

# Introducción a las Ecuaciones Integrales

Benjamín Toledo C.

27 de Noviembre de 2000

## Introducción

Existen muchos tipos diferentes de ecuaciones integrales, cada una de las cuales puede presentar problemas al resolverlas; a menudo muchos algoritmos diferentes han sido propuestos para tratar con un solo caso.

Hay una estrecha correspondencia entre las ecuaciones integrales lineales, que especifican relaciones lineales entre funciones de un espacio de funciones de dimensión infinita, y la conocidas ecuaciones lineales, que especifican relaciones análogas entre vectores de un espacio vectorial de dimensión finita. Debido a que esta correspondencia yace en el corazón de la mayoría de los algoritmos computacionales, será valioso hacerlo explícito al ver como se clasifican las ecuaciones integrales.

Las *ecuaciones de Fredholm* involucran integrales definidas con límites fijos. Una *ecuación inhomogenea de Fredholm de primera especie* tiene la forma

$$g(t) = \int_a^b K(t, s)f(s)ds \quad (1)$$

Aquí  $f(t)$  es una función desconocida a resolver, mientras que  $g(t)$  es conocida. La función de dos variables  $K(t, s)$  es llamada el *kernel*. La ecuación (1) es análoga a la ecuación matricial

$$\mathbf{K} \cdot \mathbf{f} = \mathbf{g} \quad (2)$$

cuya solución es  $\mathbf{f} = \mathbf{K}^{-1} \cdot \mathbf{g}$ , donde  $\mathbf{K}^{-1}$  es la matriz inversa. Así como la ecuación (2), la ecuación (1) tiene solución única donde quiera que  $g$  es no nula y  $K$  es invertible. Sin embargo, como veremos, esta última condición es a menudo la excepción a la regla.

El análogo del problema de autovalores en dimensión finita

$$(\mathbf{K} - \sigma \mathbf{1}) \cdot \mathbf{f} = \mathbf{g} \quad (3)$$

es llamada *ecuación de Fredholm de segunda especie*, usualmente escrita como

$$f(t) = \lambda \int_a^b K(t, s)f(s)ds + g(t) \quad (4)$$

De nuevo, las convenciones de notación no corresponden exactamente:  $\lambda$  en la ecuación (4) es  $1/\sigma$  en la ecuación (3), mientras que  $\mathbf{g}$  es  $-g/\lambda$ . Si  $g$  (o  $\mathbf{g}$ ) es nula, entonces la ecuación es *homogénea*. Si el kernel  $K(t, s)$  es acotado, entonces, como la ecuación (3), la ecuación (4) tiene la propiedad de que su forma homogénea tiene soluciones para al menos un conjunto infinito numerable  $\lambda = \lambda_n, n = 1, 2, \dots$ , los *autovalores*. Los autovalores son reales si el es kernel simétrico.

En el caso inhomogeneo, en que  $g$  (o  $\mathbf{g}$ ) es no nula, las ecuaciones (3) y (4) son solubles excepto cuando  $\lambda$  (o  $\sigma$ ) es un autovalor— porque entonces el operador integral (o matriz) es singular. En las ecuaciones integrales esta dicotomía es llamada *la alternativa de Fredholm*.

Las ecuaciones de Fredholm de primera especie presentan frecuentemente dificultades analíticas difíciles de abordar. Aplicar el kernel a una función es generalmente una operación que suaviza, así la solución, la cual requiere invertir el operador, será extremadamente sensible a pequeños cambios o errores en la entrada. Suavizar repetidamente realmente pierde información, y no hay manera de devolverla en una operación inversa. Métodos especializados han sido desarrollados para tales ecuaciones, a menudo son llamados *problemas inversos*. En general, un método debe aumentar la información dada con algún conocimiento previo de la naturaleza de la solución. Este conocimiento previo es entonces usado, de una forma u otra, para restaurar la información perdida.

Las ecuaciones inhomogeneas de Fredholm de segunda especie usualmente no presentan dificultades analíticas. La ecuación (4) puede ser escrita como

$$\int_a^b [K(t, s) - \sigma\delta(t - s)]f(s)ds = -\sigma g(t) \quad (5)$$

donde  $\delta(t - s)$  es la delta de Dirac (se ha cambiado  $\lambda$  por su reciproco  $\sigma$  por claridad). Si  $\sigma$  es suficientemente grande en magnitud entonces la ecuación (5) es, en efecto, diagonalmente dominante y por tanto bien condicionada. Solamente para  $\sigma$  pequeño volveremos al caso que se comporta mal.

