Computación Numérica

Primer Parcial A - Mayo 2016

Una máquina almacena números en punto flotante en 9 bits. El primer bit se usa para el signo del número, los cuatro siguientes para el exponente sesgado y los últimos cuatro bits para la magnitud de la mantisa. Si se sigue un criterio similar al de la norma IEEE 754:

1. Calcular el número $(100110110)_2$ en base 10.
2. ¿Cual sería el ϵ de máquina expresado en base 10?
3. ¿Cuál es el mayor entero que se puede almacenar de forma exacta?
4. ¿Cuales son el menor y el mayor real positivo que se almacena en forma normalizada? ¿Cómo se almacenarían en binario?
5. ¿Cuántos números normalizados se pueden representar con este sistema?

1.

signo	exponente	mantisa
1	0011	0110

EXPONENTE

Como el número de bits es $m = 4$, el sesgo es $2^{m-1} - 1 = 2^3 - 1 = 7$. El exponente es $E = e - sesgo = e - 7$. Como

$$e = (0011)_2 = 2^1 + 2^0 = 3,$$

tenemos $E = 3 - 7 = -4$.

MANTISA

Teniendo en cuenta el bit escondido

$$\mathbf{1}{,}0110 \times 2^{-4} = \left(1 + 2^{-2} + 2^{-3}\right) \times 2^{-4} = 0{,}0859375$$

Y teniendo en cuenta el signo el número es $-0{,}0859375$

2. El ϵ de máquina es la distancia que existe entre el número 1 y el siguiente número representable en ese sistema.
El ϵ de máquina es una cota superior del error relativo que cometemos al almacenar cualquier número con este sistema.
Como el número de dígitos de la mantisa es 10. El número 1 se representa

$$1 \quad = \quad 1{,}0000 \times 2^0$$

y el siguiente número representable es

$$1 + \epsilon \quad = \quad 1{,}0001 \times 2^0$$

Por lo que

$$\epsilon \quad = \quad \mathtt{0{,}0001} \times \mathtt{2^0} = \mathtt{2^{-4}}$$

Y en base 10 el ϵ de máquina es $2^{-4} = 0{,}0625$

3. El mayor entero que se puede almacenar de forma exacta sería el número siguiente a

$$\mathtt{11111} = \mathtt{1{,}1111} \times \mathtt{2^4}$$

es decir

$$\mathtt{100000} = \mathtt{1{,}0000} \times \mathtt{2^5}$$

porque el número siguiente habría que redondearlo

$$\mathtt{100001} \approx \mathtt{1{,}0000} \times \mathtt{2^5}$$

Por lo tanto, en base 10

$$\mathtt{2^5} = \mathtt{32}$$

4. Tenemos cuatro bits para los exponentes, es decir, $2^4 = 16$ valores distintos por lo que los exponentes posibles son

$$0, 1, 2, 3, \ldots, 13, 14, 15$$

Pero el primer valor y el último están reservados. Y si tenemos en cuenta el *sesgo* $= 7$, los valores que puede tomar son

$$R, -6, -5, -4, \ldots, 5, 6, 7, R$$

con lo que el exponente mínimo es -6 y el exponente máximo 7.

	`signo`	`exponente`	`mantisa`
menor real positivo normalizado	`0`	`0001`	`0000`

es decir

$$1{,}0000 \times 2^{-6}$$

que en decimal será

$$2^{-6} = 0{,}015625$$

	`signo`	`exponente`	`mantisa`
mayor real positivo normalizado	`0`	`1110`	`1111`

es decir

$$1{,}1111 \times 2^7$$

que en decimal será

$$(1 + 2^{-1} + 2^{-2} + 2^{-3} + 2^{-4}) \times 2^7 = 248$$

5. Con 3 bits tenemos $2^m = 2^4 = 16$ posibles exponentes. Pero le tenemos que quitar el primer valor 0000 y el último 1111 que están reservados. Por lo tanto tenemos $16 - 2 = 14$ posibles exponentes.
Por otra parte, tenemos 4 bits para la mantisa, lo que quiere decir $2^4 = 16$ posibles mantisas.
Por lo tanto $14\,\text{exponentes} \times 16\,\text{mantisas} = 224$ números positivos normalizados distintos.

Computación Numérica

Primer Parcial B - Mayo 2016

1. Sea la ecuación:
$$\frac{1}{x}+x^2-2=0$$

   a) Demostrar que en $[-2,-1]$ existe una única raíz.

   b) Aproximar la raíz haciendo tres iteraciones utilizando el Método de Bisección.

   c) Dar una cota del error cometido al calcular esta raíz.

   d) Dada la función
$$f(x)=\frac{1}{x}+x^2-2,$$
   demostrar que la raíz de $f(x)=0$ y de $g(x)=x$ en $[-2,-1]$ es la misma, siendo
$$g(x)=-\sqrt{\frac{2x-1}{x}}.$$

   e) Utilizando g, aproximar la raíz de f. Hacer 3 iteraciones partiendo de $x_0=-2$.

