TEORIA DE ERRORES.

INTRODUCCION.

Los datos de entrada en un proceso de cálculo muy rara vez son exactos ya que con frecuencia se basan en experimentos o son estimaciones y a su vez los procesos numéricos introducen errores de varios tipos haciéndose necesario el análisis del error en un resultado obtenido de un cálculo computacional o manual. En esta unidad aprenderás a practicar dicho análisis y tener un estimativo del error total cometido en cualquier proceso de cálculo.

OBJETIVOS.

Al terminar de estudiar la presente unidad debes ser capaz de:

- Saber en que consisten los diferentes métodos para estimación de los errores y poder hacer una descripción de cada uno de ellos.
- Identificar en un proceso de cálculo las fuentes de error.
- Distinguir los diferentes tipos de error de acuerdo a sus características para estar en la posibilidad de saber disminuir el error total.
- Saber redondear en forma simétrica y redondeo-truncado.
- Entender y explicar la expresión del error relativo.
- Determinar los errores absolutos y relativos por las reglas de "Redondeo Truncado" y "Redondeo Simétrico".
- Determinar el error máximo relativo en los diferentes redondeos y poder hacer comparaciones entre las estimaciones de los errores por redondeo.
- Conocer la propagación del error en algún punto del proceso de cá culo.
- Manejar las expresiones del error absoluto para las cuatro operaciones fundamentales.
- Manejar las expresiones del error relativo en las cuatro operaciones

($\oplus$, $\ominus$, $\odot$ y $\oslash$)

- Saber en qué consiste una gráfica de proceso y cuál es su utilidad.
- Elaborar gráficas de proceso para un proceso de cálculo particular.
- Calcular los errores con ayuda de las gráficas de proceso.
- Poder hacer el análisis practicado en el anexo para los casos en que tengan más de cuatro decimales.
- Aplicar esta teoría a los métodos numéricos (ec. simultáneas, Simpson, - mínimos cuadrados, etc.).

PROCEDIMIENTO DE ESTUDIO.

Estudiar fundamentalmente el anexo de la unidad II y consultar la bibliografía - recomendada para el curso.

EXAMEN DE AUTOEVALUACION.

Hacer los siguientes ejercicios.

1. La corriente pasa a través de una resistencia de 10 Ω (ohms) cuya preci-- sión esta dentro del 10%. Siendo medida la corriente con una aproximación de 0.1 amp. y su valor de 2 amperios. La ley de Ohm ($V = I \times \Omega$).

   ¿Cuáles son los errores absolutos y relativo en el voltaje calculado?

2. Supón α es un número positivo propiamente redondeado, y que el número 2 puede representarse exactamente en una computadora. Dibuja las gráfi-- cas de proceso y determina los límites en los errores relativos máximos - para demostrar que son iguales para $u = \alpha + \alpha$ y para $v = 2\alpha$.

3. Con las mismas suposiciones del ejercicio 2, demuestra que el límite en - el máximo error relativo para $u = \alpha + \alpha + \alpha$ es mayor que para - $v = 3\alpha$ ilustra esto con $\alpha = 0.6992$, conservando sólo cuatro dígitos - después de cada operación aritmética.

4. Considera la expresión 5a + b. Demuestra que en el resultado, el error re_ lativo inherente en a influye cinco veces más que el error relativo inheren- te en b.

   Puedes presentar el examen Evaluación de la unidad. Si puedes resolver - sin ayuda estos ejercicios.

Es frecuente que en este método se suponga la solución verdadera cerca de la mitad del intervalo lo cual no es válido generalmente.

Desventajas :

Requiere más del Doble de tiempo de operación de la computadora y cerca del - Doble de almacenamiento de una operación normal.

Aritmética de Dígitos Significativos.

Este método intenta no perder de vista los dígitos significativos que se pierden - al hacer operaciones en la máquina y al final del cálculo es necesario asegurarse que todos los dígitos retenidos son significativos. Es usual descartar dígitos que se piensa que no son significativos.

Desventajas :

- Se pierde información cuando se descartan dígitos.
- Los resultados obtenidos tienden a ser muy conservativos.
- Que este método aún está en marcha y la experiencia hasta ahora no es muy prometedora.

Enfoque Estadístico. En este método se adopta un modelo estocástico de la propagación del error de redondeo en el cual los errores locales se tratan como si fueran variables aleatorias y se supone que los errores locales de redondeo están uniformemente o normalmente distribuidos entre sus valores extremos. Usando la estadística se puede obtener la desviación estándar, la variancia, y estimativos del error de redondeo acumulado.

Este método implica un análisis detallado y tiempo adicional de computador, pero que ha dado buenos estimativos del error.

A continuación se muestran algunos lineamientos prácticos y funcionales para determinar la propagación de errores y estimación de errores ó un límite al tamaño máximo del error.

TEORIA DE ERRORES.

Errores de aritmética de punto flotante.

Cuando se usa aritmética de punto flotante por ejemplo de cuatro dígitos para resolver la ecuación $x^2 + 0.4002x + 0.0008 = 0$ utilizando la fórmula:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

para $x_1 = -0.00015$ En aritmética de punto flotante de cuatro dígitos.

Se introduce errores

que hace el resultado erróneo en 25%

$$\frac{0.0002 - 0.00015}{0.0002} = \frac{0.00005}{0.0002} = \frac{0.5}{2} = 0.25 \quad 25\%$$

La raíz real, determinada con aritmética de punto flotante de 8 dígitos es $x_1 = -0.0002$.

Pero no todos los errores se resuelven usando aritmética de punto flotante de 8 dígitos.

Considerando por Ej. la serie de taylor para

$$\text{Sen } x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots.$$

En problemas de ingeniería, usualmente se describe como válida para un número de términos finito.

Pero sabemos que cometemos un error, llamado error por truncamiento y se dice: Que es Menor en valor ABSOLUTO que el primer término despreciado.

