Métodos Iterativos para resolver sistemas lineales

Lautaro Ochotorena

30 de marzo de 2022

Índice

1. Preliminares

Definición 1.1. Una norma vectorial en $\mathbb{R}^n$ es una función, $\|\cdot\|$, de $\mathbb{R}^n$ a $\mathbb{R}$ con las siguientes propiedades:

i. $\|x\| \geq 0$ para todo $x \in \mathbb{R}^n$

ii. $\|x\| = 0$ sii $x = 0$

iii. $\|\alpha x\| = |\alpha| \, \|x\|$ para todo $\alpha \in \mathbb{R}$ y $x \in \mathbb{R}^n$

iv. $\|x + y\| \leq \|x\| + \|y\|$ para todo $x, y \in \mathbb{R}^n$

Definición 1.2. Las normas l_2 y l_∞ para el vector $x = (x_1, x_2, ..., x_n)$ son definidas como

$$\|x\|_2 = (\sum_{i=1}^{n} x_i^2)^{1/2} \quad \text{y} \quad \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

l_2 es llamada la norma euclídea

Definición 1.3. Una norma matricial es una función $\|\cdot\| : \mathbb{R}^{n \times n} \longrightarrow \mathbb{R}$ que satisface:
Si $A, B \in \mathbb{R}^{n \times n}$

i. $\|A\| \geq 0$

ii. $\|A\| = 0$ sii $A = 0$, la matriz con entradas 0

iii. $\|\alpha A\| = |\alpha| \, \|A\|$

iv. $\|A + B\| \leq \|A\| + \|B\|$

v. $\|AB\| \leq \|A\| \, \|B\|$

Teorema 1.0.1. Si $\|\cdot\|$ es una normal vectorial de $\mathbb{R}^n$, entonces

$$\|A\| = \max_{\|x\|=1} \|Ax\|$$

es una normal matricial que se le suele llamar normal matricial natural o inducida.
Dado cualquier $z \neq 0$, el vector $x = \frac{z}{\|z\|}$ es un vector unitario. Entonces

$$\max_{\|x\|=1} \|Ax\| = \max_{z \neq 0} \left\| A(\frac{z}{\|z\|}) \right\| = \max_{z \neq 0} \frac{\|A(z)\|}{\|z\|}$$

Corolario 1.0.2. Para cualquier $z \neq 0$, una matriz A y una norma natural, tenemos que

$$\|Az\| \leq \|A\| \, \|z\|$$

En particular, si $z = 0$ entonces $Az = 0$ y por lo tanto $\|Az\| = 0 = \|A\| \, \|z\|$

Definición 1.4. Una sucesión $\left\{x^{(k)}\right\}_{k=1}^{\infty}$ de vectores $\mathbb{R}^n$ se dice que converge a un $x \in \mathbb{R}^n$ con la norma $\|\cdot\|$ si, dado un $\epsilon > 0$, existe un entero $N(\epsilon)$ tal que:

$$\left\| x^{(k)} - x \right\| < \epsilon, \text{ para todo } k \geq N(\epsilon)$$

Definición 1.5. Si A es una matriz cuadrada, el polinomio característico de A es definido como

$$p(\lambda) = det(A - \lambda I)$$

Los ceros del polinomio característico se llaman autovalores de A y los x tales que $(A - \lambda I)x = 0$ se llaman autovectores asociados a λ

Definición 1.6. El radio espectral, $\rho(A)$, de una matriz A es definido como

$$\rho(A) = \text{máx}(|\lambda|), \text{ donde } \lambda \text{ es un autovalor de A}$$

Para los complejos $\lambda = \alpha + i\beta$, se define $|\lambda| = (\alpha^2 + \beta^2)^{1/2}$

Teorema 1.0.3. Si A es una matriz cuadrada entonces

$$\rho(A) \leq \|A\|$$

para cualquier norma natural

Demostración.
Sea λ un autovalor de A y x el autovector unitario asociado a λ, entonces $Ax = \lambda x$ y

$$|\lambda| = |\lambda| \, \|x\| = \|\lambda x\| = \|Ax\| \leq \|A\| \, \|x\| = \|A\|$$

