 0, 1,..., n) y $p_n(x)$ la función aproximada.

El tercer caso de interés es cuando el número de datos de los cuales se tiene el valores considerablemente más grande que el grado de la aproximación deseada. Es decir, si se quiere usar un polinomio de bajo orden, por ejemplo un polinomio cúbico, para una aproximación sobre un intervalo en el cual tal vez se conocen veinte valores de una función. Cuatro puntos son suficientes para determinar un polinomio cúbico único, pero a como veremos por interpolación segmentaria con tres puntos es suficiente, dadas las condiciones del método.

Una gran ventaja de los polinomios es que usando operaciones aritméticas accesibles en una computadora digital e incluso a mano, es posible evaluarlos directamente o hacer el cociente de dos polinomios. También es muy fácil evaluar integrales y derivadas de polinomios por cálculo directo y de forma manual. En cambio, el cálculo de otras funciones, como exponenciales, logarítmicas o trigonométricas, se hace con métodos de aproximación. Por estas razones, con frecuencia se usan los polinomios para aproximar a las funciones continuas.

Ajuste Polinómico por Evaluación Directa (ED)

Dados un conjunto de puntos $\{(X_0, Y_0), (X_1, Y_1),\ldots (X_n, Y_n)\}$, al sustituirlos en la función $P_k(X)_n=\sum a_i*X^i=a_0+a_1*X+a_2*X^2+a_3*X^3+\ldots+a_k*X^k$, se genera para k=n+1, el sistema X*A=B:

$$
\begin{array}{ccccc}
X & & A & = & B \\
\boxed{\begin{array}{cccccc} 1 & X_0 & X_0^2 & X_0^3 & \ldots & X_0^k \\ 1 & X_1 & X_1^2 & X_1^3 & \ldots & X_1^k \\ 1 & X_2 & X_2^2 & X_2^3 & \ldots & X_2^k \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & X_n & X_n^2 & X_n^3 & \ldots & X_n^k \end{array}} & & \boxed{\begin{array}{c} a_0 \\ a_1 \\ a_2 \\ \ldots \\ a_k \end{array}} & = & \boxed{\begin{array}{c} Y_0 \\ Y_1 \\ Y_2 \\ \ldots \\ Y_n \end{array}}
\end{array}
$$

Este Sistema de Ecuaciones Lineales con vector de incógnitas $[A]= [a_0, a_1, a_2,\ldots., a_k]$, podría reducirse si el vector de datos $(X_0, X_1, X_2,\ldots X_n)$, estuviera igualmente espaciado.

Una desventaja de interpolación simple es la de depender de la matriz de Vandermonde, la cual es mal condicionada, pues su número de condición es muy elevado ya que se tiene que el detV ≠ 0 si y sólo sí $x_i \neq x_j$, $i \neq j$, por lo que el sistema tiene solución única. La construcción del polinomio interpolador mediante la solución del sistema no es la adecuada, ya que la matriz de Vandermonde cuando se obtiene de esta forma por lo general magnifica los errores de redondeo de la solución del sistema de ecuaciones lineales, al resolverse por computador.

Ajuste Polinómico por Mínimos Cuadrados (MC)

Mínimos cuadrados es una técnica de optimización matemática aproximada, que dada una serie de mediciones o puntos (como en el caso anterior), intenta encontrar una función que ajuste a ellos minimizando la suma de cuadrados de las diferencias ordenadas (llamadas residuos) entre los puntos generados por la función y los correspondientes en los datos.

Dados los n puntos $\{(x_1, y_1), (x_2, y_2)\ldots(x_n, y_n)\}$, es asumida la forma funcional

$$P_k(X)_n= \sum a_i*X^i = a_0+a_1*X+a_2*X^2+a_3*X^3+\ldots +a_k*X^k$$

y se deriva de forma parcial cada vez la función de errores cuadráticos $L= \Sigma[Y_i-P(X_i)]^2$, con respecto a cada coeficiente a_i (k+1 veces), y se genera el así llamado Sistema de Ecuaciones Normales (SEN) M*A=B, para la matriz $M=(\boxed{X}^t*\boxed{X})$, el vector $B=(\boxed{X}^t*Y)$, y el vector de incógnitas [A]= [a_0, a_1, a_2,…., a_k], con la matrix $\boxed{X}$=[X^0, X^1, X^2,…, X^k], y los vectores de datos de longitud n: X=[x_1, x_2,…, x_n] y Y=[y_1, y_2,…, y_n]; el vector X^0 representa al vector unidad sin importar que incluya ceros, así obviando el subíndice i de X_i y Y_i se especifica el sistema de ecuaciones de la siguiente forma:

$$
\underset{M}{\begin{bmatrix} n & \Sigma X & \Sigma X^2 & \Sigma X^3 & \dots & \Sigma X^k \\ \Sigma X & \Sigma X^2 & \Sigma X^3 & \Sigma X^4 & \dots & \Sigma X^{k+1} \\ \Sigma X^2 & \Sigma X^3 & \Sigma X^4 & \Sigma X^5 & \dots & \Sigma X^{k+2} \\ \Sigma X^3 & \Sigma X^4 & \Sigma X^5 & \Sigma X^6 & \dots & \Sigma X^{k+3} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \Sigma X^k & \Sigma X^{k+1} & \Sigma X^{k+2} & \Sigma X^{k+3} & \dots & \Sigma X^{2k} \end{bmatrix}}
\underset{A}{\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \dots \\ a_n \end{bmatrix}}
\underset{=}{=}
\underset{B}{\begin{bmatrix} \Sigma Y \\ \Sigma X^*Y \\ \Sigma X^{2*}Y \\ \dots \\ \Sigma X^{k*}Y \end{bmatrix}}
$$

Un requisito implícito para el método de mínimos cuadrados es que los errores de cada medida estén distribuidos de forma aleatoria, también es importante que los datos recogidos estén bien seleccionados, para que permitan visualizar en las variables que han de ser resueltas. Si el propósito es solo de ajuste y no inferencial la aleatorización puede obviarse.

Ejemplo: Interpolar las funciones SENO y COSENO solo en el intervalo [0, π].

Para la función SENO conociendo los siete puntos X=[0 π/4 π/3 π/2 2π/3 3π/4 π] con Y=[0 √½ √¾ 1 √¾ √½ 0]. Para ajustar un polinomio de grado 4 se requiere cinco puntos, elegir aquellos igualmente espaciados X=[0 π/4 π/2 3π/4 π], Y=[0 √½ 1 √½ 0].

Para reducir cálculos re-escalar los datos X con Z=(X – C)/d, donde C= π/2 es el centro y d= π/4 es la distancia uniforme entre los datos X (al tener un número par de datos se transformaría con Z=2(X – C)/d). Así Z= [-2 -1 0 1 2]. La función de ajuste es $p(Z)=a_0+a_1{*}Z+a_2{*}Z^2+a_3{*}Z^3+a_4{*}Z^4$; la variable re-escalada es $\boxed{Z=2(2{*}X/\pi-1)}$ y los sistemas de ecuaciones lineales para cada método son:

Ec.	ED: Evaluación Directa X*A = B	MC: Mínimos Cuadrados M*A = B
1 2 3 4 5	$\begin{bmatrix} 1 & -2 & 4 & -8 & 16 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 & 16 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} 0 \\ \sqrt{½} \\ 1 \\ \sqrt{½} \\ 0 \end{bmatrix}$	$\begin{bmatrix} 5 & 0 & 10 & 0 & 34 \\ 0 & 10 & 0 & 34 & 0 \\ 10 & 0 & 34 & 0 & 130 \\ 0 & 34 & 0 & 130 & 0 \\ 34 & 0 & 130 & 0 & 514 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} 1+\sqrt{2} \\ 0 \\ \sqrt{2} \\ 0 \\ \sqrt{2} \end{bmatrix}$

Ambos parecen sistemas de ecuaciones difícil de resolver a mano, pero no lo son: observe de ED que $a_0=1$ es inmediata de la ecuación 3. Al combinar las ecuaciones (1, 5) y las ecuaciones (2, 4), resulta un SEL en [a_2, a_4]; al resolverlo lleva a la solución $a_2=(8\sqrt{2}-15)/12$; $a_4=-(2\sqrt{2}-3)/12$. Por último se resuelve las ecuaciones rezagadas para determinar $a_1=0$ y $a_3=0$. Para MC es un sistema separable: en [a_0 a_2 a_4] y [a_1 a_3].

Ambos sistemas conducen a la misma solución ya obtenida. Así podemos expresar la función de ajuste $p(Z)=a_0+a_2*Z^2+a_4*Z^4 = 1 + Z^2(a_2 + a_4*Z^2)$, que de forma compacta ya simplificada resulta en la variable original X:

$$\boxed{P_{4.sen}(X)_5 = 1-1/3*(2*X/\pi-1)^2*\{ (15-8\sqrt{2})-(12\text{-}8\sqrt{2})*(2*X/\pi-1)^2 \} \quad 0\leq X\leq\pi}$$

Se observa en esta al evaluar X en términos de π radianes, los cálculos se resumen en las operaciones aritméticas básicas (+ − ∗ /); el valor π se simplifica y el valor √2 se puede fijar en <u>1.41421</u>356, pues como veremos llegamos a alcanzar una precisión de al menos tres cifras significativas, al determinar el error máximo absoluto.

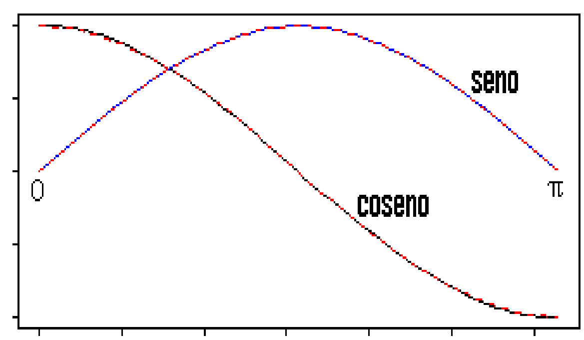
Al evaluar $X=\pi/3$ se obtiene $P_{4.sen}(\pi/3)=0.8662914$, siendo, la mejor aproximación $seno(\pi/3)= \sqrt{¾} \approx 0.8660254$, con error absoluto $\varepsilon=abs(0.8662914-0.8660254)\approx 2.66e\text{-}4$.

Para la función COSENO, tomando los mismos cinco valores para el ajuste del polinomio de grado cuatro, resulta un ajuste cúbico con la función:

$$\boxed{P_{3.cos}(X)_5 = 1/3*(2*X/\pi-1)*\{ (4\sqrt{2}-4)*(2*X/\pi-1)^2 - (4\sqrt{2}-1) \} \quad 0 \leq X \leq\pi}$$

Al evaluar $X=\pi/3$, y fijo el valor $\sqrt{2}\approx1.41412$, se obtiene $P_{3.cos}(\pi/3)=0.4969679$, siendo el valor exacto $coseno(\pi/3)=0.50$, con error $\varepsilon=abs(0.4969679-0.50) \approx 3.03e\text{-}3$.

