

ESTIMACIÓN DEL ERROR EN CIERTA CLASE DE MÉTODOS NUMÉRICOS COMBINADOS DE ESTRATEGIA RECORRECTORA

AYMÉE MARRERO* MARTA LOURDES BAGUER†

Recibido/Received: 8 Jun 2004

Resumen

El siguiente trabajo brinda un análisis del orden de los errores de aproximación y de redondeo que se cometen al combinar esquemas numéricos de estrategia recorrectora, de la familia Adams–Moulton–Bashfort. Estos métodos se utilizan en el tratamiento numérico del problema de estimación en modelos definidos por sistemas dinámicos de EDO.

Nuestro propósito es disponer de una “familia o clase” de métodos que sean eficientes o al menos competitivos al abordar la resolución numérica de este tipo de problemas,

Esta estrategia es compatible con las tendencias actuales que emplean modelos con funciones objetivos mínimo-cuadradas para el ajuste de curvas, que usan métodos de optimización sin y con restricciones, para resolver un problema discretizado por un esquema numérico que puede ser implícito, semi-implícito o explícito.

Palabras clave: Métodos numéricos de estrategia recorrectora, multipasos, rigidez, errores de aproximación y redondeo.

Abstract

This paper shows an analysis of the order for the approximation and roundoff errors in combined numerical scheme with recorrector strategy of Adams–Moulton–Bashfort family. These methods are used for the numerical solution of the estimation problem in models defined by dynamical ODE. Our interest is avail oneself of “family or kind” of methods which will be efficient or at least competitive for the numerical solutions of this type of problems. This strategy is consistent with the present tendency

*Departamento de Matemática Aplicada, Facultad de Matemática y Computación, Universidad de la Habana, San Lázaro y L, La Habana CP 10400, Cuba E-Mail: aymee@matcom.uh.cu

†Misma dirección/*same address*. E-Mail: mbaguer@matcom.uh.cu.

in the use of models with minimal square objective function for the adjustment via optimization algorithms with and without constrained and explicit, semi-explicit or implicit numerical methods.

Keywords: Recorrector strategy numerical methods, multi-step, stiffness, roundoff errors.

Mathematics Subject Classification: 65G50, 65G99.

1 Introducción

Las estrategias actuales consideran encontrar la solución numérica del siguiente problema:

$$\begin{aligned} \dot{y} &= f(x, y), \quad a \leq x \leq b \\ &\text{con } y(a) \text{ dado} \end{aligned} \tag{1}$$

empleando, por ejemplo, un método de un paso que comienza con $y(a) = y_0$ para $x_0 = a$ y genera las sucesivas aproximaciones y_n de $y(x_n)$ sobre la malla $x_0 < x_1 < \dots < x_m = b$, asumiendo generalmente que $f(x, y)$ es continua y cumple la condición de Lipschitz sobre el intervalo $[a, b]$ para todo y , con el objetivo de garantizar la existencia y unicidad de la solución $y(x)$. Para situaciones clásicas, $L(b - a)$ (L representa la constante de Lipschitz) no es "grande" y por tanto, métodos numéricos clásicos como los de Adams o Runge-Kutta son muy eficientes, porque sólo por excepción se requiere una reducción drástica del paso h_{n+1} . Situación bien diferente se produce cuando $L(b - a)$ es "grande" ya que los métodos clásicos para tratar de resolver el problema, sufren tal gama de restricciones sobre el tamaño del paso que se convierten en ineficientes. Esto es lo que usualmente se denomina rigidez. Sin embargo, la mayoría de los autores dedicados al tema ([1], [2], [6], [5], [14], [15]) dan una definición de rigidez que no sólo involucra el problema matemático. La situación de la rigidez es, desafortunadamente, más compleja y está relacionada con el método numérico que se utiliza, su estabilidad, precisión, las fórmulas que lo definen e incluso la implementación computacional. Existen esquemas numéricos eficientes para resolver sistemas de ecuaciones diferenciales rígidos, que establecen una política de cambio automático del paso de integración, basada en el control del orden del error de aproximación y de truncamiento local y utilizan derivadas de órdenes superiores. El precursor de estos estudios fue W.C.Gear, en la década del 70 (ver [9]) con su método, que utiliza las diferencias retrógradas y ha demostrado su eficiencia. En la actualidad hay dos vertientes, los métodos de un paso y los multipasos; entre los primeros se destacan los de Runge-Kutta, los llamados "Order Stars", y en general los de extrapolación; entre los multipasos se destacan los de Adams, los tipo Predictor-Corrector, el de segundo barrido de Dahlquist, los BDF y los de "One-leg". Por supuesto, que entre los multiderivadas aparecen los de Enright, los extendidos de Cash, y los que se obtienen por combinaciones y extensiones de los anteriores. Ahora bien, abordar el problema de estimación de parámetros en modelos rígidos no se restringe, por supuesto, a resolver un sistema de ecuaciones diferenciales (ver [12]) con niveles altos de rigidez dados por causas tan diversas como números de condición grandes, problemas de estabilidad, de mal planteamiento, etc; sino que es mucho más que esto.

Con la idea de validar una “familia” de métodos que se requerían fueran eficientes o al menos competitivos al abordar la resolución numérica de este tipo de problemas, se utilizó una estrategia compatible con la tendencia actual que emplea modelos con funciones objetivos mínimo-cuadradas para el ajuste de curvas y usa métodos de optimización sin y con restricciones (ver [3], [4], [7] por ejemplo), entre los que se destacan los de la familia de Newton, los quasi-Newton, el de matrices poco densas, los de tipo Broyden, entre otros para resolver un problema discretizado por un esquema numérico que puede ser implícito, semi-implícito o explícito.