Por definición del radio espectral y al tener que para todo autovalor λ se cumple la ecuación de arriba, se concluye que

$$\rho(A) \leq \|A\|$$

□

Definición 1.7. Llamamos a una matriz $A \in \mathbb{R}^{n \times n}$ convergente si:

$$\lim_{k \to \infty} (A^k)_{ij} = 0, \text{ para cada } i, j\text{=1,2,...,n}$$

Teorema 1.0.4. Los siguientes items son equivalentes:

i. A es una matriz convergente

ii. $\rho(A) < 1$

iii. $\text{lím}_{n \to \infty}(A^n x) = 0$, para todo x

Teorema 1.0.5. Todas las normas de $\mathbb{R}^n$ son equivalentes, esto es, dadas $\|\cdot\|_1$ y $\|\cdot\|_2$, existen $a, b \in \mathbb{R}^+$ tales que

$$a \|x\|_1 \leq \|x\|_2 \leq b \|x\|_1$$

para todo $x \in \mathbb{R}^n$

2. Métodos iterativos para resolver sistemas lineales

La idea es encontrar soluciones para ecuaciones del tipo $Ax = b$. En el caso en que A tiene dimensión pequeña, el método de eliminación Gaussiana funciona y da una solución exacta con eficiencia, sin embargo, a medida que A tiene más dimensión, el tiempo de cálculo y almacenamiento de este método crece por lo que se opta por métodos iterativos más eficacez como los que veremos en esta sección.
Sistemas como estos aparecen en análisis de circuitos y en soluciones numéricas de problemas de valor límite y en ecuaciones de derivadas parciales

Una técnica iterativa para resolver ecucaciones $Ax = b$ empiezan con una aproximación inicial $x^{(0)}$ a la solución x y genera una sucesion de vectores $\{x^{(k)}\}_{k=1}^{\infty}$ que converge a x

Los métodos que veremos incluyen un proceso que convierte la ecuación $Ax = b$ en una del tipo $x = Tx + c$ para alguna matriz T y un vector c. Es por ello que después de elegir un vector inicial $x^{(0)}$ se genera una sucesión de vectores haciendo

$$x^{(k)} = Tx^{(k-1)} + c$$

para cada $k = 1, 2, ...$

Veamos cuáles son esos T y c para dos métodos iterativos

2.1. Método de Jacobi

Considera $Ax = b$, con

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Y la separa en:

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} - \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ -a_{21} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ -a_{n1} & \cdots & -a_{n,n-1} & 0 \end{pmatrix} - \begin{pmatrix} 0 & -a_{12} & \cdots & -a_{1n} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -a_{n-1,n} \\ 0 & \cdots & \cdots & 0 \end{pmatrix} = D-L-U$$

Entonces la ecuación queda $(D - L - U)x = b$ por lo que

$$Dx = (L + U)x + b$$

y, si $a_{ii} \neq 0$ para cada i, entonces D^{-1} existe, por lo cual

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

Esto resulta en que consideremos

$$x^{(k)} = D^{-1}(L + U)x^{(k-1)} + D^{-1}b, \text{ para } k = 1, 2, ...$$

Donde $T_j = D^{-1}(L+U)$ y $c_j = D^{-1}b$, llegamos por lo tanto a

$$x^{(k)} = T_j x^{(k-1)} + c_j$$

El método de Jacobi requiere que la matriz A tenga $a_{ii} \neq 0$ para cada i, por ello si la matriz A es inversible entonces se pueden reordenar las ecuaciones de tal forma que los elementos de la diagonal sean no nulos.
¿Cuál es este reordenamiento? Un reodenamiento posible (permutaciones de filas) es el mismo que se da para la eliminación Gaussiana al tratar de hallar la matriz inversa. Acordarse que al hacer las permutaciones en A también hay que hacerlas en el vector b.
Para acelerar la convergencia de este método se debe arreglar las ecuaciones para que los a_{ii} sean lo más grandes posibles.

