PDF TEOREMI CALCOLO NUMERICO

Interpolazione polinomiale

Teorema

Siano $(x_0, y_0), (x_1, y_1), ..., (x_n, y_n) \in \mathbb{R}^2$ tali che $x_0, x_1, ..., x_n$ sono tutti punti distinti. Allora esiste un unico polinomio $p(x) \in \mathbb{R}_n[x]$ tale che $p(x_i) = y_i \ \ \forall i = 0, ..., n$

Dim (1)

Sia $p(x) = a_0 + a_1 x + a_2 + x^2 + ... + a_n x^n \in R_n[x]$.
Questo soddisfa la proprietà:

$$p(x_i) = y_i \ \ \forall i = 0, ..., n$$

se e solo se

$$\begin{cases} a_0 + a_1 x_0 + a_2 {x_0}^2 + ... + a_n {x_0}^n = y_0 \\ a_0 + a_1 x_1 + a_2 {x_1}^2 + ... + a_n {x_1}^n = y_1 \\ \vdots \\ a_0 + a_1 x_n + a_2 {x_n}^2 + ... + a_n {x_n}^n = y_n \end{cases}$$

Dunque se il vettore dei coefficienti $(a_0, a_1, ..., a_n)^\top$ soddisfa il sistema lineare:

$$\begin{bmatrix} 1 & x_0 & {x_0}^2 & \dots & {x_0}^n \\ 1 & x_1 & {x_1}^2 & \dots & {x_1}^n \\ 1 & x_2 & {x_2}^2 & \dots & {x_2}^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & {x_n}^2 & \dots & {x_n}^n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad (S)$$

Dimostreremo che la matrice di vandermonde $(V(x_0, ..., x_n))$ è invertibile, dimostrando che il suo determinante è uguale a:

$$\det[V(x_0, x_1, x_2, ..., x_n)] = \begin{cases} 1, \text{se } n = 0 \\ \prod_{i,j=0\ ,\ j<i}^{n} (x_i - x_j) = (x_1 - x_0) \cdot (x_2 - x_1)(x_2 - x_0) \cdot (x_3 - x_2)(x_3 - x_1)(x_3 - x_0), se\ n \geq 1 \end{cases} \quad (\star)$$

da cui segue che $\det[V(x_0, x_1, x_2, ..., x_n)] \neq 0$ in quanto $x_0, ..., x_n$ sono tutti punti distinti.
Dunque esiste un unica soluzione del sistema S, ovvero:

$$\begin{bmatrix} 1 & x_0 & {x_0}^2 & \dots & {x_0}^n \\ 1 & x_1 & {x_1}^2 & \dots & {x_1}^n \\ 1 & x_2 & {x_2}^2 & \dots & {x_2}^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & {x_n}^2 & \dots & {x_n}^n \end{bmatrix}^{-1} \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad (\alpha)$$

dunque esiste un unico polinomio $p(x) \in \mathbb{R}_n[x]$ tale che $p(x_i) = y_i \ \ \forall i = 0, ..., n$. Questo polinomio ha gli stessi coefficienti dati da (α)

Per concluedere la dimostrazione, resta solo da dimostrare $(\star)$ e lo facciamo per $n = 3$, tanto la dimostrazione è la stessa per n generico.

Dim di $(\star)$ per $n = 3$. $\forall i = 1, ..., 3$ definisco
$d_i = det[V(x_0, ..., x_i)]$. Il nostro obiettivo è calcolare d_3.

$$d_3 = \begin{bmatrix} 1 & x_0 & {x_0}^2 & {x_0}^3 \\ 1 & x_1 & {x_1}^2 & {x_1}^3 \\ 1 & x_2 & {x_2}^2 & {x_2}^3 \\ 1 & x_3 & {x_3}^2 & {x_3}^3 \end{bmatrix} =^{C[4]\to C[4]-x_3C[3]}=$$

$$\begin{bmatrix} 1 & x_0 & {x_0}^2 & {x_0}^3 - x_0^2x_3 \\ 1 & x_1 & {x_1}^2 & {x_1}^3 - x_1^2x_3 \\ 1 & x_2 & {x_2}^2 & {x_2}^3 - x_2^2x_3 \\ 1 & x_3 & {x_3}^2 & 0 \end{bmatrix} =^{C[3]\to C[3]-x_3C[2]}=\$$$

$$\begin{bmatrix} 1 & x_0 & {x_0}^2 - x_0x_3 & {x_0}^2(x_0 - x_3) \\ 1 & x_1 & {x_1}^2 - x_1x_3 & {x_1}^2(x_1 - x_3) \\ 1 & x_2 & {x_2}^2 - x_2x_3 & {x_2}^2(x_2 - x_3) \\ 1 & x_3 & 0 & 0 \end{bmatrix} =^{C[2]\to C[2]-x_3C[1]}=$$

$$\begin{bmatrix} 1 & x_0 - x_3 & x_0(x_0 - x_3) & {x_0}^2(x_0 - x_3) \\ 1 & x_1 - x_3 & x_1(x_1 - x_3) & {x_1}^2(x_1 - x_3) \\ 1 & x_2 - x_3 & x_2(x_2 - x_3) & {x_2}^2(x_2 - x_3) \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Sviluppiamo con La Place sull'ultima riga

$$(-1)^{n=3} \cdot \begin{vmatrix} x_0 - x_3 & x_0(x_0 - x_3) & {x_0}^2(x_0 - x_3) \\ x_1 - x_3 & x_1(x_1 - x_3) & {x_1}^2(x_1 - x_3) \\ x_2 - x_3 & x_2(x_2 - x_3) & {x_2}^2(x_2 - x_3) \end{vmatrix} =$$

$$(-1)^{n=3}(x_0 - x_3)(x_1 - x_3)(x_2 - x_3) \cdot \begin{vmatrix} 1 & x_0 & {x_0}^2 \\ 1 & x_1 & {x_1}^2 \\ 1 & x_2 & {x_2}^2 \end{vmatrix} \dots$$

$$\begin{aligned} &= (x_3 - x_0)(x_3 - x_1)(x_3 - x_2)d_2 \\ &= (x_3 - x_0)(x_3 - x_1)(x_3 - x_2) \cdot (x_2 - x_0)(x_2 - x_1)d_1 \\ &= (x_3 - x_0)(x_3 - x_1)(x_3 - x_2) \cdot (x_2 - x_0)(x_2 - x_1) \cdot (x_1 - x_0) \end{aligned}$$

Per ricorrenza, anche d_2 ammette uno sviluppo analogo a d_3 e quindi otteniamo

Abbiamo dimostrato $(\star)$ $\square$

Dim (2)

