Teoremi e Dimostrazioni orale Gemignani

Teorema 2.2.1 - Errore inerente

Sia $x \in \mathbb{R}$ con $\omega \leq |x| \leq \Omega$ *(no **underflow/overflow**)*, si ha:

$$\begin{aligned} |\epsilon_{in}| = \left| \frac{trn(x) - x}{x} \right| \qquad & \leq u = \beta^{1-t} \\ & = \frac{1}{2}\beta^{1-t} \text{ per } arr(x) \end{aligned}$$

Dimostrazione

Sia $x = (-1)^s \beta^p \alpha$. L'errore **assoluto** $|x - trn(x)|$ è maggiorato dalla distanza di due numeri di macchina **consecutivi**, per cui si ha

$$|trn(x) - x| \leq \beta^{p-t} \tag{1}$$

Inoltre vale $|x| \geq \beta^{p-1}$, quindi

$$|\epsilon_{in}| = \left| \frac{trn(x) - x}{x} \right| \leq \frac{\beta^{p-t}}{\beta^{p-1}} = \beta^{1-t} = u \tag{2}$$

Teorema 3.1.1 - Errore totale

$\epsilon_{tot} = \epsilon_{in} + \epsilon_{alg}$

Dimostrazione

$$\epsilon_{tot} = \frac{g(\tilde{x}) - f(x)}{f(x)} \tag{1}$$

$$= \underbrace{\frac{g(\tilde{x}) - f(\tilde{x})}{f(\tilde{x})}}_{\epsilon_{alg}} \underbrace{\frac{f(\tilde{x})}{f(x)}}_{(1+\epsilon_x)} + \underbrace{\frac{f(\tilde{x}) - f(x)}{f(x)}}_{\epsilon_{in}} \tag{2}$$

$$\begin{aligned} &= \epsilon_{alg}(1 + \epsilon_x) + \epsilon_{in} \\ &\doteq \epsilon_{alg} + \epsilon_{in} \end{aligned} \tag{3}$$

Teorema 4.4.1 - Cerchi di Gershgorin

Sia $A \in \mathbb{C}^{n \times n}$. Definiamo di **Cerchi di Gershgorin** $K_i \ \forall\, 1 \le i \le n$ come

$$K_i = \left\{ z \in \mathbb{C} : |z - a_{i,i}| \le \sum_{j=1, J \ne i}^{n} |a_{i,j}| \ \ \forall\, 1 \le i \le n \right\} \tag{1}$$

Allora

λ autovalore di $A \implies \lambda \in \bigcup_{i=1}^{n} K_i$

Dimostrazione

Sia λ autovalore di A con corrispondente autovettore **destro** x.\
La relazione $Ax = \lambda x$ implica

$$\sum_{j=1}^{n} a_{i,j} x_j = \lambda x_i \ \ \forall\, 1 \le i \le n \tag{1}$$

Estraendo dalla sommatoria l'elemento diagonale rimane

$$\sum_{j=1,\ j \ne i}^{n} a_{i,j} x_j + a_{i,i} x_i = \lambda x_i \qquad \forall\, 1 \le i \le n \tag{2.1}$$

$$\sum_{j=1,\ j \ne i}^{n} a_{i,j} x_j = \lambda x_i - a_{i,i} x_i \qquad \forall\, 1 \le i \le n \tag{2.2}$$

$$\sum_{j=1,\ j \ne i}^{n} a_{i,j} x_j = (\lambda - a_{i,i}) x_i \qquad \forall\, 1 \le i \le n \tag{2.3}$$

Preso p indice della componente di modulo massimo di x (ES: $|x_p| = \|x\|_\infty$)\
Poichè $x = 0 \implies |x_p| > 0$ e per $p = i$ otteniamo

$$\sum_{j=1,\ j \ne i}^{n} a_{p,j} x_j = (\lambda - a_{p,p}) x_p \tag{3}$$

Passando ai valori assoluti

$$|(\lambda - a_{p,p}) x_p| = |(\lambda - a_{p,p})||x_p| = \left| \sum_{j=1,\ j \ne i}^{n} a_{p,j} x_j \right| \le \sum_{j=1,\ j \ne i}^{n} |a_{p,j}||x_j| \tag{4}$$