Ya estamos en condiciones de escribir el pseudocódigo del método de Jacobi:

Input: El número de ecuaciones e incógnitas n; las entradas $a_{ij}, 1 \leq i,j \leq n$ de la matriz A; las entradas $b_i, 1 \leq i \leq n$ del vector b; las entradas $XO_i, 1 \leq i \leq n$ de $XO = x^0$; Tolerancia TOL; máximo de iteraciones N
Output: Una aproximación $x_1, x_2, ..., x_n$ o un mensaje que el número de iteraciones se superó

Paso 1: Setear $k = 1$

Paso 2: Mientras $(k \leq N)$ hacer pasos 3-6

Paso 3: Para $i = 1, ...n$

$$x_i = \frac{1}{a_{ii}}\left[-\sum_{\substack{j=1\\ j\neq i}}^{n}(a_{ij}XO_j) + b_i\right]$$

Paso 4: Si $\|x - XO\| < TOL$ entonces
Output: '$(x_1, x_2, ..., x_n)$, el procedimientio fue exitoso'
Stop

Paso 5: Setear $k = k + 1$

Paso 6: Para $i = 1, 2, ..., n$ setear $XO_i = x_i$

Paso 7: Output: 'Máximo número de iteraciones excedida'
Stop

Otro posible criterio para el paso 4 es iterar hasta que $\frac{\|x^{(k)} - x^{(k-1)}\|}{\|x^{(k)}\|} < TOL$

2.2. Método de Gauss-Seidel

Una posible mejora del algoritmo anterior consiste en utilizar las entradas x_i ya calculadas, esto es:
Si queremos calcular $x_i^{(k)}$ utilizamos las componentes de $x^{(k-1)}$, pero como para $i > 1$ ya se

calcularon $x_1^{(k)}, x_2^{(k)}, ..., x_{i-1}^{(k)}$ y posiblemente sean mejores aproximaciones a las soluciones reales $x_1, x_2, ..., x_{i-1}$, entonces consideraremos

$$x_i^{(k)} = \frac{1}{a_{ii}} \left[-\sum_{j=1}^{i-1} (a_{ij} x_j^{(k)} - \sum_{j=i+1}^{n} (a_{ij} XO_j) + b_i \right]$$

para cada $i = 1, 2, ..., n$. A esta modificación se lo llama Método iterativo de Gauss-Seidel.

Analicemos esto, si multipliicamos a_{ii} por ambos miembros de la ecuación obtenemos

$$a_{i1}x_1^{(k)} + a_{i2}x_2^{(k)} + \cdots + a_{ii}x_i^{(k)} = -a_{i,i+1}x_{i+1}^{(k-1)} - \cdots - a_{i,n}x_n^{(k-1)} + b_i$$

para cada $i = 1, 2, ..., n$ las n ecuaciones quendan:

$$\begin{array}{llll}
a_{11}x_1^{(k)} & = -a_{1,2}x_2^{(k-1)} & -a_{1,3}x_3^{(k-1)} - \cdots - a_{1,n}x_n^{(k-1)} & +b_1 \\
a_{21}x_1^{(k)} + a_{22}x_2^{(k)} & = & -a_{1,2}x_2^{(k-1)} - \cdots - a_{2,n}x_n^{(k-1)} & +b_2 \\
\vdots & & & \\
a_{n1}x_1^{(k)} + a_{n2}x_2^{(k)} + \cdots + a_{nn}x_n^{(k)} & = & & b_n
\end{array}$$

Por las deficiones de D,U,L dadas anteriormente tenemos que el método de Gauss-Seidel está representado por:

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

Y como $D - L$ es una matriz triangular inferior, el determinante es el producto de lo que está en la diagonal que resulta ser distinto de 0 ya que todos los elementos de la diagonal lo son, por ello:

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

para cada $k = 1, 2, ...n$
Llamando $T_g = (D - L)^{-1}U$ y $c_g = (D - L)^{-1}b$ se obtiene la forma

$$x^{(k)} = T_g x^{(k-1)} + c_g$$

El pseudocódigo del método de Gauss-Seidel:

Input: El número de ecuaciones e incógnitas n; las entradas $a_{ij}, 1 \leq i, j \leq n$ de la matriz A; las entradas $b_i, 1 \leq i \leq n$ del vector b; las entradas $XO_i, 1 \leq i \leq n$ de $XO = x^0$; Tolerancia TOL; máximo de iteraciones N
Output: Una aproximación $x_1, x_2, ..., x_n$ o un mensaje que el número de iteraciones se superó

Paso 1: Setear $k = 1$

Paso 2: Mientras $(k \leq N)$ hacer pasos 3-6

Paso 3: Para $i = 1, ...n$

$$x_i = \frac{1}{a_{ii}} \left[-\sum_{j=1}^{i-1} (a_{ij}x_j - \sum_{j=i+1}^{n} (a_{ij}XO_j) + b_i \right]$$

Paso 4: Si $\|x - XO\| < TOL$ entonces
Output: '$(x_1, x_2, ..., x_n)$, el procedimientio fue exitoso'
Stop

Paso 5: Setear $k = k + 1$

Paso 6: Para $i = 1, 2, ..., n$ setear $XO_i = x_i$

Paso 7: Output: 'Máximo número de iteraciones excedida'
Stop

El reordenamiento de la matriz para acelerar la convergencia y el usar otro criterio para el paso 4 también se aplica para este método

2.3. Análisis de convergencia de los métodos iterativos

Uno busca poder aplicar los métodos iterativos y saber que se aproxima a la solución exacta, es por ello que trataremos de averiguar qué requisitos tiene que cumplir una matriz T para poder asegurar esto.
Analicemos la fórmula

$$x^{(k)} = Tx^{(k-1)} + c$$

para $k = 1, 2, ..., n$ y donde $x^{(0)}$ es arbitrario

Lema 2.3.1. Si $\rho(T) < 1$ entonces $(I - T)^{-1}$ existe, y

$$(I - T)^{-1} = \sum_{j=0}^{\infty} T^j$$

Demostración.
Como λ=1 no es autovalor de T (ya que $\rho(T) < 1$) entonces $p_T(1) = det(I - T) \neq 0$. Con lo cual $(I - T)^{-1}$ existe
Sea $S_m = I + T + T^2 + \cdots + T^m$. Entonces

$$(I - T)S_m = (I + T + T^2 + \cdots + T^m) - (T + T^2 + T^3 + \cdots + T^{m+1}) = I - T^{m+1}$$

Usando el teorema 1.0.4 obtenemos que como $\rho(T) < 1$ entonces $\lim_{n\to\infty} T^n x = 0$ para todo x
Por lo cual $\lim_{m\to\infty}(I - T)S_m x = \lim_{m\to\infty}(I - T^{m+1})x = Ix$

Análogamente se puede probar que $\lim_{m\to\infty} S_m(I - T)x = Ix$

De esta forma tenemos que $(I - T)^{-1} = \lim_{m\to\infty} S_m = \sum_{j=0}^{\infty} T^j$ □

Teorema 2.3.2. Para cualquier $x^{(0)} \in \mathbb{R}^n$, la sucesión $\left\{x^{(k)}\right\}_{k=1}^{\infty}$ difinida por

$$x^{(k)} = Tx^{(k-1)} + c, \text{ para cada } k \geq 1$$

converge a la única solución de $x = Tx + c$ si y sólo si $\rho(T) < 1$

Demostración.
$\Leftarrow$) Suponer que $\rho(T) < 1$, entonces

$$\begin{aligned}
x^{(k)} &= Tx^{(k-1)} + c \\
&= T(Tx^{(k-2)} + c) + c \\
&= T^2x^{(k-2)} + (T + I)c \\
&\vdots \\
&= T^kx^{(0)} + (T^{k-1} + \cdots + T + I)c
\end{aligned}$$

y como $\rho(T) < 1$. el teorema 1.0.4 implica que T sea convergente, y

$$\lim_{k\to\infty} T^kx^{(0)} = 0$$

Por el lema anterior tenemos que

$$\lim_{k\to\infty} x^{(k)} = \lim_{k\to\infty} T^kx^{(0)} + (\sum_{j=0}^{\infty} T^j)c = 0 + (I - T)^{-1}c = (I - T)^{-1}c$$

Por lo que la sucesión $\{x^{(k)}\}$ converge al vector $x = (I - T)^{-1}c$ y llegamos a que $x = Tx + c$.

