MÉTODOS NUMÉRICOS 3006907

NOTAS DE CLASE - SEMANA 06

MINIMIZACIÓN DE ERROR Y APROXIMACIÓN DISCRETA POR MÍNIMOS CUADRADOS

Ahora, la pregunta es ¿será posible hallar un polinomio de grado n que interpole a f que posea un error ***mínimo***?

Sabemos que si $f \in \mathcal{C}^{n+1}[a,b]$ y p_n es el polinomio que interpola a f en los nodos $a \leq x_0 < \cdots < x_n \leq b$, para cada $x \in [a,b]$ existe $\widetilde{x} \in (a,b)$ tal que

$$f(x) - p_n(x) = \frac{f^{(n+1)}(\widetilde{x})}{(n+1)!} \prod_{i=0}^{n} (x - x_i),$$

así que si queremos minimizar el error de aproximación, *podemos* buscar minimizar la expresión $\prod\limits_{i=0}^{n}(x - x_i)$ y para ello, *buscamos* ***nodos*** x_i que minimicen la expresión. Estos nodos serán ceros de los polinomios de Chebyshev. Primero consideremos el intervalo $[-1,1]$ y luego el intervalo $[a,b]$.

★ Consideremos el polinomio p_n que interpola a una función f definida en $[-1,1]$ en los nodos $-1 \leq t_0 < \cdots < t_n \leq 1$. Denotemos

$$Q(t) := (t - t_0)(t - t_1)\cdots(t - t_n) = \prod_{i=0}^{n}(t - t_i),$$

luego, la cota de error esta dada por

$$|E(t)| \leq \frac{1}{(n+1)!} \max_{-1\leq t\leq 1} |f^{(n+1)}(t)| \max_{-1\leq t\leq 1} |Q(t)|.$$

El problema resuelto por Chebyshev, consistió en seleccionar el conjunto de nodos $\{t_i\}_{i=0}^{n}$ que haga mínimo el valor $\max\limits_{-1\leq t\leq 1} |Q(t)|$. Para ello emplearemos los polinomios de Chebyshev, empezamos recordando unas propiedades de los mismos.

Propiedades de los polinomios de Chebyshev

- Los polinomios de Chebyshev se definen de manera recursiva

$$\left.\begin{array}{l} T_0(t) = 1 \\ T_1(t) = t \end{array}\right\} \qquad \Rightarrow \qquad T_k(t) = 2t\,T_{k-1}(t) - T_{k-2}(t) \qquad k = 2, \ldots$$

  Por ejemplo

$$\begin{aligned} T_2(t) &= 2t\,T_1(t) - T_0(t) = 2t\,(t) - 1 = 2t^2 - 1 \\ T_3(t) &= 2t\,T_2(t) - T_1(t) = 2t\,(2t^2 - 1) - t = 4t^3 - 3t \\ T_4(t) &= 2t\,T_3(t) - T_2(t) = 2t\,(4t^3 - 3t) - (2t^2 - 1) = 8t^4 - 8t^2 + 1 \end{aligned}$$

- El coeficiente de t^n en T_n es 2^{n-1}, para $n \geq 1$.
- Cuando n es par el polinomio T_n es una función par, es decir, $T_n(-t) = T_n(t)$. Cuando m es impar el polinomio T_m es una función impar, es decir, $T_m(-t) = -T_m(t)$.
- Los polinomios de Chebyshev tienen una representación trigonométrica en $[-1,1]$

$$T_n(t) = \cos(n\cos^{-1}(t)) \qquad -1 \leq t \leq 1.$$

- Los polinomios T_n tiene n ceros distintos t_k que están todos en el intervalo $[-1,1]$. En efecto

$$T_n(t)=0 \quad \Leftrightarrow \quad \cos(n\cos^{-1}(t))=0 \quad \Leftrightarrow \quad n\cos^{-1}(t)=(2k+1)\frac{\pi}{2}$$

luego los ceros son

$$t_k=\cos\left(\frac{2k+1}{2n}\,\pi\right) \qquad k=0,\ldots,n-1. \qquad \textbf{\textit{Nodos de Chebyshev en }} [-1,1].$$

- Se cumple que $|T_n(t)|\le 1$ para $-1\le t\le 1$, debido a la representación trigonométrica $|T_n(t)|=\left|\cos(n\cos^{-1}(t))\right|\le 1$.

Para minimizar $\max\limits_{-1\le t\le 1}|Q(t)|$ Chebyshev descubrió que $t_0,\ldots,t_n$ deben ser elegidos de manera que $Q(t)=\dfrac{1}{2^n}T_{n+1}(t)$.

Teorema. *Entre todas las posibles elecciones del factor $Q(t)$ del error, es decir, entre todas las posibles elecciones de nodos distintos $\{t_i\}_{i=0}^n$ en $[-1,1]$, el polinomio $T(t)=\dfrac{1}{2^n}T_{n+1}(t)$ es la única elección que satisface*

$$\max_{-1\le t\le 1}|T(t)|\le \max_{-1\le t\le 1}|Q(t)|,$$

es más, $\max\limits_{-1\le t\le 1}|T(t)|=\dfrac{1}{2^n}$.

