LOS METODOS ITERATIVOS, ¿ MATEMATICAS PURAS O MATEMATICAS APLICADAS ?

LUIS VALDIVIA REYES

1. INTRODUCCION

Parecen constituir conceptos antinómicos, ganancia de uno sería pérdida para el otro.

Un antecedente histórico puede dar más luces para entender esta aparente disyuntiva: antes de la segunda guerra mundial los matemáticos en los Estados Unidos sólo servían las necesidades de otras disciplinas y los cursos que se impartían eran esencialmente de servicio. Según narra L. Auslander en el "Bulletin of the American Mathematical Society" (de Noviembre-1979), mientras A. Weil enseñaba en una Universidad del Este de aquella nación, fue llamado al orden por un decano y acusado de romper la estructura que no contemplaba la presentación de demostraciones en clases. Después de la guerra, las matemáticas alcanzaron personalidad propia. Las demostraciones alcanzaron aceptabilidad y los matemáticos tuvieron demanda por propio derecho, y ya no como personal de servicio. Sin embargo, este período de crecimiento trajo aparejado un efecto lateral muy desafortunado. Mientras las matemáticas se convertían en una disciplina propia, muchos de sus practicantes deseaban liberarse de su primitiva imagen de segundones. Se sentían denigrados en el papel de realizar servicio; así ellos denigraban las matemáticas de servicio. Desafortunadamente, ellos apilaron en un

mismo montón las matemáticas de servicio y las matemáticas aplicadas. Así durante este período de crecimiento de las matemáticas, nació una distinción entre matemáticas pura y aplicada. Durante estos años, los matemáticos aplicados sin tuvieron el desprecio de los matemáticos puros sobre ellos, y así la comunicación entre la matemática pura y la aplicada virtualmente se quebró.

La diferencia entre matemáticas pura y aplicada no es tan grande. En verdad, abundan ejemplos de matemáticos puros y aplicados ocupados en investigaciones similares pero por falta de comunicación ninguno sabía del trabajo de los otros. También aquí cabe una explicación del origen de estas líneas, como un ejemplo de como la brecha entre matemática pura y aplicada puede ser salvada.

Me interesé en los problemas de punto fijo, porque deseaba estructurar en un lenguaje lo más estricto y a la vez más general posible el concepto de convergencia de los métodos iterativos para la solución de sistemas de ecuaciones lineales.

Idealmente, aquellos que deseen realizar trabajos de nivel avanzado en la materia de Análisis Numérico debe - rían tener un sólido "respaldo" de la disciplina matemática conocida como Análisis Funcional, ya que es el lenguaje y me dio adecuado para muchas de las materias de nivel sobre métodos de iteración en Análisis Numérico.

2. MÉTODOS ITERATIVOS

En los sistemas de ecuaciones lineales, se distingue entre métodos directos y métodos iterativos. Aparte de errores de redondeo, los métodos directos conducen a una solución exacta luego de un número finito de pasos. Los métodos iterativos construyen una sucesión de aproximaciones, que converge a la solución exacta. Estos son generalmente mucho más simples de programar que los métodos directos, además no juegan un papel importante los errores de redondeo. Sin embargo, contrariamente a los métodos directos, requie - ren mucho tiempo de computación. Cuando se utilizan métodos directos, hay que tener cuidado del hecho que variaciones mí nimas de un método pueden presentar susceptibilidades comple tamente distintas a los errores de redondeo.

Sea $F: G \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$, F función continua. Desea mos encontrar un cero $x^* \in G$ de F , o sea una solución de la ecuación

$$(1) \quad F(x) = 0 \quad \text{yacente en } G.$$

En Análisis Funcional, se obtiene un número de condiciones suficientes para la existencia de tal cero. Luego supondremos que un cero x^* de G existe, y que existe una vecindad de x^* en la cual F no tiene otros ceros. Además supondremos que G es un conjunto abierto.