El gráfico de cada función trigonométrica y la correspondiente interpolación en [0, π], muestra lo natural del grado elegido para cada función, y aún se podría decir de ajuste casi perfecto para ambas, luego veremos este grado de perfección en la función de errores. Al ver los errores alcanzados, se llega a que el mejor polinomio para ajustar tanto a la función seno como a la función coseno, es el de grado 4.



Otros polinomios para SENO son de orden 2 y 3, que usando el método de mínimos cuadrados, es solo excluir dos o una, filas y columnas finales del SEN en MC ya construido de previo, aunque luego la solución para ambos sistemas confluyen a un mismo polinomio de grado 2 (en cinco puntos):

$$P_{2.sen}(X)_5 = 1/7*\{ (3.4+2.4\sqrt{2}) - (4+2\sqrt{2})*(2*X/\pi-1)^2 \}$$

Polinomio cuadrático en tres puntos $\{(0, 0), (\pi/2, 1), (\pi, 0)\}$ es: $P_{2.sen}(X)_3=1-(2*X/\pi-1)^2$

Similitud entre funciones dadas e interpoladas (Proposición Ixim)

La función seno es positiva en $[0, \pi]$, el valor de la integral en el dominio es a=2. Luego la función normalizada es $f_{d.sen}(X)=sen(X)/2$. Se calcula por integración directa los primeros 4 momentos centrales (discriminantes): $E(X)=\pi/2$; $Var(X)=(\pi^2-8)/4$; $Sk(X)=0$; $Kt(X)=1+2*(10-\pi^2)/(¼\pi^2-2)^2$. El polinomio interpolante de grado 4 para seno es $P_{4.sen}(X)_5 \geq 0$; fácil se calcula el valor de la integral es $a=2*\pi/45*(3+8\sqrt{2})$, y el polinomio normalizado es $P_{4.sen}(X)/a$, así se calculan los cuatro parámetros, cuya deducción explicita de las integrales respectivas se omite, pues siguen siendo polinomios de grado dos, hasta orden ocho. Para cada uno de estas funciones, la media es siempre $\mu = \pi/2$ y el sesgo es siempre $ς = 0$. También los polinomios cuadráticos en 5 y 3 puntos se incluyen sus parámetros en esta tabla:

$f_d(X)_n$	Área función (a)	σ^2	κ	$I_n(\mu)$
$f_{d.sen}(X)$	2.000000000000	0.46740110027234	2.1937502300	2.139490
$P_{d.4.sen}(X)_5$	1.998570731824	0.46643950642437	2.1940681686	10.71950
$P_{d.2.sen}(X)_5$	2.027660272889	0.49098152213010	2.1428016432	10.18368
$P_{d.2.sen}(X)_3$	2.094395102393	0.49348022005447	2.1428571429	6.079271

De la tabla anterior se señala, que el área, la varianza y la curtosis son más próximas las del polinomio de grado cuatro a las correspondientes de la función seno. Sin considerar los respectivos de la función dada, un criterio sencillo es el de información máxima, que como vemos se alcanza en el polinomio de grado 4 (la función normalizada del seno es simétrica y acampanada en el intervalo dado).

Ajuste polinómico para la función COSENO: Bien se pudo haber definido el intervalo en $[-\pi/2, \pi/2]$, para ajustar el coseno, pero es más fácil por Interpolación Segmentaria del seno para el coseno: Las funciones seno y coseno, guardan relación entre ellas, pues están desplazadas a una distancia de π tanto hacia la izquierda como hacia la derecha, por lo cual podemos definirla en dos segmentos, por desplazamiento en dos sub-intervalos del dominio [0 π] así:

$$P_{4.\cos}(X)=\begin{cases} P_{4.sen}(X+\pi/2)\ ; & [\,0 \le X \le \pi/2\,] \\ -P_{4.sen}(X-\pi/2)\ ; & [\,\pi/2 \le X \le \pi\,] \end{cases}$$

O bien se define en [0 π] $P_{4.\cos}(X) = \text{ifelse}\{x \le \pi/2, P_{4.sen}(X+\pi/2), -P_{4.sen}(X-\pi/2)\}$

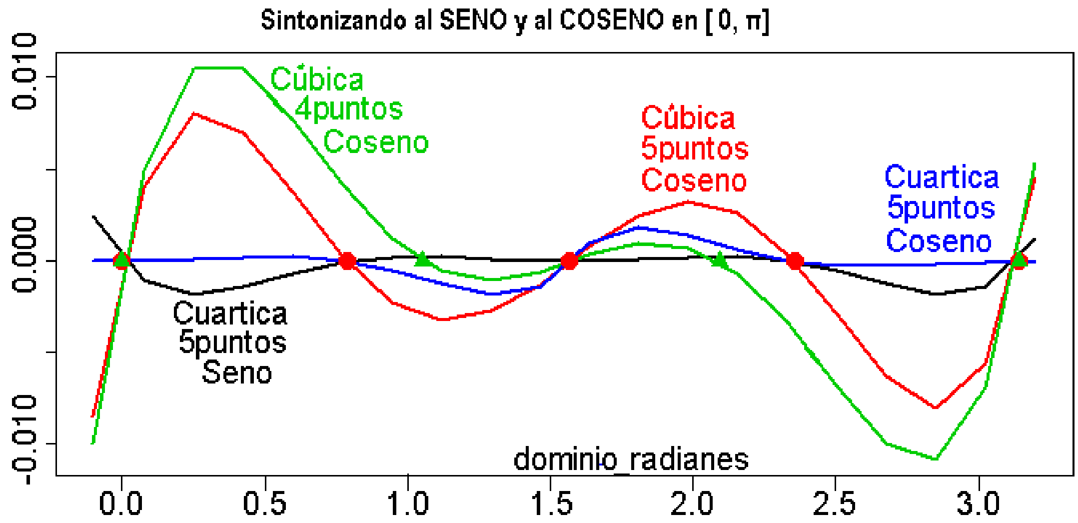
Similitudes entre las funciones interpolantes del coseno

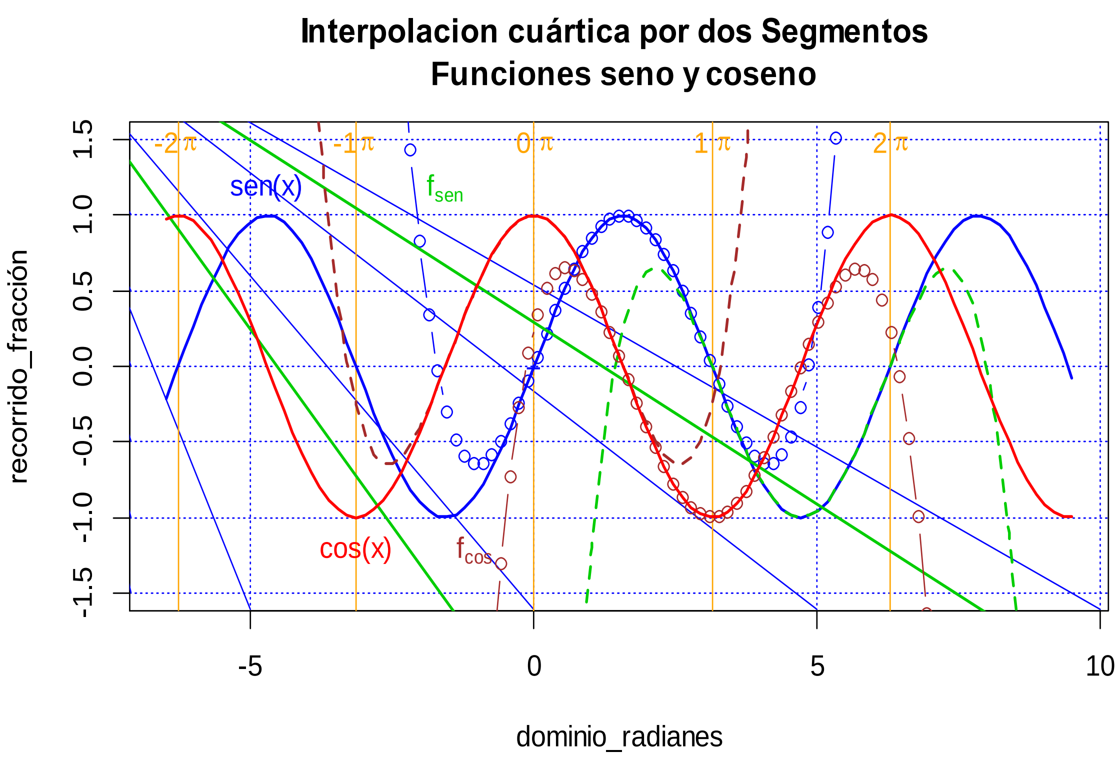
Primero notamos que la función coseno no es siempre positiva en [0, π], teniendo dos alternativas para redefinirla: sumar uno: $f(X)=1+\cos(X) \ge 0$, o bien trabajar con el valor absoluto $f(X)=ABS(\cos(X))$ en ese intervalo. Si se escogiera valor absoluto, la integral resulta con a=2, y la función normalizada es $f_d.\cos(X)=ABS(\cos(X))/2$, $\mu=\pi/2$ y $\varsigma=0$. Elijamos la asimétrica $f(X)=1+\cos(X)$, cuya integral es $a=\pi$ (incluso para cada polinomio ajustado), su normalización es $fd(X)=(1+\cos(X))/\pi$ y ahora comparémosla a las interpolaciones cúbica y cuártica ya normalizadas por incremento de 1.

$f_d(X)_n$	μ	σ^2	ς	κ	$I_n(\mu)$
$(1+\cos(x))/\pi$	0.9341766	0.4171823	0.6015864	2.59525	2.397034
$P_{3.\cos}(X)_5$	0.9315273	0.4138021	0.5911835	2.56006	12.08307
$P_{4.\cos}(X)_5$	0.9342649	0.4172948	0.6010767	2.59304	11.98194

Según información máxima se elegiría al cúbico, pero nuestra función normalizada es asimétrica. Recordemos que para un modelo simétrico y acampanado sí sería conveniente el criterio de información pero no en general. También se puede determinar la información para este caso asimétrico. Si tomáramos el intervalo [-π/2, π/2] para ajustar el coseno se llega al mismo resultado obtenido por el seno, Del gráfico (página siguiente) de las funciones residuales o de diferencias de cada polinomio interpolante correspondiente, para comparar precisiones de ajuste se debe elegir el de grado cuatro.

Por último se grafican ambas funciones seno y coseno y los polinomios de interpolación de grado cuatro respectivos en la página siguiente, estos pueden desplazarse tantas veces se quiera hacia la izquierda o hacia la derecha y se tendría la trayectoria natural de cualquiera de estas funciones, se incluyen los trazos fuera de los intervalos de los polinomios por cuestiones didácticas.





En caso de querer interpolar estas funciones trigonométricas fuera del intervalo [0, π], se debe trasladar X+mπ, unidades hacia atrás la función o bien X−mπ unidades hacia adelante, para m entero.