Este teorema puede resumirse diciendo que el menor valor de la cota del error

$$\frac{1}{(n+1)!}\max_{-1\le t\le 1}|f^{(n+1)}(t)|\max_{-1\le t\le 1}|Q(t)|$$

se alcanza si los nodos son los ceros del polinomios de Chebyshev T_{n+1}.

Ejemplo Aproximar la función $f(x)=\frac{4}{4x^2+1}$ en $[-1,1]$, por medio del polinomio interpolante de grado menor o igual a 4 empleando nodos igualmente espaciados y empleandos los respectivos nodos del polinomio de Chebyshev T_5. ¿Como se comportan los errores?

Solución: Primero empleemos nodos igualmente espaciados, es decir, queremos aproximar a la función f por medio de un polinomio interpolante de grado menor o igual a 4, para ello necesitamos 5 nodos, tomamos 5 nodos igualmente espaciados en $[-1,1]$

$$x_0=-1, \qquad x_1=-\frac{1}{2}, \qquad x_2=0, \qquad x_3=\frac{1}{2}, \qquad x_4=1,$$

empleando el polinomio interpolante de Newton

$$p_4(x)=a_0+a_1(x-x_0)+a_2(x-x_0)(x-x_1)+a_3(x-x_0)(x-x_1)(x-x_2)+a_4(x-x_0)(x-x_1)(x-x_2)(x-x_3),$$

la tabla de diferencias divididas es

x_i	D.D. cero	1^{ras} D.D.	2^{das} D.D.	3^{ras} D.D.	4^{ras} D.D.
$x_0=-1$	$\frac{4}{5}$				
$x_1=-\frac{1}{2}$	2	$\frac{12}{5}$			
$x_2=0$	4	4	$\frac{8}{5}$		
$x_3=\frac{1}{2}$	2	-4	-8	$-\frac{32}{5}$	
$x_4=1$	$\frac{4}{5}$	$-\frac{12}{5}$	$\frac{8}{5}$	$\frac{32}{5}$	$\frac{32}{5}$

luego

$$p_4(x)=\frac{4}{5}+\frac{12}{5}(x+1)+\frac{8}{5}(x+1)\left(x+\frac{1}{2}\right)-\frac{32}{5}(x+1)\left(x+\frac{1}{2}\right)x+\frac{32}{5}(x+1)\left(x+\frac{1}{2}\right)x\left(x-\frac{1}{2}\right).$$

Ahora, vamos a aproximar a la función f por medio del polinomio interpolante $\widetilde{p}$ de grado menor o igual a 4 empleando los nodos del polinomio de Chebyshev T_5 (que teóricamente, deben presentar una menor cota de error). Los ceros de T_5 están dados por

$$t_k = \cos\left(\frac{2k+1}{2(5)}\,\pi\right) = \cos\left(\frac{2k+1}{10}\,\pi\right) \qquad k = 0, 1, 2, 3, 4$$

así

- $t_0 = \cos\left(\frac{1}{10}\pi\right) \approx 0.9511 \quad \to \quad f(t_0) \approx 0.8662$
- $t_1 = \cos\left(\frac{3}{10}\pi\right) \approx 0.5878 \quad \to \quad f(t_1) \approx 1.6793$
- $t_2 = \cos\left(\frac{5}{10}\pi\right) \approx 0 \quad \to \quad f(t_2) \approx 0$
- $t_3 = \cos\left(\frac{7}{10}\pi\right) \approx -0.5878 \quad \to \quad f(t_3) \approx 1.6793$
- $t_4 = \cos\left(\frac{9}{10}\pi\right) \approx -0.9511 \quad \to \quad f(t_4) \approx 0.8662$

empleando el polinomio interpolante de Newton

$$\widetilde{p}_4(x) = \widetilde{a_0} + \widetilde{a_1}(x-t_0) + \widetilde{a_2}(x-t_0)(x-t_1) + \widetilde{a_3}(x-t_0)(x-t_1)(x-t_2) + \widetilde{a_4}(x-t_0)(x-t_1)(x-t_2)(x-t_3),$$

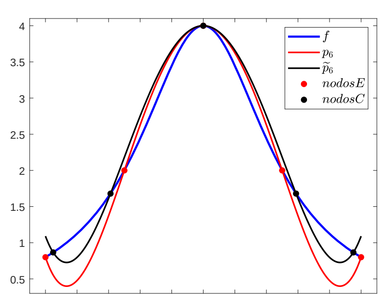
la tabla de diferencias divididas es

t_i	D.D. cero	1$^{\text{ras}}$ D.D.	2$^{\text{das}}$ D.D.	3$^{\text{ras}}$ D.D.	4$^{\text{ras}}$ D.D.
$t_0 \approx 0.9511$	0.8662				
$t_1 \approx 0.5878$	1.6793	-2.2383			
$t_2 = 0$	4	-3.9482	1.7979		
$t_3 \approx 1.6793$	1.6793	3.9482	-6.7171	5.5334	
$t_4 \approx -0.9511$	0.8662	2.2383	1.7979	-5.5334	5.8182