Dividendo per $|x|$

$$|(\lambda - a_{p,p})| \le \sum_{j=1,\ j \ne i}^{n} |a_{p,j}| \tag{5}$$

Da cui la tesi

Teorema 5.1.1 - Esistenza e Unicità di LU

Data $A \in \mathbb{R}^{n \times n}$. Se $A(1:k, 1:k)$ e' invertibile $\forall\ 1 \leq k \leq n-1$ allora esiste unica ($\exists!$) LU di A

Dimostrazione

Per induzione su n

1. Per $n = 1$ si ha $A = [a] = [1][a]$ unica fattorizzazione LU di A
2. Supponiamo la veridicita' del teorema per A di ordine $m \leq n-1$
3. Costruiamo LU per A di ordine n

$$\left[\begin{array}{c|c} A(1:n-1, 1:n-1) & z \\ \hline v^T & \alpha \end{array}\right] = \left[\begin{array}{c|c} L(1:n-1, 1:n-1) & 0 \\ \hline \omega^T & 1 \end{array}\right] \left[\begin{array}{c|c} U(1:n-1, 1:n-1) & y \\ \hline 0^T & \beta \end{array}\right]$$

Come sistema

$$\begin{cases} A(1:n-1, 1:n-1) & = L(1:n-1, 1:n-1)U(1:n-1, 1:n-1) & (1) \\ z & = A(1:n-1, 1:n-1)y & (2) \\ v^t & = \omega^T U(1:n-1, 1:n-1) & (3) \\ \alpha & = \omega^T y + \beta & (4) \end{cases}$$

Per ipotesi $A(1:n-1, 1:n-1)$ invertibile e per tale motivo $\exists! LU$ con fattori triangolari (equazione (1) del sistema)

- $L(1:n-1, 1:n-1)$
- $U(1:n-1, 1:n-1)$

Sempre dall'invertibilita' di $A(1:n-1, 1:n-1)$ segue che

1. $L(1:n-1, 1:n-1)$ e' invertibile per costruzione
2. $U(1:n-1, 1:n-1)$ e' invertibile

> **Important**
>
> $\det(A) = \det(LU) = \det(L) * \det(U)$ e $\det(L) = 1$ per costruzione. Segue che $\det(A) = \det(U)$, se $\det(A) \neq 0 \implies \det(U) \neq 0$

Di conseguenza le equazioni (2) e (3) del teorema ammetto unica soluzione. Anche l'equazione (4) ammette unica soluzione poiche' conseguenza delle equazioni precedenti

---

Teorema 6.1.2 - Convergenza dei Metodi Iterativi (Cond. Suff.)

Dato il metodo iterativo

$$\begin{cases} x^{(0)} \in \mathbb{R}^n & & (1) \\ x^{(k+1)} = Px^{(k)} + q & k \geq 0 & (2) \end{cases}$$

> **Note**
>
> Con $P = M^{-1}N, \; q = M^{-1}b, \; A = M - N$

Il metodo converse se esiste ($\exists$) una norma matriciale indotta su $\mathbb{R}^n$ t.c. $\|P\| < 1$

Dimostrazione

Prese le relazioni

$$x^{(k+1)} = Px^{(k)} + q \qquad\qquad x = Px + q$$

E sottratte tra di loro

$$\underbrace{x^{(k+1)} - x}_{e^{k+1}} = \underbrace{P(x^{(k)} - x)}_{Pe^{k}}$$

$$e^{(k+1)} = Pe^{(k)} \tag{2}$$

Stando alla relazione (2) e' facile dimostrare che

> **Important**
>
> $$\begin{aligned} e^{(k+1)} &= Pe^{(k)} \\ &= P(Pe^{(k-1)}) \\ &= P(\underbrace{P * P * \cdots * P}_{\text{k volte}} e^{(0)}) \\ &= P^{k+1}e^{(0)} \end{aligned}$$

Passando alla norma

$$\|e^{(k+1)}\| = \|P^{k+1}e^{(0)}\| \leq \|P\|^{k+1}\|e^{(0)}\|$$

> **Important**
>
> Poiche' $\|P\| < 1$ per ipotesi del teorema, $\lim_{k\to\infty} \|P\|^{k+1}\|e^{(0)}\| = 0$ e per costruzione del teorema $0 \leq \|e^{(k+1)}\|$

Segue dal teorema del confronto che $\forall e^{(0)}$ o, equivalentemente, $\forall x^{(0)}$

$$\lim_{k\to\infty} \|e^{(k+1)}\| = \lim_{k\to\infty} \|x^{(k+1)} - x\| = 0$$

Teorema 6.1.3 - Convergenza per Raggio Spettrale (Cond. Nec. e Suff.)

