Sistema de punto flotante

Definición: Un sistema de punto flotante es el conjunto de números en punto flotante normalizados en el sistema de numeración con base B, t dígitos en la parte fraccionaria y L y U valores extremos para la potencia de la base ie. $x = m\ B^e$, donde la mantisa $m = \pm\ d_0, d_1, \dots, d_t$ y $0 \le d_i < B$ con $i = 1, \dots, t$, $0 < d_0 < B$, $1 \le |m| < B$, $L \le e \le U$.

Obs: Aunque el sistema de representación de punto flotante permite representar magnitudes muy diversas, si $e > U$ se tiene error por overflow y si $e < L$ se tiene error por underflow. O no se puede representar en este sistema.

Metodo de Biseccion

El método se basa en el Teorema de valor intermedio para f continua que dice que si f es continua en el intervalo [a,b] y si $f(a)\cdot f(b)<0$ entonces f tiene al menos una raíz en el intervalo (a,b).
Si $f(a)\cdot f(b)<0$, calculamos $c=(b-a)/2$. Si $f(c)=0$, c es la raíz que buscábamos y sino vemos de descartar el a ó el b dependiendo los signos de f(a), f(b) y f(c)
Sea $x_1=c$, la aproximación de la raíz buscada, $|e_1|<|x_1-r|\le(b-a)/2$.

Error en bisección:

Si $[a_0, b_0], \dots, [a_n, b_n]$ denotan los sucesivos intervalos del método de bisección, entonces existe $\lim_{n\to\infty} a_n$ y existe $\lim_{n\to\infty} b_n$ son iguales y representan una raíz r de f. Si $c_n=(a_n+b_n)/2$ y $r=\lim_{n\to\infty} b_n$ entonces $|r-c_n|\le(b_0-a_0)/2^{n+1}$.
Dem: como $[a_j, b_j]$ con $0\le j\le n$, son los intervalos de cada iteración del método de bisección, entonces $a_0\le a_1\le\dots\le a_n\le b_0$ y $a_0\le b_n\le\dots\le b_0$ y $b_{n+1}-a_{n+1}=(b_n-a_n)/2$.
Luego, $\{a_n\}$ es creciente y acotada superiormente por b_0 entonces $\{a_n\}$ es convergente. Análogamente, como $\{b_n\}$ es decreciente y acotada inferiormente por a_0 entonces $\{b_n\}$ es convergente.
Aplicando iterativamente la última desigualdad, obtenemos que $b_n-a_n=(b_0-a_0)/2^n$

Método de Newton

La idea del método es reemplazar el problema difícil de encontrar una raíz, por una sucesión de problemas fáciles cuyas soluciones convergen bajo ciertas hipótesis.
Otra idea intuitiva del método, es suponer que f, f' y f'' son continuas alrededor de r. Si x es una aproximación de r, el desarrollo de Taylor alrededor de x es: $f(x+h) = f(x)+f'(x)\cdot h+f''(t)\cdot h^2/2 = f(x)+f'(x)+O(h^2)$. Tomando $h=r-x$ para que $x+h=x+r-x=r$.
Si h es chico, h^2 es aun mas chico por lo que lo lo podemos despreciar y queda: $O(h^2)$: $0=f(x)+f'(x)\cdot h$, despejando, $h=-f(x)/f'(x)$. Luego $x+h=x-f(x)/f'(x)$ y es la nueva aproximación a r.

Teorema de convergencia del Método de Newton:
Si f, f' y f'' son continuas en un entorno de una raíz r de f y si $f'(r)\neq0$, entonces existe $d>0$ tq si el punto inicial x_0 satisface que $|r-x_0|<d$. Luego todas las aproximaciones generadas por el método $\{x_n\}$ satisfacen $|r-x_n|\le d$, la sucesión $\{x_n\}$ converge a r y la convergencia es cuadrática.
Dem: LA HAGO???

Método de Punto Fijo

Definición: un punto fijo de una función g es un número p tal que $g(p)=p$.
La conexión entre encontrar un punto fijo de g y encontrar una raíz de f es que dado el problema de encontrar $f(p)=0$, podemos definir una función g de punto fijo que sea de la forma $g(x)=x-a\cdot f(x)$. Entonces, si g tiene un punto fijo en p, queda $f(p)=p-g(p)$ y como g tiene punto fijo en p $g(p)=p$ entonces $f(p)=p-p=0$.

