UTILIZACIÓN DEL MATLAB PARA LA REALIZACIÓN DE ANIMACIONES QUE SIMULEN PROCESOS ITERATIVOS PARA LA RESOLUCIÓN DE ECUACIONES NO LINEALES

Fernández Gutierrez, M.J., García Gonzalo, M.E.

Departamento de Matemáticas.Universidad de Oviedo
e-mail: mjfg@pinon.ccu.uniovi.es

RESUMEN: En el presente artículo se desarrollan simulaciones de procesos iterativos de punto fijo, mediante animaciones en tiempo real realizadas con MATLAB. Así se muestran las posibilidades didácticas de dicho programa en este tema ya que permite la visualización de gran variedad de casos y ejemplos, contemplando la convergencia o divergencia, velocidad y orden de convergencia, así como casos prácticos aplicados a la vida real.

1.- INTRODUCCIÓN.

Se desarrolla la presentación de un tema de análisis numérico utilizando los medios informáticos PowerPoint y Matlab. Básicamente consta de una gran variedad de ejemplos del tema de resolución de ecuaciones no lineales, comparando casos parecidos, mostrando diferencias y señalando analogías. En la bibliografía aparece algún ejemplo gráfico, pero limitándose a los casos más básicos. El Power Point se ha utilizado como un apoyo para escribir las fórmulas y desarrollo matemático. El eje principal de la exposición son las gráficas y sobre todo las animaciones, ambas realizadas con Matlab, que muestran como evolucionan las distintas sucesiones hacia la solución del problema. Se considera que en este tema en particular las animaciones son importantes debido a que así se puede visualizar como a partir de una situación inicial parecida la evolución puede ser totalmente distinta. Además las animaciones se autoexplican mientras que si se presenta el mismo ejemplo con una gráfica implica un desglose y a veces una profusa explicación de la misma.

Las animaciones dan un acceso sencillo e intuitivo a secciones del tema que habitualmente no se tratan, contribuyendo así no sólo a aclarar y fijar los conceptos habituales sino también a su ampliación y comparación con los nuevos.

Otra ventaja es que dada la sencillez y brevedad de los programas pueden ser realizados y modificados por prácticamente todos los alumnos (en nuestro caso segundo curso de ingeniería técnica informática) y repetidos tantas veces como sea necesario incluso en un ordenador con pocos requerimientos tanto de procesador como de memoria (se desarrollaron en un ordenador con procesador 486 con 8M de RAM), ya que los programas ocupan muy poco espacio y generan la animación en el instante en que se corre el programa.

2.- EL MÉTODO DE ITERACIÓN DE PUNTO FIJO.

a) Antecedentes matemáticos

La resolución de una ecuación $F(x) = 0$ es uno de los problemas matemáticos más antiguos y comunes. Fórmulas exactas, entendiendo por exactas aquellas en las que la solución puede expresarse por medio de radicales, sólo hay en un número reducido de casos, como el de las ecuaciones polinómicas hasta grado cuatro inclusive. En el resto de los casos, en general, la solución ha de obtenerse de forma aproximada. Aunque existen métodos para la resolución numérica de determinados tipos de ecuaciones, como por ejemplo las polinómicas, tiene gran interés la construcción de métodos numéricos que permitan la aproximación de raíces en el caso general. Este es el caso de los métodos conocidos como de iteración de punto fijo que consisten en la transformación de la ecuación inicial en otra del tipo $f(x) = x$, cuyas raíces se denominan puntos fijos, y la utilización para la resolución del problema de la sucesión $x_{n+1} = f(x_n)$. Si la función f es continua y la sucesión tiene límite, éste es un punto fijo de f lo que justifica la utilización del método.

Pero no todas las funciones con un punto fijo generan una sucesión que converge a dicho punto. En general interesa que $f(x)$ varíe poco al variar x: el caso ideal sería f constante; entonces, para x_0 arbitrario x_1 nos daría exactamente el valor buscado.