Polinomio interpolante de Newton

Éste método es algorítmico y resulta sumamente cómodo en determinados casos, sobre todo cuando se quiere calcular un polinomio interpolador de grado elevado. Para un polinomio de n-ésimo orden se requieren n+1 puntos: $(x_0; f(x_0))$, $(x_1; f(x_1))$ $(x_2; f(x_2))$,…$(x_n; f(x_n))$ y para determinarlo usamos la fórmula:

$$P_n(x) = b_0+b_1(x-x_0)+b_2(x-x_0)(x-x_1)+\ldots+b_n(x-x_0)(X-X_1)(X-X_{n-1})$$

o que es lo mismo: $P_n(x)= b_o + \sum_{i=1}^{n} b_i \prod_{j=1}^{i-1}(x-x_i)$

Los coeficientes se calculan mediante diferencias divididas, con su expresión general:

$$f[x_i \ldots x_{i+j+1}] = \frac{f[x_{i+1};\ldots;x_{i+j+1}]-f[x_i;\ldots;x_{i+j}]}{(x_{i+j+1}-x_i)}$$

Como se ve en la fórmula, las diferencias divididas se calculan de modo recursivo usando coeficientes anteriores. Una vez hayamos realizado todos los cálculos, notaremos que hay (muchas) más diferencias divididas que coeficientes b_i

El cálculo de todos los términos intermedios debe realizarse simplemente porqué son necesarios para poder formar todos los términos finales. Sin embargo, los términos usados en la construcción del polinomio interpolador son todos aquéllos que involucren a x_0, así: $b_o = f[x_0]$, $b_1= f[x_0; x_1]$…$b_i = f[x_0;x_1;\ldots;x_i]$. Con esta notación, podemos re expresar el polinomio p_n en dos formas (simple y anidada) como:

$$p_n(x)=f[x_0]+f[x_0;x_1](x-x_0)+f[x_0;x_1;x_2](x-x_0)(x-x_1)+\ldots+f[x_0;x_1;\ldots;x_n](x-x_0)(x-x_1)\ldots(x-x_{n-1})$$

$$p_n(x)=f[x_0]+(x-x_0)\{f[x_0;x_1]+(x-x_1)\{f[x_0;x_1;x_2]+\ldots+(x-x_{n-1})\{f[x_0;x_1;\ldots;x_n]*(x-x_{n-1})\}\ldots\}\}$$

A esta ecuación se le conoce con el nombre de fórmula de diferencias divididas interpolante de Newton.

Polinomio interpolante de Lagrange

El método de Lagrange consiste en encontrar una función que pase por n puntos dados. Para evitar el cálculo de las diferencias finitas que se hace en el polinomio de Newton, el método de Lagrange propone una fórmula más sencilla, pero que por supuesto, por ser una aproximación puede tener un ligero margen mayor de error.

Aun así, es un método muy práctico. Reformular el polinomio de Newton evita el cálculo por diferencias divididas:

$$\mathrm{P_n}(\mathrm{x}) = \sum_{i=0}^{n} L_i(x)f(x_i), \quad \text{donde} \quad L_i(\mathrm{x})=\prod_{\substack{j=0\\ j\neq i}}^{n} \frac{x-x_j}{x_i-x_j},$$

Π designa el producto de los términos desde j=0 hasta n.

Esta formulación tiene la desventaja de que, si se desea agregar un punto extra al conjunto de punto, se deben volver a realizar todos los cálculos para la obtención del polinomio. En la siguiente sección se establecerá una nueva formulación para la cual es relativamente fácil agregar un nuevo dato al conjunto de puntos de interpolación y aprovechar los cálculos efectuados.

Ejemplo: Continuando con la función trigonométrica básica (seno) en cinco puntos dados X=[0 π/4 π/2 3π/4 π] y Y=[0 √½ 1 √½ 0]. Igual que antes, se centran los datos X en la variable re-escalada $\boxed{Z = 2(2*X/\pi-1)}$; Z=[-2, -1, 0, 1, 2] y veamos cada ajuste.

Polinomio de Lagrange, en los datos simples de X: $L_{ni}(\mathrm{x})=\prod_{\substack{j=0\\ j\neq i}}^{n} \frac{x-x_j}{x_i-x_j}$

i	Términos polinómicos	$\mathrm{L_{4.i}}(\mathrm{X})$
0	$\dfrac{\left(x-\frac{\pi}{4}\right)\left(x-\frac{\pi}{2}\right)\left(x-\frac{3}{4}\pi\right)(x-\pi)}{\left(0-\frac{\pi}{4}\right)\left(0-\frac{\pi}{2}\right)\left(0-\frac{3}{4}\pi\right)(0-\pi)}$	$\dfrac{32}{3\pi^4}\left(x-\dfrac{\pi}{4}\right)\left(x-\dfrac{\pi}{2}\right)\left(x-\dfrac{3}{4}\pi\right)(x-\pi)$
1	$\dfrac{(x-0)\left(x-\frac{\pi}{2}\right)\left(x-\frac{3\pi}{4}\right)(x-\pi)}{\left(\frac{\pi}{4}-0\right)\left(\frac{\pi}{4}-\frac{\pi}{2}\right)\left(\frac{\pi}{4}-\frac{3\pi}{4}\right)\left(\frac{\pi}{4}-\pi\right)}$	$\dfrac{-128}{3\pi^4}(x-0)\left(x-\dfrac{\pi}{2}\right)\left(x-\dfrac{3}{4}\pi\right)(x-\pi)$
2	$\dfrac{(x-0)\left(x-\frac{\pi}{4}\right)\left(x-\frac{3}{4}\pi\right)(x-\pi)}{\left(\frac{\pi}{2}-0\right)\left(\frac{\pi}{2}-\frac{\pi}{4}\right)\left(\frac{\pi}{2}-\frac{3}{4}\pi\right)\left(\frac{\pi}{2}-\pi\right)}$	$\dfrac{64}{\pi^4}(x-0)\left(x-\dfrac{\pi}{4}\right)\left(x-\dfrac{3}{4}\pi\right)(x-\pi)$
3	$\dfrac{(x-0)\left(x-\frac{\pi}{4}\right)\left(x-\frac{\pi}{2}\right)(x-\pi)}{\left(\frac{3}{4}\pi-0\right)\left(\frac{3}{4}\pi-\frac{\pi}{4}\right)\left(\frac{3}{4}\pi-\frac{\pi}{2}\right)\left(\frac{3}{4}\pi-\pi\right)}$	$\dfrac{-128}{3\pi^4}(x-0)\left(x-\dfrac{\pi}{4}\right)\left(x-\dfrac{\pi}{2}\right)(x-\pi)$
4	$\dfrac{(x-0)\left(x-\frac{\pi}{4}\right)\left(x-\frac{\pi}{2}\right)\left(x-\frac{3}{4}\pi\right)}{(\pi-0)\left(\pi-\frac{\pi}{4}\right)\left(\pi-\frac{\pi}{2}\right)\left(\pi-\frac{3}{4}\pi\right)}$	$\dfrac{32}{3\pi^4}(x-0)\left(x-\dfrac{\pi}{4}\right)\left(x-\dfrac{\pi}{2}\right)\left(x-\dfrac{3}{4}\pi\right)$

Así el polinomio explícito de Lagrange de grado cuatro ya reducido es:

$$\mathrm{P_{4.L}}(\mathrm{X})= 64/(3\pi^4)*X*(X-\pi)*\{ 3(X-\pi/4)*(X-3\pi/4) - \sqrt{2}*(X-\pi/2)*(2X-\pi) \}$$

Tabla de diferencias divididas de Newton en datos $\boxed{Z = 2(2*X/\pi-1)}$ centrados

Z_i	$F[Z_0]$	$f[Z_0; Z_1]$	$f[Z_0; Z_1; Z_2]$	$f[Z_0; Z_1; Z_2 ; Z_3]$	$f[Z_0; Z_1; Z_2 ; Z_3; Z_4]$
-2	0	½ √2	-½ (√2-1)	-1/6*(3-2√2)	1/12*(3-2√2)
-1	½ √2	½ (2-√2)	-½(2-√2)	1/6*(3-2√2)	
0	1	-½ (2-√2)	-½(√2-1)		
1	½ √2	-½ √2			
2	0				

El vector de coeficientes en múltiplos de 12 es: [0, 6√2, 6-6√2, 4√2-6, 3-2√2]

Polinomio de Newton de grado cuatro, anidado en la variable Z, ya reducido:

$P_{4N}(Z) =0+(Z+2)/12*(6\sqrt{2}+(Z+1)*(6-6\sqrt{2} + (Z+0)*(4\sqrt{2}-6 + (Z-1)*(3-2\sqrt{2}))))$

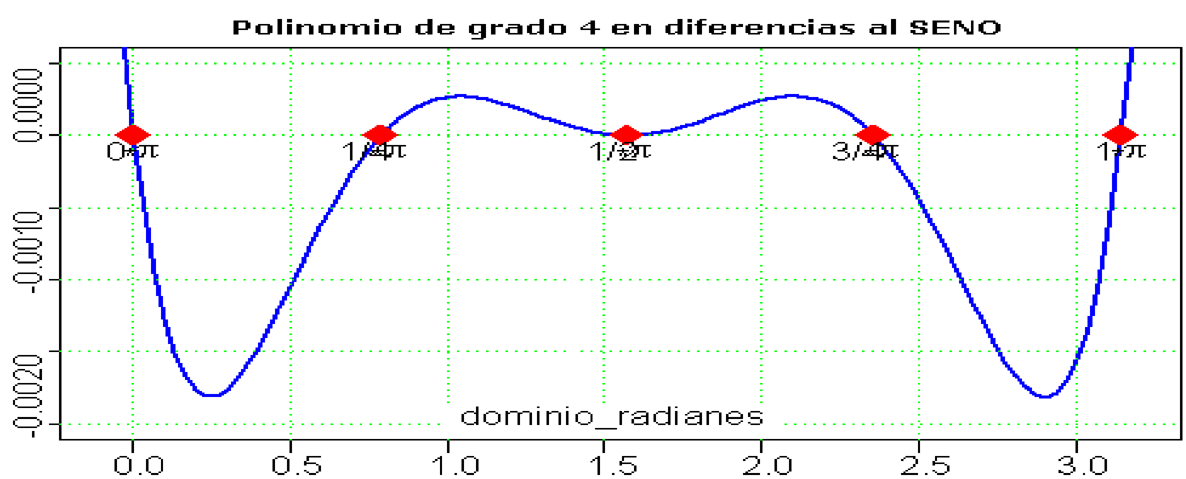
Polinomio de Newton de grado cuatro, anidado en la variable original X, ya reducido:

$P_{4N}(X)=X/(3\pi)*(6\sqrt{2} + (4*X/\pi-1)*(6-6\sqrt{2} + (4*X/\pi-2)*(4\sqrt{2}-6 + (4*X/\pi-3)*(3-2\sqrt{2}))))$

Los gráficos correspondientes a cada polinomio se omite, pues son iguales a los ya hechos para los métodos previos, pero se muestra la función de diferencias al SENO.

Observemos por último los ajustes de grado cuatro en cinco puntos dados para la función seno, por cada método en [0, π] y elíjase aquel más sencillo de evaluar.