Teorema de existencia y unicidad de punto fijo:
Sea g continua en [a,b]: 1) si $g(a)\in[a,b]$ y $g(b)\in[a,b]$ entonces existe $r\in[a,b]$ tal que $g(r)=r$.

2) si además existe $g'(x)$ $\forall x \in (a,b)$ y existe $k \in (0,1)$ tal que $|g'(x)| \leq k$ $\forall x \in (a,b)$ entonces el punto fijo es único en el intervalo $[a,b]$.

Dem: 1) si $g(a)=a$ ó $g(b)=b$ ya es punto fijo. Sino $g(a)>a$ y $g(b)<b$. Definimos $h(x)=g(x)-x$, h continua en el intervalo $[a,b]$, $h(a)=g(a)-a>0$ y $h(b)=g(b)-b<0$. Por teorema del valor intermedio, existe $r \in [a,b]$ tal que $h(r)=0$, entonces $g(r)=r$.

2) Supongamos que existen p y q en $[a,b]$ tal que $g(p)=p$ y $g(q)=q$, con $p \neq q$. Por teorema del valor medio, $g(p)-g(q)=g'(t)\cdot(p-q)$, con t entre p y q. $|p-q| = |g(p)-g(q)| = |g'(t)|\cdot|(p-q)| \leq k\cdot|p-q| \leq |p-q|$, osea que $|p-q|<|p-q|$, lo que es una contradicción, por lo tanto $p=q$, es decir, hay un solo punto fijo en el intervalo.

<u>Teorema de convergencia del método de punto fijo:</u>

Sea g tal que $g(x) \in [a,b]$ $\forall x \in [a,b]$, supongamos que existe $g'(x)$ en (a,b) y una constante $k \in (0,1)$ tal que $|g'(x)| \leq k$ $\forall x \in [a,b]$. Entonces, para cualquier $p_0 \in [a,b]$ la sucesión definida por $p_i=g(p_{i-1})$ converge al único punto fijo en el intervalo $[a,b]$.

Dem: Por el teorema de existencia y unicidad de punto fijo, existe un único punto fijo p en $[a,b]$. Como la g transforma $[a,b]$ en sí mismo, la sucesión $\{p_n\}$ está bien definida. Veamos la convergencia: $|p_n-p| = |g(p_{n-1})-p| = |g'(t)|\cdot|(p_{n-1}-p)| \leq k\cdot|p_{n-1}-p|$. Luego, $|p_n-p| \leq k\cdot|p_{n-1}-p| \leq k^2\cdot|p_{n-2}-p| \leq ... \leq k^n\cdot|p_0-p|$. Como $0<k<1$, $k->0$ cuando $n->\infty$ por lo que $\lim_{n\to\infty} |p_n-p| \leq \lim_{n\to\infty} k^n\cdot|p_0-p| \Rightarrow \lim_{n\to\infty} |p_n-p|=0$, por lo tanto, $p_n->p$ cuando $n->\infty$.

<u>Análisis del error:</u>

Dado p punto fijo de g y $\{p_n\}$ la sucesión definida por $p_{n+1}=g(p_n)$, consideremos el desarrollo de Taylor de g alrededor de p, sumaremos h a p tal que $p+h=p_n$: $p_{n+1} = g(p_n) = g(p+h) = g(p)+g'(p)\cdot(p_n-p)+...+g^{(r)}(t)\cdot(p_n-p)^r/r!$, con t entre p y p_n. Supongamos que $g'(p)=g''(p)=...=g^{(r-1)}(p)$ y $g^{(r)}\neq 0$, $p_{n-1}-p=g^{(r)}(t)\cdot(p_n-p)^r/r!$, es decir, $e_{n+1}=g^{(r)}(t)/r!\cdot e_n$. Por lo tanto, el método tiene convergencia al menos r.

Obs: El método de newton es un método de punto fijo y si f tiene una raíz simple en p, el método converge cuadráticamente.

<u>Interpolación Polinomial</u>

Se dice que el polinomio p interpola a los puntos dados si $p(x_i)=y_i$.

