

# FUNDAMENTOS SOBRE CONVERGENCIA EN MODELOS ITERATIVOS PARA SISTEMAS DE ECUACIONES LINEALES

José Gil Íñiguez

Vicerrector Académico ULS

## Introducción.

Consideremos el sistema lineal

$$Ax = b \quad (1)$$

donde  $A \in \mathbb{R}^{n \times n}$  y  $x, b \in \mathbb{R}^{n \times 1}$ . Descompongamos la matriz  $A$  como:  $A = S - T$ .  
Entonces el sistema (1) es equivalente a

$$Sx = Tx + b \quad (2)$$

por tanto podemos intentar la iteración

$$Sx^{(k+1)} = Tx^{(k)} + b \quad (3)$$

lo cual es equivalente a

$$x^{(k+1)} = S^{-1}Tx^{(k)} + S^{-1}b; \quad k = 0, 1, 2, \dots$$

La descomposición de la matriz  $A$  será adecuada cuando la sucesión  $\{x^{(k)}\}$  converja a la solución  $x$  del sistema, es decir  $\lim_{k \rightarrow \infty} \{x^{(k)}\} = x$ .

De (2) y (3) obtenemos:

$$\begin{aligned} S(x^{(k+1)} - x) &= T(x^{(k)} - x) \quad y \\ S\varepsilon^{(k+1)} &= T\varepsilon^{(k)} \end{aligned} \quad (4)$$

donde  $\varepsilon^{(k+1)}$  y  $\varepsilon^{(k)}$  son los errores en las iteraciones  $k+1$  y  $k$ .

La ecuación de diferencia (4) tendrá como solución

$$\varepsilon^{(k+1)} = (S^{-1}T)\varepsilon^{(k)}, \quad k = 0, 1, 2, \dots$$

para cualquier error inicial  $\varepsilon^{(0)}$ .

Decir que  $\{x^{(k)}\}_{k \rightarrow \infty} \rightarrow x$  es equivalente a  $\{\varepsilon^{(k)}\}_{k \rightarrow \infty} \rightarrow 0$ .

**Iteración de Jacobi.**

Sean  $S = D = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}$  y

$$T = -(L + U) = - \begin{bmatrix} 0 & a_{12} & \dots & a_{1n} \\ a_{21} & 0 & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & 0 \end{bmatrix}.$$

En la  $k+1$  iteración, el vector  $x^{(k+1)}$  se calcula de acuerdo a (2) por:

$$x^{(k+1)} = -D^{-1}(L + U)x^{(k)} + D^{-1}b. \tag{5}$$

Esta fórmula define el proceso de Jacobi o de los desplazamientos simultáneos. Desarrollando (5), obtenemos

$$x_i^{(k+1)} = -\frac{1}{a_{ii}} \left[ \sum_{j=1}^n a_{ij} x_j^{(k)} - b_i \right]; \quad j \neq i. \tag{6}$$

Como se observa, en cada ciclo iterativo se calculan aproximaciones para todas las incógnitas usando solamente aproximaciones del ciclo anterior.

**Iteración de Gauss – Seidel.**

Este proceso se denomina también de los desplazamientos sucesivos, ya que haciendo  $i = 1$  en (6), tenemos:

$$x_1^{(k+1)} = -\frac{1}{a_{11}} \left[ \sum_{j=2}^n a_{1j} x_j^{(k)} - b_1 \right].$$

Esta aproximación de  $x_1$  se usa en la aproximación de  $x_2$ , o sea:

$$x_2^{(k+1)} = -\frac{1}{a_{22}} \left[ a_{21} x_1^{(k+1)} + \sum_{j=3}^n a_{2j} x_j^{(k)} - b_2 \right]$$

y así sucesivamente, obteniendo en forma general:

$$x_i^{(k+1)} = -\frac{1}{a_{ii}} \left[ \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} + \sum_{j=i+1}^n a_{ij} x_j^{(k)} - b_i \right]; \quad j = 1, 2, \dots, n.$$

o equivalentemente en forma matricial:

$$X^{(k+1)} = -D^{-1}(LX^{(k+1)} + UX^{(k)}) + D^{-1}b$$

o también

$$X^{(k+1)} = -(D+L)^{-1}UX^{(k)} + (D+L)^{-1}b \quad (7)$$

fórmula que define el proceso de Gauss-Seidel. En este caso,  
 $S = D+L$  y  $T = -U$

