γ 4 GAMMA setembro 2004

# Entender el método de Newton

Manuel Díaz Regueiro

En el método de Newton el objetivo es resolver una ecuación f(x)=0. La idea más simple del mundo que a uno se le puede ocurrir es expresarlo $x=f^{-1}(0)$, y ya está resuelta. De hecho que cuando f es una función de inversa conocida la solución aparece en Secundaria y suspendemos a nuestros alumnos si no saben resolver *ln* x=0 o *cos* x=0. Con toda justificación.

Lo que viene a continuación tiene una traducción y un resumen visual: en el método de Newton tomamos la recta tangente a la función en un punto próximo a la solución y calculamos el punto de corte de esa tangente con el eje de las X. Pues bien, la imagen que veremos a continuación es la de la curva de orden superior deducida de la serie de Taylor de la función en un punto próximo a la solución y el cálculo del punto de corte de esa curva con el eje de las X. Esto es posible verlo tanto en la función original-*lo que es mucho más complejo de resolver*- como *en la función inversa*, en este caso aplicado a $y=f^{-1}(x)$ en un *x* próximo a cero, tratando de determinar ahora el corte de la curva tangente con el eje de las Y.

Pensar en la inversa, decía Jacobi, entusiasmado por lo que esta idea –pensar en la inversa de la función- le rentó en las integrales elípticas. Pensar en la inversa y pensar de modo simple, podriamos añadir.

**La serie de potencias de la función inversa.**

A veces, ciertos resultados como el de Lagrange de la fórmula de la inversa interfieren e impiden ver matemáticas simples desde un punto simple. Cuando escribí el artículo *Series de potencias de una función* tenía ciertos recelos de que cosas tan simples no estuvieran escritas previamente. Mirando en Internet hoy en día cuando uno busca serie de potencias de una función inversa encuentra el teorema de Lagrange-Bürmann, en variable compleja, y con restos complejos, pero sigue sin estar explicado-conocido-divulgado para una función real de variable real. Es decir, nadie leyó el citado artículo, excepto quizá la persona que lo referenció en el *Mathematical Reviews*.

¿Cuál sería esa serie de potencias si hubiese extendido unas líneas más ese artículo citado?

Partiendo de la *serie de potencias de una función* en Regueiro (1982):

$f(x)=f(a)+\mathcal{F}(f)(a)(p(x)-p(a))+\mathcal{F}^2(f)(a)(p(x)-p(a))^2/2!+\ldots+\mathcal{F}^n(f)(a)(p(x)-p(a))^n/n!+\mathcal{F}^{n+1}(f)(c)(p(x)-p(a))^{n+1}/(n+1)!$

Donde $\mathcal{F}=1/p'(x)D$ . Y una de las posibles fórmulas do resto, siendo *c* un punto intermedio entre *a* y *x* da:

$$\frac{\mathcal{F}^{n+1}(f)(c)(p(x)-p(a))^{n+1}}{(n+1)!}$$

Si y=f(x) es la función, $y_0=f(x_0)$, la función inversa tiene de serie de potencias $\mathbf{x=x_0+\mathcal{F}(x)[x_0](y-y_0)+\mathcal{F}^2(x)[x_0](y-y_0)^2/2!+\mathcal{F}^3(x)[x_0](y-y_0)^3/3!+\ldots+\mathcal{F}^n(x)[x_0](y-y_0)^n/n!+\ldots}$ donde $\mathcal{F}$ es un operador diferencial $\mathcal{F}=1/f'(x)D$, ya que estamos expresando x como serie de potencias de $f(x)-f(x_0)$.

Coincide con la serie de la fórmula de Lagrange, fórmula 3.6.6 del *Handbook* que dice que si y=f(x), $y0=f(x_0)$, $f'(x_0)\neq 0$

$$x=x_0+\sum_{K=1}^{\infty}\frac{(y-y_0)^k}{k!}\left[\frac{d^{k-1}}{dx^{k-1}}\left(\frac{x-x_0}{f(x)-y_0}\right)^k\right]_{x=x_0}$$

Ya que la serie debe ser única,

$$\left[\frac{d^{k-1}}{dx^{k-1}}\left(\frac{x-x_0}{f(x)-y_0}\right)^k\right]_{x=x_0} = F^k(x)[x_0] = \left[\left(\frac{1}{f'(x)}D\right)^k(x)\right][x_0]$$

(que como se puede comprobar para k=1 o k=2, es cierto sólo tomando el límite de la expresión de la izquerda cuando x tiende a $x_0$); límite que hace sumamente complejo el cálculo de los coeficientes y explica que lo que viene a continuación no esté contado ya desde los tiempos de Lagrange. Que su relación con el método de Newton no haya sido suficientemente expresada o conocida. Naturalmente, puestos a escoger, prefiero la mía que es máis operacional, simple, y clara- para mi. Además de más simple de calcular y probar- no más que la serie de Taylor. Tiene resto y podemos deducir conclusiones y resultados notables de ella.