<u>Teorema de unicidad y existencia del polinomio interpolante:</u>

Si $x_0,...,x_n$ números reales distintos, entonces para valores arbitrarios $y_0,...,y_n$ existe un único polinomio P_n de grado $\leq n$ tal que $P_n(x_i)=y_i$.

Dem: Unicidad: Supongamos que existen dos polinomios de gr $\leq n$ tal que $P_n(x_i)=y_i$ y $Q_n(x_i)=y_i$. Sea $h_n(x)=P_n(x)-Q_n(x)$ es un polinomio de gr $\leq n$, entonces $h_n(x_i)=P_n(x_i)-Q_n(x_i)=0$, h_n es un polinomio de gr $\leq n$ con n+1 raíces, por lo tanto $h(x)=0$. Entonces $0=P_n(x)-Q_n(x) \Rightarrow P_n(x)=Q_n(x)$.

Existencia: sea $L_k(x)=\Pi(x-x_i)/(x_k-x_i)$ el polinomio básico de lagrange de gr $\leq n$. Además, $L_k(x_j)=d_{kj}= 0$ si $k=j$ ó 1 si $k\neq j$. Sea $P_n(x)= 0 \sum_{k=0}^{n} y_k\cdot L_k(x)$ y $P_n(x_i)= \sum_{k=0}^{n} y_k\cdot L_k(x_i)= \sum_{k=0}^{n} y_k\cdot d_{kj}=y_i$.

<u>Error en el polinomio interpolante:</u>

Sean $f,f',...,f^{(n+1)}$ continuas en el intervalo (a,b) y P un polinomio de gr $\leq n$ que interpola a f en n+1 puntos distintos $x_0,...,x_n$ en $[a,b]$. Entonces para cada x_i existe un $t \in (a,b)$ tal que $f(x)-P(x)=f^{(n+1)}(t)\cdot \prod_{i=0}^{n} (x-x_i)\cdot \frac{1}{(n+1)!}$.

Dem: Si $x=x_i$ es trivial. Si $x\neq x_i$, definimos $w(x)=\Pi(x-x_i)$, $i=0,...,n$, un polinomio de grado n+1 y una constante $c= \frac{f(x)-P(x)}{w(x)}$ y una función $q(r)=f(r)-p(r)-c\cdot w(r)$. Notar que $q(x_i)=0$ y en x_1 $q(x) = f(x)-P(x)- \frac{f(x)-P(x)}{w(x)} \cdot w(x) = 0$. Luego q tiene n+2 raíces x, $x_0,..., x_n$, q' tiene al menos n+1 raíces, …, $q^{(n+1)}$ tiene al menos 1 raíz. Sea t esa raíz de $q^{(n+1)}$, así $0 = q^{(n+1)} = f^{(n+1)}(t)-P^{(n+1)}(t)-c\cdot w^{(n+1)}(t)$. Por definición de $w^{(n+1)}(r)=(n+1)! \Rightarrow w^{(n+1)}(t)=(n+1)!$. Finalmente, de la igualdad anterior, sacamos $0 = f^{(n+1)}(t)-c\cdot(n+1)! = f^{(n+1)}(t)- \frac{f(x)-P(x)}{w(x)} \cdot(n+1)!=0$, osea que $f(x)-P(x) = f^{(n+1)}(t)\cdot \frac{1}{(n+1)!} \cdot \prod_{i=0}^{n} (x-x_i)$.

Splines

Una función spline está formada por varios polinomios, preferentemente de grado bajo, que se conectan entre sí obedeciendo ciertas condiciones de continuidad.

Error en spline lineal:

Sea f definida en [a,b], derivable en (a,b) con $a=x_0<...<x_n=b$ igualmente espaciados, $x_k=a+k\cdot h$, con $h=(b-a)/n$. Sea S el spline lineal que interpola a los datos en $x_0,...,x_n$ y sea S_k el polinomio de grado ≤ 1, definido en $[x_k,x_{k+1}]$, el error de interpolación en el intervalo es $|e_k| \leq M_k/8\cdot|x_{k+1}-x_k|^2$. Luego si e(x)=f(x)-S(x), $|e(x)|\leq M/8\cdot h^2$, donde M = max M_k = max f''(x) con $x\epsilon[x_0,x_n]$.