### Convergencia.

Las fórmulas (5) y (7) se pueden escribir, en forma general:

$$X^{(k+1)} = MX^{(k)} + F \quad (8)$$

El error de truncamiento en la  $k$ -ésima iteración será:

$$\varepsilon^{(k)} = X^{(k)} - X = MX^{(k-1)} + F - MX - F,$$

y por tanto

$$\varepsilon^{(k)} = M \varepsilon^{(k-1)} \quad (9).$$

Repitiendo este proceso llegamos a:

$$\varepsilon^{(k)} = M^k \varepsilon^{(0)}$$

Cualquiera que sea el error inicial  $\varepsilon^{(0)}$ , tenemos:

$$\|\varepsilon^{(k)}\| \leq \|M^k\| \|\varepsilon^{(0)}\|.$$

Cualquier método iterativo será convergente cuando  $\lim_{k \rightarrow \infty} \{\varepsilon^{(k)}\} = 0$

o equivalentemente  $\lim_{k \rightarrow \infty} \{\|\varepsilon^{(k)}\|\} = 0$ , condición que se verificará cuando

$$\lim_{k \rightarrow \infty} \{\|M^k\|\} = 0 \quad \text{ó suficientemente cuando} \quad \lim_{k \rightarrow \infty} \{\|M\|^k\} = 0,$$

y esto se cumplirá si

$$\|M\| < 1 \quad \text{ó si} \quad |\lambda_i| < 1,$$

donde  $\lambda_i$  son los autovalores de  $M$ . Por otra parte, una norma particular para  $M$  de Jacobi,

$$\|M\| = \max_i \sum_{j=1}^n |m_{ij}| \quad j \neq i$$

$\|M\| < 1$  es equivalente a decir que la matriz A del sistema es diagonalmente dominante estrictamente, ya que:

$$\|M\| < 1 \leftrightarrow \max_i \sum_{\substack{j=1 \\ j \neq i}}^n \left| \frac{a_{ij}}{a_{ii}} \right| < 1 \leftrightarrow$$

$$\leftrightarrow \forall i \sum_{\substack{j=1 \\ j \neq i}}^n \left| \frac{a_{ij}}{a_{ii}} \right| < 1 \leftrightarrow \forall i \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| < |a_{ii}|$$

El criterio anterior de convergencia no podrá ser aplicado bajo la premisa expuesta en el libro "Métodos Numéricos y Programación Fortran" (McCracken y Dorn, Ed. Limusa-Wiley, México 1969), donde se asegura que la condición suficiente de convergencia tanto para el método de Jacobi como para el método de Gauss - Seidel es:

$$|a_{ii}| \geq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|$$

para cualquier i y cuando menos para una i,

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| ,$$

condición que nos llevará en algunos casos a que  $\|M\| \geq 1$  , lo cual contradice a la postulación de convergencia explicada anteriormente.

Analizaremos el siguiente sistema:

$$\begin{cases} 9x_1 + 2x_2 - 4x_3 = 13 \\ 3x_1 - 8x_2 + 2x_3 = \frac{1}{2} \\ x_1 + 2x_2 - 4x_3 = -\frac{1}{3} \end{cases}$$

La matriz A del sistema es diagonalmente dominante estrictamente, condición más fuerte que la exigida en el libro de McCracken y Dorn. Sin embargo, utilizando el método de Gauss-Seidel tenemos:

$$D + L = \begin{bmatrix} 9 & 0 & 0 \\ 3 & -8 & 0 \\ 1 & 2 & -4 \end{bmatrix}, \quad U = \begin{bmatrix} 0 & 2 & -4 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\text{Det}(D + L) = 288, \quad (D + L)^{-1} = \begin{bmatrix} \frac{1}{9} & 0 & 0 \\ \frac{1}{24} & -\frac{1}{8} & 0 \\ \frac{7}{144} & -\frac{1}{16} & -\frac{1}{4} \end{bmatrix},$$

$$M = -(D + L)^{-1} U = \begin{bmatrix} 0 & -\frac{2}{9} & \frac{4}{9} \\ 0 & -\frac{1}{12} & \frac{5}{12} \\ 0 & -\frac{7}{72} & \frac{23}{72} \end{bmatrix}.$$

Luego  $\|M\|_{II} = \frac{85}{72} > 1$ , por tanto concluimos que el método de Gauss – Seidel no converge con el criterio expuesto pero sí converge con el criterio de que la matriz del sistema sea diagonalmente dominante.