Las ecuaciones de Fredholm homogéneas de segunda especie no son así mismo, particularmente mal condicionadas. Si  $K$  es un operador de suavizado, entonces mapeará muchas  $f$  en cero, o cerca de cero; de este modo habrá un gran número de autovalores degenerados o casi degenerados cerca de  $\sigma = 0$  ( $\lambda \rightarrow \infty$ ), pero esto no producirá dificultades computacionales. De hecho,

podemos ver ahora que la magnitud necesaria para rescatar la ecuación inhomogénea (5) de un destino mal condicionado es generalmente mucho menor que la requerida para dominancia diagonal. Ya que el término  $\sigma$  desplaza todos los autovalores, basta que sea suficientemente grande como para desplazar a la agrupación de autovalores casi nulos del operador de suavizado fuera de cero, de modo tal que el operador resultante sea invertible (excepto, por supuesto, en los autovalores).

Las *ecuaciones de Volterra* son un caso especial de las ecuaciones de Fredholm con  $K(t, s) = 0$  para  $s > t$ . Las ecuaciones de Volterra se escriben en una forma en que el límite superior de integración es la variable independiente  $t$ .

La *ecuación de Volterra de primera especie*

$$g(t) = \int_a^t K(t, s)f(s)ds \quad (6)$$

tiene su ecuación matricial análoga (ahora escrita en componentes)

$$\sum_{j=1}^k K_{kj}f_j = g_k \quad (7)$$

Comparando con la ecuación (2), vemos que la ecuación de Volterra corresponde a una matriz  $\mathbf{K}$  que es triangular inferior, con entradas nulas sobre la diagonal. Cuando las mediciones experimentales de ruido no dominan, las ecuaciones de Volterra de primera especie tienden a comportarse bien; el límite superior de la integral introduce un paso agudo que convenientemente retira cualquier propiedad suavizante del kernel.

La ecuación de Volterra de segunda especie se escribe

$$f(t) = \int_a^t K(t, s)f(s)ds + g(t) \quad (8)$$

cuyo análogo matricial es la ecuación

$$(\mathbf{K} - \mathbf{I}) \cdot \mathbf{f} = \mathbf{g} \quad (9)$$

donde  $\mathbf{K}$  es triangular inferior. Las razones por las que no hay  $\lambda$  en estas ecuaciones son (i) en el caso inhomogéneo ( $g$  no nula) puede ser absorbida en  $\mathbf{K}$ , mientras que (ii) en el caso homogéneo ( $g = 0$ ), un teorema dice que las ecuaciones de Volterra de segunda especie con kernel acotado no tienen autovalores con autofunciones de cuadrado integrable.

De este modo hemos especializado nuestras definiciones al caso de ecuaciones integrales lineales.

Los integrandos en las versiones no lineales de las ecuaciones (1) o (6) serían  $K(t, s, f(s))$  en lugar de  $K(t, s)f(s)$ ; las versiones no lineales de las ecuaciones (4) o (8) tendrían un integrando  $K(t, s, f(t), f(s))$ . Las ecuaciones no lineales de Fredholm son considerablemente mas complicadas que sus contrapartes lineales. Afortunadamente, no se dan frecuentemente en la práctica. En contraste, resolver una ecuación no lineal de Volterra involucra usualmente solo una paqueña modificación al algoritmo para ecuaciones lineales.

Casi todos los métodos para resolver numéricamente ecuaciones integrales hacen uso de *reglas de integración*, frecuentemente reglas de integración Gaussianas.

### Reglas de integración Gaussianas

La idea de las reglas de integración Gaussianas es darnos la libertad de escoger no solo los  $W_k$  como en la integrales de Montecarlo, sino también las abscisas en que la función será evaluada: ya no tendremos una partición regular ni impuesta por  $W$ . En general, con buenas elecciones, esto permite una mejor aproximación de las integrales.

Hay, sin embargo, un aspecto adicional de las fórmulas de integración Gaussianas que amplía su utilidad: podemos arreglar la elección de los  $W_k$  y las abscisas para hacer integrales exactas de la clase:  $p(x)W(x)$ , donde  $p(x)$  es un polinómio. Además, la función  $W(x)$  puede ser escogida de modo tal que elimine alguna singularidad de la integral en cuestión. Para resumir, dada  $W(x)$  y un entero  $N$ , podemos encontrar un conjunto de  $w_k$  y abscisas  $x_k$  tal que la aproximación

$$\int_a^b W(x)f(x)dx \approx \sum_{j=1}^N w_j f(x_j) \quad (10)$$

sea exacta si  $f(x)$  es un polinómio.

Notemos que la fórmula de integración (10) puede ser escrita sin hacer claramente visible la función  $W$ : definamos  $g(x) \equiv W(x)f(x)$  y  $v_j \equiv w_j/W(x_j)$ , entonces (10) será

$$\int_a^b g(x)dx \approx \sum_{j=1}^N v_j g(x_j) \quad (11)$$

**Caso general:** las abscisas que se usarán corresponden a los ceros de un conjunto de polinómios ortogonales; para introducir polinómios ortogonales, fijemos el intervalo de interés en  $(a, b)$ . Podemos definir el producto escalar

de las funciones  $f$  y  $g$  sobre la función de ponderación  $W$  como

$$\langle f|g \rangle \equiv \int_a^b W(x)f(x)g(x)dx \quad (12)$$

El producto escalar es un número. El conjunto de funciones mutuamente ortogonales e individualmente normalizadas se llama conjunto *ortonormal*.