Cuadrados Mínimos

Consideremos el problema de estimar una función para un conjunto de datos (x_i,y_i), con i=0,..., n donde n es un valor grande. Si proponemos un modelo lineal y(t)=a·t+b deberíamos determinar los coeficientes a y b de modo que ajuste lo mejor posible a los datos, es decir, que el error sea mínimo. Como n es grande, podría no ser conveniente hacer una interpolación polinomial de grado alto y tampoco podría ser bueno hacer un spline ya que obtendriamos una función muy zigzagueante.

Pedir que el error sea 0 sería casi imposible por lo que vamos a considerar el error $E(a,b)=\Sigma(y_i-a\cdot t_i-b)^2$, con i = 1,...,n, Determinar a y b tal que E(a,b) sea mínimo. Para que exista un mínimo en dos variables debe ocurrir que $\partial E(a,b)/\partial a=0$ y $\partial E(a,b)/\partial b=0$, lo cual debe suceder simultáneamente.

Definición: Una función integrable w se llama función de pso en el intervalo I si $w\geq 0$ $\forall x\in I$, pero $w(x)\neq 0$ en cualquier subintervalo de I. El objetivo de estas funciones es darle más o menos relevancia a la aproximación en ciertas partes del intervalo I.

Definición: se dice que $\{f_0,...,f_n\}$ es un conjunto ortogonal de funciones en [a,b] de w si $\int_a^b w(x)\cdot f_k(x)\cdot f_j(x)dx$ = 0 si j=k ó $\alpha_j>0$ si $j\neq k$. Si $\alpha_j=0$, el conjunto es ortonormal.

Integración Numérica

El teorema de valor medio para integrales dice que suponiendo que f es continua en el [a,b], g integrable en [a,b] y no cambia signo en el mismo, entonces existe $c\epsilon(a,b)$ tal que $\int_a^b f(x)\cdot g(x)dx = f(c)\cdot\int_a^b g(x)dx$. En particular si g(x)=1, $\int_a^b f(x)dx = f(c)\cdot(b-a)$.

La idea es $\int_a^b f(x)dx = \Sigma a_i\cdot f(x_i)$, donde $x_i\in[a,b]$ con i=0,...,n. Sean $x_0,...,x_n\in[a,b]$ usando el polinomio interpolante de lagrange $P_n(x) = \Sigma L_k(x)\cdot f(x_k)$ con un error definido en e_n. Luego $\int_a^b f(x)dx = \int_a^b P_n(x)dx + \int_a^b e_n(x)$.

Definición: La precisión o grado de exactitud de una fórmula de cuadratura es el entero positivo más grande n tal que la fórmula sea exacta para x^k, con k=0,...,n.

Regla del trapecio

Se llama así porque cuando f es una función con valores positivos, aproximamos $\int_a^b f(x)dx$ por el área de un trapecio. Sean x_0=a y x_1=b con h=b-a. Consideremos el polinomio interpolante de Lagrange de grado ≤ 1: $P_1(x) = L_0\cdot f(x_0)+L_1\cdot f(x_1)$, luego $\int_a^b f(x)dx = \int_a^b(L_0\cdot f(x_0)+L_1\cdot f(x_1))dx+1/2\cdot\int_{x0}^{x1}f''(t)\cdot(x-x_0)\cdot(x-x_1)dx$. Como $(x-x_0)\cdot(x-x_1)$ no cambia de signo en $[x_0,x_1]$ aplicamos el TVM para integrales al término del error y queda: $\int_{x0}^{x1}f''(t)\cdot(x-x_0)\cdot(x-x_1)dx = f''(t)\cdot\int_{x0}^{x1}(x-x_0)\cdot(x-x_1)dx = f''(t)\cdot\int_{x0}^{x1}(x^2-(x_0+x_1)\cdot x-x_0\cdot x_1)dx = f''(t)\cdot(x^3/3-(x_0+x_1)\cdot x^2/2+x_0\cdot x_1\cdot x = (x_0-x_1)^3/6 = -h^3/6\cdot f''(t)$.

Luego $\int_a^b f(x)dx = h/2\cdot(f(a)+f(b))-h^3/12\cdot f''(t)$.

Error: Como el término del error tiene f'', la regla es exacta cuando se aplica a funciones con f'' nula, como los polinomios de grado ≤ 1.

