GUIDO JOSÉ PACE

MÉTODOS NUMÉRICOS

FACULTAD DE CIENCIAS EXACTAS Y NAT. Y AGRIMENSURA
UNIVERSIDAD NACIONAL DEL NORDESTE
BIBLIOTECA
FACULTAD DE CS. EXACTAS Y NAT. Y AGRIMENSURA
CORRIENTES

FACULTAD DE CIENCIAS EXACTAS
Y NATURALES Y AGRIMENSURA
U N N E
INVENTARIO PATRIMONIAL
NUMERO 00- 39864
FECHA 27:10:12

Editorial
Universitaria de la
Universidad
Nacional del
Nordeste

© 1997 EUDENE, Editorial Universitaria de la Universidad Nacional del Nordeste,
25 de Mayo 868 (3400) Corrientes, República Argentina.

Queda hecho el depósito que marca la ley 11.723.

Impreso en la Argentina Printed in Argentina

ISBN 950-656-017-X

PRÓLOGO

El cotidiano uso de las computadoras electrónicas digitales para resolver nuestros problemas profesionales, hasta los más elementales, como escribir una simple nota, nos impone una modalidad de trabajo nueva y diferente; nueva, porque recién ahora empezamos a acostumbrarnos a usarlas, a subordinar nuestras viejas metodología a ellas y a dedicar parte de nuestro tiempo a estudiar cosas relacionadas con ellas, aunque nos dediquemos a otra especialidad totalmente diferente a la computación. En una palabra, vamos desarrollando una suerte de dependencia tan fuerte que, no solo abarca cuestiones meramente formales, operativas, sino también, funcionales y lógicas.

En este orden de cosas, los que probablemente han sufrido los mayores cambios, son aquellos que la usan para resolver problemas con fuerte contenido matemático, del tipo que fuere, dado que, en un principio la computadora fue pensada para resolver operaciones aritméticas elementales y ejecutar instrucciones lógicas. A partir de esto, puede construirse una inmensa superestructura matemática, con solo saber de que manera podemos descomponerlo todo, en términos de operaciones aritméticas elementales e instrucciones lógicas ejecutables.

En ocasiones, esta tarea no es nada sencilla; es necesario recurrir a muy complejos y sofisticados modelos matemáticos que, paradójicamente, solo están compuestos de sumas, restas, multiplicaciones y divisiones. Algunas más complejas, como p. ej., un seno hiperbólico o un arco tangente, pueden hacernos creer que las cosas no son así, por lo ampuloso de sus nombres, pero, en esencia, se trata de funciones preprogramadas que, mediante el concurso de simples desarrollos en series de *Taylor* o función equivalente, están compuestas, en definitiva, por solamente las operaciones antes mencionadas.

Es así que hoy, es muy difícil prescindir de los *MÉTODOS NUMÉRICOS*, objeto de estudio en nuestro trabajo, a Matemáticos, Ingenieros, Físicos, etc., pues, esta parte de la matemática no hace más que preocuparse de elaborar modelos apropiados, destinados al procesamiento digital mediante computadoras electrónicas. Naturalmente, el estudio de los *MÉTODOS NUMÉRICOS* puede ser realizado considerando diferentes enfoques y niveles, como cualquier otra ciencia; nosotros hemos elegido como enfoque, el clásico y como nivel, el introductorio, dirigido a la aplicación directa de los mismos; vale decir, los principales beneficiarios del trabajo serán los profesionales de las distintas ramas técnicas y de las ciencias exactas que necesiten del procesamiento digital y los estudiantes, que por la especialidad de la carrera que cursan, tingan una asignatura homónima.

Finalmente, pero, no por ello menos importante, es mi deseo rendir un sentido homenaje a quien escribió, en nuestro país; el primer libro de la especialidad que llegó a mis manos, me refiero al Dr. Manuel Sadosky, de cuya obra se ha preservado, en el presente trabajo, la estructura del contenido teórico y la correspondiente ejercitación original de algunos temas clásicos.

Esperando contribuir con la pequeña parte que me corresponde en el desarrollo de nuestra región, agradezco a todos aquellos que han hecho posible la edición del presente trabajo.

El Autor

- CONTENIDO -

1 . - INTRODUCCIÓN ..... Pág. N° 11
1.1.- FUNDAMENTOS ..... " " 13
1.2.- LA RESOLUCIÓN DE PROBLEMAS ..... " " 14
1.2.1.- Estudio y Formulación del Problema ..... " " 15
1.2.2.- Desarrollo de un Modelo ..... " " 16
1.2.3.- Validación y Optimización del Modelo ..... " " 17
1.2.4.- Codificación y Procesamiento ..... " " 17
1.3.- UN EJEMPLO DE IMPLEMENTACIÓN ..... " " 18
1.4.- PROCESOS ALGORITMICOS ..... " " 21
1.5.- CLASIFICACIÓN DE LOS ALGORITMOS ..... " " 23
1.5.1.- Algoritmos Directos ..... " " 23
1.5.2.- " Indirectos ..... " " 24
1.5.3.- " Infinitos ..... " " 25
1.6.- LA FORMA DE RESOLUCIÓN DE LOS PROBLEMAS ..... " " 26
1.6.1.- Métodos Gráficos ..... " " 28
1.6.2.- " de Enumeración ..... " " 28
1.6.3.- " de Tanteos ..... " " 29
1.6.4.- " de Discretización ..... " " 29
1.6.5.- " de Simulación ..... " " 30
1.6.6.- " Iterativos ..... " " 31
1.7.- CONVERGENCIA ..... " " 34
1.7.1.- El Problema General ..... " " 34
1.7.2.- Grado de Convergencia ..... " " 35

2.- APROXIMACIONES NUMÉRICAS ..... " " 39
2.1.- INTRODUCCIÓN ..... " " 41
2.2.- NÚMEROS EXACTOS Y APROXIMADOS ..... " " 42

2.2.1.- Números Exactos ... Pág. N° 42
2.2.2.- " Aproximados ... " " 43
2.3.- CLASIFICACIÓN DE LOS ERRORES ... " " 45
2.3.1.- Errores por Abstracción ... " " 46
2.3.2.- " " Inherencia ... " " 46
2.3.3.- " " Truncamiento ... " " 48
2.3.4.- " " Redondeo ... " " 50
2.3.5.- Tipos de Redondeo ... " " 51
2.3.5.1.- Redondeo Inferior ... " " 51
2.3.5.2.- " Simétrico ... " " 51
2.3.5.3.- " Aleatorio ... " " 52
2.3.6.- El *Épsilon* de la Computadora ... " " 52
2.4.- ERROR TOTAL Y PROPAGACIÓN DE LOS ERRORES ... " " 55

3.- ERRORES EN LAS OPERACIONES ... " " 59
3.1.- INTRODUCCIÓN ... " " 61
3.2.- ERROR ABSOLUTO Y ERROR RELATIVO ... " " 62
3.3.- ERROR EN LOS ALGORITMOS INFINITOS ... " " 63
3.4.- EMPLEO DE FÓRMULAS DIFERENCIALES ... " " 66
3.5.- ERROR EN LAS OPERACIONES ELEMENTALES ... " " 68
3.6.- GRÁFICOS DE PROCESOS ... " " 71
3.6.1.- Metodología ... " " 71
3.6.2.- Aplicaciones ... " " 74
3.6.3.- Conclusiones ... " " 80
3.7.- PROBLEMA INVERSO DEL CALCULO DE ERRORES ... " " 81

4.- RAÍCES DE ECUACIONES ... " " 83
4.1.- INTRODUCCIÓN ... " " 85
4.2.- COMPORTAMIENTO DE LAS ECUACIONES ... " " 86
4.3.- MÉTODOS GENERALES DE RESOLUCIÓN ... " " 91
4.3.1.- Método de TANTEOS ... " " 91
4.3.1.1.- Error en el método de TANTEOS ... " " 92
4.3.2.- Método del INTERVALO MEDIO ... " " 92
4.3.2.1.- Error en el método del INTERVALO MEDIO ... " " 94
4.3.3.- Método de INTERPOLACIÓN LINEAL ... " " 96

4.3.3.1.- Convergencia en el método de INTERPOLACIÓN LINEAL ... Pág. N° 97
4.3.4.- Método de NEWTON-RAPHSON ... " " 98
4.3.4.1.- Convergencia del método de NEWTON-RAPHSON ... " " 99
4.3.4.2.- Error en el método de NEWTON-RAPHSON ... " " 101
4.3.5.- Método MIXTO ... " " 105

5.- TEORÍA GENERAL DE LA ITERACIÓN ... " " 107
5.1.- INTRODUCCIÓN ... " " 109
5.2.- MÉTODO DE ITERACIÓN ... " " 110
5.2.1.- Convergencia del método de ITERACIÓN ... " " 113
5.2.2.- Utilidad del método de ITERACIÓN ... " " 113
5.3.- TEORÍA GENERAL DE LA ITERACIÓN ... " " 116
5.4.- PROCESO DELTA-CUADRADO (Δ^2) DE AITKEN ... " " 123
5.5.- MÉTODO DE SEGUNDO ORDEN DE NEWTON ... " " 127
5.6.- RESOLUCIÓN DE ECUACIONES ALGEBRAICAS ... " " 129
5.6.1.- Acotación de las raíces reales ... " " 130
5.6.2.- Separación de las raíces reales ... " " 132
5.6.3.- Teorema de STURM ... " " 132

6.- ECUACIONES LINEALES SIMULTÁNEAS ... " " 135
6.1.- INTRODUCCIÓN ... " " 137
6.2.- SISTEMAS DE ECUACIONES NO HOMOGÉNEOS ... " " 140
6.2.1.- Método de ELIMINACIÓN DE GAUSS ... " " 141
6.2.2.- " " GAUSS-JORDAN ... " " 144
6.2.3.- " " las ECUACIONES DE ERROR ... " " 147
6.2.4.- " " la MATRIZ INVERSA ... " " 149
6.2.5.- " " GAUSS-SEIDEL ... " " 150
6.3.- SISTEMAS DE ECUACIONES HOMOGÉNEOS ... " " 154
6.3.1.- Generalidades ... " " 154
6.3.2.- Método de FADDEEV-LEVERRIER ... " " 157
6.3.3.- Método de las POTENCIAS ... " " 160
6.4.- SISTEMAS DE ECUACIONES ORTOGONALES ... " " 165
6.4.1.- Introducción ... " " 165
6.4.2.- El método ANALÍTICO ... " " 166
6.4.3.- Método de Ortogonalización de GRAM-SCHMIDT ... " " 168

7.- INTERPOLACIÓN NUMÉRICA ..... Pág. N° 175
7.1.- INTRODUCCIÓN ..... " " 177
7.2.- INTERPOLACIÓN ..... " " 178
7.2.1.- Tablas con Valores Equidistantes ..... " " 179
Tablas de Diferencias AVANZADAS ..... " " 180
7.3.- MÉTODOS DE INTERPOLACIÓN ..... " " 182
7.3.1.- Fórmula de NEWTON-GREGORY ASCENDENTE ..... " " 182
7.3.2.- Interpolación LINEAL y CUADRÁTICA ..... " " 186
7.3.3.- Fórmula de NEWTON-GREGORY DESCENDENTE ..... " " 188
Tablas de Diferencias ATRASADAS ..... " " 189
7.3.4.- Fórmula de LAGRANGE ..... " " 193
7.3.4.1.- Fórmula de LAGRANGE para Valores Equidistantes ..... " " 195
7.3.5.- Fórmula de INTERPOLACIÓN PARABÓLICA PROGRESIVA ... " " 197
7.3.6.- Fórmula de GAUSS ..... " " 201
7.3.7.- " " BESSEL ..... " " 204
Coeficientes BESSELIANOS ..... " " 205
7.3.8.- Fórmula de STIRLING ..... " " 206
7.3.9.- " " EVERETT ..... " " 208
7.4.- INTERPOLACIÓN INVERSA ..... " " 210
7.5.- AJUSTE ..... " " 211
7.5.1.- AJUSTE E INTERPOLACIÓN ..... " " 211
7.5.2.- AJUSTE MEDIANTE CURVAS ..... " " 212
7.5.3.- AJUSTE MEDIANTE POLINOMIOS ..... " " 213
7.5.3.1.- Metodología del Ajuste ..... " " 213
7.5.3.1.- Mecánica del Cálculo ..... " " 215

8.- DERIVACIÓN E INTEGRACIÓN ..... " " 217
8.1.- DIFERENCIACIÓN NUMÉRICA ..... " " 219
8.1.1.- Derivación Mediante Fórmulas de Interpolación ..... " " 220
8.1.2.- Algoritmo de HORNER ..... " " 222
8.1.2.1.- Introducción ..... " " 222
8.1.2.2.- Cálculo del Valor del Polinomio ..... " " 223
8.1.2.3.- " de las Derivadas Sucesivas ..... " " 224
8.1.3.- El Método de la Serie de *TAYLOR* ..... " " 224
8.2.- INTEGRACIÓN NUMÉRICA ..... " " 232
8.2.1.- Introducción ..... " " 233

8.2.2.- Fórmula de los TRAPECIOS ..... Pág. N° 233
8.2.3.- " " SIMPSON ..... " " 236
8.2.4.- Regla de los TRES OCTAVOS DE SIMPSON ..... " " 238
8.2.5.- Fórmula de EULER-MACLAURYN ..... " " 241
8.2.6.- " " GREGORY ..... " " 243

9.- RESOLUCIÓN DE ECUACIONES DIFERENCIALES .... " " 247
Métodos que no comienzan por sí mismos
9.1.- INTRODUCCIÓN ..... " " 248
9.2.- MÉTODO DE EULER ..... " " 250
9.2.1.- Error en el Método de EULER ..... " " 253
9.3.- MÉTODO MODIFICADO DE EULER ..... " " 254
9.3.1.- Error en el Método Modificado de EULER ..... " " 257
9.4.- MÉTODO DE RUNGE-KUTTA ..... " " 258
7.4.1.- Método de RUNGE-KUTTA de Primer Orden ..... " " 260
7.4.2.- Método de RUNGE-KUTTA de Segundo Orden ..... " " 260
7.4.3.- Error del Método de RUNGE-KUTTA de Segundo Orden ..... " " 263
7.4.4.- Métodos de RUNGE-KUTTA de Mayor Orden ..... " " 264
7.4.5.- Error de los Métodos de RUNGE-KUTTA de Orden Superior ..... " " 265
7.4.6.- Método de RUNGE-KUTTA para Sistemas de Ecuaciones Diferenciales ..... " " 266
7.4.7.- Método de RUNGE-KUTTA para Ecuaciones Diferenciales de Orden Superior ..... " " 267

10.- RESOLUCIÓN DE ECUACIONES DIFERENCIALES ... " " 271
Métodos que no comienzan por sí mismos
10.1.- **INTRODUCCIÓN** ..... " " 273
10.2.- MÉTODO DE MILNE ..... " " 274
10.2.1.- Error en el Método de MILNE ..... " " 277
10.2.2.- Generalización del Método de MILNE ..... " " 278
10.3.- ESTABILIDAD DE LOS MÉTODOS DE INTEGRACIÓN ..... " " 278
10.4.- MÉTODO DE HAMMING ..... " " 281
10.4.1.- Método Modificado de HAMMING ..... " " 284
10.5.- ANÁLISIS DEL ERROR ..... " " 285
10.6.- SELECCIÓN DE UN MÉTODO DE INTEGRACIÓN NUMÉRICA . " " 288

11.- RESOLUCIÓN DE ECUACIONES DIFERENCIALES ... Pág. N° 291
Problemas con valores en la frontera
11.1.- PLANTEO GENERAL ..... " " 293
11.2.- FÓRMULAS DE APROXIMACIÓN POR DIFERENCIAS CENTRALES ..... " " 295
11.3.- MÉTODO DE LAS DIFERENCIAS FINITAS ..... " " 298
11.4.- MÉTODO DE LAS DIFERENCIAS FINITAS PARA ECUACIONES DIFERENCIALES DE ORDEN SUPERIOR ... " " 302

BIBLIOGRAFÍA ..... " " 305

ÍNDICE ..... " " 307

♦♦♦ ●●● ♦♦♦

Capítulo 1

Introducción

1 . - INTRODUCCIÓN

Fundamentos. La Resolución de Problemas. Estudio y Formulación del Problema. Desarrollo de un Modelo. Validación y Optimización del Modelo. Codificación y Procesamiento. Un Ejemplo de Implementación. Procesos Algorítmicos. Clasificación de los Algoritmos. Algoritmos Directos. Algoritmos Indirectos. Algoritmos Infinitos. La Forma de Resolución de Problemas. Métodos Gráficos. Métodos de Enumeración. Métodos de Tanteos. Métodos de Discretización. Métodos de Simulación. Métodos Iterativos. Convergencia. El Problema General. Grado de Convergencia.

1.1.- FUNDAMENTOS

En la actualidad, las diversas ramas de la ciencia y la técnica, plantean complicados problemas matemáticos que, por su naturaleza compleja, no son fáciles de resolver utilizando los métodos tradicionales de la matemática pura; vale decir, la geometría, el álgebra, el análisis matemático, etc., debido principalmente a alguna de las siguientes causas:

a) que la matemática pura no provea ningún método analítico deductivo, capaz de resolver el problema, o
b) que aún existiendo un método analítico que resuelva el problema, sea de aplicación tan compleja y/o extensa que resulte poco práctico de ser utilizado con ventajas apreciables, o bien
c) que, por las características propias del problema, sea absolutamente necesario resolverlo mediante el concurso de computadoras electrónicas,

en todos los casos mencionados, es necesario valerse de métodos factibles de ser aplicados de manera fructífera y con cierta facilidad los que, por lógica consecuencia, se encuentran fuera del campo de la matemática pura.

Las circunstancias apuntadas en el párrafo anterior, dieron lugar a la aparición de una nueva rama de la matemática, denominada **MATEMÁTICA APLICADA** la que, en sus orígenes, estaba constituida exclusivamente por la materia que nos ocupa; es decir, los **MÉTODOS NUMÉRICOS**. Más recientemente se fueron incorporando otras disciplinas a la matemática aplicada, tales como el cálculo de probabilidades, la estadística, la investigación operativa, etc.

Cabe acotar que los resultados alcanzados en las soluciones obtenidas mediante métodos analítico deductivos, propios de la matemática pura, son exactos; es decir, se pueden formular, al menos teóricamente, con toda precisión, aunque ésta, no pueda ser alcanzada en la práctica. Esta precisión no es posible de obtener con los métodos numéricos, ni siquiera desde su aspecto puramente formal; es por esta razón que a los métodos numéricos se los designa también, con el nombre de **MÉTODOS APROXIMADOS**, para diferenciarlos de los primeros; aunque en rigor de verdad, con ellos sea posible alcanzar, en la práctica, el grado de aproximación que se desee con tal de satisfacer ciertas condiciones que, en cada caso particular, serán estudiadas.

No obstante que, los métodos numéricos han sido útiles para resolver un gran número de problemas que de otra manera no hubiesen podido ser resueltos, al menos desde el punto de vista práctico, el gran auge de esta disciplina se debe específicamente a lo referido en el tercer punto; vale decir, a la gran afinidad y compatibilidad que ellos tienen de ser sometidos al procesamiento mediante computadoras electrónicas digitales, las que por su naturaleza solo pueden efectuar operaciones aritméticas elementales y tomar decisiones lógicas fundamentales, elementos éstos, a los cuales se pueden reducir todos los métodos que durante el desarrollo del presente curso serán estudiados.

1.2.- LA RESOLUCIÓN DE PROBLEMAS

Según se mencionó en el punto anterior, esta materia está relacionada, fundamentalmente, con la manera de resolver problemas matemáticos para los cuales no existen métodos analíticos de solución; o bien que aún existiendo algunos de estos, el o los métodos propuestos resulten poco prácticos por su extensión y/o complejidad y que, además, se preste para ser procesado por medio de computadoras electrónicas digitales. Aspectos todos estos, respecto de los cuales nos dedicaremos con igual profundidad, con el objeto de hacer de la materia un todo funcional.

Para lograr este objetivo se expondrán ordenada y sistemáticamente, una serie de métodos que serán desarrollados a medida que avance el curso y que, además, sean aptos para ser programados en vista de su posterior procesamiento mediante computadoras electrónicas digitales.

En términos generales, es posible afirmar que la mayoría de los problemas que se presentan en la práctica, provengan del campo científico de donde provengan, se resuelven recorriendo correctamente cuatro etapas fundamentales, las que son enunciadas a continuación:

1.- Estudio y Formulación precisa del problema,

2.- Desarrollo de un modelo,

3.- Validación del modelo y análisis de Optimización, y

4.- Codificación del modelo en algún lenguaje de programación y procesamiento del mismo.

En los próximos títulos serán estudiados en forma muy somera, pero, tomando cada uno de los puntos antes detallados en particular, el significado y sus contenidos básicos.

1.2.1.- ESTUDIO Y FORMULACIÓN PRECISA DEL PROBLEMA

Esta primera etapa de la resolución de cualquier problema, tiene como objetivo primordial lograr un conocimiento lo más acabado posible del mismo; consiste en el estudio del fenómeno, las variables y parámetros que intervienen en el mismo, las relaciones que los ligan y las probables líneas de solución existentes.

La etapa del *Estudio y Formulación Precisa del Problema* es una de las causas más comunes de dificultades en la práctica, ya sea porque no se conoce en profundidad la naturaleza funcional del problema; ya sea porque proviene de campos de actividad que, por lo general, resultan muy difíciles de cuantificar, tales como los de la conducta humana, todas las ramas del arte, las ciencias relacionadas con la jurisprudencia, etc. El conocimiento insuficiente o erróneo del fenómeno o problema en estudio, conlleva a uno de los cuatro tipos de errores tipificados en la materia (como podrá estudiarse en el capítulo correspondiente a las aproximaciones numéricas), cual es el de **ABSTRACCIÓN**, que constituye un error de tipo grosero.

Dado que no existe una metodología de trabajo determinada, destinada a llevar adelante la realización del estudio y formulación de un problema, al menos dentro de los límites establecidos para esta materia, no se darán mayores detalles sobre este tema, en particular.

1.2.2.- DESARROLLO DE UN MODELO

El origen de todos los problemas es el mundo real; estos problemas nacen, fundamentalmente, de los continuos esfuerzos que realizan los hombres por comprender, describir y predecir los fenómenos del mundo real, ya sean estos, naturales o artificiales; vale decir, producidos sin la intervención de los hombres (fenómenos naturales), o con la previa planificación y la directa participación de éstos (fenómenos artificiales).

Para formular analíticamente o racionalmente los problemas, la ciencia y la técnica se valen de idealizaciones o abstracciones, más o menos complejas de la realidad, a las cuales denomina **MODELOS** y que, pueden definirse más racionalmente de la siguiente manera:

> *" Se denomina mediante el vocablo **MODELO** a una abstracción de algún fenómeno real, que tiene la posibilidad de emplearse para satisfacer objetivos de comprensión, predicción y control de aquel, con cierto grado de aproximación previamente elegido. "*

Es posible realizar una amplia clasificación de los modelos; sin embargo, a los efectos de esta materia, existe un tipo particularmente importante de ellos, que incluye los denominados **MODELOS MATEMÁTICOS** y que, son aquellos que expresan las relaciones funcionales entre variables y parámetros del problema mediante ecuaciones matemáticas y decisiones lógicas exclusivamente; son los clasificados de esta manera, aquellos modelos que serán utilizados para formalizar analíticamente la descripción de los fenómenos a tratar en el presente curso.

Cabe acotar que un determinado problema no posee un solo y único modelo matemático; sino que, por lo general, pueden ser muchos según el número de parámetros que se hagan intervenir y el tipo de ecuaciones que se adopte para relacionarlos; todo esto, en concordancia con el grado de precisión que se desea alcanzar. Así mismo se aclara, tal como se ha dicho, que no existe un conjunto de reglas generales que permitan elaborar los modelos matemáticos de un problema real particular; sino que, en base a nuestros conocimientos previos y por analogías, la experiencia indicará el camino que deberá seguirse para lograrlo.

1.2.3.- VALIDACIÓN Y OPTIMIZACIÓN DEL MODELO

Una vez que el problema ha sido formulado con toda precisión y reducido a un conjunto de ecuaciones matemáticas cuantificables y relaciones lógicas elementales; vale decir, elaborado un modelo matemático en primera instancia y, a la luz del problema real que se trata de resolver, debe ser analizado para comprobar si realmente aquel describe el fenómeno en estudio, su grado de precisión, complejidad, etc.; todo lo cual, constituye la etapa de validación del modelo.

Esta etapa comprende también, la validación de las *Muestras Artificiales* (series de valores de las variables y parámetros del modelo, que son generadas directamente por la computadora a medida que se van necesitando) las que serán utilizadas en el posterior procesamiento del modelo; sean estas, tanto de variables discretas o continuas, como de variables deterministas o aleatorias.

En segunda instancia, resulta conveniente someterlo a un severo examen desde el punto de vista del procesamiento, a los fines de reducir el modelo matemático a una forma más simple o breve mediante la aplicación de reglas y procedimientos matemáticos y de programación que tiendan a su mejor y más rápido procesamiento; esta parte de la etapa es la de optimización.

Dado que la manera de realizar esta etapa, tanto en lo que se refiere a la parte de validación como a la de optimización, no constituye un objetivo específico de nuestra materia, no se aportarán mayores detalles sobre el particular.

1.2.4.- CODIFICACIÓN Y PROCESAMIENTO

Definido el problema, elaborado y optimizado un modelo matemático, todavía se está lejos de la posibilidad de procesarlo mediante computadoras electrónicas digitales, con el fin de obtener algún resultado, pues, en general, la mayoría de los modelos matemáticos no son directamente procesables; y aunque lo fueran, se deberá completar una etapa previa que consiste en la programación; es decir, en la codificación del modelo matemático en algún lenguaje de computación comprensible por el ordenador.

1.3.- UN EJEMPLO DE IMPLEMENTACIÓN

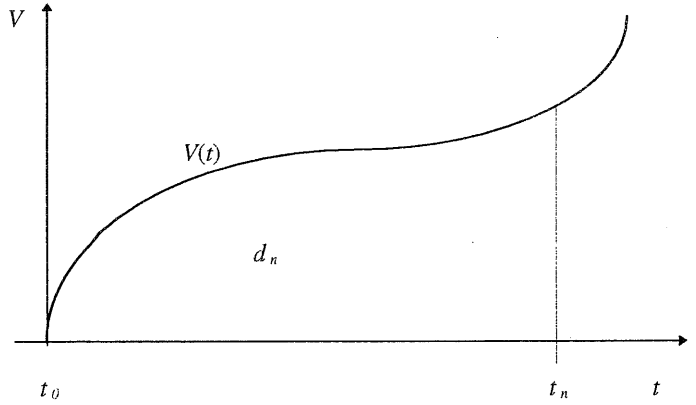
Se conoce la relación que describe la velocidad de un móvil en función del tiempo $V(t)$ y se desea conocer la distancia d_n en el instante $t=t_n$. Planteado de esta manera el problema, el modelo matemático teórico proporcionado por la mecánica es:

$$d_n = d_0 + \int_{t=t_0}^{t=t_n} V(t)\,dt \qquad (1.1)$$

donde t_0 y d_0 son el tiempo inicial y la distancia en dicho instante, respectivamente.

Esta relativamente compleja operación que pertenece al dominio del análisis matemático; más precisamente al del cálculo infinitesimal, no es en absoluto comprendida por la computadora, la que solo sabe de las operaciones aritméticas elementales. No obstante el problema tiene solución.

Para quienes están familiarizados con la expresión (1.1), el segundo término del segundo miembro representa el valor de la superficie comprendida entre la curva $V=V(t)$; el eje de las abscisas y las ordenadas extremas $t=t_0$, $t=t_n$. Suponiendo que para el instante $t_0=0$, es $d_0=0$; entonces la expresión (1.1) representa la parte interior de la figura 1.1.



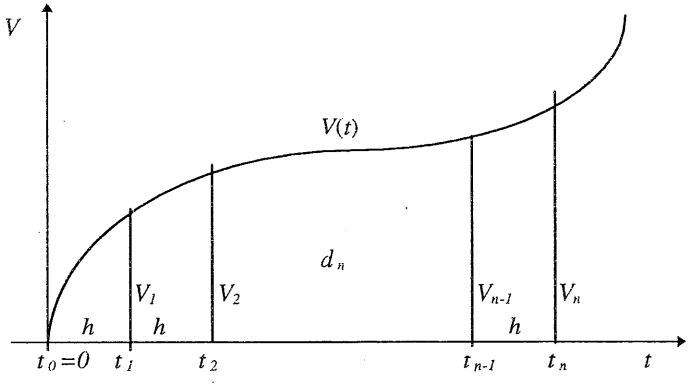
- Figura 1.1 -

Para poder resolver el problema en las condiciones descriptas en los párrafos anteriores, resulta necesario algún tipo de aproximación a la expresión (1.1), que pueda resultar comprensible por la computadora, obtenido mediante un modelo matemático basado en algún método numérico. Teóricamente, es posible realizarlo de muchas maneras diferentes; se verá una forma muy simple.

Se trata de realizar la aproximación a la integral de la figura 1.1 mediante la suma de las áreas de los trapecios obtenidos por medio de la subdivisión de la superficie dada y considerando el contorno superior como una poligonal [que reemplaza a la curva $V=V(t)$, como se muestra en la figura 1.2.

Llamando t_i al tiempo en el instante i; V_i la velocidad en dicho instante y h al ancho constante de las distintas franjas; o, lo que es lo mismo, a la altura de los trapecios formados, resulta:

$$h = t_2 - t_1 = t_3 - t_4 = \ldots = t_{n-1} - t_n$$



- Figura 1.2 -

Dadas las condiciones estudiadas, la distancia requerida d_n, puede calcularse mediante:

$$d_n = d_0 + \frac{V_0 + V_1}{2} h + \frac{V_1 + V_2}{2} h + \cdots + \frac{V_{n-1} + V_n}{2} h \qquad (1.2)$$

fórmula con la cual, dadas sus características matemáticas (solo intervienen operaciones aritméticas elementales), resulta factible de ser procesada mediante el uso de computadoras.

Respecto de la precisión de la expresión (1.2) cabe señalar que permite alcanzar la aproximación deseada, cualquiera sea ésta, con tal de tomar la altura h de los trapecios suficientemente pequeña.

Tomando en consideración la expresión (1.2), anteriormente estudiada, es posible lograr su optimización analítica, mediante el uso del operador sumatoria, resultando:

(1.3) $$d_n = d_0 + \sum_{i=1}^{n} \frac{V_{i-1} + V_i}{2} h$$

que, si bien, esta última expresión se presta muy bien para ser procesada mediante computadoras electrónicas, no resulta comprensible por ellas dado que, estas poseen un lenguaje de comunicación codificado de manera especial; por lo tanto, es absolutamente necesario realizar otro paso adicional, previo al proceso de cálculo, el cual consiste en la codificación de la fórmula (1.3) en algún lenguaje de programación cuyo núcleo esencial, en lenguaje C++ es:

```
d(0) = 0;
V(0) = 0;
{
  for (i = 1 ; i <= n ; i++){
  d(i) = d(i-1) + (V(i-1) - V(i)) * h / 2;
  }
}
d(n) = d(n) + d(0);
```

Finalmente, resta la parte correspondiente al procesamiento del programa obtenido, para completar esta etapa y consecuentemente lograr los objetivos propuestos en el planteo del problema.

♦♦♦ ●●● ♦♦♦

Como puede apreciarse a través del ejemplo presentado, entre un planteo realizado mediante un modelo matemático analítico puro, y el procesamiento del mismo existe un largo trecho que habrá que recorrer necesariamente, para lograr obtener los resultados esperados.

1.4.- PROCESOS ALGORÍTMICOS

La palabra **ALGORITMO**, deriva del nombre con que se conoció a un famoso matemático árabe del siglo IX, llamado *AL-KHUWARIZMI*, el que, estando interesado en resolver ciertos problemas de carácter aritmético y geométrico, y desconociéndose, por aquel entonces, la formulación analítica de las actuales expresiones matemáticas (notación de variables y parámetros, igualdades, ecuaciones, etc.), describió varios métodos para hacerlo de manera más o menos sistemática.

Estos métodos fueron presentados bajo la forma de listados de instrucciones específicas (al igual que hoy conocemos las denominadas recetas). Un derivado de su nombre es utilizado en la actualidad para referirse a la metodología que encierran dichos listados de cálculo o procesamiento sistemático. Demás está decir que, hoy los algoritmos no están compuestos de instrucciones literales, sino que, se trata de la codificación analítica de las matemáticas, que todos conocemos.

Un algoritmo es pues, en forma intuitiva, una receta, un conjunto de instrucciones o de especificaciones sobre un determinado proceso para realizar algo que, generalmente, es la manera de llegar a la solución de algún problema matemático. En este orden de cosas, la expresión **PROCESO ALGORÍTMICO** es entonces redundante, no obstante será utilizada con frecuencia en virtud de lo tradicional de su uso.

Visto el tema desde el punto de vista analítico, y a la luz de los conocimientos actuales, es posible definir el término algoritmo en forma más racional y estricta:

Definición: *" Para que una lista o conjunto de especificaciones o de instrucciones, pueda ser realmente considerado como un algoritmo matemático, debe poseer ciertas propiedades absolutas, las cuales serán expuestas a continuación:*

1.- estar compuesto por un conjunto de instrucciones o especificaciones matemáticamente cuantificables o lógicamente ejecutables,

2.- las instrucciones deben estar dadas en un cierto orden estricto, tal que permita resolver el problema en estudio,

3.- las decisiones lógicas y las operaciones aritméticas especificadas en cada instrucción deben estar libres de toda ambigüedad, y

4.- poseer un final cierto.

Veamos el significado y contenido de los puntos anteriores, para poder comprender su real alcance.

La primera condición enunciada, establece que para que un cierto conjunto de especificaciones conforme realmente un algoritmo debe resultar que, todas y cada una de las instrucciones que intervienen en él sean matemáticamente cuantificables o lógicamente codificables; vale decir que, considerando a las primeras como operadores, deben dar como resultado de su aplicación y evaluación, un número decimal, y las segundas, estar unívocamente determinadas para definir una alternativa, opción o camino; lo cual es equivalente a decir que, deben prestarse a ser procesadas por medio de una computadora electrónica digital. En su momento se estudiará en forma detallada de cuales instrucciones se trata y la forma en que cada una de ellas opera.

Con respecto a la segunda condición se hace referencia simplemente al orden lógico y estricto que deben guardar los distintos pasos o sentencias de un desarrollo matemático para llegar, al fin, al resultado deseado. Cabe consignar que, eventualmente, este orden no es único, dado que, algunas operaciones son permutables sin afectar en absoluto el valor verdadero del resultado final.

Es necesario destacar que, muchas de las operaciones permutables de la aritmética tradicional (sumas, restas, etc.), no son tales cuando los elementos intervinientes no son números exactos, sino aproximados (reales, con limitado número de dígitos significativos), si es que se quiere obtener la máxima precisión; dicho de otro modo, la causa de esto es que la aritmética se desarrolla bajo una hipótesis de precisión infinita, cosa que en la práctica de la matemática aplicada no ocurre, dado que, se opera con números aproximados, tal como será estudiado con mayor profundidad en los capítulos referentes a las **APROXIMACIONES NUMÉRICAS** y **ERRORES EN LAS OPERACIONES ARITMÉTICAS**.

En cuanto a la tercer condición enunciada, la misma hace referencia al hecho que si se está trabajando dentro de cierto contexto, cada instrucción o especificación del algoritmo debe poseer uno, y solamente un significado, perfectamente definido y estrictamente determinado.

Un ejemplo ilustrativo muy común consiste en el caso en el cual se debe tomar el resultado de la extracción de una raíz cuadrada, en el curso de cierto proceso de cálculo, con fines diversos, el texto de la especificación no puede ser simplemente: tomar la raíz de x, sin especificar, en cada caso, el signo que corresponda.

Otro caso típico se presenta cuando se está frente a la necesidad de calcular el área de un triángulo conociendo dos lados y un ángulo no comprendido por ellos, llamado justamente *Caso Ambiguo*.

En relación con la cuarta condición expuesta, se puede afirmar que ciertos algoritmos tienen longitud infinita o extremadamente larga, como será estudiado en el siguiente punto, en cuyo caso, el tiempo y la carga de procesamiento pueden llegar a ser muy grandes, tanto de no poder procesarse. En estos casos, se deberá agregar al algoritmo la o las instrucciones que permitan detenerlo, haciendo uso de algún criterio determinado, basado generalmente, en la precisión que se espera alcanzar. Se dirá, entonces, que el algoritmo finaliza su procesamiento por causa de una **PARADA PROGRAMADA**.

1.5.- CLASIFICACIÓN DE LOS ALGORITMOS

Dada la importancia que reviste el tema en la matemática aplicada en general, y para los métodos numéricos en particular, seguidamente se hará una clasificación de los algoritmos, desde el punto de vista del procesamiento a la vez que, se introducirán algunos conceptos importantes.

1.5.1.- ALGORITMOS DIRECTOS

Este tipo de algoritmos son sumamente simples. Veamos la siguiente:

Definición: *" Intuitivamente, se entenderá por **ALGORITMO DIRECTO**, aquellos cuyo resultado final se alcanza después de la ejecución de un número fijo y determinado de pasos conocidos y bien definidos."*

Con el objeto de fijar conceptos referidos a los algoritmos directos, se expone el siguiente:

Ejemplo N° 1.1.- Un caso simple de algoritmo directo se encuentra cuando se tiene que resolver la ecuación de primer grado:

$$a\,x + b = 0$$

Solución: Se sabe que en este caso la solución se alcanza cuando se calcula la raíz (real y única) mediante:

$$x = -\frac{b}{a}$$

lo cual se consigue en dos pasos: primero, se realiza el cociente de b sobre a; segundo, se cambia el signo del resultado obtenido en el primer paso.

♦♦♦ ●●● ♦♦♦

Así pues, algoritmo directo es aquel que consta de un número conocido de pasos elementales. Esto implica que la definición de algoritmo directo depende de lo que se entienda por paso elemental.

En este orden de ideas, se denominará **PASO ELEMENTAL** al constituido por alguna de las primeras y más simples de las operaciones aritméticas, que son: *adición, sustracción, multiplicación y división*; juntamente con las de comparación: *mayor, igual y menor.*

Es sabido que otras operaciones, tales como potenciación, radicación, etc. o trascendentes, como trigonométricas, logarítmicas, exponenciales, etc. pueden calcularse mediante pasos elementales haciendo uso de algoritmos especiales, generalmente basados en desarrollos en serie de *TAYLOR* y que, las de comparación, juntamente con las anteriores pertenecen al grupo de las **FUNCIONES PREPROGRAMADAS**. A los efectos del cómputo, en ocasiones cabe hacer otro tipo de clasificación de los algoritmos; por todo ello, la que se está realizando debe tomarse como puramente enunciativa y formal.

1.5.2.- ALGORITMOS INDIRECTOS

Existe una categoría de algoritmos más compleja que la estudiada en el punto anterior; sea la siguiente:

Definición: " *Se denominan **ALGORITMOS INDIRECTOS** todos aquellos cuyo resultado final se alcanza luego de haber procesado un número de pasos elementales, desconocido de antemano,.* "

La característica fundamental de estos algoritmos consiste en que, varias partes de él pueden ser repetidas un cierto número desconocido de veces. Es razonable preguntarse si estos algoritmos pueden llevar a la solución de un problema; la respuesta es afirmativa, como se verá durante el desarrollo del curso.

Es obvio que, utilizando ambos tipos de algoritmos; vale decir, no limitándose en la cualidad de ellos, se está en mejores condiciones para resolver problemas que, si solo se utilizan los algoritmos directos. Es más, la experiencia demuestra que, generalmente, los algoritmos indirectos son mucho más eficaces y fructíferos que los directos, aunque no se sepa a priori la cantidad de pasos que deberán ser realizados para alcanzar el resultado.

1.5.3.- ALGORITMOS INFINITOS

Aunque el uso de los algoritmos indirectos aumenta la potencia y la eficiencia de la capacidad para resolver problemas matemáticos, todavía existe una gran distancia real entre estos problemas y las soluciones numéricas primitivas que son necesarias realizar para resolverlos. Con el objeto de superar esta distancia es conveniente idealizar el concepto de algoritmo dejando de lado el requisito de que éste deba poseer un fin cierto; con esta premisa veamos la siguiente:

Definición: " *Reciben el nombre de **ALGORITMOS INFINITOS** todos aquellos cuyo resultado final, no es alcanzado jamás, por larga que sea la extensión de su procesamiento; vale decir que, teóricamente, resultaría necesario realizar infinitos pasos elementales para lograr el resultado esperado.* "

Considerando los algoritmos infinitos, anteriormente definidos, parece razonable pensar que se encontrarán soluciones cada vez más precisas si se computa indefinidamente. Este concepto implica un gran cambio, puesto que no está limitado a considerar métodos que poseen un final garantizado.

Existe un problema que surge del uso de los algoritmos que no se detienen por sí mismos. Dado que no paran jamás, tampoco llegan nunca a una respuesta de algún tipo, exacta o no.

En consecuencia, para que un algoritmo infinito sea útil, debe producir ciertas **APROXIMACIONES INTERMEDIAS** de la respuesta del problema en estudio. Se espera también, que aquellas aproximaciones intermedias sean cada vez más precisas a medida que el procesamiento del algoritmo avanza.

La consideración de la infinitud en forma precisa puede resultar muy compleja, especialmente cuando el problema consiste en precisar el significado exacto de lo que se quiere decir con: *"las estimaciones son cada vez más aproximadas al valor verdadero de la solución a medida que el procesamiento del algoritmo continúa"*.

Para poder hablar de estimaciones cada vez mejores, es necesario tener una medida de la discrepancia o error de cualquier estimación en particular; además, se debe tener una idea de la velocidad con que se vuelven buenas las estimaciones. De no ser así, no existe base alguna para decidir si una estimación, en particular, es buena o no. Intuitivamente, es posible asegurar que, para que un algoritmo resulte útil, deberá tener las propiedades expuestas.

Las características mencionadas en el párrafo anterior, serán estudiadas en general, bajo los títulos de **CONVERGENCIA** y **GRADO DE CONVERGENCIA**, respectivamente, y en particular, durante la exposición de cada uno de los métodos que se desarrollen en el transcurso de la materia.

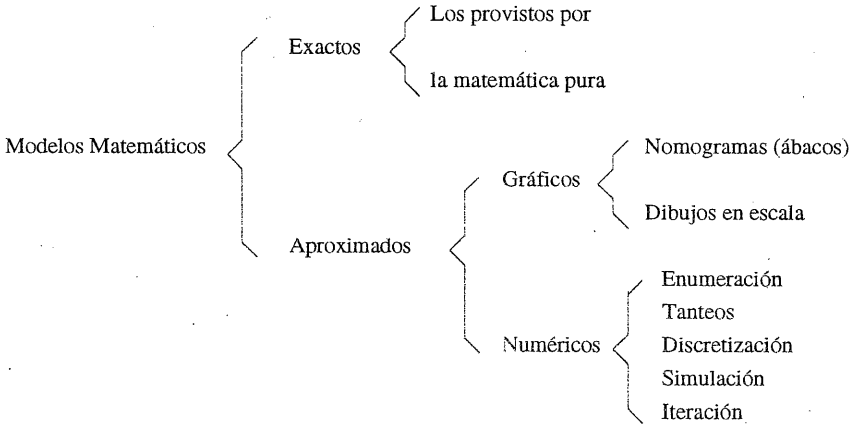
1.6.- LA FORMA DE RESOLUCIÓN DE LOS PROBLEMAS

Se ha estudiado que uno de los pasos que necesariamente se debe dar para resolver un problema matemático consiste en el desarrollo de un modelo matemático, que no es otra cosa que la idealización, más o menos abstracta, de la realidad involucrada en el problema en cuestión.

Se ha hecho mención, además, que el modelo matemático que resuelve un problema determinado no es único, sino que, cabe la posibilidad de elaborar varios modelos diferentes, ya sea por simples consideraciones respecto de sus parámetros, ya sea por la actitud adoptada al desarrollar el mismo. Por ejemplo, las expresiones 1.1 y 1.3 son dos modelos matemáticos diferentes del mismo problema de movimientos (diferente actitud adoptada en su elaboración), en los cuales no se ha hecho ningún tipo de consideración respecto del rozamiento (no fueron realizadas diferentes consideraciones respecto de los parámetros que se hacen intervenir en el modelo).

Del ejemplo anterior, consecuentemente con lo antes estudiado, se desprende claramente la multiplicidad de modelos matemáticos que son posibles elaborar, referidos al mismo problema.

Analizando más profundamente la forma en que pueden ser resueltos los problemas matemáticos, surgen, además de los factores cualitativos y cuantitativos que brindan la posibilidad de la elaboración de múltiples modelos para un solo problema, uno puramente formal; vale decir que, en forma independiente de las consideraciones que puedan hacerse sobre los parámetros y de la actitud adoptada frente al problema en estudio, existe una tercera que consiste en definir en que campo de las matemáticas se situará su expresión formal; se concluye entonces que, existe la posibilidad de clasificar a los modelos matemáticos de muy diferentes maneras, según los tipos. En el presente trabajo se realizará de la manera representada en el cuadro de la figura 1.3.



- Figura 1.3 -

Sobre los diferentes significados que existen entre los denominados modelos exactos, por un lado y los llamados modelos aproximados, por el otro, ya se han dado referencias más que suficientes, las que permitirán diferenciarlos perfectamente; por lo cual, se hará hincapié en la clasificación de los modelos aproximados, los cuales serán prioritariamente utilizados en la formalización analítica de la materia. Es necesario aclarar también que, existen métodos que por su originalidad no son factibles de encasillar totalmente, dentro de ninguno de los puntos de la clasificación realizada.

1.6.1.- MÉTODOS GRÁFICOS

Mediante esta denominación se designarán todos aquellos métodos que, valiéndose de dibujos realizados con precisión y en cierta escala, permiten obtener los resultados deseados. Las aludidas escalas, pueden ser decimales, logarítmicas o mixtas según convenga, de acuerdo a la expresión que se desea representar; por ejemplo, la curva:

$$y = c\,x^n$$

viene representada por una recta en un gráfico logarítmico. La operación que se ha descripto, que permite transformar una curva en una recta mediante un oportuno cambio de escalas, recibe la denominación de **ANAMORFÓSIS**.

Los métodos gráficos son muy utilizados en ingeniería toda vez que sea necesario obtener rápidamente la solución del problema en cuestión y no sea preciso alcanzar resultados con elevado grado de precisión o, cuando las condiciones de trabajo no permiten utilizar otros métodos o equipamiento más sofisticado; principalmente, en problemas que involucran ecuaciones algebraicas, trascendentes, integraciones, interpolaciones, etc.

Entre los métodos gráficos que se pueden encontrar, existen algunos que se limitan al uso de escalas exclusivamente; éstos constituyen una parte de la matemática aplicada denominada **NOMOGRAFÍA**, y los gráficos por ella elaborados reciben el nombre de **ÁBACOS** o **NOMOGRAMAS**.

Los métodos gráficos, por las características de esta materia, no serán objeto de estudio en el presente curso.

1.6.2.- MÉTODOS DE ENUMERACIÓN

Bajo esta denominación se clasifican todos aquellos métodos destinados a resolver problemas de tipo combinatorio con un relativamente bajo número de alternativas o soluciones posibles, estudiando en forma sistemática y exhaustiva todas y cada una de ellas. En algunos métodos incluidos en esta categoría, con el fin de reducir el número de soluciones posibles, se estudia con elevado rendimiento, un subconjunto particularmente importante de las mismas. El subconjunto así obtenido recibe el nombre de **CRÍTICO**.

Un ejemplo típico de método clasificado como de enumeración, es el método de **PROGRAMACIÓN POR CAMINO CRÍTICO** (o **PERT**, con sus múltiples variaciones), cuya metodología de trabajo consiste en descomponer toda la potencial complejidad de una voluminosa operación, en una cantidad determinada de tareas elementales o simples y con ellas formar, articuladamente, una **RED** o **ÁRBOL** orientado.

De esta *RED o ÁRBOL* que, en definitiva, es analizado y estudiado en su totalidad, se identifica un subconjunto particularmente importante de *Tareas*, las cuales reciben el nombre de *CRÍTICAS*.

1.6.3.- MÉTODO DE TANTEOS

La posibilidad cierta de contar con importantes sistemas de cómputo electrónico digital (incluyendo las hoy potentes computadoras personales), permite realizar una gran cantidad de ensayos, en tiempos relativamente breves, con costos razonablemente bajos, y por ende, encarar la solución de problemas por medio de *tanteos programados y sistemáticos* que, de otra manera demandarían un gran esfuerzo de cálculo, lo que impediría su aplicación, sobre todo cuando ésta se realiza manualmente. Esta manera de resolver problemas, que hoy constituye toda una metodología, es conocida desde la antigüedad sin haber alcanzado gran aceptación hasta el advenimiento de las computadoras.

Los métodos de tanteos se prestan particularmente para resolver algunos tipos de ecuaciones algebraicas, trascendentes y diferenciales, tanto ordinarias como en derivadas parciales; barrido de superficies de resultados, en simulación, etc.

Durante el desarrollo de la materia se estudiarán algunos métodos de resolución basados en este tipo de modelos.

1.6.4.- MÉTODOS DE DISCRETIZACIÓN

Estos métodos consisten, básicamente, en el reemplazo de los valores de las variables continuas de distintas expresiones del álgebra y del análisis matemático por valores discretos; o bien, la sustitución de funciones, representadas por curvas, mediante poligonales compuestas de un número finito de segmentos, mayor o menor, según la precisión que se desea alcanzar en el procesamiento, para la obtención de la solución.

En términos generales, es posible afirmar que, cuanto más cerca se encuentran, uno del otro los valores discretos que se toman de la variable, mayor será la precisión alcanzada en el resultado final; dicho de otro modo, cuantos más lados posea la poligonal que sustituye a la función continua, más precisa será la solución aproximada obtenida.

Un ejemplo típico de modelo clasificado como de *Discretización* es el de calcular integrales mediante trapecios inscriptos, tal como se ha expuesto en el ejemplo del punto 1.3, correspondiente al estudio de problemas de velocidad, presentado al principio de este mismo capítulo.

También pertenecen a esta categoría, todos los métodos que utilizan diferencias finitas, elementos finitos, etc.

Los procedimientos de Discretización, están destinados a resolver problemas combinatorios con infinitas soluciones posibles, o un gran número de ellas, u otros reducibles a estos.

1.6.5.- MÉTODOS DE SIMULACIÓN

Se denomina de esta manera, todo procedimiento artificial y abstracto que hace aparecer como real aquello que no lo es, no limitándose a producir vagas analogías teóricas de algún fenómeno real, sino, por el contrario, a reproducir un conjunto de características precisas que son propias del estado real, con el objeto de imitar dicho estado en todos, o casi todos sus aspectos.

Debido a que, invariablemente, estos sistemas utilizan números generados aleatoriamente, con el objeto de obtener valores con qué simular las variables del modelo, también se designa la simulación con el nombre de **MÉTODOS DE MONTECARLO**, haciendo una clara referencia al casino de aquella ciudad europea, que alberga decenas de mesas destinadas al juego de la ruleta las que, originalmente, eran destinadas a producir los números aleatorios, citados al principio de este párrafo, para ejecutar la simulación.

En la actualidad se han desarrollado métodos simples y sistemáticos para producir los números aleatorios mediante computadoras electrónicas digitales, los cuales, en rigor de verdad no son estrictamente aleatorios; por lo cual, toda serie de números generados de esta manera es clasificada como de **SEUDOALEATORIA**.

El método de simulación para la resolución de problemas, juntamente con el de enumeración cobran especial importancia en un importante campo de la matemática aplicada, denominado **INVESTIGACIÓN OPERATIVA**; no serán, por lo tanto, objeto de estudio en el presente curso.

1.6.6.- MÉTODOS ITERATIVOS

Se denomina **PROCESO ITERATIVO** a todo método mediante el cual, a partir de un valor inicial arbitrario y un proceso uniforme determinado, se obtiene una sucesión infinita de aproximaciones que tienden, al menos teóricamente, al valor de la solución deseada. Es válido, pues, para este tipo de modelos, todo lo dicho para algoritmos infinitos en lo que se refiere a la convergencia y su velocidad.

Los procesos iterativos requieren para su implementación realizar las etapas descriptas a continuación:

1.- *Definir la solución mediante una función dada por un algoritmo aplicable a una solución aproximada,*

2.- *Aplicar el algoritmo a una solución inicial razonablemente aproximada, o arbitraria,*

3.- *Utilizar el resultado obtenido luego de la aplicación de lo establecido en el paso anterior, para aplicar nuevamente el algoritmo en forma sistemática, y*

4.- *Dar por finalizado el procedimiento cuando, de la diferencia de los resultados obtenidos para la solución, en dos pasos sucesivos, no se obtiene una cifra superior en valor absoluto, a una magnitud positiva, previamente elegida.*

La primera etapa consiste en la elaboración de un algoritmo que, en términos generales, siendo x la variable del problema, puede ser expresado mediante una ecuación del tipo:

(1.4) $$x = f(x)$$

En la segunda etapa, se adopta una solución inicial $x=x_0$, que pudo haber sido elegida de manera completamente arbitraria o bien, proveniente de alguna consideración o estimación relacionada con el fenómeno o variable en cuestión (física, económica, geométrica, etc.), según el problema en estudio. En este último caso, es posible mejorar el algoritmo y acortar los tiempos de procesamiento.

Para ejecutar la tercera etapa se aplica la ecuación (1.4) haciendo $x=x_0$, con lo cual se obtiene un nuevo valor mejorado de la solución, que será designado con x_1. Iterando el procedimiento indicado se obtiene una sucesión de aproximaciones:

$$x_0 \rightarrow x_1 = f(x_0) \rightarrow x_2 = f(x_1) \rightarrow \ldots \rightarrow x_{k+1} = f(x_k) \tag{1.5}$$

donde cada iteración del procesamiento del modelo se denomina **PASO** y el resultado obtenido en cada paso **APROXIMACIÓN**. Si se tratara del *k-ésimo* paso de la iteración será la **APROXIMACIÓN DE ORDEN K.**

La cuarta etapa está constituida por una parada artificial o programada que, teóricamente, se puede realizar luego de llegar a una aproximación tan cercana de la solución verdadera como se desee. Este punto será tratado con mayor profundidad y detalle en el tema correspondiente a las **APROXIMACIONES NUMÉRICAS** y, como no es posible predecir de antemano o, a priori la cantidad de pasos o aproximaciones que serán necesarias realizar, aplicando sistemáticamente la formulación dada por la expresión (1.5) para llegar a la solución deseada, se volverá a hacer referencia al mismo, en cada método estudiado, si resultase oportuno.

Los procedimientos así descriptos son particularmente importantes por cuanto un gran número de los métodos numéricos estudiados a lo largo del curso son procesos iterativos.

Dada la gran importancia que reviste el tema y, para comprender mejor la naturaleza de los procesos iterativos se expone el siguiente:

Ejemplo 1.2.- Sea el problema que consiste en tener que calcular la raíz cuadrada de un número racional N.

Solución: Para resolver el presente problema, podría ser utilizada la muy conocida fórmula de iteración:

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{N}{x_n}\right) \tag{1.6}$$

en la cual, para $N=2$, y dando inicio al procesamiento utilizando como valor de la solución inicial $x_0 = 1$, se obtiene la sucesión:

$$x_1 = \frac{1}{2}\left(1 + \frac{2}{1}\right) = 1{,}5$$

$$x_2 = \frac{1}{2}\left(1{,}5 + \frac{2}{1{,}5}\right) = 1{,}41666\ldots$$

$$x_2 = \frac{1}{2}\left(1{,}41666\ldots + \frac{2}{1{,}41666\ldots}\right) = 1{,}414215686\ldots$$

y así sucesivamente.

En su momento se verá que la expresión (1.6) se obtiene mediante la aplicación de un método iterativo específico de resolución numérica de ecuaciones, denominado método de **NEWTON-RAPHSON**.

♦♦♦ ●●● ♦♦♦

Llegado a este punto, cabe hacer una aclaración fundamental, desde el punto de vista de la codificación o programación de un algoritmo iterativo en algún lenguaje de programación. En relación con ello, se ha clasificado a los algoritmos iterativos en un solo grupo; vale decir, sin realizar ningún tipo de distinciones, siendo que, en la realidad existen diferencias que las justifican.

En este orden de cosas, es necesario diferenciar los *Procedimientos Iterativos Estrictos,* de aquellos denominados **PROCEDIMIENTOS RECURSIVOS**.

En matemática se da el nombre de *Recursión* a la técnica de definir una función en términos de sí misma. El ejemplo más común y a la vez más sencillo que se conoce de una función definida en forma recursiva es el *Factorial* de un número, el cual para argumentos enteros positivos resulta:

$$n! = n\,(n-1)! \quad ; \quad n > 0$$

siendo 0! = 1. Así, el factorial de n es definido en términos de $(n\text{-}1)!$, el cual, a su vez, es definido en términos de $(n\text{-}2)!$, ... hasta llegar a factorial de 0, que se define explícitamente con el valor uno.

Como puede apreciarse, desde el punto de vista de los modelos matemáticos, no existen diferencias fundamentales entre los algoritmos estrictamente iterativos y los recursivos, pero, el problema se presenta en el momento de la codificación en algún lenguaje de programación, dado que, existen lenguajes que permiten la programación de estos algoritmos recursivos *MEDIANTE EL USO ANIDADO DE FUNCIONES DEFINIDAS POR EL USUARIO*, mientras que, otros lenguajes no la permiten.

1.7.- CONVERGENCIA

Se ha estudiado que, tanto en los procesos algorítmicos, en general, como en los específicamente iterativos, para que resulten útiles, deben brindar, en cada paso, aproximaciones cada vez mejores; en cuyo caso, si los sucesivos pasos producen aproximaciones que se acercan cada vez más a la solución verdadera, se dice que el proceso algorítmico o iterativo es **CONVERGENTE**; en caso contrario, se denomina **DIVERGENTE**. Esta definición básicamente intuitiva que se ha dado, puede ser formalizada más racionalmente, tal como se verá.

1.7.1.- EL PROBLEMA GENERAL

Para poder hablar de estimaciones cada vez mejores, es necesario disponer de algún parámetro, como un evaluador de la medida de la discrepancia o error de cualquier aproximación en particular. Suponiendo que la respuesta correcta de un determinado problema es v, y que, $x_1; x_2; \ldots; x_k$ son las aproximaciones sucesivas obtenidas, respectivamente, en el primero, segundo, ... *k-ésimo* pasos.

Es posible entonces, considerar las diferencias sucesivas entre el valor verdadero de la solución y las distintas aproximaciones obtenidas, dadas por la expresión:

$$|x_k - v|$$

como una medida de cuán cerca se encuentra x_k de v.

Es natural, según lo estudiado, que un proceso algorítmico o iterativo llegue a una primera aproximación, y luego a otra, y así sucesivamente. De esta manera, resulta lógico considerar que estos procesos generan una serie de estimaciones más y más precisas a medida que avanzan; pero, se pretende aún más que lo descripto, de ellos; se pretende que las aproximaciones lleguen tan cerca de la verdadera solución, como se quiera.

Recordando la definición de límite estudiada en el Análisis Matemático y considerando que E representa la aproximación que se desea alcanzar, entonces:

Definición: *" Si dado un número E positivo y arbitrario, por pequeño que sea, es posible encontrar un número* $\delta=\delta(E)$ *tal que, para todo* $k>\delta$, *resulta que:*

$$|x_k - v| < E$$

entonces se dice que la sucesión $x_1; x_2; \ldots; x_k$ *tiende al límite* v *cuando el valor de* $k \to \infty$ *"*

Tomando como base la definición dada de límite de una sucesión, se está ahora en condiciones de dar la correspondiente, más racional y estricta, del tema en estudio, y es la siguiente:

Definición: *" Los procesos algorítmicos o iterativos que generan sucesiones de números, como resultado de aproximaciones sucesivas que cumplen con la definición de límite dada, se denominan* ***CONVERGENTES****. "*

Para cada uno de los métodos numéricos que se expongan, en particular, durante el desarrollo de la materia, será necesario estudiar y demostrar su convergencia si ésta no se encuentra, de algún modo, asegurada.

1.7.2.- GRADO DE CONVERGENCIA

En el punto anterior se ha tratado el tema de la convergencia, en el cual se precisó que para que un proceso algorítmico o iterativo pudiera ser apto para resolver un problema determinado debería ser convergente; es decir, los valores obtenidos como soluciones intermedias, tiendan al valor real; en caso contrario, se puede asegurar que no resolvería el problema en cuestión.

Sin embargo, se espera más aún que la condición expresada en el punto anterior, la que debería considerarse solo como una *Condición Necesaria*; vale decir, para que un proceso algorítmico o iterativo sea verdaderamente útil, es menester que se acerque al verdadero valor de la solución mediante las aproximaciones sucesivas, pero, además, que lo haga con cierto grado de rapidez.

Según lo dicho, es necesario saber algo más sobre la convergencia, y es la rapidez o velocidad con que el algoritmo o proceso iterativo converge a la solución. En cada caso particular es posible medir la rapidez con que un proceso converge; esta rapidez o velocidad recibe el nombre de **GRADO DE CONVERGENCIA**.

Con el fin de ilustrar sobre los efectos del grado de convergencia de un algoritmo, veamos el siguiente:

Ejemplo 1.3.- Sea tener que calcular el valor de la constante absoluta *PI* con n cifras decimales exactas.

Solución: Para ello, es posible utilizar la expresión de *PI* dada por la serie:

$$\pi = 4\left(1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\cdots\right) \tag{1.7}$$

deducida a partir de un desarrollo en serie de *TAYLOR*, la que puede expresarse en forma de sumatoria:

$$\pi = 4\sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1} \tag{1.8}$$

Sin embargo, para obtener seis dígitos correctos de *PI*, es necesario iterar al rededor de 40.000 términos de la serie. Esto, obviamente, resulta poco práctico y difícil, incluso con computadoras electrónicas. Si fueran necesarios, por caso, 12 dígitos correctos, se requerirían unos 15.000.000 de términos de la serie.

♦♦♦ ●●● ♦♦♦

Como puede apreciarse, este algoritmo, si bien es convergente, resulta inútil a todas luces para obtener el valor de *PI*, debiéndose elaborar otros métodos al efecto. Cabe mencionar que existe un sinnúmero de métodos destinados a calcular el valor de *PI*.

Para cada uno de los métodos numéricos que se estudien, aún cuando su convergencia se halle asegurada, será necesario precisar sus condiciones respecto al grado o velocidad de convergencia, con el fin de estimar su posible utilidad y practicidad, o no, frente a cada problema en particular. .

Es imposible desde todo punto de vista dar reglas generales sobre el número de términos que se deben tomar o de las iteraciones que deberán realizarse durante el procesamiento, según se trate de un algoritmo o de un proceso iterativo, respectivamente, para que pueda ser considerado como útil y práctico. Será necesario estudiar cada caso en particular.

♦♦♦ ●●● ♦♦♦

Capítulo 2

Aproximaciones Numéricas

2 . - APROXIMACIONES NUMÉRICAS

Introducción. Números Exactos y Aproximados. Clasificación de los Errores. Errores por ABSTRACCIÓN. Errores por INHERENCIA. Errores por TRUNCAMIENTO. Errores por REDONDEO. Tipos de Redondeo. Redondeo Inferior, Simétrico y Aleatorio. El EPSILON de la computadora. Error Total y Propagación de los Errores.

2.1.- INTRODUCCIÓN

Prácticamente, todos los números utilizados en los distintos algoritmos que serán desarrollados durante el presente curso son números reales o complejos, desde un punto de vista matemático estricto; vale decir, son concebidos como fracciones decimales infinitas o mediante cortaduras de *Dedekind*. Para los fines computacionales tales números no son directamente utilizables, en razón de la infinita cantidad de dígitos decimales que poseen, debiéndose por tanto, aproximarse mediante otra forma de números de un tipo muy particular, conocidos como **FRACCIONES TERMINALES FINITAS**, los que tienen un número finito de cifras decimales.

El proceso de reemplazar un número real por un número racional; más precisamente por una fracción terminal finita, más las otras causas estudiadas en el capítulo anterior (diferencia entre el sistema real y su modelo matemático; imposibilidad de procesamiento infinito, etc.), produce ciertas diferencias computacionales entre los resultados verdaderos, obtenidos del sistema estudiado y aquellos derivados del cálculo, mediante la aplicación de algún método numérico sobre un modelo matemático determinado.

Estas diferencias, inevitables desde todo punto de vista en toda aplicación numérica y que, reciben el nombre genérico de **ERRORES**, una vez producidas, se propagan a todo lo largo de la parte restante del cálculo, durante cuyo proceso su influencia sobre el error es generalmente amplificada; aunque, bajo ciertas condiciones o por el cambio del modelo o algoritmo, puede ser atenuada e, incluso disminuida. El estudio de todo lo referente a este tema será realizado en el presente capítulo.

2.2.- NÚMEROS EXACTOS Y APROXIMADOS

La clasificación que será realizada en el presente capítulo, y al solo efecto del cálculo numérico o aproximado, es propia de la materia y difiere de la que se acostumbra hacer, tanto en la Matemática con fines generales, como en la Física, para describir los que se producen en las mediciones.

Para establecer cierto paralelo, es necesario recordar la manera en que se clasifican los números en las ciencias citadas. La primer gran división es la que se realiza en **CONSTANTES** y **VARIABLES**; recibiendo el nombre de constantes, todos aquellos números que no modifican su valor a través del desarrollo de algún cálculo o procesamiento; mientras que, por el contrario, se llaman variables todos aquellos números que no cumplen con la condición expresada.

Las constantes, a su vez, son clasificadas en **CONSTANTES ABSOLUTAS** y **RELATIVAS**. Siendo las absolutas, aquellas que poseen un valor determinado e invariable, sean cuales fueran las condiciones del problema en cuestión; tales como los números *PI*; 2 o la constante *e*. Las constantes relativas, por el contrario, son aquellas que tienen la posibilidad de cambiar de valor en diferentes problemas o aplicaciones; dependiendo esto, de sus condiciones particulares. Son ejemplos de constantes relativas, la aceleración de la gravedad *g*, la presión atmosférica (que dependen de la altitud y de ciertas condiciones de la atmósfera donde se desarrolla la acción del problema), etc.

Respecto de la clasificación de las variables, no se aportarán mayores detalles dado la gran amplitud y variedad con se que realiza su subdivisión, tanto en la Matemática como en la Física, las cuales no representan antecedentes importantes, a los efectos de los *Métodos Numéricos*.

Desde el punto de vista de los *Métodos Numéricos*, se mantendrá la subdivisión en constantes o variables, según el caso, de la misma manera y con el mismo criterio expresado en los párrafos anteriores, pero, haciendo la aclaración que esta clasificación no es fundamental, pues, todos los valores tratados son **NÚMEROS** (expresados en forma numérica o literal), los que deben ser considerados y tratados por igual, con una clasificación muy especial, la cual será expuesta a continuación.

2.2.1.- NÚMEROS EXACTOS

Todos aquellos números que, al ser utilizados como parámetros o variables de algún modelo matemático, no sufren ninguna modificación por causas *OPERATIVAS* a través del procesamiento del mismo, reciben el nombre de **NÚMEROS EXACTOS**.

Caen dentro de esta categoría, los enteros o decimales que pueden ser expresados con un máximo de 2 bytes o 16 bits (enteros cortos), 4 bytes o 32 bits (enteros largos y reales de simple precisión); o bien, los que pueden ser expresados con un máximo de 8 bytes o 64 bits (reales de doble precisión).

Dado que un entero es un número que no posee parte decimal, el rango de variación de los enteros cortos está comprendido en el intervalo ($-32.768 < x < 32.767$), debido a que se almacenan como números de 16 bits (-2^{15} ; $2^{15} - 1$).

De la misma manera, se puede concluir que los enteros largos, debido a que se almacenan como números de 32 bits o 4 bytes, están comprendidos en el intervalo dado por ($-2.147.483.648 < x < 2.147.483.647$); vale decir, comprendidos en (-2^{31} ; $2^{31} - 1$). Son ejemplos de números exactos: 3; 27.846.763; $^1/_2$; 324.675.825 E - 8; 0,47587321; etc.