$\forall j = 0, \dots n$ definisco il polinomio:

$$L_j(x) = \prod_{i=0 \ i \neq j}^{n} \frac{x - x_i}{x_j - x_i} = \frac{(x - x_0) \cdot \dots \cdot (x - x_{j-1}) \cdot (x - x_{j+1} \cdot \dots \cdot (x - x_n))}{(x_j - x_0) \cdot \dots \cdot (x_j - x_{j-1}) \cdot (x_j - x_{j+1}) \cdot \dots \cdot (x_j - x_n)}$$

Gli $n+1$ polinomi $L_0(x), L_1(x), \dots, L_n(x)$ hanno tutti grado n e quindi appartengono a $\mathbb{R}_n[x]$. Mostriamo che ora essi costituiscono una base di $\mathbb{R}_n[x]$ e per fare ciò ci basta dimostrare che essi sono lin. ind. in quanto essi sono in numero $n+1 = \dim(\mathbb{R}_n[x])$.
Per dimostrare che sono lin. ind, osserviamo che:

$$\forall i, j = 0, \dots, n$$

$$L_j(x_i) = \begin{cases} 1 \text{ se } i = j \\ 0 \text{ se } i \neq j \end{cases}$$

> **Caso 1 $(i = j)$:**
> Se sostituiamo x_j in $L_j(x)$ otteniamo:
>
> $$L_j(x_j) = \prod_{i=0 \ i \neq j}^{n} \frac{x_j - x_i}{x_j - x_i}$$

> ma ogni termine della frazione è del tipo:
>
> $$\frac{x_j - x_i}{x_j - x_i}$$
>
> Quindi il prodotto di tutti questi termini è 1
>
> **Caso 2** $(i \neq j)$
> Ora valutiamo $L_j(x)$ in un altro nodo x_i, con $i \neq j$
>
> $$L_j(x_i) = \prod_{k=0 \ k \neq j}^{n} \frac{x_i - x_k}{x_j - x_k}$$
>
> Nota che nel **numeratore** compare un termine $x_i - x_i$ (dato che i fa parte dell'indice del prodotto)
> Dunque essendoci nel numeratore un fattore che diventa 0, tutto il prodotto è uguale a 0:
>
> $$L_j(x_i) = 0$$

Se

$$a_0 L_0(x) + a_1 L_1(x) + ... + a_n L_n(x) = 0 \ \ \forall x \in \mathbb{R}$$

allora in particolare deve essere che

$$a_0 L_0(x_i) + a_0 L_1(x_i) + ... + a_n L_n(x_i) = 0 \ \ \forall i = 0, ..., n$$

cioè

$$a_i = 0 \ \ \forall i = 0, ...n$$

> Abbiamo la seguente combinazione lineare:
>
> $$\alpha_0 L_0(x) + \alpha_1 L_1(x) + \cdots + \alpha_n L_n(x) = 0, \quad \forall x \in \mathbb{R}.$$
>
> Questo significa che vogliamo trovare i coefficienti $\alpha_0, \ldots, \alpha_n$ per cui questa combinazione lineare è sempre il polinomio nullo.
>
> Ora valutiamo questa equazione **nei nodi di interpolazione** $x_0, x_1, \ldots, x_n$, ovvero:
>
> $$\alpha_0 L_0(x_i) + \alpha_1 L_1(x_i) + \cdots + \alpha_n L_n(x_i) = 0, \quad \forall i = 0, \ldots, n.$$
>
> Per l'osservazione vista prima:
>
> $$L_j(x_i) = \begin{cases} 1, & \text{se } i = j \\ 0, & \text{se } i \neq j. \end{cases}$$
>
> Abbiamo che il seguente sistema lineare:
>
> $$\begin{bmatrix} L_0(x_0) & L_1(x_0) & \ldots & L_n(x_0) \\ L_0(x_1) & L_1(x_1) & \ldots & L_n(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ L_0(x_n) & L_1(x_n) & \ldots & L_n(x_n) \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$
>
> è uguale a
>
> $$\begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix} \cdot \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

> E sappiamo che l'unica soluzione del sistema è:
>
> $$\alpha_0 = \alpha_1 = \cdots = \alpha_n = 0.$$

Questo mostra che $L_0(x), L_1(x), ..., L_n(x)$ sono linearmente indipendenti e sono una base di $\mathbb{R}_n[x]$

Definiamo il polinomio:

$$p(x) = y_0 L_0(x) + y_1 L_1(x) + ... + y_n L_n(x)$$

$p(x) \in \mathbb{R}_n[x]$ e per ogni $i = 0, ..., n$ si ha

$$p(x_i) = y_0 L_0(x_i) + y_1 L_1(x_i) + ... + y_n L_n(x_i) = y_i$$

Quindi abbiamo provato l'esistenza di un polinomio in $\mathbb{R}_n[x]$ che nei nodi x_i assume i valori y_i.

Supponiamo che $q(x)$ sia un altro polinomio in $\mathbb{R}_n[x]$ che nei nodi x_i assume i valori y_i. Poiché $L_0(x), ..., L_n(x)$ è una base di $\mathbb{R}_n[x]$, esistono $\beta_0, \beta_1, ..., \beta_n \in \mathbb{R}$ tali che

$$q(x) = \beta_0 L_0(x) + \beta_1 L_1(x) + ... + \beta_n L_n(x)$$

.
Valutando $q(x)$ nei nodi x_i otteniamo che $\forall i = 0, .., n$

$$y_i = q(x_i) = \beta_0 L_0(x_i) + ... + \beta_n L_n(x_i) = \beta_i$$

da cui si ricava che

$$q(x) = y_0 L_0(x) + y_1 L_1(x) + ... + y_n L_n(x) = p(x)$$

Questo prova che $p(x)$ è l'unico polinomio in $\mathbb{R}_n[x]$ che nei nodi x_i assume i valori y_i $\square$.

Teorema

Sia $f : [a, b] \to \mathbb{R}$ una funzione di classe $C^{n+1}[a, b]$ e sia $p(x)$ il polinomio d'interpolazione di $f(x)$ sui punti distinti $x_0, ..., x_n \in [a, b]$.
Allora $\forall x \in [a, b]$ esiste un punto $\xi = \xi(x) \in [a, b]$ tale che:

$$|p(x) - f(x)| = |\frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)(x - x_1)...(x - x_n)|$$

Dim

Sia $x \in [a, b]$ fissato.