Se puede encontrar un conjunto de polinómios que (i) incluya exactamente un polinomio de orden  $j$ , llamado  $p_j(x)$ , para  $j = 1, 2, \dots$ , y (ii) todos los cuales sean ortogonales sobre la función de ponderación  $W(x)$ . Un procedimiento constructivo para encontrar tal conjunto se tiene en la siguiente recurrencia

$$\begin{aligned} p_{-1}(x) &\equiv 0 \\ p_0(x) &\equiv 1 \\ p_{j+1}(x) &= (x - a_j)p_j(x) - b_j p_{j-1}(x) \quad j = 1, 2, \dots \end{aligned} \quad (13)$$

donde

$$\begin{aligned} a_j &= \frac{\langle xp_j|p_j \rangle}{\langle p_j|p_j \rangle} \quad j = 0, 1, \dots \\ b_j &= \frac{\langle p_j|p_j \rangle}{\langle p_{j-1}|p_{j-1} \rangle} \quad j = 1, 2, \dots \end{aligned} \quad (14)$$

El coeficiente  $b_0$  es arbitrario, podemos hacerlo cero.

Los polinómios definidos por (13) son mónicos, es decir, el coeficiente de la mayor potencia es la unidad. Si dividimos cada  $p_j(x)$  por la constante  $[\langle p_j|p_j \rangle]^{\frac{1}{2}}$  podemos obtener el conjunto de polinómios ortonormales. Se puede hallar polinómios ortonormales con otras normalizaciones. Siempre se pueden hacer mónicos dividiendo el polinomio  $j$  por el coeficiente  $\lambda_j$  de la mayor potencia.

Si el cuerpo sobre el que estamos trabajando es  $\mathbb{C}$ , entonces por el *teorema fundamental del álgebra*, el polinomio de orden  $j$  tiene exactamente  $j$  raíces en su dominio. Además, se puede demostrar que las raíces de  $p_j(x)$  entrelazan las  $j - 1$  raíces de  $p_{j-1}(x)$ , es decir, hay exactamente una raíz de la primera entre cada dos raíces adyacentes de la segunda. Este hecho puede resultar útil si necesitamos encontrar todas las raíces.

Recordemos que queremos encontrar estas raíces porque es precisamente en estas abscisas donde evaluaremos no sólo la función a integrar sino también la función de ponderación  $W(x)$  escogida. Una vez conocidas las abscisas, necesitamos encontrar los  $w_j$  una manera práctica de hacerlo es utilizando la

siguiente fórmula (que se deja sin demostrar)

$$w_j = \frac{\langle p_{N-1} | p_{N-1} \rangle}{p_{N-1}(x_j) p'_N(x_j)} \quad (15)$$

donde  $p'_N(x_j)$  es la derivada del polinomio ortonormal en su cero  $x_j$ .

De modo que el cálculo de una regla de integración Gaussiana tiene dos aspectos distintos: (i) la generación de polinómios ortonormales  $p_0, \dots, p_N$ , es decir, el cálculo de los coeficientes  $a_j$  y  $b_j$  en (13); (ii) la determinación de los ceros de  $p_j(x)$  y el cálculo de las ponderaciones  $w_j$  asociadas. En el caso de polinómios ortogonales “clásicos”, los coeficientes  $a_j$  y  $b_j$  son explícitamente conocidos, y se podría omitir la parte (i). Sin embargo, si nos enfrentamos a una función  $W(x)$  “no-clásica”, y no conocemos los coeficientes  $a_j$  y  $b_j$ , la construcción del conjunto de polinómios ortogonales no es trivial.

A continuación se presenta una lista con las funciones  $W(x)$ , intervalos, y relaciones de recurrencia que generan los polinómios ortogonales que se usan con mayor frecuencia y sus correspondientes fórmulas de integración Gaussiana:

*Gauss-Legendre:*

$$\begin{aligned} W(x) &= 1 & -1 < x < 1 \\ (j+1)P_{j+1} &= (2j+1)xP_j - jP_{j-1} \end{aligned} \quad (16)$$

*Gauss-Chevyshev:*

$$\begin{aligned} W(x) &= (1-x^2)^{-1/2} & -1 < x < 1 \\ T_{j+1} &= 2xT_j - T_{j-1} \end{aligned} \quad (17)$$

*Gauss-Laguerre:*

$$\begin{aligned} W(x) &= x^\alpha e^{-x} & 0 < x < \infty \\ (j+1)L_{j+1}^\alpha &= (-x + 2j + \alpha + 1)L_j^\alpha - (j + \alpha)L_{j-1}^\alpha \end{aligned} \quad (18)$$