2.2.2.- NÚMEROS APROXIMADOS

La manera de manejar los números, tanto enteros como reales, lo mismo en simple como en doble precisión cuando se encuentran fuera de los intervalos definidos en los párrafos anteriores, no resulta tan simple y deberemos remitirnos al manejo particular que, sobre este tema realiza cada compilador o lenguaje de programación en combinación con el microprocesador (*CPU*) utilizado. De todos modos, el avance de la tecnología y la velocidad de los cambios que se registran en la ciencia informática, puede descalificar todo lo dicho aquí en unos pocos años; no obstante, existen invariantes los que serán estudiados a continuación. Ver el ÉPSILON DE LA COMPUTADORA, más adelante, en este mismo capítulo.

Es necesario hacer notar que, entre las causas operativas mencionadas al principio, no se encuentran, precisamente, las variaciones propias que podrían sufrir los números a causa de su funcionalidad (variables), sino las que inevitablemente produce el hardware en aquellos casos en que le resulta imposible soportar al número en su totalidad. P. ej. cuando el resultado de una operación arroja un número que necesita más de 64 bits para ser almacenado, tales como $^1/_3$; 157^{99} ; 1.453!; etc.; o bien, cuando deben ser almacenados números tales como π, e, y otros similares. En ambos casos se trata de números que poseen infinitas cifras decimales, o al menos, son muy grandes.

En los casos mencionados en el párrafo anterior , es la máquina quien, intrínsecamente, comete el *Error*. Se estudiará a continuación en qué consiste, cómo se genera y la manera de minimizarlo.

La causa fundamental de errores en una computadora es atribuible a la diferencia que inevitablemente existe entre un número a representar y su real representación en la computadora, si aquel tuviera infinitos dígitos o cifras decimales, o al menos una cantidad muy grande, dado que, las posibilidades de soporte que brinda el hardware son siempre limitadas.

Es necesario comprender que la mayor parte de los números que se utilizan en una computadora no son exactos, y que, además, la representación que puede hacerse de ellos no es continua. P. ej., en la PC IBM original (CPU Intel 8088), el número positivo más pequeño que podía ser representado es 2,9...E-39 ($2{,}9... \cdot 10^{-39}$). Hoy en una i486 DX II es 1,08...E-324; donde los puntos suspensivos indican entre 8 y 19 dígitos, según el compilador y la precisión (simple o doble) utilizada. Por lo tanto, es absolutamente imposible representar números entre 0 (cero) y 2,9...E-39, en la PC original o, 1,08...E-324, en la i486 DX II.

Para lograr un mayor entendimiento y fijar mejor los conceptos vertidos, veamos el siguiente:

Ejemplo: 2.1.- Considerar la expresión general que permite determinar la longitud l de una circunferencia, en función del radio R de la misma:

(2.1) $$l = 2 \pi R$$

Solución: En este ejemplo es posible observar que, el segundo miembro está formado por el producto de tres números cuyas características, a los efectos de las aproximaciones numéricas, son diferentes.

Por una parte, se encuentra el 2 que, al ser utilizado en el producto no sufre ningún tipo de modificación por causas operativas; esta clase de números reciben el nombre de **NÚMEROS EXACTOS**, como se ha dicho. Por otra parte, está el número *PI* que, si bien es cierto se trata de una constante absoluta, para la matemática pura, desde el punto de vista estrictamente computacional, no opera como un número exacto, dado que contiene infinitas cifras decimales, pudiendo en la práctica hacerse uso, tan solo de una cantidad muy limitada de ellas.

Finalmente, se encuentra R que, por provenir del resultado de una medición no puede ser exacta, sino que, solo se puede expresar con una cierta aproximación que depende del grado de precisión del instrumento de medición utilizado en su estimación y de la habilidad del operador, entre otros factores; estos tipos de números, al igual que *PI*, se denominan **NÚMEROS APROXIMADOS**.

♦♦♦ ●●● ♦♦♦

Se debe notar que, a la luz de las aproximaciones numéricas, la clasificación de números realizada en los párrafos anteriores, en exactos y aproximados, se solapa o superpone completamente con la clasificación ortodoxa realizada por la matemática pura, en constantes y variables, dado que, y a título de simple ejemplo, deberán ser distinguidos los números $^1/_4 = 0{,}25$ de $^1/_3 = 0{,}333...$, donde, el primero es un número *exacto*; mientras que, el segundo es un número *aproximado*. Ambos números, para la matemática pura, son clasificados como constantes absolutas.

Es el momento de realizar una observación de suma importancia. Si fuera necesario aplicar la expresión (2.1) y habiendo realizado la medición del radio con un instrumento que permite apreciar solo hasta el orden de los centímetros, el que, por caso, resultó ser $R = 3.27$ m. Aplicando la expresión (2.1) y tomando el valor de *PI* con distintas cantidades de cifras decimales, se obtiene, de esta manera, el valor de la longitud de la circunferencia en distintas magnitudes, tal como sigue:

$l = 2 \times 3{,}27 \times 3$	$= 19{,}62$
$l = 2 \times 3{,}27 \times 3{,}1$	$= 20{,}274$
$l = 2 \times 3{,}27 \times 3{,}14$	$= 20{,}535\ 6 \approx 20{,}54$
$l = 2 \times 3{,}27 \times 3{,}142$	$= 20{,}548\ 68$
$l = 2 \times 3{,}27 \times 3{,}1416$	$= 20{,}546\ 064$
$l = 2 \times 3{,}27 \times 3{,}14159$	$= 20{,}545\ 998\ 6$

Se puede apreciar que la aplicación de la fórmula permite obtener como medida de la longitud de la circunferencia un número decimal con tantas cifras como se deseen, con tal de tomar el valor de *PI* con la cantidad de cifras decimales suficientes. Es posible comprobar, sin embargo, la imposibilidad de obtener la longitud buscada con mayor aproximación que el centímetro, pues, como será estudiado más adelante, todas las cifras posteriores a la segunda, no revisten ninguna exactitud.

Otro tanto ocurre cuando se toma el valor de *PI* con menor precisión que la de R, a partir del cual se obtienen resultados con menor aproximación que el centímetro. Se llega a la conclusión, entonces, que el mejor resultado es el obtenido cuando las aproximaciones de los elementos que intervienen en una expresión son homogéneas; vale decir, son las mismas (renglón subrayado del ejemplo), cosa que en la práctica, por lo general, no sucede con frecuencia.

Esto, que se acaba de presentar en forma totalmente intuitiva, será formalizado analíticamente a lo largo del presente capítulo.

2.3.- CLASIFICACIÓN DE LOS ERRORES

Previamente de la clasificación prometida se intentará una definición intuitiva del vocablo ERROR, la que será ampliada y profundizada más adelante. En este orden de cosas veamos la siguiente definición:

Definición: *" Se denomina con el término genérico de **ERROR**, a la diferencia que existe entre el valor verdadero de una magnitud determinada y otro valor aproximado de ella. "*

Vale decir, en el caso de tener que cuantificar una magnitud cualquiera X, y llamando X_v al verdadero valor de la magnitud en cuestión, y X_c a su correspondiente valor aproximado, resulta la expresión analítica de la estimación del error **E**, según la definición dada:

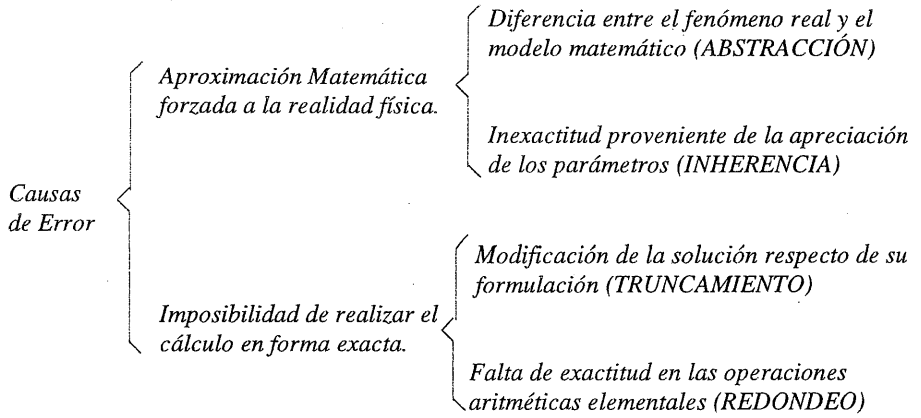
(2.2) $$\mathbf{E} = X_v - X_c$$

El origen de la diferencia dada por la expresión (2.2) puede ser debida a distintas razones, las que serán estudiadas a continuación.

En el capítulo anterior se ha dicho que los métodos tratados en el presente curso eran aproximados y que muy raramente los resultados obtenidos a partir de un método numérico coinciden con los valores exactos. Esta diferencia, denominada error, es una combinación de muchas causas, como se expone en el cuadro 2.1, tratando de dejar en claro también, sus efectos a medida que se progresa en el desarrollo de un algoritmo o proceso iterativo.

Los errores considerados en este capítulo no deben ser confundidos con los denominados *ERRORES GROSEROS* en las mediciones físicas, los cuales representan una equivocación en el algoritmo, modelo o diagrama de flujo que es utilizado para obtener la solución del problema y que, pueden ser superados solamente mediante la corrección debida, de aquellos.

Usualmente y en condiciones normales, los errores son una consecuencia del método empleado y del equipamiento utilizado para resolver cierto problema y son, generalmente, inevitables desde todo punto de vista, pero, observando ciertas reglas precisas pueden ser minimizados.



- Cuadro 2.1 -

Se ha estudiado en el capítulo anterior que la resolución de un problema pasa, inevitablemente, por una serie de etapas, tales como modelización, optimización, procesamiento, etc. Todas ellas colaboran, en mayor o menor medida, para que la solución calculada o aproximada se aleje de la real; este inevitable acontecimiento será estudiado a continuación.

2.3.1.- ERRORES POR ABSTRACCIÓN

El primer paso que deberá darse para resolver un problema, luego del estudio y formulación precisa del mismo, es la etapa de desarrollo de un modelo matemático. Esta etapa consiste en describir las condiciones del problema en estudio, mediante una simplificación o abstracción de la realidad que hace posible, en términos materiales, su posterior tratamiento.

La diferencia entre el fenómeno real o verdadero (sistema), y su descripción analítica simplificada (modelo) es la primera y principal causa de error. A estos errores se los designa como **ERRORES POR ABSTRACCIÓN** y son, salvo en casos muy simples o triviales, en los cuales se trabaja directamente con el sistema o fenómeno real, inevitables desde todo punto de vista.

Es importante notar que, los errores por abstracción son introducidos por una sola y única vez, al principio del procesamiento.

2.3.2.- ERRORES POR INHERENCIA

Se ha hecho, en puntos anteriores, una diferenciación de los factores que intervienen en la expresión (2.1), viendo que ciertas constantes tales como PI que, si bien se trataba de constantes absolutas no podían ser utilizadas como tales, debido a que solo es posible hacer uso de una cantidad limitada de cifras decimales. Por otra parte, se ha visto también, la imposibilidad de determinar la longitud del radio R con toda exactitud, debido a las imprecisiones de los aparatos utilizados en las mediciones y al grado de pericia de quien realiza la medición. Las causas apuntadas obligan a operar siempre con números aproximados; es decir, obligan a operar con cierto error, que existe y es inevitable, por pequeño que este sea.

Si los valores que componen la expresión (2.1) fueran datos de entrada en el procesamiento de algún modelo, desde el comienzo es introducido un error insalvable en el algoritmo que resuelve el problema en cuestión.

Así, las diferencias existentes entre los datos de entrada o parámetros del algoritmo respecto de sus verdaderos valores, reciben el nombre de **ERRORES INHERENTES**. Generalmente son valores desconocidos, pero, siempre puede determinarse un valor máximo estimativo de estos errores, denominado **COTA DE ERROR** y que será utilizado para obtener una cuantificación máxima del error total que puede afectar al resultado.

Llegado a este punto, es oportuno hacer notar que los errores de las mediciones, comúnmente clasificados como sistemáticos y aleatorios, para la disciplina de los métodos numéricos son errores inherentes; vale decir, están englobados en el error que se comete al apreciar los valores de las variables o parámetros que darán inicio a un determinado proceso, provengan de donde provengan.

En definitiva, como es fácil deducir, salvo en aquellos casos de problemas muy simples y que no requieran de valores complejos los errores inherentes son inevitables desde todo punto de vista y se introducen por única vez al principio del procesamiento del modelo matemático.

2.3.3.- ERRORES POR TRUNCAMIENTO

Seguidamente será estudiada una tercer fuente de errores que, al contrario de las ya detalladas, no se introduce al iniciar el procesamiento, sino que, se produce a lo largo del mismo. Para aclarar conceptos veamos el siguiente:

Ejemplo 2.2.- Sea el problema de tener que calcular el valor del número e^x, mediante su correspondiente desarrollo en serie de *TAYLOR*.

Solución: Se sabe que, mediante la fórmula de *TAYLOR* es posible desarrollar en serie el número e^x, siendo:

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \ldots = \sum_{k=1}^{\infty} \frac{x^{k-1}}{(k-1)!} \tag{2.3}$$

Dado que, a los fines computacionales, resulta completamente imposible utilizar los infinitos términos que incluye de la expresión (2.3), al realizar el cálculo mediante los n primeros términos del desarrollo dado, en realidad el valor que se calcula es una cierta $y(x)$, donde:

$$y(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \ldots + \frac{x^{n-1}}{(n-1)!} \cong e^x \tag{2.4}$$

Como puede apreciarse, aunque desde el punto de vista estrictamente teórico es posible calcular e^x en forma exacta, la realidad impone otros términos completamente diferentes.

♦♦♦ ●●● ♦♦♦

La diferencia obtenida $e^x - y(x)$ causada por despreciar en el desarrollo en serie de *TAYLOR* con infinitos términos, que expresa exactamente el valor de e^x, todos aquellos términos desde el orden $n+1$ en adelante, recibe el nombre de **ERROR POR TRUNCAMIENTO**.

Dado que, el uso de todos los términos contenidos en un algoritmo infinito presupone el requerimiento de una cantidad infinita de tiempo de procesamiento es, desde todo punto de vista, imposible de llevarse a cabo, siendo por tanto, el error por truncamiento también inevitable.

El ejemplo que acaba de darse, referido a la determinación del valor de e^x mediante el desarrollo en serie de *TAYLOR*, es un caso típico de algoritmo infinito en el cual la variable es el orden del término que está calculando.

Cabe destacar que los procesos iterativos, estudiados en el capítulo anterior, también dan origen a errores por truncamiento debido a la inevitable necesidad de detener el proceso en algún momento, despreciando de esta manera, un número muy grande, sino infinito, de iteraciones que darían lugar, al menos teóricamente, a un resultado exacto.

A pesar de lo inevitable de su aparición, cabe la posibilidad de que el error por truncamiento sea reducido cuanto se quiera, con solo hacer intervenir más términos del algoritmo infinito o, lo que es lo mismo, iterar un mayor número de veces, si se trata de un proceso iterativo, para obtener el valor del resultado con mayor precisión.

Frecuentemente, la magnitud del error por truncamiento depende del tamaño que se establezca para el incremento, el que será denotado indistintamente por h ó Δx, los cuales son utilizados en muchos algoritmos infinitos y procesos iterativos. En estos casos, este hecho induce, naturalmente, a expresar los errores por truncamiento en función de los incrementos descriptos. Más precisamente, para funciones de una sola variable independiente, es usual escribir:

$$\mathbf{E} \equiv (h^n) \qquad o \qquad \mathbf{E} = O(h^n) \tag{2.5}$$

para expresar una cota del error (máximo error) cometido al estimar un o resultado particular.

Las expresiones (2.5), que se leen: *el error máximo cometido (o cota del error) es del orden de* h^n, y significa que el valor de $\mathbf{E}$ nunca será mayor que el valor absoluto de h^n.

Si, por ejemplo, el error por truncamiento en un proceso determinado es del orden de h^3, y si el cálculo es realizado dos veces, la primera con $h=k$, y la segunda con $h=k/2$, el error por truncamiento de la primera corrida resulta: $\mathbf{E} < k^3$; mientras que, en la segunda corrida será: $\mathbf{E} < k^3/8$. Sintetizando, el error por truncamiento de la segunda corrida será, en general, menor que una octava parte del error cometido en la primera corrida.

Se utilizará también, la notación $o(h^n)$, que admite como lectura: de orden superior a h^3, tal como se hace con los infinitésimos en el análisis matemático, para designar cotas de errores, siempre que resulte apropiado.

2.3.4.- ERRORES POR REDONDEO

Una cuarta fuente de error proviene del modo en que los números son tratados, tanto durante el cálculo manual como internamente por la computadora, según sea aplicable. En el primer caso, los cálculos se realizan con números racionales, expresados en notación decimal; en el segundo, las computadoras generalmente se valen de la forma de notación científica.

En ambos casos, son soportados un número determinado de cifras significativas que depende exclusivamente de la decisión del calculista, o bien, de las posibilidades de soporte de los circuitos de la máquina empleada, respectivamente.

El error producido por la limitación de los números a una cierta cantidad de dígitos significativos recibe el nombre de **ERROR POR REDONDEO**. Es obvio que la magnitud del error por redondeo depende particularmente de la decisión del calculista o de la computadora utilizada, según el caso, y solo puede ser minimizado mediante el uso de una mayor cantidad de dígitos decimales en los cálculos. Con el objeto de aclarar ideas, veamos el siguiente:

Ejemplo 2.3.- Supóngase tener que utilizar una computadora cuyo soporte admite un máximo de cuatro dígitos, con el cual se desea realizar el producto de los números 7,379 por 9,138 ; los que se suponen exactos; vale decir, libres de errores inherentes. Determinar el error por redondeo.

Solución: El producto de los números dados es de 67,4293302 el que tiene nueve dígitos significativos y por lo tanto no pueden ser almacenados por la computadora en cuestión. En este caso, la máquina deberá redondear el resultado tomando los primeros cuatro dígitos: 67,42; o, en el mejor de los casos, en la forma tradicional de redondeo: 67,43 ; según lo permita su propio sistema operativo.

♦♦♦ ●●● ♦♦♦

2.3.5.- TIPOS DE REDONDEO

En el ejemplo presentado en el punto anterior, se ha hecho mención a dos formas diferentes de redondeo, la primera, donde se tomaban las cuatro primeras cifras directamente, y la segunda, en la cual se operaba sobre el cuarto dígito. En rigor de verdad, existen varias maneras de realizar el redondeo. Seguidamente se expondrán las principales y se calcularán cotas de error para ellas.

2.3.5.1.- Redondeo Inferior

La manera más simple de realizar el redondeo, en el caso de que los dígitos contenidos en cierto número superen la capacidad de almacenamiento de la computadora, es ignorarlos totalmente. Este tipo de redondeo donde *"se deja caer las cifras"* recibe el nombre de **REDONDEO INFERIOR** y es cada vez más utilizado debido a la automaticidad de su realización. No requiere ningún tipo de programación en el sistema operativo, no utilizando, en consecuencia, tiempo alguno de procesamiento; se puede decir, que es de aplicación automática.

Otra razón que inclina a los diseñadores de sistemas operativos a decidir su utilización, es el constante y sistemático abaratamiento del costo de los elementos que constituyen, tanto la memoria de masa como los cache de las unidades centrales de proceso, lo que hace más productivo y rendidor eludir los incrementos de tiempo, debido al exceso de procesamiento que debería realizarse después de cada operación para realizar el redondeo, a costa de un relativamente pequeño incremento del precio de la unidad, por el aumento de la capacidad de soporte de la computadora.

En este caso, es posible considerar como *cota del error por redondeo inferior*, para un sistema operativo que soporta t cifras decimales:

$$\mathbf{E} \leq 10^{-t} \tag{2.6}$$

como es obvio.

2.3.5.2.- Redondeo Simétrico

En el otro caso de redondeo, utilizado en el ejemplo 2.3, se realiza la operación incrementando en una unidad la cifra de la extrema derecha, última a retenerse, si la que le sigue (primera de la izquierda, a despreciarse) es igual o mayor que cinco; en caso contrario, se ignoran directamente, los dígitos que le siguen al último de la extrema derecha, a retenerse, al igual que en el caso del redondeo inferior.

El tipo de redondeo descripto en el párrafo anterior recibe el nombre de **REDONDEO SIMÉTRICO** y es el más comúnmente utilizado en el cálculo manual por resultar más preciso que el redondeo inferior.

Una *cota de error por redondeo simétrico*, para computadoras que soportan t dígitos significativos es:

(2.7) $$E \leq 0{,}5 \cdot 10^{-t}$$

como podría esperarse.

2.3.5.3.- Redondeo Aleatorio

Un tercer tipo de redondeo es aquel que se realiza sumando una unidad a la cifra de la extrema derecha, que no es despreciada, si la que le sigue o, ella misma, es par; en caso contrario, se ignoran los siguientes dígitos. Esta forma de realizar el redondeo es denominada **REDONDEO ALEATORIO** y es utilizado por algunas computadoras, que lo incluyen en su sistema operativo. Resulta más práctico y fácil de programar que el redondeo simétrico, en razón de que no es necesario sostener, ni aún temporariamente, la cifra decimal siguiente a la última que es soportada.

Resulta obvio que, según la teoría de las probabilidades, en la mitad de los casos sumará una unidad a la última cifra y en la otra mitad, despreciará directamente, compensándose las diferencias obtenidas.

Puede demostrarse que, la *cota del error para el redondeo aleatorio*, en computadoras que soportan un total de t dígitos significativos es:

(2.8) $$E \leq 0{,}5 \cdot 10^{-t}$$

como era de esperarse. Se debe hacer notar que en el caso del redondeo aleatorio, el error que se comete en su aplicación, resulta el mismo que en el del redondeo simétrico.

2.3.6.- EL ÉPSILON DE LA COMPUTADORA

En este orden de cosas, un concepto de fundamental importancia en el estudio del soporte de los números es el dado en la siguiente:

Definición: *" Recibe el nombre de **ÉPSILON DE LA COMPUTADORA**. y se lo designa con ε , a la magnitud del intervalo que media entre el 1 (uno) y el menor número mayor que 1 (uno), distinguible de 1 (uno), que puede representarse en la memoria de la computadora. "*

De la misma manera, es posible calcular el intervalo mencionado en la definición anterior, la que arrojará diferentes valores según el compilador utilizado y el microprocesador de la computadora empleada en su determinación. Un programa en lenguaje C++, capaz de calcular el *Épsilon* de la computadora, en su parte esencial, es:

```
{
E=1;
ciclo:
  if (E+1) > 1 {
    then printf(E);
    else goto fin;
  end if;
  }
E=E/2 : goto ciclo;
fin:
}
```

El último número impreso por este sencillo programa, es el ε de la computadora. El intervalo entre cualquier número real R y el siguiente, distinguible de él, puede representarse por:

$$\varepsilon \cdot R$$

Como puede apreciarse, la brecha entre un número real R y el siguiente, representable, va ensanchándose continuamente en proporcionalidad directa con el número involucrado y el *Épsilon*. Cuando resulte oportuno se suministrarán detalles sobre la manera en cómo influye esto en la precisión de los distintos métodos estudiados.

Siendo que el *Épsilon* de la computadora es la amplitud del intervalo que media entre el uno y el siguiente número representable y distinguible de él, esto significa que ningún número entre 1 y $1+\varepsilon$ puede ser representado en la computadora. en el caso de cualquier número $1+\alpha$, donde $0 < \alpha < \varepsilon / 2$, se redondea a uno; en el caso contrario; vale decir, si $1 > \alpha \geq \varepsilon / 2$ se redondea a $1+\varepsilon$.

De este modo, se puede considerar que $\varepsilon / 2$ es el máximo error posible de redondeo para 1; o sea que, cuando se tiene 1.0 en la memoria de la computadora, el valor original verdadero se encuentra en el intervalo $1 - \varepsilon / 2 < x < 1 + \varepsilon / 2$.

En consecuencia, el error por redondeo implicado en el almacenamiento de cualquier número real R en memoria, es aproximadamente igual a $\varepsilon R / 2$, si el número se redondea por exceso (simétrico o eventualmente, aleatorio) y εR si se redondea por defecto (inferior o eventualmente, simétrico)

Para poder apreciar desde un punto de vista más práctico todo lo estudiado, veamos el siguiente:

Ejemplo 2.4.- Supóngase tener que calcular con tres dígitos decimales exactos, la suma dada por la sucesión de términos:

$$S = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \ldots + (-1)^n \frac{1}{n!} \pm \cdots \tag{2.9}$$

Solución: En este caso, el error por truncamiento que se comete al detener el procesamiento en el *n-ésimo* término es:

$$\mathbf{E} < \frac{1}{(n+1)!}$$

Para que no influya el error por truncamiento, debe tomarse menor que:

$$5 \,.\, 10^{-4}$$

esto se consigue haciendo n=6, pues, 7!=5040, y:

$$\frac{1}{7!} \cong 2 \,.\, 10^{-4}$$

Se debe calcular, entonces, la expresión (2.9) hasta el sexto término:

$$S = 1 - \frac{1}{2} + \frac{1}{6} - \frac{1}{24} + \frac{1}{120} - \frac{1}{720}$$

tomando cuatro cifras decimales en cada sumando, para que resulte:

$$\mathbf{E} < 5 \,.\, 10^{-4}$$

a los efectos que se mantenga dentro del error permitido; con lo cual, resulta: S=0,6319, donde se ha cometido un:

- error por truncamiento $< 2 \,.\, 10^{-4}$
- error por redondeo $< 4 \,.\, 5 \,.\, 10^{-5}$

donde el error por redondeo se encuentra multiplicado solo por cuatro dado que, esa es la cantidad de términos aproximados (los dos primeros son exactos). El error total está dado por la suma de los errores estudiados, y resulta:

$$\mathbf{E} < 4 \,.\, 10^{-4}$$

Considerando, entonces S=0,632 el error será $5 \,.\, 10^{-4}$ menor que el máximo admisible: , y el resultado tiene todas sus cifras exactas.

♦♦♦ ●●● ♦♦♦

Para finalizar, es necesario decir que, en caso de tratarse de una fórmula aproximada donde intervienen, a su vez, números aproximados, habrá que considerar todas las fuentes de error estudiadas, por separado, tal como se ha hecho en los ejemplos vistos y adicionarlos para obtener el error total. Más adelante se estudiará un método sistemático para determinar una cota superior del error total cometido al tener que procesar un algoritmo determinado.

2.4.- ERROR TOTAL Y PROPAGACIÓN DE LOS ERRORES

Según se estudió, en general es posible establecer que los errores por abstracción e inherentes se introducen de una sola y única vez durante el procesamiento de un modelo matemático, coincidiendo con su principio o inicio; es decir, con la entrada de los datos y la adopción de los valores de ciertos parámetros que, por lo general, no constituyen números exactos, sino aproximados, los que, se encuentran afectados de cierto error inherente. Todo ello dentro del contexto de un modelo que a su vez, no es un fiel reflejo de la problemática real; sino el fruto de una mayor o menor simplificación de aquella, con cierto grado de abstracción.

Los errores por redondeo, en cambio, se producen prácticamente en cada operación aritmética elemental que realiza la computadora. Finalmente, los errores por truncamiento son producidos cada vez que en el procesamiento del modelo aparecen procesos iterativos o algoritmos infinitos; o bien, aunque no infinitos, lo suficientemente extensos que resulten imposibles o poco conveniente procesarlos en su totalidad.

Las cuatro fuentes de error estudiadas se adicionan; o sea que, superponen sus efectos individuales para formar el **ERROR TOTAL** que, en procesos o modelos que se aproximan por pasos o iteraciones sucesivas, predisponen al sistema para que en cada una de ellas, el error aumente de magnitud si no se toman una serie de precauciones que, juntamente con cada método serán estudiadas.

Esta particularidad del efecto conjunto de las cuatro fuentes de error o error total, que consiste en variar (generalmente, incrementándose) a medida que se avanza en el desarrollo del procesamiento de un modelo matemático, se denomina **PROPAGACIÓN DE LOS ERRORES.**

Ejemplo 2.5.- Estudiar la magnitud de un error que se propaga multiplicándose en una centésima parte de la unidad con que se está trabajando, en cada paso del procesamiento realizado.

Solución: Si la magnitud de un error se propaga multiplicándose en una centésima parte de la unidad con que se está trabajando, en cada paso, entonces, después de 100 iteraciones la magnitud acumulada del error es

$$(1{,}01)^{100} \cong 2{,}7$$

veces mayor que la original. Si bien puede considerarse aceptable, es dable pensar que cien iteraciones es un número pequeño si se tiene en cuenta la naturaleza de las computadoras electrónicas digitales y las características de los modelos matemáticos, hoy en uso normal.

En este orden de ideas, la magnitud del error que se estaba considerando, luego de 1000 pasos es de:

$$(1{,}01)^{1000} \cong 21.000$$

veces mayor que el error original; y luego de 2000 pasos es de:

$$(1{,}01)^{2000} \cong 400.000.000$$

veces mayor.

Es fácil advertir que el efecto mostrado, probablemente, invalide todo resultado después de este número de iteraciones.

♦♦♦ ●●● ♦♦♦

Debido a lo explicado; es decir, a la propagación de los errores, no es válido suponer que el efecto de pequeños errores introducidos al principio o, con anterioridad en el cálculo, más el efecto de los errores producidos durante el cálculo, se mantendrán pequeños a través de él, aunque, en ocasiones pueda lograrse este objetivo. Esto lleva a la siguiente definición intuitiva:

Definición: *" Todo método numérico, aplicado a un modelo matemático determinado y procesado dentro de cierto intervalo específico, recibe el nombre de* ***ESTABLE*** *si, a pesar del efecto de la propagación de los errores, éstos se mantienen acotados dentro de ciertos límites previamente fijados, hasta el momento de completar el procesamiento del modelo y obtener el valor de la solución buscada.*

El problema de la estabilidad en los métodos aproximados es verdaderamente complejo, pues, no solo depende del modelo que es procesado, sino también del método numérico empleado para resolverlo, como así también, del intervalo en el cual se está trabajando; lo que, en definitiva, hace que el método utilizado, si bien estable dentro de ciertos intervalos de cálculo, puede no serlo en otros.

Independientemente de la dificultad de su estudio, por no ser objeto de interés específico en la materia, cuando resulte oportuno, se harán consideraciones relativas a este tema, para cada método en particular.

♦♦♦ ●●● ♦♦♦

Capítulo **3**

Errores en las Operaciones

3 . - ERRORES EN LAS OPERACIONES

Introducción. Error Absoluto y Error Relativo. Error en los Algoritmos Infinitos. Cálculo de Errores mediante Fórmulas Diferenciales. Error en las Operaciones Elementales. Gráficos de Procesos. Metodología. Aplicaciones. Conclusiones. Problema Inverso del Cálculo de Errores.

3.1.- INTRODUCCIÓN

En el capítulo anterior se ha estudiado que, todos los números utilizados en los distintos algoritmos que son desarrollados durante el presente curso constituyen números reales o complejos, desde un punto de vista matemático estricto; vale decir que, son concebidos como fracciones decimales infinitas o mediante cortaduras de Dedekind. Para los fines de su utilización en computadoras electrónicas digitales, tales números no son directamente utilizables, en razón de la infinita cantidad de dígitos decimales que poseen, debiéndose por tanto, aproximarse mediante otra forma de números de un tipo muy particular, conocidos como **FRACCIONES TERMINALES FINITAS**, las cuales tienen un número finito de cifras decimales.

Se ha estudiado también que, en el proceso de reemplazar un número real por un número racional; más precisamente por una fracción terminal finita, más las otras causas estudiadas en el capítulo anterior (diferencia entre el sistema real y su modelo matemático; imposibilidad de procesamiento infinito, imposibilidad de soporte infinito, etc.), se producen ciertas diferencias entre los resultados verdaderos, obtenidos del sistema real, y aquellos derivados del cálculo, mediante la aplicación de algún método numérico sobre un modelo matemático determinado.

Estas diferencias, inevitables en toda aplicación numérica y que, reciben el nombre genérico de **ERRORES**, una vez producidas, se propagan a todo lo largo de la parte restante del cálculo, durante cuyo proceso su influencia sobre el error es generalmente amplificada; aunque, bajo ciertas condiciones o por el cambio del modelo o algoritmo, puede ser controlada e, incluso disminuida.

En el presente capítulo serán estudiados los errores que se producen por las causas mencionadas en segundo término; vale decir, los que se cometen en las operaciones aritméticas y, consecuentemente, en sucesiones más o menos largas de ellas, denominadas *ALGORITMOS*.

3.2.- ERROR ABSOLUTO Y ERROR RELATIVO

Hasta ahora se ha estudiado el error considerando que es la diferencia entre el *valor verdadero* X_v de una magnitud determinada, pero *desconocida*, y un *valor aproximado* de ella, que se denotará con X_c, *conocido*; y que, naturalmente, difiere en magnitud, del anterior. En rigor de verdad, este tipo de error recibe el nombre de **ERROR ABSOLUTO** y se representa por:

(3.01) $$\mathbf{E_a} = X_v - X_c$$

Como se ha mencionado, por lo general, no es conocido el valor verdadero X_v de una magnitud cualquiera a estimarse, por lo que, consecuentemente, no es posible conocer el valor verdadero del error absoluto, expresado por (3.01), que existe siempre cuando se quiere estimar una magnitud mediante un número aproximado. Por ello, y dado que, lo único que se puede estimar es un *límite máximo* del error que es posible cometer en dicha estimación. A este se lo utiliza como una aproximación del error absoluto y recibe el nombre de **COTA DEL ERROR ABSOLUTO**. Dicho en otras palabras, se adoptan las *peores condiciones* posibles para estimar el error absoluto.

Dada la imposibilidad de confusión entre los términos definidos; vale decir, en virtud de la no ambigüedad entre *Error Absoluto* y *Cota del Error Absoluto*, se utilizará, indistintamente, el término error absoluto para denominarlos y la notación $\mathbf{E_a}$, para denotar ambos errores.

Generalmente, el así definido error absoluto $\mathbf{E_a}$, no contiene toda la información suficiente para calificar o ponderar la precisión de los cálculos o mediciones por no poseer en sí mismo el marco de referencia adecuado para este propósito; en otras palabras, dos errores absolutos de igual magnitud, cometidos en la ponderación de dos valores de muy diferentes ordenes de magnitud, habla a las claras de dos mediciones muy distintas en cuanto a su calidad o grado de precisión.

En virtud de lo dicho en el párrafo anterior, se vuelve necesario desarrollar un método que permita evaluar la calidad o precisión de los cálculos o mediciones. Es muy razonable, entonces, ponderar la calidad del error por simple cociente entre el error absoluto cometido al estimar una ponderación, ya sea por cálculo o medición, respecto del valor verdadero de la magnitud que se trata de representar.

El error así definido, recibe el nombre de **ERROR RELATIVO**. Utilizando idéntica notación para los valores intervinientes, el error relativo viene dado por la expresión:

(3.02) $$\mathbf{E_r} = \frac{X_v - X_c}{X_v}$$

Dado que, tanto el verdadero valor de una magnitud, como el del error absoluto que se comete al tratar de apreciarla (el que figura en el numerador), son generalmente cantidades imposibles de conocer, se tomará como una aproximación del error relativo, al valor dado por la expresión:

(3.03) $$\mathbf{E_r} = \frac{\mathbf{E_a}}{X_c}$$

donde $\mathbf{E_a}$ representa una cota del error absoluto que, de tener pequeñas variaciones respecto del error absoluto, al igual que X_c respecto de X_v, la diferencia entre las expresiones (3.02) y (3.03) no tiene mayor influencia en la estimación del error relativo.

Es importante hacer notar, la necesidad que en la práctica tiene indicar si un error es absoluto o relativo, a menos que el significado esté claramente determinado con la notación o la expresión de la oración, a los efectos de evitar posibles ambigüedades o confusiones que, inevitablemente desembocarían en resultados seguramente inválidos, desde todo punto de vista.

3.3.- ERROR EN LOS ALGORITMOS INFINITOS

Antes de proseguir con el estudio de la metodología de la determinación de los distintos tipos de errores, es oportuno recordar que al exponer, tanto el concepto general de Procesos Iterativos como de Algoritmos Infinitos, se ha mencionado la ineludible necesidad de incluir en aquellos algoritmos, una manera de detener el procesamiento, en forma artificial, cuando para dos pasos consecutivos, en relación con la precisión con que se está trabajando, no se detecta un cambio singularmente sustancial en sus respectivas aproximaciones.

Designando mediante X_{i-1} ; X_i a las aproximaciones obtenidas en los pasos *(i-1)-ésimo* e *i-ésimo*, respectivamente, se hará detener el procesamiento por **DIFERENCIA ABSOLUTA**, si:

(3.04) $$|X_i - X_{i-1}| \leq \varepsilon$$

donde ε es una constante positiva y arbitraria, llamada **TOLERANCIA DEL ERROR** y, puede ser considerada como el máximo error admisible en la solución deseada.

La única condición que se debe imponer al valor del parámetro ε es que debe ser siempre mayor que el error total que puede cometerse al calcular el valor de la magnitud estimada; de otra manera, el proceso no resultará convergente. En el caso del cálculo realizado mediante computadoras, el ε establecido, debe ser mayor que el producto de X_i por el *Épsilon* de la computadora utilizada.

La expresión (3.04) es utilizada, generalmente, cuando se conoce de antemano el orden de magnitud de la variable X; vale decir, no existe peligro alguno de cometer grandes errores de apreciación en la estimación del valor de X.

En el caso contrario al explicado en el párrafo anterior; vale decir, cuando el orden de magnitud del valor de X es desconocido, se debe hacer uso de expresiones en términos relativos, para lo cual, será necesario detener el procesamiento por **DIFERENCIA RELATIVA**; es decir, cuando:

$$\left|\frac{X_i - X_{i-1}}{X_i}\right| \leq \varepsilon \tag{3.05}$$

En esta última expresión, son válidas todas las acotaciones y aclaraciones que han sido realizadas previamente, al definir la detención del proceso para el caso de la diferencia absoluta.

Si el problema en cuestión, en vez de involucrar un solo y único resultado X, consiste en el cálculo de los n elementos de un vector $X(j)$, en el cual es, naturalmente, $j=1;2;...;n$ es posible hacer uso de los criterios ya estudiados, generalizando de la siguiente manera:

a) En caso que todas las componentes $X(j)$ del vector sean de distinto orden de magnitud, pero, conocidos, es conveniente detener el proceso por **DIFERENCIA ABSOLUTA INDIVIDUAL** tal como se propone en la expresión siguiente:

$$\left|X(j)_i - X(j)_{i-1}\right| \leq \varepsilon(j) \quad ; \quad j=1;2;\ldots;n \tag{3.06}$$

vale decir que, el procesamiento se detiene cuando las n expresiones involucradas en la expresión (3.06) son simultáneamente satisfechas.

b) Cuando se está en presencia del caso en el cual todas las componentes $X(j)$ del vector son del mismo orden de magnitud, también conocido, es preferible detener el proceso mediante la aplicación de la expresión por **DIFERENCIA ABSOLUTA GLOBAL**, como sigue:

$$\frac{1}{n}\sum_{j=1}^{n}\left|X(j)_i - X(j)_{i-1}\right| \leq \varepsilon \tag{3.07}$$

fórmula que permite reducir mucho la extensión del cálculo aritmético, y por ende, los tiempos de procesamiento.

c) En caso que las variables involucradas sean de distinto orden de magnitud, pero, desconocidos, es muy conveniente detener el proceso por **DIFERENCIA RELATIVA INDIVIDUAL**, como se muestra:

$$\left|\frac{X(j)_i - X(j)_{i-1}}{X(j)_i}\right| \leq \varepsilon(j) \quad ; \quad j=1;2;\ldots;n \tag{3.08}$$

con la que es posible controlar cada componente $X(j)$ del vector X, individualmente.

d) Por último, si se estuviera en presencia del caso en que todos los componentes $X(j)$ del vector fueran del mismo orden de magnitud, y además, este orden de magnitud fuera desconocido, es oportuno utilizar para detener el procesamiento, la expresión denominada **DIFERENCIA RELATIVA GLOBAL**; la que viene dada por:

$$\frac{1}{n}\sum_{j=1}^{n}\left|\frac{X(j)_i - X(j)_{i-1}}{X(j)_i}\right| \leq \varepsilon \tag{3.09}$$

equivalente a la expresión (3.07), válida para algoritmos donde todos los componentes tienen un mismo orden de magnitud.

Es necesario recordar que las fórmulas estudiadas son válidas para procesos iterativos o algoritmos infinitos convergentes; pero que, si el grado de convergencia del mismo es muy lento puede darse la condición de parada o detención encontrándose el valor o valores aproximados de las variables muy lejos de los valores verdaderos de la solución del problema. Si este fuera el caso, las expresiones estudiadas operarían como condiciones necesarias, pero, no suficientes, debiéndose realizar otro tipo de controles cuando se duda sobre las bondades del algoritmo utilizado, principalmente, en lo que respecta a la velocidad de convergencia.

3.4.- CALCULO DE ERRORES MEDIANTE FÓRMULAS DIFERENCIALES

Considérese un algoritmo en el cual interviene una sola variable independiente, dada por la expresión:

(3.10) $$y = f(x)$$

para la cual, considerando una variación h de la variable independiente, se experimenta una variación k en el valor de la función. Si esta es diferenciable, su incremento k se podrá expresar mediante:

(3.11) $$k = f'(x)\,h + \varepsilon\,h$$

donde, es posible considerar que:

x : es la verdadera magnitud del valor de entrada,
h : es el error inherente de la variable x,
k : es la consecuencia en el resultado, del error h [1], y finalmente
ε : es un infinitésimo que tiende a cero cuando h tiende a cero. [2]

Planteando así las cosas y teniendo en cuenta que $\varepsilon\,h$ es un infinitésimo de orden superior, tanto a h como a k, por lo que resulta despreciable. Como consecuencia, el error cometido en el procesamiento de un algoritmo, debido al error inherente de la variable, puede expresarse por:

(3.12) $$k \cong f'(x)h$$

como se estudia en el análisis matemático.

El valor de h no incluye el error por redondeo que ocasiona el procesamiento del algoritmo. Es importante notar que, tanto h como k, son errores absolutos. Así mismo, notar que, el símbolo de aproximación es debido a que el valor de k no coincide exactamente con el de dy (diferencial de y), que es el realmente representado en la expresión (3.11). Por otra parte, tampoco el valor de x que se tome será el verdadero, pues, se lo desconoce, sabiéndose solamente que está afectado del error h.

[1] Es necesario notar que, el valor de h no incluye el valor por redondeo que ocasiona el procesamiento del algoritmo.
[2] También es importante considerar que, tanto h como k, representan errores absolutos.

La utilización de la expresión (3.12) se mostrará mediante el siguiente:

Ejemplo 3.1.- Si es necesario determinar el error cometido en el cálculo del volumen de un cubo, cuyo lado mide l = 7,3 cm., medido con un error menor de medio milímetro, siendo el volumen $V = l^3$.

Solución: Será entonces:

$$k = 3\,l^2\,h = 3\,.\,7{,}3^2\,.\,0{,}05 \cong 8\ \text{cm}^3$$

por consiguiente, el volumen será:

$$V = 7{,}3^3 \pm 8 = 389{,}017 \pm 8\ \text{cm}^3$$

en la cual se obtuvieron, al elevar 7,3 al cubo, tres cifras decimales, ninguna de las cuales se puede retener en el resultado, pues, en el mismo, solo son exactas las cifras de las decenas en adelante; vale decir, lo correcto es escribir:

$$V \cong 389 \pm 8\ \text{cm}^3$$

como podía haberse esperado.

♦♦♦ ●●● ♦♦♦

Es posible extender los conceptos estudiados y sus correspondientes resultados mediante la generalización directa de estos conceptos.

Sea una función y, derivable y con derivada continua; vale decir, diferenciable, la cual depende de diversas variables: x_1 ; x_2 ; ... ; x_n , cada una de las cuales está afectada por cierto error absoluto h_1 ; h_2 ; ...; h_n , entonces se puede escribir:

(3.13) $$k \cong \sum_{i=1}^{n} f_{x_i}(x_1;x_2;\ldots;x_n)\,h_i$$

Dado que resulta imposible poder determinar el signo de cada uno de los términos del segundo miembro, se debe aplicar la desigualdad triangular a la expresión (3.13); resulta entonces, la denominada **FORMULA FUNDAMENTAL DEL CALCULO DE ERRORES**:

(3.14) $$|k| \le \sum_{i=1}^{n} \left|f_{x_i}(x_1;x_2;\ldots;x_n)\right| \cdot |h_i|$$

Más adelante se estudiará en forma exhaustiva la manera de aplicar la fórmula fundamental del cálculo de errores, especialmente en lo que concierne a las operaciones aritméticas, las que, generalmente, componen los algoritmos.

3.5.- ERROR EN LAS OPERACIONES ELEMENTALES

Por constituir los métodos numéricos algoritmos algebraicos o procesos iterativos que, generalmente, solo involucran operaciones aritméticas elementales, no es difícil intuir la importancia fundamental que revisten para este curso los errores que se cometen al realizar estas operaciones elementales con números aproximados y, de la misma manera, como se propagan a través de secuencias de ellas.

En lo sucesivo serán consideradas $x\ ;\ y\ ;\ z\ ;\ldots$ como cantidades aproximadas y $h_x\ ;\ h_y\ ;\ h_z\ ;\ \ldots$ los valores que toman los errores absolutos, los que deberán adicionarse a los anteriores para obtener los verdaderos valores de las magnitudes en cuestión, respectivamente.

Considerando la expresión general (3.13), será:

a) Adición. Sea $z=x+y$ con derivadas $z_x = z_y = 1$; de lo que resulta entonces la siguiente expresión del **ERROR ABSOLUTO PARA LA ADICIÓN**:

$$\mathbf{E}_a^{x+y} \cong h_x + h_y \tag{3.15}$$

y aplicando la desigualdad triangular a esta última, resulta:

$$\left|\mathbf{E}_a^{x+y}\right| \le |h_x| + |h_y| \tag{3.16}$$

En este caso específico, el **ERROR RELATIVO** cometido puede calcularse mediante el artificio de multiplicar y dividir por x e y, respectivamente, el primero y segundo términos del segundo miembro, de la siguiente manera:

$$\mathbf{E}_r^{x+y} = \frac{\mathbf{E}_a^{x+y}}{x+y} = \frac{x}{x+y}\left(\frac{h_x}{x}\right) + \frac{y}{x+y}\left(\frac{h_y}{y}\right) \tag{3.17}$$

donde las cantidades encerradas entre paréntesis son los errores relativos que se cometen en las ponderaciones de las magnitudes x e y, respectivamente.

La cota de error absoluto dada por la expresión (3.16), aplicando directamente la fórmula (3.14), podrá ser obtenida para todos los casos que serán estudiados a continuación.

b) Sustracción. Sea $z = x - y$, entonces, como las derivadas son iguales al caso anterior, salvo en el signo, operando de la misma manera que la utilizada para la adición, resulta el siguiente valor del error:

$$\mathbf{E}_a^{x-y} \cong h_x - h_y \tag{3.18}$$

y, análogamente:

$$\mathbf{E}_r^{x-y} = \frac{\mathbf{E}_a^{x-y}}{x-y} = \frac{x}{x-y}\left(\frac{h_x}{x}\right) - \frac{y}{x-y}\left(\frac{h_y}{y}\right) \tag{3.19}$$

Valen en el presente caso, las mismas observaciones realizadas anteriormente para la adición.

c) Multiplicación. En este caso se considera $z = x \,.\, y$, entonces, $z_x = y\ ;\ z_y = x$, de donde es posible deducir:

$$\mathbf{E}_a^{x.y} \cong y \cdot h_x + x \cdot h_y \tag{3.20}$$

considerando el resultado de esta expresión, el error relativo cometido en la multiplicación, será entonces:

$$\mathbf{E}_r^{x.y} \cong \frac{h_x}{x} + \frac{h_y}{y} \tag{3.21}$$

con las mismas consideraciones anteriores.

d) División. Se tomará, en el caso de la división $z = x / y$, de donde pueden deducirse ambas derivadas parciales: $z_x = 1 / y\ ;\ z_y = -\ x / y^2$, luego, el error absoluto de la operación estará dado por:

$$\mathbf{E}_a^{x/y} \cong \frac{h_x}{y} - \frac{x.h_y}{y^2} \tag{3.22}$$

en la cual, tomando denominador común y^2 resulta:

$$\mathbf{E}_a^{x/y} \cong \frac{y.h_x - x.h_y}{y^2} \tag{3.23}$$

la que, dividiendo por x / y, expresa el valor del error relativo que se comete al realizar un cociente:

(3.24) $$\mathbf{E}_r^{x/y} \cong \frac{y.h_x - x.h_y}{x.y} = \frac{h_x}{x} - \frac{h_y}{y}$$

Operando de la misma manera; vale decir, determinando en primer lugar el error absoluto mediante la expresión (3.14) y luego realizando el cociente de ésta sobre el valor aproximado de la expresión que se desea ponderar, es posible deducir cotas de error relativo para otras operaciones.

Es importante comprender claramente el significado de estas fórmulas de error, dada la enorme incidencia que tienen estas expresiones sobre los resultados obtenidos en cualquier procesamiento, máxime, cuando estos involucran una cantidad relativamente grande de operaciones aritméticas elementales, tal como sucede en los algoritmos infinitos y los procesos iterativos.

En principio se parte de valores aproximados $x\ ;\ y$ de las variables, los que contienen los errores inherentes $h_x\ ;h_y$, respectivamente. Los errores pueden ser de cualquier tipo. Los valores $x\ ;\ y$ pueden provenir de resultados experimentales o de mediciones aproximadas, pueden ser, a su vez, los resultados de cálculos previos, efectuados mediante un proceso infinito y por lo tanto, pueden contener errores por truncamiento; también, pueden ser el resultado de operaciones aritméticas previamente realizadas y, por lo tanto, contener errores por redondeo.

De la misma manera, pueden ser el producido de la aplicación de un modelo aproximado; o, en fin, pueden ser los resultados de alguna combinación de alguno de aquellos errores mencionados o de todos a la vez; es decir, contener los cuatro tipos de errores estudiados.

No obstante, sea cual fuere el origen de los errores de las cantidades $x\ ;\ y$, a los efectos de su posterior tratamiento, serán considerados como errores inherentes en el paso o proceso siguiente.

Si bien las magnitudes $x\ ;\ y$ pueden contener intrínsecamente todos los tipos de errores, las expresiones (3.15) hasta (3.24), dan el error en el resultado de cada una de las operaciones indicadas *suponiendo que no existe error por redondeo* en ellas; vale decir, son considerados números exactos, a todos los efectos posteriores del cálculo de los errores. Es lógico entonces, que si se desea conocer el error total; o, en otras palabras, cómo se propaga el error en algún resultado, a otras operaciones aritméticas, se deberá agregar explícitamente el error por redondeo que, como se ha estudiado, puede ser de tres tipos diferentes, dados por las expresiones (2.6) y (2.7), según los casos.

3.6.- GRÁFICOS DE PROCESOS

En puntos anteriores se han estudiado distintas expresiones, las que permiten determinar el error relativo que se comete al efectuar cada una de las operaciones aritméticas elementales como consecuencia de los errores inherentes de los operandos.

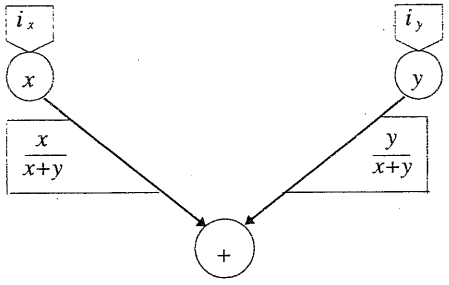
Se han estudiado también, maneras de determinar cotas de errores cometidos en diferentes tipos de redondeo que deben ser efectuados, inexorablemente, por la computadora para poder almacenar los resultados de aquellas operaciones. Se estudió, además, que tanto el error por redondeo como el error por truncamiento, se producen aún en el caso supuesto de que no hubiera errores inherentes en los operandos.

Todas las circunstancias apuntadas, hacen que si se aplican estas simples leyes a expresiones más complejas que las hasta ahora estudiadas, se estaría, seguramente, frente a insalvables dificultades en el manejo de los elementos que intervienen en el cálculo de la propagación de los errores cometidos paso a paso, en la determinación del error total.

Un método sistemático, que permite determinar la propagación de los errores relativos en los datos y de los errores relativos debidos al redondeo, a lo largo de una secuencia determinada de operaciones aritméticas elementales, se encuentra en el denominado **MÉTODO DE LOS GRÁFICOS DE PROCESO**. Esta metodología ha sido sugerida por el **Dr. Kenneth King**, Profesor de la Universidad de Columbia, el que se describe a continuación.

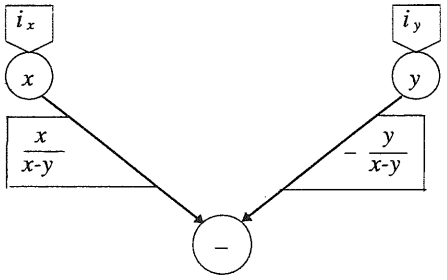
3.6.1.- METODOLOGÍA

A continuación se expondrá la técnica que permite aplicar la metodología basada en los **GRÁFICOS DE PROCESO.** Veamos la siguiente representación:



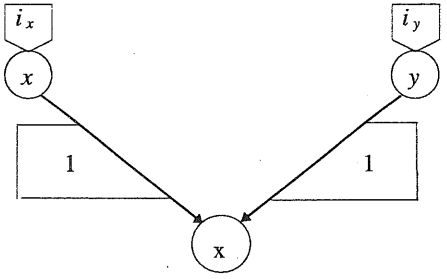
- Figura 3.1 -

El procedimiento mencionado en el punto anterior, se basa en las fórmulas estudiadas y se vale para ello, de una representación esquemática de las diferentes operaciones que constituyen la secuencia, para poder así, lograr determinar los límites superiores parciales de los errores relativos propagados, después de cada una de las operaciones que se realizan durante el procesamiento.



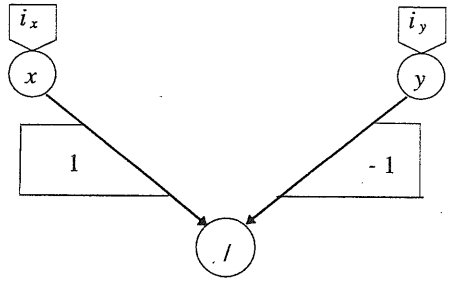
- Figura 3.2 -

El gráfico de proceso consiste, esencialmente, en encerrar dentro de pequeños círculos los valores de los operandos que intervienen en la operación (ver figura 3.1), desde los cuales parten sendas flechas etiquetadas con el factor por el que se habrá de multiplicar al error relativo del operando respectivo (indicado dentro de los pentágonos superiores), para obtener el error relativo de la operación encerrada dentro del círculo al cual conducen las flechas del gráfico. El caso que se halla representado en la figura 3.1 es, obviamente, el de la suma o adición.



- Figura 3.3 -

De la misma manera pueden lograrse los gráficos de proceso de otras operaciones, tal como se muestra en las figuras 3.2 ; 3.3 y 3.4 ; para la sustracción, multiplicación y división, respectivamente.



- Figura 3.4 -

Para mejor entender, veamos el siguiente:

Ejemplo 3.2.- Sea el problema de tener la necesidad de conocer el error relativo de una suma $x+y$, considerando que estos valores están afectados de sendos errores inherentes absolutos h_x; h_y.

Solución: La regla se aplica de la siguiente manera; llamando:

$$i_x = \frac{h_x}{x} \quad ; \quad i_y = \frac{h_y}{y}$$

a los errores relativos inherentes de las cantidades x ; y , respectivamente.

El error i_x aparece en el resultado de la operación multiplicado por el término que identifica la flecha que une x con $+$, es decir:

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

de la misma manera, el error inherente i_y de la variable y, aparece en el resultado de la operación multiplicado por el término que identifica la flecha que une y con "+"; o sea:

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

Finalmente, se deben sumar ambos términos y adicionar el error por redondeo S que se comete al almacenar el resultado. El error relativo total será entonces:

(3.25) $$\mathbf{E}_r^{x+y} = \frac{x}{x+y} i_x + \frac{y}{x+y} i_y + S$$

Una operatoria análoga puede emplearse para determinar el error relativo total, que se comete en otras operaciones elementales, sin olvidar de adicionar siempre el error por redondeo que siempre está presente al almacenar algún resultado en la computadora, sea este parcial, o total. El error calculado de esta manera, puede considerarse como un error inherente a los efectos de utilizar los resultados obtenidos en otras operaciones que deban ser realizadas a posteriori.

3.6.2.- APLICACIONES

Debido a lo fundamental de la importancia que reviste el presente tema, para una más acertada elección y una mejor y más correcta aplicación de los métodos numéricos que serán estudiados a lo largo del curso, se verán algunas aplicaciones mediante un cierto número de ejemplos, elegidos ex-profeso, para cubrir las distintas alternativas que, en general, se pueden presentar al operar con algoritmos de modelos matemáticos que incluyan números aproximados.

Ejemplo 3.3.- Determinar el error relativo que se comete al ponderar la cantidad dada por la expresión:

(3.26) $$U = (x+y).z$$

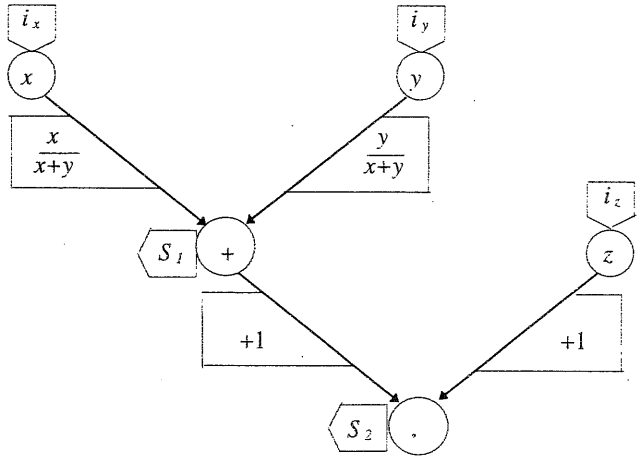
en la que los números $x\ ;\ y\ ;\ z$ son cantidades aproximadas que contienen los errores relativos inherentes $i_x\ ;\ i_y\ ;\ i_z$, respectivamente, y que, se supondrán, menores de $0{,}5\ .\ 10^{-t}$.

Solución: En la figura 3.5 es posible apreciar el gráfico de proceso de la expresión dada por (3.26). Se ha estudiado de qué manera se opera en la primera parte del diagrama; vale decir, la que corresponde a la adición, cuyo resultado está expresado en la fórmula número (2.31).

Operando con el mismo criterio para el producto que le sigue, y que constituye la segunda parte del gráfico de proceso, se concluye que:

(3.27) $$\mathbf{E}_r^U = \left(\frac{x}{x+y} i_x + \frac{y}{x+y} i_y + S_1\right).1 + i_z.1 + S_2$$

en la cual se han denominado $S_1\ ;\ S_2$, a los valores de los errores relativos por redondeo de la adición y la multiplicación, respectivamente, que si se considera que son el resultado de un redondeo simétrico, serán iguales o menores que $0{,}5\ .\ 10^{-t}$, si es que se supone que la computadora utilizado soporta solamente t dígitos decimales.



- Figura 3.5 -

En este caso será:

$$\mathbf{E}_r^U = \left(\frac{x}{x+y} + \frac{y}{x+y} \times 3\right) \times 0{,}5 \times 10^{-t}$$

de donde, por ser la suma de los dos primeros términos encerrados dentro del paréntesis, menor o igual que uno, será:

$$\mathbf{E}_r^U \leq 2 \times 10^{-t}$$

♦♦♦ ●●● ♦♦♦

Otro ejemplo, muy representativo de los cuidados que deben ser tenidos en cuenta, está dado en el siguiente:

Ejemplo 3.4.- Determinar el error total que se comete al ponderar el resultado obtenido por la suma:

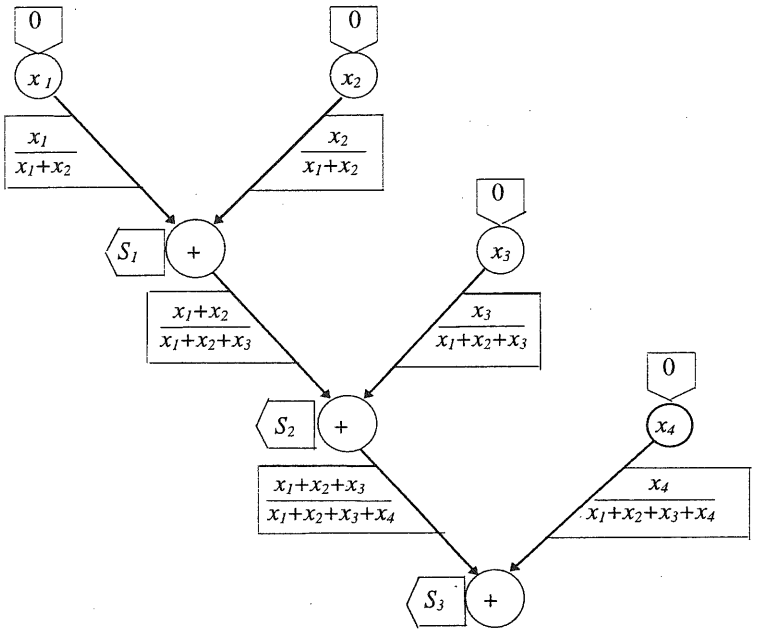
(3.28) $$y = x_1 + x_2 + x_3 + x_4$$

de términos decimales exactos.

Solución: Con el objeto de realizar ciertas consideraciones posteriores se supondrá que los valores de los x_i están libres de errores inherentes y, además que:

$$x_1 < x_2 < x_3 < x_4$$

También se supondrá que, la adición se realiza en el orden establecido; por lo tanto, el gráfico de proceso correspondiente a la expresión (3.28), es el de la figura 3.6.



- Figura 3.6 -

Bajo estas condiciones, la expresión que da el error de la primera adición es:

$$\mathbf{E}_r^{x_1+x_2} = \frac{x_1}{x_1+x_2}.0 + \frac{x_2}{x_1+x_2}.0 + S_1 = S_1$$

dado que, el valor de los errores inherentes de los sumandos es nulo; y por lo tanto, desde un punto de vista estrictamente teórico, la expresión anterior vale:

$$\mathbf{E}_r^{x_1+x_2} = S_1$$

Este último valor opera como error inherente en el paso siguiente, cuyo error relativo es:

$$\mathbf{E}_r^{x_1+x_2+x_3} = \frac{x_1+x_2}{x_1+x_2+x_3}S_1 + \frac{x_3}{x_1+x_2+x_3}.0 + S_2 = \frac{x_1+x_2}{x_1+x_2+x_3}S_1 + S_2$$

el que a su vez debe ser utilizado como error relativo en la formulación analítica del próximo paso.

Por último, aplicando una mecánica similar, y teniendo en cuenta las consideraciones realizadas, resulta la expresión del error relativo:

$$\mathbf{E}_r^{x_1+x_2+x_3+x_4} = \frac{x_1+x_2+x_3}{x_1+x_2+x_3+x_4}\left(\frac{x_1+x_2}{x_1+x_2+x_3}S_1 + S_2\right) + \frac{x_4}{x_1+x_2+x_3+x_4}.0 + S_3 =$$

$$= \frac{x_1+x_2}{x_1+x_2+x_3+x_4}S_1 \frac{x_1+x_2+x_3}{x_1+x_2+x_3+x_4}S_2 + S_3 =$$

Pero esta última expresión, como se sabe, da el valor del *ERROR RELATIVO* cometido al ponderar el resultado del valor de y por medio de la fórmula dada por la expresión (3.28); vale decir, multiplicando esta última por $x_1 + x_2 + x_3 + x_4$ es posible obtener el *ERROR ABSOLUTO* cometido en esta suma. Utilizando el símbolo sumatoria para representar la suma, vale:

$$\mathbf{E}_a^{\Sigma} = (x_1+x_2)S_1 + (x_1+x_2+x_3)S_2 + (x_1+x_2+x_3+x_4)S_3$$

Suponiendo del mismo orden de magnitud a los valores de S_1 ; S_2 ; S_3 , los que serán designados con el nombre genérico de S, resulta:

(3.29) $$\mathbf{E}_a^{\Sigma} = (3x_1 + 3x_2 + 2x_3 + x_4)S$$

Puede demostrarse, por medio de la inducción matemática, que si se trata de n sumandos, exentos de errores inherentes, la expresión (3.29) se transforma en:

(3.30) $$\mathbf{E}_a^{\Sigma} = [(n-1)x_1 + (n-1)x_2 + (n-2)x_3 + \ldots + 2x_{n-1} + x_n]S$$

Mediante consideraciones absolutamente análogas, si se realiza la adición de estos mismos elementos, ordenados en forma inversa; vale decir, de mayor a menor, resulta:

(3.31) $$\mathbf{E}_a^{\Sigma} = (3x_4 + 3x_3 + 2x_2 + x_1)S$$

y, en términos generales:

(3.32) $$\mathbf{E}_a^{\Sigma} = \left[(n-1)x_n + (n-1)x_{n-1} + (n-2)x_{n-2} + \ldots + 2x_2 + x_1\right]S$$

Comparando las expresiones (3.30) y (3.32), se observa que el error será menor en la primera que en la segunda, por intervenir en aquella los sumandos de mayor magnitud con menor peso, lo que indica la gran importancia que adquiere el ordenamiento de los sumandos cuando se debe realizar la adición de n números aproximados, de distinto orden de magnitud, pudiéndose establecer la siguiente:

Regla: *Al sumar varios números de distinto orden de magnitud, es posible minimizar el error total del resultado, ordenándolos en forma creciente; vale decir, de menor a mayor.*

Este sorprendente resultado, por causa de la formación recibida, tanto en la escuela primaria como en la secundaria, se debe al simple hecho que los Métodos Numéricos no operan bajo la **HIPÓTESIS DE PRECISIÓN INFINITA**, hipótesis ésta, implícitamente aceptada en todos los razonamientos y conclusiones de la aritmética y álgebra elementales.

Ejemplo 3.5.- Determinar el error total que se comete al ponderar la suma del ejercicio anterior (expresión 3.28) pero, realizando las operaciones asociando de a dos sumandos y considerando a éstos de igual orden de magnitud.

Solución: Para simplificar la parte resolutiva del ejercicio, se supondrá que los sumandos están libres de errores inherentes tal como se hizo en el ejercicio anterior. Se trata, pues, de realizar:

(3.33) $$y = (x_1 + x_2) + (x_3 + x_4)$$

cuyo gráfico de proceso, es el que se encuentra representado gráficamente en el diagrama de la figura 3.7.

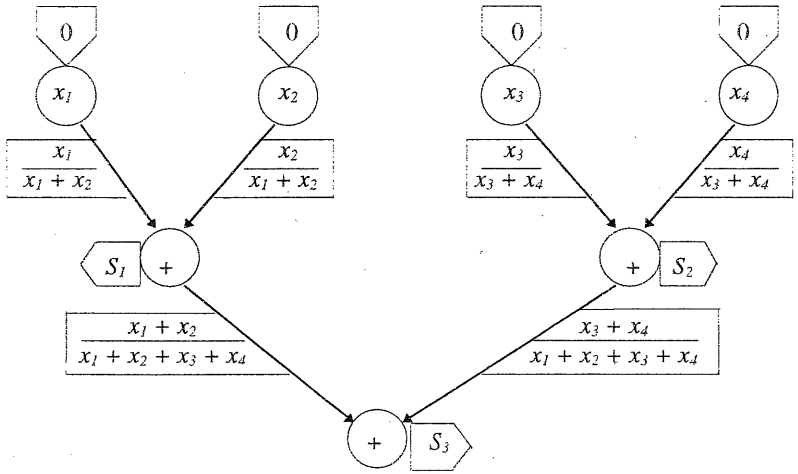
Luego de realizar el primer paso de ambas ramas, resultan los errores relativos:

$$\mathbf{E}_r^{x_1+x_2} = \frac{x_1}{x_1+x_2}.0 + \frac{x_2}{x_1+x_2}.0 + S_1 = S_1$$

y, por la otra rama:

$$\mathbf{E}_r^{x_3+x_4} = \frac{x_3}{x_3+x_4}.0 + \frac{x_4}{x_3+x_4}.0 + S_2 = S_2$$

como muestra el siguiente diagrama:



- Figura 3.7 -

Realizando el siguiente paso, considerando los resultados anteriores como errores inherentes para éste, se obtiene:

$$\mathbf{E}_r^{\Sigma} = \frac{x_1+x_2}{x_1+x_2+x_3+x_4}S_1 + \frac{x_3+x_4}{x_1+x_2+x_3+x_4}S_2 + S_3$$

de donde, considerando los errores por redondeo del mismo orden de magnitud y multiplicando el segundo miembro por la suma de las cantidades en cuestión, resulta el error absoluto:

(3.34) $$\mathbf{E}_r^{\Sigma} = \left(2x_1 + 2x_2 + 2x_3 + 2x_4\right)S = 2\left(x_1 + x_2 + x_3 + x_4\right)S$$

de la que es posible sacar algunas importantes conclusiones, como se estudiará en los siguientes puntos.