Dato il metodo iterativo

$$\begin{cases} x^{(0)} \in \mathbb{R}^n & (1) \\ x^{(k+1)} = Px^{(k)} + q \qquad k \geq 0 & (2) \end{cases}$$

> **Note**
>
> Con $P = M^{-1}N,\ q = M^{-1}b,\ A = M - N$

Se $\rho(P) < 1$ allora il metodo e' convergente

> **Note**
>
> Dove $\rho(A) = \max_i |\lambda_i|$ con $\lambda_1, \lambda_2, \ldots, \lambda_n$ autovalori di A e' detto **raggio spettrale di** A

Teorema 6.3.1 - Convergenza di Jacobi e Gauss-Seidel

Data una matrice $A \in \mathbb{R}^{n \times n}$, se questa e' predominante diagonale allora

> **Note**
>
> Una matrice $A \in \mathbb{R}^{n \times n}$ si dice **predominante diagonale** se $\forall_i \; 1 \leq i \leq n \;\; |a_{i,i}| > \sum_{j=1,\, j \neq i}^{n} |a_{i,j}|$ ovvero: i moduli degli elementi diagonali sono maggiori della somma dei moduli di ogni riga

1. A e' invertibile
2. **Jacobi** e **Gauss-Seidel** sono applicabili

   > **Note**
   >
   > I due metodi sono applicabili se $\forall_i \; 1 \leq i \leq n \;\; a_{i,i} \neq 0$

3. **Jacobi** e **Gauss-Seidel** sono convergenti

Dimostrazione

1. L'invertibilita' di A segue da ***Gershgorin*** infatti vale

   $$|0 - a_{i,i}| = |a_{i,i}| < \sum_{j=1,\, j \neq i}^{n} |a_{i,j}| \qquad\qquad \forall_i \; 1 \leq i \leq n$$

   Ma per definizione di **predominanza diagonale** le equazioni cosi' costruite non vengono mai verificate, per cui $0 \notin \bigcup_{i=1}^{n} K_i$

2. L'applicabilita' di **Jacobi** e **Gauss-Seidel** si ha poiche'

   $$|a_{i,i}| > \sum_{j=1,\, j \neq i}^{n} |a_{i,j}| \implies a_{i,i} \neq 0 \qquad\qquad \forall_i \; 1 \leq i \leq n$$

   > **Tip**
   >
   > Prendiamo $|a| > |b| + \cdots + |z|$ supponendo che la somma sia 0 questo implica che $|a| > 0 \implies a \neq 0$. Infatti una matrice che ha almeno uno 0 sulla diagonale principale sicuramente **non** e' predominante diagonale. Unendo questo ragionamento all'applicabilita' di Jacobi e GS si ottiene che per A predominanti diagonali i due metodi sono sempre applicabili

3. Per dimostrare la convergenza di **Jacobi** e **Gauss-Seidel** su matrici predominanti diagonali studiamo il polinomio caratteristico della matrice $

> **Note**
>
> - Per Jacobi $M = D$ e $N = L + U$ (M e' la diagonale di A, N quello che resta cambiato di segno)
> - Per Gauss-Seidel $M = D - L$ e $N = U$ (M e' la parte triangolare inferiore di A, N quello che resta cambiato di segno)
>
> Per entrambi $P = M^{-1}N$

$$\begin{aligned} \det(P - \lambda I_n) = & & \text{Polinomio caratteristico di } P \\ &= \det(M^{-1}N - \lambda M M^{-1}) & \text{Riscrivo } P \text{ e } I \text{ in funzione di } M \\ &= \det(M^{-1}) \det(N - \lambda M) & \text{Raccoldo e } \det(A * B) = \det(A) * \det(B) \\ &= \det(N - \lambda M) & \text{Il } \det(M^{-1}) \neq 0, \text{ cerchiamo lo } 0 \end{aligned}$$