$\Rightarrow$) Probaremos que $\lim_{k\to\infty} T^kz = 0$, para todo $z \in \mathbb{R}^n$, que es equivalente a ver que $\rho(T) < 1$. Sea $z \in \mathbb{R}^n$ y el vector x la única solución a $x = Tx + c$. Defino $x^{(0)} = x - z$ y para cada $k \geq 1$, $x^{(k)} = Tx^{(k-1)} + c$, por hipótesis $\{x^{(k)}\}$ converge a x. Además

$$x - x^{(k)} = (Tx + c) - (Tx^{(k-1)} + c) = T(x - x^{(k-1)})$$

y haciéndolo recursivamente

$$x - x^{(k)} = T(x - x^{(k-1)}) = T^2(x - x^{(k-2)}) = \cdots = T^k(x - x^{(0)})$$

Esto lleva a que $\lim_{k\to\infty} T^kz = \lim_{k\to\infty} T^k(x - x^{(0)}) = \lim_{k\to\infty}(x - x^{(k)}) = 0$.
Por lo que $\rho(T) < 1$ $\square$

Corolario 2.3.3. Si $\|T\| < 1$ para cualquier normal matricial natural y c un vector, entonces la sucesión $\{x^{(k)}\}_{k=1}^{\infty}$ definda por $x^{(k)} = Tx^{(k-1)} + c$ converge, para cualquier $x^{(0)} \in \mathbb{R}^n$, a un vector $x \in \mathbb{R}^n$ con $x = Tx + c$ y las siguientes cotas se mantienen:

(i) $\left\|x - x^{(k)}\right\| \leq \|T\|^k \left\|x^{(0)} - x\right\|$

(ii) $\left\|x - x^{(k)}\right\| \leq \frac{\|T\|^k}{1-\|T\|} \left\|x^{(1)} - x^{(0)}\right\|$

Demostración.
Primero como $\|T\| < 1$ entonces, por teorema 1.0.3, $\rho(T) \leq \|T\| < 1$, con lo cual la sucesión va a converger como vimos en el teorema anterior.
(i) Como se ve en la prueba anterior $\left\|x - x^{(k)}\right\| = \left\|T^k(x - x^{(0)})\right\| \leq \|T\|^k \left\|x - x^{(0)}\right\| = \|T\|^k \left\|x^{(0)} - x\right\|$

(ii)
Veamos ahora
$$x^{(k)} - x^{(k-1)} = T(x^{(k-1)} - x^{(k-2)}) = \cdots = T^{k-1}(x^{(1)} - x^{(0)})$$
Si consideramos $m > n \geq 1$
$$\begin{aligned}
\left\| x^{(m)} - x^{(n)} \right\| &= \left\| x^{(m)} - x^{(m-1)} + x^{(m-1)} - \cdots + x^{(n+1)} - x^{(n)} \right\| \\
&\leq \left\| x^{(m)} - x^{(m-1)} \right\| + \left\| x^{(m-1)} - x^{(m-2)} \right\| + \cdots + \left\| x^{(n+1)} - x^{(n)} \right\| \\
&\leq \left\| T^{m-1} \right\| \left\| x^{(1)} - x^{(0)} \right\| + \left\| T^{m-2} \right\| \left\| x^{(1)} - x^{(0)} \right\| + \cdots + \left\| T^{n} \right\| \left\| x^{(1)} - x^{(0)} \right\| \\
&\underset{\text{def } 1.3}{\leq} \left\| T \right\|^{n} \left\| x^{(1)} - x^{(0)} \right\| (1 + \|T\| + \|T\|^2 + \cdots + \|T\|^{m-n-1})
\end{aligned}$$