♦♦♦ ●●● ♦♦♦

Comparando esta última expresión con (3.29) y (3.31) se concluye que, cuando se trata de sumar varias cantidades de igual orden de magnitud es conveniente agruparlas en sumas parciales de, aproximadamente, igual número de sumandos, a los efectos de realizar el primer paso y luego, efectuar la suma de todos los resultados parciales.

Fundamentando en lo observado a lo largo del presente capítulo, es posible establecer la siguiente:

Regla: *Al sumar una cantidad aproximada de n^2 sumandos, de igual orden de magnitud, es conveniente agruparlos, primeramente, en n sumas de n sumandos cada una, aproximadamente, y luego, sumar los n resultados parciales obtenidos anteriormente.*

Se puede apreciar a través de este ejemplo que, tampoco se cumple la propiedad asociativa de la suma cuando se opera con números aproximados; vale decir, fuera de las condiciones establecidas por la hipótesis de precisión infinita.

3.6.3.- CONCLUSIONES

Los pocos casos expuestos en los ejemplos estudiados, en los cuales se han obtenido resultados realmente inesperados, en la práctica se cuentan por docenas; por ejemplo:

no es lo mismo	que
$x + x + x$	$3\,x$
$x_1 (x_2 - x_3)$	$x_1 x_2 - x_1 x_2$
$\frac{x_1 - x_2}{x_3}$	$\frac{x_1}{x_3} - \frac{x_2}{x_3}$

etc., como puede fácilmente demostrarse (quedando como ejercicio para el lector), dando lugar a pequeñas diferencias en los resultados de las operaciones las que, pueden acentuarse según la diferencia de magnitud entre los valores en cuestión y la complejidad de las expresiones aritméticas intervinientes.

Las poco significativas variaciones que ocasionan las expresiones estudiadas, por contener unas pocas operaciones, pueden convertirse en grandes errores cuando el procesamiento requiere cientos de miles, y a veces millones, de operaciones aritméticas, cuales son los algoritmos infinitos y los procesos iterativos.

En estas situaciones, más cercanas a la realidad, un error pequeño al principio, puede multiplicarse considerablemente a causa de las leyes de propagación estudiadas, dando como resultado cantidades completamente diferentes de las magnitudes reales.

En el capítulo primero se ha estudiado que para resolver un problema con éxito, deben agotarse de manera correcta ciertas etapas. Una de ellas es el análisis de la optimización del modelo matemático ya definido y formalizado analíticamente, el que, desde el punto de vista de las aproximaciones numéricas debe tender, en lo posible, a disminuir los efectos de la propagación de los errores, teniendo como premisas fundamentales, las siguientes:

1.- Volver a escribir las ecuaciones del modelo matemático de manera tal de minimizar el número de operaciones aritméticas que será necesario realizar por paso, para reducir al mínimo los nocivos efectos de los errores por redondeo,

2.- Redistribuir, en lo que se refiere al orden y a la forma, las distintas operaciones incluidas en cada expresión matemática del modelo con el objeto de minimizar el sistemático efecto de los errores en las operaciones aritméticas, y

3.- Cuando las expresiones matemáticas del modelo incluyan sumas y restas de varias cantidades, desarrollarlas de manera tal, que permitan trabajar primero con los números más pequeños.

Teniendo especial cuidado de observar estas simples, pero prácticas reglas de procedimiento, se obtendrán excelentes resultados, aún haciendo uso de métodos de cálculo elementales o poco perfeccionados, desde el punto de vista de la estabilidad de los métodos numéricos.

3.7.- PROBLEMA INVERSO DEL CALCULO DE ERRORES

Mirando el problema del cálculo de errores, cabe hacerse la siguiente pregunta: ¿con qué aproximación deberán tomarse los valores de las variables que intervienen como datos en una determinada función:

$$y = f(x_1 ; x_2 ; \ldots ; x_n)$$

para que sea calculada con un error menor o igual a un E arbitrario, previamente fijado ?

Como puede apreciarse, se trata de resolver el problema inverso al estudiado en el punto correspondiente a los errores inherentes. Este es un problema que se resuelve mediante la aplicación de la fórmula fundamental del cálculo de errores como puede apreciarse en el siguiente:

Ejemplo 3.6.- Se trata de determinar el volumen de un paralelepípedo rectangular cuyas aristas miden, aproximadamente, *x=35 cm.; y=40 cm.; z=45 cm.*, con un error menor que *50 cm.*

Solución: El problema, básicamente, consiste en determinar con qué precisión se deberán medir las aristas del paralelepípedo para mantenerse dentro de las condiciones fijadas por el problema. Siendo:

$$V = x \,.\, y \,.\, z$$

entonces:

$$|k| = |y.z|.|h_x| + |x.z|.|h_y| + |x.y|.|h_z| = 1800|h_x| + 1575|h_y| + 1400|h_z|$$

y admitiendo que los errores de x; y; z son del mismo orden y debiendo ser $|k| \leq 50\,cm^3$, resulta:

$$|h_x| = |h_y| = |h_z| < \frac{50\,cm^3}{4775\,cm^2} \cong 0{,}01\,cm$$

es decir que, solo pudiendo medir los lados del paralelepípedo con una aproximación de al rededor de la centésima de centímetro, es posible lograr el volumen con la aproximación deseada.

Capítulo 4

Raíces de Ecuaciones

4.- RAÍCES DE ECUACIONES

Introducción. Comportamiento de las ecuaciones. Métodos generales de resolución. Método de TANTEOS. Error en el método de TANTEOS. Método del INTERVALO MEDIO. Error en el método del INTERVALO MEDIO. Método de INTERPOLACIÓN LINEAL. Convergencia del método de INTERPOLACIÓN LINEAL. Método de NEWTON-RAPHSON. Convergencia del método de NEWTON-RAPHSON. Error en el método de NEWTON-RAPHSON. Método MIXTO.

4.1.- INTRODUCCIÓN

Uno de los problemas que con mayor frecuencia se presenta en cualquier rama de la técnica o de la ciencia consiste en tener la necesidad de determinar las raíces de una ecuación de la forma:

(4.1) $$f(x) = 0$$

donde $f(x)$ es una función de la variable real x con coeficientes reales. El proceso de resolver la ecuación (4.1) consiste en hallar aquellos valores numéricos de la variable independiente x, denominados **RAÍCES** de la ecuación, tal que reemplazados en el primer miembro la anulan.

Se sabe que el álgebra provee fórmulas de resolución para ecuaciones cuyos primeros miembros son polinomios, hasta el cuarto grado inclusive. Para ecuaciones de grado mayor no existen métodos exactos que las resuelvan [1]. Otro tanto ocurre con las ecuaciones denominadas trascendentes. En este capítulo serán estudiados varios procedimientos que permiten calcular valores **APROXIMADOS** de las raíces.

Se comenzará por definir algunos términos y estudiar el comportamiento general de las ecuaciones, desde el punto de vista de la matemática pura.

---

[1] Teorema de *RUFFINI - ABEL.*

Primeramente es necesario mencionar que no existe sobre este tema una *PANACEA UNIVERSAL*; es decir, un método que aplicado sistemáticamente, sin ningún conocimiento previo de la ecuación a resolver, y sobre todo de sus raíces, resuelva todo tipo de problemas. Este conocimiento previo al que se ha hecho referencia obliga a tratar de agotar ciertas etapas intermedias para lograr los resultados deseados.

Las distintas etapas que, inevitablemente, deben seguirse para resolver una ecuación, a las cuales se ha hecho mención en el punto anterior son:

a) *ACOTACIÓN* de las raíces,
b) *SEPARACIÓN* de las raíces, y
c) *APROXIMACIÓN* de las raíces.

El proceso de **ACOTAR LAS RAÍCES** de una ecuación consiste en determinar un intervalo abierto, denominado **INTERVALO DE ACOTACIÓN**, que contenga a todas las raíces que pudieran interesar. Existen maneras precisas de operar cuando las ecuaciones son algebraicas, para acotar raíces, las que serán estudiadas en el siguiente punto. Cuando se trata de ecuaciones trascendentes no existen métodos generales, el estudio debe realizarse para cada caso en particular.

La práctica de **SEPARAR LAS RAÍCES** equivale a realizar una partición del intervalo de acotación, tal que, en cada subintervalo cerrado se encuentre una y solamente una raíz. Para separar raíces de ecuaciones algebraicas se estudiarán algunos métodos particulares, en los puntos correspondientes. Cuando se trata de ecuaciones trascendentes el estudio debe realizarse especialmente para cada caso, en particular. Este paso es de fundamental importancia, tanto en el caso de ecuaciones algebraicas como trascendentes, para evitar pasar por alto algunas raíces o no identificar raíces múltiples, principalmente.

El paso de **APROXIMAR LAS RAÍCES** es el procedimiento que consiste en calcular el valor numérico de ellas con tanta precisión como se desee. Para poder lograrlo se estudiarán varios métodos de aplicación general; vale decir, útiles tanto para el caso de ecuaciones algebraicas como para las trascendentes. Seguidamente a los nombrados métodos generales, se estudiarán métodos especiales dedicados a ecuaciones algebraicas, los que resultan en un mayor rendimiento, velocidad y sencillez de la operatoria.

4.2.- COMPORTAMIENTO DE LAS ECUACIONES

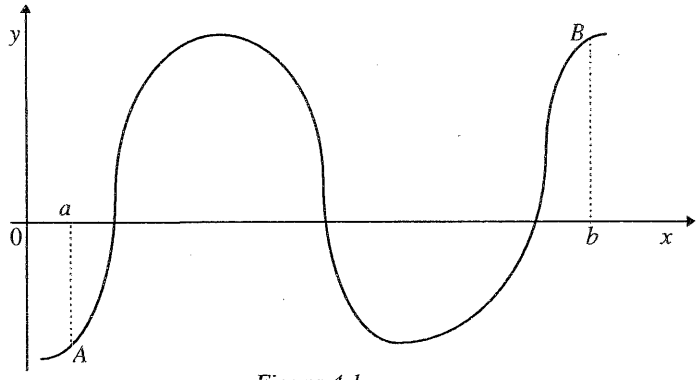
En el punto anterior se ha mencionado la conveniencia de tener cierto conocimiento previo de las ecuaciones que se desean resolver y además, que el mismo se obtiene mediante la acotación y separación previas, de las raíces. Se ha visto también, que tratándose de ecuaciones trascendentes debe ser estudiado cada caso en particular.

Este estudio se realiza teniendo presente una serie de resultados del análisis matemático que facilitan el conocimiento de la ubicación de las raíces. Considérese un intervalo $(a;b)$ en el cual una función $f(x)$ es continua, y sea AB el arco que la representa en coordenadas cartesianas, entonces es posible afirmar que:

I.- *Si $f(x)$ tiene distintos signos en dos puntos de abscisas a y b, se anula por lo menos una vez en $(a;b)$, y en general, un número impar de veces*[2]. Consultar la figura 4.1.

II.- *Si $f(x)$ tiene igual signo en dos puntos de abscisas a y b, o bien se anula un número par de veces en el intervalo $(a;b)$, o bien no tiene ninguna raíz en el mismo.*

Puede apreciarse en la figura 4.2, que entre los puntos $x=c$ y $x=d$, la función no se anula, mientras que, en el intervalo $(a;b)$ se anula cuatro veces.

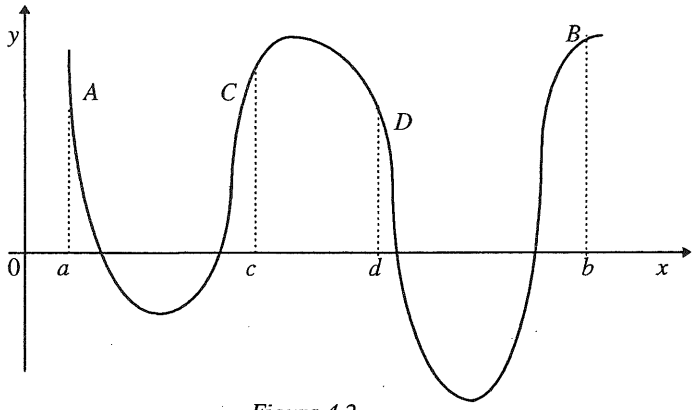


- Figura 4.1 -

III.- *Si $f(x)$ es monótona creciente (o monótona decreciente) en el intervalo $(a;b)$; vale decir que, si $f'(x)$ tiene un signo determinado (más o menos) en todo punto del intervalo, y es $sg\, f(a)$ distinto de $sg\, f(b)$, hay una sola raíz r de $f(x)=0$, mientras que, si los signos de $f(a)$ y $f(b)$ son iguales, con seguridad no hay ninguna raíz.*

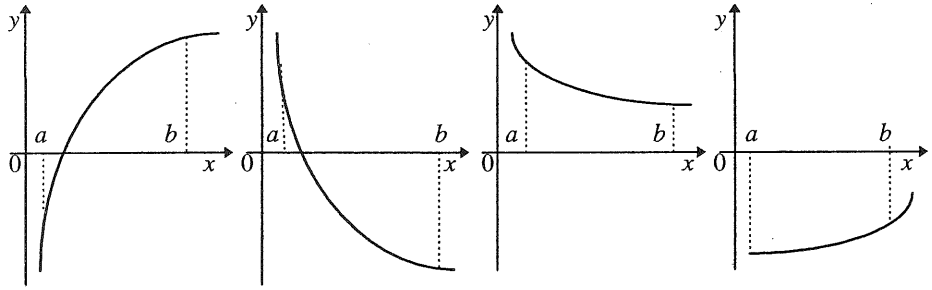
En la figura 4.3, se han representado los distintos casos que da lugar el teorema anterior, los que permiten efectuar la separación de la raíz r del intervalo $(a;b)$; es decir, asegurar cuando en dicho intervalo hay una sola raíz o, no hay ninguna raíz.

[2] Teorema de *BOLZANO*



- *Figura 4.2* -

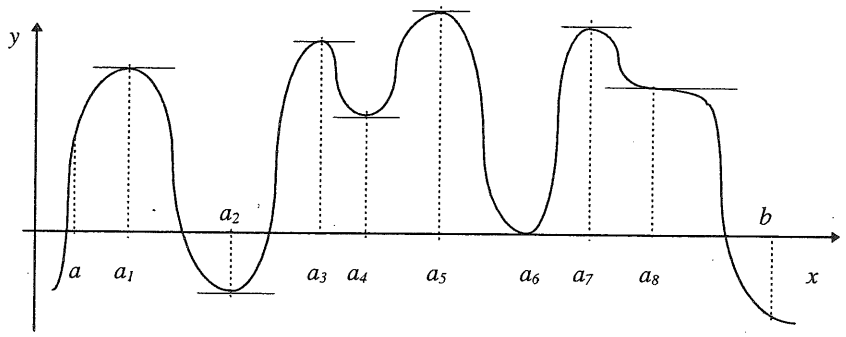
IV.- Entre cada par de raíces consecutivas de f'(x) la función f(x) es creciente o decreciente; vale decir, monótona y de acuerdo a III, tiene una sola raíz o ninguna.



- *Figura 4.3* -

En la figura 4.4, es posible apreciar gráficamente distintas opciones de estos últimos teoremas. Nótese que en a_6 , es $f(a) = f'(a) = 0$, siendo pues, una raíz de tipo especial denominada *RAÍZ MÚLTIPLE*, (en este caso doble) las que se estudiarán más adelante. En el punto a_8 , se presenta otra particularidad; en este caso se trata de un *PUNTO DE ENSILLADURA*.

Con el objeto de profundizar en la práctica del comportamiento de las funciones, se presentan los siguientes:



- *Figura 4.4* -

Ejemplo 4.1.- Sea tener que acotar y separar todas las posibles raíces de la ecuación dada por la función:

$$f(x) = x - \cos x = 0.$$

Solución: En este caso, como el valor de x está comprendido en el intervalo $(-\infty ; \infty)$; mientras que, $\cos x$, en $(-1 ; 1)$, se deduce que para $x \leq -\pi/2$, $f(x)$ es siempre negativa y que, para $x \geq \pi$, $f(x)$ es siempre positiva. Se concluye que:

$$\textit{Intervalo de Acotación} = (-\pi/2 ; \pi)$$

Para poder separar las raíces es necesario estudiar el comportamiento de la ecuación en distintos puntos del intervalo de acotación; para lo cual se construye la siguiente tabla:

x	$-\frac{\pi}{2}$	0	$\frac{\pi}{2}$	π
$f(x)$	$-\frac{\pi}{2}$	-1	$\frac{\pi}{2}$	$\pi - 2$

- *Tabla N° 4.1* -

Con respecto a los signos, puede apreciarse en la tabla N° 4.1 que:

$$sg\, f(-\pi/2) = sg\, f(0)$$

entonces, por lo expresado en los puntos *II* y *III*, se sabe que no existe ninguna raíz en el intervalo $(-\pi/2 ; 0)$. En el siguiente subintervalo,

$$sg\, f(0) \neq sg\, f(\pi/2)$$

donde, por las mismas razones dadas para el subintervalo anterior, existe una única raíz. Mediante análogas consideraciones es posible asegurar que en $(\pi/2 ; \pi)$, la ecuación no tiene ninguna raíz. Se debe notar que resulta indispensable considerar si $f'(x) = 1\text{-}sen\ x$ es monótona o no, en cada subintervalo.

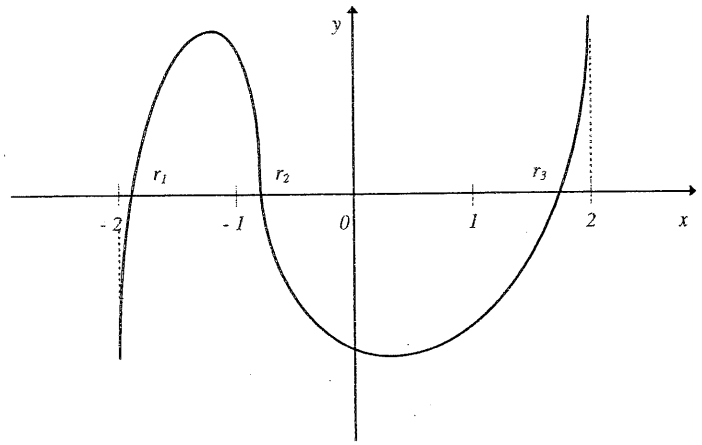
En el presente ejercicio, la única raíz que existe, tiene un valor numérico igual a $r = 0{,}73908502$.

♦♦♦ ●●● ♦♦♦

Ejemplo 4.2.- Sea el problema de acotar y separar todas las posibles raíces de la ecuación dada por la función:

$$f(x) = x^3 - 3x - 1 = 0.$$

Solución: Para separar las raíces de la ecuación $x^3 - 3x - 1 = 0$, primero, se debe calcular el valor de $f'(x) = 3(x^2 - 1)$; expresión que se anula en los puntos $x = 1$ y $x = -1$. Por ser $f(1) = -3$ y $f(-1) = 1$, hay una sola raíz en el intervalo $(-1 ; 1)$, como puede apreciarse en la figura 4.5.



- Figura 4.5 -

Además, como $f(0) = -1$, esa raíz se encuentra en el intervalo $(-1 ; 0)$. Calculando $f(-2)$ y $f(2)$, se aprecia que las tres raíces se están en los intervalos $(-2 ; -1)$, $(-1 ; 0)$ y $(1 ; 2)$, y sus respectivos valores son:

$r_1 = -1{,}532088886$; $r_2 = -0{,}347296355$ y $r_3 = 1{,}879385242$

♦♦♦ ●●● ♦♦♦

4.3.- MÉTODOS DE RESOLUCIÓN DE ECUACIONES

En los puntos que siguen se estudiarán varios métodos de resolución de ecuaciones, los que podrán ser aplicados tanto a la resolución de ecuaciones algebraicas como trascendentes. Más adelante se estudiarán métodos especializados en la resolución de ecuaciones algebraicas exclusivamente. La única razón de incluir más de un método es el hecho comprobado, mencionado en la introducción, de que no existe en este campo una *PANACEA UNIVERSAL*; vale decir, **el método** que resuelva todos los problemas que se puedan presentar en la práctica.

En los puntos que siguen se considera que se posee cierto conocimiento elemental previo de la ecuación, como ser el intervalo de acotación de las raíces o, mejor aún, subintervalos de separación de las mismas.

4.3.1.- MÉTODO DE TANTEOS

En el **MÉTODO DE TANTEOS** se determinan valores de $f(x)$ correspondientes a valores sucesivos de x hasta que se presente un cambio de signo en la evaluación de $f(x)$. El cambio de signo indica que se ha pasado por una raíz. Entonces se puede obtener una aproximación mejorada del valor de la raíz volviendo al último valor de x que precede al cambio de signo, y a partir de este valor de x, determinar nuevamente los valores de $f(x)$ correspondientes a valores sucesivos de x, reiterando el procedimiento, pero, utilizando a partir de ahora, un incremento menor al que se utilizó inicialmente, hasta que cambie nuevamente el signo de $f(x)$.

El procedimiento descrito en el párrafo anterior se reitera con incrementos de x cada vez más pequeños hasta lograr un valor suficientemente preciso de la raíz. Si se aplicara nuevamente este procedimiento para otros subintervalos de separación, se localizarán y aproximarán las sucesivas raíces de $f(x)$, en cada cambio de signo de esta.

Es importante acotar y separar las raíces antes de aplicar este método con el fin de no pasar por alto alguna de ellas. Es necesario tener la precaución de elegir cuidadosamente el valor inicial del incremento, dado que, es posible que no se identifique una de dos raíces muy próximas. Generalmente esto no sucede si se utilizan valores del incremento suficientemente pequeños en la sucesión inicial.

Otro problema que puede acarrear la falta de conocimiento previo de la función es el sobreflujo que seguramente se produce en la computadora a causa de valores muy grandes o próximos a infinito de la función para ciertos valores de x. Es necesario evitar estos intervalos, tratándolos de un modo diferente.

Este método, por demás elemental, pero muy laborioso de ser aplicado manualmente, es hoy de uso corriente debido a la gran capacidad de procesamiento de las computadoras electrónicas digitales, además de la rapidez con que lo hacen.

4.3.1.1.- Error en el Método de TANTEOS

Tal como se ha descripto en el punto anterior, el método de tanteos constituye un algoritmo infinito que deberá ser detenido indefectiblemente mediante algún procedimiento artificial.

Este procedimiento artificial puede ser utilizado para lograr la precisión deseada en cuanto al grado de aproximación requerido, si se detiene el procesamiento cuando la diferencia, en valor absoluto, entre dos valores consecutivos de la variable es menor o igual que una cierta constante **E** positiva y arbitraria, previamente establecida; es decir, cuando la *DIFERENCIA ABSOLUTA* resulta:

(4.2) $$|x_{i-1} - x_i| \leq \mathbf{E}$$

Una alternativa más eficaz y por tanto muy utilizada, dado que no demanda procesamiento adicional, consiste en operar sobre el valor de la función que, en este caso es calculada en todos los pasos para estudiar la evolución del cambio de signo de la función, de la misma manera en que se lo ha hecho con la variable; en otras palabras, puede detenerse el procesamiento y adoptar el valor de x_i para el cual resulta:

(4.3) $$|f(x_i)| \leq \mathbf{E}$$

Donde **E** será el error máximo admisible en el procesamiento, haciendo uso de esta alternativa.

♦♦♦ ●●● ♦♦♦

4.3.2.- MÉTODO DEL INTERVALO MEDIO

Supóngase que mediante algún método se ha separado una de las raíces de una ecuación dada por:

(4.1) $$f(x) = 0$$

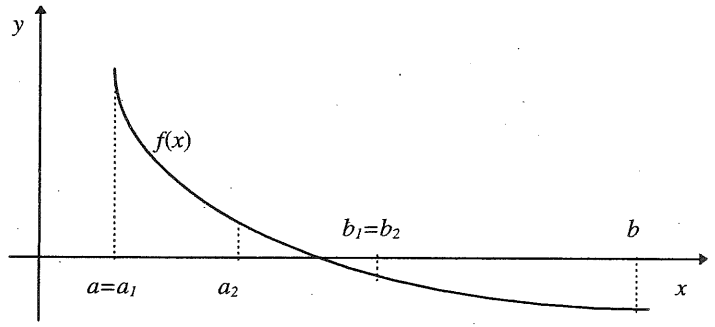
Por el teorema de Bolzano, se sabe que si $f(x)$ es continua en el intervalo $(a;b)$ y en los extremos toma valores $f(a)$ y $f(b)$, respectivamente, con signos opuestos, entonces, la función $f(x) = 0$ se anula por lo menos una vez, o un número par de veces, en un punto interior de $(a;b)$.

Una interpretación geométrica del teorema de Bolzano conduce naturalmente a dividir el intervalo dado en dos partes iguales; si no se anula en el punto de división, se tomará aquel subintervalo $(a_1 ; b_1)$ en cuyos extremos la función tiene signos diferentes (consultar la figura 4.6); es decir, si se cumple la condición:

$$f(a_1) \cdot f(b_1) < 0$$

Dividiendo nuevamente el subintervalo obtenido en el paso anterior en dos partes iguales y, si no se anula la función en el nuevo punto de subdivisión, se tomará aquel nuevo $(a_2 ; b_2)$ donde:

$$f(a_2) \cdot f(b_2) < 0$$



- Figura 4.6 -

Reiterando el procedimiento, se llegará a un punto de subdivisión en el que, al menos teóricamente, se anula $f(x)$. En general, se obtendrá un par de sucesiones indefinidas monótonas que tienden a un número u, tal que:

$$a \leq a_1 \leq \ldots \leq a_n \leq \ldots \leq u \leq \ldots \leq b_n \leq \ldots \leq b_1 \leq b$$

Considerando el procedimiento utilizado en la construcción de los subintervalos, es posible escribir:

$$b_1 - a_1 = \frac{b-a}{2}$$

$$b_2 - a_2 = \frac{b_1 - a_1}{2} = \frac{b-a}{2^2}$$

$$b_3 - a_3 = \frac{b_2 - a_2}{2} = \frac{b-a}{2^3}$$

y, en general:

(4.4) $$b_n - a_n = \frac{b-a}{2^n}$$

en la que n representa la cantidad de subdivisiones realizadas en el intervalo $(a;b)$. Tomando límites para $n \to \infty$, resulta:

(4.5) $$\lim_{n\to\infty}(b_n - a_n) = \lim_{n\to\infty}\frac{b-a}{2^n}$$

de donde es posible concluir que, en el punto $x = u$, es $f(u) = 0$, por estar u incluido en el entorno dado por $b - a$, que, por pequeño que sea, la función toma en sus extremos valores de signos opuestos.

4.3.2.1.- Error en el Método del INTERVALO MEDIO

Se ha visto cómo, mediante subdivisiones sucesivas por partes iguales, de un intervalo que contiene la raíz a determinar (intervalo de separación), se puede construir un algoritmo que permita obtener dicha raíz, con la aproximación que se desee, siempre que se verifique que:

(4.6) $$b_n - a_n \leq \mathbf{E}$$

Puede interesar conocer cuantas iteraciones son necesarias realizar para que el error con que se calcula el valor de la raíz no supere un cierto valor arbitrario $\mathbf{E} > 0$; en cuyo caso, por (4.4) y (4.6), es:

$$b_n - a_n = \frac{b-a}{2^n} \leq \mathbf{E}$$

de donde, es posible despejar el número de iteraciones que son necesarias realizar para que el error sea menor o igual a $\mathbf{E}$:

$$n \geq \frac{\log(b-a) - \log \mathbf{E}}{\log 2}$$

Para finalizar, es necesario alertar sobre la posibilidad de pasar por alto raíces, si solo se cuenta con un intervalo inicial de *ACOTACIÓN*: o bien, cuando algunas raíces son de multiplicidad superior a uno. Todos estos inconvenientes se evitan si se aplica el método sobre intervalos de *SEPARACIÓN* exclusivamente, y además, se estudia la multiplicidad de la o las raíces obtenidas.

Con el objeto de aportar mayores precisiones sobre el método del Intervalo Medio, veamos el siguiente:

Ejemplo 4.3.- Determinar el valor de la raíz de la función $f(x) = e^x - x^2 - 3$, que se halla comprendida en el intervalo (-1; 2), con una aproximación menor o igual que 0,0001; vale decir, $\mathbf{E} < 10^{-4}$.

Solución: La cantidad de veces que será necesario iterar el algoritmo para obtener la precisión deseada es:

$$n \geq \frac{\log(2+1) - \log 10^{-4}}{\log 2} \cong \frac{0{,}477 + 4}{0{,}301} \cong 14{,}87$$

Esto indica que 15 iteraciones son suficientes para obtener la aproximación deseada. El proceso de cálculo arroja los siguientes valores:

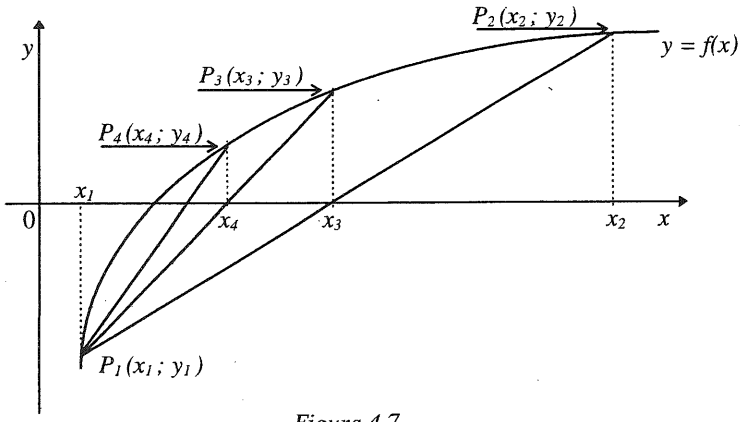
Valor de la Función		Intervalo resultante
$f(-1)$	$= -3{,}63212$	$(-1\ ;\ 2)$
$f(2)$	$= 0{,}38906$	$(0{,}5\ ;\ 2)$
$f(0{,}5)$	$= -1{,}60128$	$(1{,}25\ ;\ 2)$
$f(1{,}25)$	$= -1{,}07216$	$(1{,}625\ ;\ 2)$
$f(1{,}625)$	$= -0{,}56221$	$(1{,}8125\ ;\ 2)$
$f(1{,}8125)$	$= -0{,}15941$	

donde, puede apreciarse que el valor de la función tiende con relativa rapidez a cero y que, los sucesivos intervalos cumplen la propiedad $f(a_i)\ .\ f(b_i) < 0$. Reiterando el procedimiento mostrado hasta completar los quince pasos, se habrá obtenido el valor de la raíz con la precisión deseada, que es: $x = 1{,}8731$ con todas sus cifras decimales exactas.

♦♦♦ ●●● ♦♦♦

4.3.3.- MÉTODO DE INTERPOLACIÓN LINEAL

Supóngase tener separadas las raíces de una ecuación $f(x) = 0$, donde, uno de los subintervalos de separación viene dado por $(x_1 ; x_2)$, en el cual se cumple también la condición: $f(x_1) . f(x_2) < 0$, tal como se ilustra en la figura 4.7, para $y = f(x)$.



- Figura 4.7 -

Como $f(x_1)$ y $f(x_2)$ tienen signos diferentes, se sabe que, de ser monótona en el intervalo dado $(x_1 ; x_2)$, entonces $f(x)$ tendrá una raíz entre x_1 y x_2. La ecuación de la recta que pasa por los puntos $P_1 ; P_2$, es:

$$y - y_1 = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1)$$

en la que, para hallar la intersección con el eje de las abscisas se hace $y=0$, tomando x el valor x_3, en la cual x_3 es una mejor aproximación del verdadero valor de la raíz, como puede observarse en la figura 4.7; resulta entonces:

$$-y = \frac{y_2 - y_1}{x_2 - x_1}(x_3 - x_1)$$

y, despejando de esta última:

(4.7)
$$x_3 = \frac{x_1 y_2 - x_2 y_1}{y_2 - y_1}$$

El valor determinado de esta manera, se sustituye en la función para calcular el correspondiente valor de y_3. A esto sigue una comparación entre los signos de $f(x_3)$ con los de $f(x_1)$ y $f(x_2)$, desechando el punto para el cual la función tiene igual signo que $f(x_3)$. Seguidamente con idéntico procedimiento al ya expuesto, se reitera el primer paso entre los puntos del subintervalo así obtenido con el objeto de determinar el valor de x_4, que es aún una mejor aproximación de la raíz buscada.

Es necesario repetir el procedimiento descripto con nuevos pares de puntos que definan los correspondientes subintervalos, cada vez más pequeños, hasta que se logra una aproximación que tenga la precisión previamente fijada, lo que ocurre cuando dos valores consecutivos de x difieren en una cantidad menor que un cierto valor **E** positivo y arbitrario de prueba, y mejor aún, con la verificación de que el valor de la función se encuentra muy próximo a cero.

Este método, llamado por algunos autores **MÉTODO DE LA REGULA FALSI**, o también **MÉTODO DE LAS PARTES PROPORCIONALES**, puede ser utilizado en la computadora para determinar las raíces de una ecuación, con un grado especificado de precisión y con un número menor de iteraciones que las que se utilizan en el método de tanteos; o sea, es un algoritmo que converge más rápidamente a la solución que aquel dado en el método de tanteos.

Aún queda por estudiar su aptitud respecto de la convergencia, cosa que se realizará a continuación.

4.3.3.1.- Convergencia del método de INTERPOLACIÓN LINEAL

Para tener la plena seguridad que el *Método de Interpolación Lineal* converge, es necesario que satisfaga, en los puntos $x_1 ; x_2$ (extremos del subintervalo de separación de las raíces, tomado inicialmente) algunas condiciones específicas; estas son:

1) $f(x_1).f(x_2) < 0$

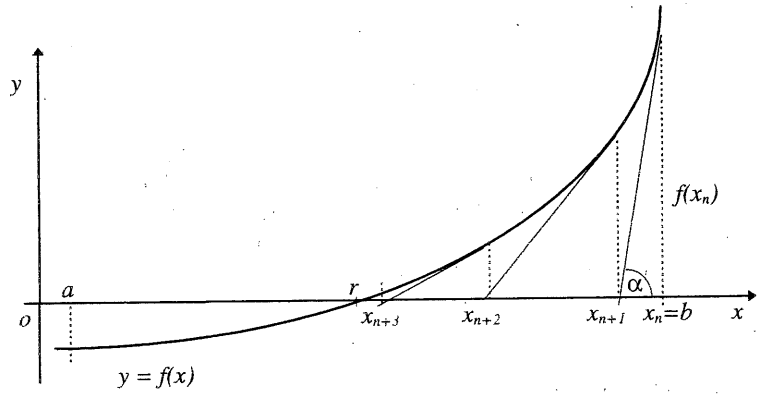
2) $f'(x) \neq 0$ *para todo* $x \in (x_1 ; x_2)$

3) *$f(x)$ debe ser monótona para todo punto del intervalo $(x_1 ; x_2)$*

♦♦♦ ●●● ♦♦♦

4.3.4.- MÉTODO DE NEWTON-RAPHSON

Este método, muy útil en el caso de tener que mejorar una primera aproximación a una raíz r de una ecuación de la forma dada por la expresión (4.1), la que pudo haber sido separada por simples tanteos, mediante algún recurso gráfico, o de cualquier otra manera posible.



- Figura 4.8 -

Supóngase que x es una primera aproximación al valor de una raíz, la que puede ser directamente uno de los extremos del intervalo cerrado de separación $[a\ ;\ b]$. Si se dibuja una recta tangente a la curva en el punto $x = x_n$, interceptará al eje de las abscisas en un valor dado por $x = x_{n+1}$, que constituye una aproximación mejorada de la raíz r. Consultar en la figura 4.8.

Se puede observar que la pendiente de la recta tangente a la curva $y = f(x)$ es:

$$\operatorname{tg}\alpha = f'(x_n) = \frac{f(x_n)}{x_n - x_{n+1}}$$

de donde:

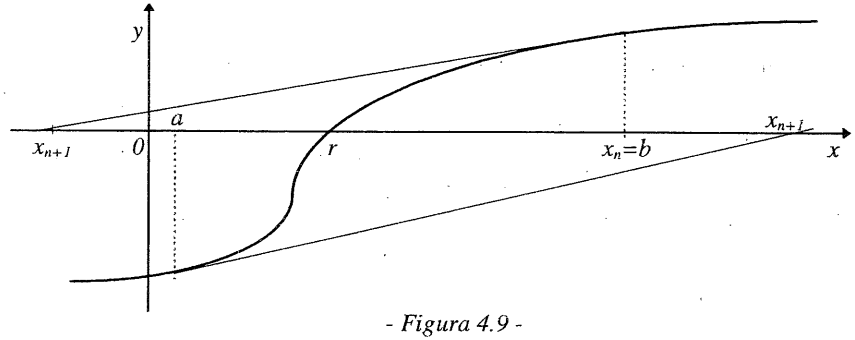
(4.8)
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

BIBLIOTECA UNIVERSITARIA ORIENTE

El valor de la función, como así también el de su derivada, son conocidas para el valor $x = x_n$, y la nueva aproximación de la raíz x_{n+1}, se obtiene utilizando la ecuación (4.8). Se repite el procedimiento descripto, partiendo de esta nueva aproximación, para obtener una mejor. Continuando de esta manera hasta que dos valores consecutivos de la raíz aproximada difieran en una cantidad igual o menor que un cierto **E** positivo y arbitrario, previamente prescrito.

4.3.4.1.- Convergencia del Método de NEWTON-RAPHSON

Tal como se puede apreciar en la figura 4.9, si se aplica el procedimiento en el punto $x_n = b$, en lugar de lograr una mejor aproximación a la raíz, el procedimiento se hubiera alejado de ella. Lo mismo sucede si el procedimiento se aplica en el punto $x_n = a$. Este problema, que fue estudiado por *Fourier*, quien estableció ciertas condiciones sobre el punto en el cual debe ser aplicado el procedimiento de *Newton-Raphson*, en función de los signos que toman $f(x)$ y $f''(x)$, pues, de lo contrario, la aplicación del método podría resultar divergente; vale decir, ampliar el intervalo en vez de estrecharlo, finalizando en un sobreflujo de la computadora.



- Figura 4.9 -

Se supondrá, entonces, que en el punto x_n, tanto la función, como sus derivadas de primero y segundo orden tienen un signo determinado, distinto de cero. Fourier estableció que el procedimiento de *Newton-Raphson* debe aplicarse en caso de que $f(x_n)$ y $f''(x_n)$ tengan igual signo; de no ser así, en general, el procedimiento es divergente.

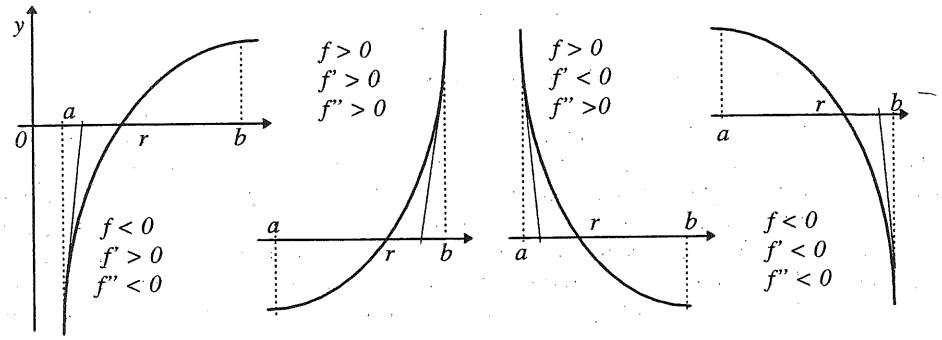
En la figura 4.10 se hallan representados gráficamente los cuatro diferentes casos que, en la práctica se pueden presentar, donde el método de *Newton-Raphson* es, con seguridad, aplicable.

A continuación se tratará de demostrar analíticamente lo que, intuitivamente, se ha expuesto. Tomando los tres primeros términos del desarrollo en serie de *Taylor*, aplicado en un entorno del punto $x = a$, resulta:

19.883

(4.9) $$f(x)=f(a)+(x-a)f'(a)+\frac{1}{2}(x-a)^2 f''(\xi) \quad ; \quad a<\xi<x$$

donde, para el valor de la variable igual al de la raíz r, se verifica que $f(r) = 0$.



- Figura 4.10 -

Suponiendo ahora que, $f'(a) \neq 0$, en el punto $x = r$, se tiene:

$$0=f(a)+(r-a)f'(a)+\frac{1}{2}(r-a)^2 f''(\xi) \quad ; \quad a<\xi<r$$

y despejando de esta última, el valor de r correspondiente al segundo término, se obtiene:

$$r=a-\frac{f(a)}{f'(a)}-\frac{1}{2}(r-a)^2\frac{f''(\xi)}{f'(a)} \quad ; \quad a<\xi<r$$

Los dos primeros términos determinan, como ya se estudió, el valor aproximado de la raíz que, si se lo designa con a', resulta:

(4.10) $$a'=a-\frac{f(a)}{f'(a)}$$

que, como puede comprobarse, es equivalente a (4.8). Se demostrará que a' está más próximo a r que el valor de a, suponiendo que la función $f(x)$ y la derivada segunda $f''(x)$ tienen igual signo en un entorno del punto $x = a$. En efecto:

(4.11) $$r-a'=-\frac{1}{2}(r-a)^2\frac{f''(\xi)}{f'(a)}$$

de donde:

(4.12) $$\frac{r-a'}{a'-a}=\frac{1}{2}\frac{(r-a)^2 f''(\xi)}{f(a)}>0$$

Esta igualdad afirma que, si la función $f''(\xi)$ tiene igual signo que $f(a)$, el numerador y el denominador del primer miembro tienen que tener, necesariamente, el mismo signo; o sea, que a' está situado entre los valores de a y r; vale decir que, a' mejora la aproximación dada por a.

La condición estudiada recibe el nombre de **OBSERVACIÓN DE FOURIER** y da origen a un Teorema que se expondrá a continuación. Luego de todo lo dicho es posible afirmar que:

Teorema: *Una función $f(x) = 0$, definida, monótona, dos veces derivable y con derivadas continuas; vale decir: ($f \in C^2$) en el intervalo (a ; b), que satisface las siguientes condiciones:*

I $f(a) \cdot f(b)<0$

II $f'(x)$ *distinta de 0 para todo* x *comprendida en (a;b)*

III $f(x_0) \cdot f''(x_0) > 0$ *con* x *comprendida en (a;b)*

entonces, dadas estas condiciones y tomando x_0 como primera aproximación de la raíz r, el método de NEWTON-RAPHSON converge a la única solución r de $f(x) = 0$.

Es necesario hacer notar que la tercer condición puede ser satisfecha en los dos extremos del intervalo de separación; en uno solo de ellos o, eventualmente, en ninguno. En la figura 4.9 se muestra el caso en que ninguno de los dos extremos del intervalo de separación permite la aplicación del método.

4.3.4.2.- Error en el Método de NEWTON-RAPHSON

Puede apreciarse que, mediante la aplicación de la fórmula del Método de NEWTON-RAPHSON, dada por la expresión 4.8, se puede cometer un error absoluto expresado por:

$$E < | r - a |$$

que según la expresión 4.11, en algún momento del cálculo, resulta:

(4.13) $$\mathbf{E} < \left| \frac{(r-a)^2 f''(\xi)}{2 f'(a)} \right|$$

Si bien no es posible conocer todos los valores de las expresiones incluidas en el segundo miembro, pueden tomarse cotas superiores de las expresiones desconocidas que intervienen en ella. Llamando h a la longitud del intervalo $(b - a)$ y, K a una cota superior del valor de $f''(x)$, resulta entonces:

$$\mathbf{E} < \left| \frac{h^2 K}{2 f'(a)} \right|$$

la que se constituye en la expresión de una cota superior del error cometido en la aplicación del método de *NEWTON-RAPHSON*.

La aplicación del método de *NEWTON-RAPHSON* se estudiará a través del siguiente:

Ejemplo 4.4.- Se trata de resolver mediante la aplicación estudiada del método de NEWTON-RAPHSON, la ecuación:

$$f(x) = x^2 - 10 \ln x - 3$$

Solución: Utilizando cualquier procedimiento para realizar la separación previa de las raíces, se encuentra que una de ellas está comprendida en el intervalo (4;5). Se sabe también que:

$$f'(x) = 2x - \frac{10}{x} \quad ; \quad f''(x) = 2 + \frac{10}{x^2}$$

Conociendo todos los elementos que intervienen en el problema, entonces deben ser verificados sus valores, a la luz de las tres condiciones previamente expuestas:

I $f(4) \cdot f(5) = -0{,}863 \cdot 5{,}906 < 0$

II $f'(x) \neq 0$; para toda x que pertenece al intervalo (4;5), por ser monótona

III Cuando $x_0 = 4 \Rightarrow f(4) \cdot f''(4) = -0{,}863 \cdot 2{,}625 < 0$ entonces se desecha por no cumplir la tercera condición establecida por el método. Cuando $x_0 = 5$; $f(5) \cdot f''(5) = 5{,}906 \cdot 2{,}4 > 0$ entonces es posible aplicarlo.

Haciendo uso de la expresión (4.10) en el punto extremo $x = 5$, resulta:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 5 - \frac{f(5)}{f'(5)} = 5 - \frac{5{,}906}{8} = 4{,}262$$

Proceso que es posible iterar hasta que, entre dos valores consecutivos de las aproximaciones de la raíz, no se observa una sensible mejoría, según criterio previamente establecido. A continuación se verá como puede calcularse una cota superior del error para el valor así obtenido. Supóngase detener el procedimiento en el primer paso; en el cual se sabe que:

$$\mathbf{E} < \left| \frac{h^2 K}{2 f'(x_0)} \right|$$

donde. $h = b - a = 5 - 4 = 1$; y

$$K = máx\,[f''(4) ; f''(5)] = máx\,[2{,}625 ; 2{,}4] = 2{,}625$$

entonces:

$$\mathbf{E} < \frac{1 \times 2{,}625}{2 \times 8} = 0{,}164$$

valor que representa una cota superior del error de la aproximación obtenida luego de realizar el primer paso.

♦♦♦ ●●● ♦♦♦

Las funciones y operaciones preprogramadas, son aquellas que mediante una sola instrucción de codificación del programa, son ejecutadas por la computadora, resolviendo una gran cantidad de pasos elementales que, de otra manera demandarían una extensa y complicada codificación. Tratándose de funciones trascendentes, generalmente, se utilizan desarrollos en serie para calcular su respectivo valor; mientras que, en el caso de algunas operaciones elementales, los desarrollos en serie se sustituyen con ventaja por fórmulas iterativas deducidas a partir del método de *NEWTON-RAPHSON*, como se podrá apreciar en el siguiente:

Ejemplo 4.5.- Supóngase tener que construir un algoritmo para calcular el inverso de un número que será denominado, en general N.

Solución: Dadas las condiciones del problema, resulta:

$$x = \frac{1}{N} \Rightarrow f(x) = \frac{1}{x} - N = 0$$

mediante la cual, luego de aplicar los principios del método de *NEWTON-RAPHSON*, se obtiene la siguiente fórmula iterativa:

$$x_{n+1} = x_n\left(2 - N\,x_n\right)$$

Se debe iniciar el procedimiento con un x_0 tal que, $N\,x_0 < 2$, en caso contrario el proceso será divergente. Para $N = 3$ y $x_0 = 0{,}5$, se genera la siguiente serie de valores:

$x_0 = 0{,}5$
$x_1 = 0{,}25$
$x_2 = 0{,}3125$
$x_3 = 0{,}33203125$
$x_4 = 0{,}3333282471$
$x_5 = 0{,}3333333333$

como era de esperarse.

♦♦♦ ●●● ♦♦♦

Con el objeto de fijar los conceptos y aportar mayores detalles del procedimiento estudiado, veamos el siguiente:

Ejemplo 4.6.- Supóngase ahora, tener que construir un algoritmo para calcular la raíz cuadrada de un número denominado, en general, *N*.

Solución: Dadas las condiciones particulares de este problema, resulta:

$$x = \sqrt{n} \Rightarrow f\left(x\right) = x^2 - N$$

de donde:

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{N}{x_n}\right)$$

expresión fácilmente aplicable. En este caso no se encuentran bien determinadas las restricciones sobre el valor inicial a ser utilizado en el procedimiento. Para calcular la raíz de 2 y tomando $x_0 = 2$, resulta.

$x_0 = 2$
$x_1 = 1{,}5$
$x_2 = 1{,}416666667$
$x_3 = 1{,}414215686$

$x_4 = 1{,}414213562$
$x_5 = 1{,}414213562$

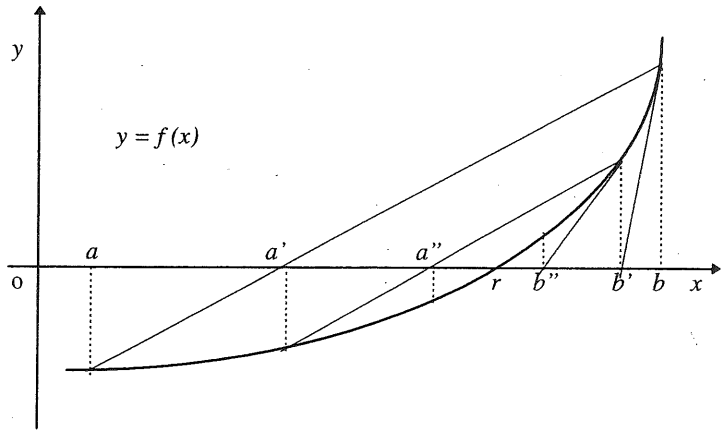
♦♦♦ ●●● ♦♦♦

4.3.5.- MÉTODO MIXTO

En este caso no se trata de un método con procedimientos originales; si no, mas bien, de la aplicación sistemática de otros dos métodos conocidos y ya estudiados con anterioridad. En este orden de cosas, diremos que el denominado **MÉTODO MIXTO** consiste en la aplicación simultánea de los métodos de las *PARTES PROPORCIONALES* y del de *NEWTON-RAPHSON*, ya estudiados en el presente capítulo.

Este método, fue sugerido en su momento por el matemático *Dandelin*, resultando sumamente práctico por la facilidad de aplicación, programación y su aceptable velocidad de convergencia; lo que se traduce en una gran agilidad en el momento del procesamiento del mismo.

El método mixto consiste, como ya se dijo, en la aplicación sucesiva de los métodos de las *PARTES PROPORCIONALES* y de *NEWTON-RAPHSON*, ya estudiados, de manera alternativa, para obtener un verdadero encaje de intervalos que contenga al valor de la raíz; vale decir, subintervalos que encierran el valor de r buscado, tal como se muestra en la figura 4.12.



- Figura 4.12 -

Para la aplicación de los distintos métodos es necesario tener en cuenta, en cada caso, las restricciones propias de cada uno de los métodos intervinientes; como así también, las condiciones de convergencia de los mismos, ya estudiadas oportunamente; del resto, todo resulta exactamente igual.

♦♦♦ ●●● ♦♦♦

Capítulo 5

Teoría General de la Iteración

5 . - TEORÍA GENERAL DE LA ITERACIÓN

Introducción. Método de Iteración. Convergencia del Método de Iteración. Utilidad del Método de Iteración. Teoría General de la Iteración. Iteración de Primer Orden de Convergencia. Iteración de Segundo Orden de Convergencia. Iteraciones de Ordenes Superiores de Convergencia. Proceso Delta-Cuadrado de Aitken. Método de Segundo Orden de Newton. Resolución de Ecuaciones Algebraicas. Acotación y Separación de las Raíces Reales. Teorema de STURM.

5.1.- INTRODUCCIÓN

En el presente capítulo, continuación natural de lo estudiado en el anterior, se proseguirá con el tratamiento de la resolución numérica de ecuaciones mediante la aplicación de métodos iterativos comunes, pero, además, haciendo uso sistemático de los resultados a los cuales se ha llegado mediante el estudio de la **TEORÍA GENERAL DE LA ITERACIÓN**.

Los resultados aludidos en el párrafo anterior, no son más que las grandes generalizaciones a las cuales pudo llegar el estudio profundo y sistemático de las características generales, con algunas de sus particularidades, que tienen en común todos los métodos recursivos o iterativos de cualquier especie; muchos de los cuales han sido tratados en el capítulo anterior; otros, serán estudiados en el presente..

En este orden de cosas, mediante la denominada **TEORÍA GENERAL DE LA ITERACIÓN**, es posible categorizar los métodos recursivos o iterativos, algunos de los cuales ya han sido estudiados en el capítulo precedente, otros serán estudiados mas adelante, juntamente con la teoría mencionada.

El elevado poder de síntesis logrado mediante los resultados obtenidos con el estudio de la **TEORÍA GENERAL DE LA ITERACIÓN**, permite elaborar métodos que pueden clasificarse como muy rápidamente convergentes, logrando de esta manera, valores altamente precisos luego de ejecutar un número relativamente bajo de pasos en el procesamiento.

Seguidamente, se estudiará el método de iteración por excelencia el cual, por tal motivo, lleva justamente el nombre de *Método de Iteración*.

5.2.- MÉTODO DE ITERACIÓN

Para poder calcular una raíz r de la ecuación

$$f(x) = 0$$

por el **MÉTODO DE ITERACIÓN**, es necesario que exista la posibilidad de escribir la expresión analítica de la ecuación dada, en la forma:

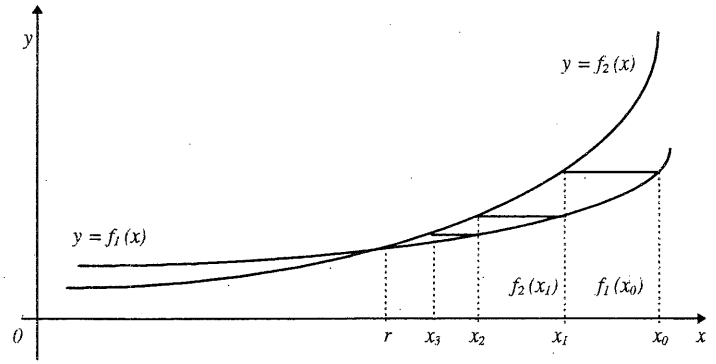
(5.1) $$f_1(x) = f_2(x)$$

lo cual, salvo muy raras excepciones, resulta posible y, generalmente, con un sencillo tratamiento algebraico de la ecuación dada.

Admitiendo de antemano que, en un entorno del punto común o de intersección de ambas curvas, las que han sido representadas por $y = f_1(x)$ e $y = f_2(x)$ en la figura 5.13, se cumpla que:

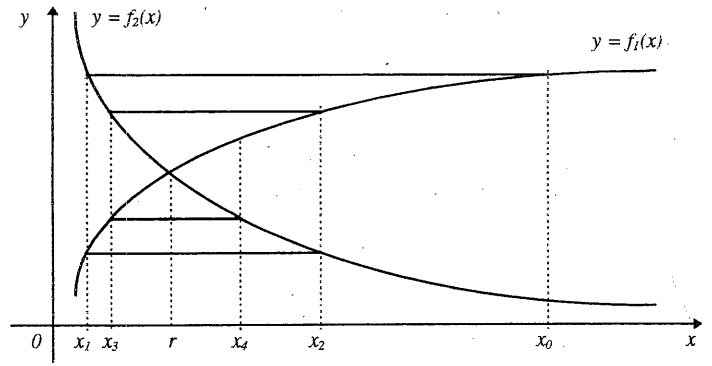
(5.2) $$|f_1'(x)| < |f_2'(x)|$$

vale decir que, en un entorno del punto $x = r$, la pendiente de la curva $f_1(x)$ sea, en valor absoluto, estrictamente menor que la pendiente de la curva $f_2(x)$, como puede apreciarse en las figuras 5.1 y 5.2.



- Figura 5.1 -

Si se parte de un valor x_0 y se procede sistemáticamente como indican las figuras 5.1 y 5.2, se obtendrán nuevos valores x_1 ; x_2 ; x_3 ;... de la variable, que convergen hacia la raíz r buscada.



- Figura 5.2 -

Como puede apreciarse, el proceso que es necesario realizar se sintetiza mediante el siguiente cuadro:

$$x_0 \Rightarrow f_1(x_0)=f_2(x_1) \Rightarrow x_1$$
$$x_1 \Rightarrow f_1(x_1)=f_2(x_2) \Rightarrow x_2$$
$$x_2 \Rightarrow f_1(x_2)=f_2(x_3) \Rightarrow x_3$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$x_n \Rightarrow f_1(x_n)=f_2(x_{n+1}) \Rightarrow x_{n+1}$$

lo cual permite obtener, si el proceso resultase convergente, el valor aproximado r, de la raíz buscada.

Si las derivadas $f_1'(x)$ y $f_2'(x)$, en un entorno del punto $x = r$, tienen igual signo, como se indica en la figura 5.1, la aproximación se realiza de la manera indicada, y recibe el nombre de *ESCALERA*. Si, en cambio, los signos de las pendientes de las curvas involucradas son diferentes, la aproximación se llama en *ESPIRAL*, como se muestra en la figura 5.2.

Ejemplo 5.1.- En este caso, se trata de calcular la raíz r de la ecuación dada por la siguiente expresión:

$$f(x) = e^x + 2\ ln\ x + x - 3 = 0$$

Solución: Para iniciar el procedimiento de iteración se puede subdividir la ecuación dada, haciendo:

$$f_1(x) = e^x + x \quad ; \quad f_2(x) = 3 - 2\ ln\ x$$

y luego aplicar a estas últimas expresiones las verificaciones y la secuencia estudiada en los párrafos anteriores.

♦♦♦ ●●● ♦♦♦

5.2.1.- CONVERGENCIA DEL MÉTODO DE ITERACIÓN

Es posible estudiar la condición de convergencia del método de *ITERACIÓN*, con solo demostrar que x_{n+1} es una mejor aproximación a la raíz r que x_n . Para ello, por ser $f_1(r) = f_2(r)$, entonces:

$$f_1(r) - f_1(x_n) = f_2(r) - f_2(x_{n+1})$$

Aplicando el teorema del valor medio a ambos miembros de la expresión anterior y tomando módulos, resulta:

$$|r - x_n| \cdot |f_1'(\xi_1)| = |r - x_{n+1}| \cdot |f_2'(\xi_2)|$$

donde, $r < \xi_1 < x_n$; $r < \xi_2 < x_{n+1}$, y como $x_n \equiv x_{n+1}$ entonces $\xi_1 \equiv \xi_2$ cuando el proceso se encuentra en estado relativamente avanzado; vale decir, cuando el valor de x_n se acerca al de la raíz r.

Dado que, en todo punto de un entorno de $x = r$, por hipótesis se cumple:

$$|f_1'(x)| < |f_2'(x)|$$

debe ser:

$$|r - x_{n+1}| < |r - x_n|$$

es decir que, la aproximación de orden $n+1$ de la raíz r es mejor que la aproximación anterior, de orden n.

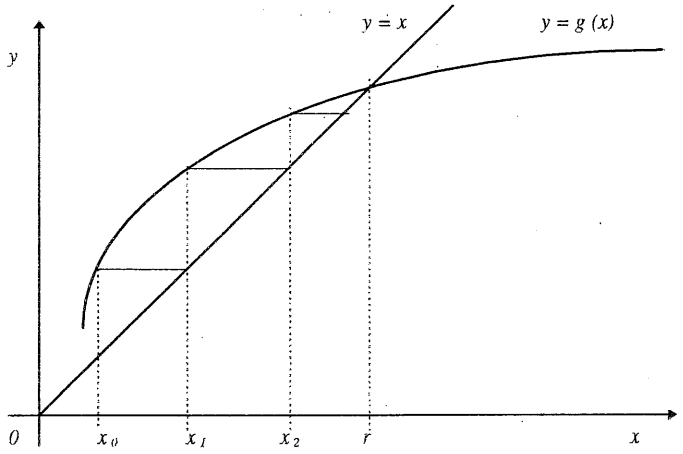
5.2.2.- UTILIDAD DEL MÉTODO DE ITERACIÓN

El uso más frecuente del método de iteración se realiza cuando la expresión $f(x)$ original se puede escribir en la forma $x = g(x)$, con lo cual, una de las derivadas es constantemente igual a uno. Suponiendo que x_0 sea una aproximación al valor de r; entonces, si es:

(5.3) $| g'(x) | < 1$

resulta la siguiente sucesión de iteraciones:

$$x_1 = g(x_0)$$
$$x_2 = g(x_1)$$
$$\cdots\cdots$$
$$x_n = g(x_{n-1})$$
$$x_{n+1} = g(x_n)$$



- Figura 5.3 -

Como puede apreciarse, bajo estas condiciones el algoritmo se facilita grandemente. Ver figura 5.3.

Es importante que la condición de convergencia previamente establecida:

$$|g'(x)| < 1$$

sea satisfecha, pues, en caso contrario, el proceso resultaría divergente, o bien, podría reducirse al caso general estudiado, ya que, debería ser interceptada primero la curva $y = x$; con lo cual se pierde la posibilidad de lograr el objetivo deseado. Observar en detalle la figura 5.3.

Ejemplo 5.2.- Utilizando el método de *ITERACIÓN* determinar la raíz comprendida en el intervalo *(0;1)*, de la ecuación dada por:

$$\backslash \quad sen\, x = 5x - 2 = 0$$

con cuatro dígitos decimales exactos.

Solución: Primeramente es necesario volver a escribir la ecuación en la forma:

$$x = g(x) = sen^{-1}\ (5x - 2)$$

de donde:

$$g'(x) = \frac{5}{\sqrt{-25x^2 + 20x - 3}}$$

en consecuencia, se da inicio al procedimiento calculando, para $x_0 = 0{,}3$ el valor:

$$g'(0{,}3) \cong 6{,}66$$

dado que, para valores menores de 0,3 la cantidad subradical resulta negativa. No obstante, se comprueba que la cantidad resultante es mucho mayor que la unidad; por lo tanto, el proceso no convergirá.

Se probarán las hipótesis correspondientes, volviendo a escribir la ecuación en la forma siguiente:

$$x = g(x) = \frac{1}{5}(2 + \operatorname{sen} x)$$

de donde:

$$g'(x) = \frac{1}{5}\cos x$$

y, finalmente, para $x = 0$, resulta:

$$g'(x) = 0{,}2$$

valor que resulta aceptable a la luz de la convergencia. Se utilizará, entonces, la relación recursiva:

$$x_{n+1} = \frac{1}{5}(2 + \operatorname{sen} x)$$

que, con $x_0 = 0$, produce:

$$x_1 = 0{,}4 \; ; \; x_2 = 0{,}40139 \; ; \; x_3 = 0{,}40140 \; ; \; x_4 = 0{,}40140 \; ; \; \ldots$$

◆◆◆ ●●● ◆◆◆

5.3.- TEORÍA GENERAL DE LA ITERACIÓN

En puntos expuestos en el presente capítulo se ha estudiado que, de manera general, una ecuación del tipo $f(x)=0$, puede ser resuelta haciendo uso de la expresión recursiva representada por:

(5.4) $$x_{n+1} = \phi(x_n)$$

Los nombrados métodos entran dentro de la categoría de los clasificados como **MÉTODOS ITERATIVOS**; y se resuelven tomando x_0 como una primera aproximación de la raíz real r, la cual se está tratando de calcular.

Mediante la relación iterativa (5.4) es posible generar una sucesión de valores $x_0 \; ; \; x_1 \; ; \; x_2 \; ; \; \ldots \; ; \; x_n$ que aproxima el valor de la raíz. Llamando E_k a la diferencia entre la aproximación x_k de orden k y el valor exacto r e la raíz, resulta:

(5.5) $$x_0 = r + E_0 \; ; \; x_1 = r + E_1 \; ; \; \ldots \; ; \; x_n = r + E_n$$

Considerando las condiciones expuestas, la sucesión $x_0 \; ; \; x_1 \; ; \; \ldots \; ; \; x_n$ tenderá al valor r si, para algún k en adelante, resulta:

(5.6) $$|E_k| > |E_{k+1}| > |E_{k+2}| > \ldots > |E_n| \to 0$$

Ahora bien, reemplazando los valores de x_{n+1} y x_n de la expresión (5.4), por los correspondientes dados en la (5.5), se obtiene:

(5.7) $$r + E_{n+1} = \phi(r + E_n)$$

y, aplicando el teorema de *TAYLOR* al segundo miembro de (5.7), resulta:

$$r + E_{n+1} = \phi(r) + E_n \phi'(r) + \frac{1}{2!} E_n^2 \phi''(r) + \ldots$$

pero, considerando que $r = \phi(r)$, dado que r es una raíz de la ecuación dada, finalmente es:

(5.8) $$E_{n+1} = E_n \phi'(r) + \frac{1}{2!} E_n^2 \phi''(r) + \frac{1}{3!} E_n^3 \phi'''(r) + \ldots$$

Considerando en detalle las consecuencias de este resultado tan importante, es posible distinguir las siguientes contingencias:

Caso 1.- $\phi'(r) \neq 0$.

En este caso, despreciando desde el término de segundo orden en adelante; vale decir, considerando solamente el primer término del segundo miembro, resulta:

(5.9) $$E_{n+1} \cong E_n \phi'(r)$$

De esta ecuación es posible deducir inmediatamente que si $|\phi'(r)| < 1$, entonces, cada término de (5.6) será menor que el anterior, de tal modo que la sucesión $x_0 \; ; \; x_1 \; ; \; x_2 \; ; \; ; \ldots ; \; x_n$ tenderá al valor de r, raíz de la ecuación.

En este caso se dice que se trata de una **ITERACIÓN DE PRIMER ORDEN** o bien que, este caso es de **PRIMER ORDEN DE CONVERGENCIA** y puede ser descripto como un proceso lineal de E_n. En otros términos, dado que no se conoce el valor de r, en la expresión (5.9), se puede reemplazar su valor por el de x_n.

De aquí que, al iniciar el proceso, el criterio práctico para saber si el método converge es que se cumpla que $|\phi'(x_0)| < 1$; y que, el número de pasos extras necesarios para aumentar la aproximación un número dado de cifras decimales es independiente del lugar en que se encuentran; vale decir que, por ejemplo, si se emplean cinco iteraciones para ir de cuatro a seis decimales exactos, también requerirá de cinco iteraciones para ir de diez a doce decimales exactos.

Ejemplo 5.3.- Sea el problema de tener que determinar el valor de las raíces positivas de la ecuación:

(5.10) $$\ln x = 1 + \frac{1}{x}$$

Solución: El primer problema que se presenta consiste en escribir la ecuación (5.10), en la forma recursiva (5.4), lo que se logra fácilmente haciendo:

(5.11) $$x = \phi(x) = \exp\left(1 + \frac{1}{x}\right)$$

A continuación es necesario formular alguna hipótesis razonable sobre el valor de r, por burda que fuera, la que permitiría iniciar el proceso iterativo; a ésta la llamaremos x_0.

Para poder satisfacer la condición $|\phi'(x_0)| < 1$; es necesario antes, determinar la derivada de la expresión (5.11), la que resulta:

(5.12) $$\phi'(x) = -\frac{1}{x^2}\exp\left(1 + \frac{1}{x}\right)$$

de donde:

$$\phi'(1) = -7{,}389056 \quad ; \quad \phi'(2) = -1{,}120422 \quad ; \quad \phi'(3) = -0{,}421519$$

Donde es posible apreciar que este último valor, no así los anteriores, cumple con la condición de convergencia establecida.

Aplicando la fórmula de recurrencia previamente establecida, dada por la expresión (5.4), tomando $x_0 = 3$, se obtiene:

$$\begin{aligned}
x_1 &= \exp\left(1 + \frac{1}{3}\right) = 3{,}700000 \\
x_2 &= \exp\left(1 + \frac{1}{3{,}7}\right) = 3{,}560000 \\
x_3 &= \exp\left(1 + \frac{1}{3{,}56}\right) = 3{,}600000 \\
x_4 &= \exp\left(1 + \frac{1}{3{,}6}\right) = 3{,}589500 \\
x_5 &= \exp\left(1 + \frac{1}{3{,}5895}\right) = 3{,}591600 \\
x_6 &= \exp\left(1 + \frac{1}{3{,}5916}\right) = 3{,}590900 \\
x_7 &= \exp\left(1 + \frac{1}{3{,}5909}\right) = 3{,}591300
\end{aligned}$$

etc.; dado que, las aproximaciones se encuentran, alternativamente, por encima y por debajo del valor de la raíz, ésta debe encontrarse necesariamente entre 3,5909 y 3,5913; cosa que se puede determinar siguiendo el procedimiento indicado.