(a) Las condiciones (suficientes, no necesarias) que ha de cumplir $f(x)=\dfrac{1}{x}+x^2-2$ en $[-2,-1]$ para que exista una única raíz en el intervalo son:
1. f continua: f es continua porque es la suma de funciones continuas, el polinomio lo es siempre y $\dfrac{1}{x}$ es continua en $[-2,-1]$ porque no se anula su denominador.
2. f tiene distinto signo en los extremos del intervalo: $f(-2)=1{,}5$ y $f(-1)=-2$
3. $f'>0$ o $f'<0$ en $(-2,-1)$:

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

Como x es negativo en $(-2,-1)$, $2x^3$ también y por lo tanto $2x^3-1$ es negativo en el intervalo. Como x^2 es siempre positivo $f'<0$ en $(-2,-1)$.

(b) Hacemos tres iteraciones por bisección

k	*a*	*m*	*b*	*f(a)*	*f(m)*	*f(b)*	*cota de error*
0	-2	(a+b)/2=(-2-1)/2=-1.5	-1	1.5	-0.42	-2	-1-(-2)=1
1	-2	(a+b)/2=(-2-1.5)/2=-1.75	-1.5	1.5	0.49	-0.42	-1.5-(-2)=0.5
2	-1.75	(a+b)/2=(-1.75-1.5)/2=-1.625	-1.5	0.49	0.025	-0.42	-1.5-(-1.75)=0.25
3	-1.625		-1.5	0.025		-0.42	-1.5-(-1.625)=0.125

Y podemos dar como raíz aproximada $-1{,}625$ (la raíz verdadera es $-1{,}61803\in[-1{,}625,-1{,}5]$).

(c) La cota de error es la longitud del último intervalo, $-1{,}5-(-1{,}625)=0{,}125$

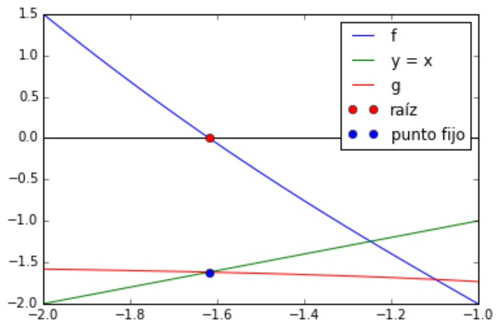
(d)

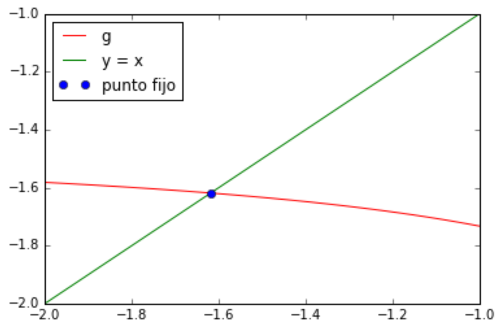
$$f(x) = 0 \Rightarrow \frac{1}{x} + x^2 - 2 = 0 \Rightarrow x^2 = 2 - \frac{1}{x} \Rightarrow x = -\sqrt{\frac{2x-1}{x}}$$

como la raíz que estamos buscando es negativa (está en [-2,-1]) tomamos la raíz cuadrada negativa. Si llamamos $g(x) = -\sqrt{\frac{2x-1}{x}}$ equivale a $x = g(x)$.

(e) Hagamos 3 iteraciones

k	$x_{k+1} = g(x_k)$
0	$x_0 = -2$
1	$x_1 = g(x_0) = g(-2) = -1{,}58114$
2	$x_2 = g(x_1) = g(-1{,}58114) = -1{,}62248$
3	$x_3 = g(x_2) = g(-1{,}62248) = -1{,}61751$





2. La función $f(x) = xe^{-x}$ tiene una única raíz $\alpha = 0$. Demostrar que para cualquier $x_0 > 1$ las iteraciones de Newton se alejan de la raíz α.

La función f tiene por derivada

$$f'(x) = e^{-x} + x\,e^{-x}(-1) = (1-x)e^{-x}.$$

Y la primera iteración definida por el método de Newton será

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = x_0 - \frac{x_0\,e^{-x_0}}{(1-x_0)\,e^{-x_0}} = \frac{x_0(1-x_0) - x_0}{1-x_0} = \frac{x_0^2}{x_0 - 1}.$$

es decir

$$x_1 = x_0 \frac{x_0}{x_0 - 1}.$$

Si $x_0 > 1$ tanto el numerador como el denominador de la fracción

$$\frac{x_0}{x_0 - 1}$$

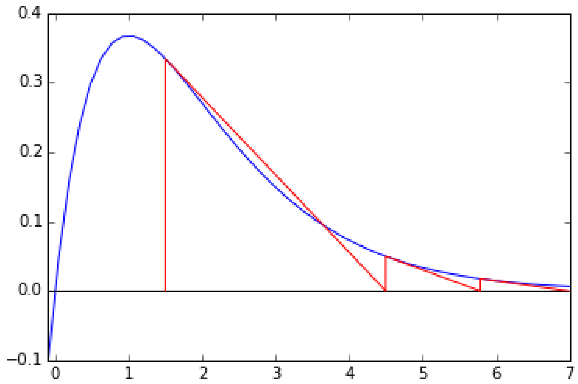
son positivos y además el numerador es mayor que el numerador, por lo que