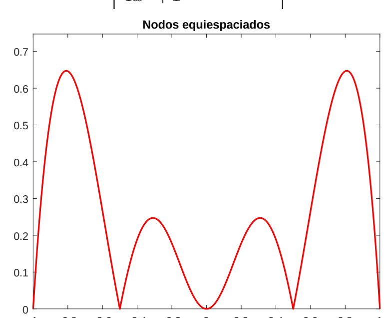
luego

$$\begin{aligned}\widetilde{p}_4(x) \approx& 0.8662 - 2.2383(x-0.9511) + 1.7979(x-0.9511)(x-0.5878)\\ &+ 5.5334(x-0.9511)(x-0.5878)x + 5.8182(x-0.9511)(x-0.5878)x(x+0.5878).\end{aligned}$$

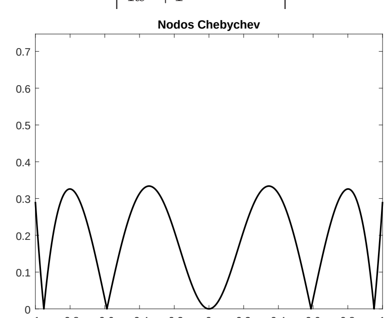
Ambos polinomios, p_4 y $\widetilde{p}_4$ interpolan a f en el intervalo $[-1, 1]$, veamos como se comportan los errores



$\left|\frac{4}{4x^2+1} - p_4(x)\right|$



$\left|\frac{4}{4x^2+1} - \widetilde{p}_4(x)\right|$



en este caso, el error que se comete con el polinomio con nodos equiespaciados es de 0.6472 y el error cometido con el polinomio con nodos de Chebyshev es de 0.3338, podemos ver que los errores en el intervalo dado por el polinomio con nodos de Chebyshev son más bajos.

Teorema. *Sea p_n el polinomio que interpola a f en los ceros del polinomio de Chebyshev T_{n+1} dados por $t_k = \cos\left(\dfrac{2k+1}{2(n+1)}\,\pi\right)$, $k = 0, \ldots, n$. Si $f \in C^{n+1}[-1, 1]$ entonces*

$$\max_{-1 \le t \le 1} |f(t) - p_n(t)| \le \frac{1}{2^n (n+1)!} \max_{-1 \le t \le 1} |f^{(n+1)}(t)|.$$

★ Consideremos ahora el polinomio que interpola a una función f definida en $[a,b]$ construido con los nodos $a \leq x_0 < x_1 < ... < x_n \leq b$, tal que $f(x) = p_n(x) + E(x)$ donde

$$E(x) = \frac{f^{(n+1)}(\widetilde{x})}{(n+1)!} \underbrace{(x-x_0)(x-x_1)...(x-x_n)}_{Q(x)}.$$

La cota de error esta dada por

$$|E(x)| \leq \frac{1}{(n+1)!} \max_{a \leq x \leq b} |f^{(n+1)}(x)| \max_{a \leq x \leq b} |Q(x)|.$$

la elección del conjunto de nodos $\{x_k\}_{k=0}^{n}$ que haga mínimo el valor $\max\limits_{a \leq t \leq b} |Q(x)|$ la obtenemos realizando un cambio de variable que me permita trasladar los nodos de Chebyshev al intervalo $[a,b]$, esto es, tomando

$$x_k = \frac{b-a}{2} t_k + \frac{a+b}{2} \qquad k = 0, \ldots, n, \qquad \textbf{\textit{Nodos de Chebyshev en }} [a,b]$$

donde $t_k = \cos\left(\frac{2k+1}{2(n+1)}\,\pi\right)$, $k = 0, \ldots, n$.

Ejemplo Aproximar la función $f(x) = \tan^{-1}(x) + \ln(x+\pi)$ en el intervalo $\left[-\frac{\pi}{4}, \frac{\pi}{3}\right]$ por medio de un polinomio de grado menor o igual a 4 que minimice la cota de error.

Solución: Para hallar p_4 empleamos los nodos de Chebyshev en el intervalo $\left[-\frac{\pi}{4}, \frac{\pi}{3}\right]$ dados por

$$x_k = \frac{\frac{\pi}{3} - \left(-\frac{\pi}{4}\right)}{2} t_k + \frac{-\frac{\pi}{4} + \frac{\pi}{3}}{2} = \frac{7\pi}{24} t_k + \frac{\pi}{24} \qquad k = 0,1,2,3,4$$

con

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

así

- $t_0 \approx 0.9511 \quad \to \quad x_0 \approx 1.0024$,
- $t_1 \approx 0.5878 \quad \to \quad x_1 \approx 0.6695$,
- $t_2 = 0 \quad \to \quad x_2 \approx 0.1309$,
- $t_3 \approx -0.5878 \quad \to \quad x_3 \approx -0.4077$,
- $t_4 \approx -0.9511 \quad \to \quad x_4 \approx -0.7406$