De ahí el interés particular de este trabajo, como parte de una estrategia para el tratamiento numérico del problema de estimación los parámetros en este tipo de modelos (ver [16], [17], [18]), en el estudio del comportamiento de los errores de aproximación y de redondeo al combinar esquemas numéricos.

2 Planteamiento del problema

Para encontrar la solución numérica aproximada del problema de Cauchy:

$$\begin{aligned} \dot{y}(x) &= f(x, y(x)), \\ y(a) &= A \end{aligned} \quad (2)$$

sobre el intervalo $[a, b]$, son muchos y muy variados los métodos que han sido desarrollados. Todos generan un conjunto de puntos $\{x_0, x_1, \dots\}$ en dicho intervalo, los que pueden tener o no igual longitud, y cuya forma general de generación de los mismos está usualmente dada por una regla del tipo:

$$\begin{aligned} x_n &= x_{n-1} + h_n, \quad n = 1, 2, \dots \\ x_0 &= a \end{aligned}$$

donde la notación y_n representa una aproximación a la solución $y(x)$ del problema anterior (2) en la malla de puntos x_n , es decir:

$$y_n = y(x_n)$$

Una solución de (2) puede ser escrita como:

$$y(x_{n+1}) = y(x_n) + \int_{x_n}^{x_{n+1}} \dot{y}(t) dt = y(x_n) + \int_{x_n}^{x_{n+1}} f(t, y(t)) dt \quad (3)$$

El conocido esquema Predictor-Corrector (**PC**) se caracteriza por su estrategia de calcular una primera aproximación predictora, denotada de forma general como p_i , por la fórmula explícita de Adams (atribuidas a Adams-Bashforth), evaluar en la función dicha predicción, aplicar la fórmula implícita correctora (atribuidas a Adams-Moulton) con la primera aproximación que generó el predictor y evaluar nuevamente la función (ver [20]).

Convencionalmente, cuando se va a realizar la implementación computacional de un método numérico se tiene muy en cuenta, el número de evaluaciones de $f(x, y)$ que requiere

el mismo. El procedimiento **PC** necesita dos evaluaciones en cada paso (más que el método de Adams, por ejemplo que sólo necesita una), pero es más preciso y mucho mejor con respecto a la propagación del error (ver capítulo 8 del [20]). Además, sus estimaciones del error son más confiables, lo que conlleva a una más efectiva selección del tamaño del paso. Para el **PC** se pueden utilizar esquemas del mismo orden para el predictor y el corrector, en este trabajo se utilizó un esquema predictor de orden k que usa los valores y_n y f_{n+1-j} para $j = 1, 2, \dots, k$ y un esquema corrector de orden $k + 1$ que usa los mismos valores además de p_{n+1} .

Con el objetivo de tratar problemas de estimación de parámetros (**EP**) definidos por sistemas de ecuaciones diferenciales ordinarias (**EDO**), este trabajo propone una estrategia combinada de tipo **PC** por dos razones fundamentales:

1. En la literatura se usan preferentemente los métodos de un paso y no se ha explotado el empleo de los multipasos para estos fines, lo que representa una línea novedosa de investigación.
2. En particular, en el tratamiento de solución de sistemas de **EDO** rígidas, donde los métodos de optimización se usan generalmente como “cajas negras”, se emplean con gran éxito los métodos multipasos del tipo **PC** (ver, por ejemplo [9]).

La idea general es combinar el clásico esquema **PC** de Adams-Bashford-Moulton con un esquema multipaso segunda derivada propuesto por Enright, con eficiencia garantizada para modelos rígidos, lo que genera un esquema tipo Predictor-Corrector-Recorrector, para el cual es preciso estimar los errores de redondeo y truncamiento por su influencia en la estabilidad y convergencia del método que se pretende proponer.

Considérese el siguiente problema discreto en el cual, el sistema de ecuaciones diferenciales del problema continuo (ver [10], [11] y [19], este último por publicar) es sustituido por un esquema multipaso de orden variable Q_i y de tamaño de paso variable h_i :

$$\begin{aligned} \min J_k(u) &= \sum_{i=0}^k \tilde{\varphi}_i [z_i, \bar{z}_i], \\ x_{i+1} &= x_i + h_i F_i(x_i, x_{i-1}, \dots, x_{i+1-Q_i}; u), \quad i = 0, 1, 2, \dots, k-1, \\ x_0 &= l(u), \\ z_i &= g_i(x_i, u), \quad i = 0, 1, 2, \dots, k. \end{aligned} \tag{4}$$

Se considera cierta partición de integración, que debe contener el conjunto de tiempos de medición $\{\tau_1, \tau_2, \dots, \tau_s\}$ de modo que, a cada índice de medición $j \in \{1, 2, \dots, s\}$ corresponde un índice de integración $i \in \{0, 1, 2, \dots, k\}$. Se define cierta correspondencia entre los índices y un conjunto que contiene los índices de integración correspondientes a las mediciones lo que permite expresar la función objetivo $\tilde{\varphi}_i$ dependiendo de la función característica de dicho conjunto de índices.

Las funciones F_i pueden ser tomadas de diferentes formas; en este caso se trabajó con las fórmulas de integración de Adams consideradas en la definición de un esquema del tipo Predictor-Corrector. Entre las muchas maneras en que se puede tratar la política de orden en un esquema numérico, se seleccionó la de incrementar el orden Q_i paso a paso hasta

alcanzar un orden máximo del Corrector Q_{\max} , y luego mantenerlo como $Q_i = Q_{\max} - 1$ para $i \geq i_0$.