### Acotación del Error.

Encontraremos ahora una cota superior del error. Si  $\Delta^{(k)}$  es la diferencia entre dos aproximaciones consecutivas, o sea:

$$\Delta^{(k)} = x^{(k-1)} - x^{(k)}$$

$$\Delta^{(k)} = \varepsilon^{(k-1)} - \varepsilon^{(k)}$$

$$\|\Delta^{(k)}\| \geq \|\varepsilon^{(k-1)}\| - \|\varepsilon^{(k)}\|.$$

Por otra parte,

$$\|\varepsilon^{(k)}\| \leq \|M\| \|\varepsilon^{(k-1)}\|, \text{ luego}$$

$$\|\varepsilon^{(k)}\| \leq \|M\| (\|\Delta^{(k)}\| + \|\varepsilon^{(k)}\|),$$

y finalmente

$$\|\varepsilon^{(k)}\| \leq \frac{\|M\| \|\Delta^{(k)}\|}{1 - \|M\|}.$$

El criterio de convergencia analizado anteriormente no es de mucha utilidad si se lo toma al pie de la letra, ya que son pocos los sistemas de ecuaciones lineales que poseen matrices de coeficientes diagonalmente dominante o matrices de iteración con norma menor que uno; sin embargo si se arreglan las ecuaciones para formar el sistema lo más

cercano posible a las condiciones de convergencia anteriormente indicadas, y postulando nuevas condiciones de convergencia como:

Si la sucesión  $\{x^{(k)}\}$  converge a la solución  $x$  del sistema, entonces la sucesión de números reales  $\{\|x^{(k+1)} - x^{(k)}\|\}$  convergerá a cero; si por el contrario esta sucesión de números diverge, entonces puede pensarse que el proceso diverge. Con esta aclaración, un criterio adicional sería detener el proceso una vez que  $\|x^{(k+1)} - x^{(k)}\| < \varepsilon$ . Por lo tanto, si el proceso iterativo diverge (según los criterios explicados al principio), un despeje adecuado de las incógnitas puede originar convergencia; por ejemplo, en lugar de despejar  $x_1$  de la primera ecuación,  $x_2$  de la segunda, etc., se podrían despejar las diferentes  $x_i$  de diferentes ecuaciones, cuidando que los coeficientes de las  $x_i$  despejadas sean distintos de cero; esta sugerencia presenta, para un sistema de  $n$  ecuaciones,  $n!$  distintas formas de reorganizar dicho sistema, pero si respetamos lo más que se pueda el criterio de matriz diagonalmente dominante, debemos despejar  $x_i$  de la ecuación  $i$  donde se cumpla que

$$|a_{ii}| \geq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|, \text{ y si } |a_{ii}| < \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \quad \forall i, \text{ despejar } x_i \text{ de la ecuación } i \text{ donde}$$

se verifique que  $|a_{ii}| - \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|$  tenga el menor valor. Del resto de las

ecuaciones hacemos el mismo análisis para  $x_2$  y así sucesivamente podemos encontrar un "buen criterio" para la convergencia si ésta existiera.

Ilustraremos con algunos ejemplos lo mencionado.