Como $\lim_{m\to\infty} x^{(m)} = x$ y la norma es continua, entonces

$$\left\| x - x^{(n)} \right\| = \left\| \lim_{m\to\infty} x^{(m)} - x^{(n)} \right\| = \lim_{m\to\infty} \left\| x^{(m)} - x^{(n)} \right\| \leq \lim_{m\to\infty} \|T\|^n \left\| x^{(1)} - x^{(0)} \right\| \sum_{i=0}^{m-n-1} \|T\|^i \leq$$
$$\leq \|T\|^n \left\| x^{(1)} - x^{(0)} \right\| \sum_{i=0}^{\infty} \|T\|^i$$

Y como $\|T\| < 1$ entonces $\sum_{i=0}^{\infty} \|T\|^i = \frac{1}{1-\|T\|}$, finalmente llegamos a

$$\left\| x - x^{(n)} \right\| \leq \frac{\|T\|^n}{1 - \|T\|} \left\| x^{(1)} - x^{(0)} \right\|$$

□

De esto se obtiene que si $\rho(T_j) < 1$ entonces el método de Jacobi converge a $x = T_j x + c_j$ para cualquier valor inicial dado, como $T_j = D^{-1}(L + U)$ y $c_j = D^{-1}b$ tenemos que

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

Por lo que $Dx = (L + U)x + b \Longrightarrow (D - L - U)x = b \Longrightarrow Ax = b$
Análogamente sucede lo mismo con T_g y c_g en el caso de que $\rho(T_g) < 1$

Hay sistemas $Ax = b$ en los que hay convergencia con un método y con el otro no. Un ejemplo es: $Ax = b$ con $b = (-1, 4, 5)^t$
$$A = \begin{pmatrix} 2 & -1 & 1 \\ 2 & 2 & 2 \\ -1 & -1 & 2 \end{pmatrix}$$
Este sistema tiene solución única, ya que $det(A) \neq 0$, la solución exacta es $(1, 2, -1)^t$,
$\rho(T_j) = \frac{\sqrt{5}}{2} > 1$ y $\rho(T_g) = \frac{1}{2} < 1$

Proposición 2.3.4. Para una matriz T cualquiera y un $\epsilon > 0$ existe una norma matricial natural $\|\cdot\|$ con la propiedad que
$$\rho(T) < \|T\| < \rho(T) + \epsilon$$

Si se tiene una matriz T tal que $\rho(T) < 1$ entonces va a existir un $\epsilon > 0$ tal que $\rho(T) + \epsilon < 1$, el ϵ se elige de esta forma para que $(\rho(T) + \epsilon)^k$ sea cada vez más chico al aumentar k, y usando la norma matricial (y en consecuencia su norma vectorial relacionada) de la proposición anterior se obtiene que

$$\left\| x^{(k)} - x \right\| \leq \|T\|^k \left\| x^{(0)} - x \right\| < (\rho(T) + \epsilon)^k \left\| x^{(0)} - x \right\|$$

lo que da una idea de la rapidez de convergencia del método y como las normas en $\mathbb{R}^n$ son equivalentes (Teorema 1.0.5) entonces la rapidez de convergencia va a ser lineal, similar para cualquier otra norma.
Es por ello que se prefiere elegir el método en el cual $\rho(T)$ sea lo más chico posible.

2.4. Matrices estrictamente diagonal dominantes

Definición 2.1. Una matriz cuadrada es diagonal dominante si

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

para todo $i = 1, 2, ..., n$
Y una matriz cuadrada es strictamente diagonal dominante si

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

para todo $i = 1, 2, ..., n$

Teorema 2.4.1. Si A es una matriz de $n \times n$ y es estrictamente diagonal dominante entonces A es no singular

Demostración.
Probaremos que A es no singular por contradicción.
Supongamos que A es singular y sea $x = (x_i)$ tal que $x \neq 0$ y $Ax = 0$ (al ser singular va a existir un autovector asociado al 0). Sea

$$0 < |x_k| = \max_{1 \leq j \leq n} |x_j|$$

Como $\sum_{j=1}^{n} a_{ij} x_j = 0$ para cada $i = 1, 2, ..., n$ tenemos que para $i = k$

$$a_{kk} x_k = - \sum_{\substack{j=1 \\ j \neq k}}^{n} a_{kj} x_j$$

Por la desigualdad triangular obtenemos que

$$|a_{kk}||x_k| \leq \sum_{\substack{j=1 \\ j \neq k}}^{n} |a_{kj}||x_j|$$

entonces, por ser x_k el máximo en norma

$$|a_{kk}| \leq \sum_{\substack{j=1 \\ j\neq k}}^{n} |a_{kj}| \frac{|x_j|}{|x_k|} \leq \sum_{\substack{j=1 \\ j\neq k}}^{n} |a_{kj}|$$

