Solución Examen - Métodos Numéricos

5 de Febrero de 2016

Problema 1 (30 puntos)

a) El Problema de Mínimos Cuadrados Lineal (PMCL) consiste en hallar el vector de coeficientes x de un modelo lineal $y = Ax$, que minimiza la distancia Euclideana entre dicho modelo y m observaciones dadas por el vector b. Es decir, consiste en hallar $x = argmin\ ||b - Ax||_2$, con $A \in R^{m\times n}$, $b \in R^m$ y $x \in R^n$. En general se tienen más observaciones que incógnitas $m > n$.

b)($\Leftarrow$) Sea $x \in \mathbb{R}^n$ solución de las ecuaciones normales. Veamos que x alcanza la norma mínima en el PMCL. Tomemos $y \in \mathbb{R}^n$ arbitrario. Como x es solución de las ecuaciones normales, tenemos que $A^t(Ax - b) = 0$. Luego:

$$\begin{aligned}
\|b - Ay\|_2^2 &= \|b - Ax + A(x-y)\|_2^2 \\
&= \|b - Ax\|_2^2 + \|A(x-y)\|_2^2 + (A(x-y))^t(b - Ax) + (b - Ax)^t(A(x-y)) \\
&= \|b - Ax\|_2^2 + \|A(x-y)\|_2^2
\end{aligned}$$

donde se ha utilizado en el último paso que:

$$(A(x-y))^t(b - Ax) = (x-y)^t[A^t(b - Ax)] = 0$$

y

$$(b - Ax)^t(A(x-y)) = (x-y)^t[A^t(b - Ax)] = 0$$

Luego, $\|b - Ay\|_2^2 \geq \|b - Ax\|_2^2$ para cualquier $y \in \mathbb{R}^n$, mostrando así que x alcanza la norma mínima en el PMCL, y es por tanto solución del PMCL.

($\Rightarrow$) Supongamos ahora que x alcanza la norma mínima en el PMCL, y veamos que entonces debe verificar las ecuaciones normales. Supongamos por absurdo que x no es solución de las ecuaciones normales. Entonces $z = A^t(Ax - b)$ es un vector no nulo. Elijamos y tal que $x - y = -hz$ para h pequeño positivo. Entonces:

$$\begin{aligned}
\|b - Ay\|_2^2 &= \|b - Ax\|_2^2 + h^2\|Az\|_2^2 - 2h(Az)^t(b - Ax) \\
&= \|b - Ax\|_2^2 + h^2\|Az\|_2^2 - 2h\|z\|_2^2 < \|b - Ax\|_2^2,
\end{aligned}$$

si se elige h suficientemente pequeño, donde se ha utilizado que $(Az)^t(b - Ax) = z^t[A^t(b - Ax)] = z^t z = \|z\|_2^2$. Luego, x no minimiza $\|b - Ay\|_2^2$, y por tanto no es solución del PMCL, lo cual es absurdo.

c) En este caso las observaciones son $b = (1, 2, 4)^T \in R^3$ y el modelo propuesto es $y = \alpha t + \beta$, de parámetros $x = (\alpha, \beta)^T \in R^2$. El residuo a minimizar se puede escribir como:

$$\begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

Por lo tanto $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix}$. Las ecuaciones normales son entonces:

$$A^T Ax = A^T b \Leftrightarrow \begin{pmatrix} 5 & 3 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 10 \\ 7 \end{pmatrix}$$

Resolviendo este sistema lineal se obtienen los parámetros $(\alpha, \beta) = (\frac{3}{2}, \frac{5}{6})$.

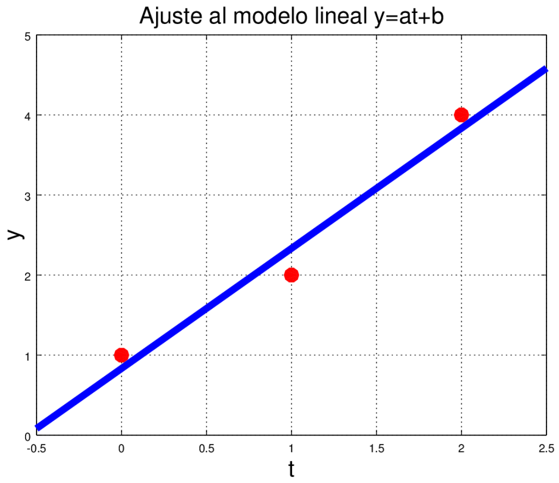

Figura 1. Ajuste mediante Mínimos Cuadrados Lineal.