el polinomio interpolante de Newton tiene la forma

$$p_4(x) = a_0 + a_1(x-x_0) + a_2(x-x_0)(x-x_1) + a_3(x-x_0)(x-x_1)(x-x_2) + a_4(x-x_0)(x-x_1)(x-x_2)(x-x_3)$$

y la tabla de diferencias es

x_i	D.D. cero	1ras D.D.	2das D.D.	3ras D.D.	4ras D.D.
1.0024	−0.0081				
0.6695	0.9625	−2.9159			
0.1309	1.8449	−1.6384	−1.4661		
−0.4077	1.5186	0.6059	−2.0834	0.4378	
−0.7406	0.7796	2.2203	−1.8525	−0.1638	0.3452

luego el polinomio de grado 4 que aproxima a f en $\left[-\frac{\pi}{4}, \frac{\pi}{3}\right]$ es

$$\begin{aligned} p_4(x) \approx\ & -0.0081 - 2.9159(x-1.0024) - 1.4661(x-1.0024)(x-0.6695) \\ & + 0.4378(x-1.0024)(x-0.6695)(x-0.1309) + 0.3452(x-1.0024)(x-0.6695)(x-0.1309)(x+0.4077). \end{aligned}$$

Teorema. *Sea p_n el polinomio que interpola a f en los ceros trasladados del polinomio de Chebyshev T_{n+1} dados por*

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

Si $f \in C^{n+1}[a,b]$ entonces

$$\max_{a\le x\le b}|f(x)-p_n(x)| \le \frac{2(b-a)^{n+1}}{4^{n+1}(n+1)!}\max_{a\le x\le b}|f^{(n+1)}(x)|.$$

Ejemplo Hallar la cota de error producida al aproximar la función $f(x) = 5\cos\left(\frac{x\pi}{5}\right)$ en $[-1,2]$ por medio del polinomio interpolante de grado menor o igual a 3 empleando los nodos de Chebyshev.

<u>*Solución:*</u> La cota de error esta dada por

$$|E(x)| = |f(x)-p_3(x)| \le \frac{2(2-(-1))^{3+1}}{4^{3+1}(3+1)!}\text{máx}\left\{|f^{(3+1)}(x)| : -1\le x\le 2\right\},$$

primero calculamos las derivadas

$$f'(x) = -\pi\,\text{sen}\left(\frac{\pi x}{5}\right), \quad f''(x) = -\frac{1}{5}\pi^2\cos\left(\frac{\pi x}{5}\right), \quad f'''(x) = \frac{1}{25}\pi^3\,\text{sen}\left(\frac{\pi x}{5}\right), \quad f^{(4)}(x) = \frac{1}{125}\pi^4\cos\left(\frac{\pi x}{5}\right),$$

así

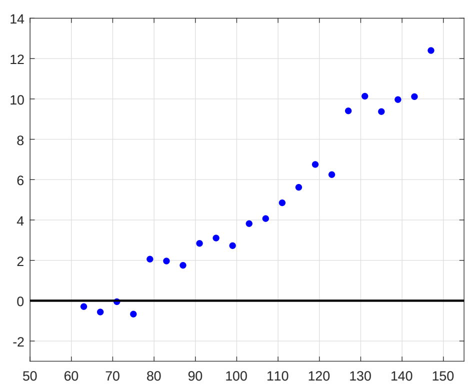
$$|E(x)| = |f(x)-p_3(x)| \le \frac{27}{1024}\text{máx}\left\{\left|\frac{1}{125}\pi^4\cos\left(\frac{\pi x}{5}\right)\right| : -1\le x\le 2\right\}$$

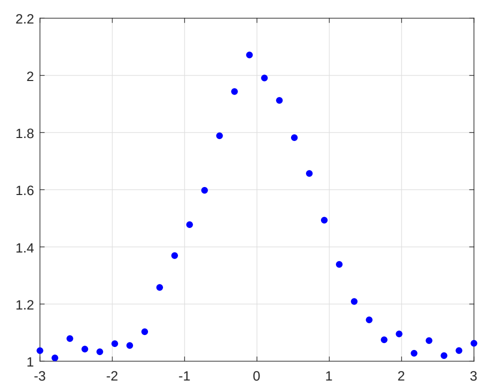
sabemos que $\left|\cos\left(\frac{\pi x}{5}\right)\right| \le 1$, por lo tanto la cota del error es

$$|E(x)| \le \frac{27}{1024}\,\frac{1}{125}\pi^4 \approx 0.020547,$$

esto significa que el mayor error que cometemos al aproximar a f por medio del polinomio interpolante de grado menor o igual a 3 empleando los nodos de Chebyshev en $[-1,2]$ es de 0.020547.

Consideremos ahora que tenemos las siguientes nubes de puntos





¿En este caso si será aconsejable hallar el polinomio interpolante? Notemos que los puntos tienen comportamiento lineal (a izquierda) y de campana de Gauss (a derecha). La respuesta a esta pregunta es no, en estos casos se aconseja buscar una curva de aproximación.