El ejemplo anterior debe resultar útil para comprobar todo lo dicho en el último párrafo, anterior al mismo.

◆◆◆ ●●● ◆◆◆

Caso 2.- $\phi'(r) = 0 \quad ; \quad \phi''(r) \neq 0.$

En este segundo caso, si se despreciara el término de tercer orden y potencias superiores de E_n, se obtiene una expresión de la forma:

(5.13) $$E_{n+1} \cong \frac{1}{2}\phi''(r)E_n^2$$

Considerando la expresión anterior es posible comprobar que, para que la sucesión $x_0 ; x_1 ; x_2 ;...$ converja a la raíz r, solo es necesario que $\phi''(r)$ sea finita y E_0 sea relativamente pequeño. En estas condiciones se puede deducir a partir de (5.13) que cada error es proporcional al cuadrado del anterior, en consecuencia, la velocidad de la convergencia resulta bastante rápida.

Dadas las condiciones descriptas en este segundo caso, puede decirse que se trata de una **ITERACIÓN DE SEGUNDO ORDEN** o bien que, es un caso de iteración de **SEGUNDO ORDEN DE CONVERGENCIA**.

El comportamiento de los métodos encasillados en este caso llevan a duplicar el número de dígitos exactos en cada iteración o paso; en otras palabras, si en un determinado paso mejora la aproximación de cuatro a ocho decimales exactos, en el paso siguiente se mejorará aquella, obteniendo de ocho a dieciséis decimales exactos.

Ejemplo 5.4.- Un acontecimiento muy importante, que es necesario destacar, es que el método de *NEWTON-RAPHSON* es un proceso iterativo de segundo orden de convergencia, como se tratará de demostrar.

Solución: La fórmula de recurrencia del método de *NEWTON-RAPHSON*, es:

(5.14) $$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

vale decir que, en este caso es:

$$\phi(x) = x - \frac{f(x_n)}{f'(x_n)}$$

de donde:

$$\phi'(x) = 1 - 1 + \frac{f(x)f''(x)}{[f'(x)]^2}$$

y, en definitiva:

(5.15) $$\phi'(x) = \frac{f(x)f''(x)}{[f'(x)]^2}$$

dado que, $f(r) = 0$, por ser r una raíz de la función f, entonces reemplazando en (5.15), resulta:

(5.16) $$\phi'(r) = \frac{f(r)f''(r)}{[f'(r)]^2} = 0$$

Por otra parte, de la expresión (5.15), es posible deducir la expresión de la segunda derivada, que vale:

$$\phi''(x) = \frac{[f'(x)f''(x) + f(x)f'''(x)][f'(x)]^2 - f(x)f''(x)[2f'(x)f''(x)]}{[f'(x)]^4} =$$
$$= \frac{[f'(x)]^3 f''(x) + f(x)[f'(x)]^2 f'''(x) - 2f(x)f'(x)[f''(x)]^2}{[f'(x)]^4} =$$
$$= \frac{[f'(x)]^2 f''(x) + f(x)f'(x)f'''(x) - 2f(x)[f''(x)]^2}{[f'(x)]^3}$$

reemplazando en esta última expresión el valor de x por el de r, y considerando que $f(r) = 0$, resulta:

(5.17) $$\phi''(r) = \frac{f''(r)}{f'(r)}$$

que, por lo general, es no nula, como se quería demostrar. En consecuencia, reemplazando (5.17) en la (5.13), la expresión dada por:

(5.18) $$E_{n+1} \cong \frac{1}{2}\frac{f''(r)}{f'(r)} E_n^2$$

expresa la relación existente entre los errores de dos iteraciones o pasos consecutivos, cometidos en un proceso iterativo de *NEWTON-RAPHSON*, en las proximidades de la raíz.

Como ejemplos, resulta suficiente remitirse a aquellos expuestos en el punto correspondiente al método de *NEWTON-RAPHSON*, haciendo notar que las condiciones descriptas no se cumplen si $f'(x) = 0$; en cuyo caso, $x = r$ será una raíz múltiple. La convergencia será entonces, de primer orden.

♦♦♦ ●●● ♦♦♦

Caso 3.- $\phi'(r)=0\,;\phi''(r)=0\,;\phi'''(r)\neq 0.$

Adoptando un criterio análogo a los anteriores y realizando toda la operatoria que involucra el desarrollo del tema, resulta la siguiente relación de errores:

(5.19) $$E_{n+1} \cong \frac{1}{3!}\phi'''(r)E_n^3$$

Este caso, el cual se presenta rara vez en la práctica, está caracterizado por obtenerse una convergencia muy rápida; no obstante, esta aparente ventaja se contrarresta por el hecho de tener, tanto la función como sus sucesivas derivadas, expresiones mucho más complejas que en los casos de convergencia de menor orden; de manera tal, que el tiempo ganado debido a la rapidez de convergencia, es perdido por la dificultad que entraña la evaluación de la función y sus derivadas.

Dadas las circunstancias apuntadas, se dice que se trata de una **ITERACIÓN DE TERCER ORDEN** o bien que, este caso es de **TERCER ORDEN DE CONVERGENCIA**, del mismo modo en que fueron denominados los casos anteriores.

Siguiendo una metodología similar, pueden ser definidos órdenes de iteración o convergencia más altos, aunque rara vez se presentan en la práctica; no obstante, la ventaja que brinda el aumento en la velocidad de convergencia de los mayores órdenes, se ve totalmente neutralizada por la engorrosa evaluación de la función y sus sucesivas derivadas.

5.4.- PROCESO DELTA - CUADRADO DE AITKEN

Un método idóneo para acelerar la convergencia de cualquier fórmula recursiva o proceso iterativo de primer orden, es el denominado **DELTA - CUADRADO de AITKEN** $\left(\Delta^2\right)$, que se detalla seguidamente.

Sean x_{n-1} ; x_n ; x_{n+1} tres aproximaciones sucesivas y consecutivas de la raíz r de la ecuación $f(x)=0$ obtenidas mediante cualquier método iterativo de *PRIMER ORDEN DE CONVERGENCIA*; en cuyo caso se sabe que los errores E_{n-1} ; E_n ; E_{n+1} correspondientes, están dispuestos aproximadamente, según una progresión geométrica; en otras palabras, es posible escribir que:

$$\frac{E_{n+1}}{E_n} \cong \frac{E_n}{E_{n-1}}$$

o, lo que resulta equivalente:

$$\frac{x_{n+1}-r}{x_n-r} \cong \frac{x_n-r}{x_{n-1}-r}$$

la que, resuelta en términos de la variable r, resulta:

$$r \cong \frac{x_{n-1}x_{n+1}-x_n^2}{x_{n+1}-2x_n+x_{n-1}}$$

Sumando y restando al segundo miembro de esta última expresión, el término x_{n+1}, se obtiene:

$$r \cong x_{n+1} + \frac{x_{n-1}x_{n+1}-x_n^2}{x_{n+1}-2x_n+x_{n-1}} - x_{n+1} = x_{n+1} - \frac{x_{n+1}^2-2x_{n+1}x_n+x_n^2}{x_{n+1}-2x_n+x_{n-1}}$$

y, en definitiva:

(5.20) $$r \equiv x_{n+1} - \frac{(x_{n+1} - x_n)^2}{x_{n+1} - 2x_n + x_{n-1}}$$

La metodología de cálculo a emplear, haciendo uso de la expresión dada por (5.20), puede resumirse en los siguientes puntos:

1.- Mediante la iniciación del procedimiento con $x=x_0$, de cualquier algoritmo iterativo de primer orden, se calculan dos aproximaciones sucesivas x_1 ; x_2 de la raíz r que, juntamente con la primera aproximación x_0 constituyen la terna de base del método de AITKEN,

2.- Haciendo uso de la expresión (5.20) se calcula una cuarta aproximación a la raíz r que, si satisface las condiciones de precisión previamente establecidas para el cálculo, se toma como tal,

3.- De no resultar satisfactoria la aproximación obtenida en el paso anterior, es utilizada como primera aproximación para hallar otros dos valores sucesivos de la raíz, mediante el método iterativo original,

4.- Se reiteran los puntos 2 y 3 hasta satisfacer las condiciones de precisión previamente establecidas para la raíz.

Con el objeto de aportar mayores detalles en cuanto a la aplicación práctica del método *DELTA - CUADRADO de AITKEN*, se expone el siguiente:

Ejemplo 5.5.- Utilizando el método de iteración, conjuntamente con la aceleración de la convergencia de AITKEN, determinar la raíz comprendida en el intervalo *(1;2)* de la ecuación:

$$e^x - x^2 - 3 = 0$$

con una aproximación de cuatro cifras decimales exactas.

Solución: Primeramente, y según las condiciones establecidas, es necesario volver a escribir la ecuación dada bajo la forma:

$$x = \phi(x) = \sqrt{e^x - 3}$$

de donde, puede deducirse que:

$$\phi'(x) = \frac{e^x}{2\sqrt{e^x - 3}}$$

En consecuencia, comenzando el procedimiento con $x_0 = 1$, es negativa la cantidad subradical del denominador, por lo tanto resulta conveniente hacer $x_0 = 1{,}1$. Con ello:

$$\phi'(x) \equiv 23{,}27$$

cantidad que, en valor absoluto, resulta mucho más grande que la unidad; por lo tanto, no se generará un proceso convergente. Dada esta circunstancia, resulta imprescindible escribir la ecuación en forma diferente. Sea:

$$x = \phi(x) = \ln\left(x^2 + 3\right)$$

de donde:

$$\phi'(x) = \frac{2x}{x^2 + 3}$$

y finalmente, tomando $x_0 = 1$, resulta:

$$\phi'(1) = 0{,}5$$

valor aceptable, pues, sólo se requiere que $|\phi'(r)| < 1$. Por lo tanto, la relación recursiva:

$$x_{n+1} = \ln (x_n^2 + 3)$$

con $x_0 = 1$, es idónea para iniciar el procedimiento descripto, resultando:

$$x_0 = 1 \quad ; \quad x_1 = 1{,}38629 \quad ; \quad x_2 = 1{,}59367$$

Utilizando los valores hallados con el objeto de la aplicación de la expresión (5.20), se obtiene:

$$x_3 = 1,83405$$

Dada la dispersión de valores obtenidos se impone la prosecución de la metodología descripta. Aplicando nuevamente el método de iteración original, da como resultado:

$$x_4 = 1,85062 \quad ; \quad x_5 = 1,86016$$

valores que, juntamente con el de x_3 y la reiteración de la aplicación de la expresión (5.20), arroja:

$$x_6 = 1,87311$$

Operando de la misma manera; vale decir, tomando este último valor como primera aproximación del método de iteración, resultan:

$$x_7 = 1,87311 \quad ; \quad x_8 = 1,87311$$

Como puede apreciarse, en los tres últimos resultados no se ha obtenido mejoría alguna, pudiéndose aceptar $r=1,87311$ como valor de la raíz con todas sus cifras decimales exactas.

♦♦♦ ●●● ♦♦♦

Si se consulta la solución de este mismo ejemplo, dado en el punto correspondiente al método del *INTERVALO MEDIO* se ve que, en aquel caso fueron necesarios doce pasos para lograr la precisión deseada; mientras que, haciendo uso de la fórmula de aceleración de *AITKEN*, solo se necesitaron seis. Si se prueba la solución, con el método de *ITERACIÓN* se verá que serán necesarias más de veinte iteraciones para obtener las cinco cifras decimales exactas.

Esta situación se hace cada vez mas notoria a medida que las condiciones del problemas son más extremas; en otras palabras, cuando por la forma de la ecuación en el intervalo de cálculo utilizado, la velocidad de convergencia es un factor crítico.

5.5.- MÉTODO DE SEGUNDO ORDEN DE NEWTON

El método descripto en el punto anterior, solo es válido si se está en presencia de una expresión recursiva de primer orden. En ocasiones resulta necesario acelerar la convergencia en procesos iterativos de segundo orden, principalmente el de *NEWTON-RAPHSON*, dado que es el método normalmente utilizado en las funciones preprogramadas de todos los lenguajes de computación, en la actualidad.

En estas circunstancias, resulta útil el *MÉTODO DE SEGUNDO ORDEN DE NEWTON*, que tiene la ventaja de una muy rápida convergencia a la solución deseada, lo que permite obtener una aproximación extremadamente cercana al valor de la raíz con un bajo número de pasos y un mínimo de cálculo. Sin embargo, como se ha estudiado, desde un punto de vista práctico, este método está limitado a su utilización en ecuaciones que tienen derivadas de mayor orden (por lo menos de segundo), relativamente fáciles de programar y calcular, ya que el tiempo que debe ser empleado en esto, puede sobrepasar de manera exagerada la ventaja que representa la rápida convergencia.

Para introducirnos en el estudio del método que nos ocupa, considérese una ecuación de la forma:

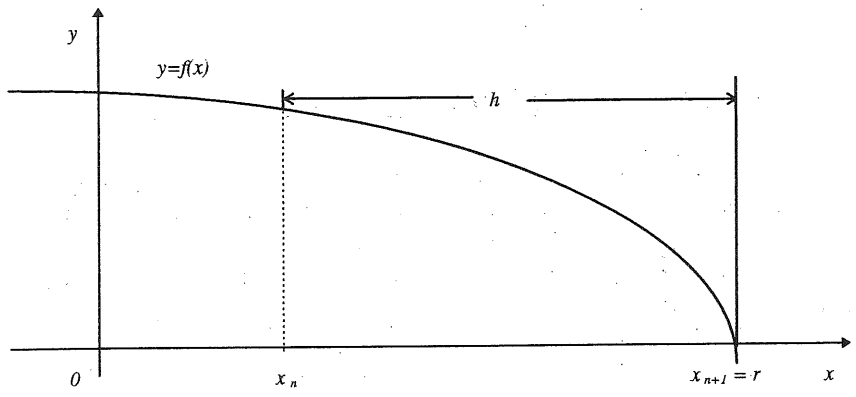
$$f(x) = 0$$

y supóngase que se ha determinado por cualquiera de los métodos estudiados, un valor aproximado de la raíz, el que puede ser uno de los extremos de algún intervalo de separación y llamando $x = x_n$ a este punto.

Desarrollando la función $f(x)$ en serie de *TAYLOR* con respecto a $x = x_n$ en el punto $x_{n+1} = x_n + h$, se obtiene:

$$f(x_{n+1}) = f(x_n) + f'(x_n)h + \frac{f''(x_n)h^2}{2!} + \ldots \tag{5.21}$$

Si h fuera el incremento particular de x para el cual la serie dada por (5.21) se redujera a cero, la cantidad $x_n + h$ sería la raíz exacta, como se muestra en la figura 5.4.



- Figura 5.4 -

Vale decir, haciendo uso de solamente los tres primeros términos de la serie dada por (5.21), resulta:

$$f(x_n)+h\left[f'(x_n)+\frac{f''(x_n)h}{2}\right]=0 \tag{5.22}$$

Un valor aproximado de h, determinado a partir de la expresión (5.22) y sumado a x_n no proporcionará el valor exacto de la raíz, ya que fueron utilizados para su cálculo, únicamente los tres primeros términos de la serie infinita (5.21); sin embargo, se obtendrá un valor muy próximo; es decir, una aproximación mucho mejor de la raíz.

Sustituyendo el valor de h encerrado dentro del corchete por la expresión dada por *NEWTON-RAPHSON*, que es:

$$h = x_{n+1} - x_n = -\frac{f(x_n)}{f'(x_n)}$$

se obtiene:

$$f(x_n)+h\left[f'(x_n)-\frac{f''(x_n)f(x_n)}{2f'(x_n)}\right]=0$$

y despejando el valor de h, resulta:

$$h = x_{n+1} - x_n = -\frac{f(x_n)}{f'(x_n)-\frac{f''(x_n)f(x_n)}{2f'(x_n)}}$$

finalmente, despejando x_{n+1}, se obtiene:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)-\frac{f''(x_n)f(x_n)}{2f'(x_n)}} \tag{5.23}$$

Mediante la aplicación sucesiva de la expresión dada por (5.23), es posible calcular en cada paso, aproximaciones cada vez más cercanas a la raíz, con una muy elevada velocidad de convergencia.

Es fácil observar que el método de *SEGUNDO ORDEN DE NEWTON* para funciones comprendidas en el segundo orden de convergencia, es el equivalente del método *DELTA - CUADRADO de AITKEN*, aplicado a ecuaciones comprendidas dentro del primer orden de convergencia, para acelerar la misma.

5.6.- RESOLUCIÓN DE ECUACIONES ALGEBRAICAS

Hasta el presente punto se ha estudiado una abundante variedad de métodos aproximados de resolución numérica de ecuaciones, sin especificar ningún tipo de condiciones formales sobre las mismas; vale decir, podría tratarse de ecuaciones trascendentes; vale decir, que incluyan funciones logarítmicas, trigonométricas, etc. Cuando se trata de ecuaciones algebraicas polinomiales de la forma:

$$P(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n = 0 \tag{5.24}$$

todos los métodos estudiados son igualmente válidos; no obstante, por ciertas características propias de las ecuaciones algebraicas del tipo (5.24), es posible aplicar algunos métodos especiales para acotar, separar y aproximar raíces.

Llegado a este punto se impone el uso de algunos conocidos resultados del álgebra y el análisis matemático, tales como la regla de *Ruffini*, la descomposición factorial de un polinomio, la relación entre los valores de los coeficientes y el de las raíces, la relación entre el orden de multiplicidad de una raíz y los valores de sus derivadas, etc.; por todo lo cual, será necesario repasar previamente, todos los teoremas referidos a los temas mencionados.

El *TEOREMA FUNDAMENTAL DEL ÁLGEBRA* y sus corolarios asociados aseguran la existencia de n raíces en una ecuación algebraica de grado n de la forma (5.24). Para calcular dichas raíces, en el caso de una ecuación de primer grado, solo son necesarias las cuatro operaciones racionales, a saber: adición, sustracción, multiplicación y división.

Para resolver las ecuaciones de segundo grado, es imprescindible agregar a la lista de operaciones mencionadas, la extracción de raíces cuadradas. De la misma manera, en las fórmulas de resolución de ecuaciones de tercer grado, aparecen raíces cúbicas; algo similar sucede cuando se trata de resolver ecuaciones de cuarto grado. Es posible demostrar que, para ecuaciones de grado superior al cuarto $n>5$, no existe ninguna solución algebraica.

La última afirmación realizada en el párrafo anterior, se encuentra demostrada por el teorema de *RUFFINI-ABEL*; la imposibilidad que racionalmente señala este teorema conduce necesariamente, a elaborar otros métodos para resolver las ecuaciones algebraicas de quinto grado o superior, en el supuesto de poseer formas particulares; en caso contrario, no existen más opciones que las estudiadas en los puntos anteriores de este mismo capítulo.

Fundamentalmente, en todos aquellos casos en que la solución de una ecuación algebraica de la forma (5.24) sea factible, deberán seguirse los mismos pasos que se han indicado al estudiar el problema general; en otras palabras, será necesario primeramente, acotar las raíces, luego separarlas y finalmente aproximarlas, como será estudiado a continuación. No obstante lo dicho, aún en el caso de ser factible resolver ecuaciones algebraicas mediante los métodos especiales, que seguidamente se estudiarán, en ocasiones aquellas pueden ser resueltas con ventaja mediante los métodos generales ya estudiados.

5.6.1.- ACOTACIÓN DE LAS RAÍCES REALES

Se debe recordar que, acotar las raíces significa determinar un intervalo $(M;N)$ tal que contenga todas las raíces reales de la ecuación dada; o al menos, todas las que se quieren calcular.

Para hallar un valor N que supere a todas las raíces de la ecuación, es suficiente descomponer el polinomio dado en varios polinomios que se conserven positivos para todo valor $x > N$. Entonces, $P(x)$ será positivo para $x > N$, y al no anularse en $(N;\infty)$, en este subintervalo no habrá raíces.

Con el objeto de realizar un aporte práctico, referido a este tema, se expone el siguiente:

Ejemplo 5.6.- Sea el problema de encontrar un valor N que acote superiormente las raíces de la ecuación:

$$P(x) = x^5 - x^4 - 6x^3 + x^2 + 6x + 2 = 0$$

Solución: En esta ecuación es posible agrupar términos de tal manera que pueda ser escrita de la siguiente forma:

$$P(x) = x^3 (x^2 - x - 6) + x^2 + 6x + 2 = 0$$

donde, el valor de lo representado dentro del paréntesis de la expresión anterior, se conserva positivo para valores de $x > 3$. Siendo los otros términos positivos para $x > 0$; por lo tanto, resulta entonces $N = 3$.

♦♦♦ ●●● ♦♦♦

Para acotar las raíces negativas es necesario hacer la sustitución $y = -x$, operando de idéntica manera que la realizada para determinar una cota superior, tal como se muestra en el siguiente:

Ejemplo 5.7.- Sea tener que acotar las raíces negativas de la ecuación dada en el ejemplo 5.6, propuesto.

Solución: Retomando el ejemplo dado, y luego haciendo la sustitución propuesta, resulta la expresión:

$$P(x) = -y^5 - y^4 + 6y^3 + y^2 - 6y + 2 = 0$$

en la que, multiplicando todos sus términos por *-1*, y agrupando convenientemente los mismos, se obtiene:

$$P(x) = y^3 \ (y^2 - 6) + y^2 \ (y^2 - 1) + 2 (3 y - 1) = 0$$

donde es posible apreciar que el valor del primer paréntesis se conserva mayor que cero para valores de $y > 3$, el segundo para valores de $y > 1$, y el tercero para valores de $y > 1$; por lo tanto, $M = mín\ [- y] = - 3$.

♦♦♦ ●●● ♦♦♦

Con los resultados que se han obtenido en los dos ejemplos anteriores, es posible afirmar que las raíces reales de la ecuación dada en aquellos, se encuentran en el intervalo *(- 3 ; 3)*.

5.6.2.- SEPARACIÓN DE LAS RAÍCES REALES

Existen varios métodos para determinar los distintos subintervalos de separación de las raíces reales de una ecuación algebraica partiendo, ya sea de un intervalo de acotación conocido, ya sea sin poseer previamente un conocimiento de él. Para este último caso, es posible apelar a un método basado en tanteos, el cual, es de muy simple programación y rápido procesamiento; sobre los mismos, por estos motivos, no se brindarán mayores detalles.

No obstante se estudiará a continuación un método general que resuelve el problema de la separación.

5.6.3.- TEOREMA DE STURM

Un método sistemático que permite separar las raíces reales de un polinomio, a partir de un intervalo de acotación previamente calculado, está dado por el denominado **TEOREMA DE STURM.**

Sea $f(x)$ una ecuación algebraica de la forma dada en la expresión (5.24), sin raíces múltiples. Calculando la derivada $f'(x)$, que será designada con $f_1(x)$; haciendo el cociente $f(x)/f_1(x)$ se obtiene un resultado que se denominará con la notación $C_1(x)$ y un resto llamado $R_1(x)$, el que, cambiado de signo será denominado $f_2(x)$. De la misma manera, se efectúa $f_1(x)/f_2(x)$, al resto de la cual, cambiado de signo, se lo designa con $f_3(x)$; y así se continúa hasta llegar a un resto constante, que no puede ser nulo si originariamente se ha supuesto que la ecuación no tiene raíces múltiples. Simbólicamente se debe proceder como se muestra en el cuadro 5.1, siguiente.

Funciones	***Cocientes***	***Restos***
$f(x)$	-	-
$f_1(x) = f'(x)$	$C_1(x) = f(x)/f_1(x)$	$R_1(x)$
$f_2(x) = - R_1(x)$	$C_2(x) = f_1(x)/f_2(x)$	$R_2(x)$
. . . . .		
$f_{n-1}(x) = - R_{n-2}(x)$	$C_{n-1}(x) = f_{n-2}(x)/f_{n-1}(x)$	$R_{n-1}(x)$
$f_n(x) = - K$	$C_n(x) = f_{n-1}(x)/f_n(x)$	-

- Cuadro N° 5.1 -

La sucesión $f(x); f_1(x);...; f_{n-1}(x); f_n(x)$ es la denominada *SUCESIÓN DE POLINOMIOS DE STURM.* Para cada valor numérico de la variable x, esta sucesión cambia un cierto número de veces de signo, y a este número se lo designa como *VARIACIÓN* de la sucesión para ese valor determinado de x. Con todos estos elementos *STURM* elaboró el siguiente:

Teorema de STURM: *" El número de raíces reales de una ecuación algebraica de la forma (5.24), sin raíces múltiples, contenidas en un intervalo (a;b) es igual a la diferencia entre el número de las variaciones de signo de la sucesión de STURM para los valores $x = a$ y $x = b$. "*

del cual, por ser de relativo interés práctico para la materia, no se expondrá la correspondiente demostración. Con el objeto de mostrar los aspectos prácticos de la aplicación del teorema de Sturm, se presenta el sencillo ejemplo siguiente:

Ejemplo 5.8.- Se trata de resolver el problema de separar las raíces reales de la ecuación:

$$P(x) = x^3 + x^2 - 2x - 1 = 0$$

que, se sabe, se encuentran acotadas en el intervalo *(-2 ; 2)*.

Solución: Por medio de la aplicación directa del algoritmo del teorema de Sturm, resulta la secuencia mostrada en el cuadro *(5.2)*.

Funciones		***Signos***			
	x	f	f_1	f_2	f_3	***Variac.***
$f(x) = x^3 + x^2 - 2x - 1$	- 2	-	+	-	+	3
$f_1(x) = 3x^2 + 2x - 2$	- 1	+	-	-	+	2
$f_2(x) = {}^{14}/_{9}x + {}^{7}/_{9}$	0	-	-	+	+	1
$f_3(x) = {}^{35}/_{9}$	1	-	+	+	+	1
	2	+	+	+	+	0

- Cuadro Nº 5.2 -

donde es posible apreciar que se pierde una variación dentro del intervalo *(-2 ; -1)*, otra se pierde en el *(-1 ; 0)*, y una tercera en *(1 ; 2)*. La ecuación propuesta tiene entonces tres raíces reales y distintas, una en cada uno de los subintervalos mencionados.

◆◆◆ ●●● ◆◆◆

Capítulo 6

Ecuaciones Lineales Simultáneas

6 . - ECUACIONES LINEALES SIMULTÁNEAS

Introducción. Sistemas no Homogéneos. Método de Eliminación de GAUSS. Método de GAUSS-JORDAN. Método de las Ecuaciones de ERROR. Método de la Matriz Inversa. Método de GAUSS-SEIDEL. Sistemas de Ecuaciones Homogéneas. Generalidades. Método de FADDEEV-LEVERRIER. Método de las Potencias. Ecuaciones Lineales Simultáneas Ortogonales. Introducción. El Método Analítico. Método de Ortogonalización de GRAM-SCHMIDT.

6.1.- INTRODUCCIÓN

Del estudio del álgebra, se deducen varios métodos para resolver *ECUACIONES LINEALES SIMULTÁNEAS*, de los cuales, los principales y más utilizados son el método de *ELIMINACIÓN o REDUCCIÓN* de las incógnitas mediante combinaciones lineales de las ecuaciones (método de *GAUSS*), como así también, haciendo uso de los determinantes (regla de *CRAMER*). En caso de ser necesario resolver sistemas de orden 3×3 (tres ecuaciones con tres incógnitas), parece ser que la regla de CRAMER tiene ventaja sobre el de reducción. Sin embargo el método de los determinantes, con desarrollo por medio de menores, es muy poco práctico o inútil cuando se deben resolver simultáneamente grandes números de ecuaciones. Es posible demostrar que la solución de un sistema de n ecuaciones con n incógnitas, mediante la regla de CRAMER, calculando los determinantes por medio de menores, requiere $(n-1).(n+1)\,!$ productos. Así, la solución de un sistema de 13 ecuaciones, con igual número de incógnitas, requerirá al rededor de 1.046.139.494.400 multiplicaciones.

Dado que, las computadoras más rápidas en la actualidad realizan al rededor de 27.000.000 de operaciones por segundo, demandarían aproximadamente 11 horas para resolverlo. A partir de estas cifras, como los productos que son necesarios realizar crecen proporcionalmente al factorial de $n+1$ no es difícil deducir la inutilidad del método para sistemas de más de seis a diez ecuaciones, con igual número de variables.

Los sistemas de ecuaciones a los que se ha hecho referencia serán de n ecuaciones con m incógnitas; vale decir, de orden $n \times m$, con el objeto de tratar el problema en forma absolutamente general. Estos sistemas, expresados en forma cartesiana vienen dados por la siguiente expresión:

$$(6.1) \qquad \begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m = C_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m = C_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m = C_n \end{cases}$$

En ocasiones, libremente será utilizada la forma de sumatorias, de la siguiente manera:

$$(6.2) \qquad \sum_{j=1}^{m} a_{ij}\, x_j = C_i \qquad para \quad i = 1;2;\ldots;n$$

o, de manera más resumida aún, haciendo uso de la notación matricial toman la forma:

$$(6.3) \qquad \mathbf{A\,X = C}$$

donde $\mathbf{A} = [a_{ij}]$ es la matriz de orden $(n \times m)$ de los coeficientes que componen el sistema; **X** y **C** son los vectores columna correspondientes a las incógnitas y los términos independientes, respectivamente.

Si los términos independientes C_i no son simultáneamente nulos, el sistema se denomina **NO HOMOGÉNEO**; en cambio, si todos los términos independientes valen cero, aquel recibe el nombre de **HOMOGÉNEO**. En el primer caso, todas las ecuaciones deben ser independientes, o más racionalmente, linealmente independientes, para que se puedan obtener soluciones *ÚNICAS*; en el segundo caso, se sabe que solo existen soluciones *NO TRIVIALES* si todas las ecuaciones no son linealmente independientes.

Si a la matriz **A** se le adiciona una columna formada por los términos independientes **C**, se forma la denominada matriz *ORLADA, AUMENTADA o AMPLIADA* del sistema.

Todo lo referente al comportamiento algebraico de los sistemas de ecuaciones, se encuentra sintetizado por el *TEOREMA DE ROUCHE-FROBENIUS*. Llamando:

r : característica de la matriz del sistema,
r' : característica de la matriz orlada,
m : número de variables del sistema, y
n : número de ecuaciones del sistema

es posible afirmar que:

"La condición necesaria y suficiente para que un sistema de **n** *ecuaciones lineales con* **m** *incógnitas sea COMPATIBLE es que la matriz* **A** *del sistema y la matriz ORLADA tengan la misma característica; en cuyo caso, si la característica es igual al número de incógnitas, el sistema es DETERMINADO; si es menor, el sistema es INDETERMINADO; vale decir, admite infinitas soluciones; cada una de ellas, queda determinada dando un sistema arbitrario de valores a* **m - r** *incógnitas. Si por el contrario, la matriz* **A** *del sistema y la matriz ORLADA no tienen idéntica característica, el sistema es INCOMPATIBLE y no tiene ninguna solución".*

Resumiendo, si:

$$\begin{cases} r \neq r' : \text{Incompatible } (\textit{el sistema no tiene solucion}) \\ r = r' : \text{Compatible} \begin{cases} r = m : \text{Determinado } (\textit{una solucion}) \\ r < m : \text{Indeterminado } (\infty\ \textit{soluciones}) \end{cases} \end{cases}$$

donde, en el último caso, $m - r$ recibe el nombre de **GRADO DE INDETERMINACIÓN**.

Puede también, darse el caso (no contemplado en el teorema de *ROUCHÉ-FROBENIUS*) que sea $r>m$, en cuyo caso se dice que el sistema es **SOBRE DETERMINADO**, existiendo infinitas soluciones representadas por los puntos encerrados dentro de un poliedro convexo del espacio de m dimensiones, las que, podrían considerarse como soluciones aproximadas del sistema.

En lo sucesivo se supondrá que **A** es una matriz $n \times m$, formada por ecuaciones linealmente independientes; en cuyo caso, el sistema será determinado cuando el determinante de la matriz **A** es no nulo.

Los métodos para resolver numéricamente, este tipo de sistemas pueden ser subdivididos en dos clases:

- MÉTODOS DIRECTOS, y
- MÉTODOS ITERATIVOS.

Los métodos directos, provistos por la matemática pura, son aquellos que llevan a una solución exacta del problema luego de un número finito de pasos. Este número depende exclusivamente de la cantidad de ecuaciones que componen el sistema. No obstante, no ser métodos numéricos, por su innegable utilidad, serán estudiados.

El error con que se obtienen los resultados haciendo uso de algún método directo se debe, si no hubiese errores inherentes en los parámetros, nada más que al redondeo que es realizado durante los cálculos. Si el proceso se realiza en forma manual, al menos teóricamente, este error no existe si se opera con todas las cifras decimales. No sucede lo mismo si el proceso es ejecutado mediante computadoras las que, son capaces de soportar solamente una cantidad determinada de cifras decimales.

Los métodos iterativos en cambio, son estrictamente numéricos y dan una solución aproximada del sistema de ecuaciones lineales, obtenida como límite de una sucesión de vectores construida mediante un proceso de aproximaciones sucesivas, como se verá más adelante.

6.2.- SISTEMAS NO HOMOGÉNEOS

Se ha definido en el punto anterior a los sistemas **NO HOMOGÉNEOS**, como aquellos en los cuales los coeficientes o términos independientes C_i de las ecuaciones no son todos simultáneamente nulos. A continuación se estudiarán algunos métodos de resolución de este tipo de sistemas.

6.2.1.- MÉTODO DE ELIMINACIÓN DE GAUSS

Dado un conjunto de n ecuaciones lineales simultáneas con n incógnitas; siendo aquellas linealmente independientes, el método de eliminación de **GAUSS** se reduce a lograr un sistema triangular equivalente, que a su vez se resuelve con facilidad mediante la denominada *SUSTITUCIÓN INVERSA*, como se verá a continuación.

El esquema de GAUSS comienza reduciendo un conjunto de ecuaciones simultáneas, tal como el (6.1), a un sistema de la forma:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1;n-1}x_{n-1} + a_{1n}x_n = C_1 \\ a'_{22}x_2 + \ldots + a'_{2;n-1}x_{n-1} + a'_{2n}x_n = C'_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a^{(n-2)}_{n-1;n-1}x_{n-1} + a^{(n-2)}_{n-1;n}x_n = C^{(n-2)}_{n-1} \\ a^{(n-1)}_{n;n}x_n = C^{(n-1)}_n \end{cases} \tag{6.4}$$

en la cual los superíndices indican los nuevos coeficientes que se forman en el proceso de reducción. En la práctica esto se obtiene de la siguiente manera:

1.- La primera ecuación de (5.1) se divide por a_{11} ; resulta entonces:

$$x_1 + \frac{a_{12}}{a_{11}}x_2 + \ldots + \frac{a_{1n}}{a_{11}}x_n = \frac{C_1}{a_{11}} \tag{6.5}$$

Esta ecuación se multiplica por a_{21} (coeficiente de x_1 en la segunda ecuación) y lo que resulta se resta de la segunda ecuación, eliminando así x_1 de ésta última. Un proceso similar se sigue con todas las otras ecuaciones del sistema, eliminando en todas ellas x_1 ; excepto en la primera ecuación, resultando:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = C_1 \\ a'_{22}x_2 + \ldots + a'_{2n}x_n = C'_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a'_{n2}x_2 + \ldots + a'_{nn}x_n = C'_n \end{cases} \tag{6.6}$$

La ecuación utilizada para eliminar las incógnitas en las ecuaciones que la siguen se denomina **ECUACIÓN PIVOTE**. En la ecuación pivote, el coeficiente de la incógnita que se va a eliminar de las ecuaciones que la siguen se denomina **COEFICIENTE PIVOTE**, siendo obvio que debe ser distinto de cero.

2.- Aplicando un procedimiento similar a todas las ecuaciones con coeficientes de la forma a'_{ij}, se elimina x_2 desde la tercera en adelante. Este procedimiento, utilizando diferentes ecuaciones pivotes, se continúa hasta que el conjunto original de ecuaciones haya sido reducido a uno de la forma (6.4).

3.- Una vez obtenido el sistema triangular de ecuaciones, la última ecuación de este conjunto suministra directamente el valor de x_n . El valor obtenido se sustituye entonces en la penúltima ecuación para obtener el valor de x_{n-1} ; y así sucesivamente. Este procedimiento es el denominado de **SUSTITUCIÓN INVERSA**, al cual se ha hecho referencia en puntos anteriores.

El efecto de la propagación de los errores por redondeo puede introducir errores relativamente grandes en la solución; por esta razón, el número de ecuaciones simultáneas que se pueden resolver satisfactoriamente con este método, utilizando un soporte de 16 a 18 dígitos significativos (precisión extendida) en las operaciones, se limita generalmente a 20 ó 25, en el mejor de los casos.

Considérese ahora el método, desde el punto de vista de la implementación en la computadora. Como ésta puede manejar solamente datos numéricos, se requiere utilizar matrices al programar. El dato es pues, la matriz orlada del sistema donde, para conservar la homogeneidad, se debe hacer $C_i = a_{i;n+1}$. Revisando el procedimiento descripto en el paso 1.-, se puede observar que los elementos de la matriz reducida, la que será denominada **A'**, se pueden describir directamente a partir de la matriz orlada, utilizando la siguiente fórmula:

$$a'_{ij} = a_{ij} - \frac{a_{ik}}{a_{kk}} a_{kj} \qquad para \quad \begin{cases} k \le j \le m \\ k+1 \le i \le n \end{cases} \tag{6.7}$$

en la que:

a = elemento de la matriz orlada,
a' = elemento de **A'**,
i = número de la fila de las matrices,
j = número de la columna de las matrices,

k = identificación de la fila pivote,
n = cantidad de filas de las matrices, y
m = cantidad de columnas de las matriz orlada $(n+1)$.

Aplicando reiteradamente el proceso descripto y respetando las convenciones establecidas, mediante la ecuación:

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - \frac{a_{kj}^{(k-1)}}{a_{kk}^{(k-1)}} a_{ik}^{(k-1)} \qquad para \quad \begin{cases} k+1 \le j \le m \\ k+1 \le i \le n \end{cases} \tag{6.8}$$

en donde se han previamente definido los elementos que intervienen.

Con esto se llega a obtener un sistema triangular de la forma (6.4). Los superíndices corresponden, simplemente, a las primas utilizadas al identificar matrices reducidas sucesivas en las explicaciones anteriores. Obsérvese que en (6.8), el extremo inferior de j, es $k+1$, porque no hay necesidad de calcular los valores iniciales nulos que aparecen en cada fila, debido al proceso de reducción.

El proceso de sustitución inversa se puede generalizar, utilizando idéntica nomenclatura, de la siguiente manera:

$$x_n = \frac{a_{nm}}{a_{nn}} \quad ; \quad x_i = \frac{a_{im} - \sum_{j=i+1}^{n} a_{ij}\, x_j}{a_{ii}} \qquad para \;\; i = n-1; n-2; \ldots; 1 \tag{6.9}$$

Antes de concluir, resulta oportuno realizar algunas consideraciones adicionales. Se ha supuesto tácitamente que, cada elemento pivote que se encuentra en el proceso de reducción ha sido no nulo. Si este no fuera el caso, el procedimiento discutido deberá ser modificado de tal modo que permita el intercambio de filas para que, la nueva no tenga elemento pivote nulo.

Otra consideración a tener en cuenta es el efecto que sobre la precisión de la solución hallada, tienen las magnitudes, en valor absoluto, de los elementos pivotes sucesivamente utilizados. Es posible demostrar que si la magnitud de los elementos pivote es apreciablemente menor que el orden de magnitud, en general, de los otros elementos que componen la matriz, la utilización de dicho elemento pivote causará una disminución general de la precisión de la solución.

En consecuencia, para obtener mayor precisión se debe realizar cada reducción utilizando como fila pivote, la fila que tenga el mayor elemento pivote. Es recomendable, por lo tanto, contemplar el intercambio de filas. Esta posibilidad debe ser siempre incorporada a la codificación del programa de computadora que se utilice para resolver sistemas de ecuaciones lineales simultáneas.

6.2.2.- MÉTODO DE GAUSS-JORDAN

El de *GAUSS-JORDAN*, es también un método directo, que constituye una variación del método de eliminación de Gauss. Permite resolver hasta 20 ó 25 ecuaciones simultáneas, utilizando de 16 a 18 dígitos significativos en las operaciones aritméticas de la computadora. Se distingue del método de Gauss en el hecho de que cuando se elimina una incógnita, esta es eliminada de todas las ecuaciones restantes; vale decir que, la incógnita es eliminada tanto de las que preceden a la ecuación pivote, así como de las que la siguen.

Se consigue resolver por aplicación del método de *GAUSS-JORDAN* cumpliendo con los siguientes pasos:

1.- Dividir la primera ecuación del sistema por el coeficiente de la primera incógnita de esa misma ecuación,
2.- Formar otro sistema con la ecuación así obtenida y esa misma multiplicada por los coeficientes de la primera incógnita de las otras ecuaciones,
3.- Formar otro sistema restando la segunda y tercera ecuaciones del segundo sistema de sus homólogas del primer sistema y la primera sin modificaciones.

Los pasos descriptos se pueden realizar sistemáticamente utilizando las fórmulas siguientes:

$$b_{i-1;j-1} = a_{ij} - \frac{a_{1j}\,a_{i1}}{a_{11}} \qquad \begin{cases} 1 < i \le n \\ 1 < j \le m \end{cases} \tag{6.10}$$

y

$$b_{n;j-1} = \frac{a_{1j}}{a_{11}} \qquad \begin{cases} 1 < j \le m \\ a_{11} \ne 0 \end{cases} \tag{6.11}$$

Se utiliza la ecuación (6.10) para determinar todos los elementos de la nueva matriz, excepto los de la última fila, que se calculan mediante las expresiones (6.11). Indicando en ellas cada nomenclador:

i = número de la fila de la matriz original,
j = número de la columna de la matriz original,
m = número máximo de columnas $(n+1)$,
n = número máximo de filas,
a = elementos de la matriz procesada, y
b = elementos de la matriz resultante.

Para ilustrar más acabadamente el método de *GAUSS-JORDAN* estudiado, se expone el siguiente ejemplo práctico:

Ejemplo 6.1.- Sea resolver mediante el método de *GAUSS-JORDAN* el siguiente sistema de ecuaciones:

$$\begin{cases} 2x_1 - 2x_2 + 5x_3 = 13 & (a) \\ 2x_1 + 3x_2 + 4x_3 = 20 & (b) \\ 3x_1 - x_2 + 3x_3 = 10 & (c) \end{cases} \tag{6.12}$$

Solución: Se comenzará por la aplicación del primer paso; vale decir, dividiendo la primera ecuación por el coeficiente de la primera incógnita en esa misma ecuación, de lo que resulta (6.13) *(a)*; luego se multiplica la ecuación así obtenida, respectivamente, por el coeficiente de la primera incógnita en cada una de las ecuaciones restantes, obteniéndose:

$$\begin{cases} x_1 - x_2 + \frac{5}{2}x_3 = \frac{13}{2} & (a) \\ 2x_1 - 2x_2 + 5x_3 = 13 & (b) \\ 3x_1 - 3x_2 + \frac{15}{2}x_3 = \frac{39}{2} & (c) \end{cases} \tag{6.13}$$

A continuación, se resta la expresión (6.13) *(b)* de la (6.12) *(b)* y la (6.13) *(c)* de la (6.12) *(c)* convirtiendo a la (6.13) *(a)* directamente en la (6.14) *(c)*; de esta manera, se obtiene:

$$(6.14)\qquad \begin{cases} 5x_2 - x_3 = 7 & (a) \\ 2x_2 - \frac{9}{2}x_3 = -\frac{19}{2} & (b) \\ x_1 - x_2 + \frac{5}{2}x_3 = \frac{13}{2} & (c) \end{cases}$$

Repitiendo el proceso, se divide la ecuación (6.14) *(a)* por el coeficiente de la primera incógnita en esa ecuación para obtener (6.15) *(a)*. Luego se multiplica la (6.15) *(a)* obtenida, respectivamente, por el coeficiente del término x_2 en cada una de las ecuaciones (6.14) *(b)* y (6.14) *(c)*. Esto produce el siguiente sistema:

$$(6.15)\qquad \begin{cases} x_2 - \frac{1}{5}x_3 = \frac{7}{5} & (a) \\ 2x_2 - \frac{2}{5}x_3 = \frac{14}{5} & (b) \\ -x_2 + \frac{1}{5}x_3 = -\frac{7}{5} & (c) \end{cases}$$

Seguidamente se resta la (6.15) *(b)* y la (6.15) *(c)* de la (6.14) *(b)* y la (6.14) *(c)*, respectivamente, y se hace que la ecuación (6.15) *(a)* se convierta en la (6.16) *(c)*, obteniéndose:

$$(6.16)\qquad \begin{cases} -\frac{41}{10}x_3 = -\frac{123}{10} & (a) \\ x_1 + \frac{23}{10}x_3 = \frac{79}{10} & (b) \\ x_2 - \frac{1}{5}x_3 = \frac{7}{5} & (c) \end{cases}$$

La repetición sistemática del proceso lleva al siguiente sistema:

$$(6.17)\qquad \begin{cases} x_3 = 3 & (a) \\ \frac{23}{10}x_3 = \frac{69}{10} & (b) \\ -\frac{1}{5}x_3 = -\frac{3}{5} & (c) \end{cases}$$

De este último sistema, finalmente, es posible rescatar los valores de las incógnitas $x_1 = 1$; $x_2 = 2$; $x_3 = 3$, lo que constituye la solución deseada del sistema de ecuaciones propuesto.

••• ●●● •••

En virtud de las posibles distorsiones que podrían producirse desde el principio o, en cualquier punto de la aplicación del método, es necesario tomar una serie de precauciones, las que, serán puntualizadas a continuación.

Es posible que el elemento pivote a_{11} de una o más de las matrices obtenidas durante el proceso de eliminación pudiera ser nulo, en cuyo caso, el programa debería incluir la lógica necesaria para verificar que a_{11} en cada una de las matrices sea diferente de cero y efectuar, consecuentemente, el intercambio de filas si fuera necesario. El intercambio de renglones no alterará el orden de las incógnitas en la matriz columna final.

En aquellas ocasiones en que a_{11} sea muy pequeño en comparación con el orden de magnitud general de los otros elementos de la columna, su utilización podría disminuir la precisión, tal como en el método de Gauss. Es conveniente, pues, agregar al programa la codificación que permita comparar los valores absolutos de los coeficientes de la primera columna y poner como ecuación pivote la que tenga el mayor coeficiente en la primera incógnita, en valor absoluto, naturalmente.

6.2.3.- MÉTODO DE LAS ECUACIONES DE ERROR

El método de las **ECUACIONES DE ERROR**, no está destinado específicamente a resolver sistemas lineales de ecuaciones simultáneas sino que, más precisamente, se lo utiliza para mejorar la aproximación con que se pueden determinar las raíces de aquellas. Se trata de un método iterativo y general, que conviene ser aplicado como culminación de aquellos métodos que pudieran sufrir una pérdida considerable de precisión en los resultados finales, debido principalmente, al error por redondeo y por truncamiento que se acumula, en razón de su propagación, durante el gran número de operaciones aritméticas que es necesario efectuar en algunos métodos, para obtener la solución.

Sea el sistema, expresado en notación matricial, siguiente:

$$(6.18)\qquad \mathbf{A\ X = C}$$

y supóngase que se han obtenido las raíces x'_1 ; x'_2 ;...; x'_n por algún otro método. Sustituyendo estos valores en la ecuación (6.18) se encontrará que los valores de los términos independientes que son obtenidos c'_1 ; c'_2 ;...; c'_n varían respecto de los coeficientes dados porque las raíces sustituidas no fueron exactas.

La sustitución explicada en el párrafo anterior puede ser expresada por:

$$(6.19)\qquad \mathbf{A\ X' = C'}$$

donde **C'** es el vector columna de componentes c'_i .

Si se denomina **H** al vector columna de componentes h_1 ; h_2 ;...; h_n las que, a su vez, se consideran que son las correcciones que deben ser agregadas a los valores aproximados de las raíces para obtener el valor exacto de las mismas, se deduce naturalmente que:

(6.20) $$x_i = x'_i + h_i \quad ; \quad i = 1; 2;...; n$$

Si se sustituyen estas expresiones para las raíces exactas en la ecuación (6.18), se obtiene:

(6.21) $$\mathbf{A\ (X' + H) = C}$$

Si las ecuaciones (6.19) son restadas de las (6.21), resulta el siguiente sistema que encierra las correcciones de las raíces:

(6.22) $$\mathbf{A\ H = C - C' = E}$$

donde, los términos independientes $e_i = c_i - c'_i$ de este sistema, denotan el error que resulta de aplicar los valores poco precisos de las raíces, y las ecuaciones reciben el nombre de **ECUACIONES DE ERROR**.

Si se deseara todavía más precisión, se pueden hacer correcciones a los valores obtenidos para las correcciones. Designando las correcciones de las correcciones mediante h' , entonces, las raíces más precisas x_i estarían dadas por:

(6.23) $$x_i = x'_i + h_i + h'_i + ...$$

Este proceso de utilizar ecuaciones de error para corregir las raíces de otras ecuaciones de error puede ser iterado tanto como sea necesario hasta obtener las raíces con la aproximación deseada.

6.2.4.- MÉTODO DE LA MATRIZ INVERSA

En aquellas ocasiones en que resulta necesario resolver un gran número de sistemas que difieren únicamente en los valores de los términos independientes, es posible utilizar con ventaja las capacidades que brinda el método de la **MATRIZ INVERSA** para reducir substancialmente el número de las operaciones aritméticas requeridas para resolverlos.

Escrito en notación matricial, considérese las siguientes ecuaciones lineales simultáneas:

(6.24) $$\mathbf{A\ X = C}$$

para el cual, si el conjunto dado de ecuaciones posee solución única, la matriz **A** de los coeficientes es no singular, y como tal, admite inversa $\mathbf{A^{-1}}$, tal que:

$$\mathbf{A\ A^{-1} = A^{-1}\ A = I}$$

donde **I** es la matriz unitaria o identidad, de orden n, cuya diagonal principal esta formada por unos, siendo los demás elementos ceros. Por otra parte, se sabe que premultiplicando cualquier matriz por la matriz identidad da como resultado la matriz original.

En consecuencia, premultiplicando ambos miembros de (6.24) por $\mathbf{A^{-1}}$, se obtiene:

$$\mathbf{A^{-1}\ A\ X = A^{-1}\ C}$$

de donde:

$$\mathbf{I\ X = A^{-1}\ C}$$

y, finalmente:

(6.25) $$\mathbf{X = A^{-1}\ C}$$

Conociendo, por tanto, el valor de la matriz inversa de coeficientes del sistema $\mathbf{A^{-1}}$, es posible determinar fácilmente los elementos de **X** para cualquier número de sistemas con idéntica matriz **A**, con solo realizar un producto de matrices.

En el capítulo correspondiente a los apéndices, del presente trabajo, puede encontrarse la metodología adecuada para realizar la inversión de matrices mediante computadoras electrónicas digitales.

6.2.5.- MÉTODO DE GAUSS-SEIDEL

Repasando los métodos de resolución de ecuaciones lineales simultáneas estudiados con anterioridad es posible comprobar que, ninguno de ellos es práctico si se trata de resolver sistemas con varias decenas o, peor aún, con centenares de ecuaciones e incógnitas, debido a la enorme cantidad de operaciones aritméticas que deben realizarse para lograr los resultados deseados, los que, consecuentemente, emplearían un tiempo excesivamente prolongado de procesamiento, a la vez que, por los efectos de la propagación de los errores que invalidan su utilización, esta resultaría poco ventajosa, como ya se ha estudiado.

El segundo factor, determinante, a ser tenido en cuenta en la elección de un método de resolución es la precisión, la cual, al aumentar el número de operaciones aritméticas disminuye drásticamente, debido a la acumulación de los efectos de los errores por redondeo y truncamiento que se producen en cada una de ellas.

En consecuencia, es posible afirmar que, ninguno de los métodos directos estudiados resulta bueno cuando se trata de resolver sistemas de orden mayor a 20 - 25 ecuaciones.

El método que aquí se presenta, el de **GAUSS-SEIDEL**, es un método clasificado como iterativo que, tiene como única desventaja, la de que no siempre converge a una solución o de que, en ocasiones lo hace muy lentamente. No obstante, convergirá siempre a una solución cuando, la matriz del sistema en cuestión cumpla con cierta condición. Primeramente veamos la siguiente:

Definición: *"Una matriz cuadrada* **A**, *recibe la denominación de* ***DIAGONALMENTE DOMINANTE*** *si:*

$$|a_{ii}| \geq \sum_{i \neq j} |a_{ij}| \quad ; \quad \forall i \tag{6.26}$$

con desigualdad estricta, para por lo menos una de las i. "

La condición para que el método de *GAUSS-SEIDEL*, aplicado a un sistema de ecuaciones lineales simultáneas, sea convergente es que, para cada fila o ecuación del sistema se cumpla que la magnitud, en valor absoluto, del coeficiente de una incógnita diferente sea mayor a la sumatoria, también en valor absoluto, de los otros coeficientes de dicha ecuación; vale decir que, la matriz **A**, de coeficientes del sistema debe ser **DIAGONALMENTE DOMINANTE**, según la definición dada.

La condición dada en el párrafo anterior es *SUFICIENTE*; es decir, de no ser satisfecha, la aplicación del método en cuestión, puede resultar convergente. Es difícil definir el margen mínimo por el que ese coeficiente debe superar a los otros para asegurar la convergencia y es aún más difícil predecir la velocidad de convergencia para alguna combinación de valores de los coeficientes cuando esa convergencia existe. Sin embargo, cuando el valor absoluto del coeficiente dominante para una incógnita diferente en cada ecuación es mayor que la suma de los valores absolutos de los coeficientes de las otras incógnitas de esa misma ecuación, la convergencia está asegurada.

Resumiendo, un sistema *DIAGONALMENTE DOMINANTE* es condición *SUFICIENTE* para asegurar la convergencia, pero, no es condición *NECESARIA*.

Se estudiarán a continuación los pasos que deben ser seguidos, necesariamente, para la aplicación del método de *GAUSS-SEIDEL*:

1.- *Asignar un valor inicial a cada incógnita. Si es posible hacer una hipótesis razonable, se mejorará substancialmente la rapidez de convergencia, si no, hágase una elección arbitraria.*

2.- *Partiendo de la primera ecuación, determinar un nuevo valor para la incógnita dominante, utilizando para las otras incógnitas los valores supuestos según lo descripto en el punto anterior.*

3.- *Pasar a la segunda ecuación y calcular el valor de la incógnita dominante, utilizando para ello, el ya calculado en el paso anterior y los valores supuestos en las otras incógnitas.*

4.- *Reiterar el procedimiento descripto, con todas las demás ecuaciones, siempre para la incógnita dominante, utilizando los últimos valores calculados. Completado este paso se dice que se ha concluido una* ***ITERACIÓN****.*

5.- Iterar los pasos 2; 3 y 4 hasta que el valor de cada incógnita difiera del valor respectivo obtenido en la iteración previa, en una cantidad menor que un E positivo y arbitrario previamente fijado.

Con el objeto de aportar mayores detalles sobre la metodología de aplicación del método de *GAUSS-SEIDEL* se expondrá el siguiente:

Ejemplo 6.2.- Sea el problema de tener que resolver haciendo uso del método de *GAUSS-SEIDEL*, con una cota de error estimada de $E \leq 0{,}02$, el siguiente sistema de ecuaciones lineales simultáneas:

$$\begin{cases} 2x_1 + 10x_2 + x_3 = 51 \\ x_1 + 2x_2 + 10x_3 = 61 \\ 10x_1 + x_2 + 2x_3 = 44 \end{cases}$$

Solución: Se puede comprobar que el sistema propuesto cumple con la condición suficiente de convergencia; es decir que, es un sistema cuya matriz de coeficientes **A** es diagonalmente dominante. Su convergencia, desde el punto de vista de la aplicación del método de *GAUSS-SEIDEL* está, por lo tanto, asegurada.

Es importante notar que, la precisión no depende exclusivamente del valor de *E*, ya que esta es una función que depende, además, de la velocidad de convergencia. O sea, mientras mayor sea la velocidad de convergencia, mayor será la precisión obtenida en los valores calculados de las incógnitas, para un *E* determinado.

Dado que es imposible, en este caso, hacer a priori una hipótesis razonable sobre el valor de las incógnitas, se supondrán para las mismas valores idénticamente nulos. De esta manera se deben realizar los siguientes pasos:

a)	$2 \cdot 0 + 10 \cdot x_2 + 0 = 51$	entonces	$x_2 = 5{,}1$
b)	$0 + 2 \cdot 5{,}1 + 10 \cdot x_3 = 61$	"	$x_3 = 5{,}08$
c)	$10 \cdot x_1 + 5{,}1 + 2 \cdot 5{,}08 = 44$	"	$x_1 = 2{,}874$

con todo lo cual, se completa la primera iteración. Una segunda iteración produce los siguientes resultados:

a)	$2 \cdot 2{,}874 + 10 \cdot x_2 + 5{,}08 = 51$	entonces	$x_2 = 4{,}0172$
b)	$2{,}874 + 2 \cdot 4{,}0172 + 10 \cdot x_3 = 61$	"	$x_3 = 5{,}00916$
c)	$10 \cdot x_1 + 4{,}0172 + 2 \cdot 5{,}00916 = 44$	"	$x_1 = 2{,}996448$

Comparando los valores de las incógnitas, obtenidos en la última iteración con los de la primera, se puede apreciar que las diferencias no son todas menores que *E*. Por lo tanto, se requiere, operando del mismo modo, otra iteración que produce los siguientes valores para las incógnitas:

$$x_1 = 2{,}999941296$$
$$x_2 = 3{,}999794400$$
$$x_3 = 5{,}000396320$$

Si ahora se realiza la comparación de estos valores con los respectivos de la iteración anterior, se puede apreciar que sus diferencias son todas menores que *E*, lo que indica que el proceso ha terminado.

En este simple ejemplo se puede encontrar que la solución exacta del problema es $x_1 = 3$; $x_2 = 4$ y $x_3 = 5$.

♦♦♦ ●●● ♦♦♦

Se ha mencionado que la precisión no depende exclusivamente del valor de *E*, sino también , de la velocidad de convergencia; es decir, que las diferencias de los resultados obtenidos en dos iteraciones consecutivas pueden diferir en menos de un *E*, aún cuando estos resultados se encuentren lejos de los verdaderos valores de las incógnitas. En consecuencia, para obtener mayor seguridad y precisión en la solución se recomienda utilizar las ecuaciones de error para concluir la aproximación de las incógnitas.

6.3.- SISTEMAS DE ECUACIONES HOMOGÉNEAS

Se considerará en este punto la solución de ecuaciones algebraicas lineales simultáneas **HOMOGÉNEAS**, las cuales tienen la forma:

$$\begin{cases} a_{11}\,x_1 + a_{12}\,x_2 + \ldots + a_{1m}\,x_m = 0 \\ a_{21}\,x_1 + a_{22}\,x_2 + \ldots + a_{2m}\,x_m = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}\,x_1 + a_{n2}\,x_2 + \ldots + a_{nm}\,x_m = 0 \end{cases} \tag{6.27}$$

el que puede ser escrito en notación matricial de la siguiente manera:

$$\mathbf{A} \,.\, \mathbf{X} = \mathbf{O}$$

Dada las especiales condiciones que revisten estos sistemas y sus muy particulares e importantes aplicaciones, serán estudiados de manera puntual, tal como se hará a continuación.

6.3.1.- GENERALIDADES

En álgebra lineal se demuestra que cualquier sistema de n ecuaciones algebraicas lineales con m incógnitas, tiene solución si, y solo si, la matriz de los coeficientes y la matriz orlada o ampliada tienen el mismo rango, como ya se ha estudiado en la introducción del presente capítulo; por lo tanto, un conjunto de ecuaciones homogéneas como el (4.27) tiene siempre una solución, ya que la matriz ampliada y la del sistema tienen, necesariamente, el mismo rango.

Si el rango r de la matriz de los coeficientes del conjunto de ecuaciones es igual al orden n, el sistema tiene una única solución, que es la denominada *SOLUCIÓN TRIVIAL* ($x_1 = x_2 = \ldots = x_n = 0$). Para este conjunto de ecuaciones, el determinante de la matriz **A** del sistema es diferente de cero; o sea que, todas las ecuaciones del sistema son linealmente independientes.

Existen soluciones *NO TRIVIALES* para un sistema si, y solo si, el rango r de la matriz **A** es menor que el orden n. Para ese sistema, el determinante de la matriz de los coeficientes debe ser igual a cero, y el sistema consta de r ecuaciones linealmente independientes y $n - r$ ecuaciones dependientes que son combinaciones lineales de las r ecuaciones independientes.

Al obtener soluciones no triviales de un sistema de ecuaciones homogéneo, no se encuentran valores únicos para las incógnitas. En su lugar, se establecen relaciones entre las incógnitas. Cualquier combinación de valores de x_i que satisface estas relaciones constituye una solución.

Los problemas más importantes que se plantean en la aplicación de ecuaciones homogéneas son aquellos denominados de **VALORES CARACTERÍSTICOS** que se presentan en diversos campos de la ciencia y de la técnica. Cuando se trata de desarrollar modelos matemáticos, las ecuaciones pueden escribirse, generalmente, de manera cartesiana, en la forma:

$$\begin{cases} (a_{11} - \lambda)\,x_1 + a_{12}\,x_2 + \ldots + a_{1n}\,x_n = 0 \\ a_{21}\,x_1 + (a_{22} - \lambda)\,x_2 + \ldots + a_{2n}\,x_n = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}\,x_1 + a_{n2}\,x_{n2} + \ldots + (a_{nn} - \lambda)\,x_n = 0 \end{cases} \tag{6.28}$$

donde, los a_{ij} son reales, las x_i son las variables del sistema y lambda es un parámetro particular del sistema que tiene valores, en general, desconocidos.

En la notación matricial, las ecuaciones (6.28) pueden ser expresadas por medio de:

$$(\mathbf{A} - \lambda\,\mathbf{I})\,\mathbf{X} = \mathbf{0} \tag{6.29}$$

en donde el uso de la matriz identidad **I**, permite utilizar a (**A** - λ **I**) como matriz de coeficientes. La matriz columna **X** recibe el nombre de **VECTOR CARACTERÍSTICO ó AUTOVECTOR ó EIGENVECTOR**; siendo las x_i las componentes de dicho vector característico. De la misma manera, los valores que se obtienen para el parámetro lambda se conocen como **VALORES CARACTERÍSTICOS ó AUTOVALORES ó EIGENVALORES** de la matriz **A**.

Considerando las ecuaciones (6.28), como el modelo matemático de un sistema físico, resulta obvio que la solución trivial no tiene ningún significado real y que el verdadero interés se encuentra focalizado sobre las soluciones no triviales. Como se indicó, existen soluciones no triviales si el determinante de la matriz de coeficientes es igual a cero.

Dado que lambda aparece como incógnita en la matriz de coeficientes, es posible, arbitrariamente, hacer que el determinante de dicha matriz sea igual a cero y, llamando **D** a éste, resulta:

$$\mathbf{D} = [\,\mathbf{A} - \lambda\,\mathbf{I}\,] = 0 \tag{6.30}$$

y encontrando valores para lambda que hagan $\mathbf{D} = 0$, se tendrá la solución.

El desarrollo algebraico del determinante **D** produce un polinomio de grado *n*, de la forma:

$$\lambda^n + p_1\lambda^{n-1} + \ldots + p_{n-1}\lambda + p_n = 0 \tag{6.31}$$

el que será necesario resolver para tener los valores de lambda que hacen $\mathbf{D} = 0$. Este polinomio recibe el nombre de **POLINOMIO CARACTERÍSTICO** ó **ECUACIÓN CARACTERÍSTICA** de la matriz **A** y las *n* raíces del polinomio se conocen, como ya se ha mencionado, con la denominación de **AUTOVALORES** o **VALORES CARACTERÍSTICOS**.

Después de determinar los valores característicos, se pueden sustituir, uno a la vez, en el conjunto original de ecuaciones para obtener un sistema correspondiente de relaciones entre las incógnitas *x* para cada sustitución. Las relaciones así obtenidas dependerán del rango *r* de la matriz (**A** - λ **I**). Si es $r = n - 1$, las relaciones serán tales que la hipótesis de un valor para una incógnita, producirá un valor correspondiente para cada una de las ecuaciones restantes; si $r = n - 2$, las relaciones serán tales que se tendrán que suponer los valores de dos incógnitas para poder obtener un valor correspondiente a cada una de las incógnitas restantes.

Cuando *n* es relativamente pequeño (2 ó 3), el desarrollo de determinantes por menores para obtener el polinomio característico no es difícil; tampoco la posterior determinación de raíces ofrece grandes dificultades. Sin embargo, cuando *n* es grande ésta determinación se vuelve muy difícil, sino imposible y se debe recurrir entonces, a algún procedimiento numérico tal como alguno de los estudiados en el capítulo correspondiente.

Seguidamente se estudiarán métodos numéricos especiales destinados a ser aplicados en la resolución de sistemas de ecuaciones lineales simultáneas.

6.3.2.- MÉTODO DE FADDEEV-LEVERRIER

Este método, constituye una técnica muy eficiente para generar los coeficientes p_i del polinomio característico dado por (6.31), de la ecuación (6.29), tanto, cuando la matriz **A** de los coeficientes es simétrica, como cuando no lo es. Tiene además, la ventaja adicional de que se obtiene automáticamente, al finalizar el proceso de generar los coeficientes del polinomio característico, la matriz inversa del sistema $\mathbf{A}^{-1}$.

