Métodos variacionales en la solución de ecuaciones diferenciales: una mejora al método de colocación

Peñuñuri, F. R.[1] y Cab Cauich, C. A.[2]

Fecha de recepción: 8 de junio de 2007 *Fecha de aceptación: 2 de agosto de 2008*

RESUMEN
En este artículo se estudian algunos métodos variacionales para la solución de ecuaciones diferenciales. Aunque estos métodos se pueden utilizar por si solos en la solución de tales ecuaciones, la simbiosis con el Método del Elemento Finito produce muy buenos resultados. Con el fin de cuantificar el error se analizan dos tipos de normas, la norma euclidiana y la norma de suma. En el contexto de reducir la norma euclidiana del error, se presenta una alternativa para mejorar el método de colocación y se muestra que la mejor solución dada por este nuevo método coincide con la del método de mínimos cuadrados. Finalmente, con la ayuda de un ejemplo sencillo, se revisan las ideas introducidas y se hace una comparación numérica entre tales métodos.

Palabras claves: Métodos variacionales, ecuaciones diferenciales, método de colocación.

Variational Methods in the solution of Differential Equations: an improvement to the collocation Method

ABSTRACT
In this paper some variational methods for the solution of differential equations are studied. Although these methods can be used by itself in the solution of differential equations, the symbioses with the Finite Element Methods produce very good results. In order to quantify the error, two different kinds of norms, the Euclidian norm and the Sum norm, were analyzed. A method to reduce the error obtained by means of the Euclidian norm for the collocation method is proposed. Comparisons are made between the proposed method and the least square method and it is proved that the best solution given by the proposed method is the same as the one obtained through the least square method. Finally, a simple example with numerical comparisons is presented.

Keywords: Variational Methods, Differential Equations, Collocation Method.

[1] Profesor de carrera. Cuerpo Académico de Física. Facultad de Ingeniería-UADY. E-mail: francisco.pa@uady.mx
[2] Ingeniero Físico. Egresado de la Facultad de Ingeniería-UADY

INTRODUCCIÓN

En general la descripción matemática de los fenómenos que ocurren en la naturaleza se presenta en forma de ecuaciones diferenciales (ED). Por tal motivo es de suma importancia tener un fuerte conocimiento de los métodos que proporcionan soluciones a tales ecuaciones. Dado que la solución analítica rara vez es posible, es necesario implementar diversos métodos que permitan obtener, en la mayor de las veces numéricamente, sus soluciones. Entre los métodos de gran popularidad el método de elemento finito (MEF) es conocido por sus buenos resultados y aplicaciones prácticas. Este método consiste en hacer una partición en elementos finitos del dominio de definición de la función para luego, usando algún método aproximado, obtener la solución en cada uno de los elementos.

Por mencionar dos de tales métodos aproximados tenemos:
a) El método de Rayleigh-Ritz
b) El método de residuos ponderados el cual se subdivide a su vez en:

- El método de Galerkin
- El método de mínimos cuadrados
- El método de colocación

Todos estos métodos son bien conocidos en la literatura y se les llama métodos variacionales. Es fácil pensar en variaciones de ellos. Por ejemplo, en el caso del método de mínimos cuadrados en vez de minimizar la norma euclidiana del error (raíz cuadrada de la suma de los cuadrados de los errores), se puede minimizar la norma de suma (la suma de los valores absolutos de los errores).

Después de un breve resumen de los métodos mencionados, en la sección II se muestra una mejora del método de colocación y se estudia el límite de tal mejora de acuerdo a ciertos criterios que permitan decir si una solución es mejor que otra. En la sección III se revisa un ejemplo en el que se comparan los errores de las soluciones obtenidas con los distintos métodos y, finalmente se presentan las conclusiones.

Consideraciones generales

Desde el punto de vista de la mecánica newtoniana, si se conocen las fuerzas que actúan sobre un cuerpo es posible determinar su posición en cualquier instante de tiempo, ya sea pasado presente o futuro. En particular se sabe que la fuerza es la derivada temporal del momento lineal lo que a su vez resulta en la segunda derivada de la posición respecto del tiempo. La descripción de la naturaleza en mecánica clásica es por medio de fuerzas, esto lleva a ecuaciones diferenciales de segundo orden. De la discusión anterior, este trabajo se limita a ecuaciones diferenciales de segundo orden, no obstante, los resultados obtenidos serán aplicables a ecuaciones diferenciales de cualquier orden, incluso ecuaciones diferenciales parciales.