*Gauss-Hermite:*

$$\begin{aligned} W(x) &= e^{-x^2} & -\infty < x < \infty \\ H_{j+1} &= 2xH_j - 2jH_{j-1} \end{aligned} \quad (19)$$

*Gauss-Jacobi:*

$$\begin{aligned} W(x) &= (1-x)^\alpha(1+x)^\beta & -1 < x < 1 \\ c_j P_{j+1}^{(\alpha,\beta)} &= (d_j + e_j x) P_j^{(\alpha,\beta)} - f_j P_{j-1}^{(\alpha,\beta)} \end{aligned} \quad (20)$$

donde los coeficientes  $c_j$ ,  $d_j$ ,  $e_j$ , y  $f_j$  son dados por

$$\begin{aligned} c_j &= 2(j+1)(j+\alpha+\beta+1)(2j+\alpha+\beta) \\ d_j &= (2j+\alpha+\beta+1)(\alpha^2-\beta^2) \\ e_j &= (2j+\alpha+\beta)(2j+\alpha+\beta+1)(2j+\alpha+\beta+2) \\ f_j &= 2(j+\alpha)(j+\beta)(2j+\alpha+\beta+2) \end{aligned} \quad (21)$$

## Ecuación de Fredholm de segunda especie

Nos interesa encontrar una solución numérica para  $f(t)$  en la ecuación

$$f(t) = \lambda \int_a^b K(t, s) f(s) ds + g(t) \quad (22)$$

el método que se describirá, uno muy básico, se conoce como *método de Nystrom*. Se requiere una regla de integración

$$\int_a^b y(s) ds = \sum_{j=1}^N w_j y(s_j) \quad (23)$$

Si aplicamos (23) a (22) resulta

$$f(t) = \lambda \sum_{j=1}^N w_j K(t, s_j) f(s_j) + g(t) \quad (24)$$

evaluando (24) en los puntos de integración

$$f(t_i) = \lambda \sum_{j=1}^N w_j K(t_i, s_j) f(s_j) + g(t_i) \quad (25)$$

Sea  $f_i$  el vector  $f(t_i)$ ,  $g_i$  el vector  $g(t_i)$ ,  $K_{ij}$  la matriz  $K(t_i, s_j)$ , y definamos

$$\tilde{K}_j = K_{ij} w_j \quad (26)$$

Ahora en notación matricial la ecuación (25) es

$$(\mathbf{1} - \lambda \tilde{\mathbf{K}}) \cdot \mathbf{f} = \mathbf{g} \quad (27)$$

Este es un conjunto de  $N$  ecuaciones lineales con  $N$  incógnitas que puede ser resuelto con técnicas estándar de descomposición triangular. La solución se comporta bien a menos que  $\lambda$  esté muy cerca de un autovalor (porque en tal caso el kernel tiende a diverger).

### Descomposición Triangular

Supongamos que somos capaces de escribir la matriz  $\mathbf{A}$  como el producto de dos matrices

$$\mathbf{L} \cdot \mathbf{U} = \mathbf{A} \quad (28)$$

donde  $\mathbf{L}$  es triangular inferior y  $\mathbf{U}$  es triangular superior (en ambos casos los elementos de la diagonal son no nulos)

Podemos usar este tipo de descomposiciones para resolver el conjunto lineal

$$\mathbf{A} \cdot \mathbf{x} = (\mathbf{L} \cdot \mathbf{U}) \cdot \mathbf{x} = \mathbf{L} \cdot (\mathbf{U} \cdot \mathbf{x}) = \mathbf{b} \quad (29)$$

mediante primero resolver para el vector  $\mathbf{y}$  tal que

$$\mathbf{L} \cdot \mathbf{y} = \mathbf{b} \quad (30)$$

y entonces resolver

$$\mathbf{U} \cdot \mathbf{x} = \mathbf{y} \quad (31)$$

La ventaja de proceder así es que la solución de un conjunto triangular de ecuaciones es muy trivial; la ecuación (30) puede ser resuelta por sustitución hacia adelante, como sigue

$$\begin{aligned} y_1 &= \frac{b_1}{\alpha_{11}} \\ y_i &= \frac{1}{\alpha_{ii}} \left[ b_i - \sum_{j=1}^{i-1} \alpha_{ij} y_j \right] \quad i = 2, 3, \dots, N \end{aligned} \quad (32)$$



donde  $\alpha_{rs}$  son los elementos de  $\mathbf{L}$ . La ecuación (31) se resuelve por sustitución hacia atrás

$$\begin{aligned} x_N &= \frac{y_N}{\beta_{NN}} \\ x_i &= \frac{1}{\beta_{ii}} \left[ y_i - \sum_{j=i+1}^N \beta_{ij} x_j \right] \quad i = N-1, N-2, \dots, 1 \end{aligned} \quad (33)$$

donde los  $\beta_{rs}$  son elementos de  $\mathbf{U}$ .