Se obtendrán condiciones para la convergencia global del proceso de iteración de punto fijo, que en síntesis se trata de la contractividad de la función en un intervalo que se aplica sobre sí mismo utilizando la función de iteración. Con estas condiciones se consigue, no sólo garantizar la convergencia de la sucesión al punto fijo del intervalo sea cual sea el punto inicial tomado dentro de este, sino la unicidad del punto fijo en él. Además, la constante de contractividad asociada a la función y al intervalo sirve para establecer una cota del error.

Sin embargo la condición de contractividad es difícil de verificar en la práctica, y se sustituye por la condición, más exigente pero más sencilla de utilizar, de que el valor absoluto de la derivada de la función de iteración sea menor que uno en el intervalo.

Localmente hablamos de puntos fijos atractivos y repulsores. Esta cualidad viene determinada por el valor de la derivada de la función de iteración f en el punto fijo $\bar{x}$. Si la derivada de f es continua en un entorno de $\bar{x}$ y $|f'(\bar{x})| < 1$, entonces, si partimos de un punto suficientemente cercano a $\bar{x}$ la sucesión converge y el punto se llama atractivo. Si $|f'(\bar{x})| > 1$ sucede lo contrario y el punto es repulsor. Si $|f'(\bar{x})| = 1$ es un caso dudoso.

Una vez determinada la convergencia de un método iterativo es interesante conocer la velocidad de convergencia del mismo, para lo cual se define el concepto de orden de convergencia y se estudia su obtención para los métodos de iteración de punto fijo. Se demuestra que el orden de iteración depende del orden de la primera derivada no nula de f en $\bar{x}$. Con esto se retoma la idea de que interesa una función que varíe lo menos posible: si $f'(\bar{x}) = 0$, la función tiene tangente horizontal en el punto y la función varía muy poco en su entorno; si además, la función tiene curvatura cero en el punto ($f''(\bar{x}) = 0$) y por lo tanto es plana en él, esta variación es aún menor; y en general, cuanto mayor sea el orden de la primera derivada no nula en $\bar{x}$, menos variará la función en un entorno de $\bar{x}$.

Las condiciones para la convergencia local exigían que $|f'(\bar{x})| < 1$. Si $f'(\bar{x}) = 0$ y f' continua, la convergencia local está garantizada sea cual sea el orden de la primera derivada que no se anule. Cuanto mayor sea el orden de convergencia menor el número de iteraciones. También influye, aunque en menor medida, el coeficiente asintótico de convergencia que sólo tiene interés si estamos comparando sucesiones del mismo orden de convergencia.

El orden de convergencia comienza a ser interesante cuando es como mínimo de orden dos, así que una vez establecidas las condiciones que ha de reunir la función de iteración f para tener convergencia de orden dos, se construirá una función de iteración con este orden. De las diferentes posibilidades se toma la función de iteración asociada al método de aproximación de raíces de Newton-Raphson.

Otro punto de partida para alcanzar el mismo resultado podría ser geométrico: partimos de una raíz aproximada x_o y para hallar la siguiente aproximación de la raíz de la curva, sustituimos la curva $F(x)$ por la recta tangente en x_o, tomando como siguiente aproximación de la raíz buscada la raíz de la recta.

La convergencia local está asegurada por la construcción del método, pero sería útil conocer las condiciones que ha de reunir la función $y = F(x)$ en un intervalo $[a,b]$ para que partiendo de un punto cualquiera de él la convergencia esté garantizada. Se observa que la presencia de puntos singulares como máximos, mínimos y puntos de inflexión perjudica en muchos casos la convergencia del método de Newton y en otros hace que converja a una raíz que no interesa, por lo que, básicamente, las condiciones son que la función tenga una única raíz en $[a,b]$, no presente puntos singulares en el intervalo y que el corte de la tangente con el eje OX en el punto de menor pendiente del intervalo esté dentro de $[a,b]$.

Estas condiciones producen cuatro casos diferentes que se pueden reducir a uno: en nuestro caso el de la función F creciente y concava hacia abajo en el intervalo $[a,b]$. Si partimos de un x_o de la zona de $[a,b]$ donde $f''(x_o)f(x_o) < 0$, se genera una sucesión creciente y acotada superiormente, por lo tanto convergente. Y además el límite es el punto fijo de f. Si por el contrario, partimos de un x_o tal que $f''(x_o)f(x_o) > 0$, x_1 está al otro lado de la raíz con lo que nos remitimos al caso anterior.