APROXIMACIÓN DISCRETA POR MÍNIMOS CUADRADOS

Consideremos el problema de hallar la ecuación de la curva $y = f(x)$ que mejor se aproxima al conjunto de m puntos $(x_1, y_1), \ldots, (x_m, y_m)$ con $x_i \neq x_j$. Para ello, primero consideremos los errores llamados *desviaciones* o *residuos*

$$e_k = f(x_k) - y_k, \qquad k = 1, \ldots, m.$$

Ahora, hay varias medidas de los errores, entre otras:

- Error máximo: $E_\infty(f) = \max\limits_{1 \le k \le m} |f(x_k) - y_k|$.
- Error medio: $E_1(f) = \dfrac{1}{m} \sum\limits_{k=1}^{m} |f(x_k) - y_k|$.
- Error cuadrático medio:

$$E_2(f) = \left(\frac{1}{m} \sum_{k=1}^{m} (f(x_k) - y_k)^2 \right)^{1/2}.$$

Para hallar la ecuación de la curva $y = f(x)$ que mejor se ajusta a los puntos (no tiene que interpolarlos), buscamos aquella que minimice el error. Recordemos que para minimizar la función de error buscamos sus números críticos, pero, los errores $E_\infty(f)$ y $E_1(f)$ no son diferenciables y minimizar $E_2(f)$ es equivalente a minimizar

$$E(f) = \sum_{k=1}^{m} (f(x_k) - y_k)^2 \qquad \textbf{\textit{Error cuadrático}}$$

así que, buscaremos la ecuación de la curva $y = f(x)$ que minimiza $E(f)$. Primero, consideremos el caso cuando la curva es una recta.

⋆ Recta de regresión

Dados m puntos $(x_1, y_1), \ldots, (x_m, y_m)$ con $x_i \neq x_j$, queremos hallar la recta $y = f(x) = Ax + B$ que mejor se ajusta a los puntos, es decir, buscamos la recta que minimiza el error

$$E(A, B) := \sum_{k=1}^{m} (A\,x_k + B - y_k)^2.$$

Para minizar la función error $E(A, B)$ respecto a las A y B, buscamos cuando $\dfrac{\partial E(A,B)}{\partial A} = 0$ y $\dfrac{\partial E(A,B)}{\partial B} = 0$

$$0 = \frac{\partial E(A,B)}{\partial A} = \sum_{k=1}^{m} 2\,x_k\,(A\,x_k + B - y_k) = 2\left(A \sum_{k=1}^{m} x_k^2 + B \sum_{k=1}^{m} x_k - \sum_{k=1}^{m} x_k\,y_k \right)$$
$$0 = \frac{\partial E(A,B)}{\partial B} = \sum_{k=1}^{m} 2(A\,x_k + B - y_k) = 2\left(A \sum_{k=1}^{m} x_k + B \sum_{k=1}^{m} 1 - \sum_{k=1}^{m} y_k \right)$$

obtenemos así el sistema de ***ecuaciones normales***

$$A \sum_{k=1}^{m} x_k^2 + B \sum_{k=1}^{m} x_k = \sum_{k=1}^{m} x_k y_k$$

$$A \sum_{k=1}^{m} x_k \quad + \quad Bm \ = \sum_{k=1}^{m} y_k$$

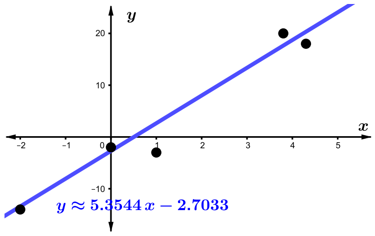
y se demuestra que la solución de este sistema minimiza la función de error $E(A, B)$. Esta recta recibe el nombre de ***recta de regresión*** o ***recta óptima en el sentido de los míminos cuadrados***.

Ejemplo Calcule la recta de regresión para los datos

x_k	-2	0	1	3.8	4.3
y_k	-14	-2	2.5	20	18

¿Cuál es el error de mínimos cuadrados en este caso?

Solución: Debemos resolver el sistema de ecuaciones normales, para ello, consideremos la siguiente tabla



x_k	-2	0	1	3.8	4.3
y_k	-14	-2	2.5	20	18
x_k^2	4	0	1	14.44	18.49
$x_k\, y_k$	28	0	2.5	76	77.4

de aquí $\displaystyle\sum_{k=1}^{5} x_k^2 = 37.93$, $\displaystyle\sum_{k=1}^{5} x_k = 7.1$, $\displaystyle\sum_{k=1}^{5} y_k = 19$ y $\displaystyle\sum_{k=1}^{5} x_k\, y_k = 183.9$, luego el sistema de ecuaciones normales a resolver es

$$37.93\,A + 7.1\,B = 183.9$$
$$7.1\,A + 5\,B = 19$$

cuya solución es $A \approx 5.3544$ y $B \approx -2.7033$, la recta de regresión para los datos es $y \approx 5.3544\,x - 2.7033$. El error de mínimos cuadrados es

$$\sum_{k=1}^{5} (5.3544\,x_k - 2.7033 - y_k)^2 \approx 11.8018.$$

Ejemplo Determine la recta de regresión para los datos $\{(x_k, y_k)\}_{k=1}^{10}$, siendo $x_k = (0.1)k$ e $y_k = x_k + \cos(\sqrt{k})$.