Antes de describir el método, recuérdese el concepto de **TRAZA** de una matriz, ya que es una parte integrante del mismo. La traza de una matriz, que será designada mediante " *tr* **A** " está dada por la siguiente:

Definición: *" En toda matriz cuadrada* **A**, *se define como* **TRAZA** *de la misma, a la suma de todos los elementos que se encuentran sobre la diagonal principal. "*

dicho analíticamente, se tiene la siguiente:

$$tr\,\mathbf{A} = a_{11} + a_{22} + \ldots + a_{nn} = \sum a_{ij} \quad ; \quad \forall\, i = j \tag{6.32}$$

El método de *FADDEEV-LEVERRIER* genera en su procedimiento, una sucesión de matrices $\mathbf{B_1}$; $\mathbf{B_2}$; ...; $\mathbf{B_n}$ de las que se determinan una serie de valores, correspondientes a aquellas, y que se denominarán b_k (k = 1; 2;...;n) que sustituidos en el **POLINOMIO DE FADDEEV-LEVERRIER**:

$$(-1)^n\left(\lambda^n - b_1\lambda^{n-1} - \ldots - b_{n-1}\lambda - b_n\right) = 0 \tag{6.33}$$

dan como resultado los coeficientes p_k del polinomio característico dado por la expresión (6.31). El factor $(-1)^n$ se utiliza solamente para dar a los términos del polinomio de *FADDEEV-LEVERRIER* los signos que tendría el polinomio característico si hubiese sido generado desarrollando el determinante correspondiente, a partir de la expresión (6.31). Los valores de b_k se obtienen como se muestra a continuación:

$$\begin{aligned} &\mathbf{B}_1 = \mathbf{A} && ; && b_1 = tr\,\mathbf{B}_1 \\ &\mathbf{B}_2 = \mathbf{A}\,(\mathbf{B}_1 - b_1\,\mathbf{I}) && ; && b_2 = {}^1\!/_2\; tr\,\mathbf{B}_2 \\ &\cdots\cdots\cdots\cdots\cdots\cdots && && \\ &\mathbf{B}_n = \mathbf{A}\,(\mathbf{B}_{n-1} - b_{n-1}\,\mathbf{I}) && ; && b_n = {}^1\!/_n\; tr\,\mathbf{B}_n \end{aligned} \tag{6.34}$$

Puede demostrarse que la inversa de la matriz **A**, cuyos valores característicos se desea hallar, se determina a partir de la ecuación:

$$\mathbf{A}^{-1} = {}^1\!/b_n\,(\mathbf{B}_{n-1} - b_{n-1}\,\mathbf{I}) \tag{6.35}$$

La aplicación del método de *FADDEEV-LEVERRIER* se puede apreciar a través del siguiente:

Ejemplo 5.3.- Supóngase que se ha obtenido, como un modelo matemático de algún problema determinado, el siguiente sistema lineal homogéneo de ecuaciones:

$$\begin{cases} (3-\lambda)x_1 + 2x_2 + 4x_3 = 0 \\ 2x_1 + (0-\lambda)x_2 + 2x_3 = 0 \\ 4x_1 + 2x_2 + (3-\lambda)x_3 = 0 \end{cases} \tag{6.36}$$

Solución: El problema consiste en determinar el polinomio característico utilizando el método de *FADDEEV-LEVERRIER*, anteriormente descripto. Para ello se deben generar las matrices $\mathbf{B_k}$ a partir de la matriz **A** del sistema. Siendo:

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

entonces:

$$\mathbf{B}_1 = \mathbf{A} = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix} \Rightarrow \quad b_1 = tr\,\mathbf{B}_1 = 3 + 0 + 3 = 6$$

de la misma manera:

$$\mathbf{B}_2 = \mathbf{A}\,(\mathbf{B}_1 - b_1\,\mathbf{I}) = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix} \left\{ \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix} - \begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix} \right\} = \begin{bmatrix} 11 & 2 & 4 \\ 2 & 8 & 2 \\ 4 & 2 & 11 \end{bmatrix} \Rightarrow$$

$$\Rightarrow \quad b_2 = {}^1\!/_2\; tr\,\mathbf{B}_2 = {}^1\!/_2\,(11 + 8 + 11) = 15$$

$$\mathbf{B}_3 = \mathbf{A}\,(\mathbf{B}_2 - b_2\,\mathbf{I}) = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix} \left\{ \begin{bmatrix} 11 & 2 & 4 \\ 2 & 8 & 2 \\ 4 & 2 & 11 \end{bmatrix} - \begin{bmatrix} 15 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 15 \end{bmatrix} \right\} = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{bmatrix} \Rightarrow$$

$$\Rightarrow \quad b_3 = {}^1\!/_3\; tr\,\mathbf{B}_3 = {}^1\!/_3\,(8 + 8 + 8) = 8$$

Sustituyendo los valores de los b_k en el polinomio de *FADDEEV-LEVERRIER*, se obtiene:

$$(-1)^3\left(\lambda^3 - 6\lambda^2 - 15\lambda - 8\right) = 0$$

del cual resulta:

$$-\lambda^3 + 6\lambda^2 + 15\lambda + 8 = 0$$

ó, multiplicando por -1:

$$\lambda^3 - 6\lambda^2 - 15\lambda - 8 = 0 \tag{6.37}$$

que es el polinomio característico del sistema (6.36). Obteniendo las raíces de la ecuación dada por (6.37) mediante cualquier método, resultan los valores propios:

$$\lambda_1 = 8 \quad ; \quad \lambda_2 = -1 \quad ; \quad \lambda_3 = -1$$

los que, sustituidos de a uno por vez en el sistema (6.36), permite el cálculo de las componentes x_j (j = 1; 2; 3) de los vectores característicos o autovectores.

Finalmente, para hallar la matriz inversa del sistema $\mathbf{A}^{-1}$, según la expresión (6.35), se debe hacer:

$$\mathbf{A}^{-1} = \frac{1}{8}\left\{\begin{bmatrix} 11 & 2 & 4 \\ 2 & 8 & 2 \\ 4 & 2 & 11 \end{bmatrix} - 15\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right\} = \begin{bmatrix} -\frac{4}{8} & \frac{2}{8} & \frac{4}{8} \\ \frac{2}{8} & -\frac{7}{8} & \frac{2}{8} \\ \frac{4}{8} & \frac{2}{8} & -\frac{4}{8} \end{bmatrix}$$

6.3.3.- MÉTODO DE LAS POTENCIAS

El método de las **POTENCIAS** es un método iterativo que, frecuentemente es utilizado en aquellos casos en que solamente se desea conocer el autovalor más pequeño y/o más grande, juntamente con sus autovectores asociados, de un sistema dado. Una ventaja adicional de este método es que los autovalores se obtienen simultáneamente con sus respectivos autovectores.

Para determinar el autovalor más grande, supóngase que tanto los elementos de la matriz como del autovalor son reales, aunque, en la práctica no tiene esta restricción. Sea el sistema:

(6.38) $$(\mathbf{A} - \lambda\, \mathbf{I})\, \mathbf{X} = \mathbf{0}$$

donde, realizando el producto indicado por el paréntesis, se obtiene:

$$\mathbf{A}\,\mathbf{X} - \lambda\, \mathbf{I}\,\mathbf{X} = \mathbf{A}\,\mathbf{X} - \lambda\, \mathbf{X} = \mathbf{0}$$

y transponiendo términos:

(6.39) $$\mathbf{A}\,\mathbf{X} = \lambda\, \mathbf{X}$$

Haciendo uso sistemático de esta última ecuación (6.39) y realizando los siguientes pasos:

I.- *Asignar valores arbitrarios a las componentes de* **X**, *y designándolo con* $\mathbf{X_0}$ *se sustituye en el primer miembro de (6.39), con lo cual se obtiene la primera aproximación del segundo miembro.*

$$\mathbf{A}\,\mathbf{X}_0 = \lambda\, \mathbf{X}_1$$

En general, resulta satisfactorio tomar los valores para $x_i = 1; 2; \ldots ;n.$

II.- *Dividir los elementos del vector* $\lambda\, \mathbf{X}_1$ *por* $\lambda\, x_1$, *para que la primera componente se reduzca a la unidad.*

III.- *Se utilizan las componentes del vector obtenido como valores mejorados de* **X**, *sustituyéndolos en el primer miembro de (6.39) para volver a obtener así, una mejor aproximación en un siguiente paso.*

IV.- *Se repiten los pasos* **II** *y* **III** *hasta que la expresión (6.39) quede esencialmente satisfecha; es decir, hasta que las componentes del autovector obtenido varíen con respecto a los del paso anterior en una cantidad menor que un cierto* **E** *positivo y arbitrario, fijado de antemano.*

Al efectuar el proceso de iteración descripto, se estaría formando una sucesión $\mathbf{A}\,\mathbf{X}_0 ; \mathbf{A}^2\,\mathbf{X}_0 ; \ldots ; \mathbf{A}^k\,\mathbf{X}_0$, en donde $\mathbf{X}_0$ es el vector arbitrario supuesto inicialmente. Las potencias de la matriz **A** que componen la sucesión, son las que le dan el nombre al método.

Se ilustrará prácticamente el método de las potencias desarrollando el ejercicio siguiente:

Ejemplo 6.4.- Considérese el sistema de ecuaciones lineales simultáneas homogéneas simultáneas siguiente:

(6.40) $$\begin{cases} (2-\lambda)x_1 + 4x_2 = 0 \\ 3x_1 + (13-\lambda)x_2 = 0 \end{cases}$$

de las que se quiere determinar el autovalor más grande y su autovector asociado.

Solución: Primeramente se deberá escribir el sistema dado (6.40) en la forma básica, mostrada en la (6.39):

$$\begin{bmatrix} 2 & 4 \\ 3 & 13 \end{bmatrix}\begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix} = \lambda \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix}$$

y realizando los pasos **I** y **II**, resulta:

$$\begin{bmatrix} 2 & 4 \\ 3 & 13 \end{bmatrix} \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} = 1 \begin{Bmatrix} 6 \\ 16 \end{Bmatrix} = 6 \begin{Bmatrix} 1 \\ \frac{16}{6} \end{Bmatrix}$$

iterando el procedimiento anterior, resulta:

$$\begin{bmatrix} 2 & 4 \\ 3 & 13 \end{bmatrix} \begin{Bmatrix} 1 \\ \frac{16}{6} \end{Bmatrix} = \begin{Bmatrix} \frac{38}{3} \\ \frac{113}{3} \end{Bmatrix} = \frac{38}{3} \begin{Bmatrix} 1 \\ \frac{113}{38} \end{Bmatrix}$$

finalmente, el siguiente paso arroja:

$$\begin{bmatrix} 2 & 4 \\ 3 & 13 \end{bmatrix} \begin{Bmatrix} 1 \\ \frac{113}{38} \end{Bmatrix} = \begin{Bmatrix} \frac{264}{19} \\ \frac{1538}{38} \end{Bmatrix} = \frac{264}{19} \begin{Bmatrix} 1 \\ \frac{769}{264} \end{Bmatrix}$$

donde $\lambda = {}^{264}/_{19} = 13{,}89$; $\mathbf{X} = (\ 1\ ;\ {}^{769}/_{264}\ ) = (\ 1\ ;\ 2{,}91\ )$. Por supuesto, una cuata iteración producirá valores más exactos. Los verdaderos son $\lambda = 14$; $x_1 = 1$; $x_2 = 3$.

••• ●●● •••

A continuación se estudiará la manera de emplear el método de las potencias para determinar el autovalor más pequeño y su autovector asociado. Para ello es necesario premultiplicar (6.39) por la inversa de **A**, resultando:

$$\mathbf{A}^{-1}\ \mathbf{A}\ \mathbf{X} = \mathbf{A}^{-1}\ \lambda\ \mathbf{X}$$

o bien:

(6.41) $$\mathbf{X} = \mathbf{A}^{-1}\ \lambda\ \mathbf{X}$$

Dividiendo esta última expresión por el valor de λ, y permutando ambos miembros, resulta:

(6.42) $$\mathbf{A}^{-1}\ \mathbf{X} = \frac{1}{\lambda}\ \mathbf{X}$$

la que producirá una convergencia al valor más pequeño del autovalor λ. Sea el siguiente:

Ejemplo 6.5.- Sea el sistema representado analíticamente por medio de la expresión (6.40) del ejemplo anterior.

Solución: Sabiendo que la matriz inversa es:

$$\mathbf{A}^{-1} = \begin{bmatrix} \frac{13}{14} & -\frac{2}{7} \\ -\frac{3}{14} & \frac{1}{7} \end{bmatrix}$$

reiterando los pasos propuestos, en base a la expresión (6.42), resulta:

$$\mathbf{A}^{-1}\ \mathbf{X} = \frac{1}{\lambda}\mathbf{X} = \begin{bmatrix} \frac{13}{14} & -\frac{2}{7} \\ -\frac{3}{14} & \frac{1}{7} \end{bmatrix} \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} = \begin{Bmatrix} \frac{9}{14} \\ -\frac{1}{14} \end{Bmatrix} = \frac{9}{14} \begin{Bmatrix} 1 \\ -\frac{1}{9} \end{Bmatrix} =$$

$$= \begin{bmatrix} \frac{13}{14} & -\frac{2}{7} \\ -\frac{3}{14} & \frac{1}{7} \end{bmatrix} \begin{Bmatrix} 1 \\ -\frac{1}{9} \end{Bmatrix} = \begin{Bmatrix} \frac{121}{126} \\ -\frac{29}{126} \end{Bmatrix} = \frac{121}{126} \begin{Bmatrix} 1 \\ -\frac{29}{121} \end{Bmatrix} =$$

$$= \begin{bmatrix} \frac{13}{14} & -\frac{2}{7} \\ -\frac{3}{14} & \frac{1}{7} \end{bmatrix} \begin{Bmatrix} 1 \\ -\frac{29}{121} \end{Bmatrix} = \begin{Bmatrix} \frac{1689}{1694} \\ -\frac{421}{1694} \end{Bmatrix} = \frac{1689}{1694} \begin{Bmatrix} 1 \\ -\frac{421}{1689} \end{Bmatrix}$$

Por lo tanto, el valor característico más pequeño es $\lambda = 1{,}0029$ y el vector característico asociado $\mathbf{X} = (1\ ;\ -0{,}2492)$. Si se siguiera iterando el proceso, se llegaría a los valores exactos, que son $\lambda = 1$; $\mathbf{X} = (1\ ;\ -0{,}25)$.

••• ●●● •••

Si los autovalores máximo e inmediatamente menor tienen valores similares; o, si otro tanto ocurre con el mínimo e inmediatamente mayor, la convergencia puede resultar lenta. Si el autovalor máximo tiene multiplicidad dos, la convergencia es hacia el autovalor máximo, pero, las componentes del autovector convergen a cualquiera de los dos que están asociados a aquel; dependiendo esto, del vector **X** supuesto inicialmente.

En caso que el vector supuesto inicialmente sea ortogonal con el autovector asociado al autovalor máximo de la matriz traspuesta, convergirá al autovalor que le sigue en magnitud en lugar de hacerlo al máximo.

Dado que se trata de un tema de la mayor importancia en la ingeniería, es necesario hacer algunas consideraciones; se ha estudiado que después de determinar los valores característicos de una matriz **A**, se pueden sustituir, uno por vez, en el sistema original homogéneo, para obtener un nuevo sistema de n ecuaciones homogéneas con n incógnitas (componentes del autovector) para cada autovalor.

Como los autovalores son cantidades que hacen que el determinante de la matriz de coeficientes $(\mathbf{A} - \lambda \mathbf{I})$ sea cero, el rango r de la matriz de los coeficientes debe ser necesariamente menor que el orden n de la matriz (no todas las ecuaciones obtenidas son independientes). Por tanto, aunque se tengan n ecuaciones con n incógnitas, el problema realmente consiste en resolver r ecuaciones con n incógnitas, ya que el número de ecuaciones linealmente independientes es igual al rango de la matriz de coeficientes.

Si el rango de la matriz de coeficientes es una unidad menor que el orden, el sistema de ecuaciones homogéneas que se va a resolver contiene n-1 ecuaciones independientes para determinar los n componentes de los autovectores. Por tanto, suponiendo un valor para una componente de un autovector, se pueden obtener valores para las componentes restantes. Habrá, pues, un autovector asociado con cada autovalor.

En ocasiones, no todos los autovalores obtenidos de la ecuación característica tendrán multiplicidad uno, como ya se ha visto en el ejemplo del punto anterior. En este caso, cuando un autovalor obtenido tiene una multiplicidad k, el rango de la matriz de los coeficientes para ese autovalor puede ser menor que el orden de la matriz con una diferencia de 1 a k unidades, y se tendrá el mismo número de autovectores linealmente independientes asociados al autovalor que tiene esa multiplicidad.

Si el rango de la matriz de coeficientes es dos unidades menor que el orden, es necesario suponer valores para dos componentes del autovector con el objeto de determinar valores para las componentes restantes. Habrá solamente n-2 ecuaciones independientes en el sistema. Se pueden obtener dos autovectores independientes y cualquier combinación lineal de esos dos autovectores constituye una posible solución. En este caso se dice que el espacio de soluciones es bidimensional.

Si n-r = 3, se deben suponer tres componentes del autovector, obteniéndose tres vectores característicos independientes, y el espacio de soluciones será tridimensional. El razonamiento es idéntico cuando el grado de indeterminación es mayor.

6.4.- ECUACIONES LINEALES SIMULTÁNEAS ORTOGONALES

En esta parte del tema se estudiarán métodos especiales destinados específicamente a la resolución de **ECUACIONES LINEALES SIMULTÁNEAS ORTOGONALES**. La importancia del tema radica en el hecho de la facilidad y rapidez con que estos sistemas pueden ser resueltos, y además, de la posibilidad de convertir un sistema cualquiera en ortogonal, si se lo desea.

6.4.1.- INTRODUCCIÓN

Un sistema de ecuaciones lineales simultáneas de la forma:

$$\mathbf{A}\,\mathbf{X} = \mathbf{C} \quad ; \quad \mathbf{A} = [\,a_{ij}\,] \tag{6.43}$$

en la que, **A** es una matriz cuadrada de orden $n \times n$, que recibe la denominación de **ORTOGONAL** si satisface la condición:

$$\sum_{r=1}^{n} a_{ij}\, a_{jr} = 0 \quad ; \quad \forall\, j \neq i \tag{6.44}$$

es decir, si la suma de los productos de coeficientes homólogos de dos filas diferentes del sistema es nulo, para todo par de filas. La expresión (6.44) recibe el nombre de **CONDICIÓN DE ORTOGONALIDAD**.

Considerando los n coeficientes de cada fila de la matriz **A** como componentes de un vector $\mathbf{A}_i$, la condición de ortogonalidad de estos vectores es precisamente la ecuación (6.44). En general es:

$$\left(\mathbf{A}_i\,;\mathbf{A}_j\right) = \sum_{r} a_{ir}\, a_{jr} = 0 \quad ; \quad \forall\, i \neq j \tag{6.45}$$

lo que expresa que, si el *PRODUCTO ESCALAR* de dos vectores $\mathbf{A}_i$; $\mathbf{A}_j$ es nulo, estos son ortogonales.

6.4.2.- EL MÉTODO ANALÍTICO

El sistema representado en la (6.43), puede ser escrito en forma cartesiana:

$$\text{(6.46)} \qquad \begin{cases} a_{11}\,x_1 + a_{12}\,x_2 + \cdots + a_{1n}\,x_n = c_1 \\ a_{21}\,x_1 + a_{22}\,x_2 + \cdots + a_{2n}\,x_n = c_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}\,x_1 + a_{n2}\,x_2 + \cdots + a_{nn}\,x_n = c_n \end{cases}$$

Se verá que la solución de este sistema, si satisface la condición (6.44), puede encontrarse fácilmente evitando la aplicación de métodos extensos y complejos. Para ello deben ser transformadas las incógnitas $x_1 ; x_2 ;...; x_n$, en otras $z_1 ; z_2 ;...; z_n$, mediante la relación:

$$\mathbf{X} = \mathbf{A}^{\mathbf{T}}\ \mathbf{Z}$$

que, en forma cartesiana se expresa:

$$\text{(6.47)} \qquad \begin{cases} x_1 = a_{11}\,z_1 + a_{21}\,z_2 + \cdots + a_{n1}\,z_n \\ x_2 = a_{12}\,z_1 + a_{22}\,z_2 + \cdots + a_{n2}\,z_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n = a_{1n}\,z_1 + a_{2n}\,z_2 + \cdots + a_{nn}\,z_n \end{cases}$$

Estos valores de x_i reemplazados en la expresión (6.46), dan como resultado:

$$\begin{cases} a_{11}\left(a_{11}\,z_1 + a_{21}\,z_2 + \cdots + a_{n1}\,z_n\right) + a_{12}\left(a_{12}\,z_1 + a_{22}\,z_2 + \cdots + a_{n2}\,z_n\right) + \cdots + \\ \qquad + a_{1n}\left(a_{1n}\,z_1 + a_{2n}\,z_2 + \cdots + a_{nn}\,z_n\right) = c_1 \\ a_{21}\left(a_{11}\,z_1 + a_{21}\,z_2 + \cdots + a_{n1}\,z_n\right) + a_{22}\left(a_{12}\,z_1 + a_{22}\,z_2 + \cdots + a_{n2}\,z_n\right) + \cdots + \\ \qquad + a_{2n}\left(a_{1n}\,z_1 + a_{2n}\,z_2 + \cdots + a_{nn}\,z_n\right) = c_2 \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\ a_{n1}\left(a_{11}\,z_1 + a_{21}\,z_2 + \cdots + a_{n1}\,z_n\right) + a_{n22}\left(a_{12}\,z_1 + a_{22}\,z_2 + \cdots + a_{n2}\,z_n\right) + \cdots + \\ \qquad + a_{nn}\left(a_{1n}\,z_1 + a_{2n}\,z_2 + \cdots + a_{nn}\,z_n\right) = c_n \end{cases}$$

Estas últimas expresiones se pueden ordenar en función de las incógnitas definidas, resultando el nuevo ordenamiento consignado en la expresión:

$$\begin{cases} \left(a_{11}^2 + a_{12}^2 + \ldots + a_{1n}^2\right) z_1 + \left(a_{11}\,a_{21} + a_{12}\,a_{22} + \ldots + a_{1n}\,a_{2n}\right) z_2 + \ldots + \\ \qquad + \left(a_{11}\,a_{n1} + a_{12}\,a_{n2} + \ldots + a_{1n}\,a_{nn}\right) z_n = c_1 \\ \left(a_{21}\,a_{11} + a_{22}\,a_{12} + \ldots + a_{2n}\,a_{1n}\right) z_1 + \left(a_{21}^2 + a_{22}^2 + \ldots + a_{2n}^2\right) z_2 + \ldots + \\ \qquad + \left(a_{21}\,a_{n1} + a_{22}\,a_{n2} + \ldots + a_{2n}\,a_{nn}\right) z_n = c_2 \\ \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\ \left(a_{n1}\,a_{11} + a_{n2}\,a_{12} + \ldots + a_{nn}\,a_{1n}\right) z_1 + \left(a_{n1}\,a_{21} + a_{n2}\,a_{22} + \ldots + a_{nn}\,a_{2n}\right) z_2 + \ldots + \\ \qquad + \left(a_{n1}^2 + a_{n2}^2 + \ldots + a_{nn}^2\right) z_n = c_n \end{cases}$$

Considerando la condición de ortogonalidad dada por (6.44) es posible apreciar que, los paréntesis que encierran sumas de productos se anulan, quedando solamente aquellos que encierran sumas de cuadrados, así:

$$\text{(6.48)} \qquad \begin{cases} \left(a_{11}^2 + a_{12}^2 + \ldots + a_{1n}^2\right) z_1 = c_1 \\ \left(a_{21}^2 + a_{22}^2 + \ldots + a_{2n}^2\right) z_2 = c_2 \\ \ldots \quad \ldots \quad \ldots \quad \ldots \\ \left(a_{n1}^2 + a_{n2}^2 + \ldots + a_{nn}^2\right) z_n = c_n \end{cases}$$

Mediante estas últimas expresiones es posible obtener, de manera directa, todos los valores de las z_i luego, haciendo uso de las ecuaciones (4.47), se calculan los valores de las x_i.

En el caso que la precisión alcanzada no fuera satisfactoria, los resultados obtenidos son susceptibles de ser refinados mediante el método de las ecuaciones de error, estudiado en su momento.

Ejemplo 6.6.- Dado el siguiente sistema, verificar su condición de ortogonalidad y resolver apropiadamente:

$$\begin{cases} x + \ y + 2w = -11 \\ 7x - \ y - 4w = \ \ 1 \\ 4x + 12y + 4w = -100 \end{cases}$$

Solución: Para verificar la condición de ortogonalidad, se deben realizar los productos escalares entre ecuaciones, resultando de esta manera:

$$(\mathbf{A}_1 ; \mathbf{A}_2) = 1 \,.\, 7 + (-1) \,.\, (-1) + 2 \,.\, (-4) = 7 + 1 - 8 = 0$$
$$(\mathbf{A}_1 ; \mathbf{A}_3) = 1 \,.\, 4 + (-1) \,.\, 12 + 2 \,.\, 4 = 4 - 12 + 8 = 0$$
$$(\mathbf{A}_2 ; \mathbf{A}_3) = 7 \,.\, 4 + (-1) \,.\, 12 + (-4) \,.\, 4 = 28 - 12 - 16 = 0$$

Aplicando al sistema dado, la fórmula resolvente de la expresión (6.48), se obtiene:

$$6\, z_1 = -11 \quad ; \quad 66\, z_2 = 1 \quad ; \quad 176\, z_3 = -100$$

vale decir:

$$z_1 = -{}^{11}/_{6} \quad ; \quad z_2 = {}^{1}/_{66} \quad ; \quad z_3 = -{}^{100}/_{176}$$

y finalmente, reemplazando estos valores en las expresiones (4.47), resulta:

$$x = -4$$
$$y = -5$$
$$w = -6$$

◆◆◆ ●●● ◆◆◆

6.4.3.- MÉTODO DE ORTOGONALIZACIÓN DE GRAM-SCHMIDT

En múltiples ocasiones puede resultar conveniente o, aún más, necesario por razones tecnológicas, convertir un sistema de ecuaciones lineales simultáneas de la forma (6.46), cualquiera, en otro equivalente (con idénticas raíces), pero, ortogonal según la definición estudiada.

En virtud de la longitud y complejidad del desarrollo teórico del tema en forma generalizada, se lo expondrá de manera concreta en un espacio de tres dimensiones. En consecuencia, para el presente caso, sea un sistema de ecuaciones lineales simultáneas de la forma:

$$\text{(6.49)} \qquad \begin{cases} a_{11}\, x_1 + a_{12}\, x_2 + a_{13}\, x_3 = c_1 \\ a_{21}\, x_1 + a_{22}\, x_2 + a_{23}\, x_3 = c_2 \\ a_{31}\, x_1 + a_{32}\, x_2 + a_{33}\, x_3 = c_3 \end{cases}$$

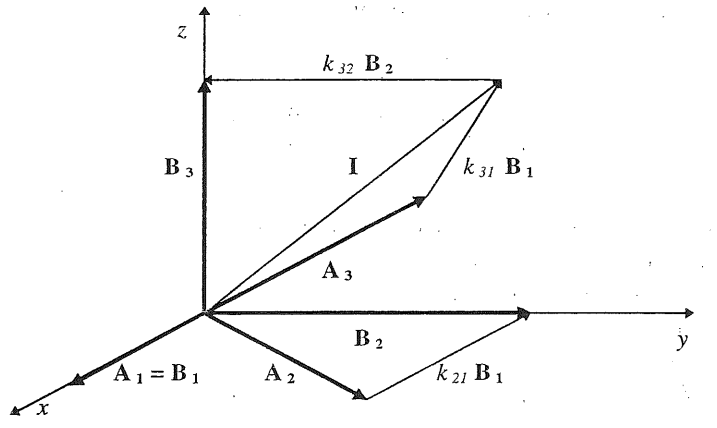
Este objetivo se puede conseguir aplicando el método de ortogonalización de **GRAM-SCHMIDT** a sistemas, cuya única condición exigida es la de que el determinante de la matriz **A** del sistema sea distinto de cero; vale decir:

$$\text{(6.50)} \qquad |\mathbf{A}| \neq 0$$

Designando mediante la notación $\mathbf{A}_1$; $\mathbf{A}_2$; $\mathbf{A}_3$ los vectores correspondientes al sistema (6.49), resulta:

$$\mathbf{A}_1 (a_{11} \,;\, a_{12} \,;\, a_{13}) \quad ; \quad \mathbf{A}_2 (a_{21} \,;\, a_{22} \,;\, a_{23}) \quad ; \quad \mathbf{A}_3 (a_{31} \,;\, a_{32} \,;\, a_{33})$$

Se estudiará a continuación, la manera de operar para obtener la ortogonalización del sistema dado. Considérese la figura 6.1, en la cual se han referido los vectores $\mathbf{A}_1$; $\mathbf{A}_2$; $\mathbf{A}_3$ a un sistema de ejes coordenados cartesianos ortogonales.



- Figura 6.1 -

Se debe notar además, que se ha hecho coincidir el vector $\mathbf{A}_1$, con el eje de las x, llamándolo $\mathbf{B}_1$; de la misma manera, se ha hecho coincidir el vector $\mathbf{A}_2$, con el plano xy; todo lo cual es siempre posible, si los vectores no son alabeados: es decir, si cumplen la condición dada por (6.50).

Sumando vectorialmente el vector $k_{21}\,\mathbf{B}_1$ (paralelo al $\mathbf{B}_1$) al vector $\mathbf{A}_2$, se obtiene el nuevo vector $\mathbf{B}_2$, coincidente con el eje de las y; es decir, ortogonal al $\mathbf{A}_1 = \mathbf{B}_1$. De lo cual resulta que:

$$\mathbf{B}_2 = k_{21}\,\mathbf{B}_1 + \mathbf{A}_2$$

como puede apreciarse en la figura.

Para obtener el vector $\mathbf{B}_3$ se debe sumar vectorialmente al vector $\mathbf{A}_3$, que se encuentra en el espacio, otro vector paralelo al $\mathbf{B}_1$, que será designado con $k_{31}\,\mathbf{B}_1$ (paralelo al $\mathbf{A}_1$). El vector resultante es el designado con $\mathbf{I}$ en la figura (6.1). El valor de k_{31} debe ser tal que haga contener a $\mathbf{I}$ en el plano yz. Adicionando a este último un vector paralelo al $\mathbf{B}_2$, tal que lo haga coincidir con el eje z, el que será designado con $k_{32}\,\mathbf{B}_2$, se obtendrá el valor de $\mathbf{B}_3$.

En definitiva, utilizando ciertos valores particulares de los parámetros k_{21}; k_{31}; k_{32}, resulta el sistema equivalente al (6.49) y ortogonal:

$$\text{(6.51)} \qquad \begin{cases} \mathbf{B}_1 = \mathbf{A}_1 \\ \mathbf{B}_2 = \mathbf{A}_2 + k_{21}\,\mathbf{B}_1 \\ \mathbf{B}_3 = \mathbf{A}_3 + k_{31}\,\mathbf{B}_1 + k_{32}\,\mathbf{B}_2 \end{cases}$$

Se sabe que la condición de ortogonalidad estudiada en el punto anterior exige que el producto de coeficientes homólogos de dos filas diferentes del sistema, sea nulo, o su equivalente, que el producto escalar de los dos vectores representativos de esas dos filas, sea nulo; en cuyo caso debe ser:

$$\text{(6.52)} \qquad (\mathbf{B}_1 ; \mathbf{B}_2) = (\mathbf{B}_1 ; \mathbf{B}_3) = (\mathbf{B}_2 ; \mathbf{B}_3) = 0$$

entonces, desarrollando cada uno de los productos escalares indicados en (6.52) por sus valores obtenidos de (6.51), resulta:

$$(\mathbf{B}_1 ; \mathbf{B}_2) = (\mathbf{A}_1 ; \mathbf{A}_2) + k_{21}\,(\mathbf{A}_1 ; \mathbf{B}_1) = 0$$

de donde:

$$k_{21} = -\frac{(\mathbf{A}_1 ; \mathbf{A}_2)}{|\mathbf{A}_1|^2}$$

por consiguiente, resulta:

$$\text{(6.53)} \qquad \mathbf{B}_2 = \mathbf{A}_2 - \frac{(\mathbf{A}_1 ; \mathbf{A}_2)}{|\mathbf{A}_1|^2}\mathbf{A}_1$$

Dado que, en la expresión de $\mathbf{B}_3$ intervienen dos parámetros, será necesario hacer uso de las dos condiciones restantes para deducir su valor; vale decir:

$$(\mathbf{B}_1 ; \mathbf{B}_3) = (\mathbf{A}_1 ; \mathbf{A}_3) + k_{31}\,(\mathbf{A}_1 ; \mathbf{B}_1) + k_{32}\,(\mathbf{A}_1 ; \mathbf{B}_2) = 0$$

$$(\mathbf{B}_2 ; \mathbf{B}_3) = (\mathbf{A}_3 ; \mathbf{B}_2) + k_{31}\,(\mathbf{B}_1 ; \mathbf{B}_2) + k_{32}\,(\mathbf{B}_2 ; \mathbf{B}_2) = 0$$

Puesto que $\mathbf{A}_1 = \mathbf{B}_1$ y que, estos últimos son ortogonales a $\mathbf{B}_2$, es posible escribir:

$$k_{31}\,|\mathbf{A}_1|^2 + (\mathbf{A}_1 ; \mathbf{A}_3) = 0$$

$$k_{32}\,|\mathbf{B}_2|^2 + (\mathbf{A}_3 ; \mathbf{B}_2) = 0$$

de donde:

$$k_{31} = -\frac{(\mathbf{A}_1 ; \mathbf{A}_3)}{|\mathbf{A}_1|^2} \qquad ; \qquad k_{32} = -\frac{(\mathbf{A}_3 ; \mathbf{B}_2)}{|\mathbf{B}_2|^2}$$

y, finalmente:

$$\text{(6.54)} \qquad \mathbf{B}_3 = \mathbf{A}_3 - \frac{(\mathbf{A}_1 ; \mathbf{A}_3)}{|\mathbf{A}_1|^2}\mathbf{B}_1 - \frac{(\mathbf{A}_3 ; \mathbf{B}_2)}{|\mathbf{B}_2|^2}\mathbf{B}_2$$

Se estudiará la aplicación práctica del método de *GRAM-SCHMIDT* mediante el siguiente:

Ejemplo 6.7.- Se trata de aplicar el método de *GRAM-SCHMIDT* para resolver el problema que consiste en ortogonalizar el sistema:

$$\begin{cases} x - y + 2z = -11 \\ 2x - y + z = -9 \\ x - 2y + z = 0 \end{cases}$$

Solución: Según la metodología, se sabe que la primer ecuación (o vector) permanece invariante, pues, se hace coincidir con el eje de las x.

Para obtener la segunda ecuación se utiliza la expresión (6.53), donde:

$$(\mathbf{A}_1 ; \mathbf{A}_2) = 1 \,.\, 2 + (-1) \,.\, (-1) + 2 \,.\, 1 = 5$$

y, por otra parte

$$|\mathbf{A}_1|^2 = 1^2 + (-1)^2 + 2^2 = 6$$

por lo tanto

$$k_{21} = -{}^5/_6 \qquad ; \qquad \mathbf{B}_2 = \frac{7}{6}x - \frac{1}{6}y - \frac{4}{6}z = \frac{1}{6}$$

Utilizando la expresión (6.54) y considerando todos los elementos que intervienen:

$$(\mathbf{A}_1 ; \mathbf{A}_3) = 1 \,.\, 1 + (-1) \,.\, (-2) + 2 \,.\, 1 = 5$$

$$(\mathbf{A}_3 ; \mathbf{B}_2) = 1 \,.\, {}^7/_6 + (-2) \,.\, (-{}^1/_6) + 1 \,.\, (-{}^4/_6) = {}^5/_6$$

$$|\mathbf{B}_2|^2 = {}^7/_6 \,.\, {}^7/_6 + (-{}^1/_6) \,.\, (-{}^1/_6) + (-{}^4/_6) \,.\, (-{}^4/_6) = {}^{11}/_6$$

de donde:

$$k_{31} = -\frac{5}{6} \qquad ; \qquad k_{32} = -\frac{5}{11}$$

siendo:

$$\mathbf{B}_3 = -\frac{4}{11}x - \frac{12}{11}y - \frac{4}{11}z = \frac{100}{11}$$

En definitiva, eliminando denominadores y multiplicando la última por -1, resulta el siguiente sistema ortogonal:

$$\begin{cases} x - y + 2z = -11 \\ 7x - y - 4z = 1 \\ 4x + 12y + 4z = -100 \end{cases}$$

el que fuera resuelto como ejemplo en el punto anterior.

♦♦♦ ●●● ♦♦♦

Capítulo 7

Interpolación y Ajuste

7. - INTERPOLACIÓN Y AJUSTE

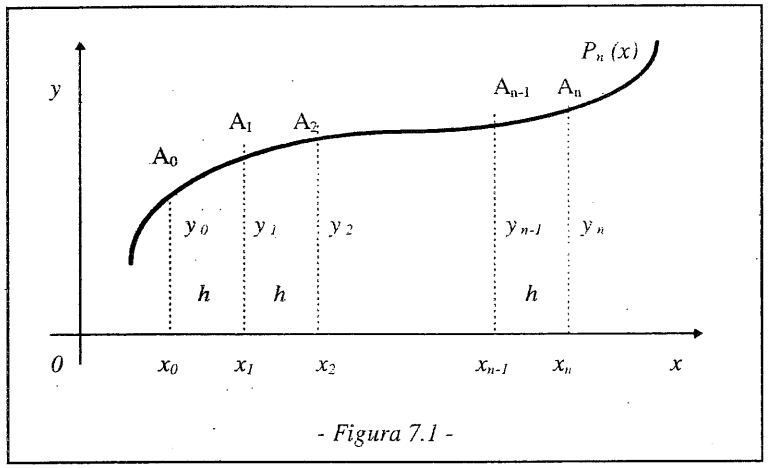
INTERPOLACIÓN. Tablas con Valores Equidistantes. Diferencias Avanzadas. Métodos de Interpolación. Fórmula de Interpolación de Newton-Gregory Ascendente. Interpolación Lineal y Cuadrática. Fórmula de Interpolación de Newton-Gregory Descendente. Diferencias Atrasadas. Fórmula de Lagrange. Tablas con Valores No Equidistantes. Fórmula de Lagrange para Valores No Equidistantes. Fórmula de Interpolación Parabólica Progresiva. Fórmula de Gauss. Fórmula de Bessel. Fórmula de Stirling. Fórmula de Everett. El Problema Inverso en la Interpolación.

AJUSTE. Ajuste e Interpolación. Ajuste Mediante Curvas. Ajuste Mediante Polinomios. Metodología del Ajuste. Mecánica del Cálculo.

7.1.- INTERPOLACIÓN

Siempre que en una tabla se busca el valor de una función para otro valor determinado de la variable que no figura explícitamente en esa tabla, se ejecuta una tarea de interpolación por medio de algunas reglas simples y muy precisas que sirven para resolver un problema aún más general; en otras palabras, de un mayor grado de generalización, como se verá más adelante.

Supóngase que se conocen los valores $y_0 ; y_1 ;...; y_n$ de la función desconocida, correspondientes a los $n+1$ valores distintos: $x_0 ; x_1 ; ... ; x_n$ de alguna variable independiente x. El problema, en este caso consiste, en determinar el valor aproximado de y que cabe asignar como correspondiente a otro valor de x, distinto de todos los x_i conocidos y comprendidos en el intervalo de trabajo $[x_0 ; x_n]$.



- Figura 7.1 -

El problema planteado, como tal no está determinado, pero, considerando que por los $n+1$ puntos A_0 ; A_1 ; ... ; A_n (consultar la figura 7.1), pasa a lo sumo una parábola representativa de un polinomio $P_n(x)$ de grado n; así, por dos puntos pasa una sola recta; por tres puntos no alineados, pasa una sola parábola de segundo grado; etc.

Por consiguiente, conocidos $n+1$ puntos, la parábola de grado n que pasa por ellos permite asignar a cada valor de x un valor de y, que será considerado como el valor de interpolación buscado.

7.2 .- INTERPOLACIÓN

Dentro de la problemática general expresada en el punto anterior, es posible diferenciar algunos casos disimiles, dados tanto por la naturaleza del problema en cuestión (problemas de interpolación directa o inversa), como por las características de las tablas dadas como datos para realizar la interpolación (tablas de valores equidistantes o no).

Para su mejor estudio, y dado que no existe un solo y único método que resuelva todos los problemas que puedan presentarse, se sistematizará su tratamiento en diferentes puntos.

7.2.1.- TABLAS CON VALORES EQUIDISTANTES

Se comenzará por estudiar el problema para el caso más frecuente: aquel en el cual se tratan problemas de interpolación cuyas tablas tienen **VALORES EQUIDISTANTES** de la variable x; vale decir que, se dará por presupuesto que:

$$x_1 - x_0 = x_2 - x_1 = \ldots = x_n - x_{n-1} = h \tag{7.1}$$

y para ello se estudiarán, primeramente, las tablas de diferencias.

Cuando los valores de y que se disponen corresponden a valores equidistantes de x; vale decir, que se ajustan a la expresión (7.1), se puede formar una tabla de diferencias, de acuerdo con el esquema general mostrado en la tabla 7.1, donde los valores de las diferencias primeras Δy_k se obtienen restando a cada valor y_{k+1} el valor y_k que le antecede en la tabla. Para obtener las diferencias segundas se opera de igual modo, partiendo en este caso, de las diferencias primeras: las terceras se obtienen a partir de las segundas, y así sucesivamente hasta completar la tabla.

- Tabla de Diferencias Avanzadas -

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$
x_0	y_0			
		$\Delta y_0 = y_1 - y_0$		
$x_1 = x_0 + h$	y_1		$\Delta^2 y_0 = \Delta y_1 - \Delta y_0$	
		$\Delta y_1 = y_2 - y_1$		$\Delta^3 y_0 = \Delta^2 y_1 - \Delta^2 y_0$
$x_2 = x_1 + h$	y_2		$\Delta^2 y_1 = \Delta y_2 - \Delta y_1$	
		$\Delta y_2 = y_3 - y_2$		$\Delta^3 y_1 = \Delta^2 y_2 - \Delta^2 y_1$
$x_3 = x_2 + h$	y_3		$\Delta^2 y_2 = \Delta y_3 - \Delta y_2$	$\vdots$
		$\Delta y_3 = y_4 - y_3$	$\vdots$	$\vdots$
$x_4 = x_3 + h$	y_4	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\Delta^3 y_{n-3} = \Delta^2 y_{n-2} - \Delta^2 y_{n-3}$
$\vdots$	$\vdots$	$\vdots$	$\Delta^2 y_{n-2} = \Delta y_{n-1} - \Delta y_{n-2}$	
$\vdots$	$\vdots$	$\Delta y_{n-1} = y_n - y_{n-1}$		
$x_n = x_{n-1} + h$	y_n			

- Tabla N° 7.1 -

Completar la tabla significa llegar a diferencias cuyos valores son poco significativos, en valor absoluto, respecto de la precisión previamente elegida, con que se está operando en el problema en cuestión. Obviamente, la tabla 7.1 se extiende más allá de lo mostrado en la misma.

Obsérvese que, siendo:

$$\Delta y_0 = y_1 - y_0$$
$$\Delta^2 y_0 = \Delta y_1 - \Delta y_0 = y_2 - 2y_1 - y_0$$
$$\Delta^3 y_0 = \Delta^2 y_1 - \Delta^2 y_0 = y_3 - 3y_2 + 3y_1 - y_0$$

Recordando la expresión de los números combinatorios, estudiada por el álgebra:

$$\binom{n}{i} = \frac{n!}{i!\,(n-i)!}$$

resulta entonces:

$$\Delta^n y_0 = \sum_{i=o}^{n} (-1)^i \binom{n}{i} y_{n-i}$$

de donde, es posible deducir que las diferencias *n-ésimas* de y_0 se forman a partir de la ordenada *n-ésima* y sus n ordenadas antecedentes afectadas por los coeficientes del *Triángulo de TARTAGLIA*. Todas las diferencias así formadas reciben el nombre de **DIFERENCIAS AVANZADAS**.

Ejemplo 7.1 .- Confeccionar la tabla de DIFERENCIAS AVANZADAS correspondiente a los valores consignados en la siguiente tabla:

x	0	1	2	3	4
y	0	1	4	9	16

Solución: Debe ser notado que, los valores de y dados en la tabla, corresponden a la función $y = x^2$. En consecuencia, la tabla de diferencias avanzadas resulta:

x	y	Δy	Δ^2 y	Δ^3 y
0	0			
		1		
1	1		2	
		3		0
2	4		2	
		5		0
3	9		2	
		7		
4	16			

- Tabla N° 7.2 -

Las diferencias terceras de esta función son siempre nulas, pues, cualesquiera que sean los valores de x y h, será:

$$\Delta^3 = (x + 3h)^2 - 3(x + 2h)^2 + 3(x + h)^2 - x^2$$

expresión, esta última, que resulta idénticamente nula.

♦♦♦ ●●● ♦♦♦

Análogamente a lo visto en el ejemplo 7.1, es posible demostrar que, en general, las diferencias de orden n+1 de x^n son nulas y por consiguiente, que las diferencias de orden n+1 de un polinomio de grado n son todas nulas.

Obsérvese que, cuando se sabe que las diferencias terceras son todas nulas, se deduce que las diferencias segundas son constantes y que, por consiguiente, una vez conocido el valor de las diferencias segundas se puede agregar siempre un valor más a la tabla, en la columna de las diferencias primeras por tratarse de simple adición, (consúltese la tabla 7.2). Conociendo una nueva diferencia primera se puede obtener un nuevo valor de la función. De esta manera se pueden construir rápidamente tablas de potencias sin más que realizar simples adiciones.

Cualquier polinomio de grado n puede ser tabulado de esta manera una vez que a partir de n+1 valores se ha logrado calcular el valor constante de sus diferencias de orden n.

No ocurre lo mismo con las funciones trascendentes, dado que tienen diferencias sucesivas, hasta las de orden infinito. Pero, considerando solamente un cierto número de cifras decimales y un intervalo h suficientemente pequeño, las diferencias sucesivas posteriores van disminuyendo hasta poder considerarlas, a los efectos prácticos, como nulas desde algún n en adelante.

De hecho, esta metodología es utilizada para construir tablas de funciones trascendentes, haciendo intervenir las diferencias hasta un orden determinado; siendo el valor del orden, mayor o menor, según la precisión que se desea alcanzar.

7.3 .- MÉTODOS DE INTERPOLACIÓN

Como en muchos otros campos de la matemática aproximada o numérica, no existe una sola y única fórmula, algoritmo o método que resuelva todos los problemas que, en la práctica, puedan presentarse. Por el contrario, a través del presente capitulo se estudiarán varios métodos, los que han sido seleccionados de entre muchos, pero que, tratan de contemplar todas las posibilidades.

Notación: en virtud de lo oportuno que pueda resultar en cada caso particular, se utilizarán indistintamente, para expresar el valor de la función en un punto determinado, de abscisa $x = x_i$, las expresiones siguientes:

$$y_i \quad ; \quad y(x_i) \quad ; \quad f(x_i) \quad ; \quad f_i$$

A continuación se estudiará uno de las más importantes métodos de interpolación.

7.3.1.- FÓRMULA DE NEWTON-GREGORY ASCENDENTE

Conocidos los valores $y_0 ; y_1 ; \dots ; y_n$ de una función, que se corresponden con los $n+1$ valores equidistantes $x_0 ; x_1 ; \dots ; x_n$ de la variable, se trata de encontrar el polinomio de grado n:

$$P_n(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \dots + a_n(x - x_0)(x - x_1)\dots(x - x_{n-1}) \tag{7.2}$$

que pase por todos los $n+1$ puntos que determinan los $n+1$ pares ordenados $(x_i ; y_i)$ que constituyen la tabla de los datos.

Los $n+1$ coeficientes $a_0 ; a_1 ; \dots ; a_n$ se determinan imponiendo a la parábola (7.2) las $n+1$ condiciones de pasar por los puntos $A_0 ; A_1 ; \dots ; A_n$ (ver figura 7.1); es decir, estableciendo que para los valores de la variable $x=x_0 ; x=x_1 ; \dots ; x=x_n$, la expresión dada por (7.2) debe ser igual a $y_0 ; y_1 ; \dots ; y_n$, respectivamente; en otras palabras, es necesario hacer:

$$\begin{aligned}
&P_n(x_0) = y_0 = a_0 \\
&P_n(x_1) = y_1 = a_0 + a_1(x_1 - x_0) \\
&\dots\dots\dots\dots\dots\dots\dots\dots \\
&P_n(x_n) = y_n = a_0 + a_1(x_n - x_0) + \dots + a_n(x_n - x_0)\dots(x_n - x_{n-1})
\end{aligned}$$

de donde, despejando resultan:

$$a_0 = y_0$$

$$a_1 = \frac{y_1 - y_0}{x_1 - x_0} = \frac{\Delta y_0}{h}$$

$$a_2 = \frac{y_2 - a_0 - (x_2 - x_0)a_1}{(x_2 - x_0)(x_2 - x_1)} = \frac{y_2 - y_0 - 2h\dfrac{y_1 - y_0}{h}}{2.1\,h^2} = \frac{\Delta^2 y_0}{2!h^2}$$

y, en general:

$$a_n = \frac{\Delta^n y_0}{n!\,h^n} \tag{7.3}$$

Sustituyendo los valores de los coeficientes, calculados a partir de la expresión (7.3), en la ecuación (7.2), se llega a la Formula de **NEWTON-GREGORY ASCENDENTE**:

(7.4) $$P_n(x) = y_0 + \frac{\Delta y_0}{h}(x-x_0) + \frac{\Delta^2 y_0}{2!\,h^n}(x-x_0)(x-x_1) + \ldots +$$

$$+ \frac{\Delta^n y_0}{n!\,h^n}(x-x_0)(x-x_1)\ldots(x-x_{n-1})$$

Haciendo $x = x_0 + hu$, y sustituyéndola en la expresión (7.4), se obtiene una fórmula de uso más práctico. Bajo estas condiciones resulta:

$$x\text{-}x_0 = hu \quad ; \quad x\text{-}x_1 = x\text{-}(x_0 + h) = hu\text{-}h = h(u\text{-}1)$$

etc.; vale decir, la formula de **NEWTON-GREGORY** toma la forma:

(7.5) $$P_n(x) = P_n(x_0 + hu) = y_0 + u\,\Delta y_0 + \frac{u(u-1)}{2!}\Delta^2 y_0 +$$

$$+ \frac{u(u-1)(u-2)}{3!}\Delta^3 y_0 + \ldots + \frac{u(u-1)(u-2)\ldots(u-n+1)}{n!}\Delta^n y_0$$

Considerando que, según la teoría y notación de los números combinatorios, se puede expresar:

$$\frac{u!}{1!(u-1)!} = \binom{u}{1} = u$$

$$\frac{u!}{2!(u-2)!} = \binom{u}{2} = u(u-1)$$

. . . . . .

$$\frac{u!}{n!(u-n)!} = \binom{u}{n} = u(u-1)(u-2)\ldots(u-n+1)$$

y considerando a $\Delta^0 y_0 = y_0$, la expresión 7.5 se puede escribir:

(7.6) $$P_n(x) = \binom{u}{0} y_0 + \binom{u}{1}\Delta y_0 + \binom{u}{2}\Delta^2 y_0 + \ldots + \binom{u}{n}\Delta^n y_0 =$$

o, como sumatoria:

$$P_n(x) = \sum_{i=0}^{n} \binom{u}{i} \Delta^i y_0$$

Con el objeto de aportar mayores conocimientos referentes a la aplicación práctica de todo lo estudiado veamos el siguiente:

Ejemplo 7.2 .- Mediante la aplicación de la formula de *NEWTON-GREGORY*, calcular el valor de la función seno, para $x = 0{,}390736$; con un error menor de una milésima, conociendo la siguiente tabla de diferencias avanzadas:

x	$\operatorname{sen} x$	Δy	$\Delta^2 y$
0,390	0,380188415		
		0,000924719	
0,391	0,381113134		−0,00000381
		0,000924338	
0,392	0,382037472		−0,00000383
		0,000923955	
0,393	0,382961427		−0,00000382
		0,000923573	
0,394	0,383885000		

- Tabla 7.3 -

Solución: Sabiendo que $x_0 = 0{,}390$, y haciendo $h = 0{,}001$; $u = \frac{x - x_0}{h} = 0{,}736$; entonces se tiene:

$$P(0{,}390736) = 0{,}380188415 + 0{,}736 \cdot 0{,}00924719 - \frac{0{,}736\,(-0{,}264)}{2} \cdot 0{,}00000381 =$$

$$= 0{,}380869045$$

◆◆◆ ●●● ◆◆◆

7.3.2.- INTERPOLACIÓN LINEAL Y CUADRÁTICA

Si en una tabla de valores, las diferencias son nulas, a partir de la segunda en adelante, la formula de *NEWTON-GREGORY* se reduce a la siguiente:

(7.7) $$P_1(x) = y_0 + \frac{\Delta y_0}{h}(x - x_0)$$

que no es más que, la expresión analítica de la recta que pasa por los puntos A_0 ; A_1 , (ver figura 7.1).

Este tipo de interpolación solo puede utilizarse cuando, como ya se ha dicho, las diferencias tabulares de orden dos y mayores son nulas; o bien, cuando dentro del orden de aproximación con que se está operando, las diferencias de orden superior Δ^2 ; Δ^3 y siguientes, son despreciables. En este caso, la curva representativa de la función es reemplazada por la poligonal que se obtiene uniendo los puntos A_0 ; A_1 ; ... ; A_n , con segmentos de recta. Es por este motivo, que la aplicación de la fórmula (7.7) recibe el nombre de **FÓRMULA DE INTERPOLACIÓN LINEAL**.

Si las diferencias tabulares de orden mayor al segundo son nulas; o bien, pueden ser consideradas como tales, en virtud del grado de aproximación con que se está trabajando, se puede utilizar la fórmula de *NEWTON-GREGORY* incluyendo los términos de segundo orden y menores:

(7.8) $$P_2(x) = y_0 + \frac{\Delta y_0}{h}(x - x_0) + \frac{\Delta^2 y_0}{2!\,h^2}(x - x_0)(x - x_1)$$

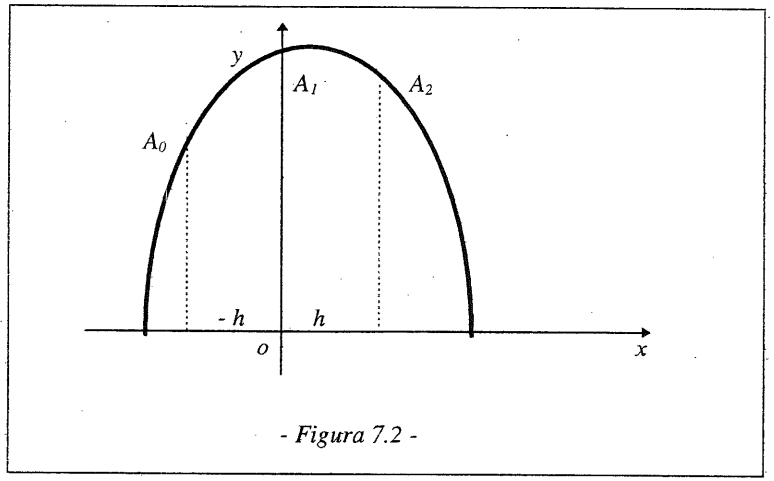
que no es otra cosa que la ecuación de la parábola de segundo grado, que pasa por los puntos A_0 ; A_1 y A_2 .

No se pierde generalidad si se considera el valor x_1 coincidiendo con el origen de coordenadas (ver figura 7.2), en cuyo caso es:

$$x_0 = -h \;\; ; \;\; x_1 = 0 \;\; ; \;\; x_2 = h$$

y la ecuación de la parábola que pasa por los puntos A_0 ; A_1 y A_2 se puede expresar, en general:

(7.9) $$y = a x^2 + b x + c$$



- Figura 7.2 -

Siendo $x_0 = -h$; $x_1 = 0$ y $x_2 = h$, la expresión (7.8) es válida para todos estos puntos, entonces, según la expresión (7.9) se puede escribir:

$$\begin{cases} y_0 = a x_0^2 + b x_0 + c = a(-h)^2 + b(-h) + c \\ y_1 = a x_1^2 + b x_1 + c = c \\ y_2 = a x_2^2 + b x_2 + c = a h^2 + b h + c \end{cases}$$

de la segunda, se obtiene directamente que $c = y_1$; y reemplazando este valor en las otras dos ecuaciones, resulta:

$$\begin{cases} a h^2 - b h + y_1 = y_0 \\ a h^2 + b h + y_1 = y_2 \end{cases}$$

de donde, sumando y restando ambas ecuaciones, se obtiene:

$$a = \frac{y_2 - 2y_1 + y_0}{2h^2} \quad ; \quad b = \frac{y_2 - y_0}{2h}$$

valores que reemplazados en la ecuación (7.9) da como resultado:

(7.10) $$y = \frac{y_2 - 2y_1 + y_0}{2h^2} x^2 + \frac{y_2 - y_0}{2h} x + y_1$$

expresión que es generalmente denominada **FÓRMULA DE INTERPOLACIÓN CUADRÁTICA**.

7.3.3.- FÓRMULA DE NEWTON-GREGORY DESCENDENTE

La fórmula de *NEWTON-GREGORY ASCENDENTE* ya estudiada, permite resolver el problema de interpolar en una tabla de intervalo constante h un valor de y correspondiente al valor x de la variable comprendida entre x_0 y x_n, pero como las diferencias sucesivas Δy_0 ; $\Delta^2 y_0$; etc., que aparecen en la fórmula se refieren al primer valor y_0, esta se emplea con mayor precisión cuando el valor de x, correspondiente al y que se desea calcular, está próximo a x_0 ; vale decir, al principio de la tabla.

Cuando, por el contrario, la interpolación debe efectuarse para un valor de x próximo a x_n o, en general, alejado de x_0, interesa conocer otras fórmulas de interpolación en las que intervengan las diferencias sucesivas que se relacionan con el último valor y_n de la tabla.

En este orden de cosas, definiendo la diferencia de primer orden mediante la expresión:

(7.11) $$\Delta y_{n-1} = y_n - y_{n-1} = \nabla y_n$$

y, operando de igual modo al realizado para definir las diferencias avanzadas, obtenidas en la tabla 7.1, las restantes se encuentran detalladas en la tabla 7.4.

- Tabla de Diferencias Atrasadas -

x	y	∇y	$\nabla^2 y$	$\nabla^3 y$
x_0	y_0			
		$\nabla y_1 = y_1 - y_0$		
$x_1 = x_0 + h$	y_1		$\nabla^2 y_1 = \nabla y_1 - \nabla y_0$	
		$\nabla y_2 = y_2 - y_1$		$\nabla^3 y_1 = \nabla^2 y_1 - \nabla^2 y_0$
$x_2 = x_1 + h$	y_2		$\nabla^2 y_{21} = \nabla y_2 - \nabla y_1$	
		$\nabla y_3 = y_3 - y_2$		$\nabla^3 y_2 = \nabla^2 y_2 - \nabla^2 y_1$
$x_3 = x_2 + h$	y_3		$\nabla^2 y_3 = \nabla y_3 - \nabla y_2$	⋮
		$\nabla y_4 = y_4 - y_3$	⋮	⋮
$x_4 = x_3 + h$	y_4	⋮	⋮	⋮
⋮	⋮	⋮	⋮	$\nabla^3 y_{n-2} = \nabla^2 y_{n-2} - \nabla^2 y_{n-3}$
⋮	⋮	⋮	$\nabla^2 y_{n-1} = \nabla y_{n-1} - \nabla y_{n-2}$	
⋮	⋮	$\nabla y_n = y_n - y_{n-1}$		
$x_n = x_{n-1} + h$	y_n			

- Tabla N° 7.4 -

Las diferencias calculadas en la tabla 7.4, reciben el nombre de **DIFERENCIAS ATRASADAS**.

Para deducir la fórmula de interpolación correspondiente, es necesario escribir la (7.12) en forma análoga a la ya utilizada para la ascendente; la expresión de la ecuación de la parábola de grado n es:

(7.12) $$P_n(x) = a_0 + a_1 (x - x_n) + a_2 (x - x_n)(x - x_{n-1}) + \ldots$$

$$+ a_n (x - x_n)(x - x_{n-1}) \ldots (x - x_1)$$

en la que, determinando sus coeficientes a_0 ; a_1 ; ...; a_n mediante las $n+1$ condiciones que impone el hecho que la parábola (7.12) tenga que pasar por los puntos A_0 ; A_1 ; ...; A_n (ver figura 7.1).

En la expresión 7.12, haciendo sucesivamente $x=x_n$; $x=x_{n-1}$; ...; $x=x_1$, se obtienen los valores:

$$P_n(x_n) = y_n = a_0$$
$$P_n(x_{n-1}) = y_{n-1} = a_0 + a_1(x_{n-1} - x_n)$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$P_n(x_0) = y_0 = a_0 + a_1(x_0 - x_n) + \ldots + a_n(x_0 - x_0)\ldots(x_0 - x_1)$$

de donde, despejando resultan los valores de los coeficientes:

$$a_0 = y_n$$

$$a_1 = \frac{y_{n-1} - y_n}{x_{n-1} - x_n} = \frac{\nabla y_n}{h}$$

y así sucesivamente. En general, será:

$$a_n = \frac{\nabla^n y_n}{n!\,h^n}$$

Sustituyendo los valores de los coeficientes calculados en la ecuación (7.12), se obtiene la Formula de **NEWTON-GREGORY DESCENDENTE**:

(7.13) $$P_n(x) = y_n + \frac{\nabla y_n}{h}(x - x_n) + \frac{\nabla^2 y_n}{2!\,h^2}(x - x_n)(x - x_{n-1}) + \ldots +$$

$$+ \frac{\nabla^n y_n}{n!\,h^n}(x - x_n)(x - x_{n-1})\ldots(x - x_1)$$

Realizando el cambio de la variable x por $x+hu$, resulta la expresión más práctica:

(7.14) $$P_n(x) = P_n(x + hu) = y_n + u\nabla y_n + \frac{u(u+1)}{2!}\nabla^2 y_n +$$

$$+ \frac{u(u+1)(u+2)}{3!}\nabla^3 y_n + \ldots + \frac{u(u+1)(u+2)\ldots(u+n-1)}{n!}\nabla^n y_n$$

pues:

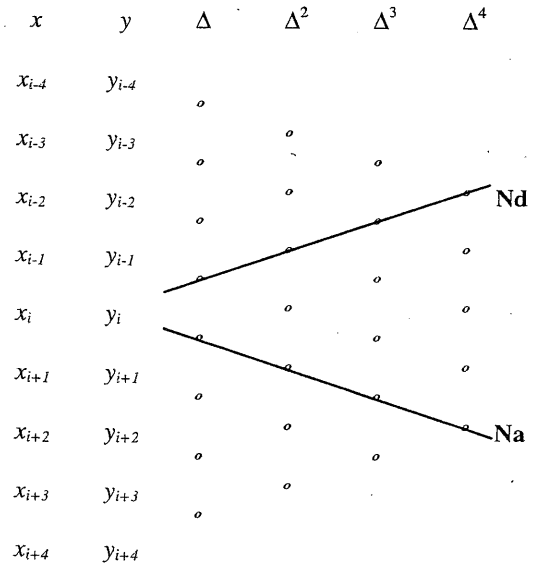
$$x - x_n = h\,u$$

$$x - x_{n-1} = x - (x_n - h) = x - x_n + h = h\,u + h = h\,(u + 1)$$

$$x - x_{n-2} = x - (x_{n-1} - h) = x - x_{n-1} + h = h\,(u + 1) + h = h\,(u + 2)$$

y así siguiendo.

En el esquema representado en la figura 7.3, se han unido mediante segmentos de rectas los valores que intervienen en las fórmulas de interpolación de *NEWTON-GREGORY ASCENDENTE* (**Na**) y en la *DESCENDENTE* (**Nd**).



- Figura 7.3 -

A los efectos de ilustrar sobre los efectos de esta última fórmula, veamos el siguiente:

Ejemplo 7.3.- Sea el problema de tener que determinar el logaritmo neperiano del número 5143,75.

Solución: Conociendo los siguientes valores de la tabla de diferencias, confeccionada a partir de alguna tabla:

x	y	Δ	Δ^2	Δ^3
5100	8,5369958187			
		19588645		
5110	8,5389446832		−38297	
		19550348		151
5120	8,5409097180		−38146	
		19512202		146
5130	8,5428609382		−38000	
		19474202		151
5140	8,5448083584		−37849	
		19436353		
5150	8,5467519937			

- Tabla 7.4 -

y considerando arbitrariamente la columna de las diferencias terceras como la de las constantes, es posible hacer:

$$x_n = 5150 \quad ; \quad h = 10 \quad ; \quad u = \frac{5143{,}75 - 5150}{10} = -0{,}625$$

y sustituyendo en la (7.14), resulta:

$$P_n(x_n + h\,u) = P_n(5143{,}75) = 8{,}5467519937 - 0{,}625 \,.\, 0{,}0019436353 +$$

$$+ {}^1/_2 (0{,}625 \,.\, 0{,}375) \,.\, 0{,}0000037849 -$$

$$- {}^1/_6 (0{,}625 \,.\, 0{,}375 \,.\, 1{,}375) \,.\, 0{,}0000000151 =$$

$$= 8{,}54553766443$$

♦♦♦ ●●● ♦♦♦

7.3.4.- FÓRMULA DE LAGRANGE

Mientras se tenga que interpolar sobre tablas numéricas confeccionadas exprofeso, difícilmente se planteará el caso de tener que considerar intervalos desiguales de la variable, pero, cuando la interpolación debe realizarse por medio de tablas obtenidas experimentalmente, es muy frecuente tener que manipular tablas que poseen intervalos no equidistantes.

Una de las fórmulas más conocidas, que resuelven este problema, es la denominada **FÓRMULA DE LAGRANGE** que puede deducirse a partir del polinomio de grado n al cual se le impone pasar por los $n+1$ puntos $A_0 ; A_1 ; \ldots ; A_n$, de la forma:

$$P_n(x) = a_0 (x - x_1)(x - x_2) \ldots (x - x_n) + a_1 (x - x_0)(x - x_2) \ldots (x - x_n) + \ldots +$$

$$+ a_n (x - x_0)(x - x_1) \ldots (x - x_{n-1})$$

donde cada uno de los términos que componen este polinomio, es a su vez, un polinomio de grado n afectado de un coeficiente a_r, el cual debe ser aún determinado.

Los $n+1$ coeficientes a_r se determinan imponiendo las $n+1$ condiciones:

$$\begin{array}{llll} para & x = x_0 & \Rightarrow & y = y_0 \\ para & x = x_1 & \Rightarrow & y = y_1 \\ \ldots\ \ldots & \ldots\ \ldots & \ldots & \ldots\ \ldots \\ para & x = x_n & \Rightarrow & y = y_n \end{array}$$

así, se obtienen las ecuaciones:

$$\begin{array}{l} y_0 = a_0 (x_0 - x_1)(x_0 - x_2) \ldots (x_0 - x_n) \\ y_1 = a_1 (x_1 - x_0)(x_1 - x_2) \ldots (x_1 - x_n) \\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots \\ y_n = a_n (x_n - x_0)(x_n - x_1) \ldots (x_n - x_{n-1}) \end{array}$$

Despejando de estas ecuaciones los valores de los a_r y sustituyéndolos en el polinomio original, resulta la denominada **FORMULA DE LAGRANGE**:

$$P_n(x) = +y_0\frac{(x-x_1)(x-x_2)\ldots(x-x_n)}{(x_0-x_1)(x_0-x_2)\ldots(x_0-x_n)}+$$

$$+y_1\frac{(x-x_0)(x-x_2)\ldots(x-x_n)}{(x_1-x_0)(x_1-x_2)\ldots(x_1-x_n)}+\ldots+ \tag{7.15}$$

$$+y_n\frac{(x-x_0)(x-x_1)\ldots(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)\ldots(x_n-x_{n-1})}$$

Dado que los numeradores de todos los términos del segundo miembro son iguales, en la práctica es más cómodo operar con la expresión que se obtiene dividiendo ambos miembros de (7.15) por el producto $(x - x_0)(x - x_1) \ldots (x - x_n)$, con lo cual resulta:

$$\frac{P_n(x)}{(x-x_0)\ldots(x-x_n)} = \frac{y_0}{(x-x_0)(x_0-x_1)(x_0-x_2)\ldots(x_0-x_n)}+$$

$$+\frac{y_1}{(x-x_1)(x_1-x_0)(x_1-x_2)\ldots(x_1-x_n)}+\ldots+ \tag{7.16}$$

$$+\frac{y_n}{(x-x_n)(x_n-x_0)(x_n-x_1)\ldots(x_n-x_{n-1})}$$

La expresión anterior es la usualmente utilizada en las aplicaciones prácticas para realizar la interpolación en tablas con valores no equidistantes.

Se puede apreciar la aplicación de esta importante fórmula mediante el siguiente:

Ejemplo 7.4.- Por medio de una experiencia de laboratorio se ha obtenido la siguiente tabla de valores:

x	0	490	840	1003
y	0	3152	5036	5773

en la que se desea calcular el correspondiente valor de y para $x = 700$.

Solución: Empleando la fórmula de *LAGRANGE*, se obtiene:

$$\frac{P_3(x)}{x(x-490)(x-840)(x-1003)} = \frac{3152}{(x-490)(490-0)(490-840)(490-1003)}+$$

$$+\frac{5036}{(x-840)(840-0)(840-490)(840-1003)}+$$

$$+\frac{5773}{(x-1003)(1003-0)(1003-490)(1003-840)}$$

en la que sustituyendo el valor de la variable x por 700, se obtiene el valor de y buscado, el que resulta:

$$P(700) = 4328$$

••• ●●● •••

7.3.4.1.- Fórmula de LAGRANGE para Valores Equidistantes

La fórmula de *LAGRANGE* puede ser utilizada, con apreciables ventajas, también para interpolar en tablas con valores equidistantes de la variable.

En estos casos, como se verá, los coeficientes de la fórmula de Lagrange resultan independientes de los valores de las abscisas de la tabla dada y de su incremento tabular h; de manera que, los coeficientes se calculan de una vez para siempre, recibiendo el nombre de **COEFICIENTES LAGRANGIANOS**. Consecuentemente, para facilitar la tarea, se han confeccionado tablas de los coeficientes *LAGRANGIANOS* que consideran los casos correspondientes a 3; 4;...; 11, etc. puntos.

Veamos el caso particular que corresponde a una tabla de coeficientes *LAGRANGIANOS* de cuatro puntos.