2.1 Formulación general del esquema Predictor-Corrector

El llamado esquema Predictor-Corrector (**P-C**) se caracteriza por su estrategia de calcular una primera aproximación predictora, denotada de forma general como p_i , por la fórmula explícita de Adams (atribuidas a Adams-Bashforth), evaluar en la función dicha predicción, aplicar la fórmula implícita correctora (atribuidas a Adams-Moulton) con la primera aproximación que generó el predictor y evaluar nuevamente la función (ver [20]). Si se tomase, por ejemplo, el método de Euler (de orden 1) como Predictor y la regla trapezoidal (de orden 2) como Corrector, la estrategia o método Predictor-Corrector tendría la forma

$$\begin{aligned} p_{n+1} &= y_n + h_{n+1} f_n, \\ y_{n+1} &= y_n + \frac{h_{n+1}}{2} [f(x_{n+1}, p_{n+1}) + f_n] \\ f_{n+1} &= f(x_{n+1}, y_{n+1}). \end{aligned}$$

2.1.1 Fórmulas de Adams-Bashforth

Los métodos de Adams aproximan la solución reemplazando $f(t, y(t))$ por un polinomio de interpolación para calcular los valores de f_i , el cual se integra. Estos métodos son de los más efectivos actualmente y tienen además, la ventaja de ser suficientemente simples, lo que los hace muy comprensibles y por ende, muy útiles.

La fórmula de Adams-Bashforth de orden k en x_n usa un polinomio $\wp_{k,n}(x_{n+1-j}) = f_{n+1-j}$ para $j = 1, 2, \dots, k$, por lo que

$$y_{n+1} = y_n + \int_{x_n}^{x_{n+1}} \wp_{k,n}(t) dt \quad (5)$$

y para valorar cuán bien y_{n+1} aproxima a $y(x_{n+1})$ desarrolla la forma de Lagrange del polinomio de interpolación:

$$\wp_{k,n}(x) = \sum_{i=1}^k l_i(x) f_{n+1-i}$$

donde:

$$l_i(x) = \prod_{j=1, j \neq i}^k \frac{x - x_{n+1-j}}{x_{n+1-i} - x_{n+1-j}}, \quad i = 1, \dots, k.$$

Nótese que si se sustituyen las expresiones anteriores en (5) se obtiene:

$$y_{n+1} = y_n + \sum_{i=1}^k f_{n+1-i} \int_{x_n}^{x_{n+1}} l_i(t) dt.$$

Todo lo anterior lleva a la forma usual de la fórmula de Adams-Bashforth,

$$y_{n+1} = y_n + h_{n+1} \sum_{i=1}^k \alpha_{k,i} f_{n+1-i}, \quad (6)$$

donde

$$\alpha_{k,i} = \frac{1}{h_{n+1}} \int_{x_n}^{x_{n+1}} l_i(t) dt = \int_0^1 l_i(x_n + sh_{n+1}) ds.$$

Nota: Esta última igualdad se cumple si consideramos el cambio de variable $s = (t - x_n)/h_{n+1}$.

Por supuesto, que por ser y_{n+1} una aproximación a la solución, tiene dos fuentes de error : la primera es la que generan los valores acumulados y_n, y_{n-1}, \dots que se conoce como error de estimación y la segunda, la llamada error de truncamiento local, que lo provoca el hecho de aproximar $\dot{y}(x)$ por un polinomio de interpolación, que se denota por ζ_{n+1}^p y se define como

$$\zeta_{n+1}^p = \int_{x_n}^{x_{n+1}} \dot{y}(t) - \wp_{k,n}(t) dt \quad (7)$$

Si $y(x) \in \mathcal{C}^{k+1}[x_{n+1-k}, x_{n+1}]$ se calcula dicho error en dependencia de una función continua que depende de x y que permite asegurar que si $y(x)$ es suficientemente suave, para todo n , el error de truncamiento local ζ_{n+1}^p es de orden $\mathbf{O}(H^{k+1})$, donde H es el paso de máxima longitud (el Teorema 2 del capítulo 2 del [20], brinda una útil expresión). Teóricamente, la idea básica que subyace en el cálculo de este error es la de comparar el resultado de un polinomio de interpolación de grado k con otro de grado $k + 1$. De modo que (7) se escribiría como:

$$\zeta_{n+1}^p = \int_{x_n}^{x_{n+1}} \wp_{k+1,n}(t) - \wp_{k,n}(t) dt + \int_{x_n}^{x_{n+1}} \dot{y}(t) - \wp_{k+1,n}(t) dt$$

y el término

$$\int_{x_n}^{x_{n+1}} \dot{y}(t) - \wp_{k+1,n}(t) dt = \mathbf{O}(H^{k+2})$$

podría ser ignorado en la estimación de ζ_{n+1}^p por ser $\mathbf{O}(H^{k+2})$. Pero en la práctica esto no es útil, pues no se conocen los $\wp_{k+1,n}(t)$ y $\wp_{k,n}(t)$ y por tal razón, se supone que

$$y(x_{n+1-i}) = y_{n+1-i}$$

entonces,

$$\wp_{k,n}(t) = \mathbf{P}_{k,n}(t) \text{ y } \wp_{k+1,n}(t) = \mathbf{P}_{k+1,n}(t)$$

lo que permite concluir que:

$$\zeta_{n+1}^p = \int_{x_n}^{x_{n+1}} \mathbf{P}_{k+1,n}(t) - \mathbf{P}_{k,n}(t) dt + \mathbf{O}(H^{k+2}).$$

Es decir, el error de truncamiento local es fácilmente estimado al hallar la diferencia entre los valores de y_{n+1} obtenidos con una fórmula de orden k y una de orden $k + 1$, especialmente si se utiliza un polinomio de interpolación basado en las diferencias retrógradas (para detalles ver, capítulo 4 del [20]) que usualmente se acota superiormente por cierta tolerancia.