$$\frac{x_0}{x_0 - 1} > 1$$

y podemos escribir

$$x_1 = x_0 \frac{x_0}{x_0 - 1} > x_0 \times 1 = x_0.$$

Es decir, $x_1 > x_0$. De forma análoga demostraríamos que $x_2 > x_1$ y así sucesivamente con lo que $x_{k+1} > x_k$ para todo $k \geq 0$ y la sucesión es creciente. Como $x_0 > 1$ nos alejamos cada vez más de la raíz $\alpha = 0$.



Computación Numérica

Segundo Parcial - Mayo 2016

1. Tenemos los siguientes datos del movimiento de un cuerpo

$t(s)$	0	2	4	6	8	10
$v(m/s)$	0	6	20	42	72	110

   a) ¿Cuál será la velocidad a los 6,2 s? Usar interpolación polinomial lineal con el polinomio interpolante en la forma de Newton.

   b) ¿Y la aceleración? Utilizar derivación numérica con $h = 0{,}2$.

Para interpolar con un polinomio de grado uno, necesitamos dos puntos. Los t más próximos a 6,2 son 6 y 8. Además el intervalo $[6, 8]$ contiene a 6,2.
Para usar la notación habitual $x = t$ y $f(x) = v(t)$. Construimos la tabla de diferencias divididas con $x_0 = 6$ y $x_1 = 8$:

x	$f(x)$	
6	42	
		$\frac{72-42}{8-6} = 15$
8	72	

Ya tenemos

$$f\left[x_0\right] = 42, \quad f\left[x_0, x_1\right] = 15$$

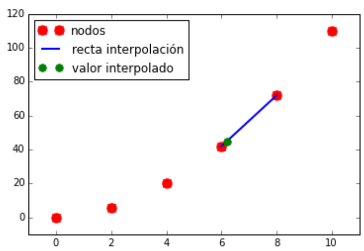
Construímos el polinomio interpolante en la forma de Newton

$$P_2\left(x\right) = f\left[x_0\right] + f\left[x_0, x_1\right]\left(x - x_0\right)$$
$$P_2\left(x\right) = 42 + 15\left(x - 6\right)$$

y entonces

$$v(6{,}2) \approx P_1\left(6{,}2\right) = 42 + 15\left(6{,}2 - 6\right) = 45\,\mathrm{m/s}$$

$$v(6{,}2) \approx 45\,\mathrm{m/s}$$



Podemos aproximar la derivada de una función usando la fórmula centrada:

$$f'(t) \approx \frac{1}{2h}(f(t+h) - f(t-h))$$

Si $f = P_1$, $t = 6{,}2$ y $h = 0{,}2$

$$a(6{,}2) \approx \frac{1}{2h}(P_1(6{,}2+0{,}2) - P_1(6{,}2-0{,}2)) = \frac{1}{2(0{,}2)}(48 - 42) = 15\,\mathrm{m/s}^2$$

2. Calcular la recta $P_1(x) = a_0 + a_1 x$ que minimiza los errores cuadráticos. Deducir la fórmula empleada para calcular esta recta a partir de los errores cuadráticos. Dada la tabla de valores

x	-2	-1	0	1
y	-0.5	0.5	1	2.1

hallar el ajuste lineal (recta de regresión) por mínimos cuadrados.

La recta será de la forma:

$$P_1(x) = a_0 + a_1 x$$

Queremos calcular la recta que minimiza la suma de los errores cuadráticos:

$$E(a_0, a_1) = \sum_{k=1}^{4}(P_1(x_k) - y_k)^2 = \sum_{k=1}^{4}(a_0 + a_1 x_k - y_k)^2.$$

Para hallar el error mínimo calculamos las derivadas parciales respecto a las dos variables y las igualamos a cero:

$$\begin{aligned}
\frac{\partial E}{\partial a_0} &= \sum_{k=1}^{4} 2(a_0 + a_1 x_k - y_k) = 0 \\
\frac{\partial E}{\partial a_0} &= \sum_{k=1}^{4} 2(a_0 + a_1 x_k - y_k)x_k = \sum_{k=1}^{4} 2(a_0 x_k + a_1 x_k^2 - x_k y_k) = 0
\end{aligned}$$

Que equivale a

$$\begin{array}{ccccc}
a_0 \sum_{k=1}^{4} 1 & + & a_1 \sum_{k=1}^{4} x_k & = & \sum_{k=1}^{4} y_k \\
a_0 \sum_{k=1}^{4} x_k & + & a_1 \sum_{k=1}^{4} x_k^2 & = & \sum_{k=1}^{4} x_k y_k
\end{array}$$

Sistema, que expresado matricialmente es:

$$\begin{pmatrix} \sum_{k=1}^{4} 1 & \sum_{k=1}^{4} x_k \\ \sum_{k=1}^{4} x_k & \sum_{k=1}^{4} x_k^2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{4} y_k \\ \sum_{k=1}^{4} x_k y_k \end{pmatrix}$$