Por todo lo anterior hemos de ver clasificación de errores

- Errores Relativos
- Errores Absolutos

El error absoluto en una cantidad, es la Diferencia entre el VALOR VERDADERO (suponiendo que se conoce) y una aproximación al valor verdadero.

así Si x = cantidad verdadera

$\bar{x}$ = una aproximación a la cantidad verdadera

e_x = error absoluto

tenemos que $x = \bar{x} + e_x \therefore e_x = x - \bar{x}$

Error Relativo. Es el cociente del error absoluto entre la aproximación

$$\frac{e_x}{\bar{x}} = \frac{x - \bar{x}}{\bar{x}}$$

todo lo que obtendremos será una estimación del Error o un límite al tamaño Máximo del error. Puesto que aunque sería razonable definirlo como el error absoluto dividido por el valor verdadero, pero generalmente no se conoce.

En una computación numérica se tienen tres tipos básicos de errores que son:

1. INHERENTES { errores en los datos }
2. POR TRUNCAMIENTO { errores debidos a la manera de efectuar los procesos numéricos }
3. POR REDONDEO { errores debidos a la manera de efectuar los procesos numéricos }

Los inherentes son debidos a la incertidumbre en las mediciones.

1) Ej. Medición de un voltaje de 6.3582753 al menos algunos de los dígitos de la derecha no tienen sentido. Por la naturaleza misma de las cantidades a representar.

2) Ej. El valor de 3.14, 3.14159265 ó 3.141592653589793 en ninguno de estos casos se tiene una representación exacta.

3) Ej. $\frac{1}{3}$ (muchas fracciones que tienen representación finita en un sistema no la tienen en otro. Ej. $\frac{1}{10} = 0.10$ (decimal) en binario 0.000110011001100...)

Por truncamiento se deben a la omisión de términos en una serie que tienen un número infinito de términos

Ej. $$\text{Sen}\, x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots$$

Por redondeo. Estos errores se introducen en los procesos de computación por el hecho de que las computadoras trabajan con un número finito de dígitos después del punto decimal y tienen que redondear.

Ej. $A = 35.78234$ y $B = 25.38345$

Si estas cantidades fueran exactas y la suma es 61.16579. Si utilizamos una máquina que trabaja con cuatro dígitos después del punto tendríamos que almacenar 35.7823 y 25.3835 siendo entonces la suma 61.1658.

En una computadora en que la mantisa tiene cuatro dígitos y el exponente tiene un dígito se hace el desarrollo siguiente:

Recordando que cada número puede representarse por una mantisa afectada por una potencia del número base, así tenemos los números de punto flotante.

$0.3864 \cdot 10^4 = 3846.$

$0.9338 \cdot 10^2 = 93.38$

$0.1872 \cdot 10^{-3} = 0.0001872$

*Se dice que un número de punto flotante está normalizado, cuando el primer Dígito de la MANTISA es $\neq$ de cero.

Hagamos de cuenta que todos los números que utilizaremos en esta parte de la explicación han sido NORMALIZADOS.

Si f representa la mantisa de un número de punto flotante x y e el exponente podemos expresar en forma general un número de punto flotante en Base Decimal como:

$$x = f \cdot 10^e$$

Para valores próximos a 1 el

$$\text{Error Relativo} \approx \text{Error Absoluto}$$

pero para valores no próximos a 1 puede existir una gran diferencia, por tal motivo es necesario indicar en cada caso a que tipo de error nos referimos.

Sabemos que el valor de f no puede ser menor que $\frac{1}{10}$ puesto que los números han sido normalizados y no puede llegar a ser 1, porque la mantisa es una fracción propia.

Ahora si realizamos la suma de

$0.1348 \times 10^3 = 134.8$ (con Mantisa de 4 dígitos y un dígito

$0.1571 \times 10^1 = 1.571$ como exponente).

con la máquina, esta se encarga de la colocación del punto y compara los exponentes para desplazar hacia la derecha el punto para alinear.

Así en el ejemplo hace lo siguiente:

$$\begin{array}{ll} 0.1348 \times 10^3 & (3 - 1 = 2 \text{ Dos lugares recorrer}) \\ 0.001571 \times 10^3 & \\ \hline 0.136371 \times 10^3 & \text{Antes de que la máquina haga la operación de redondeo.} \end{array}$$

(La mantisa de la suma tiene más de cuatro dígitos)

El resultado puede mostrarse como dos cantidades de punto flotante:

$$0.136371 \times 10^3 = 0.1363 \times 10^3 + 0.7100 \times 10^{-1}$$

Cualquier resultado proveniente de la realización de las cuatro operaciones aritméticas puede indicarse, antes de ser redondeado por la forma general:

$$y = f_y \cdot 10^{e} + g_y \cdot 10^{e-t}$$

El rango de variación de f_y por * $\quad \frac{1}{10} \leq |f_y| < 1$

El rango de variación de g_y $\quad 0 \leq |g_y| < 1$

g_y no se puede garantizar que este normalizada y tan es así que g_y puede ser cero.

En esta ecuación t es el número de dígitos que tiene f_y.

El error Relativo Máximo ocurre cuando g_y es grande y f_y es pequeño.

El valor máximo posible de g_y es < 1.0 y el valor mínimo de f_y es 0.1 por lo que el valor absoluto del valor Relativo es

$$\left| \frac{e_y}{\bar{y}} \right| = \left| \frac{g_y \cdot 10^{\ell - t}}{f_y \cdot 10^{\ell}} \right| \leq \frac{1 \cdot 10^{\ell - t}}{0.1 \cdot 10^{\ell}} = 10^{-t+1}$$

cálculo del error relativo

error relativo no máximo

Sabemos que t es el número de Dígitos en la Mantisa de cualquier número de punto flotante.

Obteniéndose un resultado interesante:

El Máximo error relativo por redondeo en el Resultado de una operación aritmética de punto flotante NO DEPENDE DEL TAMAÑO DE LAS CANTIDADES.

El redondeo más conocido es el redondeo simétrico y puede describirse la aproximación redondeada a y

como:
$$\left| \bar{y} \right| = \begin{cases} \left| f_y \right| \cdot 10^{\ell} & \text{si} \quad \left| g_y \right| < \frac{1}{2} \\ \left| f_y \right| \cdot 10^{\ell} + 10^{\ell - t} & \text{si} \quad \left| g_y \right| \geq \frac{1}{2} \end{cases}$$

$\bar{y}$ tiene el mismo signo que f_y

El segundo término (10^{e-t}) equivale a sumar 1 al último dígito retenido si el primer dígito que se pierde es igual o mayor que 5.