### Realizando la descomposición Triangular

Primero, escribimos la  $(i, j)$ -ésima componente de la ecuación (28). Esa componente es siempre una suma que empieza por

$$\alpha_{i1}\beta_{1j} + \dots = a_{ij} \quad (34)$$

El número de términos en la suma depende, sin embargo, del número menor entre  $i$  y  $j$ . Tenemos, de hecho, tres casos

$$i < j : \quad \alpha_{i1}\beta_{1j} + \alpha_{ii}\beta_{ij} = a_{ij} \quad (35)$$

$$i = j : \quad \alpha_{i1}\beta_{1j} + \alpha_{ii}\beta_{jj} = a_{ij} \quad (36)$$

$$i > j : \quad \alpha_{i1}\beta_{1j} + \alpha_{ij}\beta_{jj} = a_{ij} \quad (37)$$

Las ecuaciones (35)-(37) totalizan  $N^2$  ecuaciones para  $N^2 + N$  incógnitas  $\alpha$ 's y  $\beta$ 's (la diagonal aparece dos veces). Ya que el número de incógnitas es mayor que el número de ecuaciones, debemos especificar  $N$  de las incógnitas arbitrariamente y entonces tratar de resolver las otras. De hecho, es siempre posible tomar

$$\alpha_{ii} \equiv 1 \quad i = 1, \dots, N \quad (38)$$

Lo que sigue es usar el *algoritmo de Crout*, el que de manera muy trivial resuelve el sistema de  $N^2 + N$  ecuaciones (35)-(38) para todos los  $\alpha$ 's y  $\beta$ 's [mediante solo reordenar las ecuaciones de una cierta manera! Es como sigue:

- Sean  $\alpha_{ii} = 1$ ,  $i = 1, \dots, N$  (ecuación (38))
- Para cada  $j = 1, 2, 3, \dots, N$  realice estos dos procedimientos: Primero, para  $i = 1, 2, 3, \dots, j$ , use (35), (36) y (38) para resolver los  $\beta$ 's, esto es

$$\beta_{ij} = a_{ij} - \sum_{k=1}^{i-1} \alpha_{ik}\beta_{kj} \quad (39)$$

(cuando  $i = 1$  en (39) la sumatoria se toma igual a cero). Segundo, para  $i = j + 1, j + 2, \dots, N$  use (37) para resolver los  $\alpha_{ij}$ , esto es

$$\alpha_{ij} = \frac{1}{\beta_{jj}} \left( a_{ij} - \sum_{k=1}^{j-1} \alpha_{ik} \beta_{kj} \right) \quad (40)$$

Debemos asegurarnos de realizar ambos procedimientos antes de proseguir con el siguiente  $j$ .

Además, es necesario realizar la operación llamada *partial pivoting* que a grandes razgos consiste en hallar el mayor  $\beta$  en una fila, dividir todo por éste y mover esta fila para poner el 1 producido en la diagonal. Solo de esta manera se puede asegurar que el algoritmo sea estable.

Para resumir, este procedimiento resuelve la ecuación (27) mediante encontrar las matrices  $\mathbf{L}$  y  $\mathbf{U}$  tales que  $\mathbf{L} \cdot \mathbf{U} = \mathbf{1} - \lambda \tilde{\mathbf{K}}$ , y luego aplicar (32) y (33).

Ahora que tenemos la solución en los puntos de integración  $\{t_i\}$ , para obtener un valor aproximado de la solución en otros puntos, Nystrom observó que debemos usar la ecuación (24) como fórmula de interpolación, y así mantener la precisión ya obtenida.

Una desventaja de los métodos basados en reglas de integración Gaussianas es que no hay una manera simple de estimar el error en el resultado. El mejor método práctico consiste en incrementar el valor de  $N$ , en digamos, un 50% y tratar la diferencia entre las dos estimaciones como una aproximación del error obtenido en el resultado con  $N$  más grande.

Veamos ahora las soluciones de la ecuación homogénea. Si ponemos  $\lambda = 1/\sigma$  y  $\mathbf{g} = 0$ , entonces la ecuación (27) queda como una ecuación de autovalores estándar

$$\tilde{\mathbf{K}} \cdot \mathbf{f} = \sigma \mathbf{f} \quad (41)$$

Notemos que si el problema original tiene kernel simétrico, entonces la matriz  $\mathbf{K}$  es simétrica. Sin embargo, ya que las ponderaciones  $w_j$  no son iguales para la mayoría de las reglas de integración, la matriz  $\tilde{\mathbf{K}}$  no es simétrica. El problema de autovalores es más sencillo para matrices simétricas, y por lo tanto debemos restaurar la simetría si es posible. Dado que los  $w_j$  son positivos, podemos definir la matriz diagonal  $\mathbf{D} = \text{diag}(w_j)$  y su raíz cuadrada,  $\mathbf{D}^{1/2} = \text{diag}(\sqrt{w_j})$ . La ecuación (41) se convierte en