Problema 2 (35 puntos)

a) En muchas aplicaciones interesa considerar polinomios $P(x)$ que no solo interpolan a $f(x)$, sino que además interpolan a $f'(x)$, es decir:

$$\begin{cases} P(x_i) = y_i = f(x_i) \\ P'(x_i) = y'_i = f'(x_i) \quad \forall i = 0, 1, \dots, n \end{cases}$$

Por lo tanto, se tienen $2n+2$ condiciones a imponer, por lo que se busca un polinomio $P(x)$ de al menos grado $2n+1$, es decir tenemos $2n+2$ coeficientes a hallar. Este polinomio puede escribirse como:

$$H_{2n+1}(x) = \sum_{i=0}^{i=n} (y_i h_i(x) + y'_i \tilde{h}_i(x))$$

Con:

$$\begin{array}{ll} h_i(x_j) = \delta_{ij} & h'_i(x_j) = 0 \\ \tilde{h}_i(x_j) = 0 & \tilde{h}'_i(x_j) = \delta_{ij} \quad 0 \leqslant i, j \leqslant n \end{array}$$

Así, los $h_i(x)$ tienen la forma:

$$h_i(x) = [1 - 2l'_i(x_i)(x - x_i)](l_i(x))^2$$

$$\tilde{h}_i(x) = (x - x_i)(l_i(x))^2$$

Con $l_i(x)$ los polinomios base de Lagrange.

b) La tabla de datos es la siguiente:

x_i	y_i	y'_i
0	0	1
$\pi/4$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$

Los polinomios base de Lagrange son:

$$l_0(x) = \left(1 - \frac{\pi}{4} x\right)$$

$$l_1(x) = \frac{4}{\pi} x$$

Si recordamos la expresión hallada en la parte anterior para el polinomio interpolante, se tiene en este caso:

$$H_3(x) = 0.h_0(x) + 1.\tilde{h}_0(x) + \frac{\sqrt{2}}{2}.h_1(x) + \frac{\sqrt{2}}{2}.\tilde{h}_1(x)$$

Ahora:

$$\tilde{h}_0(x) = x\left(1 - \frac{4}{\pi} x\right)^2$$

$$h_1(x) = \left[1 - \frac{8}{\pi}\left(x - \frac{\pi}{4}\right)\right]\left(\frac{4}{\pi} x\right)^2$$

$$\tilde{h}_1(x) = \left(x - \frac{\pi}{4}\right)\left(\frac{4}{\pi} x\right)^2$$

Por lo tanto el polinomio interpolante es:

$$H_3(x) = x\left(1-\frac{4}{\pi}x\right)^2 + \frac{\sqrt{2}}{2}\left[1-\frac{8}{\pi}\left(x-\frac{\pi}{4}\right)\right]\left(\frac{4}{\pi}x\right)^2 + \frac{\sqrt{2}}{2}\left(x-\frac{\pi}{4}\right)\left(\frac{4}{\pi}x\right)^2$$

$H_3(x)$ también se puede obtener utilizando el esquema de diferencias divididas. En ese caso:

0	0			
0	0	1		
$\pi/4$	$\frac{\sqrt{2}}{2}$	$\frac{2\sqrt{2}}{\pi}$	$\frac{4(2\sqrt{2}-\pi)}{\pi^2} \approx -0{,}1269$	
$\pi/4$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{2\sqrt{2}(\pi-4)}{\pi^2}$	$\frac{\pi(2\sqrt{2}+4)-16\sqrt{2}}{\pi^2}\frac{4}{\pi} \approx -0{,}1516$

Por lo tanto el polinomio interpolante de Hermite de la función $f(x) = sen(x)$ por las abscisas $x_0 = 0$ y $x_1 = \pi/4$ es:

$$H_3(x) = x + \frac{4\left(2\sqrt{2}-\pi\right)}{\pi^2}x^2 + \frac{\pi\left(2\sqrt{2}+4\right)-16\sqrt{2}}{\pi^2}\frac{4}{\pi}x^2(x-\pi/4)$$

$$H_3(x) = x - 0{,}1269x^2 - 0{,}1516x^2(x-\pi/4)$$

c) Se puede demostrar que el error tiene la forma:

$$f(x) - H_{2n+1}(x) = \frac{f^{2n+2}(\xi(x))}{(2n+2)!}(x-x_0)^2(x-x_1)^2\dots(x-x_n)^2,\ \xi(x) \in [0, \frac{\pi}{4}]$$