Nota: Las mismas fórmulas se aplican a cantidades positivas y negativas.

cuando $g_y < \frac{1}{2}$ el error absoluto será:

$$\left| e_y \right| = \left| g_y \right| \cdot 10^{\ell - t}$$

si $g_y \geq \frac{1}{2}$ el error absoluto será:

$$\left| e_y \right| = \left| 1 - g_y \right| \cdot 10^{\ell - t}$$

el factor $\left| 1 - g_y \right|$ no es mayor que 1/2 y el error absoluto es $\left| e_y \right| \leq \frac{1}{2} \cdot 10^{\ell - t}$

error simétrico

y el valor absoluto del error relativo es entonces

$$\left| \frac{e_y}{\bar{y}} \right| \leq \left| \frac{1/2 \cdot 10^{\ell - t}}{f_y \cdot 10^{\ell}} \right| \leq \left| \frac{1/2 \cdot 10^{\ell - t}}{0.1 \cdot 10^{\ell}} \right|$$

$$= 5 \cdot 10^{-t} = \frac{1}{2} \cdot 10^{-t+1}$$

"Redondear" implica afectar de algún modo a f_y, dependiendo de g_y, entonces - surgen las preguntas ¿cómo se toma a g_y para modificar a f_y? ¿y para cada - caso cuál es el error máximo que resulta en $\bar{y}$?, ya que puede omitirse g_y lo - que significa que f_y nunca se modifica y llamaremos "redondeo truncado" con la intención de evitar cualquier confusión con el error por truncamiento que se comete al considerar sólo una parte de un proceso infinito.

Si como resultado después de que la computadora ha realizado una serie de operaciones aritméticas se obtuvo la siguiente cantidad:

$$y = 0.3873939 \cdot 10^{6} + 0.9838 \cdot 10^{-1}$$

Determinar los errores absolutos y relativos por las reglas de:

"Redondeo truncado"
"Redondeo simétrico"
y los valores de $\bar{y}$

"Redondeo Truncado"

En esta regla no existe modificación alguna de f_y ocasionada por g_y simplemente g_y desaparece.

Valor aproximado y error absoluto

$$\bar{y} = 0.3873939 \cdot 10^{6}$$

$$e_y = 0.9838 \cdot 10^{-1}$$

Error Relativo:

$$\left| \frac{e_y}{\bar{y}} \right| = \frac{.0.9838 \cdot 10^{-1}}{0.3873939 \cdot 10^{6}} = 2.539 \cdot 10^{-7}$$

comprobando si es correcto.

Este error debe ser menor que el error relativo máximo (el límite)

$$\left| \frac{e_y}{y} \right| = 10^{-t+1} = 10^{-7+1} = 10^{-6}$$

$$1 \cdot 10^{-6} > 2.539 \cdot 10^{-7} \qquad \text{OK.}$$

Redondeo Simétrico.

Para este error se aplica:

$$\left| \bar{y} \right| = \left| f_y \right| \cdot 10^{\ell} + 10^{\ell - t}$$

siendo el caso de:

$$\left| g_y \right| > \frac{1}{2}$$

$$\left| \bar{y} \right| = 0.3873939 \cdot 10^{6} + 10^{6-7} = 0.3873939 \cdot 10^{6} + 10^{-1}$$

$$\left| \bar{y} \right| = 38739.9 + 0.1 = 387394.0 = 3.873940 \times 10^{6}$$

387393.9

error absoluto para $g_y \geq \frac{1}{2}$

$$\left| e_y \right| = \left| 1 - g_y \right| \cdot 10^{\ell - t} = \left| 1 - 0.9838 \right| 10^{-1}$$

$$= \left| 0.0162 \right| \cdot 10^{-1}$$

$$= 0.00162$$

Este error tiene que ser menor que el máximo

$$\left| e_y \right| \leq \frac{1}{2} \; 10^{e-t} = 0.5 \cdot 10^{6-7} = 0.05$$

correcto

Error relativo :

$$\left| \frac{e_y}{\bar{y}} \right| = \frac{0.00162}{3.873940 \times 10^5} \cdot \approx 4.181789 \times 10^{-9}$$

el máximo error relativo

$$\frac{e_y}{\bar{y}} = 5 \times 10^{-7} \quad 1.539 \times 10^{-7}$$ y esto es mayor que 4.181789×10^{-9} OK.

como puede observarse el error "Redondeo truncado" es mayor que el de Redondeo simétrico.

El error truncado no es tan grande como su límite superior correspondiente - $(10 \cdot 10^{-4})$, como tampoco lo es para el simétrico $(5 \cdot 10^{-4})$ y puede darse el caso que el error por redondeo sea cero. Concretamente, se conocerá un límite - al tamaño del error en un cálculo, pero el error total no lo conoceremos. Po- - niéndonos del lado de la seguridad, se dirá que el error puede ser tan grande como su límite :

Como puede notarse, que los resultados se han establecido en términos de cantidades flotantes decimales.

Para efectuar cálculos científicos muchas computadoras trabajan en sistema flotante binario en lugar del sistema en base 10. En binario, cada número de punto flotante se representa por :

$$\bar{x} = f \cdot 2^{\ell}$$

$$\frac{1}{2} \leq |f| < 1$$

Un análisis, en forma semejante al efectuado anteriormente conduce a un límite en el error relativo de $2 \cdot 2^{t}$ para el error truncado y 2^{-t} para el simétrico.

Otras computadoras son hexadecimales (trabajan con números de base 16) y los límites son :

Error Relativo :	$16 \cdot 16^{-t}$	redondeo truncado
	$8 \cdot 16^{-t}$	redondeo simétrico

Aunque los resultados son diferentes en distintos sistemas. Se ilustra el error por redondeo en el Sistema que estamos más familiarizados (base 10).

En una materia como Métodos Numéricos o en un análisis numérico en que se usa la computadora, es muy importante conocer la propagación del error en algún punto del proceso de cálculo. Y dado que los errores están de alguna manera relacionados con las cantidades y las operaciones que se hacen con ellas, es necesario conocer o encontrar las expresiones para las cuatro operaciones fundamentales, tanto para el Error Absoluto como para el Error Relativo en función de los operandos y sus errores.