$$\mathbf{K} \cdot \mathbf{D} \cdot \mathbf{f} = \sigma \mathbf{f} \quad (42)$$

Multiplicando por  $\mathbf{D}^{1/2}$ , obtenemos

$$\left( \mathbf{D}^{1/2} \cdot \mathbf{K} \cdot \mathbf{D}^{1/2} \right) \cdot \mathbf{h} = \sigma \mathbf{h} \quad (43)$$

donde  $\mathbf{h} = \mathbf{D}^{1/2} \cdot \mathbf{f}$ . La ecuación (43) esta ahora en forma simétrica.

La solución a las ecuaciones (42) y (43) dará en general  $N$  autovalores, donde  $N$  es el número de puntos de integración usados. Para kernels de cuadrado integrable, darán una buena aproximación a los  $N$  autovalores más bajos de la ecuación integral. Kernels de *rango finito* (también llamados *degenerados* o kernels *separables*) tienen solo un número finito de autovalores no nulos (posiblemente ninguno). Uno puede diagnosticar esta situación al notar una acumulación de autovalores  $\sigma$  nulos en la presición de la máquina. El número de autovalores no nulos permanecerá constante al incrementar  $N$  para mejorar la presición. Aquí hay que tener algún cuidado: un kernel no degenerado puede tener un número infinito de autovalores con un punto de acumulación en  $\sigma = 0$ . Se puede distinguir los dos casos por el comportamiento de la solución al incrementar  $N$ . Si sospechamos que el kernel es degenerado, usualmente seremos capaces de resolver el problema por técnicas analíticas descritas en todos los libros de texto relacinados con el tema.

## Ecuaciones de Volterra

Una ecuación típica de Volterra de segunda especie es

$$f(t) = \int_a^t K(t, s)f(s)ds + g(t) \quad (44)$$

La mayoría de los algoritmos para ecuaciones de Volterra comienzan de  $t = a$ , construyendo la solución a medida que avanza. En este sentido, éste no solo nos recuerda la sustitución hacia adelante, sino también los problemas de condiciones iniciales para ecuaciones diferenciales. De hecho, muchos algoritmos para EDO tienen equivalentes para ecuaciones de Volterra.

La manera más simple de proceder es resolver la ecuación en una partición uniforme:

$$t_i = a + ih, \quad i = 0, 1, \dots, N, \quad h \equiv \frac{b-a}{N} \quad (45)$$

Para hacer esto, debemos escoger una regla de integración. Para una partición uniforme, la más simple es la regla trapezoidal:

$$\begin{aligned} f_0 &= g_0 \\ (1 - \frac{1}{2}hK_{ii})f_i &= h\left(\frac{1}{2}K_{i0}f_0 + \sum_{j=1}^{i-1} K_{ij}f_j\right) + g_i, \quad i = 1, \dots, N \end{aligned} \quad (46)$$

(Para una ecuación de Volterra de primera especie, el 1 a la izquierda estaría ausente, y  $g$  tendría el signo opuesto, para el resto de la ecuación los cambios son directos)

La ecuación (46) es un método explícito que da la solución en  $O(N^2)$  operaciones. A diferencia de las ecuaciones de Fredholm, no es necesario resolver un sistema lineal de ecuaciones. De modo que las ecuaciones de Volterra usualmente envuelven menos trabajo que la correspondiente de Fredholm.

Si interpretamos la ecuación (44) como una ecuación vectorial de  $m$  funciones  $f(t)$ , entonces el kernel  $K(t, s)$  es una matriz  $m \times m$ . La ecuación (46) ahora también debe ser entendida como una ecuación vectorial. Para cada  $i$ , tenemos que resolver  $m \times m$  ecuaciones lineales algebraicas por eliminación Gaussiana.

Para ecuaciones de Volterra no lineales, hay que transformar la ecuación (46) cambiando el producto  $K_{ii}f_i$  por  $K_{ii}(f_i)$ , y similarmente para los otros productos de  $K$  y  $f$ . Así para cada  $i$  resolvemos una ecuación no lineal para  $f_i$  con un lado derecho conocido. El método de Newton con una elección inicial para  $f_{i-1}$  usualmente funciona muy bien siempre y cuando el tamaño de paso no sea muy grande.

## Ecuaciones integrales con kernels singulares

Muchas ecuaciones integrales tienen singularidades en el kernel, en la solución, o en ambos. Un método simple de integración mostrará una pobre convergencia con  $N$  si ignoramos tales singularidades. No hay un procedimiento que funcione en todos los casos.