El mismo razonamiento, pero en la forma expandida: queremos minimizar la suma de los errores cuadráticos que es

$$\begin{aligned}
E(a_0, a_1) &= (P_1(-2) - (-0{,}5))^2 + (P_1(-1) - 0{,}5)^2 + (P_1(0) - 1)^2 + (P_1(1) - 2{,}1)^2 = \\
&= (a_0 + a_1(-2) - (-0{,}5))^2 + (a_0 + a_1(-1) - 0{,}5)^2 + (a_0 + a_1(0) - 1)^2 + (a_0 + a_1(1) - 2{,}1)^2
\end{aligned}$$

Derivando respecto a las variables:

$$\begin{aligned}
\frac{\partial E}{\partial a_0} &= 2(a_0 + a_1(-2) - (-0{,}5)) + 2(a_0 + a_1(-1) - 0{,}5) + 2(a_0 + a_1(0) - 1) + 2(a_0 + a_1(1) - 2{,}1) = 0 \\
\frac{\partial E}{\partial a_1} &= 2 \quad (a_0 + a_1(-2) - (-0{,}5))(-2) + 2(a_0 + a_1(-1) - 0{,}5)(-1) + \\
&+ \quad 2(a_0 + a_1(0) - 1)(0) + 2(a_0 + a_1(1) - 2{,}1)(1) = 0
\end{aligned}$$

Y sacando a_0 y a_1 factor común

$$\begin{array}{rcccl} a_0(1+1+1+1) & + & a_1((-2)+(-1)+0+1) & = & (-0{,}5)+0{,}5+1+2{,}1 \\ a_0((-2)+(-1)+0+1) & + & a_1((-2)^2+(-1)^2+0^2+1^2) & = & (-2)(-0{,}5)+(-1)(0{,}5)+(0)(1)+(1)(2{,}1) \end{array}$$

Que es

$$\begin{array}{rcrcl} 4a_0 & - & 2a_1 & = & 3{,}1 \\ -2a_0 & + & 6a_1 & = & 2{,}6 \end{array}$$

Resolvemos el sistema por Gauss: la segunda ecuación $e_2 \to e_2 - (-2/4)e_1$

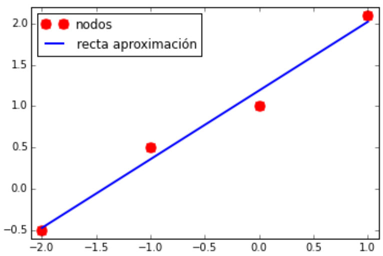
$$\begin{array}{rcrcl} 4a_0 & - & 2a_1 & = & 3{,}1 \\ & & 5a_1 & = & 4{,}15 \end{array}$$

y por sutitución reversiva

$$\begin{array}{lclcl} a_1 & = & 4{,}15/5 & = & 0{,}83 \\ a_0 & = & (3{,}1+2a_1)/4 & = & 1{,}19 \end{array}$$

Y la recta de regresión mínimo cuadrática es

$$P_1(x) = 1{,}19 + 0{,}83x$$



3. Dada la función $f(x,y) = x^2+3y$, calcular la aproximación del gradiente de f $\nabla f(x_0,y_0)$ para $(x_0,y_0) = (1,0)$ con $h_x = h_y = 0{,}2$.

El gradiente viene dado por

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$$

El primer término será:

$$\frac{\partial f}{\partial x}(x_m,y_n) \approx \frac{f(x_m+h_x,y_n) - f(x_m-h_x,y_n)}{2h_x} = \frac{f(1+0{,}2,0) - f(1-0{,}2,0)}{2\,(0{,}2)} =$$
$$= \frac{\left((1{,}2)^2+3(0)\right) - \left((0{,}8)^2+3(0)\right)}{2\,(0{,}2)} =$$
$$= \frac{1{,}44-0{,}64}{0{,}4} = \frac{0{,}8}{0{,}4} = 2$$

Y el segundo término

$$\begin{aligned}\frac{\partial f}{\partial y}(x_m, y_n) \approx &\frac{f\left(x_m, y_n+h_y\right)-f\left(x_m, y_n-h_y\right)}{2h_y} = \frac{f\left(1, 0+0{,}2\right)-f\left(1, 0-0{,}2\right)}{2\left(0{,}2\right)} = \\ = &\frac{\left(1^2+3\left(0{,}2\right)\right)-\left(1^2-3\left(0{,}2\right)\right)}{2\left(0{,}2\right)} = \\ = &\frac{1{,}2}{0{,}4} = 3\end{aligned}$$

Por lo tanto

$$\nabla f\left(1, 0\right) = \left(\partial_x f_1(1, 0), \partial_y f_2(1, 0)\right) \approx (2, 3)$$