Los métodos iterativos para determinar un cero de F se basan en una reformulación de (1) como un problema equivalente de punto fijo

$$(2) \quad x = T(x)$$

de modo que x^* es un cero de F exactamente cuando x^* es un punto fijo de T . Luego establecemos la siguiente iteración :

$$(3) \quad x^{(k)} = T[x^{(k-1)}], \quad k = 1, 2, 3, \dots$$

Se espera que la sucesión $\{x^{(k)}\}$ $k \in \mathbb{N}$ converga a x^* si el punto inicial $x^{(0)}$ está lo bastante cerca de x^* , pero esto no es suficiente. Además de la convergencia de la sucesión, debemos preocuparnos de la velocidad de esta convergencia, y de la simplicidad o complejidad de computar T .

Transformemos la ecuación (1) en el problema de punto fijo equivalente para un caso especial que aparece frecuentemente. Supongamos que la aplicación F puede ser desglorada en una suma,

$$F(x) = R(x) + S(x),$$

en la cual S dependerá "debilmente" de x y R es constructivamente invertible. Esto significa que existe un algoritmo realizable respecto al tiempo de computación, demanda de almacenamiento de memoria y sensibilidad al error de redondeo, y para el cual la ecuación $R(y) = b$ puede resolverse para todo b en una cierta vecindad de $-S(x^*)$.

Este es el caso, por ejemplo, cuando R es una transformación lineal dada mediante una matriz diagonal no-singular o mediante una matriz simétrica tridiagonal y definida positiva. Si ponemos :

$$T = R^{-1} \circ (-S)$$

entonces la ecuación $F(x) = 0$ es equivalente al problema de punto fijo $x = T(x)$. Cuando S , y por lo tanto T también, depende sólo debilmente del punto, podemos esperar que el método iterativo (3) converga a x^* para aproximaciones $x^{(0)}$ suficientemente cercas de x^* .

3. CONTRACCIONES.

Sea $T: G \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ una función y $x^* \in G$ un punto fijo de T , osea $T(x^*) = x^*$. $Y \in G$ pertenece a la "región atractiva" $I(T, x^*)$ de x^* , si la sucesión (3) para $x^{(0)} = Y$ permanece en G y converge a x^* . Al punto fijo x^* lo llamamos "atractivo" si es un punto interior de $I(T, x^*)$. La iteración (3) se llama "localmente convergente" si x^* es atractivo. La función T es una "contracción" si existe un $\alpha \in [0, 1)$ y una norma $\| \cdot \|_T$ en $\mathbb{R}^n$ tal que

$$\|T(x) - T(y)\|_T \leq \alpha \|x - y\|_T \quad x, y \in G$$

obviamente toda función contracción es continua.

Probaremos ahora el primer resultado importante para las contracciones.

Teorema : Sea $T: G \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ una función contracción, entonces :

(a) T tiene a lo mas un punto fijo $x^* \in G$. x^* es atractivo.

(b) Si $G = \mathbb{R}^n$, hay exactamente un punto fijo x^* . su región atractiva es

$$\mathbb{R}^n$$

Demostración de (a): Sean x^*, y^* dos puntos fijos de T . Ya que T es contracción, existe un $\alpha \in [0, 1)$ tal que

$$\|x^* - y^*\|_T = \|T(x^*) - T(y^*)\|_T \leq \alpha \|x^* - y^*\|_T$$

luego $x^* = y^*$.

Ahora escogemos $r > 0$ lo suficiente pequeño para que la esfera cerrada

$$K_{T,r} = \{ y \in \mathbb{R}^n \mid \|x^* - y\|_T \leq r \}$$

se encuentre incluida en G .

Luego para todo $z \in K_{T,r}$

$$\|T(z) - T(x^*)\|_T \leq \alpha \|z - x^*\|_T \leq r.$$

Por lo tanto T transforma la esfera $K_{T,r}$ en sí misma. La sucesión.

$$x^{(k)} = T(x^{(k-1)}), \quad k = 1, 2, 3, \dots$$

está definida para $x^{(0)} \in K_{T,r}$ y satisface la siguiente desigualdad

$$\|x^{(k)} - x^*\|_T \leq \alpha^k \|x^{(0)} - x^*\|_T \leq \alpha^k r,$$

$$k = 1, 2, 3, \dots$$

Luego la sucesión $\{x^{(k)}\}_{k=0}^{\infty}$ converge a x^* .

Demostración de (b) : Sea $x^{(0)} \in \mathbb{R}^n$ arbitrario.