Sea pues,

$$F(Y(x), Y'(x), Y''(x), x) = 0 \qquad (1)$$

con condiciones iniciales y de frontera (CF) apropiadas, la ecuación a resolver. La solución es $Y(x)$ tal que si se sustituye en la Ec. (1) y se verifican las (CF), la Ec. (1) se satisface idénticamente. Aquí $x \in (a,b)$ con a y b números reales.

Se representará una solución aproximada por $y(x)$. Una sustitución de $y(x)$ en la Ec. (1) da un término $R(x) \neq 0$ denominado residuo ya que la Ec. (1) no se satisface en forma exacta. Entonces

$$R(x) = F(y(x), y'(x), y''(x), x)\,. \qquad (2)$$

La representación para la solución aproximada $y(x)$ se escribe como:

$$y(x) = \sum_{j=0}^{n} c_j \phi_j(x); c_0 = 1 \qquad (3)$$

con ϕ_j funciones linealmente independientes y c_j parámetros desconocidos por determinar. Comúnmente se escogen polinomios para las funciones ϕ_j.

Cuantificación del error

Si se conocen $Y(x_i)$ y $y(x_i)$ para algún x_i dentro del intervalo (a,b), se define e_i como:

$$e_i \equiv e(x_i) = Y(x_i) - y(x_i). \qquad (4)$$

Con esto se construye un vector error **E** cuya i-ésima componente es e_i.

La norma de suma del vector **E** se define como la suma de los valores absolutos de cada componente:

$$\| \mathbf{E} \|_s = \sum_i |e_i|\,. \qquad (5)$$

La definición anterior es una muy buena alternativa para cuantificar el error. No obstante, en el límite, cuando el índice *i* se vuelve continuo, es complicado manejar los valores absolutos en forma analítica. Una alternativa muy común para cuantificar el error es la norma euclidiana, su definición es:

$$\|\mathbf{E}\|_E = \sqrt{\sum_i e_i^2}. \qquad (6)$$

Cuando se minimiza el error dado por la norma euclidiana, se dice que el error se minimiza por mínimos cuadrados.

Cuando el índice *i* se vuelve continuo,

$$\|\mathbf{E}\|_s = \int_a^b dx\,|e(x)| \qquad (7)$$

y

$$\|\mathbf{E}\|_E = \sqrt{\int_a^b dx\; e^2(x)}. \qquad (8)$$

Ya que la solución exacta $Y(x)$ es desconocida se usara el residuo, $R(x)$ en el caso continuo y $R_i \equiv R(x_i)$ en el caso discreto, para cuantificar el error. Esto es:

$$\|\mathbf{E}\|_s = \sum_i |R_i| \qquad (9)$$

$$\|\mathbf{E}\|_E = \sqrt{\sum_i R_i^2} \qquad (10)$$

$$\|\mathbf{E}\|_s = \int_a^b dx\,|R(x)| \qquad (11)$$

$$\|\mathbf{E}\|_E = \sqrt{\int_a^b dx R^2(x)}. \qquad (12)$$

De acuerdo a la discusión previa, se dice que una solución es mejor que otra dependiendo del criterio utilizado (Ecs. 9, 10, 11, 12).

Métodos variacionales

La esencia fundamental de estos métodos es optimizar cierta cantidad. En esta sección se hará una breve descripción de ellos. Para una descripción más elaborada se recomienda la siguiente bibliografía (Redy, 1993; Hutton, 2004)

El método de Rayleigh-Ritz

Este método se aplica a ecuaciones diferenciales originadas de las ecuaciones de Euler-Lagrange asociadas a cierta función lagrangiana $L(Y, Y', x)$. La solución se determina minimizando la acción:

$$S = \int_a^b dx L(Y, Y', x) \qquad (13)$$

El método de Rayleigh-Ritz consiste en sustituir la solución exacta $Y(x)$ por la aproximada $y(x)$ la que depende de parámetros $c_1, c_2, \cdots$.