Este análisis que se ha hecho sobre ζ_{n+1}^p indica que, las fórmulas de orden superior aproximan la solución mejor que las de menor orden, sobre todo para tamaños de pasos pequeños; ya que, en el caso de longitud de paso constante, las fórmulas de mayor orden usan más puntos de interpolación para definir el polinomio y en el caso de longitud de paso variable, reduciendo la longitud del paso, se reduce el error. Nótese además, que si $y(x)$ varía suavemente sobre el intervalo de interpolación entonces, las derivadas de orden superior son pequeñas y por ende, el error de truncamiento también lo es. Aunque lo contrario es cierto, cuando $y(x)$ varía rápidamente, son precisamente estas soluciones (las que varían suavemente) las que resultan relativamente fáciles de calcular numéricamente.

2.1.2 Fórmulas de Adams-Moulton

Una buena estrategia para obtener una mejor aproximación $y(x_{n+1})$ de la solución, es considerar el valor y_{n+1} de (5) como un primer intento o como un valor “predicho”, que como se explicó al inicio de esta sección usualmente se denota como p_{n+1} e incorporarlo al polinomio de interpolación. Para ello, se recomiendan las fórmulas de Adams-Moulton.

La fórmula de Adams-Moulton de orden k en x_n usa también un polinomio de interpolación, esta vez denotado por $\wp_{k,n}^*(x)$, que interpola los k valores de las derivadas,

$$\begin{aligned}\wp_{k,n}^*(x_{n+1-j}) &= f_{n+1-j}, \quad j = 1, \dots, k-1, \\ \wp_{k,n}^*(x_{n+1}) &= f(x_{n+1}, p_{n+1})\end{aligned}$$

para obtener una solución aproximada y_{n+1} , por la expresión:

$$y_{n+1} = y_n + \int_{x_n}^{x_{n+1}} \wp_{k,n}^*(t) dt \quad (8)$$

La forma de Lagrange de (8) se obtiene de forma similar a la que se mostró para el caso de las fórmulas de Adams-Bashforth y su expresión es

$$y_{n+1} = y_n + h_{n+1} \sum_{i=1}^{k-1} \alpha_{k,i}^* f_{n+1-i} + h_{n+1} \alpha_{k,0}^* f(x_{n+1}, p_{n+1}),$$

donde:

$$\alpha_{k,i}^* = \int_0^1 l_i^*(x_n + sh_{n+1}) ds.$$

y

$$l_i^*(x) = \prod_{j=1, j \neq i}^{k-1} \frac{x - x_{n+1-j}}{x_{n+1-i} - x_{n+1-j}}, \quad i = 1, \dots, k.$$

De manera análoga a la que se expuso en el caso del Predictor, se estima el error de truncamiento local para la fórmula correctora, sólo que este contendrá al error del predictor y por eso puede demostrarse (ver capítulo 6 del [20]) que el error de truncamiento del Corrector ζ_{n+1} caracteriza al error de truncamiento local del método.

Convencionalmente, cuando se va a realizar la implementación computacional de un método numérico se tiene muy en cuenta, el número de evaluaciones de $f(x, y)$ que requiere el mismo. El procedimiento Predictor-Corrector (**P-C**) necesita dos evaluaciones en cada paso (más que el método de Adams, por ejemplo que sólo necesita una), pero es más preciso y mucho mejor con respecto a la propagación del error (ver capítulo 8 del [20]). Además, sus estimaciones del error son más confiables, lo que conlleva a una más efectiva selección del tamaño del paso. Para el **P-C** se pueden utilizar esquemas del mismo orden para el predictor y el corrector. En este trabajo se utilizó un esquema predictor de orden k que usa los valores y_n y f_{n+1-j} para $j = 1, 2, \dots, k$ y un esquema corrector de orden $k + 1$ que usa los mismos valores además de p_{n+1} .

2.1.3 Esquema Predictor-Corrector

Finalmente, la expresión general que tendrá el esquema **P-C** de orden $k + 1$ (orden que define el polinomio de interpolación de la fórmula correctora) es:

$$p_{n+1} = y_n + h_{n+1} \sum_{j=1}^k \alpha_{k,j} f(x_{n+1-j}, y_{n+1-j}), \quad (9)$$

$$y_{n+1} = y_n + h_{n+1} \sum_{j=1}^k \alpha_{k+1,j}^* f(x_{n+1-j}, y_{n+1-j}) + h_{n+1} \alpha_{k+1,0}^* f(x_{n+1}, p_{n+1}) \quad (10)$$

2.1.4 Teorema de Convergencia del método P-C

Es conocido del [20], donde aparece un minucioso análisis de la convergencia y estabilidad del procedimiento **P-C** tanto para tamaño de paso constante como variable, los resultados de convergencia de este método.

A continuación, se formulan sólo algunos de estos resultados.