Conocidos los valores f_{-1} ; f_0 ; f_1 ; f_2 correspondientes a x_{-1} ; x_0 ; x_1 ; x_2 , respectivamente, el polinomio de *LAGRANGE* será:

$$f(x)=\frac{(x-x_0)(x-x_1)(x-x_2)}{(x_{-1}-x_0)(x_{-1}-x_1)(x_{-1}-x_2)}f_{-1}+$$

$$+\frac{(x-x_{-1})(x-x_1)(x-x_2)}{(x_0-x_{-1})(x_0-x_1)(x_0-x_2)}f_0+$$

(7.17)

$$+\frac{(x-x_{-1})(x-x_0)(x-x_2)}{(x_1-x_{-1})(x_1-x_0)(x_1-x_2)}f_1+$$

$$+\frac{(x-x_{-1})(x-x_0)(x-x_1)}{(x_2-x_{-1})(x_2-x_0)(x_2-x_1)}f_2+$$

donde $x_0 - x_{-1} = x_1 - x_0 = x_2 - x_1 = h$. Si se considera ahora, el cambio de variable $x = x_0 + hu$, resulta:

$$f(x)=f(x_0+hu)=-\frac{u(u-1)(u-2)}{6}f_{-1}+$$

$$+\frac{(u+1)(u-1)(u-2)}{2}f_0-$$

$$-\frac{(u+1)u(u+2)}{2}f_1+$$

$$+\frac{(u+1)u(u-1)}{6}f_2$$

Llamando L_i a los coeficientes *LAGRANGIANOS*, estos pueden ser muy fácilmente calculados. Entonces, tomando:

$$L_{-1}=-\frac{u(u-1)(u-2)}{6} \qquad L_0=\frac{(u+1)(u-1)(u-2)}{2}$$

$$L_1=\frac{(u+1)u(u+2)}{2} \qquad L_2=\frac{(u+1)u(u-1)}{6}$$

para valores de u variando de 0 a 1, con distintos intervalos o números de decimales.

Teniendo tabulados los coeficientes L_i, la interpolación se reduce a calcular la expresión:

$$f(x)=L_{-1}f_{-1}+L_0f_0+L_1f_1+L_2f_2$$

evitándose, de esta manera, la aplicación de la extensa fórmula estudiada, previa construcción de la tabla de diferencias.

7.3.5.- FÓRMULA DE INTERPOLACIÓN PARABÓLICA PROGRESIVA

Se expondrá a continuación una fórmula, atribuida a *NEWTON*, que tiene la ventaja de que, término a término, es posible obtener una cota del error que se ha cometido hasta ese momento.

Sea el polinomio (7.2) ya utilizado al deducir las fórmulas de *NEWTON-GREGORY*:

(7.2) $$P_n(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \ldots + a_n(x - x_0)(x - x_1)\ldots(x - x_{n-1})$$

y, al igual que antes, imponiéndole la condición que la curva representativa pase por los $n+1$ puntos A_0 ; A_1 ;...; A_n .

La condición que la curva pase por A_0 implica que $y_0 = a_0$; la cual puede considerarse como la primera aproximación de la fórmula de *Interpolación Parabólica Progresiva*; es decir, como el polinomio de interpolación de orden cero:

(7.19) $$P_0(x) = y_0$$

Si, en una segunda aproximación, se exige que la representación gráfica pase por los puntos A_0 y A_1 , debe verificarse, además:

(7.20) $$y_1 = a_0 + a_1(x_1 - x_0)$$

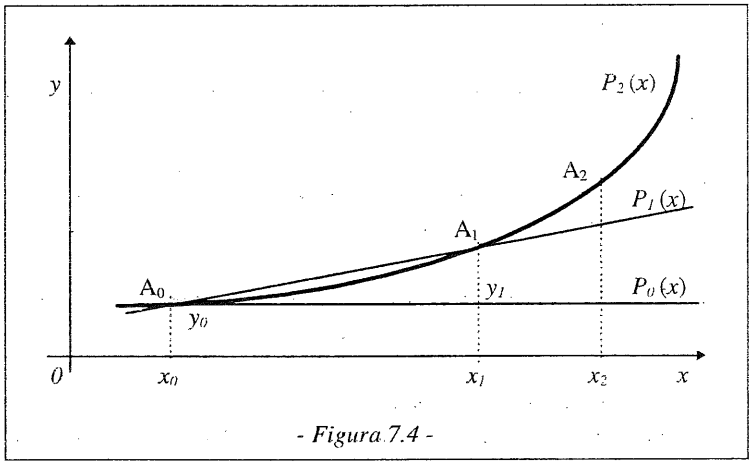
de la cual puede calcularse el valor de a_1 , que resulta:

$$a_1 = \frac{y_1 - a_0}{x_1 - x_0} = \frac{y_1 - P_0(x_1)}{x_1 - x_0}$$

Reemplazando este valor de a_1 en la expresión anterior, puede obtenerse el polinomio de interpolación de orden uno, ver figura 7.4, el que resulta:

(7.20) $$P_1(x) = y_0 + \frac{y_1 - P_0(x_1)}{x_1 - x_0}(x - x_0)$$

donde el numerador que figura en la parte fraccionaria, que vale $y_1 - P_0(x_1)$ es la magnitud del error que se comete al estimar x_1 cuando se toma como polinomio de interpolación el polinomio P_0.



- Figura 7.4 -

La condición adicional que la curva deba pasar, además, por el punto A_2, impone para poder realizar la deducción del polinomio de interpolación, la utilización de la expresión:

$$y_2 = a_0 + a_1 (x_2 - x_0) + a_2 (x_2 - x_0)(x_2 - x_1)$$

de la cual es posible deducir el valor del nuevo coeficiente a_2 que debe agregarse a la expresión general. Resulta entonces:

$$a_2 = \frac{y_2 - [a_0 + a_1 (x_2 - x_0)]}{(x_2 - x_0)(x_2 - x_1)} = \frac{y_2 - P_1(x_2)}{(x_2 - x_0)(x_2 - x_1)}$$

de donde:

(7.21) $$P_2(x) = y_0 + \frac{y_1 - P_0(x_1)}{(x_1 - x_0)}(x - x_0) + \frac{y_2 - P_1(x_2)}{(x_2 - x_0)(x_2 - x_1)}(x - x_0)(x - x_1)$$

en la cual el numerador de a_2 es el error que se cometería si se tomara como valor de y_2 el valor que proporciona el polinomio de interpolación de orden uno.

Continuando este análisis y operando de idéntica manera a la estudiada, puede obtenerse el polinomio de interpolación de orden r:

(7.22) $$P_r(x) = y_0 + \frac{y_1 - P_0(x_1)}{(x_1 - x_0)}(x - x_0) + \ldots + \frac{y_r - P_{r-1}(x_r)}{(x_r - x_0)\ldots(x_r - x_{r-1})}(x - x_0)\ldots(x - x_{r-1})$$

donde el numerador de cada coeficiente a_k mide precisamente el valor del error $y_k - P_{k-1}(x_k)$ que se comete al considerar, en lugar del valor y_k, el valor dado por el polinomio de aproximación de orden $(k-1)$ en el punto de abscisa $x=x_k$.

Es importante destacar que, en ningún momento se ha impuesto la condición que los intervalos de tabla sean regulares; vale decir, la fórmula de *NEWTON* es muy útil para interpolar en todos aquellos casos en que las tablas resultantes tienen intervalos no equidistantes.

La diferencia entre los procedimientos de *LAGRANGE* y de *NEWTON*, que se acaban de estudiar, es que en el de *LAGRANGE* todos los términos tienen el mismo grado; mientras que, en el de *NEWTON* el grado de los distintos términos va en aumento progresivamente.

En el primero, solo es posible determinar una cota del error, y en caso de ser necesaria una aproximación mayor a la obtenida es menester rehacer integralmente el cálculo. Mientras que, con el método de *NEWTON*, los numeradores de los coeficientes indican, sucesivamente, la magnitud de la precisión lograda en el paso anterior.

Para aportar mayores detalles se presenta una aplicación del método de *NEWTON* mediante el siguiente:

Ejemplo 7.5.- Haciendo uso de la siguiente tabla, determinada experimentalmente en tareas de laboratorio, se desea calcular mediante la *Fórmula de Interpolación Parabólica Progresiva* o de *NEWTON*, el valor de y correspondiente a x=10:

x	0	5	15	30
y	63	89	173	429

Solución: Para determinar el valor de los coeficientes es necesario hacer:

$$a_0 = y_0 = 63 \qquad P_0(x) = 63$$

$$a_1 = \frac{y_1 - P_0(x_1)}{(x_1 - x_0)} = \frac{89-63}{(5-0)} = 5{,}2 \qquad P_1(x) = 63 + 5{,}2x$$

$$a_2 = \frac{y_2 - P_1(x_2)}{(x_2 - x_0)(x_2 - x_1)} = \frac{173-141}{15(15-5)} = 0{,}21$$

$$P_2(x) = 63 + 5{,}2x + 0{,}21x(x-5)$$

$$a_3 = \frac{y_3 - P_2(x_3)}{(x_3 - x_0)(x_3 - x_1)(x_3 - x_2)} = \frac{429 - 376{,}5}{30.25.15} = 0{,}0047$$

$$P_3(x) = 63 + 5{,}2x + 0{,}21x(x-5) + 0{,}0047x(x-5)(x-15)$$

que para x=10, resulta y=124,325. Este valor, no es posible mejorar dado que, no se cuenta con más puntos en la tabla de valores que permitan seguir con la aplicación de la *Fórmula de Interpolación Parabólica Progresiva.*

◆◆◆ ●◆● ◆◆◆

7.3.6.- FÓRMULA DE GAUSS

Sea la expresión (7.2) del polinomio de interpolación dado:

(7.2) $$P_n(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \ldots + a_n(x - x_0)(x - x_1)\ldots(x - x_{n-1})$$

y considerando los puntos de abscisas equidistantes:

$$x_0 = x_0\,;\; x_1 = x_0 + h\,;\; x_2 = x_0 - h\,;\; x_3 = x_0 + 2h\,;\; x_4 = x_0 - 2h\,;\; \ldots$$

Sustituyendo estos valores en las expresiones de a_r de *la Fórmula de Interpolación Parabólica Progresiva*, se obtiene:

$$a_0 = y_0 = P_0(x_0) = f_0$$

$$a_1 = \frac{y_1 - P_0(x_1)}{(x_1 - x_0)}$$

de donde:

$$a_1 = \frac{f(x_0 + h) - f(x_0)}{h}$$

de la misma manera:

$$a_2 = \frac{y_2 - P_1(x_2)}{(x_2 - x_0)(x_2 - x_1)}$$

sea:

$$a_2 = \frac{f(x_0 + h) - 2f(x_0) + f(x_0 - h)}{2h^2}$$

y, finalmente:

$$a_3 = \frac{y_3 - P_2(x_3)}{(x_3 - x_0)(x_3 - x_1)(x_3 - x_2)}$$

vale decir:

$$a_3 = \frac{f(x_0+2h)-3f(x_0+h)+3f(x_0)-f(x_0-h)}{6h^3}$$

Y así sucesivamente, puede continuarse el procedimiento. A continuación, reemplazando los coeficientes obtenidos en el polinomio de interpolación, resulta la siguiente expresión:

$$P_n(x) = f(x_0) + \frac{f(x_0+h)-f(x_0)}{h}(x-x_0) +$$

$$+\frac{f(x_0+h)-2f(x_0)+f(x_0-h)}{2!h^2}(x-x_0)(x-x_1) + \tag{7.23}$$

$$+\frac{f(x_0+2h)-3f(x_0+h)+3f(x_0)-f(x_0-h)}{3!h^3}(x-x_0)(x-x_1)(x-x_2)+\ldots$$

Una expresión más práctica que la anterior, se obtiene haciendo la conocida sustitución $x=x_0+hu$, con la cual, finalmente, resulta la **FÓRMULA DE INTERPOLACIÓN DE GAUSS**:

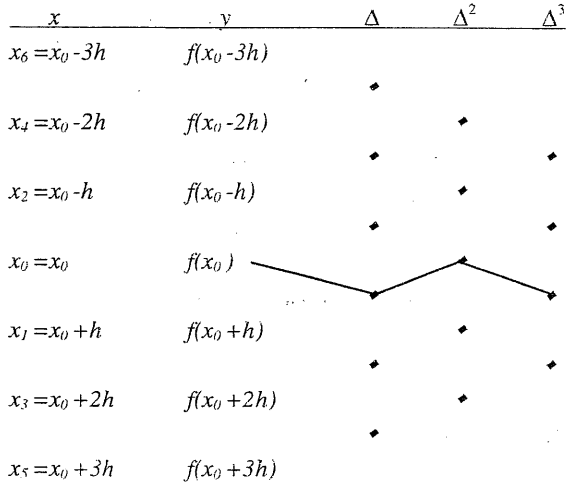
$$P_n(x) = P_n(x_0+hu) = f(x_0) + \Delta f(x_0)u +$$

$$+\frac{\Delta^2 f(x_0-h)}{2!}u(u-1) + \frac{\Delta^3 f(x_0+h)}{3!}(u+1)u(u-1) + \tag{7.24}$$

$$+\frac{\Delta^4 f(x_0-2h)}{4!}(u+1)u(u-1)(u-2)+\ldots$$

Considerando el esquema representado en la figura 7.5, es posible observar que esta fórmula puede ser recordada fácilmente si se observa que las diferencias que se utilizan están sobre la poligonal de eje horizontal.

En virtud de esta característica, es posible deducir que la fórmula de *GAUSS* es conveniente que sea utilizada en todos aquellos casos en que se trata de interpolar valores de la función que corresponden a valores de la variable cercanos al sector medio de la tabla.



- Figura 7.5 -

Para mejor comprensión del tema y fijar conocimientos, se presenta el siguiente:

Ejemplo 7.6.- Determinar el valor de e^x para x=1,5708; conociendo la tabla de diferencias dada en el siguiente cuadro:

x	$y=e^x$	Δ	Δ^2	Δ^3	Δ^4	Δ^5
1,3	3,6693					
		0,3859				
1,4	4,0552		0,0406			
		0,4265		0,0042		
1,5	4,4817		0,0448		0,0004	
		0,4713		0,0048		0,0000
1,6	4,9530		0,0496		0,0004	
		0,5209		0,0052		
1,7	5,4739		0,0548			
		0,5757				
1,8	6,0496					

Solución: Siendo $x_0 = 1{,}5$; $h = 0{,}1$; $u = x - x_0 = 0{,}708$; la aplicación de la fórmula de *GAUSS* da:

$$y \approx 4{,}4817 + 0{,}708 \,.\, 0{,}4713 + {}^1/_2[0{,}708\ (0{,}708 - 1)] \,.\, 0{,}0448 +$$

$$+ {}^1/_6[(0{,}708 + 1)\ 0{,}708\ (0{,}708 - 1)] \,.\, 0{,}0048 +$$

$$+ {}^1/_{24}[(0{,}708 + 1)\ 0{,}708\ (0{,}708 - 1)\ (0{,}708 - 2)] \,.\, 0{,}0004 = \underline{\mathbf{4.81047}}$$

♦♦♦ ●●● ♦♦♦

7.3.7.- FÓRMULA DE BESSEL

La **FÓRMULA DE INTERPOLACIÓN DE BESSEL** es una de las más utilizadas debido a las múltiples ventajas que presenta, tanto en la interpolación directa, cuanto en la inversa; debido a ello, se le dará la importancia que realmente tiene entre el amplio conjunto de expresiones destinadas a resolver el problema de la interpolación.

Es posible deducir la **FÓRMULA DE BESSEL** a partir de la fórmula de interpolación de *GAUSS*. Se tratarán de tomar de esta última, solamente los elementos de una fila horizontal de la tabla de diferencias para ser utilizadas en la deducción. Desdoblando las diferencias avanzadas de orden par, resulta la expresión:

$$f(x) = f(x_0 + hu) = \frac{1}{2}f(x_0) + \frac{1}{2}f(x_0) + u\Delta f(x_0) +$$

$$+\frac{u(u-1)}{2}\left[\frac{1}{2}\Delta^2 f(x_0 - h) - \frac{1}{2}\Delta^2 f(x_0 - h)\right] + \frac{(u+1)u(u-1)}{6}\Delta^3 f(x_0 - h) + \ldots =$$

$$= \frac{1}{2}f(x_0) + \frac{1}{2}\left[f(x_0 + h) - \Delta f(x_0)\right] + u\Delta f(x_0) +$$

$$+\frac{u(u-1)}{2}\left[\frac{1}{2}\Delta^2 f(x_0 - h) + \frac{1}{2}\Delta^2 f(x_0) - \frac{1}{2}\Delta^3 f(x_0 - h)\right] +$$

$$+\frac{(u+1)u(u-1)}{6}\Delta^3 f(x_0 - h) + \ldots$$

y, en definitiva la expresión:

$$(7.25) \qquad f(x) = \frac{1}{2}\left[f(x_0) + f(x_0 + h) + \left(u - \frac{1}{2}\right)\Delta f(x_0)\right] +$$

$$+\frac{u(u-1)}{2!}\frac{1}{2}\left[\Delta^2 f(x_0 - h) + \Delta^2 f(x_0)\right] +$$

$$+\frac{u(u-1)\left(u - \frac{1}{2}\right)}{3!}\Delta^3 f(x_0 - h) +$$

$$+\frac{(u+1)u(u-1)(u-2)}{4!}\frac{1}{2}\left[\Delta^4 f(x_0 - 2h) + \Delta^4 f(x_0 - h)\right] + \cdots$$

la que constituye la **FÓRMULA DE INTERPOLACIÓN DE BESSEL**, que puede escribirse de manera más simétrica:

$$(7.26) \qquad f(x) = f(x_0 + hu) = f(x_0) + B'\,\Delta f(x_0) + B''\left[\Delta^2 f(x_0 - h) + \Delta^2 f(x_0)\right] +$$

$$+B'''\,\Delta^3 f(x_0 - h) + B^{IV}\left[\Delta^4 f(x_0 - 2h) + \Delta^2 f(x_0 - h)\right] + \ldots$$

siendo:

$$B' = u \quad ; \quad B'' = \frac{u(u-1)}{2.2!} \quad ; \quad B''' = \frac{u(u-1)\left(u - \frac{1}{2}\right)}{3!} \quad ; \quad B^{IV} = \frac{(u+1)u(u-1)(u-2)}{2.4!}$$

y así sucesivamente. Estos coeficientes reciben el nombre de **COEFICIENTES BESSELIANOS** y se encuentran tabulados para valores de u comprendidos entre cero y uno, lo que permite efectuar con rapidez la interpolación.

Se estudiará la aplicación práctica de la fórmula de *BESSEL* mediante el siguiente:

Ejemplo 7.7.- Resolver el problema dado en el ejercicio N° 7.6, haciendo uso de la fórmula de interpolación de *BESSEL*.

Solución: Mediante la aplicación directa de la fórmula de BESSEL sobre los distintos elementos de la tabla de diferencias, se obtiene:

$$y = {}^1/_2(4{,}4817 + 4{,}9530) + (0{,}708 - 0{,}5)\,0{,}4713 +$$

$$+ {}^1/_4 0{,}708\,(0{,}708 - 1)\,(0{,}0448 + 0{,}0496) + {}^1/_6\,0{,}708\,(0{,}708 - 1)\,(0{,}708 - 0{,}5)\,0{,}0048 +$$

$$+ {}^1/_{48}(0{,}708 + 1)\,0{,}708\,(0{,}708 - 1)\,(0{,}708 - 2)\,0{,}0008 = \underline{\mathbf{4{,}81047}}$$

Como puede apreciarse arroja el mismo resultado. Si se considera que la parte correspondiente a coeficientes se encuentra tabulada, se concluye en la practicidad del método de *BESSEL*.

◆◆◆ ●●● ◆◆◆

7.3.8.- FÓRMULA DE STIRLING

Esta es otra fórmula que utiliza elementos de una sola fila horizontal de la tabla de diferencias avanzadas, como se aprecia en la figura 7.6, ofreciendo evidentes ventajas para el caso general de interpolación, ya que, es utilizable con mucha mayor precisión en un sector más amplio de la tabla dada, que las fórmulas ya estudiadas.

Escribiendo la fórmula de *GAUSS* bajo la forma que utiliza el cambio de variable $x=x_0+hu$, resulta:

$$f(x) = f(x_0 + h\,u) = f(x_0) + u\left[\Delta f(x_0) - \frac{1}{2}\Delta^2 f(x_0 - h)\right] + \frac{u^2}{2!}\Delta^2 f(x_0 - h) +$$

$$+ \frac{u(u^2-1)}{3!}\left[\Delta^3 f(x_0 - h) - \frac{1}{2}\Delta^4 f(x_0 - 2h)\right] + \frac{u^2(u^2-1)}{4!}\Delta^4 f(x_0 - 2h) + \ldots$$

y transformando las expresiones contenidas por los corchetes, homogeneizándolas respecto del orden de las diferencias avanzadas que intervienen, de modo que en el primer corchete aparezcan diferencias de primer orden; en el segundo, de segundo orden, etc., y puesto que:

$$\Delta^2 f(x_0 - h) = \Delta f(x_0) - \Delta f(x_0 - h) \quad ; \quad \Delta^4 f(x_0 - 2h) = \Delta^3 f(x_0 - h) - \Delta^3 f(x_0 - 2h)$$

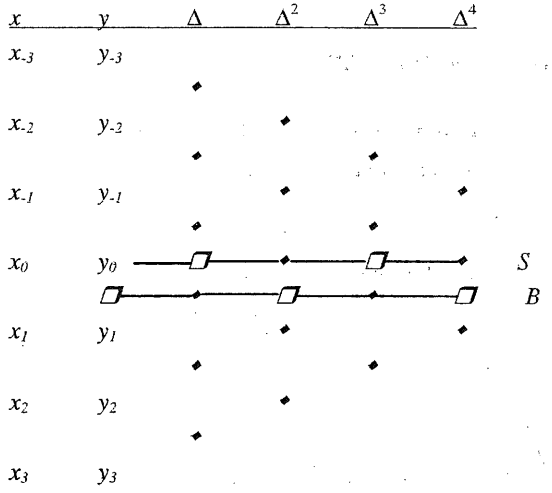
resulta la denominada **FÓRMULA DE INTERPOLACIÓN DE STIRLING**, que es la siguiente:

$$(7.27) \qquad f(x) = f(x_0 + h\,u) = f(x_0) + u\frac{\Delta f(x_0) + \Delta f(x_0 - h)}{2} + \frac{u^2}{2!}\Delta^2 f(x_0 - h) +$$

$$+ \frac{u(u^2-1)}{3!}\frac{\Delta^3 f(x_0 - h) + \Delta^3 f(x_0 - 2h)}{2} + \frac{u^2(u^2-1)}{4!}\Delta^4 f(x_0 - 2h) +$$

$$+ \frac{u^2(u^2-1)(u^2-2^2)}{5!}\frac{\Delta^5 f(x_0 - 2h) + \Delta^5 f(x_0 - 3h)}{2} + \ldots$$

En el diagrama de la figura 7.6, pueden apreciarse cuales son los valores que aparecen en las fórmulas de interpolación de *BESSEL* (B) y *STIRLING* (S), habiéndose marcado mediante rectángulos los promedios de los valores de las diferencias utilizadas en dichas fórmulas.



- Figura 7.6 -

Aprovechando todo lo expuesto en los anteriores ejemplos y haciendo uso de la tabla de diferencias dada, se presenta el siguiente:

Ejemplo 7.8.- Calcular el valor de *e* para $x = 1{,}5708$ usando la fórmula de *STIRLING*.

Solución: En este caso, es:

$$y = 4{,}4817 + {}^1/_2\ 0{,}708\ (0{,}4265 + 0{,}4713) + {}^1/_2\ 0{,}708^2\ .\ 0{,}0448 +$$

$$+ {}^1/_6 0{,}708\ (0{,}708^2 - 1)\ .\ {}^1/_2\ (0{,}0024 + 0{,}0048) +$$

$$+ {}^1/_{24}\ 0{,}708^2\ (0{,}708^2 - 1)\ .\ 0{,}0004 = \underline{\mathbf{4{,}81049}}$$

que, como puede apreciarse, tiene ciertas diferencias respecto de las anteriores fórmulas estudiadas, a partir de la quinta cifra decimal.

♦♦♦ ●●● ♦♦♦

7.3.9.- FÓRMULA DE EVERETT

La **FÓRMULA DE EVERETT** es muy utilizada en algunos campos de las ciencias; por ejemplo estadigrafía, y se caracteriza por utilizar en ella solamente las diferencias tabulares de orden par.

Si en la fórmula de *GAUSS* se hace:

$$(7.28)\qquad f(x) = f(x_0 + h\,u) = \binom{u}{0} f(x_0) + \binom{u}{1} \Delta f(x_0) + \binom{u}{2} \Delta^2 f(x_0 - h) +$$

$$+ \binom{u+1}{3} \Delta^3 f(x_0 - h) + \binom{u+1}{4} \Delta^4 f(x_0 - 2h) + \binom{u+2}{5} \Delta^5 f(x_0 - 2h) + \ldots$$

La sustitución de las diferencias de orden impar por diferencias de orden par se realiza recordando que es:

$$\Delta f(x_0) = f(x_0 + h) - f(x_0) \quad ; \quad \Delta^3 f(x_0 - h) = \Delta^2 f(x_0) - \Delta^2 f(x_0 - h)$$

y así sucesivamente. Resulta, utilizando las propiedades de los números combinatorios:

$$f(x) = f(x_0 + h\,u) = (1-u) f(x_0) + u\, f(x_0 + h) + \binom{u+1}{3} \Delta^2 f(x_0) -$$

$$- \binom{u}{3} \Delta^2 f(x_0 - h) + \binom{u+2}{5} \Delta^4 f(x_0 - h) - \binom{u+1}{5} \Delta^4 f(x_0 - 2h) + \ldots$$

Esta expresión que utiliza términos situados en dos horizontales sucesivas, puede ser escrita de la siguiente manera:

$$f(x) = f(x_0 + h\,u) = (1-u) f(x_0) - \binom{u}{3} \Delta^2 f(x_0 - h) - \binom{u+1}{5} \Delta^4 f(x_0 - 2h) - \ldots +$$

$$- u\, f(x_0 + h) + \binom{u+1}{3} \Delta^2 f(x_0) + \binom{u+2}{5} \Delta^4 f(x_0 - h) + \ldots$$

y reemplazando el valor de 1-*u* por ζ resulta:

$$(7.29)\qquad f(x) = f[x_0 + h(1-\zeta)] = \zeta\, f(x_0) + \binom{\zeta+1}{3} \Delta^2 f(x_0 - h) +$$

$$+ \binom{\zeta+2}{5} \Delta^4 f(x_0 - 2h) + \ldots + u\, f(x_0 + h) + \binom{u+1}{3} \Delta^2 f(x_0) + \binom{u+2}{5} \Delta^4 f(x_0 - h) + \ldots$$

que es la expresión desarrollada de la denominada **FÓRMULA DE INTERPOLACIÓN DE EVERETT**.

Para explicar con mayores detalles la aplicación práctica de la fórmula de *EVERETT*, veamos el siguiente:

Ejemplo 7.9.- Determinar el valor de $y = e^x$ para $x = 1{,}5708$ haciendo uso de la tabla de diferencias de los ejercicios anteriores.

Solución: En este caso es preciso calcular primeramente el valor de x, a partir de u, como sigue:

$$\zeta = 1 - u = 0{,}292$$

resulta entonces:

$$y = 0{,}292 \cdot 4{,}4817 + {}^{1}/_{6}\,[1{,}292 \cdot 0{,}292\,(-0{,}708) \cdot 0{,}0448] +$$

$$+ {}^{1}/_{120}\,[2{,}292 \cdot 1{,}292 \cdot 0{,}292\,(-0{,}708) \cdot 0{,}0004] +$$

$$+ 0{,}708 \cdot 4{,}9530 + {}^{1}/_{6}\,[1{,}708 \cdot 0{,}708\,(-0{,}292) \cdot 0{,}0496] +$$

$$+ {}^{1}/_{120}\,[2{,}708 \cdot 1{,}708 \cdot 0{,}708\,(-0{,}292) \cdot 0{,}0004 = \underline{\mathbf{4{,}81047}}$$

Resultado que coincide con los precisos métodos de *GAUSS* y *BESSEL*, ya estudiados.

••• ●●● •••

7.4.- INTERPOLACIÓN INVERSA

La interpolación resuelve, mediante métodos aproximados y valiéndose de una tabla empírica, el problema de encontrar el valor de $y=f(x)$ cuando está dado el valor de x. Pero, en principio, no resuelve el problema inverso; vale decir, el de determinar el valor de x cuando se conoce la magnitud de $f(x)$.

No obstante, este problema puede resolverse fácilmente, si se intercambia, en la tabla de valores, la variable con la función y viceversa; lo que, equivale a interpolar la función inversa:

$$x = f^{(-1)}\,(y) \tag{7.30}$$

en lugar de hacerlo sobre f. Concretamente, esto significa intercambiar los papeles de las x y las y en cuanto a la calidad de independencia y dependencia de las variables que intervienen.

Como, incluso para valores de x equidistantes, los valores correspondientes de la función no lo son, es esencial utilizar fórmulas de interpolación que se adapten a esta particularidad, tales como *LAGRANGE, Parabólica Progresiva*, etc.

••• ●●● •••

7.5.- AJUSTE

Para estudiar el comportamiento de una serie de datos obtenidos empíricamente, constituidos por puntos dados mediante pares ordenados de números, asociados con los valores de dos variables, resulta imprescindible contar con un procedimiento que permita determinar una función analítica, representada gráficamente por una curva que describa en forma simple, y lo más aproximadamente posible, la relación funcional que pudiera existir entre los datos; vale decir, una función tal, que exprese analíticamente la relación funcional que guardan las variables en cuestión.

7.5.1.- AJUSTE E INTERPOLACIÓN

El problema general estudiado en el párrafo anterior, como tal, no tiene una solución exacta en el campo de la matemática pura, según lo estudiado en la primera parte del presente capítulo, pero, como se ha visto, puede ser resuelto de manera aproximada mediante la aplicación de una variada cantidad de procedimientos numéricos esencialmente diferentes, los que a su vez, admiten ser clasificados en dos grandes grupos: los denominados de **INTERPOLACIÓN** y los llamados de **AJUSTE**.

La metodología de la **INTERPOLACIÓN**, estudiada en sus orígenes por *NEWTON*, se caracteriza por suponer que los datos que intervienen en el problema son exactos; por lo cual, en la construcción de la **FUNCIÓN DE INTERPOLACIÓN** se exige que la misma satisfaga todos y cada uno de los valores particulares que constituyen los datos. Desde el punto de vista geométrico, la **CURVA DE INTERPOLACIÓN** debe pasar por todos los puntos que representan cada par ordenado dado como dato y deben tener, en principio, tantos parámetros indeterminados como puntos se den en los datos. La condición que los datos satisfagan la ecuación permitirá calcular el valor de dichos parámetros.

Una vez determinada la función de interpolación se tiene la posibilidad de calcular nuevos valores de la variable dependiente para otros tantos valores, arbitrarios, de la variable independiente, distintos a aquellos que figuran en la tabla, dados como datos. Consultar el principio del presente capítulo.

La metodología del **AJUSTE MEDIANTE CURVAS** supone, por el contrario que los datos ingresados están afectados en cierto grado de errores debidos al modelado (abstracción) y aleatorios (inherentes); por lo que, no resulta indispensable que la **CURVA DE AJUSTE** correspondiente, pase exactamente por los puntos que representan los datos, sino que, en promedio, la aproximación sea óptima de acuerdo a un cierto y determinado criterio, denominado **CRITERIO DE AJUSTE**.

El iniciador de estos procedimientos fue *GAUSS*, quien desarrolló el tan conocido **MÉTODO DE LOS MÍNIMOS CUADRADOS**. Posteriormente fueron elaborados otros métodos de ajuste, siendo los más utilizados el de los **MOMENTOS** y el de **MÁXIMA VEROSIMILITUD**; este último, creado por *R. A. FISHER*.

7.5.2.- AJUSTE MEDIANTE CURVAS

A continuación se expondrán los principios fundamentales de la metodología del **AJUSTE** de puntos mediante curvas. Según se presentan los datos contenidos en la tabla de datos, puede convenir realizar el ajuste utilizando diferentes curvas, tales como: polinomiales logarítmicas, exponenciales, logísticas, etc. las que serán adoptadas luego de un estudio expeditivo del tipo de relación y de la forma general de la representación gráfica de los datos contenidos en la tabla dada como dato.

Demás esta decir que, no existen reglas generales que permitan de una manera absolutamente racional y sistemática, tomar una decisión categórica respecto a tal o cual elección, quedando la misma, sujeta a los conocimientos previos que se posean del fenómeno en estudio y a la experiencia en este campo de la matemática, de quien tenga que aplicar el procedimiento.

Cada uno de los tipos de curvas mencionados en el párrafo anterior, se adapta especialmente a series de datos particulares, con excepción del **MODELO POLINOMIAL** que puede ser empleado con objetivos muy generales.

Es por la circunstancia apuntada en el párrafo anterior, además de los alcances previstos para la confección de la presente obra, que se limitará el estudio al estudio de modelo polinomial.

7.5.3.- AJUSTE MEDIANTE POLINOMIOS

Como se ha estudiado, en todos aquellos casos en que los datos distan de ser valores exactos de los puntos representativos de una función desconocida; sino que, por el contrario, se suponen afectados de ciertos errores aleatorios y/o sistemáticos (inherentes) en su estimación, se impone el método de **AJUSTE** para su tratamiento. Este ajuste puede realizarse de muchos modos según sea el tipo de función elegida para realizarlo; para describir le relación funcional de los datos del problema en cuestión.

La mayoría de los fenómenos en estudio no poseen una curva que pueda definirse analíticamente ***a priori***, tal como: una exponencial, una logarítmica, etc.; se prefiere utilizar entonces, la **FUNCIÓN POLINOMIAL**, la cual puede ser empleada con muchas ventajas, en la mayoría de los casos, por lo que se la ha elegido para sistematizar el estudio del presente tema.

7.5.3.1.- Metodología del Ajuste

Para lograr el ajuste de una serie de datos, mediante una función polinomial, se propone una expresión de la forma:

$$y = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n \tag{7.31}$$

en la cual figuran $n+1$ parámetros a_i los que, deberán ser deducidos de manera tal que la expresión lograda alcance una aproximación óptima, con respecto a los valores de los datos ingresados, según un cierto criterio.

Para deducir los valores de los parámetros a_i siguiendo el **CRITERIO DE AJUSTE POR MOMENTOS**, el cual tiene como fundamento principal la suposición de que los s primeros momentos de la función son, respectivamente, iguales a los s primeros momentos de la serie empírica dada.

Se debe notar que S representa el número de parámetros que son necesarios calcular para determinar la función buscada. Para ello, será necesario partir de las siguientes igualdades:

(7.32) $$m_1 = \overline{m_1} \quad ; \quad m_2 = \overline{m_2} \quad ;\ldots; \quad m_S = \overline{m_S}$$

siendo m_i los momentos empíricos de la serie o tabla de datos, y $\overline{m_i}$ los momentos de la función que se ajusta, en cuyo caso, si las ordenadas de los datos son $y_1 ; y_2 ;\ldots; y_S$ entonces, la expresión (7.31) toma la forma:

(7.33) $$\overline{y} = a_0 + a_1 x + a_2 x^2 + \ldots + a_S x^S$$

Por consiguiente, la igualdad de momentos expresada por (7.32), lleva implícito que:

(7.34) $$\sum x_i^k \overline{y_i} = \sum \left(a_0 + a_1 x + a_2 x^2 + \ldots + a_S x^S\right)$$

siendo:

(7.35) $$\begin{cases} i = 1;2;\ldots;n \\ k = 0;1;\ldots;S \end{cases}$$

donde, cada uno de los valores que puede tomar k (naturalmente, incluyendo el cero), produce una ecuación, resultando el conjunto de ecuaciones lineales simultáneas:

(7.36) $$\begin{cases} \sum y_i = a_0 n + a_1 \sum x_i + \ldots + a_S \sum x_i^S \\ \sum x y_i = a_0 \sum x_i + a_1 \sum x_i^2 + \ldots + a_S \sum x_i^{S+1} \\ \quad \vdots \\ \sum x_i^S y_i = a_0 \sum x_i^S + a_1 \sum x_i^{S+1} + \ldots + a_S \sum x_i^{2S+1} \end{cases}$$

que es un sistema de $S+1$ ecuaciones lineales simultáneas, con un número igual de incógnitas $a_1 ; a_2 ; \ldots ; a_S$.

Resolviendo el sistema dado se obtienen, en forma directa, los coeficientes de la expresión (7.33) de grado S que resuelven el problema.

7.5.3.2.- Mecánica del Cálculo

Si se tiene cierto conocimiento previo sobre la naturaleza matemática del fenómeno en estudio, la precisión que se desea o, se debe alcanzar con la curva de ajuste y las demás condiciones del problema, pueden establecerse de primera intensión los parámetros del procedimiento.

En el caso contrario al presentado, que es el que usualmente se presenta en la práctica diaria, el cálculo debe iniciarse fijando el grado del polinomio de ajuste, dado por la expresión (7.33) en 1 (uno); vale decir, haciendo $S = 1$. Luego de ser de aplicada la metodología descripta en el punto anterior, se obtendrá como resultado una curva de ajuste de primer grado; es decir, una recta.

Es necesario destacar, llegado a este punto del procedimiento, la necesidad de verificación de la precisión con que la curva obtenida se ajusta a los datos experimentales ingresados. Esto puede realizarse mediante la evaluación de la bondad del ajuste por medio de dos parámetros estadísticos fundamentales: el **COEFICIENTE DE CORRELACIÓN** y la **VARIACIÓN NO EXPLICADA**. El significado estadístico de los parámetros mencionados excede los objetivos del presente trabajo; no obstante, para quien quiera profundizar en el conocimiento de los mismos, deberá remitirse a algún libro que trate sobre el Cálculo de Probabilidades o Estadística.

En este orden de cosas, el Coeficiente de Correlación puede ser calculado mediante la expresión:

(7.37) $$\sqrt{\frac{\sum \left(\overline{y} - y_{est}\right)^2}{\sum \left(y - \overline{y}\right)^2}}$$

también:

(7.38) $$\sum \left(y - y_{est}\right)^2$$

donde y_{est} es el valor estimado de la variable.

Si la aproximación obtenida, calculada por las expresiones (7.37) y (7.38) fuera satisfactoria, pueden aceptarse los resultados y se detiene el proceso. En caso contrario, existe la posibilidad de incrementar el grado del polinomio de ajuste en una unidad; vale decir, hacer $S = 2$, e iterar para el nuevo polinomio de ajuste los últimos pasos descriptos.

◆◆◆ ●●● ◆◆◆

Capítulo 8

Derivación e Integración

8 . - DERIVACIÓN E INTEGRACIÓN

Diferenciación Numérica. Derivación mediante Fórmulas de Interpolación. Algoritmo de Horner. Introducción. Cálculo del Valor del Polinomio. Cálculo de las Derivadas Sucesivas. Método de la Serie de Taylor. Integración Numérica. Introducción. Fórmula de los Trapecios. Fórmula de Simpson. Regla de los Tres Octavos de Simpson. Fórmula de Euler-Mac Laurin. Fórmula de Gregory.

8.1.- DIFERENCIACIÓN NUMÉRICA

Se ha estudiado en el capítulo anterior, la manera de utilizar distintas expresiones del polinomio de interpolación para hallar valores aproximados de una función $f(x)$ en puntos donde no es posible conocer el valor exacto de la función f. Otra forma de uso del polinomio de interpolación, de la misma o probablemente mayor importancia práctica, es la de aplicarlo a la resolución de problemas de diferenciación, en general y el cálculo de derivadas, en particular.

En este tipo de aplicaciones, la idea básica es sumamente simple: en lugar de efectuar la operación sobre la función f (lo que puede resultar extremadamente complicado, o incluso, desconocerse una forma analítica de su expresión), se realiza sobre un polinomio de interpolación adecuado, determinado ex-profeso.

Dicho en otras palabras, dada una tabla de valores de la función $f(x)$ para diversos valores de x, siempre se puede determinar el polinomio de interpolación que, satisfaciendo a los valores dados, represente con cierto grado de aproximación a $f(x)$. Por consiguiente, siempre es posible calcular, de manera más o menos precisa, la derivada $f'(x)$, de la función en cuestión.

Con el objeto de no repetir siempre los mismos pasos, en lugar de determinar en cada caso el polinomio de interpolación, se puede calcular, en general y por única vez, las derivadas sucesivas de la fórmula de interpolación y aplicar luego los resultados así obtenidos, a cada caso en particular.

8.1.1.- DERIVACIÓN MEDIANTE FÓRMULAS DE INTERPOLACIÓN

La metodología descripta en la introducción que, como se verá no es única, implica el uso de cualquiera de las fórmulas de interpolación estudiadas en el capítulo correspondiente. A continuación se estudiará un caso específico con los ejemplos correspondientes.

Así, en la fórmula de *NEWTON-GREGORY Ascendente*, en la cual se ha hecho la transformación $x=x_0+hu$, para facilitar su uso:

$$f(x)=f(x_0+hu)=f(x_0)+\Delta f(x_0)u+\Delta^2 f(x_0)\frac{u(u-1)}{2!}+ \tag{8.1}$$

$$+\Delta^3 f(x_0)\frac{u(u-1)(u-2)}{3!}+\dots$$

y derivando respecto de la variable u, se obtiene:

$$hf'(x_0+hu)=\Delta f(x_0)+\Delta^2 f(x_0)\frac{2u-1}{2}+\Delta^3 f(x_0)\frac{3u^2-6u+2}{6}+\dots$$

y para $x=x_0$; vale decir, para $u=0$, resulta la ecuación:

$$hf'(x_0)=\Delta f(x_0)-\frac{1}{2}\Delta^2 f(x_0)+\frac{1}{3}\Delta^3 f(x_0)-\frac{1}{4}\Delta^4 f(x_0)+\dots \tag{8.2}$$

Análogamente, para la derivada segunda se obtiene la expresión:

$$h^2 f''(x_0+hu)=\Delta^2 f(x_0)+\Delta^3 f(x_0)(u-1)+\Delta^4 f(x_0)\frac{6u^2-18u+11}{12}+\dots$$

y para $x=x_0$; o sea, haciendo $u=0$, resulta la ecuación:

$$h^2 f''(x_0)=\Delta^2 f(x_0)-\Delta^3 f(x_0)+\frac{11}{12}\Delta^4 f(x_0)-\dots \tag{8.3}$$

Este procedimiento puede ser iterado tantas veces como se necesite, para obtener derivadas de mayor orden.

Se aportaran más detalles de la aplicación práctica de las fórmulas estudiadas a través del siguiente:

Ejemplo 8.1.- Calcular las derivadas de primero y segundo orden, correspondientes a la función a continuación tabulada, en el punto $x_0=5100$.

x	$f(x)$	Δ	Δ^2	Δ^3	Δ^4
5100	8,5369958187				
		0,0019588645			
5110	8,5389546832		-0,0000038297		
		19550348		0,0000000151	
5120	8,5409097180		- 38146		- 5
		19512202		146	
5130	8,5428609382		- 38000		
		19474202			
5140	8,5448083584				

Solución: Mediante la aplicación directa de las fórmulas estudiadas y sabiendo que $h=10$, resulta:

$$10f'(5100) = 0,0019588645 + 19148 + 50 = 0,0019607843$$

$$;\qquad 100f''(5100) = -0,0000038297 - 151 - 5 = -0,0000038453$$

◆◆◆ ●●● ◆◆◆

Si se parte de la fórmula de *NEWTON-GREGORY Descendente* o, de las de *GAUSS, LAGRANGE, BESSEL*, etc., se encontrarán, análogamente, nuevas fórmulas de derivación para cada caso en particular; las cuales, ofrecerán mayor o menor precisión según la posición relativa del valor de la variable para el cual se desea calcular las derivadas.

Así, la aplicación de idéntico criterio para la fórmula de *NEWTON-GREGORY Descendente* estudiada, la que está dada por la expresión:

$$f(x_n + hu) = f(x_n) + u\nabla f(x_n) + \frac{u(u+1)}{2!}\nabla^2 f(x_n) + \frac{u(u+1)(u+2)}{3!}\nabla^3 f(x_n) + \ldots$$

da como resultado, derivando con respecto a u e igualando a cero su valor:

$$h f'(x_n) = \nabla f(x_n) + \frac{1}{2}\nabla^2 f(x_n) + \frac{1}{3}\nabla^3 f(x_n) + \ldots \tag{8.4}$$

como así también:

$$h^2 f''(x_n) = \nabla^2 f(x_n) + \nabla^3 f(x_n) + \frac{11}{12}\nabla^4 f(x_n) + \ldots \tag{8.5}$$

Tarea que es posible continuar, de la misma manera a la realizada, para poder así, obtener las derivadas de orden superior al segundo.

8.1.2.- ALGORITMO DE HORNER

Otro método idóneo para determinar el valor de las derivadas sucesivas, en el especial caso de que la función en cuestión sea una expresión algebraica, es el denominado **MÉTODO** o **ALGORITMO DE HORNER**, el cual se estudiará a continuación.

Este método, tiene la ventaja adicional que permite calcular el valor de la función para el valor de la variable en el cual se pretende determinar sus derivadas y, además, la de ser de muy sencilla aplicación:

8.1.2.1.- Introducción

En ocasiones es necesario determinar el valor numérico de una función algebraica polinómica, de la forma:

$$P(x) = a_0 x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_{n-1} x + a_n \tag{8.6}$$

ya sea, solo con el objeto mencionado; o bien, como parte de un algoritmo más complejo, en el punto $x=\alpha$; vale decir, puede tratarse de calcular simplemente el valor de $P(\alpha)$, para valores de α reales o complejos; o puede tratarse de calcular, además del valor del polinomio, sus derivadas sucesivas.

Primeramente será estudiado el caso de tener que calcular el valor de la función. Si bien es cierto que, todos los desarrollos que siguen tienen validez tanto para valores de α reales, como complejos, se entenderá que el número de operaciones a efectuarse solo es válido en el campo real.

Usando directamente su valor reemplazando el de x por el de α, resulta:

$$P(\alpha) = a_0 \alpha^n + a_1 \alpha^{n-1} + \ldots + a_{n-1}\alpha + a_n \tag{8.7}$$

en cuyo caso, se deben calcular α^2 ; α^3 ;...; α^n ; es decir, n-1 multiplicaciones. Después deben ser formados los productos $a_0 \alpha^n$; $a_1 \alpha^{n-1}$;...; $a_{n-1}\alpha$; o sea, n multiplicaciones más. Contabilizando todo se deben realizar $2n$-1 multiplicaciones y n adiciones para calcular el valor de $P(\alpha)$ a partir de la expresión (8.6), haciendo uso directo de la fórmula (8.7); vale decir, un total de $3n - 1$ operaciones.

Una alternativa dada por *HORNER*, permite disminuir prácticamente a la mitad el número de operaciones, con la posibilidad adicional de calcular también sus derivadas sucesivas en el mismo punto α: $P'(\alpha)$; $P''(\alpha)$;...; $P^{(n)}(\alpha)$, lo que será estudiado a continuación.

8.1.2.2.- Cálculo del Valor del Polinomio

Una manera recurrente de escribir la expresión (8.7) está dada por la siguiente expresión:

$$P(\alpha) = (\ldots(((a_0\alpha + a_1)\alpha + a_2)\alpha + a_3)\alpha + \ldots + a_{n-1})\alpha + a_n \tag{8.8}$$

que puede obtenerse directamente haciendo uso del algoritmo:

$$\begin{aligned} b_0 &= a_0 \\ b_1 &= b_0\alpha + a_1 \\ &\cdots\cdots\cdots \\ b_{n-1} &= b_{n-2}\alpha + a_{n-1} \\ P(\alpha) = b_n &= b_{n-1}\alpha + a_n \end{aligned}$$

donde, es importante notar que, el valor de b_n es directamente el de $P(x)$ cuando $x=\alpha$; vale decir, el de $P(\alpha)$. De este último algoritmo, es posible deducir fácilmente que, en total son necesarias solamente n multiplicaciones y n adiciones, para lograr el resultado anterior.

8.1.2.3.- Cálculo de las Derivadas Sucesivas

Para calcular las derivadas sucesivas de $P(x)$ en el punto $x=\alpha$, es necesario utilizar un simple artificio. Se sabe que la expresión (8.6) se puede escribir bajo la forma:

$$P(x) = (x-\alpha)\,[b_0\,x^{n-1} + b_1\,x^{n-2} + \ldots + b_{n-2}\,x + b_{n-1}] + b_n \tag{8.9}$$

dado que se puede reconstruir la (8.6) a partir de la (8.9); en efecto, realizando las operaciones indicadas, resulta:

$$\begin{aligned} P(x) &= b_0\,x^n - \alpha\,b_0\,x^{n-1} + b_1\,x^{n-1} - \alpha\,b_1\,x^{n-2} + \ldots + \\ &\quad + b_{n-2}\,x^2 - \alpha\,b_{n-2}\,x + b_{n-1}\,x - \alpha\,b_{n-1} + b_n = \\ &= b_0\,x^n + (b_1 - \alpha\,b_0)\,x^{n-1} + (b_2 - \alpha\,b_1)\,x_{n-2} + \ldots + \\ &\quad + (b_{n-1} - \alpha\,b_{n-2})\,x + (b_n - \alpha\,b_{n-1}) \end{aligned}$$

Comparando esta última expresión con la (8.6) e igualando sus coeficientes homólogos se pueden despejar los b_j de la siguiente manera:

$$\begin{array}{lcl} a_0 = b_0 & \Rightarrow & b_0 = a_0 \\ a_1 = b_1 - \alpha\,b_0 & \Rightarrow & b_1 = a_1 + \alpha\,b_0 \\ \ldots\ldots\ldots & & \ldots\ldots\ldots \\ a_{n-1} = b_{n-1} - \alpha\,b_{n-2} & \Rightarrow & b_{n-1} = a_{n-1} + \alpha\,b_{n-2} \\ a_n = b_n - \alpha\,b_{n-1} & \Rightarrow & b_n = a_n + \alpha\,b_{n-1} \end{array}$$

como se quería demostrar.

Llamando $Q_1(x)$ al polinomio encerrado dentro del corchete de la expresión (8.9), ésta puede escribirse de la siguiente forma:

$$P(x) = (x - \alpha)\,Q_1(x) + b_n \tag{8.10}$$

donde el resto b_n es constante e igual a $P(\alpha)$, como ya se ha visto.

Derivando respecto de x la expresión (8.10), resulta:

$$P'(x) = Q_1(x) + (x - a)\,Q_1'(x) \tag{8.11}$$

de la que es posible inferir que:

$$P'(\alpha) = Q_1(\alpha) \tag{8.12}$$

expresión, esta última, que puede ser escrita en forma recurrente y similar a la (8.8), de la manera siguiente:

$$P'(\alpha) = (\ldots(((b_0\,\alpha + b_1)\alpha + b_2)\alpha + b_3)\alpha + \ldots + b_{n-1})\alpha + b_{n-1}$$

cuyo algoritmo resolutorio, se sabe que es:

$$\begin{array}{l} c_0 = b_0 \\ c_1 = b_1 + \alpha\,c_0 \\ \ldots\ldots\ldots \\ P'(\alpha) = c_{n-1} = b_{n-1} + \alpha\,c_{n-2} \end{array}$$

Para determinar las siguientes derivadas sucesivas del polinomio $P(x)$ en el punto $x=\alpha$, es necesario hacer las siguientes consideraciones. Escribiendo el polinomio $P(x)$ en términos de la fórmula del desarrollo en serie de Taylor, resulta:

$$P(x) = P(\alpha) + (x-\alpha)P'(\alpha) + \frac{(x-\alpha)^2}{2!}P''(\alpha) + \ldots + \frac{(x-\alpha)^n}{n!}P^{(n)}(\alpha)$$

Por otra parte, de la (8.10) y considerando que $b_n=P(\alpha)$, se puede escribir que:

$$Q_1(x) = \frac{P(x) - P(\alpha)}{x - \alpha}$$

y teniendo en consideración estas dos últimas expresiones, resulta:

$$Q_1(x) = P'(\alpha) + \frac{x-\alpha}{2!}P''(\alpha) + \frac{(x-\alpha)^2}{3!}P'''(\alpha) + \ldots + \frac{(x-\alpha)^{n-1}}{n!}P^{(n)}(\alpha)$$

sacando factor común x-α, de aquellos términos del segundo miembro que lo contienen, se obtiene:

$$Q_1(x) = P'(\alpha) + (x-\alpha)\left[\frac{P''(\alpha)}{2!} + \frac{x-\alpha}{3!}P'''(\alpha) + \ldots + \frac{(x-\alpha)^{n-2}}{n!}P^{(n)}(\alpha)\right]$$

donde, finalmente, llamando $Q_2(x)$ al polinomio encerrado dentro del corchete, resulta:

$$Q_1(x) = P'(\alpha) + (x - \alpha)\, Q_2(x)$$

Derivando esta última expresión se obtiene:

$$Q_1'(x) = Q_2(x) + (x - \alpha)\, Q_2'(x)$$

que, en el punto x=α, vale:

(8.13) $$Q_1'(x) = Q_2(x)$$

Por otra parte, derivando la expresión (8.11) respecto de x, se obtiene:

$$P''(x) = 2\, Q_1'(x) + (x - a)\, Q_1''(x)$$

la cual, a su vez, en el punto x=α, toma el valor:

(8.14) $$P''(\alpha) = 2\, Q_1'(\alpha)$$

Considerando las expresiones (8.13) y (8.14), se obtiene en definitiva:

(8.15) $$P''(\alpha) = 2\, Q_2(\alpha)$$

Con mayor grado de generalización, es posible calcular las siguientes derivadas sucesivas de $P(x)$, en el punto x=α, de la manera que se indica a continuación. Iterando el procedimiento descripto anteriormente, se puede escribir, en general, lo siguiente:

(8.16) $$Q_{k-1}(x) = \frac{P^{(k-1)}(\alpha)}{(k-1)!} + (x-\alpha)Q_k(x)$$

de donde:

(8.17) $$Q_k(x) = \frac{P_k(\alpha)}{k!} + \frac{(x-\alpha)^{k+1}}{(k+1)!}P^{(k+1)}(\alpha) + \ldots + \frac{(x-\alpha)^{n-k}}{n!}P^{(n)}(\alpha)$$

Derivando la (8.16) con respecto a x, se deduce que:

$$Q_{k-1}'(x) = Q_k(x) + (x - \alpha)\, Q_k'(x)$$

en la cual, haciendo x=α, resulta:

$$Q_{k-1}'(\alpha) = Q_k(\alpha)$$

Considerando la expresión (8.17), se obtiene:

$$Q_k(\alpha) = \frac{P^{(k)}(\alpha)}{k!}$$

de donde, es posible deducir que, en general, la derivada de orden *k-ésimo* resulta:

(8.18) $$P^{(k)}(\alpha) = k!\, Q_k(\alpha)$$

que representa la expresión general de la derivada de orden k, de un polinomio $P(x)$, en el punto x=α.

Para poder apreciar la operatoria del método de *HORNER*, veamos el siguiente:

Ejemplo 8.2.- Se trata de determinar el valor del polinomio y de sus derivadas sucesivas, en el punto x=2, siendo:

$$P(x) = 3x^4 - 7x^3 + 4x^2 - 8x + 6$$

Solución: Actuando sistemáticamente, se tiene:

a) Cálculo de $P(2)$:

$b_0 = 3$
$b_1 = 2 \,.\, 3 - 7 = -1$
$b_2 = 2 \,.\, (-1) + 4 = 2$
$b_3 = 2 \,.\, 2 - 8 = -4$
$P(2) = b_4 = 2 \,.\, (-4) + 6 = -2$

b) Cálculo de $P'(2)$: $c_0 = 3$

$c_1 = 2 \,.\, 3 - 1 = 5$

$c_2 = 2 \,.\, 5 + 2 = 12$

$P'(2) = c_3 = 2 \,.\, 12 - 4 = 20$

c) Cálculo de $P''(2)$: $d_0 = 3$

$d_1 = 2 \,.\, 3 + 5 = 11$

$d_2 = 2 \,.\, 11 + 12 = 34$

$P''(2) = 2!\, d_3 = 2 \,.\, 34 = 78$

d) Cálculo de $P'''(2)$: $e_0 = 3$

$e_1 = 2 \,.\, 3 + 11 = 17$

$P'''(2) = 3!\, e_2 = 6 \,.\, 17 = 102$

Dada la simplicidad de las expresiones involucradas, estos resultados pueden ser verificados fácil y rápidamente en forma directa.

♦♦♦ ●●● ♦♦♦

8.1.3.- EL MÉTODO DE LA SERIE DE TAYLOR

Otra manera, de muy simple aplicación, destinada a realizar la diferenciación numérica, es el conocido **MÉTODO DE LA SERIE DE TAYLOR**, el cual será estudiado a continuación.

La serie de *TAYLOR* para una función $y = f(x)$ desarrollada respecto del punto $x_i + h$, es:

$$y(x_i + h) = y_i + h\, y_i' + \frac{h^2}{2!} y_i'' + \frac{h^3}{3!} y_i''' + \dots \tag{8.19}$$

donde, y_i representa la ordenada que corresponde a x_i; mientras que, $x_i + h$ se encuentra dentro del entorno de convergencia de la función.

De la misma manera, la expresión de la fórmula de *TAYLOR* desarrollada en el punto $x_i - h$ está dada por:

$$y(x_i - h) = y_i - h\, y_i' + \frac{h^2}{2!} y_i'' - \frac{h^3}{3!} y_i''' + \dots \tag{8.20}$$

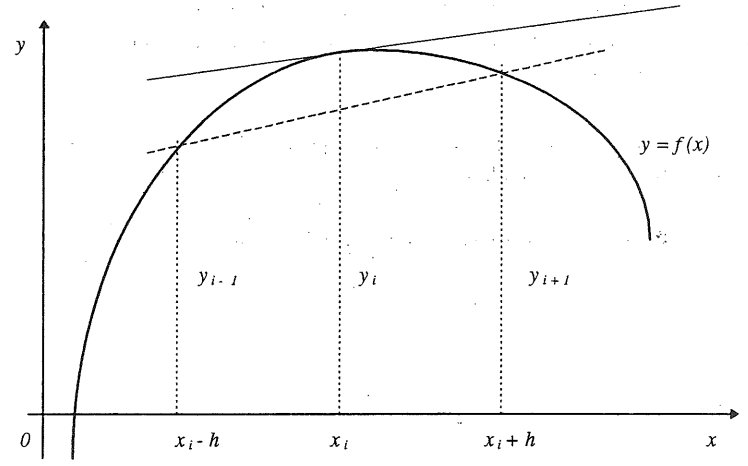
Utilizando solamente los tres primeros términos de los desarrollos en serie de *TAYLOR* dados por (8.19) y (8.20), es posible obtener una fórmula que calcule el valor de y' en el punto $x = x_i$ con solo restar la (8.20) de la (8.19), con lo cual se obtiene:

$$y_i' = \frac{y(x_i + h) - y(x_i - h)}{2h} \tag{8.21}$$

Observando la figura 8.1, es posible apreciar que designando los puntos, uniformemente esparcidos a distancias $-h$ y h del valor de x_i, con $x_i - h$ y $x_i + h$, respectivamente y, además, se identifican las ordenadas correspondientes como y_{i-1} e y_{i+1}, la ecuación (8.20) puede ser escrita en la forma:

$$y_i' = \frac{y_{i+1} - y_{i-1}}{2h} \tag{8.22}$$

Representando gráficamente los distintos elementos que comprende la situación descripta, resulta la siguiente figura:



- Figura 8.1 -

La ecuación (8.22) recibe el nombre de *PRIMERA APROXIMACIÓN POR DIFERENCIAS CENTRALES DE LA PRIMERA DERIVADA* y'. Esta aproximación representa, gráficamente, la pendiente de la recta discontinua (secante), mostrada en la figura 8.1. El valor verdadero de la derivada viene dado por la pendiente de la línea continua (tangente) a la curva $y = f(x)$.

Si, por el contrario, ahora son sumadas las expresiones (8.19) y (8.20), y haciendo uso de la misma notación, es posible escribir la siguiente expresión para la segunda derivada:

(8.23) $$y_i'' = \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2}$$

la cual es denominada *PRIMERA APROXIMACIÓN POR DIFERENCIAS CENTRALES DE LA SEGUNDA DERIVADA* y''.

Para obtener las expresiones correspondientes a la tercera, cuarta, etc. derivadas, consultar su deducción en el capítulo número 11: *Integración de Ecuaciones Diferenciales. Problemas Con Valores en la Frontera.*

La misma técnica puede ser utilizada para hallar expresiones más precisas de las derivadas con solo utilizar un mayor número de términos en los desarrollos en serie de *TAYLOR*. Sin embargo, la obtención de expresiones más precisas, especialmente para derivadas de orden superior al segundo, se torna muy laboriosa debido al gran número de ecuaciones simultáneas que deben ser resueltas.

Con el fin de simplificar la tarea y poseer elementos de control, se expone un resumen de las expresiones estudiadas:

- *Primeras Aproximaciones de las Diferencias Centrales:*

$$y_i' = \frac{y_{i+1} - y_{i-1}}{2h}$$

$$y_i'' = \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2}$$

$$y_i''' = \frac{y_{i+2} - 2y_{i+1} + 2y_{i-1} - y_{i-2}}{2h^3}$$

$$y_i^{IV} = \frac{y_{i+2} - 4y_{i+1} + 6y_i - 4y_{i-1} + y_{i-2}}{h^4}$$

- *Segundas Aproximaciones de las Diferencias Centrales:*

$$y_i' = \frac{-y_{i+2} + 8y_{i+1} - 8y_{i-1} + y_{i-2}}{12h}$$

$$y_i'' = \frac{-y_{i+2} + 16y_{i+1} - 30y_i + 16y_{i-1} - y_{i-2}}{12h^2}$$

$$y_i''' = \frac{-y_{i-3} + 8y_{i+2} - 13y_{i+1} + 13y_{i-1} - 8y_{i-2} + y_{i-3}}{8h^3}$$

$$y_i^{IV} = \frac{-y_{i-3} + 12y_{i+2} - 39y_{i+1} + 56y_i - 39y_{i-1} + 12y_{i-2} - yi - 3}{6h^4}$$

Como puede apreciarse, las expresiones por diferencias centrales para las diversas derivadas, estudiadas, hacen uso de valores de la función a ambos lados de la variable x en el que se desea conocer la derivada en cuestión.

Utilizando desarrollos en serie de *TAYLOR* convenientes, se pueden obtener expresiones para las mismas derivadas, completamente definidas en términos de valores de la función en x_i y puntos a la derecha del mismo. Estas se conocen como expresiones de *DIFERENCIAS FINITAS HACIA ADELANTE*. En forma similar, se pueden obtener expresiones para las derivadas que estén totalmente en términos de valores de la función en x_i y puntos a la izquierda de él. Estas reciben el nombre de expresiones de *DIFERENCIAS FINITAS HACIA ATRAS*.

En la diferenciación numérica, las expresiones de diferencias hacia adelante se utilizan cuando no se disponen de datos a la izquierda del punto en que se desea calcular la derivada, y las expresiones de diferencias hacia atrás, se utilizan cuando no se disponen de datos a la derecha del punto considerado.

No se expondrá una síntesis del cálculo de las fórmulas presentadas dado que, excede los objetivos propuestos en el presente trabajo.

♦♦♦ ●●● ♦♦♦

8.2.- INTEGRACIÓN NUMÉRICA

Es bien sabido que, dentro del campo analítico, perteneciente a la matemática pura, se desconoce una primitiva de la mayor parte de las funciones que ella estudia; o si esta se conoce, en muchos casos su aplicación es larga y compleja, para utilizarla con provecho en la resolución de una integral.

En muchas ocasiones, incluso, es posible que se desconozca la expresión analítica de la función sobre la cual se desea integrar, y solo se cuente con una tabla empírica, obtenida de la observación en el terreno o, del trabajo en un laboratorio. Consecuentemente, y en términos generales, es posible asegurar que la gran mayoría de los problemas que se presentan en la práctica, carecen de solución dentro del campo analítico; vale decir, de la matemática pura.

Resumiendo, la imposibilidad, en algunas ocasiones, o la inconveniencia, en otras, de la aplicación de métodos tradicionales está dada, fundamentalmente, por las siguientes causas:

I.- Que no se conozca ninguna primitiva de aquella función que es necesario integrar,

II.- Que aún conociéndose una función primitiva, su aplicación resulte excesivamente compleja o extensa su aplicación,

III.- Que, directamente, se desconozca la expresión analítica de la función que debe ser integrada.

Todo lo dicho, es el fundamento para el estudio de métodos no tradicionales más o menos precisos, fundamentalmente numéricos, que permitan el cálculo de integrales definidas aún en aquellos casos en que se desconoce la función a integrar; o bien, aunque se la conozca, no existe una primitiva de ella conocida; o, por último, se conoce tanto a la función como a una primitiva de la misma, pero, no resulta muy práctica su aplicación, por ser esta, compleja y/o extensa.

8.2.1.- INTRODUCCIÓN

Cuando el problema en cuestión consiste en calcular la integral definida de una determinada función $f(x)$, dada por:

$$I = \int_a^b f(x)\,dx \tag{8.24}$$

y se conoce una función $F(x)$, primitiva de $f(x)$; es decir, tal que $F'(x) = f(x)$, se aplica la denominada regla de *BARROW* que expresa:

$$I = \int_{x=a}^{x=b} f(x)\,dx = F(b) - F(a) \tag{8.25}$$

Pero, cuando no se conoce ninguna primitiva de la función dada, lo cual ocurre con suma frecuencia, resulta necesario apelar a métodos de cálculo aproximados. Igual proceder debe adoptarse si, aún conociéndose una primitiva, resulta poco práctico aplicarla, por su complejidad.

Lo dicho hasta ahora se refiere al caso en que se conoce la función que debe ser integrada, cosa que no siempre sucede. En ocasiones se cuenta solamente con una tabla de alguno de sus valores, proveniente de resultados experimentales; en cuyo caso, tampoco es posible aplicar la regla de *BARROW*.

Considerando que la expresión de la integral dada por (8.24) equivale a determinar el valor del área comprendida entre el eje de las x, la curva representativa de la función $f(x)$ y las ordenadas extremas $x=a$; $x=b$, es posible desarrollar diversos métodos aproximados para lograr dicho objetivo; métodos que serán estudiados a continuación.

8.2.2.- FÓRMULA DE LOS TRAPECIOS

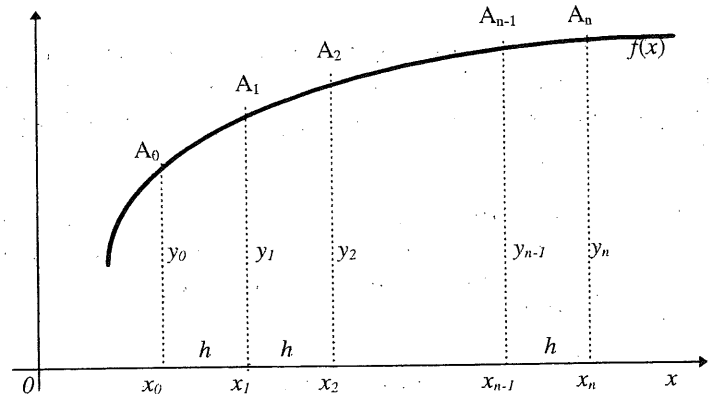
Supónganse conocidos los $n+1$ valores x_0 ; x_1 ;...; x_n deducidos a partir del procesamiento de la función $f(x)$, conocida, o por los resultados de alguna prueba de laboratorio, indistintamente, que cumplen con la condición:

$$x_k - x_{k-1} = h \qquad para \qquad k = 1; 2; \ldots ; n$$

Una primera aproximación al valor del área que se debe calcular, limitada por los puntos x_0 ; A_0 ; A_1 ; ... ; A_n ; x_n se obtiene considerando la sumatoria de las áreas que poseen los trapecios inscriptos en cada una de las superficies parciales limitadas por los siguientes conjuntos de puntos:

x_0 ; A_0 ; A_1 ; x_1
x_1 ; A_1 ; A_2 ; x_2
. . . . . .
x_{n-1} ; A_{n-1} ; A_n ; x_n

y, considerándolas como trapecios, ver figura 8.2.