Como primer ejemplo, veamos el sistema:

$$\begin{cases} x_1 + 3x_2 - x_3 = 6 \\ -x_1 + x_2 - x_3 = 0 \\ 2x_1 + x_2 + 2x_3 = 3 \end{cases}$$

En principio, cualquier método iterativo explicado diverge, pues la matriz del sistema no es diagonalmente dominante y menos estrictamente. Sin embargo, si despejamos  $x_1$  de la segunda ecuación,  $x_2$  de la primera y  $x_3$  de la tercera, tenemos:

$$\begin{cases} x_1 = x_2 - x_3 \\ x_2 = -\frac{1}{3}x_1 + \frac{1}{3}x_3 + 2 \\ x_3 = -x_1 - \frac{1}{2}x_2 + \frac{3}{2} \end{cases}$$

Utilizando como aproximación inicial al vector  $X^{(0)} = [0 \ 0 \ 0]^t$  y haciendo 18 iteraciones por Gauss – Seidel, obtenemos la siguiente tabla:

K	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\ X^{(k+1)} - X^{(k)}\ $
0	0.000000	0.000000	0.000000	-
1	0.000000	2.000000	0.500000	2.061553
2	1.500000	1.666667	-0.833333	2.034426
3	2.500000	0.888889	-1.444445	1.040655
4	2.333335	0.740741	-1.203706	0.328148
5	1.944447	0.950617	-0.919755	0.525271
6	1.870372	1.069958	-0.905351	0.141198
7	1.975310	1.039781	-0.995200	0.141405
8	2.034981	0.989941	-1.029951	0.085160
9	2.019892	0.983387	-1.011585	0.024656
10	1.994972	0.997815	-0.993880	0.033803
11	1.991695	1.004810	-0.994100	0.007727
12	1.998909	1.002331	-1.000075	0.009690
13	2.002406	0.999174	-1.001993	0.005087
14	1.999588	0.998948	-1.000641	0.001848
15	1.999588	0.999925	-1.000551	0.002153
16	1.999475	1.000326	-0.999638	0.000426
17	1.999964	1.000134	-1.000031	0.000655
18	2.000164	0.999936	-1.000132	0.000299

El proceso converge a la solución verdadera, que es:

$$X = [2 \quad 1 \quad -1]^t$$

Como un segundo ejemplo, consideremos el sistema

$$\begin{cases} 2x_1 - 4x_2 + x_3 = 1 \\ 2x_1 + x_2 - 2x_3 = 3 \\ x_1 - 2x_2 + 3x_3 = 7 \end{cases}$$

Nuevamente, la matriz del sistema no es diagonalmente dominante. Con el procedimiento explicado, despejamos  $x_1$  de la segunda ecuación,  $x_2$  de la primera ecuación y  $x_3$  de la tercera.

Tenemos entonces el sistema:

**FIDES ET RATIO**

$$\begin{cases} x_1 &= -\frac{1}{2}x_2 + x_3 + \frac{3}{2} \\ x_2 &= \frac{1}{4}x_1 + \frac{1}{4}x_3 - \frac{1}{4} \\ x_3 &= -\frac{1}{3}x_1 + \frac{2}{3}x_2 + \frac{7}{3} \end{cases}$$

La tabla de iteraciones sucesivas es:

K	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\ x^{(k+1)} - x^{(k)}\ $
0	0.000000	0.000000	0.000000	-
1	1.500000	0.125000	1.916669	2.437056
2	3.354169	1.067709	1.927095	2.080084
3	2.893240	0.955084	2.005652	0.480948
4	3.028110	1.008441	1.996267	0.145344
5	2.999406	0.999793	2.000070	0.002936
6	3.002174	1.000061	1.999993	0.000817
7	2.999406	0.999793	2.000070	0.002936
8	3.000174	1.000061	1.999993	0.000817
9	2.999962	0.999989	2.000015	0.000224
10	3.000021	1.000009	2.000009	0.000062
11	3.000005	1.000003	2.000011	0.000017
12	3.000009	1.000005	2.000010	0.000005
13	3.000008	1.000005	2.000010	0.000001
14	3.000008	1.000005	2.000010	0.000000
15	3.000008	1.000005	2.000010	0.000000

La convergencia a la solución verdadera  $x = [3 \ 1 \ 2]^t$  es rápida.