Aunque se ha construido el método de Newton buscando que su orden de convergencia sea dos, existen casos desfavorables como el de raíces múltiples donde el orden de convergencia del método se reduce a uno. Schöder propuso una modificación del método que hace que este retome su orden de convergencia cuadrático. El inconveniente es que es preciso conocer a priori el orden de multiplicidad de la raíz. Otra situación particular, aunque ya no desfavorable, se produce cuando la raíz de $F(x)=0$ es un punto de inflexión con tangente no horizontal. En este caso el orden de convergencia es superior a dos.

Aunque un método de orden dos supone, en general, la necesidad de muy pocas iteraciones para aproximar la solución con gran exactitud, se plantea la obtención de métodos de orden superior. Así, partiendo de un planteamiento geométrico análogo al del método de Newton, se describe la obtención del método de Euler: en lugar de sustituir la curva por la recta tangente la sustituimos por la parábola tangente a la curva con su misma curvatura en el punto correspondiente a la raíz aproximada. Es importante la forma en que se expresa el algoritmo, puesto que una forma inadecuada conduce a importantes errores de redondeo que invalidan la fórmula.

Se plantean como modificaciones del método de Euler los métodos de Halley y de Chebyshev, todos ellos de orden tres.

b) Desarrollo gráfico

Se comienza con una descripción gráfica del planteamiento del problema, mostrando la diferencia entre la función asociada a la ecuación y la correspondiente función de iteración. Se señala cómo la raíz de la ecuación coincide con el punto fijo de la función de iteración. Se describe brevemente en que consiste el método de iteración de punto fijo. A continuación se muestran tanto gráfica como analíticamente distintas posibilidades de funciones de iteración para una misma ecuación.

Pero no todas las funciones de iteración son válidas para resolver el problema. Es necesario que la función de iteración genere una sucesión y que esta converja al punto fijo. Esto se expresa por medio de tres condiciones. Se van estudiando los requerimientos uno a uno y se proponen condiciones para cada uno de ellos. Conforme imponemos condiciones vamos viendo gráficamente cómo se van eliminando posibles funciones de iteración. Al final, las condiciones que ha de reunir la función de iteración pueden limitarse a dos. Se expresa de formalmente el teorema de convergencia global. De él se infieren gráficamente condiciones de convergencia local. Las gráficas ejemplo permiten apreciar la importancia de la elección de un intervalo simétrico centrado en el punto fijo en la demostración del teorema correspondiente.

A continuación se ve una animación de los típicos cuatro casos (convergente/divergente, escalera/tela de araña) de sucesiones generadas por el método, resaltando la ausencia de convergencia cuando no se cumplen las condiciones de convergencia local aunque el punto inicial de la sucesión esté cerca de la solución.

Las diferencias entre distintas velocidades de convergencia y distinto orden de convergencia con sus condiciones se pueden ver en la siguiente animación. Así se observa que para orden de convergencia uno la velocidad de convergencia aumenta al "tumbar" la curva (al disminuir la derivada) y como se pasa a un orden de convergencia dos cuando la curva tiene tangente horizontal en el punto fijo (la derivada es cero). También se aprecia que cuando la curva es plana en este punto (curvatura cero, derivada segunda cero) se pasa al orden de convergencia tres. Así, a partir de una observación gráfica, se da un criterio analítico para aumentar el orden de convergencia (derivadas sucesivas de la función de iteración cero en el punto fijo).

Con las condiciones analíticas obtenidas se plantea la construcción de un método iterativo de orden dos, el de Newton-Raphson. Se deduce la función de iteración del método a partir de la función correspondiente a la ecuación. No obstante nos parece más interesante en el actual contexto, la segunda opción de construcción del método que se presenta, que tiene como no, una aproximación gráfica. En ella, en cada paso se sustituye la curva que representa la función por la aproximación que constituye una recta tangente a dicha curva.