Pero esto contradice que A sea estrictamente diagonal dominante. Absurdo. Por lo tanto, si $Ax = 0$ entonces $x = 0$ y por lo cual 0 no es un autovalor de A, con lo cual A es no singular □

Proposición 2.4.2. Si A es una matriz estrictamente diagonal dominante y $x^{(0)}$ arbitrario entonces tanto en el método de Jacobi como en el de Gauss-Seidel la sucesión $\left\{x^{(k)}\right\}_{k=1}^{\infty}$ converge a la única solución del sistema $Ax = b$.

Demostración.
Primero notemos que como A es no singular va a existir solución única al sistema $Ax = b$ y al tener que A es estrictamente diagonal dominantente se tiene que

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

y $a_{ii} \neq 0$ para todos los $i = 1, 2, ..., n$ (si algún $a_{ii} = 0$ contradeciría la ecuación de arriba)
Entonces podremos aplicar los métodos, lo que falta ver es que $\rho(T_j) < 1$ y $\rho(T_g) < 1$.
Veamos $T_j = D^{-1}(L + U)$

$$T_j = \begin{pmatrix} \frac{1}{a_{11}} & 0 & \cdots & 0 \\ 0 & \frac{1}{a_{22}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{1}{a_{nn}} \end{pmatrix} \begin{pmatrix} 0 & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -a_{n-1,n} \\ -a_{n1} & \cdots & -a_{n,n-1} & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{-a_{12}}{a_{11}} & \cdots & \frac{-a_{1n}}{a_{11}} \\ \frac{-a_{21}}{a_{22}} & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{-a_{n-1,n}}{a_{n-1,n-1}} \\ \frac{-a_{n1}}{a_{nn}} & \cdots & \frac{-a_{n,n-1}}{a_{nn}} & 0 \end{pmatrix}$$

Supongamos que λ es autovalor de T_j entonces existe $x \neq 0$ tal que $T_j x = \lambda x$

$$\begin{pmatrix} \sum_{\substack{j=1 \\ j\neq 1}}^{n} x_j \frac{-a_{1j}}{a_{11}} \\ \vdots \\ \sum_{\substack{j=1 \\ j\neq n}}^{n} x_j \frac{-a_{nj}}{a_{nn}} \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}$$

Sea

$$0 < |x_k| = \max_{1\leq j\leq n} |x_j|$$

Entonces

$$|\lambda||x_k| = |\lambda x_k| = |\sum_{\substack{j=1 \\ j\neq k}}^{n} x_j \frac{-a_{kj}}{a_{kk}}| \leq \sum_{\substack{j=1 \\ j\neq k}}^{n} |x_j||\frac{a_{kj}}{a_{kk}}|$$

Por lo que si divido por $|x_k|$ y al ser este el máximo

$$|\lambda| \leq \sum_{\substack{j=1 \\ j\neq k}}^{n} |\frac{x_j}{x_k}||\frac{a_{kj}}{a_{kk}}| \leq \sum_{\substack{j=1 \\ j\neq k}}^{n} \frac{|a_{kj}|}{|a_{kk}|}$$

y como A es estrictamente diagonal dominante

$$\frac{1}{|a_k|} < \frac{1}{\sum_{\substack{j=1\\ j\neq k}}^{n} |a_{kj}|}$$

entonces

$$|\lambda| \leq \sum_{\substack{j=1\\ j\neq k}}^{n} \frac{|a_{kj}|}{|a_{kk}|} < 1$$

por lo que $\rho(T_j) < 1$.