4. Deducir la regla del trapecio compuesta a partir de la regla del trapecio simple.

La fórmula simple del trapecio, si $h = x_j - x_i$, es

$$\int_{x_i}^{x_j} f(x)dx \approx \frac{h}{2}\left(f(x_i)+f(x_j)\right)$$

Si dividimos el intervalo $[a, b]$ en n subintervalos, cada uno de ellos de longitud h. Se verifica

$$h = \frac{b-a}{n}$$

y entonces los nodos serían

$$x_i = a + i\,h \quad \text{con} \quad i = 0, 1, 2, \ldots, n$$

Si aplicamos la fórmula simple del punto medio n veces:

$$\begin{aligned}\int_a^b f(x)dx = &\int_{x_0}^{x_1} f(x)dx + \int_{x_1}^{x_2} f(x)dx + \int_{x_2}^{x_3} f(x)dx + \cdots + \int_{x_{n-1}}^{x_{n-1}} f(x)dx + \int_{x_{n-1}}^{x_n} f(x)dx \approx \\ \approx &\frac{h}{2}\left(f(x_0)+f(x_1)\right) + \frac{h}{2}\left(f(x_1)+f(x_2)\right) + \frac{h}{2}\left(f(x_2)+f(x_3)\right) + \\ &+ \cdots + \frac{h}{2}\left(f(x_{n-2})+f(x_{n-1})\right) + \frac{h}{2}\left(f(x_{n-1})+f(x_n)\right) = \\ = &\frac{h}{2}\left(f(x_0)+2f(x_1)+2f(x_2)+2f(x_3)+\cdots+2f(x_{n-1})+f(x_n)\right) = \\ = &\frac{h}{2}\left(f(x_0)+2\left(f(x_1)+f(x_2)+f(x_3)+\cdots+f(x_{n-1})\right)+f(x_n)\right) = \\ = &\frac{h}{2}\left(f(a)+2\sum_{i=1}^{n-1} f(x_i)+f(b)\right)\end{aligned}$$

5. Calcular utilizando la regla del punto medio compuesta con 5 subintervalos la integral.

$$I = \int_1^2 \ln x \, dx.$$

La fórmula simple del punto medio, si $h = x_j - x_i$ y se tiene que $\bar{x}_j = \dfrac{x_i + x_j}{2}$

$$\int_{x_i}^{x_j} f(x)dx \approx hf(\bar{x}_j)$$

Dividimos el intervalo $[1, 2]$ en n subintervalos, cada uno de ellos de longitud h. Se verifica

$$h = \frac{b-a}{n} = \frac{2-1}{5} = \frac{1}{5} = 0{,}2.$$

Los extremos de los intervalos serán

$$\begin{aligned}
x_0 =& a = 1 \\
x_1 =& x_0 + h = 1 + 0{,}2 = 1{,}2 \\
x_2 =& x_1 + h = 1{,}2 + 0{,}2 = 1{,}4 \\
x_3 =& x_2 + h = 1{,}4 + 0{,}2 = 1{,}6 \\
x_4 =& x_3 + h = 1{,}6 + 0{,}2 = 1{,}8 \\
x_5 =& x_4 + h = 1{,}8 + 0{,}2 = 2 = b
\end{aligned}$$

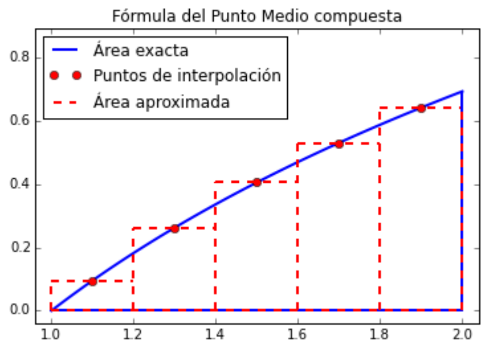
Y los puntos medios de los intervalos seránLos extremos de los intervalos serán

$$\begin{aligned}
\bar{x}_1 =& (x_0 + x_1)/2 = (1 + 1{,}2)/2 = 1{,}1 \\
\bar{x}_2 =& (x_1 + x_2)/2 = (1{,}2 + 1{,}4)/2 = 1{,}3 \\
\bar{x}_3 =& (x_2 + x_3)/2 = (1{,}4 + 1{,}6)/2 = 1{,}5 \\
\bar{x}_4 =& (x_3 + x_4)/2 = (1{,}6 + 1{,}8)/2 = 1{,}7 \\
\bar{x}_5 =& (x_1 + x_2)/2 = (1{,}8 + 2)/2 = 1{,}9
\end{aligned}$$

Aplicamos la fórmula simple del punto medio 5 veces:

$$\begin{aligned}
\int_1^2 \ln x dx =& \int_{x_0}^{x_1} f(x)dx + \int_{x_1}^{x_2} f(x)dx + \int_{x_2}^{x_3} f(x)dx + \int_{x_3}^{x_4} f(x)dx + \int_{x_4}^{x_5} f(x)dx \approx \\
\approx& hf(\bar{x}_1) + hf(\bar{x}_2) + hf(\bar{x}_3) + hf(\bar{x}_4) + hf(\bar{x}_5) = \\
=& h\left(f(\bar{x}_1) + f(\bar{x}_2) + f(\bar{x}_3) + f(\bar{x}_4) + f(\bar{x}_5)\right) = \\
=& 0{,}2\left(\ln 1{,}1 + \ln 1{,}3 + \ln 1{,}5 + \ln 1{,}7 + \ln 1{,}9\right) = 0{,}3871
\end{aligned}$$