Entonces, una cota posible para el error es $E_{max} \leq \frac{\|f^{(4)}\|_\infty}{4!}(\pi/4)^2(\pi/4)^2$. En nuestro caso $f^{(4)}(x) = sen(x)$, por lo tanto $\max_{[0,\pi/4]}|f^{(4)}(x)| = \max_{[0,\pi/4]} sen(x) = sen(\pi/4) = \frac{\sqrt{2}}{2}$. Luego, una cota superior para el error (que no es rígida) es:

$$c = \frac{\sqrt{2}}{2}\frac{1}{4!}\left(\frac{\pi}{4}\right)^4 \simeq 1{,}121x10^{-2}$$

Problema 3 (35 puntos)

a) Sea $e_n = x_n - \alpha$ el error en el paso n. Utilizando que α es punto fijo de la iteración tenemos que:

$$e_{n+1} = x_{n+1} - \alpha = Qx_n + r - (Q\alpha + r) = Q(x_n - \alpha) = Qe_n.$$

Luego, por inducción en los naturales tenemos que $e_n = Q^n e_0$. Vamos ahora a probar el directo y el recíproco en partes:

($\Rightarrow$) Supongamos por absurdo que $\rho(Q) \geq 1$. En tal caso, existe un valor propio λ de Q, con $|\lambda| \geq 1$, y un vector propio $v \neq 0$ tal que $Qv = \lambda v$. Elijamos x_0 de modo que $e_0 = v$. Esto es posible tomando $x_0 = \alpha + e_0$. Por lo anteriormente observado, tenemos que:

$$e_n = Q^n e_0 = Q^n v = \lambda^n v,$$

y tomando normas, tenemos que $\|e_n\| = |\lambda|^n \|v\|$. Tomando límites en ambos miembros, se consigue que $\lim_n \|e_n\| \neq 0$, pues $|\lambda| \geq 1$ y $\|v\| \neq 0$. Esto es decir que el error no tiende al vector nulo, o equivalentemente, que la sucesión $\{x_n\}_{n\in\mathbb{N}}$ no converge a α, en contradicción con la hipótesis.

($\Leftarrow$) Por el Teorema de Householder, el radio espectral es el ínfimo de las normas operadores. Sea ϵ igual a la mitad de la distancia entre $\rho(Q)$ y 1, es decir, $\epsilon = \frac{1-\rho(Q)}{2}$. Por definición de ínfimo, existe una norma operador $\|\|_\epsilon$ tal que $\|Q\|_\epsilon - \rho(Q) < \epsilon$. Pero entonces:

$$\|Q\|_\epsilon < \rho(Q) + \epsilon = \frac{2\rho(Q) + (1 - \rho(Q))}{2} = \frac{1 + \rho(Q)}{2} < 1.$$

Hemos conseguido así una norma $\|\|_\epsilon$ compatible con una vectorial tal que $\|Q\|_\epsilon < 1$. Como $e_n = Q^n e_0$, tomando normas en cada miembro tenemos que:

$$\|e_n\| = \|Q^n e_0\| \leq (\|Q\|_\epsilon)^n \|e_0\|,$$

y tomando límite con n tenemos que $0 \leq \lim_n \|e_n\| \leq (\|Q\|_\epsilon)^n \|e_0\| = 0$. La única opción válida es que $\lim_n \|e_n\| = 0$, y por la primera propiedad de una norma tenemos que la sucesión de vectores $\{e_n\}_{n\in\mathbb{N}}$ converge al vector nulo. Esto significa que $\{x_n\}_{n\in\mathbb{N}}$ converge a α. Obsérvese que el resultado vale independientemente del dato inicial x^0.

b) Si la matriz A es estrictamente diagonal dominante por filas (o por columnas), entonces tanto Jacobi como Gauss-Seidel son convergentes, para toda condición inicial x_0. Esto es una condición suficiente. Es decir que el método podría ser convergente aunque A no sea estrictamente diagonal dominante.

c) Una forma de asegurar la convergencia de ambos métodos es imponer condiciones para que A sea estrictamente diagonal dominante por filas:

- $3 > |-1| = 1$
- $|\beta| > 1$

Por lo tanto, si $\beta \in (-\infty, -1) \cup (1, \infty)$, tanto Jacobi como Gauss-Seidel son convergentes. Pueden existir otros valores de β para los cuales ambos métodos sean convergentes. Para determinarlos se puede hallar el radio espectral $\rho(Q)$ de Jacobi y Gauss-Seidel.

d) La matriz Q asociada al método de Jacobi es en este caso: $Q_J = \begin{pmatrix} 0 & \frac{1}{3} \\ 2 & 0 \end{pmatrix}$. Sus valores propios tienen módulo $|\lambda| = \sqrt{\frac{2}{3}} < 1$, por lo que Jacobi es convergente.