El hecho de que S sea un extremo implica que

$$dS[y(x; c_1, c_2, \cdots)] = 0 \qquad (14)$$

de donde

$$\frac{\partial S}{\partial c_1} = \frac{\partial S}{\partial c_2} = \cdots = 0 \qquad (15)$$

De la Eq. (15) se plantea un sistema de *n* ecuaciones con *n* incógnitas. Resolviendo este sistema se obtiene los parámetros c_j.

El método de residuos ponderados (Petrov-Galerkin)

Cuando se sustituye la solución (3) en la ecuación (1), resulta un residuo el cual es función de *x*, no obstante, debido a que los parámetros no son conocidos, a priori, se escribirá

$$R(x; c_j). \qquad (16)$$

En el método de residuos ponderados, los parámetros se determinan integrando el residuo multiplicado por ciertas funciones conocidas como funciones de peso.

$$\int_a^b w_i(x) R(x; c_j) = 0, \qquad (17)$$

realizando las integrales anteriores, se obtiene un sistema de *n* ecuaciones con *n* incógnitas. Resolviendo tal sistema, se obtienen los parámetros c_j que se sustituirán en la solución aproximada. Ec. (3).

El método de Galerkin

Para el caso en que las funciones de peso $w_i(x)$ de la Ec. (17) son las mismas que las de la solución aproximada (las ϕ_j), el enfoque se conoce como el método de Galerkin. Bajo ciertas circunstancias los métodos de Galerkin y Rayleigh-Ritz son equivalentes (Redy, 1993).

El método de mínimos cuadrados (MMC)
El método de mínimos cuadrados consiste en minimizar la norma euclidiana del vector error. Ec. (12):

$$I_{RE} = \int_a^b dx\ R^2(x; c_1, c_2, \cdots). \qquad (18)$$

Luego, de la condición de mínimo,

$$\frac{\partial I_{RE}}{\partial c_j} = 0 \qquad (19)$$

se obtiene de nuevo un sistema de n ecuaciones con n incógnitas que se debe resolver para obtener los parámetros c_j. Alternativamente, se puede minimizar la norma de suma del vector error:

$$I_{Rs} = \int_a^b dx \,|\, R(x; c_1, c_2, \cdots) \,|\, . \qquad (20)$$

Desafortunadamente un tratamiento analítico de tal ecuación es complicado pero se puede tratar un enfoque numérico para obtener los parámetros c_j.

El método de colocación (MC)
Este método es quizá el más simple de los mencionados hasta ahora pero es también el menos confiable.
El MC consiste en evaluar el residuo en ciertos puntos arbitrarios $x_i \in (a,b)$ e igualarlo a cero. El número de puntos a evaluar será tantos como parámetros haya en la Ec. (3) para formar un sistema de n ecuaciones con n incógnitas. Se procede como antes para obtener los c_j.
Nótese que este método es un caso especial de (17) cuando se escogen las funciones de peso como la delta de Dirac,

$$\int_a^b dx\ \delta(x - x_i) R(x; c_j) = R(x_i; c_j). \qquad (21)$$

De lo expuesto, es claro que si se desea obtener el menor valor de la norma euclidiana del error se debe usar el MMC, si lo que importa es minimizar la acción, se deberá escoger el método de Rayleigh-Ritz. La elección de un método en particular dependerá de lo que se quiera minimizar y/o de lo que sea factible de llevar a cabo. Conviene mencionar que los métodos más utilizados en el MEF son el método de Rayleigh-Ritz y el de Galerkin, por citar algunos trabajos (Jia-Yi Yeh, 2007; Pontaza, 2005; Hossne, 2007)

METODOLOGÍA

Método de colocación mejorado (MCM)
El MC es muy sencillo pero tiene serias deficiencias. Supóngase que en la Ec. (3) se tienen únicamente dos parámetros y se escogen dos puntos entre *a* y *b*, de acuerdo a la sección anterior, se tiene:

$$\begin{aligned} R(x_1; c_1, c_2) &= 0 \\ R(x_2; c_1, c_2) &= 0. \end{aligned} \qquad (22)$$

Resolviendo el sistema anterior se hace que el residuo de la solución aproximada sea cero en los puntos x_1 y x_2, pero quedan infinitos puntos donde el residuo puede ser grande.