> Si x y y son dos valores verdaderos y únicamente conocemos sus aproximaciones $\bar{x}$ y $\bar{y}$. Siendo los errores para cada uno de los valores e_x y e_y.

se tienen: EXPRESIONES PARA EL ERROR ABSOLUTO.

SUMA: $x + y = \bar{x} + e_x + \bar{y} + e_y$

el error en la suma

es: $e_{x+y} = e_x + e_y$

RESTA: En forma semejante obtenemos para la resta:

$$x - y = (\bar{x} + e_x) - (\bar{y} + e_y)$$

$$x - y = \bar{x} + e_x - \bar{y} - e_y$$

$$x - y = \bar{x} - \bar{y} + e_x - e_y$$

El error en la resta, e_{x-y}

es: $e_{x-y} = e_x - e_y$

MULTIPLICACION:

$$x \cdot y = (\bar{x} + e_x)(\bar{y} + e_y)$$

$$x \cdot y = \bar{x}\bar{y} + e_x\bar{y} + \bar{x}e_y + e_x e_y$$

$\bar{y}e_x$

suponiendo que los errores son considerablemente más pequeños que las aproximaciones, e ignorando el producto de los errores. Se tiene:

$$x \cdot y \cong \bar{x}\bar{y} + \bar{x}e_y + \bar{y}e_x$$

El error en la multiplicación $e_{x \cdot y}$

es: $e_{x \cdot y} = \bar{x}e_y + \bar{y}e_x$

DIVISION: $\frac{x}{y} = \frac{\bar{x} + e_x}{\bar{y} + e_y}$ Multiplicando el Denominador por $\frac{\bar{y}}{\bar{y}}$ y agrupando

tenemos: $\frac{x}{y} = \frac{\bar{x} + e_x}{\bar{y}\,(1 + \frac{e_y}{\bar{y}})} = \frac{\bar{x} + e_x}{\bar{y}} \left(\frac{1}{1 + \frac{e_y}{\bar{y}}}\right)$

Desarrollando en serie el factor del paréntesis por una división.

$$\frac{x}{y} = \frac{\bar{x} + e_x}{\bar{y}} \left[1 - \frac{e_y}{\bar{y}} + \left(\frac{e_y}{\bar{y}}\right)^2 - \ldots\ldots\right]$$

Efectuando la multiplicación:

$$\frac{x}{y} = \frac{\bar{x} + e_x}{\bar{y}} - \left[\frac{e_y}{\bar{y}}\left(\frac{\bar{x} + e_x}{\bar{y}}\right)\right] + \left(\frac{e_y}{\bar{y}}\right)^2 \left(\frac{\bar{x} + e_x}{\bar{y}}\right) - \ldots\ldots$$

$$= \frac{\bar{x}}{\bar{y}} + \frac{e_x}{\bar{y}} - \left(\frac{\bar{x}e_y + e_x e_y}{\bar{y}^2}\right) + \frac{\bar{x}e_y^2 + e_x^2 e_y}{\bar{y}^3}$$

$$\frac{x}{y} = \frac{\bar{x}}{\bar{y}} + \frac{e_x}{\bar{y}} - \frac{\bar{x}e_y}{\bar{y}^2} - \underbrace{\frac{e_x e_y}{\bar{y}^2} + \frac{\bar{x}e_y^2}{\bar{y}^3} + \frac{e_x^2 e_y}{\bar{y}^3} - \ldots\ldots}_{\text{Despreciando}}$$

$$\frac{x}{y} \cong \frac{\bar{x}}{\bar{y}} + \frac{e_x}{\bar{y}} - \frac{\bar{x}e_y}{\bar{y}^2}$$ por lo que

El error en la División $e_{x/y}$ está dado por la expresión siguiente :

$$e_{x/y} \cong \frac{1}{\bar{y}} e_x - \frac{\bar{x}}{\bar{y}^2} e_y$$

EXPRESIONES PARA EL ERROR RELATIVO.

Teniendo las expresiones para la propagación de los errores absolutos en las cuatro operaciones las expresiones para el Error Relativo en cada operación son :

SUMA. Sabemos que $e_{x+y} = e_x + e_y$

El error relativo $$\frac{e_{x+y}}{\bar{x} + \bar{y}} = \frac{e_x}{\bar{x} + \bar{y}} + \frac{e_y}{\bar{x} + \bar{y}}$$

Multiplicando por $\frac{\bar{x}}{\bar{x}}$ y por $\frac{\bar{y}}{\bar{y}}$ cada término del 2º miembro respectivamente y agrupando se tiene la expresión del error relativo para la suma.

$$\frac{e_{x+y}}{\bar{x} + \bar{y}} = \frac{\bar{x}}{\bar{x} + \bar{y}} \left(\frac{e_x}{\bar{x}}\right) + \frac{\bar{y}}{\bar{x} + \bar{y}} \left(\frac{e_y}{\bar{y}}\right)$$

RESTA.

$$\frac{e_{x-y}}{\bar{x} - \bar{y}} = \frac{\bar{x}}{\bar{x} - \bar{y}} \left(\frac{e_x}{\bar{x}}\right) - \frac{\bar{y}}{\bar{x} + \bar{y}} \left(\frac{e_y}{y}\right)$$

MULTIPLICACION.

$$e_{x \cdot y} = \bar{x} e_y + \bar{y} e_x$$

$$\frac{e_{x \cdot y}}{\bar{x} \cdot \bar{y}} = \frac{e_x}{\bar{x}} + \frac{e_y}{\bar{y}}$$

DIVISION.

$$\frac{e_{x/y}}{\bar{x}/\bar{y}} = \frac{e_x}{\bar{x}} - \frac{e_y}{\bar{y}}$$

Para entender claramente las expresiones de propagación del error considerar los siguientes puntos :

- Se parte de dos valores aproximados $\bar{x}$ y $\bar{y}$ que contienen los errores e_x y e_y.
- Los errores pueden ser de cualquier tipo.
- Los valores aproximados pueden provenir de un experimento, de un cálculo realizado previamente, etc. y que llevan errores inherentes ya sea por redondeo o truncamiento.
- Y que esas expresiones dan el error en el resultado de cada una de las operaciones en función de $\bar{x}$, $\bar{y}$, e_x y e_y SUPONIENDO QUE NO HAY ERROR POR REDONDEO en la operación. Si queremos saber como se propaga el error en este resultado a otras operaciones, se debe adicionar explícitamente el ERROR POR REDONDEO.