Di conseguenza $\lambda \in \mathbb{C}$ e' autovalore di P se e soltanto se $\det(N - \lambda M) = 0$

SUPPONIAMO che $|\lambda| \geq 1$ e dimostriamo che la matrice $\lambda M - N$ (posso cambiare di sengo perche' sto cercando lo 0) e' predominante diagonale

$$|a_{i,i}| > \sum_{j=1,\, j \neq i}^{n} |a_{i,j}| = \underbrace{\sum_{j=1}^{i-1} |a_{i,j}|}_{\text{per } M} + \underbrace{\sum_{j=i+1}^{n} |a_{i,j}|}_{\text{per } N} \qquad \forall_i\ 1 \leq i \leq n \tag{1}$$

$$\begin{aligned} |\lambda||a_{i,i}| &> |\lambda| \sum_{j=1}^{i-1} |a_{i,j}| + |\lambda| \sum_{j=i+1}^{n} |a_{i,j}| & (2) \\ &> |\lambda| \sum_{j=1}^{i-1} |a_{i,j}| + \sum_{j=i+1}^{n} |a_{i,j}| & (3) \end{aligned}$$

> **Important**
>
> (1) spezzo la sommatoria nella parte sotto la diagonale piu' la parte sopra la diagonale
>
> (2) moltiplico le componenti per $|\lambda|$ poiche sto lavoranto per $\lambda M - N$
>
> (3) eliminando $|\lambda|$ dalla parte della sovradiagonale rilasso il problema e la disequazione rimane verificata

Riportando quanto trovato sui due metodi ottengo

- Per **Jacobi**

$$|\lambda a_{i,i}| > \sum_{j=1}^{i-1} |a_{i,j}| + \sum_{j=i+1}^{n} |a_{i,j}|$$

> **Important**
>
> Ho rimosso $|\lambda|$ sulla prima sommatoria poiche' per Jacobi M e' composta dalla diagonale principale di A ma nella sommatoria non la considero

- Per **Gauss-Seidel**

$$|\lambda a_{i,i}| > \sum_{j=1}^{i-1} |\lambda a_{i,j}| + \sum_{j=i+1}^{n} |a_{i,j}|$$

> **Caution**
>
> Dalle relazioni segue che per $|\lambda| \geq 1$ la matrice $\lambda M - N$ rimande predominante diagonale per chui $\det(\lambda M - N) \neq 0$ a differenza di quanto supposto all'inizio. Per cui $|\lambda| \geq 1$ non sono autovalori di P e di conseguenza rimangono solo $|\lambda| < 1 \implies \rho(P) < 1$ e quindi si ha consergenza

Teorema 10.1.1 - Metodo di Bisezione

Caution

```
a(1) = a;
b(1) = b;
for k >= 1
    c(k) = (a(k) + b(k))/2;
    if (f(a(k)) * f(b(k))) <= 0
        a(k + 1) = a(k);
        b(k + 1) = c(k);
    else
        a(k + 1) = c(k);
        b(k + 1) = b(k);
    end
end
```

Sia $f : [a, b] \to \mathbb{R}$ con $f \in C^0([a, b])$ e $f(a)f(b) < 0$, per le successioni generate da metodo di **bisezione** si ha:

Note

$f \in C^0([a, b])$ indica una funzione continua su $[a, b]$ mentre $f \in C^{1,2,...,\infty}([a, b])$ indica una funzione derivabile $1, 2, \dots$ volte o ∞ (infinite volte) su $[a, b]$

$$\lim_{k\to\infty} a_k = \lim_{k\to\infty} b_k = \lim_{k\to\infty} c_k = \xi \in [a, b] \qquad \text{con } f(\xi) = 0 \tag{1}$$

Dimostrazione

Per costruzione $a_{k+1} \geq a_k$, $b_{k+1} \leq b_k$ e $c_k \in [a_k, b_k] \subset [a, b]$. Inoltre $0 \leq b_k - a_k \leq (b - a)/2^{k-1}$ e $f(a_k)f(b_k) \leq 0$ per $k \geq 1$