Ya que T es contracción existe un $\alpha \in [0, 1)$ tal que para $k = 0, 1, 2, \dots$ se cumple que

$$\|x^{(k+1)} - x^{(k)}\|_T = \|T(x^{(k)}) - T(x^{(k-1)})\|_T$$

$$\leq \alpha \left\| x^{(k)} - x^{(k-1)} \right\|_T$$

$$\leq \alpha^k \left\| x^{(1)} - x^{(0)} \right\|_T$$

Luego, con ayuda de la desigualdad triangular, para todo $i, j \in \mathbb{N}$,

$$\left\| x^{(i+j)} - x^{(i)} \right\|_T \leq \sum_{k=i}^{i+j-1} \left\| x^{(k+1)} - x^{(k)} \right\|_T \leq \left(\sum_{k=i}^{i+j-1} \alpha^k \right) \left\| x^{(1)} - x^{(0)} \right\|_T$$

$$\leq \frac{\alpha^i}{1-\alpha} \left\| x^{(1)} - x^{(0)} \right\|_T$$

de donde $\{x^{(k)}\}_{k=0}^{\infty}$ es una sucesión de Cauchy. A su valor límite lo denotamos x^* . Ya que toda contracción es continua, se sigue que

$$x^* = \lim_{k \rightarrow \infty} x^{(k)} = \lim_{k \rightarrow \infty} T(x^{(k)}) = T(x^*)$$

Por lo tanto x^* es un punto fijo de T . La región atractiva de x^* es todo $\mathbb{R}^n$.

4. APLICACIONES.

Teorema :

Sea A una matriz real $n \times n$, $b \in \mathbb{R}^n$ y sea $T(x) = A(x) + b$ una transformación afín.

entonces T es contracción si y sólo si el radio espectral $\rho(A)$ es menor que 1. En tal caso, T tiene exactamente un punto fijo. Su región atractiva $I(T, x^*)$ es todo $\mathbb{R}^n$.

La demostración de esta aserción se sigue del Teorema en 3.- y de un Lema de Teoría de Matrices ("The Theory of Matrices in Numerical Analysis", A. Householder. Blaisdel Publishing Company, 1964): "Si A es una matriz $n \times n$ con $g(A) < 1$ entonces existe una norma de matrices $\| \cdot \|_S$ tal que $\| A \|_S < 1$ ".

Demostración del Teorema :

a) Si $g(A) < 1$ entonces por el Lema, existe $\| \cdot \|_S$ tal $\| A \|_S < 1$. Sea $| \cdot |_S$ la norma vectorial consistente, correspondiente a la norma de matrices, entonces :

$$\begin{aligned} | T(x) - T(y) |_S &= | A(x) - A(y) |_S = | A(x-y) |_S \\ &\leq \| A \|_S | x-y |_S \\ &< | x-y |_S \\ &\leq \alpha | x-y |_S \\ &\alpha \in [0, 1) \end{aligned}$$

b) Si T es contracción, entonces existe una norma $| \cdot |_S$ en $\mathbb{R}^n$ tal que

$$\begin{aligned} | T(x) - T(y) |_S &\leq \alpha | x-y |_S \quad \alpha \in [0, 1) \\ | A(x) - A(y) |_S &\leq \alpha | x-y |_S \\ \| A \|_S | x-y |_S &\leq \alpha | x-y |_S \end{aligned}$$

Luego $\| A \|_S < 1$, de donde $g(A) < 1$

Ejemplos : Métodos de Jacobi y Gauss - Seidel

$$\text{Sea } T(x) = (L+U) x + b$$

donde $b \in \mathbb{R}^n$, L es estrictamente bajo triangular y U estrictamente sobre triangular, tendremos

$$x^{(k)} = T(x^{(k-1)}) = (L + U) x^{(k-1)} + b$$

Por el último Teorema, el método de Jacobi converge para $\rho(L+U) < 1$.

A su vez el método de Gauss-Seidel :

$$x^{(k)} = L x^{(k)} + U x^{(k-1)} + b$$

lo podemos escribir como :

$$x^{(k)} = (I-L)^{-1} (U x^{(k-1)} + b) ,$$

por el mismo Teorema, el método converge para

$$\rho((I-L)^{-1} U) < 1.$$