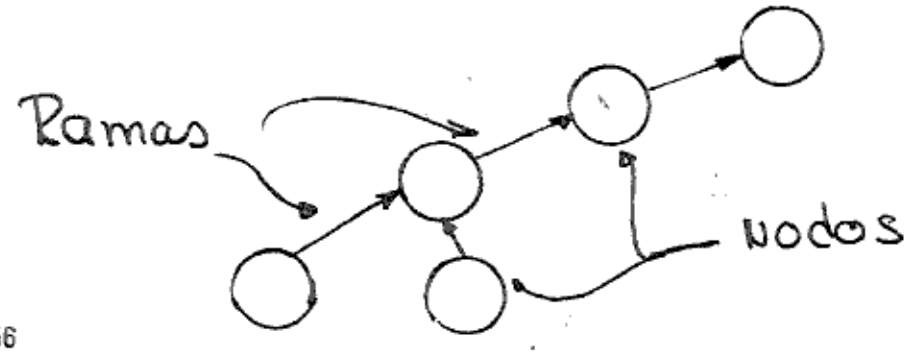
Nota :

Generalmente no conocemos el signo de un error; por lo que no debemos inferir que la suma siempre incrementa el error y que la resta siempre lo disminuye simplemente porque los errores se suman en la adición y se restan en la sustracción. Ya que si los errores tienen signos diferentes sucederá precisamente lo contrario.

GRAFICA DE PROCESOS.

Para determinar el error total al final de un proceso de cálculo es de gran utilidad las denominadas gráficas de proceso, que son una representación de la secuencia en que se efectúan las operaciones en una computación e indican la manera en que se propagan los errores y ayudan a determinar la contribución al error total de un error en cualquier paso intermedio del proceso.

La estructura de estas gráficas es de acuerdo a la siguiente figura :



Para construir y utilizar las gráficas se convendrá en lo siguiente :

1. Los nodos denotan las variables.

2. Las ramas ó flechas denotarán una ganancia y llevarán alguna identificación ó etiqueta.

3. La manera de leer las gráficas se hará de abajo hacia arriba en el sentido de las flechas.

4. Primero han de efectuarse todas las operaciones en un nivel horizontal dado y posteriormente todas las operaciones del nivel superior siguiente y así sucesivamente.

5. Los operadores de suma, resta, etc. se indicarán en el interior de los nodos.

Así la gráfica del proceso $u = (x + y) \cdot z$ es :

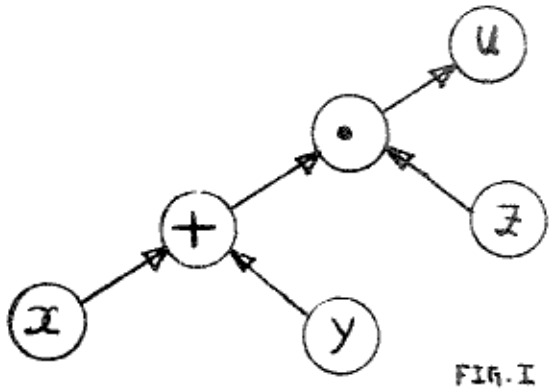


FIG. I

Para calcular el error relativo en cada etapa, adicionamos identificaciones a cada una de las flechas.

SUMA

Suponiendo que las dos flechas que conducen a un nodo de adición (nodo + ó ⊕) provienen de dos nodos cuyos resultados son X y Y (pueden ser ó datos de entrada resultados de otras operaciones) como puede observarse en la siguiente gráfica :

La flecha que va de x a ⊕ se identifica con la etiqueta:

$$\frac{x}{\bar{x} + \bar{y}}$$

y la otra con

$$\frac{x}{\bar{x} + \bar{y}}$$

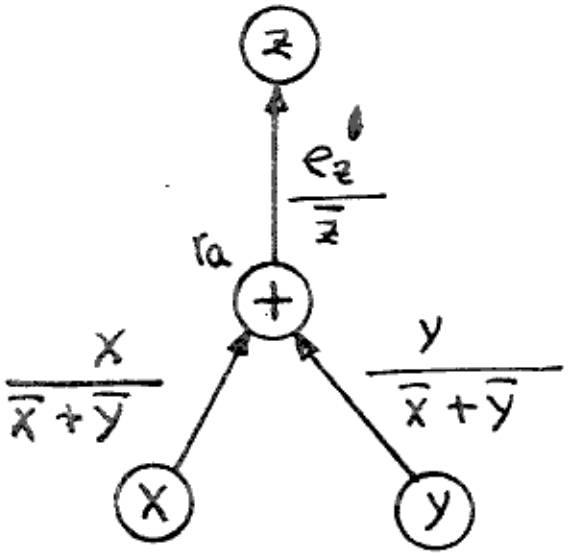


Fig. II

RESTA

Si la operación es x − y se tiene la gráfica siguiente:

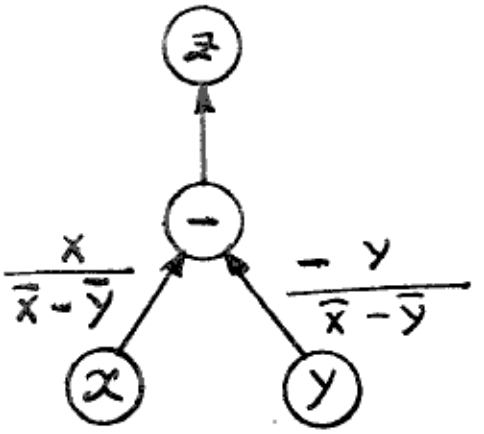


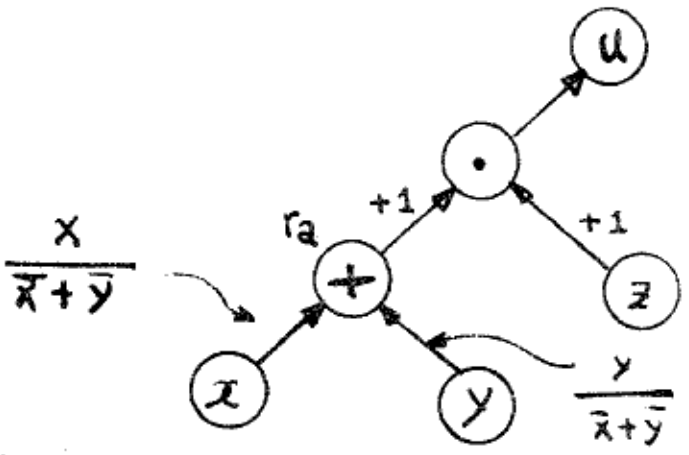
Fig. III

Siendo el Nodo de la resta ⊖

$$z = x - y$$

MULTIPLICACION

Las ramas que convergen a un nodo de multiplicación ⊙ tendrán la etiqueta +1. así para el proceso de la Fig. I. Se tiene:



$u = (x + y) \cdot z$

Fig. IV

DIVISION

Si la operación entre x y y es el cociente para el proceso $u = (\frac{x}{y})z$ la gráfica ha de ser:

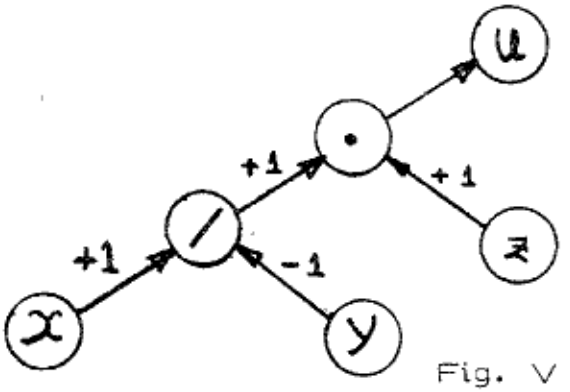


Fig. V

Habiendo convenido en lo anterior se enuncia la Regla siguiente:

"El error relativo es la suma de tres términos, dos de estos términos son los valores del error en las ramas que entran al nodo multiplicadas cada una de ellas por la etiqueta correspondiente indicada en las ramas. El tercer término es el error de redondeo creado por la operación en el nodo mismo".

Ejemplo 1.

Si queremos conocer el valor del error en la rama que sale del nodo suma en la Fig. II, es necesario calcular los factores

$$\left(\frac{e_x}{\bar{x}}\right) \frac{\bar{x}}{\bar{x} + \bar{y}}$$

esto es : el producto del error relativo en x (valor del error en la rama que sale del nodo x) por la etiqueta de dicha rama.

$$\left(\frac{e_y}{\bar{y}}\right) \frac{\bar{y}}{\bar{x} + \bar{y}}$$

y r_a el error por redondeo en la suma.

Así que el valor del error relativo al salir del nodo de suma se calcula por la expresión que sigue :

$$\frac{e_{x+y}}{\bar{x}+\bar{y}} = \left(\frac{e_x}{\bar{x}}\right) \frac{\bar{x}}{\bar{x} + \bar{y}} + \left(\frac{e_y}{\bar{y}}\right) \frac{\bar{y}}{\bar{x} + \bar{y}} + r_a$$

Ejemplo 2.

Aplicando la regla a la multiplicación para el proceso de la Fig. IV.
$(u = (x + y) \cdot z)$

$$\frac{e_u}{\bar{u}} = \left(\frac{e_x}{\bar{x}}\right) \frac{\bar{x}}{\bar{x} + \bar{y}} + \left(\frac{e_y}{\bar{y}}\right) \frac{\bar{y}}{\bar{x} + \bar{y}} + r_a + \frac{e_z}{\bar{z}} + r_m$$

en donde r_m es el error de redondeo en la multiplicación.

Si todos los resultados están correctamente redondeados (de acuerdo con el método convenido) ninguno de los errores por redondeo será mayor que $5 \cdot 10^t$.

Ejemplo 3.

Suma de cuatro números x_i , $i = 1, \ldots\ldots, 4$ tal que $x = x_1 + x_2 + x_3 + x_4$ y $0 < x_1 < x_2 < x_3 < x_4$; suponer que no hay errores inherentes en los números x_i . Observar la figura VI.