De la Ec. (22), se puede escribir el conjunto de ecuaciones lineales en forma matricial como $M_{2\times2}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$. Los puntos x_1 y x_2 son arbitrarios pero teniendo en cuenta que la matriz M sea invertible.

Para mejorar el método se propone agregar más puntos en la expresión (22). Por ejemplo si se agrega x_3:

$$\begin{aligned} R(x_1; c_1, c_2) &= 0 \\ R(x_2; c_1, c_2) &= 0 \\ R(x_3; c_1, c_2) &= 0 \end{aligned} \qquad (23)$$

el conjunto de ecuaciones lineales es ahora:

$$M_{3\times2}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}. \qquad (24)$$

Es bien sabido que tal sistema no tiene solución en el sentido usual, de hecho el sistema es claramente inconsistente, a menos que alguna de las ecuaciones sea linealmente dependiente de las otras dos lo que en última instancia se reduciría a un sistema de 2×2.

Se propone ahora, cambiar el criterio usado para encontrar los parámetros c_1 y c_2.

En lugar de tratar resolver el sistema de ecuaciones lineales (24) se pretende encontrar c_1 y c_2 de tal forma que

$$\left\| M_{3\times 2}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} - \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} \right\|_E ,$$

tenga el mínimo valor posible. Se expresa lo anterior formalmente a través del siguiente teorema:

Teorema

Sea A una matriz real $m\times n$ y $\mathbf{b} \in \mathbb{R}^m$. Entonces $A\mathbf{x}=\mathbf{b}$ siempre tiene al menos una solución $\overline{\mathbf{x}}$ en el sentido que:

$$\| A\overline{\mathbf{x}} - \mathbf{b} \|_E \leqslant \| A\mathbf{x} - \mathbf{b}\|_E,\ \forall\, \mathbf{x} \in \mathbb{R}^n$$

si las columnas de A son linealmente independientes, la solución es única y está dada por:

$$\overline{\mathbf{x}} = (A^T A)^{-1} A^T\mathbf{b}. \tag{25}$$

Para ver una demostración se recomienda la referencia (Poole, 2003).

La generalización de la ecuación (23) a una introducción de l puntos queda:

$$\begin{aligned} R(x_1;c_j) &= 0 \\ R(x_2;c_j) &= 0 \\ &\vdots \\ R(x_l;c_j) &= 0. \end{aligned} \tag{26}$$

o en forma matricial:

$$A_{l\times 2}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_l \end{pmatrix} \tag{27}$$

El método de colocación mejorado hará uso de la Ec. (25) con el fin de obtener el conjunto de parámetros que describan la mejor solución.

Nótese que un procedimiento similar se podría usar para mejorar el método de Petrov-Galerkin (desde luego que también el método de Galerkin que como ya se mencionó es un caso particular de éste), se trataría con p funciones de peso extra, tantas como se quiera, para tener un sistema de $n+p$ ecuaciones con n incógnitas, n de estas ecuaciones saldrían de (17) las otras de las funciones extra que se haya propuesto. Se procedería de la misma forma que en el MCM para obtener los c_j.

DISCUSIÓN Y RESULTADOS

Comparación con el método de mínimos cuadrados

En ocasiones no es posible realizar analíticamente el método de mínimos cuadrados y se usa un tratamiento numérico de la Ec. (18). Es claro que el resultado de la integral involucrada dependerá de la partición hecha para realizar la integral en forma numérica, pero si la partición es de dos rectángulos de igual base, el resultado será el mismo que el dado por el método de colocación si se usan los mismos puntos x_k en ambos métodos. De igual forma, existe una partición que dará los mismos resultados que el método de colocación mejorado, ésta es:

$$\int_a^b dx R^2(x;c_j) \approx \sum_k^l R^2(x_k;c_j)\Delta x. \tag{28}$$

Nótese que en la expresión anterior se ha tomado Δx y no Δx_k, la partición se toma a intervalos constantes, $\Delta x = cte$.
Una minimización de

$$\sum_k^l R^2(x_k;c_j)\Delta x \tag{29}$$

dará los mismos c_j que resolver el sistema de ecuaciones (26) o la Ec. (27).