- Caso 1: x coincide con uno dei nodi x_i
  - In tal caso posso scegliere un qualsiasi $\xi \in (a, b)$ e sono sicuro che l'uguaglianza vale perché ottengo $0 = 0$ (perché il primo membro è 0 perché $p(x_i) = f(x_i) \forall i = 0, \ldots, n$ per definizione, mentre il 2 membro è 0 perché uno dei termini è 0).
- Caso 2: x non coincide con uno dei nodi $x_0, x_1, \ldots, x_n$
  - Definiamo

$$\pi(y) = (y - x_0)(y - x_1) \ldots (y - x_n)$$

$$r(y) = f(y) - p(y)$$

$$z : [a, b] \to \mathbb{R}, \; z(y) = r(y) - \frac{r(x)}{\pi(x)}\pi(y)$$

  Questa funzione z è di classe $C^{(n+1)}$ siccome non è altro che la composizione di funzioni di classe $C^{(n+1)}$.
  Inoltre si annulla in tutti i nodi x_i e si annulla in x. Dunque si annulla in almeno $n + 2$ punti di $[a, b]$.
  Dunque per il teorema di Rolle si ha che $z'(y)$ si annulla almeno in $n + 1$ punti di (a, b), $z''(y)$ si annulla almeno in n punti di (a, b),...,

$z^{(n+1)}$ si annulla in almeno un punto , che chiamo $\xi \in (a,b)$.
Siccome $p^{(n+1)}(y)$ è nullo (siccome $p(y)$ è un polinomio di grado minore uguale a n) e $\pi^{(n+1)} = (n+1)!$ si ha:

$$\begin{aligned} 0 = z^{(n+1)}(\xi) = r^{(n+1)}(\xi) - \frac{r(x)}{\pi(x)}\pi^{(n+1)}(\xi) = \\ = f^{(n+1)}(\xi) - p^{(n+1)}(\xi) - \frac{f(x)-p(x)}{\pi(x)}(n+1)! = \\ = f^{(n+1)}(\xi) - \frac{f(x)-p(x)}{(x-x_0)(x-x_1)\cdots(x-x_n)}(n+1)! \end{aligned}$$

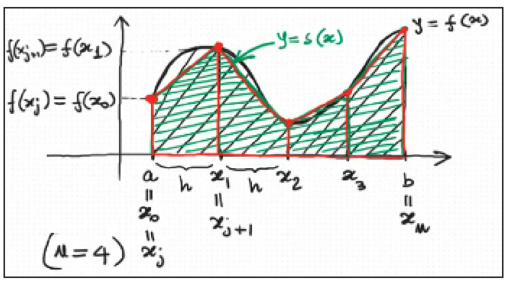
□

Integrazione numerica

Formula dei trapezi

E' data una funzione integrabile $f:[a,b]\to\mathbb{R}$ e si vuole calcolare un'approssimazione di $\int_a^b f(x)dx$. A tal fine si suddivide l'intervallo $[a,b]$ in $n \geq 1$ sottointervalli tutti della stessa ampiezza $h = \frac{b-a}{n}$ e si pone $x_j = a + jh$ per ogni $j = 0, ..., n$. Il valore che si prende come approssimazione di $\int_a^b f(x)dx$ è $\int_a^b s(x)dx$, dove

$$s:[a,b]\leftarrow\mathbb{R}, \begin{cases} s(x) = f(x_j) + \frac{f(x_{j+1} - f(x_j))}{x_{j+1}-x_j}(x-x_j), \\ \text{per } x\in[x_j, x_{j+1}], j = 0, ..., n-1 \end{cases}$$



$$\begin{aligned} I_n = \int_a^b s(x)dx = \sum_{j=0}^{n-1}\int_{x_j}^{x_{j+1}} s(x)dx = \sum_{j=0}^{n-1}\int_{x_j}^{x_{j+1}} f(x_j) + \frac{f(x_{j+1} - f(x_j))}{x_{j+1}-x_j}(x-x_j)dx = \\ = \sum_{j=0}^{n-1}\left[f(x_j)(x-x_j) + \frac{f(x_{j+1}) - f(x_j)}{x_{j+1}-x_j}\frac{(x-x_j)^2}{2}\right]_{x_j}^{x_{j+1}} \end{aligned}$$

> Quello che vienne fatto qui è integrare prima il primo termine:
>
> $$\int_{x_j}^{x_{j+1}} f(x_j)dx = f(x_j)\int_{x_j}^{x_{j+1}} 1dx = [f(x_j) + (x-x_j)]_{x_j}^{x_{j+1}}$$
>
> Ciò viene fatto perche $f(x_j)$ è una costante.
> Infine integro il secondo termine
>
> $$\begin{aligned} \int_{x_j}^{x_{j+1}} \frac{f(x_{j+1} - f(x_j))}{x_{j+1}-x_j}(x-x_j)dx = \frac{f(x_{j+1} - f(x_j))}{x_{j+1}-x_j}dx\int_{x_j}^{x_{j+1}}(x-x_j) = \\ = \left[\frac{f(x_{j+1} - f(x_j))}{x_{j+1}-x_j}\frac{(x-x_j)^2}{2}\right]_{x_j}^{x_{j+1}} \end{aligned}$$
>
> Analogo a prima

$$= \sum_{j=0}^{n-1}\left[f(x_j)(x_{j+1}-x_j) + \frac{f(x_{j+1}) - f(x_j)}{2}(x_{j+1}-x_j)\right]$$

> Qua vabbe sostituisco...

$$= \sum_{j=0}^{n-1} \frac{f(x_j)+f(x_{j+1})}{2} h = \frac{h}{2} \sum_{j=0}^{n-1} [f(x_j)+f(x_{j+1})]$$

> Senza che faccio tutti i calcoli spiego a parole come è uscita fuori quest' h.
> Sopra raccolgo (x_{j+1}, x_j) destra e sinistra. Noto alla fine che (x_{j+1}, x_j) è l'ampiezza degli intervalli ovvero proprio h!!.

$$= h/2 \left[f(a) + f(b) + 2\sum_{j=1}^{n-1} f(x_j) \right] = h \left[\frac{f(a)+f(b)}{2} + \sum_{j=1}^{n-1} f(x_j) \right]$$

Lemma

Siano $\omega, \alpha, \beta : [a,b] \to \mathbb{R}$ funzioni tali che:

- $\omega(x)$ è continua e ≥ 0 su $[a,b]$
- $\alpha(x)$ e $\beta(x)\omega(x)$ sono continue su $[a,b]$
- $m \leq \beta(x) \leq M$ per ogni $x \in [a,b]$, dove m e M sono rispettivamente il minimo e il massimo di $\alpha(x)$ su $[a,b]$

Allora esiste un punto $\eta \in [a,b]$ tale che:

$$\int_a^b \beta(x)\omega(x)dx = \alpha(\eta) \int_a^b \omega(x)dx$$

Dim

Poiché $\omega(x) \geq 0$ su $[a,b]$ e $m \leq \beta(x) \leq M$ per ogni $x \in [a,b]$, si ha $m\omega(x) \leq \beta(x)\omega(x) \leq M\omega(x)$ per ogni $x \in [a,b]$ e si ha:

$$m \int_a^b \omega(x)dx \leq \int_a^b \beta(x)\omega(x)dx \leq M \int_a^b \omega(x)dx$$

Consideriamo la funzione $z : [a,b] \to \mathbb{R}$.

$$z(y) = \alpha(y) \int_a^b \omega(x)dx$$

Questa funzione è continua su $[a,b]$ perché $\alpha(y)$ è continua su $[a,b]$.