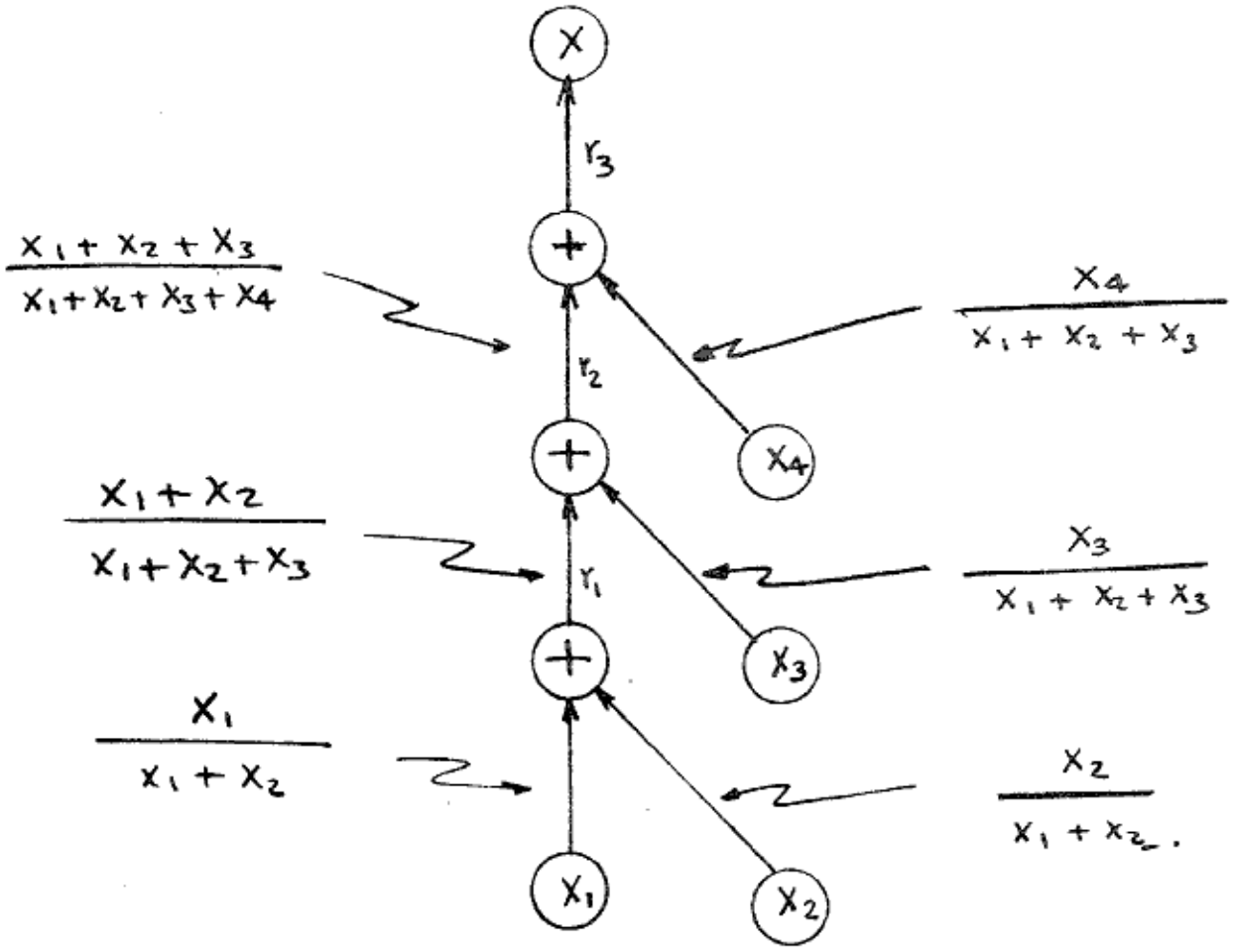


Fig. IV

$$x = x_1 + x_2 + x_3 + x_4$$

SOLUCION:

Puesto que no existe error inherente en los valores de x_i la aplicación de la regla para determinar el error total de:

$$\frac{e_x}{x} = r_1 \left(\frac{x_1 + x_2}{x_1 + x_2 + x_3}\right) \left(\frac{x_1 + x_2 + x_3}{x_1 + x_2 + x_3 + x_4}\right)$$

$$+ r_2 \left(\frac{x_1 + x_2 + x_3}{x_1 + x_2 + x_3 + x_4}\right) + r_3$$

en donde $x = x_1 + x_2 + x_3 + x_4$, por lo tanto

$$e_x = r_1(x_1 + x_2) + r_2(x_1 + x_2 + x_3) + r_3(x_1 + x_2 + x_3 + x_4)$$

ó

$$\left| e_x \right| \leq (3x_1 + 3x_2 + 2x_3 + x_4) \cdot 5 \cdot 10^{-t}$$

(suponiendo aritmética de cuatro dígitos t = 4)

Notar: Dado que no hay error inherente se ha asumido en la representación de los valores x_i que $x_i = \bar{x}_i$ por lo cual han sido suprimidas barras sobre las variables.

La ecuación para el límite del error total en la suma de n números que no tienen error inherentes es

$$\left| e_x \right| \leq \left[(n-1)x_1 + (n-1)x_2 + (n-2)x_3 + \ldots + 2x_{n-1} + x_n\right] \cdot 5 \cdot 10^{-t}$$

EJEMPLO NUMERICO. Suponer que se necesita efectuar las sumas de los siguientes números:

$0.2897 \cdot 10^0$	sumando en orden	$0.7873 \cdot 10^0$
$0.4976 \cdot 10^0$	ascendente las su-	$0.3275 \cdot 10^1$
$0.2488 \cdot 10^1$	mas parciales las	$0.1053 \cdot 10^2$
$0.7259 \cdot 10^1$	cantidades de la -	$0.2691 \cdot 10^2$
$0.1638 \cdot 10^2$	derecha (la prime-	$0.8940 \cdot 10^2$

$0.6249 \cdot 10^2$	mera suma parcial es la suma	$0.3056 \cdot 10^3$
$0.2162 \cdot 10^3$	de los dos primeros números;	$0.8289 \cdot 10^3$
$0.5233 \cdot 10^3$	la segunda suma parcial es la	$0.2232 \cdot 10^4$
$0.1403 \cdot 10^4$	suma de la primera suma par-	$0.7523 \cdot 10^4$
$0.5291 \cdot 10^4$	cial y el tercer número, etc.)	

Nota: Aunque es más usual una aritmética de 8 dígitos en una computadora se ilustra por facilidad para cuatro dígitos y cuando exceda un resultado de los cuatro se redondea. (Ejemplo: la 2a. cantidad nos da un valor de $0.32753 \cdot 10^1$ redondeado $.3275 \cdot 10$)

Si sumamos los números en orden inverso, de Mayor a Menor, las sumas parciales son:

$0.6694 \cdot 10^4$	$(.5291 \cdot 10^4 + 0.1403 \cdot 10^4)$
$0.7217 \cdot 10^4$	
$0.7433 \cdot 10^4$	La suma correcta a ocho cifras se encuentra conservando
$0.7495 \cdot 10^4$	todos los dígitos en cada suma. Siendo esta de
$0.7511 \cdot 10^4$	$0.75229043 \cdot 10^4$.
$0.7518 \cdot 10^4$	
$0.7520 \cdot 10^4$	Entonces el error en la suma ascendente es -0.1×10^4,
$0.7520 \cdot 10^4$	mientras que en la descendente es $2.9 \cdot 10^0$ (29 veces ma-
$0.7520 \cdot 10^4$	yor).

Ahora bien, los límites en los errores son del orden de 5.5×10^0 para la suma ascendente y 33.10^0 para la descendente. Siendo en los dos casos los errores actuales considerablemente menores que el error máximo posible.