$P_{4.sen}(X)_{ED} = 1-1/3*(2*X/\pi-1)^2*\{ (15-8\sqrt{2})-(12-8\sqrt{2})*(2*X/\pi-1)^2 \}$
$P_{4L}(X)= 64/(3\pi^4)*X*(X-\pi)*\{ 3(X-\pi/4)*(X-3\pi/4) - \sqrt{2}*(X-\pi/2)*(2X-\pi) \}$
$P_{4N}(X)=X/(3\pi)*\{6\sqrt{2}+(4*X/\pi-1)*[6-6\sqrt{2}+(4*X/\pi-2)*(4\sqrt{2}-6+(4*X/\pi-3)*(3-2\sqrt{2}))] \}$



Interpolación segmentaria: Trazadores cúbicos

Dados n+1 puntos $(x_0, y_0), (x_1, y_1),\ldots, (x_n, y_n)$ con $x_0, x_1,\ldots, x_n$ números reales diferentes, y f una función de valor real definida en un intervalo [a, b], que contiene a $x_0, x_1,\ldots, x_n$, se pretende aproximar la función f por segmentos o trazas. De antemano vamos a suponer que: $[x_0 < x_1 <\ldots < x_n]$. La idea es aproximar la función f en cada sub-intervalo $[x_k , x_{k+1}]$, k= 0, 1,…, n-1, usando un polinomio de grado menor o igual a tres, el cual supondremos de la forma: $P_3(X)= a_k + b_k (X-x_k) + c_k (X-x_k)^2+ d_k(X-x_k)^3$

<u>Interpolación con polinomios de Hermite</u>**:** El método de Hermite es el caso más general e interpola en un conjunto dado de puntos, no solamente a una función, sino también algunas de sus derivadas. Esto hace que la función y el polinomio interpolante tengan perfiles semejantes en el sentido que las líneas tangentes son las mismas. Precisa mencionar que los métodos vistos anteriormente, por no interpolar derivadas, se denominan interpolación de Lagrange en contraposición a Hermite.

<u>Interpolación Segmentaria Lineal</u>: Este es el caso más sencillo de interpolar una función. En él, vamos a interpolar una función f(x) de la que se nos dan un número N de puntos (x, y) por los que tendrá que pasar nuestra función polinómica P(x). La serie de funciones van a ser lineales, esto es, de grado 1, de la forma $P(X) = a_k+bX$. En un gráfico de puntos al unirlos con líneas rectas cada par de puntos adyacentes, se realiza esta interpolación lineal; para n puntos dados, de forma analítica conlleva determinar n-1 ecuaciones de la recta punto-pendiente.

Definiremos una de estas funciones por cada par de puntos adyacentes, hasta un total de (n-1) funciones, haciéndolas pasar obligatoriamente por los puntos que van a determinarlas, es decir, la función P(x) será el conjunto de segmentos que unen nodos consecutivos; es por ello que nuestra función será continua en dichos puntos, pero no derivable en general.

<u>Interpolación segmentaria cuadrática</u>: En este caso, los polinomios P(X) a través de los que construimos el Splines tienen grado 2. Esto quiere decir, que la forma funcional va a tener la forma $P(X) = aX^2 + bX + c$.

La interpolación cuadrática nos va a asegurar que la función que generemos a trozos con los distintos P(X) va a ser continua, ya que para sacar las condiciones que ajusten el polinomio, vamos a determinar cómo condiciones:

- ✓ Que las partes de la función a trozos P(x) pasen por ese punto, es decir, que las dos $P_n(x)$ que rodean al f(x), sean igual a f(x) en cada uno de estos puntos.
- ✓ Que la derivada en un punto siempre coincida para ambos "lados" de la función definida a trozos que pasa por tal punto común.

En un caso sencillo con f(x) definida en tres puntos y dos ecuaciones P(x) para aproximarla, vamos a tener seis incógnitas en total. Para resolver esto necesitaríamos seis ecuaciones, pero vamos a tener tan sólo cinco: cuatro que igualan el P(x) con el valor de f(x) en ese punto (dos por cada intervalo), y la quinta al igualar la derivada en el punto común a las dos P(x).Se necesita una sexta ecuación, esto suele hacerse con el valor de la derivada en algún punto.

Interpolación Segmentaria Cúbica: En este caso, cada polinomio P(x) a través del que construimos los Splines en [m, n] tiene grado 3. Esto quiere decir, que la forma funcional va a tener la forma $P(x) = ax^3 + bx^2 + cx + d$

En este caso vamos a tener cuatro variables por cada intervalo (a, b, c, d), y una nueva condición para cada punto común a dos intervalos, respecto a la derivada segunda:

- ✓ Que las partes de la función a trozos P(x) pasen por ese punto. Es decir, que las dos Pn(x) que rodean al f(x) que se quiere aproximar, sean igual a f(x) en cada uno de estos puntos.
- ✓ Que la derivada en un punto siempre coincida para ambos "lados" de la función definida a trozos que pasa por tal punto común (punto de nodo).
- ✓ Que la derivada segunda en un punto siempre coincida para ambos "lados" de la función definida a trozos que pasa por tal punto común. Como puede deducirse al compararlo con el caso de Splines cuadráticos, ahora no nos va a faltar una sino dos ecuaciones (condiciones) para el número de incógnitas que tenemos.

La forma de solucionar esto, determina el carácter de los Splines cúbicos. Así, podemos usar, algunas propuestas entre otras no consideradas aquí:

- Splines cúbicos naturales: La forma más típica; la derivada segunda de P se hace 0 para el primer y último punto sobre el que está definido el conjunto de Splines, esto son, los puntos m y n en el intervalo [m, n].
- Dar los valores de la derivada segunda de m y n de forma "manual", en el conjunto de Splines definidos en el intervalo [m, n].
- Incluir un punto adicional para cualesquiera dos segmentos, si se dispusiera de aquellos, para la función de ajuste, o bien del conjunto de datos obtenidos.
- Hacer iguales los valores de la derivada segunda de m y n en el conjunto de Splines definidos en el intervalo [m, n] Splines cúbicos sujetos: La derivada primera de P debe tener el mismo valor que las derivada primera de la función para el primer y último punto sobre el que está definido el conjunto de Splines, esto son, los puntos m y n en el intervalo [m, n].

<u>Interpolación por Splines cúbicos</u>: La palabra inglesa Spline define un regla sujeta a dos puntos, utilizada un dibujo técnico para pintar curvas suaves pasando por puntos especificados. Estas condiciones de suavidad y adaptabilidad son las que hacen que las funciones polinómicas a trozos adopten este nombre. Una función Spline será un polinomio a trozo de grado definido y ciertas propiedades de regularidad (derivabilidad).

La aproximación polinómica segmentaria más común utiliza polinomios de grado tres entre cada par de puntos consecutivos y recibe el nombre de Interpolación de trazadores cúbicos o Splines cúbicos. La curva realizará interpolación del conjunto de puntos de control cuando las secciones polinómicas que ajustan de modo que la curva pasa por cada punto de control. Esa curva obtenida podemos expresarla de forma matemática como una función polinómica cubica cuyas primera y segunda derivada son continuas en cada sección de la curva.

Supongamos que tenemos n+1 puntos: px (xk; yk); donde yk= f (xk); k = 0; 1;…; n
En los cuales se quiere interpolar la función f. Las abscisas no es necesario que sean equidistantes, pero se suponen ordenados, o sea, x0<x1<x2<…<xn . La idea es encontrar polinomios cúbicos qx(x) que interpolen la función f en el subintervalo: [xk; xk+1]; k =0; 1;…; n-1.

Definición de función cúbica a trozos: La función(x) se llama cúbica a trozos o por segmentos en $[x_0, x_n]$ si existen polinomios cúbicos $q_0(x), q_1(x),\ldots, q_{n-1}(x)$ tales que: $S(x)=q_x(x)$ en $[x_q, xk+1]$, para k= 0, 1,…, n-1. Para que s(x) interpole en los puntos p0, p1,…, pn los qk(x) han de verificar: $q_k(x_k)= y_k$ $q_k(x_{k+1}) =y_{k+1}$, k = 0, 1…, n−1.

Lo anterior supone 2n condiciones. Llamaremos a $q_k(x)$ Splines cúbico, o simplemente Splines, si los polinomios P(x) tienen la misma pendiente y la misma concavidad en los nodos que las unen.

Condiciones splines

1. $p_k(x_k) = f(x_k)$, $p_{n-1}(x_n) = f(y_n)$, k = 0, 1,…, n-1 (condición básica de interpolación) Esta condición supone n+1 condiciones.
2. $p_k(x_{k+1})=p_{k+1}(x_{k+1})$, k=0, 1,…, n-1 (condición de continuidad) Esta condición supone n-1 ecuaciones.
3. $p_k'(x_{k+1}) = p_{k+1}'(x_{k+1})$, k = 0, 1,…, n-1 (condición de primera derivada) Esta condición sugiere n-1 condiciones.
4. $p_k''(x_{k+1}) = p_{k+1}''(x_{k+1})$, k = 0, 1,…, n-1 (condición de segunda derivada) Esta condición sugiere n-1 condiciones.

5.a $p'(x_0)=f'(x_0)$ 5.b $p_{n+1}'(x_n)= f'(x_n)$ (condiciones a la frontera)

Para que los p_k interpolen los puntos, se deben verificar las siguientes condiciones:
Al verificar las condiciones 1, 2, 3 y 4, se asegura que los p_k tienen sus primeras y segundas derivadas en los puntos $x_0, x_1,\ldots x_n$, en este caso se dice que los p_k son **trazadores cúbicos** que aproximan la función f. Ahora, si se cumple la condición 5.a, el trazador cúbico se llama **natural**, y si cumple la condición 5.b, el trazador cúbico se llama de **frontera sujeta** (no son mutuamente excluyentes).
En forma matricial, el sistema tridiagonal que resulta es (Splines cubico Natural), los h_i representan las distancias entre abscisas de cada segmento ($h_i=x_i-x_{i-1}$), y los Δy_i lo correspondiente para las ordenadas ($\Delta y_i = y_i-y_{i-1}$), las variables a determinar son σ_i.

$$\begin{bmatrix} 2(h_0+h_1) & h_1 & \cdots & 0 \\ h_1 & 2(h_0+h_1) & \cdots & 0 \\ 0 & h_2 & \cdots & 0 \\ \cdots & \cdots & 2(h_{n-3}+h_{n-2}) & h_{n-2} \\ 0 & 0 & h_{n-2} & 2(h_{n-2}+h_{n-1}) \end{bmatrix} * \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \cdots \\ \sigma_{n-1} \end{bmatrix} = 6* \begin{bmatrix} \Delta y_1 / h_1 - \Delta y_0 / h_0 \\ \Delta y_2 / h_2 - \Delta y_1 / h_1 \\ \Delta y_3 / h_3 - \Delta y_2 / h_2 \\ \cdots \\ \Delta y_{n-1} / h_{n-1} - \Delta y_{n-2} / h_{n-2} \end{bmatrix}$$

Splines cubico natural, es una representación matemática de los Splines del dibujo técnico original (realizado a mano por trazos). Requerimos que dos secciones curvas adyacentes tengan tanto la primera como la segunda derivada igual que su frontera común; es decir exigiremos continuidad C^2

<u>Condiciones de continuidad paramétrica</u>: Para asegurar una transición suave en la curva de un intervalo al siguiente, impondremos una serie de condiciones de continuidad en los intervalos de conexión. Establecemos la continuidad al comparar las derivadas de las secciones de la curva adyacentes en su frontera común.