Abbiamo che il massimo di $z(y)$ è $M \int_a^b \omega(x)dx$ essendo M il massimo di $\alpha(y)$. Analogo per il minimo con $m \int_a^b \omega(x)dx$.

Quindi, per il teorema dei valori intermedi, $z(y)$ assume su $[a,b]$ tutti i valori compresi tra il suo minimo $m \int_a^b \omega(x)dx$ e il suo massimo $M \int_a^b \omega(x)dx$.
In particolare $z(y)$ assume il valore $\int_a^b \beta(x)\omega(x)dx$ ovvero esiste $\eta \in [a,b]$ tale che:

$$z(\eta) = \int_a^b \beta(x)\omega(x)dx$$

$\square$

Teorema

Sia $f : [a,b] \to \mathbb{R}$ di classe $C^2[a,b]$ e sia I_n la formula dei trapezi di ordine n e passo $h = \frac{b-a}{n}$ per approssimare $\int_a^b f(x)dx$.
Allora esiste un punto $\eta \in [a,b]$ tale che:

$$\int_a^b f(x)dx - I_n = -\frac{(b-a)f''(\eta)}{12} h^2$$

Dim

Poniamo $x_j = a + jh \; \forall j = 0, ..., n$ e indichiamo con $s(x)$ la funzione lineare usata precedentemente. Risulta:

$$\int_a^b f(x) - I_n = \int_a^b f(x)dx - \int_a^b s(x)dx = \int_a^b [f(x) - s(x)]dx$$

$$\sum_{j=0}^{n-1} \int_{x_j}^{x_{j+1}} f''(\xi_j(x)) \frac{(x-x_j)(x-x_{j+1})}{2}\, dx$$

per il Teorema 1.2 sull'intervallo $[x_j, x_{j+1}]$

$$= -\sum_{j=0}^{n-1} \int_{x_j}^{x_{j+1}} f''(\xi_j(x)) \frac{(x-x_j)(x_{j+1}-x)}{2}\, dx$$

$$= -\sum_{j=0}^{n-1} f''(\eta_j) \int_{x_j}^{x_{j+1}} \frac{(x-x_j)(x_{j+1}-x)}{2}\, dx$$

(Per il Lemma 2.1 applicato sull'intervallo $[x_j, x_{j+1}]$) con:

- $\alpha(x) = f''(x)$
- $\beta(x) = f''(\xi_j(x))$
- $\omega(x) = \frac{(x-x_j)(x_{j+1}-x)}{2}$

Dove η_j è un punto in $[x_j, x_{j+1}]$).

$$= -\sum_{j=0}^{n-1} f''(\eta_j) \int_0^h \frac{t(h-t)}{2}\, dt$$

> Cambio di variabile: $t = x - x_j \Rightarrow x = t + x_j$, quindi $dt = dx$:
> Dunque siccome $x = t + x_j$ avrei :
>
> $$\frac{(t)(x_{j+1} - (t + x_j))}{2} = \frac{t(x_{j+1} - x_j - t)}{2} = \frac{t(h-t)}{2}$$

Risolvendo l'integrale:

$$= -\sum_{j=0}^{n-1} f''(\eta_j) \frac{1}{2} \left[\frac{ht^2}{2} - \frac{t^3}{3}\right]_0^h = -\sum_{j=0}^{n-1} f''(\eta_j) \frac{h^3}{12} =$$

$$-\frac{h^3 n}{12} \cdot \frac{1}{n} \sum_{j=0}^{n-1} f''(\eta_j) = -\frac{h^2(b-a)}{12} f''(\eta)$$

> Abbiamo che la parte in blu è una media aritmetica di n valutazioni di f'', dunque essendo una media sarà sicuramente compresa tra il minimo e il massimo di f''. Dunque per il teorema dei valori intermedi esiste un η per cui $f''(\eta) = \frac{1}{n} \sum_{j=0}^{n-1} f''(\eta_j)$.

Analisi di matrici

Una matrice Hermitiana ha i valori sulla diagonale reali.

Una matrice Hermitiana ha autovalori reali, infatti

Una matrice $A \in \mathbb{C}^{n \times n}$ si dice **definita positiva** se

$$\mathrm{Re}(x^* A x) > 0 \quad \text{per ogni } x \in \mathbb{C}^n \setminus \{0\}.$$

Notiamo che:

$$\mathrm{Re}(x^* A x) = \frac{x^* A x + x^* A x^*}{2} = \frac{x^* A x + (x^* A x)^*}{2} = \frac{x^* A x + x^* A^* x}{2} = x^* \left(\frac{A + A^*}{2}\right) x = x^* \,\mathrm{Re}(A) x,$$

dove

$$\mathrm{Re}(A) := \frac{A + A^*}{2}, \quad \mathrm{Im}(A) := \frac{A - A^*}{2i}, \quad A = \mathrm{Re}(A) + i\,\mathrm{Im}(A).$$

$\mathrm{Re}(A)$ e $\mathrm{Im}(A)$ sono matrici hermitiane.

Infatti

$$(\mathrm{Re}(A))^* = \left(\frac{A + A^*}{2}\right)^* = \frac{A^* + A^{**}}{\overline{2}} = \frac{A + A^*}{2}$$
$$(\mathrm{Im}(A))^* = \left(\frac{A - A^*}{2i}\right)^* = -\frac{A^* - A^{**}}{2i} = \frac{A - A^*}{2i}$$

A è definita positiva se e soltanto se $\mathrm{Re}(A)$ è definita positiva

$$\begin{aligned} A \text{ è definita positiva} \quad &\iff \quad \forall \mathbf{x} \in C^{n\times n} \backslash \{\mathbf{0}\}\ \mathrm{Re}(\mathbf{x}^* A\mathbf{x}) > 0 \\ &\iff \quad \forall \mathbf{x} \in C^{n\times n} \backslash \{\mathbf{0}\}\ \mathbf{x}^* \mathrm{Re}(A)\mathbf{x} > 0 \\ &\iff \quad \forall \mathbf{x} \in C^{n\times n} \backslash \{\mathbf{0}\}\ \mathrm{Re}(\mathbf{x}^* \mathrm{Re}(A)\mathbf{x}) > 0 \\ &\iff \quad \mathrm{Re}(A) \text{ è definita positiva} \end{aligned}$$

> La terza equivalenza è data dal fatto che essendo $\mathbf{x}^*\mathrm{Re}(A)\mathbf{x}$ reale allora $\mathbf{x}^*\mathrm{Re}(A)\mathbf{x} = \mathrm{Re}(\mathbf{x}^*\mathrm{Re}(A)\mathbf{x})$ cioè sto prendendo la parte reale di un numero reale... Il numero stesso insomma.