- Figura 8.2 -

En consecuencia, resulta:

Área $(x_0 ; A_0 ; A_1 ; x_1) \approx {}^1/_2 (y_0 + y_1)\, h$
Área $(x_1 ; A_1 ; A_2 ; x_2) \approx {}^1/_2 (y_1 + y_2)\, h$
. . . . . .
Área $(x_{n-1} ; A_{n-1} ; A_n ; x_n) \approx {}^1/_2 (y_{n-1} + y_n)\, h$

Sumando las expresiones de las áreas así obtenidas, resulta:

$$(8.26) \qquad \int_{x_0}^{x_1} f(x)\,dx \cong \frac{h}{2}\left(y_0 + 2y_1 + \ldots + 2y_{n-1} + y_n\right)$$

En esta última expresión, si se denomina **E** a la suma de las ordenadas extremas; **P** e **I** a las sumas de las ordenadas de subíndices pares e impares, respectivamente, se obtiene la conocida **FORMULA DE LOS TRAPECIOS**:

$$(8.27) \qquad \int_{x_0}^{x_1} f(x)\,dx \cong h\left(\frac{\mathbf{E}}{2} + \mathbf{P} + \mathbf{I}\right)$$

La fórmula de los trapecios tiene ~~una~~ precisión suficientemente buena cuando se trata de aplicarla a determinaciones que no requieran una aproximación de orden elevado. En el caso de haberse sustituido la curva, dada por la función continua $f(x)$, mediante la poligonal inscripta, descripta mediante los puntos dados o calculados, el modelo realizado puede clasificarse como una *Discretización*; y no satisface plenamente cuando se trata de obtener una precisión muy elevada.

Para fijar conceptos y ver la metodología práctica de la fórmula de los *TRAPECIOS*, se expone el siguiente:

Ejemplo 8.3.- Sea determinar el área comprendida entre x_0 =1,27 y x_n =7,87, el eje de las abscisas, y la función dada por la siguiente tabla:

x	1,27	2,59	3,91	5,23	6,55	7,87
y	4,08	4,74	5,19	5,35	5,44	5,38

Solución: Se puede apreciar que, h=1,32; siendo los demás valores que intervienen en la fórmula de los trapecios:

E = 4,08 +5,38 =9,46
P = 4,74 + 5,35 = 10,09
I = 5,19 + 5,44 = 10,63

y, reemplazando estos valores en la **FORMULA DE LOS TRAPECIOS**, dada por la expresión (8.27), resulta:

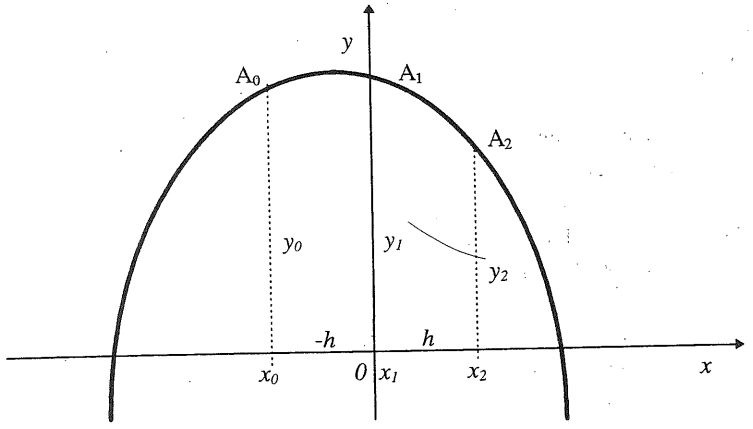
$$\text{Área} = h\,(\mathbf{E}/_2 + \mathbf{P} + \mathbf{I}) = 1{,}32\,({}^{9{,}46}/_2 + 10{,}09 + 10{,}63) = \underline{\mathbf{33{,}59}}$$

♦♦♦ ●●● ♦♦♦

8.2.3.- FÓRMULA DE SIMPSON

Utilizando segmentos de parábola para aproximar los arcos de curva, en lugar de emplear segmentos de recta, como en el método de los trapecios; vale decir, utilizando curvas en lugar de una poligonal, se obtiene una mayor precisión en el cálculo de las integrales definidas.

Primeramente se considerará el caso de la parábola de segundo grado, el que conducirá a deducir la expresión analítica de la fórmula de **SIMPSON**.



- Figura 8.3 -

El primer paso consiste en determinar el área comprendida entre el eje de las x, la parábola de eje vertical que pasa por los tres primeros puntos dados y sus ordenadas extremas. Llamando A_0 ; A_1 ; A_2 a los puntos mencionados y suponiendo que tienen abscisas equidistantes; es decir, que:

$$x_1 - x_0 = x_2 - x_1 = h$$

considerando, además que, haciendo pasar el eje y por el punto intermedio A_1 no se pierde generalidad (ver figura 8.3).

Dadas estas condiciones y teniendo en cuenta que, en general, la parábola de segundo grado es:

$$y = a x^2 + b x + c$$

pero, como debe pasar por los tres puntos A_0 ; A_1 ; A_2 , es posible escribir:

$$\begin{aligned} y_0 &= a x_0^2 + b x_0 + c = \\ &= a (-h)^2 + b (-h) + c = a h^2 - b h + c \\ y_1 &= a x_1^2 + b x_1 + c = c \\ y_2 &= a x_2^2 + b x_2 + c = \\ &= a h^2 + b h + c = a h^2 + b h + c \end{aligned}$$

sumando y restando la primera y la última de estas expresiones, y directamente de la segunda, se obtienen los siguientes valores:

$$a = \frac{y_2 - 2y_1 + y_0}{2h^2} \quad ; \quad b = \frac{y_2 - y_0}{2h} \quad ; \quad c = y_1$$

valores que serán empleados durante el cálculo, para ser reemplazados en la expresión de la integral.

Por otra parte, del análisis matemático sabemos que la expresión analítica del área buscada vale:

$$I = \int_{-h}^{h} y\,dx = \int_{-h}^{h} \left(a x^2 + b x + c\right) dx = \left[a \frac{x^3}{3} + b \frac{x^2}{2} + c x \right]_{-h}^{h} = \frac{2}{3} a h^3 + 2 c h$$

Reemplazando en esta última los valores de a y c anteriormente obtenidos, resulta:

$$I = \frac{2}{3} \frac{y_2 - 2 y_1 + y_0}{2h^2} h^3 + 2 y_1 h = \frac{h}{3} \left(y_2 - 2y_1 + y_0 + 6y_1\right) = \frac{h}{3} \left(y_0 + 4y_1 + y_2\right)$$

Así como dos ordenadas son suficientes para determinar el área que limita un segmento de recta (caso de los trapecios), el conocimiento de tres ordenadas es suficiente para determinar el área limitada por el arco de parábola cuadrática que pasa por los puntos correspondientes.

En el caso de que la curva se encuentre descripta mediante una tabla compuesta de $n+1$ puntos A_0 ; A_1 ; ...; A_n , siendo n un número par y con abscisas x_0 ; x_1 ; ...; x_n equidistantes, es siempre posible aplicar la metodología expuesta, tomando de a tres puntos por vez: (A_0 ; A_1 ; A_2); (A_2 ; A_3 ; A_4); etc. y, de este modo, obtener la expresión:

$$I = \int_{x_0}^{x_n} f(x)\,dx \cong \frac{h}{3}(y_0 + 4y_1 + y_2) + \frac{h}{3}(y_2 + 4y_3 + y_4) + \ldots + \frac{h}{3}(y_{n-2} + 4y_{n-1} + y_n)$$

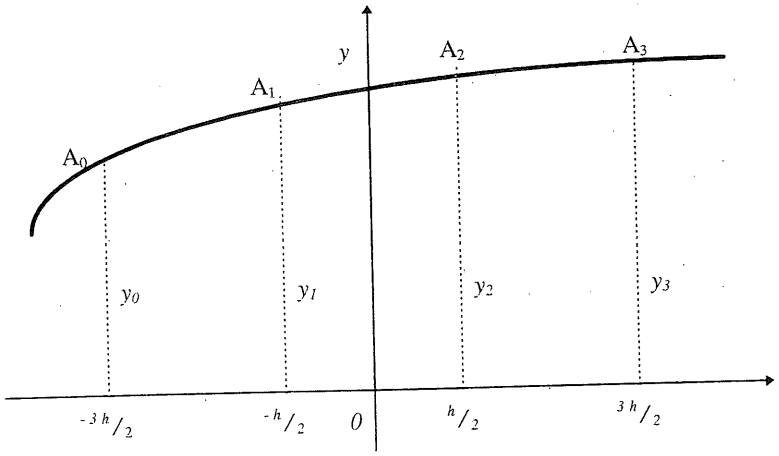
de donde, considerando a los operadores **E**, **P**, **I** con idéntico significado al establecido en el punto anterior, se obtiene:

(8.28) $$I = \int_{x_0}^{x_n} f(x)\,dx \cong \frac{h}{3}(\mathbf{E} + 4\mathbf{I} + 2\mathbf{P})$$

Esta última expresión, es la muy conocida e importante **FÓRMULA DE SIMPSON**, generalmente utilizada para determinaciones expeditivas.

8.2.4.- REGLA DE LOS TRES OCTAVOS DE SIMPSON

La fórmula de *SIMPSON* tiene una particularidad muy importante, la cual se estudiará a continuación.



- Figura 8.4 -

Como es fácil apreciar, la fórmula de *SIMPSON* estudiada en el punto anterior, solo es válida y utilizable en el caso en que se haya subdividido el intervalo de integración en un número de franjas tal, que la cantidad de puntos resultantes; vale decir, los que describen la curva $y = f(x)$, sea impar. Esto sucede cuando el número de franjas aludido es par. El mismo Simpson ha desarrollado una fórmula utilizable en el caso que el número n de franjas sea impar.

La deducción de la correspondiente fórmula es similar a la realizada para la de *SIMPSON* en el punto anterior, excepto que, para la determinación de las áreas parciales, es necesario utilizar parábolas de tercer grado que conecten cuatro puntos consecutivos de la curva en cuestión.

La forma general de la ecuación de tercer grado representada por una parábola cúbica es:

(8.29) $$y = a x^3 + b x^2 + c x + d$$

donde, para determinar los valores de los parámetros $a; b; c; d$ es necesario imponer a la expresión (8.29), la condición que pase por los cuatro puntos A_0 ; A_1 ; A_2 ; A_3 y ubicar el eje de las y como se indica en la figura 8.4, lo cual no hace perder generalidad al razonamiento; con ello el intervalo de integración resulta:

$$-\frac{3h}{2} \le x \le \frac{3h}{2}$$

entonces, se puede calcular el área buscada mediantela expresión:

$$I = \int_{-\frac{3h}{2}}^{\frac{3h}{2}} \left(a x^3 + b x^2 + c x + d\right) dx = \left[\frac{a x^4}{4} + \frac{b x^3}{3} + \frac{c x^2}{2} + d x\right]_{-\frac{3h}{2}}^{\frac{3h}{2}} =$$

$$= \frac{3^4 a h^4}{4.2^4} + \frac{3^3 b h^3}{3.2^3} + \frac{3^2 c h^2}{2^3} + \frac{3 d h}{2} - \frac{3^4 a h^4}{4.2^4} + \frac{3^3 b h^3}{3.2^3} - \frac{3^2 c h^2}{2^3} + \frac{3 d h}{2} =$$

de donde:

$$I = \frac{2.3^3 b h^3}{3.2^3} + \frac{2.3 d h}{2}$$

y, realizando las operaciones indicadas, resulta:

$$I = \frac{3^2 b h^3}{2^2} + 3 d h \tag{8.30}$$

Para calcular los valores de las constantes que intervienen en el cálculo es necesario hacer:

$$y_0 = a\left(-\frac{3h}{2}\right)^3 + b\left(-\frac{3h}{2}\right)^2 + c\left(-\frac{3h}{2}\right) + d$$
$$y_1 = a\left(-\frac{h}{2}\right)^3 + b\left(-\frac{h}{2}\right)^2 + c\left(-\frac{h}{2}\right) + d$$
$$y_2 = a\left(\frac{h}{2}\right)^3 + b\left(\frac{h}{2}\right)^2 + c\left(\frac{h}{2}\right) + d$$
$$y_3 = a\left(\frac{3h}{2}\right)^3 + b\left(\frac{3h}{2}\right)^2 + c\left(\frac{3h}{2}\right) + d$$

Resolviendo, por cualquier método, el conjunto de ecuaciones simultáneas dado y reemplazando sus valores en la expresión (8.30), de la integral I, se tiene:

$$I = \frac{9h^3}{4}\left[\frac{2(y_0 + y_3) - 2(y_1 + y_2)}{8h^2}\right] + \frac{3h}{16}\left[9(y_1 + y_2) - (y_0 + y_3)\right]$$

de lo que, en definitiva, resulta:

$$I = \int_{-\frac{3h}{2}}^{\frac{3h}{2}} f(x)\,dx = \frac{3h}{8}(y_0 + 3y_1 + 3y_2 + y_3) \tag{8.31}$$

que es la expresión analítica de la denominada **REGLA DE LOS TRES OCTAVOS DE SIMPSON**.

Al quitarle tres franjas a una zonificación constituida por una cantidad impar de ellas, da como resultado en el resto una cantidad par, a la que puede aplicarse la fórmula de *SIMPSON* ya estudiada.

Por ejemplo, si se estuviera frente al problema de calcular la superficie de un área subdividida en 47 franjas, la *REGLA DE LOS TRES OCTAVOS DE SIMPSON* se podría utilizar para aproximar el área bajo la curva ocupada por las tres primeras franjas.

El área bajo las 44 franjas restantes, luego de ser calculada mediante la fórmula de *SIMPSON*, se sumaría a la de las tres anteriores.

8.2.5.- FÓRMULA DE EULER-MACLAURIN

Mediante el agregado de términos complementarios que corrigen otras fórmulas elementales como la de los *TRAPECIOS o SIMPSON*, es posible obtener un sinnúmero de expresiones elementales de fórmulas de integración. Una de las más comunes es la que será estudiada a continuación. La misma propone adicionar una serie de términos a la fórmula de los *TRAPECIOS*, aumentando de este modo su precisión, cuado $f(x)$ es conocida.

Considérese que $F(x)$ es una primitiva de $f(x)$; vale decir que, $F'(x) = f(x)$, del mismo modo que, $F''(x) = f'(x)$; etc. Aplicando la fórmula del desarrollo en serie de *TAYLOR* a la función primitiva F, resulta:

$$F(x+h) = F(x) + h\,F'(x) + \frac{h^2}{2!}F''(x) + \frac{h^3}{3!}F'''(x) + \ldots$$

Transponiendo el primer término del segundo miembro, al primer miembro y tomando, sucesivamente, $x=x_0$; $x=x_1$; ...; $x=x_{n-1}$, resulta:

$$F(x_1) - F(x_0) = h f(x_0) + \frac{h^2}{2!}f'(x_0) + \frac{h^3}{3!}f''(x_o) + \ldots$$

$$F(x_2) - F(x_1) = h f(x_1) + \frac{h^2}{2!}f'(x_1) + \frac{h^3}{3!}f''(x_1) + \ldots$$

. . . . . .

$$F(x_n) - F(x_{n-1}) = h f(x_{n-1}) + \frac{h^2}{2!}f'(x_{n-1}) + \frac{h^3}{3!}f''(x_{n-1}) + \ldots$$

La suma miembro a miembro de estas ecuaciones da como resultado en el primer miembro $F(x_n)$ - $F(x_0)$, pero, como $F(x)$ es una primitiva de $f(x)$, es lícito aplicar la *Regla de BARROW* al primer miembro, siendo:

$$\int_{x_0}^{x_n} f(x)\,dx = h\sum_{i=0}^{n-1} f(x_i) + \frac{h^2}{2!}\sum_{i=0}^{n-1} f'(x_i) + \frac{h^3}{3!}\sum_{i=0}^{n-1} f''(x_i) + \ldots \tag{8.32}$$

Expresiones análogas a la anterior se obtienen considerando, sucesivamente, las funciones $f'(x)$; $f''(x)$; etc., resultando:

(8.33) $$\int_{x_0}^{x_n} f'(x)\,dx = f(x_n) - f(x_0) = h\sum_{i=0}^{n-1} f'(x_i) + \frac{h^2}{2!}\sum_{i=0}^{n-1} f''(x_i) + \ldots$$

(8.34) $$\int_{x_0}^{x_n} f''(x)\,dx = f'(x_n) - f'(x_0) = h\sum_{i=0}^{n-1} f''(x_i) + \frac{h^2}{2!}\sum_{i=0}^{n-1} f'''(x_i) + \ldots$$

(8.35) $$\int_{x_0}^{x_n} f'''(x)\,dx = f''(x_n) - f''(x_0) = h\sum_{i=0}^{n-1} f'''(x_i) + \frac{h^2}{2!}\sum_{i=0}^{n-1} f^{IV}(x_i) + \ldots$$

Sumando a la expresión (8.32) la (8.33) multiplicada por $C_1 h$; la (8.34) multiplicada por $C_2 h^2$; la (8.35) multiplicada por $C_3 h^3$, etc., tal como es posible apreciar a continuación, se obtiene:

$$\int_{x_0}^{x_n} f(x)\,dx + C_1 h\left[f(x_n) - f(x_0)\right] + C_2 h^2\left[f'(x_n) - f'(x_0)\right] + C_3 h^3\left[f''(x_n) - f''(x_0)\right] + \ldots =$$

$$= h\sum_{i=0}^{n-1} f(x_i) + h^2\sum_{i=0}^{n-1} f'(x_i)\left[\frac{1}{2!} + C_1\right] + h^3\sum_{i=0}^{n-1} f''(x_i)\left[\frac{1}{3!} + \frac{C_1}{2!} + C_2\right] +$$

$$+ h^4\sum_{i=0}^{n-1} f'''(x_i)\left[\frac{1}{4!} + \frac{C_1}{3!} + \frac{C_2}{2!} + C_3\right] + \ldots$$

Es necesario determinar ahora, los valores que deben tomar los coeficientes C_i de modo que se anulen los corchetes que figuran en el segundo miembro. En consecuencia, se obtiene:

$$\frac{1}{2!} + C_1 = 0 \quad \Rightarrow \quad C_1 = -\frac{1}{2}$$

$$\frac{1}{3!} + \frac{C_1}{2!} + C_2 = 0 \quad \Rightarrow \quad C_2 = \frac{1}{12}$$

$$\frac{1}{4!} + \frac{C_1}{3!} + \frac{C_2}{2!} + C_3 = 0 \quad \Rightarrow \quad C_3 = 0$$

$$\frac{1}{5!} + \frac{C_1}{4!} + \frac{C_2}{3!} + \frac{C_3}{2!} + C_4 = 0 \quad \Rightarrow \quad C_4 = -\frac{1}{720}$$

y así siguiendo se calculan los demás coeficientes. Sustituyendo estos valores en la última expresión de la integral, se obtiene la **FÓRMULA DE EULER-MACLAURIN**:

(8.36) $$\frac{1}{h}\int_{x_0}^{x_n} f(x)\,dx = \left[\frac{1}{2}f(x_0) + f(x_1) + \ldots + f(x_{n-1}) - \frac{1}{2}f(x_n)\right] -$$

$$- \frac{h}{12}\left[f'(x_n) - f'(x_0)\right] + \frac{h^3}{720}\left[f'''(x_n) - f'''(x_0)\right] -$$

$$- \frac{h^5}{30240}\left[f^{V}(x_n) - f^{V}(x_0)\right] + \ldots$$

Tal como se había dicho al principio del tema, la primera parte coincide exactamente con la fórmula de los *TRAPECIOS*. Los términos subsiguientes son correcciones de aquella, que se aplican sucesivamente con el objeto de mejorar la precisión.

8.2.6.- FÓRMULA DE GREGORY

Una fórmula que utiliza solamente los valores de la función y de las correspondientes diferencias sucesivas, interiores al intervalo $(x_0 ; x_n)$ es la denominada **FÓRMULA DE GREGORY**, la cual será deducida a partir de la ya estudiada expresión de *EULER-MACLAURYN*.

Si en la citada fórmula, las derivadas son reemplazadas por las expresiones correspondientes en términos de las diferencias; que son:

$$h f'(x_0) = \Delta f(x_0) - \frac{1}{2}\Delta^2 f(x_0) + \frac{1}{3}\Delta^3 f(x_0) - \frac{1}{4}\Delta^4 f(x_0) + \ldots$$

$$h f'(x_n) = \nabla f(x_n) + \frac{1}{2}\nabla^2 f(x_n) + \frac{1}{3}\nabla^3 f(x_n) + \frac{1}{4}\nabla^4 f(x_n) + \ldots$$

$$h^3 f'''(x_0) = \Delta^3 f(x_0) - \frac{3}{2}\Delta^4 f(x_0) + \frac{7}{4}\Delta^5 f(x_0) - \ldots$$

$$h^3 f'''(x_n) = \nabla^3 f(x_n) + \frac{3}{2}\nabla^4 f(x_n) + \frac{7}{4}\nabla^5 f(x_n) + \ldots$$

$$h^5 f^V(x_0) = \Delta^5 f(x_0) - \frac{5}{2}\Delta^6 f(x_0) + \ldots$$

$$h^5 f^V(x_n) = \nabla^5 f(x_n) + \frac{5}{2}\nabla^6 f(x_n) + \ldots$$

resulta la **FÓRMULA DE GREGORY**:

(6.37)
$$\frac{1}{h}\int_{x_0}^{x_n} f(x)dx = \left[\frac{1}{2}f(x_0) + f(x_1) + \ldots + f(x_{n-1}) + \frac{1}{2}f(x_n)\right] -$$

$$-\frac{1}{12}\left[\nabla f(x_n) - \Delta f(x_0)\right] - \frac{1}{24}\left[\nabla^2 f(x_n) - \Delta^2 f(x_0)\right] -$$

$$-\frac{19}{720}\left[\nabla^3 f(x_n) - \Delta^3 f(x_0)\right] - \frac{3}{360}\left[\nabla^4 f(x_n) - \Delta^4 f(x_0)\right] -$$

$$-\frac{863}{60480}\left[\nabla^5 f(x_n) - \Delta^5 f(x_0)\right] - \ldots$$

Dado que, la aplicación de esta expresión no resulta del todo sencilla en la práctica, se presenta el siguiente:

Ejemplo 8.4.- Se trata de tener que calcular mediante la *FÓRMULA DE GREGORY* la siguiente integral:

$$I = \int_0^{0,5} \frac{1}{\sqrt{1-x^2}}dx$$

Solución: Calculando el valor de la función para intervalos equidistantes de amplitud igual a un décimo, se obtiene la tabla de diferencias de la página siguiente.

Aplicando, entonces, la *FÓRMULA DE GREGORY* con los valores de la tabla de diferencias calculada, la integral resulta:

$$I = 0{,}524238309 - {}^1/_{12}\ 0{,}05857328 - {}^1/_{24}\ 0{,}03135158 - {}^{19}/_{720}\ 0{,}00412988 -$$

$$- {}^3/_{360}\ 0{,}00412988 = \underline{\mathbf{0{,}523600918}}$$

x	$f(x)$	Δ	Δ^2	Δ^3	Δ^4
0,0	1,00000000				
		0,00503781			
0,1	1,00503781		0,01054510		
		0,01558291		0,00153611	
0,2	1,02062072		0,01208121		0,00152317
		0,02766412		0,00305928	
0,3	1,04828484		0,01514049		0,00260671
		0,04280461		0,00566599	
0,4	1,09108945		0,02080648		
		0,06361109			
0,5	1,15470054				

Donde, el valor obtenido para la integral es 0,52360918 que difiere del valor exacto: ${}^1/_6\ \pi$, en menos de $3\text{x}10^{-6}$; vale decir que, tiene cinco cifras decimales exactas; mientras que, la fórmula de los *TRAPECIOS*, la que corresponde al primer término solamente, vale 0,5242 con solo dos dígitos decimales exactos.

◆◆◆ ●●● ◆◆◆

Capítulo 9

Integración de Ecuaciones Diferenciales

Métodos que comienzan por sí mismos

9. - INTEGRACION NUMERICA DE ECUACIONES DIFERENCIALES

PROBLEMAS DE VALORES INICIALES QUE EMPIEZAN POR SÍ MISMOS

Introducción. Método de Euler. Error en el Método de Euler. Método Modificado de Euler. Error en el Método Modificado de Euler. Método de Runge-Kutta. Método de Runge-Kutta de Primer Orden. Método de Runge-Kutta de Segundo Orden. Error del Método de Runge-Kutta de Segundo Orden. Método de Runge-Kutta de Mayor Orden. Error en los Métodos de Runge-Kutta de Orden Superior. Método de Runge-Kutta para Sistemas de Ecuaciones Diferenciales. Método de Runge-Kutta para Sistemas de Ecuaciones Diferenciales de Orden Superior.

9.1 .- INTRODUCCIÓN

El conocimiento de los métodos de resolución numérica de las ecuaciones diferenciales ordinarias resulta necesario o incluso, imprescindible, en todos aquellos casos en que, debido a la imposibilidad de aplicación de la integración exacta, por desconocerse la expresión analítica de alguna primitiva o, por las dificultades que son propias de los procedimientos de integración o, eventualmente a causa de las necesidades impuestas por algún problema en particular, se desean calcular para un grupo de valores de la variable x, los valores numéricos correspondientes de y, que satisfacen una ecuación diferencial dada, con un cierto grado de aproximación previamente establecido. Naturalmente, el proceso descripto será realizado prescindiendo del conocimiento de la expresión analítica de la función solución $y=y(x)$ que resolvería, teóricamente, el problema.

Los métodos de integración aproximada se aplican en los más diversos campos de la ciencia y la técnica, no solo en aquellos casos en que la ecuación que debe ser integrada no tiene una solución que pueda determinarse mediante funciones elementales, sino también, en aquellos otros en que la solución exacta no resuelve el problema en forma práctica. Por ejemplo, la solución de la ecuación diferencial de Bernouilli $y'+xy=y^2$ con la condición inicial $y_0=1$ para $x_0=0$, es:

$$y = \frac{e^{-\frac{x^2}{2}}}{1 - \sqrt{2}\int_0^{\frac{x}{\sqrt{2}}} e^{-t^2}\,dt}$$

en la que, aún siendo y una función explícita de la variable x, su determinación analítica resulta muy compleja y laboriosa.

Una tercera y muy importante razón para estar obligado a utilizar un método numérico o aproximado para resolver una ecuación diferencial, es el hecho de no poder prescindir, para resolverlo, del uso de una computadora electrónica digital, por las características propias del problema, las que, eventualmente, pueden comprometer un número muy grande de pasos, de complejo proceso, imposible de realizar mediante métodos analíticos.

9.2 .- MÉTODO DE EULER

La forma general de las expresiones de las ecuaciones diferenciales de primer orden, que serán tratadas en este capítulo, es:

(9.1) $$y' = f(x\,;\,y)$$

en la cual, y' no es simplemente función de la variable independiente x, sino también, de la variable dependiente y cuyos valores deben ser calculados; no es posible, pues, la integración directa de la misma mediante algunas reglas estudiadas con anterioridad.

Se presentará a continuación un método simple que permite una solución aproximada de (9.1) cuando se conoce una condición inicial $y=y_0$ para $x=x_0$.

Supóngase que, tanto la ecuación dada como la solución de la misma:

(9.2) $$y = y(x)$$

tienen la forma general de la figura 9.1 *(a)* y 9.1 *(b)*, respectivamente.

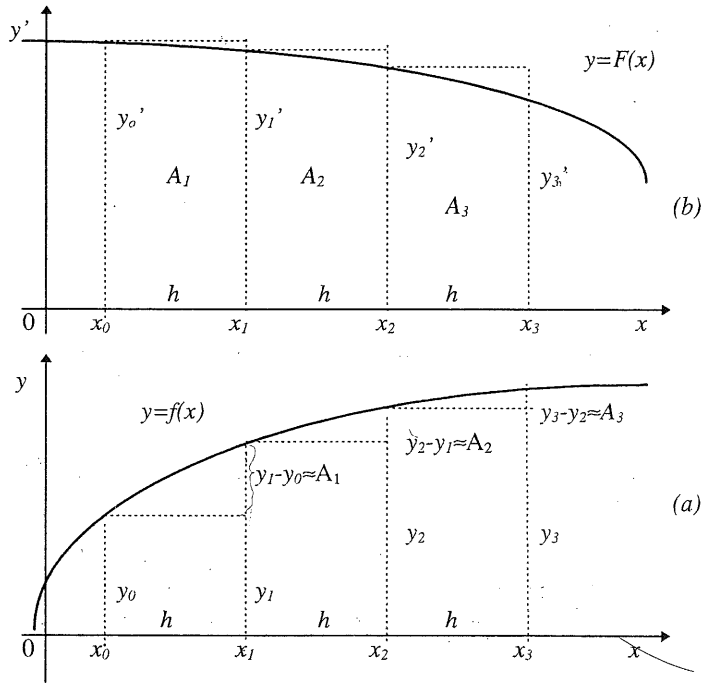
La parte *(b)* de la gráfica puede ser representada dado que por la relación (9.2), el valor de y puede ser escrito:

$$y' = f(x;y) = f[x;y(x)] = F(x)$$

Dado que la solución es conocida para un valor inicial $(x_0\,;y_0)$ es posible determinar el valor inicial de y' haciendo uso de la ecuación (9.1); vale decir, reemplazando los respectivos valores iniciales dados, en la ecuación original, resulta el valor de y_0':

$$y_0' = f(x_0\,;\,y_0)$$

La variación de y desde $x=x_0$ hasta $x=x_0+h$ se representa mediante el área bajo la curva y' entre los valores señalados de x.



- Figura 9.1 -

Teniendo en cuenta que:

$$\int_{x_0}^{x_1} y'\,dx \equiv A_1$$

y, reemplazando ambos miembros por sus valores respectivos, es:

$$[y]_{x_0}^{x_1} \equiv y_0'\,(x_1 - x_0)$$

y, en definitiva:

$$y(x_1) - y(x_0) = y_1 - y_0 \equiv y_0'\,h$$

En las tres últimas expresiones, los primeros miembros representan los valores exactos y los segundos, los valores aproximados. Despejando y_1 de la última, resulta:

(9.3) $$y_1 \cong y_0 + y_0' \, h$$

Habiéndose determinado una aproximación precisa al valor de y_1 por medio la expresión (9.3), utilizando valores pequeños de h es posible obtener una aproximación de y_1' a partir de (9.1), ya que:

$$y_1' = f(x_1 ; y_1)$$

Entonces, como la diferencia $y_2 - y_1$ es aproximadamente igual al área comprendida bajo y' desde x_1 hasta x_2 ; o sea, A_2 de la figura 9.1 *(a)*, resulta:

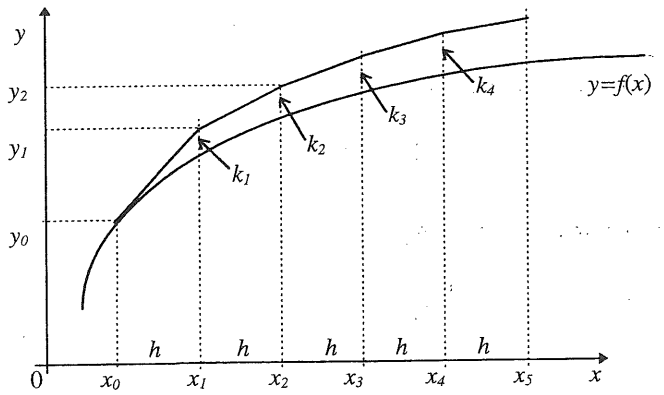
(9.4) $$y_2 \cong y_1 + y_1' \, h$$

Procediendo de la misma manera, y considerando estas expresiones como igualdades, se pueden determinar los sucesivos valores de y_i , obteniéndose la siguiente:

(9.5) $$y_{n+1} = y_n + y_n' \, h$$

la que, generalmente, es conocida como **FÓRMULA DE INTEGRACIÓN HACIA ADELANTE DE EULER.**

El *MÉTODO DE EULER* se clasifica como **MÉTODO QUE EMPIEZA POR SI MISMO**, puesto que solo requiere de un valor de la variable dependiente para hallar la solución; vale decir, el conocimiento de un único punto de la solución (valor inicial) es suficiente para dar origen al procedimiento.



- Figura 9.2 -

La representación de la solución $y=y(x)$ obtenida mediante el método de *EULER*, puede apreciarse en la figura 9.2, donde es posible observar que el método de *EULER* se aproxima, en efecto, por medio de segmentos de recta a la solución exacta del problema; vale decir que, se obtiene una aproximación a la curva representativa de la solución mediante una poligonal.

9.2 .1.- ERROR EN EL MÉTODO DE EULER

Una simple comparación de la ecuación (9.5) con la expresión del desarrollo en serie de Taylor, indica que aquella está constituida por los dos primeros términos de este; vale decir que, en la fórmula de *EULER* se desprecian los términos que contienen h^2 y potencias superiores sucesiyas de h. El error que se introduce en cada paso debido al uso de esta ecuación truncada que, como se ha visto, se conoce con el nombre de error por truncamiento y su correspondiente valor es:

(9.6) $$\mathbf{E} \leq h^2$$

La importancia de que h se mantenga lo más pequeño posible es absolutamente evidente, pues, en caso contrario, los efectos de la propagación de los errores que se cometerían, invalidaría cualquier resultado obtenido, luego de un número no muy grande de pasos.

Con el objeto de comprender más profundamente los basamentos del método de *EULER*, y en consideración al valor que posee desde el punto de vista didáctico-pedagógico, se presenta el siguiente:

Ejemplo 9.1.- Haciendo uso del método de EULER, resolver la ecuación diferencial de primer orden y primer grado:

$$y' = \frac{6x^2}{y}$$

con la condición inicial: $y_0 = 16$, para $x_0 = 4$, utilizando un incremento $h = 0{,}1$.

Solución: Las fórmulas que deberán ser empleadas son las que han sido estudiadas en el punto anterior:

$$x_{i+1} = x_i + h$$

$$y_{i+1} = y_i + y_i' \, h$$

Sabiendo que $Y = 2X\sqrt{X}$, es la solución exacta de la ecuación diferencial dada (se han utilizado mayúsculas para identificar las variables, en este caso), la fórmula resolutoria resulta:

$$Y_{i+1} = 2X_{i+1}\sqrt{X_{i+1}}$$

Confeccionando una tabla que contenga las soluciones, tanto numérica como exacta, mediante el uso de las expresiones dadas, resulta:

x_i	y_i	Y_i	y_i'	Y_i'	y_{i+1}
4,0	16,0000	16,0000	6,0000	6,0000	16,6000
4,1	16,6000	16,6037	6,0759	6,0745	17,2076
4,2	17,2076	17,2149	6,1508	6,1482	17,8227
4,3	17,8227	17,8333	6,2246	6,2209	18,4452
4,4	18,4452	. . .			

♦♦♦ ●●● ♦♦♦

Como puede apreciarse en el ejemplo que antecede, con el valor utilizado del incremento h, los errores son notables desde el principio del procesamiento. No obstante, aún tomando valores más pequeños para el incremento, el método de resolución de ecuaciones diferenciales de *EULER* solo reviste valor didáctico por la claridad y transparencia de los elementos que intervienen que, en definitiva son siempre los mismos para todos los que serán estudiados.

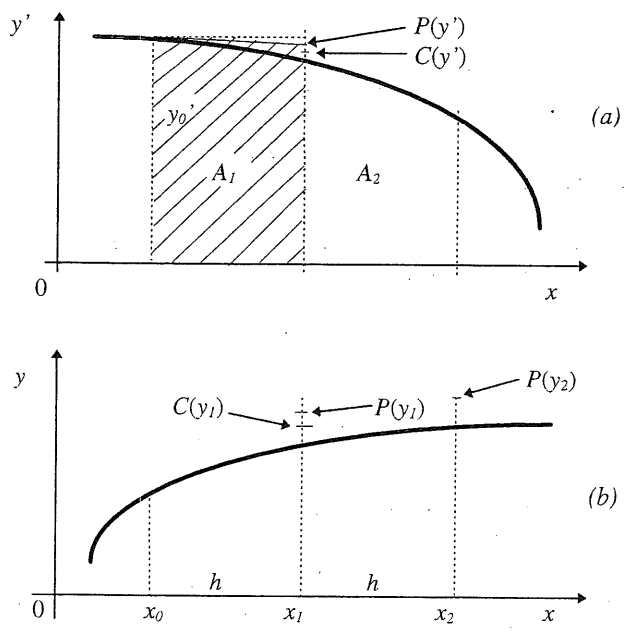
9.3 .- MÉTODO MODIFICADO DE EULER

El método de *EULER*, estudiado en el punto anterior, es muy poco preciso y su tratamiento ha sido incluido en el presente curso, solamente por razones didácticas y pedagógicas.

Una forma de mejorar su precisión es la que se estudiará a continuación. Si se considera la ecuación diferencial de primer orden primer grado:

(9.1) $$y' = f(x\,;\,y)$$

en la que se conoce el valor de $y=y_0$ cuando $x=x_0$ (denominada solución inicial); suponiendo también que, las representaciones gráficas de las funciones y e y', son las que se muestran en la figura 9.3 *(a)* y 9.3 *(b)*.



- Figura 9.3 -

Si se sustituye en la expresión (9.1) el valor inicial conocido de y, se obtiene el correspondiente valor de y' para $x=x_0$; por medio de la cual, utilizando la fórmula de *EULER*, previamente estudiada, se obtiene un valor aproximado para y_1 , mediante la siguiente expresión:

(9.7) $$P(y_1) = y_0 + y_0'\,h$$

llamado **VALOR DE PREDICCIÓN** de y_1. En la expresión (9.7), el término $y_0'\,h$ es el área rectangular A_1. Este área es obviamente diferente al área real bajo la curva dada, de manera que el valor de predicción de y_1 que se obtiene, difiere del valor real. Sin embargo, si el valor de predicción de y_1 se sustituye en (9.1), se puede obtener un valor aproximado de y_1'. Como este valor de y_1' se basa en el valor de predicción se utiliza la notación $P(y_1')$ para representarlo.

$$P(y'_1) = f[x_1\,;\,P(y_1)]$$

A continuación, utilizando el área trapezoidal rayada que se muestra en la figura 9.1 (a) como aproximación al área verdadera bajo la curva y', es posible determinar un valor corregido de y_1, el que será denotado con $C(y_1)$, en la forma:

(9.8) $$C(y_1) = y_0 + \frac{y_0' + P(y_1')}{2} h$$

la que será denominada **ECUACIÓN DE CORRECCIÓN**.

Esta característica del método es la que permite que se lo clasifique como **MÉTODO PREDICTOR-CORRECTOR**.

El valor corregido y_1 se sustituye a continuación en (9.1) para obtener un valor corregido de y_1', de la siguiente manera:

$$C(y_1') = f[x_1 ; C(y_1)]$$

Este procedimiento se continúa hasta que, para dos valores consecutivos de y_1, su diferencia en valor absoluto, sea menor que un cierto **E**, positivo y arbitrario, que satisfaga las condiciones de precisión deseadas.

Una vez obtenido el valor de y_1, con el mismo procedimiento, se avanza un intervalo para obtener y_2, y así sucesivamente.

La forma general de las ecuaciones (9.7) y (9.8), para su aplicación en cualquier paso, insertas en el proceso de cálculo, son:

(9.9)

$$P(y_{i+1}) = y_i + y_i' h \qquad (a)$$

$$P(y_{i+1}') = f[x_{i+1}; P(y_{i+1})] \qquad (b)$$

$$C(y_{i+1}) = y_i + \frac{y_i' + P(y_{i+1}')}{2} h \qquad (c)$$

$$C(y_{i+1}') = f[x_{i+1}; C(y_{i+1})] \qquad (d)$$

$$C(y_{i+1}) = y_i + \frac{y_i' + C(y_{i+1}')}{2} h \qquad (e)$$

donde, las expresiones (9.9) *(d)* y *(e)* deben ser iteradas hasta tanto dos aproximaciones sucesivas de $C(y_{i+1})$ cumplan con la condición de error, previamente establecida.

Por todo lo expuesto, y resumiendo, es posible afirmar que el *MÉTODO MODIFICADO DE EULER* es un método que *EMPIEZA POR SI MISMO* y se clasifica como de *PREDICTOR-CORRECTOR*, destinado a resolver *PROBLEMAS DE VALORES INICIALES*.

Es posible demostrar que el error por paso que resulta de la aplicación del método modificado de *EULER*, está en el orden de:

$$\mathbf{E} \leq h^3$$

por lo que, este método resulta de mayor precisión que el método de *EULER* estudiado en el punto anterior.

Para una mejor comprensión del método modificado de *EULER*, se expone el siguiente:

Ejemplo 9.2 .- Resolver la ecuación diferencial del ejemplo 9.1, calculando los respectivos valores mediante el método modificado de *EULER*, con h=0,1 ; E≤0,0001.

Solución: Aplicando sistemáticamente el conjunto de ecuaciones (9.9), se obtiene la siguiente tabla:

x_i	4,0	4,1	4,2	4,3	4,4
y_i	16,00000	16,60373	17,21487	17,83332	18,45901
y_i'	6,00000	6,07454	6,14817	6,22094	
x_{i+1}	4,1	4,2	4,3	4,4	
$P(y_{i+1})$	16,60000	17,21118	17,82969	18,45541	
$P(y'_{i+1})$	6,07590	6,14949	6,22221	6,29409	
$C(y_{i+1})$	16,60379	17,21493	17,83339	18,45907	
$C(y'_{i+1})$	6,07452	6,14815	6,22091	6,29284	
$C(y_{i+1})$	16,60373	17,21486	17,83332	18,45901	
$C(y'_{i+1})$	6,07454	6,14818			
$C(y_{i+1})$	16,60373	17,21487			

Se debe notar que la disposición de la tabla es, en este caso vertical; orientación que se ha tomado por simples razones de espacio.

Como es posible apreciar, comparando los resultados obtenidos con los del ejemplo 9.1, la corrección de los valores predichos, mejora substancialmente la precisión del método de *EULER*.

♦♦♦ ●●● ♦♦♦

9.4.- MÉTODO DE RUNGE-KUTTA

Dada una ecuación diferencial de primer orden primer grado que, en su expresión general, puede ser escrita de la siguiente manera:

(9.1) $$y' = f(x\,;\,y)$$

se denomina **MÉTODO DE RUNGE-KUTTA** a todo aquel algoritmo que utiliza como expresión de la recurrencia la fórmula:

(9.10) $$y_{i+1} = y_i + a_1 k_1 + a_2 k_2 + \ldots + a_n k_n$$

en la cual las a_i son constantes que más adelante serán calculadas y los k_i toman la forma:

(9.11)
$$k_1 = h f[x_i\,;\,y_i]$$
$$k_2 = h f[x_i + p_1 h\,;\,y_i + q_{11} k_1]$$
$$k_3 = h f[x_i + p_2 h\,;\,y_i + q_{21} k_1 + q_{22} k_2]$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$k_n = h f[x_i + p_{n-1} h\,;\,y_i + q_{n-1;1}\ k_1 + \ldots + q_{n-1;n-1}\ k_{n-1}]$$

Los parámetros $a;\ p;\ q$ deben adoptar valores tales que (9.10) suministre aproximaciones sucesivas precisas de y. Como se verá, estos valores se determinan haciendo (9.10) equivalente a cierto número especificado de términos de un desarrollo en serie de *TAYLOR* de y con respecto a x_i.

Como es posible apreciar, tener la posibilidad de adoptar arbitrariamente el valor de n, permite establecer, consecuentemente, el *ORDEN* del método de *RUNGE-KUTTA*; por lo tanto, con más propiedad se debería denominar *METODOLOGÍA DE RUNGE-KUTTA*, en vez de *MÉTODOS DE RUNGE-KUTTA*, tal como se ha hecho en el acápite.

Estos métodos empiezan por sí mismos, y teóricamente, es posible desarrollar un conjunto de ecuaciones (9.10) y (9.11) que logre cualquier grado de precisión deseado, lo que, juntamente con la facilidad de programación y procesamiento, constituyen las principales ventajas de esta metodología de cálculo.

La única desventaja que presenta el método de *RUNGE-KUTTA* es el requisito de que la función $f(x;y)$ debe ser calculada para varios valores de x e y que difieren muy poco, unos de otros, en cada paso del procesamiento.

Esta determinación repetitiva de $f(x;y)$ lo convierte, generalmente, en un método menos eficiente en lo que respecta a tiempo de procesamiento que otros métodos de precisión comparable, en los cuales los valores previamente determinados de la función, se utilizan en los pasos posteriores.

Algo que no puede ser considerado como una desventaja, pero sí, como algo inconveniente, es la gran dificultad que presenta este método para estimar el error por paso que se comete en las soluciones intermedias de los métodos de orden superior de *RUNGE-KUTTA*, tomando como referencia las soluciones obtenidas por otros métodos que serán estudiados más adelante.

Antes de exponer el método, serán repasados algunos conceptos anteriormente estudiados en el Análisis Matemático. Dada una función compuesta; vale decir, una función que a su vez depende de otra función, tal como la (9.1), en la cual es $y=y(x)$, su diferencial vale:

$$dy' = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$$

y su derivada con respecto a la variable x, es:

(9.12) $$\frac{dy'}{dx} = y'' = \frac{d^2y}{dx^2} = \frac{\partial f}{\partial x}\frac{dx}{dx} + \frac{\partial f}{\partial y}\frac{dy}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx}$$

mientras que la derivada segunda, resulta:

(9.13) $$\frac{d^2y'}{dx^2} = y''' = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx}\right)\frac{dx}{dx} + \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx}\right)\frac{dy}{dx}$$

Operando de manera similar, es posible obtener las derivadas de mayor orden. Recuérdese también que, el desarrollo en serie de *TAYLOR* del segundo miembro de (9.1), en el punto $(x_i\,;y_i)$, se puede expresar mediante:

(9.14) $$f(x_i + b; y_i + c) = f(x_i; y_i) + b\left[\frac{\partial f(x_i; y_i)}{\partial x}\right] + c\left[\frac{\partial f(x_i; y_i)}{\partial y}\right] +$$
$$\frac{1}{2!}\left\{b^2\left[\frac{\partial^2 f(x_i; y_i)}{\partial x^2}\right] + 2bc\left[\frac{\partial^2 f(x_i; y_i)}{\partial x \partial y}\right] + c^2\left[\frac{\partial^2 f(x_i; y_i)}{\partial y^2}\right]\right\} + \ldots$$

donde las constantes $b\ ;c$ deben considerarse como los incrementos de las variables $x\ ;\ y$, respectivamente.

9.4.1.- MÉTODO DE RUNGE-KUTTA DE PRIMER ORDEN

Con el fin de ilustrar, simplemente, y aportar claridad respecto de la aplicación de la metodología estudiada, se deducirá el algoritmo correspondiente al método de *RUNGE-KUTTA* de primer orden el que, en última instancia se reduce al método de *EULER* cuando $a_1 = 1$, pues, siendo en este caso $n=1$, resulta:

$$y_{i+1} = y_i + a_1 k_1$$

pero, dado que:

$$k_1 = h f(x_i ; y_i) = h y_i'$$

entonces, para $a_1 = 1$, se obtiene:

$$y_{i+1} = y_i + h y_i'$$

que no es otra cosa que la fórmula de *EULER*, como se deseaba demostrar.

Como resulta obvio, el método de *RUNGE-KUTTA*, en su expresión de primer orden, solo tiene valor didáctico y no es utilizado con fines de cálculo. Las aplicaciones del método se verán más adelante.

9.4.2.- MÉTODO DE RUNGE-KUTTA DE SEGUNDO ORDEN

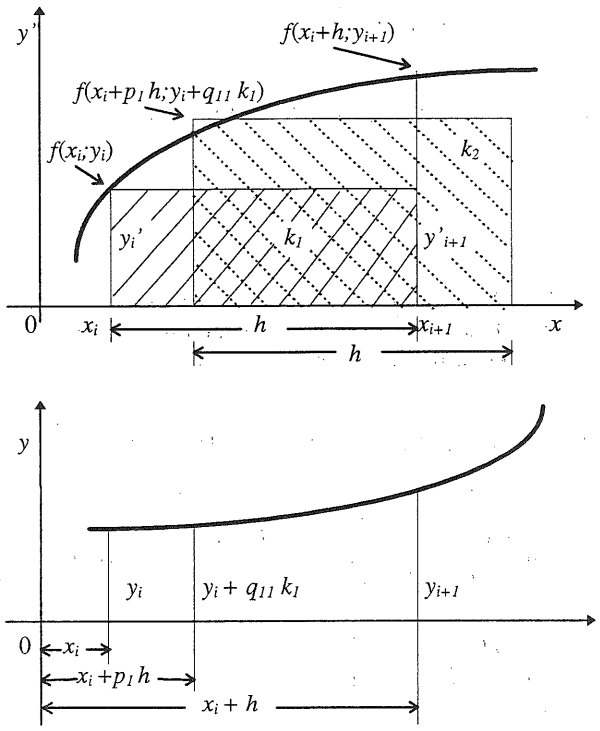
Tomando siempre como base la expresión (9.10) y, en este caso, haciendo $n=2$, aquella se convierte en:

(9.15) $$y_{i+1} = y_i + a_1 k_1 + a_2 k_2$$

donde:

(9.16)
$$k_1 = h f(x_i ; y_i)$$
$$k_2 = h f(x_i + p_1 h ; y_i + q_{11} k_1)$$

El problema consiste en determinar los valores correspondientes a los parámetros $a_1 ; a_2 ; p_1 ; q_{11}$ de manera que la expresión (9.15) suministre un valor preciso y_{i+1}. Una interpretación gráfica de las funciones k_i se ilustra en la figura 9.4. El área sombreada con trazos representa k_1 y el área sombreada con puntos k_2.



- Figura 9.4 -

Para determinar los valores de $a_1 ; a_2 ; p_1 ; q_{11}$, la expresión (9.15) se hará equivalente a un desarrollo en serie de *TAYLOR* en el punto $(x_i ; y_i)$. Para ello, primeramente se desarrolla y_{i+1} en el punto y_i, resultando:

(9.17) $$y_{i+1} = y_i + h y_i' + \frac{h^2}{2!} y_i'' + \ldots$$

De las expresiones (9.1) y (9.12), se puede apreciar que:

(9.18) $$y_i' = f(x_i ; y_i)$$

(9.19) $$y_i'' = \frac{\partial f(x_i ; y_i)}{\partial x} + \frac{\partial f(x_i ; y_i)}{\partial y} f(x_i ; y_i)$$

Sustituyendo ahora, las fórmulas (9.18) y (9.19) en (9.17), se obtiene:

$$y_{i+1} = y_i + h f[x_i ; y_i] + \frac{h^2}{2!}\left[\frac{\partial f(x_i ; y_i)}{\partial x} + \frac{\partial f(x_i ; y_i)}{\partial y} f(x_i ; y_i)\right] + \ldots \tag{9.20}$$

Observando las ecuaciones (9.15), (9.16) y (9.20) se ve que k_2 debe ser expresado en función de:

$$f(x_i ; y_i) \quad ; \quad \frac{\partial f(x_i ; y_i)}{\partial x} \quad \text{y} \quad \frac{\partial f(x_i ; y_i)}{\partial y}$$

si las ecuaciones (9.15) y (9.20) van a contener términos similares. Esto se puede lograr desarrollando k_2 en términos de una serie de *TAYLOR* para dos variables; según (9.14), considerando que $b=p_1 h$; $c=q_{11} k_1$, se puede escribir:

$$k_2 = h f(x_i + p_1 h ; y_i + q_{11} k_1) = h\left\{f(x_i ; y_i) + p_1 h \frac{\partial f(x_i ; y_i)}{\partial x} + \right. \tag{9.21}$$

$$+ q_{11} k_1 \frac{\partial f(x_i ; y_i)}{\partial y} + \frac{1}{2!}\left[p_1^2 h^2 \frac{\partial^2 f(x_i ; y_i)}{\partial x^2} + 2 p_1 q_{11} h k_1 \frac{\partial^2 f(x_i ; y_i)}{\partial x \partial y} + \right.$$

$$\left.\left. + q_{11}^2 k_1^2 \frac{\partial f(x_i ; y_i)}{\partial y^2}\right] + \ldots\right\}$$

Sustituyendo en la (9.15) la primera de las expresiones (9.16) y utilizando los tres primeros términos de (9.21), se obtiene:

$$y_{i+1} = y_i + a_1 h f(x_i ; y_i) + a_2 h f(x_i ; y_i) + a_2 h^2\left[\frac{\partial f(x_i ; y_i)}{\partial x} p_1 + \right. \tag{9.22}$$

$$\left. + q_{11} \frac{\partial f(x_i ; y_i)}{\partial y} f(x_i ; y_i)\right]$$

Igualando los coeficientes de los términos semejantes de las expresiones (9.20) y (9.22) se obtienen las tres ecuaciones independientes siguientes:

$$\begin{cases} a_1 + a_2 = 1 \\ a_2 p_1 = \frac{1}{2} \\ a_2 q_{11} = \frac{1}{2} \end{cases} \tag{9.23}$$

que tienen cuatro incógnitas. Obviamente se trata de un sistema indeterminado de grado uno. En consecuencia, asignando un valor arbitrario a una de las incógnitas y resolviendo el sistema para determinar las otras tres, es posible obtener tantos conjuntos diferentes de valores, y a la vez, otros tantos conjuntos diferentes de ecuaciones (9.15) y (9.16) como se desee.

No obstante, el procedimiento no es arbitrario. Puede demostrarse que en el método de segundo orden de *RUNGE-KUTTA*, la mayor precisión se alcanza cuando a_1 toma el valor $^1/_2$; con lo cual resulta:

$$a_1 = {}^1/_2 \quad ; \quad a_2 = {}^1/_2 \quad ; \quad p_1 = 1 \quad ; \quad q_{11} = 1$$

con los parámetros reemplazados en las ecuaciones (9.15) y (9.16) se obtiene entonces:

$$y_{i+1} = y_i + {}^1/_2 (k_1 + k_2)$$

$$k_1 = h f(x_i ; y_i) \tag{9.24}$$

$$k_2 = h f(x_i + h ; y_i + k_1)$$

expresiones que en conjunto constituyen el **MÉTODO DE RUNGE-KUTTA DE SEGUNDO ORDEN**.

9.4.3.- ERROR DEL MÉTODO DE RUNGE-KUTTA DE SEGUNDO ORDEN

Una solución obtenida a partir de la aplicación de las ecuaciones dadas por las expresiones (9.24) tendrá que cometer, forzosamente, un error por truncamiento en cada paso, del orden de h^3, ya que, para la obtención de las ecuaciones (9.24) se han utilizado fórmulas truncadas, en las cuales fueron despreciados todos los términos que contienen a h^3 y potencias mayores.

El conjunto de expresiones (9.24) se puede utilizar para resolver ecuaciones diferenciales con una precisión equivalente a la del *MÉTODO MODIFICADO DE EULER*. De hecho, reemplazando los valores de k_1 y k_2 en la expresión de y_{i+1} de las ecuaciones (9.24), se obtiene:

$$y_{i+1} = y_i + \frac{1}{2}\left\{h f(x_i ; y_i) + h f[x_i + h ; y_i + f(x_i ; y_i)]\right\} \tag{9.25}$$

considerando, además, que por (9.1), es $f(x_i ; y_i) = y_i'$, y llamando a:

$$f[x_i + h ; y_i + h f(x_i ; y_i)] = P(y'_{i+1})$$

resulta, sustituyéndolas en la (9.25):

$$y_{i+1} = y_i + \frac{y'_i + P(y'_{i+1})}{2} h$$

que es equivalente al *MÉTODO MODIFICADO DE EULER* cuando se omiten las iteraciones sucesivas, en cada paso.

9.4.4.- MÉTODOS DE RUNGE-KUTTA DE MAYOR ORDEN

De todo lo estudiado, es posible deducir que *los MÉTODOS DE RUNGE-KUTTA* de primero y segundo orden no sustituyen con ventajas apreciables a los de *EULER* y *MODIFICADO DE EULER*, respectivamente.

Para obtener métodos de *RUNGE-KUTTA* con mayor precisión es preciso hacer n=3; 4; ... y seguir un procedimiento similar al utilizado para deducir el *MÉTODO RUNGE-KUTTA* de segundo orden, conservando en los desarrollos en serie de *TAYLOR* los términos que contienen h^3; h^4; ... etc.

Por ejemplo, para n=3 se llegará a un sistema de seis ecuaciones con ocho incógnitas, y haciendo una selección particular de valores para dos de esos parámetros, de manera que las expresiones resultantes maximicen la precisión, se obtiene el siguiente conjunto de ecuaciones, las que constituyen el *MÉTODO DE RUNGE-KUTTA DE TERCER ORDEN*:

(9.26)
$$y_{i+1} = y_i + \frac{1}{6}(k_1 + 4k_2 + k_3)$$

donde:

$$k_1 = h\,f(x_i; y_i)$$
$$k_2 = h\,f\left(x_i + \frac{h}{2}; y_i + \frac{k_1}{2}\right)$$
$$k_3 = h\,f(x_i + h; y_i - k_1 + 2k_2)$$

El error que se comete al utilizar estas expresiones, resulta:

$$\mathbf{E} \le h^4$$

pues, se han despreciado en los correspondientes desarrollos en serie de *TAYLOR*, utilizados para deducirlo, todos aquellos términos que contienen h^4, como así también, sus potencias mayores.

Como puede apreciarse fácilmente, ya en el presente orden, el *MÉTODO DE RUNGE-KUTTA* sustituye con ciertas ventajas de precisión a todos los estudiados anteriormente.

El método más frecuentemente utilizado, dentro de la metodología estudiada, es el *MÉTODO DE RUNGE-KUTTA DE CUARTO ORDEN*, el que resulta de hacer n=4 en la expresión (9.10), igualar términos hasta los que contienen h^4 y seleccionar un conjunto particular de los parámetros arbitrarios, que maximicen la precisión.

Los resultados del cálculo descripto en el párrafo anterior, concluye en el siguiente conjunto de ecuaciones:

(9.27)
$$y_{i+1} = y_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

donde:

$$k_1 = h\,f(x_i; y_i)$$
$$k_2 = h\,f\left(x_i + \frac{h}{2}; y_i + \frac{k_1}{2}\right)$$
$$k_3 = h\,f\left(x_i + \frac{h}{2}; y_i + \frac{k_2}{2}\right)$$
$$k_4 = h\,f(x_i + h; y_i + k_3)$$

que, en su conjunto constituyen el método aludido.

9.4.5.- ERROR DE LOS MÉTODOS DE RUNGE-KUTTA DE ORDEN SUPERIOR

En el último algoritmo, dado por las expresiones (9.27), debido a que se han despreciado en los desarrollos en serie de *TAYLOR* correspondientes, los términos que contienen h^4, como aquellos que contienen potencias mayores, se comete un error:

$$\mathbf{E} \le h^5$$

lo que convierte a este, en un método apto para resolver problemas normales de ingeniería y otras ramas científicas.

En general, por la misma razón apuntada al estudiar el error de los métodos de *RUNGE-KUTTA* del orden primero al cuarto, para el de orden n será:

$$\mathbf{E} \le h^{n+1}$$

por lo cual, al menos teóricamente, es posible desarrollar un conjunto de ecuaciones que tenga cualquier grado deseado de precisión, con solo tomar n suficientemente grande.

9.4.6.- MÉTODO DE RUNGE-KUTTA PARA SISTEMAS DE ECUACIONES DIFERENCIALES

Será considerada primeramente, la posibilidad de solución de dos ecuaciones diferenciales de primer orden primer grado de la forma:

$$\begin{cases} \dfrac{du}{dx} = f_1[x;u(x);v(x)] \\ \dfrac{dv}{dx} = f_2[x;u(x);v(x)] \end{cases} \tag{9.28}$$

cuya solución inicial viene dada por:

$$u = u_0 \qquad ; \qquad v = v_0$$

cuando $x = x_0$.

Utilizando el método de cuarto orden de *RUNGE-KUTTA* es posible establecer el siguiente conjunto de ecuaciones para resolver el sistema (9.28):

$$u_{i+1} = u_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \tag{9.29}$$

donde:

$$\begin{aligned} k_1 &= h f_1(x_i;u_i;v_i) \\ k_2 &= h f_1\left(x_i + \frac{h}{2};u_i + \frac{k_1}{2};v_i + \frac{q_1}{2}\right) \\ k_3 &= h f_1\left(x_i + \frac{h}{2};u_i + \frac{k_2}{2};v_i + \frac{q_2}{2}\right) \\ k_4 &= h f_1(x_i + h;u_i + k_3;v_i + q_3) \end{aligned} \tag{9.30}$$

y, además:

$$v_{i+1} = v_i + \frac{1}{6}(q_1 + 2q_2 + 2q_3 + q_4) \tag{9.31}$$

donde, los nuevos elementos q_1 ; q_2 ; q_3 ; q_4 , toman los siguientes valores:

$$\begin{aligned} q_1 &= h f_2(x_i;u_i;v_i) \\ q_2 &= h f_2\left(x_i + \frac{h}{2};u_i + \frac{k_1}{2};v_i + \frac{q_1}{2}\right) \\ q_3 &= h f_2\left(x_i + \frac{h}{2};u_i + \frac{k_2}{2};v_i + \frac{q_2}{2}\right) \\ q_4 &= h f_2(x_i + h;u_i + k_3;v_i + q_3) \end{aligned} \tag{9.32}$$

La solución comienza sustituyendo los valores iniciales de x ; u ; v, en las ecuaciones diferenciales dadas, para obtener valores iniciales de f_1 ; f_2 , que multiplicados por h, dan los valores de k_1 ; q_1 de las ecuaciones (9.30) y (9.32), respectivamente.

Conocidos los valores de k_1 ; q_1 se calculan k_2 ; q_2 , después k_3 ; q_3 , y finalmente, k_4 ; q_4 . Luego se aplican las fórmulas de recurrencia (9.29) y (9.31) para determinar los valores u_1 ; v_1 para $x_1=x_0+h$. Estos valores se utilizan entonces, como valores iniciales para comenzar otro paso del proceso, que acabará calculando los valores de u_2 ; v_2 para $x_2=x_1+h=x_0+2h$, y así sucesivamente, hasta cubrir totalmente el rango de integración deseado.

Es obvio que se pueden resolver sistemas de cualquier número de ecuaciones diferenciales, con solo utilizar un conjunto de ecuaciones como los mostrados, por cada una de las variables dependientes que aparecen en el sistema dado de ecuaciones, que se trata de resolver.

9.4.7.- MÉTODO DE RUNGE-KUTTA PARA ECUACIONES DIFERENCIALES DE ORDEN SUPERIOR

Por el análisis matemático, se sabe que toda ecuación diferencial de orden n, superior al primero (n>1), que toma la forma general:

$$\frac{d^n y}{dx^n} = f\left(x;y;y';y'';\ldots;y^{n-1}\right) \tag{9.33}$$

puede ser transformada en un sistema de n ecuaciones diferenciales de primer orden con n+1 incógnitas, con solo sustituir las n-1 derivadas que aparecen en el segundo miembro de (9.33) por sendas variables dependientes.

Para ver la aplicación práctica de la técnica recientemente expuesta, se presenta el siguiente:

Ejemplo 9.3.- Transformar en un sistema de ecuaciones diferenciales de primer orden, la ecuación diferencial de segundo orden, dada por:

(9.34) $$\frac{d^2y}{dx^2} = f\left(x; y; \frac{dy}{dx}\right)$$

con la condición inicial $y=y_0$; $y' = \frac{d(y_0)}{dx}$, cuando $x=x_0$.

Solución: Haciendo la sustitución de la primera derivada que aparece en la expresión (9.34) por u, resulta:

$$u = \frac{dy}{dx}$$

La ecuación en cuestión, bien puede ser transformada en un sistema compuesto por dos ecuaciones diferenciales de primer orden con tres incógnitas, tomando la siguiente forma:

(9.35) $$\begin{cases} \frac{du}{dx} = f(x; y; u) \\ \frac{dy}{dx} = u \equiv F(x; y; u) \end{cases}$$

Aplicando, en este caso, la resolvente de *RUNGE-KUTTA* de cuarto orden resulta:

(9.36) $$u_{i+1} = u_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

donde:

(9.37) $$\begin{aligned} k_1 &= h\, f(x_i; y_i; u_i) \\ k_2 &= h\, f\left(x_i + \frac{h}{2}; y_i + \frac{k_1}{2}; u_i + \frac{q_1}{2}\right) \\ k_3 &= h\, f\left(x_i + \frac{h}{2}; y_i + \frac{k_2}{2}; u_i + \frac{q_2}{2}\right) \\ k_4 &= h\, f(x_i + h; y_i + k_3; u_i + q_3) \end{aligned}$$

también:

(9.38) $$y_{i+1} = y_i + \frac{1}{6}(q_1 + 2q_2 + 2q_3 + q_4)$$

en la cual:

(9.39) $$\begin{aligned} q_1 &= h\, F(u_i) = h\, u_i \\ q_2 &= h\, F\left(u_i + \frac{k_1}{2}\right) = h\left(u_i + \frac{k_1}{2}\right) \\ q_3 &= h\, F\left(u_i + \frac{k_2}{2}\right) = h\left(u_i + \frac{k_2}{2}\right) \\ q_4 &= h\, F(u_i + k_3) = h(u_i + k_3) \end{aligned}$$

sustituyendo las ecuaciones (9.37) y (9.39) en las (9.36) y (9.38), se obtiene:

(9.40) $$\begin{aligned} y_{i+1} &= y_i + h\, u_i + \frac{h}{6}(k_1 + k_2 + k_3) \\ u_{i+1} &= u_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \end{aligned}$$

donde:

(9.41) $$\begin{aligned} k_1 &= h\, f(x_i; y_i; u_i) \\ k_2 &= h\, f\left(x_i + \frac{h}{2}; y_i + \frac{h}{2}u_i; u_i + \frac{k_1}{2}\right) \\ k_3 &= h\, f\left(x_i + \frac{h}{2}; y_i + \frac{h}{2}u_i + \frac{h}{4}k_1; u_i + \frac{k_2}{2}\right) \\ k_4 &= h\, f\left(x_i + h; y_i + h\, u_i + \frac{h}{2}k_2; u_i + k_3\right) \end{aligned}$$

♦♦♦ ●●● ♦♦♦

De la misma manera que la descripta en el último ejemplo estudiado, es posible resolver ecuaciones diferenciales de orden superior al segundo, mediante transformaciones por sustitución. Dada la importancia que reviste el tema en estudio veamos el siguiente:

Ejemplo 9.4.- Convertir en un sistema de ecuaciones diferenciales de primer orden, la siguiente ecuación diferencial de tercer orden:

(9.42) $$\frac{d^3y}{dx^3} = f\left[x; y(x); \frac{dy}{dx}; \frac{d^2y}{dx^2}\right]$$

con la condición inicial:

$$y = y_0 \; ; \; \frac{dy}{dx} = y_0' \; ; \; \frac{d^2y}{dx^2} = y_0'' \text{ , cuando } x = x_0.$$

Solución: Tal como se ha estudiado, se deben hacer, en primer término, las siguientes sustituciones:

(9.43) $$\frac{dy}{dx} = u(x) \quad \text{con} \quad u(x) = u_0 \quad \text{cuando} \quad x = x_0$$

con lo cual la ecuación (9.42) se transforma en:

(9.44) $$\frac{d^2u}{dx^2} = f\left[x; y(x); u(x); \frac{du}{dx}\right]$$

Dado que, con la sustitución realizada, aún no se ha logrado el objetivo, a continuación es necesario hacer:

(9.45) $$\frac{du}{dx} = v(x) \quad \text{con} \quad v(x) = v_0 \quad \text{cuando} \quad x = x_0$$

y sustituir este valor en la ecuación (9.44), con lo cual se obtiene:

(9.46) $$\frac{dv}{dx} = f\left[x; y(x); u(x); v(x)\right]$$

Tomando las ecuaciones (9.46), (9.45) y (9.43), obtenidas por sustitución, se llega al sistema objeto del presente ejemplo, con la condición inicial $y(x)=y_0$; $u(x)=u_0$; $v(x)=v_0$, para $x=x_0$:

(9.47) $$\begin{cases} \dfrac{dv}{dx} = f\left[x; y(x); u(x); v(x)\right] \\ \dfrac{du}{dx} = v(x) \\ \dfrac{dy}{dx} = u(x) \end{cases}$$

♦♦♦ ●●● ♦♦♦

Como puede apreciarse, se ha convertido la ecuación diferencial de tercer orden (9.42) en un sistema de tres ecuaciones diferenciales de primer orden con cuatro incógnitas, sustituyendo las dos derivadas que aparecen en el segundo miembro de (9.42), por variables dependientes de x.