- continuidad de orden cero C^0; únicamente implica que las curvas se unen.
- Continuidad de primer orden C^1; las primeras derivadas (tangentes a la curva) son iguales en su punto de unión.
- Continuidad de segundo orden C^2; además de lo dicho en C^1 se tiene que cumplir que la derivación de los vectores tangentes según nos acerquemos al punto de unión de ambas curvas por la derecha e izquierda son equivalentes. Esto se logra forzando que la derivada sea igual en la frontera de ambas curvas.

Desventajas de Splines cúbicos: No permiten control local de la curva; si se altera cualquier punto de control de la curva afecta la curva y hay que rehacer los cálculos.

Ejemplo: para la función SENO, aplicar Spline en los tres puntos X= [π/4, π/2, 3π/4], Y= [½√2, 1, ½√2]. Hacer k=½√2, y tomar valores centrados: $X_p=[x_0=-1, x_1=0, x_2=1]$.

Por tres puntos se ajustan dos funciones cúbicas, calcular las dos primeras derivadas:

$f_0(x)=a_0+a_1x+a_2x^2+a_3x^3$	$f_0'(x)= a_1+2*a_2x+3*a_3x^2$	$f_0''(x)= 2*a_2+6*a_3x$
$f_1(x)=b_0+b_1x+b_2x^2+b_3x^3$	$f_1'(x)= b_1+2*b_2x+3*b_3x^2$	$f_1''(x)= 2*b_2+6*b_3x$

Se evalúa cada función en los puntos de cada segmento f_0 en x_0-x_1 y f_1 en x_1-x_2; la primera y segunda derivadas se evalúan el único punto de nodo $x_1=0$ y se igualan.

$f_0(x_0)=f_0(-1)$	$a_0-a_1+a_2-a_3 = k$	$\rightarrow$ $-a_1+a_2-a_3= k-1$ $\rightarrow$ $-a+b-c_0 = k-1$ [1]			
$f_0(x_1)=f_0(0)$	$a_0 = 1$	$f_0'(0)= a_1$	$\rightarrow$ $a:=a_1=b_1$	$f_0''(0)=2*a_2$	$\rightarrow$ $b:=a_2=b_2$
$f_1(x_1)=f_1(0)$	$b_0 = 1$	$f_1'(0)=b_1$		$f_1''(0)=2*b_2$	
$f_1(x_2)=f_1(1)$	$b_0+b_1+b_2+b_3= k$	$\rightarrow$ $b_1+b_2+b_3=k-1$ $\rightarrow$ $a+b+c_1 = k-1$ [2]			

De lo anterior se deducen las funciones $f_0(x)=1+ax+bx^2+c_0x^3$ y $f_1(x)=1+ax+bx^2+c_1x^3$; quedando además por especificar para c_0 y c_1, y resolver las ecuaciones [1] y [2]. Ahora veremos los dos casos: caso natural y el de puntos extremos adicionales.

Caso Spline natural: Al evaluar las segundas derivadas $f_i''(x)= 2*b+6*c_i*x$, para i=0:1, en los puntos inicial o final de cada función respectiva hacer cero: $f_0''(x_0)=0$; $f_1''(x_2)=0$, luego al sumar las ecuaciones [1] y [2], resulta el sistema de tres ecuaciones:

$f_0''(-1)=0 \rightarrow 2b+6c_0(-1) = 0$	$f_1''(1)=0 \rightarrow 2b+6c_1(1) = 0$	$2b+c_1-c_0 = 2k-2$

La solución de este sistema es: $b=¾(\sqrt{2}-2)$, $c_0=¼(\sqrt{2}-2)$, $c_1=-¼(\sqrt{2}-2)$, $a=0$. Así las funciones de cada segmento son: $f_i(x)=1+¼(\sqrt{2}-2)*x^2\{3+(-1)^i*x]\}$, para i=0:1. Ahora en la variable original $x=2*(2X/\pi-1)$, resultan los ajustes por segmentos

$$f_0(X) = 1+ (\sqrt{2}-2)*(2X/\pi-1)^2[3 + (4X/\pi-2)] \quad \text{para } ¼\pi \le X \le ½\pi$$

$$f_1(X) = 1+(\sqrt{2}-2)*(2X/\pi-1)^2[3 - (4X/\pi-2)] \quad \text{para } ½\pi \le X \le ¾\pi$$

Caso Puntos extremos adicionales $f_i(x)$, se incluye para x centrada $f_0(-2)=0$ y $f_1(2)=0$.
$f_0(-2)= 1-2a+4b-8c_0 =0$, $f_1(2)= 1+2a+4b+8c_1 =0$; se tienen cuatro ecuaciones

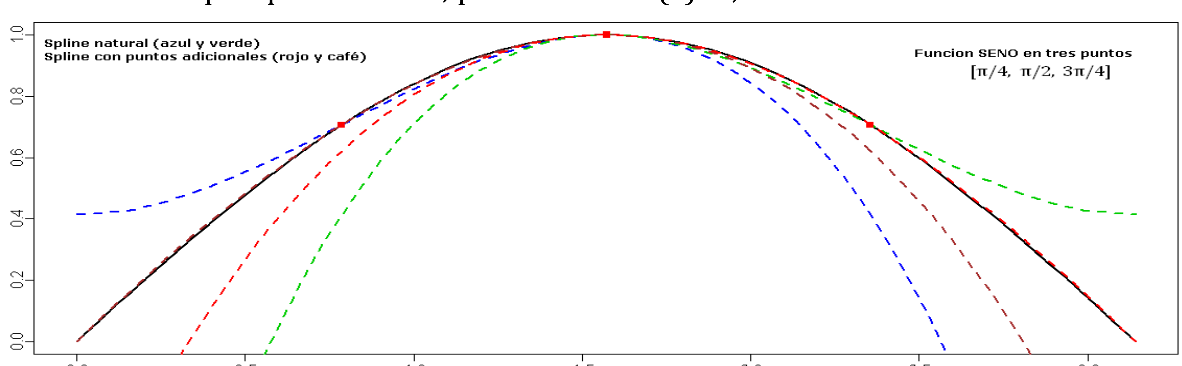
$2a-4b+8c_0 = 1$	$2a+4b+8c_1 = -1$	$-a+b-c_0 = k-1$	$a+b+c_1 = k-1$

Al relacionar las ecuaciones 1º y 3º, y las ecuaciones 2º y 4º, para eliminar c_0 y c_1 respectivamente, se genera un sistema 2x2 en los coeficientes a y b, y la solución es: $b=\sqrt{2}-7/4$; $a=0$; $c_0=¼(2\sqrt{2}-3)$; $c_1=-¼(2\sqrt{2}-3)$. En forma anidada las funciones son $f_i(x)=1+x^2*[b+c_i*x]$, para i=0:1, y luego en la X original:

$$f_0(X)=1+(2X/\pi-1)^2*[(4\sqrt{2}-7)+(4\sqrt{2}-6)*(2X/\pi-1)] \quad \text{para } ¼\pi \le X \le ½\pi$$

$$f_1(X)= 1+(2X/\pi-1)^2*[(4\sqrt{2}-7)-(4\sqrt{2}-6)*(2X/\pi-1)] \quad \text{para } ½\pi \le X \le ¾\pi$$

Abajo los gráficos de funciones de cada caso: se resalta que puntos adicionales conviene más que spline natural; para el seno $f_i''(x)=0$, no es natural en el intervalo.



<u>Similitud entre función dada e interpoladas por segmentos</u> (Proposición Ixim)

Consideraremos el Caso Puntos Extremos adicionales en el intervalo [0, π]. Cada función del segmento respectivo es dada por (expandiendo el intervalo previo):

$$f_0(X)=1+(2X/\pi-1)^2*[(4\sqrt{2}-7)+(4\sqrt{2}-6)*(2X/\pi-1)] \text{ para } 0 \le X \le \tfrac{1}{2}\pi$$

$$f_1(X)= 1+(2X/\pi-1)^2*[(4\sqrt{2}-7)-(4\sqrt{2}-6)*(2X/\pi-1)] \text{ para } \tfrac{1}{2}\pi \le X \le \pi$$

En forma condicional podemos definir la función en el intervalo [0, π] como:

$$f_{spline}(x) = \text{IFELSE}\{\ x<\tfrac{1}{2}\pi\ ,\ f_0(X),\ f_1(X)\ \}$$

Cada función cúbica es positiva en el intervalo, la integral en el dominio para $f_{spline}(x)$ es a= 2.004559755. Luego la función normalizada es $f_{3spline.sen}(X)_5=f_{spline}(X)/a$.

Se calcula por integración directa o aproximada los primeros 4 momentos centrales o discriminantes: $E(X)=\pi/2$; $Var(X)=0.4713645967$; $Sk(X)=0$; $Kt(X)= 2.190769$; observe que al ser funciones cúbicas son fácilmente integrables incluso manualmente. La información de Fisher resulta $I_n(\mu) = 10.6075$, que comparada con la de grado cuatro por Evaluación Directa nos había dado $I_n(\mu)=10.7195$.

Ahora, si quisiéramos calcular puntos extremos resulta sencillo derivar una función cúbica y aplicar la formula cuadrática, en cambio para grado cuatro se genera una función cúbica y se aplica el *Método de Cardano* (a continuación, pero sin ejemplificar). Observe que resulta posible encontrar las raíces de cualquier segmento cúbico.

Dada la ecuación cúbica $x^3 + a_1x^2 + a_2x + a_3 = 0$. Se calculan las cantidades Q, R, S_1 y S_2, para deducir las tres raíces X_1, X_2, X_3 como (el valor complejo i, es tal que $i^2=-1$):

$$Q = \frac{3a_2 - a_1^2}{9}, \qquad R = \frac{9a_1a_2 - 27a_3 - 2a_1^3}{54}$$

$$S_1 = \sqrt[3]{R + \sqrt{Q^3 + R^2}}, \quad S_2 = \sqrt[3]{R - \sqrt{Q^3 + R^2}}$$

$$\begin{cases} x_1 = S_1 + S_2 - \dfrac{a_1}{3} \\ x_2 = -\dfrac{S_1 + S_2}{2} - \dfrac{a_1}{3} + \dfrac{i\sqrt{3}}{2}(S_1 - S_2) \\ x_3 = -\dfrac{S_1 + S_2}{2} - \dfrac{a_1}{3} - \dfrac{i\sqrt{3}}{2}(S_1 - S_2) \end{cases}$$