Algunas sugerencias útiles son:

- Las singularidades integrables pueden ser removidas a veces por un cambio de variable. Por ejemplo, el comportamiento singular de  $K(t, s) \sim s^{1/2}$  o  $s^{1/2}$  puede ser removido por la transformación  $z = s^{1/2}$ . Notemos que hemos asumido que el comportamiento singular es sólo de  $K$ , mientras que la regla de integración toma el producto  $K(t, s)f(s)$ , y es este producto el que debe ser considerado. Uno siempre debería deducir la naturaleza singular de este producto antes de probar una solución numérica, y tomar las medidas que sean necesarias. Comunmente, sin embargo, un kernel singular no produce una solución singular para  $f(t)$ . Por ejemplo, el kernel  $K(t, s) = \delta(t - s)$  resulta ser el operador identidad.
- Si  $K(t, s)$  puede ser factorizado como  $w(s)\bar{K}(t, s)$ , donde  $w(s)$  es singular y  $\bar{K}(t, s)$  es suave, entonces una integración Gaussiana basada en  $w(s)$  como función de ponderación, funcionará bien. Aun cuando la factorización sea solo aproximada, a menudo la convergencia mejora dramáticamente.

Este método es un caso especial del *product Nystrom method*, donde uno factoriza un término singular  $p(t, s)$  y construye una función de ponderación conveniente para su integración Gaussiana. Los cálculos en el caso general son muy complicados porque dependen de la elección de  $\{t_i\}$  y además de la forma de  $p(t, s)$ .

- Cuando el kernel no es estrictamente singular, pero es “casi” singular, pueden ser útiles algunas fórmulas especiales de integración. Un ejemplo es cuando el kernel está concentrado en torno a  $t = s$  en una escala mucho más pequeña que la escala en la cual la solución  $f(t)$  varía. En tal caso, una fórmula de integración puede estar basada en aproximaciones locales de  $f(t)$  por polinómios, mientras calculamos los primeros *momentos* del kernel en los puntos de tabulación  $t_i$ . En tales casos el estrecho ancho del kernel es mas una ventaja que un problema: la regla de integración llega a ser exacta cuando el ancho del kernel tiende a cero.
- Un rango de integración infinita es también una forma de singularidad. Truncar el dominio en algún valor finito es solo un último recurso. Si el kernel va rápidamente a cero, las reglas de integración de Gauss-Laguerre o Gauss-Hermite funcionarán bien. A las funciones con una “cola larga” podemos aplicar la transformación

$$s = \frac{2\alpha}{z+1} - \alpha \quad (47)$$

la que mapea  $0 < s < \infty$  en  $-1 < z < 1$  así que se puede usar la integración de Gauss-Legendre. Aquí  $\alpha > 0$  es una constante que debemos ajustar para mejorar la convergencia.

- Una situación común en la práctica es que  $K(t, s)$  sea singular en la diagonal  $t = s$ . Aquí el método de Nystrom falla completamente porque el kernel queda evaluado en  $(t_i, s_i)$ . Una posible cura es *sustraer la singularidad*:

$$\begin{aligned} \int_a^b K(t, s)f(s)ds &= \int_a^b K(t, s)[f(s) - f(t)]ds + \int_a^b K(t, s)f(t)ds \\ &= \int_a^b K(t, s)[f(s) - f(t)]ds + r(t)f(t) \end{aligned} \quad (48)$$

donde  $r(t) = \int_a^b K(t, s)ds$ , que se calcula analítica o numéricamente. Si el primer término del lado derecho es ahora regular podemos usar el

método de Nystrom. En lugar de la ecuación (25), tenemos

$$f_i = \lambda \sum_{\substack{j=1 \\ j \neq i}}^N w_j K_{ij} [f_j - f_i] + \lambda r_i f_i + g_i \quad (49)$$

A veces, el proceso de sustracción de ser repetido hasta que el kernel quede completamente regularizado.

### Integración en una partición uniforme con ponderaciones arbitrarias

En general es posible encontrar reglas lineales de integración de  $n$  puntos que aproximen la integral de la función  $f(x)$ , veces una función de ponderación arbitraria  $w(x)$ , sobre un rango arbitrario de integración  $(a, b)$ , como la suma de los productos entre las ponderaciones y  $n$  valores uniformemente espaciados de la función  $f(x)$ , digamos  $x = kh, (k+1)h, \dots, (k+n-1)h$ . El esquema general para derivar tales reglas de integración consiste en escribir las  $n$  ecuaciones lineales que se deben satisfacer si la regla de integración es exacta para las  $n$  funciones  $f(x) = \text{const.}, x, x^2, \dots, x^{n-1}$ , y entonces resolver estas para los coeficientes. Esto puede ser hecho analíticamente, una vez para siempre, si los momentos de la función de ponderación

$$W_n \equiv \frac{1}{h^n} \int_a^b x^n w(x) dx \quad (50)$$

sobre el mismo rango de integración, son conocidos. Aquí el factor  $h^{-n}$  se escoge para hacer que  $W_n$  escale como  $h$  si  $b - a$  es proporcional a  $h$ .