Se una matrice A è definita positiva allora è invertibile perché i suoi autovalori sono positivi. Infatti preso $\lambda \in \mathbb{C}$ autovalore di A e $\mathbf{x} \neq \mathbf{0}$ un corrispondente autovettore ho

$$A\mathbf{x} = \lambda\mathbf{x} \iff \mathbf{x}^* A\mathbf{x} = \mathbf{x}^*\lambda\mathbf{x} = \lambda \sum_{i=0}^{n} |x_i|^2 \implies \lambda = \frac{\mathbf{x}^* A\mathbf{x}}{\sum_{i=0}^{n} |x_i|^2}$$

Sotto al denominatore ho sicuramente una quantità positiva in quanto $\mathbf{x} \neq 0$ e ogni somma è un valore al quadrato.
Sopra invece ho un valore positivo in quanto A è definita positiva e per definizione

$$\forall \mathbf{x} \neq 0\ \mathbf{x}^* A\mathbf{x} > 0$$

Teorema

Sia $A \in \mathbb{C}^{n\times n}$ una matrice hermitiana e siano $A_1, A_2, \ldots, A_n$ le sue sottomatrici principali di testa:

$$A_1 = [a_{11}], \quad A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad A_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad \ldots, \quad A_n = A.$$

Le seguenti condizioni sono equivalenti:

- A è definita positiva;
- $\mathbf{x}^* A\mathbf{x} > 0$ per ogni $x \in \mathbb{C}^n \setminus \{0\}$;
- Gli autovalori di A sono reali e positivi;
- $\det(A_k) > 0$ per ogni $k = 1, \ldots, n$.

Teorema

Se $p(\lambda)$ è un polinomio e $A \in \mathbb{C}^{n\times n}$ è una matrice con autovalori $\lambda_1, \ldots, \lambda_n$, allora gli autovalori della matrice $p(A)$ sono $p(\lambda_1), \ldots, p(\lambda_n)$.

dim

Dimostreremo il teorema solo in 3 casi:

- Caso 1: **Il polinomio $p(\lambda) = a_0$ è costante**
  - In tal caso $P(A) = a_0 I$ e i suoi autovalori sono $a_0, \ldots, a_0$ (ripetuto n volte) .

  > Questo perché $a_0 I$ è una matrice diagonale con tutti a_0 sulla diagonale

  Dunque gli autovalori di $p(A)$ sono $p(\lambda_1), \ldots, p(\lambda_n)$ e la tesi del teorema vale.

- Caso 2: **Il polinomio $p(\lambda) = a_0 + a_1\lambda$ ha grado 1**
  - In tal caso, il polinomio caratteristico di $p(A)$ e quello di A sono legati dalla seguente relazione $\forall\lambda \in \mathbb{C}$:

$$\begin{aligned} C_{p(A)}(\lambda) &= \det(\lambda I - p(A)) = \det(\lambda I - (a_0 I + a_1 A)) = \det((\lambda - a_0)I - a_1 A) \\ &= \det(a_1(\frac{\lambda - a_0}{a_1} I - A)) = a_1^n \det(\frac{\lambda - a_0}{a_1} I - A) = a_1^n C_A(\frac{\lambda - a_0}{a_1}) \end{aligned}$$

  Dunque gli autovalori di $p(A)$ sono:

$$\{\lambda \in \mathbb{C} \mid C_{p(A)}(\lambda) = 0\} = \left\{\lambda \in \mathbb{C} \mid C_A\left(\frac{\lambda - a_0}{a_1}\right) = 0\right\} = \left\{\lambda \in \mathbb{C} \mid \frac{\lambda - a_0}{a_1} = \lambda_1, \dots, \lambda_n\right\} = \\ \{\lambda \in \mathbb{C} \mid \lambda = a_0 + a_1\lambda_1, \dots, a_0 + a_1\lambda_n\} = \{a_0 + a_1\lambda_1, \dots, a_0 + a_1\lambda_n\} = \{p(\lambda_1), \dots, p(\lambda_n)\}$$

- Caso 3: La matrice A è diagonalizzabile
  - In tal caso, esistono una matrice invertibile X e una matrice diagonale

$$D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$$

(avente sulla diagonale gli autovalori di A) tali che

$$A = XDX^{-1},$$

$$A^2 = XDX^{-1}XDX^{-1} = XD^2X^{-1},$$

$$A^3 = XDX^{-1}XDX^{-1}XDX^{-1} = XD^3X^{-1},$$

$$\vdots$$

Pertanto, fissato un polinomio

$$p(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \dots + a_m\lambda^m,$$

si ha

$$p(A) = a_0 I + a_1 A + a_2 A^2 + \dots + a_m A^m.$$

Sostituendo $A = XDX^{-1}$, otteniamo

$$p(A) = X\left(a_0 I + a_1 D + a_2 D^2 + \dots + a_m D^m\right) X^{-1}.$$

Usando la notazione $p(D)$, otteniamo

$$p(A) = Xp(D)X^{-1}, \quad (3.2)$$

dove
[
p(D) = a_0 I + a_1 D + a_2 D^2 + \dots + a_m D^m =
\begin{bmatrix}
\end{bmatrix}. \quad (3.3)
]

Teorema (Primo Teorema di Gershgorin)

Gli autovalori di una matrice $A \in \mathbb{C}^{n\times n}$ stanno tutti nell'unione dei cerchi di Gershgorin di A.

dim

Sia λ un autovalore di A e sia $\vec{u} \neq \vec{0}$ un corrispondente autovettore. Si ha:

$$A\vec{u} = \lambda\vec{u} \iff (A\vec{u})_i = (\lambda\vec{u})_i \ \forall i = 1, ..., n \iff \sum_{j=1}^{n} a_{ij}u_j = \lambda u_i \ \forall i = 1, ..., n$$

Selezionando un indice i_0 corrispondente a una componente u_{i_0} di modulo massimo di $\vec{u}$, la precedente equazione i_0-esima ci dice che:

$$\sum_{j=1}^{n} a_{ij}u_j = \lambda u_i \ \forall i = 1, ..., n \iff$$
$$a_{i_0i_0}u_{i_0} + \sum_{j=1,\, j\neq i_0}^{n} a_{i_0j}u_j = \lambda u_{i_0} \iff (\lambda - a_{i_0i_0})u_{i_0} = \sum_{j=1,\, j\neq i_0}^{n} a_{i_0j}u_j$$
$$\implies |(\lambda - a_{i_0i_0})||u_{i_0}| = |\sum_{j=1,\, j\neq i_0}^{n} a_{i_0j}u_j|$$
$$\leq \sum_{j=1,\, j\neq i_0}^{n} |a_{i_0j}||u_j|$$

Per la disuguaglianza triangolare

$$\leq \sum_{j=1,\, j\neq i_0}^{n} |a_{i_0j}||u_{i_0}|$$

u_{i_0} è il massimo di tutti gli u_i

$$= |u_{i_0}| \sum_{j=1,\, j\neq i_0}^{n} |a_{i_0j}|$$

Dunque

$$|(\lambda - a_{i_0i_0})|\cancel{|u_{i_0}|} \leq \cancel{|u_{i_0}|} \sum_{j=1,\, j\neq i_0}^{n} |a_{i_0j}|$$

dunque λ dista da $a_{i_0i_0}$ (che è il centro del cerchio di G. K_{i_0}) $\leq$ del raggio $\sum_{j=1,\, j\neq i_0}^{n} |a_{i_0j}|$ di $K_{i_0} \implies \lambda \in K_{i_0} \implies \lambda \in \bigcup_{i=1}^{n} K_i \, \square$.