El mejor resultado del método de colocación mejorado será tal que el número de puntos en los que el residuo se hace cero, sea infinito. Matemáticamente,

$$\begin{aligned} R(x_1;c_j) &= 0 \\ R(x_2;c_j) &= 0 \\ &\vdots \\ R(x_l;c_j) &= 0 \\ &\vdots \end{aligned} \tag{30}$$

y es bien sabido que cuando la partición se hace infinita, la integral ya no depende de ésta, luego entonces, por el teorema previamente discutido la solución del sistema Ec. (30) coincidirá con los resultados del método de mínimos cuadrados.

Ciertamente habrá situaciones en las que una minimización de la Ec. (29) sea conveniente en vez de resolver el sistema Ec. (27), pero es claro que también habrá situaciones donde una realización del método de colocación mejorado sea más conveniente.

Aplicación de los métodos a la solución de una ecuación diferencial.

Con el propósito de hacer algunas comparaciones entre los métodos ya mencionados se realiza el siguiente ejemplo tomado de Ref. (Hoffman, 2001).

Se propone resolver la siguiente ecuación diferencial usando 4 funciones prueba en el intervalo [0,1]:

$$Y''(x) - 16Y = 0;\ Y(0) = 0;\ Y(1) = 100 \quad (31)$$

de (3)

$$y(x) = \phi_0(x) + c_1\phi_1(x) + c_2\phi_2(x) + c_3\phi_3(x), \quad (32)$$

escogiendo polinomios para las funciones $\phi_j(x)$, se tiene que $\phi_0(x) = 0$ para que $Y(x)$ y $y(x)$ satisfagan las mismas condiciones de frontera. Para las otras funciones se escogen los polinomios:

$$\phi_1(x) = x; \phi_2(x) = x(x-1); \phi_3(x) = x^2(x-1) \quad (33)$$

Entonces

$$y(x) = c_1 x + c_2 x(x-1) + c_3 x^2(x-1) \quad (34)$$

y de la segunda condición de frontera c_1=100 de donde:

$$y(x) = 100x + c_2 x(x-1) + c_3 x^2(x-1) \quad (35)$$

Aplicando el método de (R-R) con la lagrangiana:

$$L = 8Y^2 + \frac{Y'^2}{2} \quad (36)$$

se tiene

$$c_2 = \frac{21600}{377}; c_3 = \frac{5600}{29} \quad (37)$$

así que

$$y_{RR} = 100x + \frac{21600}{377}(x-1)x + \frac{5600}{29}(x-1)x^2 \quad (38)$$

Para este ejemplo, el método de Galerkin coincide con el método de (R-R). Usando las mismas funciones $\phi_j(x)$ el MMC da:

$$y_{MMC} = 100x + \frac{51630}{8961}(x-1)x + \frac{21280}{103}(x-1)x^2 \quad (39)$$

El MC, evaluando en los puntos $x_1 = \frac{1}{3}$ y $x_2 = \frac{2}{3}$ da:

$$y_{MC} = 100x + \frac{2592}{43}(x-1)x + \frac{7200}{43}(x-1)x^2 \quad (40)$$

El MCM con siete ecuaciones cuyo residuo se evaluó en el intervalo [1/8,7/8] con incrementos de 1/8 da:

$$y_{MCM} = 100x + \frac{15564800}{263577}(x-1)x + \frac{492800}{2559}(x-1)x^2 \quad (41)$$

Una minimización numérica de la norma de suma del error nos da:

$$y_{ns} = 100x + 63.9476(x-1)x + 171.09(x-1)x^2 \quad (42)$$

Los errores correspondientes se muestran en la tabla 1, éstos son respecto a la ecuación (2), no con respecto a la Ec.(4).

Tabla 1: Errores (Norma euclidiana y norma de suma)

Método	$\| E \|_E$	$\| E \|_s$
R-R	212.755	158.25
MMC	207.871	167.183
MC	239.527	156.444
MCM	211.891	158.678
ns	228.966	154.422

La solución exacta es:

$$Y(x)=100\ Csch(4)\ Senh(4x) \qquad (43)$$

y en la figura (1) se muestran las gráficas para algunas de las soluciones en el rango $0 \le x \le 0.5$.