Definiendo, para el esquema anterior (9), (10), el error global en x_i como la diferencia entre la solución exacta en x_i y la aproximación calculada

$$\mathbf{e}_i = y(x_i) - y_i, \quad (11)$$

e introduciendo una sucesión de números $\{\mathbf{E}_i\}$ para acotar el error global, de manera que

$$|\mathbf{e}_j| \leq \mathbf{E}_j, \text{ para } j = 0, 1, \dots \text{ y para todo } i,$$

se pueden formular los siguientes resultados:

Lema 1. *Sea la sucesión de números $\{\mathbf{E}_i\}$, tal que \mathbf{E}_0 es dado y $\mathbf{E}_k = \mathfrak{L} \mathbf{E}_{k-1} + \delta$ para $k = 1, 2, \dots$ y donde \mathfrak{L} y δ son constantes positivas con $\mathfrak{L} \neq 1$. Entonces*

$$\mathbf{E}_k = \mathfrak{L}^k \mathbf{E}_0 + \frac{\mathfrak{L}^k - 1}{\mathfrak{L} - 1} \delta \text{ para } k \geq 1.$$

Teorema 1. *Considérese que el esquema **P-C** ha sido aplicado para resolver el problema (2) con valores iniciales y_i que satisfacen que*

$$|y(x_i) - y_i| \leq \mathbf{E}_0, \quad i = 0, 1, \dots, k-1.$$

Si además se asume que $f_y(x, y)$ es continua y acotada por la constante L de Lipschitz, entonces para cualquier $x_n \in [a, b]$ se cumple que

$$|y(x_n) - y_n| \leq \left[\mathbf{E}_0 + \frac{\delta}{hL (\alpha^* + hL|\alpha_{k+1,0}^*|)} \right] \times \exp [(x_n - a) hL (\alpha^* + hL|\alpha_{k+1,0}^*|)] \quad (12)$$

donde

$$\begin{aligned} \delta &= \max_n |h\alpha_{k+1,0}^* f_y \zeta_{n+1}^p + \zeta_{n+1}| \\ \alpha &= \sum_{j=1}^k |\alpha_{k,j}|, \quad \alpha^* = \sum_{j=1}^k |\alpha_{k+1,j}^*|. \end{aligned}$$

Este teorema implica la convergencia del esquema **P-C**, porque la desigualdad (12) da una restricción para el peor error: $\max_{a \leq x_n \leq b} |y(x_n) - y_n|$ en $[a, b]$. Sin embargo, el siguiente corolario brinda una expresión que hace evidente que dicho error tiende a cero.

Corolario 1. *Además de las hipótesis asumidas en el Teorema anterior, si suponemos que $y(x) \in C^{k+2}[a, b]$ y que los valores iniciales satisfacen que*

$$|y(x_i) - y_i| \leq \mathbf{c}_1 h^{k+1}, \quad i = 0, 1, \dots, k-1$$

para cierta constante \mathbf{c}_1 y todo $0 < h \leq h_0$.

Entonces, existe cierta constante \mathbf{c}_2 , tal que para todo $x_n \in [a, b]$ se cumple

$$|y(x_n) - y_n| \leq \mathbf{c}_2 h^{k+1}$$

lo que significa que el esquema es uniformemente convergente de orden $k+1$.

2.2 Estrategia combinada

En la segunda mitad de la década de los 70, autores como Enright y Henrici, propusieron una serie de esquemas implícitos no lineales multipaso multiderivada de orden Q con tamaño de paso uniforme, con ciertas ventajas teóricas y prácticas para el tratamiento de sistemas de ecuaciones diferenciales ordinarias (ver [8] con continuidad en trabajos como [13]). Para este trabajo, se estudiaron los esquemas conocidos como “fórmulas de segundo orden” con el objetivo de validar una “familia” de métodos eficientes para el tratamiento de la rigidez. Específicamente se utilizó la siguiente fórmula, propuesta por Enright:

$$x_{i+1} = x_i + h \sum_{j=1}^Q \rho_j^{Q+2} f(x_{i-j+1}, u) + h \rho_0^{Q+2} f(x_{i+1}, u) + h^2 \rho_0^{Q+2} D_x f(x_{i+1}, u) \cdot f(x_{i+1}, u) \quad (13)$$

con el propósito de estructurar lo que hemos llamado método Predictor-Corrector-Recorrector (**PCR**), tratando de que fuera competitivo con fórmulas como la presentada anteriormente (13), tanto desde el punto de vista de los cálculos computacionales como al tener en cuenta las particularidades de la resolución de modelos rígidos.

La variante explícita, que se denotó por **EPCR** combina el esquema Predictor-Corrector con la fórmula explícita de Enright en un esquema corrector que evalúa, tanto el segundo término de f (o sea: $f(x_{i+1}, u)$) como el término de la derivada ($D_x f(x_{i+1}, u) \cdot f(x_{i+1}, u)$) de (13), en la aproximación que genera el Corrector. Esto es lo que lo convierte, desde el punto de vista del cálculo general, en un esquema explícito.

$$\begin{aligned} \min J(u) &= \sum_{i=0}^k \tilde{\varphi}_i [z_i, \tilde{z}_i], \\ y_{i+1} &= x_i + h \sum_{j=1}^{Q_i} \pi_j^{Q_i} f_{i-j+1}(x_{i-j+1}, u), \quad i = 0, 1, 2, \dots, k-1, \\ y_{i+1}^c &= x_i + h K_0^{Q_i+1} f_{i+1}(y_{i+1}, u) + h \sum_{j=1}^{Q_i} K_j^{Q_i+1} f_{i-j+1}(x_{i-j+1}, u), \\ x_{i+1} &= x_i + h \sum_{j=1}^{Q_i} \rho_j^{Q_i+2} f_{i-j+1}(x_{i-j+1}, u) + h \rho_0^{Q_i+2} f_{i+1}(y_{i+1}^c, u) + \\ &\quad + h^2 \rho_0^{Q_i+2} D_x f_{i+1}(y_{i+1}^c, u) \cdot f_{i+1}(y_{i+1}^c, u), \\ x_0 &= l(u), \\ z_i &= g_i(x_i, u), \quad i = 0, 1, 2, \dots, k. \end{aligned}$$

2.3 Estimación del error para la fórmula correctora

Siguiendo una estrategia similar a la que se propone en ([20]) para el esquema **PC**, estimaremos el error que se comete al combinar dicho esquema con la fórmula de Enright de segundo orden.