Teorema (Secondo Teorema di Gershgorin)

Supponiamo che l'unione di k cerchi di Gershgorin di $A \in \mathbb{C}^{n\times n}$ sia disgiunta dall'unione degli altri $n-k$ cerchi. Allora k autovalori di A stanno nella prima unione e $n-k$ nella seconda.

Teorema (Terzo Teorema di Gershgorin)

Supponiamo che $A \in \mathbb{C}^{n\times n}$ sia irriducibile. Allora i punti che stanno sul bordo di quei cerchi di Gershgorin a cui appartengono, ma non sul bordo di tutti i cerchi, non sono autovalori di A.

Teorema (Terzo Teorema di Gershgorin Debole)

Supponiamo che $A \in \mathbb{C}^{n\times n}$ sia irriducibile e sia B il bordo dell'unione dei cerchi di Gershgorin. Allora i punti di B che non stanno sul bordo di tutti i cerchi non sono autovalori di A.

Definizione: Dominanza diagonale

Sia $A \in \mathbb{C}^{n\times n}$ una matrice.

- La matrice A è detta **a diagonale dominante per righe** se:
  - $$|a_{ii}| \geq \sum_{j\neq i} |a_{ij}| \ \ \forall i = 1, \ldots, n$$
  - esiste almeno un indice $k \in \{1, \ldots, n\}$ per il quale vale la disuguaglianza stretta:

$$|a_{kk}| > \sum_{j \neq k} |a_{kj}|.$$

- La matrice A è detta **a diagonale dominante in senso stretto per righe** se:

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

- La matrice A è detta **a diagonale dominante per colonne** se:
  - $$|a_{jj}| \geq \sum_{i \neq j} |a_{ij}| \ \ \forall i = 1, \dots, n$$
  - esiste almeno un indice $\ell \in \{1, \dots, n\}$ per il quale vale la disuguaglianza stretta:

$$|a_{\ell\ell}| > \sum_{i \neq \ell} |a_{i\ell}|.$$

La matrice A è detta **a diagonale dominante in senso stretto per colonne** se:

$$|a_{jj}| > \sum_{i \neq j} |a_{ij}| \ \ \forall i = 1, \dots, n$$

Teorema

Supponiamo che la matrice $A \in \mathbb{C}^{n \times n}$ soddisfi almeno una delle seguenti condizioni:

- A è **a diagonale dominante e irriducibile**;
- A è **a diagonale dominante in senso stretto**;
- A è **a diagonale dominante per colonne e irriducibile**;
- A è **a diagonale dominante in senso stretto per colonne**.

Allora A **è invertibile**.

dim

Dimostriamo secondo l'ipotesi che A sia a diagonale dominante e irriducibile.
Mostriamo che 0 non è autovalore di A usando il terzo teorema di Gershgorin. Siccome A è a diagonale dominante, se 0 appartiene a un cerchio di Gershgorin K_i allora deve per forza stare sul bordo di K_i. Infatti non può stare all'interno, perché per l'ipotesi di dominanza diagonale si ha:

$$\text{raggio di } K_i = \sum_{j \neq i} |a_{ij}| \leq |a_{ii}| = \text{distanza}(a_{ii,0}) = \text{distanza}(\text{centro di } K_i, 0)$$

Dunque 0 sta per forza di cose sul bordo di quei cerchi. Inoltre sempre per ipotesi di dominanza diagonale, esiste un indice k per cui

$$|a_{kk}| > \sum_{j \neq k} |a_{kj}|$$

Questo significa che 0 non sta neanche sul bordo di K_k e dunque 0 non sta sul bordo di tutti i cerchi di G. di A. Siccome A è irriducibile, il terzo teorema di Gershgorin ci dice che 0 non può essere autovalore di A. Dunque è invertibile.

Norme Vettoriali

Definizione

Una funzione $|| \cdot ||$: $\mathbb{C} \to \mathbb{R}$ si dice norma vettoriale se soddisfa le seguenti proprietà:

1. $||\mathbf{x}|| \geq 0 \forall \mathbf{x} \in \mathbb{C}^n e ||\mathbf{x}|| \iff \mathbf{x} = \mathbf{0} \quad \text{Positività}$

2. $||\alpha\mathbf{x}|| = \alpha||\mathbf{x}|| \forall \mathbf{x} \in \mathbb{C}^n,\ \alpha \in \mathbb{C} \quad \text{Omogeneità}$
3. $||\mathbf{x}+\mathbf{y}|| \le ||\mathbf{x}|| + ||\mathbf{y}|| \quad \text{Disuguaglianza triangolare}$

Definizione

Sia $||\cdot|| : \mathbb{C} \to \mathbb{R}$ una norma vettoriale, definiamo la distanza tra due vettori $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ come $||\mathbf{x}-\mathbf{y}||$

Definizione

Una funzione $||\cdot|| : \mathbb{C}^{n\times n} \to \mathbb{R}$ si dice norma matriciale se soddisfa le seguenti proprietà:

1. $||A|| \ge 0 \forall \mathbf{x} \in \mathbb{C}^n e ||A|| = 0 \iff A = 0 \quad \text{Positività}$
2. $||\alpha A|| = \alpha||A|| \forall A \in \mathbb{C}^{n\times n},\ \alpha \in \mathbb{C} \quad \text{Omogeneità}$
3. $||A+B|| \le ||A|| + ||B|| \quad \text{Disuguaglianza triangolare}$

Definizione

Sia $||\cdot|| : \mathbb{C}^{n\times n} \to \mathbb{R}$ una norma matriciale, definiamo la distanza tra due matrici $A, B \in \mathbb{C}^{n\times n}$ come $||A-B||$

Definizione

Data una norma vettoriale $||\cdot|| : \mathbb{C} \to \mathbb{R}$ e una matrice $A \in \mathbb{C}^{n\times n}$ definiamo norma matriciale indotta dalla norma vettoriale $||\cdot||$:

$$||A|| = \max_{\mathbf{x}\neq 0} \frac{||A\mathbf{x}||}{||\mathbf{x}||} = \max_{\mathbf{x}\neq 0} ||A\frac{\mathbf{x}}{||\mathbf{x}||}|| = \max_{||\mathbf{y}||=1} ||A\mathbf{y}||$$