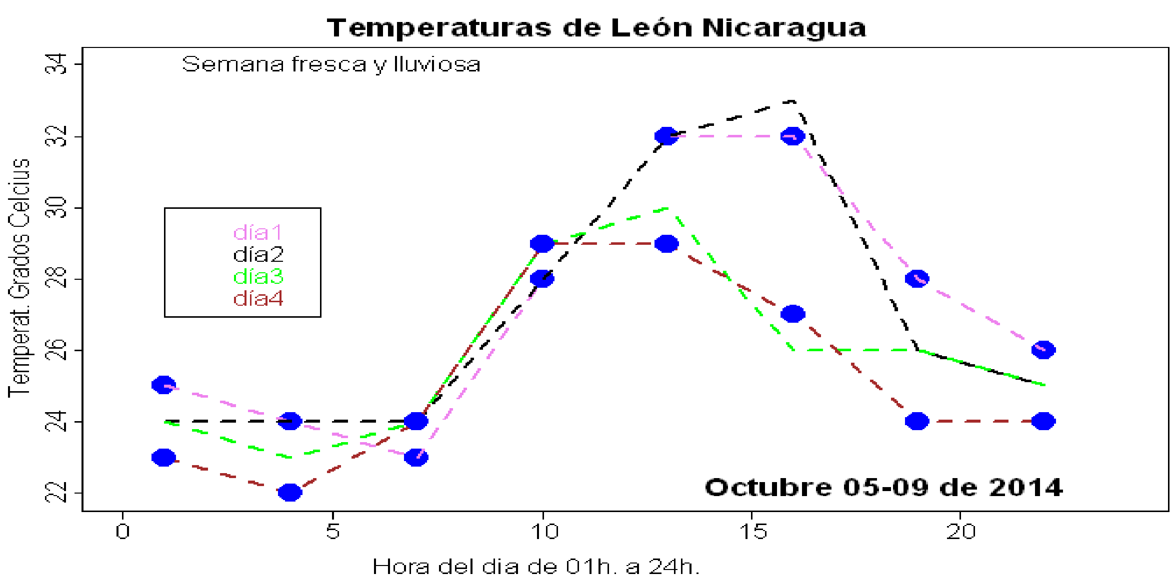
Al ser el discriminante $D = Q^3 + R^2$ se tienen las condiciones:

Si D < 0	Si D = 0	Si D > 0
todas las raíces son reales y distintas	todas las raíces son reales y al menos dos son iguales	una de las raíces es real y dos de ellas son complejas

Aplicaciones con Interpolación Segmentaria

Racionalizando lo irracional: De los ajustes de la función SENO y COSENO se tiene una aplicación de las "cuerdas medias" (significado primigenio dado al seno por matemáticos hindús) del campo trigonométrico y trascendente; hay algoritmos que permiten calcular valores del seno, coseno y tangente sin calculadora. En nuestro caso por operaciones aritméticas aproximamos de forma bastante precisa aquellas dos; relaciones de funciones trigonométricas como $sen^2(x)+cos^2(x)=1$, $sen(x)=cos(x-\pi/2)$ y $tan(x)=sen(x)/cos(x)$ permiten jugar con aquellos polinomios de ajuste en segmentos adyacentes, cerca o lejos del intervalo $[0, \pi]$; para otros ángulos se pueden calcular de forma exacta con las relaciones: $seno(x\pm y)= seno(x)cos(y) \pm cos(x)seno(y)$, $seno(3x)=3seno(x)-4seno^3(x)$; $seno(2x)=2seno(x)cos(x)$; $sen(x/2)=\sqrt{[½*(1-cos(x))]}$, pudiendo calcular para ángulos de: 5°, 10°, 7.5°, etc., resolviendo ecuaciones cúbicas.

Temperaturas extremas de una ciudad en un día: Se toman las temperaturas (página web) durante los días 05 al 09 octubre de 2014, en la ciudad de León, Nicaragua, en periodos de cada tres horas, siendo esa semana fresca y lluviosa (se pudo haber utilizado un termómetro en sombra para períodos uniformes ya fijos). Abajo se muestra el diagrama de líneas para las temperaturas de cuatro días por hora del día.



Al realizar el Análisis de Varianza para los datos de temperatura; por los factores Día (p_1=4) y Período de horas (p_2=8; ambos factores resultan significativos, generándose dos subgrupos de Días y tres subgrupos de horas. Los días 1 y 2 con temperaturas medias altas (27° y 27.25°) y los días 3 y 4 con temperaturas medias bajas (25.25° y 25.88°). Según las horas, 1h-4h-7h tienen temperaturas medias menores de 23.25° a 24°; de 10h-13h-16h con temperaturas medias mayores de 28.5° a 30.75°, y el último periodo 19h-22h con temperaturas medias menores de 25° a 26°; recordemos que esa semana era fresca y lluviosa. A continuación la tabla de Análisis de Varianza.

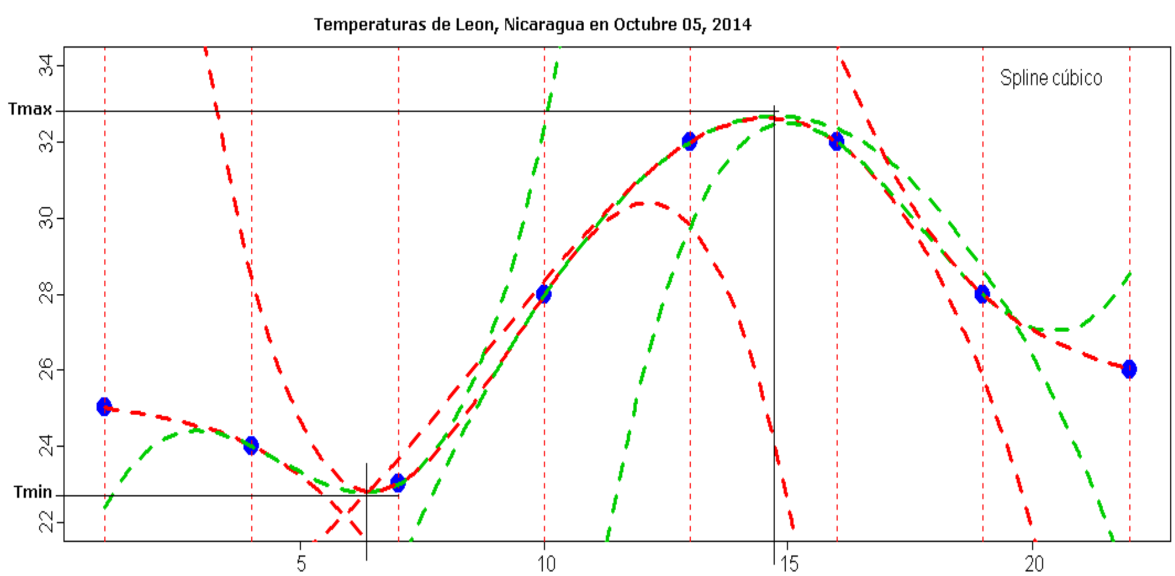
Tabla de Análisis de Varianza. Variable de Respuesta: temperatura (°C)

Fuente de Variación	Suma de Cuadrados	Grados de libertad	Media de Cuadrados	F valor calculado	Pr(>F)	Sig.
Día	21.344	3	7.115	3.8402	0.02452	*
Hora	230.969	7	32.996	17.8096	0.00000	***
Residuos	38.906	21	1.853			
Total	291.219	31				

Códigos de Significatividad: 0 '***' 0.001 ' ** ' 0.010 ' * ' 0.050 ' . ' 0.100 ' NS ' 1.000

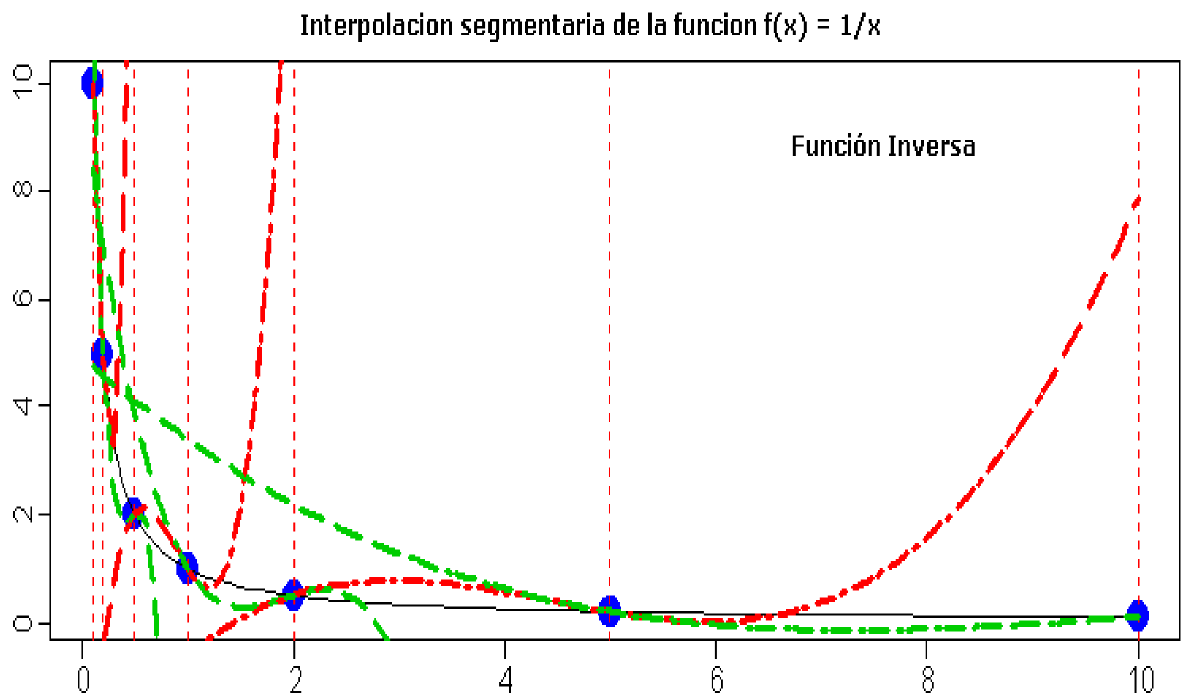
<u>Optimalidad estimada</u>: Si se quiere determinar las máximas y mínimas temperaturas de un día cualquiera, digamos el día2, al tomar los 8 puntos y aplicar interpolación segmentaria, se pueden determinar los 7 polinomios de ajuste local y elegir aquellos dos que contengan los puntos extremos a determinar; se deriva la función cúbica del segmento conteniendo al valor extremo y se resuelve la ecuación cuadrática resultante para cada punto extremo, veamos esto a continuación.

Las funciones que incluyen a las temperaturas extremas de León, al día1 son f2 en el segmento de mínimo [4h a 7h] y f6 en el segmento de máximo [13h a 16h]: punto de mínimo es 6.314 horas con 22.779ºC de temperatura y punto de máximo es 14.998 horas con 32.482ºC de temperatura; observe ambos el siguiente gráfico.



De esta interpolación se ajustó a los puntos del primer día, que corresponde al 05 de octubre del año 2014: la temperatura más baja fue de casi 23.0ºC, a las 6:19 de la mañana y la temperatura más alta fue de casi 32.5ºC cerca de las 3:00 de la tarde.

Peores escenarios en la aproximación segmentaria: se da para funciones no muy curvilíneas tales como aquellas que involucran funciones de aplicaciones científicas y de ingeniería de uso cotidianos: logaritmos, potencias, recíprocas y funciones inversas, entre otras. Veamos lo mal que se ajusta la función inversa en este gráfico.



“ Todo lo que se encuentre en tu mano realizar,

hacelo según tus fuerzas;

porque no hay obra, ni trabajo, ni ciencia,

ni sabiduría en el Seol, que es adonde vos vas. ”

El Predicador, Eclesiastés 9:10

CONCLUSION

La Interpolación segmentaria demuestra poseer una gran finura local; en cambio otros métodos de interpolación que utilizan un solo polinomio, que suele ser de grado mayor; según el número de puntos dados, se vuelven muy imprecisos a medida que se mueve hacia los extremos de aquellos o bien para extrapolación.