Regla de Simpson
Se obtiene al integrar en [a,b] el polinomio interpolante de Lagrange de grado ≤ 2 con 3 nodos: $x_0=a$, $x_1=(a+b)/2$, $x_2=b$.
Error: Como el término del error tiene $f^{(4)}$, la regla es exacta para aquellas funciones con $f^{(4)}$ nula, como los polinomios de grado ≤ 3.

Regla compuesta de simpson
Consideremos n par y subdividimos el intervalo [a,b] en n subintervalos y aplicamos la regla de simpson en cada subintervalos. El error es $h^5/180\cdot n\cdot f^{(4)}(m)$, donde $h=(b-a)/n$.

Reglas de cuadratura Gaussianas
Si en lugar de construir la regla de Simpson o la regla del Trapecio con los extremos tomamos otros puntos elegidos convenientemente, podría aumentar la precisión. Las reglas gaussianas seleccionan los puntos de evaluación de f de manera óptima, de modo de minimizar el error, aumentando así su precisión.
La selección óptima será tal que la regla sea exacta para el polinomio de grado más alto posible. En este tipo de reglas, se deben determinar $x_1,...,x_n \in [a,b]$ y $c_1,...,c_n$. En total, se deben determinar 2n parámetros, por lo tanto el polinomio tendrá grado máximo 2n-1 y esta será la precisión de la regla gaussiana basada en n puntos.

Resolución de problemas lineales

Eliminación Gaussiana
La idea es transformar [A|B] en otra matriz [U|b], donde U es triangular superior mediante operaciones elementales por fila.
Factorización LU
Sirve para resolver varios sistemas lineales de la forma $Ax=b^i$, cambiando los b^i. La idea es factorizar $A=L\cdot U$, con L triangular inferior y U triangular superior. Luego resolver dos sistemas fáciles $Ly=b$ y $Ux=y$.

Teorema: Sea $A\in R^{nxn}$ con submatrices menores principales $A_k=A(1:k,1:k)$ no singulares, con k=1,...,n-1. Entonces existen únicas matrices L y U $\in R^{nxn}$ donde L es triangular inferior con $L_{ii}=1$ y U triangular superior tal que $L\cdot U=A$. Además, $\det(A)=U_{11}\cdot...\cdot U_{nn}$.
Dem: Si n=1, $A=[a_{11}]$, $L=[1]$ y $U=[a_{11}]$, por lo tanto $L\cdot U=A$. Supongamos que existe la factorización LU hasta orden k-1 y veamos que funciona para k. COMPLETAR.

Métodos iterativos
La idea es generar una sucesión de vectores $\{x^{(k)}\}$, a partir de $x^{(0)}\in R^n$ y se espera que bajo ciertas hipótesis la sucesión tienda a x', donde x' es la solución de Ax=b. Son adecuados para sistemas lineales grandes y A con muchos ceros.

Si tengo Ax=b, escribimos A como M-N, con M no singular y queda $(M-N)x=b \Rightarrow Mx=Nx+b \Rightarrow x=M^{-1}\cdot N\cdot x+M^{-1}\cdot b$. Esto sugiere un método iterativo donde $x^{(k)}=M^{-1}\cdot N\cdot x^{(k-1)}+M^{-1}\cdot b$

Jacobi: M=D y N=-L-U
Gauss-Seidel: M=L+D y N=-U

Teorema: si A es diagonalmente dominante en sentido estricto, entonces la sucesión generada por el método iterativo converge, para cualquier vector inicial, a la solución de Ax=b.

Programación Lineal

Definición: un hiperplano es un espacio n-dimensional que está dado por $\{(x_1,...,x_n \in R^n)|a_1\cdot x_1+...+a_n\cdot x_n\leq b\}$.

<u>Definición</u>: Decimos que es un semiespacio cerrado si contiene al hiperplano si no, decimos que es un semiespacio abierto.
<u>Definición</u>: Dados A y B puntos distintos en R^n, el segmento de extremos A y B es el conjunto $\overline{AB}$ $=\{(1-t)\cdot A+t\cdot B, 0 \leq t \leq 1\}$.
<u>Definición</u>: Un subconjunto $S \in R^n$ se dice convexo si para todo par de puntos A y B en S el segmento AB está en S.