También aparece la serie inversa en otro lugar del *Handbook of Mathematical Functions,* (3.6.25) pero como tales coeficientes de la serie inversa, a los que se puede llegar resolviendo el sistema de ecuaciones resultante de substituir en la serie de Taylor de la función una serie de potencias -con coeficientes a calcular-que dé por resultado x, es decir, utilizando que $f(f^{-1}(x))=x$. El resultado es complejo e ininteligible, porque no se sabe de donde proceden esos coeficientes de expresión cada vez más complicada con *n*. Pero estábamos con el método de Newton y vamos a ver qué tiene que ver la fórmula del desarrollo en serie de la inversa con éste método. Lo que nos interesa en este artículo es iluminar si esto nos sirve para entender mejor el método de Newton.

Podemos leer en Internet como “The first attempt to use the second order expansion is due to the astronomer E. Halley (1656-1743) in 1694. “. Halley llega al segundo miembro de la serie, y en esa misma página se pueden ver los métodos de Householder (1970), sin que haya una explicación sencilla de ellos. Mathworld explica la “Halley’s Irrational Formula” mediante la intersección -con el eje de las X- de la aproximación de Taylor de segundo orden explicada. De hecho, al hacerla así tenemos que resolver una ecuación de segundo grado y aparece una raíz cuadrada, que ya no aparece en la explicación del *Householder’s Method.*

Tratemos de buscar esa explicación sencilla.

$\mathcal{F}(x)=1/f'(x)*1=1/f'(x)$, $\mathcal{F}^2(x)=-f''(x)/(f'(x))^3$, etc...

Si ahora y=0 –ya que para el *x* que estamos buscando, solución de la ecuación, debe ser y=f(x)=0, tenemos que

$\mathbf{x=x_0-f(x_0)/f'(x_0)+f(x_0)^2(-f''(x_0)f'(x_0)/(f'(x_0))^3)/2+....}$

**Voila!.** O **Eureka!**. Ya que:

-Los dos primeros términos de la solución expresada con esta serie son los de la fórmula del método de Newton. Los tres primeros son los de Halley-Householder. Ahora si ya sabemos un nuevo porqué de la efectividad del método de Newton.

-Por lo tanto, para generalizar el método de Newton sólo tenemos que ir añadiendo al método iterativo más sumandos de la fórmula de la inversa, aplicada a este caso en que y=0:

$$\mathbf{x=x_0+\mathcal{F}(x)[x_0](-y_0)+\ \mathcal{F}^2(x)[x_0]\ (-y_0)^2/2!+\mathcal{F}^3(x)[x_0](-y_0)^3/3!+.....+\ \mathcal{F}^n(x)[x_0](-y_0)^n/n!+.....}$$

En Internet podemos ver outros métodos de generalizar el método de Newton, pero con orientaciones totalmente diferentes de las que se presenta aquí.

-Si utilizamos todos los términos tenemos un método directo, no iterativo, del cálculo de la solución buscada. Teóricamente es posible calcularla con cualquier precisión, a costa de la mayor complejidad de los términos sucesivos. Esto hoy en día no es tan problemático al disponer de herramientas informáticas de cálculo simbólico.

-Se puede generalizar perfectamente a varias variables,o espacios, aún que no vamos aquí a llenar de fórmulas este pequeño artículo con algo que puede reproducir cualquier matemático. El método de Newton es un método central en este caso. Buscaríamos puntos óptimos –por poner un ejemplo- no por la dirección del mayor gradiente, sino por la dirección de mayor gradiente de la superficie tangente de 2º, 3º ... grado. Iríamos más directos a la solución , lo que en problemas extremadamente complejos de economía o bioinformática puede ser mucho más satisfactorio y eficiente.

-Mejoran los métodos de Householder.

-Pero el resultado más notable (en cuanto a entender la eficiencia de los métodos de Newton generalizados utilizando la inversa) es el siguiente- notablemente simple y rápido de demostrar en su versión general.

Teorema $\|\varepsilon_k\| \le K\|\varepsilon_{k-1}\|^{n+1}$. Siendo $\varepsilon_k$ el error cometido por la iteración de orden *k* del método de Newton de orden *n*. Es decir, en cada paso de la iteración disminuímos el error del resultado, siendo éste proporcional a una potencia n+1 del error de la iteración anterior.

Los detalles:

Sea $g(x)=x+\mathcal{F}(x)(-f(x))+\mathcal{F}^2(x)(-f(x))^2/2+\mathcal{F}^3(x)(-f(x))^3/3!+....+\mathcal{F}^n(x)(-f(x))^n/n!$ la función iterativa del método de Newton por la aproximación n-sima de la inversa. Sea *a* la solución de f(x)=0 que buscamos, es decir, f(a)=0 y g(x) cumple las condiciones del teorema de Cauchy en [a,x], esencialmente que $f\in C^{n+1}[a,x]$. Y $\mathcal{F}^{n+1}(x)$ está acotada entre *a* y *x*. Y f'(x) también, además de no anularse en [a,x].
Así $x_k=g(x_{k-1})$ y la tesis es que $|x_k-a|\leq K|x_{k-1}-a|^{n+1}$

Demostración.