Ejemplo 4.

Sumar cuatro números positivos aproximadamente iguales.
$x = (x_1 + x_2) + (x_3 + x_4)$

y

$$x_i = x_0 + \delta_i \quad , \quad i = 1, 2, 3, 4 \quad \text{en que} \quad |d_i| \ll x_0$$

$$e_{xi} = 0$$

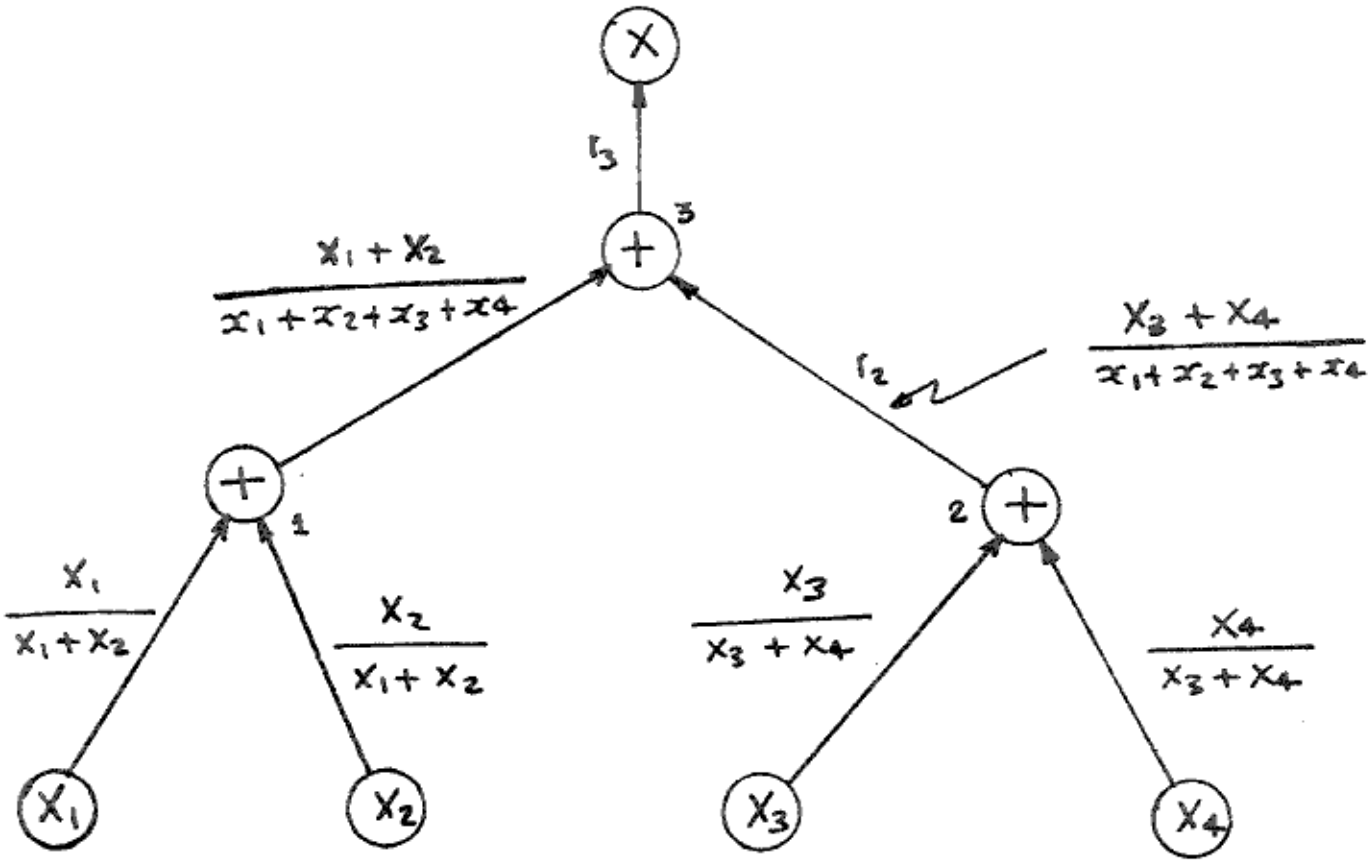


Fig. V

La salida de los nodos 1 y 2 está dada por r_1 y r_2 (asumiendo que los valores de x_i no tienen error inherente) que se multiplicarán por las etiquetas que salen de los nodos 1 y 2 respectivamente más el error de redondeo r_3 producido en la suma del nodo 3 para obtener:

$$\frac{e_x}{x} = r_1 \left(\frac{x_1 + x_2}{x_1 + x_2 + x_3 + x_4}\right) + r_2 \left(\frac{x_3 + x_4}{x_1 + x_2 + x_3 + x_4}\right) + r_3$$

6

$$\left| e_x \right| \leq (2x_1 + 2x_2 + 2x_3 + 2x_4) \cdot 5 \cdot 10^{-4}$$

expresado x_i en términos de x_0 y despreciando términos en un valor de comparados con la x_0 se obtiene finalmente

$$\left| e_x \right| \leq 4.10^{-3} \cdot x_0$$

usando la fórmula del ejemplo anterior para $\left|e_x\right|$ da

$$\left| e_x \right| \leq 4.5 \cdot 10^{-3} \cdot x_0$$

lo cual difiere del ejemplo; generalizando:

Si deseamos sumar n^2 números positivos de una magnitud aproximadamente igual el error total por redondeo se reduce si se suman en n grupos de n elementos cada uno y después se suman las n sumas parciales. Para un valor grande de n, el límite en el error es únicamente 1/n del límite correspondiente a la suma de los n^2 términos en una sola "faja" (como en el ejemplo Nº 3).

Sugerencias prácticas para lograr Mayor Precisión en los procesos de cálculo.

- En sumas y restas de números conviene trabajar primero con los números más pequeños.
- Evitar en lo posible la sustracción de dos números aproximadamente iguales.
- Reescribir la expresión para minimizar el número de operaciones haciendo primero aquellas de acuerdo a los puntos anteriores.

La observación de estas sugerencias en el peor de los casos evita complicar el problema con errores de redondeo adicionales.