Ahora veamos que $\rho(T_g) < 1$. Sea λ un autovalor de T_g entonces existe $x \neq 0$ tal que $T_g x = \lambda x$, como $T_g = (D-L)^{-1}U$ y D invertible

$$\begin{aligned} (D-L)^{-1}Ux &= \lambda x \\ Ux &= (D-L)\lambda x \\ Ux &= D\lambda x - L\lambda x \\ D^{-1}Ux &= (I - D^{-1}L)\lambda x \end{aligned}$$

$$D^{-1}U = \begin{pmatrix} \frac{1}{a_{11}} & 0 & \cdots & 0 \\ 0 & \frac{1}{a_{22}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{1}{a_{nn}} \end{pmatrix} \begin{pmatrix} 0 & -a_{12} & \cdots & -a_{1n} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -a_{n-1,n} \\ 0 & \cdots & \cdots & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{-a_{12}}{a_{11}} & \cdots & \frac{-a_{1n}}{a_{11}} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \frac{-a_{n-1,n}}{a_{n-1,n-1}} \\ 0 & \cdots & \cdots & 0 \end{pmatrix}$$

y

$$(I - D^{-1}L) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \frac{a_{21}}{a_{22}} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \frac{a_{n1}}{a_{nn}} & \cdots & \frac{a_{n,n-1}}{a_{nn}} & 1 \end{pmatrix}$$

Por lo que si se considera $x = (x_i)$, para cada $i = 1, 2, ..., n$

$$\sum_{j=i+1}^{n} \frac{-a_{ij}}{a_{ii}} x_j = \lambda \sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} x_j + \lambda x_i$$

$$-\sum_{j=i+1}^{n} a_{ij}x_j - \lambda \sum_{j=1}^{i-1} a_{ij}x_j = \lambda x_i a_{ii}$$

Siendo $\|x\|_\infty = |x_k|$ entonces puedo armar el vector $y = \frac{x}{|x_k|}$ y este va a seguir siendo autovector asociado a λ, entonces $|y_i| \leq 1$ para todo $i \neq k$ y $|y_k| = 1$. Volviendo a la ecuación anterior, tomando valor absoluto y considerando $i = k$

$$
\begin{aligned}
|\lambda||y_k||a_{kk}| &= |-\sum_{j=k+1}^{n} a_{kj}y_j - \lambda\sum_{j=1}^{k-1} a_{kj}y_j| \\
|\lambda||a_{kk}| &= |-\sum_{j=k+1}^{n} a_{kj}y_j - \lambda\sum_{j=1}^{k-1} a_{kj}y_j| \\
&\leq \sum_{j=k+1}^{n} |a_{kj}||y_j| + |\lambda|\sum_{j=1}^{k-1}|a_{kj}||y_j| \\
&\leq \sum_{j=k+1}^{n} |a_{kj}| + |\lambda|\sum_{j=1}^{k-1}|a_{kj}|
\end{aligned}
$$

Como A es estrictamente diagonal dominante se tiene que

$$
|a_{kk}| > \sum_{\substack{j=1\\ j\neq k}}^{n}|a_{kj}| = \sum_{j=1}^{k-1}|a_{kj}| + \sum_{j=k+1}^{n} |a_{kj}| \geq \sum_{j=1}^{k-1}|a_{kj}| \tag{1}
$$

por lo que $|a_{kk}| - \sum_{j=1}^{k-1}|a_{kj}| > 0$ y entonces

$$
\begin{aligned}
|\lambda|(|a_{kk}| - \sum_{j=1}^{k-1}|a_{kj}|) &\leq \sum_{j=k+1}^{n} |a_{kj}| \\
|\lambda| &\leq \frac{\sum_{j=k+1}^{n}|a_{kj}|}{|a_{kk}| - \sum_{j=1}^{k-1}|a_{kj}|}
\end{aligned}
$$

Además, por la ecuación (1) se desprende que

$$
|a_{kk}| - \sum_{j=1}^{k-1}|a_{kj}| > \sum_{j=k+1}^{n} |a_{kj}|
$$

Finalmente llegamos a que

$$
|\lambda| \leq \frac{\sum_{j=k+1}^{n}|a_{kj}|}{|a_{kk}| - \sum_{j=1}^{k-1}|a_{kj}|} < 1
$$

Por lo tanto $\rho(T_g) < 1$ $\square$