Llevando a cabo este proceso para el caso de cuatro puntos obtenemos

$$\begin{aligned} \int_a^b w(x) f(x) dx = & \\ & \frac{1}{6} f(kh) \left[ (k+1)(k+2)(k+3)W_0 - (3k^2 + 12k + 11)W_1 + 3(k+2)W_2 - W_3 \right] \\ & + \frac{1}{2} f([k+1]h) \left[ -k(k+2)(k+3)W_0 + (3k^2 + 10k + 6)W_1 - (3k+5)W_2 + W_3 \right] \\ & + \frac{1}{2} f([k+2]h) \left[ k(k+1)(k+3)W_0 + (3k^2 + 8k + 3)W_1 + (3k+4)W_2 - W_3 \right] \\ & + \frac{1}{2} f([k+3]h) \left[ -k(k+1)(k+2)W_0 + (3k^2 + 6k + 2)W_1 - 3(k+1)W_2 + W_3 \right] \end{aligned} \quad (51)$$

Mientras que los términos en paréntesis cuadrado parecen escalar como  $k^2$ , hay típicamente una cancelación en  $O(k^2)$  y  $O(k)$ . La ecuación (51) se puede utilizar para varias elecciones de  $(a, b)$ . La elección obvia es  $a = kh, b = (k+3)h$ , en cuyo caso tenemos una regla de integración de cuatro puntos que generaliza la regla de Simpson 3/8. De hecho, podemos recobrar este caso especial poniendo  $w(x) = 1$ , ahora (50) es

$$W_n = \frac{h}{n+1}[(k+3)^{n+1} - k^{n+1}] \quad (52)$$

Cada uno de los cuatro términos en paréntesis cuadrado de (51) llega a ser independiente de  $k$ , y (51) se reduce a

$$\int_{kh}^{(k+3)h} f(x)dx = \frac{3h}{8}f(kh) + \frac{9h}{8}f([k+1]h) + \frac{9h}{8}f([k+2]h) + \frac{3h}{8}f([k+3]h) \quad (53)$$

Volviendo al caso general  $w(x)$ , son útiles algunas otras elecciones para  $a$  y  $b$ . Por ejemplo, podríamos querer tomar  $(a, b)$  como  $([k+1]h, [k+3]h)$  o  $([k+2]h, [k+3]h)$ , permitiendonos completar una regla cuyo número de intervalos no es un múltiplo de tres, sin pérdida de precisión: La integral será estimada usando los cuatro valores  $f(kh), \dots, f([k+3]h)$ . Aún más útil es escoger  $(a, b)$  como  $([k+1]h, [k+2]h)$ , usando así cuatro puntos para integrar el único intervalo central. Estas ponderaciones, al ser puestas juntas en una fórmula extendida, dan esquemas de integración que tienen coeficientes suaves, i.e., sin las alternaciones del tipo Simpson 2, 4, 2, 4. Todas estas reglas tienen el mismo orden que la regla extendida de Simpson, que es, exacta cuando  $f(x)$  es un polinómio cúbico. Si se desea obtener reglas de orden más bajo, se derivan de manera similar. Las fórmula de tres puntos es

$$\begin{aligned} \int_a^b w(x)f(x)dx = & \frac{1}{2}f(kh) \left[ (k+1)(k+2)W_0 - (2k+3)W_1 + W_2 \right] \\ & + f([k+1]h) \left[ -k(k+2)W_0 + 2(k+1)W_1 - W_2 \right] \\ & + \frac{1}{2}f([k+2]h) \left[ k(k+1)W_0 - (2k+1)W_1 + W_2 \right] \end{aligned} \quad (54)$$

Aquí el caso especial simple es tomar,  $w(x) = 1$ , tal que

$$W_n = \frac{h}{n+1}[(k+2)^{n+1} - k^{n+1}] \quad (55)$$

Entonces la ecuación (54) es simplemente la regla de Simpson

$$\int_{kh}^{(k+2)h} f(x)dx = \frac{h}{3}f(kh) + \frac{4h}{3}f([k+1]h) + \frac{h}{3}f([k+2]h) \quad (56)$$

Para funciones de ponderación  $w(x)$  no constantes, sin embargo, la ecuación (54) da reglas de un orden menor que Simpson, ya que no tienen la simetría extra del caso constante.

La fórmula de dos puntos es simplemente

$$\int_{kh}^{(k+1)h} w(x)f(x)dx = f(kh)[(k+1)W_0 - W_1] + f([k+1]h)[-kW_0 + W_1] \quad (57)$$

## References

- [1] Numerical Recipes in Fortran, Second Edition  
Cap. 2, 4, 18