La idea medular de la Interpolación segmentaria es que al utilizar varios polinomios de grado bajo, son más manejables para el ajuste, tanto para interpolar, extrapolar, derivar e integrar.

Se enfatiza que entre todas las formas de ajustar puntos, los Splines cúbicos resultan ser los más adecuados casi para cualquier tipo de interpolación según sea la aplicación. Aunque la metodología de Interpolación segmentaria permite utilizar casi cualquier otra función sea polinómica o no, aparte de la cúbica, pues la condición sobre la función es que tenga primera y segunda derivada continuas por los puntos de cada segmento.

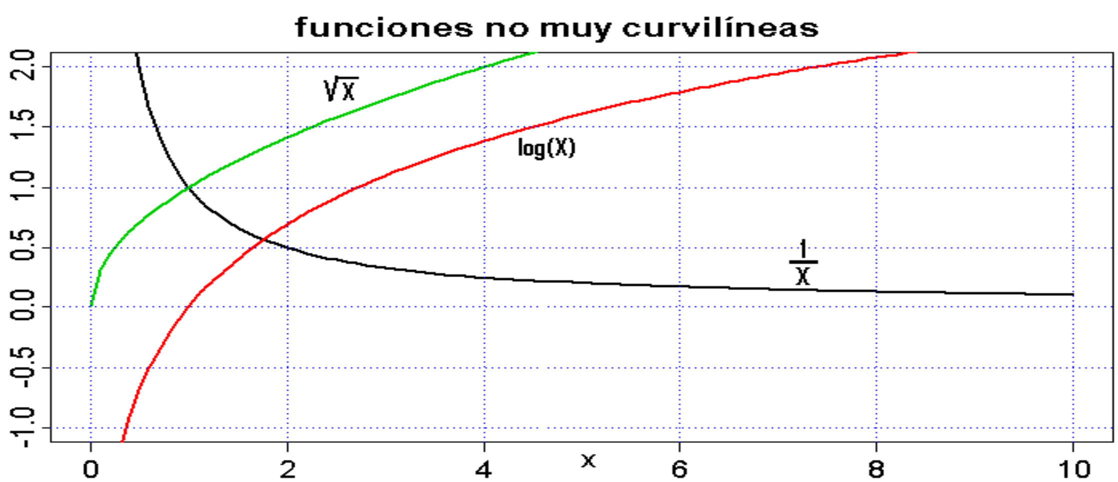
Para la acotación de errores y la similitud de funciones, tener en cuenta los cuatro parámetros de momentos centrales ($\mu \quad \sigma^2 \quad \varsigma \quad \kappa$) no es suficiente para discriminar o decidir cuan bueno es un ajuste sobre una función compleja ya “normalizada”, pues el parámetro de curtosis discrimina bien en las colas y no muy bien en las regiones centrales de las funciones.

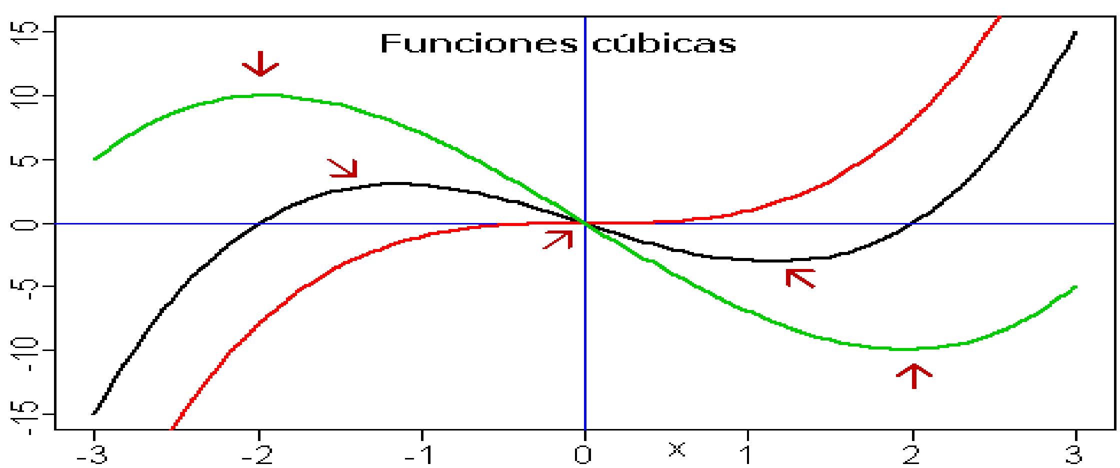
Para las aplicaciones, es posible dibujar gráficas bidimensionales, tridimensionales y aún figuras en 3D con auxilio de parametrización sobre las variables (X, Y, Z), de la forma X(t), Y(t), Z(t) para t dentro del intervalo [0, 1], o bien X(u, v), Y(u, v), Z(u, v), para ambos parámetros u y v en el intervalo [0, 1].

Recomendaciones

Para cálculos a mano, debido a que Interpolación segmentaria involucra la resolución de un sistema de ecuaciones lineales tridiagonal, es muy conveniente trabajarlo en la forma recursiva para reducir los cálculos, que también representan ahorro computacional. Para el caso Spline natural hay que tener cuidado con las segundas derivadas si no son cero de forma natural, se debe inferir que valor sí lo sea.

Hay que tener cuidado que para funciones menos curvilíneas, el ajuste por Interpolación segmentaria puede ser inadecuado, debido a que el comportamiento de una función cúbica es bastante sinuoso alrededor de sus dos puntos extremos, y aun en el caso de un punto de inflexión sigue manteniendo su forma curvilínea pronunciada entre más separados estén los puntos dados para el ajuste.





Anexos: Programas en R

I. Interpolaciones a las funciones trigonométricas seno y coseno

```
# Aproximacion de funciones trigonometricas por Interpolacion segmentaria
# spline-cubica y cuartica para el sen(x)
fp=function(x) 1-1/3*(2*x/pi-1)^2*(15-8*sqrt(2)-(12-8*sqrt(2))*(2*x/pi-1)^2)
dominio_radianes=seq(-6.5,9.5,l=20); recorrido_fracción= seq(-1.5,1.5,l=20)
plot(dominio_radianes, recorrido_fracción, type="n"); pii=expression(pi)
text(0,0,"+", col=4) ; grid(col=4); abline(v=-2:2*pi, col="orange"); e=0.2
text(-2:2*pi-e,1.5, -2:2, col="orange");text(-2:2*pi+e,1.5, pii, col="orange")
curve(sin,col=4, lwd=2, add=TRUE) ; curve(fp, col=4, type="b",add=TRUE)
text(-3*pi/2, 1.2, "sen(x)", col=4); text(-pi/2, 1.2, expression(f[sen]), col=3)
# Se desplaza ahora la funcion para el intervalo [pi, 2*pi]
fnp=function(x) -fp(x-pi)  ;  curve(fnp, col=3, lty=2, lwd=2, add=TRUE)
title(main="Interpolacion cuártica por dos Segmentos \n Funciones seno y coseno")
# Ahora para agregar el coseno cos(X)
fcp=function(x) fp(x+pi/2)  ;  curve(fcp, col="brown", lty=2, lwd=2 , add=TRUE)
fcnp=function(x) -fp(x-pi/2)  ;  curve(fcnp, col="brown", type="b", add=TRUE)
curve(cos, col=2, lwd=2 , add=TRUE);
text(-pi,-1.2, "cos(x)", col=2); text(-pi/3,-1.2, expression(f[cos]), col="brown")

# Funcion de Diferencias para sen(x)
fpdif=function(x) 1-1/3*(2*x/pi-1)^2*(15-8*sqrt(2)-(12-8*sqrt(2))*(2*x/pi-1)^2)- sin(x)
dominio_radianes=seq(-0.1, 3.2,l=20); recorrido_fracción = seq(-0.002,0.0005,l=20)
plot(dominio_radianes, recorrido_fracción, type="n"); pii=expression(pi)
curve(fpdif(x), col=4, lwd=2, add=TRUE); grid(col="green");
x0<-0:4*pi/4 ; points(x0, rep(0,5), pch=18, col=2, cex=2); e=0.05; d=-0.0001
library(MASS) ; text(x0-e, d, fractions(0:4/4)) ; text(x0,d, "*") ; text(x0+e, d, pii)
inf=optimize(fpdif, c(0.0, 0.5))   # fpdif(0.2427626) = -0.001812112 ; minimo absoluto
xl=inf[[1]]; ys=inf[[2]];points(xl,ys, pch=15, cex=0.75); text(2/3,-1/5e2,xl); text(5/6,ys,ys)
optimize(fpdif, c(0.8, 1.2), maximum = TRUE)# fpdif(1.03614)=0.00027, hace un maximo rel.
optimize(fpdif, c(2.5, 3.0))   # fpdif(2.898816) = -0.001812112, hace un minimo
# Se logra una precision de 10^-3 por la funcion cuartica; para seno y coseno.

# Para interpolar cos(x) en [0, pi] con cinco puntos se ajusta la cubica siguiente
fcos= function(x) 1/3*(2*x/pi-1)*((4*sqrt(2)-4)*(2*x/pi-1)^2 - (4*sqrt(2)-1))
dominio_radianes=seq(-6.5,9.5,l=20); recorrido_fracción= seq(-1.5,1.5,l=20)
plot(dominio_radianes, recorrido_fracción, type="n", main="funcion coseno")
text(0,0,"+", col=4) ; grid(col=4); abline(v=-2:2*pi, col="orange"); e=0.2
text(-2:2*pi-e,1.5, -2:2, col="orange");text(-2:2*pi+e,1.5, pii, col="orange")
curve(fcos, add=TRUE, type="b", col=4) ; curve(cos, add=TRUE, col=2)
f4cos= function(x) ifelse(x<=pi/2, fp(x+pi/2), -fp(x-pi/2))
curve(f4cos, add=TRUE, col="brown", lty=2)
fdif= function(x) fcos(x) - cos(x) # Se cosntruye la funcion de residuos
dominio_radianes=seq(-0.1, 3.2,l=20); recorrido_fracción = seq(-0.01,0.01,l=20)
plot(dominio_radianes, recorrido_fracción, type="n", main="Difer. coseno - cúbica")
curve(fdif(x), col=4, lwd=2, add=TRUE); grid(col="green")
rango = 0:4*pi/4 ; text(0,0,"+", col=2) ;  abline(v=rango, col="orange"); e=0.05
text(rango-e,0.01, 0:4/4, col="orange");text(rango+e,0.01, pii, col="orange")
points(rango, rep(0,5), pch=18, col=2, cex=2)
optimize(fdif, c(0, 0.5), maximum=TRUE)   # fpdif(0.2828554) = 0.00806767, hace un max.rel.
optimize(fdif, c(1, 1.5))   # fpdif(1.145) = -0.00327, hace un min.rel.
optimize(fdif, c(1.7, 2.2), maximum=TRUE)   # fpdif(1.996583) = 0.00327, hace un max.rel.
optimize(fdif, c(2.5, 3))   # fpdif(2.858735) = -0.00806767, hace un min.rel.
# Se logra una precision de 10^-2 por la funcion cubica; para seno y coseno.
```

```
#Comparemos precisión de ambas funciones en cinco puntos y la cúbica_coseno por cuatro puntos
f3cos=function(x) 1/16*(2*x/pi-1)*(9*(2*x/pi-1)^2-25) ; f3dif=function(x) f3cos(x) -cos(x)
fcos3= function(X) 1/3*(2*X/pi-1)*{ (4*sqrt(2)-4)*(2*X/pi-1)^2 - (4*sqrt(2)-1) }
fdif3=function(x) fcos3(x)-cos(x)
f4cosD= function(x) ifelse(x<=pi/2, f4sen(x+pi/2), -f4sen(x-pi/2)) - cos(x)
f2senD=function(x) 1/7*{ (3.4+2.4*sqrt(2)) - (4+2*sqrt(2))*(2*x/pi-1)^2 } -sin(x)
dr=dominio_radianes # Se excluye la cuadratica por ser de amplio recorrido
matrizfuncionesdif = cbind(fpdif(dr),  f3dif(dr),fdif3(dr), f4cosD(dr))
matplot(dr,matrizfuncionesdif , type="l", lty=1, lwd=2, col=1:5)
legend(1, 0.01, c("Cuartica5puntosSeno"), text.col=1)
legend(1, -0.005, c("Cubica5pCoseno", "Cubica4pCoseno", "Cuartica5pCoseno"), text.col=2:4)
title("Sintonizando al seno y al coseno en [0 , pi] "); rango4 = 0:3*pi/3;
points(rango, rep(0,5), pch=16, col=2, cex=1.2); points(rango4, rep(0,4), pch=17, col=3)

# Newton y Lagrange
f4L= function(x) 64/(3*pi^4)*x*(x-pi)*( 3*(x-pi/4)*(x-3*pi/4)-r2*(x-pi/2)*(2*x-pi)   )
f4N= function(x) x/(3*pi)*(6*r2+(4*x/pi-1)*(6-6*r2+(4*x/pi-2)*(4*r2-6+(4*x/pi-3)*(3-2*r2))))

# Spline natural (f0nat y f1nat) y Spline con puntos adicionales al seno (f0pad y f1pad)
f0nat= function(x) 1+(sqrt(2)-2)*(2*x/pi-1)^2*( 3 + 2*(2*x/pi-1)   )
f1nat= function(x) 1+(sqrt(2)-2)*(2*x/pi-1)^2*( 3 - 2*(2*x/pi-1)   )
f0pad= function(x) 1+(2*x/pi-1)^2*(4*sqrt(2)-7 +(4*sqrt(2)-6)*(2*x/pi-1)   )
f1pad= function(x) 1+(2*x/pi-1)^2*(4*sqrt(2)-7 -(4*sqrt(2)-6)*(2*x/pi-1)   )
```