La convergencia local del método de Newton-Raphson está asegurada por la construcción del método, pero parece interesante estudiar condiciones de convergencia global. Cómo partida se muestra en una animación cómo perjudica a la convergencia del método la existencia en las proximidades de la raíz de máximos, mínimos y puntos de inflexión. Esto permite motivar las tres primeras condiciones del teorema de convergencia global. La cuarta se puede justificar diciendo que puesto que estamos hablando de convergencia global toda la sucesión ha de estar incluida dentro de un intervalo $[a,b]$ conocido. La cuarta condición está orientada a garantizar que incluso si partimos del punto más desfavorable del intervalo, el siguiente está también dentro de él. A continuación se demuestra analíticamente el teorema ilustrando gráficamente cada uno de los pasos de ésta demostración.

Junto con las animaciones que presentan el método de Newton-Raphson con esta interpretación geométrica tradicional, es decir, con F, se presentan también las animaciónes con la función de iteración correspondiente f, que permiten apreciar mejor el orden de convergencia. Esta forma de contemplar para un mismo ejemplo la sucesión utilizando F y f permite apreciar mejor cómo aunque en general el orden de sucesión generada por la función de iteración de Newton-Raphson es dos hay casos donde esto no sucede. Así, se ilustra el caso de raíces múltiples, donde el orden de convergencia se convierte en uno y se ve como modificando de forma adecuada la función de iteración se recupera el orden de convergencia dos (modificación de Shöder). También se ve un caso más favorable, donde el orden de convergencia del método de Newton-Raphson es al menos tres, que es el caso en el que la raíz es un punto de inflexión con tangente no horizontal.

Las animaciones permiten abordar la siguiente parte del tema de forma intuitiva y por analogía con el método de Newton-Raphson. En el método de Newton-Raphson sustituíamos en cada paso la curva correspondiente a la ecuación por la recta tangente a la curva en el punto de la sucesión (aproximación de la raíz). El método de Euler aproxima la curva con la parábola tangente en el punto y que tiene en dicho punto la misma curvatura. Se ve con una animación, cómo efectivamente este método es más rápido en su convergencia que el de Newton-Raphson.
Ya de forma analítica se ven otros dos métodos de orden de convergencia tres, el de Halley y el de Chebyshev, planteados cómo aproximaciónes del método de Euler.

Por último, cómo motivación para los alumnos, se ve un caso de la vida real. Se aplica la resolución de ecuaciones al cálculo de las condiciones de una hipoteca o préstamo (quién no ha tenido que pedir un préstamo para comprar un coche o un piso). Se supone que partimos de un presupuesto mensual aproximado para afrontar el préstamo y se resuelve el problema tanto analítica como gráficamente con las condiciones de interés y cuota que se quiera. El resultado será el número de meses que tendremos que estar pagando dicho préstamo.

3.- CONCLUSIONES

La finalidad principal de este trabajo es mostrar la utilización de métodos informáticos sencillos y accesibles a la enseñanza y aprendizaje de las matemáticas.

Supone un avance sobre el papel impreso en el sentido de que básicamente se trata de animaciones que en gran parte se explican por sí mismas. Así que no es sólo la visión de la imagen estática; a ella se añade el movimiento que permite no sólo asimilar situaciones sino también procesos. El alumno también puede actuar para modificar e incluso construir las animaciones con lo que se añade una componente activa por su parte. Así no sólo se facilita el apoyo del aprendizaje de la materia principal sino que permite el acceso intuitivo y comparado con vistas a ampliaciones de la misma.

4.- BIBLIOGRAFIA

- Aubanell, A., Benseny, A. y Delshams, A.(1993) *Útiles básicos de cálculo numérico. Labor.*
- Hamming, R.W.(1986). *Numerical Methods for Scientists and Engineers.* Dover.
- Hildebrand, F.B. (1956). *Introduction to Numerical Analysis.* Dover Publications.
- Jaluria, Y.(1988). *Computer Methods for Engineering.* Allyn and Bacon.
- Pérez, C. (1996) *Matemática informatizada con MATLAB.* RA-MA.
- Sanz-Serna,(1998) J.M. *Diez lecciones de Cálculo Numérico.* Universidad de Valladolid.
- Viaño Rey, J.M. (1997). *Lecciones de Métodos Numéricos. Resolución de ecuaciones numéricas.* Tórculo.