Solución: Primero construimos los puntos

x_k	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
y_k	0.6403	0.3559	0.1394	-0.0161	-0.1173	-0.1699	-0.1796	-0.1514	-0.090	0.0002

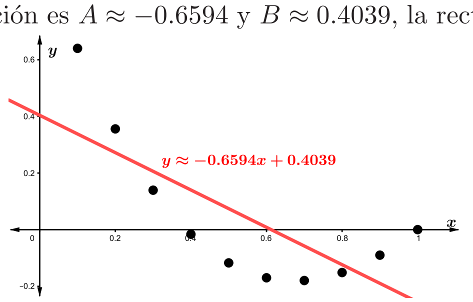
donde $y_1 = 0.1 + \cos(\sqrt{1})$, $y_2 = 0.2 + \cos(\sqrt{2})$ y así hasta $y_{10} = 1 + \cos(\sqrt{10})$. Para plantear el sistema a resolver consideremos la siguiente tabla más completa

x_k	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
y_k	0.6403	0.3559	0.1394	-0.0161	-0.1173	-0.1699	-0.1796	-0.1514	-0.090	0.0002
x_k^2	0.01	0.04	0.09	0.16	0.25	0.36	0.49	0.64	0.81	1
$x_k y_k$	0.0640	0.0712	0.0418	-0.0065	-0.0586	-0.1019	-0.1257	-0.1211	-0.0810	0.0002

de aquí $\displaystyle\sum_{k=1}^{10} x_k^2 \approx 3.85$, $\displaystyle\sum_{k=1}^{10} x_k \approx 5.5$, $\displaystyle\sum_{k=1}^{10} y_k \approx 0.4117$ y $\displaystyle\sum_{k=1}^{10} x_k y_k \approx -0.3176$, luego el sistema a resolver es

$$3.85A + 5.5B = -0.3176$$
$$5.5A + 10B = 0.4117$$

cuya solución es $A \approx -0.6594$ y $B \approx 0.4039$, la recta de regresión para los datos es $y \approx -0.6594x + 0.4039$.



Error de mínimos cuadrados

$$\sum_{k=1}^{10} (-0.6594\,x_k + 0.4039 - y_k)^2 \approx 0.286655.$$

Como podemos observar en el ejemplo anterior la recta no siempre es la curva que mejor se ajusta a los puntos, los puntos podrían tener un comportamiento parabólico o cubico.

$\star$ Ajuste polinomial

¿Comó calcular un polinomio óptimo en el sentido de los mínimos cuadrados? Consideremos un caso particular y así poder generalizar.

Hallemos el polinomio de grado 3 que mejor se ajusta a m puntos $(x_1, y_1), \ldots, (x_m, y_m)$ con $x_i \neq x_j$, esto es, la curva $y = f(x) = Ax^3 + Bx^2 + Cx + D$ debe cumplir que

$$E(A, B, C, D) := \sum_{k=1}^{m} (Ax_k^3 + Bx_k^2 + Cx_k + D - y_k)^2 \qquad \text{sea mínimo.}$$

El mínimo valor de E lo obtenemos igualando a cero las derivadas parciales

$$0 = \frac{\partial E(A, B, C, D)}{\partial A} = 2\sum_{k=1}^{m} (Ax_k^3 + Bx_k^2 + Cx_k + D - y_k)x_k^3 = 2\sum_{k=1}^{m} (Ax_k^6 + Bx_k^5 + Cx_k^4 + x_k^3 D - x_k^3 y_k)$$

procediendo de manera similar para las igualdades $\frac{\partial E(A,B,C,D)}{\partial B} = 0$, $\frac{\partial E(A,B,C,D)}{\partial C} = 0$ y $\frac{\partial E(A,B,C,D)}{\partial D} = 0$, obtenemos el sistema de *ecuaciones normales*

$$\begin{aligned}
A\sum_{k=1}^{m} x_k^6 + B\sum_{k=1}^{m} x_k^5 + C\sum_{k=1}^{m} x_k^4 + D\sum_{k=1}^{m} x_k^3 &= \sum_{k=1}^{m} x_k^3 y_k \\
A\sum_{k=1}^{m} x_k^5 + B\sum_{k=1}^{m} x_k^4 + C\sum_{k=1}^{m} x_k^3 + D\sum_{k=1}^{m} x_k^2 &= \sum_{k=1}^{m} x_k^2 y_k \\
A\sum_{k=1}^{m} x_k^4 + B\sum_{k=1}^{m} x_k^3 + C\sum_{k=1}^{m} x_k^2 + D\sum_{k=1}^{m} x_k &= \sum_{k=1}^{m} x_k y_k \\
A\sum_{k=1}^{m} x_k^3 \ + \ B\sum_{k=1}^{m} x_k^2 \ + \ C\sum_{k=1}^{m} x_k \ + \ Dm \ &= \sum_{k=1}^{m} y_k
\end{aligned}$$

El ***polinomio cúbico óptimo*** en el sentido de los mínimos cuadrados esta dado por $y = f(x) = Ax^3 + Bx^2 + Cx + D$ donde A, B, C y D son la solución del sistema de ecuaciones normales anterior.