Teorema

Sia $||\cdot|| : \mathbb{C}^{n\times n} \to \mathbb{R}$ una norma matriciale indotta dalla norma vettoriale $||\cdot||$ e siano $A, B \in \mathbb{C}^{n\times n}$. Allora valgono le seguenti proprietà:

1. $||I|| = 1$
2. $||A\mathbf{x}|| \le ||A||\ ||\mathbf{x}||\ \forall \mathbf{x} \in \mathbb{C}^n$
3. $||A||$ è la più piccola costante C che soddisfa $||A\mathbf{x}|| \le C||\mathbf{x}||\ \forall \mathbf{x} \in C^n$
4. $||AB|| \le ||A||\ ||B|| \quad \text{Submoltiplicatività}$
5. $\rho(A) \le ||A||$

Dim

1. $$||I|| = \max_{||\mathbf{y}||=1} ||I\mathbf{y}|| = \max_{||\mathbf{y}||=1} ||\mathbf{y}|| = 1$$
2. Per ogni $\mathbf{x} \neq \mathbf{0}$ si ha
$$\frac{||A\mathbf{x}||}{||\mathbf{x}||} \le \max_{\mathbf{y}\neq\mathbf{0}} \frac{||A\mathbf{y}||}{||\mathbf{y}||} = ||A|| \implies ||A\mathbf{x}|| \le ||A||\ ||\mathbf{x}||$$
3. Presa una qualsiasi costante C che soddisfa $||A\mathbf{x}|| \le C||\mathbf{x}|| \forall \mathbf{x} \in \mathbb{C}^n$ ho
$$\frac{||A\mathbf{x}||}{||\mathbf{x}||} \le C\ \ \forall \mathbf{x} \neq \mathbf{0} \implies ||A|| = \max_{\mathbf{x}\neq\mathbf{0}} \frac{||A\mathbf{x}||}{||\mathbf{x}||} \le C$$
4.
5. Sia λ un autovalore di modulo massimo e sia $\mathbf{x} \neq \mathbf{0}$ un corrispondente autovettore. Dall'equazione $A\mathbf{x} = \lambda\mathbf{x}$ otteniamo:
$$||A\mathbf{x}|| = ||\lambda\mathbf{x}|| = |\lambda|||\mathbf{x}|| = \rho(A)||\mathbf{x}|| \implies \rho(A) = \frac{||A\mathbf{x}||}{||\mathbf{x}||} \le \max_{\mathbf{y}\neq\mathbf{0}} \frac{||A\mathbf{y}||}{||\mathbf{y}||} = ||A||$$

Teorema

Sia $A \in C^{n\times n}$. Allora:

$$\lim_{k\to\infty} A^k = O \iff \rho(A) < 1$$

Dim

Dimostriamo solo nel caso in cui A è diagonalizzabile.
In tal caso esistono una matrice X invertibile e una matrice diagonale $\mathrm{diag}(\lambda_1, ..., \lambda_n)$ tali che:

$$\begin{gathered} A = XDX^{-1} \\ A^2 = XD^2X^{-1} \\ \vdots \\ A^k = XD^kX^{-1} \end{gathered}$$

($\Longleftarrow$)

Se $\rho(A) < 1$ allora dall'equazione $A^k = XD^kX^{-1}$ e dalla proprietà di submoltiplicatività applicata alla norma $||\cdot||_\infty$ si ottiene:

$$||A^k||_\infty = ||XD^kX^{-1}||_\infty \leq ||X||_\infty ||D^k||_\infty ||X^{-1}||_\infty = ||X||_\infty ||(\rho(A))^k||_\infty ||X^{-1}||_\infty \to 0$$

> Questo perche $||D^k||_\infty = ||\mathrm{diag}(\lambda_1^k, ..., \lambda_n^k)||_\infty = \max(|\lambda_1^k|, ..., |\lambda_n^k|) = \rho(A)^k$
>
> In più tende a 0 perchè $\lim_{k\to 0} \rho(A)^k = 0$, siccome $\rho(A) < 1$

Dunque $||A^k||_\infty \to 0$ e $A^k \to O$.

> Questo perché se cio che $||A^k||_\infty \to 0$ significa che
> $\max(||A^k_{[1]}||_1, ..., ||A^k_{[n]}||_1) \to 0$, ove ogni $||A^k_{[i]}||_1$ corrisponde alla somma dei moduli delle componenti della matrice A^k riga per riga.
> Se il massimo tende a 0 significa che ogni componente della matrice tende a 0 e quindi A^k tende al vettore nullo.

($\Longrightarrow$)

Se $A^k = O$
Abbiamo $A^k = XD^kX^{-1}$ da questa uguaglianza otteniamo $D^k = XA^kX^{-1}$ e

$$\rho(A)^k = ||D^k||_\infty = ||XA^kX^{-1}||_\infty \leq ||X||_\infty ||A^k||_\infty ||X^{-1}||_\infty \to 0$$

> Siccome $A^k \to O$ allora $||A^k||_\infty \to 0$

Dunque $\rho(A) < 1$ poiché $\rho(A)^k \to 0$

Risoluzione di sistemi lineari

Dato un sistema lineare:

$$A\mathbf{x} = \mathbf{b}$$

con $A \in \mathbb{C}^{n\times n}$ invertibile. Tale sistema ha un'unica soluzione $\mathbf{x} = A^{-1}\mathbf{b}$.
Vogliamo risolvere il sistema con un metodo iterativo, cioè un metodo che a partire da un vettore $\mathbf{x}^{(0)}$ costruisce una successione di vettori che convergano $\mathbf{x}$.

Consideriamo i seguenti metodi iterativi:

$$\begin{gathered} \mathbf{x}^{(0)} \in \mathbb{C}^n, \\ \mathbf{x}^{(k+1)} = P\mathbf{x}^k + \mathbf{q}, \;\; k = 0, 1, 2... \quad (4.2) \end{gathered}$$

dove $P \in \mathbb{C}^{n\times n}$ è detta *matrice d'iterazione*

Teorema (condizione necessaria e sufficiente di convergenza).

Supponiamo che il metodo (4.2) sia consistente con (4.1). Allora esso è convergente se e solo se $\rho(P) < 1$

Dim

$(\Longleftarrow)$
Poiché il metodo è consistente per ipotesi, vale:

$$\mathbf{x} = P\mathbf{x} + \mathbf{q} \ \ (4.3)$$

Inoltre vale, ovviamente anche l'equazione del metodo:

$$\mathbf{x}^{(k+1)} = P\mathbf{x}^{(k)} + \mathbf{q}, \ \ k = 0, 1, 2... \ \ (4.4)$$

Sottraendo membro a membro la (4.4) e la (4.3) si ottiene l'equazione dell'errore:

$$\mathbf{e}^{(k+1)} = P\mathbf{e}^{(k)} \quad \forall k = 0, 1, 2, \ldots \ \ (4.5)$$

dove

$$\mathbf{e}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}$$

è l'errore al passo k.