♦♦♦ ●●● ♦♦♦

Capítulo 10

Integración de Ecuaciones Diferenciales

Métodos que no comienzan por sí mismos

10.- INTEGRACIÓN NUMÉRICA DE ECUACIONES DIFERENCIALES

PROBLEMAS DE VALORES INICIALES QUE NO EMPIEZAN POR SI MISMOS

Introducción. Método de MILNE. Error en el Método de MILNE. Generalización del Método de MILNE. Estabilidad de los Métodos de Integración. Método de HAMMING. Método Modificado de HAMMING. Análisis del Error. Selección de un Método de Integración.

10.1.- INTRODUCCIÓN

En el capítulo anterior se han estudiado algunos métodos que podían ser clasificados dentro del grupo de los de *VALORES INICIALES QUE COMIENZAN POR SÍ MISMOS*, dado que, en un solo punto (la solución inicial) se encuentra toda la información necesaria para dar comienzo al procesamiento.

Por el contrario, en el presente capítulo serán estudiados algunos de aquellos que pueden ser clasificados como *PROBLEMAS DE VALORES INICIALES QUE NO EMPIEZAN POR SI MISMOS*. Estos métodos pueden ser definidos como todos aquellos para los cuales, un solo valor de la variable dependiente, dada en la solución inicial, no es suficiente para dar inicio al procedimiento de integración numérica. Dicho de otra manera, para poner en marcha el procesamiento de los algoritmos prescritos por los métodos tratados en el presente capítulo, es necesario el conocimiento de más de un punto de la solución (en general 3, 4 o más), según el caso.

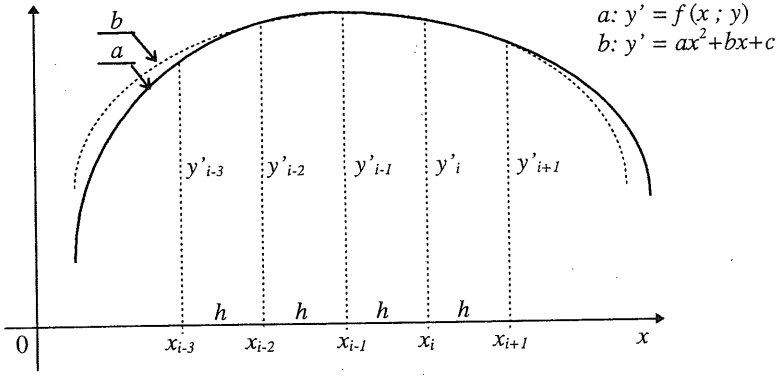
La característica descripta en el párrafo anterior, propia de aquéllos métodos que son clasificados entre los que no comienzan por sí mismos, si bien es cierto que, aparentemente, implican una cierta desventaja inicial, se verá que de hecho, ésta puede ser superada con cierta facilidad con solo realizar algunos pasos adicionales que, a la postre, pueden reportar grandes beneficios en cuanto a la precisión y practicidad, como se estudiará a lo largo del presente capítulo.

10.2.- MÉTODO DE MILNE

Sea la ecuación diferencial general de primer orden y primer grado, considerada en el capítulo anterior:

(10.1) $$y' = f(x\,;\,y)$$

donde, al menos se conoce un punto de la solución $(x_0\,;\,y_0)$, llamado *Solución Inicial.*



- Figura 10.1 -

Se ha estudiado que para cualquier método de resolución estudiado, es posible afirmar que a medida que avanza el proceso de integración numérica, la solución tiende a separarse cada vez más de la respuesta correcta, debido a la propagación de los errores que se cometen en cada paso. Si la solución buscada compromete un número relativamente grande de pasos, es muy importante que el error que se comete por paso del proceso se conserve pequeño. Se demostrará más adelante que el error por paso del método de *MILNE* es del orden de (h^5); vale decir, similar al error del método de *RUNGE-KUTTA* de cuarto orden. Sin embargo, tiene el inconveniente de no empezar por sí mismo, aunque esto, como ya se ha dicho, puede superarse con relativa facilidad.

Se empezará por dividir el área bajo un arco dado de la curva $y' = f(x\,;\,y)$ en cuatro intervalos de amplitud h (ver figura 10.1). El área real bajo esta porción de curva se aproxima entonces considerando el área de estas cuatro franjas como la abarcada por una parábola de segundo grado que tiene tres puntos en común con la curva real, según se indica con trazos discontinuos en la figura 10.1.

Si se hace coincidir el eje de las ordenadas con y'_{i-1} no se pierde generalidad en en razonamiento y los resultados y se tiene la ventaja adicional de simplificar las expresiones. De esta manera, el área bajo la parábola está dada por:

$$A = \int_{-2h}^{2h} \left(ax^2 + bx + c\right) dx$$

la que integrada resulta:

(10.2) $$A = \frac{16}{3} a h^3 + 4ch$$

siendo las constantes $a\,;\,c$:

(10.3) $$a = \frac{y'_i - 2y'_{i-1} + y'_{i-2}}{2h^2} \quad ; \quad c = y'_{i-1}$$

Sustituyendo los coeficientes dados por (10.3) en la (10.2) se obtiene, para el área A de las cuatro franjas:

(10.4) $$A = \frac{4}{3} h\left[2y'_i - y'_{i-1} + 2y'_{i-2}\right]$$

Expresión esta última, que será utilizada posteriormente como parte de la ecuación de predicción. Considerando ahora, que en la ecuación (10.1) se conoce el valor de y para $x=x_0$, esta técnica consiste básicamente en obtener valores apropiados de y utilizando una ecuación de *PREDICCIÓN* y corrigiendo luego estos valores, por el uso iterativo de una ecuación de *CORRECCIÓN.*

La ecuación de *PREDICCIÓN DE MILNE* está dada por la expresión:

(10.5) $$P\left(y_{i+1}\right) = y_{i-3} + A = y_{i-3} + \frac{4}{3} h\left[2y'_i - y'_{i-1} + 2y'_{i-2}\right]$$

que, como se ha visto, utiliza el área de cuatro franjas bajo una aproximación parabólica de segundo grado a la curva, para suministrar un valor de predicción de las ordenadas sucesivas.

La ecuación de *CORRECCIÓN DE MILNE* utiliza la fórmula de *SIMPSON* para determinar el área de dos franjas bajo una curva y suministrar así, valores corregidos de la variable dependiente. Es pues:

$$C(y_{i+1}) = y_{i-1} + \frac{h}{3}\left[y'_{i-1} + 4y'_i + P(y'_{i+1})\right] \tag{10.6}$$

Suponiendo conocidos los valores de $y_0 ; y_1 ; y_2 ; y_3$, el primer paso consiste en obtener un valor de predicción $P(y_4)$, utilizando la ecuación (10.5) con i=3. El valor de predicción que se obtiene, se sustituye en la ecuación diferencial dada (10.1) para obtener el valor de predicción de y_4', que es designado con $P(y_4')$. Este valor se utiliza, entonces, en la ecuación (10.6) para obtener un valor corregido de y_4.

Los tres pasos que se acaban de describir son los primeros en la aplicación del método de *MILNE*. Sin embargo, se ha partido del supuesto que se conocen cuatro valores de y; por ende, las derivadas primeras en los respectivos puntos. Dado que, generalmente, esto no sucede, deben ser determinados de alguna manera valores para:

$$y_1 ; y_2 ; y_3 ; y_1' ; y_2' ; y_3'$$

El uso de los desarrollos en serie de *TAYLOR*, resultan muy precisos para determinar los tres primeros valores de la función, que juntamente con el valor inicial, permiten comenzar el procedimiento. Es obvio que en este caso será necesario conocer el valor de las derivadas sucesivas primeras en el punto $x=x_0$, que intervienen en los desarrollos, para lograr su aplicación.

En caso contrario; vale decir, cuando no se conocen las derivadas primeras sucesivas que se han mencionado, es de buena práctica hacer uso del método de *RUNGE-KUTTA* para obtener aquellos valores iniciales, necesarios para la aplicación del método de *MILNE*.

La solución de la ecuación (10.1) se podría lograr utilizando únicamente el método de *RUNGE-KUTTA* de cuarto orden y con la misma precisión, ya que, ambos tienen el mismo error por paso (h^5). Sin embargo, el método de *MILNE* utiliza menos tiempo de procesamiento y la estimación del error es más sencilla y precisa que para el método de *RUNGE-KUTTA*.

La ecuación de corrección expresada en (10.6), juntamente con la ecuación diferencial dada (10.1), se utilizan en forma iterativa hasta tanto dos valores consecutivos de y_{i+1} difieran en menos de un cierto **E** previamente establecido. Habiendo sido determinado el valor de y_{i+1} con la precisión deseada, el método avanza hacia adelante incrementando el valor de x en h, repitiéndose el procedimiento para obtener el valor de y_{i+2}, y así sucesivamente.

Veamos el siguiente:

Ejemplo 10.1.- El problema consiste en resolver la ecuación diferencial dada por:

$$y' = x + y$$

con la condición inicial (0 ; 0) en el intervalo (0 ; 2,4), siendo h=0,3 y tomando $E \leq 0{,}001$.

Solución: Mediante el desarrollo en serie de *TAYLOR*, es posible calcular:

$$\begin{array}{lll} y(0{,}3) = 0{,}050 & ; & y'(0{,}3) = 0{,}350 \\ y(0{,}6) = 0{,}222 & ; & y'(0{,}6) = 0{,}822 \\ y(0{,}9) = 0{,}560 & ; & y'(0{,}9) = 1{,}460 \end{array}$$

Aplicando sistemáticamente las ecuaciones (10.5), (10.1) y (10.6) se obtiene la siguiente tabla:

x_i	$P(y_{i+1})$	$P(y'_{i+1})$	$C(y_{i+1})$	$C(y'_{i+1})$	$C(y_{i+1})$
1,2	1,119	2,319	1,119	2,319	1,119
1,5	1,979	3,479	1,982	3,482	1,981
1,8	3,248	5,048	3,249	5,049	3,249
2,1	5,062	7,162	5,066	7,166	5,066
2,4	7,618	10,018	7,622	10,019	7,623

♦♦♦ ●●● ♦♦♦

10.2.1.- ERROR EN EL MÉTODO DE MILNE

Las ecuaciones (10.5) y (10.6) de predicción y corrección del método de *MILNE*, respectivamente, resultan exactas bajo condiciones muy particulares. En general, puede demostrarse que el error cometido con la aplicación sistemática de las ecuaciones mencionadas es del orden:

$$\mathbf{E} \leq -\frac{1}{90} h^5 f^{IV}(\theta) \tag{10.7}$$

donde $x_{i-3} \leq \theta \leq x_{i-1}$.

De modo que el error no puede ser calculado con exactitud, pero, puede acotarse tomando el extremo del intervalo que hace máxima la derivada cuarta de f.

Más adelante se estudiarán expresiones avanzadas del método de *MILNE*, haciendo lo propio con sus respectivos errores, lo cual lo convierten en una poderosa herramienta para la resolución de este tipo de problemas.

10.2.2.- GENERALIZACIÓN DEL MÉTODO DE MILNE

Los principios empleados en la deducción de las fórmulas (10.5) y (10.6) del método de *MILNE* se presta a ser extendida a otros métodos, utilizando una forma de cálculo similar. De este modo, se logra toda una metodología de trabajo que extiende el método de *MILNE*. Esta manera de actuar hace aumentar la precisión de las fórmulas halladas, pero, como contrapartida, requerirán más de cuatro puntos de la curva integral para iniciar el procesamiento.

En consecuencia, para que se pueda emplear la metodología descripta, es necesario tomar 6, 8, etc. franjas bajo la curva $y'(x)$. Para el caso particular de 6 franjas, se llega a las siguientes expresiones:

$$P(y_{i+1}) = y_{i-5} + \frac{3}{10}h\left[11y'_{i-4} - 14y'_{i-3} + 26y'_{i-2} - 14y'_{i-1} + 11y'_i\right] \tag{10.8}$$

del mismo modo:

$$C(y_{i+1}) = y_{i-3} + \frac{2}{45}h\left[7y'_{i-3} + 32y'_{i-2} + 12y'_{i-1} + 32y'_i + 7P(y'_{i+1})\right] \tag{10.9}$$

Puede demostrarse que el error que se comete por paso, aplicando las expresiones (10.8) y (10.9) es del orden:

$$E \leq \left| -\frac{8}{945}h^7 f^{VII}(\theta) \right| \tag{10.10}$$

donde, $x_{i-5} \leq \theta \leq x_{i-2}$ que, como se ha estudiado, puede acotarse fácilmente con solo tomar el extremo del intervalo para el cual el valor de la derivada séptima tenga mayor valor absoluto.

10.3.- ESTABILIDAD DE LOS MÉTODOS DE INTEGRACIÓN

Primeramente, y para sistematizar el tema, se estudiará la siguiente:

Definición: *Se denominará* **PASO** *al conjunto de operaciones aritméticas elementales y decisiones lógicas que componen la parte del algoritmo de resolución de una ecuación diferencial, destinada a determinar el valor de la variable en cada punto del intervalo comprometido en el procesamiento.*

Los diversos métodos numéricos destinados a la solución de ecuaciones diferenciales que han sido expuestos hasta el presente, tienen diferentes errores por paso, los que fueron acotados desde el orden h^2 hasta h^7.-

El error por paso de un método aproximado se denomina **ERROR LOCAL**, puesto que es un error que se introduce en el paso correspondiente del proceso de integración. Los errores por redondeo son clasificados también como locales, pues, puede considerarse que ellos se introducen en cada paso. Sin embargo, el **ERROR TOTAL** que existe en la solución durante un paso particular (diferencia entre el valor verdadero y el valor calculado numéricamente) depende no solo de la magnitud de los errores locales introducidos por ese paso particular, sino también de las características de propagación de los errores locales que se han introducido en los pasos previos.

Esto es debido, a que los errores locales pueden tender a crecer durante los pasos siguientes, y cuando intervienen muchos pasos en una integración, el error total puede llegar a ser muy grande, hasta el punto de invalidar totalmente los resultados obtenidos, luego de cierto número de iteraciones, más o menos grande.

En problemas en los cuales la amplitud del intervalo de integración es muy grande, el método numérico a utilizarse debe ser tal, que el error total producido por el mismo no aumente indefinida y descontroladamente, pues, en caso contrario la solución obtenida no sería válida. En el capítulo 3, correspondiente a las Aproximaciones Numéricas, se ha ensayado una definición elemental e intuitiva del concepto de estabilidad en los métodos aproximados que a continuación serán precisados mediante dos definiciones más precisas.

Sea la ecuación diferencial:

$$y' = f(x\,;\,y) \tag{10.1}$$

en cuyo caso se da la siguiente:

Definición: *"Se dice que un método numérico es* **ESTABLE**, *si en el proceso de integrar una ecuación diferencial tal como (10.1), donde $f_y < 0$, la diferencia entre la solución real y la solución numérica (error total), tiende a disminuir en magnitud conforme la integración progresa".*

Se ve en la definición dada, que existen condiciones que no siempre son satisfechas; sin embargo, hay métodos que pueden ser aplicados con ventaja si se adaptan a la siguiente:

Definición: *"Se dice que un método numérico destinado a resolver ecuaciones diferenciales de la forma (10.1) es* **RELATIVAMENTE ESTABLE**, *si la velocidad de crecimiento del error total durante el proceso de integración es menor que la velocidad de crecimiento del valor absoluto de la solución".*

Considerando las definiciones dadas se deduce que, la *ESTABILIDAD* de un método de integración numérica no se define en el caso en que la derivada de (10.1) con respecto a y sea mayor que cero ($f_y > 0$). En este caso la solución de la ecuación diferencial crecerá en forma exponencial y el error total generalmente crecerá en la misma forma. En la solución de una ecuación de ese tipo, sobre un rango extendido de integración, se debe utilizar un método que sea *RELATIVAMENTE ESTABLE*, pues, suponiendo que la solución exacta tiene la forma:

(10.11) $$y = A\, e^{x}$$

entonces, el error total causado por el método numérico en la solución debe tener la forma general siguiente:

(10.12) $$\mathbf{E} = B\, e^{mx}$$

en cuyo caso, si es $m < 1$, el método se considera relativamente estable y su aplicación, seguramente, será ventajosa.

La estabilidad relativa de un método es también importante cuando se utiliza para resolver una ecuación tal como la (10.1), para la cual es $f_y < 0$, si la solución tiende a cero asintóticamente, porque, si se desea una solución precisa que encierre un valor de y muy cercano a cero, la velocidad de disminución del error total debe ser mayor que la de la solución. En forma analítica, si la solución tiene la forma:

(10.13) $$y = A\, e^{-x}$$

entonces el error total producido por el método debe tener la forma general:

(10.14) $$\mathbf{E} = B\, e^{-mx}$$

en el cual $m > 1$, para que el método se considere relativamente estable.

Como es posible apreciar, los conceptos de *ESTABILIDAD* y *ESTABILIDAD RELATIVA* son independientes y no excluyentes; vale decir, que de ningún modo el primero implica al segundo, como parte propia; ni tampoco que, el segundo invalida la del primero de los mencionados. Con el riesgo de resultar reiterativo, digamos que, un método de integración puede ser *estable* y *relativamente estable*, a la vez; o, puede ser estable sin ser relativamente estable.

Para finalizar, es necesario mencionar que la propagación de los errores durante una integración no es un obstáculo serio en la solución de la mayoría de los problemas prácticos, donde, con frecuencia no interesan soluciones que requieran intervalos extendidos de integración. El uso de valores relativamente pequeños de h suministrará respuestas suficientemente precisas, independientemente de las características de *estabilidad o estabilidad relativa* del método numérico empleado.

Es así que, Las soluciones obtenidas mediante los métodos de *RUNGE-KUTTA o MILNE* (que no son estables ni relativamente estables, como puede comprobarse), coinciden aproximadamente con las soluciones obtenidas mediante métodos estables y/o relativamente estables.

Sin embargo, si se desearan soluciones sobre intervalos de integración muy extensos, se deberá recurrir a métodos estables y/o relativamente estables, pues, en este caso no suministran el mismo grado de precisión luego de realizar un número relativamente grande de pasos, y el error total que se comete, crecerá conforme se extendiera arbitrariamente el intervalo de integración.

10.4.- MÉTODO DE HAMMING

El Método de **HAMMING**, al igual que todos los estudiados anteriormente, permite resolver ecuaciones diferenciales del tipo (10.1). Su tratamiento reviste gran importancia debido, principalmente, al hecho que es *ESTABLE* y *RELATIVAMENTE ESTABLE.*

Por otra parte se trata, también, de un método del tipo que se ha dado en llamar *PREDICTOR-CORRECTOR* y además, clasificado entre los que *NO COMIENZAN POR SI MISMOS.*

Utiliza, para dar inicio al procedimiento, la misma fórmula de predicción que el método de *MILNE*, vale decir:

(10.5) $$P(y_{i+1}) = y_{i-3} + \frac{4}{3}h\left[2y'_i - y'_{i-1} + 2y'_{i-2}\right]$$

deducida del área calculada debajo de la parábola de segundo grado que pasa por los tres puntos centrales, tomados de cuatro franjas de amplitud h.

Para obtener una mejor aproximación de este valor se utiliza la denominada **ECUACIÓN GENERALIZADA DE CORRECCIÓN DE HAMMING** dada por la siguiente expresión:

(10.15) $$y_{i+1} = a_i y_i + a_{i-1} y_{i-1} + a_{i-2} y_{i-2} + h\left[b_{i+1} y'_{i+1} + b_i y'_i + b_{i-1} y'_{i-1}\right]$$

en la que, sustituyendo los valores de $y_{i-2}\,;\, y_{i-1}\,;\, y'_{i-1}\,;\, y'_{i+1}$ por sus respectivos desarrollos en serie de *TAYLOR* en función de y_i, se obtiene:

(10.16)
$$y_{i+1} = a_i y_i + a_{i-1}\left[y_i + y'_i(-h) + \tfrac{1}{2!}y''_i(-h)^2 + \tfrac{1}{3!}y'''_i(-h)^3 + \ldots\right] +$$
$$+ a_{i-2}\left[y_i + y'_i(-2h) + \tfrac{1}{2!}y''_i(-2h)^2 + \tfrac{1}{3!}y'''_i(-2h)^3 + \ldots\right] +$$
$$+ h\left\{b_{i+1}\left[y'_i + y''_i h + \tfrac{1}{2!}y'''_i h^2 + \tfrac{1}{3!}y^{IV}_i h^3 + \ldots\right] + b_i y'_i +\right.$$
$$\left. + b_{i-1}\left[y'_i + y''_i(-h) + \tfrac{1}{2!}y'''_i(-h)^2 + \tfrac{1}{3!}y^{IV}_i(-h)^3 + \ldots\right]\right\}$$

Por otra parte, se sabe que desarrollando en serie de *TAYLOR* la función y_{i+1}, resulta:

(10.17)
$$y_{i+1} = y_i + h y'_i + \frac{1}{2!}h^2 y''_i + \frac{1}{3!}h^3 y'''_i + \ldots$$

Igualando los coeficientes de $y_i\,; y'_i; y''_i; y'''_i, y^{IV}_i$; respectivamente, en la expresión (10.16) con los homólogos de (10.17), se llega al siguiente conjunto de ecuaciones lineales simultáneas:

(10.18)
$$\begin{cases} a_i + a_{i-1} + a_{i-2} = 1 \\ -a_{i-1} - 2a_{i-2} + b_{i+1} + b_i + b_{i-1} = 1 \\ \tfrac{1}{2}a_{i-1} + 2a_{i-2} + b_{i+1} - b_{i-1} = \tfrac{1}{2} \\ -\tfrac{1}{6}a_{i-1} - \tfrac{4}{3}a_{i-2} + \tfrac{1}{2}b_{i+1} + \tfrac{1}{2}b_{i-1} = \tfrac{1}{6} \\ \tfrac{1}{24}a_{i-1} + \tfrac{2}{3}a_{i-2} + b_{i+1} - \tfrac{1}{6}b_{i-1} = \tfrac{1}{24} \end{cases}$$

Este sistema posee seis incógnitas. Sería necesario entonces, el concurso de una sexta ecuación para determinar el valor de las mismas. La sexta ecuación podría deducirse agregando un término adicional a cada uno de los desarrollos en serie que componen la expresión (10.16) e igualar los coeficientes de y^{IV}. Sin embargo, *HAMMING* descubrió que utilizando el valor de $a_{i-1} = 0$, conducía a una ecuación de corrección que, a la vez, era *estable y relativamente estable* cuando eran impuestas ciertas condiciones a la magnitud del incremento h.

Utilizando $a_{i-1} = 0$, el sistema (10.18) arroja los siguientes valores:

(10.19)
$$a_i = \tfrac{9}{8}\,;\, a_{i-1} = 0\,;\, a_{i-2} = -\tfrac{1}{8}\,;\, b_{i+1} = \tfrac{3}{8}\,;\, b_i = \tfrac{3}{4}\,;\, b_{i-1} = -\tfrac{3}{8}$$

La sustitución de los valores obtenidos, en la ecuación (10.15) produce la **ECUACIÓN DE CORRECCIÓN DE HAMMING**, dada por:

(10.20)
$$C(y_{i+1}) = \tfrac{1}{8}\left\{9y_i - y_{i-2} + 3h\left[P(y'_{i+1}) + 2y'_i - y'_{i-1}\right]\right\}$$

Esta ecuación es *ESTABLE y RELATIVAMENTE ESTABLE* cuando $f_y<0$, si se toma:

(10.21)
$$h < \frac{0{,}75}{|f_y|}$$

usada conjuntamente con la ecuación (10.20), para realizar la iteración hasta lograr la convergencia deseada.

La ecuación (10.20) es *RELATIVAMENTE ESTABLE* para la solución de las ecuaciones diferenciales en intervalos donde $f_y>0$, cuando:

(10.22)
$$h < \frac{0{,}4}{f_y}$$

utilizada junto con la ecuación de corrección (10.20), para ejecutar la iteración hasta lograr la convergencia deseada.

Para mantener pequeño el error por paso, el valor de h debe ser menor que los especificados por las relaciones (10.21) y (10.22). El error por paso, en este método es del orden de h^5.

La forma natural de operar, en estos casos, consiste en calcular los valores correspondientes de $y_{i-2}\,;\, y_{i-1}\,;\, y_i\,;\, y'_{i-2}\,;\, y'_{i-1}\,;\, y'_i$ para iniciar el procedimiento, del mismo modo propuesto para el método de *MILNE*, una vez obtenidos aquellos, predecir el valor de y_{i+1} por medio de la fórmula (10.5) de predicción de *MILNE*.

Luego utilizar la ecuación diferencial dada (10.1) para calcular una predicción de la derivada primera y, por último, corregir las predicciones mediante la aplicación iterativa de la fórmula de corrección de *HAMMING*, dada por (10.20), hasta obtener la convergencia deseada.

10.4.1.- MÉTODO MODIFICADO DE HAMMING

Un procedimiento alternativo, sugerido por el mismo *HAMMING*, que elimina la mayor parte del tiempo de iteración, acortando en consecuencia la duración total del procesamiento es el que se expone a continuación.

Del estudio del error por truncamiento que se produce en cada paso, mediante la aplicación de las expresiones (10.5), (10.1) y (10.20) con la metodología de *HAMMING*, se deduce que la mayor parte del error que se comete en el cálculo del valor de predicción, se puede eliminar utilizando la llamada **ECUACIÓN DE MODIFICACIÓN DE HAMMING**, dada por:

$$M(y_{i+1}) = P(y_{i+1}) - \frac{112}{121}\left[P(y_i) - C(y_i)\right] \tag{10.23}$$

donde $M\,;\,P\,;\,C$ denotan los valores de *MODIFICACIÓN, PREDICCIÓN y CORRECCIÓN*, respectivamente.

El valor modificado dado por (10.23) se sustituye luego en la ecuación diferencial dada para obtener un valor modificado de y'_{i+1} que se designa con $M(y'_{i+1})$; siendo este valor el que se utiliza en la ecuación de corrección para obtener $C(y_{i+1})$; :

$$C(y_{i+1}) = \tfrac{1}{8}\left\{9y_i - y_{i-2} + 3h\left[P(y'_{i+1}) + 2y'_i - y'_{i-1}\right]\right\} \tag{10.24}$$

Estudiando también, los errores por truncamiento en (10.5) y (10.24) se deduce que la mayor parte del error en el valor corregido dado por (10.24) se elimina utilizando la expresión:

$$F(y_{i+1}) = C(y_{i+1}) - \frac{9}{121}\left[P(y_{i+1}) - C(y_{i+1})\right] \tag{10.25}$$

donde $F(y_{i+1})$; representa directamente; es decir, sin iteraciones, el **VALOR FINAL DEL MÉTODO MODIFICADO DE HAMMING**.

Este método también resulta *estable y relativamente estable* si se aplica en el caso en que $f_y < 0$, tomando el incremento:

$$h < \frac{0{,}65}{|f_y|}$$

En caso contrario; es decir, cuando $f_y > 0$, resulta *relativamente estable* tomando al incremento:

$$h < \frac{0{,}4}{f_y}$$

10.5.- ANÁLISIS DEL ERROR

La utilización de métodos numéricos para resolver ecuaciones diferenciales produce generalmente soluciones numéricas que difieren de las soluciones exactas. La diferencia entre la solución numérica y la solución exacta en cualquier paso dado, se conoce como **ERROR TOTAL** en ese paso. Es importante que este error total se conserve dentro de límites razonablemente pequeños, a lo largo de todo el intervalo de integración comprometido. El objetivo primordial del análisis del error consiste en suministrar algunos métodos para controlar dicho error.

El error total en cualquier paso proviene, como ya se ha estudiado en los capítulos correspondientes, de las siguientes fuentes:

a) El *ERROR POR REDONDEO* que se introduce en un paso dado, al utilizar las fórmulas que involucran operaciones aritméticas con números que tienen una limitada cantidad de dígitos.

b) El *ERROR POR TRUNCAMIENTO* que se comete en un paso dado al utilizar fórmulas de desarrollos en serie o iterativas que desprecian todos los términos a partir de alguno en adelante.

c) El *ERROR ACUMULADO* que se produce debido a la *propagación de los errores* que fueron introducidos en todos los pasos anteriores, a medida que progresaba el procesamiento del algoritmo.

El error por redondeo, en general, es muy pequeño; sin embargo, si es utilizado un método *NO ESTABLE* ni *RELATIVAMENTE ESTABLE* y si el proceso de integración se realiza sobre un intervalo muy extendido, tal que comprometa un número muy grande de pasos, el efecto acumulativo de los errores por redondeo de cada paso y su propagación amplificada en los cálculos de los pasos siguientes, puede conducir a un error total excesivamente grande.

Parte del error descripto, puede controlarse con el uso de *aritmética de DOBLE PRECISIÓN o PRECISIÓN EXTENDIDA,* que es posible usufructuar en todas las computadoras actuales.

Obviamente, el error por truncamiento en cada paso es mínimo en aquellos métodos que emplean fórmulas que tienen errores por truncamiento de mayor orden. En cualquier método este error puede ser reducido con solo reducir el tamaño del paso h. Sin embargo, al reducir el error por truncamiento mediante la reducción del tamaño del paso, se llega a un límite en el cual, a partir de allí, el error por redondeo se vuelve dominante, y el error total aumentará con cualquier reducción sucesiva del tamaño del paso. En el caso específico del error por truncamiento, también los pequeños errores que van cometiéndose en cada paso, se propagarán amplificándose al calcular los valores de la variable en todos los pasos siguientes, produciendo un error total muy grande luego de un número determinado de ellos.

El error acumulado puede ser neutralizado mediante el uso de un método *ESTABLE y/o RELATIVAMENTE ESTABLE* que aseguren, bajo ciertas condiciones, que un error introducido en un paso particular del proceso, no se amplificará en los pasos posteriores. Aunque los métodos inestables no proporcionan esta garantía, se pueden utilizar satisfactoriamente si no se requiere procesar a lo largo de un intervalo extendido, que involucre un número muy grande de pasos.

Se debe recordar que, por lo general, los valores iniciales de las variables y parámetros con que se da origen al procesamiento, están afectados de ciertos errores denominados *INHERENTES*, que lógicamente, intervendrán en la formación del error total. En estos casos no se debe olvidar que no tiene ningún sentido calcular el valor de la solución con mayor precisión que la que justifican los datos utilizados, seleccionando un tamaño de paso demasiado pequeño.

De todos modos, jamás se estará libre de la difícil tarea de tener que definir razonablemente la magnitud del paso a emplearse en el procesamiento. Cabe, pues, preguntarse: ¿ sobre que base se debe seleccionar el tamaño del paso ?. Para responder a esta pregunta es necesario examinar el efecto de cada uno de los errores por paso, en función del incremento y su influencia sobre el error total.

El error por redondeo está presente en cada paso de un proceso y su magnitud depende de la capacidad del soporte de la computadora y compilador utilizados. Este error es básicamente independiente del tamaño del intervalo, y aumenta en proporción con el número de pasos comprometidos en el procesamiento.

El error por truncamiento aparece también en cada paso del proceso, pero, a diferencia del error por redondeo, es función de la magnitud del intervalo h, ya que varía con el orden del error h^n, del método que está siendo utilizado. Entonces, la única forma de controlar el error total por paso de un método, en un paso dado, consiste en controlar el error por truncamiento que se comete en ese paso; vale decir que, el error total disminuye conforme se reduce el error por truncamiento de cada paso, hasta el punto en que el incremento del número de pasos y la consecuente presencia del error por redondeo, producen un aumento del error total.

Al seleccionar inicialmente la magnitud del incremento para un método particular, el error por truncamiento de cada paso que puede admitirse, respecto del número de pasos que se anticipan en el procesamiento, se define mediante valores límites o cotas de error, como ya se ha indicado; se hace entonces, la mejor estimación posible de la magnitud de la función derivada que aparece en el término del error por truncamiento del método, y se calcula el incremento de modo tal, que este error por truncamiento se encuentre dentro de los límites definidos previamente.

Finalmente, se debe decir que, un análisis del error por truncamiento se puede hacer en cada paso de un proceso numérico y los resultados se pueden utilizar para ajustar el tamaño del paso en un entorno de ese punto.

No serán dados mayores detalles de esta metodología, por no ser objeto de estudio en el presente curso.

10.6.- SELECCIÓN DE UN MÉTODO DE INTEGRACIÓN NUMÉRICA

Habiéndose estudiado algunos de los métodos de integración numérica existentes, para resolver ecuaciones diferenciales, y el error que en cada uno de ellos se comete, se está en condiciones de compararlos desde un punto de vista tal que permita seleccionar el más apto para una aplicación determinada, siguiendo ciertos criterios definidos:

1.- Cuando resulta obvio que el rango de integración de un problema es relativamente corto, se pueden utilizar tamaños de pasos relativamente pequeños sin que por ello se deba emplear mucho tiempo de procesamiento y no existe ninguna razón que exija analizar el error por truncamiento en cada paso, o para complicar la solución utilizando uno de los métodos más precisos. Es poco probable que la estabilidad se vuelva un problema en un intervalo corto de integración, de manera que un método simple que inicie por si mismo, como el de *EULER*, es plenamente satisfactorio.

2.- En el caso en que el intervalo de integración sea suficientemente grande como para involucrar un relativamente grande número de pasos, es necesario utilizar un método cuyo error por truncamiento por paso sea suficientemente pequeño, con el objeto de minimizar el error acumulativo, tal como los de *MILNE* o *HAMMING*. Se debe hacer en cada paso un análisis del error por truncamiento por paso con el objeto de minimizar el procesamiento requerido, sin sacrificar la precisión deseada.

3.- En el caso en que el intervalo de integración sea muy grande involucrando un muy grande número de pasos, será necesario utilizar un método cuyo error por truncamiento por paso sea pequeño, con el objeto de minimizar el error. Si, además, la estabilidad estuviera comprometida, deberá preferirse el método de *HAMMING*, para valores de h que lo hagan estable. Se debe hacer en cada paso un análisis del error por truncamiento con el objeto de minimizar el tiempo de procesamiento requerido, sin sacrificar la precisión establecida.

4.- Cuando se desea un error por truncamiento por paso pequeño y no es importante el tiempo de procesamiento que demande, resulta conveniente el método de cuarto orden de *RUNGE-KUTTA*, que empieza por si mismo; en este caso la estabilidad, generalmente, no representa ningún problema y es relativamente fácil de programar.

5.- Si el problema involucra algún rango intermedio de integración, en el cual debe ser considerada la acumulación del error y además, el tiempo de procesamiento, pero ninguno como factor crítico, puede emplearse el método *MODIFICADO de EULER* o un método de *RUNGE-KUTTA* de tercer orden.

Existen muchos métodos numéricos que permiten la solución de ecuaciones diferenciales, además de los ya estudiados, sin embargo, para la mayoría de los problemas de valores iniciales comunes en las ciencias y la técnica resultará útil el uso de alguno de los métodos vistos.

♦♦♦ ●●● ♦♦♦

Capítulo 11

Integración de Ecuaciones Diferenciales

Problemas con condiciones de contorno

11.- INTEGRACIÓN NUMÉRICA DE ECUACIONES DIFERENCIALES

PROBLEMAS CON CONDICIONES DE CONTORNO

Planteo General. Fórmulas de Aproximación por Diferencias Centrales. Método de las Diferencias Finitas. Sistema Análogo de Diferencias Finitas. . Método de las Diferencias Finitas para Ecuaciones Diferenciales de Orden Superior. Ejemplos.

11.1.- PLANTEO GENERAL

Como se ha estudiado, las ecuaciones diferenciales de primer orden, de la forma $y' = f(x\ ;\ y)$, juntamente con un valor de la solución, llamado solución inicial, constituyen sistemas determinados que se conocen bajo la denominación general de *PROBLEMAS DE VALORES INICIALES*. En otro orden de cosas, muchos fenómenos de las ciencias físicas conducen a ecuaciones diferenciales de segundo orden, o mayor, donde se establecen condiciones en los extremos de un cierto intervalo a lo largo del cual se tendrá que efectuar la integración. Este tipo de modelos reciben el nombre de **PROBLEMAS CON CONDICIONES DE CONTORNO o, BORDE o, CON VALORES EN LA FRONTERA** y pueden expresarse analíticamente, en el caso de ecuaciones de segundo orden, mediante el siguiente conjunto de ecuaciones:

$$\begin{aligned} & y'' = f(x\ ;\ y\ ;\ y') \\ & y(a) = A \\ & y(b) = B \end{aligned} \qquad (11.1)$$

donde la ecuación diferencial define el comportamiento de la solución en el intervalo abierto $(a;b)$, y las dos restricciones dadas en segundo y tercer término de las (11.1) y llamadas *CONDICIONES DE CONTORNO o, CONDICIONES DE BORDE o, VALORES EN LA FRONTERA*, establecen el valor de la solución en cada extremo del intervalo.

Definición: *Un problema con valores en la frontera está formulado en términos de una ecuación diferencial de segundo orden, o mayor, sobre algún intervalo abierto que establece restricciones a los puntos frontera del intervalo dado.*

El tipo de problemas descripto en los párrafos anteriores, difiere esencialmente de aquellos con valores iniciales, donde toda restricción se establece para un solo valor de la variable independiente, logrando la aproximación a la solución mediante métodos numéricos diferentes; pues, en los problemas de valores iniciales, toda la información necesaria y suficiente para determinar unívocamente la solución en el punto siguiente, se encuentra disponible en el punto inicial.

Contrariamente, en los problemas con condiciones de contorno, se establecen las condiciones sobre la solución para diferentes valores de la variable independiente, en distintos puntos frontera del intervalo y, por lo tanto, en cada punto del intervalo en cuestión no se encuentra disponible toda la información necesaria y suficiente para determinar el valor de la solución en cualquier otro punto.

De lo dicho se concluye que la solución no puede determinarse básicamente punto a punto, como en los problemas de valores iniciales, sino que, el procedimiento numérico para resolver problemas con condiciones de contorno debe involucrar a todas las condiciones de contorno o valores en la frontera y permitir la introducción o intervención de estas restricciones en el cálculo de cada valor de la variable independiente; vale decir que, se requiere que la solución numérica se construya sobre la totalidad del rango de interés, dentro del método mismo.

11.2.- FÓRMULAS DE APROXIMACIÓN POR DIFERENCIAS CENTRALES

Las fórmulas de aproximación por diferencias centrales de las derivadas sucesivas de una función de x en el punto $x = x_i$ se obtienen utilizando las expresiones de los desarrollos en serie de *TAYLOR* de la función considerada, tomado en un entorno del punto $x = x_i$. Sea este:

$$f(x) = f(x_i + \Delta x) = f(x_i) + \Delta x\, f(x_i) + \frac{\Delta x^2}{2!} f''(x_i) + \frac{\Delta x^3}{3!} f'''(x_i) + \dots \tag{11.2}$$

Haciendo $\Delta x = h$, en un caso, y $\Delta x = -h$, en el otro, y reemplazando estos valores en la (11.2), se obtiene:

$$\begin{aligned} f(x_i + h) &= f(x_i) + h f'(x_i) + \frac{h^2}{2!} f''(x_i) + \frac{h^3}{3!} f'''(x_i) + \circ(h^4) \\ f(x_i - h) &= f(x_i) - h f'(x_i) + \frac{h^2}{2!} f''(x_i) - \frac{h^3}{3!} f'''(x_i) + \circ(h^4) \end{aligned} \tag{11.3}$$

Tomando los tres primeros términos de las fórmulas (11.3) y restando la segunda de la primera de ellas, en un caso, y sumando ambas, en el otro, se obtiene con una aproximación de $\circ(h^2)$, las siguientes expresiones:

$$f'(x_i) = \frac{1}{2h}\left[f(x_i + h) - f(x_i - h)\right] + \circ(h^2) \tag{11.4}$$

$$f''(x_i) = \frac{1}{h^2}\left[f(x_i + h) - 2f(x_i) + f(x_i - h)\right] + \circ(h^2) \tag{11.5}$$

Estas ecuaciones reciben el nombre de **FÓRMULAS DE APROXIMACIÓN POR DIFERENCIAS CENTRALES** para la primera y segunda derivadas de la función $f(x)$, respectivamente, en el punto $x = x_i$.

Haciendo en las fórmulas (11.4) y (11.5), las siguientes sustituciones, por simple cambio de notación:

$$f(x_i - h) = y_{i-1} \quad ; \quad f(x_i) = y_i \quad ; \quad f(x_i + h) = y_{i+1}$$

$$f'(x_i) = y_i' \quad ; \quad f''(x_i) = y_i''$$

se pueden escribir las siguientes expresiones:

(11.4) $$y_i' = \frac{1}{2h}(y_{i+1} - y_{i-1}) + \circ(h^2)$$

(11.5) $$y_i'' = \frac{1}{h^2}(y_{i+1} - 2y_i + y_{i-1}) + \circ(h^2)$$

Haciendo en la expresión (11.2), $x = x_i + 2h$, en un caso, y $x = x_i - 2h$, en otro, y se combinan con las ecuaciones (11.3), de la siguiente manera:

$$f(x_i + 2h) = f(x_i) + 2hf'(x_i) + \frac{1}{2}(2h)^2 f''(x_i) + \frac{1}{6}(2h)^3 f'''(x_i) + \circ(h^4)$$
$$-2f(x_i + h) = -2f(x_i) - 2hf'(x_i) - \frac{2}{2}h^2 f''(x_i) + \frac{2}{6}h^3 f'''(x_i) + \circ(h^4)$$
$$2f(x_i - h) = 2f(x_i) - 2hf'(x_i) + \frac{2}{2}h^2 f''(x_i) - \frac{2}{6}h^3 f'''(x_i) + \circ(h^4)$$
$$-f(x_i - 2h) = -f(x_i) + 2hf'(x_i) - \frac{1}{2}(2h)^2 f''(x_i) + \frac{1}{6}(2h)^3 f'''(x_i) + \circ(h^4)$$

y, por último se suman, se obtiene:

(11.6) $$f'''(x_i) = \frac{1}{2h^3}[f(x_i + 2h) - 2f(x_i + h) + 2f(x_i - h) - f(x_i - 2h)] + \circ(h^2)$$

que es la **FORMULA DE APROXIMACIÓN POR DIFERENCIA CENTRAL** de la tercera derivada de *f(x)* en el punto $x = x_i$. Operando de la misma manera, es posible obtener las fórmulas correspondientes de las diferencias centrales para la cuarta, quinta, sexta, etc. derivadas de *f(x)*, en el punto considerado. Se expone la expresión correspondiente a la cuarta derivada, la que resulta:

(11.7) $$f^{IV}(x_i) = \frac{1}{h^4}[f(x_i + 2h) - 4f(x_i + h) + 6f(x_i) - 4f(x_i - h) + f(x_i - 2h)] + \circ(h^2)$$

Con las sustituciones ya realizadas, las dos últimas expresiones pueden ser escritas de la siguiente manera:

(11.8) $$y_i''' = \frac{1}{2h^3}(y_{i+2} - 2y_{i+1} + 2y_{i-1} - y_{i-2}) + \circ(h^2)$$

(11.9) $$y_i^{IV} = \frac{1}{h^4}(y_{i+2} - 4y_{i+1} + 6y_i - 4y_{i-1} - y_{i-2}) + \circ(h^2)$$

Utilizando el mismo procedimiento, es posible determinar fórmulas de *Aproximación por Diferencias Centrales* para las derivadas de mayor orden, no calculadas.

Se puede apreciar, pues, que utilizando las *Fórmulas de Aproximación por Diferencias Centrales* en vez de las correspondientes derivadas, es posible transformar una ecuación diferencial en un sistema de ecuaciones algebraicas lineales simultáneas para valores discretos de la variable.

IMPORTANTE: Hay que evitar el uso de las fórmulas de derivación numérica estudiadas; vale decir, las de *Aproximación por Diferencias Centrales*, en los problemas de valores iniciales, pues podrían ocasionar marcadas inestabilidades en los mismos, invalidando completamente los resultados obtenidos con este método. En cambio, en los problemas con condiciones de contorno, no se debe temer la inestabilidad ya que, la función que representa la solución buscada queda sujeta por las condiciones de contorno en ambos extremos del intervalo de integración.

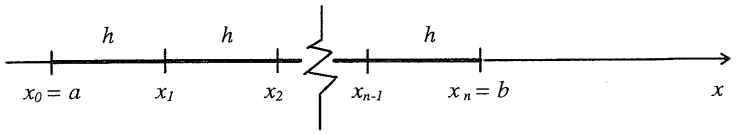
11.3.- MÉTODO DE LAS DIFERENCIAS FINITAS

En este método, como puede vislumbrarse, tanto la ecuación diferencial dada (11.1), como las condiciones de contorno que ella involucra, son reemplazadas por un sistema de ecuaciones algebraicas lineales simultáneas en la variable independiente, utilizando las *Fórmulas de Aproximación por Diferencias Centrales*, para las derivadas, el cual puede ser resuelto por cualquiera de los métodos, estudiados en su momento, destinados a resolver ecuaciones lineales simultáneas.

Supóngase que la ecuación diferencial y sus condiciones de contorno sea la siguiente:

$$\text{(11.10)} \qquad \begin{aligned} y'' &= p(x) \quad ; \quad a < x < b \\ y(a) &= A \\ y(b) &= B \end{aligned}$$

El intervalo $(a;b)$ se divide en n partes iguales de amplitud h. Representando solo la partición del intervalo, resulta:



- Figura 11.1 -

donde $h = (b - a) / n$, como se muestra en la figura 11.1; en la cual es:

$$\text{(11.11)} \qquad x_i = a + i\,h \quad ; \quad i = 0; 1; 2; ...; n$$

donde los extremos x_i de los subintervalos obtenidos de esta manera, reciben el nombre de **PUNTOS NODALES**; los puntos sobre el intervalo abierto $(a\ ;\ b)$ se llaman **PUNTOS INTERIORES**; y los puntos $x_0 ; x_n$ son llamados **PUNTOS FRONTERA**.

Si, para cada uno de los puntos nodales (excluidos los dos puntos frontera), se escriben las correspondientes fórmulas de aproximación por diferencias centrales, resultará para $x = x_i$:

$$\text{(11.12)} \qquad y_{i-1} - 2\,y_i + y_{i+1} = h^2 p_i \quad ; \quad i = 1; 2; ...; n-1$$

que es un sistema formado por n-1 ecuaciones con n+1 incógnitas o valores discretos de la variable y en $x = x_i$, para i=0; 1; 2; ...; n.

Se debe observar que la variable y toma valores desde $i = 1$, hasta $i = n - 1$; mientras que x lo hace para $i = 0$ hasta $i = n$.

Las dos condiciones de contorno de (11.10) proporcionan dos ecuaciones adicionales, con lo cual, en definitiva se obtiene un sistema determinado de n+1 ecuaciones con igual número de variables:

$$\text{(11.13)} \qquad \begin{cases} y_0 = A \\ y_0 - 2y_1 + y_2 = h^2 p_1 \\ y_1 - 2y_2 + y_3 = h^2 p_2 \\ \vdots \\ y_{n-3} - 2y_{n-2} + y_{n-1} = h^2 p_{n-2} \\ y_{n-2} - 2y_{n-1} + y_n = h^2 p_{n-1} \\ y_n = B \end{cases}$$

Este sistema de ecuaciones simultáneas recibe el nombre de **SISTEMA ANÁLOGO DE DIFERENCIAS FINITAS** que, como se puede verificar, es un conjunto de ecuaciones para valores discretos de la variable.

Sintetizando, se puede afirmar que, el *Sistema Análogo de Diferencias Finitas* se obtiene sustituyendo la ecuación diferencial dada y sus restricciones por las *Fórmulas de Aproximación por Diferencias Centrales* y luego, especificando las condiciones de contorno en términos de valores discretos de la variable introducida en el paso anterior, para poder así completar el número necesario y suficiente de ecuaciones que permita resolver el problema en cuestión.

Con el propósito de aportar mayores detalles sobre la aplicación del método de las **DIFERENCIAS FINITAS** se desarrollará a continuación el siguiente:

Ejemplo 11.1.- Sea resolver la ecuación diferencial $y'' = -12\,x^2$, en el intervalo $0<x<1$, con las condiciones de contorno $y(0) = y(1) = 0$.-

Solución: Haciendo $n = 2$; vale decir, $h = 0{,}5$ resulta la ecuación genérica:

$$y_{i-1} - 2\,y_i + y_{i+1} = -12\,x_i^2 h^2$$

y, por ende, con el agregado de las condiciones de contorno, el sistema de ecuaciones lineales simultáneas escritas en forma matricial, resulta:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{1}{2} \times 12 \times 0{,}5^2 \\ 0 \end{bmatrix} \tag{11.14}$$

En el caso de tomar $n = 4$; es decir, $h = 0{,}25$, el sistema resultante, luego de operar en forma idéntica a la mostrada, sería:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{1}{16} \times 12 \times 0{,}25^2 \\ -\frac{1}{16} \times 12 \times 0{,}5^2 \\ -\frac{1}{16} \times 12 \times 0{,}75^2 \\ 0 \end{bmatrix} \tag{11.15}$$

Los resultados obtenidos al resolver los sistemas (11.14) y (11.15), juntamente con los valores de la solución exacta, se encuentran sintetizados en la tabla 11.1.

	n=2 ; h=0,5	*n=4 ; h=0,25*	*Valores Reales*
x	*y*	*y*	*y*
0,00	0,000000	0,000000	0,000000
0,25		0,234375	0,246094
0,50	0,375000	0,421875	0,437500
0,75		0,421875	0,433594
1,00	0,000000	0,000000	0,000000

- Tabla N° 11.1 -

Se debe notar que, en la solución del punto medio, el error que se comete para el caso en que $h = 0{,}5$ es cuatro veces mayor que el que se comete en el caso en que se ha hecho $h = 0{,}25$.

♦♦♦ ●●● ♦♦♦

Con el objeto de profundizar en lo referente a la aplicación del **MÉTODO DE LAS DIFERENCIAS FINITAS**, veamos un problema que involucra otros términos de la ecuación, mediante el siguiente:

Ejemplo 11.2.- Resolver el problema que plantea un caso de oscilación, expresado por la siguiente ecuación diferencial:

$$y'' + x\,y = 1 - x \quad con \quad 0 < x < 1$$

con las condiciones de contorno:

$$y(0)=0 \;\; ; \;\; y(1)=0$$

Solución: Reemplazando los valores de y'' e y por sus correspondientes fórmulas de diferencias centrales, se obtiene:

$$\frac{1}{h^2}\left[y_{i-1} - 2y_i + y_{i+1}\right] + x_i y_i = 1 - x_i \quad ; \quad i = 1;2;\ldots;n-1 \tag{11.16}$$

Multiplicando esta última expresión por h^2, y sacando factor común, es:

$$y_{i-1} + \left(h^2 x_i - 2\right) y_i + y_{i+1} = h^2\left(1 - x_i\right) \quad ; \quad i = 1;2;\ldots;n-1 \tag{11.17}$$

que, con las condiciones de contorno, constituyen un sistema de $n+1$ ecuaciones con $n+1$ incógnitas. La solución de este problema para $n=5$ y $n=10$, está en la tabla N° 11.2.

x	*y (n=5)*	*y (n=10)*
0,00	0,000000	0,000000
0,10		- 0,029376
0,20	- 0,049750	- 0,049723
0,30		- 0,061970
0,40	- 0,067100	- 0,067031
0,50		- 0,065824
0,60	- 0,059380	- 0,059289
0,70		- 0,048397
0,80	- 0,034240	- 0,034166
0,90		- 0,017663
1,00	0,000000	0,000000

- Tabla N° 11.2 -

♦♦♦ ●●● ♦♦♦

11.4.- MÉTODO DE LAS DIFERENCIAS FINITAS PARA ECUACIONES DIFERENCIALES DE ORDEN SUPERIOR

El sistema análogo de diferencias finitas para ecuaciones diferenciales ordinarias de orden superior al segundo se puede obtener aplicando la misma operatoria, mediante el uso de las ecuaciones (11.4), (11.5), (11.8), (11.9) y otras similares a estas que pudieran determinarse para derivadas de orden superior al cuarto, haciendo uso de los mismos procedimientos que los empleados para deducir el sistema análogo de diferencias finitas.

En estos casos las condiciones de contorno incluyen casi siempre el valor de algunas derivadas de la solución en los extremos del intervalo de integración dado. Cuando las condiciones de contorno incluyen valores de las derivadas de la solución es necesario, para poder aproximarlas, el uso de las correspondientes fórmulas de derivación por diferencias centrales.

Las ecuaciones así obtenidas, constituyen parte del sistema análogo de diferencias finitas que describe el problema. Para ilustrar con mayores detalles el caso que se está tratando se presenta el siguiente:

Ejemplo 11.3.- Resolver la siguiente ecuación diferencial de cuarto orden:

$$\begin{aligned} &y^{IV} + f(x)\,y''' + g(x)\,y = p(x) \quad ; \quad 0 < x < 1 \\ &y(0) = y_0 \quad ; \quad y(l) = y_n \\ &y'(0) = y_0' \quad ; \quad y''(l) = y_n'' \end{aligned} \tag{11.19}$$

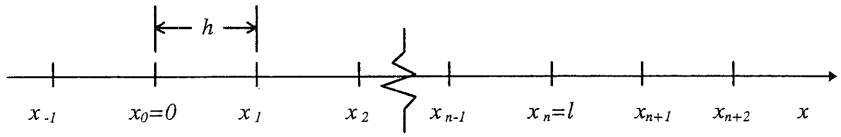
Solución: Dividiendo al intervalo de integración *(0;l)* en n partes iguales de amplitud h, será entonces:

$$x_i = i\,h \quad ; \quad i = 0;\ 1;\ 2;\ \ldots;\ n \tag{11.20}$$

Las derivadas incluidas en las condiciones de contorno del problema (11.17) se sustituyen por las fórmulas de diferencias centrales correspondientes, resultando la siguiente expresión:

$$\begin{aligned} &\left(1-\frac{h}{2}f_i\right)y_{i-2} - \left(4 - h\,f_i\right)y_{i-1} + \left(6 + h^4 g_i\right)y_i - \left(4 + h\,f_i\right)y_{i+1} + \\ &+\left(1+\frac{h}{2}f_i\right)y_{i+2} = h^4 p_i \quad ; \quad i = 1;2;\ldots;n-1 \end{aligned} \tag{11.21}$$

Esta ecuación se aproxima en cada punto interior del intervalo; de hecho, también se aproxima en los nodos; vale decir, para $i = 1;\ 2;\ \ldots;\ n\text{-}1$, como se muestra en la representación de la figura 11.2.

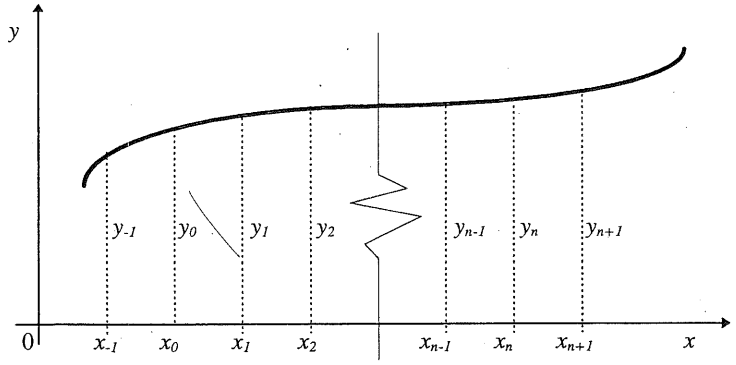


- Figura 11.2 -

Cuando $i = 1$, la ecuación (11.21) incluye un valor de y para $x = -h$, de igual modo cuando $i = n - 1$, incluye un valor de y para $x = l + h$. Estos son valores desconocidos, comprendidos fuera de la región de integración, pero, es incluido dentro del problema por el sistema análogo de diferencias finitas del problema en cuestión.

Es necesario para extrapolar el problema a una distancia h más allá de ambos extremos del intervalo de integración, incluir dos incógnitas más al sistema. La figura 11.3 muestra los extremos de integración numérica del intervalo $(0\ ;\ l)$ y cómo la solución podría extrapolarse más allá de los extremos originales.

Las propiedades de la solución en los puntos extremos están determinadas por las condiciones de contorno. La condición en $x=0$ da el valor para y_0 , además, la pendiente de la curva en el punto $(x_0\,;\,y_0)$.



- Figura 11.3 -

La fórmula por diferencia central para la primera derivada está dada en función de y_1 e y_{-1}, expresión (11.4), siendo entonces:

(11.22) $$y_i' = \frac{1}{2h}(y_{i+1} - y_{i-1}) \quad ; \quad i=0$$

En $x = l$ ya están dados los valores para $y(l)$, como así también, el de la derivada segunda $y''(l)$. La expresión para esta última se obtiene mediante la fórmula (11.5) que, en $x = l$, resulta:

(11.23) $$y_i'' = \frac{1}{h^2}(y_{i+1} - 2y_i + y_{i-1}) \quad ; \quad i = n$$

Las ecuaciones (11.22) y (11.23) junto con los valores dados para y_0 e y_n constituyen cuatro ecuaciones independientes adicionales que, con las $n - 1$ de (11.21) dan un conjunto de $n + 3$ ecuaciones algebraicas que pueden ser resueltas para las $n + 3$ incógnitas. Este conjunto es el *SISTEMA ANÁLOGO DE DIFERENCIAS FINITAS* para el problema de valores en la frontera dado por (11.19). Sintetizando, este puede ser escrito como se muestra a continuación:

(11.24)
$$\begin{cases} -y_{-1} + y_1 & = 2h y_0' \\ y_0 & = y_0 \\ \left(1 - \frac{h}{2} f_i\right) y_{i-2} - (4 - h f_i) y_{i-1} + (6 - h^4 g_i) y_i - & \\ \qquad - (4 + h f_i) y_{i+1} + \left(1 + \frac{h}{2} f_i\right) y_{i+2} & = h^4 p_i \\ y_n & = y_n \\ y_{n-1} - 2y_n + y_{n+1} & = h^2 y_n'' \end{cases}$$

para $i = 1;2;...;n-1$

El sistema propuesto puede ser resuelto por cualquiera de los métodos estudiados en el capítulo correspondiente a sistemas de ecuaciones lineales simultáneas, con un error por truncamiento del orden de $\circ(h^2)$, pues, todas las expresiones por diferencias centrales empleadas en su deducción tienen dicha aproximación.

Una manera muy sencilla de mejorar la precisión, es hacer decrecer la magnitud del incremento h.

♦♦♦ ●●● ♦♦♦

BIBLIOGRAFÍA

1.- LIBROS DE PROPÓSITO GENERAL

1.1. **ANÁLISIS NUMÉRICO**
D. Kincaid y W. Cheney — Addison - Wesley — 1993

1.2. MÉTODOS NUMÉRICOS APLICADOS CON SOFTWARE
S. Nakamura — Prentice - Hall — 1992

1.3. **INTRODUCCIÓN A LOS MÉTODOS NUMÉRICOS CON PASCAL**
L. V. Atkinson y P. J. Harley — Addison - Wesley — 1991

1.4. **MÉTODOS NUMÉRICOS**
F. Scheid y R. Di Constanzo — Mc Graw - Hill — 1991

1.5. **FÓRMULAS Y TABLAS DE MATEMÁTICAS APLICADAS**
M. Spiegel y L. Abellanas — Mc Graw - Hill — 1988

1.6. **MÉTODOS NUMÉRICOS BÁSICOS**
R. E. Scraton — Mc Graw - Hill — 1987

1.7. **MÉTODOS NUMÉRICOS PARA INGENIEROS**
S. Chapra y R. Canale — Mc Graw - Hill — 1987

1.8. **ANÁLISIS NUMÉRICO**
Burden y Faires — Grupo Ed. Iberoamérica — 1985

1.9. THEORETICAL NUMERICAL ANALYSIS
P. Linz — John Wiley & Sons — 1979

1.10. **AN INTRODUCTION TO NUMERICAL ANALISIS**
K. Atkinson — John Wiley & Sons — 1978

1.11. TÉCNICAS RECURSIVAS EN PROGRAMACIÓN
D. W. Barron — Ed. KAPELUSZ — 1976

1.12. **ANÁLISIS NUMÉRICO**
S. Conte y De Boor — Mc Graw - Hill — 1974

1.13. **CALCULO NUMÉRICO Y GRÁFICO**
M. Sadosky — Librería del Colegio — 1973

1.14. **ELEMENTOS DE ANÁLISIS NUMÉRICO**
P. Henrici — Ed. Trillas — 1972

1.15. **INTRODUCCIÓN AL ANÁLISIS NUMÉRICO**
A. Ralston — Limusa - Wiley — 1970

1.16. **MÉTODOS NUMÉRICOS APLICADOS**
M. James, G. Schmidt y otros — Repr. y Serv. de Ing. — 1970

1.17. **ANÁLISIS NUMÉRICO**
M. SALVADORI Y M. BARON — Editorial CECSA — 1969

1.18. **MÉTODOS NUMÉRICOS Y PROGRAMACIÓN FORTRAN**
D. Mc Cracken y W. Dorn — Limusa - Wiley — 1966

2.- LIBROS DE PROPOSITO ESPECIAL

2.1. Computational Methods in Ordinary Differential Equations
J. D. Lambert — John Wiley & Sons — 1977
2.2. 1130 Scientific Subroutine Package
Programmer´s Manual — IBM — 1967
2.3. Solución Mediante Computadoras de SISTEMAS ALGEBRAICOS LINEALES
G. Forsythe Y C. Moler — Ed. EUDEBA — 1966

ÍNDICE

A

AITKEN, 123
ajuste, 211
  mediante curvas, 212
  mediante polinomios, 213
    metodología, 213, 215
algoritmo/s
  clasificación de los, 23
  directos, 23
  indirectos, 24
  infinitos, 25
algorítmico/s
  procesos, 21
análisis del error en los métodos de integración numérica, 285
aproximación/es
  numéricas, 41
  intermedias, 26
autovalores, 156
autovectores, 156

B

BESSEL, 204
besselianos
  coeficientes, 207
borde,
  problemas con condiciones de, 293

C

coeficientes besselianos, 207
convergencia, 26, 34
  grado de, 26, 35
  velocidad de, 26
constante/s
  absolutas, 42
  relativas, 42
contorno,
  problemas con condiciones de, 293

D

derivación, 219
  algoritmo de *HORNER*, 222
  la serie de *TAYLOR*, 228
  por fórmulas de interpolación, 220
diferencias centrales,
  formulas de aproximación por, 295
diferencias finitas,
  método de las, 298
  sistema análogo de, 299
  para ecuaciones diferenciales de orden superior, 302

E

ecuación/es
  característica,
ecuaciones diferenciales
  métodos que comienzan por sí mismos, 249
    de *EULER*, 250
    de *RUNGE-KUTTA*, 258
    de *RUNGE-KUTTA* para ec. de orden superior, 267
    de *RUNGE-KUTTA* para sistemas de ec. diferenciales, 266
    modificado de *EULER*, 255
    predictor corrector, 256
  métodos que no comienzan por sí mismos, 273
    método de *MILNE*, 274
      generalizado de *MILNE*, 278
    método de *HAMMING*, 281
      ecuación de corrección, 283
      modificado de *HAMMING*, 284
  métodos con condiciones de contorno,
ecuaciones lineales simultáneas, 137
  compatibilidad, 139
  grado de indeterminación, 139
  indeterminación, 139

método de
la matriz inversa, 149
las ecuaciones de error, 147
eliminación de *GAUSS*, 141
GAUSS-JORDAN, 144
GAUSS-SEIDEL, 150
sistemas homogéneos, 138, 154
método/s de
las potencias, 160
FADDEEV-LEVERRIER, 157
polinomio,
característico, 156
de *FADDEEV-LEVERRIER*, 157
valores característicos, 156
vectores característicos, 156
sistemas no homogéneos, 138, 140
sistemas ortogonales, 165
condición de ortogonalidad, 165
método
analítico, 166
de *GRAM-SCHMIDT*, 168
epsilon de la computadora, 43, 52
error/es, 41, 61
absoluto, 62
cálculo de errores mediante fórmulas diferenciales, 66
clasificación de los, 45
cota de, 47
en las operaciones, 61, 68
en los algoritmos infinitos, 63
fórmula fundamental, 67
groseros, 46
local, 279
por abstracción, 47
por inherencia, 47
por redondeo, 50
por truncamiento, 48
problema inverso, 81
propagación de los, 55
relativo, 63
tolerancia del, 64
total, 55, 279
estabilidad, 56
estabilidad en los métodos de integración, 278
estabilidad relativa, 279
EULER, 241, 250, 255
EULER-MACLAURIN, 241
EVERETT, 208

F

FADDEEV-LEVERRIER, 157
formula/s de
aproximación por diferencias centrales, 295
de los $^3/_8$ de *SIMPSON*, 238
de *SIMPSON*, 236
EULER-MACLAURIN, 241
GREGORY, 243
los trapecios, 233
fracciones terminales finitas, 41, 61
FROBENIUS, 139
frontera,
problemas con valores en la, 293

G

GAUSS, 141, 144, 150, 201
GAUSS-JORDAN, 144
GAUSS-SEIDEL, 150
gráficos de procesos, 71
aplicaciones de, 74
metodología de, 71
GRAM-SCHMIDT, 168
GREGORY, 182, 188, 243

H

hipótesis de precisión infinita, 78
HAMMING, 281
HORNER, 222

I

integración, 232
fórmula de
EULER-MACLAURIN, 241
GREGORY, 243
los trapecios, 233
SIMPSON, 236
$^3/_8$ de *SIMPSON*, 238
interpolación, 177

fórmulas de
BESSEL, 204
EVERETT, 208
GAUSS, 201
LAGRANGE, 193
lineal y cuadrática, 186
NEWTON-GREGORY, 182, 188
STIRLING, 206
parabólica progresiva, 197
inversa, 210
tablas
de valores equidistantes, 179
diferencias atrasadas, 189, 231
diferencias avanzadas, 180, 231

J

JORDAN, 144

K

KUTTA, 258

L

LAGRANGE, 193
lagrangianos
coeficientes, 195
LEVERRIER, 157

M

MACLAURIN, 241
matemática
aplicada, 13, 14
pura, 13, 14
matriz/ces
diagonalmente dominantes, 151
orlada, ampliada o aumentada, 139
traza de una, 157
método/s
aproximados, 13, 14
de discretización, 29
de enumeración, 28
de simulación, 30
de tanteos, 29, 91
de las diferencias finitas, 298
exactos, 13, 14
gráficos, 28
iterativos, 31
numéricos, 13, 14
MILNE, 274
modelo/s
codificación del, 15, 17
desarrollo de un, 15, 17
optimización de un, 15, 17
procesamiento del, 15, 17
validación del, 15, 17

N

NEWTON-RAPHSON, 98
número/s
aproximados, 43
exactos, 42

P

parada
programada, 23
por diferencia absoluta, 63
global, 64
individual, 65
por diferencia relativa, 64
global, 65
individual, 65
paso/s
de orden k, 32
elementales, 24
polinomio/s
característicos, 156
problema
estudio y formulación del, 15
resolución de, 14, 26
proceso/s
algorítmicos, 21
convergentes, 35
iterativos, 31
recursivos, 33
puntos,
frontera, 298
interiores, 298
nodales, 298

R

raíces de ecuaciones, 85
  acotación de,86
    intervalo de, 86
  aproximación de, 86
  algebraicas, 129
  métodos de resolución, 91
    delta cuadrado de *AITKEN*, 123
    interpolación lineal, 96
    intervalo medio, 92
    iteración, 110
    mixto, 105
    NEWTON-RAPHSON, 98
    segundo orden de *NEWTON*, 127
    tanteos, 91
  separación de, 86
redondeo, 51
  aleatorio, 52
  inferior, 51
  simétrico, 51
RAPHSON, 98
ROUCHE-FROBENIUS, 139
RUNGE-KUTTA, 258

S

SCHMIDT, 168
selección de un método de integración
  numérica, 288
SEIDEL, 150
SIMPSON, 236
sistema/s
  análogo de diferencias finitas, 299
  lineales, 137
STIRLING, 206
STURM, 132

T

TAYLOR, 48, 228
TAYLOR
  desarrollo en serie de, 48
teorema/s de
  ROUCHE-FROBENIUS, 139
  STURM, 132
  RUFFINI-ABEL, 85
teoría general de la iteración, 109, 116
  ordenes de convergencia, 117, 120, 122

V

valore/s
  característicos, 156
variable/s, 42
vectore/s
  característicos, 156

UNIVERSIDAD NACIONAL DEL NORDESTE
FACULTAD DE CIENCIAS EXACTAS Y NATURALES Y AGRIMENSURA
B-16 | 12883
SIGNATURA TOPOGRAFICA

Se terminó de imprimir en el mes de marzo de 1997,
en los talleres de Graffit SRL,
Buenos Aires, Argentina.