Denotemos por

$$\begin{aligned} x_{i+1} &= x_i + h_i \sum_{j=1}^{Q_i} \rho_j^{Q_i+2} f_{i-j+1}(x_{i-j+1}, t_{i+1-j}) + h_i \rho_0^{Q_i+2} f_{i+1}(y_{i+1}^c, t_{i+1}) + \\ &\quad + h_i^2 \gamma_0^{Q_i+2} D_x f_{i+1}(y_{i+1}^c, t_{i+1}) \cdot f_{i+1}(y_{i+1}^c, t_{i+1}) \end{aligned} \quad (14)$$

la expresión recorrectora de tipo semi-implícito, donde y_{i+1}^c define la evaluación obtenida en la iteración $i + 1$ utilizando la fórmula del corrector de Adams-Moulton, y por

$$\begin{aligned} x(t_{i+1}) &= x(t_i) + h_i \sum_{j=1}^{Q_i} \rho_j^{Q_i+2} f(x(t_{i-j+1}), t_{i+1-j}) + h_i \rho_0^{Q_i+2} f(x(t_{i+1}), t_{i+1}) + \\ &+ h_i^2 \gamma_0^{Q_i+2} D_x f(x(t_{i+1}), t_{i+1}) \cdot f(x(t_{i+1}), t_{i+1}) + \tau_{i+1} \end{aligned} \quad (15)$$

la expresión exacta para el cálculo de la evaluación de x en la iteración $i + 1$, según la expresión del recorrector y donde τ_{i+1} representa el error de aproximación que se comete.

Restando (14) de (15) tendríamos:

$$\begin{aligned} x(t_{i+1}) - x_{i+1} &= x(t_i) - x_i + h_i \sum_{j=1}^{Q_i} \rho_j^{Q_i+2} [f(x(t_{i-j+1}), t_{i+1-j}) - f_{i-j+1}(x_{i-j+1}, t_{i+1-j})] + \\ &+ h_i \rho_0^{Q_i+2} [f(x(t_{i+1}), t_{i+1}) - f_{i+1}(y_{i+1}^c, t_{i+1}, u)] + \\ &+ h_i^2 \gamma_0^{Q_i+2} [D_x f(x(t_{i+1}), t_{i+1}) \cdot f(x(t_{i+1}), t_{i+1}) - \\ &- D_x f(y_{i+1}^c, t_{i+1}) \cdot f(y_{i+1}^c, t_{i+1})] + \\ &+ \tau_{i+1} \end{aligned}$$

Según (11), definimos el error global, entonces la expresión anterior se transforma en:

$$\begin{aligned} e_{i+1} &= e_i + h_i \sum_{j=1}^{Q_i} \rho_j^{Q_i+2} g_{i+1-j} e_{i+1-j} + h_i \rho_0^{Q_i+2} g_{i+1} [x(t_{i+1}) - y_{i+1}^c] + \\ &+ h_i^2 \gamma_0^{Q_i+2} [D_x f(x(t_{i+1}), t_{i+1}) \cdot f(x(t_{i+1}), t_{i+1}) - D_x f(y_{i+1}^c, t_{i+1}) \cdot f(y_{i+1}^c, t_{i+1})] \\ &+ \tau_{i+1} \end{aligned} \quad (16)$$

donde:

$$g_{i+1-j} = f_x(t_{i+1-j}, \xi_{i+1-j})$$

para

$$\xi_{i+1-j} \text{ entre } \begin{cases} x_{i-j+1} \text{ y } x(t_{i-j+1}), & \text{para } j > 0 \\ y_{i+1} \text{ y } x(t_{i+1}), & \text{para } j = 0 \end{cases}$$

exigiendo que : $|f_x(x, t)| \leq L$ (constante de Lipchitz), $\forall a \leq t \leq b$ y $\forall x$

Sumando y restando $D_x f(y_{i+1}^c, t_{i+1}) \cdot f_{i+1}(x(t_{i+1}), t_{i+1})$ en el cuarto término de la expresión (16), tendremos:

$$\begin{aligned} e_{i+1} &= e_i + h_i \sum_{j=1}^{Q_i} \rho_j^{Q_i+2} g_{i+1-j} e_{i+1-j} + h_i \rho_0^{Q_i+2} g_{i+1} [x(t_{i+1}) - y_{i+1}^c] + \\ &+ h_i^2 \gamma_0^{Q_i+2} [D_x f(x(t_{i+1}), t_{i+1}) \cdot f(x(t_{i+1}), t_{i+1}) - D_x f(y_{i+1}^c, t_{i+1}) \cdot f(y_{i+1}^c, t_{i+1}) + \\ &\pm D_x f(y_{i+1}^c, t_{i+1}) \cdot f_{i+1}(x(t_{i+1}), t_{i+1})] + \tau_{i+1} \end{aligned}$$

Agrupando convenientemente, sacando factores comunes y denotando $e_{i+1}^c = x(t_{i+1}) - y_{i+1}^c$, obtendremos:

$$\begin{aligned} e_{i+1} &= e_i + h_i \sum_{j=1}^{Q_i} \rho_j^{Q_i+2} g_{i+1-j} e_{i+1-j} + h_i \rho_0^{Q_i+2} g_{i+1} e_{i+1}^c + \\ &+ h_i^2 \gamma_0^{Q_i+2} \{ [D_x f(x(t_{i+1}), t_{i+1}) - D_x f(y_{i+1}^c, t_{i+1})] \cdot f(x(t_{i+1}), t_{i+1}) + \\ &+ D_x f(y_{i+1}^c, t_{i+1}) g_{i+1} e_{i+1}^c \} + \tau_{i+1} \end{aligned}$$