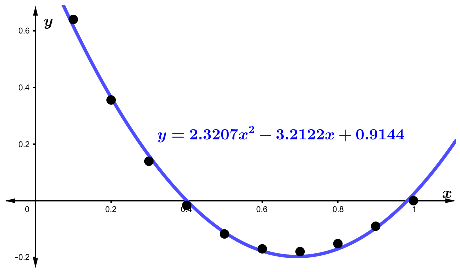
En general, el polinomio de grado menor o igual a n óptimo en el sentido de los mínimos cuadrados esta dado por $y = f(x) = a_n\, x^n + \ldots + a_1\, x + a_0$ donde $a_i,\ i = 0, \ldots, n$ son la solución del sistema de *ecuaciones normales*

$$\begin{bmatrix}
\sum_{k=1}^{m} x_k^{2n} & \sum_{k=1}^{m} x_k^{2n-1} & \ldots & \sum_{k=1}^{m} x_k^{n+1} & \sum_{k=1}^{m} x_k^{n} \\
\sum_{k=1}^{m} x_k^{2n-1} & \sum_{k=1}^{m} x_k^{2n-2} & \ldots & \sum_{k=1}^{m} x_k^{n} & \sum_{k=1}^{m} x_k^{n-1} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\sum_{k=1}^{m} x_k^{n+1} & \sum_{k=1}^{m} x_k^{n} & \ldots & \sum_{k=1}^{m} x_k^{2} & \sum_{k=1}^{m} x_k \\
\sum_{k=1}^{m} x_k^{n} & \sum_{k=1}^{m} x_k^{n-1} & \ldots & \sum_{k=1}^{m} x_k & m
\end{bmatrix}
\begin{bmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_1 \\ a_0 \end{bmatrix}
=
\begin{bmatrix} \sum_{k=1}^{m} x_k^n y_k \\ \sum_{k=1}^{m} x_k^{n-1} y_k \\ \vdots \\ \sum_{k=1}^{m} x_k\, y_k \\ \sum_{k=1}^{m} y_k \end{bmatrix}$$

Ejemplo Hallar el polinomio de grado 2 que mejor se ajusta a los datos $\{(x_k, y_k)\}_{k=1}^{10}$, siendo $x_k = (0.1)k$ e $y_k = x_k + \cos(\sqrt{k})$.

<u>*Solución:*</u> Ya vimos en el ejemplo anterior el comportamiento de los puntos y la recta de regresión no es la que mejor se ajusta a ellos. Así el polinomio de grado 2 que mejor se ajusta a los datos en el sentido de mínimos cuadrados tiene la forma $y = p_2(x) = a_2\, x^2 + a_1\, x + a_0$ donde a_2, a_1 y a_0 es la solución del sistema de ecuaciones normales

$$\left.\begin{aligned}
a_2 \textstyle\sum_{k=1}^{m} x_k^4 \quad + a_1 \sum_{k=1}^{m} x_k^3 \quad + a_0 \sum_{k=1}^{m} x_k^2 \quad &= \textstyle\sum_{k=1}^{m} x_k^2 y_k \\
a_2 \textstyle\sum_{k=1}^{m} x_k^3 \quad + a_1 \sum_{k=1}^{m} x_k^2 \quad + a_0 \sum_{k=1}^{m} x_k \quad &= \textstyle\sum_{k=1}^{m} x_k\, y_k \\
a_2 \textstyle\sum_{k=1}^{m} x_k^2 \quad + a_1 \sum_{k=1}^{m} x_k \quad + a_0\, m \quad &= \textstyle\sum_{k=1}^{m} y_k
\end{aligned}\right\} \quad \Rightarrow \quad
\begin{bmatrix} 2.5333 & 3.025 & 3.85 \\ 3.025 & 3.85 & 5.5 \\ 3.85 & 5.5 & 10 \end{bmatrix}
\begin{bmatrix} a_2 \\ a_1 \\ a_0 \end{bmatrix}
=
\begin{bmatrix} -0.3174 \\ -0.3176 \\ 0.4117 \end{bmatrix}$$



cuya solución es $a_2 \approx 2.3207$, $a_1 \approx -3.2122$ y $a_0 \approx 0.9144$, el polinomio de grado 2 que mejor se ajusta a los datos en el sentido de mínimos cuadrados es $y = p_2(x) \approx 2.3207\,x^2 - 3.2122\,x + 0.9144$ y el error de mínimos cuadrados es

$$\sum_{k=1}^{10}(2.3207\,x_k^2 - 3.2122\,x_k + 0.9144 - y_k)^2 \approx 0.0027.$$