(El valor exacto es 0.3863)



Computación Numérica

Tercer Parcial - Mayo 2016

1. Sea el sistema $A\mathbf{x} = \mathbf{b}$ donde

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 1 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix}$$

   a) ¿Es A diagonal dominante por filas? Con este resultado ¿qué podemos concluir acerca de la convergencia del método de Jacobi?

   b) Calcular la matriz de iteración de Jacobi B_J.

   c) Calcular la norma infinito de la matriz de iteración de Jacobi B_J. Con este resultado ¿qué podemos concluir acerca de la convergencia del método de Jacobi?

   d) Calcular la norma uno de B_J. Con este resultado ¿qué podemos concluir acerca de la convergencia del método de Jacobi?

   e) Calcular los autovalores de B_J. Con este resultado ¿qué podemos concluir acerca de la convergencia del método de Jacobi?

   f) Realizar 2 iteraciones con Jacobi. Utilizar $\mathbf{x}^{(0)} = \mathbf{0}$

(a)

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 1 \end{pmatrix} \quad \begin{array}{l} |1| \not> |1| + |0| \\ |4| > |1| + |1| \\ |1| \not> |0| + |1| \end{array}$$

No, puesto que para la primera fila no se verifica $|1| > |1| + |0|$ y para la tercera fila $|1| \not> |0| + |1|$ y como A no es diagonal dominante, no podemos concluir nada respecto a la convergencia del método de Jacobi.

(b) Se tiene que $B_J = -D^{-1}(L + U)$. O también:

1. Dividimos cada fila por el correspondiente elemento de la diagonal.

$$\begin{pmatrix} 1 & 1 & 0 \\ 1/4 & 1 & 1/4 \\ 0 & 1 & 1 \end{pmatrix}$$

2. Cambiamos todos los elementos de signo.

$$\begin{pmatrix} -1 & -1 & 0 \\ -1/4 & -1 & -1/4 \\ 0 & -1 & -1 \end{pmatrix}$$

3. Ponemos ceros en la diagonal principal.

$$B_J = \begin{pmatrix} 0 & -1 & 0 \\ -1/4 & 0 & -1/4 \\ 0 & -1 & 0 \end{pmatrix}$$

(c)

$$B_J = \begin{pmatrix} 0 & -1 & 0 \\ -1/4 & 0 & -1/4 \\ 0 & -1 & 0 \end{pmatrix} \begin{array}{l} 0+1+0=1 \\ 1/4+0+1/4=1/2 \\ 0+1+0=1 \end{array}$$

$$||B_J||_\infty = \text{Max}\,(1, 1/4, 1) = 1$$

Y como $||B_J||_\infty \nless 1$ no podemos concluir nada respecto a la convergencia del Método de Jacobi.

(d)

$$B_J^T = \begin{pmatrix} 0 & -1/4 & 0 \\ -1 & 0 & -1 \\ 0 & -1/4 & 0 \end{pmatrix} \begin{array}{l} 0+1/4+0=1/4 \\ 1+0+1=2 \\ 0+1/4+0=1/4 \end{array}$$

$$||B_J||_1 = \text{Max}\,(1/4, 2, 1/4) = 2$$

Y como $||B_J||_1 \nless 1$ no podemos concluir nada respecto a la convergencia del Método de Jacobi.

(e)

$$\begin{vmatrix} 0-\lambda & -1/4 & 0 \\ -1 & 0-\lambda & -1 \\ 0 & -1/4 & 0-\lambda \end{vmatrix} = -\lambda^3 - \left(-\frac{1}{4}\lambda - \frac{1}{4}\lambda\right) = \frac{1}{2}\lambda - \lambda^3 = \lambda\left(\frac{1}{2} - \lambda^2\right) = 0$$

Y los autovalores son

$$\lambda_1 = 0, \quad \lambda_{2,3} = \pm\sqrt{\frac{1}{2}} = \pm 0{,}71$$

Como todos los autovalores de B_J son menores que uno en valor absoluto podemos concluir que el Método de Jacobi será convergente para cualquier valor inicial.