Haciendo $\mathbf{v} = g_x$, con $|g_x(x, t)| \leq M$ (cierta constante lipchitziana), tendremos que

$$\begin{aligned} \mathbf{v}_{i+1} &= g_x(t_{i+1}, \xi_{i+1}), \\ \mathbf{v}_{i+1}(x(t_{i+1}) - y_{i+1}^c) &= g(x(t_{i+1}), t_{i+1}) - g(y_{i+1}^c, t_{i+1}) = \\ &= D_x f(x(t_{i+1}), t_{i+1}) - D_x f(y_{i+1}^c, t_{i+1}) \end{aligned}$$

De modo que:

$$\begin{aligned} e_{i+1} &= e_i + h_i \sum_{j=1}^{Q_i} \rho_j^{Q_i+2} g_{i+1-j} e_{i+1-j} + h_i \rho_0^{Q_i+2} g_{i+1} e_{i+1}^c + \\ &+ h_i^2 \gamma_0^{Q_i+2} [\mathbf{v}_{i+1} e_{i+1}^c f(x(t_{i+1}), t_{i+1}) + D_x f(y_{i+1}^c, t_{i+1}) g_{i+1} e_{i+1}^c] + \tau_{i+1} \\ &= e_i + h_i \sum_{j=1}^{Q_i} \rho_j^{Q_i+2} g_{i+1-j} e_{i+1-j} + h_i \left(\rho_0^{Q_i+2} + h_i \gamma_0^{Q_i+2} D_x f(y_{i+1}^c, t_{i+1}) \right) g_{i+1} e_{i+1}^c + \\ &+ h_i^2 \gamma_0^{Q_i+2} \mathbf{v}_{i+1} e_{i+1}^c f(x(t_{i+1}), t_{i+1}) + \tau_{i+1} \end{aligned}$$

Según la estrategia del ([20]), podemos denotar como:

$$e_{i+1}^p = e_i + h_i \sum_{j=0}^{Q_i-1} \pi_{j+1}^{Q_i} g_{i-j} e_{i-j} + \tau_{i+1}^p$$

el error del esquema predictor,

$$e_{i+1}^c = e_i + h_i \sum_{j=1}^{Q_i-1} K_j^{Q_i+1} g_{i+1-j} e_{i+1-j} + h_i K_0^{Q_i+1} g_{i+1} e_{i+1}^p + \tau_{i+1}^c$$

el del corrector.

Estas expresiones las sustituimos en la estimación del error del esquema combinado Predictor-Corrector-Recorrector y asociando convenientemente los términos de las suma-

torias, obtenemos:

$$\begin{aligned}
e_{i+1} = & e_i + h_i \sum_{j=0}^{Q_i} \rho_j^{Q_i+2} g_{i+1-j} e_{i+1-j} + h_i^2 \gamma_0^{Q_i+2} D_x f(y_{i+1}^c, t_{i+1}) g_{i+1} e_i + \\
& + h_i^3 \gamma_0^{Q_i+2} D_x f(y_{i+1}^c, t_{i+1}) g_{i+1} \sum_{j=1}^{Q_i} K_j^{Q_i+1} g_{i+1-j} e_{i+1-j} + \\
& + h_i^3 \gamma_0^{Q_i+2} D_x f(y_{i+1}^c, t_{i+1}) g_{i+1}^2 K_0^{Q_i+1} e_i + \\
& + h_i^4 \gamma_0^{Q_i+2} K_0^{Q_i+1} D_x f(y_{i+1}^c, t_{i+1}) g_{i+1}^2 \sum_{j=1}^{Q_i} \pi_j^{Q_i} g_{i+1-j} e_{i+1-j} + \\
& + h_i^3 \gamma_0^{Q_i+2} K_0^{Q_i+1} D_x f(y_{i+1}^c, t_{i+1}) g_{i+1}^2 \tau_{i+1}^p + h_i^2 \gamma_0^{Q_i+2} D_x f(y_{i+1}^c, t_{i+1}) g_{i+1} \tau_{i+1}^c + \\
& + h_i^2 \gamma_0^{Q_i+2} \mathbf{v}_{i+1} e_i f(x(t_{i+1}), t_{i+1}) + \\
& + h_i^3 \gamma_0^{Q_i+2} \mathbf{v}_{i+1} \sum_{j=1}^{Q_i} K_j^{Q_i+1} g_{i+1-j} e_{i+1-j} f(x(t_{i+1}), t_{i+1}) \\
& + h_i^3 \gamma_0^{Q_i+2} K_0^{Q_i+1} \mathbf{v}_{i+1} g_{i+1} e_i f(x(t_{i+1}), t_{i+1}) + h_i^4 \gamma_0^{Q_i+2} K_0^{Q_i+1} \mathbf{v}_{i+1} g_{i+1} \times \\
& \times \sum_{j=1}^{Q_i} \pi_j^{Q_i} g_{i+1-j} e_{i+1-j} f(x(t_{i+1}), t_{i+1}) + \\
& + h_i^3 \gamma_0^{Q_i+2} K_0^{Q_i+1} \mathbf{v}_{i+1} g_{i+1} \tau_{i+1}^p f(x(t_{i+1}), t_{i+1}) + \\
& + h_i^2 \gamma_0^{Q_i+2} \mathbf{v}_{i+1} \tau_{i+1}^c f(x(t_{i+1}), t_{i+1}) + \tau_{i+1}
\end{aligned}$$

donde

- τ_{i+1}^p es el error de truncamiento de la expresión predictora, de orden $k + 1$,
- τ_{i+1}^c es el error de truncamiento de la expresión correctora, de orden $k + 2$,
- τ_{i+1} es el error de truncamiento de la expresión recorrectora, de orden $k + 3$.