Por el teorema de Cauchy

$$\frac{g(x)-g(a)}{f(x)-f(a)}=\frac{g'(c)}{f'(c)}$$

siendo *c* un punto intermedio a *a* y *x*. f(a)=0 Y g(a)=a por ser *a* solución de f(x)=0.

$f(x)-f(a)=f'(c_1)(x-a)$ por el teorema del valor medio. Y $g'(c)/f'(c)=\mathcal{F}(g(x))[c]$.

Por lo que calculamos $\mathcal{F}(g(x))=\mathcal{F}(x)+\mathcal{F}^2(x)(-f(x))+\mathcal{F}(x)(-1)+\mathcal{F}^3(x)(-f(x))^2/2+\mathcal{F}^2(x)f(x)+\mathcal{F}^4(x)(-f(x))^3/3!-\mathcal{F}^3(x)f(x)^2/2....+\mathcal{F}^{n+1}(x)(-f(x))^n/n!=\mathcal{F}^{n+1}(x)(-f(x))^n/n!$ Puesto que los demás sumandos se anulan dos a dos(como en una suma telescópica).

Por lo que

$$g(x)-g(a)=\left[\frac{\mathcal{F}^{n+1}(x)(-f(x))^n}{n!}\right][c](f(x)-f(a))$$

$$=\left[\frac{\mathcal{F}^{n+1}(c)}{n!}\right](f(a)-f(c))^n(f(x)-f(a))=$$

$$=(-1)^n\left[\frac{\mathcal{F}^{n+1}(c)}{n!}\right](f(c)-f(a))^n(f(x)-f(a))$$

Aplicando ahora el hecho de que $\mathcal{F}^{n+1}(x)$ está acotada entre *a* y *x* por $K_1$. Y que f'(x) lo está por $K_2$, aplicando el teorema del valor medio a f(c)-f(a) y a f(x)-f(a) llegamos a que

$$\|g(x)-g(a)\|\leq\frac{K_1K_2^{n+1}}{n!}\|x-a\|^{n+1}$$ por lo que

$$\|g(x_{k-1})-g(a)\|\leq K\|x_{k-1}-a\|^{n+1}$$ es decir

$$\|x_k-a\|\leq K\|x_{k-1}-a\|^{n+1}$$ c.q.d.

Demostrar esto para n=1 utilizando Taylor es casi tan complejo como esta demostración. Utilizar la fórmula de Lagrange de la inversa para esto es poco menos que imposible.

Y, para finalizar, una aplicación al cálculo de raíces supereficiente.

Dado f(x)=0, tomemos un valor ε pequeño; tenemos que f(x)+ ε=ε. Entonces llamando g(x)=f(x)+ ε, la solución es $x=g^{-1}(\varepsilon)$. Si conocemos un $x_0$ tal que $g^{-1}(x_0)=0$, o bien -simplemente-próximo a la raíz buscada x, podemos hacer el desarrollo en serie de la inversa, de $g^{-1}$,en $x_0$, y calcularla con $x=g^{-1}(\varepsilon)$ .

Un ejemplo lo aclarará. Calculemos una raíz de $x^{27}-2x+1{,}01=0$. $g(x)=x^{27}-2x+1$ e g(1)=0. El desarrollo de la inversa de g en 1 es:

x=1 + 1/25·y - 351/15625·y^2 + 173277/9765625·y^3- 19770426/1220703125y^4+ 61188852114/ 3814697265625 y^5

Con seis sumandos e y= -0,01 tenemos que x = 0.9995977356926588170428791247328786840237307810960155204634384877862616893413086061119795060869434757699048162008433485268593472066072615740067869208777050868150137416863032285226960522774817316681 4038..., que da una precisión de más de 300 decimales, comprobado con Derive. Sin embargo, hay que resaltar que con un término menos sólo aproximaba 11 decimales!. La conclusión, Derive utiliza la serie hasta el quinto término. Los coeficientes salen de $1/(27x^{26}-2)$ [1]=1/25 $(1/(27x^{26}-2)D(1/(27x^{26}-2)D](x)[1]/2=$ -351/15625, etc, etc.

Decir, por último, que estas series permiten calcular tablas de funciones inversas de alta precisión. O valores numéricos. ¿Alguno se atreve con el número de oro?.

**Bibliografía.**


Manuel Díaz Regueiro (1982). Series de potencias de una función en *Actas IX Jornadas Matemáticas Hispano-Lusas, Vol. I.* Universidad de Salamanca.

*Newton's method and high order iterations* en http://numbers.computation.free.fr/Constants/Algorithms/newton.html

Newton, (1664-1671), *Methodus fluxionum et serierum infinitarum*

Abramowitz e Stegun (1972-décima edición). *Handbook of Matemátical functions*. Dover. New York.