Segue che $\exists \xi,\ \eta \in [a, b]$ tali che

$$\lim_{k\to\infty} a_k = \xi \qquad e \qquad \lim_{k\to\infty} b_k = \eta \tag{2}$$

Confrontando con (1) per il teorema del confronto si ottiene

$$\begin{aligned} \xi = \lim_{k\to\infty} a_k = \lim_{k\to\infty} b_k &= \eta \\ &= \lim_{k\to\infty} c_k \end{aligned} \tag{2}$$

Di conseguenza

$$\lim_{k\to\infty} f(a_k)f(b_k) = \lim_{k\to\infty} f(\xi)f(\xi) = f(\xi)^2 \leq 0$$

$$\implies f(\xi) = 0$$

Important

L'unico numero che elevato al quadrato fa 0 e' lo 0 stesso

Teorema 10.2.2 - Teorema del Punto Fisso

> **Caution**
>
> Il metodo di iterazione funzionale fornito e'
>
> $$\begin{cases} x_0 \in [a,b] \\ x_{k+1} = g(x_k) & k \geq 0 \end{cases}$$
>
> Con $g(x)$ funzione **equivalente** di $f(x)$
>
> Data una funzione $f(x) = 0$ e $g(x) - x = 0$ questa si dicono **equivalenti** se $f(\xi) = 0 \iff g(\xi) = \xi$ e ξ, radice dell'equazione $f(x) = 0$, e' detta **punto fisso** di $g(x)$

Sia $g : [a,b] \to \mathbb{R},\ g \in C^1([a,b]),\ g(\xi) = \xi,\ \xi \in (a,b)$. Se $\exists \rho > 0$ tale che $|g'(x)| < 1\ \forall x \in [\xi - \rho, \xi + \rho] = I_\xi \subset [a,b]$ allora $\forall x_0 \in I_\xi$ la successione generata dal metodo di iterazione funzionale soddisfa

1. $x_k \in I_\xi$ per ogni $k \geq 0$
2. $\lim_{k \to \infty} x_k = \xi$

Dimostrazione

Dal teorema di Weierstrass essendo $g'(x)$ continua e limitata su I_ξ (intervallo chiuso e limitato) abbiamo

$$\lambda = \max_{k \in I_\xi} |g'(x)| < 0 \tag{1}$$

Pertanto si dimostra che la successione geenerata a dartire da $x_0 \in I_\xi$ soddisfa

$$|x_k - \xi| \leq \lambda^k \rho \qquad k \geq 0 \tag{2}$$

> **Note**
>
> In parole stiamo dicendo che alla k-esima iterazione la distanza tra ila x generata e la soluzione ξ ($|x_k - \xi|$) e' minore $\lambda^k \rho$
>
> Nota bene che $\lambda < 1$ (per (1)) e per cui $\lambda^k \to 0$ e ρ e' il raggio dell'intorno

Per cui $\lambda^k \rho \leq \rho$ ne segue da (2) che le x_k generate appartengono ad un intorno di ξ piu' piccolo $\implies x_k \in I_\xi$ ovvero il punto 1. del teorema

Poiche' $\lambda^k \to 0$ e $0 \leq |x_k - \xi| \leq \lambda^k \rho$ avremo che $0 \leq |x_k - \xi| \leq 0$ e quindi $\lim_{k \to \infty} |x_k - \xi| = 0$ ovvero il punto 2. del teorema

Per dimostrare (2) procediamo per induzione su k

1. Per $k = 0$ da $x_0 \in I_\xi$ si ha

$$|x_k - \xi| \leq \lambda^0 \rho = \rho$$

   Ok per ipotesi (2)

2. Assumiamo l'ipotesi vera fino all'indice k, per il teorema di **Lagrange**

   > **Note**
   >
   > Lagrange dice che $f'(\eta) = \frac{f(a) - f(b)}{a - b}$
   >
   > Nel nostro caso $g'(\eta)(x_k - \xi) = \frac{g(x_k) - g(\xi)}{x_k - \xi}(x_k - \xi) = g(x_k) - g(\xi)$

$$|x_{k+1} - \xi| = |g(x_k) - g(\xi)| = |g'(\eta_k)(x_k - \xi)| = |g'(\eta_k)||(x_k - \xi)|$$