También existe ajuste potencial, exponencial o diferentes curvas de aproximación. Si por ejemplo, buscamos una curva de aproximación de la forma $y = be^{ax}$ que mejor se ajuste a los datos $\{(x_k, y_k)\}_{k=1}^{m}$, el sistema no lineal de ecuaciones normales a resolver será

$$\sum_{k=1}^{m}(be^{ax_k} - y_k)e^{ax_k} = 0$$
$$\sum_{k=1}^{m}(be^{ax_k} - y_k)bx_k e^{ax_k} = 0$$

sistema para el que no siempre es posible encontrar una solución exacta. En estos casos, se propone realizar una linealización del problema y hallar la recta de regresión. En este caso, la linealización (siempre que $y_k > 0$) será $\ln y = \ln b + ax$. Es importante mencionar que la aproximación que se obtiene no es la de mínimos cuadrados del problema original.

Ejemplo Cuando una población $P(t)$ no puede crecer más allá de un cierto valor límite L, la gráfica de la función $P(t)$ es una curva, llamada curva logística, de ecuación $y = \dfrac{L}{1 + Ce^{At}}$. Calcule A y C para los siguientes datos, siendo $L = 5000$

t	0	1	2	3	4
y	500	1000	1800	2800	3700

Solución: Buscamos introducir un cambio de variables que me permita linealizar la ecuación, la resolvemos por medio de la recta de regresión y luego volvemos a las variables originales.

$$y = \frac{L}{1 + Ce^{At}} \quad \Leftrightarrow \quad 1 + Ce^{At} = \frac{5000}{y} \quad \Leftrightarrow \quad Ce^{At} = \frac{5000}{y} - 1 \quad \Leftrightarrow \quad \ln(Ce^{At}) = \ln\left(\frac{5000}{y} - 1\right)$$
$$\ln(C) + At = \ln\left(\frac{5000}{y} - 1\right)$$

si llamamos $Y = \ln\left(\frac{5000}{y} - 1\right)$, $X = t$ y $B = \ln(C)$, nuestro problema es hallar A y B tales que $Y = AX + B$ sea la recta de regresión para los datos

X	0	1	2	3	4
Y	$\ln(\frac{5000}{500} - 1)$	$\ln(\frac{5000}{1000} - 1)$	$\ln(\frac{5000}{1800} - 1)$	$\ln(\frac{5000}{2800} - 1)$	$\ln(\frac{5000}{3700} - 1)$

completando la tabla

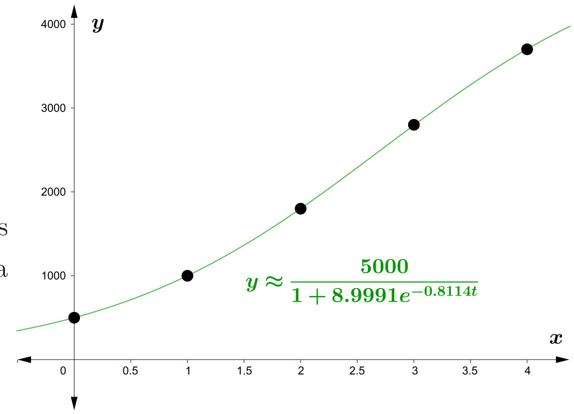
X_k	0	1	2	3	4
Y_k	2.1972	1.3863	0.5754	-0.2412	-1.0460
X_k^2	0	1	4	9	16
X_kY_k	0	1.3863	1.1507	-0.7235	-4.1839

el sistema de ecuaciones normales es

$$30A + 10B = -2.3703$$
$$10A + 5B = 2.8718$$

y su solución es $A \approx -0.8114$ y $B \approx 2.1971$, representa a la variables originales, $C = e^B \approx e^{2.1971} \approx 8.9991$. La curva que mejor se ajusta a los datos es

$$y \approx \frac{5000}{1 + 8.9991e^{-0.8114t}}$$



y el error de mínimos cuadrados es

$$\sum_{k=1}^{10} \left(\frac{5000}{1 + 8.9991e^{-0.8114t_k}} - y_k \right)^2 \approx 32.8142.$$

Otros cambios de variables que permiten linealizar curvas

Función $y = f(x)$	**Función** $Y = F(X)$	**Cambio**
$y = Ce^{Ax}$	$Y = a_0 + a_1 X$	$X = x,\ Y = \ln y,\ C = e^{a_0},\ A = a_1$
$y = Cx^A$	$Y = a_0 + a_1 X$	$X = \ln x,\ Y = \ln y,\ C = e^{a_0},\ A = a_1$
$y = \dfrac{1}{Cx + A}$	$Y = a_0 + a_1 X$	$X = x,\ Y = \dfrac{1}{y},\ C = a_1,\ A = a_0$
$y = \dfrac{A}{x} + C$	$Y = a_0 + a_1 X$	$X = \dfrac{1}{x},\ Y = y,\ C = a_0,\ A = a_1$