(f) El sistema es

$$\begin{array}{rrrcl} x & +y & & = & 3 \\ x & +4y & +z & = & 4 \\ & y & +z & = & 2 \end{array}$$

En la primera ecuación despejamos la primera incógnita, x, en la segunda, la segunda incógnita, y, y finalmente, z.

$$\begin{array}{rcl} x & = & 3-y \\ y & = & (4-x-z)/4 \\ z & = & 2-y \end{array}$$

Realizamos la primera iteración

$$\begin{array}{rcl} x^{(1)} & = & 3-y^{(0)} = 3-0 = 3 \\ y^{(1)} & = & \left(4-x^{(0)}-z^{(0)}\right)/4 = (4-0-0)/4 = 1 \\ z^{(1)} & = & 2-y^{(0)} = 2-0 = 2 \end{array}$$

y la segunda

$$\begin{array}{rcl} x^{(2)} & = & 3-y^{(1)} = 3-1 = 2 \\ y^{(2)} & = & \left(4-x^{(1)}-z^{(1)}\right)/4 = (4-3-2)/4 = -1/4 = -0{,}25 \\ z^{(2)} & = & 2-y^{(1)} = 2-1 = 1 \end{array}$$

2. Minimizar la función $f(x, y) = 4x - 3y + 1$ sujeta a las restricciones.

$$\begin{cases} x + y & \leq 2, \\ x + 2y & \leq 3, \\ x, y & \geq 0. \end{cases}$$

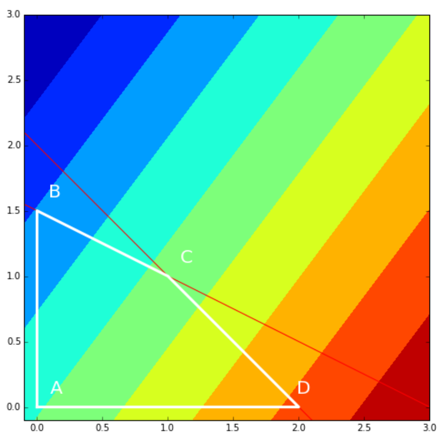
Las condiciones $x \geq 0$ e $y \geq 0$ significan que la región del plano que describen las condiciones está en el primer cuadrante. Dibujamos las demás fronteras de la región donde ha de encontrarse la solución. Tenemos dos rectas, r_1 y r_2

$$r_1 \qquad x + y = 2 \qquad \rightarrow \qquad \frac{x}{2} + \frac{y}{2} = 1$$

$$r_2 \qquad x + 2y = 3 \qquad \rightarrow \qquad \frac{x}{3} + \frac{y}{1{,}5} = 1$$

y entonces r_1 corta al eje OX en $x = 2$ y al eje OY en $y = 2$. Análogamente r_2 corta al eje OX en $x = 3$ y al eje OY en $y = 1{,}5$.

Cada una de estas rectas divide el plano en dos regiones, una que verifica la condición y otra que no. Basta por lo tanto con probar con un punto y si este punto cumple la condición, la región del plano donde se encuentra este punto es pertenece a la región. Por simplicidad, probamos con el origen $(0, 0)$. El origen cumple la condición $x + y \leq 2$ porque $0 + 0 \leq 2$. Y también la condición $x + 2y \leq 3$ porque $0 + 2(0) \leq 3$. Así que la región será la representada en la siguiente figura.



El mínimo estará en un de los vértices del polígono blanco que son

Vértice	x	y	$f(x, y)$
A	0	0	1
B	0	1.5	-3.5
C	1	1	2
D	2	0	9

El mínimo está en B puesto que es el punto de la región donde la función toma el valor más pequeño.

3. $\boxed{\text{Proponer una función objetivo para minimizar la función } f(x,y) = 22 - 4x + x^2 - 12y + 2y^2 \text{ con las restricciones del ejercicio anterior utilizando el método de penalización.}}$

La idea del método de penalización es reemplazar la función objetivo f por otra función

$$F(x,y) = f(x,y) + cP(x,y)$$

y resolver el problema sin restricciones. Para ello tomamos c como una constante positiva y P satisfaciendo:

- P es continua en el dominio de f.
- $P(x,y) \geq 0$ para todo punto del domino de f, y
- $P(x,y) = 0$ si y solo si el punto (x,y) satisface las restricciones.

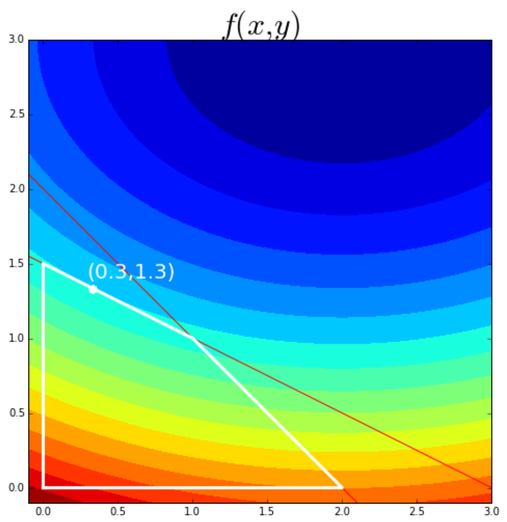
Una posible función para aproximar el mínimo con restricciones sería