   Per ipotesi induttiva $\eta_k \in I_\xi$ e dunque

$$|x_{k+1} - \xi| = |g'(\eta_k)||(x_k - \xi)| \leq \lambda \lambda^k \rho = \lambda^{k+1} \rho$$

Teorema 10.2.3 - Convergenza Locale dei Metodi di Iterazione Funzionale

Sia $g : [a, b] \to \mathbb{R},\ g \in C^2([a, b]),\ g(\xi) = \xi,\ \xi \in (a, b)$ se $|g'(\xi)| < 1$ allora il metodo e' localmente convergente in ξ

Dimostrazione

Sia $h : [a, b] \to \mathbb{R},\ h(x) = |g'(x)| - 1.$

Si ha che $h \in C^0([a, b]),\ h(\xi) = |g'(\xi)| - 1 < 0$ e dunque per il teorema della permanenza del segno $\exists I_\xi = [\xi - \rho, \xi + \rho] \subset [a, b]$ tale che $h(x) = |g'(x)| - 1 < 0\ \forall x \in I_\xi$

Pertanto la tesi segue dal teorema 10.2.2

Teorema 10.3.1 - Convergenza Locale per Radici Semplici

> **Caution**
>
> Metodo di iterazione fornito (**tangenti**)
>
> $$\begin{cases} x_0 \in [a,b] \\ x_{k+1} = g(x_k) = x_k - \frac{f(x_k)}{f'(x_k)} & \quad k \geq 0 \end{cases}$$

Sia $f : [a,b] \to \mathbb{R},\ f \in C^2([a,b]),\ f(\xi) = 0,\ f'(\xi) \neq 0,\ \xi \in (a,b)$

Allora il metodo e' localmente convergente in ξ e $\exists \rho > 0$ tale che $\forall x_0 \in [\xi - \rho, \xi + \rho] = I_\xi \subset [a,b]$ la successione generata soddisfa

1. $x_k \in I_k$ per ogni $k \geq 0$
2. $\lim_{k \to \infty} x_k = \xi$

Se inosltre la successione verifica $x_k \neq \xi,\ k \geq 0$ allora la convergenza e' almeno quadratica

$$\lim_{k \to \infty} \frac{|x_{k+1} - \xi|}{|x_k - \xi|^2}$$

Dimostrazione

Da $f'(\xi) \neq 0$ per il teorema della permanenza del segno segue che $\exists I'_\xi = [\xi - \rho', \xi + \rho'] \subset [a,b]$ tale che $f'(x) \neq 0\ \forall x \in I'_\xi$

> **Important**
>
> Stiamo dicendo che la derivata nel punto ξ non si annulla e per tali motivi esiste un intorno di ξ in cui la derivata non si annulla

Da questo segue che $g : I'_\xi \to \mathbb{R},\ g(x) = x - \frac{f(x)}{f'(x)}$ e' ben definita e soddisfa $g \in C^1(I'_\xi)$ con $g'(x) = \frac{f(x)f''(x)}{f'(x)^2}$

Poiche' $g'(\xi) = \frac{f(\xi)f''(\xi)}{f'(\xi)^2} = \frac{0}{f'(x)^2} = 0$ (per costruzione di g') i punti 1. e 2. sono verificati dal teorema **10.2.3** (Per $|g'(\xi)| < 1$ ok 1., ok 2.)

Per la stima della velocita' di convergenza dallo sviluppo di tailor al secondo ordine

$$0 = f(\xi) = f(x_k) + f'(x_k)(\xi - x_k) + \frac{f''(\eta_k)(\xi - x_k)^2}{2}$$

$$\text{SEGUE} \qquad x_{k+1} - \xi = \frac{f''(\eta_k)(\xi - x_k)^2}{2f'(x_k)}$$

Per continuita' di $f'(x)$ e $f''(x)$

$$\lim_{k \to \infty} \frac{|x_{k+1} - \xi|}{|x_k - \xi|^2} = \left| \frac{f''(\xi)}{2f'(\xi)} \right| \in \mathbb{R}$$