Como tercer ejemplo, consideremos el sistema

$$\begin{cases} x_1 + 2x_2 + 3x_3 &= 9 \\ x_1 - x_2 + 2x_3 &= 9 \\ 2x_1 + 2x_2 - x_3 &= -1 \end{cases}$$

Si despejamos  $x_1$  de la tercera ecuación,  $x_2$  de la segunda y  $x_3$  de la primera, resulta:

$$\begin{cases} x_1 &= -x_2 + \frac{1}{2}x_3 - \frac{1}{2} \\ x_2 &= x_1 + 2x_3 - 9 \\ x_3 &= -\frac{1}{3}x_1 - \frac{2}{3}x_2 + 9 \end{cases}$$



La tabla de iteraciones es la siguiente:

K	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\ X^{(k+1)} - X^{(k)}\ $
0	0.000000	0.000000	0.000000	-
1	-0.500000	-9.500000	15.500003	18.186536
2	16.750001	38.750007	-22.416680	63.744024
3	-50.1458347	-104.291706	95.347271	197.094500
4	151.465342	333.1159884	-263.595097	600.81.3361
5	-465.457433	-1001.647628	831.917741	1833.699317

No es necesario realizar más iteraciones: El proceso diverge rápidamente y se aleja de la solución verdadera, que es  $X = [2 \ -1 \ 3]^t$   
 Ensayando otra forma, despejamos  $x_1$  de la tercera ecuación,  $x_2$  de la primera y  $x_3$  de la segunda, con lo que tenemos

$$\begin{cases} x_1 = -x_2 + \frac{1}{2}x_3 - \frac{1}{2} \\ x_2 = -\frac{1}{2}x_1 - \frac{3}{2}x_3 + \frac{9}{2} \\ x_3 = -\frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{9}{2} \end{cases}$$

La tabla de iteraciones es la siguiente:

K	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\ X^{(k+1)} - X^{(k)}\ $
0	0.000000	0.000000	0.000000	-
1	-0.500000	4.750000	7.125000	8.577769
2	-1.687500	-5.343750	2.671875	11.096138
3	6.179688	-2.597656	0.111328	8.717228
4	2.153320	3.256348	5.051514	8.653694
5	-1.230591	-2.461975	3.884308	6.746291
...	...	...	...	...
57	1.998498	-1.001427	3.00037	0.004004
58	2.001445	-1.000779	2.998888	0.003229
59	2.000223	-0.998443	3.000667	0.003180
60	1.998777	-1.000389	3.000417	0.002437

Vemos en este caso que la convergencia es muy lenta pero que se aproxima a la solución verdadera del sistema original.

Para los métodos directos, Gauss - Doolittle, el número de multiplicaciones y divisiones  $P_n$  y el número de sumas y restas  $S_n$  para triangulizar la matriz del sistema, están dadas por:

$$P_n = \frac{1}{6}(n-1)n(2n-1) + (n-1)n$$

$$S_n = \frac{1}{6}(n-1)n(2n-1) + \frac{1}{2}(n-1)n$$

y el número de multiplicaciones y divisiones  $p_n$  y el número de sumas y restas  $s_n$  para resolver el sistema triangular, están dadas por:

$$p_n = \frac{1}{2}n(n+1)$$

$$s_n = \frac{1}{2}(n-1)n$$

Por tanto, el número total de multiplicaciones y divisiones  $A$  y el número total de sumas y restas  $B$  son:

$$A = P_n + p_n = \frac{1}{3}n^3 + n^2 - \frac{1}{3}n$$

$$B = S_n + s_n = \frac{1}{3}n^3 + \frac{1}{2}n^2 - \frac{5}{6}n$$

Obviamente, el trabajo computacional para resolver un sistema lineal de orden  $n$  por un método directo, es función del número de operaciones necesarias proporcional a  $n^3$  y por otro lado, las necesidades de memoria de máquina serán proporcionales a  $n^2$ , lo que constituye mucha exigencia para sistemas que gracias a sus características poseen muchos elementos nulos y por ende son aconsejables los métodos iterativos.

#### Referencia bibliográfica:

##### - Métodos Numéricos y Programación Fortran

McCracken y Dorn

Limusa-Wiley

##### - Introducción al Análisis Numérico

Anthony Ralston

Limusa-Wiley

##### - Análisis Numérico

Gerald

Alfaomega

##### - Análisis Numérico

Burden Richard

Alfaomega

**- Métodos Numéricos**

Burden-Faires

Thomson Paraninfo S.A.

**- Análisis Numérico**

FJ. Scheid

McGraw-Hill