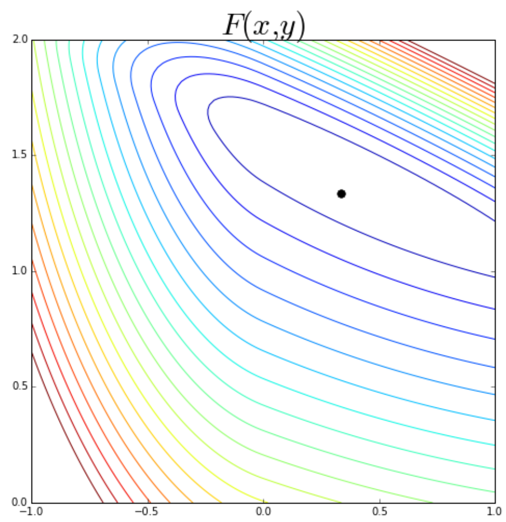
$$F(x,y) = f(x,y) + 10g_1(x,y) + 10g_2(x,y) + 10g_3(x,y) + 10g_4(x,y)$$

$$g_1(x,y) = \begin{cases} 0 & \text{si } x \geq 0 \\ x^2 & \text{si } x < 0 \end{cases} \qquad g_2(x,y) = \begin{cases} 0 & \text{si } y \geq 0 \\ y^2 & \text{si } y < 0 \end{cases}$$

$$g_3(x,y) = \begin{cases} 0 & \text{si } x + y \leq 2 \\ (x+y-2)^2 & \text{si } x + y > 2 \end{cases} \qquad g_4(x,y) = \begin{cases} 0 & \text{si } x + 2y \leq 3 \\ (x+2y-3)^2 & \text{si } x + 2y > 3 \end{cases}$$

Si representamos la función $f(x,y)$ el mínimo está donde el contorno es tangente a la curva de nivel en el punto (0,3, 1,3). Si representamos la función $F(x,y)$ vemos que tiene un mínimo cerca de (0,3, 1,3).





4. Dada la función

$$f(x,y) = (x-1)^2 + y^2$$

a) Hallar un mínimo local de f.

b) Probar que dicho mínimo es, de hecho, global.

c) Aproximar el mínimo de la función comenzando por el punto inicial $(0,1)$ y utilizando el método del gradiente (una iteración).

(a) La condición necesaria de mínimo para un punto (x_m, y_m) es que $\nabla f(x_m, y_m) = \left(f'_x, f'_y\right)(0,0)$. Es decir

$$\begin{cases} f'_x = 0 \\ f'_y = 0 \end{cases} \qquad \begin{cases} 2(x-1) = 0 & \to \\ 2y = 0 & \to \end{cases} \qquad \begin{cases} x = 1 \\ y = 0 \end{cases}$$

Y el punto $(x_m, y_m) = (1,0)$ cumple las condiciones necesarias de mínimo. Veamos si también cumple la condición suficiente que es que la matriz Hessiana en el punto (x_m, y_m) sea definida positiva. La matriz Hessiana para la función f en cualquier punto (x,y) es

$$H_f = \begin{pmatrix} f''_{xx} & f''_{xy} \\ f''_{yx} & f''_{yy} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Y en particular, en el punto $(x_m, y_m) = (1,0)$

$$H_f(1,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Veamos si es definida positiva. Para ello los determinantes de los menores principales han de ser estrictamente positivos. Es decir

$$|H_f(1,0)| = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4 > 0 \quad \text{y} \quad |2| = 2 > 0.$$

Por lo tanto la matriz es definida positiva y se cumple la condición suficiente de mínimo.

(b) Se tiene que:

- Cualquier mínimo local de una función convexa es también un mínimo absoluto.
- Si para todo elemento del dominio de f, la matriz hessiana $H_f(a)$ es definida positiva, entonces, la función f es estrictamente convexa.

En este caso particular, la matriz Hessiana f en cualquier punto coincide con la matriz Hessiana en $(1,0)$ y ya hemos demostrado que es definida positiva, por lo que la función f es convexa y el mínimo local $(1,0)$ es también mínimo global.

(c) Realizaremos una iteración por el método del gradiente usando la fórmula

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} - h\nabla f_{(x_0,y_0)} \tag{1}$$

y buscaremos h para que

$$g(h) = f\left(x_0 - hf_x(0,1), y_0 - hf_y(0,1)\right)$$

tenga un valor mínimo.
Tenemos que

$$\nabla f = (f_x, f_y) = \left(2\left(x-1\right), 2y\right)$$

y

$$\nabla f_{(x_0,y_0)} = \nabla f_{(1,1)} = \left(-2, 2\right).$$

$g(h) = f\left(x_0 - hf_x(0,1), y_0 - hf_y(0,1)\right) = f\left(0+2h, 1-2h\right) = \left(2h-1\right)^2 + \left(1-2h\right)^2 = 2\left(1-2h\right)^2$

Optimicemos esta función:

$$g'(h) = 4\left(1-2h\right)\left(-2\right) = 0 \Longrightarrow 1-2h = 0 \Longrightarrow h = \frac{1}{2}$$

y usando este valor en la ecuación (1)

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} -2 \\ 2 \end{pmatrix} = \begin{pmatrix} 0+1 \\ 1-1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Hemos llegado al mínimo porque $f\left(x_0, y_0\right) = f(0,1) = 2$ y $f\left(x_1, y_1\right) = f(1,0) = 0$ que es el menor valor posible puesto que f es siempre positivo.