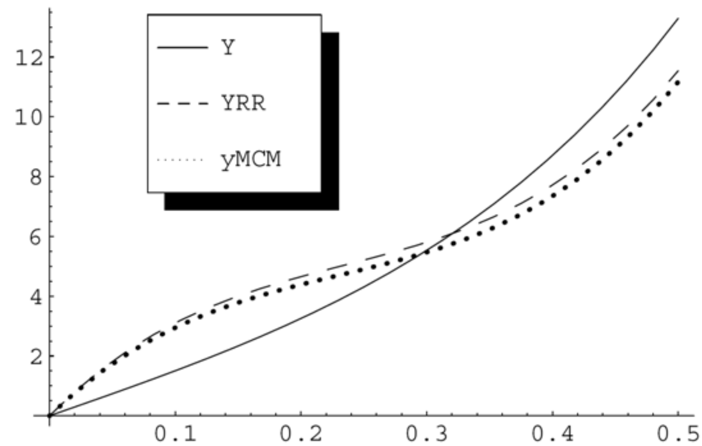

Figura 1: Gráficas; solución exacta, R-R y MCM

En la tabla 2, se compara la norma euclidiana del error para los métodos MMC y MCM. Se observa que si el número de ecuaciones se incrementa, las normas euclidianas de los errores para ambos métodos coinciden.

Tabla 2: Comparación entre las normas euclidianas de los errores para (MCM) y (MMC)

No. de ecuaciones	$\|\mathbf{E}_{MCM}\|_E - \|\mathbf{E}_{MMC}\|_E$	$\|\mathbf{E}_{MMC}\|_E$
7	4.01966	207.87115
99	0.02395	207.87115
999	0.00023	207.87115
9999	$2´10^{-6}$	207.87115

CONCLUSIONES

Después de un breve resumen de algunos métodos variacionales, se escogió el método de colocación para realizar mejoras sobre éste con lo que se obtuvo un nuevo enfoque al cual se le ha llamado método de colocación mejorado (MCM). Con el propósito de ilustrar las ideas expuestas se analizó un ejemplo sencillo y se vio que con una introducción de siete ecuaciones en el MCM se tiene un error por mínimos cuadrados menor que en el método de Rayleigh-Ritz. Se hicieron comparaciones entre ellos y se mostró que la mejor solución del MCM (en el sentido de minimizar la norma euclidiana del error) es la misma que en el método de mínimos cuadrados. Numéricamente, para una introducción de 9999 ecuaciones en el MCM se obtuvo una diferencia entre $\| \mathbf{E}_{MCM} \|_E$ y $\| \mathbf{E}_{MMC} \|_E$ del orden de 2×10^{-6}.

REFERENCIAS BIBLIOGRÁFICAS

Reddy, J.N (1993). An Introduction to the Finite Element Method. McGraw-Hill.

Hutton, D.V (2004). Fundamentals of Finite Elements Analysis. McGraw-Hill.

Jia-Yi Yeh (2007). Vibration analyses of the annular plate with electrorheological fluid damping treatment. Finite Elements in Analysis and Design, Volume 43, Issues 11-12, 965-974.

Pontaza, J.P (2005). Least-squares variational principles and the finite element method: theory, formulations, and models for solid and fluid mechanics. Finite Elements in Analysis and Design, Volume 41, Issues 7-8, 703-728.

Hossne, A.G (2007). Lagrangiano de un elemento finito plano elástico con ocho grados de libertad. Ingeniería, Revista Académica de la FI-UADY,11-1,pp.25-36,ISNN:1665-529X.

Poole, D (2003). Linear Algebra. A modern Introduction. Thomson Brooks/Cole.

Hoffman, J.D (2001). Numerical Methods for Engineers and Scientists. McGraw-Hill.

Este documento se debe citar como:
Peñuñuri, F. R. y Cab Cauich, C. A. (2008**). Métodos variacionales en la solución de ecuaciones diferenciales: una mejora al método de colocación.** Ingeniería, Revista Académica de la FI-UADY, 12-2, pp. 37-44, ISSN: 1665-529X