Sviluppando per ricorrenza la (4.5) si ottiene:

$$e^{(k+1)} = Pe^{(k)} = P^2 e^{(k-1)} = P^3 e^{(k-2)} = \cdots = P^{k+1} e^{(0)}, \quad k = 0, 1, 2, \ldots$$

Per ipotesi $\rho(P) < 1$, e un noto teorema dell'analisi numerica ci dice che $P^k \to O$.
Dunque si deduce che $\mathbf{e}^{(k)} \to 0$ e dunque ricordando a cosa è uguale $\mathbf{e}^{(k)}$ si ha:

$$\mathbf{x}^{(k)} - \mathbf{x} \to 0 \iff \mathbf{x}^{(k)} \to \mathbf{x}$$

$\square$

Corollario (Condizioni sufficienti di convergenza)

Supponiamo che il metodo (4.2) sia consistente con (4.1). Se esiste una norma matriciale indotta $\|\cdot\|$ tale che $\|P\| < 1$ allora il metodo è convergente.

Dim

Poiché $\rho(P) \leq \|P\|$ per un teorema, la condizione $\|P\| < 1$ implica che $\rho(P) < 1$ e dunque il metodo è convergente.

Corollario (Condizioni necessarie di convergenza)

Supponiamo che il metodo (4.2) sia consistente con (4.1).

- Se $|\text{traccia}(P)| \geq n$ allora il metodo non è convergente.
- Se $|\det(P)| \geq 1$ allora il metodo non è convergente.

Quindi le condizioni $|\text{traccia}(P)| < n$ e $|\det(P)| < 1$ sono necessarie per la convergenza del metodo (4.2).

Dim

- Se $|\text{traccia}(P)| \geq n$ allora esiste almeno un autovalore di P di modulo ≥ 1.
  Infatti, se tutti gli autovalori $\lambda_1, \ldots, \lambda_n$ di P fossero di modulo < 1, allora avremmo:

  $$|\text{traccia}(P)| = |\lambda_1 + \cdots + \lambda_n| \leq |\lambda_1| + \cdots + |\lambda_n| < n.$$

  Poiché esiste almeno un autovalore di P di modulo ≥ 1, si deduce che:

  $$\rho(P) = \max(|\lambda_1|, \ldots, |\lambda_n|) \geq 1$$

  e dunque il metodo non è convergente.

- Se $|\det(P)| \geq 1$ allora esiste almeno un autovalore di P di modulo ≥ 1.
  Infatti, se tutti gli autovalori $\lambda_1, \ldots, \lambda_n$ di P fossero di modulo < 1, allora avremmo:

  $$|\det(P)| = |\lambda_1 \cdot \cdots \cdot \lambda_n| = |\lambda_1| \cdot \cdots \cdot |\lambda_n| < 1.$$

  Poiché esiste almeno un autovalore di P di modulo ≥ 1, si deduce che:

  $$\rho(P) = \max(|\lambda_1|, \ldots, |\lambda_n|) \geq 1$$

  e dunque il metodo non è convergente.

Teorema

Supponiamo che la matrice $A \in \mathbb{C}^{n \times n}$ soddisfi almeno una delle seguenti condizioni:

- A è a diagonale dominante e irriducibile;
- A è a diagonale dominante in senso stretto;
- A è a diagonale dominante per colonne e irriducibile;
- A è a diagonale dominante in senso stretto per colonne.

Allora i metodi di Jacobi e Gauss-Seidel per risolvere un sistema lineare di matrice A sono convergenti.

Dim

Dimostriamo il teorema per il metodo di Gauss-Seidel sotto l'ipotesi di dominanza diagonale e irriducibilità.
Dobbiamo dimostrare che $\rho(G) < 1$, ove $G = I - E^{-1}A$ è la matrice d'iterazione del metodo di Gauss-Seidel. Per l'osservazione "Smart", gli autovalori di G sono le radici del polinomio:

$$\det(\lambda E + A - E) = \begin{vmatrix} \lambda a_{11} & a_{12} & a_{13} & a_{14} \\ \lambda a_{21} & \lambda a_{22} & a_{23} & a_{24} \\ \lambda a_{31} & \lambda a_{32} & \lambda a_{33} & a_{34} \\ \lambda a_{41} & \lambda a_{42} & \lambda a_{43} & \lambda a_{44} \end{vmatrix} \quad (\star)$$

Nessun numero di λ di modulo ≥ 1 può essere radice di questo polinomio. Infatti se $|\lambda| \geq 1$ allora la matrice $\lambda E + A - E$ è a diagonale dominante e irriducibile esattamente come A.
La dimostrazione di ciò si basa su due osservazioni:

- $\lambda E + A - E$ è a diagonale dominante come A. Infatti $\forall i = 0, \ldots, n$

  $$|\lambda a_{ii}| = |\lambda||a_{ii}| \geq |\lambda|(|a_{i1}| + \ldots + |a_{i\,i-1}| + |a_{i\,i+1}| + \ldots + |a_{in}|)$$

  > Questo perché $|a_{ii}| \geq |a_{i1}| + \ldots + |a_{i\,i-1}| + |a_{i\,i+1}| + \ldots + |a_{in}|$ poiché A è a diagonale dominante.

  $$= |\lambda a_{i1}| + \ldots + |\lambda a_{i\,i-1}| + |\lambda| \cdot |a_{i\,i+1}| + \ldots + |\lambda| \cdot |a_{in}|$$

  $$\geq |\lambda a_{i1}| + \ldots + |\lambda a_{i\,i-1}| + |a_{i\,i+1}| + \ldots + |a_{in}|$$

  > Siccome $|\lambda| \geq 1$.

  Dunque è a diagonale dominante poiche:

  $$|\lambda a_{ii}| \geq |\lambda a_{i1}| + \ldots + |\lambda a_{i\,i-1}| + |a_{i\,i+1}| + \ldots + |a_{in}|$$

- $\lambda E + A - E$ è irriducibile perché gli elementi nulli di $\lambda E + A - E$ si trovano nelle stesse posizioni in cui si trovano gli elementi nulli di $A \implies \lambda E + A - E$ ed A hanno lo stesso grafo $\implies \lambda E + A - E$ è irriducibile perché A è irriducibile.

Dunque nessun numero λ di modulo ≥ 1 può essere radice di $(\star)$, altrimenti la matrice $\lambda E + A - E$ è a diagonale dominante e irriducibile, per cui è invertibile per un teorema e dunque $\det(\lambda E + A - E) \neq 0$. In definitiva gli autovalori di G devono essere per forza in modulo < 1 e perciò $\rho(G) < 1$.

$\square$

me so dimenticato de salva l'ultimo teorema....