II. Interpolación segmentaria y Aplicaciones

```
## Los Splines cubicos son adecuados segun el method definido "natural"
## para interpolar algunas funciones trascendentes como las trigonometricas,
## para otras funciones mas simples quizas sea conveniente por otros method 1/x, log(x), ...
## Comparar integrales de linea para la funcion original y la de splines como un criterio de ajuste
# Funcion definida es la funcion inversa y = 1/x y puntos definidos (Peor escenario para spline)
f=function(x) 1/(x) ; xp=c(0.1,0.2,0.5,1,2,5,10); (yp=f(xp) ) # Ejemplo para funcion inversa
# Desde aqui correr cualquier xp yp o funcion hasta la linea del ciclo for
titulo = "funcion definida" ; subtitulo="Spline cúbico" ;    # Gráfico de la funcion
etiqueta.x = "variable X"  ;  plot(xp, yp, pch=19, cex=2, col=4, xlab=etiqueta.x );
k=length(xp); abline(v=xp, col=2, lty=2); curve(f, add=TRUE) ; title(main= titulo, sub=subtitulo)
h=0 ; for(i in 2:k) h[i-1] <- xp[i]-xp[i-1] ; h    # Incrementos de x son las h´s

# Matriz Tridiagonal del Sistema de Ecuaciones Lineales
coef.diag=0 ; for(i in 2:(k-1)) coef.diag[i-1]<-2*(h[i-1]+h[i]) ; coef.diag
require(Matrix); A <- spMatrix(k-2,k-2,h[2:(k-2)], i= 1:(k-3), j=2:(k-2)) # Kronecker
 TriDiag= A+t(A)+diag(coef.diag) ; d = det(TriDiag) ; Adj = det(TriDiag)*solve(TriDiag)
 TriDiag; names(d) <- "determinante"; d ; Adj # Matrices Tridiagonal, Adjunta y determinante
dy = 0       # Valores de las pendientes
for(i in 1:(k-2)) dy[i] <- ((yp[i+2]-yp[i+1])/(xp[i+2]-xp[i+1])-((yp[i+1]-yp[i])/(xp[i+1]-xp[i])))
dy
# Solucion al Sistema Tridiagonal
 (sigma <- solve(TriDiag, 6*dy) ) ; sigma0 = sigmak = 0 # Para Spline cubico natural
( sigma=c(0, sigma[,1], 0 ) )
# Generador de funciones tricubo
qfun = function(x, p)  sigma[p ]/6*( (xp[p+1]-x)^3/h[p] -h[p]*(xp[p+1]-x) ) +
                 sigma[p+1]/6*( (x-xp[p])^3/h[p] - h[p]*(x-xp[p]) ) +
                  yp[p]*( xp[p+1]-x )/h[p] + yp[p+1]*( x-xp[p] )/h[p]
color=c(2,3,2,3, 2,3,2,3)
for(i in 1:k) curve(qfun(x,i), add=TRUE, col=color[i], lty=3, lwd=3, )
# En qfun(x, i) se define cada función spline cúbica
```

```
## Probar con las funciones sin(x), sin(1/x), 1/sin(x), log(x), log10(x)
## Para la propia de R: spline(xp, yp)
## Compare log10 en xp= c(0.1,0.2,0.5,1,2,5,10) e incluir 100, y luego 1000; !que mal ajuste!
## Usar  xp=c(-7, -3,-2,-1,0,1,2,3,7); f= function(x) 1/(x^2+1); yp=f(xp)
## Resultando yp= 0.02 0.10 0.20 0.50 1.00 0.50 0.20 0.10 0.02
## Usar colores simetricos en el ciclo for color=c(2:4, 5, 4:2)
## Para seno(x) usar xp=seq(0, 2*pi, length=9); f= function(x) sin(x) ; yp=f(xp)
## Para log10(x) usar xp=c(0,  )

##### Aplicaciones a temperaturas diarias de Leon Nicaragua
### Octubre 05-09, 2014 ("semana fresca y lluviosa")
# agregar al último grafico

# Temperaturas de la ciudad de León Nicaragua
hora = c(1,4,7,10,13,16,19,22);  xp=hora ; temperat1=c(25,24,23,28,32,32,28,26); yp=temperat1;
  titulo = "Temperaturas de León Nicaragua \n Octubre 05-09 de 2014"
  subtitulo= "Spline cúbico" ; etq.x="Hora del dia de 01h. a 24h."; etq.y= "Temperat. Grados Celcius"
# Gráfico de la funcion
  plot(seq(0,24, by=1), seq(22, 34, length=25), pch=19, cex=2, col=4, xlab=etq.x, ylab=etq.y,type="n");
 points(xp, yp, pch=19, cex=2, col=4, xlab=etiqueta.x ); title(titulo)
  text(5, 34, "Semana fresca y lluviosa")
# Correr el programa desde la linea 8 hasta antes de las lineas del rotulos ## luego las siguientes:
lines( xp, yp, lwd=2, lty=2, col="violet") # Esto agregarlo despues de corrido el Spline
temperat2=c(24,24,24,28,32,33,26,25); yp=temperat2; lines( xp, yp, lwd=2, lty=2, col="black")
temperat3=c(24,23,24,29,30, 26,26,25);yp=temperat3; lines( xp, yp, lwd=2, lty=2, col="green")
temperat4=c(23,22,24,29,29,27,24,24) ;yp=temperat4; lines( xp, yp, lwd=2, lty=2, col="brown")
legend(1,30, c("día1", "día2", "día3", "día4"), text.col=c("violet","black", "green", "brown"))
# Analisis de Varianza en dos factores sin repeticion (DBCA)
temperatura=c(temperat1,temperat2,temperat3,temperat4); Hora=factor(rep(hora,4))
Dia=factor(sort(rep(6:9,8)));  anova(lm(temperatura~Dia+Hora)) ;
TempLeon=data.frame(temperatura, Dia, Hora)

# f0nat y f1nat: spline natural y f0pad y f1pad con puntos adicionales al seno:
f0nat= function(x) 1+(sqrt(2)-2)*(2*x/pi-1)^2*( 3 + 2*(2*x/pi-1)   )
f1nat= function(x) 1+(sqrt(2)-2)*(2*x/pi-1)^2*( 3 - 2*(2*x/pi-1)   )
f0pad= function(x) 1+(2*x/pi-1)^2*(4*sqrt(2)-7 +(4*sqrt(2)-6)*(2*x/pi-1)   )
f1pad= function(x) 1+(2*x/pi-1)^2*(4*sqrt(2)-7 -(4*sqrt(2)-6)*(2*x/pi-1)   )

# Aplicaciones Binomial y Normal
xp=-3:3 ; yp=dnorm(xp);  curve(dnorm, -3,3); points(xp, yp, pch=22, col=2)
xp=-5:5; yp=pnorm(xp); curve(pnorm, -5,5); points(xp, yp, pch=22, col=2)
points(spline(xp, yp,method ="natural"), pch=18, cex=0.5 , col=4)
```

Bibliografía

- Tutorial de Análisis Numérico, Interpolación: Splines cubico. Jesús García Quesada. Departamento de Informática y Sistemas. Universidad de Las Palmas de Gran Canaria, 35017 Campus de Tafira, España.
- Análisis Numérico: Matrices y Métodos directos de resolución de SEL. Manuel Díez Minguito.
- Análisis numérico avanzado Ahmed Ould Khaoua, Ph. D.
- Splines curvas y superficies. Introducción al dibujo de curvas de aproximación e interpolación por computador. Carlos Gonzales Morcillo.
- Análisis numérico José Alberto Gutiérrez Robles. Miguel Ángel Olmos Gómez, Juan Martín Casillas González. Universidad de Guadalajara
- Probabilidad y Estadística, Morris H. DeGroot. Addison-Wesley Iberoamericana.

Páginas web examinadas

- http://www.qualitydigest.com/inside/quality-insider-article/problems-skewness-and-kurtosis-part-one.html
- http://www.qualitydigest.com/inside/quality-insider-article/problems-skewness-and-kurtosis-part-two.html
- http://www.qualitydigest.com/inside/quality-insider-article/problems-skewness-and-kurtosis-part-one.html
- http://www.qualitydigest.com/inside/quality-insider-article/problems-skewness-and-kurtosis-part-two.html
- http://es.wikipedia.org/wiki/Seno_(trigonometría)
- http://es.wikipedia.org/wiki/Ecuación_de_tercer_grado