3 Conclusiones

Todo el análisis anterior, nos permite concluir que al combinar la fórmula de segundo orden, con las conocidas fórmulas de Adams-Bashfordt-Moulton, para proponer una estrategia combinada de tipo Predictor-Corrector-Recorrector, estamos tratando con un método numérico que está avalado por resultados teóricos que lo hacen eficientes para problemas rígidos (ver [8]) y con una estimación global de aproximación del orden de h^2 , dada por

la expresión:

$$\begin{aligned}
 e_{i+1} = & e_i + h_i \sum_{j=1}^{Q_i} \rho_j^{Q_i+2} g_{i+1-j} e_{i+1-j} + h_i \rho_0^{Q_i+2} g_{i+1} e_{i+1}^c + \\
 & + h_i \left[\left(\rho_0^{Q_i+2} + h_i \gamma_0^{Q_i+2} D_x f(y_{i+1}^c, t_{i+1}) \right) g_{i+1} e_{i+1}^c + \right. \\
 & \left. + h_i^2 \gamma_0^{Q_i+2} \mathbf{v}_{i+1} e_{i+1}^c f(x(t_{i+1}), t_{i+1}) \right] + \tau_{i+1}
 \end{aligned}$$

Referencias

- [1] Alexander, R. (1997) “Reliability of software for stiff initial-value problems”, *SIAM J.,Sci. Comput.*, **IV**(10).
- [2] Arévalo, C.; Fuhrer, C.; Soderlind, G. (1996) “Stabilized multistep methods for index 2 Euler-Lagrange DAEs”, *BIT* **36**: 1–13.
- [3] Bock, H.G. (1991) “Recent advances in parameter identification techniques for ODE in modern methods of optimization”, in: W. Krabs & J. Zowe (Eds.) *Proceedings of the Summer School*, Bayreuth, 1990, Springer Verlag, New York.
- [4] Broyden, C.G. (1967) “Quasi-Newton methods and their application to function minimization”, *Math.Comp.* **21**: 368–381.
- [5] Butcher, J.C. (1990) “Order, stepsize and stiffness switching”, *Computing* **44**: 209–220.
- [6] Dahlquist, G. (1974) “Problems related to the numerical treatment of stiff differential equations”, in: A. Gunther, B. Levrat, & H. Lipps (Eds.) *International Computing Symp.*, Amer. Elsevier, New York, p.307.
- [7] Dennis, J.E.; Schnabel, R.B. (1983) *Numerical Methods for Unconstrained Optimization and Non Linear Equations*. Prentice-Hall Series in Computational Mathematics, New Jersey.
- [8] Enright, W.H. (1976) “Second derivate multi-step methods for stiff ordinary differential equations”, *Numerical Analysis* **2**(2): 351–362.
- [9] Gear, W.C. (1971) *Numerical Initial Value Problems in O.D.E.* Prentice-Hall, New Jersey.
- [10] Gómez, J.A.; Marrero, A. (2000) “Computing gradients of inverse problems in ODE models”, *Revista de Investigación Operativa, ALIO* **9**(1,2,3): 179–206
- [11] Gómez, J.A.; Marrero, A. (2000) “Convergence of discrete aproximations of inverse problems in ODE models”, *Revista de Investigación Operativa, ALIO* **9**(1,2,3): 207–224.

- [12] Hairer, E.; Wanner, G.; Norsett, S.P. (1996) *Solving Ordinary Differential Equations I, II*. Springer Verlag, New York.
- [13] Higham, D.J. (1989) “Analysis of the Enright-Kamel partitioning methods for stiff ordinary differential equations”, *IMA J.Numerical Anal.* **9**: 1–14.
- [14] Lebedev, V.I. (1994) “How to solve stiff systems of differential equations by explicit methods”, in: G.I. Marchuk (Ed.) *Numerical methods and Applications*, CRC Press: 45-80.
- [15] Lebedev, V.I.; Medovikov, A.A. (1994) “Explicit methods of second order for the solution of stiff systems of ODEs”, *Russian Academy of Science* **IV**(2), Moscow.
- [16] Marrero, A.; Gómez, J.A. (1990) “Estimación de parámetros en modelos dinámicos no lineales con empleo de métodos numéricos”, *Investigación Operacional* **XI**(2): 101–118.
- [17] Marrero, A.; Gómez, J.A. (1994) “Problemas de estimación de parámetros en modelos dinámicos no lineales”, *Ciencias Matemáticas* **15**(1): 41–54.
- [18] Marrero, A.; Gómez, J.A. (1997) Experiencias numéricas con métodos multipaso-multiderivadas para el problema de estimación de parámetros en modelos dinámicos”, *Investigación Operacional* **18**(1-3): 3–10.
- [19] Marrero, A. (2000) *Un Enfoque para la Solución Numérica del Problema de Estimación de Parámetros Definido por Ecuaciones Diferenciales Ordinarias*. Tesis doctoral, Universidad de La Habana.
- [20] Shampine, L.F.; Gordon, M.K. (1991) *Computer Solution of Ordinary Differential Equations. The Initial Value Problem*. W.H. Freeman & Company, San Francisco.